Expert Systems with Applications 40 (2013) 4054–4064

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A two-stage approach for discriminating melanocytic skin lesions using standard cameras

Pablo G. Cavalcanti [a], Jacob Scharcanski [a,*], Gladimir V.G. Baranoski [b]

[a] Instituto de Informática, Universidade Federal do Rio Grande do Sul., Avenida Bento Gonçalves, RS 91501-970, Porto Alegre, Brazil
[b] Natural Phenomena Simulation Group, School of Computer Science, University of Waterloo, Ontario, Canada

ARTICLE INFO

Keywords:
Computer-aided diagnosis
Melanocytic skin lesion
Melanoma
Melanin
Classification
Standard cameras

ABSTRACT

In this paper, we propose a novel approach to discriminate malignant melanomas and benign atypical nevi, since both types of melanocytic skin lesions have very similar characteristics. Recent studies involving the non-invasive diagnosis of melanoma indicate that the concentrations of the two main classes of melanin present in the human skin, eumelanin and pheomelanin, can potentially be used in the computation of relevant features to differentiate these lesions. So, we describe how these features can be estimated using only standard camera images. Moreover, we demonstrate that using these features in conjunction with features based on the well known ABCD rule, it is possible to achieve 100% of sensitivity and more than 99% accuracy in melanocytic skin lesion discrimination, which is a highly desirable characteristic in a prescreening system.

1. Introduction

Malignant melanoma is a type of melanocytic skin lesion, and it is among the most dangerous forms of cancer. According to World Health Organization (2011), approximately 132,000 melanoma cases occur globally each year. The early diagnosis of melanomas is essential for the patient prognosis since most malignant skin lesion cases can be treated successfully in their initial stages. However, a benign atypical melanocytic nevus shares at least some, and sometimes all, of the clinical characteristics of a malignant melanoma, and discriminating benign from malignant cases is often challenging, sometimes in the extreme (Fikrle & Pizinger, 2007; Rao et al., 1997).

It would be desirable that methods for detecting this malignancy approach to 100% of sensitivity, meaning that almost no malignant melanomas are missed. Ideally, this high degree of sensitivity should be coupled with a high degree of specificity, and consequently a high degree of accuracy (Rao et al., 1997). Currently, the most reliable method is the histopathology analysis. However, this requires that the patient undergoes surgical excision. Although this may not be a significant problem in some cases, this is a significant problem for patients with the atypical-mole syndrome, presenting several melanocytic nevi (i.e. 100 nevi or more (Rao et al., 1997), not viable of doing biopsy in all nevi). Moreover, this option is very time consuming, considering the patient and medical staff dislocation, the surgery preparation and the posterior patient monitoring.

To help diagnosing melanocytic skin lesion, physicians often use dermoscopy. This non-invasive technique allows the magnification of submacroscopic structures through the use of an optical device (a dermoscope) and liquid immersion. According to Mayer (1997), the use of dermoscopy can increase the diagnosis sensitivity in 10–27% with respect to the clinical diagnosis. Moreover, some computer-aided diagnosis systems have been proposed in the literature to help the analysis of dermoscopy images. An example is the approach proposed by Celebi et al. (2007), which achieved 92.34% and 93.33% of sensitivity and specificity, respectively. Iyatomi et al. (2008) proposed a web-based system, obtaining a sensitivity of 85.9% and a specificity of 86%. More recently, Ruiz, Berenguer, Soriano, and Snchez (2011) combined three classification algorithms and obtained classification rates of 87.76%. However, even with the assistance of dermoscopy, differentiating malignant and benign lesions is a difficult task. In fact, specialists state that in the early evolution stages of malignant lesions, dermoscopy may not be helpful since it often does not improve the diagnosis accuracy Skvara, Teban, Fiebiger, Binder, and Kittler (2005).

Still considering early stage cases, there are practical situations where a non-specialist (e.g., a physician not trained on dermatology) wishes to have a qualified opinion about a suspect skin lesion, but only standard camera imaging is available on site. In such situations, telemedicine is justifiable, and the non-specialist can capture a macroscopic image of the suspect melanocytic skin lesion and send it to a specialist, who can analyze it in higher detail. In this particular situation, a teledermatology consultation brings

* Corresponding author. Tel.: +55 5133087128.
E-mail addresses: pgcavalcanti@inf.ufrgs.br (P.G. Cavalcanti), jacobs@inf.ufrgs.br (J. Scharcanski), gvgbaran@curumin.cs.uwaterloo.ca (G.V.G. Baranoski).

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.01.002

benefits such as the easier access to health care and faster clinical results (Massone, Wurm, Hofmann-Wellenhof, & Soyer, 2008). Besides these benefits, recent results also suggest that teledermatology tends to be effective and reliable (Whited, 2006).

We also can find in the literature systems designed to assist the diagnosis of melanocytic skin lesions using standard camera images. Manousaki et al. (2006) proposed a system that achieved a sensitivity of 60.9% and a specificity of 95.4% in their experiments. The approach proposed by Tabatabaie, Esteki, and Toossi (2009) resulted in 85% and 92.5% of sensitivity and specificity, respectively. Alcon et al. (2009) described a methodology to obtain a lesion classification with sensitivity of 89.72% and specificity of 75.56%. Cavalcanti and Scharcanski (2011) recently proposed a system even more accurate, achieving a sensitivity of 96.2% and a specificity of 97.7%.

Regardless of the type of image and the devices employed in the data (image) acquisition, the systems proposed to assist the diagnosis of skin lesions usually perform the following operations: (1) preprocessing, in which image artifacts, such as hair or uneven illumination, are eliminated; (2) segmentation, in which the lesion boundaries are determined; (3) feature extraction, in which a quantitative representation for the lesion area is generated; and (4) classification, in which an estimate of whether the lesion is benign or malignant is produced. In general, we can find very different methodologies to perform the first two operations. On other hand, all systems employ the same methodology to perform the last two operations. It consists in reproducing ABCD rule of dermoscopy, a medical criterion that tries to differentiate malignant and benign melanocytic lesions using a set of lesion characteristics, namely Asymmetry, Border irregularity, Color variation and Differential structures.

Beyond the ABCD rule of dermoscopy, there are other dermatological methods, such as the Menzies Scoring Method or the 7-point Checklist (Johr, 2002), that specialists can employ to identify malignant cases. These criteria share some similarities, and seek to determine which features should be visually identified by the physician to perform the diagnosis. However, even experienced dermatologists do not obtain 100% of accuracy, independent of the method in use (Argenziano, Zalaudek, & Soyer, 2004).

In opposite side to these methods of visual analysis, recent studies have been made using features that are not visible to the naked eye. Melanomas are malignant tumors of melanocytes, the cells responsible for the biochemical process that leads to the production of melanin pigments. And, this process (known as melanogenesis) takes place in highly specialized organelles known as melanosomes, and it results in the two main classes of melanin present in the human skin, namely the eumelanin (brown and black) and the pheomelanin (red and yellow). The color of human skin is largely determined by the presence of these pigments, whose relative concentration may vary significantly in skin lesions, notably in melanocytic tumors (Jimbow, Reszka, Schmitz, Salopek, & Thomas, 1995). Accordingly, relevant efforts have been directed toward the identification of variations of eumelanin and/or pheomelanin in melanocytic skin lesions with the purpose of improving the non-invasive diagnosis of these lesions. For example, Marchesini, Bono, and Carrara (2009) performed a retrospective analysis on 1671 lesions using diffuse reflectance spectroscopy. They observed a decrease in pheomelanin and an increase in eumelanin concentration in passing from benign to malignant cases. Their results are consistent with results later obtained by Zonios, Dimou, Carrara, and Marchesini (2010) from reflectance measurements performed on 1379 lesions. However, Zonios et al. observed that the spectral responses obtained from different melanin concentrations do not appear strong enough to provide a definite criterion for the diagnosis of melanocytic skin lesions. However, Zonios et al. also suggested that these responses could be potentially useful for the characterization of melanocytic lesions and the early diagnosis of melanoma if used in combination with other parameters and diagnostic criteria. More recently, Matthews, Piletic, Selim, Simpson, and Warren (2011) proposed a pump–probe imaging system. Examining slices from 42 lesions, they also observed that melanomas have higher amounts of eumelanin in comparison to benign nevi, concluding that melanin features can improve diagnostic accuracy if used in conjunction with current diagnostic techniques.

The main goal of this work is to propose a two-stage classification algorithm that potentially can obtain highly accurate results, and achieve 100% of sensitivity, which is, as already mentioned, a very desirable characteristic for prescreening systems. Firstly, in Section 2 we present the algorithms for the first two initial processing stages: preprocessing and segmentation. In Section 3 is presented the feature extraction process, first computing 52 features to reproduce the ABCD rule of dermoscopy, and then computing 12 additional features that represent the eumelanin and pheomelanin lesion contents. As far as we know, this is the first time that melanin features are used in melanocytic skin lesion image analysis and classification, specially using standard camera images only. In Section 4 we present our proposed two-stage melanocytic skin lesion discrimination scheme, designed to improve on the ABCD rule for malignant cases identification. Section 5 presents a general overview of the processing steps of our melanocytic skin lesion pre-screening scheme, while Section 6 presents experimental results demonstrating how accurate this scheme can be, specially because melanin-based features are used to complement the ABCD rule-based features. Finally, Section 7 presents our conclusions and outlines directions for future work.

2. Skin lesion detection

Before we start quantifying medical or physiological characteristics to differentiate benign from malignant cases, it is crucial that we identify the lesion area in the input image. This task is performed by the segmentation process, which is usually preceded by a image preprocessing step. In the following subsections, these two procedures are detailed.

2.1. Preprocessing

The input image acquired with a standard camera may be affected by illumination artifacts, and if used directly in the segmentation process, shading and lesion regions could be confused. Therefore, shading is attenuated in the input image before the image segmentation.

Our preprocessing step is based on the approach proposed by Cavalcanti, Scharcanski, and Lopes (2010), which is based on the idea that a physician (or a person responsible for taking the lesion picture) focuses a camera on the skin lesion, placing the lesion area in the center of the image, and healthy skin pixels appear at the image corners.

The first step of our shading attenuation method is to convert the image from the original RGB color space to the HSV color space, and retain the Value channel V_{orig}, which presents the higher visibility of the shading effects. We extract 20×20 pixels in each V corner and define S as the union of these four sets. This pixel set with 1600 pixels is used to adjust the quadric function $z(x,y) = P_1(x,y) = P_1x^2 + P_2y^2 + P_3xy + P_4x + P_5y + P_6$, where the six quadric function parameters $P_i (i = 1,\ldots,6)$ are chosen to minimize the quadratic error.

Calculating the quadric function $z(x,y)$ for each image spatial location (x,y), we get an estimate $z(x,y)$ of the local illumination intensity in the image V_{orig}. Dividing the original V_{orig} channel by

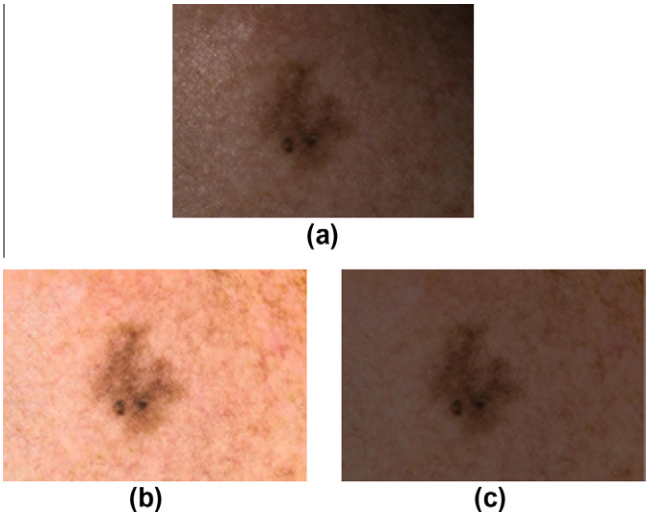

Fig. 1. Illumination and color corrections of a melanocytic skin lesion image. (a) The original image. (b) Fig. (a) after the shading attenuation step. (c) Fig. (b) after the color correction step.

$z(x,y)$, we obtain a new Value channel V_{proc} where the shading effects have been attenuated. The final step consists in replacing V_{orig} by V_{proc}, and converting the image from the HSV color space to the original RGB color space. This shading correction method has limited effect on local cast shadows, and may fail to reduce shading artifacts in areas that are not locally smooth since it is based on a second-order quadric function. However, as can be seen in Fig. 1(a) and (b), this algorithm tends to relight the healthy skin area without loosing lesion information. In this way, we obtain an image where the skin lesion can be segmented more easily.

We consider the healthy skin region surrounding the lesion to compute lesion characteristics, specially the melanin features. The shading attenuation method may generate a highly contrasted image, with a bright skin color, and that may influence negatively the melanin features estimation. Thus, to facilitate the feature extraction stage, we perform one more image preprocessing step, correcting the illumination to obtain more realistic skin colors (more similar to those in the original image), while trying to obtain an even illumination condition. We extract again the Value channels (V from the HSV color space) from the original image (V_{orig}) and from the processed image (V_{proc}), and generate a new Value channel (V_{new}):

$$V_{new}(x,y) = \frac{V_{proc}(x,y) * \mu_{V_{orig}}}{\mu_{V_{proc}}}, \tag{1}$$

where (x,y) indicates the pixel coordinates, and μ_V represents the mean value of the referred V channel. After that, V_{new} replaces V_{proc}, and we obtain an image with even illumination and correct colors representation by converting from the HSV to the RGB color space. Fig. 1 shows an example of this process, presenting an original melanocytic skin lesion image, the result of applying the shading attenuation method in this image, and also shows the image obtained by applying Eq. (1) (to improve the image quality for the feature extraction process).

2.2. Segmentation

Several techniques have been proposed for segmenting melanocytic skin lesions in dermoscopy images (Celebi et al., 2008; Celebi, Iyatomi, Schaefer, & Stoecker, 2009; Gomez, Butakoff, Ersboll, & Stoecker, 2008; Iyatomi et al., 2008; Wighton, Sadeghi, Lee, & Atkins, 2009; Wighton, Lee, Lui, Mclean, & Atkins, 2011). However, dermoscopy obtains skin lesion images with constant illumination. Consequently, these methods do not perform well on standard camera images, and other approaches have been proposed specifically for these images. Usually, the input standard camera image is converted to grayscale and a thresholding-based algorithm is used to identify the lesion area (Alcon et al., 2009; Manousaki et al., 2006; Ruiz et al., 2011). Also, segmentation approaches have been proposed using the original color image obtained with a standard camera (Tang, 2009; Wong, Scharcanski, & Fieguth, 2011). However, these image representations may confuse healthy and lesion pixels due to its color or grayscale similarity. So, in our experiments we used a multichannel image representation that maximizes the discrimination between healthy and unhealthy skin regions as proposed in Cavalcanti and Scharcanski (2011).

As proposed in Cavalcanti and Scharcanski (2011), we create a new 3-channel normalized image $\bar{I}_i^N$ (i.e., $\bar{I}_i^N(x,y) \subset [0,1],\ \forall i$) based on the normalization of the RGB channels $\bar{I}_i^C$ of the input image. The first channel is a representation of the image darkness, relying on the fact that lesion areas are hyper-pigmented skin regions. Each pixel is defined as $\bar{I}_1^N(x,y) = 1 - \bar{I}_1^C(x,y)$, i.e., the complement of the normalized Red channel.

The second channel is a intensity variation representation, since the intensity variability usually is higher in lesions than in healthy skin areas. Being $\bar{L}$ a normalized Luminance image defined by the average of the three $\bar{I}_i^C$ channels, we quantify the textural variability in $\bar{L}(x,y)$ by computing $\tau(x,y,\sigma)$:

$$\tau(x,y,\sigma) = \bar{L}(x,y)\frac{\widetilde{S}(x,y,\sigma)}{S(x,y,\sigma)}, \tag{2}$$

where $S(x,y,\sigma) = \bar{L}(x,y) * G(\sigma)$ (i.e., the Luminance image $\bar{L}$ is smoothed by a Gaussian filter with standard deviation σ), and $\widetilde{S}(x,y,\sigma)$ represents its complement. In this way, if an image region is dark, its intensity variation information is emphasized; if the region is bright, it is de-emphasized. However, since a single Gaussian filter may not be sufficient to capture the textural variability, $\tau(x,y,\sigma)$ is calculated for different σ values,[1] and we select its maximum value at each pixel. Finally, the $\bar{I}_2^N(x,y)$ channel is obtained normalizing these values to the range [0,1].

The third channel $\bar{I}_3^N(x,y)$ of the representation describes the color variation, assuming that healthy and unhealthy skin regions present different color distributions. The Principal Component Analysis (PCA) method is applied on the normalized colors of the image $\bar{I}_i^C(x,y)$. We observed that the first principal component (i.e., the component that maximizes the local data variance) tends to project the healthy skin pixels nearer to zero and lesion pixels tend to have larger magnitudes. So, the $\bar{I}_3^N$ channel is computed normalizing the magnitudes of the first component, and then filtering with a 5×5 median filter to reduce noise.

Once obtained this multichannel representation, the Otsu's thresholding method (Otsu, 1979) is used to segment the image. A pixel is defined as part of a lesion region if its value is higher than the threshold in at least two of the three channels $\bar{I}_i^N(x,y)$. After thresholding, the remaining skin artifacts (such as freckles and hair) are eliminated more easily. The perimeter and the area of all thresholded connected pixel sets are computed, and then this set of regions is partitioned in two clusters. All regions in the cluster with smaller areas (in average) are eliminated, and their correspondent mask pixels are set to zero. At the end, the resultant mask is filtered by a 5×5 median filter, eliminating any possible remaining artifacts that may originate rim imperfections. In Fig. 2, we present the results for all steps of the skin lesion segmen-

[1] We tested different σ values, and based on our experiments we suggest using $\sigma = 1, \frac{11}{7}, \frac{15}{7}, \ldots, \frac{43}{7}$, and filter window sizes of $7\sigma \times 7\sigma$.

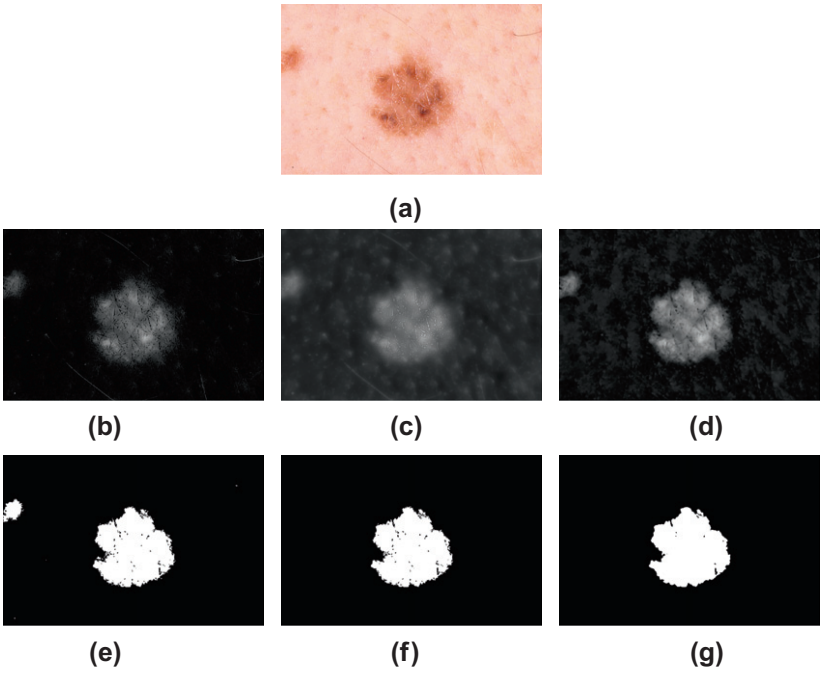

Fig. 2. Segmentation process for the image shown in (a). In (b)–(d), respectively, the $\bar{I}_i^N$ channels representing darkness, texture variation and color variation. In (e)–(g), respectively, binary masks after thresholding, artifacts elimination and filtering.

tation method, including the multichannel representation generation, the thresholding and post-processing steps.

3. Melanocytic skin lesion characterization

The ABCD rule of dermoscopy is the methodology normally used to differentiate benign and malignant melanocytic skin lesions. Therefore, we propose in Section 3.1 a set of 52 features aiming at capturing the ABCD rule of dermoscopy criteria, namely, Asymmetry, Border irregularity, Color variation and Differential structures. As mentioned before, although prescreening systems usually try to reproduce the ABCD rule of dermoscopy criteria, the proposed approaches usually are not able to achieve 100% of sensitivity, as it would be desirable. So, in addition to the ABCD features, we propose 12 features based on the concentrations of eumelanin and pheomelanin, the two kinds of melanin that can potentially be used for the discrimination of melanocytic lesions.

3.1. ACBD rule characterization

The acronym ABCD refers to the criteria commonly used by physicians to differentiate malignant melanomas from benign melanocytic nevi; that is, malignant lesions tend to exhibit **A**symmetry, **B**order irregularity, **C**olor variation, and a **D**iameter greater than 6 mm (Rao et al., 1997). However, we do not adopt any kind of constraint during the image acquisition, and the diameter measurement become impracticable. So, our ABCD rule characterization is based on the ABCD rule of dermoscopy (Nachbar et al., 1994), which changes the D letter to **D**ifferential structures, submacroscopic morphologic and vascular structures (e.g., pigment network, dots, globules) that occur more frequently in malignant cases. And, to perform this characterization, we extract 52 features as described next. It is important to observe that the $\bar{I}_i^N$ channels refer to the same channels that have been created during the segmentation process (see Section 2.2).

We start by quantifying the lesion shape, characterizing the asymmetry and irregularity characteristics of the lesion. Let A, Ac, Ab be, respectively, the areas of the segmented lesion, the convex hull and the bounding box. Moreover, p is the lesion perimeter, $L1$ the major axis of the lesion aligned with its longest diameter, and L_2 is the minor axis orthogonal to L_1. Considering these measurements, we consider the following 11 parameters (Alcon et al., 2009; Cavalcanti & Scharcanski, 2011; Celebi et al., 2007; Fikrle & Pizinger, 2007): A/Ac (solidity), A/Ab (extent), $4A/(L_1\pi)$ (equivalent diameter), $4\pi A/(L_1 p)$ (circularity), p/A, ratio between the principal axes (L_2/L_1), ratio between sides of the axis L_1 and of the axis L_2, the difference between the areas of each side of the axes L_1 and L_2 normalized by the whole area A, and ratio between sides of the lesion bounding box.

The boundary sharpness is quantified by the magnitude of the gradient using the Sobel operator. We dilate the lesion rim by 2 pixels, obtaining a 5 pixels wide region, and compute the following six features: the average and the variance of the gradient magnitudes of all pixels in the extended rim in each one of the three $\bar{I}_i^N$ channels.

To reproduce closely the application of the ABCD rule in terms of asymmetry and border irregularity, the next features are based on a sub-division of the lesion rim in eight symmetric regions (Nachbar et al., 1994). In addition to the two principal axes L_1 and L_2, we rotate these orthogonal axes by 45° and obtain two additional axes, obtaining eight symmetric regions $R = 1,\ldots,8$. For each channel $\bar{I}_i^N$, the average gradient magnitudes of the extended rim pixels $\mu_{R,i}$ $(R = 1,\ldots,8)$ are computed. Therefore, in this way we calculate six more features: the average and the variance of the 8 $\mu_{R,i}$ gradient magnitude values in each one of the three $\bar{I}_i^N$ channels.

To represent the C letter of the ABCD rule, we start by computing maximum, minimum, mean and variance of the intensities in the color variation channel $\bar{I}_3^N$. The same four features are also computed for each one of the three original $\bar{I}_i^C$ channels. We also consider the ratios between the mean values: $mean\left(\overline{I_1^C}\right)/mean\left(\overline{I_2^C}\right)$, $mean\left(\overline{I_1^C}\right)/mean\left(\overline{I_3^C}\right)$ and $mean\left(\overline{I_2^C}\right)/mean\left(\overline{I_3^C}\right)$, totalizing 19 features.

Table 1
Six possible colors of a lesion on the RGB color space (Alcon et al., 2009).

Color	Red	Green	Blue
White	1	1	1
Red	0.8	0.2	0.2
Light Brown	0.6	0.4	0
Dark Brown	0.2	0	0
Blue–Gray	0.2	0.6	0.6
Black	0	0	0

Furthermore, physicians usually identify six distinct hues in skin lesions: white, red, light and dark brown, blue–gray, and black (Nachbar et al., 1994). So, we compute the occurrence of these typical hues within the lesion. Given a pixel in the lesion segment, we find the nearest reference color (associated with a typical hue, see Table 1 (Alcon et al., 2009)) using the Euclidean distance to the pixel color in $\overline{I_i^C}$. A hue occurrence counter is created, one cell per typical hue. For each lesion pixel, the nearest typical hue counter is increased by 1. Finally, typical hue counters are normalized/divided by the lesion area A, and generate 6 additional features.

Finally, we quantify the differential structures. The differential structure characteristics are not easily discriminated based on macroscopic images, however these structures tend to generate different intensity variations in benign and malignant lesions. So, we extract four features, namely the maximum, minimum, mean and variance of the intensity variation channel $\overline{I_2^N}$, considering only the pixels inside the lesion segment.

3.2. Melanin content variation characterization

As already mentioned in Section 1, recent studies have analyzed the melanin variation in melanocytic skin lesion to differentiate benign and malignant cases through spectroscopy. In this work, we follow a different approach. Instead of using spectroscopy, we propose to estimate the melanin variation in skin lesions using standard camera images. In order to perform this task, we employ a biophysically-based spectral model of light interaction with human skin developed by Krishnaswamy and Baranoski (2004) and Baranoski and Krishnaswamy (2010) known as BioSpec (see more details about BioSpec in Appendix A). This model simulates the reflectance spectra of skin specimens within the visible spectral domain (i.e., from 400 to 700 nm). It takes into account the specimens' biophysical characteristics described through a set of biophysical parameters. In this work, unless otherwise stated, we employed the default values assigned for these parameters in the online version of the model (Natural Phenomena Simulation Group, 2011). This set of parameters includes the concentration of eumelanin and pheomelamin. Although BioSpec was developed for healthy skin tissue, we believe that by varying the concentration of eumelanin and pheomelanin we can obtain sound estimates for the reflectance spectra of melanomas since the main factor responsible for their distinct spectral appearance, notably in their initial stages, is their high concentration of melanins (Lazova & Pawelek, 2009).

In Fig. 3, we present a plate showing the skin colors that were obtained by converting the reflectance spectra provided by BioSpec given a wide range of parameter values (see Appendix B for more details about how this conversion is performed). We varied the concentration of eumelanin from 20 to 300 g/L, and the concentration of pheomelanin from 4 to 60 g/L, both with steps of 4 g/L, and used the default values for the other model parameters. The choice of using eumelanin predominantly in these simulations is supported by optical scanning studies in which the spectrum of malignant melanomas tend to show an eumelanin predominance in their spectral characteristics (Dimitrow et al., 2009; Zonios et al., 2008). Thus, we used 71 and 15 values for the eumelanin and pheomelanin parameters, respectively, obtaining 1065 skin colors with BioSpec. The colors appearing in the lesion images (photographs) are matched with the skin colors derived from the reflectance spectra provided BioSpec, and the estimates of the eumelanin and pheomelanin parameters are obtained.

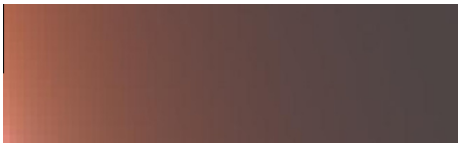

Fig. 3. The possible colors of skin, varying the concentration of eumelanin from 20 to 300 g/L (in the horizontal axis) and the concentration of pheomelanin from 4 to 60 g/L (in the vertical axis).

3.2.1. Extracting melanin variation features from skin lesion images captured with standard cameras

Given a pixel, its color is compared to each one of the 1065 possible modeled colors obtained using BioSpec. This comparison, consists of using the minimum Euclidean distance to find the modeled color nearest to the pixel color. This information allows us to estimate the concentration of eumelanin and pheomelanin in the skin location that corresponds to this pixel. To compare the variation of the pigment concentrations within the lesion with the concentrations in healthy skin, we also extract these concentration values in the healthy skin in the surround area outside the lesion. In order to locate the healthy skin in the area surrounding the lesion, we dilate the binary mask obtained in the segmentation step with a disk structuring element of 30 pixels of radius. Then, we subtract the original segmented lesion binary mask from the dilated segmentation binary mask, obtaining just the region surrounding the lesion with a width of approximately 30 pixels. An example of this operation can be seen in Fig. 4. Located the healthy skin region surrounding the lesion, we estimate the concentrations of eumelanin and pheomelanin in each inner pixel in the way described above. Given a healthy skin pixel, we also estimate these values by minimum Euclidean distance to the 1065 modeled colors as mentioned before.

After computing the estimates of the eumelanin and pheomelanin concentrations associated with a set of pixels of the given image, we extract the following melanin features mf_i:

- mf_1: Average of the eumelanin concentration of all pixels inside the lesion segment;
- mf_2: Average of the eumelanin concentration of all pixels inside the lesion segment as well as all pixels inside the healthy skin surrounding area;
- mf_3: Standard deviation of the eumelanin concentration of all pixels inside the lesion segment;
- mf_4: Standard deviation of the eumelanin concentration of all pixels inside the lesion segment as well as all pixels inside the healthy skin surrounding area;
- mf_5: The difference between the average of the eumelanin concentration inside the lesion segment and the average of

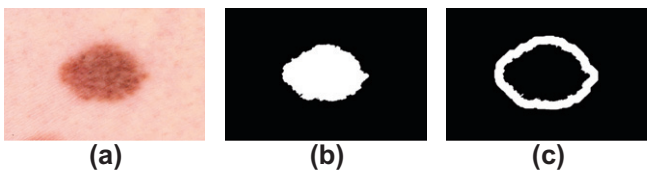

Fig. 4. Examples of the binary masks used in the melanin features computation. (a) The original image. (b) The lesion segmentation binary mask. (c) The binary mask corresponding to the healthy skin area surrounding the lesion.

the eumelanin concentration inside the healthy skin surrounding area;

mf_6: The ratio between the average of the eumelanin concentration inside the lesion segment and the average of the eumelanin concentration inside the healthy skin surrounding area;

We also consider six more pheomelanin features mf_7–mf_{12}, similar to mf_1–mf_6, but instead of using eumelanin concentration we use the pheomelanin concentration.

4. Proposed two-stage melanocytic skin lesion discrimination scheme

Once the features that characterize the melanocytic lesions are extracted, we use these information to discriminate the malignant and the benign cases. The following Sections detail our two-stage discrimination approach, first using the ABCD rule-based features, and next refining by the discrimination results using the melanin variation features. The melanocytic skin lesion discrimination approaches proposed in the literature usually contain a single discrimination step based on the ABCD rule, and as already mentioned in Section 1, are not able to obtain the desirable 100% of sensitivity. So, we add a second discrimination stage to our scheme, trying to correct cases that incorrectly received a benign label in the initial stage. The melanin variation features improve the discrimination results obtained with the ABCD rule (first stage), helping achieve a higher degree of sensitivity (i.e. so virtually no malignant case would be missed).

4.1. First stage: melanocytic skin lesion discrimination using the ABCD rule

Based on the 52 ABCD features that have been extracted (see Section 3.1), we generate a preliminary discrimination of benign and malignant melanocytic skin lesions. However, classifiers tend to be more efficient if these feature values are normalized (i.e. produce values within a specified range). We normalize feature values with the z-score transformation (Aksoy & Haralick, 2000; Celebi et al., 2007):

$$Z_{i,j} = \frac{((v_{i,j} - \mu_j)/(3\sigma_j) + 1)}{2}, \tag{3}$$

where $v_{i,j}$ is the value of the jth feature of the ith sample (image), μ_j and σ_j are the mean and standard deviation of the jth feature, respectively, of all training set samples. After the z-score transformation, most of the $Z_{i,j}$ values are in the $[0,1]$ range. The out-of-range values are saturated to either 0 or 1.

To create the training set, half of the samples in each class (benign or malignant) are randomly selected. Additionally, since we have a relatively small image data set, new samples are added to the training set using the Smoothed Bootstrap Resampling method (i.e., original samples are randomly selected, and new ones are created adding a small amount of zero-centered noise to their feature values, enlarging the data set (Young, 1990)). In our experiments, we used zero mean Gaussian noise with $\sigma = 0.1$, obtaining from 500 to 2500 samples for each class so the classes are balanced (i.e., have the same number of samples). Using Normal probability plots, we observed that Normality is a reasonable assumption for our features, and by introducing small Gaussian noise we avoid "ties" that may lead to overfitting in random oversampling procedures.

Finally, we apply a classifier to determine the class (benign or malignant) of each sample/image. We used a K-Nearest Neighbor Classifier (KNN) with $K = 1$, or simply a Nearest Neighbor Classifier, because this is a very simple method where each sample/image is assigned to its closest neighboring class in Euclidean (feature) space. Besides its simplicity, this method has been used in melanocytic skin lesion image classification research (Burroni et al., 2004; Dreiseitl et al., 2001), producing very accurate results (Cavalcanti & Scharcanski, 2011). Also, this classifier also has been successfully used for the classification of non-melanoma skin lesions (Ballerini, Fisher, Aldridge, & Rees, 2012a, 2012b).

4.2. Second stage (beyond the ABCD rule): melanocytic skin lesion discrimination with melanin variation features

A desirable characteristic for prescreening systems is to achieve 100% of sensitivity, i.e. to not identify mistakenly any malignant case as benign. However, we observed in our experiments that considering the proposed melanin variation features mf_1–mf_{12} in the first classification stage, i.e. when using the KNN classifier, is not efficient to achieve this goal (i.e. of obtaining high sensitivity). So, in order to reach high sensitivity rates, we propose to include a second classification stage. After obtaining the first melanocytic skin lesion classification result, we re-classify all samples/images that have been initially identified as benign using the melanin variation features mf_1–mf_{12}. This second stage is based on the Bayes's Classifier (Alpaydin, 2004) (which provided the best classification results in our experiments), and assigns a sample to the most likely class based on the extracted features. Let L be a set of melanin features $[mf_1,\ldots,mf_{12}]$ extracted from a lesion, $D = 1$ the malignant class and $D = 0$ the benign class. Using Bayes' rule, the posterior probability $P(D|L)$ of assigning a lesion to class D can be written as:

$$P(D|L) = \frac{P(L|D)P(D)}{P(L)}, \tag{4}$$

where, $P(L|D)$ is the conditional probability of finding a lesion with the L characteristics in class D, $P(D)$ is the prior probability of class D, and $P(L)$ is the evidence and it represents the probability that the features L are seen. To the best of our knowledge, $P(D)$ is not known, and considering that our training sets have an equal number of samples per class, we consider $P(D = 0) = P(D = 1)$ and this term can be discarded from Eq. (4). Since $P(L)$ is constant with respect to $argmax_D(P(L|D)/P(L))$ and therefore has no effect on the solution, it can be also discarded, and a lesion can be considered malignant if:

$$P(L|D = 1) > P(L|D = 0) \tag{5}$$

and it is benign otherwise. This classification method is usually referred to as the Maximum Likelihood Classifier. So, the challenge is to properly estimate $P(L|D)$ in a way that minimizes the probability classification error. To perform this task and generate accurate classification results, we propose the following discrimination scheme:

- To generate the training set, we use half of the samples of each class (the same samples that have been randomly selected to train the first discrimination stage). Since the number of sample is small, specially in the benign class, we use again Smoothed Bootstrap Resampling (see Section 4.1) to generate 2500 samples for each class (i.e., benign and malignant), adding Gaussian noise ($\mu = 0$ and $\sigma = 0.1$) to the original mf_i features. In this way, we obtain more statistical significant data;
- Given the training set and a subset ψ containing the mf_i features, we generate two multivariate probabilities density functions (pdfs): one for the benign class $P(L|D = 0)$ and one for the malignant class $P(L|D = 1)$. We model these pdfs using an unsupervised modified EM (Expectation–Maximization) algorithm to identify finite Gaussian mixture models as proposed by Figueiredo and Jain (2002);
- Each sample L that has been previously classified as benign is re-tested, and it is re-classified by a Maximum Likelihood Clas-

sifier as malignant if its ψ features indicate a higher probability of being malignant (see Eq. (5)).

We tested this two-stage discrimination scheme with different ψ sets of features, and the obtained results are presented in Section 6.

5. Method overview

In the last three Sections we presented all the steps of the proposed two-stage discrimination scheme, and Fig. 5 presents the overall system workflow used to identify a melanocytic skin lesion as benign or malignant. We start by preprocessing the input image, i.e. attenuating the shading effects, and segmenting the skin lesion. After that, the first discrimination stage is performed, using the 52 features based on the ABCD rule and a KNN classifier, assigning a benign/malignant label to the imaged lesion.

The second discrimination stage tries to enhance as much as possible the melanocytic skin lesion classification sensitivity. All images that have been initially classified as benign are re-classified now using the melanin variation features extracted from inside the lesion and from the healthy skin area surrounding the lesion. The final classification result is provided by a Maximum Likelihood Classifier.

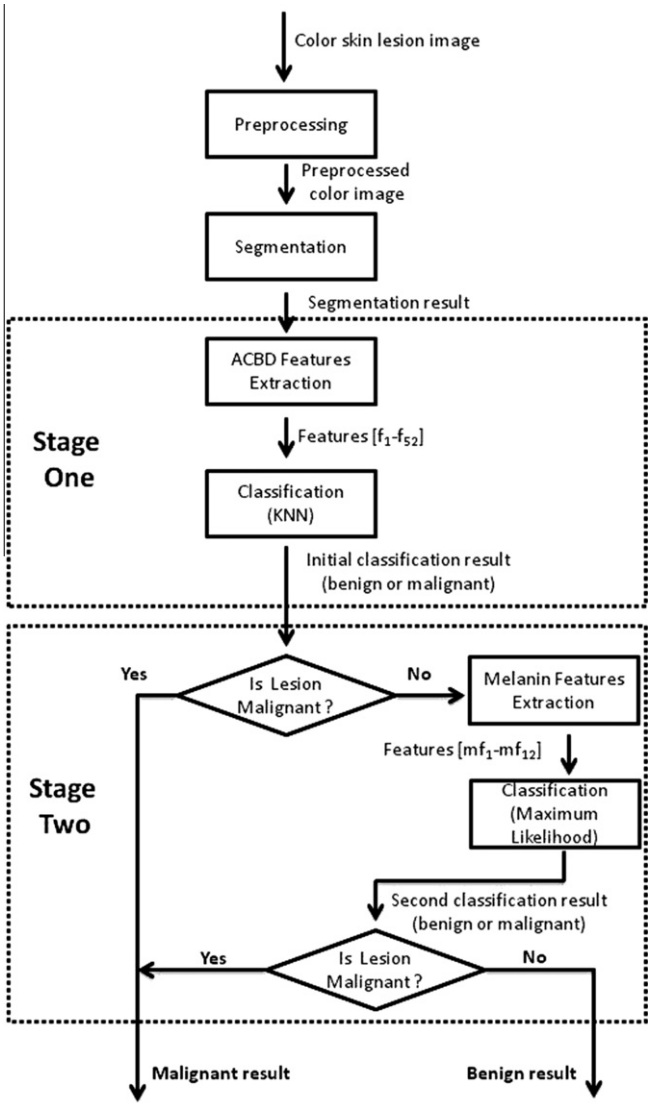

Fig. 5. Overall system workflow.

6. Experimental results

We performed our experiments on the same image dataset that have been used in the experiments of Alcon et al. (2009). They collected an image set containing 45 benign atypical nevi (or Clark nevi) and 107 melanomas from the Dermnet dataset (Dermnet Skin Disease Image Atlas, 2009).

Considering that training our proposed two-stage discrimination scheme requires a random resampling step, we repeated the whole process 50 times, seeking to obtain a training set that characterizes well the two classes, i.e., the benign and malignant cases. Furthermore, we tested different ψ sets of melanin variation features (see Section 4.2), trying to identify the feature set that obtains less false negatives. We started the selection of the ψ feature sets by taking each mf_i feature individually, and then increased the set size by adding one more feature at a time, until we reached six features per set, when we observed that the obtained results were decreasing in quality. We tested all the possible feature set combinations: all the 12 features individually, the 66 combinations of two features, the 220 combinations of three features, the 495 combinations of four features, the 792 combinations of five features, and 924 combinations of six features. In Table 2 we present the ψ feature sets considering these number of features that resulted in the best classification results, i.e. the sets that obtained 100% of sensitivity associated with high specificity, and using the number of times that results was obtained as a tie-breaker. We also tested the 6 eumelanin features, the six pheomelanin features, and the combination of all 12 melanin variation features as three feature sets, and the corresponding results also are presented in Table 2.

Considering that for each feature set ψ we generate randomly 50 training and testing sets by Smoothed Bootstrap Resampling, we only present the average results, the best obtained results in terms of accuracy and the number of times that these results were obtained (when testing with the 50 different sets). However, although we use randomly generated training sets to improve the chance of characterizing well the lesion classes, it shall be observed that we only use the original data (and not the synthetic data) for testing.

It is important to observe that in the experiments of Alcon et al. (2009) with the same dataset, but not including bootstrapping in their training step, the best obtained results were: sensitivity of 89.72% and specificity of 75.56%. Later, Cavalcanti and Scharcanski (2011) included the bootstrapping step, and presented a system even more accurate in this dataset, achieving a sensitivity of 96.2% and a specificity of 97.7%. So, comparing the results of these approaches with our results shown in Table 2, we may conclude that our proposed discrimination scheme potentially can enhance the sensitivity (i.e., reduce the number of false negatives), while maintaining or even increasing the accuracy, although the specificity may decrease slightly. The set of features containing all 12 melanin variation features obtained very good results in average, and achieved 100% of sensitivity in 11 of the 50 test sets. However, the best classification results have been obtained using the feature set $\psi = \{mf_6, mf_8, mf_{10}, mf_{12}\}$. This set containing one feature based on the eumelanin concentration and three based on the pheomelanin concentration achieved 100% of sensitivity in 24 of the 50 test sets, combined with high specificity and, consequently, high accuracy.

Important to observe the presence of mf_6 and mf_{12} (i.e., the ratio between the eumelanin concentration inside the lesion and in the healthy skin, and the ratio between the pheomelanin concentra-

Table 2
Discrimination results in 50 trials.

		Sensitivity (%)	Specificity (%)	Accuracy (%)	Number of times the result was obtained
$\psi = \{mf_{12}\}$	Average	100	73.02	92.01	–
	Best result	100	80	94.08	1
$\psi = \{mf_{10}, mf_{12}\}$	Average	99.83	85.87	95.70	–
	Best result	100	95.56	98.68	2
$\psi = \{mf_6, mf_{10}, mf_{12}\}$	Average	99.68	90.76	97.04	–
	Best result	100	97.78	99.34	6
$\psi = \{mf_6, mf_8, mf_{10}, mf_{12}\}$	**Average**	**99.70**	**96.18**	**98.66**	–
	Best result	**100**	**97.78**	**99.34**	**24**
$\psi = \{mf_6, mf_8, mf_{10}, mf_{11}, mf_{12}\}$	Average	99.57	95.96	98.50	–
	Best result	100	97.78	99.34	16
$\psi = \{mf_3, mf_6, mf_8, mf_{10}, mf_{11}, mf_{12}\}$	Average	99.64	94.31	98.07	–
	Best result	100	97.78	99.34	9
$\psi = \{mf_1, \ldots, mf_6\}$	Average	99.35	90.18	96.63	–
	Best result	99.07	97.78	98.68	8
$\psi = \{mf_7, \ldots, mf_{12}\}$	Average	99.66	92.98	97.68	–
	Best result	100	97.78	99.34	4
$\psi = \{mf_1, \ldots, mf_{12}\}$	Average	99.48	96.31	98.54	–
	Best result	100	97.78	99.34	11

Discrimination results in 50 trials (best results are in bold face).

tion inside the lesion and in the healthy skin) in the feature set that obtained the best classification results. That indicates that although the concentration of melanin inside the lesion is an important feature to discriminate malignant and benign cases, it is also important to consider the characteristics of the healthy skin to identify if the lesion pigmentation variation is comparable to a melanoma or not.

We also analyzed the performance of each classification stage of our approach separately. We observed that when our system achieves its best result (i.e., sensitivity of 100%, specificity of 97.78% and accuracy of 99.34%), the first stage which uses the ABCD features only generated a sensitivity of 96.26% and specificity of 97.78%, and a total accuracy of 96.71%. Considered independently, the second stage which is based on melanin features, obtained a lower sensitivity to 94.39%, high specificity of 100%, and 96.05% of accuracy. However, it is important to recall that we only use this second stage for the samples that have been previously classified as benign (as a verification stage). Consequently, our two-stage approach achieves 100% of sensitivity and a very high level of accuracy (i.e. 99.34%). Although the second stage individually obtains worse results in terms of sensitivity and accuracy, our experiments demonstrate that the melanin features potentially can help identify melanomas that have been erroneously classified as benign initially.

It shall be observed that our results were obtained using the 1065 different skin colors, that have been modeled using the average characteristics of human skin. Our classification results suggest that our approach potentially could estimate the melanin variation in different types of skin. Despite of the fact that our image set does not retain any information about patients, the variability of skin characteristics is noticeable in this set. However, this dataset does not include melanocytic lesions occurring in dark (black) skin, and computing a different color table may be necessary for these skin tones. The very low incidence of melanomas (melanocytic lesions) in dark (black) skin, and the focus on the most common cases, justifies the absence of such examples of in our dataset. In fact, the incidence of melanomas in dark (black) skin is a rare phenomenon, and recent statistics show this incidence to be more than 20 times lower than in other skin tones (Howlader et al., 2012).

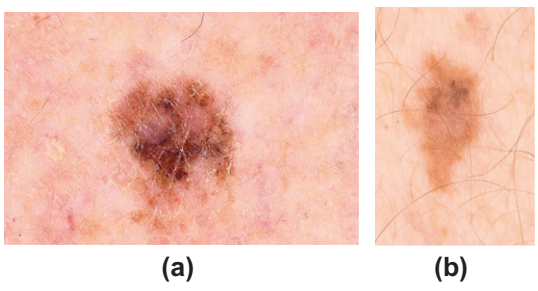

Fig. 6. Examples of lesions that have been previously identified as benign and then were correctly re-classified as malignant. These images are displayed with their illumination corrected for better visualization.

In Fig. 6, we illustrate two examples of melanocytic skin lesions that have been incorrectly identified as benign nevi in the first classification stage, and then were correctly classified as malignant melanomas in the second classification stage using the proposed melanin variation features.

7. Conclusions

According to recent studies (Marchesini et al., 2009; Matthews et al., 2011; Zonios et al., 2010), the concentration of the two main classes of melanin present in human skin, eumelanin and pheomelanin, can potentially be used to identify melanocytic skin lesions as benign or malignant if employed in conjunction with current diagnostic techniques. Based on these observations, we presented in this paper a new method that uses only a standard camera image (a simple photograph) to compute classification features based on the estimation of eumelanin and pheomelanin contents in melanocytic skin lesions, and employs these melanin features in a two-stage classification scheme.

In medical applications, it is especially desirable that a prescreening system achieves 100% of sensitivity, i.e., all malignant cases are correctly classified and no false negatives occur. However, as far as we know, this goal still poses a challenge for researches. With this objective in mind, we proposed a scheme that starts with a discrimination stage based on the ABCD rule, which are well known medical criteria that prescreening systems usually try to implement. Next, a second discrimination stage based on melanin variation features is used to help improve the

discrimination sensitivity and accuracy. Our preliminary experimental results suggest that the proposed framework which combines ABCD rule-based and melanin variation features potentially can be very accurate and obtain high sensitivity values (near 100% in our experiments). In addition, we remark that although the proposed framework is designed for pre-screening purposes, which ideally should involve melanocytic skin lesions (e.g., melanomas) in their initial stages, our experiments were performed on images depicting melanomas in different stages of development. Hence, the high level accuracy of our preliminary results also illustrates the robustness of the proposed framework with respect to different input data (skin images).

As future work, we plan to test our framework more extensively in clinical trials, and make it part of a complete teledermatological system. As more biophysical data on melanomas becomes available, we also intend to further enhance its predictive capabilities, notably with respect of melanomas in advanced stages, by incorporating into our simulations other factors affecting the spectral signatures of melanocytic lesions such as abnormalities in the distribution and morphology of the melanosomes. Nevertheless, other features could improve the reproduction of the ABCD rule and we intend to develop and test it according to necessity. For example, fuzzy counters to represent the color presence (i.e., assigning weight distances to each reference color) and textural descriptors trying to identify the dermatological structures may possibly contribute for the differentiation of malignant and benign cases.

Acknowledgments

We would like to thank CNPq (Brazilian National Council for Scientific and Technological Development) for partially funding this project. Also, we thank the anonymous reviewers for their valuable contributions that helped improve this work.

Appendix A. BioSpec: a spectral model of light interaction with human skin

For completeness, we summarize the main characteristics of the BioSpec model in this section. We remark that detailed descriptions of BioSpec can be found in the related publications (Baranoski & Krishnaswamy, 2010; Krishnaswamy & Baranoski, 2004) and its source code is openly available online (Natural Phenomena Simulation Group, 2011).

The BioSpec model uses Monte Carlo based algorithms to simulate the processes of light propagation (surface reflection, subsurface reflection and transmission) and absorption in the skin tissues. It considers the stratification of skin into four semi-infinite main layers: stratum corneum, epidermis, papillary dermis, and reticular dermis. The model parameter space includes: the refractive index and thickness of each layer, the refractive index and the diameter of collagen fibrils, the extinction coefficient, concentration, and volume fraction of the main chromophores present in the skin tissues (e.g., eumelanin, phaeomelanin, oxyhemoglobin, deoxyhemoglobin, β-carotene and bilirubin) and the aspect ratio of the stratum corneum folds.

The propagation of light in the skin tissues is simulated by the BioSpec model as a random walk process (that rely on the generation of random numbers ξ_j, for $j = 1,2,\ldots,9$, uniformly distributed in the interval $[0,1]$) using ray optics. In this random walk process, the transition probabilities are associated with Fresnel coefficients computed at each interface between the layers, and the termination probabilities are determined by the ray free path length.

Once a ray hits a skin specimen at the air/stratum corneum interface, it can be reflected back or refracted into the stratum corneum. In the former case, the BioSpec model computes the distribution of the reflected light taking into account the aspect ratio (or oblateness) of the stratum corneum folds. In the BioSpec model, these mesostructures are represented as ellipsoids. The aspect ratio, $\varsigma \in [0,1]$, of these ellipsoids is defined as the quotient of the length of the vertical axis by the length of the horizontal axis, which are parallel and perpendicular to the specimen's normal respectively. As the folds become flatter (lower ς), the reflected light becomes less diffuse. In order to account for this effect, the reflected rays are perturbed using angular displacements obtained from the surface-structure function proposed by Trowbridge and Reitz (1975), which represents rough air-material interfaces using microareas randomly curved. These displacements are given in terms of the polar perturbation angle:

$$\theta_s = \arccos\left[\left(\left(\frac{\varsigma^2}{\sqrt{\varsigma^4 - \varsigma^4\xi_1 + \xi_1}} - 1\right)b\right)^{\frac{1}{2}}\right], \tag{A.1}$$

where b corresponds to $1/(\varsigma^2 - 1)$, and the azimuthal perturbation angle, ϕ_s, is given by $2\pi\xi_2$.

If the ray penetrates the skin specimen, then it can be reflected and refracted multiple times within the skin layers before it is either absorbed or propagated back to the environment through the air/stratum corneum interface. Since the subcutaneous tissue is a highly reflective medium, for body areas characterized by the presence of hypodermis, the BioSpec model assumes total reflection at the reticular dermis/hypodermis interface. In the epidermis and stratum corneum, scattering is simulated using angular displacements measured by Bruls and van der Leun (1984).

Every ray entering one of the dermal layers is initially tested for Rayleigh scattering. If the test fails or the ray has already been bounced off one of the dermal interfaces, then the ray is randomized around the normal direction using a warping function based on a cosine distribution in which the polar perturbation angle, α_c, and the azimuthal perturbation angle, β_c, are given by:

$$(\alpha_c, \beta_c) = \left(\arccos((1-\xi_3)^{\frac{1}{2}}), 2\pi\xi_4\right). \tag{A.2}$$

In order to perform the Rayleigh scattering test, the spectral Rayleigh scattering amount, $\mathcal{S}(\lambda)$, is computed using the appropriate expression for Rayleigh scattering involving particles (McCartney, 1976). Next, a random number ξ_5 is generated. If $\xi_5 < 1 - e^{-\mathcal{S}(\lambda)}$, then the ray is scattered using an azimuthal perturbation angle, β_R, given by $2\pi\xi_8$, and a polar perturbation angle, α_R, obtained using the following rejection sampling algorithm based on the Rayleigh phase function (McCartney, 1976):

do

$$\alpha_R = \pi\xi_6$$

$$\chi = \frac{3}{2}\xi_7$$

while $\left(\chi > \frac{3\sqrt{6}}{8}(1+\cos^2\alpha_R)\sin\alpha_R\right)$

Once a ray has been scattered, it is tested for absorption. The absorption testing is performed probabilistically every time a ray starts a run in a given layer, and it consists in estimating the ray free path length using an expression based on Beer's law (Tuchin, 2000), and considering the total absorption coefficient, $\mu_{ai}(\lambda)$, of a given layer i, which takes into account the extinction coefficient and the concentration of each pigment present in this layer. This test consists in estimating the ray free path length, $p(\lambda)$, through the following expression:

$$p(\lambda) = -\frac{1}{\mu_{ai}(\lambda)}\ln(\xi_9)\cos\theta, \tag{A.3}$$

where θ corresponds to the angle between the ray and the specimen's normal. If $p(\lambda)$ is greater than the thickness of the layer, then the ray is propagated, otherwise it is absorbed.

Appendix B. Converting the modeled spectral reflectance values to skin colors

Unfortunately, we cannot directly access the original reflectance values associated with the pixels of standard camera images. Therefore, we first transform the modeled reflectance functions $R(\lambda)$ (in our experiments, obtained using BioSpec) to colors that can be matched straightforwardly with the lesion colors appearing in a standard image. The color tristimulus values XYZ can be calculated by the additive law of color matching (Lee, 2005):

$$X = N\sum_{\lambda} R(\lambda)\ S(\lambda)\ \bar{x}(\lambda)\ \Delta\lambda, \tag{B.1}$$

$$Y = N\sum_{\lambda} R(\lambda)\ S(\lambda)\ \bar{y}(\lambda)\ \Delta\lambda, \tag{B.2}$$

$$Z = N\sum_{\lambda} R(\lambda)\ S(\lambda)\ \bar{z}(\lambda)\ \Delta\lambda, \tag{B.3}$$

where $S(\lambda)$ is the relative spectral power distribution of the illuminant, $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ are the spectral sensitivity functions, and $\Delta\lambda$ represents the wavelength intervals. In our experiments, we used the CIE Standard Illuminant D65 (Commission Internationale de l'Eclairage, 2011), the CIE 1931 sensitivity functions of the standard observer for 2°, and $\Delta\lambda = 5nm$. The constant N is defined as follows:

$$N = \sum_{\lambda} S(\lambda)\ \bar{y}(\lambda)\ \Delta\lambda. \tag{B.4}$$

Obtained the color tristimulus XYZ, we convert XYZ to the RGB space using the sRGB standard (Lee, 2005):

$$R' = 3.2410X - 1.5374Y - 0.4986Z, \tag{B.5}$$

$$G' = -0.9692X + 1.8760Y + 0.0416Z, \tag{B.6}$$

$$B' = 0.0556X - 0.2040Y + 1.0570Z. \tag{B.7}$$

As shown in Lee (2005), to obtain the RGB colors in the correct range, an additional non-linear transform is required. If $R',G',B' \leqslant 0.003040$, then we obtain the corrected $\overline{RGB}$ color representation by:

$$\overline{R} = 12.92R', \tag{B.8}$$

$$\overline{G} = 12.92G', \tag{B.9}$$

$$\overline{B} = 12.92B', \tag{B.10}$$

otherwise,

$$\overline{R} = 1.055R'^{(1/2.4)} - 0.055, \tag{B.11}$$

$$\overline{G} = 1.055G'^{(1/2.4)} - 0.055, \tag{B.12}$$

$$\overline{B} = 1.055B'^{(1/2.4)} - 0.055. \tag{B.13}$$

Finally, the obtained $\overline{RGB}$ are normalized, i.e., are transformed to the range [0,1], and then the correct eight-bit values RGB are obtained by multiplying $\overline{RGB}$ by 255.

References

Aksoy, S., & Haralick, R. M. (2000). Probabilistic vs. geometric similarity measures for image retrieval. In *Proc. IEEE conference on computer vision and pattern recognition* (Vol. 2, pp. 357–362).

Alcon, J. F., Ciuhu, C., ten Kate, W., Heinrich, A., Uzunbajakava, N., Krekels, G., et al. (2009). Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis. *IEEE Journal of Selected Topics in Signal Processing, 3*, 14–25.

Alpaydin, E. (2004). *Introduction to machine learning (adaptive computation and machine learning)*. The MIT Press.

Argenziano, G., Zalaudek, I., & Soyer, H. (2004). Which is the most reliable method for teaching dermoscopy for melanoma diagnosis to residents in dermatology? *British Journal of Dermatology, 151*, 512–513.

Ballerini, L., Fisher, R. B., Aldridge, B., & Rees, J. (2012). Non-melanoma skin lesion classification using color image data in a hierarchical k-nn classifier. In *ISBI* (pp. 358–361).

Ballerini, L., Fisher, R. B., Aldridge, B., & Rees, J. (2012a). A color and texture based hierarchical k-nn approach to the classification of non-melanoma skin lesions. In M. E. Celebi & G. Schaefer (Eds.), *Color medical image analysis. Lecture notes in computational vision and biomechanics* (Vol. 6). Springer.

Baranoski, G., & Krishnaswamy, A. (2010). *Light & skin interactions: Simulations for computer graphics applications*. Morgan Kaufman.

Bruls, W., & van der Leun, J. (1984). Forward scattering properties of human epidermal layers. *Photochemistry and Photobiology, 40*, 231–242.

Burroni, M., Corona, R., DellEva, G., Sera, F., Bono, R., Puddu, P., et al. (2004). Melanoma computer-aided diagnosis: Reliability and feasibility study. *Clinical Cancer Research, 10*, 1881–1886.

Cavalcanti, P. G., Scharcanski, J., & Lopes, C. B. O. (2010). Shading attenuation in human skin color images. In *Proc. sixth international symposium on visual computing (ISVC 2010)*.

Cavalcanti, P. G., & Scharcanski, J. (2011). Automated prescreening of pigmented skin lesions using standard cameras. *Computerized Medical Imaging and Graphics, 35*, 481–491.

Celebi, M., Iyatomi, H., Schaefer, G., & Stoecker, W. V. (2009). Lesion border detection in dermoscopy images. *Computerized Medical Imaging and Graphics, 33*, 148–153.

Celebi, M. E., Kingravi, H. A., Iyatomi, H., Aslandogan, Y. A., Stoecker, W. V., Moss, R. H., et al. (2008). Border detection in dermoscopy images using statistical region merging. *Skin Research and Technology, 14*, 347–353.

Celebi, M. E., Kingravi, H. A., Uddin, B., Iyatomi, H., Aslandogan, Y. A., Stoecker, W. V., et al. (2007). A methodological approach to the classification of dermoscopy images. *Computerized Medical Imaging and Graphics, 31*, 362–373.

Commission Internationale de l'Eclairage. (2011). Selected colorimetric tables. <http://www.cie.co.at>.

Dermnet skin disease image atlas. (2009). <http://www.dermnet.com>.

Dimitrow, E., Riemann, I., Ehlers, A., Koehler, M., Norgauer, J., Elsner, P., et al. (2009). Spectral fluorescence lifetime detection and selective melanin imaging by multiphoton laser tomography for melanoma diagnosis. *Experimental Dermatology, 18*, 509–515.

Dreiseitl, S., Ohno-Machado, L., Kittler, H., Vinterbo, S., Billhardt, H., & Binder, M. (2001). A comparison of machine learning methods for the diagnosis of pigmented skin lesions. *Journal of Biomedical Informatics, 34*, 28–36.

Figueiredo, M. A. F., & Jain, A. K. (2002). Unsupervised learning of finite mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 24*, 381–396.

Fikrle, T., & Pizinger, K. (2007). Digital computer analysis of dermatoscopical images of 260 melanocytic skin lesions; perimeter/area ratio for the differentiation between malignant melanomas and melanocytic nevi. *Journal of the European Academy of Dermatology and Venereology, 21*, 48–55.

Gomez, D. D., Butakoff, C., Ersboll, B. K., & Stoecker, W. (2008). Independent histogram pursuit for segmentation of skin lesions. *IEEE Transactions on Biomedical Engineering, 55*, 157–161.

Howlader, N., Noone, A., Krapcho, M., Neyman, N., Aminou, R., Altekruse, S., Kosary, C., Ruhl, J., Tatalovich, Z., Cho, H., Mariotto, A., Eisner, M., Lewis, D., Chen, H., Feuer, E., & Cronin, K. (2012). SEER cancer statistics review, 1975–2009 (Vintage 2009 populations), National cancer institute. <http://seer.cancer.gov/csr/1975_2009_pops09/>.

Iyatomi, H., Oka, H., Celebi, M., Hashimoto, M., Hagiwara, M., Tanaka, M., et al. (2008). An improved internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm. *Computerized Medical Imaging and Graphics, 32*, 566–579.

Jimbow, K., Reszka, K., Schmitz, S., Salopek, T., & Thomas, P. (1995). Distribution of eu- and pheomelanins in human skin and melanocytic tumors, and their photoprotective vs phototoxic properties. In M. C. L. Zeise & T. Fitzpatrick (Eds.), *Melanin: Its role in human photoprotection* (pp. 165–175). Overland Park, KS, USA: Valdenmar Publishing Company.

Johr, R. H. (2002). Dermoscopy: Alternative melanocytic algorithms-the ABCD rule of dermatoscopy, menzies scoring method, and 7-point checklist. *Clinics in Dermatology, 20*, 240–247.

Krishnaswamy, A., & Baranoski, G. (2004). A biophysically-based spectral model of light interaction with human skin. *Computer Graphics Forum, 23*, 331–340.

Lazova, R., & Pawelek, J. (2009). Why do melanomas get so dark? *Experimental Dermatology, 18*, 934–938.

Lee, H.-C. (2005). *Introduction to color imaging science*. New York, NY, USA: Cambridge University Press.

Manousaki, A. G., Manios, A. G., Tsompanaki, E. I., Panayiotides, J. G., Tsiftsis, D. D., Kostaki, A. K., et al. (2006). A simple digital image processing system to aid in melanoma diagnosis in an everyday melanocytic skin lesion unit: A preliminary report. *International Journal of Dermatology, 45*, 402–410.

Marchesini, R., Bono, A., & Carrara, M. (2009). In vivo characterization of melanin in melanocytic lesions: Spectroscopic study on 1671 pigmented skin lesions. *Journal of Biomedical Optics, 14*, 014027.

Massone, C., Wurm, E. M. T., Hofmann-Wellenhof, R., & Soyer, H. P. (2008). Teledermatology: An update. *Seminars in Cutaneous Medicine and Surgery, 27*, 101–105.

Matthews, T. E., Piletic, I. R., Selim, M. A., Simpson, M. J., & Warren, W. S. (2011). Pump-probe imaging differentiates melanoma from melanocytic nevi. *Science Translational Medicine, 3*, 71ra15.

Mayer, J. (1997). Systematic review of the diagnostic accuracy of dermatoscopy in detecting malignant melanoma. *The Medical Journal of Australia, 167*, 206–210.

McCartney, E. (1976). *Optics of the atmosphere: Scattering by molecules and particles*. John Wiley & Sons.

Nachbar, F., Stolz, W., Merkle, T., Cognetta, A. B., Vogt, T., Landthaler, M., et al. (1994). The abcd rule of dermatoscopy: High prospective value in the diagnosis of doubtful melanocytic skin lesions. *Journal of the American Academy of Dermatology, 30*, 551–559.

Natural phenomena simulation group. (2011). Run BioSpec Online. <http://www.npsg.uwaterloo.ca/models/biospec.php>.

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics, 9*, 62–66.

Rao, B. K., Marghoob, A. A., Stolz, W., Kopf, A. W., Slade, J., Wasti, Q., et al. (1997). Can early malignant melanoma be differentiated from atypical melanocytic nevi by in vivo techniques? *Skin Research and Technology, 3*, 8–14.

Ruiz, D., Berenguer, V., Soriano, A., & Snchez, B. (2011). A decision support system for the diagnosis of melanoma: A comparative approach. *Expert Systems with Applications, 38*, 15217–15223.

Skvara, H., Teban, L., Fiebiger, M., Binder, M., & Kittler, H. (2005). Limitations of dermoscopy in the recognition of melanoma. *Archives of Dermatology, 141*, 155–160.

Tabatabaie, K., Esteki, A., & Toossi, P. (2009). Extraction of skin lesion texture features based on independent component analysis. *Skin Research and Technology, 15*, 433–439.

Tang, J. (2009). A multi-direction GVF snake for the segmentation of skin cancer images. *Pattern Recognition, 42*, 1172–1179.

Trowbridge, T., & Reitz, K. (1975). Average irregularity representation of a rough surface for ray reflection. *Journal of the Optical Society of America, 65*, 531–536.

Tuchin, V. (2000). *Tissue optics light scattering methods and instruments for medical diagnosis*. Bellingham, WA, USA: The International Society for Optical Engineering.

Whited, J. D. (2006). Teledermatology research review. *International Journal of Dermatology, 45*, 220–229.

Wighton, P., Lee, T. K., Lui, H., Mclean, D. I., & Atkins, M. S. (2011). Generalizing common tasks in automated skin lesion diagnosis. *IEEE Transactions on Information Technology in BioMedicine*, 622–629.

Wighton, P., Sadeghi, M., Lee, T. K., & Atkins, M. S. (2009). A fully automatic random walker segmentation for skin lesions in a supervised setting. In *Proceedings of the 12th international conference on medical image computing and computer-assisted intervention: Part II MICCAI '09* (pp. 1108–1115). Berlin, Heidelberg: Springer-Verlag.

Wong, A., Scharcanski, J., & Fieguth, P. (2011). Automatic skin lesion segmentation via iterative stochastic region merging. *IEEE Transactions on Information Technology in Biomedicine, 15*, 929–936.

World Health Organization. (2011). How commom is skin cancer? <http://www.who.int/uv/faq/skincancer/en/index1.html>.

Young, G. A. (1990). Alternative smoothed bootstraps. *Journal of the Royal Statistical Society. Series B (Methodological), 52*, 477–484.

Zonios, G., Dimou, A., Bassukas, I., Galaris, D., Tsolakidis, A., & Kaxiras, E. (2008). Melanin absorption spectroscopy: New method for noninvasive skin investigation and melanoma detection. *Journal of Biomedical Optics, 13*, 014017-1–014017-8.

Zonios, G., Dimou, A., Carrara, M., & Marchesini, R. (2010). In vivo optical properties of melanocytic skin lesions: Common nevi, dysplastic nevi and malignant melanoma. *Photochemistry and Photobiology, 86*, 236–240.