

Shape

ASCI course A20, feb 2003, part 2

Remco Veltkamp

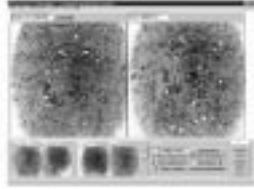


Why shape

- User interviews show:
for object retrieval, shape is more important than color and texture
- Problem: effective shape is more difficult than effective color and texture
- Check yourself: e.g. with Blobworld:
<http://elib.cs.berkeley.edu/photos/blobworld/start.html>

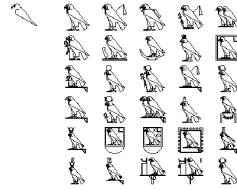
Typical problems

- What is the matching transformation?
- No one-to-one correspondence
- Occlusion
- Noise



Typical problems

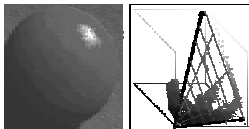
- Partial match: only part of query appears in part of database shape



What is shape?

Plato, "Meno", 380 BC:

*"figure is the only existing thing that is found
always following color"*



What is shape?

"terms employed in geometrical problems":

"figure is limit of solid"



Levels of Shape

- Global:
 - image
 - object
- Local (weak):
 - image
 - object

Types of Object Shape

- Point pattern
- Curve
- Contour
- Region

Global Image

- Fourier transform
- Discrete cosine transform (used in jpeg)
- Wavelet transform



Global Image

- Transformation from color in spatial domain to color variation in frequency domain
- No explicit shape encoding, but intensity transitions at object boundaries
- Because whole image is encoded, no matching of individual shapes

Global Object

- Matching of a single, whole, object
- Features:
 - Elementary descriptors
 - Moments
 - Modal matching
 - Fourier descriptors
 - Curvature scale space
- Drawbacks:
 - Relies on perfect segmentation
 - Sensitive to noise and occlusion

Local Object

- Matching local descriptors
- Weak segmentation
- Features:
 - Corners, relative position
 - Corners, curvature

Global, Region
Elementary descriptors

- Area A
- Perimeter l
- Compactness $c=l^2/(4\pi A)$
- Circularity, roundness l/c
- Centroid (center of mass)
- Major and minor axes λ_1, λ_2
- Eccentricity $\|\lambda_1\|/\|\lambda_2\|$
- Minimal bounding box area $A_m = h b$
- Rectangularity A/A_m

Global, Object
Moments

- Object uniquely defined by infinite sequence of moments $M_{p,q}$:

$$m_{p,q} = \int_{Obj} x^p y^q dx dy$$

- In terms of pixels $[1,n] \times [1,m]$ image $f(x,y)$:

$$m_{p,q} = \sum_{x=1}^n \sum_{y=1}^m f(x,y) x^p y^q$$

Global, Object
Moment invariants

- Translation invariant: central moments

$$\mu_{p,q} = \int_{Obj} \sum_x x^p - \frac{m_{1,0}}{m_{0,0}} \sum_x x^p \sum_y y^q - \frac{m_{0,1}}{m_{0,0}} \sum_x x^p \sum_y y^q dx dy$$

- Invariant under uniform scaling with factor α :

$$\eta_{p,q} = \frac{\mu_{p,q} / \alpha^{p+q+2}}{\mu_{0,0}^{(p+q+2)/2}}$$

Global, Object
Moment invariants

Rotation invariant:

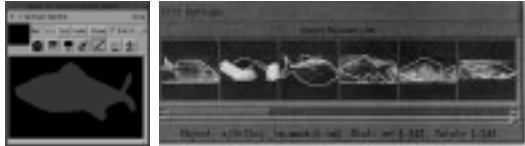
$$\begin{aligned} \phi_1 &= \mu_{20} + \mu_{02} \\ \phi_2 &= (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 \\ \phi_3 &= (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \\ \phi_4 &= (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2 \\ \phi_5 &= (\mu_{30} - 3\mu_{12}) + (\mu_{30} + \mu_{12}) [(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + \\ &\quad (3\mu_{21} - \mu_{03}) + (\mu_{21} + \mu_{03}) [3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \\ \phi_6 &= (\mu_{20} - \mu_{02}) [(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11} (\mu_{30} + \mu_{12}) (\mu_{21} + \mu_{03}) \end{aligned}$$

Reflection and rotation invariant:

$$\begin{aligned} \psi_7 &= (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2 + \\ &\quad (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03}) [3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \end{aligned}$$

Global, Object
Example: QBIC

Features:
Area, compactness, eccentricity, major axis orientation, moment invariants up to degree 8



Curves, Contours
Modal Matching

- Take n samples along contour
- Matrix $D=(d_{ij})$ describes interaction between points i and j
- Determine n modes, or eigenshapes, the eigenvectors of D :

$$D e_i = \lambda e_i$$

Curves, Contours
Modal Matching

- n modes e_i of query, n modes e_j' of target
- Some dissimilarity measures between modes: $m(e_i, e_j')$
- For fixed i_0 , determine value j_0 of j for which $m(e_{i_0}, e_{j_0}')$ is minimal
- If i for which $m(e_i, e_{j_0}')$ is minimal is i_0 , then point i of query and point j of target match

Curves, Contours
Fourier Descriptors

- FD of some shape signature:
 - Complex Coordinates: $z(t)$
 - Central Distance: $r(t)$
 - Chordlength: $r^*(t)$
 - Curvature: $K(t)$
 - Cumulative Angles: $\varphi(t)$
 - Area function: $A(t)$

Complex Coordinates

$$z(t) = [x(t) - x_c] + i[y(t) - y_c]$$

$$x_c = \frac{1}{N} \sum_{t=0}^{N-1} x(t), \quad y_c = \frac{1}{N} \sum_{t=0}^{N-1} y(t)$$

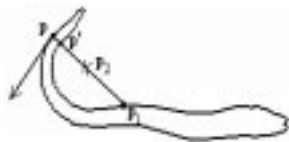
Central Distance

$$r(t) = ([x(t) - x_c]^2 + [y(t) - y_c]^2)^{1/2}$$



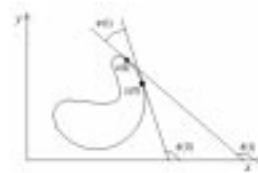
Chordlength

$r^*(t)$ = length of chord in object perpendicular to tangent at p , as a function of p



Cumulative Angular Function

- Also called turning angle function
- $\varphi(t) = [\theta(t) - \theta(0)] \bmod(2\pi)$

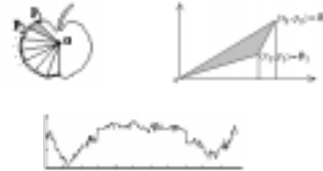


Curvature Function

- $K(t) = \theta(t) - \theta(t-1)$
- $\theta(t) = \arctan \frac{y(t) - y(t+w)}{x(t) - x(t+w)}$
- w is jumping step in selecting next pixel

Area Function

$$A(t) = \frac{1}{2} |x_1(t)y_2(t) - x_2(t)y_1(t)|$$



Fourier Descriptors

- **Fourier transform of the signature $s(t)$**

$$u_n = \frac{1}{N} \sum_{t=0}^{N-1} s(t) \exp\left(\frac{-j2\pi nt}{N}\right)$$

- $u_n, n = 0, 1, \dots, N-1$, are called FD denoted as FD_n

- **Normalised FD**

$$\mathbf{f} = \left[\frac{|FD_1|}{|FD_0|}, \frac{|FD_2|}{|FD_0|}, \dots, \frac{|FD_m|}{|FD_0|} \right]$$

Where $m=N/2$ for central distance, curvature and angular function
 $m=N$ for complex coordinates

FD Convergence Speed

- Finite number of coefficients are used to approximate the signal. Partial Fourier sum of degree n of $u(t)$ is given by

$$(S_n u)(t) = \sum_{|k| \leq n} \hat{u}(k) e^{jkt}$$

- For piecewise smooth function $u(t)$, there exists one-to-one correspondence between $u(t)$ and the limit of their Fourier series expansion

$$\lim_{n \rightarrow \infty} (S_n u)(t) = u(t)$$

- For shape retrieval application, number of coefficients to represent shape should not be large, therefore, the convergence speed of the Fourier series derived from the signature function is crucial

FD Convergence Speed

- How fast get the FD coefficients below a threshold?

Signature functions	Number of normalized FD coefficients > 0.1	Number of normalized coefficients > 0.01
$r(t)$	15	120
$r^2(t)$	40	360
$A(t)$	20	210
$\theta(t)$	10	50
$\Psi(t)$	40	280
$\Phi(t)$	∞	∞

← fast

← slow

FD Matching

- Similarity between a query shape and a target shape in the database is

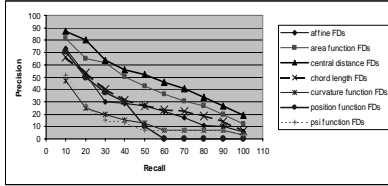
$$d = \left(\sum_{i=1}^m (f_i^q - f_i^t)^2 \right)^{1/2}$$

where $\mathbf{f}_q = (f_q^1, f_q^2, \dots, f_q^m)$ and $\mathbf{f}_t = (f_t^1, f_t^2, \dots, f_t^m)$ are the feature vectors of the two shapes respectively

FD Performance

- Precision P is the ratio of the number of relevant retrieved shapes r to the total number of retrieved shapes n
- Recall R is the ratio of the number of relevant retrieved shapes r to the total number m of relevant shapes in the whole database

$$P = \frac{r}{n} \quad R = \frac{r}{m}$$



Curves, Contours

Local descriptors

To process the shape locally on saliency: search for local maximum $\{s(t)\}$

Supported by perceptual evidence that:

- humans focus on high curvature points
- humans keep the memory of figure by a few, salient points
- humans are very poor in quantitative measurement of overall shape

Curves, Contours Local descriptors

Location: $\mathbf{x}(t)$

Tangent: $\phi(t) = \mathbf{x}'(t)$

Curvature: $\kappa(t) = \phi'(t) = \mathbf{x}''(t)$



Curves, Contours

Curvature Scale Space

- Smooth curvature along contour



Curves, Countours Curvature Scale Space

- Contour $C(s) = (x(s), y(s))$
- Convolution with Gaussian kernel of width σ :

$$x_\sigma(s) = x(s) * \phi_\sigma(t) = \int x(s)\phi_\sigma(t-s) dt$$

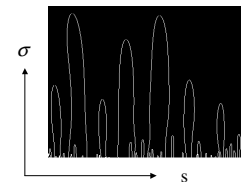
$$\phi_\sigma(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{t^2}{2\sigma^2}}$$

same for $y(s)$

Curvature Scale Space

- Increasing σ : positions of zero crossings move together, then annihilate
- Number of zero crossings decreases until contour is convex
- Matching: match points of annihilation in (s, σ) -plane:

<http://www.ee.surrey.ac.uk/Research/VSSP/imagedb/demo.html>



Hausdorff distance

- Finite point sets A, B:
 $d(A,B) = \max_{a \in A} \min_{b \in B} d(a,b)$
- Infinite point sets (curves, regions):
 $d(A,B) = \sup_{a \in A} \inf_{b \in B} d(a,b)$
- Hausdorff distance $H(A,B) = \max\{d(A,B), d(B,A)\}$

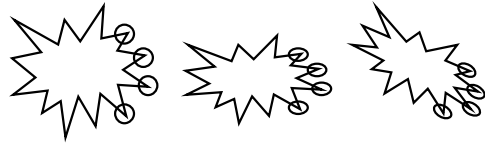
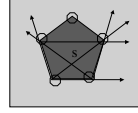


Local Object Affine Invariant Ratio

Affine invariance :
 given 4 points a, b, c, d and
 s the intersection of ac and bd,
 then are

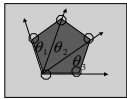
$$\frac{|as|}{|ab|} \text{ and } \frac{|cs|}{|cd|}$$

invariant



Local Object Projective Invariant Ratio

Projective invariance: cross ratio



Projective invariance (5 points) :
 $\frac{\sin(\theta_1 + \theta_2) \sin(\theta_2 + \theta_3)}{\sin(\theta_2) \sin(\theta_1 + \theta_2 + \theta_3)}$



Which similarity?

discrete metric:

$$d(A,B) = \begin{cases} 0 & \text{if A equals B} \\ 1 & \text{otherwise} \end{cases}$$

exact congruence matching
 metric, invariant under all homeomorphisms!
 but lacks robustness properties

Which algorithm?

class of algorithms:

- voting schemes
 - alignment (Ullman, Huttenlocher)
 - geometric hashing (Wolfson et al.)
 - generalized Hough transform (Ballard)
 - pose clustering (Stockman)

Which algorithm?

class of algorithms:

- subdivision schemes
 - decision problem, translation+scaling (Huttenlocher, Rucklidge)
 - optimization problem, affine transformation (Hagedoorn, Veltkamp)
 - combination with alignment (Mount)
 - combination with matchlist (Breuel)

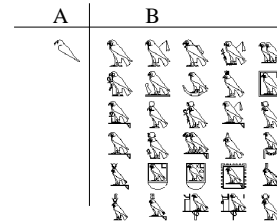
Example 1: application

6800 hieroglyphs, 72000 polylines
from Center for Computer-aided
Egyptological Research:



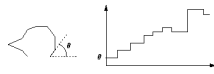
Example 1: problem

find B's for which there is g: d(g(A),B) ≤ ε ?



Example 1: similarity

turning function distance

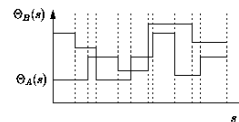


similarity is function distance

Example 1: similarity

turning function distance

$$d(A, B) = \sum_{\theta} \min_{t, \theta} \int_0^1 |\Theta_A(s+t) - \Theta_B(s) + \theta|^p ds$$

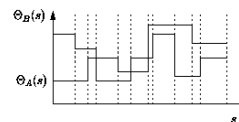


Example 1: properties

- triangle inequality
- translation, rotation invariant
- deformation robust

Example 1: algorithm

- $\theta^*(t) = \int g(s) ds - \int f(s) ds - 2\pi t$
- Naive evaluation: compute each of $O(mn)$ shifts in $O(m+n)$ time
- Incremental evaluation: $O(mn \log mn)$ time



Example 2: Weighted Point Set

- Certainty of position
low weights should match easier
- Certainty of existence
high weights should match easier
- Amount of some property
match one distribution with another
- This example: amount of curvature

Example 2: similarity

Earth Mover's Distance between
 $\{(p_1, w_1), \dots, (p_m, w_m)\}$ and $\{(q_1, u_1), \dots, (q_m, u_m)\}$

$$d = \frac{\min_F \sum_{i=1}^n \sum_{j=1}^m f_{ij} d_{ij}}{\min(\sum w_i, \sum u_j)}$$

$$f_{ij} \geq 0$$

$$\sum_j f_{ij} \leq w_i, \sum_i f_{ij} \leq u_j$$

$$\sum_i \sum_j f_{ij} = \min(\sum w_i, \sum u_j)$$

no triangle inequality for unequal total weights

Example 2: similarity

transportation distance between
 $\{(p_1, w_1), \dots, (p_m, w_m)\}$ and $\{(q_1, u_1), \dots, (q_m, u_m)\}$

$$d = \frac{\min_F \sum_{i=1}^n \sum_{j=1}^m f_{ij} d_{ij}}{\sum_i w_i}$$

$$f_{ij} \geq 0, \sum_i \sum_j f_{ij} = \sum_i w_i$$

$$\sum_i f_{ij} = \frac{u_j \sum_i w_i}{\sum_j u_j}, \sum_j f_{ij} = w_i$$

triangle inequality for unequal total weights

Example 2: PTD Result



Example 2: EMD result



With EMD, 2nd apple not as 2nd match

Example 2: PTD Result

demo



Example 2 Counter example



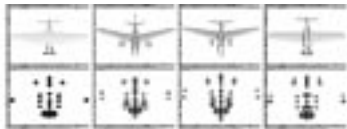
Example 2 3D models

- Polyhedral models
- In fixed grid, select those vertices with high curvature

Experimental results



"Princeton database": 133 models classified by function into 25 classes



"Utrecht database": 512 models classified by shape into 6 classes

Example 2:
PTD
Result

demo

Model ID	Class	Result
Model 1413 (Airplane)	1413	
Model 1414 (Airplane)	1414	
Model 1415 (Airplane)	1415	
Model 1416 (Airplane)	1416	
Model 1417 (Airplane)	1417	