Exp Brain Res (2008) 189:109–120
DOI 10.1007/s00221-008-1406-9

RESEARCH ARTICLE

Natural textures classification in area V4 of the macaque monkey

F. Arcizet · C. Jouffrais · P. Girard

Received: 1 November 2007 / Accepted: 25 April 2008 / Published online: 28 May 2008

Abstract Natural texture of an object is an important cue for recognition. In real conditions, the incidence angle of light on natural textures leads to a complex pattern of micro-shading that modifies 3D rendering of surfaces. Little is known about visual processing of material properties. The present work aims to study the coding of natural textures by the neurons of area V4 of the awake macaque monkey. We used patches of natural textures issued from the CURET database and illuminated with two or three different angles with their corresponding controls (scrambled Fourier phase). We recorded the responses of V4 neurons to stimuli flashed in their receptive fields (RFs) while the macaques performed a simple fixation task. We show that a large majority of V4 neurons responded to texture patches with a strong modulation across stimuli. The analysis of those responses indicate that V4 neurons integrate first and second order parameters in the image (mean luminance, SNR, and energy), which may be used to achieve texture clustering in a multidimensional space. This clustering was comparable to that of a pyramid of Gabor filters and was not affected by illumination angles. Altogether, these results suggest that the V4 neuronal population acts as a set of filters able to classify textures independently of illumination angle. We conclude that area V4 contains mechanisms that are sensitive to the aspect of textured surfaces, even in an environment where illumination changes continuously.

Keywords 3D vision · Natural texture · Awake monkey · Electrophysiology · V4

Introduction

Almost every object has a surface covered with a distinctive texture. In addition to 3D geometry, humans can use this important material property to recognize objects or classify them (Humphrey et al. 1994; Price and Humphreys 1989; Rossion and Pourtois 2004). Texture is characterized by the repetitive occurrence of a pattern that may appear more or less regular at different scales. This important cue reflects the physical composition of the objects and the characteristics of their surface. The texture that covers any object has a 3D aspect that can be more (rough) or less (smooth) pronounced. In real conditions, illumination casts a micro-shading pattern on the object surface that depends on its texture. When the illumination angle varies, the micro-shading pattern is modified, leading to an important change in the global aspect of the object surface. Despite these marked visual differences, it is still possible to identify material properties of real objects in a visual scene.

Electronic supplementary material The online version of this article (doi:10.1007/s00221-008-1406-9) contains supplementary material, which is available to authorized users.

F. Arcizet · C. Jouffrais · P. Girard (✉)
Faculté de Médecine de Rangueil, Université de Toulouse, CerCo, UPS, CNRS, UMR5549, 31062 Toulouse, France
e-mail: pascal.girard@cerco.ups-tlse.fr

C. Jouffrais
IRIT, CNRS-Université Toulouse 3-INPT-Université Toulouse 1, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse CEDEX 9, France

Present Address:
F. Arcizet
UCLA, Department of Neurobiology, CHS 58-114, Bisley's Lab, 10833 Le Conte Ave, Los Angeles, CA 90095-1763, USA
e-mail: farcizet@mednet.ucla.edu

Humans can even recognize natural textures on 2D pictures of objects that have been taken under various lighting conditions.

The results of several neurophysiological studies in the primate visual system show that texture difference can be used as a cue for objects segmentation (Lamme 1995), or boundaries detection (von der Heydt et al. 1984). However, virtually nothing is known about neuronal processes that underlie recognition of material properties, even less in different lighting environments. The ventral visual pathway in the primate cortex is devoted to object recognition but it is relatively unclear at which level the information concerning material properties is extracted. Area V4 is likely to be crucial for texture encoding since several studies have identified deficits in texture segmentation following lesions of macaque area V4 (Schiller 1993; De Weerd et al. 1996; Merigan 1996). V4 neurons are selective to orientation, length, and width of simple bars (Desimone and Schein 1987), but elementary constituents of textures (textons) are more complex than those simple attributes (Julesz and Schumer 1981). It has been shown that V4 neurons also respond to complex shapes (Kobatake and Tanaka 1994), non-Cartesian gratings (Gallant et al. 1996; Hegde and VanEssen 2007) and complex boundary features at specific locations within larger shapes (Pasupathy and Connor 1999, 2001). Consequently, textons should efficiently stimulate V4 receptive fields (RFs) that have complex inner structure (Pollen et al. 2002). Finally, V4 is a major source of projections to the inferotemporal cortex where it has been recently shown that neurons are selective to texture-defined gradients (Liu et al. 2004). In addition, one study by Hanazawa and Komatsu (2001) showed that V4 neurons are selective to the gradient of illumination and the density and size of artificial texture elements that are regularly spaced. The question remains as to how complex features could be integrated as elementary components of textures and combined by the visual system to characterize the material properties of an object, even with regularly changing illumination conditions.

The aim of the present study was to examine the neuronal encoding of natural textures in area V4 of the awake macaque monkey. We used a large set of stimuli representing 2D pictures of textures illuminated with different angles. Our goal was twofold: first, to determine if the V4 neuronal population is selective enough to code for natural texture classification; second, to determine if this classification is invariant to lighting angles. Our set of stimuli was chosen to allow a comparison with computational studies that have designed texture classifiers. Similarly, we show that the V4 neurons population resembles a set of filters able to classify textures independently from illumination angle.

Methods

Single-cell recording

Single-cell activity was recorded in two adult rhesus monkeys, one female (M1) and one male (M2), weighing 3 and 6 kg, respectively. A first surgery, performed under general anesthesia and sterile conditions, consisted of implanting a head post (CRIST®). A pain reliever (Ketofen, 20 mg/kg) and systemic antibiotics were administrated just before the surgery started. Induction of anesthesia was performed by an injection of ketamine (16 mg/kg IM). Deeper anesthesia was achieved with a mixture of alphadolone/alphaxolone (Saffan, 15 mg/kg/h IV, rate adjusted if required). A recording chamber (CRIST®) was implanted during a second surgery, which was performed under the same conditions. In addition, corticoids were injected (Solumedrol, 1 mg/kg, IM) to prevent brain edema. During training and recording sessions the animals were seated in a primate chair, with their head restrained, in front of a computer monitor at a distance of 57 cm. An ISCAN infrared eye-tracking system (120 Hz) monitored eye positions by tracking the corneal reflection of a focused infrared LED through a CCTV camera with a 250-mm lens.

We recorded V4 neurons in the lower left parafoveal representation of the visual field. The location of the recording chamber (2 cm diameter) over the prelunate gyrus was based on stereotaxic coordinates as mentioned in Girard et al. (2002), and on skull landmarks and sulci positions that were perceptible during the surgery, through the intact dura. Single cell activity was recorded using tungsten in glass-coated electrodes with impedances of 0.5–1.5 MΩ (Trec®). Their position was controlled with a stepping motor microdrive (Trec®). The electrical signal was amplified and filtered and single unit activity was recorded on-line with a spike sorting software (Alpha-Omega MSD®) and oscilloscope.

For each isolated cell, the coarse location of the RF was first determined with dark, light or colored hand-moved bars while the animal was fixating a central spot. In order to roughly map out the RF of the cells, we used a computer-controlled sequence of dark and light 1° squares briefly flashed (25 ms) with a 25 ms inter-stimulus interval. Each square was presented ten times following a pseudo-random order, in the left inferior quadrant of the visual field within a 12° square region centered on five degrees eccentricity, where we expected to find the RFs. In some cases, the size and the position of the black and white squares were modified in order to refine the RF plotting. Using forward correlation of spike times with stimulus times, it was possible to determine the RF of the neuron within a few minutes. The classical RF of the cell was delineated as a region in which reliable excitatory

responses could be evoked by either black and/or white squares. We then placed the stimuli (texture patches) at the center of the RF. RF size was estimated as a function of eccentricity from Gattass et al. (1988) formula (size = 1° + 0.625X eccentricity). After each recording session, we used the RF centers to draw retinotopic maps of the recorded regions for both monkeys. We superimposed these maps on the drawings of sulcal landmarks that were estimated during surgery (see suppl. Fig. A).

Stimuli and protocol

Stimuli were 2D patches of natural texture images from the columbia-utrecht reflectance and texture (CUReT) database. A detailed description of this database can be found at http://www1.cs.columbia.edu/CAVE/software/curet/index.php. The stimuli that we selected from the database were natural textures photographed in frontal view with a light source coming from the right side. For each texture, we chose two (or three) patches illuminated with incidence angles of 22.5, 45, (and 67.5°) that we termed lighting angles L1, L2, (and L3). Our stimuli were squares (2–5 square degrees) cut at the center of these pictures, converted in 256 gray-levels images. Stimuli size was chosen among four predefined sets (2, 3, 4, and 5 degrees) to be included in the interval [0.5–1] * RF estimated size. The stimuli were always completely contained in the RF (corresponding in average to 64% of the RF). For the sake of readability, we indicated the name and/or the number of the texture patches (this number being used in figures).

In a first set of stimuli (set A), we used 41 textures illuminated from angles L1 and L2 (total 82 stimuli). We designed a second set of stimuli (set B) including 12 textures from the set A, and adding the corresponding exemplars illuminated with L3 (Fig. 1). For each stimulus, we designed a control image (coined scrambled-phase) with the same power spectrum but random redistribution of phases (Fig. 1b). Hence, the set B contained 72 different stimuli consisting in 12 patches * 3 lighting angles * 2 phase spectra (original vs. scrambled). The stimuli were gamma corrected on a 21″ CRT monitor (Iiyama vision master pro512) placed 57 cm in front of the eyes of the monkeys. Mean luminance of the stimuli was 24.15 cd/m^2 and background luminance was 11.5 cd/m^2.

CORTEX software (courtesy of NIMH) controlled behavior, stimulus presentation and data acquisition. During each trial, the monkey was required to maintain fixation on a 0.1 degree gray central spot for a variable delay (400–600 ms) within a 2-degree square window, before a stimulus was flashed for 250 ms. The trial was completed after a second random duration fixation (200–300 ms) without the stimulus. Trials completed without breaking fixation were rewarded with a drop of water. The inter stimulus interval between successive correct trials ranged from 600 to 900 ms. Stimuli were pseudo-randomly selected and were presented 5–10 times each. The monkeys were weighted every working day and additional water was given if needed. All animal procedures complied with guidelines of the European Ethics Committee on Use and Care of Animals.

Texture patches analysis

The determination of relevant parameters to fully describe a texture is a wide computational field. A review of this topic is beyond the scope of this paper. Among four major categories of methods for texture analysis (statistical, geometrical, model-based, and signal processing), we chose the statistical method that describes the relationships between one pixel and its neighbors.

In this method, first-order statistics describe the properties of individual values of pixels in the picture, ignoring their relative spatial locations. We computed several first-order parameters such as mean and variance of luminance, root mean square (RMS) contrast, and signal to noise ratio (SNR) from the gray-levels histogram.

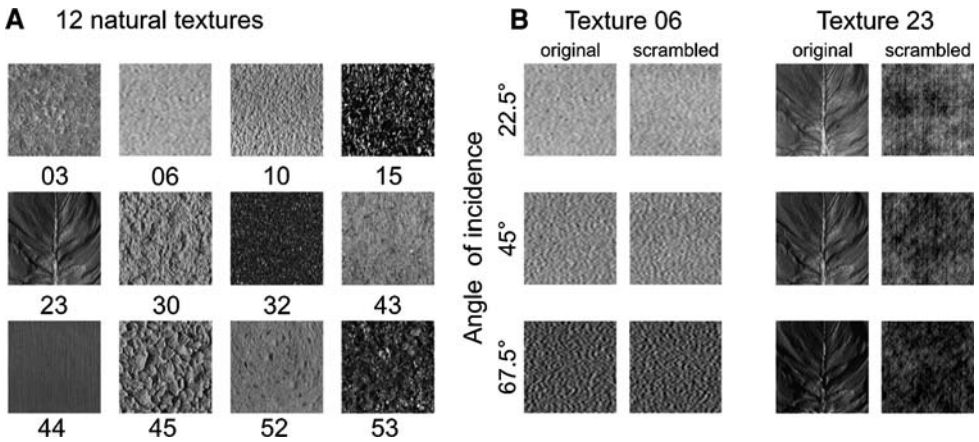

Fig. 1 **a** The 12 natural textures belonging to both sets of stimuli (set A and set B). Illumination angle displayed here is 22.5° (L1). Stimulus number is indicated below each texture patch. **b** Two textures patches (#6 and #23) illuminated with the three different angles from the frontal axis (L1 = 22.5°, L2 = 45°, and L3 = 67.5°) and their corresponding scrambled-phase patches

Second and higher-order statistics describe the properties of two or more pixels values at a specific location. A widely used second-order method relies on co-occurrence matrices that contain the number of occurrences of two gray-level values, separated by a given pixel distance in a given direction in the image (Haralick et al. 1973). From these matrices, one can compute up to 14 parameters that characterize roughness, granularity or homogeneity of the textures. We focused on four of second-order parameters—energy, inertia, entropy, and homogeneity, which we computed across four different orientations (0°, 45°, 90°, and 135°). As our textures are mostly isotropic, second order parameters varied little across orientations. Table 1 of supplementary material summarizes the formulas used to compute first and second order parameters. Except energy, first and second order parameters do not change significantly with phase scrambling (ANOVA, $P < 0.05$). In set B, mean and SNR statistically decrease with illumination angle (IA, $P = 3.4 \times 10^{-5}$ and $P = 0.007$), whereas RMS increases with IA ($P = 0.029$). Luminance variance and second order parameters do not vary with IA.

One of the major fields of image analysis consists in building algorithms that achieve texture classification (Hayman et al. 2004; Rao and Lohse 1996; Varma and Zisserman 2003). To classify our stimuli, we used a multiscale 2D Gabor filters bank, keeping the ratio between size and frequency of the filters constant. The 12 filters (six even and six odd-symmetric) had frequencies increasing with one octave step from 0.5 to 16 cycles per degree. For each filter, we used eight orientations from 0 to 157.5°, evenly increasing by 22.5°. Each texture was convolved with each individual filter. We then computed the energy of each convolution (square root of the squared mean filtered image) and used this value in the multidimensional analysis (see below).

Data analysis

We recorded 148 V4 neurons in the right hemisphere of two monkeys, 100 in monkey M1 and 48 in monkey M2. Response rates were computed during a period ranging from 50 to 300 ms after stimulus onset. The reference period started during the fixation period, 400 ms before the onset of the stimulus, and lasted 250 ms. All analyses were performed on both the raw activity and on the discharge rates on which we subtracted the baseline activity (spontaneous discharge rate computed during the reference period). As the results were consistent, only the results obtained with the raw discharge rates are described in the paper.

We used a non-parametric Mann–Whitney test to select the neurons that presented a significant difference between the response to patches and the baseline activity. Then, we used a Kruskal–Wallis test to compare single cells response across textures. Finally, the responses of the neurons were analyzed with two complementary multivariate exploratory technique. First, we produced hierarchical cluster dendrograms (using Ward method) on standardized responses (Sary et al. 2004). Second, we computed Euclidian distance response matrices for all possible image comparisons. We used multidimensional scaling (MDS) final configurations to represent these distances in a low-dimensional space (Young and Yamane 1992). We applied MDS to V4 cells responses and to the energies computed from the multiscale 2D Gabor filters bank. We used scree plot (stress against the number of dimensions) analysis to compute the dimensionality of the MDS configurations (Kruskal and Wish 1978). We then compared the two distances matrices (cells response and filters energy) with a Procrustes rotation procedure, which consists in translating, dilating, and rotating one matrix to find the best fit with the other one. We used a statistical test (PROTEST) based on a Monte Carlo simulation (Matlab code courtesy of Pedro Peres-Neto, Dept. of Zoology, University of Toronto) to check that the congruence between the two distance matrices was not obtained by chance (Jackson 1995).

Results

Neuronal database

We recorded 148 neurons (52 with set A and 96 with set B) in the right dorsal V4. Accordingly, the RFs were located in the inferior left quadrant of the visual field. The average eccentricity was 5.6° with a range of 1.4° to 11.3°. A vast majority of neurons (97%, $n = 144$) had a significant response to stimulus presentation compared to the baseline activity (Mann–Whitney test, $P < 0.05$). These 144 responsive neurons—97 in monkey M1 and 47 in monkey M2—were kept for subsequent analysis (51 with set A and 93 with set B).

Texture selectivity

We explored single cell response patterns to the 41 textures included in set A. Texture stimuli were very efficient in driving V4 neurons. The responses of 49 neurons (97%) were significantly modulated according to natural texture identity (Kruskal–Wallis, $P < 0.05$). We called these neurons "texture selective (TS) neurons". Figure 2 illustrates the response pattern of a typical TS neuron. When compared to the baseline activity, this neuron had a significant response to 38 textures (only responses to textures #07, #34, and #41 are not significant). Similarly, the majority of TS neurons (70%, $n = 34$) responded significantly to at least

20 textures among the 41. Seven cells were responsive to the 41 textures of set A. The narrowest response pattern was observed for two cells that had significant responses to five textures only.

The typical example of Fig. 2 shows a response pattern with a strong response modulation across textures. In order to describe this modulation, we computed the selectivity index and sparseness index for each TS neuron. Selectivity index (SI) is $[R_{max} - R_{min}]/[R_{max} + R_{min}]$, where R_{min} is the minimal response and R_{max} the maximal response. Sparseness index (SpI) is $[\Sigma_{i=1,n}(R_i/n)]^2/[\Sigma_{i=1,n}(R_i^2/n)]$, where R_i is the mean response to texture i and $n = 41$. A SpI value of 1.00 indicates similar responses to all stimuli (Rolls and Tovee 1995). Figure 3 shows a scatter plot and distribution histograms for the SI and SpI computed on TS population. The SI histogram shows that TS neurons present a strong selectivity for natural textures. The SpI histogram shows that TS neurons significantly respond to many natural textures (large bandwidth); the sparseness of the representation of the 41 natural textures had an average of 0.81 across the TS population. Hence, typical V4 neurons have a marked preference for a wide range of textures (see example in Fig. 2). In order to check that texture preference was not due to a specific position of local cues (e.g., leaf vein in #23), we designed a control test in which the whole set of stimuli was shifted by 1° within the RF. Figure 4a shows an example of a neuron among the ten recorded with this control. This plot clearly shows that texture selectivity is preserved in spite of the shift of the stimuli in the RF. The ranking of texture preference was confirmed among the population of ten neurons (Spearman rank, $r = 0.881$, $P < 0.001$; see Fig. 4b).

In order to go beyond a mere qualitative description of textures that cannot explain the data, we searched for significant linear regression between statistical parameters of the patches (first order and second order parameters listed in Table 1 of supplementary material) and single cells responses. Only 18, 14, 22, and 16% of TS cells responses were correlated to mean, variance, RMS, and SNR, respectively; and less than 18% were correlated to second-order parameters such as energy, entropy, inertia, and homogeneity ($R^2 > 0.15$, $P < 0.001$). These results show that a single parameter does not account for the response pattern of the cells. Hence, we used multivariate exploratory techniques (hierarchical cluster analysis and MDS) to reveal potential clusters of stimuli that evoke similar population responses. Figure 5 shows the dendrogram in which each terminal (horizontal) branch of the tree represents the

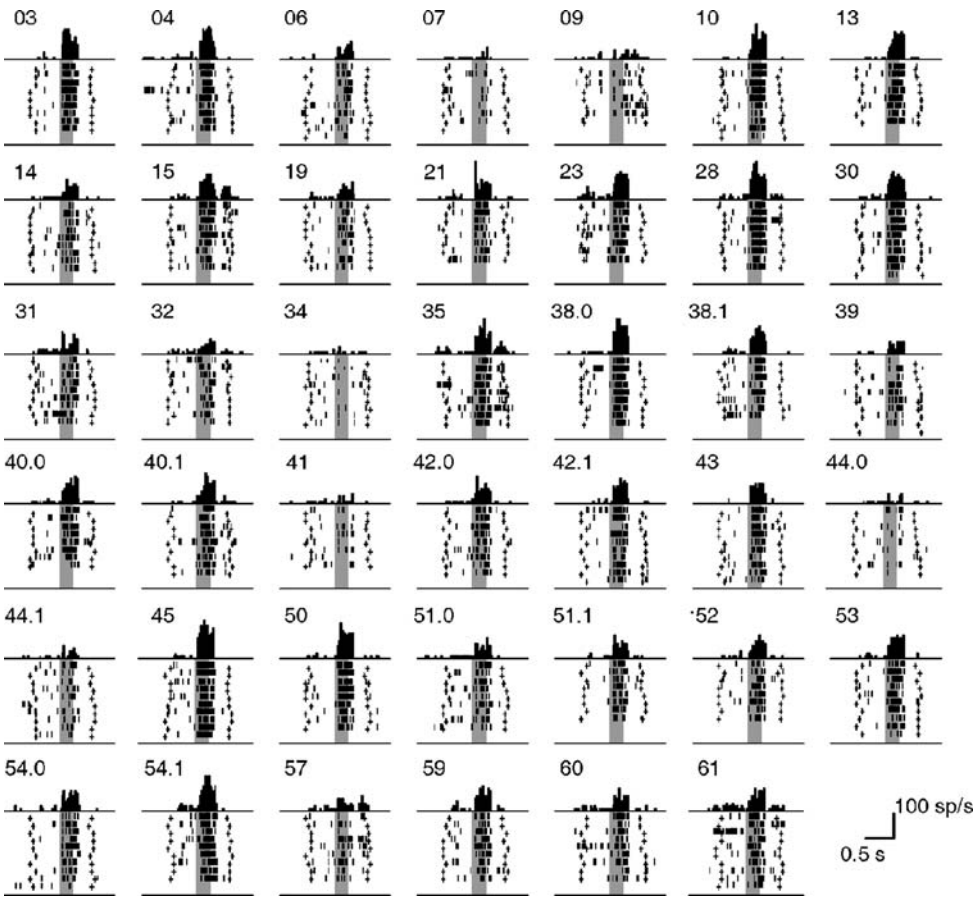

Fig. 2 Example of a V4 neuron selective to natural textures (TS neuron). Each *histogram* represents the responses to a given texture. Illumination angles L1 and L2 are pooled together. All trials are lined-up with the onset of the stimulus. The *gray bar* corresponds to the ON period of the stimulus (250 ms duration). The bin width of the histograms is 25 ms. The *number* in the upper-left corner of each PSTH indicates the texture patch number

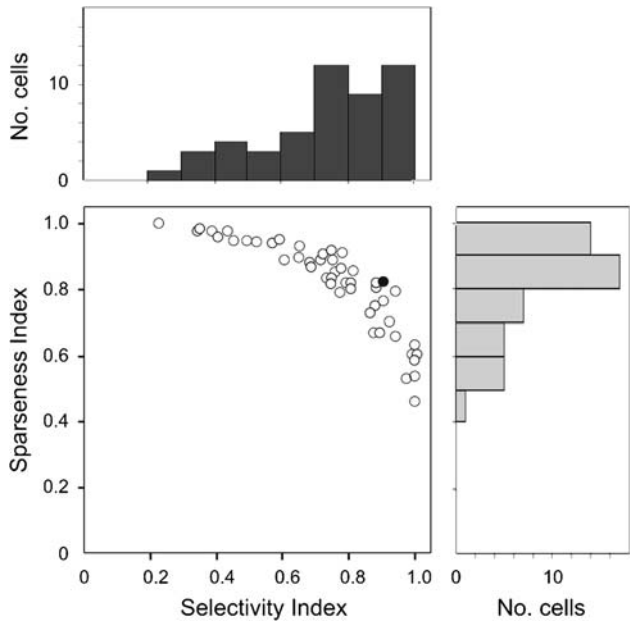

Fig. 3 TS neurons selectivity and sparseness for natural textures. SpI of 49 TS cells are plotted in a scatter plot (*open circles*) as a function of SI. The *black dot* represents the neuron shown in Fig. 2. Indices distributions are shown on the *top* (SI) and on the *right side* (SpI)

population response to one stimulus. A small linkage distance between two stimuli indicates similar responses. The tree splits into two distinct clusters (A and B) at its highest level. The A branch contains two clusters (A1 and A2) grouping, respectively, 26 and 20 stimuli. The B branch contains two clusters: B1 containing 20 stimuli and B2 containing 16 stimuli. Textures belonging to clusters A and B are respectively, heterogeneous and homogeneous as they differ in variance of luminance (mean variance = 0.027 and 0.008, respectively, ANOVA $P = 0.000007$). The second level contains clusters of textures that differ according to the mean luminance. Textures in A1 and B2 have a similar high range of luminance values (mean 0.49 and 0.48, respectively, $P = 0.986$) while textures in A2 and B1 have a similar range of lower luminance (mean 0.23 and 0.28, $P = 0.483$). Moreover, mean luminance in A1 and B2 significantly differs from that in A2 and B1 ($P = 9.04 \times 10^{-16}$). Second order parameters are not different across clusters ($P > 0.05$). Finally, there is a marked tendency to observe the same texture illuminated with L1 and L2 gathered in the same cluster: 35 pairs of L1–L2 textures out of 41 belong to the same cluster. The percentage of paired L1-L2 textures is 85% in A1, 90% in A2, 70% in B1, and 100% in B2. Within these clusters, the members of 15 pairs of stimuli were the closest neighbors.

We also performed a MDS analysis (see suppl. Fig. B). The stress value (0.13) and elbow on the scree plot (not shown) restricted the analysis to three dimensions. Residual sum of squares (RSQ) values indicated that this 3D configuration accounted for 89% of the variance (1D and 2D solutions accounted for 64 and 83%, respectively). In order to make a visual comparison, we reported the colors of the clusters identified in the hierarchical cluster analysis onto the MDS plots. Figure 6a shows that the four clusters (A1, A2, B1, and B2) are also visible in the MDS plots. Dimension one of the MDS is significantly correlated (Spearman rank order correlation) with SNR ($r = 0.738$, $P = 1 \times 10^{-6}$) and variance of luminance ($r = 0.639$, $P = 1 \times 10^{-6}$). Other parameters are not significantly represented in dimension one ($r < \pm 0.4$). Dimension two

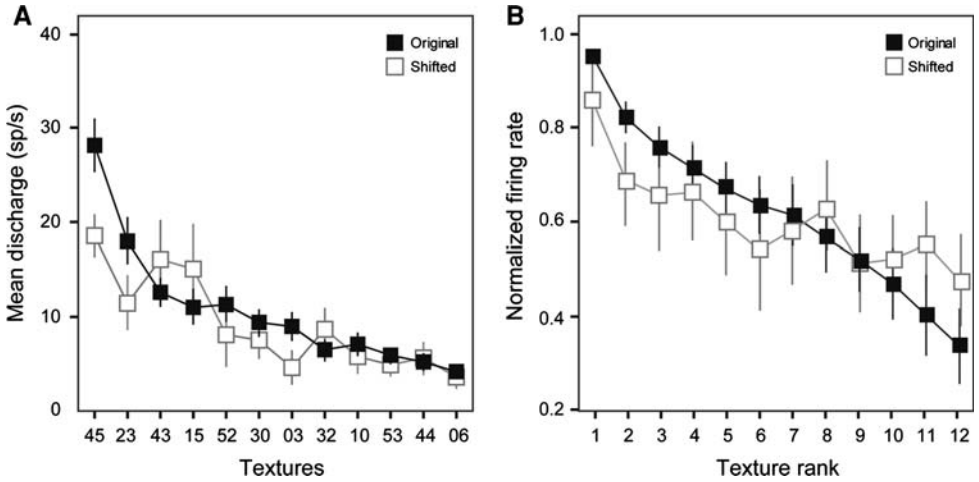

Fig. 4 Position test. **a** Example of a TS neuron preferring texture #45. The selectivity observed for the original position (*black dots*) is preserved when the stimulus position is shifted by 1° within the RF (*open dots*). Textures are ranked in descending order according to mean responses to original position. This rank is preserved to plot the corresponding responses to the shifted patches (*open dots*). **b** Rank order of the population ($n = 10$) tested with the shifted condition. The *black curve* illustrates the ranking of average normalized responses to patches in original position. Same rank for the corresponding responses to the shifted patches (*open dots*)

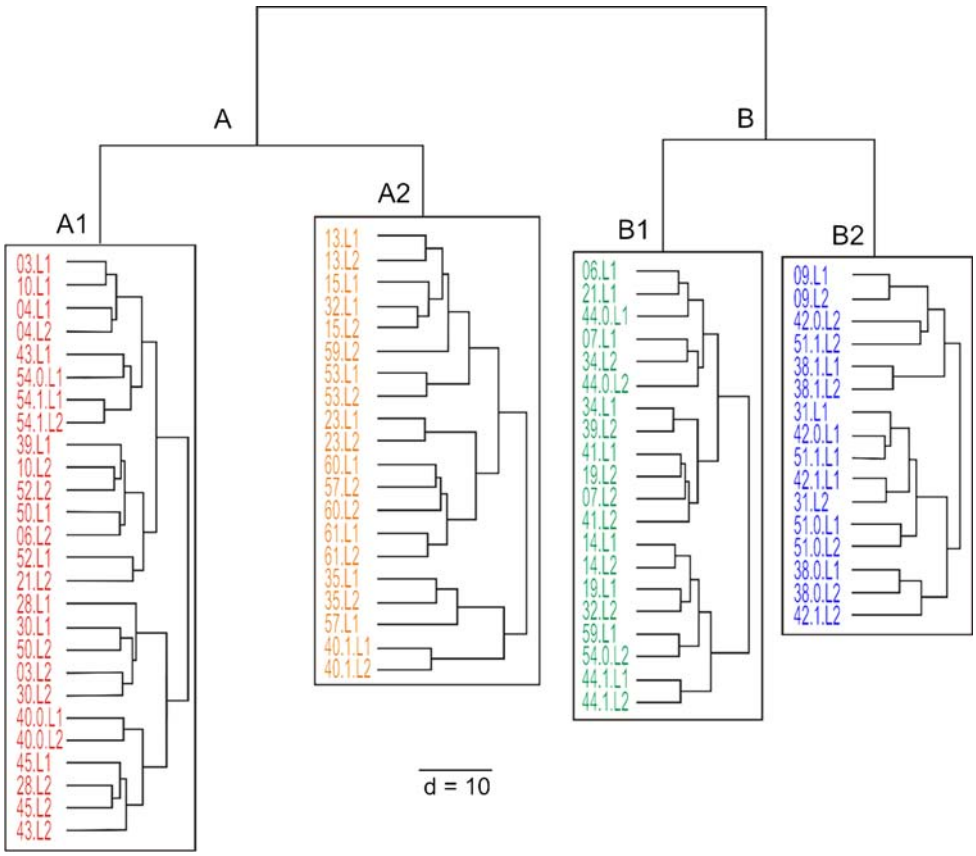

Fig. 5 Dendrogram from the hierarchical cluster analysis of normalized single unit responses to the 82 original stimuli (41 textures × 2 lighting angles). Each horizontal line represents one stimulus (texture# and IA—L1 or L2—are mentioned). Arbitrary colors correspond to clusters and will be used in subsequent figures. For the sake of readability, we represent clusters vertically. Linkage distances (*d*) are respected

exhibits a correlation with mean luminance only ($r = 0.670$, $P = 1 \times 10^{-6}$). Dimension three is correlated with maximum values of second order parameters (energy max: $r = 0.514$, $P = 1 \times 10^{-6}$; entropy max: $r = 0.482$, $P = 4.23 \times 10^{-6}$; homogeneity max: $r = 0.467$, $P = 9.69 \times 10^{-6}$).

Finally, we compared population response and energy of convolutions. Figure 6c and d show the MDS computed from filters output. The three dimensions solution explained 99% of the variance with a very low stress value (0.03). A Procrustes rotation test demonstrated that there is a significant correlation between the neuronal response space and the filters output space ($m^2 = 0.68$, $P < 0.001$).

Response modulation to phase scrambling

The 2D distribution of phases is crucial to determine the visual aspect of images (Tadmor and Tolhurst 1993). If V4 cells classify textures, as suggested from MDS analysis above, they should be sensitive to marked alterations in the distribution of Fourier phases of texture images. We examined the responses of 66 TS cells[1] to control stimuli that contain a random redistribution of the phases (set B, see methods). Although they share identical spatial frequency content, original and scrambled phase stimuli may have different visual aspect. Indeed, the 3D appearance due to micro-shading patterns disappears in scrambled phase stimuli (cf. Fig. 1b). The neuron shown in Fig. 7a robustly responds to texture #23 and not to its scrambled-phase version. The Kruskal–Wallis test shows that 42% ($n = 28$) of the TS cells have a significant different response according to the phase of the stimuli ($P < 0.05$). For the majority of these cells ($n = 27$), the response to original textures was stronger than to scrambled-phase textures. Figure 7b shows the mean TS population response according to the stimuli phase content. This plot shows that TS population response to original textures is stronger than the response to corresponding scrambled-phase patches. The significance of the effect is delayed with respect to the initial part of the response (150–300 ms, *t*-test, $P < 0.001$). The normalized TS population response ranked according to single cells preference to original textures shows that population response to the preferred original texture is much stronger than the response to the corresponding scrambled-phase patch, and that preference order is quite disrupted (see Fig. 7c). However, if the response to scrambled phase stimuli were completely independent from the response to original stimuli, the second curve should be flat. In fact, the gray curve shows that TS

[1] The proportions of TS cells recorded with sets A and B are not significantly different (chi^2, $P = 0.237$).

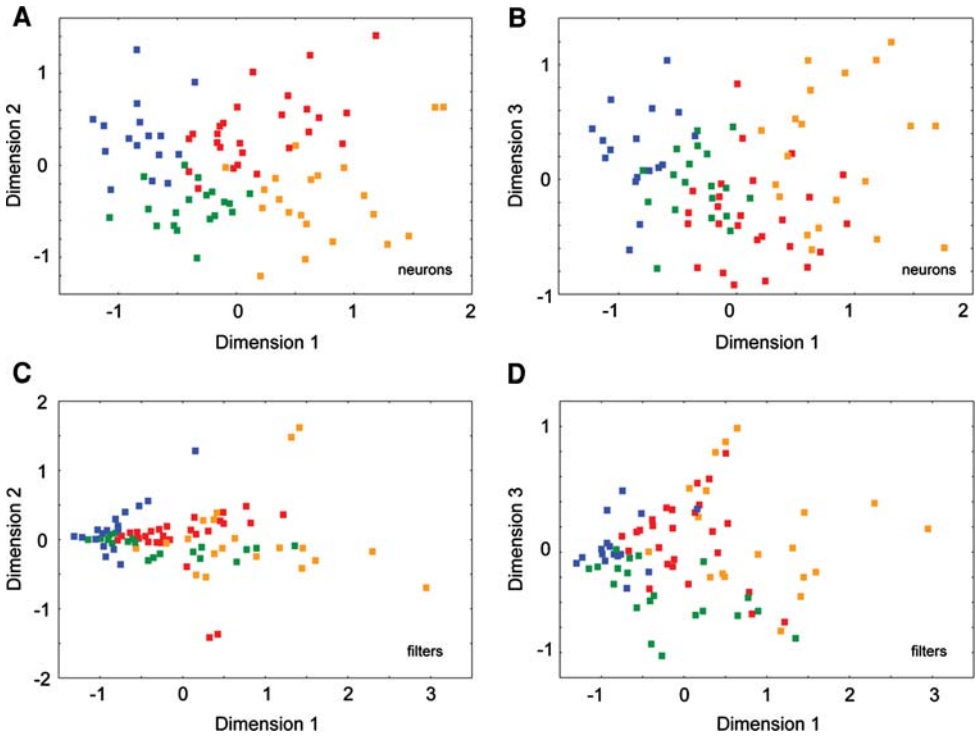

Fig. 6 MDS plot obtained from single cells responses (**a** dimensions 1 vs. 2, and **b** dimensions 1 vs. 3). The MDS has RSQ and stress values of 89% and 0.13.MDS plots obtained from filters output (**c** dimensions 1 vs. 2, and **d** dimensions 1 vs. 3). The MDS has RSQ and stress values of 99% and 0.03, respectively. Colors obtained from the clusters identified in the dendrogram (Fig. 5) are reported here

neurons rather prefer the scrambled phase stimulus corresponding to the preferred original stimulus (Spearman rank order correlation, $P = 1 \times 10^{-6}$; $r = 0.943$). These results show that TS cells do not merely encode phase spectrum of the texture images, though they are sensitive to alterations in the distribution of phases.

Response modulation to illumination angle

An efficient texture coding would allow recognition in a constantly varying visual environment, and hence should be invariant to lighting angle. Invariance to illumination angle (IA) was already suggested in Fig. 5 where L1 and L2 patches are very often close to each other in the dendrogram. To assess the invariance, we searched for an illumination angle invariant coding in the 66 TS cells recorded with set B. Figure 8a illustrates a typical IA independent neuron. This neuron is selective to texture #15 with a response invariant to IA (Kruskal–Wallis, $P = 0.359$). We computed a modulation index (MI) on the responses to the best texture only (see Vogels and Biederman 2002). For each cell, MI is defined as the subtraction of the minimal response (R_{min}, to the less preferred illumination angle) from the maximal response (R_{max}, preferred illumination angle), divided by the mean standard deviation (SD) of those responses. The formula of the index is given by:

$$\mathrm{MI} = 2 * (R_{max} - R_{min})/(\mathrm{SD}_{max} + \mathrm{SD}_{min})$$

Hence, MI of 0 indicates a complete invariance to lighting angle, while value of 1 indicates that the difference between the maximal and the minimal response is similar to the mean standard deviation. The neuron illustrated in Fig. 8a had a MI of 0.75 indicating a non-significant difference between responses to preferred and non preferred IA. The index distribution for the TS population is shown in Fig. 8b. The median of MI distribution is 0.99 indicating that modulation due to lighting angle is equivalent to the mean standard deviation of the single cells responses. Statistical significance of MI (Fig. 8b) shows that a vast majority of TS cells are invariant to IA (72%, $n = 48$; Kruskal–Wallis, $P > 0.05$) although a subpopulation of neurons is influenced by IA ($n = 16$, Kruskal–Wallis, $P < 0.05$). As a matter of comparison, we also computed a contrast ratio index ($[R_{max} - R_{min}]/[R_{max} + R_{min}]$). As expected from the MI distribution, the contrast ratio index shows that the modulation is modest (median = 0.2). We also quantified the effect of IA on heterogeneous (variance > 0.015) and homogeneous (variance < 0.01) textures respectively. We obtained two similar distributions (Mann–Whitney test, $P = 0.721$), indicating that IA effect is similar for TS neurons preferring heterogeneous vs. homogenous textures.

Figure 8c shows the rank order plot for TS population ($n = 66$). There is a tendency that when a given neuron prefers a texture illuminated with L1, the same texture illuminated with L2 and L3 is preferred. The Spearman's rank correlation confirmed this observation (correlation between L1 and L2: $r = 0.942$, $P < 0.001$; L1 and L3: $r = 0.923$, $P < 0.001$; and L2 and L3: $r = 0.848$, $P < 0.001$), which shows that texture preference does not depend on IA.

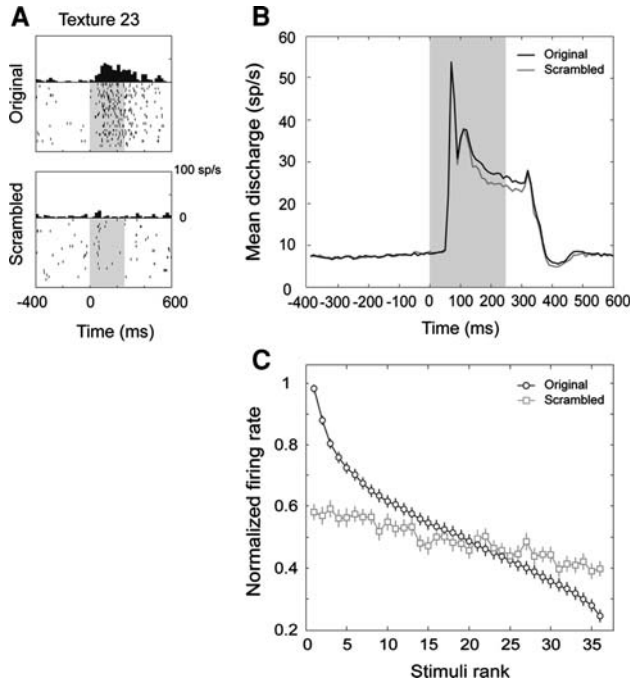

Fig. 7 Response modulation according to phase. **a** Example of a TS neuron. *Upper* and *lower quadrants* show responses recorded with the original- and scrambled-phase stimuli, respectively, (responses to the three lighting angles are pooled together). **b** Mean discharge of TS population to original versus scrambled phase textures. Bin width is 20 ms. The *gray bar* indicates the ON period of the stimuli. **c** Rank order plot of the TS population as a function of original-phase texture preference. The *black curve* illustrates the ranking of average normalized responses ($n = 66$). Rank 1 corresponds to the preferred texture (that could be different from cell to cell). The rank order is preserved to plot the responses to the scrambled-phase stimuli (*gray curve*)

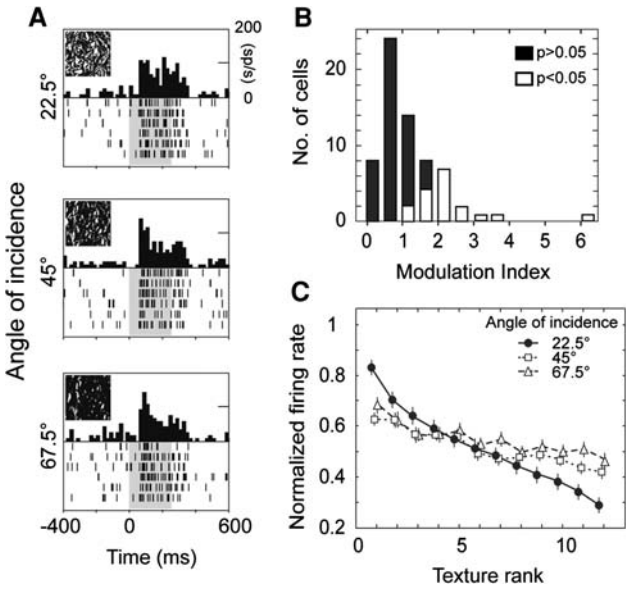

Fig. 8 Effect of illumination angle. **a** V4 neuron invariant to illumination angle (*IA*). Thumbnails of textures patches (#15) are displayed at the upper left corner of each PSTH. The bin width of the histogram is 25 ms. **b** Distribution of the modulation indices (*MI*) computed for the preferred texture of each TS cell ($n = 66$). The distribution of MI has a median value of 0.99. *Filled bars* indicate non-significant effect of IA (Kruskal–Wallis, $P > 0.05$). **c** Rank order plot as a function of IA. The mean normalized firing rates for the L1 stimuli (IA = 22.5, *solid dots*) are computed for each TS cell and ranked in descending order. Corresponding mean normalized firing rates to L2 (IA = 45, *empty squares*) and L3 (IA = 67.5, *empty triangles*) stimuli are plotted on the same graph. *Error bars* represent the standard error of the mean (*sem*)

Discussion

Selectivity for natural textures in V4

Many studies have shown that neurons selectivity becomes more and more complex along the hierarchy of cortical visual areas (i.e., Kobatake and Tanaka 1994). V4 neurons are selective to basic attributes like orientation, length, and width of bar stimuli, as well as orientation and spatial frequency of gratings (Schein and Desimone 1990). Subsequent studies have shown that V4 neurons are selective to spatial complex stimuli such as non-Cartesian gratings (Gallant et al. 1993, 1996), and contour curvatures (Pasupathy and Connor 1999, 2001). These complex stimuli may be considered as an alphabet-like set of descriptors of boundary elements. They may also be seen as individual elements of a textured surface (textons). Indeed, the fMRI study by Puce et al. (1996) in humans revealed responses to natural textures in the region of V4 (but it is difficult to ascertain that the V4 area that they localized is the homologous of the macaque area V4 identified in Denys et al. 2004). Moreover, Hanazawa and Komatsu (2001) showed that V4 neurons are selective to density and size of points mimicking a texture-like surface.

In agreement with these studies, our results indicate that many neurons (more than 97%) in area V4 are selective to natural textures. Along with the broad selectivity observed in the studies of Gallant et al. (1996) and Pasupathy and Connor (1999), we show that a majority of V4 neurons strongly respond to a wide range of natural textures. One could first question whether the selectivity to texture results from artefactual responses from a local cue such as the border of the stimuli or, for instance, the vein of the leaf falling in a RF singularity. One pervasive aspect of textures is that a given pattern is more or less regularly repeated over space. We would expect selective neuronal responses to be invariant to shifts of a textural stimulus in the RF. We recorded a number of neurons that kept their selectivity when the texture is shifted in the RF. This result suggests that V4 cells selectivity for textures does not depend on texture position in the RF. It goes along with that of Gallant et al. (1996) who observed position invariant responses to Cartesian and non-Cartesian stimuli.

Alone, these response patterns do not necessarily mean that V4 neurons are indeed selective to natural textures. TS cells could be tuned to phase or orientation of spatial frequency components (see Desimone and Schein 1987), and may exhibit a similar response to several textures because they contain identical spatial frequency components. Since both Fourier phase and amplitude spectra are essential to determine the visual aspect of an image (Tadmor and Tolhurst 1993), we designed control stimuli with a random redistribution of Fourier phases that disrupted their visual aspect. Original images contain important 3D cues such as microshading patterns that disappear on scrambled-phase controls. TS population strongly responded to original images compared to control ones, and did not only reflect the phase spectrum in texture images (which is identical in original and control patches). Moreover, we showed that neuronal selectivity remains stable in spite of stimuli's shifting into the RF and we provided further evidence from the responses of V4 neurons to the same texture illuminated with different angles. The illumination angle dramatically changes the micro-shading patterns, and consequently, the Fourier spectra of the images; despite this, we found that most of TS cells were not significantly modulated by IA. Altogether, these observations suggest that TS selectivity cannot be only explained by the selectivity to spatial frequency.

The coding of angle and direction of illumination

We showed that most V4 neurons response to their preferred texture is invariant to illumination angles along one direction. This is first reflected in the MI computed from single cells response and the fact that the texture preference is kept for different illumination angles. Consistent with it, the V4 neuronal population had a marked tendency to cluster pairs of textures illuminated with different angles. We also showed that a small proportion only of TS cells are correlated to first or second order parameters, hence a single image parameter invariant to IA cannot explain TS cells invariance to IA. Most, if not all, of the papers dealing with illumination focus on the influence of illumination direction (and not angle) on object recognition, and lead to controversial results. On the one hand, the structural theory sets that object recognition does not vary according to illumination direction because it relies on stored parts of objects that do not contain this information (Biederman 1987; Marr and Nishihara 1978). Psychophysical studies reinforce this theory as human subjects can easily recognize faces and objects when illumination direction varies (Nederhouser et al. 2001). Electrophysiological studies (Vogels and Biederman 2002) show that invariance to illumination direction may be implemented in IT cortex and additional studies are necessary to determine whether illumination invariance would be present in earlier stages of the visual system (V1 and V2). On the other hand, image-based theory proposes that illumination direction is encoded in internal face and object representations (Poggio and Edelman 1990; Ullman 1989). The theory is corroborated by psychophysical data showing that recognition of faces and objects can vary with illumination (Braje 2003; Gauthier and Tarr 1997; Troje and Bulthoff 1998). In a single cell recording study, Hanazawa and Komatsu (2001) showed that V4 cells tuned to illumination direction were preferentially tuned along the vertical axis only. Our results show that V4 single cells responses to preferred natural textures are invariant to illumination angle. However, they are not at odds with Hanazawa and Komatsu's results. We used three angles of incidence but only one direction of illumination coming from the right side; therefore the two results are not contradictory. Indeed, separate subpopulations of V4 cells may exhibit preferred direction of illumination, but with invariance to illumination angle along the preferred direction. Such mechanisms would fit the image-based theory.

Comparison with computational studies

Recent computational studies of texture classification or recognition tend to extend models to naturalistic viewing conditions by including surface normal variations and, consequently, the effect of illumination direction. Leung and Malik (2001) define a "3D dictionary of textons" (by reference to Julesz's 2D textons) based on responses of a bank of filters. Their algorithm successfully (87%) recognizes learned textures under novel lighting and viewing conditions. Recently, several groups have developed sophisticated algorithms (always based on filter responses) that are more efficient than previous ones in that they also classify patches of textures under new lighting or viewing conditions (Hayman et al. 2004; Varma and Zisserman 2002). Our results show that V4 TS cells responses are comparable to the output of a pyramid of Gabor filters, and are not modulated by lighting angle, which is in agreement with such models.

Individual components of a 2D Gabor filters bank rather match RFs of V1 simple cells than that of V4 cells (Jones and Palmer 1987). Consequently, the question is unanswered as to how texture processing follows hierarchical steps from V1 to V4. Is there just a difference in scaling between areas performing the same operations (V4 RFs pooling many V1 and V2 RFs)? V1 neurons can extract high order spatial correlations from visual textures and could detect redundancy in retinal images (Purpura et al. 1994). Lamme (1995) showed that texture segregation in V1 allows extraction of figure from ground. Hegde and VanEssen (2007) recently questioned the increasing

complexity of stimulus encoding from V1, V2, to V4 using contour and grating stimuli of low or intermediate complexity. This study shows that shape selectivity is far from being hierarchically built from V1 to V4 and that stimuli defined by internal patterns like gratings were substantially more effective in activating V4 neurons than contour stimuli. The difference was quantitatively much smaller in V2. In V1, they observed a "substantial intermixing" between grating and contour stimuli. Moreover, Pollen et al. (2002) demonstrate that V4 cells have a RF divided into several subfields with similar properties such as selectivity of orientation or same spatial frequency. Consequently, texture could be encoded on the basis of repetitive textons; each subfield of the RF corresponding to bottom-up inputs from V1 and V2.

Conclusion

Natural textures are complex stimuli that cannot be simply described by a simple set of parameters (Haralick et al. 1973). One can then wonder if V4 responses really reflect textural processing per se or simply reflect differential responses to an arbitrary complex set of stimuli. Humans qualify or classify visual textures using many terms related to the visual aspect of their material properties (Bhushan et al. 1997; Rao and Lohse 1996). Many terms related to the tactile aspect are also employed for somatosensory classification (Picard et al. 2003). All these studies, using MDS, showed that several dimensions are required to explain the classification data. Our multidimensional analyses revealed that three dimensions were sufficient to decipher V4 population response. Two dimensions were identified as first order parameters in the images: mean luminance in one dimension, and SNR or variance of luminance (both describe heterogeneity) in the second dimension. The third dimension was correlated with second order parameters (mainly energy) that describe the heterogeneity at the second order level. Functionally, these three dimensions could be interpreted as a counterpart of the perceptual dimensions required for classifying textures.

We do not suggest that V4 is the only locus of texture perception, but rather that it is included in a network integrating texture and other cues for object recognition. Furthermore we do not have evidence that V4 is a locus of conscious texture perception. For example, it would be interesting to know if the invariance to illumination angles is even broader in an active recognition task, in which a monkey has been trained to recognize textures under various lighting conditions. This generalization of texture recognition to "difficult" illumination is plausible in V4 since neurons of this area have been demonstrated to be prone to perceptual learning (Rainer et al. 2004).

Acknowledgments This work was supported by grants from the Information Society Technologies (INSIGHT2+, #2000 29688, Neuronal basis of coding of 3D shape and material properties for recognition) and the Fondation de France. The authors thank Dr. Rufin Vogels for useful discussions; Renaud Lestringant for advice in data processing; Drs. Jean-Michel Hupé, Rufin VanRullen, Lionel Nowak, James Bisley, and Pascal Barone for comments on various drafts of the manuscript; and Franck Lefèvre and Sebastien Aragones for animal care.

References

Biederman I (1987) Recognition-by-components: a theory of human image understanding. Psychol Rev 94:115–147

Bhushan N, Rao AR, Lohse GL (1997) The texture lexicon: understanding the categorization of visual texture terms and their relationship to texture images. Cogn Sci 21:219–246

Braje WL (2003) Illumination encoding in face recognition: effect of position shift. J Vis 3:161–170

De Weerd P, Desimone R, Ungerleider LG (1996) Cue-dependent deficits in grating orientation discrimination after V4 lesions in macaques. Vis Neurosci 13:529–538

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

Desimone R, Schein SJ (1987) Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. J Neurophysiol 57:835–868

Gallant JL, Braun J, Van Essen DC (1993) Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. Science 259:100–103

Gallant JL, Connor CE, Rakshit S, Lewis JW, Van Essen DC (1996) Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey. J Neurophysiol 76:2718–2739

Gattass R, Sousa AP, Gross CG (1988) Visuotopic organization and extent of V3 and V4 of the macaque. J Neurosci 8:1831–1845

Gauthier I, Tarr MJ (1997) Orientation priming of novel shapes in the context of viewpoint- dependent recognition. Perception 26:51–73

Girard P, Lomber SG, Bullier J (2002) Shape discrimination deficits during reversible deactivation of area V4 in the macaque monkey. Cereb Cortex 12:1146–1156

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

Haralick RM, Shanmugan K, Dinstien I (1973) Textural features for image classification. IEEE Trans Man Cybern 3:610–621

Hayman E, Caputo B, Fritz M, Eklundh J (2004) On the significance of real-world conditions for material classification. In: Proceedings of 8th conference on computer vision, Czech Republic

Hegde J, Van Essen D (2007) A comparative study of shape representation in macaque visual areas V2 and V4. Cereb Cortex 23:749–763

Humphrey GK, Goodale MA, Jakobson LS, Servos P (1994) The role of surface information in object recognition: studies of a visual form agnosic and normal subjects. Perception 23:1457–1481

Jackson DA (1995) PROTEST: a PROcrustean randomization TEST of community environment concordance. Ecoscience 2:297–303

Jones JP, Palmer LA (1987) The two-dimensional spatial structure of simple receptive fields in cat striate cortex. J Neurophysiol 58:1187–1211

Julesz B, Schumer RA (1981) Early visual perception. Annu Rev Psychol 32:575–627

Kobatake E, Tanaka K (1994) Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. J Neurophysiol 71:856–867

Kruskal JB, Wish M (1978) Multidimensional scaling. In: Uslaner EM (ed) Sage Publications, Beverly hills

Lamme VA (1995) The neurophysiology of figure-ground segregation in primary visual cortex. J Neurosci 15:1605–1615

Leung T, Malik J (2001) Representing and recognizing the visual appearance of materials using three-dimensional textons. Int J Comp Vis 43:29–44

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

Marr D, Nishihara HK (1978) Representation and recognition of the spatial organization of three-dimensional shapes. Proc R Soc Lond B 200:269–294

Merigan WH (1996) Basic visual capacities and shape discrimination after lesions of extrastriate area V4 in macaques. Vis Neurosci 13:51–60

Nederhouser M, Mangin MC, Biederman I, Subramaniam S, Vogels R (2001) Is object recognition invariant to direction of illumination and direction of contrast [Poster]. In: 8th Joint Symposium on Neural Computation, La Jolla, CA

Pasupathy A, Connor CE (1999) Responses to contour features in macaque area V4. J Neurophysiol 82:2490–2502

Pasupathy A, Connor CE (2001) Shape representation in area V4: position-specific tuning for boundary conformation. J Neurophysiol 86:2505–2519

Picard D, Dacremont C, Valentin D, Giboreau A (2003) Perceptual dimensions of tactiles textures. Acta Psychol 1114:164–185

Poggio T, Edelman S (1990) A network that learns to recognize three-dimensional objects [see comments]. Nature 343:263–266

Pollen DA, Przybyszewski AW, Rubin MA, Foote W (2002) Spatial receptive field organization of macaque V4 neurons. Cereb Cortex 12:601–616

Price CJ, Humphreys GW (1989) The effects of surface detail on object categorization and naming. Q J Exp Psychol A 41:797–827

Puce A, Allison T, Asgari M, Gore JC, McCarthy G (1996) Differential sensitivity of human visual cortex to faces, letter-strings, and textures: a functional magnetic resonance imaging study. J Neurosci 16:5205–5215

Purpura KP, Victor JD, Katz E (1994) Striate cortex extracts higher-order spatial correlations from visual textures. Proc Natl Acad Sci USA 91:8482–8486

Rainer G, Lee HK, Logothetis NK (2004) The effect of learning on the function of monkey extrastriate visual cortex. PLoS Biol 2:275–283

Rao AR, Lohse GL (1996) Toward a texture naming system: Identifying relevant dimensions of texture. Vision Res 36:1649–1669

Rolls ET, Tovee MJ (1995) Sparseness of the neural representation of stimuli in the primate temporal visual cortex. J Neurophysiol 73:713–726

Rossion B, Pourtois G (2004) Revisiting Snodgrass and Vanderwart's object pictorial set: the role of surface detail in basic-level object recognition. Perception 33:217–236

Sary G, Chabaide Z, Tompa T, Kovacs G, Köteles K, Boda K, Raduly L, Benedeck G (2004) Relationship between stimulus complexity and neuronal activity in the inferotemporal cortex of the macaque monkey. Cogn Brain Res 22:1–12

Schein SJ, Desimone R (1990) Spectral properties of V4 neurons in the macaque. J Neurosci 10:3369–3389

Schiller PH (1993) The effects of V4 and middle temporal (MT) area lesions on visual performance in the rhesus monkey. Vis Neurosci 10:717–746

Tadmor Y, Tolhurst DJ (1993) Both the phase and the amplitude spectrum may determine the appearance of natural images. Vision Res 33:141–145

Troje NF, Bulthoff HH (1998) How is bilateral symmetry of human faces used for recognition of novel views? Vision Res 38:79–89

Ullman S (1989) Aligning pictorial description; an approach to object recognition. Cognition 32:193–254

Varma M, Zisserman A (2002) Statistical approaches to material classification. Paper presented at the Proc Ind Conf Comp Vis Graph Im Proc

Varma M, Zisserman A (2003) Texture classification: are filter banks necessary? Paper presented at the IEEE Conf Comp Vis Patt Recog

Vogels R, Biederman I (2002) Effects of illumination intensity and direction on object coding in macaque inferior temporal cortex. Cereb Cortex 12:756–766

von der Heydt R, Peterhans E, Baumgartner G (1984) Illusory contours and cortical neuron responses. Science 224:1260–1262

Young MP, Yamane S (1992) Sparse population coding of faces in the inferotemporal cortex. Science 256:1327–1331