# Using Grid Cells for Coordinate Transforms

Marcus Lewis
mlewis@numenta.com

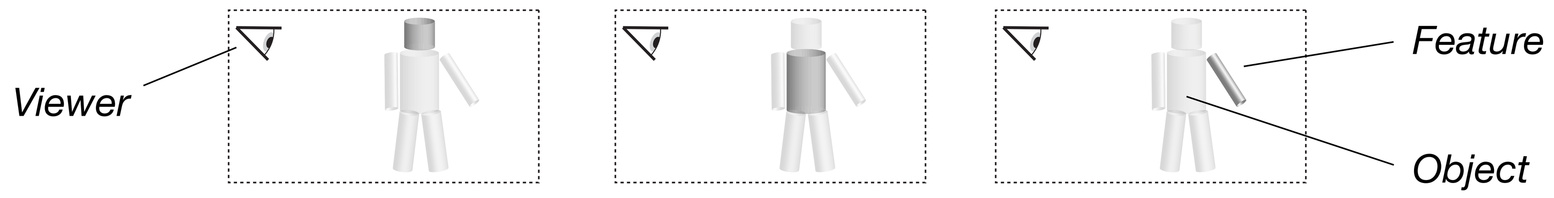
## Summary

The brain could use a grid cell code to represent sensed structures at locations in viewer-centric coordinates.

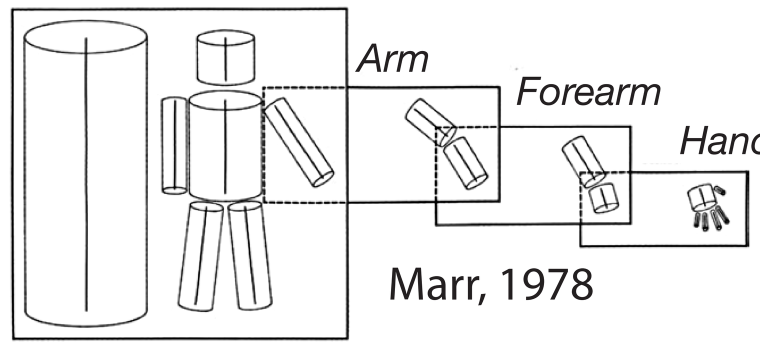
This grid cell code would make it possible for the brain to calculate the location of each structure relative to some higher structure, enabling object recognition.

## Object recognition requires a coordinate transform

1. Using sensory input, it's most natural for the brain to build a representation of the sensed features and their locations in **viewer-centric** coordinates.




2. But to learn / recognize objects, the brain needs to build a representation of the sensed features and their locations in **object-centric** coordinates.



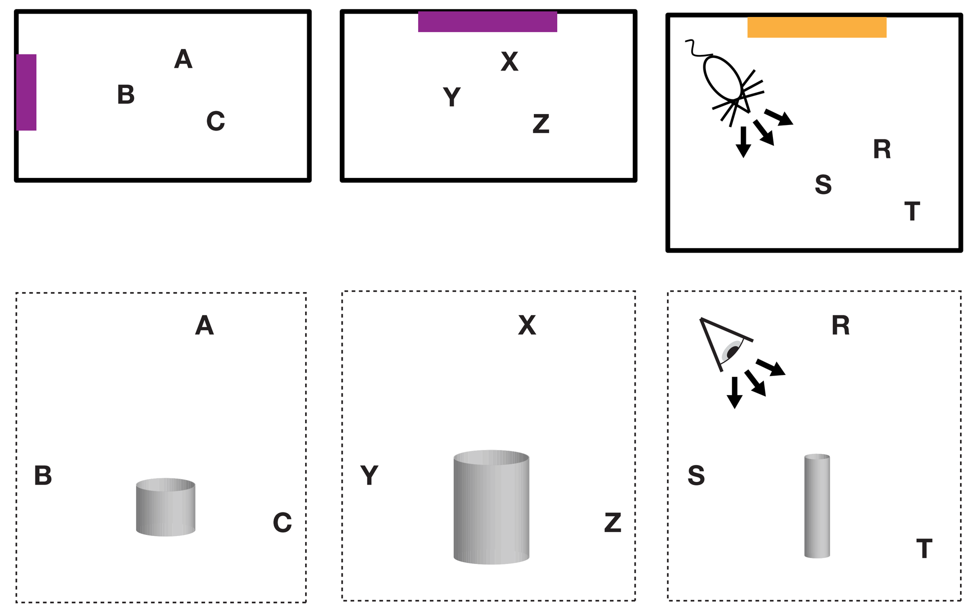

Marr, 1978

***Object recognition seems to require a transform from viewer-centric to object-centric coordinates.***

## 1. Using grid cells for viewer-centric representations

Entorhinal cortex seems to represent an animal's location in a specific environment.

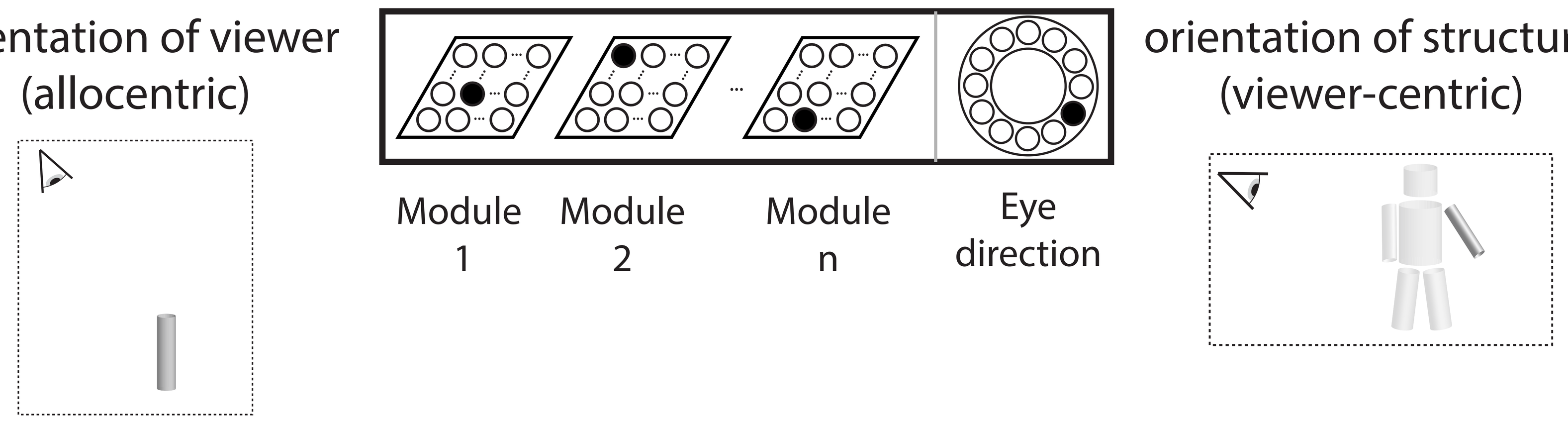
Neocortex could use this same grid code to represent the animal's location relative to the sensed feature or object.




**There's an alternate way to describe a grid cell representation.**

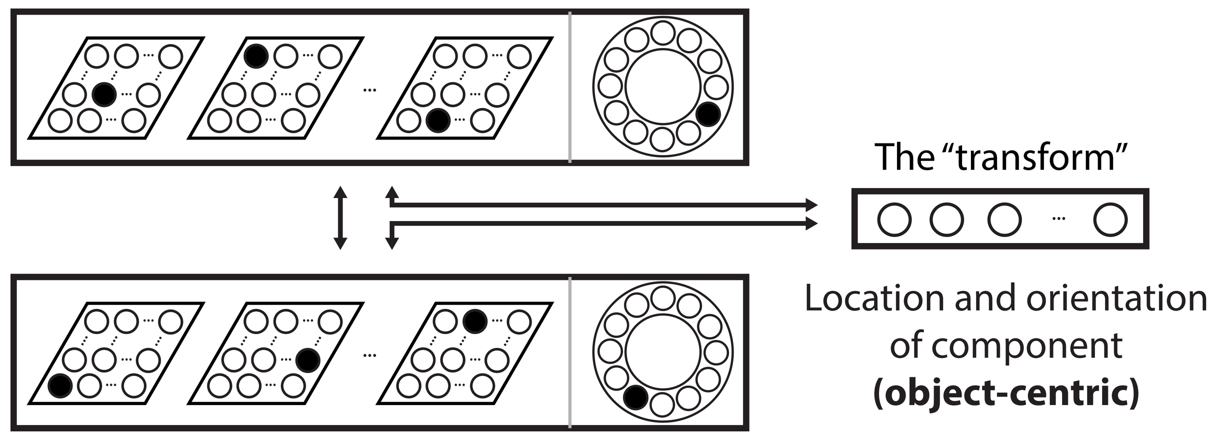
*Conventional:* Location and orientation of viewer (allocentric)

*Alternate:* Location and orientation of structure (viewer-centric)

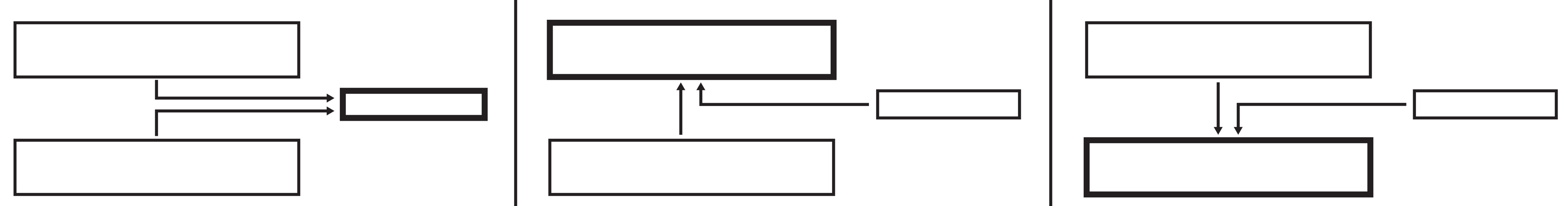



***The neocortex could use grid and head direction cell analogs to represent sensed features and objects.***

## 2. Detecting transforms between grid populations




**This coordinate transform circuit supports 3 operations.**



Build representation of object

- Detect the transform between component and object.
- Represent the object as a set of these transforms.

Predict component at location

- Activate representation of object.
- Predict each component and its location / orientation.

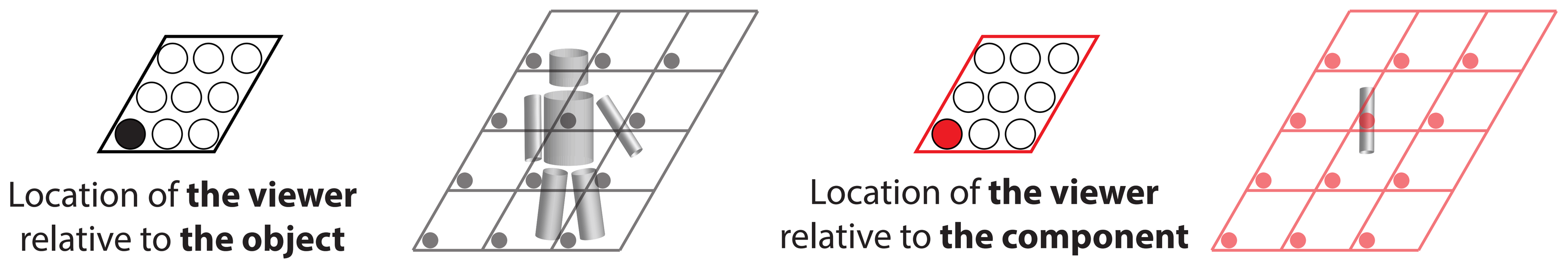
Recognize object at location

- Recall transforms for this component.
- Determine the location of these candidate objects.
- Repeat with multiple components until one candidate object "wins".

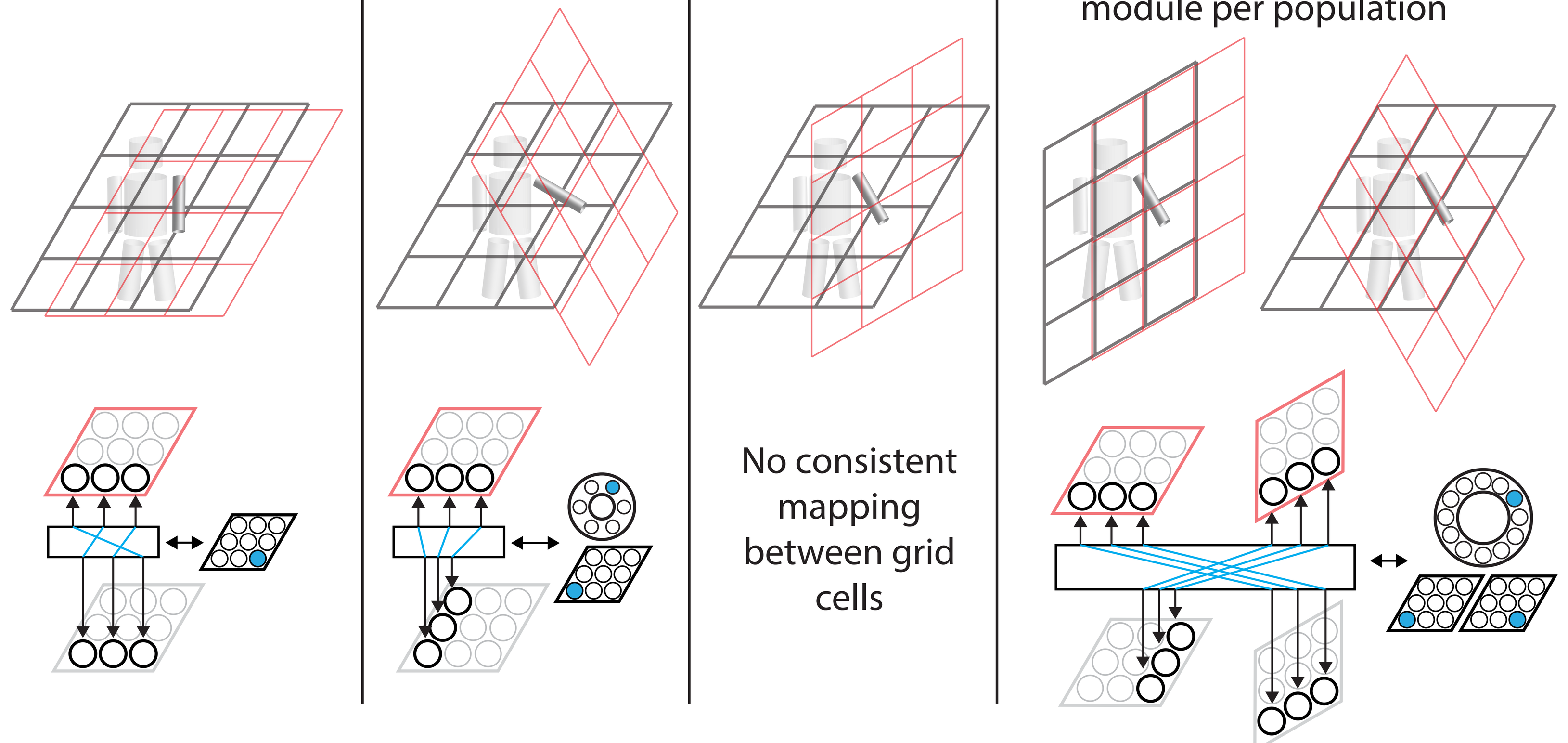
***Using this grid cell code, another population of cells could detect the object-centric location of a sensed feature.***

## Intuition: "Which grid cells fire together?"

Imagine two grids, one anchored to an object and another to a component.

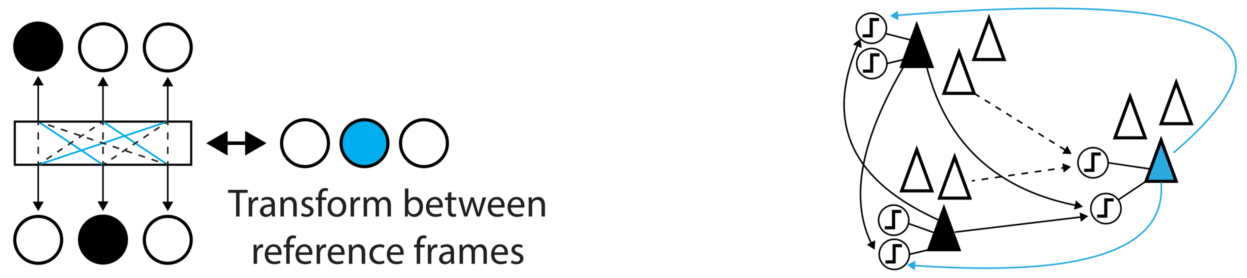



In some spatial arrangements of the two structures, each grid cell in one of these modules will consistently fire with a particular cell in the other module.

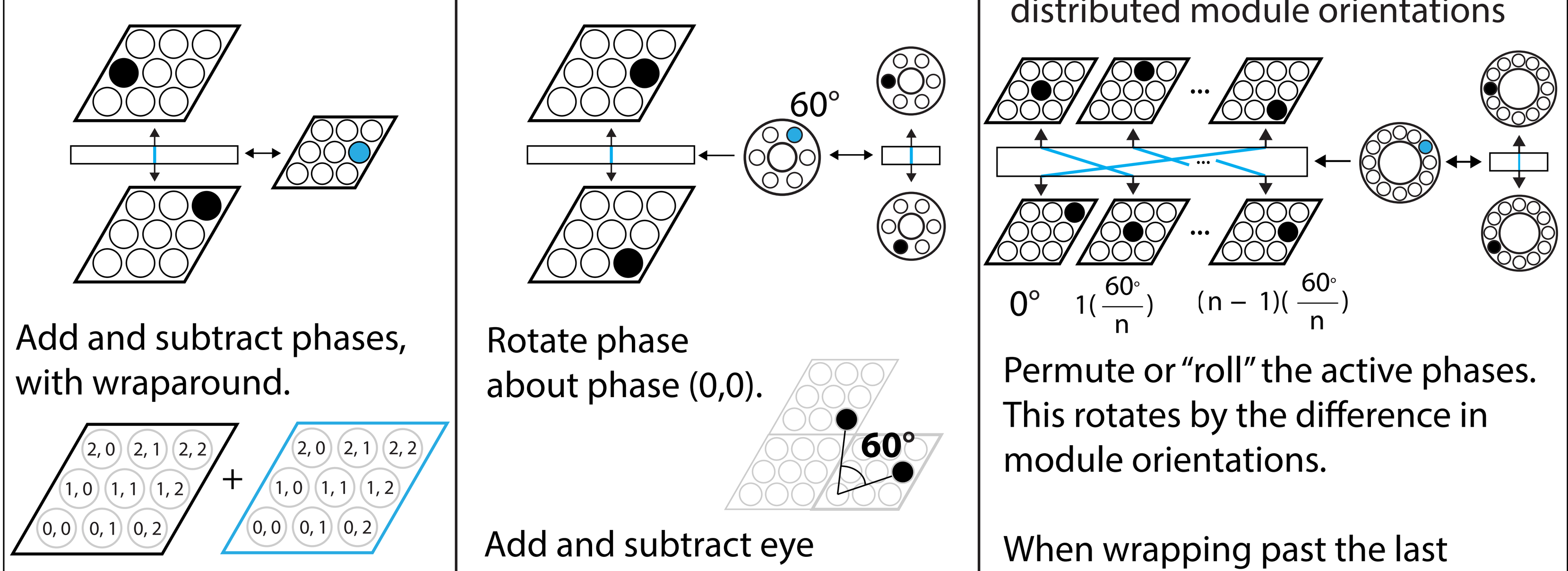



The transform represents which grid cells fire together. (Sampling shown above)

**Detailed mechanism: The "transform" routes between cells.**




**This neural circuit can perform three metric operations.**



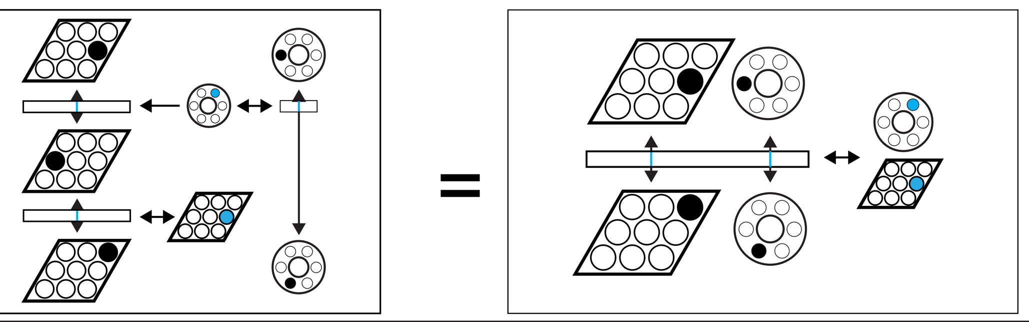

Translation: Add and subtract phases, with wraparound. (Similar to Bush et al. 2015)

Rotation by increments of 60°: Rotate phase about phase (0,0). Add and subtract eye directions, with wraparound.

Arbitrary rotations: Permute or "roll" the active phases. This rotates by the difference in module orientations. When wrapping past the last module, rotate by 60°.

**These operations can be chained and flattened into a single synapse hop.**

1. Rotate active phase about phase (0,0)
2. Translate active phase, correcting for the rotation and adding additional displacement.

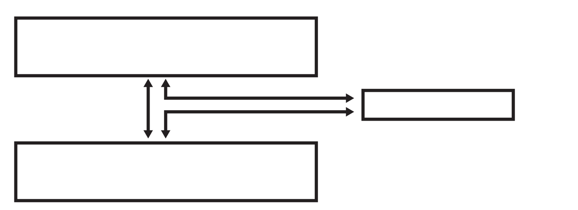


## Predictions of the theory

The cortex uses analogs of grid cells and head-direction cells to represent the processed sensory input. In this way, a single population can represent a specific sensed feature at a specific viewer-centric location and orientation.

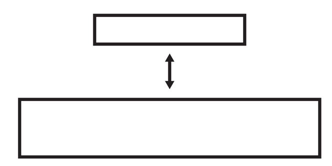
Grid cells connect to a population of cells that represents a viewer-invariant arrangement of components.

*Possibility #1: Hierarchical*



Multiple populations of grid cells could represent locations of higher and lower structures.

*Possibility #2: Oscillatory / attentional*



A single population of grid cells could represent locations of higher and lower structures at different times.

In areas that specialize in building object-centric representations of objects, grid cell modules will have a distribution of orientations.

## References

Marr, D., & Nishihara, H. K. (1978). Representation and Recognition of the Spatial Organization of Three-Dimensional Shapes. Proceedings of the Royal Society B: Biological Sciences.
Bush, D., Barry, C., Manson, D., Burgess, N. (2015). Using Grid Cells for Navigation, Neuron, 87(3), 507–520
Hawkins, J., Ahmad, S. & Cui, Y. (2017). A theory of how columns in the neocortex enable learning the structure of the world. Front. Neural Circuit., 11, 81.