Multimedia

## What is multimedia?



▸ Text , sound , image , animation , movie, …
▸ What does CS do about multimedia?
  ▸ Create, manage(such as store, search etc), recognition, and enhance (digital) multimedia content.
  ▸ New applications and new devices: games, 3D movies,…
▸ Outline
  ▸ Image compression
  ▸ Computer vision: 3D image/movie
  ▸ Pattern recognition
▸

## Digital data compression

▸ Compression is to reduce the data size.
  ▸ Benefits: data can be stored and transmitted more efficiently.
▸ Examples,
  ▸ 00000000001111111111 can be compressed as (10,0,9,1)
  ▸ 123456789 can be compressed as (1,1,9)
  ▸ AABAAAABAAC can be compressed as 11011111011100, where A, B, C are encoded as 1, 01, and 00 respectively.
▸ Image data
  ▸ GIF: Good for cartoons
  ▸ JPEG: Good for photographs
  ▸ TIFF: Good for image archiving
▸

## JPEG (Joint Photographic Experts Group)



100%, size:315K
In BMP format

50%, size:13K
In JPG format

10%, size:5K
In JPG format

▸ Typical compression ratio is 12:1.

▸ American gothic by Grant Wood ,1930

## How does JPEG compress images?

- The main algorithm that used in JPEG is **discrete cosine transform (DCT)** $f_m = \sum_{k=0}^{n-1} x_k \cos\left[\frac{\pi}{n}m\left(k+\frac{1}{2}\right)\right]$
  - It transforms signals from spatial domain to frequency domain
    - Most images have large low frequency signals
  - A variation of Fourier transform (FT)
    - FT, DCT will be taught in 工程數學，通訊概論，數值方法
    - Space transformation will be taught in 線性代數
    - More image compression method will be taught in 影像處理

---

## 3D images

- Why can we have stereo sound?
  - The same reason that we can have stereo views
- The same objects viewed from different angles



  - Epipolar geometry (線性代數, 電腦視覺)
- Giving two images of the same object from different positions (known), one can reconstruct the depth information of the object

- http://en.wikipedia.org/wiki/Epipolar_geometry

---

## 3D movies

- Videos are shot using a special camera to record the images as seen from two perspectives.
- When the movie is played, a special projection hardware are needed to play both perspectives synchronously.
- When watching the movies, viewers need to ware special glasses to get the illusion of depth.
  - Glasses with colored filters
  - Polarized glasses (普物)



---

## Pattern recognition
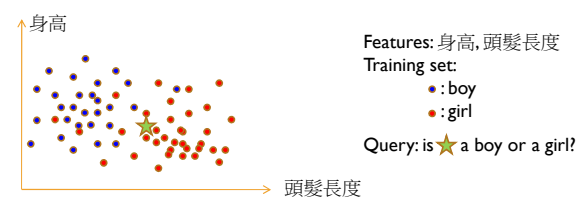
- Which one is a dog? and which one is a cat?



  - How do you tell?
- Two important pre-processing
  - Observation(feature selection): ear, nose, eyes, mouth, color…
  - Training: you must have seen hundreds (images) of cats and dogs before, and someone had told you which ones are cats and which ones are dogs.

## Machine learning

▶ How to make machine recognize things?
  ▶ Feature selection: measurable by computer or sensors
  ▶ Training: (there are tons of methods…)
▶ Example: statistical method



身高

Features: 身高, 頭髮長度
Training set:
  ● : boy
  ● : girl

Query: is ⭐ a boy or a girl?

頭髮長度

▶     Background: 線性代數，機率統計