

Feasibility Study of Optical Spectroscopy as a Medical Tool for Diagnosis of Skin Lesions

Asad Safi and Sheikh Ziauddin
COMSATS Institute of Information
Technology Islamabad, Pakistan

Alexander Horsch
Department of Medical Statistics
and Epidemiology, Technische Universität
München, Germany

Mahzad Ziai
Klinik und Poliklinik für Dermatologie
und Allergologie, Technische Universität
München, Germany

Victor Castaneda
Universidad de Chile Independencia 1027
Santiago, Chile

Tobias Lasser and Nassir Navab
Chair for Computer Aided Medical Procedures (CAMP),
Technische Universität München, Germany

Abstract—Skin cancer is one of the most frequently encountered types of cancer in the Western world. According to the Skin Cancer Foundation Statistics, one in every five Americans develops skin cancer during his/her lifetime. Today, the incurability of advanced cutaneous melanoma raises the importance of its early detection. Since the differentiation of early melanoma from other pigmented skin lesions is not a trivial task, even for experienced dermatologists, computer aided diagnosis could become an important tool for reducing the mortality rate of this highly malignant cancer type.

In this paper, a computer aided diagnosis system based on machine learning is proposed in order to support the clinical use of optical spectroscopy for skin lesions quantification and classification. The focus is on a feasibility study of optical spectroscopy as a medical tool for diagnosis. To this end, data acquisition protocols for optical spectroscopy are defined and detailed analysis of feature vectors is performed. Different techniques for supervised and unsupervised learning are explored on clinical data, collected from patients with malignant and benign skin lesions.

Keywords—Melanoma; Classification; Supervised Learning; Computer-Aided Diagnosis; Machine Learning; Optical Spectroscopy

I. INTRODUCTION

Skin cancer is among the most frequent types of cancer and one of the most malignant tumors. The incidence of melanoma in the general population is increasing worldwide [1], especially in countries where the ozone layer is thinning. Its incidence has increased faster than that of almost all other cancers, and the annual rates have increased by 3% to 7% in the fair-skinned population in recent decades [1]. Currently, between 2 and 3 million non-melanoma skin cancers and 132,000 melanoma skin cancers occur globally each year [2].

New technologies to assist the dermatologists in identifying and diagnosing skin lesion have been introduced, such as hand-held magnification devices and computer-aided image analysis. Colored image processing methods have been introduced for detecting the melanoma [3] which focused on non-constant visual information of skin lesions. Neural network diagnosis of skin lesion has been applied by classifying extracted features from digitized dermoscopy images of lesions [4] [5].

The extracted features are based on geometry, colors, and texture of the lesions, involving complex image processing techniques. Many other attempts have been made to automate the detection and classification of melanoma from the digital color and surface reflectance images [6][7][8][9][10]. Those attempts involve the initial segmentation of the skin lesion from the surrounding skin followed by the calculation of classification features [11][12][5][13][14][15]. Accurate description and measurement of image features cannot be achieved without accurate image segmentation. Therefore, a wide range of algorithms have been proposed in the past for color image segmentation [16], broadly categorized as pixel-based segmentation, region-based segmentation [17], region-based segmentation and edge detection [18]. However, in the case of optical spectral reflectance images, the research is still limited due to the late introduction of the imaging technology in dermatology.

A. Why Optical Spectroscopy?

One of the substantial features for the diagnosis of malignant melanoma is the skin lesion color [19]. In most of the related research, skin lesion color was investigated to disintegrate malignant melanoma lesions from benign lesions in clinical images [20]. Human skin is a variegated surface, with fine scale geometry, which makes its appearance difficult to model. Furthermore, the conditions under which the skin surface is viewed and illuminated greatly affect its appearance.

As we know that light of different wavelengths access the skin in different depths (as shown in Figure 1). This fact led the researchers to evaluate pigmented lesions under specific wavelengths of light from visible spectrum to near infrared range. Through multi-spectral imaging we can capture light from frequencies beyond the visible light range which allows us to extract additional information that the human eye fails to capture with its receptors for red, green and blue. Furthermore, the spectral information can be employed for the analysis and the information retrieval about the consistence and the concentration of absorbers and reflectors in the skin. Different pigments of the skin absorb different wavelengths of optical spectrum, which helps in determining the reflectance coefficient of the area of the skin.

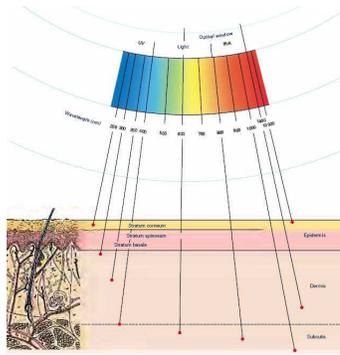


Fig. 1. Different wavelengths penetrate the skin to different depths. Visible light and near infrared penetration in skin is more than other wavelengths (Image source: [21])

One of the most significant features of spectral reflectance is the property that the spectral reflectance curve is based on the material composition of the object's surface, color, biochemical composition and cellular structure. This property can be utilized for recognizing objects and segment regions. Currently there exist only a small number of systems, e.g. spectrophotometric intracutaneous analysis (SIA) scope [22], Melafind [23] and SpectroShade [24], which use multispectral dermoscopic images as the inputs for subsequent computer analysis. To the best of our knowledge, the systems which have already been developed for the analysis of skin lesion from multispectral images, are based on the images of selected wavelength without keeping record of reflectance spectra. However, as different skin lesions can be investigated more in detail by observing their reflectance, we analyze the feasibility of spectroscopy as a tool to distinguish benign and malign skin lesions.

B. Introduction of Spectroscopy

Spectroscopy is a new imaging technology which is increasingly used to derive significant information about tissue. Due to its multi-spectral nature, this imaging method allows to detect and classify multiple physiological changes like those associated with increased vasculature, cellular structure, oxygen consumption or edema in tumors [25]. The hardware setup for data acquisition is explained in more detail in section III.

Optical spectra in different wavelengths and amplitude is shown in Figure 2 which shows the differences between four colors (red, green, yellow and blue). The experiment is performed on the phantom, where chalks colored with four different inks are used in the experiment. Figure 2 clearly shows that variation in color produces difference in optical spectroscopy.

We design an experiment, to observe the difference between objects based on internal structure. In this experiment we gather six different fruits (Apple, Blueberry, Kiwi, Strawberry, Plum and Orange). Data was collected from each fruit after 12 hours for 7 days consecutively. Due to the change in the internal structure of the fruits the curve was changed, but the main shape of the curve was always constant. Standard deviation of each fruit in wavelength and amplitude is shown in Figure 3.

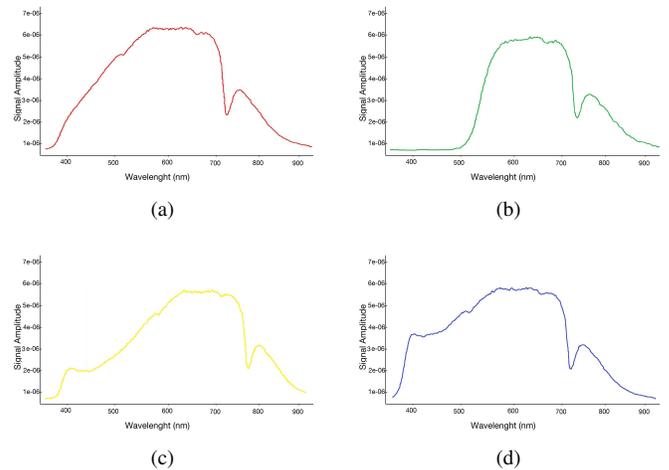


Fig. 2. Optical spectra from chalk with color inks in wavelength and amplitude. (a) Red color, (b) Green color, (c) Yellow color, (d) Blue color

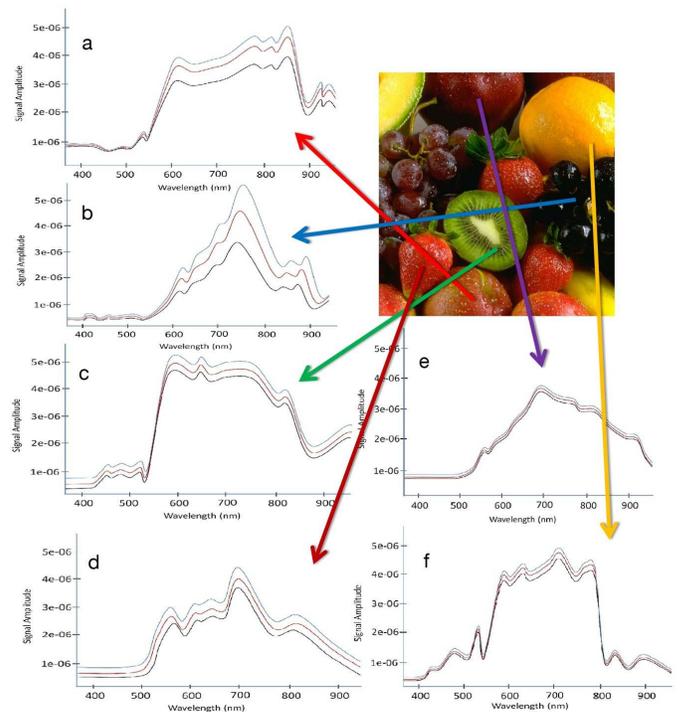


Fig. 3. Spectral standard deviation of each fruit in wavelength and amplitude. Lower curve (Black colored) is the minimum, the upper curve (Blue colored) represents the maximum and the middle curve (Red colored) represents the mean. (a) Apple, (b) Blueberry, (c) Kiwi, (d) Strawberry, (e) Plum, (f) Orange

II. STATE OF THE ART

Skin color measurement through reflectance spectroscopy has received significant attention in the literature [19][26][27][28]. It has been used to provide a numerical index for color, which in turn allows for the study of constriction of a blood vessel and abnormal redness of the skin due to local congestion, such as in inflammation [29]. Dawson et al. [30] worked on the reflectance spectroscopy for the measurement of skin tissue to exemplify the spectral properties. Farrell et al. [31] and Kienle et al. [32] addressed the problem of reflectance measurements to determining in

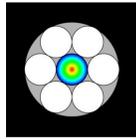


Fig. 4. Schematic representation of the fiber arrangement in the spectroscopy probe: $6 \times 200\mu\text{m}$ illumination fibers arrayed around one $600\mu\text{m}$ acquisition fiber.

vivo tissue optical properties. Another approach for measuring the optical reflectance over a broad range of wavelengths spectroscopy has been utilized for assessing the skin type and gestation age of newborn infants by Lynn et al. [33].

The first work to evaluate the possibilities of using reflectance spectrophotometry for discriminating between benign and malignant skin lesions was done by Marchesini et al. [34]. Their experiments show that the wavelengths between 400 and 800 nm were highly significant to show the differences between the reflectance spectra of benign and malignant melanomas. Consequently, the authors report a discrimination between 31 primary melanoma and 31 benign lesions with a sensitivity of 90.3% and a specificity of 77.4%, a stepwise discriminate analysis of reflectance spectral features [35].

Moreover, Bono et al. [36] conclude that color is the most important parameter in discriminating melanomas from benign in spectrophotometric imaging of skin lesions using 420-1020 nm . Recently with Raman spectroscopy the molecular structure of skin lesions are explored [37], but due to the side effects of the laser beam on the sensitive skin surface, this technique is not preferred in the dermatology practice.

III. SYSTEM SETUP

A hand-held reflectance spectroscopy probe (StellarNet Inc., Oldsmar, FL, USA) (see Figure 5), consisting of $6 \times 200\mu\text{m}$ illumination fibers arrayed around one $600\mu\text{m}$ acquisition fiber as shown in Figure 4, was attached to an infrared optical tracking target in order to be able to determine its position and orientation in real-time. The selected tracking system consists of four ARTtrack2 infrared cameras (A.R.T. GmbH, Weilheim, Germany) positioned to be able to track a volume of $2 \times 2 \times 2\text{m}^3$. According to the manufacturer the positional accuracy for such a configuration is 0.4 mm with a maximum error of 1.4 mm (for angle 0.002 rad and 0.007 rad respectively).

A 178 – 1132 nm , 2048 px , 12 bit CCD spectrometer (StellarNet Inc., Oldsmar, FL, USA) was connected to the acquisition fiber, and a 12 W tungsten lamp was connected to the illumination fibers as a light source. The spectrometer was controlled by a data processing unit to acquire spectra synchronously with the tracking information of the probe. The data-processing unit was also used to run the augmented reality application that combined spectra, positions and orientations. An overview of the entire setup is displayed in Figure 5.

A. Data Acquisition Protocol

In our protocol, the the mole selection for the data acquisition is purely based on the doctor's (or physician's) choice based on a visual examination. The labeling of mole is

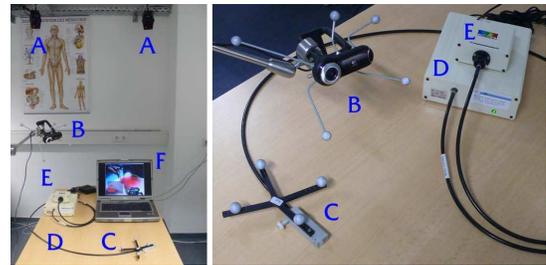


Fig. 5. System setup: (a) tracking cameras, (b) augmented camera, (c) tracked probe, (d) spectrometer, (e) light source, and (f) data-processing unit.

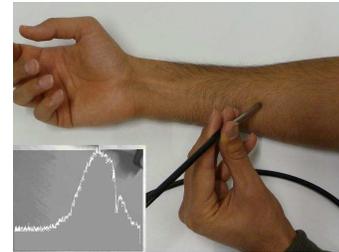


Fig. 6. Covering all the surface of prob tip by contacting skin surface.

performed using two classes: suspicious skin lesion (possibility of malignant melanoma) and normal skin moles based on physician's diagnosis.

The data is stored as a plot of wavelength and amplitude (as shown in Figure 2) by spectrometer without taking into account the mole structure. The time of data acquisition and the number of measurements depend on the number of moles defined on patients, where the time for whole body skin checkup was approximately 20 minutes.

The spatial resolution of sampling region is 1 mm diameter which permits the study of smaller lesions and sampling of several regions within bigger lesions. For mole size bigger than 3 mm and smaller then 6 mm we take 5 measurements (4 from the edges, 1 from the center). If the mole sizes exceeds 6 mm then we take 7 measurements (6 from the edges, 1 from the center). To make sure that the database is consistent and not biased, we only use the measurements which were taken once per mole.

The data acquisition time for one mole is 100 ms . It is important to contact the surface of the mole by the probe tip and keep the probe in a way that no light goes in from outside to ensure that the spectra are only obtained from the lesion itself, as shown in Figure 6. Hair, nails and tattoos are avoided during data acquisition.

B. Data Acquisition

The data collection for this study was performed in collaboration with the dermatology department at Klinikum Rechts der Isar Munchen; Germany. All lesions in this analysis were selected by dermatology experts. In total, 3072 spectroscopic data vectors were collected from 148 patients, where 2926 measurements were of normal skin moles and 146 measurements from malignant skin lesions. The schematic representation of data acquisition system is shown in Figure 7. Out of 146 malignant skin lesions, 9 cases were histological proven

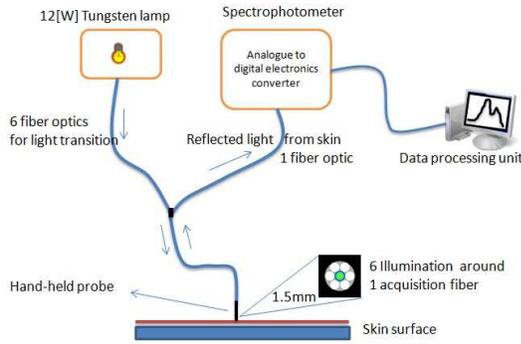


Fig. 7. Schematic diagram of data acquisition system.

melanoma. The remaining 137 are kept under observation. The details of the 9 cases of melanoma are: average Breslow thickness was 1.1 mm, the minimum being 0.1 mm and the maximum 2.8 mm, the average diameter of the lesions was 3 mm, the minimum being 2 mm and the maximum 5 mm. The average age of patients was 40, where the youngest and oldest patients were 2 and 82 years old, respectively. 70% of the examined patients were female. The collected data consists of the following clinical cases:

- Normal skin: spectra were obtained from the inside of the upper arm, groin and inside thigh, a region defined as skin that is not normally exposed to sunlight (i.e. not tanned).
- Normal skin moles: in average 19 spectra per patient were obtained from benign skin moles. Normal skin moles can be visually very similar to malignant moles, as illustrated in Figure 8.
- Malignant skin mole: one spectra was obtained from middle positions on the lesion. Multiple spectra were taken depending on size of the mole as discussed in data acquisition protocol (Section III-A).

Immediately prior to each patient's data collection session, the spectrophotometer probe end was placed in the disinfectant substance to prevent migration of any diseases.

To make sure of reproducibility and accuracy of data acquisition, one concern was that the pressure of the probe on the skin might cause blanching by forcing blood out of local vessels. To test a novel approach to reduce this effect and to assess the magnitude of this problem, a study was performed by Osawa et al. [38]. In their study the probe was held in contact with a flat area of skin and the pressure slowly increased beyond that which would be applied normally for taking skin reflectance measurements. Increasing the pressure caused a decrease in overall reflectance. Osawa et al. suggested three methods for eliminating the effect: (a) a sensor to determine the pressure being applied, (b) an adhesive pad to just hold the probe against the skin, and (c) an electrical contact sensor to feed back information on when the probe makes contact with the skin. In our study the pressure on the skin was reduced by increasing the surface area of contact with a probe holder that was designed to slide in the probe which was also used to keep the tracking points (see Figure 5).

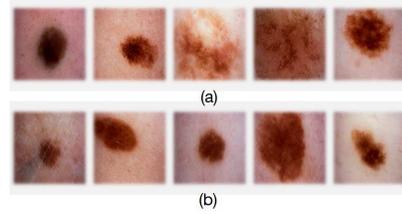


Fig. 8. Skin lesions: (a) Malignant skin lesions, (b) Normal skin lesions.

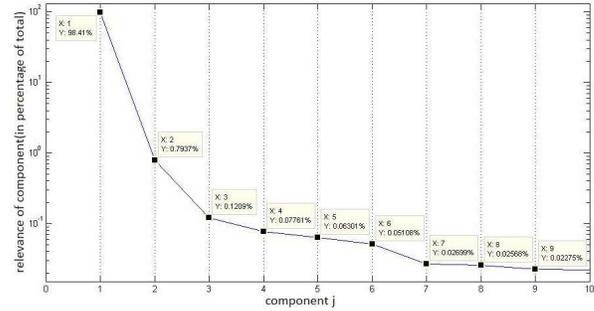


Fig. 9. Representative example of the first part of the sorted PCA eigenvalue spectrum (e_j^i), the y -axis shows the values of the component as a percentage of the total in log scale.

C. Data Processing

The spectral data is acquired as a $2048D$ vector of the floating points values $x_i \in R^{2048}$, $i = 1, \dots, n$ where n denotes the number of measurements. Each x_i represents the discretized reflective spectrum from 178 nm to 1132 nm (due to limitation of hardware) of the i th measurement and is stored normalized as

$$\hat{x}_i = \frac{x_i}{\|x_i\|_2} \text{ where } i = 1, \dots, n \quad (1)$$

To reduce the dimensions of the input data, principal components analysis (PCA) is applied. The resulting spectrum of eigenvalues $(e_j^i)_{j=1, \dots, 2048}$ is sorted descending by magnitude. Since the highest eigenvalues represent the most relevant components, a cut-off value C_{PCA} is chosen, such that the final input data y_i for the classification algorithm from measurement x_i ($i = 1, \dots, n$) is

$$y_i = (e_j^i)_{j=1, \dots, C_{PCA}} \quad (2)$$

The cut-off value C_{PCA} is chosen empirically from the data. Figure 9 is showing a representative example of $(e_j^i)_{j=1, \dots, 2048}$ from which C_{PCA} was selected as one of $\{2, 3, 4, 5\}$.

IV. CLASSIFICATION

Classification is performed by a support vector machine (SVM) [39]. SVM was selected as the method of choice as it allows to linearly classify data in a high-dimensional feature space that is non-linearly related to the input space via the use of specific kernel functions, such as polynomial functions or

radial basis functions (RBF). This way we can build complex enough models for skin lesion classification while still being able to compute directly in the input space.

The SVM classifier needs to be trained first before using it, thus we partition our already reduced input data $(y_i), i = 1, \dots, n$ into two partitions, $T \subset \{1, \dots, n\}$ the training set and $V \subset \{1, \dots, n\}$ the testing (or validation) set with $T \cup V = \{1, \dots, n\}$ and $T \cap V = \{\}$. The training data set T is labeled manually into two classes with the ground truth, $l(y_i) = \pm 1$. Once the classifier is trained, a simple evaluation of the decision function $d(y_i) = \pm 1$ will yield the classification of any data y_i .

In detail, SVM is trying to separate the data $\phi(y_i)$ mapped by the selected kernel function ϕ by a hyperplane $w^T \phi(y_i) + b = 0$ with w the normal vector and b the translation. The decision function then is $d(y_i) = \text{sgm}(w^T \phi(y_i) + b)$. Maximizing the margin and introducing slack variables $\xi = (\xi_i)$ for non-separable data, we receive the primal optimization problem:

$$\min_{w,b,\xi} = \frac{1}{2} w^T w + C \sum_{i \in T} \xi_i \quad (3)$$

with constraints $l(y_i)(w^T \phi(y_i) + b) \geq 1 - \xi_i, \xi_i \geq 0$ for $i \in T$. C is a user-determined penalty parameter. Switching to the dual optimization problem allows for easier computation,

$$\min_{\alpha} = \frac{1}{2} \alpha^T Q \alpha - e^T \alpha \quad (4)$$

with constraints $0 \leq \alpha_i \leq C$ for $i \in T, \sum_{i \in T} y_i \alpha_i = 0$. The $\alpha = (\alpha_i)$ are the so-called support vectors, $e = [1, \dots, 1]^T$ and Q is the positive semidefinite matrix formed by $Q_{jk} = l(y_j)l(y_k)K(y_j, y_k)$, and $K(y_j, y_k) = \phi(y_j)^T \phi(y_k)$ is the kernel function built from ϕ . Once this optimization problem is solved, we determine the hyperplane parameters w and b , w directly as $w = \sum_{i \in T} \alpha_i l(y_i) \phi(y_i)$ and b via one of the Karush-Kuhn-Tucker conditions as $b = -l(y_i) y_i^T w$, for those i with $0 < \alpha_i < C$. Thus the decision function of the trained SVM classifier ends up as

$$\begin{aligned} d(y_i) &= \text{sgn}(w^T \phi(y_i) + b) \\ &= \text{sgn} \left(\sum_{j \in T} \alpha_j l(y_j) K(y_j, y_i) + b \right) \end{aligned} \quad (5)$$

V. EXPERIMENTS

Data collection of 3072 spectroscopic instances is defined as $(x_i), i = 1, \dots, 3072$ labeled into two classes: normal skin $l(x_i) = 1$ and lesion $l(x_i) = -1$. The 3072 data points were randomly separated into a training data set T and a testing (validation) data set V with $|T| = 2072$ and $|V| = 1000$, however retaining the balance of both sets containing 50% from each of the two classes. A color-coded representation of the normalized skin spectra $\hat{x}_i, i \in T$ of the training data set T is shown in Figure 10.

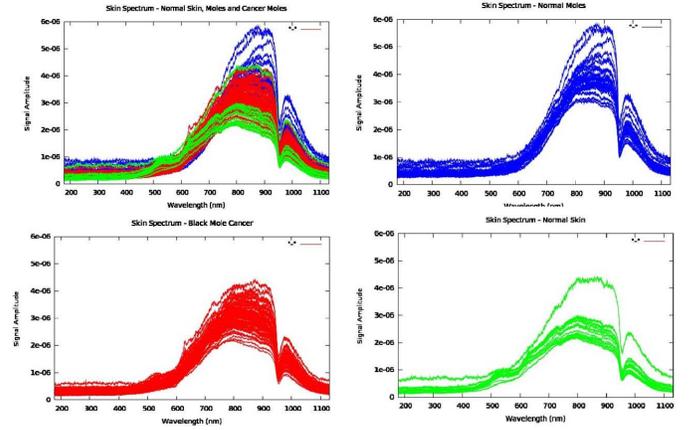


Fig. 10. Plot of all normalized spectra \hat{x}_i from the training data set T , color-coded as blue for normal skin moles, red for malignant mole and green for normal skin. One curve represents one skin lesion data.

TABLE I. RESULTS OF THE CROSS-VALIDATION USING THE TRAINING DATASET T .

Parameters	SVM Training			
	Linear Kernel	Poly Kernel	RBF Kernel	Sigmoid Kernel
$C_{PCA} = 2$	95 ± 9.2	96 ± 8.3	95 ± 7.5	95 ± 9.1
$C_{PCA} = 3$	95 ± 8.3	96 ± 6.7	97 ± 9.5	96 ± 9.5
$C_{PCA} = 4$	95 ± 9.6	97 ± 7.2	97 ± 8.7	96 ± 8.6
$C_{PCA} = 5$	96 ± 9.2	97 ± 9.7	97 ± 8.3	97 ± 7.7

Before classification, PCA was applied to the \hat{x}_i for dimension reduction to yield our classification input y_i . The eigenvalue cut-off C_{PCA} was empirically chosen as one of $C_{PCA} \in \{2, 3, 4, 5\}$.

The SVM classifier (we used LibSVM, [40]) was then trained using the training data set T . As there are multiple parameters to be selected, like for example the penalty parameter C , we performed a cross-validation of 10 folds via parallel grid search. The average accuracy on the prediction of the validation fold is the cross validation accuracy.

A. Discussion

The cross-validation of the training data set T determined, among others, the parameters $C = -5$ and $\gamma = -7$. For the further parameters C_{PCA} and the choice of the kernel (linear, polynomial, radial basis function (RBF) or sigmoid) we performed cross validation of the training data set T , the results are shown in Table I. The best results were received consistently by using the RBF kernel, while for C_{PCA} the value of 5 turned out to be the best choice with an accuracy of 97 ± 8.3 , where 8.3 is standard deviation.

With the training of the classifier completed, we studied the accuracy of the testing (validation) data set V . We compared the manual ground truth labeling $l(y_i)$ for data point y_i with the computed decision function $d(y_i)$ to compute the accuracy as follows

$$\begin{aligned} \text{Accuracy}(\%) &= \frac{\# \text{ of correctly predicted data}}{\# \text{ of total data}} \times 100 \\ &= \frac{|d(y_i)|}{|V|} \times 100 \end{aligned} \quad (6)$$

TABLE II. CLASSIFICATION ACCURACY RESULTS USING THE TESTING DATASET V.

Parameters	SVM Training			
	Linear Kernel	Poly Kernel	RBF Kernel	Sigmoid Kernel
$C_{PCA} = 2$	86.8%	90.3%	89.9%	88.8%
$C_{PCA} = 3$	89.3%	92.5%	91.8%	90.3%
$C_{PCA} = 4$	91.9%	92.9%	94.9%	94.1%
$C_{PCA} = 5$	92.1%	93.6%	94.9%	94.6%

The results are shown in Table II. We achieve the same accuracy of 94.9% for the kernels RBF with C_{PCA} values of 4 and 5. This corresponds to Figure 9, where it is clear that between C_{PCA} 4 and 5 there is only very little difference. In total we received the best results using the RBF kernel and $C_{PCA} = 5$.

VI. MANIFOLD LEARNING FOR DIMENSIONALITY REDUCTION OF SKIN LESIONS USING OPTICAL SPECTROSCOPY DATA

Most recent applications of machine learning in data mining, computer vision, and in other fields require deriving a classifier or function estimate from a large data set. Modern data sets often consist of a large number of examples, each of which is made up of many features. Though access to an abundance of examples is purely beneficial to an algorithm attempting to generalize from the data, managing a large number of features (some of which may be irrelevant or even misleading) is typically a burden to the algorithm. Overwhelmingly complex feature sets will slow the algorithm down and make finding global optima difficult. To lessen this burden on standard machine learning algorithms (e.g. classifiers, function estimators), a number of techniques have been developed to vastly reduce the quantity of features in a dataset, i.e. to reduce the dimensionality of the data.

Dimensionality reduction has other, related uses in addition to simplifying data so that it can be efficiently processed. The most obvious is visualization; if data lies, for instance, in a 100-dimensional space, one cannot get an intuitive feel for what the data looks like. However, if a meaningful two or three dimensional representations of the data can be found, then it is possible to analyze it more easily. Though this may seem like a trivial point, many statistical and machine learning algorithms have very poor optimality guarantees, so the ability to see the data and the output of an algorithm is of great practical interest. In our case, spectroscopic data is typically acquired as a high dimensional vector (in our case a 2048 element vector); this high-dimensionality, however, creates difficulties for visualization and classification of the data. Manifold learning has a significant role in dimensionality reduction and clustering due to its nature of unsupervised learning [41].

There are many approaches to dimensionality reduction based on a variety of assumptions and used in a variety of contexts. We will focus on an approach initiated recently based on the observation that high-dimensional data is often much simpler than the dimensionality would indicate. In this work, we present results of applying different manifold learning techniques such as Isomap [42], Laplacian Eigenmaps [43] and Diffusion Map [44] to spectroscopy data from 48 patients with normal and malignant lesions to reduce the dimensionality,

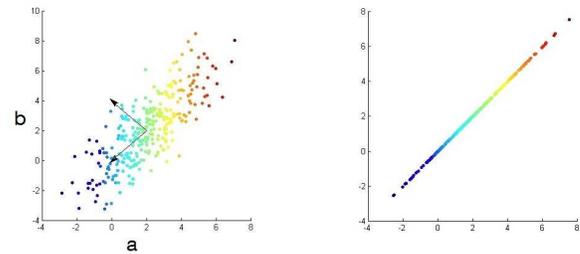


Fig. 11. Working example of PCA. The left image shows a Gaussian distribution together with the two principal components. The coloring is dependent on values of a and b . The right side shows the projection on the eigenvector corresponding to the largest eigenvalue [46].

and compare them to traditional linear technique Principal Component Analysis. Clustering results after dimensionality reduction are shown in Table III for each technique, where some of the method/parameter combinations yield excellent results on the patient data compared to the diagnosis of the treating physicians.

A. Principal Component Analysis

A linear method such as PCA ignores protrusion or concavity of the data [45]. In order to demonstrate the shortcomings of purely linear methods, we will show results using PCA and compare with nonlinear manifold learning. PCA finds a subspace i.e. which finds an optimal subspace that best preserves the variance of the data [46].

The goal of PCA is to find an optimal subspace i.e. the variance of the data is maximized. In general, manifold learning methods do not care about the variance of the data. Non-linear methods in particular, typically famous on preserving neighborhood properties within the data [46]. The input and output of PCA are defined as in equation 7, given N input points.

Figure 11 shows a Gaussian distribution together with the first (and only) two principal components, calculated by the method described above. The vectors are therefore the eigenvectors of the matrix C .

The coloring is linearly dependent on the values of a and b . The right side shows the projection on the eigenvector corresponding to the largest eigenvalue. As one can see, the variance of the data is preserved.

Figure 12 shows that PCA cannot handle non-linear datasets. The left image shows a spiral distribution (2-d Swiss roll) together with the two principal components. The coloring is dependent on the values of t , where the function is given as $f(t) = (t \cos(t), t \sin(t))$. The right side of Figure 12 shows the overlapping projection on the eigenvector corresponding to the largest eigenvalue. One can observe that blue, red and yellow points are all overlapping in the center of the projected line [46].

This means that most geometric information of the data is lost through this projection. In most cases distances are only meaningful in local neighborhoods, following Non-linear manifold learning methods address this problem.

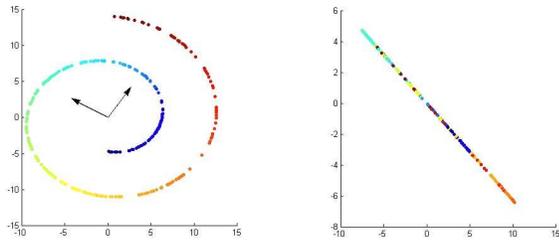


Fig. 12. PCA cannot handle non-linear datasets. The left image shows a spiral distribution (2-d Swiss roll) together with the two principal components. The coloring is dependent on the values of t , where the function is given as $f(t) = (t\cos(t), t\sin(t))$. The right side shows the overlapping projection on the eigenvector corresponding to the largest eigenvalue [46].

B. Non-linear Manifold Learning Methods

Typical non-linear manifold learning methods are graph-based and perform the following three basic steps.

- 1) Build undirected similarity graph $G = (V, E)$, where the vertices V are give by the data points x_i
- 2) Estimate local properties, i.e. the weight matrix W to define the weighted similarity graph $G = (V, E, W)$, where $w_{ij} \geq 0$ represents the weight for the edge between vertex i and j . Weights are obtained by means of a kernel. A weight of 0 means that the vertices are not connected.
- 3) Derive an optimal global embedding Ψ which preserves these local properties.

There are three often used techniques for building the similarity graph G . First, there is the ϵ -neighborhood graph which connects all vertices with distance $\|x_i - x_j\|^2$ smaller than ϵ . The ϵ graph is naturally symmetric [47] [46].

Contrary to this local connection is the fully connected graph which uses a similarity function that incorporates local neighborhood relations such as the Gaussian function: $w_{ij} = \exp(-\|x_i - x_j\|^2 / (2\sigma^2))$. This leads directly to the third step, since it implicitly defines the weights [46].

k -nearest neighbor (kNN) graphs combine both worlds by connecting each vertex only to its k -nearest neighbors.

C. Manifold Learning

In the field of machine learning, a very popular research area is manifold learning, which is related to the algorithmic techniques of dimensionality reduction. Manifold learning can be divided into linear and nonlinear methods. Linear methods, which have long been part of the statistician's toolbox for analyzing multivariate data, include Principal Component Analysis (PCA) and multidimensional scaling (MDS). Recently, researchers focus on techniques for nonlinear manifold learning, which include Isomap, Locally Linear Embedding, Laplacian Eigenmaps, Hessian Eigenmaps, and Diffusion Maps [46]. The algorithmic process of most of these techniques consists of three steps: a nearest-neighbor search, a computation of distances between points, and an eigen-problem for embedding the D -dimensional points in a lower-dimensional space. The manifold learning: Isomap, Laplacian Eigenmaps and Diffusion Maps will be compared and contrasted with the linear

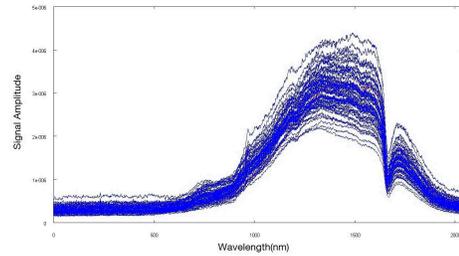


Fig. 13. Normalized spectral graph data sets, malignant skin lesions. Each curve is the vector, representing one skin lesion. Without labeling of the data the overlaps curves are difficult to separate

method PCA for a spectroscopic dataset. The goal is to find a mapping function Ψ from the original D -dimensional data set X to a d -dimensional dataset Y in which distances and information are preserved as much as possible and $d < D$:

$$\Psi : \mathbb{R}^D \rightarrow \mathbb{R}^d \quad (7)$$

In our case, we have $D = 2048$ and thus

$$\Psi : \hat{x}_i \in \mathbb{R}^{2048} \rightarrow y_i \in \mathbb{R}^d \quad (8)$$

where x_i and x_i are vectors and \mathbb{R}^d is a space.

D. System Experiments

We collected 372 spectroscopic data vectors from 48 patients, 326 measurements were of normal skin moles, 46 measurements were malignant skin lesion (as diagnosed by the treating physician). 13 cases out of 46 malignant skin lesions were pathologically verified by the laboratory. All lesions for this experiment were selected by only well-experienced physicians (not by newly joined dermatologists). This was the only additional protocol to the data acquisition protocols as discussed in section III-A. A color-coded representation of the normalized skin spectra data set is shown in Figure 13 and Figure 14. Figure 13 shows malignant skin lesions and Figure 14 shows malignant skin lesions combined with normal skin mole. In Figure 14 one can observe the overlap between two classes of data set.

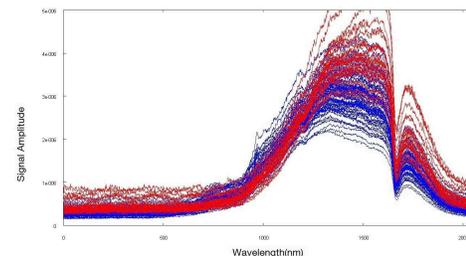


Fig. 14. Normalized spectral graph data sets combined form, blue for malignant skin lesions and red for normal skin mole.

The proposed methods were implemented in Matlab 10.1 using libraries for the dimensionality reduction. Clustering was performed by selecting a separating hyperplane in the processed three-dimensional data.

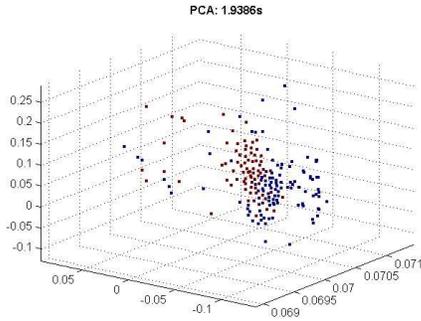


Fig. 15. PCA 3D representation of 2048D dataset. The best possible angle to visualize the data points. Blue for malignant skin lesions and red for normal skin mole. PCA:1.9386s is the runtime of method

TABLE III. CLUSTERING ACCURACY WITH DIFFERENT METHODS AND PARAMETERS. WHERE k IS k -NEAREST NEIGHBORS, A IS FOR ALPHA AND S IS REPRESENTING SIGMA PARAMETER

Parameters	Isomap	Laplacian Eigenmaps	Diffusion Maps
$k = 15, A = 2, S = 20$	88%	0%	10%
$k = 20, A = 2, S = 30$	90%	87%	81%
$k = 30, A = 1, S = 20$	86%	92%	90%
$k = 35, A = 1, S = 20$	94%	96%	92%

Before applying manifold learning we need to elucidate some parameters that play a significant role in producing meaningful data representation. The parameters for the non-linear dimensionality reduction techniques are:

- k : The k -nearest neighbors specify the number of nearest neighbors used to build the graph for the Isomap, Laplacian eigenmaps and Diffusion maps methods. If k is chosen too large or too small, the local geometry may not be interpreted correctly. Here we used the values of $k = 15, 20, 30, 35$.
- Alpha: This parameter controls the normalization.
- Sigma: This specifies the width of the Gaussian kernel. The larger Sigma is, the more weight far-away points will exert on the weighted graph. We used Sigma = 20, 30.

E. Discussion

All four studied methods (PCA, Isomap, Laplacian Eigenmaps and Diffusion maps) were applied independently. PCA is applied on 2048 dimensional data vectors, and the first three most significant components are taken. Each point represents one skin lesion (malignant or benign). The data set is labeled which is represented by two colors red and blue. Red points are malignant and blue are benign. It is clear from the 3D representation of the data shown in Figure 15 that the data is not clearly distinguishable into two clusters. The main reason PCA could not perform well is because PCA maximizes the variance of the data and in our case direction of the variance helps to distinguish between the two classes. The best clustering accuracy PCA achieved is 63%.

The 3D representation of the 2048D data vector after applying Isomap is shown in Figure 16. It is clear from the figure that some area of the data is very nicely clustered.

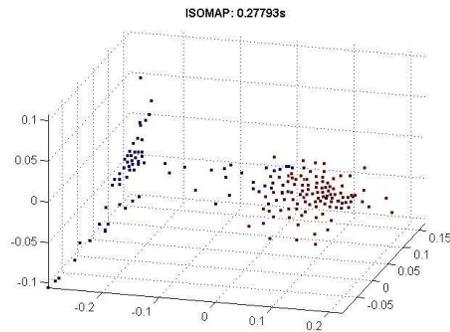


Fig. 16. Applying manifold learning by using Isomap and the output 3D representation as a result. Blue for malignant skin lesions and red for normal skin mole. The points that corresponds to malignant data examples, are well separated from those points corresponds to benign.

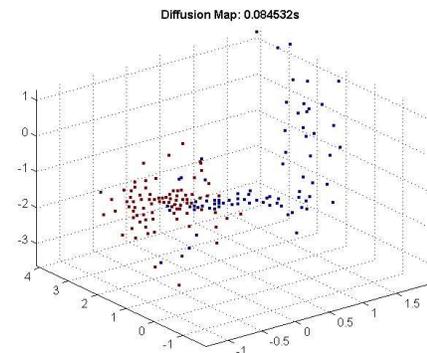


Fig. 17. Diffusion maps 3D data representation. The clusters are clearly visible. Blue for malignant skin lesions and red for normal skin mole.

The Isomap is governed by the geodesic distances between distant points, which causes distortions in local neighborhoods so maybe that is one reason that the data set is not clustered perfectly. Overall Isomap produce better results than PCA.

Figure 17, shows that the Diffusion maps is able to preserve the order of clusters in three dimensions similar as Isomap. Choosing the right parameter(s) is a difficult stage in manifold learning. Experiments are performed with different parameters as in table III. The results were computed as the number of

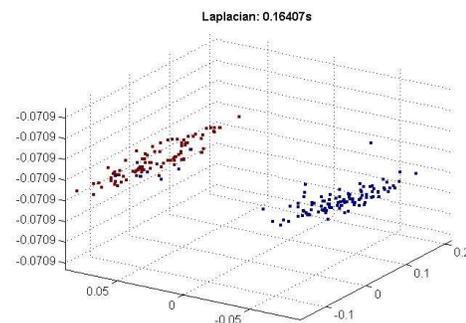


Fig. 18. Laplacian Eigenmaps 3D representation of 2048D dataset. Apart from few points which are in wrong cluster, the two clusters are well separated. Blue for malignant skin lesions and red for normal skin mole.

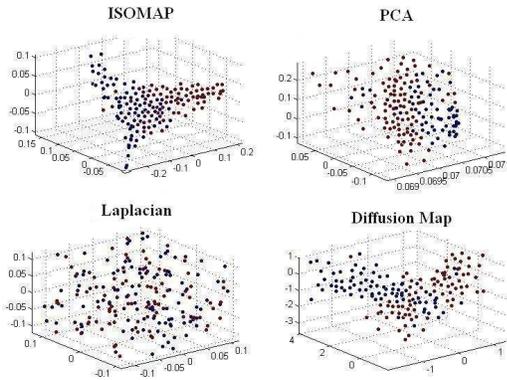


Fig. 19. A reduced 3D representation of spectroscopy 2048D dataset. The worst selection of parameters for all four methods. None of the method produced clear clustering of the dataset. Blue for malignant skin lesions and red for normal skin mole

correctly classified points over the total number of points and as a ground truth we have labeling provided by dermatologist. According to the literature [48], Diffusion maps perform better as compared to other manifold learning techniques but in our case Laplacian eigenmaps produces best results by choosing the right parameters shown in Figure 18. Laplacian eigenmaps preserve local neighborhood of the points which reflect the geometric structure of the manifold.

In Figure 19, all four methods are shown with worst parameters selection. The figure shows that the dataset is not easily distinguishable into two clusters. Variation in parameters for non-linear manifold learning methods are shown in Table III.

Isomap capture local geometry correctly and the dataset is clustered into two parts with an accuracy of 94% as shown in Table III. By increasing the neighborhood size to 20 and Sigma to 30, Laplacian eigenmaps and Diffusion maps perform better. Adding even more neighborhood information, Laplacian eigenmaps clustering accuracy improves to 96%. The parameters shown in the table are the only best combination for our dataset.

Four manifold learning techniques are applied to the problem of dimensionality reduction and clustering of optical spectroscopic data in dermatology. In contrast to the linear method PCA, all studied manifold learning techniques were able to perform satisfactorily in clustering normal skin mole from malignant skin lesions, provided the parameters were chosen correctly. In particular, Laplacian Eigenmaps look very promising for the intended dermatological application.

VII. CONCLUDING REMARKS ON FEASIBILITY OF OPTICAL SPECTROSCOPY FOR SKIN LESIONS CLASSIFICATION

In this part of the paper, optical spectroscopy for skin lesions classification is analyze. Optical spectroscopy by itself produces data, which, due to its high-dimensionality, cannot be directly utilized for classifying skin lesions. In other words, distinguishing between malignant and benign skin lesions is difficult. First the dimension of the data needs to be reduced in a meaningful way. In this respect, the application of manifold learning techniques to the problem of dimensionality

reduction and clustering of spectroscopic data in dermatology is introduced. One other problem in dermatology is about quantifying the progress of skin lesions. For this purpose, one needs to be able to numerically compare two or more images of e.g. the same lesion taken during different sessions. This involves accurate registration of all those images. Combination of optical spectroscopy with tracking as a solution to this problem is presented. In our approach, this combination is used as a guidance for acquiring spectral measurements at the same positions and orientations as the first acquisition. We defined spectroscopic data acquisition protocol in section III-A for using our system optimally. We also evaluated a patient dataset with an SVM-based classification of skin lesions.

The system opens a new way for utilizing the real potential of optical spectroscopy for noninvasive diagnosis of skin lesions. In taking optical spectroscopy even one step further using the system, it is a promising technique for the discrimination of malignant skin lesions from benign ones. Spectroscopy could form the basis of a clinical method to diagnose skin lesions due to the accuracy and reproducibility of its measurements. Acquisition of spectroscopic data causes little or no patient discomfort, does not alter the basic physiology of the skin, poses no hazard to the patient and does not interfere with any other standard clinical diagnostic practices. The scan could be performed by a non-specialist and therefore might be a useful tool for the prescreening of skin lesions. However, before full integration of spectroscopy into the clinical workflow, some further challenges need to be addressed:

- From our experience, there is need for several spectroscopic probes with different diameter sizes in order i) to cover only the area relevant to the lesion during the acquisition, i.e. to avoid getting measurements from the healthy skin region around the lesion and ii) to avoid multiple scans of the same lesion.
- In our experiments, we have observed that different samples taken from the same mole led to different spectral readings. A method is required to create a representative measurement from multiple spectroscopic readings for each mole.
- Optical spectroscopy based skin lesion diagnosis systems should be patient specific, since every patient has his/her own individual pattern of lesions which can be monitored throughout his/her body moles. In our study, we have observed that it is important to perform the classification within patient specific data in order to build a reliable system.
- Combining optical spectroscopy with other imaging technologies, e.g. dermoscopy imaging, multispectral imaging and hyperspectral imaging, can improve the diagnosis further, since the optical spectroscopy provides complementary information to these techniques.
- Patient age is an important factor which needs to be taken into account during the acquisition of optical spectroscopy data. As the cellular structures can change according to the age of the patient, differences in spectroscopic readings have been observed between young and elderly people, which can be addressed by creating groups of patients accordingly.

- Accurate data acquisition requires constant contact of the probe with the surface of the lesion which is hindered in some cases by ragged skin lesions. Further studies are required to investigate new techniques for data acquisition without touching the skin surface.
- A more in-depth study on data sets with larger variation is required to demonstrate general utility of optical spectroscopy in the clinical setting. Especially, data accompanied by pathological verification of malignant melanoma would be highly desirable to demonstrate the reliability of the presented methods.

REFERENCES

- [1] R. Marks, Epidemiology of melanoma, clinical excremental dermatology, PMID: 11044179 (PubMed - indexed for MEDLINE) 25 (2000) 395–406.
- [2] W. H. Organization, Ultraviolet radiation and the intersun programme, Website, <http://www.who.int/uv/faq/skincancer/en/> (2007).
- [3] O. Colot, R. Devinoy, A. Sombo, D. de Brucq, A color image processing method for melanoma detection, Medical Image Computing and Computer-Assisted Intervention MICCAI 1496 (1998) 562.
- [4] M. H. Madsen, L. Hansen, J. Larsen, K. Drzewiecki, A probabilistic neural network framework for detection of malignant melanoma, in artificial neural networks in cancer diagnosis prognosis and patient management, in: 18th IEEE Symposium on Computer-Based Medical Systems, 2001, pp. 141–183.
- [5] P. Rubegni, G. Cevenini, M. Burrioni, R. Perotti, G. Dell’eva, P. Sbrano, C. Miracco, P. Luzi, P. Tosi, P. Barbini, L. Andreassi, Automated diagnosis of pigmented skin lesions, Publication of the International Union Against Cancer 101 (2002) 576–580.
- [6] J. Sanders, B. Goldstein, D. Leotta, K. Richards, Image processing techniques for quantitative analysis of skin structures, Comput. Methods Programs Biomed. 59 (4) (1999) 167–180.
- [7] A. Bono, S. Tomatis, C. Bartoli, The invisible colors of melanoma. a telespectrophotometric diagnostic approach on pigmented skin lesions, Eur. J. Cancer 34 (10) (1996) 727–729.
- [8] S. Dreiseitl, L. O. Machado, H. Kittler, S. Vinterbo, H. Billhardt, M. Binder, A comparison of machine learning methods for the diagnosis of pigmented skin lesions, Journal of Biomedical Informatics 34 (2001) 28–36.
- [9] A. Blum, H. Luedtke, U. Ellwanger, R. Schwabe, G. Rassner, C. Garbe, Digital image analysis for diagnosis of cutaneous melanoma. development of a highly effective computer algorithm based on analysis of 837 melanocytic lesions, Br. J. Dermatol. 151 (5) (2004) 1029–1038.
- [10] R. J. Stanley, R. H. Moss, W. V. Stoecker, C. Aggarwal, A fuzzy based histogram analysis technique for skin lesion discrimination in dermatology clinical images, Comput. Med. Imag. Graph. 27 (2003) 387–396.
- [11] F. Ercal, A. Chawla, W. V. Stoecker, H. Lee, R. H. Moss, Neural network diagnosis of malignant melanoma from color images, IEEE Trans. Biomed. Eng. 14 (9) (1994) 837–845.
- [12] J. Boldrick, C. Layton, J. Ngyuen, S. Swtter, Evaluation of digital dermoscopy in a pigmented lesion clinic: Clinician versus computer assessment of malignancy risk, J. Amer. Acad. Dermatol. 56 (3) (2007) 417–421.
- [13] K. Hoffmann, T. Gambichler, A. Rick, M. Kreutz, M. Anschuetz, T. Grunendick, A. Orlikov, S. Gehlen, R. Perotti, L. Andreassi, J. N. Bishop, J. P. Cesarini, T. Fischer, P. J. Frosch, R. Lindskov, R. Mackie, D. Nashan, A. Sommer, M. Neumann, J. P. Ortonne, P. Bahadoran, P. F. Penas, U. Zoras, P. Altmeyer, Diagnostic and neural analysis of skin cancer (danaos). a multicentre study for collection and computer-aided analysis of data from pigmented skin lesions using digital dermoscopy, Br. J. Dermatol. 149 (10) (2003) 801–809.
- [14] G. Surowka, K. Grzesiak-Kopec, Different learning paradigms for the classification of melanoid skin lesions using wavelets, in: in Proc. 29th Annu. Int. Conf. IEEE EMBS, 2007, pp. 3136–3139.
- [15] Z. Zhang, R. H. Moss, W. V. Stoecker, Neural networks skin tumor diagnostic system, in: in Proc. IEEE Int. Conf. Neural Netw. Signal Process., 2003, pp. 191–192.
- [16] M. Kudo, J. Sklansky, Comparison of algorithms that select features for pattern classifiers, Pattern Recognit. 33 (2000) 25–41.
- [17] S. Umbaugh, Y. Wei, M. Zuke, Feature extraction in image analysis, IEEE Eng. Med. Biol. 16 (4) (1997) 62–73.
- [18] S. E. Umbaugh, R. H. Moss, W. V. Stoecker, Applying artificial intelligence to the identification of variegated coloring in skin tumors, IEEE Eng. Med. Biol. Mag. 10 (4) (1991) 57–62.
- [19] E. A. Edwards, S. Q. Duntley, The pigments and color of living human skin, American Journal of Anatomy 65 (3) (2005) 1–33.
- [20] S. Nischic, C. Forster, Analysis of skin erythema using true color images, IEEE Trans. Med. Imag. 16 (6) (1997) 711–716.
- [21] S. D. Depths, www.missionignition.net/bms/led_heal_clip_image008.jpg
- [22] M. Moncrieff, S. Cotton, E. Claridge, P. Hall, Spectrophotometric intracutaneous analysis: a new technique for imaging pigmented skin lesions, Br J Dermatol. 146 (3) (2002) 448–57.
- [23] P. name: MelaFind(R), Mela sciences, inc., Address: 50 South Buckhout Street, Suite 1, Irvington, NY 10533, <http://www.melasciences.com/> (2004).
- [24] P. name: SpectroShade(R), Mht s.p.a. headquarter, Production and Commercialization Via Milano 12, 37020 Arbizzano di Negrar (VR) Italy, www.mht.it (2007).
- [25] A. Kim, U. Kasthuri, B. Wilson, A. White, A. L. Martel, Preliminary clinical results for the in vivo detection of breast cancer using interstitial diffuse optical spectroscopy, MICCAI Workshop on Biophotonics Imaging for Diagnostics and Treatment 75 (2006) 1601–2321.
- [26] H. E. Kuppenheim, R. aymond Heer, Spectral reflectance of white and negro skin between 440 and 1000 μ m, Journal of Applied Physiology 4 (10) (1952) 800–806.
- [27] R. R. Anderson, J. A. Parrish, The optics of human skin, Journal of Investigative Dermatology 77 (1) (1981) 13–19.
- [28] J. W. Feather, M. Hajizadeh-Saffar, G. Leslie, J. B. Dawson, A portable scanning reflectance spectrophotometer using visible wavelengths for the rapid measurement of skin pigments, Phys. Med. Biol. 34 (1) (1989) 807–20.
- [29] B. L. Diffey, R. J. Oliver, P. Farr, A portable instrument for quantifying erythema induced by ultraviolet radiation, Br. J. Dermatol. 111 (1984) 663–72.
- [30] J. B. Dawson, D. J. Barker, D. J. Ellis, E. Grassam, J. A. Cotterill, G. W. Fisher, J. W. Feather, A theoretical and experimental study of light absorption and scattering by in vivo skin, Phys. Med. Biol. 25 (1980) 695–709.
- [31] T. J. Farrell, M. S. Patterson, B. Wilson, A diffusion theory model of spatially resolved, steady state diffuse reflectance for the non-invasive determination of tissue optical properties in vivo, Med. Phys. 19 (1992) 879–88.
- [32] A. Kienle, L. Lilje, M. S. Patterson, R. Hibst, R. Steiner, B. C. Wilson, Spatially resolved absolute diffuse reflectance measurements for non-invasive determination of the optical scattering and absorption coefficients of biological tissue, Appl. Opt. 35 (1996) 2304–14.
- [33] C. J. Lynn, I. S. Saidi, D. G. Oelberg, S. L. Jacques, Gestational age correlates with skin reflectance in newborn infants of 24 – 42 weeks gestation, Biol. Neonate 75 (1993) 9–64.
- [34] R. Marchesini, M. Brambilla, C. Clemente, M. Maniezso, A. E. Sichirollo, A. Testori, D. R. Venturoli, N. Cascinelli, In vivo spectrophotometric evaluation of neoplastic and non-neoplastic skin pigmented lesions–i. reflectance measurements, Photochem Photobiol 53 (1991) 77–84.
- [35] R. Marchesini, N. Cascinelli, M. Brambilla, C. C. L. M. E. Pignoli, A. Testori, D. R. Venturoli, In vivo spectrophotometric evaluation of neoplastic and non-neoplastic skin pigmented lesions–2. discriminant analysis between nevus and melanoma, Photochem Photobiol 55 (1992) 515–22.
- [36] A. Bono, S. Tomatis, C. Bartoli, G. Tragni, G. Radaelli, A. Maurichi, R. Marchesini, The abcd system of melanoma detection: a spectrophotometric analysis of the asymmetry, border, color, and dimension, Cancer 85 (1999) 7–72.

- [37] S. Sigurdsson, P. A. Philipsen, L. K. Hansen, J. Larsen, M. Gniadecka, H. C. Wulf, Detection of skin cancer by classification of raman spectra, in: *IEEE Trans. Biomed. Eng.*, 2004, pp. 1784–1793.
- [38] M. Osawa, S. Niwa, A portable diffuse reflectance spectrophotometer for rapid and automatic measurement of tissue, *Meas. Sci. Technol.* 4 (1993) 668–76.
- [39] N. Cristianini, J. S. Taylor, *An Introduction to Support Vector Machines and Other Kernel based Learning Methods*, 1st Edition, Cambridge University Press, 2000.
- [40] C.-C. Chang, C.-J. Lin, Libsvm: A library for support vector machines, Open source, <http://www.csie.ntu.edu.tw/~cjlin/libsvm> (2001).
- [41] L. Cayton, Algorithms for manifold learning, TR 2008–0923, University of California, Franklin St., Oakland, (June 15 2005).
- [42] J. B. Tenenbaum, V. de Silva, J. C. Langford, A global geometric framework for nonlinear dimensionality reduction, *Science* 290 (5500) (2000) 2319–2323.
- [43] M. Belkin, P. Niyogi, Laplacian eigenmaps and spectral techniques for embedding and clustering. in: *Advances in neural information processing systems*, MIT Press, Cambridge 14 (1) (2002) 585–591.
- [44] R. Coifman, S. Lafo, Diffusion maps: Applied and computational harmonic analysis, *Special Issue: Diffusion Maps and Wavelets* 21 (1) (2006) 5–30.
- [45] T. Hastie, R. Tibshirani, J. H. Friedman, *The elements of statistical learning: Data Mining, Inference and Prediction*, Springer, 2001.
- [46] R. Socher, M. Hein, *Manifold Learning and Dimensionality Reduction with Diffusion Maps*, Tech. report, 2008.
- [47] U. von Luxburg, A tutorial on spectral clustering, *Stat Comput*, Springer Science and Business Media 17 (5) (2007) 395–416.
- [48] B. Bah, *Diffusion maps: Analysis and applications*, dissertation 66, Mathematical Modeling and Scientific Computing, Wolfson College, University of Oxford, UK (September 2008).