nature neuroscience

Perceptual training continuously refines neuronal population codes in primary visual cortex

Yin Yan[1,2], Malte J Rasch[1,2], Minggui Chen[1,2], Xiaoping Xiang[1], Min Huang[1], Si Wu[1] & Wu Li[1]

Perceptual learning substantially improves visual discrimination and detection ability, which has been associated with visual cortical plasticity. However, little is known about the dynamic changes in neuronal response properties over the course of training. Using chronically implanted multielectrode arrays, we were able to capture day-by-day spatiotemporal dynamics of neurons in the primary visual cortex (V1) of monkeys trained to detect camouflaged visual contours. We found progressive strengthening and accelerating in both facilitation of neurons encoding the contour elements and suppression of neurons responding to the background components. The enhancement of this figure-ground contrast in V1 was closely correlated with improved behavioral performance on a daily basis. Decoding accuracy of a simple linear classifier based on V1 population responses also paralleled the animal's behavioral changes. Our results indicate that perceptual learning shapes the V1 population code to allow a more efficient readout of task-relevant information.

Perceptual learning, improvement with training in perceptual abilities, has been used as a model to study experience-dependent cortical plasticity in adults. The focal contentions in perceptual learning studies in the last couple of decades have been the cortical loci in which the plastic changes occur and the forms of cortical changes that can account for the improved behavioral performance[1–3]. Some electrophysiological studies have shown that training in simple discrimination and detection tasks can alter neuronal response properties in visual cortical areas such as V1 (refs. 4–7) and V4 (refs. 8–11). However, it has been argued that perceptual training mainly modifies downstream cortical areas that are engaged in the allocation of attentional resource[12] and integration of sensory evidence for decision-making[13–15]. Thus, it is still a matter of debate whether the sensory codes for the trained stimulus can be refined in early visual cortex through repeated practice.

More controversial observations of learning-induced changes come from imaging studies, which have shown that training can result in increased[16–19] or decreased[20,21] activation in early cortical areas, that an initial increase is followed by a drop of activation back to pre-training level[22], and even that the overall activation is not affected, whereas the discrimination abilities of individual voxels are enhanced[23]. These mixed results suggest that perceptual training can cause complex changes in early visual cortex and that different response patterns may reflect different phases of learning.

Previous single-electrode recordings have shown learning-induced changes in V1 by comparing different neuronal samples before and after training. It is unclear how perceptual learning dynamically affects neural coding at the population level over the entire course of training and to what extent neuronal dynamics can account for the diverse imaging results. We sought to address this issue, which is important for understanding the neural mechanisms of perceptual learning. To this end, we tracked possible dynamic changes in spatiotemporal properties of V1 neurons in monkeys implanted with multielectrode arrays. By characterizing learning-induced daily changes across more than 2 weeks of extensive training on a visual contour detection task, we found that V1 population responses were markedly remodeled in space and time. Our observations suggest that perceptual learning refines the representation of visual information in V1 for enabling and facilitating a simple readout of task-relevant signals.

RESULTS

Two naive adult male monkeys (*Macaca mulatta*, named MG and MJ) participated in this study. Multiunit activities of V1 superficial layer neurons were recorded with chronically implanted microelectrode arrays (6 × 8 electrodes, 0.5 mm in length, 0.4-mm spacing). The receptive field (RF) centers and preferred orientations of V1 recording sites were very stable across training days (**Fig. 1a**), suggesting that we had been recording from the same set of cortical columns.

The monkeys were trained to perform a contour detection task using a two-alternative forced-choice task[6,24]. A pair of stimulus patterns was displayed simultaneously for a brief amount of time (**Fig. 1b**). These two patterns were circular in outline, 8.5° (in monkey MG) or 7.5° (MJ) in diameter, divided by hidden rhomboid grids. Each rhomboid compartment, with equal and constant base and height of 0.5° and therefore equal and constant area, contained a small bar of 0.25 × 0.05°. In one pattern (referred to as the noise pattern, NP), the orientation of each component bar was randomly assigned. In the other pattern (referred to as the contour pattern, CP; **Fig. 1c**), a contour was formed by collinear alignment of adjacent bars along a diagonal axis of the hidden compartments, whereas the remaining bars were exactly identical to those in the noise pattern (**Fig. 1b**, see also refs. 24,25). In a block of trials, either the CP or NP

[1]State Key Laboratory of Cognitive Neuroscience and Learning, IDG/McGovern Institute for Brain Research, and Center for Collaboration and Innovation in Brain and Learning Sciences, Beijing Normal University, Beijing, China. [2]These authors contributed equally to this work. Correspondence should be addressed to W.L. (liwu@bnu.edu.cn).

Received 16 April; accepted 13 August; published online 7 September 2014; doi:10.1038/nn.3805

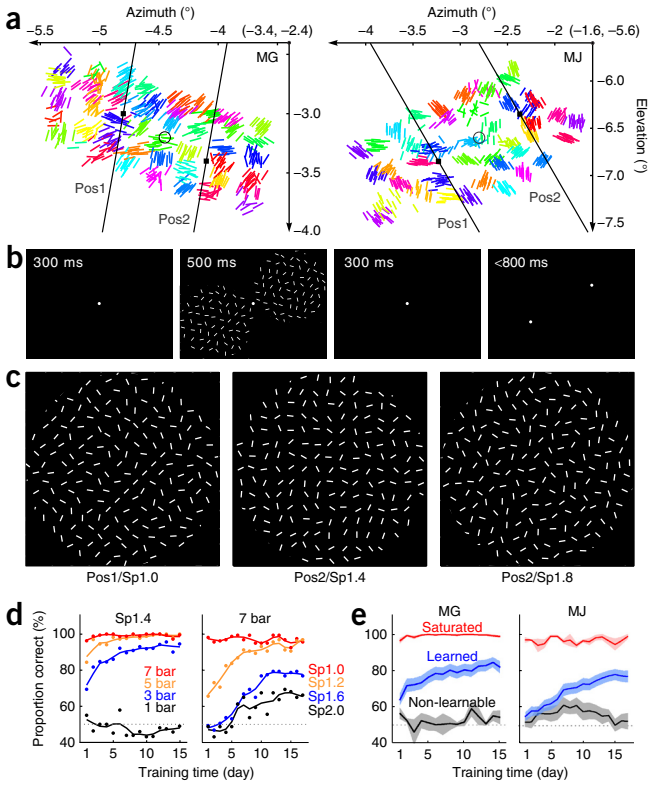

Figure 1 Experimental design. (**a**) Visual-field locations of the recorded RFs in MG and MJ, respectively. Each small cluster of lines of the same color indicates a recording site, with the center and orientation of each line representing the RF center and preferred orientation measured on each training day by the same electrode. Different recording sites are marked by different colors. The two long black parallel lines show the two possible positions of the embedded contour and its orientation, with the black dot on each line indicating the center of the global collinear contour. The empty circle in the panel center shows the center of the entire stimulus pattern (MG, −4.5°, −3.2°; MJ, −2.8°, −6.6°). Note that the total number of electrodes that were able to pick up recordable spiking activity slightly fluctuated across days (Online Methods). (**b**) The timing of a trial in the contour detection task. When the animal was maintaining fixation, the contour and noise patterns were displayed simultaneously. The monkey indicated the contour pattern position by making a saccade to either of the two targets displayed at the end of the trial. (**c**) Sample contour patterns showing two possible contour positions (Pos1 and Pos2) and three spacings (Sp1.0, Sp1.4 and Sp1.8) between collinear bars (seven bars here). (**d**) Dependence of contour detection performance and learning rate on the number of collinear bars (left, sample conditions from MG) and on the spacing between collinear bars (right, sample conditions from MJ). (**e**) The two animals' learning curves averaged across different conditions (red, $n = 6$ in MG, $n = 3$ in MJ; blue, $n = 27$ in MG, $n = 23$ in MJ; black, $n = 3$ in MG, $n = 10$ in MJ). The shaded areas represent s.e.m.

stimulus—with equal probability—covered all of the simultaneously recorded RFs; the other stimulus was placed symmetrically around the fixation point in the opposite visual field. After a stimulus presentation, the monkeys were required to indicate the CP by making a saccade to either of two dots corresponding to the two stimulus locations (**Fig. 1b**).

To minimize sampling bias in V1, the contour was embedded at either of two fixed positions in the randomly oriented bars (**Fig. 1a**,**c**). These two possible positions and the contour orientation were arbitrarily set in such a way that a considerable number of RFs lay on the contour as well as on the background with varying distances from the contour (**Fig. 1a**). To cover a range of task difficulties, we changed the number of collinear bars forming the contour and varied the spacing between them while keeping the density of bars in the stimulus pattern constant (**Fig. 1c** and Online Methods).

Six center-to-center spacings between collinear bars (1.0, 1.2, 1.4, 1.6, 1.8 and 2.0× the height of a rhomboid compartment), four different contour lengths (1, 3, 5 and 7 collinear bars), and two possible locations for embedding the contour were combined into 48 conditions with different degrees of contour saliency (**Fig. 1c**). Once the CP and the corresponding NP were defined for each condition, all stimuli were kept constant across all training sessions. Note that, in the one-bar contour conditions, the CP was identical to the NP (12 of the 48 conditions). In these control conditions, the monkey could only guess, with 50% chance level, which side was the CP. All 48 stimulus conditions were randomly mixed in a block of trials, in which each stimulus condition was repeated ten times. Given that the CPs (and NPs) were presented in equal amounts of trials on either the RF side or the opposite hemifield, the total number of trials in a training session was $48 \times 10 \times 2 = 960$. The animals underwent four training sessions a day.

Before we collected the perceptual training data, the animals underwent an operational training stage using highlighted contours with a luminance contrast against the complex background. After the animals fully understood the task (Online Methods), perceptual training and data collection began. The training lasted about 2 weeks, until the animals' contour detection performance reached plateaus in most of the stimulus conditions (15 d for MG, 17 d for MJ).

Changes in behavioral performance with training

Expectedly, the animals' performance on contour detection and the rate of learning depended on the geometry of the contour stimuli (**Fig. 1d**). For data analyses, we separated the 48 stimulus conditions into three groups by comparing the animal's performance across days in each condition.

For contours of very high or very low saliencies, no significant change in detection performance was observed during the training (detection rates tested for significant increase using linear regression over days, F statistics of mean squared error fit, $P > 0.05$). These conditions were further categorized into three groups: the control conditions (that is, the one-bar contour conditions), the saturated conditions (in which the animals were able to reliably detect the contours in more than 90% of trials throughout the training; MG, $n = 6$; MJ, $n = 3$; **Fig. 1e**) and the non-learnable conditions (in which the animals' performance was still near the chance level at the end of training; MG, $n = 3$; MJ, $n = 10$; **Fig. 1e**). Note that there were more non-learnable conditions and fewer saturated conditions in MJ than in MG. This could be a result of a larger eccentricity of stimulus in MJ, rendering contour detection more difficult[26]. The remaining conditions with intermediate contour saliencies were classified as the learned conditions, in which the animals showed a monotonic increase in contour detection performance (linear regression, F statistics, $P < 0.05$; MG, $n = 27$; MJ, $n = 23$; **Fig. 1e**).

Spatial layout of learning-induced changes in V1

As a pre-training control, we recorded V1 neuronal responses to the entire set of stimuli while the animal simply maintained fixation. After the operational training when the animals fully understood the detection task, we continuously recorded neuronal activity while the animals were being trained in the contour detection task for about 2 weeks.

At the end of perceptual training, we observed that the majority of recording sites responded differently to CP versus NP stimuli: most

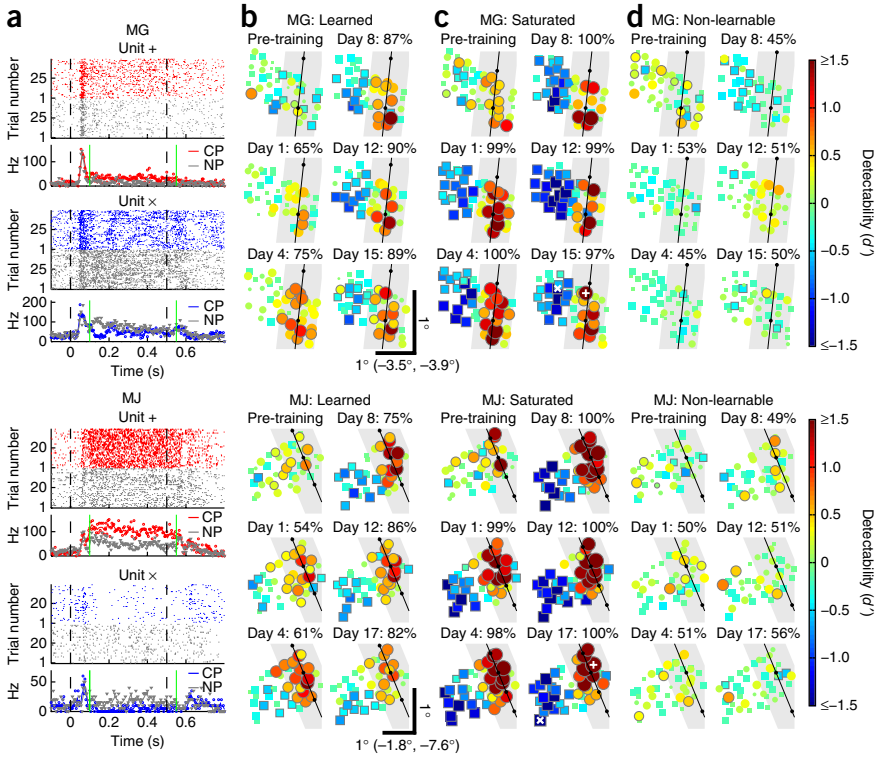

Figure 2 Contour-related signals in V1 and their changes with training. Shown here are sample data from MG (upper row) and MJ (lower row). (**a**) Spike raster plots and PSTHs (5-ms bins smoothed by sliding a boxcar filter of three-bin width) from the two sample recording sites marked with + and × in **c**. Responses to the contour pattern (CP, red or blue) and noise pattern (NP, black) were compared for each site. Dashed lines indicate the stimulus exposure duration (0–500 ms); green lines on the PSTH mark the time window (100–550 ms) used for calculating the contour-related signals shown in **b**–**d**. (**b**) Contour-related signals defined as the *d′* value between CP and NP responses for each recording site in a typical learned condition. The six panels for each animal correspond to six selected time points (days of training and corresponding behavioral performance are indicated on top of each panel). The squares and circles in each panel represent the recorded V1 sites showing contour-induced inhibitory and facilitatory modulation, respectively, with symbol size and color indicating the modulation strength. Symbols with black outlines indicate those sites showing a significant difference in response to the CP and NP stimuli (Wilcoxon rank-sum test, $\alpha = 0.05$). Black lines show the position and orientation of the embedded contour; equally spaced dots along the line indicate the centers of collinear bars forming the middle segment of the contours. (**c**) Same recording sites as in **b** in a saturated condition. (**d**) A non-learnable condition.

V1 sites with RFs lying on the contour responded more vigorously to the CP than NP, whereas most sites on the background were suppressed by the CP relative to the NP (**Fig. 2a**).

To quantify these contour-related signals, neuronal responses to the CP stimulus were contrasted with those to the NP using the detectability measure *d′*. Defined by the difference of the mean responses to CP and NP stimuli normalized by the magnitude of the noise fluctuations (Online Methods), a *d′* value of 0 indicates no contour-induced activity, and a positive or negative value indicates facilitation or suppression of neuronal responses induced by the embedded contour, respectively. This measure for the strength and sign of the contour modulation in individual V1 site was monitored across days to examine learning-induced changes.

We observed that perceptual learning systematically affected contour modulation strength across the recorded neuronal population. Taking a typical learned stimulus condition as an example (**Fig. 2b**), we found that few V1 recording sites showed discernible contour-related responses before training, regardless of whether their RFs were located on the contour or background. When the animals started to perform the task and their performance progressively improved with training, the contours elicited increasingly stronger modulations of neuronal responses over the course of training: neurons with RFs lying over the contour were gradually facilitated (referred to as contour facilitation), whereas neurons with RFs on the background were gradually suppressed (referred to as background suppression). The contour facilitation and background suppression were generally observed for V1 sites with different orientation preferences, consistent with recent findings[27].

Taking a typical saturated stimulus condition as another example (**Fig. 2c**), a similar modulatory pattern was weakly expressed before training, but it immediately became very pronounced on the first day, when perceptual learning just began. Subsequent extensive training had little effect on the strength of contour facilitation, but the background suppression was further enhanced over the population (see below). Compared with the learned and saturated conditions, for the non-learnable conditions, no clear or only very weak contour-induced modulatory effects evolved in V1 with training (**Fig. 2d**).

Quantification of learning-induced changes in V1

Binning the recording sites according to the distances between their RF centers and the contour (**Fig. 3a**), the separation into two opposite groups was evident: V1 sites on the contour showed facilitatory effects (defined as the near sites) and those on the complex background showed inhibitory effects (defined as the far sites). A boundary between the near and far groups of recording sites was quantitatively defined by fitting a difference of Gaussians curve to the modulatory profile, resulting in distances of 0.48° and 0.58° from the contour for MG and MJ, respectively (**Fig. 3a**). This boundary coincides with about half of the mean RF size, confirming that whether a V1 site is facilitated or inhibited by the global contour depends on whether its RF lies on the embedded contour or on the background.

Combining recording sites into the near and far groups allowed us to quantify the contour facilitation and background suppression over the course of training. For the learned conditions (**Fig. 3b**), the average contour modulation strengths on the first training day were significantly stronger than the pre-training values for both near and far sites in both animals (averaged *d′* value, pre-training versus first day, one-tailed unpaired *t* test: MG near, 0.03 ($n = 594$) versus 0.11 ($n = 570$), $P = 3 \times 10^{-6}$; MJ near, 0.01 ($n = 459$) versus 0.29 ($n = 436$), $P < 10^{-6}$; MG far, −0.03 ($n = 621$) versus −0.08 ($n = 564$), $P = 0.001$; MJ far, 0.01 ($n = 461$) versus −0.08 ($n = 461$), $P < 10^{-6}$). The same set of recording sites was included multiple times as a result of averaging

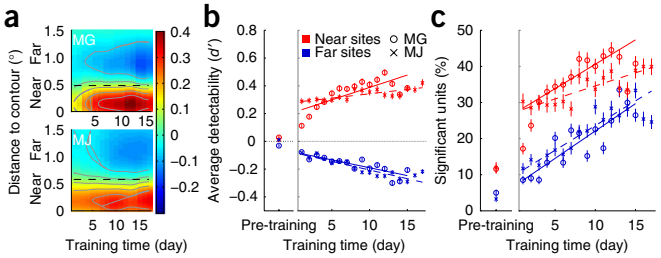

Figure 3 Quantification of learning-induced changes in V1 for the learned conditions. (**a**) Daily changes in contour modulation (d' value) for recording sites binned by RF-to-contour distances (vertical axis, 0.005° bin, smoothed with a Gaussian of 0.2° s.d.). A clear separation and strengthening of contour facilitation and background suppression was noticeable in both monkeys. Dashed line indicates the border between near and far sites (Online Methods). (**b**) A progressive increase with training in the mean strength of contour modulation for near and far sites, respectively (averaged for the learned conditions; error bars indicate s.e.m.). The isolated data points on the left show the pre-training modulation strengths. The slope of a linear regression of data points across days was used as a measure of learning-induced changes in facilitation of near sites and suppression of far sites. (**c**) A progressive increase with training in percentage of V1 sites showing significant contour facilitation and background suppression (Wilcoxon rank sum test, $\alpha = 0.05$). Percentages were calculated by pooling all the learned conditions.

across multiple learned conditions. The contour-induced modulations gradually increased over the course of training (linear regression using a robust fitting method, *robustfit* in Matlab, two-tailed one-sample t test whether the slope is different from zero: MG near, slope = 1.8×10^{-2}, $P = 2 \times 10^{-3}$, $t = 3.9$; MG far, slope = -1.1×10^{-2}, $P = 2 \times 10^{-4}$, $t = -5.2$; MJ near, slope = 0.6×10^{-2}, $P = 3 \times 10^{-4}$, $t = 4.6$; MJ far, slope = -1.3×10^{-2}, $P < 10^{-6}$, $t = -8.7$; degrees of freedom in linear regression: 13 for MG, 15 for MJ; **Fig. 3b**).

For the saturated conditions, the background suppression of the far sites also tended to increase with training regardless of the unchanged behavioral performance (t test as above: MG far, slope = -0.8×10^{-2}, $P = 0.12$, $t = -1.7$; MJ far, slope = -1.6×10^{-2}, $P = 3 \times 10^{-4}$, $t = -4.6$); however, the contour facilitation of the near sites was not significantly affected (MG near, $P = 0.46$, $t = -0.7$; MJ near, $P = 0.33$, $t = 1.0$). These results suggest that perceptual learning is able to enhance V1 coding efficiency even when the behavioral performance has reached the ceiling and that this is mainly achieved by suppression of the background.

For the non-learnable conditions, the contour-induced modulations slowly evolved with training, but they were much weaker than the learned conditions (t test as above: MG near, slope = 1.2×10^{-2}, $P = 0.01$, $t = 2.8$; MG far, slope = -0.4×10^{-2}, $P = 0.17$, $t = -1.5$; MJ near, slope = 0.4×10^{-2}, $P = 0.04$, $t = 2.2$; MJ far, slope = -0.3×10^{-2}, $P = 0.06$, $t = -2.0$).

These results were derived from pooling stimulus conditions. To assess the changes of contour modulation strength for individual conditions, we defined the figure-ground contrast by subtracting the mean d' value (with sign) of far sites from that of the near sites for each stimulus condition. The figure-ground contrast increased significantly (measuring the slope across days as above, $P < 0.05$) in 24 of 27 (MG) and 14 of 23 (MJ) of the learned conditions.

The arising of a clear profile of figure-ground contrast over the course of training in individual learned conditions (**Fig. 2b**) suggests that the increase of the site-averaged modulation strength (**Fig. 3a**,**b**) is a result of a gradual response change of the whole population of recorded cells rather than a marked change in a few selected and fixed units. To quantify the contribution of individual sites, we counted, for each day, the percentage of sites that showed a significant difference in response to CP and NP stimuli (Wilcoxon rank sum test, $P < 0.05$). Before training, only a small proportion of near sites was weakly modulated by the contours in the learned conditions, and an even smaller proportion of far sites were modulated (**Fig. 3c**). The percentage of significant sites, both near and far, progressively increased during learning (slope significantly larger than 0, t test as above: MG near, slope = 1.4×10^{-2}, $P = 4 \times 10^{-5}$, $t = 6.1$; MG far, slope = 1.5×10^{-2}, $P = 6 \times 10^{-5}$, $t = 5.8$; MJ near, slope = 0.7×10^{-2}, $P = 7 \times 10^{-5}$, $t = 5.4$; MJ far, slope = 1.4×10^{-2}, $P < 10^{-6}$, $t = 9.8$; **Fig. 3c**).

Changes in the temporal dynamics of V1 responses

We next examined the extent to which the temporal dynamics of the figure-ground contrast was systematically modified by perceptual learning. We grouped V1 data into three training intervals (days 1–5, 6–10 and 11–15) and constructed the population peri-stimulus time histograms (PSTHs) for the near and far sites in response to the CP and NP stimuli, respectively (**Fig. 4a**). Consistent with previous analyses, training led to an increase of near sites' responses and a decrease of far sites' responses (**Fig. 4b**). These contour-induced facilitatory and inhibitory modulations, which are measured relative to the baseline control of NP stimulation, seemed to covary after stimulus onset. However, the raw population responses of the near and far sites to the CP stimulus were actually very similar in their temporal structures (**Fig. 4a**): the initial peaks of their population PSTHs were nearly superimposed and their delayed response components were approximately parallel with a constant difference (**Fig. 4b**). Conversion of the differences of mean firing rates into the discriminability measure (d') generated very similar temporal profiles (**Fig. 4c**). These data indicate that, despite the complex temporal dynamics of contour facilitation and background suppression relative to the NP baseline, the temporal evolution of the figure-ground contrast during CP stimulation was simple and almost constant, and was systematically elevated by training (**Fig. 4b**,**c**). The strengthening of the figure-ground contrast was highly correlated with improved behavioral performance (Pearson correlation across days and learned conditions: MG, $r = 0.91$, $P < 10^{-6}$; MJ, $r = 0.76$, $P < 10^{-6}$).

In addition to enhancing the figure-ground contrast, perceptual learning also substantially shortened the latency of the contour integration process in V1. By estimating the time point relative to stimulus onset when the contour modulation first became significant (**Fig. 4d**), we found that the onset latency for both the contour facilitation of near sites and the background suppression of far sites decreased progressively with training. The latency decrease over days could be well fitted with an exponential decline (**Fig. 4d**). When comparing the first and the last available day on the basis of the exponential fit (that is, the two end points of the exponential curves), the latency of contour facilitation decreased by 35 ms in MG (from 128 to 93 ms) and 16 ms in MJ (from 104 to 88 ms). The latency of background suppression, which was considerably longer than contour facilitation, also became shorter with training (decreased by 51 ms in MG, from 178 to 127 ms; 69 ms in MJ, from 189 to 120 ms).

Training improves both encoding and readout

From the perspective of information processing, the visual signals encoded by V1 should be accessible to higher order areas, and a hypothetical decision unit should be able to decode task-relevant stimulus

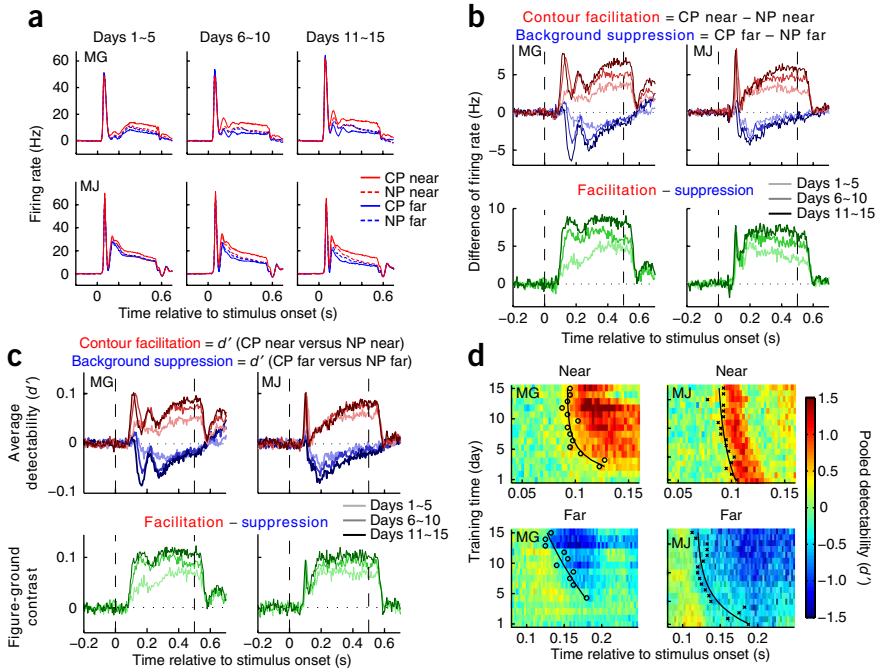

Figure 4 Learning-induced changes in temporal dynamics of V1 responses. (**a**) Population PSTHs (5-ms bin with spontaneous activity subtracted) averaged over the learned conditions for the near (red) and far (blue) sites in responses to the CP (solid curves) and NP (dashed curves). The data are grouped into three training periods (corresponding to three columns) for each animal. (**b**) Contour-induced modulations directly measured as firing rate changes by comparing the PSTHs shown in **a**. Top, differential PSTHs between the CP and NP stimulus conditions showing the contour facilitation of near sites (upward curves) and background suppression of far sites (downward curves). Three curves in each group represent different training periods. Vertical dashed lines delimit the stimulus exposure duration. Bottom, the overall modulation strength measured as the distance between corresponding upward and downward curves shown in the top panels. (**c**) Data presented as in **b**, but the contour-induced modulations were measured as the *d′* values rather than firing rate changes. (**d**) Shortening of latencies of contour-related V1 responses. The time course (2.5-ms bin) of *d′* values (color coded) was calculated by pooling all available spike trains across the learned conditions on each day for the near (upper) and far (lower) sites, respectively (Online Methods). Symbols (o and ×) indicate the time at which the contour modulation first became significant (determined by bootstrapping, Online Methods). This time point was defined as the latency of contour-related responses. The data points are fitted by an exponential function.

information based on the output signals from V1 neurons. In principle, there are two possible mechanisms for how perceptual learning could improve this readout process. One way is selective re-weighting of the sensory outputs from V1 (ref. 13): information about the stimuli could remain unchanged in V1 throughout behavioral training, whereas downstream cortical areas—initially limited in their readout capabilities—gradually adjust their weighting to more efficiently read out the existing task-relevant information entangled in V1 responses (**Fig. 5a**). Another way is input selection in V1 (ref. 28): training could selectively increase information conveyed by V1 neurons about task-relevant stimulus features (**Fig. 5b**). These two mechanisms for selection of task-related signals are not mutually exclusive.

To examine these mechanisms, we trained artificial classifiers to distinguish the CP from the NP responses and thereby simulated the trial-by-trial choice of a hypothetical decision unit. If perceptual learning would be solely a process of readout re-weighting in higher areas (**Fig. 5a**) and available information in V1 would not change, an artificial classifier trained on the V1 data would show constant performance across days. On the other hand, if perceptual learning could improve sensory encoding so that more and more information about task-related stimulus features is made available in V1 (**Fig. 5b**), the decoding performance of an artificial classifier would increase accordingly.

We used the mean responses (100–550 ms after stimulus onset) of individual recording sites as inputs to the classifiers (Online Methods). We first used a powerful nonlinear classifier (support vector machine, SVM, with a radial-basis function kernel), which is capable of incorporating information from individual recording sites as well as from their potential nonlinear interactions. The overall classification performance of the SVM classifier averaged across the learned conditions and all training days well matched the behavioral data in monkey MG (SVM, 74.6% ± 0.8%, mean ± s.e.m.; behavior, 77.5% ± 0.9%; **Fig. 5c**); the classification performance was on average even better than the behavioral performance in MJ (SVM, 73.5% ± 0.8%; behavior, 68.9% ± 0.8%; **Fig. 5d**).

By comparing the classification performance of SVM with the animal's performance day by day (**Fig. 5c**,**d**), evidence for both hypotheses of the neural mechanisms underlying perceptual learning could be found. On the one hand, the classifier's performance increased significantly with training in both monkeys, as was seen in data averaged over the learned conditions (**Fig. 5c**,**d**), and in the great majority of individual conditions (Wilcoxon rank-sum test of whether the median performance of the first 5 d was smaller than that of the last 5 d, $P < 0.05$; 26 of 27 significant conditions in MG, 20 of 23 in MJ). This indicates that the contour information contained in V1 responses indeed greatly increased during perceptual learning. On the other hand, at least in MJ, the classifier's performance was much better than the animal's behavior during the first few days of training (**Fig. 5d**), indicating that the readout system was not adjusted well enough initially to access all existing contour information in V1. However, MJ's performance gradually caught up with the classifier, suggesting that, in addition to enhanced V1 population code, the readout process is improved by training as well.

Training enhances signal, but does not reduce internal noise

Each neuron's responses to repeated stimulus presentations are noisy in nature, fluctuating around a mean value (**Fig. 6a**). In principle, there are two simple strategies to enhance the signal-to-noise ratio for improving the classification performance: increasing the distance between the mean values of the responses to CP and NP stimuli (that is, increasing the signal strength), and reducing the trial-to-trial fluctuations of the responses along the direction orthogonal to the decision line that separates the CP and NP responses (that is, decreasing the harmful components of noise fluctuations). We asked what strategy was chosen in V1 during perceptual learning.

If one assumes that the noise fluctuations of neuronal responses to both CP and NP stimuli are normally distributed, a simple linear classification method, such as Fisher linear discriminant (FLD), can

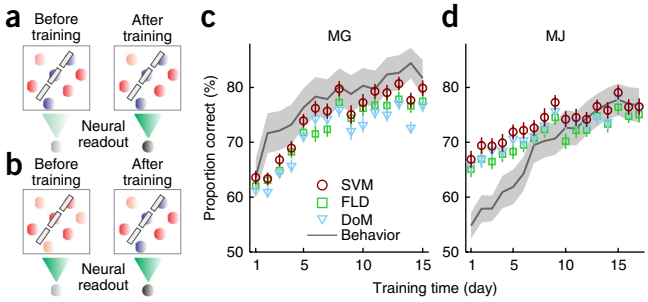

Figure 5 Decoding V1 population responses. (**a**,**b**) Two hypothetical mechanisms of perceptual learning. The readout reweighting scheme is shown in **a**: information encoded in V1 is unaffected by training, but the downstream readout process is adjusted to access the task-relevant information. The input selection scheme is shown in **b**: task-relevant information in V1 is increased with training. (**c**,**d**) Comparison of the decoding performance with the animals' behavior in the learned stimulus conditions. Classification accuracy of the SVM is presented here together with the results from two other types of simple linear classifiers, FLD and DoM (**Fig. 6**). Note the close similarities between the much simpler (DoM) and more complex (SVM and FLD) decoders over the entire course of training. The classifiers were retrained for each day. Error bars indicate s.e.m. (across classifier's training-testing repeats over the learned conditions, see Online Methods).

be used to optimize the signal-to-noise ratio of the neural population for the detection task (**Fig. 6a**,**b**). FLD projects multi-dimensional data (CP and NP responses of all V1 sites) onto a one-dimensional space (termed the FLD direction) that maximizes the ratio of the signal strength to the noise hampering the signal (the s.d. of the noise fluctuations along the FLD direction). In other words, FLD finds a set of optimum weights (or synaptic strengths) for the recorded V1 neurons connecting to a hypothetical linear readout neuron. As a result, this readout neuron would have the highest detectability (d' value), among all possible weight configurations, in decoding CP and NP stimuli on the basis of the outputs of V1 neurons.

We performed the FLD analysis for each of the learned conditions in each day and plotted the mean signal strength (averaged across the learned conditions) and the corresponding mean noise fluctuation over days (**Fig. 6c**). Notably, we found that perceptual learning in the contour detection task predominantly enhanced the signal strength rather than decreased the noise fluctuations. In both monkeys, the increase of the signal strength was highly significant (slope of linear regression is larger than 0, t test, $P = 10^{-6}$ for MG, $P = 7 \times 10^{-5}$ for MJ), but no change in the noise fluctuations was detected (MG, $P = 0.28$; MJ, $P = 0.59$). Note that the noise fluctuation magnitudes measured here are only those components—among the noise fluctuations in all possible directions—that hamper the decision, as they interfere with reading out the task-relevant information from the population activity. The structure of noise fluctuations as a whole might nevertheless be affected by perceptual learning in other dimensions that are irrelevant to the readout process. Indeed, we found that the noise correlation between any two recording sites markedly decreased over the course of training (Online Methods), consistent with a recent study[29].

Simple readouts suffice to extract information in V1

It is well known that a single neuron has the computational capabilities to build a simple linear classifier: the synaptic strengths, together with a spiking threshold, form a weighted linear summation of inputs and a subsequent binary classification[30,31]. Could a downstream readout neuron with such a simple summation and classification property extract the task-relevant information from V1 to the same degree as the powerful nonlinear SVM?

The FLD method introduced above can be used to implement such a linear classifier when one adds a simple mechanism to classify the projected one-dimensional data (**Fig. 6a**,**b** and Online Methods). The FLD performance averaged across the last 3 d and the learned conditions was very close to the SVM method (only 2.5% and 2.3% lower in MG and MJ, respectively, although the small difference was statistically significant, paired t test, $P < 10^{-6}$ in MG, $P = 1 \times 10^{-5}$ in MJ; **Figs. 5c**,**d** and **6d**). This indicates that task-relevant information is indeed accessible to simple linear readouts and that high-dimensional nonlinear interactions between neurons potentially have only a minor role in the V1 population code.

Given that enhancing signal strength was more important than reducing the noise fluctuations, we tested whether even simpler readouts could perform the task as well. Instead of finding the FLD direction, and thus an optimal decision line, readout units might alternatively ignore the noise fluctuation structure altogether and only focus on maximizing the signal strengths. To test this, we projected

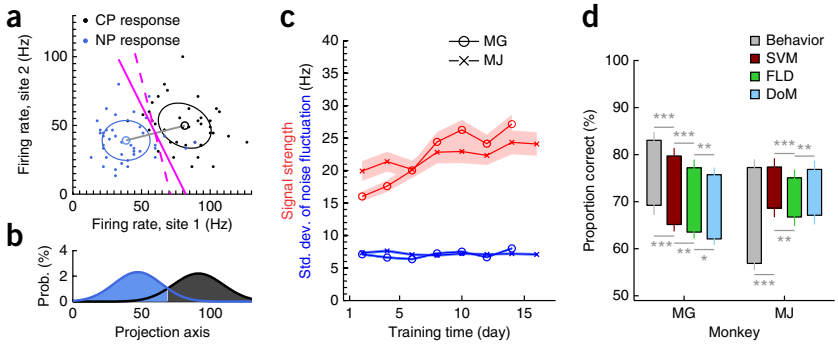

Figure 6 Further decoding analyses. (**a**) Analysis of the contributions of signal enhancement and internal noise reduction to perceptual learning. A simplified two-dimensional example (that is, two recording sites) showing how firing rates to CP and NP stimuli fluctuated around the mean values (two small circles) across trials (individual dots). Noise fluctuations were fitted by a two-dimensional normal distribution (ellipses). Although the FLD method adjusts the decision line (magenta, solid) to the noise variance, the DoM decision line (magenta, dashed) ignores it, simply orthogonal to the line connecting the two means (solid gray line). (**b**) Data in **a** were projected onto an axis orthogonal to the decision line and then fitted with two Gaussians. After this one-dimensional projection of the CP and NP responses, trials could be classified on the basis of their likelihood ratios. Note that one could either increase the distance between the two Gaussians (the signal strength) or decrease the variance of the Gaussians (the noise fluctuations) to get a better detection performance. (**c**) Changes with training in signal strength (red) and noise fluctuation (blue) estimated with the FLD method described in **a** and **b** (averaged for the learned conditions and every 2 d, shading indicates s.e.m.). (**d**) Quantitative comparison of detection performance among three different classifiers and the animals' behavior. The lower and upper ends of each bar indicate the mean performance averaged for the first 3 and last 3 d of training, respectively (for daily performance, see **Fig. 5c**,**d**). Error bars indicate s.e.m. (one-tailed paired t test, $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$).

the data onto the axis simply connecting the CP and NP mean responses and calculated the difference of means (DoM; **Fig. 6a**), thereby ignoring any potential information provided by the shape of the noise fluctuations. This method resulted in a decision line that was normally different from the other more powerful classifiers (**Fig. 6a**). The performance of this simpler DoM classifier was very similar to the FLD method: only 1.5% worse in MG (one-tailed paired *t* test, $P = 0.006$) and even 1.8% better in MJ ($P = 0.001$) when averaged across the learned conditions in the last 3 training days (**Figs. 5c**,**d** and **6d**). The comparable performance of DoM to the FLD and SVM suggests that this simplest readout method suffices to extract contour-related information contained in V1 population responses.

DISCUSSION

Learning-induced visual cortical changes remain poorly understood and controversial. Here, by chronic recording from multiple sites in V1, we revealed how spatiotemporal properties of neuronal populations were progressively modified by learning a visual contour detection task. Our results indicate that perceptual learning of contour detection selectively modifies the activity of neuronal populations in V1. The consequence is an increment of task-relevant information in V1 as well as an acceleration of the integration process, which could enable and facilitate simple readouts at later processing stages. Behavioral improvement is largely driven by an increase in signal strength rather than by a decrease of trial-by-trial noise fluctuations.

Possible role of feedback and top-down modulations

Even after 2 weeks of extensive training, the onset of contour modulation was still significantly delayed relative to visual response latencies of V1 neurons, suggesting that contour integration involves feedback or recurrent mechanisms as opposed to a simple feedforward process and that modification of synaptic weights of reentrant connections are responsible for learning-induced improvement. Delayed response components have been generally observed in V1 of awake monkeys during processes that involve grouping and segregating of complex image components, such as contour integration[24,32,33] and surface segmentation[34–36]. These delayed response components are closely correlated with the animal's performance on target detection in the presence of a complex background[6,24,37–39]. It has been suggested that these figure-ground signals derive from complex interactions between feedback projections to V1 and intrinsic connections in V1 (refs. 33,36,39,40). In particular, using simultaneous recording from monkey V1 and V4, we recently found interdependent feedforward and feedback processes that operate synergistically to enhance the contour signals, resulting in a parallel increment of contour information in both cortical areas[27].

Given the presence of feedforward and feedback interactions, we cannot exclude the possibility that the learning-induced changes in V1 are also associated with changes in higher order cortical areas. It has been shown that perceptual learning of orientation discrimination can selectively enhance neuronal selectivity for the trained stimulus orientation in both V1 (refs. 4,23) and V4 (refs. 10,11). Task-dependent top-down influences mediated by feedback connections are known to dynamically affect contextual modulation in V1 by selectively enhancing task-relevant stimulus features and suppressing irrelevant ones[5,28]; thus, a refinement in top-down control may also contribute to the learning-induced changes in V1. In fact, top-down modulation and perceptual learning show some similar effects on visual cortical processing (see reviews, refs. 3,41,42), although they operate on different timescales. Moreover, selective attention to task-related stimulus feature is usually required for learning-induced behavioral improvement[43,44] and corresponding cortical changes[4,6]. Thus, a plausible account of the perceptual learning mechanisms is that short-term cortical dynamics under top-down influences, when repeatedly exercised, will result in long-term enhancement in efficiency of sensory encoding, signal readout and attentional control, leading to improvement in perceptual ability as well as automatization of the perceptual task.

Despite a possible chain of interconnected changes across multiple processes during perceptual training, our findings indicate that V1 substantially contributes to the transformation of these changes into a more efficient and informative sensory representation of the learned stimulus. These changes at the earliest stage of visual cortical processing would in turn contribute to a more efficient and less effortful readout at subsequent processing stages and to better behavioral performance.

We found that contour-induced modulation of V1 responses was already significantly stronger on the first training day than the pre-training strength, especially for the saturated conditions. This rapid change may reflect a fast learning phase[45,46], which is very unlikely to be mediated by massive plastic changes of synaptic connections; rather, it could result from task-dependent top-down influences, which have been shown to be able to dynamically alter response properties of V1 neurons to meet the requirements of the perceptual task[28,33]. In addition to this fast learning phase, extensive training continuously strengthened the figure-ground contrast over the course of training, suggesting a second and relatively slow phase of perceptual learning.

The neural code for perceptual learning

Earlier human imaging studies have shown that perceptual learning increases[16–19] or decreases[20,21] activation in early retinotopic areas. It has also been reported that training initially enhances V1 activation, and that after the observers' behavioral performance is saturated further training results in decreased V1 activation[22]. Our current study showed that perceptual learning both facilitated neurons encoding the target and inhibited those responding to the background; therefore, the overall activation in V1 would depend on the balance of excitation and inhibition. In particular, for contours of high saliencies the facilitation of near sites was little affected by training but the suppression of far sites was gradually increased, suggesting that task difficulty would affect the overall activation in V1. Our findings, together with previous mixed imaging results, suggest that an overall increase and decrease in V1 population activity are not reliable neural correlates of perceptual learning. It is the enhancement in selectivity of the neuronal ensemble that contributes to the improved discrimination ability; increasing the signal strength for a particular task might involve facilitation as well as suppression of neural responses.

An ongoing debate on perceptual learning is whether the improvement takes place at the stage of sensory encoding or readout[47,48]. Our decoding analyses suggest that perceptual training does increase the task-relevant information encoded at the earliest stage of visual cortical processing. Notably, monkey MJ's performance at the initial training stage was worse than an artificial classifier, implying that the readout system in the animal was suboptimal. However, the convergence of MJ's performance with the classifier at the end of training suggests that the readout and decision processes are refined as well. Our observations therefore support the idea that perceptual learning modifies multiple stages of information processing, from encoding to readout.

Another debate on the neural code of perceptual learning is about the role of internal noise reduction. Some behavioral and modeling

studies suggest that a reduction in internal noise could contribute to improved perceptual ability[48,49], but it has been argued that perceptual learning does not affect the internal noise[50]. Our series of decoding analyses revealed that enhancement in neural signals associated with population figure-ground contrast largely accounts for the learning-induced improvement in contour detection and that a decrease in noise fluctuations has a minimal role. This finding is consonant with a recent study showing that a general reduction of noise correlation in trained animals has little contribution to decoding of population responses in a multisensory cortical area[29].

Taken together, our findings shed light on the neural mechanisms of visual perceptual learning by showing that training can shape neural population code in early visual cortex, leading to a marked increase in task-related signals at the earliest stage of sensory processing and allowing for a more efficient readout of task-relevant information.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank X. Xu and F. Wang for technical assistance. This work was supported by the National Basic Research Program of China (973 Program 2014CB846101, 2011CBA00400), the National Natural Science Foundation of China (31125014, 31371109 and 30970983) and the Fundamental Research Funds for the Central Universities of China.

AUTHOR CONTRIBUTIONS

M.C., Y.Y. and W.L. designed the experiments. Y.Y., M.C. and X.X. conducted the experiments. M.J.R., Y.Y., M.H. and S.W. analyzed the data. M.J.R. performed the population decoding analyses. M.J.R. and Y.Y. prepared the figures. M.J.R., Y.Y. and W.L. wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Sasaki, Y., Nanez, J.E. & Watanabe, T. Advances in visual perceptual learning and plasticity. *Nat. Rev. Neurosci.* **11**, 53–60 (2010).
2. Sagi, D. Perceptual learning in vision research. *Vision Res.* **51**, 1552–1566 (2011).
3. Gilbert, C.D. & Li, W. Adult visual cortical plasticity. *Neuron* **75**, 250–264 (2012).
4. Schoups, A., Vogels, R., Qian, N. & Orban, G. Practising orientation identification improves orientation coding in V1 neurons. *Nature* **412**, 549–553 (2001).
5. Crist, R.E., Li, W. & Gilbert, C.D. Learning to see: Experience and attention in primary visual cortex. *Nat. Neurosci.* **4**, 519–525 (2001).
6. Li, W., Piech, V. & Gilbert, C.D. Learning to link visual contours. *Neuron* **57**, 442–451 (2008).
7. Hua, T. *et al.* Perceptual learning improves contrast sensitivity of V1 neurons in cats. *Curr. Biol.* **20**, 887–894 (2010).
8. Rainer, G., Lee, H. & Logothetis, N.K. The effects of learning on the function of monkey extrastriate visual cortex. *PLoS Biol.* **2**, e44 (2004).
9. Yang, T. & Maunsell, J.H.R. The effect of perceptual learning on neuronal responses in monkey visual area V4. *J. Neurosci.* **24**, 1617–1626 (2004).
10. Raiguel, S., Vogels, R., Mysore, S.G. & Orban, G.A. Learning to see the difference specifically alters the most informative V4 neurons. *J. Neurosci.* **26**, 6589–6602 (2006).
11. Adab, H.Z. & Vogels, R. Practicing coarse orientation discrimination improves orientation signals in macaque cortical area V4. *Curr. Biol.* **21**, 1661–1666 (2011).
12. Bartolucci, M. & Smith, A.T. Attentional modulation in visual cortex is modified during perceptual learning. *Neuropsychologia* **49**, 3898–3907 (2011).
13. Petrov, A.A., Dosher, B.A. & Lu, Z.L. The dynamics of perceptual learning: An incremental reweighting model. *Psychol. Rev.* **112**, 715–743 (2005).
14. Law, C.-T. & Gold, J.I. Neural correlates of perceptual learning in a sensory motor, but not a sensory, cortical area. *Nat. Neurosci.* **11**, 505–513 (2008).
15. Zhang, G.L., Cong, L.J., Song, Y. & Yu, C. ERP P1-N1 changes associated with Vernier perceptual learning and its location specificity and transfer. *J. Vis.* **13**, 19 (2013).
16. Schwartz, S., Maquet, P. & Frith, C. Neural correlates of perceptual learning: a functional MRI study of visual texture discrimination. *Proc. Natl. Acad. Sci. USA* [comment] **99**, 17137–17142 (2002).
17. Furmanski, C.S., Schluppeck, D. & Engel, S.A. Learning strengthens the response of primary visual cortex to simple patterns. *Curr. Biol.* **14**, 573–578 (2004).
18. Kourtzi, Z., Betts, L.R., Sarkheil, P. & Welchman, A.E. Distributed neural plasticity for shape learning in the human visual cortex. *PLoS Biol.* **3**, e204 (2005).
19. Sigman, M. *et al.* Top-down reorganization of activity in the visual pathway after learning a shape identification task. *Neuron* **46**, 823–835 (2005).
20. Schiltz, C. *et al.* Neuronal mechanisms of perceptual learning: changes in human brain activity with training in orientation discrimination. *Neuroimage* **9**, 46–62 (1999).
21. Mukai, I. *et al.* Activations in visual and attention-related areas predict and correlate with the degree of perceptual learning. *J. Neurosci.* **27**, 11401–11411 (2007).
22. Yotsumoto, Y., Watanabe, T. & Sasaki, Y. Different dynamics of performance and brain activation in the time course of perceptual learning. *Neuron* **57**, 827–833 (2008).
23. Jehee, J.F.M., Ling, S., Swisher, J.D., van Bergen, R.S. & Tong, F. Perceptual learning selectively refines orientation representations in early visual cortex. *J. Neurosci.* **32**, 16747–16753 (2012).
24. Li, W., Piech, V. & Gilbert, C.D. Contour saliency in primary visual cortex. *Neuron* **50**, 951–962 (2006).
25. Li, W. & Gilbert, C.D. Global contour saliency and local colinear interactions. *J. Neurophysiol.* **88**, 2846–2856 (2002).
26. Nugent, A.K., Keswani, R.N., Woods, R.L. & Peli, E. Contour integration in peripheral vision reduces gradually with eccentricity. *Vision Res.* **43**, 2427–2437 (2003).
27. Chen, M. *et al.* Incremental integration of global contours through interplay between visual cortical areas. *Neuron* **82**, 682–694 (2014).
28. Li, W., Piech, V. & Gilbert, C.D. Perceptual learning and top-down influences in primary visual cortex. *Nat. Neurosci.* **7**, 651–657 (2004).
29. Gu, Y. *et al.* Perceptual learning reduces interneuronal correlations in macaque visual cortex. *Neuron* **71**, 750–761 (2011).
30. Rosenblatt, F. The perceptron: a perceiving and recognizing automaton (Report 85–460–1) (Cornell Aeronautical Laboratory, 1957).
31. Maass, W., Natschlager, T. & Markram, H. Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Comput.* **14**, 2531–2560 (2002).
32. Bauer, R. & Heinze, S. Contour integration in striate cortex. Classic cell responses or cooperative selection? *Exp. Brain Res.* **147**, 145–152 (2002).
33. McManus, J.N.J., Li, W. & Gilbert, C.D. Adaptive shape processing in primary visual cortex. *Proc. Natl. Acad. Sci. USA* **108**, 9739–9746 (2011).
34. Zipser, K., Lamme, V.A. & Schiller, P.H. Contextual modulation in primary visual cortex. *J. Neurosci.* **16**, 7376–7389 (1996).
35. Roelfsema, P.R., Tolboom, M. & Khayat, P.S. Different processing phases for features, figures and selective attention in the primary visual cortex. *Neuron* **56**, 785–792 (2007).
36. Poort, J. *et al.* The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. *Neuron* **75**, 143–156 (2012).
37. Roelfsema, P.R. & Spekreijse, H. The representation of erroneously perceived stimuli in the primary visual cortex. *Neuron* [see comments] **31**, 853–863 (2001).
38. Supèr, H., Spekreijse, H. & Lamme, V.A. Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). *Nat. Neurosci.* **4**, 304–310 (2001).
39. Lee, T.S., Yang, C.F., Romero, R.D. & Mumford, D. Neural activity in early visual cortex reflects behavioral experience and higher-order perceptual saliency. *Nat. Neurosci.* **5**, 589–597 (2002).
40. Piëch, V., Li, W., Reeke, G.N. & Gilbert, C.D. Network model of top-down influences on local gain and contextual interactions in visual cortex. *Proc. Natl. Acad. Sci. USA* **110**, E4108–E4117 (2013).
41. Byers, A. & Serences, J.T. Exploring the relationship between perceptual learning and top-down attentional control. *Vision Res.* **74**, 30–39 (2012).
42. Gilbert, C.D. & Li, W. Top-down influences on visual processing. *Nat. Rev. Neurosci.* **14**, 350–363 (2013).
43. Shiu, L.P. & Pashler, H. Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Percept. Psychophys.* **52**, 582–588 (1992).
44. Ahissar, M. & Hochstein, S. Attentional control of early perceptual learning. *Proc. Natl. Acad. Sci. USA* **90**, 5718–5722 (1993).
45. Poggio, T., Fahle, M. & Edelman, S. Fast perceptual learning in visual hyperacuity. *Science* **256**, 1018–1021 (1992).
46. Karni, A. & Sagi, D. The time course of learning a visual skill. *Nature* **365**, 250–252 (1993).
47. Dosher, B.A. & Lu, Z.L. Mechanisms of perceptual learning. *Vision Res.* **39**, 3197–3221 (1999).
48. Bejjanki, V.R., Beck, J.M., Lu, Z.L. & Pouget, A. Perceptual learning as improved probabilistic inference in early sensory areas. *Nat. Neurosci.* **14**, 642–648 (2011).
49. Dosher, B.A. & Lu, Z.L. Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proc. Natl. Acad. Sci. USA* **95**, 13988–13993 (1998).
50. Gold, J., Bennett, P.J. & Sekuler, A.B. Signal but not noise changes with perceptual learning. *Nature* **402**, 176–178 (1999).

ONLINE METHODS

Animal preparations. The animals were prepared under general anesthesia induced with ketamine (10 mg per kg of body weight) and maintained, after intubation, by ventilation with O_2 (100%) mixed with isoflurane (1.5–2.5%). A titanium post was attached to the skull with bone screws for immobilizing the animal's head during training and experiments. The monkeys were first trained to perform a simple fixation task by maintaining their fixation at a small fixation point (FP) and responding to a slight change in luminance of the FP. After the fixation training, a 6 × 8 microelectrode array (Blackrock Microsystems) was implanted in the superficial layers of V1. The microelectrodes were 0.5 mm long and 0.4 mm apart. All experimental procedures were conducted in compliance with the US National Institutes of Health Guide for the Care and Use of Laboratory Animals and were approved by the Institutional Animal Care and Use Committee of Beijing Normal University.

Visual stimuli and behavioral tasks. Visual stimuli were generated by a stimulus generator (ViSaGe, Cambridge Research Systems) on a 22-inch CRT monitor (Iiyama Vision Master Pro 514, 1,200 × 900 pixels, 100-Hz refresh rate, 100-cm viewing distance).

The stimuli were circular patterns with a diameter of 8.5° (for monkey MG) or 7.5° (for MJ), composed of white (13.32 cd m^{-2}) line segments on a gray (4.44 cd m^{-2}) background (**Fig. 1c**). A visual contour was formed by 1, 3, 5 or 7 collinear bars embedded in a background of randomly oriented bars. Each bar was 0.25° long and 0.05° wide, distributed in hidden rhomboid grids with equal base and height of 0.5°. The center-to-center spacing between collinear bars was 1.0, 1.2, 1.4, 1.6, 1.8 or 2.0× the height of a rhomboid compartment, which was adjusted by introducing a skew angle to the hidden rhomboid grids dividing the stimulus pattern (for details, see ref. 24). In addition to the stimulus pattern containing an embedded contour (CP), another stimulus pattern without any embedded contour (NP) was also displayed simultaneously in a trial (**Fig. 1b**). With equal probability and in a pseudorandom sequence, either of the two patterns was centered on the cluster of RFs recorded by the microelectrode array while the other pattern was presented in the opposite visual field. Note that the location of the stimulus patterns was fixed in the visual field; the orientation and position of the embedded contour were precisely adjusted by rotation and translation of the invisible rhomboid grids.

On each trial a 0.1° FP was displayed in the CRT center. Eye positions were sampled at 30 Hz by an infrared tracking system (Matsuda, K., Nagami, T., Kawano, K. & Yamane, S. A new system for measuring eye position on a personal computer. *Soc. Neurosci. Abstr.* **744.2**, 2000). A trial began when the animal fixated at the FP in an invisible circular window of 0.6° in radius around the FP. After the animal maintained its fixation for 300 ms, a pair of CP and NP stimuli was displayed for 500 ms, followed by a blank interval of 300 ms. A trial was aborted if the animal's fixation moved outside the fixation window. After the fixation period, two bright dots were displayed at the two locations where the CP and NP stimuli had been presented (**Fig. 1b**). The animal was required to make a saccade to the dot corresponding to the CP location in 800 ms in exchange for a reward of a drop of juice. If the number of collinear lines in the CP was one (the control conditions), the two stimulus patterns were exactly identical; the animal was randomly rewarded with equal probability by choosing either of the two targets.

Operational training. To let the monkeys understand the contour detection task, they had to go through a stage of operational or procedural training (for details, see ref. 6). Two measures were taken to minimize the effects of perceptual learning before data collection. First, the operational training was carried out in the other two visual-field quadrants orthogonal to those two used for examining the perceptual learning effects (based on a previous finding that contour detection learning in monkeys is specific to the trained visual-field location[6]). Second, the embedded collinear contours were highlighted by a noticeable luminance contrast with the noise background. About a week later, when the monkeys could reliably choose highlighted contours of various orientations and lengths, the two stimuli were moved to the visual-field quadrants to be examined, and the luminance of both contour and background components was set and fixed at 13.32 cd m^{-2}. To ensure that perceptual learning indeed took place at the new stimulus locations, we monitored the animal's daily mean detection rate in each stimulus condition. Compared with the first day at the new stimulus locations, if the detection rate in any of the stimulus conditions significantly increased (χ^2 test, $P < 0.05$), the preceding day was considered to be the first day of perceptual training. For monkey MG, a significant learning effect occurred in the second day at the new stimulus locations, so the first day was taken as the start point of perceptual learning. For monkey MJ, no significant learning effect was seen until the sixth day; that is, there was no significant improvement during the first 5 d in any of the 48 stimulus conditions, so the fifth day was taken as the start of perceptual learning (that is, day 1 in all the figures).

Electrophysiological recording. The spiking signals were amplified and bandpass filtered (250–7,500 Hz) by a data acquisition system (Cerebus, Blackrock Microsystems). We set a voltage threshold to include as many spike waveforms as possible, which were sampled at 30 kHz for further offline denoising. To remove noise from the recorded waveforms, we adopted a robust method[51], which uses mixtures of multivariate *t* distributions to model the statistics of spike waveforms, and uses an expectation-maximization based mixture decomposition algorithm to sort neural spikes from noise. No attempt was made to isolate single units, but this would not affect examination of learning-induced changes in V1.

RF mapping. The RF location and size along the horizontal (azimuth) and vertical (elevation) axes for each recording site were quantitatively mapped using a narrow (0.3° × 7°) band of drifting gratings[27]. Neuronal responses as a function of stimulus position were fitted with a Gaussian function. The RF center was defined as the Gaussian center, and the RF width and length as 2 × 1.96 s.d. (95% confidence interval). The mean RF sizes along the horizontal and vertical axes were 0.98 ± 0.39° (mean ± s.d.) and 0.75 ± 0.26° in MG at eccentricities between 4.8° and 6.2°, and were 1.08 ± 0.34° and 1.04 ± 0.30° in MJ at eccentricities between 6.4° and 8.2° (**Fig. 1a**). Orientation tuning curves of all the sites were determined by presenting a circular patch of drifting gratings (6° in diameter) centered on the cluster of recorded RFs. Only recording sites that showed clear RF profiles and orientation tuning curves (the R^2 of the Gaussian fitting ≥ 0.7) were included in data analyses. The total number of electrodes that were able to pick up recordable spiking activity slightly fluctuated across days (MG, 37–43; MJ, 39–43; over more than 2 weeks of experimental period, likely due to slight changes of electrode depth). Nevertheless, the RF centers and preferred orientations, as long as measurable, were very stable over the course of training (**Fig. 1a**).

Calculating the *d′* value. The detectability d', which measures the discriminability between two distributions of neuronal responses S_1 and S_2, is defined as

$$d' = \frac{E[S_1] - E[S_2]}{\sqrt{0.5(V[S_1] + V[S_2])}}$$

where E denotes the mean firing rate (100–550 ms after stimulus onset) averaged across trials, and V represents the corresponding variance. In this study, S_1 and S_2 correspond to the CP and NP responses, respectively.

Separating the near and far sites. To determine which sites belonged to the near or far group, we calculated the d' value of individual sites and binned them according to the RF-contour distance (**Fig. 3a**). Aggregating all data from the learned conditions in the last 5 d of training, we fitted a Difference of Gaussians (DoG) function to the response profile across the near and far sites. The point where the DoG curve intersected the line corresponding to zero d' value was taken as the boundary of the near and far sites. Defined in this manner, the total numbers of near and far sites were comparable for each contour position despite a small fluctuation across days (Pos1: MG near, 18–22; MG far, 16–23; MJ near, 17–23; MJ far, 20–24; Pos2: MG near, 16–23; MG far, 19–24; MJ near, 19–23; MJ far, 17–22; **Fig. 1a**).

Calculating the onset time of contour modulation. To assess the time point when the contour modulation became significant after stimulus onset, we pooled all available spike trains across the learned conditions and recording sites (separately for near and far sites, and for CP and NP responses). We binned the spike trains into 2.5-ms consecutive intervals and computed the d' values between CP and NP responses for each time bin and each day (**Fig. 4d**). To determine whether the binned d' values were significantly different from zero, we randomly shuffled the labels of CP and NP stimuli and computed bootstrapped d' values (1,000 iterations). The original d' value was regarded as significant if it was beyond the 95%

doi:10.1038/nn.3805

confidence interval of the bootstrapped data. This process was repeated for each time bin. We defined the onset latency of contour facilitation and background suppression as the time point when three adjacent time bins first became significantly larger or smaller than zero (**Fig. 4d**).

Decoding V1 population responses using different classifiers. Mean firing rates (100–550 ms) were extracted from individual electrodes, with the population mean subtracted to remove any overall bias. CP and NP responses were labeled to train the classifiers on data collected from each condition and on each day (that is, the classifiers were retrained for each condition and each day). This re-training process ensured that the decoder was capable of adjusting to changed population code. Data was randomly divided into ten subsets to avoid overfitting. One set was used for testing and the others for training. After cycling through the data sets so that each subset of data was used once for testing, the mean classification rate was calculated. We repeated this entire procedure five times and averaged the resulting mean performances to reduce variability. For SVM calculations using Matlab's *svmtrain* function, the radial-basis function kernel widths were set to half of the median class distance estimated from the training data set, which has been shown to yield satisfactory results[52,53]. The FLD direction was computed by solving a generalized eigenvector equation[54]. DoM (difference of means) and FLD directions were estimated from the training data set. Since the inverse of a covariance matrix determines the FLD, we needed a robust estimate of the covariance. We thus neglected those dimensions with very low fluctuations by preprocessing the data with a principal component analysis (PCA) and keeping the 30 largest components (from ~40 recording sites per day). After projection of the data according to the DoM and FLD methods, the data points were fitted by normal distributions, and classification was made based on the likelihood ratios (using Matlab's *classify* function).

Noise correlation. We first binned the responses of each site in each trial into 50-ms bins and computed the trial-to-trial noise correlation using the correlation coefficient between binned spike counts from any two recording sites. We then averaged the resulting values between 100–550 ms after stimulus presentation and across the learned conditions, for pairs of sites less than 0.5° apart. The paired sites were classified into three groups according to their RF locations relative to the contour: near-near (N-N), near-far (N-F) and far-far (F-F) sites. Using an iterative least-squares algorithm (Matlab's *robustfit*), we determined the slope of noise correlation change over days. We found that the noise correlations in all cases about halved toward the end of training (one sample *t* test, whether slope is different from zero): for MG, the noise correlation on the first day was 0.14 (N-N), 0.16 (N-F) and 0.14 (F-F), and it decreased by 35% ($P = 7 \times 10^{-4}$), 44% ($P = 2 \times 10^{-4}$) and 64% ($P = 0.009$), respectively, on the last day of training; similarly, for MJ, the mean noise correlations decreased from 0.06 (N-N), 0.07 (N-F) and 0.06 (F-F) by 57% ($P = 4 \times 10^{-5}$), 52% ($P = 5 \times 10^{-5}$) and 77% ($P = 0.001$), respectively.

A **Supplementary Methods Checklist** is available.

51. Shoham, S., Fellows, M.R. & Normann, R.A. Robust, automatic spike sorting using mixtures of multivariate t-distributions. *J. Neurosci. Methods* **127**, 111–122 (2003).
52. Gretton, A., Borgwardt, K.M., Rasch, M.J., Scholkopf, B. & Smola, A. A kernel two-sample test. *J. Mach. Learn. Res.* **13**, 723–773 (2012).
53. Rasch, M.J., Gretton, A., Murayama, Y., Maass, W. & Logothetis, N.K. Inferring spike trains from local field potentials. *J. Neurophysiol.* **99**, 1461–1476 (2008).
54. Fisher, R.A. The use of multiple measurements in taxonomic problems. *Ann. Eugen.* **7**, 179–188 (1936).