Available online at www.sciencedirect.com
SCIENCE DIRECT
Vision Research 46 (2006) 1982–1995
Vision Research
www.elsevier.com/locate/visres

The perceived transmittance of inhomogeneous surfaces and media

Barton L. Anderson [a,*], Manish Singh [b], Jeanette Meng [c]

[a] University of New South Wales, Sydney, Australia
[b] Department of Psychology and Center for Cognitive Science, Rutgers University, New Brunswick Campus, Piscataway, NJ, USA
[c] Massachusetts Institute of Technology, Cambridge, MA, USA

Received 19 August 2005; received in revised form 7 November 2005

Abstract

A series of experiments was performed to determine how the visual system computes the transmittance of inhomogeneous surfaces and media. Previous work (Anderson, B. L. (1999) Stereoscopic surface perception. *Neuron, 26*, 919–928; Anderson, B. L. (2003) The role of occlusion in the perception of depth, lightness, and opacity. *Psychological Review, 110*, 762–784) has suggested that the visual system employs a *transmittance anchoring principle* in determining when transparency is perceived. This principle states that the visual system interprets the highest contrast region along contours and surfaces as a region in plain view and uses this anchor as a reference point for transparency computations. In particular, recent work has shown that the transmittance of homogeneous transparent surfaces is well described by a ratio of contrasts model (Singh, M., & Anderson, B. L. (2002). Toward a perceptual theory of transparency. *Psychological Review, 109*, 492–519). In this model, the transmittance of a transparent surface is determined by the contrast of a transparent image region normalized by the contrast of the region in plain view. Here, a series of experiments is reported that assesses this model for inhomogeneous transparent surfaces that vary in both space and time. The results of these experiments reveal that transmittance anchoring has both a spatial and temporal component, and that the perceived transmittance of transparent surfaces is well described by a ratio of perceived contrasts model.

Keywords: Depth; Transparency; Surface perception; Opacity; Anchoring

1. Introduction

A fundamental problem in vision science is determining how the visual system decomposes images into the separate physical causes that are responsible for their formation. One of the most phenomenally salient forms of decomposition occurs in the perception of transparency. To recover scene structure in such contexts, the visual system must decompose the image along common lines of sight and compute the properties of both the transparent media and the underlying surfaces. During the past few decades, a growing body of data has shed light on how the visual system accomplishes this complex task. Metelli's (1970, 1974; Metelli et al., 1985) pioneering work provided a simple and elegant physical model of the image properties that must be present for an image to be consistent with transparency, and inspired a host of studies into the perceptual abilities that support the perception of transparency (Anderson, 1997, 1999, 2003; Beck & Ivry, 1988; Beck, Prazdny, & Ivry, 1984; D'Zmura, Colantoni, Knoblauch, & Laget, 1997; Faul & Ekroll, 2002; Gerbino, Stultiens, Troost, & de Weert, 1990; Kasrai & Kingdom, 2001; Robillotto, Khang, & Zaidi, 2002; Robillotto & Zaidi, 2004; Singh, 2004; Singh & Anderson, 2002a, 2002b). Metelli's model was based on a simple physical setup: a two-toned background visible through a rotating fan blade (an episcotister). Metelli derived simple algebraic relationships between the reflectance values generated by surfaces in plain view, to the reflectance and transmittance values of the regions covered by the episcotister. The majority of these studies concluded that the perception of transpar-

* Corresponding author. Tel.: +61 2 9385 2441.
E-mail address: bart.a@unsw.edu.au (B.L. Anderson).

0042-6989/$ - see front matter
doi:10.1016/j.visres.2005.11.024

ency is well predicted by Metelli's model, suggesting that the visual system essentially inverts the image generation process to recover scene structure under conditions of transparency.

More recently, however, work from our (Singh & Anderson, 2000, 2002a, 2002b, 2005) and subsequently other labs (Robillotto et al., 2002; Robillotto & Zaidi, 2004) has revealed some fundamental failures of Metelli's model. Although Metelli derived separate expressions for the transmittance and reflectance of a homogenous (or "balanced") transparent surface, none of the previous perceptual experiments performed to assess his model required observers to independently judge these two properties. Recently, however, we experimentally uncoupled these dimensions and had observers perform separate judgments of the transmittance and lightness of transparent surfaces. These experiments revealed large and systematic departures from the predictions of Metelli's model. Our most robust finding involved the perception of surface transmittance in transparency displays with sinusoidal textures: whereas Metelli's model predicts the transmittance of transparent surface is determined by the ratio of luminance differences between the region of transparency to the surfaces in plain view, our work revealed that perceived transmittance scaled instead with Michelson contrast—thereby indicating that Metelli's formula involving the ratio of luminance differences should be replaced with the ratio of Michelson contrasts. Similar results were subsequently obtained by Robillotto et al. (2002) and Robillotto and Zaidi (2004). However, in their displays, Michelson contrast (and other tested measures of contrast) failed to capture the perceived contrast of their targets; nonetheless, they found that *perceived* contrast did a good job in capturing perceived transmittance of their displays. It is important to emphasize that the results from both groups demonstrated the inadequacy of Metelli's model in predicting human perception[1], and show that the visual system employs computational strategies that generate systematic errors in computing the properties of transparent surfaces. The importance of this for perceptual theory cannot be underestimated. Although it is common to treat visual perception as an "inverse optics" problem, these transparency results reveal at least one domain where such approaches fail to capture perceptual experience.

Within the domain of transparency perception, the finding that the visual system uses contrast to compute the transmittance of transparent layers raises a number of questions. All natural scenes generate contrast variations. How, then, does the visual system determine whether a scene is being viewed in plain view or through a transparent layer or medium? A simple thought experiment reveals the problem confronted by the visual system once the possibility of transparent interpretations is allowed. In principle, *any* image could arise from either viewing a scene in plain view, or by viewing a higher-contrast scene through a contrast-reducing transparent layer. How, then, does the visual system determine when a transparent surface is present? More specifically, what image properties must be present for the visual system to infer the presence of a transparent layer? One of us (Anderson, 1999, 2003) has recently proposed that the visual system employs a *transmittance anchoring principle* (TAP) to compute transparency. The intuition shaping the TAP is that the visual system embodies a bias to assume the fewest surfaces necessary to account for the image data. It is, in a sense, an "Occam's razor" in image interpretation. Specifically, the TAP states that the visual system treats the highest-contrast regions along contours (and surfaces) as regions in plain view, and uses contrast reductions along contours or surfaces to compute the presence of an overlying transparent surface[2]. In this principle, the highest-contrast contour segment (or surface region) serves as an "anchor" that is interpreted as a surface region in plain view. All other contrast values along the contour or surface are compared to this anchor region, and reductions in contrast relative to the anchor value are used to infer the presence of a transparent layer.

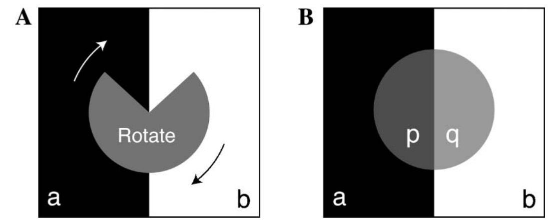

Fig. 1. A depiction of Metelli's episcotister display. (A) A disk with a missing sector is rapidly rotated over a two-toned background, generating a four-toned image that generates a percept of a transparent layer similar to that depicted in (B).

To understand the role of the TAP in predicting when transparency is or is not perceived, consider a variant of the display generated by Metelli's episcotister model (Fig. 1). This stimulus generates four luminance values, labeled *a*, *b*, *p*, and *q*. Metelli's model was based on knowledge of his physical setup: reflectances *p* and *q* were written as a weighted sum of the reflectances of the background and the episcotister, whereas regions *a* and *b* are the reflectances of the surfaces in plain view. Metelli's model predicts that transparency of the central disk should only occur when the reflectance difference $p–q$ has the same sign as, and a magnitude no greater than,

[1] Ironically, Metelli (1974a) himself reported a significant failure of his model in explaining perception. He noted that a light episcotister appears less transmissive than a dark episcotister even when they have identical physical transmittances. However, no further mention was made about the negative impact this fact had on the veracity of his model.

[2] Strictly speaking, transparent surfaces or media need not cause a reduction in the contrast of underlying features. This occurs when the transparent surface does not contribute any luminance to the image. Note, however, that this is a degenerate case that is physically equivalent to changes in illumination (such as shadows).

the difference *a–b*. Simply stated, these rules imply that transparent layers cannot reverse the polarity of underlying contours, and can only reduce (or leave unchanged) the luminance difference across the underlying contour. Metelli's model focused on predicting when the central region in his display appears to contain a transparent layer, but it remains mute in explaining why transparency is never reported in his displays in *both* the center and the surround. Indeed, it is perfectly possible that different transparent layers exist in the center and surround simultaneously, yet no explanation is offered for why this is never perceived. The TAP fills this conceptual gap. Unlike Metelli's model, the TAP is a principle of perceptual inference, not a description of a particular physical setup, and makes an explicit statement about why transparency is not a ubiquitous percept in all images, or all image regions. It states that the highest-contrast region of an image serves as a transmittance anchor, and hence, predicts that there should always be regions in an image that are perceived in plain view.

In its original formulation, the TAP was applied to static images containing sharply defined contours that separated regions of transparency and regions in plain view (Anderson, 1999, 2003). However, the anchoring principle is not restricted to a particular geometric context, and can be extended to more complex scenes than those described by Metelli's model, or related models that emphasize the ordinal luminance relationships across contours and contour junctions (Adelson & Anandan, 1990; Anderson, 1997; Beck & Ivry, 1988). The TAP only requires the presence of polarity-preserving contrast variations along surfaces or contours to compute transparency, and asserts that the highest contrast along a contour or surface provides a quantitative reference point that can be used to compute the transmittance of transparent layers. So stated, this principle can be used to predict when transparency will or will not be perceived in images depicting *any* transparent or surface media, including inhomogeneous media such as smoke or fog. Moreover, our recent computational model of transparency (Singh & Anderson, 2002a) allows us to quantitatively predict the perceived transmittance of any transparent surface: it states that the perceived transmittance of a transparent region should equal the ratio of the contrast of the transparent region to the contrast of the region seen in plain view. Although recent work has shown that the TAP is capable of qualitatively explaining percepts of inhomogeneous transparency (Anderson, 1999, 2003), to our knowledge no previous study has explicitly measured the perceived opacity of continuously varying media. One of the goals of this paper is to fill this experimental gap.

Although the concept of transmittance anchoring was developed in the context of static images, there is nothing in this principle that is inherently limited to such contexts. Indeed, in natural scenes, transparent media can emerge or disappear in time. The question naturally arises as to whether the transmittance anchoring plays any role in the perception of transparency that varies in opacity over both time and space. In particular, can temporal reductions in image contrast induce percepts of transparency? If so, will the highest-contrast region in the temporal sequence serve as an anchor point that is perceived as a surface in plain view? The question is the second focus of this paper.

2. Preliminary observations and stimuli

To explore these questions, we generated a stimulus in which the spatial and temporal contrast could be manipulated in a continuous manner. To this end, a random-dot pattern was generated and its contrast was modulated in both space and time. The images consisted of a random-dot pattern modulated by a sinusoidal contrast envelope (similar to those initially introduced to study second-order motion; see Chubb & Sperling, 1988). To enhance the percept of transparency and facilitate observers' judgments of surface transmittance, binocular disparity was added to the contrast modulation so that it appeared in front of the random-dot pattern, and this contrast envelope was drifted over a (stationary) random-dot texture (see Fig. 2 for a static example). Our preliminary observations with this stimulus revealed two significant findings. First, consistent with the spatial form of the TAP, the region in the random-dot pattern containing the highest contrast was perceived as a surface region in plain view, *independent* of its absolute contrast. Second, if the contrast of the entire pattern was spatially modulated in time (such that all contrast values of the contrast-modulated random-dot pattern changed over time), the highest contrast region in this spatiotemporal sequence appeared as a surface region in plain view, and subsequent reductions in contrast were perceived as the appearance and disappearance of a partially transmissive medium into this region (i.e., regions that were initially clear appeared to become filled with fog). These preliminary observations are consistent with a temporal form of the TAP that states that the highest-contrast region in space *and* time is used to determine the presence or absence of transparent surfaces and media. The experiments that follow were performed to quantitatively assess these observations and our contrast-ratio model of perceived transmittance. To anticipate, we find that the TAP correctly predicts when transparency is and is not perceived, and that it has both a spatial and temporal component. However, in keeping with recent data (Robillotto et al., 2002; Robillotto & Zaidi, 2004), we find that the ratio of Michelson contrasts fails to capture perceived transmittance, because Michelson contrast does not adequately capture perceived contrast in our experimental displays. Nonetheless, the ratio of perceived contrasts (i.e., local perceived contrast normalized by the perceived contrast of the transmittance anchor) does an excellent job of quantitatively modeling the transmittance matches in our displays.

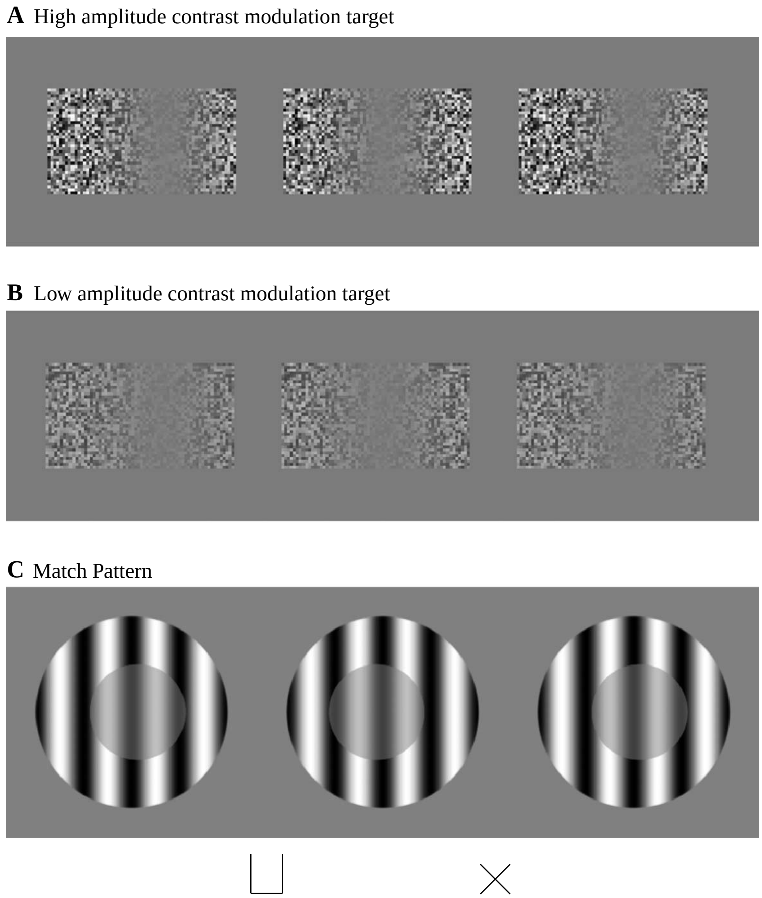

Fig. 2. The main stimuli and method used in the experiments reported here. (A) An 8-bit random-dot pattern is contrast-modulated. To enhance the percept of transparency, binocular disparity is added to the contrast envelope, and this envelope is translated across a stationary random-dot pattern at a constant speed (cross fusers should fuse the right two images, divergers the left two). This generated vivid percepts of a transparent layer that varied in opacity. (B) A low contrast variant of the stimulus. (C) An example of the matching pattern used to measure perceived transmittance. A sinusoidal grating containing a lower-contrast central patch was used as a matching pattern. The phase and spatial frequency of the patch was identical to the surround, but a disparity was added to the outer borders of the central patch to cause it to appear in front of the higher-contrast surround. Observers adjusted the luminance range of the central sinusoidal patch until it appeared to match the perceived transmittance of the region in the target.

3. Experiment 1: Spatial anchoring of inhomogeneous transparency

The purpose of Experiment 1 was to quantitatively measure perceived transmittance in stimuli that give rise to spatially inhomogeneous percepts of transparency. Previous experiments on the perception of transparency have used stimuli depicting homogenous transparent layers. In this experiment, we investigated the perception of transparency in stimuli that give rise to percepts of transparent surfaces that appear to vary in perceived transmittance. Our previous work (Singh & Anderson, 2002a) indicated that the visual system uses the ratio of contrasts between the region of transparency and the region seen in plain view to compute the transmittance of transparent surfaces; We therefore used patterns that were contrast modulated with a fixed mean luminance. These images were viewed binocularly, and an interocular phase offset was introduced in the contrast envelope to give it a disparity that caused it to appear clearly in front of the random-dot pattern. To further enhance the percept of transparency, the contrast modulation was spatially drifted across the random-dot pattern. These manipulations generated a vivid percept of an inhomogeneous transparent medium drifting across and in front of a random-dot background.

3.1. Methods

3.1.1. Observers

Three experienced psychophysical observers participated in the experiments, two of the authors (J.M. and M.S.), and one naive observer. All had normal or corrected-to-normal vision.

3.2. Apparatus

The experiments were performed using a Macintosh G4 (500 MHz, dual processor) using VisionShell software to generate the stimuli, control and execute each trial, and to record responses.

The stimuli were presented on a 22 in. CRT (LaCie electron 22blueII) viewed through a stereoscope at a distance of 100 cm. The resolution of the monitor was set to 1024×768, and the refresh rate to 85 Hz. Before each session, the monitor was calibrated by a LaCie blue eye calibrator. The gamma of the monitor was set to 1.0, and the gray levels were linearized from 0.7 cd/m^2 (minimum luminance or L_{min}) to 100.5 cd/m^2 (maximum luminance or L_{max}) and stored in an 8 bit lookup table. The laCie blue eye calibrator operates through a video cable that uses the DDC 2bi VESA standard, which allows a two-way information exchange between the Macintosh CPU and the electron22blue monitor. The three Red, Green, and Blue guns are calibrated individually inside the monitor.

3.3. Stimuli

The stimuli were contrast-modulated random-dot texture patterns. The contrast was varied in the horizontal direction, creating vertical bands of contrast modulation (Fig. 2). Two separate ranges of contrast were used that we will refer to as "high" and "low" amplitudes (defined below). Fig. 2A depicts the stimulus with a high-amplitude contrast modulation, and Fig. 2B shows the stimulus with the low-amplitude modulation. The target stimulus and the matching pattern were viewed against a uniform gray background of luminance 50.6 cd/m^2.

The target consisted of a rectangular 8-bit random-dot texture patch that subtended 4.8° horizontally and 2.5° vertically. Each texture element was first randomly assigned a luminance value L_{pix} by sampling from a uniform-distribution between 0.7 cd/m^2 ("black") and 100.5 cd/m^2 ("white"). The contrast of the texture was then modulated in a sinusoidal fashion, using a contrast envelope defined by:

$$\mathrm{Env}(x,t) = \mathit{min\_contrast}/2 + (A/4) * [1 + \sin(f * x + \theta(t))]. \quad (1)$$

Here x is the horizontal location of a texture element; f is the frequency of the sinusoidal envelope; $\theta(t)$ is the phase of the envelope (which depends on time t) and A is its amplitude; *min_contrast* is the trough of the envelope, which was set to .05 (or 5% contrast). The amplitude A was set to ones of two values: 0.9 (for the "high" amp.) or 0.4 (for the "low" amp.). The spatial frequency f of the contrast modulation was 0.2083 c/deg of visual angle, which resulted in one complete cycle of modulation along the length of the target. For convenience, we will express the horizontal location x in radians with $0 \leqslant x \leqslant \pi$ (i.e., in terms of an element's location along the sinusoidal cycle of modulation). The transformation applied to the luminance values $\mathrm{L}_{\mathrm{pix}}$ of the uniform-distribution texture in order to generate the contrast-modulated display was thus given by:

$$\Phi(L_{\mathrm{pix}})_{x,t} = L_{\mathrm{mean}} + 2 * (L_{\mathrm{pix}} - L_{\mathrm{mean}}) * \mathrm{Env}(x,t), \quad (2)$$

where $L_{\mathrm{mean}} = 50.6$ cd/m^2. The vertical gray band in the targets (see Fig. 2) corresponds to the trough of the sinusoidal modulation where the contrast is lowest (5%). The location where the texture elements have the highest contrast (95% or 45%) corresponds to the peak of the modulation. For each frame (235 ms), the phase θ decreased by $\pi/24$ radians, leading to a shift of 0.1° of visual angle to the left. The contrast modulation thus drifted leftward (2nd order motion) as the frames were shown in succession, whereas the underlying random-dot pattern remained stationary. The drift speed was 0.43° of visual angle per second.

A phase shift of 0.35π was introduced to the sinusoidal contrast modulation between the left and the right eye, thereby generating a near disparity of 50.4 min of arc. When viewing the binocular image sequence through a stereoscope, observers perceived a layer of inhomogeneous transparent medium drifting to the left over a stationary higher-contrast random-dot texture background.

3.4. Task

The observers' task was to adjust the perceived transmittance of a homogeneous transparent filter to match that of the inhomogeneous transparent targets at various locations (cf. Singh & Anderson, 2002a, 2002b). The matching pattern (Fig. 2C) consisted of two concentric circular discs of vertical sine wave gratings (0.8 c/deg) with identical orientations, spatial frequency, and phases. The diameters of the outer and inner discs subtended a visual angle of 5.0° and 2.5°, respectively. The contrast of the outer disc was set to the maximal contrast available to the monitor (99%). The contrast (or luminance range) of the inner disc was adjustable by moving a computer mouse. The mean luminance of the inner and the outer discs were the same (50.6 cd/m^2). The inner disc presented to the left and the right eye also had a π phase shift relative to the underlying texture. This caused the inner disc to appear to float in front of the outer disc when viewed through the stereoscope. Since the inner disc also had a lower-contrast, observers perceived it as a homogeneous transparent filter floating in front of a circular patch of high-contrast sine-wave grating.

At the beginning of each trial, a pair of vertical markers was presented, above and below the contrast-modulated random-dot pattern, indicating the horizontal location along the target where the observer was to match the transmittance of the transparent layer. The observers were instructed to adjust the luminance range within the small disc in the matching stimulus so that it appeared to have the same perceived transmittance as the transparent media floating between the markers in the target. Subjects were allowed to view the stimuli as long as needed to make the judgment. Sixteen locations along the horizontal extent of the sinusoidal contrast modulation were randomly presented twice in each session. Each observer completed eight experimental sessions: four sessions of high-amplitude contrast modulation, and four sessions of low-amplitude contrast modulation.

3.5. Results

The results of Experiment 1 are depicted in Fig. 3. Individual observer data were very similar, so the results were averaged across the three observers. Each data point thus represents the mean of 24 transmittance matches at each of 16 locations along the axis of the contrast modulation, and the error bars depict 95% confidence intervals. Each location is represented by its phase along the sinusoidal modulation. Fig. 3A shows the matches to the high-amplitude contrast modulation, and Fig. 3B shows the matches to the low-amplitude contrast modulation. The transmittance values plotted on the y-axis are defined as the ratio of Michelson contrasts between the small disc, set by the observers, divided by the surround (99%) in the match (Singh & Anderson, 2002a). The solid lines in the figure represent the sinusoidal contrast modulation of the random-dot pattern. Note that the contrast at the peaks of the modulation π differs by more than a factor of 2 in these two conditions (95% in a, and 45% in b). However, subjects consistently assign essentially identical transmittance values to these regions: 99.4% and 97.4%, respectively, demonstrating that they are not simply matching the contrast of the test and match patterns. In essence, then, observers report these contrast peak regions to be unobscured by a transparent surface region. We refer to these locations as the "Spatial Transmittance Anchors." They correspond to the highest-contrast regions along the surface, and in keeping with the TAP, they are perceived as being unobscured by an intervening transparent layer.

In addition to providing experimental support for the TAP, these data also provide information on how transmittance values are scaled relative to the transmittance anchor. The TAP predicts that transmittance should be computed relative to the highest contrast in the target. If perceived contrasts in this experimental context are determined by Michelson contrast, the ratio of contrasts model would predict that the perceived transmittance of transparent layer should simply follow the Michelson contrast of each target image normalized by the Michelson contrast of its respective anchor. The dotted curves in Figs. 3A and B depict the predicted transmittance values obtained by normalizing the modulated contrast (solid curves) by 95% and 45%, respectively (i.e., relative to the maximal contrast of each image). Note that the high and the low amplitude modulation (Figs. 3A and B) have very different absolute contrasts (solid curves), but the normalized transmittance predictions (dotted curves) are very similar. Indeed, the matching data from these two conditions are essentially identical, even though the target image's maximal contrast varied by a factor of 2.2. This provides unambiguous evidence that observers are not simply matching the contrast (or the individual luminance values) of the target and match patterns, as these would yield very different patterns for the 95% and 45% contrast conditions. Indeed, to match the perceived transmittance of the maximal contrast region of the low contrast target to the high contrast match, observers set the luminance range of the match to full contrast ("black" and "white"), whereas the pixel values in the target appeared to be two mid shades of grey.

Note, however, that apart from the spatial transmittance anchor where the contrast was at its peak, subjects consistently *overestimate* the transmittance of the media at all other locations (i.e., the transparent layer is perceived as more transmissive than predicted by the ratio

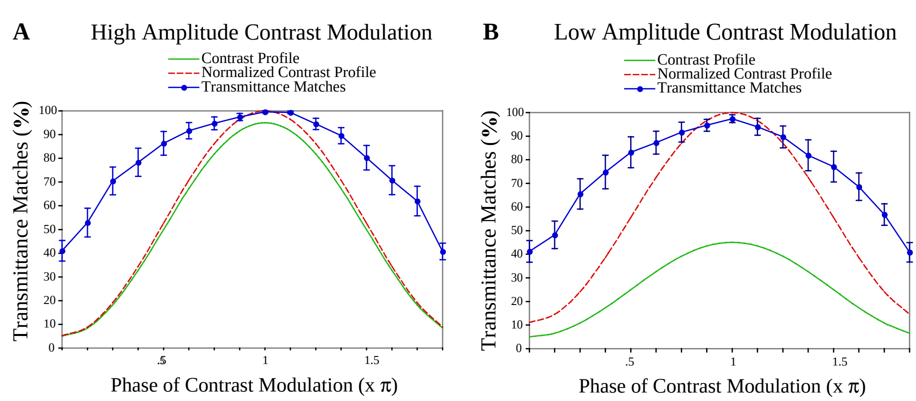

Fig. 3. Data from Experiment 1. The x-axis depicts the phase of the contrast modulation in radians. The peak corresponds to the highest contrast region in the target pattern. The y-axis depicts the transmittance values measured as the ratio of center contrast of the matching pattern, normalized by the contrast of the matching pattern's surround (hence this scale varies between 0 and 1). Two contrast ranges were tested, corresponding to the stimuli depicted in Figs. 2A and B. The solid green lines depict the contrast values in each display, and the dotting red lines depict the contrast values normalized by their peak contrast. Note that although the contrast ranges in (A) and (B) differed by more than a factor of 2, the matched transmittance curves were essentially identical, and both contrast peaks were perceived to have ~ 100% transmittance (i.e., they appeared in plain view). Note, however, that the perceived transmittance of all other regions of the test image were perceived as more transmissive than predicted by a ratio of Michelson contrasts model (dashed red curve). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

of Michelson contrasts). This overestimation increases as distance from the anchor increases. This pattern of overestimation is observed in every subject's individual data, and also in all subsequent experiments. We will explore the possible cause(s) of this pattern of transmittance matches in a series of control experiments described below.

In sum, the data from this experiment support the TAP's prediction that the highest-contrast region in an image is perceived as a region in plain view. But the results of this experiment also raise an intriguing question. When the low-amplitude contrast modulation was viewed within a separate spatial context, its highest-contrast region (45% contrast) was perceived to be fully transmissive (Fig. 3B). What will happen then if the low-amplitude contrast modulation is embedded in a temporal sequence of higher-amplitude contrast modulations? Will the perception of transmittance vary as a function of time, or will transmittance remain unaffected, and only the underlying surface lightness values appear to change? The purpose of Experiment 2 is to address this question.

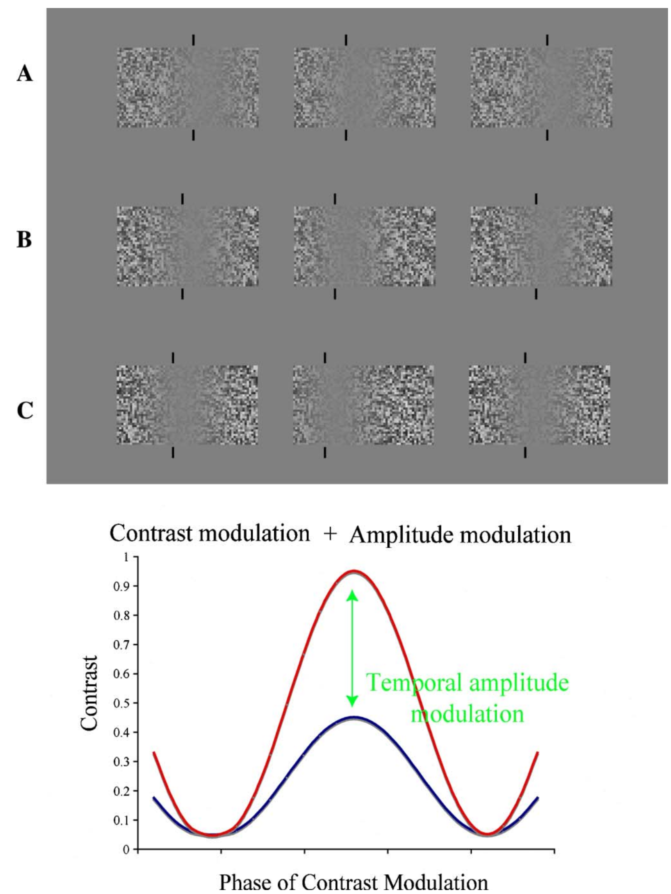

Fig. 4. The stimuli used in Experiment 2. The two contrast ranges used in Experiment 1 served as the endpoints of a temporal sequence in which the entire amplitude of the contrast envelope varied in time. Three steps in this temporal sequence are depicted in (A–C), and the two endpoints of the contrast envelope are depicted in (D). In this stimulus, observers report the appearance and disappearance of a transparent haze into the highest contrast region of the random-dot texture when the contrast of this region decreased and increased, respectively. Their task was to match the perceived transmittance of the target when the contrast profile was at its minimum.

4. Experiment 2: Temporal anchoring of perceived transmittance

Experiment 1 revealed that the highest-contrast region in an image is used to compute the surface region in plain view. In Experiment 2, the two stimuli used in Experiment 1 serve as the endpoints of a temporal sequence in which the contrast of the entire image is continuously modulated over time. The primary question investigated is whether the highest-contrast region in the lowest-contrast image will still appear as a surface in plain view (that changes its perceived lightness values), or whether the contrast changes will induce a percept in which the underlying surface appears to be partially obscured by a transparent surface.

4.1. Methods

4.1.1. Observers

The same three observers that participated in Experiment 1 served as observers in Experiment 2.

4.2. Stimuli

Fig. 4 depicts one of the eight amplitudes of contrast modulation used in this experiment. The visual angle of the target was 7.2° horizontally, and 2.5° vertically. The sinusoidal contrast modulation was the same as in Experiment 1, except that with each frame, the amplitude of the modulation of the entire pattern changes. The amplitude in Figs. 4A and C corresponds to the low-amplitude ($A_{min} = 0.4$) and the high-amplitude ($A_{max} = 0.9$) used in Experiment 1. The trough of the modulation was clamped at 5% contrast. Eight contrast amplitude values in total were used cycling between A_{max} and A_{min} with 0.0625 increments or decrements. For example, if the first frame had a contrast modulation amplitude of 0.4, the next frame would have a higher amplitude at 0.4625. The contrast amplitude would continue to increase in increments of 0.0625 with successive frames (235 ms/frame) until it reached A_{max}; it would then decrease in decrements of 0.0625. The contrast amplitude of the entire pattern was therefore varied as a triangle wave. It took a total of 3.76 s for a single cycle of contrast modulation (i.e., for the amplitude of the modulation to rise from A_{min} to A_{max}, and back down to A_{min}). Fig. 4D depicts the profile of the sinusoidal contrast modulations plus the amplitude modulation in time. Notice that as the amplitude was lowered in time, the slope along the contrast modulation becomes shallower. This stimulus change induced a percept of drifting bands of transparent media spreading vividly into the clear region every time the amplitude was lowered. Such intrusions receded every time the amplitude was increased. The drift speed of the contrast modulation was the same as in Experiment 1 (0.43° of visual angle per second).

4.3. Task

The observers' task was similar to that in Experiment 1. Observers adjusted the perceived transmittance of a small homogeneous filter in the center of the matching pattern to equal that of the media between the markers above and below the target. However, matches were restricted to the point in the temporal sequence of the contrast modulation where the amplitude was at its minimum ($A_{min} = 0.4$). This particular moment in time was also accompanied by an auditory tone to remind observers which time slice in the sequence they were to judge. Subjects were instructed to match the transmittance at the point in the sequence where the peak contrast was minimal, which would also be indicated by a tone. They were allowed to view as many cycles as needed until they were satisfied with their settings. Sixteen locations along the contrast modulation were repeated randomly twice in each session. Four sessions were completed by each observer.

4.4. Results

Fig. 5 shows the mean transmittance settings for the three observers when the amplitude of the contrast modulation was at its minimum. Sixteen locations along the contrast modulation were measured and the error bars represent 95% confidence intervals. The solid lines depict the contrast profiles of the maximum and minimum amplitude displays used in the temporal sequence. The upper dashed line depicts the transmittance match for the low-amplitude ($A_{min} = 0.4$) contrast modulation from Experiment 1 (same as in Fig. 3B). The critical test of the temporal version of the TAP occurred at the spatial contrast peak of the sinusoidal modulation. In this experiment, the matched transmittance dropped from 97.4% to 78.1%. Thus, by embedding the low contrast pattern in a temporal sequence of higher-contrast images, the perceived transmittance of the modulation was substantially reduced. This implies that transmittance is not only anchored to the highest-contrast region in space, but is also anchored to higher-contrast regions that precede it in time. We will return to the issue of how transmittance is scaled quantitatively in time below.

5. Experiment 3

To measure the time course of temporal anchoring, we repeated Experiment 2 and varied the time it took to cycle between the highest and lowest contrasts in the sequence (i.e., from A_{min} to A_{max}, and back to A_{min}). We used a faster cycle that took 0.94 s, and a slower one that took 22.56 s (and a 2 day delay marked "infinite" on the graph). Fig. 6 shows the resulting decrease in perceived transmittance that occurred by embedding the lower-contrast image in a temporal sequence of higher-contrast images, as a function of cycle time (averaged over the same three observers that participated in Experiments 1 and 2). We refer to this change in transmittance as the "temporal anchoring effect." As in the previous experiments, transmittance is defined as the ratio of Michelson-contrast of the matching pattern relative to its surround. It can be seen that temporal anchoring significantly decreases as the duration of the sequence increases. This makes intuitive sense; in the limit as the temporal change becomes infinitely slow, the images become equivalent to a series of static images, and temporal integration does not occur. In such contexts, the visual system treats the changes in contrast as arising from changes in the reflectance of the underlying surface rather than arising from the appearance and disappearance of partially transmissive media. Thus, when the change in contrast takes place over an extended time period, observers re-anchor their estimates of transmittance, which approach the

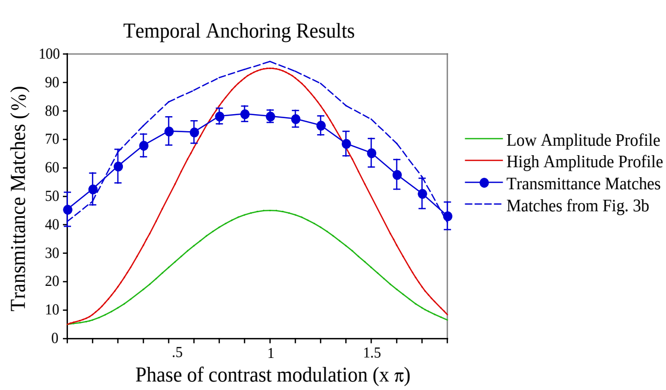

Fig. 5. The results of Experiment 2. The solid lines depict the two contrast ranges present at ends of the temporal modulation of the contrast envelope. The dashed line depicts the results of the data from Experiment 1 for the low contrast stimuli (i.e., the stimulus that was identical to the contrast values the observers judged in Experiment 2). Note the reduction (downward shift) in perceived transmittance of the target when it is embedded in a higher-contrast temporal sequence. We refer to this reduction in perceived transmittance as the "temporal anchoring effect."

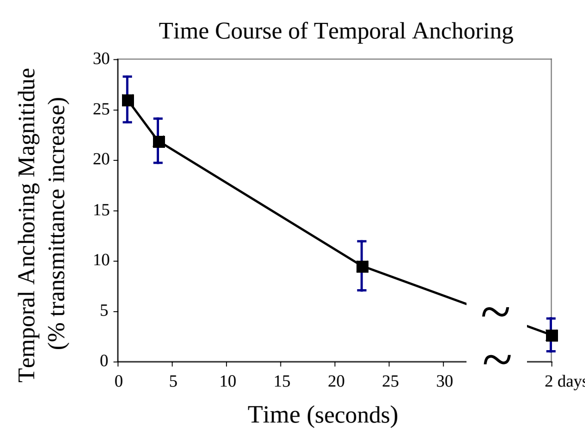

Fig. 6. The temporal anchoring effect measured as a function of time. Experiment 2 was repeated for a number of different temporal intervals of the (triangle-wave) oscillation between the highest and lowest contrast envelope. The data demonstrate that there is a clear dependence on the time taken for the contrast changes to occur, monotonically decreasing as the temporal interval of the contrast change increases.

values obtained in the low-amplitude contrast-modulation condition in Experiment 1.

In sum, the results of Experiments 1–3 reveal that the anchoring of perceived transmittance has both a spatial and temporal component. In Experiment 1, the highest-contrast region in space served as the transmittance anchor in these images, and in Experiments 2 and 3, we found that higher-contrast regions in the temporal sequence also had an impact on the perceived transmittance in the lowest contrast image (we will return to the problem of quantitatively estimating the amount of temporal anchoring that occurs below). However, the measured transmittance values at the lower-contrast locations in Experiment 1 raise the puzzle as to how transparency is scaled in these images. Singh and Anderson (2002a) showed that the perceived transmittance of a transparent filter layer with sharp edges, over a sinusoidal-grating pattern (such as the matching display used here) is well captured by the Michelson contrasts in the transparent region normalized by that of its surround. However, in the experiments reported here, we observed a systematic overestimation relative to these predictions for every value except the region of highest contrast predicted by the TAP to appear as a surface in plain view. What is responsible for these deviations? The experiments that follow are designed to answer this question.

5.1. Determining the cause of transmittance overestimation

There are two broad explanations for the pattern of transmittance matches observed in the first two experiments. One possibility is that deviations from the contrast-ratio model of transmittance arose because the contrast model fails to capture perceived transmittance in complex spatial patterns, i.e., patterns containing inhomogeneous percepts of transparency. Alternatively, the matching errors could be due to a misperception of contrast in *either* the target or in the matching stimulus (with respect to the Michelson-contrast prediction), giving rise to systematic errors in transmittance matches. We therefore performed a number of experiments to determine whether the pattern of transmittance matches observed in these experiments could be attributed to systematic differences in the perception of contrast in different spatial patterns.

6. Experiment 4

Our first control experiment was designed to investigate whether the difference between the textures used for the target and the matching pattern played any role in the transmittance matches. In Experiments 1–3, the textures of the target and the matching displays had been deliberately chosen to be different to prevent observers from making adjustments based solely on local image similarities. However, this raises the possibility that the contrast of different texture patterns were perceived differently. Indeed, contrast sensitivity for the target and the match may differ significantly because of their spatial frequency content (Campbell & Robson, 1968; Carlson, Cohen, & Gorog, 1977; Georgeson, 1991; Robson, 1966).

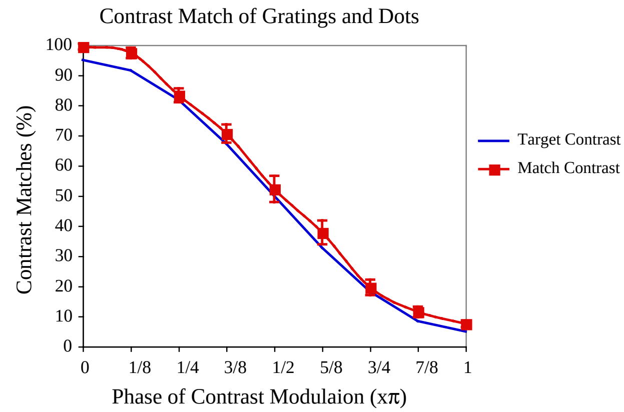

Fig. 7. Results of the control experiment (Experiment 4) to determine whether the difference in the spatial frequency content of the target and matching pattern was responsible for the apparent overestimation of transmittance values (relative to the predictions of the ratio of Michelson contrast model) of the target textures at values other than the anchor point. Observers simply matched the contrast of the target to the matching pattern. No reliable differences in contrast matches were observed in the regions where transmittance overestimation was observed in Experiment 1.

To test this possibility, two naïve observers were recruited to perform a simple contrast matching experiment. The target consisted of a random-dot texture pattern with a single contrast randomly chosen from the values measured in Experiment 1. The match was a single circular patch of a sine-wave grating with the same spatial frequency and size used in the experiments. The grating was also randomly assigned a contrast value between 0% and 95%. Both patterns were viewed against a uniform gray background with a mean luminance of 50.6 cd/m^2. The observers were instructed to adjust the disc's contrast to match that of the target. Fig. 7 shows the result of the contrast match between the two textures tested. It can be seen that there is no systematic bias in the perception of contrast across the two different texture patterns used; the curves are essentially identical. We therefore conclude that the difference in texture is not playing a significant role in the pattern of transmittance matches we observed.

7. Experiment 5

Our second control experiment involved testing the influence of transmittance anchoring on the matching display used in Experiments 1 and 2. The transmittance anchoring principle applies to both the target and match displays used in our experiments. In the low-amplitude contrast modulation (Fig. 2B), the highest contrast in the target was 45%, but the highest contrast in the match pattern was 99% (which was the same in all experiments). Our previous work (Singh & Anderson, 2002a) revealed that the visual system uses Michelson contrast to scale perceived transmittance, but in those experiments, the surround contrast was the same in the target and match patterns. Does the high-contrast surround in the *match* pattern affect the perceived transmittance anchor in a lower-contrast target?

To address this question, the low-amplitude contrast-modulation condition in Experiment 1 was repeated by two of the original three observers (R.F. and J.M.). The contrast of the outer disc in the match was set to 45% (Fig. 8A) instead of the value of 99% used in Experiments 1 and 2. Thus, the observers had to set the contrast of the inner small disc lower than 45% to perceive the central region to be transparent. The same 16 locations were measured along the contrast modulation. The results of the transmittance matches at these locations are plotted in Fig. 8B with 95% confidence interval across the two subjects. The bottom solid line represents the profile of the contrast modulation of the target. The upper dashed line marks the transmittance matches recorded in Experiment 1, where the outer disc of the match pattern was set at 99% contrast. Note that the raw data from these experiments are markedly different. This is to be expected, as transparency will only be perceived in the central target of the current experiment for contrast values below 45%, whereas transparency will be perceived for any contrast values below 99% in the first experiment. However, if the contrast matches of the first experiment are rescaled (by multiplying by ~0.45 (or more precisely, (45/99) = .455)), it can be seen that the data from these two experiments are indistinguishable. The dashed line with asterisk symbols depicts the result of this normalization. The identity of the normalized data from these two experiments indicates that the high-contrast surround in the match pattern does not affect the perceived-transmittance matches in the target pattern.

8. Experiment 6

Our last experiment was designed to assess the possibility that the contrast in the matching pattern was misperceived. One reason to be concerned about the perception of contrast in the matching pattern is that such patterns have been shown to give rise to transformations in perceived contrast. Indeed, our matching pattern was one of the original stimuli used to study the *contrast–contrast* effect (wherein a medium-contrast texture appears to be lower in contrast when surrounded by a higher-contrast texture; see Ejima & Takahashi, 1985; Chubb, Sperling, & Solomon, 1989). Yu, Klein, and Levi (2001) have recently reported a series of experiments using stimuli that were identical to those used in our matches to study this contrast effect. They reported that the perceived contrast of the center was lowered when the higher-contrast sine-wave grating from the surround shared the orientation with the center (Fig. 9B). This kind of misperception could cause the inner disc of the match pattern to be set to a higher Michelson contrast, which would be tantamount to an overestimation in transmittance in our studies.

To assess this possibility, the low-amplitude contrast-modulation condition was repeated (Fig. 9). The equipment used and the experimental conditions were identical to Experiment 1. In this experiment, observers were required to simply adjust the contrast of the inner disc of the matching display to match the contrast of a sinusoidally contrast-modulated random-dot pattern at a location indicated by markers above and below a region of the figure (Fig. 9A). In this context, we wanted to minimize the percept of transparency and any effects such percepts could have on observer's adjustments, so these images were

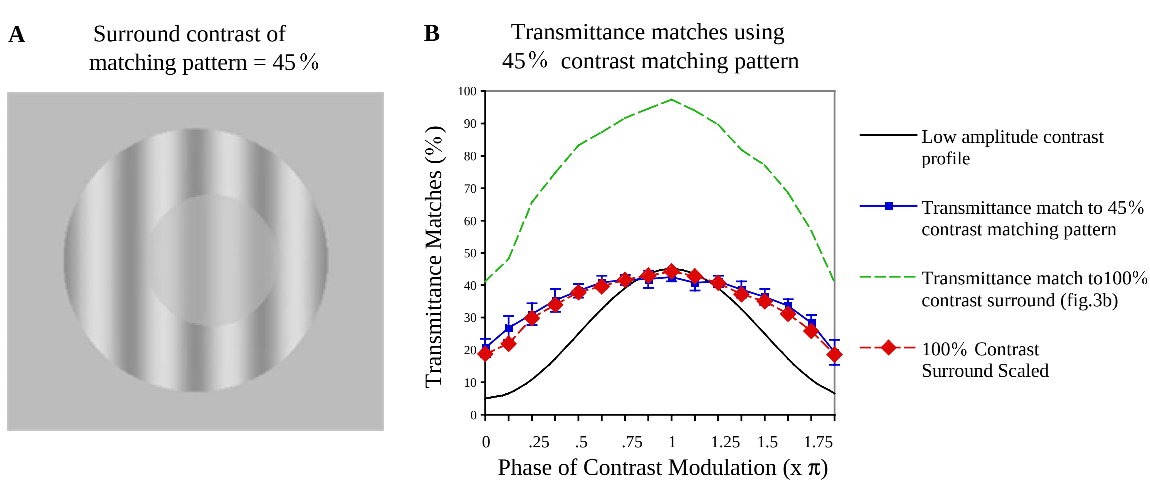

Fig. 8. Results from a control experiment (Experiment 5) to determine whether the high contrast surround used in the matching pattern influenced the judgments of transmittance in the target. The low contrast matching condition of Experiment 1 was performed using a matching pattern (A) with a surround contrast of 45%. (B) The results of Experiment 5. The solid blue lines indicate the contrast values of the target texture, and the red curve represents the data from Experiment 5. The dashed green curve depicts the data obtained when using the matching pattern with the high contrast surround (Experiment 1). Note, however, that the matching pattern in the present display will only appear transparent for contrast values lower than the surround (i.e., lower than 45%). When the data from the first Experiment are multiplied by this scale factor, it can be seen that the data are indistinguishable, demonstrating that the high contrast in the surround used in Experiment 1 did not influence transmittance judgments. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

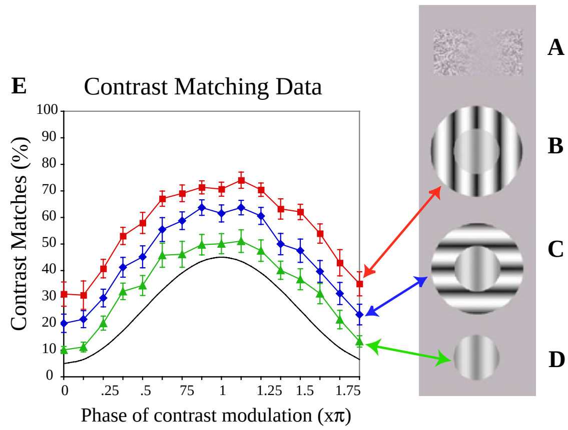

Fig. 9. The results from Experiment 6. Observers matched the perceived contrast of a nonstereoscopic variant of the target (A) using one of three matching patterns (B–D). In (B), the central target patch was of the same orientation, frequency and phase as the surround; in (C), the orientation of the surround was rotated 90°; and in (D) the surround grating was removed. Observers had to set the matching pattern to a much higher-contrast for the types of matching patterns used in our experiments (B), indicating that the contrast of the matching pattern used in our experiments was systematically underestimated.

viewed without any binocular disparity. Three kinds of match patterns were used (Figs. 9B–D). The first match pattern (panel B) was the same as in the previous experiments, where the orientation of the match's outer disc was the same as the inner one. In the second match pattern (panel C), the orientations of the outer and inner discs were orthogonal. The third match (panel D) was simply the inner disc alone. The same three observers used in Experiments 1–3 participated in this experiment.

Fig. 9E shows the data from this experiment. The bottom solid curve represents the (Michelson) contrast of the low-amplitude sinusoidal modulation at 16 different locations along the horizontal extent of the target. The three lines, with 95% confidence intervals, show the results of the contrast matching using the three different match displays. It can be seen from Fig. 9E that the presence of the outer disc elevates the contrast matches of the center. This effect is strongest when the outer and inner discs share the same orientation (panel B), and less so when the orientations of the two discs are orthogonal (panel C). When the inner disc was used alone (panel D), the matched contrast was closest to the actual contrast of the target. Thus, the contrast of the matching pattern used in Experiments 1 and 2 is systematically underestimated. On the face of it, the pattern of overestimation observed in the contrast matches in the present experiment is quite different from that of in the transmittance matches in Experiments 1 and 2, in that the overestimation of contrast is the same at all locations. This also demonstrates (again) that observers are not simply matching contrast per se when they are making transmittance matches, as this would result in the pattern of data observed in Fig. 9E. The question remains, however, whether this pattern of contrast matches accounts for the transmittance matches in Experiments 1 and 2.

Our model provides an explicit means to transform the data in this experiment to account for the transmittance matches in our previous experiments. Specifically, the TAP and our contrast-ratio model predict that the perceived transmittance of a transparent layer is determined by the ratio of contrasts between the region of transparency and the region seen in plain view. In our previous work (Singh & Anderson, 2000, 2002a), our test and match displays contained the same geometric properties and the same background, so any misperception of contrast in these displays would be the same for the test and matching displays. In the ratio-of-contrasts model, this implies that any misperception of the central-patch contrast would be represented by a common scale factor that would simply cancel out when computing the ratio of these values. However, in the present experiments, we cannot assume such a simple relationship in the perception of contrast in our test and matching patterns. The question, then, is whether the ratio of *perceived* contrasts in the region of transparency to the region seen in plain view accounts for the pattern of transmittance matches observed in Experiments 1–3 (cf. Robillotto et al., 2002; Robillotto & Zaidi, 2004). To answer this question, we need to transform the data of contrast matches observed in Fig. 9E by scaling these matches by the highest perceived contrast in these images, i.e., by dividing each data point in Fig. 9E by the peak contrast setting of this curve. This is tantamount to expressing the ratio-of-contrasts model in terms of the ratio of perceived contrasts.

To test this hypothesis, we normalized the contrast matches from Experiment 6 (condition b; see Fig. 9E) by the peak contrast of these contrast matches (i.e., we divided each data point in Fig. 9E by the peak contrast of this curve). The peak was estimated based on the best-fitting sinusoidal curve to the contrast data (i.e., of the form $y = a + b\sin(x - \pi/2)$)[3]—yielding a value of 0.76. When the contrast data were normalized by this value, the resulting curve (solid blue curve in Fig. 10) was found to be essentially indistinguishable from the best-fitting sinusoidal curve to the transmittance data from Experiment 1 (dashed red curve in Fig. 10). Thus, a model based on the ratio of perceived contrasts provides an excellent account of the perceived transmittance of inhomogeneous transparent surfaces. The visual system both anchors the spatially maximal contrast in an image to full transmittance, and it scales perceived transmittance in other image regions by normalizing perceived contrast by this spatial anchor.

8.1. Spatiotemporal anchoring

The analysis in the preceding section revealed that a ratio of perceived contrast model, together with the transmittance anchoring principle, provides an excellent account

[3] The best-fitting curve was determined by linear regression of the contrast/transmittance matches y on the transformed variable $s = \sin(x - \pi/2)$.

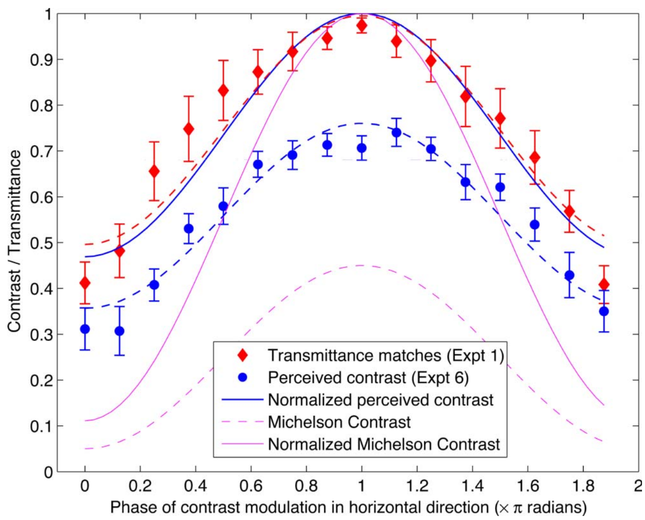

Fig. 10. Analysis revealing that a ratio of perceived contrasts model fully explains the transmittance matches from Experiment 1. Best fitting sinusoids were fit to the data from Experiment 1 (dashed red curve) and from Experiment 6 (dashed blue curve). When the perceived contrast of the matching pattern is normalized by its highest contrast value (solid blue curve), these data are indistinguishable from the transmittance matches from Experiment 1 (i.e., the dashed red curve is essentially identical to the solid blue curve). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

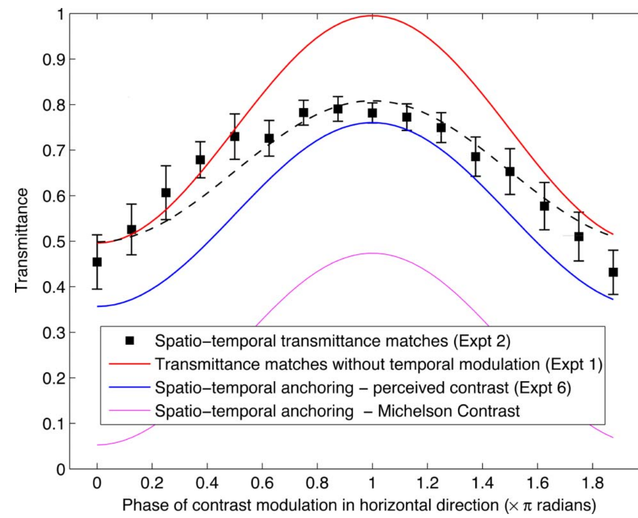

Fig. 11. A curve estimating the magnitude of the temporal anchoring effect. See text for details.

of the perceived transmittance of spatially inhomogeneous transparent media. What can be said about the temporal anchoring observed in Experiments 2 and 3? According to the spatiotemporal version of the anchoring principle, the highest contrast in a spatiotemporal sequence is anchored to full transmittance. The scaling of perceived transmittance is then computed by normalizing image contrast with respect to this spatiotemporal anchor.

To quantify the extent of temporal anchoring in Experiments 2 and 3, we express the observed temporal anchoring effect as a percentage of the maximal anchoring possible—i.e., based on normalizing perceived contrast with respect to the highest contrast in the entire temporal sequence. To do this, we first specify the two ends of the scale. If there is *no* temporal anchoring (i.e., the anchor is computed de novo on each frame), the perceived transmittance would be scaled simply by the spatially highest contrast within the current frame. This means that, in the lowest-modulation frame in which measurements were obtained, the perceived transmittance at its contrast peak would be 100%—since this is the spatial anchor for this frame. (The results of Experiment 1 showed that this is indeed what happens when this frame is presented in isolation; see the red curve in Fig. 11.) At the other end of the spectrum, if transmittance is scaled by normalizing perceived contrast by the spatiotemporally highest contrast in the entire image sequence, the predicted transmittance would be given by the perceived contrast at the lowest-modulation peak normalized by the perceived contrast at the highest-modulation peak. Given the substantial overestimation of apparent contrast observed in Experiment 6, one may reasonably assume that the perceived contrast at the highest-modulation peak is essentially 100%.[4] As a result, the perceived transmittance at the lowest-modulation peak is predicted simply by its perceived contrast-shown in the solid blue curve in Fig. 11 (fitted to the contrast-matching data from Experiment 6, condition b). Its peak, as we noted above, is 0.76. Thus, the maximal temporal effect possible in the current context is the difference between the two ends of the scale, namely, 0.24.

To measure the extent of temporal anchoring in Experiment 2, we estimate the peak of transmittance-matching data by fitting a sinusoidal curve to it (dashed black line in Fig. 11). This peak is found to be 0.8082, which implies that the obtained magnitude of temporal anchoring in Experiment 2 is 0.1918. This is 79.92% of the maximal temporal anchoring possible (0.24). Thus, normalization with respect to the spatiotemporally highest contrast in the image sequence explains a large proportion of the temporal anchoring observed. As the results of Experiment 3 revealed, the magnitude of temporal anchoring varies systematically with the timing parameters of the image sequence. Applying the above analysis to the data with different cycle durations, the extent of temporal anchoring was found to increase to 95.24% of maximal possible for the higher-speed condition (cycle time = 0.94 s), and decrease to 34.73% of maximal possible for the lower-speed condition (cycle time = 22.56 s).

[4] We are in fact being conservative in our estimate of temporal anchoring by assuming that the perceived contrast at the highest-modulation peak is 100%. Under the assumption of no overestimation in this case (i.e., perceived contrast at the highest-modulation peak = 95%), the estimated extent of temporal anchoring in Experiment 2 would be 95.9% rather than the 79.92% derived above. (The maximal temporal effect would be 1–0.76/0.95, i.e., 0.2, and the estimated extent of temporal anchoring obtained would therefore be 0.1918/0.20, i.e., 95.9% of maximal effect.)

9. General discussion

The results described herein reveal new insights into the computations underlying the perception of transparency. One of the primary goals of the reported experiments was to test the TAP's ability to predict when image regions will be perceived in plain view. In Experiment 1, we found that the TAP correctly predicted that the highest-contrast region of a texture appears in plain view, independently of its absolute contrast. In this experiment, observers provided the same pattern of transmittance matches for test images that varied in contrast by more than a factor of 2, which clearly would not have occurred if observers were simply matching contrasts of the two regions. Experiment 2 revealed that transmittance anchoring also has a temporal component: the highest-contrast region of the low-contrast pattern used in Experiment 1 no longer appeared in plain view when it was embedded within a temporal sequence that contained a series of higher-contrast regions. We found that the temporal anchoring effect varied as a function of the time elapsed between the highest-contrast frame and the measurement frame (or equivalently, by the speed of the contrast modulation). Although more research is required to determine the precise cause and characterization of these temporal effects, the main point that can be concluded here is that transmittance anchoring also has a temporal component. Finally, we found that a model based on the ratio of perceived contrasts provides an excellent quantitative fit to observer's transmittance matches in displays generating percepts of spatially inhomogeneous patterns.

The concept of transmittance anchoring was shaped and partially inspired by the concept of lightness anchoring discussed by Land and Mcann (1971) and developed by Gilchrist and colleagues (see, e.g., Gilchrist et al., 1999). In the lightness domain, this idea refers to the mapping from image luminance onto perceived surface reflectance. This problem is difficult because image luminance is a product of illumination and reflectance; some computational strategy is required for the visual system to determine how to parse image luminance into these two distinct sources. Gilchrist and colleagues have argued that the visual system accomplishes this goal by applying grouping principles to decompose images into regions, and then applying anchoring principles to these grouped regions ("frameworks") to infer surface reflectance. There is now a substantial body of data that show that the visual system appears to have a bias to interpret the most reflective surface in a scene as white, even when it is not (see Gilchrist et al., 1999 for a review). This "highest luminance is white" rule is conceptually related to the "highest contrast region of the scene is a surface in plain view" expressed by the TAP. There is, however, a principled difference between the two types of anchoring principles. The anchoring of the highest luminance to white in perceived reflectance is simply an empirically derived principle; no justification has yet been given for why this particular form of anchoring is used by the visual system over any other form of anchoring (such as interpreting the lowest luminance to black, or the average luminance as grey, or any other imaginable rule). In contradistinction, the TAP can be viewed as a form of "Occam's razor" in transparency perception (Anderson, 2003). It asserts that the visual system assumes the minimal number of surfaces needed to explain the image data.

The core intuition shaping the TAP is the idea that the visual system will only infer the presence of a transparent surface or medium if there is some *perturbation* in the contrast of surfaces or contours that provides evidence for its presence. So expressed, it is not simply a photometric constraint on image interpretation (i.e., one that relies solely on relative image intensities), but rather, is *a.photo-geometric* constraint (i.e., one that relies on a combination of geometric and photometric properties). In some displays, this constraint can be well captured by considering the contrast relationships that occur along contours (Anderson, 1999, 2003) or along both contours and textures (Singh & Anderson, 2002a, 2002b). Crucially, however, it is not a principle that is applied over the entire image. This can be readily appreciated by considering a scene containing a variety of patterned objects that generate different image contrasts. In such contexts, lower contrasts are not (necessarily) perceived as veiled, but rather, as possessing lower lightness contrasts. Thus, the TAP must act relatively locally in distinguishing portions of a scene that are in plain view, from those that are veiled by transparent surfaces or media.

One of the main predictions of the TAP is that the visual system uses relative contrast in space and time to determine whether a transparent layer is present or not. An intriguing clinical example that reveals the consequences of the TAP occurs in the formation of cataracts. The clouding of the lens during the formation of cataracts takes years to develop, leading to a very slow degradation of the contrast of the image over a span of roughly 20 years. Patients afflicted with cataracts do not usually report perceiving cloudy media in their visual world unless the cataract covers only a restricted portion of their visual field. Immediately following surgery, however, patients typically report a dramatic increase in their perceived contrast of the world. The relative homogeneity of the cataract, combined with the temporal sluggishness of its onset, apparently prevents patients from being aware of this slowly developing disability.

In sum, the TAP and the ratio of perceived contrast models provide a coherent quantitative account of the perception of surface transmittance in partially transmissive media and surfaces. Further research is needed to determine how other properties of transparent surfaces are computed (including their lightness), as well as determining the critical variables responsible for the temporal effects observed in the experiments reported herein.

References

Adelson, E. H., & Anandan, P. (1990) Ordinal characteristics of transparency. *AAAI-90 Workshop on Qualitative Vision,* July 29, 1990, Boston, MA.

Anderson, B. L. (1997). A theory of illusory lightness and transparency in monocular and binocular images. *Perception, 26*, 419–453.

Anderson, B. L. (1999). Stereoscopic surface perception. *Neuron, 26*, 919–928.

Anderson, B. L. (2003). The role of occlusion in the perception of depth, lightness, and opacity. *Psychological Review, 110*, 762–784.

Beck, J., & Ivry, R. (1988). On the role of figural organization in perceptual transparency. *Perception and Psychophysics, 44*, 585–594.

Beck, J., Prazdny, K., & Ivry, R. (1984). The perception of transparency with achromatic colors. *Perception and Psychophysics, 35*, 407–422.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology (London), 197*, 551–566.

Carlson, C. R., Cohen, R. W., & Gorog, I. (1977). Visual processing of simple two-dimensional sine-wave luminance gratings. *Vision Research, 17*, 351–358.

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A, 5*, 1986–2007.

Chubb, C., Sperling, G., & Solomon, J. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Sciences of the USA, 86*, 9631–9635.

D'Zmura, M., Colantoni, P., Knoblauch, K., & Laget, B. (1997). Color transparency. *Perception, 26*, 471–492.

Ejima, Y., & Takahashi, S. (1985). Apparent contrast of a sinusoidal grating in the simultaneous presence of peripheral gratings. *Vision Research, 25*, 1223–1232.

Faul, F., & Ekroll, V. (2002). Psychophysical model of chromatic perceptual transparency based on substractive color mixture. *Journal of the Optical Society of America A, 19*, 1084–1095.

Gerbino, W., Stultiens, C. I., Troost, J. M., & de Weert, C. M. (1990). Transparent layer constancy. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 3–20.

Georgeson M. A. (1991). Over the limit: Encoding contrast above threshold in human vision. In J. Kulikowski, V. Walsh, I. J. Murray (Eds.), *Limits of Vision*. Vol. 5, J. Cronly-Dillon (Ed.), *Vision and Visual Dysfunction*, London: Macmillan, pp. 106–119.

Gilchrist, A. L., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, R., Li, X., et al. (1999). An anchoring theory of lightness perception. *Psychological Review, 106*, 795–834.

Kasrai, R., & Kingdom, F. (2001). Precision, accuracy, and range of perceived achromatic transparency. *Journal of the Optical Society of America A, 18*, 1–11.

Land, E. H., & Mcann, J. J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America A, 61*, 1–11.

Metelli, F. (1970). An algebraic development of the theory of perceptual transparency. *Ergonomic, 13*, 59–66.

Metelli, F. (1974a). Achromatic color conditions in the perception of transparency. In R. B. MacLeod & H. L. Pick (Eds.), *Perception: Essays in Honor of J. J. Gibson*. Ithaca, NY: Cornell University Press.

Metelli, F. (1974b). The perception of transparency. *Scientific American, 230*, 90–98.

Metelli, F., Da Pos, O., & Cavedon, A. (1985). Balanced and unbalanced, complete and partial transparency. *Perception and Psychophysics, 38*, 354–366.

Robillotto, R., Khang, B., & Zaidi, Q. (2002). Sensory and physical determinants of perceived achromatic transparency. *Journal of Vision, 2*, 388–403.

Robillotto, R., & Zaidi, Q. (2004). Perceived transparency of neutral density filters across dissimilar backgrounds. *Journal of Vision, 4*, 183–195.

Robson, J. G. (1966). Spatial and temporal contrast-sensitivity functions of the visual system. *Journal of the Optical Society of America, 56*, 1141–1142.

Singh, M. (2004). Lightness constancy through transparency: Internal consistency in layered surface representations. *Vision Research, 44*, 1827–1842.

Singh, M., & Anderson, B. L. (2000). Toward a perceptual theory of transparency. *Investigative Ophthalmology and Visual Science Supplement, 41*, 1141.

Singh, M., & Anderson, B. L. (2002a). Toward a perceptual theory of transparency. *Psychological Review, 109*, 492–519.

Singh, M., & Anderson, B. L. (2002b). Perceptual assignment of opacity to translucent surfaces: The role of image blur. *Perception, 31*, 531–552.

Singh, M., & Anderson, B. L. (2006). Photometric determinants of perceived transparency. *Vision Research*, in press.

Yu, C., Klein, S. A., & Levi, D. M. (2001). Surround modulation of perceived contrast and the role of brightness induction. *Journal of Vision, 1*(1), 18–31. doi:10.1167/l.l.3, http://journalofvision.Org/l/l/3/.