

5. Pattern recognition: Segmentation and texture analysis

5.1. Image segmentation

(Reading: Castleman, 1996, pp. 447-469, 499-501)

- segmentation represents a critical step in the analysis of an image (see Figure 1.1) and is the prerequisite to object analysis or more advanced operations; essentially, segmentation comprises the partitioning of an image into non-overlapping features or regions, or the delineation of the boundaries separating these regions

5.1.1. Thresholding and watershed transforms

- the simplest approach to segmentation is through global thresholding, i.e. by specifying a grey value G_{thr} and then binarizing the image about this value, such that $G'(x,y) = 255 \Leftrightarrow G(x,y) > G_{thr}$, $G'(x,y) = 0 \Leftrightarrow G(x,y) \leq G_{thr}$; G_{thr} can be derived from the image histogram in a variety of ways, Fig. 5.1 shows an example for which the grey values associated with the two feature classes (ice, pores) are well and narrowly defined and can be derived from the image histogram; here, an areal-fraction criterion that defines a threshold based on the areal fraction of the two feature classes present can be derived

- for images that exhibit a bias (uneven lighting etc.) across the image, local (adaptive) thresholds can be derived based on the histogram obtained for subsets of the overall scene

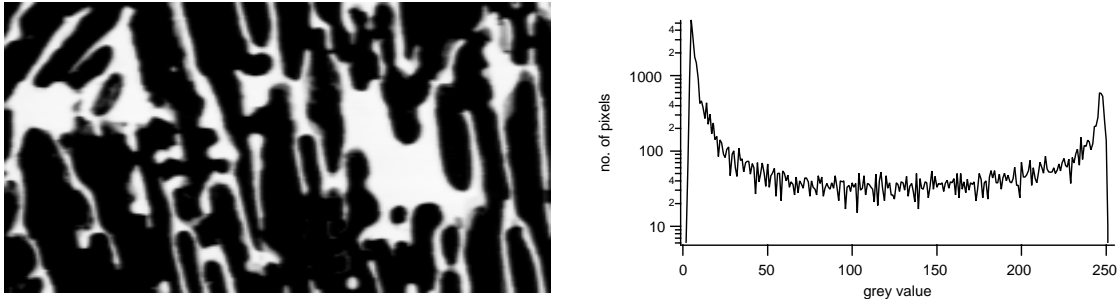


Fig.5.1: Image of sea-ice thin section (approximately 8 mm across) recorded in incident light, with pores shown as white features and ice shown black (left). At right, the image histogram is shown (note the logarithmic scale on the vertical axis).

- in many cases, it is not only illumination gradients across the image but variability in the magnitude of the threshold for different boundaries that causes problems in a particular scene; this is addressed by implementing a watershed segmentation algorithm based on a Euclidean distance map of a preprocessed image (see Section 3.3, Fig. 3.14); essentially, the boundaries between neighbouring objects are derived from the local ridges in the Euclidean distance map of a scene (Fig. 5.2)

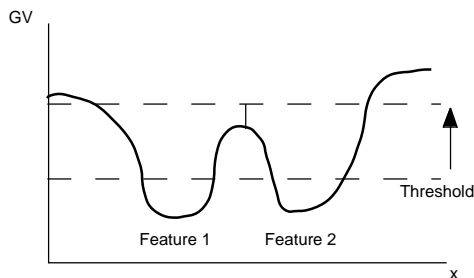


Fig. 5.2: Illustration of the watershed segmentation algorithm (after Castleman, 1996): the threshold is sequentially increased across the entire grey value depth range of two features identified as separate entities at a low threshold setting, with the crest between the features marking the boundary.

5.1.2. Boundary delineation through gradient filtering

- rather than identifying the individual features or regions present within an image, segmentation can also be accomplished by delineating the boundaries between different features through gradient filtering; the principles of gradient filters have been discussed in Section 3.2 (Figs. 3.10 and 3.11) for the spatial domain and in Section 4 for the frequency domain

- in determining the location of such feature boundaries, the computation of the Laplacian transform proves useful in identifying the exact boundary location through the zero crossings of the 2nd derivative (see also Fig. 3.10); however, the sensitivity to noise (such as associated with scan lines in a video image in Fig. 5.3) requires some image enhancement through low-pass or median filtering prior to computing the Laplacian transform, with either one of the following convolution kernels:

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad (5.1)$$

- the Sobel filter, on the other hand, is more robust as the central line/column in the convolution kernel is empty; the resulting image is derived by recombination of two composite images convolved with the following kernels:

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (5.2)$$

- Fig. 5.3 shows examples of these two techniques that demonstrate the sensitivity of the Laplacian transform to high-frequency noise; note also that due to the empty row/column, the Sobel transform generates a broader boundary that may require subsequent thinning/skeletonizing

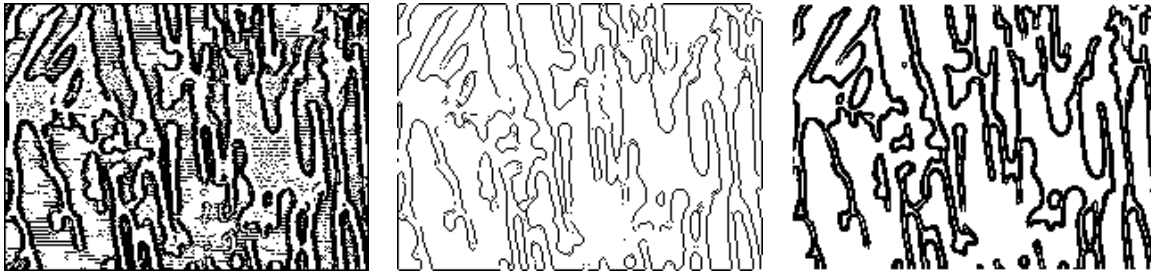


Fig. 5.3: Edge detection of pore image shown in Fig. 5.1 with Laplacian transform (left, note horizontal striping parallel to video scan lines), Laplacian transform after median filtering (center) and Sobel transform (right).

5.2. Analysis of image texture

- the texture of an image, i.e. the grey-value variability at different spatial scales and orientations, is of importance both in the process of segmentation, where textural differences between different features can form the basis for the segmentation process, as well as in image analysis where textural measures convey important information about a particular scene; here, two different types of textural measures will be discussed briefly, those that are invariant and those that are variant with rotation and scale

5.2.1. Rotation and scale invariant measures of image texture

- the simplest measure of texture are the moments of the grey value distribution of a given sub-sample in an image; in particular the variance (see also Section 2.1) provides a measure of the degree of grey-value variability

- rather than preselecting particular regions in an image, a variance operator can be implemented such that a window W is translated across an image for which the variance is computed according to

$$\sigma_w^2 = \frac{1}{N-1} \sum_{m,n \in W} (G_{m,n} - \langle G \rangle_w)^2 \quad (5.3)$$

- based on the magnitude of the variance (or other, higher moments of the grey value distribution) boundaries can be identified between different regions in an image

5.2.2. Rotation and scale variant measures of image texture

- textural measures can be derived from the Fourier spectrum of an image, which contains information both about the directionality as well as periodicity of the image (for a detailed description of such an approach see Jähne, 1995)

- a set of measures commonly employed (particularly in remote-sensing applications) are based on the co-occurrence matrix (COM) C_{ij} ; the COM is defined for a given separation of l pixels in a given direction d (commonly N, NE, E, SE etc. on a rectangular grid, see Fig. 5.4), for a particular value of l and d , the elements i,j of the COM are given by the number of co-occurrences of pixels of grey value g_i and g_j divided by the number M of pixel pairs covered by the matrix for two pixels separated by l in direction d ; thus, the COM has the size $N \times N$ where N is the number of grey levels resolved in the image (to reduce the size and sparsity of the matrix, an image is requantized from, e.g., 8 bits to 4 bits or less)

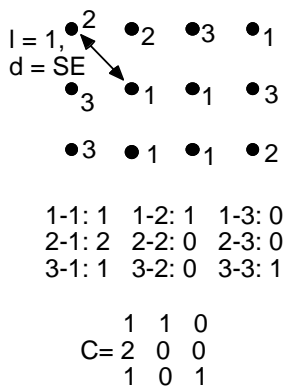


Fig. 5.4: Computation of the co-occurrence matrix C.

- from the COM, texture measures such as the entropy H , inertia I and energy E can be derived (among a larger number of others, see, e.g., Shokr M. E., Evaluation of second-order texture parameters for sea ice classification from radar images, *J. Geophys. Res.*, 96, 10625-10640, 1991, for details)

$$H = \sum_{i=1}^N \sum_{j=1}^N C_{ij} \log C_{ij}$$

$$I = \sum_{i=1}^N \sum_{j=1}^N (i - j)^2 C_{ij} \tag{5.4}$$

$$E = \sum_{i=1}^N \sum_{j=1}^N C_{ij}^2$$

5.3. Model-based extraction of features: The Hough transform

- a common problem that can be considered part of the segmentation of an image consists of the extraction of size parameters describing a set of geometric objects (such as linear or circular features) present in an image but not particularly well defined (e.g., a set of points falling on a line but separated by wider gaps inbetween); in such a case the computational effort required to translate, e.g., all sets of lines that could run through pairs of points in an image is prohibitive

- the Hough transform presents us with an elegant alternative solution; in the case of points arranged along a line, it is based on the fact that any straight line $y = a x + b$ can be represented in polar coordinates by a vector of length ρ enclosing an angle θ with the x axis (Fig. 5.5) such that

$$\rho = x \cos \theta + y \sin \theta \tag{5.5}$$

- conversely a single point in the x,y plane is represented by a harmonic function (eq. 5.5) in the Hough plane; for points falling on a straight line, these harmonic functions all intersect in a single point that defines the normal to that particular line in polar coordinates (Fig. 5.4); by computing the Hough transform for an image, lines can be identified based on the number of sinusoids intersecting in different points in the image

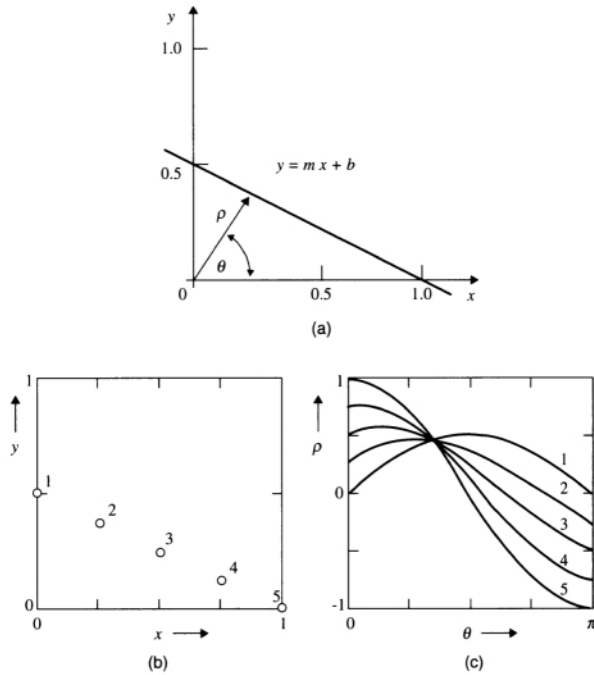


Fig. 5.5: Representation of a line in polar coordinates (a) and representation of points on a line (b) in Hough space (c), where the intersection corresponds to the polar coordinates of the normal on the line joining the points (Castleman, 1996).

- this approach is applicable to other functions as well; thus, for a circle given by

$$r_0^2 = (x - a)^2 + (y - b)^2 \tag{5.6}$$

an entire set of Hough transforms exists in a 3-parameter space; around every point located on the circumference of a circle in the x,y plane, a set of circles with increasing radius r_i can be constructed that intersect in the center of the circle for $r_i = r_0$ (Fig. 5.6)

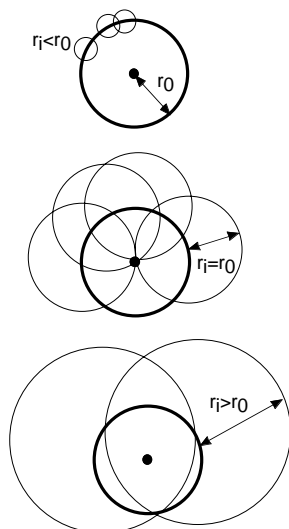


Fig. 5.6: Hough transformation for circle with radius r_0 into Hough space; the radius and center location of the circle are derived from sets of circles with variable radius r_i constructed on the circumference of the circle which intersect in its center for $r_i = r_0$.