# Fourier Descriptors

and

*Object Recognition Overview*



Computer Vision Laboratory

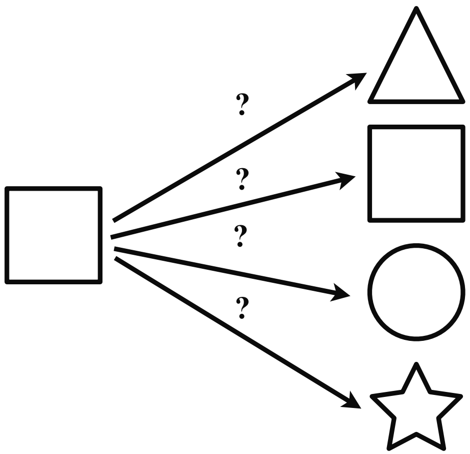
## Outline

- Part I: Fourier Descriptors with Applications
  - Shape matching
  - Fourier Descriptors (FDs)
  - Matching by Correlating FDs
  - Traffic Sign Recognition
- Part II: Object Recognition

LiU

Computer Vision Laboratory

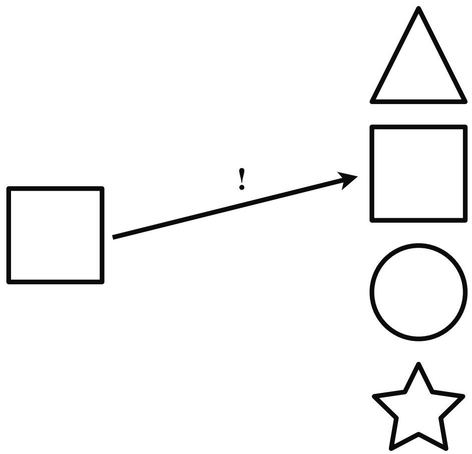
## Shape Matching




LiU

Computer Vision Laboratory

## Shape Matching




LiU

# Shape Matching

First problem in a real world scenarios is how to acquire the geometrical shapes, i.e. how to acquire the contours.

Second problem is how to match the geometrical shapes

We will only address the second part here. We assume that we have some means of extracting contours.





# Shape Matching

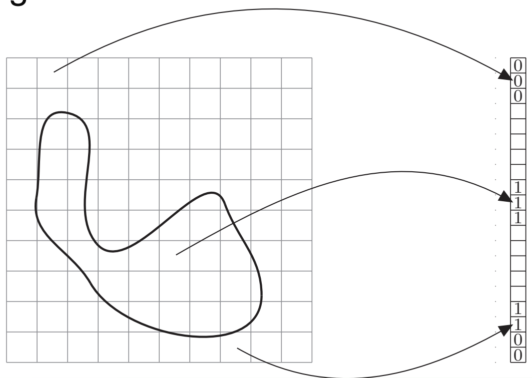
- Region based approaches
  - Capturing information inside the boundary
- Contour based approaches
  - Capturing information regarding the boundary **only**





# Shape Matching

- Region based approaches
  - E.g. Grid based




Match shapes by estimate distance between vectors





# Shape Matching

- Contour based approaches
  - E.g. Fourier Descriptors (p. 818 - 821)

# Fourier Descriptors

- Apply the Fourier transform to a periodic 1D parameterization of the contour.
- Results in a shape descriptor in the frequency domain.

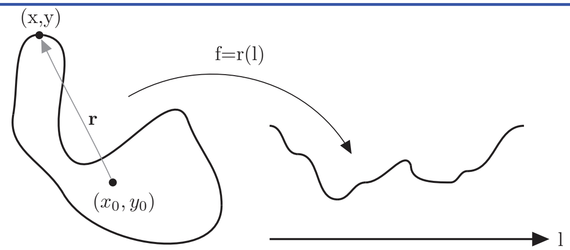
Granlund,G.H.: Fourier Preprocessing for Hand Print Character Recognition. IEEE Trans. on Computers C–21(2)(1972)195–201

Zahn, C.T and Roskies, R.Z.: Generic Fourier Descriptor for Shape Based Image Retrieval. IEEE Trans. on Computers C–21(2)(1972)269–281





# Shape Signature




- Distance to the centroid
- Tangent direction
- Complex valued

$$c(l) = c(l+L) = x(l) + iy(l)$$





# Fourier Descriptors

- Given a contour c(l) the n:th FD coefficient is given according to:

$$f_n = \frac{1}{N}\int_{l=0}^{L} c(l) exp(-\frac{2\pi inl}{L})dl$$

Match shapes by estimate distance between vectors





# Reasons for popularity

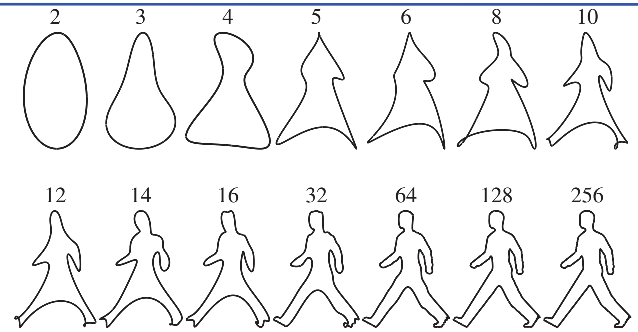
- Good matching performance
- Easy to achieve invariance to common transformations
- Easy to implement
- Easy to interpret

# Easy to Interpret

- **Low frequency** components contain information about the **general shape** of a contour
- **High frequency** components contain information about the **fine details** of a contour



# Easy to Interpret






# FD Invariances

- Invariance to translation, scale, rotation and index-shift can be obtained easily.



# FD Invariances

- Invariance to translation, scale, rotation and index-shift can be obtained easily.

$$\text{Translation: } c(l) + T$$

$$f_n^T = f_n + \frac{1}{N}\int_{l=0}^{L} T\exp(-\frac{2\pi i n l}{L})dl$$

$$\frac{1}{N}\int_{l=0}^{L} T\exp(-\frac{2\pi i n l}{L})dl = \begin{cases} \frac{1}{N}\int_{l=0}^{L} T dl \neq 0 & n = 0 \\ [-\frac{L}{2\pi i n} T\exp(-\frac{2\pi i n l}{L})]_0^L = 0 & n \neq 0 \end{cases}$$

Computer Vision Laboratory

# FD Invariances

- Invariance to translation, scale, rotation and index-shift can be obtained easily.

Translation: $c(l) + T$

Translation affects only the dc-component.
Set dc = 0 to achieve translation invariance

$$\frac{1}{N}\int_{l=0}^{L} T\exp(-\frac{2\pi inl}{L})dl = \begin{cases} \frac{1}{N}\int_{l=0}^{L} T dl \neq 0 & n = 0 \\ [-\frac{L}{2\pi in} T\exp(-\frac{2\pi inl}{L})]_0^L = 0 & n \neq 0 \end{cases}$$

LiU

Computer Vision Laboratory

# FD Invariances

- Invariance to translation, scale, rotation and index-shift can be obtained easily.

$$Scaling : Ac(l)$$

$$f_n^A = \frac{A}{N}\int_{l=0}^{L} c(l)\exp(-\frac{2\pi inl}{L})dl = Af_n$$

LiU

Computer Vision Laboratory

# FD Invariances

- Invariance to translation, scale, rotation and index-shift can be obtained easily.

$Scaling : Ac(l)$

Scaling affects the magnitude of each FD-coefficient.
Normalize the signal energy to be invariant to scale.

LiU

Computer Vision Laboratory

# FD Invariances

- Invariance to translation, scale, rotation and index-shift can be obtained easily.

$$Rotation : exp(i\phi)c(l)$$

$$f_n^{\phi} = \frac{\exp(i\phi)}{N}\int_{l=0}^{L} c(l)\exp(-\frac{2\pi inl}{L})dl = \exp(i\phi)f_n$$

LiU

# FD Invariances

- Invariance to translation, scale, rotation and index-shift can be obtained easily.

Rotation affects the phase of each FD-coefficient.
Only look at magnitudes to be invariant to rotation.

Computer Vision Laboratory

# FD Invariances

- Invariance to translation, scale, rotation and index-shift can be obtained easily.

$$Index-shift : c(l+\Delta l)$$

$$\begin{aligned} f_n^{\Delta l} &= \frac{1}{N}\int_{l=0}^{L} c(l+\Delta l)\exp(-\frac{2\pi inl}{L})dl \\ &= \exp(\frac{2\pi in\Delta l}{L}) f_n \end{aligned}$$

Computer Vision Laboratory

# FD Invariances

- Invariance to translation, scale, rotation and index-shift can be obtained easily.

Index-shift affects the phase of each FD-coefficient.
Only look at magnitudes to be invariant to index-shift.

Computer Vision Laboratory

# FD Invariances

- Invariance to translation, scale, rotation and index-shift can be obtained easily.

- *Translation affects the dc-component only* ➡ *Remove the dc-component*
- *Scaling affects the magnitude of each coefficient* ➡ *Normalize with the (remaining) signal energy*
- *Rotation and index-shift affects the phase of each coefficients* ➡ *Use only the magnitude of each FD-coefficient*

# The phase is important



# Matching with Phase

- Different approaches to incorporate phase information have been used
  - Try to find the transformation T that minimize the error norm.

    Persoon, E. and Fu, K.S. : Shape discrimination using Fourier descriptors. *IEEE Transaction on Systems, Man and Cybernetics*, 7(3):170-179, 1977

  - Normalizing the descriptors.

    Arbter,K., Snyder, W.E and Burkhardt,H. Application of affine-invariant Fourier descriptors to recognition of 3D-objects. *PAMI*, 12(7):640-647, 1990



# Correlation Based Matching

Remove dc-component and normalize with respect to (remaining) energy. Keep the phase. Then

$$\min_{\mathcal{T}} \|c_1 - \mathcal{T} c_2\|^2 \approx 2 - 2 \max_{l} |r_{12}(l)|$$

when the transformation T is restricted to rotation and index shift.



# Correlation Based Matching

Assume that $c_2$ is a shifted and rotated version of $c_1$. Then the cross correlation $r_{12}$ is given as:

$$\begin{aligned} r_{12} &= \mathcal{F}^{-1}\{\bar{C}_1 \cdot C_2\} \\ &= \mathcal{F}^{-1}\{\bar{C}_1(n) \exp i\phi\, C_1(n) \exp(-\frac{i2\pi n \Delta l}{L})\} \\ &= \exp i\phi\, \mathcal{F}^{-1}\{|C_1(n)|^2 \exp(-\frac{i2\pi n \Delta l}{L})\} \\ &= \exp i\phi\, r_{11}(l - \Delta l) \end{aligned}$$

# Correlation Based Matching

$$\|c_1 - \mathcal{T}c_2\|^2 = \|c_1\|^2 + \|\mathcal{T}c_2\|^2 - 2(c_1 \star \mathcal{T}c_2)(0)$$
$$= 2 - 2\exp(-i\phi)\,\boxed{(c_1 \star c_2)(\Delta l)}$$
$$= 2 - 2exp(-i\phi)\boxed{exp(i\phi)r11(\Delta l - \Delta l)}$$
$$= 2 - 2r11(0)$$
$$= 2 - 2\max_l |\boxed{r12(l)}|$$
$$\exp i\phi\, r_{11}(l - \Delta l)$$





# Correlation Based Matching

Remove dc-component and normalize with respect to (remaining) energy. Keep the phase. Then

$$\min_{\mathcal{T}} \|c_1 - \mathcal{T}c_2\|^2 \approx 2 - 2\max_l |r_{12}(l)|$$

when the transformation T is restricted to rotation and index shift.

Rotation dependent matching is achived by using the maximum real value instead

$$\min_{\mathcal{T}} \|c_1 - \mathcal{T}c_2\|^2 \approx 2 - 2\max_l \text{real}(r_{12}(l))$$





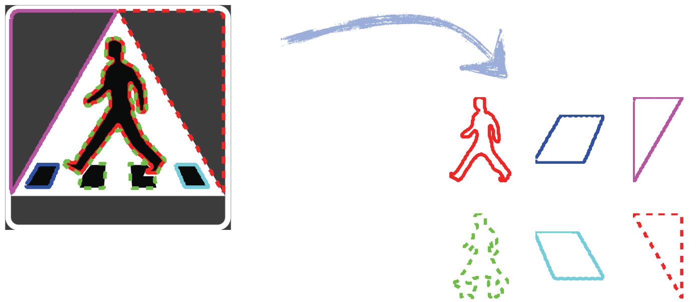
# Summary FD

- Apply the Fourier transform to a periodic 1D parameterization of the contour.
- Normalize with respect to translation
- Normalize with respect to scale
- KEEP THE PHASE!
- Apply the correlation based matching cost



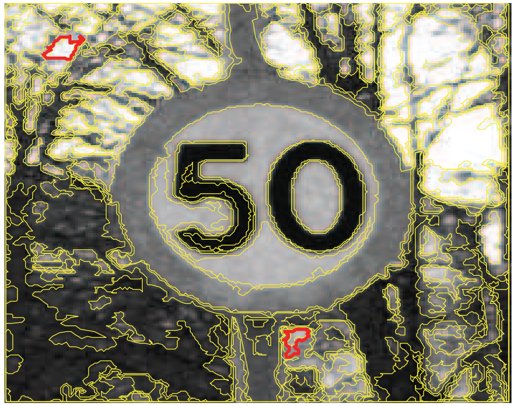


# Traffic Sign Recognition



The prototype for each traffic sign consists of a number of contours. Each described by the FD

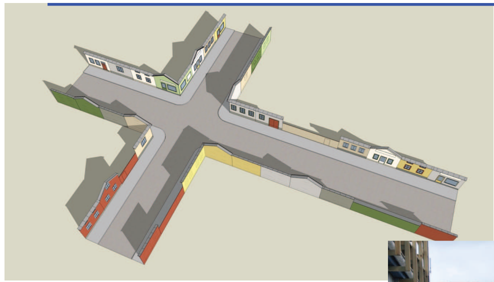
# Traffic Sign Recognition



- Extract contours using MSER
- Describe contours using FDs
- Match contours against all prototypes
- Report found signs if any


Computer Vision Laboratory

# Video



Computer Vision Laboratory

# Video


Computer Vision Laboratory

End of Part I