Subspace mapping of the three-dimensional spectral receptive field of macaque MT neurons

Mikio Inagaki\textsuperscript{1,2}, Kota S. Sasaki\textsuperscript{1,2}, Hajime Hashimoto\textsuperscript{1}, Izumi Ohzawa\textsuperscript{1,2}

\textsuperscript{1}Graduate School of Frontier Biosciences, Osaka University, Suita, Osaka 565-0871, Japan
\textsuperscript{2}Center for Information and Neural Networks, Suita, Osaka 565-0871, Japan

Author contributions:
M.I., K.S.S., H.H., and I.O. conception and design of research; M.I., K.S.S., and H.H. performed experiments; M.I. analyzed data; M.I. and I.O. interpreted results of experiments; M.I. prepared figures; M.I. drafted manuscript; M.I. and I.O. edited and revised manuscript; M.I., K.S.S., H.H., and I.O. approved final version of manuscript.

Running Head:
Subspace mapping of MT 3-D spectral receptive field

Address for Correspondence:
Izumi Ohzawa
Graduate School of Frontier Biosciences, Osaka University, 1-4 Yamadaoka, Suita, Osaka 565-0871, Japan
Phone: +81-6-6879-4434
Fax: +81-6-6879-4438
E-mail: ohzawa@fbs.osaka-u.ac.jp
ABSTRACT

Neurons in the middle temporal (MT) visual area are thought to represent the velocity (direction and speed) of motion. Previous studies suggest the importance of both excitation and suppression for creating velocity representation in MT; however, details of the organization of excitation and suppression at the MT stage are not understood fully. Here, we examine how excitatory and suppressive inputs are pooled in individual MT neurons by measuring their receptive fields in a three-dimensional (3-D) spatiotemporal frequency domain. We recorded the activity of single MT neurons from anesthetized macaque monkeys. To achieve both quality and resolution of the receptive field estimations, we applied a subspace reverse correlation technique in which a stimulus sequence of superimposed multiple drifting gratings was cross-correlated with the spiking activity of neurons. Excitatory responses tended to be organized in a manner representing a specific velocity independent of the spatial pattern of the stimuli. Conversely, suppressive responses tended to be distributed broadly over the 3-D frequency domain, supporting a hypothesis of response normalization. In spite of the nonspecific distributed profile, the total summed strength of suppression was comparable to that of excitation in many MT neurons. Furthermore, suppressive responses reduced the bandwidth of velocity tuning, indicating that suppression improves the reliability of velocity representation. Our results suggest that both well-organized excitatory inputs and broad suppressive inputs contribute significantly to the invariant and reliable representation of velocity in MT.

(231 words / 250 words max)
New & Noteworthy

Neurons in the middle temporal (MT) visual area are thought to represent the velocity of visual motion. Our findings suggest that both well-organized excitatory inputs and broad suppressive inputs contribute significantly to the invariant and reliable representation of velocity in MT. Although suppression tends to be broadly distributed, total sum of suppression often exceed that of excitation. Furthermore, suppression sharpened the velocity tuning, and appears to improve the reliability of velocity representation.

KEYWORDS: motion, velocity, spatiotemporal frequency, suppression, reverse correlation

INTRODUCTION

Neurons in the middle temporal (MT) visual area of macaque monkeys show selectivity for the direction and speed of moving stimuli (Zeki 1974; Maunsell and Van Essen 1983; Albright 1984; Mikami et al. 1986; Lagae et al. 1993; DeAngelis and Uka 2003). Compared with V1, the bandwidths of direction tuning are broader (Albright 1984) and movement in the opposite direction tends to elicit suppressive responses relative to spontaneous activity in MT (Mikami et al. 1986). Therefore, pooling of both excitation and suppression beyond V1 must underlie the response properties of MT neurons. Indeed, a cascaded linear-nonlinear model incorporating both broad excitatory inputs from V1 and opponent suppression can capture a wide variety of responses to plaid stimuli (consisting of two drifting gratings) observed in MT (Rust et al. 2006).
However, this model does not provide details of excitation and suppression, such as whether excitation and suppression prefer the same speed but opposite directions. To this end, it is necessary to measure a receptive field in a three-dimensional (3-D) frequency domain spanned by spatial ($f_x$ and $f_y$) and temporal ($f_t$) frequencies (i.e., a 3-D spectral receptive field).

For the 3-D spectral receptive field of MT neurons, Simoncelli and Heeger (1998) proposed a computational model to represent velocity. As spatiotemporal frequency components of a 2-D pattern moving at a constant velocity lie on a plane passing through the origin of a 3-D spatiotemporal frequency domain, the receptive field of a neuron representing a particular velocity is expected to lie on the plane (Fig. 1). Since each V1 neuron is thought to be a spatiotemporal filter covering only a pair of small spherical regions, collecting output from many V1 neurons whose receptive fields lie on a specific velocity plane is necessary for the MT stage. In their model, suppression is postulated to occur for every spatiotemporal frequency that lies off the velocity plane to balance the total amounts of excitation and suppression. To date, whether that is also the case for actual MT neurons has not been examined experimentally.

Although the 3-D spectral receptive field, including excitation and suppression, of MT neurons has been examined (Nishimoto and Gallant 2011), it is still unclear how excitatory and suppressive inputs are organized in the 3-D frequency domain. Specifically, an outstanding question is their contributions to the velocity tuning profiles. To address these issues, we have
measured the 3-D frequency domain receptive fields of MT neurons by extending a subspace-mapping method that has been used in several laboratories. Compared with excitation, suppressive responses were less concentrated on a single velocity plane and tended to be scattered over the 3-D frequency domain. Furthermore, our analysis suggests that suppression contributes to sharpening the velocity tuning of MT neurons in agreement with Mikami et al. (1986) in which contributions of suppression are observed for direction selectivity and speed tuning separately.

MATERIALS AND METHODS

We recorded the neuronal activity of single MT neurons from two rhesus monkeys (*Macaca mulatta*; monkey N, male, 6 kg; monkey H, male, 6 kg). All animal care and experimental procedures conformed to the guidelines of the National Institutes of Health and were approved by the animal experiment committee of Osaka University.

Animal preparation

Before recording experiments, a head holder and recording chamber were surgically attached to the skull to allow for head restraint and electrode insertion. The head holder was positioned at the top of the skull and was attached with acrylic resin. The recording chamber was centered at approximately 17 mm lateral and 14 mm dorsal to the occipital ridge with a 25° elevation from the horizontal (Uka and DeAngelis 2003). A portion of the skull under the chamber was removed and the dura was exposed.
All surgical procedures were performed under anesthesia and aseptic conditions. The monkeys were sedated with ketamine hydrochloride (Ketalar, 5 mg/kg, i.m.) and medetomidine (Domitor, 0.05 mg/kg, i.m.). After starting inhalation of isoflurane (1–3% in 70% N₂O and 30% O₂) through an intratracheal cannula, the effects of medetomidine were reversed with atipamezole (Antisedan, 0.25 mg/kg, i.m.). Surgical anesthesia was maintained with isoflurane and local anesthesia was applied with lidocaine (2% Xylocaine) as needed. A vasotrop drug (Adona, 1 mg/kg, i.m.) and anti-plasmin agent (Transamine, 17 mg/kg, i.m.) were given to reduce bleeding during surgery. Lactated Ringer’s solution (Solulact-D, 10 mL/h) containing atropine sulfate (0.005 mg/kg/h) and riboflavin (Bisulase, 0.05 mg/kg/h) was infused through an intravenous tube. An electrocardiogram, arterial oxygen-saturation level, heart rate, body temperature, and end-tidal CO₂ were monitored continuously throughout surgery (BSM-3592 and GF-119; Nihon Kohden). A heating pad was used to maintain body temperature at 37–38°C. After surgery, an antibiotic (Pentcilin, 40 mg/kg) and anti-inflammatory and analgesic agent (Voltaren, 1 mg/kg) were given to the animals. During the first postoperative week, the monkeys were treated with the antibiotic and an anti-inflammatory and analgesic agent (Megeide, 0.8 mg/kg).

After a recovery period of 1 week, corneal curvature and the refracting power of the eyes were measured using an auto kerato-refractometer (KR-7100; Topcon) to select appropriate contact lenses for our experimental setup. Fundus photographs were taken using a retinal camera (TRC-50X; Topcon) to record the blood vessel patterns and find the position of the fovea. The correspondence of the positions on the retina and those of the monitor screen was confirmed by back-projection of the blood vessel patterns with the aid of an ophthalmoscope (BXα-13A;
The monkeys were sedated with ketamine hydrochloride (5 mg/kg, i.m.) and medetomidine (0.05 mg/kg, i.m.). After starting inhalation of isoflurane (1–3% in 70% N₂O and 30% O₂) through an intratracheal cannula, the effects of medetomidine were reversed with atipamezole (0.25 mg/kg, i.m.). Then, the opioid fentanyl citrate (Fentanyl, 12–25 mg/kg/h) in lactated Ringer’s solution (10 mL/h) containing atropine sulfate (0.005 mg/kg/h) and riboflavin (0.05 mg/kg/h) was infused through an intravenous tube. Under inhalation anesthesia, the superficial layer of the dura was stripped to insert an electrode safely. After surgery, inhalation of isoflurane was stopped. An electrocardiogram, arterial oxygen-saturation level, heart rate, body temperature, and end-tidal CO₂ were monitored continuously throughout the experiment. Body temperature was maintained at 37–38°C.

During recording, vecuronium bromide (Musculax, 0.1–0.2 mg/kg/h) was added to the solution to prevent eye movements and the monkeys were artificially ventilated. The pupils were dilated and the lenses were relaxed by the application of 0.5% tropicamide and 0.5% phenylephrine hydrochloride (Mydrin-P). The corneas were covered with the appropriate contact lenses selected according to the auto kerato-refractometer measurements. An artificial pupil (3 mm diameter) was added to the contact lenses. Each recording experiment lasted up to 8 h.

After recording, the infusion of the opioid and muscle relaxant was terminated and spontaneous
respiration recovered within 1.5 h. Respiration recovery was boosted by neostigmine methylsulfate (Vagostigmin, 0.05 mg/kg, i.m.). The eyes were washed with saline and treated with an antibiotic (Tarvid Ophthalmic Solution) and vitamin B₂ (Flavitan Eye Drops). An antibiotic (Pencilin, 40 mg/kg) and anti-inflammatory and analgesic agent (Megeide, 0.8 mg/kg) were provided to the monkeys before they were returned to their home cages. After a recovery period of 1–2 weeks, the next recording experiment was conducted.

Electrophysiology.

Tungsten electrodes (1.0 MΩ; Alpha-Omega) were used to record neuronal activity extracellularly. The voltage signals were amplified (×10000) and band-pass filtered (300–5000 Hz) by an amplifier (Model 1800; A-M Systems). Each spike was sorted by a custom-made on-line spike sorting system with 40 μs resolution (Ohzawa et al. 1996). With the aid of a guide tube, visual area MT was accessed from the recording chamber located over the occipital cortex. The transition patterns of spike occurrence during electrode penetration were consistent with the standard brain atlas of rhesus monkeys (Saleem and Logothetis 2007).

Visual stimulation.

Visual stimuli were presented on a CRT monitor (resolution, 1600 × 1024 pixels; refresh rate, 76 Hz; GDM-FW900; Sony) monocularly (the right eye in most cases). The monitor was positioned 57 cm away from the eyes of the monkeys. The screen subtended 23° × 30° in the visual angle. Only the green channel was used to avoid color misconvergence. The nonlinear relationship between the actual luminance and input voltage of the CRT monitor was measured.
using a photometer (CS-100; Minolta), and then corrected to a linear relationship with lookup tables.

We developed new stimuli to examine the full 3-D spectral receptive field of MT neurons within a reasonable recording time, typically 30 min to 1 h. The challenge in our approach is measuring responses to each of a large number of stimuli defined in 3-D space (fx, fy, and ft), which requires a very long recording time. To reduce the total time for presenting all stimuli, six drifting sinusoidal gratings were superimposed and presented simultaneously at any instant. The duration of each grating was 237 ms (18 video frames), which is brief but sufficiently long enough to define the required range of temporal frequencies. The superimposed grating stimuli were refreshed one component at a time every 39 ms (3 video frames). Therefore, each grating component persisted for 237 ms and was replaced by a new stimulus chosen randomly from the pool of stimuli. Since only one out of six gratings was updated at a time, the appearance of the stimuli was dynamic, but generally smooth without abrupt periodic changes. The construction of the stimuli is illustrated schematically in Fig. 2, and a short segment of the stimuli is presented as a movie in Supplemental Video S1.

The temporal frequencies (ft) of the stimulus set were 0, 4, 8, 12, 16, 20, and 24 Hz. The spatial frequencies (fx and fy) comprised a 13 × 13 matrix for all neurons, and the actual frequencies were adjusted for each neuron based on a preliminary search. After determining the cut-off spatial frequency with a drifting grating manually, we used it as the highest spatial frequency and evenly spaced the rest of the frequencies on the linear axis (e.g., if the cut-off was 3
cycles/deg, both fx and fy were -3, -2.5, -2, -1.5, -1, 0.5, 0, 0.5, 1, 1.5, 2, 2.5, and 3 cycles/deg).

The initial phase of the gratings was set to 0°, 90°, 180°, and 270°. The total number of stimuli was 4732 conditions (13 fx × 13 fy × 7 ft × 4 phases). With blank (uniform, mean-luminance screen) conditions (32 times) included, the grand total was 4764 conditions. This stimulus set constituted a single stimulus block. A block of stimuli was presented as a sequence of superimposed gratings by randomizing the order of the stimuli in the block. The contrast of each grating was 16% (1/6), so that the sum at any point in space and time will not exceed 100%. The stimulus block was repeated 5–10 times (approximately 30 min to 1 h) each time with a reshuffled order. The stimuli covered an area slightly larger than the classical receptive field determined by a manual search using drifting gratings.

For a subset of neurons, we tested direction selectivity using a single drifting sinusoidal grating. The contrast of the grating was 50% and an appropriate set of spatial and temporal frequencies was selected for each neuron. The number of the drifting direction was 12 (30° apart) or 24 (15° apart). Each drifting grating was presented for 4 s and repeated 5 times in a randomized order.

Data analysis

We applied a standard subspace reverse correlation (Ringach et al. 1997b) to the spike trains in response to our new stimuli (Fig. 2). The subspace was spanned by spatial (fx and fy) and temporal (ft) frequencies. The initial phase of the drifting gratings was ignored in this study, and the responses to four conditions were summed. The spike-triggered average was computed in the 3-D frequency domain by assigning a spike equally to every drifting grating that appeared in
a stimulus frame. To estimate the baseline response, the spike-triggered average for the
five-grating condition (in which one of six superimposed gratings did not appear) was
calculated and subtracted from the raw average. In this study, we refer to the positive part of the
subtracted result as an excitatory response and the negative part as a suppressive response
because they increased or decreased neuronal activity compared with baseline. Note that we
only measured net excitation and suppression: not the absolute excitation and suppression
strengths but the relative balance of them. The correlation delay was varied from 0 to 240 ms
with 15 ms spacing. The optimal delay was selected so that the variance of the excitatory
receptive field became largest.

To check the reliability of the estimated responses, we examined a two-fold cross validation
(2CV) by dividing the full dataset into odd and even trials. For each neuron, the similarity of the
estimated responses in odd trials and those in even trials was summarized by the goodness-of-fit
($r^2$ value).

From the excitatory and suppressive response profiles defined in the 3-D ($f_x$, $f_y$, and $f_t$) spectral
domain as above, we computed a number of parameters to characterize each neuron. A
suppression/excitation ratio was calculated as follows:

$$
Suppression/Excitation Ratio = \frac{\sum_{f_x} \sum_{f_y} \sum_{f_t} R_S(f_x, f_y, f_t)}{\sum_{f_x} \sum_{f_y} \sum_{f_t} R_E(f_x, f_y, f_t)}
$$
where \( f_{x_i}, f_{y_i}, \) and \( f_{t_i} \) are stimulus indices of each frequency (-6 to 6 for \( f_x \) and \( f_y \), 0 to 6 for \( f_t \)), \( HR[\ast] \) is half-wave rectification, and \( R \) represents response strength (after baseline subtraction) for specific stimulus conditions at a particular correlation delay. Both summed responses were computed at the optimal delay of the excitatory responses. If the ratio was 1, the excitatory and suppressive receptive fields were balanced. If the ratio was less/more than 1, the excitatory/suppressive receptive field was dominant, respectively.

To quantify how excitatory responses concentrate on a velocity plane and how they extend angularly on the plane, we used the on-plane ratio and horizontal-vertical ratio, respectively (Nishimoto and Gallant 2011). First, we determined an optimal velocity plane by using the maximum coverage and symmetry constraints (Nishimoto and Gallant 2011). As a velocity plane in the 3-D frequency domain is characterized by its azimuth (direction) and elevation (speed), we made a 2-D response map with direction and speed to examine response coverage as follows:

\[
RespMap_E(\phi, \psi) = \sum_{f_{x_i}} \sum_{f_{y_i}} \sum_{f_{t_i}} OP_E(\phi, \psi, f_{x_i}, f_{y_i}, f_{t_i}),
\]

\[
OP_E(\phi, \psi, f_{x_i}, f_{y_i}, f_{t_i}) = \begin{cases} R_E(f_{x_i}, f_{y_i}, f_{t_i}), & \text{if } |Z(\phi, \psi, f_{x_i}, f_{y_i}) - f_{t_i}| < 1.25, \\ 0, & \text{otherwise} \end{cases}
\]

\[
Z(\phi, \psi, f_{x_i}, f_{y_i}) = -f_{x_i} \tan \psi \cos \phi - f_{y_i} \tan \psi \sin \phi,
\]

where \( \phi \) and \( \psi \) are azimuth (direction) and elevation (speed), respectively. Essentially, on-plane excitation (\( OP_E \)) is defined as the excitation that lies within a "slab" of thickness 1.25 (5Hz).
along the \( f_t \) (temporal frequency) axis. To compute response map \( \text{RespMap}_E \), \( \text{OP}_E \) is then integrated within the slab for combinations of 48 directions (7.5° apart in azimuth, covering 360°) and 32 speeds (2.5° apart in elevation). The response map was smoothed by taking the average of the 3 × 3 neighbors. The range of speed depends on the spatial frequency range used for a recorded neuron. The response map was penalized by asymmetry in the summed responses on the two sides of the azimuth (direction) of the plane as the following form (Nishimoto and Gallant 2011):

\[
\text{AsymMap}_E(\phi, \psi) = |\text{RespMap}_E R(\phi, \psi) - \text{RespMap}_E L(\phi, \psi)|,
\]

where \( \text{RespMap}_E R \) is a response map created by summing responses on one side of the plane, and \( \text{RespMap}_E L \) is that of the other side. They were defined as follows:

\[
\text{RespMap}_E R(\phi, \psi) = \sum_{f_{x_t}} \sum_{f_{y_i}} \sum_{f_{t_i}} \text{OP}_E(\phi, \psi, f_{x_t}, f_{y_i}, f_{t_i}),
\]

\[
\text{OP}_E R(\phi, \psi, f_{x_t}, f_{y_i}, f_{t_i}) = \begin{cases} O E(\phi, \psi, f_{x_t}, f_{y_i}, f_{t_i}), & \text{if } \phi \geq \tan^{-1}\frac{f_{y_i}}{f_{x_t}} \\ 0, & \text{otherwise} \end{cases}
\]

\[
\text{RespMap}_E L(\phi, \psi) = \sum_{f_{x_t}} \sum_{f_{y_i}} \sum_{f_{t_i}} \text{OP}_E(\phi, \psi, f_{x_t}, f_{y_i}, f_{t_i}),
\]

\[
\text{OP}_E L(\phi, \psi, f_{x_t}, f_{y_i}, f_{t_i}) = \begin{cases} O E(\phi, \psi, f_{x_t}, f_{y_i}, f_{t_i}), & \text{if } \phi \leq \tan^{-1}\frac{f_{y_i}}{f_{x_t}} \\ 0, & \text{otherwise} \end{cases}
\]

The rationale of this constraint is that the response coverage map should have a single peak on the 2-D domain. We calculated a corrected map \( \text{CorrMap}_E \) by incorporating the maximum coverage and symmetry constraints as follows:

\[
\text{CorrMap}_E(\phi, \psi) = \text{RespMap}_E(\phi, \psi) - \text{AsymMap}_E(\phi, \psi).
\]

Based on the corrected map, the optimal direction and speed (i.e., the optimal velocity plane)
were determined for the neuron.

Having determined the optimal plane (i.e., optimal velocity), we quantify how much excitation is concentrated on or near the optimal plane. This is defined by the on-plane ratio for excitation as the sum of the responses on and near the optimal plane divided by the total response across the entire 3-D frequency domain as the following form:

\[ \text{On Plane Ratio (Exc)} = \frac{\sum \sum \sum OP_E(\phi_{opt}, \Psi_{opt}, f_{x_i}, f_{y_i}, f_{t_i})}{\sum \sum \sum R_E(f_{x_i}, f_{y_i}, f_{t_i})}. \]

If the on-plane ratio is 1, all responses lie on the plane. If the on-plane ratio is near 0, the majority of the responses lie off the plane. We also calculated the on-plane ratio for suppression by using \( R_S \) instead of \( R_E \).

We examined the angular distribution of the excitatory receptive field on the optimal plane by calculating the horizontal-vertical ratio as follows:

\[ \text{Horizontal Vertical Ratio (Exc)} = \frac{\sum \sum \sum OP_E^H(\phi_{opt}, \Psi_{opt}, f_{x_i}, f_{y_i}, f_{t_i})}{\sum \sum \sum OP_E^V(\phi_{opt}, \Psi_{opt}, f_{x_i}, f_{y_i}, f_{t_i})}, \]

\[ OP_E^H(\phi_{opt}, \Psi_{opt}, f_{x_i}, f_{y_i}, f_{t_i}) = \begin{cases} OP_E(\phi_{opt}, \Psi_{opt}, f_{x_i}, f_{y_i}, f_{t_i}), & \text{if } \tan^{-1}\frac{f_{y_i}}{f_{x_i}} < \phi_{opt} - \frac{\pi}{4}, \phi_{opt} + \frac{\pi}{4} < \tan^{-1}\frac{f_{y_i}}{f_{x_i}}; \\ 0, & \text{otherwise} \end{cases} \]
This essentially quantifies the completeness of the circular excitatory receptive field presented in Fig. 1, because it tends to be incomplete and non-annular for actual neurons with weak response weights near \( f_t = 0 \). The response weights around the optimal azimuth (±45°) were summed and compared with those of the rest (summed across -90° to -45° and 45° to 90°). If the horizontal-vertical ratio is 0, the receptive field is localized near the optimal direction. If the ratio is 1, the receptive field extends broadly and makes a complete ring, as expected by the Simoncelli-Heeger model. We also calculated the horizontal-vertical ratio for suppression by using \( R_s \) instead of \( R_E \).

To test the reliability of the velocity plane, we examined the tuning curve for rotation angle from the optimal velocity plane. An ideal velocity detector should be sensitive to a subtle deviation from the optimal velocity, and therefore its tuning curve should be sharp. Conversely, a detector for a particular set of spatial and temporal frequencies is not tuned for a unique velocity because it is possible that many velocity planes pass through the receptive field, making its tuning curve broad. To plot the tuning curve for the excitatory responses, we varied the rotation angle about the on-plane azimuth axis and calculated a series of summed responses (see Fig. 6) as follows:

\[
OP_EV(\phi_{\text{opt}}, \psi_{\text{opt}}, f_x, f_y, f_t) = \begin{cases} 
OP_E(\phi_{\text{opt}}, \psi_{\text{opt}}, f_x, f_y, f_t), & \text{if } \phi_{\text{opt}} - \frac{\pi}{4} \leq \tan^{-1} \frac{f_y}{f_x} \leq \phi_{\text{opt}} + \frac{\pi}{4}. \\
0, & \text{otherwise} 
\end{cases}
\]

This equation defines the excitatory part of the velocity field, where \( OP_E \) is a function that computes the excitatory response given the optimal angles and frequency weights. The condition inside the cases ensures that the response is non-zero only within a certain range of azimuth angles defined by \( \phi_{\text{opt}} \) and its deviation limits.
where $\theta$ is the rotation angle. A simple Gaussian function was fitted to the tuning curve and its standard deviation was used as a parameter representing tuning width. Tuning width is limited to less than 75° because the rotation angle ranged from -75° to 75°. Additionally, we calculated the tuning curve incorporating both the excitatory and suppressive responses by using $R$ instead of $R_E$.

To visualize a 3-D spectral receptive field, isosurfaces were drawn for the excitatory and suppressive responses. The original results of spike-triggered average analysis were smoothed using a Gaussian kernel (sigma: 0.65 relative to the data point interval), and then interpolated linearly so that the data points become 49 $f_x \times 49 f_y \times 25 f_t$ (originally 13 $f_x \times 13 f_y \times 7 f_t$). Data for negative temporal frequencies, when needed, were duplicated from the corresponding positive part symmetric about the origin. The isosurfaces were set to 50, 70, and 90% of the peak of the two receptive fields. The thresholds at 50, 70, and 90% were only for illustration purposes, and were not applied for other analyses.

To validate the receptive field recovered with our new stimuli, a direction-tuning curve made from the receptive field was compared with that of single grating measurements. The data points used for the single grating test were set in the 3-D frequency domain, and then the responses...
(smoothed and interpolated data) around them were averaged separately. A window of a sphere
(radius: 2 relative to the data point interval after interpolation) was used for the averaging. The
averaged responses were half-wave rectified to produce non-negative values resembling firing
rates. After normalization, the two direction-tuning curves were compared and the
goodness-of-fit ($r^2$ value) was calculated to quantify their similarity (grating validation: GRV).

RESULTS

We recorded from 78 MT neurons in two monkeys (37 in monkey N, 41 in monkey H).
Superimposed drifting sinusoidal gratings were used to examine the 3-D receptive field in the
frequency domain ($f_x$, $f_y$, and $f_t$). Excitatory and suppressive receptive fields were recovered
separately after subtracting baseline activity. The excitatory/suppressive receptive field indicates
that the relative balance of excitation is higher/lower than suppression, respectively. They do not
indicate the absolute strengths of excitation and suppression. For example, if excitation and
suppression are of equal strength for a given ($f_x$, $f_y$, $f_t$), they will cancel each other and neither
will be recovered.

3-D spectral receptive field of MT neurons

We observed a wide variety of profiles for 3-D receptive fields in the MT population that were
consistent with known diversities reported in previous studies using gratings (Perrone and
Thiele 2001; Priebe et al. 2003), plaids (Movshon et al. 1985), random-dot motion (Kumano
and Uka 2013), and natural movies (Nishimoto and Gallant 2011). The neuron shown in Fig. 3
responded to many gratings (Fig. 3A) and was also suppressed by many other gratings (Fig. 3B).
The distribution of the raw responses was positively skewed (Fig. 3C), indicating that some stimulus conditions elicited strong excitatory responses. In the excitatory receptive field (Fig. 3D, top row), this neuron demonstrated a considerable dependence on the combination of spatial and temporal frequencies, producing a slightly elongated egg-shaped profile in the side view perpendicular to the optimal direction (Fig. 3D, top row, second from right). The axis of elongation appeared to go through the origin. As the speed was constant along an oblique line through the origin, this neuron was tuned for a constant speed in a manner relatively independent of the spatial and temporal frequencies of the grating stimuli. The excitatory receptive field extended along the orthogonal axis as this neuron responded to a broad range of drifting directions. These properties partly matched the ideal velocity detector model (Fig. 1) in which the receptive field makes a tilted ring in the 3-D frequency domain. However, this neuron did not respond to gratings near zero temporal frequency, making the ring a partial one.

The suppressive receptive field of this neuron was centered on the opposite direction relative to the optimal direction of the excitatory profile (Fig. 3D, middle row). As the strength of suppression was weaker than that of excitation, the suppressive receptive field seemed to be small in the isosurface plots. Unlike the excitatory receptive field, there was no clear oblique elongation in the side view perpendicular to the optimal direction (Fig. 3D, middle row, second from right). Although direction preference was roughly opposite for the excitatory and suppressive receptive fields, the preferred speeds of excitation (the elevation of the red dashed
line) and suppression (the elevation of the blue dashed line) slightly differed (Fig. 3D, bottom row, second from right). The two receptive fields did not demonstrate exact mirror symmetry of each other in the 3-D frequency domain (Fig. 3D, bottom row).

We checked the validity of the 3-D spectral receptive field in two different ways. First, we examined similarity of two receptive fields (raw data) each recovered from odd or even trials separately (two-fold cross validation; 2CV). The $r^2$ of the 2CV was 0.41 for this MT neuron (Fig. 3D). Second, for a subset of MT neurons, we tested direction selectivity using a single sinusoidal drifting grating. We compared a direction-tuning curve estimated from the 3-D spectral receptive field (smoothed and interpolated data) with that of a grating measurement (grating validation; GRV). Data points for the grating measurement were positioned in a circle in the 3-D frequency domain floating at the height corresponding to the temporal frequency of the grating, as shown by the small black spheres in Fig. 3E. The strengths of the excitatory and suppressive responses were averaged locally around each data point (within the extent of each sphere) and then half-wave rectified to produce non-negative values (Fig. 3E). The direction-tuning curves were similar and the $r^2$ was very high (0.97) for this neuron (Fig. 3F).

For the majority of the MT neurons tested (38/62), the $r^2$ value of GRV was higher than 0.8 (Fig. 3G). This result suggests that our receptive field estimation using superimposed gratings is comparable to that of conventional single grating measurements.

The degree of joint dependence on the spatial and temporal frequencies (i.e., speed tuning) varied across MT neurons. The degree of angular extent (i.e., perfectness of the ring shape) also
differed among MT neurons. The three example neurons shown in Fig. 4 demonstrate these
diversity across the MT population. The neuron shown in Fig. 4A had a very dominant
excitatory receptive field and small suppressive receptive field. The excitatory receptive field
was elongated along an oblique line, but did not extend angularly compared with the previous
example in Fig. 3. Therefore, this neuron was tuned for a constant speed, but collected motion
information over a rather limited range of spatiotemporal frequencies. For the second neuron
shown in Fig. 4B, an oblique profile was not obvious in either the excitatory or suppressive
receptive fields. The two receptive fields did not demonstrate clear symmetry in the 3-D
frequency domain, since the preferred speeds were different for excitation and suppression, as
indicated by the different slopes of the red and blue dashed lines (Fig. 4B, center). The optimal
speed for suppression appeared to be substantially higher for suppression than excitation. The
optimal spatial and temporal frequencies also appeared to be different. The third neuron shown
in Fig. 4C had a separable profile for spatial and temporal frequencies without joint dependence.
This neuron responded to a small number of gratings and its direction tuning was sharp. This
profile is suitable for the detection of a specific spatial frequency irrespective of temporal
frequency, such as a typical V1 simple cell (Priebe et al. 2006), but is far from that of an ideal
velocity detector.

Properties of the excitatory receptive field of MT neurons

In this section, we quantitatively examined the characteristics of the excitatory receptive field of
MT neurons with respect to the distributions of the preferred velocities and sharpness of tuning.

An optimal velocity plane was determined for the excitatory receptive field (raw data) of each neuron by using maximum coverage and symmetry constraints (see MATERIALS AND METHODS). The direction and speed of the optimal velocity plane were plotted for the population of MT neurons in Fig. 5A. There was no clear bias in optimal direction. The distribution of optimal speed was continuous and ranged broadly (approximately 1–50 deg/s).

These results were consistent with previous studies (e.g., DeAngelis and Uka 2003), indicating that our sampling was not biased to a particular subset of the whole MT population.

The example MT neurons shown in Figs. 3 and 4 demonstrated a variety in their degree of joint dependence between the spatial and temporal frequencies (i.e., speed tuning). The degree of angular extent also varied across the example neurons. To examine these properties over the MT population, we quantified the on-plane and horizontal-vertical ratios for the excitatory receptive field (raw data) of each neuron (see MATERIALS AND METHODS). The on-plane ratio is related to the joint dependence between the spatial and temporal frequencies. It is essentially a metric of “goodness-of-fit” for the Simoncelli-Heeger model (1998; see Fig. 1). If a receptive field is elongated obliquely in the 3-D frequency domain, such as the neuron shown in Fig. 3, the on-plane ratio has a higher value. If a receptive field is separable and elongated vertically, such as the neuron shown in Fig. 4C, the value becomes lower. The horizontal-vertical ratio quantifies another aspect of the receptive field of MT neurons, i.e., how complete an MT
receptive field is. Compared with the idealized MT model (Fig. 1), actual MT neurons appear to be missing input from V1 for regions near \( ft = 0 \) (i.e., near the horizontal plane). The horizontal-vertical ratio quantifies variations in this respect. If a receptive field extends broadly, such as the neuron shown in Fig. 3, the horizontal-vertical ratio is close to 1. If a receptive field is localized in a narrow range, such as the neuron shown in Fig. 4C, the value becomes lower. As an ideal velocity detector has higher values near 1 for both ratios, it should be plotted at the upper right corner in the scatter plot spanned by the two ratios. Actual MT neurons did not cluster near the upper right, but were distributed broadly (Fig. 5B), indicating that MT neurons implement the Simoncelli-Heeger model in an incomplete manner. In other words, few neurons in visual area MT had the ideal receptive field of the Simoncelli-Heeger model. Note that there was no correlation between the two ratios \((r = -0.006, p = 0.956)\), suggesting that these parameters are separate.

So far, we determined the optimal velocity for the excitatory responses for each neuron, and examined their distributions for the population. We next examined the sharpness of velocity tuning. This may be determined from a velocity-tuning map spanned by direction and speed axes for each neuron (see MATERIALS AND METHODS). If the velocity tunings of MT neurons are simple, i.e., where speed and direction tunings are mutually independent (or separable), we could simply measure the respective bandwidths. However, these two dimensions are not separable, and are known to exhibit complex behavior in general (Kumano and Uka 2013). We found similar complexity for our MT neurons. For example, the velocity-tuning map of the neuron shown in Fig. 3 is simple, and is essentially circular and...
separable for direction and speed (Fig. 6A). However, in another example, the velocity-tuning map of the neuron shown in Fig. 4A exhibits a “U-shape” and the direction-tuning curves (horizontal cross-sections) became bimodal at higher speeds (Fig. 6B). With these variations, meaningful comparisons of velocity tuning widths across neurons and constructing a summary for the population would be difficult if we adopt traditional bandwidth metrics in the speed and direction domains.

For the reasons noted above, we defined velocity-tuning bandwidth for a receptive field (raw data) as follows. Since a theoretically minimal MT neuron receives input from a single V1 complex cell, its 3-D spectral receptive field is the same as that of a V1 complex cell, as illustrated in Fig. 6C, with two spheres positioned symmetrically about the origin (Simoncelli and Heeger, 1998). The velocity plane (through the origin) is constrained, leaving one free parameter, rotation angle \( \theta \), as shown in Fig. 6D. The velocity plane for actual MT neurons is further constrained with respect to \( \theta \) due to additional V1 complex cell input. Therefore, this allows us to define bandwidth for velocity tuning using \( \theta \) in a manner applicable to all MT neurons, from the most V1-like to the most MT-like, without the complications of dealing with double peaks. Fig. 6E depicts the prediction of a velocity-tuning profile for a model V1 complex cell, and it clearly shows the origin of the U-shaped response in Fig. 6B. Although this definition of bandwidth may appear somewhat unintuitive initially, it is related to the strength and extent of additional V1 input and allows a natural definition in the simplest manner.

Using the definition of velocity tuning bandwidth described above, if a neuron is sensitive to a
deviation from its optimal velocity, the optimal velocity plane is reliable and the tuning curve for rotation angle should be sharp (Fig. 6F, black curve). Otherwise, the optimal velocity plane is unreliable and the tuning curve should be broad (Fig. 6F, gray curve). In the extreme case of a model V1 complex cell, the curve would be nearly flat. For the neuron that had a partial ring excitatory receptive field (Fig. 3), the tuning curve for $\theta$ was sharp and its tuning width was relatively small (Fig. 6G; tuning width = 35.6). For the neuron that had a limited extent on the velocity plane (Fig. 4A), tuning for $\theta$ was broad and its tuning width was relatively large (Fig. 6H; tuning width = 44.9). The tuning width was negatively correlated with the on-plane ratio (Fig. 6I; $r = -0.590$, $p < 0.001$) and also correlated with the horizontal-vertical ratio (Fig. 6J; $r = -0.565$, $p < 0.001$). Note that negative correlation coefficients were also observed for scrambled data in which the stimulus-response relations were randomized within each neuron (tuning width vs. on-plane ratio, $r = -0.254$, $p = 0.025$; tuning width vs. horizontal-vertical ratio, $r = -0.104$, $p = 0.363$). Thus, a part of the negative correlation was explained by intrinsic correlation among the metrics. However, stronger negative correlations were observed in the actual data, suggesting that the sharpness of velocity tuning, and hence the reliability of velocity representation, depends on having as much excitatory input on a single velocity plane as possible and on the extent of pooling on this plane.

Properties of the suppressive receptive field of MT neurons

Next, we investigated the properties of the suppressive responses of MT neurons. We quantified
the strength of the suppressive responses by the suppression/excitation ratio in which the total
excitatory and suppressive responses (raw data) were compared at the optimal delay of
excitation. The median of the suppression/excitation ratio was 0.68 across the MT neurons (see
Fig. 8D). Roughly one third of neurons (22/78) had a suppression/excitation ratio higher than 1,
indicating that those neurons showed stronger suppression than excitation. We also calculated
the suppression/excitation ratio using two separate correlation delays each optimized for
excitation and suppression, respectively. On average, the optimal delays were 65 ± 1.9 ms
(mean ± SEM) for excitatory response and 58.7 ± 4.2 ms for suppressive responses, respectively.
Using separate optimal delays for excitation and suppression, the median of the
suppression/excitation ratio was 0.79 and higher than original one, but the difference was not
statistically significant (Mann-Whitney U test, p = 0.157).

We also examined how the baseline firing-rate (the response to the five-grating condition)
relates to the suppression/excitation ratio. We did not observe significant correlation (n = 78, r =
-0.05, p = 0.67). Additionally, we tested how the reliability of the receptive field estimation
relates to the suppression/excitation ratio. There was no significant correlation between the r²
value of 2CV and the suppression/excitation ratio (n = 78, r = -0.19, p = 0.10). We thus
concluded that neither the baseline firing-rate nor the estimation reliability systematically biases
the estimated suppression strength.

Some MT neurons demonstrated clear suppression, such as that shown in Fig. 3. For this neuron,
the peak of suppression exceeded 50% of the excitatory peak (Fig. 7A). To examine the
suppressive profiles in detail, a contour plot was made for a 2-D cross-section spanned by spatial frequency along the optimal direction and temporal frequency (Fig. 7B). Unlike the excitatory profile, the suppressive profile was not elongated obliquely toward the origin. We calculated the on-plane ratio to check how suppressive responses concentrate on a single plane. The on-plane ratios of the suppressive responses (raw data) were lower than those of the excitatory responses (Fig. 7C; Wilcoxon signed rank test, p < 0.001). This suggests that the suppressive responses were concentrated less on a single velocity plane than the excitatory responses in MT neurons. We also compared the horizontal-vertical ratios of the excitatory and suppressive responses (raw data). The horizontal-vertical ratios were higher for the suppressive responses (Fig. 7D; Wilcoxon signed rank test, p < 0.001), indicating that the suppressive responses were distributed relatively broadly on the optimal plane. Taken together, the suppressive responses tended to be scattered in the 3-D frequency domain compared with the excitatory responses.

We next examined the effects of the suppressive responses on the reliability of the optimal velocity plane determined with excitatory responses. The neuron shown in Fig. 4B had relatively broad tuning for plane rotation when the suppressive responses were excluded (Fig. 8A). If the suppressive receptive field (raw data) was also taken into account, the tuning curve became sharp (Fig. 8B). The tuning width was reduced from 43.9 to 25.1 by incorporating suppression. In most MT neurons, suppressive responses contributed to a reduction of tuning
width (Fig. 8C; Wilcoxon signed rank test, p < 0.001). Furthermore, the degree of tuning-width
reduction was correlated with the suppression/excitation ratio (Fig. 8D; r = 0.615, p < 0.001),
suggesting that stronger suppression contributes to making the velocity tuning sharper, thus
increasing the reliability of velocity representation in MT.

-------------------- Figure 8 near here ---------------------

Finally, we calculated the pattern index (Smith et al. 2005) to test whether the neurons respond
to component- or pattern-direction of plaid stimuli (Movshon et al. 1985). For the neurons
validated with single gratings, we made predictions of component and pattern tunings using the
tuning for the single gratings. We compared them with a tuning for plaid stimuli (grating
separation: 120 degrees) estimated from the spectral receptive field (smoothed and interpolated
data), and computed the pattern index. A pattern index < 0 indicates component
direction-selective, and > 0 pattern direction-selective. For our population (n = 62), 50 neurons
had a value < 0, and 12 neurons > 0. Moreover, 28 neurons were classified as "component cell"
(< -1.28), and 2 neurons "pattern cell" (> 1.28). The median of the pattern index was -1.04, and
similar to previously reported value of -1.01 (Kumano and Uka 2013).

DISCUSSION

In this study, we examined how excitation and suppression are pooled in single MT neurons by
recovering a 3-D spectral receptive field. We developed a new method to obtain a full 3-D
spectral receptive field within a reasonable recording time. The activity of 78 MT neurons was
recorded from anesthetized monkeys and subspace reverse correlation was applied to the spike trains. The excitatory responses of MT neurons were well organized in a manner partly consistent with the Simoncelli-Heeger model (1998). Conversely, the suppressive responses were less organized and scattered over the 3-D frequency domain. Nonetheless, the presence of suppression contributed substantially to the reliability of velocity representation in MT.

**Shape of the excitatory receptive field in the 3-D frequency domain**

If each MT neuron is tuned to a specific velocity that is represented by a velocity plane (tilted plane going through the origin of the 3-D spatiotemporal frequency domain), we would expect the excitatory receptive field to be concentrated on this plane. We quantified how excitatory responses lie on a single velocity plane and how they extend angularly by calculating the on-plane ratio and horizontal-vertical ratio, respectively. If a receptive field forms a tilted ring-shape fully consistent with the Simoncelli-Heeger model (1998), both ratios will have large values. Conversely, for a separable receptive field responding to a particular combination of spatial and temporal frequencies, both ratios will have small values. In reality, the distributions were broad and unimodal, and only a single cluster was found in the scatter plot spanned by the two ratios (Fig. 5B). Consistent with previous studies (Movshon et al. 1985; Perrone and Thiele 2001; Priebe et al. 2003; Nishimoto and Gallant 2011; Kumano and Uka 2013), MT neurons were not sorted into discrete extreme classes based on their response properties. Although a theoretical model proposes a complete ring shape for the excitatory receptive field (Simoncelli and Heeger 1998; see Fig. 1), we did not find such a complete ring, but found a partial ring in a subset of MT neurons. These results are largely consistent with those of a previous study using
awake animals, natural movies, and a model-fitting framework (Nishimoto and Gallant 2011).

The lack of responses at zero temporal frequency mainly made the receptive field a partial ring. This property matches a previous finding tested with bar stimuli (Albright 1984). Albright reported that approximately 30% of MT neurons respond to a static (zero temporal frequency) bar parallel to their optimal direction, but the response was weaker by a factor of three on average compared with the response to an optimal moving bar. Taken together, MT neurons rarely have a clear ring shape for the excitatory receptive field suggested by the theoretical model.

Both on-plane and horizontal-vertical ratios were negatively correlated with the bandwidth of velocity tuning (Fig. 6I, J). Importantly, there was no correlation between the on-plane and horizontal-vertical ratios (Fig. 5B), indicating that these two ratios were independently correlated with tuning width without spurious correlations. These results indicate that coverage on a single velocity plane and its angular extent are not associated with each other and contribute separately to the sharpness of velocity tuning.

**Functional roles of suppression in velocity tuning**

We identified suppressive receptive fields as well as excitatory ones for MT neurons. Although suppressive input from the opposite direction is important for creating pattern direction selectivity in MT (Rust et al. 2006) and a suppressive receptive field was recovered in a previous study (Nishimoto and Gallant 2011), the organization of the suppressive receptive field in the 3-D frequency domain is not understood fully. At the preceding stage, i.e., V1 complex
cells, responses are thought to be generated by motion energy computations. In its general form, motion energies are calculated for the forward and opposite directions and then subtracted to produce an opponent motion signal. Such opponent signaling is thought to contribute to noise reduction (Adelson and Bergen 1985; Watson and Ahumada 1985; Rust et al. 2005; see a review, Bradley and Goyal 2008). If opponent motion encoding is the case for MT neurons, their excitatory and suppressive receptive fields should be mutually symmetric in the 3-D frequency domain. Therefore, it could be predicted that the on-plane and horizontal-vertical ratios are comparable between excitatory and suppressive responses. Contrary to that prediction, the on-plane ratio of the suppressive responses was significantly lower than that of the excitatory ones (Fig. 7C), and the horizontal-vertical ratio of the suppressive responses was significantly higher than that of the excitatory ones (Fig. 7D). Instead, the suppressive responses were distributed broadly in the 3-D frequency domain compared with the excitatory responses.

Note also that the total summed strength of suppression was quite strong. As Fig. 8D indicates, the sum total of suppression was even stronger than that of excitation in approximately one-third of our samples (suppression/excitation ratio > 1). Although the individual examples presented in Figs. 3D and 4 may give the impression that suppression is generally weak, that is because these plots only show peaks and not widely distributed weak suppression. In other words, we may conclude that excitation is generally concentrated on a velocity plane, but suppression tends to be everywhere within the 3-D frequency domain. Such a scheme may play an analogous role to the well-known response normalization proposed by Heeger (1992; see a review, Carandini and Heeger 2012). Indeed, it has been proposed as unselective suppression in
a model of MT neurons (Simoncelli and Heeger 1998). Moreover, the strength and the spread of the suppression we measured might be underestimated, because inhibition is significantly weakened and localized in V1 receptive field of mice under anesthesia (Haider et al. 2013).

As to the effects of suppression on the tuning characteristics of MT neurons, incorporating suppressive responses significantly reduced the bandwidth of velocity tuning, indicating that suppression improves velocity representation in MT. For example, the neuron presented in Fig. 8A and B showed a reduction of tuning width (in rotation angle) from 44 to 25 degrees when suppression was incorporated. Although we adopted a bandwidth metric based on the plane rotation angle because of its uniform applicability to all MT neurons, the metric is not intuitive. For the take-home image of the improvement of bandwidth in conventional parameter space, the bandwidths (full-width-at-half-height) of direction and speed tuning for the above example were made narrower from 195 degrees and 6.67 degrees/s without suppression to 113 degrees and 6.39 degrees/s with suppression, respectively. Obviously, this kind of unit conversion can only be applied to a subset of well-behaved MT neurons for the reasons described for Fig. 6.

Extension of the subspace mapping method

The method we used here is an extension of a subspace reverse correlation technique in which stimulus space is limited to a particular subspace that experimenters are interested in (Ringach et al. 1997b). The signal to noise ratio is higher than that of white noise analysis because stimuli have random but clear dynamic “wavy” shapes and stimulus energy is concentrated in a limited number of sinusoidal components. A drawback is that responses are measured at a limited
number of sampled points in the 3-D frequency domain, whereas with white noise stimuli, the
domain can be explored more thoroughly provided that a sufficiently long time is allowed for
measurements. Subspace reverse correlation has been applied to orientation tuning (Ringach et
al. 1997a), spatial frequency tuning (Bredfeldt and Ringach 2002), chromaticity tuning (Cottaris
and De Valois 1998), 2-D joint tuning of orientation and spatial frequency (Mazer et al. 2002;
Shapley et al. 2003; Nishimoto et al. 2005), and binocular disparity selectivity (Baba et al.
2015). We extended this technique into 3-D joint tuning of spatial (fx and fy) and temporal (ft)
frequencies to study MT response properties for motion stimuli.

A major problem of this extension is temporal correlation among neighboring stimulus frames
because of the nature of motion, i.e., it requires a certain duration by definition. Therefore, we
did not focus on the temporal dynamics of the MT response, although previous studies using
subspace reverse correlation revealed dynamic changes of tuning along the temporal
(correlation delay) axis. Another potential concern is the superimposition of the drifting gratings
adopted to reduce the total time required for stimulus presentation. In a subset of MT neurons,
called pattern direction-selective neurons, responses to plaid stimuli (consisting of two drifting
gratings) significantly differ from the sum of responses to component gratings (Movshon et al.
1985). In addition, such pattern direction-selective neurons greatly reduce direction selectivity
for tri-plaid stimuli in which three drifting gratings were superimposed (Jazayeri et al. 2012).
Since their tri-component stimuli always had at least one component off the velocity plane for
the neuron, the reduction of responses was generally consistent with our finding that many
neurons have suppressive regions that can extend over a wide volume in the 3-D frequency
However, it should be noted that the suppressive interactions described above may affect the estimation of the spectral receptive field when we use superimposed drifting grating. To examine this point, we compared a direction-tuning curve calculated from a recovered spectral receptive field with that of single drifting grating measurements. The two tuning curves nicely matched each other in most of our samples (Fig. 3G). We conclude that receptive field estimation using superimposed drifting grating is valid in the sense that it is comparable to estimation using single drifting grating stimuli.

ACKNOWLEDGEMENTS

We thank members of our laboratory and also members of Dr. Ichiro Fujita’s laboratory for help with the experiments and discussions.

GRANTS

This work was supported by grants to I.O. from the Ministry of Education, Culture, Science, Sports, and Technology (KAKENHI 22135006, 15H05921).

DISCLOSURES

The authors declare no conflict of interest, financial or otherwise.
REFERENCES


Ohzawa I, DeAngelis GC, and Freeman RD. Encoding of binocular disparity by simple cells in


Smith MA, Majaj NJ, Movshon JA. Dynamics of motion signaling by neurons in macaque area


FIGURE LEGENDS

Fig. 1. Simoncelli and Heeger model for the 3-D spectral receptive field of MT neurons. A: The MT receptive field forms a tilted ring-shape on a plane in the 3-D spatiotemporal frequency domain. B: A top view of the fx-fy profile. opt, optimal direction; orth, orthogonal direction relative to optimal. C: A side view of the fx-ft profile that highlights the dependency between spatial and temporal frequencies. D: Another side view of the fy-ft profile that captures a ring-shaped profile extending annularly.

Fig. 2. Subspace reverse correlation in a spatiotemporal frequency domain. For each correlation delay, the spike-triggered average was computed across all spikes elicited during stimulus presentation. The averaging was done in a parameter space spanned by spatiotemporal frequencies (fx, fy, and ft) of drifting grating stimuli. Stimulus parameters of all gratings that appeared in each spike-preceding frame were equally fed into the averaging. DG: drifting grating.

Fig. 3. Three-dimensional spectral receptive field of a representative MT neuron. A, B: Spike-triggered average in the 3-D frequency domain. After subtracting baseline activity, the positive (A) and negative (B) parts were plotted separately. We referred to these as excitatory and suppressive responses, respectively. The radius of each sphere is linearly related to response magnitude. C: Distribution of raw responses. Red histogram is for excitatory responses, blue for suppressive ones. Arrow indicates the response for the five-grating condition ("baseline"). D: Isosurface plots of the 3-D spectral receptive field. After smoothing the raw data, isosurfaces
were created for 50, 70, and 90% of the maximum response for both excitatory (top row) and suppressive (middle row) responses. They are overlaid in the bottom row to highlight their positional relationship. Only an isosurface of 50% is available in the suppressive profiles because the suppressive responses did not reach 70% of the maximum of the excitatory responses. To help understand the 3-D shape of the receptive fields, an overlook view and top and side views were created from the same data. Two side views are created from viewing directions perpendicular to the optimal direction axis and perpendicular to the orthogonal direction axis, respectively. 2CV, two-fold cross validation. E: Schematic illustration of tuning-curve estimation from a spectral receptive field. The black spheres define a region around each spatiotemporal frequency of drifting gratings where the response weights, including both excitatory (red) and suppressive (blue), are summed together to produce a direction-tuning curve. F: Normalized direction-tuning curves. The black line represents a tuning curve estimated from the spectral receptive field (dashed line: a tuning curve without half-wave rectification). The gray line represents a tuning curve measured by drifting gratings. G: Distribution of goodness-of-fit (r^2) between direction-tuning curves from the two different measurements.

Fig. 4. Examples of 3-D spectral receptive field for three additional MT neurons. A: A neuron showing joint dependence on spatial and temporal frequencies. The conventions are same as in the bottom row of Fig. 3C. The range of spatial frequency was -1.4 to 1.4 cycles/deg for this neuron. B: A neuron with an intermediate joint dependence on spatial and temporal frequencies. The range was -4.8 to 4.8 cycles/deg. C: An example with an excitatory receptive field
separable for spatial and temporal frequencies. The range was -3.0 to 3.0 cycles/deg.

Fig. 5. Summary of the basic characteristics of the excitatory receptive fields of MT neurons. A: Distributions of the optimal direction and optimal speed determined by the excitatory receptive field. Circles indicate top 50% validated neurons in 2CV (Top50), and triangles bottom 50% (Btm50). B: Distributions of the on-plane and horizontal-vertical ratios calculated for the excitatory receptive field (circles: Top50, triangles: Btm50).

Fig. 6. Relationship between the reliability of velocity tuning and characteristics of excitatory receptive fields. A, B: Examples of 2-D velocity tuning maps of the excitatory responses. Progressively darker regions indicate higher responses and each contour level corresponds to 30% to 90% (10% step) of the maximum response. For clarity, the optimal direction was set to 0 degrees. C: The 3-D spectral receptive field of a model V1 complex cell (gray spheres) that represents a particular spatiotemporal frequency. The optimal velocity plane (gray plane) passes through the centers of the receptive field and the origin. Note that, for a V1 complex cell, the plane is not constrained for rotation about the axis connecting the two spheres. D: A schematic illustration for estimating the reliability of velocity tuning. Reliability is essentially a measure of additional constraints beyond that for the V1 complex cell in C, and is defined as sensitivity to the plane-rotation angle around the optimal velocity plane. E: 2-D velocity map of the model V1 complex cell shown in C. A white dot represents the optimal velocity and a white line corresponds to the trajectory for rotating the plane about the on-plane optimal direction axis. On the trajectory, the tick interval is 15 degrees in the rotation angle. F: Tuning curve for the
plane-rotation angle. Bandwidth of the tuning is a metric of velocity-tuning reliability. $G, H$:

Examples of a tuning curve for two MT cells calculated for an excitatory receptive field. Each circle represents a summed response on and near a rotated plane. The solid curve is a Gaussian function fitted to the data $I, J$: Relationships between the tuning width and on-plane ratio ($I$), and the tuning width and horizontal-vertical ratio ($J$) (circles: Top50, triangles: Btm50).

Fig. 7. Characteristics of suppressive receptive fields. $A$: Side view of the spectral receptive field of the neuron shown in Fig. 3. This is a re-plot of the second figure from the bottom of Fig. 3C. $B$: 2-D cross-section of the spectral receptive field. The plane for the cross-section goes through the origin, and is parallel to both the optimal direction axis and temporal frequency axis. For clarity, the excitatory and suppressive responses are normalized separately before making the contour plots in this figure. $C$: Comparison of the on-plane ratios of the excitatory and suppressive responses (circles: Top50, triangles: Btm50). $D$: Comparison of the horizontal-vertical ratios of the excitatory and suppressive responses (circles: Top50, triangles: Btm50).

Fig. 8. Effects of suppression on velocity-tuning reliability. $A, B$: Tuning curves for the plane-rotation of the neuron shown in Fig. 4B. Only excitatory responses are used in $A$. Both excitatory and suppressive responses are used in $B$. The conventions are the same as in Fig. 6G, H. The dashed line in $B$ is a re-plot of the tuning curve in $A$. $C$: Comparison of tuning widths calculated with excitatory responses only and those of all responses (circles: Top50, triangles: Btm50). $D$: Relationship between the suppression effects on tuning width and total suppression.
strength relative to that for excitation (circles: Top50, triangles: Btm50).
Inagaki et al.

Fig. 1
Inagaki et al.
Fig. 2
Inagaki et al.
Fig. 3
Inagaki et al.
Fig. 4
Inagaki et al.
Fig. 5
Inagaki et al.

Fig. 6
Fig. 8

A N0905_ep193_005_ch0_sp1
Tuning width (Exc) = 43.9

B N0905_ep193_005_ch0_sp1
Tuning width (Exc + Sup) = 25.1

C

D

N = 78

r = 0.615
p < 0.001

Inagaki et al.
Fig. 8