

Combining Visual Information to Perceive 3d Layout

Martin S. Banks

University of California at Berkeley USA

November 28, 2005 at 15:00

Abstract: 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 binocular disparity (differences in the two eyes' images) and the texture gradient (perspective) are used to infer 3d structure. The combination of disparity and texture cues to a surfaces slant is well modeled by weighted summation. The combination approaches statistical optimality and thereby minimizes the uncertainty of the slant estimate. We find that disparity and texture slant cues are also combined subtractively for the purpose of assessing texture homogeneity; this hypothesized process is closely related to shape constancy .

