Title
The visibility of color breakup and a means to reduce it

Journal
Journal of Vision, 14(14)

ISSN
1534-7362

Authors
Johnson, Paul V
Kim, Joohwan
Banks, Martin S

Publication Date
2014-12-19

DOI
10.1167/14.14.10

Peer reviewed

The visibility of color breakup and a means to reduce it

Paul V. Johnson University of California, San Francisco, and University of California, Berkeley, Berkeley, CA, USA

Joohwan Kim Vision Science Program, University of California, Berkeley, Berkeley, CA, USA

Martin S. Banks Vision Science Program, School of Optometry, University of California, Berkeley, Berkeley, CA, USA

Color breakup is an artifact seen on displays that present colors sequentially. When the eye tracks a moving object on such a display, different colors land on different places on the retina, and this gives rise to visible color fringes at the object's leading and trailing edges. Interestingly, color breakup is also observed when the eye is stationary and an object moves by. Using a novel psychophysical procedure, we measured breakup both when viewers tracked and did not track a moving object. Breakup was somewhat more visible in the tracking than in the non-tracking condition. The video frames contained three subframes, one each for red, green, and blue. We spatially offset the green and blue stimuli in the second and third subframes, respectively, to find the values that minimized breakup. In the tracking and non-tracking conditions, spatial offsets of $\Delta x/3$ in the second subframe (where Δx is the displacement of the object in one frame) and $2\Delta x/3$ in the third eliminated breakup. Thus, this method offers a way to minimize or even eliminate breakup whether the viewer is tracking or not. We suggest ways to implement the method with real video content. We also developed a color-breakup model based on spatiotemporal filtering in color-opponent pathways in early vision. We found close agreement between the model's predictions and the experimental results. The model can be used to predict breakup for a wide variety of conditions.

Introduction

Many commercial projectors, such as the popular Digital Light Processing projector (DLP; Hornbeck, 1997), display colors sequentially. The most common implementations present red, green, and blue components sequentially in a given frame. In this case, the presentation rate (the rate at which image data is displayed) is three times the capture rate (the rate at which images are captured by the camera or produced in computer-generated imagery). When such projectors present a bright moving object on a dark background, colored fringes are often seen at the object's leading and trailing edges (Arend, Lubin, Gille, & Larimer, 1994; Cheng & Shieh, 2009; Post, Monnier, & Calhoun, 1997; Post, Nagy, Monnier, & Calhoun, 1998; Zhang & Farrell, 2003). These fringes are referred to as color breakup.

For a projector presenting red, green, and then blue (RGB), the appearance of color breakup depends on the object's motion and color, and the viewer's eye movements. When the viewer tracks a moving white object, the three colors land in different places on the retina. Figure 1A shows the colored fringes that might be seen with an RGB-sequential protocol when a white object moves rightward across a dark background and the viewer tracks it. Due to the eye movement, the images from later subframes are seen displaced leftward relative to those from earlier subframes. Thus blue and cyan (green + blue) fringes are seen near the left trailing edge, and red and yellow (red + green) fringes are seen near the right leading edge. Figure 1B is a video of an object moving right and left and illustrates the breakup phenomenon.

Figure 2 depicts the individual subframes presented to the eyes in the RGB protocol (R in the first subframe, G in the second, and B in the third). The stimulus is a four-pixel-wide white bar moving at constant speed. When the viewer fixates a stationary point (non-tracking), the images of the moving object from the three subframes land on the same position on the retina, but at different times. When the viewer tracks the object (tracking), the three subframes present images at the same location on the display, but at different locations on the retina.

Citation: Johnson, P. V., Kim, J., & Banks, M. S. (2014). The visibility of color breakup and a means to reduce it. *Journal of Vision*, *14*(14):10, 1–13, http://www.journalofvision.org/content/14/14/10, doi:10.1167/14.14.10.

doi: 10.1167/14.14.10 Received July 11, 2014; published December 19, 2014 ISSN 1534-7362

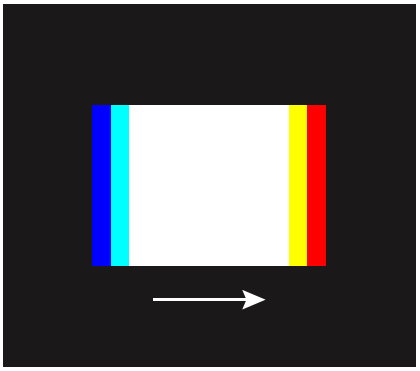

Figure 1A. Color breakup with an RGB-sequential display protocol when the viewer tracks a moving stimulus. Left: Depiction of a white object moving rightward across a dark background. Red and yellow fringes appear at the leading edge, and blue and cyan fringes at the trailing edge. The width of the fringes depends on the object's speed and display's presentation rate.

Figure 1B. Video illustrating color breakup. A bright object is presented with three subframes (red, green, and then blue presented at the same screen location) per frame. As the object moves rightward and you track it, red and yellow fringes appear at the right (leading edge) and blue and cyan fringes appear at the left (trailing edge). As it moves leftward, red and yellow appear at the left edge and blue and cyan at the right. If you fixate a stationary point, you still see color breakup but its visibility is reduced.

To understand the color-sequential stimulus, it is useful to consider the spatiotemporal differences between a bright moving stimulus with R, G, and B presented sequentially versus simultaneously: that is, when R, G, and B are presented together in each of three subframes. The sequential presentation requires stimulation in all three subframes, so we made the same assumption for simultaneous color presentation. Let the frame duration be Δt and the displacement of the stimulus on the screen Δx. Figure 3 shows the spatiotemporal modulation of red and green when the stimulus is composed of equiluminant R, G, and B. The left column plots the modulation when the viewer does not track the stimulus and the right column the modulation when the viewer does track it. The upper row shows the modulation in R as a function of space and time and the lower row the modulation in G. (The panel for B would be very similar, just shifted one subframe.) In the figure, gray represents no difference in luminance between the sequential and simultaneous presentations. White represents space–time positions in which the luminance of R (or G) is greater in the sequential than in the simultaneous stimulus. Black represents space–time positions in which R (or G) is lower. As you can see, the differences with tracking are identical to those without tracking except for a vertical shear parallel to the position axis; the shear, of course, is due to the eye movement. Thus, from the physical differences in luminance—that is, from the differences without consideration of spatiotemporal filtering by the visual system—one cannot predict whether tracking or non-tracking will produce more apparent color breakup.

With tracking, introducing a spatial offset of $\Delta x/3$ in the image in the second subframe relative to the first and an offset of $2\Delta x/3$ in the third subframe relative to the first will cause the subframes to be presented at the same location on the retina. Figure 4 shows differences in the modulation of red and green between sequential and simultaneous presentation for an offset of $\Delta x/3$ in the G (second) subframe. The modulation differences with and without tracking are identical except for a vertical shear parallel to the position axis. In fact, the differences are identical to those without the spatial offset except for shifts parallel to the position axis due to the offset. The modulation differences in space–time, therefore, do not reveal whether spatial offsetting should reduce or eliminate color breakup, nor whether offsetting should have a different effect with tracking than without tracking. This means that color breakup is not an attribute of the display or video content, but rather is created in visual processing. In particular, without low-pass filtering in early vision, color breakup would always occur when red, green, and blue are presented sequentially and breakup would be unaffected by spatial offsets of the sort illustrated in Figure 4.

To understand the causes of color breakup and particularly ways to minimize its visibility, we need to incorporate spatiotemporal filtering in color-opponent channels in early vision (Figure 5). We constructed a model based on measured spatial and temporal properties of color-opponent processes in the human vision. The input was a bright achromatic object moving across a dark background. Red, green, and blue subframes were presented sequentially. Details of the model are provided in Appendix A.

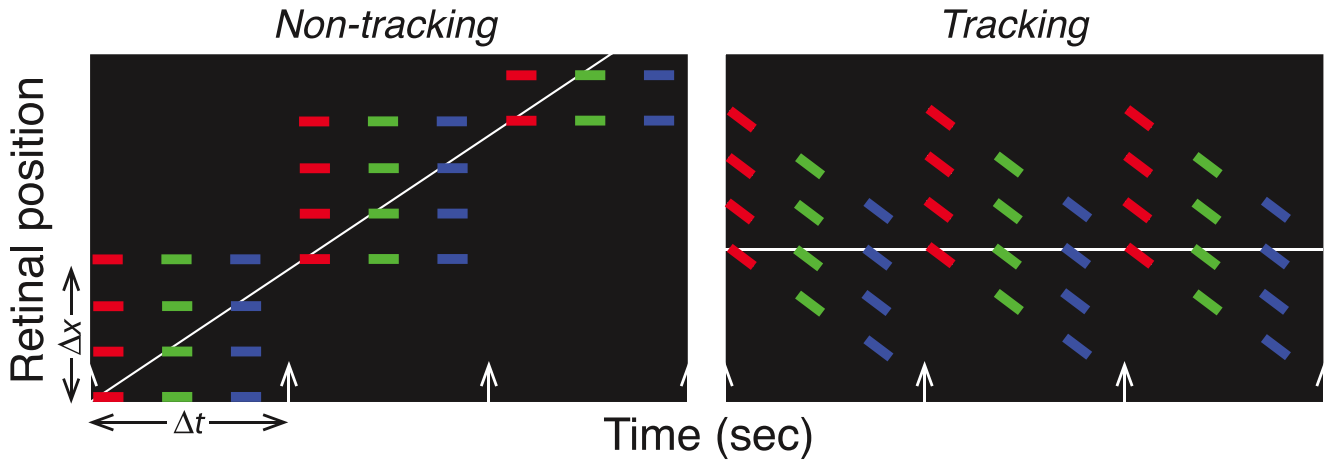

Figure 2. Space-time plot of the retinal images generated by a narrow white bar moving at constant speed. Position on the retina is plotted as a function of time. The object is four pixels wide and moving at three pixels per frame. The object's displacement on the screen is Δx for each frame of duration Δt. Left: The viewer fixates a stationary point while the object moves. The subframes within a frame fall on the same retinal position but at different times. Right: The viewer tracks the object. The red component spatially leads the green by $\Delta x/3$ due to the eye movement during the first subframe. The green component also leads the blue component by $\Delta x/3$.

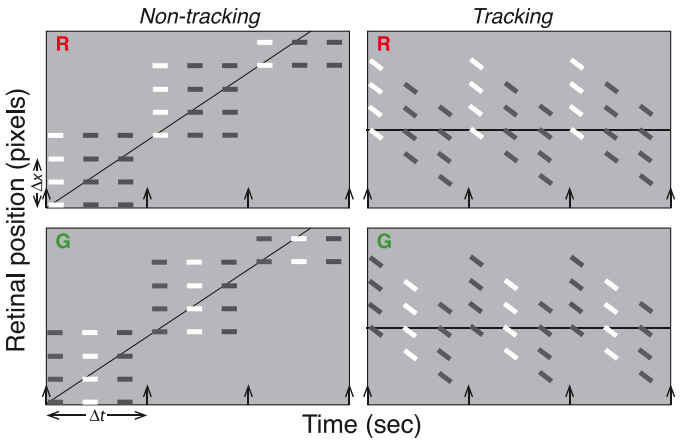

Figure 3. Space-time differences in color modulation between simultaneous and sequential color presentation. Each panel plots the variation in red or green stimulation as a function of time and retinal position when a white object moving at a speed of $\Delta x/\Delta t$ is presented (two pixels per frame). The object is four pixels wide. Vertical arrows indicate frames, each consisting of three subframes. In the sequential presentation, the subframes are R, then G, and then B. In the simultaneous presentation, R, G, and B are presented in all three subframes. The left column shows the differences when the viewer does not track the object and the right column the differences when the viewer does track. The upper and lower rows show the differences for R and G, respectively. Gray represents no difference. White represents cases in which R (or G) is more luminous in the sequential presentation. If the luminance of R (and G) is 1 integrated across a frame, then white represents +2/3. Black represents cases in which R (or G) is less luminous, and corresponds to −1/3.

Figure 6 shows some of the modeling results. Each panel plots predicted modulation among color-opponent channels as a function of retinal position and time; when the value is zero there is no modulation, so breakup should not be seen. The left half of the figure shows the responses when the viewer's eyes are

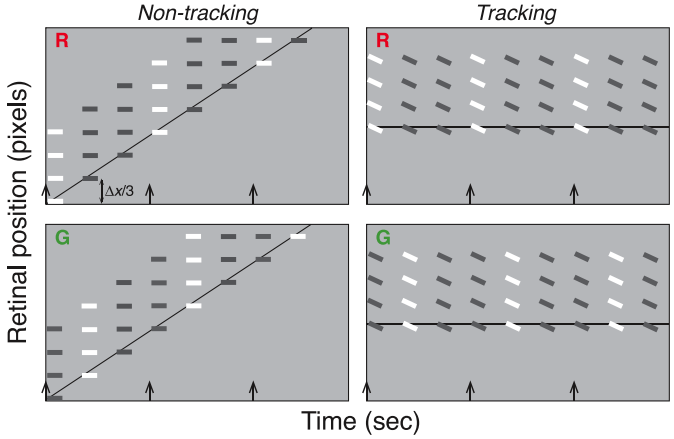

Figure 4. Space-time color modulation when a spatial offset of $\Delta x/3$ is applied to the G subframe. Each panel plots the variation in red or green as a function of time and retinal position when a moving white object is presented. In the sequential presentation, the subframes are R, then G, and then B. In the simultaneous presentation, R, G, and B are presented in all three subframes. The left and right columns show the variation without and with tracking, respectively. The upper and lower rows show the variation for the R and G channels, respectively. Gray represents no difference. White represents cases in which R (or G) is more luminous in the sequential presentation. Black represents cases in which R (or G) is less luminous.

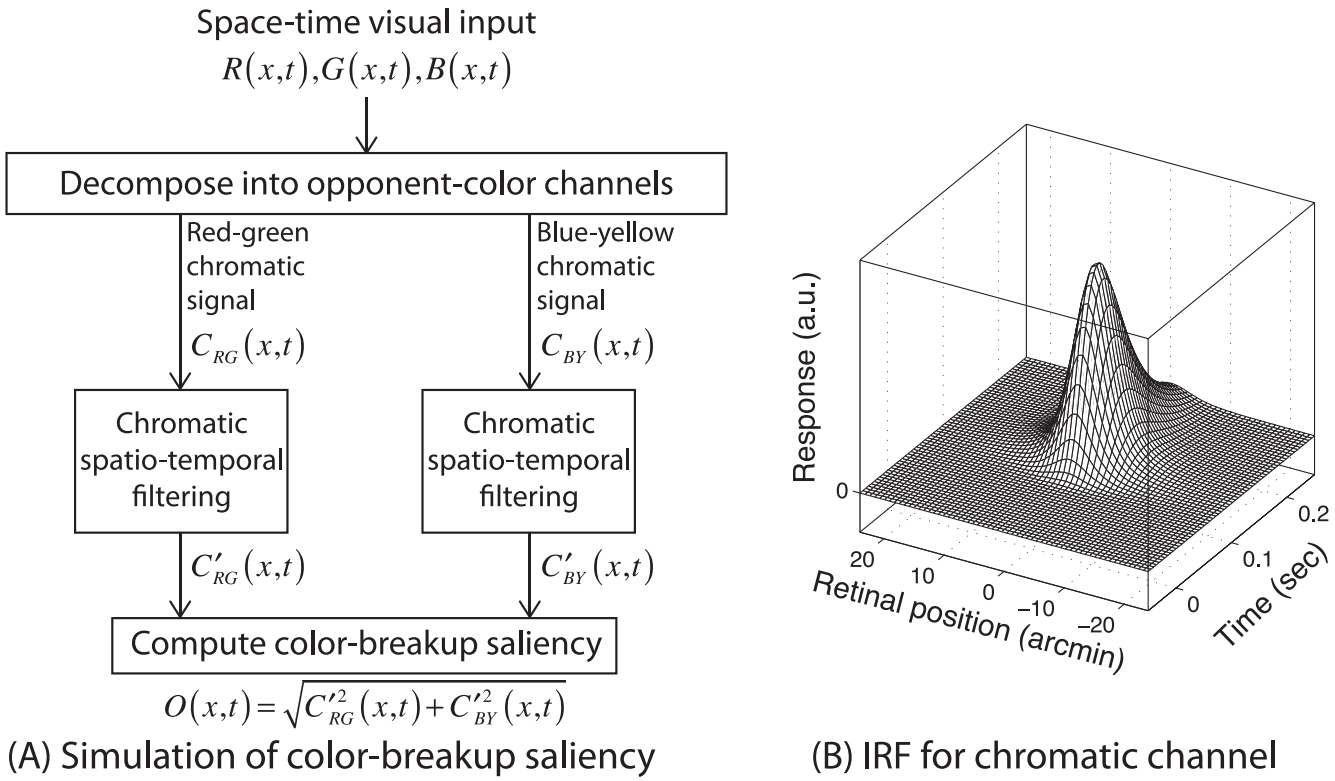

Figure 5. Model of color breakup. The model estimates the visibility of color breakup based on spatio-temporal filtering in color-opponent channels in early vision. The left side of the figure depicts the model's components. The right side shows the impulse-response function (IRF) for the color-opponent channels used in the model. For details, see Appendix A.

stationary, and the right half shows the responses when the viewer tracks the stimulus. The panels from top to bottom show the responses when the second (green) subframe is displaced relative to the first (red) subframe by $-\Delta x/3$, 0, $\Delta x/3$, $2\Delta x/3$, and Δx, respectively, and the third (blue) subframe is displaced by twice those values. As expected from consideration of low-pass temporal filtering, a spatial offset of $\Delta x/3$ eliminates color breakup in the model when the viewer tracks the object. Interestingly, the same offset minimizes predicted breakup when the eyes are stationary. The range of offsets for which breakup should not be visible is wider in the non-tracking case than in the tracking case.

As we said earlier, the cause of visible breakup during tracking when the offset is zero is obvious: Different subframes fall on different positions on the retina and therefore appear spatially displaced. The cause of breakup with zero offset when the eyes are stationary is less obvious. When a frame is presented, three subframes land on one position on the retina, but at different times. When the next frame appears, the visual impression from the first subframe of the last frame has faded more than the impression from the second and third subframes. Therefore, a preceding frame affects the appearance of the first subframe the most, and a succeeding frame affects the last subframe the most. At the trailing edge, a black frame follows a white frame leaving the visual impression from the last subframe, which is blue. At the leading edge, a black frame precedes a white frame, so the visual impression from the first subframe, which is red, is more salient than from the other subframes. As a result, one sees blue and red at the trailing and leading edges, respectively: the same ordering as observed when tracking the white object. A spatial offset of $\Delta x/3$ for the second subframe and $2\Delta x/3$ for the third causes the trailing and leading edges to oscillate mostly in hue. Sensitivity to rapid variations in hue is low, so the visibility of breakup is reduced or eliminated.

We predict that a spatial offset of $\Delta x/3$ for the second subframe and $2\Delta x/3$ for the third subframe can minimize color breakup regardless of whether the viewer is tracking a moving object or fixating a stationary point while an object moves past. We also predict that the range of offsets that eliminate or minimize color breakup will be wider when the eyes are stationary than when they are tracking. We conducted psychophysical experiments to evaluate these predictions.

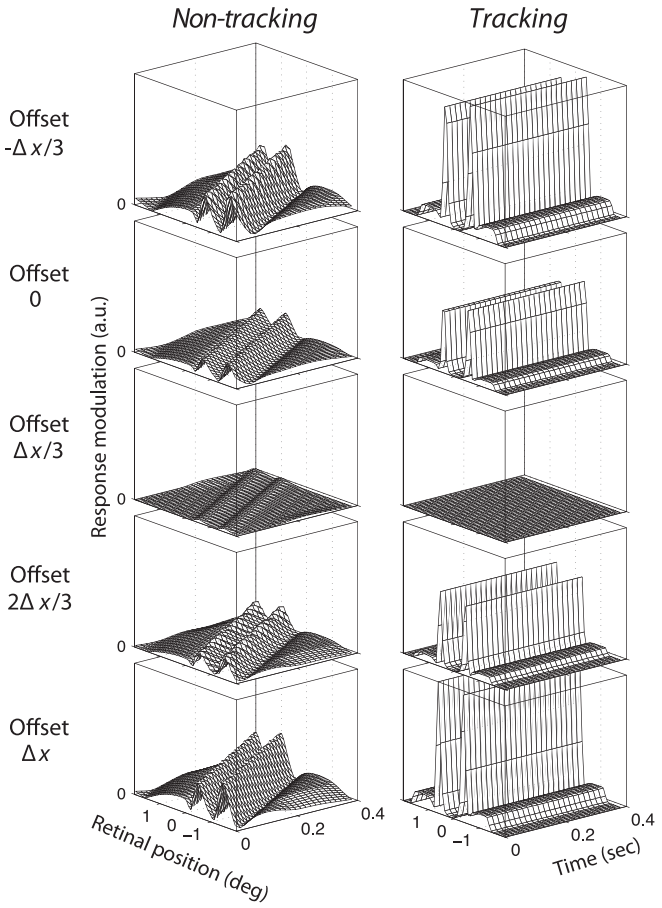

Figure 6. Model predictions. The panels plot the output of the model described in Figure 5 for various conditions. Each panel plots predicted response modulation among color-opponent channels as a function of retinal position and time. Zero represents no response modulation and therefore no predicted breakup. Greater values correspond to greater predicted modulation and therefore breakup. The left and right columns represent respectively the responses when the viewer's eyes are stationary and when the viewer tracks the stimulus. From top to bottom, the spatial offsets of the second subframe are −3, 0, 3, 6, and 9 arcmin, corresponding to $-\Delta x/3$, 0, $\Delta x/3$, $2\Delta x/3$, Δx, respectively. The offsets for the third subframe are twice those values. The simulated stimulus was a 1°-wide white target traveling at 6°/s on a black background. The capture rate was 40 Hz.

Methods

Subjects

Three subjects, 22–32 years of age, participated. All had normal or corrected-to-normal visual acuity, stereo acuity, and color vision. Two were authors; one was not aware of the experimental hypotheses. Appropriate consent and debriefing were done according to the Declarations of Helsinki.

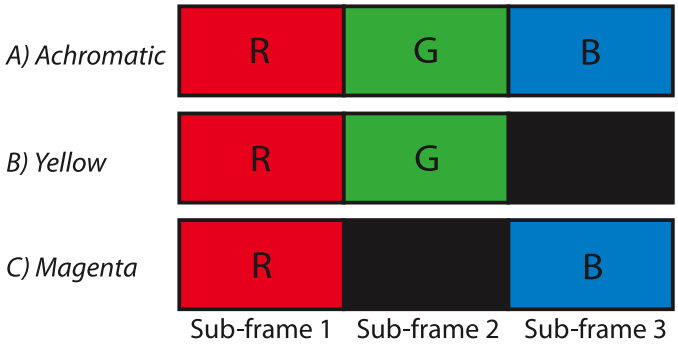

Figure 7. Protocols tested in the color-breakup experiment. The same sequential RGB protocol was used in all experiments, except that the blue or green channel was dropped in some conditions to create a yellow or magenta stimulus, respectively.

Apparatus

Stimuli were presented on a cathode-ray tube (CRT, ViewSonic G225f, Walnut, CA, USA; 1280 × 960 pixels, providing 640 × 960 pixels per each-eye view) and viewed binocularly. Viewing distance was 107.3 cm where each pixel subtended 1 arcmin. Refresh rate (or presentation rate) was 120 Hz. The capture rate was 40 Hz.

Stimuli and procedure

We tested the three presentation protocols depicted in Figure 7. All three decomposed color into red, green, and blue.

Yellow and magenta stimuli (Figure 7B, C) were used to test whether an offset in the second subframe relative to the first had the same effect as twice that offset in the third subframe relative to the first. The yellow stimulus presented light during the first and second subframes (RGX) and the magenta stimulus presented light during the first and third (RXB).

We tested two eye-movement conditions. In the non-tracking condition, we presented five object speeds (2, 6, 10, 20, and 30°/s, or 1, 3, 5, 10, and 15 pixels/frame); in the tracking condition, we presented only three (2, 6, and 10°/s) because tracking becomes inaccurate at speeds greater than 10°/s (Robinson, Gordon, & Gordon, 1986).

The stimulus consisted of two groups of bright rectangles on a dark background (Figure 8). Each rectangle was 1° high. The widths were 0.5, 1, 2, or 4°. The distance between adjacent rectangles was increased in proportion to stimulus speed so that no spatiotemporal aliasing occurred. In the tracking condition, the two groups of rectangles traveled horizontally at equal speed in the same direction. A fixation cross, located between the two groups, moved with the rectangles. To provide enough time for the subject to fixate, we

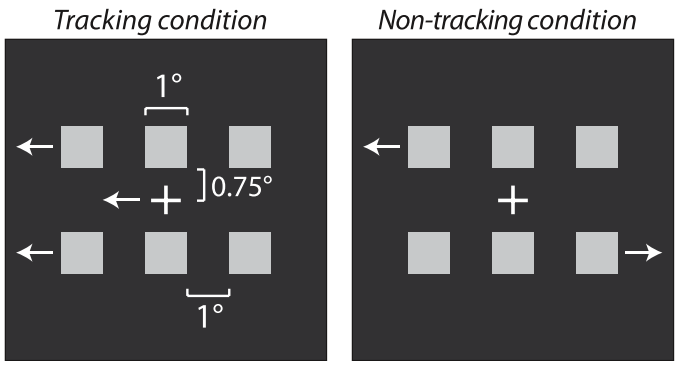

Figure 8. Stimulus used to measure color breakup. A series of bright rectangles moved across a dark background. In the tracking case, subjects tracked a cross that moved with the rectangles. In the non-tracking case, subjects fixated a stationary cross and the rectangles moved in opposite directions above and below the cross. In the figure, stimulus width is 1°, but the width varied between 0.5° and 4°.

presented the cross at its initial position for 0.5 s before the onset of the rectangles. When the moving rectangles appeared, the cross moved with them. The rectangles stimulus was presented for 1 s. In the non-tracking condition, the cross was presented midway between the rows of rectangles and was stationary. The two groups of rectangles traveled in opposite directions. By presenting opposing motions, we made it easier for subjects to hold the eyes stationary. In pilot testing, we noticed transient color breakup at the beginning and end of some presentations, probably due to reflexive tracking eye movements. To minimize such tracking, we faded the rectangles in and out at the beginning and end of each trial, for the non-tracking condition only. The fade in and fade out was not included in the model because the sensation of color breakup is strongest when the stimulus is brightest. After each presentation, subjects indicated with a key press whether or not they had perceived color breakup. They were instructed to respond positively only when they perceived hue variation at the edges. Thus, they did not respond if they perceived other motion artifacts such as judder or edge banding (Hoffman, Karasev, & Banks, 2011).

We examined how spatial offsets added to the second and third subframes affect breakup visibility. One offset was added to the green component, and twice that offset to the blue. For each condition, we presented nine different offsets according to the method of constant stimuli. For a tracked stimulus moving at speed v, the offset δ for the second subframe that should yield no breakup is:

$$\delta = \frac{v}{f}, \quad (1)$$

where f is the presentation rate. The predicted offset for the third subframe is 2δ. Using the same logic, we

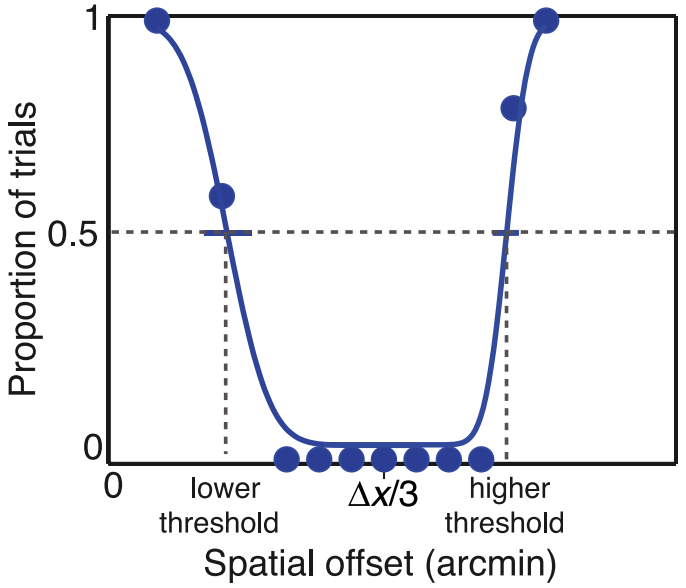

Figure 9. Psychometric function for perception of color breakup. Spatial offsets were added to the second and third subframes as described in the text. The proportion of trials in which color breakup was reported is plotted as a function of the offset of the second subframe. The data are from the achromatic condition with a stimulus speed of 10°/s and stationary fixation. The value of the offset that should eliminate breakup is $\Delta x/3$ (Equation 1). The curves are two cumulative Gaussians fit to the data using a maximum-likelihood criterion.

predict that the nulling offset for a yellow stimulus (in which only the first and second subframes are illuminated) would be δ added to the second subframe, and that the nulling offset for magenta (in which the first and third subframes are illuminated) would be 2δ added to the third subframe. Expressed as proportions of the displacement Δx per frame, δ and 2δ correspond to $\Delta x/3$ and $2\Delta x/3$, respectively.

Figure 9 illustrates typical data for one speed and presentation protocol. The proportion of trials in which color breakup was reported is plotted as a function of the spatial offset. We fit the data with two cumulative Gaussians, one on each side, using a maximum-likelihood criterion (Fründ, Haenel, & Wichmann, 2011; Wichmann & Hill, 2001a, 2001b). In all conditions and with all observers, there was a spatial offset that resulted in color breakup being reported on fewer than 20% of the trials, so we were able to reliably locate the descending and ascending portions of the psychometric data.

If we had tested all conditions, the experiment would have consisted of 10,800 trials per subject: 3 presentation protocols × 2 eye-movement conditions × 4 stimulus widths × 5 speeds × 9 offsets × 10 trials. Fewer trials were actually presented because some of those combinations of conditions were not realizable.

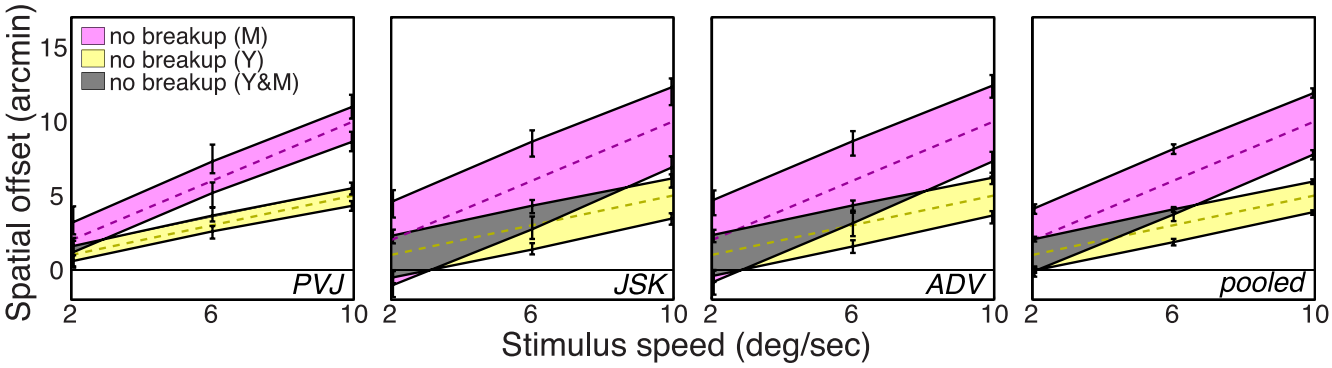

Figure 10. Color breakup when subjects tracked a yellow or magenta stimulus. Spatial offsets are plotted for different stimulus speeds. The shaded regions indicate the combinations of offsets and speeds that yielded breakup on fewer than half the trials. The first three panels show individual subject data, and the last panel shows data averaged across subjects. The yellow and magenta dashed lines correspond to the predicted nulling offsets for yellow and magenta, respectively (Equation 1). The yellow region corresponds to the range of offsets that produced color breakup on fewer than half the trials for the yellow stimulus, while the magenta region corresponds to the range of offsets that produced breakup on fewer than half the trials for the magenta stimulus. The gray region is the range of offsets that produced no color breakup for both stimuli. Error bars represent 95% confidence intervals.

About 5 hrs were required for each subject to complete the whole experiment.

Results

For each stimulus speed, we found the range of spatial offsets that yielded no or few reports of color breakup. The data figures indicate those ranges with shaded regions. The black lines at the edges of the shaded regions are the offsets that yielded reports of breakup on half the trials.

Figure 10 shows the results with the yellow and magenta stimuli. Recall that the first and second subframes were illuminated for the yellow stimulus (RGX) and the first and third for the magenta stimulus (RXB). These stimuli were used to test our assumption that a spatial offset of δ in the second subframe had the same effect on breakup as an offset of 2δ in the third subframe. Subjects tracked the stimulus during this experiment. The yellow and magenta shaded regions represent the offsets for which viewers perceived color breakup on fewer than half the trials for the yellow and magenta stimuli, respectively; the gray regions represent offsets for which viewers perceived breakup less than half the time for both stimuli. The dashed lines represent the predicted offsets required to eliminate breakup from Equation 1. As you can see, the experimental results are quite consistent with the predictions from that equation. In particular, the offset that eliminated breakup with the magenta stimulus was, as predicted, twice the offset that eliminated breakup with the yellow stimulus. In all subsequent figures, we will not plot the data for the third subframe because it is redundant.

Having confirmed that an offset of 2δ in the third subframe has an equivalent effect to an offset of δ in the second subframe, we proceeded with the main experiment with bright achromatic stimuli (i.e., R, G, and B illuminated in each frame). For each stimulus speed, we found how often color breakup was reported for different offsets of the second and third subframes relative to the first. Figure 11 shows the results for the individual subjects. The shaded regions represent the offsets for which viewers perceived color breakup on fewer than half the trials. As predicted, breakup was minimized in the tracking condition when the offset was $\Delta x/3$. That offset also minimized breakup in the non-tracking condition when the speed was 10°/s or slower. At faster speeds, the offset that minimized breakup was smaller than $\Delta x/3$. The upper row of Figure 12 shows the data averaged across subjects. The lower panels show the predictions from the model. The ranges of offsets that minimize breakup according to the model were quite similar to the range we observed empirically. They were essentially identical in the tracking condition and were quite similar in the non-tracking condition. Thus, the model seems to be an accurate predictor of the visibility of color breakup.

We observed in model simulations that the size of the moving stimulus had little effect on breakup visibility whether the eyes were moving or not. We investigated whether these predictions are consistent with viewers' percepts by presenting different stimulus widths (0.5, 1, 2, and 4.0°) and measuring breakup with stationary fixation. The upper row of Figure 13 shows those experimental results and the lower row

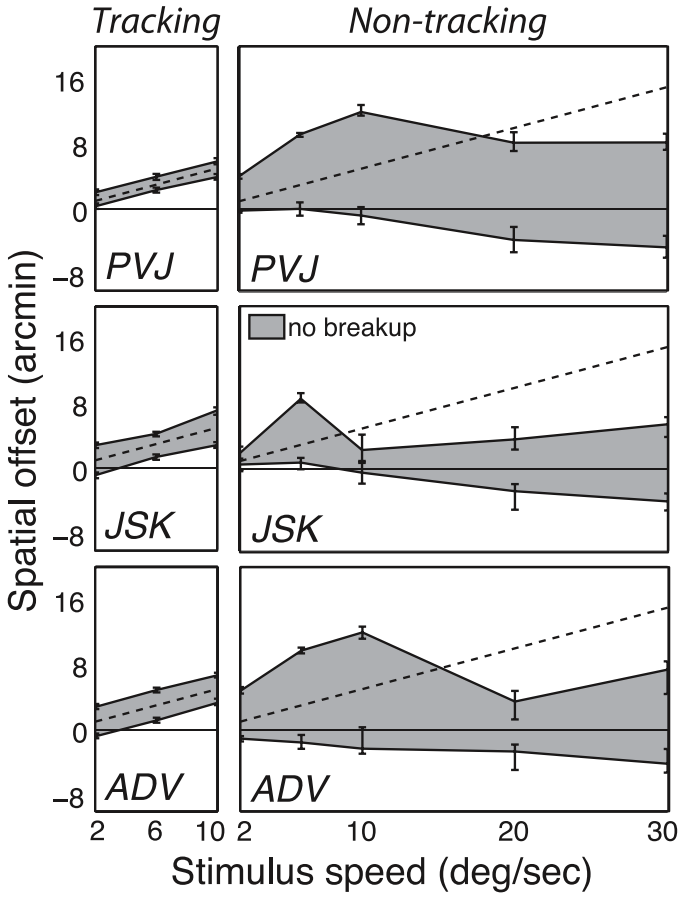

Figure 11. Color breakup as a function of stimulus speed and spatial offset of the second subframe relative to the first. The offset added to the third subframe was twice that of the second subframe. The left column shows the data for the tracking condition and the right column the data for the non-tracking condition. Each row represents the data from a different subject. The gray regions represent the offsets that yielded breakup on fewer than half the trials. The dashed lines represent the predictions of Equation 1. Error bars represent 95% confidence intervals.

the model predictions. As you can see, there was very little effect of stimulus size, confirming the model's predictions.

There is one condition for which the model's predictions and experimental behavior differ noticeably: where the stimulus width was 0.5° and speed was 20°/s. In that condition, the model predicts that breakup will be much less salient than we observed experimentally. That condition corresponds to stimulus movement such that Δx (the distance moved from one frame to the next) equals the object's width. In that special case, the luminances of R, G, and B remain constant but shifted spatially by Δx in each frame. With temporal low-pass filtering, each color component would be almost uniformly blurred across time, resulting in less salient breakup. However, small fixational eye movements

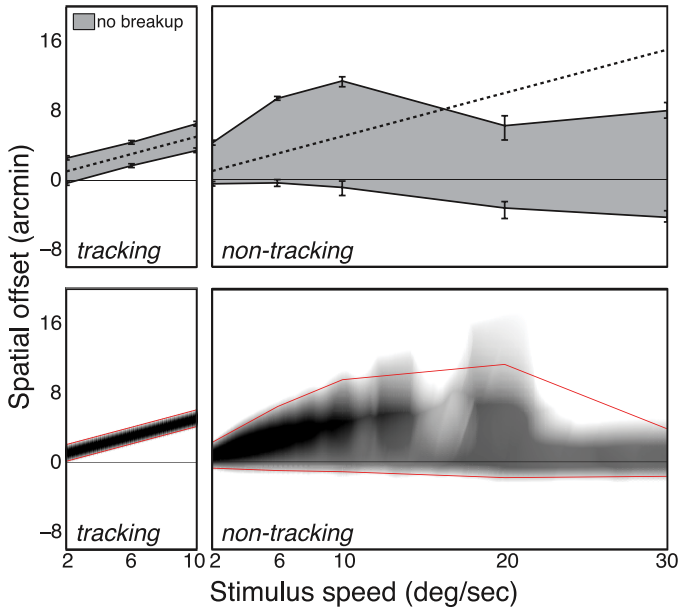

Figure 12. Top: Color-breakup data averaged across subjects. The spatial offset of the second subframe relative to the first is plotted as a function of stimulus speed. Left: Subjects tracked the stimulus. Right: Subjects maintained stationary fixation while the stimulus moved. The gray region represents the values of the spatial offset that yielded reports of color breakup on fewer than half the trials. Dashed lines represent the predictions of Equation 1. Error bars represent 95% confidence intervals. Bottom: Output of the color-breakup model. The dark regions correspond to the range of spatial offsets predicted to minimize breakup. The red contours represent the combinations of offsets and speeds that yielded a criterion amount of modulation at the speeds tested in the experiment.

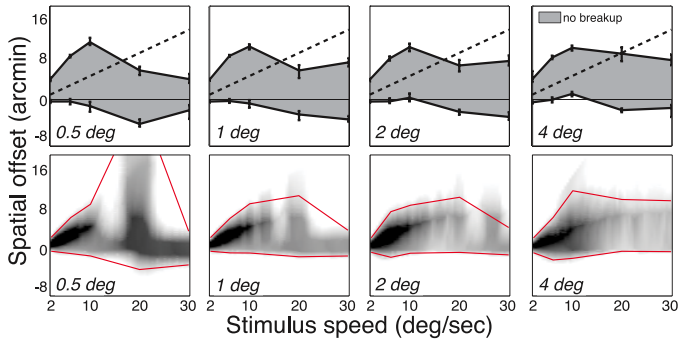

Figure 13. Top: Color breakup for four different stimulus widths (0.5°, 1.0°, 2.0°, 4.0°), non-tracking condition, pooled across subjects. The gray region represents the range of spatial offsets that yielded reports of color breakup on fewer than half the trials. Bottom: Color-breakup predictions for the same four stimulus widths. The dark regions correspond to the range of spatial offsets predicted to minimize color breakup. The red contours represent the combinations of offsets and speeds that yielded a criterion amount of modulation at the speeds tested in the experiment.

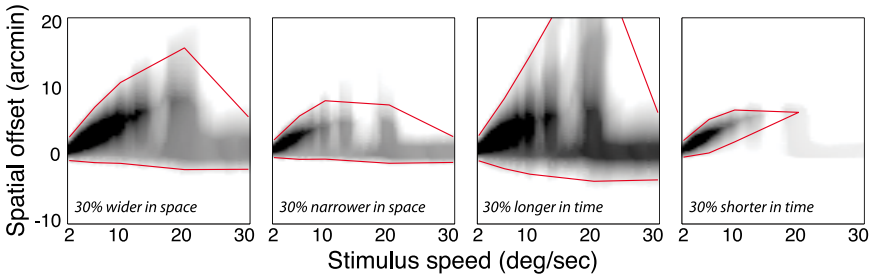

Figure 14. Properties of the impulse-response function and visibility of color breakup. The predicted visibility of color breakup from the output of the model is plotted for different spatial offsets and stimulus speeds, for the non-tracking condition. The stimulus was moving at 6°/s. We varied the spatial width and temporal duration of the impulse-response function (Figure 5) by ±30% to generate four different functions. From left to right, the panels show the results for impulse-response functions that are 30% wider in space, 30% narrower in space, 30% longer in time, and 30% shorter in time. The red contours represent the combinations of offsets and speeds that yielded a criterion amount of modulation at the speeds tested in the experiment.

(jitter, drift and refixation saccades) are ever-present, and would cause subframes to be displaced slightly on the retina, thereby avoiding the overlap that occurs with no eye movement. We examined the possibility that subjects' fixational movements caused breakup by incorporating such movements in the model. We found that these movements can indeed cause an increase in predicted breakup for that special condition. We did not, however, incorporate fixational movements into the model because their properties are difficult to measure and because their effect is probably only present in the special condition in which Δx equals the stimulus width.

The retinal eccentricity of the stimulus ranged from 1–10°. We wondered whether breakup was more visible in the peripheral or central visual field. One might argue that it would be most noticeable in the peripheral field where high temporal frequencies are more visible. But one might also argue that breakup would be most noticeable near the fovea where spatial resolution and sensitivity to hue variation are greatest. All subjects reported that breakup was most visible when the rectangles passed nearest the fovea suggesting that the poor resolution and reduced sensitivity of the retinal periphery make color breakup less noticeable in that part of the visual field.

Discussion

When a moving stimulus is presented by a color-sequential projector, color breakup is often observed. Consideration of the spatiotemporal filtering in color-opponent channels leads to the prediction that inserting specific spatial offsets to the second and third subframes will minimize or even eliminate breakup whether the viewer is tracking the stimulus or not. We confirmed this prediction in psychophysical experiments.

Shape of the impulse-response function

The similarity of the predictions and empirical observations supports the hypothesis that the visibility of color breakup is dependent on spatiotemporal filtering in early vision. We asked whether the similarity of predictions and results depends on specific filtering parameters or is more robust than that. To do this, we varied the spatial and temporal properties (width and duration) of the impulse-response function by ±30%, yielding four different functions. We then convolved those functions with the same input stimuli and reran the simulations. The main results were unaffected by these variations: A spatial offset of $\Delta x/3$ always minimized color breakup with tracking and non-tracking, and breakup saliency was always more dependent on spatial offset when the viewer was tracking than when fixation was stationary. Therefore, slightly different spatiotemporal impulse-response functions yield behavior that is consistent with our main conclusions. However, the quantitative agreement between the predictions and experimental observations is clearly dependent on the parameters of the assumed impulse-response function. This is illustrated in Figure 14, which shows some of the non-tracking results. The individual panels show the results for different impulse-response functions. The quantitative agreement between predictions and observations was best for the parameters derived from Burbeck and Kelly (1980) (Figure 12).

Pursuit gain

When people track an object with a smooth eye movement, the velocity of the eye rotation does not always match the velocity of the object, so retinal slip occurs. Such errors of smooth movements can be quantified by the pursuit gain. A gain of 1 means of

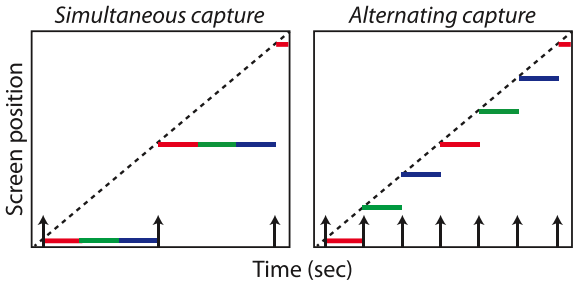

Figure 15. Simultaneous versus alternating capture at triple the rate. Left: In simultaneous capture, the red, green, and blue components of the image are captured at the same time (black arrows), and presented sequentially. Right: In alternating capture, the red, green, and blue components are captured alternately over time and presented sequentially.

course that the eye speed is perfectly matched to the object speed. With predictable object motion, as in our experiments, the gain is typically close to 1 for object speeds up to 10°/s; it falls at greater speeds (Robinson et al., 1986). In modeling our experimental data, we assumed a gain of 1 for the tracking condition because we did not present object speeds greater than 10°/s.

Saccades and breakup

Saccades are rapid, ballistic eye movements in which the eye rotates 100–500°/s (Becker, 1991). Such rapid movements should create quite salient color breakup even when the stimulus is stationary. Yohso and Ukai (2006) examined the effect of saccades of up to 300°/s and found that the width of the color breakup percept was narrower than predicted based simply on the retinal loci of stimulation. The reduced visibility is almost certainly due to *saccadic suppression*, the reduction in sensitivity to incoming visual information during a saccade (Matin, 1974). Although breakup is somewhat less visible during saccades relative to expectation, such eye movements remain an issue in color-sequential displays, particularly with high-contrast content. The problem is that saccades come at unpredictable times with unpredictable magnitudes and directions, so one cannot predict and compensate for them.

Triple capture to create spatial offsets

When an object's position changes by Δx every frame, adding spatial offsets of $\Delta x/3$ and $2\Delta x/3$ to every second and third subframe minimizes or even eliminates color breakup. In realistic video content, however, the value of Δx may vary from one object to another. Thus, different offsets may be needed for different objects depending on their speeds.

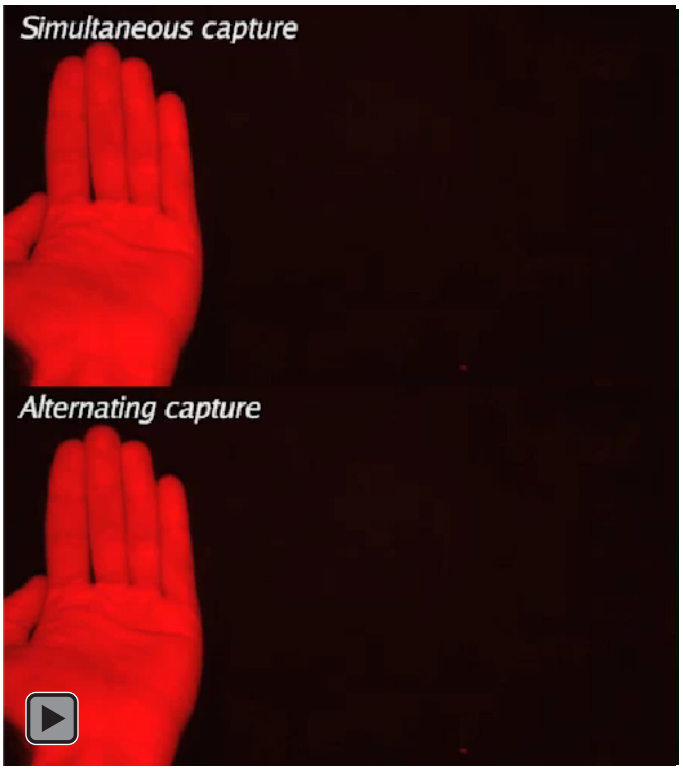

Figure 16. Demonstrations of color breakup with simultaneous and alternating capture. In the upper and lower videos, a bright hand moves left and right over a dark background. Upper: Simultaneous capture. The red, green, and blue components of the image are captured at the same time and presented sequentially. Color breakup is apparent, particularly when you track the object: Red and yellow fringes are visible at the leading edge, and blue and cyan fringes are visible at the trailing edge. Lower: Alternating capture. The red, green, and blue components are captured alternately at three times the rate as with simultaneous capture in the upper video. The captured images are then presented sequentially. Color breakup is minimized or eliminated.

One could in principle create the required offsets for each speed by using motion-compensated, frame-rate conversion (MCFRC; Kang, Cho, & Kim, 2007). MCFRC algorithms estimate motion between existing frames and create appropriate positions for intermediate frames. Thus, one could use such an algorithm to interpolate the appropriate second and third subframes and then use those interpolated frames to create the desired offsets. But MCRFC algorithms are computationally expensive and prone to artifacts when run in real time (Lee & Nguyen, 2010).

A simpler approach for creating the appropriate offsets would be to triple the capture rate. Figure 15 schematizes the idea. Instead of capturing content at f

and presenting at 3*f* (three subframes per frame), one would capture at 3*f* and present the appropriate colors in each of the three subframes. By tripling the capture rate, one would generate the first, second, and third subframes from the captured data and could use those frames to implement the appropriate spatial offsets for every speed. Thus, frames 1, 4, and 7 would be sent to the red channel; 2, 5, and 8 to the green channel; and 3, 6, and 9 to the blue channel. The desired spatial offsets would be baked into the content so no further manipulation would be needed.

This approach of capturing subframes at their appropriate spatial locations was used a long time ago in television. The first color television, demonstrated by Baird in 1928, used color-sequential presentation with a rotating color filter (Baird, 1929). The filter rotated in synchrony with the content such that the capture and presentation rates were identical. This method in principle could have eliminated color breakup, but the early color-sequential systems had frame rates that were too low to do so (Goldmark, Dyer, Piore, & Hollywood, 1942). The sequential technique for color presentation became obsolete when simultaneous color displays were introduced in 1947 (Kell, 1947).

Examples of the conventional protocol (simultaneous capture and alternating presentation) and the proposed protocol (alternating capture and presentation) are provided in Figure 16. The two videos display the same bright object moving across a dark background. The upper one depicts the conventional protocol in which red, green, and blue are captured simultaneously at a rate of *f* and then displayed alternately at a rate of 3*f* (red, green, and then blue). Color breakup is quite visible at the leading and trailed edges especially when you track the object. The lower video depicts the suggested protocol in which red, green, and blue are captured and displayed alternately at 3*f*. Breakup is now barely visible if at all. This shows that the suggested method could greatly reduce the visibility of color breakup.

Keywords: color breakup, DLP projector, motion, eye movements, color-opponent processing

Acknowledgments

Funding: NIH grant EY012851.

Commercial relationships: none.
Corresponding author: Paul V. Johnson.
Email: pvjohn98@gmail.com.
Address: University of California, Berkeley, Berkeley, CA, USA.

References

Arend, L., Lubin, J., Gille, J., & Larimer, J. (1994). Color breakup in sequentially-scanned LC displays. *SID Symposium Digest of Technical Papers*, *25*, 201–204.

Baird, J. L. (1929). *U.S. Patent No. 1,735,946*. Washington, DC: U.S. Patent and Trademark Office.

Becker, W. (1991). Saccades. In R. H. S. Carpenter (Ed.), *Eye movements: vision and visual dysfunction*. (pp. 95–137). Basingstoke, UK: Macmillian Press.

Burbeck, C. A., & Kelly, D. H. (1980). Spatiotemporal characteristics of visual mechanisms: Excitatory-inhibitory model. *Journal of the Optical Society of America A*, *70*, 1121–1126.

Burr, D. C., & Morrone, M. C. (1993). Impulse-response functions for chromatic and achromatic stimuli. *Journal of the Optical Society of America A*, *10*, 1706–1713.

Cheng, Y.-K., & Shieh, H.-P. D. (2009). Relative contrast sensitivity for color break-up evaluation in field-sequential-color LCDs. *Journal of Display Technology*, *5*, 379–384.

Fründ, I., Haenel, N. V., & Wichmann, F. A. (2011). Inference for psychometric functions in the presence of nonstationary behavior. *Journal of Vision*, *11*(6):16, 1–19, http://www.journalofvision.org/content/11/6/16, doi:10.1167/11.6.16. [PubMed] [Article]

Goldmark, P. C., Dyer, J. N., Piore, E. R., & Hollywood, J. M. (1942). Color television—Part 1. *Proceedings of the Institute of Radio Engineers*, *30*, 162–182.

Hoffman, D. M., Karasev, V. I., & Banks, M. S. (2011). Temporal presentation protocols in stereo displays: Flicker visibility, perceived motion, and perceived depth. *Journal of the Society for Information Display*, *19*, 255–281.

Hornbeck, L. J. (1997). Digital light processing for high-brightness high-resolution applications. *Proceedings of SPIE*, *3013*, 27–40.

Kang, S. J., Cho, K. R., & Kim, Y. H. (2007). Motion compensated frame rate up-conversion using extended bilateral motion estimation. *IEEE Transactions on Consumer Electronics*, *53*, 1759–1767.

Kell, R. D. (1947). An experimental simultaneous color-television system: Part I-Introduction. *Proceedings of the Institute of Radio Engineers*, *35*, 861–862.

Kelly, D. H. (1969). Diffusion model of linear flicker

responses. *Journal of the Optical Society of America*, *59*, 1665–1670.

Kelly, D. H. (1983). Spatiotemporal variation of chromatic and achromatic contrast thresholds. *Journal of the Optical Society of America*, *73*, 742–750.

Lee, Y.-L., & Nguyen, T. (2010). High frame rate motion compensated frame interpolation in high-definition video processing. *Acoustics Speech & Signal Processing, 2010 IEEE International Conference,* 858–861. IEEE.

Mason, S. J., & Zimmerman, H. J. (1960). *Electronic circuits, signals and systems*. New York: Wiley.

Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin*, *81*, 899–917.

Mullen, K. T. (1985). The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. *Journal of Physiology*, *359*, 381–400.

Poirson, A. B., & Wandell, B. A. (1993). Appearance of colored patterns: pattern—color separability. *Journal of the Optical Society of America A, 10*(12), 2458–2470.

Post, D. L., Monnier, P., & Calhoun, C. S. (1997). Predicting color breakup on field-sequential displays. *Proceedings of SPIE*, *3058*, 57–65.

Post, D. L., Nagy, A. L., Monnier, P., & Calhoun, C. S. (1998). Predicting color breakup on field-sequential displays: Part 2. *SID Symposium Digest of Technical Papers*, *29*, 1037–1040.

Robinson, D. A., Gordon, J. L., & Gordon, S. E. (1986). A model of the smooth pursuit eye movement system. *Biological Cybernetics*, *55*, 43–57.

Stork, D. G., & Falk, D. S. (1987). Temporal impulse responses from flicker sensitivities. *Journal of the Optical Society of America A, 4*(6), 1130–1135.

Tong, X., Heeger, D., & Van den Branden Lambrecht, C. J. (1999). Video quality evaluation using ST-CIELAB. *Proceedings of SPIE*, *3644*, 185–196.

Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, *63*, 1293–1313.

Wichmann, F. A., & Hill, N. J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics*, *63*, 1314–1329.

Winkler, S. (2005). *Digital video quality: Vision models and metrics*. Chichester, UK: John Wiley & Sons.

Yohso, A., & Ukai, K. (2006). How color break-up occurs in the human-visual system: The mechanism of the color break-up phenomenon. *Journal of the Society for Information Display*, *14*, 1127–1133.

Zhang, X., & Farrell, J. E. (2003). Sequential color breakup measured with induced saccades. *Proceedings of SPIE*, *5007*, 210–217.

Zhang, X., & Wandell, B. A. (1997). A spatial extension of CIELAB for digital color-image reproduction. *Journal of the Society for Information Display*, *5*, 61–63.

Appendix A

Here we provide the details of the spatiotemporal filtering model we developed. It is designed to predict the color-opponent response modulation that the stimuli in our experiments would create. We assume that the visibility of color breakup is monotonically related to the response modulation.

The stimulus was a bright achromatic rectangle on a dark background. The rectangle moved horizontally, so we considered only the horizontal dimension of space. We first defined the input stimulus in two-dimensional space-time coordinates as in Figure 2. The input is in retinal coordinates. We were interested mainly in the chromaticity of the percept, so we converted the RGB values of the input stimuli into values in opponent color space, which approximates the luminance and chromatic channels of the visual system (Poirson & Wandell, 1993; Zhang & Wandell, 1997). First, we linearized the RGB values. Because the stimulus in our simulation was a white target on a black background, we set the linearized RGB to [1 1 1] in the target and [0 0 0] in the background. We then transformed the linearized RGB values into tristimulus values in the CIE XYZ color space. If the values conform to sRGB IEC61966-2.1, a standard color space, the transformation is (Winkler, 2005):

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412 & 0.358 & 0.180 \\ 0.213 & 0.715 & 0.072 \\ 0.019 & 0.119 & 0.950 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \text{(A1)}$$

The transformation from CIE XYZ color space to opponent color space is (Zhang & Wandell, 1996):

$$\begin{bmatrix} A \\ C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 0.279 & 0.720 & 0.107 \\ -0.449 & 0.290 & 0.077 \\ 0.086 & 0.590 & 0.501 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad \text{(A2)}$$

As the displayed image changes over time, the luminance of pixels changes accordingly. Ideally, temporal changes would be instantaneous, but the display has its own temporal response. We incorporated the temporal properties of the display (ViewSonic

G225f) by measuring its temporal impulse-response function (IRF). The function was very close to an ideal exponential decay with a time constant of 1.5 msec. All simulation results included the display's IRF. (When we replaced the IRF with a delta function, the simulation results were nearly identical, which means that the display's IRF was short relative to the assumed human IRF.)

The resulting values in opponent color space, passed through the display's IRF, were then convolved with the IRFs of the two chromatic channels. To calculate IRFs, we adopted the red-green channel's contrast sensitivity function (CSF) as described by Burbeck and Kelly (1980) and Kelly (1983) (Equation A3). The CSF is the sum of excitatory (E) and inhibitory (I) components; it is not separable in the space-time domain even though the excitatory and inhibitory components are separable. E_s, E_t, I_s, and I_t are defined, respectively, as the spatial response of the excitatory component, temporal response of the excitatory component, spatial response of the inhibitory component, and temporal response of the inhibitory component. The last equation under A3 is commonly used in deriving the other equations in A3. f_{s1}, f_{s2}, f_{t1}, and f_{t2} are constants that depend on individual variances. We adopted values for f_{s1}, f_{s2}, f_{t1}, and f_{t2} measured by Burbeck and Kelly (1980): 10 cpd, 0.5 cpd, 19 Hz, and 1 Hz.

$$
\begin{aligned}
CSF(f_x,f_t) &= E(f_x,f_t) + I(f_x,f_t)\\
&= \frac{E_s(f_x)E_t(f_t)}{E_s(f_{x2})} + \frac{I_s(f_x)I_t(f_t)}{I_s(f_{x2})}
\end{aligned}
$$

$$
E_s(f_x) = \begin{cases} \dfrac{K(f_{x1},f_{t2})}{K(f_{x1},f_{t1})}K(f_x,f_{t1}) & \text{for } f_x \le f_{x1} \\ K(f_x,f_{t2}) & \text{for } f_x > f_{x1} \end{cases}
$$

$$
E_t(f_x) = \begin{cases} \dfrac{K(f_{x2},f_{t1})}{K(f_{x1},f_{t1})}K(f_{x1},f_t) & \text{for } f_t \le f_{t1} \\ K(f_{x2},f_t) & \text{for } f_t > f_{t1} \end{cases}
$$

$$
I_s(f_x) = E_s(f_x) - K(f_x,f_t)
$$

$$
I_t(f_x) = E_t(f_x) - K(f_x,f_t)
$$

$$
K(f_x,f_t) = 4\pi^2 f_x f_t \left[6.1 + 7.3 \left| \log_{10}\left(\frac{f_t}{3f_x}\right)\right|^3 \right] e^{-4\pi\frac{2f_x+f_t}{45.9}} \tag{A3}
$$

The CSF has no phase information, so we had to reconstruct phase, which we did by extending the assumption of a minimum-phase filter (Stork & Falk, 1987; Kelly, 1969; Mason & Zimmermann, 1960) to the space-time domain. Such an assumption is reasonable for chromatic channels (Burr & Morrone, 1993). We assume a complex transfer function, H, which is the Fourier transform of the IRF of the visual system, h.

$$
\begin{aligned}
&\iint h(x,t)e^{-2\pi i(f_x x + f_t t)}dxdt\\
&\quad = H(f_x,f_t) = |H(f_x,f_t)|e^{i\theta(f_x,f_t)}.
\end{aligned} \tag{A4}
$$

Here, the modulus is the same as the CSF. Our goal was to estimate $\theta(f_x, f_t)$ in order to reverse-engineer h. First, we computed the logarithm of the transfer function. The log of the modulus and the unknown phase become, respectively, the real and the imaginary parts:

$$
\log[H(f_x,f_t)] = \log[|H(f_x,f_t)|] + i\theta(f_x,f_t) \tag{A5}
$$

If these real and imaginary parts are a Hilbert transform pair, then h is real and causal (Mason & Zimmermann, 1960). The Hilbert transform pair is not the only solution for causal and real h, but it is the solution that satisfies minimum phase. For a signal k whose independent variable is t, the Hilbert transform is

$$
K_{Hi}(t) = \frac{-1}{\pi t} * k(t) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{k(t')dt'}{t'-t}, \tag{A6}
$$

where the Cauchy principal value takes care of the integration around the singularity at $t = t'$. In our case, we want a causal signal in time such that

$$
h(x,t) = 0 \quad \text{for } t < 0 \tag{A7}
$$

However, we do not need such a constraint in the space dimension. Thus we apply the Hilbert transform along the temporal frequency dimension only. The convolution and transforming integration that we used were:

$$
\begin{aligned}
\theta(f_x,f_t) &= \left[\frac{-\delta(f_x)}{\pi f_t}\right] * \log[|H(f_x,f_t)|]\\
&= \frac{1}{\pi}\iint \frac{\delta(f_x - f_x')}{f_t' - f_t}\log\left[|H(f_x',f_t')|\right]df_x'f_t',
\end{aligned} \tag{A8}
$$

where δ is the Kronecker delta function. With the phase term estimated, we calculated the IRF $h(x,t)$ by taking the inverse Fourier transform of $H(f_x, f_t)$. The calculated IRF was causal and real. We also used the same IRF for the yellow-blue channel because it is very similar to its red-green analog (Mullen, 1985; Tong, Heeger, & van den Branden Lambrecht, 1999).