

Assignment 4
Computational Vision
U. Minn. Psy 5036
Daniel Kersten

```
In[1]:= Off[General::spell1];
```

Themes

Modeling image variation: Scene-based vs. image-based Geometric vs. photometric variation

Different generative models make different predictions for recognition inference

Humans easily recognize familiar 3D objects from unfamiliar views. Humans can also recognize a familiar object under an unfamiliar illumination. How is this done? There are many possible models. Thinking about how to model image variation--i.e. the generative process, suggests some basic hypotheses to test, even if we don't have a precise model of recognition.

One can model viewpoint and illumination change using "true" 3D scene-based models of image formation. Perhaps the visual system in effect "knows" how the 3D structure of the object interacts with viewpoint and illumination, and uses this knowledge to allow for image variations.

Alternatively, image variation can be modeled using 2D image-based knowledge. One way this could be done is if the visual system has 2D knowledge of a few instances of images it has seen before, and then interpolates between them to allow for variation (Ullman & Basri, 1991).

In the first two exercises, you are going to see how to model **viewpoint** change using 3D (scene-based) and 2D (image-based) manipulations. This second (2D) method is just one of many ways to model 2D variations (e.g. if one has 2D "affine" operations for scale or rotation, or shear that are workable approximations to model 3D depth or 3D rotations over small domains, e.g. Liu & Kersten, 1998). In these exercises, you will be modeling "geometrical" variation.

In the second two exercises, you will model **illumination** change using 3D (scene-based) and 2D (image-based) manipulations. In these second two exercises, you will be modeling "photometric" variation.

3D Scene-based modeling of viewpoint change

2D Image-based modeling of viewpoint change

3D scene-based modeling of illumination change

2D image-based modeling of illumination change

QUESTIONS

You will be graded on the following

- **1. Evaluate and try to understand each step in this notebook.**

Then type your answers to the following questions:

- 2a. Explain in words the differences between how the images are generated using 3D vs. 2D models of viewpoint change. Describe one way the 2D approximation seems to depart from the "ground truth" represented by the 3D views.
- 2b. Suppose an experimentalist tells you that his human subjects could recognize an object equally well from all views (i.e. from any point on the 3D viewing sphere), even though they had initially been trained on only two views. As a modeler, would you be inclined to favor a 3D scene-based or a 2D image-based theory? Why?
- 3a. Explain in words the differences between how the images are generated using 3D vs. 2D models of illumination change.
- 3b. Suppose an experimentalist tells you that after training subjects on two different but specific illuminations (e.g. point light source on the left and then one on the right), her human observers recognize the object given "in-between" illuminations equally well, but not others (e.g. ones generated with light sources from above or below).
Would this support the 2D image-based model of illumination change described above or not? Why?
- 4. There are two major theoretical/computational problems with the image-based model for dealing with view-point change in non-wire objects: a) self-occlusion; b) identifying geometrical feature points.

Explain why these are minor problems for the wire-objects above, but are major problems for a typical "everyday" object.

References

© 2004, 2006, 2008 Daniel Kersten, Computational Vision Lab, Department of Psychology, University of Minnesota.