Computer Vision & Digital Image Processing

Image Segmentation

Image segmentation

- Segmentation divides an image into its constituent parts or objects
- Level of subdivision depends on the problem being solved
- Segmentation stops when objects of interest in an application have been isolated
- Example:
  - For an air-to-ground target acquisition system interest may lie in identifying vehicles on a road
    - Segment the road from the image
    - Segment contents of the road down to objects of a range of sizes that correspond to potential vehicles
    - No need to go below this level, or segment outside the road boundary

Image segmentation (continued)

- Autonomous segmentation is one of the most difficult tasks in image processing - largely determines the eventual failure or success of the process
- Segmentation algorithms for monochrome images are based on one of two basic properties of gray-level values
  - Discontinuity
  - Similarity
- For discontinuity, the approach is to partition an image based on abrupt changes in gray level
- The principal areas of interest are:
  - detection of isolated points
  - detection of lines and edges in an image

Image segmentation (continued)

- For similarity, the principal approaches are based on
  - thresholding
  - region growing
  - region splitting
  - merging
- Using discontinuity and similarity of gray-level pixel values is applicable to both static and dynamic (time varying) images
- For dynamic images, the concept of motion can be exploited in the segmentation process

Discontinuity detection

- Detecting discontinuities (points, lines and edges) is generally accomplished by mask processing (much as in the spatial domain filter examples)
- Use the response equation

$$R = w_1 z_1 + w_2 z_2 + \ldots + w_9 z_9 = \sum_{i=1}^{9} w_i z_i$$

- A mask used for detecting isolated points (different from a constant background would be

-1	-1	-1
-1	8	-1
-1	-1	-1

Isolated point detection

- Detection of isolated points is accomplished by using the previous mask
- An isolated point is detected if the response of the mask is greater that a predetermined threshold

$$|R| > T$$

- This measures the weighted difference between a center point and its neighbors
- The mask is the same as the high frequency filtering mask
- The emphasis here is on the detection of points
  - Only differences that are large enough to be considered isolated points in an image are of interest

Line detection

- Line detection would involve the application of several masks
- In the creation of masks, the intent is to form a mask (or set of masks) that will respond to a 1-pixel thick line in a given orientation
  - Horizontal, Vertical, +45º, -45º

-1	-1	-1
2	2	2
-1	-1	-1

-1	2	-1
-1	2	-1
-1	2	-1

-1	-1	2
-1	2	-1
2	-1	-1

2	-1	-1
-1	2	-1
-1	-1	2

Line detection (continued)

- With a constant background, the maximum response occurs when the line is "lined up" with the center of the mask
- Note that the preferred direction of each mask is weighted with a larger coefficient than other possible directions
- Let R_1, R_2, R_3 and R_4 denote the responses of the masks
- If, at a certain point in the image,

$$|R_i| > |R_j| \text{ for all } j \neq i$$

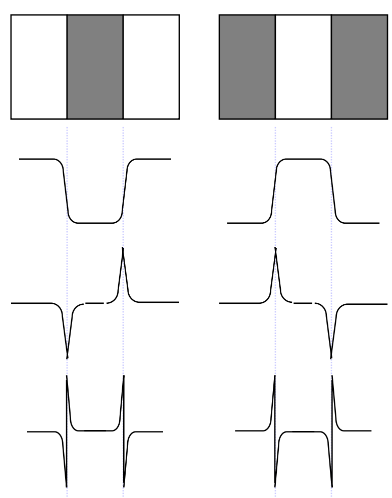
- that point is said to be more likely associated with a line in the direction of mask *i*

Edge detection

- Edge detection is by far the most common approach for detecting discontinuities in gray levels
  - Isolated points and 1-pixel thin lines are not common in most practical applications
- Basic formulation and initial assumptions
  - An edge is a boundary between two regions with relatively distinct gray-level properties
  - Regions are sufficiently homogeneous so that the transition between the regions can be determined on the basis of gray-level discontinuities alone
  - If this is not valid, some other techniques will be used
- The basic idea behind most edge detection techniques is the computation of a local derivative operator

Derivative operators

- An image of a dark stripe on a light background (and visa versa)
- A profile of the lines in the image (modeled as a gradual rather than sharp transition)
  - Edges in images tend to be slightly blurred as a result of sampling
- The first derivative: the magnitude detects the presence of an edge
- The second derivative: the sign tells the type of transition (light-to-dark or dark-to-light) Note also the presence of a zero-crossing at each edge



Gradient operators

- The first derivative at any point in an image is computed using the magnitude of the gradient

$$\nabla \mathrm{f} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \partial f / \partial x \\ \partial f / \partial y \end{bmatrix}$$

- Where the *magnitude* is

$$\nabla f = \mathrm{mag}(\nabla \mathrm{f}) = \left[G_x^2 + G_y^2\right]^{1/2}$$
$$\approx |G_x| + |G_y|$$

- The *direction* of the gradient vector is the angle α(x,y) given by

$$\alpha(x, y) = \tan^{-1}\left(\frac{G_y}{G_x}\right)$$

Gradient operators (continued)

- The derivatives may be digitally implemented in several ways, but the Sobel operators are commonly chosen as they provide both a differencing and a smoothing
  - The smoothing is advantageous as derivative operators enhance noise
- The gradient computation using Sobel operators is given as

$$G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)$$
$$G_y = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)$$

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

The Laplacian

- The Laplacian is a second order derivative operator given by

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

- As with the gradient, this may be implemented digitally
- With a 3x3 mask, the most common form is

$$\nabla^2 f = 4z_5 - (z_2 + z_4 + z_6 + z_8)$$

- The basic requirement for the digital Laplacian is that the center coefficient be positive, the other coefficients be negative (or zero), and that the sum of the coefficients be zero (indicating a zero response over a constant area)

The Laplacian (continued)

- Although the Laplacian responds to changes in intensity, it is seldom used in edge detection for several reasons
  - As a second derivative operator it is typically unacceptably sensitive to noise
  - The Laplacian produces double edges
  - Unable to detect direction
- As such, the Laplacian is used in the secondary role of detector for establishing whether a pixel is on the light or dark side of an edge

The Laplacian (continued)

- A more general use of the Laplacian is to find the location of edges using the zero-crossings property
- Basic idea is to convolve an image with the Laplacian of a 2-D Gaussian function of the form

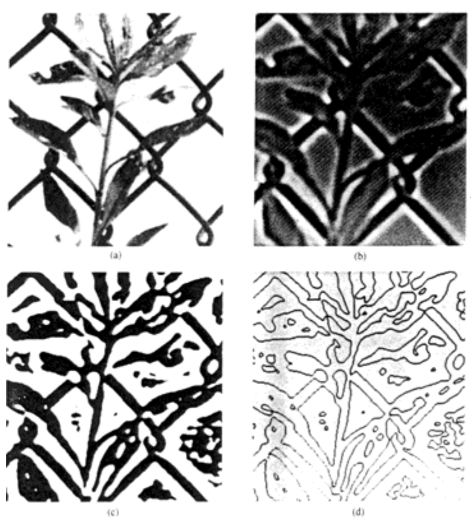
$$h(x,y) = \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$$

- σ = standard deviation.
- If $r^2=x^2+y^2$, then the Laplacian is then

$$\nabla^2 h = \left(\frac{r^2-\sigma^2}{\sigma^4}\right)\exp\left(-\frac{r^2}{2\sigma^2}\right)$$

Example using the Laplacian

Original Image



Original Image convolved with the Laplacian

Thresholding the convolved image to yield a binary image

Zero crossings from the binary image