Image Denoising

Homework n°5 - Exercises

Exercise 1.2

Let X be an input grayscale image, of size $M \times N$.
We consider a 4×4 patch of X, denoted by P.
By definition, its DCT is given by

$$DCT(P)_{i,j} = \alpha_i \alpha'_j \sum_{k=0}^{3} \sum_{l=0}^{3} P_{k,l} \cos\left(\pi \left(k + \frac{1}{2}\right) \frac{i}{M}\right) \cos\left(\pi \left(l + \frac{1}{2}\right) \frac{j}{N}\right)$$

where

$$\alpha_i = \begin{cases} \sqrt{1/(4M)} & \text{if } i = 0 \\ \sqrt{1/(2M)} & \text{if } i = 1, 2, 3 \end{cases}, \alpha'_j = \begin{cases} \sqrt{1/(4N)} & \text{if } j = 0 \\ \sqrt{1/(2N)} & \text{if } j = 1, 2, 3. \end{cases}$$

We consider a convolution of weights w (and without bias), and Q_w the output of P through w:

$$Q_w(i, j) = \sum_{s,t} P(i + s, j + t) w(s, t).$$

For each pair of frequencies $(i, j) \in \{0, 1, 2, 3\} \times \{0, 1, 2, 3\}$, we define a convolution by $w_{i,j}(s, t) = \alpha_i \alpha'_j \cos\left(\pi \left((i + s) + \frac{1}{2}\right) \frac{i}{M}\right) \cos\left(\pi \left((j + t) + \frac{1}{2}\right) \frac{j}{N}\right)$ so that $DCT(P)_{i,j} = Q_{w_{i,j}}(i, j)$.

We thus have $4 \times 4 = \boxed{16 \text{ convolution filters}}$.

So to compute the DCT of a given patch P, for each frequency (i, j) we apply the convolution kernel $w_{i,j}$, and retrieve $Q_{w_{i,j}}(i, j)$.

The patch-wise DCT transform (with 4×4 patches) of a grayscale image can be implemented and represented using convolutions.

Exercise 1.4

$$\mathcal{F}(x) = f_3(y;\theta_3),\ \ y = f_2(f_1(x;\theta_1);\theta_2)$$
$$\mathcal{G}(x) = y + f_3(y;\theta_3)$$

We also set $\boxed{z = f_1(x;\theta_1)}$, so $y = f_2(z;\theta_2)$.
We then have:

$$\begin{aligned}
\frac{\partial \mathcal{F}}{\partial \theta_1}(x) &= \frac{\partial f_3(y;\theta_3)}{\partial \theta_1} \\
&= \frac{\partial f_3(y;\theta_3)}{\partial y} \times \frac{\partial y}{\partial \theta_1} \text{ (chain rule)} \\
&= \frac{\partial f_3(y;\theta_3)}{\partial y} \times \frac{\partial f_2(z;\theta_2)}{\partial \theta_1} \\
&= \frac{\partial f_3(y;\theta_3)}{\partial y} \times \frac{\partial f_2(z;\theta_2)}{\partial z} \times \frac{\partial z}{\partial \theta_1} \text{ (chain rule)} \\
&= \frac{\partial f_3(y;\theta_3)}{\partial y} \times \frac{\partial f_2(z;\theta_2)}{\partial z} \times \frac{\partial f_1(x;\theta_1)}{\partial \theta_1}.
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \mathcal{F}}{\partial \theta_2}(x) &= \frac{\partial f_3(y;\theta_3)}{\partial \theta_2} \\
&= \frac{\partial f_3(y;\theta_3)}{\partial y} \times \frac{\partial y}{\partial \theta_2} \text{ (chain rule)} \\
&= \frac{\partial f_3(y;\theta_3)}{\partial y} \times \frac{\partial f_2(z;\theta_2)}{\partial \theta_2}
\end{aligned}$$

Similarly, we have

$$\frac{\partial \mathcal{F}}{\partial \theta_3}(x) = \frac{\partial f_3(y;\theta_3)}{\partial \theta_3}.$$

As for $\mathcal{G}(x) = y + \mathcal{F}(x)$, we have $\dfrac{\partial \mathcal{G}(x)}{\partial \theta_i} = \dfrac{\partial \mathcal{F}(x)}{\partial \theta_i} + \dfrac{\partial y}{\partial \theta_i}$. We then use the above results, and notice that we have already computed $\dfrac{\partial y}{\partial \theta_1}$ and $\dfrac{\partial y}{\partial \theta_2}$, and that $\dfrac{\partial y}{\partial \theta_3} = 0$ so:

$$\frac{\partial \mathcal{G}}{\partial \theta_1}(x) = \frac{\partial f_2(z;\theta_2)}{\partial z} \times \frac{\partial f_1(x;\theta_1)}{\partial \theta_1} \times \left(1 + \frac{\partial f_3(y;\theta_3)}{\partial y}\right),$$

$$\frac{\partial \mathcal{G}}{\partial \theta_2}(x) = \frac{\partial f_2(z;\theta_2)}{\partial \theta_2} \times \left(1 + \frac{\partial f_3(y;\theta_3)}{\partial y}\right),$$

$$\frac{\partial \mathcal{G}}{\partial \theta_3}(x) = \frac{\partial f_3(y;\theta_3)}{\partial y}.$$

We see that compared to $\mathcal{F}$, the derivatives of $\mathcal{G}$ with respect to θ_1 and θ_2 have more terms. This may prevent **gradient vanishing**.