Matching
ASCI course A20, feb 2003, part 1

Remco Veltkamp

Which algorithm?

depends on

- which similarity measure, depends on
- which required properties, depends on
- which particular matching problem, depends on
- which application

Which application?

- retrieval
- recognition and classification
- alignment, registration
- approximation

Which problem?

- computation problem: d(A,B)
- decision problem: d(A,B) $\leq \varepsilon$?
- decision problem: is there g: d(g(A),B) $\leq \varepsilon$?
- optimization problem: find g: min d(g(A),B)
- approximate optimization problem

Which properties?

- Metric properties
- Continuity
- Invariance

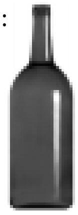
Metric Properties

- S set of patterns
- Metric: $d: S \times S \rightarrow R$ satisfying
  1. Self-identity: $\forall\ x \in S,\ d(x,x)=0$
  2. Positivity: $\forall\ x \neq y \in S,\ d(x,y)>0$
  3. Symmetry: $\forall\ x,\ y \in S,\ d(x,y)=d(y,x)$
  4. Triangle inequality: $\forall\ x,\ y,\ z \in S,\ d(x,z) \leq d(x,y)+d(y,z)$
- Semi-metric: 1, 2, 3
- Pseudo-metric: 1, 3, 4
- *S* with fixed metric *d* is called metric space

Symmetry

d(A,B) = d(B,A)

not always so for human perception

variant A:  prototype B: 

d(A,B) < d(B,A)

Triangle inequality

d(A,B)+d(B,C) ≥ d(A,C)

not always so for human perception,

in particular for partial matching:

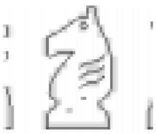
A:  B:  C: 

Continuity

Robustness

arbitrary small changes:

- deformation
- blurring
- cracks
- noise

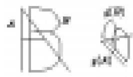


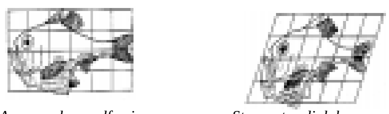
lead to arbitrary small change in similarity

Invariance

d(g(A),g(B)) = d(A,B)

for all g in transformation group G





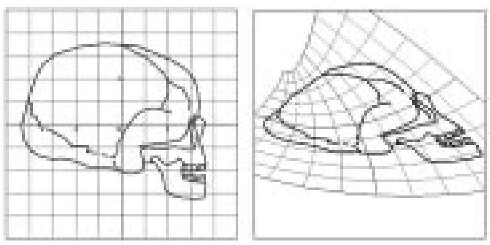
Argyropelecus olfersi *Sternoptyx dialphana*

(D'Arcy Thompson, 1911)

Invariance?

Invariance?

for large enough transformation group ...



Matching Feature Vectors

- Result of feature extraction: numerical values $x_1,\ldots,x_n$ assembled in a "feature vector" $x=(x_1,\ldots,x_n)$
- Example: 64 values for hue histogram, 8 for edge directions histogram, 16 for wavelet coefficients, 16 for Fourier coefficients of contour $\Rightarrow n=104$
- n-dimensional feature space

L_p metrics

- Also called Minkowski distance
- $x=(x_1,\ldots,x_n)$, $y=(y_1,\ldots,y_n)$

$$L_p = \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{1/p}$$

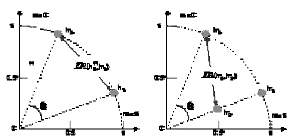
- Metric for $p \geq 1$ (otherwise no triangle inequality)
- L_1: taxicab, city block, Manhattan, rectilinear distance
- L_2: Euclidean distance
- L_∞: max, chess board distance, $\max_i |x_i - y_i|$

Cosine Distance

- $x=(x_1,\ldots,x_n)$, $y=(y_1,\ldots,y_n)$

$$d(x,y) = 1 - \cos(\angle(x,y)) = 1 - \frac{x\,y}{|x|\,|y|}$$

- Metric
- Only angle relevant, not vector lengths



Quadratic Form Distance

$$d(x,y) = \sum_{i=1}^{n}\sum_{j=1}^{n} |x_i - y_i|\, w_{ij}\, |x_j - y_j| = |x-y|^{\mathrm{T}}\, W\, |x-y|$$

- Metric if $w_{ij}=w_{ji}$ and $w_{ii}=1$

Earth Mover's Distance

- $A=\{(x_i,w_i)\}$, $\sum w_i=W$, $B=\{(y_j,u_j)\}$, $\sum u_j=U$
- f_{ij} flow from x_i to y_j over d_{ij}
- $f_{ij} \geq 0$
- $\sum_j f_{ij} \leq w_i$
- $\sum_i f_{ij} \leq u_i$
- $\sum_i \sum_j f_{ij} = \min(W,U)$

$$EMD(A,B) = \frac{\min_F \sum_{ij} f_{ij} d_{ij}}{\min(W,U)}$$

Properties EMD

- Invariant under rigid motion
- Respects scaling
- Metric if d metric, and $W=U$
- If $W \neq U$:
  - No positivity, surplus not taken into account
  - No triangle inequality

Proportional Transportation Dist

- $A=\{(x_i,w_i)\}$, $\sum w_i=W$, $B=\{(y_j,u_j)\}$, $\sum u_j=U$
- f_{ij} flow from x_i to y_j over d_{ij}
- $f_{ij} \geq 0$
- $\sum_j f_{ij} = w_i$
- $\sum_i f_{ij} \leq u_i W/U$
- $\sum_i \sum_j f_{ij} = W$

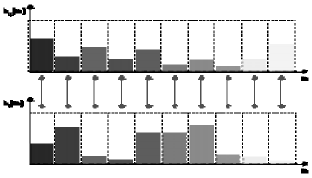
$$PTD(A,B) = \frac{\min_F \sum_{ij} f_{ij} d_{ij}}{W}$$

Properties PTD

- Invariant under rigid motion
- Respects scaling
- PTD is pseudo-metric:
  - Triangle inequality holds
  - No positivity
    - but only when same relative weights
    - surplus taken into account

Histogram Matching

- Histogram seen as feature vector e.g. $d(H_1,H_2)$ is Euclidean distance

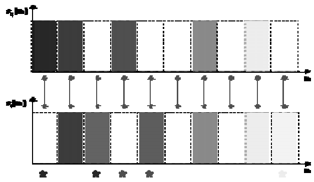


Hamming Distance

- Binary vectors/histograms

$$d(H_1,H_2) = \sum_{i=1}^{n} |H_1[i] - H_2[i]| = \sum_{i=1}^{n} (H_1 \ XOR \ H_2)$$

- Normalization: divide by n



Histogram Intersection

- To check occurrence of object in region
- Typically $\sum H_{obj}[i] < \sum H_{reg}[i]$
- Non-metric form (not symmetric):

$$d(H_{obj},H_{reg}) = 1 - \frac{\sum_{i=1}^{n} \min(H_{obj}[i], H_{reg}[i])}{\sum_{i=1}^{n} H_{obj}[i]}$$

Histogram Intersection

- Metric form:

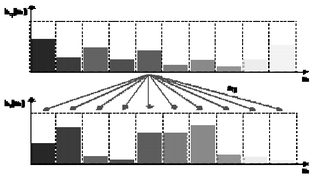
$$d(H_{obj},H_{reg}) = 1 - \frac{\sum_{i=1}^{n} \min(H_{obj}[i], H_{reg}[i])}{\min\left(\sum_{i=1}^{n} H_{obj}[i], \sum_{i=1}^{n} H_{reg}[i]\right)}$$

- If $\sum H_{obj}[i] = \sum H_{reg}[i]$
  $d(H_{obj},H_{reg}) = L_1(H_{obj},H_{reg})$

Histogram Matching

- Cross bin matching:

$$d(H_1,H_2)=\left(\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\,|H_1[i]-H_2[j]|^p\right)^{1/p}$$



2D Histogram Matching

- Instead of $\sum_{i=1}^{n} \ldots H_1[i] \ldots H_2[i] \ldots$ (2n terms)

  do $\sum_{i=1}^{n}\sum_{j=1}^{n} \ldots H_1[i,j] \ldots H_2[i,j] \ldots$ (n^2 terms)

Earth Mover's Distance

On histograms

- $H_1=\{(i,H_1[i])\}$, $\sum H_1[i]=W$, $H_2=\{(i,H_2[i])\}$, $\sum H_2[i]=U$
- f_{ij} flow from $H_1[i]$ to $H_2[j]$ over $d_{ij}=|i-j|$
- $f_{ij} \geq 0$
- $\sum_j f_{ij} \leq H_1[i]$
- $\sum_i f_{ij} \leq H_2[i]$
- $\sum_i \sum_j f_{ij} = \min(W,U)$

$$EMD(A,B)=\frac{\min_F \sum_{ij} f_{ij} d_{ij}}{\min(W,U)}$$

Kullback-Leibler

- Consider histograms as distibrutions: $\Sigma\, H_1[i] = \Sigma\, H_2[i] = 1$, $H_1[i], H_2[i] \geq 0$

$$d(H_1,H_2)=\sum_{i=1}^{n}\left(H_1[i]\log\left(\frac{H_1[i]}{H_2[i]}\right)\right)$$

- No metric: not symmetric

Divergence

- Symmetrized Kullback-Leibler

$$d(H_1,H_2)=d_{KL}(H_1,H_2)+d_{KL}(H_2,H_1)$$
$$=\sum_{i=1}^{n}\left((H_1[i]-H_2[i])\log\left(\frac{H_1[i]}{H_2[i]}\right)\right)$$

Mahanalobis Distance

- Quadratic form with inverse covariance matrix: d(x,y)= (x-y) Σ^{-1}(x-y)
- Of N feature vectors $\mathbf{x}^1,\ldots,\mathbf{x}^N$, each of length n, compute

$$\mu_i=\frac{1}{N}\sum_{j=1}^{N}\left(\mathbf{x}^j\right)_i,\ i=1,\ldots,n$$

$$\sigma_{ij}=E((x_i-\mu_i)(x_j-\mu_j))=\sum_{k,l=1}^{n}\frac{(x^k{}_i-\mu_i)(x^l{}_j-\mu_j)}{n^2}$$

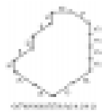
- $\Sigma = (\sigma_{ij})$

Edit Distance

- Feature interpreted as string of characters
- Edit distance operations
  - Insertion, where an extra character is inserted into the string
  - Deletion, where a character has been removed from the string
  - Transposition, in which two characters are reversed in their sequence
  - Substitution, which is an insertion followed by a deletion

Edit Distance

- Strings with a small edit distance are likely to be similar
- Edit distance is number of edit distance operations from one string to another
- Example: chaincodes
  12345678
  123845677 have distance 3



Edit Distance

- Use dynamic programming
- Given two strings $x = x_1x_2 .. x_n$ and $y = y_1y_2 ..y_m$
- edit distance *f(i, j)* is computed as best match of two substrings $x_1x_2 .. x_i$ and $y_1y_2 ..y_j$ where
  - $f(0,0) = 0$
  - $f(i, j) = min[f(i-1, j) + 1,$
    $f(i, j-1) +1, f(i-1, j-1) + d(x_i,y_j)]$

Merging Similarities

Linear weighting

- Combine *k* different feature distances d_i, e.g. color, texture and shape distances
- Linear weighting: $d = \sum_{i=1}^{k} w_i d_i$
  (weighted average)
- Affine weighting: $\sum_{i=1}^{k} w_i = 1$
- Convex weighting: $\sum_{i=1}^{k} w_i = 1, w_i \geq 0$

Merging Similarities

Non-linear weighting

- α-trimmed mean: weight only α percent highest of the *k* values
- Rank-based merging: sort values in decreasing order $d_1', \ldots, d_k'$

$$d = \frac{1}{k}\sum_{j=1}^{k} \frac{1}{j}\sum_{i=1}^{j} d_k'$$