Practical Image Matching

CS510
Lecture #18
March 22, 2002

2D Fourier Transform

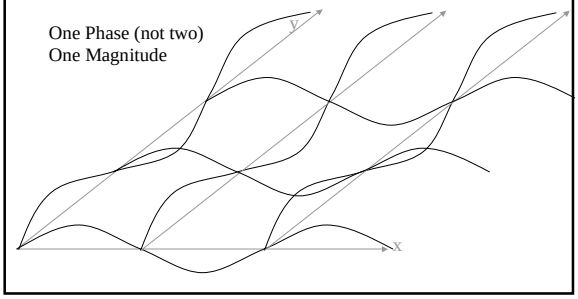
- *In answer to Aaron's question…(this slide is now part of the Sampling lecture, #8)*
- For 2D signals, the Fourier transform is essentially the same:

$$F(u,v) \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)[\cos(2\pi(ux+vy)) - i\sin(2\pi(ux+vy))]$$

- Note that frequencies are now two-dimensional (u= frequency in x, v = frequency in y)
- For every frequency (u,v), there is a single real and imaginary value

2D Frequency (Illustrated)

One Phase (not two)
One Magnitude



Practical Image Matching

- Idealized mugshot scenario:
  - All images of frontal human faces
  - Two diffuse illuminants, one on each side
    - But not always consistent
  - Blank backgrounds
    - Ideally, although not perfect. "Big Hair" also a problem
  - Scale, position, rotation approximately fixed
    - But subjects shift slighty, so images not registered
  - Task: match new mugshot to known fugitives, ex-convicts

Mugshot Processing

- General Approach: Subspace Projection
  - One-to-many
  - Efficiency matters (large data base)
- Problem #1: Images not registered
  - Near registration → similarity transform (rotation, translation, scale) is sufficient
  - Two points are enough
  - Common practice: identify eye centers by hand

Mugshot (II)

- Problem #2: Hair, background imperfections as distractors
  - All faces are approximately the same shape
  - After registration, apply an oval mask that eliminates all background and most hair pixels
- Problem #3: Inconsistent lighting
  - Created by non-uniform set-ups, minor changes in subject positions
  - Ideal solution: zero-mean, unit-length images
  - More often used: histogram equalization (similar results)

Mugshot Images (Processed)

- As a result, the actual images look like:

- Hand registration not totally accurate
- Preprocessing seems to obscure some useful information….
- A hack: many systems dismiss the first eigenvector, because it is often lighting

Mugshot Matching

- Once the subspace is computed
  - All gallery images projected into subspace
  - Probe (test) image projected into subspace
  - Nearest neighbor retrieval in subspace
    - Distance measures
      - L1: $\frac{1}{n}\sum_i |P_i - G_i|$
      - L2: $\frac{1}{n}\sqrt{\sum_i (P_i - G_i)^2}$
      - Angle: $\frac{P \cdot G}{|P \cdot G|}$
      - Mahalanobis: $\frac{1}{n}\sqrt{\sum_i \frac{(P_i - G_i)^2}{\sigma_i^2}}$

Opportunities for Improvement

- Better registration:
  - Use Fourier matching (to average face?) to register images.
  - Increase/decrease mask size
  - Add original average intensity/st dev as a feature.
  - Select subset of eigenvectors (search)
  - Select distance measure
  - Use alternate classifier (other than nearest neighbor)

Cat & Dog

- Cat & Dog is a harder scenario
  - Images are registered (by hand)
    - Not as “head on” as mugshots
  - Lighting is uncontrolled
  - No masking: cat & dog faces have different shapes!

HumanID at a distance

- Current Challenges:
  - Subjects at a distance (and non-cooperative)
    - In essence, just walking by
  - Uncontrolled (possibly outdoor) lighting
  - Occlusion
  - Subjects in motion

HumanID needs

- Face detection
  - Separate faces from other scene clutter
  - Scale, angle unknown
- Automatic face registration
  - Localize three or more points for registration
- Face recognition
  - Similar to mugshot scenario
- How?