scientific reports

OPEN

Response-locked classification image analysis of perceptual decision making in contrast detection

Hironori Maruyama, Natsuki Ueno & Isamu Motoyoshi✉

In many situations, humans make decisions based on serially sampled information through the observation of visual stimuli. To quantify the critical information used by the observer in such dynamic decision making, we here applied a classification image (CI) analysis locked to the observer's reaction time (RT) in a simple detection task for a luminance target that gradually appeared in dynamic noise. We found that the response-locked CI shows a spatiotemporally biphasic weighting profile that peaked about 300 ms before the response, but this profile substantially varied depending on RT; positive weights dominated at short RTs and negative weights at long RTs. We show that these diverse results are explained by a simple perceptual decision mechanism that accumulates the output of the perceptual process as modelled by a spatiotemporal contrast detector. We discuss possible applications and the limitations of the response-locked CI analysis.

While humans and animals can immediately recognize objects and scenes at a glance[1–3], in many situations they assemble information in a sequence to take more appropriate decisions[4–6]. In cognitive psychology, such dynamic information processing has been investigated mainly by measuring the reaction times and correct rates of observers. However, the reaction time alone is not powerful enough to reveal what kind of information in the stimuli led the observers to make a decision at that moment in time, unless data obtained under various conditions are compared.

In visual neuroscience, reverse correlation analysis is widely applied to reveal the information in stimuli that determines the system responses[7,8]. This analysis has been applied not only to the responses of cortical neurons[9], but also to the analysis of the behavioral responses of human observers[10,11]. The classification image (CI) method, one such technique, visualizes what information in the stimuli observers consider important for a given perceptual judgement[12–14]. In typical experiments, the observer's responses to a visual target embedded in white noise are collected, and the information in the stimulus that affected the observer's response is mapped out by analyzing the correlation between the noise and the response in each trial. The CI technique has been used to reveal the spatial distribution of information, or perceptive field, that determines the observer's judgments for a variety of visual tasks[15–17].

With dynamic stimuli, the CI method can also yield spatiotemporal perceptive fields[18–20]. Neri and Heeger[18] analyzed the correlation between spatiotemporal noise and responses in each trial in a contrast detection task for luminance bars that slowly appear in dynamic noise. They found CI profiles with biphasic weights in time and space, similar to the spatiotemporal impulse response of the early visual system. Recently, similar psychophysical reverse correlation with dynamic stimuli has been applied to the judgement on the average of time-varying visual information to investigate the mechanisms of perceptual decision making[21–28].

In the aforementioned studies, however, observers made decisions after the visual stimuli had been shown. Such a judgment, which is usually based on visual working memory, may be different from on-the-fly judgments that we make in the real world. To clarify when observers make decisions and what information observers rely on to make decisions during observation, one can analyze correlations at each time point of the stimulus locked to the reaction time of the observer during the presentation of the stimulus rather than the stimulus onset. This response-locked reverse correlation has been employed in several studies[10,29–31]. For example, Caspi et al.[30] examined visual features that trigger saccadic eye movements by analyzing the noise at time points locked to the onset of the saccade while the observers views a multi-element display. Okazawa et al.[31] adopted a reverse

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

---
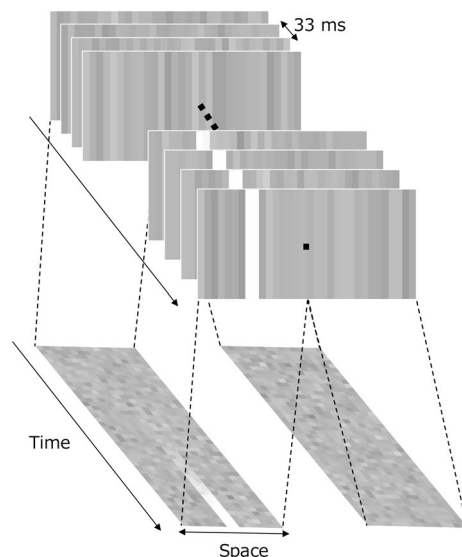
Figure 1. Schematic diagram of the visual stimuli used in the experiment. The bright bar that appears on the left is the target.

correlation analysis locked to button-press responses to stochastic motion to explore the properties of global-motion detectors and decision-making mechanisms.

In the present study, we applied the response-locked CI analysis to the most basic visual task, luminance contrast detection. Specifically, we used stimuli similar to those used by Neri and Heeger[18] to measure responses and reaction times for target stimuli that emerge slowly in dynamic noise, and we then analyzed the correlation between the noise and response at each time point backward, locked to the observer's reaction time. This protocol allowed us to examine what signals and what point in the stimulus determined the observer's decision about the target and the observer's reaction time. The results revealed spatiotemporally biphasic CIs similar to those reported by Neri and Heeger[18]. On the other hand, we also found that the profile of the CI substantially varied depending on the response time of the observer in a way that was unpredictable from the response properties of the early visual system. These apparently complicated results, however, were quantitatively described by a simple computational model incorporating a perceptual process approximated by a spatiotemporal filter and a decision process (drift–diffusion) that accumulates its output[32,43].

Methods

Observers.
Five naïves and two of the authors (average age: 22.8 years) with corrected-to-normal vision participated in the experiment. All experiments were conducted with permission from the Ethics Committee of the University of Tokyo. Observers gave written informed consent. The study followed the Declaration of Helsinki guidelines.

Apparatus.
Visual stimuli were displayed on a gamma-corrected LCD monitor (BENQ XL2735) controlled by a PC. The refresh rate was 60 Hz, and the pixel resolution was 0.04 deg/pixel at the viewing distance of 50 cm that we used. The mean luminance of the uniform background was 88.9 cd/m^2. All experiments were conducted in a dark room.

Stimuli.
The visual stimulus was square dynamic one-dimensional noise (4.8 × 4.8 deg) comprising 16 vertical bars with a width of 0.3 deg (Fig. 1). The contrast ($C_{noise}(t)$) of each bar was switched at a frame rate of 30 Hz according to Gaussian noise with an RMS contrast of 0.1. The total duration was 8000 ms. Two independent 1D-noise fields were presented adjacent to the fixation point.

$$\log C_{target}(t) = min(\alpha t - 3, 0)$$

(1)

Here, t is the frame number (33 ms per unit) from stimulus onset. α is the rate of increase, which was set at three levels: 0.05, 0.1, and 0.2. The contrast of each bar was clipped in the range of −1 to +1. The two fields, with and without the target signal, were transformed into luminance images using the relation $L(t) = L_{mean}(1 + C(t))$, where L_{mean} is the mean luminance of the uniform background (88.9 cd/m^2).

Procedure.
In each trial, observers viewed the stimulus at a fixation point binocularly and indicated by pressing a button whether the target appeared in the left or right noise field as quickly as possible. If an observer's response exceeded the deadline (8000 ms) or was an error, auditory feedback was given, and the data recorded in that trial were excluded from the analyses. The next trial started no less than 0.5 s after the observer's response. The average error rates were 0.03, 0.02, and 0.02 for contrast increases (α) of 0.05, 0.1, and 0.2, respectively. In
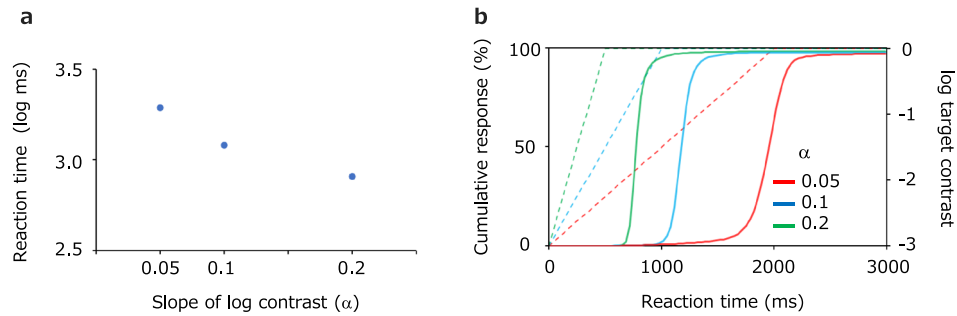
Figure 2. (**a**) Average logarithmic reaction time as a function of the contrast increase rate. Error bars represent ± s.e. across observers (invisibly small). (**b**) Cumulative histogram of reaction times of all observers (solid line). Dashed lines represents the logarithmic contrast of the target stimulus over time. Red, blue, and green lines represent a contrast increase of 0.05, 0.1, and 0.2, respectively.
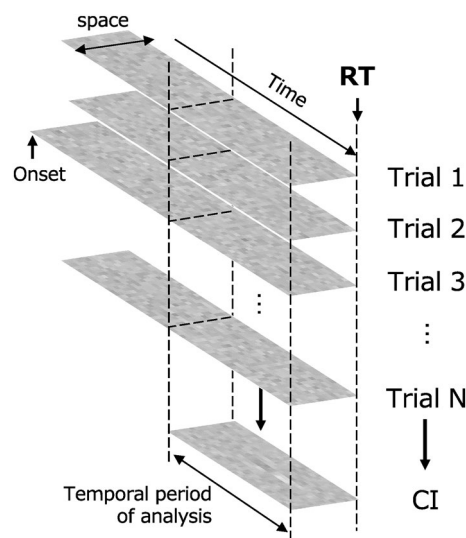
Figure 3. Reverse correlation analysis locked to the reaction. The classification image (CI) was calculated for each bar contrast at each time from the reaction time.

each trial, the contrast values of all individual bars ($C_{noise}(x,t)$), the observer's response (left, right), and the reaction time were recorded. Each session of the experiment comprised 150 trials for a single condition. For each observer, sessions were repeated until at least 1200 trials were conducted for each condition.

Ethics approval. All experimental protocols were approved by Ethics Committee of the University of Tokyo.

Approval for human experiments. Written informed consent was obtained from all participants/observers.

Results
Reaction time. Figure 2a shows the average logarithmic reaction time of the observer, plotted as a function of the contrast increase (α). On a linear scale, the reaction times were 1949 ms (s.e. = 36.1), 1207 ms (s.e. = 20.7), and 818 ms (s.e. = 22.7) for a contrast increase (α) of 0.05, 0.01 and 0.2, respectively. Figure 2b is a cumulative histogram of the reaction times of all observers. Figure 2b shows that a slower rate of increase in the contrast resulted in a longer average reaction time. One-way repeated-measure ANOVA on the average reaction time with increasing contrast showed a significant effect of the rate of contrast increase on the reaction time (F(2, 12) = 1574.6, p < 0.001).

Reverse correlation analysis locked to the response time. We conducted a reverse correlation analysis between the contrast of each bar and the observer's response (left, right) at each time (t) back from the reaction time to characterize the noise common to the time before the reaction. Figure 3 is a diagram of the analysis. As in Neri and Heeger[18], $\mu_1(x,t)$ is the mean of the noise contrast in the region where the observer responded that
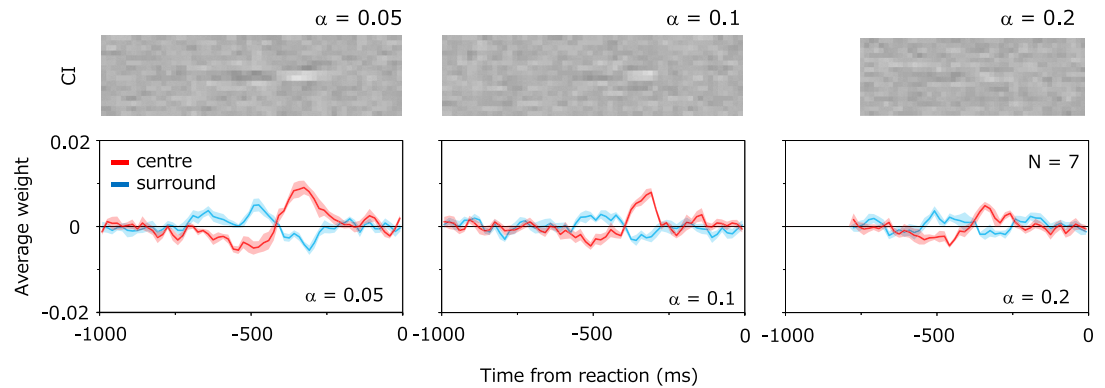
Figure 4. Results of response-locked reverse correlation analysis. The upper panels show the classification image (CI). The vertical axis represents the space and the horizontal axis represents the time from the reaction time. Each pixel represents a positive (bright) or negative (dark) weight. The lower panels show the average of the weights in the central two bars (red curves) and the average of the weights in the two adjacent bars (blue curves). The vertical axis represents the weights and the horizontal axis represents the time from the reaction time. Error bars represent ± s.e. across observers. Individual panels show the results for a contrast increase (α) of 0.05, 0.1, or 0.2.

the target was present and $\mu_0(x,t)$ is the mean of the noise contrast in the region where the observer responded that the target was not present. Since the reaction time varied across trials, the number of trials used to calculate the mean at each time point was not constant and tended to decrease as time increased backward from the response. The results were calculated as follows.

$$MeanKernel(x, t) = \mu_1(x, t) - \mu_0(x, t) \tag{2}$$

here Mean Kernel refers to the effect of the noise contrast on the response.

The upper panels in Fig. 4 show the classification image (i.e., Mean Kernel) obtained in the reverse correlation analysis locked to the reaction time. The horizontal axis represents the time (t) back from the reaction time, and the vertical axis represents the spatial position of each bar (x). In comparison with the grey background, the brighter points represent positive weights and the darker points represent negative weights. Individual panels show results for a contrast increase (α) of 0.05, 0.1, or 0.2. The lower panels show the mean of the weights of the two central bars (red) and the mean of the weights of the two adjacent bars (blue) in the CI. The vertical axis represents the weight and the horizontal axis represents the time from the reaction time. We refer to the plots as impact curves.

The above plots show characteristic temporal and spatial variations in the weights before the target detection response. At the center of the stimulus where the target appeared, a large positive weight was found about 300 ms before the response, and a negative weight was found about 500 ms before the response. For the spatial variation, we find that the weights around the target are reversed from the center. This means that the decrease in luminance in the bars adjacent to the target are also useful for detecting the target. It is also found that the absolute magnitude of the weights tends to decrease as the contrast increase rate increases. We also conducted the same analysis for contrast variance, as was done in a previous study[18,44], but found no clear CI profile. This discrepancy could be due to the fact that the target appeared gradually in the present study, whereas the target was abruptly flashed in the previous study.

Relationship between the RT and CI. As shown in Fig. 2, the reaction time of the observer varied even under the same conditions. Each individual observer responded quickly in some trials and took a long time in others. Taking advantage of this fact, we investigated if and how the CI changes with the reaction time. To this end, we divided the observer's data into 50% trials with short reaction times, 50% trials with intermediate reaction times, and 50% trials with long reaction times for each condition of the contrast increase rate, and carried out the reverse correlation analysis for each group.

Figure 5 shows the CIs and impact curves obtained for each reaction time group. The results were surprisingly different across the groups. The positive weights about 300 ms before the reaction are larger for the shorter reaction time group. Conversely, the negative weights 500 ms before the reaction are larger for the longer reaction time group. This tendency is constant regardless of the contrast increase rate (α). The results indicate that the spatiotemporal profile of the weights of information correlated with the response is remarkably different depending on the reaction time.

Discussion

The present study examined the information utilization strategy adopted in dynamic decision making during stimulus observation in a simple contrast detection task. Applying the classification image method, we calculated the weights of the embedded noise at each time point retrospectively from the reaction time for the target. The resulting CIs indicate that observers responded by utilizing the biphasic luminance change and the central
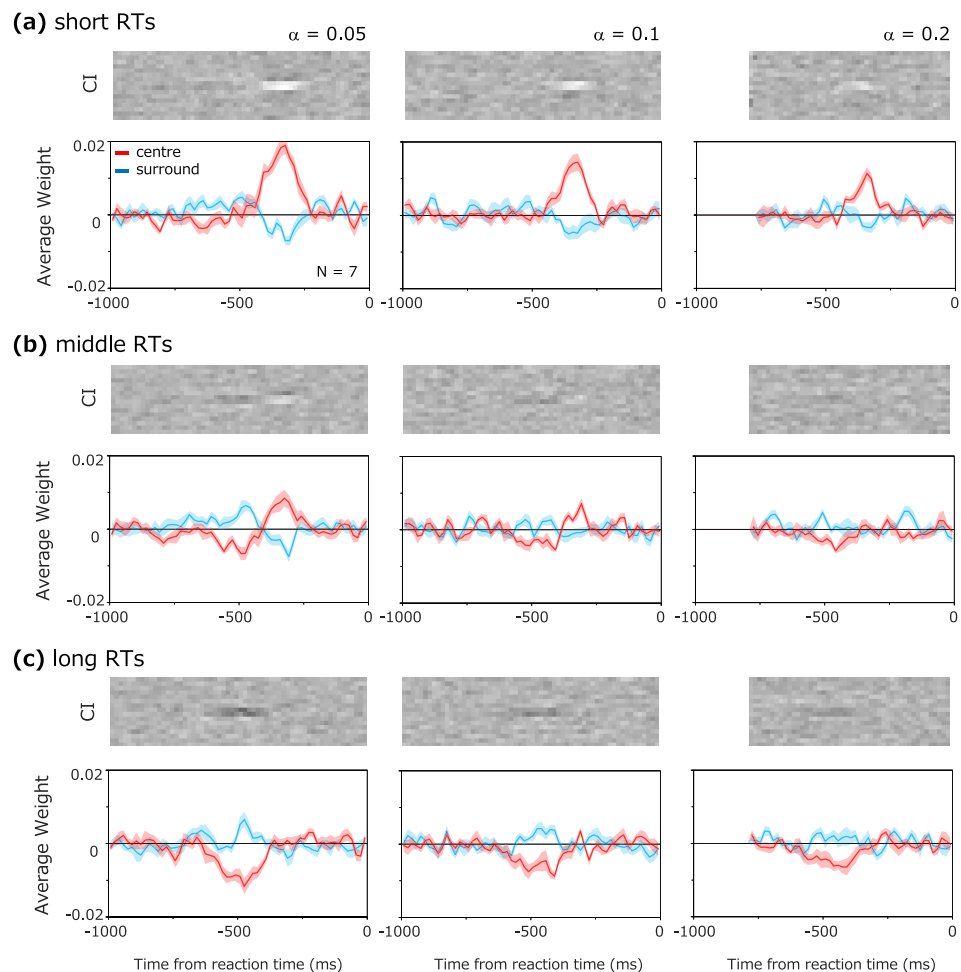
Figure 5. Results for each reaction time group. Results for (**a**) short reaction times, (**b**) intermediate reaction times, and (**c**) long reaction times. Upper panels show the CIs and lower panels show the impact curves.

antagonistic spatial contrast before the response. In addition, we found that these spatiotemporal profiles of CIs varied significantly depending on the reaction time.

The complex diversity of the results depending on the reaction time appears to be difficult to understand intuitively. This may seem to indicate that the observers were so flexible that they use different strategies for utilizing information depending on whether they could respond quickly to the target or not. However, we should first consider a simple explanation—that the results are a natural consequence of an interplay between the sensory system and the decision process. Therefore, we tested a simple model consisting of the early visual process (linear filtering model) and the perceptual decision process (drift–diffusion model). As a result, we found that this conservative model with a fixed set of parameters successfully duplicated the human data for all conditions and RT ranges.

Computational model. Figure 6 shows an outline of the model, which is inspired by a previous study on spatiotemporal ensemble perception (Yashiro et al.[28]). The model compares the spatially summarized outputs of the perceptual process, which is approximated using linear spatiotemporal filters, between the two regions. The decision process accumulates the differential signal between the two regions as sensory evidence over time and makes a decision when the evidence reaches a given boundary. The basic structure of the perceptual process follows that of a previous CI study (Neri and Hereger[18]), and the computation of the decision making follows traditional drift–diffusion model (DDM) for a two-alternative forced-choice task[32–35]. Figure 6 shows each step of the process graphically for the case that a target appears on the left. The calculation of each step is described in detail below.

Following previous studies[18], the perceptual system is approximated as a space–time separable linear filter, $F_{st}(x,t)$, as follows.

$$F_{st}(x,t) = F_s(x) \cdot F_t(t) \tag{3}$$

Here $F_s(x)$ is the spatial filter and $F_t(t)$ is the temporal filter. The spatial filter $F_s(x)$ is given as a DoG function, which has been widely used as a first-order approximation for contrast detectors in the visual system[36]
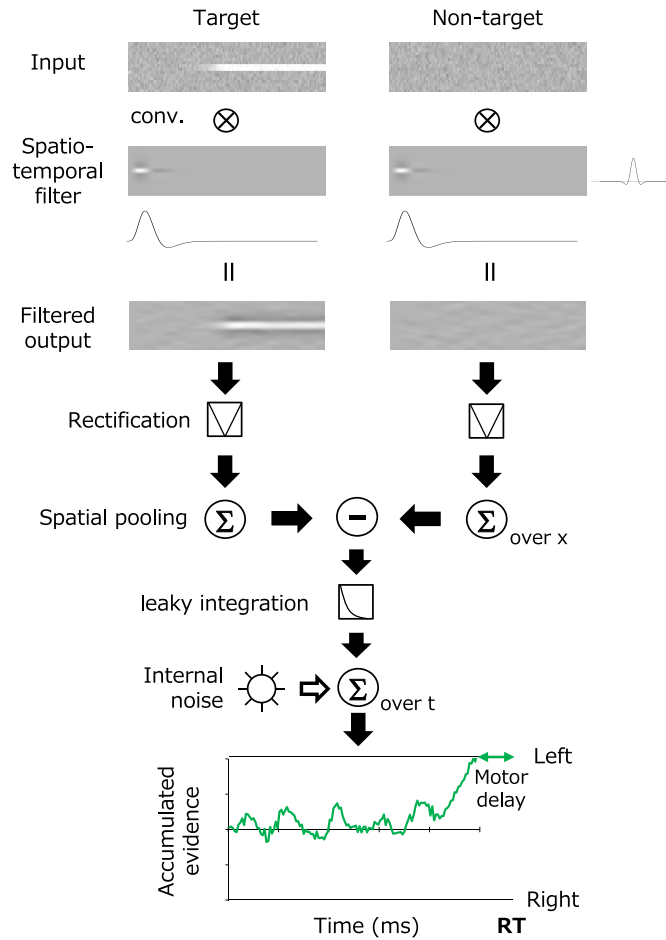
Figure 6. Schematic diagram of a model based on spatiotemporal filtering and the accumulation of sensory evidence.

(Enroth-Cugell and Robson 1966). Note that t is the frame number (33 ms per unit) and x is the pixel (0.04 deg per unit).

$$F_s(x) = exp\left(\frac{-x^2}{2\sigma_c^2}\right) - \frac{\sigma_c^2}{\sigma_s^2} \cdot exp\left(\frac{-x^2}{2\sigma_s^2}\right) \tag{4}$$

Here σ_c is the standard deviation for the central region and σ_s is the standard deviation for the adjacent region. The temporal filter $F_t(t)$ is given as the following biphasic function[37,38].

$$F_t(t) = \left(\frac{1}{n!} - B\frac{(t/\tau)^2}{(n+2)!}\right) \cdot (t/\tau)^n \exp(-t/\tau) \tag{5}$$

Here n is the number of stages in the time integrator, t is the transient factor, and B is a parameter that defines the amplitude ratio of the positive and negative phases.

The response of the perceptual system was obtained by convolving the above spatiotemporal filter $F_{st}(x,t)$ with the stimulus input $I(x,t)$.

$$R(x,t) = I(x,t) * F_{st}(x,t) \tag{6}$$

Decisions concerning whether the target presented in the left or right region were made by comparing the spatial sum of the absolute values of the responses in each region between the left and right. Thus, the model observer continually monitored the difference $\Delta R(t)$ between the left and right responses at time t from the stimulus onset. Here, $\Delta R(t)$ is regarded as the sensory evidence at time t in the decision-making model.

$$\Delta R(t) = \sum_x |R_{\text{left}}(x,t)| - \sum_x |R_{\text{right}}(x,t)| \tag{7}$$
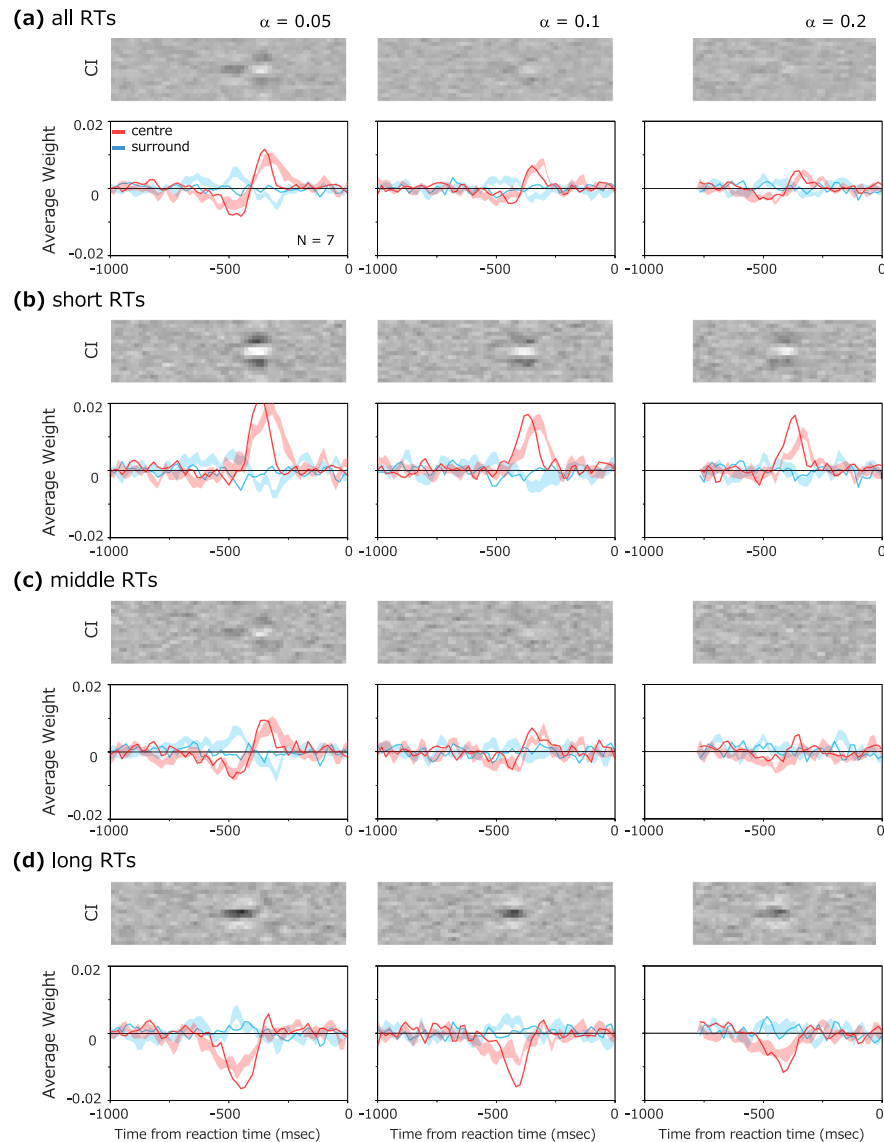
Figure 7. Results of the model simulations. (**a**) Impact curves and CIs derived from the data for all RTs. (**b**–**d**) Impact curves and CIs when divided into reaction time groups, with solid lines representing results for model observers and light-colored bands representing the ± 1 s.e. range for human data.

Decisions for targets are based on evidence accumulated over time. However, a number of decision-making studies suggest that sensory evidence decays with time; that is, the evidence weakens as it ages[22,39,40]. This property is practically described as a leaky temporal integration, and it is potentially a product of the adaptive gain control of evidence signals[41,42]. According to these findings, the present modeling assumes that the cumulative evidence $S(T)$ at time T is given by the following equation, which approximates the noisy leaky integration of $\Delta R(t)$.

$$S(T) = \sum_{t=1}^{T} (\gamma^{(T-t)} \Delta R(t) + \epsilon_t) \tag{8}$$

Here g is the time constant of evidence integration and ε_t is the internal noise following a normal distribution. The model observer makes a decision about whether the target is on the left or right when $S(T)$ exceeds a certain decision boundary, that is, b or $-b$, respectively. The observer was assumed to execute a manual response after a constant motor delay of 250 ms from T.

In this modeling, the perceptual process part has five parameters: the standard deviations of the spatial filter (σ_c and σ_s, in pixels), number of biphasic temporal filter integration stages (n), time constant (τ, in frames), and ratio of positive to negative phases (B). The decision-making process part has three parameters: the decision boundary (b), internal noise (ε_t), and time constant for evidence reduction (γ, in frames).

Model simulation. We analyzed the CI and impact curves of the model observer using the image input data that were presented to each observer in the experiment. In the simulations, for all data in the condition of $\alpha = 0.05$, the model parameters were optimized for each observer to minimize the squared error between the impact curve obtained for the model observer and that of the human observer. To achieve the steady fitting, only the number of integration steps of the biphasic temporal filter (n) was fixed to 5, for all model observers.

Figure 7 shows the simulation results. The thick impact curve represents the average of results obtained for the optimized model for each observer, and the light-colored bands represent the ± 1 se range of the average for the human observer data. Estimated parameters and the s.e. across model observers were $[\sigma_c, \sigma_s, B, \tau, \gamma, b, \varepsilon_t] = [4.15, 17.4, 0.47, 1.360, 0.046, 496.5, 49.6]$ (s.e. = 0.18, 1.67, 0.07, 0.158, 0.002, 3.58, 2.64). For all values of the contrast increase (α), we find that the model successfully duplicated both the CI and the impact curve of the observers. For the three different reaction time groups (Fig. 7b–d), the model duplicated the observers' data, reflecting the characteristic differences of RT-dependent CI and impact curves. The root-mean-square error (i.e., difference) between the fitted models and the model observers' data averaged over all observers was 0.005 (s.e. = 0.0001). The difference in the behavior of the impact curves for each reaction time group can be intuitively explained as follows. In the short reaction times, the positive weights are larger about 300 ms before the response because there is not enough time for the negative part of the biphasic temporal impulse response to activate. On the other hand, in the long reaction times, the negative weight of the biphasic temporal impulse response acts and becomes visible, but the positive weight is thought to occur because the effect decreases via leaky integration.

To investigate the importance of the functional processes assumed in the model in Fig. 6, we simulated the model without some of the functional processes. We found that (1) the unique shape of the observers' CIs and impact curves could not be simulated if even one of the parameters of the spatiotemporal filter was omitted and (2) without the leaky integration property being assumed in the decision making, the effect of the early stage of stimulus presentation did not decrease even after a long observation in some RT ranges. On the other hand, modifying the model to accumulate the responses in each domain separately as two pieces of evidence and then calculate those differences, instead of accumulating the differences in responses between the two domains as evidence, did not change the behavior of the model, because the model essentially accumulates evidence linearly[43].

The present results support the idea that on-the-view behavioral responses to visual stimuli can be explained by a simple combination of the conventional perceptual model and the standard perceptual decision-making model. This finding may allow us to perform a response-locked reverse correlation analysis of human responses to sensory stimuli during observation, rather than after observation, to explore the characteristics and strategies of human information use in various cognitive tasks. In further investigations, a similar framework may be used to understand the mechanisms for attentional selection and for high-level visual cognition. The present computational model can be used as a baseline account in these investigations.

It should be noted that psychophysical analysis cannot reliably separate the properties of decision making from the low-level perceptual process[31]. Although one can partially overcome this limitation by making full use of various aspects of data, such as by dividing the data into different RT ranges as in the present study, it is difficult to distinguish between some properties such as the latency of the perceptual sensors and the motor delay in the decision process.

References

1. Thorpe, S., Fize, D. & Marlot, C. Speed of processing in the human visual system. *Nature* **381**(6582), 520–522 (1996).
2. Motoyoshi, I., Nishida, S. Y., Sharan, L. & Adelson, E. H. Image statistics and the perception of surface qualities. *Nature* **447**(7141), 206–209 (2007).
3. Whitney, D. & Yamanashi Leib, A. Ensemble perception. *Annu. Rev. Psychol.* **69**, 105–129 (2018).
4. Bergen, J. R. & Julesz, B. Parallel versus serial processing in rapid pattern discrimination. *Nature* **303**(5919), 696–698 (1983).
5. Treisman, A. M. & Gelade, G. A feature-integration theory of attention. *Cogn. Psychol.* **12**(1), 97–136 (1980).
6. Wolfe, J. M. (2015). Visual search.
7. Neri, P. & Levi, D. M. Receptive versus perceptive fields from the reverse-correlation viewpoint. *Vis. Res.* **46**(16), 2465–2474 (2006).
8. Neri, P., Parker, A. J. & Blakemore, C. Probing the human stereoscopic system with reverse correlation. *Nature* **401**(6754), 695–698 (1999).
9. DeAngelis, G. C., Ohzawa, I. & Freeman, R. D. Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. II. Linearity of temporal and spatial summation. *J. Neurophysiol.* **69**(4), 1118–1135 (1993).
10. Ringach, D. L. Tuning of orientation detectors in human vision. *Vis. Res.* **38**(7), 963–972 (1998).
11. Solomon, J. A. Noise reveals visual mechanisms of detection and discrimination. *J. Vis.* **2**(1), 7–7 (2002).
12. Ahumada, A. J. Jr. Perceptual classification images from Vernier acuity masked by noise. *Perception* **25**, 2 (1996).
13. Abbey, C. K. & Eckstein, M. P. Classification image analysis: Estimation and statistical inference for two-alternative forced-choice experiments. *J. Vis.* **2**(1), 5–5 (2002).
14. Abbey, C. K. & Eckstein, M. P. Classification images for simple detection and discrimination tasks in correlated noise. *JOSA A* **24**(12), B110–B124 (2007).
15. Eckstein, M. P., Shimozaki, S. S. & Abbey, C. K. The footprints of visual attention in the Posner cueing paradigm revealed by classification images The Bayesian Observer. *Cueing* **20**, 25–45 (2002).
16. Gold, J. M., Murray, R. F., Bennett, P. J. & Sekuler, A. B. Deriving behavioural receptive fields for visually completed contours. *Curr. Biol.* **10**(11), 663–666 (2000).
17. Rajashekar, U., Bovik, A. C. & Cormack, L. K. Visual search in noise: Revealing the influence of structural cues by gaze-contingent classification image analysis. *J. Vis.* **6**(4), 7–7 (2006).
18. Neri, P. & Heeger, D. J. Spatiotemporal mechanisms for detecting and identifying image features in human vision. *Nat. Neurosci.* **5**(8), 812–816 (2002).
19. Mareschal, I., Dakin, S. C. & Bex, P. J. Dynamic properties of orientation discrimination assessed by using classification images. *Proc. Natl. Acad. Sci.* **103**(13), 5131–5136 (2006).
20. Neri, P. & Levi, D. Temporal dynamics of directional selectivity in human vision. *J. Vis.* **8**(1), 22–22 (2008).

21. De Gardelle, V. & Summerfield, C. Robust averaging during perceptual judgment. *Proc. Natl. Acad. Sci.* **108**, 13341–13346 (2011).
22. Hanks, T. D. & Summerfield, C. Perceptual decision making in rodents, monkeys, and humans. *Neuron* **93**, 15–31 (2017).
23. Li, V., Castañón, S. H., Solomon, J. A., Vandormael, H. & Summerfield, C. Robust averaging protects decisions from noise in neural computations. *PLoS Comput. Biol.* **20**, 1–19 (2017).
24. Vandormael, H., Herce Castañón, S., Balaguer, J., Li, V. & Summerfield, C. Robust sampling of decision information during perceptual choice. *Proc. Natl. Acad. Sci.* **114**, 2771–2776 (2017).
25. Sato, H. & Motoyoshi, I. Distinct strategies for estimating the temporal average of numerical and perceptual information. *Vis. Res.* **174**, 41–49 (2020).
26. Summerfield, C. & Tsetsos, K. Do humans make good decisions?. *Trends Cogn. Sci.* **19**(1), 27–34 (2015).
27. Yashiro, R. & Motoyoshi, I. Peak-at-end rule: Adaptive mechanism predicts time-dependent decision weighting. *Sci. Rep.* **10**(1), 1–8 (2020).
28. 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).
29. Busse, L., Katzner, S., Tillmann, C. & Treue, S. Effects of attention on perceptual direction tuning curves in the human visual system. *J. Vis.* **8**(9), 2–2 (2008).
30. Caspi, A., Beutter, B. R. & Eckstein, M. P. The time course of visual information accrual guiding eye movement decisions. *Proc. Natl. Acad. Sci.* **101**(35), 13086–13090 (2004).
31. Okazawa, G., Sha, L., Purcell, B. A. & Kiani, R. Psychophysical reverse correlation reflects both sensory and decision-making processes. *Nat. Commun.* **9**(1), 1–16 (2018).
32. Ratcliff, R. & McKoon, G. The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Comput.* **20**(4), 873–922 (2008).
33. Gold, J. I. & Shadlen, M. N. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
34. Kiani, R., Hanks, T. D. & Shadlen, M. L. Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J. Neurosci.* **28**(12), 3017–3029 (2008).
35. Huk, A. C. & Shadlen, M. N. Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *J. Neurosci.* **25**, 10420–10436 (2005).
36. Enroth-Cugell, C. & Robson, J. G. The contrast sensitivity of retinal ganglion cells of the cat. *J. Physiol.* **187**(3), 517–552 (1966).
37. Watson, A. B. Temporal sensitivity. *Handb. Percept. Human Perform.* **1**(6), 1–43 (1986).
38. Adelson, E. H. & Bergen, J. R. Spatiotemporal energy models for the perception of motion. *Josa a* **2**(2), 284–299 (1985).
39. Usher, M. & McClelland, J. L. The time course of perceptual choice: The leaky, competing accumulator model. *Psychol. Rev.* **108**, 550–592 (2001).
40. Yashiro, R., Sato, H. & Motoyoshi, I. Prospective decision making for randomly moving visual stimuli. *Sci. Rep.* **9**, 3809 (2019).
41. Cheadle, S. *et al.* Adaptive gain control during human perceptual choice. *Neuron* **81**(6), 1429–1441 (2014).
42. Li, V., Michael, E., Balaguer, J., Castañón, S. H. & Summerfield, C. Gain control explains the effect of distraction in human perceptual, cognitive, and economic decision making. *Proc. Natl. Acad. Sci.* **115**(38), E8825–E8834 (2018).
43. Bogacz, R., Brown, E., Moehlis, J., Holmes, P. & Cohen, J. D. The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* **113**(4), 700 (2006).
44. Pascucci, D., Megna, N., Panichi, M. & Baldassi, S. Acoustic cues to visual detection: A classification image study. *J. Vis.* **11**(6), 7–7 (2011).

Acknowledgements

Author contributions

I.M. conceived the study. H.M., N.U., and I.M. designed the experiment. H.M. and N.U. performed experiments and analyzed data. H.M., N.U., 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.