1 Introduction
View-based and Cartesian representations provide rival accounts of visual navigation [1–3]. Here we show that the distribution of errors in navigation on the scale of a room are strongly influenced by the scene geometry, in ways that can be modelled using only simple view-based features.

2 Experiment
Participants were asked to complete a homing task in an immersive virtual reality environment. In interval one they were shown three very long coloured vertical poles from one viewing location. They were then transported (virtually) to another location, and in interval two, they tried to navigate to the initial viewing point, where they pressed a button (‘end-point’).

3 Data
The distributions of end-point errors on the ground plane differed significantly in shape and extent depending on pose configuration and goal location. In condition 2, where the three poles and the goal point almost lie on a single circle, the errors tend to be distributed around this arc. Conditions 36 and 45 show how a relatively small visual angle between two of the poles tends to lead to elongated errors on the ground plane in such a way as to preserve the ratio of angles between poles from these viewpoints, even if they are not extremely accurate.

4 Features
Simple visual features were used to describe the views available to the participants. Some are monotonic, single-view features, such as angles between the poles as measured from the cyclopean point. Others are inherently two-view features (stereo or motion) such as disparity, relative disparity, and disparity gradient.

5 Model Construction
We chose a subset of the candidate features (not necessarily independent — see Features) to describe the visual information at the goal and at each endpoint. The difference between these two vectors gave a single vector describing the error at the end point.

6 Model Comparison
Almost any combination of our visual features can be used to learn a model, so to compare the models, the maps were turned into true probability distributions by normalisation. A good model puts more of the probability mass around the regions in which data points are to be found.

7 Conclusion
Using visual features (such as angles between objects and disparity between objects) accounts well for participants’ behaviour in this navigation task. Preliminary modelling using a 3D reconstruction algorithm (not shown here) has shown a quite different pattern of predicted behaviour. We will develop these predictions further to test between Cartesian and view-based models of scene representation and navigation.

References