



HAL
open science

Analysis of human skin hyper-spectral images by non-negative matrix factorization

July A Galeano Z, Romuald Jolivot, Franck Marzani

► **To cite this version:**

July A Galeano Z, Romuald Jolivot, Franck Marzani. Analysis of human skin hyper-spectral images by non-negative matrix factorization. 10th Mexican International Conference on Artificial Intelligence, Nov 2011, Puebla, Mexico. pp.431-442. hal-00790466

HAL Id: hal-00790466

<https://hal.science/hal-00790466>

Submitted on 20 Feb 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Analysis of Human Skin Hyper-Spectral Images by Non-negative Matrix Factorization

July Galeano, Romuald Jolivot, and Franck Marzani

Université de Bourgogne,
LE2I Laboratoire Électronique, Informatique et Image,
UFR Science et Techniques,
BP 47870,21078 Dijon Dedex, France.

{july.galeano-zea,romuald.jolivot,franck.marzani}@u-bourgogne.fr

Abstract. This article presents the use of Non-negative Matrix Factorization, a blind source separation algorithm, for the decomposition of human skin absorption spectra in its main pigments: melanin and hemoglobin. The evaluated spectra come from a Hyper-Spectral Image, which is the result of the processing of a Multi-Spectral Image by a neural network-based algorithm. The implemented source separation algorithm is based on a multiplicative coefficient update. The goal is to represent a given spectrum as the weighted sum of two spectral components. The resulting weighted coefficients are used to quantify melanin and hemoglobin content in the given spectra. Results present a degree of correlation higher than 90% compared to theoretical hemoglobin and melanin spectra. This methodology is validated on 35 melasma lesions from a population of 10 subjects.

Keywords: Blind source separation algorithms, Non-negative Matrix Factorization, human skin absorbance spectrum, Multi/Hyper-Spectral imaging

1 Introduction

Human skin is a complex multilayered structure composed of several particles that imply different physical phenomena. From an optical point of view, the dominant effects correspond to scattering and absorption. This last effect mostly occurs at melanocyte and erythrocyte cells which are known to contain chromophores. The main light-absorbing pigments present in those cells are melanin and hemoglobin respectively [1]. Melanin is the chromophore of human skin charged mainly for the protection from solar radiation, and in the assessment of skin color. It is also involved in several human skin pathologies such as malignant melanoma, albinism, vitiligo and melasma [2]. Hemoglobin is the pigment related to red blood cells, which are mostly present in vascular densities. This fact has shown the importance of hemoglobin in the study of gastroenterological diseases by improving the classification of colic polypus [3].

Melanin and hemoglobin seem to be the clue for the analysis of different diseases underlying human skin. Nevertheless, since human skin spectrum is the result of the complex light-skin interaction, the obtention of melanin and hemoglobin in a separately way is not an easy task. Due to this, it is important to have a device able to perform two main tasks: to detect human skin absorption/reflectance spectrum, and to discern melanin from hemoglobin.

Several works related to the achievement of those both components can be found. Most of them use spectrometers as the instrument for the acquisition of human skin absorption spectrum. Their difference remains in the way as data is analysed. Among those ways, ones can enumerate Monte Carlo simulations, Kubelka-Munk theory, linear regression, and statistical approaches [4–6].

This paper presents a Multi-Spectral approach for the obtainment of human skin reflectance spectrum, and propose the use of a simple Blind Source Separation (BSS) method for the estimation of the amount of melanin and hemoglobin. The human skin data used for the performed analysis was obtained from 10 volunteers presenting with melasma lesions. The later allows us to corroborate the potential use of BSS method in the discernment between melasma and healthy skin zones. The device used for the data acquisition, a Hyper-Spectral system, is presented in section 2. The BSS algorithm, used for the analysis of the obtained Hyper-Spectral data, is presented in section 3. Finally the obtained results are presented and discussed in section 4.

2 Hyper-Spectral Data Acquisition: ASCLEPIOS System

Among the optical instruments used for the study of human skin, it is possible to find color-based instruments (as color cameras) and spectrometers [4]. The first one allows the analysis of a huge area of interest based on the information given by three wide spectral bands, RGB. Although spectrometers allow performing a deeper analysis about skin-light interactions, they allow the analysis in only small areas of interest. The tradeoff is then between size of the area to be evaluated (region of interest ROI) and the accuracy of spectral information for each pixel. This last fact determines the kind of analysis which can be performed.

Multi-Spectral and Hyper-Spectral Imaging (MSI/HSI) systems are presented then as a way to overcome the problematic size of ROI-kind of analysis. Since it acquires images at different spectral bands, shorter than the conventional RGB systems, MSI-HSI systems allow then a spectral analysis in a big area of interest of the object under study. This information gives not only a merely information and improvement about skin color, but also information about how skin's components interact with light.

For the purpose of the present work, we use a system called ASCLEPIOS (Analysis of Skin Characteristics by Light Emission and Processing of Images Of Spectrum), which is an innovative system since it evolves from a MSI into a HSI system without the need for an increased number of wavelengths [7]. As depicted in figure 1, the device consists of a camera, and a rotating wheel with a

set of 10 filters. Those filters are between the wavelengths of 430 to 780 nm, each one with Full Width at Half Maximum (FWHM) of 80 nm. In this way the object of interest is illuminated at 10 different spectral bands leading to a MSI. Then, by using a software based on artificial neural networks [8], an Hyper-Spectral cube of the object under study is obtained. The cube is then the HSI. The x,y coordinates of the HSI correspond to the spatial dimension giving the reflectance value at each pixel coming from the camera. The z coordinate is the spectral dimension corresponding to 36 different spectral bands with a FWHM of 10 nm. In this way, a reflectance spectrum composed of 36 values between 430 and 780 nm, is obtained at each pixel of the HSI.

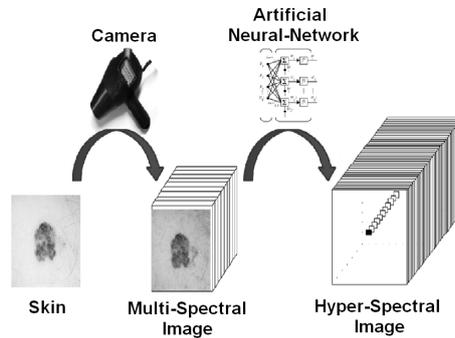


Fig. 1. Scheme of the used system called ASCLEPIOS.

This Hyper-Spectral device has been validated on a population composed of 150 healthy participants from five different Skin Photo Types (SPT) at different body locations. The validation was performed by comparing data acquired from a commercial spectrophotometer with the spectra processed from averaging the obtained HSI. The results revealed that the system is able to provide HSI with a Goodness of Fit Coefficient (GFC) superior to 0,997 for the average of all SPT for each location. This means that ASCLEPIOS system provides accurate Hyper-Spectral Images which can be effectively used for analysis of skin reflectance spectra [9].

3 Hyper-Spectral Data Modeling and Analysis

The acquired Hyper-Spectral data are analysed by a linear BSS method. Such method represents the given data as a weighted sum of source components. The later implies linearity in the phenomena to be studied. Since our goal is to obtain the principal components of human skin from a measured reflectance spectrum, it is important to represent the physical phenomena of light-skin interaction in a linear way. For that representation, the study presented in this article is based on the modified Beer-Lambert law.

3.1 Physical approach: Light-skin Interaction

Human skin is presented as a scattering multi-layered media composed of different pigments. As depicted in figure 2, the principal pigments of human skin, melanin and hemoglobin, are present in epidermis and dermis respectively [10, 11].

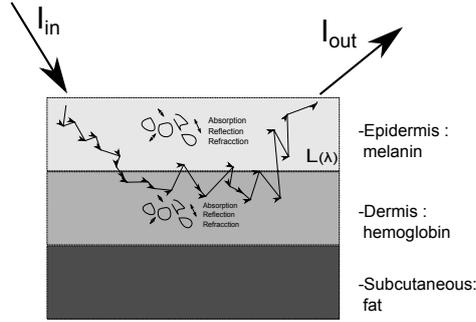


Fig. 2. Light-skin interaction: when light interacts with the different layers present in skin, absorption, reflection and refraction occurs at the pigments. The principal pigments present in human skin are considered to be melanin and hemoglobin.

When light interacts with skin, light travels through the different layers where scattering, absorption, and reflection occurs at the pigments. As a result, light travels inside human skin through a geometrical path dependant of the wavelength ($L(\lambda)$). Since ASCLEPIOS system detects reflectance $R(\lambda)$ spectrum as the ratio of incident to reflected energy, absorbance spectrum $A(\lambda)$ can be deduced from reflectance by equation 1 [12]:

$$A(\lambda) = -10 \log(R(\lambda)). \quad (1)$$

In optical absorption terms, the modified Beer-Lambert law holds. The total absorption of human skin can be determined as the contribution of the absorption values present at the different layers. This is expressed by equation 2 [11].

$$A(\lambda) = \sum_{i=1}^n \Delta A_i(\lambda) = \sum_{i=1}^n C_i \epsilon_i(\lambda) L(\lambda) + G. \quad (2)$$

According to the relation 2, the change in absorbance at layer i ($\Delta A_i(\lambda)$) is related to the molar absorption coefficient of pigment i ($\epsilon_i(\lambda)$), the concentration C_i , and the mean path length $L(\lambda)$. G are the losses due to the components not considered in the model.

In this work, BSS allow then to obtain the spectral absorption of dermis and epidermis. Those spectra are considered then as an approximation to the molar absorption coefficient of melanin and hemoglobin respectively, which average spectra are presented in figure 3.

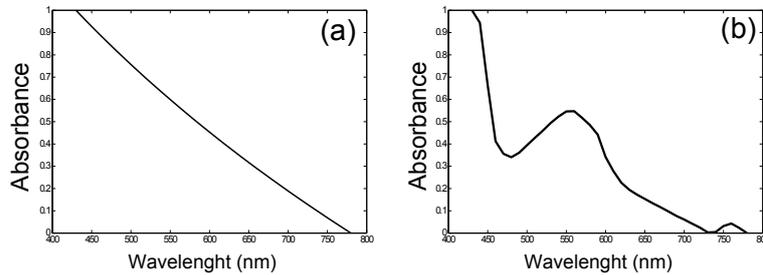


Fig. 3. Average absorption spectrum of: (a) melanin, and (b) hemoglobin.

3.2 Blind Source Separation applied to dermatology

The goal of BSS algorithms is to decompose a given signal in its main sources. Some of their uses are for example the analysis of biomedical signals, telecommunications, and Multi-Hyper-Spectral Imaging. The points to be considered before using linear BSS algorithms correspond to a previous knowledge of the expected results, and to ensure linearity in the physical phenomena to be represented [13].

Since human skin absorption spectra can be represented as a linear combination of components, linear BSS techniques could be useful in separating them. For such purpose, different methods can be found such as Independent Component Analysis (ICA) and Non-negative Matrix Factorization (NMF). Their implementation depends on the kind of materials used for the data acquisition. If color based instruments are used, a spatial approach is more suitable. In the case of Hyper-Spectral systems a spectral approach is more convenient. In spatial approach, images obtained at different spectral bands can be seen as a linear combination of source images. As an example one can mention the work of Tsumura et al, who used ICA methods to separate the spatial distributions of melanin and hemoglobin in skin from a color image [6].

In the case of Hyper-Spectral systems, more than 3 values are obtained for each pixel. As depicted in figure 4, the spectrum at each pixel of the Hyper-Spectral cube can be seen as the linear combination of spectral sources, in this case melanin and hemoglobin.

Although ICA can be useful in skin component decomposition from a spectral point of view [14, 15], here we center the discussion on the use of simple BSS methods such as NMF. The latter have been widely used for the study of geological components [13, 16]. Nevertheless, to our knowledge, NMF are not so much used in the study of human skin components. Due to this lack, here we evaluate the potential use of NMF in the decomposition of HSI.

The implementation of NMF algorithms is supported on the non-negativity of

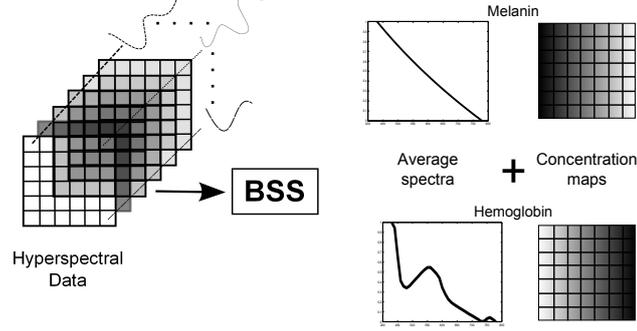


Fig. 4. Spectral approximation in BSS algorithms. The spectrum obtained at each pixel of a Hyper-Spectral cube is observed as the weighted sum of principal components. In our case, these components are considered to be melanin and hemoglobin. BSS algorithms obtain the average spectra of those principal components, together with their respective quantification at each pixel of the Hyper-Spectral cube (concentration maps).

the data to be evaluated. In our case, this constrain is related to the physical meaning of the HSI obtained with ASCLEPIOS system.

From a mathematical point of view, the idea in NMF is to approximate a given $n \times m$ matrix Y , with $Y_{ij} \geq 0$, to the product of two non-negative matrices $W \in R^{n \times r}$ and $H \in R^{r \times m}$ ($Y \approx WH$) [17–19].

The typical way to find those non-negative matrices W and H is minimizing the difference between Y and WH by:

$$f(W, H) \equiv \frac{1}{2} \|Y - WH\|_F^2 \quad (3)$$

where $\|\cdot\|$ is the Forbenious norm.

As it is well known in the domain, a multiplicative update rule has been proposed by Lee and Seung [17] to solve the difference denoted by equation 3. This multiplicative update is giving by:

$$\begin{aligned} H_{a,u} &\leftarrow H_{a,u} \frac{(W^T Y)_{a,u}}{(W^T W H)_{a,u}} \\ W_{i,a} &\leftarrow W_{i,a} \frac{(Y H^T)_{i,a}}{(W H H^T)_{i,a}}. \end{aligned} \quad (4)$$

The function denoted by 3 can be modified in several ways according to the application. So that, penalties can be added in order to enforce sparseness or smoothness in the obtained matrices W and/or H [16, 20]. In our case we used smoothness pelnaty in matrix H . In this way the multiplicative update presented in relation 4 becomes:

$$\begin{aligned}
H_{a,u} &\leftarrow H_{a,u} \frac{(W^T Y)_{a,u} - H_{a,u}}{(W^T W H)_{a,u}} \\
W_{i,a} &\leftarrow W_{i,a} \frac{(Y H^T)_{i,a}}{(W H H^T)_{i,a}}.
\end{aligned} \tag{5}$$

On the scope of this work, the $n \times m$ matrix Y is the bidimensional representation of the Hyper-Spectral cube obtained with ASCLEPIOS system. The number of columns of matrix Y corresponds to the number of spectral bands, 36 in this case. Each column of this matrix represents the spatial distribution of absorption values at the given spectral band.

Each line of matrix H contains the calculated absorption spectra of melanin and hemoglobin respectively. The theoretical spectra are considered to be the ones presented at figure 3. Finally, matrix W presents in each column the estimated quantification of melanin and hemoglobin at each pixel of the ROI.

In the following, matrix Y is considered as the measured Hyper-Spectral cube, and the multiplication $W \times H$ as the estimated one.

4 Results and Analysis

NMF algorithm is used for the study of different skin spectra. Those spectra are considered as the linear combination of melanin and hemoglobin components. The analysis is done in melasma lesions and healthy skin areas from 10 patients. The results are evaluated in three different ways: comparing the measured and estimated Hyper-Spectral cubes, comparing the theoretical and calculated absorption spectra of melanin and hemoglobin, and analysing the quantification of melanin with respect hemoglobin.

In this paper we present in figures 5 and 6, the results obtained for 2 patients. These two patients are considered to be representative of the total evaluated population. For each one of these figures, two sets of four pictures are presented. Set (a) corresponds to the results obtained for the healthy ROI, and set (b) are the results for the melasma lesion. For each set, the upper left image denotes by a white square the ROI of the skin's area under evaluation. The upper right image presents by continuous and dashed line the measured and estimated absorbance spectra in one pixel of the ROI. In most of cases it is possible to observe the fine congruence between both curves. The calculated absorption spectra of melanin and hemoglobin are presented in left down image. In the case of melanin, it is possible to observe how the obtained curve presents a decay of almost 50% in absorbance around 550 nm, which is in coherence with the theoretical result. For hemoglobin, results present the characteristic absorption peaks around 450 and 570 nm, as presented in theoretical spectrum. Finally, the right down image presents in a histogram the normalized concentration of melanin and hemoglobin in the evaluated area. Since the model presents skin absorption spectrum as a linear combination of melanin and hemoglobin, components' concentration values are interpreted to be relative. In this way, ROI corresponding to melasma lesions present for melanin a histogram with a peak at a concentration value

higher than the histogram's peak of hemoglobin. This fact suggests us that for the presented method, melasma lesions present higher concentration of melanin with respect hemoglobin (figures 5b and 6b). In the case of ROI corresponding to healthy skin areas, the histogram's peak of hemoglobin is at a higher (6a) or same (figure 5a) concentration values than the ones from melanin. The results are coherent with the histological cause of melasma: an increased amount of melanin component [21].

In a numerical way, three coefficients of correlation are calculated for each analysed area: a first one corresponds to the degree of correlation between the calculated and the measured Hyper-Spectral cubes; second and third one are calculated between the theoretical and estimated melanin-hemoglobin absorbance spectra. Correlation was evaluated using equation 6 [22].

$$r = \frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^N ((X_i - \bar{X})^2)} \sqrt{\sum_{i=1}^N ((Y_i - \bar{Y})^2)}} \quad (6)$$

where X is the estimated data and Y is the theoretical one. \bar{X} and \bar{Y} are the mean value of the estimated and theoretical data respectively. The results, which are given in table 4, present in most cases a degree of correlation higher than 0.9.

Table 1. Correlation Coefficients

<i>Sample Number</i>	<i>Correlation Result</i>		
	Hyper-Spectral	Melanin	Hemoglobin
ROI 1	0.99	0.93	0.91
ROI 2	0.99	0.95	0.94
ROI 3	0.99	0.94	0.94
ROI 4	0.99	0.98	0.94

5 Conclusions and Further Work

Non-negative Matrix Factorization algorithm has been applied in the study of human skin absorbance spectrum from 10 patients presenting melasma lesions. The data correspond to a Hyper-Spectral cube obtained with ASCLEPIOS system. The use of a multiplicative update approach demonstrated its capacity in estimating in a ROI, two of the principal pigments present in skin: melanin and hemoglobin. Also, the relative quantity or concentration of these two pigments is estimated with the mentioned algorithm. The estimated pigments together with their relative concentrations lead to a stimated Hyper-Spectral cube. In most cases, this cube presents a degree of correlation close to 90% with respect to the

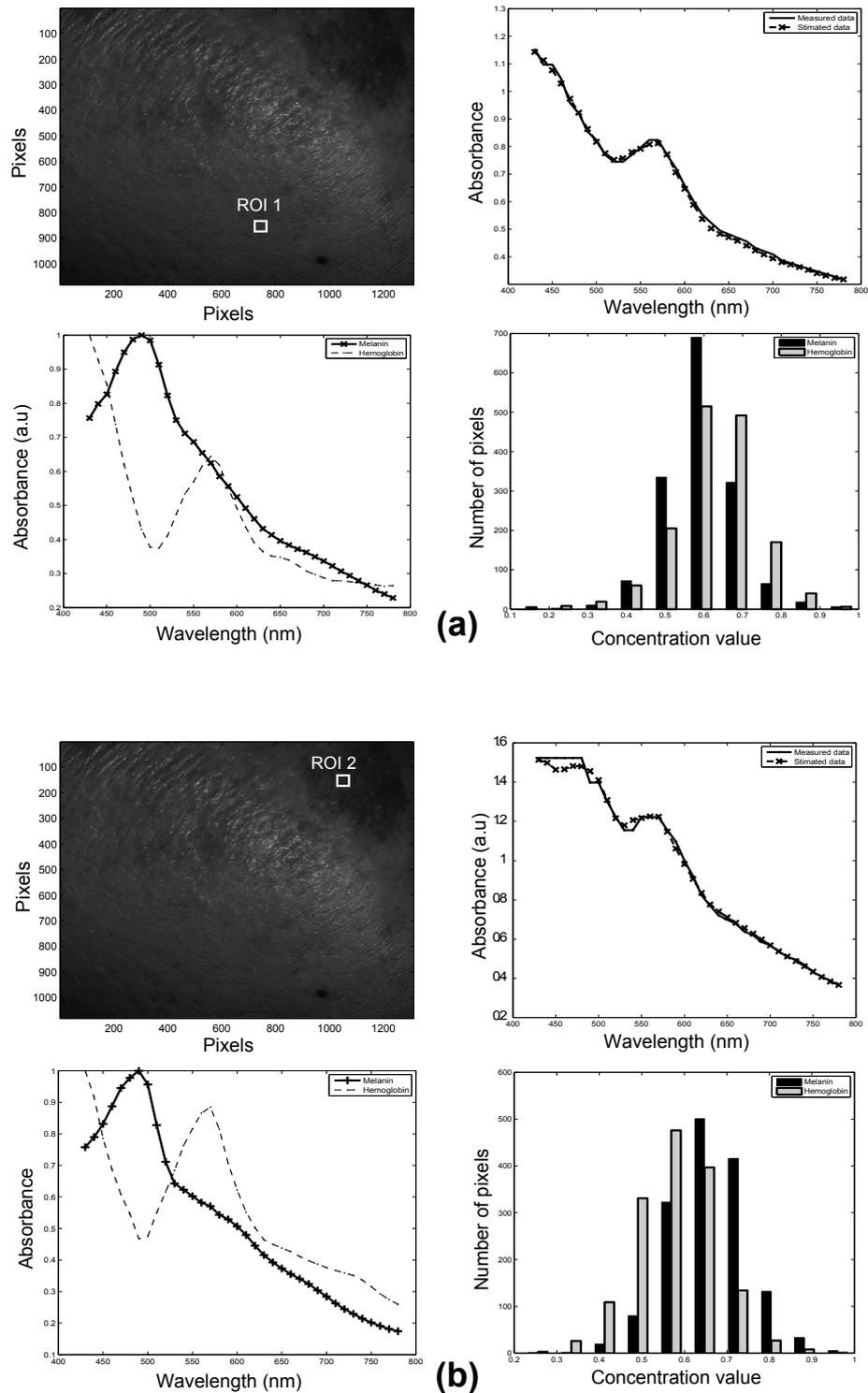


Fig. 5. Results in (a).Healthy skin area, and (b) Melasma lesion. (a-b)*i* Region of interest ROI. (a-b)*ii* Measured and estimated absorption spectra in one pixel of the Hyper-Spectral cube. (a-b)*iii* Estimated melanin and hemoglobin. (a-b)*iv* Histogram of melanin and hemoglobin concentrations.

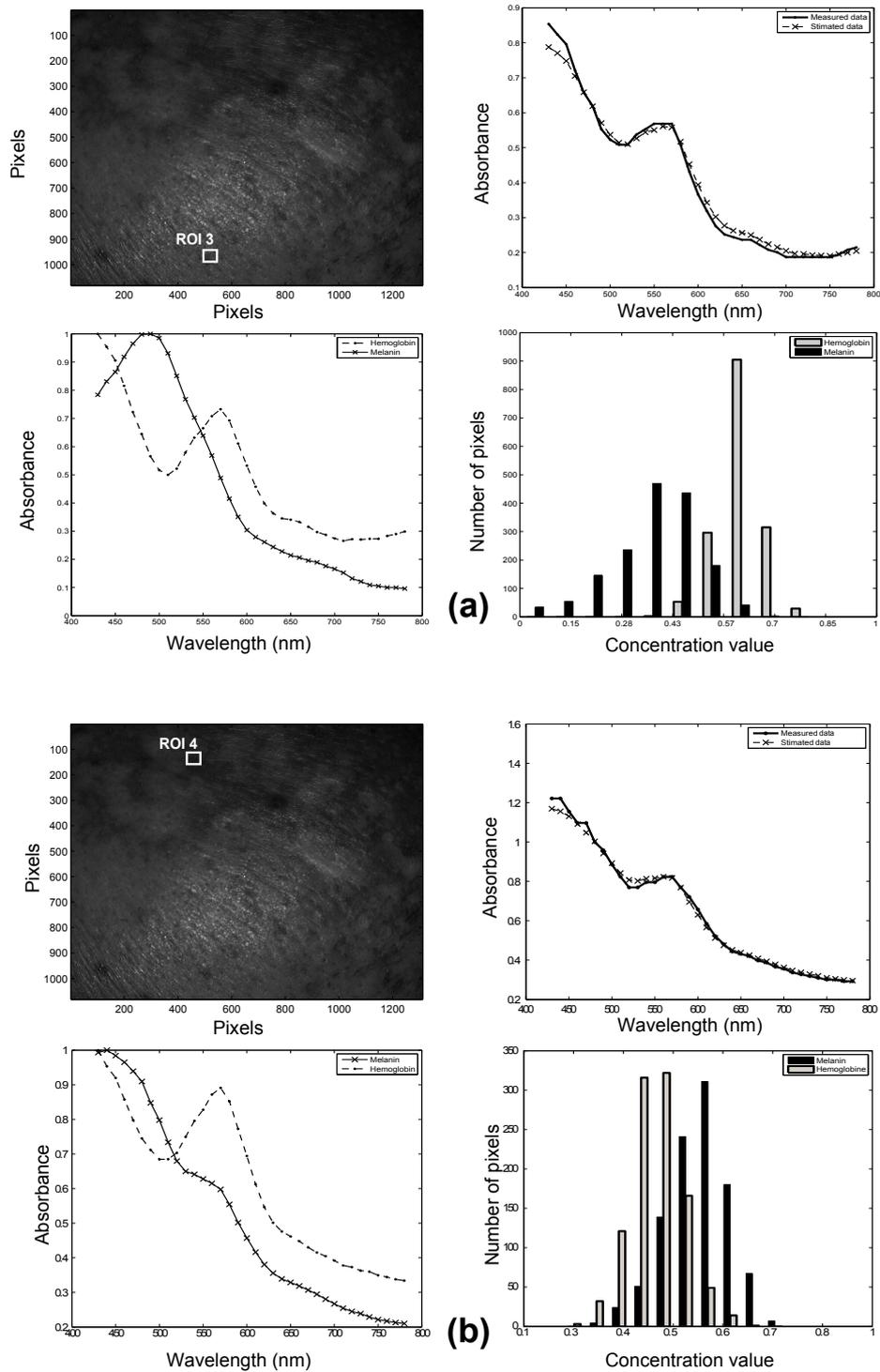


Fig. 6. Results in (a).Healthy skin area, and (b) Melasma lesion. (a-b)*i* Region of interest ROI. (a-b)*ii* Measured and estimated absorption spectra in one pixel of the Hyper-Spectral cube. (a-b)*iii* Estimated melanin and hemoglobin. (a-b)*iv* Histogram of melanin and hemoglobin concentrations.

one obtained from ASCLEPIOS system. In the same way, a degree of correlation higher than 90% is obtained between the estimated and theoretical absorption spectra of melanin and hemoglobin. It has been also shown that melasma lesions present higher concentrations of melanin with respect to hemoglobin. In the case of healthy skin areas, hemoglobin concentration is higher or equal to melanin one. Results agree with the histological cause of melasma: melasma is an hyperpigmentation caused by an increment in melanin [21]. The use of such spectral decomposition of the data obtained from Hyper-Spectral system seems to be a useful tool for the study of human skin illnesses.

Nevertheless, the presented work needs further analysis for the study of the two major components underlying melanin and hemoglobin pigments: eumelanin- pheomelanin, and oxy-deoxy hemoglobin respectively. Human skin phantoms can be also evaluated with the aim to corroborate the effectiveness of NMF algorithms.

Acknowledgments

We would like to thank the cooperation given by Dr. Roshidah baba and Dr. Noorlaily from the Department of Dermatology at Hospital Kuala Lumpur. Also we would like to acknowledge the assistance given by Professor Ahmad Fadzil and Hermawan Nugroho from Universiti Teknologi Petronas; and Dr. Norashikin from the faculty of Medicine and Health Sciences at Universiti Putra Malaysia.

The authors would like also to thank the financial support provided by *Conseil Régional de Bourgogne-France*, and *Fond Européen de Développement Régional (FEDER)-France*.

References

1. T. Igarashi, K. Nishino, and S.K. Nayar. The appearance of human skin. *Columbia Univ., New York, Tech. Rep. CUCS-02405*, (2005).
2. G. Zonios, A. Dimou, I. Bassukas, D. Galaris, A. Tsolakidis, and E. Kaxiras. Melanin absorption spectroscopy: new method for noninvasive skin investigation and melanoma detection. *Journal of biomedical optics*, 13:014017, (2008).
3. T. Stehle, R. Auer, S. Gross, A. Behrens, N. Karssemeijer, and ML Giger. Classification of colon polyps in nbi endoscopy using vascularization features. In *Proc SPIE*, volume 7260, pages 2s1–12, (2009).
4. V.V. Tuchin. Tissue optics: light scattering methods and instruments for medical diagnosis. SPIE-International Society for Optical Engineering, (2007).
5. M. Shimada, Y. Masuda, Y. Yamada, M. Itoh, M. Takahashi, and T. Yatagai. Explanation of human skin color by multiple linear regression analysis based on the modified Lambert-Beer law. *Optical Review*, 7(4):348–352, (2000).
6. N. Tsumura, H. Haneishi, and Y. Miyake. Independent-component analysis of skin color image. *JOSA A*, 16(9):2169–2176, (1999).
7. R. Jolivot, P. Vabres, and F. Marzani. Reconstruction of hyperspectral cutaneous data from an artificial neural network-based multispectral imaging system. *Computerized Medical Imaging and Graphics*, 35(2):85–88, (2011).

8. A. Mansouri, FS Marzani, and P. Gouton. Neural networks in two cascade algorithms for spectral reflectance reconstruction. In *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, volume 2, pages II-718, (2005).
9. R. Jolivot, P. Vabres, and F. Marzani. Validation of a 2d multispectral camera: application to demartology/cosmetology on a population covering five skin phototypes. *SPIE 2011*, (2010).
10. F. Martelli, S. Del Bianco, A. Ismaelli, and G. Zaccanti. Light Propagation through Biological Tissue and Other Diffusive Media, (2010).
11. M. Shimada, Y. Yamada, M. Itoh, and T. Yatagai. Melanin and blood concentration in human skin studied by multiple regression analysis: experiments. *Physics in Medicine and Biology*, 46:2385, (2001).
12. RR Anderson, J. Hu, and JA Parrish. Optical radiation transfer in the human skin and applications in in vivo remittance spectroscopy. In *Bioengineering and the skin: based on the proceedings of the European Society for Dermatological Research symposium*, page 253. Springer, (1981).
13. N. Keshava and J.F. Mustard. Spectral unmixing. *Signal Processing Magazine, IEEE*, 19(1):44-57, (2002).
14. J. Mitra, R. Jolivot, P. Vabres, and FS Marzani. Source separation on hyperspectral cube applied to dermatology. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7624, page 116, (2010).
15. J. Mitra, R. Jolivot, P. Vabres, and F. Marzani. Blind source separation of skin chromophores on a hyperspectral cube. In *ISBS International Symposium*, (2009).
16. V.P. Pauca, J. Piper, and R.J. Plemmons. Nonnegative matrix factorization for spectral data analysis. *Linear Algebra and its Applications*, 416(1):29-47, (2006).
17. D.D. Lee and H.S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788-791, (1999).
18. C.J. Lin. Projected gradient methods for nonnegative matrix factorization. *Neural Computation*, 19(10):2756-2779, (2007).
19. W. Wang and X. Zou. Non-negative matrix factorization based on projected non-linear conjugate gradient algorithm. In *ICA Research Network International Workshop (ICARN 2008)*, pages 5-8, (2008).
20. P. Comon and C. Jutten. *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic Press, (2010).
21. Y. Miyachi. *Therapy of Skin Diseases*. Springer Verlag, (2009).
22. J.V. Stone. *Independent component analysis: a tutorial introduction*. The MIT Press, (2004).