https://doi.org/10.2352/ISSN.2470-1173.2018.8.MAAP-152

Three-dimensional hyperspectral imaging: a new method for human face acquisition

Lou GEVAUX [1], Cyprien ADNET [2], Pierre SÉROUL [2], Raphael CLERC [1], Alain TRÉMEAU [1], Jean Luc PERROT [3], Mathieu HÉBERT [1];
[1]Univ Lyon, UJM-Saint-Etienne, CNRS, Institut d'Optique Graduate School, Laboratoire Hubert Curien UMR 5516, F-42023, Saint-Etienne, France.
[2]Newtone Technologies, 13 bis Place Jules Ferry, F-69006 Lyon, France
[3]Department of Dermatology, University Hospital of Saint Etienne, 42000 Saint-Etienne, France

Abstract

Hyperspectral imaging is an emerging non-invasive method for the optical characterization of human skin, allowing detailed surface measurement over a large area. By providing the spectral reflectance in each pixel, it enables not only color simulation under various lighting conditions, but also the estimation of skin structure and composition. These parameters, which can be correlated to a person's health, are deduced from the spectral reflectance in each pixel thanks to optical models inversion. Such techniques are already available in 2D images for flat skin areas, but extending them to 3D is crucial to address large scale and complex shapes as in the human face. The requirements for accurate acquisition are a short acquisition time for in vivo applications and uniform lighting conditions to avoid shadowing. The proposed method combines wide field hyperspectral imaging and 3D surface acquisition, using a single camera, with an acquisition time of less than 5 seconds. Complete color consistency can be achieved by computationally correcting irradiance non-uniformities using 3D shape information.

1. INTRODUCTION

Skin is a complex layered organ composed of several types of cells, fibers, glands, vessels, pigments and blood, which together generate a heterogeneous distribution of light absorption and scattering [1]. The measurement of skin optical properties can offer crucial information about a person's health, leading to a growing demand for optical devices in the fields of cosmetology and dermatology. In recent years, research in the measurement of skin optical properties has focused on non-invasive in vivo methods as ex vivo methods are known to alter the properties of the skin, and as invasive analysis may cause patient discomfort. Several optical methods have emerged, such as diffuse reflectance spectroscopy [2], modulated imaging [3], spectral imaging, and color imaging [4] [5]. Data retrieved from these methods have been analyzed using various models such as the Monte Carlo method [6], the radiative transfer equation [7], the diffusion approximation [3], flux models such as the Kubelka Munk model [5], and independent component analysis [8].

Hyperspectral imaging [9] is a non-invasive and contactless method that provides more information than standard color imaging and may reveal details invisible to the human eye. Moreover, it can provide information related to color appearance under various illuminants as well as physicochemical composition. By analyzing the acquired spectral reflectance distribution over a large area using an optical model combined with a skin structure model, various skin composition properties can be estimated pixel per pixel, then displayed in the form of images or maps. Since the capacity of this technique to reveal fine details is directly related to the image resolution, an imaging system with high spatial and spectral resolution is preferable. However, as the computation time needed to analyze the spectral information increases proportionally to the number of pixels, the skin model and the optical model chosen should be as simple as possible, to make the time needed for data analysis acceptable. In a previous work [5], a Kubelka-Munk inspired model applied onto a two-layer skin structure was shown to give satisfactory results by computing up to five skin composition maps of 1120 × 900 pixels in less than one hour. The hyperspectral images were captured by using the SpectraCam® camera [5] [10] developed by Newtone Technologies. This technique, however, is limited to small planar skin areas, as the effect of shadows on complex three-dimensional objects is not taken into account. The combination of this technique with 3D images of the observed skin area thus appears as a solution to surpass the aforementioned limitations and address large, non-flat areas such as a human face.

The few existing near field 3D-spectral imaging systems do not seem really appropriate to high spectral and spatial measurement on living organs. Either they are multispectral systems (e.g. [11] [12] [13]) providing too low spectral resolutions for optical analysis, or in the case of those 3D-spectral imaging systems that do provide high spectral resolutions (e.g. [14]), the acquisition process is not adapted to living subjects.

This is the reason why we designed the 3D-spectral measurement setup dedicated to skin presented in this paper, which gives a satisfactory tradeoff between high resolution and short acquisition time. In addition to this device, solutions for ensuring good color consistency and reducing shadowing effects are investigated. We focus here on the human face, the part of the body that presents the biggest challenge for 3D scanning, as well as having the highest interest for cosmetology, to create a robust device that is capable of scanning every part of the human body.

The paper is structured as follows: the measurement setup is presented first in Section 2. In the following, the 3D acquisition (Section 3), the hyperspectral acquisition (Section 4) and the optical treatment of the measured data (Section 5) are introduced, before presenting experimental results in Section 6. The question of irradiance correction is briefly discussed in Section 7, before reaching conclusions.

2. EXPERIMENTAL SETUP

The proposed 3D hyperspectral acquisition system is an extension of the existing hyperspectral camera, SpectraCam® [5] [10], incorporating the acquisition of depth map information and allowing wide field acquisition. It is composed of a monochrome CMOS camera of 2048 × 2048 pixels and a Liquid Cristal Tunable Filter (LCTF).

The setup comprises two lighting systems, used sequentially: one for the depth map acquisition, one for the hyperspectral acquisition. The first lighting system is a portable digital LED projector used to project fringes during the 3D scanning step of the acquisition process. The angle between the camera and the projector

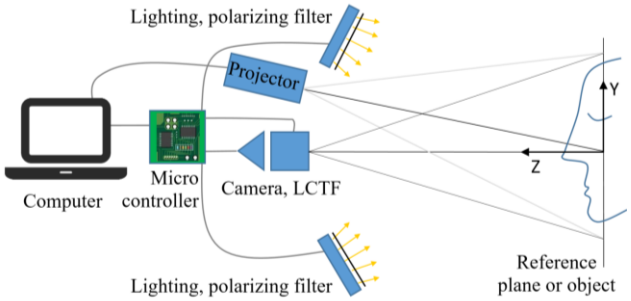

Figure 1. Schematic representation of the acquisition system.

is approximately 10 degrees, to enable a full face acquisition while minimizing the occluded areas. The second lighting system comprises two LED lighting units designed with blue and white LEDs to provide sufficient irradiance on the skin over the whole visible spectrum.

Linear polarizing filters are added on each lighting unit. They are oriented in order to create a cross polarized (CP) configuration with the LCTF polarization direction, this latter playing the role of an analyzer for the reflected light. The CP configuration is used to remove the gloss component of the reflected light: when light interacts with a diffusing medium, a portion of the flux is reflected at the air-medium interface according to the Fresnel law, without any change in polarization. This specular reflection of light, related to the visual attribute called gloss, is removed by the analyzer in a CP configuration. The remaining portion of flux enters the medium where it is scattered, progressively losing its polarization state, and finally emerging from the medium as an unpolarized light that can partially cross the analyzer. In the rest of the paper, the specularly reflected light will be discarded and only the unpolarized light component will be considered.

The camera, LCTF and lighting are controlled with a micro controller in order to shorten acquisition time. The acquisition is driven with a dedicated software written in C++.

A plane is also needed for reference and calibration. This plane is positioned perpendicularly to the camera's optical axis; together they define the reference coordinates system (X, Y, Z) shown in Figure 1, where Z corresponds to the optical axis. The camera and projector pupils are located in the same plane $Z = L$. The reference plane is mounted on a translation system which allows precise translations along the Z axis.

Hyperspectral images over the visible spectrum (410 nm – 700 nm) and data for 3D reconstruction are sequentially acquired within five seconds, a duration that does not totally obviate the risk of the subject moving, but remains fairly acceptable.

3. THREE-DIMENTIONAL SCANNING

Various optical and imaging technologies provide non-contact methods for collecting the geometrical shape of an object [15]. These methods can be categorized into those that are passive and active.

Passive methods don't require any controlled illumination and rely on ambient light. Stereoscopic imaging, inspired by binocular vision, is one of the most commonly employed 3D passive techniques: by capturing a scene from at least two points of view, the 3D geometry can be estimated using the triangulation principle. A given point has to be located in both views as a first step of the 3D reconstruction. Therefore, binocular and multi-view stereovision consist in solving a correspondence problem.

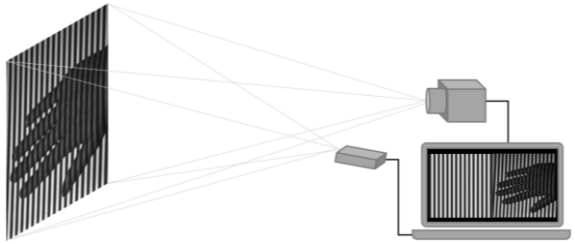

Figure 2. Illustration of 3D scanning based on structured light projection.

Active methods are based on the same triangulation problem. In contrast with stereovision, one camera is replaced with a controllable illumination source that projects a spot, a line or a structured light onto the object. Single point laser scanning emerged in the 1970s, allowing a 3D acquisition without any ambiguity. A similar concept using a sweeping line remains a popular solution for 3D acquisition of static scenes or objects that can be placed on a turntable. However, as the laser scanning process is slow and limited to static scenes, structured light projection techniques are now widely used [16] [17].

The use of a digital projector to project structured light into a full field-of-view replaces the temporal sweeping of the laser beam. This method is faster but not ambiguity-free. Many suggestions have been proposed to overcome the correspondence problem: for example, coded pattern projection (known geometries of spots, color coded fringes, binary coded fringes, etc.) or sinusoidal fringe projection (Fig. 2) with phase unwrapping algorithms [18].

The scanning system that we developed is based on a fringe projection method, offering an excellent compromise between speed, quality and cost. This method has already proven its effectiveness in biomedical applications such as 3D intra-oral dental measurements, human body shape measurements, body deformation measurements and in vivo measurements [19]. Structured light projection also allows the use of a single camera for both hyperspectral and 3D acquisition, which ensures a good pixel correspondence between the 3D shape and the texture.

3.1 Triangulation principle

Sinusoid fringes are projected onto an object and deformed in function of the object geometry. The measurement of the phase deformation yields a depth map of the scene from the triangulation principle [17].

Once the phase deformation is known, the triangulation principle is applied to retrieve the depth information between the object and the reference plane. Let us consider a perfect projector whose pupil is centered in a point E, and a perfect camera whose pupil is centered in a point D (Fig. 3). Let us also consider a reference plane where the optical axes of the camera and the projector intercept in points G and C, respectively. Let us now consider a point F on the object, separated from the reference plane by a length h called the depth. Point F is illuminated by a ray from the projector that would intercept the reference plane in a point A were it not for the presence of the object; point F is also viewed by the same pixel (x,y) in the camera as the point B on the reference plane. According to the pinhole model, the depth h satisfies the following equation:

$$\frac{\mathrm{AB}}{\mathrm{DE}} = \frac{h}{\mathrm{DG} - h} \tag{1}$$

By denoting the distance DG between the reference plane and the camera's pupil as L; the distance DE between the camera and projector pupils as b; and the fringe frequency on the reference plane as f, the phase difference between points A and B is:

$$\Delta\varphi = \varphi_{obj} - \varphi_{ref}$$

where φ_{obj} and φ_{ref} are respectively the phase at point F and at point B.

From equation (1), it is possible to write the relationship between the depth and the phase difference:

$$h = \frac{L}{1 + \frac{2\pi f b}{\Delta\varphi}} \tag{2}$$

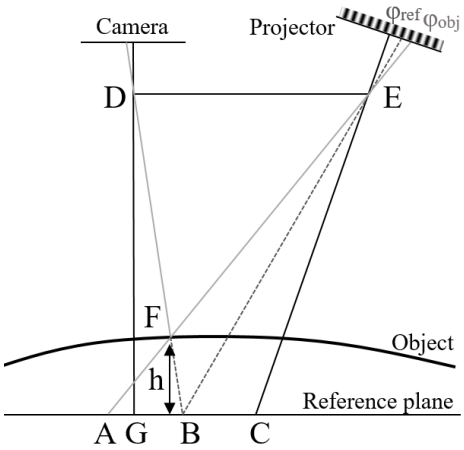

Figure 3. Illustration of the triangulation principle.

3.2 Phase shift principle

In order to get information on each pixel (x,y), the phase shift method is used. N fringe images, defined by the following equation, are projected:

$$I_n = I' + I''\cos\left(2\pi f x + \frac{2\pi(n-1)}{N}\right), \quad n = 1,\ldots \tag{3}$$

where I' is the average intensity and I'' is the intensity modulation.

The projected images represent identical fringes, each of which are successively shifted by $2\pi/N$. On the projected images, the phase, $2\pi fx$, depends on the pixel location x. On the images J_n, $(n=1,\ldots N)$ recorded by the camera, the phase is modified and can be computed using:

$$\varphi = \arctan\left(\frac{\sum_{n=1}^{N} J_n \sin\left(\frac{2\pi(n-1)}{N}\right)}{\sum_{n=1}^{N} J_n \cos\left(\frac{2\pi(n-1)}{N}\right)}\right) \in [-\pi, \pi] \tag{4}$$

The selection of the N value depends on the application requirements. When N increases, the acquisition time also increases but noise is reduced [20]. The observed noise, a consequence of the non-linearity of the projector-camera grey scale, is specific to fringe projection 3D acquisition: it is periodic, with a period of N times the period of the projected fringes. A N value around 5 or 6 is a good compromise between reduced noise (which is already much lower than with $N = 3$) and acquisition time.

3.3 Phase unwrapping algorithm

The phase calculated using (4) is discontinuous since its values are restricted in the interval $[-\pi, \pi]$. The continuous phase is retrieved through a phase unwrapping algorithm. Phase unwrapping techniques can be classified into two categories: temporal phase unwrapping [21] and spatial phase unwrapping. Temporal phase unwrapping relies on the projection of a large number of images (binary coded fringes, multiple frequency acquisition…) to acquire the necessary information to solve the unwrapping problem. Spatial phase unwrapping uses the continuity between pixels to retrieve the original phase. Therefore, an intuitive method to correct spatial discontinuities is to assume that the phase difference between two neighboring pixels is always less than π: the phase has to be unwrapped each time the difference between two neighboring pixels is higher than π. However, this method is not robust to noise and errors propagate, necessitating more complex solutions. Since a temporal method would increase the acquisition time, the chosen solution is Ghilia's "2D unweighted phase unwrapping" [22] based on the discrete cosine transform (DCT), which is a spatial method sufficiently robust to noise for facial acquisition (Fig. 4).

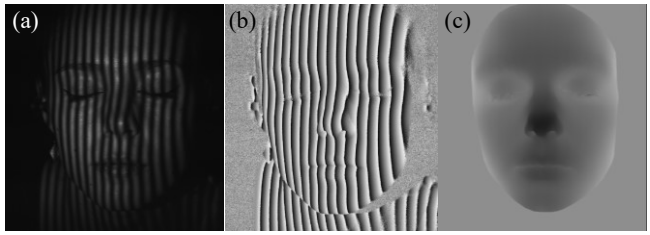

Figure 4. 3D acquisition intermediate results: recorded fringe image (a), wrapped phase computed from 6 fringe images (b), unwrapped phase on a selected zone of interest (c).

3.4 Calibrations

The acquisition process comprises three calibration steps for 3D scanning: a radiometric calibration which has to be done prior to generating the projected fringes, a depth calibration which allows the conversion of the phase onto a depth map, and an XY-plane calibration which allows the conversion of the depth map into (x,y,z) coordinates expressed in millimeters.

Radiometric calibration

The projected fringes must have a perfectly sinusoidal profile for optimal 3D reconstruction quality. However, the sinusoidal profile defined in the digital image that is projected is often transformed into another profile due to the radiometric nonlinearity of the projector and the camera. The radiometric calibration consists in measuring and correcting the grey scale non-linearity of the projector-camera system. Various methods of radiometric calibration have been proposed [20]; the method chosen consists in creating corrected fringes to be projected. These fringes are not sinusoidal but designed in order to correspond to the chosen shape when seen through the camera after projection. To computationally correct the projected fringes, 20 grey levels are projected onto the reference plane. The best polynomial of degree two matching with the experimental curve is identified using a least-squares algorithm.

Finally, a lookup table is created to link each targeted grey level of an 8-bit scale with the value that has to be input before projection.

When projected onto skin or any turbid media, light diffusion also affects the shape of the projected fringes. Since estimating this phenomena is complex, the 3D scanning is done, when possible, using a wavelength for which the media is as opaque as possible [19]. As the opacity of skin is highest in the short visible wavelengths, we have selected the blue channels for our imaging systems.

Depth calibration

The conversion of phases into depth information from the triangulation principle described by (2) requires knowledge of the system geometrical parameters L, b and f. These parameters, in addition to being dependent on the pixel location for the fringe frequency f, are difficult to measure accurately. Therefore, a calibration must be performed. The simplest calibration is linear, based on the approximate assumption that the phase difference is directly proportional to the depth when L is much larger than h. A non-linear calibration which does not rely on this approximation is also possible. The implemented non-linear calibration, described by Jia [23], requires the acquisition of several phase maps on planes located at several positions (h_1, h_2 …) along the Z axis (Figure 5) to find two coefficients M and N such as:

$$h(x,y) = \frac{\Delta\varphi(x,y)}{M(x,y) + N(x,y)\Delta\varphi(x,y)} \tag{5}$$

A minimum of three acquisitions is required to identify the coefficients, but more measurements can be performed to give a better estimate of M and N by applying a least-squares algorithm.

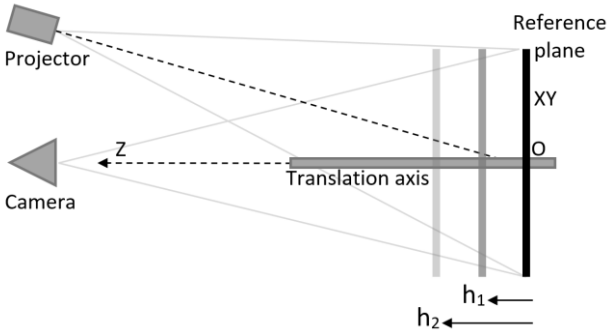

Figure 5. Scheme of the depth calibration setup.

XY-plane calibration

The depth calibration yields the coefficients M and N in mm^{-1} that are used in equation (5) to compute the depth map. The XY-plane calibration step consists in converting the depth map into a point cloud that represents the object surface in the real world coordinates system, which can also be described as the "2.5D" information.

Several methods [24] [25] have been developed to fully characterize the camera and projector intrinsic parameters (focal length, pixel sizes, and optical distortion) in order to perform an accurate XY-plane conversion. The method that we implemented is basic and assumes that the zone of interest is not affected by optical distortion and that the acquired depth range is sufficiently small to have the same pixel size for the whole object. A ruler is imaged in order to obtain the pixel-to-millimeter ratio.

4. HYPERSPECTRAL IMAGING

Hyperspectral imaging is an extension of classical RGB imaging where the number of channels is much higher, corresponding to a number of narrow spectral bands continuously covering the visible spectrum (and possibly UV and IR domains). Hyperspectral image data are also called "hypercubes" since they are three dimensional: two dimensions for the spatial domain, one dimension for the spectral domain. The information captured in each pixel is proportional to the spectral radiance coming from one direction in the camera's field. By dividing the spectral radiance measured from the object by the spectral radiance measured from a perfectly white diffuser that is illuminated in the same way as the object, one obtains in each pixel its spectral reflectance factor. This latter is often considered as a spectral reflectance, which is theoretically true only when the object is a Lambertian reflector.

Many hyperspectral imaging systems have been developed based on various technologies and physical principles, with different specificities in terms of spatial resolution, spectral resolution, surface shape constraints and acquisition time [9]. The staring approach, which is the method implemented in the proposed system, records the full spatial information for each single spectral band and requires a temporal scan of the spectral dimension. It relies either on active illumination or on adding a filtering system in front of the sensor, which can be a filtering wheel, a tunable filter or an interferometer. Active illumination methods seem promising [26] in terms of cost and acquisition time, given the availability of a wide range of LEDs with different emission spectra. However, since we are focusing on the face, consistent lighting over a large area is important, which seems difficult to achieve with active illumination methods. Therefore, we prefer a method using filters.

The chosen method is based on a Liquid Cristal Tunable Filter (LCTF) which offers the possibility of a rapid filtering scan within the visible spectrum with a sufficiently high spectral resolution for skin study. This kind of filter is based on the association of several Lyot filters [27], and is equivalent to high quality interference filters with a band width of 10 nm and a precision of 1 nm, between 400 and 720 nm.

4.1 Spectral calibration

Once the hyperspectral acquisition is performed, the luminance L_{obj} measured on each pixel for each wavelength depends on the surface reflectance properties of the object R_{object}, the incident irradiance $E_{lighting}$, the sensor sensitivity S_{sensor}, the camera transmittance T_{camera} and the background noise n:

$$L_{obj} = R_{obj} \cdot E_{lighting} \cdot S_{sensor} \cdot T_{camera} + n \tag{6}$$

In the case of hyperspectral imaging, the quantities in equation (6) are functions of wavelength and pixel location. Hyperspectral calibration aims at retrieving the surface reflectance information from the measured luminance. $E_{lighting}$, S_{sensor}, T_{camera}, and n are unknown quantities that are calibrated through the acquisition of black and white reference Lambertian samples of respective albedo ρ_{black} and ρ_{white}. The calibration results in the object reflectance for each pixel and each wavelength after applying the formula:

$$R_{obj} = \frac{1}{\pi}\left(\frac{L_{obj} - L_{black}}{L_{white} - L_{black}}\left(\rho_{white} - \rho_{black}\right) + \rho_{black}\right) \tag{7}$$

Figure 6. Acquired and corrected hyperspectral images converted into RGB color (illuminant D65): (a) Non uniform illumination on the white reference plane, (b) uncalibrated image (c) calibrated image

The calibrated image (Fig. 6.c) seems darker than the usually perceived visual appearance of skin but the resulting color is closer to the actual skin reflectance.

This calibration method gives satisfactory results when applied on flat objects. However, for 3D objects, the incident irradiance on the reference samples is not the same as that on the object, with the difference resulting in an error in the calibration output. For non-planar objects, the actual surface reflectance can be assessed through a more complex photometric calculation that requires the surface orientation and location, and knowledge of the incident lighting intensity in space. As the developed system provides both the spectral luminance and the 3D surface information, a "3D irradiance calibration", mentioned later, can be implemented for a consistent color measurement.

5. OPTICAL ANALYSIS

The optical analysis relies on two models: one for the skin structure that represents the skin tissues as a stack of two layers, and one for the light-skin interaction which predicts the reflectance of the skin according to the absorption and scattering properties of the two layers.

5.1 Skin model

Our model considers the skin as a two-layer medium, corresponding roughly to the epidermis and the dermis, whose composition and optical properties are noticeably different. The real skin structure is much more complex, as the epidermis and dermis can be further divided into several sub-layers, each of which have their own specific cell structures and chemical compositions [1]. The upper layer, the epidermis, is composed of the stratum corneum, the stratum lucidum, the stratum granulosum, the stratum spinosum and the stratum basal. Keratinocytes and melanocytes, two kind of cells that produce melanin and keratin fibers, can be found in the epidermis. The second layer, the dermis, is composed of the papillary layer and the reticular layer. They contain collagen fibers, elastin fibers, fibroblasts and erythrocytes (or red blood cells) carrying hemoglobin. They are traversed by blood vessels and capillaries. Beneath the dermis, the hypodermis is an elastic layer that includes a large amount of fat cells. Moreover, the dermal-epidermal junction in not flat: the dermis forms upward finger-like projections into the epidermis called dermal papillae [28].

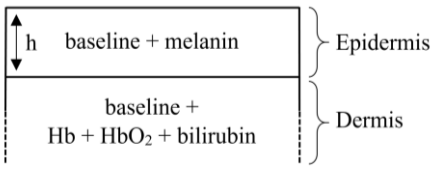

Figure 7. Skin model used in the optical analysis of hyperspectral data.

The two-layer model for skin (Fig. 7) is a simplification in which each layer is flat and homogeneous. The epidermis is composed of a baseline and melanin, and the dermis is composed of a baseline, deoxyhemoglobin (Hb), oxyhemoglobin (HbO_2) and bilirubin.

Skin, modeled as a two-layer turbid medium, can be characterized by the absorption and scattering coefficients of each layer. The absorption coefficients depend on the spectral absorption properties of chromophores contained in the layer and their volume fraction: according to the Beer-Lambert-Bouguer law, the spectral absorption coefficients of the epidermis $K_e(\lambda)$ and dermis $K_d(\lambda)$ can be written according to additive linear laws:

$$K_e(\lambda) = (1 - c_m)K_b(\lambda) + c_m K_m(\lambda) \tag{8}$$

and

$$\begin{aligned} K_d(\lambda) = {} & (1 - c_{Hb} - c_{HbO2} - c_{bi})K_b(\lambda) \\ & + c_{Hb}K_{Hb}(\lambda) + c_{HbO2}K_{HbO2}(\lambda) + c_{bi}K_{bi}(\lambda) \end{aligned} \tag{9}$$

We refer to reference [29] for the spectral absorption coefficient values of the baseline $K_b(\lambda)$, of melanin $K_m(\lambda)$, of hemoglobin $K_{Hb}(\lambda)$, of oxyhemoglobin $K_{HbO2}(\lambda)$ and of bilirubin $K_{bi}(\lambda)$, as well as for the spectral scattering coefficients of the epidermis $S_e(\lambda)$ and dermis $S_d(\lambda)$. The volume fractions of melanin c_m, of hemoglobin c_{Hb}, of oxyhemoglobin c_{HbO2} and of bilirubin c_{bi} are unknown quantities that can be estimated through an optimization based on the Kubelka-Munk theory.

5.2 The Kubelka Munk theory

The hyperspectral reflectance results are analyzed using a two-flux inverse model described in Ref. [5], where the light propagation in the skin layers is modeled according to the Kubelka-Munk theory [30], whose advantage in comparison with other light scattering models is that it provides analytical formulas for the reflectance and transmittance of a multilayer scattering structure. The epidermis is characterized by its thickness h_e, spectral absorption coefficient $K_e(\lambda)$ and spectral scattering coefficient $S_e(\lambda)$. Its proper reflectance R_e and transmittance T_e, expressed according to the Kubelka-Munk model [32], are (for each wavelength of light):

$$R_e = \frac{\sinh(b_e S_e h_e)}{b_e \cosh(b_e S_e h_e) + a_e \sinh(b_e S_e h_e)} \tag{11}$$

and

$$T_e = \frac{b_e}{b_e \cosh(b_e S_e h_e) + a_e \sinh(b_e S_e h_e)} \tag{12}$$

with

$$a_e = (K_e + S_e)/S_e \quad \text{and} \quad b_e = \sqrt{a_e^2 - 1} \tag{13}$$

The dermis is assumed to be opaque, therefore infinitely thick from an optical point of view. It is characterized by its spectral absorption coefficient $K_d(\lambda)$ and spectral scattering coefficient $S_d(\lambda)$, and its reflectance R_d predicted according to the Kubelka-Munk model is given by:

$$R_d = a_d - b_d = \frac{K_d + S_d - \sqrt{K_d(K_d + 2S_d)}}{S_d} \quad (14)$$

where parameters a_d and b_d are functions of $K_d(\lambda)$ and $S_d(\lambda)$ defined similarly as in equation (13).

The total spectral reflectance of the skin, superposition of the dermis and epidermis, is given by Kubelka's formula [30]:

$$R_{skin} = R_e + \frac{T_e^2 R_d}{1 - R_e R_d} \quad (15)$$

5.3 Inverse problem and optimization

For each pixel, the volume fraction of each component is estimated through an optimization based on a spectral angle similarity metric, the Spectral Angle Mapping (SAM) [31] (also equivalent to a normalized correlation).

The search for chromophore volume fractions requires the calculation of a theoretical reflectance spectrum $R_{th}(\lambda)$. The distance d defined by the SAM between $R_{th}(\lambda)$ and the measured spectrum $R_m(\lambda)$ is computed as:

$$d = \arccos\left(\frac{\sum R_{th}(\lambda) R_m(\lambda)}{\sqrt{\sum R_{th}^2(\lambda) \sum R_m^2(\lambda)}} \right) \quad (16)$$

The inverse problem consists in searching for the parameters that minimize d:

$$\{c_m, c_{Hb}, c_{HbO2}, c_{bi}, h\} = \underset{c_m, c_{Hb}, c_{HbO2}, c_{bi}, h}{\arg\min} \; (d) \quad (17)$$

Thanks to the use of the SAM metric, results of the skin analysis are very good even on shadowed region in the input data. Indeed, this metric improves the algorithm robustness by reducing its sensitivity to variations due to lighting non uniformities. As an example (illustrated by Figure 9), let's compare three regions of skin. The first is a reference region characterized by the measured spectral luminance $S_0(\lambda)$. The second region has the same chromophore composition as the first but with a lower incident irradiance, which implies the perceived color is darker and the spectral luminance is $S_1(\lambda) = kS_0(\lambda)$ with $k < 1$. The third region of skin differs in chromophore composition from the first two and is characterized by the measured spectral luminance $S_2(\lambda)$.

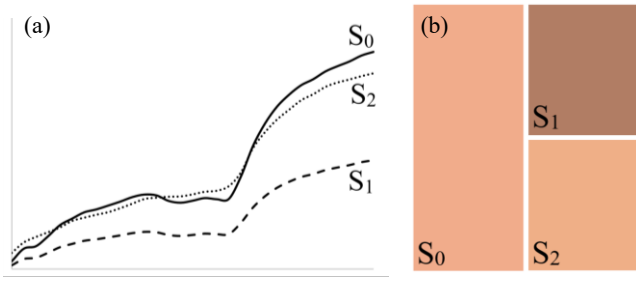

Figure 9. Spectral luminances (a) and corresponding RGB colors under D65 illumination (b). Distances to the reference spectrum S_0 according to the RMS metric and the SAM metric are: $d_{RMS}(S_1)=14$; $d_{RMS}(S_2)=2$; $d_{SAM}(S_1)=0$; $d_{SAM}(S_2)=6.10^{-2}$.

When comparing the three spectral luminances using the root-mean-square (RMS) metric, the distance between $S_0(\lambda)$ and $S_1(\lambda)$ is high while the distance between $S_0(\lambda)$ and $S_2(\lambda)$ is low due to the invariance of this metric in respect to lightness. This comparison is consistent with perceived color, which can be affected by varying irradiance, but doesn't represent the skin chromophore properties. When comparing the three spectral luminances using the SAM metric, on the other hand, the distance between $S_0(\lambda)$ and $S_1(\lambda)$ is null, while there is a noticeable distance between $S_0(\lambda)$ and $S_2(\lambda)$. This metric is more sensitive to changes in the spectrum shape and insensitive to irradiance non-uniformities, and therefore yields better results for skin analysis optimization.

6. RESULTS AND PERFORMANCE

We propose in this section to present the experimental results achieved with our method, considering in turn the different steps developed above: the 3D acquisition, the hyperspectral acquisition, and the optical analysis. Lastly, 3D maps of skin parameters will be presented and discussed.

6.1 3D meshes

3D scanning based on fringe projection results in a point cloud representing the surface of the object. This point cloud has to be converted into a mesh in order to add a texture on it and display it as a 3D object.

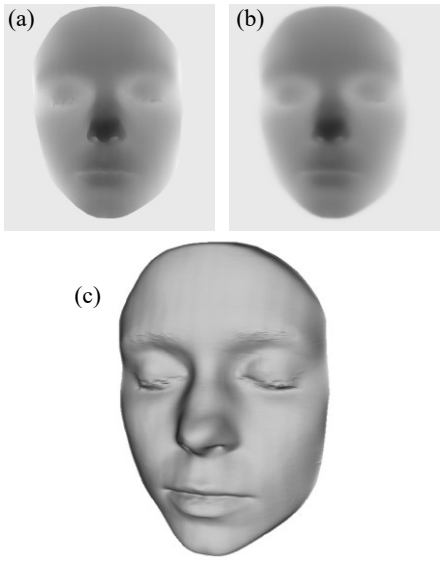

Figure 10. 3D hyperspectral acquisition results: (a) depth map before smoothing, (b) depth map after smoothing, (c) 3D mesh without texture.

Before conversion, the depth map is filtered with a low pass Butterworth filter in order to reduce noise (Fig. 10.b). After noise reduction, the filtered depth map is converted into a .vtk 3D file using a function from The Visual Toolkit (VTK) [32] library (see Figure 10.c).

The precision of the acquisition system has been defined by measuring a tilted plane at several inclinations, with and without smoothing. The maximum, average and standard deviations are presented in Table 1, before and after smoothing (Fig. 11). The smoothing step consists in reducing the periodic noise and filtering high frequencies. Table 2 shows the tilt angle between the XY plane and planes A and B, measured from the 3D data and manually.

Table 1: 3D scanning precision on the tilted plane A

Tilted plane A measurement	Maximum deviation	Average deviation	Standard deviation
Before smoothing	1.39 mm	0.25 mm	0.32 mm
After smoothing	0.51 mm	0.17 mm	0.20 mm

Table 2: Tilted plane A and B, angle measurement

	3D acquisition	Manual
Plane A	17.52°	17.35° ± 0.35°
Plane B	33.23°	34.02° ± 0.71°

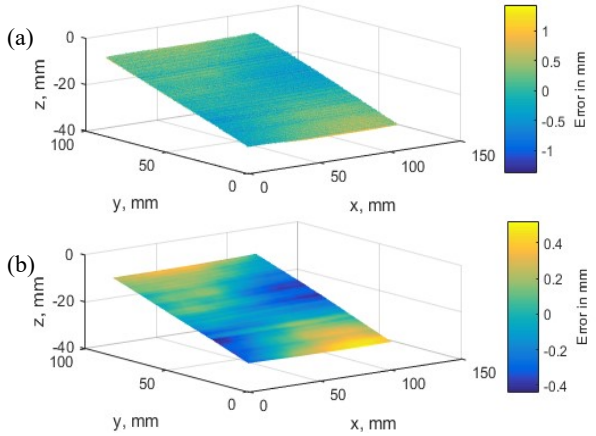

Figure 11. Point cloud display of the measured titled plane before smoothing (a) and after smoothing (b). The color corresponds to the deviation amplitude.

The analysis of the 3D results shows that the noise is not white but is the sum of white noise, periodic noise and deformation (due to errors in the conversion of phase map into depth map). On an opaque plane such as the one used for the measurement, the noise is mainly the result of the projector-camera non linearity. This measured deviation has to be interpreted as a minimum value characterizing the acquisition method. For in vivo skin acquisition, the noise can also be caused by: unintentional movements during acquisition; blurred fringes due to the limited depth of field of the projector [20]; diffuse fringes due to the acquired object not being opaque but turbid. The acquisition setups have to be optimized to minimize noise: the acquisition time has to be short, the projected fringes have to be blue and the camera focus has to be carefully adjusted. The acquisition precision of the current method is sufficient for full face acquisition. However, for more detailed applications, such as measurement of wrinkles, greater precision would be required.

6.2 Hyperspectral acquisition results

Images obtained with the hyperspectral camera are hypercubes of 2048 x 2048 pixels x 30 wavelengths covering the visible spectrum (410 to 700 nm) with an acquisition time of approximately 2.5 seconds. From each hypercube, it is possible to calculate the RGB colors perceived under a given illuminant after spectral calibration (Fig. 13.d). Images obtained at particular wavelengths show the specific skin characteristics: melanin stains are visible at low wavelengths (Fig. 13.a), blood vessels appear between 560 nm and 600 nm (Fig. 13.b), and finally, when seen in infrared (Fig. 13.c), the skin appears uniform, more transparent and more diffusing.

Figure 12. Example of hyperspectral image: (a) Image at 420 nm, (b) image at 590 nm, (c) image at 700 nm, (d) RGB color reconstruction under a D65 illuminant.

The hypercube resolution is good enough for analysis, however the LCTF and the camera bring chromatic distortion, resulting in spectral errors on small objects, such as spots and hairs. Chromatic errors can also result from small movements of the subject during the acquisition process. This phenomenon is especially visible on the eyelid, which is susceptible to involuntary micro-movements caused by the lighting used during acquisition.

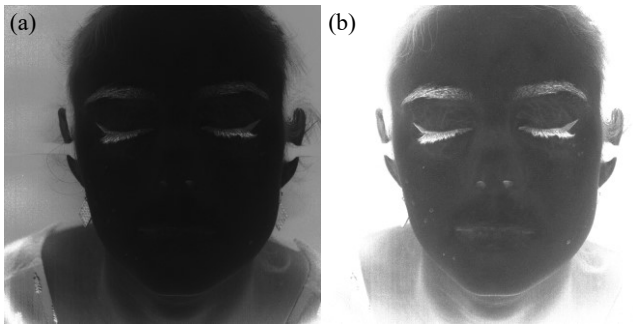

Figure 13. Error map from the skin optical analysis, (a) full dynamic (SAM ∈ [0, 0.5]) and (b) optimized dynamic (SAM ∈ [0, 0.1]) to observe errors on skin.

6.3 Optical analysis results

The skin analysis optimization yields chromophore concentrations in each pixel according to the method summarized in Section 5, then displayed under the form of maps that show the relative volume fractions of blood, oxygen and melanin (Fig. 14).

When the optical analysis is performed, an error map showing for each pixel the residual error (defined by Equation (16)) between the measured spectrum and the theoretical spectrum is created. The full dynamic error map (Fig. 13.a) shows that the optical model is good at distinguishing skin (low error) from other elements such as hair, fabric, make-up and background (high error). The error map on Fig. 13.b, in which the display dynamic has been optimized to allow

more details to be seen on the face shows that there is larger error on small elements, such as blood vessels on the eyelid. This is caused by the chromatic aberration of the acquisition system: on small elements or on borders, the pixel spectrum is modified by the chromatic aberration, impacting the pixel-by-pixel chromophore map. For now, we consider that the number of pixels impacted by this error is small enough to be negligible in the final image.

The error is also higher in areas where the incident irradiance varies, as visible on the neck, ears, lips and nostril sides. This shows that the SAM metric (part 5.3) is not sufficient to avoid the influence of varying irradiance. Another hypothesis for explaining errors in regions such as nostril sides and ears is that the geometry of the surface presents convex zones where interreflections [33] can occur. Interreflections consist of multiple light reflections within a cavity that modifies the perceived spectral luminance as a nonlinear function of the spectral reflectance of the concave surface. These spectral modifications impact the chromophore map reconstruction and results in an error. We intend to investigate this topic in the near future.

Finally, the optical analysis is based on the assumption that skin is Lambertian (ignoring the gloss component) and that the Kubelka Munk model can be applied in the same way, whether the observed surface is normal to the observation direction or not. However, for non Lambertian surfaces, the quantity of light reflected depends on the angle between the normal and the observation direction. An extended model taking into account the angular dependence of the skin reflectance would be necessary to further improve the accuracy of the method.

6.4 3D hyperspectral results

When the depth map is converted into a 3D mesh, texture is added: this can represent the reflectance information for a given wavelength, the color information, or the chromophore information calculated with the skin analysis algorithm (Fig. 15). Figure 15.b, c and d represent the results of the entire process, comprising 3D scanning, hyperspectral imaging and skin analysis. The visualization of the face in 3D can supplement the chromophore map information, potentially aiding the study of skin properties that implicate shape.

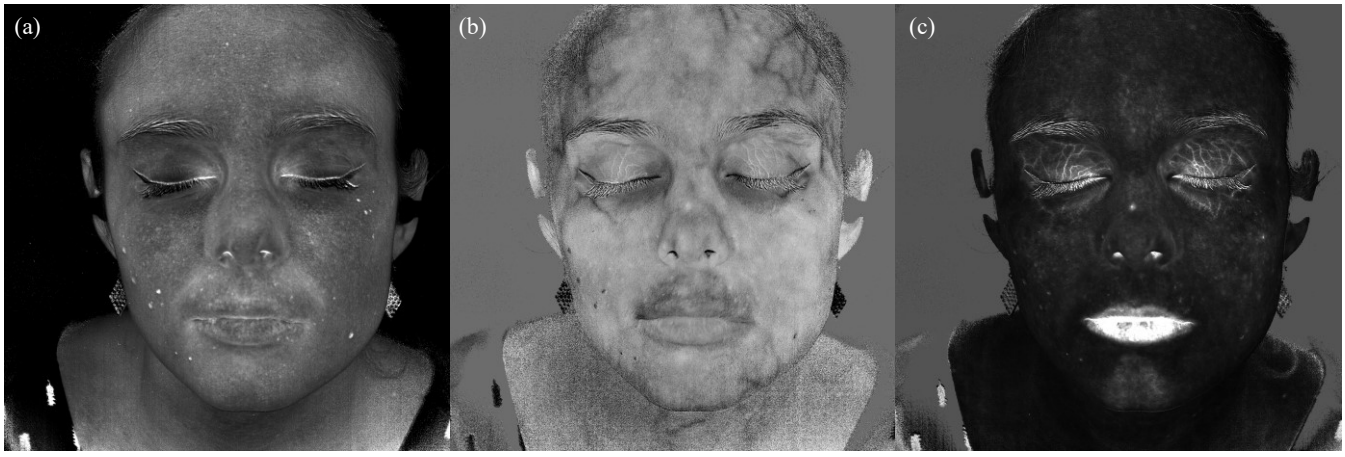

Figure 14. Optical analysis results, volume fraction map of: melanin (a), oxygen (b), blood (c).

Figure 15. 3D hyperspectral acquisition results: 3D visualization of the color (a), 3D visualization of the melanin map (b), 3D visualization of the oxygen rate map (c), 3D visualization of the blood map (d).

7. TOWARD IRRADIANCE CORRECTION

As mentioned in section 5.2, a simple luminance correction based on a flat reference sample is not sufficient to provide accurate spectral measurement. When a hyperspectral acquisition is performed on a full face, the reconstructed color is usually a faithful representation of the actual spectral reflectance only on the forehead

and the cheeks. To overcome this limitation, a calibration method taking into account the full 3D geometry of the object and lighting must be implemented.

The proposed method is based on two assumptions. Firstly, the surface is assumed to be Lambertian. Secondly, light sources are approximated as punctual sources. In the proposed system, each lighting unit is built with ten LEDs, each of which is considered as a punctual light source. The total incident irradiance is the sum of each LED contribution.

Let's consider a small area of the object surface that corresponds to one pixel (x, y) when imaged on the camera sensor. For each pixel, Bouguer's law and Lambert's law result in equation (18), which gives the luminance $L_{x,y}$ of the pixel as a function of: the incident intensity $I_k(\theta_{x,y,k}, \varphi_{x,y,k})$ of each LED k in the direction of the pixel; the 3D surface (i.e., the distance to the LED $d_{x,y,k}$ and the angle $\alpha_{x,y,k}$ between the surface normal and the LED axis); and the albedo $\rho_{x,y}$ of the surface:

$$L_{x,y} = \frac{\rho_{x,y}}{\pi} \sum_{LEDk} I_k(\theta_{x,y,k}, \phi_{x,y,k}) \frac{\cos \alpha_{x,y,k}}{d^2_{x,y,k}} \qquad (18)$$

The surface albedo, which is the parameter that has to be determined, can be retrieved by performing a lighting calibration using a black and a white flat sample to provide the missing information $I_k(\theta_{x,y,k}, \varphi_{x,y,k})$ for each LED.

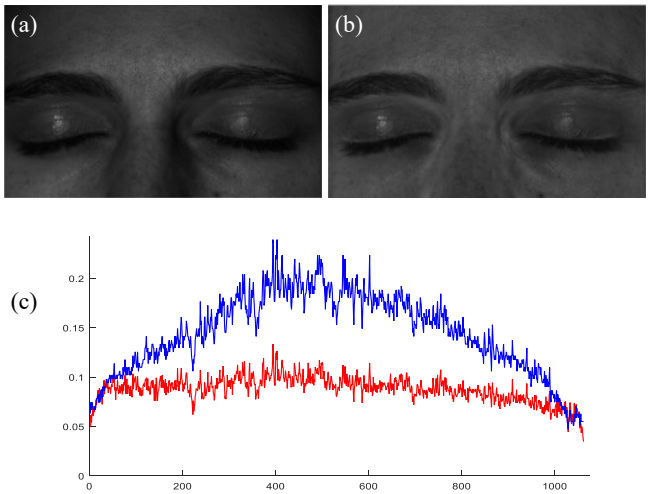

Figure 16. Original image with incident irradiance non-uniformities (a), image after correction (b), profile of the luminance intensity on the forehead for the non-corrected image (blue) and the corrected image (red) (c).

Figure 16 shows the results obtained in a simple configuration based on a single monochromatic source, before (Fig. 16.a) and after irradiance correction (Fig. 16.b). Figure 16.c shows the irradiance horizontal profile on the forehead before and after correction. In this example, the gloss component of skin reflectance is visible as no polarizing filter has been used. The implemented method corrects the irradiance variation observed on the forehead. The curve of the blue line on Figure 16.c is the consequence of the irradiance non uniformities on the forehead but doesn't describe the skin reflectance on this zone. The red line shows that the reflectance profile is more homogeneous after applying the irradiance correction. Local variations are due to skin texture on the forehead.

8. CONCLUSION AND DISCUSSION

The system developed is, as far as we could see, one of the first 3D hyperspectral imaging systems conceived for living tissues, satisfying good spectral and spatial resolutions and short acquisition time. The combination of acquisition system, shadow correction and optical models is also original and brings a new dimension to skin analysis. Moreover, the proposed system contributes to the general trend of replacing an object with its numerical mockup, allowing virtual analysis and interaction with low constraints.

The use of 3D information for correcting irradiance non-uniformity and improving the measurement of spectral reflectance was the first motivation for implementing a 3D scanning system. However, it has demonstrated high potential for other applications, such as providing the shape in addition to the color, or investigating the Bidirectional Reflectance Distribution Function (BRDF) of objects without shape constraints.

Although the chromophore maps show very interesting results, the implemented method for skin analysis does not currently take into account scattering, although it is an important parameter for skin. The quantification of scattering could widen the range of applications for full face skin analysis, providing a better description of skin properties such as inflammation and aging, which are of particular interest to dermatology and cosmetology.

The acquisition system quality could also be improved by using a brighter projector to shorten acquisition time and by offering 3D scanning from more than one point of view in order to ensure good 3D reconstruction on all parts of a human face.

Acknowledgement

This work was supported by the Région Auvergne-Rhône-Alpes within the program ARC 6, and the French National Research Agency (ANR) within the program "Investissements d'Avenir" (ANR-11-IDEX-0007), in the framework of the LABEX MANUTECH-SISE (ANR-10-LABX-0075) of Universite de Lyon.

References

[1] T. Igarashi, K. Nishino, S. K. Nayar and others, "The appearance of human skin: A survey," *Foundations and Trends® in Computer Graphics and Vision,* vol. 3, no. 1, pp. 1-95, 2007.

[2] G. Zonios, J. Bykowski and N. Kollias, "Skin melanin, hemoglobin, and light scattering properties can be quantitatively assessed in vivo using diffuse reflectance spectroscopy," *Journal of Investigative Dermatology,* vol. 117.6, pp. 1452-1457, 2001.

[3] D. J. Cuccia, F. Bevilacqua, A. J. Durkin, F. R. Ayers and B. J. Tromberg, "Quantitation and mapping of tissue optical properties using modulated imaging," *Journal of biomedical optics,* vol. 14, no. 2, pp. 024012-024012, 2009.

[4] M. Doi, A. Kimachi, S. Nishi and S. Tominaga, "Estimation of Local Skin Properties from Spectral Images and its Application to Appearance Reproduction," *Journal of Imaging Science & Technology,* vol. 60, no. 5, 2016.

[5] P. Seroul, M. Hébert, M. Cherel, R. Vernet, R. Clerc and M. Jomier, "Model-based skin pigment cartography by high-resolution hyperspectral imaging," *Journal of Imaging Science and Technology,* vol. 60, no. 6, pp. 60404-1, 2016.

[6] I. Nishidate, Y. Aizu and H. Mishina, "Estimation of melanin and hemoglobin in skin tissue using multiple regression analysis aided by Monte Carlo simulation," *Journal of Biomedical Optics,* vol. 9, no. 4, pp. 700-710, 2004.

[7] C. Magnain, M. Elias and J.-M. Frigerio, "Skin color modeling using the radiative transfer equation solved by the auxiliary function

method: inverse problem," *JOSA A,* vol. 25, no. 7, pp. 1737-1743, 2008.

[8] N. Tsumura, H. Haneishi and a. Y. Miyake, "Independent-component analysis of skin color image," *JOSA A,* vol. 16.9, pp. 2169-2176., 1999.

[9] G. Lu and B. Fei, "Medical hyperspectral imaging: a review," *Journal of biomedical optics,* vol. 19, no. 1, pp. 010901-010901, 2014.

[10] A. Nkengne, J. Robic, P. Seroul, S. Gueheunneux, M. Jomier and K. Vie, "SpectraCam®: A new polarized hyperspectral imaging system for repeatable and reproducible in vivo skin quantification of melanin, total hemoglobin, and oxygen saturation," *Skin Research and Technology,* 2017.

[11] V. C. Paquit, K. W. Tobin, J. R. Price and F. Mériaudeau, "3D and multispectral imaging for subcutaneous veins detection," *Optics express,* vol. 17, no. 14, pp. 11360-11365, 2009.

[12] K. Hirai, R. Nakahata and T. Horiuchi, "Measuring Spectral Reflectance and 3D Shape Using Multi-primary Image Projector," in *International Conference on Image and Signal Processing*, Springer, 2016, pp. 137-147.

[13] C. Zhang, M. Rosenberger, A. Breitbarth and G. Notni, "A novel 3D multispectral vision system based on filter wheel cameras," in *IEEE International Conference on Imaging Systems and Techniques (IST)*, IEEE, 2016, pp. 267-272.

[14] M. H. Kim, "3D Graphics Techniques for Capturing and Inspecting Hyperspectral Appearance," in *Ubiquitous Virtual Reality (ISUVR), 2013 International Symposium on*, IEEE, 2013, pp. 15-18.

[15] F. Remondino and S. El-Hakim, "Image-based 3D modelling: A review," *The Photogrammetric Record,* vol. 21, no. 115, pp. 269-291, 2006.

[16] S. S. Gorthi and P. Rastogi, "Fringe projection techniques: whither we are?," *Optics and lasers in engineering,* vol. 48, no. IMAC-REVIEW-2009-001, pp. 133-140, 2010.

[17] D. Lanman and G. Taubin, "Build your own 3D scanner: 3D photography for beginners," in *ACM SIGGRAPH 2009 Courses*, ACM, 2009, p. 8.

[18] J. Geng, "Structured-light 3D surface imaging: a tutorial," *Advances in Optics and Photonics,* vol. 3, no. 2, pp. 128-160, 2011.

[19] Y. Wang, J. I. Laughner, I. R. Efimov and S. Zhang, "3D absolute shape measurement of live rabbit hearts with a superfast two-frequency phase-shifting technique," *Optics express,* vol. 21, no. 5, pp. 5822-5832, 2013.

[20] Z. Li and Y. Li, "Gamma-distorted fringe image modeling and accurate gamma correction for fast phase measuring profilometry," *Optics letters,* vol. 36, no. 2, pp. 154-156, 2011.

[21] C. Zuo, L. Huang, M. Zhang, Q. Chen and A. Asundi, "Temporal phase unwrapping algorithms for fringe projection profilometry: A comparative review," *Optics and Lasers in Engineering,* vol. 85, pp. 84-103, 2016.

[22] D. C. Ghiglia and L. A. Romero, "Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods," *JOSA A,* vol. 11, no. 1, pp. 107-117, 1994.

[23] P. Jia, J. Kofman and C. English, "Comparison of linear and nonlinear calibration methods for phase-measuring profilometry," *Optical Engineering,* vol. 46, no. 4, pp. 043601-043601, 2007.

[24] S. Zhang and P. S. Huang, "Novel method for structured light system calibration," *Optical Engineering,* vol. 45, no. 8, pp. 083601-083601, 2006.

[25] D. Moreno and G. Taubin, "Simple, accurate, and robust projector-camera calibration," in *3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2012 Second International Conference on*, IEEE, 2012, pp. 464-471.

[26] M. Goel, E. Whitmire, A. Mariakakis, T. S. Saponas, N. Joshi, D. Morris, B. Guenter, M. Gavriliu, G. Borriello and S. N. Patel, "HyperCam: hyperspectral imaging for ubiquitous computing applications," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM, 2015, pp. 145-156.

[27] N. Gat, "Imaging spectroscopy using tunable filters: a review," in *Proc. Spie*, 2000, pp. 50-64.

[28] V. Newton, R. Bradley, P. Seroul, M. Cherel, C. Griffiths, A. Rawlings, R. Voegeli, R. Watson and M. Sherratt, "Novel approaches to characterize age-related remodelling of the dermal-epidermal junction in 2D, 3D and in vivo," *Skin Research and Technology,* vol. 23, no. 2, pp. 131-148, 2017.

[29] S. L. Jacques, "Optical properties of biological tissues: a review," *Physics in medicine and biology,* vol. 58, no. 11, p. R37, 2013.

[30] P. Kubelka, "New contributions to the optics of intensely light-scattering materials. Part II: Nonhomogeneous layers," *JOSA,* vol. 44, no. 4, pp. 330-335, 1954.

[31] R. Yuhas, A. Goetz and J. Boardman, "Discrimination among semiarid landscape endmembers using the spectral angle mapper (sam) algorithm," in *Summaries of the Third Annual JPL Airborne Geoscience Workshop, I*, JPL Publication, 1992, p. 147–149.

[32] Kitware, "The Visual Toolkit," [Online]. Available: https://www.vtk.org/.

[33] D. Forsyth and A. Zisserman, "Mutual illumination," in *Computer Vision and Pattern Recognition, 1989. Proceedings CVPR'89., IEEE Computer Society Conference on*, IEEE, 1989, pp. 466-473.

Author Biography

Lou Gevaux received her B.S. in optical sciences from the University of Saint-Etienne (France, 2014) and an engineering degree from the Institut d'Optique Graduate School (France, 2014). She is currently pursuing a Ph.D. at the University Jean Monnet of Saint-Etienne (France), advised by Professor Mathieu Hébert. In association with Newtone Technologies (France, Lyon) and the Department of Dermatology, Hôpital Nord (Saint-Etienne, France), her research focuses on optical methods for skin analysis.