



OPEN

Ensemble perception without phenomenal awareness of elements

Taisei Sekimoto & Isamu Motoyoshi

Humans efficiently recognize complex scenes by grouping multiple features and objects into ensembles. It has been suggested that ensemble processing does not require, or even impairs, conscious discrimination of individual element properties. The present study examined whether ensemble perception requires phenomenal awareness of elements. We asked observers to judge the mean orientation of a line-based texture pattern whose central region was made invisible by backward masks. Masks were composed of either a Mondrian pattern (Exp. 1) or of an annular contour (Exp. 2) which, unlike the Mondrian, did not overlap spatially with elements in the central region. In the Mondrian-mask experiment, perceived mean orientation was determined only by visible elements outside the central region. However, in the annular-mask experiment, perceived mean orientation matched the mean orientation of all elements, including invisible elements within the central region. Results suggest that the visual system can compute spatial ensembles even without phenomenal awareness of stimuli.

The human visual system can easily and rapidly recognize objects, scenes, and materials from complex natural images^{1–3}. In order to achieve such efficient information processing, the visual system is thought to represent the statistical—or “collective”—aspects of spatially distributed information as ensembles^{4–6}. Indeed, increasing psychophysical evidence shows that humans can perceive not only ensembles of low-level image features such as orientation and size^{4,7} but also ensembles of high-level objects such as faces, biological motion, etc.^{8–11}. Moreover, another line of evidence points out that ensemble coding can explain visual phenomena such as search and crowding^{12–16}. In the cognitive-processing literature, ensemble representations have been proposed as helpful compact codes for handling large numbers of features and objects in working-memory systems with highly limited capacities¹⁷.

A defining characteristic of ensemble perception is that it does not require conscious access to individual elements⁶. A number of studies suggest that human observers can accurately discriminate averages and other statistics across stimuli without conscious access to individual element properties (e.g. orientation)^{8,9,18,19}. On the other hand, if a stimulus is perceived as an ensemble, it becomes difficult to discriminate the properties of its constituent elements at any given location, especially in the peripheral visual field^{20–23}. This ‘crowding’ phenomenon has been investigated extensively, and it is generally believed to result from the impairment of conscious access to the spatial location or properties of the target elements^{21–23}. In this respect, ensemble perception is inextricably linked to crowding²⁴.

In a nutshell, ensemble perception does not require ‘access’ awareness of stimulus elements. But it is unclear whether ensemble perception requires ‘phenomenal’ awareness of stimulus elements²⁵. Many psychophysical studies have introduced visual displays such as masking, adaptation, and binocular rivalry^{26–30}, in which human observers are not only unable to access stimulus properties but are also unaware of the presence of the stimulus itself (i.e., the stimulus is phenomenally ‘invisible’), despite implicit underlying neural processing of the stimulus^{31–35}. In such displays, do human observers can correctly perceive an ensemble (e.g., spatial average) if a portion of the stimulus’ elements are invisible?

Given vision’s functional requirement that vast amounts of information be processed instantaneously, it is plausible to assume that ensemble perception does not require phenomenal awareness of elements. Alternatively, it has been shown that phenomenal awareness of elements is necessary for crowding, a percept closely related to ensemble perception. Wallis and Bex (2011), for instance, showed that crowding does not occur if a proportion of elements are made invisible by adaptation³⁶ (but see also³⁷). This result would seem to imply that phenomenal awareness of elements is in fact necessary for ensemble perception.

Department of Life Sciences, The University of Tokyo, Tokyo, Japan. email: imotoyosi.ac@gmail.com

To elucidate the role of phenomenal awareness in ensemble perception, we examined whether perceived mean orientation of line-segment textures is affected by elements made invisible via dichoptic backward masking. In Experiment 1, we briefly presented a texture pattern whose central region was tilted relative to the surrounding region while the central region was backward-masked with a Mondrian pattern. Observers were asked to estimate mean texture orientation as a whole. With the central texture region almost totally invisible, observers reported mean orientations close to those of visible elements in the surrounding region. In Experiment 2, we used an annular contour mask that was designed to make central elements invisible without spatially overlapping with them. We found that the central region was perceived as a uniform hole with no visible elements, but the estimated mean orientation was consistent with the mean orientation of all elements including the invisible elements in the central region. These results suggest that, at least in the perception of average orientation, the visual system can compute spatial ensembles without phenomenal awareness of elements.

Experiment 1

Methods. *Observers.* We initially recruited 9 participants whose results showed that the expected effect—an interaction between the mask condition and the center orientation—was statistically significant. Based on the effect size in this result, we conducted an a priori power analysis using G*Power 3.1 which revealed that a sample size of 10 should be sufficient to achieve 0.90 power³⁸. We then excluded 1 author and recruited 2 more participants, and finally had 10 naïve students participated in the experiment (3 females, 7 males, 21.8 years old on average). All observers had normal or corrected vision. All experiments were conducted in accordance with the Declaration of Helsinki and with the permission of the Ethical Review Committee for Experimental Research on Human Subjects, Graduate School of Arts and Sciences, the University of Tokyo. All observers provided filled informed consent forms.

Apparatus. Visual stimuli were displayed on an LCD monitor (BenQ XL2735-B) with a frame rate of 60 Hz and a pixel resolution of 0.022 deg/pixel. The luminance of the monitor was gamma-corrected based on careful measurements with a colorimeter (ColorCal II CRS). In line with previous studies^{39,40}, all stimuli were presented dichoptically through a mirror stereoscope on two backgrounds (16(W) * 16(H) deg) displayed on the left and right sides of the monitor. The dichoptic presentation was effective to prevent low-level monocular (eg, retinal) interference²⁸. To facilitate binocular fusion, both backgrounds were surrounded by thin frames made of random dots.

Stimuli. Target stimuli were texture patterns consisting of 64 line segments scattered in a circular area of 3.7 deg in diameter (Fig. 1). Each line segment had a width of 0.1 deg, a length of 0.4 deg, and a luminance of 30 cd/m². The elements were randomly arranged with the minimum distance of 0.3 deg. Element orientation was determined according to a Gaussian distribution with a particular mean and a s.d. of 4 deg. Elements within the circular central region of a 2.0 deg diameter were varied by -30, 0, or 30 deg with respect to elements in the surrounding region. The mean orientation of the entire texture was determined for each trial in accordance with a staircase procedure. The mask was a circular Mondrian pattern (2.0 deg diameter) consisting of overlaid discs of random sizes and luminances. In order to check whether the observer properly reported the visibility of the central region, a texture with no central region (i.e. holed) was presented in 1 out of 5 trials.

Procedure. Perceived mean texture orientation was measured with a two-alternative forced-choice (2AFC) method. In each trial, the target texture was presented to either eye for 17 ms, and the mask was presented to the other eye for 17 ms following a 50 ms SOA. The target and mask were presented 4.9 deg to the right of a black fixation point (0.2 × 0.2 deg) shown continuously in the center of the background. No mask was presented in half the trials. While viewing the display with steady fixation, observers indicated by the press a button whether the mean orientation of the whole texture was tilted clockwise or counter-clockwise. Observers then pressed a button to indicate whether or not they perceived texture elements in the central region. Observers were instructed to answer “perceived” if they could perceive any line segments in the central region. The next trial was started approximately 0.5 s after the observer’s response.

In accordance with the staircase method, mean texture orientation was varied every 2 deg depending on the observer’s response. In each session, multiple staircase sequences corresponding to different experimental conditions were randomly interleaved. For each observer, at least 250 trials of data were collected for each central-orientation condition. The mean orientation of elements in the surrounding region giving a 50% clockwise response rate, defined as the subjective point of equivalence (PSE), was estimated via logit analysis and maximum likelihood^{41,42}. Trials in which masking did not occur were not used in the estimation of PSE in the masked condition. In order to cancel any bias in the perception of absolute vertical orientation for each observer, the PSE shift was calculated by subtracting the PSE at a relative central orientation of 0 deg from the PSE obtained for each condition.

Results. Figure 2a shows the proportion of trials in which observers perceived any line segments in the central region. In the no-mask condition (red), observers correctly discriminated between the physically filled and holed textures. In the masked condition (blue), observers rarely perceived the filled texture in the central region, thereby indicating a robust masking effect. Figure 2b shows the PSE shift as a function of the relative orientation of the central region. In the no-mask condition (red), the PSE shift is fairly close to the physical mean orientation of all elements as denoted by the gray dashed line. Alternatively, in the masked condition (blue), the PSE shift is close to zero regardless of the central region’s relative orientation.

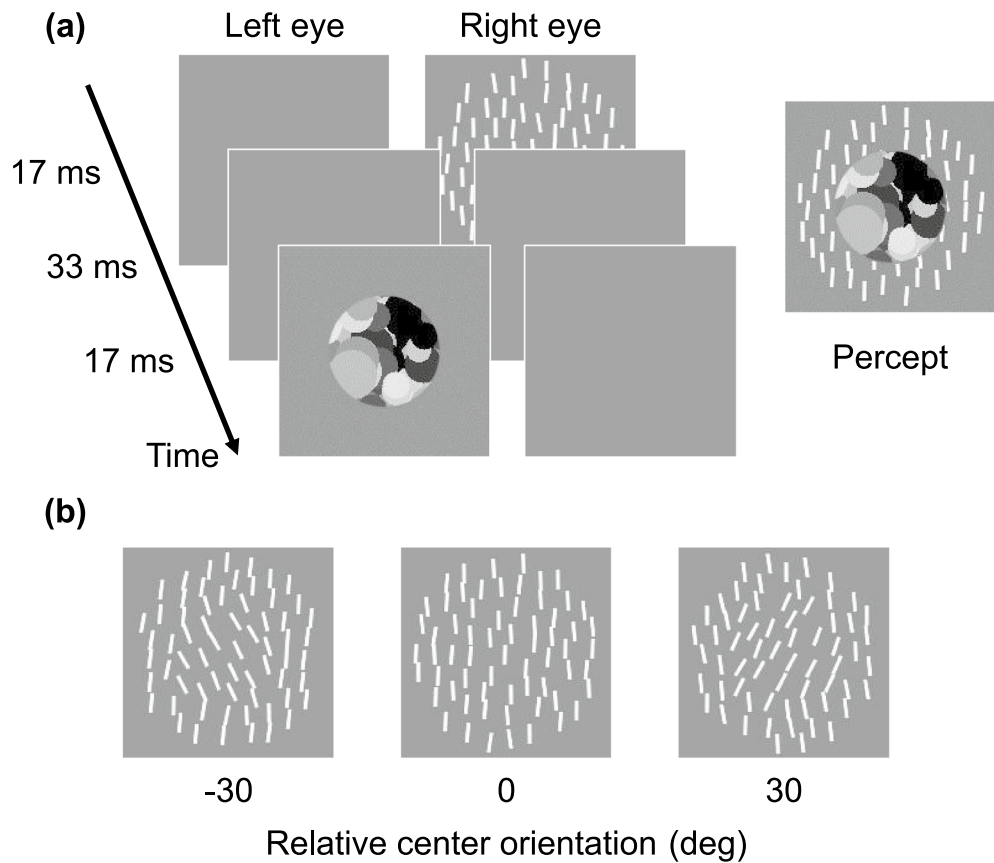


Figure 1. Stimuli used in Experiment 1. (a) Schematic of the temporal sequence of the stimuli. A line-based texture was briefly shown followed by a Mondrian pattern mask with an SOA. The right image illustrates a typical percept of the display. (b) The central region of the texture was varied by -30 , 0 , and $+30$ deg relative to the surrounding region.

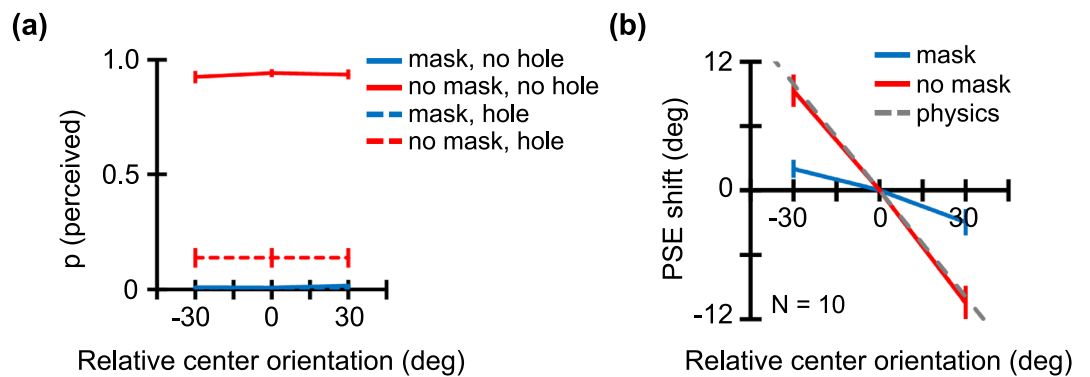


Figure 2. Results for Experiment 1. The red line shows results for the no-mask condition, and the blue line shows results for the masked condition. (a) The proportion of trials in which any line segments were perceived in the central region as a function of the relative orientation of the central region. The dashed line represents results for the texture with no central region. (b) The PSE shift as a function of the central region's relative orientation. The gray dashed line denotes the physical mean orientation of all elements. Error bars represent ± 1 SEM.

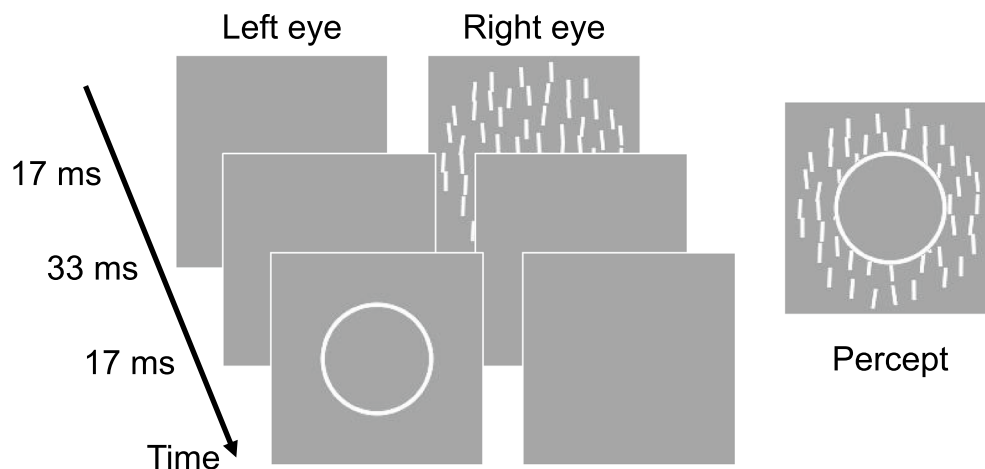


Figure 3. Stimuli used in Experiment 2. Schematic of the temporal sequence of the stimuli. A line-based texture was briefly shown followed by an annular contour mask with an SOA. The right image illustrates a typical percept of the display.

A two-way repeated measures ANOVA was performed with mask condition and relative center orientation as factors. To improve the estimate of the probability of obtaining the results, we report the key interaction for both the initial and full samples, as well as the $p_{\text{augmented}}$ statistic⁴³. After running the first 9 participants, results showed significant main effects of the relative center orientation ($F(1, 32) = 80.2, p < 0.0001, \eta^2 = 0.56, 1-\beta = 0.99$), an interaction between the two factors ($F(1, 32) = 30.2, p < 0.0001, \eta^2 = 0.21, 1-\beta = 0.89$), and no significant main effects of the mask ($F(1, 32) = 0.11, p = 0.74, \eta^2 < 0.01, 1-\beta = 0.05$). For the full sample of 10 participants, results showed significant main effects of the relative center orientation ($F(1, 36) = 88, p < 0.0001, \eta^2 = 0.57, 1-\beta = 0.99, p_{\text{augmented}} = [0.01, 0.01]$), an interaction between the two factors ($F(1, 36) = 31, p < 0.0001, \eta^2 = 0.20, 1-\beta = 0.90, p_{\text{augmented}} = [0.01, 0.01]$), and no significant main effects of the mask ($F(1, 36) = 0.008, p = 0.93, \eta^2 < 0.01, 1-\beta = 0.05, p_{\text{augmented}} = [0.87, 0.93]$). In the no-mask condition, the PSE is significantly shifted for both the -30 deg (two-tailed paired t test; $t(9) = 6.26, p = 0.0002, d = 1.98, 1-\beta = 0.98$) and 30 deg ($t(9) = -6.56, p = 0.0001, d = 2.07, 1-\beta = 0.99$) relative orientations of the central region. In the masked condition, on the other hand, the PSE is not significantly shifted for both the -30 deg (two-tailed paired t test; $t(9) = 2.34, p = 0.04, d = 0.74, 1-\beta = 0.17$) and 30 deg ($t(9) = -2.41, p = 0.04, d = 0.76, 1-\beta = 0.19$) relative orientations of the central region. Furthermore, between the mask and no-mask conditions, the PSE shift is significantly different for both the -30 deg (two-tailed paired t test; $t(9) = -4.55, p = 0.001, d = 2.00, 1-\beta = 0.99$) and 30 deg ($t(9) = 4.79, p = 0.001, d = 1.68, 1-\beta = 0.99$) relative orientations of the central region.

In summary, our data indicate that observers estimated the mean orientation of the texture similar to those of visible elements in the surrounding region. Results also support the conclusion that observers discounted the orientation of elements in the central region made invisible by the Mondrian mask. When presented with a texture with no central region (i.e. holed), observers reported ‘perceived’ on trials with a probability of 13% even in the no mask condition. This is likely due to the strong instructions which encouraged observers to respond whenever they saw anything other than a mask in the central regions.

Experiment 2

The results of Experiment 1 suggest that ensemble perception depends on the phenomenal awareness of masked elements. However, this conclusion discounts the possibility that orientation information in the central region was compromised by spatially overlapping luminance / pattern signals in the Mondrian mask. For example, edges in the Mondrian could have suppressed local orientation codes for individual elements. Even if we allow that neural representation of such low-level image features could survive under backward-masking conditions^{31–35}, it is still possible that information in texture and mask confounded the formation of ensemble representations. Therefore, in Experiment 2, we used a circular contour mask that prevented such spatial interference. Indeed, it is known that circular-contour masks strongly suppress conscious texture perception inside the annulus³⁹ but preserves the neural processing responsible for texture segregation and pop-out⁴⁰.

Method. In this experiment, the mask in Experiment 1 was replaced by a circular contour with a diameter of 2.0 deg (Fig. 3). The annulus consisted of a thin contour with a width of 0.2 deg and a luminance of 35 cd/m². To ensure that there was no phenomenal awareness, observers were given strict instructions to respond “perceived” if they saw anything other than a uniform hole in the central region of the texture. The others were the same as in Experiment 1. We initially recruited 9 participants, but two out of nine who could not discriminate texture orientation at all were excluded from the analysis. Based on sample size of observers in Experiment 1, we added four more observers to guard against power loss due to planned data exclusions referring to the early stage of the

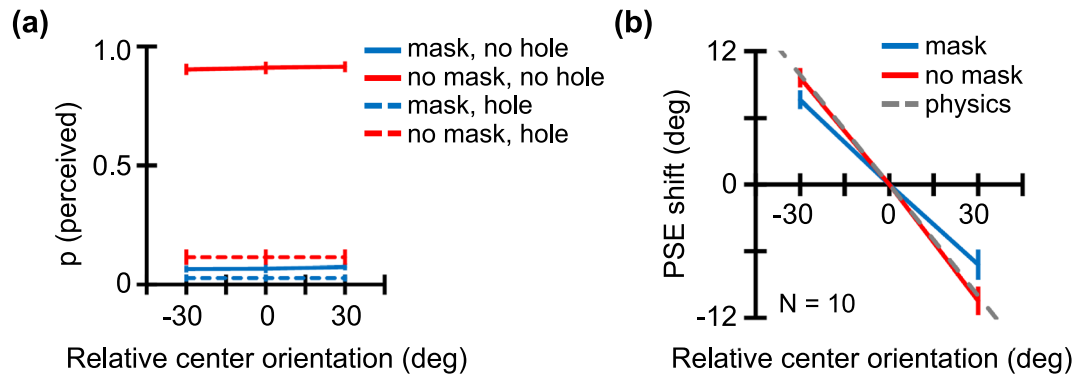


Figure 4. Results for Experiment 2. The red line shows results for the no-mask condition, and the blue line shows results for the masked condition. **(a)** The proportion of trials for which anything other than a uniform hole was perceived in the central region as a function of the central region's relative orientation. The dashed line represents results for the texture with no central region. **(b)** The PSE shift as a function of the central region's relative orientation. The gray dashed line denotes the mean physical orientation of all elements. Error bars represent ± 1 SEM.

experiment. In the end, thirteen observers, including twelve naïve students and one of the authors, participated in the experiment, but three observers were excluded from the analysis.

Results. Figure 4a shows the proportion of trials in which anything other than a uniform hole was perceived in the central region. We found that the texture in the central region (i.e., the one corresponding to the inside of the annular mask) was invisible in almost 100% of the trials in the masked condition. Figure 4b shows the PSE shift as a function of the central region's relative mean orientation. In contrast to Experiment 1, the PSE in both the no-mask (red) and mask (blue) conditions was shifted remarkably close to the physical mean of all elements, including the invisible elements in the central region.

A two-way repeated measures ANOVA was performed with mask condition and relative center orientation as factors. After running the first 7 participants, results showed a significant main effect of relative center orientation ($F(1, 24) = 196, p < 0.0001, \eta^2 = 0.86, 1-\beta = 0.99$), a non-significant main effect of mask condition ($F(1, 24) = 1.20, p = 0.29, \eta^2 = 0.01, 1-\beta = 0.06$), and a weak interaction between the two factors ($F(1, 24) = 5.89, p = 0.02, \eta^2 = 0.03, 1-\beta = 0.18$). For the full sample of 10 participants, results showed a significant main effect of relative center orientation ($F(1, 36) = 246, p < 0.0001, \eta^2 = 0.86, 1-\beta = 0.99, p_{\text{augmented}} = [0.01, 0.01]$), a non-significant main effect of mask condition ($F(1, 36) = 0.34, p = 0.57, \eta^2 < 0.01, 1-\beta = 0.05, p_{\text{augmented}} = [0.41, 0.57]$), and a weak interaction between the two factors ($F(1, 36) = 5.49, p = 0.02, \eta^2 = 0.02, 1-\beta = 0.13, p_{\text{augmented}} = [0.01, 0.03]$). In the no-mask condition, the PSE is significantly shifted for both the -30 deg (two-tailed paired t test; $t(9) = 10.9, p < 0.0001, d = 3.45, 1-\beta = 0.99$) and 30 deg ($t(9) = -8.16, p < 0.0001, d = 2.58, 1-\beta = 0.99$) relative orientations of the central region. Similarly, in the masked condition, the PSE is shifted significantly in both the central region's -30 deg ($t(9) = 9.20, p < 0.0001, d = 2.91, 1-\beta = 0.99$) and 30 deg ($t(9) = -5.30, p = 0.0005, d = 1.67, 1-\beta = 0.90$) relative-orientation conditions.

In summary, results of Experiment 2 indicate that observers assessed mean texture orientation using all elements, even without phenomenal awareness of elements in the central region made invisible by the annular mask. Our data support the notion that ensembles are readily computed without conscious detection of elements, provided that there is no interference from other stimuli such as the mask used in Experiment 1. In the no mask condition, observers still reported 'perceived' in 11% of trials in which textures had no central region (i.e. holed). This finding is probably attributable to our strict instruction to respond "perceived" whenever the observers saw anything other than a uniform hole in the central region.

General discussions

The present paper has sought to clarify the role of phenomenal awareness of texture elements in ensemble perception. To this end, we used a backward masking paradigm to measure whether mean perceived orientation of line-based textures is affected by invisible elements. A first experiment with a Mondrian mask suggested that perceived mean orientation was computed mainly from visible elements. However, a second experiment with a spatially non-overlapping annular contour mask revealed that observers computed mean orientation using all elements, including those elements made invisible by masking. These results suggest that human vision does not necessarily require phenomenal awareness of elements to spatially average the orientation of an ensemble.

These results are in favor of the idea that the visual system can efficiently recognize and memorize complex scenes by representing multiple elements of information as ensembles^{5,6}. In order to achieve such tiered functionality, the visual system must rapidly compute ensembles before the stimulus reaches conscious visual awareness. Indeed, there is much evidence that textures composed of low-level image features, such as orientation, are processed implicitly in the early visual cortex such as V1 and V2⁴⁴⁻⁴⁷. It has also been suggested that conscious awareness of visual stimuli is mediated by feedback loops from higher-order visual areas⁴⁸⁻⁵⁰, and these

feedback signals are known to be suppressed by backward masking or transcranial magnetic stimulation (TMS) applied slightly later than stimulus onset^{32,51}. According to these findings, visual cortex could generate ensemble representations in the feed-forward process prior to recurrent inputs from the feedback loop. Presumably, it is this feedback loop that our backward-masking display managed to disrupt.

Our data appear inconsistent with previous findings of crowding—a percept closely related to ensemble perception—in that ensemble perception does not always require phenomenal awareness whereas crowding is strongly correlated to phenomenal awareness³⁶ (but see also³⁷). However, crowding is defined as the difficulty in discriminating the property of a specific local element in a set (i.e. a problem of access)⁵². Reconciling apparent discrepancies between our results and the crowding literature may therefore hinge on the involvement of higher-order attentive processes rather than pre-attentive processes in ensemble computation. Indeed, crowding is often attributed to the spatial distribution of attention associated with localization errors and eye movements rather than to lower-order processes such as lateral interactions between elements^{22,53,54}. In a word, the visual system may transform the visual input into ensembles or textures codes, and these representations may be computed before conscious stimulus detection. In such a functional architecture, ensemble representations would enter consciousness via feedback loops. One may hypothesize that crowding occurs because information about an element at a given location is no longer accessible in compact ensemble representations.

In our experiments, observers utilized invisible elements to estimate mean orientation in the annular-mask condition but not in the Mondrian-mask condition. We take this as evidence that luminance and pattern information inside the Mondrian mask interferes with orientation information in the central region of the texture. There are several possible interpretations as to the specific level of visual information processing at which interference occurs. One possibility is that the Mondrian mask inhibited local orientation signals of individual texture elements, but this account is inconsistent with the findings that stimulus features are encoded even if stimulus perception is lost due to masking or binocular rivalry^{31–35}. Given that observers reported they could not see any line segments inside the central region, it is also unlikely that observer consciously perceived a mixture of texture and Mondrian. At the unconscious level, however, it is possible that information in the Mondrian was temporally averaged with (local and/or ensemble) representations of the central texture region. With the contour mask, on the other hand, invisible elements also contributed to the ensemble, since stimulus visibility was suppressed without any direct interference or averaging from the mask signal. Whereas the pattern mask has been widely used as a visual stimulus to impair the conscious perception of a target stimulus²⁸, the results of Exp. 1 may bring a caution about the nature of mask interference on the signal⁵⁵ when considering different classes of implicit representations.

The results of Experiment 2 imply that observers estimated mean orientation of all texture elements, including those in the masked central region. Yet, it is also true that observers achieved this judgment by seeing only the visible surrounding region. One explanation for this paradoxical observation is that observers had access to the neural representation of mean orientation of the whole texture which was formed at the unconscious level, then made a judgement independently how they perceived the visible surrounding regions. Another possible account is that observers made a judgment upon the apparent orientation of the visible surrounding regions as spatially averaged with the orientation of the masked central region. This explanation appears to be inconsistent with recent psychophysical data showing that perceptual assimilation between two clearly distinct uniform stimuli requires phenomenal awareness of the stimuli⁵⁶. On the other hand, it has been reported that compulsory averaging of the apparent orientation of local elements occurs in textures with noisy oriented elements like our stimuli²⁴. Our results could also be interpreted as indicating that such averaging, which can be regarded as partially equivalent to ensemble perception, takes place before conscious stimulus detection.

The results of the present study suggest the visual system can compute spatial ensembles without phenomenal awareness for low-level image features such as orientation. Yet, it is still unclear that the same would hold true for complex features such as faces and objects. Several psychophysical studies have pointed out that the computation of face ensembles may be a product of rapid serial encoding in a process with limited capacity⁵⁷. It has also been suggested that the perception of temporal ensembles of serially sampled visual information involves distinct processes from the computation of spatial ensembles⁵⁸. Importantly, though, it is known that the perception of temporal ensembles requires conscious access to the property of individual samples⁵⁹. Therefore, it is possible that spatial ensemble perception of high-level attributes (e.g., faces) requires conscious awareness of elements. Testing this possibility will be useful for comprehensive understanding of multi-level mechanisms for ensemble perception.

Data availability

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Received: 26 January 2022; Accepted: 30 June 2022

Published online: 13 July 2022

References

1. Thorpe, S., Fize, D. & Marlot, C. Speed of processing in the human visual system. *Nature* **381**(6582), 520–522 (1996).
2. Rousselle, G. A., Fabre-Thorpe, M. & Thorpe, S. J. Parallel processing in high-level categorization of natural images. *Nat. Neurosci.* **5**(7), 629–630 (2002).
3. Motoyoshi, I., Nishida, S. Y., Sharan, L. & Adelson, E. H. Image statistics and the perception of surface qualities. *Nature* **447**(7141), 206–209 (2007).
4. Chong, S. C. & Treisman, A. Representation of statistical properties. *Vis. Res.* **43**(4), 393–404 (2003).

5. Whitney, D., Haberman, J. & Sweeny, T. D. 49 from textures to crowds: multiple levels of summary statistical perception. In *The New Visual Neuroscience* (ed. J. S. Werner, L. M. Chalup.) 695–710 (The MIT Press, 2014).
6. Whitney, D. & Yamanashi Leib, A. Ensemble perception. *Annu. Rev. Psychol.* **69**, 105–129 (2018).
7. Dakin, S. C. & Watt, R. J. The computation of orientation statistics from visual texture. *Vis. Res.* **37**(22), 3181–3192 (1997).
8. Haberman, J. & Whitney, D. Rapid extraction of mean emotion and gender from sets of faces. *Curr. Biol.* **17**(17), R751–R753 (2007).
9. Haberman, J. & Whitney, D. Seeing the mean: Ensemble coding for sets of faces. *J. Exp. Psychol. Hum. Percept. Perform.* **35**(3), 718 (2009).
10. Leib, A. Y., Kosovicheva, A. & Whitney, D. Fast ensemble representations for abstract visual impressions. *Nat. Commun.* **7**(1), 1–10 (2016).
11. Sweeny, T. D., Haroz, S. & Whitney, D. Perceiving group behavior: Sensitive ensemble coding mechanisms for biological motion of human crowds. *J. Exp. Psychol. Hum. Percept. Perform.* **39**(2), 329 (2013).
12. Balas, B., Nakano, L. & Rosenholtz, R. A summary-statistic representation in peripheral vision explains visual crowding. *J. Vis.* **9**(12), 13–13 (2009).
13. Keshvari, S. & Rosenholtz, R. Pooling of continuous features provides a unifying account of crowding. *J. Vis.* **16**(3), 39–39 (2016).
14. Oliva, A. & Torralba, A. Building the gist of a scene: The role of global image features in recognition. *Prog. Brain Res.* **155**, 23–36 (2006).
15. Rosenholtz, R., Huang, J. & Ehinger, K. A. Rethinking the role of top-down attention in vision: Effects attributable to a lossy representation in peripheral vision. *Front. Psychol.* **3**, 13 (2012).
16. Rosenholtz, R., Huang, J., Raj, A., Balas, B. J. & Ilie, L. A summary statistic representation in peripheral vision explains visual search. *J. Vis.* **12**(4), 14–14 (2012).
17. Brady, T. F. & Alvarez, G. A. Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychol. Sci.* **22**(3), 384–392 (2011).
18. Alvarez, G. A. & Oliva, A. Spatial ensemble statistics are efficient codes that can be represented with reduced attention. *Proc. Natl. Acad. Sci. U. S. A.* **106**(18), 7345–7350 (2009).
19. Treisman, A. How the deployment of attention determines what we see. *Vis. Cogn.* **14**(4–8), 411–443 (2006).
20. Bouma, H. Interaction effects in parafoveal letter recognition. *Nature* **226**(5241), 177–178 (1970).
21. Levi, D. M. Crowding—An essential bottleneck for object recognition: A mini-review. *Vis. Res.* **48**(5), 635–654 (2008).
22. Pelli, D. G., Palomares, M. & Majaj, N. J. Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *J. Vis.* **4**(12), 12–12 (2004).
23. Whitney, D. & Levi, D. M. Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends Cogn. Sci.* **15**(4), 160–168 (2011).
24. Parkes, L., Lund, J., Angelucci, A., Solomon, J. A. & Morgan, M. Compulsory averaging of crowded orientation signals in human vision. *Nat. Neurosci.* **4**(7), 739–744 (2001).
25. Block, N. Two neural correlates of consciousness. *Trends Cogn. Sci.* **9**(2), 46–52 (2005).
26. Blake, R. & Logothetis, N. K. Visual competition. *Nat. Rev. Neurosci.* **3**(1), 13–21 (2002).
27. Bonneh, Y. S., Cooperman, A. & Sagi, D. Motion-induced blindness in normal observers. *Nature* **411**(6839), 798–801 (2001).
28. Breitmeyer, B. & Ögmen, H. *Visual Masking: Time Slices Through Conscious and Unconscious Vision* (Oxford University Press, 2006).
29. Motoyoshi, I. & Hayakawa, S. Adaptation-induced blindness to sluggish stimuli. *J. Vis.* **10**(2), 16–16 (2010).
30. Tsuchiya, N. & Koch, C. Continuous flash suppression reduces negative afterimages. *Nat. Neurosci.* **8**(8), 1096–1101 (2005).
31. Blake, R. & Fox, R. Adaptation to invisible gratings and the site of binocular rivalry suppression. *Nature* **249**(5456), 488–490 (1974).
32. Fahrenfort, J. J., Scholte, H. S. & Lamme, V. A. Masking disrupts reentrant processing in human visual cortex. *J. Cogn. Neurosci.* **19**(9), 1488–1497 (2007).
33. Fang, F. & He, S. Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nat. Neurosci.* **8**(10), 1380–1385 (2005).
34. Tong, F., Meng, M. & Blake, R. Neural bases of binocular rivalry. *Trends Cogn. Sci.* **10**(11), 502–511 (2006).
35. Williams, M. A., Morris, A. P., McGlone, F., Abbott, D. F. & Mattingley, J. B. Amygdala responses to fearful and happy facial expressions under conditions of binocular suppression. *J. Neurosci.* **24**(12), 2898–2904 (2004).
36. Wallis, T. S. & Bex, P. J. Visual crowding is correlated with awareness. *Curr. Biol.* **21**(3), 254–258 (2011).
37. Ho, C. & Cheung, S. H. Crowding by invisible flankers. *PLoS ONE* **6**(12), e28814 (2011).
38. Faul, F., Erdfelder, E., Lang, A. G. & Buchner, A. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**(2), 175–191 (2007).
39. Caputo, G. Texture brightness filling-in. *Vis. Res.* **38**(6), 841–851 (1998).
40. Motoyoshi, I. Texture filling-in and texture segregation revealed by transient masking. *Vis. Res.* **39**(7), 1285–1291 (1999).
41. Burns, E. J., Yang, W. & Ying, H. Friend effects framework: Contrastive and hierarchical processing in cheerleader effects. *Cognition* **212**, 104715 (2021).
42. Ying, H., Burns, E., Choo, A. & Xu, H. Temporal and spatial ensemble statistics are formed by distinct mechanisms. *Cognition* **195**, 104128 (2020).
43. Sagarin, B. J., Ambler, J. K. & Lee, E. M. An ethical approach to peeking at data. *Perspect. Psychol. Sci.* **9**(3), 293–304 (2014).
44. Freeman, J. & Simoncelli, E. P. Metamers of the ventral stream. *Nat. Neurosci.* **14**(9), 1195–1201 (2011).
45. Freeman, J., Ziemba, C. M., Heeger, D. J., Simoncelli, E. P. & Movshon, J. A. A functional and perceptual signature of the second visual area in primates. *Nat. Neurosci.* **16**(7), 974–981 (2013).
46. Ziemba, C. M., Freeman, J., Movshon, J. A. & Simoncelli, E. P. Selectivity and tolerance for visual texture in macaque V2. *Proc. Natl. Acad. Sci. U. S. A.* **113**(22), E3140–E3149 (2016).
47. Zipsper, K., Lamme, V. A. & Schiller, P. H. Contextual modulation in primary visual cortex. *J. Neurosci.* **16**(22), 7376–7389 (1996).
48. Lamme, V. A. Neural mechanisms of visual awareness: A linking proposition. *Brain Mind* **1**(3), 385–406 (2000).
49. Lamme, V. A. & Roelfsema, P. R. The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci.* **23**(11), 571–579 (2000).
50. Lamme, V. A., Super, H., Landman, R., Roelfsema, P. R. & Spekreijse, H. The role of primary visual cortex (V1) in visual awareness. *Vis. Res.* **40**(10–12), 1507–1521 (2000).
51. Pascual-Leone, A. & Walsh, V. Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science* **292**(5516), 510–512 (2001).
52. Manassi, M. & Whitney, D. Multi-level crowding and the paradox of object recognition in clutter. *Curr. Biol.* **28**(3), R127–R133 (2018).
53. He, S., Cavanagh, P. & Intriligator, J. Attentional resolution and the locus of visual awareness. *Nature* **383**(6598), 334–337 (1996).
54. Yeshurun, Y. & Rashal, E. Precueing attention to the target location diminishes crowding and reduces the critical distance. *J. Vis.* **10**(10), 16–16 (2010).
55. Ross, J. & Speed, H. D. Contrast adaptation and contrast masking in human vision. *Proc. Biol. Sci.* **246**(1315), 61–70 (1991).
56. Tomassini, A. & Solomon, J. A. Awareness is the key to attraction: Dissociating the tilt illusions via conscious perception. *J. Vis.* **14**(12), 15–15 (2014).

57. Ji, L., Chen, W., Loeys, T. & Pourtois, G. Ensemble representation for multiple facial expressions: Evidence for a capacity limited perceptual process. *J. Vis.* **18**(3), 17–17 (2018).
58. Yashiro, R., Sato, H., Oide, T. & Motoyoshi, I. Perception and decision mechanisms involved in average estimation of spatiotemporal ensembles. *Sci. Rep.* **10**(1), 1–10 (2020).
59. De Lange, F. P., Van Gaal, S., Lamme, V. A. & Dehaene, S. How awareness changes the relative weights of evidence during human decision-making. *PLoS Biol.* **9**(11), e1001203 (2011).

Acknowledgements

The present study was supported by JSPS KAKENHI Grant-In-Aid JP20H01782.

Author contributions

I.M. conceived the study. T.S. and I.M. designed the study. T.S. performed experiments and analyzed data. T.S. and I.M. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to I.M.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022