1 2D-DCT Transform

(a) For $A^{4\times4}$, N=4. Therefore, the DCT matrix T should be:

$$T[i,j] = \begin{cases} \sqrt{\frac{1}{N}} = \frac{1}{2}, & \text{if } i = 0 \\ \sqrt{\frac{2}{N}} \cos\left(\frac{(2j+1)i\pi}{2N}\right) = \frac{1}{\sqrt{2}} \cos\left(\frac{(2j+1)i\pi}{8}\right), & \text{if } 0 < i < N,\ 0 \le j < N \end{cases}$$

.

$$T = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2}\cos\frac{\pi}{8} & \frac{1}{2}\cos\frac{3\pi}{8} & \frac{1}{2}\cos\frac{5\pi}{8} & \frac{1}{2}\cos\frac{7\pi}{8} \\ \frac{1}{2}\cos\frac{2\pi}{8} & \frac{1}{2}\cos\frac{6\pi}{8} & \frac{1}{2}\cos\frac{10\pi}{8} & \frac{1}{2}\cos\frac{14\pi}{8} \\ \frac{1}{2}\cos\frac{3\pi}{8} & \frac{1}{2}\cos\frac{9\pi}{8} & \frac{1}{2}\cos\frac{15\pi}{8} & \frac{1}{2}\cos\frac{21\pi}{8} \end{bmatrix} \approx \begin{bmatrix} 0.5 & 0.5 & 0.5 & 0.5 \\ 0.653281 & 0.270598 & -0.270598 & -0.653281 \\ 0.5 & -0.5 & -0.5 & 0.5 \\ 0.270598 & -0.653281 & 0.653281 & -0.270598 \end{bmatrix}.$$

As for the DCT transform consequence $D = TAT^T$

Therefore, the consequence D= $\begin{bmatrix} 10.0000 & 9.2388 & 0.0000 & -3.8268 \\ 9.2388 & 8.5355 & 0.0000 & -3.5355 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ -3.8268 & -3.5355 & 0.0000 & 1.4645 \end{bmatrix}$.

(b)Comparing with the 2D-DFT, the 2D-DCT is easier to compute. The DCT only consider the cosine computation. The DFT must consider cosine and sine computation at same time. They will have similar performance on sjowing the color frequency on images. However, the DCT could save half of computation cost. Therefore, DCT performs better from my perspective.

2 Controllable Vedio Generation

(a).

(b).

References

[1] Chan, C. et al., “Everybody Dance Now”, CVPR 2019.

[2] Siarohin, A. et al., “First Order Motion Model for Image Animation”, NeurIPS 2019.

[3] Zhang, X. et al., “Pose-Guided Person Image Generation”, ICCV 2019.