A model of visual working memory in PFC

Frank van der Velde, Marc de Kamps

Cognitive Psychology Unit, Leiden University
Leiden, The Netherlands

Abstract

We present a model of visual working memory in ventral prefrontal cortex. Activation in ventral PFC consists of reverberating activity. Representations in ventral PFC are conjunctions of location and (partial) identity representations. With many objects, representations in ventral PFC interfere, which results in loss of information. Therefore, the number of objects in working memory is limited. However, because ventral PFC is connected to the 'identity' levels in the visual cortex, the number of features for each object is unlimited. The 'blackboard' architecture of ventral PFC results in a unification (binding) of the feature representations of the objects maintained in memory.

Keywords: Blackboard architecture; Binding problem; Prefrontal cortex; Working memory

1. Introduction

Recent investigations [1] have shown that humans have the ability to maintain a number of visual objects in visual working memory. A remarkable characteristic of this finding is that the number of objects that can be maintained in working memory without interference (i.e., loss of information) is limited (to about four), but the number of object features (e.g., color, form, location) is unlimited for each of the objects. Functional neuroimaging studies [2] have shown that in particular the ventral prefrontal cortex (V-PFC) is involved in the maintenance of information in visual working memory. In contrast with animal studies [3], these neuroimaging studies did not reveal a distinction between spatial information and object (feature) information within PFC. However, a distinction between V-PFC and dorsal PFC (D-PFC) was found. Whereas V-PFC was mainly involved in maintenance of information, D-PFC appeared to be involved in operations that manipulate the information maintained in visual working memory.

In line with these observations, we present a model of visual working memory in PFC. The model is based on a neural 'blackboard' architecture that we used in our simulation of object-based attention in the visual cortex [4], illustrated in figure 1. The visual cortex consists of the ventral parthway, which includes the areas V_2, V_4, the posterior infero-temporal cortex (PIT) and the anterior infero-temporal cortex (AIT), and the dorsal pathway, which leads into the parietal cortex (PG). The ventral pathway is involved in the processing and selection of 'object features' (e.g., shape and color) and the dorsal pathway is involved in the processing and selection of spatial information (e.g., positions of objects in a display) and spatial transformations (e.g., for making eye movements). Both pathways start from the primary visual cortex (V_1), but they are also interconnected on the levels of V_2, V_4 and PIT. Both pathways project to the prefrontal cortex. In the ventral pathway, objects are identified through a feedforward network of areas, going from the primary visual cortex (V_1) to the higher areas in the temporal cortex (AIT). In this network, the retinotopic representation in the lower areas is gradually transformed (in V_2 to PIT) into a location-invariant identity (e.g., shape, color) representation (in AIT).

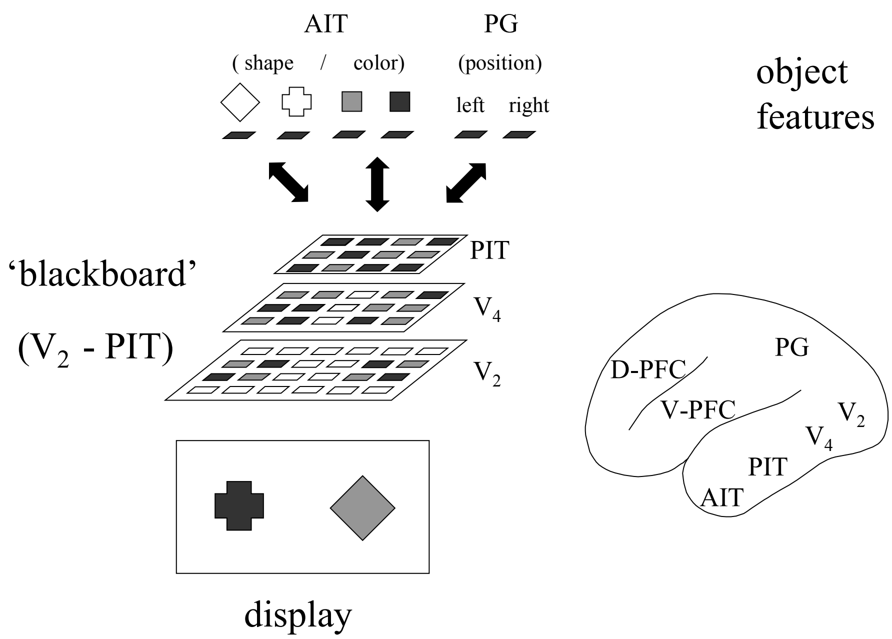

Fig. 1. Blackboard architecture of visual cortex. PIT = posterior infero-temporal cortex; AIT = anterior infero-temporal cortex; PG = parietal cortex; V-PFC = ventral prefrontal cortex; D-PFC = dorsal prefrontal cortex. For explanation see text.

In object-based attention, an object is selected as a target, and a feedback process is initiated in the networks that produce object identification. This feedback process carries information about the identity of the object to the lower areas in the visual cortex (V_2 - PIT). The feedback process interacts in these areas with the feedforward process described above, by means of microcircuits that connect the feedforward and feedback networks [4]. This interaction results in the enhanced activation of target representations in these areas, which results in the selection of target-related location information [4]. The representations in these lower 'in-between' areas consist of conjunctions of (partial) identity information and location information. This is a direct consequence of the fact that these areas gradually transform the retinotopic information in the primary visual cortex into identity-based information. By means of the interaction process described above, information about the location of an identified object can be recovered, even though that information was lost at the level of object identification. In a converse manner, the selection of a location in the 'in-between' areas can be used to select the information about the identity of an object, which will result in the identification of the object on the given location. In computational terms, the 'in-between' areas provide a ‘blackboard’ architecture because they link different ‘processors’ to one another [5]. The ‘processors’ in this case are networks for feature identification (e.g., shape, color, location, motion, etc.). The blackboard serves to unify (bind) the information processed in each of the specialized processors.

2. Blackboard architecture of working memory in PFC

Here, we propose a similar 'blackboard' architecture for the maintenance of information in visual working memory (figure 2).

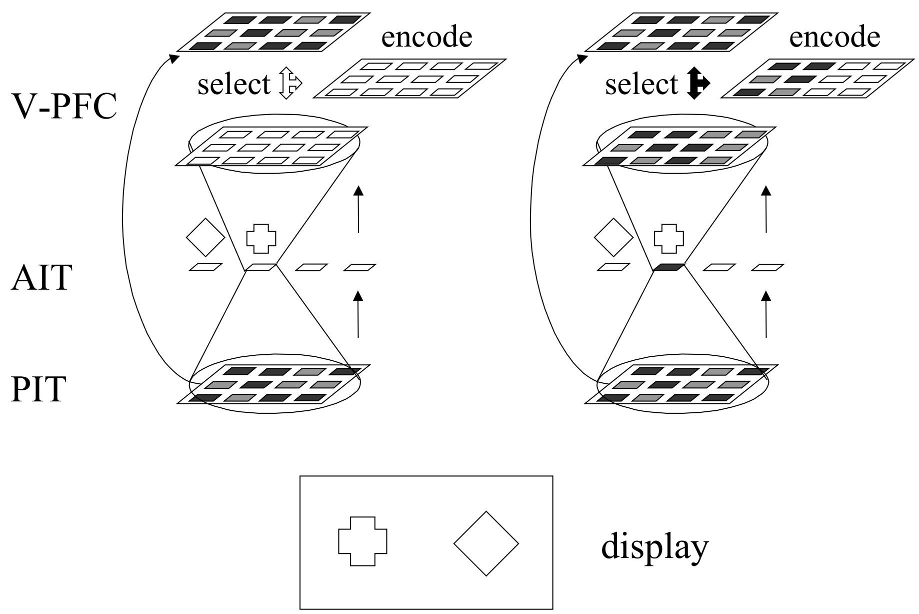

Fig. 2. A blackboard architecture in prefrontal cortex (PFC). PIT = posterior infero-temporal cortex; AIT = anterior infero-temporal cortex; V-PFC = ventral prefrontal cortex. For explanation see text.

In this model, V-PFC has a layered structure with representations similar to the representations in the visual (temporal) cortex (figure 1). First, PIT connects to one of the layers in V-PFC (for the purpose of illustration: the top layer in figure 2). As in PIT itself, the representations in this layer of V-PFC consist of conjunctions of location and (partial) identity (object-feature) representations. In turn, another layer of V-PFC (the bottom layer in figure 2) is connected to the higher-level areas in the visual cortex, in which (location-invariant) object identity and location information is processed and represented (in figure 2 illustrated for AIT). These connections are similar to the connections of the feedback network of the visual cortex [4]. Thus, they have the same 'fan-out' structure, in contrast with the 'fan-in' structure of the connections in the feedforward network of the visual cortex (as illustrated with the cones in figure 2). As a result, the representations in the bottom layer in V-PFC consist of distributed 'identity' representations, in line with the representations in the feedback network of the visual cortex [4]. The bottom and top layer in V-PFC interact in a manner similar to the interaction between the feedforward and feedback networks of the visual cortex, using similar microcircuits [4]. This interaction results in the activation of a third layer in V-PFC (the 'encode' layer in figure 2). The representations in this layer are similar to the representations in the top layer. In this model, the encode layer is the layer in which the sustained (reverberating) activity occurs, typical of working memory in the cortex [3].

Figure 2 illustrates that a similar selection process occurs in the V-PFC model as in the blackboard model of the visual cortex. In figure 2 (left), two objects are processed in the visual cortex, and their PIT representations also activate the representations in the top layer in V-PFC. However, none of the objects activates its identity representation in AIT (e.g., for shape in figure 2). As a result, the interaction between the top and bottom layer in V-PFC does not produce activation in the encode layer. In figure 2 (right), one of the objects (the cross) does activate its identity representation in AIT, which activates the bottom layer in V-PFC. As a result, the interaction between the top and bottom layer activate the representations in the encode layer that are selective for the features (e.g., shape, color, position) of the cross. Thus, objects have to activate their identity representations in AIT (i.e., have to be 'attended') before they can be stored in working memory.

3. Feature binding in working memory

The nature of the representations in ventral PFC and the connections with the higher-level areas in the visual cortex produces the behavioral effects described above. When too many objects are present in a display, their representations in ventral PFC interfere, which results in loss of information. This interference results in a limitation of the number of objects that can be maintained in working memory. However, because the ventral PFC is connected to the 'identity' levels in the visual cortex, the number of features for each object is unlimited. The blackboard architecture of ventral PFC results in a unification (binding) of the feature representations of the objects maintained in memory, similar to the blackboard architecture in the visual cortex. In this way, all features of an object can be maintained in working memory as long as the representations of the object in ventral PFC do not interfere.

The binding of feature representations in V-PFC is illustrated in figure 3. In a display of two objects, a blue cross and a yellow diamond, the features are cross, diamond, blue, yellow, left and right. To maintain both objects in (working) memory, a correct binding between these features is needed (e.g., a blue cross instead of a yellow cross, a diamond on the right instead of a diamond on the left). The hippocampus and surrounding areas, referred to as the 'hippocampus complex' (HC) in figure 3, form an episodic memory of the feature units that are concurrently active in the visual cortex. However, representations in HC are conjunctive representations [6], as indicated with the 'cone' in figure 3. This entails that they can represent a collection of features, but not their relations (binding). For instance, on the basis of the representation illustrated in figure 3 (left), one cannot determine whether the display consisted of a blue cross and yellow diamond or of a yellow cross and blue diamond.

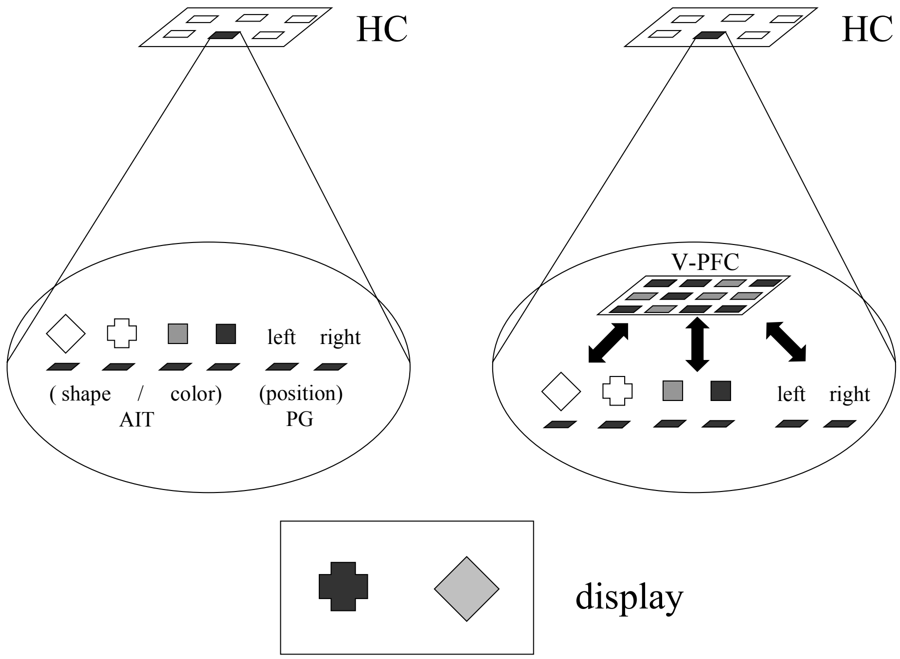

Fig. 3. Feature binding in working memory. AIT = anterior infero-temporal cortex; PG = parietal cortex; V-PFC = ventral prefrontal cortex; HC = hippocampus complex. For explanation see text.

Binding (or relations) between features can be represented with a compositional form of representation, which will result from embedding the individual feature representations in a computational architecture such as a 'blackboard' architecture. The relations between the features can then be retrieved from the interactions within the 'blackboard' [5], as in case of the visual cortex illustrated in figure 1 [4,7]. In the model illustrated in figure 3, the encode layer in V-PFC subserves the role of a 'blackboard' in visual working memory, with representations comparable to those in PIT (or perhaps V_4). Relations between features in working memory are retrieved by the interactions between V-PFC and the individual feature representations, as illustrated in figure 3 (right). Figure 3 (right) also illustrates that the HC can form a conjunctive ('cone') representation that preserves relations (binding) between features if it includes the V-PFC. This HC representation can be the basis for an episodic memory in which relations between objects are preserved.

4. Event coding in working memory

The (temporal) manipulation of information in working memory will depend on D-PFC in de model presented here. Briefly stated, V-PFC represents (maintains) the feature information in working memory (as discussed and illustrated above), whereas D-PFC is used to represent the temporal structure of information in working memory. HC now forms a conjunctive representation of V-PFC and D-PFC. In this way, two displays, for instance, can be stored in sequence in working memory. The order of the sequence is represented in D-PFC and the content of each display is represented in the blackboard representation of V-PFC (similar to figure 3, right). D-PFC will determine which of the two display representations is active in V-PFC. In this way, an object in one display can be used as the target in a visual search task, without an interference with the objects in the other display stored in working memory [8].

References

[1] E. K. Vogel, G. F. Woodman, S. J. Luck, Storage of features, conjunctions, and objects in visual working memory, Journal of Exp. Psychol.: HPP 27 (2001) 92-114.
[2] M. D'Esposito, Functional neuroimaging of working memory, in R. Cabeza and A. Kingstone eds., Handbook of functional neuroimaging of cognition, (MIT Press, Cambridge, MA, 2001) 293-327.
[3] F. A. W. Wilson, S. P. O. Scalaidhe, P. S. Goldman Rakic, Dissociation of object and spatial processing domains in primate prefrontal cortex, Science 260 (1993) 1955-1957.
[4] F. van der Velde, M. de Kamps, From knowing what to knowing where: Modeling object-based attention with feedback disinhibition of activation, J. Cognitive Neurosci. 13 (4) (2001) 479-491.
[5] F. van der Velde, On the use of computation in modelling behavior, Network: Computation in Neural Systems 8 (1997) 1-32.
[6] R. C. O'Reilly, Y. Munakata, Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain (MIT Press, Cambridge, MA, 2000).
[7] M. de Kamps, F. van der Velde, Using a recurrent network to bind form, color and position into a unified percept, Neurocomputing 38-40 (2001) 523-528.
[8] G. F. Woodman, E. K. Vogel, S. J. Luck, Visual search remains efficient when visual working memory is full, Psychological Science 12 (2001) 219-224.

Corresponding author: Frank van der Velde, Cognitive Psychology Unit, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands. Email addresses: vdvelde@fsw.leidenuniv.nl, kamps@fsw.leidenuniv.nl.