

Statistical characterization of tactile scenes in three-dimensional environments reveals filter properties of somatosensory cortical neurons

Authors: Nadina O. Zweifel¹, Sara A. Solla^{2,3}, and Mitra J. Z. Hartmann^{1,4*}

Affiliations:

1. Department of Biomedical Engineering
Northwestern University
Evanston, IL, 60208, USA
2. Department of Physics and Astronomy
Northwestern University
Evanston, IL, 60208, USA
3. Department of Neuroscience
Northwestern University
Chicago, IL, 60611, USA
4. Department of Mechanical Engineering
Northwestern University
Evanston, IL, 60208, USA

***Corresponding author:** Prof. Mitra J. Z. Hartmann
Department of Biomedical Engineering
Department of Mechanical Engineering
Northwestern University
2145 Sheridan Road
Evanston, IL, 60208
hartmann@northwestern.edu
847-467-4633

Abstract

Natural scenes statistics have been studied extensively using collections of natural images and sound recordings. These studies have yielded important insights about how the brain might exploit regularities and redundancies in visual and auditory stimuli. In contrast, natural scenes for somatosensation have remained largely unexplored. Here we use three-dimensional scans of natural and human-made objects to quantify natural scene statistics at the scale of the human fingertip. Using measurements of distance, slope, and curvature from the object surfaces, we show that the first order statistics follow similar trends as have been observed for images of natural and human-made environments. In addition, independent component analysis of curvature measurements reveals Gabor-like basis vectors similar to those found in natural images. A simple neural model using these filters showed responses that accurately capture the statistics of responses in primate primary somatosensory cortex.

1. Introduction

An animal's only access to the environment is through the signals provided by the activation of its sensors as they interact with the external world. These sensory signals typically represent compressed and nonlinear transformations of physical quantities in the environment. Nonetheless, nervous systems can efficiently and reliably infer relevant information about the original environmental variables from the sensory input signals they evoke. It has been hypothesized that these successful inferences are achieved by exploiting regularities and redundancies present both in the environment and in the resulting sensory input signals [1, 2].

Over the last five decades, a large body of research on the statistics of visual and auditory natural scenes has generated strong evidence in support of this hypothesis [3-21]. Statistical models that predict the coding schemes and representations characteristic of cortical neurons (basis functions, basis vectors, receptive fields) have received considerable attention, and different algorithms to fit the models to sensory data have been proposed [4, 5, 14, 22-27]. The underlying assumption of these models is that sensory systems process incoming information by simultaneously activating a small subset of neurons while the majority remain silent. Some modeling approaches directly impose this constraint by maximizing the sparseness of the optimized representations (sparse coding) [22], while others indirectly achieve similar results by maximizing the independence of the predicted basis functions (independent component analysis) [4, 28].

The data used to train these models typically originate from a collection of natural scenes in the form of visual images or sound recordings; these represent the ensemble of inputs that the sensory system may encounter. For images, a large number of small regions within these images is then used to mimic receptor-like sampling of the scene. The basis functions that emerge from the fitted model thus represent the features of the sensory signal encoded by individual cortical neurons; these have been found to be very similar to cortical receptive fields recorded experimentally [4, 22, 28]. The features thus selected can give insight into the neural tuning that gives rise to an efficient code. In vision, these features resemble edge detectors and Gabor filters [22, 26, 28].

In contrast to vision and audition, the natural scene statistics of somatosensation have remained largely unexplored [19-21, 29]. Tactile sensors are unique in that the signals they acquire result from direct physical contact with the environment: it is impossible to “feel” an object without touching it. The interactive nature of touch makes tactile environmental statistics extremely challenging to quantify, as evidenced by the small number of studies that have attempted to measure these parameters *in vivo* [21] or *in silico* [19, 20, 30]. To date, psychophysical, neurophysiological, and self-report experiments have primarily focused on the physical and qualitative dimensions of tactile perception in the environment (e.g., roughness, smoothness, compliance, curvature, etc.) [31-38]. Despite substantial groundwork from these studies, the environmental statistics of somatosensation are still poorly documented [29].

The present study aims to shed light on the environmental statistics of the natural *tactile* scene by quantifying geometrical properties of the natural environment in three-dimensional (3D) space. We focus specifically on three different Cartesian-based metrics: local distance, slope, and curvature. We begin by quantifying the statistics of each of these metrics at the level of individual objects as well as the entire population of objects, and then use independent component analysis (ICA) in two different ways. First, ICA is used to find the leading statistically independent dimensions in local distance measurements, revealing that 95% of the variance in local distance data is predominantly accounted for by curvature. Second, ICA is used to estimate the filter characteristics of somatosensory cortical neurons. This analysis shows that the ICA basis vectors estimated from curvature patches have Gabor-like filter properties and are consistent with orientation sensitivity found in neurons in areas 1 and 3b of primary somatosensory cortex (S1).

2. Results

We collected 3D representations of 96 natural objects such as rocks, leaves, bark, fruits, vegetables, etc., as well as 41 human-made objects including scissors, knives, purses, mugs, etc. (Fig. 1ab). To quantify the local shape of these data, we randomly sampled 100-150 circular surface patches with a radius of 6 mm, similar to the size of a human fingertip [39] (Fig. 1c, *left*). This procedure resulted in a total of 12,309 and 5,788 patches for natural and human-made objects, respectively. For each surface patch we computed maps of distance, slope, and curvature; details for calculating these metrics are found in *Methods*. Briefly, the distance of the surface points was computed with respect to a reference plane (Fig. 1c, *middle*), yielding a circular distance map $P(D)$ with a total of 448 points for each patch. From $P(D)$ we computed a 421-dimensional slope map $P(S)$ using first-order forward difference, and a 384-dimensional curvature map $P(C)$ using second-order central difference. Examples of $P(D)$, $P(S)$, $P(C)$ for an example patch are shown on the right in Fig. 1c.

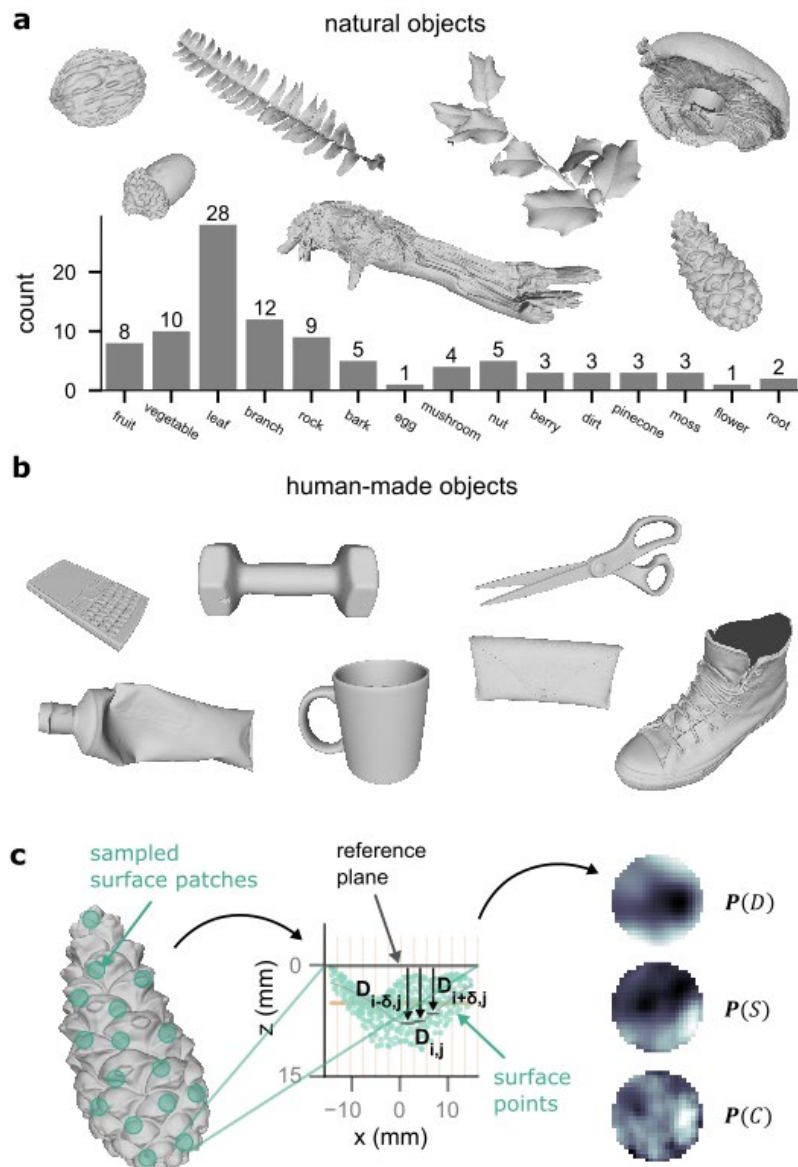


Fig. 1. The 3D datasets include a total of 96 natural objects and 41 human-made objects, cleaned and processed following procedures described in *Methods*. **(a)** Examples of 3D scanned natural objects, and a histogram of the number of objects in each of 15 categories. **(b)** Examples of 3D scanned human-made objects. **(c)** To sample the 3D data, approximately 100-150 circular surface patches were randomly selected from each object (left). For each surface patch, the distance of the surface points was measured with respect to a reference plane (middle). From these measurements, a circular distance map $P(D)$ of 448 points was created (right). Using forward and central differencing, a 421-dimensional slope map and a 384-dimensional curvature map, $P(S)$ and $P(C)$, were computed for each patch (see *Methods*).

2.1. First-order statistics of natural and human-made objects

We examined the first-order statistics of natural and human-made objects in terms of the local metrics: distance, slope, and curvature (D , S , and C). Fig. 2a compares the probability distributions of these metrics for two natural objects (bark, leaf) and two human-made objects (mug, purse). The distributions were computed from all patches for each object. As expected, human-made objects tend to have larger flat and low-curvature regions (e.g., the body of the mug); the probability distributions for all three metrics are correspondingly skewed towards zero. At the same time, small distinct regions of human-made objects exhibit higher curvature values (e.g., rim and handle of the mug) that affect the shape of the corresponding distributions. The effect of the rim and handle of the mug is particularly obvious in the distribution of C , in which additional modes appear at large positive and negative values. In contrast, the distributions for natural objects are clearly unimodal with a smooth decay towards the tails, indicating that the measured local metrics are more continuous across the surface than for human-made objects.

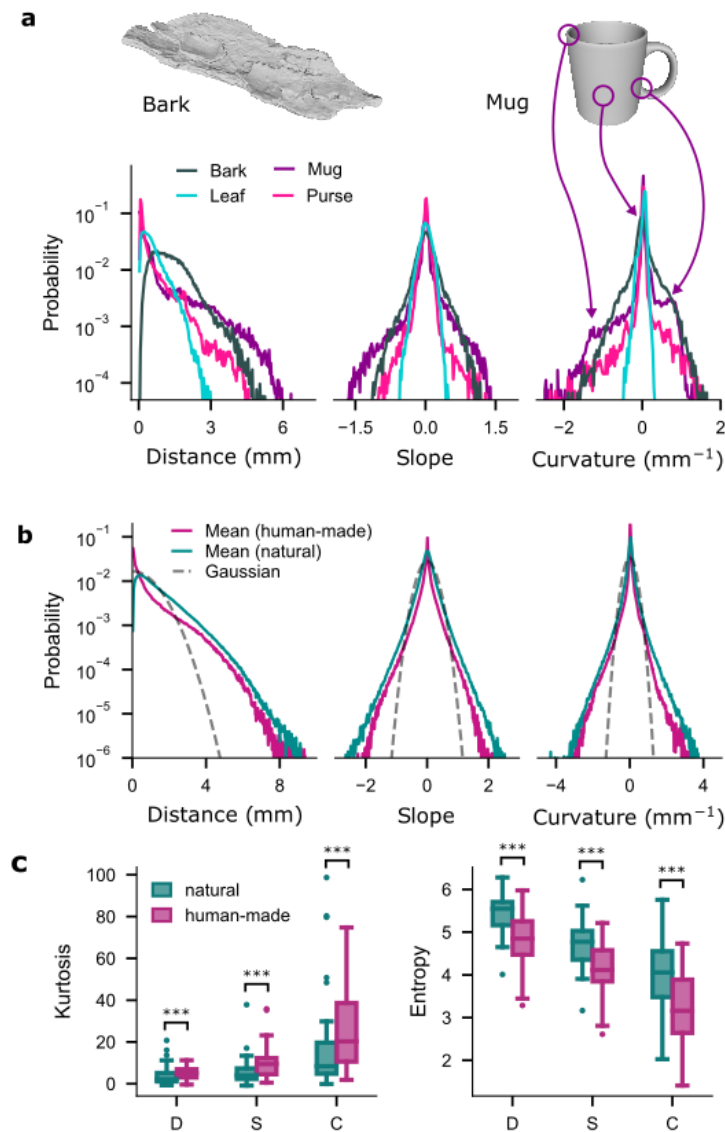


Fig. 2. Probability distributions of the three metrics D , S , and C for both natural and human-made objects. **(a)** Four example objects (two illustrated, not to scale) and their corresponding distributions of D (left), S (middle), and C (right). Distributions were computed from the ensemble of patches with $R = 6$ mm associated with each object. Note that the vertical axis is logarithmic. **(b)** Average probability distribution of D (left), S (middle), and C (right) across all sampled patches with $R = 6$ mm for natural (turquoise) and human-made (magenta) objects. The black dashed lines show a zero-mean Gaussian distribution of the same variance for comparison. Only the positive half of the Gaussian is used for distance D . **(c)** Distributions of kurtosis (left) and entropy (right, measured in bits) for the distribution for each object class and each metric.

To generalize this analysis, we computed the average probability distributions across all $P(\cdot)$ for each metric (Fig. 2b). Both natural and human-made objects result in super-Gaussian distributions, i.e., they have higher peaks at zero and heavier tails than a Gaussian distribution with the same mean and variance. Although the average distributions for human-made objects smooth over irregularities present in the distributions for individual human-made objects, the average distributions for human-made objects remain heavily skewed towards zero and are significantly different from the distributions for natural objects distance D (Kolmogorov-Smirnov test: $D = 0.303$, $p < 0.001$), slope S (Kolmogorov-Smirnov test: $D = 0.100$, $p < 0.001$), and curvature C (Kolmogorov-Smirnov test: $D = 0.149$, $p < 0.001$).

Differences between natural and human-made objects can also be quantified in terms of kurtosis: human-made objects result in significantly higher kurtosis values than natural objects (Fig 2c, left) in terms of D (Mann-Whitney rank test: $U = 1029$, $p < 0.001$), S (Mann-Whitney rank test: $U = 892$, $p < 0.001$), and C (Mann-Whitney rank test: $U = 984$, $p < 0.001$). The average distributions of D , S , and C have kurtosis values of 2.8, 6.1, and 17.9 for natural objects, and 4.4, 8.9, and 24.3 for human-made objects, respectively. Kurtosis is related to regularity (correlations) in the data and is inversely related to entropy, i.e., high kurtosis values correspond to low entropy, shown in the right panel of Fig. 2c. Entropy is significantly lower for human-made objects than for natural objects for all three metrics D (Mann-Whitney rank test: $U = 508$, $p < 0.001$), S (Mann-Whitney rank test: $U = 791$; $p < 0.001$), C (Mann-Whitney rank test: $U = 696$, $p < 0.001$). From an information-theoretical perspective, this result suggests that fewer bits are necessary to encode information about human-made objects than for natural objects when the objects are characterized through the chosen metrics D , S , and C . In addition, the right plot in Fig. 2c also indicates that entropy decreases with the order of the derivative: D results in higher entropy and C in lower entropy across individual objects. This result implies that a metric such as C may provide more regularities that could potentially be exploited by a sensory system to efficiently encode information about tactile images present in the world.

2.2. ICA analysis for the various metrics

The first-order statistics of three local metrics that characterize the shape of $R = 6\text{mm}$ surface patches: distance, slope, and curvature are highly regular and super-Gaussian, congruent with the statistics observed in natural images [40, 41]. However, it remains unclear which of these metrics may be the best candidate to provide information about the world to a tactile sensory system. To investigate this question, we used an unsupervised learning algorithm, Independent Component Analysis (ICA), to identify the fundamental features of natural data in terms of distance D , the least processed metric that characterizes local shape.

As described in [42], ICA is a linear model that maps a data vector x consisting of a set of features $[x_1, \dots, x_n]$ to a set of components $s = [s_1, \dots, s_n]$ through a square mixing matrix A , such that the sensory signal x can be linearly reconstructed from the components by

$$x = As. \quad \text{eq. (1)}$$

Both the mixing matrix A and the components s are unknown. The matrix A is estimated from the statistics of the observed data vectors x under the assumptions that A is square, the components s are statistically independent, and x can be reconstructed linearly [43]. The independent components s are then obtained by using the matrix W , the inverse of A :

$$s = Wx. \quad \text{eq. (2)}$$

The columns in A are called basis vectors; each basis vector is associated with one of the components in s . The rows in W play the roles of filters that act on the data vector x to obtain the corresponding component [28]. The matrix W contains the coefficients or weights for each of the features in x to recover each independent component in s . The set of weights associated with a particular component s_i , which correspond to the i -th row of W , can thus be interpreted as the receptive field for that component [28]. In contrast, the column basis vectors in A represent fundamental features that can be linearly combined to compose the observed signal x . The amplitude of each such feature is determined by the coefficients of the corresponding component of s ; these components are typically sparsely distributed [28]. There are many different ways to implement ICA; we used one of the most popular and computationally efficient algorithms, FastICA, proposed by Hyvarinen and Oja [43].

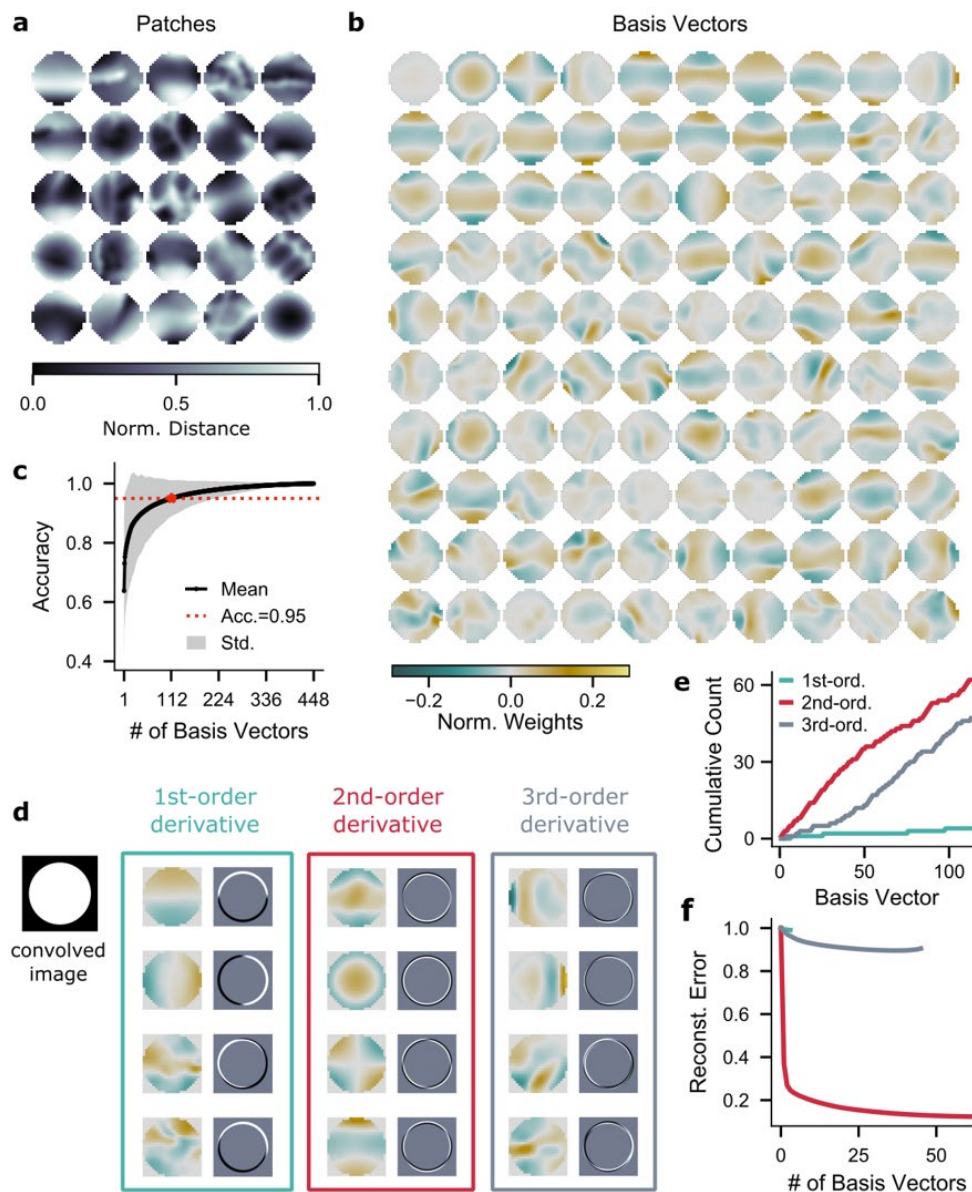


Fig. 3. Basis vectors estimated by ICA for surface patches characterized by distance D to the reference plane. **(a)** Example surface patches ($R = 6\text{mm}$) randomly sampled from natural 3D data that show spatial maps of the distance D . Grey scale is normalized to maximum distance within each patch. **(b)** The leading 100 basis vectors (columns of the mixing matrix A) as estimated by ICA from the distance D data. The mean was subtracted from each basis vector, and the norm of each basis vector was scaled to 1. The basis vectors are sorted by their contribution to reconstruction accuracy as shown in (c). **(c)** Cumulative reconstruction accuracy of the sorted basis vectors. **(d)** The basis vectors are classified as 1st, 2nd, or 3rd order by convolving each with a simple image: a white circle on black background (left). Representative examples selected from the first 113 basis vectors are shown in the three boxes. *Left box*: basis vectors with 1st-order derivative characteristics in detecting the change in distance D (slope). *Middle box*: basis vectors with 2nd-order derivative characteristics in detecting the rate of change in distance D (curvature). *Right box*: basis vectors with 3rd-order derivative characteristics (change in curvature). **(e)** Cumulative count of basis vectors for each category shown in (d). **(f)** Reconstruction error of the basis vectors for each category identified in (d). The basis vectors in each category are sorted by their contribution to the reconstruction error.

In our scenario, x represents a 448-dimensional vector of distances from a map $P(D)$ sampled from a natural object. Examples of $P(D)$ used for fitting the ICA model are shown in Fig. 3a. The FastICA algorithm does not sort the estimated basis vectors (columns of A), so we used a sorting algorithm proposed by Cheung and Xu [44] to order the basis vectors according to their contribution to reconstruction accuracy. The first 100 of the sorted basis vectors are shown in Fig. 3b. The cumulative reconstruction accuracy as sorted basis vectors are added is shown in Fig. 3c; this plot indicates that the leading 113 basis vectors suffice to achieve a reconstruction accuracy of 0.95. Based on this result, we determined the filtering properties of the leading 113 basis vectors by convolving each of them with a test image (Fig. 3d, *left inset*). The convolution output allowed for a categorization of the basis vectors into 1st, 2nd, and 3rd-order derivative filters. Examples of each category with the corresponding convolved images are shown in Fig. 3d. Note that 1st-order filters, associated with slope, are notably scarce, rising slowly in number as basis vectors are added. Within the first 113 basis vectors, only four 1st-order filters were present (Fig. 3e). In contrast, 2nd-order filters, associated with curvature, were the most common.

To further assess the relevance of the different types of basis vectors, each category was separately tested for reconstruction accuracy. The 2nd-order category resulted in significantly smaller reconstruction error than either the 1st-order (Mann-Whitney rank test: $U = 18042$, $p < 0.001$) or 3rd-order categories (Mann-Whitney rank test: $U = 1507754$, $p < 0.001$). Unsurprisingly, reconstruction accuracy for the 1st-order category was by far the lowest of the three (Median = 0.003)

Within the first 113 leading basis vectors, the 1st-order category includes a total of 4 basis vectors, the 2nd-order category includes 62, and the 3rd-order category includes 47 basis vectors. Given that reconstruction accuracy typically correlates with the number of basis vectors, one possibility is that the differences in reconstruction error between categories may just be an artifact of the different number of basis vectors in each category. However, sorting the basis vectors in each category by their contribution to the reconstruction error, we find that there is an intrinsic difference in the amount of information carried by the basis vectors in each category. The results in Fig. 3f show that the reconstruction error drops the fastest with additional basis vectors for those in the 2nd-order category while the slowest for those in the 1st-order category. This result indicates that the 2nd-order basis vectors carry more information about the original patch than the 1st or 3rd order basis vectors. Thus, even though the number of basis vectors in each category does affect the reconstruction accuracy, the order of the basis vector plays a significant role in the observed reconstruction errors shown in Fig. 3f.

2.3. Tactile coding model based on curvature information

Sparse coding and ICA are powerful statistical models that seem to capture aspects of the coding strategies in visual as well as auditory cortices [4, 22, 28]. The sparse coding algorithm originally introduced by Olshausen and Field (1997) used data from natural images to estimate basis functions that had striking similarities to the receptive fields of simple cells in the primary visual cortex [45].

Because the ICA results identified curvature as a good candidate to represent 3D shape information, we used a similar approach and applied the FastICA algorithm to the 384-dimensional curvature maps $P(C)$ of natural objects. Examples of $P(C)$ used to obtaining the ICs are shown in Fig. 4a. The leading 100 basis vectors for the curvature map, sorted again by the method proposed in [44], are shown in Fig. 4b. In contrast to the results for the distance-based analysis (Fig. 3b), the basis vectors for the curvature map show localized Gabor-like properties similar to those found for natural images [22, 28]. This result is consistent with experimental work that identified oriented Gabor filters as good predictors for the filter properties of S1 neurons [46, 47].

Even more importantly, the same analysis applied to data for human-made objects (Fig. 4c) does not generate basis vectors with localized, Gabor-like patterns (Fig. 4d). Instead, basis vectors estimated from human-made objects have very regular characteristics, extending across the entirety of the patches and are even sometimes curved. Notably, these basis vectors seem to mimic the geometrical properties of the original curvature patches sampled from human-made objects (see Fig. 4c and 4d). Thus, although the first-order statistics of natural and human-made objects are similar (Fig. 2), these objects clearly differ in their second-order statistics.

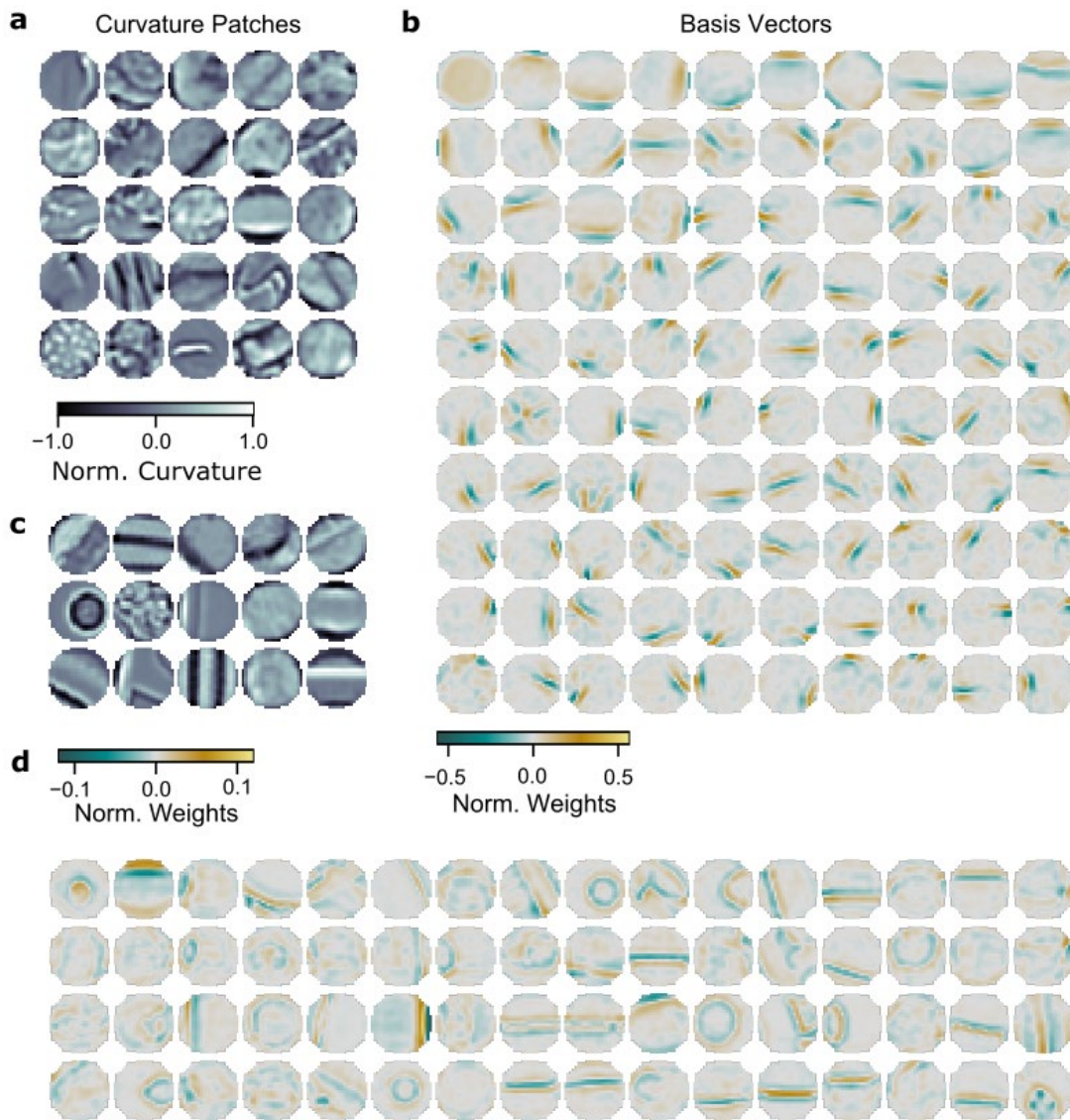


Fig. 4. Basis vectors estimated by ICA for surface patches characterized by curvature C . **(a)** Example surface patches ($R = 6$ mm) randomly sampled from natural 3D data that show spatial maps of the curvature C . Grey scale is normalized to the maximum absolute value of the curvature within each patch. **(b)** The leading 100 basis vectors (columns of mixing matrix A) as estimated by ICA from the curvature C of natural 3D data shown in (a). The mean of each basis vector was subtracted and the norm of the basis vectors was scaled to 1. **(c)** Same as (a) but for human-made 3D data. **(d)** Subset of 64 basis vectors (columns of mixing matrix A) as estimated by ICA from curvature C of human-made 3D data shown in (c). The mean of each basis vector was subtracted and the norm of the basis vectors was scaled to 1.

2.4. Filter properties of S1 model neurons

To test whether the curvature-based basis vectors identified above might be associated with filter properties of S1 neurons, we used the neural model of Bensmaia et al. [46] to investigate the orientation selectivity of the corresponding ICA filters (rows of unmixing matrix W). The proposed neural model [46] is described by the equation:

$$R(x) = |\alpha (f_{\phi} * x) + \beta| \quad \text{eq. (3)}$$

where x is the stimulus, $R(x)$ is the neural response, f_ϕ is a filter characterized by a set of parameters ϕ , and α and β are scalar coefficients. The $*$ operator symbolizes convolution; the parameters ϕ , α , and β are fit to model the response of each neuron. Bensmaia et al. [46] showed that Gabor filters for f_ϕ accounted for 57% to 68% of the variance in individual neural responses.

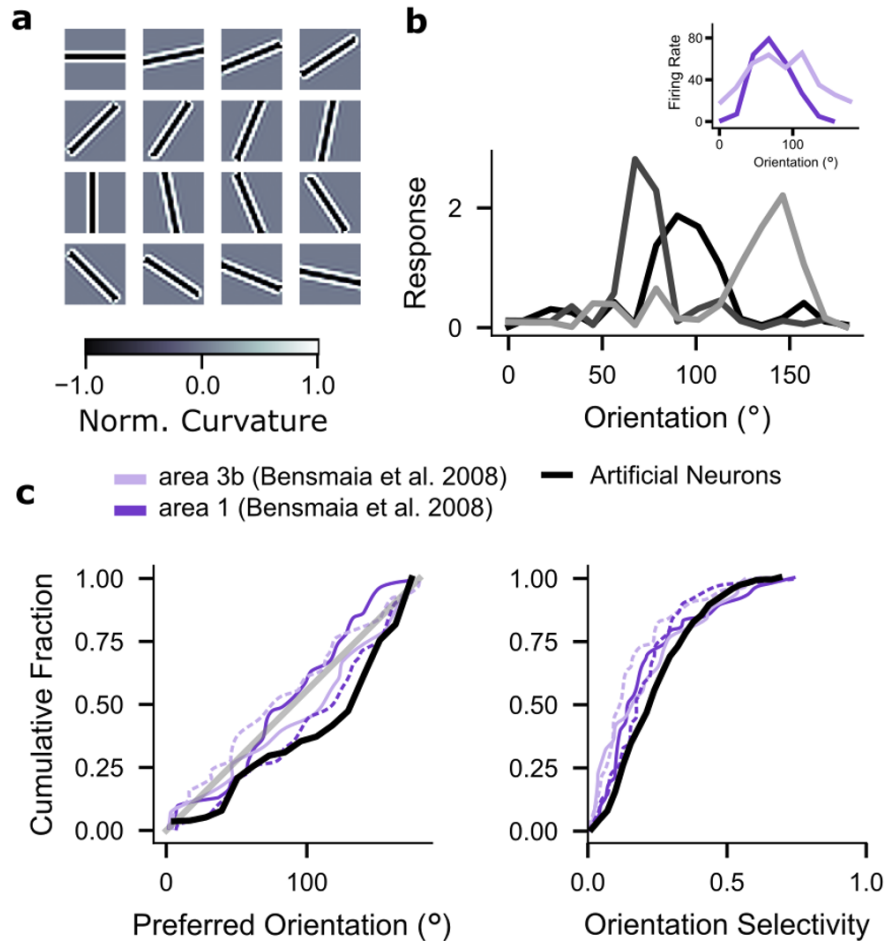


Fig. 5. Orientation selectivity of model neurons based on ICA filters. (a) Curvature maps of the bar stimuli used to test artificial neurons based on ICA filters for the curvature analysis of natural objects. The orientation varies between 0 (top left) and 168.75 degrees (bottom right) in steps of 11.25 degrees. (b) Responses of three example artificial neurons with the highest orientation selectivity as a function of bar orientation in degrees. *Inset*: Example neurons in area 1 and 3b of S1 from experiments conducted by Bensmaia et al. [46]. (c) Cumulative distributions of neural responses across all 384 artificial neurons, each based on one of the ICA filters. *Left*: Cumulative fraction of artificial neurons as a function of their preferred orientation (solid black). Experimental results from area 1 and area 3b in S1 by Bensmaia et al. are shown in purple for bars indented into the fingertip (solid lines) and bars scanned over the fingertip (dashed lines). The 45-degree line representing a uniform distribution is shown in solid grey. *Right*: Same as *left* but cumulative fraction as a function of orientation selectivity.

To test our estimated filters, we generated model neurons by replacing f_ϕ with each of the rows of W as estimated by ICA, and set $\alpha = 1$ and $\beta = 0$. This procedure resulted in a population of 384 artificial neurons, each one based on one of the estimated ICA filters corresponding to the basis vectors for the curvature analysis of natural objects (the leading 100 of these filters are shown in Fig. 4b). To characterize the response of these artificial neurons we generated a set of bar stimuli with 16 different orientations between 0 and 180 degrees, similar to those used in the experiments of Bensmaia et al. (Fig. 5a). These stimuli, represented in terms of curvature, were used as input to the artificial neurons. For each artificial neuron, we measured the preferred direction, defined as the orientation resulting in maximum neural response, and the orientation selectivity, defined

as the ratio of the neuron's response in the preferred direction to its response in all other directions (see *Methods*).

Fig. 5b shows three artificial neurons that are strongly orientation selective (95th percentile); they have preferred directions at approximately 65, 90, and 145 degrees. Examples of cortical neurons in area 1 and 3b from Bensmaia et al. [46] are shown in the inset for comparison. The cumulative fraction of the preferred orientation across all artificial neurons is shown in the left panel of Fig. 5c (black). The experimental results by Bensmaia et al. [46] are shown in purple, with solid lines indicating neural responses elicited by bars indented into the fingertip and dashed lines indicating neural responses activated by bars scanned across the fingertip. The selectivity of the artificial neurons closely approximates the experimental data in Bensmaia et al. [46]. In the right panel of Fig. 5c, we compare the cumulative fraction of orientation selectivity of the artificial neurons (black) with that of the recorded S1 neurons (purple, same as in left panel). Although the model neurons have not been fit to any neural data and result only from the statistical analysis of curvatures in natural objects, the cumulative fraction of orientation selectivity shows trends similar to those of neurons recorded in areas 1 and 3b in S1 [46].

3. Discussion

The statistical properties of visual scenes have been extensively studied, including studies of the global statistics of images in terms of luminance [6, 40, 48], contrast [6, 40, 49], and color [50, 51], as well as contours [52, 53], occlusion [52, 54-56], and eye movement [52]. However, all these studies primarily focus on the statistics of two-dimensional (2D) images, as proxies for 2D representations of the retinal image. Only a few attempts have been made to capture and describe the statistics of the 3D visual world [41].

For somatosensation, it is crucial to consider the 3D geometry of the environment with which the sensors interact. The work presented here has taken the first steps to quantify the statistics of 3D natural and human-made tactile scenes at the scale of human touch. The first-order statistics of local shape measurements (distance, slope, curvature) follow trends similar to those previously quantified for both camera and range images. Previous work showed that both these types of images exhibit different first-order statistics for natural (vegetation) and human-made (urban) scenes, mainly showing higher kurtosis for human-made scenes [55]. Other work showed that spatial low-frequency components are much more prominent in human-made scenes than in natural scenes [57]. We observed similar trends in our data.

The ICA analysis of higher order statistics reveals that basis vectors with second-order filter characteristics contribute significantly to accurate reconstruction of the data. These results suggest that curvature is an important metric to represent the 3D shape of objects at the spatial scale of a fingertip. Previous experimental work supports this hypothesis. Experiments in monkeys have shown that the spike rate of cutaneous mechanoreceptors in the finger pad respond to the amount of curvature and its rate of change, both for slowly adapting and rapidly adapting afferent fibers [33, 36, 38]. In addition, areas in human somatosensory cortex have been found to be active during a shape discrimination task that specifically tested for surface curvature [58]. A study in macaque inferior temporal cortex found neurons that are tuned to the amount of curvature and to curvature direction [59]. Correlations between curvature direction and neural activity were also found in a study that examined shape coding in visual V4 and somatosensory cortices, suggesting that these areas use similar mechanisms to encode shape [60].

The ICA of natural curvature data revealed basis vectors with Gabor-like characteristics similar to those found for natural images [45]. Further analysis with a simple model based on artificially generated neurons that use the IC filters showed neural responses that accurately capture the statistics of S1 neural responses in primates [46]. These results suggest that efficient coding algorithms such as ICA may yield insights about the filter properties of cortical neurons that go beyond those identified by receptive field measurements [46].

Although the present analysis tactile scenes has focused on the scale relevant to human touch, the methods used are not restricted to that scale or to a specific strategy for sampling the environment. Our approach can be extended to patches of different sizes explored with different degrees of spatial resolution. As such, the data presented in this study is best considered as a small subset of the data that an organism might acquire over a lifetime or even over generations. Another important extension to the framework developed here is to incorporate the temporal component that is crucial to tactile systems. Movement allows animals to gather information over time, as tactile sensors are moved over a surface. Here we have limited our analysis to a static component that

describes the aggregated tactile information obtained over such a sweep by concentrating on the intrinsic properties of object patches.

This work examined the statistics of three metrics: distance, slope, and curvature, as intrinsic properties of external objects; patches were randomly sampled from each objects' surface. However, animals and humans sample the world with a goal in mind (such as identifying an object) and combine past and current sensory information both to predict future sensory input as well as increase search efficiency. Therefore, we expect that subjects would sample a set of patches that differs from the randomly sampled ones used here. This hypothesis could be tested by asking human subjects to identify objects based only on tactile input in a situation where the object is behind a screen and thus visually inaccessible. To approximate the conditions of the work presented here, the subject would be allowed to use only one finger to touch the surface of the objects. An efficient exploration could be enforced by limiting the time available for object identification. The objects would be the same ones used in the presented study; here we have characterized the statistics $P(W)$ of these objects as representative of the world W . In the experiment with human subjects, the places on the object's surface explored by the subject as they try to identify the object would be tracked. This data would allow for a characterization of the statistics of the world when sampled by volitional exploration E . This $P(W|E)$ could then be compared to the $P(W)$ analyzed here.

4. Methods

4.1. Data acquisition

We used a *EinScan Pro 2X* scanner to capture 3D point cloud representations of the surface of both natural objects and human-made objects. The scanner was mounted on a *Shining 3D* desktop tripod while the objects were placed on the *Shining 3D* turntable. The *EXScan Pro* software was used to capture the scans in "Fixed Scan" mode as "Non-Texture Scans" (without color information). For each scan, the turntable performed 30-50 steps for one rotation and one rotation per scan. Typically, each object was scanned in 2-3 different orientations and the scans were then combined into a single point cloud. One of the objects was very small in size and was therefore scanned using "Handheld Rapid Scan" mode with highest resolution.

4.2. Data cleaning

Each point cloud was imported to *Geomagic Design X* software for cleaning and meshing. First, the standard meshing algorithm of *Geomagic Design X (Mesh Buildup Wizard™)* was used to triangulate the point cloud. Outlier data (non-manifold triangles, small clusters, and isolated triangles) were removed manually or using the *Healing Wizard* software. Holes in the scan were filled using the *Fill Holes* tool. For larger holes or for holes at the boundary, the *Add Bridge* tool was used to segment the holes to ensure better reconstruction with the subsequent use of the *Fill Holes* tool. After all holes were filled and unusable parts of the mesh had been removed, the scan was remeshed targeting an average edge length of 0.1mm. After remeshing, the mesh was enhanced (smoothed and sharpened) with medium settings and optimized to improve curvature flow at medium settings with a maximum of 10 iterations. Finally, the centroid of the mesh was redefined to be the origin and exported in *Binary STL* format, which stores each triangle as a collection of three edge points and the corresponding face normal. The resulting dataset consisted of 137 scans, of which 96 represented natural scenes and 41 represented human made objects (SI Dataset).

4.3. Sampling and computation of shape metrics

The method used to compute the distance (D), slope (S), and curvature (C) metrics is illustrated in Fig. 6. First, we randomly sampled 100-150 circular surface patches of radius R from each object (Fig. 6a). The point cloud for each sampled surfaced patch was analyzed using Principal Component Analysis (PCA). The plane spanned by the two leading PCs was found, and the points in the surface cloud were projected onto this reference plane (Fig. 6b). The location of the projected points within the plane was digitized to a 24x24 grid (Fig. 6c). The position of the grid cell in the i^{th} row and j^{th} column is denoted as $p_{i,j}$; the width of the grid cells is $\delta = R/12$.

The plane was then translated in the direction of its outward normal vector, away from the surface of the object just until all surface sampled points were below it. For each point (i, j) on the surface cloud, the distance $D_{i,j}$ to the reference plane was computed as the orthogonal distance from that point to its projection (Fig. 6d). The distances D were then interpolated over the entire grid, smoothed, and cropped to a circular patch of radius R to

obtain a map $P(D)$ over 448 grid points. Patches that filled less than 75% of the circular area before interpolation were rejected.

The metrics S and C were computed from the interpolated and smoothed square grid before cropping it to a circular patch. The slope S was calculated as the forward difference in D between neighboring points in the grid (Eq. 1), resulting in a circular map of 421 slope values $P(S)$. Similarly, C was calculated as the second-order central difference in D (Eq. 2), resulting in a circular map of 384 curvature values $P(C)$. The resulting maps for one example patch are shown in Fig. 6e.

Since the square grid size is fixed at 24x24 before cropping it to a circular patch, the spatial resolution of the resulting circular maps is determined by the patch radius R . Each circular patch can be viewed as a sensory surface with a fixed number of receptors represented by the grid. A larger R leads to a grid that covers a larger area sampled at a lower spatial resolution. Based on the average contact area of a finger pad ($\sim 120 \text{ mm}^2$ at 1N normal load [39]), we chose R to be 6 mm (corresponding to 113 mm^2 of circular contact area) for the results presented here. The choice of a grid size of 24x24 is based on the average edge length of the 3D meshes of the data ($\sim 0.1 \text{ mm}$) rather than on biologically plausible receptor density. With an R of 6 mm, the datasets for natural and human-made objects consisted of a total of 12,309 and 5,788 patches, respectively.

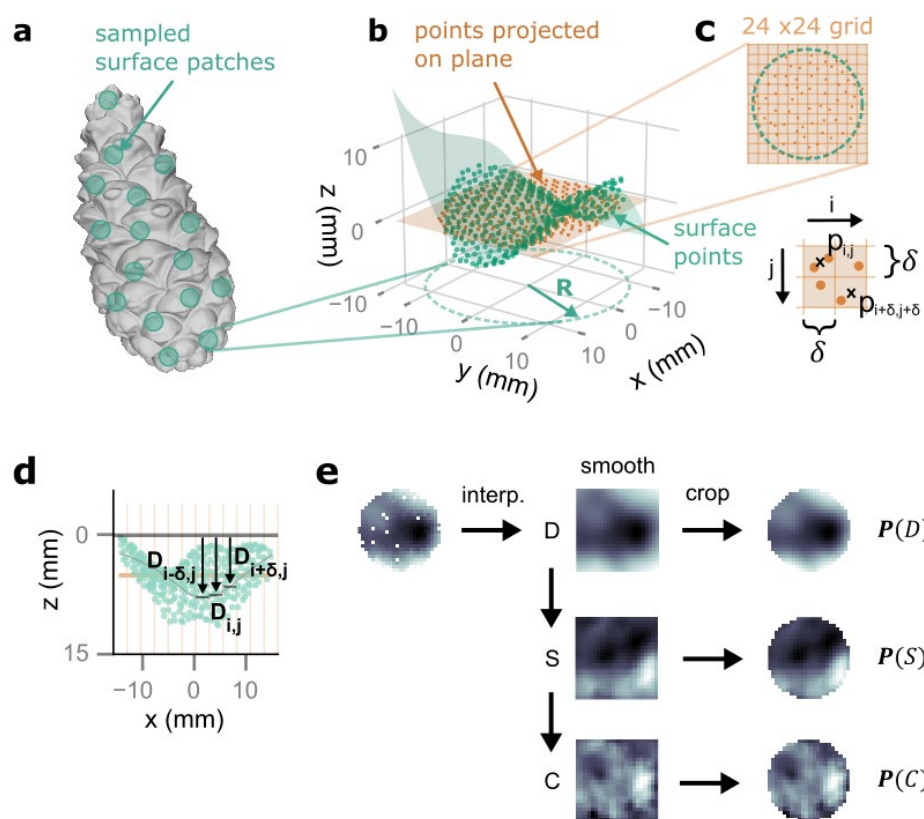


Fig. 6. Processing of 3D data. **(a)** Approximately 100-150 circular surface patches of radius R were randomly sampled from each object scan. **(b)** For each surface patch, the average plane of the point cloud (orange) was found by using Principal Component Analysis (PCA). **(c)** The points in the cloud were then projected onto the plane, and their location within the plane digitized to a 24x24 grid. The position of the grid cell in the j^{th} row and i^{th} column was indicated as $p_{i,j}$. The cells are of linear size $\delta = R/12$. **(d)** For each cell (i,j) the distance $D_{i,j}$ to the reference plane describes the local shape of the patch. **(e)** The distances D were interpolated to the entire grid, smoothed, and cropped to a circular patch. The metrics slope S and curvature C were computed from the interpolated and smoothed grid before cropping it to a circular patch.

Because the slope of a surface depends on the direction along which it is taken, its value for a given grid cell was computed by averaging the slope in the x direction ($S_{i,j}^x$) and the slope in the y direction ($S_{i,j}^y$). The equations for $S_{i,j}^x$ and $S_{i,j}^y$ are:

$$\begin{aligned} S_{i,j}^x &= \frac{D_{i+\delta,j}-D_{i,j}}{\delta} \\ S_{i,j}^y &= \frac{D_{i,j+\delta}-D_{i,j}}{\delta} \end{aligned} \quad (\text{Eq. 1})$$

The curvature of a surface is also direction dependent, and was again computed as the average of the curvature in the x direction ($C_{i,j}^x$) and the curvature in the y direction ($C_{i,j}^y$). The equations for $C_{i,j}^x$ and $C_{i,j}^y$ are:

$$\begin{aligned} C_{i,j}^x &= \frac{D_{i+\delta,j}-2D_{i,j}+D_{i-\delta,j}}{\delta^2} \\ C_{i,j}^y &= \frac{D_{i,j+\delta}-2D_{i,j}+D_{i,j-\delta}}{\delta^2} \end{aligned} \quad (\text{Eq. 2})$$

4.4. Neural model

The orientation selectivity os is computed as:

$$os = \frac{\sqrt{[\sum_i R(\theta_i) \sin(2\theta_i)]^2 + [\sum_i R(\theta_i) \cos(2\theta_i)]^2}}{\sum_i R(\theta_i)} \quad (\text{Eq. 3})$$

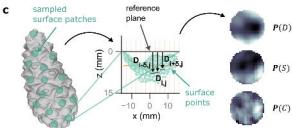
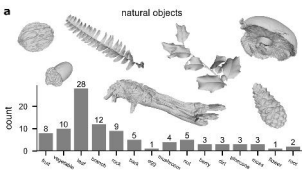
where $R(\theta_i)$ represents the response of the artificial neuron to a bar stimulus with orientation θ_i . Stimulus orientation ranged from 0 to 180 degrees in intervals of 11.25 degrees, for $i = \{1, 2, \dots, 16\}$. The orientation selectivity of a neuron ranges from 0 to 1, where 0 indicates a uniform response to all orientations and 1 indicates that the neuron responds only to a single stimulus orientation. For more details about this measure of orientation selectivity see [46].

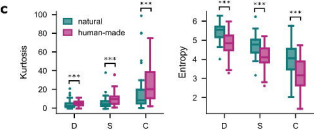
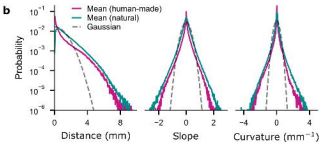
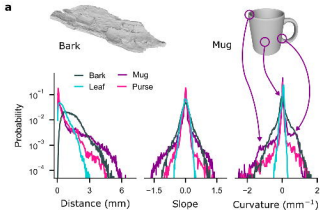
References:

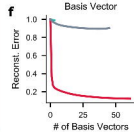
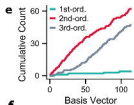
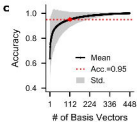
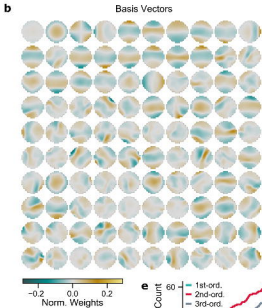
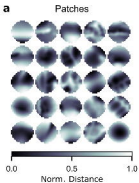
1. Attneave, F., *Some informational aspects of visual perception*. Psychological Review, 1954. **61**(3): p. 183-193.
2. Barlow, H.B., *Possible principles underlying the transformations of sensory messages*. Sensory communication, 1961. **6**(2): p. 57-58.
3. Olshausen, B.A. and D.J. Field, *Emergence of simple-cell receptive field properties by learning a sparse code for natural images*. 1996. p. 607-609.
4. Lewicki, M.S., *Efficient coding of natural sounds*. Nature neuroscience, 2002. **5**(4): p. 356-363.
5. Hyvärinen, A., *Statistical models of natural images and cortical visual representation*. Topics in Cognitive Science, 2010. **2**(2): p. 251-264.
6. Field, D.J., *Relations between the statistics of natural images and the response properties of cortical cells*. Journal of the Optical Society of America, 1987. **4**(12): p. 2379-2394.
7. Ruderman, D.L., T.W. Cronin, and C.-C. Chiao, *Statistics of cone responses to natural images: implications for visual coding*. Journal of the Optical Society of America A, 1998. **15**(8): p. 2036-2045.
8. Atick, J.J. and A.N. Redlich, *What does the retina know about natural scenes?* Neural computation, 1992. **4**: p. 196-210.
9. Barnett, P.D., K. Nordstrom, and D.C. O'Carroll, *Motion adaptation and the velocity coding of natural scenes*. Current Biology, 2010. **20**(11): p. 994-9.
10. Hausler, C., A. Susemihl, and M.P. Nawrot, *Natural image sequences constrain dynamic receptive fields and imply a sparse code*. Brain Research, 2013. **1536**: p. 53-67.
11. Stansbury, D.E., T. Naselaris, and J.L. Gallant, *Natural scene statistics account for the representation of scene categories in human visual cortex*. Neuron, 2013. **79**(5): p. 1025-34.
12. Dyakova, O. and K. Nordstrom, *Image statistics and their processing in insect vision*. Current Opinion in Insect Science, 2017. **24**: p. 7-14.
13. Samonds, J.M., W.S. Geisler, and N.J. Priebe, *Natural image and receptive field statistics predict saccade sizes*. Nature Neuroscience, 2018. **21**: p. 1591-1599.
14. Smith, E.C. and M.S. Lewicki, *Efficient auditory coding*. Nature, 2006. **439**(7079): p. 978-82.
15. Parise, C.V., K. Knorre, and M.O. Ernst, *Natural auditory scene statistics shapes human spatial hearing*. Proceedings of the National Academy of Sciences, 2014. **111**(16): p. 6104-8.
16. Traer, J. and J.H. McDermott, *Statistics of natural reverberation enable perceptual separation of sound and space*. Proceedings of the National Academy of Sciences, 2016. **113**(48): p. E7856-E7865.
17. Huang, N. and M. Elhilali, *Auditory salience using natural soundscapes*. The Journal of the Acoustical Society of America, 2017. **141**(3): p. 2163.
18. Dodds, E.M. and M.R. DeWeese, *On the sparse structure of natural sounds and natural images: Similarities, differences, and implications for neural coding*. Frontiers in Computational Neuroscience, 2019. **13**: p. 39.
19. Evans, M.H., *Efficient coding in the whisker system: biomimetic pre-processing for robots?*, in *Biomimetic and Biohybrid Systems: Second International Conference, Living Machines 2013, London, UK, July 29 -- August 2, 2013. Proceedings*, N.F. Lepora, et al., Editors. 2013, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 364-367.
20. Hafner, V.V., et al., *Optimal coding for naturally occurring whisker deflections*. Artificial Neural Networks and Neural Information Processing - Ican/Iconip 2003, 2003. **2714**: p. 805-812.
21. Manfredi, L.R., et al., *Natural scenes in tactile texture*. Journal of neurophysiology, 2014. **111**: p. 1792-802.
22. Olshausen, B.A. and D.J. Field, *Emergence of simple-cell receptive field properties by learning a sparse code for natural images*, in *Nature*. 1996. p. 607-609.
23. Hyvärinen, A. and P. Hoyer, *Emergence of phase- and shift-invariant features by decomposition of natural images into independent feature subspaces*. Neural Computation, 2000. **12**(7): p. 1705--1720.
24. Hoyer, P.O. and A. Hyvärinen, *Independent component analysis applied to feature extraction from colour and stereo images*. Network: Computation in Neural Systems, 2000. **11**(3): p. 191--210.
25. Hyvarinen, A., M. Gutmann, and P.O. Hoyer, *Statistical model of natural stimuli predicts edge-like pooling of spatial frequency channels in V2*. BMC Neurosci, 2005. **6**: p. 12.
26. Hyvärinen, A. and P.O. Hoyer, *A two-layer sparse coding model learns simple and complex cell receptive fields and topography from natural images*. Vision Research, 2001. **41**(18): p. 2413-2423.

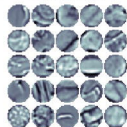
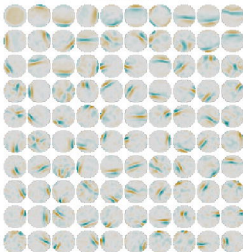
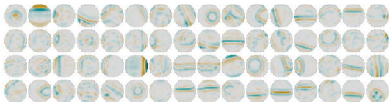
27. Van Hateren, J.H. and D.L. Ruderman, *Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex*. Proceeding The Royal Society, 1998. **265**: p. 2315-2320.
28. van Hateren, J.H. and A. van der Schaaf, *Independent component filters of natural images compared with simple cells in primary visual cortex*. Proceedings of the Royal Society B-Biological Sciences, 1998. **265**(1394): p. 359-366.
29. Hobbs, J.A., R.B. Towal, and M.J. Hartmann, *Probability distributions of whisker-surface contact: quantifying elements of the rat vibrissotactile natural scene*. Journal of Experimental Biology, 2015. **218**(Pt 16): p. 2551-62.
30. Saal, H.P., et al., *Simulating tactile signals from the whole hand with millisecond precision*. Proceedings of the National Academy of Sciences of the United States of America, 2017: p. 201704856-201704856.
31. Shao, F., et al., *A novel tactile sensation measurement system for qualifying touch perception*. Proceedings of the Institution of Mechanical Engineers, Part H, 2010. **224**(1): p. 97-105.
32. Sakamoto, M. and J. Watanabe, *Exploring tactile perceptual dimensions using materials associated with sensory vocabulary*. Frontiers in Psychology, 2017. **8**: p. 569.
33. Goodwin, A.W., V.G. Macefield, and J.W. Bisley, *Encoding of object curvature by tactile afferents from human fingers*. Journal of Neurophysiology, 1997. **78**(6): p. 2881-2888.
34. Cheeseman, J.R., J.F. Norman, and A.M. Kappers, *Dynamic cutaneous information is sufficient for precise curvature discrimination*. Scientific Reports, 2016. **6**: p. 25473.
35. van der Horst, B.J. and A.M. Kappers, *Using curvature information in haptic shape perception of 3D objects*. Experimental brain Research, 2008. **190**(3): p. 361-7.
36. Khalsa, P.S., et al., *Encoding of shape and orientation of objects indented into the monkey fingerpad by populations of slowly and rapidly adapting mechanoreceptors*. Journal of Neurophysiology, 1998. **79**(6): p. 3238-3251.
37. Lamotte, R.H., et al., *Cutaneous neural codes for shape*. Canadian Journal of Physiology and Pharmacology, 1994. **72**(5): p. 498-505.
38. Srinivasan, M.V. and R.H. Lamotte, *Encoding of shape in cutaneous mechanoreceptors*, in *Information processing in the somatosensory system*, O. Franzen and J. Westman, Editors. 1991, Macmillan International Higher Education.
39. van Kuilenburg, J., M.A. Masen, and E. van der Heide, *A review of fingerpad contact mechanics and friction and how this affects tactile perception*. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 2013. **229**(3): p. 243-258.
40. Ruderman, D.L., *The statistics of natural images*. Network: Computation in Neural Systems, 1994. **5**: p. 517-548.
41. Huang, J., A.B. Lee, and D. Mumford, *Statistics of range images*. Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2000. **1**: p. 324-331 vol.1.
42. Bell, A.J. and T.J. Sejnowski, *An information-maximization approach to blind separation and blind deconvolution*. Neural Computation, 1995. **7**(6): p. 1129-1159.
43. Hyvärinen, A. and E. Oja, *Independent component analysis: algorithms and applications*. Neural Networks, 2000. **13**(4): p. 411-430.
44. Cheung, Y.-m. and L. Xu, *Independent component ordering in ICA time series analysis*. Neurocomputing, 2001. **41**(1): p. 145-152.
45. Olshausen, B.A. and D.J. Field, *Strategy Employed by V1 ?* Vision Research, 1997. **37**(23): p. 3311-3325.
46. Bensmaia, S.J., et al., *The representation of stimulus orientation in the early stages of somatosensory processing*. The Journal of Neuroscience, 2008. **28**(3): p. 776.
47. DiCarlo, J.J. and K.O. Johnson, *Spatial and temporal structure of receptive fields in primate somatosensory area 3b: Effects of stimulus scanning direction and orientation*. The Journal of Neuroscience, 2000. **20**(1): p. 495.
48. Fine, I., D.I. MacLeod, and G.M. Boynton, *Surface segmentation based on the luminance and color statistics of natural scenes*. Journal of the Optical Society of America. A, Optics, image science, and vision, 2003. **20**(7): p. 1283-91.
49. Brady, N. and D. Field, *Local contrast in natural images: normalisation and coding efficiency*. Perception, 2000. **29**(9): p. 1041-1055.
50. Nascimento, S.M.C., K. Amano, and D.H. Foster, *Spatial distributions of local illumination color in natural scenes*. Vision Research, 2016. **120**: p. 39-44.

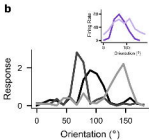
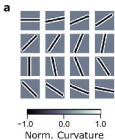
51. Su, C.C., et al. *Natural scene statistics of color and range*. in *18th IEEE International Conference on Image Processing (ICIP)*. 2011. Brussels, BELGIUM.
52. Geisler, W.S. and J.S. Perry, *Contour statistics in natural images: grouping across occlusions*. *Visual Neuroscience*, 2009. **26**(1): p. 109-21.
53. Geisler, W.S., et al., *Edge co-occurrence in natural images predicts contour grouping performance*. *Vision Research*, 2001. **41**: p. 711–724.
54. Huang, J., A.B. Lee, and D.B. Mumford, *Occlusion models for natural images: A statistical study of a scale-invariant dead leaves model*. *International Journal of Computer Vision*, 2001. **41**(1).
55. Lee, A., J. Huang, and D. Mumford, *Random-collage model for natural images*. *International Journal of Computer Vision*, 2000.
56. Vilankar, K.P., et al., *Local edge statistics provide information regarding occlusion and nonocclusion edges in natural scenes*. *Journal of Vision*, 2014. **14**(9).
57. Torralba, A. and A. Oliva, *Statistics of natural image categories*. *Network: Computation in Neural Systems*, 2003. **14**(3): p. 391-412.
58. Bodegård, A., et al., *Hierarchical processing of tactile shape in the human brain*. *Neuron*, 2001. **31**(2): p. 317-328.
59. Kayaert, G., et al., *Tuning for shape dimensions in macaque inferior temporal cortex*. *European Journal of Neuroscience*, 2005. **22**(1): p. 212-24.
60. Pasupathy, A., et al., *Analogous intermediate shape coding in vision and touch*. *Proceedings of the National Academy of Sciences*, 2009. **106**(38): p. 16457-16462.







a Curvature Patches**b** Basis Vectors**c**



c

— area 3b (Bensmaia et al. 2008)

— area 1 (Bensmaia et al. 2008)

— Artificial Neurons

