Cognitive Modeling
Homework 2: Bayesian Models of Visual Perception

Aleksander Frese (s163859) & Sebastian Sbirna (s190553)

1. Bayes – Exercise 1

This homework allows us to solve a perceptual problem using Bayesian inference rules and establish a mathematical model for how visual sight and image projection processing happens within our cognitive boundaries.

- **Question 1**

The first question involves trying to map out a reliable shape for a projection of the Necker cube. More specifically, we have an image *I* of a cube with 8 points. This is regarded as a 2D image projection of a shape, which our minds interpret as a 3D cubic shape due to the relative vertices of the cube in relation with each other, and also due to the specific angles that this image gives out.

What we are interested in is to try to maximize the posterior probability of a shape given the specific image, $P(S|I)$, knowing the Bayesian rules which link this probability to the likelihood of receiving an image given a certain shape, $P(I|S)$, and to the prior probability of the shape existing (i.e. being found) in the real world, $P(S)$. In mathematical terms, $P(S|I) = \frac{P(I|S)*P(S)}{P(I)}$. By maximizing the posterior probability of a shape determining our cubic-style image, we are asking to find the real-world shape which is closest (i.e. resembles best) to our spatial perception of that image. The magnitude of the likelihood can be very small, so instead, we may minimize the negative log likelihood instead, which converts the small values to larger negative ones with better finite precision.

In our first situation, we will assume a uniform prior probability which will make all shape probabilities within the real-world equal in value, thereby allowing us to remove out the prior as parameter from the model. The equation now becomes: $P(S|I) = P(I|S)$. In other words, the posterior probability is now equal to the likelihood.

Using this assumption, we compute the minimal value for negative log likelihood of a 3D shape, to give my Necker cube 2D image, and we get results as found in Figure 1 down below.

Interpreting the two results, we can see in Figure 1 (on the left side) that the shape which is supposed to project the sought-after cubic image, and also the real 3D cube shape, match identically in terms of their projected points, specifically in that initial state of vertices and edges placement (left side, Fig. 1). However, the moment you spatially rotate the image with a certain angle, you will see the true form of the underlying shape in blue, which is tremendously different than the cubic structure (shown in red) that we were looking for (right side, Fig. 1).

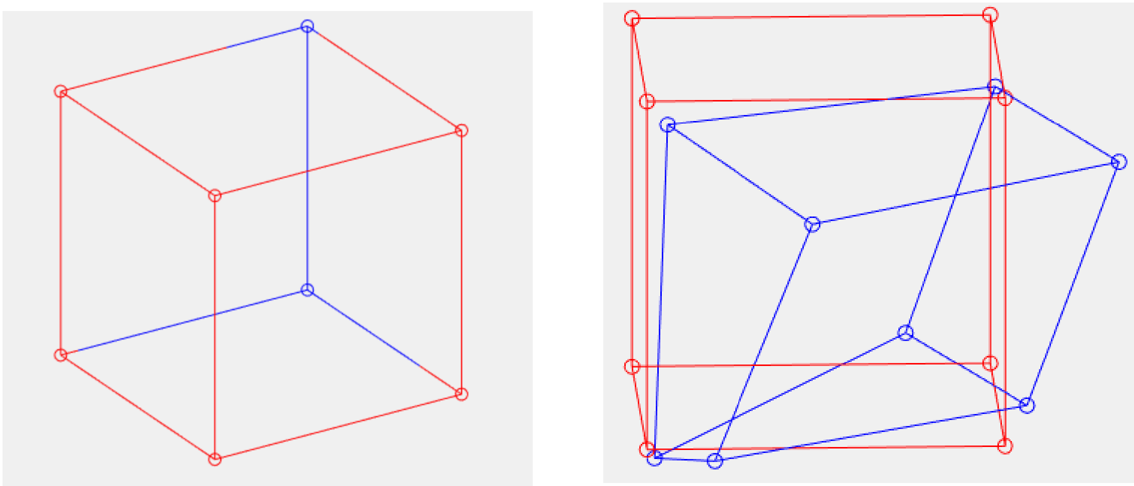

Fig. 1. *Visual representation of the 3D best fit scene (**blue**) and the real cube scene (**red**) in the initial perceived rotation (**left**) and in a random rotation (**right**)*

What could be the cause for this development? It is likely due to our initial ungrounded assumption of a uniform prior. When we allowed each 8-point 3D shape to have equal probability of existing, we have allowed the minimization problem to be entirely dependent on our σ_{noise}, the noise of the visual perception of the image. Important assumptions of a perceiver regarding expectations of a certain shape, with certain angles and vertices as the one shown in the Necker image, were not taken into consideration.

To find a shape which much more closely represents our real cube, we need to make good assumption of the underlying dependencies of the angles and vertices within the cube. For example, we will take a prior where angles need to be of 90°, and we will use weights for assumptions of point dependencies (σ_{noise}) and assumptions for right angle measures (σ_{prior}).

By weighting the importance of edges and angles within the shape, we get a better representation of the original image from any rotational viewpoint, not just the original. We (i.e. our minds) need to tinker with setting a reasonable expectancy for different priors (σ_{prior}), so that the balance between holding the cubic structural integrity and holding the angles at a 90° position is being well-matched. Below, in Figure 2, are the results from using this method with $\sigma_{prior} = 100$ (left) and $\sigma_{prior} = 10000$ (right). The results with a higher weight for the prior better reflect the actual cube shape. For the $\sigma_{prior} = 100$ situation, the maximization algorithm puts too much emphasis in keeping right angle integrity in all its corners, neglecting the importance of shape. For

$\sigma_{prior} = 10000$, both the cube's dimensions and edges are well-reflected in the guessed shape, which is simply a translation of the real cube shape.

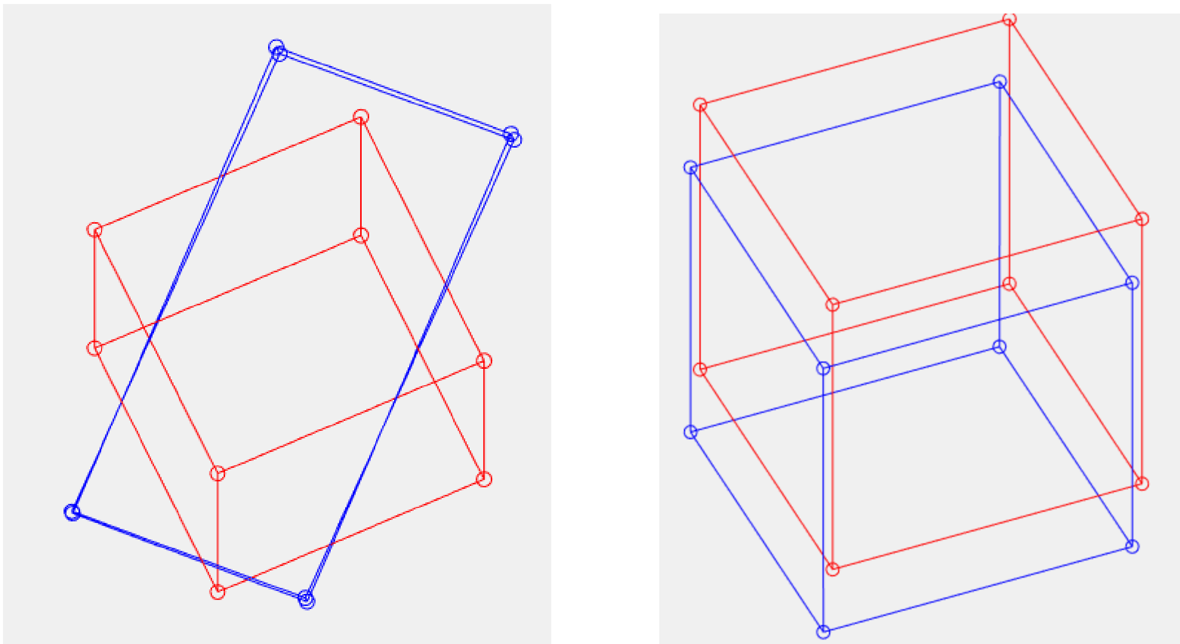

Fig. 2. *Visual representation of the 3D best fit scene* ***(blue)*** *and the real cube scene* ***(red)*** *through using a weight for priors of* $\sigma_{prior} = 100$ ***(left)*** *and* $\sigma_{prior} = 10000$ ***(right)***

- **Question 2**

For a hexagonal cylinder projection, we need to understand the angle between the base and the length (90°) and the angle between each vertex of the base (120°). By setting a prior as in the previous exercise, where all the edge angles would need to be 90°, this prior would be wrong since the hexagon is not bound by this rule. By using the previous prior again, we would be giving high probabilities to guessed shapes which resemble a parallelepipedal structure.

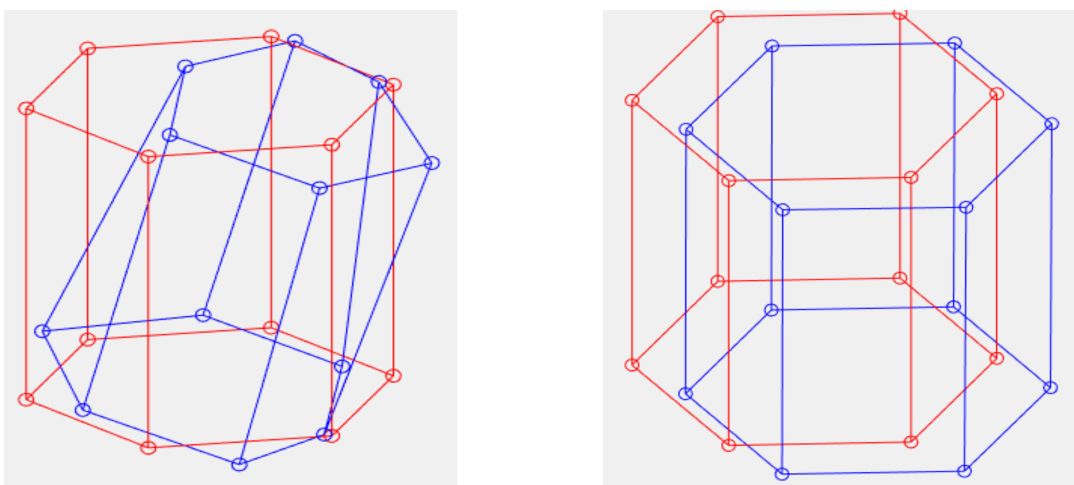

Fig. 3. *Visual representation of the 3D best fit scene* ***(blue)*** *and the real cylinder scene* ***(red)*** *through using uniform prior probability* ***(left)*** *and non-uniform with* $\sigma_{prior} = 10000$ ***(right)***

Instead, we will use a prior that minimizes differences between angles in the base and differences between angles around the junction between base and length. In this case, we can be sure to keep the structural integrity of the guessed shape to be close to that of the original shape. This would probably give a better resemblance to how our brain works. In Figure 3, we have rotated versions of the guessed shape and the real hexagon cylinder shape, on the left being shown the case in which we would have a uniform prior probability, and on the right being used the new prior mentioned above, with a weight of $\sigma_{prior} = 10000$.

- **Question 3**

When we consider understanding and finding the correct shape for a projection that we see in our retinas, we most likely will perceive such a shape from two different viewpoints, representing the images projected upon the two eyes of our anatomy.

$$P(S_1, S_2|I) = \frac{P(I|S_1, S_2) * P(S_1, S_2)}{P(I)}$$

This introduces the following consideration: we have a nuisance parameter S_2 that is affecting our model of vision: we are only interested in the presence of a certain shape (S_1), and not on the viewpoint from which we see it (S_2). Ideally, we would want to discount for additional viewpoints and have a generic viewpoint assumption that is independent of binocular disparity:

$$P(S_1|I) = \sum_{S_2} P(S_1, S_2|I) = \sum_{S_2} \frac{P(I|S_1, S_2) * P(S_1, S_2)}{P(I)}$$

Since the prior will give us similar information to having multiple viewpoints determine our perception of angles and vertices, we may assume a uniform prior distribution in order to link the posterior with the likelihood (which will be dependent on the number of viewpoints we have; in our human vision's case, 2 viewpoints):

$$P(S_1|I) = \sum_{S_2} P(I|S_1, S_2)$$