LETTERS

Image statistics and the perception of surface qualities

Isamu Motoyoshi[1], Shin'ya Nishida[1], Lavanya Sharan[2] & Edward H. Adelson[2]

The world is full of surfaces, and by looking at them we can judge their material qualities. Properties such as colour or glossiness can help us decide whether a pancake is cooked, or a patch of pavement is icy. Most studies of surface appearance have emphasized textureless matte surfaces[1–3], but real-world surfaces, which may have gloss and complex mesostructure, are now receiving increased attention[4–7]. Their appearance results from a complex interplay of illumination, reflectance and surface geometry, which are difficult to tease apart given an image. If there were simple image statistics that were diagnostic of surface properties it would be sensible to use them[8–11]. Here we show that the skewness of the luminance histogram and the skewness of sub-band filter outputs are correlated with surface gloss and inversely correlated with surface albedo (diffuse reflectance). We find evidence that human observers use skewness, or a similar measure of histogram asymmetry, in making judgements about surfaces. When the image of a surface has positively skewed statistics, it tends to appear darker and glossier than a similar surface with lower skewness, and this is true whether the skewness is inherent to the original image or is introduced by digital manipulation. We also find a visual aftereffect based on skewness: adaptation to patterns with skewed statistics can alter the apparent lightness and glossiness of surfaces that are subsequently viewed. We suggest that there are neural mechanisms sensitive to skewed statistics, and that their outputs can be used in estimating surface properties.

Figure 1 shows two renderings of a three-dimensional model of Michelangelo's sculpture of St Matthew[12]. The version on the left appears darker and glossier than the one on the right. This is true even though the two images have been scaled to have the same mean luminance. We are unaware of any theories that will predict the changes in lightness or gloss that we observe.

The image of a surface arises from the combination of the surface geometry, the surrounding illumination, and the surface optics. Each of these components can be complex (for example, the reflectance at each point is characterized by a four-dimensional function known as the bidirectional reflectance distribution function[13]). Each is typically unknown, and estimating any one using 'inverse optics' requires knowing the others. To bypass this problem, we have looked for simple statistical image measurements that can provide information that is useful even if not complete. Any two-dimensional image measurements that are statistically related to properties of the three-dimensional scene are potentially useful[8–11].

We made a set of patches of stucco-like material. The values of albedo and glossiness were uniform within each patch, but they were varied systematically from one patch to another by changing paint pigmentation and acrylic media coating, respectively. We photographed these objects, linearized the pixel values and normalized the mean luminance by multiplicative scaling. We found that changes in albedo and glossiness were accompanied by characteristic changes in the luminance histogram. Consider the two stucco patches of

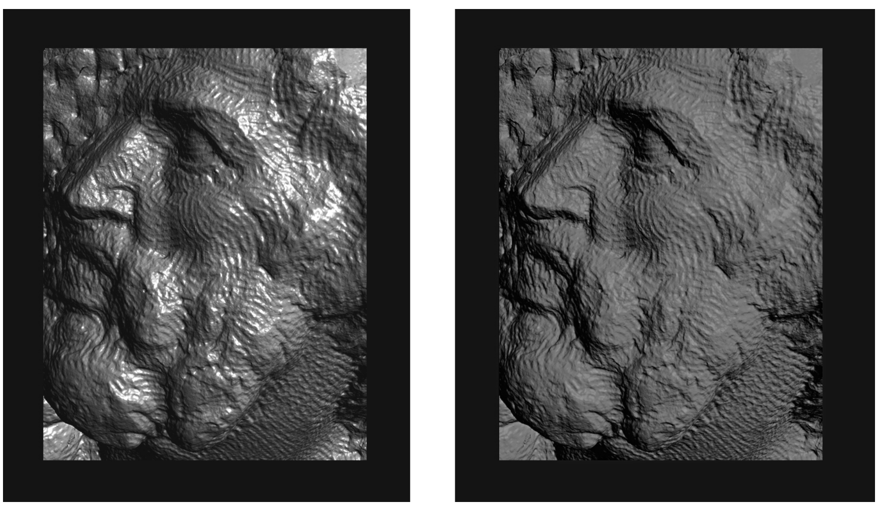

Figure 1 | These two synthetic images of Michelangelo's St Matthew sculpture have the same mean luminance. The one on the left looks darker and glossier than the one on the right.

[1]Human and Information Science Lab, NTT Communication Science Labs, Nippon Telegraph and Telephone Corporation, 3-1 Morinosato-Wakamiya, Atsugi 243-0198, Japan. [2]Department of Brain and Cognitive Sciences and Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 43 Vassar Street, 46-4115, Cambridge, Massachusetts 02139, USA.

Fig. 2a. In comparison with a light matte surface (left), a dark glossy surface (right) has a long positive tail. In general, as the albedo of glossy surfaces is decreased, or as the glossiness is increased regardless of the albedo, the histogram's skewness tends to increase (Fig. 2b; black circles). These changes make sense given the influence of specular and diffuse reflectance on the appearance of specular highlights. Highlights are stronger and sharper on glossy surfaces, and they have higher contrast when viewed on darker surfaces, because they are seen against a body surface that has a lower luminance.

Having observed this physical relationship, we next looked for a corresponding psychophysical relationship. We showed these stucco images, one by one, to human observers, presenting them against a dark background on a monitor at constant mean luminance, and asked the observers to rate the lightness (perceived diffuse reflectance) or glossiness of each surface. The judgments were well correlated with the corresponding physical properties, as shown in Fig. 2b (red circles). Both the lightness and glossiness ratings were also well correlated with the skewness of the luminance histogram (Fig. 2c) to a degree comparable with, or even higher than, the correlations with corresponding physical properties ($r = -0.87$ for correlation with skewness of lightness ratings, and 0.89 for glossiness ratings, respectively).

We next chose a set of images of three materials (stucco, black cotton fabric and crumpled white paper, all of which were surfaces of uniform albedo and glossiness) and used a lookup table to force the luminance histograms to have specific skewness values. As expected, the lightness rating showed a strong negative dependency on skewness, whereas the glossiness rating showed a strong positive dependency. This was true for each image class (Fig. 2d). Further tests of a wide variety of materials gave similar results, described in Supplementary Data A.

In addition to the effects of skewness, we found a minor effect of the standard deviation of the luminance histogram on both lightness and glossiness. The mean luminance had a significant effect on lightness[1–3], but not on glossiness. We found little, if any, effect of kurtosis (Supplementary Data B).

The above results indicate that skewness or a similar measure of histogram asymmetry is useful in estimating surface qualities, and that humans may indeed use it. How might such statistics be computed at the neural level? The early stages of vision are dominated by neurons that represent luminance variation in certain sub-bands of spatial frequency. These cells do not have direct access to raw luminance, but there is a strong correlation between sub-band skewness and luminance skewness for the uniform albedo surfaces we used ($r \geq 0.86$ for sub-bands at spatial frequencies ranging from 4 to 64 cycles per image, obtained with two-octave gaussian bandpass filters). We note that sub-band statistics also have certain advantages over luminance statistics, because they reflect spatial image structure whereas luminance statistics do not (see also Supplementary Data D).

Skewness is a measure of the asymmetry of a distribution; it indicates the balance between the positive and negative tails. Various definitions have been used[14], the most popular being based on the third standardized moment. If X is a random variable with zero mean and unit variance, then skewness is the expected value of X^3. If X is the output of an array of neurons that act as sub-band filters, then a neural implementation needs normalization, cubing, and summation over a region; these are easy to implement in neural hardware. A bandpass neuron's output already has zero mean, and local gain control[15] will tend to normalize the response variance over a given region. Cubing and summing over a region are straightforward.

In more concrete terms, we suggest the flow diagram of Fig. 3. As input, we use an image constructed from the two St Matthew images. The image is filtered with on-centre and off-centre receptive fields and split into separate (all positive) streams. The on-centre and off-centre responses each pass through an accelerating nonlinearity. These responses are summed over a region, and these summed responses are subtracted. The difference signal is an estimate of local

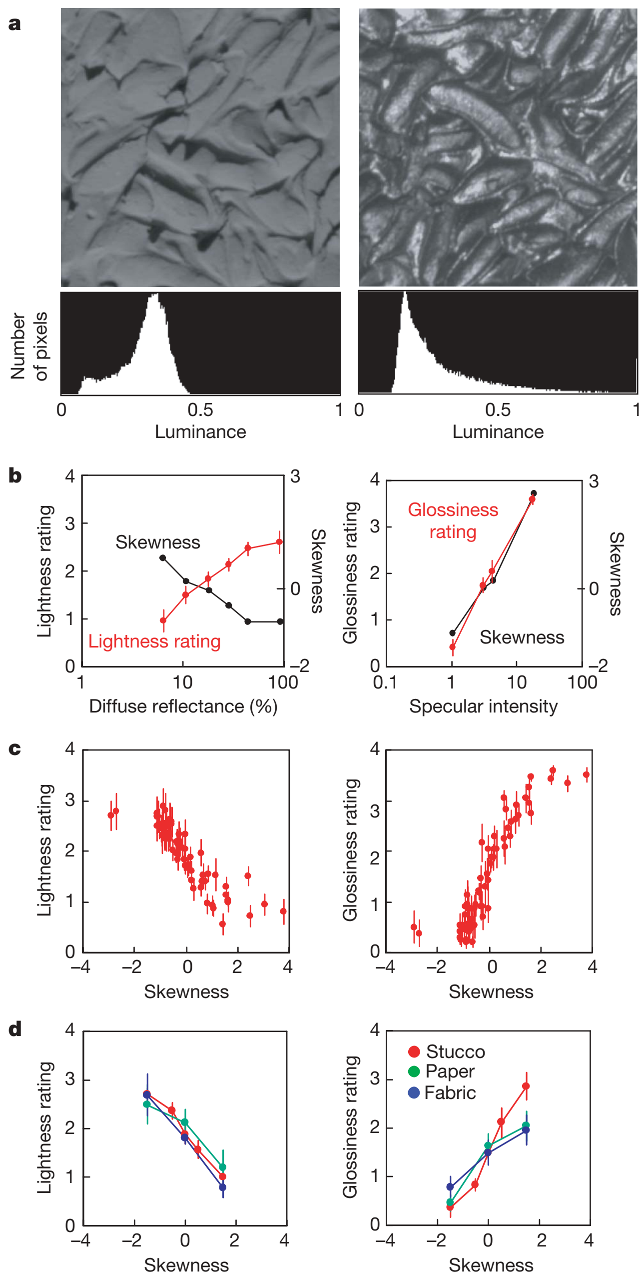

Figure 2 | Perceived lightness and glossiness may be based on the skewness of the luminance histograms. **a**, The stucco-like surface on the right looks darker and glossier than one on the left even though the mean luminance of both images is equal. We note that the luminance histogram is negatively skewed for the left image (skewness $= -1.34$), and positively skewed (skewness $= 2.40$) for the right image. **b**, In the left panel, as the diffuse reflectance increases, the lightness ratings given by human observers increase (red circles, y-axis scale on the left side) and the skewness of the image histogram decreases (black circles, y-axis scale on the right side). The data were obtained with medium-glossy surfaces. In the right panel, we see that as the specular reflectance increases, both the rated glossiness (red circles) and the histogram skewness (black circles) increase. The data were obtained with dark-grey surfaces. **c**, The rated lightness (left panel) and glossiness (right panel) of the 63 surface images of varying diffuse and specular reflectance under different illumination conditions. The human ratings correlate well with the skewness of the image histogram. Error bars represent $\pm$ 1 s.e.m across six observers. **d**, The effect of histogram manipulation on perceived lightness and glossiness. For images of stucco, crumpled paper, and fabric, the luminance histogram was matched to a positively or negatively skewed beta distribution. The rated lightness (left panel) and glossiness (right panel) varied according to the final skewness of the manipulated image (mean $= 8.2$ cd m^{-2}, s.d./mean $= 0.1$, See Supplementary Data B for data obtained with other parameters). Error bars represent $\pm$ 1 s.e.m. across seven observers.

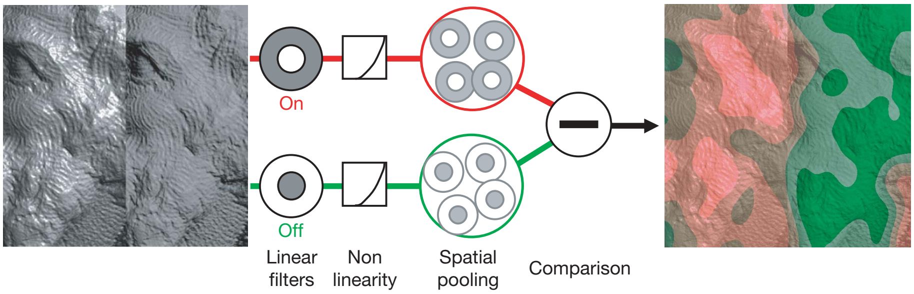

Figure 3 | A proposed neural mechanism for encoding the sub-band skewness by early visual units. The image is analysed by on-centre and off-centre filters followed by an accelerating nonlinearity (square, cube and so on). The outputs are then separately pooled over space. The difference between the pooled outputs of the on- and off-centre channels provides an estimate of the sub-band skewness.

skewness. See Supplementary Discussion for a more formal description of the model.

Could this computation be supported with known physiology? The on-centre and off-centre cells of the brain's lateral geniculate nucleus (LGN) would work as the initial stage, as would even-symmetric ('bar detector') simple cells. Next we require a cortical cell that pools and compares the outputs of a set of such cells after a nonlinearity. The classic complex cell will not suffice, because it is insensitive to contrast sign. Our putative skewness cells would be selective for contrast sign, but not for position and not necessarily for orientation. Such cells would be excited by bright (or dark) dots or lines anywhere within their receptive fields, but not by ones of the opposite sign. There are various reports of cells in areas V1 and V2 of the brain that are selective for contrast sign[16–20], and these could participate in the processing chain we are proposing.

The notion of 'skewness detectors' suggested a psychophysical experiment. Suppose we adapt to a pattern with positive skewness, and thereby shift the balance of sensitivities in the positive and negative skewness mechanisms. Would that shift our judgements of lightness and glossiness of a surface subsequently viewed?

We had subjects adapt to the patterns shown in Fig. 4a, which consisted of quasi-randomly placed blurred spots. Subjects fixated at a position between the two images. The adaptor with bright spots had positive skewness and the one with dark spots had negative skewness. After adaptation, two images of the same stucco surface that differed only in the sign of their skewness values were presented, side by side, as shown in Fig. 4b, and we asked subjects to judge their relative lightness and glossiness. Both judgments were shifted in the expected direction, as shown in Fig. 4d. We also used stucco images with positive or negative skewness as adapters, as in Fig. 4c, and found a similar shift. Thus, the effect is similar whether the adapting stimulus looks like a surface or merely a set of random spots.

We determined that the skewness adaptation did not affect the apparent brightness of the uniform grey test field of various luminance values. This ruled out the possibility that the lightness after-effect was due to a simple change in the luminance transducer function of the visual system.

We tested the inter-ocular transfer of the after-effects and found that the after-effect caused by the different-eye adaptor was 76% ± 6.1% (95% confidence interval) of that by the same-eye adaptor (estimated from the averaged results of three observers). Such incomplete transfer points to a partial involvement of monocular

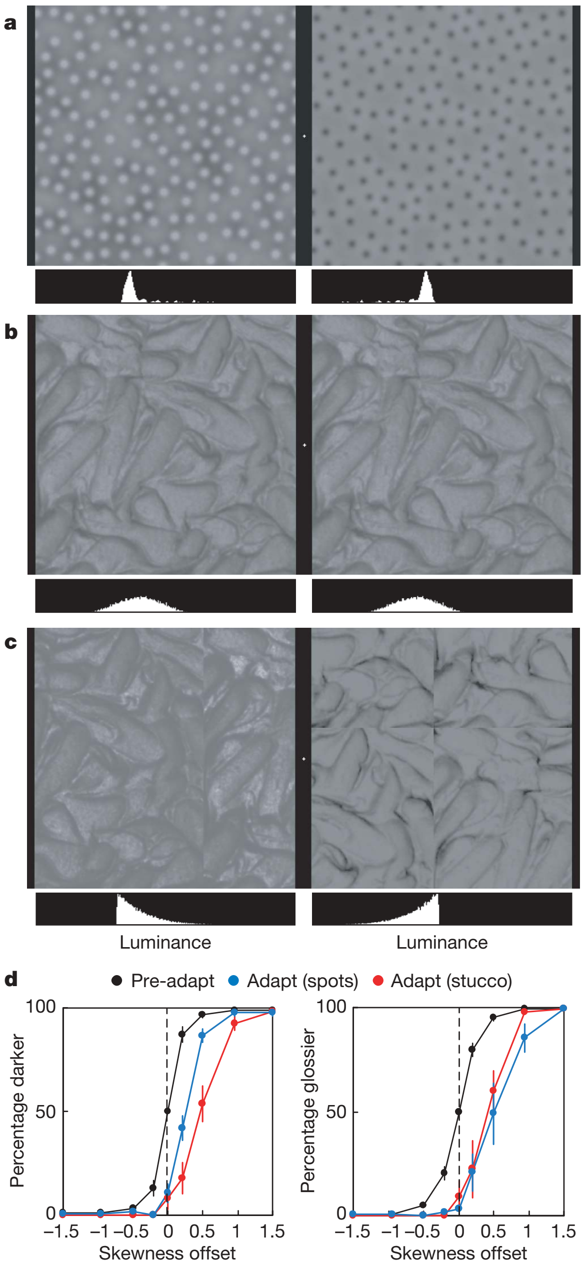

Figure 4 | After-effects of perceived lightness and glossiness. Observers adapted to the artificial textures shown in **a**. These textures consist of on-centre (left) and off-centre (right) difference-of-gaussian elements. After the prolonged observation of images in **a** observers were shown a pair of test images (**b**). Although the test surfaces shown in **b** are physically the same, observers saw the right surface as darker and glossier than the left one. Similar after-effects were obtained after adaptation to natural surfaces with skewed histograms, as shown in **c**. (See Supplementary Movies.) **d**, The probability that the subjects judged one test image (on the side that adapted to positive skewness), relative to the other test surface (on the side that adapted to negative skewness) as darker (lightness judgement, left panel) or glossier (glossiness judgement, right panel). Error bars represent ±1 s.e.m. across six observers. There were three adaptation conditions: no adaptation (black circles), adaptation to artificial textures (difference-of-gaussian patterns, blue circles), and adaptation to natural surfaces (stucco images, red circles). The horizontal axis indicates the difference in skewness of the luminance histogram of the two test images. Both plots in **d** correspond to the case of adaptation to positive skewness. The shift of the psychometric function indicates the after-effect.

sensors that are only evident in peripheral processing stages such as retina, LGN and V1. Thus, some of the adaptable processing stages might occur quite early in the visual system.

While skewness is predictive of perceived surface qualities, it can of course be computed on arbitrary images, whether or not they look like surfaces. A picture of fireworks against the night sky will be positively skewed, but one cannot meaningfully judge its albedo or gloss; the same is true of the adapting stimulus of Fig. 4a. Our findings were made in the case where the image is perceived as a surface of uniform albedo with some highlights. We do not know what aspects of image structure determine 'surfaceness' or 'highlights'. When our images are phase-scrambled so as to retain sub-band power, but not phase structure, they are typically seen as plausible but not convincing surfaces. The lightness effects are retained, but glossiness is lost. When the images are pixel-scrambled they are seen as two-dimensional noise patterns without a unitary albedo or gloss. These manipulations and the effects of spatial structure are discussed further in Supplementary Data C and D.

Malik and Perona[21] proposed that even-symmetric filters underlie human sensitivity to contrast sign in texture discrimination. Chubb *et al.*[22], working with unstructured random noise textures, found evidence for a 'blackshot' mechanism sensitive to dark outliers. Our stimuli (which are seen as surfaces rather than two-dimensional random patterns) and our task (judging surface quality rather than discriminating textured regions) are quite different, but the processing could involve similar computations.

The present study gives an interesting perspective on neural computation and natural image statistics. Variance and kurtosis, which are even-order statistics, have been vigorously studied[23–25]. Skewness, an odd-order statistic, has been largely ignored. Even-order statistics are always the same for an image and its negative, so that they are blind to any asymmetries in light and dark (such as those that occur with highlights and shadows). Skewness is specifically sensitive to these asymmetries. It is easily computed, and we find psychophysical evidence that it is used in human vision.

METHODS

Photographs of real surfaces (24 handmade stuccos, fabric, and crumpled paper) were taken by a 16-bit linear camera (Bitran BS-42N). The standard deviation (s.d.) and skewness of the luminance histogram were defined as:

$$\text{s.d.} = \sqrt{\frac{\sum (I(x,y)-m)^2}{N}}$$
$$\text{skewness} = \frac{\sum (I(x,y)-m)^3}{N(\text{s.d.})^3} \qquad (1)$$

where $I(x,y)$ is the luminance of a pixel, m the mean luminance, and N the number of pixels (256×256). Surface images were presented on a CRT monitor (Sony GDM-F500R, refresh rate 100 Hz, luminance range of 0.1–82 cd m^{-2}) through a graphics card (Cambridge Research System, VSG2/5), with 8-bit luminance resolution for the luminance range of each image. The mean luminance of all images was normalized to 16.3 cd m^{-2}, and the background luminance was kept below 0.1 cd m^{-2}. In the first experiment (Fig. 2a–c), various stucco images were presented in random order, and subjects rated the lightness or the glossiness using a five-level physical scale (0 to 4). Physical samples (Optical Society of America patches or stucco patches) were shown to the subject as reference. In the second experiment (Fig. 2d), the skewness of the luminance histogram of surface images was varied by the procedure of histogram matching to a beta distribution, given by:

$$f(l) = \frac{1}{B(p,q)} l^{p-1}(1-l)^{q-1}$$
$$B(p,q) = \int_0^1 l^{p-1}(1-l)^{q-1}\,\mathrm{d}l \qquad (2)$$

where $q = 10 - p$, l is the luminance and p is the parameter that controls skewness. In the after-effect experiment, subjects viewed an adaptation image pair (Fig. 4a or c) for 100 s at the beginning of a session. In each trial, after 4 s of top-up adaptation, a pair of oppositely skewed stucco images with a given skewness magnitude was presented for 0.5 s, and the subjects indicated which of the two surfaces appeared darker (or glossier). Methods are described more in detail in Supplementary Methods.

Received 21 December 2006; accepted 26 February 2007.
Published online 18 April 2007.

1. Land, E. H. & McCann, J. J. Lightness and retinex theory. *J. Opt. Soc. Am.* **61,** 1–11 (1971).
2. Gilchrist, A. *et al.* An anchoring theory of lightness perception. *Psych. Rev.* **106,** 795–834 (1999).
3. Brainard, D. H. Color constancy in the nearly natural image. 2. Achromatic loci. *J. Opt. Soc. Am. A* **15,** 307–325 (1998).
4. Dana, K. J. *et al.* Reflectance and texture of real-world surfaces. *ACM Trans. Graph.* **18,** 1–34 (1999).
5. Todd, J. T., Norman, J. F. & Mingolla, E. Lightness constancy in the presence of specular highlights. *Psych. Sci.* **15,** 33–39 (2004).
6. Pont, S. C. & Koenderink, J. J. Bidirectional texture contrast function. *Int. J. Comp. Vis.* **62,** 17–34 (2005).
7. Robilotto, R. & Zaidi, Q. Lightness identification of patterned three-dimensional, real objects. *J. Vis.* **6,** 18–36 (2006).
8. Nishida, S. & Shinya, M. Use of image-based information in judgments of surface-reflectance properties. *J. Opt. Soc. Am. A* **15,** 2951–2965 (1998).
9. Fleming, R. W., Dror, R. O. & Adelson, E. H. Real-world illumination and the perception of surface reflectance properties. *J. Vis.* **3,** 347–368 (2003).
10. Dror, R. O., Willsky, A. S. & Adelson, E. H. Statistical characterization of real-world illumination. *J. Vis.* **4,** 821–837 (2004).
11. Fleming, R. W. & Bülthoff, H. H. Low-level image cues in the perception of translucent materials. *ACM Trans. Appl. Percept.* **2,** 346–382 (2005).
12. Levoy, M, *et al.* The Digital Michelangelo Project. ⟨http://graphics.stanford.edu/projects/mich⟩ (2004).
13. Nicodemus, F. Directional reflectance and emissivity of an opaque surface. *Appl. Opt.* **4,** 767–773 (1965).
14. MacGillivray, H. L. Skewness and asymmetry: measures and orderings. *Ann. Stat.* **14,** 994–1011 (1986).
15. Heeger, D. J. Modeling simple-cell direction selectivity with normalized, half-squared, linear operators. *J. Neurophysiol.* **70,** 1885–1898 (1993).
16. Schiller, P. H., Finlay, B. L. & Volman, S. F. Quantitative studies of single-cell properties in monkey striate cortex. I. Spatiotemporal organization of receptive fields. *J. Neurophysiol.* **39,** 1288–1319 (1976).
17. Baizer, J. S., Robinson, D. L. & Dow, B. M. Visual responses of area 18 neurons in awake, behaving monkey. *J. Neurophysiol.* **40,** 1024–1037 (1977).
18. Shipp, S. & Zeki, S. The functional organization of area V2, I: specialization across stripes and layers. *Vis. Neurosci.* **19,** 187–210 (2002).
19. Kagan, I., Gur, M. & Snodderly, D. M. Spatial organization of receptive fields of V1 neurons of alert monkeys: comparison with responses to gratings. *J. Neurophysiol.* **88,** 2257–2274 (2002).
20. Mata, M. L. & Ringach, D. L. Spatial overlap of ON and OFF subregions and its relation to response modulation ratio in macaque primary visual cortex. *J. Neurophysiol.* **93,** 919–928 (2005).
21. Malik, J. & Perona, P. Preattentive texture discrimination with early vision mechanisms. *J. Opt. Soc. Am. A* **5,** 923–932 (1990).
22. Chubb, C., Landy, M. S. & Econopouly, J. A visual mechanism tuned to black. *Vision Res.* **44,** 3223–3232 (2004).
23. Olshausen, B. A. & Field, D. J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **381,** 607–609 (1996).
24. Simoncelli, E. P. & Olshausen, B. A. Natural image statistics and neural representation. *Annu. Rev. Neurosci.* **24,** 1193–1216 (2001).
25. Kingdom, F. A. A., Hayes, A. & Field, D. J. Sensitivity to contrast histogram differences in synthetic wavelet-textures. *Vision Res.* **41,** 585–598 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank Y. Li for discussions. L.S. and E.H.A. were supported by NTT and by a grant from the National Science Foundation to E.H.A.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to I.M. (motoyosi@apollo3.brl.ntt.co.jp).