A computational mechanism for seeing dynamic deformation

Takahiro Kawabe* & Masataka Sawayama

NTT Communication Science Laboratories

*Corresponding author
3-1, Morinosato Wakamiya, Atsugi, 243-0198, Kanagawa, Japan
E-mail: takkawabe@gmail.com
Facsimile: +81-46-240-4716

Abstract

Human observers are able to perceptually discriminate dynamic deformation of materials in the real world. However, the psychophysical and neural mechanisms responsible for the perception of dynamic deformation have not been fully elucidated. By using a visual illusion wherein a static bar apparently deforms against a background grating's orientation and spatial frequency, we show that the spatial and directional patterns of the simulated responses of direction-selective units (i.e., MT pattern motion cells) consistently explain an observer's perception of dynamic deformation. Simulating the direction-selective unit's responses to stimuli, we found that the perception of dynamic deformation was based on the unit's higher-order mechanism monitoring the spatial modulation of direction responses. The results indicate that the perception of dynamic deformation is dependent on a high-level neural mechanism monitoring the spatial patterns of direction-selective units of MT areas.

Author Summary

Human observers easily detect the deformation of objects in the scene. How do the observers perceive the deformation from image signals? The present study tried to answer the question from the psychophysical and computational points of view. In the psychophysical experiment, we used a novel visual illusion wherein a bar apparently deformed in front of tilted drifting gratings and consequently specified image parameters to cause deformation perception. In the computational analysis, we showed that in order to explain psychophysical data for illusory deformation perception it was necessary to assume an unknown additional unit that monitored spatial frequency of the pattern of direction responses of MT cells that were sensitive to image motion. Seeing real deformations could also be explained by the model assuming the additional unit. We also showed that the mechanism was sensitive to luminance-based motion signals, not to contrast-based motion signals. The results indicate that the mechanism that tunes to spatial frequency of the directional responses mediates the perception of deformation.

Introduction

Materials in the real world are often non-rigid. This material non-rigidity dynamically produces image deformations of the material when the materials physically deform with time. The dynamic deformation of retinal images is a rich source of visual information allowing the visual system to assess material properties in the real world. For example, from the dynamic deformation of retinal images, human observers can recognize transparent liquid [1], transparent gas [2], the elasticity and/or stiffness of materials [3-6], the stiffness of fabrics [7-9], and more.

Importantly, however, the visual mechanism for seeing dynamic deformation itself has not been thoroughly examined. Previous studies have phenomenally reported that dynamic deformation was perceived when local motion integration did not provide evidence for rigid motion [10,11] at a layer-represented level [12]. Though successfully addressing when rigid motion perception was violated, these previous studies have not mentioned the conditions accompanying the perception of dynamic deformation. Although some studies have addressed the computational model for the detection of two-dimensional motion patterns containing shearing and/or rotating motion [13,14], the model is not directly applicable in order to explain deformation perception because, as shown in the previous study[10], a shearing motion pattern does not always cause deformation perception. That is, the detection of the shearing motion itself does not lead to shearing deformation perception; deformation patterns with higher deformation frequency do not produce deformation perception [10]. No previous computational model does not account for the dependency of deformation perception on the spatial frequency of deformation. Hence, it is necessary to assume additional mechanism in order to explain deformation perception in human observers. Moreover, it remained unclear what visual information could be employed by the visual system to generate the representation of dynamic deformation. A previous study [15] has shown that motion is important information for discerning the shape of non-rigidly deforming objects. The present study thus focuses on how motion information contributes to the perception of dynamic deformation.

Here we propose a novel visual illusion in which a static bar with solid edges apparently deforms against a slightly tilted drifting grating (See Figures 1a and 1b). We presume that moiré patterns [15-17], which were generated between the bar's edge and background grating, produced motion signals that were related to the apparent deformation. Phenomenally, we observed that manipulating the spatial frequency as well as the orientation of the background grating strongly affected the appearance of dynamic deformation. Thus we expected that the mechanism responsible for seeing dynamic deformation might be clarified by closely investigating the associated circumstances when an apparent deformation occurred in the illusion.

The purpose of this study was to psychophysically and computationally specify the mechanism that underlies the perception of dynamic deformation. In Experiment 1, we show that the orientation and spatial frequency of the background grating are critical to the illusory perception of deformation. By simulating responses of direction-selective units (i.e., MT pattern motion cells [19-21]), we show how the spatial pattern of the direction responses of the units is the critical component that determines the perception of dynamic deformation. In Experiment 2, using stimuli with an actually deformed bar, we present results essentially identical to the results of Experiment 1. In Experiment 3, we show that the perception of dynamic deformation is luminance-based, not contrast-based. We then discuss how these results establish that the visual perception of dynamic deformation, whether the deformation is actual or illusory, is mediated by a high-level mechanism that monitors the spatial and directional patterns of responses of velocity-selective units.

Results and Discussion

Experiment 1

The purpose of Experiment 1 was to specify psychophysical parameters to induce the illusion of dynamic deformation and look into a possible relationship between the psychophysical data and stimulus data for the responses of direction-selective units (MT pattern motion cells). In a psychophysical experiment, the orientation and spatial frequency of drifting background gratings were systematically

controlled (Supplementary videos 1-3), while a bar, which was located in front of the gratings, was completely static. We asked observers to report whether the bar dynamically deformed or not. We also simulated the responses of the MT area's direction-selective units to the stimuli as used in the psychophysical experiment, and explored the relationship between the perception of dynamic deformation and the responses of the direction-selective units.

We calculated the proportion of trials in which the observer reported the dynamic deformation of the bar and plotted them as a function of grating orientation for each spatial frequency condition in Figure 1c. We conducted a two-way repeated measures ANOVA with grating orientation and grating spatial frequency as within-subject factors. The main effect of the grating spatial frequency was significant [$F(2,22) = 27.440$, $p < .0001$, $\eta^2_p = 0.72$]. The main effect of the grating orientation was also significant [$F(2,22) = 255.772$, $p < .0001$, $\eta^2_p = 0.96$]. Interaction between the two factors was significant [$F(12,132) = 15.362$, $p < .0001$, $\eta^2_p = 0.58$).

Simulation of MT responses

Although it was shown that the orientation and spatial frequency of the drifting background grating were critical to the perception of dynamic deformation, it was still unclear how the visual system utilized image features in the moiré patterns that were generated between the bar and background gratings. To specify the mechanism for seeing dynamic deformation, we decided to check the relationship between the perception of dynamic deformation and the simulated responses of direction-selective units to the stimuli. On the basis of the previous literature [19-21,23,24], we constructed the simulation by employing a standard procedure with the following steps (see also Figure 2b): 1) convolving stimuli with spatiotemporal filters to get spatiotemporal energy [19,23,24], 2) half-wave rectification [19], 3) divisive normalization[19], 4) spatial pooling [19,23], 5) the weighted summation of spatiotemporal energy in an opponent fashion [19-21], 6) half-wave rectification [19], and 7) divisive normalization (See Methods for details).

Figure 2d shows the responses of direction-selective units as functions of the preferred direction of the units and spatial position, for each of the orientation and spatial frequency conditions of background grating. There are several important features. First, the spatial pattern of the unit responses got finer as the grating orientation increased. Second, the spatial pattern of the unit responses got relatively finer as the grating spatial frequency increased. Third, with smaller grating orientation, the unit responses showed high activities in a spatially alternate manner between leftward and rightward directions while with larger grating orientation, the responses showed high activities between downward and rightward directions.

Properties of units monitoring the spatial pattern of direction responses

Based on the simulation, we suggested that the psychophysical results could be well accounted for by assuming higher-order units that monitored the spatial pattern of the direction responses. That is, we predicted that the higher-order units would determine whether the input signal came from deformation or not, depending on the spatial pattern of the direction responses. To test this prediction, we conducted a pattern-matching analysis using a kernel as shown in Figure 3a. From the results of the simulation of the direction-selective units (Figure 2d), it was plausible to assume that the higher-order units could tune to spatially sinusoidal modulation of motion direction as shown in Figure 3a. We call the pattern having this modulation a kernel. To assess the similarity between each spatial pattern of direction responses and the pattern of the kernel, manipulating the spatial frequency of the modulation in the kernel, we calculated the normalized cross-correlation (NCC) between the kernel and the spatial pattern of the unit responses and checked how the calculated NCC was related to the psychophysical responses. It was expected that the higher NCC for kernels would lead to the higher proportions of deformation responses only when the kernels had specific range of spatial frequency of direction responses.

We evaluated how an exponential function was fitted to the psychophysical data as a function of the NCC for each stimulus (Supplementary Figure 1) and show a plot of the coefficient of determination (r^2) in Figure 3b. Interestingly, at the specific ranges of the spatial frequency of the modulation in the kernel, the coefficient of determination was high. We assumed that the brain monitors all kernels with

different weights and hence by using the coefficients r^2 as weights, we summed NCC values for each stimulus, and the results of the weight-summed NCC are plotted in the left panel of Figure 3c. The pattern of the NCC as a function of the parameters of the stimuli was similar to the pattern of the psychophysical performances. We calculated the coefficient of determination (r^2) of the exponential function to the psychophysical data as a function of weight-summed NCC and found a high coefficient ($r^2 = 0.92$), which indicates that the observers' performance is explained well by the pattern of the weight-summed NCC.

The results suggest that the deformation perception in our illusion is likely mediated by the higher-order unit monitoring the spatial pattern of the responses of the direction-selective units. Moreover, the higher-order unit seems to tune to the specific band of the spatial frequency of motion responses. We suggest that the higher-order unit is responsible for the deformation perception. In turn, the results indicate that the mechanism for deformation perception may be selective for the spatial frequency of motion responses, in a sense consistent with the previous study showing that the spatial frequency of image deformation determines the appearance of image deformations [1].

Experiment 2

The purpose of this experiment was to confirm whether the higher-order units we proposed in the previous analyses could well explain the deformation perception due to stimuli wherein a bar was actually deformed (Supplementary video 4, *https://bit.ly/2JanDT0*). A previous study [10] showed that when a sinusoidally modulated line translated, the line was seen as deforming when the spatial frequency of the sinusoidal modulation was low, consistent with the results in our previous experiment. Here we asked the observers to report whether the bar (which was actually deformed) was seen as deforming, and examined the relationship between the observers' responses and the pattern of velocity-selective units.

Figure 4b shows the proportion of trials wherein the observers reported the bar as dynamically deforming. By using the proportion, we conducted a repeated-measures one-way ANOVA with the spatial frequency of modulation as a within-subject factor. The main effect was significant [$F(6,36) = 38.358$, $p < .0001$, $\eta^2_p = 0.86$]. Multiple comparison showed that the proportions in the 0.1, 0.2, 0.4, and 0.8

cycles per degree(cpd) conditions were significantly higher than the proportions in the 1.6, 3.2, and 6.4 cpd conditions ($p < .05$).

As in the previous analyses, we simulated the responses of direction-selective units for the stimuli as used in this experiment, and confirmed that the obtained spatial pattern of direction responses (Figure 4c) was similar to the patterns (Figure 2d) for the stimuli as used in the previous experiment. In a similar way to the previous analysis, by manipulating the spatial frequency of the kernel (Figure 3a) we calculated the NCC for each stimulus and fitted an exponential function to the psychophysical data as a function of the NCC for each stimulus. The coefficients of determination (r^2) of the fitting are shown in Figure 4d. As in the previous analysis, r^2 peaked at the specific spatial frequency bands of the spatial pattern of direction responses. We calculated weight-summed NCC for each spatial frequency of modulation in the stimuli (Figure 4e) and got a high coefficient ($r^2 = 0.96$) which indicates that the proportion of trials with deformation reports was exponentially correlated with the weight-summed NCC.

The results indicate that for the stimuli causing a deformation illusion, the higher-order units tuning the spatial frequency of direction responses may operate in the perception of deformation for physically deformed stimuli.

When the spatial frequency of sinusoidal modulation of a bar is high, the simulated responses of direction-selective units showed preferences for downward motion (3.2 and 6.4 cpd conditions of Figure 4c). This pattern of responses is consistent with the psychophysical data in a previous study [10] which showed that the sinusoidal modulation of a line perceptually resulted in unidirectional translation when it had a high spatial frequency of modulation. The results of our simulations indicate that the model we employed could precisely capture the properties of human motion perception.

Experiment 3

The purpose of this experiment was to additionally confirm whether the deformation perception could be obtained on the basis of the contrast-based moiré pattern. In Experiment 1, we used dark and bright bars against a background drifting grating, so the moiré pattern generated between the bar and the background grating was always defined by luminance. By testing the condition with the bar luminance at a neutral gray (Supplementary video 5, *https://bit.ly/2JanDT0*), we

investigated whether the contrast-based moiré pattern also contributed to the perception of dynamic deformation. If the deformation perception was operated by the mechanism that had a function equivalent to the velocity-selective units, because the unit was assumed to have selectivity to luminance, no strong influence of contrast-based moiré pattern would be observed.

In Figure 5, the proportion of trials with reports of bar deformation is plotted for each background orientation condition as a function of the luminance of the bar. As in Experiment 1, the deformation perception was more often reported with a grating orientation of 1° than 16°. Interestingly, the proportion suddenly dropped when the luminance of the bar was set at the level of neutral gray. By using the proportion, we conducted a two-way repeated measures ANOVA with the bar luminance and background grating orientation as within-subject factors. The main effect of the bar luminance was significant [$F(8,88) = 16.36$, $p < .0001$, $\eta^2_p = 0.60$]. The main effect of the background grating orientation was also significant [$F(1,11) = 1463.35$, $p < .0001$, $\eta^2_p = 0.99$]. Interaction between the two factors was also significant [$F(8,88) = 17.763$, $p < .0001$, $\eta^2_p = 0.62$]. The simple main effect based on the significant interaction showed that the proportion for 76 cd/m^2 was significantly lower than the proportions for other luminance conditions when the background orientation was 1° ($p < .05$). Moreover, the proportion for 76 cd/m^2 was the chance level. The results showed that the deformation perception was attenuated when the bar luminance was neutral gray, suggesting that the mechanism for the deformation perception tunes to luminance-defined rather than contrast-defined features in our illusion. Moreover, the illusory deformation was still reported even when the luminance of the bar itself was outside of the luminance contrast range of the background grating. The results suggest that the deformation illusion triggered by moiré patterns occurs with a flexible relationship between the bar and the background grating.

General Discussion

This study investigated the mechanism responsible for deformation perception by human observers. Our data showed that the perception of dynamic deformation was mediated by higher-order units tuning to the spatial pattern of the responses of

the direction-selective units. Moreover, it was shown that the units were selective for luminance-defined features, rather than contrast-defined features.

Previous studies have reported that the human visual system is sensitive to direction-defined stripes [25,26] and/or direction-defined gratings [27]. Does the mechanism responsible for motion-defined structures also mediate the perception of deformation? Previous studies have consistently shown that the mechanism for the detection of a motion-defined structure is low-spatial-frequency selective. On the other hand, as shown in Figures 3c and Figure 4c, the mechanism for deformation perception may have band-pass properties. We therefore suggest that the mechanism for deformation perception is not simply equivalent to the mechanism for detecting motion-defined structures, though it is possible that some processing procedures are shared between them.

What is the possible neural mechanism for seeing deformation? It is well known that velocities are processed in the MT area. In general, velocities are captured via a large receptive field and hence processed globally [30, 31]. On the other hand, some studies have reported that an identical receptive field of the MT area could locally respond to independent velocities [32,33]. The kind of local velocity extraction may mediate the perception of deformation, though further clarification is necessary to evaluate this possibility, since the spatial frequency properties of the local velocity extraction in the MT area are still an open issue. Structure from motion, which is occasionally involved with a complex motion structure, is also processed in the MT area [34,35]. There is thus a possibility that the MT area is responsible for deformation perception, which is also involved with complex motion structure.

In this study, we did not closely investigate the role of speed in the perception of dynamic deformation. Specifically, we assessed only direction parameters in the computation model, while the model proposed in previous studies could assess both direction and speed [19]. This was because the most revealing speed of image motion signals in our stimuli with both illusory and real bar deformations was easily anticipated, and so we could determine the optimal speed parameter of the model in advance. On the other hand, it is known that local motion speeds determine the appearance of image deformation [36], and thus unclear how deformation speed influences the appearance of dynamic deformation. Manipulating the speed of deformation as well as checking the speed parameters in the MT model need to be tested in future investigations.

So far, vision scientists have investigated the appearance of moiré patterns themselves [16-18]. In this study, we investigated an illusion wherein moiré patterns were generated between a static bar and a drifting background grating. The phenomenon indicates that local moiré patterns can alter the overall appearance of an object that is related to the pattern. As shown in Figure 1, the moiré-based deformation occurs even with a static frame. On the other hand, dynamic rather than static versions of stimuli caused a stronger effect. We suggest that higher-order units monitoring the spatial pattern of direction responses here operate in the deformation perception of dynamic stimuli. For static stimuli, orientation-based rather than direction-based mechanisms may be responsible for the deformation perception.

In this study, we focused only on shearing deformation [37] but not on compressive deformation. Because the background grating in our illusion stimuli contained luminance variation along a one-dimensional space, only shearing deformation could be induced. On the other hand, in the image deformation of natural materials such as the flow of a transparent liquid, both shearing and compressive deformations seem to exist [1]. However, just how shearing and compressive deformation are processed and interact with each other in the visual system remains an open question. Psychophysical and computational investigation of simultaneous shearing and compressive deformations will lead to further understanding of how the visual system detects and interprets image deformation in natural scenes.

Methods

Experiment 1

Observers. 12 people (10 females and 2 males) participated in this experiment. Their mean age was 38.2 (SD: 7.63). All observers in this study reported having normal or corrected-to-normal visual acuity. They were recruited from outside the laboratory and received payment for their participation. Ethical approval for this study was obtained from the ethics committee at Nippon Telegraph and Telephone Corporation (Approval number: H28-008 by NTT Communication Science Laboratories Ethical Committee). The experiments were conducted according to principles that have their origin in the Helsinki

Declaration. Written, informed consent was obtained from all observers in this study.

Apparatus. Stimuli were presented on a 21-inch iMac (Apple Inc. USA) with a resolution of 1280 x 720 pixels and a refresh rate of 60 Hz. A colorimeter (Bm-5A, Topcon, Japan) was used to measure the luminance emitted from the display. A computer (iMac, Apple Inc., USA) controlled stimulus presentation, and data were collected with PsychoPy v1.83 [21-22].

Stimuli. As shown in Figure 1a (and Supplementary videos 1-3), a vertical bar (0.6 deg wide × 5.00 deg high) was presented in front of a drifting grating. For each stimulus, the luminance of the bar was randomly chosen from two levels (37 and 112 cd/m^2). The orientation of the background grating was selected from the following 7 levels (0.5, 1, 2, 4, 8, 12, and 16°). The spatial frequency was selected from the following 3 levels (6.4, 12.9, and 25.8 cycles per degree). Drift temporal frequency was kept constant at 1 Hz. Drift direction was randomly determined. The luminance contrast of the grating was set at 0.75, and thus the luminance level of the grating ranged between 37 and 112 cd/m^2. The drifting grating was windowed by a horizontal Gaussian envelope with a standard deviation of 0.62 deg.

Procedure. Each observer was tested in a lit chamber. The observers sat 102 cm from the display. With each trial, a stimulus clip having a static bar and drifting grating was presented for 3 seconds. After the disappearance of the clip, white visual noise (with each cell subtending 0.16 deg × 0.16 deg) was presented until the observer's response. The task of the observers was to judge whether the static bar dynamically deformed or not. The judgment was delivered by pressing one of the assigned keys. Each observer had four sessions, each consisting of 3 spatial frequencies × 7 orientations × 5 repetitions. Within each session the order of trials was pseudo-randomized. Thus, each observer had 420 trials in total. It took 30–40 minutes for each observer to complete all of four sessions.

Simulation of MT responses

As described in the main text, we simulated the responses of direction-selective units on the basis of previous studies [16-18,20,21]. We extracted a stimulus area near the right vertical edge of the bar that was presented against a background grating tilted rightward and drifting rightward (Figure 2a) and simulated the responses of the direction-selective units to the area. Based on the spatial

frequency of the background grating, we changed the width of the extracted area for analysis (0.16, 0.08, 0.04, deg for 6.4, 12.9, and 24.8 cpd conditions). The height of the extracted area was constant at 4.97 deg. The extracted area was first analyzed by a set of spatiotemporal filters. The spatial size and spatial wavelength of the spatiotemporal filters were also consistent with the width of extracted area, that is, 0.16 deg × 0.16 deg, 0.08 deg × 0.08 deg, and 0.04 deg × 0.04 deg for 6.4, 12.9, and 24.8 cpd condition, respectively. The number of filter orientations was 24 (i.e., 15° steps). The position of the filters did not spatially overlap, and hence 32, 64, and 128 filters covered the extracted area for 6.4, 12.9, and 24.8 cpd conditions, respectively. The temporal size and temporal wavelength of the filters were kept constant at 30 frames (0.5 seconds). Because the temporal frequency of modulation in the stimuli was 1 Hz, the filters could capture a half-cycle of modulation. We adopted the temporal properties because we wanted to extract a one-way modulation of the moiré pattern. The responses of the filters were half-wave rectified and normalized as reported in the previous study [19]. The normalized responses were spatially pooled among four adjacent filters, yielding 29, 61, and 125 responses. The pooled responses were filtered by a direction-tuned filter [20-21] that tuned to the motion direction using a cosine function. That is, at this level, the rectified and normalized outputs of the spatiotemporal filters were summed with weightings in an opponent fashion. The filters' preferred directions numbered 24 (i.e., 15° steps). The filtered responses were half-wave rectified and then normalized as in a previous study [21], yielding the responses of direction selective units as functions of the preferred direction of units and spatial position, as shown in Figure 2d.

Properties of units monitoring the spatial pattern of direction responses

The kernel was defined by the product of the spatially sinusoidal pattern *S* and the directionally sinusoidal pattern *D* (Figure 3a). *S* was defined by the following formula,

$$S(x) = sin(2fx + \varphi) \qquad (1),$$

wherein f denotes spatial frequency, ϕ denotes phase, and *x* denotes spatial position. *D* was defined by the following formula,

$$D(\theta) = cos(2\pi(\theta - \alpha)) \quad (2),$$

wherein θ denotes the motion direction and ranges from 0 to 2π, and a denotes the preferred direction of the kernel. Here, α was set to 0 deg (leftward direction) on the basis of the spatial motion of the velocity-selective units as shown in Figure 2d. Thus, a kernel K was defined as the product of S and D,

$$K(x, \theta) = S(x)D(\theta) \qquad (3).$$

To see how the kernel matched the spatial pattern of motion direction in Figure 2d, we calculated the normalized cross-correlation between them using a standard procedure [23].The f of (1) took one of the following levels: 0.1, 0.2, 0.4, 0.8, 1.6, 2.4 and 3.2 cpd. The ϕ was tested in 32 steps (each step = 0.0625 π), and the maximum value among the 32 outcomes based on the 32 steps was considered to be the NCC of the kernel.

Experiment 2

Observers. 7 people (5 females and 2 males) participated in this experiment. Although all of them had participated in the previous experiment, none was aware of the specific purpose of the experiment because there was no preliminary explanation or debriefing provided.

Apparatus. Apparatus was identical to that used in Experiment 1.

Stimuli. In stimuli (Figure 4a and Supplementary video 4), the edge of a vertical bar (0.6 deg width × 5.0 deg height) was horizontally deformed at one of the following 7 spatial frequencies (0.1, 0.2, 0.4, 0.8, 1.6, 3.2, and 6.4 cpd). The amplitude was kept constant at 0.04 cpd. With each stimulus, upward or downward drifting was randomly given to the modulation. The modulation temporal frequency was 1 Hz. As in the previous experiment, the luminance of the bar was randomly chosen as one of two levels (38 and 114 cd/m^2). The luminance of background was 76 cd/m^2.

Procedure. Procedure was identical to that in the previous experiment except for the following. With each trial, a stimulus clip having a deforming bar was presented for 3 seconds. After the disappearance of the clip, visual white noise (each cell subtending 0.16 deg × 0.16 deg) was presented until the observer's response. The task of the observers was to judge whether the static bar dynamically deformed or not. Each observer had two sessions, each consisting of 7 spatial frequencies of the modulation × 10 repetitions. Within each session the order of trials was pseudo-randomized. Thus, each observer had 140 trials in total. It took approximately 20 minutes for each observer to complete all of both sessions.

Experiment 3

Observers. 12 people who had participated in Experiment 1 again participated in this experiment. Still, none was aware of the specific purpose of the experiment.

Apparatus. Apparatus was identical to that used in Experiment 1.

Stimuli. Stimuli were identical to those used in Experiment 1 except for the following. The background grating orientation was 1° or 16°, which respectively produced strong deformation and non-deformation responses in Experiment 1. The grating spatial frequency was kept constant at 12.88 cpd. As shown in Supplementary video 5, the luminance of the bar was randomly chosen from the nine levels (0.0, 17.5, 37, 58, 76, 95, 112, 132, and 148 cd/m^2 wherein 76 cd/m^2 was the luminance of a neutral gray level).

Procedure. Procedure was identical to that used in Experiment 1 except for the following. Each observer had four sessions, each consisting of 2 levels of grating orientation × 9 luminance levels of the bar × 5 repetitions. Within each session the order of trials was pseudo-randomized. Thus, each observer performed 360 trials in total. It took approximately 40 minutes for each observer to complete all of both sessions.

References

1. Kawabe, T., Maruya, K., & Nishida, S. (2015). Perceptual transparency from image deformation. Proceedings of the National Academy of Sciences, 112(33), E4620–E4627. https://doi.org/10.1073/pnas.1500913112
2. Kawabe, T., & Kogovšek, R. (2017). Image deformation as a cue to material category judgment. Scientific Reports, 7, 44274. https://doi.org/10.1038/srep44274
3. Masuda, T., Matsubara, K., Utsumi, K., & Wada, Y. (2015). Material perception of a kinetic illusory object with amplitude and frequency changes in oscillated inducer motion. Vision Research, 109, Part B, 201–208. https://doi.org/10.1016/j.visres.2014.11.019
4. Masuda, T., Sato, K., Murakoshi, T., Utsumi, K., Kimura, A., Shirai, N., … Wada, Y. (2013). Perception of Elasticity in the Kinetic Illusory Object with

Phase Differences in Inducer Motion. PLoS ONE, 8(10), e78621. https://doi.org/10.1371/journal.pone.0078621

5. Paulun, V. C., Schmidt, F., Assen, J. J. R. van, & Fleming, R. W. (2017). Shape, motion, and optical cues to stiffness of elastic objects. , 17(1):20, 1–22. https://doi.org/10.1167/17.1.20
6. Schmidt, F., Paulun, V. C., Assen, J. J. R. van, & Fleming, R. W. (2017). Inferring the stiffness of unfamiliar objects from optical, shape, and motion cues. , 17(3):18, 1–17. https://doi.org/10.1167/17.3.18
7. Bi, W., Jin, P., Nienborg, H., & Xiao, B. (2018). Estimating mechanical properties of cloth from videos using dense motion trajectories: Human psychophysics and machine learning. Journal of Vision, 18(5):12, 1–20. https://doi.org/10.1167/18.5.12
8. Bi, W., & Xiao, B. (2016). Perceptual Constancy of Mechanical Properties of Cloth Under Variation of External Forces. Proceedings of the ACM Symposium on Applied Perception, 19–23. https://doi.org/10.1145/2931002.2931016
9. Bi, W., Jin, P., Nienborg, H., & Xiao, B. (2019). Manipulating patterns of dynamic deformation elicits the impression of cloth with varying stiffness. Journal of Vision, 19(5):18, 1–18, https://doi.org/10.1167/19.5.18.
10. Nakayama, K., & Silverman, G. H. (1988a). The aperture problem--I. Perception of nonrigidity and motion direction in translating sinusoidal lines. Vision Research, 28(6), 739–746. https://doi.org/10.1016/0042-6989(88)90052-1
11. Nakayama, K., & Silverman, G. H. (1988). The aperture problem—II. Spatial integration of velocity information along contours. Vision Research, 28(6), 747–753. https://doi.org/10.1016/0042-6989(88)90053-3
12. Weiss, Y., & Adelson, E. H. (2000). Adventures with Gelatinous Ellipses—Constraints on Models of Human Motion Analysis. Perception, 29(5), 543–566. https://doi.org/10.1068/p3032
13. Zhang, K., Sereno, M. I., & Sereno, M. E. (1993). Emergence of Position-Independent Detectors of Sense of Rotation and Dilation with Hebbian Learning: An Analysis. Neural Computation, 5(4), 597–612. https://doi.org/10.1162/neco.1993.5.4.597

14. Sachtler, W. L., & Zaidi, Q. (1995). Visual processing of motion boundaries. Vision Research, 35(6), 807–826. https://doi.org/10.1016/0042-6989(94)00160-N
15. Jain, A., & Zaidi, Q. (2011). Discerning nonrigid 3D shapes from motion cues. Proceedings of the National Academy of Sciences, 108(4), 1663–1668. https://doi.org/10.1073/pnas.1016211108
16. Oster, G. (1965). Optic art. Applied Optics, 4(11), 1359-1369.
17. Spillmann, L. (1993). The Perception of Movement and Depth in Moiré Patterns. Perception, 22(3), 287–308. https://doi.org/10.1068/p220287
18. Wade, N. J. (2007). The Stereoscopic Art of Ludwig Wilding. Perception, 36(4), 479–482. https://doi.org/10.1068/p3604ed
19. Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. Vision Research, 38(5), 743–761.
20. Perrone, J. A., & Krauzlis, R. J. (2014). Simulating component-to-pattern dynamic effects with a computer model of middle temporal pattern neurons. Journal of Vision, 14(1):19, 1–17. https://doi.org/10.1167/14.1.19
21. Perrone, J. A., & Krauzlis, R. J. (2008). Spatial integration by MT pattern neurons: A closer look at pattern-to-component effects and the role of speed tuning. , 8(9):1, 1–14. https://doi.org/10.1167/8.9.1
22. Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. Journal of Neuroscience Methods, 162(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017
23. Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. Frontiers in Neuroinformatics, 2, 10. https://doi.org/10.3389/neuro.11.010.2008
24. Nishimoto, S., & Gallant, J. L. (2011). A Three-Dimensional Spatiotemporal Receptive Field Model Explains Responses of Area MT Neurons to Naturalistic Movies. Journal of Neuroscience, 31(41), 14551–14564. https://doi.org/10.1523/JNEUROSCI.6801-10.2011
25. Mante, V., & Carandini, M. (2005). Mapping of Stimulus Energy in Primary Visual Cortex. Journal of Neurophysiology, 94(1), 788–798. https://doi.org/10.1152/jn.01094.2004
26. Lewis, J. P. (1995). Fast Template Matching, Vision Interface 95, Canadian Image Processing and Pattern Recognition Society, Quebec City, Canada, May 15-19, p. 120-123.

27. van Doorn, A. J., & Koenderink, J. J. (1982). Spatial properties of the visual detectability of moving spatial white noise. Experimental Brain Research, 45(1–2), 189–195.
28. van Doorn, A. J., & Koenderink, J. J. (1982). Visibility of movement gradients. Biological Cybernetics, 44(3), 167–175. https://doi.org/10.1007/BF00344272
29. Nakayama, K., Silverman, G. H., MacLeod, D. I., & Mulligan, J. (1985). Sensitivity to shearing and compressive motion in random dots. Perception, 14(2), 225–238.
30. Newsome, W. T., Wurtz, R. H., Dursteler, M. R., & Mikami, A. (1985) Deficits in visual motion processing following ibotenic acid lesions of the middle temporal visual area of the macaque monkey. Journal of Neuroscience, 5, 825-840.
31. Dubner, R., & Zeki, S. M. (1971) Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. Brain Research, 35, 528-532.
32. Majaj, N. J., Carandini, M., & Movshon, J. A. (2007). Motion Integration by Neurons in Macaque MT Is Local, Not Global. Journal of Neuroscience, 27(2), 366–370. https://doi.org/10.1523/JNEUROSCI.3183-06.2007
33. Perrone 2018
34. Bradley, D. C., Chang, G. C., Andersen, R. A. (1998) Encoding of three-dimensional structure-from-motion by primate area MT neurons. Nature 392, 714-717.
35. Grunewald, A., Bradley, D. C., Andersen, R. A. (2002) Neural correlates of structure-from-motion perception in macaque V1 and MT. Journal of Neuroscience, 22, 6195–6207.
36. Kawabe, T. (2018). Linear Motion Coverage as a Determinant of Transparent Liquid Perception. I-Perception, 9(6), 2041669518813375. https://doi.org/10.1177/2041669518813375
37. Nakayama, K., & Tyler, C. W. (1981). Psychophysical isolation of movement sensitivity by removal of familiar position cues. Vision Research, 21(4), 427–433. https://doi.org/10.1016/0042-6989(81)90089-4

Figure legends

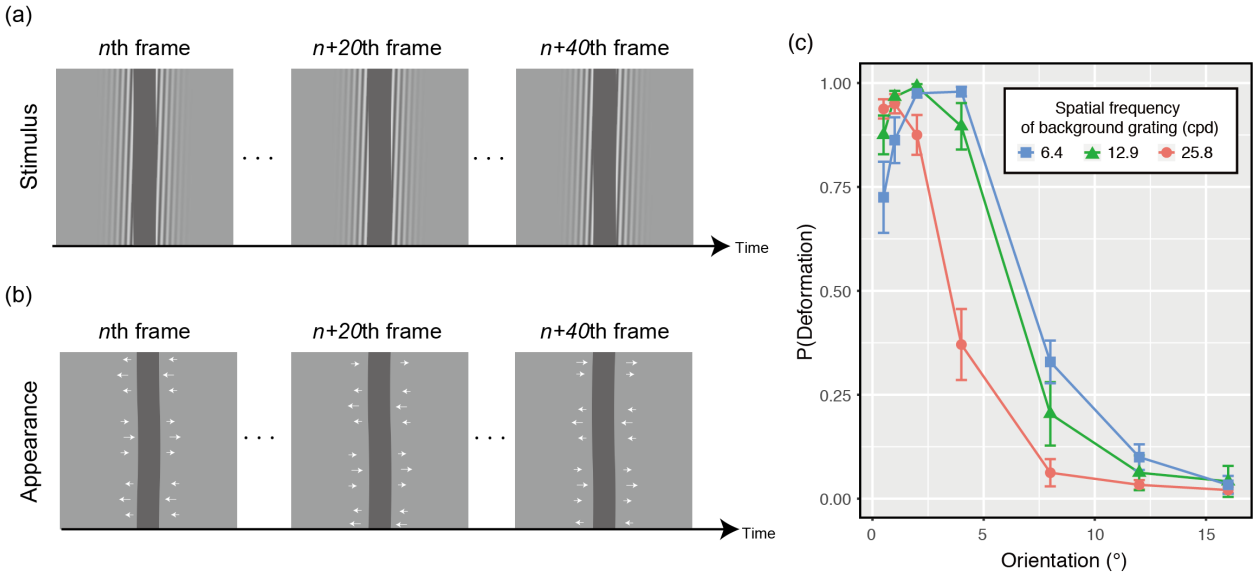


Figure 1 (a,b) Schematic illustrations of the appearance of the illusion we found. (a) Some snapshots of a stimulus clip causing the illusion (see also Supplementary videos 1-3, https://bit.ly/2JanDT0). (b) Some examples of the appearance of the black bar in (a). A bar with straight edges apparently deforms when a tilted background grating drifts. White arrows indicate vectors of apparent deformation. (c) Experiment 1 results. Error bars denote standard errors of the mean (N = 12).

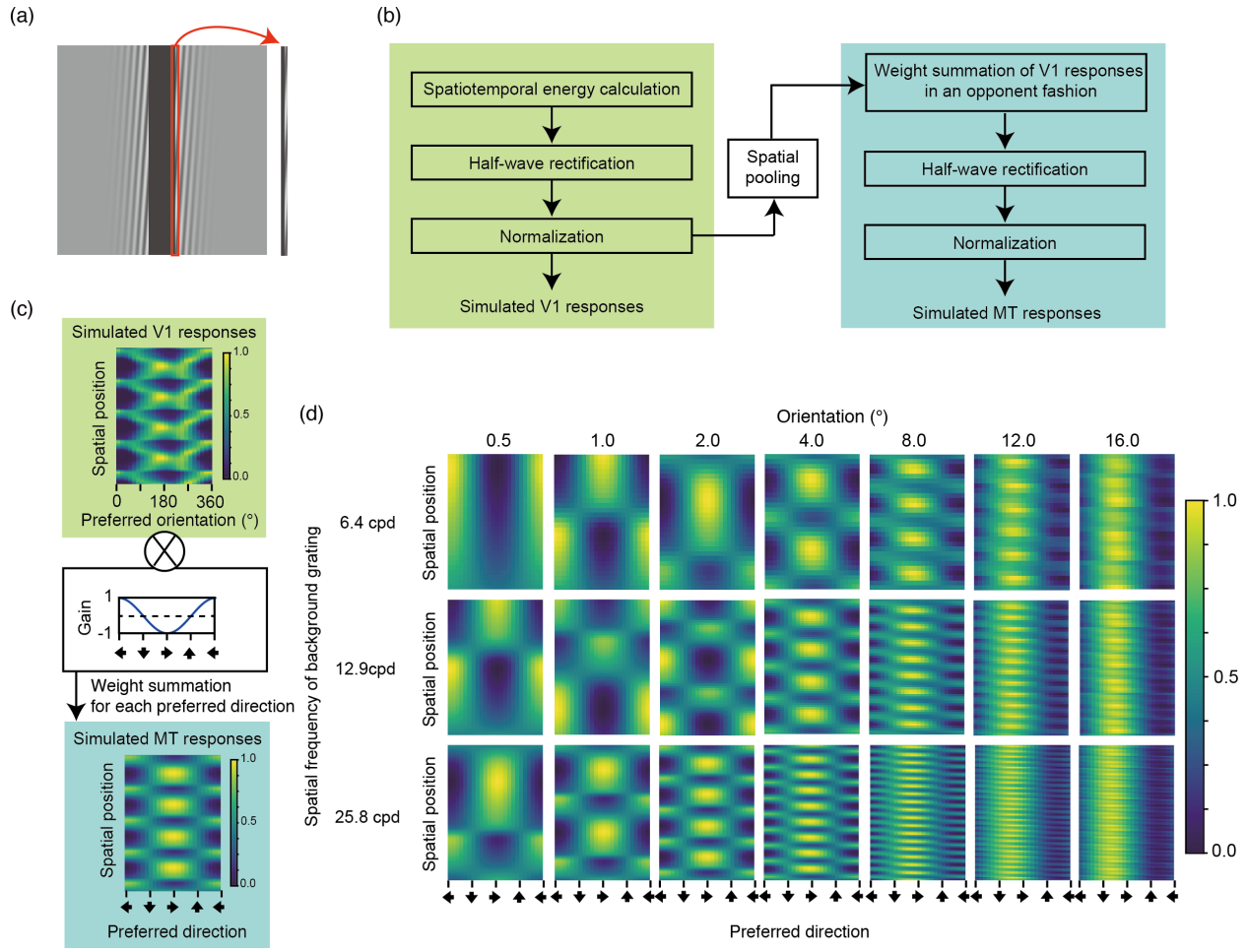


Figure 2 (a) Extracted area for simulation. The red-bound area was used to simulate the response of the direction-*selective units. (b) A pipeline of our model. (c) The product of simulated V1 responses and opponent filtering is weight-summed for each preferred orientation, and this returns simulated MT responses. In this figure, opponent filtering has a preference for the leftward direction. d) The simulated responses of the velocity-selective units for each spatial frequency of background grating are visualized as a function of spatial position and preferred motion direction. In this figure the range of each density plot is normalized between 0 and 1. Raw values were used for further analysis.*

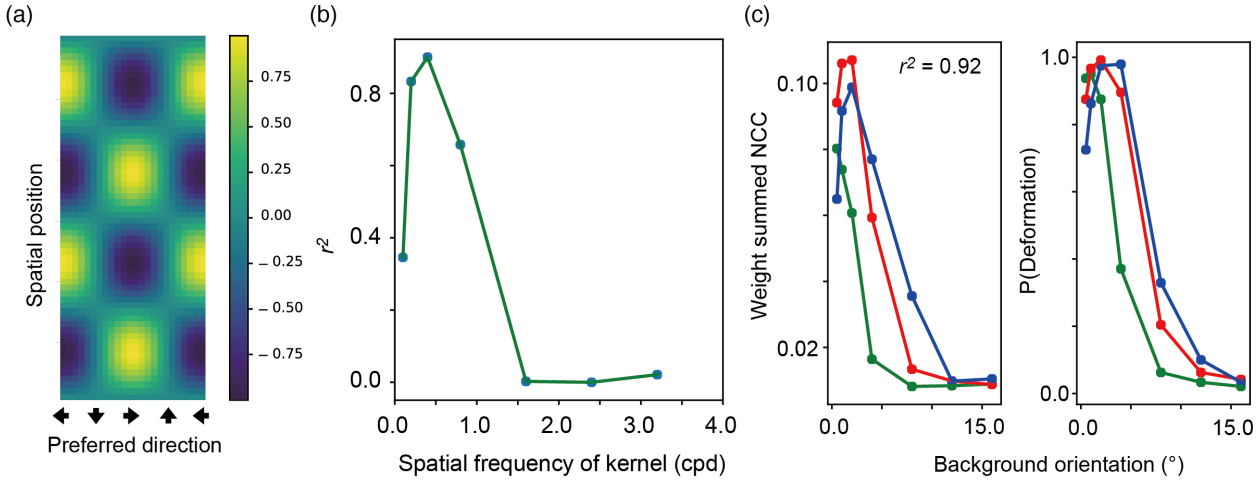


Figure 3 (a) An example of the kernel which was employed here. (b) The vertical axis denotes the coefficient determination (r^2) for the fitting of an exponential function to the proportion of trials with deformation reports as a function of NCC, and the horizontal axis denotes the spatial frequency of modulation of a kernel. See Supplementary Figure 1 for further information. (c) Left: Weight-summed NCC for each stimulus condition as a function of background orientation. Right: Proportion of trials with deformation reports (here Figure 1b is replotted for comparison).

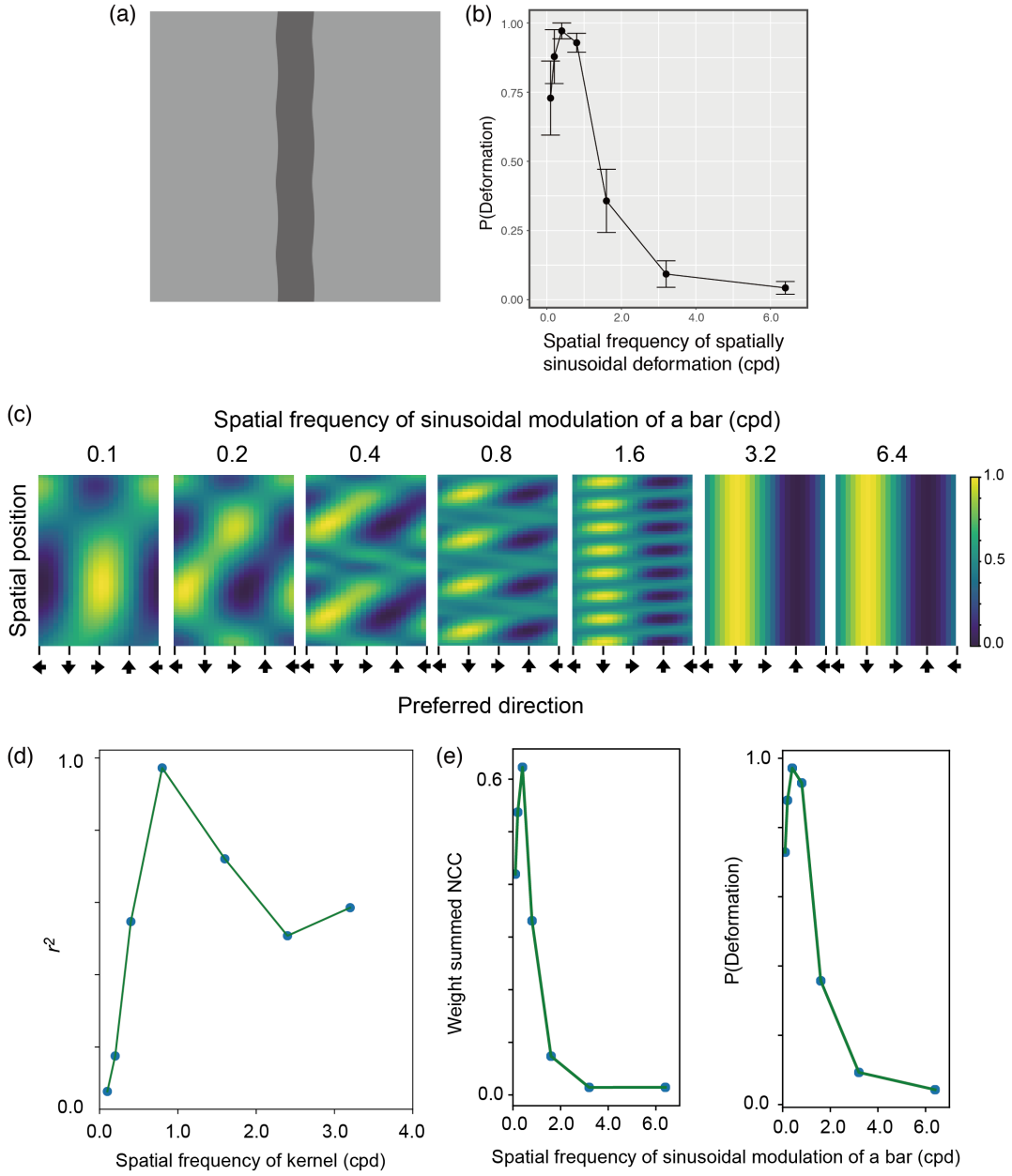


Figure 4 (a) A snapshot of a stimulus clip as used in Experiment 2 (Supplementary video 4). (b) Experiment 2 results. Error bars denote standard errors of the mean (N = 7). (c) Simulated responses of direction-selective units for the stimuli of Experiment 2. Here, the spatial size of the spatiotemporal filters employed for analysis was 0.08 x 0.08 deg. The temporal frequency of the filter was 2 Hz. As in the analysis of Experiment 1, we used seven levels (0.1, 0.2, 0.4, 0.8, 1.6, 2.4 and 3.2 cpd) of the modulation spatial frequency of a kernel. In this panel the range of

each density plot is normalized between 0 and 1. Raw values were used for further analysis. (d) For each of the spatial frequencies of the kernel, the coefficient of determination (r^2) o*f the fitting of an exponential function to the proportion of trials with deformation reports as a function of NCC is plotted. (e) Left: Weight-summed NCC as a function of background orientation. Right: Proportion of trials with deformation reports (here Figure 4b is replotted for comparison).*

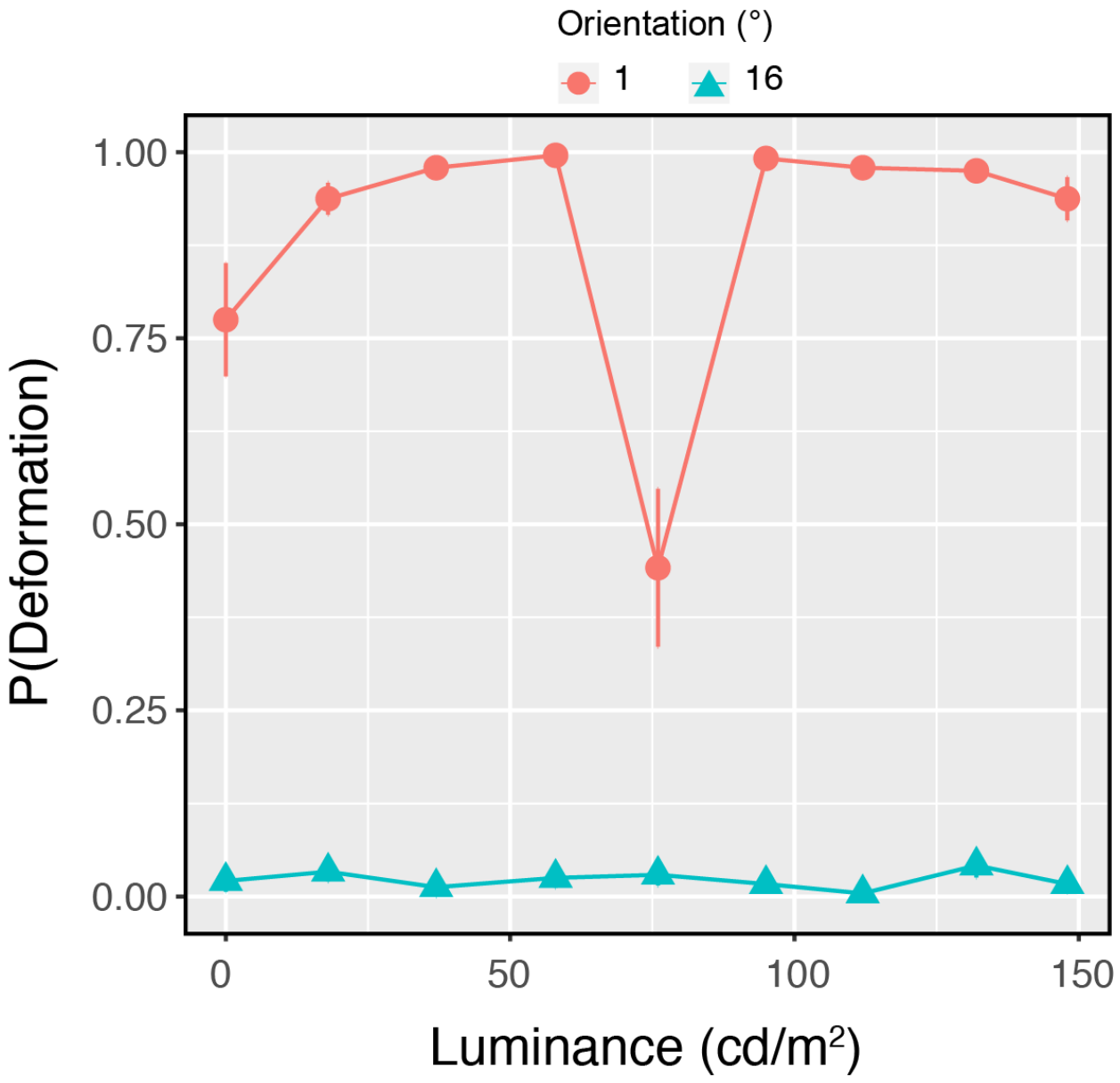
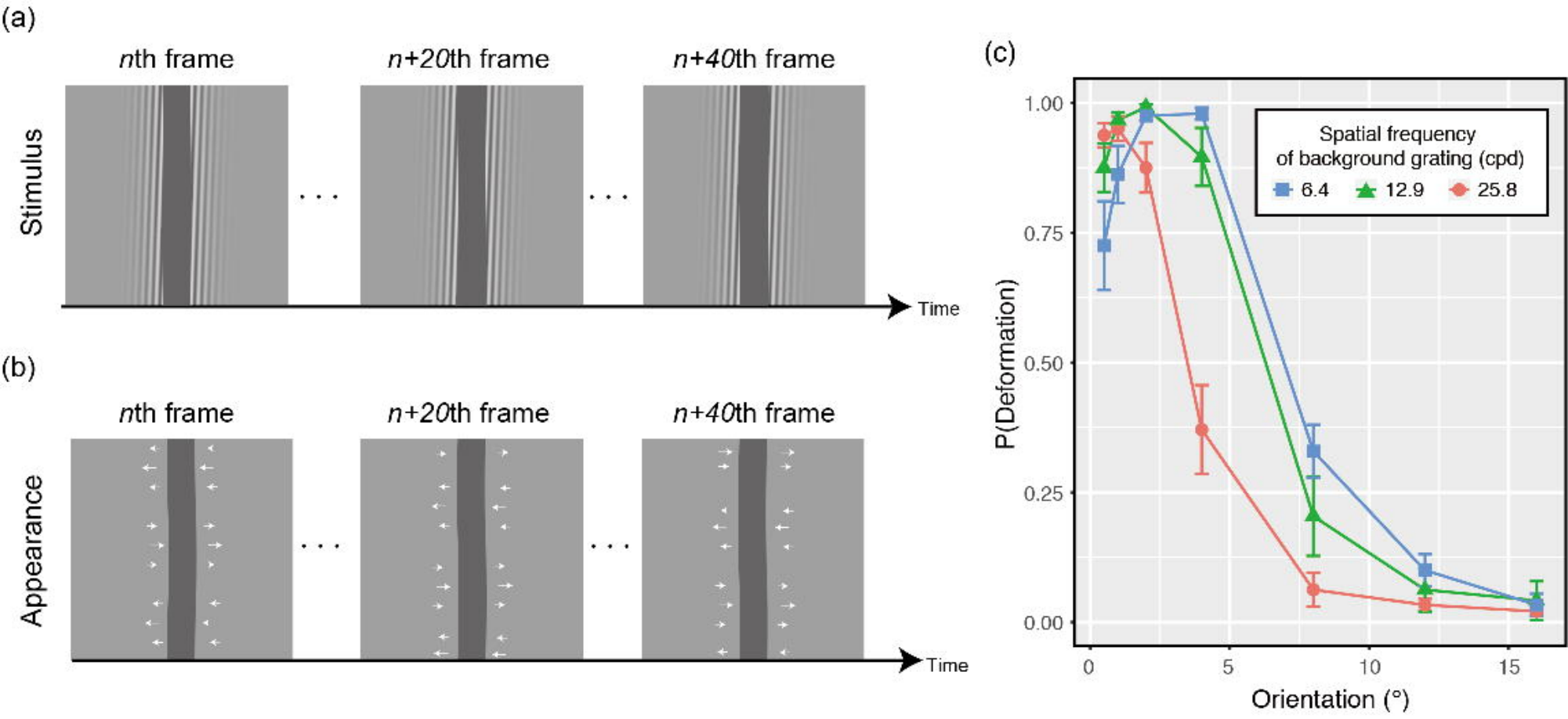
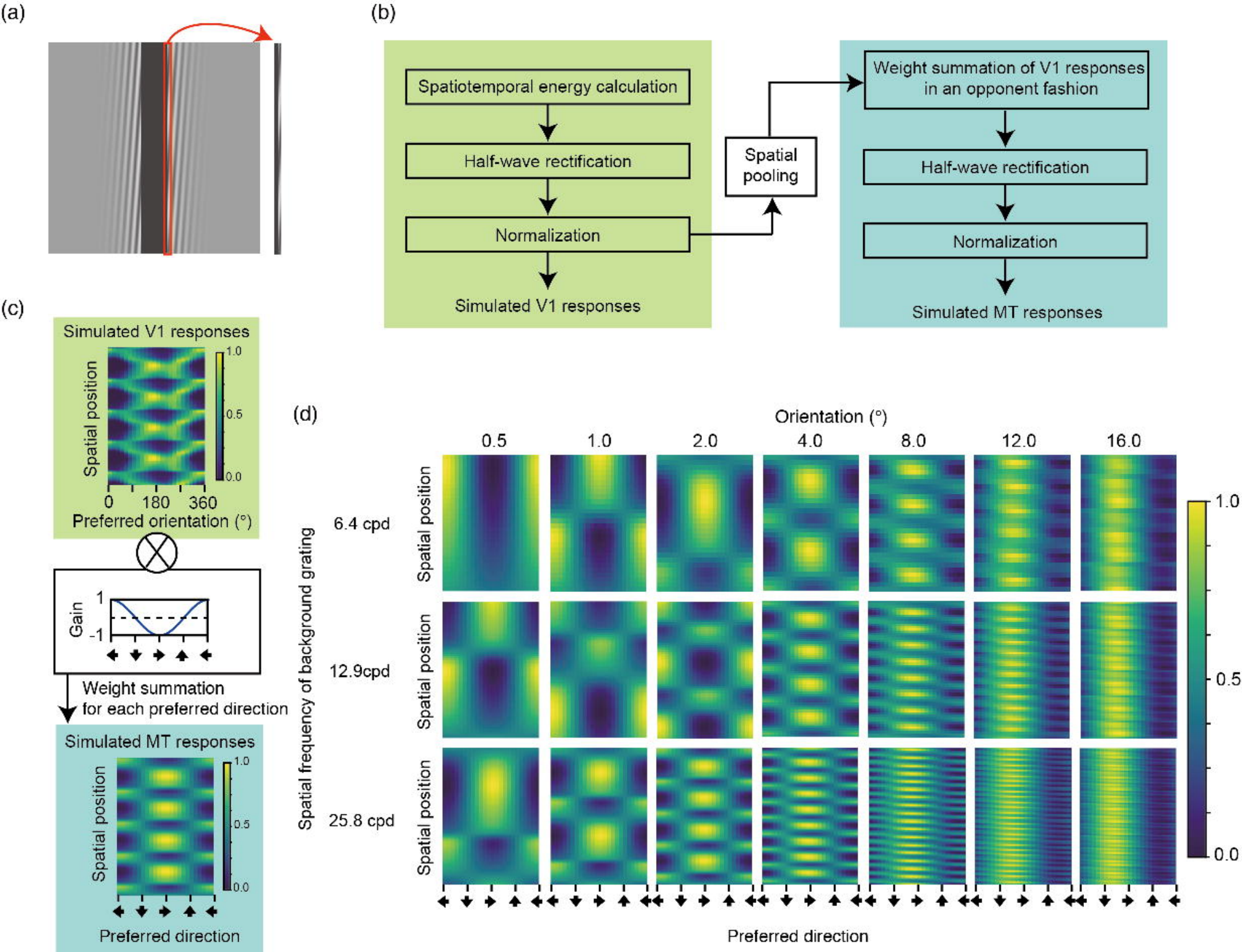
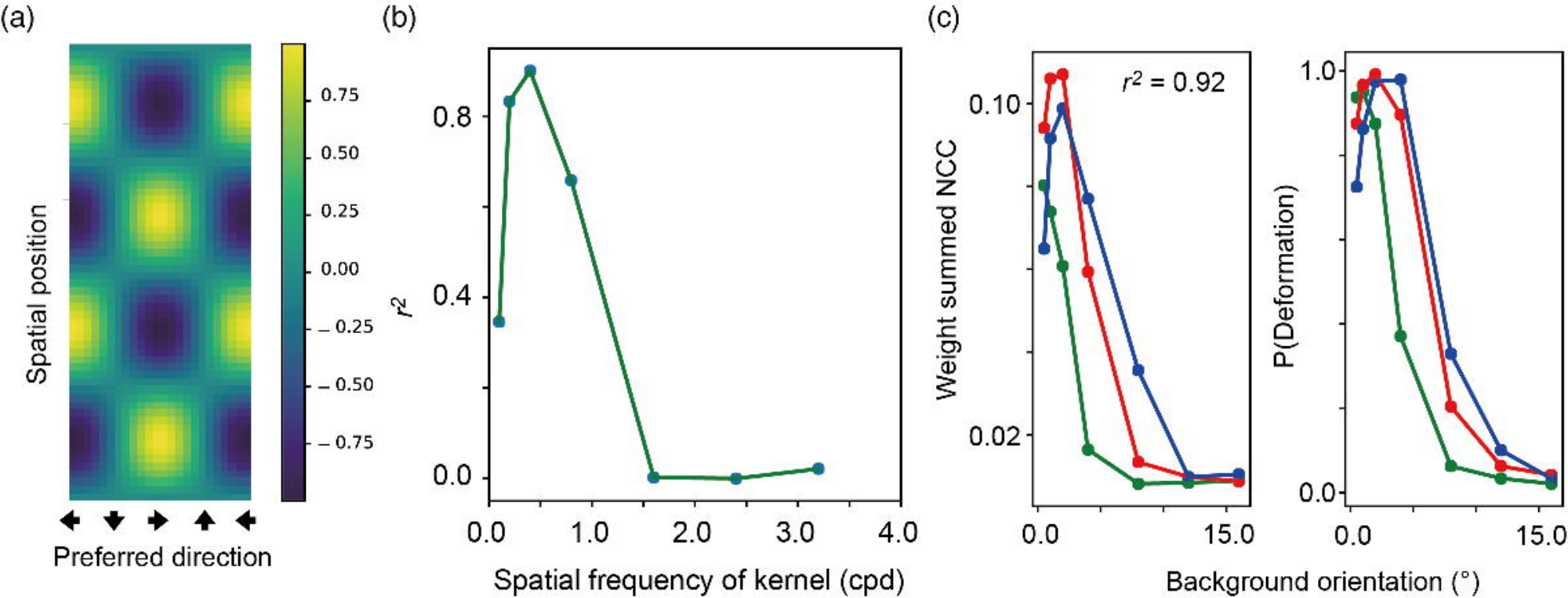
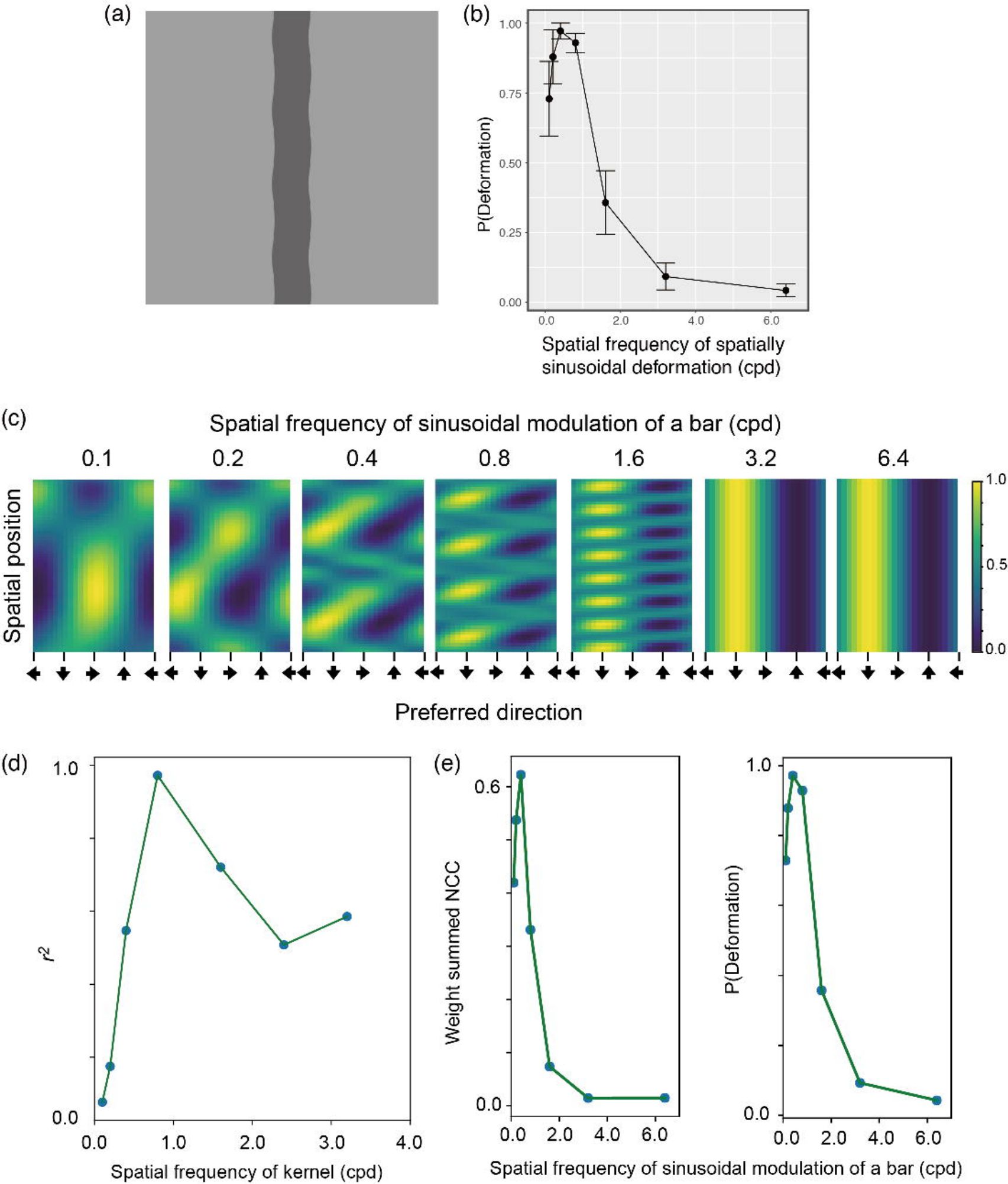


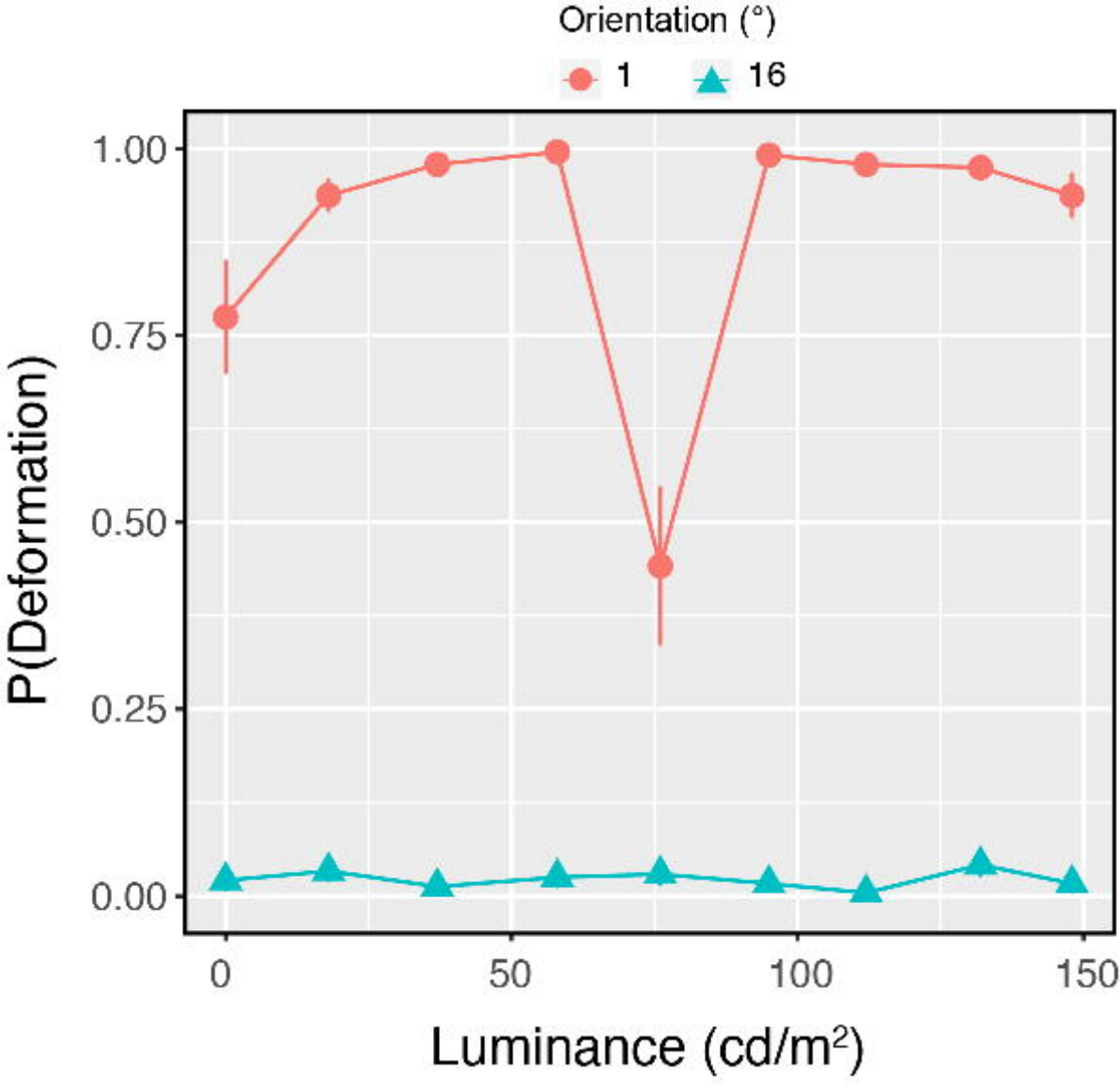
Figure 5. Experiment 3 results. Error bars denote standard errors of the mean (N = 12).


(a)
nth frame
n+20th frame
n+40th frame
Stimulus
Time
(b)
nth frame
n+20th frame
n+40th frame
Appearance
Time
(c)
Spatial frequency
of background grating (cpd)
6.4
12.9
25.8
P(Deformation)
1.00
0.75
0.50
0.25
0.00
0
5
10
15
Orientation (°)


(a)
(b)
Spatiotemporal energy calculation
Half-wave rectification
Normalization
Simulated V1 responses
Spatial pooling
Weight summation of V1 responses in an opponent fashion
Half-wave rectification
Normalization
Simulated MT responses
(c)
Simulated V1 responses
Spatial position
1.0
0.5
0.0
0
180
360
Preferred orientation (°)
Gain
1
-1
Weight summation
for each preferred direction
Simulated MT responses
Spatial position
1.0
0.5
0.0
Preferred direction
(d)
Orientation (°)
0.5
1.0
2.0
4.0
8.0
12.0
16.0
Spatial frequency of background grating
6.4 cpd
12.9cpd
25.8 cpd
Spatial position
1.0
0.5
0.0
Preferred direction


(a)
Spatial position
Preferred direction
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
(b)
r^2
0.8
0.4
0.0
0.0
1.0
2.0
3.0
4.0
Spatial frequency of kernel (cpd)
(c)
r^2 = 0.92
Weight summed NCC
0.10
0.02
P(Deformation)
1.0
0.0
0.0
15.0
0.0
15.0
Background orientation (°)


(a)
(b)
P(Deformation)
1.00
0.75
0.50
0.25
0.00
0.0
2.0
4.0
6.0
Spatial frequency of spatially sinusoidal deformation (cpd)
(c)
Spatial frequency of sinusoidal modulation of a bar (cpd)
0.1
0.2
0.4
0.8
1.6
3.2
6.4
Spatial position
1.0
0.5
0.0
Preferred direction
(d)
1.0
0.0
r^2
0.0
1.0
2.0
3.0
4.0
Spatial frequency of kernel (cpd)
(e)
0.6
0.0
Weight summed NCC
1.0
0.0
P(Deformation)
0.0
2.0
4.0
6.0
Spatial frequency of sinusoidal modulation of a bar (cpd)


Orientation (°)
1
16
1.00
0.75
0.50
0.25
0.00
P(Deformation)
0
50
100
150
Luminance (cd/m²)