INTRODUCTION TO ARTIFICIAL INTELLIGENCE

EPISODE 9: DIGITAL SIGNAL PROCESSING/PATTERN RECOGNITION
TODAY’S MENU

1. PATTERN RECOGNITION

2. ININVARIANT FEATURES

3. "STRUCTURE FROM MOTION"

Image (C): Steve McCurry/National Geographic
TERMINATOR AFTER ALL?
DIGITAL SIGNALS

• An image can be represented as a function $f(x,y)$ of the x- and y-coordinates

• A color image has three "bands", (RGB), i.e., three functions

• Audio signals can be represented as a function of time, $f(t)$, or equivalently, in the frequency domain, as a function of frequency, $f(f)$ [Hz]
DIGITAL SIGNALS

• An image can be represented as a function $f(x,y)$ of the x- and y-coordinates

• A color image has three "bands", (RGB), i.e., three functions

• Audio signals can be represented as a function of time, $f(t)$, or equivalently, in the frequency domain, as a function of frequency, $f(f)$ [Hz]
AI CHALLENGES WITH SIGNALS

• Digital signals may easily contain millions of individual entries
  – e.g., $1000 \times 1000\ \text{pixels} = 1000\ 000\ \text{pixels}$
  + in images: resolution [dpi]
  + in audio: sampling rate [Hz]

• Their information content, however, is limited:
  – redundancy
  – noise

• Signals also depend on external conditions:
  – camera angle, distance, lighting, ...
  – echo, microphone, background, ...

=> Recognizing objects is HARD!
DIGITAL SIGNAL PROCESSING

• What we mean by the field of Digital Signal Processing includes a wide range of signal related problems that are related to:
  – computer graphics (rendering)
  – signal enhancement (e.g., denoising, sharpening, and all kinds of "Photoshopping", Instagram, Prisma, ...)

DIGITAL SIGNAL PROCESSING

- What we mean by the field of Digital Signal Processing includes a wide range of signal related problems that are related to:
  - computer graphics (rendering)
  - signal enhancement (e.g., denoising, sharpening, and all kinds of "Photoshopping", Instagram, Prisma, ...)

[Images of six different faces with various artistic effects applied, labeled PRISMA]
DIGITAL SIGNAL PROCESSING

What we mean by the field of Digital Signal Processing includes a wide range of signal related problems that are related to:

- **computer graphics** (rendering)
- **signal enhancement** (e.g., denoising, sharpening, and all kinds of "Photoshopping", Instagram, Prisma, ...)
- **imaging** (e.g., computer tomography)
- **motion capture**
- **pattern recognition**
- ...
The task is to recognize a given object in a signal.

Mustn't be distracted by external conditions.

To do so, we need to identify features that are not sensitive to external conditions:

- in images:
  + shapes
  + relative distances
- in audio (Shazam!):
  + frequency
  + frequency changes (up, down)
  + rhythm
Invariant Features in Images

- **Invariant feature** = property of signals that is insensitive to external conditions

- Goal is to find features that are invariant w.r.t.:
  - *scale* (distance)
  - *orientation* (angle)
  and that can be computed
  - *efficiently*

- On the other hand, the features should also be identifiable
  - A point on a flat *surface* is *not* identifiable
  - A *corner* is identifiable
INVARIANT FEATURES IN IMAGES

- Scale Invariant Feature Transform (SIFT)

- Speeded-Up Robust Features (SURF)

We'll focus on SURF but the basic idea is similar (and SIFT is patented)

Reference:
**SURF**

- Stage 1: Choose **interest points**:
  - choose extrema of pixel intensity values:
    - no flat surface or edge, yes corner, "blob"
  - repeat at different scales to make features **scale-invariant**
SURF

• Stage 2: Construct **feature descriptors**
  – focus on the local neighborhood around each interest point
  – find dominant direction of intensity (=> **rotation-invariance**)
  – construct a 64-dimensional **descriptor vector** based on intensity variation
SURF

Stage 2:
- Focus on the local neighborhood around each interest point.
- Find the dominant direction of intensity (\(\Rightarrow\) rotation-invariance).
- Construct a 64-dimensional descriptor vector based on intensity variation.
SURF

- Stage 2: Construct **feature descriptors**
  - focus on the local neighborhood around each interest point
  - find dominant direction of intensity
    (=> rotation-invariance)
  - construct a 64-dimensional **descriptor vector** based on intensity variation

- Outcome:
  - (x, y, scale, orientation, descriptor vector)
SURF IN ACTION

REAL-TIME TRACKING
PATTERN RECOGNITION WITH SURF

• Stage 1: Choose interest points in images A and B

• Stage 2: Construct feature descriptors from both images

• Stage 3: Match features based on descriptor vectors
  – based on Euclidean distance
  – eliminate false positives (bad matches)
PATTERN RECOGNITION WITH SURF
PATTERN RECOGNITION WITH SURF
PATTERN RECOGNITION WITH SURF
PATTERN RECOGNITION WITH SURF

SURF + Bruteforce = Object Matching

qwook 107 views
COMPUTER VISION: "STRUCTURE FROM MOTION"

- **3D rendering**: Given a 3D model and the camera position \((x,y,z,u,v,w)\), project points on a 2D canvas.

- Inverse problem: Given a 3D model and the 2D projections, infer the camera position.
COMPUTER VISION:
"STRUCTURE FROM MOTION"

- **3D reconstruction**: Given 2D projections in several 2D images, infer the 3D model and the camera positions.

Source: Julien Michot
COMPUTER VISION:
"STRUCTURE FROM MOTION"

- **3D reconstruction**: Given 2D projections in several 2D images, infer the 3D model and the camera positions.

How to identify the same points in different images? **SURF!**