Neural information processing and its neuronal basis supporting visual object cognitive functions

Graduate School of Science and Engineering
Kagoshima University
Interdisciplinary Department of Science and Engineering

RIDEY HSIAO WANG

INTRODUCTION

In our daily life, an object or a face can be easily recognized from others. Understanding the underlying neural mechanism of visual recognition is one of the most challenging topics in neuroscience.

Research background

Visual object recognition is the ability to identify objects based on visual input. The most difficult and challenging part of visual object recognition is to understand the underlying neuronal mechanism on recognition invariance. We can usually identify an object from others regardless of changes in illumination, object pose, background context, and so on. Different illumination, different pose, or different background context results different input retinal images. The underlying neural basis and information processing creating the object recognition invariance across such changes is critical to understand the brain function of object recognition. In my study, attempts have been paid to understand the invariance across the change of object pose, or viewing angle.

Neuronal representation of object

Our visual system finds objects in the real world from retinal images using object model. In cognitive psychology, two types of models for object representation in the brain have been proposed, the object-centered representation and viewer-centered representation[1-4]. Viewer-centered representation captures shapes at a particular view, whereas object-centered representation represents the intrinsic 3D shape. Although the models were proposed several decades ago and extensive successive studies have been conducted after that, it still remains controversial. A large part of recent study suggests that object shape is represented in a viewer-centered fashion, however, others still report the existence of object-centered representations.

One of the most critical factors affecting object recognition or discrimination performance is similarity between targeted objects. As we can easily

discriminate an object from others if they appear largely different, lower degree of inter-object similarity leads better performance[4-8]. Such recognition can be conducted by build an internal description of the object based on its component parts. On the other hand, similar objects share largely with their part-based description. To achieve enough performance for object recognition with high inter-object similarity objects, it usually requires additional learning. The recognition performance often depends on changes in object pose or viewpoint.

Familiarity of targeted objects is another critical factor affecting object recognition performance. A novel object often cannot be distinguished from similar distractors when the viewing angle changes, recognition across changes in viewing angle develops through the familiarization of the objects or learning[9]. Object recognition learning is thought for the objects to become increasingly differentiated from each other, and for different views of the same object to become increasingly associated.

Electrophysiological findings

With the help of single cell recording technique, electrophysiological studies have been conducted extensively in the past several decades to investigate activities of single cells in various cortical areas. Inferotemporal cortex, the cortical area locates as the last stage of ventral cortical pathway, has been demonstrated repeatedly to be critical for object recognition and discrimination[10-16]. In monkey inferotemporal cortex, Perrett et al.[17] demonstrated responsibility of a small proportion of cells to all views of a particular person's head or of a particular object. The response selectivity of object across different views might underlie the tolerance of perception to changes in viewing angle. In other studies, inferotemporal cells did not show perfect selectivity to all the viewing angles but moderately tunings for viewing angle of 15-50 deg from the optimal[18-20]. Such cells occupied a large proportion of inferotemporal cells, which might also contribute to the tolerance of perception to changes in viewing angle.

In addition, as a noninvasive means evaluating electrophysiological

response to presentation of object image, even-related potential (ERP) was introduced in my research with human subject. Object- or face-sensitive responses have been reported at the first negative component of N1 at the occipito-temporal area[21-24]. A previous study[25] showed that the N1 was significantly modulated by the tolerance of perception to changes in viewing angle. Such result demonstrates the sensitivity of posterior N1 to the association of different viewpoint images of the same object during three-dimensional object recognition. However, underlying mechanism remains largely unknown.

Objectives

My work is focused on object as well as face recognition, particularly the neuronal basis for view-invariant recognition. With the use of the electrophysiological research tools, my experiments can be divided into two parts. The first part of research is with the subject of animal. I targeted the question of how fully invariant object recognition is completed, if indeed it is, in IT. I clarified the processing of view-invariant object discrimination in IT by using population activity, in addition to the analysis at the single-cell level. With human subjects, the purpose of my study is to investigate if ERP is sensitive enough to detect the view-invariant object recognition component.

METHODS

All studies were carried under the regulation of Kagoshima University. Animals were cared for in accordance with Guiding Principles for Care and Use of Animals in the Field of Neuroscience from Japan Neuroscience Society. Both animal and human studies were approved by Kagoshima University.

Electrophysiological study in early visual cortices

All experimental procedures were conducted under anesthesia. Data were obtained from cat area 17 and 18. The activity of the exposed cortical area was measured first with optical imaging based on intrinsic signals and then

electrophysiological recording with electrode array. Optical imaging was conducted once for each hemisphere at the beginning of the experiment to map the functional organization of the exposed cortical area. An array electrode with 5×5 recording points (Blackrock microsystems, USA) were used. Receptive fields together with their sizes and preferred orientations for all the encountered cells were determined. Signals were recorded by using CerePlex Direct (Blackrock Microsystems, USA), and stored on a computer hard disc for offline analysis. Spikes were sorted using Spike2 software (CED, UK), and cross-correlations of the spikes recorded from different recording points were analyzed using programs coded with MATLAB (Mathworks, USA).

Visual stimuli were generated and presented using ViSaGe (Cambridge Research Systems, UK). Fig. 1 shows an example of the visual stimuli used in the experiment. Despite the change in orientation, there were two types of stimuli. Full-field stimuli of square-wave gratings were in the size of 30° × 40° (Fig. 1a). The orientations of the gratings were 0° (horizontal), 22.5°, 45°, 67.5°, 90° (vertical), 112.5°, 135°, and 157.5°. Another type of stimuli were center-surround stimuli

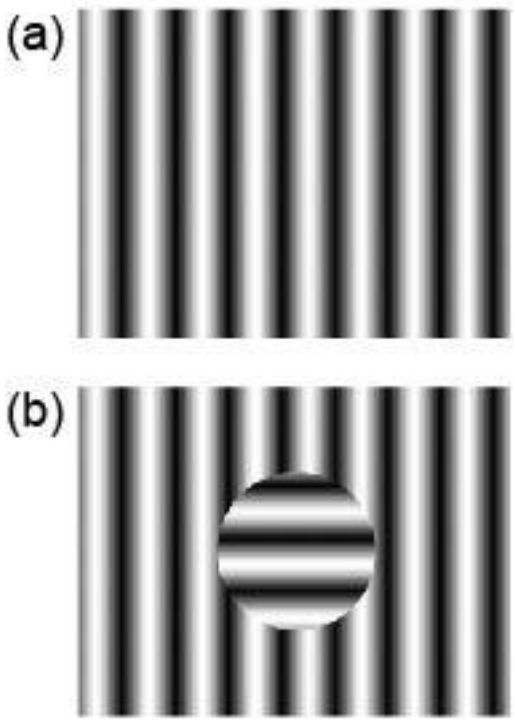

Figure 1. Schematic images of examples of a full-field stimulus (a) and a center-surround stimulus (b).

which consisted of a 5° central circular patch of a square-wave grating and a surround square-wave grating with different orientations (Fig. 1b). The orientation and moving direction of the gratings at the central patch and its surround were controlled separately. The initial phase of the center patch sinusoidal gratings was randomized over the same stimulus repetitions. In the central patch, the orientations of the gratings were 0°, 45°, 90°, and 135°. The orientations of the surround gratings were either the same as or 90° different from the orientations of the gratings presented in the central patch. During recording, the gratings were moved perpendicularly to their orientations at a temporal frequency of 4 Hz.

Electrophysiological study in inferotemporal cortex

Electrophysiological recordings with electrodes were conducted in monkey inferotemporal cortex. Monkeys were trained for months to have prior experience with the object images before electrophysiological recording. For each object set (Fig. 2), four artificial objects were created by changing the parameters defining a three-dimensional prototype in different ways[25]. Four views were created by rotating the object in depth at 30° intervals. Therefore, an object set included 16 images. The difficulty of discrimination was adjusted sot that the performance was 80% for human subjects. The average size of object images was 6.5°.

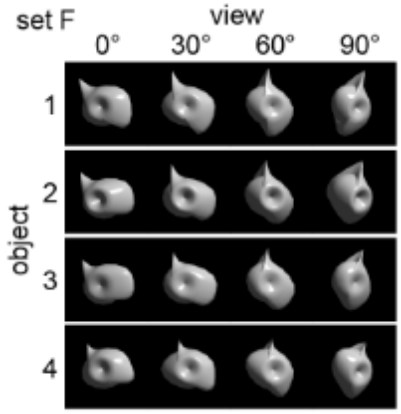

Figure 2. Images of an example stimulus object set.

Before electrophysiological recording, the monkeys were trained to have prior experience of the object images. In the tasks for training, what the monkeys had to do was the same, from a series of successively presented images detecting a

change in object identity but ignoring the viewing angle change of the same object. A trial was initiated by press-down of a lever placed in front of the monkey chair. After fixation point, two to five images appeared sequentially. In a trial, the last image was always an image of a different object from previously viewed object. Stimulus presentation period and inter-stimulus interval were set to 0.5 s. During a trial, monkeys had to keep pressing the lever during the presentation of the first object image(s) and release the lever immediately if a second object appeared. In a trail, the animals had to keep their eyes on a ± 2.5° area around the fixation point.

The prior experiences were applied in three different ways with object task, Within-set Image task, and Across-set Image task respectively. For the Within-set Image task, an identical image of the first object appeared repeatedly for one to four times before the image of a second object at the same viewing angle appeared. Thus, in the Within-set Image task, the monkey made discrimination of objects at the same viewing angle but had no chance to experience different views in the same trial. The first and second objects were selected from the same set. The Object task was the same as the Within-set Image task but the view changed randomly during the repetition of the first object. The Object task required association across views. The difference of the Across-set Image task from the Within-set Image task was the selection of second objects. In the Across-set Image task, the second object was selected from a different object set. Thus, the monkey discriminated tiny differences in object shapes within a set in the Within-set Image task, but to detect only a large difference in the Across-set Image task.

Electrophysiological recording of single cells was conducted with Tungsten electrodes to an area lateral to the anterior middle temporal sulcus between 17 and 23 mm anterior to the ear bar position. Analysis focused on neuronal response to the first stimulus image in each trial. To achieve enough number of cells, the data previously obtained in the lab were also included[26-29]. A machine learning algorithm of support vector machine (SVM) was introduced to achieve classifiers for object discrimination. First, response vectors to all the 16 images in an object set were constructed by pooling all the responses of individual cells to the images.

Vectors for the 16 object images of the same object set were grouped. Since an object image was usually repeated more than 10 times, the 16 × 10 vectors were used to train the object discrimination classifier. Randomly, 90% of trials were used for training the classifier and the remaining 10% of trials were used for testing. Cross-validation was used to validate the model's accuracy.

In the present study, the neural distance between a pair of stimulus images was defined by using the similarity of elicited responses in a population of cells. The responses of a population of cells to each of 16 stimuli were arranged in the vectors to calculate Pearson's correlation coefficients (r) for pairs of stimuli. The neural distance between two stimuli was defined as 1 - r. For normalization, the value of the neural distance was then further subtracted by that obtained during the first 60 ms after the stimulus onset.

Electrophysiological study with human subjects

The participants were male university students aged 18 to 26 years with a mean age of 22.6 years. They were divided into two groups. The participants were limited to those with no previous history of neurological or psychiatric disease and were not taking any type of medication. All participants had normal or corrected-to-normal vision.

The face stimulus images used in the experiment are shown in Figure 3. Face images were created based on the human face model, using FaceGen Modeler (Inversions Inc., Canada). First, an upright prototype face was altered by changing the parameter set, defining a 3D face in four different ways to create four daughter faces (Fig. 3a). Four different views of each face were created by rotating the face at 30° intervals around an axis perpendicular to the visual axis that connected the viewer's eyes to the face image. The actual values of the parameters defining a 3D face were adjusted based on the results of a face recognition task performed by a different group of participants (n=7) who were of the same age, cultural, and educational background as the participants enrolled in the electrophysiological study. From the same face prototype, a set of faces with high similarity and a set

with low similarity were created. The face set with high similarity was created first, followed by that with low similarity. For faces with high similarity, the actual final values for each feature were those performing 75–85% on average and with the same discrimination performance across any pair of face image members. Faces with low similarity were created by increasing the actual size of each parameter four times as large as the difference between the high-similarity faces and the prototype face (Fig. 3b).

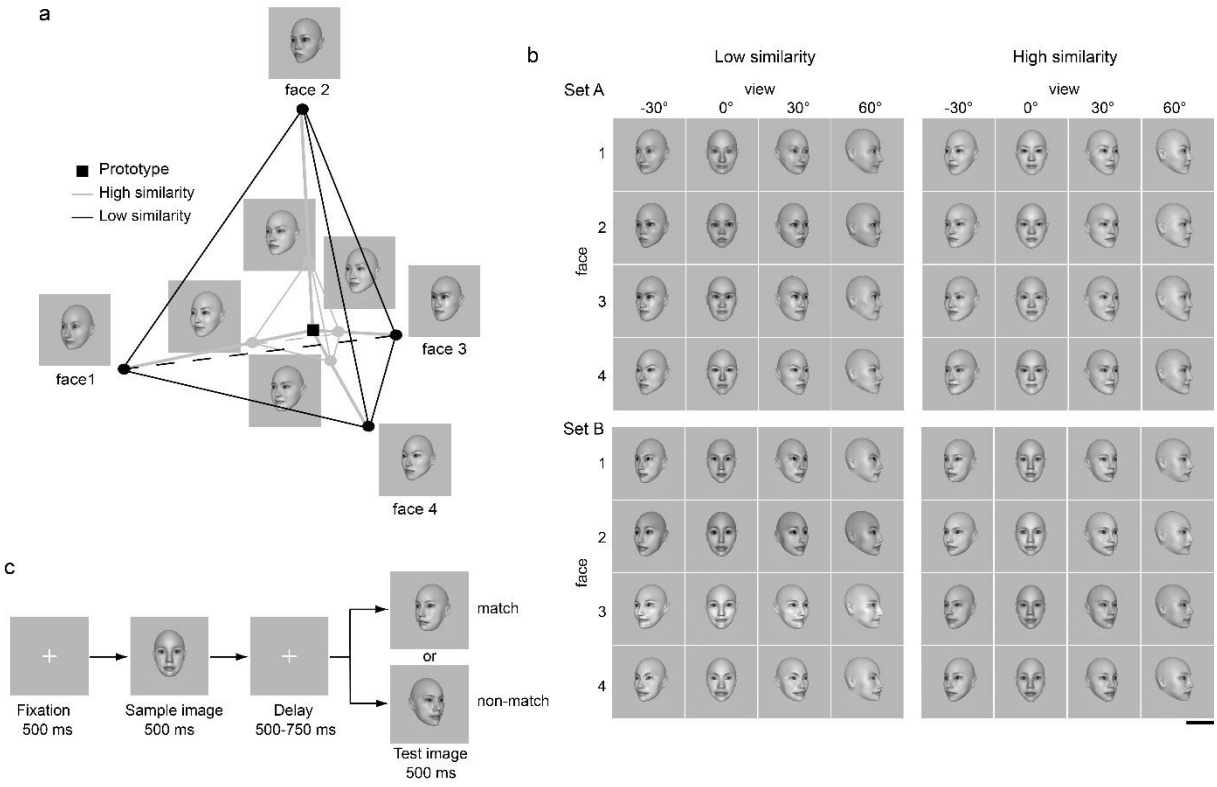

Figure 3. Formation of face sets and task flowchart. (a) Determination of high-similarity and low-similarity faces. (b) Stimulus images are used in the experiment. Bar: 5°. (c) Time sequence of events in the task.

E-Prime with a series input box (Psychology Software Tools, Inc., USA) was used to control the presentation of the stimulus images. As shown in Fig. 3c, a trial started with a mouse click and began with the presentation of the "+" symbol for fixation. Next, a sample image was presented for 500 ms, followed by a 500–

750-ms random interval. The test image was presented with either a matching or nonmatching image for a duration of 500 ms. In each trial, sample and test images were randomly selected from the 16 images included in the same stimulus set. Participants pressed the left mouse button if the test image matched the sample image or the right mouse button if nonmatching. The chances for matching and nonmatching trials were the same (50% for each trial). The image presented as a test always had a different viewpoint from the face image presented as a sample. Participants had to press the mouse button as quickly as possible within a 1-s time window that started from the presentation of the test image.

Participants were divided into two groups. In one group, participants performed the task using high-similarity faces in set A, followed by faces with low similarity in set B. The other group participants were asked to perform the task using the high-similarity faces of set B first, and then the task using the low-similarity faces of set A. The percentage of correct trials was computed for every block, which usually included 300 trials. Participants had to train themselves until their performance was 80% or higher. To balance image exposure, participants were instructed to conduct the same number of trials for tasks using low-similarity and high-similarity faces. Each participant had to complete at least three blocks. Electrophysiological recordings were conducted during the first and last blocks.

EEG was recorded in a darkened, sound-attenuated room using a digital EEG recording system (Neurofax EEG-2110, Nippon Kohden, Japan), using a 19-channel electrode cap (ECI, Electro-Cap International, USA). The recording locations were fixed according to the recognized standard international 10-20 system sites (American Electroencephalographic Society, 1994). The recordings were referenced to linked earlobes. Facial electrodes were used to monitor horizontal and vertical electrooculograms (EOGs). Electrode impedances were kept below 5 kΩ during the recording. EEG signals were filtered with a bandpass filter (0.1–100 Hz) and sampled at a rate of 500 Hz. Trials with excessive eye and body movements (>50 μV) were excluded. The response codes and reaction times were recorded together with the EEG signals.

RESULTS

Electrophysiological findings in early visual cortices

Fig. 4 shows an example of recordings from a hemisphere. Receptive fields

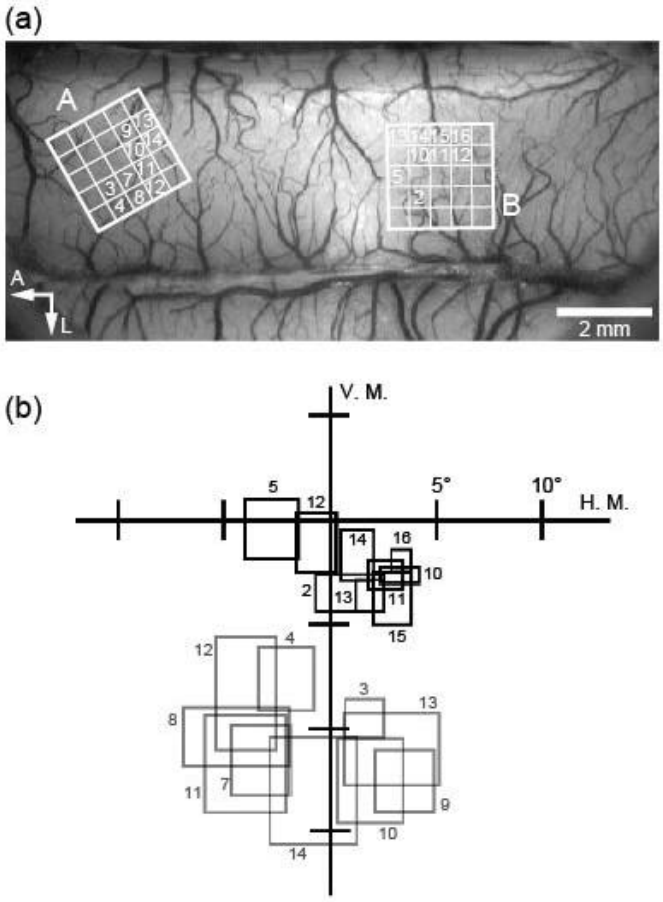

Figure 4. Cortical image and receptive fields of recorded cells. (a) A blood vessel image of the cortical area and positions of array electrodes. L: Lateral, A: Anterior. (b) Receptive fields of the cells recorded in the array electrodes. Gray and black rectangles indicate receptive fields of the cells recorded in the array electrodes A and B, respectively. Numbers correspond to each recording point in (a). V. M., vertical meridian; H. M., horizontal meridian.

of cells recorded with 5 × 5 array electrodes were plotted. The value of cross-correlation between the spike activities of two cells was used to express the interaction between the cells. In order to statistically evaluate the significance of cross-correlation value, a threshold of mean + 5 SD (Standard Deviation) of cross-correlogram value was set. In the peripheral visual field, 11.0% of cell pairs showed higher cross-correlation values than the threshold. In the central visual field, 8.6% of cell pairs showed higher cross-correlation values. It is interesting that higher percentage of cell pairs was confirmed in the periphery of the visual field than in the center.

Cell pairs with higher-than-threshold values were sorted by the extent of overlap of receptive fields of the cells. Fig. 5 plots the percentage of the pairs in either the center or the periphery of the visual field. The percentage was smaller than others with the receptive field overlap of 0-25%. The values tended to be larger as the receptive field overlap increased, however, statistical significance confirmed

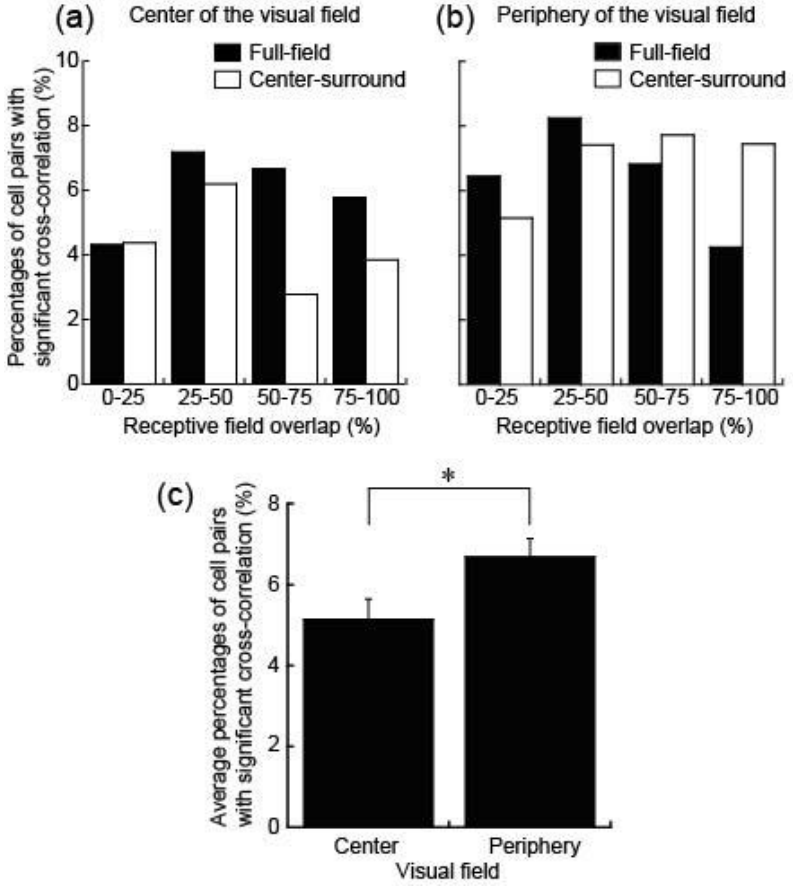

Figure 5. Percentage of cell pairs showing higher-than-threshold cross-correlation vs. overlap of receptive fields plots in

neither in the center (full-field stimuli, ρ = 0.20, P = 0.80, Spearman's rank correlation coefficient; center-surround stimuli, ρ = - 0.60, P = 0.40; Fig. 5a) nor the periphery (full-field stimuli, ρ = - 0.40, P = 0.60; center- surround stimuli, ρ = 0.80, P = 0.20; Fig. 5b) of the visual field. Interestingly, a significantly larger percentage of the cell pairs showing larger-than-threshold was found in the peripheral visual field than that in the center of the visual field ($P < 0.05$, Wilcoxon rank sum test; Fig. 5c).

Temporal response characteristics of inferotemporal cells

Depending on the prior experienced tasks, cells recorded from each animal were divided based on the object sets. The presentation of the 16 images usually evoked different responses of the cell (Fig. 6). Analysis was on the temporal

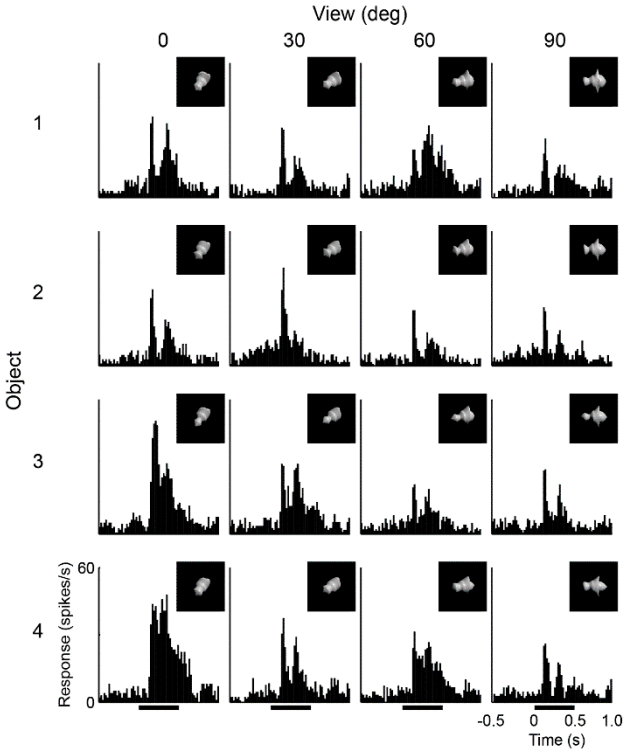

Figure 6. Peri-stimulus time histograms (PSTHs) of a sample IT cell in response to the 16 images in an object set. Images on the upper-right corner of each plot represent stimulus images. The horizontal bar indicates stimulus presentation (0.5 s).

property during the period of response immediately after stimulus presentation. Presentation of a stimulus activated a rapid increase in spike rate peaked, on average, at 107 ms after stimulus onset. A relatively high spike rate usually remained until several hundred milliseconds after removal of stimulus. From the

observation of the shape for histogram, one could realize the difference between the early phase and late phase of responses. In the study, the early phase denoted the period of response in the time period of 100 – 260 ms with a conical shape peaked at 140 ms after stimulus onset. The late phase temporally immediate followed the early phase in the time period of 260 – 660 ms.

In the early phase and late phase of response, investigation was performed on the difference in stimulus selectivity. Preferred stimulus images were determined independently in the early and late response phases. Cells were classified into four types. Type I was defined as the cells with the same stimulus preference in both the early response phase and the late response phase. There was no preference change during response period. Others were further classified into type II, type III and type IV. Type II cell preferred the same object in the late phase but different view from that in the early phase. Type III cells showed a preference for the same views but in different objects. Remaining cells were classified as type IV, which preferred to the images of different objects at different views in the early and late phases.

For the object images experienced in the object task, type I cells constituted 34% of all the cells, and type II cells 32%, while the percentages for type III and type IV cells were 12% and 22%, respectively. For the across-set image task, type I, type II, type III, and type IV cells occupied 33%, 13%, 20% and 34% respectively. The percentage of type I cell was comparable to that found in the object task. Interestingly, the type II cell percentage for the object task was significantly higher than that for the across-set image task. In contrast, the type III cells for the across-set image task occupied significantly larger percentage than that for the object task. The cell type distribution with prior experience of across-set image task significantly differed from the distribution in cell types with prior experience of the object task ($p<0.0001$, Chi-square test). A statistically significant higher percentage of type II cells and a significant lower percentage of type III cells were confirmed.

In inferotemporal cells, response histograms appeared largely different in

their shapes in the early and late phases. To understand their involvement in object recognition across views, I trained object classifiers using the data averaged in the early phase and late phase separately. In the early phase of the response, the performance for responses to the images prior experienced in the object task was 61.6 ± 9.8%, while the performance for the responses to the images prior experienced in the across-set image task was 64.1 ± 7.8%, not being statistically different ($t=0.708$, $df=9$, $p=0.497$). In the late phase, the performance for responses to the images prior experienced in the object task was as high as 79.6 ± 7.4%, significantly different from that for the responses in the across-set image task (66.5 ± 7.4%, $t=4.392$, $df=9$, $p<0.005$).

Effects of overtraining on inferotemporal cell responses

To understand the overtraining induced change, I computed the differences in neural distance by using cell populations pooled over 10 days with a 5-day overlap by recording date, sequentially. Fig. 7 shows the change in difference

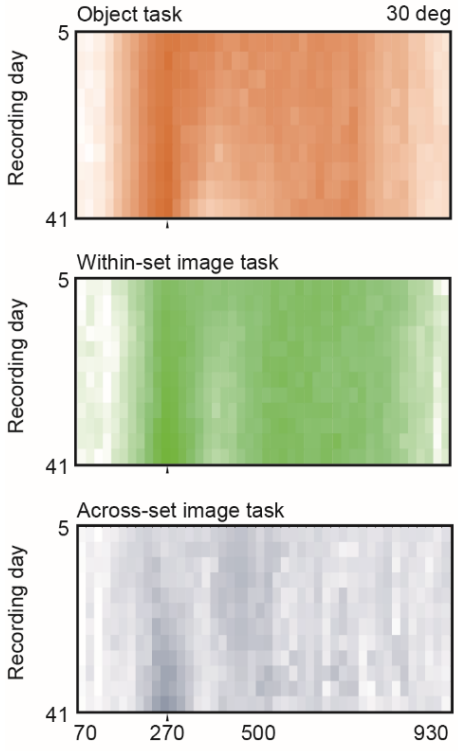

Figure 7. The change in the time course of differences in neural distance during the overtraining period, at a viewing angle separation of 30°.

between the neural distance for views of different objects and the distance for views of the same object by recording date.

The largest difference in distances for views of different objects and the distance for views of the same object was observed for the images experienced in the Object task (Fig. 7, upper). At the beginning of the recording, the difference started at about 100 ms and remained continuously large until 800 ms after stimulus onset. Interestingly, at the end of the recording period, a lowering of the difference could be found at 370 ms after stimulus onset. The curve showed a large peak at about 270 ms after stimulus onset, a transient lowering at about 370 ms, and then another peak at 450–630 ms. For images experienced in the Across-set Image task (Fig. 7, lower), an obvious increase in the difference between neural distances for different object views and the same object views was observed at 270 ms. The dynamics for the images experienced in the Within-set Image task showed similar shape as those of the Object task (Fig. 7, middle). The change was smaller than that in the Object task, however, a significant difference in neural distance could be confirmed in a relatively long period of time. At the beginning of recording, the difference kept continuously high during the period of response, but with a clear drop at 370 ms at the end of the recording.

Event-Related Potentials accompanying to object discrimination learning

We focused on the ERP component N170 in response to the presentation of the sample images in each trial. ERP waveforms at T6, which showed the largest N170 amplitude and voltage topography at the N170 peak delay, are shown in Fig. 8. Topographically, the sample stimulus evokes negativity in a large area, including T5, T6, O1, O2, P3, and P4. The T6 electrode exhibited the largest negativity at the N170 peak latency. We averaged and quantitatively compared the amplitudes and

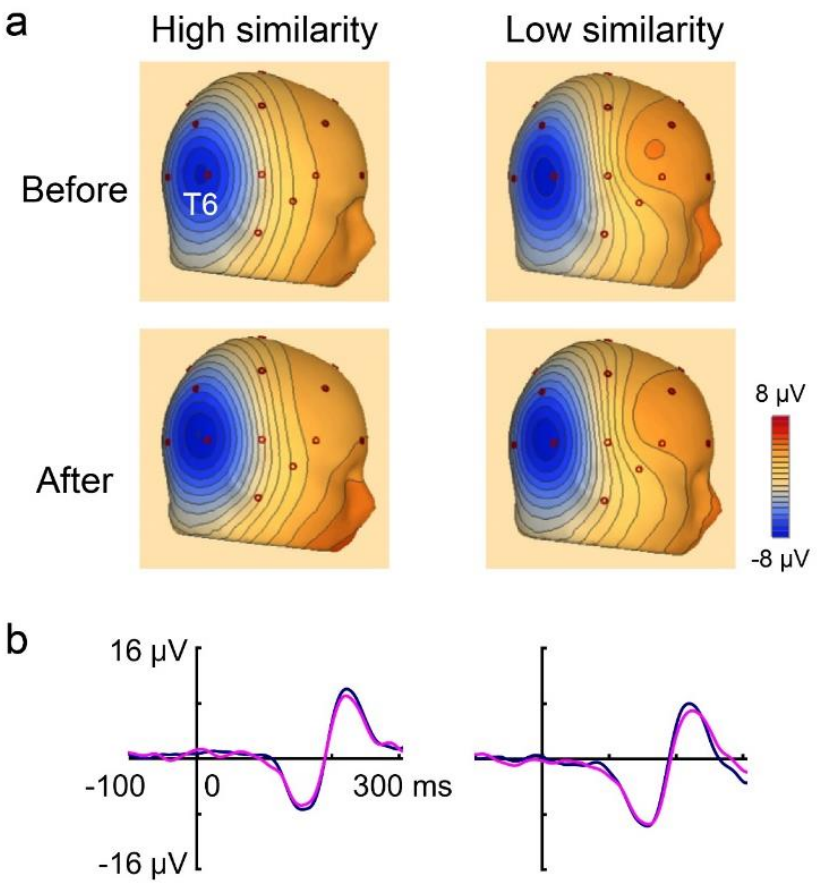

Figure 8. Topography for N170 and grand averaged ERPs activated by the presentation of sample images. (a) Voltage topography for N170 peak. (b) Grand averaged ERP waveforms at T6 for the group of participants. Blue lines show the ERP waveforms obtained before training, and pink lines denote the waveforms obtained after training.

delays obtained before and after training in each participant group.

In addition to the averaged ERPs for all face stimuli in a set, I investigated changes across faces and views in the face recognition learning process. In response to the 16 images in a face set, ERPs were grouped according to whether they belonged to the same face or from the same viewpoint. Variation of N170 exhibited variations in both the amplitude and delay across the faces, as well as across the four viewpoints. The measure of face scatter and the measure of viewpoint scatter were introduced to quantitatively evaluate such variations. The measures were

defined as the standard deviation across the faces and standard deviation across the viewpoints.

Fig. 9 shows the values for face and viewpoint scatter in the N170 amplitude. The results demonstrate the dependence of the similarity. When using the high-similarity faces, the averaged face scatter at the beginning of the training was smaller than that obtained after the training. In both face sets, statistically significant differences were confirmed between the values obtained before and after the training (set A: Fisher's PLSD, $p<0.0001$; set B: Fisher's PLSD, $p<0.0001$). In contrast, if low-similarity faces were used, no significant difference could be confirmed between the values obtained before and after training (set A: Fisher's PLSD, $p=0.8671$; set B: Fisher's PLSD, $p=0.7458$). Although not statistically significant, the scatter obtained when using high-similarity faces before training was consistently smaller than that observed when using low-similarity faces. Interestingly, unlike object learning, viewpoint scatter did not change significantly with face recognition training, regardless of face similarity.

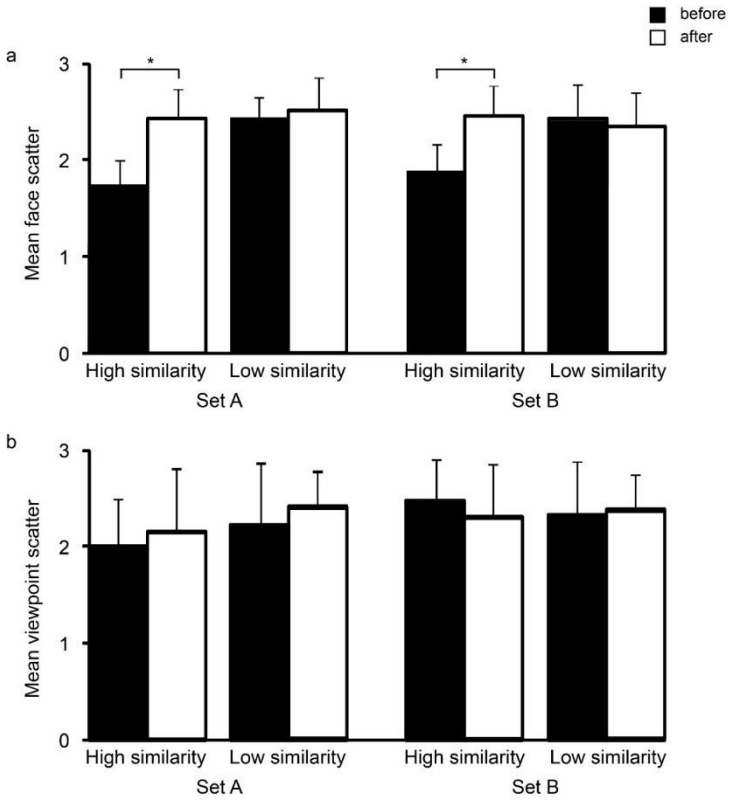

Figure 9. Mean face scatter (a) and viewpoint scatter (b) for the N170 amplitudes obtained before and after training. *$p<0.005$.

In contrast to the N170 amplitude, statistical analyses showed no significant changes for the N170 latency. Training appears to have no significant impact on the variation in the N170 peak latency.

DISCUSSION

Activity of cells in early visual cortical areas corresponding to the central an peripheral visual fields

Cross-correlations of pairs of cells with their receptive fields in the central visual field were compared to those of the cells with their receptive fields in the peripheral visual field. Plenty of studies have been conducted to investigate the cross-correlation of spikes of the cell pairs in the early visual cortex of cats and monkeys[30-38], but few of them were focused on the difference between the cell pairs with their receptive fields in the center and periphery of the visual field. In this study, comparing to the central visual field, a larger percentage of the cell pairs demonstrated cross-correlation in the peripheral visual field. This suggests a kind of difference in representation of surround modulation between the central and peripheral visual field. The cells in the cortical area corresponding to the peripheral visual field were with close interaction with each other.

Comparing to the cell in inferotemporal cortex, the cell in the early visual cortex is with small receptive field. The main function of early visual cortex is to extract contour information through detection of line orientation. The perception of orientation can be modified by the presentation of surround stimulus[39]. Psychophysical experiments demonstrate that the effect of the surround stimulus depends on the location in the visual field. It is larger in the periphery of the visual field than in the center[40]. Another evidence demonstrating the difference between the periphery and the center visual field was reported by Tangen et al.[41], the face distortion effect that the eye-aligned faces presented rapidly in the periphery of the visual field were perceived to be distorted, but after inspecting the faces individually, the faces appeared normal. Anatomically, the neuronal connectivity

for the center of the visual field has been frequently reported to be different from that for the periphery of the visual field[42-45]. Electrophysiological experiments demonstrated that spatial frequency preferences of the cells in the early visual cortex referred to the receptive field eccentricity[46,47]. The cortical magnification factor differs depending on the visual field eccentricity[48]. However, it is repeatedly reported that the cortical magnification factor does not explain the degraded recognition performance in the peripheral vision[49-51].

Temporal characteristics of inferotemporal cell response

Activities in the early phase and the late phase contribute significantly different in object recognition across viewing angles. The effect of two types of object discrimination tasks, one required object recognition across viewing angles and no need for the other, was investigated. With the classifiers trained by the respective responses, the object discrimination performance of the classifier created by the response data to the images experienced in the object task was comparable to that created for the across-set image task in the early phase. In the late phase, the performance for the object task was significantly better than that for the across-set image task. Those results imply that in the late phase, the activity of inferotemporal cells may reflect the neural processing necessarily to achieve generalization across views of the same object. Consistent with previous findings on the investigation of inferotemporal cell selectivity[14,15,26], the activity of inferotemporal cells in the early response phase is more sensitive to the change of two-dimensional image shape. There were several researches on the difference between the early and later response time windows[52-56]. Global categorical information was conveyed in the early time period, and fine information in the late time period[52,55].

View-invariant object discrimination learning did not significantly change the percentage of type I cells. It remains at the level of about 30% of all the recorded cells, regardless of the training task for prior experience. The remaining approximately 70% cells changed their stimulus images in the object set during the

response period. The object task associated the views of the same objects, in the same time, differentiated the images of different objects. Such experience had more of IT cells starting to respond to different views of the same objects in the early and late response phases. This is in the same line of the pairing learning[57] and a finding with long-term shape discrimination learning[58].

When the objects are similar, it is difficult to discriminate from each other if there is change in viewing angle[5,8,59]. Usually, further learning in object recognition is required. There are several models for object recognition across views. Viewing rotating objects is proposed to be critical to establish the capability of view-invariant recognition. Viewing can be done actively through association learning of different views, or it can be passive through experiencing successive views of the object[60-63]. It becomes a consensus among a part of researchers. Such way for association of object views is repeatedly mentioned and discussed as the underlying mechanism in their studies[64-79].

Overtraining induced changes

In the present study, I demonstrated response tolerance in almost full range of viewing angle after overtraining with the Within-set Image task. Overtraining increased the difference between neural distances of the representations of the views for the same object and that for different objects. There exists a threshold in neural distance for separation of the images of the same object from images of other objects.

View-variant object recognition should need generalization of representations of the same object views, and, in the same, differentiation from other objects. Training effect on IT cell responsibility has been studied previously. Object discrimination significantly sharpened the stimulus selectivity of single IT cells[71,72,80] with a decreased response to suboptimal stimuli, rather than as an increase of the response to the best stimulus[81]. Immediately after saturation of object discrimination performance in the Within-set Image task, IT cells showed response tolerant within a range of 30°[26]. Overtraining further broadened the

response tolerance around the trained views, so that the representations of trained neighboring views merged with each other.

The effect of long-term visual experience on single cell responses remains controversial. Learning enhanced the responses of IT cells to the learned stimulus images more than those to unlearned stimulus images[18,58,82]. In other studies, responses to familiar stimuli were reported with comparable or even weaker as compared to novel stimuli[71,81,83,84]. A following study reported different effects of long-term visual experience on the broad-spiking (putative excitatory) cells and narrow-spiking (putative inhibitory) cells[85]. Visual experience increased the responses of putative excitatory neurons but decreased the respones of putative inhibitory neurons. A recent interesting studey demonstrated a wide rage neural network in the temporal and prefrontal cortex to maintain memories for valuable objects[86]. Such memory representation could last for months after the last experience. An alternative explanation of the findings revealed by the current study could be the case that the monkeys were able to find valuable objects efficiently in the course of overtraining.

The image on retina changes drastically if rotating an object in depth. However, we can recognize the object almost effortlessly despite changes in viewing angle[20,66,87-93]. Inferotemporal cortex play a critical role in object recognition[94]. The responses of IT neurons are found to be tolerance to some object attributes[15]. Neuronal responses to familiar natural objects or faces are tolerant in almost all the object views[19,95,96], whereas the response tolerance is narrower for unfamiliar artificial objects[97].

ERP changes accompanying discrimination learning

This study was designed to investigate whether face recognition learning could induce similar electrophysiological changes in object recognition learning [98-100]. This is similar to the N1 change in the object discrimination learning process, regardless of similarity across faces; the ERP component, N170, exhibited no significant change in either amplitude or amplitude delay accompanying

discrimination learning if we measured the mean values of the amplitudes and delays. However, I found, accompanying the face recognition learning, a significant increase in the variation of N170 amplitudes for different faces in the case of high-similarity faces, but not the variation of those for views of the same faces. In contrast, object recognition learning was found to increase in both N1 amplitude and latency variation across objects and, at the same time, decreased N1 amplitude and latency variation across the views of the same objects when using high-similarity faces. These differences between object and face recognition learning suggest differences in neuronal representation for object recognition and face recognition.

Because the high-similarity faces were not all different from each other, discrimination performance across novel high-similarity faces was closer to that by chance. Accompanying the improvement of the performance, the amplitudes of the N170s that were averaged for the views of each of the faces began to differ from each other. The view-invariant face recognition learning used in this study differentiates different faces and, at the same time, generalizes across views of the same face. The N170 variation across faces changed with respect to differentiation of the faces, which could have been achieved during view-invariant face discrimination learning. These findings are consistent with those of our previous studies on high-similarity objects[98-100]. If faces had low-similarity faces, learning was not required to recognize novel faces. Since the recognition performance was good enough from the beginning, the learning process did not lead to any further significant improvement in behavioral performance. Face recognition training did not significantly change the variation across faces. Accompanying the learning, no significant difference was found in the means of the N170 amplitude and delay. Instead, we focused our analyses on the variation of the N170 response to different faces and different views of the faces. In addition to the use of means of the ERP component, the results of this study demonstrate that the measure of variation may provide a way to investigate the difference in neural representation of faces. Instead of real face images, a computer-generated artificial face was used in this study. To

control the similarity level across these faces, the faces used in the study were created by quantitatively altering an artificial prototype in different ways, which made it possible to define the similarity across faces. In addition, using visually novel computer-generated 3D faces, a participant's prior exposure to the stimuli could be both easy and well-controlled. When using this method, the distance in the image base within views of the same object was comparable to that across different faces from the same viewpoint. Therefore, the use of these prototypes excluded the possibility of discrimination based on similarities in the image base. All these can only be realized by using an artificial face, but not with a real face image. Artificial appearance compromises most aspects of facial processing. Compared to real faces, artificial faces are remembered more poorly[101,102] and discriminated less efficiently[101]. Physiologically, N170 sensitivity to species is affected by artificial appearance[103], and the effects of face animacy were evidenced in face processing networks by fMRI[104]. However, clear effects of face animacy were not always evident at the N170[104] but may emerge at later components[106].

The N170 changes accompanying face recognition learning were different from those induced by object recognition learning[100]. Event-related potentials demonstrate different sensitivities to low-level and high-level visual processing. The ERP component, P100, is more sensitive to low-level stimulus differences[107,108]. Sensitivity to race[109,110] or age[110,111] has been reported for the face-selective EPR component N170 and later components[109]. The response contrast of N170 in response to upright and inverted faces evidenced the engagement of high-level processing[112]. Inversion effects on the N170 were larger for own-race faces than other-race faces[113], which reflected disruption of the early stages of high-level face processing. Electroencephalography provides an efficient means to uncover the time course of face versus object processing[114-129].

References

1. Marr, D. (1982) Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. Henry Holt and Co., Inc.,

USA.

2. Marr, D. and Nishihara, H.K. (1978) Representation and recognition of the spatial organization of three-dimensional shape. Philosophical Transactions of the Royal Society of London, B200, 269-294.
3. Biederman, I. (1987) Recognition-by-components: A theory of human image understanding. Psychological Review, 94, 115–147.
4. Biederman, I. and Gerhardstein, P.D. (1993) Recognizing depth-rotated objects and conditions for three-dimensional viewpoint invariance. Journal of Experimental Psychology, 8, 510-523.
5. Tarr, M.J. (1995) Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects. Psychonomic Bulletin and Review, 2, 55-82.
6. Tarr, M.J. and Pinker, S. (1989) Mental rotation and orientation dependence in shape recognition. Cognitive Psychology 21, 233-282.
7. Humphrey, G.K. and Khan, S.C. (1992) Recognizing novel views of three-dimensional objects. Canadian Journal of Psychology, 46, 170-190.
8. Bülthoff, H.H. and Edelman, S. (1992) Psychophysical support for a two-dimensional view interpolation theory of object recognition. Proceedings of the National Academy of Sciences of the United States of America, 89, 60-64.
9. Poggio, T. and Edelman, S. (1990) A network that learns to recognize three-dimensional objects. Nature, 343, 263-266.
10. Gross, C.G., Rocha-Miranda, C.E. and Bender, D.B. (1972) Visual properties of neurons in inferotemporal cortex of the Macaque. Journal of Neurophysiology, 35, 96-111.
11. Tanaka, K., Saito, H., Fukada, Y. and Moriya, M. (1991) Coding visual images of objects in the inferotemporal cortex of the macaque monkey. Journal of Neurophysiology, 66, 170-189.
12. Tarr, M.J., Kersten, D. and Bülthoff, H.H. (1999) Object recognition in man, monkey and machine. MIT Press, Cambridge.
13. Rolls, E.T. (2000) Functions of the primate temporal lobe cortical visual areas

in invariant visual object and face recognition. Neuron, 27, 205-218.

14. Tanaka, K. (1996) Inferotemporal cortex and object vision. Annual Review of Neuroscience, 19, 109-139.
15. Wang, G., Tanaka, K. and Tanifuji, M. (1996) Optical imaging of functional organization in the monkey inferotemporal cortex. Science, 272, 1665-1668.
16. Wang, G., Tanifuji, M. and Tanaka, K. (1998) Functional architecture in monkey inferotemporal cortex revealed by in vivo optical imaging. Neuroscience Research, 32, 33-46.
17. Perrett, D.I. , Smith, P.A.J and Potter, D.D. (1984) Neurons responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception. Human Neurobiology, 3, 197–208.
18. Logothetis, N.K., Pauls, J. and Poggio, T. (1995) Shape representation in the inferior temporal cortex of monkeys. Current Biology, 5, 552-563.
19. Perrett, D.I., Oram, M.W., Harries, M.H., Bevan, R., Hietanen, J.K., Benson, P.J. and Thomas, S. (1991) Viewer-centred and object-centred coding of heads in the macaque temporal cortex. Experimental Brain Research, 86, 159–173.
20. Logothetis, N.K. and Sheinberg, D. L. (1996) Visual object recognition. Annual Review of Neuroscience, 19, 577-621.
21. Bentin, S., Allison, T., Puce, A., Perez, A. and McCarthy, G. (1996) Electrophysiological studies of face perception in humans. Journal of Cognitive Neuroscience, 8, 551-565.
22. Thorpe, S., Fize, D. and Marlot, C. (1996) Speed of processing in the human visual system. Nature, 381, 520-522.
23. Rossion, B. (2014) Understanding face perception by means of human electrophysiology. Trends in Cognitive Science, 18, 310–318.
24. Johnson, J.S. and Olshausen, B.A. (2003) Time-course of neural signatures of object recognition. Journal of Vision, 3, 499-512.
25. Wang, G. and Kameda, S. (2005) Event-related potential component associated with the recognition of three-dimensional objects. NeuroReport, 16, 767-771.

26. Okamura, J.Y., Yamaguchi, R., Honda, K., Wang, G. and Tanaka, K. (2014) Neural substrates of view-invariant object recognition developed without experiencing rotations of the objects. Journal of Neuroscience, 34, 15047-15059.
27. Yamaguchi, R., Okamura, J.Y. and Wang, G. (2016) Dynamics of population coding for object views following object discrimination training. Neuroscience, 330, 109-120.
28. Zhao, C., Wang, R.H. and Wang, G. (2018) Long-term object discrimination at several viewpoints develops neural substrates of view-invariant object recognition in inferotemporal cortex. Neuroscience, 392, 190-202.
29. Okamura, J.Y., Uemura, K., Saruwatari, S. and Wang, G. (2018) Difference in the generalization of response tolerance across views between the anterior and posterior part of the inferotemporal cortex. European Journal of Neuroscience, 48, 3552-3566.
30. Ts'O, D.Y., Gilbert, C.D. and Wiesel, T.N. (1986) Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis. Journal of Neuroscience, 6, 1160-1170.
31. Smith, M.A. and Kohn, A. (2008) Spatial and temporal scales of neuronal correlation in primary visual cortex. Journal of Neuroscience, 28, 12591-12603.
32. Tamura, H., Sato, H., Katsuyama, N., Hata, Y. and Tsumoto, T. (1996) Less segregated processing of visual information in V2 than in V1 of the monkey visual cortex. European Journal of Neuroscience, 8, 300-309.
33. Hung, C.P., Ramsden, B.M. and Roe, A.W. (2007) A functional circuitry for edge-induced brightness perception. Nature Neuroscience, 10, 1185-1190.
34. Roe, A.W. and Ts'o, D.Y. (2005) Specificity of V1-V2 orientation networks in the primate visual cortex. Cortex, 72, 168-178.
35. Uchimura, K., Okamura, J.Y. and Wang, G. (2012) Surround modulation in cortical orientation map revealed by optical imaging and its dependency on receptive field eccentricity. European Journal of Neuroscience, 36, 3344-3355.
36. Gilbert, C.D. and Wiesel, T.N. (1979) Morphology and intracortical projections of functionally characterised neurons in the cat visual cortex. Nature, 280, 120-

125.

37. Gilbert, C.D. and Wiesel, T.N. (1983) Clustered intrinsic connections in cat visual cortex. Journal of Neuroscience, 3, 1116-1133.
38. Okamura, J., Wang, R.H., Kawano, K. and Wang, G. (2021) Differences in the neuronal correlation between the central and peripheral vision. NeuroReport, 32, 899-906.
39. Gibson, J.J. and Radner, M. (1937) Adaptation, after-effect and contrast in the perception of tilted lines. I Quantitative studies. Journal of Experimental Psychology, 20, 453-467.
40. Xing, J. and Heeger, D.J. (2000) Center-surround interactions in foveal and peripheral vision. Vision Research, 40, 3065-3075.
41. Tangen, J.M., Murphy, S.C. and Thompson, M.B. (2011) Flashed face distortion effect: grotesque faces from relative spaces. Perception, 40, 628-630.
42. Baizer, J.S., Ungerleider, L.G. and Desimone, R. (1991) Organization of visual inputs to the inferior temporal and posterior parietal cortex in macaques. Journal of Neuroscience, 11, 168–190.
43. Sousa, A.P., Pinon, M.C., Gattass, R. and Rosa, M.G.P. (1991) Topographic organization of cortical input to striate cortex in the Cebus monkey: a fluorescent tracer study. Journal of Comparative Neurology, 308, 665–682.
44. Palmer, S.M. and Rosa, M.G.P. (2006) A distinct anatomical network of cortical areas for analysis of motion in far peripheral vision. European Journal of Neuroscience, 24, 2389–2405.
45. Ungerleider, L.G., Galkin, T.W., Desimone, R. and Gattass, R. (2008) Cortical connections of area V4 in the macaque. Cerebral Cortex, 18, 477–499.
46. Movshon, J.A., Thompson, I.D. and Tolhurst, D.J. (1978) Spatial and temporal contrast selectivity of neurons in area 17 and 18 of the cat's visual cortex. Journal of Physiology, 283, 101–120.
47. Foster, K.H., Gaska, J.P., Nagler, M. and Pollen, D.A. (1985) Spatial and temporal frequency selectivity of neurons in visual cortical area V1 and V2 of the macaque monkey. Journal of Physiology, 365, 331–363.

48. Tusa, R.J., Palmer, L.A. and Rosenquist, A.C. (1978) The retinotopic organization of area 17 (striate cortex) in the cat. Journal of Comparative Neurology, 177, 213-235.
49. Strasburger, H., Rentschler, I. and Harvey, L.O. (1994) Cortical magnification theory fails to predict visual recognition. European Journal of Neuroscience, 6, 1583-1587.
50. Jüttner, M. and Rentschler, I. (1996) Reduced perceptual dimensionality in extrafoveal vision. Vision Research, 36, 1007-1022.
51. Milleret, C., Buisseret, P. and Gary-Bobo, E. (1988) Area centralis position relative to the optic disc projection in kittens as a function of age. Investigative Ophthalmology and Visual Science, 29, 1299-1305.
52. Sugase, Y., Yamane, S., Ueno, S. and Kawano, K. (1999) Global and fine information coded by single neurons in the temporal visual cortex. Nature, 400, 869–873.
53. Brincat, S.L. and Connor, C.E. (2006) Dynamic shape synthesis in posterior inferotemporal cortex. Neuron, 49, 17–24.
54. Brincat, S.L. and Connor, C.E. (2004) Underlying principles of visual shape selectivity in posterior inferotemporal cortex. Nature Neuroscience, 7, 880–886.
55. Matsumoto, N., Okada, M., Sugase-Miyamoto, Y., Yamane, S. and Kawano, K. (2005) Population dynamics of face-responsive neurons in the inferior temporal cortex. Cerebral Cortex, 15, 1103–1112.
56. Tamura, H. and Tanaka, K. (2001) Visual response properties of cells in the ventral and dorsal parts of the macaque inferotemporal cortex. Cerebral Cortex, 11, 384–399.
57. Sakai, K. and Miyashita, Y. (1991) Neural organization for the long-term memory of paired associates. Nature, 354,152-155.
58. Kobatake, E., Wang, G. and Tanaka, K. (1998) Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. Journal of Neurophysiology, 80, 324-330.
59. Logothetis, N.K., Pauls, J., Bülthoff, H.H. and Poggio, T. (1994) View-

dependent object recognition by monkeys. Current Biology, 4, 401-414.

60. Foldiak, P. (1991) Learning invariance from transformation sequences. Neural Computation, 3, 194-200.
61. Wiskott, L. and Sejnowski, T.J. (2002) Slow feature analysis: unsupervised learning of invariances. Neural Computation, 14, 715-770.
62. Wyss, R., König, P. and Verschure, P.F. (2006) A model of the ventral visual system based on temporal stability and local memory. Plos Biology, 4, e120.
63. Masquelier, T. and Thorpe, S.J. (2007) Unsupervised learning of visual features through spike timing dependent plasticity. Plos Computational Biology, 4, e31.
64. Wallis, G. and Rolls, E.T. (1997) Invariant face and object recognition in the visual system. Progress in Neurobiology, 51, 167-194.
65. Wallis, G. and Bülthoff, H. (1999) Learning to recognize objects. Trends in Cognitive Science, 3, 22-31.
66. Palmeri, T.J. and Gauthier, I. (2004) Visual object understanding. Nature Reviews Neuroscience, 5, 291-303.
67. Connor, C.E., Brincat, S.L. and Pasupathy, A. (2007) Transformation of shape information in the ventral pathway. Current Opinion in Neurobiology, 17, 140-147.
68. DiCarlo, J.J., Zoccolan, D. and Rust, N.C. (2012) How does the brain solve visual object recognition? Neuron, 73, 415-434.
69. Riesenhuber, M. and Poggio, T. (2000) Models of object recognition. Nature Neuroscience, 3, Suppl: 1199-1204.
70. Kravitz, D.J., Saleem, K.S., Baker, C.I., Ungerleider, L.G. and Mishkin, M. (2013) The ventral visual pathway: an expanded neural framework for the processing of object quality. Trends in Cognitive Science, 17, 26–49.
71. Baker, C.I., Behrmann, M. and Olson, C.R. (2002) Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. Nature Neuroscience, 5, 1210-1216.
72. Sigala, N. and Logothetis, N.K. (2002) Visual categorization shapes feature selectivity in the primate temporal cortex. Nature, 415, 318-320.

73. De Baene, W., Ons, B., Wagemans, J. and Vogels, R. (2008) Effects of category learning on the stimulus selectivity of macaque inferior temporal neurons. Learning and Memory, 15, 717-727.
74. Li, N., Cox, D.D., Zoccolan, D. and DiCarlo, J.J. (2009) What response properties do individual neurons need to underlie position and clutter "invariant" object recognition? Journal of Neurophysiology, 102, 360-376.
75. Wallis, G. and Bülthoff, H. (2001) Effects of temporal association on recognition memory. Proceedings of the National Academy of Sciences of the United States of America, 98, 4800-4804.
76. Cox, D.D., Meier, P., Oertelt, N. and DiCarlo, J.J. (2005) 'Breaking' position-invariant object recognition. Nature Neuroscience, 8, 1145-1147.
77. Optican, L.M. and Richmond, B.J. (1987) Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. III. Information theoretic analysis. Journal of Neurophysiology, 57, 162–178.
78. Sripati, A.P. and Olson, C.R. (2009) Representing the forest before the trees: a global advantage effect in monkey inferotemporal cortex. Journal of Neuroscience, 29, 7788-7796.
79. Wang, R.H., Dai, L., Okamura, J., Fuchida, T. and Wang, G. (2021) Object discrimination performance and dynamics evaluated by inferotemporal cell population activity. IBRO Neuroscience Report, 10, 171-177.
80. Freedman, D.J., Riesenhuber, M., Poggio, T. and Miller, E.K. (2003) A comparison of primate prefrontal and inferior temporal cortices during visual categorization. The Journal of Neuroscience, 23, 5235-5246.
81. Freedman, D.J., Riesenhuber, M., Poggio, T. and Miller, E.K. (2006) Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. Cerebral Cortex, 16, 1631-1644.
82. Miyashita, Y. (1993) Inferior temporal cortex: where visual perception meets memory. Annual Review of Neuroscience, 16, 245-263.
83. Anderson, B., Mruczek, R.E., Kawasaki, K. and Sheinberg, D. (2008) Effects of familiarity on neural activity in monkey inferior temporal lobe. Cerebral Cortex,

18, 2540-2552.

84. Op de Beeck, H.P., Deutsch, J.A., Vanduffel, W., Kanwisher, N.G. and DiCarlo, J.J. (2008) A stable topography of selectivity for unfamiliar shape classes in monkey inferior temporal cortex. Cerebral Cortex, 18, 1676-1694.
85. Woloszyn, L. and Sheinberg, D.L. (2012) Effects of long-term visual experience on responses of distinct classes of single units in inferior temporal cortex. Neuron, 74, 193-205.
86. Ghazizadeh, A., Griggs, W., Leopold, D.A. and Hikosaka, O. (2018) Temporal-prefrontal cortical network for discrimination of valuable objects in long-term memory. Proceedings of the National Academy of Sciences of the United States of America, 115, E2135-E2144.
87. Tarr, M.J. and Bulthoff, H.H. (1995) Is human object recognition better described by geon structural descriptions or by multiple views? Comment on Biederman and Gerhardstein (1993). The Journal of Experimental Psychology: Human Perception and Performance, 21, 1494-1505.
88. Bar, M. and Ullman, S. (1996) Spatial context in recognition. Perception, 25, 343-352.
89. Edelman, J.A. and Keller, E.L. (1998) Dependence on target configuration of express saccade-related activity in the primate superior colliculus. Journal of Neurophysiology, 80, 1407-1426.
90. Hayward, W.G. (2003) After the viewpoint debate: where next in object recognition? Trends in Cognitive Science, 7, 425-427.
91. Kersten, D., Mamassian, P. and Yuille, A. (2004) Object perception as Bayesian inference. Annual Review of Psychology, 55, 271-304.
92. Pinto, Y., Olivers, C.N. and Theeuwes, J. (2008) The detection of temporally defined objects does not require focused attention. The Quarterly Journal of Experimental Psychology, 61, 1134-1142.
93. Poggio, T. and Ullman, S. (2013) Vision: are models of object recognition catching up with the brain? Annals of the New York Academy of Sciences, 1305, 72-82.

94. Aggleton, J.P. and Mishkin, M. (1983) Visual recognition impairment following medial thalamic lesions in monkeys. Neuropsychologia, 21, 189-197.
95. Booth, M.C. and Rolls, E.T. (1998) View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. Cerebral Cortex, 8, 510-523.
96. Freiwald, W.A. and Tsao, D.Y. (2010) Functional compartmentalization and viewpoint generalization within the macaque face-processing system. Science, 330, 845-851.
97. Yamane, Y., Carlson, E.T., Bowman, K.C., Wang, Z. and Connor, C.E. (2008) A neural code for three-dimensional object shape in macaque inferotemporal cortex. Nature Neuroscience, 11, 1352-1360.
98. Wang, G. and Suemitsu, K. (2007) Object recognition learning differentiates the representations of objects at the ERP component N1. Clinical Neurophysiology, 118, 372-380.
99. Wang, G. and Bingo, J. (2010) Three-dimensional object recognition learning alters an early ERP component of N1. Clinical Neurophysiology, 121, 1473-1480.
100. Tokudome, W. and Wang, G. (2012) Similarity dependency of the change in ERP component N1 accompanying with the object recognition learning. International Journal of Psychophysiology, 83, 102-109.
101. Crookes, K., Ewing, L., Gildenhuys, J.D., Kloth, N., Hayward, W.G., Oxner, M. and Rhodes, G. (2015) How well do computer-generated faces tap face expertise? PloS One, 10, 1–18.
102. Balas, B. and Pacella, J. (2015) Artificial faces are harder to remember. Computers in Human Behavior, 52, 331–337.
103. Balas, B. and Koldewyn, K. (2013) Early visual ERP sensitivity to the species and animacy of faces. Neuropsychologia, 51, 2876–2881.
104. Looser, C.E., Guntupalli, J.S. and Wheatley, T. (2013) Multivoxel patterns in face-sensitive temporal regions reveal an encoding schema based on detecting life in a face. Social Cognitive and Affective Neuroscience, 8, 799–805.

105. Balas, B., van Lamsweerde, A.E., Saville, A. and Schmidt, J. (2017) School-age children's neural sensitivity to horizontal orientation energy in faces. Developmental Psychobiology, 59, 899–909.

106. Wheatley, T., Weinberg, A., Looser, C., Moran, T. and Hajcak, G. (2011) Mind perception: real but not artificial faces sustain neural activity beyond the N170/VPP. PloS One, 6, 1–7.

107. Ganis, G., Smith, D. and Schendan, H.E. (2012) The N170, not the P1, indexes the earliest time for categorical perception of faces, regardless of interstimulus variance. Neuroimage, 62, 1563–1574.

108. Rossion, B., Gauthier, I., Tarr, M.J., Despland, P.A., Bruyer, R., Linotte, S. and Crommelinck, M. (2000) The N170 occipito-temporal component is enhanced and delayed to inverted faces but not to inverted objects: an electrophysiological account of face-specific processes in the human brain. NeuroReport, 11, 69-74.

109. Stahl, J., Wiese, H. and Schweinberger, S.R. (2008) Expertise and own-race bias in face processing: an event-related potential study. Neuroreport 19, 583–587.

110. Wiese, H. (2012) The role of age and ethnic group in face recognition memory: ERP evidence from a combined own-age and own-race bias study. Biological Psychology, 89, 137–147.

111. Wiese, H., Schweinberger, S.R. and Hansen, K. (2008) The age of the beholder: ERP evidence of an own-age bias in face memory. Neuropsychologia, 46, 2973–2985.

112. Maurer, D., Le Grand, R. and Mondloch, C.J. (2002) The many faces of configural processing. Trends in Cognitive Science, 6, 255–260.

113. Caharel, S., Montalan, B., Fromager, E., Bernard, C., Lalonde, R. and Mohamed, R. (2011) Other-race and inversion effects during the structural encoding stage of face processing in a race categorization task: an event-related brain potential study. International Journal of Psychophysiology, 79, 266–271.

114. Yin, R. (1969) Looking at upside down faces. Journal of Experimental

Psychology, 81, 141–145.

115. Kanwisher, N. and Yovel, G. (2006) The fusiform face area: a cortical region specialized for the perception of faces. Philosophical Transactions of the Royal Society B Biological Sciences, 361, 2109–2128.

116. Gauthier, I., Williams, P., Tarr, M.J. and Tanaka, J. (1998) Training 'greeble' experts: a framework for studying expert object recognition processes. Vision Research, 38, 2401–2428.

117. Gauthier, I., Behrmann, M. and Tarr, M.J. (1999) Can face recognition really be dissociated from object recognition? Journal of Cognitive Neuroscience, 11, 349–370.

118. Davies, G.M., Shepherd, J.W. and Ellis, H.D. (1979) Similarity effects in face recognition. The American Journal of Psychology, 92, 507-523.

119. Wang, G., Obama, S., Yamashita, W., Sugihara, T. and Tanaka, K. (2005) Prior experience of rotation is not required for recognizing objects seen from different angles. Nature Neuroscience, 8, 1768-1775.

120. Wang, R.H., Okamura, J. and Wang, G. (2022) Change in event-related potential accompanying view-invariant object discrimination learning. IEEJ Transactions on Electronics, Information and Systems, 142, 522-529.

121. Jeffreys, D.A. (1996) Evoked potential studies of face and object processing. Visual Cognition, 3, 1–38.

122. Rossion, B., Joyce, C., Cottrell, G. and Tarr, M. (2003) Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. NeuroImage, 20, 1609–1624.

123. Tanaka, J.W. and Curran, T. (2001) A neural basis for expert object recognition. Psychological Science, 12, 43–47.

124. Rossion, B., Curran, T. and Gauthier, I. (2002) A defense of the subordinate-level expertise account for the N170 component. Cognition, 85, 189–196.

125. Gauthier, I., Curran, T., Curby, K.M. and Collins, D. (2003) Perceptual interference supports a non-modular account of face processing. Nature Neuroscience, 6, 428–432.

126. Halgren, E., Raij, T., Marinkovic, K., Jousmaki, V. and Hari, R. (2000) Cognitive response profile of the human fusiform face area as determined by MEG. Cerebral Cortex, 10, 69–81.

127. Liu, J., Harris, A. and Kanwisher, N. (2002) Stages of processing in face perception: an MEG study. Nature Neuroscience, 5, 910–916.

128. Xu, Y., Liu, J. and Kanwisher, N. (2005) The M170 is selective for faces, not for expertise. Neuropsychologia, 43, 588–597.

129. Krigolson, O.E., Pierce, L.J., Holroyd, C.B. and Tanaka, J.W. (2009) Learning to become an expert: Reinforcement learning and the acquisition of perceptual expertise. Journal of Cognitive Neuroscience, 21, 1833-1840.