Memory formation of object representation : natural scenes

Eiichi Hoshino[12], Fumihiko Taya[2], and Ken Mogi[12]

[1] Tokyo Institute of Technology, eiichi@brn.dis.titech.ac.jp
[2] Sony Computer Science Laboratories

Abstract. It has been suggested that the parahippocampal region collects spatial information in relation to navigational objects, in a joint encoding of space (Janzen and van Turennout (2004)). The navigational object may play a role of "landmark" when the episodically dispersed snapshots are combined into comprehensive spatial information of an individual space. Here we study the nature of object recognition in the long-term memory in human cognition that is learnt during scene-integration as regards the viewing angle under attention control. Based on the result, it is suggested that objects in a scene without attention may be processed into a 2-D representation bound to the background scene as a texture.

1 Introduction

The interpretation of spatial relationships is an important factor in contextual behaviours and the formation of episodic memory. The classic key idea of spatial information processing has been analysed by O'Keefe and Nadel [1]. Together with discovery of "place cells" and concept of "cognitive map" [2], they proposed the cognitive map theory in which place cells, dead reckoning system and landmark navigation are combined into allocentric map-based navigation in hippocampus formation. On the other hand, Yamaguchi et al. have proposed a mechanism of hippocampal memory encoding of episodic events from novel temporal inputs caused by theta phase precession [3]. More recently, Janzen and van Turennout demonstrated that the parahippocampal region collects spatial information in relation to various objects, in a joint encoding of space and objects [4]. To produce allocentric long-term memory (LTM), it is necessary that the egocentric representations primarily obtained from perceptual information are combined together. In a case of absence of self motion, the production of allocentric LTM crucially relies on the integration of different scenes. If scenes are episodically dispersed, it is required that there be navigational landmarks in the scenes to combine these scenes into coherent allocentric representations.

Current study focuses on the observation of object recognition underlying human cognition after episodically dispersed views are combined into comprehensive spatial information of an individual space. Additionally, attention enhances the visual LTM (VLTM) of previously attended objects embedded in a natural scene [5], which is supported by a dynamic evolution model on attention and memory [6], suggesting that object representations in LTM may also be affected by attention. Thus, we also study the effect of attention on the object representations revealed in a later recognition task.

2 Experiments

To investigate the nature of object representations within human cognition in LTM that is learnt during scene-integration, experiment 1 was designed especially as regards the viewing angle with attention control, in the context of integration of the spatial information. In the leaning phase, participants were, in turn, viewed two dispersed views in which the several objects were located. They were instructed to remember objects on green bases and their position while those on blue bases were distractors at the moment. After the learning phase, they were required to conduct a two-alternative forced-choice recognition test. The objects presented in the test phase were divided into three types; i.e. objects viewed from same angle as the learning phase, those from different angles and novel objects. The objects were chosen from everyday use objects in same basic-level category (i.e. chairs, in this study) to make the different visual angle recognition easy because the aim of the experiment is to study the nature of object representations in LTM, but not in short-term memory (STM) in which the multi-angle object representations would be more easily established than those in LTM. Consequently, experiment 2 was carried out as the same as experiment 1 but the background in learning phase were changed to monochromatic.

2.1 Results

The rate of correct judgement (RCJ) (familiar or new) was significantly higher for the attended objects (green bases) compared to the unattended objects (blue bases). A 2 by 2 repeated-measures analysis of variance (ANOVA) (attention by visual angle) on the RCJ showed a significant effect of attention ($F(1,8) = 7.977, P = 0.022 < 0.05$), whereas no significant effect of visual angles and no interaction between same and different visual angles ($F(1,8) = 0.397, P = 0.546$) (Fig. 1 (a)). The RCJ on unattended objects viewed from the same visual angle was higher than that for rotated objects, although the difference was not statistically significant (Fig. 1 (a)). The RCJ on unattended objects viewed from different visual angle was significantly lower than that on attended objects, ($p < .05$), while that on attended objects had no significance. The RCJ on objects located at a certain spatial configuration, such as right- or left turn corner and crossroads, was not consistent in conditions with any particular visual angles. A few participants reported that remembering the objects and their location was too difficult and could not confidentially discriminate familiar or new for the most of the objects. However the tendency in less RCJ on the unattended objects presented from different visual angle was observed among participants. A motivated participant showed that good performances in both types of RCJ for attended objects, but again his RCJ for unattended objects viewed from different angles was significantly poor.

In experiment 2, no statistically significant difference was observed among the 4 conditions, although both types of RCJ of objects from different visual angles appeared to be higher than that of objects from same visual angles (Fig. 1(b)).

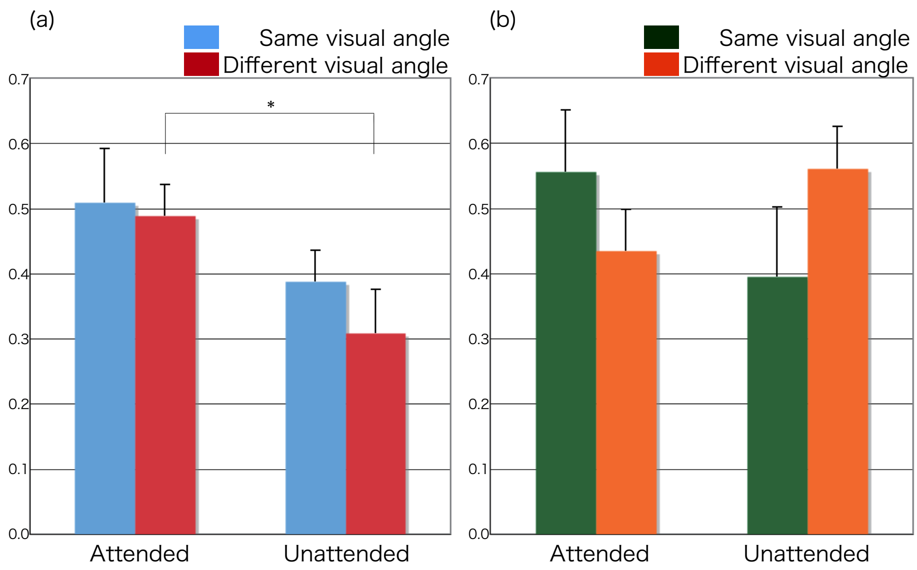

Fig. 1. Rate of correct judgement (RCJ) for the objects in experiment 1 and 2. (a) RCJ for attended and unattended objects viewed from same and different visual angles in experiment 1. (b)That in experiment 2. Error bars are standard errors across participants.

3 Discussion

The result that overall VLTM of attended objects was relatively well established compared with that of unattended objects is consistent with previous work by Hollingworth [5], supporting that attention to objects in a scene enhanced the consolidation of memory for navigational landmark representations. Moreover, the accuracy on attended object recognitions from different visual angles indicates the role of attention is producing the angle invariant object representations.

The inconsistency between the succeeded recognition for object from particular visual angles and the spatial configurations suggests that the memory of object representations do not include the spatial configurations around the object. (But see Hollingworth 2006 [8]). This supports Mallot and Gillner's finding that the local views and objects are recognised individually and not recognised as configurations among objects when navigating in a large-scale environment [9]. The selective activation for navigational objects previously showed in a scene but not for non-navigational objects, found by Janzen and van Turennout, in the parahippocampal gyrus [4] may not support encoding of spatial configurations around the objects. The result that the RCJ vary according to attention could predict that the nature of object representations would reflect only object intrinsic representations, again consistent with Mallot and Gillner's finding. The VLTM of object representations is strongly bound to a presented scene [8]. If attention plays a role in extracting focal information, a cognitive stage of object representations may be produced by attention. Therefore, attention may support to obtain the view-invariant or 3-D representations of objects from a scene, but in absence of attention, such representations are never obtained. Rather the object-to-scene binding gives rise to a perception of objects as a texture in the scene. In this view, the failure of recognition in unattended and depth rotation may arise from the object-to-scene binding.

Whereas unattended objects in a scene might be primarily treated as a texture bound to the scene in the VLTM without any 3-D structural information. The failure to recognise rotated objects whose orientation is novel for subjects may reflect the strong object-to-scene binding where an object is regarded as a part of plain surface of a scene, resulting in inhibition priming.

References

1. O'Keefe, J., Nadel, L.: The Hippocampus as a Cognitive Map. Oxford, Clarendon (1978)
2. O'Keefe, J., Dostrovsky, J.: The Hippocampus as a Spatial Map. Preliminary Evidence from Unit Activity in the Freely-Moving Rat. Brain Res Vol. 34. (1971) 171-175
3. Yamaguchi, Yoko.: A Theory of Hippocampal Memory Based on Theta Phase Precession. Biological Cybernetics. 89 (2003) 1-9
4. Janzen, G., Turennout, van, M.: Selective Neural Representation of Objects Relevant for Navigation. Nature Neuroscience. Jun. Vol. 7. No. 6. (2004) 673-7
5. Hollingworth, A., Henderson, J. M.: Accurate Visual Memory for Previously Attended Objects in Natural Scenes. Journal of Experimental Psychology. Vol. 28. No. 1. (2002) 113-136
6. Wang, R., Yu, J., Zhang, Z.: A Neural Model on Cognitive Process. Lecture Notes in Computer Science. ISNN 2006. LNCS3971. (2006) pp50-58
7. Hollingworth, A.: The Relationship Between Online Visual Representation of a Scene and Long-Term Scene Memory. Journal of Experimental Psychology. Vol. 31. No. 3. (2005) 396-411
8. Hollingworth, A.: Scene and Position Specificity in Visual Memory for Objects. Journal of Experimental Psychology. Vol. 32. No. 1. (2006) 58-69
9. Mallot, H, A., Gillner, S, Route., Navigating without Place Recognition: What Is Recognised in Recognition-Triggered Responses?
10. Hoshino, E., Taya, F., Mogi, K., Egocentric Space and Object Perception. Posted on : SfN2006 (Society for Neuroscience). 18th. Oct. (2006)