

Stimulus features coded by single neurons of a macaque body category selective patch

Ivo D. Popivanov^{a,b}, Philippe G. Schyns^c, and Rufin Vogels^{a,1}

^aLaboratorium voor Neuro- en Psychofysiologie, KU Leuven Medical School, 3000 Leuven, Belgium; ^bDepartment of Cognitive Science and Psychology, New Bulgarian University, 1618 Sofia, Bulgaria; and ^cInstitute of Neuroscience and Psychology, University of Glasgow, Glasgow G12 8QQ, United Kingdom

Edited by Terrence J. Sejnowski, Salk Institute for Biological Studies, La Jolla, CA, and approved March 18, 2016 (received for review October 14, 2015)

Body category-selective regions of the primate temporal cortex respond to images of bodies, but it is unclear which fragments of such images drive single neurons' responses in these regions. Here we applied the Bubbles technique to the responses of single macaque middle superior temporal sulcus (midSTS) body patch neurons to reveal the image fragments the neurons respond to. We found that local image fragments such as extremities (limbs), curved boundaries, and parts of the torso drove the large majority of neurons. Bubbles revealed the whole body in only a few neurons. Neurons coded the features in a manner that was tolerant to translation and scale changes. Most image fragments were excitatory but for a few neurons both inhibitory and excitatory fragments (opponent coding) were present in the same image. The fragments we reveal here in the body patch with Bubbles differ from those suggested in previous studies of face-selective neurons in face patches. Together, our data indicate that the majority of body patch neurons respond to local image fragments that occur frequently, but not exclusively, in bodies, with a coding that is tolerant to translation and scale. Overall, the data suggest that the body category selectivity of the midSTS body patch depends more on the feature statistics of bodies (e.g., extensions occur more frequently in bodies) than on semantics (bodies as an abstract category).

body perception | body selective area | feature selectivity | macaque | inferior temporal cortex

The body category-selective regions in the human occipito-temporal cortex are defined as those that respond to images of bodies (1–8). We previously identified two bilateral regions in the macaque inferotemporal cortex that respond stronger to monkey, human, and animal bodies in comparison with other stimuli, including faces (6). Subsequent single-unit recordings in the posterior body patch [i.e., the middle superior temporal sulcus (midSTS) body patch] demonstrated that indeed the average spiking activity of the neuron population was greater to images of bodies compared with other objects. However, the responses of single neurons showed a strong selectivity for particular body—and sometimes nonbody—images (7). However, it is still unknown what particular stimulus features single body patch neurons respond to. Moreover, we still do not know how those neurons code information about different animate and inanimate stimuli.

The Bubbles technique (9), in which parts of the image of an object are sampled by trial-unique randomly positioned Gaussian apertures, has been used successfully in many psychophysical studies to reveal the features critical for certain perceptual tasks such as face identification, gender discrimination, emotional discrimination, and so forth (e.g., refs. 9–13). Although this technique has been used in neuroimaging [functional MRI (fMRI), EEG, magnetoencephalography (MEG), and electrocorticography] studies (11, 12, 14), it has rarely been exploited in single-unit studies (15, 16), and this only for face stimuli.

Here, we used the Bubbles technique to reveal the image fragments that drive single midSTS body patch neurons. Bubbles provides an unbiased method for sampling the images with the advantage that it requires no prior specification of stimulus features to which the neurons are supposed to be selective. With fMRI, we

first defined the midSTS body patch in two monkeys. Then, in this identified body patch we recorded the spiking activity of well-isolated single neurons in response to 100 images of various categories. Based on the spiking activity to the 100 images, we selected for each neuron a response-eliciting image. Then, we sampled the selected image at five different spatial scales with randomly positioned Gaussian apertures and recorded the responses of the neuron to a large number of these trial-unique Bubbles stimuli. Following the experiment, we applied reverse correlation to relate the excitatory and inhibitory neural responses to particular image fragments.

Furthermore, we assessed whether the revealed image fragments tolerated changes in spatial location and size of the Bubbles stimuli or instead reflected spatially localized image regions. We showed before that many midSTS body patch neurons respond to silhouettes of bodies (17). Silhouettes isolate shape contours, removing texture and shading information. Thus, in a subset of neurons, we applied Bubbles to a silhouette version of the selected image.

Results

Using an fMRI block design localizer (contrast: monkey bodies – objects) in two monkeys, we isolated the midSTS body patch—the posterior one of the two body patches in the macaque inferior temporal cortex (6, 7). In the localized body patch, we searched for single neurons that responded to any of 100 images depicting monkey and human bodies, human and monkey faces, four-legged mammals, birds, fruits, sculptures, and manmade objects. Forty percent of these search stimuli contained bodies (human or monkey bodies, mammals, and birds). For each responsive neuron, we selected its preferred image (i.e., the image that elicited the greatest firing rate). Next, we tested each responsive neuron with Bubbles

Significance

A fundamental and understudied question is that of the higher-level features that actually drive neuronal responses to complex images. Although previous single-unit studies developed our knowledge of the features driving face-selective regions, little is known about the feature selectivity of neurons in body-selective regions. Using a reverse correlation technique, we reveal for the first time to our knowledge the image fragments coded by single neurons of a body-selective region defined by functional imaging. We found that the neurons respond to local fragments, such as extremities (e.g. leg fragments) and curved boundaries (e.g. shoulder), tolerating position and scale changes, with evidence of opponent body coding in a few neurons. Thus, our data offer new insights on how the brain codes higher-level features.

Author contributions: I.D.P., P.G.S., and R.V. designed research; I.D.P. performed research; P.G.S. contributed new reagents/analytic tools; I.D.P., P.G.S., and R.V. analyzed data; and I.D.P., P.G.S., and R.V. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

¹To whom correspondence should be addressed. Email: rufin.vogels@med.kuleuven.be.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1520371113/-DCSupplemental.

stimuli derived from its selected image—either the original textured and shaded image, or its black silhouette version (*Methods*).

To implement Bubbles (9, 18) we first decomposed the image into five spatial frequency (SF) bands of one octave each with peak SF ranging from 11 to 0.7 cycles per degree (Fig. 1A). Within each band, circular image fragments (information “bubbles”) were sampled with randomly located Gaussian windows (with Gaussian

diameter and number of bubbles scaled according to the peak SF of the band). To produce a trial-unique Bubble stimulus, we summed the Gaussian information samples across the five SF bands (Fig. 1A). Note that this procedure presents image features at different spatial scales, acknowledging that early visual cortex applies a decomposition of the image across different SF bands (19). For each neuron, we generated a large number (>800) of such Bubble

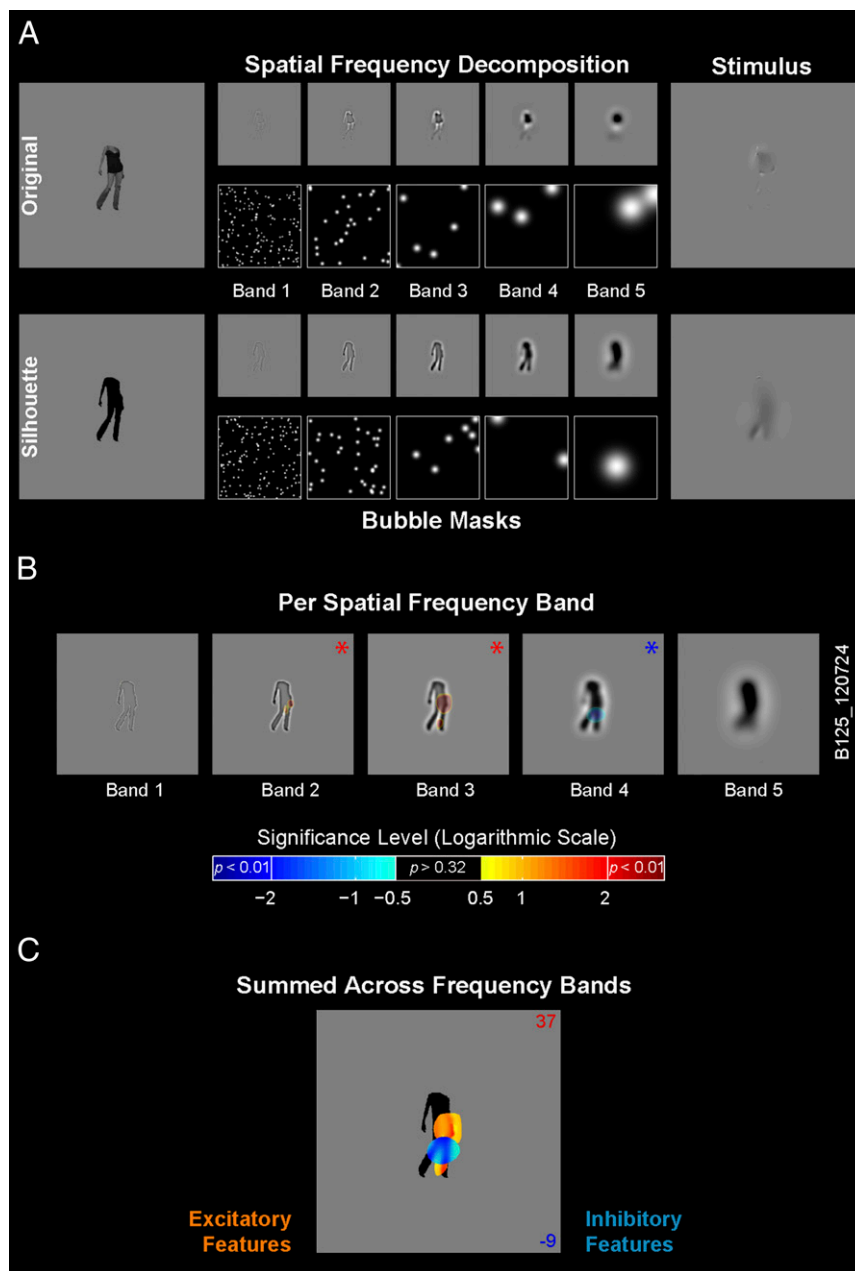


Fig. 1. Bubbles: procedure and single neuron example. (A) Procedure for original and silhouette stimuli. We decomposed the original and silhouette image into five nonoverlapping SF bands (bands 1–5) of one-octave bandwidth. Gaussian apertures (bubbles) positioned at random spatial locations within each band sampled different fragments of the spatially filtered image. We scaled bubble size and number across bands. The stimulus comprised the image fragments summed across SF bands. For display purposes in this figure (but not in the actual stimuli), we adjusted the contrasts for each band. (B) Bubbles results from one example neuron tested with the silhouette of a human body, plotted per SF band. Classification images [signed significance level of difference scores (DS), logarithmic scale] are overlaid at each SF band. The sign of the logarithm corresponds to the sign of DS (positive excitatory fragment indicated in red; negative inhibitory fragment indicated in blue). Red and blue asterisks denote significant positive and negative DS , respectively, in this band ($P < 0.01$). DS values with $P > 0.32$ are not displayed. (C) Bubble results for the same example neuron where the band-passed image fragments that included a significant DS in a band were summed across SF bands and are shown on top of the silhouette image. The color of the blobs corresponds to the sign of DS (positive excitatory fragment indicated in orange; negative inhibitory fragment indicated in blue). The two numbers in the right top and bottom corners indicate the mean net firing rates (in spikes per s) for the high (16% top; red) and low (16% bottom; blue) response trials, respectively.

Table 1. Number of neurons with significant features with Bubbles

Stimulus version	No. of tested neurons	With significant <i>DS</i>
Original	56	44
Silhouette	49	37
Original and silhouette	105	81

The second column shows the total number of tested neurons with original and silhouette versions of the stimuli, and the third column presents the number of neurons in which there was significant *DS* ($P < 0.01$) in at least one of the SF bands.

stimuli to broadly sample the visual information present in the selected image without bias.

To measure single neuron responses to the information samples, first we determined an effective location defined based on mapping the receptive field (RF) of the neuron (*Methods*). Then, at this effective location, we stimulated the neuron with up to 1,600 Bubble stimuli visually presented in a random sequential order. Following the experiment, for each SF band we reverse correlated the baseline-subtracted net firing rate of the neuron to the Bubbles stimuli (*Methods*). This procedure resulted in a difference score (*DS*) at each image pixel of each SF band. A positive *DS* indicates that the neuron's net firing rate was greater when the bubbles revealed the image pixel in question than when they did not (that is, this image fragment excited the neuron). A negative *DS* indicates a suppressive influence of the revealed image fragment on the response of the neuron. Based on a permutation test, we computed the significance of each pixel *DS* value (*Methods*).

The Large Majority of Body Patch Neurons Respond to Local Body Fragments. Fig. 1 *B* and *C* illustrate the statistically significant image fragments revealed by the Bubbles procedure per SF band (Fig. 1*B*) and across SF bands (Fig. 1*C*) for one example neuron tested with Bubbles stimuli derived from the silhouette version of its preferred image. Red and blue colors depict the image fragments with significant positive and negative *DS*, respectively. This single neuron responded preferentially to the image of a human body. Bubbles further attributed this response to significant excitatory image fragments (red blobs) representing a part of the torso around the waist as well as a small blob around the right foot. There was also a significant inhibitory component (blue blob) related to the hips and the upper part of the leg.

In total, we applied Bubbles in 105 single body patch neurons of two monkeys (85 neurons in monkey E). We tested 56 of these neurons with Bubbles stimuli computed from their preferred image, whereas we tested the other 49 neurons with Bubbles derived from the silhouette version of their preferred image. The median body selectivity index (BSI) (*Methods*) was 0.47 ($n = 105$), a number somewhat larger than that reported in our previous study [0.38 (7)]. Fifty-seven percent of the neurons had a BSI >0.33 , a conventional but arbitrary criterion to define category selectivity, whereas 86% has a BSI larger than zero, corresponding to a greater response to bodies compared with objects. Because we aimed to examine the feature selectivity of body patch neurons in general, we examined all neurons irrespective of their BSI. Dividing neurons according to their BSI (Tables S1–S5) did not reveal a significant difference between these groups, likely due to the small number of neurons that responded only weakly to bodies. Indeed, the preferred image of the large majority of neurons (92%) was a human body, a monkey body, another four-legged mammal, or a bird. The Bubbles procedure revealed statistically significant features in the large majority (77%) of the tested neurons and this both for neurons tested with the original shaded and textured image versions (79%) or their silhouettes (76%; Table 1). For the 81 neurons with significant features, the median net firing rate averaged across the 16% top and 16% bottom Bubble stimuli that

we used to compute *DS* was 33.8 spikes per s (quartiles: 20.9 and 51.6) and -8.3 spikes per s (quartiles: -15.4 and -3.8), respectively. The top Bubbles stimuli produced sizable responses that were statistically indistinguishable from those obtained with the original, full images in the independent search test (38.3 spikes per s; $P = 0.25$, Wilcoxon matched pairs test; $n = 81$).

Across the population of neurons, significant features were revealed at all SF bands but with a strong bias for the middle bands (Table 2). This bias suggests that the neurons are most sensitive to fragments that correspond to identifiable body (or object) parts. Indeed, the low SF band corresponds to large blobs in which body parts are poorly identifiable. The fewer significant fragments revealed at the highest SF band may be related to the lower energy and/or the small bubble size in this band. Bubbles revealed both excitatory and inhibitory features, with the large majority of the significant features being excitatory (175/186, pooled across SF bands; see Table 2 for a split per SF band). For both the original images and their silhouettes, we found that eight neurons (10% of the cells with significant features) coded different fragments of the preferred stimulus that either excited or inhibited responses (e.g., Fig. 1*C*), implying opponent feature coding. Such opponent coding was present within the same SF band or more often across different bands.

We used an “inverse mapping analysis” (*Methods*) to estimate the net response produced by the significant feature. For the 79 neurons with excitatory features, the median net firing rate to the Bubble stimuli revealing the feature was 20.3 spikes per s (quartiles: 11.3 and 30.3), which for each neuron was greater than the response to the Bubbles stimuli that did not reveal the feature (median: 8.8 spikes per s; $P < 0.00001$; Wilcoxon matched pairs test). For the inhibitory features, the median net response to the Bubble stimuli revealing the feature was 4.9 spikes per s ($n = 8$; quartiles: 0 and 10.6), which was for each neuron lower than the response in those neurons to the Bubble stimuli without the feature (12.3 spikes per s; $P < 0.005$). The nonnegative median net response to the inhibitory features can be explained by the likely possibility that the other Bubbles in some of those stimuli revealed other parts of the body/object that produced (weak) responses in the neuron.

Fig. 2 summarizes the variety of image fragments that the Bubbles reverse correlation procedure revealed. For illustration, we summed the fragments of those bands that showed a significant feature ($P < 0.01$) and superimposed the summed fragments on top of the original image, or on top of the silhouette version when the neuron was tested with the silhouettes. Bubbles revealed fragments related to the limbs or extremities, relatively small curved boundary segments of the torso or the head, or a larger part of the torso or even the whole body. We classified the fragments as belonging to one of six classes (*Methods*): extremities (as in Fig. 2*A*), curved segments (Fig. 2*B*), parts of the torso (Fig. 2*C*), torso combined with an extremity (Fig. 2*D*), the full body (Fig. 2*E*), or a sixth class of features that did not fall in any of the other classes. Table 3 shows the number of neurons with excitatory or inhibitory fragments for each of these six classes. The total number of fragments (100) in Table 3 is greater than the number of neurons with a significant feature because Bubbles revealed more than one

Table 2. Number of neurons with significant excitatory or inhibitory features in their preferred image per SF band

Stimulus version	Band 1		Band 2		Band 3		Band 4		Band 5	
	+	-	+	-	+	-	+	-	+	-
Original	7	1	24	1	27	1	23	0	11	2
Silhouette	7	2	14	1	21	0	23	2	18	1
Original and silhouette	14	3	38	2	48	1	46	2	29	3

Significant excitatory and inhibitory features are symbolized by + and -, respectively. The total number of neurons tested are shown in Table 1.

fragment in 18 neurons. Note that these multiple fragments can even be coded in an opponent manner because, as described above, eight neurons had both excitatory and inhibitory fragments in their receptive fields. We found that the majority of neurons responded to features that are prominent in bodies such as extensions (limbs), curved contours along part of the body, and parts of the torso. We computed the amount of overlap between the revealed fragment (defined using a lenient $P < 0.32$ threshold; *Methods*) and the body or object image as a percentage of the area of the full image. The median overlap was 46%, with only 16 neurons (20%) having a fragment larger than 80% of the image area. Fourteen of these neurons were recorded with silhouette images. Indeed, the median overlap was greater for the silhouettes (58%) than for the original images (34%; $P =$

0.007, Mann–Whitney U test) Note that because we used a lenient threshold of $P < 0.32$ to define the fragment, the degree of overlap is likely to be an upper bound of the relative size of the feature the neuron responds to. Thus, the large majority of midSTS neurons responded to parts of a whole body.

We defined a fragment as corresponding to the whole body when the overlap between revealed fragment and body was at least 90%. With that criterion, Bubbles revealed the whole body in nine neurons (Table 3 and Fig. 2E), which corresponds to only 11% of the neurons with significant features. Interestingly, five of these images were silhouettes of birds, three were monkey bodies (two silhouettes and one original image), and one was an original image of a human body. Simulations with model neurons responding to the full body template suggested that the Bubbles procedure was

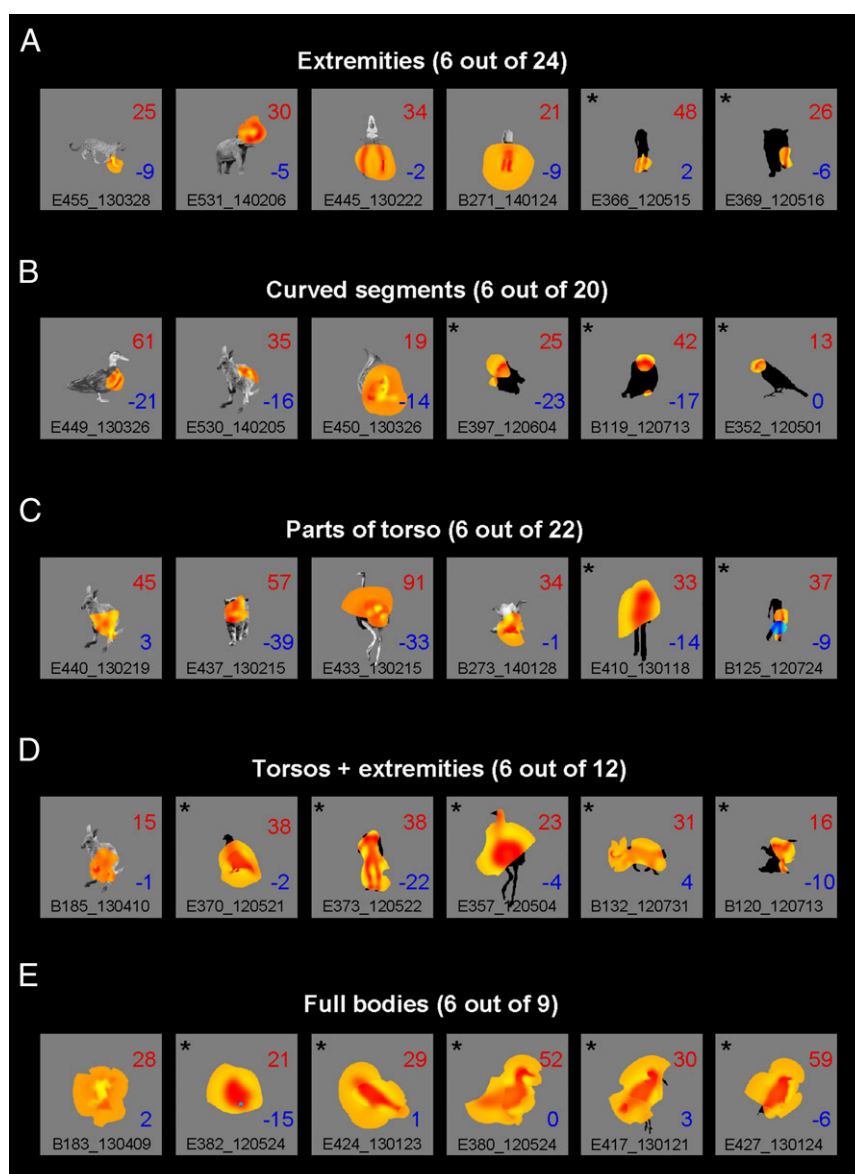


Fig. 2. Classification images of example neurons grouped by the class of the fragment revealed by the Bubbles. (A) Extremities, usually related to limbs. (B) Curved boundary segments. (C) Larger fragments that include part of the torso. (D) Fragments that encompass the torso and one or more extremities. (E) Neurons for which Bubbles revealed the whole body, defined as at least 90% overlap between fragment and stimulus. The actual feature summed across the significant SF bands as revealed by the Bubbles is shown in yellow (excitatory fragments) or blue (inhibitory fragments) and overlaps the original image or its silhouette. The asterisks indicate silhouettes. The mean net response for the high- and low-response trials (16% top and 16% bottom stimuli) are shown in each panel, following the conventions in Fig. 1. Note that due to limitation of the figure size A–E do not show the results of all neurons. The total number of neurons having a feature of a particular class is shown above each panel.

Table 3. Classification of the features revealed by the Bubbles in the preferred images of $n = 81$ neurons with significant features

Feature type	Original		Silhouette		Total	
	+	−	+	−	+	−
Extremities	14	1	7	2	21	3
Curved segments	10	1	8	1	18	2
Part of the torso	14	1	7	0	21	1
Torso + extremity	2	0	10	0	12	0
Full body	2	0	7	0	9	0
Unclassified	7	1	2	3	9	4

The left column represents the feature types (see Fig. 2 for examples). Significant excitatory and inhibitory features are symbolized by + and −, respectively. Each number denotes the count of neurons in which a particular feature can be found. The neurons marked with “Full body” have more than 90% coverage between the revealed features and the entire stimulus. Note that the sum of the total numbers (100) in the two rightmost columns exceeds the number of neurons with significant fragments reported in Table 1. This is due to multiple significant feature types revealed for the same neuron ($n = 18$ neurons). Seventy-five of the neurons were tested with bodies, four with faces, and two with manmade objects. One neuron that was tested with a face had a feature on the lower part of the face; this was tabulated as “Part of the torso” (*Supporting Information*).

sufficiently sensitive to reveal full bodies (*Supporting Information*). When presenting the same Bubbles stimuli to these model neurons as those used with the real neurons, only 5% of the model neurons showed an overlap smaller than 90% between the fragment revealed by Bubbles and the whole image. This strongly contrasts with the observed percentage of real neurons (89%) having an overlap smaller than 90% (Fig. S1).

We tested 25 neurons with multiple response-eliciting original gray-level (18) or silhouette (seven) images (number of multiple images per neuron ranged from two to seven, median three; $n = 25$). Bubbles revealed responses to image fragments from multiple images in 16 of these neurons (12 with original and 4 with silhouette stimuli). Fig. 3 illustrates that some of these 16 neurons (indicated by a rectangular frame) responded to corresponding features in the different images [e.g., an extremity (marked with blue neuron identifiers in Fig. 3) and curved boundary (marked with red identifiers)]. Less-clear instances of correspondence were present, however, in some other neurons. This could be due to single neurons responding to multiple features that may not overlap across the small sample of images that we tested with Bubbles.

As reported before (7), some body patch neurons also responded well to nonbody object images. As shown in Fig. 3 (fourth row, neuron E445_130222) for a neuron tested with a nonbody object, Bubbles suggested that selectivity for limb-like extremities (e.g., the handle of pliers) can explain at least some of the responses of body patch neurons to nonbody objects.

The midSTS Feature Code Tolerates Changes in Position and Scale.

The Bubbles data suggest that the majority of midSTS body patch neurons code relatively local image fragments. Here, we test whether this coding is specific to the image coordinates of the fragment (i.e., a specific location within the RF) or whether it generalizes to different retinal positions and sizes of the fragments. Note that a degree of invariance would be expected because midSTS neurons respond in a translation- and size-tolerant manner to whole body images (17). To test for translation tolerance, we reverse-correlated neuron ($n = 46$) responses with Bubble stimuli presented in two different locations of the neurons' RF. To test for size tolerance, we reverse-correlated neuron responses ($n = 21$) for two different image sizes and proportional bubbles sizes. Whenever we found a significant fragment in at least one RF location or for one size for a particular neuron and SF band, we computed

pixelwise correlations between the *DS* obtained from the two RF locations, or from the two stimulus sizes for that band—conforming the two *DS* to align them in object-centered coordinates (*Methods*). The distribution of Pearson correlation coefficients was significantly greater than 0 at each band for position changes (Wilcoxon test; $P < 0.002$) and size changes, but only for the middle SF bands in the latter (Wilcoxon test; $P < 0.002$; Fig. 4). When summing the *DS* across bands, the median correlation coefficients were 0.38 and 0.67 for the position and size tolerance, respectively, with both median values significantly greater than 0 (Wilcoxon test; $P < 0.0001$). Thus, coding of image fragments showed translation and size tolerance, demonstrating that midSTS body patch neurons code the feature content of image fragments, not merely their absolute spatial locations in the image or in the spatial RF of the neurons.

Independent Test That Neurons Indeed Code Fragments. The Bubbles data suggest that midSTS body patch neurons code image fragments smaller than the whole object. In a follow-up experiment we compared in an independent sample of 34 body-patch neurons (same guide tube locations) from both monkeys the response to the full preferred image with the responses to its parts. We presented a response-eliciting full body image interleaved with presentations of single parts of that image (Fig. S2). For each neuron, we determined the “best” and “worst” body part using an unbiased procedure (*Methods*). We then compared responses to these parts with the full body image using two contrast indices (best part: PFI(B); worst part: PFI(W); *Methods*). A contrast index of 0 indicates an equal response to the single part and the full body, whereas an index smaller than -0.33 corresponds to a twofold greater response to the full body than to the body part. The median PFI(B) (-0.04) did not differ significantly from 0 (Wilcoxon test; $P = 0.59$), indicating a similar average response to the best part and the full body (Fig. S2). We found that 65% of the best body parts were extremities. The median PFI(W) was -0.93 , indicating a negligible net response to at least one part of the full body. These data show that, overall, single body patch neurons respond as well to a single body part as to the full body, in agreement with the conclusion of Bubbles.

Discussion

We previously showed that the average responses of midSTS body patch are stronger for images of bodies than for those of faces, fruits, and inanimate manmade objects (7). However, although the large majority of these neurons respond to bodies, it was unclear what stimulus features they respond to. Using Bubbles, we found that most midSTS body patch neurons respond to local body features, such as extremities and outer contour fragments of the torso. The presence of such features in some manmade objects (e.g., the handle of pliers) might explain why midSTS neurons also respond to some inanimate objects. This suggests that category selectivity of the midSTS body patch depends more on the feature statistics of bodies (e.g., extensions occur more frequently in bodies) than on semantics (bodies as an abstract category). One advantage of the Bubbles procedure applied to brain data is to dissociate the interpretation of the origin of response (i.e., feature probabilities in a category vs. the semantics of the category) (20).

In a minority of the tested neurons, Bubbles did not isolate the fragment that drove the neuron's response. This could arise from a number of reasons, including a too-low number of trials, noisier responses, or the more interesting possibility that the whole image was necessary to elicit a response. Bubbles revealed the whole image in only 11% of the tested body patch neurons with a significant Bubbles outcome. Note that Bubble stimuli present parts of an image and not the complete image. This suggests that the neurons for which Bubbles revealed the whole image respond to multiple body parts that do not need to be all simultaneously present to drive the neuron (an OR instead of a wholistic AND operation).

The midSTS body patch is a possible homolog of the human extrastriate body area (EBA) (8). Our data can assist interpretations

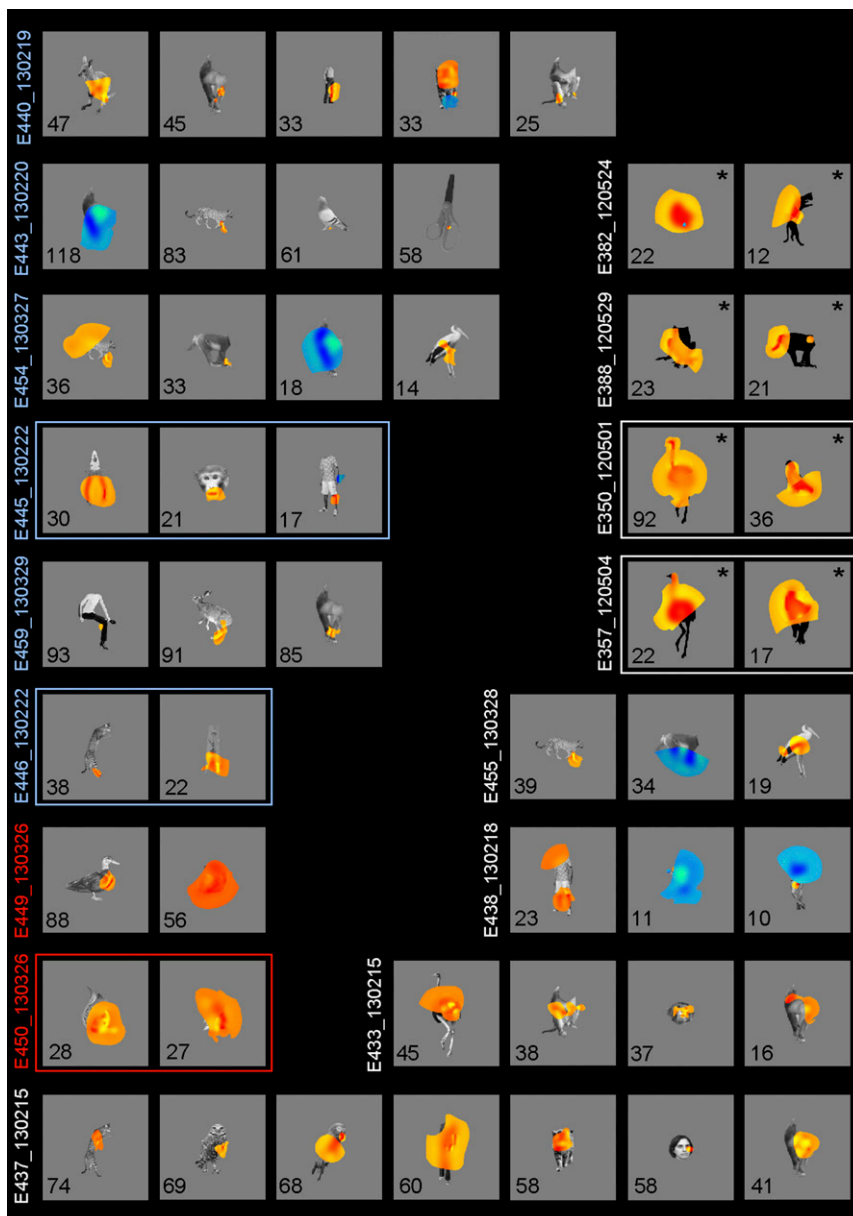


Fig. 3. Neurons tested with multiple stimuli showing significant fragments for more than one stimulus. Classification images are computed and presented using the same conventions as in Figs. 1C and 2. Left-to-right order of the tested stimuli corresponds to the ranking from the most preferred to least preferred for each neuron. Numbers in the bottom left corners of the panels correspond to the net firing rate (spikes per s) to the original stimuli during the search test. For each neuron, we show only images for which Bubbles revealed a significant feature (three images rejected) and for which the net firing rate to that image in the search test was at least 10 spikes per s (four images rejected). The asterisks indicate silhouettes. The blue and red color of the identifiers denotes neurons that responded to extremities and curved segments, respectively. Other neurons are marked with white identifiers. Neurons for which Bubbles revealed corresponding features in multiple images are marked by rectangles.

of human fMRI data of the EBA. EBA activation for single body parts is smaller than for whole bodies (21) and scales in proportion to body visibility in the image (22). This can be readily explained by our observation that the large majority of single midSTS neurons respond to local body fragments and that those differ between neurons. Indeed, the more the images contain such local fragments, the stronger the population response will be. By comparing fMRI activations to different configurations of body parts, it has been suggested more recently that the EBA represents bodies as a whole (23). However, the fMRI activations to the simultaneous presentation of randomly positioned body parts (“sum of parts” stimulus) may have been subject to divisive normalization (24) in that study, explaining at least partially the smaller activa-

tions compared with full bodies. Also, the former stimulus contained a greater number of ipsilaterally located parts than the full body, which also may have contributed to the effect. Because our Bubble stimuli consisted of several fragments at different spatial scales, the location of an image fragment relative to other image parts can be approximately computed in many of our stimuli. Thus, we cannot exclude that a configurational or relative position code (e.g., a particular feature at a particular location relative to the center of mass of the global shape) is present in the midSTS body patch, and this requires further investigation. Such a configurational coding could explain at least part of the EBA fMRI data reported by Brandman and Yovel (23) because in their “sum of parts” stimulus the natural configuration of body parts is

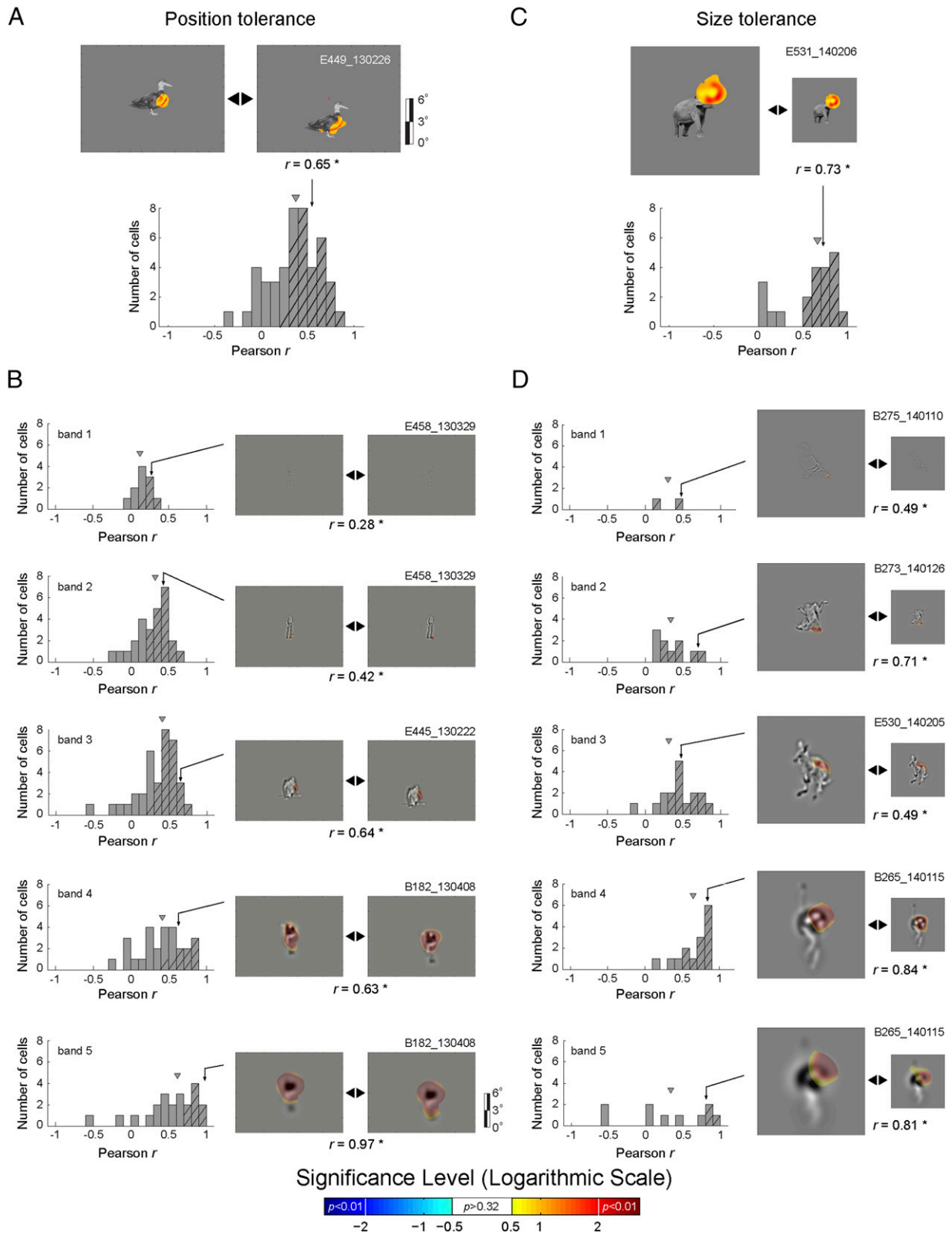


Fig. 4. Translation and size tolerances. (A) Translation tolerance per SF band. (Bottom) The distribution of the Pearson correlation coefficients r (in object-centered coordinates) between the raw DS values for the Bubbles stimuli presented at two nonoverlapping positions. (Top) A representative example neuron. (B) r distributions for translation tolerance per SF band. (Left) Distributions of Pearson correlation coefficients r (in object-centered coordinates) for each band separately. (Right) Representative examples of neurons with significant features per band. The gray rectangle depicts the screen and full spectrum (A) or band-pass filtered (B) images are presented in screen coordinates. (C) Size tolerance across SF bands. (Bottom) The Pearson correlation coefficients r between the raw DS values for the Bubbles stimuli presented at two different sizes. (Top) A representative example neuron. (D) r distributions for the size tolerance per SF band. (Left) Distributions of Pearson correlation coefficients r for each band separately. (Right) Representative examples of neurons with significant features per band. Number of neurons per band varies because we only plot neurons with significant positive or negative DS ($P < 0.01$). Hatching denotes a significant correlation. Triangular marks denote the median of the distributions.

destroyed, which may have reduced the response of body part neurons. However, note that another human fMRI study failed to find the smaller activations to the “sum of parts” stimulus in the EBA (25).

Bubbles revealed inhibitory and excitatory image fragments in a few neurons, suggesting opponent coding. Opponent coding as a computational principle is present at different levels of the visual system. Well-known examples include the center-surround organization of retinal ganglion cells (26) and the visual cortical simple cell receptive fields (27). Modeling of shape tuning functions of single neurons suggested the presence of inhibitory subunits in the macaque posterior inferior temporal cortex (28). Here we provide direct evidence for opponent coding in some neurons of an fMRI-defined body patch, using a method that is well-known for having little bias (9).

In agreement with our previous study (17) that showed similar responses and selectivity of midSTS body patch neurons for the original and silhouette versions of the objects, Bubbles revealed similar fragments for these two image versions. However, Bubbles revealed larger fragments and more frequently whole bodies for the silhouettes compared with the full images. A difference between silhouettes and the shaded, textured images is that unlike in the latter, silhouettes have uniform luminance within their boundaries. The variation in luminance and contrast within the original shaded stimuli may have reduced the fragment size revealed by the Bubbles in those stimuli relative to the silhouette versions.

Some body patch neurons also respond well to faces, as reported by (7). When testing body patch neurons with Bubble stimuli derived from a face, we revealed fragments of the cheek, ear, and the curved back of the head. The image features that Bubbles isolated for body-patch neurons differ from the face features reported for face-selective neurons in the neighboring face patches PL (posterior lateral), ML (middle lateral), and MF (middle fundus). First, unlike in the face patches PL and ML (15), Bubbles rarely disclosed the eyes as significant features. Second, face cells in ML and MF are tuned to the internal shape and contrast polarity features of a face (29, 30), unlike the extensions and boundary features of body-patch neurons. Together, the different feature selectivity of the face and body-selective patches reveals an important computational coding principle in the ventral visual stream: Different sets of features code different object categories such as faces, bodies, and inanimate objects. This agrees with computational (31) and behavioral research (32, 33) suggesting that category exemplars are represented in terms of category-specific features, including when reverse correlating from integrated EEG and MEG brain signals (11–13). These features can be viewed as the building blocks whose combinatorics can represent the variety of different images within the object category (31). This principle is similar to those of some deep neural network models of the ventral visual stream that consist of a mixture of “submodels” that are each effective at a particular categorization but suboptimal for others (34).

In summary, with Bubbles we showed that the majority of midSTS body patch neurons respond to relatively local image fragments that often are limb-like extremities or parts of a body boundary instead of a whole body, with a coding that is tolerant to translation and scale, and is sometimes opponent. Further work is needed to determine how this coding is transformed in the anterior body patch. Because the midSTS body patch neurons respond to fragments of a body (or an object that contains the feature), we are reluctant to label these neurons “body neurons.” Nonetheless, because these neurons respond to features that are present frequently in images of bodies, they can contribute to the perception of bodies. Causal experiments are needed to answer this crucial question.

Methods

Stimuli.

Search stimulus set. We searched visually responsive neurons with 100 stimuli that included 10 classes of achromatic images: monkey and human bodies (excluding the head), monkey and human faces, four-legged mammals (with

head), birds (with head), fruits/vegetables, body-like sculptures, and two classes of manmade objects. These stimuli were identical to those of the main test of ref. 7. We equated low-level image characteristics (i.e., mean luminance, contrast, and aspect ratio) across stimulus classes (see ref. 7). We controlled the difference between the mean aspect ratio of the monkey and human bodies by using two classes of manmade objects: one matching the aspect ratio of the monkey bodies (objectsM) and another matching the aspect ratio of the human bodies (objectsH). Images were resized so that the average area per class was matched across all classes, except for the objectsH and human bodies, while allowing for some variation in area [range: 3.7° to 6.7° (square root of the area)] among the exemplars in each class. The mean vertical and horizontal extent of the images was 8.3° and 6.7°, respectively. All images were embedded in a pink noise background having the same mean luminance as the stimuli that filled the entire display (40° × 30°). Each image was presented on top of nine different backgrounds that varied randomly across stimulus presentations. Stimuli were gamma-corrected.

Bubbles stimuli. The Bubbles procedure was applied on the achromatic stimuli of the search test (discussed above) or on their black silhouette versions, for which all object pixels were black (17). We first decomposed the original or silhouette image into five SF bands of one octave each, centered at 11.3, 5.65, 2.8, 1.4, and 0.7 cycles per degree (using the matlabPyrTools for MATLAB, www.cns.nyu.edu/lcv/software.php). At each SF band, we applied Gaussian apertures (bubbles) at random image positions to sample different image fragments (i.e., contiguous pixels). Bubble size was normalized to reveal the same number of cycles (three) per bubble per band; SDs of the Gaussian kernels were 0.23°, 0.45°, 0.90°, 1.81°, and 3.62° for the highest to the lowest SF band, respectively. The number of bubble apertures was adapted in an exponential manner across bands so that the high SF bands were sampled with more and the low SF bands were sampled with fewer bubbles (Fig. 1A). The exact number of bubbles per band differed slightly across images due to a random factor in the formula, whereas the total number of bubbles across bands per image, 188, was kept constant. Next, we summed the sampled images across SF bands to obtain the stimulus and then placed it on a uniform gray background with the same mean luminance as the stimuli in the search test. Except for the silhouette-based Bubble stimuli, the contrast of the resulting stimulus was enhanced, maximizing the range between the minimum and maximum gray values in the stimulus. For each neuron, we repeated this procedure 800–1,600 times to compute many trial-unique stimuli, which were gamma-corrected.

Body part stimuli. We created stimuli that isolated body parts (e.g., limbs, tail, torso, and head) and other smaller fragments (such as limb extremities and curved boundaries of the torso) that were shown at their original position (Fig. S2). They were placed on a uniform gray background, covering the entire display and having the same luminance as the mean luminance of the search stimulus set. The edges where the body parts were disconnected from the rest of the body were smoothed and faded to the background. Stimuli were gamma-corrected.

Subjects. Two adult male rhesus monkeys (*Macaca mulatta*, 6–8 kg) took part in this study. They were implanted with an MR-compatible headpost, for immobilizing the head during training and recording, and a recording cylinder targeting the left midSTS body patch. These were the same animals as in ref. 7. Animal care and experimental procedures complied with the national and European laws and the study was approved by the Animal Ethical Committee of the KU Leuven.

Recordings. Standard single-unit recordings were performed with tungsten microelectrodes (from FHC, impedance measured in situ ranging between 0.8 and 1.7 M Ω ; for details see refs. 7 and 35). Briefly, the electrode was lowered into the brain with a Narishige microdrive using a metal guide tube fixed in a standard Crist grid positioned within the recording chamber. After amplification and filtering between 540 Hz and 6 KHz, single units were isolated online using a custom amplitude- and time-based discriminator.

The position of one eye was continuously tracked by means of an infrared video-based tracking system (EyeLink from SR Research; sampling rate 1 kHz). Stimuli were presented on a CRT display (Philips Brilliance 202 P4; 1,024 × 768 screen resolution; 75-Hz vertical refresh rate) at a distance of 57 cm from the monkey's eyes. The on- and offset of the stimulus was signaled by a photodiode detecting luminance changes of a square in the corner of the display that was invisible to the animal. A digital signal processing-based computer system controlled stimulus presentation, event timing, and juice delivery while sampling the photodiode signal, eye positions, spikes, and behavioral events. Time stamps of the spikes, eye positions, stimulus, and behavioral events were stored for offline analyses.

We recorded from the left midSTS body patch, as defined by fMRI in the same subjects by contrasting images of headless monkey bodies and control objects (for details see ref. 7). The Crist grid locations targeting the body patch were identical to those of ref. 7.

Tests. In all tests, stimuli were presented for 200 ms each, with an interstimulus interval of ~400 ms, during passive fixation on a small red target (0.2°) superimposed on the stimuli in the middle of the screen. Fixation was required in a period from 100 ms prestimulus to 200 ms poststimulus onset (square fixation window size: 2° × 2°), and if fixation was aborted during this period, the trial was discarded from the analysis. Juice rewards were given with decreasing time intervals (2,000 ms to 1,350 ms) as long as the monkeys maintained fixation.

Search test. Neurons were tested with pseudorandom, interleaved presentations of the 100 images (see ref. 7 for details). The mean number of unaborted presentations per stimulus was 5.5, averaged across neurons. The pink noise background was present throughout the task and was refreshed simultaneously with stimulus onset. Based on this test, images were selected for the subsequent tests.

RF mapping. For some neurons, a 4° sized response-eliciting image was presented at 35 positions ranging from 3° ipsilateral to 9° contralateral and from 9° below to 9° above the horizontal meridian. Adjacent positions differed by 3°, horizontally or vertically. The positions were tested randomly interleaved. The fixed pink noise background was present throughout the task. The mean number of unaborted presentations per position was 6.3 ($n = 68$ neurons); see ref. 17.

Bubbles test. Trial-unique bubble stimuli, originating from a response-eliciting image or its silhouette, were presented foveally ($n = 87$ tests) or in another location within the RF ($n = 18$ tests, mean eccentricity for the nonfoveal tests: 4.5°) of the neuron. The uniform gray background covered the entire display. If eye fixation was aborted, the trial was not analyzed and the same Bubble stimulus was presented later. If we could hold the neuron long enough, a new Bubbles test was run using another response-eliciting image. The mean number of unaborted presentations across tests was 872.

Position tolerance Bubbles test. Bubble stimuli originating from a response-eliciting image were presented at two positions within the RF of the neuron. There was minimal or no overlap between the stimuli at the two positions (mean eccentricity of the two positions: 0.8° and 6°, mean distance between the positions: 5.6°, $n = 46$ tests). The Bubble stimuli shown at the two positions were drawn randomly from the same pool. Thus, they were trial-unique at a given position, but some stimuli were repeated across positions (depending on the number of generated Bubble stimuli in the pool and the number of trials shown to the monkey). The presentations at the two positions were randomly interleaved and the mean number of unaborted trials per position was 866 ($n = 46$ tests).

Size tolerance Bubbles test. Bubble stimuli originating from a response-eliciting image for the neuron were presented at their original size and at 50% of their original size. In all tests, the stimuli were presented at the fovea. Similar to the position tolerance test, the Bubble stimuli for the two sizes came from the same pool, so they were trial-unique per size but partially overlapped across sizes. The presentations of the two sizes were randomly interleaved and the mean number of unaborted trials per size was 822 ($n = 21$ tests).

Body part test. The full body and the separate body parts of a response-eliciting image (mammal, bird, or human body) were shown randomly interleaved. The stimuli were presented at the fovea. The mean number of unaborted trials per stimulus condition was 10 ($n = 34$).

Data Analysis. Firing rate was computed for each unaborted stimulus presentation in two analysis windows: a baseline window ranging from 100 to 0 ms before stimulus onset and a response window ranging from 50 to 250 ms after stimulus onset. The responses in the search and RF-mapping tests were assessed for statistical significance by a split-plot ANOVA with repeated measure factor baseline vs. response window and between-trial factor stimulus condition. A test for a neuron was included only when either the main effect of the repeated factor or the interaction of the two factors was significant ($P < 0.05$) for that test in that neuron. All further analyses (unless otherwise stated) were based on baseline subtracted net responses. We computed for each neuron a BSI, defined as: $BSI = \frac{\bar{R}_{body} - \bar{R}_{nonbody}}{|\bar{R}_{body} + \bar{R}_{nonbody}|}$, where \bar{R}_{body} and $\bar{R}_{nonbody}$ are the mean net firing rates to body (monkey and human bodies, mammals, and birds) and nonbody stimuli (monkey and human faces, fruits/vegetables, and manmade objects) presented in the search test.

The fragments correlated with excitatory or inhibitory neural responses were revealed through classification images depicting DS (9, 18). First, we sorted the trial-unique Bubbles stimuli according to the net response they elicited. We then fitted a cumulative Gaussian function to the sorted net

responses and estimated the mean and the SD of the Gaussian. Next, Bubbles stimuli were assigned to three groups: high response (response strength above mean + SD), medium response (within 1 SD around the mean response strength), and low response images (responding below mean - SD). $DS(b, i)$ were computed for each SF band b and pixel i (excluding the background) separately as follows:

$$DS(b, i) = \sum_h B(b, i) - \sum_l B(b, i),$$

where $B(b)$ are the bubble masks (the Gaussian bubbles apertures) of the band b of the image, with h and l corresponding to high response and the low response Bubbles stimuli, respectively. Note that h and l groups contained an equal number of stimuli. The Bubble masks $B(b)$ were down-sampled with a factor of $2^{(b-1)}$ (where $b = 1$ is the highest SF band and $b = 5$ is the lowest one) to overcome the effect of correlations between neighboring pixels due to Bubbles sampling. The background was eliminated by a mask that coincided with pixels that contained nonzero values of the filtered image (before applying the bubbles) per SF band. Those masks contained a different number of pixels per original image and per SF band (lower SF bands including more pixels than higher ones).

Similar to ref. 36, the significance of the DS was assessed by means of a permutation test for each SF band separately. We randomly permuted the trial order of the Bubble stimuli, thus destroying the relationship between the stimuli and the neural responses. Then we computed the DS using the same procedure as above and took the maximum (DS_+) and minimum of (DS_-) of the computed DS values of the permuted data. This procedure was performed 500 times, yielding null distributions of the 500 DS_+ and 500 DS_- values, respectively. For each pixel of a band, we computed two P values using the percentile of the observed DS value in the DS_+ and DS_- null distributions. P values smaller than 0.01 were considered significant.

In the figures we show image fragments summed across the SF bands that revealed a significant feature. To do this, we selected per SF band for each fragment of which at least one pixel passed the $P < 0.01$ significance threshold, the pixels with a $P < 0.32$. Then, similar to the way the Bubbles stimuli were created, we summed the selected pixel values across the SF bands to obtain the image fragment. Note that to maintain the information about excitatory and inhibitory responses (positive and negative DS), we changed the sign of the values (multiplying by 1 for DS_+ and -1 for DS_- , and the values of the pixels with $P \geq 0.32$ were set to 0).

We also analyzed the Bubbles data by computing the mutual information between the cell's response and the bubble masks for each SF band, following ref. 37. This yielded qualitatively similar results but the advantage of computing DS is that it is a signed metric, distinguishing excitatory from suppressive image fragments.

We performed an "inverse mapping analysis" to estimate, for each neuron and stimulus, the net response produced by the feature that the Bubbles procedure isolated. Following ref. 11, we first selected the Bubbles stimuli that revealed the significant features—that is, the images where at least 15% of the pixels of the significant feature (defined as pixels with a DS corresponding to $P < 0.01$) are visible (defined as at least 50% of the maximum of the Gaussian bubble) in at least one of the SF bands. This was done separately for the significant excitatory (positive DS) and inhibitory (negative DS) features. We then took the averaged net responses to the selected Bubbles images as an estimate of the net firing rate elicited by the presentation of the feature. For comparison, we also computed the average net firing rate for the remaining Bubbles images where the significant feature was not visible.

For each Bubbles test, we inspected visually the image fragment(s) revealed by Bubbles. We then classified the fragment(s) for the neuron's preferred image into six classes, based on the location and coverage of the fragments on the image. The six classes were extremity (such as limbs or elongated extensions of an object), part of torso, part of torso combined with extremity (in the cases when the revealed part of the torso extended to a limb), curved segment (such as the curved arc of the back of an animal or a human's shoulder), full body (in the cases when the revealed fragment was covering more than 90% of the object), and unclassified (fragments that did not fall in any of the previous five classes). The percentage of overlap/coverage between the revealed fragment(s) and the whole image was computed as the ratio between the pixels revealed by Bubbles (using the $P < 0.32$ threshold) and the total number of pixels in a given image.

Position tolerance was assessed by means of the Pearson correlation r between the $DS(b, i)$ at the two tested positions. The correlation was computed across the pairs of pixels i within the mask for a particular band b in object-centered coordinates. We tested the significance of r for each single neuron and SF band with a permutation test. To do so, we randomly

permuted the trial order of the Bubble stimuli presented at one of the two positions, thus destroying the relationship between the stimuli and the neural responses, followed by a computation of the DS (DS_p). Then we computed the Pearson correlation between DS_p and the unshuffled DS of the other position. We performed this procedure 100 times to produce a null distribution. We used the same procedure to test the significance of the DS summed across the five SF bands. The procedure for testing the significance of the size tolerance was identical, except that the $DS(b)$ for the big size was downscaled to 50% to match the $DS(b)$ dimensions of the smaller size. The null distribution consisted of 100 r values, computed between the unshuffled $DS(b)$ for the big size and the permuted stimulus labels DS_p for the small size. Note that this permutation test preserves the spatial correlations that are present in the Bubble stimuli and thus is more conservative and appropriate than a test in which the DS values themselves are permuted (the latter also destroys the spatial correlations inherent to the Bubbles stimuli).

For the body part test, we compared the mean net responses for the best or worst body part and the whole body. For each neuron, we computed a part vs. full index, $PFI = \frac{\bar{R}_p - \bar{R}_f}{|\bar{R}_p| + |\bar{R}_f|}$, where \bar{R}_p and \bar{R}_f are the mean net responses to body

parts and full bodies, respectively. We computed two separate indices: one for the “best” [PFI(B)] and one for the “worst” [PFI(W)] body part. In order to select the best and worst body parts in an unbiased manner, the rank of the stimuli (best or worst) was determined from half of the trials and then we computed the average response to these parts from the other half of the trials. We performed this procedure twice, alternating which half of the trials was used for selection or response computation. The response to a part that entered the index was the average of the two computations.

ACKNOWLEDGMENTS. We thank B. Cooreman, M. Docx, I. Puttemans, C. Ulens, P. Kayenbergh, G. Meulemans, W. Depuydt, S. Verstraeten, and M. De Paep for technical support; Dr. P. Downing and Dr. M. Tarr for providing some of the stimuli; and Dr. R. Ince for helping with the Mutual Information analysis. This study was supported by Fonds voor Wetenschappelijk Onderzoek Vlaanderen, GOA (Gemeentelijk Onderwijsachterstandbeleid), IUAP (Inter-University Attraction Pole), and PF (Programma Financiering) grants. P.G.S. is supported by Wellcome Investigator Grant 107802/Z/15/Z. This work was also supported by a fellowship from the Agentschap voor Innovatie door Wetenschap en Technologie, Grant 101071 (to I.D.P.).

- Downing PE, Jiang Y, Shuman M, Kanwisher N (2001) A cortical area selective for visual processing of the human body. *Science* 293(5539):2470–2473.
- Tsao DY, Freiwald WA, Knutsen TA, Mandeville JB, Tootell RB (2003) Faces and objects in macaque cerebral cortex. *Nat Neurosci* 6(9):989–995.
- Bell AH, Hadj-Bouziane F, Frihauf JB, Tootell RB, Ungerleider LG (2009) Object representations in the temporal cortex of monkeys and humans as revealed by functional magnetic resonance imaging. *J Neurophysiol* 101(2):688–700.
- Pinsk MA, DeSimone K, Moore T, Gross CG, Kastner S (2005) Representations of faces and body parts in macaque temporal cortex: A functional MRI study. *Proc Natl Acad Sci USA* 102(19):6996–7001.
- Pinsk MA, et al. (2009) Neural representations of faces and body parts in macaque and human cortex: A comparative fMRI study. *J Neurophysiol* 101(5):2581–2600.
- Popivanov ID, Jastorff J, Vanduffel W, Vogels R (2012) Stimulus representations in body-selective regions of the macaque cortex assessed with event-related fMRI. *Neuroimage* 63(2):723–741.
- Popivanov ID, Jastorff J, Vanduffel W, Vogels R (2014) Heterogeneous single-unit selectivity in an fMRI-defined body-selective patch. *J Neurosci* 34(1):95–111.
- Caspari N, et al. (2014) Fine-grained stimulus representations in body selective areas of human occipito-temporal cortex. *Neuroimage* 102(Pt 2):484–497.
- Gosselin F, Schyns PG (2001) Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Res* 41(17):2261–2271.
- Rousselet GA, Ince RA, van Rijsbergen NJ, Schyns PG (2014) Eye coding mechanisms in early human face event-related potentials. *J Vis* 14(13):7.
- Smith ML, Gosselin F, Schyns PG (2004) Receptive fields for flexible face categorizations. *Psychol Sci* 15(11):753–761.
- Smith ML, Fries P, Gosselin F, Goebel R, Schyns PG (2009) Inverse mapping the neuronal substrates of face categorizations. *Cereb Cortex* 19(10):2428–2438.
- Schyns PG, Petro LS, Smith ML (2007) Dynamics of visual information integration in the brain for categorizing facial expressions. *Curr Biol* 17(18):1580–1585.
- Tang H, et al. (2014) Spatiotemporal dynamics underlying object completion in human ventral visual cortex. *Neuron* 83(3):736–748.
- Issa EB, DiCarlo JJ (2012) Precedence of the eye region in neural processing of faces. *J Neurosci* 32(47):16666–16682.
- Rutishauser U, et al. (2011) Single-unit responses selective for whole faces in the human amygdala. *Curr Biol* 21(19):1654–1660.
- Popivanov ID, Jastorff J, Vanduffel W, Vogels R (2015) Tolerance of macaque middle STS body patch neurons to shape-preserving stimulus transformations. *J Cogn Neurosci* 27(5):1001–1016.
- Smith ML, Gosselin F, Schyns PG (2006) Perceptual moments of conscious visual experience inferred from oscillatory brain activity. *Proc Natl Acad Sci USA* 103(14):5626–5631.
- De Valois RL, Albrecht DG, Thorell LG (1982) Spatial frequency selectivity of cells in macaque visual cortex. *Vision Res* 22(5):545–559.
- Schyns PG, Jentzsch I, Johnson M, Schweinberger SR, Gosselin F (2003) A principled method for determining the functionality of brain responses. *Neuroreport* 14(13):1665–1669.
- Peelen MV, Downing PE (2007) The neural basis of visual body perception. *Nat Rev Neurosci* 8(8):636–648.
- Taylor JC, Wiggett AJ, Downing PE (2007) Functional MRI analysis of body and body part representations in the extrastriate and fusiform body areas. *J Neurophysiol* 98(3):1626–1633.
- Brandman T, Yovel G (2014) Bodies are represented as wholes rather than their sum of parts in the occipital-temporal cortex. *Cereb Cortex*, 10.1093/cercor/bhu205.
- Zoccolan D, Cox DD, DiCarlo JJ (2005) Multiple object response normalization in monkey inferotemporal cortex. *J Neurosci* 25(36):8150–8164.
- Soria Bauser D, Suchan B (2015) Is the whole the sum of its parts? Configural processing of headless bodies in the right fusiform gyrus. *Behav Brain Res* 281:102–110.
- Kuffler SW (1953) Discharge patterns and functional organization of mammalian retina. *J Neurophysiol* 16(1):37–68.
- Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 160:106–154.
- Brincat SL, Connor CE (2004) Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nat Neurosci* 7(8):880–886.
- Freiwald WA, Tsao DY, Livingstone MS (2009) A face feature space in the macaque temporal lobe. *Nat Neurosci* 12(9):1187–1196.
- Ohayon S, Freiwald WA, Tsao DY (2012) What makes a cell face selective? The importance of contrast. *Neuron* 74(3):567–581.
- Ullman S, Sali E, Vidal-Naquet M (2001) A fragment-based approach to object representation and classification. *International Workshop on Visual Form*, eds Arcelli A, Cordella LP, Sanniti di Baja G (Springer, Berlin), pp 85–100.
- Schyns PG, Rodet L (1997) Categorization creates functional features. *J Exp Psychol Learn Mem Cogn* 23(3):681–696.
- Schyns PG, Bonnar L, Gosselin F (2002) Show me the features! Understanding recognition from the use of visual information. *Psychol Sci* 13(5):402–409.
- Yamins DL, et al. (2014) Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc Natl Acad Sci USA* 111(23):8619–8624.
- Sawamura H, Orban GA, Vogels R (2006) Selectivity of neuronal adaptation does not match response selectivity: A single-cell study of the fMRI adaptation paradigm. *Neuron* 49(2):307–318.
- Alemi-Neissi A, Rosselli FB, Zoccolan D (2013) Multifunctional shape processing in rats engaged in invariant visual object recognition. *J Neurosci* 33(14):5939–5956.
- Schyns PG, Thut G, Gross J (2011) Cracking the code of oscillatory activity. *PLoS Biol* 9(5):e1001064.