Texture Analysis and its role in Radiomics

Many thanks to Ulas Bagci (Northwestern University)
for sharing his experience and course material

Χρήσιμο Υλικό

- Από το βιβλίο Gonzalez/Woods κεφ. 11 Εξαγωγή χαρακτηριστικών
- Συγκεκριμένα 11.4.3 Υφή

- Συμπληρωματικά (όχι υποχρεωτικά):

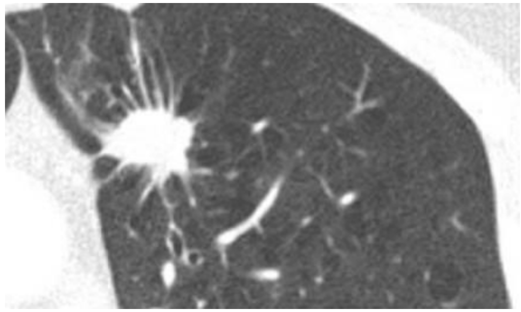
https://prism.ucalgary.ca/bitstream/handle/1880/51900/texture%20tutorial%20v%203_0%20180206.pdf

Texture analysis in practise

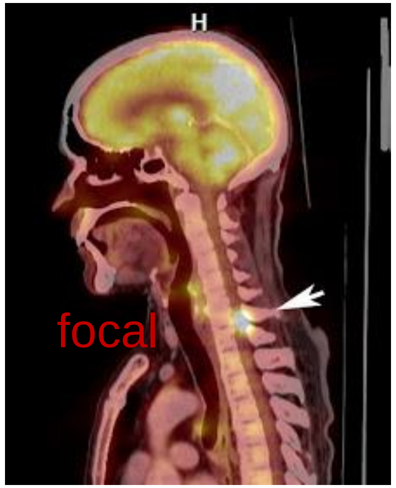
- Texture information of lesion (heterogeneous vs homogeneous)

TEXTURE ANALYSIS?

- Technique used in image processing to identify, characterize, and compare regions with distinct patterns
- Measure and capture local image properties which are not necessarily based on intensity properties



spiculated



focal

What to Measure in PET/SPECT/.. Images?

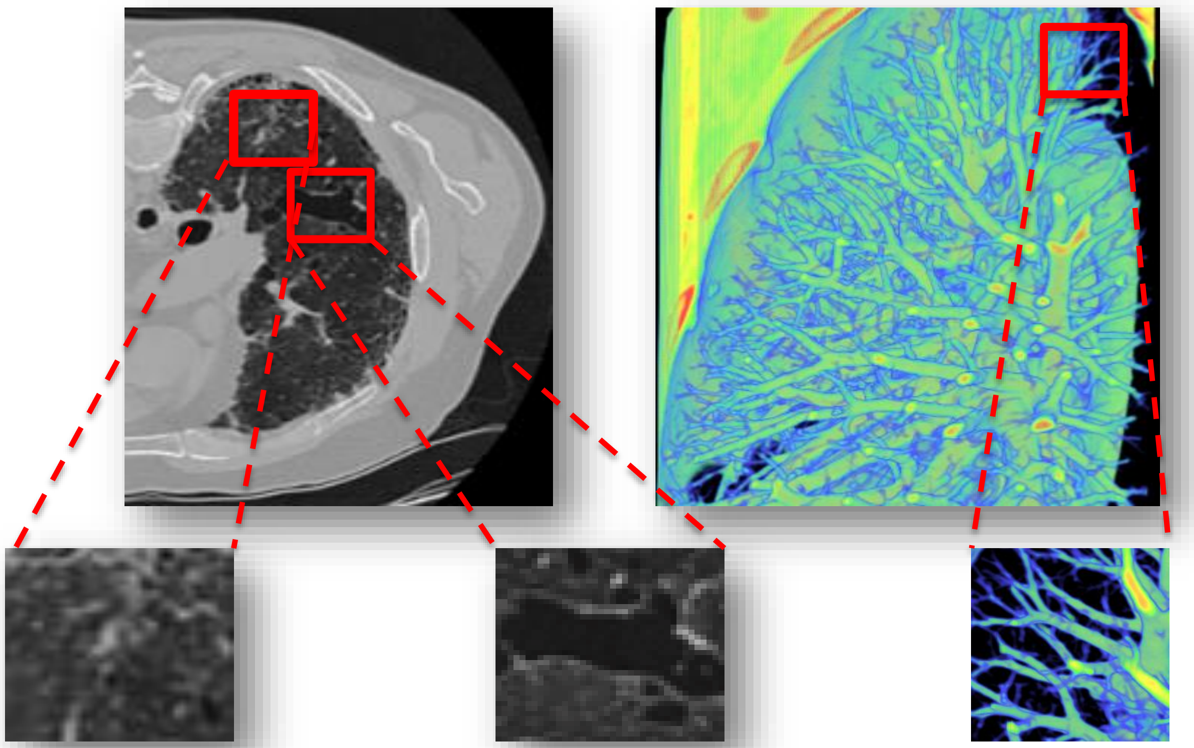
- Texture information of lesion (heterogeneous vs homogeneous)

TEXTURE ANALYSIS?

- Technique used in image processing to identify, characterize, and compare regions with distinct patterns
- Measure and capture local image properties which are not necessarily based on intensity properties

What to Measure in PET/SPECT/.. Images?

- Texture information of lesion (heterogeneous vs homogeneous)



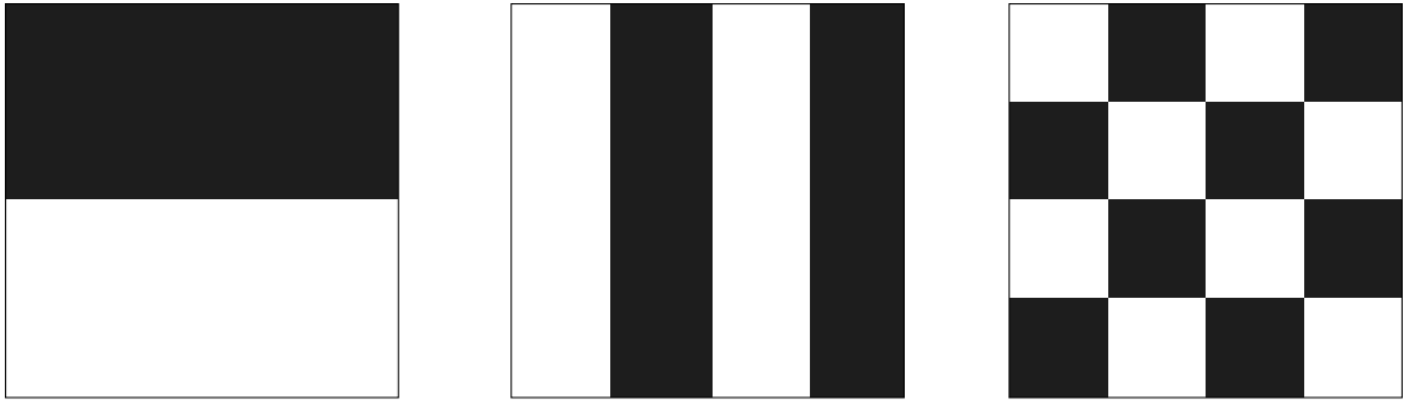
A spatial arrangement of a predefined number of voxels allowing the extraction of complex image properties.

Credit to:

Bagci EMBC 2011, RSNA 2011, 2012, ISBI 2012, PlosOne 2013.

Simple Histogram can’t tell the whole story

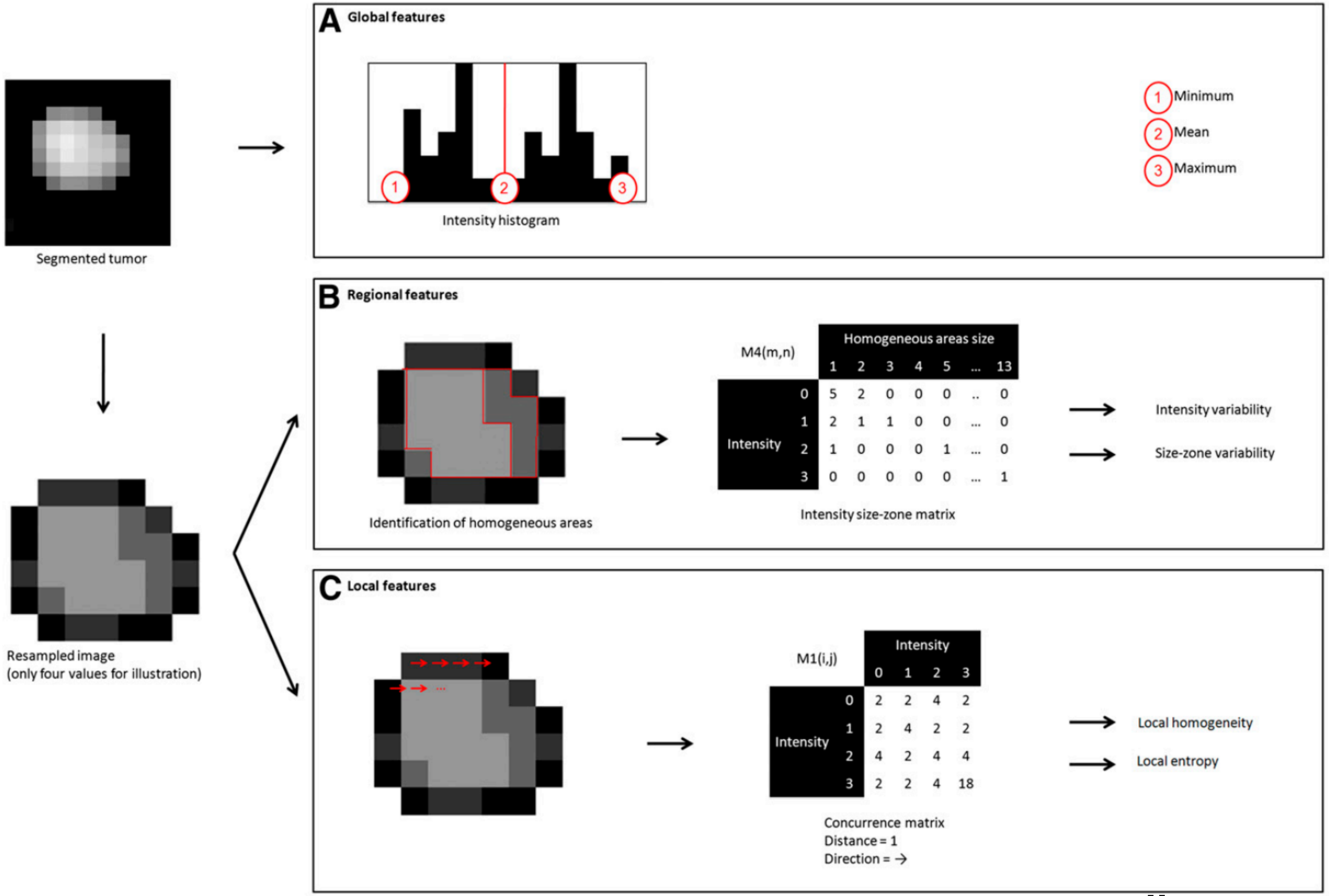
- For example, an image has a 50% black and 50% white distribution of pixels.



- Three different images with the same intensity distribution, but with different textures.

Credit to: Michael A. Wurth from Purdue Univers

Example Texture Analysis Strategy



Credit to: *M. Hatt, JNM*

Texture Analysis tutorial

UNIVERSITY OF CALGARY

LIBRARIES AND CULTURAL RESOURCES

Research | Services | Hours & Locations | Ask Us | My Account

DSpace Αρχική / Arts / Arts Research & Publications / Προβολή τεκμηρίου

Menu

GLCM Texture: A Tutorial v. 3.0 March 2017

Abstract

This tutorial describes both the theory and practice of the use of Grey Level Co-occurrence Matrix (GLCM) textures as originally described by Haralick and others in 1973. It leads users through the practical construction and use of a small sample image, with the aim of deep understanding of the purpose, capabilities and limitations of this set of descriptive statistics. Explanations and examples are concentrated on use in a landscape scale and perspective for enhancing classification accuracy, particularly in the cases where spatial arrangement of tonal (spectral) variability provides independent data relevant to the class identification. Background information is provided to answer the questions arising from 15 years of use of the tutorial, and increased practical experience of the author in teaching and research. Some information is provided to make the material accessible to specialists in fields other than remote sensing, for example medical imaging and industrial quality control. However the author is not an expert in these fields and texture's use there is not covered in detail. A basic bibliography is provided for research that has promoted the field of remote sensing GLCM texture; research projects that simply make use of it are not systematically covered.

Refereed

No

Of use generally for students of intermediate or advanced undergraduate remote sensing classes, and graduate classes in remote sensing, landscape ecology, GIS and other fields using rasters as the basis for analysis. Also useful for researchers undertaking the use of texture in classification and other image analysis fields. May be of use for algorithm and app developers serving these communities.

Προβολή

Texture tutorial including illustrations, examples and exercises with answers

Download

Texture tutorial including illustrations, examples and exercises with answers (1.939Mb)

Author

Hall-Beyer, Mryka

Accessioned

2017-04-04T17:21:09Z

Available

2017-04-04T17:21:09Z

Issued

2017-03

Subject

remote sensing

spatial descriptors

Department

Geography

Faculty

Arts

Αναζήτηση

Αναζήτηση στο DSpace

Αυτή η συλλογή

ΠΛΟΗΓΗΣΗ

Όλο το DSpace

Κοινότητες & Συλλογές

Ανά ημερομηνία δημοσίευσης

Συγγραφείς

Τίτλοι

Λέξεις κλειδιά

Αυτή η συλλογή

Ανά ημερομηνία δημοσίευσης

Συγγραφείς

Τίτλοι

Λέξεις κλειδιά

Ο ΛΟΓΑΡΙΑΣΜΟΣ ΜΟΥ

Σύνδεση

Εγγραφή

STATISTICS

Most Popular Items

What is texture?

Everyday texture terms - rough, silky, bumpy - refer to *touch*. They are most easily understood in relation to a topographical surface with high and low points, and a scale compatible with a finger or other.

A texture that is ***rough*** to touch has:

- a large difference between high and low points, as compared to the size of a fingertip, and
- a space between highs and lows approximately the same size as a fingertip.

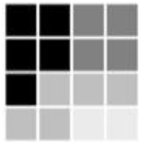
Rough textures can of course occur at any spatial scale, not just what we could touch. The illustration above would be considered rough whether it represents 1 cm or 100 km in horizontal dimension. To probe it with anything except our scale-informed eyes, however, we would have to adapt the "fingertip" used to the appropriate scale.

Silky or ***smooth*** has

- little difference between high and low points, and
- the differences would be spaced very close together relative to fingertip size.

Gray-level Co-occurrence Matrix Texture

The image as it appears:



The GL (digital numbers) associated with each pixel:

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

Original image:

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

The filled in east (1,0) spatial relationship GLCM. The entries in each cell are the sum of the counts (check boxes) as described above. The entries in colour refer to the three examples in the three matrices below.

0 0 1 1 0 0 1 1 0 2 2 2 2 2 3 3	0 0 1 1 0 0 1 1 0 2 2 2 2 2 3 3	0 0 1 1 0 0 1 1 0 2 2 2 2 2 3 3

neighbour pixel value -> ref pixel value:	0	1	2	3
0	2	2	1	0
1	0	2	0	0
2	0	0	3	1
3	0	0	0	1

Texture equations

Energy feature

$$Energy = \sum_{i,j=0}^{N-1} (P_{ij})^2$$

Entropy feature

$$Entropy = \sum_{i,j=0}^{N-1} -\ln(P_{ij})P_{ij}$$

Contrast feature

$$Contrast = \sum_{i,j=0}^{N-1} P_{ij}(i-j)^2$$

Homogeneity feature

$$Homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{ij}}{1+(i-j)^2}$$

Correlation feature

$$Correlation = \sum_{i,j=0}^{N-1} P_{ij} \frac{(i-\mu)(j-\mu)}{\sigma^2}$$

Shade feature

$$Shade = \operatorname{sgn}(A)|A|^{1/3}$$

Prominence feature

$$Prominence = \operatorname{sgn}(B)|B|^{1/4}$$

- aka *second order statistics*
- *The computation of large number of texture features lead to the advancement of Radiomics approaches*

https://prism.ucalgary.ca/handle/1880/51900

Ανάλυση υφής Gonzalez Woods

- Διαβάστε τις σελίδες 723 με 730 από το βιβλίο σχετικά με ανάλυση υφής!

A technical view on texture analysis

Statistical Approaches (1st order)

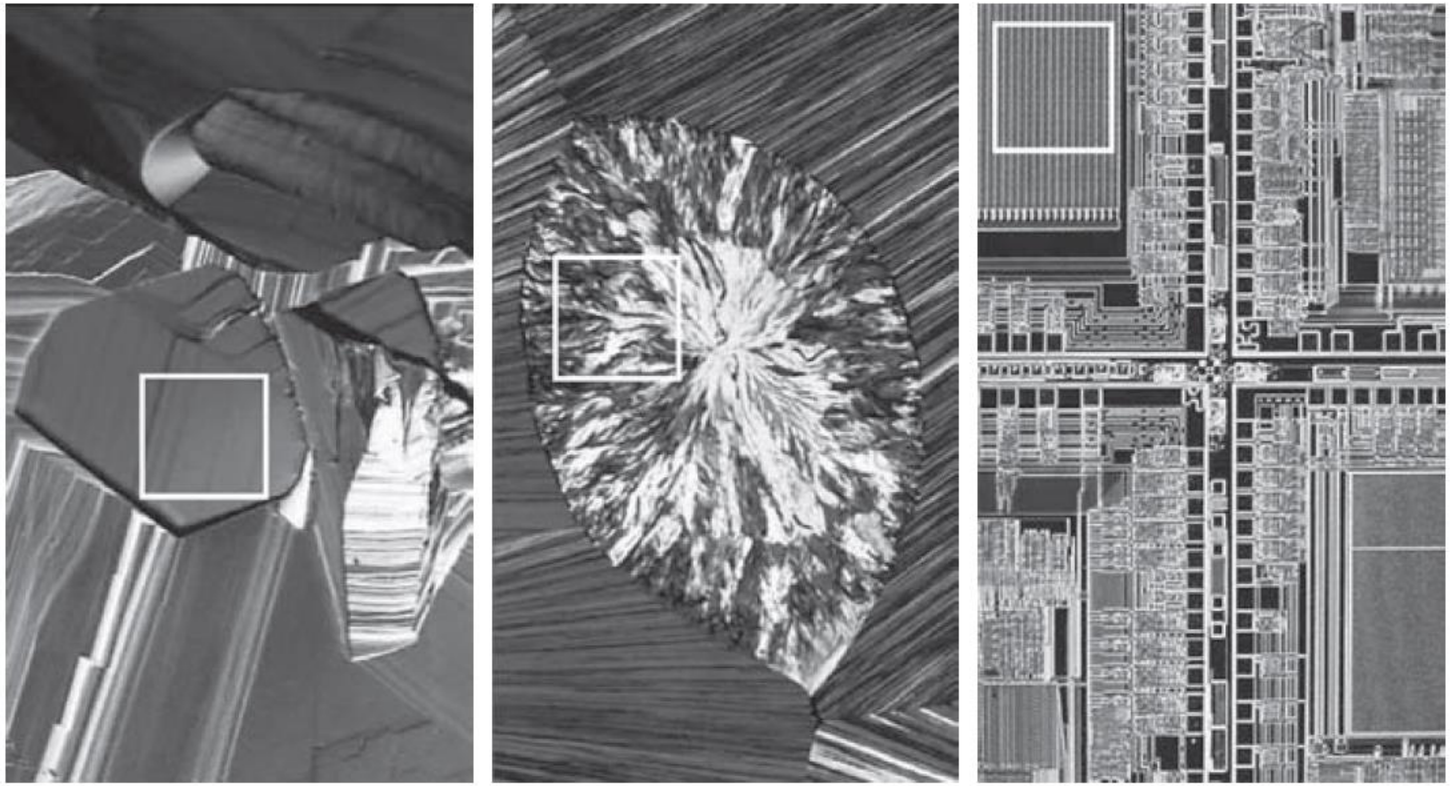
- One of the simplest approaches for describing texture is to use statistical moments of the intensity histogram of an image or region.
- Let *z* be a random variable denoting intensity, and let $p(z_i)$, $i = 0,1,2,\ldots,L-1$, be the corresponding normalized histogram, where *L* is the number of distinct intensity levels.
- the *n*th moment of *z* about the mean is

$$\mu_n(z) = \sum_{i=0}^{L-1} (z_i - m)^n p(z_i)$$

- where *m* is the mean value of *z* (i.e., the average intensity of the image or region):

$$m = \sum_{i=0}^{L-1} z_i p(z_i)$$

- The second moment [the variance $\sigma^2(z) = \mu_2(z)$] is particularly important in texture description. It is a measure of intensity contrast that can be used to establish descriptors of relative intensity smoothness.
- Η συνέχεια στο βιβλίο Gonzalez and Woods



a b c

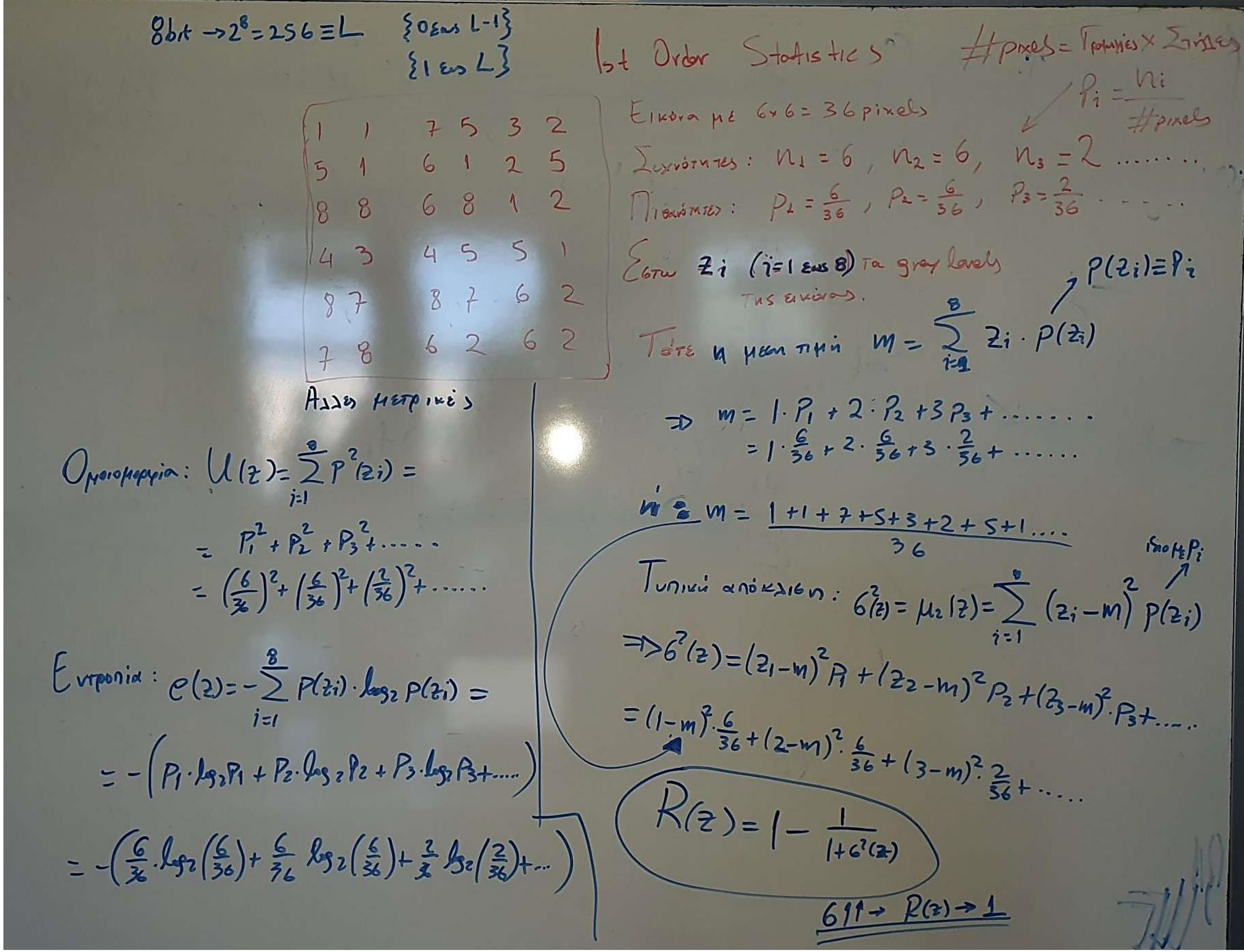
FIGURE 11.29 The white squares mark, from left to right, smooth, coarse, and regular textures. These are optical microscope images of a superconductor, human cholesterol, and a microprocessor. (Courtesy of Dr. Michael W. Davidson, Florida State University.)

TABLE 11.2
Statistical texture measures for the subimages in Fig. 11.29.

Texture	Mean	Standard deviation	*R* (normalized)	3rd moment	Uniformity	Entropy
Smooth	82.64	11.79	0.002	−0.105	0.026	5.434
Coarse	143.56	74.63	0.079	−0.151	0.005	7.783
Regular	99.72	33.73	0.017	0.750	0.013	6.674

Παράδειγμα από το μάθημα (1D)

Εδώ για να πάμε από συχνότητες σε πιθανότητες διαιρούμε με τον συνολικό αριθμό pixels (δηλ. γραμμές x στήλες εικόνας=36)



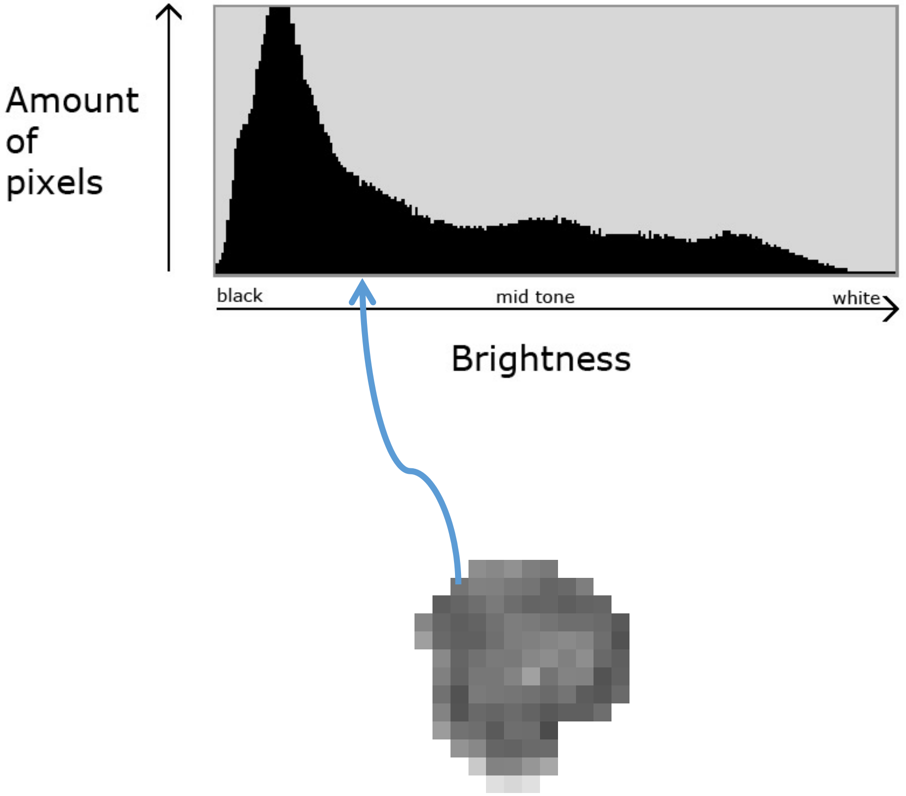
Statistical Approaches (2nd order)

Texture based on GLCM

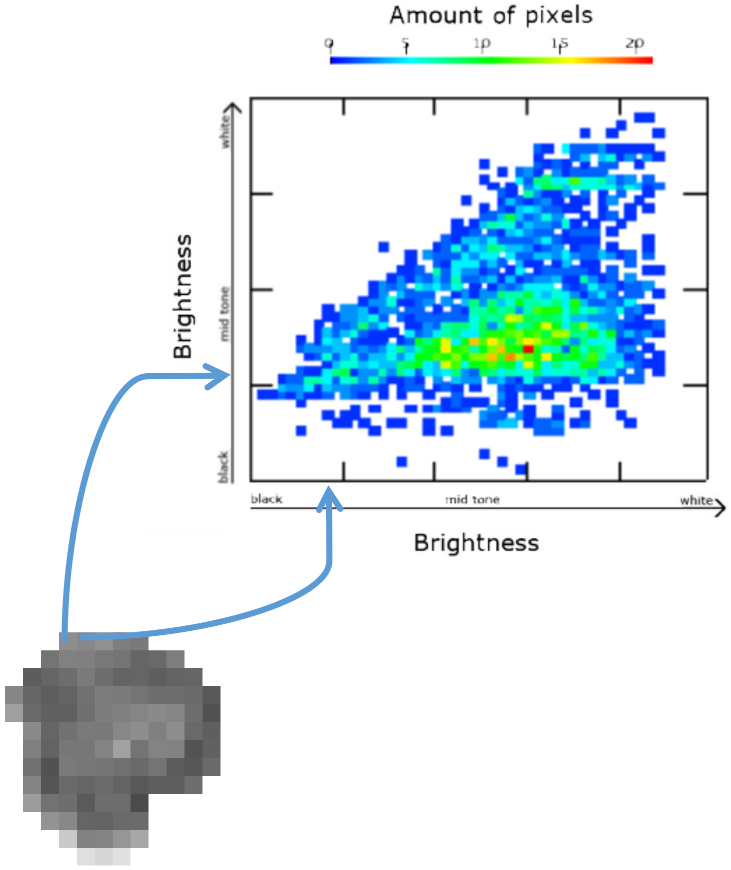
- Haralick *[1973 Haralick Robert M., Shanmugam K., Dinstein]*

Compared to what?

Method 1: Haralick features and feature maps.



Gray Level Histogram



Gray Level Coocurence Matrix (GLCM)

1	1	7	5	3	2
5	1	6	1	2	5
8	8	6	8	1	2
4	3	4	5	5	1
8	7	8	7	6	2
7	8	6	2	6	2

Image f

	1	2	3	4	5	6	7	8
1	1	2	0	0	0	1	1	0
2	0	0	0	0	1	1	0	0
3	0	1	0	1	0	0	0	0
4	0	0	1	0	1	0	0	0
5	2	0	1	0	1	0	0	0
6	1	3	0	0	0	0	0	1
7	0	0	0	0	1	1	0	2
8	1	0	0	0	0	2	2	1

Co-occurrence matrix **G**

Steps to create the GLCM of the ROI with orientation $\Theta = 0^0$ (→_), and distance d = 1.

Let **G** be a matrix whose element g_{ij} is the number of times that pixel pairs with intensities z_i and z_j occur in image f in the position specified by Q, where $1 < i, j < L$. A matrix formed in this manner is referred to as a *graylevel* (or *intensity*) *co-occurrence matrix*. When the meaning is clear, **G** is referred to simply as a *co-occurrence matrix*.

The number of possible intensity levels in the image determines the size of matrix **G**. For an 8-bit image (256 possible intensity levels), **G** will be of size 256×256. This is not a problem when working with one matrix but, as you will see in as Example 11.13, co-occurrence matrices sometimes are used in sequences. One approach for reducing computations is to quantize the intensities into a few bands in order to keep the size of **G** manageable. For example, in the case of 256 intensities, we can do this by letting the first 32 intensity levels equal to 1, the next 32 equal to 2, and so on. This will result in a co-occurrence matrix of size 8×8.

The total number, n, of pixel pairs that satisfy Q is equal to the sum of the elements of **G** ($n = 30$ in the example of Fig. 11.30). Then, the quantity

$$p_{ij} = \frac{g_{ij}}{n}$$

is an estimate of the probability that a pair of points satisfying Q will have values (z_i, z_j). These probabilities are in the range $[0, 1]$ and their sum is 1:

$$\sum_{i=1}^{K} \sum_{j=1}^{K} p_{ij} = 1$$

where K is the row and column dimension of square matrix **G**.

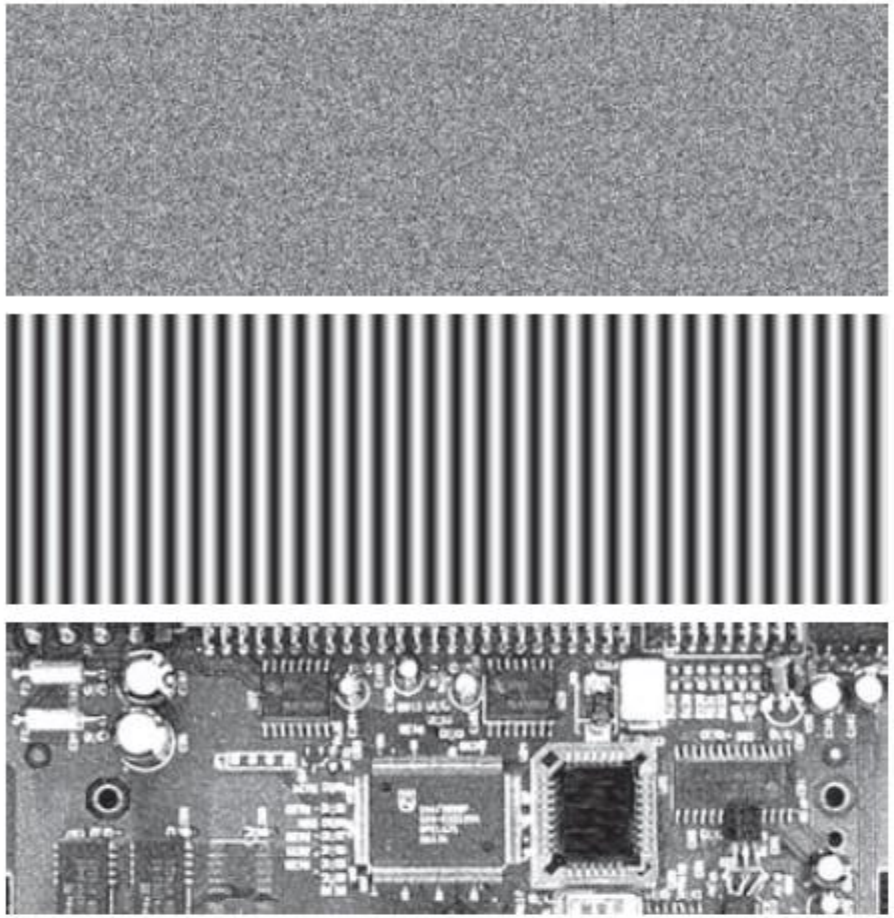
Because **G** depends on Q, the presence of intensity texture patterns can be detected by choosing an appropriate position operator and analyzing the elements of **G**. A set of descriptors useful for characterizing the contents of **G** are listed in Table 11.3. The quantities used in the correlation descriptor (second row) are defined as follows:

$$m_r = \sum_{i=1}^{K} i \sum_{j=1}^{K} p_{ij}$$

$$m_c = \sum_{j=1}^{K} j \sum_{i=1}^{K} p_{ij}$$

TABLE 11.3 Descriptors used for characterizing co-occurrence matrices of size $K \times K$. The term p_{ij} is the ij-th term of $\mathbf{G}$ divided by the sum of the elements of $\mathbf{G}$.

Descriptor	Explanation	Formula
Maximum probability	Measures the strongest response of $\mathbf{G}$. The range of values is [0, 1].	$\max_{i,j}(p_{ij})$
Correlation	A measure of how correlated a pixel is to its neighbor over the entire image. The range of values is 1 to −1 corresponding to perfect positive and perfect negative correlations. This measure is not defined if either standard deviation is zero.	$\sum_{i=1}^{K}\sum_{j=1}^{K}\frac{(i-m_r)(j-m_c)p_{ij}}{\sigma_r \sigma_c}$ $\sigma_r \neq 0;\ \sigma_c \neq 0$
Contrast	A measure of intensity contrast between a pixel and its neighbor over the entire image. The range of values is 0 (when $\mathbf{G}$ is constant) to $(K-1)^2$.	$\sum_{i=1}^{K}\sum_{j=1}^{K}(i-j)^2 p_{ij}$
Uniformity (also called Energy)	A measure of uniformity in the range [0, 1]. Uniformity is 1 for a constant image.	$\sum_{i=1}^{K}\sum_{j=1}^{K}p_{ij}^2$
Homogeneity	Measures the spatial closeness to the diagonal of the distribution of elements in $\mathbf{G}$. The range of values is [0, 1], with the maximum being achieved when $\mathbf{G}$ is a diagonal matrix.	$\sum_{i=1}^{K}\sum_{j=1}^{K}\frac{p_{ij}}{1+\lvert i-j\rvert}$
Entropy	Measures the randomness of the elements of $\mathbf{G}$. The entropy is 0 when all p_{ij}'s are 0, and is maximum when the p_{ij}'s are uniformly distributed. The maximum value is thus $2\log_2 K$.	$-\sum_{i=1}^{K}\sum_{j=1}^{K}p_{ij}\log_2 p_{ij}$



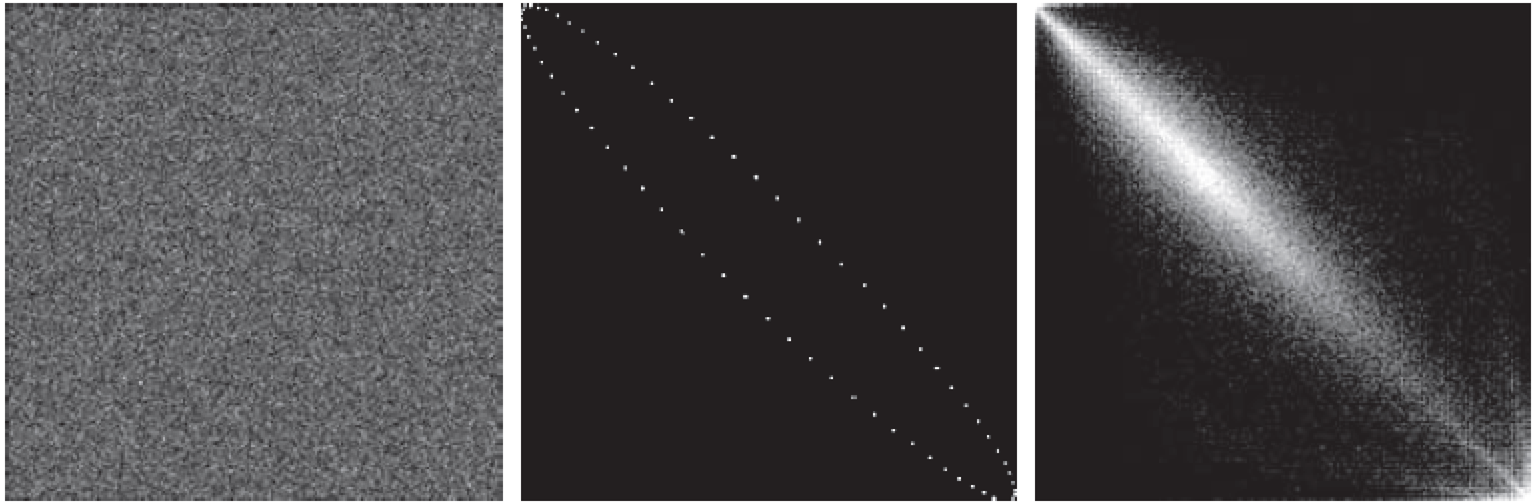
a
b
c

FIGURE 11.31 Images whose pixels have (a) random, (b) periodic, and (c) mixed texture patterns. Each image is of size 263×800 pixels.

variability in intensities. The high transitions in intensity occur at object boundaries, but these counts are low with respect to the moderate intensity transitions over large areas, so they are obscured by the ability of an image display to show high and low values simultaneously, as we discussed in Chapter 3.

The preceding observations are qualitative. To quantify the "content" of co-occurrence matrices, we need descriptors such as those in Table 11.3. Table 11.4 shows values of these descriptors computed for the three co-occurrence matrices in Fig. 11.32. To use these descriptors, the co-occurrence matrices are low with respect to the moderate intensity transitions over large areas, so they are obscured by the ability of an image display to show high and low values simultaneously, as we discussed in Chapter 3.

The preceding observations are qualitative. To quantify the "content" of co-occurrence matrices, we need descriptors such as those in Table 11.3. Table 11.4 shows values of these descriptors computed for the three co-occurrence matrices in Fig. 11.32. To use these descriptors, the co-occurrence matrices must be normalized by dividing them by the sum of their elements, as discussed earlier. The entries in Table 11.4 agree with what one would expect from the images in Fig. 11.31 and their corresponding co-occurrence matrices in Fig. 11.32. For example, consider the Maximum Probability column in Table 11.4. The highest probability corresponds to the third co-occurrence matrix, which tells us that this matrix has the highest number of counts (largest number of pixel pairs occurring in the image relative to the positions in Q) than the other two matrices. This agrees with our analysis of $\mathbf{G}_3$. The second column indicates that the highest correlation corresponds to $\mathbf{G}_2$, which in turn tells us that the intensities in the second image are highly correlated. The repetitiveness of the sinusoidal pattern in Fig. 11.31(b) indicates why this is so. Note that the correlation for $\mathbf{G}_1$ is essentially zero, indicating that there is virtually no correlation between adjacent pixels, a characteristic of random images such as the image in Fig. 11.31(a).

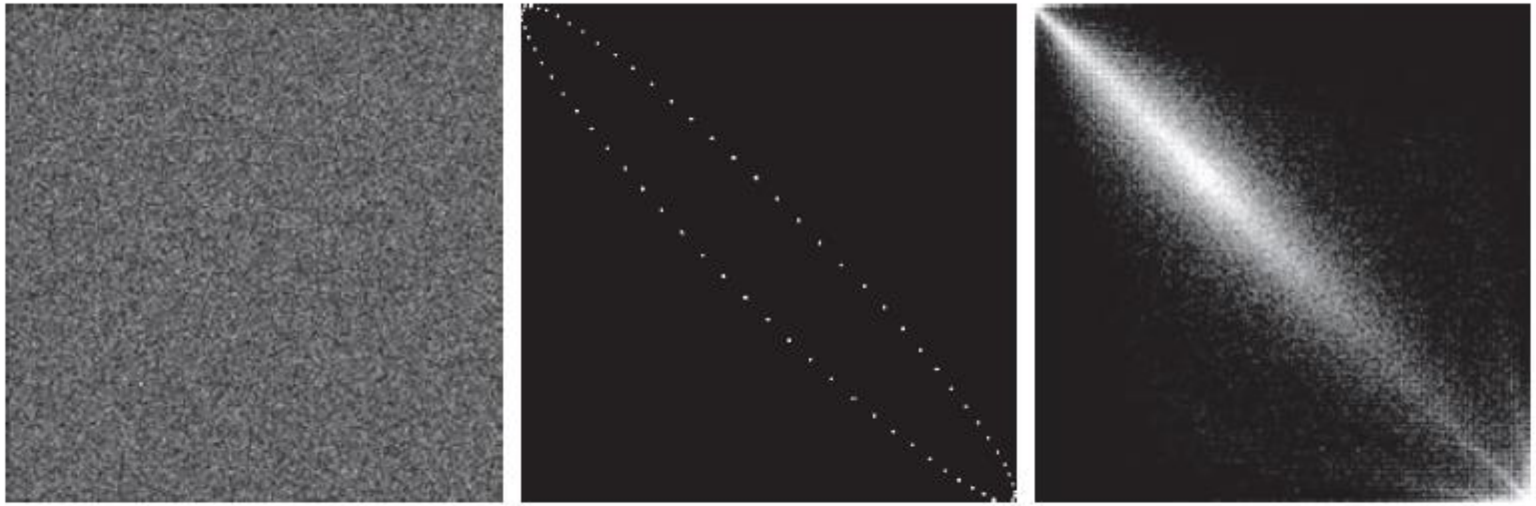


a b c

FIGURE 11.32 256×256 co-occurrence matrices $\mathbf{G}_1$, $\mathbf{G}_2$, and $\mathbf{G}_3$, corresponding from left to right to the images in Fig. 11.31.

TABLE 11.4
Descriptors evaluated using the co-occurrence matrices displayed as images in Fig. 11.32.

Normalized Co-occurrence Matrix	Maximum Probability	Correlation	Contrast	Uniformity	Homogeneity	Entropy
$\mathbf{G}_1/n_1$	0.00006	–0.0005	10838	0.00002	0.0366	15.75
$\mathbf{G}_2/n_2$	0.01500	0.9650	00570	0.01230	0.0824	06.43
$\mathbf{G}_3/n_3$	0.06860	0.8798	01356	0.00480	0.2048	13.58



a b c

FIGURE 11.32 256×256 co-occurrence matrices $\mathbf{G}_1$, $\mathbf{G}_2$, and $\mathbf{G}_3$, corresponding from left to right to the images in Fig. 11.31.

The contrast descriptor is highest for $\mathbf{G}_1$ and lowest for $\mathbf{G}_2$. Thus, we see that the less random an image is, the lower its contrast tends to be. We can see the reason by studying the matrix displayed in Fig. 11.32. The $(i-j)^2$ terms are differences of integers for $1 \leq i, j \leq L$, so they are the same for any $\mathbf{G}$. Therefore, the probabilities of the elements of the normalized co-occurrence matrices are the factors that determine the value of contrast. Although $\mathbf{G}_1$ has the lowest maximum probability, the other two matrices have many more zero or near-zero probabilities (the dark areas in Fig. 11.32). Because the sum of the values of $\mathbf{G}/n$ is 1, it is easy to see why the contrast descriptor tends to increase as a function of randomness.

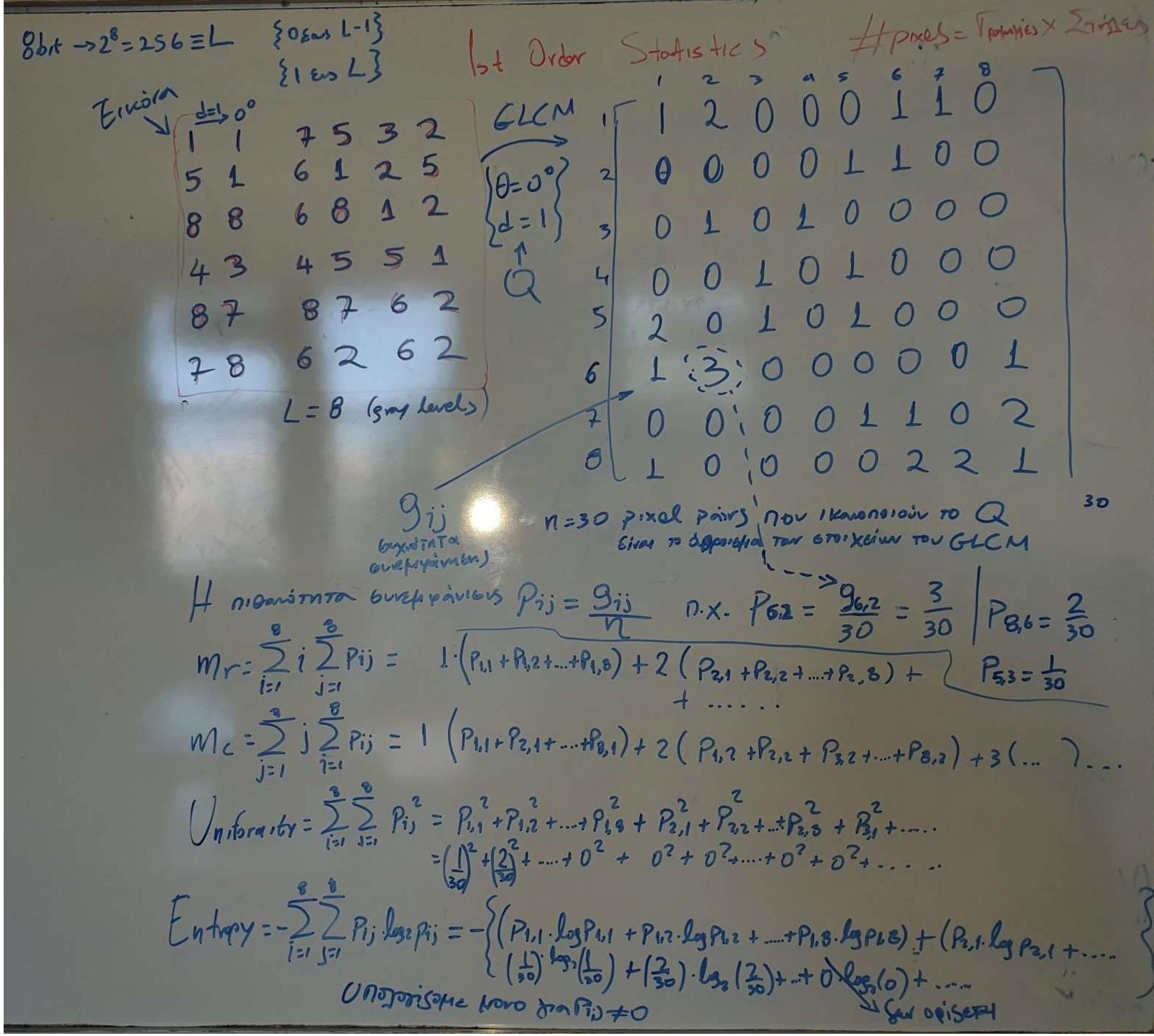
The remaining three descriptors are explained in a similar manner. Uniformity increases as a function of the values of the probabilities squared. Thus, the less randomness there is in an image, the higher the uniformity descriptor will be, as the fifth column in Table 11.4 shows. Homogeneity measures the concentration of values of $\mathbf{G}$ with respect to the main diagonal. The values of the denominator term $(1+|i-j|)$ are the same for all three co-occurrence matrices, and they decrease as i and j become closer in value (i.e., closer to the main diagonal). Thus, the matrix with the highest values of probabilities (numerator terms) near the main diagonal will have the highest value of homogeneity. As we discussed earlier, such a matrix will correspond to images with a "rich" gray-level content and areas of slowly varying intensity values. The entries in the sixth column of Table 11.4 are consistent with this interpretation.

The entries in the last column of the table are measures of randomness in co-occurrence matrices, which in turn translate into measures of randomness in the corresponding images. As expected, $\mathbf{G}_1$ had the highest value because the image from which it was derived was totally random. The other two entries are self-explanatory. Note that the entropy measure for $\mathbf{G}_1$ is near the theoretical maximum of 16 $(2\log_2 256 = 16)$. The image in Fig. 11.31(a) is composed of uniform noise, so each intensity level has approximately an equal probability of occurrence, which is the condition stated in Table 11.3 for maximum entropy.

Example of 2d texture extraction from class

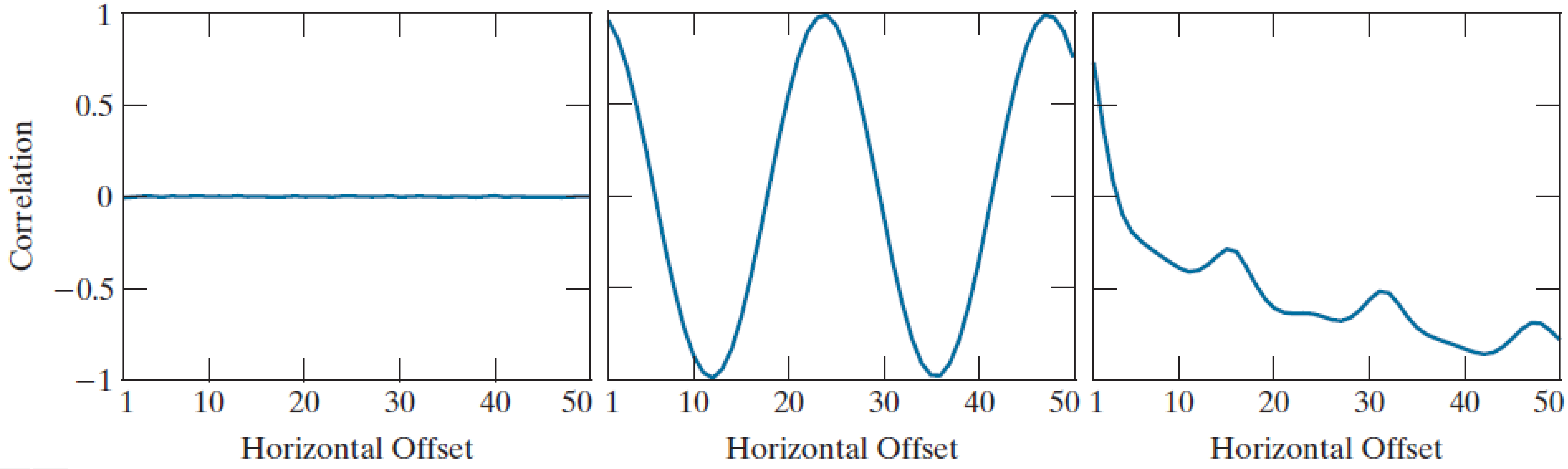
Παράδειγμα από το μάθημα (2D)

Εδώ για να πάμε από συχνότητες σε πιθανότητες διαιρούμε με τον συνολικό αριθμό συν εμφανήσεων (co-occurrences) δηλ το άθροισμα του πινακα GLCM=30)



Periodic/directional texture (from Gonzalez/woods)

- Thus far, we have dealt with single images and their co-occurrence matrices. Suppose that we want to “discover” (without looking at the images) if there are any sections in these images that contain repetitive components (i.e., periodic textures).
- One way to accomplish this goal is to examine the correlation descriptor for sequences of co-occurrence matrices, derived from these images by increasing the distance between neighbors.
- As mentioned earlier, it is customary when working with sequences of co-occurrence matrices to quantize the number of intensities in order to reduce matrix size and corresponding computational load.
- The following results were obtained using $L = 8$.
- Figure 11.33 shows plots of the correlation descriptors as a function of horizontal “offset” (i.e., horizontal distance between neighbors) from 1 (for adjacent pixels) to 50. Figure 11.33(a) shows that all correlation values are near 0, indicating that no such patterns were found in the random image.
- The shape of the correlation in Fig. 11.33(b) is a clear indication that the input image is sinusoidal in the horizontal direction. Note that the correlation function starts at a high value, then decreases as the distance between neighbors increases, and then repeats itself.
- Figure 11.33(c) shows that the correlation descriptor associated with the circuit board image decreases initially, but has a strong peak for an offset distance of 16 pixels. Analysis of the image in Fig. 11.31(c) shows that the upper solder joints form a repetitive pattern approximately 16 pixels apart (see Fig. 11.34). The next major peak is at 32, caused by the same pattern, but the amplitude of the peak is lower because the number of repetitions at this distance is less than at 16 pixels. A similar observation explains the even smaller peak at an offset of 48 pixels.



a b c

FIGURE 11.33 Values of the correlation descriptor as a function of offset (distance between "adjacent" pixels) corresponding to the (a) noisy, (b) sinusoidal, and (c) circuit board images in Fig. 11.31.

End of today’s lecture

Thank you for your attention!

Credits to Ulas Bagci and Michael A. Wurth from Purdue University
http://www.cyto.purdue.edu/cdroms/micro2/content/education/wirth06.pdf