COMP 448/548: Medical Image Analysis

Feature extraction

Çiğdem Gündüz Demir
cgunduz@ku.edu.tr

1

Handcrafted feature extraction for medical images

- Intensity features
  - Quantify color/grayscale distribution of pixels
  - First-order statistical texture
- Texture features
  - Similar structures repeated over and over again (repeated patterns)
  - Statistical approach is to quantify spatial arrangement of pixel intensities
  - Structural approach is to define texture on objects/primitives

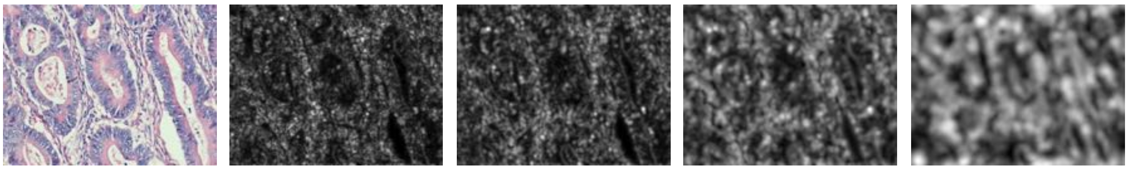
LAST LECTURE

- Structural features to quantify spatial distribution of objects/primitives
- Morphological features to quantify the shape/size of a segmented object

2

Intensity/texture features to represent an entire image

1. Calculate them on the entire image
   - e.g., Haralick features of the cooccurrence matrix calculated on all image pixels
   - e.g., first order statistical features (mean, standard deviation, entropy, etc.) of filter responses calculated on all image pixels

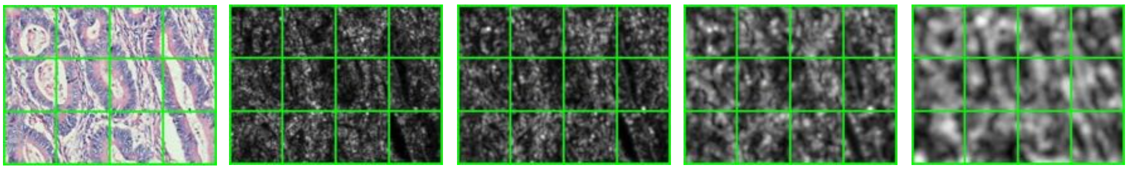


Example: The grayscale image is convolved with Gabor filters of six orientations and four scales. For each scale, the filter responses of different orientations are averaged to obtain rotation-invariant features. Statistical features will be separately calculated on each of these four average response maps.

3

Intensity/texture features to represent an entire image

2. Grid-based approach
   - Divide an image into grids
   - Calculate features on each grid
   - Aggregate the feature values

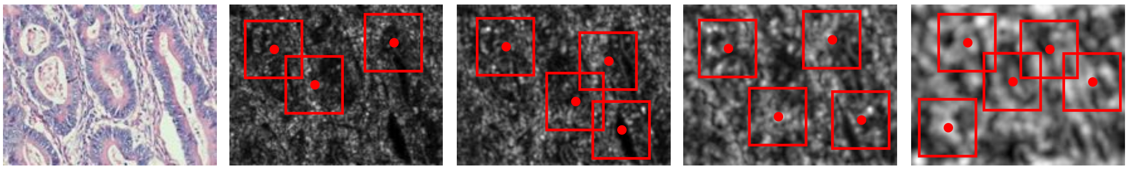


Example: The grayscale image is convolved with Gabor filters of six orientations and four scales. For each scale, the filter responses of different orientations are averaged. Then, for each grid entry of each average map, statistical features will be separately calculated. Afterwards, the statistical features will be aggregated for the same map.

4

Intensity/texture features to represent an entire image

3. Calculate them on the keypoints
   - Find the key points
   - Locate a window on each of these key points
   - Calculate features on each window
   - Aggregate the feature values



Example: The grayscale image is convolved with Gabor filters of six orientations and four scales. For each scale, the filter responses of different orientations are averaged. Then, for each window located on a keypoint, statistical features will be separately calculated. Afterwards, the statistical features will be aggregated for the same map.

5

Keypoint localization by SIFT

- SIFT (scale-invariant feature transform)
  - Convolves an image with Gaussian filters at different scales
  - Identifies keypoints as local maxima of the difference of these Gaussians (DoG)
  - Discards low-contrast keypoints and edge response points along edges

DoG maxima

Discard low-contrast

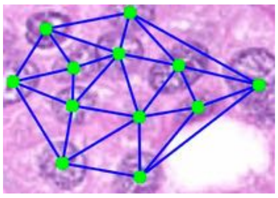
Discard edge responses

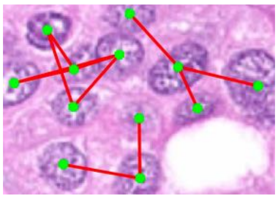
Lowe, "Distinctive image features from scale-invariant keypoints," Int Journal of Computer Vision, 2004. https://www.cs.ubc.ca/~lowe/keypoints/

6

Structural features

- Quantify spatial distribution of objects/primitives
  - By defining texture metrics (see the previous slide set)
  - By making use of graphs
    - ***Graphs can be used to represent other types of relations between the objects/primitives as well (they are not just for spatial relations)***
- A graph G = (V, E) consists of a set of vertices (nodes) V and a set of edges E, each of which is defined between a pair of vertices

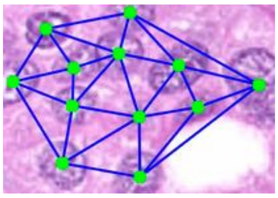


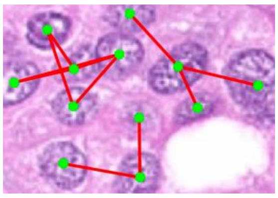


7

Structural features

- To construct a graph on a medical image, you need to
  - First define its vertices
  - Then assign edges in between these vertices
    - Edges may consist of ordered or unordered pairs of vertices (*directed* vs *undirected* graphs)
    - Edges may be associated with unit weights, continuous-valued weights or discrete values (*unweighted* vs *weighted* vs *attributed* graphs)
- Graphs can be used to represent any relations between the vertices
  - e.g., Delaunay triangulation
  - e.g., probabilistic graph where the edges are probabilistically set with respect to, for example, geometric distances or similarities between vertices

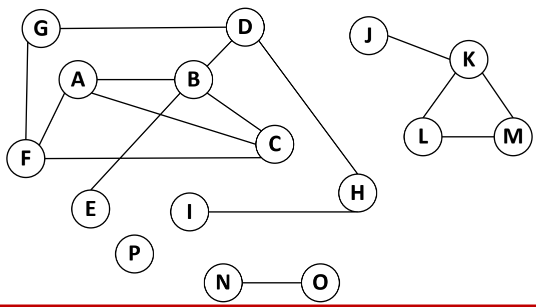




8

Graphs – basic definitions

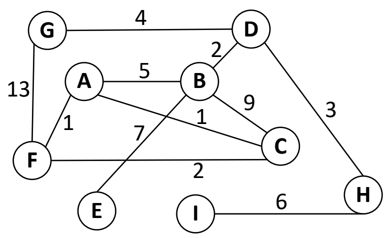
- Two vertices are adjacent if they are the endpoints of the same edge
  - *F and C are adjacent, but A and H are not*
- A vertex is reachable from another vertex if there exists a path between these two vertices
  - *A is reachable from I, but A is not reachable from M*
- An undirected graph is connected if there exists at least one path from every vertex to every other vertex
  - *This graph is unconnected*
  - *It has four connected components*

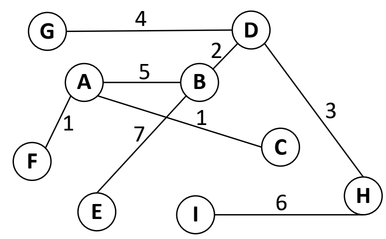


9

Useful graph algorithms for feature extraction

- *Breadth-first traversal (BFT)* to calculate shortest paths in an unweighted graph
  - Shortest path distance from A to G → 2 (AFG when edge weights are ignored)
- *Dijkstra's shortest path algorithm* to calculate shortest paths in a weighted graph (for nonnegative weights)
  - Shortest path distance from A to G → 11 (shortest path consisting of ABDG)
- *Prim's and Kruskal's algorithm* to find a minimum spanning tree in a weighted graph
- Random walks to generate (multiple) sequences from a given graph

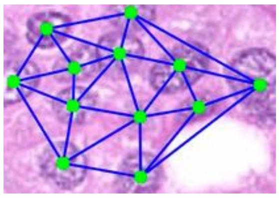




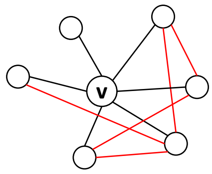
Minimum spanning tree

10

Local graph features



- Defined for each graph vertex v
  - *Degree* d_v is the number of its neighbors (adjacent vertices)
  - *Clustering coefficient* $CC_v = {}^{E_v}/_{(d_v(d_v-1)/2)}$, where E_v is the number of the existing edges between these neighbors
  - *Closeness* is the average length of the shortest paths between vertex v and every other vertex reachable from v
  - *Eccentricity* is the maximum length of these shortest paths



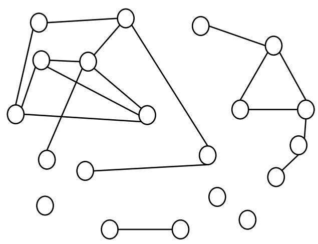
$CC_v = {}^{5}/_{(6\cdot 5/2)} = 1/3$

- All these metrics are to quantify connectivity information
  - Degree and clustering coefficient consider vertices in a close neighborhood
  - Closeness and eccentricity consider more distant vertices

11

Global graph features

- Defined for an entire graph (and thus, for an entire image on which the graph is constructed)
  1. Statistics on the local graph features:
     - Average degree
     - Average clustering coefficient
     - Average closeness
     - Average eccentricity
     - Diameter (maximum eccentricity)
  2. Using connected components
     - Ratio of the size of the giant (largest) connected component
     - Percentage of isolated vertices (whose degree is 0) and end vertices (whose degree is 1)

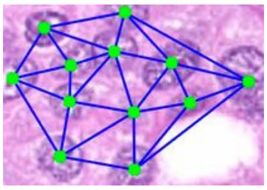


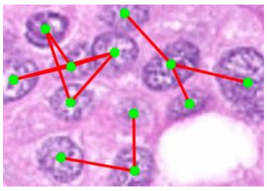
Giant connected comp. ratio = 9/20
Isolated vertex percentage = 3/20
End vertex percentage = 1/20

12

Global graph features

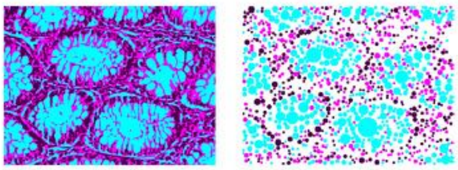
- Defined for an entire graph (and thus, for an entire image on which the graph is constructed)
  3. Using graph spectrum (eigenvalues of the graph's adjacency matrix)
     - Spectral radius: Absolute value of the largest eigenvalue in the spectrum
     - Eigen exponent: Slope of the sorted eigenvalues (sometimes n largest eigenvalues) in the spectrum as a function of their orders in log-log scale
  4. Additional metrics such as
     - Statistics on the graph edge lengths
     - Statistics on the triangle areas, if the graph is constructed using Delaunay triangulation





Example: Graphs for tissue classification

- Represent an image with a graph of cytological tissue components
  - Approximate their locations with circles
  - Construct a Delaunay triangulation on the circles
  - Color each triangle edge according to the circle (component) type of its end-nodes
- Considering edge colors, redefine average degree, average clustering coefficient, and diameter
- Classify the image based on these features using a support vector machine classifier



(a) (b)

Fig. 3. (a) Pixels after color quantization and (b) circular primitives located on these quantized pixels. Here, purple, pink, and cyan correspond to nuclear, stromal, and luminal components, respectively.

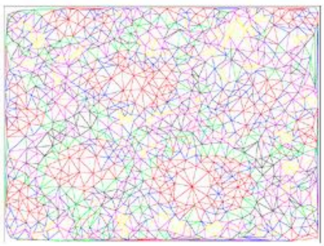


Fig. 4. Edges that are obtained by using Delaunay triangulation and that are colored with one of the six colors depending on the component types of their end nodes.

Altunbay, Cigir, Sokmensuer, and Gunduz-Demir, "Color graphs for automated cancer diagnosis and grading," IEEE T Biomedical Engineering, 2010.

Example: Graphs for tissue classification

- Relies on locating and characterizing biological structures in a tissue
  - Define an attributed graph for an entire image
  - Define a set of query graphs as a reference to the normal biological structure
  - Locate key regions (points) that are most similar to a normal biological structure by searching the query graphs over the entire tissue graph
    - Uses inexact subgraph matching (as exact graph matching is an NP-complete problem)

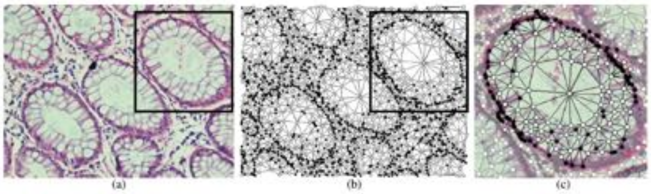


Fig. 2. An illustration of the graph generation step: (a) an example normal tissue image, (b) the tissue graph generated for this image, and (c) a query graph generated to represent a normal gland.

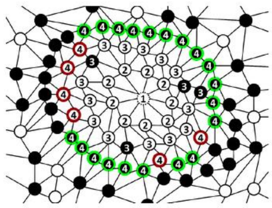


Fig. 3. An illustration of generating a query graph. The node labels are indicated using four different representations and the orders in which the nodes are expanded are given inside their corresponding objects.

Ozdemir and Gunduz-Demir, “A hybrid classification model for digital pathology using structural and statistical pattern recognition,” IEEE T Medical Imaging, 2013.

15

Example: Graphs for tissue classification

- After locating key regions that are most similar to a normal biological structure
  - Extract texture features on these key regions
  - Use graph edit distances between each query graph and the corresponding region as additional features
  - Classify the image based on these features using a support vector machine classifier

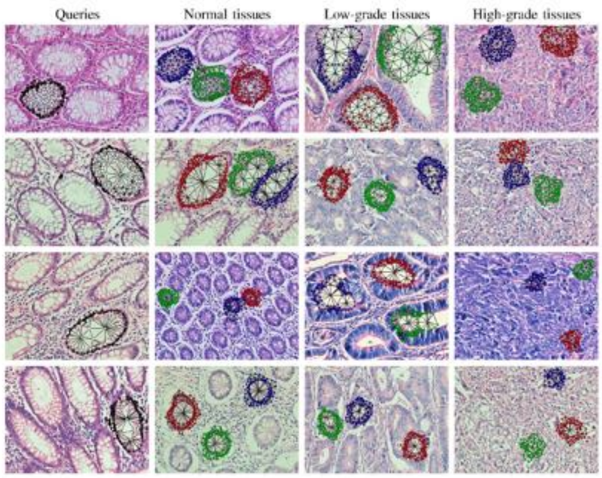


Fig. 4. The query graphs generated as a reference for a normal gland structure and the subgraphs located in example normal, low-grade cancerous, and high-grade cancerous tissue images. The first image of each row shows the query graph on the image from which it is taken whereas the remaining ones show the three-most similar subgraphs to the corresponding query graph. In this figure, the subgraphs of the same image are shown with different colors (red for the most similar subgraph, green for the second-most similar subgraph, and blue for the third-most similar subgraph).

Ozdemir and Gunduz-Demir, “A hybrid classification model for digital pathology using structural and statistical pattern recognition,” IEEE T Medical Imaging, 2013.

16

Example: Graphs for nucleus segmentation

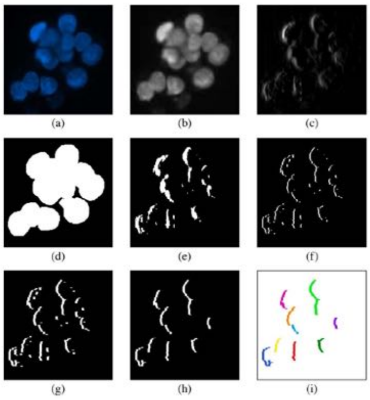


Fig. 2. Defining left boundary primitives: (a) original subimage, (b) blue band I_b of the image, (c) response map $R_{\mathtt{left}}$ obtained by applying the Sobel operator $S_{\mathtt{left}}$, (d) mask used to determine local Sobel threshold levels, (e) binary image $B_{\mathtt{left}}$ after thresholding, (f) boundaries obtained after taking the leftmost pixels [here there are discontinuities between the boundaries because of their one-pixel thickness], (g) boundary map $P_{\mathtt{left}}$ obtained after taking the d-leftmost pixels, (h) $P_{\mathtt{left}}$ after eliminating small connected components, (i) left boundary primitives, each of which is shown in a different color.

- Define four types of primitives to represent nucleus boundaries at different orientations
- Construct an attributed relational graph on the primitives to represent their spatial relations

Arslan, Ersahin, Cetin-Atalay, and Gunduz-Demir, "Attributed relational graphs for cell nucleus segmentation in fluorescence microscopy images," IEEE T Medical Imaging, 2013.

17

Example: Graphs for nucleus segmentation

- Reduce the nucleus identification problem to finding predefined structural patterns in the constructed graph

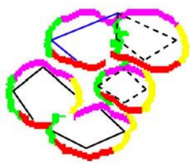


Fig. 5. Two structural patterns used for nucleus localization: 4PRIM and 3PRIM. Corresponding edges are shown in black and blue, respectively. The 4PRIM pattern has two subtypes that correspond to subgraphs, forming a loop (dashed black edges) and a chain (solid black edges).

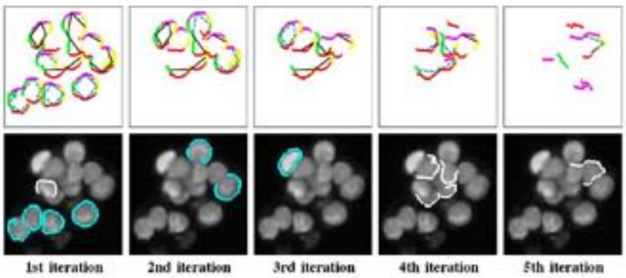


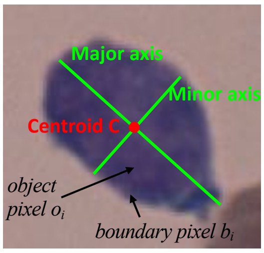
Fig. 8. Illustration of graphs constructed in different iterations and selected primitive segments in these iterations. In the first row, the graph edges of the 4PRIM and 3PRIM patterns that satisfy the standard deviation constraint of the corresponding iteration are indicated as dashed blue lines and the others as black solid lines. In the second row, selected segments of these 4PRIM and 3PRIM patterns are shown in cyan and white, respectively.

Arslan, Ersahin, Cetin-Atalay, and Gunduz-Demir, "Attributed relational graphs for cell nucleus segmentation in fluorescence microscopy images," IEEE T Medical Imaging, 2013.

18

Morphological (geometric) features

- Mostly defined for an individual object (for each segmented region)
- Quantify the shape and size characteristics of this object (region)
- *Size*
  - *Area* is the number of pixels in the object
  - *Perimeter* is the number of pixels on the object's boundaries (based on 4-connectivity or 8-connectivity)
  - *Major axis* is the longest line that goes through the centroid
  - *Minor axis* is the line that is perpendicular to the major axis and goes through the centroid
  - *Average radius* is average length of the radial lines from the centroid to every boundary pixel

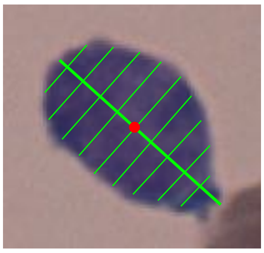


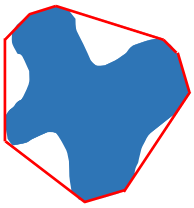
$$C = \frac{1}{N}\sum_{i=1}^{N} coord(o_i)$$

$$Avg\ radius = \frac{1}{N}\sum_{i=1}^{N} \|b_i - C\|$$

Morphological (geometric) features

- *Shape*
  - *Circularity* $= 4\pi \frac{area}{perimeter^2}$ (0 < metric <=1, 1 for circles)
  - *Symmetry* is based on the length difference of two segments on the same line perpendicular to the major axis
  - *Smoothness* is the sum of the smoothness of the boundary pixels
    - For each boundary pixel, it is the difference between its radius (radial line from the centroid to this pixel) and the average of the radii of its surrounding (or close enough) boundary pixels
  - *Eccentricity* is the ratio of the minor to the major axis length
- Can be quantified also using the convex hull of the object



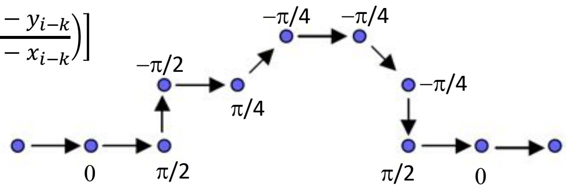


Convexity $= \frac{perimeter\ (object)}{perimeter(convex\ hull)}$ *Solidity* $= \frac{area\ (object)}{area(convex\ hull)}$ (0 < metric <=1)

Morphological (geometric) features

- *Boundary descriptors*
  - *Curvature* is to quantify the amount by which the boundary contour deviates from being a straight line
    - ➢ *K-curvature* of a boundary pixel $b_i = (x_i, y_i)$ is estimated from the change in the k-slope (considering preceding and succeeding boundary points)

$$kSlope(b_i) = \left[\tan^{-1}\left(\frac{y_{i+k} - y_i}{x_{i+k} - x_i}\right) - \tan^{-1}\left(\frac{y_i - y_{i-k}}{x_i - x_{i-k}}\right)\right]$$

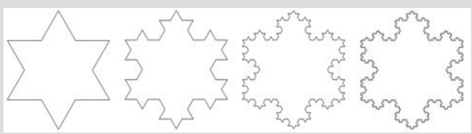


    - ➢ *Bending energy* is the sum-of-squares of the K-curvatures of boundary pixels

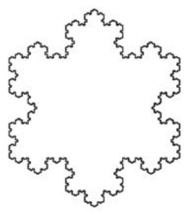
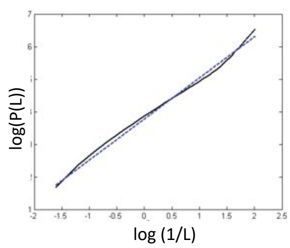
Morphological (geometric) features

- *Boundary descriptors*
  - *Fractal dimension* of the boundary contour quantifies the rate at which its length (perimeter) increases as the measurement scale decreases
    - ➢ Walk along the boundary with a given stride length L and measure the perimeter P(L)
    - ➢ Smaller stride lengths allow visiting paths with more irregularities, which increases P(L)
    - ➢ Plot the measured perimeter as a function of the stride length on log-log axes
    - ➢ Fractal dimension is the slope of the fitted line

Fractals exhibit similar patterns at different scales
This property is known as *self-similarity*




Koch snowflake and cauliflower as examples of fractals

Thank you!

Next time:

Basics of classifiers

23