COMP3204 — Lecture 2 Notes: Image Formation & Frequency Space

1 Lecture 2 Scope

- Topics covered in the lecture slides:
  - What is inside an image?
  - Restrictions on image formation (resolution + brightness sampling)
  - Moving to a different representation: **Fourier / frequency domain**

2 What is Inside an Image? (Digital Image Representation)

- **Image as a matrix**
  - A grayscale image can be treated as an $N \times N$ **matrix** of numbers (pixel intensities).
- **Pixel values and bit-depth**
  - In these slides, each pixel is stored as an **8-bit unsigned integer**.
  - **8-bit** $\Rightarrow$ there are $2^8 = 256$ possible values.
  - **Unsigned** $\Rightarrow$ values are non-negative.
  - Therefore pixel intensity range: $[0, 255]$.
  - Bits are weighted by powers of 2:

$$\text{value} = \sum_{k=0}^{7} b_k 2^k, \quad b_k \in \{0, 1\}$$

    where b_0 is the **least significant bit (LSB)** and b_7 is the **most significant bit (MSB)**.
- **Bit-plane decomposition**
  - A **bit-plane** is the binary image formed by taking one fixed bit position from every pixel.
  - Visual interpretation from the slide examples:
    * Higher-order bits (especially the **MSB**) contain most of the **recognisable structure**.
    * Lower-order bits (near the **LSB**) often look **noise-like**.
    * Some mid-level planes can capture effects like **lighting variation**.
  - Exam takeaway: **Most perceptual content is carried by higher-order bits.**

3 Restrictions on Image Formation

- **Resolution (spatial sampling)**
  - If an image is sampled on an $N \times N$ grid, then N controls spatial detail.
  - Low N (e.g. 64×64): blocky, detail is lost.
  - Higher N (e.g. 128×128, 256×256): more detail preserved.
  - Trade-off stated in slides:
    * Larger N increases storage and computation.
    * Motivating question: **How do we choose an appropriate value for N?**
- **Brightness sampling (intensity quantisation)**
  - Brightness is discretised by limiting intensity to a finite set of values (here: 0–255).
  - Key framing from slides: **sampling affects both space and brightness, and is not as simple as it appears.**

4 Frequency Viewpoint: Waves and Why Fourier Matters

- **Waves**
  - A sine/cosine wave is a repeating pattern over a variable (often time).
  - For images, we later generalise this idea to **2D patterns** over spatial coordinates (x, y).
- **Complex exponential representation**
  - A compact way to represent sinusoids is via the complex exponential:

$$e^{j\omega t} = \cos(\omega t) + j \sin(\omega t)$$

  - Interpretation (as in the diagram):
    * $\cos(\omega t)$ is the real-axis component,
    * $\sin(\omega t)$ is the imaginary-axis component,
    * the vector rotates in the complex plane as t changes.
- **Period and frequency**
  - If a pattern repeats every T seconds (period), its frequency is:

$$\xi = \frac{1}{T}$$

  - High frequency $\Rightarrow$ repeats quickly; low frequency $\Rightarrow$ repeats slowly.

5 Fourier Transform: From Time/Space Domain to Frequency Domain

- **Core idea (Fourier)**
  - Signals can be expressed as sums of sinusoids at different frequencies.

- **Fourier transform (definition used in slides)**
  - For a signal $p(t)$:
  $$F_p(\omega) = \int_{-\infty}^{\infty} p(t)\, e^{-j\omega t}\, dt$$
  - This produces a complex-valued function telling how much of each angular frequency ω is present.
  - The exponential can be expanded as:
  $$e^{-j\omega t} = \cos(\omega t) - j\sin(\omega t)$$
  (so the transform measures cosine and sine content, packaged together).
- **Angular frequency vs ordinary frequency**
  - Relationship:
  $$\omega = 2\pi\xi$$
- **Inverse Fourier transform (reconstruction)**
  - Recover the original signal by integrating over all frequencies:
  $$p(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F_p(\omega)\, e^{j\omega t}\, d\omega$$
  - Interpretation: $p(t)$ is rebuilt by summing all sinusoidal components weighted by $F_p(\omega)$.
- **Equivalent form in terms of ξ (cycles per unit time)**
  - Slides also show:
  $$F_p(\xi) = \int_{-\infty}^{\infty} p(t)\, e^{-j2\pi\xi t}\, dt, \qquad p(t) = \int_{-\infty}^{\infty} F_p(\xi)\, e^{j2\pi\xi t}\, d\xi$$
  - Same concept; the factor 2π moves into the exponent.
- **Key conceptual mapping highlighted in slides**
  - Fourier transform: move from a representation where components are "mixed" to one where frequency components are separated/identifiable.
  - Inverse transform: combine frequency components to reconstruct the original.

6 Worked Example in Slides: Rectangular Pulse → Fourier Transform

- **Signal**
  - A rectangular pulse $p(t)$ with amplitude 1 over a finite window and 0 elsewhere.
- **How the integral is evaluated (the key slide step)**
  - Start:
  $$F_p(\omega) = \int_{-\infty}^{\infty} p(t)e^{-j\omega t}\, dt$$

– Since $p(t) = 0$ outside the pulse duration, the integral reduces to where $p(t) = 1$:

$$F_p(\omega) = \int_{-T/2}^{T/2} e^{-j\omega t}\, dt$$

– Integrate:

$$\int e^{-j\omega t} dt = \frac{e^{-j\omega t}}{-j\omega}$$

– Apply limits:

$$F_p(\omega) = \left[\frac{e^{-j\omega t}}{-j\omega}\right]_{-T/2}^{T/2} = \frac{e^{-j\omega(T/2)} - e^{j\omega(T/2)}}{-j\omega}$$

– Use $e^{-ja} - e^{ja} = -2j\sin(a)$:

$$F_p(\omega) = \frac{2\sin(\omega T/2)}{\omega}$$

7 Reconstruction Intuition (Inverse Transform Visualisation in Slides)

- Each frequency ω contributes an oscillatory component.
- Adding more frequencies (integrating over a wider range of ω) improves the reconstruction of the original pulse.
- Slide illustration shows contributions for specific ω values and how combining them approximates the pulse.

8 Complex Output: Magnitude and Phase

- **Fourier outputs are complex**

$$F_p(\omega) = \mathrm{Re}(F_p(\omega)) + j\ \mathrm{Im}(F_p(\omega))$$

- **Magnitude**
  – Measures the strength of frequency ω:

$$|F_p(\omega)| = \sqrt{\mathrm{Re}(F_p(\omega))^2 + \mathrm{Im}(F_p(\omega))^2}$$

- **Phase**
  – Encodes alignment/shift information of components:

$$\arg(F_p(\omega)) = \tan^{-1}\left(\frac{\mathrm{Im}(F_p(\omega))}{\mathrm{Re}(F_p(\omega))}\right)$$

9 Example in Slides: Fourier Transform of a Musical Chord

- A chord is a mixture of multiple frequencies.
- Its Fourier transform shows multiple components corresponding to the tones/harmonics present.

10 Why Phase Matters (Highly Examinable Slide Result)

- Slide demonstration: reconstruct images by mixing
  - magnitude from one image and phase from another.
- Observation: the reconstructed image tends to resemble the image that provided the **phase** (structure is strongly phase-driven in this example).
- Exam takeaway:
  - **Magnitude** $\Rightarrow$ "how much of each frequency"
  - **Phase** $\Rightarrow$ "where/structure alignment" (crucial for recognisable reconstruction)

11 Fourier for Reconstruction / Coding (Compression Intuition in Slides)

- Slide pipeline: transform $\rightarrow$ keep/encode selected components $\rightarrow$ inverse transform.
- Slide example: keeping only a small fraction of transform components (e.g. **5%**) can still yield a visually close reconstruction, with an error image showing what was lost.

12 Main Slide Summary Points

- **Sampling is not as simple as it appears.**
- **Sampling affects space and brightness.**
- **Fourier helps us understand frequency.**
- **Fourier enables coding (compression) and more.**
- Next framing in slides: Fourier will be used to understand sampling.

13 Other Transforms Mentioned (Names Only; Not Expanded)

- Discrete Cosine (Sine) Transform
- Discrete Hartley Transform
- Wavelet families/types (named in slides): continuous, discrete, complex, stationary, dual, Haar, Daubechies, Morlet, Gabor, etc.
- Other named transform families: curvelets, shearlets, bandelets, contourlets, fresnelets, chirplets, noiselets, ...

14 Wavelet Transform (What the Slide Example Shows)

- Slides show an example of a **2D discrete wavelet transform** used in **JPEG2000**.
- Output appears as multiple sub-images capturing information at different:
  - **scales** (coarse-to-fine), and
  - **detail patterns** (different types of local changes).

- Key intuition consistent with the slide visual:
  - wavelets can represent content in a way that is both **spatially localised** (where it happens) and **scale-sensitive** (coarse structure vs fine detail).