

Article

# Visuospatial Working Memory for Autonomous UAVs: A Bio-Inspired Computational Model

José-Antonio Cervantes <sup>1</sup>, Sonia López <sup>1,\*</sup>, Salvador Cervantes <sup>1</sup>, Adriana Mexicano <sup>2</sup>  
and Jonathan-Hernando Rosales <sup>3</sup>

<sup>1</sup> Department of Computer Science and Engineering, Universidad de Guadalajara, Ameca 46600, Mexico; antonio.alvarez@academicos.udg.mx (J.-A.C.); salvador.cervantes7964@academicos.udg.mx (S.C.)

<sup>2</sup> Division of Graduate Studies and Research, Instituto Tecnológico de Ciudad Victoria, Ciudad Victoria 87010, Mexico; adriana.ms@cdvictoria.tecnm.mx

<sup>3</sup> Department of Computer Science, Universidad Autónoma de Guadalajara, Zapopan 45129, Mexico; jonathan.rosales@edu.uag.mx

\* Correspondence: sonia.lopez@academicos.udg.mx

**Abstract:** Visuospatial working memory is a fundamental cognitive capability of human beings needed for exploring the visual environment. This cognitive function is responsible for creating visuospatial maps, which are useful for maintaining a coherent and continuous representation of visual and spatial relationships among objects present in the external world. A bio-inspired computational model of Visuospatial Working Memory (VSWM) is proposed in this paper to endow Autonomous Unmanned Aerial Vehicles (UAVs) with this cognitive function. The VSWM model was implemented on a low-cost commercial drone. A total of 30 test cases were designed and executed. These test cases were grouped into three scenarios: (i) environments with static and dynamic vehicles, (ii) environments with people, and (iii) environments with people and vehicles. The visuospatial ability of the VSWM model was measured in terms of the ability to classify and locate objects in the environment. The VSWM model was capable of maintaining a coherent and continuous representation of visual and spatial relationships among interest objects presented in the environment even when a visual stimulus is lost because of a total occlusion. The VSWM model proposed in this paper represents a step towards autonomous UAVs capable of forming visuospatial mental imagery in realistic environments.

**Keywords:** cognitive functions; visuospatial working memory; bio-inspired computational models; computer vision; autonomous unmanned aerial vehicles



**Citation:** Cervantes, J.-A.; López, S.; Cervantes, S.; Mexicano, A.; Rosales, J.-H. Visuospatial Working Memory for Autonomous UAVs: A Bio-Inspired Computational Model. *Appl. Sci.* **2021**, *11*, 6619. <https://doi.org/10.3390/app11146619>

Academic Editor: Jozsef Katona

Received: 6 May 2021

Accepted: 1 July 2021

Published: 19 July 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Unmanned Aerial Vehicles (UAVs), also known as drones, are becoming more and more popular in the research community because they have been widely studied in order to know their potential usage in different areas such as entertainment [1], marketing [2], healthcare [3], agriculture [4], and security [5]. The Artificial Intelligence (AI) research community has made a great effort to develop “intelligent” UAVs capable of doing tasks in an autonomous or semi-autonomous way [6]. Creating an “intelligent” UAV capable of navigating and interacting autonomously in the real world involves facing formidable challenges because they need to be able to process multiple tasks simultaneously, such as constantly sensing the environment; identifying and classifying both static and dynamic obstacles and targets; generating an internal representation of the real world, where the spatial relationship information between the environment’s objects must be constantly updated; reasoning and making right decisions to react appropriately when unexpected events appear. New interdisciplinary fields such as cognitive computing [7] and cognitive infocommunication [8] aim to create various bio-inspired engineering applications such as brain–computer interfaces [9] and computational systems capable of mimicking the

intelligence of living beings (e.g., insects, rodents, primates, and humans) [10–14]. The computing research community focused on developing bio-inspired systems considers that nature offers a rich source of inspiration for developing “intelligent” systems. Undoubtedly, one of the most challenging bio-inspired systems is the system that tries to mimic human intelligence. This paper presents a bio-inspired computational model of Visuospatial Working Memory (VSWM) to endow autonomous UAVs with this cognitive function. The VSWM is a fundamental cognitive component of human beings for exploring the visual environment. The VSWM is a memory buffer that allows observers to retain visuospatial information for a short period of time when visual stimuli are no longer in view as a consequence of object occlusion and saccadic eye movements [15]. The VSWM is formed by two types of Working Memory (WM): Visual Working Memory (VWM) and Spatial Working Memory (SWM) [16]. Whereas the VWM is a short-term memory involved in the active representation of the visual appearance of relevant objects [15,16], the SWM is a short-term memory involved in the active representation of the locations of relevant objects [16,17]. Therefore, the VSWM integrates sensory information coming from both the VWM and the SWM to generate visuospatial maps, also known as visuospatial mental imagery [16,18,19]. These maps offer a coherent and continuous representation of visual and spatial relationships among objects present in the external world, which are then further processed for the generation of specific actions such as spatial shifting, visual search, and target selection.

The bio-inspired VSWM computational model proposed in this paper is not a unique model. Other bio-inspired models have been proposed and reported in the literature. In [10], a bio-inspired model known as RatSLAM was described. This model has been inspired by current knowledge of how rodents’ brains work. RatSLAM has been implemented on ground robots and tested in both indoor and outdoor environments. In [20], a computational model for visual processing and object recognition was proposed. This computational model has been based on the current knowledge of the human brain areas related to the visual object recognition cognitive function. However, this computational model was focused only on one part of the human visual system, which is involved in the visual object recognition process, without considering the human brain areas involved in the spatial information. Additionally, this computational model is limited to work with a single visual object formed by lines, which is presented in an environment without noise and distractors. In [12], a computational model of spatial memory was proposed. This computational model was implemented in LIDA cognitive architecture and tested in static virtual environments. This cognitive architecture is a model of the human mind (based on global workspace theory), not of the human brain [12]. The computational model of the VSWM proposed in this paper aims to offer an algorithm biologically inspired by the functionality of the human brain. This computational model aims to endow autonomous UAVs with the visuospatial working memory cognitive function to use these aerial robots in real environments. Table 1 shows a comparative chart between the bio-inspired VSWM computational model proposed in this paper and related work. Relevant characteristics of these computational models are identified and highlighted, such as research approach, biological inspiration, type of world, environments, type of maps (visual and spatial maps), and classification tasks. This research is expected to be useful to the AI computing research community to improve their knowledge and understanding of human beings’ cognitive functions from an abstract representation of such cognitive functions. Additionally, this research can be useful to guide future research on bio-inspired intelligent systems capable of exhibiting human-like behavior.

This paper is structured as follows. In Section 2, design details of the modules that make up the bio-inspired VSWM computational model proposed in this paper are offered. Additionally, neuroscientific evidence is included as part of the description of each module to support the bio-inspired approach. In Section 3, the material and methods used for validating the bio-inspired computational model are described. An initial version of the computational model proposed was implemented and tested on a commercial drone.

In Section 4, a detailed analysis of the results and behavior exhibited by the drone is presented. Finally, Section 5 provides a description of experimental results and some concluding remarks.

**Table 1.** Relevant characteristics of the proposed model and related work.

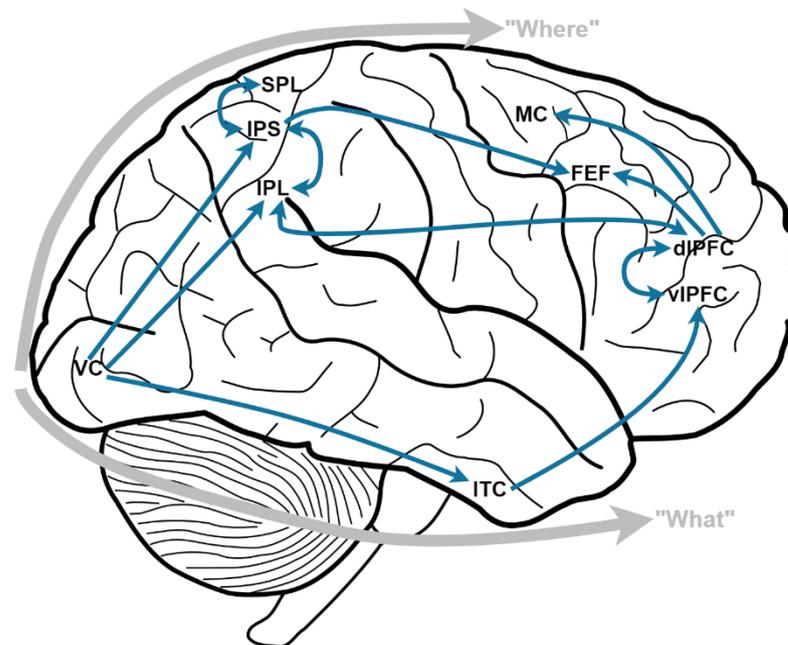
Computational Models	Research Approach	Biological Inspiration	Agent	World	Environment	Type of Maps	Classification Task
RatSLAM [10]	Neuroscience and Artificial Intelligence	The rodent brain	Ground robots	Real	Static and dynamic environments	Spatial and visual maps	Landmark-based places
A computational model for visual processing and object recognition [20]	Neuroscience and Artificial Intelligence	The human brain	Virtual human characters	Virtual	Static environments	Visual maps	2D basic geometry figures and simple 2D objects formed by lines
A computational model of spatial memory [12]	Psychology and Artificial Intelligence	The human brain	Virtual human characters	Virtual	Static environments	Spatial and visual maps	Object-based places
A bio-inspired VSWM computational model	Neuroscience and Artificial Intelligence	The human brain	Aerial robots	Real	Static and dynamic environments	Spatial and visual maps	People and vehicles

## 2. A Bio-Inspired Computational Model for Autonomous UAVs

The VSWM of human beings combines stimuli identification with information about their location [16,21]. There are different cognitive processes underlying the building of visuospatial maps in the VSWM. These cognitive processes can be grouped into two major blocks. Such grouping is based on the division of cortical visual processing identified in both humans and non-humans (primates) as a dorsal and a ventral pathway (see Figure 1) [21–23]. The human brain areas belonging to the ventral pathway (commonly known as the “What” pathway) have been associated with cognitive tasks focused on identifying and recognizing visual stimuli to generate visual maps of them. On the other hand, the human brain areas belonging to the dorsal pathway (commonly known as the “Where” pathway) have been associated with cognitive tasks focused on processing spatial information of visual stimuli to generate spatial maps of them. Anatomically, the ventral pathway has been described as a multisynaptic pathway projecting from the visual cortex (VC) to the anterior temporal target (TE area) in the inferior temporal cortex (ITC), with a further projection from the ITC to the ventrolateral prefrontal cortex (vlPFC), whereas the dorsal pathway has been described as a multisynaptic pathway projecting from the visual cortex (VC) to the posterior parietal cortex (PPC), with a further projection from the PPC to the dorsolateral prefrontal cortex (dlPFC) [21,24]. The PPC itself is divided into an upper and lower portion: the superior parietal lobe (SPL) and inferior parietal lobe (IPL), respectively. These two lobes are separated from one another by a sulcus called the intraparietal sulcus (IPS). Processes involved in the ventral pathway focus on identifying and classifying objects in the environment in order to offer visual information to the VSWM [16,21], whereas processes involved in the dorsal pathway focus on identifying locations of objects in the environment in order to offer spatial information to the VSWM [16,21].

The bio-inspired VSWM computational model proposed in this paper has taken inspiration from the underlying neural correlates to the VSWM that have been identified in both human and non-human (primate) brains. Therefore, modules of the computational model have been labeled with the name of the brain area that they represent. Figure 1 shows the brain areas that have been considered to propose the bio-inspired VSWM computational model. These brain areas are part of the ventral and dorsal pathways, respectively. The frontal eye fields (FEF) are a frontal brain region that receives converging inputs from both the dorsal and the ventral streams. This brain region contributes to the guidance of saccades. Additionally, the motor cortex (MC) has also been considered as part of the dorsal pathway in order to generate motor behaviors. Figure 1 shows

connections that have been considered to propose the computational model of the VSWM. These connections show how visual information flows through these pathways in order to generate visuospatial maps. This does not mean that the bio-inspired computational model proposed in this paper considers all brain areas involved in the VSWM of human beings, or that it implements exact neural mechanisms. This bio-inspired computational model just offers an artificial and abstract representation of major brain areas related to the VSWM.



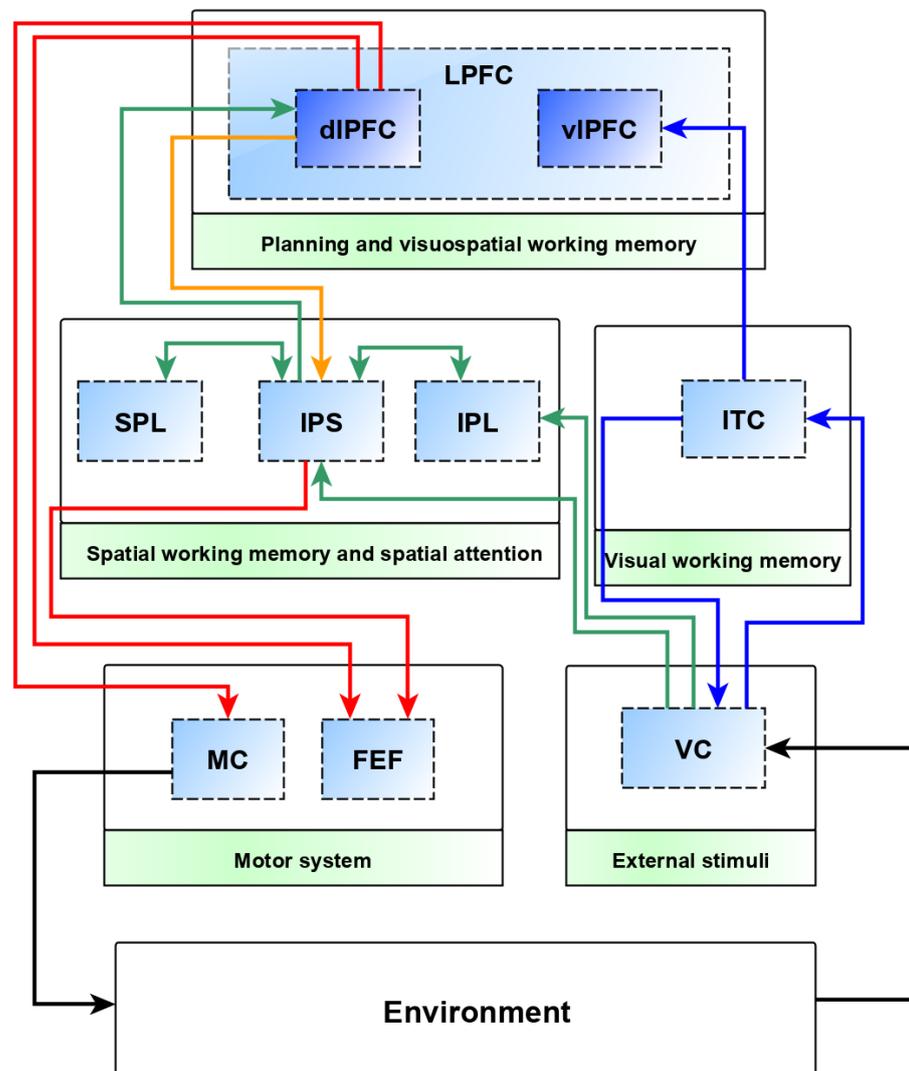
**Figure 1.** Identification of brain areas associated with the dorsal and ventral visual pathways, respectively.

Figure 2 shows the architectonic design of the bio-inspired VSWM computational model proposed in this paper. It also uses color codes for identifying the type of information that each module receives, processes, and sends. This information is classified as visual information (blue arrows), spatial information (green arrows), highly cognitive control information (yellow arrows), and sensorimotor information (red arrows). Finally, black arrows indicate the drone's interaction with the environment. The following subsections offer a detailed description of inputs, processes, and outputs associated with the bio-inspired computational model's modules. Description is presented from a bio-inspired approach, but details about their computational implementation are also offered.

### 2.1. External Stimuli

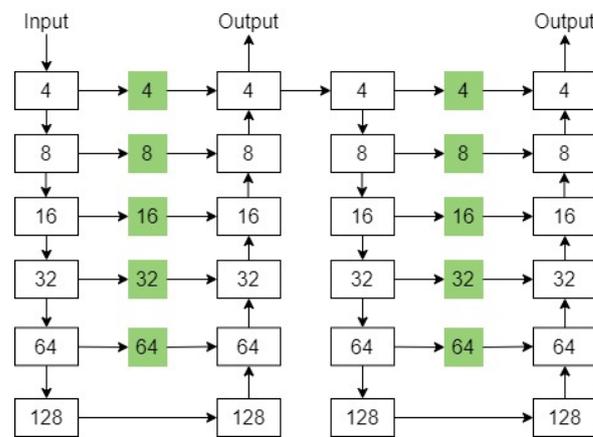
- **VC module.** This module represents the visual cortex. The VC is a complex system consisting of different brain areas encompassing V1, V2, V3, and V4 [25,26]. Offering a detailed computational model of the VC is outside the scope of this paper. Therefore, the VC has been simplified as a single module within the bio-inspired VSWM computational model. Despite our limited knowledge of neuronal mechanisms and processes involved in the VC, studies have demonstrated that visual perception starts in early visual areas with the detection of diverse low-level visual features such as angles, size, orientation, motion, and color [25–28]. This piecemeal analysis is very different from our subjective perception. However, these low-level visual features encoded through large distributed neuronal populations in the VC are fundamental for identifying complex objects in posterior visual processing brain areas [25]. Additionally, neuroscientific evidence suggests that the process of segmenting images into objects and background starts in the primary visual cortex (area V1) [25]. Therefore, the VC module proposed in this paper is responsible for separating relevant visual stimuli from the background and then sends visual information to the ITC, IPL, and IPS

modules for future processes related to spatial and visual identification. Additionally, a connection from the ITC module to the VC module has been considered in order to offer feedback. This feedback allows decreasing the number of likely regions of interest (ROIs) early identified by the VC module. The VC module implements a heatmap-based visual stimuli detection method. This method is an adapted implementation of the heatmap-based object detection method used in CenterNet [29,30]. This method was adapted for making a rough separation of visual stimuli belonging to the foreground from the background. The algorithm implemented in this module works as follows:



**Figure 2.** The architectonic design of the bio-inspired VSWM computational model.

The VC module processes an input RGB image  $I$  of width  $W$  and height  $H$ , where  $I \in R^{W \cdot H \cdot 3}$ . The VC module aims to produce a keypoint heatmap  $\hat{Y} \in [0, 1]^{\frac{W}{R} \cdot \frac{H}{R} \cdot C}$  by using a stacked hourglass network that allows downsampling of the input image with an output stride  $R = 4$ , followed by two sequential hourglass modules. Each hourglass module is a symmetric 5-layer down- and up-convolutional network, as shown in Figure 3.



**Figure 3.** The hourglass network architecture. White squares represent stage nodes and green squares represent sum nodes.

A heatmap head, dimension head, and offset head are obtained for each image and each C class (our C class includes two keypoint types that represent relevant visual stimuli). Heatmap head is used for the estimation of keypoints on an input image. A heatmap focal loss is defined for training the heatmap head. Focal loss improves the keypoint detection by weighting the keypoints detections. Equation (1) shows the focal loss function, where a value 1 for  $Y_{xyc}$  represents a positive keypoint for a class and a different value represents a negative keypoint.

$$L_k = \frac{-1}{N} \sum_{xyc} \begin{cases} (1 - \hat{Y}_{xyc})^\alpha \log(\hat{Y}_{xyc}) & \text{if } Y_{xyc} = 1 \\ (1 - Y_{xyc})^\beta (\hat{Y}_{xyc})^\alpha \log(1 - \hat{Y}_{xyc}), & \text{otherwise} \end{cases} \quad (1)$$

$\alpha$  and  $\beta$  are hyper-parameters with values of 2 and 4, respectively, and  $\hat{Y}_{xyc}$  represents a keypoint prediction. On the other hand, the offset head is used to recover the discretization error caused by the output stride. An L1 Norm Offset Loss is defined for training the offset head. Equation (2) describes the L1 Norm Offset Loss.

$$L_{off} = \frac{1}{N} \sum_p \left| \hat{O}_{\tilde{p}} - \left( \frac{p}{R} - \tilde{p} \right) \right| \quad (2)$$

Dimension head predicts the dimension boxes of the keypoints using an L1 Norm Dimension Loss (see Equation (3)) for training, where  $\hat{S}$  represents the predicted dimensions and  $s$  is actual ground truth dimensions.

$$L_{size} = \frac{1}{N} \sum_{k=1}^N \left| \hat{S}_{pk} - s_k \right| \quad (3)$$

The overall training objective is defined in Equation (4). This equation represents the total loss of the network, where  $\lambda_{size}$  and  $\lambda_{off}$  were set to 0.1 and 1, respectively.

$$L_{det} = L_k + \lambda_{size} L_{size} + \lambda_{off} L_{off} \quad (4)$$

An extended explanation of the heatmap-based object detection method can be found in [29,30].

### 2.2. Visual Working Memory

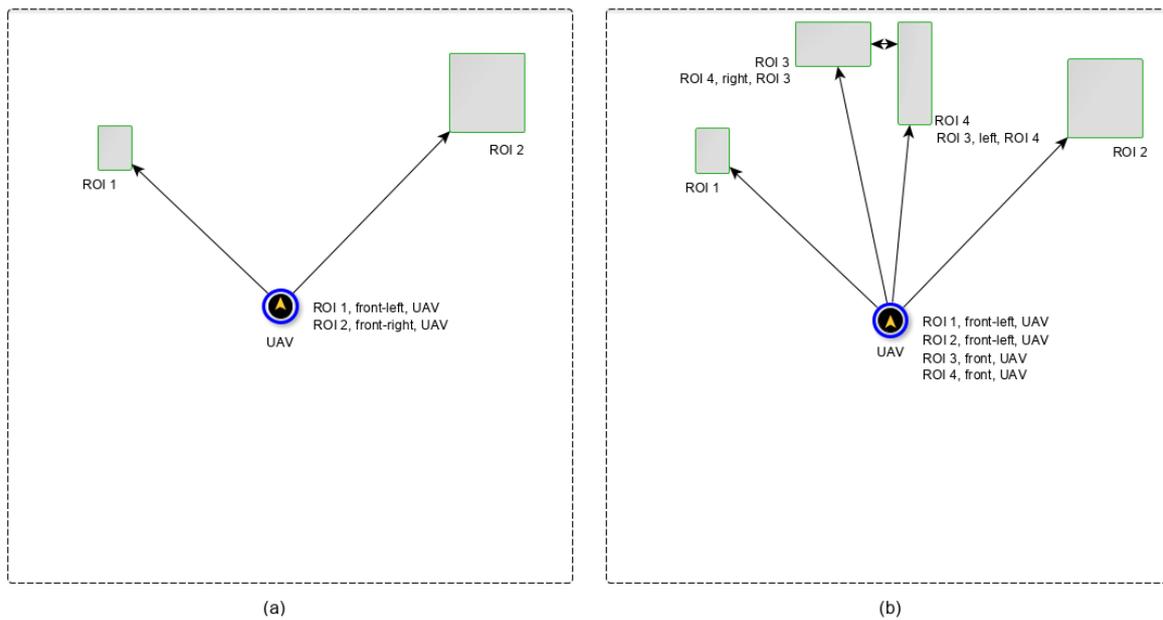
- ITC module. This module represents the inferior temporal cortex. This brain area has been associated with the process of identifying complex objects in the environment, such as animate and inanimate objects [31,32]. Findings in neuroscience show that vi-

sual information starts in the visual area V1 and passes through a sequence of processing stages in the visual cortex until complex object representations are formed in the anterior part of the ITC [24]. Therefore, the ITC module proposed in this paper receives a constant stream of visual stimuli coming from the VC module to identify complex objects in the environment. As part of the implementation of this module, we have included a convolutional neural network (CNN) for classifying relevant stimuli according to a set of category clusters. The ITC module receives the heatmaps produced on the VC module and extracts the peaks for each class independently. All responses whose value is greater or equal to its 8-connected neighbors are detected and keep the top 100 peaks. Let  $\hat{P}_c$  be the set of  $n$  detected center points  $\hat{P} = \{(\hat{x}_i, \hat{y}_i)\}_{i=1}^n$  of class  $c$ . Each keypoint location is given by integer coordinates  $(x_i, y_i)$ . Keypoint values  $\hat{Y}_{x_i, y_i, c}$  are used as a measure of its detection confidence, and produce a bounding box at location  $(\hat{x}_i + \delta\hat{x}_i - \frac{\hat{w}_i}{2}, \hat{y}_i + \delta\hat{y}_i - \frac{\hat{h}_i}{2}, \hat{x}_i + \delta\hat{x}_i + \frac{\hat{w}_i}{2}, \hat{y}_i + \delta\hat{y}_i + \frac{\hat{h}_i}{2})$ , where  $(\delta\hat{x}_i, \delta\hat{y}_i) = \hat{O}_{\hat{x}_i, \hat{y}_i}$  is the offset prediction and  $(\hat{w}_i, \hat{h}_i) = \hat{S}_{\hat{x}_i, \hat{y}_i}$  is the size prediction. The resulting data are used to feed a CNN with a  $3 \times 3$  max pooling operation. This operation allows avoiding detection of an object as multiple objects. After classifying visual stimuli, the ITC module sends highly processed visual information to the vLPFC module. According to Madl et al. [12], CNNs trained with real-world images are highly similar to those recorded in the ITC of humans and non-humans (primates).

### 2.3. Spatial Working Memory

- **SPL module.** This module represents the superior parietal lobe. This brain area has been associated with tasks related to spatial working memory, maps of coordinates, spatial shifting, and spatial attention [17,33,34]. Therefore, the SPL module is responsible for decoding visual cues sent by the VC module through the IPS module. Information of these cues is used by the SPL module for creating maps of coordinates. These maps are encoded in an egocentric fashion, according to relative coordinates of the UAV and its orientation. Allocentric representations, relative to environmental references, are not considered yet in this first version of the bio-inspired VSWM computational model. Neurophysiological studies suggest that the superior parietal lobe is part of the networking involved in visual search tasks [33,34]. Therefore, the SPL module helps with spatial shifting and spatial attentional tasks in order to detect displacement of visual stimuli. Additionally, this module shares visuospatial information with the IPS module.

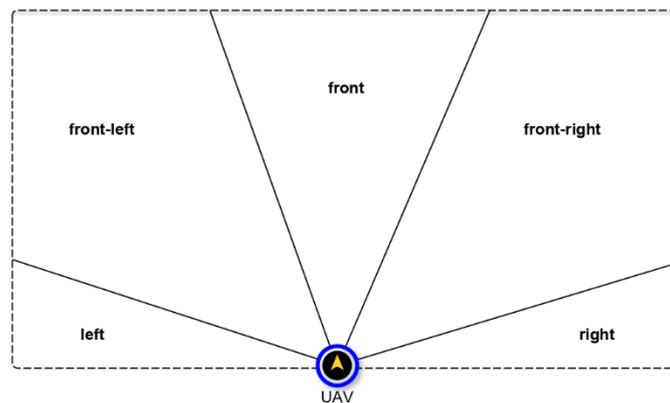
Figure 4 shows an example of coordinate maps created by the SPL module. Fuzzy spatial information of the relationship between the UAV and ROIs identified in the environment is processed and included in these maps. For instance, Figure 4a shows that in a first scene two ROIs have been identified. According to the ROIs' coordinates, the SPL module has considered that ROI 1 is located to the front left of the UAV, and ROI 2 is located to the front right of the UAV. On the other hand, Figure 4b shows a second scene, where two new ROIs have been identified. When two or more ROIs are close to each other, the SPL module generates additional fuzzy spatial information of the relationship between these ROIs.



**Figure 4.** (a) A coordinate map that represents the spatial relationship between the UAV and ROIs identified in the environment. (b) A second coordinate map with additional spatial information between ROIs that are close to each other.

In order to establish the positions of the identified ROIs relative to the UAV, Equations (5) and (6) are defined for segmenting the UAV’s field of view into five regions as shown in Figure 5. These trivial equations allow the UAV to situate ROIs in the environment.

$$f(x) = \begin{cases} -2x, & x < 0 \\ 2x, & x \geq 0 \end{cases} \tag{5}$$



**Figure 5.** The UAV’s view field segmentation.

Equation (5) establishes the limit between the region classified as front and the regions classified as front left and front right when  $x$  and  $f(x)$  represent the  $x$ -coordinate and  $y$ -coordinate, respectively.

$$f(x) = \begin{cases} -\frac{x}{4}, & x < 0 \\ \frac{x}{4}, & x \geq 0 \end{cases} \tag{6}$$

Equation (6) establishes the limit between the region classified as front left and left and front right and right. Moreover, a local matrix of  $N \cdot M$  dimensions is applied over each ROI to generate additional fuzzy spatial information when two or more ROIs are close to each other. Equations (5) and (6), and their reflecting functions in relation to the  $x$  axis are applied for segmenting this local matrix to establish spatial relationships between ROIs.

- **IPS module.** This module represents the intraparietal sulcus. In monkeys and humans, the IPS subdivides the PPC into a superior and an inferior parietal lobe [17]. The IPS has been associated with the creation of attentional priority maps, also known as saliency maps. This brain area is part of networking involved in the visual search task [33]. Therefore, the IPS module has bidirectional communication with the SPL and IPL modules, which are involved in the visual search task too. The IPS module is responsible for decoding visual cues coming from the VC module in order to represent spatial information for the SWM. The creation of attentional priority maps is part of the tasks done by this module. An attentional priority map is a topographic representation of the distribution of attentional weights. In order to create an attentional priority map, the IPS module integrates both bottom-up (perceptual features of the stimulus) and top-down (high-level representations of expectations and action-goals) information [35,36]. Bottom-up information comes from the VC module, whereas top-down information comes from the dlPFC module. The attentional priority map is useful in the presence of distractors because the calibration of attentional weights allows the computational model to resolve the competition between stimuli presented in the environment. This module sends visuospatial information generated by itself and other modules (such as the SPL, and IPL modules) to the dlPFC module. Additionally, this module has direct communication with the FEF module in order to support visually guided actions. However, a mechanism for generating priority maps has not been implemented yet in this first version of the bio-inspired VSWM computational model.
- **IPL module.** This module represents the inferior parietal lobe. This brain area has been considered a relevant brain area involved in the interruption of the current cognition activity and the reorientation of attention when a salient stimulus of high behavioral relevance appears at an unexpected position [17]. However, the physical appearance of a stimulus is determined not only by perceptual factors but also by top-down processes such as expectations or behavioral goals. The IPL has also been associated with maintaining alertness, sustaining attention, and detecting novelty [17]. Therefore, the current IPL module proposed in this paper has been limited to sending cues for indicating or alerting when a new ROI appears in the environment. The IPL module receives visual stimuli coming from the VC module. Additionally, the IPL module has bidirectional communication with the IPS module to share information.

#### 2.4. Visuospatial Working Memory

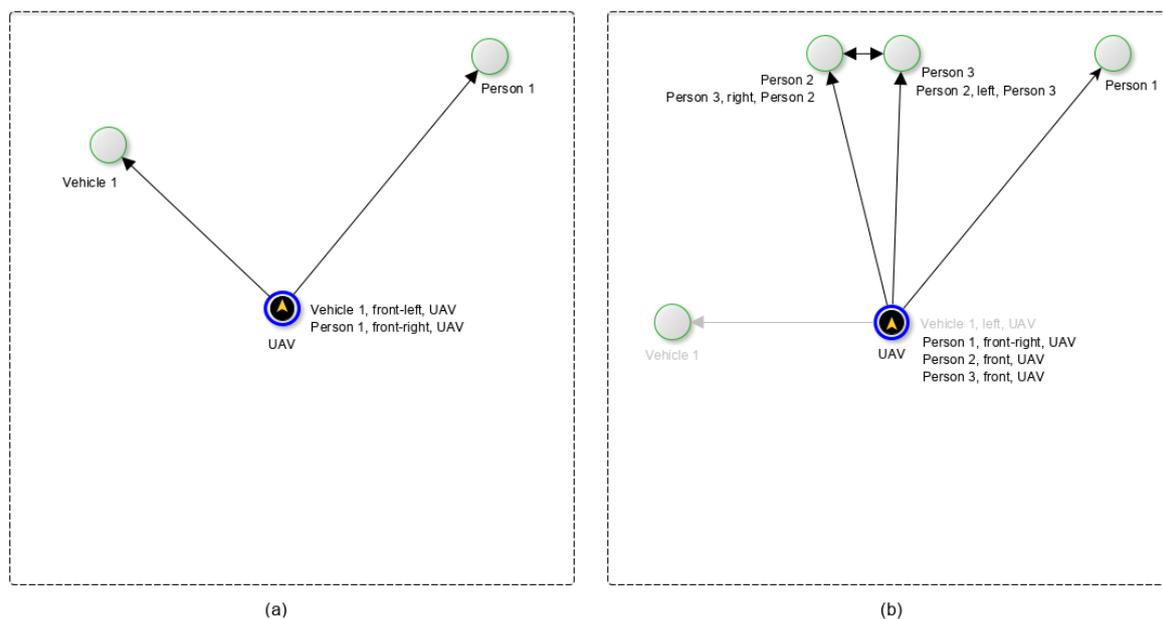
- **LPFC module.** This module represents the lateral prefrontal cortex (LPFC). Findings in neuroscience indicate that broadly, but strictly within the domain of vision, the entire LPFC is involved in maintenance and manipulation tasks such as attention, working memory, and switching task sets [24]. The LPFC module includes the dlPFC and vlPFC modules. Therefore, this module integrates visual and spatial information coming from the dlPFC and vlPFC modules, respectively, to create visuospatial maps.

Figure 6 shows an example of visuospatial maps created by the LPFC module. These maps offer a coherent and continuous representation of interest objects that have been identified, along with their spatial relationship in the environment. For instance, Figure 6a shows a first scene where a vehicle and a person have been identified and labeled as vehicle 1 and person 1, respectively. Spatial information indicates that vehicle 1 is located in front, but to the left side of the UAV, and person 1 is located in front, but to the right side of the UAV. On the other hand, Figure 6b shows a second scene, where two additional persons (labeled person 2 and person 3, respectively) have appeared and they are close to each other. In this case, when two or more interest objects are close to each other, additional spatial information is generated and included. This spatial information indicates the spatial relationship between these objects. Additionally, Figure 6b shows what happens when an object has gone out of the UAV's visual field. In this case, a selective removal process is activated. However, this process can be stopped if the object reappears in the environment.

Therefore, whereas the VSWM tries to maintain information about the objects that are no longer present in the environment, the selective removal process operates on outdated information to limit the VSWM's load and hence facilitates the maintenance of relevant information [37,38]. To implement this behavior in the bio-inspired VSWM computational model, Equation (7) was coded as part of the selective removal process:

$$R = e^{-\frac{t}{s}} \quad (7)$$

where  $R$  is retrievability (a measure of how easy it is to retrieve a piece of information from the VSWM),  $s$  is stability of memory (how fast  $R$  falls over time in the absence of a relevant stimulus), and  $t$  is time. As shown in Figure 6b, when a stimulus that has been previously identified is no longer present in the environment, its representation in the VSWM is degraded until it disappears. The graph represents this process by fading its node until it disappears, simulating how irrelevant stimuli that are no longer present in the environment are removed from the UAV's VSWM.



**Figure 6.** (a) A visuospatial map with information about the identified objects in the environment and their spatial relationship with the UAV. (b) A visuospatial map with additional spatial information about the relation between two close objects. Additionally, a graphical representation is shown of an object (vehicle 1) that is being forgotten.

- **dIPFC module.** This module represents the dorsolateral prefrontal cortex. There is neuroscientific evidence that shows that the dorsolateral prefrontal cortex is responsible for spatial selectivity [24]. The dIPFC has also been associated with planning, problem solving [39] and top-down executive control [21]. Currently, the dIPFC module is responsible for integrating spatial information coming from the IPS module on the visuospatial maps and sending motor commands to the FEF and MC modules in order to exhibit top-down behavior.
- **vIPFC module.** This module represents the ventrolateral prefrontal cortex. Neuroscientific evidence shows that the vIPFC is responsible for object selectivity [24]. Currently, the vIPFC module is responsible for integrating visual information coming from the ITC module on the visuospatial maps.

### 2.5. Motor System

- **FEF module.** This module represents the frontal eye field. Neuroscientific evidence indicates that the frontal eye field receives converging inputs from many cortical areas

involved in bottom-up and top-down attentional control. Currently, the computational module that represents the frontal eye field proposed in this paper considers two connections coming from the dlPFC and IPS modules [17,36,40]. Motor commands implemented in this module allow the UAV's camera to pan and tilt a specific number of degrees.

- MC module. This module represents the motor cortex. This brain area can be divided into three major areas: the primary motor cortex, the premotor cortex, and the supplementary motor area. Voluntary movements depend critically on these motor areas [41]. However, proposing a detailed computational model that considers all motor areas involved in generating movements is outside the scope of this paper. Therefore, the motor cortex has been simplified as a single module labeled MC. This module has a set of motor commands such as take-off, land, rotate clockwise, rotate counterclockwise, fly forward, fly backward, fly up, fly down, fly left, and fly right. These commands can be invoked by the dlPFC module.

### 3. Materials and Methods

The bio-inspired VSWM computational model proposed in this paper was implemented on a Bebop 2. This is a low-cost commercial drone equipped with a 14 mega-pixel camera with a fish-eye lens, dual-core processor with quad-core GPU, GPS, and 8-GB flash storage system. Nevertheless, the bio-inspired VSWM computational model was not embedded in the drone. Instead, a laptop (with an Intel core i7 processor, an external NVIDIA GeForce GTX 1060 video card, and 16 GB of RAM) was used to host the bio-inspired VSWM computational model. Therefore, the drone and the bio-inspired VSWM computational model communicate through a Wi-Fi connection. Currently, the bio-inspired VSWM computational model can identify people and vehicles (cars and trucks) in the real world. The following research questions were studied in order to identify the performance and accuracy of the bio-inspired VSWM computational model proposed in this paper:

1. What is the visuospatial ability of the bio-inspired VSWM computational model when the drone flies between 9.84 and 16.40 feet?
2. Can the bio-inspired VSWM computational model offer a coherent and continuous representation of visual and spatial relationships among interest objects present in the environment?

To answer these questions, a total of 30 test cases were designed and executed. These test cases were grouped into three scenarios (i) environments with static and dynamic vehicles, (ii) environments with people, and (iii) environments with people and vehicles. Detailed information on these test cases is available at [CogniDron/VSWM/test\\_report](#). In all the test cases, the drone did not have a specifically defined target in the visual task. Therefore, visual attention was driven by bottom-up processes for generating visuospatial maps of the environment.

#### 3.1. First Test Scenario: Environments with Static and Dynamic Vehicles

The main objective of this test scenario was to test the bio-inspired VSWM computational model's visuospatial ability to identify vehicles present in the environment and then to generate 2D visuospatial maps of them. A second objective of this test scenario was to test the bio-inspired VSWM computational model's capability for retaining visuospatial information for a short period of time when visual stimuli were no longer present in the environment. Therefore, 12 test cases were designed and implemented in this test scenario. Table 2 shows a summary of the test cases implemented in this scenario. This summary includes the test case ID, the flight altitude reached by the drone for executing each test case, as well as a brief description of each test case. Additionally, there was no specific target defined in this visual task for all the test cases. Therefore, the drone's camera was fixed at a specific point in the environment in all the test cases.

**Table 2.** Test cases with static and dynamic vehicles.

Test ID	Altitude	Description
1	9.84 feet	Visual stimuli involved in this test case were two parked vehicles. All vehicles were always presented in the drone's field of view.
2	13.12 feet	
3	16.40 feet	
4	9.84 feet	Visual stimuli involved in this test case were a parked vehicle and a vehicle in motion. All vehicles were always presented in the drone's field of view.
5	13.12 feet	
6	16.40 feet	
7	9.84 feet	Visual stimuli involved in this test case were a parked vehicle and a vehicle in motion. After the two vehicles were identified by the bio-inspired computational model, the vehicle in motion was intentionally driven outside of the drone's field of view and returned after a few seconds to test the bio-inspired computational model's ability to retain visuospatial information for a short period of time when visual stimuli are no longer present.
8	13.12 feet	
9	16.40 feet	
10	9.84 feet	Visual stimuli involved in this test case were a parked vehicle and a vehicle in motion. After the two vehicles were identified by the bio-inspired computational model, the vehicle in motion was intentionally driven outside of the drone's field of view and returned after a few seconds to test the bio-inspired computational model's ability to retain visuospatial information for a short period of time when visual stimuli are no longer present. Finally, the vehicle in motion was permanently left outside of the drone's field of view to test the bio-inspired computational model's selective removal process responsible for degrading visuospatial information of visual stimuli that are no longer present in the environment until such visuospatial information is deleted from the VSWM's buffer.
11	13.12 feet	
12	16.40 feet	

### 3.2. Second Test Scenario: Environments with People

Like the first test scenario, the main objective of this test scenario was to test the bio-inspired VSWM computational model's visuospatial ability. However, this second test scenario involved people instead of vehicles. Therefore, 9 test cases were designed and implemented in this test scenario. Participants were strategically located in the environment in order to obtain images of them at different scales. Table 3 shows a summary of the test cases implemented in this test scenario. Like the first test scenario, there was no specific target defined in this second scenario. Therefore, the drone's camera was fixed at a specific point in the environment in all the test cases.

**Table 3.** Test cases with people.

Test ID	Altitude	Description
13	9.84 feet	Visual stimuli involved in this test case were three persons in the environment. The persons walked aimlessly around the environment, but they always avoided a partial or total occlusion between them.
14	13.12 feet	
15	16.40 feet	
16	9.84 feet	Visual stimuli involved in this test case were three persons in the environment. These persons were located diagonally in front of the drone. After the three persons were identified by the bio-inspired computational model, the person who was located to the right of the drone walked around the environment and crossed behind another person in order to generate a partial or total occlusion between them for a few seconds.
17	13.12 feet	
18	16.40 feet	
19	9.84 feet	Visual stimuli involved in this test case were three persons in the environment. The persons walked aimlessly around the environment, but they always avoided a partial or total occlusion between them. After the three persons were identified by the bio-inspired computational model, one of them went outside of the drone's field of view. After that, a second person went outside of the drone's field of view too.
20	13.12 feet	
21	16.40 feet	

### 3.3. Third Test Scenario: Environments with People and Vehicles

The objective of this test scenario was the same one proposed in the previous test scenarios. However, in this test scenario, both people and vehicles were presented simultaneously in the environment. In total, 9 test cases were designed and implemented in this test scenario. Table 4 shows a summary of these test cases. Like previous test scenarios,

there was no specific target defined in this test scenario. Therefore, the drone's camera was fixed at a specific point in the environment in all the test cases.

**Table 4.** Test cases with people and vehicles.

Test ID	Altitude	Description
22	9.84 feet	Visual stimuli involved in this test case were two persons and two parked vehicles in the environment. After the persons and vehicles were identified by the bio-inspired computational model, the persons walked aimlessly around the environment. Persons always avoided a partial or total occlusion between themselves and the vehicles.
23	13.12 feet	
24	16.40 feet	
25	9.84 feet	Visual stimuli involved in this test case were two persons and two parked vehicles in the environment. After the persons and vehicles were identified by the bio-inspired computational model, the persons walked aimlessly around the environment and then they generated a partial or total occlusion between themselves for a few seconds.
26	13.12 feet	
27	16.40 feet	
28	9.84 feet	Visual stimuli involved in this test case were two persons and two parked vehicles in the environment. After the persons and vehicles were identified by the bio-inspired computational model, the persons walked aimlessly around the environment and then they went outside of the drone's field of view, but at different times.
29	13.12 feet	
30	16.40 feet	

#### 4. Results

This paper considers visuospatial ability as a person's capacity to identify visual and spatial relationships among objects. Therefore, the visuospatial ability of the bio-inspired VSWM computational model was measured in terms of the ability to classify and locate objects in the environment. Images and videos were recorded in each test case to analyze off-line the bio-inspired VSWM computational model's visuospatial ability. (Supplementary Materials are available at CogniDron/VSWM/videos) In order to estimate the bio-inspired computational model's accuracy for generating spatial relationships among relevant stimuli in the environment, 100 frame samples were taken from the video of each test case. We observed that spatial relationships among relevant stimuli present in the environment were always right in the three test scenarios. Frames taken by the bio-inspired computational model were stored in order to compute the bio-inspired computational model's accuracy for classifying relevant visual stimuli (such as people and vehicles) present in the environment. Results of each test scenario are reported in Tables 5–7, respectively.

**Table 5.** Results of the first test scenario.

Test Id	Test Title	Feet	Parked Vehicles	Vehicles in Motion	Images	Total of Vehicles	Detected Vehicles	False Positive	False Negative	Visual Accuracy
1	Visuospatial information	9.84	2	0	135	270	270	0	0	100%
2		13.12	2	0	79	158	158	0	0	100%
3		16.4	2	0	101	202	202	0	0	100%
4		9.84	1	1	113	226	224	0	2	99.12%
5		13.12	1	1	114	228	223	0	5	97.81%
6		16.4	1	1	87	174	166	0	8	95.40%
7	Retaining visuospatial information for a short period of time	9.84	1	1	155	283	283	0	0	100%
8		13.12	1	1	198	368	362	0	6	98.37%
9		16.4	1	1	118	222	214	0	8	96.40%
10	Selective removal process for the visuospatial working memory buffer	9.84	1	1	123	176	176	0	0	100%
11		13.12	1	1	113	191	189	0	2	98.95%
12		16.4	1	1	79	134	129	0	5	96.27%

**Table 6.** Results of the second test scenario.

Test Id	Test Title	Feet	People	Images	Total of People	Detected People	False Positive	False Negative	Visual Accuracy
13	Visuospatial information	9.84	3	158	474	473	0	1	99.79%
14		13.12	3	87	261	260	0	1	99.62%
15		16.4	3	65	195	194	0	1	99.49%
16	Retaining visuospatial information for a short period of time	9.84	3	114	338	338	0	0	100%
17		13.12	3	101	299	289	1	10	96.33%
18		16.4	3	98	288	285	0	3	98.96%
19	Selective removal process for the visuospatial working memory buffer	9.84	3	113	249	249	0	0	100%
20		13.12	3	177	364	361	0	3	99.18%
21		16.4	3	150	266	260	0	6	97.74%

**Table 7.** Results of the third test scenario.

Test Id	Test Title	Feet	People	Parked Vehicles	Images	Total of People	Total of Vehicles	Detected People	Detected Vehicles	People		Vehicles		Visual Accuracy
										False Positive	False Negative	False Positive	False Negative	
22	Visuospatial information	9.84	2	2	129	258	258	258	258	0	0	0	0	100%
23		13.12	2	2	107	214	214	214	214	0	0	0	0	100%
24		16.4	2	2	95	190	190	190	190	0	0	0	0	100%
25	Retaining visuospatial information for a short period of time	9.84	2	2	103	202	206	201	206	0	1	0	0	99.75%
26		13.12	2	2	130	256	260	256	260	0	0	0	0	100%
27		16.4	2	2	116	228	232	228	232	0	0	0	0	100%
28	Selective removal process for the visuospatial working memory buffer	9.84	2	2	102	156	204	156	204	0	0	0	0	100%
29		13.12	2	2	114	153	228	153	228	0	0	0	0	100%
30		16.4	2	2	106	140	212	140	212	0	0	0	0	100%

Table 5 shows the results obtained in the first test scenario. The bio-inspired computational model's accuracy for classifying vehicles was from 95.40% to 100% in this test scenario, whereas the bio-inspired computational model's accuracy for classifying people was from 96.33% to 100% in the second test scenario (see Table 6).

Finally, Table 7 shows the results obtained in the third test scenario. The bio-inspired computational model's accuracy for classifying both people and vehicles simultaneously was from 99.75% to 100% in this third test scenario.

The selective removal process proposed in the bio-inspired VSWM computational model was useful for maintaining a coherent and continuous representation of visual and spatial relationships among interest objects presented in the environment. Figure 7 shows how visuospatial information is maintained even when a visual stimulus is lost as a consequence of a total occlusion. As a result of this occlusion, the bio-inspired VSWM computational model keeps the last position where the visual stimulus was identified, and the selective removal process is activated. This removal process allows degrading information of the visual stimulus that disappeared. However, this process can be stopped if the visual stimulus reappears in the environment. When a visual stimulus reappears in the environment, the visuospatial map updates its location otherwise, the visual stimulus information is removed from the visuospatial map. The same behavior was observed when the bio-inspired VSWM computational model had problems identifying visuospatial information from one frame to another.

Figure 8 shows an example of how the bio-inspired VSWM computational model generates additional fuzzy spatial relationship information between relevant stimuli when they are close to each other. For instance, four people are present in the environment shown in Figure 8. The bio-inspired computational model identified a first person to the front-right of the drone, a second person in front of the drone, and two more persons to the front-left of the drone. Additionally, the bio-inspired computational model considers that these two last persons are close to each other, which shows that additional fuzzy spatial information is generated to establish the spatial relationship between these persons. Finally, we are aware that the visuospatial performance of the bio-inspired computational model proposed in this paper can vary according to the environment's features, such as lights and shadows, occlusions, the number of visual stimuli, and background clutter.



Figure 7. Generating visuospatial maps when a total occlusion happens.

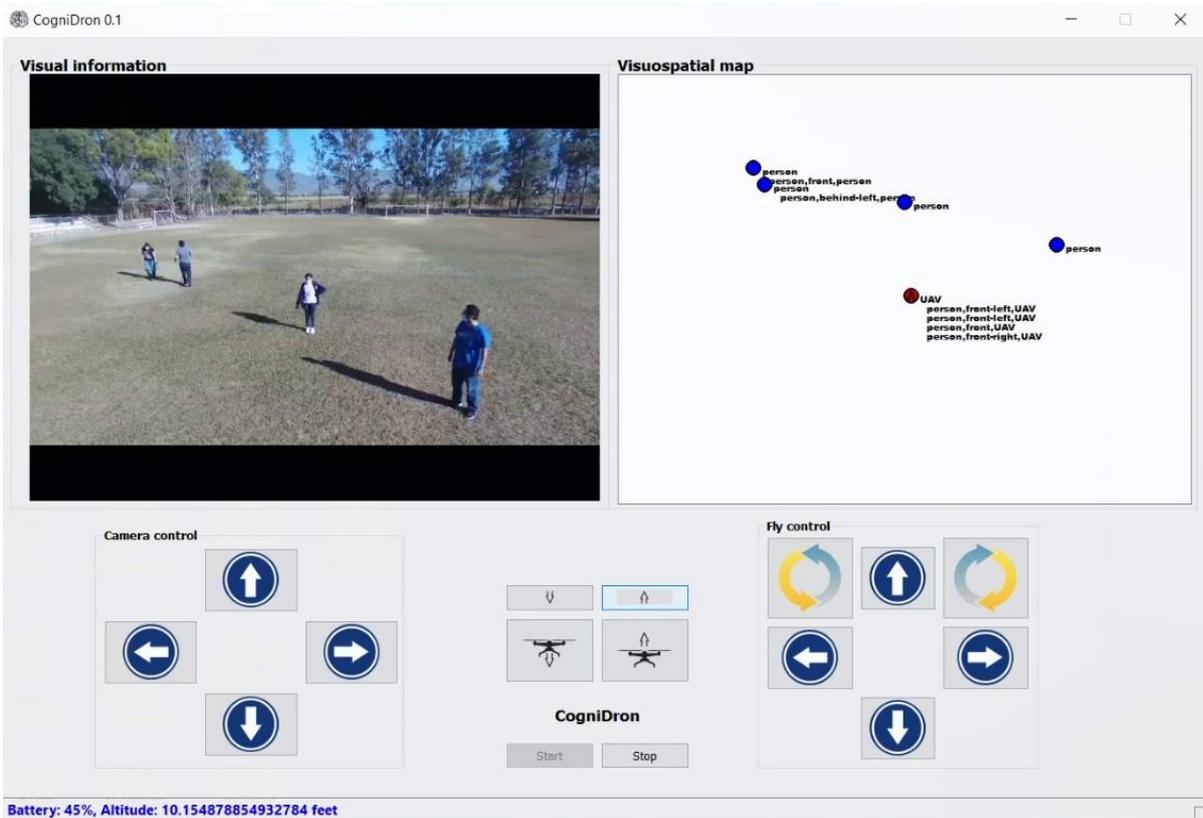


Figure 8. Additional fuzzy spatial relationship information between relevant stimuli when they are close to each other.

## 5. Discussion

VSWM is a fundamental cognitive capability of human beings needed for exploring the visual environment. This cognitive function is responsible for creating visuospatial maps, which are then further processed for the generation of specific actions such as spatial shifting, visual search, and location of targets. Obtained results after executing the test cases show that the bio-inspired VSWM computational model proposed in this paper is capable of maintaining a coherent and continuous representation of visual and spatial relationships among interest objects presented in the environment even when a visual stimulus is lost because of a total occlusion. We consider that this bio-inspired computational model represents a step towards autonomous UAVs capable of forming visuospatial mental imagery in realistic environments. As future work, we are going to implement our bio-inspired VSWM computational model on UAVs with a stereo vision to add depth information to the visuospatial maps. We believe that designing autonomous UAVs capable of forming visuospatial mental imagery can be very useful for doing different tasks such as surveillance and rescue missions where VSWM is a fundamental cognitive capability for exploring the visual environment. On the other hand, the bio-inspired VSWM computational model proposed in this paper is part of a new brain-inspired cognitive architecture named CogniDron. Therefore, future work also includes modeling and designing other cognitive functions such as creating allocentric spatial maps, emotional system, planning, decision-making, and learning.

**Supplementary Materials:** The following are available online at <https://drive.google.com/drive/folders/1b0yWNkhVh08OhaPkEzrzTjzpv5Mm2-eP>. The videos and images of the 30 test cases implemented for testing the bio-inspired VSWM computational model, as well as additional information regarding these test cases to support the computational model proposed in this paper, are available at CogniDron/VSWM/videos and images.

**Author Contributions:** Conceptualization, J.-A.C., S.L., S.C., A.M. and J.-H.R.; methodology, J.-A.C., S.L. and S.C.; software, J.-A.C., S.C., A.M. and J.-H.R.; validation, J.-A.C., S.L. and S.C.; writing—original draft, J.-A.C. and S.L.; writing—review and editing, S.C., A.M. and J.-H.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Kim, S.J.; Jeong, Y.; Park, S.; Ryu, K.; Oh, G. A survey of drone use for entertainment and AVR (augmented and virtual reality). In *Auagmented Reality and Virtual Reality*; Jung, T., Tom Dieck, M.C., Eds.; Springer: Berlin/Heidelberg, Germany, 2018; pp. 339–352. [[CrossRef](#)]
2. Stankov, U.; Kennell, J.; Morrison, A.M.; Vujičić, M.D. The view from above: The relevance of shared aerial drone videos for destination marketing. *J. Travel Tour. Mark.* **2019**, *36*, 808–822. [[CrossRef](#)]
3. López, S.; Cervantes, J.-A.; Cervantes, S.; Molina, J.; Cervantes, F. The plausibility of using unmanned aerial vehicles as a serious game for dealing with attention deficit-hyperactivity disorder. *Cogn. Syst. Res.* **2020**, *59*, 160–170. [[CrossRef](#)]
4. Calvario, G.; Sierra, B.; Alarcón, T.E.; Hernandez, C.; Dalmau, O. A multi-disciplinary approach to remote sensing through low-cost UAVs. *Sensors* **2017**, *17*, 1411. [[CrossRef](#)]
5. Stepinac, M.; Gašparović, M. A review of emerging technologies for an assessment of safety and seismic vulnerability and damage detection of existing masonry structures. *Appl. Sci.* **2020**, *10*, 5060. [[CrossRef](#)]
6. Boubeta-Puig, J.; Moguel, E.; Sánchez-Figueroa, F.; Hernández, J.; Preciado, J.C. An autonomous UAV architecture for remote sensing and intelligent decision-making. *IEEE Internet Comput.* **2018**, *22*, 6–15. [[CrossRef](#)]
7. Hassabis, D.; Kumaran, D.; Summerfield, C.; Botvinick, M. Neuroscience-inspired artificial intelligence. *Neuron* **2017**, *95*, 245–258. [[CrossRef](#)]

8. Katona, J. A Review of Human-Computer Interaction and Virtual Reality Research Fields in Cognitive InfoCommunications. *Appl. Sci.* **2021**, *11*, 2646. [[CrossRef](#)]
9. Katona, J.; Ujbanyi, T.; Sziladi, G.; Kovari, A. Speed control of Festo Robotino mobile robot using NeuroSky MindWave EEG headset based brain-computer interface. In Proceedings of the 2016 7th IEEE international conference on cognitive infocommunications (CogInfoCom), Wroclaw, Poland, 16–18 October 2016; pp. 251–256. [[CrossRef](#)]
10. Milford, M.; Jacobson, A.; Chen, Z.; Wyeth, G. RatSLAM: Using models of rodent hippocampus for robot navigation and beyond. In *Robotics Research*; Inaba, M., Corke, P., Eds.; Springer: Cham, Switzerland, 2016; pp. 467–485. [[CrossRef](#)]
11. Ramos Corchado, F.F.; López Fraga, A.C.; Salazar Salazar, R.; Ramos Corchado, M.A.; Begovich Mendoza, O. Cognitive Pervasive Service Composition Applied to Predatory Crime Deterrence. *Appl. Sci.* **2021**, *11*, 1803. [[CrossRef](#)]
12. Madl, T.; Franklin, S.; Chen, K.; Montaldi, D.; Trapp, R. Towards real-world capable spatial memory in the LIDA cognitive architecture. *Biol. Inspired Cogn. Archit.* **2016**, *16*, 87–104. [[CrossRef](#)]
13. Metta, G.; Natale, L.; Nori, F.; Sandini, G.; Vernon, D.; Fadiga, L.; Von Hofsten, C.; Rosander, K.; Lopes, M.; Santos-Victor, J. The iCub humanoid robot: An open-systems platform for research in cognitive development. *Neural Netw.* **2010**, *23*, 1125–1134. [[CrossRef](#)]
14. Zhao, F.; Zeng, Y.; Wang, G.; Bai, J.; Xu, B. A brain-inspired decision making model based on top-down biasing of prefrontal cortex to basal ganglia and its application in autonomous UAV explorations. *Cogn. Comput.* **2018**, *10*, 296–306. [[CrossRef](#)]
15. Xu, Y. Reevaluating the sensory account of visual working memory storage. *Trends Cogn. Sci.* **2017**, *21*, 794–815. [[CrossRef](#)]
16. Van der Stigchel, S.; Hollingworth, A. Visuospatial working memory as a fundamental component of the eye movement system. *Curr. Dir. Psychol. Sci.* **2018**, *27*, 136–143. [[CrossRef](#)]
17. Ptak, R. The frontoparietal attention network of the human brain: Action, saliency, and a priority map of the environment. *Neuroscientist* **2012**, *18*, 502–515. [[CrossRef](#)]
18. Pisella, L. Visual perception is dependent on visuospatial working memory and thus on the posterior parietal cortex. *Ann. Phys. Rehabil. Med.* **2017**, *60*, 141–147. [[CrossRef](#)]
19. Jerde, T.A.; Merriam, E.P.; Riggall, A.C.; Hedges, J.H.; Curtis, C.E. Prioritized maps of space in human frontoparietal cortex. *J. Neurosci.* **2012**, *32*, 17382–17390. [[CrossRef](#)]
20. González-Casillas, A.; Parra, L.; Martin, L.; Avila-Contreras, C.; Ramirez-Pedraza, R.; Vargas, N.; del Valle-Padilla, J.L.; Ramos, F. Towards a model of visual recognition based on neurosciences. *Procedia Comput. Sci.* **2018**, *145*, 214–231. [[CrossRef](#)]
21. Kravitz, D.J.; Saleem, K.S.; Baker, C.I.; Mishkin, M. A new neural framework for visuospatial processing. *Nat. Rev. Neurosci.* **2011**, *12*, 217–230. [[CrossRef](#)] [[PubMed](#)]
22. Ludwig, K.; Sterzer, P.; Kathmann, N.; Hesselmann, G. Differential modulation of visual object processing in dorsal and ventral stream by stimulus visibility. *Cortex* **2016**, *83*, 113–123. [[CrossRef](#)]
23. Freud, E.; Plaut, D.C.; Behrmann, M. ‘What’ is happening in the dorsal visual pathway. *Trends Cogn. Sci.* **2016**, *20*, 773–784. [[CrossRef](#)]
24. Kravitz, D.J.; Saleem, K.S.; Baker, C.I.; Ungerleider, L.G.; Mishkin, M. The ventral visual pathway: An expanded neural framework for the processing of object quality. *Trends Cogn. Sci.* **2013**, *17*, 26–49. [[CrossRef](#)]
25. Poort, J.; Raudies, F.; Wannig, A.; Lamme, V.A.; Neumann, H.; Roelfsema, P.R. The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. *Neuron* **2012**, *75*, 143–156. [[CrossRef](#)]
26. Skalicky, S.E. The primary visual cortex. In *Ocular and Visual Physiology*; Skalicky, S.E., Ed.; Springer: Singapore, 2016; pp. 207–218. [[CrossRef](#)]
27. Garg, A.K.; Li, P.; Rashid, M.S.; Callaway, E.M. Color and orientation are jointly coded and spatially organized in primate primary visual cortex. *Science* **2019**, *364*, 1275–1279. [[CrossRef](#)]
28. Ghose, G.M.; Daniel, Y. Integration of color, orientation, and size functional domains in the ventral pathway. *Neurophotonics* **2017**, *4*, 031216. [[CrossRef](#)]
29. Xu, Z.; Hrustic, E.; Vivet, D. CenterNet Heatmap Propagation for Real-Time Video Object Detection. In Proceedings of the Lecture Notes in Computer Science, European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; Vedaldi, A., Bischof, H., Brox, T., Frahm, J.M., Eds.; Springer: Cham, Switzerland, 2020; pp. 220–234. [[CrossRef](#)]
30. Zhou, X.; Wang, D.; Krähenbühl, P. Objects as points. *arXiv* **2019**, arXiv:1904.07850.
31. McKee, J.L.; Riesenhuber, M.; Miller, E.K.; Freedman, D.J. Task dependence of visual and category representations in prefrontal and inferior temporal cortices. *J. Neurosci.* **2014**, *34*, 16065–16075. [[CrossRef](#)]
32. Kriegeskorte, N.; Mur, M.; Ruff, D.A.; Kiani, R.; Bodurka, J.; Esteky, H.; Tanaka, K.; Bandettini, P.A. Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* **2008**, *60*, 1126–1141. [[CrossRef](#)] [[PubMed](#)]
33. Vandenberghe, R.; Molenberghs, P.; Gillebert, C.R. Spatial attention deficits in humans: The critical role of superior compared to inferior parietal lesions. *Neuropsychologia* **2012**, *50*, 1092–1103. [[CrossRef](#)] [[PubMed](#)]
34. Gertz, H.; Lingnau, A.; Fiehler, K. Decoding movement goals from the fronto-parietal reach network. *Front. Hum. Neurosci.* **2017**, *11*, 84. [[CrossRef](#)]
35. Katsuki, F.; Constantinidis, C. Early involvement of prefrontal cortex in visual bottom-up attention. *Nat. Neurosci.* **2012**, *15*, 1160–1166. [[CrossRef](#)]
36. Bowling, J.T.; Friston, K.J.; Hopfinger, J.B. Top-down versus bottom-up attention differentially modulate frontal—parietal connectivity. *Hum. Brain Mapp.* **2020**, *41*, 928–942. [[CrossRef](#)]

37. Souza, A.S.; Czoschke, S.; Lange, E.B. Gaze-based and attention-based rehearsal in spatial working memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **2020**, *46*, 980. [[CrossRef](#)]
38. Lewis-Peacock, J.A.; Kessler, Y.; Oberauer, K. The removal of information from working memory. *Ann. N. Y. Acad. Sci.* **2018**, *1424*, 33–44. [[CrossRef](#)]
39. Nejati, V.; Salehinejad, M.A.; Nitsche, M.A. Interaction of the left dorsolateral prefrontal cortex (l-DLPFC) and right orbitofrontal cortex (OFC) in hot and cold executive functions: Evidence from transcranial direct current stimulation (tDCS). *Neuroscience* **2018**, *369*, 109–123. [[CrossRef](#)] [[PubMed](#)]
40. Hutchison, R.M.; Gallivan, J.P.; Culham, J.C.; Gati, J.S.; Menon, R.S.; Everling, S. Functional connectivity of the frontal eye fields in humans and macaque monkeys investigated with resting-state fMRI. *J. Neurophysiol.* **2012**, *107*, 2463–2474. [[CrossRef](#)] [[PubMed](#)]
41. Svoboda, K.; Li, N. Neural mechanisms of movement planning: Motor cortex and beyond. *Curr. Opin. Neurobiol.* **2018**, *49*, 33–41. [[CrossRef](#)] [[PubMed](#)]