COMP3204 — Lecture 3 Notes: Image Sampling (DFT, FFT, Sampling, Aliasing)

1 Why frequency space matters in Computer Vision

- An image can be treated as a **2D signal**: intensity varies over **space** (pixel coordinates).
- The Fourier viewpoint: represent a signal as a sum of **sinusoids** at different **frequencies**.
- In images, frequency means **spatial frequency**:
  - **Low spatial frequency**: intensity changes slowly (smooth shading, broad structure).
  - **High spatial frequency**: intensity changes rapidly (edges, fine texture, noise).
- Why it is useful (as stated in the lecture material):
  - **Understanding/analysis** of image content at different scales.
  - **Speeding up computation** (many operations become efficient using FFT).
  - **Representation properties** (e.g., shift effects become simple).
  - **Coding/compression** (store/keep important components).
  - **Recognition/understanding** (e.g., texture).

2 1D Discrete Fourier Transform (DFT)

2.1 Forward 1D DFT

- Input: $p[i]$ for $i = 0, \ldots, N-1$ (a sampled 1D signal).
- Output: complex frequency coefficients $Fp[u]$ for $u = 0, \ldots, N-1$.
- Definition:

$$Fp[u] \;=\; \frac{1}{N}\sum_{i=0}^{N-1} p[i]\; e^{-j\frac{2\pi}{N}ui}. \tag{1}$$

- Meaning of symbols:
  - i: sample index (position in the signal).
  - u: frequency index (which discrete sinusoid is being measured).
  - j: imaginary unit, $j^2 = -1$.
  - $e^{-j\theta} = \cos\theta - j\sin\theta$: complex sinusoid basis.
- The $1/N$ scaling is chosen so that the **DC term** becomes the **average** of the samples.

2.2 Magnitude and phase: what the complex output means

- Each $Fp[u]$ is complex and can be written as:

$$Fp[u] = \mathrm{Re}(Fp[u]) + j\ \mathrm{Im}(Fp[u]).$$

- **Magnitude** (strength of that frequency):

$$|Fp[u]| = \sqrt{\mathrm{Re}(Fp[u])^2 + \mathrm{Im}(Fp[u])^2}.$$

- **Phase** (alignment/shift information):

$$\arg(Fp[u]) = \tan^{-1}\left(\frac{\mathrm{Im}(Fp[u])}{\mathrm{Re}(Fp[u])}\right).$$

- DC component:
  - For $u = 0$, $e^{-j(0)} = 1$, so

    $$Fp[0] = \frac{1}{N}\sum_{i=0}^{N-1} p[i],$$

    which is the **mean value** of the signal.

2.3 Inverse 1D DFT (reconstruction)

- You can reconstruct the original samples from all frequency coefficients:

$$p[i] \ = \ \sum_{u=0}^{N-1} Fp[u]\ e^{+j\frac{2\pi}{N}ui}. \tag{2}$$

- Interpretation: the signal is rebuilt by adding all sinusoidal components (including DC).

3 Transform-pair intuition (pulse example)

- A short, sharp pulse in the sample domain contains many frequencies.
- Key exam takeaway:
  - **Sharp changes** in a signal require **high-frequency** components.
  - In images, **edges** are high-frequency phenomena.

4 Reconstruction using only some frequencies

- If you reconstruct using only:
  - **Low frequencies**: you keep coarse structure but lose detail (blurry/smooth).
  - **More frequencies**: you progressively recover sharper details.
- Conceptual message from the lecture visuals:
  - Low frequencies carry global structure.
  - High frequencies add fine detail and sharpness.

5 2D DFT for images

5.1 Forward 2D DFT

- Image samples: $P[x, y]$ for $x, y = 0, \ldots, N-1$.
- Frequency coefficients: $FP[u, v]$, where u and v are horizontal/vertical spatial frequency indices.
- Definition:

$$FP[u, v] \;=\; \frac{1}{N^2} \sum_{y=0}^{N-1} \sum_{x=0}^{N-1} P[x, y]\; e^{-j\frac{2\pi}{N}(ux+vy)}. \tag{3}$$

- DC component $FP[0, 0]$ equals the **average intensity** of the entire image.

5.2 Inverse 2D DFT

- Reconstruction:

$$P[x, y] \;=\; \sum_{v=0}^{N-1} \sum_{u=0}^{N-1} FP[u, v]\; e^{+j\frac{2\pi}{N}(ux+vy)}. \tag{4}$$

- Interpretation: the image is a sum of 2D sinusoidal "patterns" (basis functions).

5.3 Visualising the Fourier transform

- In practice, displays often show:
  - **Magnitude** (often using a log scale so the large DC component does not dominate).
  - **Phase** (contains crucial alignment/structure information even if it looks unintuitive).
- Keeping only low-frequency coefficients reconstructs a recognizable but blurry image; adding higher frequencies restores detail.

6 FFT: fast computation of the DFT

- Direct DFT computation is expensive because you compute a sum for each frequency coefficient.
- **FFT (Fast Fourier Transform)** computes the same DFT much faster.
- Core algorithmic idea (as in the lecture material):
  - **Divide-and-conquer**: split into smaller DFTs, then combine results efficiently (often visualised as repeated add/subtract "butterfly" patterns).
- Why it matters in vision: makes frequency-domain processing feasible on real image sizes.

7 Fourier properties used in vision

7.1 Shift effect (magnitude invariance)

- Shifting a signal/image in space changes **phase** but preserves **magnitude**.
- 1D property (conceptual form):

$$\mathcal{F}\{p(t-\tau)\} = e^{-j\omega\tau} P(\omega).$$

- Since $|e^{-j\omega\tau}| = 1$, the magnitude is unchanged:

$$|\mathcal{F}\{p(t-\tau)\}| = |P(\omega)|.$$

- Exam implication:
  - Magnitude-based features can be robust to translation.
  - Phase carries location/alignment information.

7.2 Rotation

- Rotating an image rotates its Fourier transform by the same angle.
- Intuition: orientation of structures (e.g., stripes/edges) corresponds to oriented energy in frequency space.

8 Filtering using the Fourier transform

- Filtering idea: modify the frequency coefficients, then apply the inverse transform.
- **Low-pass filtering**:
  - Keep low frequencies, suppress high frequencies.
  - Effect: smoothing / blur, reduces fine detail and noise.
- **High-pass filtering**:
  - Suppress low frequencies, keep high frequencies.
  - Effect: emphasizes edges and fine details.

9 Sampling theory

9.1 What sampling is

- Sampling converts a continuous signal into discrete measurements.
- In images: the sensor/ADC measures intensity on a discrete pixel grid.
- Good sampling: enough samples to capture the fastest variations you care about.
- Poor sampling: too few samples can make high-frequency content appear as a different (lower-frequency) pattern.

9.2 Sampling in frequency domain: spectrum replication idea

- Lecture key relation (stated conceptually):

$$\mathcal{F}\{x(t)\,\delta(t)\} = \mathcal{F}\{x(t)\} * \mathcal{F}\{\delta(t)\}.$$

- Takeaway used in sampling theory:
  - Sampling causes **copies** of the spectrum to repeat in frequency.

9.3 Nyquist sampling theorem

- To reconstruct a signal from samples:

$$f_s \geq 2f_{\max},$$

  where f_s is the sampling frequency and $f_{\max}$ is the highest frequency present in the original signal.

- Frequency-domain interpretation:
  - If f_s is high enough, repeated spectra do **not overlap** $\Rightarrow$ reconstruction is possible.
  - If $f_s < 2f_{\max}$, the repeated spectra **overlap** $\Rightarrow$ aliasing occurs.

- Practical guideline from the slides: "two pixels for every pixel of interest" (sample densely enough to represent the smallest detail you care about).

10 Aliasing

- **Aliasing** happens when sampling is too low, causing high-frequency content to appear as lower frequency content (false patterns).
- Spatial aliasing example (as shown): fine repetitive textures (e.g., blinds) can produce large-scale incorrect patterns when downsampled.
- Temporal aliasing (wagon-wheel effect):
  - Sampling in time is the frame rate.
  - If motion changes too much between frames, a wheel can appear to rotate backwards or stand still.
- Conceptual avoidance (as implied by the sampling theorem material):
  - Increase sampling rate (higher resolution / higher frame rate), and/or
  - Remove high frequencies before sampling (pre-filtering / anti-aliasing).

11 Mentioned (not expanded in this lecture)

- Compressed sensing
- Sparsity
- Regularisation