

1 Online course

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

2 Exercises

2.1 Part 1: Introductory math for computational vision and initial steps of efficient coding

A Choose a pair of stereo images

Here I selected two photos shown in Fig. 1 where I cut down the size of the original images to 200×200 pixels from 1000×1000 for easier computation and visualization. The luminance channel of the images is also shown for reference, where the pixel values are mapped to the grayscale intensity among $[0, 255]$.

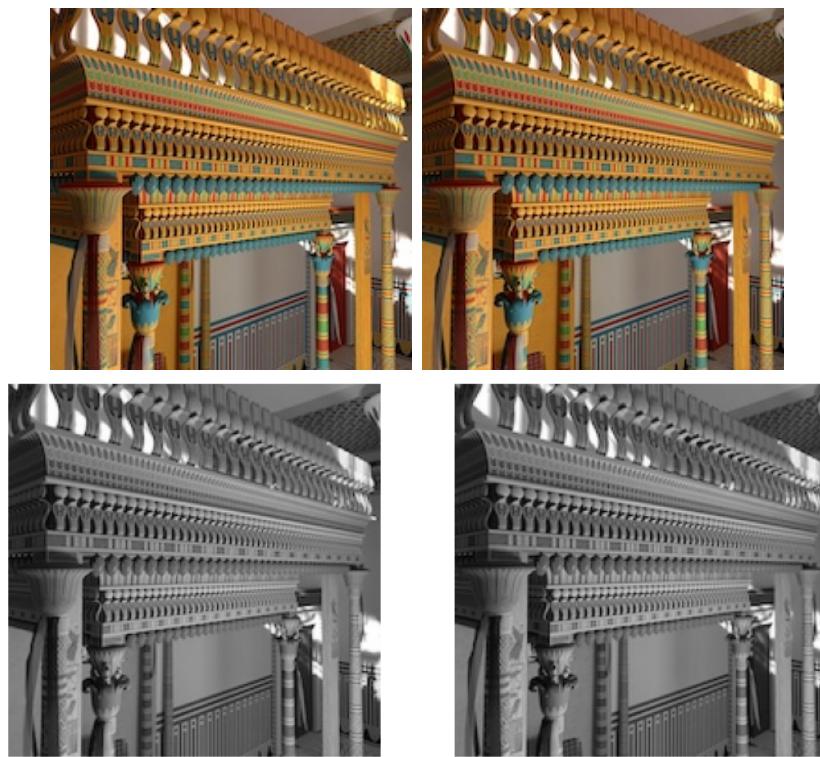


Fig. 1: Selected photos (Top: the original images; Bottom: luminance channel of the image).

B Normalize and calculate the probability $P(S)$

The probability density distributions of both $P(S_L)$ and $P(S_R)$ are shown in Fig. 2. The two pictures are quite similar as there is only slight difference between two distributions.

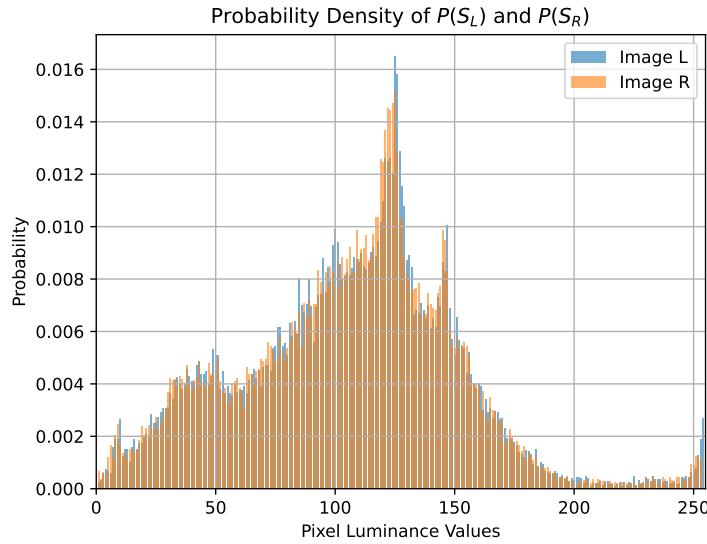


Fig. 2: Probability density distributions of $P(S_L)$ and $P(S_R)$.

C Pixel entropy, joint probability, joint entropy, mutual information, and redundancy

The pixel entropies of the two images are computed as:

$$H(S_1) = 7.4300 \text{ bits}, H(S_2) = 7.4194 \text{ bits}$$

Then the joint probability distribution is computed as shown in Fig. 3.

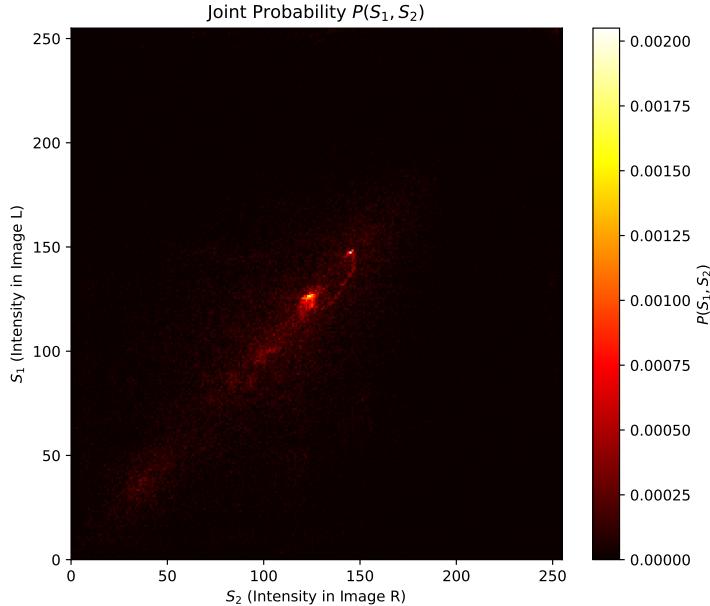


Fig. 3: Joint probability distribution $P(S_L, S_R)$.

The joint entropy and mutual information are computed as:

$$H(S_1, S_2) = 13.5912 \text{ bits}$$

$$I(S_1; S_2) = 1.2582 \text{ bits}$$

The mutual information computed here refers to the mutual information between the two images *per pixels*, but not the whole images, which is interesting to note and may result in the respective small value.

Here the redundancy is computed as:

$$R(S_1; S_2) = 0.0926 = 9.26\%$$

D Different n

Here different n values ($n = 1, 2, 3, \dots, 8$) are tested for scaling the images. The results are shown in Fig. 4 and Fig. 5.

As shown in Fig. 4, with the increase of n , the entropies all increase, which is reasonable as larger n means more possible pixel values, thus more uncertainty and more information. One interesting observation is that the mutual information increases with n more slightly respectively, and it remains a low value compared to the entropies. It indicates that even with larger n , the two images still share limited information per pixel.

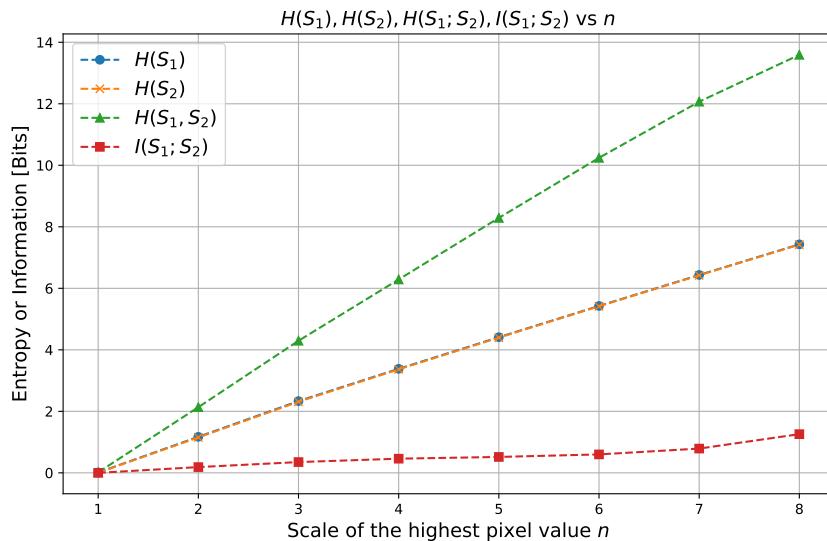


Fig. 4: Entropy, Joint Entropy, and Mutual Information vs n .

According to Fig. 5, the redundancy rates show a ‘U’ curve with the increase of n . It first shows a significantly low value at $n = 1$, which may because the information is abstracted too much at this stage and the selected image collapse into mostly one value, leading to very small entropy and further less shared information between the two images

than the joint information. Then it decreases progressively from $n = 2$ to $n = 6$, as the pixel values become more precise and contain more detailed difference pixels-wise, as expected.

After that it increases to $n = 8$, which is quite interesting. It may be because the image size doesn't support sampling with larger n values. Given the definition that $R(S_1, S_2) = [H(S_1) + H(S_2)]/H(S_1, S_2) - 1$, where $H(S_i) = -\sum_{S_i} P(S_i) \log P(S_i)$ and $H(S_1, S_2) = -\sum_{S_1, S_2} P(S_1, S_2) \log P(S_1, S_2)$, it is possible that when n increase over 6, the joint probability distribution is thinned by added values with very low probability, and this thinning effect is stronger than that to edge probability $P(S_1)$ and $P(S_2)$. Then as a result of this '*oversampling*', the joint entropy $H(S_1, S_2)$ decrease more than the summation of two pixel entropies $H(S_1) + H(S_2)$.

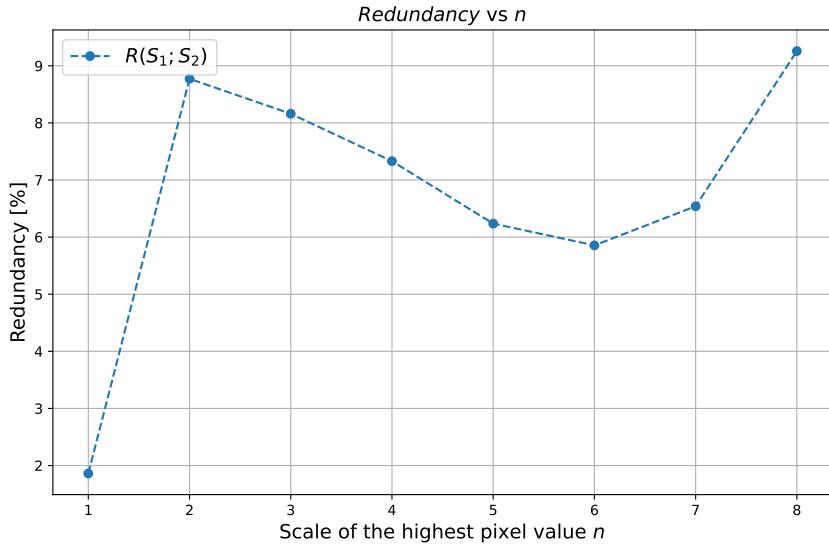


Fig. 5: Redundancy vs n .

Also, the scaled images by different n values ($n = 1, 2, 3$) are shown in Fig. 6. Here just three n values are shown for illustration purpose, as the images with larger n values are quite similar to the original images. It can be seen that the image scaled with $n = 1$ loses most of the details, while the image scaled with $n = 3$ already looks quite similar to the original images. So larger n values are preferred for better representation of the images.

E Correlation matrix \mathbf{R}^S

The correlation matrix of the shifted images is computed as:

$$\mathbf{R}^S \approx \begin{pmatrix} 2057.30 & 1140.37 \\ 1140.37 & 1990.28 \end{pmatrix}$$

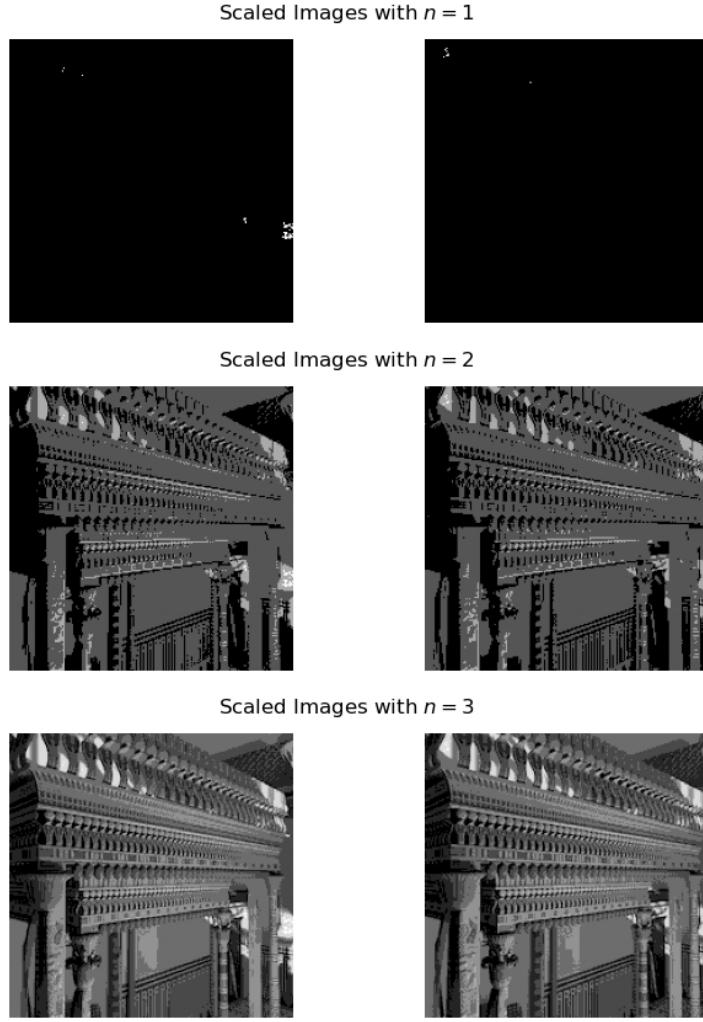


Fig. 6: Images scaled with different n (from top to bottom: $n = 1, 2, 3$).

where $\mathbf{R}_{ij}^S = \langle S_i, S_j \rangle$, $i, j = 1, 2$.

Entities in this matrix refer to *covariance*, instead of correlation coefficient, that's why they are obviously larger than 1.

F Scatter plot of S_L versus S_R

The scatter plot of S_L versus S_R is shown in Fig. 7 (left). It can be seen that there seems to be a positive correlation between the two images, implied by the oval area in the center. There are also some outliers far away from the center area, which may be caused by other special features in the images that are not aligned between two images, make it not so similar to the nice one show in 'Figure 1' of the reference paper.

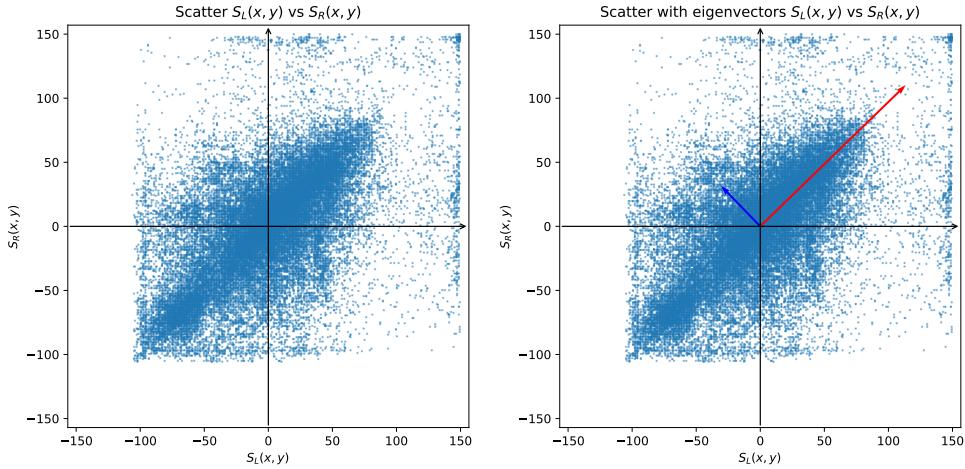


Fig. 7: Scatter plot of S_L versus S_R with eigenvectors. (Left: without eigenvectors; Right: with eigenvectors)

G Eigenvalues and eigenvectors of \mathbf{R}^S

The eigenvalues and eigenvectors of the correlation matrix \mathbf{R}^S are computed as:

$$\lambda_1 \approx 3164.65, \lambda_2 \approx 882.93$$

$$\mathbf{v}_1 \approx \begin{pmatrix} 0.7174 \\ 0.6966 \end{pmatrix}, \mathbf{v}_2 \approx \begin{pmatrix} -0.6966 \\ 0.7174 \end{pmatrix}$$

The scatter plot with eigenvectors is shown in Fig. 7(right). It can be seen that the first eigenvector \mathbf{v}_1 points to the direction with the largest variance (nearly $\pi/4$), while the second eigenvector \mathbf{v}_2 is orthogonal to the first one and points to the direction with the second largest variance (nearly $3\pi/4$). The two eigenvectors are orthogonal to each other, as expected.

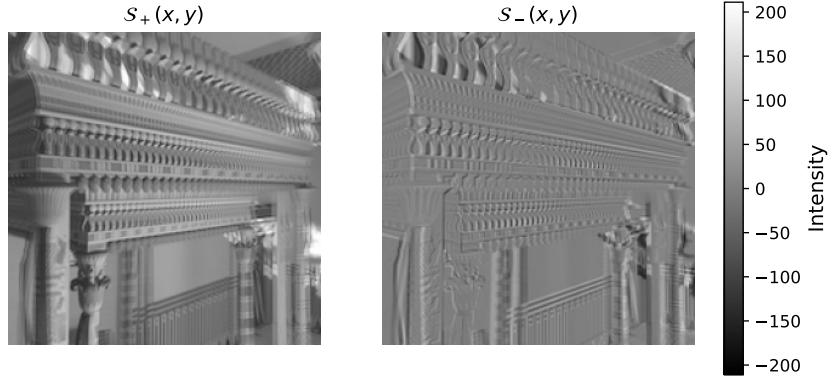
H The \mathcal{S}_+ and \mathcal{S}_- decorrelated from S_L and S_R

The transformed pictures \mathcal{S}_+ and \mathcal{S}_- are shown in Fig. 8. It can be seen that \mathcal{S}_+ performs like an average of the two images, while \mathcal{S}_- looks like the difference between the two images.

The correlation matrix of \mathcal{S}_+ and \mathcal{S}_- is computed as:

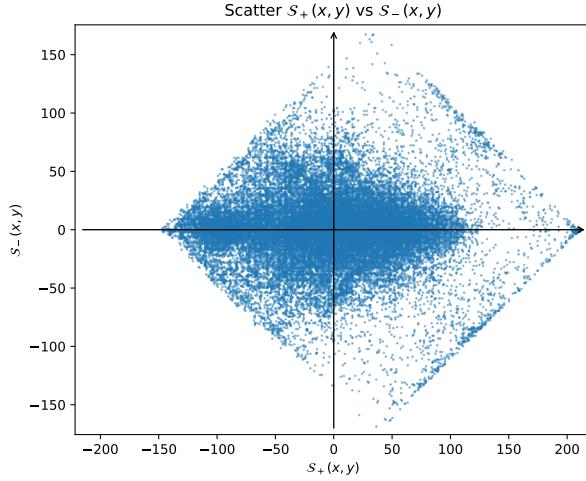
$$\mathbf{R}^S \approx \begin{pmatrix} 3164.16 & -33.51 \\ -33.51 & 883.42 \end{pmatrix}$$

It can be seen that the off-diagonal terms are much smaller than the diagonal terms, i.e., $R_{+-}^S = R_{-+}^S \approx 0$. It means that \mathcal{S}_+ and \mathcal{S}_- are nearly uncorrelated. Also, $R_{++}^S >$

Fig. 8: The \mathcal{S}_+ and \mathcal{S}_- images.

$R_{--}^{\mathcal{S}}$, which means that the \mathcal{S}_+ image has larger variance than the \mathcal{S}_- image, consistent with the eigenvalues computed before.

The scatter plot of \mathcal{S}_+ versus \mathcal{S}_- is shown in Fig. 9. It can be seen that the points are more aligned with the \mathcal{S}_+ axis, indicating larger variance along this direction, while the variance along the \mathcal{S}_- axis is smaller. This is consistent with the correlation matrix computed above.

Fig. 9: Scatter plot of \mathcal{S}_+ versus \mathcal{S}_- .

I The \mathcal{O}_+ and \mathcal{O}_- after gain control

Here we let the gain values be $g_k = \frac{1}{\sqrt{R_{kk}^{\mathcal{S}}}}$ ($k = \pm$), which means to whiten the signals to strengthen the weak contrast signals \mathcal{S}_- and suppress the strong contrast signals \mathcal{S}_+ .

The output images \mathcal{O}_+ and \mathcal{O}_- after gain control are shown in Fig. 10. It can be seen that the \mathcal{O}_+ image becomes brighter overall respective to \mathcal{O}_- , compared to the \mathcal{S}_+ image in Fig. 8. This is because the gain control suppresses the signal with larger variance (\mathcal{S}_+) and enhances the signal with smaller variance (\mathcal{S}_-), to make two images have roughly

the same variance. It can also be implied by the correlation matrix R^O computed below.

$$R^O \approx \begin{pmatrix} 1 & -0.02 \\ -0.02 & 1 \end{pmatrix}$$

It shows that the variances of \mathcal{O}_+ and \mathcal{O}_- are both normalized to 1, and the off-diagonal terms are close to 0, indicating that they are nearly uncorrelated. The gain control works as *whitening*.

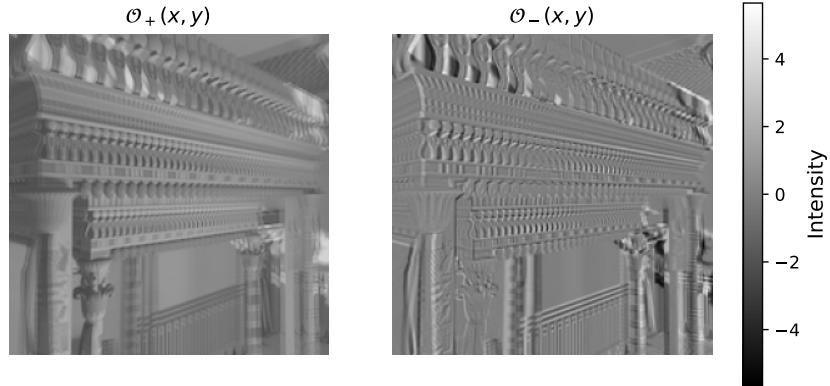


Fig. 10: Output images O_+ and O_- .

Also, the scatter plot of \mathcal{O}_+ versus \mathcal{O}_- is shown in Fig. 11 (Left). It can be seen that the central area is more circularly distributed compared to Fig. 9, indicating that the variances along both axes are more similar now, consistent with the correlation matrix R^O computed above.

J New images O_1 , O_2 reconstructed from \mathcal{O}_+ and \mathcal{O}_-

The new images O_1 and O_2 reconstructed from \mathcal{O}_+ and \mathcal{O}_- are shown in Fig. 12. It can be seen that they look quite similar to the original images shown in Fig. 1, but with some differences in brightness and contrast.

The scatter plot of O_1 versus O_2 is shown in Fig. 11 (Right). From the scatter plot, it is hard to tell whether there is not correlation between the two images, as the points still somehow tend to align along a diagonal direction. Then again, the correlation matrix R^O is computed as:

$$R^O \approx \begin{pmatrix} 0.98 & 0 \\ 0 & 1.02 \end{pmatrix}$$

It can be seen that the off-diagonal terms are nearly 0, indicating that O_1 and O_2 are indeed uncorrelated. Also, the variances of both images are close to 1.

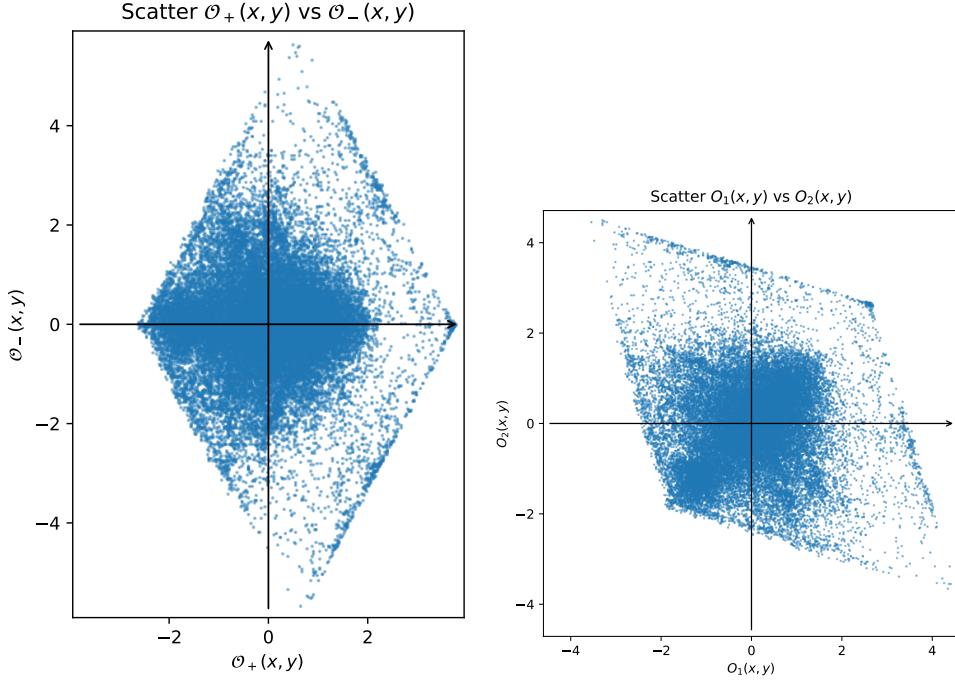


Fig. 11: Scatter plot of \mathcal{O}_+ versus \mathcal{O}_- (Left) and O_1 versus O_2 (Right).

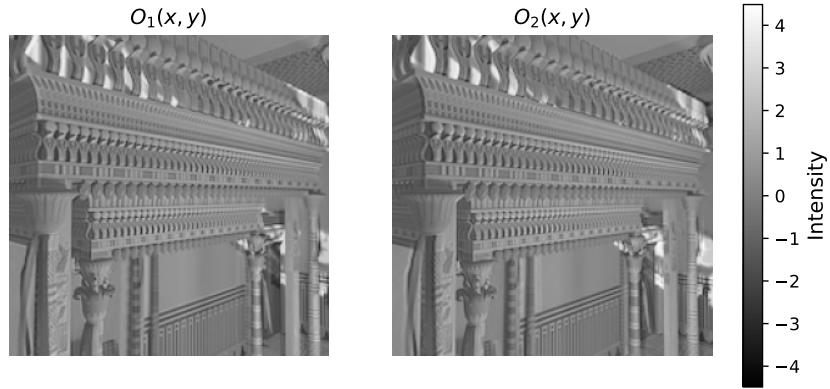


Fig. 12: Output images O_1 and O_2 .

K Add some noise!

Here random noise with range of $[-N_{max}, N_{max}]$ are added to the S_L and S_R images, where $N_{max} = \sqrt{(\mathbf{R}_{11}^S + \mathbf{R}_{22}^S)/2}$. The new images with noise added are shown in Fig. 13.

The probability density distributions and joint probability distribution, shown in Fig. 14, shows that the noise makes the distributions more spread out and close to uniform distribution. The correlation matrix of the noisy images is computed as:

$$\mathbf{R}^{S'} \approx \begin{pmatrix} 878.15 & 259.95 \\ 259.95 & 969.37 \end{pmatrix}$$

The scatter plot of noisy images S_L versus S_R with eigenvectors is shown in Fig. 15.

It can be seen that the points are also more spread out due to the noise. The eigenvectors are also plotted, which are similar to those computed before in direction, but the values decrease a bit due to the noise.

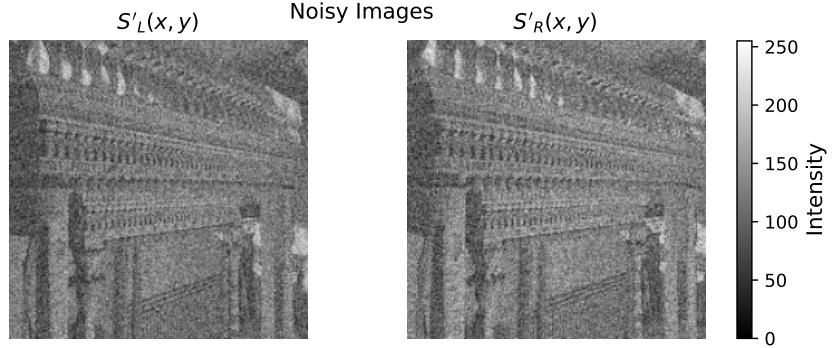


Fig. 13: Noisy images S'_L and S'_R .

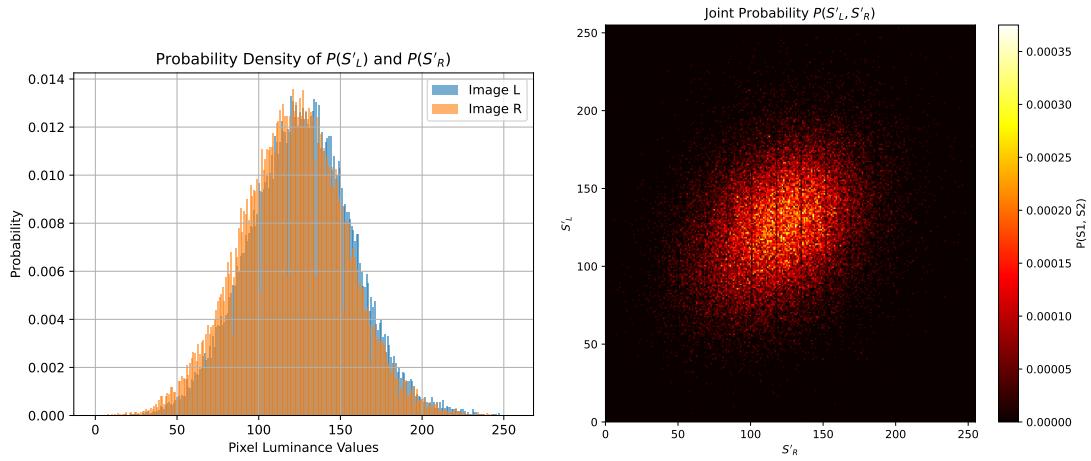


Fig. 14: Probability density and joint probability distributions of noisy images S_L and S_R .

Then get the \mathcal{S}'_+ and \mathcal{S}'_- images and the scatter plot, shown in Fig. 16 and Fig. 17(left) respectively. The correlation matrix of \mathcal{S}'_+ and \mathcal{S}'_- is computed as below. The off-diagonal terms are still small compared to the diagonal terms, indicating that \mathcal{S}'_+ and \mathcal{S}'_- are still nearly uncorrelated. However, the variances of both images decrease due to the noise.

$$R^{\mathcal{S}'} \approx \begin{pmatrix} 1183.71 & 45.61 \\ 45.61 & 663.81 \end{pmatrix}$$

Apply the same gain control to get \mathcal{O}'_+ and \mathcal{O}'_- images, shown in Fig. 18. The scatter plot is also shown in Fig. 17(right). The correlation matrix of \mathcal{O}'_+ and \mathcal{O}'_- is computed

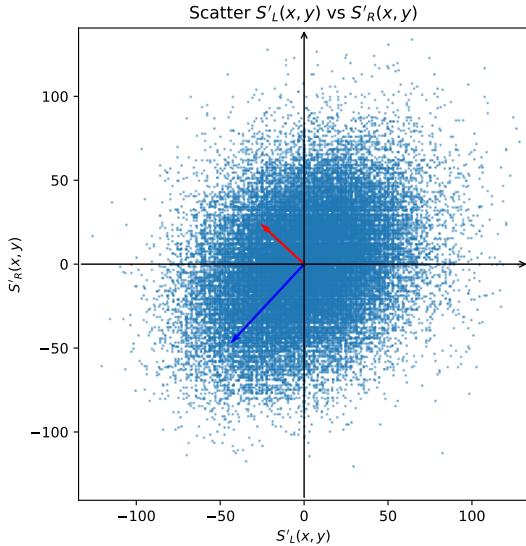


Fig. 15: Scatter plot and eigenvectors of noisy images S'_L and S'_R .

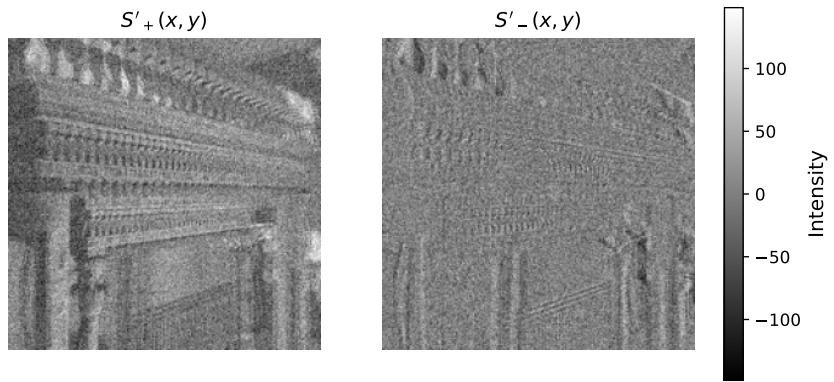


Fig. 16: Noisy images \mathcal{S}'_+ and \mathcal{S}'_- .

as:

$$R^{O'} \approx \begin{pmatrix} 1 & 0.05 \\ 0.05 & 1 \end{pmatrix}$$

Then reconstruct the noisy images O'_1 and O'_2 , shown in Fig. 19. The correlation matrix of O'_1 and O'_2 is computed as below. The off-diagonal terms are zero and the variances of both images are close to 1, similar to the $R^{O'}$

$$R^{O'} \approx \begin{pmatrix} 1.05 & 0 \\ 0 & 0.95 \end{pmatrix}$$

The result so far is quite similar to the noise-free case. The noise affects the variances of the intermediate images \mathcal{S}'_+ and \mathcal{S}'_- , but the final output images O'_1 and O'_2 are still uncorrelated with variances close to 1 after gain control and reconstruction. As shown below, the off-diagonal terms of the correlation matrix of the new noisy images

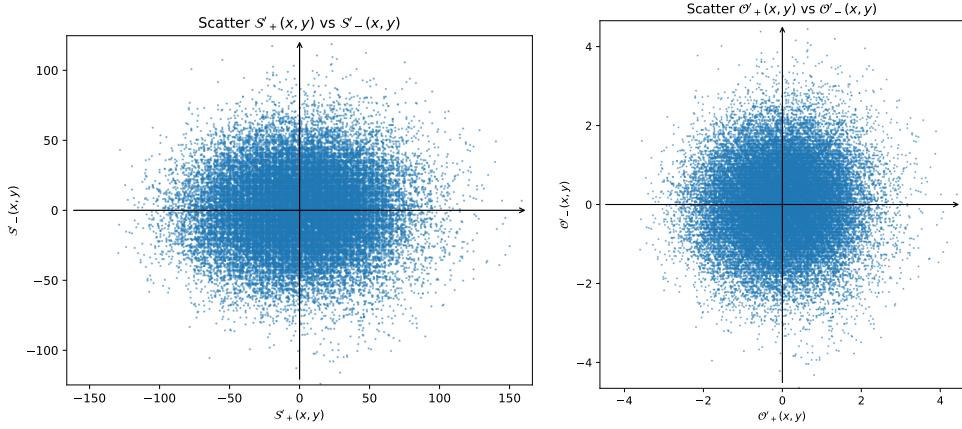


Fig. 17: Scatter plot of noisy images \mathcal{S}_+ and \mathcal{S}_- , also \mathcal{O}'_+ and \mathcal{O}'_- .

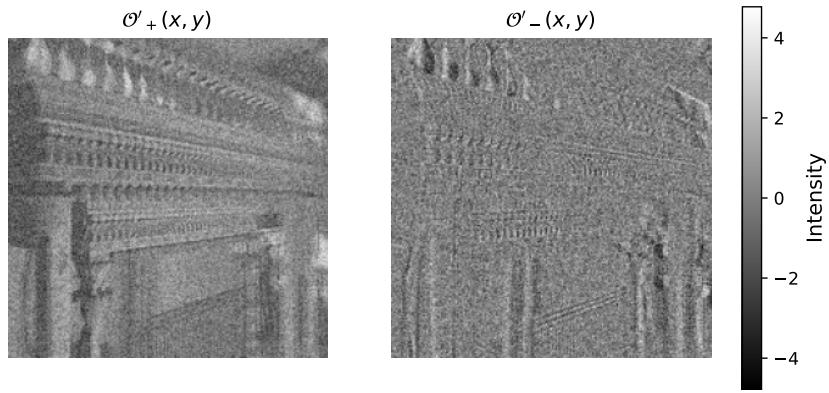


Fig. 18: Noisy images \mathcal{O}'_+ and \mathcal{O}'_- .

are nearly 0 already, which indicates that images with larger noise amplitude are more uncorrelated.

Noise with larger amplitude was also tested, the noisy images and the scatter plot are shown in Fig. 20 and Fig. 21 respectively. It can be seen that with larger amplitude, the scatter plot becomes more circular and the eigenvectors can seldomly indicate the correlation between the two images, such as the feature of 'smoothing' and 'contrast'. The correlation matrix shown below also indicates that the two images are nearly uncorrelated now, as the off-diagonal terms are close to zero compared to the diagonal terms.

$$R^{S''} \approx \begin{pmatrix} 1058.44 & 3.00 \\ 3.00 & 896.76 \end{pmatrix}$$

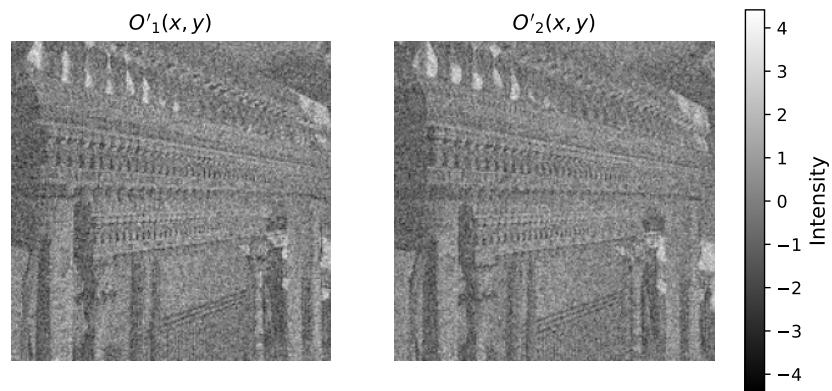


Fig. 19: Noisy images O'_1 and O'_2 .

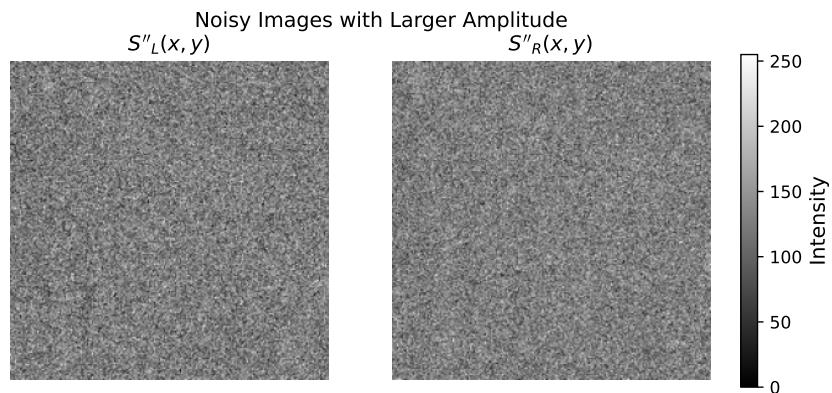


Fig. 20: Noisy images S'_L and S'_R with larger amplitude.

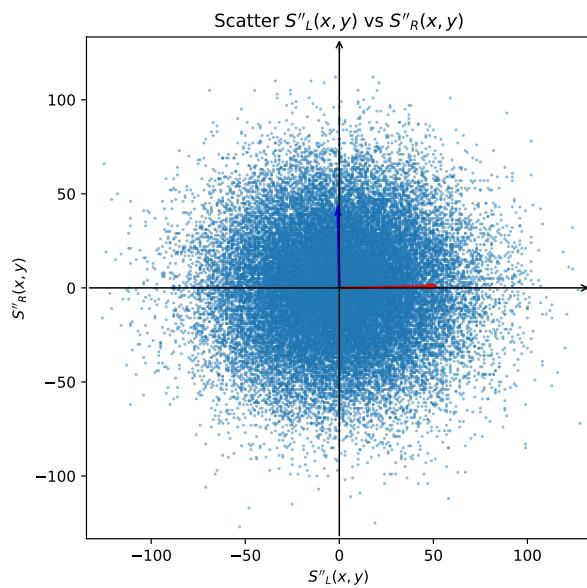


Fig. 21: Scatter plot of noisy images S'_L and S'_R with larger amplitude.

2.2 Part 2: Efficient spatial coding

A Choose a natural picture

The natural image and its gray-scaled luminance are shown in Fig. 22.



Fig. 22: Natural image and its luminance version. (Left: original image; Right: luminance image)

B Zero-mean and approximate spatial correlation

Normalize the luminance image to make it zero mean, and then compute the spatial correlation matrix $\mathbf{R}^S(d)$ as a function of pixels distance d , shown in Fig. 23 (left). Also, here tried considering the image to be 'periodic', i.e., horizontally mirror the image and consider the pixel at the right edge is neighboring the pixel at the left edge, to keep the same number of pixel pairs for all distances. The result is shown in Fig. 23 (right).

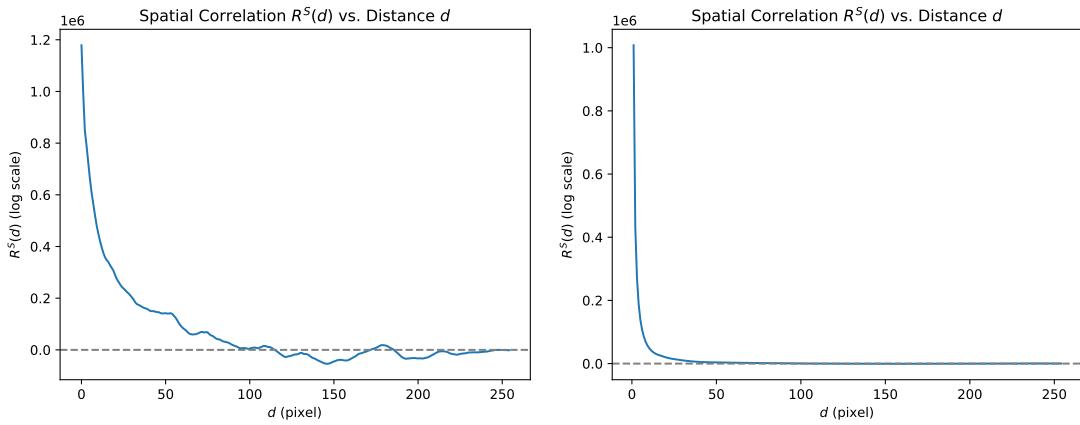


Fig. 23: Spatial correlation as a function of pixel distance d . (Left: non-periodic; Right: periodic)

It can be seen that, as distance d of pixels increase, the approximate spatial correlation $R^S(d)$ decays rapidly, exhibiting an roughly exponential decline on a logarithmic scale.

For some d there are even negative correlation, which might due to the limited data sample of the image and the specific features in this image. It indicates that pixels far away from each other tend to be less correlated, which is reasonable as the features in natural images are usually locally continuous. From the aspects of efficient coding, it means that pixels close to each other contain more redundant information, thus can be compressed with fewer bits more effectively.

C 2-D Fourier transform

Use ‘numpy.fft.fft2’ to compute the 2-D Fourier transform of the zero-mean luminance image, plot the frequency spectrum in Fig. 24 (Top), and then compute the power spectrum as the squared norm of the Fourier coefficients $k = \sqrt{k_x^2 + k_y^2}$, shown in Fig. 24 (Bottom).

It can be seen that the frequency spectrum shows obviously large value near the zero frequency axes, which indicates the low frequency hold larger power in the image, and the power decreases with frequency k roughly as $\sim |1/k|^2$.

D Fourier transform with noise

Here random noise with range of $[-N_{max}, N_{max}]$ are added to the zero-mean luminance image, where $N_{max} = 100$ is set. The noisy image is shown in Fig. 25.

The frequency spectrum and power spectrum are shown in Fig. 26. It can be seen that with random noise, the frequency spectrum show clear ‘x’ pattern along the zero frequency axes as before, and the power at high frequency k doesn’t decrease along $1/k^2$ as expected, but tends to a constant value, which is aligned with the pattern shown in Figure 4B in the instruction. It may because the random noise contains all frequency components equally.

E Gain control

Apply gain control to the Fourier components $\mathcal{S}(k)$ with gain values $g(k) = (|k| + |k_o|) \cdot \exp(-|k|^2/k_{low}^2)$ to get the output $\mathcal{O}(k)$, setting $k_o = 10^{-8}$ and $k_{low} = 0.15$. $k_{low} = 0.15$ is selected due to the observation in Fig. 26 (Right), where the power spectrum starts to deviate from the $\sim 1/k^2$ line around this value. It is also the *optimal* gain where the signal and noise power are comparable. The gain versus frequency k under the optimal gain is shown in Fig. 27 (left), and the gain matrix versus k_x and k_y is shown in Fig. 27 (right).

Then do the inverse Fourier transform to $\mathcal{O}(k)$ to get the output image $O(x, y)$ under optimal gain control. In addition, two output images with ‘higher-pass’ ($k_{low} = 1$) and

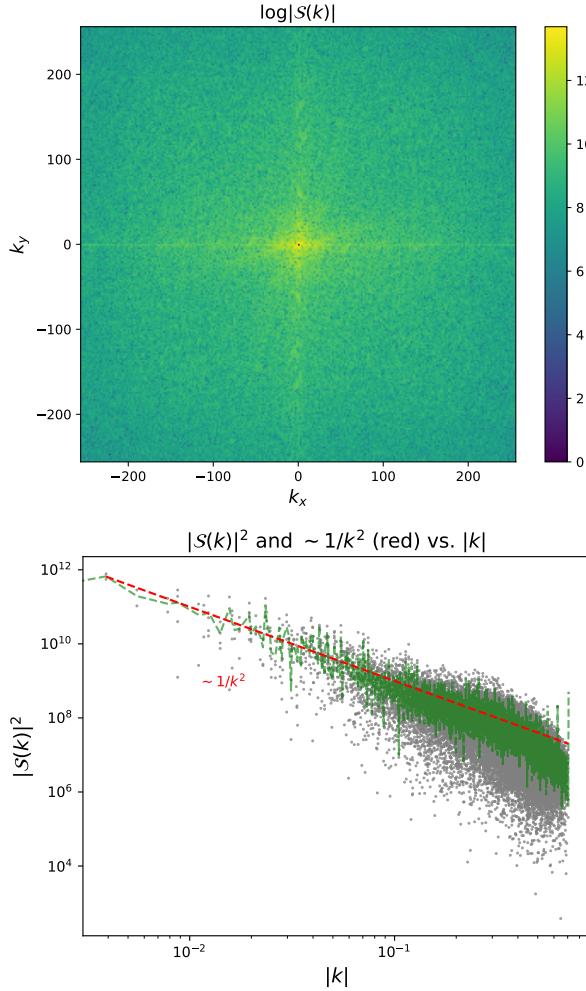


Fig. 24: Top: Fourier transform frequency spectrum $\mathcal{S}(k)$ of signal $S(k)$, visualized as $\log|\mathcal{S}(k)|$. Bottom: Power $|\mathcal{S}(k)|^2$ as a function of spatial frequency k . (Red line: $\sim |1/k|^2$, Green line: the average power for each $|k|$)

'lower-pass' ($k_{low} = 0.03$) gain control are also computed for comparison. The three output images are shown in Fig. 28. The results are similar to Figure 4CDE in the instruction, where the optimal gain control output image shows enhanced edges and details, the higher-pass gain control output image shows very sharp edges but lost most of the other details, while the lower-pass gain control smooths the signal but blurs the edges.

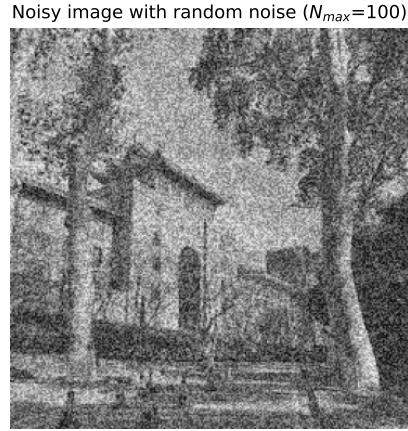


Fig. 25: Noisy luminance image.

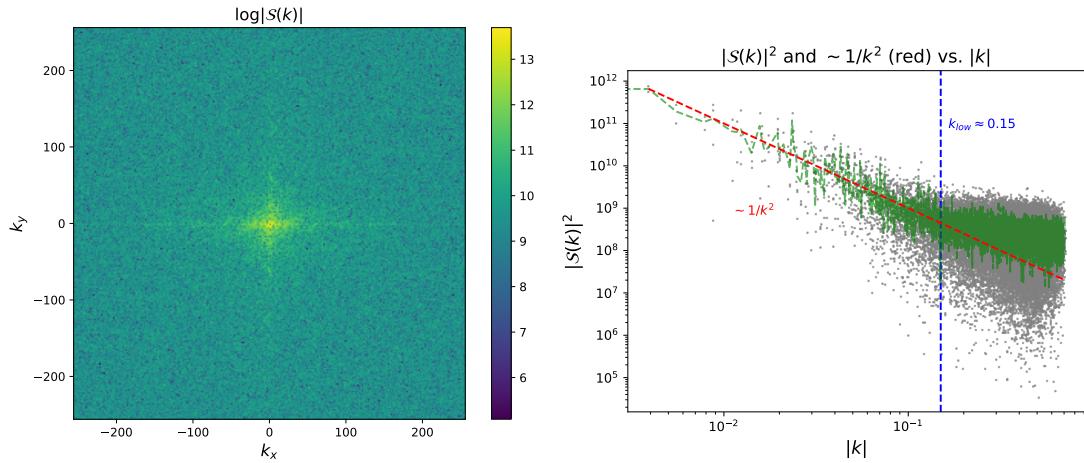


Fig. 26: Fourier transform frequency spectrum $S(k)$ (Left) and power spectrum (Right) of noisy luminance image.

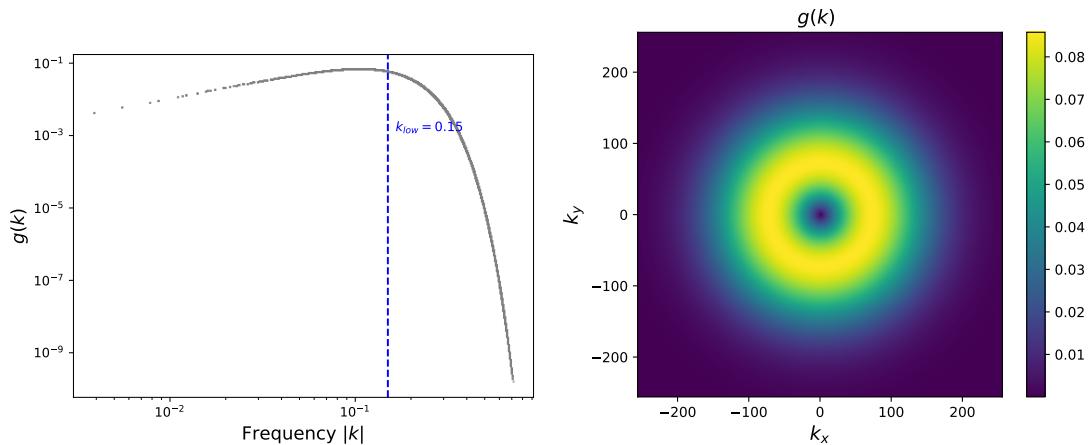


Fig. 27: The optimal gain $g(k)$ versus frequency k (Left), and versus k_x and k_y (Right)

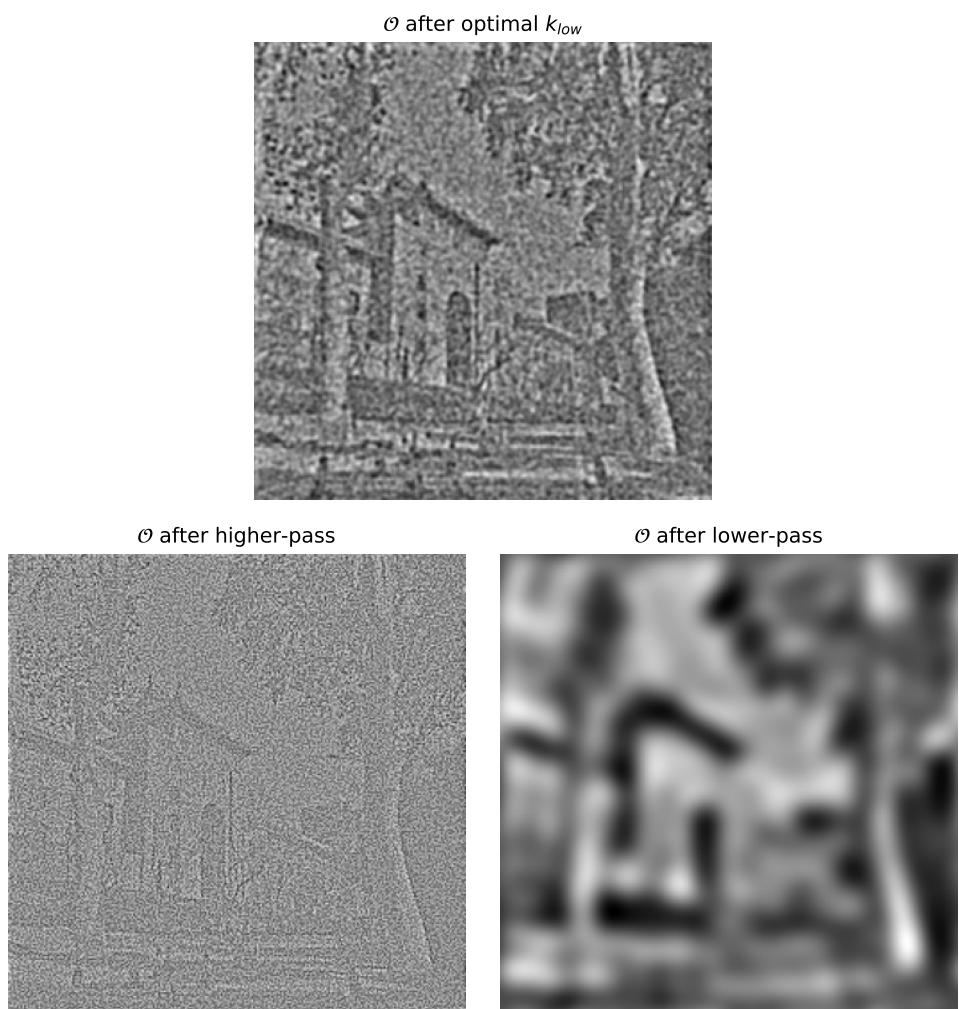


Fig. 28: Output images after gain control. (Top: optimal gain control; Bottom left: higher-pass gain control; Bottom right: lower-pass gain control)

F Spatial receptive field from $g(k)$

Here study the spatial receptive field corresponding to the gain $g(k)$, by doing the inverse Fourier transform of $g(k)$. The receptive field under optimal gain $k_{low} = 0.15$ is shown in Fig. 29.

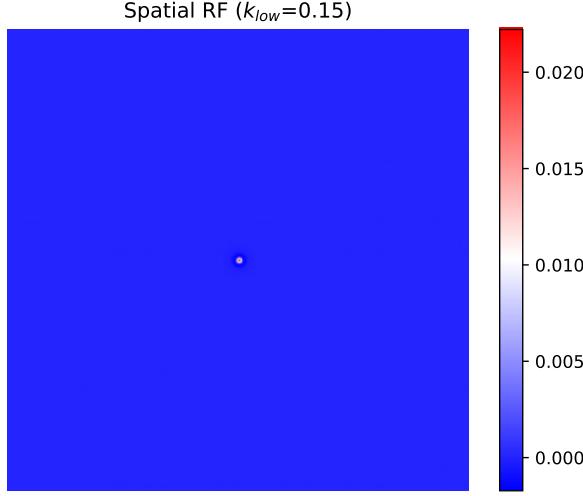


Fig. 29: Spatial receptive field corresponding to the optimal gain $g(k)$.

If we change k_{low} to be smaller value and make gain to be lower-pass filter, e.g. $k_{low} = 0.02, 0.01, 0.005$, as shown in Fig. 30, the receptive field spread out with lower the peak and weakened negative-surround (according to the colorbars).

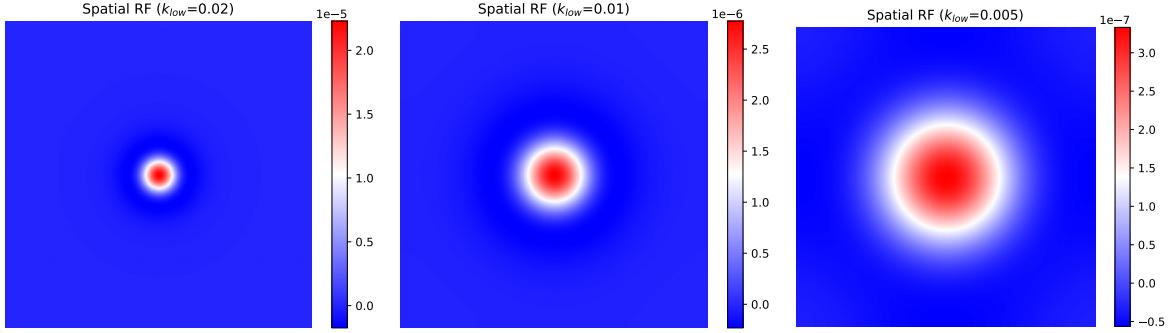


Fig. 30: Spatial receptive fields corresponding to lower-pass gain $g(k)$ with $k_{low} = 0.02$ (Left), $k_{low} = 0.01$ (Middle), and $k_{low} = 0.005$ (Right).

As the image become noisier, the Signal-to-Noise rate decrease, then the k_{low} where the signal and noise are comparable become smaller. Thus the gain control tends to be more like a lower-pass filter, to collect worthwhile low-frequency information and suppress high-frequency noise. Correspondingly, lower frequency would lead to more spread-out spatial receptive field with weakened negative-surround, to collect information from larger spatial area.