Lecture 8 AND 9

Fourth stage

Medical Physical Department

Medical Imaging Processing

Image Statistics and Image Features

By

Asst. Prof. Dr. Mehdi Ebady Manaa

Image Statistics and Image Features

Medical images, on the whole, appear to be rather complex. They are filled with objects and shadows and various surfaces containing various patterns at a wide range of orientations. Amid this complexity, it may seem surprising that such images share any consistent statistical features. Consider the six images shown in Fig. 1, such images may seem widely different, but as a group they can be easily distinguished from a variety of other classes of image. A fundamental task in many statistical analyses is to characterize the location and variability of a data set.

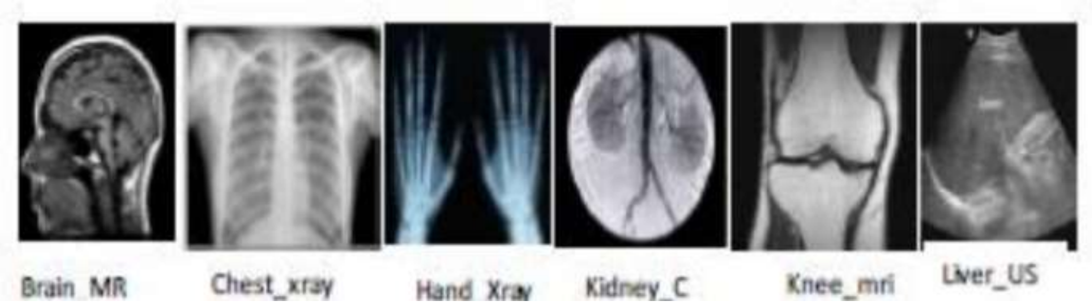

Figure1: Bench mark test medical images

Histogram features

The histogram features that we are considered are statically based features where the histogram is used as a model of the probability distribution of the gray levels. These statistical features provide us with information a bout the characteristic of the gray – level distribution for the image or sub image. Image statistics are investigated and studied based on first and higher order statistics that capture certain statistical regularities of medical images. We define the first – order histogram probability as follow:-

1. Mean

This is a average which indicates the general brightness of the image and is given by

$$\text{Mean}(\mu) = \frac{1}{M*N} \sum_{x=1}^{M} \sum_{y=1}^{N} p(x,y)$$

where ∑∑ p(x,y) represents the summation of all pixel values of the image and M*N is the size of the image. An image with a high mean indicates that the image is bright image and the dark image will have a low mean.

2. Standard deviation (σ)

The measure of the frequency distribution of a pixel value of an image is known as standard deviation of that image. The standard deviation can be calculated by given formula. [1 2 3; 4 5 6; 7 8 9]

$$\text{std dev}(\sigma) = \sqrt{\frac{1}{M * N} \sum_{x=1}^{M} \sum_{y=1}^{N} (p(x,y) - \mu)^2}$$

where ∑∑ p(x,y) represents the summation of all pixel values of the image and M*N is the size of the image. It describe the spread in the data, so a high contrast image will have a high variance, and a low – contrast image will have a low variance.

3. Variance

The variance is the square of the standard deviation and is calculated by the given formula.

$$\text{variance} = (\text{std dev})^2$$

An image with a high variance means that the image has a high contrast and an image with low variance indicates that the image has low contrast.

4. Energy

Energy is defined based on a normalized histogram of the image. Energy shows how the gray levels are distributed. When the number of gray levels is low then energy is high. If the energy is high, it tells us that the number of gray levels in the image is few, that is, the distribution is concentrated in only a small number of different gray levels.

Sometimes the energy can be a negative measure to be minimized and sometimes it is a positive measure to be maximized. The energy can be calculated in the given formula. [1 2 3; 4 5 6; 7 8 9]

$$\text{Energy (E)} = \frac{1}{M * N} \sum_{x=1}^{M} \sum_{y=1}^{N} p(x,y)^2$$

where MxN is the size of the image and p(x,y) is the value of the pixel of the image.

5. Entropy

Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image. Entropy is calculated for gray scale image and a

scalar value representing the entropy of gray scale image. If an image is thought of as a source of symbols or gray levels, then entropy is the measure of information content. The entropy of an image in this research is measured by $\log_2(4) = \log_2(2^2) = 2$

$\text{Log}_2(4) = \ln(4)/\ln(2) = 2$

(i) Normal entropy

$$\text{Normal Entropy } (E_n) = \sum_{x=1}^{M} \sum_{y=1}^{N} p(x,y) \log_2 (p(x,y))$$

(ii) Shannon entropy

$$\text{Shannon Entropy } (E_s) = -\frac{1}{M * N} \sum_{x=1}^{M} \sum_{y=1}^{N} p(x,y)^2 \log_2 (p(x,y)^2)$$

(iii) log energy

$$\text{Log Energy } (E_l) = -\frac{1}{M * N} \sum_{x=1}^{M} \sum_{y=1}^{N} \log_2 (p(x,y)^2)$$

the entropy is a measure that tells us how many bits we need to code the image data as the pixel values in the image are distributed among more gray levels, the entropy increases. This measure tends to vary inversely with the energy.

6. Measures of Skewness and Kurtosis

A further characterization of the data includes skewness and kurtosis. The histogram is an effective graphical technique for showing both the skewness and kurtosis of data set.

6.1. Skewness

Skewness is a measure of symmetry, or more precisely, the lack of symmetry. A distribution, or data set, is symmetric if it looks the same to the left and right of the center point. The skewness for a normal distribution is zero, and any symmetric data should have a skewness near zero. Negative values for the skewness indicate data that are skewed left and positive values for the skewness indicate data that are skewed right. By skewed left, we mean that the left tail is long relative to the right tail. The skewness is calculated by using the following formula. [1 2 3]

= - 1/64[(12-12)/5)^3+……………………..]

$$\text{Skewness}(s) = -\frac{1}{M*N}\sum_{x=1}^{M}\sum_{y=1}^{N}\left[\frac{p(x,y)-\mu}{\sigma}\right]^3$$

where

$$\text{Mean }(\mu) = \frac{1}{M*N}\sum_{x=1}^{M}\sum_{y=1}^{N}p(x,y)$$

$$\text{std dev}(\sigma) = \sqrt{\frac{1}{M*N}\sum_{x=1}^{M}\sum_{y=1}^{N}(p(x,y)-\mu)^2}$$

This method of measuring skew is more computationally efficient, especially considering that, typically, the mean and standard deviation have already been calculated. One of these techniques is to calculate the skewness of the data set. The normal distribution is perfectly symmetrical with respect to the mean, and thus any deviation from perfect symmetry indicates some degree of non-normality in the measured distribution.

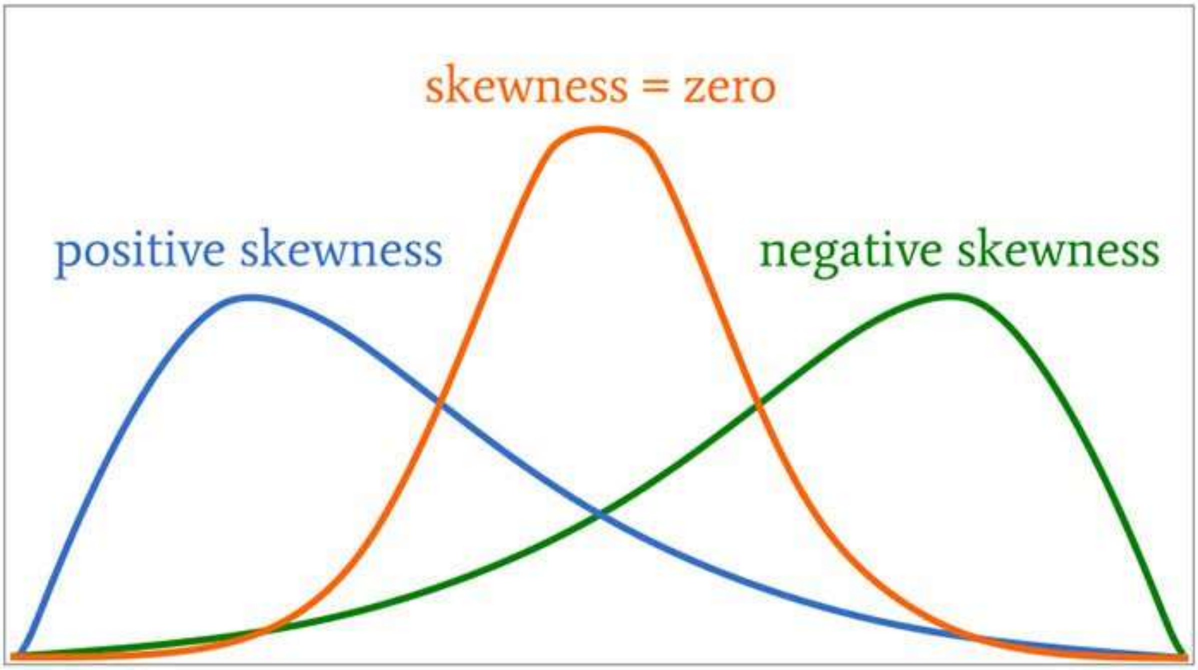

Figure 2: Skewness in positive, normal and negative distribution

6.2 Kurtosis

Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution. Data sets with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails. Data sets with low kurtosis tend to have a flat top near the mean rather than a sharp peak. A uniform distribution would be the extreme case see the figure (3) below.

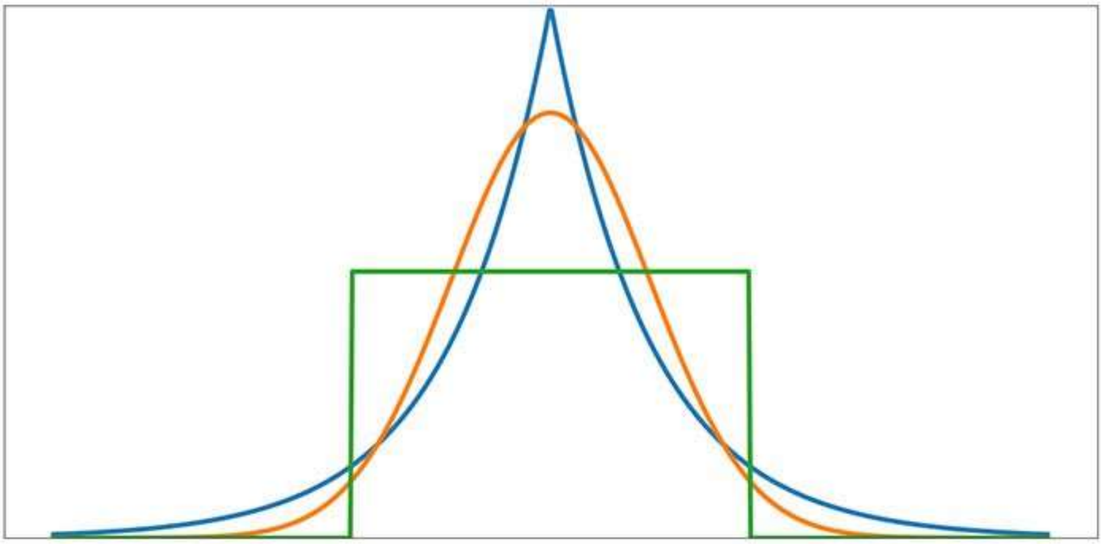

Figure 3: Three Types of Kurtosis Distribution

Kurtosis (k) is calculated by using the given formula

$$k = -\frac{1}{M * N}\sum_{x=1}^{M}\sum_{y=1}^{N}\left[\frac{p(x,y)-\mu}{\sigma}\right]^4 - 3$$

<u>Exercises</u>

Suppose we have the following small gray value image in the range from 0 to 19. Compute the following Image Statistics and Image Features:

a. Mean
b. standard Deviation
c. Variance
d. Energy
e. Entropy

12	6	5	13	14	14	16	15
11	10	8	5	8	11	14	14
9	8	3	4	7	12	18	19
10	7	4	2	10	12	13	17
16	9	13	13	16	19	19	17
12	10	14	15	18	18	16	14
11	8	10	12	14	13	14	15
8	6	3	7	9	11	12	12

f. Skewness

g. Kurtosis