From *AAAI-90 Workshop on Qualitative Vision*, July 20, 1990, Boston, MA.

Ordinal characteristics of transparency.

Edward H. Adelson* and P. Anandan**

*Media Laboratory and Department of Brain and Cognitive Sciences, MIT, Cambridge MA 02139, and **Department of Computer Science, Yale University, New Haven CT 06520

Figure 1 shows an example of visual transparency. The image could arise from a number of different physical causes. For example, a square of tissue paper could be in front of a dark grey circle; or a circular shadow could be cast on a plane containing a light grey square; or a dark circular filter could be lying on top of a light grey square. Although the physics is uncertain, one can perceive the image as a combination of two more primitive images.

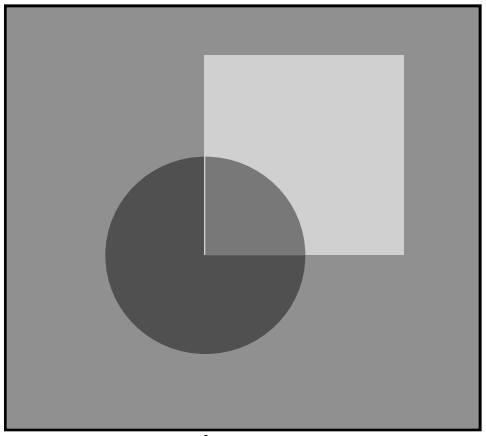

Figure 1

We use the term "transparency" to cover the general case of such image combination, including what would be called "translucency" in ordinary language. Many physical phenomena can produce transparency. For example, dark filters, specular reflections, puffs of smoke, gauze curtains, and cast shadows, all combine with patterns behind them in a transparent manner.

When an image has been formed by the combination of two primitive images, then it is usually more parsimonious to describe the image in terms of that combination; thus it is advantageous for a visual system to parse the image into the primitive images along with a combination rule. This parsimony does not depend on assigning a unique physical interpretation to the primitive images; figure 1 can be parsed into a circle and a square, even in the absence of a decision about the underlying physics.

We suggest that visual transparency may be initially analysed at a "pre-physical" level, which does not include the physical specificity of a full intrinsic image analysis [1]. The representation at this level consists of a set of primitive image layers which are ordered in depth. Each layer contains filled regions which modify the appearance of the layers beneath them, and unfilled regions which are perfectly clear. The filled regions of different layers combine with

each other according to simple rules such as multiplication and addition. Figure 2 shows an example of the layers which might give rise to figure 1. In the simplest layered model luminance only propogates from the lower layers to the higher ones (i.e. toward the viewer).

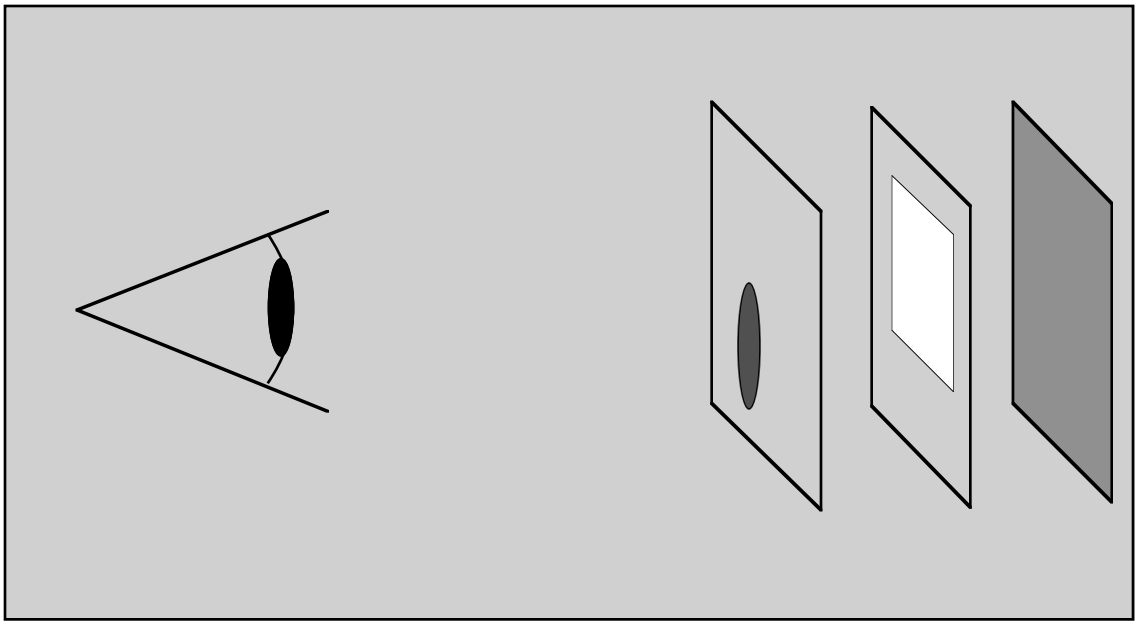

Figure 2

<u>The significance of X junctions.</u>

When patterns on distinct layers overlap, they typically give rise to X junctions in the image, which have an important influence on the perception of transparency by humans [2 - 4]. These X junctions can be quite diagnostic of the nature of the transparent interaction, and the depth ordering of layers. For example, figures 3(a-c) contain three images, which the human visual system interprets in three different ways. Transparency can be seen in figure 3(a), which is interpreted as containing two dark filters; however, the depth ordering is ambiguous: either square can be seen as lying in front of the other. Transparency can also be seen in figure 3(b), but in this case the depth ordering is unambiguous: the square on the lower-left appears to be in front. Finally, transparency is not seen in figure 3(c), which is commonly seen as a painted pattern lying in a single layer.

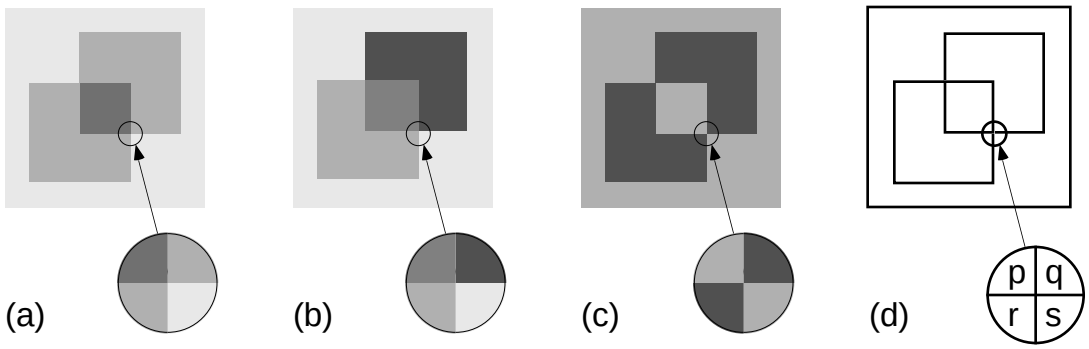

Figure 3

The qualitative characteristics of the transparency can be related to the qualitative characteristics of the X junctions. Let *p,q,r,s* be the luminances in the four regions surrounding the X junction, as indicated in figure 3(d). In figure 3(a) $p<q$ and $r<s$, which is to say that the vertical edge retains the same sign in both halves of the X junction. Similarly, $p<r$ and $q<s$, which is to say that horizontal edge, also retains the same sign in both halves of the X junction. We call this a "non-reversing" junction because both edges retain their sign. In figure 3(b) the vertical edge changes sign across the X junction, whereas the horizontal edge retains its sign. We call this a "single-reversing" junction. Finally, in figure 3(c) both the horizontal and the vertical edges change sign across the X junction. We call this a "double-reversing" junction.

The human visual system seems to employ heuristics related to these different categories of X junctions. Non-reversing junctions support the perception of transparency, while leaving the depth ordering of the layers ambiguous. Single-reversing junctions also support transparency, and in addition impose a unique depth ordering. Double-reversing junctions do not support transparency. The junctions thus offer pieces of local evidence which may be propogated through the figure to the interpretation of transparency and depth order.

<u>Computational analysis</u>

We may examine transparency from a computational point of view, to understand the basis for the heuristics described above. We begin with a framework to characterize the combination of transparent layers; the layers will be denoted $I_1 \ldots I_n$. Each layer may attenuate the luminance from the layer beneath it by a factor a, $0 < a \leq 1$, and may contribute its own emission of quantity e, $e \geq 0$. The attenuation and emission are functions of position, $a(x,y)$ and $e(x,y)$. An unfilled region has $a = 1$ and $e = 0$. (This formulation is slightly different from Metelli's, but the resulting restrictions are similar to those derived from Metelli's rules).

If layer n-1 contains a luminance pattern $I_{n-1}(x,y)$, then the luminance pattern at layer n is:

$$I_n(x,y) = a_n(x,y).I_{n-1}(x,y) + e_n(x,y) \qquad (1)$$

Since both multiplicative and additive interactions are allowed, a wide range of (p,q,r,s) values are legal examples of transparency. However, there are constraints on those values. The allowed ranges of a and e imply that a filled region must reduce or leave unchanged the amplitude of the luminance variation in a lower layer. This allows us to establish some inequalities concerning the X junction of figure 3(d). We assume that an X-junction results from the overlap of filled regions in two layers; it remains to determine whether the frontal layer's edge is vertical or horizontal, and which half of the edge is filled.

The four possible local hypotheses about the filled frontal region are: (i) it lies above the horizontal line, (ii) it lies below the horizontal line, (iii) it lies to the left of the vertical line, and (iv) it lies to the right of the vertical line. The conditions on the attenuation factor translate into the following inequality conditions,

(1) hypothesis (i) is physically plausible iff $0< (p-q)/(r-s) \leq 1$,
(2) hypothesis (ii) is plausible iff $0< (r-s)/(p-q) \leq 1$,
(3) hypothesis (iii) is plausible iff $0 < (p-r)/(q-s) \leq 1$, and
(4) hypothesis (iv) is plausible iff $0 < (q-s)/(p-r) \leq 1$.

Note that conditions (i) and (ii) are mutually exclusive unless the ratio is unity, likewise for conditions (iii) and (iv).

The fact that these ratios are non-negative leads to the edge-reversal heuristics noted above. Thus an edge which is tranparently occluded cannot reverse sign, while an edge which is in front may or may not reverse sign. It follows that double-reversing junctions have two consistent interpretations, and single-reversing junctions have only one. A double-reversing junction would require that both the vertical and horizontal edges be in front of the other, which is impossible; therefore no transparent interpretation is allowed.

Conclusion

Transparency can arise in images due to a number of different physical phenomena. We have proposed a pre-physical level of representation in which a number primitive images organized as layers combine together to form an observed image. The ordinal relationships between the luminances at an X junction can be used categorize the X junction as non-,single-,and double-reversing junctions. These categories can be determined without precise measurements, and are robust against point nonlinearities in luminance sensitivity. Non- and single-reversing junctions support transparency; single-reversing junctions lead to an unambiguous interpretation of depth-order of the layers, while non-reversing junctions leave the depth-order ambiguous. Double-reversing junctions do not support transparency. Propogation of these constraints can be used to rapidly restrict the set of the legal interpretations of an image.

References:

[1] Barrow, H. G., and Tenenbaum, J. M., "Recovering Intrinsic Scene Characteristics from Images." In Hanson, A. R., and Riseman, E. M., **Computer Vision Systems**, 3-26 (1978).

[2] Beck, J., Prazdny, K., and Ivry, R., "The perception of transparency with achromatic colors." *Perception and Psychophysics*, **35**, 407-422 (1984).

[3] Kersten, D., "Transparency and the Cooperative Computation of Scene Attributes", to appear in Landy, M., and Movshon, J., eds., **Computational Models of Visual Processing**, MIT Press, 1990.

[4] Metelli, F., "The perception of transparency," *Scientific American*, **230**, (4), 91-98