Classification of skin hyper-pigmentation lesions with multi-spectral images

Sylvain Prigent, Xavier Descombes, Didier Zugaj, Laurent Petit, Anne-Sophie Dugaret, Philippe Martel, Josiane Zerubia

Classification of skin hyper-pigmentation lesions with multi-spectral images

Sylvain Prigent, Xavier Descombes, Didier Zugaj , Laurent Petit, Anne-Sophie Dugaret, Philippe Martel, Josiane Zerubia

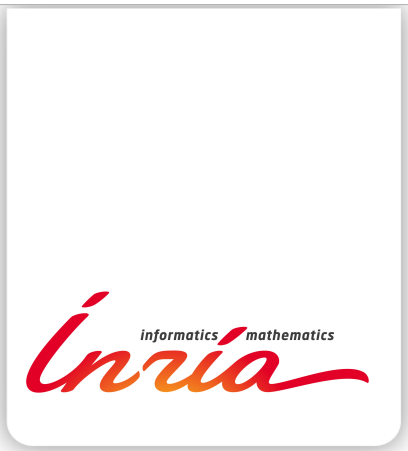
RESEARCH REPORT

N° 8105

October 2012

Project-Teams Morphème and Ayin

ISSN 0249-6399 ISRN INRIA/RR--8105--FR+ENG



Classification of skin hyper-pigmentation lesions with multi-spectral images

Sylvain Prigent*, Xavier Descombes†, Didier Zugaj ‡,
Laurent Petit‡, Anne-Sophie Dugaret‡, Philippe Martel‡,
Josiane Zerubia§

Project-Teams Morphème and Ayin

Research Report n° 8105 — October 2012 — 20 pages

Abstract: According to clinical protocols, skin diseases are quantified by dermatologists throughout a treatment period, and then a statistical test on these measures allows to evaluate a treatment efficacy. The first step of this process it to classify pathological interest areas. This task is challenging due to the high variability of the images in one clinical data set. In this report, we first review algorithms that exist in the literature and adapt them to our problem. Then we choose the more appropriate algorithm to design a classification strategy. Thereby, we propose to use data reduction combined with SVM to do a first classification of the disease. Then we associate the obtained classification map with a segmentation map in an "interactive classification tool" in order to compromise between operator dependency and algorithm robustness.

Key-words: skin, hyper-pigmentation, SVM, data reduction, multi-spectral images

* Morphème / Ayin
† Morphème
‡ Galderma R&D
§ Ayin

RESEARCH CENTRE
SOPHIA ANTIPOLIS – MÉDITERRANÉE

2004 route des Lucioles - BP 93
06902 Sophia Antipolis Cedex

Classification de lésions d'hyper-pigmentation cutanée à partir d'images multi-spectrales.

Résumé : Lors des protocoles cliniques actuels, les maladies de peau sont quantifiées par les dermatologues tout au long d'une période de traitement. Puis un test statistique sur ces diagnostiques permet d'évaluer l'efficacité d'un traitement. Afin d'automatiser un tel processus à l'aide de l'imagerie spectrale, la première étape est d'extraire les zones pathologiques d'intérêt. Cette tache est difficile en raison de la grande variabilité des images dans un ensemble de données cliniques. Dans ce rapport, nous examinons d'abord les algorithmes qui existent dans la littérature et qui peuvent être adaptés à notre problème. Puis, nous choisissons l'algorithme le plus approprié afin de concevoir une stratégie de classification. Ainsi, nous proposons d'utiliser la réduction de données combinée avec un séparateur à vaste marge (SVM) pour faire une première classification de la pathologie. Ensuite, nous associons la carte de classification obtenue avec une carte de segmentation grâce à un outil de "classification interactive" afin de trouver un compromis entre la dépendance à un opérateur et la robustesse de l'algorithme.

Mots-clés : peau, hyper-pigmentation, SVM, réduction de dimension, imagerie multi-spectrale

Contents

1 Introduction

In dermatology, one of the last steps to validate a new treatment is to test it on a sufficient number of patients and assess if its efficacy is statistically better than a standard treatment or a placebo. Nowadays, for such a study, dermatologists have to clinically measure the severity of all the patients along the treatment period. Then, statistical tests such as the t-test [1] are applied on the dermatologist measures to determine if the treatment is statistically efficient. In practice, specialists do not use an absolute measure to quantify the severity evolution of the studied treatment using a statistical test but, on the contrary, a relative measure with respect to a reference treatment or a placebo. Let's call this reference a comparator. For example, in the clinical study we consider in this report, patients are selected to present the same severity degree on the two cheeks. Then, one cheek receives the tested treatment that we will call "active" whereas the second cheek receives the comparator. It is the severity on the active treated area normalised by the comparator treated area that is used for the statistical test.

In order to automatise as much as possible the treatment evaluation process and to make it more objective, we propose to use images associated with a processing methodology. Images we use are multi-spectral. They consist of a collection of monochrome images taken at different wavelengths. With these images, we have both spatial and spectral informations of the skin. Indeed, spectral information is crucial to analyse the skin. Figure 1 shows the theoretical absorptions of melanin and haemoglobin. In this figure, it appears clearly that without the spectral information it is difficult to unmix these two components.

To employ images for skin hyper-pigmentation evaluation, we propose the following scheme. First, the pathological areas have to be segmented on images at the baseline visit. Then, a severity criterion has to be drawn and estimated to quantify the severities of pathological areas at each measurement time. These severity measurements can be used by clinicians to evaluate treatment efficacy or patient evolution. In this report, we focus on the classification part.

In the literature, methods have been proposed to segment skin lesions on color or spectral images [2, 3, 4, 5, 6, 7, 8]. The most popular solution is to transform the RGB representation of a color image to the standard CIE L^*a^*b representation. Then the luminance band L^* or the

ITA index [2, 3] computed with the L^* and b^* bands of the CIE L^*a^*b representation can be thresholded to give a classification map of the disease. Concerning the spectral analysis of the skin, several methods have been proposed to evaluate the quantity of a skin component. This problem can be referred to as unmixing or source separation, applied on the measured spectrum. A theoretical absorbency spectrum of the components of the skin is presented on figure 1. The simplest method to evaluate the pigmentation of a skin area is to select only one spectral value in an area of interest or to compute a ratio between two spectral measurements [3]. Stamatas et al. proposed an algorithm [4] to analyse the spectrum and determine the proportion of the principal skin components which are melanin and haemoglobin. The method consists in modelling the melanin absorbency as linear in the spectrum range of 630-700 nm:

$$A_m(\lambda) = a\lambda + b, \tag{1}$$

where A_m represents melanin absorbance, a and b are linear coefficients and λ is the wavelength. Then, the spectrum can be corrected by subtracting the influence of the estimated melanin:

$$A_c(\lambda) = A(\lambda) - A_m(\lambda). \tag{2}$$

Concentrations of oxy-hemoglobin and deoxy-hemoglobin can be calculated by solving a system given by the Beer-Lamber law applied to two specified wavelengths ($\lambda_1 = 560$ nm , $\lambda_2 = 580$ nm):

$$A_c(\lambda) = [oxy\_Hb] * \epsilon_{oxy\_Hb}(\lambda) + [deoxy\_Hb] * \epsilon_{deoxy\_Hb}(\lambda), \tag{3}$$

where ϵ_{oxy_Hb} and ϵ_{deoxy_Hb} represent extinction coefficients which are tabulated values. Finally, corrected melanin concentration cartography can be extracted by correcting the affine estimation of melanin concentration taking into account the deoxy-hemoglobin influence in the 630-700 nm range:

$$[Melanin]_c = [Melanin] - \left([deoxy\_Hb] * a[\epsilon_{deoxy\_Hb}] + b[\epsilon_{deoxy\_Hb}]\right), \tag{4}$$

where $a[\epsilon_{deoxy\_Hb}]$ and $b[\epsilon_{deoxy\_Hb}]$ are coefficients from the linear model of the deoxy-hemoglobin absorbance in the range 630-700 nm. Considering the problem of pigmentation, with multi-spectral data, this algorithm provides a quantification map of the melanin. Thresholding this map gives a classification map of the pigmentation. Other algorithms, initially proposed within the signal processing community, can provide unmixing on multi-spectral data. In this report, we propose to use a data reduction and classification scheme that we compare to the Stamatas et al. algorithm. We perform data reduction, and classification using Support Vector Machine (SVM). A lot of algorithms have been proposed to reduce the data dimension. The most popular are unmixing ones like PCA [9], Kernel PCA [10], Diffusion maps [11], ICA [12], endmember extraction [13]. These algorithms tend to combine the spectral data into a new lower dimensional space by minimizing a criterion. Another family of algorithms aim to partition the spectrum and combine neighbouring spectral bands. This can be done manually by selecting spectral areas of interest and averaging neighbouring redundant bands or automatically with algorithms like the projection pursuit (PP) [14]. Both types of techniques have been tested in this study.

This report is organized as follows: section 2 describes the chosen algorithm to classify the skin areas of interest and shows results on real data validated by experts. In section 3 we propose a classification strategy that tends to compromise between operator dependency and robustness. Finally, in section 4, we conclude about the proposed method and obtained results.

2 classification

In order to classify pathological areas on multi-spectral images we use the classical scheme which consists in first, reducing the spectrum and then classifying the reduced data with a supervised

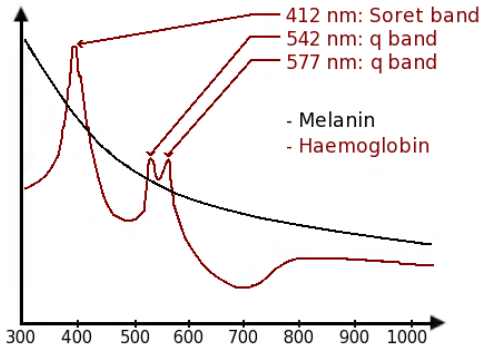


Figure 1: Theoretical absorption of melanin and haemoglobin (Y-axe) depending on the wavelength (X-axe)

algorithm. The interest of data reduction is threefold. First it avoids the Hughes phenomenon [15]. Second it redistributes the data within the reduced data cube in order to avoid redundancy, and then it increases the classes separability. The third interest is linked to the two first ones since lower dimensionality and less redundant information make a data classifier faster. For skin application, we propose to study two different data reduction approaches. The first one called Projection Pursuit (PP), partitions the image spectrum into groups of bands and projects each group to obtain one band per group. The second approach called the independent component analysis (ICA), belongs to the unmixing algorithm familly. ICA tends to redistribute the spectral information to obtain a new data cube where bands have minimum common information from a statistical point of view. The selection of the n most relevant bands allows to reduce the data cube with ICA.

2.1 Data reduction with Projection Pursuit

The projection pursuit (PP) approach was originally proposed in [14]. An adaptation to multi-spectral images or hyper-spectral data has been proposed in [16, 17]. The PP method can be decomposed in two steps.

The first step consists in partitioning the spectrum into groups. Originally each group contained the same number of bands. In [17] a method that partitions the spectrum into various band number groups is proposed. It tends to minimize the variance σ inside each group. Such a partitioning is important in our application since we want to keep an interpretation of the spectral data due to the absorbency of melanin or haemoglobin (see Figure 1). The method in [17] requires the input parameter K (number of groups). In [18] we proposed an alternative method. From an index I such as the Kullback-Leibler divergence, we measure the deviation between consecutive bands. It results in a function F_I which describes the spectrum variations:

$$F_I(k) = I(k-1, k), \qquad k = 2, ..., N_b, \tag{5}$$

with N_b the number of bands. $F_I(k)$ is low if the bands $k-1$ and k are redundant and high if the bands $k-1$ and k exhibits less redundancy. Thresholding $F_I(k)$ with a threshold T [18] allows to select the most significant local maxima of F_I. The location of these maxima gives the spectral bands to be chosen to partition the spectrum. The advantage of the parameter T is that it can be interpreted as the degree of redundancy accepted between each group.

The second step is the estimation of the vectors that allow to project each group into one band. These vectors are estimated by maximising a distance between groups [16]. To estimate

this distance, the most common indices are the Bhattacharyya distance and the Kullback-Leibler divergence [17]. For classification purpose more specific indices can be used. For example, an index based on histogram analysis is proposed in [19]. Nevertheless, such indices require manual parameter specification, and do not improve significantly the classification accuracy. This is why we prefer to use the symmetrical Kullback-Leibler deviation:

$$D_{kb}(a,b) = \frac{H_{kb}(a,b) + H_{kb}(b,a)}{2}, \tag{6}$$

such as

$$H_{kb}(a,b) = \int f_a(x) ln \left(\frac{f_a(x)}{f_b(x)} \right) \mathrm{d}x, \tag{7}$$

where a and b are the projected groups of bands and f the density function of the distribution a (respectively b).

2.2 Data reduction with ICA

ICA is one of the most famous unmixing algorithm. It had been detailed in [12]. For skin hyper-pigmentation application, we expect that a source will be representative of the melanin concentration. Mathematically, ICA models the unmixing problem as linear:

$$X_{i,j} = AS_{i,j} + N_{i,j}, \tag{8}$$

where $X_{i,j}$ is a single pixel at the spatial position (i,j). This pixel $X_{i,j}$ is considered as a linear combination of independent components given by A. $S_{i,j}$ is the vector containing the linear mixing coefficients (i.e. proportion of each component per pixel). The noise is modelled as additive by the vector $N_{i,j}$. The coefficients are quantities that should be positive and the sum of the proportions of all the components have to be equal to 1:

$$\forall k \in [0.N_b], S_{i,j}(k) \geq 0 \tag{9}$$

and

$$\sum_{k=1}^{N} S_{i,j}(k) = 1 \tag{10}$$

The unmixing model of equation (8) shows one indetermination. In fact, each permutation of the columns of matrix A gives a different result. Nevertheless, this indetermination only changes the order of the sources.

In order to solve the unmixing problem for the skin hyper-pigmentation with multi-spectral images, we compare two algorithms. FastICA [20] and JADE [21]. The goal of these algorithms is to filter the noise $N_{i,j}$, and then to find $\hat{A}$ (an estimate of A). The proportion of each component per pixel (i,j) is then obtained by the inverse product:

$$S_{i,j} = PX^d_{i,j}, \tag{11}$$

where $X^d_{i,j}$ represents the denoised $X_{i,j}$, and P the pseudo-inverse matrix of $\hat{A}$.

FastICA performs the estimation of A in a similar way to the projection pursuit. In [20], the author proposes to get the component by maximazing the non-Gaussianity through a neg-entropy index. This index aims at measuring the deviation between the studied signal X and the entropy of a Gaussian distribution with the same correlation matrix. Indeed, each component (i.e. each column of A) is iteratively estimated with the constraint that each component has to

be orthogonal to all the other ones. To compare the data to a Gaussian distribution with the same correlation matrix, whitening is used in order to make the data covariance matrix equal to the identity matrix:

$$E(XX^T) = Id_{N_b} \tag{12}$$

To compute the data whitening, the Principal Component Analysis (PCA) is used:

$$Z_{i,j} = WX_{i,j}, \tag{13}$$

with W the whitening matrix obtained by $W = \sqrt{D}^{-1}E^T$, where D is the diagonalized covariance matrix (XX^T) by PCA and E the eigen-vector matrix. At this step, one can keep only the first components that give enough explained variance. This allows to reduce the data and the noise which is contained in low variance PCA obtained bands.

The JADE (Joint Approximate Diagonalisation of Eigenmatrices) algorithm proposed in [21] consists in minimizing the fourth order cumulants in order to solve the linear ICA problem. The author demonstrates that minimizing cumulants of order 1,2 and 4 (i.e. mean, covariance and fourth order cumulants) is mathematically equivalent to maximize the neg-entropy between components. JADE algorithm starts with a data whitening using PCA. Then, the problem can be written as follows:

$$Z_{i,j} = WAS_{i,j} \tag{14}$$

Where W is the whitening matrix and $Z_{i,j}$ the whitened data. The new mixing matrix WA is a rotation matrix since it links two white data. The last part of the JADE algorithm is to perform this rotation transformation optimizing the criterion:

$$\phi(z) = \sum_{ilkj \neq iikl} (Q^Z_{i,j,k,l})^2 \tag{15}$$

where $Q^z_{i,j,k,l}$ is the fourth order cumulant of Z:

$$Q^Z_{i,j,k,l} = E(Z_iZ_jZ_kZ_l) - E(Z_i)E(Z_j)E(Z_k)E(Z_l) - E(Z_iZ_k)E(Z_jZ_l) - E(Z_iZ_l)E(Z_jZ_k) \tag{16}$$

2.3 SVM classification

SVM [22] is the most wildly used method for classification on multi-spectral images. SVM classify a data set into two classes by computing a separating hyperplane. A class is associated to each pixel depending on its position w.r.t the hyperplane. This algorithm is widely used in various domains to classify images (see for instance [10, 23, 24, 25, 26]). The condition for a good classification is to find the optimal hyperplane that separates at best two classes defined by training pixels. To do this, we seek to maximize the margin between the separating hyperplane and the points of the two training sets. If we denote the margin $\frac{2}{||\vec{\omega}||}$, the hyperplane equation is:

$$\vec{\omega}.\vec{x} + b = 0. \tag{17}$$

To estimate the hyperplane separator (i.e. estimate $\vec{\omega}$ and b), a label $y = -1$ or $y = +1$ is associated to each training pixel depending on its class. Then the hyperplane estimation problem reduces to minimize $\frac{||\vec{\omega}||}{2}$ such that:

$$\begin{cases} \vec{\omega}.\vec{x} + b \geq +1 \text{ if } y_i = +1 \\ \vec{\omega}.\vec{x} + b \leq -1 \text{ if } y_i = -1 \end{cases} \tag{18}$$

It is not straightforward to minimize the expression of equation (18). Nevertheless, the transformation into the dual space with Lagrange multipliers makes it a quadratic problem easy to solve:

$$\max_{\lambda} W(\lambda) = -\frac{1}{2}\sum_{i=1}^{N_p}\lambda_i\lambda_j\vec{x_i}.\vec{x_j} + \sum_{i=1}^{N_p}\lambda_i. \tag{19}$$

with

$$\sum_{i=1}^{N_p}\lambda_i y_i = 0,\ 0 \leq \lambda_i, \forall i \in [1, N] \tag{20}$$

where N is the number of training pixels.

In order to classify non linear problem, a kernel is introduced in the equation (20):

$$\max_{\lambda} W(\lambda) = -\frac{1}{2}\sum_{i=1}^{N_p}\lambda_i\lambda_j K(\vec{x_i}, \vec{x_j}) + \sum_{i=1}^{N_p}\lambda_i \tag{21}$$

We choose the Gaussian kernel because, in our experiments [27], it outperforms linear, polynomial or angular kernels [23].

2.4 Face volume compensation

An example of classification by PP and SVM is shown in Figure 4(b), and a reference manual classification is shown in Figure 2(f). As one can see, the classification is quite accurate in the flat areas whereas there is no detection of the pathology in areas corrupted by the volumes of the face. Concerning ICA, experimentations show that a single component is sufficiently representative of the melanin concentration in the skin. This is the reason why we perform a classification in a single component. This classification can be done by SVM. Nevertheless, we observe that thresholding this component provides accurate results. In Figure 5(b) the classification is accurate in the flat areas, but there are a lot of false detections in areas concerned by shading.

Whatever the classification method we use, it needs an additional filter to compensate the volumes of the face. The literature proposes several minimisation methods to solve the shape from shading inverse problem [28, 29, 30] introduced by Horn in [31]. Most of these methods tend to solve the eikonal equation. For image processing, with the Lambertian surface hypothesis the eikonal equation for a pixel at the position x, y is:

$$I(x, y) = \frac{I_0(x, y)}{\sqrt{p^2(x, y) + q^2(x, y) + 1}} \tag{22}$$

where p^2 and q^2 are the derivatives of the Lambertian surface in x and y directions respectively, I_0 the albedo and I the image intensity.

In our application, we need to compensate the shading on multi-spectral data. Then, the idea is to analyse the spectrum to find which spectral area is better adapted to estimate the shading variations. Figure 2 shows the five bands obtained after the reduction of an 18 spectral band data cube. As one can see the last band depicted on 2(e) mainly contains the shading variations. In fact, it is justified by Figure 1. For near infra-red wavelengths (900nm), melanin and haemoglobin absorptions are very low. So the main skin components do not react to this wavelength and then the obtained image contains mainly the shading. If we assume the skin in the near infra-red image to react as a Lambertian surface, equation (22) gives that we can compensate the spectral band shading by the near infra-red image. Figure 7(a) shows the second band of the reduced data cube. Figure 3(b) shows the result obtained when dividing the second

band of the reduced cube by the near infra-red band of the reduced cube. As one can see, shading on the border of the image are well compensated. They became white in the compensated image and are now not interfering with the dark gray of the pathological areas. But the compensation is not efficient in the areas where the shading is caused by the volume of the face. If we perform such a compensation and compute a classification we obtain false classification (see Figure 4(c) and Figure 5(c))We thus have rejected the Lambertian assumption.

Consider that the light illuminating the skin is partly reflected by the surface and partly penetrates the skin before being reflected. The signal reflected by the skin surface can be used as a model of the shading. The interference theory states that in this case, the signal amplitudes exhibit the following relationship:

$$A_0^2 = A_{01}^2 + A_{02}^2 + 2A_{01}A_{02}cos(\phi_1 - \phi_2), \tag{23}$$

where A_0 is the amplitude of the global signal, A_{01} and A_{02} the amplitudes of the two reflected light rays, and $cos(\phi_1 - \phi_2)$ the phase angle between the two light rays. As a first approximation, if we neglect the phasing angle, the relation between the shading and the skin absorption information is additive. This is why we propose to compensate the shading by subtracting the near infra-red band to all the others bands:

$$I_\lambda^c = I_\lambda + z[max(I_{IR}) - I_{IR}] \tag{24}$$

with

$$z = \frac{max(I_\lambda) - min(I_\lambda)}{max(I_{IR}) - min(I_{IR})} \tag{25}$$

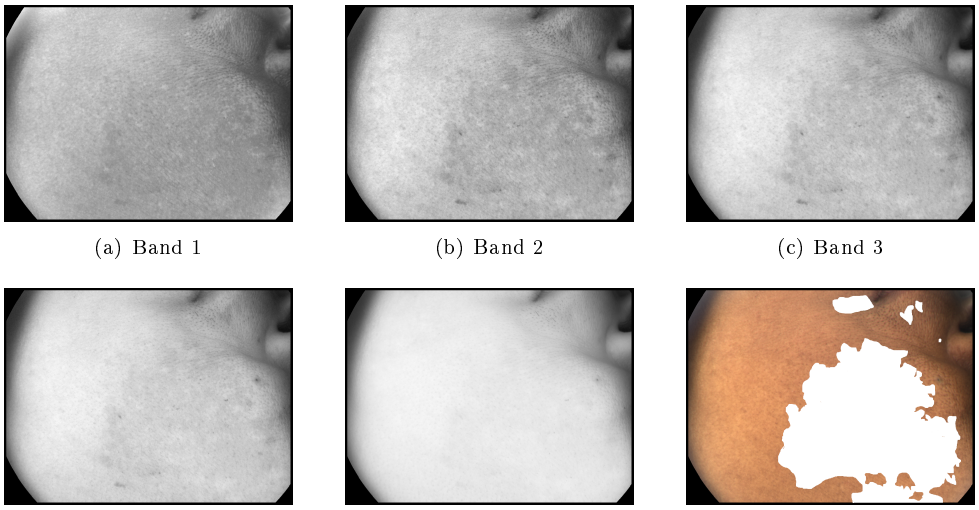
where I_λ^c and I_λ denote respectively the compensate band and the band at the wavelength λ, and I_{IR} the near infra-red band. The factor z allows to take into account how deep the light penetrates the skin depending on the wavelength. So the interference term that we neglect is partially taken into account in this z term. Figure 3(c) shows the effect of the subtraction based compensation on a spectral band. With this method, both shadows on the border of the image and the face volume shading are compensated. Then, when we apply this method in the PP and SVM scheme or on the ICA and thresholding classification, we have an accurate detection both on the flat areas and on areas where artefacts are caused by the volume shading.

2.5 Experimental results

The proposed classification methods have been tested on 96 images. These are multi-spectral images that contain 18 spectral bands of 1000*1000 pixels. On this data set the PP plus SVM method and the ICA based method are competitive and give equivalent quality results. The main difference between these two methods is the operator dependency. In fact SVM ask the operator to select training pixels whereas the ICA thresholding ask the operator to choose the value of the threshold. The second method is easier for an operator to use but include subjectivity.

To evaluate the classification on these 96 images, we partition them into 3 groups. The first one contains 27 images that are qualified as "easy" in the sense that these images contain mainly skin in the flat area of the cheek. The second group qualified as "intermediate" contains more artefacts like eye, hair or nose. This group represents 37 of the 96 images. The third group contains 31 images with a lot of artefacts.

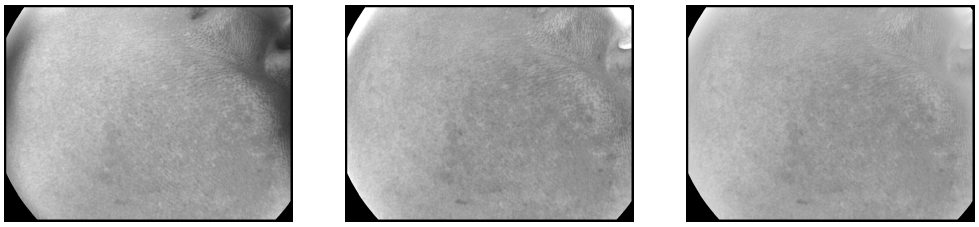
As the Stamatas et al. algorithm and the ICA based methods are equivalent from an operator input point of view, we propose to compare them first. For both method, we estimate the threshold that maximizes the correspondence with the dermatologist analysis in term of surface. Figure 6(b) shows the correlation curves between the dermatologist estimation of the pathological



(a) Band 1 (b) Band 2 (c) Band 3

(d) Band 4 (e) Band 5 (f) Manual classification

Figure 2: Spectral bands of the reduced data cube by projection pursuit (a-e), and a manual classification of the hyper-pigmentation (f).



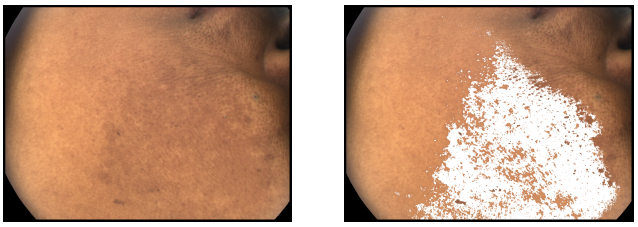
(a) 2^{nd} band on the reduced data cube (b) Band after division shading compensation (c) Band after subtraction shading compensation

Figure 3: Influence of the shading compensation on the second band of the reduced data cube

area and the algorithm calculation on the 27 images of the first group. The dermatologist scale is discrete from 0 (no pathology) to 6 (100% of the area is pathological). The perfect correlation is represented by the red diagonal line. Figure 6(b) shows that the ICA based method outperforms the Stamatas et al. algorithm. It has to be notice that the measured correlation in this case is artificially heigh since the dermatologist analysis is taken into account to measure the surface.

Now, we compare the surface evaluation of ICA based method and SVM on the whole set (ie. 96 images). Table 1 shows the percentage of correspondence between the dermatologist area estimation in a scale from 0 to 6 and the algorithm surface estimation in percentage for both the ICA based method and PP-SVM. It is natural in the table 1 that the percentages for the ICA based method are better since this method use the dermatologist clinical analysis to calculate the pathological area size. Although the correspondence between obtained classifications and dermatologist measurement is quite satisfactory, some over-estimation of the pathological area appears mainly for the groups 2 and 3 where the correspondence percentage decrease.

Finally, experiments show that the ICA based method and the PP-SVM suit well our problem. However, these techniques are not robust enough to the data variability appearing in a clinical



(a) Reconstructed color image

(b) SVM Segmentation

(c) SVM segmentation with shading compensation by the division based method

(d) SVM segmentation with shading compensation by the substraction based method

Figure 4: Classification results with projection pursuit depending on the shading compensation method

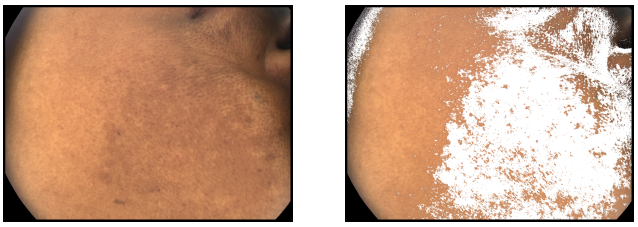
Table 1: Correspondence of surface measurement between dermatologist and algorithms

	ICA	**PP-SVM**
Group 1	81.5%	63%
Group 2	75.7%	62.2%
Group 3	62.5%	52%
Total	73%	58.7%

study.

3 Generalisation to a complete study

The classification scheme proposed in the previous section is not usable on a full study. First, the algorithms are not robust enough to image set variability. Second, for the ICA based method, the operator has to select a source and a threshold for this source. For the SVM, the operator has to select training pixels on each image. This is not compatible with a practical use on a whole study, containing about 100 subjects. In this section, we propose a protocol that allows both to automatize the algorithms as much as possible and improves the robustness of the classification. We focus on the SVM based method since it is the one that seems to be the easier to automatize. To this end, we propose a scheme to train a SVM once for all the images, and then, we add a interactive classification step to avoid possible over or miss-classification.



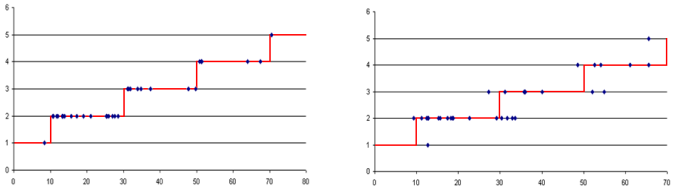
(a) Reconstructed color image

(b) ICA based segmentation

(c) ICA based segmentation with shading compensation by the division based method

(d) ICA based segmentation with shading compensation by the substraction based method

Figure 5: Classification results with the ICA based method depending on the shading compensation method



(a) (b)

Figure 6: Correlation of the ICA based method (left) and the Stamatas algorithm (right), with the dermatologist quantification. x-axis contains the algorithm estimated surface, and the y-axis the dermatologist surface estimation.

3.1 Global training

To make a global training for all the images, we first need to normalize them. Images have more or less shading and non skin elements. Moreover, natural skin darkness is different from one patient to another. Let's detail the two steps:

3.1.1 Fourier filtering

In order to remove as many artefacts as possible in all the images, we propose to add a Fourier filtering step between the projection pursuit and the face volume compensation. Indeed, in our images, we observe that the main undesired parts are in the low frequencies areas. The eyes, the nose, or the lips are inside the face volume (in opposition to the flat area of the cheek), and the face volume represents low frequency variations in our images. Another remaining artefact not

completely corrected by the face volume compensation step is the lighting inhomogeneity which is also contained in low frequency in our images. To make the Fourier filtering, let's consider I_R a reduced image by projection pursuit. We first apply a Hamming 2D window to each band $I_R(b)$ of I_R to avoid oscillations [32]:

$$\forall b \in [0, n_b], HI_R(b) = I_R(b).HW, \tag{26}$$

with n_b the total number of bands of I_R, and

$$HW(i,j) = 0.54 - 0.46 * \cos\left(\pi * \sqrt{\left(\frac{i - \frac{n_l}{2}}{n_l}\right)^2 + \left(\frac{i - \frac{n_c}{2}}{n_c}\right)^2}\right) \tag{27}$$

where n_l and n_c are respectively the number of lines and columns of I_R. Then, we apply a 2D Fourier transform to each band b of HI_R [32]:

$$F(HI_R)(u,v,b) = \frac{1}{nl * nc} \sum_{x=1}^{nl} \sum_{y=1}^{nc} HI_R(x,y,b) e^{-2i\pi\left(\frac{ux}{nl} + \frac{vy}{nc}\right)} \tag{28}$$

Finally we apply a low frequency mask M in the Fourier space and compute the Fourier inverse transform to obtain I_{R_F} [32]:

$$\forall b \in [0, nb], I_{R_F}(b) = F^{-1}\left(F(HI_R)(b).M\right), \tag{29}$$

with

$$F^{-1}(I)(x,y) = \frac{1}{nl * nc} \sum_{u=1}^{nl} \sum_{v=1}^{nc} F_I(u,v) e^{2i\pi\left(\frac{ux}{nl} + \frac{vy}{nc}\right)} \tag{30}$$

where I represents a monochrome image and F_I its Fourier transform. An example of the filtered image is shown in Figure 7(c). As one can see, the surrounding of the eye and the nose are better filtered than with the face volume compensation method alone shown in Figure 7(b).

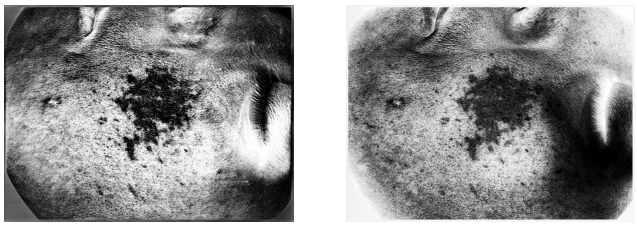
3.1.2 Global training

In this section, we propose to use SVM with a single training for the whole studies. SVM is applied on the filtered images obtained with the method described in section 3.1.1. To this end, each band has to be radiometrically normalized. In fact, the SVM will define a single separator for the whole data set. To perform this normalization, we use a histogram specification [32] that makes each band of I_{R_F} having a flat histogram. An example of a filtered and normalized band is shown in Figure 7(d). The image 7(e) shows how important the Fourier filtering is. If there is not such a filtering, the histogram specification doesn't eliminate enough shading. When images are filtered and normalized, we train the SVM on selected training images. These images have to be selected carefully. The images must not be too atypical otherwise the SVM will not be able to generalize to the whole study. Then, on each chosen image, an operator specifies healthy and pathological areas. Finally, all the marked pixels are concatenated and used to train a SVM. All the images can then be classified with this SVM.

(a) Initial band

(b) Band with infra-red compensation

(c) Band with infra-red compensation and Fourier filtering



(d) Band with infra-red compensation, Fourier filtering and histogram specification

(e) Band with infra-red compensation and histogram specification

Figure 7: Influence of filtering on one band.

3.2 Interactive classification tool

The classification proposed in the previous section can still fail for non typical images. If there are too many artefacts in an image or if the contrast between healthy and pathological area is to low, the trained separator will not be adapted to classify correctly the image. Furthermore, the SVM classifier has a non zero over or miss-classification rate. In our application, we need to be sure that the classified areas are pathological to correctly evaluate a treatment. This is why we propose a protocol that we call "interactive classification" to allow an operator to easily make corrections w.r.t the classification.

There are two types of corrections that can be chosen by the operator: remove over-segmentation and add miss-classification.

To enable the operator to remove over-segmentation, we perform a connected component analysis [32] on the SVM classification map. Then the operator can click on a component he/she wants to add or remove from the final classification.

If the SVM miss-classifies an area, a segmentation map is needed to allow the operator to add a segment to the classification. For our application we perform the segmentation on I_{R_F}. In the literature a lot of segmentation algorithms have been proposed to deal with multi-spectral images [33]. The two main methods are split and merge segmentation or unsupervised classification associated with connected component analysis [32]. To reduce the computation time as much as possible we use the second approach. We use the Gaussian mixture classification algorithm as proposed in [33]. We first perform a classification on I_{R_F} with a Gaussian mixture model. Each obtained component is partitioned with connected component analysis. The final segmentation map is obtained by concatenation of the segments of each Gaussian component on a single map. The global classification scheme is drown in Figure 8. Figure 9 and Figure 10 show screen-shots of the classification GUI we developed. To maximize the chance to segment the area of interest,

we perform twice the segmentation: one with a Fourier high-pass filter that cut 20% of the low frequencies and one with a Fourier high-pass filter that cut 40% of the low frequencies. The operator can click on the segment of these two maps to add it in the final classification, or chose the SVM classification. On the Figure 9 it is the SVM classification that have been chosen and drawn in the bottom right quadrant. If the SVM do over-classification, the operator can click on the regions he/she wants to keep or remove. On the Figure 10 segments from the segmentation maps in the two top quadrants have been selected. They appear in white on the segmentations maps. The final classification is shown in the bottom right quadrant. In the worst cases where both SVM classification and the segmentation fail to detect the interest area, the operator can select it by hand.

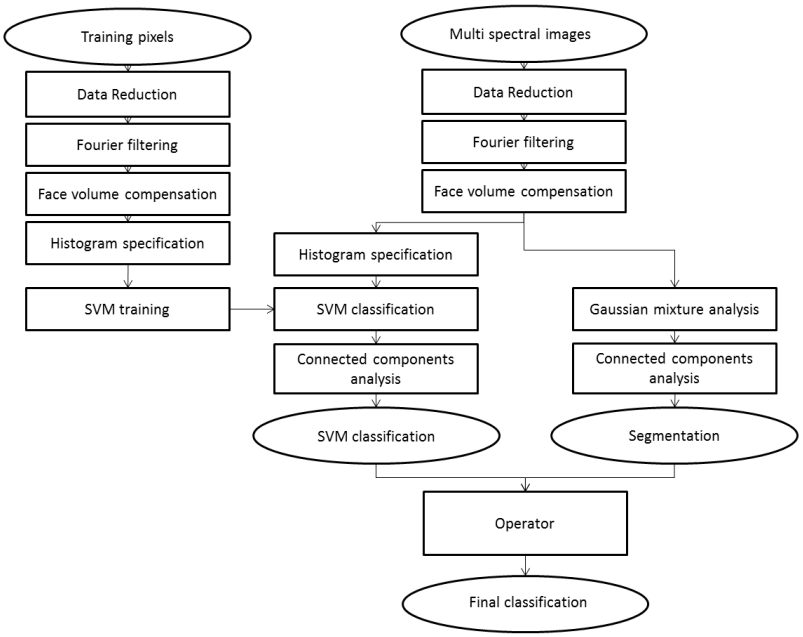


Figure 8: Final classification scheme.

3.3 Obtained results

We use the proposed classification approach on two clinical studies. The first one contains 48 patients, with two images per patient. The second one contains 44 patients, with two images per patient. We trained the SVM only once for both studies with three patients selected from these studies. These three patients (see Figure 11) were selected as follow. The first patient 11(a) is an "easy" image in the sense that it has few shading effect and the pathology contrast is intermediate. The two others patients are selected because they contain shading effect. The patient 11(b) has a contrasted located pathological area with a shading area around the eye. The patient 11(c) has a diffuse pathological area with shading areas around the nose and the chin. Filtered bands and training areas of these three images are shown in Figures 11(d), 11(e) and 11(f)

We launched the SVM classification and the segmentation on both studies keeping the same parameters. Then an operator did a interactive classification. During the interactive classification, 91% of the images was classified with the SVM and/or the segmentation. The 9% images

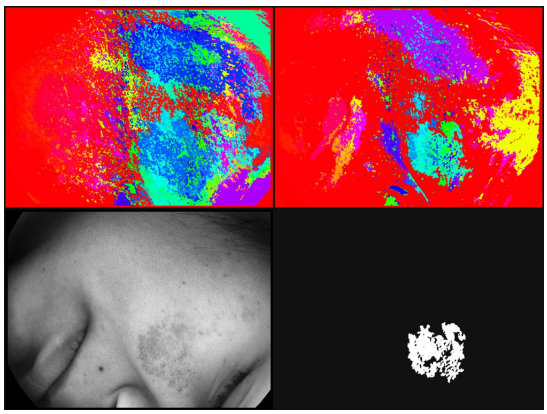


Figure 9: Screenshot of the software that allows to do the interactive classification, first example. Top left: segmentation obtained with the Fourier filter that cut 20% of the low frequencies, top right: segmentation obtained with the Fourier filter that cut 40% of the low frequencies, bottom left one band of the original image, bottom right SVM chosen component

left were classified by hand. On the 91% of images classified by the proposed method, the operator need to use the segmentation map in about 40% of the images, and in about 30% of the cases either the SVM classification or the segmentation map can be use. The interactive classification took one hour for the operator to classify about 100 images. If the operator has to do it manually it takes more time with a less precise classification. To quantify the accuracy of the classification, we compare it with a clinical measurement of the pathological area made by a dermatologist with a scale from 0 to 6. We compare this measurement with the percentage of the images area classified as pathological. As the dermatologist does not consider the holes inside pathological areas, we fills holes in the binary masks. Correspondence curves between this dermatologist and the algorithm are shown in Figure 12. For the first study the correlation between clinical and algorithm classifications is 76.3% and 71.3% for the second. The obtained correlations outperform the SVM algorithm without the proposed procedure with global training and the interactive selection tool. In fact, for the first study we got 58.7% of correlation between the SVM algorithm (see table 1) and the clinical analysis. For the second study (see Figure 13) we got only 45.1% of correlation between the SVM algorithm and the clinical analysis. Let add two remarks about the presented results. An important point is that the "interactive classification tool" makes the operator work at the output of the algorithm. This insure a minimisation of over and miss classification not possible with the SVM algorithm alone. A second remark is that, even though the correlation curves in Figure 6 seems better than the one in Figure 12, the classification in the second case fits better the pathology. This difference is due to the method used to compute the threshold to obtain the results in Figure 6 knowing the dermatologist analysis.

4 Conclusion

In this report, we have proposed a new scheme that allows to classify hyper-pigmentation lesions of skin using multi-spectral images. The obtained results on two clinical studies tend to show that the method is robust. The fact that the operator interacts at the last step of the classification can be seen as a positive thing. It allows both to make a miss or over-classification control and avoid the classification to be based only on machine learning techniques which are often contested

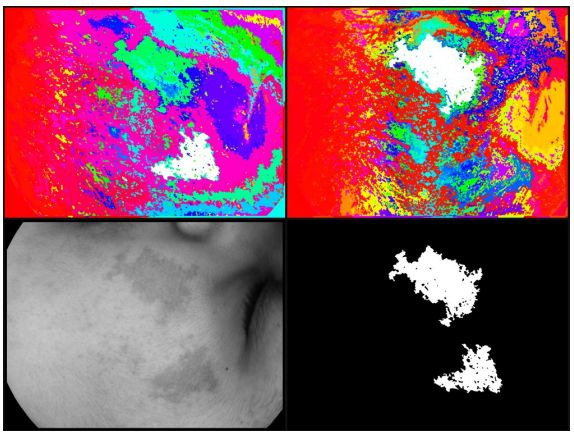


Figure 10: Screenshot of the software that allows to do the interactive classification, first example. Top left: segmentation obtained with the Fourier filter that cut 20% of the low frequencies, top right: segmentation obtained with the Fourier filter that cut 40% of the low frequencies, bottom left one band of the original image, bottom right final classification. The final classification has been obtained by clicking two regions in the segmentations map (regions in white on the top quadrants).

in medical applications. From the obtain classification, severity criteria can be designed to characterize the disease and evaluate a patient evolution or a treatment efficacy. For example, we designed a contrast criterion in [34] and are currently working on spatial and homogeneity criteria.

Acknowledgment

S. Prigent, X. Descombes and J. Zerubia would like to thank Galderma R&D for partial funding this research work and for providing the data. Material presented in this report has been partially published earlier in [27, 18].

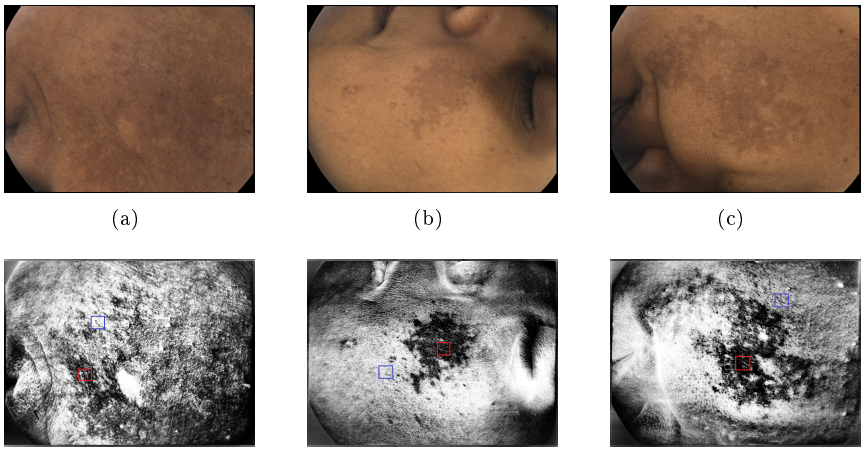
References

[1] A. M. Mood, F. A. Graybill, and D. C. Boes. *Introduction to the theory of statistics.* McGraw-Hill, 1974.

[2] I.L. Weatherall and B.D. Coombs. Skin color measurements in terms of CIELab color space value. *Journal of Investigative Dermatology*, 99:468–473, 1992.

[3] G. N. Stamatas, B. Z. Zmudzka, N. Kollias, and J. Z. Beer. Non-invasive measurements of skin pigmentation in situ. *Pigment cell res*, 17:618–626, 2004.

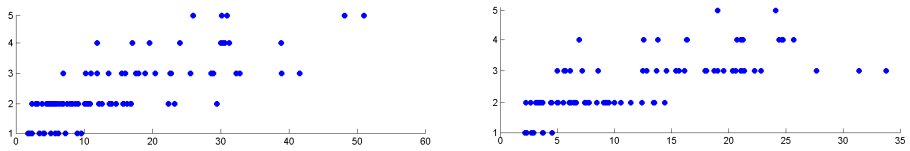
[4] G. N. Stamatas, B. Z. Zmudzka, N. Kollias, and J. Z. Beer. In vivo measurement of skin erythema and pigmentation: new means of implementation of diffuse reflectance spectroscopy with a commercial instrument. *British Journal of Dermatology*, 159:683–690, 2008.



(a) (b) (c)

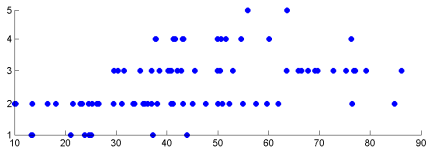
(d) (e) (f)

Figure 11: Images used for the global training. The first row shows color reconstructions of the initial images. The second row shows a filtered band with the location of the training areas. Red rectangles show training pathological areas, and blue rectangles show training healthy areas.



(a) (b)

Figure 12: Correlation between the clinician and algorithm measurements on (a) first study containing 48 patients and (b) the second study containing 44 patients.



(a)

Figure 13: Correlation between the clinician and surface measurements based on a SVM algorithm trained on each images of the second study containing 44 patients.

[5] C. Grana, G. Pellacani, R. Cucchiara, and S. Seidenari. A new algorithm for border description of polarized light surface microscopic images of pigmented skin lesions. *IEEE Transactions on Medical Imaging*, 22:959–964, 2003.

[6] M. Burroni, L. Alparone, and F. Argenti. Comments on "a new algorithm for border description of polarized light surface microscopic images of pigmented skin lesions". *IEEE Transactions on Medical Imaging*, 25:1655–1656, 2006.

[7] B.F. Jones. A reappraisal of the use of infrared thermal image analysis in medicine. *IEEE*

Transactions on Medical Imaging, 17:1019–1027, 1998.

[8] P. Schmid. Segmentation of digitized dermatoscopic images by two-dimensional color clustering. *IEEE Transactions on Medical Imaging*, 18:164–171, 1999.

[9] I. T. Jolliffe. *Principal Component Analysis*. Springer-Verlag, 1986.

[10] M. Fauvel. Spectral and spatial methods for the classification of urban remote sensing data. In *PhD Thesis (in English)*. Institut National Polytechnique de Grenoble. France, 2007.

[11] S. Lafon and A.B. Lee. Diffusion maps and coarse-graining: a unified framework for dimensionality reduction, graph partitioning, and data set parameterization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28:1393–1403, 2006.

[12] P. Comon. Independent Component Analysis, a new concept ? *Signal Processing, Elsevier*, 36(3):287–314, April 1994. Special issue on Higher-Order Statistics.

[13] J. Bioucas-Dias. A variable splitting augmented lagrangian approach to linear spectral unmixing. In *IEEE Workshop on Hyperspectral Image and Signal Processing (Whispers)*, Grenoble, France, 2009.

[14] S. Mallat and Z. Zhang. Matching pursuit with time-frequency dictionaries. *IEEE Trans. on Signal Processing*, 41:3397–3415, 1993.

[15] G. F. Hughes. On the mean accuracy of statistical pattern recognizers. *IEEE Trans. on Information Theory*, 14:55–63, 1968.

[16] L.O. Jimenez and D.A Landgrebe. Hyperspectral data analysis and supervised feature reduction via projection pursuit. *IEEE Trans. on Geoscience and Remote Sensing*, 37:2653–2667, 1999.

[17] G. Rellier, X. Descombes, F. Falzon, and J. Zerubia. Texture feature analysis using a Gauss-Markov model in hyperspectral image classification. *IEEE Trans. on Geoscience and Remote Sensing*, 42:1543–1551, 2004.

[18] S. Prigent, X. Descombes, D. Zugaj, and J. Zerubia. Spectral analysis and unsupervised SVM classification for skin hyper-pigmentation classification. In *IEEE Workshop on Hyperspectral Image and Signal Processing (Whispers)*, Reykjavik, Iceland, June 2010.

[19] D. Gomez, L. K. Harder, B. Ersbøll, and J. M. Carstensen. Precise acquisition and unsupervised segmentation of multi-spectral images. *Computer Vision and Image Understanding*, 106:183–193, 2007.

[20] A. Hyvarinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans. on Neural Networks*, 10:626–634, 1999.

[21] J.F. Cardoso. High-order contrasts for independent component analysis. *Neural Computation*, 11:157–192, 1999.

[22] V. Vapnik. Statistical learning theory. *John Wiley and sons, inc.*, 1998.

[23] O. Zammit, X. Descombes, and J. Zerubia. Burnt area mapping using support vector machines. *International Conference on Forest Fire Research, Figueira da Foz, Portugal*, 2006.

[24] I. El-Naqa, Y. Yang, Wernick M.N., P.N. Galatsanos, and R.M. Nishikawa. A support vector machine approach for detection of microcalcifications. *Trans. on Medical Imaging*, 21:1552–1563, 2002.

[25] P. Padilla, M. López, J. M. Górriz, J. Ramírez, D. Salas-González, and Álvarez I. Nmf-svm based cad tool applied to functional brain images for the diagnosis of alzheimer's disease. *Trans. on Medical Imaging*, 31:207–215, 2012.

[26] J.L. Rojo-Álvarez, J. Bermejo, V. M. Juárez-Caballero, R. Yotti, C. Cortina, M. A. García-Fernández, and J. C. Antoranz. Support vector analysis of color-doppler images: A new approach for estimating indices of left ventricular function. *Trans. on Medical Imaging*, 25:1037–1043, 2006.

[27] S. Prigent, X. Descombes, D. Zugaj, P. Martel, and J. Zerubia. Multi-spectral image analysis for skin pigmentation classification. In *Proc. IEEE International Conference on Image Processing (ICIP)*, Hong-Kong, China, September 2010.

[28] Q. Zheng and Chellappa R. Estimation of illuminant direction, albedo, and shape from shading. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:680–702, 1991.

[29] M. Bichsel and A. P. Pentland. A simple algorithm for shape from shading. *IEEE Proceedings of Computer Vision and Pattern Recognition*, pages 459–465, 1992.

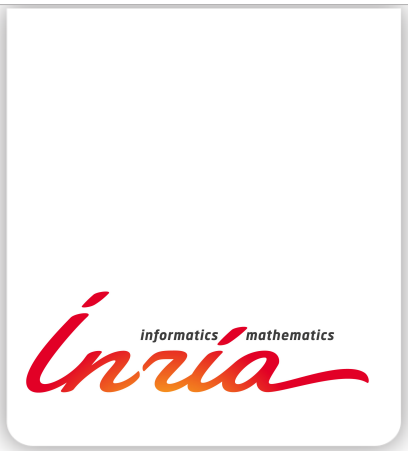
[30] A. Crouzil, X. Descombes, and J.D. Durou. A multiresolution approach for shape from shading coupling deterministic and stochastic optimization. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(11):1416–1421, November 2003.

[31] B. K. P. Horn. Shape from shading: A method for obtaining the shape of a smooth opaque object from one view. In *PhD thesis*. MIT, 1970.

[32] R.C. Gonzalez and R.E. Woods. *Digital image processing*. Prentice Hall, 2008.

[33] Y. Tarabalka, J. A. Benediktsson, and J. Chanussot. Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques. *IEEE Trans. Geos. and Remote Sens.*, 47(8):2973–2987, 2009.

[34] S. Prigent, D. Zugaj, X. Descombes, P. Martel, and J. Zerubia. Estimation of an optimal spectral band combination to evaluate skin disease treatment efficiency using multi-spectral images. In *Proc. IEEE International Conference on Image Processing (ICIP)*, Brussels, Belgium, September 2011.



RESEARCH CENTRE
SOPHIA ANTIPOLIS – MÉDITERRANÉE

2004 route des Lucioles - BP 93
06902 Sophia Antipolis Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399