Modeling the Joint Distribution of Scene Events at an Edge

Edges in an image arise from discontinuities in scene variables, namely reflectance (R), illumination (I), depth (D) and surface orientation (O). Early visual coding in primate is selective for edges and edge detection forms an early stage for many computer vision algorithms. However, image interpretation depends not only on detection, but *discrimination* of the scene variables that gave rise to the edge.

Prior human and computer vision studies suggest that edge discrimination is a more difficult problem than edge detection. These studies have focused primarily on discriminating depth edges from other forms of edge, although illumination (shadow) edges have also been examined. Typically, edge discrimination is viewed as a binary classification problem: each edge is assumed to arise from one of two disjoint categories (e.g., depth or not depth, shadow or not shadow).

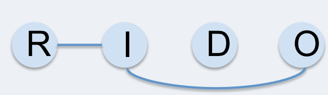
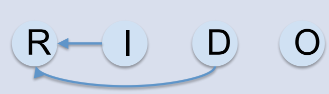
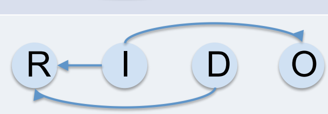
Here we suggest an alternate view in which an edge may signal discontinuities in any combination of the scene variables (RIDO). While this generates up to 16 disjoint categories, we note that depth and surface orientation changes are only visible if they entail changes in reflectance or illumination. Thus restricting our attention to visible edges (excluding modal and amodal completions) imposes the constraint that $R \vee I = 1$, reducing the number of feasible categories to 12.

To explore this model, we had 4 trained observers label one randomly selected edge in each of 1,000 randomly selected images drawn from the McGill Calibrated Colour Image Database and from Flickr; the averaged data provide an estimate of the joint distribution over the 4 binary variables RIDO (Table 1).

We first observe that the distribution is far from uniform, suggesting that accurate modeling of this distribution might improve inference. While the full distribution could be modeled exactly, identifying independence relationships between the 4 variables RIDO could yield insight into the problem.

We consider an exhaustive set of 42 undirected and 72 directed graphical models of low-order (<9 parameters); Figure 1 shows the model with lowest KL divergence from the full empirical distribution, for each level of complexity. A number of qualitative observations emerge. The models first link the RI variables, capturing the $R \vee I = 1$ constraint but also the fact that the two variables change together less often than would be predicted by independence – it's usually one or the other. The IO variables are then linked, which captures the fact that illumination is more likely to be changing if surface orientation is changing. More subtle relations are most efficiently captured by directed models. For example, conditioning R on D captures the increased likelihood of reflectance change at occlusion boundaries.

R	I	D	O	Prob.
1	0	0	0	0.20
1	0	1	0	0.16
1	1	1	0	0.16
0	1	0	1	0.10
0	1	0	0	0.08
1	1	1	1	0.08
1	0	1	1	0.07
0	1	1	0	0.05
1	0	1	0	0.03
0	1	1	1	0.03
1	1	0	0	0.02
1	1	0	1	0.02

Model	Parameters	KL Divergence
	5	0.19
	6	0.16
	7	0.08
	8	0.04