Relative spatial frequency tuning and its contrast dependency in human perception

Tomoyuki Naito

Laboratory of Cognitive and Behavioral Neuroscience, Graduate School of Medicine, Osaka University, Osaka, Japan

Naofumi Suematsu

Laboratory of Cognitive and Behavioral Neuroscience, Graduate School of Frontier Biosciences, Osaka University, Osaka, Japan

Eriko Matsumoto

Graduate School of Intercultural Studies, Kobe University, Kobe, Japan

Hiromichi Sato

Laboratory of Cognitive and Behavioral Neuroscience, Graduate School of Medicine, Osaka University, Osaka, Japan

Laboratory of Cognitive and Behavioral Neuroscience, Graduate School of Frontier Biosciences, Osaka University, Osaka, Japan

Several physiological studies in cats and monkeys have reported that the spatial frequency (SF) tuning of visual neurons varies depending on the luminance contrast and size of stimulus. However, comparatively little is known about the effect of changing the stimulus contrast and size on SF tuning in human perception. In the present study, we investigated the effects of stimulus size and luminance contrast on human SF tuning using the subspace-reverse-correlation method. Measuring SF tunings at six different stimulus sizes and three different luminance contrast conditions (90%, 10%, and 1%), we found that human perception exhibits significant stimulus-size-dependent SF tunings. At 90% and 10% contrast, participants exhibited relative SF tuning (cycles/image) rather than absolute SF tuning (cycles/8) at response peak latency. On the other hand, at 1% contrast, the magnitude of the size-dependent-peak SF shift was too small for strictly relative SF tuning. These results show that human SF tuning is not fixed, but varies depending on the stimulus size and contrast. This dependency may contribute to size-invariant object recognition within an appropriate contrast range.

Introduction

Spatial frequency (SF) tuning is a fundamental property of the visual system. It is often studied in standard psychophysical methods for human perception by modulating the contrast threshold (Anderson & Burr, 1987; Blakemore & F. W. Campbell, 1969; B. Y. F. W. Campbell & Robson, 1968). This approach, however, does not allow for analysis of the stimulus contrast effect on SF tuning, because the experimental design implicitly assumes contrast invariance. In contrast, animal studies have shown that such effects are not negligible. In monkey and cat, it has been reported that the SF tuning of visual neurons in the lateral geniculate nucleus (LGN) and primary visual cortex (V1) varies with stimulus contrast (Kimura, Shimegi, Hara, Okamoto, & Sato, 2013; Nolt, Kumbhani, & Palmer, 2004; Sceniak, Hawken, & Shapley, 2002). At the same time, these studies show great disagreement on the effects. In the macaque V1, Sceniak et al. (2002) reported that lowering the stimulus contrast decreased the SF tuning bandwidth but made no change to the peak SF of the response, whereas in the cat LGN, Nolt et al. (2004) reported that...
increasing the contrast increased the peak SF while the bandwidth remained contrast-invariant. Similarly and more recently, Kimura et al. (2013) reported that in the cat LGN the peak SF and SF bandwidth were contrast invariant.

Another factor that affects SF tuning is stimulus size. A recent study on neurons in the macaque inferior-temporal cortex (IT) reported that a population of IT neurons decreased their peak SF with increasing stimulus size (Inagaki & Fujita, 2011). That study concluded that IT neurons tuned not to absolute SF (cycles/°) but to relative SF (cycles/image). Weak but significant stimulus-size-dependent SF tuning was also reported in cat V1 (Osaki, Naito, Sadakane, Okamoto, & Sato, 2011; Teichert, Wachtler, Michler, Gail, & Eckhorn, 2007).

In humans as well, several psychophysical experiments have shown evidence for size-dependent SF tuning. For example, priming experiments that tested subjects on line drawings or gray-scale images of ordinary objects showed no significant change in the reaction time for the target detection if the line width changed with stimulus size (Biederman & Cooper, 1991; Fiser & Biederman, 1995). A perceptual learning task with gray-scale images found similar results (Furmanski & Engel, 2000). Ashida (2002) investigated the relationship between SF tuning and stimulus size using the Ouchi illusion, which produces an illusory motion and depth sensation by using a checkerboard pattern surrounded by a second, orthogonally oriented checkerboard pattern. Ashida reported that the optimal fundamental SF for this illusion is lower for a larger stimulus, affirming that human SF tuning does indeed depend on the stimulus size. Anderson and Burr (1987) showed by the motion direction detection task that different optimal SFs have different receptive field sizes.

Although these results indicate that the human SF tuning is not fixed but rather varies with stimulus size, there is, to our knowledge, no systematic research that quantitatively examines the size dependency of human SF tuning, nor any that considers effectively the contrast dependency. In the present study, we measured SF tunings in human perception at various stimulus sizes with three different luminance contrasts using a psychophysical subspace-reverse-correlation method (Ringach, 1998). This technique allowed us to efficiently measure SF tunings at various stimulus sizes and their temporal dynamics in a short time. Additionally, the method makes it possible to measure the effects of stimulus contrast on SF tuning, because it uses not a contrast threshold but a response probability to sequential stimuli. Our results show that human SF tuning depends on both stimulus size and luminance contrast. The functional significance of the dependencies will be discussed.

Methods

Participants

Five naïve, paid volunteers and two of the authors (TN and NS) participated in the experiments (N = 7; mean age ± standard deviation [SD] = 24.9 ± 2.7). All had normal or corrected-to-normal visual acuity (≥ 1.0). Each participant gave written informed consent prior to the experiments, which were performed in accordance with the Declaration of Helsinki.

Apparatus and stimuli

All stimuli were generated using custom-made MATLAB (Mathworks, USA) programs with Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and presented on a gamma-corrected CRT monitor (CPD-G500J, SONY; mean luminance, 70 cd/m²; screen size, 40 cm × 30 cm; resolution, 1280 × 960 pixels; refresh rate, 100 Hz) placed 57 cm in front of the participants’ eyes. Luminance linearity was confirmed within the experimental range. A chin rest was used to stabilize the head position. Participants’ response onset time was measured using a custom-made response box with 1 ms resolution. Eye movement was not monitored during the experiments.

Psychophysical subspace reverse correlation

We applied the subspace-reverse-correlation procedure (Kimura & Ohzawa, 2009; Nishimoto, Arai, & Ohzawa, 2005; Ringach, Sapir, & Shapley, 1997) used previously for electrophysiological experiments (Figure 1). We used the onset timing of a button-pressing response for the target detection task as the signal (Ringach, 1998). The visual stimulation and analysis were equivalent to those used in previous physiological studies (Kimura & Ohzawa, 2009; Nishimoto et al., 2005; Ringach et al., 1997). Sinusoidal gratings of six orientations (0°–150° in 30° steps), six SFs (0.1–4.0 cycles/° equally spaced on log-scale), six sizes (1°–30° in diameter equally spaced on log-scale), and four phases (0°, 90°, 180°, and 270°) were presented for 100 ms (10 video frames) in a randomized order. Stimulus sets, each of which included 864 randomized sequences (6 Orientations × 6 SFs × 6 Sizes × 4 Phases), were presented 20 or 30 times. Thus, one particular combination of orientation, SF, and stimulus size was presented 80 or 120 times to a participant. Participants were required to detect a predetermined target orientation (horizontally, 0°, or vertically, 90°, oriented gratings randomly selected for each participant) by pressing a button on a
custom-made response box that generated TTL pulses with 1 ms resolution. The onset of the participants’ response and the stimulus sequence were then cross-correlated to obtain a 2-D SF and stimulus size selectivity map (SF × SIZE Map) at each orientation and for several correlation delays (τ) (Figures 1 and 2). We calculated the SF × SIZE maps for τ values ranging from 0 to 1000 ms in 25-ms steps. From the amplitude of the SF × SIZE maps, we could extract the response probability for a particular stimulus (a combination of SF, size, and orientation) as follows:

$$P_{SF, size, orientation}(τ) = \frac{A_{SF, size, orientation}(τ)}{T_{SF, size, orientation}}.$$  \hspace{1cm} (1)

$P_{SF, size, orientation}(τ)$ represent the response probability for a stimulus with a particular combination at τ. $A_{SF, size, orientation}(τ)$ and $T_{SF, size, orientation}$ show the number of times a stimulus with a given parameter combination at τ appeared before a response was made and the total number of stimulus presentations, respectively. Similarly, we also calculated the response probability for an orientation ($P_{orientation}$) as follows:

$$P_{orientation}(τ) = \frac{A_{orientation}(τ)}{T_{orientation}}.$$ \hspace{1cm} (2)

where $A_{orientation}(τ)$ and $T_{orientation}$ show the number of times a stimulus with a given orientation at τ appeared before a response was made and the total number of stimulus presentations, respectively. All subjects observed the three contrast conditions (1%, 10%, and 90%) in a randomized order.

To make an objective criterion for significant responses, we estimated the noise level in the data by computing the SDs of the response probability that had noncausal SDs (baseline). The baseline was calculated as the mean of $P_{SF, size, orientation}$ (−150 to 0, 25-ms steps). All participants exhibited significant response probabilities (baseline + 5 SD) for stimuli with target orientations around 400–600 ms after the stimulus onset. For all response probability curves, the baseline was subtracted before the analysis.

**Absolute and relative SF tunings**

To characterize the dependence of the SF tuning on stimulus size, we fitted an SF × SIZE map with six Gaussian functions using a method similar to that of previous neurophysiological studies (Inagaki & Fujita, 2011; Priebe, Lisberger, & Movshon, 2006) for each contrast condition. Each Gaussian function characterized SF tuning for a particular stimulus size as follows:
\[ R_i(sf) = B + A_i \exp \left\{ - \left( \frac{\log_{10} sf - (\log_{10} sf_0 + SI_i)}{2\sigma_i^2} \right)^2 \right\}, \]

where \( i \) is the index \((-2, -1, 0, 1, 2, 3)\) of the stimulus size \((1.0^\circ, 2.0^\circ, 3.9^\circ, 7.7^\circ, 15.2^\circ, \text{and} 30.0^\circ)\). \( R_i(sf) \) denotes the response probability to a stimulus with an absolute SF at the corresponding size, and \( A_i \) and \( \sigma_i \) are the peak response probability and tuning width, respectively. \( B \) is the baseline response probability, \( sf_0 \) is the preferred absolute SF at a stimulus size of \( 7.7^\circ \), and \( SI \) (shift index) is the parameter that quantifies the dependency between the preferred SFs and sizes. When \( SI \) is zero, the preferred SF does not change across stimulus size (i.e., strictly tuned to absolute SF). When \( SI \) is one, the preferred SF is proportional to the stimulus size, while the preferred number of cycles, relative SF (cycles/image), does not change (i.e., strictly tuned to relative SF). \( A_i \) and \( \sigma_i \) were independent across stimulus sizes, while \( B, sf_0, \) and \( SI \) were the same for different sizes. Finally, \( B \) was subtracted from \( R_i(sf) \) to set the baseline response to zero. All parameters were estimated from the data. We used the MATLAB “lsqcurvefit” function (Mathworks, USA) to fit the SF tuning curves.

\section*{Results}

\subsection*{Probability of target detection}

Participants were required to detect targets at a predetermined orientation \((0^\circ \text{ or} 90^\circ)\). Using the psychophysical subspace-reverse-correlation method, we measured \( SF \times SIZE \) maps and the corresponding temporal dynamics (see Methods) for each participant. Figure 2 shows representative \( SF \times SIZE \) maps for participant NS at 90\% contrast. The stimulus-specific increment of the response probability (Equation 1) was observed (indicated by the reddish color) at the target orientation, 90\° in this case, at 400 and 600 ms after the stimulus onset (Figure 2, \( \tau = 400 \text{ and} 600 \text{ ms})\). There was a clear SF and size dependency in the response probability and interaction between the two. In contrast, at nontarget orientations, \( SF \times SIZE \) maps did not show discriminable changes in the response probability at any \( \tau \), suggesting that participant NS clearly detected the target gratings and that the target detection probability depended on both the SF and size of the stimuli. We confirmed that all participants exhibited the orientation detection, which was well tuned to the predetermined target orientation for each participant at all contrast conditions.

\subsection*{Effect of luminance contrast on target detection}

We examined the time course and contrast dependency of the response probability to the target and nontarget orientation. For this purpose, we calculated the response probability along the response time course as the number of responses to the target or nontarget orientations divided by the total number of stimulus presentations of each orientation for all combinations of the other parameters (SFs, sizes, and phases) for each contrast (Equation 2). Figures 3A and 3B show temporal profiles of the response probabilities at target orientation 0\° and nontarget orientation 90\° for participant SH, respectively. The red, green, and blue curves indicate the luminance contrasts of the stimulus (90\%, 10\%, and 1\%). As described above, at target orientation 0\°, participant SH exhibited high response probability from \( \tau = \sim 400 \text{ to} \sim 600 \text{ ms} \). In contrast, at orthogonal orientation to the target (90\°), the response probability was slightly suppressed below the baseline, especially at 90\% and 10\% contrasts, which is consistent with a previous study (Ringach, 1998).

To quantitatively evaluate the temporal dynamics of the response probability, we estimated response peak, onset, and offset times from the raw response probability curve at the target orientation. We defined the response peak time as \( \tau \) at maximum response probability; response onset time as \( \tau \) at which mean response probability first exceeded 50\% that at peak response time; and response offset time as the response probability first falling below 50\% after peak response time. Figure 3C shows the mean response onset, peak, and offset times at the three contrast conditions. The three parameters all significantly increased as the stimulus contrast decreased, one-way repeated analysis of variance (ANOVA), onset, \( F(2, 12) = 12.9, p = 0.001, \eta_p^2 = 0.996 \); peak, \( F(2, 12) = 15.7, p = 0.0004, \eta_p^2 = 0.724 \); offset, \( F(2, 12) = 8.15, p = 0.006, \eta_p^2 = 0.576 \), suggesting significant effects of luminance contrast on the target detection latency.

Figure 3D shows mean response probabilities at response peak time (maximum response probability) under the three contrast conditions. The maximum response probability significantly decreased as stimulus contrast decreased, one-way repeated ANOVA, \( F(2, 12) = 4.87, p = 0.03, \eta_p^2 = 0.448 \), and post-hoc analysis (Tukey’s HSD, significance level, \( p < 0.05 \)) revealed that response probability at 1\% contrast was significantly lower than the other conditions.
Size dependency of SF tuning

We reconstructed absolute SF tunings for each stimulus size. Figure 4A shows an example of an SF × SIZE map from participant NS for response peak time at 90% contrast. Figure 4B shows the absolute SF tuning curves for each stimulus diameter (indicated by the different colors and symbols), which were reconstructed from Figure 4A. The optimal SF showed strong stimulus-size dependency, as the optimal SF decreased with increasing stimulus size, suggesting that SF tuning in human perception is stimulus-size dependent.

We also calculated relative SF tunings (cycles/image), which are a product of the absolute SF (cycles/image) and stimulus size (%). Figure 4C shows relative SF tunings for the stimulus sizes in Figure 4B. The relative SF that gave maximum response probability was very similar among stimulus sizes.

To quantitatively evaluate how well the obtained tuning approximated perfect relative or absolute SF tuning, we calculated SIs from the SF × SIZE maps (see Methods). When human perception is strictly tuned to absolute SF, SI is zero; when strictly tuned to relative SF, SI is one. The SI for participant NS under 90% contrast was 0.81, suggesting participant NS responded with near perfect relative SF tuning.

Figures 4D–F show an SF × SIZE map from the same participant from the mean response probability...
between response-onset and offset times (400–600 ms) at 90% contrast and the corresponding absolute and relative SF. Here, the SI for the response peak was 0.79, which is very similar to that calculated from the peak response time.

Figures 5A–C shows an SF × SIZE map from participant NS at peak latency under 1% contrast and the corresponding absolute and relative SF tunings. The peak SF was slightly lower than at high contrast (Figure 4). The stimulus-size-dependent peak SF shift (Figure 5B) was relatively small compared to that under the 90% contrast condition (Figure 4B). Therefore, relative peak SFs (Figure 5C) would appear to vary with stimulus size. Consistent with this observation, the SI for 1% contrast was 0.33, which is much closer to absolute SF tuning (SI = 0) and unlike that at 90% contrast. We also calculated the SI from the mean response probabilities at 1% contrast (Figures 5D–F). The SI was 0.39, which is very similar with that calculated from the response peak (0.33).

**Temporal dynamics of size-dependent SF tuning**

Figure 6 shows the temporal profiles of mean SIs (N = 7) at the three contrast conditions. At 90% contrast, the mean SI at the response onset, peak, and offset times were 0.62, 0.81, and 0.46, respectively, although the differences were not significant (Tukey’s HSD post-hoc analysis after two-way repeated ANOVA; onset vs. peak, p = 0.61; onset vs. offset, p = 0.88; peak vs. offset, p = 0.06). A similar tendency was also observed at 10% contrast, where the mean SIs were 0.54, 0.94, and 0.53, respectively (onset vs. peak, p = 0.03; onset vs. offset, p = 0.06).
We observed under these two contrast conditions a temporal increase in SI from the response onset time to the response peak time and a temporal decrease of SI from the response peak time to the response offset time.

At 1% contrast, however, this tendency was not seen. The mean SI at the response onset, peak, and offset times were 0.51, 0.49, and 0.35, respectively, with no significant difference in SIs (onset vs. peak, \( p = 0.99 \); onset vs. offset, \( p = 0.89 \); peak vs. offset \( p = 0.93 \)). The SF tuning in human perception was almost constant from the response onset to offset time, indicating it was far from perfect relative SF tuning.

At the response onset and offset times, SI values were very similar among the three contrasts, with no significant differences between them (Tukey’s HSD post-hoc analysis after two-way repeated ANOVA; 90% onset vs. 10% onset, \( p = 0.99 \); 90% onset vs. 1% onset, \( p = 0.98 \); 10% onset vs. 1% onset, \( p = 0.99 \); 90% offset vs. 10% offset, \( p = 0.99 \); 90% offset vs. 1% offset, \( p = 0.98 \); 10% offset vs. 1% offset, \( p = 0.81 \)). In contrast, at the response peak time, the SI at 1% contrast was lower than those at the 90% and 10% contrast conditions (90% peak vs. 10% peak, \( p = 0.36 \); 90% peak vs. 1% peak, \( p = 0.005 \); 10% peak vs. 1% peak, \( p = 0.0006 \)).

For the entire population (all participants and all conditions), SIs calculated from the mean response probability and from the response peak time showed strong and significant correlation (Pearson linear correlation coefficient, \( r = 0.81, p < 0.00001 \)). On the other hand, there were no significant correlations between SIs at the response onset or offset times with that for the mean response probability SI (Pearson’s linear correlation coefficients; onset vs. mean, \( r = 0.42, p = 0.56 \); offset vs. mean; \( r = 0.21, p = 0.36 \)).

From our results, we concluded that human SF tuning is not fixed, but varies depending on the

Figure 5. Absolute and relative SF tunings at 1% contrast. (A)–(C) SF x SIZE map of participant NS, absolute SF tunings, and relative SF tunings, respectively, at peak response time. (D)–(F) The same as A–C but calculated from the mean response (mean of 425–600 ms response probabilities). The baseline level was subtracted from each response probability. Details are same as in Figure 4.
stimulus size. This property corresponds to a transformation from more absolute to more relative SF tuning with increasing stimulus contrast. Our results therefore suggest that to achieve relative SF tuning in human perception, high luminance contrast in the stimuli is necessary.

Discussion

In the present study, we successfully measured SF tuning in human perception and the corresponding temporal dynamics using the psychophysical subspace-reverse-correlation method. Because this method does not use a contrast detection threshold as its sensitivity index, we were able to measure the effect of luminance contrast on SF tuning. It is worth mentioning that it took only 30–45 min to obtain reliable SF × SIZE maps (6 SF × 6 Size over 0–1000 ms in 25 ms steps) for each contrast, meaning that for one particular stimulus size, 5–7 min was enough to measure SF tuning.

The obtained peak SF range (for example, 3.2 cycles/° at 2.0° in diameter and 90% contrast; Figure 4) is consistent with the 2–6 cycles/° previously reported (F. W. Campbell & Robson, 1968; Patching & Jordan, 2005; Tsuruhara, Nagata, Suzuki, Inui, & Kakigi, 2013). However, having used the subspace reverse-correlation method, we found that the peak SF shifted to a lower value as the interstimulus interval became shorter (data not shown). We used a 10 Hz (10 video frames of a CRT monitor with 100 Hz refresh rate) continuous stimulus presentation for all participants, as this frequency gave a good tuning signal to noise ratio for short and efficient measurements. Gratings with longer interstimulus intervals (for example, 5 Hz; 20 video frames) resulted in slightly higher peak SF tuning (data not shown). Therefore, the peak SF in human perception may depend not only on stimulus size and contrast, but also on the temporal frequency of the stimuli, as previously suggested (Kelly & Burbeck, 1980).

Underlying mechanisms for size-dependent SF tuning

We showed that human SF tunings are significantly stimulus-size dependent, as increasing the stimulus size decreased the peak SF tuning, a relationship that corresponds to relative SF tuning. Assuming the low-level detector hypothesis, which assumes low-level target detectors have relative SF tuning, and the high-level detector hypothesis, which assumes that a high-level target detection mechanism (e.g., attention) controls the contrast gain with relative SF tuning (Di Russo, Spinelli, & Morrone, 2001), we speculate that both low and high-level detectors contributed to this phenomenon.

Another important finding from our study was that the magnitude of the size-dependent peak SF shift changed in a contrast-dependent manner: at 1% contrast, the SF tuning was intermediate, somewhere between relative and absolute SF, whereas at higher contrasts (10% and 90%), the tuning was much closer to strictly relative SF tuning. The low-level detector hypothesis explains the contrast dependency of size dependency by contrast-gain-control mechanisms in the low-level detectors. On the other hand, Di Russo et al. (2001) reported that attention increased the cortical activity for target detection at high contrast more than that at low contrast, which suggests that attention controls contrast gain differently at different contrasts. Therefore, the high-level detector hypothesis suggests that our results indicate a weak attentional effect at 1% contrast such that SF tunings at this contrast reflect a low-level detector profile of imperfect relative SF tuning.

We found that at the response onset, SF tuning behaved as an intermediate between strictly absolute and strictly relative SF tuning, but approached strictly relative SF tuning 400–600 ms later (Figure 6). Several studies have reported that rapid responses occur when the observer accumulates sufficient sensory evidence of the target presentation to make a response (Amano et al., 2006; Cook & Maunsell, 2002; Miller & Schwarz, 2006). Accordingly, at the response onset, stimuli could effectively trigger a response by imperfect relative SF tuning. In contrast, at the response peak time, the
response seems based on near-perfect relative SF tuning.

In the low-level-detector hypothesis, temporal dynamics reflect temporally varying SF tuning in the low-level detectors. Several psychophysical studies have demonstrated that a stimulus detector is composed of a small number of low-level detectors (Humanski & Wilson, 1993; H. Wilson & Bergen, 1978; H. R. Wilson, McFarlane, & Phillips, 1983). It was reported that V1 neurons exhibit coarse to fine temporal SF tuning (Bredfeldt & Ringach, 2002; Nishimoto et al., 2005), suggesting that low-level detectors show a temporal SF tuning shift. Therefore, if low-SF tuned detectors show a larger shift in peak SF than high-SF tuned detectors, an integrative high-level detector will show a temporal increase in its size dependency.

It has been suggested that attention enhances the perceptual saliency (Huang & Dobkins, 2005; Treue, 2004) and orientation discrimination performance (Carrasco, Ling, & Read, 2004). Therefore, the temporal profiles of the SF × SIZE maps we observed probably reflect the SF and size tuning of the attentional modulation, at least to some extent. According to the high-level-detector hypothesis, stimuli might be filtered by attention before passing the response threshold. Consequently, responses are enhanced to stimuli with particular SF-SIZE combinations, which may contribute to the size-dependent SF tuning of human perception.

Comparison with visual neurons

We found that peak SF was lower at low contrast (1%) than at higher contrasts (10% and 90%). This result is not consistent with previous physiological studies in V1 and the LGN (Kimura et al., 2013; Sceniak et al., 2002) in which contrast dependent peak SF shifts were negligible. Yet we also found that the tuning width was slightly narrower at 1% contrast than at 90% contrast (Figures 4 and 5), which is consistent with a previous V1 study (Sceniak et al., 2002) but not an LGN study (Kimura et al., 2013), suggesting that the correlation between SF tuning and contrast is common for cortical neurons but not subcortical ones.

Other physiological studies have reported that a subpopulation of V1 neurons and multiunit activities of V1 neurons exhibit stimulus-size dependent SF tuning by showing that increasing the stimulus diameter decreases the SF tuning peak (Osaki et al., 2011; Teichert et al., 2007). This trend is consistent with our results, especially at 1% contrast, where the size dependency is weak and far from perfect relative SF tuning.

The observed relative SF tuning in human perception is consistent with a report that used band-pass-filtered face images as stimuli and found SF tuning in awake macaque IT neurons (Inagaki & Fujita, 2011). There, most IT neurons that responded to the face images tuned not to the absolute SF but relative SF of the images. Because the IT is the highest association area in the dorsal visual stream of the cortex, it is reasonable to assume that the observed relative SF tuning in humans reflects the SF tuning property of IT neurons, at least to some extent. If so, the observed contrast dependency of relative SF tuning in the present paper indicates that the magnitude of the relative SF tuning of macaque IT neurons may depend on the stimulus contrast of the images. By combining these works with ours, we infer that the transformation of retina-based image processing (absolute SF tuning) to object based image processing (relative SF tuning) is an objective of the visual cortex.

Watson and Ahumada (2005) demonstrated that simple models could account for the visibility of a wide variety of spatial stimuli. They further showed that it was not the individual filter shape, but rather how the outputs of different filters were combined that resulted in the different contrast sensitivity filter shapes. Therefore, it is possible that the SF tuning and contrast sensitivity function of individual neurons is different from those from human psychophysical experiments.

Functional significance of size- and contrast-dependent SF tuning

Although we found that SF tuning significantly changed depending on stimulus size and contrast, the functional significance of this property is unclear. One plausible explanation is that stimulus-size dependent SF tuning contributes to object recognition invariant changes in the retinal image size. That is, though the size and SF of the visual images vary according to changes in the viewing distance, the perceived object sizes are rather stable in a normal visual environment. This scenario represents the so-called size invariance of object recognition.

Evidence for size-invariant object recognition comes from several psychophysical experiments (Biederman & Cooper, 1991; Fiser & Biederman, 1995; Furmanski & Engel, 2000). Neurophysiological studies have also clarified that visual cortical areas, including cat V1 (Osaki et al., 2011) and monkey IT (Rolls, 2000, 2012; Tovee, Rolls, & Azzopardi, 1994), exhibit dynamic SF tuning in response to changes in stimulus size, which could then contribute to size-invariant visual processing. For example, Ito, Tamura, Fujita, and Tanaka...
(1995) measured the receptive field properties of neurons in the anterior part of the macaque IT cortex. They found that IT neurons exhibited large receptive fields up to about 50° in diameter and that 57% of the neurons had a stable response to preferred objects within stimulus sizes ranging over 2 octaves, showing that the response to visual stimuli is size invariant. Additionally, it has been argued that the relative SF tuning of IT neurons is an essential mechanism for generating size-invariant object recognition in primates (Inagaki & Fujita, 2011).

Finally, Georgeson and Sullivan (1975) reported that for contrast strength that sufficiently exceeds visual detection thresholds, the perceived contrast magnitude is independent of spatial frequency, size, or viewing distance, a property known as contrast constancy. They suggested that the visual system corrects for contrast losses caused by blurring of the image early in the visual process. Accordingly, at high contrast (10% or 90%), a participant will perceive the contrast of a stimulus as identical among a wide range of SFs and sizes. However, we found that the response probability varies with SF and size, suggesting the response probability is based on the physical luminance contrast rather than the perceived contrast. Therefore, different mechanisms would appear responsible for contrast constancy and size invariance.

Conclusions

Using the psychophysical subspace-reverse-correlation technique, we measured stimulus contrast and size effects on SF tuning in human perception. Our results show that human SF tuning depends on stimulus size. Increasing the stimulus size decreased the peak SF tuning in a manner that suggests relative SF tuning, which may contribute to size-invariant object recognition. Furthermore, the observed size-dependency of human SF tuning depended on the stimulus contrast. At 90% and 10% contrast, participants exhibited near perfect relative SF tuning, while at 1% contrast the stimulus size-dependent peak SF shift was too small for strictly relative SF tuning.

Keywords: psychophysical subspace-reverse-correlation; relative spatial frequency tuning; size-invariant object recognition

Acknowledgments

We are grateful to Dr. P. Karagiannis for his helpful comments on earlier versions of this manuscript and improving the English. This work was supported by JSPS KAKENHI Grant Numbers 22700339, 23530952, 24500460, 25380981.

Commercial relationships: none.
Corresponding author: Tomoyuki Naito.
Email: naito@vision.hss.osaka-u.ac.jp.
Address: Laboratory of Cognitive and Behavioral Neuroscience, Graduate School of Medicine, Osaka University, Osaka, Japan.

References


Campbell, B. Y. F. W., & Robson, J. G. (1968). Gratings by F. W. Campbell and J. G. Robson. From the Physiological Laboratory, University of Cambridge (Received 10 November 1967), 551–566.


Teichert, T., Wachtler, T., Michler, F., Gail, A., &


