

1 Online course

The summary of the online course is attached to the end of this report.

2 Exercises

A Choose a photo

The original photo and its gray-scaled are as followed:



Fig. 1: Original photo (left) and its grayscale version (right).

B Filter the photo with toy center-surround models

With $L \in \{10, 50, 100\}$ and $v \in \{1, 5, 10\}$, toy models were built as center-surround receptive fields and used to filter the grayscale photo.

According to the responses map shown in Fig. 2, as L and v get larger, boundaries between different regions have become blurred and smoother. A small receptive field size L can benefit seeing precisely, as more fields are used to respond to the photo in more detail. Also, a larger v enhances the effect of the ‘off-region’ of the receptive field, causes the response at a certain location to be more affected by the surrounding area, resulting in the boundary shapes of these areas becoming smoother.

C Difference-of-gaussian filter

Set $w_c = 11, w_s = 10, \sigma_c = 1, \sigma_s = 5$ as the selected parameters to build a DoG model, with kernel size $L = 6 \max\{w_c, w_s\}$. Then change them one by one to see how each

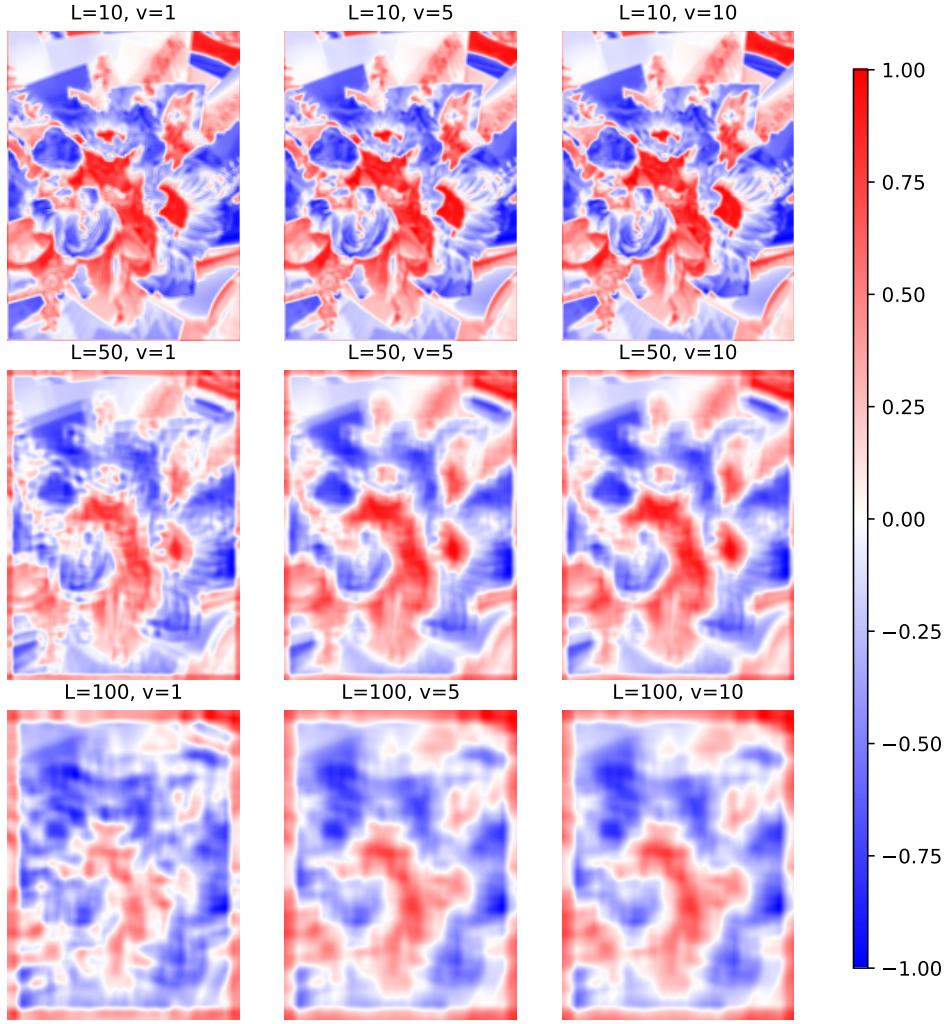


Fig. 2: Response maps of different parameters L and v

parameter affects the kernel and response map. The receptive fields of these DoG models are shown in Fig. 3.

As we can see from the figure, changing the parameters alters the shape and size of the receptive fields. For instance, increasing w_s or w_c while keeping other parameters constant, results in larger areas of on-region or off-region respectively. Similarly, increasing σ_c or σ_s leads to a broader spread of the respective regions with lower values.

Thus, w_s and w_c control the size of the receptive field's on and off regions, while σ_c and σ_s influence the spread and intensity distribution within those regions.

Then use the selected DoG model to filter the grayscale photo, and the outcome is shown in Fig. 4, which contains more details and sharp boundaries, much better than the toy model's results.

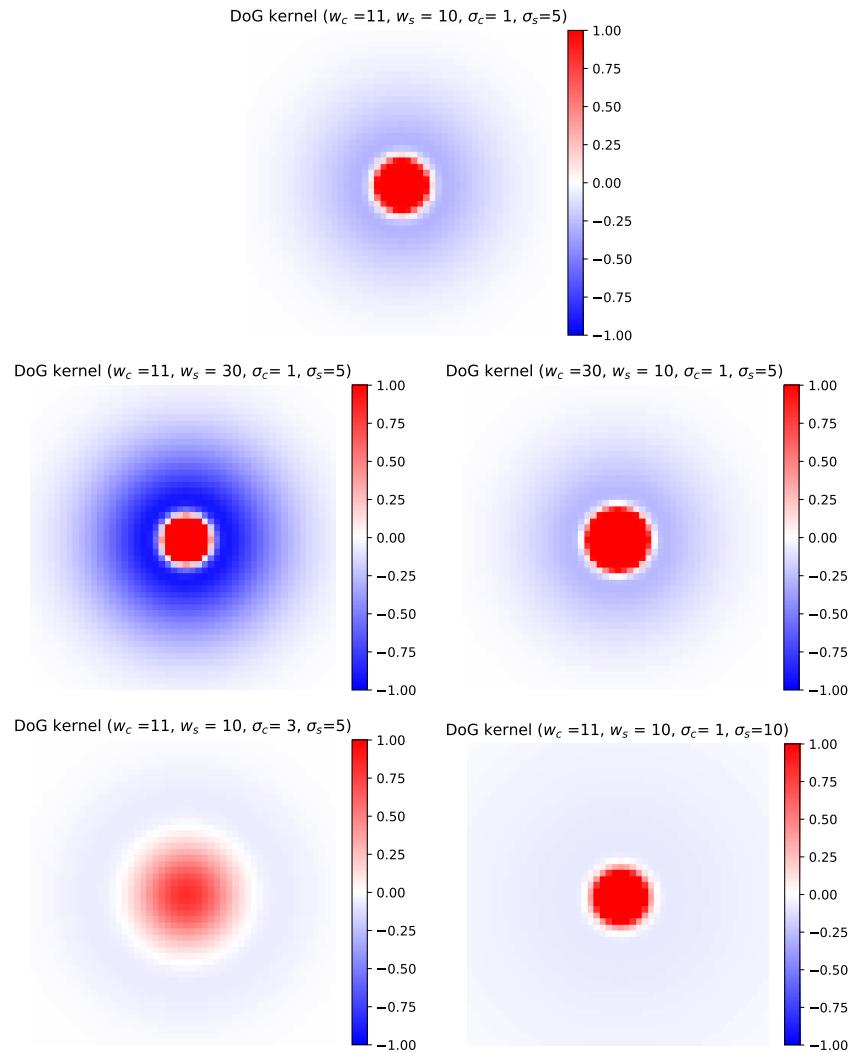


Fig. 3: Receptive field of DoG models with different parameters (Top: selected parameters, Middle: varied w_c and w_s , Bottom: varied σ_c and σ_s)

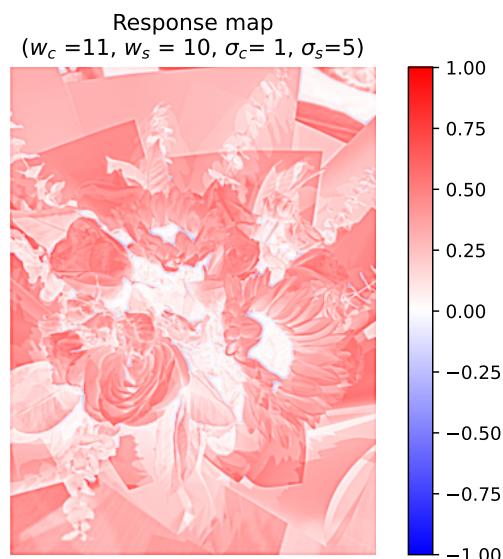


Fig. 4: Outcome of the selected DoG model

D V1's simple cell model (tuned to vertical orientation)

A vertical orientation selective neuron of V1 can be modeled by Gabor filter. Set $\sigma_x = 1$, $\sigma_y = 1.5$, $\hat{k} = 2\pi/3$, $\phi = 0$ as the selected parameters to build a Gabor filter, with kernel size $L = 6 \max\{w_c, w_s\}$. Then change them one by one to play with these parameters. The receptive fields of these Gabor filters are shown in Fig. 5.

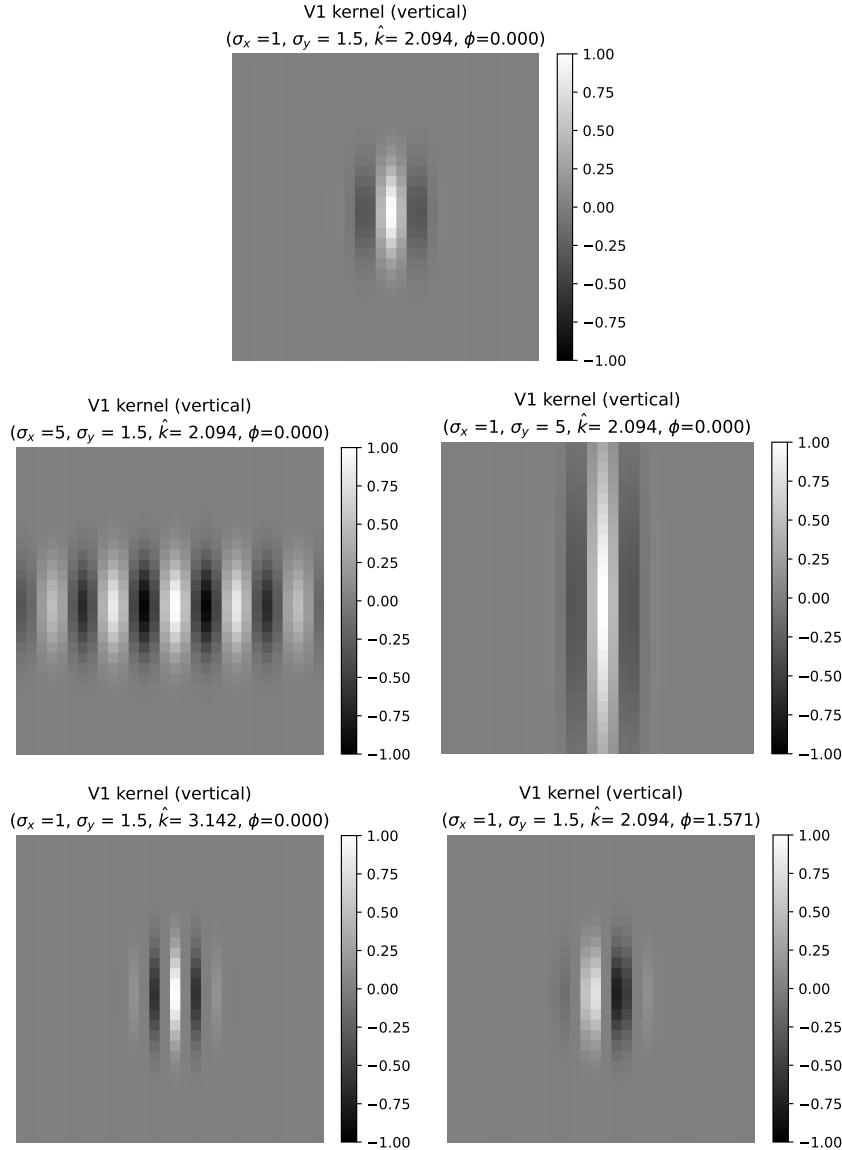


Fig. 5: Receptive field of V1 models with different parameters. (Top: selected parameters, Middle: varied σ_x and σ_y , Bottom: varied \hat{k} and ϕ)

When σ_x or σ_y increases, the receptive field becomes wider horizontally or vertically, respectively. Then we can learn that σ control the size of the gaussian function, which determines the area of the receptive field. A larger σ means a larger receptive field, leading to a broader area of influence.

When \hat{k} get larger (but not more than π), the preferred frequency of the filter increases,

making the receptive field contain more part of one cycle of the sinusoidal function. It will result in lower response in flat regions and higher response at boundaries.

When ϕ changes from 0 to $\pi/2$, the receptive field shifts from being cosine-modulated to sine-modulated. This phase shift alters the alignment of the receptive field with respect to the visual stimulus. When $\phi = \pi/2$, the receptive field prefers vertical edges.

Then use the selected V1 model to filter the grayscale photo, and the outcome is shown in Fig. 6.

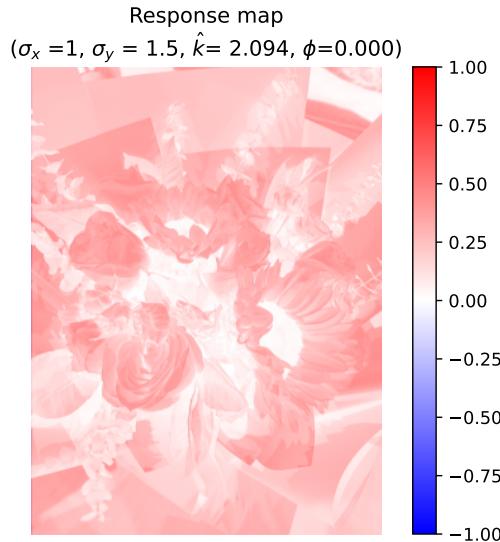


Fig. 6: Outcome of the selected V1 model (vertical)

E V1's simple cell model (tuned to horizontal orientation)

Set $\sigma_x = 1.5, \sigma_y = 1, \hat{k} = 2\pi/3, \phi = \pi/2$ as the selected parameters to build the filter. Plot the receptive fields of filters tuned with horizontal and vertical orientations with the same parameters in Fig. 7. It clearly show that they focus on different orientations.

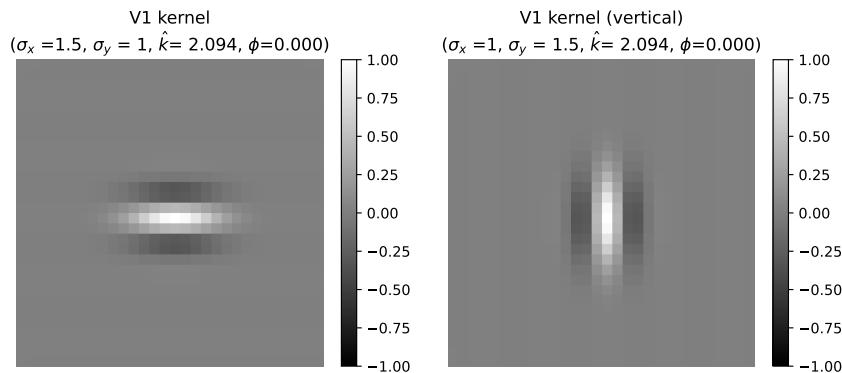


Fig. 7: Receptive fields of V1 models tuned to horizontal (left) and vertical (right) orientations with the same parameters.

Then use this horizontal orientation selective V1 model to filter the grayscale photo, and compare the outcome with the vertical model in Fig. 8. With the similar parameters, the responses of these two models do not show much difference. It may because the photo does not contain strong orientation features, making both models respond similarly.

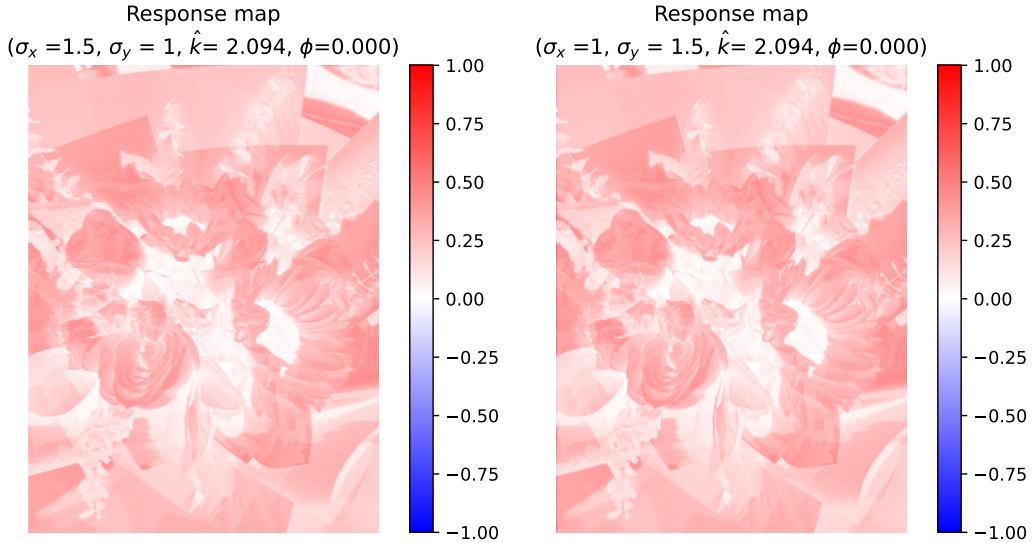


Fig. 8: Outcomes of the selected V1 model tuned to horizontal (left) and vertical (right) orientations.

Therefore, a picture of baboon with strong orientation features is used to further compare these two models, as shown in Fig. 9. The horizontal model responds strongly to the vertical edges features of the baboon, while would blur the horizontal features. The vertical model is just the opposite.

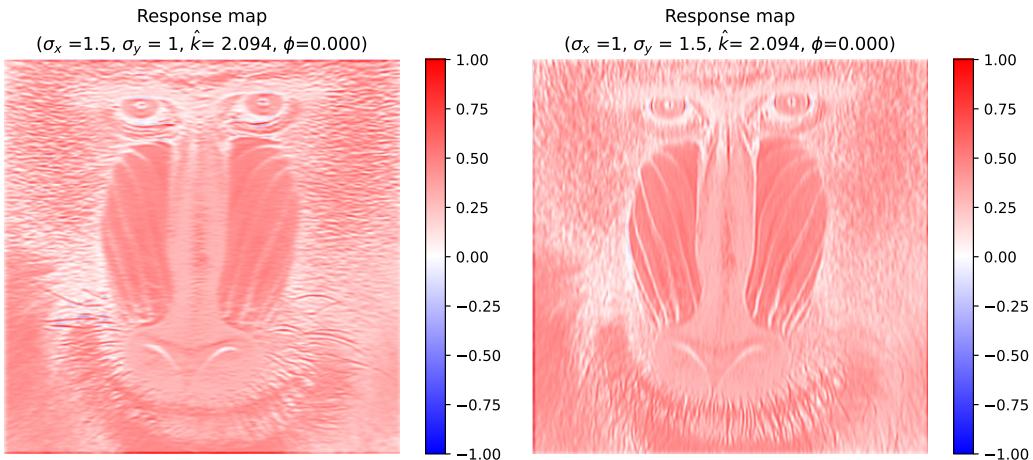


Fig. 9: Outcomes of the selected V1 model tuned to horizontal (left) and vertical (right) orientations, to a picture of baboon.

Also compare the outcomes of these two models when $\phi = \pi/2$. In this case, the models will only focus on edges. As is shown in Fig. 10, the horizontal model responds

strongly to ‘vertical’ edges, while the vertical model responds to ‘horizontal’ edges, which is worth noticing.

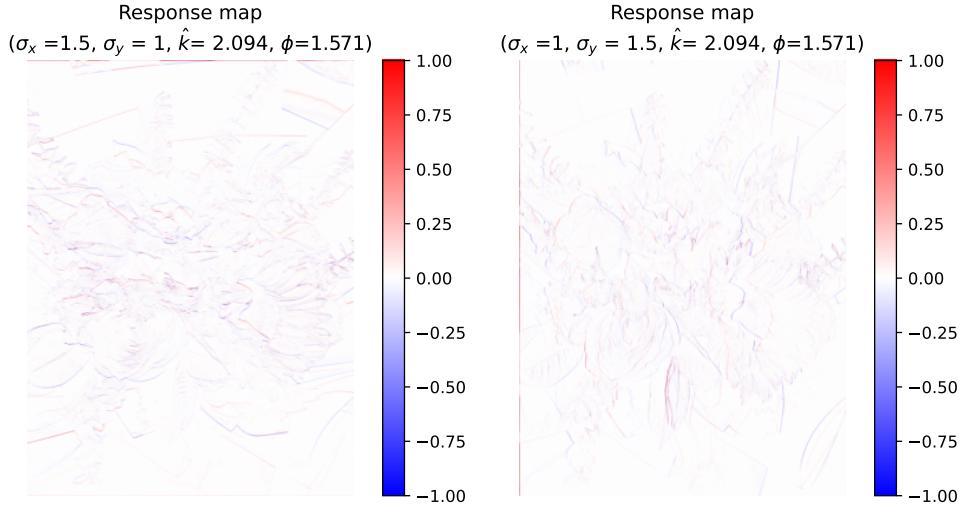


Fig. 10: Outcomes of the selected V1 model tuned to horizontal (left) and vertical (right) orientations (with $\phi = \pi/2$).

F Contrast sensitivity function of DoG receptive field

Using the DoG model with $w_c = 11, w_s = 10, \sigma_c = 1, \sigma_s = 5$ to calculate the contrast sensitivity function, response of the model as a function of the spatial frequency k . It is shown in Fig. 11, with a peak at $k = 0.52$. Here the kernel size L is modified to get more k points, but will not affect the CSF result, as it won’t change the structure of the receptive field.

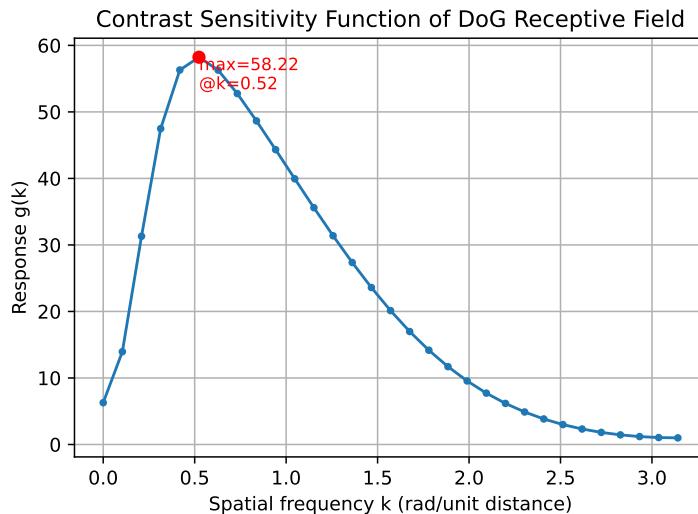


Fig. 11: Contrast Sensitivity Function of DoG Receptive Field

G Contrast sensitivity function of V1 receptive field

Using the V1 model with $\sigma_x = 1, \sigma_y = 1.5, \hat{k} = 2\pi/3, \phi = 0$ to calculate the contrast sensitivity function, as shown in Fig. 12 (top) with a peak at $k_p \approx 1.86 (\neq \hat{k} \approx 2.09)$. The value of L is also modified here to get more k points, but will not affect the CSF result.

Here change parameters σ ($\sigma_x = 2, \sigma_y = 3$) and calculate the contrast sensitivity function, as shown in Fig. 12 (bottom) with a peak at $k_p \approx 2.09 = \hat{k}$. The reason why the CSF has a peak at \hat{k} is simply that it is the frequency of the Gabor filter itself.

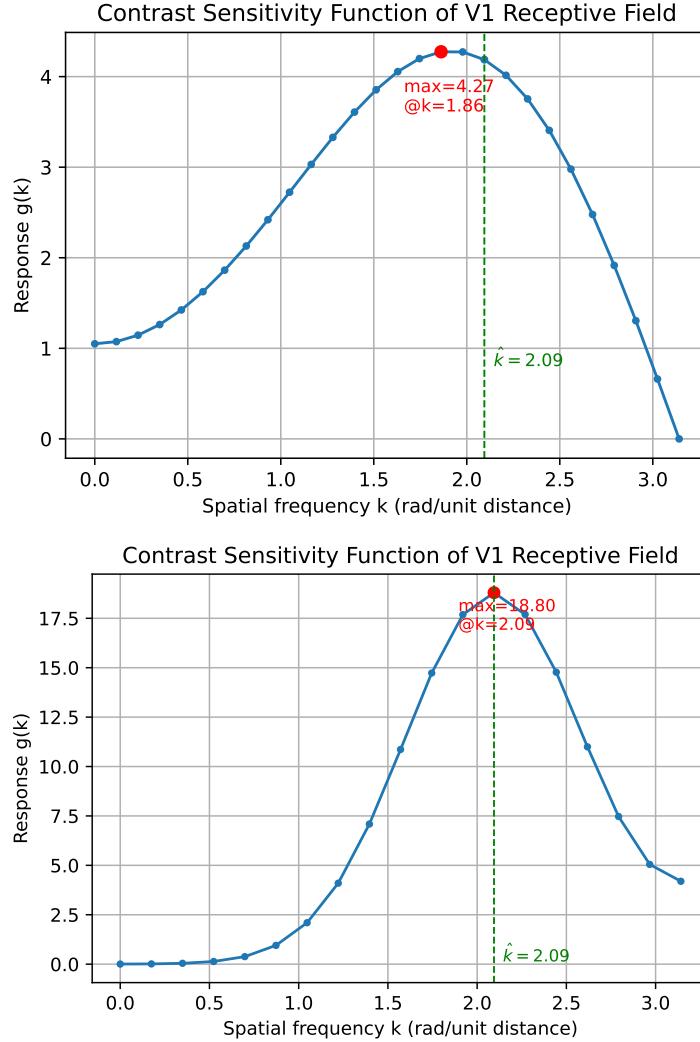


Fig. 12: Contrast Sensitivity Function of V1 Receptive Field (Top: $\sigma_x = 1, \sigma_y = 1.5, \hat{k} = 2\pi/3, \phi = 0$; Bottom: $\sigma_x = 2, \sigma_y = 3, \hat{k} = 2\pi/3, \phi = 0$)

But for Fig. 12 (top), why it has $k_p \neq \hat{k}$? Here calculate the CSF under different values of σ_x , while $\sigma_y = 1.5$, just to play around. As shown in Fig. 13, when σ_x get larger, the k_p where CSF has a peak get closer to \hat{k} . The reason behind this phenomenon may be that a larger σ_x makes the receptive field cover more cycles of the sinusoidal function along the x-axis, thus making the filter more frequency-selective and aligning its

peak sensitivity closer to its inherent frequency \hat{k} .

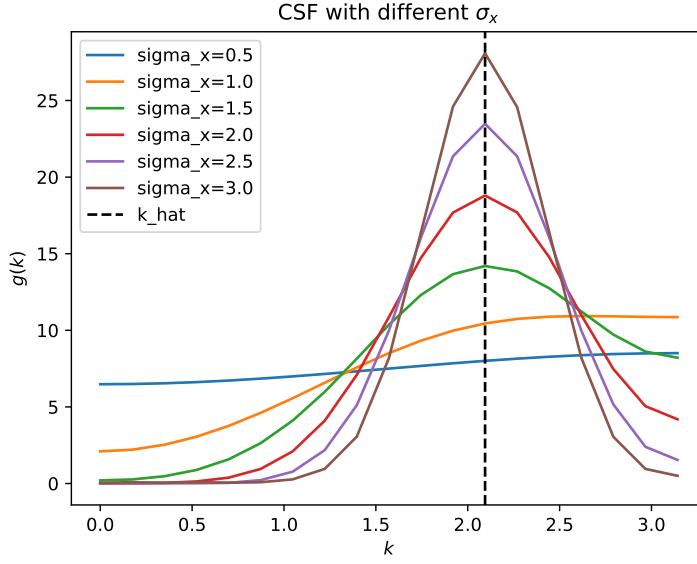


Fig. 13: CSF under different σ_x

Compared to the DoG model, the V1 model shows a higher peak frequency and a sharper decline in sensitivity at higher frequencies. Also, V1 model shows a stronger preference for high-frequency components than the DoG model, which may reflect the more complex processing capabilities of V1 neurons compared to the simpler center-surround structure of DoG receptive fields.

H Power spectrum of photos

Replace the kernel by photos, extend them according to the method provided in the instruction, and calculate the power spectrum ($|S_K^2|$ as a function of $|k|$) of each photo. Then average the results of seven photos and plot them together with $1/|k|^2$ for comparison, as shown in Fig. 14.

Here I collect seven photos with size 256×256 pixels as the input signals.

The average power spectrum of the seven photos exhibits a trend that is similar to the $1/|k|^2$ line, i.e. $\langle |S_K|^2 \rangle \propto 1/|k|^2$. This suggests that the power spectrum of natural images tends to follow a $1/|k|^2$ distribution, indicating that lower spatial frequencies (smaller $|k|$ values) generally contain more power compared to higher spatial frequencies.

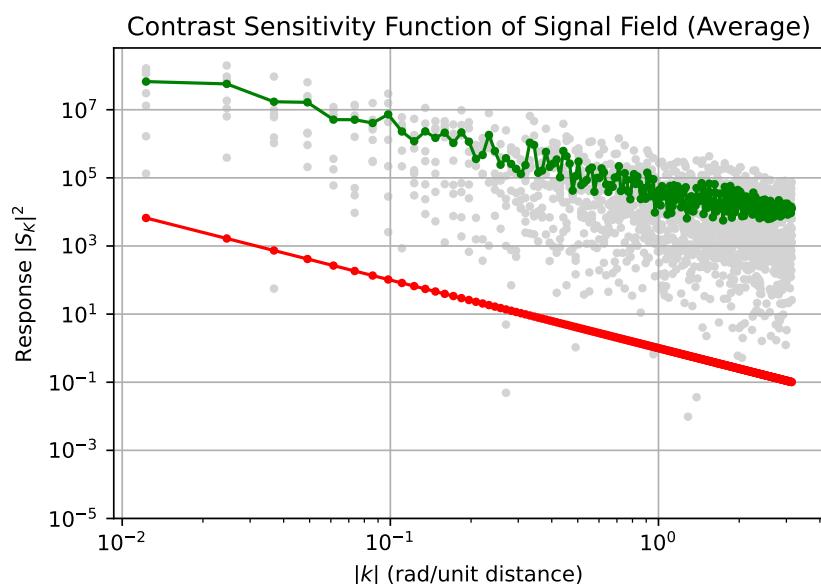


Fig. 14: Power $|S_K|^2$ versus $|k|$ for input $S(x, y)$ (Green line: the average of $|S_K|^2$ over seven photos; Red line: $1/|k|^2$)