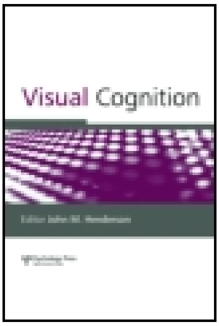

Visual Cognition
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pvis20

Introduction to real-world scene perception

John Henderson [a]
[a] Michigan State University, East Lansing, MI, USA

Version of record first published: 01 Oct 2010.

VISUAL COGNITION, 2005, 12 (6), 849–851

Psychology Press
Taylor & Francis Group

Introduction to real-world scene perception

John M. Henderson
Michigan State University, East Lansing, MI, USA

Real-world scene perception involves sensory and cognitive processing of visual input that in some important sense is like that typically encountered during the natural course of everyday human activity. Unlike objects, which are viewed from a vantage point outside of themselves, scenes are specific views of the environment within which the viewer is embedded. Described in this way, real-world scene perception differs from perception of the visual stimuli that are often used in psychophysical and visual cognition experiments. For example, real-world scenes fill the entire visual field and produce specific wavelength, intensity, and spatial frequency profiles that are constrained by properties of the world itself. These regularities in image statistics arise in part because real-world scenes are composed of objects and backgrounds that have specific semantic and spatial constraints that in turn give rise to the semantics of the scene as a whole.

Scene can be defined as a semantically coherent (and often nameable) human-scaled view of a real-world environment comprising background elements and multiple discrete objects arranged in a spatially licensed manner (Henderson & Ferreira, 2004; Henderson & Hollingworth, 1999). Background elements are larger scale, immovable surfaces and structures, and objects are smaller scale discrete entities that move or can be moved within the scene. Real-world scenes have a hierarchical spatial structure. For example, a kitchen viewed from the vantage point of the doorway would likely include floor, ceiling, and walls as background elements, and a stove, refrigerator, and dishwasher as a few prominent objects among others, spatially arranged according to the laws of physics and the constraints associated with the functions of a kitchen. But, from the vantage point of standing in front of the counter, the countertop might be considered a scene, with its surface forming most of the background and a loaf of bread, knife, peanut butter jar, and plate serving as individuated objects, themselves spatially arranged according to the same physical laws and the semantics of sandwich preparation. One way to bound the

Please address all correspondence to: John M. Henderson, 285B Psychology Building, Michigan State University, East Lansing, MI 48824-1116, USA. Email: john@eyelab.msu.edu

http://www.tandf.co.uk/journals/pp/13506285.html DOI:10.1080/13506280444000544

concept of *scene* is to limit scale to that affording human locomotion and motor interaction (Henderson & Ferreira, 2004).

Spatial licensing includes adherence to the physical constraints of the universe, including gravity, space, and time, and to the semantic constraints imposed by object identity and function (Biederman, Mezzanotte, & Rabinowitz, 1982). Examples of the former constraint include the fact that objects generally cannot float unsupported in the air and that two objects cannot occupy the same region of space at the same time. Examples of the latter constraint include the fact that a fire hydrant does not belong on top of a mailbox and that couches are not typically found outdoors (Biederman et al., 1982). Some objects of course can float without visible support, and some objects do allow other objects to be seen through them (i.e., violate standard rules of interposition), so determining whether a physical constraint has been violated often requires knowledge of the identity and semantic attributes of the objects (Henderson, 1992). Nevertheless, a well-formed scene must ultimately conform to a set of basic physical and semantic constraints that impose structure on the environment.

Henderson and Ferreira (2004) noted two orthogonal dimensions along which the stimuli used to study real-world scene perception can differ. First, the stimuli in scene experiments can either be the real environment itself, or some form of depiction of it. Second, stimuli can be complete or can be degraded in some way. Henderson and Ferreira (2004) referred to the former as true real-world scenes, and the latter as ersatz scenes. Examples of ersatz scenes used in visual cognition experiments include simple arbitrary arrays of objects (either in the environment or in depictions) with no natural structure or semantic interpretation. To the extent that such exhortations are persuasive, we have urged investigators to reserve the term *scene* for true scenes, because there are likely to be important differences in how true scenes versus (for example) object arrays are processed visually and cognitively. For example, as argued in several of the papers in this issue, true real-world scenes are likely identified as coherent meaningful entities using global image properties that transcend the specific objects in them. Arrays, in contrast, have no route to semantic interpretation beyond that which can be inferred from the identities of the objects they contain, and so cannot benefit from this additional global level of visual analysis. In addition, perceptual factors such as the functional field of view are likely to differ for objects in simple arrays and in scenes because scenes are semantically denser and also because objects in scenes tend to obscure each other via lateral masking. Scene depictions capture some of the important properties of real-world scene perception by including what are thought to be the important properties of scenes such as visual complexity, structural and semantic constraint, and meaning, while allowing for control over factors that would be difficult if not impossible to hold constant in the actual environment. At the same time, scene depictions such as drawings or photographs necessarily reduce the amount of information normally available to the visual system.

The current volume brings together an eclectic group of investigators, all of whom study critical issues in the perception of true real-world scenes. Topics include the rapid acquisition of scene gist, scene recognition, spatial layout and spatial scale, distance perception in scenes, updating of scene views over time, visual search for meaningful objects in scenes, scene context effects on object perception, scene representation in memory, the allocation of attention including eye fixations during scene viewing, and the neural implementation of these representations and processes in the brain. Because the study of real-world scene perception benefits from an interdisciplinary approach, contributors to this special issue use a variety of research methods including psychophysical and behavioural techniques, eyetracking, functional neuroimaging (including fMRI and ERP), and mathematical and computational modelling. While much has been learned from studying simplified visual stimuli, many of the papers in this special issue make the important point that understanding the functional and neural architectures of the visual system requires studying how that system operates when faced with the types of real-world stimuli that evolution crafted it to handle.

REFERENCES

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*, 143–177.

Henderson, J. M. (1992). Object identification in context: The visual processing of natural scenes. *Canadian Journal of Psychology*, *46*, 319–341.

Henderson, J. M., & Ferreira, F. (2004). Scene perception for psycholinguists. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, *50*, 243–271.