Color Perception in Natural Images

Christopher W Tyler and Joshua A Solomon

City, University of London, Islington, London EC1, UK

christopher.tyler.1@city.ac.uk

Abstract

This review of color processing in natural image viewing – rather than artificial laboratory images – addresses the role of color edges. Much of the color variation in nature is a result of evolutionary processes in complex organisms that have developed eye-brain systems that use color signals for a variety of biological functions. One aspect of human color processing is the tendency to attribute the appearance of extended color fields to a process of filling-in from the differential color signals at color edges, where one color transitions to another. Does such a process account for the appearance of extended color fields in natural images? Some form of color filling-in must underlie the color filling-in percept known as the Watercolor Effect, but this effect is too weak to account for the appearance of extended color fields in natural images. Moreover, natural images do not look very colorful when their color is restricted to edge transitions. Conversely, purely chromatic images with maximally graded (‘edgeless’) transitions look fully colorful, leading to the conclusion that color filling-in makes no more than a minor contribution to the appearance of extended color regions in natural images. Other effects, such as the selective enhancement of perceived image color by luminance contours coordinated with the color contours and color image structure, also play a role the color perception of natural images.

Word count: 2055
The spectral variations that give rise to color are a ubiquitous feature of nature and natural images. Organisms, in particular, have evolved to use color as a signaling system in a wide variety of ways – species recognition among members of a species; species identification to other species; signaling of gender and of sexual receptiveness; camouflage against predators; signaling of readiness to as a food source in the case of fruit and flowers; etc. All these functions of color in the natural world are bound up with the visual perception of color by a neural system of some kind, implying that much of the color we find in nature is a result of evolutionary processes in complex organisms that have developed an eye-brain system that can appreciate the color signals.

These aspects of evolved color are distinct from the adventitious physical processes such as minerals, crystals and rainbows, which would be occurring in nature regardless of any organic evolution. These kinds of optically refrangent colors are also appreciated at least by human eyes, in the form of precious gemstones, and by certain bird species such as magpies and bower birds, but there seems to be no evolutionary aspect to the existence of such physical processes per se.

Color is, of course, well established as having three dimensions of variation to the typical human eye (though perhaps many more in species with many more receptor types – such as the mantis shrimp (12 receptor types, Bok, Roberts & Cronin, 2018) or the Japanese swallowtail butterfly (9 receptor types; Arikawa, 2018). The dimensionality of color representation may be sliced in many ways, but
one of the most revealing is to separate the color variations from the grayscale variations in an image (Fig. 1, see Appendix I).

Doing so makes it clear how much information about object structure and boundaries is conveyed by color in natural images – without knowing what a parrot is, it is challenging to complete the overall outlines of the two parrots from bottom-up information in the grayscale image alone (or from the color image alone). Indeed, Hansen & Gegenfurtner (2008) show that pure color borders are as common in natural images as luminance borders, but that the two types carry statistically independent information. This illustrates the kinds of difficulties faced by computer science in real-world object recognition.

Fig. 2. Balancing the chromatic information for equal color contrast across the color spectrum in the macaw image. A: the image of Fig. 1A chromatically balanced. B: balanced chromatic component with the luminance contrast removed. C: grayscale image of the luminance component alone after chromatic balancing.

The color image also illustrates the second dimension of color, which is the variation in the intensity of the color (otherwise known as saturation or color contrast) as well as its hue. The reds and blues in Fig. 1B are very intense, while the greens are almost grayed out. These variations are adjusted in a second version of the image (Fig. 2), which was designed to balance the saturation levels across the color spectrum (see Appendix I). When the luminance contrast is removed (Fig. 2B), the remaining hue boundaries now appear much less salient, making it clear that color contrast plays as great a role in making color regions stand out from their backgrounds as their hue. The contrast balanced colors,
though still perceptually evident, do not stand out as expressing discrete shapes as strongly as those in the original color channel including color contrast variations (Fig. 1B).

In a retina containing about 5 million cones of three varieties (selective for the red-, green- and blue-regions of the color spectrum), it may seem self-evident that the perception of color is mediated by the outputs, or combinations of outputs, of those cone pathways at a local scale comparable to that of the receptors. However, color processing continues in the visual cortex, where a subpopulation of the neurons combines signals from the individual cones into color-edge detectors (Livingstone & Hubel 1984; Hubel & Livingstone, 1987; Shapley & Hawken, 2011), which inherently emphasize the local changes of color at color edges, at the expense of signals in regions of uniform color fields. In fact, such results are the basis for the widespread idea that much of color processing operates by identifying the colors at color edges (Larimer & Piantanida, 1988; Pinna, Brelstaff, & Spillmann, 2001; Von der Heydt et al., 2003; Shapley & Hawken, 2011; Paramei & van Leeuwen, 2016; Gerardin et al., 2018). As Shapley & Hawken (2011) write, “many ... perceptual results indicate that the color appearance of a region sometimes is even more dependent on color contrast at the boundary of the region than it is on the spectral reflectance of the region.” Perhaps the most famous perceptual result attributed to color edges is the strong color-orientation specificity in studies of joint color/orientation adaptation known as the McCollough Effect (McCollough, 1965). The argument proceeds from the proposal by Craik (1958), O’Brien (1966), and Cornsweet (1970) that lateral inhibition reduces or eliminates the neural response to uniform fields, leaving as the most salient features the edges, or regions of rapid luminance change in the image (Nieves, Nascimento & Romero, 2012). Edge information is then propagated into uniform regions to produce the full color impression (e.g., Gerrits & Vendrik, 1970; Grossberg & Mingolla, 1985; Ma & Manjunath, 1997; Iakovidou et al., 2019). This somewhat mysterious process, attributed to higher cortical levels, is known as “filling-in.”

The filling-in concept is then generalized from luminance to color fields, forming is the conceptual basis of the 'Watercolor Effect" of Pinna (1987), in which faint color is indeed seen to propagate across the white spaces between the double lines of two contrasting colors (Fig. 3). However, although there is demonstrable color filling-in, it is markedly unsaturated, as quantified by Wachtler & Wehrhahn (1997) and Gerardin et al. (2018), who show that it averaged about 10% of the strength of the luminance filling-in (or about a 0.5% effect in terms of cone contrast). This result raises the
question whether this weak form of color filling-in can account for the normal range of rich color perception of extended color surfaces in natural images, such as that of Fig. 1A.

Fig. 3. Left: An example of the Watercolor Effect of Pinna (1987), in which color is induced in the white regions between the double-color bounding lines. (From Pinna, Werner & Spillmann, 2003)

An interesting property of color filling-in is that it is weak or non-existent when the two colors of the border are equiluminant with each other (Pinna, Brelstaff & Werner, 2001; Devinck et al., 2005). Thus, it is not a pure color effect analogous to the filling-in of the luminance domain, which is a pure luminance effect. Instead, the color filling-in must be parasitic on the luminance filling-in, as though the activation of luminance propagation pathway (whatever that is) to carry the color signals along with it.

A primary argument against the logic of color filling-in as a general phenomenon is that most color-selective receptive fields of cortical neurons do not exhibit surround inhibition (see Komatsu, 2006, for review). Moreover, there is little or no psychophysical evidence for lateral inhibition in the color domain. The type of low frequency fall-off that is taken as psychophysical evidence for lateral inhibition in the luminance domain (de Lange, 1953; Campbell & Robson, 1968) is not found for pure color modulation in either the spatial (van der Horst & Bouman, 1967; van der Horst, de Weert, & Bouman, 1967; Mullen, 1985) or temporal (van der Horst, 1969; Regan & Tyler, 1971) domains. These lines of evidence give every reason to expect only minimal filling-in process to operate for color
images. Even if color filling-in did operate under laboratory conditions, would it be an effective operating mode for the viewing of natural images? It seems most appropriate to answer this question under natural viewing conditions, in which the eye is free to move.

To evaluate the question of color filling-in, Tyler & Solomon (2018) used a color-border version of the equiluminant, balanced contrast color image of Fig. 2B, computed by filtering out its low spatial frequencies with an isotropic high-pass filter (Fig. 4; Tyler & Solomon, 2018). It can be seen that the image appears as narrow lines of high contrast color borders with predominantly uniform gray between them (although there may be a hint of perceived color as expected from the Watercolor Effect). It is evident, therefore, that under free-viewing conditions the color borders *per se* in this natural image are not sufficient to provide for the full-color impression of the original image.

![Fig. 4](image)

**A. Color border**  **B. With grayscale overlay**

Fig. 4. A: Color-border versions of the equiluminant, balanced contrast color image of Fig. 2, showing some degree of Watercolor Effect filling-in. B: the balanced grayscale overlay of Fig. 2 added to the color-border image, which reduces both the border color and the color filling-in percepts.

Now, it could be argued that full-color equiluminant images are unnatural, even when derived from a natural image, because they lack the natural joint occurrence of luminance and color borders that seems to be critical for the watercolor effect (Pinna, Brelstaff & Spillmann, 2001). Fig. 4 (right panel)
therefore shows the same color border information added back to the grayscale image of Fig. 2 (right), which shows that, if anything, the luminance information tends to mask the color information rather than to induce enhanced filling-in, since the color now appears less vivid both at the borders and in the interior regions. This reduction in color filling-in is supported by recent demonstrations that luminance borders can selectively influence color perception and integration (van Lier, Vergeer, & Anstis, 2009; Anstis, Vergeer & van Lier, 2012a,b; Coia & Crognale 2018; Hong & Tong, 2017). An illustration of this effect for colored afterimages is found in Anstis, Vergeer & van Lier (2012b), based on the ‘natural images’ of paintings by Gainsborough (‘The Blue Boy’, ~1770) and Ingres (‘La Source’, 1856). They showed that adding the luminance component from each image to a 50:50 meld of the chromatic components from both images selectively suppressed the other image (Fig. 5). The meld is simply the sum of the balanced chromatic components (see Appendix I).

![Fig. 5. Demonstration of the selective effect of luminance structure on color perception (recast from Anstis, Vergeer & van Lier, 2012b), using the same logic as in Fig. 2. Upper right: two classic paintings. Upper left: their luminance components. Lower left: their balanced chromatic components. Lower right: the color meld of the two chromatic components alone, and with the addition of each of the two luminance components. Note that the addition of each luminance component selectively switches the same color pattern in the direction of the respective originals.](image)

One objection to this analysis might be that the equiluminant balancing procedure departs from the unbalanced nature of the color across natural images in general, and that color filling-in might be more effective for unbalanced color edges. To address this point, Tyler & Solomon (2018) included a further figure of the color edge manipulation for the unbalanced equiluminant case from Fig. 2 and
four other diverse natural images (Fig. 6). It is evident that only the smallest color patches at the scale of the edge-filtering show pronounced color percepts. For more extended regions, filling-in is essentially absent, despite the variety of natural border structure. Except for small features like the blue car, which is on the scale of the border-color information, the brightly colored uniform regions all appear gray between the borders, implying that color filling-in failed to operate under these conditions. These results provide further support for the contention that color-border filling-in is not the primary mechanism for our color perception in natural images.

Interestingly, Hamburger et al. (2006, 2012) report a remarkably strong variety of color propagation, under the decidedly unnatural viewing condition wherein images are stabilized on the retina. In this case, surface colors can propagate across strong color borders to invade interior regions of any other color. The resulting color is not a mixture but completely dominated by the stronger surface color. Thus, color propagation under stabilized retinal adaptation exhibits remarkable higher-order interactions that cannot be explained by the kind of basic border-to-border filling-in examined in this
review of color processing in natural image viewing. These effects may be related to the afterimage (and hence stabilized) color filling-in reported by Shimojo, Kamitani & Nishida (2001). Color is induced laterally in the Varin illusion (a color version of the Kanisza illusion) is replicated in the afterimage of the stimulus. The wrinkle in the afterimage version is that there is reciprocal fading of the inducing pattern and induced color field, implying that the filled-in afterimage color was an afterimage of the induced color filling-in (rather than a filling-in color from the afterimage of the illusion color).

**Conclusion**

There are four types of filling-in identified in this survey of color filling-in effects in natural images. The classic type of filling-in is Craik-O’Brien-Cornsweet form of region filling-in the luminance information from the edges, which only partially accounts for the perceived brightness of the region. The Watercolor Effect is the chromatic version of the same effect, but is approximately 10x weaker than its luminance counterpart. The third type is the Hamburger-type figural regional filling-in of color across stabilized-image color borders, which is based on the figural relations of an outer surrounding region filling-in over a different-colored inner region. And the final type is the Shimojo-Kamitani-Nishida selective enhancement of color filling-in (or filling-out) to luminance defined borders, with complementary suppression of color regions that are incompatible with the luminance border structure.

**Figure Credits**

Figs. 1,2,4: dima266f ; Fig. 6 l to r: dima266f, pixelio.de, Fedor Yakubovich, Tony Hisgett, Allan Warren; Fig. 7, Jim Kasson.

**References**


Anstis S, Vergeer M, Van Lier R (2012a). Luminance contours can gate afterimage colors and “real” colors. *J Vision*, 12(10):2, 1–13, www.journalofvision.org/content/12/10/2, doi:10.1167/12.10.2. Although termed “color filling-in”, this is the opposite to center filling-in from edge colors. Instead it is the blockage of field color perception by achromatic edges, allowing for selection of different color percepts from the field-colors in the same mixed-color images or afterimages.


Describe a contour adaptation paradigm that reduces the brightness contrast of a surface inside the area bounded by the contour, but has no effect on the color of the bounded area for either achromatic or chromatic contours. These results support the idea that a significant proportion of the perceived brightness of bounded regions but a negligible proportion of the perceived color is carried by propagation of the differential information from the edges to the bounded region. Indeed, achromatic contour adaptation even tended to enhance the color of colored regions.


The inner contour of the watercolor effect generated an assimilative filling-in for hue judgments but an inverse contrast effect for brightness judgments. Both effects increased with increasing luminance of either the enclosed or the surrounding region.


Activity patterns to selective contours that induced different color percepts from the same mixed-color afterimage (Anstis, Vergeer, & Van Lier, 2012a,b) allowed for reliable discrimination of the filled-in colors percepts in human cortical areas V3 and V4, but not in earlier visual areas.


Inclusion of color edge information in a computational image descriptor that calculates region derivatives at different scales and orientations enhanced the mean average precision of the descriptor relative to other descriptors based on luminance contours alone.


These authors show how color is perceived to spread from opposing-color borders into the empty white space between them, in a striking demonstration of long-range color filling-in effects. Curiously, these work best when the opposing-color border is wiggly rather than straight, suggesting that they are specific to the vagaries of natural images rather than artificial stimuli.


Appendix I: Equiluminance Transforms

All of the transformations were performed in accordance with international standards (see IEC, 1999). The chosen sRGB images were downloaded from the internet. Each sRGB image was represented by an $m \times n \times 3$ matrix $S$, where $m$ and $n$ are the number of pixels in each row and column, where each matrix entry is a real number $u \in [0,1]$. The original nonlinear gamma correction was linearized by replacing each value $u$ with $v$, where

$$v = \begin{cases} 
  u/12.92 & u < 0.04045 \\
  (u+0.055)^{2.4}/1.055 & \text{otherwise}
\end{cases}.$$

Each resulting "linear" pixel $V = \begin{bmatrix} r & g & b \end{bmatrix}$ was then converted to greyscale via matrix multiplication with the transpose of vector $L = \begin{bmatrix} 0.2126 & 0.7152 & 0.0722 \end{bmatrix}$. Each pixel $y$ in the greyscale matrix $Y$ (Fig. 1c) has been gamma corrected, i.e.

$$y = \begin{cases} 
  12.92vL' & vL' \leq 0.0031308 \\
  1.055(vL')^{1/2.4} - 0.055 & \text{otherwise}
\end{cases}.$$
Residual chromatic images were obtained from the difference between original and greyscale versions (both normalized to have zero mean). Specifically, if we let $\mathbf{X}$ and $m_X$ represent the scalar, arithmetic mean and maximum absolute value of every entry in an arbitrary matrix $X$, then the 'equiluminant' image in Fig. 1b is described by a matrix in which each element $c$ is scaled to lie within the interval $[0,1]$, i.e., $c = \left[ \left( \frac{d}{m_b} \right) + 1 \right] / 2$, where $d$ is each element in the normalized difference matrix $D$ defined by $D = S - S - Y - Y$. The meld of the two equiluminant images for Fig. 5 was defined by $D_M = (D_1 + D_2) / 2$, with the addition of the luminance component by $D_{MY} = (D_M + Y) / 2$.

The logic of the chromatic balancing was to shift the centroid of the equiluminant pixel gamut to the neutral (gray) point of the color space. In practice, this was implemented by adjusting the red-cyan balance of the equiluminant image to its perceptual balance point for the authors, as illustrated in Fig. 2B.
AUTHOR DECLARATION

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us.

We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property.

We understand that the Corresponding Author is the sole contact for the Editorial process (including Editorial Manager and direct communications with the office). He/she is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs. We confirm that we have provided a current, correct email address which is accessible by the Corresponding Author and which has been configured to accept email from cwtyler2020@gmail.com.

Signed by all authors as follows:

C. W. Tyler

J. A. Solomon

Signed by all authors as follows:

22/2/19

February 23, 2019