

Objectives

List the main objectives of the proposed research in order of priority [up to 4000 chars]

It is not known how animals, including humans, build up a representation of the 3D scene around them as they move. The answer, if we could find it, would have important implications in industry, for example affecting the way that engineers programme autonomous mobile robots as they explore and map novel terrain. Increasingly, experimental evidence casts doubt on the prevailing model, which is based on 3D reconstruction. Indeed, an alternative approach using 'view-based' representations has been implemented in navigating robots and can provide an adequate account of navigational behaviours in humans and other animals such as route-following. There is also good evidence for view-based approaches underlying human object recognition. These abilities are, in some ways, at opposite ends of a spectrum of representation, from the scale of a single object to long-range navigation. However, as yet there is no adequate explanation of how a view-based model would help represent the location of objects in a room. This project sets out to test the hypothesis that view-based approaches could span a wide range of spatial abilities from object recognition to large scale navigation.

The specific goals of the project are as follows: a) We will determine whether view-based approaches could, in theory, account for performance in tasks that at present are assumed to require the brain to generate a 3D reconstruction a scene. b) We will test these models against the performance of human participants using immersive virtual reality. Our laboratory has one of the best-calibrated, low-latency virtual reality systems in the world for the study of visuo-spatial performance in humans and we will continue to devote effort to maintaining our position at the forefront of high-fidelity virtual reality display technology in this regard. Immersive virtual reality allows us to generate environments under tight experimental control, limit the amount of information available to participants where appropriate, and to transport participants 'virtually' to another location without proprioceptive feedback.

We will apply the following tests to evaluate both view-based and 3D reconstruction hypotheses. i) The errors in navigating to a previously-viewed location ('homing') will be measured while introducing subtle, unnoticed changes in the structure of the scene (so that there is no correct answer in the homing task). Because we have established models of both the view-based and 3D-reconstruction-based strategies in this task, we can arrange for the changes in scene structure to maximise the discriminative power of the data in separating the two models. ii) The accuracy and precision of pointing to a once-viewed object that is now out of sight can be predicted under a range of conditions for both the 3D reconstruction and view-based models and compared to human performance in the same task. iii) Participants will attempt to navigate to a location they have never visited but which they have seen from another location. iv) Finally, we will test participants' navigation and pointing performance using a novel stimulus that preserves information about corresponding points. The changes in the stimulus as observers explore the environment should not affect a 3D reconstruction algorithm based on corresponding points but will markedly alter the view-based information.

Thus, we aim to develop view-based models sufficiently for them to provide a credible alternative to 3D reconstruction for spatial navigation and spatial memory tasks, even in cases where this has not been considered a possibility in the past. Having detailed, quantitative models of both view-based and 3D-reconstruction-based strategies allows us to choose the most discriminative experiments when testing between them.

Summary

Describe the proposed research in simple terms in a way that could be publicised to a general audience [up to 4000 chars]. Note that this summary will be automatically published on EPSRC's website in the event that a grant is awarded.

The way that animals use visual information to move around and interact with objects involves a highly complex interaction between visual processing, neural representation and motor control. Understanding the mechanisms involved is of interest not only to neuroscientists but also to engineers who must solve similar problems when designing control systems for autonomous mobile robots and other visually guided devices.

Traditionally, neuroscientists have assumed that the representation delivered by the visual system and used by the motor system is something like a 3D model of the outside world, even if the reconstruction is a distorted version of reality. Recently, evidence against such a hypothesis has been mounting and an alternative type of theory has emerged. 'View-based' models propose that the brain stores and organises a large number of sensory contexts for potential actions. Instead of storing the 3D coordinates of objects, the brain creates a visual representation of a scene using 2D image parameters, such as widths or angles, and information about the way that these change as the observer moves. This project examines the human representation of three-dimensional scenes to help distinguish between these two opposing hypotheses.

To do this, we will use immersive virtual reality with freely-moving observers to test the predictions of the 3D reconstruction and 'view-based' models. Head-tracked virtual reality allows us to control the scene the observer sees and to track their movements accurately. Certain spatial abilities have been taken as evidence that the observer must create a 3D reconstruction of the scene in the brain. For example, people are able to view a scene, remember where objects are, walk to a new location and then point back to one of the objects they had seen originally even if it is no longer visible (i.e. people can update the visual direction of objects as they move). However, this capacity does not necessarily require that the brain generate a 3D model of the scene and, as evidence, we will extend view-based models to include this pointing task and others like it. We will then test the predictions of both view-based and 3D reconstruction models against the performance of human participants carrying out the same tasks.

As well as predicting the pattern of errors in simple navigation and pointing tasks, we will also measure the effect of two types of stimulus change. 3D reconstruction uses 'corresponding points' which are points in an image that arise, for example, from the same physical object (or part of an object) as a camera or person moves around it. Using a novel stimulus, we will keep all of these 'corresponding points' in a scene constant yet, at the same time, changing the scene so that the images alter radically when the observer moves. This manipulation should have a dramatic effect on a view-based scheme but no effect at all on any system based only on corresponding points.

Overall, we will have a tight coupling between experimental observations and quantitative predictions of performance under two types of model. This will allow us to determine which of the two models most accurately reflects human behaviour in a 3D environment. One potential outcome of the project is that view-based models will provide a convincing account of performance in tasks that have previously been considered to require 3D reconstruction, opening up the possibility that a wide range of tasks can be explained within a view-based framework.

Academic Beneficiaries

Describe who will benefit from the research [up to 4000 chars].

The academic communities that will benefit directly from this work fall into three main categories. First, neuroscientists interested in the coherence of visual perception across eye and head movements and in the visual representation of space; second, computer vision and robotics experts interested in efficient, robust representations for mapping and localisation; and third, academics interested in using high-end virtual reality facilities for their science.

In neuroscience, there is a relatively recent realisation that explaining the function of the cortex in a freely-moving, freely-fixating animal may require a radical re-think of established assumptions about brain processing. By analogy, it may be that the cortex provides one small brush-stroke per saccade contributing to a much larger canvas rather than creating monolithic representations (eye-, body-, hand-, world-centred) during each period of fixation. Such a re-think throws much greater weight on establishing the coordinate frame and characteristics of the 'larger canvas'. Our project will contribute to that re-appraisal.

In the computer vision and robotics community, there is increasing interest in finding coherent and robust alternatives to the use of 3D metric frames as the representation on which simultaneous localisation and mapping (SLAM) is based. A