UNIT V  IMAGE REPRESENTATION AND RECOGNITION


PART A

1. What is skeletonizing? (Nov./Dec.2016)

In digital image processing, morphological skeleton is a skeleton (or medial axis) representation of a shape or binary image, computed by means of morphological operators.

Examples of skeleton extraction of figures in the binary image

Morphological skeletons are of two kinds:

Those defined by means of morphological openings, from which the original shape can be reconstructed,

Those computed by means of the hit-or-miss transform, which preserve the shape's topology.


An image texture is a set of metrics calculated in image processing designed to quantify the perceived texture of an image. Image texture gives us information about the spatial arrangement of color or intensities in an image or selected region of an image.

3. Does the use of chain code compress the description information of an object contour?

Chain codes are the most size-efficient representations of rasterised binary shapes and contours. A chain code is a lossless compression algorithm for monochrome images. The basic principle of chain codes is to separately encode each connected component, or "blob", in the image.

For each such region, a point on the boundary is selected and its coordinates are transmitted. The encoder then moves along the boundary of the region and, at each step, transmits a symbol representing the direction of this movement.

This continues until the encoder returns to the starting position, at which point the blob has been completely described, and encoding continues with the next blob in the image.

This encoding method is particularly effective for images consisting of a reasonably small number of large connected components.
4. What is meant by pattern classes?
A pattern classes are family of patterns that share some common properties. Pattern classes are denoted by \( w_1, w_2, \ldots, w_N \) where \( N \) is the number of classes.

PART B

1. Explain in detail any two boundary representation schemes. (Nov./Dec. 2016)

After image segmentation the resulting collection of regions is usually represented and described in a form suitable for higher level processing. Most important representations are based on shape or texture. Desirable property: descriptors should be insensitive to changes in size, translation or rotation. Actual measurement of features in digital images makes use of many of the techniques discussed earlier, such as linear or morphological image operators.

Put boundary points of a binary region \( R \) (or its boundary) in a clockwise-sorted order. Needed for computation of boundary descriptors.

Chain Codes

Freeman chain codes: strings of integers representing a boundary by a connected sequence of straight-line segments of specified length and direction. The direction of each line segment is coded using a numbering scheme adapted to the connectivity. The accuracy of the straight-line representation depends on the spacing of the sampling grid.

Chain codes
Chain codes: first difference

First difference of a chain code: count number of direction changes (e.g., counterclockwise) that separate two adjacent code elements. chain code: 10103322 (start lower left) first difference: 33133030 (circular sequence)

Polygon representation

A digital boundary can also be approximated by a polygon, possibly with minimum length (MPP: minimum-perimeter polygon). Such rubber band approximations directly within the grey value image are known as active contour models, or snakes.
Minimum-perimeter polygon (MPP)

(a) Light grey: region containing the boundary curve. (b) Convex (white) and concave (black) corner vertices. (c) Concave vertices moved to diagonal mirror locations. The MPP is indicated.

Signatures

1-D representation of a boundary

Example: distance \( r(\theta) \) of centroid to boundary as a function of angle \( \theta \). Often the boundary is first smoothed.


Template Matching is a high-level machine vision method that distinguishes the parts on an image that match a predefined layout.

The algorithm is:

An image having texts (might be in degraded form) or objects is taken as input and converted into gray scale image.

It is passed through Gaussian filter in order to smoothen the broken edges and noise.
It is passed through other pre-processing filters like dilation, noise pixel removal step, thresholding, etc. (texts or objects in white).

All the separate white regions are marked as different objects and counted, cropped to its minimum size. A bounding box is created around each object.

After that, the object region is resized to the size of templates and then each object is compared to all the templates pre-saved in a matrix. corr2 (template {1, n}, char) is a function that calculates the correlation in the template image and the test object image. The template image which has the highest correlation coefficient is marked as identified object or text.

Similarly all the texts or objects are compared and the results are stored in a text file which is displayed at the end of the program. Template Matching methods are relied upon to address the prerequisite of distinguishing all input picture areas at which the template image article is available. Contingent upon the particular issue close by, the client may (or may not) have any desire to recognize the pivoted or scaled events

Let the template T(xt, yt), where (xt, yt) represent the coordinates of each pixel in the template.

Then simply move the center (or the origin) of the template T(xt, yt) over each (x, y) point in the search image and calculate the sum of products between the coefficients in S(x, y) and T(xt, yt) over the whole area spanned by the template. As all possible positions of the template with respect to the search image are considered, the position with the highest score is the best position. This method is sometimes referred to as ‘Linear Spatial Filtering’ and the template is called a filter mask. Above Figure shows an example with a real optical image. The first image shows a cup in a simple background. The second image shows the estimated position of the template with proper application of geometric parameters. It is seen here that the algorithm settled into this estimate after 42 iterations and that it is an accurate estimate. Improvements can be made to the matching method by using more than one template, these other templates can have different scales and rotations. Template matching has various different applications and is used in such fields as face recognition and medical image processing. Systems have been developed and used in the past to count the number of faces that walk across part of a bridge within a certain amount of time. Other systems include automated calcified nodule detection within digital chest X-rays. Template matching is a Hierarchical process, the required number of comparisons can be significantly reduced by clustering.
3. Explain the various boundary descriptors in detail with a neat diagram. (Apr./May.2017)

There are many features that depend on boundary descriptors of objects such as bending energy, curvature etc. For an irregularly shaped object, the boundary direction is a better representation although it is not directly used for shape descriptors like centroid, orientation, area. Consecutive points on the boundary of a shape give relative position or direction. A 4- or 8-connected chain code is used to represent the boundary of an object by a connected sequence of straight line segments. 8 connected number schemes are used to represent the direction in this case. It starts with a beginning location and a list of numbers representing directions such as ddd N ,, 21 ..... . Each direction provides a compact representation of all the information in a boundary. The directions also represent the slope of the boundary. In Fig. below an 8 connectivity chain code is displayed where the boundary description for the boxes with red arrows will be 2-1-0-7-7-0-1-1.

![Chain Code Diagram](image)

Curvature:

The rate of change of a slope is called the curvature. As the digital boundary is generally jagged, getting a true measure of curvature is difficult. The curvature at a single point in the boundary can be defined by its adjacent line segments. The difference between slopes of two adjacent (straight) line segments is a good measure of the curvature at that point of intersection. The curvature of the boundary at can be estimated from the change in the slope is given by,

$$\kappa = \tan^{-1}\left(\frac{y_{i+k} - y_i}{x_{i+k} - x_i}\right) - \tan^{-1}\left(\frac{y_{i} - y_{i-k}}{x_{i} - x_{i-k}}\right) \mod 2\pi$$

Curvature (κ) is a local attribute of a shape. The object boundary is traversed clockwise for finding the curvature. A vertex point is in a convex segment when the change of slope at that point is positive;
Bending Energy

The descriptor called bending energy is obtained by integrating the squared curvature \( \kappa(\rho) \) through the boundary length \( L \). It's a robust shape descriptor and can be used for matching shapes.

\[
E_c = \frac{1}{L} \sum_{p=1}^{L} \kappa(\rho)^2 \quad \frac{2\pi}{R} \leq E_c \leq \infty
\]

The value \( 2\pi / R \) will be obtained as its minimum for a perfect circle with radius \( R \) and the value will be higher for an irregular object.

Total Absolute Curvature

Total absolute curvature is the curvatures added along the boundary points and divided by the boundary length.

\[
\kappa_{\text{total}} = \frac{1}{L} \sum_{p=1}^{L} |\kappa(\rho)| \quad 2\pi \leq \kappa_{\text{total}} \leq \infty
\]

As the convex object will have the minimum value, a rough object will have a higher value.