Introduction

In this project, I studied how we perceive density of objects. The first part of the project is a replication study based on the paper "*A common visual metric for approximate number and density*"[1], which investigated how humans estimate the density and number of objects. The paper found that our sense of density and sense of number are intertwined and proposed a spatial frequency (SF) response ratio model for predicting density and number. I only tried to reproduce results for the density experiment to verify the density estimation model as it is the basis for the number estimation model. My replication results are in accordance with the results of the paper, which found that human observers' abilities to judge density is influenced by the size of stimulus patches, with larger patches appearing denser. And density can be estimated using the relative response of spatial filters tuned to low and high SFs, since high SF energy is largely determined by the number of objects while low SF energy depends more on the area occupied by objects.

In the second part of the project, I design my own experiment to study how we perceive density of front/back layer of black/white circles when occlusion and left/right lateral motion is involved. The experiment results from six subjects (myself included) indicate that when both occlusion level and density level were high, we were biased to judge the front layer of circles as being denser.

I. Reproduce the density experiment and the SF response ratio model

I first reproduced the density experiment described in "*A common visual metric for approximate number and density*", and then compared to the results estimated by the proposed SF response ratio model.

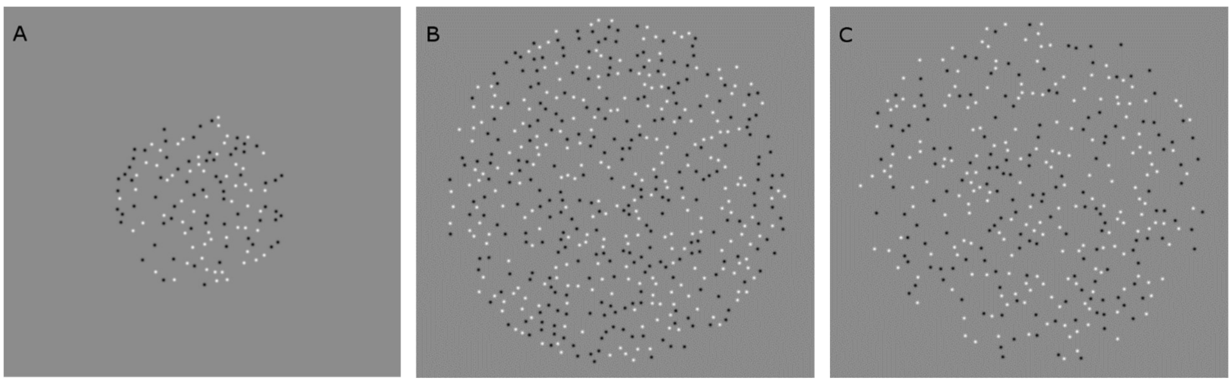
Experiment

Images were implemented with PsychoPy[2, 3] and displayed using 1536×864 Dell laptop monitor. Pairs of stimuli (a test and a reference patch) were randomly presented at $\pm 5^{o}$ left/right of central fixation for 250ms. Patches were composed of a variable number of small 2D Gaussian blobs (diameter 0.15^{o} of visual angle circle masked with 'gauss') that uniformly distributed within a circular region of specific radius. The color of each blob was randomly chosen between black and white, resulting in a mixture of black/white blobs for each stimulus patch. The distance between blobs was larger than the blob size so that there were no overlapping blobs. The task is to choose which patch was denser. A 3×3 design, independently varying the size of the test and reference patches (radii: 2^{o}, 2.8^{o}, or 4^{o} of visual angle), was used to measure observer's performance. The reference always contained 128 blobs. The density of the variable-size test was set using a method of ***constant stimuli*** and varied over a range of 50-200% in seven steps, which were 50, 63, 79, 100, 126, 159, or 200% of reference density. Each sub-session of experiment consisted of 63 trials (9 trials at seven stimulus levels) and 5 sub-sessions were performed for the whole experiments. Subject binocularly judged which of the left/right stimulus patch was denser by pressing down left/right key with eyes positioned 45cm directly in front of the screen.

Results

Sample stimulus patches from the density experiment are shown in Fig. 1 below (N is number of Gaussian blobs and Radius is radius of the stimulus patch). We can see that varying the patch size affects our perception of density. Patch B has the same physical density as patch A, but the larger patch size of B makes it appear considerably denser. In contrast, although patch C has lower physical density than patch A, it perceptually matches the density of A. Therefore, one question the paper asked is how mismatching patch size impacts on observers' ability to discriminate denser patch between a pair of stimuli. The density experiment was thus designed to quantify this impact.

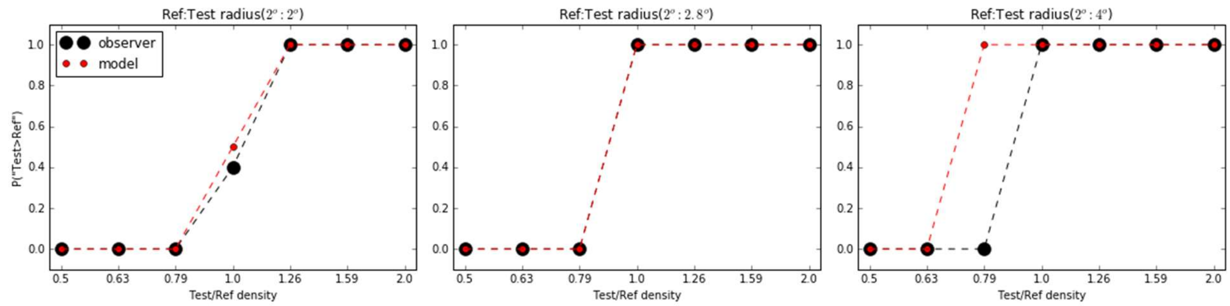
Fig. 1: Sample stimulus patches



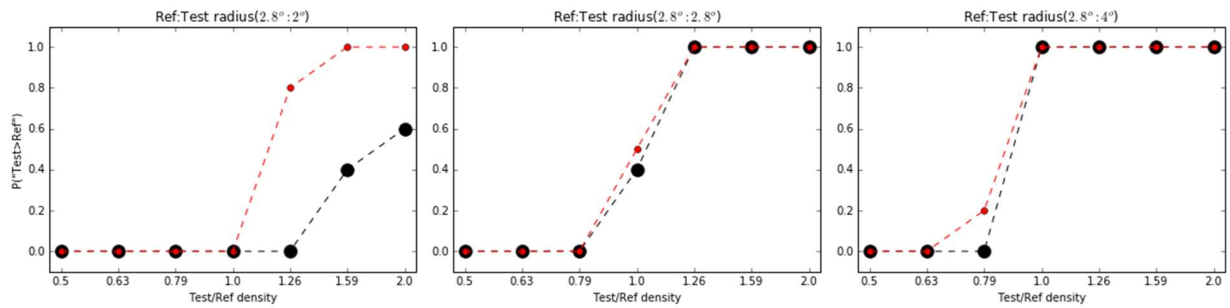
N=128 Radius=2° N=512 Radius=4° N=365 Radius=4°

I did the density experiment on myself and my responses are plotted as ***observer*** data points in Fig. 2 below by the 3×3 conditions. The graphs plot the proportion of times I categorized a test as being denser than the reference as a function of the ratio of test density and reference density. The reference patches (Ref) in each row have the same radii but the radii of test patches (Test) are increased from left to right. Within each column, the Test radii are the same while the Ref radii are increased from top to bottom. Data were fitted with cumulative Gaussian functions to derive ***bias***, which is the proportion of test density required to produce a subjective match between test and reference so that the subject is 50% likely to choose either reference or test patch as denser. The values below each condition show how observer bias varies with patch-size mismatch. We can see that observer bias is decreased from left to right in each row with increasing Test radius. Bias equals 1 when Test and Ref have the same radii and is smaller than 1 when Test radius is larger than Ref radius. Conversely, bias is bigger than 1 if Ref radius is larger. Bias also increased together with the radius of Ref from top to bottom in each column. This result is consistent with the finding of the paper (Fig. 3), which found that *for a small test patch to be perceptually matched to a variable-size reference the test's density must increase as the reference grows (small red symbols increase from left to right). In contrast, a large test paired with a small reference (leftmost blue point) can be sparser and still be perceptually matched.*

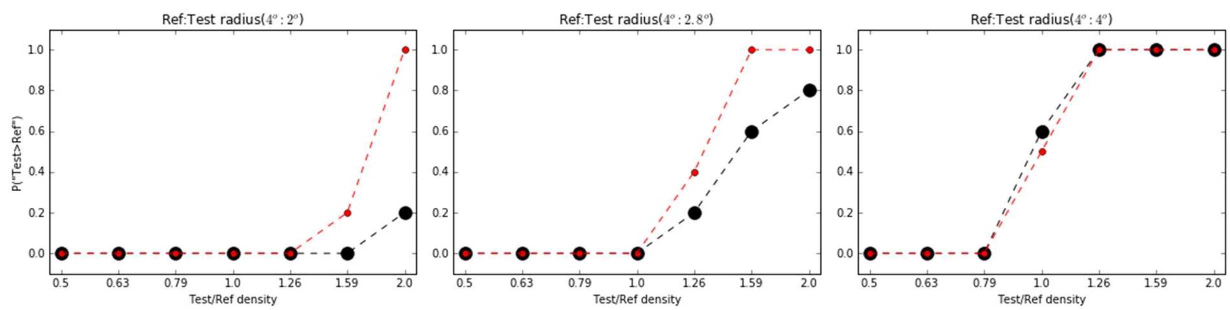
Fig. 2: Observer responses and model prediction results



Bias: observer=1 model=1 observer=0.9 model=0.9 observer=0.9 model=0.71

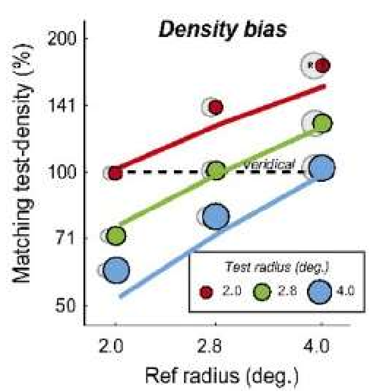


Bias: observer=1.84 model=1.26 observer=1 model=1 observer=0.9 model=0.79



Bias: observer=2 model=1.59 observer=1.61 model=1.26 observer=1 model=1

Fig. 3: Figure from the paper



Model

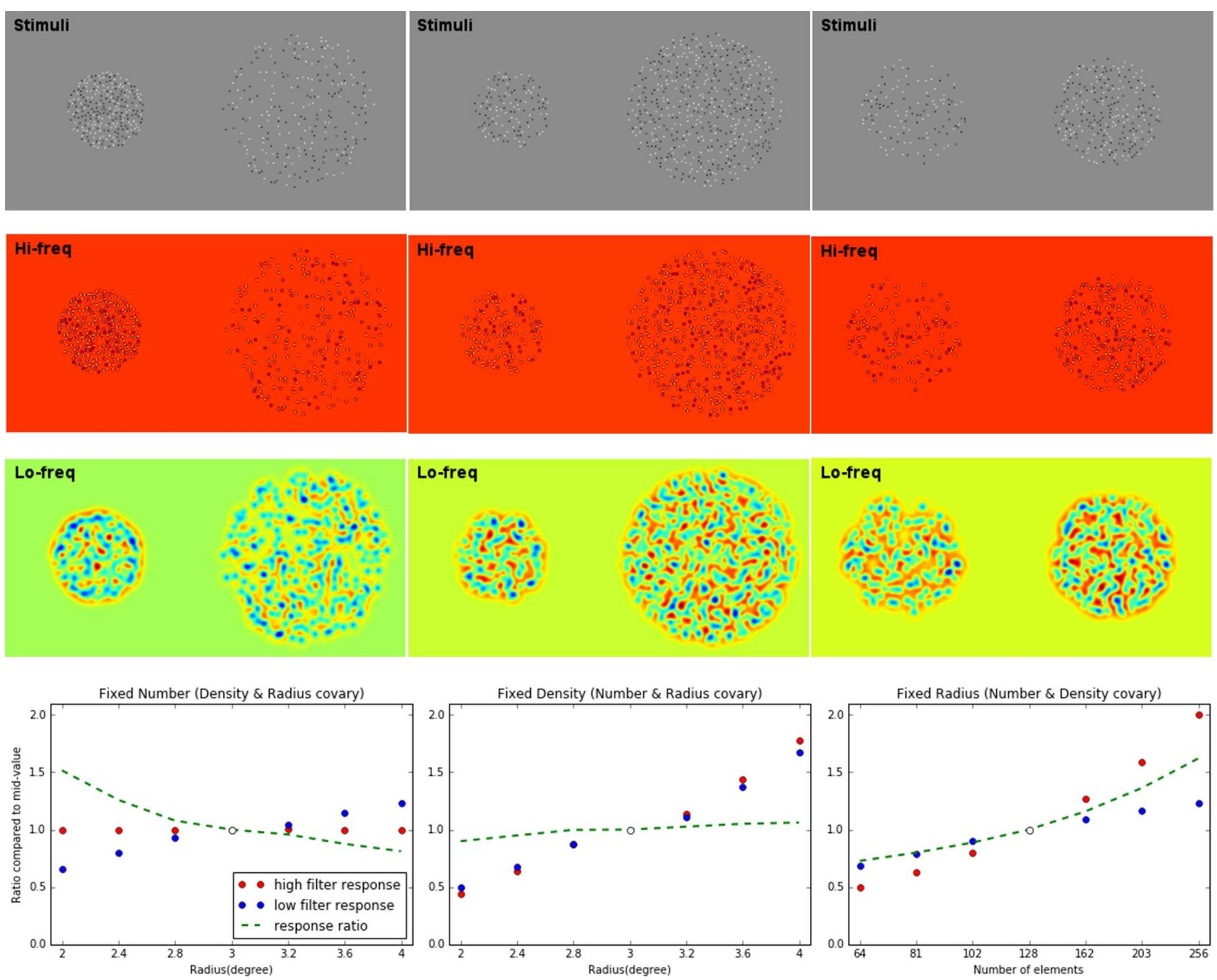
The paper proposed a SF response ratio model for estimating density. Specifically, the stimuli were full-waved rectified (color of all blobs was converted to black) and then convolved with Laplacian of Gaussian (LOG), a center-surround filter constructed from the combination of a Gaussian filter and a second derivative (Eq. 1), to estimate the filter response pooled across all image locations (Eq. 2).

$$\nabla^2 G_\sigma(x,y) = \frac{1}{\pi\sigma^4}\left(1 - \frac{x^2+y^2}{2\sigma^2}\right)\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (1)$$

$$R_\sigma = \sum_{x,y} |\nabla^2 G_\sigma \otimes |\mathrm{I}|| \quad (2)$$

Convolving with LOG of different spatial frequency produced different responses. Small filters generated isolated responses to individual elements while large filters responded to clusters of elements and their responses are limited by the stimulus patch size. These distinct responses are illustrated in Fig. 4 with example stimuli showing the effects of number, density and radius by fixing one of these parameters (from left to right) and allowing the other two to covary.

Fig. 4: Example stimuli, high-SF(Hi-freq) and low-SF(Lo-freq) versions of rectified versions of the stimuli and the effect of fixed number, density, radius on SF response ratio



Bottom panels in Fig. 4 plot the responses from the high-frequency filters (red symbols) and low-frequency filters (blue symbols) respectively, averaged across all pixels in 10 image examples. All responses were normalized relative to the response to the midrange stimulus. We can see that the red symbols (high-frequency response) closely follow the change of numbers. In contrast, although the blue symbols (low-frequency response) rise together with number, they also rise as a function of patch radius when number is fixed (bottom left panel). Based on these observations, the paper proposed using the ratio of high- and low-frequency filter responses to predict density (Eq. 3: γ_σ is Gaussian random noise).

$$C = 2^{\gamma_\sigma} \frac{R_{hi}}{R_{lo}} \qquad (3)$$

The response ratio $\frac{R_{hi}}{R_{lo}}$ is plotted as green dashed line in bottom panels of Fig. 4. Among all possible pairings of filter spatial frequencies, the filter pair that maximized the slope of straight-line fitted to the response ratio (maximizing sensitivity to density change) was selected to use in the SF response ratio model. For Eq. 1 in my experiment, σ is 1 pixel for the high-frequency filter and 11 pixels for the low-frequency filter.

The relative density of test and reference stimulus can then be estimated using Eq. 4. Since both the numerator and denominator contain the same noise term that can be cancelled out, the value of $d_{test,ref}$ thus only depends on the response ratio of test stimulus relative to reference stimulus. Test is denser than reference if $d_{test,ref}$ is greater than 1. Otherwise, reference is denser if $d_{test,ref}$ is smaller than 1.

$$d_{test,ref} = \frac{C_{test}}{C_{ref}} \qquad (4)$$

This model was used to predict relative density for stimulus image pairs that generated with the same way as in the density experiment. Simulation was carried out by collecting 5 SF response ratios for each Test stimulus in the 3×3 conditions and computing $d_{test,ref}$ values by comparing to the corresponding Ref SF response ratio that derived from 15 Ref stimuli. The model prediction results were plotted as red symbols in Fig. 2. We can see that model responded similarly to observer in the different conditions. Model biases derived from cumulative Gaussian functions fitting demonstrated similar changes as those of observer, with bias equals 1 when Test patch and Ref patch are the same size and model bias smaller than 1 when Test patch is larger. Conversely, model bias is bigger than 1 if Ref patch is larger. Within each column, the increas in radius of Ref from top to bottom was also accompanied by an increase in model bias.

Remark

The paper didn't describe how to apply the random noise term in Eq. 3. Luckily this term is cancelled out in Eq. 4 so that I can compute the density estimation model by ignoring the noise. However, the noise term needs to be incorporated into the calculation of number estimation model. Therefore, I couldn't replicate the number model without proper knowledge of how to apply the noise.

II. Density discrimination involving occlusion and lateral motion

In this part of the project, I set up my own experiment to study how we perceive density when objects are moving and occlusion between objects occurs.

Experiment

Images were implemented with PsychoPy[2, 3] and displayed using 1536×864 Dell laptop monitor. Each stimulus composed of a front layer and a back layer of circles (diameter of 0.2° of visual angle) randomly distributed across a squared field of size 10°×10° by visual angle. One layer was black and the other was white, with color randomly chosen each time a stimulus was presented. To avoid total overlap of two circles, the distance between each pair of circle centers was larger than 0.1° such that overlapping region between two circles of the same layer was no more than 50%. The front-layer circles were on top of the back-layer circles and thus having occlusion occurring. Both layers were moving left and right in opposite directions with a sinusoidal pattern. The position of a circle at each frame can be calculated using the formula:

$$P = P_0 + Amp \times sin(\frac{\pi t}{T})$$

Where P and P_0 refers to the new and old position respectively. *Amp* is the amplitude (0.03°), t is the frame number and T is the time of a half cycle movement (0.75 seconds). Each stimulus was displayed on the screen for 3 seconds (2 cycles of movement).

Each layer of the stimulus was defined by a mean density η, which is the number of circles per squared degree of visual angle, and the area A of each circle. The total number of circles in each layer was $N = \eta S$, where S is the fixed area of the square field. The value N was rounded to the nearest integer. A 2×2 design was used in the experiment. Table 1 shows the values of η, A, and N that used in the different conditions. The values of η and A were chosen so that their product

$$\lambda \equiv \eta A$$

is constant within each column and increases from left to right. The parameter λ is called the *occlusion factor*, which is the expected total area of circles per squared degree of field size. The greater the value of λ, the more likely occlusions will occur. While the area A decreases and mean intensity η increases from top to bottom in each column of Table 1, the value of λ is constant in each column. Also, η increases and A is constant from top left to bottom right, and A increases and η is constant from bottom left to top right.

Table 1: Values of mean density η, area A, and number N used in the experiment

$\lambda = 0.15$	$\lambda = 0.62$
$\eta = 2.5$, $A = 0.0615$, $N = 250$	$\eta = 5$, $A = 0.126$, $N = 500$
$\eta = 5$, $A = 0.0314$, $N = 500$	$\eta = 10$, $A = 0.0615$, $N = 1000$

The experiment assesses how well observers can estimate the density of circles in the front versus back layer. The levels of $\Delta\eta$ in the two layers were chosen separately for each mean density value η, with 9 density difference levels for each η:

$$\Delta\eta = \{0, \pm\frac{\eta}{20}, \pm\frac{2\eta}{20}, \pm\frac{3\eta}{20}, \pm\frac{4\eta}{20}\}$$

Therefore, for each stimulus the density of the two layers was $\eta \pm \frac{\Delta\eta}{20}$.

The task is to choose which layer is denser: the front or back. Subjects pressed one of two keys on the keyboard in down-up arrangement (for instance 'm' for front and 'i' for back) to input the answer. Although subjects were explicitly asked to judge density, they may use other strategy like numerosity to do the task. Since both layers of circles were bounded within a squared field of the same size, performing the task using either strategy should thus lead to the same results.

The experiment targeted a 50% probability of choosing each of the two layers as being denser (*point of subjective equality*). Stimuli were presented at the center of monitor using a method of ***constant stimuli***. The whole experiment included 6 sub-sessions, with 36 trials in each sub-session (4 different η values with 9 levels of $\Delta\eta$ for each η) and took about 30 minutes to complete. Subjects were seated so that their eye was positioned at 45 cm directly in front of the screen and viewed the stimuli monocularly through the dominant eye. The non-dominant eye was covered with an eye-patch.

Difficulties encountered

1. The display of stimuli involves lateral motion with frame rate greatly affected by CPU usage. Jumpy frame transition occurred when using my personal laptop to display motion for large number of circles with different sizes and speeds. To avoid this problem, I had to keep the stimuli as simple as possible by using uniform circle size and only one speed for each layer.

2. Although the method of constant stimuli used here has the advantage of preventing subjects from predicting the level of next stimulus by randomly presenting stimuli, it suffers the limitation of longer experiment time since each condition of stimulus was presented to the subjects with equal number of times. And longer experiment might make the subjects feel exhaustive and increase the chance of guessing answer. On the other hand, shortening experiment time by decreasing the number of trails will lead to decrease in accuracy and sensitivity of the experiment. The alternative staircase procedure can overcome this disadvantage by adaptively presenting stimuli and clustering responses around the psychometric threshold. However, staircase procedure was prohibited from being used in PsychoPy with the above design of density differences because PsychoPy's built-in staircase procedure can only change the task to easier levels in one direction while the bidirectional property of density differences in the experiment requires changing task to easier levels in two directions.

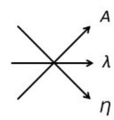
3. The stimuli are highly repetitive, which easily make subjects lose focus during the test. To overcome this, subjects were asked to rest at each interval of sub-sessions.

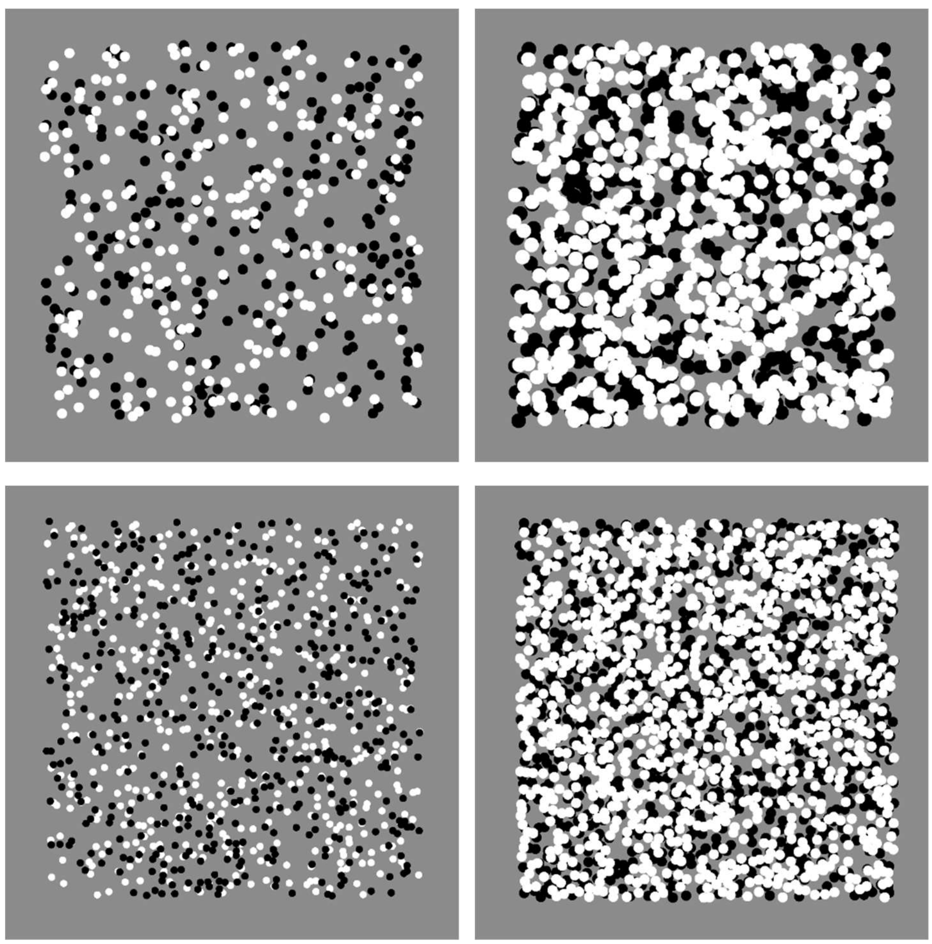
4. Some subjects reported that they perceived depth reversals on some trials. This introduced unquantifiable error to the results.

Results

Fig. 1 are example stimuli from the experiment at level $\Delta\eta = 0$. The 2×2 layout corresponds to Table 1, with arrows indicating the direction in which the variables A, λ, η increase.

Fig. 1: Example stimuli





To explore how *occlusion factor* λ varied in these different stimuli, I computed *image occupancy* fractions, which is the fraction of the pixels corresponding to visible surfaces in the front and back layer. Fig. 2 shows mean image occupancy fractions for the 2×2 conditions of Table 1 for the case of uniform density ($\Delta\eta = 0$). The error bars show the standard deviations multiplied by 10 for better illustration. Within each condition for each experiment, theses values are similar from stimulus to stimulus. For each row of Fig. 2, the means of front and back image occupancy fractions over 3000 stimuli increase with λ, and the difference between the means increases as well. Therefore, for a larger occlusion factor λ, a more negative density difference $\Delta\eta$ is needed to obtain an equal image occupancy for both the front and back layers.

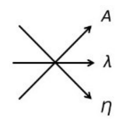
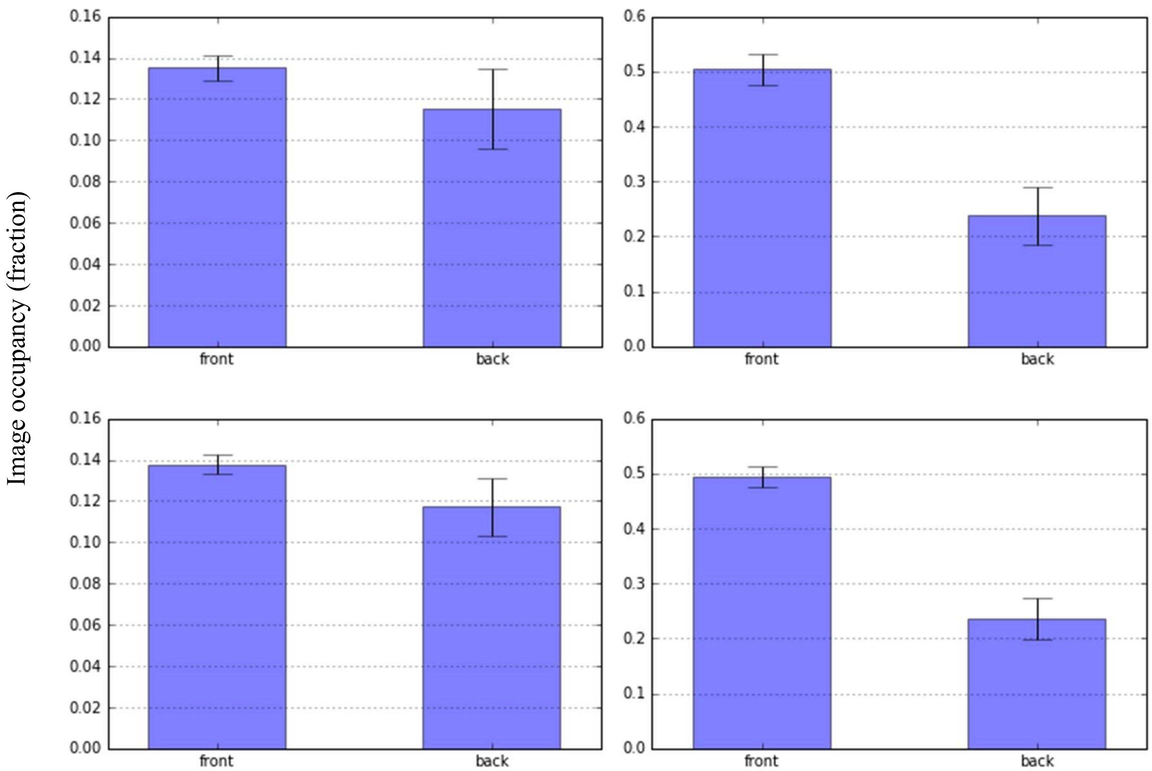


Fig. 2: Mean image occupancy



Six subjects participated in the experiment (myself included). Each subject's responses were written to Excel files by PsychoPy. From the experiment results, we can compute the fraction of trials in which the subject chose the front layer as being denser. Using the Palamedes toolbox [4] to fit a logistic function to these fractions,

$$p(x;\alpha,\beta) = \frac{1}{1+\exp\left(-\beta(x-\alpha)\right)}$$

where x is one of the nine density difference levels $\Delta\eta$, we can obtain the ***bias*** α and slope β.

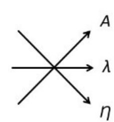
If we define the *just noticeable difference* (JND) of density in the two halves to be value of $x - \alpha$ so that x is the 75% correct threshold, then JND $\equiv \frac{ln\,(3)}{\beta}$. ***Weber fraction***, which measured the density discrimination performance, can then be derived by plugging JND into the following quantity:

$$\frac{\text{JND}}{\eta+\frac{\alpha}{2}}$$

The denominator $\eta + \frac{\alpha}{2}$ is the density of the front layer at the *point of subjective equality* (PSE) such that when $\Delta\eta = \alpha$ the front and back layer densities are $\eta + \frac{\alpha}{2}$ and $\eta - \frac{\alpha}{2}$ respectively. The density of the front layer at the PSE was used as an estimate of the observer's perceived density. Therefore, the Weber fraction is the JND between the density of the two layers relative to the perceived density.

Fig. 3 shows the mean of the ***bias*** with error bars denoting the standard deviations. It is difficult to reach a general conclusion using statistics since bias for the different conditions varied a lot between subjects. Nonetheless, with one-tailed T-Test, the bottom right condition is significantly different from the bottom left condition (p = 0.016) and has near significant difference from the top left condition (p = 0.066), suggesting that subjects were biased to the front when both the values of mean intensity η and occlusion factor λ increase.

Fig. 3: Bias α for human observers



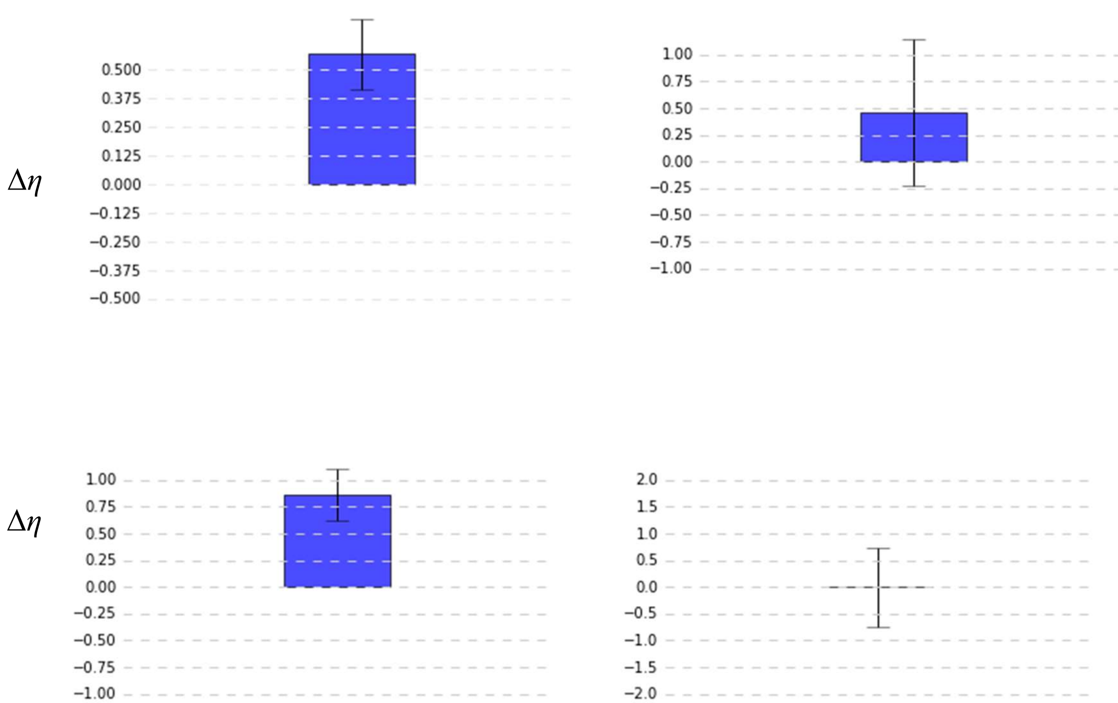
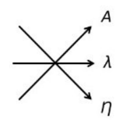
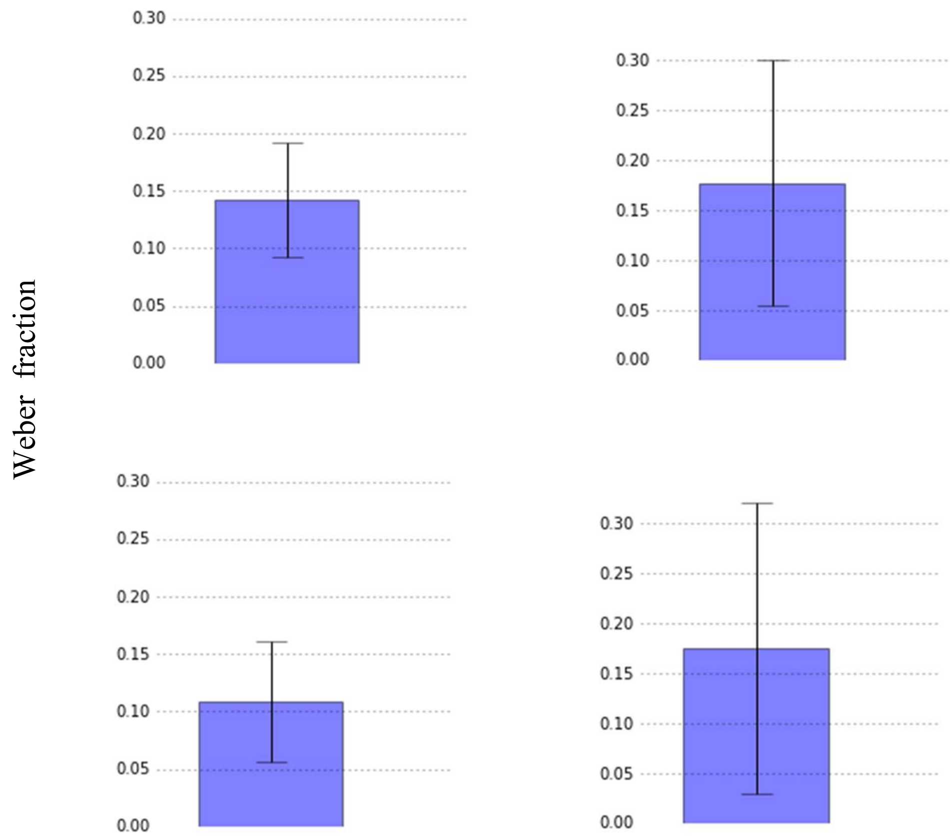


Fig. 4 shows the plots of mean ***Weber fraction***. With the great variability in bias, it is not surprising to see that Weber fraction also varied a lot between subjects in the different conditions (see the long standard deviation error bars in the following plots). Specifically, Weber fractions were found to increase with occlusion factor λ in 50% of the subjects while decrease with λ in the rest of the subjects. Thus, there are no statistically difference between Weber fractions in the 2×2 conditions.

Fig. 4: Mean Weber fractions for human observers





Conclusion

The experiment here failed to provide conclusive result for us to understand how human perceive density when circles are moving and occlusion between circles occurs. The values of bias and Weber fraction differ a lot between subjects in each condition. This may be caused by the limited number of trials in the whole experiment and the low number of participants. Depth reversal ambiguity might also exaggerate the error. More trials should be added to the experiment and more subjects are needed to explore the mechanism involved in this type of density perception.

Software used

The experiments were created using PsychoPy2 Experiment Builder (v1.90.3)[2, 3].

The data were fitted with cumulative Gaussian function or logistic function from the Palamedes toolbox[4] using Matlab.

The density estimation model, image occupancy and other data processing codes were implemented with Python2 in Spyder.

Reference

1. Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. & Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences*, *108(49), 19552-19557*.

2. Peirce, J. W. (2007) PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods*, *162(1-2):8-13*.

3. Peirce, J. W. (2009) Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2:10. doi: 10.3389/neuro.11.010.2008*

4. Prins, N., & Kingdom, F. A. (2018). Applying the model-comparison approach to test specific research hypotheses in psychophysical research using the Palamedes Toolbox. *Frontiers in psychology*, *9*.