Information for perceiving blurry events: Optic flow and color are additive

Hongge Xu¹ · Jing Samantha Pan^{1,2} · Xiaoye Michael Wang³ · Geoffrey P. Bingham⁴

Accepted: 1 September 2020 / Published online: 1 October 2020
© The Psychonomic Society, Inc. 2020

Abstract

Information used in visual event perception includes both static image structure projected from opaque object surfaces and dynamic optic flow generated by motion. Events presented in static blurry grayscale displays have been shown to be recognized only when and after presented with optic flow. In this study, we investigate the effects of optic flow and color on identifying blurry events by studying the identification accuracy and eye-movement patterns. Three types of color displays were tested: grayscale, original colors, or rearranged colors (where the RGB values of the original colors were adjusted). In each color condition, participants identified 12 blurry events in five experimental phases. In the first two phases, static blurry images were presented alone or sequentially with a motion mask between consecutive frames, and identification was poor. In Phase 3, where optic flow was added, identification was comparably good. In Phases 4 and 5, motion was removed, but identification remained good. Thus, optic flow improved event identification during and after its presentation. Color also improved performance, where participants were consistently better at identifying color displays than grayscale or rearranged color displays. Importantly, the effects of optic flow and color were additive. Finally, in both motion and postmotion phases, a significant portion of eye fixations fell in strong optic flow areas, suggesting that participants continued to look where flow was available even after it stopped. We infer that optic flow specified depth structure in the blurry image structure and yielded an improvement in identification from static blurry images.

Keywords Optic flow · Event perception · Color · Eye movement

Public significance statement In everyday life, we perceive and interact with moving objects (known as visual events). In this study, we presented blurry images and blurry movies in their original colors, altered colors, or in grayscale. Observers better identified blurry events when incorporating motion to the otherwise unrecognizable events. Furthermore, color helped event identification in static images as well as in movies. Thus, both motion-generated optic flow and color facilitate event identification, but the two do not interact. Results and implications of this study are relevant to the functional evaluation and rehabilitation of extreme blurry vision, such as in AMD and amblyopic patients.

✉ Jing Samantha Pan
panj27@mail.sysu.edu.cn

✉ Xiaoye Michael Wang
xywang01@yorku.ca

¹ Department of Psychology, Sun Yat-sen University, Guangzhou, China

² Guangdong Provincial Key Laboratory of Social Cognitive Neuroscience and Mental Health, Guangzhou, China

³ Center for Vision Research, York University, Toronto, ON, Canada

⁴ Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA

Perceiving events, which consist of objects in motion, is ubiquitous in one's daily life. Under natural viewing conditions, two sources of optical information specify events. First, objects' opaque surfaces reflect light and produce image structure, including contrasts and gradients of luminance and chromaticity. Image information is static and therefore is weak in specifying depth and 3D properties of objects or events. Observers normally need to combine various types of image information, such as relative size and texture gradients, to recover depth (Riesenhuber & Poggio, 2000; Tarr & Bulthoff, 1998). Ambiguities may arise in figure-ground relations and depth scaling (e.g., Norman, Todd, & Phillips, 1995; Todd, Christensen, & Guckes, 2010). However, image information is stable: It exists as long as objects and surfaces in a scene are visible. This stability is important for guiding actions, because online guidance requires that the relevant surfaces remain in view (Pan, Bingham, & Bingham, 2013; Zhao & Warren, 2017).

Second, object motion yields continuous changes in the optics. This transformation in light structure, or optic flow, lawfully corresponds to properties of objects and of their

motion. Optic flow, in turn, specifies the characteristics of 3D motions (e.g., Li & Warren, 2000, 2002; Loula, Prasad, Harber & Shiffrar 2005; Neri, Morrone, & Burr, 1998; Warren & Hannon, 1990), even without image-based information (e.g., Beardworth & Bucker, 1981; Jokisch & Troje, 2003; Pavlova & Sokolov, 2000). In addition, motion can also specify an object's 3D structure through a process known as structure from motion (SFM; Koenderink & Van Doorn, 1991). SFM specifies surface slant (Wang, Lind, & Bingham, 2018) and object shapes (e.g., Jiang, Jiang, & Parasuraman 2015; Lee, Lind, Bingham, & Bingham, 2012; Norman, Beers, Holmin, & Boswell, 2010; Norman & Todd, 1993), and facilitates grasping (Lee & Bingham, 2010). The drawback of optic flow is its lack of stability—that is, information is only available when objects are moving and ceases to exist immediately when motion stops.

In visual event perception, both image structure and optic flow are available. Image structure is weak in specifying object motion and depth structure, but is temporally stable. Optic flow is powerful in specifying motion and depth, but is transient. We propose that the visual system combines them to achieve perceptual accuracy and stability (Pan, Bingham, & Bingham, 2013). Specifically, when an object moves, optic flow specifies the spatiotemporal properties of objects and events and connects discrete image structure. Consequently, static images were enriched with spatiotemporal meaning. After motion stops, the calibrated image structure remains and preserves depth structure and motion properties that optic flow previously specified. Offloading perceived properties to stable image structure allows observers to access and act upon such perceived properties without having to hold them in the mind. In this way, the interaction between optic flow and image structure elicits spatially accurate and temporally stable visual perception.

The optic-flow–image-structure interaction makes perception accurate when image information is blurry. Norman et al. (2010) asked normally sighted participants to wear 2, 2.5, and 3-diopter convex lenses to look at objects rotating in depth and performed a 3D shape discrimination task. They found that when motion-generated information was available during rotation, discrimination accuracy increased for all levels of the simulated blur, where the d' value for 3-diopter blur ($d' = 3.05$) was similar to the d' value with no blur ($d' = 3.65$). Moreover, Bochsler and colleagues (Bochsler, Legge, Gage, & Kallie, 2013; Bochsler, Legge, Kallie, & Gage, 2012) found that both normal observers with simulated visual blur and low-vision observers were better at identifying surface orientation and telling whether the floor ahead of them was a flat surface, a ramp (tilting up or down), or a step (rising or dropping) if they first walked 10 ft and then attempted the orientation discrimination task.

Furthermore, the optic-flow–image-structure interaction makes perception robust and tolerant of image distortions

when, for example, images are blurry (Pan & Bingham, 2013), shifted (Pan, Bingham, & Bingham, 2017a), or temporarily invisible (Pan, Bingham, & Bingham, 2013). In this context, image structure does not specify spatiotemporal properties in the world per se, but instead, it preserves such properties revealed through optic flow, rendering high image quality unnecessary. For example, in the case of blurry images, normally sighted observers as well as age-related macular degeneration and amblyopia patients were unable to identify events if only given blurry static images, but could readily do so if images were in motion (Pan & Bingham, 2013; Pan et al., 2017b). More interestingly, after having identified the events from blurry videos, normal and clinical observers were able to identify events from similar static blurry pictures. The successful identification lasted for a week and was independent of whether observers remembered what they had previously seen. These results supported the theoretical proposal that optic flow and image structure are functionally complementary, and their interaction yields accurate and stable event perception.

The current study uses a similar experimental paradigm to further explore this interaction between optic flow and image structure and uses eye tracking to investigate whether the eyes would fixate on areas of an event with optic flow information, when the optic flow is available and after it calibrated static image structures. Event identification based on static images requires individuating objects through image-based information, which requires relatively high spatial frequency. Without this, observers may be unable to discriminate depth edges and objects from other high-contrast elements in an image. For example, when seeing a blurry picture of someone wearing a graphic T-shirt, it might be difficult to differentiate the depth edges around the head, arm, or body from high-contrast patterns on the T-shirt. Such confusion weakens the ability to segment objects and surfaces in depth (Greene & Oliva, 2009). The ability to separate surfaces in depth and group edges is a prerequisite for recognition and identification (Nakayama, Shimojo, & Silverman, 1989; Palmer & Brooks, 2008). Thus, observers viewing blurry images should be expected to scan an entire image looking for areas of high contrast to identify objects and events.

However, when multiple static blurry frames of an event are played continuously, motion is detectable (given that image blur does not generally affect monocular motion detection; Jobling, Mansfield, Legge, & Menge, 1997), generating optic-flow information. The magnitude of optic flow covaries with depth and thus specifies the spatial layout of surfaces in the scene (i.e., depth edges and surface boundaries). Once motion stops, the blurry images remain and the relevant elements in image structure have been identified and thus can preserve the spatial order that was once specified by optic flow. So, optic flow separates surfaces in depth, and image structure might preserve such information after flow

stops. Accordingly, observers should fixate on depth edges rather than arbitrary lines of high contrast in blurry stimuli *while* and *after* being exposed to optic flow.

Moreover, we also explore the role of color, as a part of image structure, in the interaction between optic flow and image structure. In previous studies, we only used grayscale images that contained luminance, contrast, and motion information (Pan & Bingham, 2013; Pan et al., 2017b). These are sufficient for visual motion perception, but in addition, luminance-defined contrasts are necessary for visual motion detection. Motion displays consisting of only isoluminant or equiluminant colors (e.g., hues) have been found to be entirely ineffective, allowing no visual motion perception (Bilodeau & Faubert, 1997; Cavanagh & Anstis, 1991; Lu, Lesmes, & Sperling, 1999; Ramachandran & Gregory, 1978; Ruppertsberg, Wuerger, & Bertamini, 2003; Teller & Lindsey, 1993; Wuerger & Landy, 1993). Thus, color is a dimension, independent from motion, that carries informative structure in images. Given this, we investigated what color might add to an observer's ability to identify events in blurry images without or with visual motion information.

Research in early visual processing suggests that luminance, chromaticity, and motion are three basic types of information for visual recognition (Livingstone & Hubel 1987). Color, which includes luminance and chromaticity, improves object recognition (Bramão, Reis, Petersson, & Faisca 2011; Shevell & Kingdom, 2008), facial recognition (Yip & Sinha, 2002), and memory retrieval (Gegenfurtner, Wichmann, & Sharpe, 1998). However, it remains unclear whether and how color, motion, and image blur might interact when viewing events. Color is semantically rich, and color diagnosticity, which refers to the strong association between a particular color and a particular object (e.g., a banana is yellow, and a leaf is green), may exert top-down effects on event identification. Some studies show that high color diagnosticity leads to better object recognition (Tanaka & Presnell, 1999), and recognition of blurry scenes (Oliva & Schyns, 2000). For example, Oliva and Schyns (2000) showed that when categorizing scenes from stimuli containing only low spatial frequencies, adding diagnostic colors to luminance-only displays enhanced performance. On the other hand, some studies found that color aided object recognition, but not because of color diagnosticity; rather, it is because color facilitates feature and contour extraction in an image (De Valois & Switkes, 1983) and improves object segmentation and hence recognition (Gegenfurtner & Rieger, 2000). Yip and Sinha (2002) presented blurry faces in their original color, in pseudocolor (created by turning a color wheel to randomize colors), or in grayscale. They found no difference in gender identification performance between blurry faces presented in original colors and in pseudocolors, but

performance was poorer when faces were blurry and grayscale. The authors argued that the presence of color, albeit unnatural, helped item segmentation within an image that was subsequently used to perform the recognition task. In this regard, color and motion play similar roles.

In the current study, we used 12 blurry event videos of common, everyday events. Each video was 20 frames in length. We used individual frames or static images from each video as well as the entire video as such. We presented static images (premotion phase), videos (motion phase), and static images again (postmotion phase), to investigate two questions. First, using eye tracking, we explored how optic flow and blurry image structure interact. We extracted optic flow from each event and collected eye movement data when observers looked at static images in the postmotion phase to study whether improvement in performance occurs because observers know where to look. If, in the postmotion phases, observers become better at identifying blurry events as compared with the premotion phases, and if observers look at parts of the static images where optic flow is strong in the previous motion phase, then it would support the hypothesis that spatial relations specified by the optic flow information are preserved in the static image structure information, and continues to be accessible after motion stops. Second, we examined how color interacts with optic flow, and how it facilitates event identification (as being “diagnostic” or enhancing image segmentation). We applied color manipulations and presented blurry events in grayscale, original colors, or rearranged colors. If performance in all color treatment types is equivalent, then there is no color advantage. If performance with colored stimuli is better than with rearranged color or grayscale stimuli, then color helps event identification because of color diagnosticity. If performance with colored and with rearranged color stimuli are equivalent and both are better than that with grayscale stimuli, then color helps event identification by aiding image segmentation. Moreover, looking at how performance varies in the premotion, motion, and postmotion phases for the three color types would inform us about the nature of the relationship between optic flow and color.

Method

Participants

Thirty-six participants from Sun Yat-sen University (mean age 21.5 ± 2.0 years, 18 females) completed this experiment and were compensated at the rate of ¥30/hr. All participants had normal or corrected-to-normal vision (mean visual acuity 0.2 ± 0.23 , intact color vision). This study was approved by the Institutional Review Board of the Department of Psychology, Sun Yat-sen University. Informed consent was obtained from all participants.

Stimuli and apparatus

Twelve daily events were used. We either recorded (Nikon D7000 digital video camera) or purchased (from <https://www.paixin.com>) footage of daily events, including a person bowling, two people dancing tango, a man walking toward the camera and waving, children skipping rope, a basketball flying to the hoop, a top spinning on a chessboard, eggs being beaten, a cue ball striking another ball, three empty swings swinging, dandelion being blown, flags flying in the wind, and a campfire with dancing flames (see Fig. 1).

All videos had a resolution of $2,056 \times 1,156$ and were processed using MATLAB 2014a (The MathWorks Inc., Natick, MA). We first applied a Gaussian blur with a standard deviation of 28 pixels and a cutoff frequency of 0.006 cycles/pixel (or, equivalently, 0.28 cycles/degree visual angle). The blurry images used in the experiment were similar to someone with visual acuity of 20/200.¹ Subsequently, we selected 20 frames from each video and adjusted colors in them. Colored stimuli retained the original colors. Grayscale stimuli retained the original luminance, but hues were removed. Rearranged color stimuli had the RGB values randomly switched. In rearranged color stimuli, all 20 frames from each event went through the same color switching, and the total luminance was not lower than in color or grayscale images.

Stimuli were displayed using MATLAB Psychophysics Toolbox (Brainard 1997; Kleiner et al., 2007; Pelli & Vision, 1997) on a 27-inch ASUS PG279Q monitor with resolution of $2,560 \times 1,440$ at 144 Hz refresh rate and in a room with the luminance of 31.4 cd/m^2 . The stimuli were $43.6^\circ \times 25.4^\circ$ at 60-cm viewing distance.

We tracked eye movement using an EyeLink 1000 Plus Desktop Mount eye tracker (SR Research Inc., Mississauga, Canada) controlled using a native EyeLink host PC. This system included a chin and forehead rest to minimize head movement. We collected monocular eye position from participants' dominant eye at a sampling frequency of 1000 Hz. We used a nine-point target display to calibrate and validate the system. We would recalibrate if validation showed a $\pm 1.5^\circ$ deviation. For each trial, we also performed drift correction using a single target before presenting the actual stimuli.

¹ Four normally sighted observers looked at the static blurry testing stimuli and then looked at unblurred images extracted from events used in the experiment through positive lenses. The lenses were placed in front of the dominant eye, one at a time, until observers felt the blur to be equivalent to the blur in the experimental stimuli. We repeated the lens fitting three times and computed average lens power for each observer. Then, looking through the average lens, the observer's near and distant visions were tested by using the Standard Near Vision eye chart and the Tumbling E eye chart. Results showed that the blurry images in the experiment were similar to someone with visual acuity of approximately 20/200.

Procedures

We randomly assigned participants to one of the three color conditions. All conditions had identical testing procedures. In all conditions, participants completed five experimental phases, with the first four phases in the first session, and the fifth phase in the second session conducted after 5 days. Participants sat in front of a computer monitor and placed their chins on a chin rest, at a viewing distance of 60 cm. The EyeLink Desktop Mount was attached to the chin rest. We calibrated and validated the eye-tracking system at the beginning of each session and after breaks. We informed participants that they would see blurry images or videos of some daily events, and that they should identify and describe those events. During each trial, participants viewed the stimuli, and once they had an answer, they pressed “J” on the keyboard to stop eye tracking, and verbally reported their answers. An experimenter recorded participants' responses.

In Phase 1, we randomly selected three static frames from each of the 12 events and presented one frame on a given trial (36 trials). Viewing time was not restricted, and the frame remained onscreen until the participant pressed the “J” key. In Phase 2, we presented all frames of an event sequentially and inserted a 1,000 ms random dot noise map (motion mask) between consecutive frames in each trial (12 trials). Each frame was displayed for 250 ms, and total duration of an event in this phase was 25 s. In this phase, the stimuli contained full image information, but no motion information. In Phase 3, we removed the motion mask and presented frames consecutively at 250-ms/frame. The total duration of an event in this phase was 5 s. Both image information and optic flow were available in this phase. We looped each event sequence in Phases 2 and 3 until participants pressed “J” and provided a response. Finally, Phase 4 was similar to Phase 1, where three different frames from each event were randomly selected and presented. There was a total of 96 trials in the first session. After 5 days, participants came back for the second session. Phase 5 followed the same procedure as Phases 1 and 4. In Phases 4 and 5, stimuli were onscreen until participants pressed “J” to end the trial. Experimental demo is uploaded (<https://youtu.be/LQIFO8MH-1I>).

Identification data

Two raters watched both blurred experimental stimuli and the original videos, and independently evaluated participants' responses. We instructed the raters that a correct response should capture the gist of an event or describe its details. For example, for the bowling image in Fig. 1, responses like “a person bowling” or “a person standing with the arm out holding something” would both be correct. The interrater reliability for this assessment is 0.984 (Cohen's kappa; Fleiss & Cohen, 1973). We used a third rater to code trials with rater

Fig. 1 Example frames from two events (a person bowling and flags flying in the wind) in the unprocessed (first column), blurry grayscale (second column), blurry colored (third column), and blurry rearranged

color (fourth column) footages. See online video for more information: <https://youtu.be/LQIFO8MH-1I>

discrepancy. We used a mixed-design analysis of variance (ANOVA) to analyze the correct rate of identification with three factors: experimental phase (five levels, within subject), events (12 levels, within subject), and color treatment type (three levels, between subject). We also analyzed viewing time (the time taken to look at the stimulus and complete one trial) in different phases and with different types of color treatment to see whether and how viewing time and correctness of identification were related.

Eye-movement data

We used EyeLink Data Viewer to process raw fixation data. We set saccade velocity threshold at $30^\circ/\text{s}$, acceleration threshold at $8,000^\circ/\text{s}^2$, and saccadic motion threshold at 0.1° . Areas of interest (AOIs) were hand-drawn around event-related regions. The event-related locations in each frame was defined by several free-hand drawing closed graphs, which included both moving objects and objects interacting with them (e.g., a moving basketball and the hoop).

We used linear mixed-effect models to analyze fixation patterns in the correct trials because this method handles unequal correct trials in the various experimental phases (Bates, Mächler, Bolker, & Walker 2015). We used proportion of fixation counts and proportion of fixation duration in AOIs as dependent variables, and treated experimental phase and color treatment type as fixed effects, and event and participant as random effects.

Optic flow processing

To extract optic flow information, we used Sun and colleague's Classic+NL optic flow estimation method (MATLAB open implementation; Sun, Roth, & Black, 2010, 2014). We used every two consecutive frames in each event to compute one optic flow map, resulting in 19 flow maps for each event. We applied the min-max normalization method to scale the flow data and set one standard deviation

above the mean as the flow threshold. If the flow's magnitude on a pixel exceeded the threshold, it was coded as 1, or otherwise 0. We then computed the overlap between fixation points and the derived optic flow matrix and computed the proportion of such overlap over total fixation points. The resulting percentage (called overlap proportion) reflected how much looking occurred in areas where optic flow was or had been present. We analyzed how overlap proportion changed in Phases 3, 4 and 5 with the three color types. Optic flow had never been available in Phases 1 or 2, so overlap was not defined in these phases. In Phases 4 and 5, optic flow had calibrated the static images (by hypothesis), so we used the flow maps computed using these images' subsequent frames (or previous frame if the frame was at the end of the sequence) to indicate where optic flow had been. We analyzed optic flow and looking behaviors in trials where the responses were correct.

Results

Identification results

In this experiment, we tested the identification of blurry grayscale, color or rearranged color events in five experimental phases and found significant main effects of experimental phase, $F(1.6, 53.3) = 376.165$, $p < .001$, $\eta_p^2 = 0.92$; event, $F(7.8, 258.2) = 9.12$, $p < .001$, $\eta_p^2 = 0.22$; and color treatment type, $F(2, 33) = 39.16$, $p < .001$, $\eta_p^2 = 0.70$, all with Greenhouse–Geisser corrections for sphericity (see Fig. 2a). The Experimental Phase $\times$ Event interaction, $F(13.37, 441.2) = 5.99$, $p < .01$, $\eta_p^2 = 0.15$, and Color Treatment Type $\times$ Event interaction, $F(22, 363) = 4.08$, $p < .001$, $\eta_p^2 = 0.20$, were both significant. Although event affected identification rate and interacted with other factors, the effect sizes were much smaller than those of phase and color treatment type. The Experimental Phase $\times$ Color Treatment Type interaction was

Fig. 2. **a** Mean proportion of correct event identification for the three color types across five experimental phases. **b** Mean improved proportion of correct identification in Phases 2, 3, 4, and 5 compared with Phase 1. **c** Mean fixation counts decreased from Phase 1 to Phases 4 and 5. **d** Mean fixation duration decreases from Phase 1 to Phases 4 and

5. Error bars in **a** and **b**: within-subjects repeated-measures 95% CIs, calculated using the Cousineau–Morey method (Baguley, 2012; Morey, 2008); error bars in **c** and **d**: 95% CIs calculated using the traditional method

not significant ($p > .30$). Post hoc analysis with Bonferroni correction showed that Phase 1 (static images) had a significantly lower identification rate than all other phases ($p < .001$); Phase 2 (motion mask) had a significantly lower identification rate than Phases 3, 4, and 5 ($p < .001$); Phase 3 had higher identification rate than all other phases. Identification rates did not differ between Phases 4 and 5 ($p > .05$), and both were higher than in Phases 1 and 2 ($p < .001$).

Unaltered stimulus color improved performance. Correct identification in the color condition was consistently higher by about 30% than in the grayscale and rearranged color conditions ($p < .01$). There was no difference between the grayscale and rearranged color conditions ($p > .30$). When treating Phase 1's performance as baseline and looking at improvement in performance in subsequent phases, there was no difference among color treatment type, $F(2, 33) = 0.54$, $p = 0.59$, $\eta_p^2 = 0.03$ (see Fig. 2b). Thus, natural stimuli color aided event recognition in static images, while motion equally improved event recognition across all color treatment type.

In addition to being more accurate, viewing time in the correct trials of postmotion Phases 4 and 5 was shorter than

in premotion Phase 1 (Phase 1 mean = 9.91 s, 95% CI [8.48, 11.33]; Phase 4 mean = 2.97 s, 95% CI [2.72, 3.23]; Phase 5 mean = 2.69 s, 95% CI [2.43, 2.96]). Furthermore, comparing between correct and incorrect responses in Phases 1, 4, and 5, viewing time was shorter in correct than in incorrect trials, $F(1, 3827.645) = 100.79$, $p < .001$. Viewing time and correctness were negative correlated ($r = -0.40$, -0.59 , and -0.77 , respectively; $p < .01$ in all cases), and longer viewing time was associated with poorer performance. Finally, color treatment type did not affect viewing time ($p > .05$).

Eye fixations

The total number of eye fixations and number of fixations in AOIs were both smaller in the correct trials in Phases 4 and 5 than in Phase 1 (see Fig. 2c). The proportion of fixation counts in AOIs over total fixation counts increased from Phase 1 to Phases 4 and maintained through Phase 5. A linear mixed-effect model, with event and participants as random effects, showed that there was a significant relationship between fixation count proportion and experimental phase, $F(2, 2048.27)$

$= 5.29$, $p < .01$, $\eta_p^2 = 0.005$, but not between count proportion and color treatment type ($p > .10$). Pairwise, the differences between Phases 1 and 4 and between Phases 1 and 5 were significant ($p < .05$).

Convergently, the total fixation time and fixation time in AOIs were both smaller in the correct trials in Phases 4 and 5 than in Phase 1 (see Fig. 2d). The proportion of fixation duration in AOIs over total fixation duration increased from Phase 1 to Phase 4 and maintained through Phase 5. The relationship between fixation duration proportion and experimental phase was significant, $F(2, 2048.20) = 4.91$, $p < 0.01$, $\eta_p^2 = 0.005$, but that between fixation duration proportion and color treatment type was not ($p > .12$). Pairwise, the differences between Phases 1 and 4 and between Phases 1 and 5 were significant ($p < .05$). In sum, participants fixated on the AOIs more in the postmotion phases, and this looking pattern was consistent with the identification performance.

Optic flow and eye movement

We extracted optic flow from each event (see Fig. 1) and computed the overlap proportion (fixation in high flow areas divided by total fixations) in the correct trials of Phases 3, 4, and 5. A linear mixed-effect model that treated subject and event as random effects showed that overlap proportion varied significantly between experimental phases, $F(2, 2221.26) = 4.11$, $p < .05$, $\eta_p^2 = 0.004$, but this was actually because overlap in Phase 4 was significantly, but only slightly (difference of means $= 0.05$; $p < .05$), higher than that in Phase 3. Overlap proportion was not affected by color treatment type ($p > 0.3$) or Color Treatment Type $\times$ Phase interaction ($p > .40$). Thus, when there was ongoing optic flow, or when motion stopped and optic flow was no longer available, participants still looked at areas where optic flow was or had been detected.

Discussion

In this study, we explored how optic flow and image structure information interact to enable event perception when images are blurry, and how color affects event perception when images are blurry. We analyzed (1) identification accuracy in phases before, during, and after optic flow was available in combination with blurry images; (2) eye fixations; and (3) optic flow strength in the stimuli. We found that perception of blurry events was affected by the experimental phase (which reflected the interaction of optic flow and blurry image structure) and by color treatment type, but not by their interaction. Specifically, our results showed that participants became better at identifying blurry events during and after motion (in Phases 3, 4, and 5). The optic-flow-improved performance was maintained after optic flow stopped in Phases 4

and 5, and the proportion of eye fixations on areas where strong optic flow had previously been detected in Phase 3 persisted. For a given event, identification accuracy, watch time, and eye fixations were all different between premotion Phase 1 and postmotion Phases 4 and 5, but they were not different between Phases 4 and 5. This means after receiving optic flow and allowing it to interact with image structure, the combination of information leads to more effective and highly stable perception of events. Observers were able to offload perceived events to blurry image structures and reidentify them in real time upon encountering the image information again, even after 5 days.

We argue that performance improvement from Phase 1 to Phases 4 and 5 was due to the detection of optic flow and not simply a practice effect. There are three reasons. First, there were 20 frames for each event and in each of Phases 1, 4, and 5, three frames were randomly selected and presented. So, participants did not view the same blurry image nine times. Second, in Phase 1, participants viewed three random frames from each event, and identification accuracy was not higher when they viewed the third frame than when they viewed the first frame (correct rate difference $= 0.01$), $t(35) = 0.74$, $p = .46$. However, when they viewed the fourth frame in Phase 4, there was a significant improvement from when they viewed the third frame (correct rate difference $= 0.61$), $t(35) = 15.96$, $p < .001$. Then, when participants viewed the fifth frame of the same event in Phase 4, there was no more improvement (correct rate difference $= 0.02$), $t(35) = 0.37$, $p = .71$. In other words, the improvement was stage-like, as a result of having seen the event with optic flow; the improvement was not smooth, as a result of practice. Third, in a previous study using a similar task, Pan and Bingham (2013) explicitly showed that only doing Phase 1 and Phase 4 did not improve event identification. Thus, higher event identification in Phases 4 and 5 was not due to repeated viewing of the blurry images, but instead to the detection of optic flow.

In this experiment, we did not restrict viewing time, because we wanted the participants to take as much time as needed to acquire ample information to make fair responses. The results that viewing time was shorter and identification correctness was higher in postmotion Phases 4 and 5 than in premotion Phase 1, and that longer viewing time was associated with incorrect responses and shorter viewing time was associated with correct responses could be interpreted as when there was event-specifying information, an observer picked it up quickly and made a correct response; when there was little or no information, one looked and ponder before making a response, which turned out to be incorrect.

Viewing time was not likely to confound identification performance, because, first, although in theory participants could watch the stimuli indefinitely, in practice they did not. The viewing time in Phase 2 was the longest, followed by that in Phase 1 and then in Phase 3. Viewing time in Phases 2 and 3

were long because the stimuli were long (25 s and 5 s, respectively) and participants must complete one loop before moving on to the next trial. A comparison between viewing time data and stimulus durations in these phases showed that participants did not watch the stimuli for more than two loops: Phase 2, $t(431) = -32.63, p < .001$; Phase 3, $t(431) = 0.97, p = .16$. Second, looking at the stimuli for longer did not improve identification performance because viewing time and correctness were negative correlated. Third, we used seemingly unrelated regression (Zellner, 1962) models to recompute how color and phase predicted viewing time and predicted identification accuracy, if viewing time and accuracy were related. Results showed that even considering that viewing time and accuracy had correlated error terms, these predictors were still significant ($p < .01$ in both cases).

Regarding the effect of color, in all experimental phases, performance was consistently better with unaltered natural colors (see Fig. 2a). Many previous studies have reported that color had a positive effect on various visual identification tasks (e.g., Greene & Oliva, 2009; Wurm, Legge, Isenberg, & Luebker, 1993; Yip & Sinha, 2002) and attributed this to either color diagnosticity or better image segmentation. In the current study, we created stimuli with rearranged, unnatural colors as a control for the color diagnosticity effect. We did not find a difference between the rearranged color and grayscale stimuli, while colored stimuli were better identified than the two. This implied that once the accustomed association between colors and objects was broken, as with the color rearrangement manipulation, identification became less effective. Current results support the color diagnosticity hypothesis, and natural colors helped event identification with blurry images.

Importantly, we found, with novelty, that color did not interact with optic flow in event identification. Optic flow improved performance when it was available and such improvement was equivalent across all color treatment types (see Fig. 2b). Further analyses suggested that, after optic flow was introduced, observers in all three color conditions fixated more on event-related AOIs and strong optic flow regions to the same extent. Thus, superior performance in the natural color condition was *not* because color aided object segmentation, in which case, natural colors would cause more fixation on those regions. Therefore, when perceiving blurry events, color and optic flow were both useful information, but their effects were independent and additive, which is what we had expected, given the relative independence of hue and motion measurement.

Nonetheless, the current study did not explicitly test what component of color, luminance or hue, aided blurry event identification. In the current experiment, the purpose of the rearranged color stimuli was to eliminate color diagnosticity. And it did. When color diagnosticity was removed, the color advantage disappeared. By randomly swapping the RGB

values, however, we could have changed luminance distribution in the images, and thus the local luminance contrast in the rearranged color displays could be different from (either lower or higher than) that in the color and grayscale displays. If the local luminance contrast in the rearranged color stimuli was tuned too low, then it was possible that color, in addition to being diagnostic, also aided object segregation, but observers could not detect and hence use the contrast from the rearranged color images because it was artifactually made too low. However, paired with the eye movement data, this was not likely, because in all eye movement analyses, there was no difference between looking patterns in the three-color treatment conditions. Contrast edges in the rearranged color images did not seem harder (or easier) to detect. Alternatively, if the local luminance contrast in the rearranged color stimuli was higher than in color and grayscale stimuli, this would not affect the current interpretation of results. We have begun new experiments to systematically investigate the individual contributions of luminance and hue to the color advantage in blurry event identification. Yet the current results validly support the primary message of this study that color, independent of optic flow, improved blurry event perception.

Optic flow speed covaries with surface depth (Simpson, 1993) and provides information about spatial relations and depth edges in a blurry event. After having been exposed to optic flow, participants knew where to look to locate depth edges or boundaries of objects and surfaces. This claim is supported first by the differences in watch time, number of fixations, and fixation durations without and with/after optic flow, which implied that after being exposed to optic flow in Phase 3, observers no longer scanned over the entire display but knew where to look to acquire event-specifying information. Second, this claim is consistent with the result that proportions of the number of fixations in AOIs and proportions of fixation durations in AOIs were both higher when viewing blurry images in the postmotion phases as compared to the premotion phase. Third, the overlap of high optic flow and eye fixation areas did not drop from Phase 3 to Phases 4 and 5. This means that participants looked at the optic flow regions after motion stopped (that is, in static images) as much as when motion was present. These eye movement patterns, along with the behavioral outcome in terms of accuracy and stability, all support that observers perceived depth relations among the otherwise ambiguous blurry contours and identified events using optic flow and continued to identify events from blurry images after motion stopped, because they were able to acquire, in real time, the spatial properties that were once specified by optic flow and preserved in the blurry contours.

In conclusion, color and motion provide useful information for event perception when events were otherwise unrecognizable due to extreme blurriness. As has been shown in previous studies of the effects of optic flow in perceiving and

identifying events under conditions often experienced in low vision, this study points out additional ways that low vision individuals might be able to recognize what is happening around them.

Acknowledgements This study is funded by the National Natural Science Foundation of China (General Programs 31970988), Guangdong Basic and Applied Basic Research Foundation (2020A1515010630) and Sun Yat-sen University (19wkzd22). The funders had no role in the study design, data collection, and preparation of this manuscript.

Open practices statements The data are available (<https://osf.io/ga5wc/>). Images and videos used in this study are too big for uploading, but they are available upon request to the corresponding author. None of the experiments was preregistered.

References

- Baguley, T. (2012). Calculating and graphing within-subject confidence intervals for ANOVA. *Behavior Research Methods*, 44(1), 158–175.
- Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Beardworth, T., & Bukner, T. (1981). The ability to recognize oneself from a video recording of one's movement without seeing one's body. *Bulletin of the Psychonomic Society*, 18, 19–22.
- Bochsler, T. M., Legge, G. E., Gage, R., & Kallie, C. S. (2013). Recognition of ramps and steps by people with low vision. *Investigative Ophthalmology & Visual Science*, 54(1), 288–294.
- Bochsler, T. M., Legge, G. E., Kallie, C. S., & Gage, R. (2012). Seeing steps and ramps with simulated low acuity: Impact of texture and locomotion. *Optometry and Vision Science*, 89(9), E1299–E1307. doi:<https://doi.org/10.1097/OPX.0b013e318264f2bd>
- Bilodeau, L., & Faubert, J. (1997). Isoluminance and chromatic motion perception throughout the visual field. *Vision Research*, 37(15), 2073–2081.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436.
- Bramão, I., Reis, A., Petersson, K. M., & Faisca, L. (2011). The role of color information on object recognition: A review and meta-analysis. *Acta Psychologica*, 138(1), 244–253.
- Cavanagh, P., & Anstis, S. (1991). The contribution of color to motion in normal and color-deficient observers. *Vision Research*, 31(12), 2109–2148.
- De Valois, K. K., & Switkes, E. (1983). Simultaneous masking interactions between chromatic and luminance gratings. *Journal of the Optical Society of America*, 73(1), 11–18.
- Fleiss, J. L., & Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement*, 33(3), 613–619.
- Gegenfurtner, K. R., & Rieger, J. (2000). Sensory and cognitive contributions of color to the recognition of natural scenes. *Current Biology*, 10(13), 805–808.
- Gegenfurtner, K. R., Wichmann, F. A., & Sharpe, L. T. (1998). The contribution of color to visual memory in X-chromosome-linked dichromats. *Vision Research*, 38(7), 1041–1045.
- Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological Science*, 20(4), 464–472.
- Jiang, X., Jiang, Y., & Parasuraman, R. (2015). The visual priming of motion-defined 3D objects. *PLOS ONE*, 10(12).
- Jobling, J. T., Mansfield, J. S., Legge, G. E., & Menge, M. R. (1997). Motion parallax: Effects of blur, contrast, and field size in normal and low vision. *Perception*, 26(12), 1529–1538.
- Jokisch, D., & Troje, N. F. (2003). Biological motion as a cue for the perception of size. *Journal of Vision*, 3(4), 1–1.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, 36(14), 1.
- Koenderink, J. J., & Van Doorn, A. J. (1991). Affine structure from motion. *Journal of the Optical Society of America: A-Optics Image Science and Vision*, 8(2), 377–385.
- Lee, Y. L., & Bingham, G. P. (2010). Large perspective changes yield perception of metric shape that allows accurate feedforward reaches-to-grasp and it persists after the optic flow has stopped! *Experimental Brain Research*, 204, 559–573.
- Lee, Y.-L., Lind, M., Bingham, N., & Bingham, G. P. (2012). Object recognition using metric shape. *Vision Research*, 69(C), 23–31.
- Li, L., & Warren, W. H., Jr. (2000). Perception of heading during rotation: Sufficiency of dense motion parallax and reference objects. *Vision Research*, 40(28), 3873–3894.
- Li, L., & Warren, W. H., Jr. (2002). Retinal flow is sufficient for steering during observer rotation. *Psychological Science*, 13(5), 485–490.
- Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, 7(11), 3416–3468.
- Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 210–220.
- Lu, Z. L., Lesmes, L. A., & Sperling, G. (1999). Perceptual motion standstill in rapidly moving chromatic displays. *Proceedings of the National Academy of Sciences of the United States of America*, 96(26), 15374–15379.
- Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, 4, 61–64.
- Nakayama, K., Shimojo, S., & Silverman, G. H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception*, 18(1), 55–68.
- Neri, P., Morrone, M. C., & Burr, D. C. (1998). Seeing biological motion. *Nature*, 395(6705), 894.
- Norman, J. F., Beers, A. M., Holmin, J., & Boswell, A. M. (2010). Effective 3-D shape discrimination survives retinal blur. *Attention, Perception, & Psychophysics*, 72(6), 1569–1575.
- Norman, J. F., & Todd, J. T. (1993). The perceptual analysis of structure from motion for rotating objects undergoing affine stretching transformations. *Perception & Psychophysics*, 53, 279–291.
- Norman, J. F., Todd, J. T., & Phillips, F. (1995). The perception of surface orientation from multiple sources of optical information. *Perception & Psychophysics*, 57(5), 629–636.
- Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology*, 41(2), 176–210.
- Palmer, S. E., & Brooks, J. L. (2008). Edge-region grouping in figure-ground organization and depth perception. *Journal of Experimental Psychology: Human Perception and Performance*, 34(6), 1353–1371.
- Pan, J. S., & Bingham, G. P. (2013). With an eye to low vision: Optic flow enables perception despite image blur. *Optometry and Vision Science: Official Publication of the American Academy of Optometry*, 90(10), 1119–1127.
- Pan, J. S., Bingham, N., & Bingham, G. P. (2013). Embodied memory: Effective and stable perception by combining optic flow and image structure. *Journal of Experimental Psychology: Human Perception and Performance*, 39(6), 1638–1651.
- Pan, J. S., Bingham, N., & Bingham, G. P. (2017a). Embodied memory allows accurate and stable perception of hidden objects despite

- orientation change. *Journal of Experimental Psychology: Human Perception and Performance*, 43(7), 1343–1358.
- Pan, J. S., Li, J., Chen, Z., Mangiaracina, E. A., Connell, C. S., Wu, H., ... Hassan, S. E. (2017b). Motion-generated optical information allows event perception despite blurry vision in AMD and amblyopic patients. *Journal of Vision*, 17(12), 13–16.
- Pavlova, M., & Sokolov, A. (2000). Orientation specificity in biological motion perception. *Perception & Psychophysics*, 62(5), 889–899.
- Pelli, D. G., & Vision, S. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Ramachandran, V. S., & Gregory, R. L. (1978). Does colour provide an input to human motion perception? *Nature*, 275(5675), 55–56.
- Riesenhuber, M., & Poggio, T. (2000). Models of object recognition. *Nature Neuroscience*, 3, 1199–1204.
- Ruppertsberg, A. I., Wuerger, S. M., & Bertamini, M. (2003). The chromatic input to global motion perception. *Visual Neuroscience*, 20(4), 421–428.
- Simpson, W. A. (1993). Optic flow and depth perception. *Spatial Vision*, 7, 35–75.
- Shevell, S. K., & Kingdom, F. A. (2008). Color in complex scenes. *Annual Review of Psychology*, 59, 143–166.
- Sun, D., Roth, S., & Black, M. J. (2010). Secrets of optical flow estimation and their principles. *IEEE International Conference on Computer Vision and Pattern Recognition*. doi:<https://doi.org/10.1109/CVPR.2010.5539939>
- Sun, D., Roth, S., & Black, M. J. (2014). A quantitative analysis of current practices in optical flow estimation and the principles behind them. *International Journal of Computer Vision*, 106(2), 115–137.
- Tarr, M. J., & Bulthoff, H. H. (1998). Image-based object recognition in man, monkey and machine. *Cognition*, 67, 1–20.
- Tanaka, J. W., & Presnell, L. M. (1999). Color diagnosticity in object recognition. *Perception & Psychophysics*, 61(6), 1140–1153.
- Teller, D. Y., & Lindsey, D. T. (1993). Motion at isoluminance: Motion dead zones in three-dimensional color space. *Journal of the Optical Society of America A*, 10(6), 1324–1331.
- Todd, J. T., Christensen, J. C., & Guckes, K. M. (2010). Are discrimination thresholds a valid measure of variance for judgments of slant from texture? *Journal of Vision*, 10(2), 20–20.
- Warren, W. H., & Hannon, D. J. (1990). Eye movements and optical flow. *Journal of the Optical Society of America A*, 7(1), 160–169.
- Wang, X. M., Lind, M., & Bingham, G. P. (2018). Large continuous perspective change with noncoplanar points enables accurate slant perception. *Journal of Experimental Psychology: Human Perception and Performance*, 44(10), 1508.
- Wuerger, S. M., & Landy, M. S. (1993). Role of chromatic and luminance contrast in inferring structure from motion. *Journal of the Optical Society of America A*, 10(6), 1363–1372.
- Wurm, L. H., Legge, G. E., Isenberg, L. M., & Luebker, A. (1993). Color improves object recognition in normal and low vision. *Journal of Experimental Psychology: Human Perception and Performance*, 19(4), 899.
- Yip, A. W., & Sinha, P. (2002). Contribution of color to face recognition. *Perception*, 31(8), 995–1003.
- Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association*, 57(298), 348–368.
- Zhao, H., & Warren, W. H. (2017). Intercepting a moving target: On-line or model-based control? *Journal of Vision*, 17(5), 12–13.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.