Detecting jaundice by using digital image processing

J. Castro-Ramos[1], C. Toxqui-Quitl[2], F. Villa Manriquez[1], E.Orozco-Guillen[3], A. Padilla-Vivanco[2], [4]JJ. Sánchez-Escobar

[1] Instituto Nacional de Astrofísica, Óptica y Electrónica; [2] Univ. Politécnica de Tulancingo; [3] Univ. Politécnica de Sinaloa, [4]Centro de Enseñanza Técnica Industrial (México).

ABSTRACT

When strong Jaundice is presented, babies or adults should be subject to clinical exam like "serum bilirubin" which can cause traumas in patients. Often jaundice is presented in liver disease such as hepatitis or liver cancer. In order to avoid additional traumas we propose to detect jaundice (icterus) in newborns or adults by using a not pain method. By acquiring digital images in color, in palm, soles and forehead, we analyze RGB attributes and diffuse reflectance spectra as the parameter to characterize patients with either jaundice or not, and we correlate that parameters with the level of bilirubin. By applying support vector machine we distinguish between healthy and sick patients.

Keywords: Digital Image Processing, Diffuse Reflectance, Bilirubin, Support Vector Machine.

1. INTRODUCTION

Many methods have been developed to determine the level of bilirubin in newborns or adults both invasive and non invasive, Penhaker[1] designed an electronic instrument based on transmitted light through the skin. The skin photo-diagnostic handle 450 nm - 575 nm monochromatic light. Buttitta et al [2] develop a non invasive bilirubin monitor which uses two wavelengths of light radiation directed to an infant arterial system. The reflectivity or backscatter of the light from the infant`s bloodstream from the light sources is detected an measured to determine the bilirubin level of that arterial system of the infant. Alla Suresh et al [3] develop a method to isolate the intravascular and extra vascular bilirubin starting with the diffuse reflectance spectrum. A nonlinear optimization algorithm was adopted to extract the optical properties including bilirubin concentration from the skin reflectance spectrum. A support vector machine (SVM) is a method for classifying multivariate data. The possibility of using SVM for develop any diagnostic algorithms is also attracting attention. While Palmer et al.[4] used a linear SVM classifier for classifying auto fluorescence and diffuse reflectance spectra of breast tissues in vitro, Lin et al.[5] classified in vivo auto fluorescence spectra from tissues by using both the linear and the nonlinear SVM classifier with RBF kernel. In the reports of both groups, the tissue spectra were dimensionally reduced by applying linear PCA algorithms prior to using the SVM approach for classification. Lin et al.[5] 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. In this paper we develop a novel idea by taking an image we get the RGB attributes and by using a spectrometer we get the diffuse reflectance spectra from patients with low and high level of bilirubin we

Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XXI, edited by Thomas G. Brown, Carol J. Cogswell, Tony Wilson, Proc. of SPIE Vol. 8949, 89491U

correlate the values with a BILITEST®, hence by applying support vector machine we were able to develop an algorithm to distinguish between low and high level of bilirubin for any patient.

2. HYPERBILIRUBINEMIA

The bile a liquid green substance produced by the liver it is continuously segregate by the liver and in the inter-digestive periods it is stored in the gallbladder and liberate to the duodenum after meals consume. In many times the liver cannot break down the red blood cells which produce hyperbilirubinemia, it increases the level of bilirubin too much in the organism causing damage in the brain and death in few cases [6]. In the figure 1 we show the diagram the digestive system where is mentioned the responsible organs to produce hyperbilirubinemia.

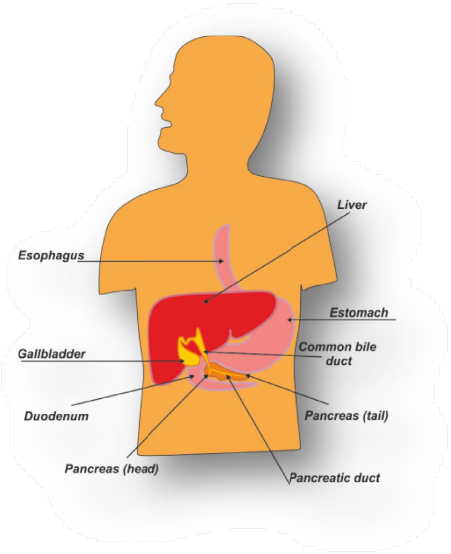


Figure 1. digestive system with the responsible organs to produce hyperbilirubinemia.

3. DIGITAL IMAGE PROCCESING

The most commonly employed color space in computer technology is the RGB color space, which is based on the additive mixture of the three primary colors *R, G, B* [7, 8]. The terms red, green, and blue are a standardization to provide descriptions for the primary colors. The primary colors are used as reference colors in the majority of the imaging sensors. They form the base vectors of a three-dimensional orthogonal color-vector space, where the zero-vector represents black. The origin is also described as black point.

On the other hand, any color can be viewed as a linear combination of these base vectors in the RGB space. In one such accepted RGB color space, a color image is mathematically treated as a vector function with three components. Specifically, for an image

$$I(x,y) = (R(x,y), G(x,y), B(x,y))^T.$$

These values are referred to as tristimulus values.

4. SUPPORT VECTOR MACHINE

The foundation of Support Vector Machines (SVM) have been developed by Vapnik [9] and is gaining popularity due to many attractive features, and promising empirical performance. It is based on the principle of minimization of structural risk in constructing an optimally separating hyper plane 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 hyper plane [10].

A classification task is usually involved with training and testing data which consist of some data instances. Each instance in the training set which contains one ``target value" (class labels) and several ``attributes" (features). The goal of SVM is to produce a model which predicts target values of data instances in the testing set which are given only by the attributes.

Mathematically, a hyper plane is defined by $w.x + b = 0$ in the feature space of the sample data, where w is the norm to the hyper plane and b is a plane constant, in figure 2 a representation is shown.

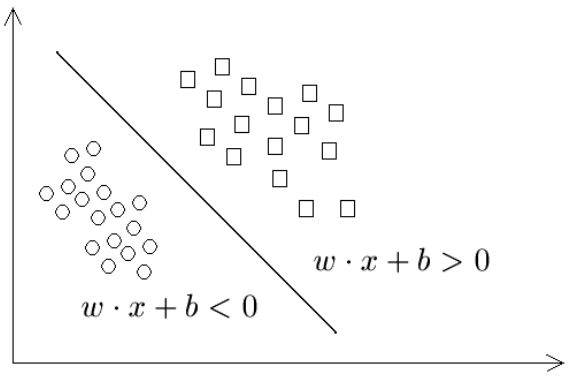


Figure 2. Separating Hyper plane

Given a training set of instance-label pairs $(x_i, y_i), i = 1, \ldots m$ where $x_i \in R^n$ and $y_i \in \{-1,1\}^{\mathbf{m}}$. Where m and n are the number and dimension of the sample vectors. If we consider that a hyper plane can separate the positive samples $(y_i = +1)$ from the negative samples $(y_i = -1)$ in the feature space of sample data; then $y_i\ (w.x + b) \geq 0$. The margin of a separating hyper plane, defined as the sum of the shortest distances from the positive and negative samples to the hyper plane, equals $\frac{2}{\|w\|}$. The SVM algorithm is simply looking for an optimal hyper plane with the maximal separating margin by

$$\text{Minimizing} \quad \frac{\|w\|^2}{2}$$
$$\text{subject to } y_i\ (w \cdot x + b) \geq 1.$$

When two classes of training samples are not linearly separable, the constraint is relaxed to allow misclassification:

$$y_i\ (w \cdot x + b) \geq 1 - \zeta_l, \quad \zeta_i \geq 0,\ i = 1, \ldots l$$

The optimization problem becomes

minimizing $\frac{\|w\|^2}{2} + C \sum \zeta_i$ subject to $y_i\ (w \cdot x + b) \geq 1$

or

maximizing $\sum \alpha_i - \frac{1}{2} \sum \alpha_i\, y_i y_j \left(x_i, x_j\right)$ subject to $\sum \alpha_i\, y_i = 0,\ 0 \leq \alpha_i \leq C$

where C is a parameter chosen by the user to define the cost of constraint relaxation. A higher value C of corresponds to a higher penalty for the assigned errors. Each α_i is a Lagrange multiplier corresponding to a sample $\left(x_i, x_j\right)$ in the training set.

The optimization process determines support vectors that are the training sample vectors with nonzero Lagrange multipliers. A linear SVM classifier can then be constructed using these support vectors as

$$f(x) = sgn[\sum_{i=1}^{k} \alpha_i y_i (x_i \cdot x_j) + b]$$

where k is the number of support vectors and x is the sample to be classified.
If the separating boundary is nonlinear in the feature space of the data sample and the linear SVM classifier cannot separate the training data well, improved classification results may be obtained using a nonlinear SVM method. Assuming the nonlinear separating boundary can be linearized in a higher dimensional feature space the nonlinear SVM classifier can be constructed as

$$f(x) = sgn\left[\sum_{i=1}^{k} \alpha_i y_i K(x_i, x_j) + b\right], \qquad \alpha_i \geq 0$$

Where $K(x_i, x_j)$ is called the kernel function. SVM and Data Mining reported in literature give the following four basics kernel:
linear:

$$K(x_i, x_j) = x_i^T x_j$$

polynomial:

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \qquad \gamma > 0$$

radial basis function (RBF):

$$K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|}, \quad \gamma > 0$$

sigmoid:

$$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$$

Here γ, r and d are kernel parameters.

$$w \cdot x + b > 0$$

5. EXPERIMENTAL SETUP AND RESULTS

We use a spectrometer ocean optics USB 4000, to get the diffuse reflectance spectra, with the idea to have a useful device, we acquire the images with a smart phone camera, we use a camera color 640X480X3 it was enough to acquire differences between patients with low level of bilirubin or with high level of bilirubin. We measure a set of 20 Mexican patients, 10 with high level of bilirubin and 10 with low level of bilirubin. Either both health and ill patients were brown skin color Mexican race, body temperature was taken. Fluorescent lamps was used which luminance of light was controlled at 180 lux inside the room. By one side we measure the level of bilirubin with a bilirubinometer BILITEST from technomedica to distinguish between high and low level of bilirubin. For other side we measure rms of the RGB, the diffuse reflectance spectra and for reference the level of bilirubin for each patient. Due a best signal of bilirubin, we measure four parts of the body palm, sole, forehead and arm. In figure 3, we can see the experimental setup. In figure 4, we can see any differences between the diffuse reflectance spectrum of several patients, also. figure 5 shows the differences in RGB values in patients with high and low level of bilirubin,

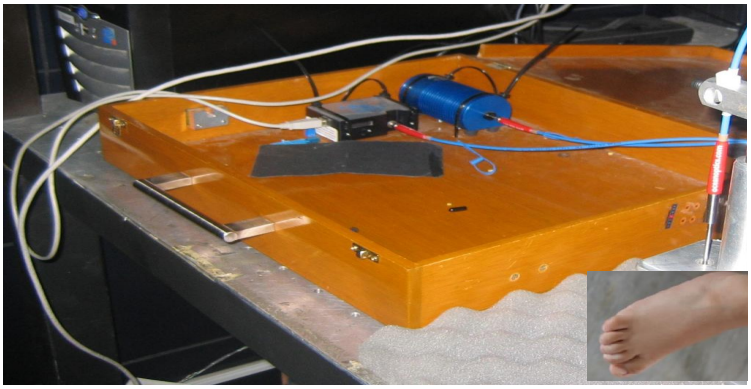
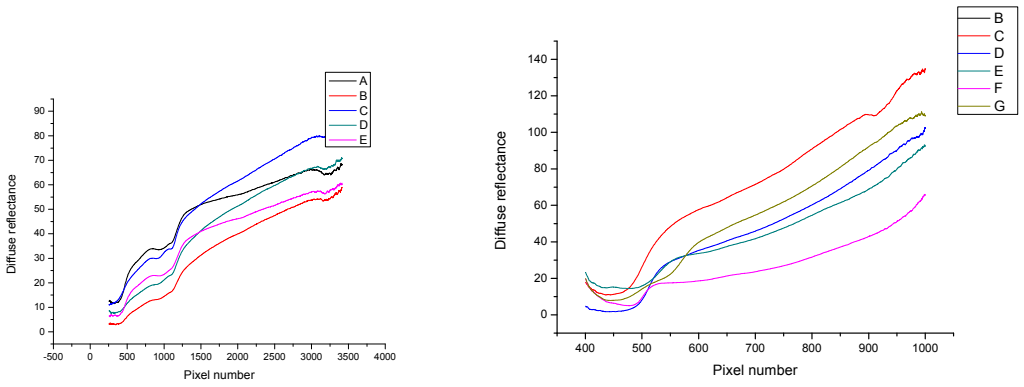
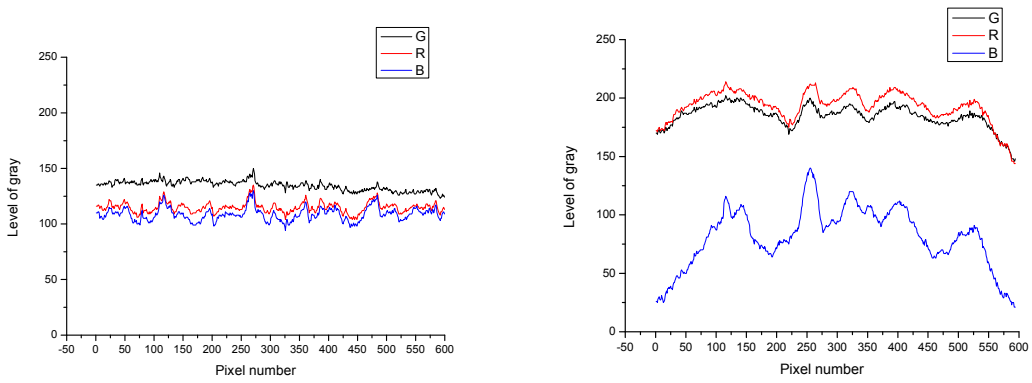


Figure 3. Experimental setup which shows the diffuse reflectance spectrometer, the test probe and the sole of the foot of any patient.



a) b)

Figure 4. Diffuse reflectance curves of a) normal health patients and b) with high level of bilirubin in the visible range.



a) b)

Figure 5. RGB values a)for a normal patient and the other one with hyperbilirrubin, in all cases we note that in normal patients BR channels appear together and in patients with high level of bilirubin channels GR appear together.

A good statistical parameter in a sample is the rms it was take with each channel RGB also the rms of the reflectance was taken and by applying support vector machine described above and taking a training set of 20 patients, we were able to get a classifier, figure 6 shows the separating hyper plane for patients with high and low level of bilirubin.

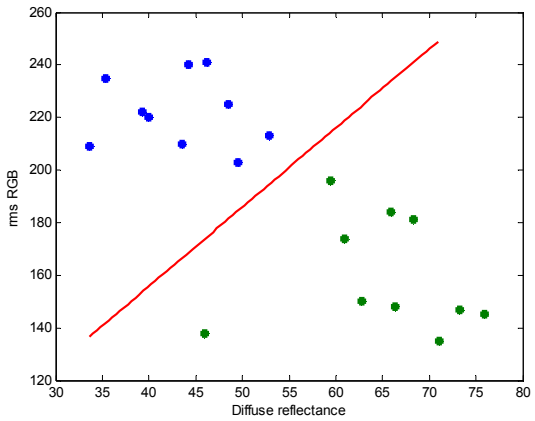


Figure 6. Classification of the support vector machine for high level of bilirubin (blue dot) and for the low level of bilirubin (green dot).

We can see in figure 6, a good classification using the support vector machine for high level of bilirubin and for the low level of bilirubin.

6. CONCLUSION

Support vector machine together with digital image processing seems to be a good technique to distinguish between low and high levels of bilirubin we obtain a sensitivity of 71.8 % and a specificity of 78.8 obtained with 20 spectra, 10 with high bilirubin and 10 with low level of bilirubin. As future work we will work with more patients to improve the sensitivity and specificity and we are will working with an app for smart phone to develop a predictive method for hyperbilirubinemia in that equipments.

REFERENCES

[1] M, Penhaker, V. Kasik, and B. Hrvolova, "Advanced Bilirubin Measurement by a photometric Method", Elektronika ir elektrotechnika, vol. 19, 3, 2013.
[2] Anthony Buttitta and Steven Falk, "Non invasive bilirubin monitor", European Patent Application, EP0747002A1, 11.12.1996.
[3] Clinical system model for monitoring the physiological status of jaudience by extracting bilirubin components from skin diffuse reflectance spectra, Advanced Biomedical and Clinical Diagnostic Systems VII, edited by Anita Mahadevan-Jansen, Tuan Vo-Dinh, Warren S. Grundfest, Proc. of SPIE Vol. 7169, 2009.
[4] G. M. Palmer, C. Zhu, T. M. Breslin, F. Xu, K. W. Gilchrist, and N. Ramanujam, “Comparison of multiexcitation fluorescence and diffuse reflectance spectroscopy for the diagnosis of breast cancer” IEEE Trans. Biomed. Eng. 50, 1233-1242 (2003).

[5] W. Lin, X Yuan, P Yuen, WI Wei, J Sham. PC Shi. “Classification of in vivo autofluorescence spectra using support vector machines”. J Biomed. Opt.9, 180-186(2004).
[6] Lee Goldman, Andrew Schafer, Goldman`s Cecil Medicine, Section XIII, Diseases of the liver, Gallbladder and Bile Ducts, Saunders of Elsevier Inc, 2000.
[7] González, R.C., Wintz, P. (1996). *Procesamiento digital de imágenes*. Addison-Wesley.
[8] J.J. Baez Rojas, M. Guerrero, J. Conde Acevedo, A. Padilla vivanco and G, Urcid Serrano, Segmentacion de imagenes de color, Rev. Mex. Fis, 50, 6, pp. 579, 2004.
[9] V. Vapnik, The Nature of Statistical Learning Theory, Chap. 4, Springer-Verlag, New York (1995, 2000).
[10] A Practical Approach to Neonatal Jaundice, Sarah k. Moerschel, Lauren b. Cianciaruso, and Lloyd r. Tracy, *Am Fam Physician*. 2008 May 1,77(9),1262.