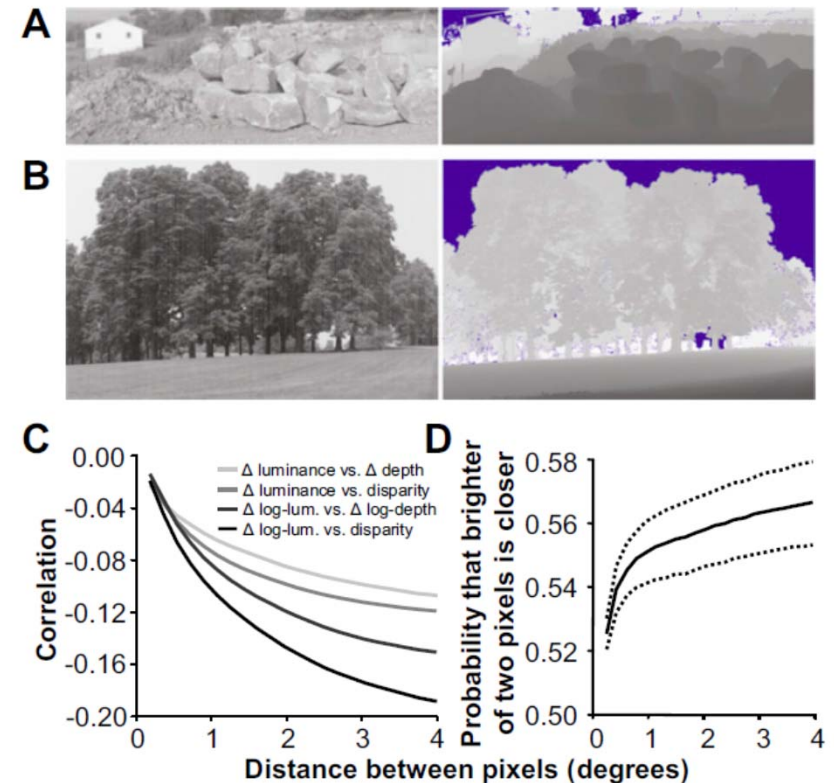


# Depth: Stereo and Lightness

Alan Yuille

# Lightness Depth Cues

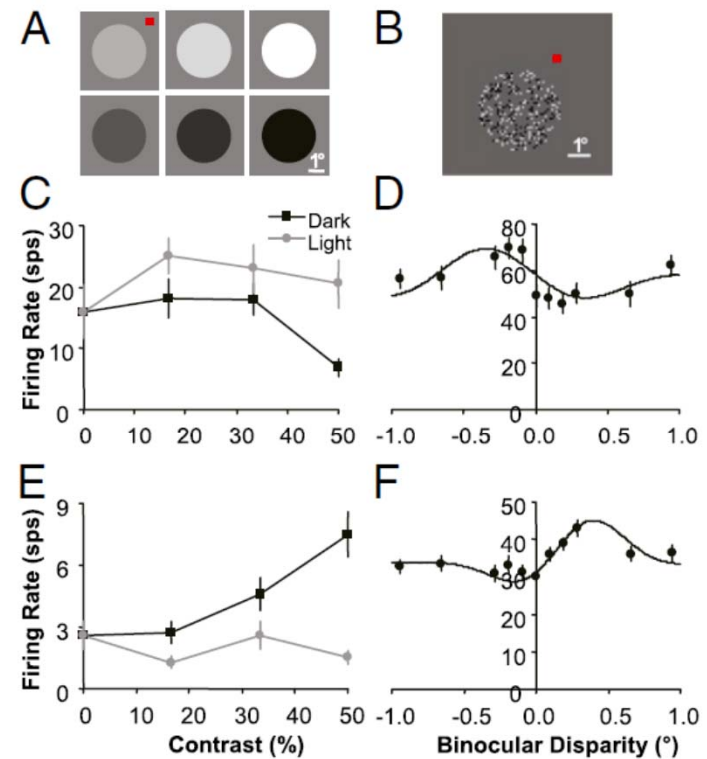
- Da Vinci.
- Sammonds, Potetz, Lee (PNAS)



**Fig. 1.** Examples of the correlation between intensity and depth in natural scenes. (A) Shadows (Left) are produced in the crevasses (range data, Right) of piles of objects. Lighter means farther for the range data images (blue regions were not recorded because they were beyond the range of the scanner). (B) There are also shadows seen in the more distant sections of foliage. (C) Correlation between relative luminance and relative depth using four different metrics (*Materials and Methods*). (D) Given two pixels, the one

# Neuron Response: Tuned to lightness+ depth

- Sammonds, Potetz, Lee.



**Fig. 2.** Single-neuron examples of relative luminance and binocular disparity tuning. (A) Disk stimuli used to measure luminance tuning. (B) DRDS stimuli used to measure binocular disparity tuning (image for one eye shown). (C and D) A neuron that responds better to light surfaces and near

# Neg Correlations

- Sammonds, Potetz, Lee.

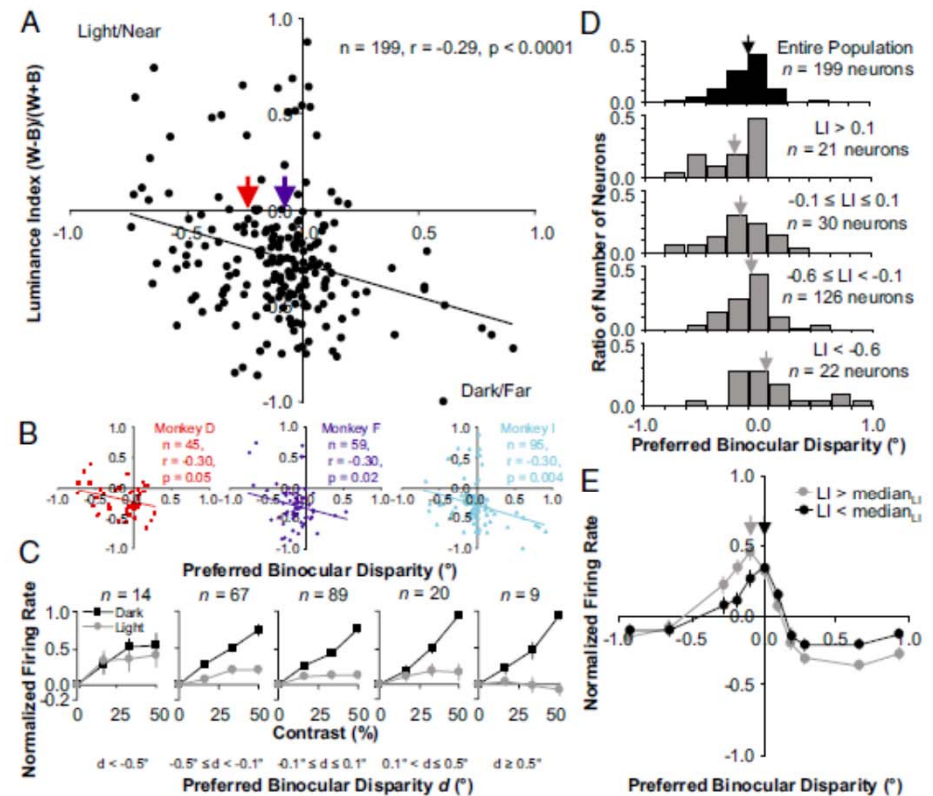


Fig. 3. Negative correlation between luminance index and preferred binocular disparity. (A) Aggregate scatter plot of measurements made from neurons recorded from all three monkeys. (B) Scatter plot of measurements made from neurons recorded from each monkey individually. (C) Population average of contrast response curves for light and dark disks. Each plot uses a different group of neurons on the basis of their preferred disparity  $d$ . The response to a blank gray screen (0%) was subtracted from the contrast response curve for each neuron and then the curve was normalized by the peak firing rate before averaging. Error bars are population SE. (D) Histograms of preferred binocular disparity. Each histogram uses a different group of neurons on the basis of their luminance index  $LI$ , ranging from preferring white to preferring black (top to bottom). Arrows are mean preferred binocular disparity. (E) Population average of disparity tuning curves. Each plot uses a different group of neurons on the basis of their luminance index  $LI$ . The mean response to all disparities was subtracted from the disparity tuning curve for each neuron and then the curve was normalized by the peak firing rate before averaging. Error bars are population SE.

# Sammonds, Potetz, Lee. Summary 1.

- In this article, we described the perceptual phenomenon of lighter surfaces appearing to be nearer than darker surfaces. We then described the correlations between image and 3D natural scene statistics, which might provide an ecological basis for this phenomenon. This correspondence suggests that the behavior is evidence of a statistical trend that humans make use of when inferring 3D shape in images. Most previous studies that sought to understand the visual system by analyzing the statistics of natural scenes have explained neurophysiological properties that were already well known, such as the center-surround antagonistic and wavelet-like receptive field structures (2–4), as well as contextual modulation of receptive field responses to contour segments outside the classical receptive field (6, 7, 39). Our study is one of a few studies (5) that instead confirm a prediction made by theoretical studies of natural scenes, using neurophysiological experiments.



## Sammonds, Potetz, Lee; Summary 2

Within a given image region, darker surfaces are more likely to be part of a shadow and are thus more likely to be farther away than nearby lighter surfaces. The comparative statistics (Fig. 1) illustrate the tendency for shadowed regions to lie farther from the observer. In addition, these comparative statistics are related to response properties of neurons in V1. Neural responses to luminance in V1 are relative to both the absolute intensity of light striking the retina and also the local relative intensity of the region, due to the center-surround receptive field structure found throughout the early visual system. Likewise, neurons are selective to absolute binocular disparity, which is not a measure of absolute depth from the observer. It is instead a measure of relative depth from the fixation plane, which is commonly focused to minimize stereo disparity for the object fixated at the fovea.

# Sammonds, Potetz, Lee: Summary 3

- With the neurophysiology experiments, we demonstrated that there is a significant negative correlation between relative luminance preference and preferred binocular disparity among a population of V1 neurons (Fig. 3A). Neurons that respond to near binocular disparities also respond relatively better to lighter disks compared with darker disks than neurons that respond to far disparities. The negative correlation observed is invariant to changes in several disk and aperture sizes that we tested (Figs. S3 and S4). The trend is also clear in the population averages of light and dark contrast response curves as the composition of neurons in each population varied in tuning from near to far preferred binocular disparities (Fig. 3C). Regardless of how we defined the relative luminance preference or ratio between the responses to light disks vs. far disks, we always observed a significant negative correlation so the trend did not depend on any specific choice of a luminance index (SI Text). Overall, the neurophysiological results were robust and consistent with the prediction derived from the analysis of natural scene statistics.

## Sammonds, Potetz, Lee: Summary 4.

- The subtle shifts in disparity estimated from our population response in V1 due to relative luminance preference (Fig. 3 A, D, and E) are comparable to those measured in perception. For example, in psychophysical experiments, a white disk on a gray background at a distance of 125 cm is perceived as 5 cm nearer than a black disk on a gray background (18). Similarly the shift in preferred disparity that we measured between disparity-tuned neurons preferring white and black disks (Fig. 3A, red and blue arrows) would correspond to a difference in depth of 10 cm at distance of 125 cm. The association we measured does not imply that the actual depth-decoding process, including the neural correlate of the perceptual bias, would be located in V1. Because discs were presented binocularly and the disparity of the disk was not ambiguous, the bias would not play a role in the subject's perception (18). More sophisticated experiments would be necessary to uncover the neural correlate of the perceptual bias. Our results reveal that the association in V1 is a possible component of the mechanism of depth inference rather than a by-product of the perception.



## Sammonds, Potetz, Lee: summary 5

Determining depth and identifying 3D shape from images is a difficult problem that our visual system handles very efficiently. Features, structures, and patterns in an image can have numerous potential 3D interpretations, which necessitates that depth perception is solved by inference using a multitude of visual cues to gather as much evidence as possible. Using inference as our foundation, we have approached the issue of full-cue depth perception by first understanding the statistical relationships between images and depth to formulate hypotheses for neurophysiological experiments. We have now identified the neurophysiological basis for one form of cue coupling as early as the primary visual cortex. However, understanding the full scope of depth-cue integration requires studies extending throughout the visual hierarchy. In this work, we have focused on the link between relative intensity and relative depth, a powerful form of cue coupling, and elucidated how this statistical trend in natural scenes might be encoded in a neuronal population in V1 to support the perceptual inference of depth.