
Automated Breast Cancer Detection Using FISH Spectral Linear Unmixing

Issa Ibraheem

Biomedical Engineering, Al-Andalus Private University for Medical Sciences, Al-Qadmus, Tartus, Syria

Email address:

issa.ibraheem@gmail.com (Issa Ibraheem)

To cite this article:

Issa Ibraheem. Automated Breast Cancer Detection Using FISH Spectral Linear Unmixing. *American Journal of biomedical and Life Sciences*. Vol. 3, No. 2-3, 2015, pp. 1-7. doi: 10.11648/j.ajbls.s.2015030203.11

Abstract: Fluorescence microscopy plays an important role in the classification of cancerous Tissue. The dramatic increase in multicolor fluorescence microscopy applications witnessed over the past decade is due, in part, to the significant advances in instrument and detector design. A number of advanced microscopy techniques have been applied using multi-color fluorescence labeling, including fluorescence recovery after photo bleaching (FRAP), fluorescence correlation spectroscopy (FCS), fluorescence resonance energy transfer (FRET), fluorescence in situ hybridization (FISH), and fluorescence lifetime imaging (FLIM). Many of these methods benefit significantly from the ability to use specifically targeted fluorescent proteins in live-cell imaging experiments. In addition, live-cell imaging has been revolutionized by the introduction of ever increasingly useful genetically encoded fluorescent proteins spanning the entire visible spectral region. However, the problem of fluorescence microscopy is the crosstalk between the channels caused by the overlap of the emission spectra of the different fluorophores, The crosstalk cannot be solved on the filter level, and not by specialized fluorophores. To eliminate the crosstalk the hyperspectral imaging using the spectra unmixing (algorithmically reduce the overlap of spectra) can be the possible way to reduce the errors in the classification of the tissue. Spectral imaging is the combination of commuter vision and spectroscopy. In addition, because every object of interest consists of more than one pixels, every pixel is dependent on its neighboring pixels. Thus, the spatial context of the image contains useful information for a classification and increase the sensitivity and specificity of a spectral classification.

Keywords: Fluorescence microscopy, for breast cancer, fluorescence in situ hybridization, FISH

1. Introduction

Breast carcinoma has been the most common form of cancer in women since the end of the nineteen seventies. Almost 10% of malignancies in women are diagnosed as a breast carcinoma, which represents 22% of all cancer cases in women [1]. 5% to 10% of the breast carcinoma are genetically conditioned. The risk for women, whose mother or sister had a breast carcinoma, is twice that of women without a positive family anamnesis. This risk increases by a factor between four and six if two family members developed cancer [2].

Fluorescence in situ hybridization (FISH) is a well established molecular cryptogenic method for genetic analysis. DNA probes are labeled with fluorescent dyes and hybridized to chromosomes in metaphase preparations or interphase nuclei. The probe hybridized to a defined target nucleotide sequence of DNA in the cell, and the dye emits a fluorescence in a particular wavelength, when excited by an excitation light

source. The emitted fluorescences using different fluorochromes with different wavelengths can be detected as multicolored pixels. Normally, to classify the cell parts, three fluorochrome are used. The major common problem with these multicolored measurements is the overlap of the emission spectra of the different fluorochrome[4]. The overlap increases with the number of the used fluorochromes. Other problem is the Autofluorescence of the tissue. Auto-fluorescence originates from tissue components as elastin and collagen. The substances show a unspecific broad banded fluorescence emission overlapping the useful signal and thus causing an decrease in image quality. Currently the pathologists use RGB color images of (Multi -FISH) samplesto make their diagnosis. The quality of these images could be enhanced by acquiring hyperspectral data. Moreover, applying spectral unmixing (SU) methods, within the hyperspectral imaging system measures the spectrum at each pixel in the image. The information content of these hyperspectral images is higher than in standard traditional color image (RGB- Images)

enabling SU method to unmix the overlapping emission spectra more effectively and reduce tissue auto fluorescence by 55% [3]. This allows classification algorithms to count characteristic fluorescent signals more reliably and support experts in their diagnosis.

2. Material and Methods

2.1. Fluorophores and Spectral Acquisition

Three fluorophores “Rhodamin 6G, Fluorescein, and Rhodamin B” were applied. For the data analysis with the linear spectral unmixing methods of the fluorophores rhodamine 6G and rhodamine B were mixed with the conditions Rh_6G / Rh_B of 1:3 or 3:1 with each other: Figure 1 shows the measured multispectral information. Here, a LCTF to generate the spectral image information is used. The Figure 1 (a) shows the Fluoreszenzreaktionen of Rh_6G (green circle), Rh_B (red circle) and the mixing ratios Rh_6G / Rh_B of 1:3 (yellow circle) and 3:1 (blue circle). The corresponding emission spectra are shown in Figure 1(b).

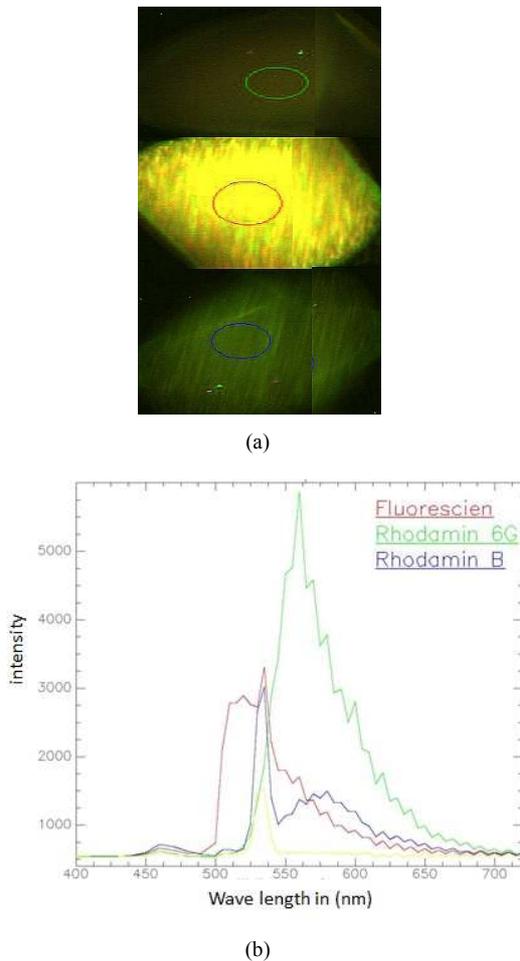


Figure 1. the fluorophores “Rhodamin 6G, Fluorescein, and Rhodamin B” and its spectrums acquired with the microscope used LCTF.

This captures individual images of the test measurement sample over the spectral range of 400-720nm in 10nm steps. For the measurements, rhodamine B and rhodamine 6G used

as fluorophores and a green laser with 532nm for the excitation.

2.2. Epidemiology of breast Cancer

Breast Carcinoma is caused by a malfunction in the cellular mechanisms which regulate growth[1]. Proto-oncogenes are normal genes that are responsible for the development and differentiation can cause these proto-oncogenes to change their behavior and become hyperactive and even non-physiological. HER2/neu (from human epidermal growth factor receptor 2) is proto-oncogenes belong to the family of tyrosine kinase receptors. It is a member of the epidermal growth factor receptor (EGFR/ERBB) family. Amplification or overexpression of this oncogene is shown to play an important role in the development and progression of certain aggressive types of breast cancer. In recent years the protein has become an important biomarker and target of therapy for approx. 30% of breast cancer patients Which has four subtypes HER1, HER2/neu, HER3 and Her4[2]. These receptors are involved in the growth and differentiation of cells. HER2/neu is one of a few evidence-based features for the diagnosis of breast carcinoma[1]. Normal breast epithelial cells have two HER2/new gene copies and between 20000 and 40000 HER2/new receptors. In the early stage 20% of breast carcinoma the HER2/neu is over expressed because of gene amplification [5]. These increases the number of HER2/neu receptors on the surface relative to normal breast epithelial cells [8].

CEP17 (chromosome 17 centromere enumeration probe). Increased CEP17 signals detected in invasive breast carcinomas may lead to discordant interpretation of gene amplification in a significant proportion of the cases, depending on which criterion is used for interpretation. However, increased gene, regardless of the evaluation method, is positively correlated with HER2 protein expression[9]. Figure 2 illustrates the different between HER2 and CEP 17 parts in breast normal and Cancer cells

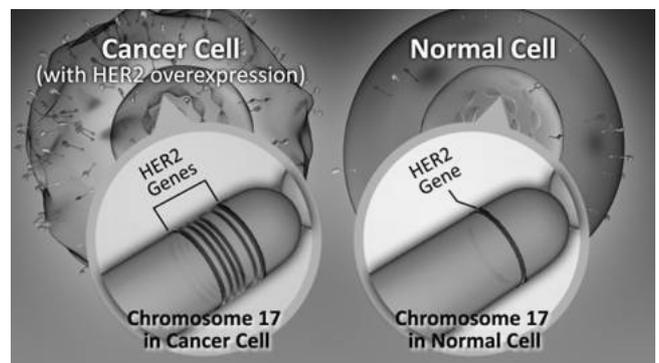


Figure 2. HER2 and CEP 17 in human normal and cancer cell. From (National Cancer Institute) [15]

2.3. Spectral imaging

There are different methodologies to acquire a hyper-spectral image. The classical approach is to spatial scan a sample with

a single point probe while recording spectral data for each point. This approach provides both spectral and spatial resolution, but, due to the acquisition time is not applicable for real-time application. As it is not generally possible to simultaneously record two spatial plus the spectral dimension of a spectral image using a 2D detector, i.e. some form of camera, either the wavelength information or one spatial dimension must be acquired sequentially. Two major spectral imaging principles have emerged wavelength scanning spectral imaging, in remote sensing better known as “staring imager”, and spatial scanning spectral imaging, also known as “push-broom scanning”.

For high quality detection and classification of florescent images, the spectral imaging coupled with image analysis using linear immixing can be employed to segregate mixed florescent signals and more clearly resolve the spatial contribution of each fuorophore. Microscopes are now available that have been specifically designed to accommodate spectral imaging and, although the technique bestows significant advantages, it increases the complexity and purchase price of the instrument. Spectral imaging merges the two well-established technologies of spectroscopy and imaging to produce a tool that has proven useful in a variety of disciplines that rely on various forms of optical microscopy.

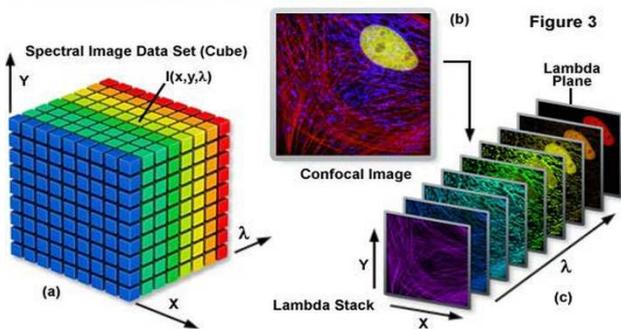


Figure 3. The spectral imaging Lambda stack. (a) data cube, (b) confocal image (c) lambda sock.

2.4. Spectral Unmixing

Linear spectral mixing. This methods is sometimes also called: Spectral Mixture Analysis (SMA: [10]) Or Mixture Modeling (MM: [11]). Each surface component within a pixel is sufficiently large enough such that no multiple scattering exists between the components (Singer and McCord, 1979). Each surface component within a pixel is sufficiently large enough such that no multiple scattering exists between the components. The linear scattering approximation is valid when the size of the pixel is smaller than the typical ‘patch’ or component being sensed - i.e. linear mixing occurs at the macroscopic scale.

2.5. Linear Unmixing Model

For an image with: “N” bands, “C” different cover types, $X_i = \{x^1, x^2, \dots, x^N\}^T$ are the observed image values in the i_{th} band, and $f_i = \{f_1, f_2, \dots, f_c\}^T$ the proportions of each pixel within

each cover type “c”.

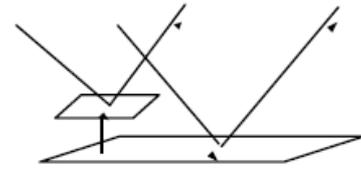


Figure 4. Florescent shifting of avelength

The Linear Unmixing Model is defined as: $x = Mf + e$, where, M is an $\{n \times c\}$ matrix.

The columns of this matrix represent the spectra of the different end members.

Spectral unmixing is a common way to resolve crosstalk in the emission spectra and reduce tissue auto fluorescence of fluorescence measurements. The method assumes that every pixel consist of a linear combination of overlapping emission spectra. In the last years there have been several implementation of spectral unmixing methods. Each fuorophore or absorbing dye, regardless of the degree of spectral overlap with other probes, has a unique spectral signature or emission fingerprint that can be determined independently and used to assign the proper contribution from that probe to individual pixels in a wavelength (lambda) stack. The result of the linear unmixing technique is the generation of distinct emission fingerprints for each fuorophore used in the specimen. To optimize the classification of the cell parts signals (nuclide) and the CEP17 signal spectral unmixing was used to reduce the effects of noise and auto fluorescence. Assuming that X is a $(n \times c)$ matrix with n rows containing the spectra and c columns related with the number of wavelengths. The rows of X are denoted by and represent the spectra x_{ref} is the reference spectrum of the kth class. Y is a matrix that contains the number of normalized reference spectrum x_{ref} .

$$x_i^T = (x_{i,1}, x_{i,2}, \dots, x_{i,c}) \quad (1)$$

$$X = \begin{bmatrix} X_{ref,1}^T \\ X_{ref,2}^T \\ \cdot \\ \cdot \\ X_{ref,l}^T \end{bmatrix} \quad (2)$$

The dispersion matrix

$$Y_i = \begin{bmatrix} Y \\ X_i \end{bmatrix} \quad (3)$$

Contains all the spectra of Y and the ith spectrum xi in the last row.

The pure spectra are the linear compensation of all spectra from the data set. The value is called ‘‘simplicity’’ can be used to estimate how close the eigenvectors to the pure spectrum

$$S_i = \frac{1}{N} \left(\sum_{i=1}^N X_i^4 \frac{\left(\sum_{i=1}^N X_i^2 \right)}{N} \right) \quad (4)$$

In order to use a linear mixture model there is a need to measure the spectral reflectance of the ‘pure’ end members[17-19]. Ideally ground-based spectra would be acquired to produce accurate end-members, since end members taken from even very high spatial-resolution imagery may contain multiple surface components. However, such errors can be minimized through sampling the image from within the center of known features or from known locations during prediagnosis.

Like PCA (principal Component analysis) the optimization procedure iteratively performs planar rotations of the loading matrix to maximize simplicity. The algorithm does not guarantee a global maximum and thus has to be repeated several times. There are relative and absolute stopping criterion for the iteration. The absolute value which is optimized until the difference $v_r v_{i+1} / v_i < \epsilon$ [12]

Assumed that spectra values and abundances are positive and that there is a set of candidate spectra that could be obtained (e.g. by orthogonal projection analysis). If they were definitely pure spectra then the abundance can be found by solving the least square problem [20-24].

Z_i is the pure concentration of aq component and can be calculated by:

$$Z_1 = (X \ Y_0^T) \left(Y_0^T \ Y_0^T \right)^T \quad (5)$$

However, in practice one has to deal with imperfect candidates what lead to situation that some abundances obtained by the least squares algorithm are negative, Simple factors are orthogonal and usually contain positive values, so some abundances will be almost for sure negative. In Alterating least square they are clipped to zero $Z_{1,c}$ and used again to recomputed spectra candidates with the help pf least square approach. Another problem are negative values. With

$$X = Z_{1,c} Y_1 \quad (6)$$

And

$$Y_1 = \left(Z_{1,c}^T Z_{1,c}^{-1} \right) \left(Z_{1,c}^T X \right) \quad (7)$$

They are clipped to zero resulting in $Y_{i,c}$. The algorithm is iterated until reaches the convergence criterion:

$$\partial Y_1 = (Y_{1,c} - Y_0) - \frac{tr(\partial Y_1^T \cdot \partial Y_1)}{tr(\partial Y_{1,c}^T \cdot \partial Y_1, c)} \quad (8)$$

Where tr is the trace of a matrix. It is the sum of the elements on the main diagonal of the determinat.

If the algorithm converged or the number of defined iterations is reached then $Y=Y_{1,c}$, $Z=Z_{1,c}$. Otherwise Y_0 is set to $Y_{1,c}$ and iteration go on.

2.6. Segmentation of CEP 17 signal and HER2-neu

After spectral unmixing the image containing the HER-2-neu signals and the image containing the CEP 17 signal are converted to binary images by the application of Otsu’s method, which is an automated thresholding procedure based on automatically histogram shape. The Algorithm assumes that the image to be thresholded contains two classes of pixels. Then calculates the optimum threshold separating those two classes so that their combined spread is minimum[12]. The core point in Otsu’s is the searching for the threshold value that minimizes the

$$Z = Z_1 \cdot Y_0$$

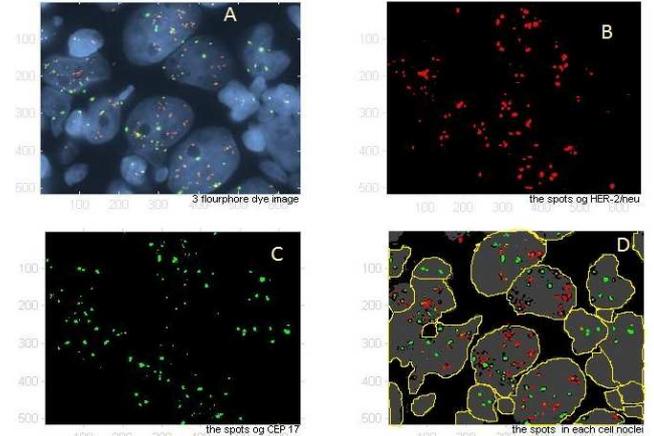


Figure 5. shows a ample of female breast cancer tissue dyed with three different fluorescent.

within-class variance, defined as the weighted sum of variances of the two classes.

$$\sigma_w^2(t) = w_1(t) \sigma_1^2(t) + w_2(t) \sigma_2^2(t) \quad (9)$$

Where w_i are weights, or the probability of two classes, t is the threshold and σ_i^2 are variance of the class i . The threshold and the variance used to separate the classes. Otsu shows that minimizing the within-class variance is the same as maximizing of the between class variance

Figure 5. Segmented HER-2/new signals, Segmented CEP 17 signals and cell nuclei. A the acquired image; B Intensity image of HER-2/neu signals; C Intensity image of CEP signals and D the Intensity image of HER-2/neu signals, CEP signals and cell nuclei. The borders of the cell nuclei are shown in yellow..

$$\sigma_b^2(t) = \sigma^2 - \sigma_w^2(t) = w_1(t) w_2(t) [\mu_1(t) - \mu_2(t)]^2 \quad (10)$$

$$D_{x,y} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (12)$$

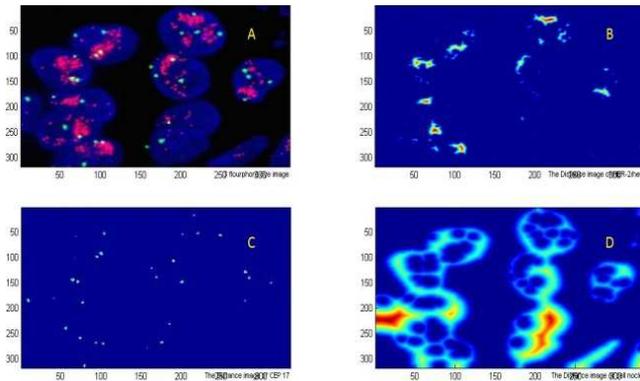


Figure 6. (A) the acquired image of breast cancer; (B) Distance images of the HER-2/new; (C) CEP17 and (D) cell nuclei Which is expressed in term of class probabilities w_i and class means, which in turn can be updated iteratively.

After the reduction of the unmixing result into binary image, morphological operations were applied to fit holes inside the areas of the HER-2/neu signals and CEP17 signals and to remove single pixels from the borders of the signals.

2.7. Spot Counting

The term spot counting refers to the determination of the number of HER-2/neu and CEP 17 signals and the calculation of HER-2/neu to CEP 17 ratio per cell nucleus. The final HER-2/neu to CEP 17 ratio is determined by dividing the HER-2/neu to CEP 17 ratio of all cell nuclei by the number of cell nuclei[3]. The results in a mean HER-2/neu to CEP 17 ratio \bar{R} .

$$\bar{R} = \frac{1}{N} \sum_{n=1}^N \frac{S_{HER}(n)}{S_{CEP17}(n)} \quad (11)$$

Where N is the number of nuclei, $S_{HER}(n)$ is the number of HER-2/neu signals and $S_{CEP17}(n)$ the number of CEP 17 signals in nucleus number n .

The number of nuclei, which have to be counted, is dependent on the HER-2/neu to CEP 17 ratio \bar{R} .

In case where the ratio is below 1.8 or above 2.2 it is recommended to count the signals of 20 nuclei (dependent on the role of Vysis, the company which manufactures the FISH staining kit)[13]. And if the ratio beyond 2.0 will receive Trastuzumab or “Hercepting” therapy, it is a humanized monoclonal antibody that acts on the HER-2-neu (erB2) receptors[14] After the calculation of distance images of the spots a median filter was used to avoid over segmentation of the different areas as it shown in figure 6.

The final segmentation was done using the watershed algorithm to separate the nuclei from the background and from each other. The resulting image is shown in figure 5 (D).

3. Results and conclusion

To determinate the HER-2/new to CEP n17 ratio, it was calculated by calculating the mean HER-2/neu to CEP 17 ratio for the founded nuclei in the sample data set. To calculate the number of signals for each nuclei, the classification results were the basic. Due to the fact that the CEP 17 signals do not form spatial cluster, because the spot size is small (less than 30 pixels), the counting of CEP 17 signal was directly performed on the classification result shown in figure 5(c). For the determination of HER-2/neu signal per nucleus the mean size of signals of HER-2/neu was defined (at least 100 pixel in each HER-2/neu cluster).

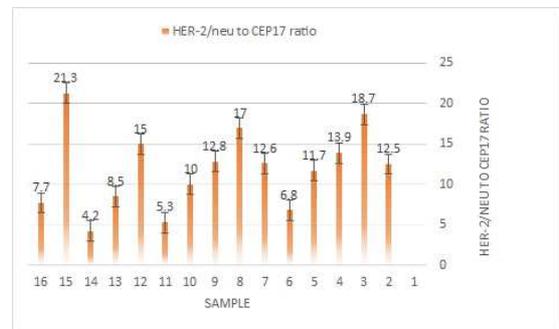


Figure 7. HER-2 to CEP 17 Ratue

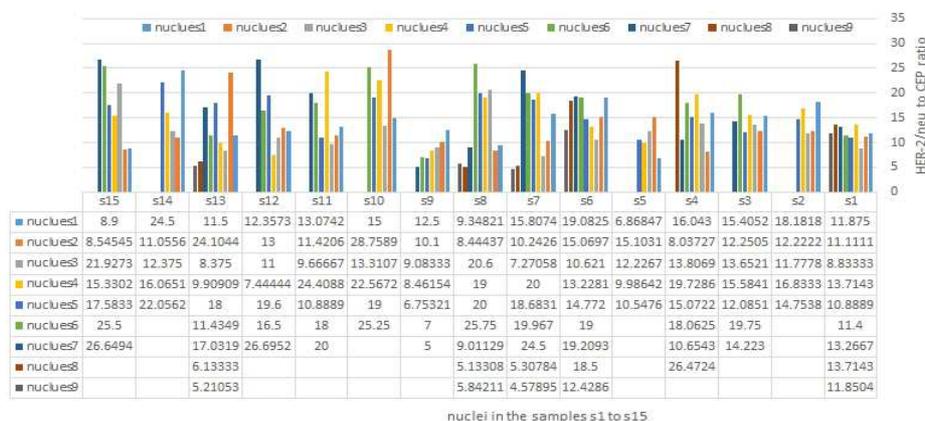


Figure 8. the group's results of the classification methods of the HER-2/CEP17 ratio

This was necessary to divide the spatial clusters of HER-2/neu signals into a number of single signals. The single signals of HER-2/neu and CEP 17 were counted to determine the ratio of HER-2/neu and CEP 17 by dividing the number of HER-2/neu to the number of CEP17 per cell nucleus. The final HER-2/neu and CEP 17 ratio is then determined by

HER-2/neu to CEP 17 ratios of all nuclei in the sample by the number of cell nuclei. The images of the cancer samples were converted to binary images by the application of Otsu method. to improve

The results, using the linear spectral unmixing for spot counting of HER-2/neu, CEP 17 and nuclei of 15 stained breast cancer samples (3-phlorophor) are shown, in the

Table 1 as statistical values on nuclei numbers, CEP 17 and HER-2/neu in each sample (different nucleus number) and the calculated ratio HER-2/neu to CEP 17 of each sample.

Table 1. The counted HER-2/neu, CEP 17 and the calculated ratio HER-2/neu to CEP 17 of each sample

Sample	Nucleus nr.	CEP 17	HER-2/neu	HER-2/neu to CEP17 ratio
1	8	16	200	12.5
2	4	6	112	18.7
3	6	9	125	13.9
4	8	18	210	11.7
5	6	17	115	6.8
6	9	16	201	12.6
7	4	6	102	17.0
8	5	9	115	12.8
9	4	9	90	10
10	6	20	105	5.3
11	7	14	210	15.0
12	7	12	102	8.5
13	8	25	105	4.2
14	5	10	213	21.3
15	6	15	115	7.7

Acknowledgements

I well thank Damascus university and Al-Andalus private university for medical sciences for supporting this work.

References

- [1] N. Harbeck, W. Eiermann, J. Engel, I. Funke, A. Lebeau, W. Parmanetter, and M. Mutsch. Prognosefaktoren beim primären Mammakarzinom, Vol 8, Zuckerschwendt, Muenchen, Bern, Wien, New York, 2001.
- [2] R. Lupu, M. Cardillo, M. Hijazi, L. Harris and K. Rosenberg. Interaction between erbB receptors and heregulin in breast cancer tumor progression and drug resistance. *Sem Cancer Biol*, 6: 135-145, 1995
- [3] Z. Mitri, T. Constantine, R. O'Regan (2012). "The HER2 Receptor in Breast Cancer: Pathophysiology, Clinical Use, and New Advances in Therapy". *Chemother Res Pract* 2012: 743193.
- [4] J. R. Lakowicz. Principles of Fluorescence Spectroscopy, volume 3. Springer Media LLC, 233 Spring Street, New York, NY 10013, USA, 2006.
- [5] L. Coussens, TL Yang-Feng, YC Liao, E Chen, A Gray, J McGrath, PH Seeburg, TA Libermann, J Schlessinger, U Francke (December 1985). "Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene". *Science* 230 (4730): 1132–9.
- [6] V Roy, EA Perez (November 2009). "Beyond trastuzumab: small molecule tyrosine kinase inhibitors in HER-2-positive breast cancer". *Oncologist* 14 (11): 1061–9.
- [7] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in *Magnetism*, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
- [8] M Tan, D Yu (2007). "Molecular mechanisms of erbB2-mediated breast cancer chemoresistance". *Adv. Exp. Med. Biol.* 608: 119–29
- [9] Ji-Won Kim, Jee Hyun Kim, Seock-Ah Im, Yu Jung Kim, Hye-Suk Han, Jin-Soo Kim, Kyung-Hun Lee, Tae-Yong Kim, Sae-Won Han, Yoon Kyung Jeon, Do-Youn Oh, Tae-You Kim, HER2/CEP17 ratio and HER2 immunohistochemistry predict clinical outcome after first-line trastuzumab plus taxane chemotherapy in patients with HER2 fluorescence in situ hybridization-positive metastatic breast cancer, July 2013, Volume 72, Issue 1, pp 109-115 OTSU N. *IEEE Trans. Syst. Man Cybern*, 1979, 9: 62-66.
- [10] Drake, N.A., Mackin, S. and Settle, J.J., 1999, Mapping vegetation, soils, and geology in semiarid shrublands using spectral matching and mixture modelling of SWIR AVIRIS imagery, *Remote Sensing of Environment*, 67, pp 12-25.
- [11] Wessman, C.A., Bateson, C.A. and Benning, T.L., 1997, Detecting fire and grazing patterns in tallgrass prairie using spectral mixture analysis, *Ecological Applications*, 7, 2, 493-511.
- [12] OTSU N. "A Threshold Selection Method from Gray-level Histograms," *IEEE Trans. Syst. Man Cybern*, 1979, 9: 62-66.
- [13] Path Vysion HER-2 DNA Probes Kit Package insert, Vysis Inc.
- [14] G. Viani, S.L. Afonso, E. J. Stefano, L.I. De Fendi and F. V. Soares, Adjuvant trastuzumab in the treatment of HER-2 Positive early breast cancer: a meta analysis of published radioized *BMC Cancer* [page 7:153, 2007.
- [15] Targeted Therapies for Breast Cancer Tutorial National Cancer Institute, at the National Health institute, http://www.cancer.gov/cancertopics/understandingcancer/targetedtherapies/breastcancer_htmlcourse/page3
- [16] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer; Auflage: first ed. 2006. Corr. 2nd printing 2011 (2007)
- [17] R.O.Duda, P.E.Hart and D.G. Strock, *Pattern Classification*, John Wiley & Sons; Auflage: 2. Auflage (21. November 2000)
- [18] T.M. Lillesand, R.W. Kiefer and J.W. Chipman, *Remote Sensing and Image Interpretation*, John Wiley & Sons, Hoboken, NJ, USA, 2004.
- [19] Richards, J.A. (1993), *Remote Sensing Digital Image Analysis*, 2nd ed. Springer Verlag, 1993.

- [20] Rissanen, J. (1978), "Modeling by shortest data description," *Automatica*, vol. 14, pp. 465-471, 1978.
- [21] Settle, J.J. (1996), "On the relationship between spectral unmixing and subspace projection," *IEEE Trans on Geoscience and Remote Sensing*, vol. 34, no. 4, pp. 1045-1046, July 1996.
- [22] Settle, J.J. and N.A. Drake (1993), "Linear mixing and estimation of ground cover proportions," *Int. J. Remote Sensing*, vol. 14, no. 6, pp. 1159-1177, 1993.
- [23] Zhao, X. (1996), *Subspace Projection Approach to Multispectral/Hyperspectral Image Classification Using Linear Mixture Modeling*, Master Thesis, Department of Computer Sciences and Electrical Engineering, University of Maryland Baitimare County, MD, May 1996.
- [24] Ren, H. and C.-1 Chang (2000b), "Target-constrained interference-minimized approach to subpixel target detection for hyperspectralimagery," *Optical Engineering*, vol. 39, no. 12, pp. 3138-3145, December 2000.