



# Human visual cortical responses to specular and matte motion flows

Tae-Eui Kam<sup>1</sup>, Damien J. Mannion<sup>2,3,4</sup>, Seong-Whan Lee<sup>1,2\*</sup>, Katja Doerschner<sup>5,6,7</sup> and Daniel J. Kersten<sup>2,4</sup>

<sup>1</sup> Department of Computer Science and Engineering, Korea University, Seoul, South Korea, <sup>2</sup> Department of Brain and Cognitive Engineering, Korea University, Seoul, South Korea, <sup>3</sup> School of Psychology, UNSW Australia, Sydney, NSW, Australia, <sup>4</sup> Department of Psychology, University of Minnesota, Minneapolis, MN, USA, <sup>5</sup> Department of Psychology, Bilkent University, Ankara, Turkey, <sup>6</sup> National Magnetic Resonance Research Center, Bilkent University, Ankara, Turkey, <sup>7</sup> Department of Psychology, Justus-Liebig-University Giessen, Giessen, Germany

Determining the compositional properties of surfaces in the environment is an important visual capacity. One such property is specular reflectance, which encompasses the range from matte to shiny surfaces. Visual estimation of specular reflectance can be informed by characteristic motion profiles; a surface with a specular reflectance that is difficult to determine while static can be confidently disambiguated when set in motion. Here, we used fMRI to trace the sensitivity of human visual cortex to such motion cues, both with and without photometric cues to specular reflectance. Participants viewed rotating blob-like objects that were rendered as images (photometric) or dots (kinematic) with either matte-consistent or shiny-consistent specular reflectance profiles. We were unable to identify any areas in low and mid-level human visual cortex that responded preferentially to surface specular reflectance from motion. However, univariate and multivariate analyses identified several visual areas; V1, V2, V3, V3A/B, and hMT+, capable of differentiating shiny from matte surface flows. These results indicate that the machinery for extracting kinematic cues is present in human visual cortex, but the areas involved in integrating such information with the photometric cues necessary for surface specular reflectance remain unclear.

## OPEN ACCESS

### Edited by:

Michael A. Silver,  
University of California, Berkeley, USA

### Reviewed by:

Axel Kohler,  
University of Osnabrück, Germany  
Gouki Okazawa,  
New York University, USA

### \*Correspondence:

Seong-Whan Lee  
sw.lee@korea.ac.kr

**Received:** 12 May 2015

**Accepted:** 04 October 2015

**Published:** 21 October 2015

### Citation:

Kam T-E, Mannion DJ, Lee S-W,  
Doerschner K and Kersten DJ (2015)  
Human visual cortical responses to  
specular and matte motion flows.  
*Front. Hum. Neurosci.* 9:579.  
doi: 10.3389/fnhum.2015.00579

**Keywords:** visual perception, surface materials, motion flow, functional magnetic resonance imaging (fMRI), classification

## 1. INTRODUCTION

Experiencing visual qualities, such as the glossiness of polished marble or the smoothness of silk, are an integral part of human conscious experience. The automaticity with which this perceptual process occurs belies the computational difficulty that the brain is faced with in its task to extract meaningful information from the ambiguous retinal signal. The ambiguity lies in whether the pattern of light arriving at the retinae originates from variations in illumination, shape, mesoscale geometrical structure, or the material of the object. Despite this computational challenge humans can effortlessly visually sense dynamic physical properties such as viscosity, elasticity, or stiffness and optical properties such as transparency, glossiness, shininess, or roughness and easily discriminate between material classes. Yet, surprisingly little is understood about how the brain recognizes materials.

While there is a growing body of research on how the visual system extracts optical material qualities such as surface glossiness, roughness or translucency (e.g., Nishida and Shinya, 1998;

Adelson, 2001; Dror et al., 2001; Fleming et al., 2004, 2013; te Pas and Pont, 2005; Ho et al., 2006; Motoyoshi et al., 2007; Anderson and Kim, 2009; Doerschner et al., 2010, 2011b; Kim and Anderson, 2010; Olkkonen and Brainard, 2010; Wijntjes and Pont, 2010; Kim et al., 2011, 2012; Marlow et al., 2011, 2012; Zaidi, 2011; Fleming, 2012; Gkioulekas et al., 2013), the majority of this research has focused on visual information available in static images (but see Sakano and Ando, 2008; Wendt et al., 2010). However, image motion can also convey optical material qualities. In a seminal demonstration, Hartung and Kersten (2002) showed that image motion influences how we perceive the material of a rotating object: when surface features are rigidly attached to the object it appears matte, when surface features slide across the shape, consistent with specular flow, the object looks shiny. Critically, when the object was not moving it was ambiguous whether it was shiny or matte (<http://gandalf.psych.umn.edu/users/kersten/kersten-lab/demos/MatteOrShiny.html>). In subsequent work (Doerschner et al., 2011a), using a combination of image and optic flow analysis, pattern classification and psychophysics, identified three motion cues that the brain could rely on to distinguish between matte and shiny surfaces. Their results revealed a previously unknown use for optic flow in the perception of optical surface material properties. How the brain processes material-dependent image motion, however, is currently unstudied.

In fact, only a few studies have investigated the neural basis of material perception, and these have focused on static image cues to surface material appearance. Electrophysiological studies found neurons in the superior temporal sulcus and anterior inferior temporal cortex to be responsive to surface material properties (e.g., Nishio et al., 2012), whereas neuroimaging studies have identified several loci in low and high level visual areas including V1, V2, V3, V4, posterior inferior temporal cortex, and ventral higher-order visual areas (Peuskens et al., 2004; Cant and Goodale, 2007; Köteles et al., 2008; Cant and Goodale, 2009, 2011; Cavina-Pratesi et al., 2010a,b; Hiramatsu et al., 2011; Okazawa et al., 2012; Wada et al., 2014) with responses in higher level visual areas correlating with perceived similarities of surface material categories, and responses in early visual areas correlating with concurrent changes in simple image features such as spatial frequency or color (Hiramatsu et al., 2011). The first hint that there may be specialized neural mechanisms sensitive to material-specific motion cues came from an experiment by Kam et al. (2012) which found that visual adaptation to a specular rotating object biases subsequently presented objects toward matte appearance.

The study of perceptual or neural responses to optic flow produced by object motion is complicated by the fact that both kinematic and photometric factors contribute to flow. The kinematic deals with the geometric relation between objects and their image projections and is critical for inferring structure. The photometric deals with the relation between the material reflective properties of objects and their images given possibly varying illumination conditions, and is crucial for perceiving material appearance. Reliable estimates of the shape of a rigidly rotating object require the identification of geometric features that are uniquely tied to surface points. Thus, studies of

perceived “structure-from-motion” have traditionally relied on kinematic displays composed of moving dots which have a unique relationship to corresponding surface points. However, on inspection such displays convey no information about surface properties such as reflectance or shininess. On the other hand photometric flows are characterized by spatio-temporal changes in intensity which provide information about material properties as well as geometrical structure. While photometric and kinematic factors are not independent (Zang et al., 2010) one can assess the effect of added photometric information.

Here, our aim was to investigate the brain processing underlying the perception of specular reflectance from motion in human observers. We used functional magnetic resonance imaging (fMRI) to infer the magnitude of brain activation across posterior visual cortex to rotating blob-like objects. Following Doerschner et al. (2011a), we rendered the objects in image sequences that yield a perceptual impression of a matte or shiny specular reflectance when set in motion. We also identified the flow structure of such sequences and used dots to create presentation conditions that mimic the matte and specular flows but lack the perception of surface specular reflectance. Our key prediction was that areas of human visual cortex that are sensitive to specular reflectance from motion would show an interaction between the rendering type (image or dot) and the flow type (matte or shiny).

## 2. MATERIALS AND METHODS

### 2.1. Participants

Nine participants, each with normal vision, participated in the current study. Each participant gave their informed written consent and the study conformed to safety guidelines for MRI research and was approved by the Institutional Review Board at Korea University. One participant was excluded from analysis due to difficulties in defining his/her retinotopic visual areas, and the analysis presented here is derived from the remaining eight participants.

### 2.2. Apparatus

Functional imaging was conducted using a Siemens 3T-Trio magnet (Erlangen, Germany) with a 32-channel head coil. To allow participants to have unrestricted viewing of the display through each eye, the coil was operated with the lower 20 elements only. Images were collected with a  $T_2^*$  sensitive gradient echo imaging pulse sequence (TR = 3 s, TE = 30 ms, delay in TR = 0.8 s, flip angle = 90°, matrix = 96 × 96, GRAPPA acceleration factor = 2, FOV = 192 × 192 mm, partial Fourier =  $\frac{7}{8}$ , voxel size = 2 mm isotropic) in 35 interleaved oblique coronal slices covering the occipital lobes. Stimuli were displayed on an LCD monitor (“BOLDscreen,” Cambridge Research Systems, Kent, UK) with a spatial resolution of 1920 × 1200 pixels, temporal resolution of 60 Hz, and mean luminance of 450 cd/m<sup>2</sup>. The monitor output was linearized via correction of luminance values measured with a ColorCAL MKII colorimeter (Cambridge Research Systems, Kent, UK). The screen was viewed through a mirror mounted on the head coil at a distance of 112 cm, giving a viewing angle of 26.0° ×

16.4°. Stimuli were displayed using PsychToolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) on a Macbook Pro driving an Intel HD Graphics 4000 video card. As detailed below, analyses were performed using FreeSurfer 5.1.0 (Dale et al., 1999; Fischl et al., 1999), FSL 4.1.6 (Smith et al., 2004), and AFNI/SUMA (2013/05/22; Cox, 1996; Saad et al., 2004).

## 2.3. Stimuli

Four stimulus conditions, each based on a single blob-like object, were developed: a *rendering* class with two conditions corresponding to photometric and kinematic displays, which we call “image” and “dot” flows, respectively. These conditions were crossed with a *material flow* class with two conditions referred to as “matte” and “shiny.”

“Image” renderings were created with Radiance 3D (Larson et al., 1998) using the environment-mapping techniques described in Debevec (2002). Objects were either rendered as specularly-reflecting or as textured and diffusely reflecting. For the latter, the specular reflection from one particular view point was “stuck on” to the object’s surface creating a matte, textured appearance when the object was rotating. See Doerschner et al., 2011a for details of the object generation and rendering procedure, and see **Figure 1** for stimulus examples.

Five environment maps were generated by combining several maps of the Debevec (2002) database using an array of techniques in Adobe Photoshop such as blending, duplicating, and reflecting over the horizontal axis. This was done in order to maximize spatial complexity in the resulting environment map, and to obtain optimal optic flow estimates. The object was assigned a material property of either diffuse or specular reflectance. Such material properties are difficult to distinguish when viewed statically, as evident in **Figure 1** and quantified in Supplementary Figure 1, and this ambiguity was further enhanced by histogram equalization and contrast reduction. The luminance histogram was equalized using the method proposed by Gonzalez and Woods (2008). This approach essentially reduces the skewness in the luminance histogram, and thus removes a (static) image cue that has been shown to correlate with a glossy appearance (Motoyoshi et al., 2007). As a next step, we decreased the intensity of pixels above a value of 128 using the following procedure: we computed the difference that a given pixel intensity has with respect to 128, and multiplied this difference by 0.4. The new pixel value was then obtained by adding the attenuated difference to 128, and rounding the result. As an example a pixel value of 255 would be reduced to 179 ( $128 + (255 - 128) \times 0.4$ ). This manipulation reduced image contrast, by attenuating the intensity of particular bright regions on the object (such as specular highlights), thus further reducing the effects of static cues to perceived glossiness (Marlow et al., 2011). By these two processing steps the mean luminance and contrast of all stimuli (“shiny” and “matte”) was effectively equated to a value of 141 and 0.2, respectively.

We define photometric cues on Section 1 as those, signaling the relation between material reflective properties of objects and their images given varying illumination conditions. Luminance histogram skewness and image contrast are only two potential

correlates of shiny appearance (see Chadwick and Kentridge, 2015), thus while we might have reduced photometric cues to glossiness we did not eliminate all. In particular the idiosyncratic spatial structure (e.g., compression at high curvature points) in the images of specular objects remained intact (e.g., Fleming et al., 2004).

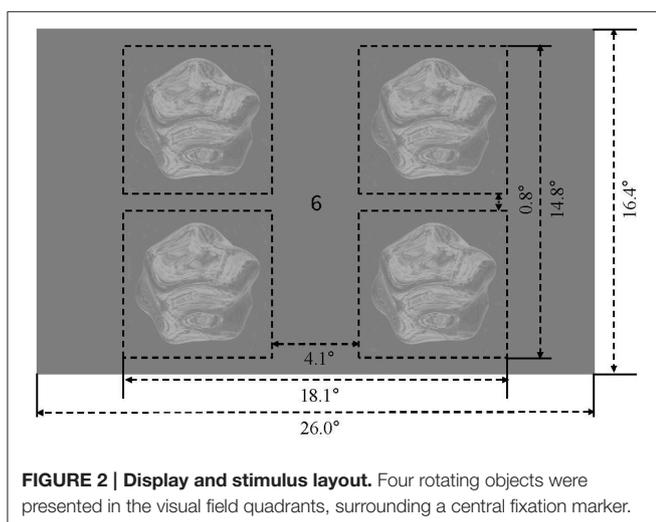
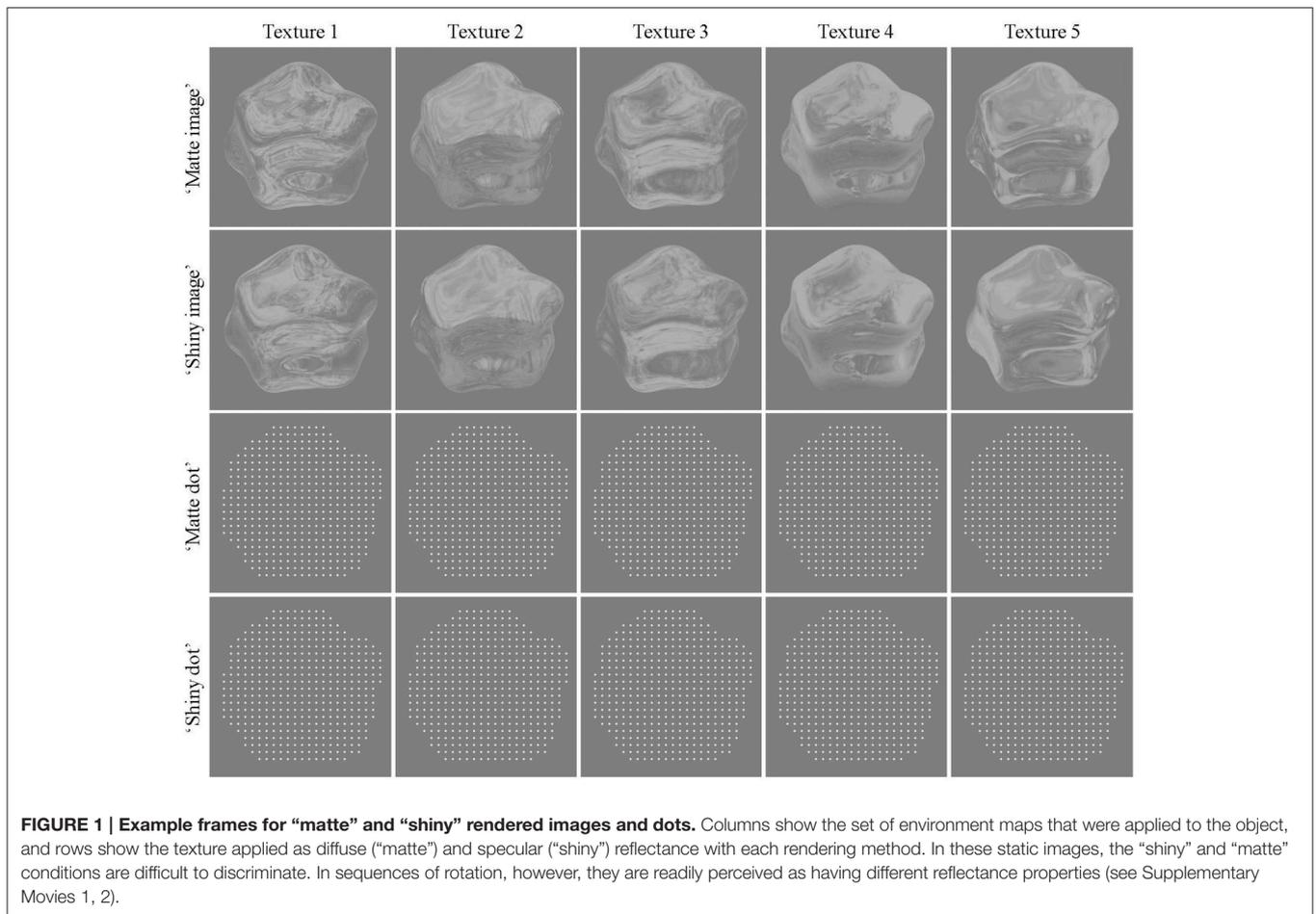
More importantly, we wanted to create a baseline stimulus where the surface material is ambiguous in the static case [as in the Hartung and Kersten (2002) demo, i.e., both interpretations, “shiny” and “matte,” are equally likely; also see Supplementary Figure 1]. Had our objects already looked shiny in the static case, we would have introduced an asymmetry for the moving stimuli: the motion of “sticky” reflections patterns (matte appearance) would have been surprising and in conflict with the photometric cues (say high positive skewness), whereas the specular motion would not have this conflicting information. In order to avoid this confounding asymmetry we performed the above described image manipulations.

The kinematic dot renders were created by applying a phase-based optical flow method (Gautama and Van Hulle, 2002) to the image renders. Dots (white diamonds, 0.1° visual angle in diameter) were initially placed at uniform locations in the two-dimensional area occupied by the object (9 dots/° visual angle<sup>2</sup> density), and each dot was then moved on each frame by an estimate of its underlying flow field (see Supplementary Movies 3, 4). Due to the lack of photometric information, the shiny/dot and matte/dot conditions are not perceived as shiny or matte. Both image and dot renders do, however, retain the differences between shiny and matte stimuli in coverage, divergence, and shape reliability cues that were reported by Doerschner et al. (2011a); see Supplementary Figures 2–4 and Supplementary Tables 1–3. The details of the methods for calculating the features are described in Doerschner et al. (2011a).

## 2.4. Experimental Procedures

The four stimulus conditions (matte/shiny image flows, matte/shiny dot flows) were presented in a block design. Within each stimulus block, four objects, each rendered with the same environment map (assigned from a set of five maps in pseudorandom order across blocks) and being approximately 7.0° in diameter, were centered at 6.9° eccentricity within the visual field quadrants (see **Figure 2**). Each object rotated back and forth eight times over the course of a 15 s block, with each rotation traveling 15°. The object in a given visual field quadrant rotated about the same axis over the course of the experiment, and the three cardinal axes and one oblique axis were assigned to the four visual field quadrants (see Supplementary Movie 5). Blocks were ordered in sequences in which the four stimulus condition blocks were followed by a blank block, with the arrangement of stimulus blocks chosen such that each condition was preceded an equal number of times by each of the other conditions. There were four such sequences per run, and an additional blank block was appended to the run order, giving a run duration of 315 s (105 volumes). Each participant completed 12 runs, collected within a single session.

A demanding foveal task was used to control fixation, and divert attention from the appearance of the objects. Throughout



each run, a sequence of digits, which were randomly chosen from zero to nine and of random polarity (black or white), was presented at center of the screen. Participants were instructed to press a button when one of two target digits appeared in the sequence. The foveal targets, which had opposite polarities

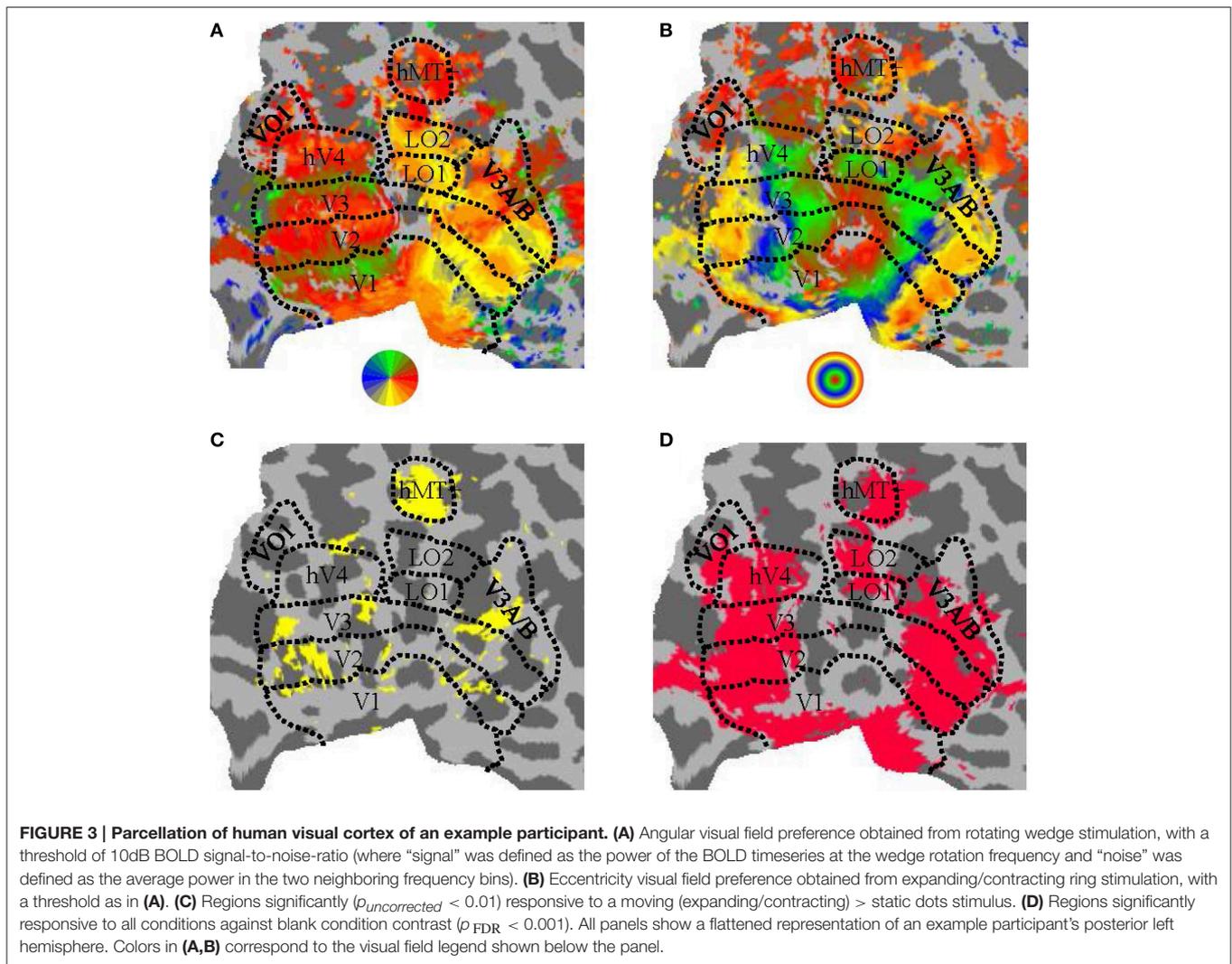
and differing digits, were introduced to the participants at the beginning of each run.

## 2.5. Anatomical Acquisition and Processing

A  $T_1$ -weighted anatomical image (sagittal MP-RAGE, 1mm isotropic resolution) was collected from each participant in a separate session. FreeSurfer (Dale et al., 1999; Fischl et al., 1999) was used for segmentation, cortical surface reconstruction, and surface inflation and flattening of each participant’s anatomical image.

## 2.6. Visual Area Definition

In a separate session, standard protocols (Serenio et al., 1995; DeYoe et al., 1996; Engel et al., 1997; Larsson and Heeger, 2006; Hansen et al., 2007; Schira et al., 2007; Bressler and Silver, 2010) were used for defining retinotopic visual areas of the brain. Participants observed a clockwise/anti-clockwise rotating wedge stimulus during four runs and expanding/contracting ring stimulus during two runs. We manually defined the visual areas V1, V2, and V3 (Dougherty et al., 2003), hV4 (Wade et al., 2002; Goddard et al., 2011), LO1/2 (Larsson and Heeger, 2006), and VO1 (Brewer et al., 2005) based on the visual field preferences established from phase-encoded analysis (see **Figure 3**).



We also functionally defined the human MT complex (hMT+; Tootell et al., 1995) and Lateral Occipital Complex (LOC; Malach et al., 1995) via two runs of moving (expanding/contracting) and static low-contrast dots and two runs of a Face-Place-Object (FPO) paradigm (Cant and Goodale, 2011), respectively. For the FPO paradigm, we collected face and object images from the Carnegie Mellon University database (Harvey and Burgund, 2012) and place images from the Caltech database (Andrews et al., 2010).

The relevant localizer contrast from a GLM analysis (motion > static for hMT+, objects > faces and places for LOC) was visualized on the flattened cortical surface of each participant and hemisphere. The threshold on the contrast statistic was manually adjusted to estimate the extent of hMT+ and LOC, with the location of hMT+ and LOC identified as the cluster that matched the expected anatomical location and the expected location with respect to the retinotopically-defined visual areas. The ROIs were then drawn manually as a closed and filled region on the cortical surface (see **Figure 3C** for an example ROI definition for hMT+).

## 2.7. Pre-processing

Functional images were motion corrected using AFNI, with reference to the volume acquired closest in time to a within-session fieldmap image, and resampled with heptic interpolation before being unwarped using FSL to correct geometric distortions introduced by magnetic field inhomogeneities. No slice-timing correction was applied. The participant’s anatomical image was then coregistered with a mean of all functional images via AFNI’s `align_epi_anat.py`, using a local Pearson correlation cost function (Saad et al., 2009) and six free parameters (three translation, three rotation). Coarse registration parameters were determined manually and passed to the registration routine to provide initial estimates and to constrain the range of reasonable transformation parameter values. The motion-corrected and unwarped functional data were then projected onto the cortical surface by averaging between the white matter and pial boundaries (identified with FreeSurfer) using AFNI/SUMA. No specific spatial smoothing was applied. All analysis was performed on the nodes of this surface domain representation in the participant’s native brain space.

## 2.8. Univariate Analysis

Stimulus blocks for each condition were modeled as boxcars and convolved with SPM's canonical haemodynamic response function. Legendre polynomials up to the third degree were included as additional regressors, for each run, to detrend for low-frequency noise. The GLM was estimated via AFNI's `3dREMLfit`, which accounts for noise temporal correlations via a voxelwise ARMA(1,1) model.

The stimulus condition beta weights obtained from the GLM were converted to Percent Signal Change (PSC) via division by the average of the synthesized baseline timecourse (derived from the Legendre polynomial regressors). For each visual area, such percent signal change values were then averaged across the cortical surface nodes within the area that showed above-baseline responses to visual stimulation (identified by an all stimulus > 0 contrast,  $p_{\text{FDR}} < 0.001$ ).

The mean PSC value across participants within each area were normalized by subtracting each participant's mean response across stimulus conditions and adding the grand mean across participants and conditions (Cousineau, 2005). A Two-way within-subjects ANOVA was then conducted for each visual area, with flow (matte, shiny) and rendering (image, dots) as fixed factors and participants as a random factor.

## 2.9. Multivariate Pattern Analysis (MVPA)

The timeseries for each participant and run were first high-pass filtered with Legendre polynomials up to the third degree. An amplitude was then estimated for each block as the mean signal within its five volumes (15 s), shifted by two volumes (6 s) to compensate for the delayed hemodynamic response. The amplitude estimates within each run were then normalized ( $z$ -scored). This procedure produced 192 responses per participant for each node on the cortical surface; four responses for each of four conditions in each of 12 runs.

The MVPA was performed separately for each participant and visual area, and was implemented using a 12-fold leave-one-run-out strategy in which the responses from a given run were designated (in turn) to form the "test" set and the remaining runs to form the "training" set. In each fold, separate linear support vector machines (SVMs) were trained for image and dot renderings on labeled examples of matte and shiny flow response patterns. Each training set thus consisted of 88 examples, with matte and shiny flow examples equally represented. Flow discrimination accuracy was then estimated by using the trained SVMs to predict the flow condition of test set examples from the same (within-class) or different (between-class) rendering. SVMs were implemented with `libsvm` 3.17 (Chang and Lin, 2011) via Matlab 8.1.0.604 (The Mathworks Inc., Natick, MA). The accuracy was based on the proportion of hits and false alarms after aggregation of the 12-folds, and expressed in  $d'$  units. The  $d'$  calculation included an addition of 0.5 to all hit and false alarm counts and an addition of 1 to the number of trials in each condition class, in order to accommodate extreme hit or false alarm rates (Stanislaw and Todorov, 1999).

Features were selected for inclusion in the response pattern for a given visual area based on the  $t$ -value of the all stimulus > blank screen contrast performed in the univariate analysis. The surface

nodes within a given visual area were ranked in descending order based on the magnitude of this localizing  $t$ -value, and the MVPA procedure was performed with patterns formed from including increasing numbers of such ranked nodes (from  $n = 10$  to  $N$ , where  $N$  is the number of nodes with statistically significant  $t$ -values at  $p < 0.05$ , one-tailed, uncorrected) in 10 node increments. The variation in MVPA performance with increasing nodes was summarized via a least-squares fit to the function:

$$p = a(1 - e^{-\frac{n}{c}}) \quad (1)$$

where  $p$  is the performance level ( $d'$ ),  $n$  is the number of included nodes, and  $a$  and  $c$  are fitted parameters that describe the asymptotic performance ( $a$ ) and curvature ( $c$ ). The classification accuracy for a given participant, visual area, training class (image, dot), and testing class (within-class, between-class) was then taken as either the fitted asymptote or, in the case of unsuccessful fit, the mean accuracy over nodes. The classification accuracy and fitted performance levels with increasing nodes are shown in Supplementary Figures 5–13.

## 2.10. Multiple Comparisons Correction

The largely exploratory nature of the research question investigated in this study motivates the testing of hypotheses in multiple candidate regions in the posterior vision-sensitive area of the brain. This approach necessitates consideration of the inflationary effect of performing statistical analyses in multiple regions-of-interest (visual areas) on the rate of false-positive inference. Here, we adopted a false discovery rate (FDR; Benjamini and Hochberg, 1995) strategy to help reduce this multiple comparisons problem. Specifically, we adjusted the probability outcomes of each statistical test conducted across multiple visual areas via the FDR procedure, and the resulting  $p_{\text{FDR}}$  values were then evaluated at a criterion of 0.05 for statistical significance (with the exception of the localizer contrast, which was evaluated at 0.001).

## 3. RESULTS

We examined the response characteristics of the low and mid-level regions of human visual cortex during observation of blob-like objects with motion properties consistent with different surface attributes (shiny and matte). The objects were rendered either as dynamic intensity flows (the photometric or "image" condition) or as moving light points (the kinematic or "dot" condition). Photometric presentation supports the perception of matte or shiny surface attributes, whereas kinematic presentation, despite containing similar flow patterns, does not evoke a perception of matte or shiny surface structure. The renderings were presented in four quadrants of the visual field (see **Figure 2**) while fMRI was used to measure the BOLD activity from within the posterior region of human visual cortex. This stimulus presentation evoked significant activity levels ( $p_{\text{FDR}} < 0.001$ ), aggregated over shiny and matte renderings and dot flows and compared with a blank screen baseline, within low-level visual areas V1, V2, and V3, dorsal areas LO1, LO2, V3A/B, and hMT+, and ventral areas hV4 and

VO1. We also evaluated the anterior subdivision of the LOC, defined functionally as preferring images of objects over images of faces and houses, but it was not activated consistently across participants and was therefore not analyzed further.

### 3.1. Response Amplitude

We conducted a Two-way ANOVA (flow class: matte, shiny; rendering: image, dot) on the mean response amplitude elicited in each visual area. We predicted that the perceptual sensitivity to material properties from motion would be related to the activity levels in the low and mid-level regions of human visual cortex, and that this would be demonstrated via an interaction between the motion flow class (matte, shiny) and the rendering type (image, dot). However, as shown in **Figure 4**, such an interaction was not evident in any of the investigated visual areas (all  $p_{\text{FDR}} \gg 0.05$ ; see **Table 1**, for details).

We then investigated the potential main effects of flow (matte, shiny) and rendering (image, dot). There was a significant main effect of flow class for the mean BOLD response amplitudes in V1 [ $F_{(1,7)} = 37.62, p_{\text{FDR}} = 0.004$ ], V2 [ $F_{(1,7)} = 20.22, p_{\text{FDR}} = 0.013$ ], LO1 [ $F_{(1,7)} = 9.72, p_{\text{FDR}} = 0.038$ ], and hMT+ [ $F_{(1,7)} = 13.41, p_{\text{FDR}} = 0.024$ ]. Responses were greater for shiny than matte flows in V1 (difference mean = 0.26, SEM = 0.04) and V2 (difference mean = 0.18, SEM = 0.04) while response were greater for matte than shiny in LO1 (difference mean = 0.09, SEM = 0.03) and hMT+ (difference mean = 0.13, SEM = 0.04). The main effect of flow class was not significant in the other visual areas under investigation ( $p_{\text{FDR}} > 0.05$ ; see **Table 1**).

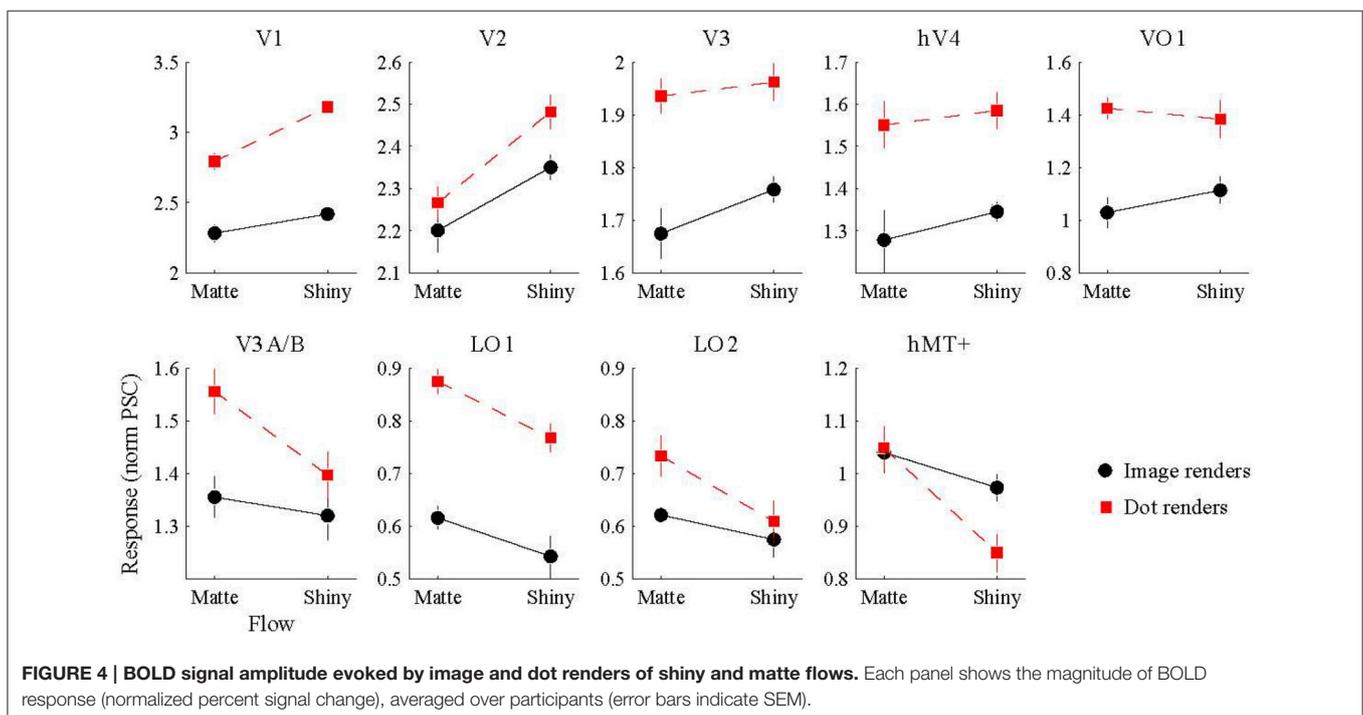
There was also a significant main effect of rendering class for the mean BOLD response amplitudes in V1 [ $F_{(1,7)} = 89.87, p_{\text{FDR}} < 0.001$ ], V3 [ $F_{(1,7)} = 17.72, p_{\text{FDR}} = 0.012$ ], hV4 [ $F_{(1,7)} = 8.69, p_{\text{FDR}} = 0.039$ ], VO1 [ $F_{(1,7)} = 10.43,$

$p_{\text{FDR}} = 0.033$ ], and LO1 [ $F_{(1,7)} = 43.42, p_{\text{FDR}} = 0.001$ ]. Responses were greater for dots than images in each of these areas (V1 difference mean = 0.64, SEM = 0.07; V3 difference mean = 0.23, SEM = 0.06; hV4 difference mean = 0.26, SEM = 0.09; VO1 difference mean = 0.33, SEM = 0.10; LO1 difference mean = 0.24, SEM = 0.04). The main effect of rendering class was not significant in the other visual areas under investigation ( $p_{\text{FDR}} > 0.05$ ; see **Table 1**).

### 3.2. Response Pattern

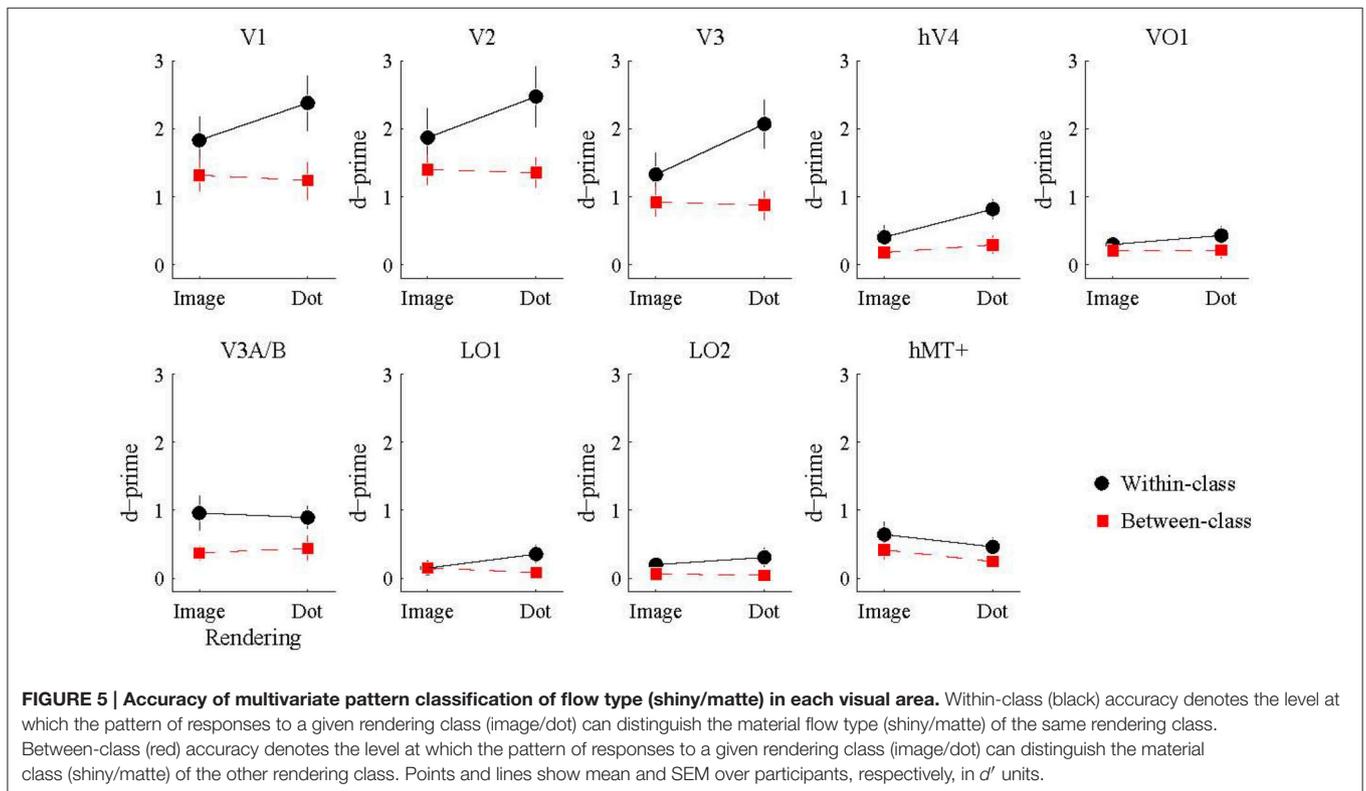
We also investigated whether the multivariate pattern of responses within each visual area could discriminate between shiny and matte flows. We employed a pattern classification procedure to test for visual areas with a representation of surface material from motion flows. For each visual area, a classifier was trained with activation patterns from a particular rendering class (image or dot) and then tested with either the same or different rendering class (within or between-class classification). The rationale was that a visual area representing surface material would be significantly better at discriminating shiny and matte flows for within-class classifications with image renderings, since that is the only combination in which surface material perception differs between shiny and matte in the same way between training and testing. However, as shown in **Figure 5**, this interaction between rendering class (image, dot) and classification type (within, between) was not statistically significant in any of the visual areas under consideration (all  $p_{\text{FDR}} > 0.05$ ; see **Table 2**, for details).

We then considered the main effects of rendering class (image, dot) and classification type (within, between) on the observed classification accuracies. None of the visual areas under consideration had a significant main effect of rendering



**TABLE 1 | Results of Two-way ANOVA for each visual area, conducted on the average signal amplitude for each participant.**

Area	Flow (shiny, matte)			Rendering (image, dots)			Flow × Rendering		
	$F_{(1,7)}$	$p$	$p_{FDR}$	$F_{(1,7)}$	$p$	$p_{FDR}$	$F_{(1,7)}$	$p$	$p_{FDR}$
V1	37.62	< 0.001	0.004	89.87	< 0.001	< 0.001	3.85	0.090	0.509
V2	20.22	0.003	0.013	3.03	0.126	0.161	0.43	0.531	0.710
V3	4.49	0.072	0.092	17.72	0.004	0.012	0.39	0.552	0.710
hV4	1.54	0.254	0.286	8.69	0.021	0.039	0.11	0.747	0.747
VO1	0.31	0.593	0.593	10.43	0.014	0.033	2.28	0.175	0.509
V3A/B	4.81	0.064	0.092	4.58	0.070	0.105	1.76	0.226	0.509
LO1	9.72	0.017	0.038	43.42	< 0.001	0.001	0.21	0.659	0.741
LO2	6.33	0.040	0.072	2.69	0.145	0.163	0.84	0.389	0.701
hMT+	13.41	0.008	0.024	1.42	0.272	0.272	1.98	0.202	0.509



**FIGURE 5 | Accuracy of multivariate pattern classification of flow type (shiny/matte) in each visual area.** Within-class (black) accuracy denotes the level at which the pattern of responses to a given rendering class (image/dot) can distinguish the material flow type (shiny/matte) of the same rendering class. Between-class (red) accuracy denotes the level at which the pattern of responses to a given rendering class (image/dot) can distinguish the material class (shiny/matte) of the other rendering class. Points and lines show mean and SEM over participants, respectively, in  $d'$  units.

class (all  $p_{FDR} > 0.05$ ). The main effect of classification type was significant in V1 [ $F_{(1,7)} = 25.26$ ,  $p_{FDR} = 0.011$ ], V2 [ $F_{(1,7)} = 11.49$ ,  $p_{FDR} = 0.021$ ], V3 [ $F_{(1,7)} = 21.16$ ,  $p_{FDR} = 0.011$ ], hV4 [ $F_{(1,7)} = 18.10$ ,  $p_{FDR} = 0.011$ ], V3A/B [ $F_{(1,7)} = 14.28$ ,  $p_{FDR} = 0.016$ ], and LO2 [ $F_{(1,7)} = 7.08$ ,  $p_{FDR} = 0.049$ ]. As shown in **Figure 5**, accuracies were significantly greater for within-class than between-class classification in these areas.

Finally, we examined the ability of each visual area and classification condition to discriminate shiny and matte flows at levels significantly greater than chance. All of the visual areas under consideration were able to perform within-class classification (image → image, dot → dot) significantly greater than chance (all  $p_{FDR} < 0.05$  except LO2 where  $p_{FDR} =$

0.057; see **Table 3**, for details). Visual areas V1, V2, V3, V3A/B, and hMT+ were also able to perform between-class classification (image → dot, dot → image) significantly greater than chance (all  $p_{FDR} < 0.05$ ). The between-class classification performance was less reliable for hV4 and VO1 ( $p_{FDR}$  between 0.051 and 0.067), and not significantly greater than chance in LO1 and LO2 (all  $p_{FDR} > 0.05$ ).

## 4. DISCUSSION

We measured the responses of visual areas to rotating 3D objects, with photometric or kinematic flows corresponding to shiny or matte objects, while observers were engaged in a

**TABLE 2 | Results of Two-way ANOVA for each visual area, conducted on the accuracy ( $d'$ ) for each participant.**

Area	Classification (within, between)			Rendering (image, dots)			Classification $\times$ Rendering		
	$F_{(1,7)}$	$\rho$	$\rho_{FDR}$	$F_{(1,7)}$	$\rho$	$\rho_{FDR}$	$F_{(1,7)}$	$\rho$	$\rho_{FDR}$
V1	25.26	0.002	0.011	3.02	0.126	0.313	10.32	0.015	0.133
V2	11.49	0.012	0.021	2.72	0.143	0.313	4.91	0.062	0.187
V3	21.16	0.002	0.011	4.61	0.069	0.310	6.27	0.041	0.184
hV4	18.10	0.004	0.011	10.62	0.014	0.125	2.60	0.151	0.226
VO1	3.52	0.103	0.132	0.97	0.358	0.537	2.96	0.129	0.226
V3A/B	14.28	0.007	0.016	0.00	0.998	0.998	0.26	0.627	0.705
LO1	1.13	0.323	0.323	0.64	0.451	0.580	2.90	0.132	0.226
LO2	7.08	0.032	0.049	0.42	0.538	0.605	0.42	0.539	0.693
hMT+	2.31	0.173	0.194	2.29	0.174	0.313	0.00	0.980	0.980

**TABLE 3 | Results of a one-sample  $t$ -test against a chance level of  $d'$  accuracy of 0 for each visual area, conducted on the accuracy of multivariate pattern classification.**

Area	Image $\rightarrow$ Image			Dot $\rightarrow$ Dot			Image $\rightarrow$ Dot			Dot $\rightarrow$ Image		
	$t_7$	$\rho$	$\rho_{FDR}$	$t_7$	$\rho$	$\rho_{FDR}$	$t_7$	$\rho$	$\rho_{FDR}$	$t_7$	$\rho$	$\rho_{FDR}$
V1	4.94	0.001	0.008	5.55	<0.001	0.001	5.28	0.001	0.003	4.23	0.002	0.006
V2	4.05	0.002	0.009	5.19	0.001	0.001	5.86	<0.001	0.003	5.78	<0.001	0.003
V3	3.89	0.003	0.009	5.40	0.001	0.001	4.10	0.002	0.007	3.87	0.003	0.007
hV4	2.32	0.027	0.030	5.32	0.001	0.001	2.06	0.039	0.054	2.16	0.034	0.051
VO1	2.32	0.027	0.030	2.90	0.011	0.015	2.01	0.042	0.054	1.87	0.052	0.067
V3A/B	3.57	0.005	0.010	5.08	0.001	0.001	3.32	0.006	0.014	2.32	0.027	0.048
LO1	2.63	0.017	0.025	2.66	0.016	0.018	1.50	0.088	0.099	1.39	0.104	0.117
LO2	1.81	0.057	0.057	2.12	0.036	0.036	1.07	0.160	0.160	1.07	0.161	0.161
hMT+	3.22	0.007	0.013	3.45	0.005	0.008	2.75	0.014	0.026	5.16	0.001	0.003

demanding fixation task. With almost all of the previous work investigating the cortical processing of material properties with static images (Cant and Goodale, 2007, 2009; Hiramatsu et al., 2011), we focused here on dynamic information for surface reflectance (Doerschner et al., 2011a). While we find no obvious candidates in human low and mid-level visual cortex for an explicit representation of specular reflectance from motion, we do find several areas with response properties that are modulated by the structure of matte and shiny flows.

#### 4.1. Representation of Specular Reflectance from Motion

We did not observe an interaction between rendering type (image, dot) and flow (matte, shiny) in the mean response amplitudes of any of the visual areas under consideration. We also did not observe a difference in the accuracy of pattern classification performance that would be indicative of a representation of specular reflectance from motion in any of the visual areas.

However, the results outlined above do not necessarily indicate that the visual areas investigated in the current study are unable to represent the perception of specular reflectance from motion. Such a representation may produce differences in response amplitude that are too small to be recovered with

the current design, or may be present at a spatial scale that is not visible to the resolution of our fMRI measurements. Furthermore, it is possible that the extraction of specular reflectance from motion requires observer attention to be directed at the stimulus. Our wariness of an attentional confound, in which shiny surfaces are potentially more engaging than matte surfaces, led us to direct our observers' attention to an unrelated task at fixation. However, a potential consequence of this attentional focus is a lack of activation of cortical pathways that are involved in extracting the surface properties from the motion flows. This could perhaps be assessed behaviorally in future studies using an adaptation paradigm (Kam et al., 2012) with differing attentional demands.

#### 4.2. Correlates of Specular Reflectance Motion

We did find several areas in low and mid-level human visual cortex that were affected by whether the flow was consistent with a shiny or matte surface. Shiny and matte flows differ on many dimensions, which raises the question of which aspects of the motion signals are affecting the response properties of the different visual areas. For example, shiny and matte flows can differ in the statistical distribution of velocity parameters (mean and variance of speed and direction; cf. Doerschner et al.,

2011b), luminance contrast, and higher-order regularities. We next consider the potential contribution of such dimensions to the visual regions under investigation.

#### 4.2.1. Effect of matte vs. shiny flow class on average responses

While areas V1, V2, LO1, and hMT+ all had significantly different levels of mean activity for shiny and matte flows, V1 and V2 showed larger responses to shiny than to matte flows, whereas LO1 and hMT+ showed the opposite pattern (Figure 4).

If shiny and matte objects contain different levels of motion energy, that could potentially account for the differences in activity levels we observed in early visual areas. We checked this possibility by computing motion energy for all of our stimuli using Gabor filters (Adelson and Bergen, 1985). While the motion energy between shiny and matte photometric flows did not differ significantly, we found that motion energy for matte dot flows were significantly larger than for shiny dot flows (Supplementary Figure 14 and Supplementary Tables 4, 5). If cortical areas were sensitive solely to motion energy we would expect a significant interaction term for rendering type and material class for identified cortical areas. This was not, however, what we found. Thus, while motion energy might explain the general larger responses of cortical areas to dot flows<sup>1</sup>, they cannot account for the overall observed response pattern.

Macaque V1 (Priebe et al., 2006) and MT (Priebe et al., 2003) contains neurons selective for stimulus speed, and a difference in speed distributions between shiny and matte stimuli could potentially introduce differences in response magnitude into the observed fMRI responses. We computed speed histograms to evaluate such a possibility and find no difference in mean speed between shiny and matte flows (Supplementary Figure 15), making it unlikely that speed is responsible for the effects observed in low-level areas.

Inspecting the corresponding motion direction histograms in Supplementary Figure 15 it becomes apparent that the distributions for matte flows form a cluster around a dominant direction of motion whereas shiny flows tend to be distributed more uniformly—owing to the large variations of flow direction in specular flow (Doerschner et al., 2011a). Considering the different receptive field sizes in V1/V2 compared to hMT+/LO1 (Mikami et al., 1986; Amano et al., 2009) we might be able to partially explain the opposing response patterns to matte and shiny flows in these areas: due to their small receptive field sizes, non-overlapping neural populations in V1/V2 might respond their preferred motion, creating large net response. In hMT+, however, due to the larger receptive field sizes, and opponent processes (Heeger et al., 1999) a given neural population may have preferred and non-preferred direction in their receptive field present, thus creating an overall decreased response. By the same argument we would expect decreased net response in V1/V2 to matte flows and a relatively increased net response in hMT+/LO1. Though variability in motion direction can be related to motion coherency these two are not the same

<sup>1</sup>With the exception of area hMT+ which is known to saturate at lower contrasts.

concept<sup>2</sup>. Thus, we consider the latter as a potential explanation of our results next.

Shiny and matte flows, both kinematic and photometric, differ with respect to the degree of motion coherency. There is large literature which has studied behavioral and neural responses to structure-from-motion stimuli composed of dots whose coherency is manipulated by assigning random velocities to varying proportions of the dots. While this dimension does not map simply on to matte vs. shiny flows, coherency is correlated with distortions in the appearance of specular flows (Doerschner et al., 2011a). Specifically Doerschner et al. (2011a) showed that distortion manifests itself in the degree of expansion/contraction of optic flow (“divergence”), the degree to which image features could be unambiguously tracked across time (“trackability”), and the degree to which features were consistent with rigid-body rotation (“shape reliability”), for matte and shiny surfaces. Both decreased trackability and shape reliability (and perhaps to a lesser extent divergence) are also characteristic of decreased coherence in random dot displays. It has been shown that V1 shows greater activation for random dot motion than for translational coherent motion (Braddick et al., 2001, but see Morrone et al., 2000), whereas areas V5/hMT+ responses increase with increases in motion coherence (Rees et al., 2000; Braddick et al., 2001; Händel et al., 2007 but also see McKeefry et al., 1997; Paradis et al., 2000; Smith et al., 2006) raising the possibility that these stimulus features account in part for the larger response of early visual cortex to shiny over matte, and the opposite pattern found in hMT+.

While this finding confirms the potential usefulness of these three motion cues in identifying reflectance properties, the current experimental design does not permit to identify the relative weight that these cues are given at this early stage of the visual analysis. This will be the topic of future experiments.

It is likely that simple cues like coherence or divergence play a less important role in motion-based material identification at intermediate levels of visual analysis compared to the early stages of visual processing. This idea is supported by the observed larger responses of LO1 to matte flows. However, this increased response might not necessarily indicate sensitivity to motion characteristics—such as coherence—*per se* (Larsson and Heeger, 2006), but rather reflect this area’s sensitivity to 3D structure (Freeman et al., 2012). One hallmark of specular moving objects is that they frequently yield a non-rigid percepts. Thus, these stimuli may provide less information for computing 3D rigid shape from motion than diffusely reflecting moving objects—which might account for the observed pattern of responses.

<sup>2</sup>We equate variability with estimated variance of the motion direction histogram. Coherence is defined as similarity in motion direction *and* speed. The component that relates the two concepts is motion direction, however one can think of all possible constellations of high and low motion direction variability and high and low motion coherence. For example: A given dot motion field can have low motion direction variability and low coherence (if dots move at different speeds, but in the same direction) or low motion direction variability and high coherence (if dots move at the same speeds, in the same direction) or can have high motion direction variability and high coherence (half of the dots move at same speed in one direction and the other half at the same speed at the opposite direction, with non-overlapping dot trajectories) or can have high motion direction variability and low coherence (random direction dot motion).

Taken together, differences in motion coherence in shiny and matte flows appear to offer a fairly good account of the observed pattern of responses in areas V1 and hMT+, keeping in mind though that previous findings on coherence preferences of areas V1, V2, and hMT+ have been mixed. It is likely that it is a combination of factors such as direction variability, coherence and kinetic structure that is responsible for the observed pattern of cortical responses—either directly or potentially indirectly via altering other stimulus properties such as perceived surface brightness, contrast, or motion detection thresholds.

#### 4.2.2. Discrimination of Matte vs. Shiny Flows Across Visual Areas

Visual areas V1, V2, V3, V3A/B, and hMT+ were able to discriminate shiny and matte flows at levels significantly above chance, for both image and dot renders and within and between-class classification. What flow information is common to both photometric and kinematic flows that might explain the discriminative capabilities these visual areas? From the above discussion and stimulus measurements it appears that motion coherence, direction variability, and shape reliability are viable candidates, since these computations would all require to discount the intensity variations in photometric flow and thus be consistent with the extraction of a motion field. In line with this argument, we ruled out motion energy—a computation that does depend on photometric properties (Adelson and Bergen, 1985)—as an explanatory factor.

The response profile of these cortical areas thus suggests that the computation of surface material cannot simply discount structure, but that these two computations are interrelated. Intuitively this makes sense, given that 3D structure is needed, for example, to explain away (via boundary motion) a non-rigid interpretation of optic flow (see Supplementary Movies 6,7), or the interesting perceptual trade-off between perceived shininess and perceived rigidity of moving objects (Doerschner and Kersten, 2007; Zang et al., 2009, 2010; Doerschner et al., 2011a; Doerschner, 2013). Conversely, in some cases material inference may also influence perceived structure (Kersten et al., 1992).

### 4.3. Conclusion

Natural image flows carry information not only about structure, but also about material. Traditional studies of object motion have focused on structure-from-motion and have typically used very simple experimental manipulations such as varying coherence by randomizing the motions of subsets of dots. Dot flows largely miss the space of flows normally experienced and in particular which provide information for material. While were unable to

identify regions of low and mid-level human visual cortex that respond preferentially to material structure from motion, we report the presence of several visual areas that modulate their activity with changes in specular reflectance flows.

Future work needs to be done to study brain mechanisms involved in the interaction between image flow properties and the conscious decisions about material. For example, Cant and Goodale (2007) have shown that attention to static object shape or material modulates different regions of the ventral stream, and more recently Wada et al. (2014) identified ventral and dorsal areas involved in both, the processing of image cues to glossiness as well as the perception of gloss in static images. More work is also needed to continue to quantify higher-order image regularities that capture characteristic motion patterns that support the dimensions along which humans can perceive object material qualities. For example, simple inspection of a rotating shiny object through an aperture illustrates the importance of boundary information and thus shape (Supplementary Movies 6, 7): without shape to “explain away” non-rigid interpretations, one often perceives non-rigid fluid flow.

## ACKNOWLEDGMENTS

TK and SL were supported by ICT R&D program of MSIP/IITP. [R0126-15-1107, Development of Intelligent Pattern Recognition Softwares for Ambulatory Brain-Computer Interface]. DK was supported by ONR (N000141210883), the World Class University program funded by the Ministry of Education, Science, and Technology through the National Research Foundation of Korea (R31-10008). KD was supported by a Marie Curie International Reintegration Grant (239494) within the 7th European Community Framework Programme, an EU Marie Curie Initial Training Network “PRISM” (FP7 - PEOPLE-2012-ITN, Grant Agreement: 316746), a grant of the Scientific and Technological Research Council of Turkey (TUBITAK 1001 Grant No 112K069) and a Turkish Academy of Sciences (TUBA) GEBIP Award for Outstanding Young Scientists. The study was conceived by KD and DK, with experimental design by TK, DM, and DK. Data was collected and analyzed by TK and DM. Lee provided feedback and additional support throughout the project. The paper was written by TK, DM, KD, and DK.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fnhum.2015.00579>

## REFERENCES

- Adelson, E. H. (2001). “On seeing stuff: the perception of materials by humans and machines,” in *Proceedings of the SPIE*, Vol. 4299 (San Jose, CA: Citeseer), 1–12.
- Adelson, E. H., and Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A* 2, 284–299. doi: 10.1364/JOSAA.2.000284

- Amano, K., Wandell, B. A., and Dumoulin, S. O. (2009). Visual field maps, population receptive field sizes, and visual field coverage in the human mt+ complex. *J. Neurophysiol.* 102, 2704–2718. doi: 10.1152/jn.00102.2009
- Anderson, B. L., and Kim, J. (2009). Image statistics do not explain the perception of gloss and lightness. *J. Vis.* 9:10. doi: 10.1167/9.11.10

- Andrews, T. J., Clarke, A., Pell, P., and Hartley, T. (2010). Selectivity for low-level features of objects in the human ventral stream. *Neuroimage* 49, 703–711. doi: 10.1016/j.neuroimage.2009.08.046
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57, 289–300.
- Braddick, O. J., O'Brien, J. M., Wattam-Bell, J., Atkinson, J., Hartley, T., and Turner, R. (2001). Brain areas sensitive to coherent visual motion. *Perception* 30, 61–72. doi: 10.1068/p3048
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vis.* 10, 433–436. doi: 10.1163/156856897X00357
- Bressler, D. W., and Silver, M. A. (2010). Spatial attention improves reliability of fMRI retinotopic mapping signals in occipital and parietal cortex. *Neuroimage* 53, 526–533. doi: 10.1016/j.neuroimage.2010.06.063
- Brewer, A. A., Liu, J., Wade, A. R., and Wandell, B. A. (2005). Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nat. Neurosci.* 8, 1102–1109. doi: 10.1038/nn1507
- Cant, J. S., and Goodale, M. A. (2007). Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cereb. Cortex* 17, 713–731. doi: 10.1093/cercor/bhk022
- Cant, J. S., and Goodale, M. A. (2009). Asymmetric interference between the perception of shape and the perception of surface properties. *J. Vis.* 9:13. doi: 10.1167/9.5.13
- Cant, J. S., and 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. doi: 10.1523/JNEUROSCI.6113-10.2011
- Cavina-Pratesi, C., Kentridge, R. W., Heywood, C. A., and Milner, A. D. (2010a). Separate channels for processing form, texture, and color: evidence from fMRI adaptation and visual object agnosia. *Cereb. Cortex* 20, 2319–2332. doi: 10.1093/cercor/bhp298
- Cavina-Pratesi, C., Kentridge, R. W., Heywood, C. A., and Milner, A. D. (2010b). Separate processing of texture and form in the ventral stream: evidence from fMRI and visual agnosia. *Cereb. Cortex* 20, 433–446. doi: 10.1093/cercor/bhp111
- Chadwick, A. C., and Kentridge, R. (2015). The perception of gloss: a review. *Vis. Res.* 109, 221–235. doi: 10.1016/j.visres.2014.10.026
- Chang, C.-C., and Lin, C.-J. (2011). LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2, 27:1–27:27. doi: 10.1145/1961189.1961199
- Cousineau, D. (2005). Confidence intervals in within-subject designs: a simpler solution to Loftus and Masson's method. *Tutorial Quant. Methods Psychol.* 1, 42–45.
- Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173. doi: 10.1006/cbmr.1996.0014
- Dale, A. M., Fischl, B., and Sereno, M. I. (1999). Cortical surface-based analysis. I: segmentation and surface reconstruction. *Neuroimage* 9, 179–194. doi: 10.1006/nimg.1998.0395
- Debevec, P. (2002). Image-based lighting. *IEEE Comput. Graph. Appl.* 22, 26–34. doi: 10.1109/38.988744
- DeYoe, E. A., Carman, G. J., Bandettini, P., Glickman, S., Wieser, J., Cox, R., et al. (1996). Mapping striate and extrastriate visual areas in human cerebral cortex. *Proc. Natl. Acad. Sci. U.S.A.* 93, 2382–2386. doi: 10.1073/pnas.93.6.2382
- Doerschner, K. (2013). "Image motion and the appearance of objects," in *Handbook of Experimental Phenomenology: Visual Perception of Shape, Space and Appearance*. (Chichester, UK: John Wiley & Sons), 223–242.
- Doerschner, K., Boyaci, H., and Maloney, L. (2010). Estimating the glossiness transfer function induced by illumination change and testing its transitivity. *J. Vis.* 10:8. doi: 10.1167/10.4.8
- Doerschner, K., Fleming, R. W., Yilmaz, O., Schrater, P. R., Hartung, B., and Kersten, D. (2011a). Visual motion and the perception of surface material. *Curr. Biol.* 21, 2010–2016. doi: 10.1016/j.cub.2011.10.036
- Doerschner, K., and Kersten, D. (2007). Perceived rigidity of rotating specular superellipsoids under natural and not-so-natural illuminations. *J. Vis.* 7:838. doi: 10.1167/7.9.838
- Doerschner, K., Kersten, D., and Schrater, P. R. (2011b). Rapid classification of specular and diffuse reflection from image velocities. *Pattern Recognit.* 44, 1874–1884. doi: 10.1016/j.patcog.2010.09.007
- Dougherty, R. F., Koch, V. M., Brewer, A. A., Fischer, B., Modersitzki, J., and Wandell, B. A. (2003). Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *J. Vis.* 3, 586–598. doi: 10.1167/3.10.1
- Dror, R., Adelson, E., and Willsky, A. (2001). "Estimating surface reflectance properties from images under unknown illumination," in *SPIE Photonics West: Human Vision and Electronic Imaging VI* (San Jose, CA), 231–242.
- Engel, S. A., Glover, G. H., and Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cereb. Cortex* 7, 181–192. doi: 10.1093/cercor/7.2.181
- Fischl, B., Sereno, M. I., and Dale, A. M. (1999). Cortical surface-based analysis. II: inflation, flattening, and a surface-based coordinate system. *Neuroimage* 9, 195–207. doi: 10.1006/nimg.1998.0396
- Fleming, R. W. (2012). Human perception: visual heuristics in the perception of glossiness. *Curr. Biol.* 22, R865–R866. doi: 10.1016/j.cub.2012.08.030
- Fleming, R. W., Torralba, A., and Adelson, E. H. (2004). Specular reflections and the perception of shape. *J. Vis.* 4, 798–820. doi: 10.1167/4.9.10
- Fleming, R. W., Wiebel, C., and Gegenfurtner, K. (2013). Perceptual qualities and material classes. *J. Vis.* 13:9. doi: 10.1167/13.8.9
- Freeman, E., Sterzer, P., and Driver, J. (2012). fMRI correlates of subjective reversals in ambiguous structure-from-motion. *J. Vis.* 12:35. doi: 10.1167/12.6.35
- Gautama, T., and Van Hulle, M. (2002). A phase-based approach to the estimation of the optical flow field using spatial filtering. *IEEE Trans. Neural Netw.* 13, 1127–1136. doi: 10.1109/TNN.2002.1031944
- Gkioulekas, I., Xiao, B., Zhao, S., Adelson, E. H., Zickler, T., and Bala, K. (2013). Understanding the role of phase function in translucent appearance. *ACM Trans. Graph.* 32:147. doi: 10.1145/2516971.2516972
- Goddard, E., Mannion, D. J., McDonald, J. S., Solomon, S. G., and Clifford, C. W. G. (2011). Color responsiveness argues against a dorsal component of human V4. *J. Vis.* 11, 1–21. doi: 10.1167/11.4.3
- Gonzalez, R. C., and Woods, R. E. (2008). *Digital Image Processing, 3rd Edn.* Upper Saddle River, NJ: Prentice Hall.
- Händel, B., Lutzenberger, W., Thier, P., and Haarmeier, T. (2007). Opposite dependencies on visual motion coherence in human area mt+ and early visual cortex. *Cereb. Cortex* 17, 1542–1549. doi: 10.1093/cercor/bhl063
- Hansen, K. A., Kay, K. N., and Gallant, J. L. (2007). Topographic organization in and near human visual area V4. *J. Neurosci.* 27, 11896–11911. doi: 10.1523/JNEUROSCI.2991-07.2007
- Hartung, B., and Kersten, D. (2002). Distinguishing shiny from matte. *J. Vis.* 2, 551–551. doi: 10.1167/2.7.551
- Harvey, D. Y., and Burgund, E. D. (2012). Neural adaptation across viewpoint and exemplar in fusiform cortex. *Brain Cogn.* 80, 33–44. doi: 10.1167/2.7.551
- Heeger, D. J., Boynton, G. M., Demb, J. B., Seidemann, E., and Newsome, W. T. (1999). Motion opponency in visual cortex. *J. Neurosci.* 19, 7162–7174.
- Hiramatsu, C., Goda, N., and Komatsu, H. (2011). Transformation from image-based to perceptual representation of materials along the human ventral visual pathway. *Neuroimage* 57, 482–494. doi: 10.1016/j.neuroimage.2011.04.056
- Ho, Y., Landy, M., and Maloney, L. (2006). How direction of illumination affects visually perceived surface roughness. *J. Vis.* 6:8. doi: 10.1167/6.5.8
- Kam, T.-E., Kersten, D., Fleming, R. W., Lee, S.-W., and Doerschner, K. (2012). Visual adaptation to reflectance-specific image motion. *J. Vis.* 12:871. doi: 10.1167/12.9.871
- Kersten, D., Bühlhoff, H. H., Schwartz, B. L., and Kurtz, K. J. (1992). Interaction between transparency and structure from motion. *Neural Comput.* 4, 573–589. doi: 10.1162/neco.1992.4.4.573
- Kim, J., and Anderson, B. L. (2010). Image statistics and the perception of surface gloss and lightness. *J. Vis.* 10:3. doi: 10.1167/10.9.3
- Kim, J., Marlow, P. J., and Anderson, B. L. (2011). The perception of gloss depends on highlight congruence with surface shading. *J. Vis.* 11:4. doi: 10.1167/11.9.4
- Kim, J., Marlow, P. J., and Anderson, B. L. (2012). The dark side of gloss. *Nat. Neurosci.* 15, 1590–1595. doi: 10.1038/nn.3221
- Kleiner, M., Brainard, D., and Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception* 36:14.
- Köteles, K., De Maziere, P. A., Van Hulle, M., Orban, G. A., and Vogels, R. (2008). Coding of images of materials by macaque inferior temporal cortical neurons. *Eur. J. Neurosci.* 27, 466–482. doi: 10.1111/j.1460-9568.2007.06008.x
- Larson, G. W., Shakespear, R., Ehrlich, C., and Mardaljevic, J. (1998). *Rendering with Radiance: the Art and Science of Lighting Visualization*. San Francisco, CA: Morgan Kaufmann.

- Larsson, J., and Heeger, D. J. (2006). Two retinotopic visual areas in human lateral occipital cortex. *J. Neurosci.* 26, 13128–13142. doi: 10.1523/JNEUROSCI.1657-06.2006
- Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc. Natl. Acad. Sci. U.S.A.* 92, 8135–8139. doi: 10.1073/pnas.92.18.8135
- Marlow, P., Kim, J., and Anderson, B. (2011). The role of brightness and orientation congruence in the perception of surface gloss. *J. Vis.* 11:16. doi: 10.1167/11.9.16
- Marlow, P. J., Kim, J., and Anderson, B. L. (2012). The perception and misperception of specular surface reflectance. *Curr. Biol.* 22, 1909–1913. doi: 10.1016/j.cub.2012.08.009
- McKeefry, D. J., Watson, J. D., Frackowiak, R., Fong, K., and Zeki, S. (1997). The activity in human areas v1/v2, v3, and v5 during the perception of coherent and incoherent motion. *Neuroimage* 5, 1–12. doi: 10.1006/nimg.1996.0246
- Mikami, A., Newsome, W. T., and Wurtz, R. H. (1986). Motion selectivity in macaque visual cortex. ii. spatiotemporal range of directional interactions in mt and v1. *J. Neurophysiol.* 55, 1328–1339.
- Morrone, M. C., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, G., and Burr, D. C. (2000). A cortical area that responds specifically to optic flow, revealed by fmri. *Nat. Neurosci.* 3, 1322–1328. doi: 10.1038/81860
- Motoyoshi, I., Nishida, S., Sharan, L., and Adelson, E. H. (2007). Image statistics and the perception of surface qualities. *Nature* 447, 206–209. doi: 10.1038/nature05724
- Nishida, S., and Shinya, M. (1998). Use of image-based information in judgments of surface-reflectance properties. *J. Opt. Soc. Am. A* 15, 2951–2965. doi: 10.1364/JOSAA.15.002951
- Nishio, A., Goda, N., and Komatsu, H. (2012). Neural selectivity and representation of gloss in the monkey inferior temporal cortex. *J. Neurosci.* 32, 10780–10793. doi: 10.1523/JNEUROSCI.1095-12.2012
- Okazawa, G., Goda, N., and Komatsu, H. (2012). Selective responses to specular surfaces in the macaque visual cortex revealed by fmri. *Neuroimage* 63, 1321–1333. doi: 10.1016/j.neuroimage.2012.07.052
- Olkkonen, M., and Brainard, D. (2010). Perceived glossiness and lightness under real-world illumination. *J. Vis.* 10:5. doi: 10.1167/10.9.5
- Paradis, A. L., Cornilleau-Péres, V., Droulez, J., Van De Moortele, P. F., Lobel, E., Berthoz, A., et al. (2000). Visual perception of motion and 3-d structure from motion: an fmri study. *Cereb. Cortex* 10, 772–783. doi: 10.1093/cercor/10.8.772
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vis.* 10, 437–442. doi: 10.1163/156856897X00366
- Peuskens, H., Claeys, K. G., Todd, J. T., Norman, J. F., Van Hecke, P., and 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. doi: 10.1162/0899892904323057371
- Priebe, N. J., Cassanello, C. R., and Lisberger, S. G. (2003). The neural representation of speed in macaque area mt/v5. *J. Neurosci.* 23, 5650–5661.
- Priebe, N. J., Lisberger, S. G., and Movshon, J. A. (2006). Tuning for spatiotemporal frequency and speed in directionally selective neurons of macaque striate cortex. *J. Neurosci.* 26, 2941–2950. doi: 10.1523/JNEUROSCI.3936-05.2006
- Rees, G., Friston, K., and Koch, C. (2000). A direct quantitative relationship between the functional properties of human and macaque v5. *Nat. Neurosci.* 3, 716–723. doi: 10.1038/76673
- Saad, Z. S., Glen, D. R., Chen, G., Beauchamp, M. S., Desai, R., and Cox, R. W. (2009). A new method for improving functional-to-structural MRI alignment using local Pearson correlation. *Neuroimage* 44, 839–848. doi: 10.1016/j.neuroimage.2008.09.037
- Saad, Z. S., Reynolds, R. C., Argall, B., Japee, S., and Cox, R. W. (2004). “SUMA: an interface for surface-based intra- and inter-subject analysis with AFNI,” in *Proceedings of IEEE International Symposium on Biomedical Imaging: Nano to Macro* (Arlington, TX), 1510–1513.
- Sakano, Y., and Ando, H. (2008). Effects of self-motion on gloss perception. *Perception* 37:77.
- Schira, M. M., Wade, A. R., and Tyler, C. W. (2007). Two-dimensional mapping of the central and parafoveal visual field to human visual cortex. *J. Neurophysiol.* 97, 4284–4295. doi: 10.1152/jn.00972.2006
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268, 889–893. doi: 10.1126/science.7754376
- Smith, A. T., Wall, M. B., Williams, A. L., and Singh, K. D. (2006). Sensitivity to optic flow in human cortical areas mt and mst. *Eur. J. Neurosci.* 23, 561–569. doi: 10.1111/j.1460-9568.2005.04526.x
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23(Supp. 1), S208–S219. doi: 10.1016/j.neuroimage.2004.07.051
- Stanislaw, H., and Todorov, N. (1999). Calculation of signal detection theory measures. *Behav. Res. Methods Instrum. Comput.* 31, 137–149. doi: 10.3758/BF03207704
- te Pas, S., and Pont, S. (2005). “A comparison of material and illumination discrimination performance for real rough, real smooth and computer generated smooth spheres,” in *Proceedings of the 2nd Symposium on Applied Perception in Graphics and Visualization* (New York, NY: ACM), 75–81.
- Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., et al. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J. Neurosci.* 15, 3215–3230.
- Wada, A., Sakano, Y., and Hiroshi, A. (2014). Human cortical areas involved in perception of surface glossiness. *Neuroimage* 98, 243–257. doi: 10.1016/j.neuroimage.2014.05.001
- Wade, A. R., Brewer, A. A., Rieger, J. W., and Wandell, B. A. (2002). Functional measurements of human ventral occipital cortex: retinotopy and colour. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 357, 963–973. doi: 10.1098/rstb.2002.1108
- Wendt, G., Faul, F., Ekroll, V., and Mausfeld, R. (2010). Disparity, motion, and color information improve gloss constancy performance. *J. Vis.* 10:7. doi: 10.1167/10.9.7
- Wijntjes, M. W., and Pont, S. C. (2010). Illusory gloss on Lambertian surfaces. *J. Vis.* 10:13. doi: 10.1167/10.9.13
- Zaidi, Q. (2011). Visual inferences of material changes: color as clue and distraction. *Wiley Interdiscipl. Rev.* 2, 686–700. doi: 10.1002/wcs.148
- Zang, D., Doerschner, K., and Schrater, P. R. (2009). “Rapid inference of object rigidity and reflectance using optic flow,” in *Proceedings of the 13th International Conference on Computer Analysis of Images and Patterns* (Münster: Springer-Verlag), 888.
- Zang, D., Schrater, P. R., and Doerschner, K. (2010). “Object rigidity and reflectivity identification based on motion analysis,” in *Image Processing (ICIP), 2010 17th IEEE International Conference on* (Hong Kong: IEEE), 4573–4576.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Kam, Mannion, Lee, Doerschner and Kersten. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.