Graph Theory and Video

So we've considered graphs where nodes are images, and edges are similarity relationships between images.

And we've considered graphs where nodes are pixels, and edges are similarities between pixels.

These both lead to graph theory based algorithms for finding clusters of related pixels or paths between distant images.

Graph theoretic algorithms are also appropriate for analyzing pre-processed images.

One hard task:

- Written character recognition.

- Why is it hard?

Astoundingly poorly defined.

- Why is it hard?
- Why is it hard?
  - Why is it hard?
- Why is it hard?
  - Why is it hard?
  - Why is it hard?
  - Why is it hard?
  - Why is it hard?

Strange definition

- Human handwriting is taught not to be "sufficiently close to some standard", but rather to be "interpretable by human readers" --- that is, (sometimes with great pain), we learn to write better and better until others can usually read our writing.

- There are some properties that seem consistent across different writing styles. Arms of letters, loops, "below the line" protrusions, etc.

MNIST Handwritten Digits

Matching with original and deformed prototypes

Prototype | Test | Error
---|---|---

Next 15 slides from a talk by Serge Belongie at ICCV 2001.
Deforming Biological Shapes

- D’Arcy Thompson: *On Growth and Form*, 1917
  - studied transformations between shapes of organisms

Finding correspondences between shapes

- Each shape is represented by a set of sample points
- Each sample point has a descriptor - the shape context
- Define cost $W_{ij}$ for matching point $i$ on first shape with point $j$ on second shape.
- Solve for correspondence as optimum assignment.

Shape Context

- Compact representation of distribution of points relative to each point

Comparing Shape Contexts

- Compute matching costs using Chi Squared Test:
  $$C_{ij} = \frac{1}{2} \sum_{k=1}^{n} (h_{ij}(k) - \hat{h}_{ij}(k))^2$$

- Recover correspondences by solving linear assignment problem with costs $C_{ij}$

[Jonker & Volgenant 1987]
What’s the linear assignment problem?

*Linear Assignment is weighted bipartite matching.*

Formal mathematical definition
The formal definition of the assignment problem (or linear assignment problem) is:

Given two sets A and B of equal size, together with a cost function C: A x B -> R
Find the bijection \( \pi \) such that

\[ \sum_{(a,b) \in \pi} C(a,b) \]

is minimized.

The problem is “linear” because the cost function to be optimized as well as all the constraints can be expressed as linear equations.

**Why is bipartite matching appropriate?**

Iterative Matching Example

Trademark Similarity

Measuring Shape Similarity

- Image appearance around matched points
  - color or gray-level window
  - orientation
- Shape context differences at matched points
- Bending Energy

COIL Object Database
Prototypes (cuz you don’t know any 3D information)

Prototype Selection: Coil-20

Hand-written Digit Recognition
- **MNIST 60 000**
  - linear: 12.0%
  - 40 PCA+ quad: 3.3%
  - 1000 RBF +linear: 3.6%
  - K-NN: 5%
  - K-NN (deskewed): 2.4%
  - K-NN (tangent dist.): 1.1%
  - SVM: 1.1%
  - LeNet 5: 0.95%

- **MNIST 600 000 (distortions):**
  - LeNet 5: 0.8%
  - SVM: 0.8%
  - Boosted LeNet 4: 0.7%

- **MNIST 20 000**
  - K-NN, Shape context matching: 0.63%

Problems with Shape contexts and graph matching?

So, what is actually used?

(If you can’t solve the problem, change it)

Pen Gesture Commands
- Might mean delete

Define a series of (hopefully) simple drawing gestures that mean different commands in a system, or characters.
Quikwriting.

Typing: “the”

Never caught on.

Special Alphabets

- Graffiti - Unistroke alphabet on Palm PDA
- Other alphabets or purposes
  - Gestures for commands

Reverse Engineered Palm Code?

- The principle for the Graffiti demo is that every stroke is normalized by bringing it back into a unit square and with a constant number of interpolations.
- The reference alphabet is pre-normalized to be the same.
- Then the input stroke is compared to each reference stroke, by computing the distance between them, point by point.
- The reference stroke that is found to be the closest is the match.
- A manual refinement step is then applied to improve the result for letters that looks alike (like L and h, or D and P).

Graffiti: What ambiguity?

- Sidelakes the issue completely
- You write the stroke the way you're supposed to
- If it mis-recognizes, the assumption is that it was your fault
- Blaming the user is argued by some to be the cause of its success.
- Closes the circle --- you keep changing the way that you write until your characters are recognized.

Human Handwriting Recognition:
Current State of Technology

- Generally based on neural network technology
- 5-6% error rate typical for isolated letters
- Good typists tolerate up to 1% error rate on keyboards that generate random errors
- Human subjects make 4-8% errors in isolated character reading, and 1.5% errors given context