Physiological Characterization of Skin Lesion using Non-linear Random Forest Regression Model

Daniel S. Cho[1] and Shahid Haider[1] and Robert Amelard[1] and Alexander Wong[1] and David Clausi[1]

Abstract— **The current diagnostic technique for melanoma solely relies on the surface level of skin and under-skin information is neglected. Since physiological features of skin such as melanin are closely related to development of melanoma, the non-linear physiological feature extraction model based on random forest regression is proposed. The proposed model characterizes the concentration of eumelanin and pheomelanin from standard camera images or dermoscopic images, which are conventionally used for diagnosis of melanoma. For the validation, the phantom study and the separability test using clinical images were conducted and compared against the state-of-the art non-linear and linear feature extraction models. The results showed that the proposed model outperformed other comparing models in phantom and clinical experiments. Promising results show that the quantitative characterization of skin features, which is provided by the proposed method, can allow dermatologists and clinicians to make a more accurate and improved diagnosis of melanoma.**

I. INTRODUCTION

With the advancement of imaging acquiring devices, the use of standard camera images is not limited at the surface level of objects, but extended to extract more information beyond what people can see. Skin imaging is one of the active research areas, where takes an advantage from the current imaging technology by finding physiological features underneath skin [1], [2]. Skin is composed of multiple layers and each layer has a different combination of pigments, such as melanin and hemoglobin [3], [4]. These molecules or physiological features play a key role to determine the colour of skin [5], and thus, identification and quantitative measurement of physiological features persistently attracts interests of researchers from image processing [6], cosmetics [7], and medical diagnosis [8], [9].

Medical diagnosis, particularly for melanoma, can be benefited from physiological feature extraction. Melanoma is the deadliest form of skin cancer [10]. It typically originates from melanocytes, which is responsible for producing melanin. What makes melanoma dangerous is the ability to metastasis. Therefore, it is crucial to diagnose this disease as early as possible before it spreads to other parts of body [11]. Initial diagnosis of melanoma is typically conducted with the naked eyes of clinicians or dermatologists using diagnostic criteria. While this has been the standard procedure for diagnosing melanoma, the limitation of this method lies in individual subjectivity of diagnosis of clinicians, which is solely based on the visual inspection of the lesion. Some studies showed that by simply quantifying the diagnostic criteria, the accuracy of the diagnosis of melanoma has been improved [12], [13]. To move a step further, incorporating physiological features, for example melanin and hemoglobin, which are known to be related to melanoma, would improve results even more.

Melanin, which is produced from melanocytes, is one of the most prevalent pigments in skin, and has two subtypes: eumelanin and pheomelanin. Eumelanin gives brown-blackish colour while pheomelanin colors red-yellowish. Skin color is dominated by eumelanin [5], and the ratio between the concentration of pheomelanin and eumelanin present in human skin varies greatly from individual to individual. While a major function of melanin is providing a protection from ultraviolet (UV) radiation, pheomelanin is known to be more vulnerable than eumelanin to DNA damages or mutations, caused by UV radiation [14]. This vulnerability of pheomelanin suggest that pheomelanin plays an important role to develop a cancer. Salopek *et al.* [15] and Hu *et al.* [16] found that the concentrations of pheomelanin increased in melanomic cells, when compared with normal cells.

Since typical imaging modalities used by clinicians and dermatologists are standard camera images or dermoscopic images, the same modalities were adopted into physiological feature extraction. Colour determination of skin from various pigments is very complicated process involving reflection, scattering and absorption of light. Moreover, from the best of our knowledge, the most skin feature extraction models use a linear model based on Lambert-Beer law, which accounts only for the absorption of skin, leaving the scattering and reflection ignored [7], [17]. Cavalcanti *et al.* [18] proposed a non-linear nearest-neighbour model, that extracts the concentrations of eumelanin and pheomelanin, and the model includes the scattering and reflection as well as absorption. However, it is a computationally expensive model, and thus if faces a challenge when more features added to the model or dealing with a large set of data.

The contribution of the paper is to propose a non-linear random forest regression model to extract important physiological features from standard camera or dermoscopic images. For the scope of this paper, eumelanin and pheomelanin are employed. For validation, phantom study and separability test with clinical skin lesion images are conducted and compared with the state-of-the-art non-linear and linear feature extraction model.

*This research was undertaken, in part, thanks to funding from Agfa Healthcare Inc., the Natural Sciences and Engineering Research Council (NSERC) of Canada, and Canada Research Chairs program.

[1]D. Cho, R. Amelard, A. Wong, and D. Clausi are with Department of Systems Design Engineering, University of Waterloo, Ontario, Canada, N2L 3G1 $\{s34cho, ramelard, a28wong, dclausi\}$@uwaterloo.ca

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

II. METHOD

A. Problem Formulation

When light hits the surface of skin, the light is remitted based on the composition of the physiological features under the skin. The remitted light can be decomposed into three major spectral bands, known as red, green, and blue band, and the intensities of each band are captured by the image acquiring device (*i.e.*, camera). For explaining the above phenomenon mathematically, the concentrations of the physiological features of skin is expressed,

$$p = \{p_k\}_{k=eumelanin,pheomelanin,\dots}, p \in P \tag{1}$$

where subscript k represents different physiological skin features such as eumelanin, pheomelanin, oxygenated hemoglobin, or deoxygenated hemoglobin. P is the space of all possible physiological features inside skin tissue. Given p, obtaining skin lesion image can be formulated as follows:

$$i = f(p) \tag{2}$$

where f computes the intensities in the three spectral bands, i. Moreover, i can be written as follows:

$$i = \{i_n\}_{n=r,g,b}, i \in I \tag{3}$$

where I describes the space of the acquired image.

Based on the above formulation, the inverse model function, which estimates the concentrations of physiological features from skin images can be stated as

$$p = f^{-1}(i) \tag{4}$$

The forward model function, f, cannot be described as a simple linear model, because interaction between light and the different layers of skin is complex process, which involves different degree of reflection, scattering, and absorption at each skin layer. Therefore, in this paper, a non-linear regression model was proposed for the inverse model with the assumption that the interaction between light and skin is non-linear.

B. Non-Linear Random Forest Regression Modeling

As aforementioned, the proposed model estimates the concentration of physiological features from skin lesion images via random forest regression. The basic principle of random forest is to generate many decision trees, and aggregate their results [19]. One advantage of this model is that results are more robust and less biased than other regression models that depends on single or few classifiers [20]. For each tree, every node is split using the best among a subset of predictors, which are randomly chosen at that node. Let X be the p-dimensional vector of variables or predictors that $X = x_1, x_2, \cdots, x_p$, then an ensemble of N trees be $h_1(X, \theta_1), h_2(X, \theta_2), \cdots, h_n(X, \theta_n)$ where θ_i is an independently distributed random vector. Each classifier computes the output, $\hat{Y}_i$, as a form of numerical values,

$$\hat{Y}_i(X) = h_i(X, \theta_i) \tag{5}$$

and the final output for random forest regression is the average of the outputs from trees as following:

$$\hat{Y} = \frac{1}{N} \sum_{i=1}^{N} \hat{Y}_i(X) \tag{6}$$

In this study, the intensities from the three spectral bands (red, green, and blue band for standard camera and dermoscopic images) are used as predictors, and the concentrations of eumelanin and pheomelanin are computed as outputs of the algorithm.

C. Experimental Design

To train the proposed model, the forward model, proposed by Baranoski *et al.* [21] (*i.e.*, biophysically-based spectral model of light interaction with human skin) was employed to generate training data, because the concentrations of physiological skin features is not easily measured or extracted from clinical images. Given the concentrations of physiological skin features, Baranoski's forward model calculated the corresponding reflectance values by simulating the light propagation via Monte Carlo simulation, and the reflectance values were converted to RGB values using simple linear function [22]. For this study, the concentrations of eumelanin and pheomelanin were changed from 20 to 300 g/L, and 4 to 60 g/L with step of 4 g/L, respectively, while other model parameters were kept at default values.

The validation of the proposed method was conducted using cross-validation, phantom study and separability test. Non-linear nearest neighbor model (CNN) [18] was used for comparison as well as linear regression model (LRM) [23] to be used as a baseline model. Both models were trained using the same dataset, which was generated by Baranoski's forward model.

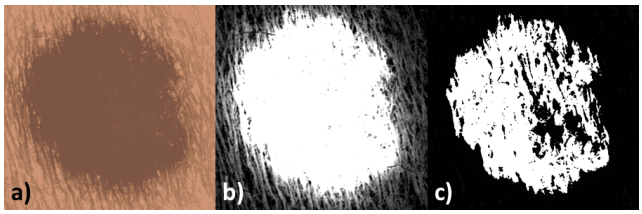

Fig. 1. a) The phantom image created based on melanomic image, and its concentration map of b) eumelanin and c) pheomelanin

1) Phantom Study: In order to investigate the proposed algorithm in more clinical setting, a phantom image was created. First, a malignant lesion was delineated from melanomic image. The concentrations of eumelanin and pheomelanin were then extracted using the random forest regression model. Baranoski's forward model was followed to simulate the RGB values for the corresponding physiological feature concentrations, and the calculated colours were applied back to delineated lesion as shown in Fig. 1. The proposed RF method, CNN, and LRM extracted the concentrations of eumelanin and pheomelanin from the phantom image.

2) Separability Test: Since the actual concentrations of eumelanin and pheomelanin cannot be acquired from clinical images, the direct validation of the proposed method is not feasible. To bypass this issue, the separability test was performed. The purpose of the separability test is to examine how each feature can uniquely separate a malignant lesion from a benign lesion, and thus the strength of features from different extraction techniques can be quantitatively measured. The Fisher criterion scores were computed based on the performance of the proposed model and CNN as a measure of separability [24]. A dataset of 112 clinical images (39 melanoma, 63 non-melanoma) from DermIS [25] and DermQuest [26] was used for the experiment.

III. RESULTS

To compare the performance of the proposed method with two other state-of-the-art techniques, three different validations were conducted. First of all, cross-validation was performed. Samples that contains various combination of eumelanin and pheomelanin concentrations and their corresponding RGB values, was generated according to II-C. With a total of 1065 samples, 90% of them were randomly chosen as training set, and the rest 10% as testing set. Each algorithm (i.e. RF, CNN and LRM) were trained and tested with a total of 50 iterations, and the root-mean-square-error (RMSE) of predicted physiological concentrations from models are shown in Table I.

TABLE I
CROSS-VALIDATION: COMPARING RMSE OF EUMELANIN AND PHEOMELANIN THAT WAS PREDICTED FROM RF, CNN, AND LRM (G/L)

	RF (g/L)	**CNN (g/L)**	**LRM (g/L)**
Eumelanin	1.2	4.0	5.0
Pheomelanin	2.3	8.0	8.8

Secondly, the models were validated via phantom, which was created based on a clinical image. Unlike cross-validation test, feature extraction model was built based on the entire colour map. RMSE of eumelanin and pheomelanin was used as a measure of comparison (Table II).

TABLE II
PHANTOM STUDY: COMPARING RMSE OF EUMELANIN AND PHEOMELANIN THAT WAS PREDICTED FROM RF, CNN, AND LRM (G/L)

	RF (g/L)	**CNN (g/L)**	**LRM (g/L)**
Eumelanin	1.9	2.0	8.0
Pheomelanin	3.2	5.2	15.8

Lastly, Fisher criterion scores were computed for RF and CNN from a total of 112 clinical images in Table III.

TABLE III
COMPARING FISHER CRITERION SCORES FROM RF AND CNN USING CLINICAL MELANOMIC AND NON-MELANOMIC IMAGES

	RF	**CNN**
Fisher criterion score	0.0439	0.0046

IV. DISCUSSION

In this paper, we proposed a novel technique to extract physiological features (*i.e.*, eumelanin and pheomelanin) from standard camera images. The importance of physiological features is that the concentration map of those features can provide additional information to clinicians and dermatologists, which is typically neglected in current diagnostic system. For example, some studies showed that the increase in eumelanin and pheomelanin concentrations could be used as bio-markers for melanoma [15], [16]. The concentration maps, which were generated by the proposed model, are consistent with their findings as the elevation of eumelanin and pheomelanin concentrations was observed in melanomic image, compared to non-melanomic image (Fig 2).

For the validation, the proposed method via random forest regression model was compared and validated against CNN, since Cavalcanti's nearest neighbor model is the only non-linear model found in literature. Moreover, the linear regression model was also compared to serve as a baseline model. From both of cross-validation and phantom study, RF model outperformed over CNN and LRM when root-mean-square-error was compared with predicted concentrations of eumelanin and pheomelanin with ground truth. The performance of LRM was inferior to other methods particularly in the phantom study. This implies that the optical properties of physiological features under skin are far more complex, and thus, it is difficult to describe the interaction between light and the physiological features in linear fashion. When two non-linear models were compared, the proposed method yielded the superior results not only in RMSE comparison but also in computation time. For phantom study, RF predicted about 15% faster than CNN as the computation time for RF and CNN was measured 7.9s and 9.1s, respectively, using the machine equipped with Intel Core i7-4770 @ 3.40GHz, 16.0GB ram, and Intel HD Graphics 4600. Given that the dimension of the phantom used is 520 by 460, and it is a single image, the computation time of CNN will likely increase exponentially when larger dataset is used and more features are taking into account for the system.

To validate the quality of the extracted features from CNN and RF, the Fisher criterion score was employed. The Fisher criterion score is calculated by the distance between two classes over the sum of sparsity of each class. Therefore, it indicates the ability to separate benign and malignant cases for the given features, and the higher scores represent stronger ability for classification. From two non-linear models, RF scored 0.04 and 0.0044 for CNN. This results show that the features extracted using RF more effectively discriminate benign and malignant melanoma than the ones using CNN.

While the proposed physiological extraction model showed the promising results compared to the current state-of-the-art, it requires further improvements in order to be adopted in the clinical setting. First of all, more skin features

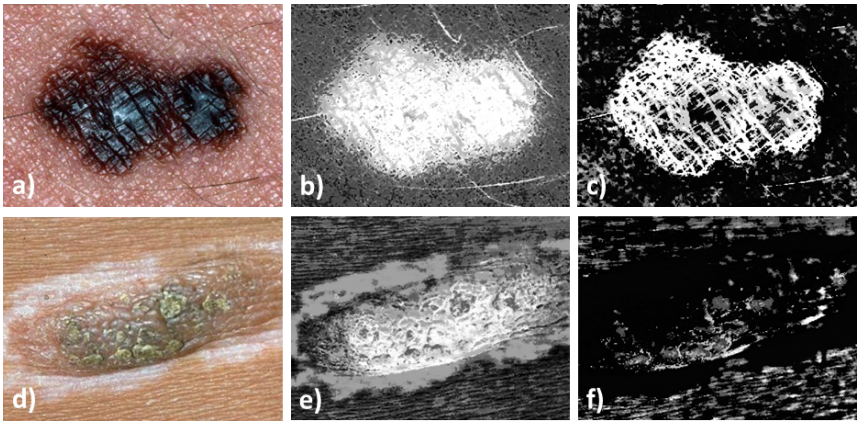

Fig. 2. Top: a) melanomic skin lesion image, and its concentration map of b) eumelanin and c) pheomelanin. Bottom: d) non-melanomic skin lesion image, and its concentration map of e) eumelanin and f) pheomelanin

need to be extracted to examine melanoma. For example, hemoglobin also plays an important role for the interaction between light and human skin as it is prevalent in skin and a major absorber and scatterer. Moreover, given that the ground truth used for this study was synthetically generated based on computer model, Baranoski's forward model, the clinical validation of the proposed method is required.

V. CONCLUSIONS

We have proposed a novel physiological skin feature extraction model. A non-linear model was created based on random forest regression, and it extracts the concentration of two skin features, eumelanin and pheomelanin, on skin lesion. The proposed method was validated and compared with the non-linear nearest neighbor model and the linear regression model, and the proposed technique outperformed over other models in both cross-validation and phantom study. The proposed method can provide an additional information of the important physiological features, that closely linked to melanoma, on top of the conventional diagnosis, and we believe that it would ultimately help clinicians and dermatologists for better diagnosis of melanoma.

REFERENCES

[1] Stamatas, Georgios N., et al. "NonInvasive Measurements of Skin Pigmentation In Situ." Pigment Cell Research 17.6 (2004): 618-626.
[2] Wallace, V. P., et al. "Classification of reflectance spectra from pigmented skin lesions, a comparison of multivariate discriminant analysis and artificial neural networks." Physics in medicine and biology 45.10 (2000): 2859.
[3] Doi, Motonori, and Shoji Tominaga. "Spectral estimation of human skin color using the Kubelka-Munk theory." Electronic Imaging 2003. International Society for Optics and Photonics, 2003.
[4] A. Krishnaswamy and G. V. Baranoski, "A study on skin optics," Natural Phenomena Simulation Group, School of Computer Science, University of Waterloo, Canada, Technical Report 1, 1-17 (2004).
[5] A. Thody, E. Higgins, K. Wakamatsu, S. Ito, S. Burchill, and J. Marks, "Pheomelanin as well as eumelanin is present in human epidermis," Journal of Investigative Dermatology 97, 344-344 (1991).
[6] Ramamoorthi, Ravi, and Pat Hanrahan. "A signal-processing framework for inverse rendering." Proceedings of the 28th annual conference on Computer graphics and interactive techniques. ACM, 2001.
[7] Tsumura, Norimichi, et al. "Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin." ACM Transactions on Graphics. Vol. 22. No. 3, 2003.
[8] Madooei, Ali, et al. "Intrinsic melanin and hemoglobin colour components for skin lesion malignancy detection." MICCAI 2012. Springer Berlin Heidelberg, 2012.
[9] Claridge, Ela, et al. "From colour to tissue histology: physics-based interpretation of images of pigmented skin lesions." Medical Image Analysis 7.4 (2003): 489-502.
[10] Melanoma PubMed Health, U.S. National Library of Medicine [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0001853
[11] Ascierto, Paolo A., et al. "The role of spectrophotometry in the diagnosis of melanoma." BMC dermatology 10.1 (2010): 5.
[12] Amelard, Robert, et al. "Melanoma decision support using lighting-corrected intuitive feature models." Computer Vision Techniques for the Diagnosis of Skin Cancer. 2014. 193-219.
[13] Cavalcanti, Pablo G., and Jacob Scharcanski. "Automated prescreening of pigmented skin lesions using standard cameras." Computerized Medical Imaging and Graphics 35.6 (2011): 481-491.
[14] Ye, Tong, et al. "Photoionization Thresholds of Melanins Obtained from Free Electron LaserPhotoelectron Emission Microscopy, Femtosecond Transient Absorption Spectroscopy and Electron Paramagnetic Resonance Measurements of Oxygen Photoconsumption." Photochemistry and photobiology 82.3 (2006): 733-737.
[15] SALOPEK, THOMAS G., et al. "Dysplastic melanocytic nevi contain high levels of pheomelanin: quantitative comparison of pheomelanin/eumelanin levels between normal skin, common nevi, and dysplastic nevi." Pigment cell research 4.4 (1991): 172-179.
[16] Hu, Dan-Ning, et al. "Comparison of eumelanin and pheomelanin content between cultured uveal melanoma cells and normal uveal melanocytes." Melanoma research 19.2 (2009): 75-79.
[17] Gong, Hao, and Michel Desvignes. "Hemoglobin and melanin quantification on skin images." Image Analysis and Recognition. Springer Berlin Heidelberg, 2012. 198-205.
[18] Cavalcanti, Pablo G., Jacob Scharcanski, and Gladimir VG Baranoski. "A two-stage approach for discriminating melanocytic skin lesions using standard cameras." Expert Systems with Applications 40.10 (2013): 4054-4064.
[19] Breiman, Leo. "Random forests." Machine learning 45.1 (2001): 5-32.
[20] Pardo, Carlos, et al. "Rotation Forests for regression." Applied Mathematics and Computation 219.19 (2013): 9914-9924.
[21] Krishnaswamy, Aravind, and Gladimir VG Baranoski. "A biophysicallybased spectral model of light interaction with human skin." Computer Graphics Forum. Vol. 23. No. 3, 2004.
[22] Lee, Hsien-Che. Color Imaging Science. Cambridge: Cambridge University Press, 2005.
[23] Lai, T. L., and Ching, Z. W. "Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems." The Annals of Statistics (1982): 154-166.
[24] Yang, Jian, and Jing-yu Yang. "Why can LDA be performed in PCA transformed space?." Pattern recognition 36.2 (2003): 563-566.
[25] Dermatology Information System [Online], Available: http://www.dermis.net
[26] DermQuest [Online], Available: http://www.dermquest.com