Prediction of skin ages by means of multi-spectral light sources

H. Seker*, V. Uslan, A.B. Orun, E. Goodyer and G. Smith

Abstract— **Assessment of skin aging is a complex biological process that depends on various internal and external factors but has become important due to personalized skin health and cosmetic treatments. Although there are a small number of attempts to assess the skin aging, they identify only one of the previously classified skin aging groups. The methods used to achieve it are also based generally on highly expensive measurement devices. This work therefore proposes novel low-cost skin aging assessment apparatus by using light back-scatter intensity level of Red, Blue, Green and Infrared bands. This is further enhanced by using a machine learning method to accurately predict actual skin age. The proposed method appears to be highly capable of capturing multi-layer cellular changes exhibited by the biological skin aging process and predicting skin ages with a root-mean-square error of as low as 0.160 by using only four features based on the four multi-spectral light sources. This assessment kit seems to be the first of its kind, which is expected to bring great benefit to both personalized skin healthcare and cosmetic sectors.**

I. INTRODUCTION

Skin aging is a complex biological process which affects different skin layers, structures and components. Particularly, in the basal layer (*stratum basale*), many remarkable age-related changes may occur such as the size of cells which are called *"basal keratinocytes*" [1]. The other age related changes are also observed in skin blood flow and skin blood content (blood capillary loops) [2].

Modelling of skin aging process is important in many different aspects including skin health and cosmetics [1, 3, 4]. This can help better understand skin growth, damage and diseases as well as manage personalized care of skin.

One of the major structural changes through the skin aging process occurs in the cell sizes at the basal layer [1] which is located at approximately 0.1*mm* depth of the skin. It was reported that Red light (λ=620-740 nm) can penetrate up to 2*mm* downward inside the skin [5] and may interact with those cells (*basal keratinocytes)* to generate light back-scatter effects due to changes in the cell network. This effect of light normally conveys indirect information about the changes in cell size and their structural characteristics by the series of comparative observations. More complex textural changes in blood flow (affecting the textural intensity of blood cells) were studied by Ryan [2]. It was concluded that over an age group from 20 to 72 skin blood flow falls by 40% by aging. In addition, the alternative Red laser light used in our earlier tests is found to have optimal wavelength of 650 nm for red blood cells reflections to detect them [3, 6]. These effects that appear on skin images were built due to laser speckle phenomenon [7, 8]. However, in our present work only light back-scatters are used as it was also studied by previous researchers [3, 13, 14].

In the previous studies carried out within our group, the laser-optical and textural analysis of sub-skin layers exploiting the textural characteristics of the skin were studied and found to be highly capable of analyzing the skin structure [3] and accurate classification of two skin age groups, namely, below/above 30 years [9]. These proposed methods and equipment were found to be quite cost-effective alternative to some of the functions performed by high-cost confocal microscopy (being around $100K) or similar expensive instrumentations [3, 9]. However, the system developed requires full laser equipments and not used with a machine learning method for prediction. This work therefore aims to investigate the feasibility of more basic low-cost and non-invasive solutions to be used for the analysis of skin-aging process in corporation with support vector regression as a statistical machine learning method, which has been used to be more capable of modelling complex data sets [11, 12].

The proposed system utilizes multi-spectral light sources based on the light back-scatter intensity levels of Red, Blue, Green and Infrared lights. Using the four-feature set based on the four light back-scatter intensity levels, this current study is particularly concerned with the prediction of actual skin age, which would be achieved by utilizing a machine learning method.

It should be noted that the methods are not aimed at substituting all aspects of such high-cost instruments and laser based systems. However, it is anticipated to be effectively used to undertake some of their important functions to assess skin aging progress at earlier stages by means of such a low-cost reliable and easy-to-use instrumentation and method, which is one of the most

The authors are with the Bio-Health Informatics Research Group within the Centre for Computational Intelligence, De Montfort University, Leicester, UK (e-mails: aorun@dmu.ac.uk, hseker@dmu.ac.uk, eg@dmu.ac.uk and gsmith02@dmu.ac.uk). G. Smith is also with the Pharmaceutical Technologies Research Group, Leicester School of Pharmacy, De Montfort University, Leicester, UK.

*Corresponding author: hseker@dmu.ac.uk

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

desirable components of current healthcare systems around the world.

To the best of our knowledge, this study and proposed kit are the first of its kind, where the prediction of actual skin ages is presented. The method and low-cost equipment developed are therefore expected to have a great potential in both healthcare and cosmetic sectors, particularly for the personalized care aspect and objective assessment of anti-aging cosmetic products.

II. METHODS AND MATERIALS

To prove the above hypothesis, 136 participants of broad range of age groups (between the ages of 19 and 60) have been selected to collect skin data set by utilizing back scatter readings of Red, Green, Blue and Infrared bands (RGB & IR). The samples also include participants with five different types of skin colors graded as from light to dark. All the back scatter samplings were taken from the same inner-arm location. The measurements were then analyzed and modelled by using the support vector regression method in order to perform an automated prediction of skin age. These two stages will be discussed further in the following sections.

A. Multi-spectral light sources (Red, Green, Blue and Infrared bands (RGB & IR)) for the skin data collection

The multi-spectral light sources, namely Red, Green, Blue and Infrared bands/lights (RGB & IR), are used to investigate the role of their back-scatter effect to detect any skin change as their penetration distance into the skin is up to 2.5 *mm* [5] beyond the basal layer.

The multi-spectral light sources (RGB & IR) are used for pinpoint skin analysis (Fig.1), whereby the back-scatters are recorded and analyzed for skin aging modelling, which also includes the prediction of actual skin age being the main concern of this paper.

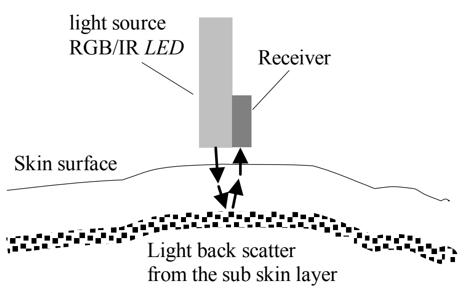

Fig.1. Experimental setup to measure light back scatters where Red (λ=640 nm), Green (λ=568 nm), Blue (λ=480 nm) and IR (λ=880 nm) *LED* types were used.

The instrumentation set consists of four types (R, G, B & IR bands) of transmitter diodes (LED) and receiver sensors to measure the light back scatter intensity levels from skin sub-surface. The wavelengths of the emitters used are Red (λ=640 nm), Green (λ=568 nm), Blue (λ=480 nm) and IR (λ=880 nm). Each instrument is calibrated before each measurement in order to avoid the effects of changes in environmental light intensity level. For the back scatter analysis, simple hand held devices were designed that enabled the team to place an emitter adjacent to a detector on the skin surface. The received signal was calibrated using a mirror to represent 100%, and a highly absorbent black material to represent 0%. All the four readings were taken from the same position in order to maintain consistency.

It should be noted that there were five different measurements taken for each subject for each light. Subsequently, the average of the five readings was used to represent each subject for each light.

Fig.2 shows how the values of four readings (R, G, B & IR bands) are distributed over the skin ages. The figure suggests that there is no single spectral light source that seems solely capable of determining the skin age. The values obtained from B and G seem to be more compact whereas those of IR and R are scattered. Therefore, the results suggest that it is necessary to explore all or combinations of these four light sources through a predictive model, for which the support vector regression-based predictive model will be utilized.

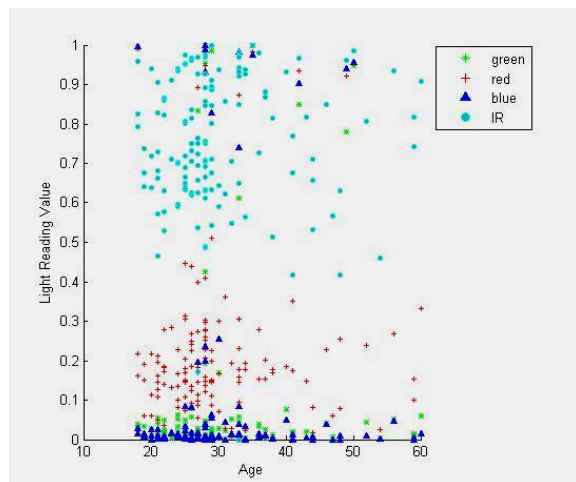

Fig.2. The normalized distribution of the values of four back-scatter readings (R, G, B & IR) with respect to skin ages (small number of 100% reflection values for IR region is likely to be due to instrumental calibration error that was also taken into account in the study).

B. Support Vector Regression (SVR)

Support Vector Machine (SVM) is a supervised statistical learning approach that learns structure from data using structural risk minimization [10]. Growing amount of classification and real-value estimation tasks are conducted by utilizing this recent method due to its generalization capabilities. The fundamental difference it suggests from other regression-based approaches is that it attempts for the minimization of an upper bound on the empirical regression error rather than minimization of an objective function.

SVR has been shown to be efficient particularly on modelling high-dimensional data sets and uses support vectors which are the chosen training points during the learning process [10]. It should be noted that, when the real-value estimation is conducted, SVM can be referred to as Support Vector Regression (SVR).

SVR approximates a linear function *h(x)* in the following form:

$$h(x) = w^T x + b \tag{1}$$

where w and b are the coefficients that denote weight vector and bias term, respectively.

The optimization problem is to be able to find optimum values of the parameters and can be defined for SVR as follows:

$$\min \frac{1}{2}\|w\|^2 + C\sum \xi_+ + \xi_- \tag{2}$$

where ξ_+, ξ_- denote the slack variables and the constant C > 0 is a parameter that helps to control the tradeoff between the function complexity and margin which deviates greater than ϵ. Fig.3 depicts ϵ-insensitive loss function graphically [11].

The function used for minimization of the optimization problem defined in Eq.(2) is subject to:

$$\begin{aligned} &y' - (w^T x + b) \leq \epsilon + \xi_+ \\ &(w^T x + b) - y' \leq \epsilon + \xi_- \\ &(\xi_+, \xi_-) \geq 0 \end{aligned} \tag{3}$$

One of the important parameters used to optimize SVR is type of kernel functions. The common kernel functions used for SVR are linear, polynomial, radial and sigmoid functions [10, 11]. The formulations given in the mathematical expressions (1-3) are only valid for the linear kernel function where two optimization parameters (C and ϵ) are used. For other three kernel functions (polynomial, radial and sigmoid), there is one more optimization parameter (γ) that has been shown to be an important parameter for the optimization of SVR [12].

In order to implement SVR, LIBSVM tool was utilized in this study [12].

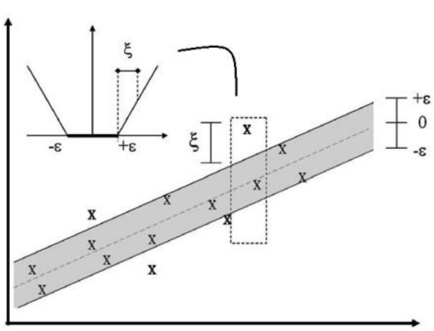

Fig.3. Illustration of ϵ-insensitive loss function for SVM with a linear kernel function [11]

III. RESULTS AND DISCUSSION

As depicted in Fig.2., it is observed that there is no single particular feature (light source) that could be sufficiently utilized to help determine a skin age. The analysis was therefore carried out by using the four features, namely the readings of four multi-spectral light sources (R, G, B & IR).

In order to develop a predictive model for skin age, SVR was utilized. As SVR has a number of parameters that help optimize the predictive model, various combinations of these parameters have been studied in order to obtain the best possible predictive model. There are four different kernel types (linear, polynomial, radial basis function and sigmoidal functions). In addition, there are two optimization parameters (C and ϵ) for the linear kernel and three parameters (C, ϵ and γ) for other three kernel functions. Each kernel function was studied using various numbers of the values of these parameters in a grid search.

In order to assess generalization ability of the proposed approach, five-fold cross validation was implemented. In this case, the data set was divided into five subsets, each one of which consists of more or less equal number of samples. Four of these folds (sub-sets) were used to build up a SVR-based predictive model, and the remaining fold was then used to test the model. This procedure was repeated until all the five folds were used for testing. As the sub-sets are randomly generated, this process was repeated for a few times (e.g., 10 times) in order to show if or how the predictive models would be sensitive to different samples.

The results presented in Table 1 include the average of root-mean square error (RMSE) that can be calculated as

$$\text{RMSE} \;) = \sqrt{\frac{\sum_{t=1}^{n}(\hat{y}_t - y_t)^2}{n}}. \tag{4}$$

where $\hat{y}_t$ and y are the predicted and actual skin ages, respectively, and n is the number of samples. In addition, standard deviation value is also given to demonstrate variation of the prediction over different sample sub-sets.

As a result of all the analyses, the mean predictive RMSE is found to be between 0.160 and 0.301 as presented in Table 1. Among all the kernel types, the results appear to suggest that the radial basis function kernel is the best representative of the skin age data set yielding the lowest RMSE of 0.160 for the testing case. In addition, the following parameter sets were obtained for each kernel function, which have resulted in the best predictive SVR model with the values of parameters as

- C=7.5, ϵ=0.01 for the Linear Kernel function
- C=1.5, ϵ=0.05, γ=0.5 for the Polynomial Kernel
- C=3.5, ϵ=0.03, γ=1 for the Radial Basis function
- C=2.0, ϵ=1.00, γ=1 for Sigmoidal Basis function

TABLE I

RESULTS FOR PREDICTION OF ACTUAL SKIN AGES BASED ON RGB & IR LIGHT BACK SCATTER READINGS AND SUPPORT VECTOR REGRESSION (SVR)

SVR Kernel Type	Training RMSE (Standard Deviation)	Testing RMSE (Standard Deviation)
Linear	0.229 (0.008)	0.173 (0.046)
Polynomial	0.227 (0.012)	0.181 (0.051)
Radial Basis Function	**0.225 (0.004)**	**0.160 (0.023)**
Sigmoidal Basis Function	0.301 (0.002)	0.294 (0.007)

The results appear to suggest that the predictive models have performed better for the testing than at the training stages as RMSE for the testing is found to be lower than those obtained at the training stages. However, the standard deviation is higher for all the testing samples suggesting that not all the testing samples were predicted as accurately as seen in the average RMSE. The standard deviation figure is therefore an important value that helps assess if and how the prediction would be affected over different sub-sets of the samples. It is also observed that the results suggest that all the methods are not that much sensitive to the variations over the sub-sets of the samples studied, particularly at the training stages. However, the sigmoidal kernel function yielded the lowest standard deviation in both the training and testing cases although it generated the least accurate model.

As a conclusion, the SVR-based predictive model with the radial basis function kernel and the values of optimization parameters ($C=3.5$, $\epsilon=0.03$, $\gamma=1$) are found to form the optimal predictive model for skin age.

IV. Conclusions

The study presented in this paper introduces a cost effective and non-invasive unified measurement technique that can be used in not only skin based researches carried out by the academic institutions but also the assessment of skin health and identification of actual skin age (similar to assessment of biological age) by healthcare and cosmetic practitioners, who cannot afford high-cost instrumentations such as confocal microscopy. The proposed work also provides an advantage of cost-effectiveness and practicality of rapid inspection over the high-cost microscopy techniques [1] which are currently used for the comprehensive analysis of skin-aging process.

To the best of our knowledge, this successful pilot study is the first of its kind, and the computational method and very low-cost equipment developed have resulted in a promising result that demonstrates the capability of accurately predicting actual skin age rather than assigning a subject to one of the previously determined skin age classes. This is therefore expected to have a great potential in both healthcare and cosmetic sectors, particularly for personalized skin care. This is also further expected to help objectively assess the ultimate results of anti-aging cosmetic products.

Given the encouraging results, further research is now being geared towards utilizing similar and different light-scatter methods to be able to better assess the skin age. This will include the laser-speckle effect along with textural information of skin as similar to our previous studies where the detection of skin abnormalities [3] and classification of skin age groups [9] were studied. In addition, machine learning methods such as SVR will be further investigated and optimized to be able to improve the predictive performance of skin age as it has been shown in recent studies [15].

References

[1] A.N. Rice, "Laser microscopy tests for skin damage," *Bio Photonics*, 2013.

[2] T. Ryan, "The ageing of the blood supply and the lymphatic drainage of the skin," *Micron*, vol. 35, no. 3, pp. 161–171, 2004.

[3] A.B. Orun, E. Goodyer, H. Seker, G. Smith and V. Uslan, "Optimized parametric skin modelling for diagnosis of skin abnormalities by combining light back-scatter and laser speckle imaging" (Accepted for publication) *Skin Research and Technology*, 2014. doi:10.1111/srt.12142.

[4] C. Longo, A. Casari, F. Beretti, A.M. Cesinaro and G. Pellacani, "Skin aging: In vivo microscopic assessment of epidermal and dermal changes by means of confocal microscopy," *Journal of the American Academy of Dermatology*, vol. 68, no.3, pp.e73 – e82, 2013.

[5] A.N Bashkatov , E.A. Genina, V.I. Kochubey and V.V. Tuchin, "Optical properties of human skin, subcutaneous and mucous tissues in the wavelength range from 400 to 2000nm," *J. Phys. D: Appl. Phys*.,vol.38, pp.2543–55, 2005.

[6] K.R. Forrester, "Blood-flow imaging adapted for endoscopy," *Biophotonics International*, 2003.

[7] G. Mahé, P. Rousseau, S. Durand, S. Bricq, G. Leftheriotis, P. Abraham, "Laser speckle contrast imaging accurately measures blood flow over moving skin surface,", *Microvascular Research*, vol. 81, no. 2, pp. 183-188, 2011.

[8] C. S. Agarwal, J. Allen, A. Murray and I. F. Purcell, "Laser Doppler assessment of dermal circulatory changes in people with coronary artery disease", *Microvascular Research*, vol. 84, no.1, pp.55-59, 2012.

[9] A.B. Orun, H. Seker, E. Goodyer, G. Smith and V. Uslan, "An improvement of skin aging assessment by non-invasive laser speckle effect: A comparative texture analysis", (Accepted) Proc. of the IEEE-EMBS International Conferences on Biomedical and Health Informatics, Valencia, Spain, 1-4 June 2014.

[10] V. N. Vapnik, "An overview of statistical learning theory", IEEE Transactions on Neural Networks, vol.10, no. 5, pp. 988 – 999, 1999.

[11] B. Scholkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond*. MIT press, 2002.

[12] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines", ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, 2011.

[13] Hojjatoleslami, S.A., E. Claridge and M. Moncrieff. *Accuracy of the skin model in quantifying blood and epidermal melanin*, London, MIUA2000, 10-11 July 2000.

[14] Claridge E, Cotton S, Moncrieff M, Hall P, "Spectrophotometric Intracutaneous Imaging (SIAscopy): Method and clinical applications". Chapter 37 in *Handbook of Non-Invasive Methods and the Skin* (2nd ed). Serup J, Jemec GBE, Grove G (Eds), pp. 315-325: CRC Press, 2006.

[15] V. Uslan and H. Seker, "Support vector-based fuzzy system for the prediction of mouse class I MHC peptide binding affinity", Proc. of the 13th IEEE International Conference on BioInformatics and BioEngineering (IEEE BIBE2013), Chania, Crete, Greece, 10-13 November 2013, pp:1-4.