Retinotopy of visual masking and non-retinotopic perception during masking

Babak Noory · Michael H. Herzog · Haluk Ogmen

Published online: 14 March 2015
© The Psychonomic Society, Inc. 2015

Abstract Due to the movements of the observer and those of objects in the environment, retinotopic representations are highly unstable during ecological viewing conditions. The phenomenal stability of our perception suggests that retinotopic representations are transformed into non-retinotopic representations. It remains to show, however, which visual processes operate under retinotopic representations and which ones operate under non-retinotopic representations. Visual masking refers to the reduced visibility of one stimulus, called the target, due to the presence of a second stimulus, called the mask. Masking has been used extensively to study the dynamic aspects of visual perception. Previous studies using Saccadic Stimulus Presentation Paradigm (SSPP) suggested both retinotopic and non-retinotopic bases for visual masking. In order to understand how the visual system deals with retinotopic changes induced by moving targets, we investigated the retinotopy of visual masking and the fate of masked targets under conditions that do not involve eye movements. We have developed a series of experiments based on a radial Ternus-Pikler display. In this paradigm, the perceived Ternus-Pikler motion is used as a non-retinotopic reference frame to pit retinotopic against non-retinotopic visual masking hypothesis. Our results indicate that both metacontrast and structure masking are retinotopic. We also show that, under conditions that allow observers to read-out

effectively non-retinotopic feature attribution, the target becomes visible at a destination different from its retinotopic/spatiotopic location. We discuss the implications of our findings within the context of ecological vision and dynamic form perception.

Keywords Visual masking · Motion perception · Reference frames

Introduction

Retinotopic and non-retinotopic representations in human vision

The optics of the eye map neighboring points in the environment to neighboring photoreceptors in the retinae, and these neighborhood relations, known as retinotopic organization, are preserved in early visual cortical areas. Under normal viewing conditions, due to the movements of the observer's body, head, eyes, and due to the movements of objects, the stimuli impinging on retinotopic representations are highly dynamic and unstable. Thus, understanding ecological vision requires an understanding of how visual processes operate under these dynamic conditions. In order to explain the phenomenal stability of our environment, it is often postulated that the brain constructs non-retinotopic representations wherein the ego-centric representations (i.e., based on the observer, such as retinotopic representations) are transformed into exo-centric representations (i.e., based outside of the observer, such as spatiotopic representations). However, determining whether a given visual process operates in retinotopic or non-retinotopic representations and which visual processes operate in non-retinotopic representations remains one of the fundamental challenges in understanding ecological vision.

B. Noory · H. Ogmen (✉)
Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77204-4005, USA
e-mail: ogmen@uh.edu

B. Noory · H. Ogmen
Center for Neuro-Engineering and Cognitive Science, University of Houston, Houston, TX, USA

M. H. Herzog
Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Retinotopy of visual masking assessed with the Saccadic Stimulus Presentation Paradigm

Saccadic eye movements constitute a major source for retinotopic instability. However, during these eye movements, the world appears phenomenally stable suggesting that retinotopic shifts caused by saccades are either dismissed or compensated by the visual system. Theories of dismissal propose that very little information is kept from one saccade to another and vision starts *tabula rasa* after each saccade. Theories of complete compensation propose that all information is remapped across the saccade by taking into account the global shift caused by the saccade. Theories that take intermediate positions between these two extremes have also been proposed (Bridgeman, van der Heijden, & Velichkovsky, 1994). In general, compensation theories rely on three mechanisms: 1) before the initiation of the saccade, retinotopic information is stored in memory, 2) during and after the saccade, retinotopic information is suppressed to prevent inappropriate integration of pre- and post-saccadic images, 3) after the saccade, the new image is integrated with the contents of the memory by taking into account the retinotopic shift caused by the saccade. Because saccadic shifts take in general few tens of milliseconds, sensory (iconic) memory¹ and backward masking² have been viewed as the major candidates to perform the memorization and suppression tasks, respectively.

Several studies investigated whether sensory memory, masking, and information integration occur in retinotopic or non-retinotopic coordinates across saccadic eye movements.³ The Saccadic Stimulus Presentation Paradigm (SSPP) has been the classical experimental paradigm for these studies (Davidson, Fox, & Dick, 1973; Irwin, 1991; McRae, Butler, & Popiel, 1987; Melcher & Morrone, 2003). Figure 1 shows the SSPP paradigm used by Davidson et al. (1973) to investigate retinotopic versus non-retinotopic bases of backward masking.

The observer is asked to make a saccade from one fixation point to a second one. Target stimuli (five letters) are presented briefly before the saccade, followed by a mask stimulus (either a nonoverlapping ring, as in Fig. 1, or an overlapping pattern) presented after the saccade. As depicted in Fig. 1, the mask

Fig. 1 Saccadic Stimulus Presentation Paradigm (SSPP) used by Davidson et al. (1973). The observer makes a saccade from the first to the second fixation. Target stimuli, consisting of five letters are presented briefly just before the initiation of the saccade. A ring mask is presented after the saccade. The light gray target letters at the bottom show the relative position of the mask with respect to the target letters. In the actual stimulus, letters were only presented before the saccade. As one can see from the figure, the ring mask surrounds letter V according to spatiotopic coordinates and letter Y according to retinotopic coordinates. The non-overlapping ring mask corresponds to the metacontrast condition. The experiments also had an overlapping pattern to examine pattern masking by structure

stimulus surrounds (or covers) different letters according to retinotopic and non-retinotopic (spatiotopic) coordinates. By measuring which of the two letters is suppressed from perception, one can infer whether this mask operates in retinotopic or non-retinotopic coordinates. Davidson et al. (1973) reported that the mask suppressed the letter that shared its retinotopic coordinates but appeared to occupy the same position as the letter that shared its spatiotopic coordinates. They proposed that there is retinotopic visible persistence at which trans-saccadic masking occurs, and a spatiotopic sensory memory at which trans-saccadic integration occurs. In a follow-up study, McRae et al. (1987) reported not only retinotopic but also spatiotopic masking. They suggested that the transition from retinotopic to spatiotopic representations takes time and the reason Davidson et al. (1973) did not find evidence for spatiotopic masking could be the relatively shorter Inter-Stimulus Interval (ISI) used by Davidson et al. (ca. 80 ms) compared with the ISI used in McRae et al.'s study (153 ms).

¹ Sensory (iconic) memory is a visual storage mechanism with relatively high capacity and a relatively short time span (rev., Haber, 1983).

² Backward masking refers to the reduction in the visibility of a target stimulus caused by a mask stimulus that follows the target in time. When the mask surrounds but does not spatially overlap with the target, it is called metacontrast. When the mask spatially overlaps and shares structural similarities to the target, it is called backward structure masking (Bachmann, 1994; Breitmeyer & Ögmen, 2006).

³ White (1976) used smooth pursuit eye movements to study retinotopic versus non-retinotopic aspects of visual masking. He reported spatiotopic masking (White, 1976). However, a subsequent study where eye movements were monitored showed that masking during pursuit was retinotopic and not spatiotopic (Sun & Irwin, 1987).

That masking is retinotopic at short ISIs also was confirmed by Irwin, Brown, and Sun (1988). These authors also presented evidence for spatiotopic memory integrating information across saccades. However, their data suggested that spatiotopic integration of information were rather abstract depending on position and identity information rather than detailed image-like fusion of trans-saccadic stimuli (van der Heijden, Bridgeman, & Mewhort, 1986).

The observation of shifts of neuronal receptive fields in the direction of intended saccades (Duhamel, Colby, & Goldberg, 1992) generated a renewed interest for the problem of visual stability across saccades from this perspective (Melcher & Colby, 2008; Wurtz, 2008; Cavanagh, Hunt, Afraz, & Rolfs, 2010; Melcher, 2011). This “remapping of receptive fields” has been associated with shifts in the perceived positions of peri-saccadically presented targets (Ross, Morrone, Goldberg, & Burr, 2001). De Pisapia, Kaunitz, and Melcher (2010) suggested that these shifts, in turn, can help a target stimulus escape from masking. Moreover, they have presented evidence for spatiotopic masking for ISIs shorter (48 ms) than the ISIs reported in previous studies. Hunt and Cavanagh (2011) presented a brief target before the saccade followed by a long-duration mask that turned on before the saccade and remained on after the saccade until the subject responded. With this paradigm, Hunt and Cavanagh (2011) showed masking when the mask was presented at the *post-saccadic retinotopic coordinates* of the location where the target was presented. Taken together, these studies paint a complex picture for the retinotopy of masking. Part of the reason for this complexity may be due to the fact that many of the studies used different types of targets, masks and widely different parameters. It is known that masking is not a unitary phenomenon (Bachmann, 1994; Breitmeyer & Ögmen, 2006), and the differences between the studies may be due to differences in the types of masking functions and mechanisms evoked in different studies. Notwithstanding this issue, these studies show that SSPP provides a powerful method for exploring retinotopy of visual masking across saccades. However, SSPP involves eye-movement related processes, such as saccadic suppression and efference copy, and cannot be employed to study retinotopy of visual masking independent of eye movements.

Retinotopy of visual masking in the absence of eye movements

In a recent study, Lin and He (2012) investigated the retinotopy of masking in the absence of eye movements by using a modified version of object-specific reviewing paradigm (Kahneman, Treisman, & Gibbs, 1992). We use an alternative method that directly pits retinotopic predictions against non-retinotopic predictions. We will first introduce

our approach and compare in Section “Discussion” our methods and findings to those of Lin and He (2012).

The method that we have proposed for exploring non-retinotopic processing is based on the Ternus-Pikler paradigm (Ögmen, Otto, & Herzog, 2006; Boi, Ogmen, Krummenacher, Otto, & Herzog, 2009). The Ternus-Pikler display is an apparent motion stimulus, introduced by Gestalt psychologists about a century ago, and employed extensively since then to study the spatio-temporal aspects of human vision (Petersik & Rice, 2006; Pikler, 1917; Ternus, 1926). Figure 2 shows how the Ternus-Pikler stimulus has been adopted for studying non-retinotopic attribution of stimulus features (Ögmen et al., 2006). The basic Ternus-Pikler paradigm (Fig. 1a) consists of two display frames separated by a blank frame (ISI). The two display frames are identical, except all elements in Frame 2 are shifted by one inter element distance with respect to the elements in the first frame.

Depending upon the ISI, two different types of motion are perceived (Pantle & Picciano, 1976). For short ISIs (e.g., 0 ms) observers perceive Element Motion, in which the leftmost element in Frame 1 is perceived to be moving to the position of the rightmost element in Frame 2 (Fig. 2b). In this case, no motion is perceived for the other two elements. For long ISIs (e.g., 100 ms) observers perceive Group Motion, in which all elements are perceived to be moving together as a group (Fig. 2c). To study non-retinotopic feature attribution, a simple feature called a vernier offset is inserted into the central element of the first frame (Fig. 2d). A purely retinotopic hypothesis predicts that features of the central element in Frame 1 should be integrated into the leftmost element of Frame 2 for all ISIs within the window of temporal integration. Hence, performance should be well above chance level when subjects are asked to report the direction of the probe vernier perceived in the leftmost element in Frame 2, and near chance for the other elements in Frame 2. However, it was shown that performance depends on the ISI. When group motion is established between the Ternus-Pikler elements (ISI = 100), performance is well above chance when subjects are asked to report the perceived direction of the vernier offset in the central element in Frame 2, and near chance for other elements (Fig. 2d). On the other hand, in the case of element motion (ISI = 0), performance is higher when subjects report the vernier offset perceived in the leftmost element in Frame 2. The illusory attribution of the vernier offset also depends critically on the elicitation of a motion percept. If the leftmost line of the first frame and the rightmost line of the second frame are omitted (Fig. 2e), no apparent motion is induced since the remaining elements spatially overlap. In this control display, percentage of responses in agreement with the probe-vernier is high only for the element labeled 1 and at chance level for element 2 for both ISIs. These results indicate that feature attribution between elements of consecutive Ternus-Pikler display frames is governed according to motion-induced grouping; i.e., according to the dashed arrows in Fig. 2b and c. In other words,

Fig. 2 Ternus-Pikler paradigm for exploring non-retinotopic feature processing. **(a) Ternus-Pikler Display:** two display frames are separated by a blank interval called Inter Stimulus Interval (ISI). The two display frames are identical, except all elements in Frame 2 are shifted by one inter element distance with respect to the elements in the first frame. **(b) Element Motion:** For short ISIs (e.g., 0 ms) observers perceive the leftmost element in Frame 1 to be moving to the position of the rightmost element in Frame 2. In this case, no motion is perceived for the two central elements. **(c) Group Motion:** For long ISIs (e.g., 100 ms) observers

perceive all elements to be moving as a group. **(d) Results:** A Ternus-Pikler display presented with an ISI of either 0 or 100 ms. The central element in the first frame included a small offset called the “probe-vernier.” Observers attended to one of the elements of the second frame labeled as 1, 2, or 3. **e. Control Stimulus:** Only the elements that overlapped in the two Ternus-Pikler frames were shown, i.e., the leftmost element of the first and the rightmost element of the second frame of the stimulus shown in D were not displayed. No motion percept was elicited (adapted from Ögmen et al., 2006)

perceived motion of the Ternus-Pikler elements serves as a non-retinotopic reference frame for feature attribution between elements of the Ternus-Pikler display frames.

In the present study, we use a similar Ternus-Pikler paradigm to probe retinotopic and non-retinotopic bases of visual masking and to assess non-retinotopic perception during masking.

General methods and materials

All visual stimuli were generated via a Visual Stimulus Generator (VSG 2/5) card manufactured by Cambridge Research Systems. The stimuli were displayed on a 22-inch color monitor set at a resolution of 800×600 with a refresh rate of 100 Hz. Subject responses were collected by means of

a joystick connected to the computer hosting the VSG card. The distance between the observer and the monitor was fixed at 0.5 m, and a head/chin rest was utilized to minimize subject head motion during the experiments. Observers were asked to maintain a stable gaze at a fixation cross that remained visible throughout the experiment at the center of the monitor. Our previous experiments indicate that observers are able to keep a stable fixation while viewing the Ternus-Pikler displays (Boi et al., 2009). Nevertheless, to completely rule out the involvement of eye movements, we ran control experiments with eye movement monitoring and the results of these experiments are provided in Appendix. All experiments were conducted in a dimly lit room. Background and Ternus-Pikler disk/square luminances for all experiments were set at 25 cd/m^2 and 10 cd/m^2 respectively. Target and mask luminance levels were chosen at 30 or 40 cd/m^2 depending upon subject thresholds for optimum masking. Four participants with age range from 25 to 37 years, of whom three were naïve to the purpose of the study, took part in the experiments. The experiments were conducted according to a protocol approved by the University of Houston Committee for the Protection of Human Subjects. Informed consent was obtained from every participant, and practice trials were conducted to familiarize the observers with experimental procedures. The results of practice trials were not included in the data analysis.

Experiment 1: Retinotopy of metacontrast masking

Methods

In this experiment, we utilized a radial Ternus-Pikler display to study retinotopy of metacontrast masking. Two display frames, each of which contained two discs and a central square aligned on the perimeter of an invisible circle centered at the fixation, were displayed sequentially to create perception of radial Ternus-Pikler apparent motion (Fig. 3a and b). The radius of this virtual circle was fixed at 2.5 degrees in all experiments. The target-mask combination shown in Fig. 3c was displayed at variable Stimulus Onset Asynchronies (SOAs). The target was predictably presented at the center of the square in the first frame of the Ternus-Pikler sequence, and subjects were asked to attend and report the location of the missing corner on the target diamond (left/right). Depending upon the spatial location of the mask within Frame 2, retinotopic and non-retinotopic masking effects were distinguished. Figure 3a displays the case of *Retinotopic Mask Condition*, where the mask in Frame 2 was presented at the same retinotopic location as that of the target diamond in Frame 1. Note that in the absence of eye movements, retinotopic and spatiotopic masking conditions are equivalent. Figure 3b, on the other hand, depicts the case of *Non-Retinotopic*

Mask Condition. In this case, the mask was displayed in the central square of Frame 2. The two squares presented in Frames 1 and 2 of the Ternus-Pikler sequence correspond to one another only when group motion is established. In addition to the retinotopic and non-retinotopic mask experiment conditions, two control conditions were included in this experiment. In the *Static Control Condition*, masking functions were obtained for individual subject in the absence of Ternus-Pikler motion. Under this condition, the Ternus-Pikler elements remained visible throughout the experiment at the same spatial location as that of Frame 1 in Fig. 3a or b. In the *No-Mask Control Condition*, target was shown in the absence of the mask.

Spatial parameters of the target and mask are displayed in Fig. 3c. Variable “x” represents the size of the probe gap, which was varied in the range of 12° to 25° to meet each individual subject’s masking threshold requirements. The square and the disk in the Ternus-Pikler stimulus had, respectively, a side and a diameter equal to 1.5° . Figure 3d displays the timing diagram of a typical trial. The ISI was fixed (0 or 40 ms) for each experimental block, and the target always appeared just before the ISI. Target and mask stimuli were presented for 10 and 20 ms, respectively. Mask onset time was randomized from trial to trial to allow for different ISI-SOA combinations per block. The Ternus-Pikler disks were displayed at an eccentricity of 4° and for a duration of 120 ms. As shown Fig. 3d, the Ternus-Pikler ISI limits the shortest masking SOA that can be used. Therefore, eccentricity, background luminance, Ternus-Pikler element shapes, and target/mask/disk contrasts were chosen in such a way that Ternus-Pikler group motion was perceived by all observers at a relatively short ISI (40 ms), whereas strong masking effect was observed at the corresponding SOA (50 ms). In this study, we used only one contrast-polarity for targets and masks. The luminance values for the target, mask, disk, and background were 40, 40, 10, and 50 cd/m^2 , respectively. Thus, with respect to the disk within which they were presented, the target and mask had positive contrast polarity. Based on previous research, we would expect quantitatively different but qualitatively similar results for other contrast polarity combinations (Breitmeyer, Tapia, Kafalıgönül, & Ögmen, 2008). Ternus-Pikler radial motion (upward or downward) also was randomized from trial to trial. In a two-alternative forced-choice design, three naïve observers as well as one of the authors reported the perceived location of the missing corner of the target diamond (left/right).

Note that in the non-retinotopic mask condition (Fig. 3b), the target and mask always stimulate distinct retinal areas. The retinotopic hypothesis predicts no masking effect for this condition, regardless of stimulus timing and Ternus-Pikler grouping. The non-retinotopic hypothesis predicts that in such a case, masking effect follows stimulus timing and the perceived Ternus-Pikler motion. As

Fig. 3 Experiment 1: Two display frames, each of which contained two disks and a square, were displayed sequentially to create the perception of radial apparent motion. The blank ISI frame is not displayed in this figure for the sake of simplicity. Subjects were asked to report the location of the missing corner on the target diamond, shown at the center of the middle square in the first frame. **(a)** *Retinotopic Mask Condition*: The mask was displayed in Frame 2, at the same spatial location as that of the target diamond in Frame 1. **(b)** *Non-Retinotopic Mask Condition*: The mask

was displayed in the central square of Frame 2, which corresponded to the central square of Frame 1, only when disks were perceived to be in group motion. **(c)** Spatial Parameters of the Target and Mask: Variable “x” represents the size of the probe gap, which was varied to meet individual subject threshold requirements. **(d)** Timing Diagram: The ISI was fixed (0 or 40 ms) per block, and the target predictably appeared just before the ISI. The mask presentation time was randomized from trial to trial in order to allow for different ISI-SOA combinations per block

such, if the Ternus-Pikler elements are perceived to be in element motion, the non-retinotopic prediction is same as the retinotopic hypothesis. However, when Ternus-Pikler elements are perceived to be in group motion, the non-retinotopic hypothesis predicts masking effect for non-retinotopic mask condition instead. Figure 4 summarizes the respective predictions of retinotopic and non-retinotopic hypotheses, based on the perceived motion grouping of the Ternus-Pikler disks.

Results

Figure 5a shows performance as a function of the SOA for the static control condition. Repeated measures ANOVA indicates significant masking effect ($F_{1,3} = 77.06$; $p = 0.003$; $\eta_p^2 = 0.963$). Metacontrast masking function dips at $\text{SOA} = 40$ ms, indicating type-B masking function ($F_{5,15} = 5.71$; $p = 0.004$; $\eta_p^2 = 0.658$).

Figure 5b shows performance as a function of SOA for element motion condition ($\text{ISI} = 0$ ms). Two-way repeated

Fig. 4 Predictions of *Retinotopic* and *Non-Retinotopic* Hypotheses for Experiment 1: Panels (a) and (c) depict predictions of the retinotopic hypothesis. Masking effect is expected only in retinotopic mask experiment condition, regardless of the perceived Ternus-Pikler motion

(group or element). Panels (b) and (d) depict predictions of the non-retinotopic hypothesis. Masking effect for each experiment condition is expected to depend on perceptual grouping of Ternus-Pikler disks

measures ANOVA shows significant effect of the mask condition ($F_{2,6} = 29.4$; $p = 0.012$; $\eta_p^2 = 0.907$), as well as the SOA ($F_{8,24} = 45.89$; $p = 0.007$; $\eta_p^2 = 0.939$). However, when retinotopic mask condition is removed from the analysis, metacontrast masking effect ($F_{1,3} = 0.214$; $p = 0.675$), as well as the SOA effect become insignificant ($F_{8,24} = 36.06$; $p = 0.497$). Figure 5c shows performance as a function of the SOA, when Ternus-Pikler disks are perceived to be in group motion (ISI = 40 ms). Once again, significant effect of the mask ($F_{2,6} = 126.09$; $p < 0.001$; $\eta_p^2 = 0.977$) and the SOA ($F_{4,12} = 5.49$; $p = 0.009$; $\eta_p^2 = 0.647$) are observed. Removal of the retinotopic mask condition from the analysis, once again, renders both masking ($F_{1,3} = 1.000$; $p = 0.391$) and SOA ($F_{4,12} = 1.876$; $p = 0.179$) effects insignificant.⁴

⁴ It is interesting to note that retinotopic mask in the element motion condition exerts stronger masking at SOAs shorter than 40 ms when compared to the static control condition. In other words, the metacontrast masking function appears to shift from type-B to type-A. The expression of type-A versus type-B metacontrast functions has been attributed to differences in stimulus parameters, criterion contents, or individual differences (Breitmeyer & Ögmen, 2006; Duangudom, Francis, & Herzog, 2007; Francis & Cho, 2008; Albrecht, Klapotke, & Matler, 2010; Maksimov, Murd, & Bachmann, 2011). In our study, the same subjects were used in the two conditions and the inspection of individual subject data showed that all subjects' metacontrast functions consistently showed increased masking at short SOAs in the Ternus-Pikler condition compared to the static control condition. Ruling out individual differences, stimulus configuration becomes a more likely explanation for this change. In the static condition, target and mask stimuli are embedded inside the same static square. In the dynamic Ternus-Pikler configuration, the square changes into a disk. Additional experiments can determine if this change can account for the increased masking at short SOAs.

Discussion

These results show clearly that metacontrast masking in the absence of eye movements is retinotopic. Regardless of the perceived motion of the Ternus-Pikler disks, the retinotopic mask significantly masks the target at optimum SOAs, whereas the non-retinotopic mask has no significant effect on performance. To generalize this result across mask and masking function types, we used a *spatially overlapping* mask that shared structural similarity with the target in the next experiment. In this structure masking paradigm, we chose a strong mask to generate a Type-A (i.e., monotonic) backward masking function instead of the Type-B (i.e., non-monotonic) backward masking function obtained in the metacontrast experiment.

Experiment 2: Retinotopy of structure masking with type-a masking function

Methods

Experimental design and procedures of Experiment 2 were identical to those of Experiment 1, with the exception of the target and mask. The target consisted of a square outline missing one side. Three bars were aligned on the screen, as depicted in Fig. 6, to form the target. The missing bar was randomly placed at the top or bottom of the square. The mask consisted of a collection of random horizontal and vertical

Fig. 5 Metacontrast masking. Percentage of correct responses in detecting the missing corner of the target diamond (left/right), averaged across the four observers. **(a)** Static control condition. Performance is near chance at an SOA of 40 ms with a Type-B masking function. **(b)** Element-Motion (ISI = 0 ms). Performance is near chance at SOAs near 40 ms only

in the retinotopic mask condition. **(c)** Group-Motion (ISI = 40 ms). Masking is observed only for the retinotopic mask. Error bars correspond to ± 1 SEM. In the case of No Mask condition, ± 1 SEM are shown by gray horizontal lines

bars, i.e., shared the same structural components as the target, to generate masking by structure (Fig. 6).

Figure 6d displays the timing diagram of a typical trial in Experiment 2. As in Experiment 1, the ISI was fixed (0 or 40 ms) for each experimental block and the target always appeared just before the ISI. Mask presentation time was again randomized from trial to trial to allow for different ISI-SOA

combinations per block. Background luminance, Ternus-Pikler element shapes, and target/mask/disk contrasts were chosen as explained in Experiment 1. Ternus-Pikler radial motion (upward or downward) also was randomized from trial to trial. In a two-alternative forced-choice design, three naïve observers and one of the authors reported the perceived location of the missing side of the target square (up/down).

Fig. 6 Stimuli and Respective Parameters for Experiment 2: (a) Retinotopic and (b) Non-retinotopic masking conditions. (c) The target consisted of a square outline missing one side. Three bars were aligned on the screen to form the target. The missing bar was randomly placed at the

top or bottom of the square. The mask consisted of a collection of random horizontal and vertical bars. (d) Stimulus timing was identical to that of Experiment 1

Results

Figure 7a shows performance as a function of the SOA for the static control condition. As expected, and supported by the significant effect of the mask condition ($F_{1,3} = 64.58$; $p = 0.004$; $\eta_p^2 = 0.958$), as well as the SOA ($F_{5,15} = 20.24$; $p < 0.001$; $\eta_p^2 = 0.871$), a strong type-A masking function was found. Figure 7b shows performance as a function of the SOA for the element motion condition (ISI = 0 ms). Two-way repeated measures ANOVA shows significant effect of the mask condition ($F_{2,6} = 76.45$; $p < 0.001$; $\eta_p^2 = 0.962$), as well as the SOA ($F_{4,12} = 5.42$; $p = 0.010$; $\eta_p^2 = 0.644$). However, when

retinotopic mask condition is removed from the analysis, metacontrast masking effect ($F_{1,3} = 0.109$; $p = 0.763$), as well as the SOA effect become insignificant ($F_{4,12} = 1.06$; $p = 0.415$). Figure 7c shows performance when Ternus-Pikler disks are perceived to be in group motion (ISI = 40 ms). Once again, two-way repeated measures ANOVA shows significant effect of the mask condition ($F_{2,6} = 26.29$; $p < 0.001$; $\eta_p^2 = 0.898$), as well as the SOA ($F_{4,12} = 8.02$; $p = 0.002$; $\eta_p^2 = 0.728$). However, when retinotopic mask condition is removed from the analysis, metacontrast masking ($F_{1,3} = 0.134$; $p = 0.738$), as well as the SOA ($F_{4,12} = 2.151$; $p = 0.137$) effect become insignificant.

Fig. 7 Masking by structure. Percentage of correct responses in detecting the missing side of the target square (up/down), averaged across the four observers. **(a)** Static control condition. Performance is near chance at SOA of 0 ms with a Type-A masking function. **(b)** Element-Motion

(ISI = 0 ms). **(c)** Group-Motion (ISI = 40 ms). Error bars correspond to ± 1 SEM. In the case of No Mask condition, ± 1 SEM are shown by gray horizontal lines

Discussion

Together, the results of Experiments 1 and 2 show that backward masking is retinotopic and this finding holds for metacontrast and structure masking, as well as, for type-A and type-B masking functions.

In a recent study, Lin and He (2012) investigated the retinotopy of masking by using a modified version of object-specific reviewing paradigm (Kahneman et al., 1992). A rectangular object (frame) was presented for a preview period of 200 ms. The target was presented during the last 10 ms of this preview period in one of the two sides of the rectangle. This

rectangular frame was then shifted to a new location and displayed for another 200 ms. The mask stimuli were presented during the first 30 ms of the shifted frame. One side of the frame contained a weak mask and the other side contained a strong mask. Neither mask occupied the same retinotopic location as the target but one of the masks occupied the same rectangle-relative position as the target (i.e., the same side). Observers performed worse when the strong mask occupied the same relative position as the target. Lin and He interpreted this finding as evidence for non-retinotopic frame-centered backward masking. While this interpretation is plausible, it is difficult to make inferences about masking without observing the complete masking functions and comparing directly retinotopic, non-retinotopic, and baseline conditions. At the single short SOA of 10 ms (corresponding to ISI = 0 ms) used in the experiment, it is difficult to assess whether the difference in performance across the two mask types is due to masking *per se* or other factors. In our experiments, we included baseline no-mask measures, multiple SOA values to reveal the full typical type-A and type-B masking functions and compared directly retinotopic and non-retinotopic masking conditions according to two different motion grouping conditions. Our results reveal only retinotopic masking.

Previous studies showed that features of a masked target can be observed as being part of the mask stimulus (Werner, 1935; Wilson & Johnson, 1985; Herzog & Koch, 2001; Otto, Ogmen, & Herzog, 2006; Ögmen et al., 2006; Breitmeyer, Herzog, & Ogmen, 2008). As indicated in Fig. 1, the vernier offset of the target in the first frame can be observed on the mask stimulus shown in the second frame even though no vernier is presented at this element nor at this retinotopic location. Similarly, by using the sequential metacontrast paradigm, we have shown that features of a target, whose visibility is suppressed, can nevertheless be perceived along motion streams to which the target belongs (Otto et al., 2006; Herzog, Otto, & Ogmen, 2012). Our studies showed that the attribution of the target's features to the mask stimulus is a consequence of motion grouping rather than masking itself (Ögmen et al., 2006; Breitmeyer et al., 2008). The goal of the next experiment was to study this motion-dependent non-retinotopic feature attribution in masking.

Experiment 3: Non-retinotopic feature attribution

In some trials of Experiments 1 and 2, subjects informally reported perceiving the target to be moving with the Ternus-Pikler elements, as one would expect from non-retinotopic feature attribution. In such cases, the target could be perceived at spatial locations different from where the target stimulus was actually presented. To formally study this, we removed the motion ambiguity from Experiments 1 and 2, and instructed our subjects to spread their attention as discussed in the following section, so as to facilitate the read-out of non-retinotopic feature attribution.

Methods

Experimental design and procedures of Experiment 3 were identical to those of Experiments 1 and 2, with the exception of the Ternus-Pikler motion. The Ternus-Pikler motion was made predictably upwards in all trials, and the subjects were instructed to spread their attention to *the central Ternus-Pikler square in both display frames*. The target and mask design was identical to those of Experiments 1 and 2 for the respective metacontrast and structure masking conditions. Stimulus timing was also chosen to match those of the previous two experiments. Once again, the ISI was fixed (0 or 40 ms) for each experimental block, and the target always appeared just before the ISI. Mask presentation time, was again randomized from trial to trial to allow for different ISI-SOA combinations per block. Background luminance, Ternus-Pikler element shapes, and target/mask/disk contrasts were chosen to match those of the previous experiments. The Ternus-Pikler radial motion was fixed (upward) in all trials to remove motion ambiguity. In a two-alternative forced-choice design, three naïve observers as well as one of the authors reported the perceived missing corner of the target diamond or the location of the missing side of the target square for metacontrast and structure masking conditions, respectively.

Results A: metacontrast masking

Figure 8a shows performance as a function of the SOA, when Ternus-Pikler disks are perceived to be in element motion (ISI = 0 ms). Two-way repeated measures ANOVA shows significant effect of the mask condition ($F_{2,6} = 102.09$; $p < 0.001$; $\eta_p^2 = 0.971$), as well as the SOA ($F_{8,24} = 4.587$; $p = 0.002$; $\eta_p^2 = 0.605$) when all mask conditions are included in the analysis. When retinotopic mask condition is removed from the analysis, metacontrast masking effect ($F_{1,3} = 3.183$; $p = 0.172$), as well as the SOA effect become insignificant ($F_{8,24} = 0.972$; $p = 0.480$). These findings were in accordance with our findings reported in our previous experiments. However, when the disks are perceived to be in group motion (Fig. 8b), masking effect ($F_{2,6} = 1.509$; $p = 0.294$) as well as the SOA effect ($F_{4,12} = 0.738$; $p = 0.584$) become insignificant, even when both retinotopic and non-retinotopic mask conditions are included in the analysis.

Results B: masking by structure

Similar pattern of results was observed when a structure mask was utilized in the presence of predictable Ternus-Pikler disk motion. Figure 9a shows performance as a function of the SOA when the disks are perceived to be in element motion (ISI = 0 ms). Two-way repeated measures ANOVA shows significant effect of the mask condition ($F_{2,6} = 10.250$; $p = 0.012$; $\eta_p^2 = 0.774$), as well as the SOA ($F_{8,24} = 3.569$; $p = 0.007$; $\eta_p^2 = 0.543$) when all mask conditions are included in the analysis. When retinotopic mask condition is removed

Fig. 8 Metacontrast masking with predictable Ternus-Pikler motion. The observers attended to the central Ternus-Pikler square in both display frames. Percentage of correct responses in detecting the missing side of the target diamond (left/right), averaged across the four observers. **(a)**

Element-Motion (ISI = 0 ms). Performance is near chance at SOA of 10 ms only in the retinotopic mask condition. **(b)** Group-Motion (ISI = 40 ms). No masking is observed. Error bars correspond to ± 1 SEM. In the case of No Mask condition, ± 1 SEM are shown by gray horizontal lines

from the analysis, metacontrast masking effect ($F_{1,3} = 0.283$; $p = 0.631$), as well as the SOA effect ($F_{8,24} = 0.852$; $p = 0.568$) become insignificant. These findings were in accordance with our findings reported in our previous experiments. However, when the disks are perceived to be in group motion (Fig. 9b), masking effect ($F_{2,6} = 4.351$; $p = 0.068$) as well as the SOA effect ($F_{4,12} = 1.937$; $p = 0.169$) become insignificant, even when both retinotopic and non-retinotopic mask conditions are included in the analysis.

Discussion

In agreement with the results found in experiments 1 and 2, retinotopic masking is observed when the Ternus-Pikler disks are perceived in element motion (ISI = 0 ms). However, in contrast to the results found in our previous two experiments, when observers can focus their attention to the Ternus-Pikler element in the second frame which is grouped with the Ternus-Pikler element in the first frame containing the target, they can identify the target based on its continued appearance along the motion

path of the element containing the target. This finding is in agreement with our previous results from sequential metacontrast (Otto et al., 2006; Herzog et al., 2012) and Ternus-Pikler display (Fig. 1). Informal reports of our subjects state that a faded, but complete copy of the target is perceived at the non-retinotopic destination, in accordance with the motion of the stimulus.

General discussion

The functional significance of retinotopic masking in the absence of eye movements can be understood by considering how the visual system analyzes the form of moving targets. Under normal viewing conditions, a briefly presented stimulus remains visible for approximately 120 ms after the stimulus offset (Haber & Standing, 1970; Coltheart, 1980). Due to this visible persistence, one would expect moving objects to appear highly blurred with a comet-like trailing smear. Yet our normal perception of objects in motion is relatively clear and sharp (Ramachandran, Rao, & Vidyasagar, 1974; Bex, Edgar,

Fig. 9 Masking by structure with predictable Ternus-Pikler motion. The observers attended to the central Ternus-Pikler square in both display frames. Percentage of correct responses in detecting the missing side of the target square (up/down), averaged across the four observers. **(a)**

Element-Motion (ISI = 0 ms). Performance is near chance at SOA of 10 ms only in the retinotopic mask condition. **(b)** Group-Motion (ISI = 40 ms). No masking is observed. Error bars correspond to ± 1 SEM. In the case of No Mask condition, ± 1 SEM are shown by gray horizontal lines

& Smith, 1995; Westerink & Teunissen, 1995; Hammett, 1997), a phenomenon known as motion deblurring (Burr, 1980; Hogben & Di Lollo, 1985; Chen, Bedell, & Ogmen, 1995; Burr & Morgan, 1997). We have proposed a theory according to which dynamic form computation relies on a synergy between retinotopic masking and motion-based reference frames (Ögmen, 2007; Ögmen & Herzog, 2010). According to this theory, masking and motion mechanisms play *complementary* roles: Masking operates in retinotopic representations to control motion blur and motion mechanisms provide the reference frame used to compute non-retinotopically features of moving targets. This is in contrast to theories suggesting that motion deblurring, dynamic form perception, and masking all result from motion mechanisms (Burr, 1980; Burr, Ross & Morrone, 1986) and those suggesting that computation of features and masking are linked by the common process of object substitution or updating (Enns & Di Lollo, 1997; Di Lollo, Enns, & Rensink, 2000; Enns, 2002; Enns, Lleras, & Moore, 2010). Whereas some models of backward masking causally linked backward masking and motion mechanisms (Kahneman, 1967; Burr, 1984), various studies

showed that these two processes are largely independent and can be dissociated from each other (Weisstein & Grownay, 1969; Stoper & Banffy, 1977; Breitmeyer & Horman, 1981). Below we discuss independent but complementary roles that these processes play in the perception of dynamic form.

Mechanisms of motion deblurring: Metacontrast, and not motion, mechanisms

Several models have been proposed to explain motion deblurring (Anderson & Van Essen, 1987; Burr, 1980; Burr, Ross & Morrone, 1986; Martin & Marshall, 1993). According to Burr (1980), motion estimation is achieved by the spatio-temporally oriented receptive fields of motion mechanisms (such as the Reichardt motion detector or equivalent motion energy models). These motion-based models predict that an isolated moving target should not produce motion blur provided that it sufficiently stimulates the motion mechanisms. However, this prediction does not agree with findings from various studies that show the perception of extensive blur for moving isolated targets (Bidwell, 1899; Chen et al., 1995;

Fig. 10 Schematic representation of retinotopic and non-retinotopic interactions. The retinotopic space (retinotopic visual areas) is shown at the bottom of the figure. Three dots are moving in the rightward directions while four dots are moving in the upward direction. Masking takes place within these retinotopic representations according to spatio-temporal properties of the stimuli. Dots are grouped into two groups according to their motion vectors. A reference motion vector is extracted and serves as the reference frame whereby dots are mapped into non-retinotopic representations. Features of the dots are computed and attributed according to these reference frames in the non-retinotopic space. For example, the

rightmost dot of the group on the left has the same motion vector as the reference frame motion vector. Therefore, according to this reference frame, it is represented as a stationary dot in the non-retinotopic manifold. As this dot moves in the retinotopic space, its features are mapped to the same non-retinotopic locus yielding non-retinotopic integration of its features. Although the activity corresponding to this map can be suppressed at the retinotopic representations, observers can report the features of the dot by reading-out its non-retinotopic representation. For more details, see (Öğmen, 2007; Öğmen & Herzog, 2010)

Lubimov & Logvinenko, 1993; McDougall, 1904; Smith, 1969a, b). By using several paradigms directly tailored to test the predictions of motion-based models for deblurring, we showed that the activation of motion mechanisms is not a sufficient condition for motion deblurring. These findings argue against aforementioned motion-based models of deblurring. Our theoretical (Ogmen, 1993), experimental (Chen et al., 1995), and computational (Purushothaman, Ogmen, Chen, & Bedell, 1998) studies suggest that, metacontrast, and not motion, mechanisms underlie motion deblurring. In agreement with these findings, several studies showed strong correlation between motion smear and metacontrast in terms of their dependence on spatial separation, timing, and eccentricity (Castet, Lorenceau, & Bonnet, 1993; Chen et al., 1995; Di Lollo & Hogben, 1985; Farrell, 1984). Motion deblurring is closely related to “sequential metacontrast masking” (Herzog et al., 2012; Otto et al., 2006; Piéron, 1935), which is an extended form of metacontrast (Breitmeyer & Öğmen, 2006).

Mechanisms of non-retinotopic feature attribution: Motion, and not metacontrast, mechanisms

However, masking mechanisms solve only partly the motion blur problem. They can make motion streaks appear shorter thereby reducing the amount of blur in the picture. Yet, although deblurred, moving objects would still suffer from having a ghost-like appearance (Öğmen, 2007). This is because, in the retinotopic space, a moving object will stimulate each retinotopically localized receptive-field briefly and

incompletely processed form information would spread across the retinotopic space just like the ghost-like appearances of moving objects in pictures taken at relatively slow shutter speeds. As a solution to this “moving ghosts” problem, we suggested that features of moving objects are processed according to motion-based non-retinotopic reference frames (Öğmen et al., 2006; Otto et al., 2006; Öğmen, 2007; Öğmen & Herzog, 2010). As depicted in Fig. 10, stimuli are grouped in the retinotopic space according to their motion; a common motion vector is extracted to serve as a reference-frame; and this reference-frame is used to map features of the stimuli onto non-retinotopic representations (manifolds), a process that we have termed *non-retinotopic feature attribution*. In agreement with this proposal, recent studies have indicated that visual attributes of a stimulus such as form (Öğmen et al., 2006), luminance (Shimozaki, Eckstein, & Thomas, 1999), color (Nishida, Watanabe, Kuriki, & Tokimoto, 2007), size (Kawabe, 2008), and motion (Boi et al., 2009) are computed according to motion-based non-retinotopic reference frames.

In backward masking, it has been long known that features of the target can be perceived as being part of the mask, an effect termed feature transposition, feature inheritance, or feature attribution (Werner, 1935; Stewart & Purcell, 1970; Wilson & Johnson, 1985; Hofer, Walder, & Groner, 1989; Herzog & Koch, 2001; Enns, 2002; Sharikadze, Fahle, & Herzog, 2005; Öğmen et al., 2006; Otto et al., 2006). The close relationship between feature attribution and masking led some researchers to suggest that processing of features for moving stimuli and masking are linked causally by a

common process, viz., the process of object substitution or updating (Enns & Di Lollo, 1997; Di Lollo et al., 2000; Enns, 2002; Enns et al., 2010). To test whether feature attribution results from masking or motion mechanisms, we measured magnitudes of feature attribution, motion, and masking and computed correlations between these variables (Breitmeyer, Herzog, et al., 2008). Our results showed that, when apparent motion occurs *without* masking, it correlates positively with feature attribution. Furthermore, when apparent motion occurs *with* masking, feature attribution remains positively correlated with apparent motion after the contribution of masking is factored out, but does not correlate with masking after the contribution of apparent motion is similarly factored out. Taken together, these findings support the view that feature attribution is based on motion and not on masking mechanisms.

In our theory, masking operates in retinotopic representations while features of a stimulus can escape masking and become visible in non-retinotopic representations. Little is known in terms of neural correlates of visibility and non-retinotopic representations; however, we can propose an outline of how this dissociation between retinotopic masking and non-retinotopic visibility can take place following two general neural representation schemes.

In one, that we call *structural correlates*, different types of neurons represent different types of information. For example, a group of neurons may represent the *visibility of a stimulus in retinotopic coordinates*, whereas another group of neurons may represent the *visibility of features of a stimulus in motion-based coordinates (i.e., non-retinotopic representations)*. As shown in Fig. 11, architecture in which masking operates on retinotopic neurons and feature information flows upstream before masking can explain our results in Experiment 3. With predictable motion, observers can reliably readout activity in non-retinotopic representations and this

Fig. 11 Schematic representation of retinotopic masking and non-retinotopic visibility. Circles represent populations of neurons and dashed circles represent those that underlie visibility. Feature information is sent to both retinotopic and non-retinotopic representations, wherein they are represented according to retinotopic and motion-based reference frames, respectively. The locus of masking is at retinotopic representations

activity is immune to masking, which operates at the level of retinotopic representations.

In a second scheme, that we call *activational correlates*, the same population of neurons can carry out both types of information but through different activation patterns. For example, it has been suggested that synchronization between neurons at pre-determined frequency bands (e.g., alpha, gamma) may underlie conscious awareness of a stimulus. In this case, results of Experiment 3 can be explained by masking mechanisms that disrupt synchrony at retinotopic, but not at non-retinotopic, level. Of course, structural and activational correlates are not mutually exclusive and it is possible to find a mixture of the two. In general, our approach highlights the importance of motion-based reference-frames and suggests a broad and distributed neural representation that requires coordination between ventral and dorsal streams to process features in terms of motion-based reference frames.

All of these observations are in agreement with the findings of retinotopic masking under conditions where the observer is stationary with the eyes under steady fixation. However, under normal viewing conditions, both the subject and objects move. Future studies will determine how reference frames generated by ego-motions (as in the case of eye movements) and exo-motions (as in the case of moving objects, studied herein) are coordinated to work in synergy.

Acknowledgments Michael Herzog is supported by the *Swiss National Science Foundation (SNF)* project: “*Basics of visual processing: from retinotopic encoding to non-retinotopic representations.*” We thank the action editor and the reviewers for helpful comments and suggestions.

Appendix

Our previous experiments indicate that observers are able to keep a stable fixation while viewing the Ternus-Pikler displays (Boi et al., 2009). Nevertheless, to completely rule out the involvement of eye movements, we ran control experiments with eye movement monitoring.

Methods

The experimental procedures and stimulus parameters for the eye movement control experiments were identical to those of the original experiments, with the exception of eye monitoring procedures as explained below. Three observers (one new and two from the original participants; one author and two naive) participated in the replication of the first two experiments; since predictable motion is more conducive to eye movements, we have recruited seven observers (five new and two from the original participants; one author and six naive) for the replication of Experiment 3. Eye position was sampled at a rate of 250 Hz using an SR-Research Eyelink II eye tracker.

Saccades were detected with Eyelink's default saccade detection algorithm. A nine-point calibration was conducted at the beginning of each experiment block to map observer eye position to screen coordinates. Drift correct routine was conducted before every trial to account for minor observer or headgear movement during each block. Observer eye movements were analyzed online throughout all trials as well as offline. Trials during which a saccade was detected were rejected and

repeated as a new trial. Similarly, trials during which observer's gaze moved outside an imaginary circle of 1° diameter (centered around the fixation point) were rejected and repeated as a new trial. Furthermore, offline analysis was conducted on the eye movement data to rule out any correlation between eye movement and performance. Analysis of eye movement data indicates that subjects were able to maintain highly accurate fixation throughout the trials.

Fig. 12 Metacontrast masking with unpredictable Ternus-Pikler direction and with eye movement monitoring. Percentage of correct responses in detecting the missing corner of the target diamond (left/right), averaged across the three observers. (a) Static control condition. Performance is near chance at an SOA of 40 ms with a Type-B masking function. (b)

Element-Motion (ISI = 0 ms). Performance is near chance at SOAs near 40 ms only in the retinotopic mask condition. (c) Group-Motion (ISI = 40 ms). Masking is observed only for the retinotopic mask. Error bars correspond to ± 1 SEM. In the case of No Mask condition, ± 1 SEM are shown by gray horizontal lines

Results

Metacontrast masking with unpredictable Ternus-Pikler direction

The results are shown in Fig. 12. In comparing these results with those in Fig. 5, one can see that, while there are some quantitative differences, the general findings are the same.

As seen in Fig. 12b and c, two-way repeated measures ANOVA shows significant masking effect when all conditions are included in the analysis under element motion ($F_{2,6} = 6.42$; $p = 0.032$; $\eta_p^2 = 0.682$) as well as group motion ($F_{2,6} = 5.82$; $p = 0.039$; $\eta_p^2 = 0.668$) of the Ternus-Pikler disks. However, when retinotopic mask condition is removed from the analysis, masking effect becomes insignificant, whether disks are in element ($F_{1,3} = 6.00$; $p = 0.093$) or group motion ($F_{1,2} = 1.28$; $p = 0.374$).

Fig. 13 Masking by structure with unpredictable Ternus-Pikler direction and with eye movement monitoring. Percentage of correct responses in detecting the missing side of the target square (up/down), averaged across the three observers. (a) Static control condition. Performance is near

chance at SOA of 0 ms with a Type-A masking function. (b) Element-Motion (ISI = 0 ms). (c) Group-Motion (ISI = 40 ms). Error bars correspond to ± 1 SEM. In the case of No Mask condition, ± 1 SEM are shown by gray horizontal lines

Masking by structure with unpredictable Ternus-Pikler direction

Results are shown in Fig. 13 (compare with Fig. 7).

Similar results were found in the case of masking by structure (Fig. 13b and c). Two-way repeated measures ANOVA shows significant masking effect when all conditions are included in the analysis under element motion ($F_{2,6} = 37.90$; $p = 0.003$; $\eta_p^2 = 0.950$) as well as group motion ($F_{2,6} = 17.09$; $p = 0.011$; $\eta_p^2 = 0.895$) of the Ternus-Pikler disks. However, when retinotopic mask condition is removed from the analysis, masking effect becomes insignificant, whether disks are in element ($F_{1,2} = 1.84$; $p = 0.308$) or group motion ($F_{1,2} = 0.33$; $p = 0.622$).

Metacontrast masking with predictable Ternus-Pikler direction

Furthermore, under predictable Ternus-Pikler reference motion, the results of eye movement control experiments remain in agreement with the findings reported in Experiment 3.

Figure 14a (compare with Fig. 8) shows performance as a function of the SOA, when Ternus-Pikler disks are perceived to be in predictable element motion (ISI = 0 ms). When disks are perceived to be in element motion, two-way repeated measures ANOVA shows significant masking effect when all conditions are included in the analysis ($F_{2,4} = 1.77$; $p = 0.015$; $\eta_p^2 = 0.877$), but not when the retinotopic mask condition is removed from the analysis ($F_{1,2} = 1.77$; $p = 0.227$). On the other hand, when Ternus-Pikler disks are perceived to be in predictable group motion, masking effect is found to be insignificant, even when retinotopic mask condition is included in the analysis ($F_{2,4} = 0.51$; $p = 0.633$).

Masking by structure with predictable Ternus-Pikler direction

Results are shown in Fig. 15 (compare with Fig. 9).

Similar effects were observed in the masking by structure experiments. Under predictable Ternus-Pikler element motion, two-way repeated measures ANOVA shows significant structure masking effect when all conditions are included in the analysis ($F_{2,4} = 25.316$; $p = 0.005$; $\eta_p^2 = 0.927$), but not

Fig. 14 Metacontrast masking with predictable Ternus-Pikler direction and with eye monitoring. The observers attended to the central Ternus-Pikler square in both display frames. Percentage of correct responses in detecting the missing side of the target diamond (left/right), averaged across the three observers. (a) Element-Motion (ISI = 0 ms). Performance

is near chance at SOA of 10 ms only in the retinotopic mask condition. (b) Group-Motion (ISI = 40 ms). No masking is observed. Error bars correspond to ± 1 SEM. In the case of No Mask condition, ± 1 SEM are shown by gray horizontal lines

Fig. 15 Masking by structure with predictable Ternus-Pikler motion and with eye monitoring. The observers attended to the central Ternus-Pikler square in both display frames. Percentage of correct responses in detecting the missing side of the target square (up/down), averaged across the three observers. (a) Element-Motion (ISI = 0 ms). Performance is near

chance at SOA of 10 ms only in the retinotopic mask condition. (b) Group-Motion (ISI = 40 ms). No masking is observed. Error bars correspond to ± 1 SEM. In the case of No Mask condition, ± 1 SEM are shown by gray horizontal lines

when the retinotopic mask condition is removed from the analysis ($F_{1,2} = 5.684$; $p = 0.140$). However, when Ternus-Pikler disks are perceived to be in predictable group motion, masking effect is found to be insignificant, even when retinotopic mask condition is included in the analysis ($F_{2,4} = 4.927$; $p = 0.083$).

As an additional control experiment for Experiment 3, we recruited 4 new naïve subjects. In our control experiments described in this Appendix, we have used five samples for the SOA values in element and group motion conditions. Because the group motion condition is limited between 50-ms and 90-ms SOA values, these five samples provide 10-ms sampling interval. On the other hand, in the element motion case, the SOAs ranged from 10 ms to 90 ms with a sampling interval of 20 ms. With our new subjects, we used an alternate sampling for the element motion case so as to cover values not covered in the previous control experiment (20, 40, 60, 80, 100 ms instead of 10, 30, 50, 70, 90 ms). All other procedures, including eye movement monitoring, were identical to the previous control experiment.

Results for metacontrast masking are shown in Fig. 16. In the static control condition (top panel), we obtained the expected type-B masking function with performance near

chance at an SOA of 40 ms. Under predictable element motion of the Ternus-Pikler disks (ISI = 0; middle panel) significant masking is observed when all mask conditions are included in the two-way repeated measures ANOVA ($F_{2,6} = 17.05$; $p = 0.003$; $\eta_p^2 = 0.850$), but not when the retinotopic mask condition is removed from the analysis ($F_{1,3} = 0.018$; $p = 0.903$). On the other hand, when Ternus-Pikler disks are perceived to be in predictable group motion, masking effect is found to be insignificant, even when retinotopic mask condition is included in the analysis ($F_{2,6} = 0.633$; $p = 0.563$).

Results for masking by structure are shown in Fig. 17. In the static control condition (top panel), we obtained the expected Type-A masking function. Under predictable element motion of the Ternus-Pikler disks, (ISI = 0; middle panel) significant masking is observed when all mask conditions are included in the two-way repeated measures ANOVA ($F_{2,6} = 59.72$; $p < 0.001$; $\eta_p^2 = 0.952$), but not when the retinotopic mask condition is removed from the analysis ($F_{1,3} = 0.138$; $p = 0.735$). On the other hand, when Ternus-Pikler disks are perceived to be in predictable group motion, masking effect is found to be insignificant, even when retinotopic mask condition is included in the analysis ($F_{2,6} = 1.93$; $p = 0.224$).

Fig. 16 Metacontrast masking with predictable Ternus-Pikler direction and with eye monitoring. Four new observers were recruited. All methods and procedures were the same as in the previous control experiment with the exception of sampling of the SOA values in the element motion case. The top, middle, and bottom panels show the results the static, element motion, and group motion conditions, respectively. Error bars correspond to ± 1 SEM. In the case of No Mask condition, ± 1 SEM are shown by gray horizontal lines

Fig. 17 Masking by structure with predictable Ternus-Pikler direction and with eye monitoring (same observers as in Fig. 16). All methods and procedures were the same as in the previous control experiment with the exception of sampling of the SOA values in the element motion case. The top, middle, and bottom panels show the results the static, element motion, and group motion conditions, respectively. Error bars correspond to ± 1 SEM. In the case of No Mask condition, ± 1 SEM are shown by gray horizontal lines

These results collectively support our earlier findings and rule out eye movement as an explanation for elimination of masking effect under predictable group motion condition. As

mentioned, there have been reports of individual differences in masking functions; for example, for the same stimulus parameters, some subjects exhibit Type-A, whereas others have

Type-B masking functions (Albrecht et al., 2010; Maksimov et al., 2011). Whereas much of this difference can be explained by different criterion contents used by subjects, there remains a possibility of inherent differences between subjects. In our main experiments, we had four subjects (1 author, 3 naïve) and the control experiments had in total seven subjects (2 from the main experiment and 5 new naïve). There is still a possibility that some individual subjects may exhibit Type-A instead of Type-B masking function and vice-versa. Because the retinotopy of masking that we report holds for both Type-A and Type-B functions, we suggest that the individual differences in masking types should not affect our main findings.

References

- Albrecht, T., Klapotke, S., & Matler, U. (2010). Individual differences in metacontrast masking are enhanced by perceptual learning. *Consciousness and Cognition*, 19, 656–666.
- Anderson, C. H., & Van Essen, D. C. (1987). Shifter circuits: a computational strategy for dynamic aspects of visual processing. *Proceedings of the National Academy of Sciences of the United States of America*, 84(17), 6297–6301.
- Bachmann, T. (1994). *Psychophysiology of visual masking: The fine structure of conscious experience*. Commack, NY: Nova Science Publishers.
- Bex, P. J., Edgar, G. K., & Smith, A. T. (1995). Sharpening of blurred drifting images. *Vision Research*, 35, 2539–2546.
- Bidwell, S. (1899). *Curiosities of light and sight*. London: Swan Sonnenschein.
- Boi, M., Ogmen, H., Krummenacher, J., Otto, T. U., & Herzog, M. H. (2009). A (fascinating) litmus test for human retino- vs. non-retinotopic processing. *Journal of Vision*, 9(13), 11.
- Breitmeyer, B. G., & Horman, K. (1981). On the role of stroboscopic motion in metacontrast. *Bulletin of the Psychonomic Society*, 17(1), 29–32.
- Breitmeyer, B. G., Herzog, M. H., & Ogmen, H. (2008). Motion, not masking, provides the medium for feature attribution. *Psychological Science*, 19, 823–829.
- Breitmeyer, B. G., & Ögmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision* (2nd ed.). Oxford: Oxford University Press.
- Breitmeyer, B. G., Tapia, E., Kafalgönül, H., & Ögmen, H. (2008). Metacontrast masking and stimulus contrast polarity. *Vision Research*, 48, 2433–2438.
- Bridgeman, B., van der Heijden, A. H. C., & Velichkovsky, B. M. (1994). A theory of visual stability across saccadic eye movements. *Behavioral and Brain Sciences*, 17, 247–292.
- Burr, D. (1980). Motion smear. *Nature*, 284, 164–165.
- Burr, D. C. (1984). Summation of target and mask metacontrast stimuli. *Perception*, 13, 183–192.
- Burr, D. C., & Morgan, M. J. (1997). Motion deblurring in human vision. *Proceedings of the Royal Society of London B*, 264, 431–436.
- Burr, D. C., Ross, J., & Morrone, M. C. (1986). Seeing objects in motion. *Proceedings of the Royal Society Series B-Biological Sciences*, 227(1247), 249–265.
- Castet, E., Lorenceau, J., & Bonnet, C. (1993). Inverse intensity effect is not lost with stimuli in apparent motion. *Vision Research*, 33, 1697–1708.
- Cavanagh, P., Hunt, A. R., Afraz, A., & Rolfs, M. (2010). Visual stability based on remapping of attention pointers. *Trends in Cognitive Sciences*, 14(4), 147–153.
- Chen, S., Bedell, H. E., & Ogmen, H. (1995). A target in real motion appears blurred in the absence of other proximal moving targets. *Vision Research*, 35, 2315–2328.
- Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics*, 27, 183–228.
- Davidson, M. L., Fox, M. J., & Dick, A. O. (1973). Effect of eye-movements on backward masking and perceived location. *Perception & Psychophysics*, 14(1), 110–116.
- De Pisapia, N., Kaunitz, L., & Melcher, D. (2010). Backward masking and unmasking across saccadic eye movements. *Current Biology*, 20, 613–617.
- Di Lollo, V., & Hogben, J. H. (1985). Suppression of visible persistence. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 304–316.
- Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General*, 129, 481–507.
- Duangudom, V., Francis, G., & Herzog, M. H. (2007). What is the strength of a mask in visual metacontrast masking? *Journal of Vision*, 7, 1–10.
- Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, 255(5040), 90–92.
- Enns, J. T. (2002). Visual binding in the standing wave illusion. *Psychonomic Bulletin & Review*, 9, 489–496.
- Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science*, 8, 135–139.
- Enns, J. T., Lleras, A., & Moore, C. M. (2010). Object updating: A force for perceptual continuity and scene stability in human vision. In R. Nijhawan & B. Khurana (Eds.), *Space and time in perception and action* (pp. 503–520). Cambridge: Cambridge University Press.
- Farrell, J. E. (1984). Visible persistence of moving objects. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 502–511.
- Francis, G., & Cho, Y. (2008). Effects of temporal integration on the shape of visual backward masking functions. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 1116–1128.
- Haber, R. N. (1983). The impending demise of the icon: A critique of the concept of iconic storage in visual information processing. *Behavioral and Brain Sciences*, 6, 1–54.
- Haber, R. N., & Standing, L. (1970). Direct estimates of the apparent duration of a flash. *Canadian Journal of Psychology*, 24, 216–229.
- Hammett, S. T. (1997). Motion blur and motion sharpening in the human visual system. *Vision Research*, 37, 2505–2510.
- Herzog, M. H., & Koch, C. (2001). Seeing properties of an invisible object: Feature inheritance and shine through. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 4271–4275.
- Herzog, M. H., Otto, T. U., & Ogmen, H. (2012). The fate of visible features of invisible elements. *Frontiers in Psychology*, 3, 119.
- Hofer, D., Walder, F., & Groner, M. (1989). Metakontrast: Ein beru“hmtes, aber schwer messbares Pha“nomen. *Schweizer Zeitschrift für Psychologie*, 48, 219–232.
- Hogben, J. H., & Di Lollo, V. (1985). Suppression of visible persistence in apparent motion. *Perception & Psychophysics*, 38(5), 450–460.
- Hunt, A. R., & Cavanagh, P. (2011). Remapped visual masking. *Journal of Vision*, 11(1), 13.
- Irwin, D. E. (1991). Information integration across saccadic eye-movements. *Cognitive Psychology*, 23(3), 420–456.

- Irwin, D. E., Brown, J. S., & Sun, J.-S. (1988). Visual masking and visual integration across saccadic eye movements. *Journal of Experimental Psychology: General*, 117(3), 276–287.
- Kahneman, D. (1967). An onset-onset law for one case of apparent motion and metacontrast. *Perception & Psychophysics*, 2, 577–584.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24(2), 175–219.
- Kawabe, T. (2008). Spatiotemporal feature attribution for the perception of visual size. *Journal of Vision*, 8(8), 9.
- Lin, Z., & He, S. (2012). Automatic frame-centered object representation and integration revealed by iconic memory, visual priming, and backward masking. *Journal of Vision*, 12(11):24, 1–18.
- Lubimov, V. & Logvinenko, A. (1993). Motion blur revisited. *Perception* 22(Suppl.), 77.
- Maksimov, M., Murd, C., & Bachmann, T. (2011). Target-mask congruence impacts the type of metacontrast masking. *Scandinavian Journal of Psychology*, 52, 524–529.
- Martin, K. E., & Marshall, J. A. (1993). Unsmearing visual motion: Development of long-range horizontal intrinsic connections. In Hanson, S., Cowan, J., & Giles, C. (Eds.), *Advances in neural information processing systems* (Vol. 5). San Mateo, Calif.: Morgan Kaufmann.
- McDougall, W. (1904). The sensations excited by a single momentary stimulation of the eye. *British Journal of Psychology*, 1, 78–113.
- McRae, K., Butler, B. E., & Popiel, S. J. (1987). Spatiotopic and retinotopic components of iconic memory. *Psychological Research Psychologische Forschung*, 49(4), 221–227.
- Melcher, D. (2011). Visual stability. *Philosophical Transactions of the Royal Society B*, 366, 468–475.
- Melcher, D., & Colby, C. L. (2008). Trans-saccadic perception. *Trends in Cognitive Sciences*, 12(12), 466–473.
- Melcher, D., & Morrone, M. C. (2003). Spatiotopic temporal integration of visual motion across saccadic eye movements. *Nature Neuroscience*, 6(8), 877–881.
- Nishida, S., Watanabe, J., Kuriki, I., & Tokimoto, T. (2007). Human visual system integrates color signals along a motion trajectory. *Current Biology*, 17(4), 366–372.
- Ogmen, H. (1993). A neural theory of retino-cortical dynamics. *Neural Networks*, 6, 245–273.
- Ögmen, H. (2007). A theory of moving form perception: Synergy between masking, perceptual grouping, and motion computation in retinotopic and non-retinotopic representations. *Advances in Cognitive Psychology*, 3, 67–84.
- Ögmen, H., & Herzog, M. H. (2010). The geometry of visual perception and the representation of information in the human visual system. *Proceedings of the IEEE*, 98, 479–492.
- Ögmen, H., Otto, T. U., & Herzog, M. H. (2006). Perceptual grouping induces non-retinotopic feature attribution in human vision. *Vision Research*, 46(19), 3234–3242.
- Otto, T. U., Ogmen, H., & Herzog, M. H. (2006). The flight path of the phoenix: The visible trace of invisible elements in human vision. *Journal of Vision*, 6, 1079–1086.
- Pantle, A., & Picciano, L. (1976). Multistable movement display - evidence for 2 separate motion systems in human-vision. *Science*, 193(4252), 500–502.
- Petersik, J. T., & Rice, C. M. (2006). The evolution of explanations of a perceptual phenomenon: A case history using the ternus effect. *Perception*, 35(6), 807–821.
- Piéron, H. (1935). Le processus du métacontrast. *Journal de Psychologie Normale et Pathologique*, 32, 1–24.
- Pikler, J. (1917). Sinnesphysiologische Untersuchungen.
- Purushothaman, G., Ogmen, H., Chen, S., & Bedell, H. E. (1998). Motion deblurring in a neural network model of retino-cortical dynamics. *Vision Research*, 38, 1827–1842.
- Ramachandran, V. S., Rao, V. M., & Vidyasagar, T. R. (1974). Sharpness constancy during movement perception. *Perception*, 3, 97–98.
- Ross, J., Morrone, M. C., Goldberg, M. E., & Burr, D. C. (2001). Changes in visual perception at the time of saccades. *Trends in Neurosciences*, 24(2), 113–121.
- Sharikadze, M., Fahle, M., & Herzog, M. H. (2005). Attention and feature integration in the feature inheritance effect. *Vision Research*, 45, 2608–2619.
- Shimozaki, S. S., Eckstein, M., & Thomas, J. P. (1999). The maintenance of apparent luminance of an object. *Journal of Experimental Psychology: Human Perception and Performance*, 25(5), 1433–1453.
- Smith, V. C. (1969a). Scotopic and photopic functions for visual band movement. *Vision Research*, 9, 293–304.
- Smith, V. C. (1969b). Temporal and spatial interactions involved in the band movement phenomenon. *Vision Research*, 9, 665–676.
- Stewart, A. L., & Purcell, D. G. (1970). U-shaped masking functions in visual backward masking: Effects of target configuration and retinal position. *Perception & Psychophysics*, 7, 253–256.
- Stoper, A. E., & Banffy, S. (1977). Relation of split apparent motion to metacontrast. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 258–277.
- Sun, J.-S., & Irwin, D. E. (1987). Retinal masking during pursuit eye movements: Implications for spatiotopic visual persistence. *Journal of Experimental Psychology: Human Perception and Performance*, 13(1), 140–145.
- Ternus, J. (1926). Experimentelle Untersuchungen über Phänomene der Identität. *Psychologische Forschung*, 7, 81–136.
- van der Heijden, A. H. C., Bridgeman, B., & Mewhort, D. J. K. (1986). Is stimulus persistence affected by eye movements? A critique of Davidson, Fox, and Dick (1973). *Psychological Research*, 48, 179–181.
- Weisstein, N., & Growney, R. (1969). Apparent movement and metacontrast: A note on Kahneman's formulation. *Perception & Psychophysics*, 5, 321–328.
- Werner, H. (1935). Studies of contour. I: Qualitative analysis. *American Journal of Psychology*, 47, 40–64.
- Westerink, J. H. D. M., & Teunissen, K. (1995). Perceived sharpness in complex moving images. *Displays*, 16, 89–97.
- White, C. W. (1976). Visual masking during pursuit eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 2(4), 469–478.
- Wilson, A. E., & Johnson, R. M. (1985). Transposition in backward masking: The case of the traveling gap. *Vision Research*, 25, 283–288.
- Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision Research*, 48, 2070–2089.