

## The Extraction of 3D Shape from Texture and Shading in the Human Brain

Svetlana S. Georgieva<sup>1</sup>, James T. Todd<sup>2</sup>, Ronald Peeters<sup>3</sup> and Guy A. Orban<sup>1</sup>

<sup>1</sup>Laboratorium voor Neuro- en Psychofysiologie, Katholieke Universiteit Leuven School of Medicine, Campus Gasthuisberg, B-3000 Leuven, Belgium, <sup>2</sup>Department of Psychology, The Ohio State University, Columbus, OH 43210, USA and <sup>3</sup>Department of Radiology, Universitair Ziekenhuis Gasthuisberg, UZ Gasthuisberg, B-3000 Leuven, Belgium

**We used functional magnetic resonance imaging to investigate the human cortical areas involved in processing 3-dimensional (3D) shape from texture (SfT) and shading. The stimuli included monocular images of randomly shaped 3D surfaces and a wide variety of 2-dimensional (2D) controls. The results of both passive and active experiments reveal that the extraction of 3D SfT involves the bilateral caudal inferior temporal gyrus (caudal ITG), lateral occipital sulcus (LOS) and several bilateral sites along the intraparietal sulcus. These areas are largely consistent with those involved in the processing of 3D shape from motion and stereo. The experiments also demonstrate, however, that the analysis of 3D shape from shading is primarily restricted to the caudal ITG areas. Additional results from psychophysical experiments reveal that this difference in neuronal substrate cannot be explained by a difference in strength between the 2 cues. These results underscore the importance of the posterior part of the lateral occipital complex for the extraction of visual 3D shape information from all depth cues, and they suggest strongly that the importance of shading is diminished relative to other cues for the analysis of 3D shape in parietal regions.**

**Keywords:** 3D shape, fMRI, human, shading, texture

### Introduction

There are many different aspects of optical stimulation that are known to provide perceptually useful information about an object's 3-dimensional (3D) form. The most frequently studied of these 3D depth cues involve systematic transformations among multiple views of an object, such as the optical deformations that occur when an object is observed in motion, or the disparity between each eye's view in binocular vision. However, there are other important sources of information about 3D shape that are available within individual static images. To demonstrate this more clearly, it is useful to consider some example images that are presented in the upper row of Figure 1. Note in particular how a compelling impression of 3D surface structure can be obtained from gradients of optical texture (Fig. 1A) or from patterns of image shading (Fig. 1B, see also Fig. S1).

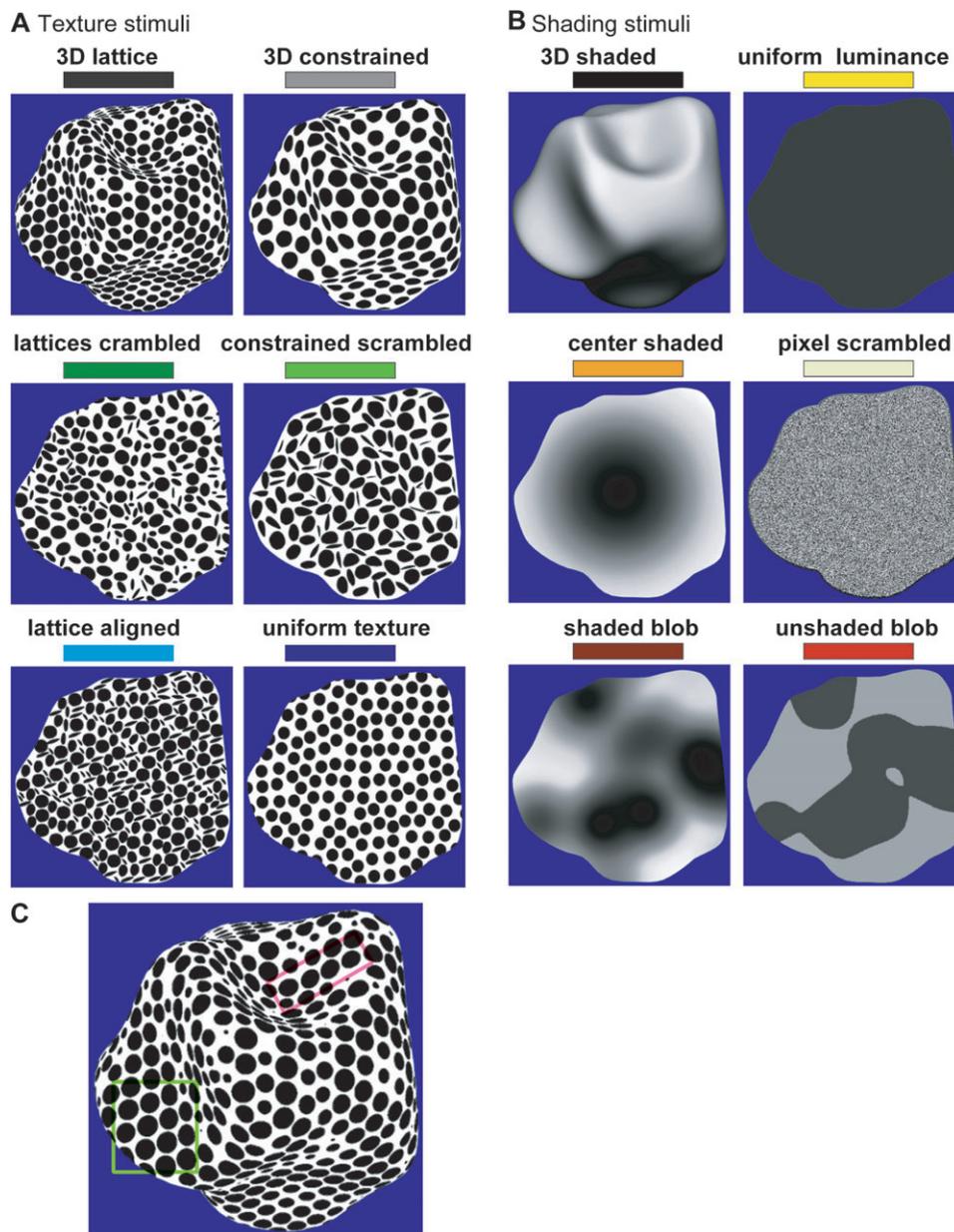
There have been numerous psychophysical studies reported in the literature that have investigated the perceptual analysis of 3D shape from texture (3D SfT) (e.g., Blake et al. 1993; Knill 1998a, 1998b; Todd et al. 2004, 2005). For example, in 1 recent study by Todd et al. (2004) observers judged the surface depth profiles of randomly shaped textured objects like the one shown in the upper panels of Figure 1A. The results revealed that observers' judgments about the magnitude of surface relief were systematically underestimated, but that the overall

apparent shapes of the surfaces were almost perfectly correlated with the ground truth. This pattern of results is quite similar to those obtained for the perception of 3D shape from motion (3D SfM) or stereo (e.g., Todd and Norman 2003).

Image shading is a more complicated source of information than other cues because it is affected not only by the surface geometry, but also the surface reflectance function and the overall pattern of illumination (Koenderink and van Doorn 2004). Moreover, whereas motion, stereo and texture can provide useful information about planar surfaces, shading is only relevant for the analysis of curved surfaces. Another interesting aspect of the perception of shape from shading (SfS) is that observers' judgments are less tightly coupled to the ground truth than is the case for other cues. For example, several studies have shown that the perceived shape of a surface can be systematically distorted by a shearing transformation relative to the ground truth (Koenderink et al. 2001; Cornelis et al. 2003; Todd 2004).

Unfortunately, there has been relatively little research on the functional neural mechanisms by which 3D shape is determined from texture and shading within the human visual system. 3D shape in this context refers to systematic variations of surface depth, orientation and curvature, as opposed to a 2-dimensional (2D) shape that is defined by a discontinuity in depth (e.g., Mendola et al. 1999; Gilaie-Dotan et al. 2001; Kourtzi et al. 2003). To effectively study neuronal mechanisms of 3D shape extraction from texture or shading the experimental design must meet a number of criteria. First, the texture or shading information should not be confounded by other sources of information about 3D shape such as patterns of polyhedral edges or vertices. Second, appropriate controls should be included to show that the mechanisms are sensitive to the appearance of 3D shape as opposed to covarying 2D properties that are always present in monocular images of 3D surfaces. Third, experimental stimuli should include a variety of topographical surface features such as hills, ridges, valleys and dimples in order to ensure the generality of the findings.

The only previous imaging study that has investigated the neural processing of 3D SfT was reported by Shikata et al. (2001). Using functional neural imaging, they compared the patterns of activation that occurred while subjects judged either the color or 3D slant of planar surface patches defined by gradients of optical texture. Their results suggest that surface orientation discrimination activates caudal and anterior regions of the intraparietal sulcus (IPS), though they found no significant activations in the ventral cortex as one might have expected on the basis of prior research using other depth cues (e.g., Orban et al. 1999; Moore and Engel 2001; Vanduffel et al. 2002; Kourtzi et al. 2003; Peuskens et al. 2004). Two potential



**Figure 1.** Visual stimuli of the main experiments. (A) Texture stimuli: 3D lattice (black), 3D constrained (gray), constrained scrambled (light green), lattice scrambled (dark green), uniform texture (dark blue); lattice aligned (light blue). (B) Shading stimuli: 3D shaded (black), center shaded (orange), shaded blob (dark red), unshaded blob (light red), uniform luminance (dark yellow), and pixel scrambled (light yellow). (C) Parts of image with aligned dots (red rectangle) and patch of uniform dots (green square) in 3D lattice stimulus.

problems with this study are 1) the stimulus set was restricted to just planar surfaces, which are a very small subset of possible 3D shapes; and 2) no 2D controls were provided to show that subjects were not using 2D properties, such as the position of the coarsest dots or the orientation of the dot alignments in order to perform the slant judgments. One indication that subjects may have indeed used 2D orientation is that the pattern of parietal activation was similar to that obtained by Faillenot et al. (2001) for grating orientation discrimination.

There have been 4 previous studies that have investigated the neural processing of 3D Sfs (Humphrey et al. 1997; Moore and Engel 2001; Taira et al. 2001; Kourtzi et al. 2003). One of these by Humphrey et al. (1997) was focused exclusively on the primary visual cortex (V1) as a limited area of interest. The

stimulus set was restricted to just 4 stimuli in which shading and direction of illumination were confounded. Two other studies by Kourtzi et al. (2003) and Moore and Engel (2001) had a similarly limited focus on the lateral occipital complex (LOC). Their results indicate that there is greater activation in this area when a stimulus is perceived as a 3D volumetric object rather than as a 2D pattern of contours (e.g., a silhouette). Figure S2 shows a multipart object similar to the ones used in those studies. Note in this figure that there are 2 sources of information about 3D shape: the smooth gradients of shading and the discontinuities that define the pattern of edges and vertices. Moreover, this latter source of information is sufficient to produce a compelling perception of 3D shape, even in the absence of smooth shading gradients. Because of

this confound, it cannot be determined which of the available depth cues is responsible for the reported results. Taira et al. (2001) are the only previous researchers who have investigated the neural processing of 3D Sfs in both dorsal and ventral regions of the human cortex. They compared the patterns of activation that occurred when observers made judgments about the sign of surface curvature from shading relative to a color discrimination control task. A multisubject conjunction analysis of 3D shape versus color processing revealed significant activation in the right parietal cortex, but there were no significant activations in the LOC or any other region within the ventral visual pathways. The problems with this study are the absence of any 2D controls, and a limited set of just 6 stimulus surfaces that all had the same shape except for the sign and magnitude of relief.

In light of the shortcomings described above and the striking discrepancies in the reported results, the research described in the present article was designed to provide a more rigorously controlled investigation of the neural processing of 3D SFT and 3D Sfs using whole-brain scans to measure possible activation sites throughout the visual system, with a wide variety of 3D surfaces and a wide variety of control stimuli to eliminate possible confounds with low-level 2D stimulus attributes. The goals of this research were 3-fold. The first and foremost aim of the study was the identification of the cortical regions that are involved in the processing of 3D SFT and 3D Sfs. To that end we scanned a large number of subjects who were either passive with respect to the stimuli presented or performed a task with the stimuli. Second, by presenting the same objects with both sources of information to the same subjects, we intended to determine the extent to which there is convergence or specialization in the neural processing of different depth cues. Finally, we addressed the question of whether 3D shape is processed in the ventral cortex throughout the LOC as suggested by some researchers (Moore and Engel 2001), or only in the posterior regions of this complex, as suggested for other cues (Orban et al. 1999; Vanduffel et al. 2002; Peuskens et al. 2004).

## Materials and Methods

### Subjects

We performed functional magnetic resonance (MR) measurements on 18 right-handed healthy human volunteers (8 males and 10 females, mean age 25 years, range: 20–33). All of them participated in the main experiments and subsets of them in each of the 3 control experiments. In addition, 5 psychophysical experiments were also executed outside the scanner. The 18 subjects and an additional 10 subjects participated in the first psychophysical experiment (*depth magnitude estimation task*), all 28 subject in its shading version and 12 of them in its texture version. Six out of the 18 subjects participated in the second experiment in which a *3D shape adjustment task* was used to evaluate the different conditions of each main experiment. Nine out of the 18 subjects participated in the third psychophysical experiment, in which the same *3D shape adjustment task* was used to compare the strength of the shading and texture cues. Finally, the 6 subjects of experiment 2 also participated in 2 additional psychophysical studies (experiments 4 and 5) in which the sensitivity for the 2 cues was further investigated.

All subjects had normal or corrected-to-normal vision using contact lenses, and were drug free. None of them had any history of mental illness or neurological disease. All subjects were given detailed instructions for the experiments. They provided written informed consent before participating in the study in accordance with the

Helsinki Declaration and the study was approved by the Ethical Committee of the K.U. Leuven Medical School.

All subjects wore an eye patch over their right eye to eliminate conflicting 3D information from binocular vision (except in the localizer scans and the retinotopy mapping), and their head movements were immobilized using an individually molded bite-bar and by means of small vacuum pillows.

Subjects were asked to maintain fixation on a small red target ( $0.45^\circ \times 0.45^\circ$ ) in the center of the screen during all experiments, except when performing a high-acuity task (Vanduffel et al. 2001) in which the target was replaced with a red bar and in a 1-back task in which the fixation target was smaller ( $0.2^\circ \times 0.2^\circ$ ). Eye movements were recorded (60 Hz) during all of the functional magnetic resonance imaging (fMRI) experiments using the MR-compatible ASL eye tracking system 5000 (Applied Science Laboratories, Bedford, MA).

### Stimuli and Tasks

Visual stimuli were projected from a liquid crystal display projector (Barco Reality 6400i,  $1024 \times 768$ , 60 Hz refresh frequency) onto a translucent screen positioned in the bore of the magnet at a distance of 36 cm from the point of observation. Subjects viewed the stimuli through a  $45^\circ$  tilted mirror attached to the head coil.

### Main Experiments

**Texture and shading stimuli.** The visual stimuli were created and rendered using 3D Studio Max. They depicted 11 randomly generated complex 3D surfaces, representing the front surface of meaningless 3D objects, with a large assortment of variably shaped hills, ridges, valleys, and dimples, at multiple scales (see Norman et al. 1995, 2004; Fleming et al. 2004; Todd et al. 2004). The images of these complex surfaces were presented on a blue background ( $34^\circ \times 16.5^\circ$ ,  $27.6 \text{ cd/m}^2$ ). To quantitatively assess the variety of 3D structure in these displays we aligned all the surfaces in terms of size and position, and calculated a depth map for each image based on the 3D scene geometry that had been used to render it. We then correlated the depths at corresponding positions for each pairwise combination of surfaces. The resulting correlations produced  $r^2$  values that ranged from 0.02 to 0.44. The median of the distribution had an  $r^2$  of 0.184, and the first and third quartiles were 0.133 and 0.243, respectively. In other words, the different shapes we employed were largely independent of one another, with less than 20% overlap on average. This indicates that even if the overall 3D shape of the surfaces was convex, typical of most small objects, the variations around this average were large enough to create largely different 3D shapes. Additional variation was also created by presenting the displays at a variety of different sizes as is shown in Figure S1. All of the surfaces were smoothly curved, so they did not provide information from configurations of edges and vertices (e.g., Moore and Engel 2001; Kourtzi et al. 2003). Examples of the different stimulus types are presented in Figure 1A and B and a complete set of 3D shapes (with shading) is shown in Figure S1. When projected onto the translucent display screen in the bore of the magnet, the sizes of the depicted surfaces in the shading and texture stimuli averaged  $10^\circ$ .

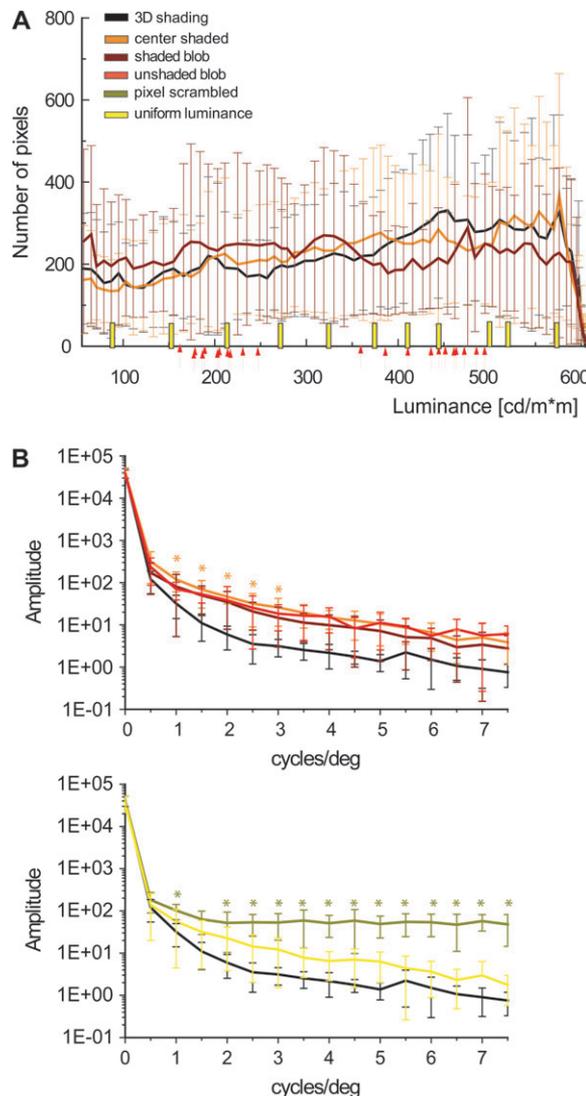
In the 3D SFT experiment, the shapes were presented with 2 different types of volumetric texture that will be referred to, respectively, as the 3D lattice and 3D constrained conditions (Fig. 1A). In both cases, the texture was composed of a set of small spheres that were distributed without overlapping in a 3D volume. Any region of the depicted surface that cut through a sphere was colored black, and any region that cut through the space between spheres was colored white. In the 3D lattice condition, the spheres were arranged in a hexagonal lattice within the texture volume. Note in this case that a local region of an object could cut through the center of a sphere, which would produce a large black dot on the object's surface, or it could just graze through the periphery of the sphere, which would produce a much smaller black dot. Thus, in the 3D lattice conditions, the depicted surfaces were covered with a pattern of circular polka dots that varied in size, and could be systematically aligned along the symmetry axes of the texture lattice. To eliminate these systematic alignments and variations of size, we also employed a 3D constrained condition, in which the spheres were distributed in 3D space such that

their centers were constrained to lie on the depicted surface at randomly selected positions. The impact of this constraint is that all of the polka dots on a depicted surface had the same size, and they were not systematically aligned with one another.

We also included several control conditions in which the patterns of texture did not produce a compelling perception of a 3D surface. These included transformed versions of the 3D lattice and constrained conditions, in which the positions of the texture elements were randomly scrambled within the boundaries of each object. These will be referred to, respectively, as the (2D) lattice-scrambled condition and the (2D) constrained-scrambled conditions (Fig. 1A). The 3D lattice condition differs from its scrambled version not only by the presence of gradients but also by alignments of identical elements or patches of identical elements (boxes in Fig. 1C). In an effort to disentangle these properties, a (2D) lattice-aligned condition was included that eliminated the systematic texture gradients of the 3D lattice displays, but had a similar pattern of texture element alignments. Each stimulus contained 3–4 alignments of 3–6 identical elements. In the (2D) uniform-texture condition all of the projected texture elements had the same circular or elliptic shape (Fig. 1A). Although constant within a stimulus, the elements differed across stimuli: their size ranged from  $0.12^\circ$  to  $1^\circ$ , their elongation ranged from circular to 4/1 ratio and when elongated, elements also differed in orientation. In all texture conditions the 11 stimuli were presented ranging in size from  $8^\circ$  to  $12^\circ$ .

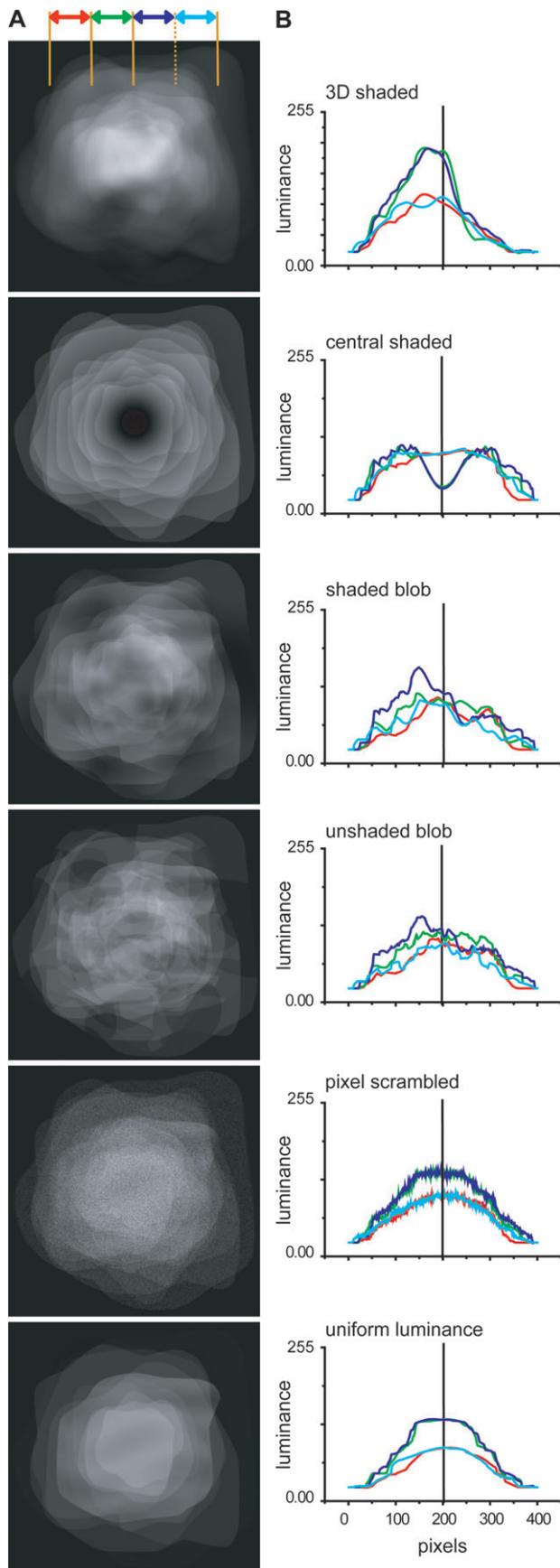
In the SfS experiment (Fig. 1B and Fig. S1), the surfaces in the 3D shaded condition (Fig. S1) were illuminated by a rectangular area light at a  $22^\circ$  angle directly above the line of sight, and they were rendered using a standard Blinn reflectance model, in which the shading at each point is determined as a linear combination of its ambient, diffuse and specular components (mean luminance  $367 \text{ cd/m}^2$ ). In the main experiment the reflectance was Lambertian, with no specular component. A number of control conditions were included in which the patterns of shading did not produce a compelling perception of a 3D surface, yet they had luminance histograms and/or Fourier amplitude spectra that were closely matched to those of the 3D displays (Fig. 2). The first method we employed for eliminating the appearance of depth in the (2D) pixel-scrambled condition was to randomly reposition the pixels ( $2.3 \times 2.3$  minarc) within the boundary of each object. The luminance histograms in these displays were identical to those in the 3D shaded condition, but the local luminance gradients were quite different. Note in Figure 2 that the 3D shaded stimuli contained relatively large regions of nearly uniform luminance. The 2D uniform-luminance condition was designed to create flat looking stimuli that shared this aspect of the 3D displays. The stimuli in that condition included 11 silhouettes of different uniform luminance covering the same luminance range as in the 3D shaded condition (Fig. 2A, vertical straight yellow bars). Two additional control conditions were created that attempted to mimic the pattern of shading gradients in the 3D displays without eliciting the appearance of a 3D surface. In the center-shaded condition, all stimuli had a luminance pattern that increased radially from the center of each silhouette. In the (2D) shaded-blob condition each silhouette contained 3–5 randomly shaped ovals with blurred edges on a light background. A 1-way analysis of variance (ANOVA) revealed that the luminance histograms in these latter 2 conditions did not differ significantly from that of the 3D condition. Finally, a 2D unshaded-blob condition was included that was identical to the shaded blobs, except that all of the smooth luminance gradients were eliminated. This was achieved by thresholding the image intensities to contain just 2 possible luminance values. In all shading conditions, the 11 different stimuli were presented ranging in size from  $5^\circ$  to  $15^\circ$  (Fig. S1).

It is important to keep in mind when evaluating the different control conditions for investigating the perception of 3D SfS or 3D SfT that it is not possible to create stimuli that are perceived as flat and share all the low-level 2D properties of the images used in the 3D conditions. The only viable solution to this problem is to use a wide battery of controls that collectively match the low-level properties of the 2D displays, and that is the approach that was adopted in this study. As will be described later in the results section, this also makes it possible to employ regression analyses to measure the extent to which variations in BOLD (blood oxygen level-dependent) response between different pairs of conditions can be accounted for by low-level differences in image structure between those conditions.



**Figure 2.** Luminance histograms and amplitude spectra of the stimuli used in the main experiment. (A) Luminance histograms averaged over all 11 shapes (the error bars indicate standard deviations [SDs]) of 3D shaded, center-shaded, and shaded-blob stimuli. Yellow bars indicate the luminance of the uniform-luminance stimuli and the arrows indicate the light and dark gray values for each of the unshaded-blob shapes (see Materials and Methods). (B) Amplitude spectra averaged over the 11 shapes (the error bars indicate SDs). Upper panel: spectra of 3D shaded (black line), center-shaded (orange line), shaded-blob (dark red), and unshaded-blob (light red) shapes; lower panel: spectra of 3D shaded (black line), uniform luminance (yellow line), and pixel scrambled (olive line). The stars indicate frequencies at which spectra differed significantly (1-way ANOVA,  $P < 0.05$ ) between 3D shaded stimuli and the 2D controls (orange: center shaded; olive: pixel scrambled). Notice that the amplitude spectra were not calculated on the interior of the shapes but on the central  $15.4^\circ \times 14.5^\circ$  part of the display.

Two quantitative analyses were performed on the stimuli from the SfS experiment in order to compare their low-level properties. First, we compared the amplitude spectra of the 2D conditions with those of the 3D shaded condition. The amplitude spectra were calculated for each complete stimulus image, including the outline of the surface and the central part ( $15.4^\circ \times 15.4^\circ$ ) of the background, using a 2D discrete fast Fourier transform (MATLAB). The 2D output (amplitude as a function of spatial frequency and orientation) was reduced to 1 dimension by collapsing across orientations. Spatial frequency was computed in terms of cycles per image, which was then converted to cycles per degree (cycles/deg) by assuming a viewing distance of 36 cm. In general, the spectra of the 2D conditions differed slightly from that of



**Figure 3.** Average luminance distribution across images in the 6 conditions of the SFS experiment. (A) The luminance distributions in the central  $14.5^\circ \times 14.5^\circ$  part of

the 3D condition, but the difference reached significance only for the pixel-scrambled and center-shaded conditions (Fig. 2B). Differences were most significant for the pixel-scrambled condition at high spatial frequencies, which is a typical feature of scrambling.

A second analysis examined the mean luminance distribution across the images in the different conditions (Fig. 3). The 3D shading condition had a clear upper-lower asymmetry introduced by the position of the light source. This was much less the case for the 2 blob conditions. The center-shaded condition had luminance distribution that was much lower than that of the 3D shaded condition in the central part of the image. The luminance distributions of the last 2 conditions were also more symmetrical, as the center-shaded condition, but luminance levels were more similar to those of the 3D shaded conditions in the center of the images. These 2 analyses numerically show that it is impossible to create a single condition that is perceived as flat and shares all the low-level features of the images in the 3D condition. Even the shaded-blob condition does not meet all the requirements. This underscores the strength of our approach combining several control conditions.

*Testing schedules.* In addition to the conditions described above, all of the experiments included a fixation-only condition to provide a baseline level of activation. All the experiments used block designs with an epoch duration of 30 s, corresponding to 10 functional volumes or scans. Within an epoch, the 11 stimuli were presented twice for 1400 ms with no gaps. Each of the conditions ( $n = 7$ ) was presented in a different epoch and was repeated once in a time series yielding 140 functional volumes per time series. Eight time series were recorded in each subject and presentation order of the conditions was randomized between these time series.

The initial 12 subjects were tested in 3 sessions of 8 or 6 time series, 1 devoted to each cue and a third session for the motion and 2D shape localizers. The last 6 subjects were tested in 8 scanning sessions consisting each of 7 time series: 2 time series for the main experiment, 1 for each cue - shading and texture, 2 time series (1 for shading and 1 for texture cue) in a 1-back matching task (see below), 2 in a passive viewing using the same control stimuli as in the 1-back task, and 1 localizer. The order of the experiments was changed between scan sessions and between subjects.

#### Control Experiments

*Control experiment 1: specular surfaces ( $n = 4$ , shading cue).* In the main experiment, the shading patterns were almost entirely diffuse (i.e., Lambertian) (Fig. S1). A first control experiment, in which 4 of the subjects participated, aimed at comparing the neural substrates for 3D SFS for Lambertian and specular reflectances. Therefore, in the 3D shaded condition, surface reflectance included a 30% specular component (mean luminance  $309 \text{ cd/m}^2$ ) (Fig. 12D). The control stimuli for this experiment were built following the same procedure described in the main shading experiment, but matched to 3D shaded condition with specular surfaces. In these 4 subjects, 8 time series were tested in a single scan session, in order to be compared with the 8 series obtained in the main experiment. Subjects were required only to fixate the fixation target. Subjects made on average 4 or 5 saccades per block in the specular and Lambertian shading runs, respectively. The number of saccades was not significantly different amongst conditions (both 1-way ANOVAs  $P > 0.8$ ).

*Control experiments 2: 1-back task ( $n = 6$ , shading and texture cues).* This task was designed to control for variations in attention among the different conditions that could conceivably have played a role in the results of the main study. In order to compel their attention, observers were required to judge whether each successive

the display, averaged across the 11 stimuli used in each condition. (B) luminance plotted as a function of vertical position to highlight the upper-lower asymmetries in the stimuli. The 4 lines (color code see inset) correspond to  $2.5^\circ$  wide vertical strips, centered on the fixation point, over which luminance was integrated.

pair of images was the same or different. To prevent these judgments from being based on the 2D bounding contours of the objects, the stimuli from the main experiment in the *3D shaded*, *shaded-blob*, *unshaded-blob*, *uniform-luminance*, *3D lattice*, *uniform-texture*, and *lattice-aligned* conditions were modified so that they all had identical 6.5° circular boundaries (see Figs 14 and 15, insets). All 11 images of a condition were first resized to subtend approximately 10° of visual angle. Then four 6.5° circular cutouts within the boundary of each object were created that were positioned 0.5° from the center along each diagonal. Thus, with 4 distinct circular patches for each of the original images, there were 44 stimuli per condition. The aim of increasing the stimulus set size was to increase the difficulty of the task and to avoid rote learning. In the uniform-texture condition task difficulty was increased by rotating the images either 20° or -20° in the image plane. The uniform-luminance condition was used in both shading and texture versions of this control experiment as a low-level condition, as well as a fixation-only condition as a baseline. The uniform-luminance condition consisted of 22 different gray-level stimuli with a luminance range from 0 (black) to 255 (white) and an average luminance equal to that of the 3D shaded condition (175 cd/m<sup>2</sup>). The complete set of stimuli was divided over 2 runs to achieve equal probability (50%) of same and different stimulus sequences. Two or more consecutive repetitions occurred in each epoch. Nevertheless the observers never saw the same stimuli (except in the uniform-luminance condition) in these runs.

Each image was presented for 800 ms with a blank interval of 600 ms between images (note that 1400 ms is the presentation time of each of the original stimuli). Subjects were required to respond within these blank intervals by interrupting an infrared beam with minimal movements of their left thumb for “same” responses and right thumb otherwise. Left-right responses were interchanged between subjects. Before scanning each of the participants underwent a training session outside the scanner to ensure their ability to perform the task and to compose individual sequence according to their correct responses.

The 5 conditions of the shading (Fig. 15, inset) or texture (Fig. 14, inset) control experiments were presented in different epochs (of 10 functional volumes) and were replicated twice, yielding 150 functional images per time series. Eight different time series were collected for the 2 cues (shading and texture) and the 2 tasks (passive-viewing/1-back matching task). This collection was interleaved with that of the main experiments and localizers, as described above.

Among the 6 subjects, the percent correct responses were very similar between the conditions. Average values were 87.5% (standard deviation [SD] 1.4%) and 85.5% (SD 1.1%) for shading and texture experiments, respectively, with no statistically significant differences among the different conditions (1-way ANOVA,  $P > 0.5$  and  $P > 0.7$  for shading and texture). Subjects made on average 6 or 7 eye movements per block in the passive and active versions of the shading control experiments. They made on average 6 and 5 eye movements per block in the passive and active versions of the texture control experiment, respectively. The numbers of saccades were not significantly different between conditions (all 1-way ANOVAs  $P > 0.5$ ) in any of these experiments.

*Control experiments 3: high-acuity task (n = 3, shading and texture cues).* As an additional control for attention, 3 subjects were scanned while performing a high-acuity task (Vanduffel et al. 2001) during presentation of the stimuli of the 2 main experiments. They were required to interrupt a light beam with their right thumb when a small red bar, presented in the middle of the screen changed orientation from horizontal to vertical (Sawamura et al. 2005). Psychophysics tests performed in the scanner indicated that performance level (% correct detection) decreased and reaction time increased slightly with decreasing bar size, suggesting that they are sensitive indicators of the subjects’ attentional state (Sawamura et al. 2005). For each of the individual subjects we selected the bar size (typically 0.2° x 0.05°) that corresponded to 80% correct detection. Again, 8 time series were tested in a single scan session for each subject, to be compared with the 8 series obtained in the main experiment.

In the shading experiment, the mean (across subjects) percent correct response in the high-acuity task was 86.4% (SD 1.9%). The

performance level was lower in the scrambled condition (79% correct) compared with the other conditions (85–90%). As a consequence, performance differed amongst conditions (1-way ANOVA,  $P < 0.05$ ). The lower performance in the scrambled condition can be explained by the difficulty of seeing the bar because of the fine structure in the stimuli. The mean reaction time equaled 404 ms (SD 24 ms), with no statistically significant difference amongst conditions (1-way ANOVA,  $P > 0.05$ ).

In the texture experiment, the mean performance and reaction time were 79.4% (SD 7.5%) and 391 ms (SD 34 ms), respectively. These values were not significantly different amongst conditions (1-way ANOVA  $P > 0.40$  and  $P > 0.90$ , respectively).

We also examined how well subjects fixated during the high-acuity task experiments. They made few saccades per block: 4 saccades on average in the shading test and 5 saccades the texture test. These numbers did not significantly (1-way ANOVA) differ amongst conditions of the shading ( $P > 0.60$ ) or the texture ( $P > 0.60$ ) experiment.

#### *Localizer Tests*

Motion and 2D shape or LOC localizer tests were performed on all of the subjects. For the 2D shape localizer scans, we used grayscale images and line drawings (12 × 12 visual degrees) of familiar and unfamiliar objects as well as scrambled versions of each set (Kourtzi and Kanwisher 2000; Denys et al. 2004). Motion localizer scans contrasted a moving with a static random texture pattern (7° diameter, Sunaert et al. 1999). For the 2D shape localizer 4 time series were tested, and 2 for the motion localizer.

In 4 of the subjects, retinotopic mapping was performed using 4 types of patterns, black and white checkerboards, colored checkerboards, moving random dots and moving random lines, to stimulate the horizontal and vertical visual field meridian, the upper and lower visual field and the central and peripheral representations of visual cortex (Fize et al. 2003; Claeys et al. 2004). We mapped the borders of areas V1, V2, V3, V3A dorsally and V1, V2, V3, and V4 ventrally. In 4 subjects the kinetic occipital (KO) region was localized by comparing kinetic gratings to transparent motion (Van Oostende et al. 1997).

#### *Psychophysical Experiments*

*Experiment 1: depth magnitude estimation task.* Observers’ judgments of the relative magnitudes of relief were measured outside the scanner by presenting all of the stimuli to the volunteers and asking them to estimate the apparent depth of each object on a scale of 1–10, such that higher numbers indicated greater amounts of depth. This evaluation was performed for all of the stimuli of the main shading and texture experiments. We did not find statistically significant differences between the 11 stimuli (1-way ANOVA, with a factor shape) of any of the texture conditions ( $P > 0.86$ ). The same analysis in the shading experiment revealed no differences between the stimuli ( $P > 0.8$ ), except in the shaded-blob condition ( $P > 0.013$ ).

*Experiments 2–4: 3D shape adjustment task.* Observers performed a shape adjustment task adapted from previous psychophysical investigations by Koenderink et al. (2001) and Todd et al. (2004). On each trial a single image of a 3D shape or a control stimulus, 10° in size, was presented together with 8 equally spaced dots along a horizontal scan line through the center of the depicted surface. An identical row of dots was also presented on a separate part of the same display screen (below the shape), each of which could be moved in a vertical direction with a hand held mouse. Observers were instructed to adjust the dots of the second row in order to match the apparent surface profile in depth along the designated scan line. Once they were satisfied with their settings, observers initiated a new trial by pressing the enter key on the computer keyboard. The shapes were presented in alternating blocks with either *Lambertian* shading or a *lattice* volumetric texture. Each shape was presented once in a random order in a block. All participants wore an eye patch over the right eye and their head movements were restricted using a chin rest. This task was used to compare the 2 cues, but also to evaluate the perceptual differences between 3D and 2D control conditions. In the main assessment of the 2 cues (experiment 3), observers were presented

with the same 11 3D shapes employed in the imaging studies, except that their sizes were adjusted so that they all subtended  $10^\circ$  of visual angle. Three blocks were performed for each cue in an alternating sequence, so that we obtained shape judgments for 66 scan lines per subject, 33 lines per cue. To compare the relative sensitivity of the 2 cues (experiment 4), these measurements were repeated but with stimuli in which the depth magnitude was set at either 100%, 64%, or 28% of the original value in the main experiments (Fig. S3). Adjustments were made along either a horizontal or a vertical line. In total we obtained shape judgments for 132 scan lines per subject, 22 lines per condition. Finally this task was also used to compare the depth magnitude perceived in the 2D control conditions (experiment 2). For each cue 5 conditions were tested, leaving out the uniform conditions. Adjustments were made only along a horizontal line. In total, shape judgments were obtained for 220 scan lines per subject, 22 lines per condition.

**Experiment 5: successive discrimination of 3D shapes.** In this experiment subjects compared 2 successively presented 3D shapes, one which was used in the main experiment and another with either the same 3D shape or a 3D shape obtained by displacing each surface point in depth as a sinusoidal function of its vertical position (Fig. S3). By using 5 different amplitudes of this sinusoidal shape perturbation we were able to compare observer sensitivity to shape changes for SFs and SFT. Shapes were presented for 800 ms, interstimulus interval was 600 ms, intertrial interval 1400 ms. The subjects indicated in the 2200 ms following the start of the second stimulus, whether or not the 2 stimuli were the same by pressing 1 of 2 response buttons (S and M on keyboard). The buttons were assigned same or different randomly across subjects. To prevent subjects from basing their judgments on 2D stimulus properties, the original shapes and the 5 deformed versions of each one were presented in either of 3 orientations in depth, obtained by turning the object in depth around the vertical axis by  $-5^\circ$ ,  $0^\circ$ , or  $+5^\circ$ .

#### MRI Data Collection

Data were collected with a 3T MR scanner (Intera, Philips Medical Systems, Best, The Netherlands). The functional images consist of gradient-echo-planar whole-brain images with 50 horizontal slices (2.5-mm slice thickness; 0.25 mm gap), acquired every 3.0 s (time repetition [TR]), echo time (TE) 30 ms, flip angle  $90^\circ$ ,  $80 \times 80$  acquisition matrix ( $2.5 \times 2.5$  mm in plane resolution), with a SENSE reduction factor of 2. A 3D high-resolution  $T_1$ -weighted image covering the entire brain was used for anatomical reference (TE/TR 4.6/9.7 ms; inversion time 900 ms, slice thickness 1.2 mm;  $256 \times 256$  matrix; 182 coronal slices; SENSE reduction factor 2.5). The scanning sessions lasted between 90 and 120 min, including shimming, anatomical and functional imaging.

In total 99 904 functional volumes were sampled: 40 320 volumes in 18 subjects for the main shading and texture experiments, 4480 volumes in 4 subjects for the control experiment 1 (specular surfaces), 28 800 volumes in 6 subjects for the control experiment 2 (1-back task), 3360 volumes in 3 subjects for the control experiment 3 (high-acuity task), 10 944 and 4320 volumes in 18 subjects, respectively, for the 2D shape and motion localizers, 7680 volumes in 4 subjects for retinotopic mapping and 4608 volumes in 4 subjects for KO localizers.

#### Data Analysis

##### Preprocessing

fMRI data were analyzed with SPM software (Wellcome Department of Cognitive Neurology, London, UK). The preprocessing (SPM2) steps included realignment, coregistration of the anatomical images to the functional scans, and spatial normalization into the standard Montreal Neurological Institute (MNI) space. The data were sub sampled in the normalization step to  $2 \times 2 \times 2$  mm for the random effect analysis (main experiments and localizers) analysis and for the single-subject analysis and to  $3 \times 3 \times 3$  mm for fixed effect analysis (control experiments). They were spatially smoothed with an isotropic Gaussian kernel (group random effect and single-subject analysis: 8 mm; group fixed effect analysis: 10 mm) before statistical analysis.

##### Main Experiments

We used SPM2 for the statistical analysis in both single-subject and random effect analysis of the main experiments. To obtain the regions activated by the 3D conditions compared with 2D conditions in the main experiments, we performed a random effects analysis that yields the voxels that are activated in each of the subjects (taking into account the variability between sessions and subjects), rather than the fixed effects model which gives the areas that were activated on average across the subjects (Penny and Holmes 2003). The group data of the shading or texture main experiments were analyzed as follows: first, for each subject all contrast images between the 3D conditions and each of the 2D control conditions were calculated (5 contrasts in the shading and 4 in the texture experiment). Second, the contrast images from all 8 subjects were grouped by subtraction and entered into a 1-way ANOVA across contrasts (Henson and Penny 2003). Finally, a second level conjunction analysis (Friston et al. 1999) was applied to find activation common to these subtractions. The same analysis but at the first level (omit step 2) were applied within all individual subjects. The level of significance was set at  $P < 0.001$  uncorrected for multiple comparisons for both single subjects and the main random effect group analysis (conjunction null analysis, Friston et al. 2005). We determined the number of single subjects in whom a given region of activation revealed by the random effect analysis was activated. We consider that an activation site revealed by the random effect analysis is present in a single subject if the difference between the MNI coordinates of the group and subject activation sites does not exceed 3 voxels (i.e., 7.5 mm) in each direction for any coordinate.

The overlap between shading- and texture-sensitive areas was revealed by inclusively masking the conjunction of the texture related subtractions with those of shading. An additional analysis was performed on the results of the main shading experiment. In an attempt to find regions for which the activation correlated with the strength of the 3D shape perception, we derived linear contrasts involving the four 2D shading conditions evoking intermediate degrees of 3D perception of the stimuli (center-shaded, shaded-blob, unshaded-blob, and pixel-scrambled conditions). For each individual subject the average judged depths of these 4 conditions were used as a weight for the different conditions in the contrast equation. Setting the sum of the differences equal to zero yielded a contrast for each subject. For example, in subject RD the average judged depths for these conditions was 4.6, 4.8, 2, and 3 for a grand mean of 3.6, thus the contrast entered in SPM2 was: 0, 1, 1.2, -1.6, -0.6, 0, 0, respectively, for 3D shaded, center-shaded, shaded-blob, unshaded-blob, pixel-scrambled, uniform-luminance, and fixation conditions. A random effect analysis was applied with the resulting contrast images defined by such linear contrasts and the result ( $P < 0.001$  uncorrected) was masked inclusively with the conjunction of the 5 main subtractions (3D minus each of the 2D condition) ( $P < 0.001$  uncorrected) to restrict the correlation to voxels processing 3D shape information.

The standard contrasts of the motion (moving vs. stationary) and 2D shape or LOC (intact vs. scrambled objects images) localizer tests were also subjected to a random effects analysis. For the 2D shape and motion localizers the threshold was set at a  $P < 0.0001$  uncorrected for the analysis, corresponding to  $T$ -score  $> 4.71$  (Fig. 6).

##### Control Experiments

All fixed effect analyses were performed using SPM 99. This is because of the limitations of the amount of data we could use in SPM2. The thresholds of the fixed effect analysis were taken at  $P < 0.05$  corrected for multiple comparisons ( $T$ -score  $> 4.99$ ) except for the high-acuity experiment wherein it was  $P < 0.005$  uncorrected. The analysis for the 1-back control experiments ( $n = 6$ ) follows the logic described above for the main experiments. Three contrasts (3D vs. each of the 2D control conditions) were entered in a conjunction analysis for each cue (shading, texture) and task (passive/1-back task) in order to reveal the common cortical activation pattern. The local maxima from the task experiments ( $P < 0.05$  corrected) were used to plot the percent MR responses in both passive and task experiments. We grouped the data from the passive and task experiments (separately for the 2 cues) in 1 statistical test to explore the interaction (uniform luminance - fixation) task - (uniform luminance - fixation) passive.

When regions, activated by 3D conditions relative to 2D conditions, were compared in 2 different datasets, that is, when comparing passive and high-acuity task ( $n = 3$ ), or specular and Lambertian shading ( $n = 4$ ), care was taken to compare MR activation in regions (centered on local maxima) defined according to the activation pattern common to the 2 sets compared. This was obtained by integrating the 2 datasets of all subjects into a single statistical analysis and taking a conjunction between subtractions to define the local maxima.

#### Flatmaps

The fMRI data ( $T$ -score maps) were mapped onto the human Population averaged, landmark and surface based (PALS) atlas surface in SPM-MNI space as well as onto the individual flattened cortical reconstruction of the single cortices of the 4 subjects in which we performed the retinotopic mapping using a volume-to-surface tool in Caret (Van Essen 2005). Caret and the PALS atlas are available at <http://brainmap.wustl.edu/caret> and [http://sumsdb.wustl.edu:8081/sums/archivist.do?archive\\_id=6356424](http://sumsdb.wustl.edu:8081/sums/archivist.do?archive_id=6356424) for the left hemisphere and [http://sumsdb.wustl.edu:8081/sums/archivist.do?archive\\_id=6358655](http://sumsdb.wustl.edu:8081/sums/archivist.do?archive_id=6358655) for the right hemisphere.

#### Activity Profiles

**Standard (local) activity profiles.** The raw MRI data were converted to percent signal change plotting the response profiles relative to fixation condition. In the standard procedures, activity profiles were calculated for a small region surrounding the local maximum in the SPM by averaging the data from the most significant voxel and the 4 closest neighbors. For data analyzed with random effects analysis, profiles from the single-subjects analysis were averaged and SEs were calculated. For data analyzed with fixed effect analysis (SPM 99), activity profiles were obtained directly from the group analysis.

**LOC profiles.** In 1 analysis we contrasted intact versus scrambled object images (LOC localizer) for defining a region of interest (ROI). Voxels belonging to LOC were defined individually per subject using different high thresholds  $P < 10^{-6}$  ( $T$ -score between 11 and 35), adjusted to yield approximately 150 voxels in at least 1 hemisphere. The average number of voxels in the 2 hemispheres included per subject was  $133 \text{ voxels} \pm 67$  (SD between 18 subjects), with local maxima  $-46 \pm 3$ ,  $-78 \pm 4$ ,  $-9 \pm 5$  in the left hemisphere and  $43 \pm 5$ ,  $-76 \pm 4$ ,  $-10 \pm 4$  on the right site (Fig. S7). The fMRI response was extracted by averaging the data from all the voxels within the ROIs of the 2 hemispheres. The magnitude of the responses in these ROIs was measured for each subject in each condition and then averaged between subjects (Table S5). This method of defining ROIs was intended to emulate the procedure followed in most imaging studies, in which only limited localizer data are acquired.

To characterize LOC more fully, using the strength of our extensive data collection we used an average of the different local activity profiles. These profiles were taken at the local maxima of the LOC localizer subtraction that exceed the standard  $P < 0.0001$  uncorrected threshold and applying these coordinates in single subjects. This was the case in Figure 11 and Table S4.

**V1 profiles.** The coordinates for this early visual area were defined on the available retinotopic maps of 4 of the subjects probing in the left and right hemispheres the ventral and dorsal borders of the central stimulus representation (2.5 visual degrees size) between the horizontal and vertical meridian representations of V1. The coordinates of these dorsal and ventral probes were averaged across the 4 subjects and applied within all individual subjects in the shading and texture experiments, respectively. They are shown as white dots on the flatmaps in Figure 6. The activity profiles for these local probes were averaged across subjects and hemispheres (Fig. 11).

## Results

### Behavioral Results

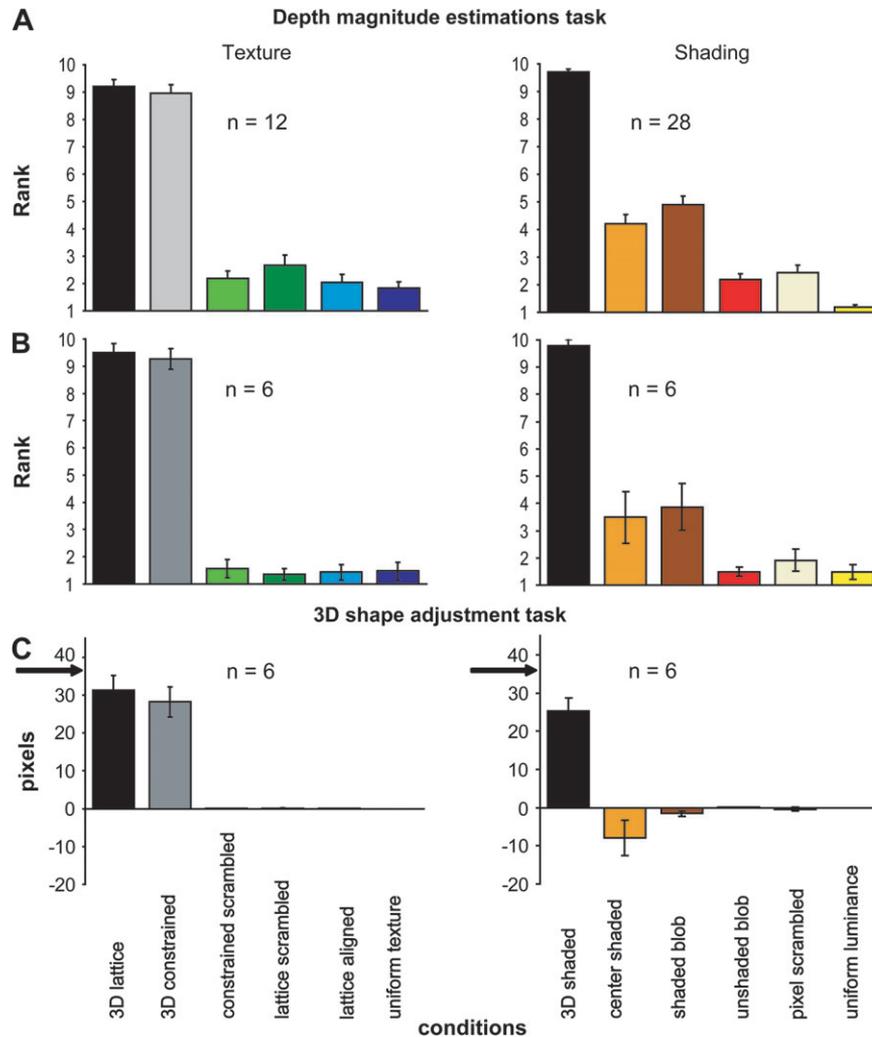
Twenty-eight volunteers (including all 18 subjects from the main fMRI experiment) performed the *depth magnitude estimation task* (psychophysical experiment 1) with the

shading stimuli and 12 of them performed the same task with the texture stimuli. Subjects estimated the apparent depth of each object on a 1–10 scale, such that higher numbers indicated greater amounts of depth. The average judgments for each of the different conditions are presented in Figure 4A, and these findings confirm that there were large differences in apparent depth between the 3D shapes and the 2D controls for the texture stimuli. For the shading stimuli 2 control conditions, center-shaded and shaded blob evoked more depth than the 3 other control conditions (unshaded blob, uniform luminance, and pixel scrambled). This magnitude estimation procedure was validated by comparing it in 6 subjects with the 3D shape adjustment task. To that end the 8 settings obtained for each shape in the 3D shape adjustment task were averaged. In general, the 2 tests were in excellent agreement (Fig. 4B,C): correlation was 0.99 for texture conditions and 0.86 for shading conditions. The lower correlation in the shading conditions was due to the shaded-blob condition, which produced some apparent depth in the magnitude estimation task, but relatively little in the adjustment task.

Subjects maintained fixation well during all fMRI measurements. They made few saccades during scanning: 5 per block on average in both the shading and texture main experiment. Importantly, the number of saccades did not differ significantly between the conditions (1-way ANOVA) in either the shading ( $P = 0.47$ ) or texture experiment ( $P = 0.46$ ). The same was true for the various additional experiments (see Materials and Methods). These control measurements indicate that eye movements did not influence the functional results.

### Main Experiment: 3D SFT, Group Analysis

In order to identify the human brain regions involved in the extraction of 3D SFT we contrasted the 3D shape conditions to each of their 2D controls and at a second level performed a conjunction null analysis of these 4 contrasts. The voxels reaching significance ( $P < 0.001$  uncorrected) are shown in Figure 5A on the rendered brain, with the local maxima listed in Table 1. These voxels are also plotted on the flatmaps in Figure 6 to indicate the relationships with motion and 2D shape sensitivity, listed in Table 2. As in the extraction of 3D SfM (Orban et al. 1999) activation extended, in both hemispheres, from occipito-temporal cortex, onto the lateral occipital cortex, into the occipital part of the IPS, finally reaching the dorsal part of that sulcus. The occipito-temporal site (caudal inferior temporal gyrus [caudal ITG], 1, 2 in Figs 5A and 6) is located ventral of human homologue of the middle temporal (MT)/V5 complex (hMT/V5+) and the lateral occipital activation (3, 4 in Figs 5A and 6) largely avoids, in its dorsal part, the motion-sensitive V3A area (Fig. 5A). This was confirmed in single subjects whose retinotopic regions were mapped (Fig. 7). Hence the lateral region is referred to as lateral occipital sulcus (LOS), following Denys et al. (2004). The IPS activation sites (5–10 in Figs 5A and 6) were located close to but were not always identical to those involved in 3D SfM extraction (Orban et al. 1999, 2003), as can be inferred from their motion sensitivity (Table 2) and their location. To remain on the conservative side we followed the same labeling policy as in Denys et al. (2004), and used the labels of Orban et al. (1999), but added a suffix (t) when the average location was too different from that of previous studies (Orban et al. 1999; Sunaert et al. 1999; Denys et al. 2004). As shown in Table 2, almost all parietal activation sites were also shape-sensitive (Denys et al. 2004). These included the motion-sensitive



**Figure 4.** Results from psychophysical experiments 1 and 2. Depth magnitude estimation task in complete population of subjects ( $n = 28$  in the shading and  $n = 12$  in the texture experiments, *A*) and the 6 subjects participating in experiment 2 (*B*), and 3D shape adjustment task (*C*). Apparent depth (*A*, *B*) and average setting of 8 points (*C*), averaged over stimuli and subjects, is plotted as function of conditions. Error bars indicate SEs across participants. In (*C*) arrows indicate ground truth.

ventral IPS (VIPS) region in the right hemisphere, which extended further forward into a less motion-sensitive part. In the left hemisphere only this latter activation (VIPSt) was observed. In both hemispheres the motion-sensitive dorsal IPS medial (DIPSM) region was activated (Table 1). In the right hemisphere sites close to the parieto-occipital IPS (POIPS) region and to the dorsal IPS anterior (DIPSA) region were observed but lacked motion sensitivity. Furthermore, the posterior site was located about 10 mm too rostral compared with POIPS in earlier studies and was labeled POIPSt, whereas the anterior one was referred to as shape-sensitive DIPSA (DIPSA, Denys et al. 2004).

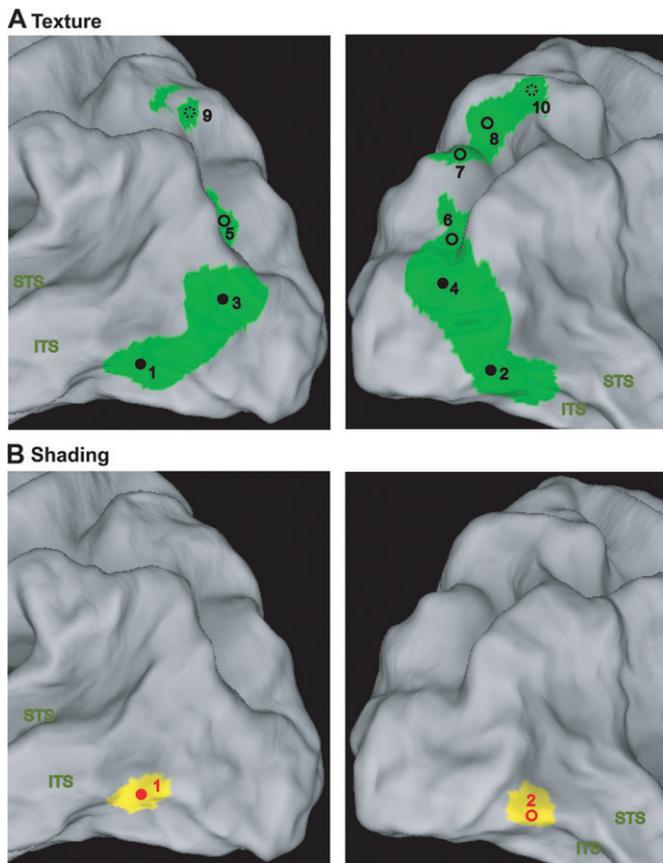
Figure 8A plots the activity profiles of the 3D SfT activation sites, obtained by averaging, across all 18 subjects, the profiles of the group local maximum. In general, MR activity is 30–50% greater in the 3D conditions than in the 2D control conditions. The significance of the individual contrast is listed in Table 1. As could be expected from a conjunction null analysis, these contrasts all reach  $P < 0.001$  uncorrected in all regions significant in the conjunction.

To understand the importance of the conjunction analysis involving several contrasts compared with a single contrast

with the most critical 2D condition, the activation patterns of the conjunction and its 4 constitutive contrasts are plotted on flatmaps in Figure 9. Somewhat counter intuitively the activation patterns for the 2 contrasts using the most crucial control conditions, the scrambled conditions, are fairly extensive (Fig. 9B,C). In particular contrasting the 3D lattice to the 2D scrambled lattice condition yields activation in fusiform regions which are not removed by the 2D lattice-aligned control (Fig. 9D), but are eliminated by the uniform-texture condition (Fig. 9E). This suggests that these regions are driven by patches of uniform texture, which are present in the 3D conditions (Fig. 1C). Indeed an earlier study showed that judgments about 2D texture involve these ventral regions (Peuskens et al. 2004). Notice also that this figure indicates that the activation pattern in the conjunction corresponds indeed to what one expects.

#### Main Experiment: 3D Sfs, Group Analysis

In order to identify the cortical region involved in extracting 3D Sfs we used a procedure similar to that for texture. We contrasted the 3D condition with each of the five 2D conditions



**Figure 5.** 3D SFS and 3D SFT sensitive regions rendered on the fiducial brain. Shading and texture specific areas shown on the posterior parts of left and right hemispheres (ventral-lateral view) of the fiducial PALS atlas (Van Essen 2005). Green (A) and yellow (B) patches represent voxels significantly ( $P < 0.001$  uncorrected, Conjunction null analysis, random effect,  $n = 18$ ) activated by 3D shape defined by texture and shading, respectively. The local maxima are indicated with black (texture) and red (shading) circles (filled circles: significance at  $P < 0.05$  corrected; open circles: significance  $P < 0.0001$  uncorrected), labeled with the numbers used in Table 1: (1) and (2) L and R caudal ITG, (3) and (4) L and R LOS, (5) L VIPSt, (6) R VIPs, (7) R POIPSt, (8) R DIPSM, (9) L DIPSM, and (10) R DIPSA.

and made a conjunction null analysis of these 5 contrasts at the second level. The resulting activation was much more reduced, almost to a tenth, than that obtained for the texture cue. In fact, only a single region, the caudal ITG region was activated in both hemispheres (Figs 5B and 6). This region is similar to the one involved in 3D SFT because the number of voxels activated by both cues amounted to about two-thirds (216/347) of the voxels activated in 3D Sfs.

The average activity profile of the caudal ITG region involved in 3D Sfs is shown in Figure 8B. The 3D condition activates both regions more than the 2D control conditions but the difference is smaller than in the texture case and amounts to only 20–30%. Activity in caudal ITG was equally strong in the center-shaded, shaded-blob and unshaded-blob conditions, indicating a strong 2D shape sensitivity, not surprisingly, given the overlap with the 2D shape-sensitive regions assessed with the LOC localizer stimuli (orange lines in Fig. 6). Yet all contrasts used in the conjunction reached  $P < 0.001$  uncorrected in the 2 caudal ITG regions (Table 1).

Again comparing the activation patterns for the conjunction and the 5 individual contrasts used in the conjunction (Fig. 10)

underscores the importance of using several control conditions. Indeed the contrast between 3D shaded and 2D shaded blob, which seem the most closely matched, in fact yielded an activation of caudal ITG and also of early visual areas, including V1 (Fig. 10C). This was also true for unshaded-blob condition, although V1 activation was reduced to the right hemisphere, in line with larger luminance differences between the conditions left of the fixation point (green line in Fig. 3). Early activation was more extensive in the contrast with central shaded control, in line with the larger luminance differences near the center of the image (Fig. 3). This control however constrained the anterior part of the caudal ITG activation, perhaps because of the large luminance gradient. The pixel-scrambled control on the other hand is the most effective in eliminating the early visual activation (Fig. 10E). This is not surprising as it is the most efficient stimulus of all 6 to drive V1 (Fig. 11) and yielded similar activation levels in dorsal and ventral V1. In fact the contrast 3D shaded minus 2D pixel scrambled was significant in V1 in the other direction (Table S4). In line with the dorsal ventral asymmetries in the luminance distributions of a number of shading conditions (Fig. 3), the activity profiles of dorsal and ventral V1 were different (Fig. 11). A 2-way ANOVA with conditions and parts as factors yielded significant effects of both factors (condition  $P < 10^{-6}$ , part  $P < 0.01$ ) and a significant interaction ( $P < 0.002$ ). In the texture experiment only the factor condition was significant ( $P < 10^{-6}$ ). Finally the uniform-luminance control used by Humphrey et al. (1997), was the least efficient control for early visual regions (Fig. 10F), even if the mean luminance distribution was very similar to that of the pixel-scrambled conditions (Fig. 3). This probably reflects the operation of surround mechanisms (Lamme 1995; Lee et al. 1998; Zenger-Landolt and Heeger 2003; Boyaci et al. 2007) rejecting stimuli with large uniform-luminance regions.

In the main experiment the 3D shaded objects contained no specularity. To control for possible effects of the different shading cues, we compared, in 4 subjects, the extraction of 3D Sfs in the Lambertian and the specular case (control experiment 1). Both shading cues yielded activation of only caudal ITG in the 2 hemispheres and ITG MR activity differed little between the 2 variants of shading (Fig. 12), in agreement with the psychophysical results (Nefs et al. 2006).

#### **Main Experiments: Additional Analyses of Group Data**

Additional analyses were performed in an effort to confirm that the differences in activation between the 2D and 3D conditions were due to the apparent depths of the surfaces as opposed to other low-level differences among the stimulus images in the various experimental and control conditions. First, from the observers' judgments in the depth magnitude estimation task, we calculated the average difference in apparent depth between each of the 15 pairwise combinations of conditions in both the Sfs and Sft experiments. Next we calculated the difference in BOLD response for each pairwise combination of conditions in both studies for all of the areas designated in Table 1 that had significant contrasts between the 2D and 3D displays. For each of these areas, a regression analysis was then performed to see how closely the differences in the BOLD response conformed with differences in apparent depth. The results, shown in Table 3, reveal that the BOLD response in these areas was highly correlated with the magnitude of perceived depth, that is, the average correlation was 0.93 for Sft regions and 0.62 for Sfs regions. Removing the shaded-blob

**Table 1**

Cue	Region	Coordinates			T-score					N of sub			
		x	y	z	Conj	Contrast				Text	Shad		
						CS	LS	LA	UT				
3D SFT	1	L caudal ITG	-50	-76	-12	<b>6.28</b>	<b>8.62</b>	<b>6.70</b>	<b>8.80</b>	<b>6.28</b>	16	16	
			-44	-88	-4	<b>5.54</b>	<b>8.61</b>	<b>6.46</b>	<b>7.88</b>	<b>5.54</b>			
	2	R caudal ITG	42	-82	-10	<b>5.66</b>	<b>9.15</b>	<b>5.73</b>	<b>8.29</b>	<b>5.66</b>	15	17	
			44	-66	-8	4.26	<b>5.70</b>	4.26	<b>6.34</b>	4.89			
	3	L LOS	-40	-90	12	<b>7.51</b>	<b>8.98</b>	<b>7.51</b>	<b>8.05</b>	<b>7.64</b>	17	6	
			-34	-86	4	<b>7.20</b>	<b>9.05</b>	<b>7.20</b>	<b>8.36</b>	<b>8.04</b>			
	4	R LOS	36	-90	14	<b>6.57</b>	<b>7.26</b>	<b>6.57</b>	<b>7.11</b>	<b>6.91</b>	15	5	
			28	-88	10	<b>5.75</b>	<b>7.25</b>	<b>5.75</b>	<b>6.94</b>	<b>6.22</b>			
	5	L VIPSt	-24	-72	30	4.11	4.43	4.21	<b>5.69</b>	4.11	9	4	
			28	-80	20	4.99	<b>5.20</b>	4.99	<b>7.01</b>	5.13			
6	R VIPS	28	-80	20	4.99	<b>5.20</b>	4.99	<b>7.01</b>	5.13	11	2		
		24	-72	32	4.34	<b>5.90</b>	4.34	<b>5.56</b>	4.81				
7	R POIPSt	20	-76	48	4.21	4.50	4.21	4.33	4.40	8	—		
		22	-62	56	4.91	<b>6.32</b>	4.91	<b>5.53</b>	<b>5.69</b>				
8	R DIPSM	22	-62	56	4.91	<b>6.32</b>	4.91	<b>5.53</b>	<b>5.69</b>	10	—		
		26	-64	56	3.57	5.15	3.57	4.95	4.10				
9	L DIPSM	-26	-64	56	3.57	5.15	3.57	4.95	4.10	9	3		
		26	-54	64	3.36	3.37	3.36	3.74	3.39				
10	R DIPSAAs	26	-54	64	3.36	3.37	3.36	3.74	3.39	7	4		
3D SFS	1	L caudal ITG	-48	-76	-6	<b>6.09</b>	<b>6.09</b>	<b>6.73</b>	4.81	<b>11.2</b>	<b>12.84</b>	15	16
			46	-70	-8	4.79	4.79	<b>5.35</b>	4.81	<b>7.67</b>	<b>9.08</b>		
2	R caudal ITG	46	-70	-8	4.79	4.79	<b>5.35</b>	4.81	<b>7.67</b>	<b>9.08</b>	12	17	

Note: The values in the T-score columns indicate significance at: **P < 0.05 corrected (bold)**, *P < 0.0001 uncorrected (normal)* and *P < 0.001 uncorrected (italic)*. Conj: conjunction; N of sub: number of subjects reaching *P < 0.001 uncorrected (conjunction)*. Shad, shading; text, texture; LS, lattice scrambled, CS, constrained scrambled, LA, lattice aligned, UT, uniform texture, CS, central shaded, ShB, shaded blob, USHB, unshaded blob, PS, pixel scrambled, UL, uniform luminance.

**Table 2**

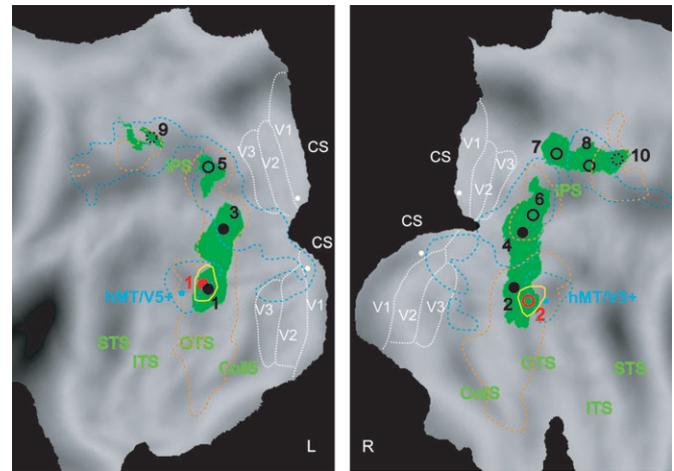
2D shape and motion significance in 3D SFT and 3D SFS sensitive regions, random effect (*n* = 18)

Cue	Region	Coordinates			2D shape, T-score	Motion, T-score	
		x	y	z			
3D SFT	1	L caudal ITG	-50	-76	-12	<b>9.75</b>	ns
			-44	-88	-4	<b>9.62</b>	ns
	2	R caudal ITG	42	-82	-10	<b>10.19</b>	ns
			44	-66	-8	<b>7.97</b>	3.97
	3	L LOS	-40	-90	12	<b>8.59</b>	4.19
			-34	-86	4	<b>8.69</b>	3.79
	4	R LOS	36	-90	14	5.54	7.37
			28	-88	10	6.75	7.46
	5	L VIPSt	-24	-72	30	3.95	ns
			28	-80	20	6.60	4.97
6	R VIPS	24	-72	32	3.67	ns	
		20	-76	48	ns	ns	
8	R DIPSM	22	-62	56	5.39	4.59	
		26	-64	56	5.39	6.89	
9	L DIPSM	-26	-64	56	5.39	6.89	
		26	-54	64	4.36	ns	
3D SFS	1	L caudal ITG	-48	-76	-6	<b>8.88</b>	4.90
			46	-70	-8	<b>8.96</b>	3.82

Note: The values in the T-score columns indicate significance at: **P < 0.05 corrected (bold)**, *P < 0.0001 uncorrected (normal)* and *P < 0.001 uncorrected (italic)*; ns: nonsignificant.

condition from the correlation had little effect: the average correlation was still 0.60 for the SFS regions.

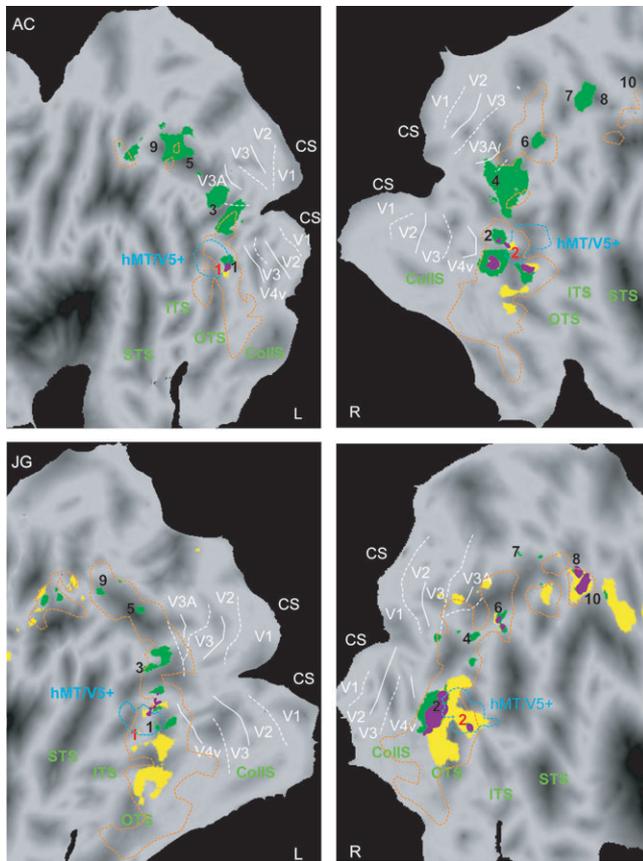
A similar analysis was also performed to assess the impact of low-level image differences on the pattern of activations in these areas. The luminance structure of each image was measured using an analysis designed by Lades et al. (1993) to approximate the representation of image structure in Area V1 of the visual cortex. The analysis, implemented in MATLAB, employs a bank of log Gabor filters, called a jet, that is, centered on each pixel. The filters in each jet include 6 different orientations with a separation and band width of 30°, 5 different scales with a separation and band width of 1.4 octaves, and 2 different phases (even- and odd-symmetric) in all possible combinations. The selection of 5 scales was constrained so that the wavelength of the smallest filter would



**Figure 6.** 3D Sfs and 3D Sft sensitive regions plotted on the flattened hemispheres (posterior part). Green patches and yellow outlines represent voxels significantly (*P* < 0.001 uncorrected, Conjunction null analysis, random effect, *n* = 18) activated by 3D shape defined by texture and shading, respectively. Same conventions as in Figure 5. The 2D shape-sensitive regions (i.e., voxels significant in subtraction intact versus scrambled images) from the LOC localizer (Denys et al. 2004; Kourtzi and Kanwisher 2000) are indicated by dashed orange outlines. Motion-sensitive areas are outlined with blue dashed lines and the blue stars indicate the motion-sensitive region hMT/V5+ (Sunaert et al. 1999). The threshold for LOC and motion localizers is set to *P* < 0.0001 uncorrected (*n* = 18, random effect). The early retinotopic areas V1, V2 and V3 (from Caret atlas) are indicated by dotted white lines. White dots indicate location of the V1 probes used in Figure 11 and Table S4. R, right; L, left; CS, calcarine sulcus; STS, superior temporal sulcus; ITS, inferior temporal sulcus; OTS, occipital temporal sulcus; CollS, collateral sulcus.

cover at least 3 pixels, and the wavelength of the largest filter would be no larger than the size of the image. The output of each filter at each image location was computed in the Fourier domain as described by Kovessi (1999).

It is important to recognize that the set of filter outputs for any given image can be thought of as a vector in a high dimensional space, where each individual Gabor filter defines



**Figure 7.** 3D SFS and 3D SFT sensitive regions (single-subject analysis, same conjunction analysis as in Figs 5 and 6). Statistical parametric maps, indicating voxels significantly ( $P < 0.001$  uncorrected) active when viewing 3D shapes defined by shading (yellow patches) and 3D objects defined by texture (green patches) shown on the flattened surface representation of the left and right hemispheres (posterior part) of 2 of the subjects (A.C. and J.G.). Purple voxels indicate overlap between shading- and texture-sensitive regions. Dashed orange and blue outlines correspond to LOC and motion localizers at threshold  $P < 0.05$  corrected. Solid and dashed white lines indicate the projection of horizontal and vertical meridians in a given hemisphere (Claeys et al. 2004; Fize et al. 2003). For other conventions see Figure 6.

a dimension, and the output of the filter defines a specific position along that dimension. The differences between a pair of images can therefore be measured as the Euclidean distance between their corresponding vector endpoints. Using this procedure, we measured the average difference in low-level image structure between each of the 15 pairwise combinations of conditions in both the Sfs and Sft experiments, and we correlated those with the differences in BOLD response for each combination of conditions for all of the areas that had significant contrasts between the 2D and 3D displays. The results of this analysis are presented in Table 3, and they reveal quite clearly that the differences in BOLD response among the different experimental conditions in the depth sensitive areas were largely independent of the differences in low-level image structure. Indeed, the average correlation was only 0.17 for Sft regions and negative for Sfs regions. When considered in combination, these regression analyses provide strong evidence that the conjunction of significant contrasts between the 2D and 3D displays was primarily due to systematic variations in apparent depth, and that low-level image differences had little or no impact on the overall pattern of activation.

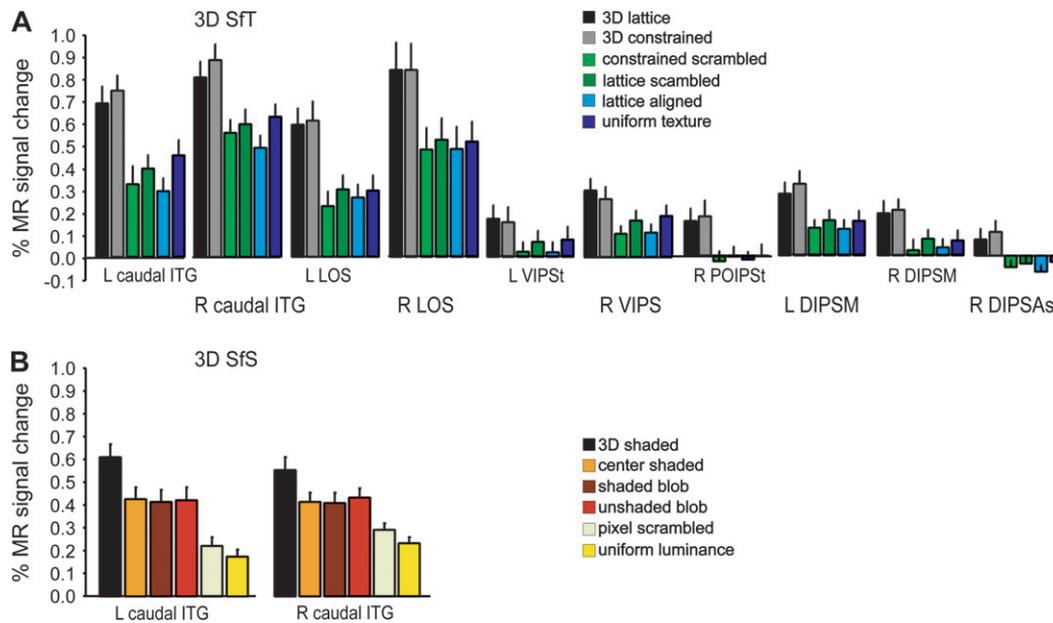
The previous analysis showed that activity in regions discovered by contrasting 3D and 2D conditions correlated with the variations of depth in the stimuli. We took advantage of the fact that subjects perceived some depth in 2D shading stimuli to look more directly for the voxels correlating most with perception in the 3D Sfs (Fig. S4). We correlated the activity of individual subjects with their psychophysical score in the 4 middle conditions: center shaded, shaded blob, unshaded blob, and pixel scrambled. The resulting local maximum was located in left caudal ITG (-50, -72, -12,  $T$ -score 3.78) with a nonsignificant trend in the symmetric region (40, -69, -10,  $T$ -score 2.72). The activity profiles indicate that activation in these voxels was strongest in the 3D condition, which was not included in the analysis, confirming indeed the relationship between MR activity in caudal ITG and 3D perception (Fig. S4). Even these voxels, however, were activated at intermediate if not strong levels by the 2D conditions containing 2D shape elements, confirming the profiles of Figure 8. Removing the shaded-blob condition from this correlation, still yielded the caudal ITG regions but with reduced significance: -54,-72,-10,  $T$ -score 2.65 and 52, -66, -10,  $T$ -score 2.06.

#### Main Experiments: Single-Subject Analysis

In the single-subject analysis we followed the same procedure of conjunction analyses as used in the group analysis. Figure 7 shows the voxels significantly activated by 3D SFT in green and in 3D Sfs in yellow for the 2 hemispheres of 2 subjects. Subject AC was typical in the sense that the 3D SFT activation was more extensive than that of 3D Sfs in the majority of subjects (13/18), whereas subject JG is representative of the 5 other subjects. On average across subjects 762 voxels were activated by 3D Sft, whereas 209 were activated by 3D Sfs. In all 4 hemispheres, the 3 main components of the 3D SFT activation were present: caudal ITG, LOS and IPS. In all 4 hemispheres 3D Sfs engaged caudal ITG, but in subject JG 3D Sfs also activated IPS regions. Also, in all 4 hemispheres the 3D Sft and 3D Sfs activations overlapped (purple voxels) in caudal ITG.

As indicated by the numbers in Figure 7, most activation sites in single subjects could be identified as 1 of the sites revealed by the group analysis (i.e., were located within 7.5 mm of the group local maximum). Table 1 lists, for each group activation site, the number of subjects in which that site was identified. Because the group analysis revealed an overlap between the 3D Sft and 3D Sfs activation in caudal ITG, it is not surprising that this region was activated in both main experiments in most single subjects, as illustrated in Figure 7. LOS was also activated by 3D Sft in the majority of subjects, but in only one-third of the subjects by 3D Sfs. Finally, the IPS regions were involved in 3D Sft in about half the subjects but in only a few subjects in 3D Sfs. Thus the single-subject analysis confirms that 3D Sft and 3D Sfs systematically overlap only in caudal ITG.

In the 6 subjects who had performed the depth adjustment task for the texture and shading stimuli (Fig. 4C), we correlated for the different regions obtained in the group analysis (Table 1), the differences in MR activity between the conditions in the individual subjects with the differences in their depth adjustment in those conditions. The result of this analysis, which was restricted to those regions in which we had observations in at least 4 subjects, is provided in Table S1. For



**Figure 8.** Activity profiles plotting averaged MR signal changes compared with fixation condition over all subjects ( $n = 18$ ). (A) 3D SFT sensitive regions: L and R caudal ITG, L and R LOS, L VIPst, R VIPS, R POIPst, L and R DIPSM, and R DIPsAs. (B) 3D Sfs sensitive regions: L and R caudal ITG. The error bars indicate SEs between subjects. Color bars indicate different conditions following the same convention as in Figure 1.

all regions, the median correlation exceeded 0.72, reaching frequently values close to 0.90.

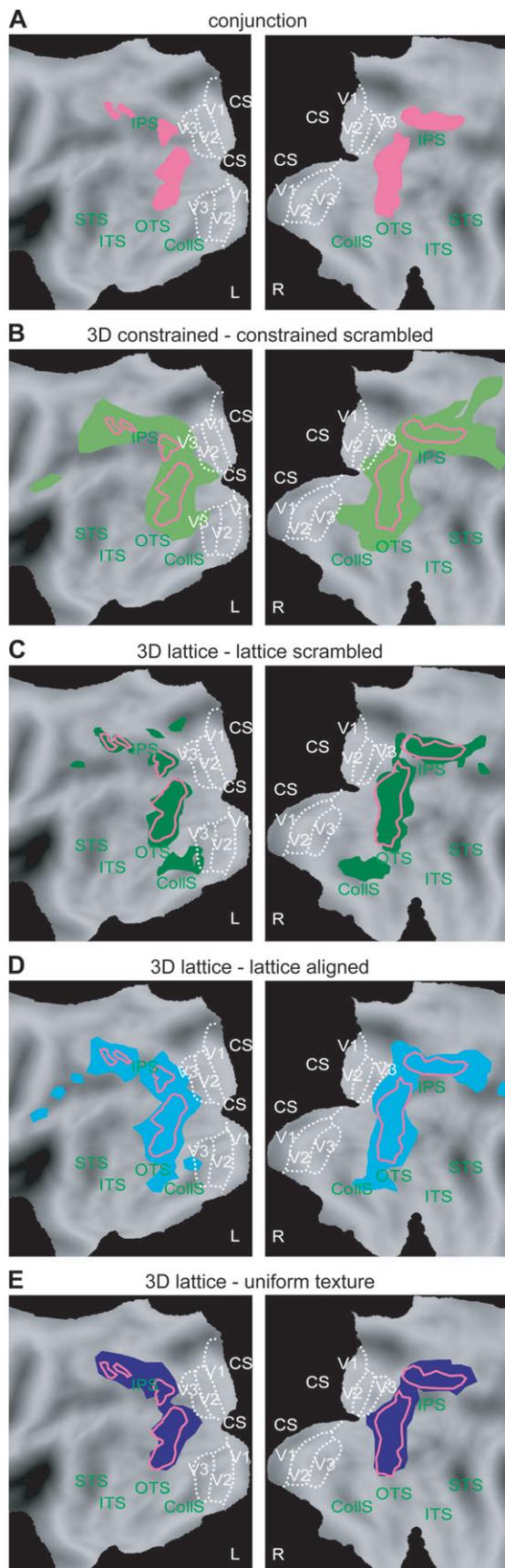
Figure 7 also illustrates the close association between the 3D SFT and 3D Sfs activation with 2D shape-sensitive regions. In particular caudal ITG is located in the posterior part of LOC, just ventral to hMT/V5+. In 4 subjects we had mapped the retinotopic regions and Figure 7 illustrates the main finding in these subjects: the 3D SFT and 3D Sfs activations were located largely anterior to the early retinotopic regions, although small invasions of anterior V3A were observed. The LOS activation by 3D Sft is located in between motion-sensitive regions V3A and hMT/V5+, the general vicinity in which the KO region is located (Van Oostende et al. 1997). It has been suggested that KO is specific for depth structure (Tyler et al. 2006), although controls for relative disparity were limited in that study. In 4 subjects, we had localized KO (Fig. S5). On average, KO is located somewhat more ventral, that is, closer to hMT/V5+ than the LOS 3D Sft activation, which is located more dorsally and anterior to V3A and might correspond to V3B (Press et al. 2001; Larsson and Heeger 2006). In 2 subjects the overlap between the kinetic and 3D SFT activations was extensive (66%), as in subject AC, whereas in the 2 other subjects it was minimal (23%), as in subject JG.

#### Main Experiments: Psychophysical Control Experiments

Because there were differences in the extent of cortical activations between 3D Sfs and 3D SFT, we conducted an additional psychophysical control experiment to ensure that this finding could not have been caused by differences in the relative effectiveness of these cues at eliciting the perceptual appearance of 3D shape. Nine of the subjects who had no parietal activation for extraction of 3D Sfs performed the depth adjustment task for both cues. A fixed effect analysis of their data show that indeed extracting 3D Sfs did not activate parietal cortex, not even at very low threshold ( $P < 0.1$  uncorrected), whereas extraction of 3F SFT did (Fig. S6). An

analysis of linear regression revealed that the average adjusted depth was highly correlated with the ground truth (on average  $r = 0.84$  for shading, and  $r = 0.86$  for texture), although the magnitude of surface relief was systematically underestimated by approximately 40% for both depth cues, which is consistent with prior studies (see Todd et al. 2004). The median slope of the correlation with ground truth equaled 0.55 (Q1 0.43 and Q3 0.77, first and third quartiles) and 0.62 (Q1 0.58, Q3 0.71) for shading and texture, respectively. There was a statistically nonsignificant trend ( $Z = 1.955$ ,  $P > 0.05$  Wilcoxon matched pairs test) for the texture adjustments to exceed those of shading, but plotting the texture adjustment directly as a function of the shading adjustments yielded a median slope of 0.91 (median  $r = 0.82$ ), indicating a trend in the opposite direction. Thus, the 3D shape adjustment task indicates that these cues were equally effective in the main experiment at eliciting the perceptual appearance of 3D shape.

Yet when examining the stimuli in Figure 1, many observers report anecdotally that the shaded objects are perceptually more compelling than those specified by texture. These impressions are borne out by testing the sensitivity for the 2 cues, as suggested by 1 of the anonymous reviewers. A first test was to use the 3D shape adjustment task for stimuli that were flattened compared with the original stimuli of the main experiments. Plotting the texture adjustments as a function of the shading adjustments yielded median slopes over 0.9 for the original stimuli and those flattened to 64% depth (Fig. 13A). However this slope decreased to 0.23 when the depth was further reduced to 28% of the original value. Similar results were obtained with the same-different task probing sensitivity for small perturbations in 3D shape. The sensitivity for 3D shape indexed by the  $d'$  obtained from signal detection analysis was significantly larger for the Sfs stimuli than the SFT stimuli (Fig. 13B). A 2-way ANOVA (repeated measures) with perturbation step and cue as factors yielded significant main effects of cue and steps (both  $P < 10^{-6}$ ) and significant



**Figure 9.** Comparison of activation patterns in conjunction and individual contrasts of the SFT experiment on flattened hemispheres. SPM plotting the voxels significant at the  $P < 0.001$  unc level (random effects) in the conjunction (A), the contrast 3D

interaction ( $P < 0.001$ ). The cue also had a small effect on the reaction times, which were marginally longer for the SFT stimuli than the SFS stimuli. Two-way ANOVA revealed a significant main effect of perturbation step ( $P < 10^{-6}$ ) and a significant interaction ( $P < 0.0002$ ). These findings are important for interpreting the results of the main experiment because they indicate quite clearly that the reduced pattern of activation for Sfs cannot be due to a lower sensitivity for shading information relative to the information provided by texture.

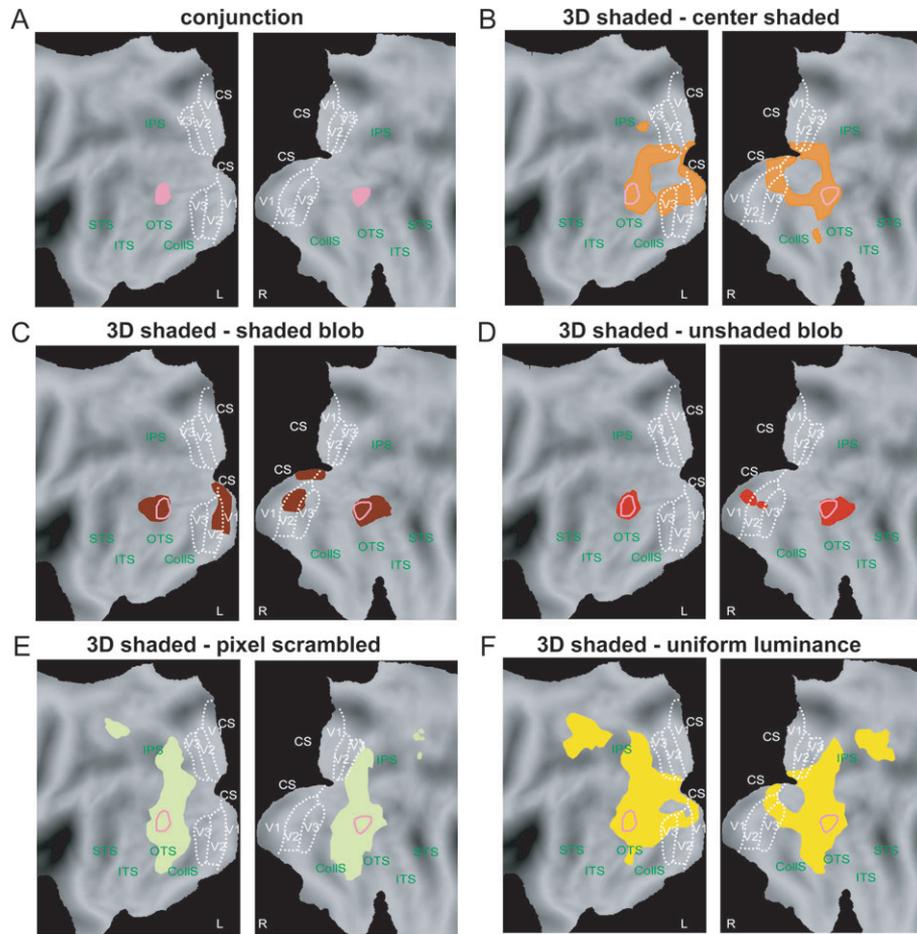
### Control Experiment 2: 1-Back Task

It could be argued that in the passive conditions subjects do not attend to the relevant characteristics of the stimuli and therefore we underestimate the extent of the 3D SFT and 3D Sfs activation patterns. For 3D Sfs in particular, this might explain the discrepancy with an earlier study (Taira et al. 2001), although it is unclear why this should apply to the shading and not to the texture cue, for which we obtained parietal activations in agreement with Shikata et al. (2001). In order to meet this criticism we tested 6 subjects while they were engaged in a 1-back task applied to the stimuli of the 3D condition and the 2D control conditions. To avoid the possibility that the subjects use the outlines of the stimuli we presented 88 different sub-regions from the stimulus objects in the main experiment that were all contained within a fixed circular aperture (see Materials and Methods). Also, in both experiments (texture and shading) we included a uniform-luminance condition as a low-level control. Subjects performed the task equally well in all conditions, reaching 84–87% correct in the different conditions of the texture 1-back task and 86–89% correct in those of the shading 1-back task (see Materials and Methods).

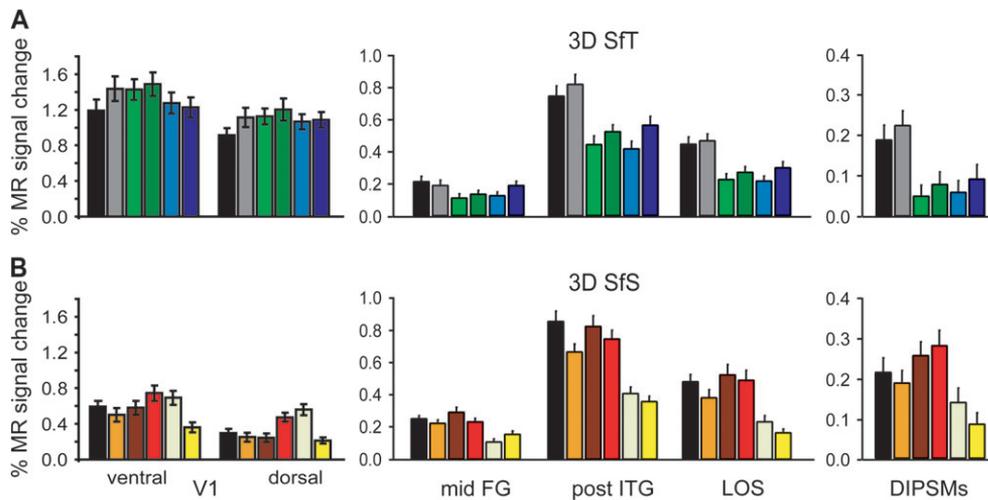
Contrasting the 3D texture condition with the three 2D controls and taking the conjunction yielded a large activation extending from occipito-temporal cortex to the anterior part of the IPS (green voxels in Fig. 14A). The local maxima are listed in Table S2 and include those of the main experiment plus a number of small ventral activation sites such as fusiform, collateral and lingual sites and parietal regions such as SPL (superior parietal lobule) and post central sulcus regions. Most of these regions remained significant in the passive condition (blue voxels and overlap light blue voxels in Figure 14A, black stars in Fig. 14B) with exception of the ventral fusiform, collateral and lingual sites and the post central sites. Those 7 sites losing significance exhibited small differences between the 3D and control conditions in the task condition, suggesting an activity profile with little selectivity. A number of the sites losing their significance were strongly activated by the uniform-luminance condition in the 1-back task, as indicated by a significant interaction between the factors task and stimulus (yellow asterisks in Fig. 14C).

Contrasting the 3D shading condition to each of the 2D controls and taking the conjunction yielded an activation (red voxels) centered on LOS and caudal ITG with small ventral sites

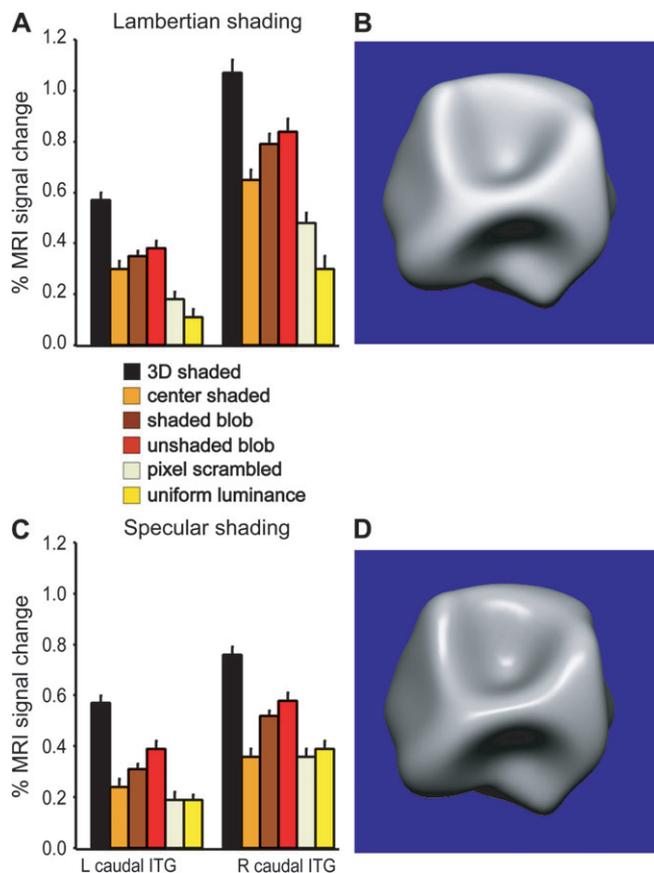
constrained minus 2D constrained-scrambled (B), 3D lattice minus 2D lattice scrambled (C), 3D lattice minus 2D lattice-aligned (D), and 3D lattice minus 2D uniform texture (E). In (B–E) the red outlines reproduce the activation in the conjunction (A).



**Figure 10.** Comparison of activation patterns in conjunction and individual contrasts of the SfS experiment on flattened hemispheres. SPM plotting the voxels significant at the  $P < 0.001$  unc level (random effects) in the conjunction (A), the contrast 3D shaded minus 2D center shaded (B), 3D shaded minus 2D shaded blob (C), 3D shaded minus 2D unshaded blob (D), 3D shaded minus 2D pixel scrambled (E), and 3D shaded minus 2D uniform luminance (F). In (B-F) the red outlines reproduce the activation in the conjunction (A).



**Figure 11.** Activity profiles, averaged across subjects ( $n = 18$ ) and local maxima, of the 2D shape-sensitive areas and dorsal and ventral V1 in the main experiments: (A) 3D SFT and (B) 3D SfS. Mid-FG (LOa), post-ITG (LO), LOS and DIPSMs (Denys et al. 2004; Kourtzi and Kanwisher 2000) were defined by the subtraction intact versus scrambled images (at  $P < 0.0001$  uncorrected random effect). In 4 subjects V1 was defined retinotopically and the averaged coordinates were applied across all subjects. Error bars indicate SEs between subjects. The locations of the V1 probes are indicated in Figure 6. Same color conventions as in Figure 1.

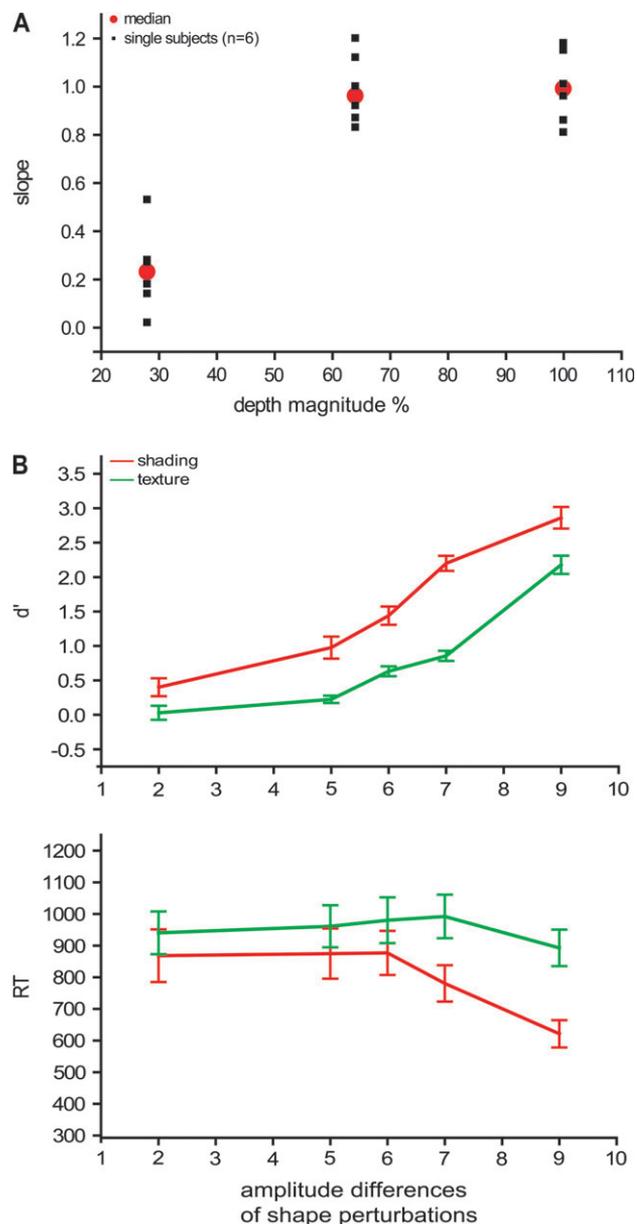


**Figure 12.** (A and C) Activity profiles plotting percent MR signal change compared with fixation condition in the main 3D Sfs sensitive regions (L and R caudal ITG) when the same visual stimuli were presented with Lambertian (A) and specular (C) shading surfaces (fixed effect,  $n = 4$ ). (B and D) Examples of a 3D shape with, respectively, Lambertian (B) and specular (D) shading.

such as fusiform and collateral sites in both hemispheres and, in 1 hemisphere, small IPS activation sites (Table S3). These shading sites were far less specific than those of texture as the activity evoked by the control conditions was proportionally larger in the shading than in the texture case (Fig. 15). Again, a number of sites were strongly activated by the uniform-luminance condition just as in the texture experiment: fusiform and dorsal IPS regions. Not surprisingly most of the sites obtained in the analysis of the active condition lost their significance in the passive condition, with the exception of caudal ITG and LOS in the left hemisphere. Thus this control experiment confirmed that parietal cortical regions are substantially more involved in 3D Sft than in 3D Sfs.

### Control Experiment 3: Central High-Acuity Task

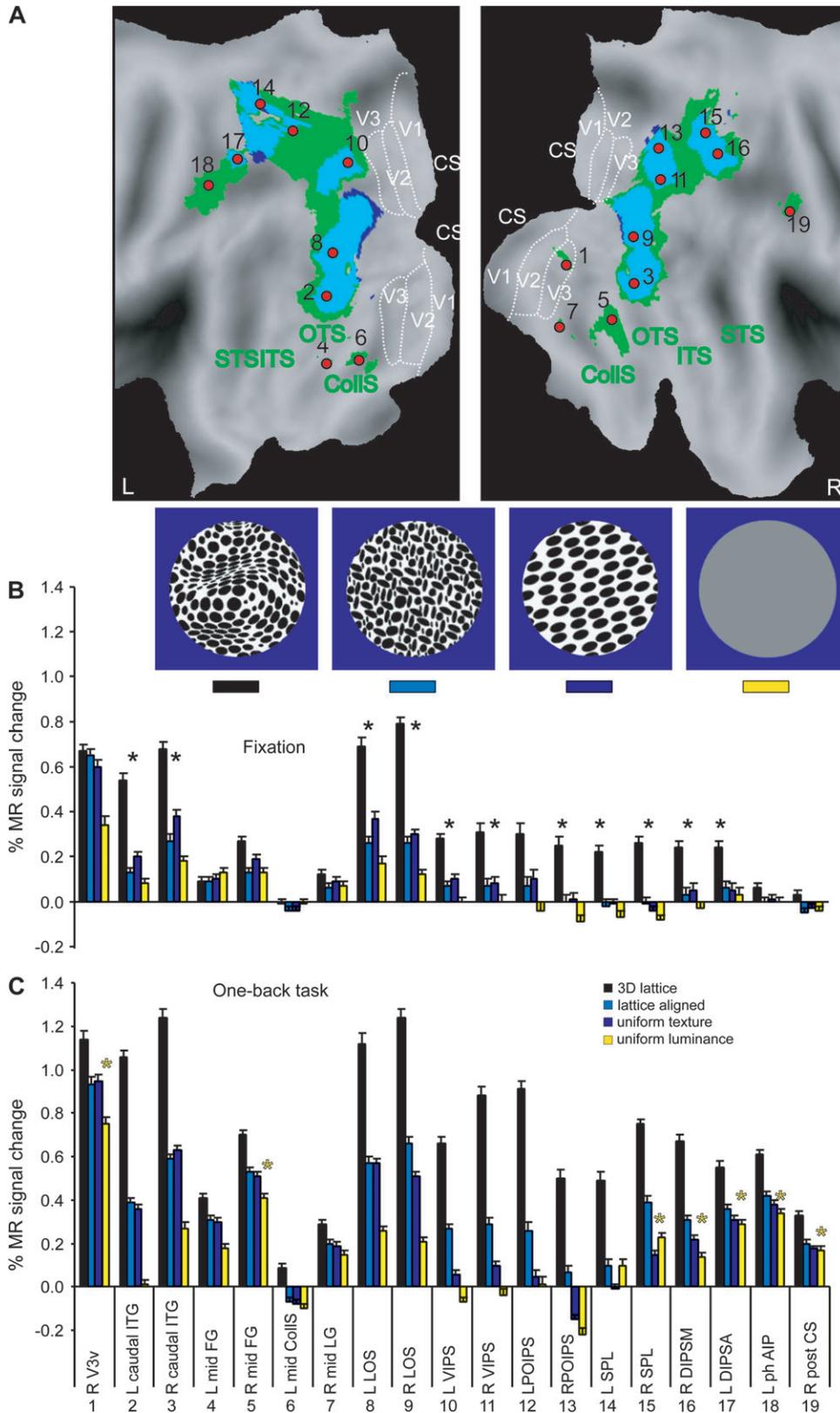
The 1-back control experiment also indicates that the activation by 3D Sft and 3D Sfs are maintained when attention is equalized across conditions, because the subjects were paying attention to the stimuli in all conditions. In 3 subjects, we removed attention from the stimuli by engaging them in a demanding central high-acuity task. In these subjects, only caudal ITG and left LOS were activated in both attention conditions. Most importantly, in all sites except right caudal ITG for shading, the activation was more significant in the high-acuity task than in the passive condition (Table 4), confirming that the activations depended little on attention.



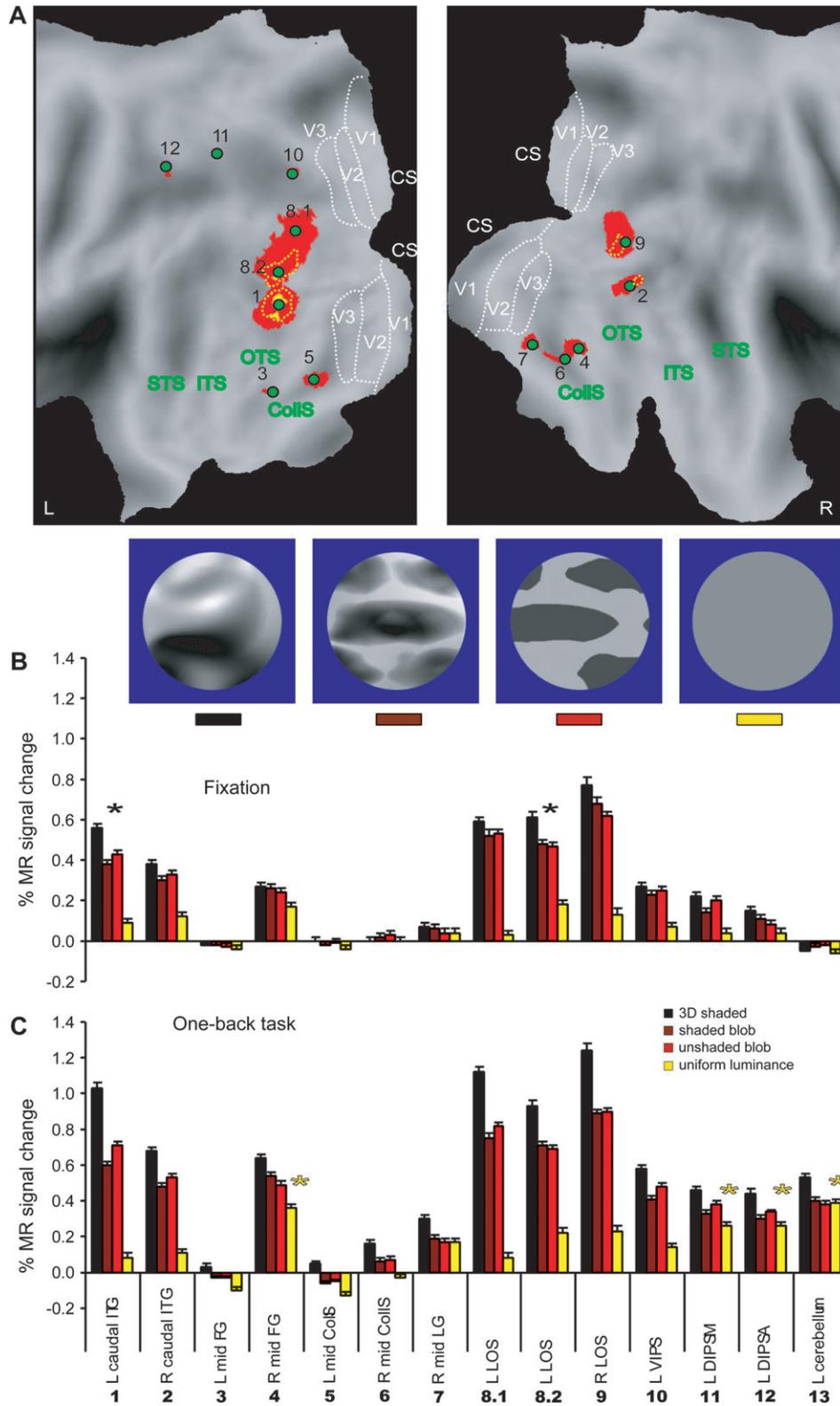
**Figure 13.** Results of psychophysical experiments 4 and 5. (A) Slope of line relating texture settings and shading settings in the 3D shape adjustment task plotted as a function of depth magnitude (in % of original). Individual values (black) and medians (red) are plotted. (B and C) Mean and SE of  $d'$  and reaction time in the same-different task plotted as a function of the amplitude of the shape perturbation. In (A) the median correlations were 0.28, 0.95, and 0.98, respectively.

### Relationship with LOC

Both the group and the single-subject analyses revealed that the regions involved in extracting 3D Sft and 3D Sfs were 2D shape-sensitive (Figs 6–8, Table 2). The question then arises as to whether the reverse also holds: are all 2D shape-sensitive regions and particularly LOC, also involved in extracting 3D shape? Table 5 indicates that that is not the case. In this table we probe the different local maxima of the lateral occipital (LO) complex, including those in the post-ITG, or LO region, in the mid-FG, or lateral occipital anterior (LOa) region, and in the LOS region (Denys et al. 2004), for significance in the conjunction of contrasts comparing 3D Sfs or 3D Sft with their respective 2D controls. With the exception of a single



**Figure 14.** 3D SFT sensitive regions (control experiment 2, fixed effect,  $n = 6$ ). (A) Texture specific areas significantly ( $P < 0.05$  corrected) activated during passive viewing (dark blue voxels, light blue overlap with “active” voxels) and when the participants performed 1-back task (green voxels) shown on the PALS flattened representations of left and right hemispheres (posterior part). The significant local maxima defined by the active runs are indicated with red dots (1) R V3v, (2) and (3) L and R caudal ITG, (4) and (5) L and R mid-FG (middle fusiform gyrus), (6) L mid CollS (collateral sulcus), (7) R mid LG (lingual gyrus), (8) and (9) L and R LOS, (10), and (11) L and R VIPS, (12) and (13) L and R POIPS, (14) and (15) L and R SPL (superior parietal lobule), (16) R DIPSM, (17) L DIPSA, (18) L phAIP (Binkofski et al. 1999; Greffkes et al. 2002), and (19) R post CS (central sulcus) (Table S2). (B and C) Activity profiles of the 19 local maxima in the passive (fixation) and active (1-back task) runs, respectively. Black stars indicate the areas significant in passive conjunction analysis (see Materials and Methods). Yellow stars indicate areas significant in the interaction: (uniform luminance minus fixation) task – (uniform luminance minus fixation) passive. Inset: examples of the visual stimuli used in the experiment.



**Figure 15.** 3D SFS sensitive regions (control experiment 2, fixed effect,  $n = 6$ ). (A) Shading specific areas significantly ( $P < 0.05$  corrected) activated during passive viewing (yellow voxels) and when the participants performed 1-back task (red voxels) shown on the PALS flattened representations of left and right hemispheres (posterior part). The yellow dotted outlines indicate voxels significant at  $P < 0.001$  uncorrected from the passive-viewing experiment. These yellow outlines and voxels are contained in the “active” red voxels. The significant local maxima defined from the active runs are indicated by green dots: (1) and (2) L and R caudal ITG, (3) and (4) L and R mid-FG (middle fusiform gyrus), (5) and (6) L and R mid CollS (collateral sulcus), (7) R mid LG (lingual gyrus), (8.1, 8.2) and (9) L and R LOS, (10) L VIPs, (11) L DIPSM, (12) L DIPSA (Table S3). For other conventions see Figure 6. (B and C) Activity profiles of all local maxima in the passive (fixation) and active (1-back task) runs. Black stars indicate the areas significant in passive conjunction analysis (see Materials and Methods). Yellow stars indicate areas significant in the interaction: (uniform luminance - fixation) task - (uniform luminance - fixation) passive. Inset: examples of the visual stimuli used in the experiment.

**Table 3**

Correlation between 3D SFT and 3D Sfs sensitive regions with Gabor and psychophysics differences between all possible combinations of conditions ( $n = 15$ )

Cue	Region	Coordinates			Corr. with Gabor diff	Corr. with psycho. diff
		x	y	z		
3D SFT	1 L caudal ITG	-50	-76	-12	0.13	0.91*
		-44	-88	-4	0.22	0.86*
	2 R caudal ITG	42	-82	-10	0.14	0.86*
		44	-66	-8	0.004	0.82*
	3 L LOS	-40	-90	12	0.18	0.98*
		-34	-86	4	0.18	0.95*
	4 R LOS	36	-90	14	0.09	0.99*
		28	-88	10	0.11	0.97*
	5 L VIPSt	-24	-72	30	-0.04	0.88*
	6 R VIPS	28	-80	20	-0.1	0.81*
3D Sfs	1 L caudal ITG	24	-72	32	0.17	0.89*
		20	-76	48	0.17	0.98*
	8 R DIPSt	22	-62	56	0.19	0.94*
	9 L DIPSt	-26	-64	56	0.22	0.95*
2 R caudal ITG	26	-54	64	0.18	0.93*	
	-48	-76	-6	-0.24	0.63 (0.62)*	
		46	-70	-8	-0.34	0.61 (0.52)*

Note: The asterisks indicate significant correlations. The numbers between the brackets indicate correlations when shaded-blob condition is excluded from the analyses.

**Table 4**

3D SFT and 3D Sfs sensitive regions, control experiment 3, fixed effect ( $n = 3$ )

Cue	Region	Coordinates			T-score	
		x	y	z	Passive viewing	High-acuity task
Texture	L caudal ITG	-42	-75	-12	3.58	<b>4.62</b>
	R caudal ITG	42	-84	-3	2.93	3.91
	L LOS	-30	-90	18	2.34	3.74
Shading	L caudal ITG	-45	-75	-6	<b>5.63</b>	<b>5.72</b>
	R caudal ITG	46	-76	-6	<b>6.32</b>	2.47

Note: The values in the T-score columns indicate significance at:  $P < 0.05$  corrected (**bold**),  $P < 0.001$  uncorrected (normal) and  $P < 0.005$  uncorrected (*italic*).

maximum in post-ITG (LO), no 2D shape-sensitive site is significantly involved in 3D Sfs. A number of 2D shape-sensitive sites in post-ITG (LO) and LOS reach significance in the conjunction defining 3D Sft sensitivity, but none in the mid-FG (LOa) region (Table 5).

We calculated the activity profiles of the shape-sensitive regions by averaging profiles over local maxima belonging to that region over the subjects. The results are shown in Figure 11 and Table S4. In none of the 3 parts of LOC was 3D Sfs more active than in any of the five 2D controls. The same was true for the 4 parietal shape-sensitive regions (Denys et al. 2004). Finally we also tested MR activity in V1 (see Materials and Methods) and again failed to obtain significance in most of the contrasts. For the extraction of 3D Sft, the results are quite different, insofar as all 4 contrasts comparing 3D Sft with 2D controls reached significance in both post-ITG (LO) and LOS, but not in mid-FG (LOa) (Table S4, Fig. 11). Also, the 3 posterior parietal shape-sensitive regions were significantly involved in 3D Sft, because all 4 contrasts reached significance in these regions. V1, on the contrary, failed to reach significance in any of the contrasts comparing 3D and 2D texture control conditions.

The previous analysis demonstrated that only a part of LOC is involved in the extraction of 3D shape from monocular cues: only the posterior subpart of the post-ITG (LO) region in 3D Sfs and only the post-ITG (LO) region but not the mid-FG (LOa)

**Table 5**

3D SFT and 3D Sfs sensitivity in the LOC local maxima

Region	LO local maxima			LO T-score	3D SFT T-score	3D Sfs T-score	
	x	y	z				
Mid-FG (LOa)	-38	-34	-28	<b>14.04</b>	ns	ns	
	-34	-32	-26	<b>13.3</b>	ns	ns	
	-38	-42	-22	<b>9.9</b>	ns	ns	
	-38	-58	-18	<b>9.65</b>	ns	ns	
	38	-54	-26	<b>16.62</b>	ns	ns	
	38	-44	-28	<b>14.79</b>	ns	ns	
	44	-56	-12	<b>10.61</b>	ns	ns	
	32	-56	-14	<b>9.52</b>	ns	ns	
	Post-ITG (LO)	-44	-74	-14	<b>13.09</b>	4.44	ns
		-50	-78	-10	<b>10.3</b>	<b>5.9</b>	3.9
-50		-68	-20	<b>10.3</b>	ns	ns	
42		-80	-14	<b>12.57</b>	4.24	ns	
50		-72	-14	<b>11.06</b>	3.3	ns	
-42		-90	-2	<b>11.34</b>	<b>5.78</b>	ns	
LOS	-34	-88	4	<b>9.7</b>	<b>6.57</b>	ns	
	-40	-88	12	<b>8.87</b>	<b>7.06</b>	ns	
	-32	-82	16	<b>8.34</b>	<b>5.49</b>	ns	
	-38	-74	0	6.8	ns	ns	
	44	-72	2	<b>8.45</b>	ns	ns	
	48	-82	-2	7.2	3.67	ns	
	44	-74	10	6.59	ns	ns	
	44	-78	8	6.48	ns	ns	

Note: The values in the T-score columns indicate significance at:  $P < 0.05$  corrected (**bold**),  $P < 0.0001$  uncorrected (normal) and  $P < 0.001$  uncorrected (*italic*). ns: nonsignificant.

region in 3D Sft. Yet a number of studies have reported the involvement of LOC in the extraction of 3D shape. We can obtain the same result in our data if we select only the most significant shape-sensitive voxels of LOC. Restricting the definition of LOC to those voxels reaching  $P < 10^{-6}$  in the comparison intact versus scrambled images yields indeed significance in all contrasts comparing 3D conditions to 2D controls both for shading and texture (Table S5). This sort of selection, which in fact selects only the most posterior voxels in LOC (Fig. S7), would be performed when only a limited number of runs are sampled in the localizer test.

## Discussion

The research described in the present article was designed to investigate the neural processes throughout the entire brain involved in the analysis of 3D Sft and 3D Sfs. Observers were presented with over 100 different 3D surfaces with a wide variety of randomly shaped hills, ridges, valleys and dimples. The experimental design included both passive and active conditions, and numerous 2D controls in order to isolate the neural activations that are specifically involved in the processing of 3D shape. Our results reveal that: 1) processing of 3D Sft activates a large swath of ventral and dorsal visual cortical regions; 2) that processing 3D Sfs activates mainly caudal ITG regions in ventral cortex, which 3) largely overlap with those involved in 3D Sft and 4) are restricted to the post-ITG part of LOC or LO proper. Control studies indicate that these results could not be explained by attention, eye movements or low-level visual differences between the conditions.

One of the aims of the present study was to investigate to what extent there is convergence or specialization in the neural processing of different depth cues to extract 3D shape. This is an issue that can only be addressed empirically, because there are no a priori reasons why different cues must be processed in the same regions of the brain. Our results reveal

that there is indeed a high degree of cue convergence, with 1 notable exception. In the ventral cortex, the activations obtained for SfT and SfS were largely overlapping in the posterior regions of LOC, which are similar to areas identified in previous studies that are involved in the analysis of 3D structure from motion and binocular disparity (Orban et al. 1999; Janssen et al. 2002). In the parietal cortex, the activations obtained for SfT were again similar to those reported previously for motion (Orban et al. 1999), but there were no significant activations for SfS. Does that mean that we failed to isolate a general factor of 3D shape processing in our study, as 1 anonymous reviewer suggested? Not necessarily, given that a modular architecture for dealing with different depth cues is a quite plausible processing strategy. Moreover, had we obtained complete convergence of all depth cues, it still would not guarantee that we had isolated a general factor of 3D shape processing. Convergence at the level of the population, as measured with fMRI is no guarantee for convergence at the single cell level, as has been observed in the caudal intraparietal or the lower bank of the superior temporal sulcus (STS) of the monkey (Tsutsui et al. 2002; Liu et al. 2004).

The processing of 3D SfT activates a large part of human visual cortex, including caudal ITG, LOS, VIPS, and DIPSM. This pattern of activation is much wider than that reported by Shikata et al. (2001), but there are significant methodological differences between the studies. Shikata et al. required subjects to judge the tilt directions of slanted planes, which were textured with a regular lattice of circular texture elements. One problem with this procedure is that the correct directions were all clearly specified by the 2D alignment of texture elements within the optical projection of each surface, so that successful performance did not necessarily require a perceptual analysis of 3D shape. The significance of this problem is reinforced by the fact that the activations reported by Shikata et al. are similar to those obtained by Faillenot et al. (2001) for judgments of 2D grating orientations. Parietal activations in discrimination tasks are difficult to interpret, as they can represent either processing of the sensory evidence or sensory-motor processing related to decisions. Indeed, even color discrimination activates human IPS regions, in all likelihood because of the decision processes involved (Claeys et al. 2004). In contrast to the slanted planes used by Shikata et al., the stimuli in the present studies depicted complex surfaces with randomly distributed regions of concavity and convexity. Because the differences in parietal activations between the 3D and 2D conditions were similar when subjects viewed the stimuli passively or performed the 1-back task, it is likely that these differences reflect sensory processing rather than sensory-motor coupling. It is important to point out, however, that our DIPSM region is very close (22, -62, 56) to the right posterior IPS activation site reported by Shikata et al. (2001) (20, -67, 58). It is also worth noting that the 3D SfT activation pattern shares many regions with 3D SfM: LOS, VIPS and DIPSM reported by Orban et al. (1999) and Murray et al. (2003) for the neural processing of 3D SfM. In SfT, the caudal ITG activation replaces the activation of MT/V5+ complex in 3D SfM, which extends ventrally into ITG (Vanduffel et al. 2002; Peuskens et al. 2004). Also the 3D SfM, compared with 3D SfT, extends further forward in IPS engaging DIPSA more fully (Orban et al. 1999).

Whereas other depth cues including texture, motion and stereo activate numerous regions along the IPS (Orban et al. 1999; Janssen et al. 2002; Vanduffel et al. 2002; Peuskens et al.

2004), no such activations were observed for 3D SfS. There was a small number of voxels near DIPSA in the left hemisphere that were activated in the 3D shape condition during the 1-back task, and this is close to the site identified previously for active judgments of 3D SfS by Taira et al. (2001). However, we also obtained activation in that site for the detection of luminance differences (see Fig. 10), and previous results have shown that regions in this vicinity are active in decision processes for several other low-level 2D properties including color, line orientation or direction of motion (Claeys et al. 2004). Because no activations were obtained in that region under passive-viewing conditions, there is little evidence to suggest that this region is involved in the sensory processing of 3D SfS.

With respect to the ventral processing stream, our results indicate that 3D SfS activations are restricted to the posterior part of the LOC complex, which is consistent with other findings for the processing of 3D SfT, motion and stereo (Orban et al. 1999; Janssen et al. 2002; Vanduffel et al. 2002; Peuskens et al. 2004). This result conflicts, however, with those obtained by Kourtzi et al. (2003), who found greater sensitivity to the sign of surface curvature in the anterior part of LOC. One important factor that could potentially account for this difference is that the Kourtzi et al. stimuli included another depth cue in addition to shading that was provided by patterns of edges and vertices (see Fig. S2), and it is possible that these 3D edge structures are processed in a different location than the analysis of smoothly curved surfaces.

The present study indicates that the 2 monocular static 3D shape cues (i.e., texture and shading) converge at the level of the caudal ITG regions. As mentioned above, this region is also engaged by the 3D SfM both in passive and active subjects (Vanduffel et al. 2002; Peuskens et al. 2004). These results suggest that several different sources of information about 3D shape converge at this level (Janssen et al. 2002; Welchman et al. 2005), and this view is supported by the finding that 3D shape from tactile information also converges with visual information in this region (Amedi et al. 2002; James et al. 2002). Thus, the caudal ITG region may provide a general, cue invariant representation of the 3D geometry of objects in the environment.

Prior to the present investigation, there have been only 2 previous studies that have employed whole-brain scans to map the cortical mechanisms involved in the processing of 3D SfS (Taira et al. 2001) or 3D SfT (Shikata et al. 2001). Because neither of these experiments included any 2D control conditions, and the stimuli they employed were not comparable to one another (i.e., a rounded square bump vs. a slanted plane), they do not allow any useful comparisons about how these important depth cues are processed within the visual cortex. The present experiment overcomes these difficulties by mapping the cortical mechanisms involved in the processing of 3D SfS and 3D SfT in the same subjects using comparable stimuli with appropriate 2D controls. The results reveal that there are substantial differences between the neural processing of 3D SfS and 3D SfT. Whereas the shading cue is processed only in ventral regions, the neural analysis of texture is performed both dorsally and ventrally, in regions that are in close proximity to those involved in the analysis of 3D SfM and stereo (Orban et al. 1999; Janssen et al. 2002; Vanduffel et al. 2002; Peuskens et al. 2004). Thus the neural substrate underlying SfS and SfT differs both in size and in location. One possible explanation for the reduced SfS activation pattern

is that observers are less sensitive to shading than they are to other depth cues. However, the results of our psychophysical control experiments indicate quite clearly that is not the case. Subjects with widely different activation patterns for SfT and SFS perceived comparable magnitudes of surface relief from those cues. Moreover, observers were actually more sensitive to SFS than SfT when they were required to detect small perturbations of 3D shape in the same-different task, or small magnitudes of relief in the profile adjustment task. These findings provide strong evidence that the reduced pattern of activation for SFS cannot be due to a lower sensitivity for shading information relative to other depth cues.

There are several possible reasons why the analysis of shading might be performed in different regions compared with the processing of other depth cues. First, information from shading is more ambiguous than that provided by other cues. Whereas motion, stereo and texture are computationally ambiguous up to a scaling in depth (see Todd and Norman 2003), shaded images have an additional shearing ambiguity (Belhumeur et al. 1999), which has also been documented empirically in several psychophysical experiments (Koenderink et al. 2001; Cornelis et al. 2003; Todd 2004). Second, shading differs from other cues in that it cannot provide information about the slant of a planar surface. It is only relevant to the analysis of curved surfaces. Surface 3D orientation is important for visual guidance of action, an important function of parietal cortex (Andersen and Buneo 2002; Durand et al. 2007). Finally, the pattern of image shading is influenced by a wider variety of environmental factors, including the surface geometry, the pattern of illumination and the surface material properties. This last difference is likely to be especially significant for the neural mechanisms of 3D shape perception. Because the ventral pathway contains mechanisms involved in the analysis of material properties (Peuskens et al. 2004), and discounting variations in illumination (Bartels and Zeki 2000; Hanazawa and Komatsu 2001), it makes sense that the analysis of Sfs would also be primarily localized in the ventral pathway so that these other properties can more easily be taken into account.

Given the extra work that is required for disentangling the effects of surface geometry, surface reflectance and the pattern of illumination in order to determine 3D Sfs, it is quite surprising that the brain performs this analysis with only a subset of the neural resources that are employed to estimate 3D shape from other depth cues. One possible solution to this conundrum that was suggested by an anonymous reviewer is that the level of neural resources employed could be determined in part by the apparent realism of the stimuli. Although the textured stimuli in our experiments produced vivid perceptions of 3D shape, most observers agree that they appear less like natural objects than the shaded stimuli, perhaps because the depth structure does not become apparent as readily. This would also be the case for the random dot patterns that are typically used to study the perception of 3D structure from motion or stereo. It is interesting to note in this context that others have related naturalness of stimuli to reduced activation in visual cortex (Murray et al. 2002), and that may well have played a role in the present studies. Another potential solution to this conundrum is that the difficulties involved in the analysis of shading make it an unreliable source of information for the online control of motor actions, which is likely to be a primary function for the processing of 3D shape

within the parietal cortex. Additional research will obviously be required to sort out these issues.

### Conclusions

In the present investigation, we mapped human cortical regions involved in extracting 3D SfT and 3D SFS using a large set of over 100 3D surfaces with randomly shaped hills, ridges, valleys and dimples, and a large battery of control conditions to distinguish the processing of 3D shape from the analysis of 2D image properties. The results reveal that 3D Sfs involves the caudal infero-temporal gyrus in both hemispheres; 3D SfT involves in addition to the caudal ITG region, LOS, and several IPS sites in both hemispheres. These activation patterns were consistent across subjects. These results underscore the importance of the posterior part of the LOC for the extraction of visual 3D shape information from several depth cues, and they suggest strongly that the importance of shading is diminished relative to other cues for the analysis of 3D shape in parietal regions.

### Supplementary Material

Supplementary material can be found at: <http://www.cercor.oxfordjournals.org/>.

### Funding

GOA 2005/18, FWO G 0151.04, Insight 2+ (IST-2001-29688) and Neurobotics (IST-001917).

### Notes

We acknowledge Y. Celis, M. De Paep, W. Depuydt, P. Kayenbergh, and G. Meulemans for the technical support. We also thank Prof. R. Vandenberghe for suggesting control experiment 2, Prof. R. Vogels for help with the same-different task and Prof. P. Van Hecke for access to the scanner. *Conflict of Interest:* None declared.

Address correspondence to Prof. Dr Guy A. Orban, Laboratorium voor Neuro- en Psychofysiologie, O&N2, Herestraat 49, bus 1021, Katholieke Universiteit Leuven School of Medicine, Campus Gasthuisberg, B-3000 Leuven, Belgium. Email: [guy.orban@med.kuleuven.be](mailto:guy.orban@med.kuleuven.be).

### References

- Amedi A, Jacobson G, Hendler T, Malach R, Zohary E. 2002. Convergence of visual and tactile shape processing in the human lateral occipital complex. *Cereb Cortex*. 12:1202-1212.
- Andersen RA, Buneo CA. 2002. Intentional maps in posterior parietal cortex. *Annu Rev Neurosci*. 25:189-220.
- Bartels A, Zeki S. 2000. The architecture of the colour centre in the human visual brain: new results and a review. *Eur J Neurosci*. 12:172-193.
- Belhumeur PN, Kriegman D, Yuille A. 1999. The bas-relief ambiguity. *Int J Comput Vis*. 23:149-168.
- Binkofski F, Buccino G, Posse S, Seitz RJ, Rizzolatti G, Freund HJ. 1999. A fronto-parietal circuit for object manipulation in man: evidence from an fMRI study. *Eur J Neurosci*. 11:3276-3286.
- Blake A, Bulthoff HH, Sheinberg D. 1993. Shape from texture: ideal observers and human psychophysics. *Vision Res*. 33:1723-1737.
- Boyaci H, Fang F, Murray SO, Kersten D. 2007. Responses to lightness variations in early human visual cortex. *Curr Biol*. 17:989-993.
- Claeys KG, Dupont P, Cornette L, Sunaert S, Van Hecke P, De Schutter E, Orban GA. 2004. Color discrimination involves ventral and dorsal stream visual areas. *Cereb Cortex*. 14:803-822.
- Cornelis EVK, van Doorn AJ, de Ritter H. 2003. Mirror-reflecting the shape of an object: what happens to the shape percept? *Percept Psychophys*. 65:1110-1125.

- Denys K, Vanduffel W, Fize D, Nelissen K, Peuskens H, Van Essen D, Orban GA. 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.
- Durand JB, Nelissen K, Joly O, Wardak C, Todd JT, Norman JF, Janssen P, Vanduffel W, Orban GA. 2007. Anterior regions of monkey parietal cortex process visual 3D shape. *Neuron*. 55:493-505.
- Faillenot I, Sunaert S, Van Hecke P, Orban GA. 2001. Orientation discrimination of objects and gratings compared: an fMRI study. *Eur J Neurosci*. 13:585-596.
- Fize D, Vanduffel W, Nelissen K, Denys K, Chef d'Hotel C, Faugeras O, Orban GA. 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 RW, Torralba A, Adelson EH. 2004. Specular reflections and the perception of shape. *J Vision*. 4:798-820.
- Friston KJ, Holmes AP, Price CJ, Buchel C, Worsley KJ. 1999. Multisubject fMRI studies and conjunction analyses. *Neuroimage*. 10:385-396.
- Friston KJ, Penny WD, Glaser DE. 2005. Conjunction revisited. *Neuroimage*. 25:661-667.
- Gilaie-Dotan S, Ullman S, Kushnir T, Malach R. 2001. Shape-selective stereo processing in human object-related visual areas. *Hum Brain Mapp*. 15:67-79.
- Grefkes C, Weiss PH, Zilles K, Fink GR. 2002. Crossmodal processing of object features in human anterior intraparietal cortex: an fMRI study implies equivalencies between humans and monkeys. *Neuron*. 35:173-184.
- 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.
- Henson RNA, Penny WD. 2003. ANOVAs and SPM. Technical report. Wellcome Department of Imaging Neuroscience.
- Humphrey GK, Goodale MA, Bowen CV, Gati JS, Vilis T, Rutt BK, Menon RS. 1997. Differences in perceived shape from shading correlate with activity in early visual areas. *Curr Biol*. 7:144-147.
- James TW, Humphrey GK, Gati JS, Servos P, Menon RS, Goodale MA. 2002. Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia*. 40:1706-1714.
- Janssen P, Pareto D, Peuskens H, Sunaert S, Vogels R, Orban GA. 2002. Higher order disparity sensitive regions in human cortex? An fMRI study. *Soc Neurosci Abstr*. 56:11.
- Knill DC. 1998a. Ideal observer perturbation analysis reveals human strategies for inferring surface orientation from texture. *Vision Res*. 38:2635-2656.
- Knill DC. 1998b. Discrimination of planar surface slant from texture: human and ideal observers compared. *Vision Res*. 38:1683-1711.
- Koenderink JJ, van Doorn AD. 2004. Shape and shading. *Vis Neurosci*. 2:1090-1105.
- Koenderink JJ, van Doorn AJ, Kappers AML, Todd JT. 2001. Ambiguity and the "Mental Eye" in pictorial relief. *Perception*. 30:431-448.
- Kourtzi Z, Erb M, Grodd W, Bulthoff HH. 2003. Representation of the perceived 3-D object shape in the human lateral occipital complex. *Cereb Cortex*. 13:911-920.
- Kourtzi Z, Kanwisher N. 2000. Cortical regions involved in perceiving object shape. *J Neurosci*. 20:3310-3318.
- Kovesi P. 1999. Image features from phase congruency. *J Comput Vision Res*. 1:1-28.
- Lades M, Vorbruggen C, Buhmann J, Lange J, von der Malsburg C, Wurtz R, Konen W. 1993. Distortion invariant object recognition in the dynamic link architecture. *IEEE Trans Comput*. 42:300-311.
- Lamme VA. 1995. The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci*. 15:1605-1615.
- Larsson J, Heeger DJ. 2006. Two retinotopic visual areas in human lateral occipital cortex. *J Neurosci*. 26:13128-13142.
- Lee TS, Mumford D, Romero R, Lamme VA. 1998. The role of the primary visual cortex in higher level vision. *Vision Res*. 38:2429-2454.
- Liu Y, Vogels R, Orban GA. 2004. Convergence of depth from texture and depth from disparity in macaque inferior temporal cortex. *J Neurosci*. 24:3795-3800.
- Mendola JD, Dale AM, Fischl B, Liu AK, Tootell RB. 1999. The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *J Neurosci*. 19:8560-8572.
- Moore C, Engel SA. 2001. Neural response to perception of volume in the lateral occipital complex. *Neuron*. 29:277-286.
- Murray SO, Kersten D, Olshausen BA, Schrater P, Woods DL. 2002. Shape perception reduces activity in human primary visual cortex. *Proc Natl Acad Sci USA*. 22:15164-15169.
- Murray SO, Olshausen BA, Woods DL. 2003. Processing shape, motion and three-dimensional shape-from-motion in the human cortex. *Cereb Cortex*. 13:508-516.
- Nefs HT, Koenderink JJ, Kappers AML. 2006. Shape-from-shading for matte and glossy objects. *Acta Psychol Amst*. 121:297-316.
- Norman JF, Todd JT, Orban GA. 2004. Perception of three-dimensional shape from specular highlights, deformations of shading, and other types of visual information. *Psychol Sci*. 15:565-570.
- Norman JF, Todd JT, Phillips F. 1995. The perception of surface orientation from multiple sources of optical information. *Percept Psychophys*. 57:629-636.
- Orban GA, Fize D, Peuskens H, Denys K, Nelissen K, Sunaert S, Todd J, Vanduffel W. 2003. Similarities and differences in motion processing between the human and macaque brain: evidence from fMRI. *Neuropsychologia*. 41:1757-1768.
- Orban GA, Sunaert S, Todd JT, Van Hecke P, Marchal G. 1999. Human cortical regions involved in extracting depth from motion. *Neuron*. 24:929-940.
- Penny WD, Holmes AP. 2003. Random effect analysis. In: Frackowiak RSJ, Friston KJ, Frith CD, Dolan R, Price CJ, Zeki S, Ashburner J, Penny WD, editors. *Human brain function*. 2nd ed. New York: Academic Press. p. 843-850.
- Peuskens H, Claeys KG, Todd JT, Norman JF, Van Hecke P, Orban GA. 2004. Attention to 3-D shape, 3-D motion, and texture in 3-D structure from motion displays. *J Cogn Neurosci*. 16:665-682.
- Press WA, Brewer AA, Dougherty RF, Wade AR, Wandell BA. 2001. Visual areas and spatial summation in human visual cortex. *Vision Res*. 41:1321-1332.
- Sawamura H, Georgieva S, Vogels R, Vanduffel W, Orban GA. 2005. Using functional magnetic resonance imaging to assess adaptation and size invariance of shape processing by humans and monkeys. *J Neurosci*. 25:4294-4306.
- Shikata E, Hamzei F, Glauche V, Knab R, Dettmers C, Weiller C, Buchel C. 2001. Surface orientation discrimination activates caudal and anterior intraparietal sulcus in humans: an event-related fMRI study. *J Neurophysiol*. 85:1309-1314.
- Sunaert S, Van Hecke P, Marchal G, Orban GA. 1999. Motion-responsive regions of the human brain. *Exp Brain Res*. 127:355-370.
- 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.
- Todd JT. 2004. The perception of 3D shape. *Trends Cogn Sci*. 8:115-121.
- Todd JT, Norman JF. 2003. The perception of 3D shape from multiple cues: are observers capable of perceiving metric structure? *Percept Psychophys*. 65:31-47.
- Todd JT, Oomes AH, Koenderink JJ, Kappers AM. 2004. The perception of doubly curved surfaces from anisotropic textures. *Psychol Sci*. 15:40-46.
- Todd JT, Thaler L, Dijkstra TM. 2005. The effects of field of view on the perception of 3D slant from texture. *Vision Res*. 45:1501-1517.
- Tsutsui K, Sakata H, Naganuma T, Taira M. 2002. Neural correlates for perception of 3D surface orientation from texture gradient. *Science*. 298:409-412.
- Tyler CW, Likova LT, Kontsevich LL, Wade AR. 2006. The specificity of cortical region KO to depth structure. *Neuroimage*. 30:228-238.
- Vanduffel W, Fize D, Mandeville JB, Nelissen K, Van Hecke P, Rosen BR, Tootell RB, Orban GA. 2001. Visual motion processing investigated using contrast agent-enhanced fMRI in awake behaving monkeys. *Neuron*. 32:565-577.
- Vanduffel W, Fize D, Peuskens H, Denys K, Sunaert S, Todd JT, Orban GA. 2002. Extracting 3D from motion: differences in human and monkey intraparietal cortex. *Science*. 298:413-415.

- Van Essen DC. 2005. A population-average, landmark- and surface-based (PALS) atlas of human cerebral cortex. *Neuroimage*. 28:635-662.
- Van Oostende S, Sunaert S, Van Hecke P, Marchal G, Orban GA. 1997. The kinetic occipital (KO) region in man: an fMRI study. *Cereb Cortex*. 7:690-701.
- Welchman AE, Deubelius A, Conrad V, Bulthoff HH, Kourtzi Z. 2005. 3D shape perception from combined depth cues in human visual cortex. *Nat Neurosci*. 8:820-827.
- Zenger-Landolt B, Heeger DJ. 2003. Response suppression in v1 agrees with psychophysics of surround masking. *J Neurosci*. 23:6884-6893.