NeuroImage 63 (2012) 1321–1333

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Selective responses to specular surfaces in the macaque visual cortex revealed by fMRI

Gouki Okazawa, Naokazu Goda *, Hidehiko Komatsu

Division of Sensory and Cognitive Information, National Institute for Physiological Sciences, Okazaki 444-8585, Japan
Department of Physiological Sciences, The Graduate University for Advanced Studies (Sokendai), Okazaki 444-8585, Japan

ARTICLE INFO

Article history:
Accepted 23 July 2012
Available online 1 August 2012

Keywords:
Inferior temporal cortex
Surface properties perception
Ventral visual pathway

ABSTRACT

The surface properties of objects, such as gloss, transparency and texture, provide important information about the material characteristics of objects in our visual environment. However, because there have been few reports on the neuronal responses to surface properties in primates, we still lack information about where and how surface properties are processed in the primate visual cortex. In this study, we used functional magnetic resonance imaging (fMRI) to examine the cortical responses to specular surfaces in the macaque visual cortex. Using computer graphics, we generated images of specular and matte objects and prepared scrambled images by locally randomizing the luminance phases of the images with specular and matte objects. In experiment 1, we contrasted the responses to specular images with those to matte and scrambled images. Activation was observed along the ventral visual pathway, including V1, V2, V3, V4 and the posterior inferior temporal (IT) cortex. In experiment 2, we manipulated the contrasts of images and found that the activation observed in these regions could not be explained solely by the global or local contrasts. These results suggest that image features related to specular surface are processed along the ventral visual pathway from V1 to specific regions in the IT cortex. This is consistent with previous human fMRI experiments that showed surface properties are processed in the ventral visual pathway.

Introduction

The various materials from which objects in our visual world are made, such as metal, ceramic, paper and plastic, have different surface reflectance properties, which enable us to visually discriminate these materials based on their surface characteristics (Adelson, 2001). Among an object's surface properties, gloss perception has been investigated by many psychophysical studies (Doerschner et al., 2010; Fleming et al., 2003; Kim and Anderson, 2010; Motoyoshi et al., 2007; Olkkonen and Brainard, 2010; Wendt et al., 2010). Importance of specular highlights has been recognized for the long time (Beck and Prazdny, 1981) and recent studies have shown that spatial relationships between specular highlights and diffuse shadings or the object shape provide important clue for gloss perception (Kim et al., 2011; Marlow et al., 2011). However, the brain mechanisms underlying gloss perception are still unknown. To understand them, physiological experiments to study actual neural processing of realistic surface properties are required (In this paper, we use the term "realistic surface properties" to refer to complex surface features such as the gloss and texture of real materials that are normally seen in the real world or computer graphics. This is in contrast to rather simple surface properties such as color and luminance).

In a number of neurophysiological studies, the brain responses to the realistic surface properties of objects are now being investigated using fMRI in humans (Peuskens et al., 2004). For example, a series of human fMRI studies by Cant et al. (Cant and Goodale, 2007, 2011; Cant et al., 2009) and by Cavina-Pratesi et al. (2009, 2010) has shown that the ventral higher-order visual areas, around the collateral sulcus and the fusiform gyrus, are possible brain loci in which the surface properties of materials are processed. In line with this, Hiramatsu et al. (2011) showed that patterns of fMRI responses in ventral higher-order visual areas correlate with perceptual similarities of categories of surface materials measured using the semantic differential scaling method, whereas patterns of responses in early visual areas can be explained by simple image features such as spatial frequency or color (Hiramatsu et al., 2011). These results suggest that information regarding surface properties is processed along the ventral stream.

Many earlier electrophysiological studies in the macaque monkey have also shown that neurons in the ventral stream are selective for artificial textures or colors (Komatsu and Ideura, 1993; Komatsu et al., 1992; Liu et al., 2004; Richmond et al., 1987; Wang et al., 2003). Until recently, however, there had been few studies that addressed the representation of realistic surface properties such as gloss or material (Köteles et al., 2008; Nishio et al., 2012). Nishio et al. (2012)

* Corresponding author at: Division of Sensory and Cognitive Information, National Institute for Physiological Sciences, Myodaiji, Okazaki 444-8585, Japan. Fax: +81 564 55 7865.
E-mail address: ngoda@nips.ac.jp (N. Goda).

1053-8119/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2012.07.052

found that neurons in the lower bank of the superior temporal sulcus (STS), within the inferior temporal (IT) cortex, a ventral higher-order visual area in the macaque, exhibit selectivity for stimuli defined by particular parameters of gloss (specular reflectance, diffuse reflectance and roughness) that could not be explained by responses to colors or luminance (Nishio et al., 2012). In addition, Köteles et al. (2008) identified neurons selective for surface materials in the anterior IT cortex, while these neurons were also modulated by the direction of illumination. In other brain regions including the early visual areas, no studies directly examined the responses to realistic surface properties, except for a study of V4 by Arcizet et al. (2008). Consequently, little is known about which parts of visual cortex play major roles in the perception of realistic surface properties.

In the present study, we endeavored to use fMRI to examine the responses to realistic surface properties in the macaque visual cortex. Our focus was on surface gloss because it is one of the most basic surface properties which strongly affects the appearance of surfaces and is commonly attractive to humans, animals and even infants (Yang et al., 2011). Furthermore, gloss perception has received a lot of attention in the literature recently (Doerschner et al., 2010; Fleming et al., 2003; Kim and Anderson, 2010; Motoyoshi et al., 2007; Olkkonen and Brainard, 2010; Wendt et al., 2010). There are several aspects describing surface gloss (Hunter and Harold, 1987) and in the common physical reflectance model (Ward, 1992), three parameters (diffuse reflectance, specular reflectance, and roughness) are used to characterize surface reflection. Among them, the specular reflectance is very important in controlling the glossy appearance. Increasing the specular reflectance makes the object's appearance glossy or, with higher values, shiny, while objects with no specular reflectance never appear glossy. We therefore especially focused on the specular reflectance in this study and hypothesized that in some brain regions neurons preferentially respond to images with high specular reflectance. To test this hypothesis, we designed an experiment to compare the brain responses to images of specular objects with the responses to images of non-specular objects. The results showed that activations sensitive to specular surfaces are observed in wide regions along the ventral visual pathway including early visual areas and some regions in the IT cortex.

Materials and methods

Subjects

Two male macaque monkeys were used in this study (M1 and M2; *Macaca fuscata*, 6–7 kg). Before the experiments, each monkey was implanted with a magnetic resonance-compatible headpost, which was anchored to the skull using dental acrylic and small ceramic screws (Uwe Thomas Recording, Giessen, Germany). All surgical procedures were performed while the animal was under general anesthesia with intravenous injection of Nembutal (the total dose was 39 mg/kg for monkey M1 and 24 mg/kg for monkey M2) following sedation with intramuscular injection of ketamine hydrochloride. After surgery, the monkey was allowed to recover for at least one week before training began. During this period, an antibiotic (Cefazolin, 60 mg) was given every 12 h. All experimental procedures were conducted in accordance with NIH guidelines and were approved by the Animal Experiment Committee of Okazaki National Research Institutes.

Each monkey was trained to fixate on a central fixation spot (a square fixation window 2–3° on a side) for several seconds (typically for 3 s). The monkey had to maintain fixation to get a liquid reward; if the monkey made a saccade, the trial was terminated without a reward.

Apparatus

During experiments, each monkey was seated in the “sphinx” position in a horizontally oriented, custom-made monkey chair as originally described by Vanduffel et al.(2001). Using the headpost, the monkey's head was rigidly fixed to a head holder on the chair. Visual stimuli were generated using a VSG 2/5 graphics board with 15-bit resolution (Cambridge Research Systems Ltd., Rochester, England) and projected from a LCD projector (800×600 pixels, 60 Hz refresh rate; Victor, Yokohama, Japan) onto a screen positioned in front of the monkey's eyes at a distance of 54 cm. The display system was calibrated by measuring the spectral power distributions of the red, green and blue primaries using a spectrophotometer (PhotoResearch PR-650 SpectraScan, Chatsworth, CA, USA). Eye movements were recorded using an infrared (IR) eye tracking system (Matsuda et al., 2000), which computed horizontal and vertical eye positions based on the center of the pupil in an image of the eye captured using an IR-sensitive CCD video camera (60 Hz interlace; Sony, Tokyo, Japan). IR light was directed at an eye using a fiber optic cable. Body movements were automatically detected in images of the monkey's body captured by a CCD camera (Akizuki Denshi Tsusho Co., Ltd., Tokyo, Japan). The eye tracking and motion-detection systems sent signals to a computer, which ran custom-made software controlling the behavioral task and stimulus display.

Visual stimuli

Experiment 1

In experiment 1, we prepared four stimulus conditions, specular (S), matte (M), scrambled specular (SS) and scrambled matte (SM), to identify the brain regions preferentially responding to specular surface (Fig. 1A).

Specular and matte stimuli. The specular and matte stimuli were created using a LightWave renderer (NewTek, San Antonio, TX, USA). A three-dimensional amorphous shape created by the experimenter was rendered under high dynamic range illumination (Eucalyptus Globe) photographed in a real scene obtained from Paul Debevec's database (Debevec, 1998). The surface properties of the objects were controlled using the surface editor in LightWave to vary the diffuse reflectance, specular reflectance, and roughness. Diffuse reflectance determines the lightness of surfaces, whereas specular reflectance determines the strength of the specular highlights on the surfaces. The roughness defines the blur of specular highlights. To obtain specular stimuli, we set the diffuse and specular reflectances at 0% and 100%, respectively, which made the surface appear metallic. The roughness was fixed at zero to render the object with sharp highlights. To obtain matte stimuli, we set the diffuse and specular reflectances at 90% and 0%, respectively. The stimuli were shown as a 2-s movie in which the object rotated leftwards or rightwards for about 70° around a vertical axis through the center of the object. The rotating object continuously changed the position of its contour and specular highlights. It is also known that moving objects are perceived to be glossier than static ones, even when they are rendered with the same reflectance parameters (Doerschner et al., 2011; Sakano and Ando, 2010; Wendt et al., 2010). The mean luminance of each frame was set 20 cd/m^2. The images were 5×5° squares and pixels outside the object were filled with white Gaussian noise with a mean luminance of 20 cd/m^2 and a standard deviation (SD) of 5 cd/m^2. The background of the images was uniformly gray (20 cd/m^2).

Phase-scrambled stimuli. The specular surface contains spatial patterns with high luminance contrast and high spatial frequency components in restricted local regions with high 3D curvatures, whereas the matte surface is relatively uniform in luminance. To dissociate the activities evoked by specular surface from those due to the presence of such local pattern on the surface, we designed non-specular stimuli whose amplitudes of spatial frequencies were largely maintained at the local level but the phases were randomized for both the specular and matte stimuli. Conventionally, Fourier transformation is used to randomize the

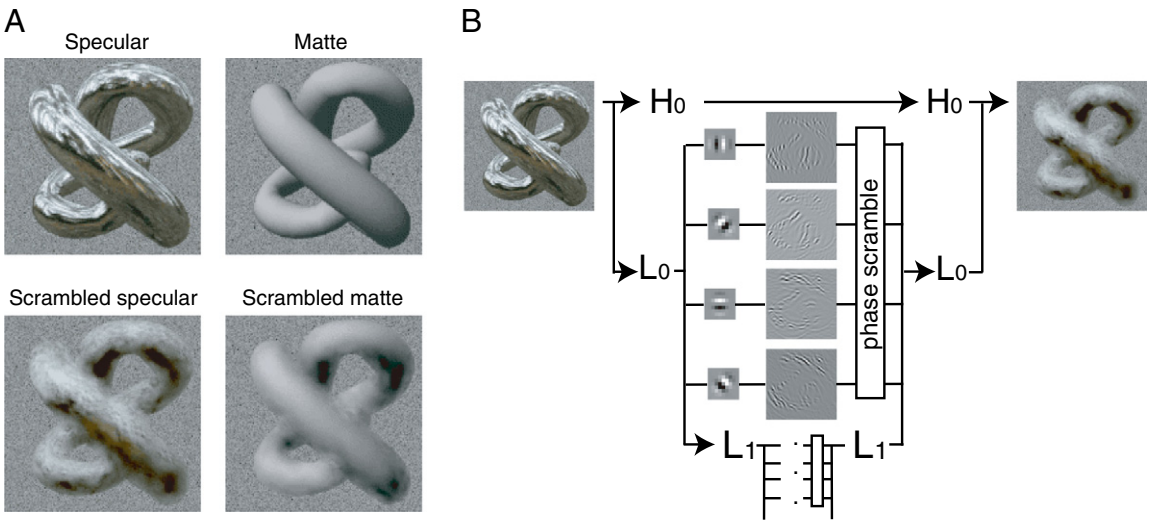

Fig. 1. (A) Specular (S), matte (M), scrambled specular (SS) and scrambled matte (SM) stimuli used in experiment 1. These stimuli were shown as movies in which the object rotated (see text for details). (B) Method for generating scrambled stimuli. An input frame was first high-pass (H_0) and low-pass (L_0) filtered; then the low-passed image was filtered using gabor-like oriented filters. The phases of the filtered images were then scrambled, after which the final image was obtained by reconstruction from the scrambled images.

phase of images as a control stimulus in fMRI experiments (Andrews et al., 2010; Malach et al., 1995). We did not think this procedure fit as a control condition for specular stimuli because the Fourier-based method works on images globally and thus dramatically changes object shape and local spatial frequency contents. We therefore sought to scramble images locally using the wavelet pyramid method (Adelson et al., 1984; Heeger and Bergen, 1995; Portilla and Simoncelli, 2000), which was also employed in a recent electrophysiological study (Rust and Dicarlo, 2010). In brief, the wavelet pyramid method enables one to extract the powers and phases of the spatial frequencies of images locally by using wavelet filters with several different scales. In this study, we employed gabor-like oriented filters with four orientations and five scales using steerable pyramid tools provided by Portilla and Simoncelli (2000) (Fig. 1B). The central spatial frequencies of the filters were 0.97, 1.9, 3.8, 7.6, and 15 cycles/°. We first divided the images using each filter to generate each wavelet band. We then scrambled the phase of each band by adding a random value to each phase angle and reconstructed the images using these phase-randomized bands. The high and low frequency residuals were not phase-randomized. Because this manipulation blurs the outline of the object, we masked the scrambled images with the object's outline from its original image and filled pixels outside the object with white Gaussian noise. Fig. 1A shows examples of images generated for a specular stimulus (top left) and for a matte stimulus (top right) as well as for their scrambled counterparts (bottom left and right). The example movies of these stimuli can be found in Supplementary movies 1, 2, 3, and 4 for specular, matte, scrambled specular, and scrambled matte conditions, respectively. The procedure described here marginally affected amplitudes of spatial frequencies, partly because the steerable filters are, unlike orthogonal wavelet transforms, overcomplete (Simoncelli et al., 1992). The effect of this mismatch of amplitudes is discussed in the Discussion section and Supplementary analysis 2.

As described above, specular and matte stimuli were presented as a movie. We also presented scrambled specular and scrambled matte stimuli as movies by phase-randomizing all of the frames of the original movies. If each frame were scrambled independently, however, the generated movies would appear to flicker at high frequency because the phase relationships between different frames would be broken. Therefore, we maintained these phase relationships by adding the same random value to the phases of all the frames. The mean luminance of each frame was 20 cd/m^2 under all four conditions. To evaluate the contrasts of the images, we calculated the RMS (root mean square) contrast of each image as

$$\text{RMS contrast} = \text{SD}/\text{I}_{\text{mean}}$$

where SD and I_{mean} indicate the standard deviation and mean of the image histograms, respectively. The white Gaussian noise surrounding the object was not included in this computation. RMS contrasts averaged across all the frames in the movies were 1.1, 0.76, 0.81 and 0.65 for the specular, matte, scrambled specular and scrambled matte conditions, respectively. The visual angle of the images was 5°.

Experiment 2

The phase scrambling method described above breaks the phase relationships between different scales, which results in blurred images with reduced image contrast. Therefore, images in the scrambled specular condition have lower RMS contrasts than those in the specular condition. To examine the effect of contrast, in experiment 2 we directly manipulated the RMS contrast of images by adjusting the SD of the image luminance using the following equation,

$$\text{I}_o = (\text{I}_i - \text{I}_{\text{mean}})/\text{SD}_i * \text{SD}_o + \text{I}_{\text{mean}}$$

where I_i and I_o are input and output images, I_{mean} is the mean luminance (set to 20 cd/m^2), SD_i is the standard deviation of the input images, and SD_o is the targeting standard deviation. We set SD_o to 15 and 11 cd/m^2 to prepare two levels of RMS contrasts (0.75 and 0.55, respectively). Again white Gaussian noise surrounding the object was not included in the calculation. Each level of RMS contrast was employed to generate the specular and scrambled specular stimuli, and a total of four types of stimuli were made: high (H) and low (L) contrast images in both the specular (S) and scrambled specular (SS) conditions (HS, LS, HSS, and LSS, see Fig. 4A). These RMS contrasts were selected so as not to degrade the shiny appearance in the specular condition. The visual angle, viewing distance and time of presentation were the same as in experiment 1.

Scanning

Images were acquired using a Siemens 3 T Allegra scanner (Siemens, Erlangen, Germany). Functional images were collected using a gradient-

echo echo-planar (GE-EPI) pulse sequence sensitive to blood oxygenation level-dependent (BOLD) contrast (TR 2 s, TE 20 ms, flip angle 80°, 1.25 mm in-plane resolution, slice thickness 1.6 mm, slice gap 0.32 mm, superior–inferior phase-encoding direction). A surface coil (9×11 cm inner diameter; Takashima Seisakusho, Tokyo, Japan) was positioned immediately over the head. Each volume consisted of 31 coronal slices, covering the occipital and temporal cortices. T2-weighted anatomical images (inversion recovery turbo spin-echo, 0.75 mm in-plane resolution) were scanned at the same locations as those used for the functional images. High-resolution anatomical images (0.5 mm isotropic voxel) were also collected in a separate session using 3D magnetization-prepared rapid-acquisition gradient-echo (MPRAGE) pulse sequence while the monkey was anesthetized with Nembutal. The high-resolution anatomical images were placed in stereotaxic space (the origin was placed at the middle of the interaural line (Saleem and Logothetis, 2007)) and used to reconstruct the cortical surface of each hemisphere using CARET software (Van Essen et al., 2001).

The monkeys performed a fixation task. Each trial began with the onset of a small central fixation spot, on which the monkeys had to fixate. After they had maintained fixation within a 2°×2° square fixation window for 400 ms, a stimulus was presented for 2000 ms. The fixation spot then disappeared 200–400 ms after the stimulus offset. Each trial was continued even when a saccade or body movement was detected so that stimulus duration was kept constant across trials. During the experiments, a reward was given in all trials to maintain the motivation of monkeys; during training, however, a reward was given only when the monkey maintained fixation during the fixation period. The fixation performance and head movements were analyzed offline, and data from runs in which the performance was poor or there was an excessive amount of head movement were discarded (see Data analysis section).

In all experiments, we used a block design in which each stimulus block consisted of four trials with the same stimulus condition. There was an intertrial interval (ITI) of more than 700 ms, during which no fixation spot appeared. The four stimulus conditions were arranged in a pseudorandom order and interleaved with blank blocks consisting of four trials in which only the fixation spot appeared. Each stimulus block was repeated two times within each run.

Localizer fMRI experiments

To identify the areal boundaries within the early visual cortex and between known functional subdivisions in the IT cortex, we conducted a retinotopic mapping experiment and face/place/object localizer experiments. The retinotopic-mapping run consisted of four blocks of horizontal and four blocks of vertical wedges interleaved with blank blocks. The wedges were uniform or checkerboard and extended 0.5 to 9° in eccentricity. Based on these data, the areal boundaries of the early retinotopic areas (V1, V2, V3 and V4) were determined using horizontal and vertical meridians (Fize et al., 2003). The face/place/object localizer run consisted of separate 16-s blocks in which achromatic images of monkey faces, places (landscapes), objects (fruits and man-made tools) and grid-scrambled objects were presented (5°×5°) interleaved with 16-s fixation-only blocks. Each block consisted of four trials and two of the sixteen stimuli prepared for each condition were presented within a trial. The presentation period of one stimulus was 900 ms interleaved by 200 ms blank period. We also conducted an experiment to localize motion-responsive regions by contrasting responses to moving random dots with responses to stationary dots. The moving dots expanded or contracted in a 0.5 Hz cycle at a speed of 4.5°/s. The dots covered 4.5° of visual angle centered at the fixation point and were presented for 2 s in each trial. In all experiments, monkeys performed the fixation task and completed at least 10 runs of each experiment.

Data analysis

Functional data were analyzed using SPM2 (Wellcome Department of Imaging Neuroscience, London, UK) as well as custom Matlab codes. The first and last several volumes (in blank block) in each run were eliminated to allow for stabilization of the magnetization. All functional images were motion-corrected, aligned with a reference functional image from the first session, and registered to the anatomical image of the same-slice position. The functional–anatomical registration was improved by adjusting the offset in the phase-encoding direction. The images were then registered to the high-resolution anatomical image in the stereotaxic space and resampled in 1.0-mm isotropic voxels. The data were temporally filtered with a 3-point temporal median filter to remove movement-related outliers (Mazaika et al., 2007), and were spatially filtered using a Gaussian filter (3 mm full width at half maximum) to improve signal-to-noise ratios.

The functional data from each run were used for analysis only if the monkey's head did not move too much (less than 2% of volumes were affected by head movements whose translation/rotation was over 6 times larger than the standard deviation of the total translations/rotations), and the fixation performance was good throughout the run (the overall percentage fixation from the onset of the fixation spot to its offset was more than 95%). With these criteria, we retained more than 22 runs for each monkey in each experiment (M1: 52 for Exp 1 and 23 for Exp 2; M2: 58 for Exp 1 and 29 for Exp 2).

Whole brain analysis

We conducted voxel-wise statistical analysis for each monkey based on a general linear model (Friston et al., 1995). To detrend the data, they were high-pass filtered (2 cycles/total scan time in each run). The signal time course was modeled using a boxcar function convolved with the macaque BOLD hemodynamic response function measured by Leite et al. (2002) and a run effect. For each stimulus condition, the regressors were set at the onset of the stimulus presentation. Head-movement parameters were also included as regressors of no interest. To test hypotheses involving regionally specific condition effects, the estimates for each model parameter were compared with the linear contrasts. The resulting set of voxel values constituted a statistical parametric map of the t statistic. The statistical parametric map was projected onto the surface of each hemisphere using the CARET 'average voxel' method with a 1.25-mm averaging radius.

In experiment 1, we prepared four stimulus conditions: specular (S), matte (M), scrambled specular (SS), and scrambled matte (SM) (Fig. 1A). To identify the brain regions preferentially responding to specular surface, we carried out a stringent conjunction analysis in which voxels were labeled as being active only when their activities in the specular condition were significantly higher than in all other conditions ((S>M) and (S>SS) and (S>SM)). The conjunction we performed relies on the conservative "conjunction null hypothesis," in which a significant conjunction means that all contrasts were individually significant (Friston et al., 2005; Nichols et al., 2005). Although the three contrasts are not independent of one another, the method remains valid (Nichols et al., 2005). The t statistic obtained by the analysis of each voxel was thresholded at $P<0.001$ and corrected for multiple comparisons at the cluster-level ($P<0.05$, minimum cluster size; 57 voxels for M1, 55 for M2). To exclude voxels that did not show positive responses to visual stimuli, the voxels were masked when their visual responses (the average responses under all four conditions) were below the threshold level ($P=0.001$, uncorrected for multiple comparisons).

Regions of interest (ROIs) analysis

We conducted region of interest (ROI) analyses to further examine the responses in regions identified in experiment 1 and in independent localizer experiments. For regions identified in the conjunction analysis in experiment 1, we collected the voxels whose activation

was significant in each retinotopically-defined region (V1, V2, V3 and V4) and in the posterior and central IT cortices. The significance level was set as $P<0.05$, corrected for multiple comparisons at the cluster level unless otherwise stated. The border between the posterior and central IT cortices was determined based on the areal partitioning scheme of Felleman and Van Essen (1991). Average time courses sorted by block onsets from the experimental data were then extracted from these ROIs, and the mean response amplitude (%BOLD change) for each stimulus condition was computed by averaging signals between 6 and 20 s after the block onset.

In experiment 2, we extracted the %BOLD change for each of the four stimulus conditions; i.e., high contrast specular (HS), low contrast specular (LS), high contrast scrambled specular (HSS) and low contrast scrambled specular (LSS) (Fig. 4A). To dissociate the effects of specularity and contrast on cortical responses, we conducted a two-way repeated-measures analysis of variance (ANOVA) with luminance contrast (high vs. low contrasts) and specularity (specular vs. scramble) as within-subject factors. As a measure of the amount of modulation by each factor, we also calculated specularity and contrast indices as follows.

$$\text{Specularity index} = ((HS + LS) - (HSS + LSS))/(HS + LS + HSS + LSS)$$
$$\text{Contrast index} = ((HS + HSS) - (LS + LSS))/(HS + LS + HSS + LSS)$$

where HS, LS, HSS and LSS respectively indicate %BOLD changes averaged across hemispheres for the high contrast specular, low contrast specular, high contrast scrambled specular and low contrast scrambled specular conditions in experiment 2.

In this experiment, we also examined the responses in other brain regions, including the face/object-responsive regions of the IT cortex (identified in the localizer experiment), the motion-responsive region in the STS and the visually-responsive region in the intraparietal sulcus (IPS). Face-responsive regions were identified by contrasting responses to faces with those to objects. Object-responsive regions were identified by contrasting responses to objects with those to scrambled objects. Motion-responsive regions were identified by contrasting the responses to moving dots with those to stationary dots in another localizer experiment (Vanduffel et al., 2001). The motion-responsive regions peaked at the floor of posterior bank of the STS (about 1 mm posterior from the interaural line). Because these regions would contain several subdivisions (Nelissen et al., 2006), we call this as MT+ here. To calculate %BOLD signal changes for these regions, we collected the voxels within 2 mm distance from the activation peak for each contrast. The visually-responsive region in the IPS was defined as the voxels within 2 mm distance from the peak visual responses (the average responses under all four stimulus conditions) in the IPS.

Calculation of local luminance contrasts of images

To further examine the possibility that the difference in cortical responses observed in experiment 2 could be explained by the difference in local luminance contrasts of the images, even when the global RMS contrasts were closely matched, we calculated the local contrasts within images using the definition provided in Mante et al. (2005). In this definition, a weighting function was defined for each pixel (x_i, y_i) to calculate the local contrast at a particular position (x_c, y_c) within the image

$$w_i = \cos\left(\frac{2\pi}{d}\sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}\right) + 1$$

where d is the diameter of the patch. After normalizing this weighting function to sum to 1, the local contrast at that position can be calculated as follows.

$$c = \sqrt{\sum_{i=1}^{N} w_i \frac{(L_i - L)^2}{L^2}}$$

where L_i indicates the luminance at pixel (x_i, y_i) and L indicates the mean local luminance

$$L = \sum_{i=1}^{N} w_i L_i.$$

By changing the parameter d, we can compute the local contrast at different spatial frequencies. To determine the amount of local contrast under each stimulus condition, we averaged the local contrasts for all positions and stimulus frames in each condition. Because earlier studies showed that early visual areas are responsive within the range between 0.5 and 10 cycles/° (Henriksson et al., 2008), we set d such that the spatial frequencies of the weighting function were 0.5, 1, 2, 5 and 10 cycles/°.

Results

Experiment 1: activation elicited by specular objects observed in conjunction analysis

To localize the brain regions responding to specular surface, we presented specular (S), matte (M), scrambled specular (SS), scrambled matte (SM) images on a screen while monkeys fixated on a central fixation spot. If an area solely represents the three-dimensional (3D) shape or contours of a stimulus, it should respond equally under both the specular and matte conditions when the same 3D shape was presented. On the other hand, if an area is sensitive to differences in low-level image features, such as the presence/absence of intricate patterns formed by specular reflection on the surface, the responses in the specular and scrambled specular conditions should be similar because the stimuli in those two conditions have patterns with similar spatial frequencies on the surface. If an area preferentially responds to specular surface, the responses in the specular condition should be stronger than those in the other conditions. We therefore carried out the conjunction analysis which select voxels whose activities in the specular condition were significantly higher than in all three other conditions to identify the voxels responsive to specularity, but not surface shapes or surface patterns. The conjunction is based on the conservative "conjunction null hypothesis"; therefore a significant conjunction means that all contrasts were individually significant. Fig. 2 shows a representative result obtained from the left hemisphere of monkey M1. Activation detected through conjunction analysis was mapped onto the reconstructed cortical surface in Fig. 2A. The activities overlaid on the surface show the voxels whose responses were significantly stronger in the specular condition than in the other three conditions ($P<0.05$, corrected for multiple comparisons at the cluster-level). The areal boundaries of V1, V2, V3 and V4 (green dashed lines) were determined in a separate retinotopic mapping experiment (see Materials and methods section). The PIT/CIT border was determined based on the areal partitioning scheme of Felleman and Van Essen (1991).

Activation was observed in the early retinotopic regions (V1, V2, V3 and V4), at the lip of the STS in the posterior IT cortex (PITd, green arrow), in the lateral convexity of the IT cortex anterior to the posterior middle temporal sulcus (PMTS) (PITv, blue arrow), and at the bank/lip of the STS in the central IT cortex (CIT, red arrow). Fig. 2B shows the average time courses of BOLD responses from the onset of blocks around the voxel showing significant activation in each region in the left hemisphere of M1. The activation in all the regions clearly shows the responses to be strongest in the specular condition; responses elicited under other conditions were similar, and were smaller in amplitude than the specular condition.

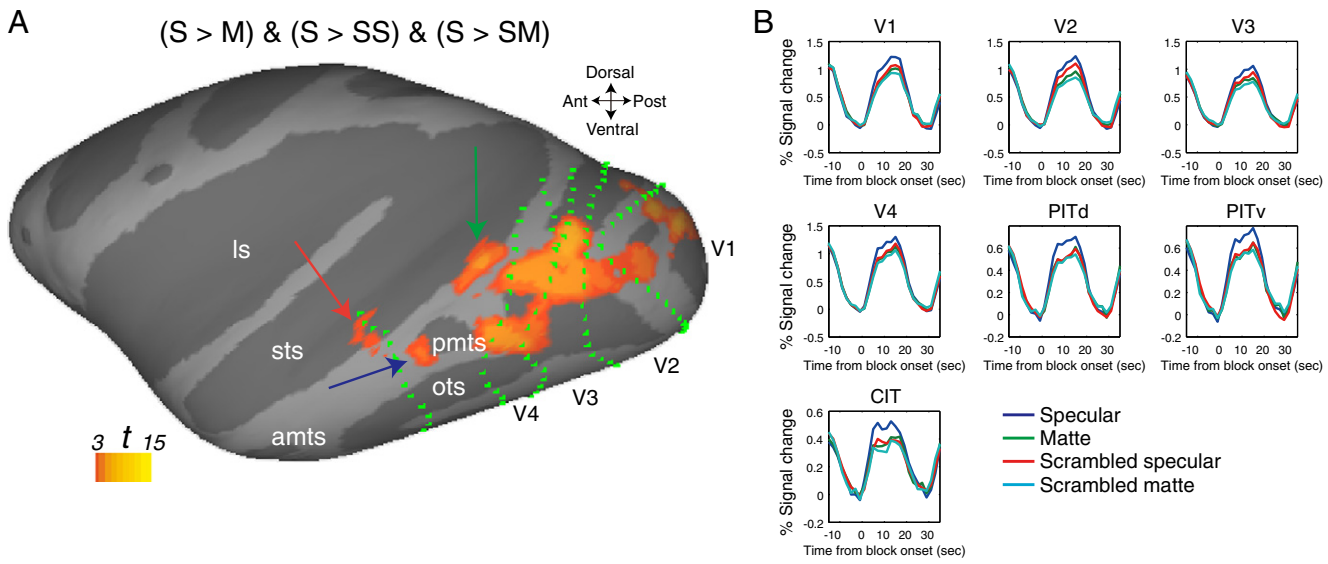

Fig. 2. (A) Regions more responsive to specular stimuli than to other stimulus types identified by conjunction analysis in the left hemisphere of monkey M1. The statistical parametric map was superimposed on the inflated surface. The statistical threshold was $P<0.05$, corrected for multiple comparisons at the cluster-level. The color scale indicates the *t*-score. Green dashed lines indicate the areal boundaries of V1, V2, V3, V4, PIT and CIT. All of the boundaries except the PIT/CIT border were determined in a separate retinotopic mapping experiment. The PIT/CIT border was determined based on the areal partitioning scheme of Felleman and Van Essen (1991). Sulci: ls: lateral sulcus, sts: superior temporal sulcus, amts: anterior middle temporal sulcus, pmts: posterior middle temporal sulcus, ots: occipito-temporal sulcus. (B) Time courses of %BOLD signals in voxels whose *t*-statistics exceeded the statistical threshold ($P<0.05$, corrected for multiple comparisons) for V1, V2, V3, V4, PITd, PITv and CIT in the left hemisphere of M1. Each colored line indicates the averaged time course of the responses in each condition.

Fig. 3 shows the spatial distributions of the activity elicited in the visual cortex in all four hemispheres of the two monkeys. In the early visual areas, activation was consistently seen in V1, V2, V3 and V4 in all four hemispheres. In the IT cortex, activation in the right hemisphere of M1 was observed over large regions that included the three locations responsive to specular surface (PITd, PITv, and CIT) in the opposite hemisphere. In the other monkey (M2), activation in the IT cortex was restricted in the dorsal part of the posterior IT, around the lip of the STS, which coincides with the PITd observed in M1. Fig. 3B shows coronal sections centered at each activation site in the IT cortex. The activation in the PITd was located about 4 mm anterior to the interaural lines in all of the hemispheres, whereas activation in the CIT of M1 was about 12 mm anterior to the interaural lines. By lowering the threshold ($P<0.005$, uncorrected for multiple comparisons), activation could also be seen in the CIT in the left hemisphere of M2 (Fig. 3C right panel), while activation in the PITv could be seen in both hemispheres (Fig. 3C left panel). In summary, we observed robust activation in V1, V2, V3, V4 and the PITd in all four hemispheres. In the IT cortex, there are likely two additional foci sensitive to specular surface (PITv and CIT), but activation in these regions appears weaker than in the PITd.

To obtain more detailed profiles of the activities, we plotted the %BOLD signal changes for each activated area averaged across all four hemispheres (Fig. 3D). To compute the %BOLD signal change, we collected the voxels whose *t*-statistic in the conjunction analysis was larger than the statistical threshold $P<0.05$ corrected for multiple comparisons at the cluster level from each area, except for PITv of both hemispheres and CIT of left hemisphere in M2 where we set the statistical threshold $P<0.005$ uncorrected for multiple comparisons. Note that in this analysis, the same data were used to localize voxels and to compute the %BOLD signals, so the figure is shown just for the illustrative purpose. Open circles indicate the responses in the specular or scrambled specular condition, while open squares indicate the responses in matte or scrambled matte conditions. In all regions, the lines connecting the specular and scrambled specular conditions (open circles) were steeper than those connecting the matte and scrambled matte conditions (open squares), which indicates that scrambling more strongly affected responses to specular stimuli than matte stimuli. There was also a tendency that the responses to the three conditions other than the specular condition (i.e. M, SS, and SM) were relatively similar in higher visual areas, but dissimilar in early visual areas. When we conducted one-way repeated-measures ANOVA across these three conditions for each area, the significant main effect of conditions was observed in V1, V2, V3, and V4 ($P<0.01$), while it was not observed in PITd, PITv, and CIT ($P>0.05$). This implies that regions in IT cortex reflect the strength of specular reflectance more faithfully than those in earlier visual areas.

We also examined the activation revealed by other comparisons. We found no region that responded more strongly in the matte condition than the specular condition. When we contrasted responses in the specular and matte conditions with those in the scrambled specular and matte conditions ($(S+M)>(SS+SM)$), regions in the IT cortex similar to those activated in the conjunction analysis were activated. Contrasting the matte with the scrambled matte condition did not elicit consistent responses across hemispheres. These are as expected because the regions identified in the conjunction analysis showed stronger responses in the specular condition than in other conditions (Fig. 2B).

Experiment 2: dissociating responses to specularity and contrast

In experiment 1, we localized brain regions that responded to specular surface. As noted in the Materials and methods section, however, our method of phase scrambling reduces the luminance contrast of images both in the global and local levels. The global contrast can be numerically represented as RMS contrast. The value of RMS contrast averaged across all the frames of the movies presented in the specular, matte, scrambled specular, scrambled matte conditions was 1.1, 0.76, 0.81 and 0.65, respectively. It therefore remains possible that some parts of the responses associated with specularity observed in experiment 1 could be explained by differences in the RMS contrasts of the images. To examine this possibility, in experiment 2 we directly matched the RMS contrast of the images and designed experiments to dissociate responses to specularity from those to contrasts by generating high (H) and low (L) contrast images in both the specular (S) and

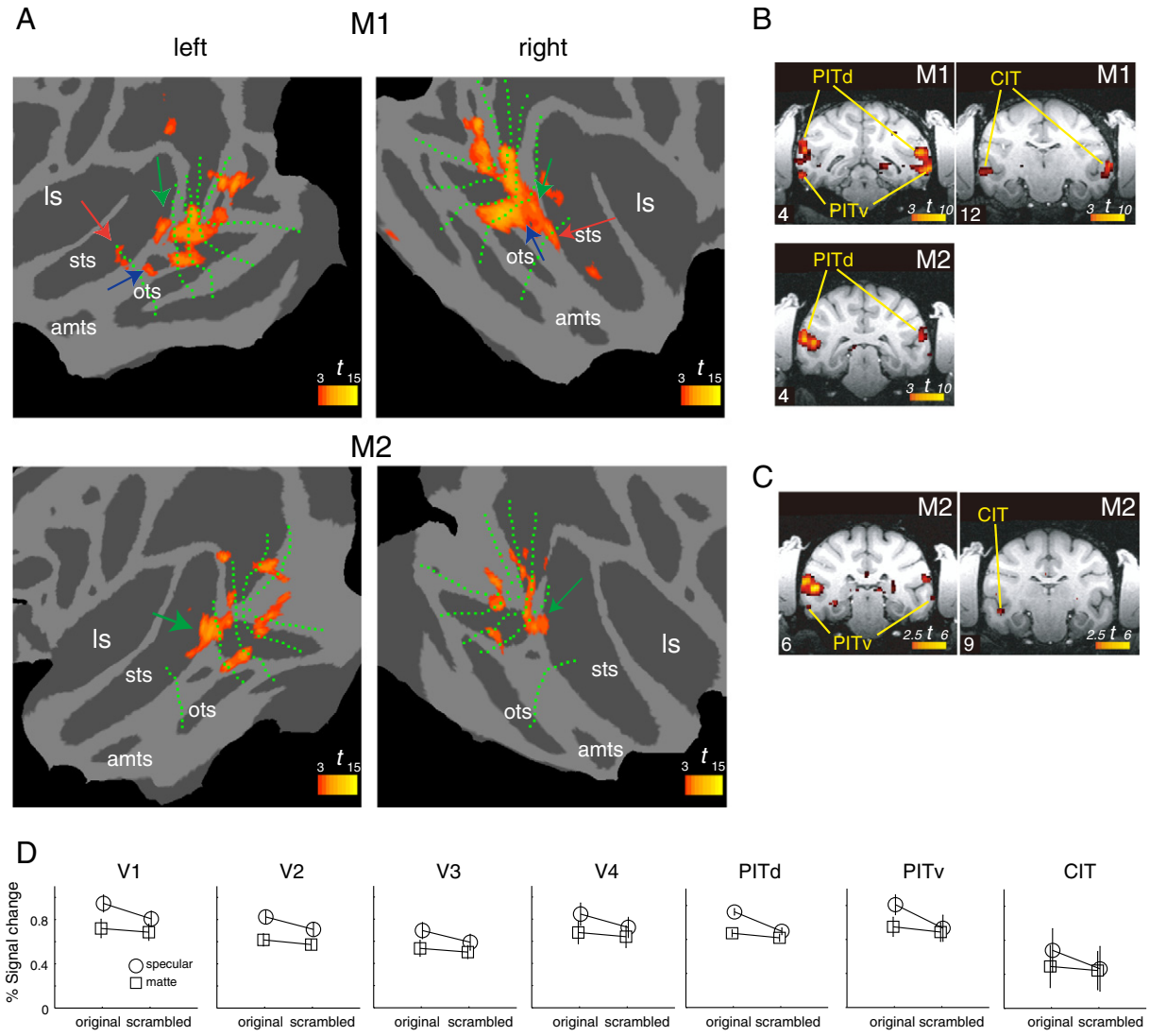

Fig. 3. (A) Results of a conjunction analysis of all four tested hemispheres. The activations were overlaid on the flattened surface of the cortex. The statistical threshold was $P<0.05$, corrected for multiple comparisons at the cluster-level. See Fig. 2 for abbreviations. (B and C) Statistical parametric maps superimposed on coronal sections of each monkey. The statistical threshold was $P<0.05$, corrected for multiple comparisons at the cluster-level (B) or $P<0.005$ uncorrected for multiple comparisons (C). Numbers labeled in the lower left of each panel indicate the locations of the coronal sections in millimeters relative to the interaural line (positive values are anterior). (D) Mean %BOLD signal for each condition taken from each area. Error bars indicate the standard error of the mean across hemispheres. The data from the CIT are averages of three hemispheres; the data from the other regions are averages of all four hemispheres.

scrambled specular (SS) conditions (Fig. 4A). The four conditions (HS, LS, HSS, and LSS) were presented in a block design to the fixating monkeys.

Fig. 4B shows an example of results from the right hemisphere of M1. The activation in the left panel is the response contrast between the sum of the high and low contrast specular conditions and the sum of the high and low contrast scrambled specular conditions ((HS + LS) > (HSS + LSS)), which gives the response to specularity independent of the RMS contrast of the images. On the other hand, the activation in the right panel is the contrast between the sum of the high contrast specular and scrambled specular conditions and the sum of the low contrast specular and scrambled specular conditions ((HS + HSS) > (LS + LSS)), which gives the response to the RMS contrast independent of the specularity. Both activations are shown without statistical threshold and masked with the result of the conjunction analysis in experiment 1. The activation was generally above zero (colored yellow to red) in both maps, indicating that regions identified in experiment 1 actually respond to both specularity and contrast, although there appeared to be a tendency for responses to specularity to be stronger (reddish colors) in the more anterior region, whereas the responses to contrast were stronger in more posterior region. To quantitatively assess the strength of the responses to specularity and contrast, we extracted the %BOLD signal changes for each condition, as noted in the Materials and methods section.

We first examined the responses in the regions where activation to specular surface was observed in experiment 1. From V1, V2, V3, V4 and the PITd, we collected voxels whose t values were significant in the conjunction analysis of experiment 1 ($P<0.05$, corrected for multiple comparisons). Then in experiment 2, we averaged the %BOLD signal changes recorded in these voxels. Fig. 5A shows the %BOLD signal change for all stimulus conditions in all regions. The results from the PITv and CIT, where the activations were not significant at the corrected level in some hemispheres, are shown in the Supplementary analysis 1. We conducted a 2-way repeated-measures ANOVA to examine the main effects of specularity (specular vs. scramble) and contrast (high vs. low contrast). The asterisks in Fig. 5A indicate the level of significance for the main effect of specularity, which was significant in all of these areas ($P<0.05$ for V1, V3, V4 and the PITd; $P<0.01$ for V2). A significant main effect of contrast was observed in V1, V2, V3 and V4 ($P<0.05$ for V1, and V4; $P<0.01$ for V2 and V3) but not in the PITd ($P>0.05$).

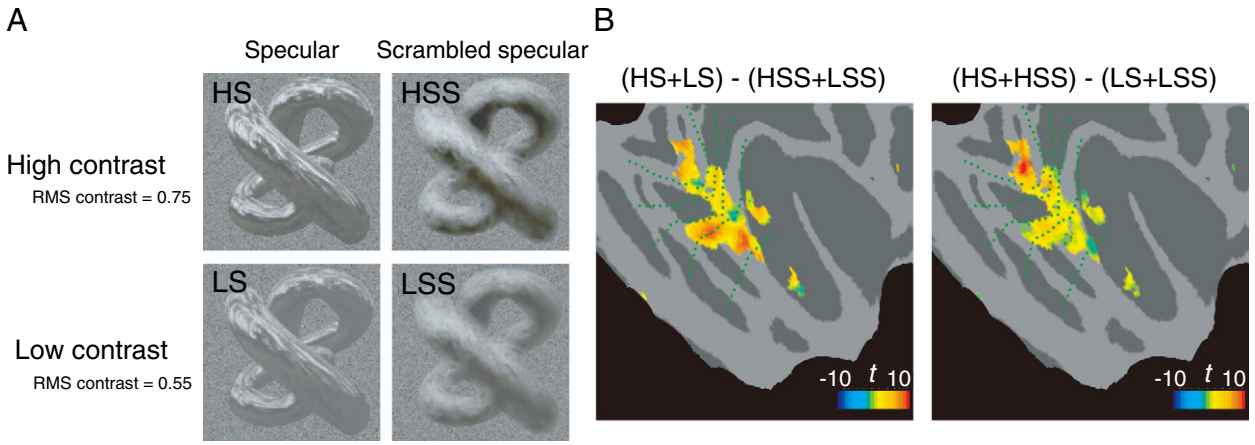

Fig. 4. (A) The four conditions used in experiment 2: high contrast specular (HS), low contrast specular (LS), high contrast scrambled specular (HSS) and low contrast scrambled specular (LSS). (B) Examples of responses observed in the right hemisphere of M1. The color scale indicates the *t*-score. Activation in the left panel is the contrast between the high/low contrast specular and high/low contrast scrambled specular, while activation in the right panel is the contrast between the high contrast specular/scrambled specular and low contrast specular/scrambled specular. Both activations are shown without statistical threshold and are masked with the result of the conjunction analysis in experiment 1.

To visualize the amount of modulation elicited by each factor (specularity or contrast), we computed specularity and contrast indices to compare the relative changes in responses across areas. In brief, the specularity index was the difference between the responses in the specular and scrambled specular conditions normalized by the sum of the responses across all four stimulus conditions, while the contrast index was the difference between the responses in the high and low contrast conditions normalized by the sum of the responses across the four conditions. Fig. 5B shows the obtained specularity and contrast indices. We found that the specularity index was relatively high in V1, was lower in V2 to V4, and increased again in the PITd. On the other hand, the contrast index remained relatively constant across all of these areas. This trend is more clearly seen in Fig. 5C, which shows the relationships between the specularity and contrast indices. V2, V3 and V4 are all positioned near the diagonal line, whereas V1 and PITd are shifted upward, indicating those regions are more sensitive to specularity. Although the activations in the PITv and CIT elicited by specular surface in experiment 1 were only significant ($P<0.05$ corrected with multiple comparisons) in two hemispheres of one monkey (M1), we found that the specularity

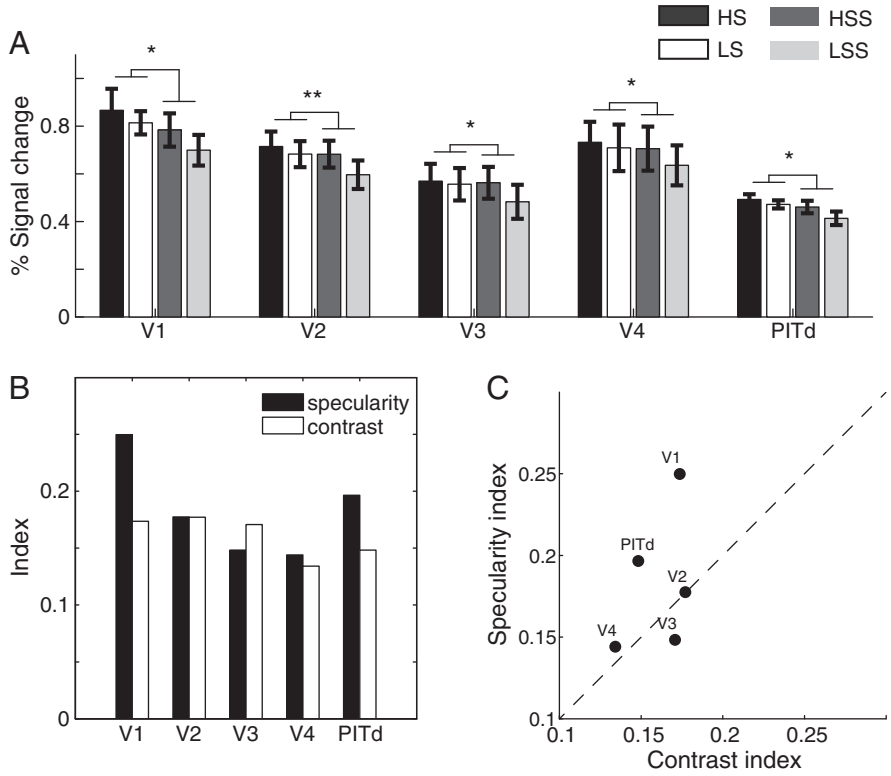

Fig. 5. %BOLD signal changes and the specularity and contrast indices computed from the responses in experiment 2 for each region localized using the activations in experiment 1. (A) Mean %BOLD signal changes for the four conditions in each region localized by collecting all voxels whose activity was above the threshold level in the conjunction analysis in experiment 1. Error bars indicate the standard error of the mean across hemispheres. The asterisks indicate the significance of main effect of specularity (specular vs. scramble). *$P<0.05$, **$P<0.01$ (ANOVA). (B) Specularity and contrast indices for each region. (C) Scatter plot of the specularity and contrast indices for each region.

indices for these areas were large, and the main effect of specularity in the CIT was significant (Supplementary analysis 1 and Supplementary Fig. 1).

The results summarized above indicate that the difference in global contrast of an image cannot explain the responses associated with specularity in experiment 1. However, the procedure used to manipulate the global RMS contrast does not ensure that the contrasts will also be matched at the local levels. For that reason, we verified that the local contrasts cannot explain the responses to specularity by comparing the cortical responses to local contrasts in each stimulus condition. We computed the local contrasts of images using the definition provided in Mante et al. (2005), as noted in the Materials and methods section. Fig. 6A shows the local contrasts of images in all four stimulus conditions at each of the spatial frequencies used for their calculation. When the spatial frequency was low (0.5 cycles/°), there were only tiny differences in the contrasts between the HS and HSS conditions, or between the LS and LSS conditions, as is the case with the global RMS contrasts. But when the spatial frequencies were higher (5 or 10 cycles/°), contrasts in the specular conditions became higher, as expected. Nevertheless, when we looked at the LS and HSS conditions, we found that the local contrasts in the LS condition never exceeded those in the HSS condition. This suggests to us that, if an area is mainly responsive to local contrast, the response should be stronger in the HSS than the LS condition, but if an area is mainly responsive to specular surface, the response should be stronger in the LS than the HSS condition. To determine which one of these two scenarios is actually the case, we replotted the responses in V1 and the PITd identified in experiment 1 and compared the bias observed in the local contrasts with the responses observed in the regions sensitive to specular surface (Fig. 6B, the data are the same as those in Fig. 5A). We found that the cortical responses were consistently higher in the LS condition than the HSS condition, which is opposite to the bias observed in the local contrasts, and the responses in the specular conditions (HS and LS) in these regions were stronger than expected from the local contrasts. This indicates that the significant effects of specularity observed in experiment 2 cannot be explained solely by the local contrasts of the images.

As observed in Fig. 5A, the main effects of specularity were all positive in the early visual areas. However, because the data presented in Fig. 5A were collected from voxels activated in response to specular stimuli in experiment 1, there could be a sampling bias such that the data were selectively sampled from the voxels that correspond to the visual field where image contrasts were locally higher in the specular condition than in the scrambled specular condition. To test that possibility, we reanalyzed the data from the early visual areas by sampling the signal in all voxels showing significant visual responses (the average of four conditions in the experiment, $P<0.001$, uncorrected for multiple comparisons) in the early retinotopic regions (Fig. 7). Using this analysis we found that the main effects of specularity were significantly positive in all of these areas ($P<0.05$, ANOVA), which confirms that the early visual areas are actually responding to features associated with the specular condition.

Comparison with face/object localizer

It has been shown that there are regions specialized for individual categories such as faces and objects within IT cortex (Denys et al., 2004; Pinsk et al., 2005; Tsao et al., 2003, 2006). To examine the spatial relationships between those modules and the specularity-responsive regions observed in the IT cortex, we conducted a face/object localizer experiment (see Materials and methods section). A representative result of this localizer experiment for the left hemisphere of M1 is shown in Fig. 8A. When we contrasted the responses to faces with the responses to objects, two activations were observed in the IT cortex, one located around the central IT cortex and another around the anterior IT cortex. These activation patterns are very similar to those previously described by Bell et al. (2009). When we contrasted the responses to objects with those to scrambled objects, activations were observed in larger regions of the IT and surrounded the responses to faces as found by Tsao et al. (2003). Because place-responsive regions were not clearly observed in our dataset, place responses were not used in the following analyses.

Fig. 8B shows the results of comparing the localizer responses and the activations elicited by specular surface. The activations illustrated in Fig. 8B show the responses to specular surface, which are the same as in Fig. 3. Green and blue lines indicate the outlines of the face- and object-responsive clusters, respectively. For the responses to specular surface in the PITd, the activations elicited by specularity overlapped the posterior part of the object responses (blue lines) but did not overlap the face responses (green lines) in all four hemispheres tested. In the CIT observed in two hemispheres, the activation elicited by specularity overlapped the posterior part of the face responses.

To directly examine the responses to specularity within the localized regions, we first identified peaks for the face and object responses (Fig. 9A). Points labeled face1 and face2 are the peaks of the face responses, with face1 located more posteriorly. Similarly, points labeled obj1 and obj2 are the peaks of the object responses, with obj1 located more posteriorly. We extracted the responses for all four conditions in experiment 2 in these regions (Fig. 9B). Both object peaks (obj1 and obj2) and both face peaks (face1 and face2) did not yield significant main effects of specularity ($P>0.1$, ANOVA). This

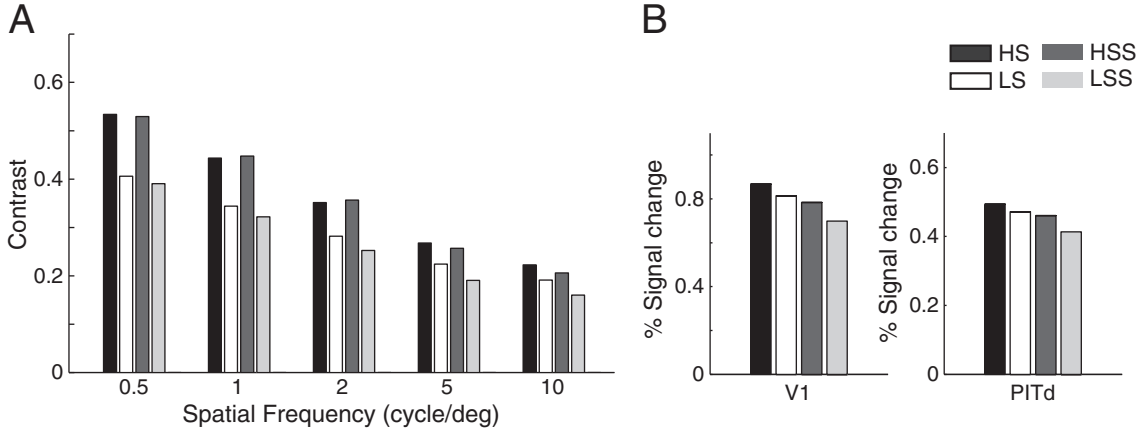

Fig. 6. Comparison of local contrasts of images with responses observed in specularity-responsive regions. (A) Local contrasts of stimuli in experiment 2 at several spatial frequencies (cycles/°). The bars indicate each condition. The method for calculating local contrasts is described in Mante et al. (2005) and in the main text. (B) The %BOLD signal changes for each condition in experiment 2 observed in V1 and the PITd. The data are the same as those presented in (A).

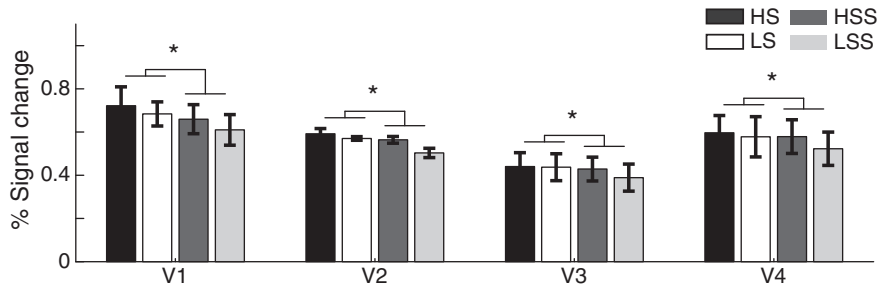

Fig. 7. Mean %BOLD signal changes for the four conditions in experiment 2 for early visual areas localized by collecting all the voxels with significantly positive visual responses ($P<0.001$, uncorrected for multiple comparisons).

implies that, although responses to specular surface overlap both the object and face responses to some degree, regions giving strong responses to object, face and specularity are segregated, which suggests the underlying processes are separated. Significant main effect of contrast was observed in face2.

We also examined the responses in other higher visual areas (Fig. 9C), including area MT+ and the IPS. Although the stimuli presented in experiment 2 were movies of rotating objects, the main effect of specularity in the MT+ was not significantly positive ($P>0.1$, ANOVA). This might be because object contour motions which dominate in the image motions were preserved in the scramble conditions. We also extracted the peak of the visual responses within the IPS. If specular stimuli drew more attention than scrambled specular stimuli, we would expect differential responses in the IPS, because the IPS activity is known to correlate with attentional states and eye movements (Bisley and Goldberg, 2010). However, the data obtained from the IPS showed significant modulation by contrast ($P<0.05$, ANOVA), but not by specularity ($P>0.05$, ANOVA), which suggests attentional modulation did not differ between the specular and scrambled specular conditions.

Discussion

Earlier fMRI experiments carried out in humans showed that there are regions in the ventral higher-order visual areas that are specifically responsive to the realistic surface properties of objects (Cant and Goodale, 2007, 2011; Cant et al., 2009; Cavina-Pratesi et al., 2009,

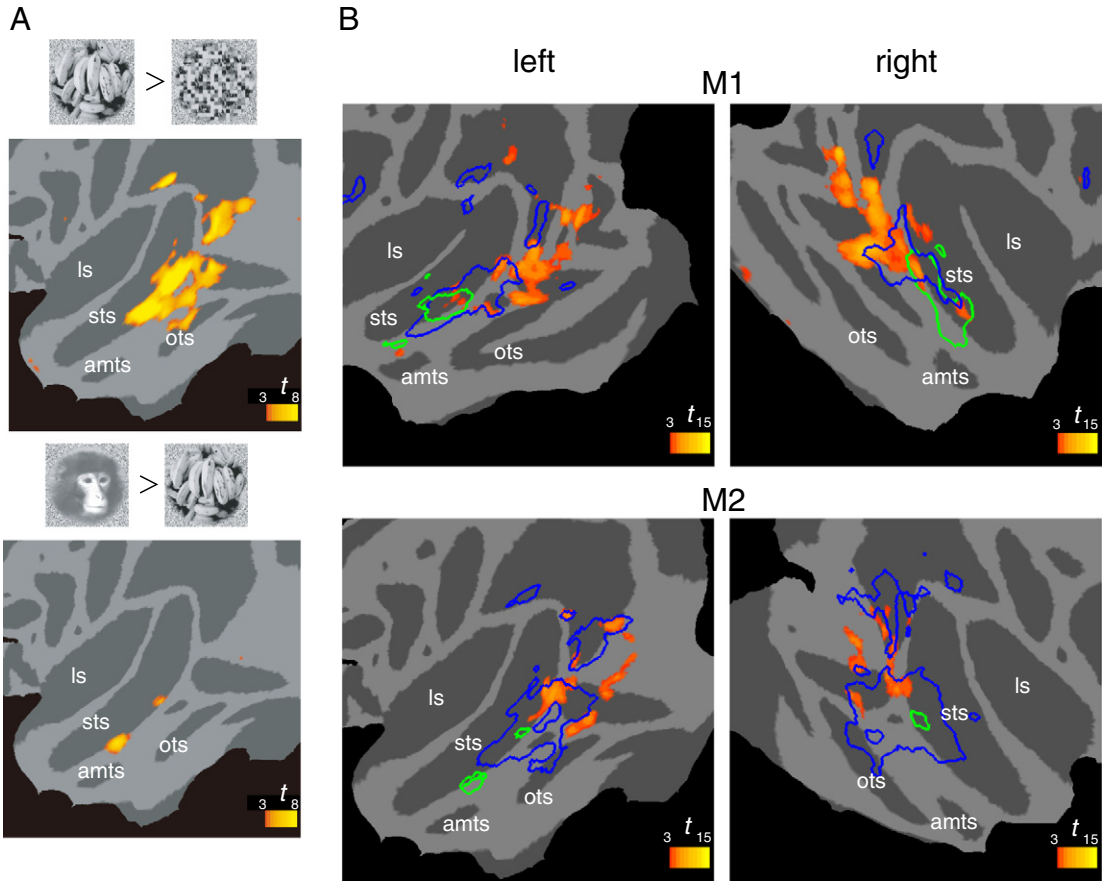

Fig. 8. Comparison of activation elicited by specular surface with those elicited by faces and objects. (A) Representative results of face/object localizer activation in the left hemisphere of M2. The upper panel illustrates the activation revealed by contrasting objects with scrambled objects. Images above the activation map show an example of image used in object and scrambled object conditions. The lower panel illustrates the activation revealed by contrasting monkey faces with objects. Images above the activation map show an example of image used in face and object conditions. See Fig. 2 for abbreviations. (B) The localizer activations are overlaid as green (face) and blue (object) contours on the activation by specular surface observed in experiment 1 for all four hemispheres tested (same as Fig. 3A).

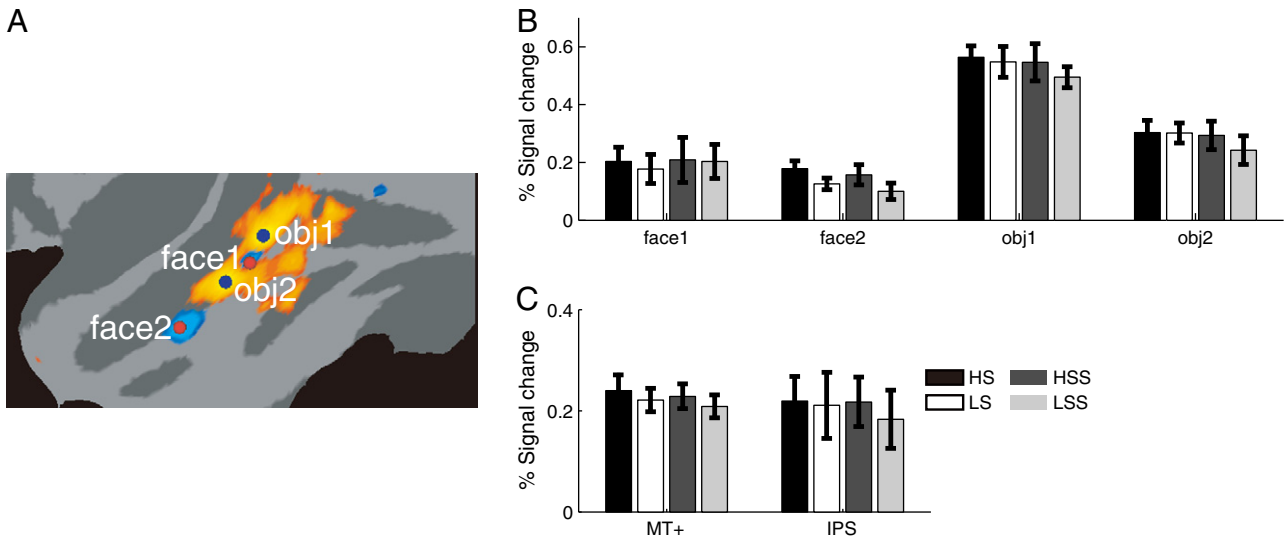

Fig. 9. Responses of the regions defined in the localizer experiment and other visual areas in experiment 2. (A) Two peaks for object responses (obj1 and obj2, blue dots) and two peaks for face responses (face1 and face2, red dots) are overlaid on the activation in the localizer experiment in the left hemisphere of M2. Yellow activations indicate object responses, while blue activations indicate face responses (same as in (A)). (B) Mean %BOLD signal changes for the four conditions in experiment 2 for each peak in the localization experiment. No area showed significant main effect of specularity in ANOVA. (C) Mean %BOLD signal changes for the four conditions in other areas. The MT+ peak was localized in the experiment in which moving dots were presented, and the IPS peak was localized as the peak of visual responses (the average responses under all four stimulus conditions) in IPS. The data are the averages of all four hemispheres. No area showed significant main effect of specularity in ANOVA.

2010; Peuskens et al., 2004) and that representations of surface materials in those areas correlated with human categorical perception of the materials (Hiramatsu et al., 2011). Correspondingly, there are neurons in the macaque IT cortex that selectively responds to images of particular materials (Köteles et al., 2008) or to stimuli with particular parameters of gloss (Nishio et al., 2012). However, because there had been few studies in macaque that directly tested the brain responses to realistic surface properties (Arcizet et al., 2008; Köteles et al., 2008; Nishio et al., 2012), it was not clear which regions of the macaque visual cortex play a major role in the processing and representation of surface properties. Using fMRI, we have now localized the regions responsive to specular surface in the macaque visual cortex. We hypothesized that there are brain regions that preferentially respond to specular surface because the processing of gloss would require unique mechanisms (Fleming et al., 2003; Motoyoshi et al., 2007). To test that hypothesis, in experiment 1 we contrasted the responses to specular images with those to matte and scrambled images, and found that specular sensitive activations distribute in wide regions along the ventral visual pathway including early visual areas (V1, V2, V3, and V4) and some regions in the IT cortex. In experiment 2, we manipulated the RMS contrast of images to dissociate responses to image specularity from responses to image contrast and confirmed that the responses in the regions identified in experiment 1 could not be explained by the global image contrast. We also compared the local contrast across stimuli and concluded that the responses cannot be explained by the difference in local contrasts. To our knowledge, this is the first study that directly identified regions responsive to a realistic surface property in the macaque visual cortex using fMRI.

With respect to the early visual areas, we found responses related to specularity in V1, V2, V3 and V4. As mentioned above, these responses cannot be explained by the global contrast or by local contrasts in the images, although there remains a possibility that certain combination of these low-level features may explain the responses. We also have to note that the amplitudes of spatial frequencies are not precisely matched between the specular and scrambled specular conditions because of the properties of the scrambling method we used (see Materials and methods section). To evaluate the mismatch of amplitudes, we calculated the amplitude of each spatial frequency band for stimuli in experiment 2 (see Supplementary analysis 2 for details). The evaluated normalized amplitudes for high contrast specular condition (HS) were 1.0, 0.91, 0.66, 0.58, and 0.58 for 0.97, 1.9, 3.8, 7.6, and 15 cycles/°, respectively (the values are normalized by the amplitude of coarsest scale (0.97 cycles/°) for HS condition). The normalized amplitudes for high contrast scrambled specular condition (HSS) were 1.3, 1.1, 0.41, 0.33, and 0.37. Therefore, the amplitude tended to decrease by the scrambling in the higher frequency bands, whereas increase in the lower frequency bands if we compared stimuli with the matched RMS contrast. This raises a possibility that the reduction of high spatial frequency components may explain the difference in responses observed in this study. However, this is not likely the case considering the spatial frequency tunings of BOLD responses in the visual cortex. Several fMRI studies have shown that early visual areas are most sensitive to relatively low spatial frequency (typically peaked at about 1–2 cycles/°) and are less sensitive to high spatial frequency (7 cycles/° or higher) (Henriksson et al., 2008; Rajimehr et al., 2011; Sasaki et al., 2001; Singh et al., 2000). Moreover, the amplitudes of higher frequency were relatively smaller in our stimuli. Thus, the reduction of higher frequency component is likely to have little effect on the difference in the responses observed in this study. Furthermore, an additional analysis confirmed that the specularity-induced activations were not restricted in the regions representing central visual field which are more sensitive to high spatial frequencies than those representing more peripheral visual field (Henriksson et al., 2008; Rajimehr et al., 2011; Sasaki et al., 2001); the effect of the scrambling on the responses was even consistently larger in the peripheral regions than in the central regions (Supplementary analysis 2 and Supplementary Fig. 2). Taken together, the reduction of high frequency component alone could not explain the difference in the responses observed in this study.

One major possible image features explaining the responses in the visual cortex may be phase congruencies across different spatial scales that should be destroyed by the phase-scrambling. In terms of gloss perception, Kim et al. (2011) and Marlow et al. (2011) have found that glossiness depends on the congruency of positions and orientations between specular highlights and shadings. Because previous fMRI studies have shown that early visual areas have sensitivity to cross-frequency spatial phase relationships (Henriksson et al., 2009;

Perna et al., 2008; Rainer et al., 2001), features related to highlight congruencies may be processed in the early visual areas, which explain the higher responses to specular surface observed in this study. The control stimuli used in the current study were scrambled images whose spatial phases were randomized or matte images whose spatial complexities are lower. Therefore, it remains open whether the cortical regions respond to the phase congruency of complex spatial patterns per se or they respond to the phase information particularly related to gloss. To understand which of these alternative possibilities is more likely, it would be interesting to examine cortical responses using stimuli such as matte, textured ("sticky" reflection) objects (Doerschner et al., 2011; Hartung and Kersten, 2002) or matte objects with misaligned highlights (Kim et al., 2011) in the future.

Motoyoshi et al. (2007) have recently found that the skewness of luminance histogram is positively correlated with the gloss perception in humans. Indeed the skewness of specular stimuli used in this study was larger (1.9) than matte (0.51) and scrambled stimuli (SS: 0.78, SM: 0.48). Although there are controversies about the correlation between the skewness and glossiness (counter examples are shown in Anderson and Kim, 2009), this higher order image statistics may also account for the responses in early visual areas (Olman et al., 2008).

Within the IT cortex, we found activation in small regions of the PITd in all four hemispheres in the two monkeys tested in experiment 1 which supports the idea that information regarding surface specularity is represented in restricted regions within the IT cortex. In experiment 2, we found that activity in the PITd region is modulated by specularity, not by image contrast. This region has been shown to respond preferentially to low spatial frequency (1 cycle/° or lower) (Rajimehr et al., 2011), suggesting that the difference in the spatial frequency contents between our stimuli could not account for the results in this study (see above). Previous neuroimaging studies in monkeys suggest the IT cortex contains a number of regions or modules specialized for individual categories such as faces (Tsao et al., 2003), bodies (Pinsk et al., 2005), objects (Denys et al., 2004; Tsao et al., 2003) and for colors (Conway et al., 2007; Harada et al., 2009). Our present results are consistent with the observations that representation within the IT cortex is not uniformly distributed; instead, particular categories or aspects of visual stimuli are processed in distinct regions, and surface properties are one of the attributes represented by specific neural substrates in the higher order visual area. Comparison with the results of the face/object localizer experiment revealed that activation in the PITd is located outside the face patches but partly overlapped with the posterior part of object-responsive regions. However, the activation peaks for objects were not selective for specularity (Fig. 9B), which suggests that separate mechanisms work on the processing of objects and specularity.

We also found activation in the CIT in the bilateral hemispheres of one monkey (M1) and, by lowering the threshold, activation was also observed at a similar position within the CIT in one hemisphere of another monkey (M2). Similarly, we found activation in the PITv in the bilateral hemispheres of M1 and, by lowering the threshold, found activation in the bilateral PITv of M2. In the Supplementary analysis 1, we show that the main effect of specularity measured in experiment 2 was significant for the CIT of M1 (Supplementary Fig. 1), suggesting that region of the CIT may also be involved in processing specularity. Nishio et al. (2012) recently found that there are neurons in the central part of the IT cortex, close to the activation site in the CIT identified in the present study, that selectively respond to stimuli with particular parameters of gloss. The neural substrates of activation we found in the CIT may correspond to those gloss-selective neurons. Nishio et al. (2012) also reported that some neurons preferred a glossy surface while others preferred a matte surface, and that as a population the neurons preferred glossy to matte surfaces. This bias may explain the specularity-induced activation observed in the CIT in the present study. However, the preference bias for different surfaces may vary across monkeys or hemispheres, which may explain why the response to specular surface was not detected in one hemisphere of M2, even after lowering the threshold.

We did not find any other regions responding to specular surfaces in the IT cortex, although activities in the ventral parts of the IT cortex and some parts of the lateral convexity cannot be measured using fMRI due to signal drop-out caused by susceptibility-induced field inhomogeneities (Ojemann et al., 1997). It is possible that regions responsive to surface properties would be found in the more ventral part of the IT cortex, given that human fMRI studies indicate that regions responsive to materials are found in the medial regions of the ventral pathway (Cant and Goodale, 2007, 2011; Cant et al., 2009; Cavina-Pratesi et al., 2009, 2010; Peuskens et al., 2004) that presumably correspond to more ventral parts of the IT cortex in macaque.

Surface shading is similar to specular highlights in the sense that both features are generated through the interaction of illumination, surface reflectance and surface orientation. Several earlier studies examined the cortical responses to shading in humans (Georgieva et al., 2008; Gerardin et al., 2010; Humphrey et al., 1997; Kourtzi et al., 2003; Moore and Engel, 2001; Taira et al., 2001) and macaques (Arcizet et al., 2009; Hanazawa and Komatsu, 2001; Nelissen et al., 2009), mainly in relation to 3D shape perception from shading. Some of the human fMRI studies reported activation in the higher visual areas, such as the LOC (Kourtzi et al., 2003; Moore and Engel, 2001) and right intraparietal area (Taira et al., 2001), while others also reported response modulation in the early visual areas (Humphrey et al., 1997; Taira et al., 2001), which is consistent with our finding that the early areas responded to surface cues. When various control stimuli were used to isolate activity related to shading, activation was observed only in the caudal inferior temporal gyrus (Georgieva et al., 2008). In macaques, neurons in V4 have been shown to respond to shading and to be selective for the direction of illumination (Arcizet et al., 2009; Hanazawa and Komatsu, 2001). Recently, Nelissen et al. (2009) used fMRI to detect shading-induced activation in V4 and the STS lip in the right TEO. The region appears to be located in the proximity of the region sensitive to specular surface in the PITd identified in the present study. However, because we contrasted the responses to specular stimuli with responses to matte stimuli that had shading, the activations we found cannot be explained simply by the responses to shading. Instead, we suggest that both of the surface cues, specularity and shading, are processed in similar regions of the IT cortex.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2012.07.052.

Acknowledgments

We are grateful to K. Matsuda for providing the eye tracking software. We also thank M. Takagi for technical assistance and T. Ohta for his assistance with animal training. This work was supported by a Grant-in-Aid for Scientific Research on Innovative Areas (no. 22135007) from the Ministry of Education, Culture, Sports, Science and Technology, Japan to H.K., and a Grant-in-Aid for JSPS Fellows from the Japan Society for the Promotion of Science to G.O, and Grants-in-Aid for Scientific Research (no. 22500248) from the Japan Society for the Promotion of Science to N.G.

References

Adelson, E.H., 2001. On seeing stuff: the perception of materials by humans and machines. In: Rogowitz, B.E., Pappas, T.N. (Eds.), Proceedings of SPIE, pp. 1–12.
Adelson, E.H., Anderson, C.H., Bergen, J.R., Burt, P.J., Odgen, J.M., 1984. Pyramid methods in image processing. RCA Eng. 29, 33–41.
Anderson, B.L., Kim, J., 2009. Image statistics do not explain the perception of gloss and lightness. J. Vis. 9 (11), 10.
Andrews, T.J., Clarke, A., Pell, P., Hartley, T., 2010. Selectivity for low-level features of objects in the human ventral stream. NeuroImage 49, 703–711.
Arcizet, F., Jouffrais, C., Girard, P., 2008. Natural textures classification in area V4 of the macaque monkey. Exp. Brain Res. 189, 109–120.
Arcizet, F., Jouffrais, C., Girard, P., 2009. Coding of shape from shading in area V4 of the macaque monkey. BMC Neurosci. 10, 140.
Beck, J., Prazdny, S., 1981. Highlights and the perception of glossiness. Percept. Psychophys. 30, 407–410.

Bell, A.H., Hadj-Bouziane, F., Frihauf, J.B., Tootell, R.B., Ungerleider, L.G., 2009. Object representations in the temporal cortex of monkeys and humans as revealed by functional magnetic resonance imaging. J. Neurophysiol. 101, 688–700.
Bisley, J.W., Goldberg, M.E., 2010. Attention, intention, and priority in the parietal lobe. Annu. Rev. Neurosci. 33, 1–21.
Cant, J.S., Goodale, M.A., 2007. Attention to form or surface properties modulates different regions of human occipitotemporal cortex. Cereb. Cortex 17, 713–731.
Cant, J.S., Goodale, M.A., 2011. Scratching beneath the surface: new insights into the functional properties of the lateral occipital area and parahippocampal place area. J. Neurosci. 31, 8248–8258.
Cant, J.S., Arnott, S.R., Goodale, M.A., 2009. fMR-adaptation reveals separate processing regions for the perception of form and texture in the human ventral stream. Exp. Brain Res. 192, 391–405.
Cavina-Pratesi, C., Kentridge, R.W., Heywood, C.A., Milner, A.D., 2009. Separate processing of texture and form in the ventral stream: evidence from fMRI and visual agnosia. Cereb. Cortex 20, 433–446.
Cavina-Pratesi, C., Kentridge, R.W., Heywood, C.A., Milner, A.D., 2010. Separate channels for processing form, texture, and color: evidence from fMRI adaptation and visual object agnosia. Cereb. Cortex 20, 2319–2332.
Conway, B.R., Moeller, S., Tsao, D.Y., 2007. Specialized color modules in macaque extrastriate cortex. Neuron 56, 560–573.
Debevec, P.E., 1998. Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques. ACM, pp. 189–198.
Denys, K., Vanduffel, W., Fize, D., Nelissen, K., Peuskens, H., Van Essen, D., Orban, G.A., 2004. The processing of visual shape in the cerebral cortex of human and nonhuman primates: a functional magnetic resonance imaging study. J. Neurosci. 24, 2551–2565.
Doerschner, K., Boyaci, H., Maloney, L.T., 2010. Estimating the glossiness transfer function induced by illumination change and testing its transitivity. J. Vis. 10 (4), 8.
Doerschner, K., Fleming, R.W., Yilmaz, O., Schrater, P.R., Hartung, B., Kersten, D., 2011. Visual motion and the perception of surface material. Curr. Biol. 21, 2010–2016.
Felleman, D.J., Van Essen, D.C., 1991. Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex 1, 1–47.
Fize, D., Vanduffel, W., Nelissen, K., Denys, K., Chef d'Hotel, C., Faugeras, O., Orban, G.A., 2003. The retinotopic organization of primate dorsal V4 and surrounding areas: a functional magnetic resonance imaging study in awake monkeys. J. Neurosci. 23, 7395–7406.
Fleming, R.W., Dror, R.O., Adelson, E.H., 2003. Real-world illumination and the perception of surface reflectance properties. J. Vis. 3 (5), 347–368.
Friston, K.J., Holmes, A.P., Poline, J.B., Grasby, P.J., Williams, S.C., Frackowiak, R.S., Turner, R., 1995. Analysis of fMRI time-series revisited. NeuroImage 2, 45–53.
Friston, K.J., Penny, W.D., Glaser, D.E., 2005. Conjunction revisited. NeuroImage 25, 661–667.
Georgieva, S.S., Todd, J.T., Peeters, R., Orban, G.A., 2008. The extraction of 3D shape from texture and shading in the human brain. Cereb. Cortex 18, 2416–2438.
Gerardin, P., Kourtzi, Z., Mamassian, P., 2010. Prior knowledge of illumination for 3D perception in the human brain. Proc. Natl. Acad. Sci. U. S. A. 107, 16309–16314.
Hanazawa, A., Komatsu, H., 2001. Influence of the direction of elemental luminance gradients on the responses of V4 cells to textured surfaces. J. Neurosci. 21, 4490–4497.
Harada, T., Goda, N., Ogawa, T., Ito, M., Toyoda, H., Sadato, N., Komatsu, H., 2009. Distribution of colour-selective activity in the monkey inferior temporal cortex revealed by functional magnetic resonance imaging. Eur. J. Neurosci. 30, 1960–1970.
Hartung, B., Kersten, D., 2002. Distinguishing shiny from matte [Abstract]. J. Vis. 2 (7), 551.
Heeger, D.J., Bergen, J.R., 1995. Pyramid-based texture analysis/synthesis. Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques. ACM, pp. 229–238.
Henriksson, L., Nurminen, L., Hyvarinen, A., Vanni, S., 2008. Spatial frequency tuning in human retinotopic visual areas. J. Vis. 8 (10), 5.
Henriksson, L., Hyvarinen, A., Vanni, S., 2009. Representation of cross-frequency spatial phase relationships in human visual cortex. J. Neurosci. 29, 14342–14351.
Hiramatsu, C., Goda, N., Komatsu, H., 2011. Transformation from image-based to perceptual representation of materials along the human ventral visual pathway. NeuroImage 57, 482–494.
Humphrey, G.K., Goodale, M.A., Bowen, C.V., Gati, J.S., Vilis, T., Rutt, B.K., Menon, R.S., 1997. Differences in perceived shape from shading correlate with activity in early visual areas. Curr. Biol. 7, 144–147.
Hunter, R.S., Harold, R.W., 1987. The Measurement of Appearance. Wiley, New York.
Kim, J., Anderson, B., 2010. Image statistics and the perception of surface gloss and lightness. J. Vis. 10 (9), 3.
Kim, J., Marlow, P., Anderson, B.L., 2011. The perception of gloss depends on highlight congruence with surface shading. J. Vis. 11 (9), 4.
Komatsu, H., Ideura, Y., 1993. Relationships between color, shape, and pattern selectivities of neurons in the inferior temporal cortex of the monkey. J. Neurophysiol. 70, 677–694.
Komatsu, H., Ideura, Y., Kaji, S., Yamane, S., 1992. Color selectivity of neurons in the inferior temporal cortex of the awake macaque monkey. J. Neurosci. 12, 408–424.
Köteles, K., De Maziere, P.A., Van Hulle, M., Orban, G.A., Vogels, R., 2008. Coding of images of materials by macaque inferior temporal cortical neurons. Eur. J. Neurosci. 27, 466–482.
Kourtzi, Z., Erb, M., Grodd, W., Bülthoff, H.H., 2003. Representation of the perceived 3-D object shape in the human lateral occipital complex. Cereb. Cortex 13, 911–920.
Leite, F.P., Tsao, D., Vanduffel, W., Fize, D., Sasaki, Y., Wald, L.L., Dale, A.M., Kwong, K.K., Orban, G.A., Rosen, B.R., Tootell, R.B., Mandeville, J.B., 2002. Repeated fMRI using iron oxide contrast agent in awake, behaving macaques at 3 Tesla. NeuroImage 16, 283–294.
Liu, Y., Vogels, R., Orban, G.A., 2004. Convergence of depth from texture and depth from disparity in macaque inferior temporal cortex. J. Neurosci. 24, 3795–3800.
Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc. Natl. Acad. Sci. U. S. A. 92, 8135–8139.
Mante, V., Frazor, R.A., Bonin, V., Geisler, W.S., Carandini, M., 2005. Independence of luminance and contrast in natural scenes and in the early visual system. Nat. Neurosci. 8, 1690–1697.
Marlow, P., Kim, J., Anderson, B.L., 2011. The role of brightness and orientation congruence in the perception of surface gloss. J. Vis. 11 (9), 16.
Matsuda, K., nagami, T., Kawano, K., Yamane, S., 2000. A new system for measuring eye position on a personal computer. Soc. Neurosci. Abstr. 7442.
Mazaika, P., Gabrieli, S.W., Reiss, A., 2007. Artifact repair for fMRI data from high motion clinical subjects. 13th Annual Meeting of the Organization for Human Brain Mapping, Chicago, IL.
Moore, C., Engel, S.A., 2001. Neural response to perception of volume in the lateral occipital complex. Neuron 29, 277–286.
Motoyoshi, I., Nishida, S., Sharan, L., Adelson, E.H., 2007. Image statistics and the perception of surface qualities. Nature 447, 206–209.
Nelissen, K., Vanduffel, W., Orban, G.A., 2006. Charting the lower superior temporal region, a new motion-sensitive region in monkey superior temporal sulcus. J. Neurosci. 26, 5929–5947.
Nelissen, K., Joly, O., Durand, J.B., Todd, J.T., Vanduffel, W., Orban, G.A., 2009. The extraction of depth structure from shading and texture in the macaque brain. PLoS One 4 (12), e8306.
Nichols, T., Brett, M., Andersson, J., Wager, T., Poline, J.B., 2005. Valid conjunction inference with the minimum statistic. NeuroImage 25, 653–660.
Nishio, A., Goda, N., Komatsu, H., 2012. Neural selectivity and representation of gloss in the monkey inferior temporal cortex. J. Neurosci. 32, 10780–10793.
Ojemann, J.G., Akbudak, E., Snyder, A.Z., McKinstry, R.C., Raichle, M.E., Conturo, T.E., 1997. Anatomic localization and quantitative analysis of gradient refocused echo-planar fMRI susceptibility artifacts. NeuroImage 6, 156–167.
Olkkonen, M., Brainard, D.H., 2010. Perceived glossiness and lightness under real-world illumination. J. Vis. 10 (9), 5.
Olman, C., Boyaci, H., Fang, F., Doerschner, K., 2008. V1 responses to different types of luminance histogram contrast [Abstract]. J. Vis. 8 (6), 345.
Perna, A., Tosetti, M., Montanaro, D., Morrone, M.C., 2008. BOLD response to spatial phase congruency in human brain. J. Vis. 8 (15), 11–15.
Peuskens, H., Claeys, K.G., Todd, J.T., Norman, J.F., Van Hecke, P., Orban, G.A., 2004. Attention to 3-D shape, 3-D motion, and texture in 3-D structure from motion displays. J. Cogn. Neurosci. 16, 665–682.
Pinsk, M.A., DeSimone, K., Moore, T., Gross, C.G., Kastner, S., 2005. Representations of faces and body parts in macaque temporal cortex: a functional MRI study. Proc. Natl. Acad. Sci. U. S. A. 102, 6996–7001.
Portilla, J., Simoncelli, E.P., 2000. A parametric texture model based on joint statistics of complex wavelet coefficients. Int. J. Comput. Vision 40, 49–71.
Rainer, G., Augath, M., Trinath, T., Logothetis, N.K., 2001. Nonmonotonic noise tuning of BOLD fMRI signal to natural images in the visual cortex of the anesthetized monkey. Curr. Biol. 11, 846–854.
Rajimehr, R., Devaney, K.J., Bilenko, N.Y., Young, J.C., Tootell, R.B., 2011. The "parahippocampal place area" responds preferentially to high spatial frequencies in humans and monkeys. PLoS Biol. 9, e1000608.
Richmond, B.J., Optican, L.M., Podell, M., Spitzer, H., 1987. Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. I. Response characteristics. J. Neurophysiol. 57, 132–146.
Rust, N.C., Dicarlo, J.J., 2010. Selectivity and tolerance ("invariance") both increase as visual information propagates from cortical area V4 to IT. J. Neurosci. 30, 12978–12995.
Sakano, Y., Ando, H., 2010. Effects of head motion and stereo viewing on perceived glossiness. J. Vis. 10 (9), 15.
Saleem, K.S., Logothetis, N.K., 2007. A Combined MRI and Histology Atlas of the Rhesus Monkey Brain in Stereotaxic Coordinates. Elsevier, London.
Sasaki, Y., Hadjikhani, N., Fischl, B., Liu, A.K., Marrett, S., Dale, A.M., Tootell, R.B., 2001. Local and global attention are mapped retinotopically in human occipital cortex. Proc. Natl. Acad. Sci. U. S. A. 98, 2077–2082.
Simoncelli, E.P., Freeman, W.T., Adelson, E.H., Heeger, D.J., 1992. Shiftable multiscale transforms. IEEE Trans. Inf. Theory 38, 587–607 (Special Issue on Wavelets).
Singh, K.D., Smith, A.T., Greenlee, M.W., 2000. Spatiotemporal frequency and direction sensitivities of human visual areas measured using fMRI. NeuroImage 12, 550–564.
Taira, M., Nose, I., Inoue, K., Tsutsui, K., 2001. Cortical areas related to attention to 3D surface structures based on shading: an fMRI study. NeuroImage 14, 959–966.
Tsao, D.Y., Freiwald, W.A., Knutsen, T.A., Mandeville, J.B., Tootell, R.B., 2003. Faces and objects in macaque cerebral cortex. Nat. Neurosci. 6, 989–995.
Tsao, D.Y., Freiwald, W.A., Tootell, R.B., Livingstone, M.S., 2006. A cortical region consisting entirely of face-selective cells. Science 311, 670–674.
Van Essen, D.C., Drury, H.A., Dickson, J., Harwell, J., Hanlon, D., Anderson, C.H., 2001. An integrated software suite for surface-based analyses of cerebral cortex. J. Am. Med. Inform. Assoc. 8, 443–459.
Vanduffel, W., Fize, D., Mandeville, J.B., Nelissen, K., Van Hecke, P., Rosen, B.R., Tootell, R.B., Orban, G.A., 2001. Visual motion processing investigated using contrast agent-enhanced fMRI in awake behaving monkeys. Neuron 32, 565–577.
Wang, Y., Fujita, I., Murayama, Y., 2003. Coding of visual patterns and textures in monkey inferior temporal cortex. Neuroreport 14, 453–457.
Ward, G.J., 1992. Measuring and modeling anisotropic reflection. Comput. Graphics 26, 265.
Wendt, G., Faul, F., Ekroll, V., Mausfeld, R., 2010. Disparity, motion, and color information improve gloss constancy performance. J. Vis. 10 (9), 7.
Yang, J., Otsuka, Y., Kanazawa, S., Yamaguchi, M.K., Motoyoshi, I., 2011. Perception of surface glossiness by infants aged 5 to 8 months. Perception 40, 1491–1502.