Review of techniques to analyze spectra obtained across diffuse reflectance for diagnosis of skin lesions

E. E. Orozco-Guillén*,†, J. A. Delgado-Atencio†,**, S. Vázquez-Montiel†, J. Castro-Ramos† and M. Cunill-Rodríguez†,**

*Facultad Experimental de Ciencia y Tecnología, Departamento de Física, Universidad de Carabobo, C.P. 2005 Valencia-Venezuela
†Instituto Nacional de Astrofísica Óptica y Electrónica, C.P 72000 Puebla, Pue. México.
**Department of Physics, Centre for Technological Applications and Nuclear Development, CEADEN-CITMA. P.O. Box 6122. Havana, Cuba.

Abstract. Determination of tissue optical properties across of diffuse reflectance spectroscopy is an effective and extensively used technique for the non-invasive study of tissues. The methods to obtain the spectra are practically uniform, however, approaches for signal detection and data processing have varied widely. In some cases it is used a physical model (Monte Carlo model, the diffusion approximation to the transport equation), in other cases an empiric model (principal component analysis) accompanied by tools (Support Vector Machine, Neural Network, Partial Least Squares, etc.) that allow to analyze the obtained spectra. In this work, we present a critical review of the available methods used to analyze the measured diffuse reflectance spectra of human skin for diagnostic purposes.

Keywords: optical properties; reflectance; simulations; tissue optics, pattern recognition

PACS: 87.64.Cc, 87.85.Lf, 87.10.Rt, 42.30.Sy

INTRODUCTION

Diffuse reflectance spectroscopy is a non-invasive optical technique that could augment the current standard of cancer diagnosis, and provide real-time feedback on the optimal biopsy location before tissue is removed. The diffuse reflectance spectrum reflects the absorption and scattering properties of the tissue. This technology is fast, quantitative and sensitive to alterations in tissue structure and biochemistry [1]. Several studies have demonstrated that diffuse reflectance spectroscopy can diagnose early epithelial cancers with high sensitivity and specificity [2, 3].

DIFFUSE REFLECTANCE

Light is incident on the surface of tissue, which is assumed to be a homogeneous semi-infinite turbid medium characterized by absorption and reduced scattering coefficients, \( \mu_a \) and \( \mu_s' \) respectively. The tissue is also characterized by a refractive index, \( n \). Part of the incident light is absorbed in tissue while the non-absorbed part is subjected to multiple scattering and eventually emerges from the tissue surface as diffuse reflectance.
Instrumentation

The experimental setup employed is essentially, a spectrophotometer capable of collecting spectra in the 400-1100 nm range. Light is delivered and collected by means of a fiber optic probe, the illumination is provided by a tungsten-halogen light source. All spectra are referenced to a reflectance calibration standard and normalized dividing by the diffuse reflectance spectrum of the calibration standard.

DATA PROCESSING TECHNIQUES

Determination of tissue optical properties using reflectance is well established, however, approach for signal detection and data processing have varied widely[4]. A common approach that has been shown to be highly successful in estimating optical properties involves development of a predictive empirical model through three essential steps: 1. generation of steady-state spatially resolved reflectance data for model calibration through modeling or experimental approaches, 2. development of an inverse model by multivariate calibration, and 3. application of the trained model to unknown samples to predict the optical properties [5].

Monte Carlo Method

By considering the skin as an optically turbid medium, by knowing the absorption and diffusion coefficients and also by considering the diffusion of probability and absorption per unit of distance traveled for each wavelength. Bohnert, and Walter[6] developed a mathematical model of the skin by means of Monte Carlo simulations. By giving the coefficients absorption of the hemoglobin monoxide of carbon (CO-hb) in the skin model they investigate whether the method is suitable to determine the concentration of CO-hb from the reflectance spectral curves of liver from death patients[6]. From reflectance time resolved for a biological medium Alwin Kienle [7] et al, had shown with only one simulation of Monte-Carlo it is possible to obtain absorption and scatter coefficients. They obtain errors less than 1% and 2% in the value of coefficients respectively.

Solved reflectance has been studied by Frederic Bevilacqua et al, for turbid media, by analyzing the separation source - detector by using Monte-Carlo simulations[8]. They found that moments of the phase function play an important role on the reflectance curves. Q. Liu and N. Ramanujam[9] have developed a scaling method to calculate the diffuse reflectance of mediums compose by multi-layers with many optical properties for several wave length range from ultraviolet to visible. Scaling method employ information trajectory of the photon generated by Monte Carlo simulations in an homogeneous medium in order to scale the exit distance and weight of the exit photon to generate a new set of optical properties en the multi layer medium, it is useful to generate diffuse reflectance spectrum to extract the optical properties of each layer.
Principal Component Analysis

Principal component analysis is a mathematical tool that reduces the dimensions of a dataset to a set of informative principal components (PCs) that account for most of the variance of the original dataset. The first PC accounts for as much of the variability in the dataset as possible, and each succeeding component accounts for as much of the residual variability as possible [10]. Therefore the PCs are normally arranged in the order of their contributions to the variance of entire dataset. In principle, PCA is an operation that rotates the coordinates of the original data to form new coordinates using the PCs. Most of the information carried in the data is distributed in the first few PCs of the new coordinates. The contributions from the rest of the PC coordinates are negligible. By presenting the original data in new PC coordinates formed with a few informative PCs, the dimensions of the data can be significantly reduced without losing important information.

The PCA is used to process diffuse reflectance spectra, it transforms wavelengths, the original spectral variables, into a set of PC spectra. Each original spectrum is a combination of the PC loading spectra that are orthogonal to each other. The PCs with negligible contributions to the variance of the dataset are eliminated. The dimensions of the dataset for developing the diagnostic algorithm can then be significantly reduced without losing useful information.

Support Vector Machine Algorithms

A support vector machine is a new method for classifying multivariate data. It is based on the principle of minimization of structural risk in constructing an optimally separating hyperplane that separates different classes of data. When the separating boundary is nonlinear, SVM maps the sample data with specific kernel functions to a higher dimensional feature space to linearize the boundary and generate the optimal separating hyperplane. Compared with other multivariate statistical methods, SVM was designed to be particularly effective in developing a reliable classifier from a training set with a small sample size [11].

The possibility of using SVMs for developing diagnostic algorithms is also attracting attention. While Palmer et al.[12] used a linear SVM classifier for classifying autofluorescence and diffuse reflectance spectra of breast tissues in vitro, Lin et al.[13] classified in vivo autofluorescence spectra from tissues by using both the linear and the nonlinear SVM classifier with RBF kernel. In the reports of both the groups, the tissue spectra were dimensionally reduced by applying linear PCA algorithms prior to using the SVM approach for classification. Lin et al.[13] showed that the classification performance of an SVM classifier trained on the full spectral data was comparable to that obtained with the classifier trained on the diagnostically relevant principal components only. Their combined PCA-SVM approach was reported to have reduced computational complexity.
Artificial Neural Networks

The application of neural networks in biomedical optics has been spread abruptly, specially to solve the inverse problem of determination of the optical properties from spatially resolved diffuse reflectance data and to classify lesions from diffuse reflectance spectra measurements [14]. Two major experimental approach has been employed to carry out the classification of pigmented skin lesions in combination with neural network analysis, namely camera based systems that analysis images taken at multiple wave-lengths and point spectral measurement of diffuse reflectance using fiber optics probe in contact with the skin lesion. An example of the first approach is the diagnostic and neural analysis of skin cancer (DANOS) yielding an sensitivity/specificity of 82/85% respectively during a multicentre study of pigmented skin lesion using digital dermoscopy [15]. A comparison of multivariate discriminant analysis and artificial neural networks in the classification of reflectance spectra is done by Wallace et al [16]. They report a classification performance of 83/88% using a neural network classifier. In our opinion, the use of artificial neural network could be a good starting point to establish classification algorithms in skin lesion using fiber optic diffuse reflectance spectroscopy. However attention must be paid to the applicability of this technique regarding issues such as the number of input data, overfitting, and generalization to data outside the training set.

CONCLUSION

This work has given an overview of the techniques to analyze spectra obtained across diffuse reflectance for diagnosis of skin lesions, the current state of research .......... INAOE, UC, etc.

REFERENCES

detector separations close to one transport mean free path”, J. Opt. Soc. Am. A, 16, 12, 2935-2945,
(1999).
10. W. R. Dillon, and M. Goldstein, Multivariate analysis: methods and applications (Wiley, New York,
1984).
2000).
multiexcitation fluorescence and diffuse reflectance spectroscopy for the diagnosis of breast cancer”
No12, 2281-2286.
of skin cancer (DANAOS). A multicentre study for collection and computer-aided analysis of data
reflectance spectra from pigmented skin lesions, a comparison of multivariate discriminant analysis