Affine-Invariant Curve Normalization for Shape-Based Retrieval

Yannis Avrithis, Yiannis Xirouhakis and Stefanos Kollias

Image, Video & Multimedia Systems Laboratory
Department of Electrical and Computer Engineering
National Technical University of Athens
Heroon Polytechniou 9, 157 73 Zographou, Greece

Problem Statement

- Two-dimensional curve *normalization* with respect to affine transformations
- Affine-invariant curve representation *without loss of information* on the original curve
- *Decouple* affine-invariant description from feature extraction and pattern matching
- *Pre-processing* for shape representation, classification, recognition or retrieval (e.g. shape matching using deformable templates)

Existing Solutions

- Several affine-invariant methods available (B-splines polygonal approximation, chain coding, moments, Fourier descriptors etc.)
- In most of them, invariance *embedded* in the process of matching, recognition, or similarity measure estimation
- Alternatively, *matching* two given curves by optimally evaluating their affine parameters: high computational cost and requirement of *a priori* knowledge of both shape instances

The Proposed Technique

- Curves estimated from object contours modelled by cubic *B-splines* : shape simplified and segmentation noise reduced
- Several *normalization* steps to eliminate translation, scaling, skew, starting point, rotation and reflection transformations
- Normalization based on a combination of curve features including *moments* and *Fourier descriptors*
- All features *globally* estimated
- Computational complexity negligible

Assumptions

- Object contour shape available as a set of ordered points forming a 2-D planar closed curve
- Shapes obtained from image data by means of manual or automatic segmentation
- *M-RSST color segmentation* algorithm employed, combined with motion segmentation in case of video sequences
- *No occlusion* between objects

B-Spline Curve Modelling

- Employed to reduce segmentation noise and obtain *uniform sampling* in terms of arc length
- *Control points* determined by fitting the B-spline to data points in a MMSE sense
- *Knot points* derived from a linear combination of the estimated control points
- Parametric value obtained using the *Chord Length* (CL) method; knot points re-allocated with equal spacing in terms of the estimated parametric value

Curve Orthogonalization

- Normalization with respect to translation, skew and scaling - reduces affine transformations to orthogonal ones
- 2-D curve $\mathbf{s} = [\mathbf{s}_0\ \mathbf{s}_1\ \dots\ \mathbf{s}_{N-1}]$ represented by its horizontal and vertical coordinates $\mathbf{x} = [x_0\ x_1\ \dots\ x_{N-1}]$ and $\mathbf{y} = [y_0\ y_1\ \dots\ y_{N-1}]$
- (p, q)-order moments of order up to 2 used:

$$m_{pq}(\mathbf{s}) = \frac{1}{N}\sum_{i=0}^{N-1} x_i^p y_i^q$$

Orthogonalization Steps

- 1. Translation $\mathbf{x}_1 = \mathbf{x} - \mu_x, \quad \mathbf{y}_1 = \mathbf{y} - \mu_y$
- 2. Scaling $\mathbf{x}_2 = \sigma_x \mathbf{x}_1, \quad \mathbf{y}_2 = \sigma_y \mathbf{y}_1$
- 3. Rotation $\mathbf{s}_3 = \mathbf{R}_{\pi/4} \mathbf{s}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{x}_2 - \mathbf{y}_2 \\ \mathbf{x}_2 + \mathbf{y}_2 \end{bmatrix}$
- 4. Scaling $\mathbf{x}_4 = \tau_x \mathbf{x}_3, \quad \mathbf{y}_4 = \tau_y \mathbf{y}_3$

where

$$\sigma_x = 1 \Big/ \sqrt{m_{20}(\mathbf{s}_1)} \qquad \sigma_y = 1 \Big/ \sqrt{m_{02}(\mathbf{s}_1)}$$

$$\tau_x = 1 \Big/ \sqrt{m_{20}(\mathbf{s}_3)} \qquad \tau_y = 1 \Big/ \sqrt{m_{02}(\mathbf{s}_3)}$$

Orthogonalization Results

- Normalized curve $n_a(\mathbf{s})$ has the properties

$$m_{10}(n_a(\mathbf{s})) = m_{01}(n_a(\mathbf{s})) = m_{11}(n_a(\mathbf{s})) = 0$$

$$m_{20}(n_a(\mathbf{s})) = m_{02}(n_a(\mathbf{s})) = 1$$

- For two curves related through an affine transformation

$$\mathbf{s}' = \mathbf{A}\mathbf{s} + \mathbf{t} = \begin{bmatrix} \mathbf{x}' \\ \mathbf{y}' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$$

there exists an orthogonal 2×2 matrix **Q** s.t.

$$n_a(\mathbf{s}') = \mathbf{Q}\, n_a(\mathbf{s})$$

Starting Point Normalization

- Complex vector notation employed:
  $\mathbf{z} = \mathbf{x} + j\,\mathbf{y} = [z_0\ z_1\ \ldots\ z_{N\text{-}1}]^T$
- Calculate the Discrete Fourier Transform

$$u_k = \sum_{i=0}^{N-1} z_i w^{-ki}, \quad k = 0,1,\ldots,N-1$$

- Define a *standard* circular shift using the first and last of the Fourier phases $a_k = \operatorname{Arg} u_k$

$$p(\mathbf{z}) = \left\lfloor \frac{N}{4\pi}(a_1 - a_{N-1}) \right\rfloor \bmod N/2$$

Starting Point Results

- Given a curve circularly shifted with respect to **z** by m samples, $m \in \{0,1,\ldots,N-1\}$

$$\mathbf{z}' = S_m(\mathbf{z}): \quad z'_i = z_{(i+m)\,\mathrm{mod}\,N}, \quad i = 0,1,\ldots,N-1$$

the following hold for $n_p(\mathbf{z}) = S_{-p(\mathbf{z})}(\mathbf{z})$:

$$p(n_p(\mathbf{z}')) = p(n_p(\mathbf{z})) = 0$$

$$n_p(\mathbf{z}') = \begin{cases} n_p(\mathbf{z}), & 0 \le p(\mathbf{z}) + m < N/2 \\ S_{N/2}(n_p(\mathbf{z})), & N/2 \le p(\mathbf{z}) + m < N \end{cases}$$

Rotation/Reflection Normalization

- Rotation normalization: $\mathbf{z}_1 = \mathbf{z}\, e^{-jr(\mathbf{z})}$
  where
  $$r(\mathbf{z}) = \left(\frac{1}{2}(a_1 + a_{N-1})\right) \bmod \pi$$
- Reflection normalization:
  $$n_r(\mathbf{z}) = \mathbf{z}_2 = v_x(\mathbf{z}_1)\mathbf{x}_1 + jv_y(\mathbf{z}_1)\mathbf{y}_1$$
  where

$$v(\mathbf{z}_1) = v_x(\mathbf{z}_1) + jv_y(\mathbf{z}_1) = \operatorname{sgn} m_{12}(\mathbf{z}_1) + j \operatorname{sgn} m_{21}(\mathbf{z}_1)$$

Rotation/Reflection Results

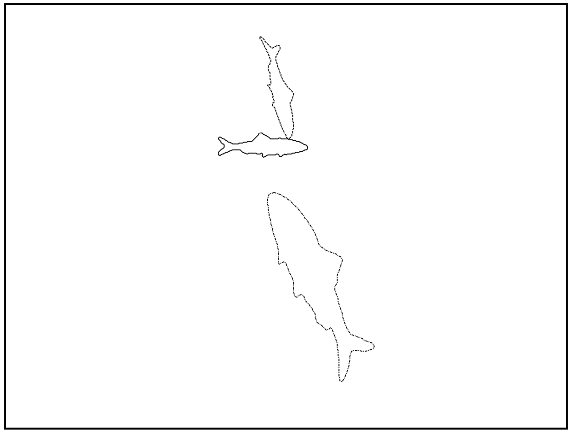
- For two curves **z**, **z′** orthogonalized and normalized w.r.t. starting point:
  $\mathbf{z}' = (s_x \mathbf{x} + j s_y \mathbf{y}) e^{j\theta}$
  where $s_x = \pm 1$, $s_y = \pm 1$ and $\theta \in [0, \pi)$, the following hold:

$$r(n_r(\mathbf{z}')) = r(n_r(\mathbf{z})) = 0$$

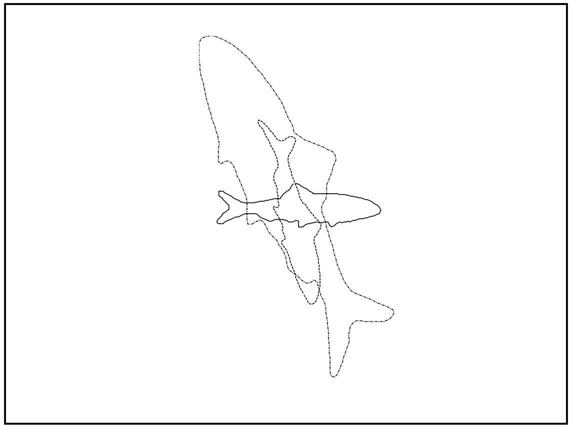
$$v_x(n_r(\mathbf{z}')) = v_y(n_r(\mathbf{z}')) = v_x(n_r(\mathbf{z})) = v_y(n_r(\mathbf{z})) = 1$$

$$n_r(\mathbf{z}') = n_r(\mathbf{z})$$

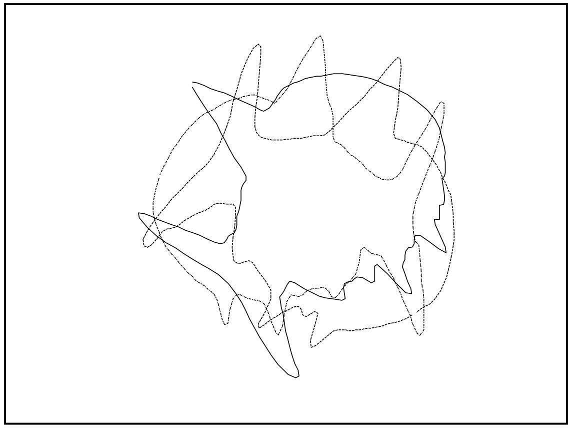
Results: Same Object



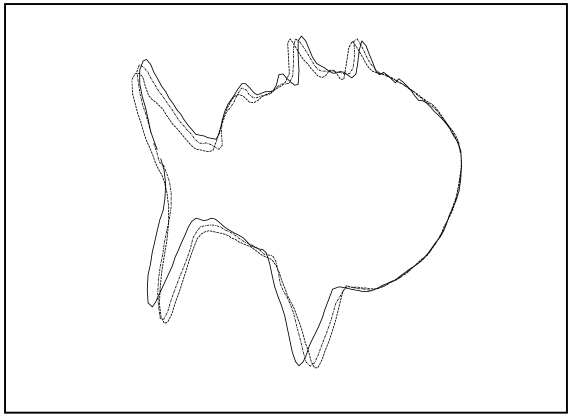
Original



Translation

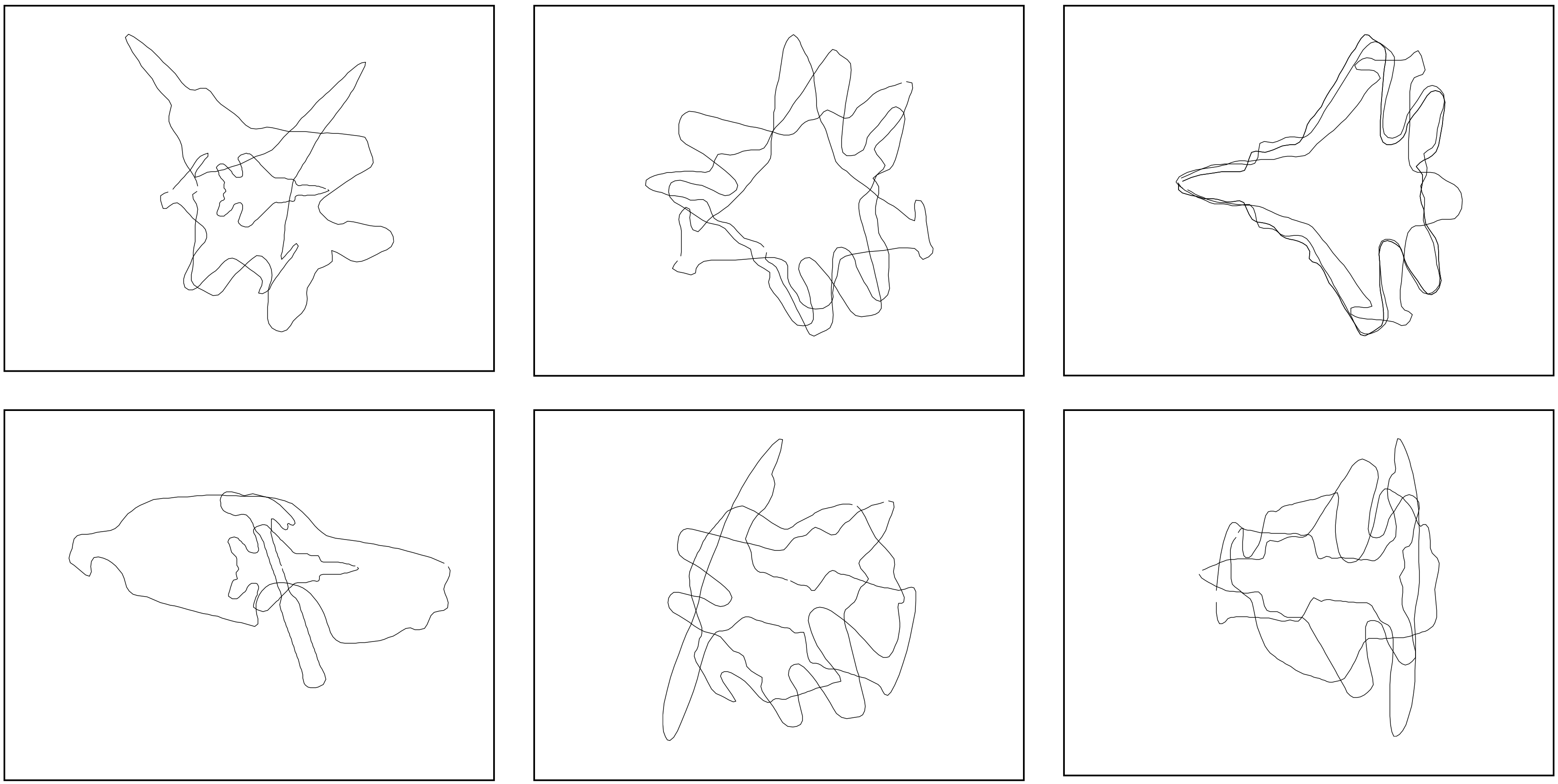


Skew/Scaling



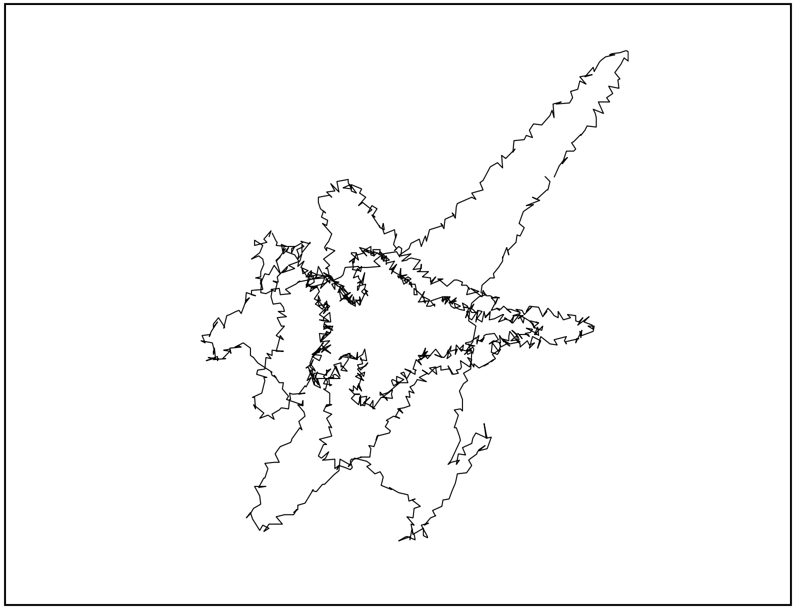
Rotation

Results: Similar/Different Objects

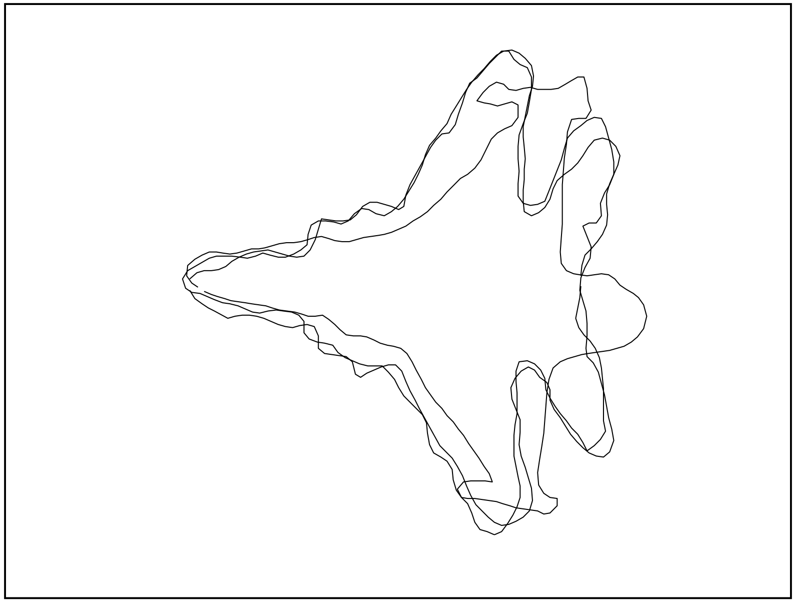


Translation Skew/Scaling Rotation

Results: Noise Effect



Original



Normalized

Contour Distance Measures

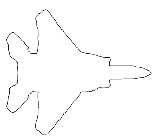
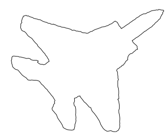
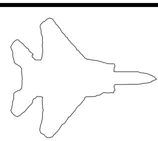
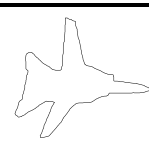
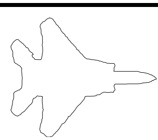
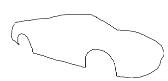
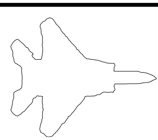
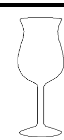
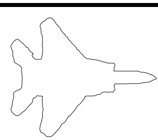
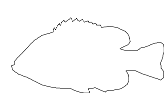
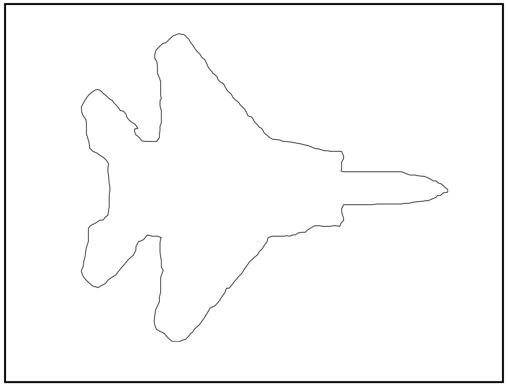
Contour 1	Contour 2	Points	FD	MFD
		0.01	0.02	0.01
		0.19	0.12	0.11
		0.75	0.41	0.57
		0.89	0.62	0.65
		0.76	0.25	0.32

Image Retrieval Results

Input Image



Extracted Contour

Retrieved images in descending contour similarity

Conclusions

- Shapes normalized to *standard* position: all affine transformations of the same object also normalized to the same position
- Apart from affine transformation parameters, *no other information discarded*
- Successful for *content-based retrieval* from image / video databases employing a number of curve similarity measures
- Considerably *robust* to noise and shape deformations
- Easy integration into *real-time systems*