

Combining Information to Perceive 3d Layout

Martin S. Banks

Vision Science, University of California, Berkeley,
360 Minor Hall, Berkeley, CA 94720-2020, USA

A fundamental problem faced by the visual system is providing information about the 3d environment from the 2d retinal images. Several sources of information--usually called depth cues--are used in this inference process. We have examined the means by which various cues are used to infer 3d structure. The combination of disparity and texture cues to a surface's slant is well modeled by weighted summation. The combination approaches statistical optimality and thereby minimizes the uncertainty of the slant estimate. We find that the system behaves robustly when the conflict between disparity and texture is large. We have also examined the combination of disparity, a metric depth cue, and configuration, a nominally ordinal cue. We find that configuration cues provide more than ordinal depth information and that the combination of disparity and configuration approaches optimality considering the natural statistics of the environment.