Visual Attention in The Fovea and The Periphery during Visual Search

Authors

Jie Zhang[1, 2, 3, 7], Xiaocang Zhu[1, 2, 7], Shanshan Wang[4], Zhengyu Ma[3], Hossein Esteky[5], Yonghong Tian[3], Robert Desimone[6] & Huihui Zhou[1, 3, *]

[1]Brain Cognition and Brain Disease Institute (BCBDI), Shenzhen-Hong Kong Institute of Brain Science-Shenzhen Fundamental Research Institutions, Guangdong Provincial Key Laboratory of Brain Connectome and Behavior, CAS Key Laboratory of Brain Connectome and Manipulation, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China.

[2]University of Chinese Academy of Sciences, Beijing 100049, China.

[3]Peng Cheng Laboratory, Shenzhen 518000, China.

[4]Paul C. Lauterbur Research Center for Biomedical Imaging, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China.

[5]Research Group for Brain and Cognitive Sciences, School of Medicine, Shahid Beheshti University, Tehran, Iran.

[6]McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

[7]Co-first author

[*]Correspondence: zhouhh@pcl.ac.cn

Summary

Visual search depends on both the foveal and peripheral visual system, yet the foveal attention mechanisms is still lack of insights. We simultaneously recorded the foveal and peripheral activities in V4, IT and LPFC, while monkeys performed a category-based visual search task. Feature attention enhanced responses of Face-selective, House-selective, and Non-selective foveal cells in visual cortex. While foveal attention effects appeared no matter the peripheral attention effects, paying attention to the foveal stimulus dissipated the peripheral feature attentional effects, and delayed the peripheral spatial attentional effects. When target features appeared both in the foveal and the peripheral, feature attention effects seemed to occur predominately in the foveal, which might not distribute across the visual field according to common view of distributed feature attention effects. As a result, the parallel attentive process seemed to occur during distractor fixations, while the serial process predominated during target fixations in visual search.

Keywords

Feature-based Attention, Foveal, Face, Visual Cortex, Prefrontal Cortex, Visual search,

Introduction

In naturalistic environment, our eyes generally move 2-3 times per second to gather information from different locations of the world. Visual search has been widely investigated to understand the neural mechanisms of visual attention during active vision. Feature attention enhances the visual responses of neurons to stimuli sharing features with the target in V4, IT, lateral intra-parietal cortex (LIP), and the prefrontal cortex (PFC) (Bichot et al., 2005, 2015, 2019; David et al., 2008; Mirpour et al., 2018; Motter, 2018; Sapountzis et al., 2018; Zhou and Desimone, 2011), and shifted the tuning of V4 neurons to more closely match the spectral properties of the target (Mazer and Gallant, 2003). The feature attention effects seem to occur throughout the visual field, independently of the locus of spatial attention (Bichot et al., 2005; Cohen and Maunsell, 2011; Maunsell and Treue, 2006; Mazer and Gallant, 2003; McAdams and Maunsell, 2000; Motter, 2018; Saenz et al., 2002; Sapountzis et al., 2018; Painter et al., 2014; Treue and Martinez-Trujillo,1999; Zhou and Desimone, 2011), and the amplitude of the attentional modulation is related to saccade behaviors during search (Motter, 2018; Sapountzis et al., 2018; Zhou and Desimone, 2011). Visual responses are also enhanced by spatial attention during free-gaze visual search, and the spatial and feature attentional processes seem to occur in parallel temporally in V4, PFC, LIP (Bichot et al., 2005, 2015; Motter, 2018; Sapountzis et al., 2018; Zhou and Desimone, 2011). The prefrontal cortex and the parietal cortex might modulate the responses in visual cortex during the attentional processes in visual search (Bichot et al., 2015, 2019; Zhou and Desimone, 2011). However, these findings are based on analysis of activities of neurons with the peripheral receptive field (RF), thus, reflecting the attentional mechanisms in the peripheral. There is still a lack of insights into the mechanisms of visual attention through the foveal visual system in visual search, although the primate visual system is designed to preferentially analyze the foveal stimuli in the visual field.

It is known that both the peripheral and foveal visual systems play important roles in visual search. In particular, masking the foveal visual field interferes severely with visual exploration behaviors, including decreases in search accuracy, increases in search time, elimination of the search facilitation in repeated displays (Bertera and Rayner, 2000; Cornelissen et al., 2005; McIlreavy et al., 2012; Murphy and Foley-Fisher, 1988), and these behavioral effects is comparable to, or larger than the effects caused by masking the peripheral field (Bertera and Rayner, 2000; Cornelissen et al., 2005). Currently, most studies involving central vision are focused on mechanisms of object recognition and categorization in high-level visual areas such as V4, IT (Bao et al., 2020; Bashivan et al., 2019; Chang and Tsao, 2017; Hong et al., 2016; Yamins et al., 2014), but not in the free-gaze visual search. Wang et al. (2018) show a target-selective enhanced response in human medial temporal lobe (MTL) and medial frontal cortex (MFC) during visual search, but this study doesn't map the RFs of these recorded neurons, leaving open the question on foveal attention mechanisms.

The co-existence of attentional modulation in different locations have been reported, including feature attention effects across the peripheral visual field, spatial and feature attention across multiple peripheral locations in the visual and prefrontal cortex (Bichot et al., 2005, 2015; Mirpour et al., 2018; Motter, 2018; Sapountzis et al., 2018; Zhou and Desimone 2011), while it is still an open question about parallel attentional modulation in the peripheral and the foveal. A number of studies suggests different relationship between the foveal and peripheral attentional processes. High level of attentional load in the foveal reduce the detection accuracy of stimuli in the peripheral (Macdonald and Lavie, 2008), and directing attention to peripheral regions reduced the EEG response to foveal stimuli (Lissa et al., 2020), and attending one feature in the central task cost the performance in the periphery (Morrone et al., 2002; VanRullen et al., 2004), while other studies show independent relationship between the foveal and peripheral processes (Ludwig et al., 2014; Morrone et al., 2002; Shen et al., 2003; VanRullen et al., 2004). Simultaneously recording of cells with foveal RF and peripheral RF allows us to obtain further insights into this relationship at neuronal level.

In this study, we recorded from foveal and peripheral cells simultaneously in area V4, IT cortex and lateral prefrontal cortex (LPFC), while monkeys performing a category-based visual search task. We found that foveal cells exhibit stronger Face or House selectivity than that of the peripheral cells. These Face-selective and House-selective cells showed stronger feature attentional enhancement to their preferred stimulus category, while the attentional effects on different level responses to stimuli within the same category were similar. Paying attention to the foveal stimulus dissipated the feature attentional effects in the peripheral, and delayed spatial attentional effects in V4, IT and LPFC peripheral cells. Thus, feature attentional enhancement in the peripheral and the foveal seemed not occur in parallel during visual search. This study extended our understanding of distribution of attention in active vision and feature attention toward complex stimulus features.

Results

In the category-based visual search task, the monkeys (Macaca mulatta) were free to find either one of the two targets in the search array and required to fixated on it for 800 ms (Figure 1A-B). The targets were indicted by an early-appeared cue stimulus that was always different from the targets. The targets and the cue stimulus belonged to the same category. Both monkeys performed well in this free-gaze visual search task, with 92% correct by monkey S and 86% correct by monkey E.

We recorded both single unit and multiunit activity in area V4, inferior temporal (IT), and LPFC simultaneously in two monkeys. Figure 1C shows representative MRI sections through V4, IT, and LPFC. The estimated V4, IT, and LPFC recording sites in the two monkeys are shown in Figure S1. We recorded 1898 foveal units, and 765 peripheral units with increased visual response in V4. These foveal units showed significantly increased responses to the cue stimulus presented in the foveal (Wilcoxon rank-sum test, $p < 0.05$), but not to the search array covering the peripheral field before the first saccade (Figure S2 A-C, E-G). These peripheral units responded to the search array, but not to the cue stimulus in the foveal (Figure S2 I-K, M-O, Q-S). The RFs of these peripheral units were further mapped using a visually guided saccade task. In IT, we recorded 1511 foveal units, and 239 peripheral units. In LPFC, we recorded 35 foveal units, and 507 peripheral units. Further analyses were based on these foveal units and peripheral units. The results were qualitatively similar in both monkeys and were therefore combined.

Feature attentional modulation in V4 and IT foveal cells during visual search

For the 1898 V4 foveal units, 266 of them were defined as Face-selective units, 304 as House-selective units, and 1051 as Non-selective units (see Methods). For the 1511 IT foveal units, there were 518 Face-selective units, 340 House-selective units, and 558 Non-selective units. Figure S2 D and H shows distributions of the selectivity indices of all foveal cells in V4 and IT.

We found that the responses of the foveal cells in both V4 and IT were modulated by feature attention. Figure 2 A-B shows normalized firing rates averaged across the populations of Face-selective foveal cells during "Face Target", "Face Distractor", "House Target", and "House Distractor" fixations (see methods) in IT and V4, respectively. The responses to the target Face stimuli in the foveal were significantly larger than responses to the same stimuli when the Face stimuli were distractors (Wilcoxon signed-rank test, $p < 0.05$). Feature attention also enhanced responses to the House stimuli in the House-selective foveal cells in IT and V4 (Figure 2 C-D), respectively. For stimuli in the non-preferred category, these feature attentional effects were weak (Figure 2 A-D). Thus, feature attention seemed to selectively enhance responses to the stimuli in the preferred category of these foveal selective cells in IT and V4. For Non-selective foveal cells in IT and V4, feature attention enhanced responses to the two categories of stimuli non-selectively (Wilcoxon signed-rank test, $p < 0.05$; Figure 2 E-F; Figure S3).

We calculated an Attention Indices to quantify the attention effects for each unit, which was the difference divided by the sum of the firing rates in the two attention conditions. Figure S3 A-B show distributions of the Attention Indices of all foveal Face-selective units in IT and V4 calculated from their responses to Face stimuli (IT mean Attention Indices 0.058, Wilcoxon signed-rank test, $p < 0.05$; V4 mean Attention Indices 0.023, Wilcoxon signed-rank test, $p < 0.05$). The attention

indices were also significantly larger than zero in foveal House-selective cells in IT and V4 calculated from their responses to House stimuli (Figure S3 C-D; IT mean Attention Indices 0.03, Wilcoxon signed-rank test, $p < 0.05$; V4 mean Attention Indices 0.03, Wilcoxon signed-rank test, $p < 0.05$). The Attention Indices calculated from responses to the preferred category was significantly larger than the index calculated from responses to the non-preferred category in the IT and V4 Face- and House- selective cells (Wilcoxon signed-rank test, $p < 0.05$). In addition, the attention indices calculated from their responses to both Face and House stimuli were significantly larger than zero in foveal Non-selective cells in IT and V4 (Figure S3 E-F; IT mean Attention Indices 0.026, Wilcoxon signed-rank test, $p < 0.05$; V4 mean Attention Indices 0.0075, Wilcoxon signed-rank test, $p < 0.05$).

Attentional modulation in the peripheral and temporal relationship between attentional modulation in the foveal and the peripheral.

In the 765 V4 peripheral units, there were 19 Face-selective units, 13 House-selective units, and 730 Non-selective units. In IT, the 239 peripheral units included 11 Face-selective units, 10 House-selective units, and 216 Non-selective units. In the 507 LPFC peripheral units, there were 15 Face-selective units, 23 House-selective units, and 466 Non-selective units. Figure S2 L, P, and T show distributions of the stimulus category selectivity indices of all peripheral cells in V4, IT and LPFC. Because the numbers of selective cells were very limited, we focused our analysis on these Non-selective cells in V4, IT and LPFC.

Feature attention enhanced visual responses of V4, IT and LPFC peripheral cells (Wilcoxon signed-rank test, $p < 0.05$; Figure 3 A-C), when monkeys were planning an eye movement to a stimulus out of the RF. For the spatial attention effects in the peripheral cells, we compared the responses to a stimulus in the RF when the animal was planning a saccade to that stimulus (Attention In) with responses to the same stimulus when the animal was planning a saccade out of the RF (Attention Out). Figure 3 D-F show that the Attention In response was significantly larger than the Attention Out response (Wilcoxon signed-rank test, $p < 0.05$) in V4, IT and LPFC, respectively.

We analyzed the temporal relationship of these attentional effects. For the foveal cells, the feature attentional modulation became significant (Wilcoxon signed-rank test, $p < 0.05$) at 148 ms and 170 ms after fixation onset in IT and V4 Face-selective cells, respectively. The attentional effect was significantly earlier in IT than that in V4 (two sided permutation test, $p < 0.05$). The attentional latencies were similar in House-selective cells (IT: 139 ms; V4: 140 ms) and Non-selective cells (IT:142 ms; V4: 137 ms). Figure 4A show the cumulative distribution of feature attention latencies of these foveal units. Overall, the latencies in V4 and IT foveal cells were similar, except for the late effect in V4 Face-selective cells. For the peripheral cells, the latencies of feature attention effects were 137 ms, 147 ms and 58 ms in V4, IT and LPFC, respectively. The attentional effect was significantly earlier in LPFC than in V4 and IT foveal and peripheral cells (Figure 4 B-C; two-sided permutation test, $p < 0.05$), consistent with previous findings (Zhou & Desimone 2011; Bichot et al., 2015). We further compared the time courses of attentional modulations between the peripheral and the foveal within the same area. Overall, the latencies of feature attention effects were similar in the two parts of visual field in V4 and IT (two-sided

permutation test, $p > 0.05$; Figure 4 B-C), except for the late effects in V4 foveal Face-selective cells.

Influence of foveal feature attention state on the peripheral feature attentional modulation

The temporal overlap of these attention effects suggested that feature attention effects might appear in parallel in the foveal and the peripheral. To test this predication, we analyzed the feature attention effects in the peripheral cells when the features of the stimulus in the foveal was either attended or not (illustrated in Figure 5). Feature attention enhanced responses in V4, IT and LPFC peripheral cells (Wilcoxon signed-rank test, $p < 0.05$; Figure 5 B, E, H) during fixation on a distractor. However, fixating on a target in the foveal seemed to dissipate this peripheral feature attentional enhancement (Wilcoxon signed-rank test, $p > 0.05$; Figure 5 C, F, I), suggesting that feature attentional enhancements appeared either in the foveal or in the peripheral, but not in both areas. In contrast to the peripheral attention modulation, the foveal attention effects recorded in this study occurred when there was always a peripheral target that could cause response enhancement in cells with RF covering the peripheral target, suggesting that the foveal feature attention process might dominate the peripheral attention process when target features appeared both in the foveal and the peripheral.

Influence of foveal feature attention state on the peripheral spatial attentional modulation

We investigated influence of foveal feature attention state on the peripheral spatial attentional process (illustrated in Figure 6). When the features of foveal stimulus were not attended, we observed significant spatial attention effects in peripheral V4, IT, and LPFC units (Wilcoxon signed-rank test, $p < 0.05$; Figure 6 B, E, H), while fixating on a target in the foveal seemed to reduce the spatial attention enhancements in peripheral V4 and LPFC units (Figure 6 C, I), and dissipate the spatial attention enhancements in IT (Wilcoxon signed-rank test, $p > 0.05$; Figure 6F). However, the peripheral spatial attentional effects (when the features of foveal stimuli were attended) became significant around saccade onset in IT (Wilcoxon signed-rank test, $p < 0.05$; Figure 7K, right), suggesting that foveal feature attention might mainly delay the spatial attention enhancements in the peripheral.

Feature and spatial attention distribution during search

There were four types of fixations throughout visual search before the last target fixation: the "D-D" fixation in which monkeys fixated on a distractor followed by a saccade to a distractor (Figure 7 1st row), the "D-T" fixation in which monkeys fixated on a distractor followed by a saccade to a target (Figure 7 2nd row), the "T-D" fixation in which monkeys fixated on a target followed by a saccade to a distractor (Figure 7 3rd row), and the "T-T" fixation in which monkeys fixated on a target followed by a saccade to another target (Figure 7 4th row). Feature attention enhanced the responses to all peripheral targets on the screen and spatial attention enhanced the response to the saccade target around saccade onset during D-D fixations in V4, IT and LPFC peripheral cells (Wilcoxon signed-rank test, $p < 0.05$; Figure 7 A-C), and during D-T fixations (Wilcoxon signed-rank test, $p < 0.05$; Figure 7 D-F). Feature attention did not enhance the response to the peripheral target during T-D fixations in those cells (Wilcoxon signed-rank test, $p > 0.05$; Figure 7 G-I left) and during T-T fixations (Wilcoxon signed-rank test, $p > 0.05$; Figure 7 J-L left). However, spatial attention enhanced the response to the saccade target in V4, IT and LPFC during

T-D fixations (Wilcoxon signed-rank test, $p < 0.05$; Figure 7 G-I right), and during T-T fixations (Wilcoxon signed-rank test, $p < 0.05$; Figure 7 J-L right). Thus, the distribution of attention in the peripheral whether fixation was on a target or a distractor in the foveal. There were parallel feature attention to target stimuli and spatial attention to a saccade target in the peripheral during distractor fixation. During target fixation, no feature but spatial attention to the saccade target in the peripheral, it seemed like a serial attention shift during visual search.

Dependence of feature attentional modulation on the stimulus category

Neural mechanisms of feature attention based on simple visual features such as colors, simple shapes, luminance, motion direction have been widely investigated, while study on attention to complex visual features such as high-level category features is still absent. We observed larger attentional modulation on responses to the stimuli of the preferred category that evoked higher visual responses in these foveal selective cells (Figure 2). To further clarify the role of category features in this attention process, we analyzed attentional modulations on different levels of responses to different stimuli in the same category and the modulations on responses to different categories. The stimuli in the same category were classified into 4 subsets based on their response amplitudes. Figure 8 A-H show the attention effects on responses from low to high to different subsets of House stimuli and Face stimuli in IT House-selective cells. The attentional effects across responses to different subsets seemed similar although the visual responses evoked by the four subsets were very different, suggesting the feature attentional modulation might depend on stimulus category, rather than on the response level alone. Figure 8I shows relationship between averaged attention effects during a time window of 150 to 225 ms after fixation onset and the averaged magnitudes of visual response (time window: 50-225 ms after fixation onset) to different subsets of stimuli in the IT House-selective cells. The visual response to the four different subsets of House stimuli were significant different when they were attended (one-way ANOVA, $F_{3, 335} = 20.33$, $P < 0.001$) and not attended (one-way ANOVA, $F_{3, 335} = 19.57$, $P < 0.001$), while the attention effects on responses to the four different subsets were similar (one-way ANOVA, $F_{3, 335} = 1.09$, $P > 0.05$). The attention effects on responses to different categories (House vs Face) were significantly different (one-way ANOVA, $F_{1, 1354} = 19.09$, $P < 0.001$). These tendencies also appeared in V4 House-selective cells (within category: one-way ANOVA, $F_{3, 297} = 0.33$, $P > 0.05$; Across categories: one-way ANOVA, $F_{1, 1202} = 10.73$, $P < 0.001$; Figure 8K, Figure S4 I-P), and V4 Face-selective cells (within category: one-way ANOVA, $F_{3, 262} = 0.46$, $P > 0.05$; Across categories one-way ANOVA, $F_{1, 1062} = 15.97$, $P < 0.001$; Figure 8L, Figure S4 Q-X). For the IT Face-selective cells, the attention effects within category and between categories were both significantly different (one-way ANOVA, $P < 0.05$; Figure 8J, Figure S4 A-H), but the effects across categories ($F_{1,1918} = 125.65$) were substantially larger than the effects within category ($F_{3,476} = 3.77$).

Interestingly, in V4 Face-selective cells, the attention effects for the subset of Face stimuli with the smallest visual response was still larger (Wilcoxon signed-rank test, $p < 0.05$) than the attention effects for the subset of House stimuli with the largest visual response, while the visual response to the Face subset was similar (Wilcoxon signed-rank test, $p > 0.05$) to the response to the House subset (Figure 8L).

For these Non-selective cells, the response magnitude and the attention effect between categories were similar. The Attention Indices in responses to the Face and House stimuli were not different (IT, Wilcoxon signed-rank test, $N = 558$, $p > 0.05$; V4, Wilcoxon signed-rank test, $N = 1051$, $p > 0.05$; Figure S3 G-J). Thus, feature attention enhanced responses to the two categories of stimuli non-selectively in these Non-selective foveal cells. Together, it seemed that the category dependence of feature attention modulation was determined by selectivity of foveal cells.

Discussion

We have developed a category-based visual search task and recorded both foveal and peripheral cells simultaneously to investigate the attentional process in V4, IT and LPFC. Foveal cells exhibited stronger Face- or House- selectivity than that of the peripheral cells in area V4 and IT cortex. These selective cells showed stronger feature attentional enhancement to their preferred stimulus category, while the attentional effects on different level responses to stimuli within the same category were similar. While the foveal attention effects occurred when there was always a peripheral target that could be attended, the peripheral feature attentional enhancement in V4, IT, and LPFC disappeared when the foveal stimulus features were attended. Paying attention to the foveal features also delayed spatial attentional effects to peripheral locations in these areas. Thus, when target features appeared both in the foveal and the peripheral, feature attention effects seemed to occur predominately in the foveal, but not to distribute across the visual field according to common view of distributed feature attention effects. This study also further clarified the distribution of feature attention and overt spatial attention throughout visual search.

Although foveal visual and attentional processing plays an important role in the visual search, understanding of its neural mechanisms is still very limited. In visual cortex, studies on the foveal visual system have been mostly focused on mechanisms of object recognition and categorization (Bao et al., 2020; Bashivan et al., 2019; Chang and Tsao, 2017; Hong et al., 2016; Yamins et al., 2014) in tasks other than the free-gaze visual search. We recorded foveal and peripheral cells simultaneously in this study. Consistent with preferentially analyzing the foveal stimuli in the visual field in the primate visual system, we found that foveal cells exhibited stronger Face- or House- selectivity than that of the peripheral cells. About 57% IT and 30% V4 foveal units were Face- or House-selective, but only 8% IT and 4% V4 peripheral units were selective, which would be helpful for the foveal system to perform visual analysis of different stimuli during visual search. Previous studies show that masking the foveal visual field results in decreases in search accuracy, increases in search time (Bertera and Rayner, 2000; Cornelissen et al., 2005; McIlreavy et al., 2012; Murphy and Foley-Fisher, 1988; Nuthmann, 2014; Shen et al., 2003). We found that feature attention enhanced responses of Face-selective, House-selective, and Non-selective foveal cells in V4 and IT, and the levels of response to the attended stimuli were related to saccadic search behaviors.

The foveal feature attention seemed to engage stimulus processing in the foveal at a cost of peripheral stimulus processing. Behaviorally, fixations on a target stimulus tended to be longer than fixations on a distractor in this study. Consistent with this, previous studies show that

attention toward a task in the foveal such as visual discrimination substantially degraded the overall visual search performance, including longer reaction time, more fixations, and longer fixation duration (Shen et al., 2003; Hooge & Erkelens, 1999). Further, we observed that the peripheral feature attentional enhancement disappeared when the features of foveal stimulus were attended, and the spatial attention effects in the peripheral were also delayed. Feature attention in the foveal seemed to delay search for next stimulus. The feature attentional enhancements in the peripheral appeared when features of foveal stimulus were not attended, which facilitated search for next stimulus in the peripheral. Thus, the efficient visual search might depend on coordination of the foveal and peripheral attentional processes.

Numerous studies in non-human primates suggest that feature attention enhances stimuli with attended features across the visual field in visual search (Bichot et al., 2005; Mazer and Gallant, 2003; Motter, 2018; Sapountzis et al., 2018; Zhou and Desimone, 2011) and other tasks (Cohen and Maunsell, 2011; Maunsell and Treue, 2006; McAdams and Maunsell, 2000; Treue and Martinez-Trujillo,1999). EEG and fMRI studies also show that relevant color or category was saliently represented in parallel across the visual field when that feature was attended during visual search (Bartsch et al., 2018; Painter et al., 2014; Peelen et al., 2009). These studies were based on the responses to stimuli in the peripheral without considering feature attention to the foveal stimuli. Our study extended the understanding of feature attention distribution by considering both the foveal and peripheral visual field. We found that feature attention enhancements appeared either in the foveal or in the peripheral in our visual search task, but not in both, which were different from the common idea that feature attention is deployed throughout the visual field. It seemed that the global feature attention effects suggested by previous studies are confined to the peripheral visual field. In addition, because the foveal attention effects in this study occurred when there was always a peripheral target that could cause response enhancement in cells with RF covering the peripheral target, and the feature attentional enhancements in the peripheral disappeared when feature of foveal stimulus were attended, suggesting that foveal feature attentional processing might dominate peripheral feature attentional processing in our task.

Bichot et al. (2005) show evidence for both the parallel and serial attention engagement during visual search, with the parallel feature attention processes in the visual field and serially overt spatial attention on one stimulus each time. Our study further clarified the attention distribution throughout visual search, showing the parallel feature attention on multiple stimuli and spatial attention on the saccade target in the peripheral during distractor fixations, and spatial attention to the saccade target in the peripheral during target fixations. Previous studies show evidences for serial shifts of covert attention across items in the search array (Buschman and Miller, 2009; Woodman and Luck, 2003). Guided search model suggests that serial selection is guided parallelly by multiple source of information including the “top-down” guidance such as target features. Studies show that different features (color, size, orientation) cause very different patterns of guidance. For example, color is very effective in guiding search (Friedman-Hill and Wolfe, 1995; Lindsey et al., 2010; Palmer et al., 2019), while orientation information is less effective or even harmful for efficient search (Hulleman, 2020; Hulleman, Lund, & Skarratt, 2019; Olds & Fockler, 2004). We used the naturalistic complex stimuli in our visual task, and these stimuli are more similar to objects we met in daily life than the simple stimuli, such as

combination of sample shapes and colors. Further study is needed to investigate whether the pattern of the attention deployment found in our study is applicable to visual search using simple stimuli, and the serially covert attention shift during search.

Feature attention based on simple features such as color, shape, motion, etc. (Bichot et al., 2005, 2019; Cohen and Maunsell, 2011; Maunsell and Treue, 2006; McAdams and Maunsell, 2000; Motter, 2018; Sapountzis et al., 2018; Treue and Martinez-Trujillo,1999; Zhou and Desimone, 2011) has been widely investigated. Feature attention enhancement of neuronal responses to naturalistic complex (object, Face, or natural photograph patches) has been reported in a few studies (Hayden and Gallant, 2009; Bichot et al., 2015), while the dependence of the attention effects on the stimulus category is still an open question. We found that Face-selective and House-selective cells showed stronger feature attentional enhancement to their preferred stimulus category, similar to feature attention effects based on simple feature in corresponding selective cells. Moreover, the attentional effects on different level responses to stimuli within the same category were similar, suggesting that the feature attentional modulation depended on stimulus category, rather than the response level alone. Similar to simple features, feature attention could also be based on the complex features of naturalistic stimuli.

Acknowledgments
Supported by Program of Ministry of Science and Technology of China (PCL2021A13), National Natural Science Foundation of China grant (31671108 and 62027804), International Partnership Program of Chinese Academy of Sciences (No.172644KYSB20160175), and Shenzhen Basic Research Program (JCYJ20200109114805984).

Declaration of interests
The authors declare no competing interests.

Figure 1

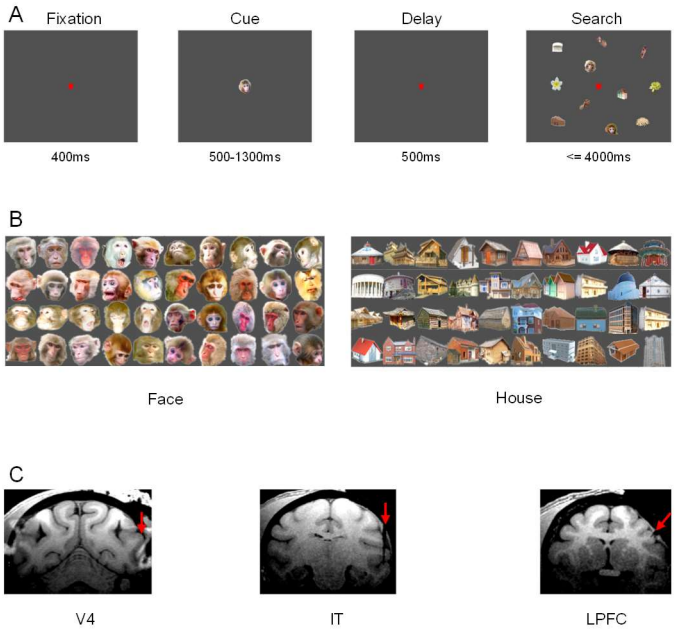

Figure 1. Task and recording sites.
(A) Illustration of the behavioral tasks. A central cue was presented first to indicate the category of the searched-for target, then a search array with eleven stimuli including two target stimuli and nine distractors appeared on the screen. The cue and the two targets belonged to the same category, but the targets were always different from the cue. Monkeys were rewarded for fixating on either one of the targets for ⩾800 ms.
(B) Stimuli of Face category and House category.
(C) MRI images showing the typical recording regions of V4, IT and LPFC. Red arrows indicated the electrode trace directions.

Figure 2

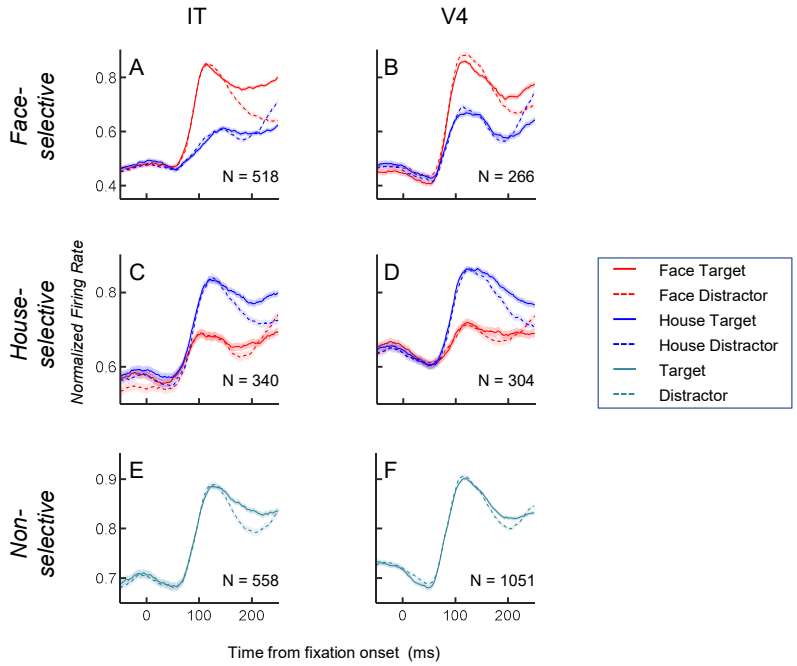

Figure 2. Foveal feature attentional modulation in V4 and IT.
(A) Normalized firing rates averaged across the IT foveal Face-selective cells during Face Target, Face Distractor, House Target, and House Distractor fixations (see methods). All firing rates were normalized to the maximum rates of the attended responses of preferred category. Shading around average firing rates indicates the SEM (±). (B)-(F) show the normalized population responses in V4 Face-selective cells, IT House-selective cells, V4 House-selective cells, IT Non-selective cells and V4 Non-selective cells, respectively. For non-selective cells, Face Target and House Target fixations were combined into "Target" fixations, and Face Distractor and House Distractor fixations were combined into "Distractor" fixations. Firing rates were normalized to the maximum rates of the "Target" responses.

Figure 3

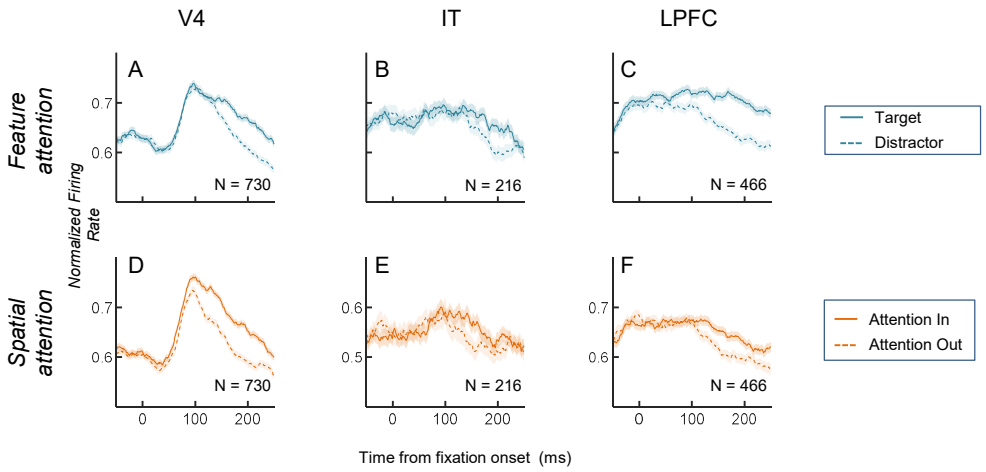

Figure 3. Peripheral feature and spatial attentional modulation in V4, IT and LPFC. (A)-(C) show population responses of peripheral Non-selective cells to target stimuli ("Target") and to matched distractor stimuli ("Distractor") in V4, IT and LPFC, respectively. (D)-(F) show population responses of these cells to stimuli followed by saccades into their RF (Attention In) and out of their RF (Attention Out) in V4, IT and LPFC, respectively.

Figure 4

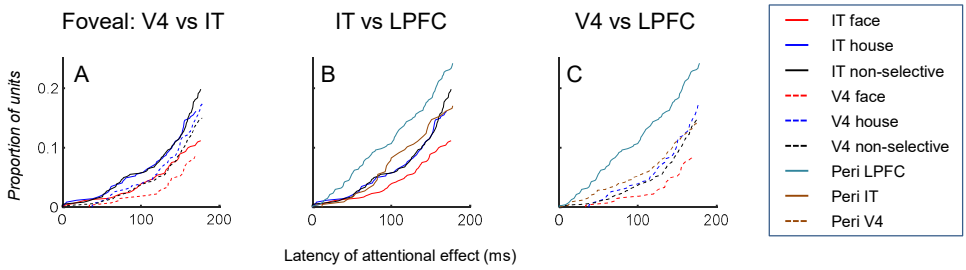

Figure 4. Temporal relationship of attentional modulation in the foveal and the peripheral. (A) Cumulative distribution of feature attention effect latencies in V4 and IT, computed from individual foveal Face-, House- and Non-selective units and represented as proportions of the total units. (B)-(C) show cumulative distributions of feature attention effect latencies of the foveal and peripheral units in IT and V4, respectively.

Figure 5

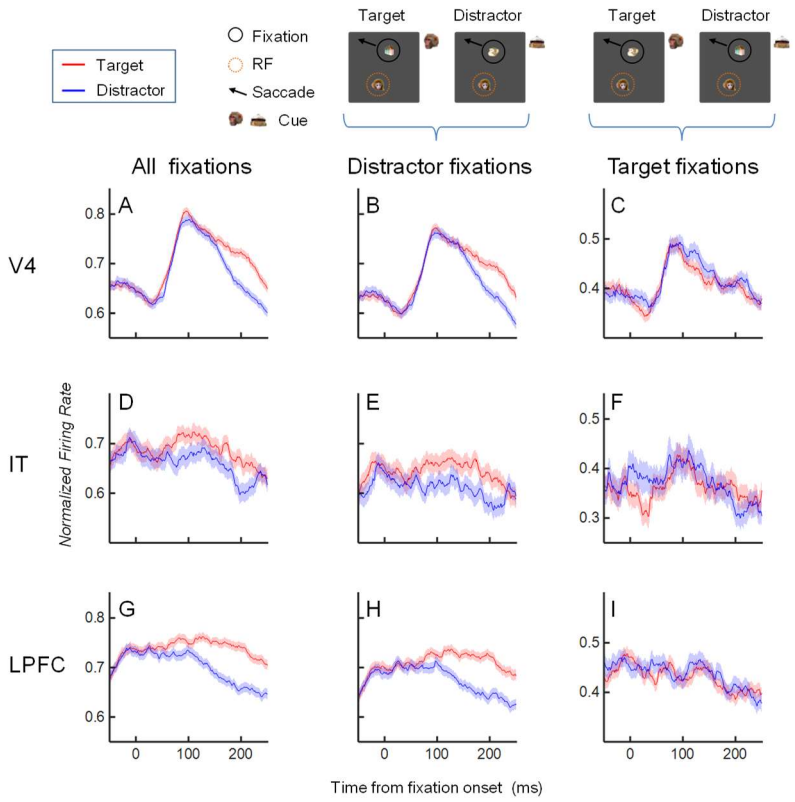

Figure 5. Influence of foveal feature attention state on the peripheral feature attentional modulation.

(A), (D), (G) show population responses of peripheral Non-selective cells to target stimuli ("Target") and to matched distractor stimuli ("Distractor") without considering central fixations in V4 (N = 396), IT (N = 123), and LPFC (N = 350), respectively. (B), (E), (H) show responses of these cells to target stimuli and to matched distractor stimuli during distractor fixations. The Target and Distractor conditions during distractor fixations are illustrated above (B). (C), (F), (I) show responses of these cells to target stimuli and to matched distractor stimuli during target fixations. The Target and Distractor conditions during target fixations are illustrated above (C).

Figure 6

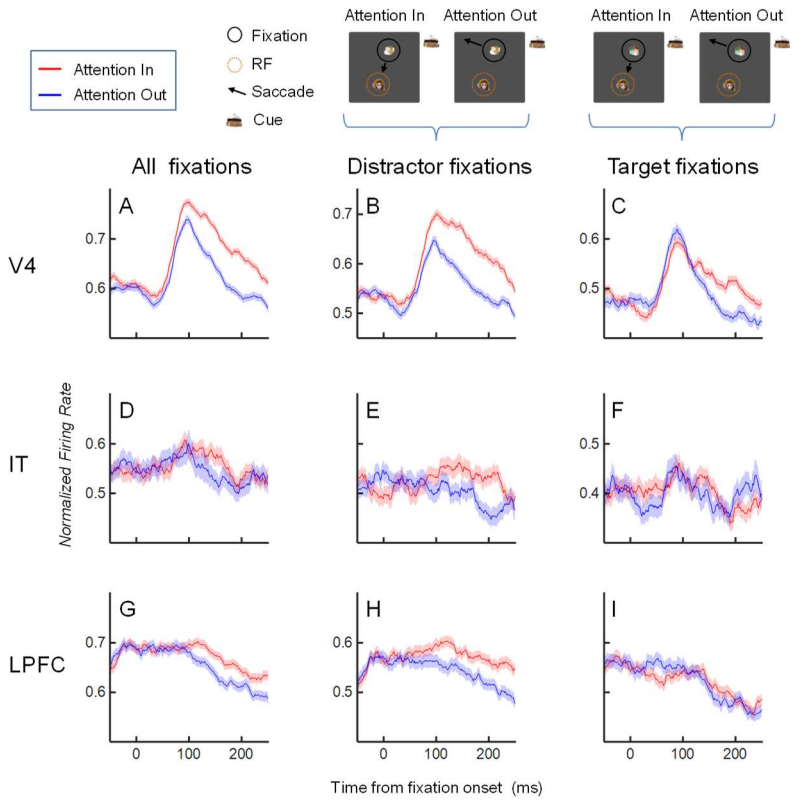

Figure 6. Influence of foveal feature attention state on the peripheral spatial attentional modulation.

(A), (D), (G) show spatial attention effects of V4 peripheral Non-selective cells (N = 651), IT peripheral Non-selective cells (N = 191), and LPFC peripheral Non-selective cells (N = 425), respectively. (B), (E), (H) show spatial attention effects of these cells when the features of foveal stimulus were not attended. The Attention In and Attention Out conditions during distractor fixations are illustrated above (B). (C), (F), (I) show spatial attention effects of these cells when the features of foveal stimulus were attended. The Attention In and Attention Out conditions during target fixations are illustrated above (C).

Figure 7

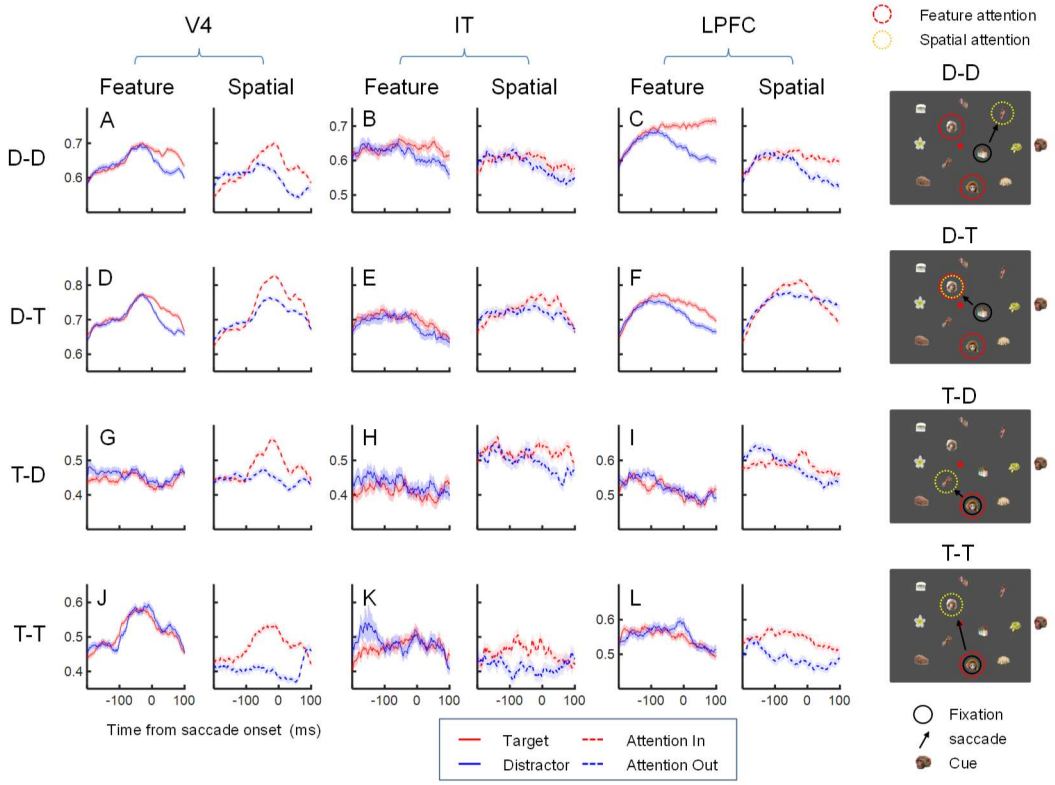

Figure 7. Feature and spatial attention distribution during search
(A)-(C) show feature and spatial attention effects of Non-selective units during D-D fixations in V4 (N = 727), IT (N = 216), and LPFC (N = 464), respectively. Illustration of feature and spatial attention distribution during D-D fixations is on the right of (C). (D)-(F) show feature and spatial attention effects of these units during D-T fixations in V4 (N = 730), IT (N = 216), and LPFC (N = 465), respectively. Illustration of feature and spatial attention distributions during D-T fixations is on the right of (F). (G)-(I) show feature and spatial attention effects of these units during T-D fixations in V4 (N = 600), IT (N = 212), and LPFC (N = 423), respectively. Illustration of feature and spatial attention distributions during T-D fixations is on the right of (I). (J)-(L) show feature and spatial attention effects of these units during T-T fixations in V4 (N = 676), IT (N = 200), and LPFC (N = 436), respectively. Illustration of feature and spatial attention distributions during T-T fixations is on the right of (L).

Figure 8

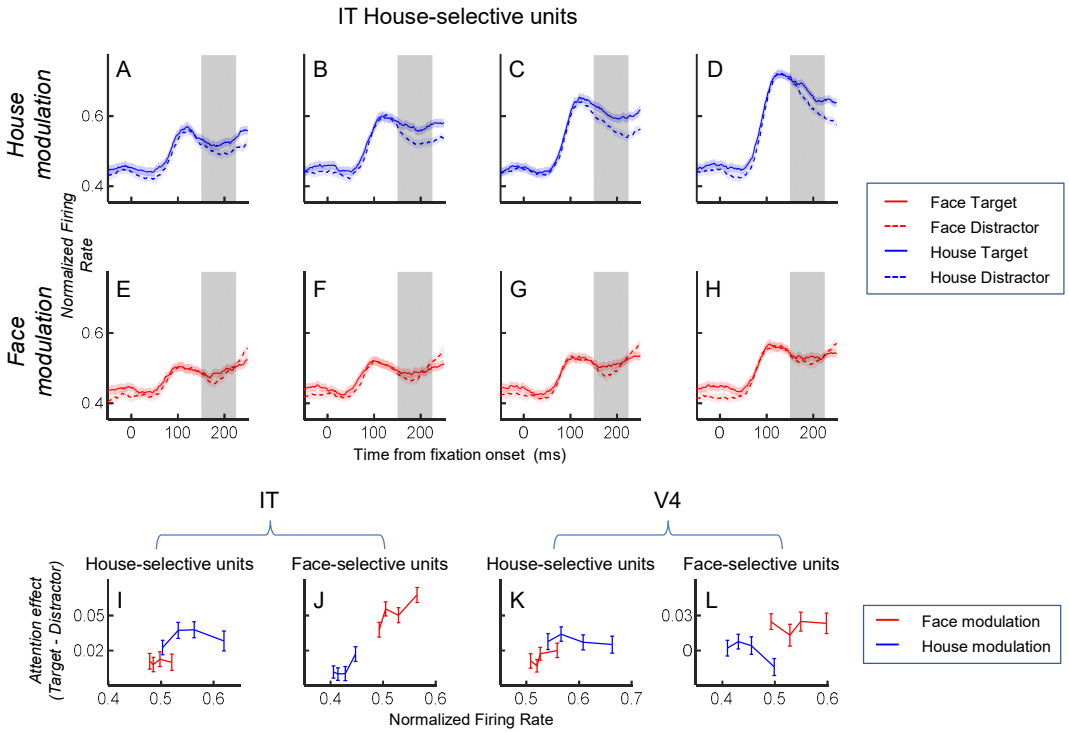

Figure 8. Influence of stimulus category on the attentional modulation.
(A)-(D) show the feature attentional effects on responses from low to high to 4 subsets of House stimuli in IT House-selective cells (N = 339), respectively. The rectangle shading indicates the time window (150 - 225 ms after fixation onset) used for analyzing the attentional effects. (E)-(H) show the feature attentional effects on responses from low to high to 4 subsets of Face stimuli in IT House-selective cells (N = 339), respectively. (I) The attention effects on responses to subsets of Face stimuli and House stimuli in IT House-selective cells. X axis: the amplitude of normalized visual response to subsets of House and Face stimuli; Y axis: the amplitude of attentional effects (Attended – Unattended) on response to these stimulus subsets. Visual response was calculated in a window of 50 - 225 ms after fixation onset, while the subsets of stimuli were not attended. (J)-(L) show the attention effects on responses to stimulus subsets in IT foveal Face-selective cells (N = 480), V4 foveal House-selective cells (N = 301), and V4 foveal Face-selective cells (N = 266), respectively.

Methods

General Procedures

Two male rhesus monkeys weighing 12 and 15 kg were used. Monkeys were implanted under aseptic conditions with a post to fix the head and recording chambers over areas V4, IT and LPFC. Localization of the chambers was based on MRI scans obtained before surgery. The behavioral experiments were under the control of a computer using MonkeyLogic software (University of Chicago, IL; Asaad et al., 2013), which presented the stimuli, monitored eye movements, and triggered the delivery of the reward. All animal procedures were approved by the Animal Care and Use Committees of Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences (No. SIAT-IRB-160223-NS-ZHH-A0187-003).

Behavioral tasks

Monkeys were trained to perform a free-gaze visual search task. After 400 ms fixation on a center spot, a cue stimulus replaced the center spot and presented in the central of screen. The cue stimulus was replaced by the center fixation spot after a 500 - 1300 ms random period. Following another 500 ms fixation on the center spot, a search array with 11 items including two target stimuli was presented in 11 randomly selected locations from a total of 20 pre-defined locations. The two target stimuli and the cue stimulus belonged to the same category, although the target stimuli were always different from the cue stimulus. The cue stimulus was selected randomly from the House or Face stimuli with equal probability. The other 9 stimuli in the search array belonged to the other three categories. The stimuli consisted of 160 natural object images including 4 categories (Face 40; House 40; Flower 40; Hand 40), subtended an area of approximately 2 × 2 degree. The aspect ratio, luminance, hue and saturation in HSV color space of these images were matched across categories. Monkeys were required to find either one of two targets within 4000 ms and keep fixation on the target for 800 ms to receive a juice reward. No constraints were placed on their search behavior to allow animals to conduct the search naturally. Before the search array onset, monkeys were required to keep central fixation. The 20 locations, covering the visual field of eccentricities from 5 to 11 degree, included 18 locations located symmetrically in the left and right visual field with 9 in each side, and 2 locations on the vertical middle line.

A visually guided saccade task was used to map the peripheral receptive fields (RFs) of recorded cells. After central fixation for 400 ms, one stimulus appeared randomly in 1 of the 20 locations, monkeys were required to make a saccade to the stimulus within 500 ms and fixate on it for 300 ms to get a reward.

Neural Recording

Single unit and multi-unit spikes were recorded from V4, IT and LPFC through 24 or 32-contact electrodes (V-Probe or S-Probe, Plexon Inc, Dallas, USA) in a 128 channel Cerebus System (Blackrock Microsystems, Salt Lake City, UT, USA). In most sessions, we recorded activities in

two of the above three areas simultaneously. Neural signals were filtered between 250 Hz and 5 kHz, amplified and digitized at 30 kHz to obtain spike data. The location of recordings in V4, IT and LPFC was verified with MRI. Eye movements were recorded by an infrared eye tracking system (iViewX Hi-Speed, SensoMotoric Instruments (SMI), Teltow, Germany) at a sampling rate of 500 Hz.

Data Analysis

Firing rate analysis.

Measurements of neural activity were derived from spike density functions generated by convolving the time of action potentials with a function that projects activity forward in time (Growth = 1 ms, Decay = 20 ms) and approximates an EPSP (Thompson et al., 1996).

Receptive field analysis.

Visual response to the Cue stimulus and to the search array in free-gaze visual search task, which were detected by comparing the firing rates during a post-stimulus period (50 to 200 ms after stimulus onset) with the baseline firing rates during a pre-stimulus period (-150 to 0 ms before stimulus onset) using a Wilcoxon rank-sum test, were first used to separate cells with foveal RFs and cells with peripheral RFs. Foveal cells were defined as these that only responded to the Cue stimulus in the foveal (Wilcoxon rank-sum test, $P < 0.05$), but not to the search array appeared in the peripheral visual field (Wilcoxon rank-sum test, $P > 0.05$). Peripheral cells were defined as these that responded only to the search array (Wilcoxon rank-sum test, $P < 0.05$), but not to the Cue stimulus (Wilcoxon rank-sum test, $P > 0.05$). There were also other cells that respond significantly to both the Cue stimulus and the search array, which were not further investigated. The RFs and stimulus selectivity of these peripheral cells were further mapped based on their activities in the visually guided saccade task.

Category Selectivity analysis.

We determined the selectivity of cells based on a selectivity index similar to the index used in previous studies on IT (Freiwald et al., 2009; Freiwald and Tsao, 2010). For foveal cells, the responses to Face stimuli (R_{Face}) or House stimuli (R_{House}) were determined by subtracting baseline activity during -150 to 0 ms from the Cue stimulus onset from the firing rates during 50 to 200 ms after the onset in visual search task. For peripheral cells, the responses were determined by subtracting baseline activity during -150 to 0 ms from the peripheral stimulus onset from the firing rates during 50 to 200 ms after the onset in the visually guided saccade task. The selectivity index (SI) was defined as $(R_{Face}-R_{House}) / (R_{Face}+R_{House})$. SI was set to 1 when $R_{Face} > 0$ and $R_{House} < 0$, and to -1 when $R_{Face} < 0$ and $R_{House} > 0$. For Face-selective cells, their R_{Face} was at least 130% of their R_{House}, that is, their SIs were larger than 0.13, and R_{Face} were significantly higher than R_{House} (Wilcoxon rank-sum, $p < 0.05$). Similarly, in these House-selective cells, their R_{House} were at least 130% of their R_{Face}, and R_{House} were significantly higher than R_{Face}. Cells were defined as Non-selective cells if their R_{Face} and R_{House} were similar (Wilcoxon rank-sum, $p > 0.05$). The remaining cells that did not fit into any above types of cells were classified as Un-defined cells.

Attention effect analysis.

To investigate the feature attention in foveal cells, we compared responses to a target stimulus in the foveal and responses to the same stimulus in the foveal when it was a distractor, while monkey was preparing a saccade away from the stimulus. For Face- or House- selective cells, fixations during search period (before the last fixation on the target at the end of search) were sorted into four types: "Face Target", "Face Distractor", "House Target", and "House Distractor". In the Face Target fixations, the stimulus in the foveal was a Face stimulus and the monkey was searching for a Face target. In the Face Distractor fixations, the stimulus in the foveal was a Face stimulus and the monkey was searching for a House target. In the House Target fixations, the stimulus in the foveal was a House stimulus and the monkey was searching for a House target. In the House Distractor fixations, the stimulus in the foveal was House stimulus and the monkey was searching for a Face target. For Non-selective cells, Face Target and House Target fixations were combined into "Target" fixations, and Face Distractor and House Distractor fixations were combined into "Distractor" fixations. The stimulus in the foveal was matched across the attended and unattended conditions. Neural activities in V4 and IT during these fixations were calculated and compared to show the feature attention effects.

For feature attention effects in peripheral cells, we sorted fixations during the search period according to a similar approach in our pervious study (Zhou and Desimone, 2011). In the "Target" fixations, one target stimulus was in the cell's RF. In the "Distractor" fixations, the same stimulus was in the same location of the cell's RF, but now it was a distractor. Only fixations followed by a saccade away from the RF were included for this analysis. For spatial attention effects in the peripheral cells, we compared responses in "Attention In" and "Attention Out" fixations, which followed by saccades to one stimulus in the RF and out of the RF of a cell, respectively. The saccade target stimulus in RF during Attention In fixations were matched with a stimulus in the same location in RF during Attention Out fixations.

The latency of attention effects at population average level was determined based on averaged responses of each cells using a sliding window method. If a significant difference (Wilcoxon signed-rank test, $p < 0.05$) was found successively for 35 ms between the "Target" and "Distractor" responses or between the "Attention In" and "Attention Out" responses, the first time point of the 35 ms window was defined as the starting point of attentional modulation. To test whether a latency difference at the population level was significant, we ran a two-sided permutation test with 1000 repeats as described in our previous study (Zhou and Desimone, 2011). The Attention Indices to quantify the magnitude of attention effects was defined as the difference divided by the sum of the firing rates in the two attention conditions based on the averaged firing rates in a time window of 150 - 225 ms after fixation onset.

References

Asaad WF, Santhanam N, McClellan S, Freedman DJ (2013). High-performance execution of psychophysical tasks with complex visual stimuli in MATLAB. J. Neurophysiol. 109, 249-260.

Bao P, She L, McGill M, and Tsao DY (2020). A map of object space in primate inferotemporal cortex. Nature. 583, 103-108.

Bartsch MV, Donohue SE, Strumpf H, Schoenfeld MA, Hopf JM. Enhanced spatial focusing increases feature-based selection in unattended locations (2018). Sci Rep. 8(1):16132.

Bashivan P, Kar K, and DiCarlo JJ (2019). Neural population control via deep image synthesis. Science. 364, eaav9436.

Bertera JH, Rayner K (2000). Eye movements and the span of the effective stimulus in visual search. Percept Psychophys. 62, 576-85.

Bichot NP, Rossi AF, and Desimone R (2005). Parallel and serial neural mechanisms for visual search in macaque area V4. Science. 308, 529-534.

Bichot NP, Heard MT, DeGennaro EM, and Desimone R (2015). A source for feature-based attention in the prefrontal cortex. Neuron. 88, 832-844.

Bichot NP, Xu R, Ghadooshahy A, Williams ML, Desimone R (2019). The role of prefrontal cortex in the control of feature attention in area V4. Nat Commun. 16;10(1):5727.

Buschman TJ, Miller EK (2009). Serial, covert shifts of attention during visual search are reflected by the frontal eye fields and correlated with population oscillations. Neuron. 63(3):386-96.

Chang L, and Tsao DY (2017). The Code for Facial Identity in the Primate Brain. Cell. 169, 1013-1028.

Cornelissen FW, Bruin KJ, and Kooijman AC (2005). The Influence of artificial scotomas on eye movements during visual search. Optom Vis Sci. 82, 27-35.

Cohen MR, and Maunsell JH (2011). Using neuronal populations to study the mechanisms underlying spatial and feature attention. Neuron. 23;70(6):1192-204.

David SV, Hayden BY, Mazer JA, and Gallant JL (2008). Attention to stimulus features shifts spectral tuning of V4 neurons during natural vision. Neuron. 59, 509-521.

Freiwald WA, and Tsao DY (2010). Functional compartmentalization and viewpoint generalization within the macaque Face-processing system. Science 330, 845-851.

Freiwald WA, Tsao DY, and Livingstone MS (2009). A Face feature space in the macaque temporal lobe. Nature neuroscience 12, 1187-1196.

Friedman-Hill S, Wolfe JM (1995). Second-order parallel processing: visual search for the odd item in a subset. J Exp Psychol Hum Percept Perform. 21(3):531-51.

Hayden BY, Gallant JL (2009). Combined effects of spatial and feature-based attention on responses of V4 neurons. Vision Res. 49(10):1182-7.

Hong H, Yamins DLK, Majaj NL, and DiCarlo JJ (2016). Explicit information for category-orthogonal object properties increases along the ventral stream. Nat Neurosci. 19, 613-22.

Hooge IT, Erkelens CJ (1999). Peripheral vision and oculomotor control during visual search. Vision Res. 39(8):1567-75.

Hulleman J (2020). Quantitative and qualitative differences in the top down guiding attributes of visual search. J. Exp. Psychol: Human Perception and Performance, on-line.

Hulleman J, Lund K, Skarratt PA (2019). Medium vs. difficult visual search: how a quantitative change in the functional visual field leads to a qualitative difference in performance. Atten Percept Psychophys, on-line first, 1-22.

Lissa PD, Caldara R, Nicholls V, and Miellet S (2020). In pursuit of visual attention: SSVEP frequency-tagging moving targets. PLoS One 15, e0236967.

Lindsey DT, Brown AM, Reijnen E, Rich AN, Kuzmova Y, and Wolfe JM (2010). Color Channels, not Color Appearance or Color Categories, Guide Visual Search for Desaturated Color Targets. Psychol Sci, 21(9), 1208-1214.

Ludwig CJH, Davies JR, and Eckstein MP (2014). Foveal analysis and peripheral selection during active visual sampling. Proc Natl Acad Sci U S A 111, E291-9.

Macdonald JSP, and Lavie N (2008). Load induced blindness. Journal of Experimental Psychology: Human Perception and Performance 34, 1078-1091.

Maunsell JH, Treue S (2006). Feature-based attention in visual cortex. Trends Neurosci. 29(6):317-22.

Mazer JA, and Gallant JL (2003). Goal-related activity in V4 during free viewing visual search: evidence for a ventral stream visual salience map. Neuron. 40, 1241-1250.

McAdams CJ, Maunsell JH (2000). Attention to both space and feature modulates neuronal responses in macaque area V4. J Neurophysiol. 83(3):1751-5.

McIlreavy L, Fiser J, and Bex PJ (2012). Impact of simulated central scotomas on visual search in natural scenes. Optom Vis Sci. 89, 1385-1394.

Mirpour K, Bolandnazar Z, and Bisley JW (2018). Suppression of frontal eye field neuronal responses with maintained fixation. Proc Natl Acad Sci U S A 115, 804-809.

Morrone MC, Denti V, and Spinelli D (2002). Color and luminance contrasts attract independent attention. Curr Biol. 12, 1134-1137.

Motter MC (2018). Saccadic momentum and attentive control in V4 neurons during visual search. Journal of Vision 18, 1-21.

Murphy KSJ, and Foley-Fisher JA (1988). Visual search with non-foveal vision. Ophthal. Physiol. Opt. 8, 345-8.

Nuthmann A (2014). How do the regions of the visual field contribute to object search in real-world scenes? Evidence from eye movements. J Exp Psychol Hum Percept Perform. 40(1):342-60.

Olds ES, Fockler KA (2004). Does previewing one stimulus feature help conjunction search? Perception, 33(2), 195-216.

Painter DR, Dux PE, Travis SL, Mattingley JB (2014). Neural responses to target features outside a search array are enhanced during conjunction but not unique-feature search. J Neurosci. 34(9):3390-401.

Palmer EM, Van Wert MJ, Horowitz TS, Wolfe JM (2019). Measuring the time course of selection during visual search. Atten Percept Psychophys. 81(1):47-60.

Peelen MV, Fei-Fei L, Kastner S (2009). Neural mechanisms of rapid natural scene categorization in human visual cortex. Nature. 460(7251):94-7.

Saenz M, Buracas GT, Boynton GM (2002). Global effects of feature-based attention in human visual cortex. Nat Neurosci. 5(7):631-2.

Sapountzis P, Paneri S, and Gregoriou GG (2018). Distinct roles of prefrontal and parietal areas in the encoding of attentional priority. Proc Natl Acad Sci U S A 115, E8755-E8764.

Shen J, Reingold EM, Pomplun M, and Williams DE (2003). Saccadic selectivity during visual search: The influence of central processing difficulty. In The mind's eye: Cognitive and applied aspects of eye movement research, J. Hyona, R. Radach, H. Deubel, ed. (Elsevier), pp. 65-88.

Thompson KG, Hanes DP, Bichot NP, and Schall JD (1996). Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. Journal

of Neurophysiology. 76, 4040-4055.

Treue S, Martinez Trujillo JC (1999). Feature-based attention influences motion processing gain in macaque visual cortex. Nature. 399(6736):575-9.

VanRullen R, Reddy L, and Koch C (2004). Visual search and dual tasks reveal two distinct attentional resources. J Cogn Neurosci. 16, 4-14.

Wang S, Mamelak AN, Adolphs R, and Rutishauser U (2018). Encoding of target detection during visual search by single neurons in the human brain. Curr Biol. 28, 2058-2069.e4.

Woodman GF, Luck SJ (2003). Serial deployment of attention during visual search. J Exp Psychol Hum Percept Perform. 29(1):121-38.

Yamins DLK, Hong H, Cadieu CF, Solomon EA, Seibert D, and DiCarlo JJ (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. Proc Natl Acad Sci U S A 111, 8619–8624.

Zhou H, and Desimone R (2011). Feature-based attention in the frontal eye field and area V4 during visual search. Neuron 70, 1205-1217.