Maximum Likelihood and Spectral Angle Mapper and K-means algorithms used to detection of Melanoma

Issa Ibraheem

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

Email address:
issa.ibraheem@gmail.com

Abstract: Melanoma is a leading fatal illness responsible for 80% of deaths from skin cancer. It originates in the pigment-producing melanocytes in the basal layer of the epidermis. Melanocytes produce the melanin, (the dark pigment), which is responsible for the color of skin. As all cancers, melanoma is caused by damage to the DNA of the cells, which causes the cell to grow out of control, leading to a tumor, which is much more dangerous, if it cannot be found or detected early. Only biopsy can determine exact malformation diagnose, though it can rise metastasizing. When a melanoma is suspected, the usual standard procedure is to perform a biopsy and to subsequently analyze the suspicious tissue under the microscope. In this Paper, we provide a new approach using methods known as "Imaging Spectroscopy" or "Spectral Imaging" for early detection of melanoma. Spectral imaging can fill this gap of the classical imaging, which carries little spectral information while spectroscopy is severely limited in terms of measuring (potentially) inhomogeneous samples. Three different classifiers were applied, Maximum Likelihood ML and Spectral Angle Mapper SAM and K-Means. SAM rests on the spectral "angular distances" and the conventional classifier ML rests on the spectral distance concept. SAM and ML are two methods of the supported classification routines and K-Means is the known unsupported classification (clustering) algorithm.

Keywords: Melanoma; Spectral imaging; spectral spectroscopy; Maximum Likelihood; Spectral Angle Mapper, classification, K-Means clustering, Supported classification, unsupported classification, cancer detection

1. Introduction

Melanoma is the most serious form of skin cancer. It originates in melanocytes, i.e. pigment cells within the skin, which turn malignant and develop into a tumor. Malignant melanoma can be diagnosed by clinical and histological means. The first step usually is a clinical examination, in-vivo and non-invasive. Here, the discrimination between melanoma and e.g. benign nevi is performed based on visual features like Asymmetry (A), Boundary (B), Color (C), and Depth (D), what is known as "ABCD-Diagnostic Rule" for melanoma detection[1],[2]. This examination is relatively cheap but frequently not sufficient for a reliable diagnosis. In many cases, the results are used as an indicator whether a patient should be referred to a biopsy of a suspect skin region. Here the application of Spectral imaging to detect the Melanoma has a number of advantages. First, the spectroscopic measurement allows to reliably contactless, non-invasive and in-vitro measure spectra for each pixel in the melanoma object, second it is purely harmless optical methods, additionally the spectral data contain information about the color, material and concentration of the tissue. Furthermore, when using the spectral imaging sys-tem, scanning e.g. a 2 x 5 cm² area of the skin takes about 30 s, with the detection-results being available practically instantaneously. This short detection time resolution allows monitoring the development of the melanoma over time, thus providing even more information. Practically, two major Spectral Imaging (SI) principles have emerged wavelength scanning SI, in remote sensing better known as "staring imaging", and spatial scanning SI, also known as "push-broom scanning imaging" [4]5[6]

1.1. Wavelength Scanning S I

This method is essentially based on acquiring a number of single 2D-images of an identical sample, at different wavelengths. Hence, both spatial dimensions are acquired simultaneously, while the spectral information is acquired sequentially. Practically, the wavelength selection can be done either by a number of discrete filters, by tunable filters, namely acousto-optical tunable filters (AOTF) or liquid crystal tuna-

ble filters (LCTF) or by illumination of the sample at selected, discrete wavelengths. This method is highly useful in particular when only a few images at char-acteristic wavelengths have to be recorded, [7][8][9].

1.2. Push-broom Imaging SI

More suitable for many high throughput applications would be spatial scanning SI. The frequently used term "push broom scanning" originates from remote sensing and implies the line-wise acquisition of the image data, making use of a constant, relative movement (linear feed) between sample (skin) and imager (Camera) as it shown in Figure 1. Instead of recording a two-dimensional image, a line across the sample, perpendicular to the direction of the relative movement, is projected into an imaging spectrograph. The radiation originating along this observation line is spectrally analyzed and the spectral information for each pixel along the investigate line projected along the second axis of the two-dimensional detector chip. The spectral encoding can be provided either by dispersive optics forming an imaging spectrograph5 or by linearly variable filters. Since the spatial information along the line is retained, the computerized images contain the spatial information along the first axis and the full spectral wavelength information along the second axis. The spectral and the first spatial dimension are simultaneously acquired, while the second spatial dimension is recorded sequentially due to the movement of the sample relative to the SI sensor. By combining the slices, the second spatial axis can be derived, resulting in a full image. [7][8][9].
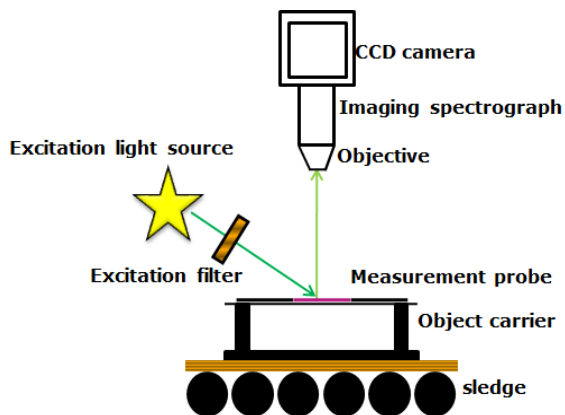

Figure 1. spectral imaging system- setup

In contrast to the stop-motion requirement of wavelength scanning SI, spatial scanning SI has a motion requirement, i.e. a continuous relative movement between imager and sample is a necessary pre-requisite for the operation. In case this is not provided as part of the process to be monitored, opting for a staring image may well be a better choice, as no moving mechanical parts would have to be added [3].
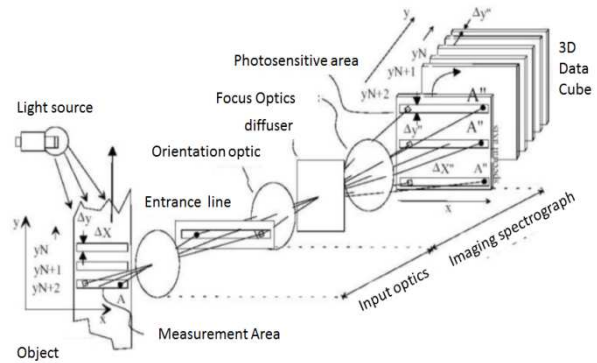

Figure 2. Principal function of spectral imaging using the imaging spectrograph (push broom imaging)

However independent on the acquiring methods (Wavelength scanning or push prom method) the Spectral Data consist of 3D-Data Matrix (Spectral Data Cube) (x, y, λ), where x, y are the special information and the third dimension λ refers to the spectral information as it shown in Error! Reference source not found.
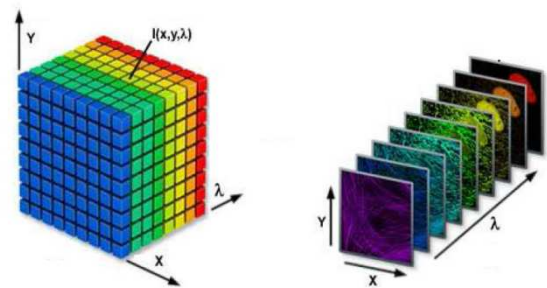

Figure 3. the spectral data cube

2. Methodology

The acquiring system is capable of capturing an image with a spatial axis of 480 pixels and a spectral axis of 480 pixels. Therefore, the spectral range from 380 nm to 780 nm is divided to 270 locations (bands), with spectral resolution of (10 nm). The SI system is designed so that the object table is moved by a linear table to implement the necessary relative movement between camera and sample. The region of the image, which will be examined, is typically traversed in 400 lines. Theoretically, each pixel of the acquired images corresponds to a rectangular area of approximately 0.1 μm x 0.1 μm. The effectively achievable spatial resolution is physically limited by the diffraction limitation to the order of magnitude of the wavelength of the transmitted light, i.e. 380- 780 nm.[3]. The system acquires the reflectivity of the light wave length, it is an indicator of the optical tissue properties in the wavelength range (in our study in VIS wavelength range). The reflectivity of each pixel in the measured object $R(x,y)$ can be calculated using the following calibration equation:

$$R(x, y) = \frac{I(x, y) - I_{Black}(x, y)}{I_{White}(x, y) - I_{Black}(x, y)} \tag{1}$$

where I(x,y) is the Intensity of measured pixel in the image, $I_{Black}(x,y)$ and $I_{White}(x,y)$ are the intensities of black- and white current consequently. Black current is the intensity if zero illumination (lens is covered) comes into the camera chip, while the white current is the intensity if the maximum illumination comes into the camera chip[5].

2.1. Skin, Melanoma and Moll Spectral Signatures

Reflectance spectra in the wavelength region from 380 nm to 700 nm were measured from 200 volunteers as training data and from 300 volunteers as test data[8]. The spectral signature of the Melanoma, healthy skin and moll are shown in the following figure
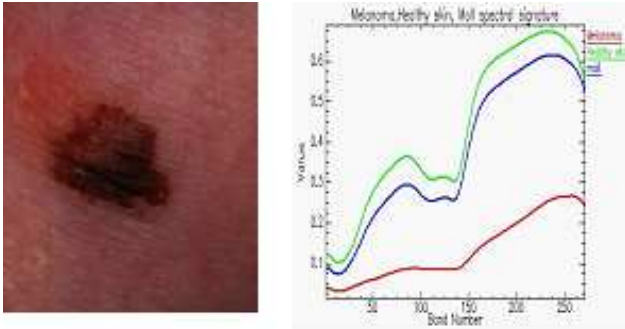

Figure 4. *2x2 cm-image of melanoma object (left); spectral signature of melanoma, Moll and Healthy skin (right)*

2.2. Detection Algorithms

Spectral classification methods were developed specifically for use on hyperspectral data, but they provide an alternative method for classifying multispectral data, often with improved results that can easily be compared to spectral properties of materials. In this Paper, the supervised as well as the unsupervised classification were used to cluster pixels in a dataset into classes corresponding to user defined training classes. It requires, using the supported classification, a training set, which must be defined for use as the basis for machine learning to build the discrimination function (recognition model) . Two supervised methods are then applied in this study to determine if a specific pixel qualifies as a class member [5]. The first one is the Maximum Likelihood (ML) while the other is the Spectral Angle Mapper (SAM) [5].

The k-means as unsupervised classification routine is used to order automatically each pixel in the spectral image in one class of different classes based on the squared Mahalanobis distance of each pixel to the centers of each clusters.

$$(\mathbf{x}_k - \hat{\mu}_i)^t \hat{\Sigma}_i^{-1} (\mathbf{x}_k - \hat{\mu}_i) \qquad (2)$$

Maximum Likelihood (ML)

Maximum likelihood classification is a supervised classification method derived from the Bayes theorem, which assumes that the statistics for each class in each band are normally distributed and calculates the probability that a given pixel belongs to a specific class[8]. The probability that a pixel with feature vector ω belongs to class *i*, is given by:

$$P(x|i) = \frac{P(w|i)P(i)}{P(w)} \qquad (3)$$

where $P(\omega|i)$ is the likelihood function, $P(i)$ is the a priori information, i.e., the probability that class *i* occurs in the study area and $P(\omega)$ is the probability that ω is observed, which can be written as:

$$P(w) = \sum_{i=1}^{M} P(w|i)P(i) \qquad (4)$$

Where M is the number of classes. ML often assumes that the distribution of the data within a given class i obeys a multivariate Gaussian distribution. It is then convenient to define the log likelihood (or discriminant function)

$$g(w) = \ln P(w|i) = -\frac{1}{2}(w - \mu_i)^T C_i^{-1}(w - \mu_i) - \frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln(|C_i|) \quad (5)$$

Pixel x is assigned to class *i* by the rule:

$$X \in i \quad if \quad g_i(w) \; \rangle \; g_j(w) \quad for \quad all \quad j \neq i \qquad (6)$$

Each pixel is assigned to the class with the highest value of g. Each pixel is assigned to the class with the highest likelihood or labelled as unclassified if the probability values are all below a threshold set by the user [6].

2.2.1. Spectral Angle Mapper (SAM)

Spectral Angle Mapper algorithm computes the "spectral angle" between the pixel spectrum and the training's pixel spectrum, i.e. (SAM), is a common distance metric, which compares an unknown pixel spectrum t to the reference spectra di, i = 1, ..,K, for each of K references and assigns t to the material having the smallest distance: This technique is comparatively insensitive to illumination and albedo effects. Smaller angles represent closer matches to the reference training's spectra. The result indicates the radian of the spectral angle computed using the following equation[6].

$$\alpha = \cos^{-1}\left(\frac{V_1^T V_2}{\|V_1\| \cdot \|V_2\|}\right) = \cos^{-1}\left(\frac{\sum_{i=1}^m T_i \cdot d_i}{\sqrt{\left(\sum_{i=1}^m T_i^2\right) \cdot \left(\sum_{i=1}^m d_i^2\right)}}\right) \quad (7)$$

Where m = the number of bands; Ti = (txi,tyi) is the i-pixel spectrum ; di =reference spectrum in training's data and α = radian of the spectral angle, (see figure 4).

The spectral angle classifiers we applied here rests on the spectral "angular distances," while the conventional classifier maximum likelihood rests on the spectral distance concept.[5]

$$\text{If } \alpha_i = \min(\alpha_{j,k}), \text{ then } x_i \in x_{j,k} \qquad (8)$$

where: x_i the spectral angel of the pixel x in test set. $x_{j,k}$ spectral angel of the pixel x the class k in training set.
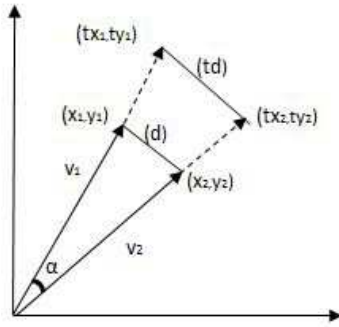
Figure 5. Spectral angle and spectral distance [6]

We can measure the similarity between two spectra **x** and **y** by using the Euclidean distance measure

$$\|\mathbf{x} - \mathbf{y}\| = \sqrt{\sum_{k=1}^{K} \left(x_k - y_k\right)^2} \qquad (9)$$

2.2.2. Training Set for the Supervised Classification

Using the spectral date of clinical diagnosed melanoma objects of 200 volunteers, we built a training set to learn the classification machine (classification routine) the scatterplot of training data is shown in the following figure:
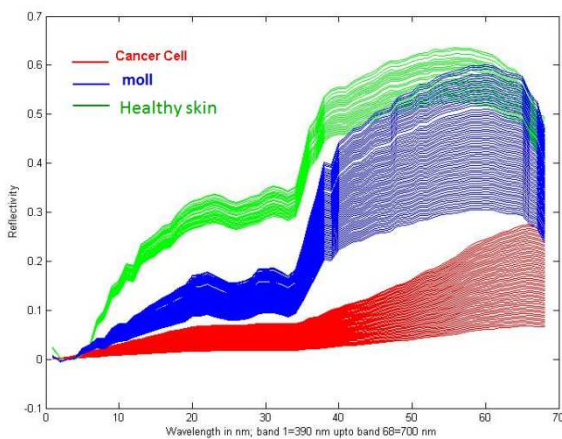

Figure 6. Spectrums of the Training data

2.2.3. k-means Unsupervised Algorithm

K-Means unsupervised classification calculates initial class means evenly distributed in the data space, then iteratively clusters the pixels into the nearest class using a minimum-distance technique. Each iteration recalculates class means and reclassifies pixels with respect to the new means. All pixels are classified to the nearest class unless a standard deviation or distance threshold is specified, in which case some pixels may be unclassified if they do not meet the selected criteria. This process continues until the number of pixels in each class changes by less than the selected pixel change threshold or the maximum number of iterations is reached. it is clear that the probability in the equation (4) is large when the squared Mahalanobis in equation (2) is small. Suppose that we merely compute the squared Euclidean distance $|x_k - \mu_i|^2$, find the center of the cluster (the mean μ_m) nearest to x_k and approximate the probability as

$$\hat{P}(\omega_i|\mathbf{x}_k, \hat{\theta}) \approx \left\{ \begin{array}{ll} 1 & \text{if } i = m \\ 0 & \text{otherwise.} \end{array} \right. \qquad (10)$$

It is to minimize the function of the square distance in each iteration and compare it with its previous value up to reach the smallest different between the actual and previous values of the distance as it illustrated in the following figure
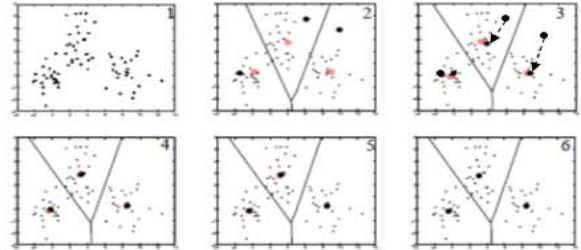

Figure 7. The mean value of the centroids in each iteration

From a statistical point of view, it may be inappropriate to use K-Means clustering since K-Means cannot use all the higher order information that PCA or ICA provides. There are several approaches that avoid using K-means,. However, for large images this algorithm fails to converge. A 2-stage K-means clustering strategy is developed that works particularly well with skin data:

1. Drop spectral data that contain only noise or correspond to artifacts.
2. Perform K-Means clustering with 5 clusters.
3. Those clusters that correspond to healthy skin are taken together into one cluster. This cluster is labelled as *skin*.
4. Perform a second run of K-Means clustering on the remaining clusters (inflamed skin, lesion, etc.). This time use 3 clusters. Label the clusters that correspond to the mole and melanoma center as *mole* and *melanoma*. The remaining clusters are considered to be 'regions of normal skin or unclassified regions'.
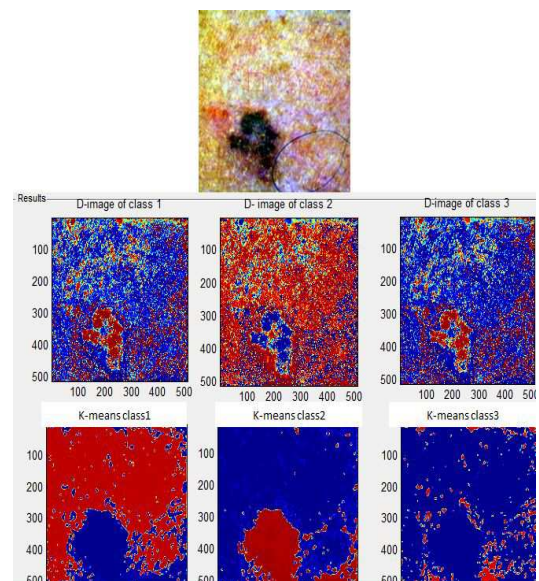

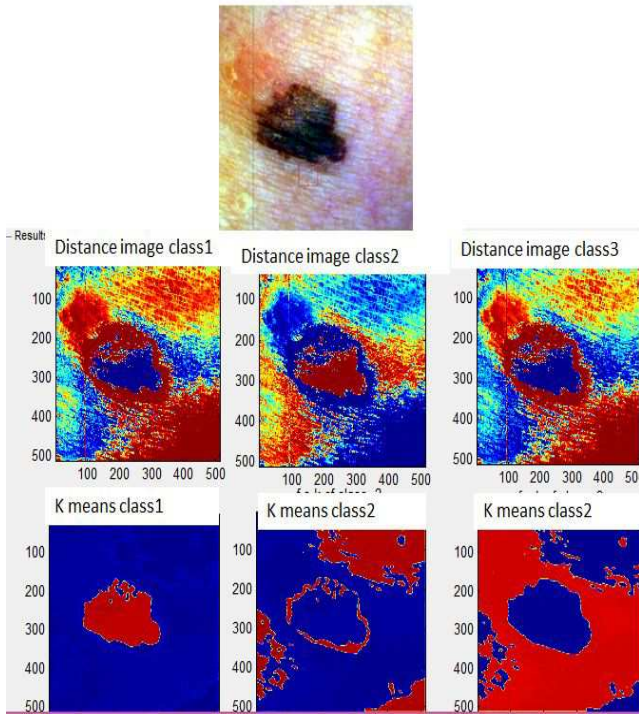*Figure 8. k-means classification of melanoma object*

Figure 9. *k-means classification of melanoma object*

In table3 the confusion matrix of the classification of each class using the unsupervised K-means algorithm, based on the truth-values used in the training set (diagnosed by dermatologist).

Table 1. *confusion matrix of k-means unsupervised classification*

Confusion Matrix (Memory1) 512x512x1				
Overall Accuracy (192858/286961) 67.2070%				
Ground Truth (Pixels)				
Clases	Class1	Class2	Class3	Total
Unclassified	0	0	0	0
Class1	16117	2227	7	18351
Class2	0	1288438	14738	170176
Class3	0	50131	48303	98434
Total	16117	180796	90048	286961
Ground Truth (Percent)				
Clases	Class1	Class2	Class3	Total
Unclassified	0	0	0	0
Class1	100	1.23	0.01	6.39
Class2	0	71.04	46.35	59.30
Class3	0	27.73	53.64	34.30
Total	100	100	100	100

2.2.4. Test Set

300 objects were tested using the Maximum Likelihood. (ML) , Spectral Angle Mapper (SAM) and K-means.

The results show that the ML and SAM classifiers were for pixel as well as for object classification more efficient than K-means. However, K-means was more flexible because it does not need to be trained. Some result-samples shown in

To compare the results of the applied ML, SAM and K-means, we build the confusion matrix of the tested classes, Melanoma, Moll and Healthy skin.
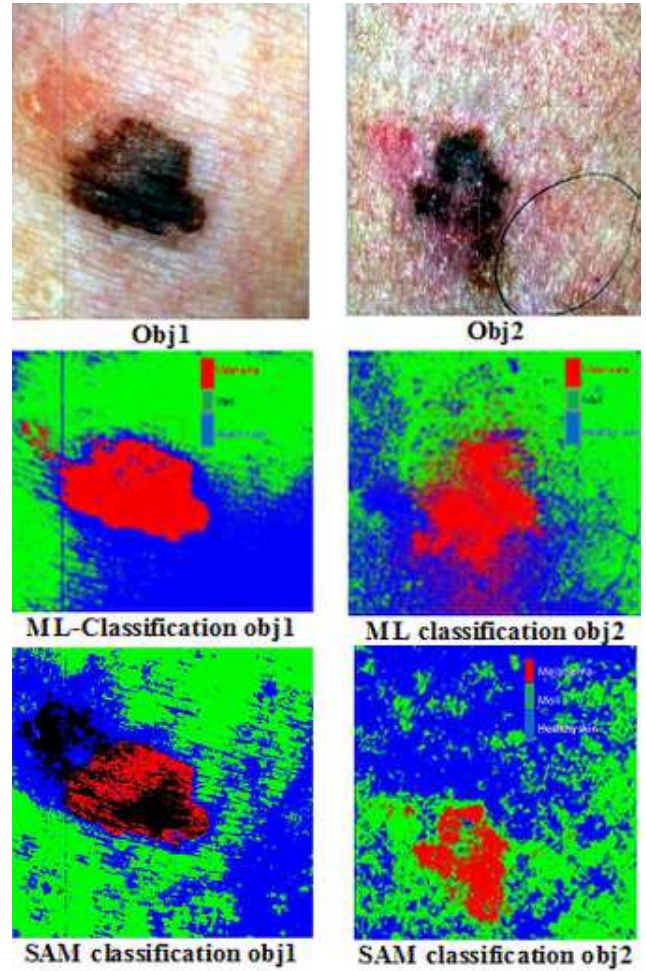

Figure 10. *Classification results using ML and SAM of melanoma object*

Table 2. *confusion matrix of LM supervised classification*

Confusion Matrix (Memory1) 512x512x1				
Overall Accuracy (192858/286961) 67.2070%				
Ground Truth (Pixels)				
Clases	Class1	Class2	Class3	Total
Unclassified	0	0	0	0
Class1	16114	0	0	16114
Class2	3	114798	30645	145441
Class3	0	66003	59403	125406
Total	16117	180796	90048	286961
Ground Truth (Percent)				
Clases	Class1%	Class2%	Class3%	Total %
Unclassified	0	0	0	0
Class1	99.98	0	0	5.62
Class2	0.02	63.51	34.03	50.68
Class3	0.0	36.51	65.97	43.7
Total	100	100	100	100

Table 3. *confusion matrix of SAM supervised classification*

Confusion Matrix (Memory1) 512x512x1				
Overall Accuracy (192858/286961) 67.2070%				
Ground Truth (Pixels)				
Classes	Class1	Class2	Class3	Total
Unclassified	1788	374	259	2421
Class1	13252	2689	731	16672
Class2	1058	123456	35870	160384
Class3	1938	64023	54323	88346

Confusion Matrix (Memory1) 512x512x1			
Overall Accuracy (192858/286961) 67.2070%			
Total	16217	188766	96548	301531
	Ground Truth (Percent)			
Classes	Class1%	Class2%	Class3%	Total %
Unclassified	11.03	0.20	0.27	11.49
Class1	81.72	1.42	0.76	83.90
Class2	0.97	85.00	11.26	97.24
Class3	11.95	2.13	87.34	89.47
Total	100	100	100	100

In table1 and table 2 the confusion matrix of the classification of each class using the supervised LM and SAM algorithms, based on the truth-values used in the training set (diagnosed by dermatologist). In table 3 the confusion matrix of k-means algorithm for pixel classification.

In table 4 the true positive classification of ML, SAM and K-means for each class (Melanoma, Moll and Healthy skin)

Table 4. Cl.ssifier true positive results using ML, SAM and K-menas

	ML	SAM	Kmeans
Melanoma	88.28%	81.83	79
Moll	92.28%	86.98	84
Healthy skin	93.17 %	87.92	85

The sensitivity, specificity, positive predictive value and negative predictive value of each class is calculated using the true positive, true negative, false positive and false negative arguments

Table 5. confusion matrix of ML Classification

		Classificationion results ML		
Ground Truth			Melanoma	Moll	Healthy skin
	Melanoma	%88.28	%6.12	%5.6
	Moll	%6.49	%92.28	%1.23
	Healthyskin	5.23	%1.6	%93.17

Table 6. confusion matrix of SAM Classification

		Classification results SAM		
Ground Truth		Melanoma	Moll	Healthy skin
	Melanoma	81.72%	0.97%	0%
	Moll	1.42%	85.00%	2.13%
	Healthy skin	0.76	11.26%	87.34%

Table 7. confusion matrix of K-means Classification

		Classification results K-means		
Ground Truth		Melanoma	Moll	Healthy skin
	Melanoma	79 %	8%22.	%11.95
	Moll	14.12%	84.8%	%1.2
	Healthy skin	4.5	10%	85.5%

3. Results

From confusion matrix of ML-, SAM K-means classifiers in table 4, it is clear that the ML-true-positive of Melanoma 88.28% is higher than SAM-true-positive 81.83% and K-means true positive 79%. These results show that, the difference between ML, SAM and K-means is not too high. The value of the "false negative" using ML in table 5 (Melanoma classified as Moll =6.12%), and (Melanoma classified as

healthy skin = 5.6%), while false negative using SAM in table 6 (Melanoma classified as Moll =1.42 %) and false negative (Melanoma classified as healthy skin =11.95 %). False negative, using SAM, is two wise greater than it using ML. False Negative ratio is a danger factor, because it very dangerous to classify a melanoma object as a Moll or as a healthy skin, "melanoma is not detected!"

False Positive using ML (Moll classified as Melanoma = 6.49% and Healthy skin classified as Melanoma = 1.23%), while False positive using SAM (Moll classified as Melanoma = 12.89% and Healthy skin classified as Melanoma = 0.13%). The False Positive using K-means is 1.2%

The Values of confusion matrix mean that the ML- Classifier is more robust to detect and classify the skin Melanoma. Because the true positive of ML is higher and the false negative is lower than SAM, as it shown in Table 1, Table 2 and Table 3.

Despite the quite small data set, the results are promising and a second follow-up study in University clinic of Damascus with a larger number of patients has been started yet to support these results and to find, if we, using this approach, could also detect and evaluate other skin abnormalities like psoriasis or and cartisuma a.o.

4. Conclusion

In this report, we have proposed a new scheme that allows to classify melanoma as pigmentation lesions of skin using multi-spectral images applying three different classification algorithms: ML, SAM as supervised classifiers and K-means as unsupervised classifier. The obtained results on 300 melanoma objects in clinical study tend to show that the spectral imaging method as new technology is robust and usable in Vivo and non-invasive diagnostic method.

The fact that the supervised classification algorithms interacts at the last step of the classification can be seen as a benefit tool compared with the unsupported classification algorithms. Because it allows to both make a miss or over classification control and make the classification to be based only on machine learning techniques, which are often controllable and evaluable.

In a possible application, where the physician is assisted by a system which pre-screens patients, we have to take care about high sensitivity which is typically accompanied with a loss in specificity. Preliminary experiments showed that a true positive of 88% using ML or 81% using SAM is possible at the cost of less than 15% false-positives using MI and SAM.

K-Means provided only 7% true positive.

Acknowledge

Reference

[1] I. Ibraheem, R. Leitner, H. Mairer, L. Cerroni and, J. Smolle, Hyperspectral analysis of stained histological preparations for the detection of melanoma. Proceeding of third International Workshop on Spectral Imaging, Graz, 13 May 2006.

[2] I. Ibraheem, Novel approach for the automated detection of allergy test using spectral imaging, J. Biomedical Science and Engineering, 2012, 5, 416-421

[3] R. Leitner, I. Ibraheem, A. Kercek, Proc. 2nd International Workshop on Spectral Imaging (2005)

[4] I. Ibraheem Linear and quadratic classifier to detection of skin lesions "epicutaneus". Fifth International Conference on Bio-informatics and Biomedical Engineering, Wuhan, 31 May 2011, 1-5.

[5] C. M. Bishop, Pattern Recognition and Machine Learning, Springer; Auflage: first ed. 2006. Corr. 2nd printing 2011 (2007)

[6] R.O.Duda, P.E.Hart and D.G. Strock, Pattern Classification, John Wiley & Sons; Auflage: 2. Auflage (21. November 2000)

[7] T.M. Lillesand, R.W. Kiefer and J.W. Chipman, Remote Sensing and Image Interpretation, John Wiley & Sons, Hoboken, NJ, USA, 2004.

[8] R. Bhargava, I. Levin (Eds.) Spectrochemical Analysis Using Infrared Multichannel Detectors, Blackwell Publishing (2005)

[9] E. Fix and J.L. Hodges, (1989) Discriminatory analysis, non-parametric discrimination: Consistency proper-ties. International Statistical Review, 57, 238-247.

[10] N. Eisenreich, T. Rohe in Encyclopedia of Analytical Chemistry, 7623 – 7644, Wiley & Sons (2000)

[11] Zenzo, S.D., R. Bernstein, S.D. Degloria and H.C. Kolsky (1987b), "Gaussian maximum likelihood and contextua1 classification algorithms for multicrop

[12] Du, Q. (2000), Topics in Hyperspectral Image Analysis, Department of Computer Science and Electrical Engineering, University of Matyland, Baltimore County,MD, May 2000.

[13] Du, Q. and C.-1 Chang (1998), "Radial basis function neural networks approach tohyperspectral image classification," 1998 Conference an Information Science andSystems, Princeton University, Princeton, NJ, pp. 721-726, March 1998.

[14] Du, Q. and C.-I Chang (1999), "An interference rejection-based radial basis function neural network approach to hyperspectral image classification," International JointConference on Neural Network, Washington DC, pp. 2698-2703, July 1999.

[15] Du, Q. and C.-1 Chang (2000), "A hidden Markov model-based spectral measure for hyperspectral image analysis," SPIE Conf Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VI, Orlando, FL, pp. 375-385, April 2000.

[16] Du, Q. and C.-1 Chang (2001a), "A linear constrained distance-based discriminantanalysis for hyperspectral image classification," Pattern Recognition, vol. 34, no. 2, 2001.

[17] Du, Q. and C.-1 Chang (2001b), "An interference subspace projection approach tosubpixel target detection," SPIE Conf an Algorithms for Multispectral,Hyperspectral and Ultraspectral Imagery VII, Orlando, Florida, pp. 570-577, 20- 24 April, 200 I.

[18] Du, Q. and H. Ren (2002), "On relationship between OSP and CEM," SPIE Conf. anAlgorithms for Multispectral, Hyper-spectral and Ultraspectral Imagery VIII,Orlando, Florida, 20-24 April, 2002.

[19] Du, Q., C.-1 Chang, D.C. Heinz, M. L.G. Althause and I.W. Ginsberg (2000), "Hyperspectral image compression for target detection and classification," IEEE 2000 International Geos-cience and Remote Sensing Symp., Hawaii, USA, July 24- 28, 2000.

[20] Fano, R.M. (1961), Transmission of Information: A Statistical Theory of Communication, John Wiley & Sons, N.Y., 1961.

[21] Farrand, W., and J.C. Harsanyi (1997), "Mapping the distribution of mine tailing in the coeur d'Alene river valley, Idaho, through the use of constrained energy minimization technique," Remote Sensing of Environment, vol. 59, pp. 64-76,1997.

[22] Friedman, J.H. and J.W. Tukey (1974), "A projection pursuit algorithm for exploratotydata analysis," IEEE Transactions an Computers, vol. c-23, no. 9, pp. 881-889,1974.

[23] Friedman, J.H. (1987) "Exploratoty projection pursuit," Journal of American Statistical Association, 82, pp. 249-266, 1987.

[24] Frost Ill, O.L. (1972), "An algorithm for linearly constrained adaptive array processing," Proc. IEEE, vol. 60, pp. 926-935, 1972.

[25] Fukunaga, K. (1982), "Intrinsic Dimensionality Extraction", Classification, Pattern Recognition and Reduction of Dimen-sionality, Handbock of Statistics, vol. 2, P .R.

[26] Krishnaiah and L.N. Kanal eds., Amsterdam: North-Holland Publishing Company,1982, pp. 347-360.

[27] Fukunaga, K (1992), Statistical Pattern Recognition, 2nd ed., New York: Academic Press, 1992.

[28] Pa!, N.R. and S. K. Pa! (1989), "Entropie thresholding," Signal Processing, Vol. 16,pp. 97-108, 1989.

[29] Poor, H.V. (1994), An Introduction to Detection and Estimation Theory, 2nd. ed., NewYork: Springer-Verlag, pp. 58-59, 1994.

[30] Rabiner, L. and B.-H. Juang (1993), Fundamentals ofSpeech Recognition, Prentice-Hall,1993.

[31] Reed, I.S. and X. Yu (1990), "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribu-tion," IEEE Trans. on Acoustic, Speech andSignal Process., vol. 38, no. 10, pp. 1760-1770, Oct. 1990.

[32] Resmini, R.S., M.E. Kappus, W.S. Aldrich, J.C. Harsanyi and M. Anderson (1997), "Mineral mapping with HYperspectral Digital lmagery Collection Experiment (HYDICE) sensor data at Cuprite, Nevada, U.S.A.," Int. J. Remote Sensing, vol.18, no. 17, pp. 1553-1570, 1997.

[33] Ren, H. (1998), A Comparative Study of Mixed Pixel Classi-fication Versus Pure Pixel Classification for Mul-ti/Hyperspectral Imagery, Department.of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, MD, May 1998.

[34] Ren, H. (2000), Unsupervised and Generalized Orthogonal Subspace Projection andConstrained Energy Minimization for Target Detection and Classification inRemotely Sensed Imagery, Department of Computer Science and ElectricalEngineering, University ofMaryland, Baltimore County, MD, May 2000.

[35] Ren, H. and C.-1 Chang (1998), "A computer-aided detection and classification methodfor concealed targets in hyperspectral imagery," IEEE 1998 International Geoscience and Remote Sensing Symposium, Seattle, WA, pp. 1016-1018, July 5- 10, 1998.

[36] Ren, H. and C.-I Chang (1999), "A constrained least squares approach to hyperspectral image classification," 1999 Conference on Information Science and Systems, pp. 551-556, Johns Hopkins University, Baltimore, MD, March 17-19, 1999.

[37] Ren, H. and C.-I Chang (2000a), "A generalized orthogonal subspace projection approach to unsupervised multispectral image classification," IEEE Trans. on Geoscience and Remote Sensing, vol. 38, no. 6, pp. 2515-2528, November 2000.

[38] 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.

[39] Richards, J.A. (1993), Remote Sensing Digital Image Analysis, 2nd ed. Springer Verlag, 1993.

[40] Rissanen, J. (1978), "Modeling by shortest data description," Automatica, vol. 14, pp. 465-471, 1978.

[41] Roger, R.E. (1996), "Principal components transform with simple, automatic noise adjustrnent," International Journal Remote Sensing, vol. 17, no. 14, pp. 2719- 40 2727, 1996.

[42] Roger, R.E. and J.F. Amold (1996), "Re1iably estimating the noise in AVIRIShyperspectral imagers," *Int. J. Remote Sensing,* vol. 17, no. 10, 1951-1962, 1996.

[43] Sabol, D.E., J.B. Adams and M.O. Smith (1992), "Quantitative sub-pixel spectral detection oftargets in multispectral images," *J. Geophys. Research,* 97, pp. 2659-2672, 1992.

[44] Sahoo, P.K., S. Soltani, A.K.C. Wong and Y.C. Chen (1988), "A survey ofthresholding techniques," *Computer Vision, Graphics and Image Process.(CVGIP),* vol. 41, pp. 233-260, 1988.

[45] Schalkoff, R. (1992), *Pattern Recognition: Statistical, Structure and Neural Network,*New York: John Wiley and Sons, 1992.

[46] Scharf, L.L. (1991), *Statistical Signal Processing,* MA: Addison-Wesley, 1991.

[47] Schowengerdt, R.A. (1997), *Remote Sensing: Models and Methods for ImageProcessing,* 2nd ed., Academic Press, 1997.

[48] Schwarz, G. (1978), "Estimating the dimension of a model," *Ann. Stat.,* vol. 6, pp. 461- 464, 1978.

[49] Settle, J.J. (1996), "On the relationship between spectral un-

mixing and subspace projection," *IEEE Trans on Geoscience and Remote Sensing,* vol. 34, no. 4, pp. 1045-1046, July 1996.

[50] 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.

[51] Shahshahani, B.M. and D.A. Landgrebe (1994), "The effect of unlabeled samples in reducing the small sample size problem and mitigating the Hugh phenomenon,"*IEEE Trans. Geoscience and Remote Sensing,* vol. 32, no. 5, pp. 1087-1095,September 1994.

[52] Shimabukuro, Y.E. (1987), *Shade Images Derived from Linear Mixing Models of Multispectral Measurements of Forested Areas,* Ph.D. dissertation, Department of Forestand Wood Science, Colorado State University, Fort Collins, 1987.

[53] Shimabukuro, Y.E. and J.A. Smith (1991), "The least-squares mixing models to generate fraction images derived from remote sensing multispectral data," *IEEE Trans on Geoscience and Remote Sensing,* vol. 29, pp. 16-20, 1991.

[54] Singer, R.B. and T.B. McCord (1979), "Mars: !arge scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance," *Proc. Lunar Planet. Sei. Conf 10th,* pp. 1835-1848, 1979.

[55] Smith, M.O., J.B. Adams and D.E. Sabol (1994), "Spectral mixture analysis-new strategies for the analysis of multispectral data," *Image Spectroscopy-a tool for Environmental Observations* edited by J. Hili and J. Mergier, Brussels and Luxembourg, pp. 125-143, 1994.

[56] Smith, M.O., D.A. Roberts, J. Hili, W. Mehl, B. Hosgood, J. Verdebout, G. Schmuck, C. Koechler and J.B. Adams (1994), "A new approach to quantifying abundances of materials in multispectral images," *Proc. IEEE Int. Geoscience and Remote Sensing Symposium'94,* Pasadena, CA, pp. 2372-2374, 1994.

[57] Soltanian-Zadeh,H., J.P. Windham, D.J. Peck (1996), "Optimal linear transformation for MRl feature extraction," *IEEE Trans. on Medical Imaging,* vol. 15, pp. 749- 767, 1996.

[58] Stark, H. and J. Woods (1994), *Probability, Random Processes, and Estimation for Engineers,* 3rd ed., Prentice-Hall, 2002.

[59] Steams, S.D., B.E. Wilson and J.R. Peterson (1993), "Dimensionality reduction by optimal band selection for pixel classification of hyperspectral imagery," *Applications of Digital Image Processing XVI,* SPIE, vol. 2028, pp. 118-127,

[60] Zhao, X. (1996), Subspace Projection Approach to Multispectral!Hyperspectral Image Classification Using Linear Mixture Modeling, Master Thesis, Departrnent of Computer Seiences and Electrical Engineering, University of Maryland Baitimare County, MD, May 1996.