Skin lesion evaluation from multispectral images

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

Skin lesion evaluation from multispectral images

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

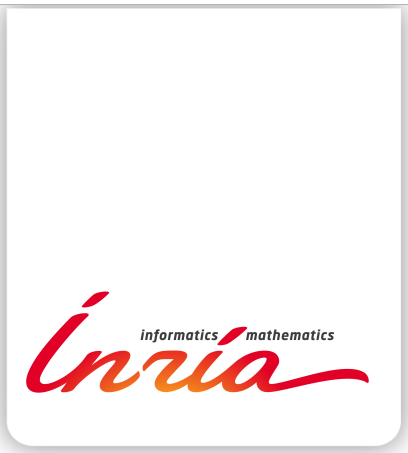
RESEARCH REPORT

N° 8136

November 2012

Project-Teams Morphème and Ayin

ISSN 0249-6399 ISRN INRIA/RR--8136--FR+ENG



Skin lesion evaluation from multispectral images

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

Project-Teams Morphème and Ayin

Research Report n° 8136 — November 2012 — 20 pages

Abstract: During evaluation of skin disease treatments, dermatologists have to clinically measure the pathology severity of each patient during treatment periods. Such a process is sensitive to intra- and inter- dermatologist diagnosis. To make this severity measurement more robust we propose to use image processing to quantify the pathology severity. We focus on a hyperpigmentation disorder called melasma. During a treatment period, multispectral images are taken on patients receiving the same treatment. After co-registration and classification steps, we propose algorithms to measure the intensity, the size and the homogeneity evolution of the pathological area. Obtained results are compared with a dermatologist diagnosis using statistical tests on a full clinical study.

Key-words: skin, hyperpigmentation, change detection, statistical inferences, multispectral imaging.

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

RESEARCH CENTRE
SOPHIA ANTIPOLIS – MÉDITERRANÉE

2004 route des Lucioles - BP 93
06902 Sophia Antipolis Cedex

Evaluation de lésions cutanées par imagerie multispectrale

Résumé : Lors de l'évaluation des traitements des maladies de peau, les dermatologues doivent mesurer la sévérité de la pathologie de chaque patient tout au long d'une période de traitement. Un tel procédé est sensible aux variations intra- et inter- dermatologues. Pour rendre cette mesure de sévérité plus robuste, nous proposons d'utiliser l'imagerie spectrale. Nous nous concentrons sur une pathologie d'hyperpigmentation cutanée appelée mélasma. Au cours d'une période de traitement, des images multispectrales sont acquises sur une population de patients sous traitement. Après des étapes de recalage des séries temporelles d'images et de classification des régions d'intérêt, nous proposons une méthodologie permettant de mesurer, dans le temps, la variation de contraste, de surface et d'homogénéité de la zone pathologique pour chaque patient. Les résultats obtenus sont comparés à un diagnostique clinique à l'aide de tests statistiques réalisés sur une étude clinique complète.

Mots-clés : peau, hyperpigmentation, détection de changements, inférences statistiques, images multispectral

Contents

1 Introduction

In dermatology, specific protocols are established to quantify the evolution of lesions severity along a treatment period. In this report, we focus on a hyperpigmentation disorder called melasma. This disease shows darker and irregular spots on the face caused by an abnormal melanocytes activity in response to a hormonal reaction. Melasma has the specificity to be symmetric in the two cheeks. Thereby, for melasma, the protocol consists in applying a studied treatment in one cheek, and a comparator treatment in the second cheek [1]. The comparator is often chosen as the vehicle used for the studied treatment without active product. To make a clinical study reliable, patients are randomized [1]. The efficacy of a treatment is clinically measured by dermatologists. To this end, a clinical score is evaluated by one or several dermatologists in each cheek of each patient at several time moments during the treatment. Finally, a statistical test such as the Student test [2] or the Wilcoxon test [2] is performed on these measurements to obtain a quantification of the studied treatment efficacy relatively to the comparator treatment [3]. The most widely used score for melasma severity evaluation is the "Melasma Area and Severity Index" (MASI) [4, 5, 6]. It has been introduced in [7]. The MASI split the face into 4 regions: forehead, left cheek, right cheek, and chin. For each region, three scores are evaluated:

- *Area (A):* Measures the area affected by the disease. It is estimated on a scale ranging from 0 to 6 as: $0 = 0\%$; $1 \leq 10\%$; $2 = 10-29\%$; $3 = 30-49\%$; $4 = 50-69\%$; $5 = 70-89\%$ and $6 = 90-100\%$.

- *Darkness (D):* Color of the pathological area compared with normal skin. It is graded on a scale from 0 to 4 as follows: 0=normal skin color without evidence of hyperpigmentation; 1=barely visible hyperpigmentation; 2=mild hyperpigmentation; 3=moderate hyperpigmentation; 4=severe hyperpigmentation.
- *Homogeneity (H):* Homogeneity of the hyperpigmentation, graded on a scale from 0 to 4 as follows: 0=normal skin color without evidence of hyperpigmentation; 1=specks of involvement; 2=small patchy areas of involvement < 1.5 cm diameter; 3=patches of involvement > 2 cm diameter; 4=uniform skin involvement without any clear areas).

To integrate these three criteria measured on the four regions: forehead (F), left cheek (LC), right cheek (RC) and chin (C), the following equation is proposed in [7]:

$$MASI = 0.3(D_F + H_F)A_F + 0.3(D_{LC} + H_{LC})A_{LC} + 0.3(D_{RC} + H_{RC})A_{RC} + 0.1(D_C + H_C)A_C, \quad (1)$$

where D_F, D_{LC}, D_{RC}, D_C are darknesses measured respectively on the forehead, left cheek, right cheek and chin, H_F, H_{LC}, H_{RC}, H_C homogeneities measured respectively on the forehead, left cheek, right cheek and chin and A_F, A_{LC}, A_{RC}, A_C areas measured respectively on the forehead, left cheek, right cheek and chin. This MASI definition has been discussed in [3]. The authors claim that the H criterion has significant variations from one dermatologist to another. They then propose remove the H criterion from Equation (1):

$$MASI = 0.3D_F A_F + 0.3D_{LC}A_{LC} + 0.3D_{RC}A_{RC} + 0.1D_C A_C. \quad (2)$$

In this report, we propose a electronic score called "differential score". The idea is to use images with an associated image processing algorithm in order to automatically evaluate a severity score containing area (A), darkness (D), and homogeneity (H) measurements. Such a computer based score avoids the variability on the measures obtained by human inspection. To describe the proposed protocol, we introduce multispectral imaging for skin analysis, present the clinical study used for experimental results, and describe the proposed scheme.

1.1 Multispectral imaging

Multispectral images are data cubes where each slice corresponds to a monochrome image taken at a specified wavelength. Such images allow the combination of spectral and spatial information of the imaged area. A multispectral image can be seen as a $2D$ spectrocolorimeter. Several algorithms have been proposed to analyse the spectrum of the skin, and showed that the skin spectra contain useful information. The most popular sprectral analysis algorithm for the skin is probably the one proposed by Stamatas et al. in [8, 9]. The authors model the skin absorbency as a mixture of the melanin, deoxy-haemoglobin and oxy-haemoglobin components. A linear regression associated with the Beer-Lamber law provides the proportions of the components mentioned above for each pixel of a multispectral image. The images we use are 1200*960 pixels, and contain 18 spectral bands with wavelengths from 405 nm to 970 nm.

1.2 Clinical study

To evaluate the proposed "differential score", we use a clinical study that contains two treatments denoted T and A. Each treatment is tested on a population of 22 patients. For each patient, one cheek receives one of the treatments T or A and the second cheek receives a comparator that is the vehicle used for treatments T and A. The study lasts 3 months and two multispectral images

are taken for each patient (one per cheek) every month. Thus, we have $2 * 22 * 2 = 88$ series of four images corresponding to times t_0 = baseline, t_1 = after one month, t_2 = after two months and t_3 = after three months.

1.3 Proposed methodology

To evaluate the efficacy of a treatment, we need to measure the severity of the melasma on each image taken along the clinical study. To quantify the severity for an image, we first classify the pathological area in the image, and then extract features describing the disease.

As we are interested in the pathology evolution, we measure change features from t_0 to t_k according to the clinical protocol. Thus, the proposed scheme is drawn is Fig. 1. We first classify each image taken at baseline (t_0). The classification step is not described in this report. In [10] we compare classification methods for skin hyperpigmentation. In [11] we proposed a method based on support vector machines with a global training and an interactive software to perform the classification on all the baseline images. In the second step, we extract a spectral feature describing the studied disease. We define this feature as the linear combination of bands that maximizes the darkness of the disease. To compute the change detection, time series of images need to be co-registered. This is done with the algorithm described in [12] initialized by an affine registration. To evaluate the spatial changes of the disease, we propose two change detection algorithms between monochrome images obtained by the spectral feature. These two algorithms aim at quantifying criterion A and H, respectively. Finally, criteria calculated on the spectral feature map and the change map give an "differential score" for each patient at each measurement time. To detail the presented steps, the report is organized as follows. Section 2 describes the spectral feature extraction method and the darkness criterion. Section 3 describes the proposed method to quantify the changed area between t_0 and t_k. Section 4 proposes to use a multi-scale change detection analysis method based on statistical inferences to get the homogeneity criterion. Section 5 proposes an "differential score" from the criteria D, A and H computed in Section 2, 3 and 4. The last section concludes and discusses the proposed differential score.

Figure 1: Proposed scheme to evaluate a treatment efficacy on a clinical study

2 Darkness criterion

To extract a spectral feature of the studied disease, let's assume that we have extracted a region of healthy skin and a region of pathological skin for each series of co-registered multispectral images. For a given patient, at a given time t_k, we thus have four regions: $I_{t_k}^{pS}$ and $I_{t_k}^{hS}$ the pathological and healthy regions on the cheek treated with the studied treatment, and $I_{t_k}^{pC}$ and $I_{t_k}^{hC}$ the pathological and healthy regions on the cheek treated with the comparator treatment. We propose to use these four regions to compute the weights of a spectral feature M defined by:

$$M(I) = \sum_{b=1}^{N_b} m_b I(b), \tag{3}$$

where m_b is the weight of the b^{st} spectral band $I(b)$ of the image I and N_b the number of spectral bands in the initial multispectral image. We look for features that give the highest contrast between the healthy areas and the pathological areas.

In the literature, feature extraction for spectral images has been widely studied. Methods of features extraction considering a linear combination of spectral bands are PCA [13], NAPCA [14], SVD [15], or ICA [16, 17] that tend to extract features by minimizing the correlation or higher order statistics. The MNF approach [18] searches features by factorizing the signal matrix into two non negative matrices. This factorization can be done by minimizing the mean square error or the Kullback-Leibler divergence for example [19, 20]. Algorithms such as N-FINDR [21, 22, 23], PPI [24, 25] or VCA [26] analyse the geometry of the data set in the spectral N_b dimensional space to extract features defined as extreme points in this set. We can also cite supervised methods like DAFE [27], NWFE [28, 29] or maximum likelihood based approach [30].

2.1 Proposed methods

For skin spectral feature extraction, we focus on two approaches: one involving an objective function optimization, and one based on independent component analysis (ICA). Before going to these two approaches we first introduce a normalization step that combines the four regions $I_{t_k}^{pS}$, $I_{t_k}^{hS}$, $I_{t_k}^{pC}$, $I_{t_k}^{hC}$. The first normalization combines the healthy and the pathological areas. We can either ignore the healthy area or compute the difference between healthy and pathological areas. We do not consider dividing healthy and pathological areas since division is more sensitive to noise. The second normalization combines the studied treatment S and the comparator treatment C. We use the method proposed in clinical protocols that calculate the difference $S-C$. Finally, depending on the use of the healthy area we have two possible normalizations:

$$D_{t_k}^e = (\mu_{M(I_{t_k}^{pS})}) - (\mu_{M(I_{t_k}^{pC})}) \tag{4}$$

and

$$D_{t_k}^e = (\mu_{M(I_{t_k}^{hS})} - \mu_{M(I_{t_k}^{pS})}) - (\mu_{M(I_{t_k}^{hC})} - \mu_{M(I_{t_k}^{pC})}), \tag{5}$$

where $D_{t_k}^e$ denotes the darkness for the patient e at t_k and $\mu_{M(I_{t_k}^{hS})}$ (respectively $\mu_{M(I_{t_k}^{pS})}$, $\mu_{M(I_{t_k}^{hC})}$, $\mu_{M(I_{t_k}^{pC})}$) is the spatial average of the spectral feature M on the region $I_{t_k}^{hS}$ (respectively $I_{t_k}^{pS}$, $I_{t_k}^{hC}$, $I_{t_k}^{pC}$).

2.1.1 Objective function

In this section, we propose an objective function that maximizes both the distance between pathological and healthy areas, and the deviation between measurement times. The distance between healthy and pathological areas can be measured by a difference to avoid noise introduced by divisions. Thus, the objective function is:

$$\begin{aligned} f = \sum_{b=1}^{N_b} m_b \Bigg\{ \sum_{t=t_1}^{t_n} \sum_{e=1}^{N_e} \Big[& \mu_{I_{t_0}^{e,hS}} - \mu_{I_{t_0}^{e,pS}} \\ & - \mu_{I_{t_0}^{e,hC}} + \mu_{I_{t_0}^{e,pC}} \\ & - \mu_{I_t^{e,hS}} + \mu_{I_t^{e,pS}} \\ & + \mu_{I_t^{e,hC}} - \mu_{I_t^{epC}} \Big] \Bigg\}. \end{aligned} \tag{6}$$

The function f is linear with respect to the parameters m_b of M. The optimization is then trivial. The solution is the spectral band with the highest contrast if we consider only positive m_b. If we consider also negative m_b, the solution is given by two bands. These solutions are optimal with respect to the darkness criterion.

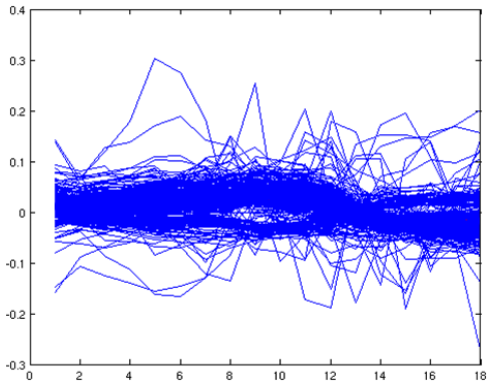


Figure 2: Feature obtained by ICA on all the images at t_0 on a clinical study of melasma. **X-axis:** indexes of spectral bands, **Y-axis:** coefficient given to spectral bands.

2.1.2 ICA-based method

In this section we propose to use ICA to get a spectral feature of melasma. Performing ICA on a multispectral image gives, at most, as many features as the number of bands in the image. We then select the most representative feature of the pathological area. The selected component is the one that maximizes the deviation between a healthy reference area and the pathological area of interest. If we calculate ICA separately on all the images we get slightly different features from one images to another. On Fig. 2 one can see that features obtained in all the images of a clinical study have similar shape. Several features are negative due to the ICA model indetermination. To get a single feature for all the images, we propose to calculate the average feature with a -1 coefficient for the negatives features. It can be noticed that the coefficients in the near infra-red (IR) bands are negative. During experiments we observe that this near infra-red information biased the contrast measurements. The IR coefficients for the ICA based feature are then set to zero.

2.2 Obtained results

Proposed feature extraction methods have been tested in a clinical study involving three active treatments that we denote S_t, A_{d1} and A_{d2}. Clinical conclusions state that S_t has no efficacy whereas A_{d1} and A_{d2} have efficacy. Furthermore, A_{d2} has a stronger effect than A_{d1}. The protocol of this study is the same as the one presented in section 1.2.

We calculate a spectral feature for each treatment S_t, A_{d1} and A_{d2} individually. Then we compute the features in all the images of the study on the four regions $I_{t_k}^{pS}$, $I_{t_k}^{hS}$, $I_{t_k}^{pC}$, $I_{t_k}^{hC}$ manually segmented. It results in a darkness measurement $D_{t_k}^{e}$ for each patient. We finally apply a Wilcoxon hypothesis test for each treatment between t_k and t_0 to get a quantification of the treatment efficacy. Table 1 shows the p-values obtained by the Wilcoxon test with the proposed methods and with the luminance feature that we use as a reference feature. Equation (5) is used for normalization. As one can see, all the features agree with the clinical conclusions. The difference is in how fast features detect an efficacy of the treatment. The method based on the objective function optimization gives the fastest efficacy detection compared to the ICA-based method and the luminance. To understand these differences, we plot the features obtained with all the methods in Fig. 3. As the luminance coefficient is not linear, we use its linear approximation Y. We observe that the feature obtained by f selects a specific band at 590 nm

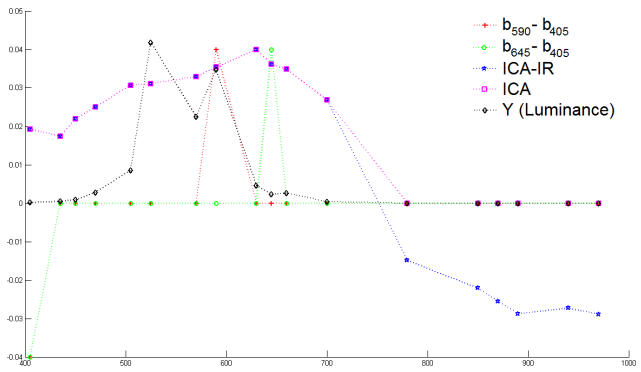


Figure 3: Spectral features obtained by the studied methods. **X-axis:** wavelengths in nanometers, **Y-axis:** coefficients m_b.

Table 1: Results obtained by the Wilcoxon test on a clinical study with the normalization of Equation (5) for the darkness criterion.

Feature	Traitement	$t_1 - t_0$	$t_2 - t_0$	$t_3 - t_0$
L	S_t	$7.959\ 10^{-1}$	$5.014\ 10^{-1}$	$7.173\ 10^{-1}$
	A_{d1}	$3.793\ 10^{-1}$	$\mathbf{1.128\ 10^{-2}}$	$8.793\ 10^{-2}$
	A_{d2}	$9.798\ 10^{-2}$	$\mathbf{3.204\ 10^{-3}}$	$\mathbf{5.312\ 10^{-4}}$
f	S_t	$1.476\ 10^{-1}$	$1.089\ 10^{-1}$	$7.563\ 10^{-1}$
	A_{d1}	$5.571\ 10^{-2}$	$\mathbf{3.400\ 10^{-2}}$	$\mathbf{4.373\ 10^{-2}}$
	A_{d2}	$\mathbf{8.360\ 10^{-3}}$	$\mathbf{3.204\ 10^{-3}}$	$\mathbf{9.350\ 10^{-4}}$
ICA	S_t	$5.349\ 10^{-1}$	$8.767\ 10^{-1}$	$9.176\ 10^{-1}$
	A_{d1}	$4.080\ 10^{-1}$	$\mathbf{2.987\ 10^{-2}}$	$9.798\ 10^{-2}$
	A_{d2}	$7.873\ 10^{-2}$	$\mathbf{5.233\ 10^{-3}}$	$\mathbf{9.350\ 10^{-4}}$

for the treatments S_t, and A_{d2} and the band 645 nm for the treatment A_{d1}. ICA selects spectral bands mainly around 600 nm, but sets significant coefficients to wavelengths between 400 and 600 nm. The Y feature selects bands in the 500-600 nm wavelengths. Due to the Stamatas et al. model for pigmentation we should focus on the wavelength between 600 and 700 nm. Thus, heigh coefficients in the spectral ranges out of 600-700 nm explain why the ICA based method and the Y features detect later the treatments efficacy than the feature estimated by f. Table 2 shows the results obtained with the Wilcoxon tests on the same studies with the normalization of Equation (4). The results are comparable to the ones with the normalization of Equation (5). This means that the calibration of the camera is robust and that the normalization with the healthy area is not mandatory. As this healthy normalization can introduce noise due to the variation of the healthy area, we prefer to use the normalization of Equation (4). To validate the spectral features calculated on the test clinical study, we apply them on the validation study described in section 1.2. The p-values of the Wilcoxon tests on the validation study are shown in Table 3. For f we choose the feature $b_{590} - b_{405}$. Obtained results are coherent with the clinical analysis.

Table 2: Results obtained by the Wilcoxon test on a clinical study with the normalization of Equation (4) for the darkness criterion.

Feature	Traitement	$t_1 - t_0$	$t_2 - t_0$	$t_3 - t_0$
L	S_t	$7.959\ 10^{-1}$	$6.791\ 10^{-1}$	$6.416\ 10^{-1}$
	A_{d1}	$\mathbf{1.306\ 10^{-2}}$	$\mathbf{2.289\ 10^{-2}}$	$9.798\ 10^{-2}$
	A_{d2}	$1.788\ 10^{-1}$	$\mathbf{7.169\ 10^{-3}}$	$\mathbf{1.918\ 10^{-3}}$
f	S_t	$7.173\ 10^{-1}$	$5.349\ 10^{-1}$	$5.014\ 10^{-1}$
	A_{d2}	$\mathbf{2.289\ 10^{-2}}$	$\mathbf{1.737\ 10^{-2}}$	$\mathbf{3.400\ 10^{-2}}$
	A_{d3}	$\mathbf{1.997\ 10^{-2}}$	$\mathbf{1.609\ 10^{-3}}$	$\mathbf{9.725\ 10^{-3}}$
ICA	S_t	$6.416\ 10^{-1}$	$5.694\ 10^{-1}$	$5.014\ 10^{-1}$
	A_{d1}	$\mathbf{1.997\ 10^{-2}}$	$\mathbf{3.400\ 10^{-2}}$	$8.793\ 10^{-2}$
	A_{d2}	$1.476\ 10^{-1}$	$\mathbf{7.169\ 10^{-3}}$	$\mathbf{1.123\ 10^{-3}}$

Table 3: Results obtained by the Wilcoxon test on the validation clinical study for the darkness criterion.

Test	Treatment	$t_1 - t_0$	$t_2 - t_0$	$t_3 - t_0$
L	A	$8.455\ 10^{-1}$	$7.210\ 10^{-1}$	$1.270\ 10^{-1}$
	T	$9.741\ 10^{-1}$	$\mathbf{9.397\ 10^{-3}}$	$\mathbf{1.637\ 10^{-3}}$
f	A	$1.116\ 10^{-1}$	$5.373\ 10^{-1}$	$9.777\ 10^{-2}$
	T	$4.750\ 10^{-1}$	$\mathbf{3.859\ 10^{-3}}$	$\mathbf{9.337\ 10^{-3}}$
ICA	A	$7.210\ 10^{-1}$	$5.589\ 10^{-1}$	$1.627\ 10^{-1}$
	T	$8.966\ 10^{-1}$	$\mathbf{6.389\ 10^{-3}}$	$\mathbf{3.478\ 10^{-3}}$

3 Area criterion

To calculate an area criterion, we propose to compute the spatial changes between t_k and t_0. Lets denote $C^e_{t_0}$ a classification mask at t_0 for the patient e obtained by the method proposed in [31]. We denote by $B^e_{t_{k/0}}$ a binary change map between t_0 and t_k for the patient e. Then, the area criterion can be calculated by:

$$A^e_{t_{k/0}} = \frac{\text{card}\{x:\ B^e_{t_{k/0}}(x) = 1\}}{\text{card}\{x:\ C^e_{t_0}(x) = 1\}}. \tag{7}$$

In the literature, numerous methods have been proposed to quantify the changes between two images. Two state of the art articles are [32, 33]. Change detection methods consider the image either globally by subtraction or division [33, 34, 35, 36] or locally with a window that browses the image [37, 38, 39, 40, 41, 42].

Using a local window tends to a loss of resolution. Furthermore, the procedure of dividing images is sensitive to noise. That is why, we choose to perform change detection by subtraction:

$$F^e_{t_{k/0}} = M(I^e_{t_k}) - M(I^e_{t_0}), \tag{8}$$

where $M(I^e_{t_k})$ (respectively $M(I^e_{t_0})$) is a gray scale image obtained by integrating the spectral image $I^e_{t_k}$ (respectively $I^e_{t_0}$) with the spectral feature M. A threshold τ on $F^e_{t_{k/0}}$ allows to get the binary change map:

$$B^e_{t_{k/0}}(x) = \begin{cases} 1 & \text{if } F^e_{t_{k/0}}(x) > \tau \text{ and } C^e_{t_0}(x) = 1 \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

where x denotes the position of any pixel in the image. The choice of τ is crucial. We propose a method that computes one threshold τ per image series estimated on the histogram of $F^e_{t_{1/0}}$. First, we assume that the histogram of $F^e_{t_{1/0}}$ is monomodal symmetric and centred in zero in the case of no changes $F^e_{t_{k/0}}$. This assumption is justified since $F^e_{t_{1/0}}$ is a centred noise in zero in the case of no changes. Then, we assume that there are no significant changes in the negative part (i.e. the disease does not become worst). Thus, in the case of no disease decreasing, the number of positive and negative pixels in $F^e_{t_{1/0}}$ should be the same. With this model, the area of changes in $F^e_{t_{1/0}}$ is $A = \frac{(\sharp>0)-(\sharp\leq 0)}{(\sharp>0)+(\sharp\leq 0)}$, where $(\sharp > 0)$ (respectively $(\sharp \leq 0)$) denote the number of positive (respectively negative) pixels in $F^e_{t_{1/0}}$. To calculate τ from $F^e_{t_{1/0}}$, we sort the positive pixels of $F^e_{t_{1/0}}$ and take the threshold τ such as: $(\sharp > 0 \text{ and } < \tau) = (\sharp \leq 0)$.

An inconvenience of thresholding methods for change detection is their sensitivity to noise. To decrease noise, we filter the change map $B^e_{t_{k/0}}$ with opening and closing mathematical morphology operators using a disk of two pixels as a structural element. Fig. 9 d-f) shows obtained results for an image time series. For this patient, dermatologists clinically measure a decrease of the pathology. As one can see, the area measured by the algorithm shows a significant evolution at time t_1 and then a gradual evolution of the pathological area. To validate this area measurement, we calculate change areas $A^e_{t_{k/0}}$ for all the patients of the validation clinical studies. As for the darkness measurement, we compute the Wilcoxon test for each treatment between t_k and t_0 in the areas measurements. Results are shown in Table 4. Results show that the treatment T has an effect whereas the treatment A does not. This is in agreement with the dermatologists diagnosis.

Table 4: p-values of the Wilcoxon tests on the validation clinical study for the area criterion.

Criterion	Treatment	$t_1 - t_0$	$t_2 - t_0$	$t_3 - t_0$
Area	A	$2.840\ 10^{-1}$	$3.986\ 10^{-1}$	$1.045\ 10^{-1}$
	T	$4.552\ 10^{-1}$	$\mathbf{3.478\ 10^{-3}}$	$\mathbf{2.438\ 10^{-4}}$

4 Homogeneity criterion

With the proposed darkness (section 2) and area (section 3) criteria, we obtained a definition of a differential score with two components D and A like in the clinical MASI defined in [3]. In this section we study a differential and numerical equivalent of the homogeneity clinical criterion proposed in [7]. This criterion jointly measures the distribution and the size of brown areas constituting the pathology area. To get such information, we adapt the methodology of "Statistical Parametric Mapping (SPM)" [43] developed for functional MRI. This approach models a change map as a Gaussian random field. Then a multi-scale analysis partition this map into regions and a probability is assigned to each region. These probabilities measure the likelihood for the regions to be a realization of the random field. For our application, we apply this methodology on each change map $F^e_{t_{k/0}}$. To model $F^e_{t_{k/0}}$ as a Gaussian random field, we need to normalize it. Then we apply the multi-scale analysis to get regions and calculate probabilities.

4.1 Change map normalization

First, we smooth the image $F^e_{t_{k/0}}$ with a Gaussian convolution filter to increase its spatial correlation. Second, as the interest area of $F^e_{t_{k/0}}$ comes from a unique class, its histogram can be modelled as monomodal in the hypothesis of no changes. We use a histogram specification to transform it into a Gaussian distribution. We center and reduce the specified histogram to standardize it. In the following we will denote $F^{Ne}_{t_{k/0}}$ the normalized change map.

4.2 Change map thresholding

$F^{Ne}_{t_{k/0}}$ satisfies the conditions to apply the Gaussian field theory [44]. An interesting characteristic of a Gaussian field F of size S is the DT (Differential Topology) characteristic defined by:

$$\chi(A_u) = (-1)^{(d-1)} \sum_{k=0}^{d-1} (-1)^k \chi_k(A_u) \tag{10}$$

with

$$A_u = \{s \in S, F(s) \geq u\}, \tag{11}$$

where u is a given threshold, d the dimension of the field and $\chi_k(A_u)$ the number of points $s \in S$ satisfying the conditions: (a) $F(s) = u$, (b) $F^{(i)}(s) = 0, \forall i \in [1, d-1]$, (c) $F^{(d)}(s) > 0$ and (d) the $(d-1) \times (d-1)$ matrix of second order partial derivatives of F(s) has exactly k negative eigenvalues. In the above expression, $F^{(i)}$ denotes the i^{th} derivative of F. In the 2D case (i.e. $d = 2$), for a given threshold u the expectation of the DT characteristic of a centred and reduced Gaussian field can be written as follows [44]:

$$\mathbb{E}_{\chi(A_u)} = \mathcal{S}(2\pi)^{\frac{3}{2}} |\Lambda|^{\frac{1}{2}} \sigma^{-3} u e^{-\frac{u^2}{2\sigma^2}} \tag{12}$$

with $\mathcal{S}$ the surface of the field, Λ the 2×2 correlation matrix of the field derivatives, and $\sigma^2 = \mathbb{E}(F^2)$.

In order to estimate the likelihood of a pixel of $F^{Ne}_{t_{k/0}}$ and its neighbourhood under the Gaussian field assumption, we compare its characteristic at a threshold u with the DT characteristic. Two parameters can be used to compute this likelihood: the maximum intensity and the spatial extent of the considered neighbourhood above u.

Let x_0 denotes the maximum intensity of $R^{x_0}_u$ a region above the threshold u. The likelihood of this region under the Gaussian field assumption is [45]:

$$P(R^{x_0}_u \in F) = \frac{\mathbb{E}_{\chi(A_{x_0})}}{\mathbb{E}_{\chi(A_u)}} = \frac{x_0}{u} e^{\frac{u^2 - x_0^2}{2}}. \tag{13}$$

Let S_0 denotes the spatial extent of $R^{S_0}_u$ a region above the threshold u. The expectation for this region under the Gaussian field assumption is [45]:

$$\mathbb{E}(R^{S_0}_u \in F) = \frac{\mathbb{E}(N_u)}{\mathbb{E}_{\chi(A_u)}}, \tag{14}$$

where $\mathbb{E}(N_u)$ is the expectation to have N_u pixels above the threshold u in a Gaussian random field F. As F follows a standard normal distribution,

$$\mathbb{E}(N_u) = \mathcal{S} \int_u^\infty (2\pi)^{\frac{1}{2}} e^{-\frac{x^2}{2}} = \mathcal{S}\Phi(-u) \tag{15}$$

then,

$$\mathbb{E}(R^{S_0}_u \in F) = \frac{\Phi(-u)}{(2\pi)^{\frac{3}{2}} |\Lambda|^{\frac{1}{2}} u e^{-\frac{u^2}{2}}}. \tag{16}$$

Thus, the number of pixels in a region above the threshold u has an exponential distribution with the parameter $\lambda_u = 1/\mathbb{E}(R^{S_0}_u \in F)$. We can then write the probability of a region $R^{S_0}_u$ to appear in a realization of the Gaussian field by:

$$P(R^{S_0}_u \in F) = e^{-\lambda_u S_0} = e^{\frac{(2\pi)^{\frac{3}{2}} |\Lambda|^{\frac{1}{2}} S_0 u e^{-\frac{u^2}{2}}}{\Phi(-u)}} \tag{17}$$

Finally, for a spatial region above u with the maximum intensity x_0 and the spatial extend S_0, its probability to be in F is [45]:

$$P(R^{x_0,S_0}_u \in F) = min(P(R^{x_0}_u \in F), P(R^{S_0}_u \in F)). \tag{18}$$

Thereby, to detect changes between t_0 and t_k we define a family of n_u thresholds $U = \{u_1, ..., u_{n_u}\}$. For each u_i, $F^e_{t_{k/0}}$ is segmented into regions denoted $R^{xo,S_0}_{u_i}$. This segmentation is obtained by a connected components analysis of the class defined by the pixels of $F^e_{t_{k/0}}$ above the threshold u_i. Then we assign a probability $P(R^{xo,S_0}_{u_i} \in F)$ to each region $R^{xo,S_0}_{u_i}$. This process gives a statistical map of the changes denoted $SM^e_{t_k}$ between t_0 and t_k. The algorithm 1 summarizes the procedure for obtaining SM_{t_k} between I_0 and I_k for a given patient.

4.3 Final normalization

The statistical approach proposed in Section 4.2 makes the hypothesis that changes are rare events. In our images, a large part of the pathological area can change. Thus, if we apply directly the methodology describes above, we will detect only a small part of the changes (see

Algorithm 1 Multi-scale analysis

Require: I_0, I_k, M_{p_0}
1: $F^e_{t_{k/0}} = I_k - I_0$
2: Smooth and normalize $F^e_{t_{k/0}}$ to get $F^{Ne}_{t_{k/0}}$ (see section 4.1)
3: Choose a family $U = u_1, ..., u_{n_u}$
4: **for** $i = 1$ to n_u **do**
5: Compute the connected components above u_i
6: **for** each connected component **do**
7: Calculate $P(R^{x_0,S_0}_u \in F)$ with eq. 18
8: **end for**
9: **end for**
10: Merge the obtained connected clusters obtained with the family U to get $SM^e_{t_k}$

figure 9 g-i)). Homogeneity variations cannot be quantified in these maps. We propose then to introduce the normalization:

$$F^{N_d e}_{t_{k/0}} = F^{Ne}_{t_{k/0}} + \frac{1}{N_r} \sum_{r=t_0}^{t_k} \left(\mu_{C_{t_0}}(F^{Ne}_{t_{r/0}}) - \mu_{C_{t_0}}(F^{Ne}_{t_{1/0}}) \right) \tag{19}$$

where N_r is the number of time measurements between t_0 and t_k. $\mu_{C_{t_0}}(F^{Ne}_{t_{r/0}})$ (respectively $\mu_{C_{t_0}}(F^{Ne}_{t_{1/0}})$) denote the average of the pixels intensities in the interest area of $F^{Ne}_{t_{r/0}}$ (respectively $F^{Ne}_{t_{1/0}}$) delimited by C_{t_0}. If the normalized data with Equation (19) do not respect the rare event hypothesis, the calculated statistics will be biased. Nevertheless, we do not use the absolute values of these statistics but their spatial repartition. Fig. 9 j-l) shows $SM^e_{t_k}$ cartographies obtained with the normalization of Equation (19). The pathology evolution is now visible during the time sequence.

To quantify a treatment efficacy from homogeneity maps, we integrate them into a scalar criterion. To do so, we define a function $f^e_s(p)$ that represents the spatial repartition of the changes by:

$$f^e_s(p) = \frac{card\{x : SM^e_{t_k}(x) \leq p\}}{card\{SM^e_{t_k}\}}. \tag{20}$$

$f^e_s(p)$ is an increasing function from $[0,1]$ to $[0,1]$. As the interpretation of $f^e_s(p)$ is not straightforward, we prefer to work with $f^{pe}_{t_k}(s)$ that is the inverse function of $f^e_s(p)$. A linear interpolation is used to calculate $f^{pe}_{t_k}(s)$ with a regular sampling along s. $f^{pe}_{t_k}(s)$ is also an increasing function from $[0,1]$ to $[0,1]$. In the following, we use $f^{pe}_{t_k}(s)$ for both calculating a graphical representation of the homogeneity criterion (see section 5), and a scalar $H^e_{t_k}$ describing the homogeneity of the patient e at time t_k.

To integrate $f^{pe}_{t_k}(s)$ in a scalar $H^e_{t_k} \in [0,1]$, we can compute:

$$H^e_{t_k} = 2 \int_0^1 |f^{pe}_{t_k}(s) - s| ds. \tag{21}$$

$H^e_{t_k}$ represents the area between the function $f^{pe}_{t_k}(s)$ and the identity function $Id(s) = s$. Id corresponds to the most heterogeneous repartition that can be expected. Then, more homogeneous changes yield higher values $H^e_{t_k}$. To calculate $H^e_{t_k}$ with discrete data, we use the Riemann integral.

To validate the proposed homogeneity criterion, we apply it on the validation study. The homogeneity measure makes sense only when some changes appear in the image between t_0 and

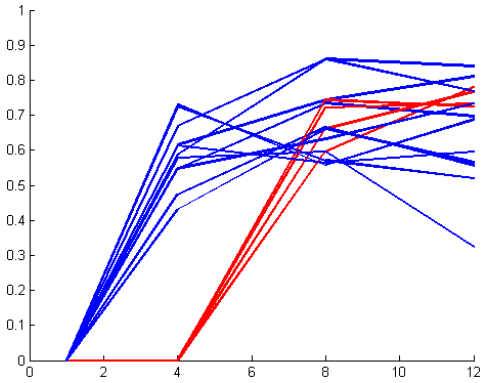


Figure 4: Time sequences of the homogeneity measurements for the patients of the group where an evolution has been measured by the area and darkness criteria. **X-axis:** time in weeks ($t_0 = 1$, $t_1 = 4$, $t_2 = 8$, $t_3 = 12$), **Y-axis:** homogeneity value.

t_k. A margin is then fixed at 10% of the pathological area changes below which we do not consider the criterion $H^e_{t_k}$. To qualitatively evaluate the homogeneity criterion, we plot its evolution for the population of patients whose pathology evolve. Due to the darkness and the area criteria, the considered patients are those who receive the treatment T in the validation study. Fig. 4 shows time plots of the homogeneity measurement for each patient of the group receiving the treatment T. If the treatment has an effect, we expect to see the homogeneity measurement growing. In this population of 22 patients, we observed that for only 7 patients the H criterion is not interpretable. Indeed, for five patients less than 10% of the pathological area changes at time $t_2 = 8$ or $t_3 = 12$ and for two patients there were less than 10% of the pathological area that change in all the time sequence. For the others fifteen patients we globally observe the expected growth of the homogeneity with time. For most patients (in blue) main changes occurred in the second measurement time $t_1 = 4$, and for 4 patients (in red), main changes occurred in time $t_2 = 8$. These results show that the homogeneity criterion gives relevant interpretation of disease progression during treatment. This criterion is interpretable only on patients whose pathology has evolved. This is why we recommend to use $H^e_{t_k}$ as an "extra" criterion to analyse the evolution of a patient pathology.

5 Graphical representation

The clinical MASI as described in [3] and [7] calculates a linear combination of darkness, area and homogeneity measurements (see Equations (1) and (2)). For the differential score, such linear combination does not have any meaning. Indeed, the three criteria have different dynamics, and combining them can bias the final measurement. That is why we prefer to keep the three measurements independent. To make the differential score easily readable for an individual patient we propose the following graphical representation using colored bars.

5.1 Darkness representation

The proposed darkness criterion is a differential measure between two monochrome images. Its theoretical magnitude is then $[-255, 255]$ for 8 bits images. In the experiments, we observe that the interval is $[-12, 12]$. To make a graphical representation of D, we use a function to transform

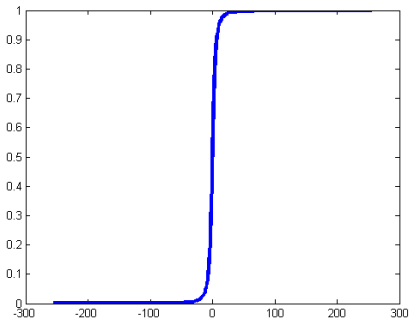


Figure 5: Plot of the darkness transformation of Equation (22). **X-axis:** D in the scale values [-255, 255], **Y-axis:** D_G values in the scale [0,1] .

(a) (b) (c)

Figure 6: Darkness graphical representation. a) for a patient whose pathology became darker, b) for a patient with stationary darkness, c) for a patient whose pathology decrease.

the maximum of dynamic from the interval $[-12, 12]$ into the interval $[0, 1]$:

$$D_G = \frac{D}{\sqrt{1 + D^2}} \tag{22}$$

where D_G is the darkness value used for the graphical representation. Fig. 5 shows a plot of the function of Equation (22). The darkness bar contains a chip with a color from red (-12) to blue (+12). This chip position on the bar depends on the D_G value from -12 to 12. Fig. 6 shows three examples of darkness representation bars.

5.2 Area representation

The area criterion we proposed has values in the interval $[0, 1]$. The graphical representation we choose is then a percentage bar. Fig. 7 shows examples of graphical representation of the area criterion.

(a) (b) (c)

Figure 7: Area graphical representation. a) for a patient with stationary pathological area, b) for a patient whose pathological area decrease by 26%, c) for a patient whose pathological area decrease by 86%.

(a) (b) (c)

Figure 8: Homogeneity graphical representation. a) for a patient with without evolution of the pathology, b) for a patient with heterogeneous variation of the pathology, c) for a patient with homogeneous variation of the pathology.

5.3 Homogeneity representation

For the graphical representation of the homogeneity criterion, we propose to use $f_{t_k}^{pe}(s)$. In fact this information is richer than $H_{t_k}^{e}$. The graphical bar of homogeneity is then a plot of $f_{t_k}^{pe}(s)$ with a gradient color. Values of $f_{t_k}^{pe}(s)$ close to zero are in blue and values of $f_{t_k}^{pe}(s)$ close to one are in red. Intermediate values are represented with the "jet" colormap. If the gradient is all red, that means there is no homogeneity variation. The slower the gradient changes from blue to red, the more heterogeneous the variation is. Fig. 8 shows examples of the homogeneity bar.

5.4 Differential score representation

Finally, the differential score representation is the combination of the three bars of darkness, area and homogeneity. An example for a patient whose pathology decrease is shown in Fig. 9 and an example for a patient whose pathology does not decrease is shown in Fig. 10. As one can see, with this graphical representation, we can both quantify and qualify the patient pathology evolution.

6 Conclusion

In this report, we proposed a strategy to calculate the severity of skin hyperpigmentation using multispectral images. The proposed criterion called "differential score" has been developed to be a computer score related to the clinical MASI criterion. The proposed differential score has been tested on a full clinical study and gave relevant results with the clinical analysis. The proposed methodology does not take into account any specificity of the studied pathology. We can thus adapt it to other pathologies like vitiligo, scars or rosacea.

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.

References

[1] A. Pandya, M. Berneburg, J. Ortonne, and M. Picardo. Guidelines for clinical trials in melasma. *Br J Dermatol*, 156 Suppl 1, 2006.

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

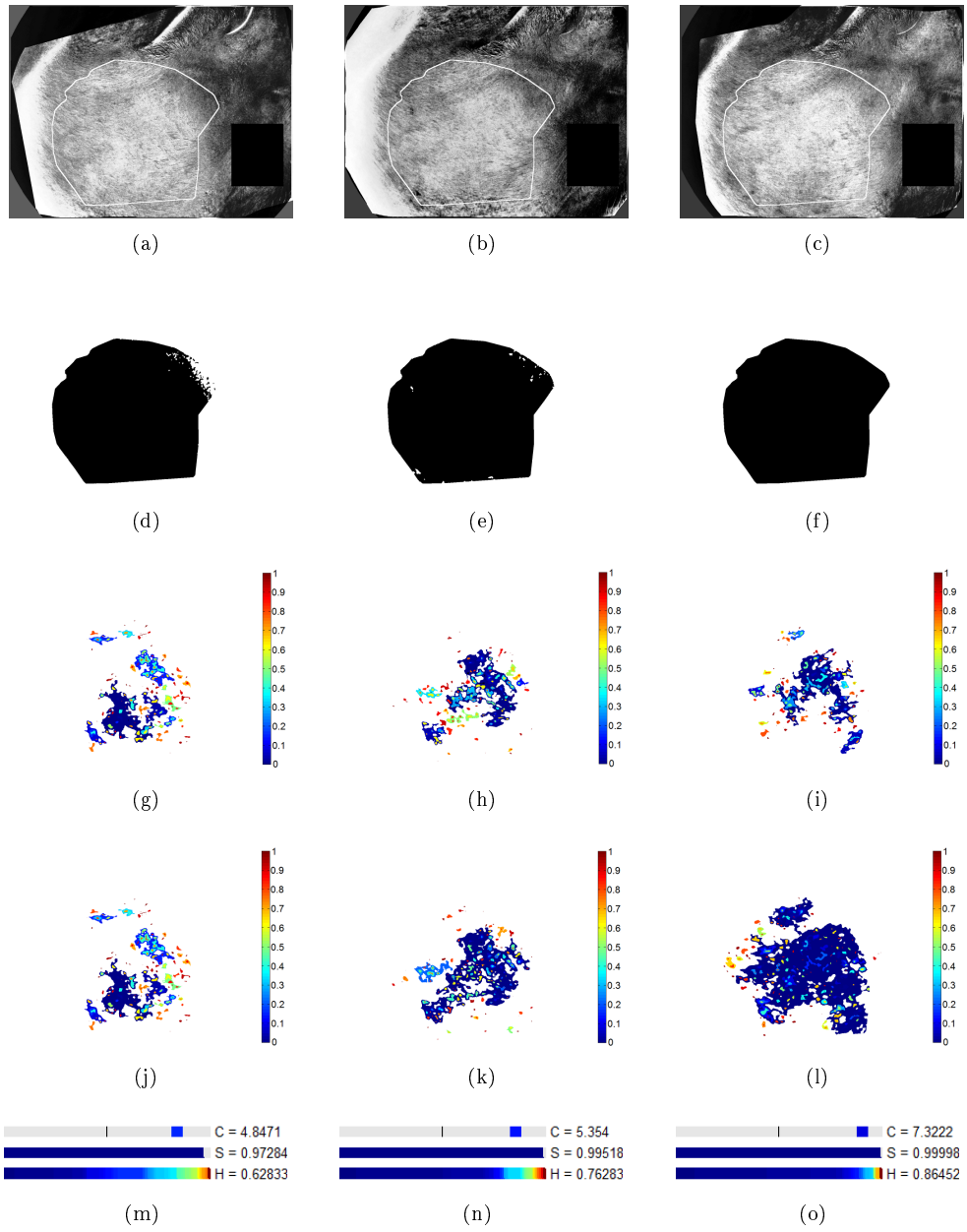


Figure 9: Changes map and graphical representation of the differential score for a patient whose pathology decreases during the treatment period. a-c) equalized difference image on the spectral feature map between each time t_1, t_2 et t_3 and t_0, d-f) change detection map used to measure the area for each times t_1, t_2 et t_3, g-i) homogeneity maps $SM^e_{t_k}$ for the three time t_1, t_2 et t_3 without using the normalization of Equation (19), j-l) homogeneity maps $SM^e_{t_k}$ for the three time t_1, t_2 et t_3 using the normalization of Equation (19), m-o) graphical representation of the differential score for the three times t_1, t_2 et t_3.

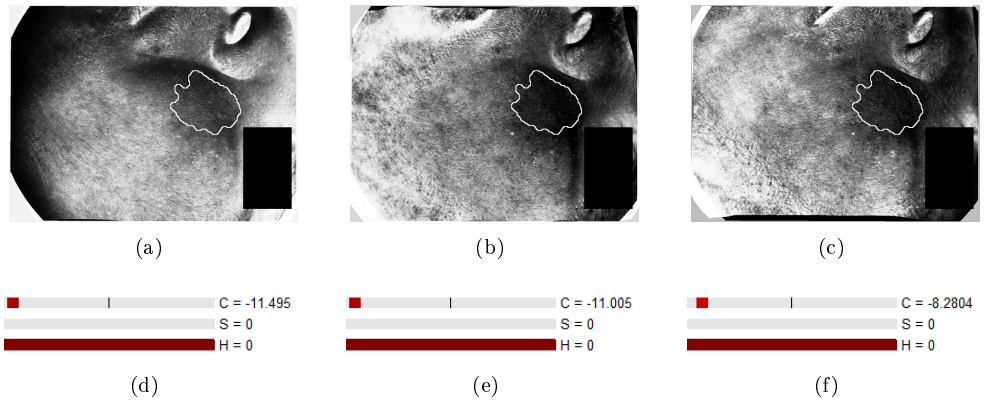


Figure 10: Graphical representation of the differential score for a patient whose pathology did not decrease during the treatment period. a-c) equalized difference image on the spectral feature map between each time t_1, t_2 et t_3 and t_0, d-f) graphical representation of the differential score for the three times t_1, t_2 et t_3.

[3] A.G. Pandya, L.S. Hynan, R. Bhore, F.C. Riley, I.L. Guevara, P. Grimes, J.J. Nordlund, M. Rendon, S. Taylor, R.W. Gottschalk, N.G. Agim, and J.P. Ortonne. Reliability assessment and validation of the melasma area and severity index (masi) and a new modified masi scoring method. *Journal of the American Academy of Dermatology*, 64(1):78 – 83.e2, 2011.

[4] R. Balkrishnan, A.J. Mcmichael, F.T. Camacho, F. Saltzberg, T.S. Housman, S. Grummer, S.R. Feldman, and M.M. Chren. Development and validation of a health-related quality of life instrument for women with melasma. *British Journal of Dermatology*, 149:572–577, 2003.

[5] R.M.P. Manaloto and T. Alster. Erbium:yag laser resurfacing for refractory melasma. *Dermatologic Surgery*, 25:121–123, 1999.

[6] M.E. Hurley, I.L. Guevara, R.M. Gonzales, and A.G. Pandya. Efficacy of glycolic acid peels in the treatment of melasma. *Archives of Dermatology*, 138:1578–1582, 2002.

[7] CK Kimbrough-Green, CE Griffiths, LJ Finkel, TA Hamilton, SM Bulengo-Ransby, CN Ellis, and et al. Topical retinoic acid (tretinoin) for melasma in black patients. a vehicle-controlled clinical trial. *Arch Dermatol*, 130:727–33, 1994.

[8] 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.

[9] 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.

[10] 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.

[11] Sylvain Prigent, Xavier Descombes, Didier Zugaj, Laurent Petit, Anne-Sophie Dugaret, Philippe Martel, and Josiane Zerubia. Classification of skin hyper-pigmentation lesions with multi-spectral images. Research Report RR-8105, INRIA, October 2012.

[12] A. Myronenko and X. Song. Intensity-based image registration by minimizing residual complexity. *IEEE Trans. on Medical Imaging*, 29(11):1882–1891, 2010.

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

[14] A.A. Green, M. Berman, P. Switzer, and M.D. Craig. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Trans. Geosciences Remote Sensing*, 26, 1988.

[15] J.E. Ball, L.M. Bruce, and N.H. Younan. Hyperspectral pixel unmixing via spectral band selection and dc-insensitive singular value decomposition. *Geoscience and Remote Sensing Letters, IEEE*, 4, Juillet 2007.

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

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

[18] P. Paatero and U. Tapper. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5:111–126, 1994.

[19] I.S. Dhillon and S. Sra. Generalized nonnegative matrix approximations with bregman divergences. In *NIPS*, 2005.

[20] D.D. Lee and H.S. Seung. Learning the parts of objects by non-negative matrix factorization. 401:788, 1999.

[21] M. Winter. Fast autonomous spectral end-member determination in hyperspectral data. In *Conf. on Applied Geologic Remote Sensing*, volume 2, pages 337–344, 1999.

[22] N. Dobigeon, S. Moussaoui, M. Coulon, J.Y. Tourneret, and A.O. Hero. Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery. *IEEE Trans. Signal Processing*, 57(11):4355–4368, Nov. 2009.

[23] C. Chang, C.C. Wu, and C.T. Tsai. Random n-finder (n-findr) endmember extraction algorithms for hyperspectral imagery. *IEEE Trans. on Image Processing*, 20, March 2011.

[24] J. W. Boardman, F.A. Kruse, and R.O. Green. Mapping target signatures via partial unmixing of aviris data. In *in Summaries of JPL Airborne Earth Science Workshop*, 1995.

[25] A. Plaza and C.I. Chang. A fast iterative algorithm for implementation of pixel purity index. *IEEE Geoscience and Remote Sensing Letters*, 3, Jan. 2006.

[26] J. M. P. Nascimento and J. M. Dias. Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE Trans. Geosci. Remote Sens.*, 43:898–910, Apr. 2005.

[27] D.A. Landgrebe. *Signal Theory Methods in Multispectral Remote Sensing*. Hoboken, New Jersey,, 2003.

[28] C. Lee and D.A. Landgrebe. Feature extraction based on decision boundaries. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 15:388–400, Apr 1993.

[29] C. Lee and D.A. Landgrebe. Decision boundary feature extraction for neural networks. *IEEE Trans. on Neural Networks*, 8:75–83, Jan 1997.

[30] M. Riedmann and E.J. Milton. Supervised band selection for optimal use of data from airborne hyperspectral sensors. In *IEEE IGARSS*, volume 3, 2003.

[31] S. Prigent, X. Descombes, D. Zugaj, L. Petit, A.S. Dugaret, P. Martel, and J. Zerubia. Classification of skin hyper-pigmentation lesions with multi-spectral images. *Submitted to MedIA*, 2012.

[32] D. Lu, P. Mausel, E. Brondizio, and E. Moran. Change detection techniques. *Int. J. Remote Sensing*, 25, juin 2004.

[33] R.J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam. Image change detection algorithms: a systematic survey. *IEEE Trans. on Image Processing*, 14(3):294 – 307, 2005.

[34] K. Skifstad and R. Jain. Illumination independent change detection for real world image sequences. *Comput. Vis. Graph. Image Process.*, 46:387–399, 1989.

[35] S. Liu, C. Fu, and S. Chang. Statistical change detection with moments under time-varying illumination. *IEEE Trans. Image Process.*, 7:1258–1268, Sep. 1998.

[36] L. Li and M. K. H. Leung. Integrating intensity and texture differences for robust change detection. *IEEE Trans. Image Process.*, 11:105–112, Fev. 2002.

[37] H. V. Poor. *An Introduction to Signal Detection and Estimation, 2nd ed.* New York: Springer-Verlag, 1994.

[38] T. Aach and A. Kaup. Statistical model-based change detection in moving video. *Signal Process.*, 31:165–180, Mar. 1993.

[39] E. Rignot and J. van Zyl. Change detection techniques for ers-1 sar data. *IEEE Trans. Geosci. Remote Sens.*, 31:896–906, Juil. 1993.

[40] L. Bruzzone and D. F. Prieto. An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images. *IEEE Trans. Image Processing*, 11:452–466, Avr. 2002.

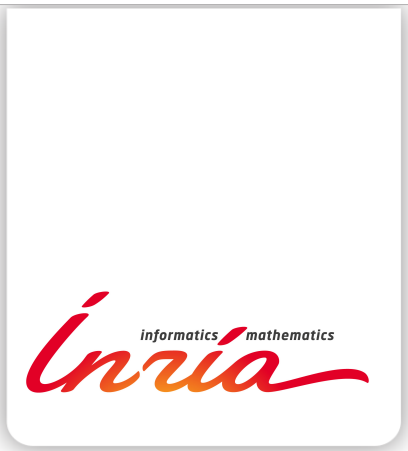
[41] J. Inglada and G. Mercier. A new statistical similarity measure for change detection in multitemporal sar images and its extension to multiscale change analysis. *IEEE Trans. on Geoscience and Remote Sensing*, 45:1432–1445, mai 2007.

[42] Y. Yakimovsky. Boundary and object detection in real world images. *J. ACM*, 23:599–618, Oct. 1976.

[43] K.J. Friston, J. Ashburner, S.J. Kiebel, T.E. Nichols, and W.D. Penny, editors. *Statistical Parametric Mapping: The Analysis of Functional Brain Images.* Academic Press, 2007.

[44] R.J. Alder. *The Geometry of Random Fields.* Wiley, 1981.

[45] J.B. Poline, K.J. Worsley, A.C. Evans, and K.J. Friston. Combining spatial extent and peak intensity to test for activations in functional imaging. *NeuroImage*, 5(2):83–96, 1997.



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