Automated characterization of skin aging using in vivo confocal microscopy

Julie Robic

UNIVERSITÉ PARIS-EST

To obtain the degree delivered by

Université Paris Est

Doctoral school n° 532 : Mathématiques et Sciences et Technologie de l'Information et de la Communication

A thesis presented by

Julie Robic

To obtain the degree of

Doctor of philosophy

Field of research "Computer Science"

June 2018

Automated quantification of the skin aging process using in-vivo confocal microscopy

Thesis director : **Michel Couprie**
Thesis Co-supervisor : **Benjamin Perret**
Thesis Co-supervisor : **Alex Nkengne**

Jury

Mme. Elsa Angelini,	Senior Scientist, Imperial College, London	Rapporteur
M. Grégoire Malandain,	Directeur de Recherche, INRIA	Rapporteur
M. Giovanni Pellacani,	Dean of the Faculty of Medicine, University of Modena and Reggio Emilia	Examinateur
M. Hugues Talbot,	Professeur, ESIEE Paris	Examinateur

Acknowledgements

I would like to start by thanking **Pr. Elsa Angelini**, **Pr. Grégoire Malandain**, **Pr Hugues Talbot** and **Pr. Giovanni Pellacani** for taking the time to read my dissertation and for being part of my defense committee. It is a great honor for me to have such esteemed and respected members of our research community to evaluate and give feedback of my work.

Je remercie **Benjamin Perret** de m'avoir guidée tout au long de cette thèse, pour nos échanges et pour ses précieux conseils et idées.

Je remercie **Michel Couprie** pour son accompagnement et sa bienveillance.

Je remercie une nouvelle fois **Hugues Talbot** de m'avoir fait découvrir cette discipline grâce à laquelle j'ai fait d'une passion, un métier. Merci de ton écoute et de ta présence depuis toutes ses années.

Je remercie **Alex Nkengne** qui a fait naître ce sujet de thèse et qui m'a permise de grandir et d'évoluer pendant ces 3 années. Plus qu'un responsable, c'est un nouveau mentor avec qui j'ai la chance de travailler.

Je remercie **Katell Vié**, Directrice Recherche et Développement Evaluation, pour sa confiance et de son engagement pour porter notre projet et nous ouvrir les portes.

Je remercie **Nathalie Issachar**, Directrice Recherche et Développement, pour son soutien de toujours.

Merci à **Richard Fitoussi** d'avoir été mon premier maitre de stage dans cette belle entreprise qu'est Clarins. Merci de cette confiance mutuelle que nous avons su créer et qui m'as permise d'avancer si loin.

Je tiens à remercier l'ensemble du laboratoire d'Informatique Gaspard Monge. Un grand merci à **Thibault**, **Diane**, **Clara**, **Elodie**, **Odyssée**, **Eloïse**, **Laurent**, **Sylvain**, **Kacper**, **Bruno** et **Laurent Najman** pour leur bonne humeur. Je n'aurais pas pu rêver meilleure équipe pour partager cette expérience !

Je remercie mes premiers collègues et amis, **Laxmee**, **Natacha** et **Patrick**. Quel chemin depuis notre bureau partagé ! Un grand merci, pour tout. Je tiens également à remercier l'ensemble de l'équipe du Laboratoire de Biologie Cellulaire pour leur soutien et générosité. Merci à **Gallic**, **Christelle**, **Eric**, **Nathalie** et **Gaëlle**.

Cette thèse m'a donnée l'opportunité de d'intégrer une nouvelle équipe, le Laboratoire d'Evaluation Clinique, avec des gens formidables. Merci à **Armelle**, **Caroline**, **Aurore**, **Clara**, **Sabine**, **Michele**, **Sonia**, **Laura**, **Annick** et **Louisa**. Merci de m'avoir accompagnée et de toujours avoir partagé cette expérience avec moi. Un dernier merci à l'Open

Space et à ses protagonistes. Sans vous, le quotidien ne serait pas si agréable, et il serait surtout beaucoup moins drôle. Soyez prêtes, maintenant, vous allez me voir tous les jours.

Je tiens à remercier ma mère **Pascale**, pour m'avoir toujours soutenue, contre vents et marées, et pour avoir accueilli avec bienveillance tous les choix que j'ai pu faire. Merci pour nos conversations si pertinentes et dévouées. Merci à ma sœur, **Emilie**, pour son soutien indéfectible et tous ses encouragements.

Merci à mon amour et compagnon, **Nicolas**, pour m'avoir toujours soutenue et encouragée. Partager avec toi cette expérience particulière, qu'est la thèse, a été une vraie force et une source d'inspiration. Depuis la Sicile jusqu'à Paris en passant par Strasbourg et El Salvador, je suis très fière et très heureuse de la belle équipe que nous formons.

Abstract

This research work deals with the application of Reflectance Confocal Microscopy to the study of in-vivo skin structures, with a view to understand and characterize markers associated with aging. However, skin aging has been studied using this image modality but so far only by experienced dermatologists. The objective of this thesis is to develop new methods to quantify automatically the phenomenon of skin aging using in vivo reflectance confocal microscopy. In our work we first address the quantification of the epidermal state. Then, we characterize the Dermal-Epidermal Junction (DEJ). Finally, we validate the proposed methods through both clinical and cosmetic product efficacy studies.

The epidermal layer appears on RCM images as a honeycomb pattern. Its regularity decreases with age. We propose an algorithm composed of two steps: 1) the image is segmented into individual cells, 2) each cell is classified as regular or irregular by machine learning based on spatial features. Then, we propose two measures to quantify the regularity of the honeycomb pattern on each image stack: 1) the percentage of regular cells, 2) the average size of the regular regions. The aggregated scores defined by the classification results show significant difference among groups of different ages and photo-exposition sites.

The DEJ is a complex, surface-like, 3D structure separating the epidermis from the dermis. We provide a method for segmenting the 3D confocal images into three regions with reduced uncertainty: Epidermis, DEJ and Dermis. The proposed approach relies on a 3D Conditional Random Field to model the skin biological properties and impose regularization constraints. Classical methods for analyzing the DEJ shape rely on the characterization of its peaks and valleys. The inclusion relation between the level-lines of the DEJ forms a tree that can be analysed to extract relevant attributes. The choice of attributes is driven by the dermatologists expertise, they have identified the loss of circularity in the dermal papillae with aging. We show that, with aging, the DEJ is composed of more irregular objects, which is consistent with the dermatologists expertise.

We carry out a clinical validation study involving 160 subjects from 4 different ethnic background. Clinical annotations are performed by experienced dermatologists. Using our proposed measurements, we are able to retrieve the statistical differences between the different annotated sets. When predicting the clinical annotations, we obtain, for the epidermis, the following scores: 80% accuracy, 81% sensibility, 81% specificity and for the DEJ 83% accuracy, 76% sensibility, 81% specificity.

We finally perform a cosmetic product efficacy study. After only 2 weeks of product application, early signs of the the product's anti-aging action are detected thanks to the proposed methods.

French long summary

Introduction

Le vieillissement cutané est défini par l'ensemble des altérations de la peau résultant de l'accumulation, au fil des années, des modifications progressives de ses différents constituants. Il est régi par des paramètres intrinsèques et extrinsèques. Les méthodes classiques de caractérisation du vieillissement cutané regroupent des techniques invasives comme la biopsie, des techniques indirectes comme l'évaporimètre, ou des évaluations cliniques, comme le score SCINEXA, qui ne permettent de détecter que des altérations avancées de la peau. Depuis une dizaine d'années, des techniques de microscopie non-invasives ont vu le jour, notamment la microscopie multiphoton et la microscopie confocale. Cette dernière offre des avantages indéniables et s'est révélée être un outil efficace dans l'exploration cutanée. La microscopie confocale à balayage laser ouvre une fenêtre sur la peau permettant de voir l'épiderme et le derme de façon non-invasive jusqu'au derme superficiel. Pour produire des images confocales, un rayon laser dans le proche infrarouge (830 nm) est utilisé. Le laser pénètre dans la peau et est réfléchi par le tissu cutané. Seule la lumière réfléchie provenant de la région focale est détectée. Le contraste sur les images repose sur la différence d'indices de réfraction des tissus qui dépendent de leurs structures chimiques et moléculaires respectives. En raison de ces variations, seule une certaine partie de la lumière est réfléchie. Les structures ayant un indice de réfraction plus élevé, telles que la mélanine et la kératine, apparaissent brillantes sur les images. Les variations de l'indice de réfraction des microstructures des tissus fournissent les images avec contraste. Les images sont de deux types :

- des images de mosaïques qui permettent à une profondeur donnée, d'explorer de grandes surfaces (jusqu'à 8 mm^2) par juxtaposition d'images prises dans le même plan focal ; et
- des piles images dites des images de « stack », qui permettent une exploration en 3 dimensions de la peau jusqu'au derme superficiel par empilement d'images prises à écarts réguliers de la surface vers la profondeur.

Les images obtenues, de manière non invasive, ont une résolution proche de l'histologie. Toutefois, l'interprétation de ces images est complexe et leur lecture nécessite l'expertise d'un dermatologue expérimenté. Les applications du microscope confocal sont nombreuses

dans le domaine du diagnostic médical notamment dans la détection et caractérisation des mélanomes. Ce domaine du confocal est plus mature que la caractérisation du vieillissement cutané et a permis l'automatisation de certaines méthodes de diagnostic. Concernant le vieillissement cutané, des descripteurs du vieillissement cutané ont été mis en évidence. Ces descripteurs ont été corrélés avec des données histologiques ainsi qu'avec le score SCINEXA décrit plus haut. Ils ont aussi été identifiés dans le cas d'un vieillissement cutané dû à l'exposition au soleil. Les descripteurs du viellissement cutané ayant été identifiés sont :

- l'altération du motif épidermique apparaissant comme un nid d'abeilles ;
- la circularité des papilles correspondant à la jonction dermo-épidermique ;
- l'altération des fibres de collagène du derme papillaire.

Un score semi-quantitatif du vieillissement a été établi, nécessitant à ce jour l'évaluation visuelle des images par des dermatologues expérimentés.

L'objectif de cette thèse est de développer et de valider une méthode de quantification automatique des descripteurs du vieillissement cutané à partir d'images de microscopie confocal in-vivo.

Bases de données

Afin de répondre à notre problématique, nous avons constitué 3 jeux de données.

1. Un premier jeu de données a été acquis dans les premiers mois de la thèse après le suivi d'une formation de 3 jours sur l'utilisation du microscope confocal. Il est composé d'images provenant de 15 personnes réparties dans deux groupes d'âges : un groupe de 7 personnes âgées de 15 à 35 ans et un autre de 8 personnes âgées de 55 à 65 ans. Ce jeu de données est utilisé pour le développement des méthodes de traitements d'images.

2. Une seconde base de données clinique a été acquise par l'équipe du Docteur Giovanni Pellacani dans le cadre d'une étude clinique. Elle comporte des images acquises auprès de 160 sujets. Cette base de données a été annotée par des dermatologues expérimentés et est utilisée pour valider les méthodes développées pendant la thèse.

3. La troisième base de données a été acquise lors d'une étude sur l'efficacité de produits cosmétiques. Elle comporte des images provenant de 17 sujets âgées de 45 à 60 ans. Quatre conditions ont été testées : une condition témoin sans application de produits, une condition avec l'application d'un produit hydratant, les deux dernières conditions correspondent à l'application de produits cosmétiques anti-âge.

Caractérisation de l'état de l'épiderme

L'épiderme est caractérisé par un motif en nid d'abeilles sur les images de microscopie confocale. Avec l'âge, ce motif perd en régularité.

Nous proposons un algorithme, pour quantifier la régularité du motif épidermique, comportant deux étapes.

1. Les cellules épidermiques sont segmentées individuellement en utilisant un algorithme de ligne de partage des eaux. La ligne de partage des eaux considère l'image ou son gradient comme un relief topographique dont on va étudier les bassins versants. L'idée consiste à placer des marqueurs (sources d'eau) dans les minima régionaux du relief topologique. L'eau entre à travers les marqueurs et inonde la surface. Lors de l'inondation, des barrières sont établies à l'endroit où les eaux provenant de deux minima différents se rencontrent. La ligne de partage des eaux conduit à une partition du domaine de l'image en bassins versants. Le principal défi de cette méthode est de contraindre l'inondation avec des marqueurs pertinents. Des a priori biologiques sur la taille des cellules épidermiques nous permettent de réduire le nombre de marqueurs initiaux et de ne conserver que ceux correspondant à l'intérieur des cellules épidermiques.

2. Chaque cellule est ensuite classée comme étant « régulière » ou « irrégulière » par un algorithme de classification intégrant des forêts d'arbres aléatoires. La racine d'un arbre de décision rassemble l'ensemble des données, tandis que les feuilles représentent les données appartenant aux mêmes classes. Chaque nœud intérieur correspond à l'un des descripteurs à partir duquel une règle de décision est créée pour diviser l'ensemble de données en sous-ensembles appartenant aux mêmes classes. Le processus continue jusqu'aux feuilles. Des descripteurs tels que la taille des cellules, leur élongation, leur nombre de voisins et leur compacité sont utilisés.

A partir des résultats de la classification, nous proposons deux mesures pour quantifier la régularité du motif épidermique : le pourcentage de cellules régulières et la taille moyenne des régions formées par les cellules régulières. La méthode proposée permet de mesurer des différences entre nos deux groupes d'âges. Le pourcentage de cellules régulières et la taille des régions régulières diminuent significativement avec l'âge.

Caractérisation de la Jonction Dermo Épidermique

La Jonction Dermo-Épidermique (JDE) est une surface 3D complexe qui sépare l'épiderme du derme. Ses pics et vallées, appelées des papilles dermiques, sont dues à des projections du derme dans l'épiderme. Avec le vieillissement de la peau, l'apparence de la JDE s'aplatit, ce qui pourrait avoir des conséquences importantes telles qu'une

adhérence épidermique plus faible. La JDE subit de multiples changements sous des conditions pathologiques ou de vieillissement. Les dermatologues ont identifié que la forme des papilles dermiques, les pics et vallées de la JDE, devient polycyclique avec l'âge. La première étape pour la caractérisation de la JDE est sa segmentation. Nous fournissons une méthode pour segmenter les images confocales 3D en trois classes : épiderme, JDE et derme. L'approche proposée s'appuie sur un champ aléatoire conditionnel 3D (CRF) pour modéliser les propriétés biologiques de la peau et imposer des contraintes de régularisation. Le modèle CRF prédit la classe d'un pixel selon 1) ses probabilités d'appartenir à l'une des couches de peau, et 2) les classes de ses voisins. La paramétrisation du modèle CRF est inspirée par des informations a priori sur la structure de la peau. L'architecture de la peau est modélisée par les probabilités de transitions d'une classe à une autre entre des pixels voisins. Les relations entre les pixels imitent l'organisation spatiale des couches de la peau en 3D : l'épiderme est au-dessus de la JDE qui se trouve au-dessus du derme. Une fois la JDE segmentée, nous pouvons nous intéresser à sa forme. Les méthodes classiques se concentrent sur la hauteur et le nombre de pics et de vallées de la JDE pour quantifier sa forme. Nous fournissons une représentation arborescente de la surface 3D modélisant les relations d'inclusions de ses lignes de niveaux. Les feuilles de l'arbre représentent les lignes de niveaux les plus hautes et les plus basses. Dans notre cas, les feuilles correspondent aux pics et vallées de notre surface 3D. Des attributs pertinents, tels que la compacité, et des profils d'attribut sont extraits de l'arbre. Le choix des attributs est inspiré de l'expertise des dermatologues, ces derniers ont montré une perte de circularité des papilles dermiques avec l'âge. Nous montrons qu'avec l'âge, la JDE est composé d'objets plus irréguliers, ce qui est cohérent avec l'expertise des dermatologues.

Caractérisation des fibres de collagène

Des résultats préliminaires sur la caractérisation des fibres de collagène sont présentés. Parmi la base de données cliniques, les fibres de collagène sont classées en deux catégories : réticulées et grossières. Nous proposons une méthode de classification par une forêt d'arbres aléatoire basée sur des classifieurs de textures afin de classer les fibres de collagène en fibres réticulées ou grossière. Les images annotées comme réticulées par les dermatologues ont un pourcentage significativement plus élevé de fibres réticulées que les images annotées comme grossières. Ce résultat est en accord avec l'analyse visuelle par des dermatologues qui évaluent l'étendue des fibres réticulées parmi les images.

Validation

Nous réalisons une étude de validation clinique à partir de la base composée de 160 sujets issus de 4 ethnies différentes. Les annotations cliniques sont effectuées par des dermatologues expérimentés. Nous montrons des différences significatives entre

les ensembles annotés en utilisant les mesures proposées. De plus, des algorithmes d'apprentissage supervisés sont entrainés sur les mesures proposées afin d'évaluer leurs capacité à prédire les annotations cliniques. Lors de la prédiction des annotations cliniques, nous obtenons, pour l'épiderme, 80% de précision, 81% de sensibilité, 81% de spécificité et pour la JDE 83% de précision, 76% de sensibilité, 81% de spécificité. À notre connaissance, ce sont les premiers résultats les résultats d'une approche automatique aux annotations produites par un dermatologue pour l'évaluation du vieillissement cutané à l'aide de la microscopie confocale in vivo. Puisque quatre ethnies sont présentes dans l'étude, nos méthodes se révèlent ethnie-indépendantes.

Nous présentons finalement nos résultats de l'étude sur l'efficacité de produits cosmétiques anti-âge après deux semaines d'application. Les mesures de la régularité du motif épidermique ne sont pas significativement améliorées. La forme de la JDE est significativement plus régulière. Quant à l'état des fibres de collagènes, il est légèrement amélioré. Nous obtenons des résultats encourageants étant donné la petite taille de notre population (17 sujets) et la courte période d'application du produit. Les méthodes proposées sont en mesure de détecter les premiers signes d'efficacité des produits anti-âges.

Conclusion

Le système de cotation développé par le dermatologue a été validé pour évaluer le vieillissement de la peau sur des images confocales. Les méthodes proposées sont non seulement capables d'évaluer les changements dus au vieillissement cutané, mais aussi d'être en accord avec les annotations cliniques de dermatologues expérimentés. De plus, nous avons prouvé que le microscope confocal a le pouvoir de suivre les changements dans le temps tout en fournissant la mesure de l'efficacité des produits, et que cette tâche peut être automatisée. Nous proposons plusieurs perspectives dans l'utilisation du microscope confocal. D'une part, un support a été spécialement conçu afin de réaliser des acquisitions sur des explants de peau ex-vivo. Ce support permettra de réaliser des validations histologiques de nos méthodes ainsi que d'explorer des structures cutanées plus profondes en retournant l'explant, tels que le derme profond ou le tissu adipeux. D'autre part, un fantôme optique de la peau a été développé. Il est composé de plusieurs couches, avec des indices de réflexion proches des couches de la peau, et contient des billes dont l'indice de réfraction est proche de celui de la mélanine. Ces billes apparaissent alors comme des points lumineux sur les images acquises offrant la possibilité de calculer la fonction d'étalement du point du microscope confocal utilisé afin de restaurer les images acquises.

List of Publications

International conferences with proceedings

Robic, Julie, A. Nkengne, Benjamin Perret, Michel Couprie, and Hugues Talbot. "Automated quantification of the epidermal aging process using in-vivo confocal microscopy." In Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on, pp. 1221-1224. IEEE, 2016.

Robic, Julie, Benjamin Perret, Alex Nkengne, Michel Couprie, and Hugues Talbot. "Classification of the dermal-epidermal junction using in-vivo confocal microscopy." In Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on, pp. 252-255. IEEE, 2017.

Abstracts presented in national conferences

Robic, Julie, A. Nkengne, Benjamin Perret, Michel Couprie, and Hugues Talbot. "Automated quantification of the skin aging process using in-vivo confocal microscopy." In Cosmetic Measurement and Testing (COMET), 2017, Oral presentation.

Robic, Julie, A. Nkengne, Benjamin Perret, Michel Couprie, and Hugues Talbot. "Automated quantification of the skin aging process using in-vivo confocal microscopy." In Société Francophone d'ingénierie et d'imagerie Cutanée (SF2iC), 2017, Poster presentation.

International journals

Mound, Abdallah, Vesela Lozanova, Céline Warnon, Maryse Hermant, Julie Robic, Christelle Guere, Katell Vie et al. "Non-senescent keratinocytes organize in plasma membrane submicrometric lipid domains enriched in sphingomyelin and involved in re-epithelialization." Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids 1862, no. 9 (2017): 958-971.

Boutrand, Laetitia-Barbollat, Amélie Thépot, Charlotte Muther, Aurélie Boher, Julie Robic, Christelle Guéré, Katell Vié, Odile Damour, and Jérôme Lamartine. "Repeated short climatic change affects the epidermal differentiation program and leads to matrix

remodeling in a human organotypic skin model." Clinical, cosmetic and investigational dermatology 10 (2017): 43.

Nkengne, A., 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 24, no. 1 (2018): 99-107.

Contents

List of Figures

List of Tables

1 Introduction

Contents

1.1 Objectives and motivation

The skin has been described as the "monumental facade of the human body" by Comel's in 1953. In fact, the skin is a key feature of our identification as individual human beings and plays a major role in social interactions. The importance of aging lies in the enormous consumer demand for agents or treatments that can prevent or reverse its effects. In light of this, it is essential to detect early skin aging signs when the process can be readily reversed or, at least, minimized.

The skin intends to protect human beings against environmental aggression. The external layer, i.e. the epidermis, acts as a protective barrier and ensures the continuous regeneration of the skin surface and the skin pigmentation. The dermis is located below the epidermis and plays an important role, as a nutritional and supportive structure involved in the maintenance of the epidermis and the skin elasticity properties.

Skin aging is defined by the set of alterations of the skin resulting from the accumulation, over the years, of diverse modifications of its various components. It is governed by both intrinsic and extrinsic parameters. Examples of the former are genetics. Examples of the latter are UV exposure or pollution [1]. The appearance of an elderly skin is

usually affected by a combination of these factors. Clinical signs of skin aging include color changes, loss of elasticity, wrinkles etc. Beside genetic aspects, environmental factors are responsible for the largest inter-individual and intra-individual variations of the aging process within the population.

In order to study and quantify the skin aging process, numerous in vivo methods and noninvasive devices have been developed.

The skin surface has been long studied. On one hand, clinical scores, such as the SCINEXA score (score for intrinsic and extrinsic skin Aging), are able to quantify macroscopic changes [2]. On the other hand, skin color and texture can be quantified on high-resolution images under standardized illumination conditions [3]. Those techniques are able to quantify clinical skin alterations, but they do not provide direct information on the skin inner structure.

A way to study the skin deeper structures in-vivo is through the study of the variations of its biophysical parameters.

The most important biophysical parameters characterizing the epidermal state are the skin pH, epidermal hydration, transepidermal water loss and sebum excretion [4]. As instance, the skin hydration can be estimated through an electrical method which measures skin impedance [5, 6].

The dermis, located below the epidermis, can be characterized through the mechanical properties of the skin which it is mainly responsible for. The skin elasticity can be measured using a device called a cutometer [7] which creates a negative pressure to deform the skin mechanically. The resistance of the skin to the negative pressure and its ability to return into its original position allow getting information about the elastic and mechanical properties of the skin. The use of ultrasounds can provide information on the skin dermal echogenicity, which decreases with age [4, 8].

These characterization techniques are indirect measurements of the skin aging alterations.

There is a need for bioengineering devices, capable of providing structural and direct information on the skin structure and its evolution with age. Over the last decade, new instrumental devices such as multiphoton microscopy [9], Optical Coherence Tomography (OCT) [10] and Reflectance Confocal Microscopy (RCM) have emerged. The latter offers undeniable advantages and has proved to be an effective tool in skin exploration, as it produces skin images at a cellular resolution in real-time.

In this scenario, In vivo reflectance confocal microscopy is the first in-vivo imaging system that provides sufficient resolution and information to enable an "optical biopsy" of the skin.

Confocal laser scanning microscopy opens a window into the skin to see the epidermis and the dermis noninvasively up to a depth of 200 μm. The representation of the cells in a layer with a thickness less than 5.0 μm can be achieved with keratin, melanin and dermal fibers working as natural contrast agents [11]. The images obtained noninvasively have a resolution of 0.5 to 1.0 μm in the lateral axis, i.e. close to that of histology [12].

Confocal microscope applications are numerous in the field of medical diagnosis, especially in the detection and characterization of melanoma [13]. The power of RCM for assessing change at the microscopic level is a great opportunity for longitudinal studies while providing the measurement of the product efficacy. Several descriptors related to skin aging have been highlighted from confocal images [14]. These descriptors have been correlated with histological data and the SCINEXA score, described above. A semiquantitative score of aging has been established, requiring visual assessment of the images by experienced dermatologists [15]. The descriptors have also been identified in the case of skin aging with sun-exposure [15, 16]. The following descriptors have been identified for the study of skin aging:

- the thickness of the epidermis;
- the epidermal state characterized by a regular or irregular honeycomb pattern;
- the shape of the interface between the epidermis and the dermis, i.e. the Dermal-Epidermal Junction (DEJ);
- the pigmentation; and
- the alteration of the collagen fibers in the dermis.

As described above, RCM has the capacity to produce in vivo images at a cellular resolution, making morphological information available. However, intensive operator training and experience are necessary for a successful interpretation of the confocal images. To reduce the need for training, computer aided diagnostic systems and image processing procedures have been investigated over the past few years. Because RCM reports endogenous morphology, one of the first applications for image processing is the identification and classification of differential diagnostic criteria. Several approaches to automatize confocal image analysis have been proposed focusing on quantifying the epidermal state [17, 18], performing computer-aided diagnostic of skin lesions, [19, 20], and identifying the layers of human skin [21–23]. With further advances in automated diagnostic aids, RCM is expected to become easier to use and therefore to become an ideal tool for rapid diagnosis and skin physiology understanding.

The automatic detection and characterization of the skin aging descriptors has not been investigated yet. It could represent the emergence of a new objective method, which would be able to characterize the skin aging process at a cellular resolution.

The objective of this thesis is the development of a new technology to automatically quantify the phenomenon of skin aging using in vivo reflectance confocal microscopy.

1.2 Thesis contribution

In this manuscript, we propose new computational methods focusing on the quantification of the skin aging process using in-vivo confocal microscopy. Moreover, we report on several studies proving the robustness and usefulness of our proposed methods, from a small study dedicated to the development of image analysis methods to a clinical validation study to confront methods to the dermatologists expertise. The overall contributions of this thesis are summarized below.

- The quantification of the epidermal state is addressed through the quantification of its honeycomb pattern regularity. A segmentation procedure is proposed using a watershed-based approach, followed by a classification of the epidermal cells as regular or irregular. Two measurements are obtained from the classification results. They enable the quantification of both chronological and photo-aging.

- We propose a method to segment the dermal-epidermal junction (DEJ) using 3D Conditional Random Fields. Our contribution lies on exploiting the additional depth and 3D information of the skin architecture to provide spatial regularization over label distribution and to model skin biological properties. We take advantage of a biological information on the skin structure, i.e. its layers organization, to provide a robust segmentation method. It is an essential step in the automation of the skin aging characterization on RCM images.

- The DEJ shape is described through its representation as a level line tree. Two analysis strategies are proposed to characterize its shape: 1) the development of local and topological tree descriptors, 2) the study of the response profile of the DEJ surface evolution when undergoing a filtering process. Our proposed methods are able to discriminate the two age groups. They also agree with the dermatologists' analyses.

- Three studies are conducted. The three databases cover a large range of intra and inter-individual variability. Our proposed methods, which have been developed on a small database (15 subjects), are evaluated on a large database (160 subjects from four different ethnic background) annotated by experienced dermatologists. The proposed methods performance enable us to evaluate their robustness when dealing with intra and inter-individual variability. Furthermore, we perform a cosmetic product efficacy study with a two-weeks application period. Our proposed methods are able to measure the product efficacy and to further understand the different pathways of action of active ingredients in vivo.

- We propose several perspectives in the use of the confocal microscope. We have built an ex-vivo device in order to perform ex-vivo validation of our methods. Moreover, we show that this device can image deeper skin layers simply by imaging them from the other side. We discuss the development of a skin phantom, which possess the same optical properties as the skin layers. The skin phantom is filled with spherical beads which can help the measurement of the Point Spread Function of the confocal microscope.

The work carried out in this thesis contributes to improve the monitoring of the skin aging process under a variety of conditions. This novel approaches should help convert the use of the RCM from a purely qualitative to a quantitative modality.

1.3 Structure of the thesis

The thesis is structured into seven chapters.

- **Chapter 2**: is dedicated to the contextual and biological background of this thesis. We first define the RCM technology and the use of the confocal microscope. Then, we detail the skin cellular organization in order to present and understand the skin pattern on RCM images. As the RCM images are acquired at a cellular resolution, they are a direct visualization of the skin structure.

- **Chapter 3**: we first present the foundations of our problematic, i.e. the skin aging patterns established by dermatologists. We bring together the cellular modifications that the skin undergoes with aging and the resulting changes of the skin patterns in RCM images. Then, we highlight our first contribution, i.e. a database generation. We aim to reduce the inter and intra-individual variability by proposing a reliable and accurate data collection protocol.

  From now on, the following chapters represent our main contributions.

- **Chapter 4**: we present a method to quantify the epidermal state. The epidermal state is characterized by a regular or irregular honeycomb pattern on RCM images. We propose a method that reproduces the visual expertise of dermatologists. The honeycomb pattern is segmented using a watershed-based method and a regularity measure is proposed from the segmentation results. The results presented in this chapter are obtained from a small database used for algorithm development.

- **Chapter 5**: we present a method to segment the skin layers using 3D Conditional Random Fields (CRF).

- **Chapter 6**: we propose a tree-based method to measure the DEJ shape irregularity in 3D, which has been identified as a factor of aging by dermatologists.

- **Chapter 7**: previous results are obtained on a small database with an age criterion. In this chapter, we present a clinical validation of our methods on a database including 160 subjects. We correlate our results with ground-truth annotations provided by an experienced dermatologist. We also present the results of a cosmetic product efficacy study.

- **Conclusion**: includes the conclusion of our work and several perspectives on an ex-vivo validation protocol and a skin phantom development for image restoration.

2

Normal Skin Imaging using RCM

Contents

This chapter is dedicated to the biological and contextual background of this thesis. We first review the emergence and evolving use of the confocal microscope and detail its technical principle and practical aspects. Then, we describe the skin structure in order to understand the skin patterns on RCM images.

2.1 The confocal story

The need for in vivo non-invasive skin imaging techniques was first expressed in the dermatological field. Because melanoma in advanced stages is still difficult to cure, early detection is essential for reducing mortality. Several diagnostic techniques have emerged in the last decades such as dermoscopy, OCT and RCM.

All of these techniques provide additional information to the clinician while preserving the integrity of the skin tissue. Among these techniques, Reflectance Confocal Microscopy

is the only one that enables the en-face (horizontal) imaging of the skin at a cellular resolution (0.5 to 1.0 μm in the lateral axis and 4-5 μm in the axial axis).

The general principles of RCM were described by Marvin Minksy in 1957, but it took several decades until the technology improved for its application to the imaging of human skin. Reflectance model relies on the variation in refractive indices of skin components. In 1967, a tandem scanning reflected light microscope was developed by the group Petran and Hadravsky [24], using a mercury lamp as a light source. It was first used for imaging of excised tissues and on organs of living animals in the late 80's such as rats kidney [25] and cats eyes [26]. Features of human skin, using RCM, were first described in the early 90's [27, 28]. In 1995, Rajadhyaksha and al. imaged the human skin with a laser beam at multiple wavelengths in the visible and near-infrared band (400-900nm) [29]. As the resolution is inversely related to wavelenght of illumination source, commercial confocal microscope nowadays use laser beams in the near-infrared band.

The initial field of research involving the confocal microscope is the diagnosis of melanocytic lesions [30, 31], i.e, pigmented lesions that contains a naevus. As RCM provides images at a cellular resolution, the morphology of naevi and melanoma was successfully described. The diagnosis of melanoma using RCM was shown to be widely facilitated and has been compared to histological sections [13, 32]. Pellacani et al. [33] have shown that the use of RCM can improve the diagnosis of pathological lesions, while reducing the number of unnecessary excisions. Other fields of confocal research include non-melanocytic lesions [34], corneal disease [35] and inflammatory skin disease [36].

As the use of the confocal microscope increases, so does the understanding of the skin structure and physiology. Therefore, others applications of the confocal microscope emerge, such as the study of normal skin [37–39]. The confocal findings in skin aging will be further described in Chapter 3.

2.1.1 Technical principles of RCM

The imaging system of the confocal microscope is presented in Fig. 2.1. The microscope uses a laser as a monochromatic and coherent light source. The light penetrates the skin and illuminates the tissues. The reflected light is then gathered through a small pinhole and forms an image on the detector. In other words, the pinhole does not allow the reflected light (reflectance) from another tissue point to reach the detector. Only reflected light from the focal region (confocal) is detected. Light from out-of-focus planes is rejected by the pinhole.

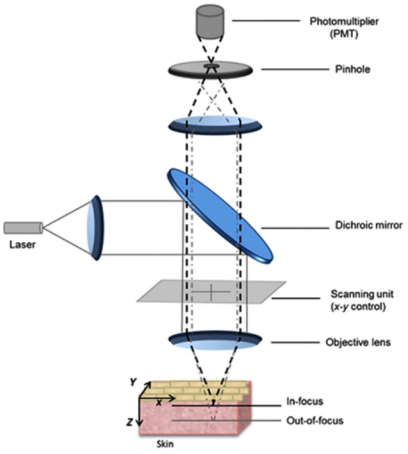


Figure 2.1: Schematic illustration of a reflectance confocal microscope [40].

The contrast on RCM images relies on the difference of refractivity of the tissue which depends on chemical and molecular structures. Due to these variations, only a certain portion of light is reflected. Structures with a higher refractive index, such as melanin and keratin, appears bright on RCM images.

Since the skin is only weakly optically transparent, higher laser power is needed to image deeper skin sections.

The commercial device Vivascope 1500, used for the image acquisition in this thesis, uses a laser with a near infrared wavelength of 830 nanometers. At that wavelength, laser light is non-invasive, as it causes no damage to tissues. However, imaging depth is limited to 200-300 μm. Depending on the skin site, RCM is able to image the skin down to the upper dermis. Higher laser power could provide deeper imaging with higher contrast but would be hazardous for the skin or the eye.

2.1.2 Practical aspects

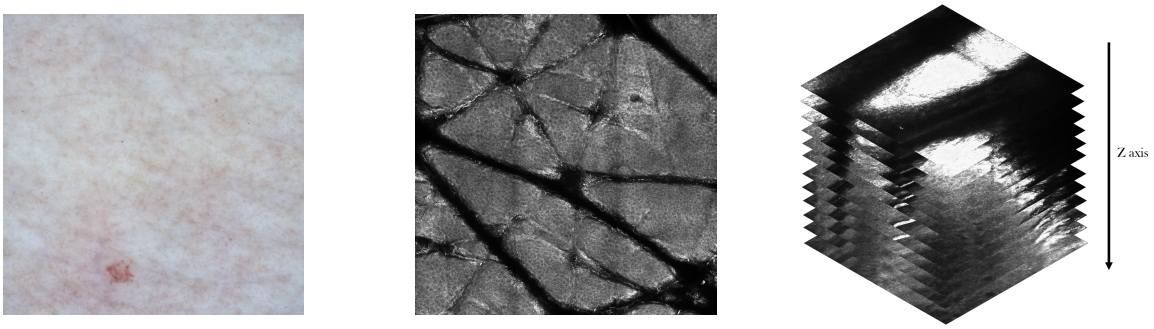
The examination of a skin site with RCM is an easy, fast and painless procedure compared to other examination techniques, such as biopsies.

A polymer window is attached to a metal ring fixed to the skin with a special adhesive to reduce skin movement. Before the ring is attached, a small amount of immersion oil is applied to optically couple the window to the skin. A macroscopic image is taken using a dermoscopic camera. This image is used as an interactive map to move the laser beam and record the acquisitions coordinates. Afterwards, the objective lens is magnetically connected to the ring. A water-based immersion gel, with a refractive index close that of the epidermis (~ 1.34) is put between the ring and the objective lens to reduce the spherical aberration due to the beam passing through air.

As mentioned above, RCM images are en-face sections of the skin. Each image is composed of 1000×1000 pixels, displaying a 500×500 μm field of vue. RCM provides two types of images modes:

Block: This is a two-dimensional composite mosaic to image a large area of 8×8 μm with a grid of continuous horizontal images.

Z-stack: This is an automated vertical sequence of images, each 500×500 μm field of view, providing a three-dimensional view of a selected area. The step between each section can be adjusted. The imaging modalities are presented in Fig. 2.2.



(a) (b) (c)

Figure 2.2: On the left, the macroscopic image (a) serves as an interactive map to choose the acquisition zone. The mosaic image (b) is used to explore large areas at a chosen depth. A z-stack (c) is used for a three-dimensional view of a selected area.

In the next section, the skin structure is described and its appearance on confocal images is discussed. As the confocal images have a resolution near that of histology, the comprehension of the skin cellular organization helps with understanding confocal images patterns.

2.2 Normal Skin Imaging using RCM

The human skin is the largest organ of the body. This tegument, about 2mm thick, has an area of about 2m^2, weights about 10kg, and handles many biological functions due to its interface with the environment. It provides protection against external aggression, thermoregulation, has an immune function (Langerhans cell), a metabolic function (vitamin D synthesis) and a sensory function (touch, pain, heat). Due to its interface with the environment, the skin is a constantly changing organ. It has a remarkable ability to adapt to its surrounding environment. Its main roles are protection of the body against injuries and prevention of bodily fluids loss. At the same time, the skin is responsible for the uninterrupted communication with the environment to provide an adapted response.

In this section, we will review the skin structure and its visualization using RCM. As mentioned in the previous section, contrast on RCM images relies on the difference

of refractivity indexes of the tissue. The objective here is to describe the skin cellular organization in order to understand skin patterns in RCM images.

The skin envelope is composed of three major layers. From the outer to the ineer layer: the epidermis, the dermis and the hypodermis, see Fig. 2.3. The interfaces between layers also have specific properties.

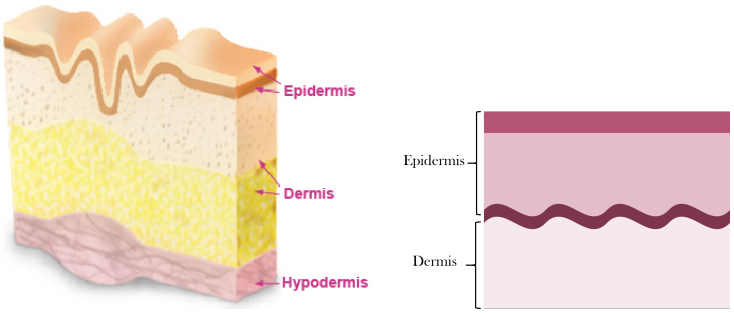


Figure 2.3: Schematic views of the skin layers organization. The right schema will be used to reference each layer in the descriptions below.

2.2.1 The epidermis

The epidermis is a stratified epithelium. It contains different sub-layers of epidermal cells, called keratinocytes, each at a different state of differentiation. The epidermis is the outer layer of the skin, which serves as a physical and chemical barrier against the environment. Keratinocytes are produced in the lower layer of the epidermis, the basal layer, and migrate upward to the outer layers. The cells mature throughout this process. They become larger and ultimately end up as a compact anucleated barrier, i.e. the stratum corneum (SC). The outer cells regularly come off the skin surface, this process is known as cell desquamation. This process is compensated by the renewal of the epidermis, a process undertaken by the keratinocytes, also responsible for maintaining skin homeostasis. The basal layer also contains melanocytes, the cells responsible for melanin production in the skin. Melanin absorbs radiant energy from the sun, allowing the epidermis to protect the skin and the organs underneath from ultraviolet radiations.

The epidermis can be sub-divided into several layers.

The stratum corneum

The Stratum Corneum (SC) is the top layer of the epidermis. In the rest of the thesis we will associate the SC to the skin surface. The SC is about 15-20 μm thick. It is the final state of the maturation of keratinocytes, in which the keratinocytes are called corneocytes. The corneocytes are large, polygonal and anucleated cells. The keratin composing them is then condensed. As keratin is a strong contrast agent on RCM images, the cells in the

skin surface show a very bright pattern. The skin surface is the brightest image compared to other epidermal layers not only due to this keratinocyte migration but also because of the back-scattered light related to the difference in the refractive indexes at the interface between the microscope immersion medium and the skin surface.

The skin folds are visible on the skin surface and appear as non-refractile or dark furrows between groups of corneocytes forming a network-like structure on the skin surface. The corneocytes appear as bright polygonal structures with dark outlines on RCM images, see Fig. 2.4.

Other skin structures appear on the skin surface, such as hair and pores. Hair shaft are composed of keratin, they appear as cylindrical or tubular structures with a uniform high refractivity and no cellular components. Pores and hair follicules appear as oval to round structures, see Fig. 2.5.

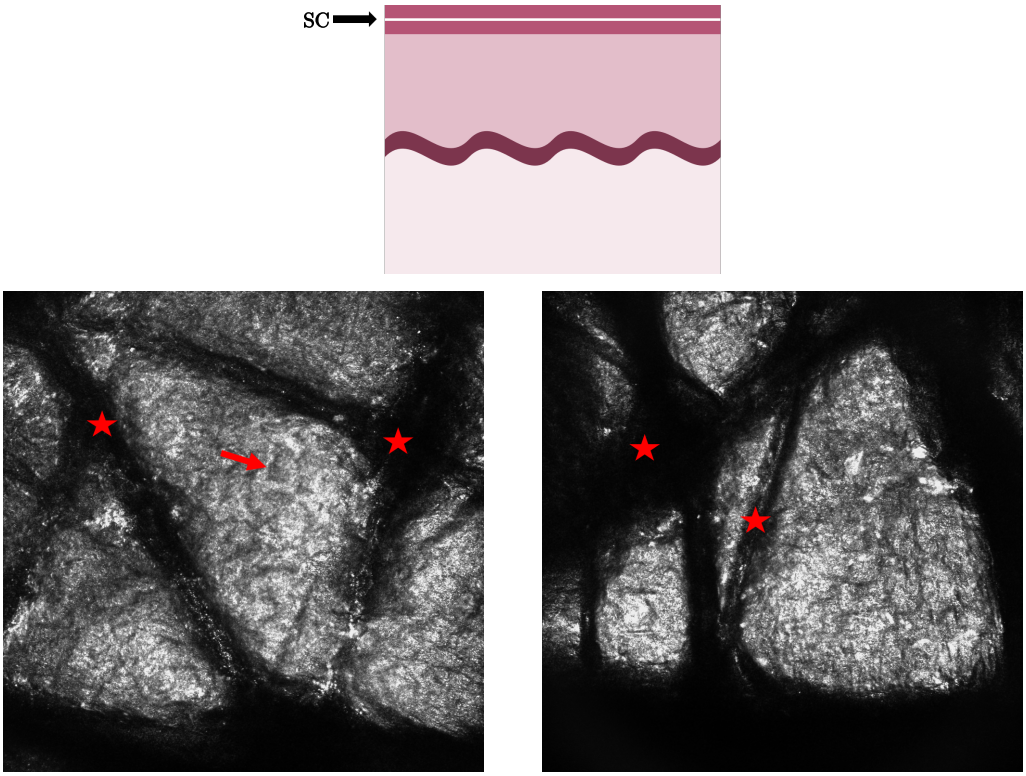


Figure 2.4: RCM images of the stratum corneum in normal skin. Large anucleated corneocytes (red arrow) and skin folds (red stars).

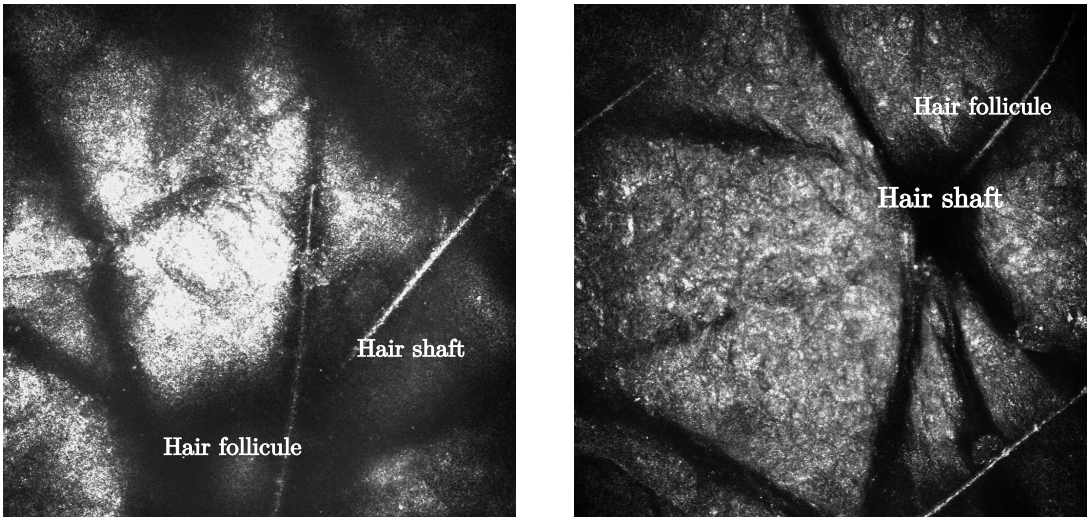


Figure 2.5: Examples of artefacts. Hair shaft and hair follicules are visible on the skin surface.

The stratum granulosum

The stratum granulosum is located under the SC. It is the first epidermal layer containing viable cells: the granular cells. These approach their final state of maturation. Therefore, they are large polygonal keratinocytes with a diameter of 25-35 μm. They contain a nucleus and other granular components which appear highy contrasted in RCM images.

Due to the amount of keratin the cells contain, their membranes show a bright pattern.

The keratinocytes have well-demarcated outlines and a grainy interior. They form a honeycomb pattern, which is the main pattern of the epidermal layers. Some RCM images of the stratum granulosum are presented in Fig. 2.6.

The stratum spinosum

The deeper we observe in the epidermial layers, the smaller the keratinocyces appear. A spinous keratinocyte has a diameter of 15-25 μm. The spinous layer extend from 20 to 100 μm below the skin surface. The RCM pattern of spinous cells is the same as the granular ones, i.e. a honeycomb pattern. The only difference is that the honeycomb pattern is tighter, see Fig. 2.7.

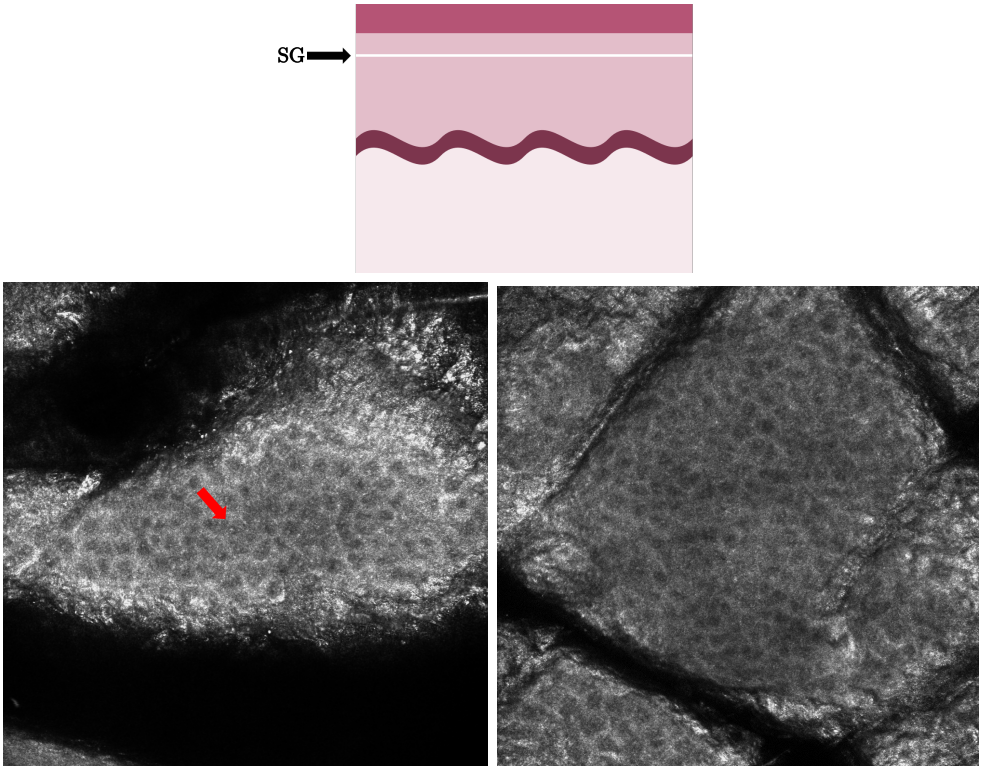


Figure 2.6: RCM images of the stratum granulosum in normal skin. Honeycomb pattern with grainy structures (red arrow).

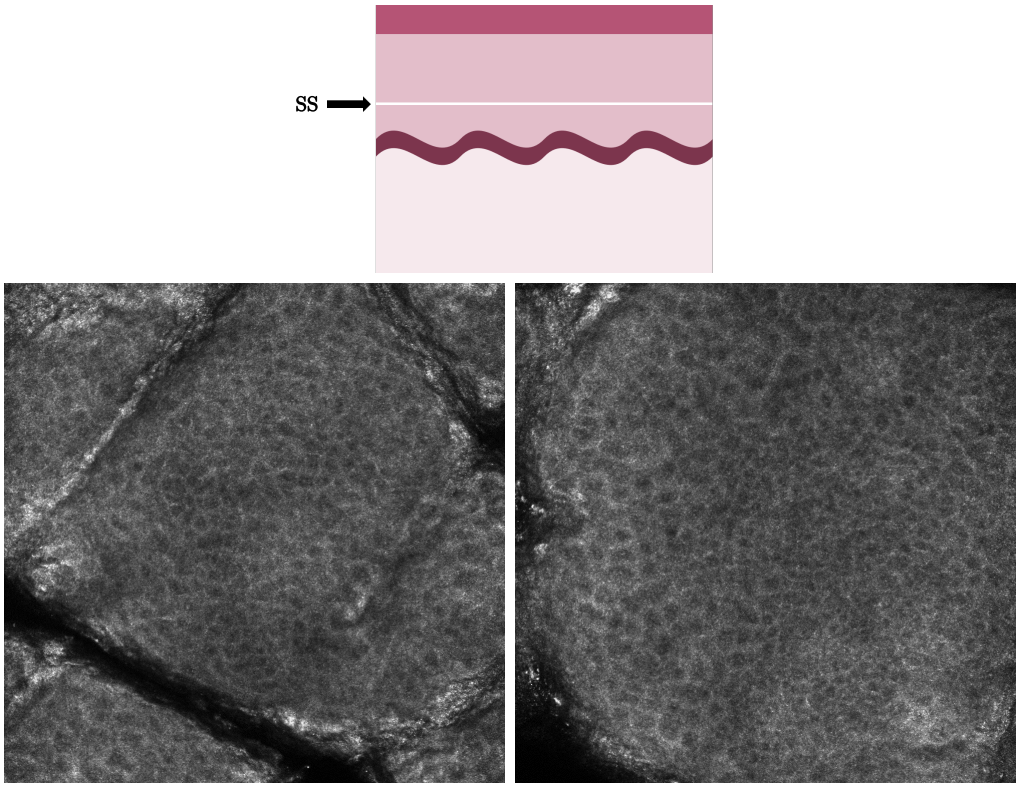


Figure 2.7: RCM images of the stratum spinosum in normal skin. Tightened honeycomb pattern.

The stratum basalis

Below the spinous layer, between 50-100 μm under the surface, there is a single layer of cells, the basal cells located at the dermal epidermal junction. Basal cells have a diameter about 7-12 μm and contain melanin caps, which are strong contrast agents.

The visibility of basal cells on RCM images is highly dependent of the skin photo-type. In dark skin, with a high content in melanin, basal keratinocytes are arranged in clusters forming a cobblestone pattern on RCM images. Conversely, in fairer skins, in which the basal keratinocytes have a lower refractivity index, they can be difficult to distinguish. Examples of the basal keratinocytes RCM patterns are presented in Fig. 2.8.

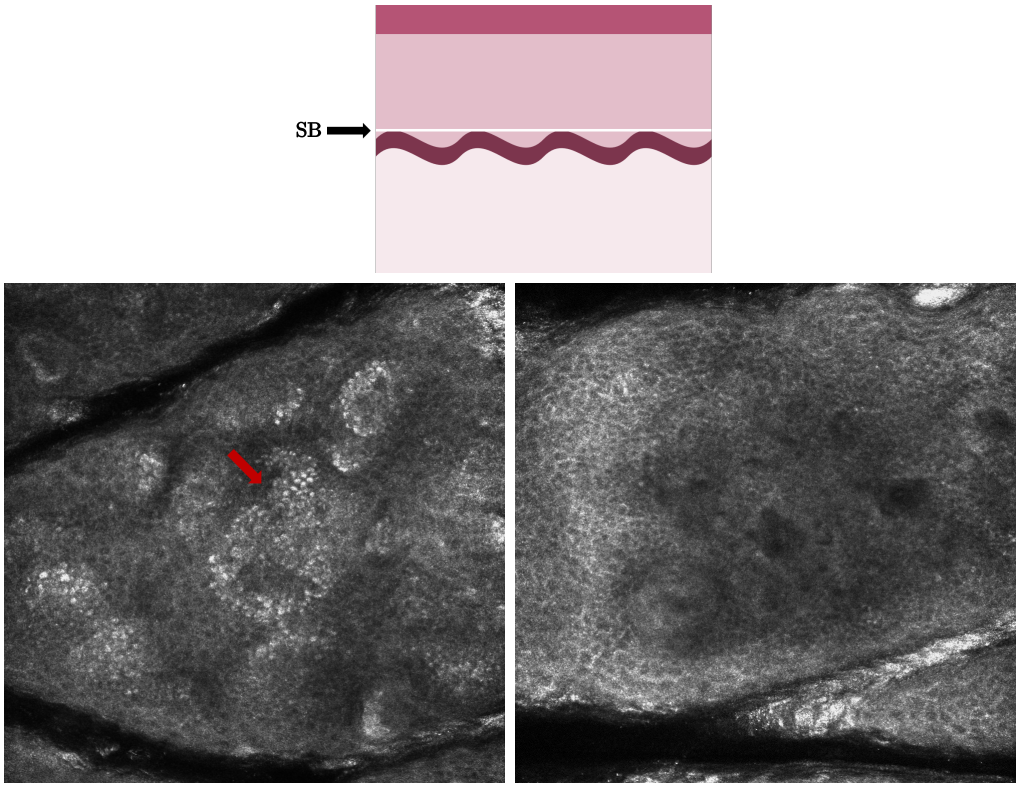


Figure 2.8: RCM images of the stratum basalis in normal skin. A cobblestone pattern (red arrow) is visible on the left image, corresponding to high refractive keratinocytes. On the right, the basal keratinocytes have a lower refractivity index and are hardly visible.

2.2.2 The dermal-epidermal junction

The dermal-epidermal junction (DEJ) is a complex, wave-like, 3D structure separating the epidermis from the dermis. Its peaks and valleys, called dermal papillae, are caused by the projection of the dermis into the epidermis. The DEJ serves as a nutrient and signaling molecules carrier between the epidermal and the dermal layers. It also has a role in the adhesion of the epidermis.

As the basal layer may show different refractivity indexes, the DEJ can have multiple confocal patterns, especially in fair skin. It can appear as an amorphous and low-contrasted structure or as circular rings. Confocal sections at the DEJ showing dark oval areas surrounded by bright basal keratinocytes correspond to a "dermal papillae". Collagen fibers in the dermis may be observed inside the dermal papillae. The low-contrasted pattern is due to a lower refractive index of basal cells, making the precise DEJ localization difficult, see Fig. 2.9.

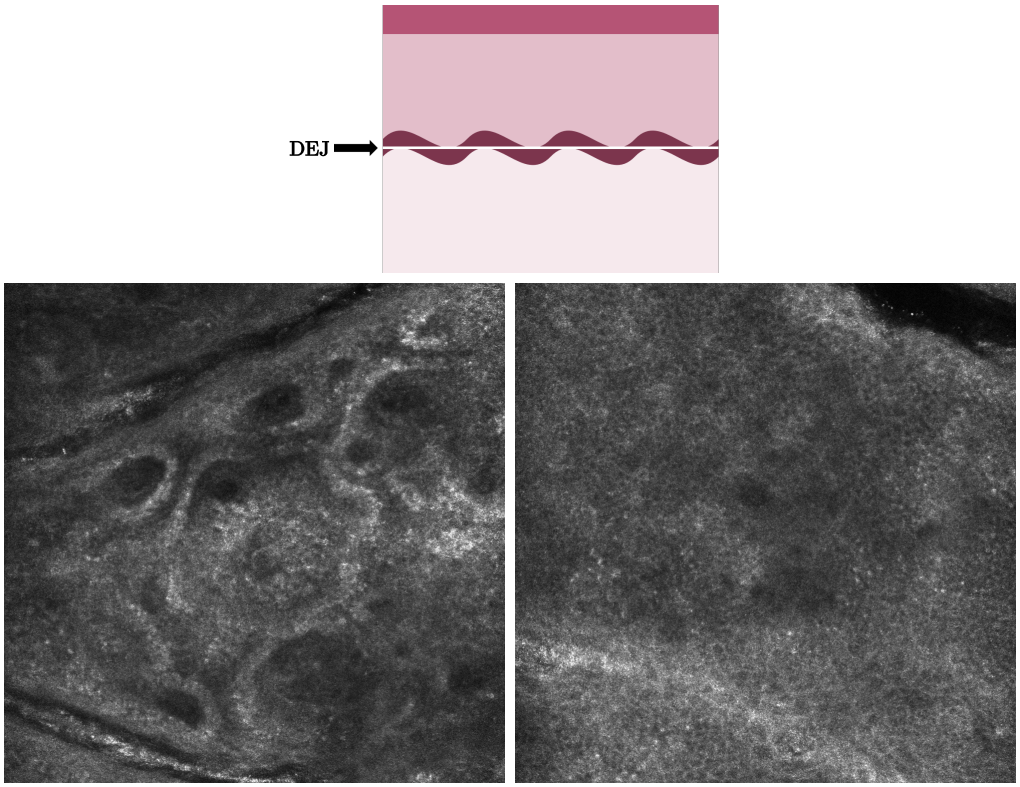


Figure 2.9: RCM images of the DEJ in normal skin. Dermal papillae are visible on the left image, while a blurry pattern is present on the right image.

2.2.3 The dermis

The dermis is a dense, fibrous and elastic connective tissue. It is divided into two areas:

- The papillary dermis: it is located just below the epidermis with which it performs nutritious exchanges. It contains collagen fibers oriented perpendicularly to the epidermis and the elastic fibers.

- The reticular dermis: it corresponds to most of the dermis and contains large collagen fibers arranged in thick network. It is in contact with the subcutaneous tissue.

The tensile strength and resistance of the skin to other mechanical forces is mainly due to the collagen fibers located in the dermis.

Collagen and elastin fibers appears as bright elongated fibrillar structures with no cellular component. They can be arranged in a reticular network or as bright bundles. Collagen fibers viewed by RCM are presented in Fig. 2.10.

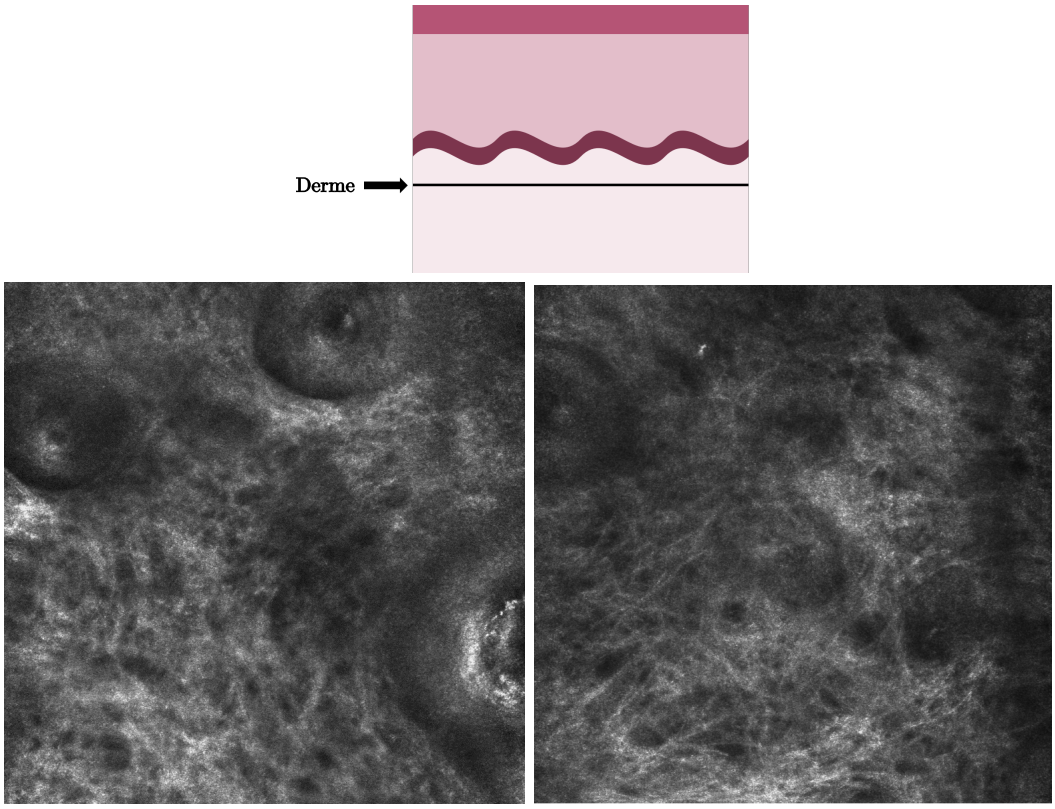


Figure 2.10: RCM images of the dermis in normal skin. Network of collagen fibers.

2.2.4 The subcutaneous tissue

The subcutaneous tissue is a layer made of fat and connective tissues containing larger blood vessels and nerves. This layer is important for the regulation of the skin and the body temperature. The hypodermis stores lipids into adopicytes cells, making it the largest energetic tank of the human body. As mentionned in Sec. 2.1, the confocal microscope maximal imaging depth is 200 μm. Therefore, the subcutaneous tissue can not be imaged using the in vivo confocal microscopy used in this thesis.

2.3 Summary

In the next page, we present an overview of the skin structure using RCM through Fig. 2.11 to Fig. 2.16. In the next chapter, we will review the skin aging pattern on RCM images.

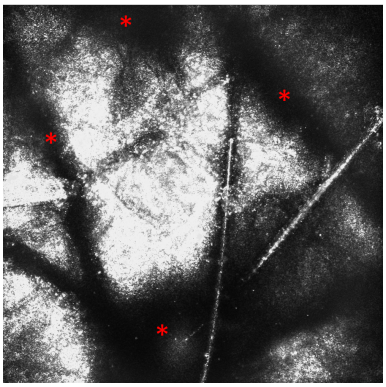


Figure 2.11: The SC is top and brightest image compared to other epidermal layers. The skin folds and hair shaft are visible on the skin surface (*).

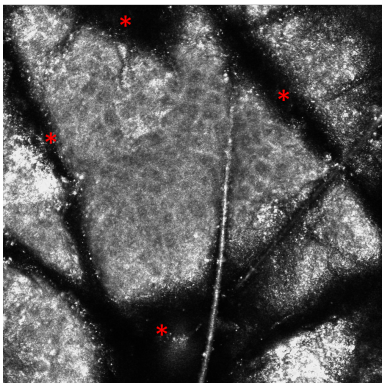


Figure 2.12: The stratum granulosum is located below the SC. The keratinocytes are outlined, forming a honeycomb pattern. The skin surface artifacts are still visible.

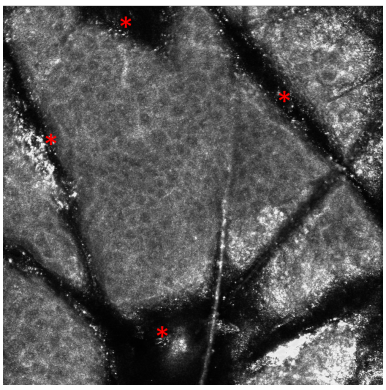


Figure 2.13: As we go deeper, the keratinocytes become smaller. The honeycomb pattern tightens.

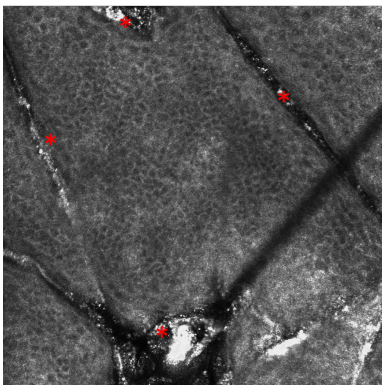


Figure 2.14: The stratum spinosum is composed of the smaller keratinocytes. Notice the image distortion located below the skin surface artifacts.

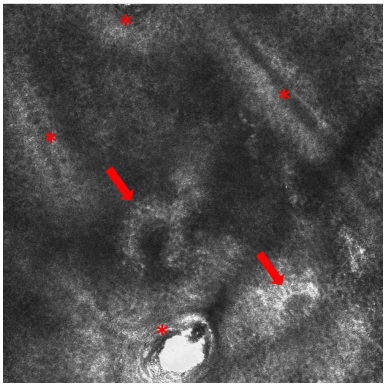


Figure 2.15: The dermal-epidermal junction is characterized by the loss of the honeycomb pattern replaced by a blurry pattern. Dermal papillae may appear (red arrows).

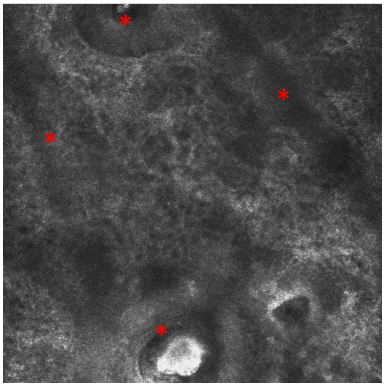


Figure 2.16: The dermis is composed of a variable amount of refractive collagen fibers.

3

Effects of aging as observed with RCM

Contents

In this chapter, we first describe the confocal aging features identified by dermatologists. Then, we detail the generation of three databases that were used to develop and validate the proposed methods.

3.1 Epidermal layers

As previously described, the epidermis is composed of several sub-layers. We present the aging pattern from the stratum corneum and the deeper epidermal layers.

3.1.1 Stratum corneum

With aging, the cutaneous lines, composing the stratum corneum, show a more linear appearance [41, 42]. The "furrow aspect" of these lines has been described on RCM

images as the dark folds between groups of keratinocytes. They were described for their orientation and intersections [14].

Toward skin aging, they can present a "rhomboidal" or "linear" pattern. In [15], the authors have shown that the the rhomboidal pattern of the skin furrows were inversely correlated with wrinkles grading with the SCINEXA score.

In this thesis, we will not focus on the characterization of the furrow pattern as confocal imaging does not bring additional information compared to other existing methods [42–44].

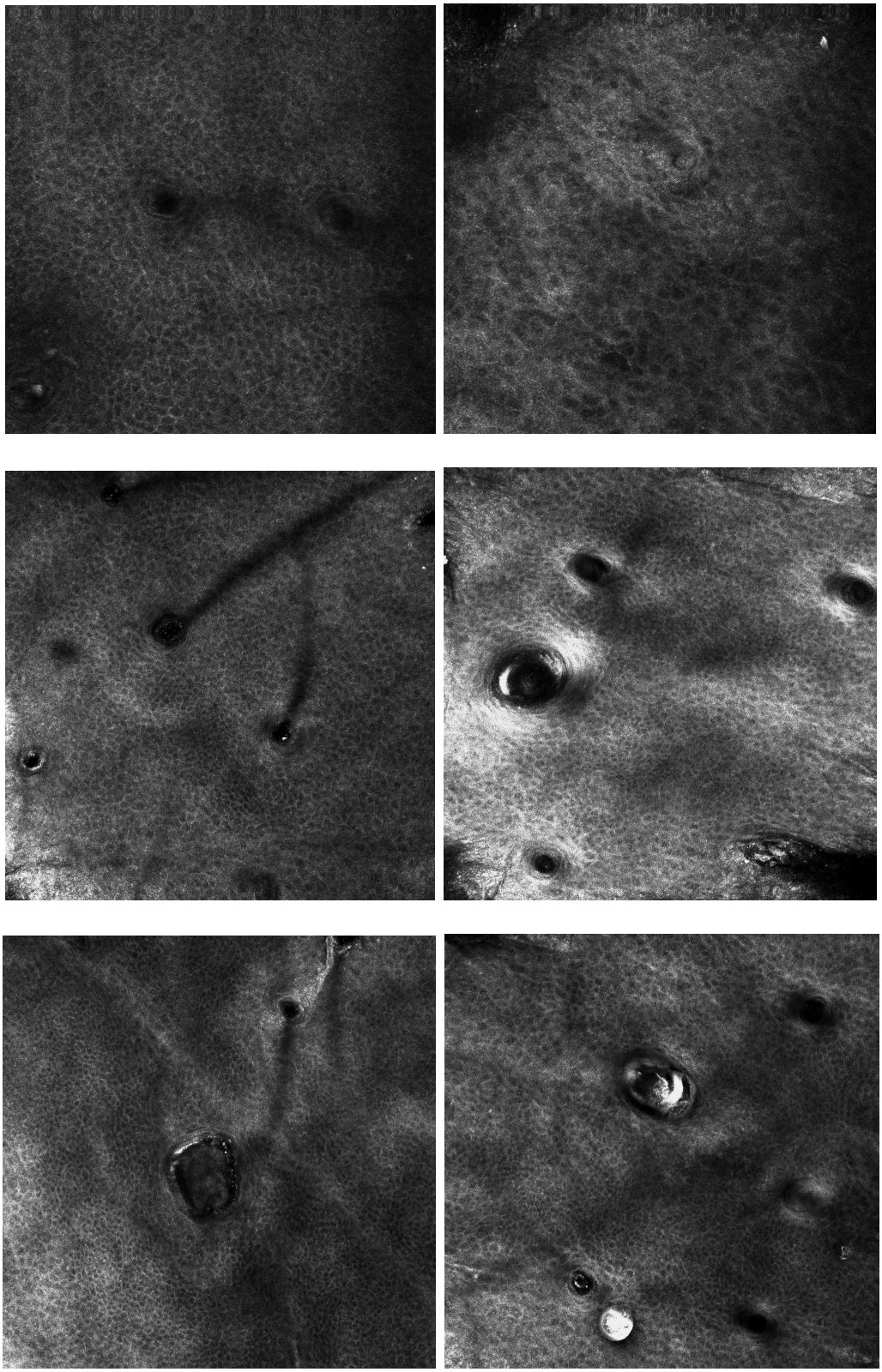
3.1.2 Deeper epidermal layers

During aging, the renewal process undertaken by the keratinocytes is altered [45]. This phenomenom compromises its protective role, but also contributes to the appearance of skin disorders, including excessive dryness, as well as increased pre-disposition to the formation of wrinkles [46].

Keratinocytes diameter has been studied using RCM [47], as it provides cellular information on the skin components. It has been shown that keratinocyte diameter is only marginally influenced by age. Epidermal thickness has also been studied [14, 38, 39] using RCM, showing a thinning of the epidermis over the years.

In [14, 15], the authors described the regularity of the epidermal honeycomb pattern as a sign of skin aging. With skin aging, the keratinocytes show increased variability in size and shape. Cell borders are not always well outlined and preserved, leading to a distortion of the honeycomb pattern which progressively loses its regularity.
This confocal descriptor showed an increasing trend among different age groups, which correlates with the clinical grading score SCINEXA [2]. The authors performed histopathologic examinations, showing a correlation between the irregularity of the honeycomb pattern and variable degree of dyskeratosis (an abnormal keratinization process of the epidermis).

The honeycomb pattern is described as regular or irregular by dermatologists. Examples of regular and irregular honeycomb patterns are presented in Fig. 3.1.



(a) Regular honeycomb pattern (b) Irregular honeycomb pattern

Figure 3.1: Epidermal aging. On the left column, the epidermis exhibits a regular honeycomb pattern compared to the images on the right column. One can notice that the images are affected by several defects: noise, blur, non-homogeneous intensity.

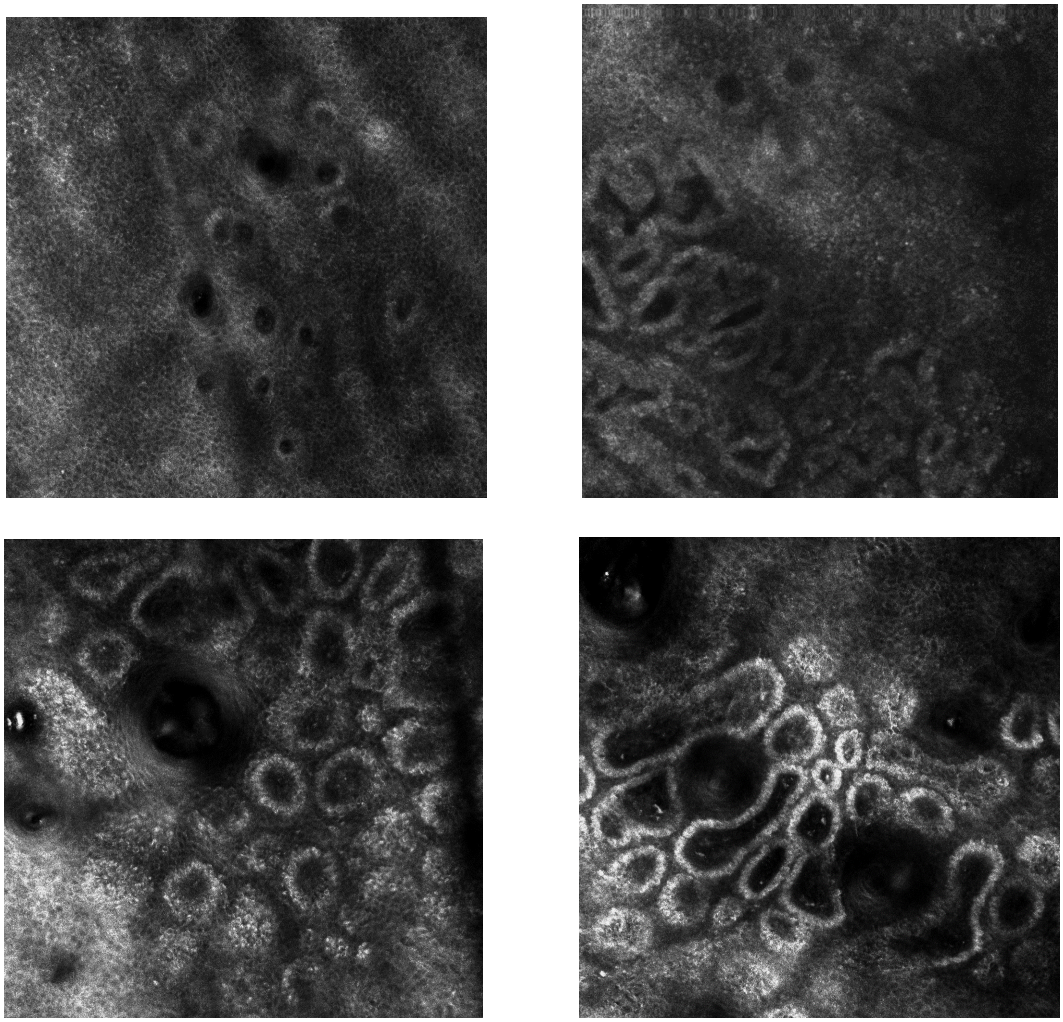
3.2 Dermal-epidermal junction

Alteration of the epidermal and dermal layers within skin aging induce a flattening of the dermal-epidermal junction [48, 49], reducing the surface for exchange of water and nutrient from the dermis to the epidermis.

The first confocal descriptor identified on the DEJ aging was the number of dermal papillae per area. In [47, 50], the authors showed a dramatic reduction in the number of dermal papillae with aging, particularly in subjects with fair skin, which is linked to the flattening of the DEJ.

Apart from the number of dermal papillae, the shape of the dermal papillae had been previously studied in [51] for the characterization of lentigo. In [14, 15], the authors have extended this confocal descriptor to the study of skin aging. The shape of the dermal papillae has been described as ringed or poly-cyclic. Poly-cyclic dermal papillae are present in over 50% of subjects between 45 and 65 years old, whereas they are virtually non-existent among young subjects (15-30 years old). The shape of the dermal papillae is now the descriptor of reference when studying the DEJ aging [52].

As the visibility of dermal papillae is highly dependent on the skin phototype and acquisition zone, most confocal images cannot be analyzed for the DEJ aging pattern. Most of the dermal-epidermal junctions on RCM images appear as low-contrasted and blurry. Examples of the DEJ patterns are presented in Fig. 3.2.



(a) Ringed dermal papillae (b) Poly-cyclic dermal papillae

Figure 3.2: DEJ aging: ringed dermal papillae can be observed on the left column. Their shape becomes poly-cyclic with age (right column). The second line of images correspond to a darker skin, with higher melanin content, making the DEJ pattern more visible.

3.3 Dermal layers

The dermis suffers the most dramatic changes with age. The collagen production progressively decreases over the years while the remaining collagen fibers undergo an alteration of their structures and organization [53]. Progressive fragmentation of the dermal collagenous matrix has an important consequence: it decreases overall skin strength, favoring wrinkle formation.

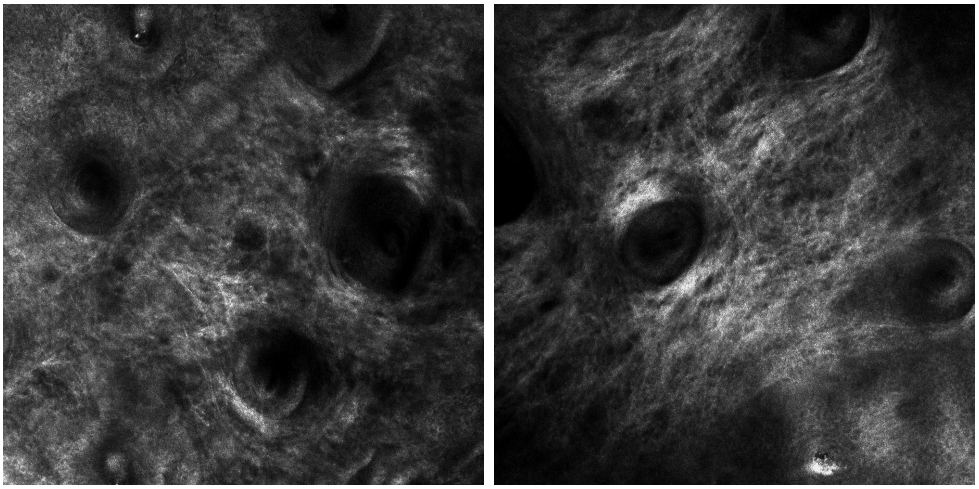
On RCM images, four main types of collagen fibers organisation have been described [14, 15], each one corresponding to a state of alteration.

- Thin reticulated collagen: bright thin fibrillar structures forming a delicate web-like pattern.
- Coarse collagen structure: coarse filamentous structures with a tendency to concentrate. a web-like pattern is still observable but with larger and irregularly spaced meshes.
- Huddled collagen: large hyporefractive blotches of amorphous material. individual collagen fibers are no longer visible.
- Curled bright collagen: highly refractive thick and short wavy fibers, when severe sun damage is present.

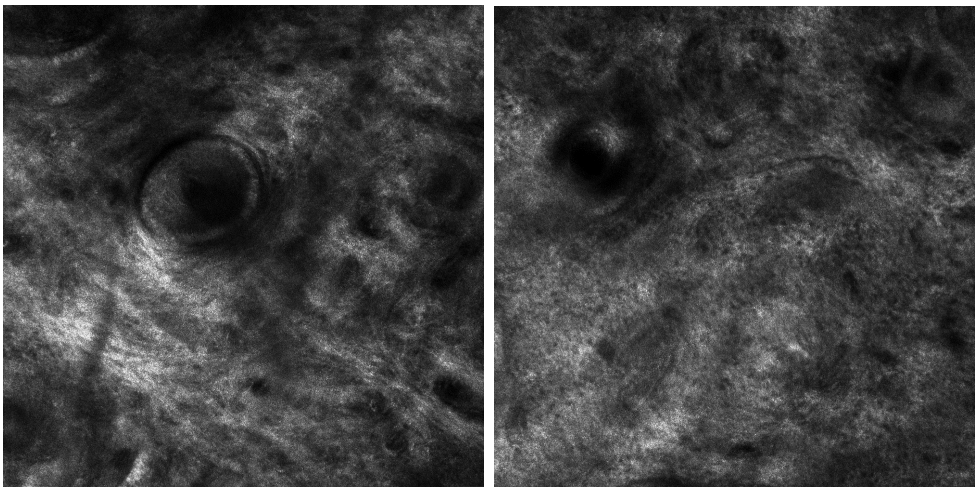
Thin reticulated collagen fibers characterize young subjects. With age, they are replaced by huddled and coarse collagen fibers. Curled fibers are present in subjects over 65 years old.

The confocal descriptor associated with dermal aging is the classification of the dermal architecture into the four types of collagen fibers. Distinct collagen aspects have been correlated to the SCINEXA score and histological data [15]. Examples of the different collagen fiber type are presented in Fig. 3.3.

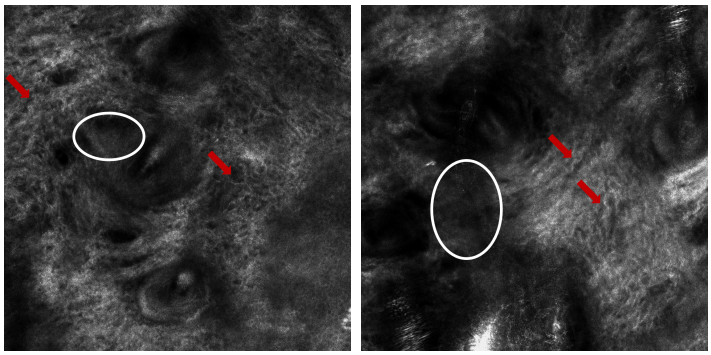
We have reviewed the most important skin aging patterns on RCM images. In the next section, we present our study protocols to generate several databases.



(a) Reticulated collagen fibers



(b) Coarse collagen fibers



(c) Curled collagen fibers

Figure 3.3: Collagen fiber types: (a) reticulated fibers characterized by tiny reflective fibers orderly disposed to form weblike structures, (b) coarse fibers composed by thick fibers grossly arranged to form nets, (c) curled fibers showing up as undulated highly reflective fibers (red arrows). Huddles of collagen: constituted by amorphous hyporeflective material (white circles).

3.4 Database generation

We aim to provide an automatic tool to study the skin under variable conditions such as aging, the environment, or cosmetic product application. Therefore, a standardized image acquisition protocol must be put in place to reduce the inter- and intra-individual variability and to produce repeatable and reproducible results. We focus on acquisition of z-stacks sequences from the surface up to the upper dermis, which are similar to optical biopsies of the skin.

In the following, we present three databases acquired during the thesis. The first was acquired during the first months of our research for image analysis development. The second one was acquired during a clinical study carried out in the second year for clinical validation, and the last one was acquired during the last year to test if the algorithms developed were able to show evidence of a skin care product efficacy.

For each of the databases, the study protocol is presented through the following items: the population of interest, the inclusion criteria. Then, the methodology is detailed through the definition of the evaluation criteria, the randomization procedure, the image acquisition protocol and the statistical analysis strategy.

3.4.1 Database 1: image analysis development

The first database was acquired during the first months of our research after a 3-days training on RCM images and their visual characterization. Its purpose is the development of image analysis algorithms allowing to highlight differences between age groups.

Population

The first database is composed of 15 subjects with fair skin who were assigned to two groups: a 7-persons group aged from 18 to 25 and another 8-persons group aged from 55 to 65.

Database 1 was acquired in-house. Subjects were selected among a large panel containing around 3000 volunteers.

Inclusion criteria

Meta-data were collected, in order to decide whether a subject could be included in each study.

For inclusion in the study, subjets must:

- not have undergone any plastic nor aesthetic surgery on the acquired areas;
- not present any skin lesions nor skin disorder on the acquired areas;
- not be pregnant nor breast feeding;

- respect the study criteria (such as age);
- not be prescribed any medication that can interfer with the study parameters (such as corticoids);
- have signed the legal waivers;
- be able to follow the study protocol; and
- must be medically insured.

Appropriate consent is obtained from all subjects before imaging. Subjects are informed of their right to stop participation or withdraw consent at any time without incurring any penalty.

The inclusion criteria described above remain the same for each of the study.

Main evaluation criteria

This database is used for the development of image analysis methods. The first objective is the validation of the proposed methods. Therefore, ground-truth annotations are performed to assess the robustness of the developed methods.

- The regularity of the honeycomb pattern was visually classified as regular or irregular on 30 images.
- Ground-truth segmentations of the skin layers (Epidermis, DEJ, Dermis) was performed on 23 images.

Then, differences between age group were assessed to evaluate the sensitivity of our methods to detect signs of aging.

Randomization

Each subject entering the study is assigned a number.

Method

RCM images were acquired using a near-infra-red reflectance confocal laser scanning microscope (Vivascope 1500; Lucid Inc, Rochester, NY, USA) [29]. Each image corresponds to an horizontal section with a $500 \times 500\,\mu$m field of view and a resolution of 1000×1000 pixels.

Image acquisition was carried out on the cheek to assess chronological aging, the dorsal forearm (a photo-exposed area) and the volar arm (a photo-protected area) to assess photo aging. On each imaged site, three stacks were acquired from the skin surface to the reticular dermis with a step of $5\,\mu$m. Overall, the database is composed of 134 z-stacks.

In order to optimize imaging, it is essential to find an appropriate patient position that allows a good microscope access to the selected skin area. Depending on the number of acquisitions and areas, one imaging session can take from 15 minutes (for one or two areas) to 1 hour (more than three areas). It is therefore crucial to maintain the patient in a comfortable position to avoid immediate changes in the imaging depth due to body movement. Moreover, the skin is a constantly changing organ in response to outdoor environment, especially to temperature and humidity. Patients must be kept in a room, with controlled temperature and humidity, for at least 15 minutes before imaging for the skin to return to a neutral state. We followed the following imaging protocol: zero depth (z = 0) is set at the skin surface for each z-stack acquisition as the skin thickness may be uneven (i), a 5 μ step is chosen to avoid overlapping depth imaging (ii), at least three z-stacks should be acquired (iii).

Several parameters are needed during the acquisition process, such as the setup of the zero depth (z = 0), from which the acquisitions will start, and the maximal depth acquisition. In order to properly interpret and acquire RCM images, one must refer to the cellular and molecular structure of the skin. As seen in the Chap. 2, melanin and keratin provide natural contrast in RCM images. The skin surface is the brightest image compared to other epidermal layers not only due to this keratinocyte migration but also because of the back-scattered light caused by the difference in the refractive indexes at the interface between the microscope immersion medium and the skin surface.

By understanding the skin physiology and the optical path, the surface of the skin may be properly found to setup the start of acquisitions. Zero depth should be set when the bright reflection of the skin surface gets higher at the center of the image, correlating to the level of the outer layer of the epidermis. An example of zero depth is presented in Fig. 3.4.

Furthermore, when choosing a protocol for the z-stack, it should be considered that imaging depth is limited to 150-200 μm, usually correlating with the upper dermis, see Fig. 3.5. The presence of artifacts altering the image formation (hair, air bubbles, etc.) should be considered and avoided.

The imaging protocol stands for each of the databases.

Statistical analysis

The robustness of the proposed methods is assessed by comparing our results to the ground-truth annotations using metrics derived from the confusion matrix.

To assess the ability of our proposed methods to detect signs of aging, results from the two age groups are compared. Where data are parametric, mean and standard deviation are analyzed using Student's unpaired t-test to test the hypothesis that the means of two

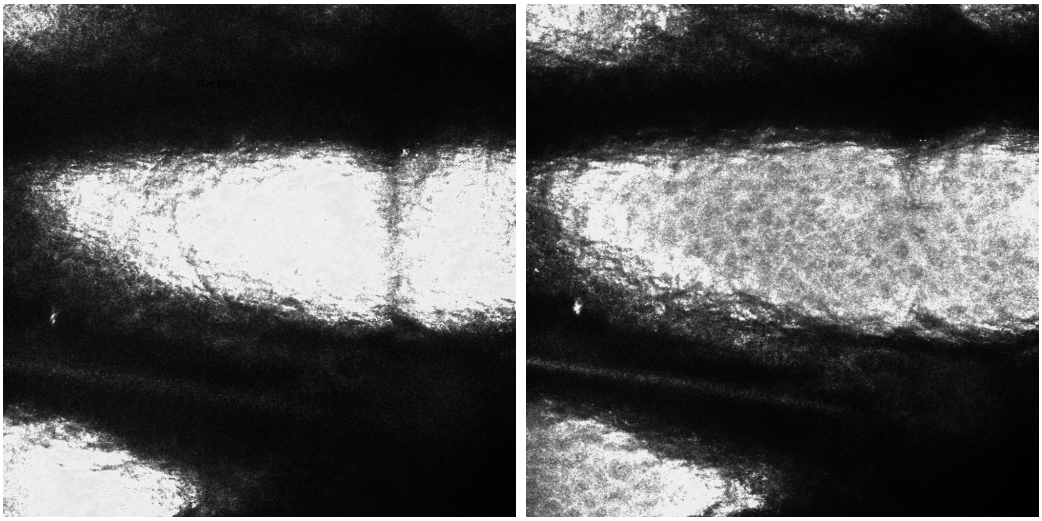


Figure 3.4: Example of setting zero depth. The zero depth should be set for the image on the left which shows a higher reflection. One can notice the presence of epidermal cells on the right image.

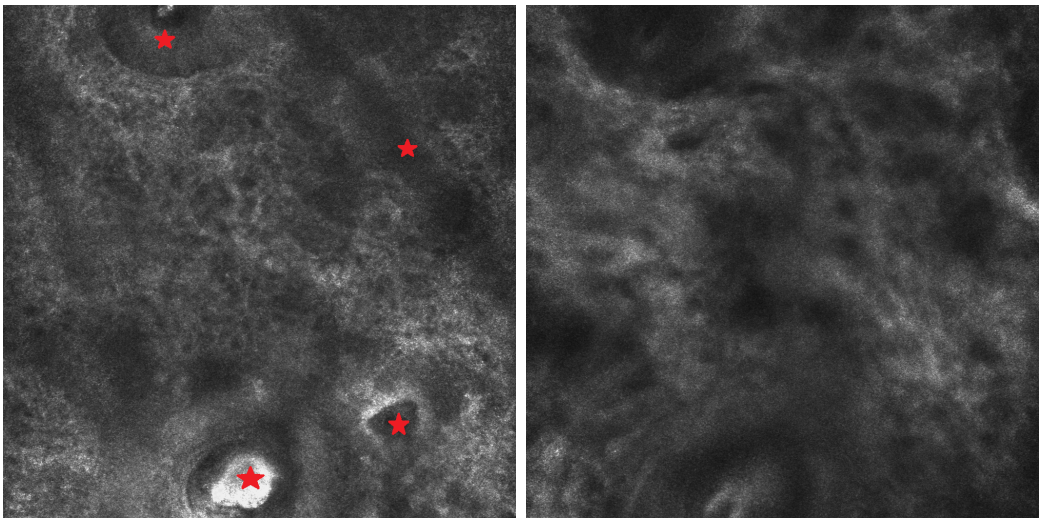


Figure 3.5: Imaging depth limitation. Left image (z = 100 μ). Right image (z = 150 μ) The red stars outline the presence of artifacts.

groups are the same. Non-parametric data are subjected to a Mann–Whitney test which does not rely on normally distributions of the two groups. Distributions are compared using a two-sample Kolmogorov–Smirnov test.

P-values lower than 0.05 are considered significant. The statistics significance is defined as follow:

- * : $0.01 < P\text{-}values \leq 0.05$;
- ** : $0.001 < P\text{-}values \leq 0.01$;

- *** : *P-values* ≤ 0.001.

All image analysis and statistical analysis are carried out blind.

3.4.2 Database 2: clinical validation

The database 2 is dedicated to the clinical validation of our methods on a large dataset.

Population

The study enrolled 160 subjects from 4 different ethnic groups (Caucasian, Hispanic, African and Asian). For each group, two age groups are defined: a 80-persons group aged from 15 to 35 and another 80-persons group aged from 45 to 65.

An illustration of the studied population is presented in Fig. 3.6.

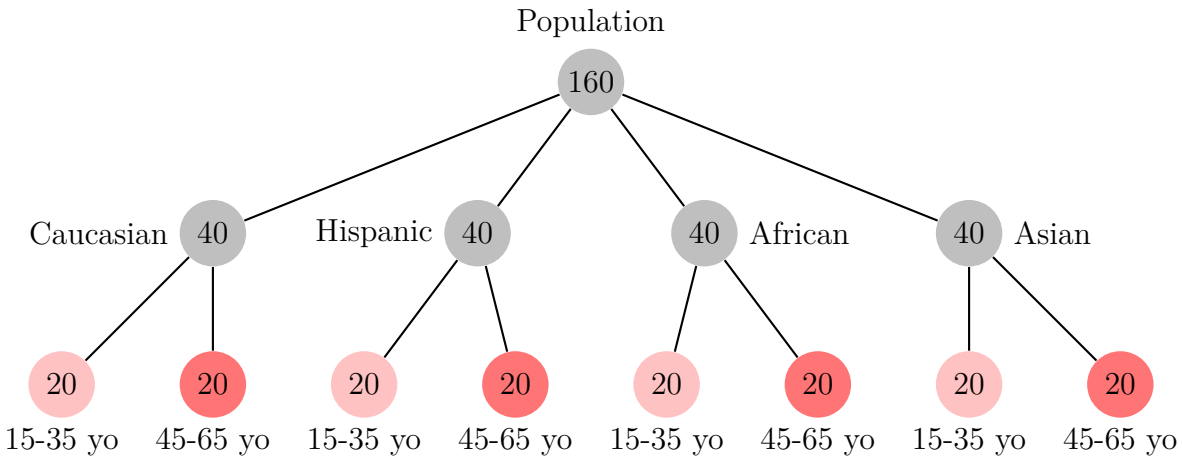


Figure 3.6: Clinical study population. The numbers inside each circle represents the number of subjects.

This clinical database was acquired by Professor Giovanni Pellacani and his team from the Department of Dermatology University of Modena and Reggio Emilia.

Inclusion criteria

The same inclusion criteria are applied as in database 1.

Main evaluation criteria

This objective of this study is to evaluate the performance of our proposed methods on a large dataset annotated by experienced dermatologists. The following descriptors were used:

- the regularity of the epidermal honeycomb pattern (0: regular, or 1: irregular);

- the shape of the dermal papillae (0: ringed and 1: poly-cyclic);
- the collagen fibers dominant types (0: reticulated, 1: coarse, 2: huddles, 3: curled).

For the DEJ annotations, only 55% of the database was used, due to the poor visualization of the DEJ in the remaining images. 89 subjects were analyzed for the DEJ parameter.

The evaluation criteria is to assess whether our proposed methods agree with the dermatologists' annotations.

Method

RCM images are acquired using a near-infra-red reflectance confocal laser scanning microscope (Vivascope 1500; Lucid Inc, Rochester, NY, USA) [29]. Each image corresponds to a horizontal section with a $500 \times 500\,\mu$m field of view and a resolution of 1000×1000 pixels.

Image acquisition are carried out on the cheek. On each imaged site, three stacks were acquired from the skin surface to the reticular dermis with a step of $5\,\mu$m. The same imaging protocol, as in database 1, is applied.

The second database is composed of 480 z-stacks.

Statistical analysis

We propose two types of analysis:

1. the statistical analysis of our measures according to the clinical annotations. Results from different ground-truth sets are compared;
2. some metrics (confusion matrix metrics and ROC curve analysis) to assess the performance of a predictive model build from the annotated data.

The same tests and criteria are used as in the first analysis

All image analysis and statistical analysis are carried out blind.

We assess how a predictive model performs when predicting the clinical annotations from our set of features. We use Random Forest classifiers with bootstrapping and feature bagging procedures (see Chap. 4 for further explanation). We perform a 10-fold cross validation. From the confusion matrix, we extract several metrics: global accuracy, sensitivity and specificity.

The sensitivity measures the proportion of actual positives that are correctly identified as such. The specificity measures the proportion of actual negatives that are correctly identified as such.

At each fold of the cross-validation, we calculate the Receiver Operating Characteristic (ROC) and the Area Under Receiver Operating Characteristic (AUROC).

The ROC is a graphical plot that illustrates the diagnostic ability of a binary classifier when its discrimination threshold varies. The ROC is created by plotting the true positive rate, i.e., the sensitivity, against the false positive rate, i.e., 1-specificity. This method considers both sensitivity and specificity while taking into account the discrimination threshold used during the classification, i.e., the probability above which an outcome is considered as positive by the classifier.

Intuitively, a perfect classifier reaches the top left corner of the graph, i.e. it maximizes the true positive rate while minimizing the false positive rate for a given discrimination threshold.

The AUROC is the area under the ROC. It summarizes the classifier performance but does not account for the optimum discrimination threshold. The AUROC indicates whether a randomly chosen True Positive has a higher probability to be classified as positive than a randomly chosen True Negative.

3.4.3 Database 3: cosmetic product efficacy

The last database acquired during this thesis is dedicated to a cosmetic product efficacy study. In most cases, the effects of an active ingredient or an anti-aging product are clinically visible after 1 or 2 months of product application. However, we would like to be able to predict these effects after only two weeks of application.

Population

The third database is composed of 17 subjects aged from 45 to 60 years old.

Inclusion criteria

The same inclusion criteria are applied as in database 1.

Specific instructions for this study

In order to reduce the inter and intra-individual variability, additional instructions are necessary. Subjects undergo a 2-weeks period of "wash-out" before the study began, i.e. they are asked to use a specific cosmetic product containing no active ingredient in order to set a neutral state for every subject. Four areas are delineated on the volar arms of each subject, a sun protected area, to test the four different conditions. For two weeks after the wash-out period, three cosmetic products are applied on the subjects' arms every morning by one of Clarins study monitor. No other cosmetic products nor skin treatment are allowed on the day of any of the acquisitions. During the study, subjects are asked to control their sun-exposure on the acquisition areas, i.e. on their volar arm.

Main evaluation criteria

The last database was not annotated, as its goal is not to validate the algorithms but to assess whether a cosmetic product effect can be measured or not.

Randomization

The areas are randomized for each subject and the analysis carried out in a double-blind fashion. All image and statistical analyses are carried out blind. The product application areas are randomized. The randomization is only disclosed at the end of all analysis.

Method

RCM images were acquired using a near-infra-red reflectance confocal laser scanning microscope (Vivascope 1500; Lucid Inc, Rochester, NY, USA) [29]. Each image corresponds to a horizontal section with a $500 \times 500\,\mu$m field of view and a resolution of 1000×1000 pixels.

Acquisitions were carried out at three times: initial time $T0$, one-week of application $T1$ and two-weeks of application $T2$. At each time, three z-stacks are acquired. The same imaging protocol, as in database 1, is applied.

Four different conditions are tested:

1. P1: A controlled condition with no product application
2. P2: Application of a moisturizing cosmetic product (product number 1)
3. P3: Application of an anti-aging cosmetic product (product number 2)
4. P4: Application of an anti-aging cosmetic product (product number 3)

The two anti-aging products have different cellular action pathways. One contains a high amount of Retinol, a well-known anti-aging substance which stimulates collagen production. The other one contains Alpha Hydroxy Acid (AHA) agents, which stimulate skin renewal. The goal is to measure the product efficacy and to further understand the different pathways of action of active ingredients in vivo.

When it comes to measuring a product efficacy, subjects should be repositionned in the exact same manner before and after the product application to avoid any bias.

As the confocal microscope achieve a lateral resolution around 1 micron, repositioning the probe requires a high precision. To achieve this, masks were made out from plastic sheets.

For the first venue of the subjects, the acquisition areas and confocal probe position are delineated and cut into the plastics sheets. One plastic sheet is used for cosmetic product application and the other one for the confocal probe repositioning. Visual landmarks

(such as nevi or scars) are annotated on the sheets to help for the repositionning. An illustration of the repositionning strategy is presented in Fig. 3.7.

Once the confocal probe is in place, the laser beam is positionned at the coordinates of the previous acquisition and the operator searches for visual landmarks on the confocal monitor to find the first acquired zone, such as wrinkles or any other recognizable features. Since the acquisition procedure is time-consuming, the last step may not be completed at each acquisition. However, the use of masks and acquisition coordinates ensures us that the acquisitions are nearly made in the same skin area. An example of the laser beam positioning is presented in Fig. 3.8.

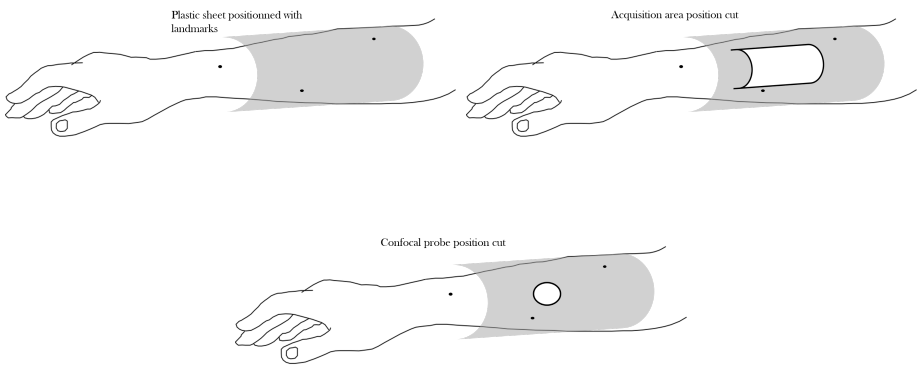
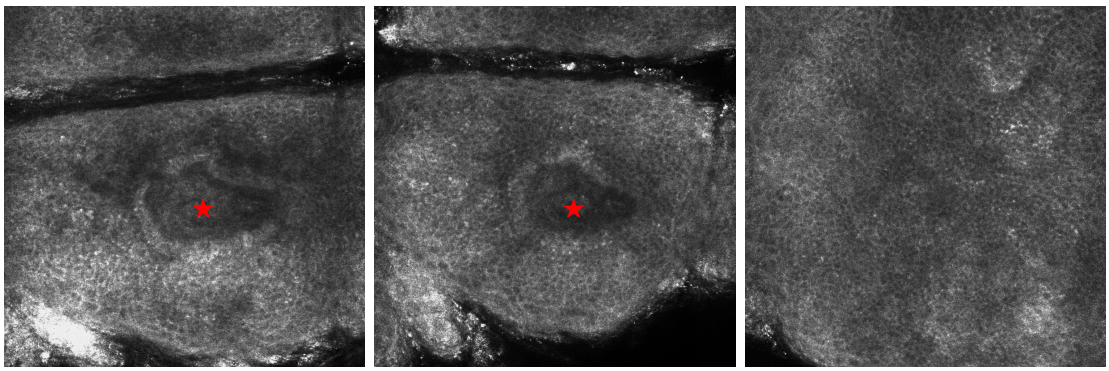


Figure 3.7: Plastic sheets used for repositionning.



TO $T1$ $T2$

Figure 3.8: Repositionning of the laser beam. At $T0$, a noticeable dermal papillae is visible (red star). At $T1$, the laser beam is almost perfectly repositionned. Due to timing constrains, it could not be found at $T2$.

Statistical analysis

Statistical analyses are performed to assess the significance of the variations of our proposed measurements over time on each of the acquired area. Our first concern is to evaluate if there are any significant variations over time for the controlled condition (condition P1 with no product application). As we do our best to maintain the controlled conditions stable over time, our hypothesis is that the significant changes observed in the patient's skin are due to the action of the product. Therefore, the measures are compared to the $T0$ acquisition for each product.

If there is a significant variation over time in the controlled condition, we assume that the skin intrinsic structures were modified due to external factors, such as the environment. In this case, the difference between the products and the controlled conditions is analyzed at each time, i.e. the variations due to external factors are discarded.

Our goal is to assess whether we are able to measure early signs of product efficacy. We search, not only for statistically significant changes, but also for early trends of product activity. As a two-weeks application is a short period of time for any cosmetic product to operate and the study population is rather small, P-values lower than 0.1 will be considered limit significant.

The statistics significance are defined as follow:

- LS : $0.05 < \textit{P-values} \leq 0.1$;
- * : $0.01 < \textit{P-values} \leq 0.05$;
- ** : $0.001 < \textit{P-values} \leq 0.01$;
- *** : $\textit{P-values} \leq 0.001$.

The conditions are then designated as follow:

1. P1: Controlled condition;
2. P2: Moisturizing product;
3. P3: Product containing Retinol;
4. P4: Product containing AHA;

3.5 Conclusion

In this chapter, we have reviewed the skin aging pattern imaged by in-vivo confocal microscopy.

We have presented the three studies conducted during this thesis.

It is well known that database design and production is an essential part of the development of image analysis algorithm. It covers areas such as content, size, ground truth and standardization. The significance of a study is directly linked to the content of the database which may be a source of bias in development.

Several bias are of concern regarding the database generation. The first one is related to the database size, which may not be large enough to adequatly represent the entire study population, especially when inter-individual variability is high, such as in skin studies.

Furthermore intra-individual variability may also be high in our field of research. The skin is a constantly changing organ, which adapts itself in response to the exterior environment. Moreover, depending on the acquisition area, the skin may show different patterns, which increases the intra-individual variability.

Our three databases cover a large range of intra- and inter-individual variability. Our proposed methods, which have been developed on a small database (15 subjects), are evaluated on a larger one (160 subjects from four different ethnic backgrounds) annotated by experienced dermatologists. Their representativity will enable us to evaluate their robustness when dealing with intra- and inter- individual variability.

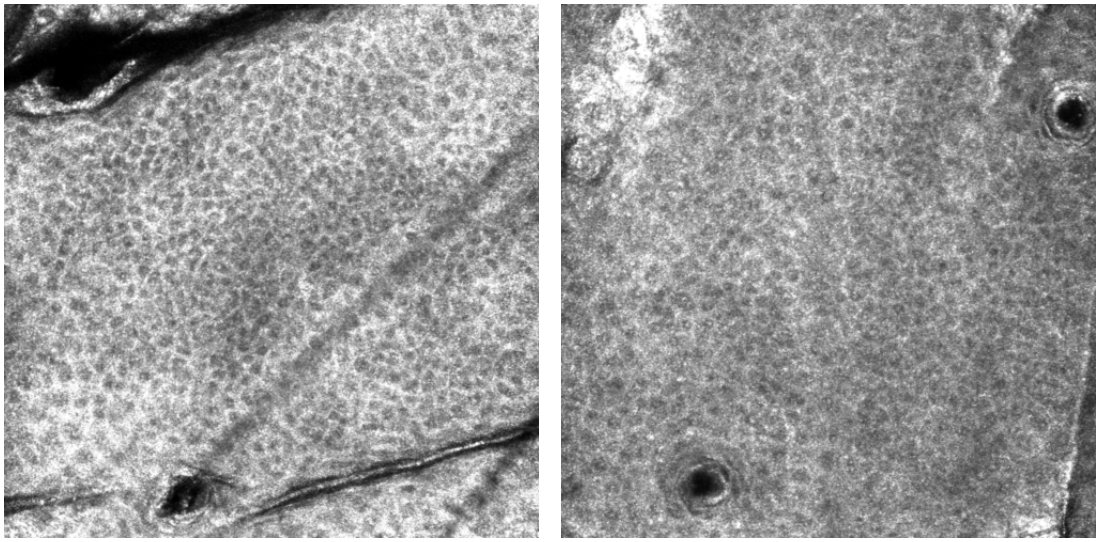
4

Quantification of the epidermal state

Contents

As seen in Chapter 2.2, the epidermis exhibits a honeycomb pattern on RCM images due to keratinocytes organization. As observations progress from the superficial to the deeper epidermal layers, the keratinocytes become smaller. Therefore, the honeycomb pattern tends to tighten with depth.

As the skin ages, the keratinocytes show increased variability in size and shape, leading to a distortion of the honeycomb pattern which progressively loses its regularity, see Fig. 4.1. When analyzing such pattern, dermatologists tend to identify variations in terms of size, cellular outlines, and architecture of the honeycomb pattern [14, 15]. Our goal is to provide a method for automatically quantifying the regularity of the epidermal honeycomb pattern using in vivo reflectance confocal microscopy.



(a) Young epidermis (b) Aged epidermis

Figure 4.1: Epidermal honeycomb pattern. One can notice that the images are affected by various defects: noise, blur, non-homogeneous intensity, and artifacts.

In this chapter, we first review mathematical preliminaries on graphs and connected operators. Then, we describe our proposed method for artifacts detection, as the skin surface is composed of several structures that influence the image formation at deeper layers. Once the artifacts are segmented, we address the epidermal cells segmentation with a watershed-based method to reconstruct the cell borders. Then, the honeycomb pattern is analyzed and cells are classified as *regular* or *irregular* with a supervised learning algorithm. Aggregated measures are finally defined to provide a global score for each image and results toward skin aging are presented.

4.1 Preliminaries

In this section, we review some notions on graphs and connected operators.

4.1.1 Graphs

We review the image definition as an undirected graph and some of the properties of the graph framework. The notions and notations presented below will be used in the following chapters of this thesis.

An undirected graph $G = (V, E)$ is defined by V, the finite set of *vertices*, and $E \subseteq \{\{x, y\} | x \in V, y \in V\}$ the set of *edges*. In our case, each element v in V represents a *pixel* or a *voxel* of the image domain and each edge e in E is a pair of elements of V that models the neighborhood relationship between two vertices (usually 4 or 8-connectivity for

2D images and 6 or 26-connectivity for 3D images). If there exists an edge $e = \{x, y\} \in E$, then we say that x and y are *adjacent*.

A graph $G = (V, E)$ is said to be *connected* if for any two pixels x, y in V there exists a path from x to y, i.e., a sequence of $n > 1$ vertices $\{x_0 = x, \ldots, x_n = y\}$ such that every pair $\{x_i, x_{i+1}\}$ is an edge of G. Classically, the entire image domain is connected.

Given a subset of pixels $S \subset V$, the subgraph *induced by* S is the graph whose vertex set is S and whose edge set contains any edge of G wich is made of two elements in S. Let $X = (V(X), E(X))$ and $Y = (V(Y), E(Y))$ be two graphs. If $V(Y) \subseteq V(X)$ and $E(Y) \subseteq E(X)$, then Y is a *subgraph of* X, denoted $Y \subseteq X$.

We say that Y i a *connected component of* X if Y is a connected subgraph of X and if Y is maximal, i.e., for any connected subgraph Z of X, $Y \subseteq Z \subseteq X$ implies $Z = Y$. An example of a connected subgraph which is maximal is presented in Fig. 4.2.

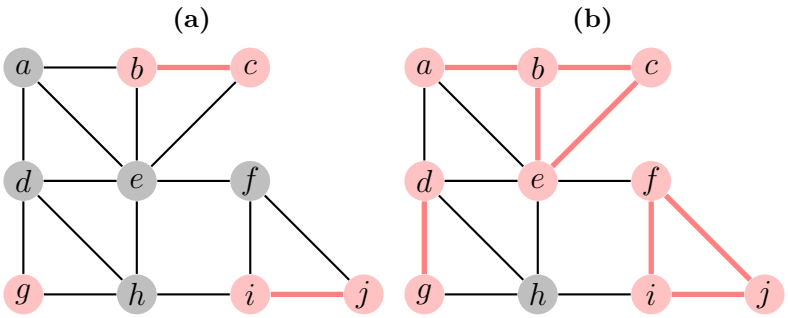


Figure 4.2: A graph G. In (a) a subgraph X of G (in red). Its extension in (b).

Images

A binary image X is a subset of the image domain V. We denote by X^c the complement of X, that is, $X^c = V \setminus X$. We consider the subgraph (X, E_X) of G induced by X, the set X is said to be connected if the subgraph (X, E_X) is connected. We define $C(X)$ as the set of connected components of (X, E_X).

A grayscale image is a function f from the set of vertices V (of G) to $\mathbb{R}$.

4.1.2 Connected operators

The notion of connected operator has been discussed in [54–56] for sets. A binary operator on a set V is a mapping from $V \times V$ to V. A connected operator is defined as follows:

A binary operator Ψ is *connected* when, for any binary image X, the set difference $X \setminus \Psi(X)$ is exclusively composed of connected components of X.

The extension of connected operators to grayscale images relies on the concept of partition. A *region* is a non empty connected subset of the image domain. A partition

(in regions) of an image I is a set of regions C_i, which are disjoint, and such that their union is the entire image domain.

A grayscale image can be decomposed into partitions through the notion of *flat zones* [57].

The flat zones $F_h(f)$ of a grayscale image f at level $h \in \mathbb{R}$ are the connected components of the level set of f at level h:

$$F_h(f) = C(\{x \in V, f(x) = h\}) \tag{4.1}$$

The set of flat zones of a gray scale image is a partition. It leads to the definition of gray-level connected operators:

An operator Ψ acting on gray-level images is *connected* if, for any image f, the partition of flat zones of f is finer than the partition of flat zones of $\Psi(f)$.

Connected operators operate by merging flat zones without splitting them. They preserve the contour and shape of flat zones. This characteristic makes connected operators useful for image simplification and other applications where contour preserving is essential, such as image segmentation [58–60].

One popular technique to create connected operators relies on a reconstruction process after a filtering step. In this section, we present some notions on connected operators through a reconstruction process.

Connected filters

Lets us start by discussing *connected openings* that are filters that can simplify images by removing some of their maxima.

Let $x \in V$ be a pixel, let X be a binary image, the binary opening Γ_x associates to X the connected component of X that contains x if $x \in X$ and $\emptyset$ otherwise. Γ_x is an anti-extensive operator, i.e. it can only preserve or remove connected components.

The most classical way to construct connected openings is to use an *anti-extensive reconstruction process*.

Let $X \subseteq V$ and $Y \subseteq X$ be two images. The anti-extensive reconstruction of X under Y is given by:

$$\rho_Y(X) = \bigcup_{y \in Y} \Gamma_y(X) \tag{4.2}$$

The upper level set $\mathcal{U}_h(f)$ of a grayscale image f at level $h \in \mathbb{R}$ is the set of pixels of V such that f is greater than h:

$$\mathcal{U}_h(f) = \{x \in V \mid f(x) \geq h\}. \tag{4.3}$$

Let f and g be two grayscale images (called the *reference* and the *marker* image, respectively). The grayscale anti-extensive reconstruction of f under g is obtained by the reconstruction of every upper level sets of f under the upper level sets of g:

$$\forall x \in V, \rho_g(f)(x) = \bigvee \{h \in \mathbb{R} \mid x \in \rho_{\mathcal{U}_h(g)}(\mathcal{U}_h(f))\} \tag{4.4}$$

In practice, useful connected operators are obtained by considering that the marker image g is a transformation $\tau(f)$ of the input image f.

For example, g can be obtained by a subtraction of a constant h from the original image f to obtain a contrast simplification of the image f:

$$\rho_g(f) = \rho_{f-h}(f) \tag{4.5}$$

This operator is known as the *h-maxima operator*, or h-max for short. The maxima of low contrast are removed and the contours of the maxima of high contrast are well preserved. One common strategy is to subtract the resulting reconstruction from the original image to retrieve the significant "domes" (h-domes) of the image. An example of the h-max operator is presented in Fig. 4.3.

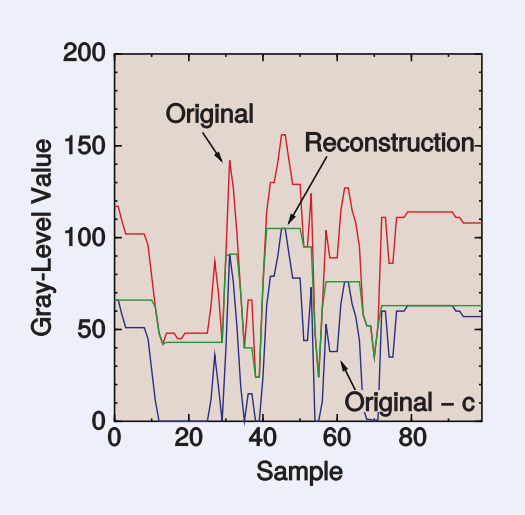


Figure 4.3: Example of the hmax transform. A constant c is subtracted from the original image f (in red) to create the marker image g (in blue). Illustration is taken from [56].

To retrieve the significant "basins" of an image, the extensive reconstruction by erosion $\rho_{f+h}(f)$ can be used. This operator, dual of the h-maxima operator, is called the *h-minima* operator.

Attribute filters

Connected operators can also be used for image simplification though the notion of attribute filter [58, 61–64].

An *attribute criterion* T is a predicate on connected component based on an attribute value. A binary attribute filter Φ_T removes the connected components of a binary image X that do not satisfy the predicate T:

$$\Phi_T(X) = \bigcup\{Y \in C(X) \mid T(Y)\} \tag{4.6}$$

A classical example of attribute filter is the area filter of size k which removes the connected components containing less than k pixels. An example of area filter is presented in Fig. 4.4.

(a) Original Image (b) $\lambda = 100$ pixels (c) $\lambda = 400$ pixels (d) $\lambda = 700$ pixels

Figure 4.4: Binary area closing.

The extension of attribute filters to grayscale image is based the decomposition of the image into upper level-sets. The attribute filter for a grayscale image f results in the preservation or removal of the peak components. It can be defined as follow:

$$\forall x \in V, \phi_T(f)(x) = \bigvee\{h \mid x \in \Phi_T(\mathcal{U}_h(f))\}. \tag{4.7}$$

The extension of area filters to grayscale images suppresses maxima whose area is less than a size threshold k. An example of a grayscale area filter is presented in Fig. 4.5.

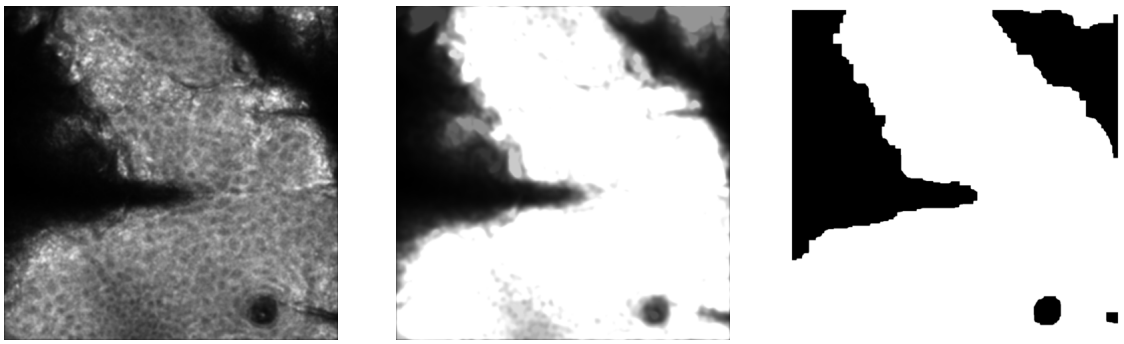
(a) Original Image (b) $\lambda = 100$ pixels (c) $\lambda = 400$ pixels (d) $\lambda = 700$ pixels

Figure 4.5: Grayscale area closing.

4.2 Artifact detection

The skin surface includes several structures such as hair or pores, which influence the image formation at deeper layers, distorting and shadowing the honeycomb pattern underneath them. Moreover, the skin topography is not smooth. Skin folds appear on the skin surface as dark regions in confocal images, which also disturb the image formation. Artifacts locations remain unchanged though all depths layers. Thus, the artifacts can be detected on the skin surface and applied as a mask on deeper images. The detection is performed just under the stratum corneum, at 20 μm below the surface, where the artifacts are the most contrasted.

The artifacts detection is presented on Fig. 4.6.



(a) Filtered image (b) Morphological reconstruction (c) Otsu threshold

Figure 4.6: Artifacts detection using connected reconstruction.

We apply a bilateral filter [65] with a large kernel to smooth the image while preserving edges. A bilateral filter is an edge-preserving and noise-reducing filter. It replaces the intensity of each pixel with a weighted average of intensity values from its neighbors while taking into account not only their spatial distance but also their similarity in terms of gray-level. Let $\mathcal{N}_i$ be the set of neighbors of a pixel i and $I[i]$ the gray-level of pixel i. The bilateral filtering of i is defined as follow:

$$I_{filtered}[i] = \frac{1}{W_i} \sum_{j \in \mathcal{N}_i} G_{\sigma_s}(\parallel i - j \parallel) G_{\sigma_r}(I[i] - I[j]) I[j] \tag{4.8}$$

with W_i is a normalization factor: $W_i = \sum_{j \in \mathcal{N}_i} G_{\sigma_s}(\parallel i - j \parallel) G_{\sigma_r}(I[i] - I[j])$ where:

- G_{σ_s} is a spatial Gaussian that decreases the influence of distant pixels. This distance is defined by is a parameter defining the extension of the neighborhood; and
- G_{σ_r} is a range Gaussian that decreases the influence of pixel j when its gray level differs from the one at pixel i.

In order to merge the brightest areas, we perform an anti-extensive reconstruction. Then, the reconstructed image is thresholded adaptively using the Otsu criterion. The Otsu criterion consists of calculating the optimum threshold separating two classes by maximizing the intra-class variance, or equivalently minimizing the inter-class variance [66]. The Otsu criterion for a threshold t for the inter-class variance maximization is defined as follow:

$$\sigma_B^2[t] = w_0[t] \times w_1[t] \times (\mu_0[t] - \mu_1[t])^2 \tag{4.9}$$

with w_0 and w_1 the class probabilities and μ_0, μ_1 the class means.

The Otsu threshold is the value T that maximises the inter-class variance, i.e. $T = \underset{t}{\text{argmax}}\sigma_B^2[t]$.

Validation of the artifact detection was by visual means on all images of the database.

4.3 Epidermal cells segmentation

Segmentation of the epidermal cells from RCM images is an essential step for studying their morphology, distribution and organization regarding the epidermal state. Epidermal cells have a dark central nuclei surrounded by refractive white thin cytoplasm.

In this section, we first review some state-of-the-art methods for segmenting the cells borders. Then, we describe our watershed-based segmentation of the epidermal cells. We finally present segmentation results and their validation.

4.3.1 State-of-the-art

Characterization of the honeycomb pattern has been studied on corneal cells using specular microscopy in [67, 68], and RCM in [69]. Analysis has long been performed by visual inspection of images acquired by means of confocal or specular microscope, possibly aided by digital tools to count the cells in order to evaluate their approximate density. Several automatic methods have been proposed. The main challenge is the location of the cell boundaries in the gray level image. Most of them rely on the segmentation of the cells contour based on relevant markers detection.

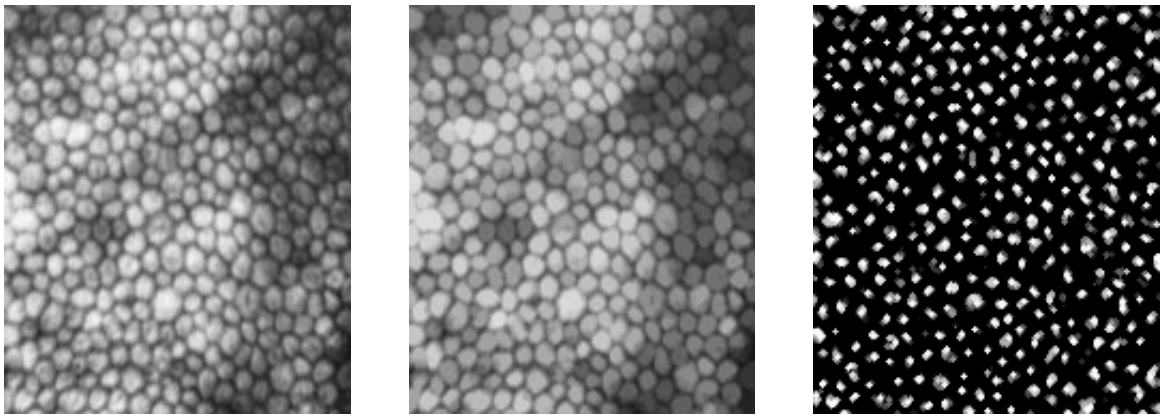
Watershed based methods

The watershed transformation was introduced as a tool for segmenting images and is now used as a fundamental step in many segmentation procedures where an image can be viewed as a topological relief [70–73]. The basic idea consists in placing markers (water sources) in the regional minima of the topological relief. Water enters through

the markers and floods the surface. During the flooding, barriers are set where the water coming from two different minima meet. The watershed transform leads to a partition of the image domain into catchments basins.

The main challenge using the watershed is to constrain the flooding with relevant markers. In [67, 68, 74], the authors propose several methods based on advanced morphological transformation, including grayscale reconstruction to retrieve relevant minima and to perform a watershed segmentation.

Vincent and Masters [67] perform a dome extractor based on morphological grayscale reconstruction to detect relevant minima. They chose not to take any size information into account but to rely on contrast information. In corneal images, the cells boundaries appear as a thin dark line. The authors subtract a constant h from the image I and compute a grayscale reconstruction of $I - h$ under I. An example is presented in Fig. 4.7.

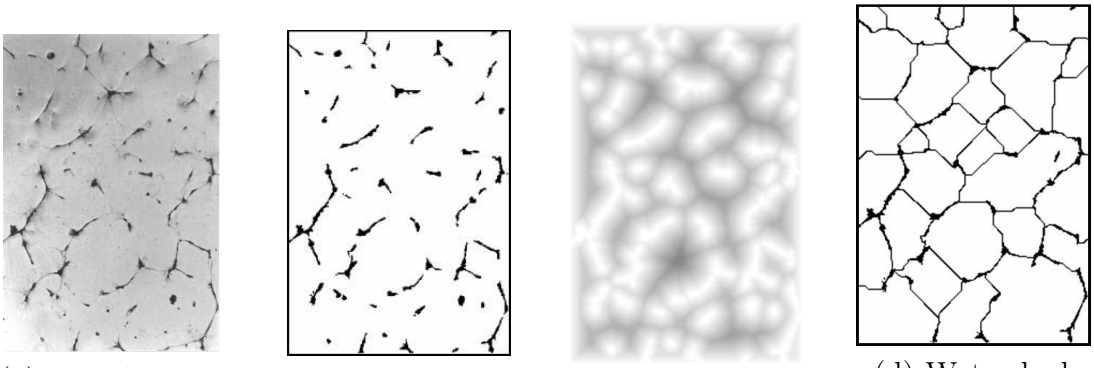


(a) Initial image (b) Morphological reconstruction (c) Extracted domes

Figure 4.7: Example of the hmax transform. Illustrations are taken from the work presented in [67].

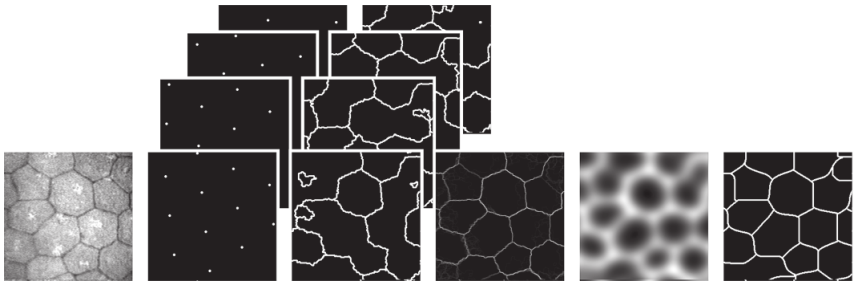
In both [68, 74], authors perform a similar dome extractor procedure but they work on a distance map, obtained from the image binarization, rather than on the original image. The distance map [75] indicates, for each pixel in a binary image X, the shortest distance to the nearest pixel in the complementary image X^c. The maxima of the distance map correspond to the interior of the cells and are used as markers, see Fig. 4.8.

Another strategy is used in [69], where authors first perform a stochastic watershed from which they extract the markers. A stochastic watershed [76] is an iterative procedure. At each step, n markers are randomly chosen on the image and a watershed is performed. The results of m iterations are added together, yielding a probability density function (PDF) of the cells contours. Pixels chosen more frequently are more likely to be part of an actual cell boundary. They suppress small minima with an h-minima transform. They



(a) Initial image (b) Binary image (c) Distance map (d) Watershed segmentation

Figure 4.8: Example of the watershed segmentation from the distance map. Illustrations are taken from the work presented in [74].



(a) (b) (c) (d) (e) (f)

Figure 4.9: Example of the stochastic segmentation. Illustrations are taken from the work presented in [69]. (a) Original Image (b) n random seeds (c) n randomly seeded watersheds (d) PDF (e) blur & h-minima (f) watershed

finally refine their segmentation with a classical watershed using the regions interior of their first segmentation as markers. An example of the method is presented in Fig. 4.9.

The choice of the parameter h is a crucial step in these methods, as it defines the selection criteria for the marker detection. When ground-truth is available, authors evaluate the resulting segmentation for several combination of parameters to find the optimal one. Other strategies rely an visual analysis of the segmentation results for several parameters.

Active contour

Active contour [77], also called snakes, are a special case of the general technique of matching a deformable model to an image by means of energy minimization. It is influenced by the image force that pull it toward specific features, such as contours.

Charłampowicz, Reska, and Bołdak [78] use active contours in order to segment the corneal cells. They define an energy containing two terms : the *image energy*, i.e., the intensity level at each position of the snake since the cell borders are dark lines, and a

regularization term to smooth the snake appearance. Because multiple snakes may be evolving on the image, a constrain is added, forbidding snakes to intersect or become nested.

Active contours require an initialization step, i.e. a marker of the cells border. In [78], the authors use an adaptative threshold followed by the detection of the largest (locally) circular regions containing almost only black pixels. The initial approximation of the cell borders are fitted to the actual cell borders using their active contour model.

Bayesian model

The method proposed in [79] relies on the introduction of a corneal cells shape model.

One of the theoretical frameworks for the expression and use of a-priori information is the Bayesian modeling framework [80]. It requires the expression of a statistical description of the model. Given a statistical description of a model, a Bayesian model can express a posterior distribution of the model.

This choice is related to the physiological assumption that a normal cell has a regular hexagonal shape, while the statistical description of the deviation from this shape is related to a pathological state of the cells.

In [79], define a cell field model $\mathcal{O} = \{o_1, \ldots, o_N\}$ which describes a set of N single cells models o_i with $i = 1, \ldots, N$. The posterior probability density function is expressed as:

$$p(\mathcal{O}) = \prod_{i=1}^{N} p(\mathcal{N}_i \mid o_i) p(o_i) \tag{4.10}$$

with $\mathcal{N}_i$ the set of neighboring cells of i.

The density function expresses two levels of information: the likelihood of a single cell o_i, and the likelihood of the cells mutual positions. A single cell model is defined with the density function $p(o_i)$ that expresses the prior information on the size, shape and cardinality of the cells. The cell interaction model is defined by $p(N_j \mid o_i)$ that takes into account the gradient information at the cells border.

The described framework strongly relies on an appropriate choice of starting values for parameters. In this case, the positioning parameters are the most critical. In [80], a first seed-point is manually chosen to initialize a starting cell model.

Conclusion

Of all the methods presented above, they all have in common the choice of relevant markers.

The methods based on active contours and Bayesian modeling both introduces prior information on the cells shape.

The watershed based methods give good results while relying on few parameters. They do not require an optimization procedure to retrieve the optimal solution.

For this reasons, we choose a watershed-based method.

4.3.2 Watershed segmentation of the epidermal cells

Now, we present our watershed-based method to reconstruct the cell borders, with prior information on the relevant minima.

Definitions

We follow the definition in [73] that states the notion of watershed in an edge-weighted graph, following the drop-of-water principle.

The intuitive idea underlying the notion of a watershed comes from the field of topography: a drop of water falling on a topographic surface follows a descending path and eventually reaches a minimum. The regions of a watershed are associated with the regional minima of the map. Each catchment basin contains a unique regional minimum, and conversely, each regional minimum is included in a unique catchment basin: the regions of the watershed extend the minima [81].

The catchment basins are separated by "curves", defined on the edges of the graph.

Following the definition in [73] an edge-weighted graph is a pair (G, f), with $G = (V, E)$ a graph and with f a mapping from E into $\mathbb{R}$ or $\mathbb{Z}$. If e is an edge of a graph G, $f(e)$ is the *altitude of* e. Let $X \subseteq V$ and $k \in \mathbb{Z}$. The subgraph $G_X = (X, E_X)$ of G induced by X is a *minimum of* f (at altitude k) if:

- G_X is connected,
- k is the altitude of any edge of G_X,
- the altitude of any edge of $E \setminus E_X$ containing a vertex of X is strictly greater than k.

We denote by $M(f)$ the graph whose vertex set and edge set are, respectively, the union of the vertex sets and edge sets of all minima of f. In the following, we denote $S \subseteq E$ as a subset of E, and $\overline{S}$ the *complementary set* of S, i.e. $\overline{S} = E \setminus S$.

The definition of a watershed following the drop of water principle relies on the notions of *extension* and *graph cuts*. Let X and Y be two subgraphs of $G = (V, E)$. We say that X is an extension of Y if $Y \subseteq X$ and if any connected component of X contains one component of Y.

A set $S \subseteq E$ is said to be a *graph cut* if there exists a partition of V such that S is the set of all edges of G whose extremities are in two distinct sets of the partition. If each component of the subgraph X induced by the separation S is connected, then S is a cut for this subgraph. An example is presented in Fig. 4.10.

Let $\rho = \{x_0, \ldots, x_n\}$ be a path in an edge weighted graph $G = (V, E)$. The path ρ is *descending* if, for any $i \in [1, n-1], F(\{x_{i-1}, x_i\}) \geq F(\{x_i, x_{i+1}\})$.

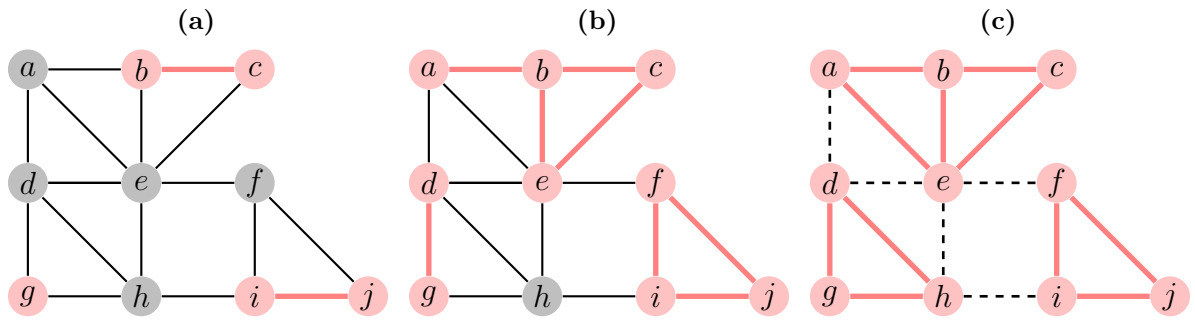


Figure 4.10: A graph G. In (a) a subgraph X of G (in red). One extension of X in (b). In (c) one maximal extension. The separation induced by the dashed edges in (c) is a graph cut for the graph G.

We say that $S \subseteq E$ satisfies the drop of water principle if $\overline{S}$ is an extension of $M(F)$ and if for any $v = \{x_0, y_0\} \subseteq S$, there are two descending paths $\rho_0 = \{x_0, \ldots, x_n\}$ and $\rho_1 = \{y_0, \ldots, y_m\}$ in $\overline{S}$ such that x_n and y_m are vertices belonging to two distinct minima of F and $F(v) \geq F(\{x_0, x_1\})$, resp. $F(v) \geq F(\{y_0, y_1\})$, whenever ρ_0, resp. ρ_1, is not trivial.

A *watershed cut* is a graph cut that satisfies the drop of water principle.

An example of a watershed cut is presented in Fig. 4.11.

The subgraph X induced by a watershed cut S, or a *watershed* S is a partition of the graph G. The connected components of X are the catchments basins of G.

In this formalism, the borders of the regions are defined on the edges of the pixels, thus each pixel is assigned to a unique region.

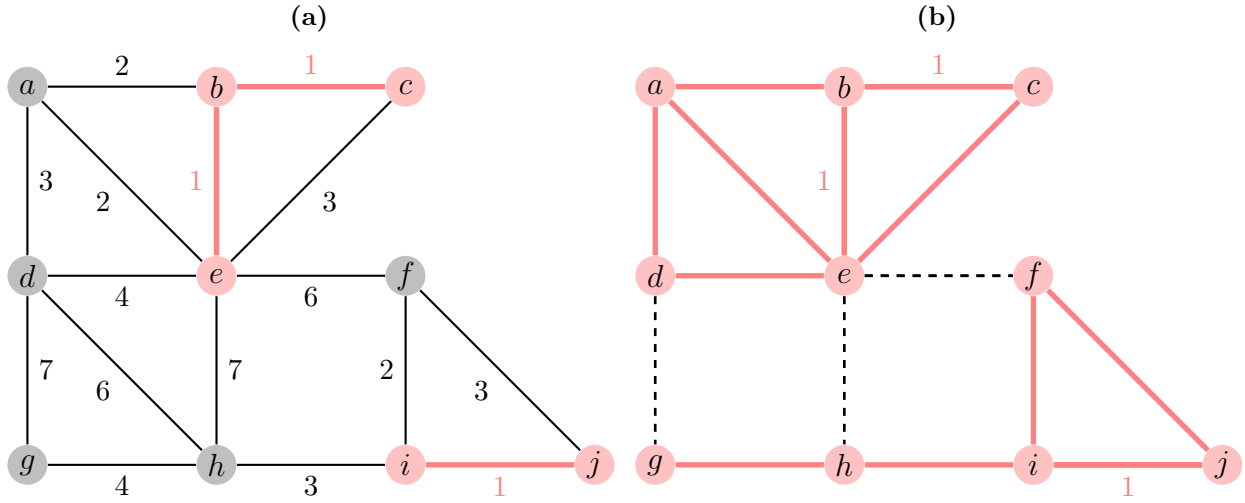


Figure 4.11: A graph G. In (a) the minima $M(G)$ are presented in red. The separation induced by the dashed edges in (b) is a watershed cut for the graph G. The red graph in (b) represents the catchmen bassins of G.

Graph from RCM images

Due to body motion during the acquisition process, such as subject respiration, 3D confocal stacks suffer from lateral distorsion. Each 2D confocal image is then segmented.

A 2D confocal image I is a map from the set of pixels v to a subset of the positive integers, with V a rectangular subset of $\mathbb{R}^2$. For any $x \in V$, $I(x)$ is the intensity at pixel x. In order to define a graph over the set of pixels, we consider the 4-connectivity relation between pixels.

To extract a watershed-cut from this graph, a map f, which weight the edges of $G = (V, E)$, must be defined. In our case, an edge u linking pixels x and y is weighted with the mean of the intensity values of x and y, i.e., $f(\{x, y\}) = \frac{I(x)+I(y)}{2}$. Each pixel is assigned to the closest local minimum by following a steepest path descent.

Due to the large amount of noise and blur, the direct application of the watershed transform leads to over-segmentation. An example is presented in Fig. 4.12.

The state-of-the-art methods presented above rely on either contrast or cell shape to select the relevant minima for the cells segmentation. As contrast on RCM images is highly dependent on the skin melanin content, we choose not to rely on contrast information as a selection criteria.

With aging, epidermal keratinocytes vary in size and shape as the keratinization process may be altered. However the typical range of cell diameters vary between 15 to 35 μm in the epidermal layers.

In order to remove non-relevant minima, and prevent over-segmentation, we use a connected filter that removes the connected components of the original image for which the area is smaller than a parameter λ. This filter is called an *area closing*.

Epidermal cells shape typically look like hexagons, and so are close to disks. In order to approximate the minimal cells area, we calculate the area of the circle in which a cell of minimal diameter fits. The resulting area is obtained as follow $A = \pi * \frac{15}{2}^2 = 176\,\mu\text{m}^2$. The corresponding area criterion then corresponds to 353 pixels, conservatively down to 300 pixels.

4.3.3 Segmentation results

Segmentation results are presented in Fig. 4.12. The positive impact of the minima filtering is noticeable.

We do not possess a manual segmentation of the honeycomb pattern as its manual delineation is a difficult and very time-consuming task. To validate our segmentation, manual counting of the epidermal cells is performed and compared to the number of cells segmented with our method. Manual counting was performed three times, means and

standard deviations were compared to the number of cells deduced from the watershed segmentations. The results are presented in Tab. 4.1. We remind that each confocal image is composed of 1000×1000 pixels, displaying a 500×500 μm field of vue. We achieve a mean cell density of 1132 cells per mm^2 with our watershed segmentation and a mean cell density of 1087 ± 6.6 cells per mm^2 with manual counting.

We also compare our cell density measurement to previous data reported in [38]. Our results agree with visual counts of nearly 1000 epidermal cells per mm^2.

Overall, our results are close to manual counting and previously reported data. The largest differences are observed on irregular honeycomb patterns. Our watershed segmentation tends to over-segment these patterns. This effect is also visible in Fig. 4.12. This is also true for the manual counting procedure, where higher standard deviation is mostly found on images characterized by an irregular honeycomb pattern. Well outlined cells are correctly segmented, whereas irregular contours lead to smaller segmented regions. Moreover, irregular epidermal cells show grainy interior patterns which can disturb the segmentation procedure.

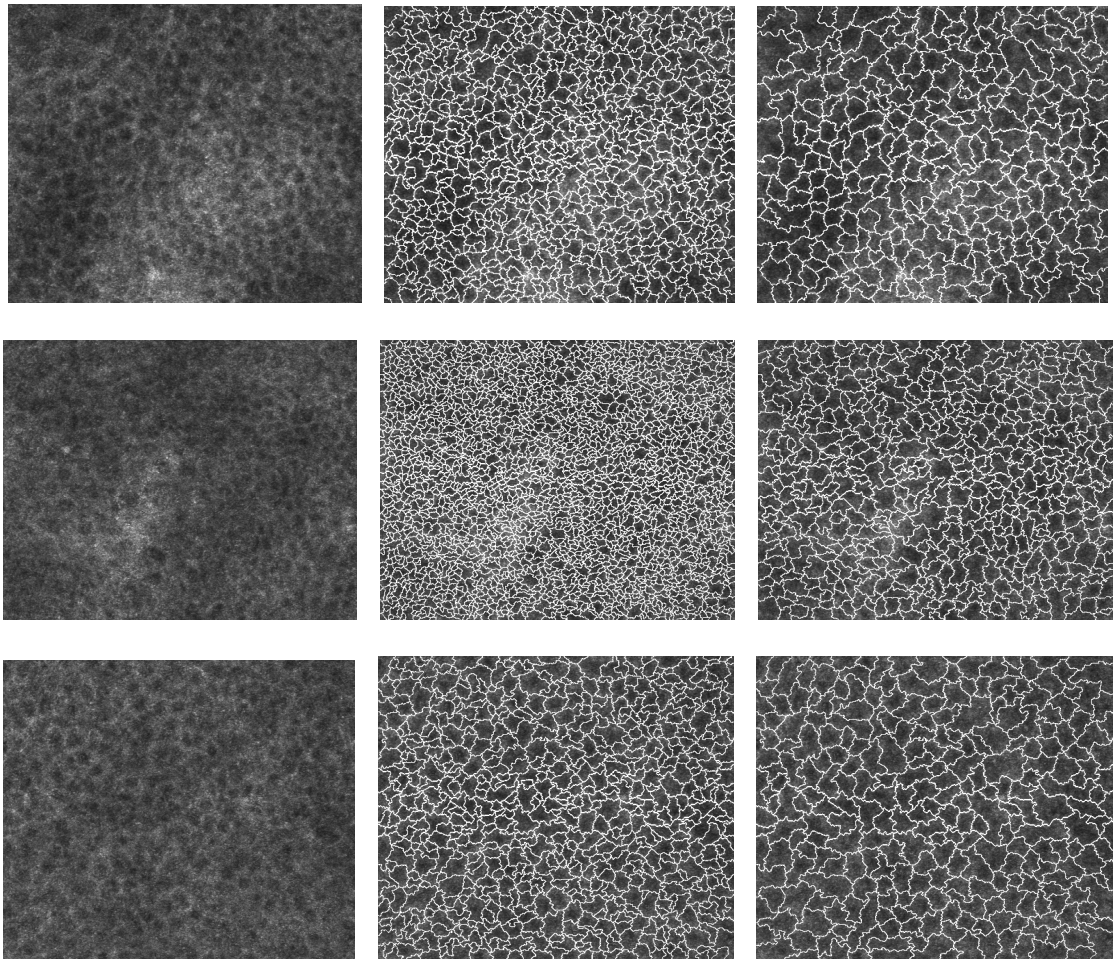
One way to improve our segmentation results would be to adjust the minima filtering criterion according to the depth of the processed image as the keratinocytes size varies with depth. Moreover, additional attribute criteria can be added to retrieve the minima of the cells such as volume, contrast etc.

Table 4.1: Cell density results

	Manual counting : mean ± standart deviation	Watershed segmentation
1	580.2 ± 13.5	660
2	585.7 ± 9.1	612
3	555.9 ± 4.4	564
4	504.3 ± 11.2	540
5	591.6 ± 5.7	608
6	456.4 ± 2.6	464
7	648.1 ± 7.3	672
8	528.5 ± 2.3	536
9	487.2 ± 1.8	498
10	497.5 ± 8.5	506

4.4 Epidermal cells classification

In this section, we present our method to extract information from the image segmentations. With aging, we expect the epidermal cells to show more variability in size and shape. We propose to classify each cell as regular or irregular based on its shape and neighborhood.



(a) (b) (c)

Figure 4.12: Segmentation results from the original image (a), coarse watershed segmentation without minima filtering (b) and watershed segmentation using the *area closing* (c).

When analyzing a honeycomb pattern, dermatologists focus on the size and shape of the epidermal cells to determine the regularity score of a confocal image.

In order to classify each cell, we propose the following features inspired from both the literature [67, 82] and the dermatologists expertise:

- area: the number of pixels in the cell;
- compactness: $4\pi \times (\text{area}/\text{perimeter}^2)$;
- elongation: λ_1/λ_1 with λ_1 and λ_2 the eigenvalues of the inertia matrix of the region;
- number of neighbors (computed from the region adjacency graph);
- average distance between cells centroids.

The features are computed from the region adjacency graph (RAG) of the cells segmentation. In a RAG, each region of an image, in our case each cell, can be represented as a node. Two nodes are connected by an edge if the corresponding regions are adjacent. Besides, each node in the RAG keeps the features of the corresponding region such as parameters, size etc. An example of a segmentation and its corresponding RAG is presented in Fig. 4.13.

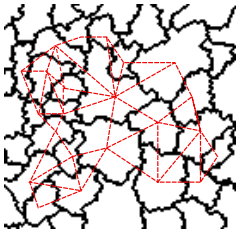


Figure 4.13: Example of a region adjacency graph. Each cell is represented by a node and is connected to its neighbors by edges (dashed red lines).

In order to perform a supervised classification, a training set was constructed to include about 1 500 cells taken at four different depths: 30 μm, 40 μm, 50 μm and 60 μm below the surface to cover all the epidermal cells diameters (as presented in Chap. 2). The training set was constructed so that it contained roughly the same number of regular and irregular cells. Examples of the cell labeling as regular or irregular are presented in Fig. 4.14.

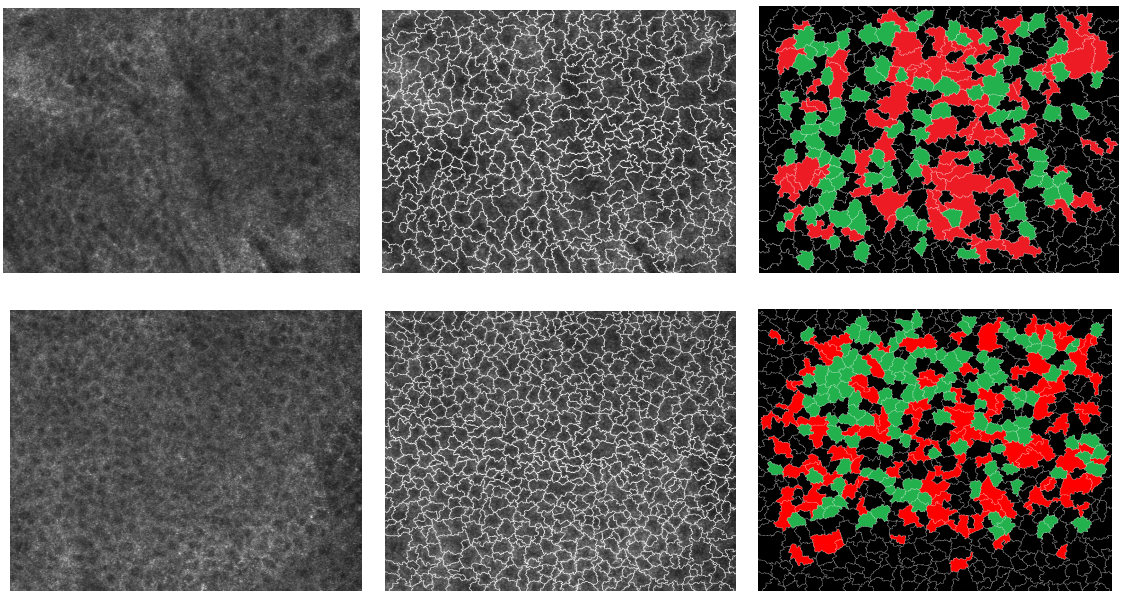


Figure 4.14: Cells ground-truth annotations. In red, the cells are labeled irregular. In green, they are labeled regular. The remaining cells (in black) are not labeled and do not enter the training set to keep the classes balanced or because their label could not be assigned with sufficient confidence.

4.4.1 Classification method

We use a Random Forest classifier [83] which is a supervised classification algorithm that integrates a number of decision tree classifiers. The idea is to combine several weaker classifiers to create a stronger one.

The root of a decision tree contains the entire dataset, while the leaves represents each class label. Each interior node correspond to one of the input features from which a decision rule is created to split the dataset into subsets with consistent labels. The process continues down to the leaves.

A Random Forest classifier is an ensemble of N decision trees classifiers learned on various subsets of the dataset, this process is called *bootstrap aggregating* or *tree bagging*. Moreover, at each candidate split, a random subset of features is selected to choose the most discriminant feature for the split, i.e. the feature that minimizes label misclassification. The misclassification rate for a split is called the *gini impurity*. This procedure is called *feature bagging*. Let M be the number of features, each subset of features usually consists in $\sqrt{M}$ features.

The final decision is taken by averaging the decision trees predictions.

Bootstraping and feature bagging are used to improve the model performance by decreasing the correlation between the decision trees which can introduce bias in the final decision. Bootstraping decreases the inter-tree correlation by learning on different training sets. Feature bagging prevents a very strong feature, if such a feature exists, to be chosen by too many trees, a phenomenon that would induce strong correlation.

4.4.2 Application

Since the cell size is depth-dependent, we train an independent Random Forest classifier for each depth. The four classifiers achieved accuracies above 90% on a 10-fold cross-validation test. The confusion matrices are presented in Tab. 4.2, 4.3, 4.4, 4.5. We achieve sensitivity and specificity above 90%, for each depth, see Tab. 4.6.

Table 4.2: Confusion matrix for the cells classification at depth 1.

	Regular	Irregular
Regular	170	5
Irregular	5	169

Table 4.3: Confusion matrix for the cells classification at depth 2.

	Regular	Irregular
Regular	179	8
Irregular	10	173

Table 4.4: Confusion matrix for the cells classification at depth 3.

	Regular	Irregular
Regular	213	8
Irregular	7	239

Table 4.5: Confusion matrix for the cells classification at depth 4.

	Regular	Irregular
Regular	116	9
Irregular	10	131

Table 4.6: Sensitivity and specificity results for the cells classification as regular or irregular

	Sensitivity	Specificity
Depth 1	0.97	0.97
Depth 2	0.94	0.95
Depth 3	0.97	0.95
Depth 4	0.92	0.94

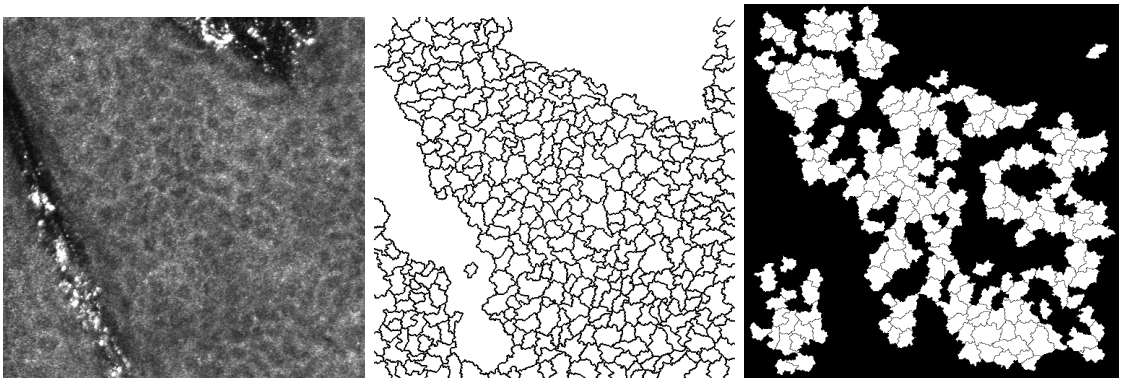


Figure 4.15: Epidermal cell segmentation, from left to right: initial image, segmentation, and regular cells class.

An example of the classification of regular cells is presented in Fig. 4.15.

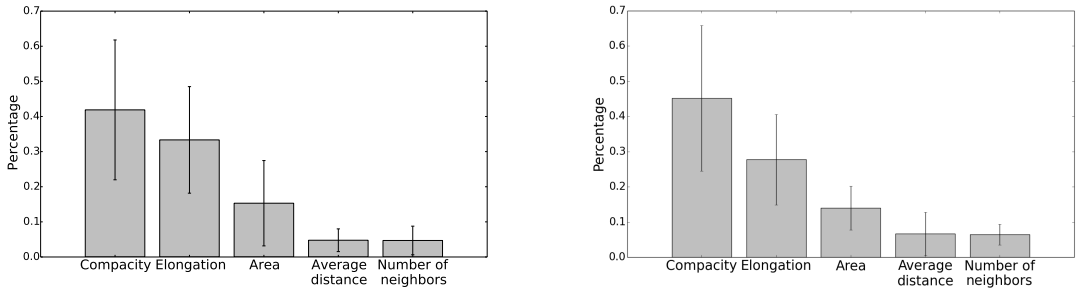
The most discriminant features are presented in Fig. 4.16. Bars represent the features importance in the forest prediction and the error bars the inter-trees variabilities using the *gini importance*. Features of more importance are more likely to split a large sample of the dataset.

We can observe that the compactness and elongation parameters are the most discriminant features, explaining 60% of the classification. The cells size appears to be relevant for the classification but not as discriminant as the elongation parameter. This result agrees with previous work reported in [47], which shows that keratinocytes diameter are only marginally influenced by age.

4.5 Measurements

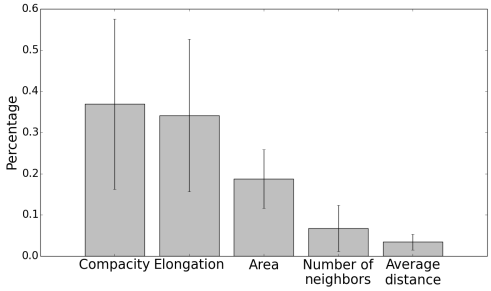
To date, image scoring has been performed by dermatologists, who estimate the proportion of irregular honeycomb pattern [15]. We propose to assess the irregularity of the honeycomb pattern on a 2D confocal image using these indicators:

1. $\text{WDC} = \dfrac{\text{Number of regular cells}}{\text{Number of cells}}$

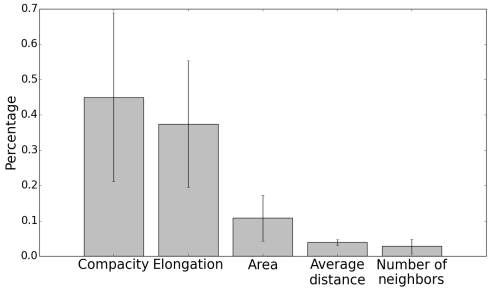


(a) Depth 1

(b) Depth 2



(c) Depth 3



(d) Depth 4

Figure 4.16: Feature importance in the Random Forest classifier for each depth. Percentages represent the feature importance in the Random Forest prediction, the higher the value, the more important is the contribution of the feature to the prediction function.

2. $\text{WDR} = \dfrac{\text{Number of regular cells}}{\text{Number of regular connected components}}$

WDC is the percentage of regular cells in the image.

The number of regular connected components corresponds to the number of region formed by neighbouring regular cells. WDR is an indicator of the size of the regular regions formed by neighbouring regular cells. The larger the regular regions are, i.e. the smaller WDR is, the more coherent the epidermis is, leading to preserved physiological functions. When computing WDR, we allow three regular cells to be 1-cell apart from each other while still forming a regular region, i.e. we consider that two cells are adjacent if there is a path in the RAG connecting them which is composed of less than three edges. For example, in Fig. 4.15, there is 3 regular regions formed by the cells classified as regular.

4.6 Results

In this section, we present our results obtained from the first database acquired during our research work, consisting in 15 subjects assigned to two age groups: a 7-persons group aged from 18 to 25 and another 8-persons group aged from 55 to 65. Image acquisition is carried out on the cheek to assess chronological aging, the dorsal forearm (a photo-exposed

area) and the volar arm (a photo-protected area) to assess photo aging. On each imaged site, three stacks are acquired from the skin surface to the reticular dermis with a step of 5 μm. The database is described in Chap. 3 and referred as database 1. The clinical validation results will be further presented in Chap. 7.

Each stack is represented by its mean scores for all depth. Each acquisition area is then represented by the mean scores of the three stacks.

Results from the method are expressed as mean and standard deviation. A non-parametric Wilcoxon test is used to compare the two groups of volunteers and image acquisition areas. *P-values* of less than 0.05 are considered significant. Fig. 4.17 shows box-and-whisker plots of all the measurements at all depths.

WDC is significantly higher on the cheek in the younger group (0.59 $\pm$ 0.05) than in the older group (0.57 $\pm$ 0.04) and on the volar arm (0.61 $\pm$ 0.03 for the younger group against 0.59 $\pm$ 0.03 for the older group) but do not show any changes in other location. WDR is significantly higher on the cheek and on the volar arm in the younger group. In the younger group, WDR is significantly higher on the volar arm (10.8 $\pm$ 2.3) than on the dorsal arm (8.4 $\pm$ 2.6). No significant difference is seen on the dorsal forearm, which tends to provide credence to the fact that the distribution difference on the other areas are indeed significant.

Similar results are found when comparing depths pairs instead of means for all depths.

Finally, in order to test if our method is able to predict the class (18-25 years old or 55-65 years old) of a subject, we train a Random Forest classifier on the 15 subjects.

Each subject is characterized by its mean score value on the cheek and volar arm. A leave-one-out cross-validation gives an average precision score of 73% of good classification with the Random Forest classifier.

4.7 Conclusion

Our first results show significant differences between young and old populations, most significantly on the volar arm. The percentage of regular cells remains stable with age and location, unlike the average size of well-defined regions. This may indicate that there is a link between region-forming in the keratinization process and epidermal aging. Moreover, the classification of a cell is highly dependent of the shape of neighboring cells, as their share a common border. The shape parameters involved might influence the classifier to label as regular neighboring cells sharing a regular border, which could explain the differences in WDR among age groups.

The honeycomb pattern irregularity is increased by 22% in sun-exposed area among young subjects, which indicates that irregular honeycomb pattern due to sun exposition

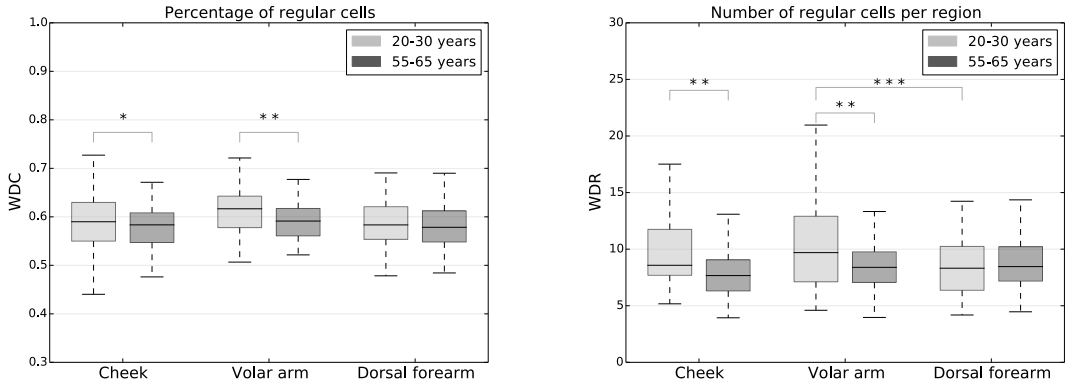


Figure 4.17: The box-and-whisker plots represent the data distributions through their quartiles. The vertical lines indicate the variability outside the upper and lower quartiles. Medians of distributions are represented as horizontal lines inside the boxes. Statistical results are represented as : * : $0.01 < P\text{-}values \leq 0.05$; ** : $0.001 < P\text{-}values \leq 0.01$; *** : $P\text{-}values \leq 0.001$. Results for the young (light grey) and old population (dark grey) are compared for each of the acquisition area.

can potentially be quantified using our method. To support this, in previous studies [16], it was found that irregular honeycomb pattern correlates with chronological aging and was exacerbated on UV-exposed areas. The classification results of our subjects in age groups are encouraging, given the small size of the population and intra-class variability.

In this chapter, we have presented a new method for the quantification of the regularity of the epidermal honeycomb pattern. A segmentation procedure was proposed using a watershed-based approach, followed by a classification of the epidermal cells as regular or irregular. Our two measures (the percentage of regular cells WDC and the average size of regular regions WDR) enable the quantification of both chronological and photo-aging.

One way to improve our method is to refine the cells classification along with the optimization of the Random Forest parameters to detect more subtle changes in the epidermal state. The honeycomb segmentation could also benefit from the inclusion of additionnal minima filtering, based on a volume attribute for example, to reduce the over-segmentation in the irregular cases.

The inclusion of new measures of the epidermal state for aging characterization should allow us to further improve on our results in future work.

In the next chapter, we present the segmentation of the dermal-epidermal junction and the characterization of its shape.

5

Segmentation of the dermal-epidermal junction

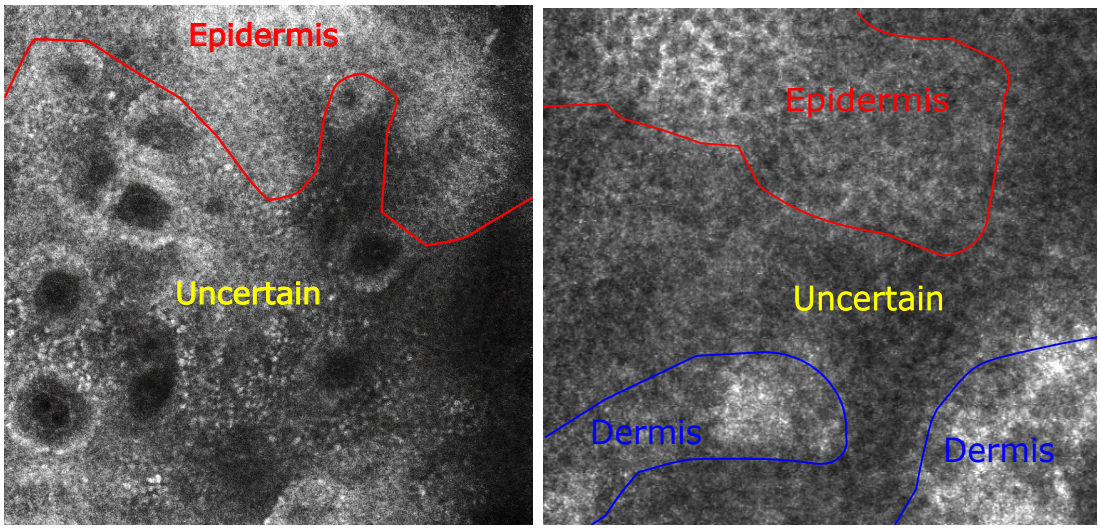
Contents

The DEJ is a complex, surface-like, 3D structure separating the epidermis from the dermis. Its peaks and valleys, called dermal papillae, are due to projections of the dermis into the epidermis. The DEJ undergoes multiple changes under pathological or aging conditions. In [14, 15], the authors have defined the shape of the dermal papillae as the confocal descriptor to study skin aging. The shape of the dermal papillae has been described as ringed or poly-cyclic.

In fair skin, the DEJ can have multiple confocal patterns. It can appear as an amorphous and low-contrasted structure or as circular rings which corresponds to the 2D vertical view of a dermal papillae, see Fig 5.1.

One particular difficult point is the visualization of the DEJ, which is hard to identify purely by visual means. Segmenting the DEJ is a useful task because it allows clinicians to better locate it, and opens the way to quantify its appearance. It is also an essential step in the automation of the skin aging characterization on RCM images. It could further improve the understanding of the skin physiological response toward aging or other environmental conditions.

We propose a segmentation method of the DEJ using 3D Conditional Random Fields (CRF). CRFs are a graphical models family that can allow us to encode biological information into the DEJ segmentation without having to statistically describe a DEJ shape model, which would be a challenging task. Our contribution lies on exploiting the additional depth and 3D information of the skin architecture to provide spatial regularization over label distribution and to model skin biological properties.



(a) (b)

Figure 5.1: Examples of different DEJ pattern. The circular rings pattern in (a) provides an easy way to identify the DEJ, whereas the identification is much less certain in (b). Unfortunately, the second case is the more frequent, especially on the cheeks.

As the DEJ appearance varies significantly and constrast can be weak, the precise localization of the DEJ is difficult. In practice ground-truth annotations often consist of 3 thick layers: the epidermis (E), a thick uncertain area containing the DEJ (U), and the dermis (D).

We aim to segment the confocal images in three classes: epidermis (E), uncertain (U) and dermis (D).

Our approach starts with a Random Forest classifier (RF), providing the probabilities of a pixel to belong to one of our three classes, with no assumptions on the

dependencies between pixels. The RF output encodes the textural information and defines the CRF potentials.

The CRF model parametrization is inspired by prior information on the skin structure. The skin architecture is modeled by the conditional relationship between pixels. The relations between pixel neighbors mimic the skin layers behavior in 3D by imposing a specified potential function according to their location in depth and their relative position to each other. The CRF potentials are derived from annotated skin RCM data.

The CRF model allows us not only to incorporate resemblance between neighbors, but also to specify biological information.

In this chapter, we first review state-of-the-art methods on the DEJ segmentation. Then, we shortly describe the CRF framework in order to present our CRF model used in the DEJ segmentation and we finally present some experiments and results.

5.1 Related work

Kurugol et al. [21] propose a machine learning-based method using textural features to automatically locate the DEJ location. They reproduce the key aspects of human visual expertise, who rely on texture and contrast differences between layers of the epidermis and dermis, in order to locate the DEJ. They have developed several algorithms to detect the DEJ in fair and dark skin. In fair skin, the DEJ ring-like pattern shows low contrast. Therefore, experts use local textural features to delineate the DEJ. They use a LS-SVM method, a variant of Support Vector Machine, which takes into account the expected similarity between neighboring tiles within images to include spatial relationship between neighbors and to increase robustness. A large set of features is proposed, with wavelet and Gabor features being the most discriminative. They achieve 64%, 41%, and 75% of correct classification of tiles for epidermis, uncertain region, and dermis, respectively. They propose a second approach in [84] which incorporates a mathematical shape model for the DEJ using a Bayesian model. The DEJ shape is modeled using a Marked Poisson Process. Their model can account for uncertainty in number, location, shape and appearance of the dermal papillae. Their output segmentation presents sharp deformations which could withhold the subtle deformations the DEJ suffers with aging.

Hames et al. [85] address the problem of identifying all of the anatomical strata of human skin using a 1D linear chain Conditional Random Field and structured support vector machines to model the skin structure. They show an improvement in the classification scores with the use of such a model. Their model achieves 85.6% classification accuracy on a test set of 100 RCM stacks.

The following methods perform an image-level classification instead of a pixel-level classification.

Another Bayesian approach is proposed in [86]. The authors focus on the classification of human skin lentigo, a clinical disorder of the skin. They propose a Bayesian approach with a Markov chain Monte Carlo (MCMC) algorithm to estimate posterior probabilities which they use for their classification.

Somoza et al. [87] use an unsupervised clustering method to classify a whole confocal image as a single distinct layer, resulting in a correlation coefficients of 0.84 to 0.95 between their predictions and the ground truth.

Kaur et al. [88] use a hybrid of classical methods in texture recognition and recent deep learning methods, which gives good results on a moderate size database of 15 stacks. They classify each confocal image as one of the skin layers. They introduce a hybrid deep learning approach which uses traditional feature vectors as input to train a deep neural network. Their approach performs with a test accuracy of 82%.

Bozkurt et al. [89] propose the use of deep recurrent conventional neural networks (RNN) [90, 91] to delineate the skin strata. The dataset used in this study is composed of 504 RCM stacks. Each confocal image is labeled in one of 3 classes : epidermis, DEJ or dermis. They first extract the last layer of a convolutional neural network (CNN) to obtain feature representations of each image. Then, a recurrent neural network is trained on a sequence of image features to classify each image. They make the assumption that the skin layers are ordered with depth: the epidermis is the top layer of the skin, followed by the DEJ and the dermis. The use of deep recurrent conventional neural networks on a large dataset allows them to take the sequential dependencies between different images into account. They have trained numerous models with varying networks architectures to overcome the computational memory issues. First, they trained the RNN on a partial sequence of the RCM stack containing a local neighborhood of 3 images around the RCM image to classify. Then, they performed a sequence-to-sequence training. The model processes entire RCM stacks and outputs predictions for each image in the stack. This approach is potentially more flexible, as it provides the model with the complete RCM stack and allows it to learn what information is useful for slice-wise classification. Their best models obtained 87.97% accuracy on their test set. With partial sequence training, the authors showed no improvement when enlarging the neighborhood of images around the RCM image to be classified.

Due to the uncertainty in visual labeling and inter-subject variability, state-of-the-art methods tend to combine textural information and prior information on the DEJ shape in their methods, either by modeling the DEJ shape or by using a linear regularization of the segmentation through depth. The use of a RNN enables the representation of the dependencies between images.

Most methods focus on the localization of the DEJ in depth, with no consideration toward the characterization of its shape. Our goal is to segment the DEJ shape in order

to quantify the modifications it undergoes during aging. In order to do so, we seek to combine textural information and 3D dependencies between pixels within a RCM stack to perform a pixel-level segmentation.

Graphical models appear to be well adapted and useful tools toward this purpose.

5.2 Graphical modeling with Conditional Random Fields

Segmenting boundaries of interest in 3D microscopy images is often challenging due to high intra and inter-subject variability, and the complexity of the boundary structures. In such applications, we wish to predict an output vector $\mathbf{x}$ of random variable given an observed feature vector $\mathbf{y}$. Our goal is to maximize the number of correctly classified labels x_i by learning a classifier that maps $\mathbf{y} \rightarrow \mathbf{x}$. Output variables can have complex dependencies. A natural way to represent the manner in which output variables depend on each other is provided by Graphical Models which include Bayesian networks, factor graphs, Markov random Fields, ..., and Conditional Random Fields.

With graphical modeling, a distribution over variables can be represented as a product of local functions that depend on a smaller subset of variables. The nodes in the graph are identified with random variables, and joint probability distributions are defined by taking products over functions defined on connected subsets of nodes. A detailed review of Graphical Models can be found in [92, 93].

Most works in graphical models have focused on models that explicitly attempt to model a joint probability distribution $p(\mathbf{y}, \mathbf{x})$. These models are fully generative, they identify dependent variables and define the strength of their interactions. The dependencies between features can be quite complex, and the construction of the probability distribution over them can be challenging.

A solution to this problem is the modeling of the conditional distribution, yielding a discriminative model. The main conceptual difference between discriminative and generative models is that a conditional distribution $p(\mathbf{y} \mid \mathbf{x})$ does not include a model of $p(\mathbf{x})$ which often contains many highly dependent features that are difficult to model.

This is the approach taken by Conditional Random Fields (CRFs) [94]. A detailed review can be found in [93].

Conditional Random Fields are popular techniques for image labeling due to their flexibility in modeling dependencies between neighbors and image features. Linear chain CRFs are the simplest and the most widely used. They have become very popular in natural language processing [95, 96] and bio-informatics [97]. Applications of 2D CRFs have also proved to be useful in computer vision applications [98, 99]. Medical imaging has been a field of interest in applying CRFs to many segmentation problems such as brain and liver tumor segmentation [100, 101].

5.3 DEJ segmentation using CRFs

An image $\mathbf{I}$ consists of M pixels $\mathrm{i} \in V = [1, M]$ with observed data $\mathbf{y_i}$, i.e., $\mathbf{y} = (y_1, \ldots, y_M)$. Pixels are organized in layers (en-face images) forming a 3D volume. We seek to assign a discrete label x_i to each pixel i from a given set of classes $\mathcal{C} = \{E, U, D\}$, with the epidermis (E), a thick uncertain area containing the DEJ (U), and the dermis (D). The classification problem can be formulated as finding a configuration $\hat{\mathrm{x}}$ that maximizes $p(\mathbf{x} \,|\, \mathbf{y})$, the posterior probability of the labels given the observations.

A CRF is a model of $p(\mathbf{x} \,|\, \mathbf{y})$ that can be represented with an associated graph $G = (V, E)$ where V is the set of vertices representing the image pixels and E the set of edges modeling the interaction between neighbors, [102]. Here, E is the usual 3D 6-connectivity.

We use a model with pairwise interactions defined by:

$$p(\mathbf{x} \,|\, \mathbf{y}) \propto \prod_{i \in M} \varphi_i(x_i, \mathbf{y}) \times \prod_{i,j \in E} \psi_{ij}(x_i, x_j, \mathbf{y}), \tag{5.1}$$

where $\varphi_i(x_i, \mathbf{y})$ is the *node potential* linking the observations to the class label at pixel i and $\psi_{ij}(x_i, x_j, \mathbf{y})$ the *interaction potentials* modeling the dependencies between the labels of two neighboring pixels i and j.

The CRF model is represented in Fig. 5.2.

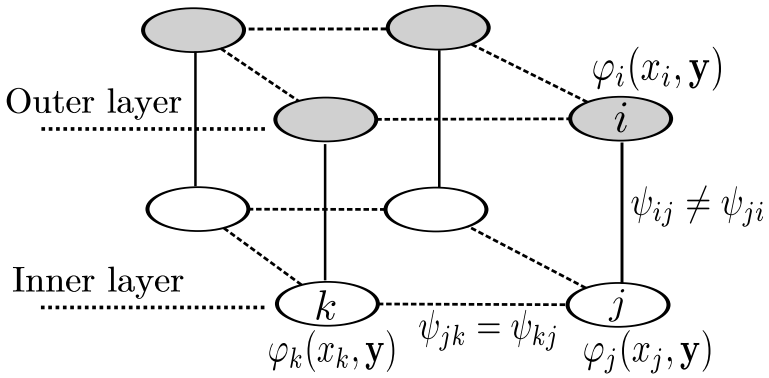


Figure 5.2: 3D CRF modelisation. The set of nodes in gray and in white belong to two different confocal sections. The edge potentials of each confocal sections ψ_{jk} are learned at each depth. Edge potentials between en-face sections ψ_{ij} impose biological transition constraints.

We propose to specify the CRF potentials in order to incorporate biological information. The skin layers are strictly ordered according to their depth. The epidermis is the top skin layer, followed by the uncertain area (containing the DEJ) and the dermis. Therefore, a pixel located near the surface will have a higher probability to belong to the epidermis than to the uncertain area or to the dermis. Moreover, the pixels labels must respect the skin layers order, i.e. if a pixel is classified as uncertain, the pixels located below

it can not belong to the epidermis and conversely, a pixel classified as dermis cannot be located above neither the epidermis nor the uncertain area. The classification has to respect the biological skin structure.

The CRF parameters are depth-dependent. We define D the set of all depths of the image $\mathbf{I}$. For each $d \in D$, we define:

- M_d the set of pixels at depth d,
- E_d the set of edges at depth d,
- $E_{d \to d\pm1}$ the set of edges between depth d and $d+1$, resp. $d-1$.

The notation $\mathbb{1}_{x=c}$ represents the indicator function, which takes the values 1 when $x = c$ and 0 otherwise. We denote by $\mathrm{x}^\top$ the transposed vector of x. The operators $\circ$ and $\cdot$ denote respectively the Hamadar product (element-wise multiplication) and the dot product of vectors.

5.3.1 Node potential

The node potential is defined as the probability of a label x_i to take a value c given the observed data $\mathbf{y}$ by:

$$\varphi_i(x_i = c, \mathbf{y}) = p(x_i = c \mid f_i(\mathbf{y})) \tag{5.2}$$

with $f_i(\mathbf{y})$ a feature vector computed at pixel i from the observed data.

In our case, each node potential $\varphi_i(x_i, \mathbf{y})$ is associated with the predicted class probability vector $f_i(\mathbf{y})$ produced by a Random Forest (RF) classifier, [83]. The product of the node potentials $\varphi_i(x_i, \mathbf{y})$ is expressed as:

$$\prod_{i \in M} \varphi_i(x_i = c, \mathbf{y}) = \prod_{d \in D} \prod_{i \in M_d} \mathbb{1}_{x_i = c} \cdot \theta_d \circ f_i(\mathbf{y})^\top \tag{5.3}$$

The parameter $\theta_d = [\theta_{epidermis}, \theta_{uncertain}, \theta_{dermis}]$ balances the bias introduced by labels appearing more often in the training data, i.e., the epidermal and the dermal one.

The feature vector $f_i(\mathbf{y})$ is a 1×3 vector of probability for a pixel to belong to each label. It is produced by a random forest classifier. We use the following textural features inspired by [21]:

1. first and second order statistics;

   We calculate statistical metrics mean, variance, skewness and kurtosis. As a second order statistic, we compute the power spectrum [103] of the image defined as:

   $$S(i,j) = \frac{\mid F(i,j) \mid^2}{M} \tag{5.4}$$

with F the Fourier transform of the image and M the number of pixels. The power spectrum is related to the autocorrelation of an image which describes how closely two points of an image are as a function of their distance and orientation.

2. gray level co-occurrence matrix contrast, energy and homogeneity;

   A co-occurrence matrix is the distribution of co-occurring pixel values of an image at a given offset according to a given orientation [104]. The gray level co-occurrence matrix G is a square matrix of order N (N being the number of gray levels in the image), it is defined as:

$$G(m,n)_{(\Delta m,\Delta n),\Theta} = \frac{P(m,n)_{(\Delta m,\Delta n),\Theta}}{\sum_{m,n=1}^{N} P(m,n)_{(\Delta m,\Delta n),\Theta}} \tag{5.5}$$

   with $P(m,n)$ the number of occurrence of grey level m and n within a given offset $(\Delta x, \Delta y)$ and a given orientation Θ. The sum in the denominator represents the total number of grey level pairs (m,n) within the offset.

   In order to compute textural feature from the co-occurrence matrix, we calculate the co-occurrence matrix for four orientations (0, 45, 90 and 135 degrees) with an offset of 60 pixels, i.e. large enough to include two epidermal cells forming the epidermal honeycomb pattern as their diameter varies from 15 μm to 35 μm. The feature are averaged over the four orientations.

   The contrast feature is the measure of the intensity contrast between a pixel and its neighbors. The energy feature is the sum of squared elements in the co-occurrence matrix. The homogeneity feature measure the closeness of the distribution of elements in the co-occurrence matrix to the diagonal matrix.

3. Gabor response filter output;

   A Gabor filter can be viewed as a complex sinusoidal wave of particular frequency and orientation, modulated by a Gaussian kernel [105, 106]. The complex Gabor function is defined as follow: $g(x,y) = s(x,y)w_r(x,y)$ where $s(x,y)$ is a complex exponential and $w_r(x,y)$ is a 2D Gaussian function. A bank of Gabor filters is generated by varying the spatial frequency and the orientation of the sinusoid.

   A 2D Gabor filter can be expressed in the spatial domain as follow:

$$g(x,y) = w_r(x,y)\exp(-j2\pi(u_0x + v_0y)) \tag{5.6}$$

   where w_r is a 2D Gaussian-shaped function, while u_0 and v_0 define the Gabor filter's spatial frequency in Cartesian coordinates.

These filters are then convolved with the image to obtain the Gabor feature image $r(x, y)$.

When a Gabor filter is applied to an image, it yields its highest response at edges and at points where texture changes. Therefore, they have been widely applied for texture characterization [107]. We compute a bank of Gabor filters with 4 levels of frequency and 4 orientation, with a 32×32 Gaussian kernel.

As features, we use the local energy, i.e. the sum of the squared values of the Gabor feature image $r(x, y)$, and the mean amplitude of the response, i.e. the sum of the absolute value of $r(x, y)$.

4. and Laplacian features;

   The 2D Laplacian L of an image I is given by:

$$L = \frac{\partial I^2}{\partial x} + \frac{\partial I^2}{\partial y} \tag{5.7}$$

   In practice, it can be computed using a convolution filter. We use the classical kernel: $D_{xy} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

   The Laplacian is related to the curvature of intensity changes. In classical edge detection theory [108], the zero-crossings of the Laplacian indicate contour locations. High values in the Laplacian are also associated with rapid intensity changes. The DEJ is an amorphous area compared to the epidermis, which appears as a honeycomb pattern, and the dermis, which contains collagen fibers. Thus, we expect low values in the Laplacian variance in confocal sections around the DEJ location.

   We propose new features to estimate the distance of the current pixel to the DEJ.

   For a pixel i at a given confocal section p, we calculate the Laplacian variance for every confocal section at its location within a 50×50 pixel window. A feature vector is computed containing the Laplacian variance at pixel i coordinates for all depth. The pixel i is characterized by its distance along the z axis to its closest minimum in the feature vector.

   We add to the set of features: the Laplacian variance of pixel i, its distance to its closest minimum as described above and the features (Laplacian variance and depth) of its closest minimum within the Laplacian feature vector. An example is presented in Fig. 5.3.

These features were chosen for their ability to discriminate texture from blurry patterns, which mostly corresponds to the DEJ pattern. The ringed DEJ pattern can be identified via its strong contrast and specific spatial arrangement. A summary of the proposed features is presented in Tab. 5.1.

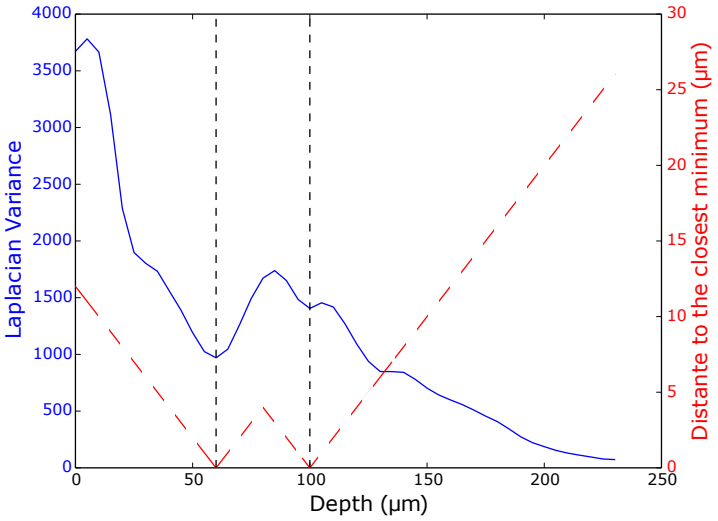


Figure 5.3: Laplacian variance and distance to the closest minimum for pixel. The blue line represents the Laplacian variance at coordinates (i, j) at all depth. The red dashed line corresponds to the distance to the closest minimum.

Table 5.1: Set of the 52 feature used to produce the node potentials.

Parameter	Number of features	Presumed use
Statistics	5	Low intensity of the blurry DEJ pattern
Gray-level co-occurrence matrix	12	Epidermal pattern, ringed DEJ pattern
Gabor output filter	32	Blurry pattern of the DEJ, epidermal and dermal contrasted pattern
Laplacian features	3	Blurry pattern of the DEJ and its location in depth

5.3.2 Interaction potential

The interaction potential describes how likely x_i is to take the value c given the label c' of one of its neighboring pixel j:

$$\psi_{ij}(x_i = c, x_j = c', \mathbf{y}) = p(x_i = c \mid x_j = c') \tag{5.8}$$

Prior information on skin structure is essential to determine efficiently the interaction potentials in our CRF model. The interaction potentials are modeled by 3×3 matrices representing the transition probabilities between classes.

We define two types of transitions: the transitions within a layer as H_d, which are symmetrical and depth-dependent and the inter layers transitions V_{+d} and V_{-d}, which are directional and also depth-dependent.

The product of the pairwise interaction potentials $\psi_{ij}(x_i, x_j, \mathbf{y})$ is expressed as:

$$\begin{aligned}\prod_{i,j\in E} \psi_{ij}(x_i = c, x_j = c', \mathbf{y}) = \prod_{d\in D} \prod_{i,j\in E_d} & \mathbb{1}_{x_i=c} \cdot H_d \cdot \mathbb{1}^{\top}_{x_j=c'} \\ \circ \prod_{i,j\in E_{d\to d+1}} & \mathbb{1}_{x_i=c} \cdot V+_d \cdot \mathbb{1}^{\top}_{x_j=c'} \\ \circ \prod_{i,j\in E_{d\to d-1}} & \mathbb{1}_{x_i=c} \cdot V-_d \cdot \mathbb{1}^{\top}_{x_j=c'}\end{aligned} \tag{5.9}$$

The within-layer interaction potential models the behavior of the skin at a given depth. We know that several skin layers can co-exist in a single confocal section. In a confocal section, edges are modeled symmetrically, i.e. $\psi_{ij} = \psi_{ji}$ (dashed lines in Fig 5.2). The within-layer interaction potential H_d is then a symmetrical matrix. The horizontal transition matrix H_d at depth d is defined in Tab. 5.2.

Table 5.2: Horizontal transition matrix with neighboring pixels i and j where $a, \ldots, f$ are the learned probabilities of transition. The symmetric, non-zero values ensure that transitions in both ways are equally possible.

H_d =

$label_i$ \ $label_j$	epidermis	uncertain	dermis
epidermis	a_d	b_d	c_d
uncertain	b_d	d_d	e_d
dermis	c_d	e_d	f_d

The skin layers follow a specific order from the surface to inner layers: the epidermis is the top skin layer, followed by the uncertain area (containing the DEJ) and the dermis. We define an *incoherent transition* as a transition not following this specific order. Between confocal sections, only coherent transitions are allowed, the edge potentials thus depend of their direction, i.e. $\psi_{ij} \neq \psi_{ji}$.

To impose the biological transition order in depth, constraints are added to the transition matrix according to the edge direction. $V+_d$ is defined in Tab. 5.3 as the vertical transition matrix from pixels i to j, with i above j. The reverse transition matrix $V-_d$ from j to i is defined in Tab. 5.4.

Table 5.3: Vertical transition matrix with i above j where $m, \ldots, s$ are the learned probabilities of transition. The null values ensure that incoherent transitions are impossible.

$V+_d =$

i \ j	Epidermis	Uncertain	Dermis
Epidermis	m_d	n_d	p_d
Uncertain	0	q_d	r_d
Dermis	0	0	1

Table 5.4: Vertical transition matrix with j below i where $m, \ldots, s$ are the learned probabilities of transition. The null values ensure that incoherent transitions are impossible.

$V-_d =$

j \ i	Epidermis	Uncertain	Dermis
Epidermis	1	0	0
Uncertain	n_d	q_d	0
Dermis	p_d	r_d	m_d

5.3.3 Parameter optimization

The parameters described above are learned from the ground-truth dataset. We need to optimize the parameters to improve the classification robustness.

Our model can be expressed as:

$$\log(\, p(\mathbf{x} \,|\, \mathbf{y})\,) = \sum_{i \in M} \log(\varphi_i(x_i, \mathbf{y})) + \sum_{(i,j) \in E} \log(\psi_{ij}(x_i, x_j)) \tag{5.10}$$

We define Ω_d the set of parameters at depth d and $\Omega = \bigcup_{d \in D} \Omega_d$ the set of parameters for all depths. The set of parameters is summarized in Tab. 5.5.

Our goal is to find the set of parameters $\overline{\Omega}$ that maximizes the log-likelihood:

$$\overline{\Omega} = \underset{\Omega}{\operatorname{argmax}} \sum_{i \in V} \log\,(\, p(\, x_i = c_i \,|\, \mathbf{y}\,)\,) \tag{5.11}$$

with c_i the class of i.

The parameters θ_d, H_d, $V+_d$ and $V-_d$ are depth dependent. For each depth d, we estimate the parameters of Ω_d, which leads us to 300 parameters for 20 depths. The parameters Ω are optimized using the Powell search method, an iterative optimization algorithm [109] that does not require estimating the gradient of the objective function.

Table 5.5: Set of parameters of the CRF model.

Parameter	Form	Number of parameters	Use
θ_d	1×3 vector	3	Weight bias vector used to balance the labels occurrences
H_d	3×3 symmetrical matrix	6	Matrix transition between classes at depth d
$V+_d$	3×3 upper triangular matrix	6	Matrix transition between classes between d and $d+1$

We use a Loopy Belief Propagation (LBP) method to estimate $p(\mathbf{x} \mid \mathbf{y})$. The computation is performed using the library developed in [110].

The value of θ_d is initialized at $[0.3, 0.3, 0.3]$ for all depths and optimized to increase the model accuracy. The initial values of H_d and $V+_d$ are estimated from the frequency of co-occurrence of classes (c, c') between neighboring pixels i and j in the ground-truth images. Co-occurrence frequencies are estimated at each depth $d \in D$ of confocal sections.

5.4 Experimentation

5.4.1 Database

Our dataset consists of 23 annotated stacks of confocal images acquired from fifteen healthy volunteers with fair skin, see Chap. 3. Image acquisition was carried out on the cheek to further assess chronological aging. No cosmetic products nor skin treatment were allowed on the day of the acquisitions. Appropriate consent was obtained from all subjects before imaging. RCM images were acquired using a near-infra-red reflectance confocal laser scanning microscope (Vivascope 1500; Lucid Inc, Rochester, NY, USA) [29]. On each imaged site, stacks were acquired from the skin surface to the reticular dermis with a step of $5\,\mu$m. Visual labeling of the DEJ is not easy to perform, therefore an expert was asked to delineate the stacks in 3 zones: epidermis (E), uncertain (U) and dermis (D). We segmented confocal images between depths $20\,\mu$m and $150\,\mu$m, the images above $20\,\mu$m belonging to the epidermis with high confidence. We used a 10-fold cross validation test to evaluate the segmentation results.

5.4.2 Feature evaluation

To evaluate our proposed set of features used in the Random Forest classification, we compare the mean accuracy of our classification results to the state-of-the-art methods.

The mean accuracies of the RF classifications are presented in Tab. 5.6. In Kurugol et al. [21], the authors propose a SVM classification method using features similar to ours. They achieved 64%, 41%, and 75% of correct classification of tiles for epidermis, transition region, and dermis, respectively. Hames et al. [85] used a Random Forest classifier followed by a regularization with a 1D CRF. They achieved 82%, 78% and 88% of correct classification for the Epidermis, DEJ and Dermis respectively. With the proposed set of features, we are able to achieve 93%, 56% and 75% accuracy respectively for the epidermal, uncertain and dermal classification. These results suggest that our set of features is relevant to identify the three skin labels according to the experts' visual inspection. However, the result of our initial classification still contains 11% of incoherent transitions (not following the expected biological order), see Tab. 5.8, motivating the introduction of spatial constraints with the CRF regularization.

Table 5.6: Results for the unregularized experiments. Mean accuracy of the RF classifications of the three labels. The RF classification provides the node potentials for the CRF model.

	Epidermis	Uncertain	Dermis	Number of RCM stacks
Proposed features	**0.93**	0.56	0.75	23
Kurugol et al. [21]	0.64	0.41	0.75	15
Hames et al. [23]	0.82	**0.78**	**0.88**	308

5.4.3 CRF parameters evaluation

To evaluate the chosen parameters and the relevance of our proposed CRF model, we compare three cases:

1. CRF_{2D}, the CRF with only horizontal regularization. Each confocal section is regularized independently using the horizontal transition matrix H_d;

2. $\text{CRF}_{\text{3D Sym}}$, the CRF with horizontal regularization H_d and symmetrical vertical regularization, i.e. $V{+}_d = V{-}_d$;

3. $\text{CRF}_{\text{3D Asym}}$, the CRF with horizontal regularization H_d and asymmetrical vertical regularization, i.e. $V{+}_d \neq V{-}_d$. This is our proposed model where the skin layers order is imposed.

We will first comment the optimization results for the three cases. Then we will discuss the incoherent transition percentages for the three cases and their sensitivity and specificity results after the optimization of the parameter Ω.

Fig. 5.4 shows the convergence behavior of the Powell method for training the parameter Ω for the three cases. It shows that the procedure converges more slowly for $CRF_{3D\ Sym}$ and $CRF_{3D\ ASym}$ due to the addition of vertical constrains. After convergence, a larger value of the objective function is achieved with the CRF_{2D} case but the resulting segmentation with CRF_{2D} contains 17% of incoherent transitions (transitions not following the biological order) which is even higher than the direct segmentation after the Random Forest classification without any regularization.

The segmentation results for the $CRF_{3D\ ASym}$ experiment before and after the optimization are presented Tab. 5.7. Without parameter optimization, we obtain a very high sensitivity for the epidermal and dermal classification (>0.80%) but a rather low sensitivity for the uncertain area. The algorithm tend to under-estimate the uncertain area. The optimization of the parameter Ω improves the sensitivity of the uncertain area classification from 66% to 70% and the epidermal and dermal specificity up to 96%.

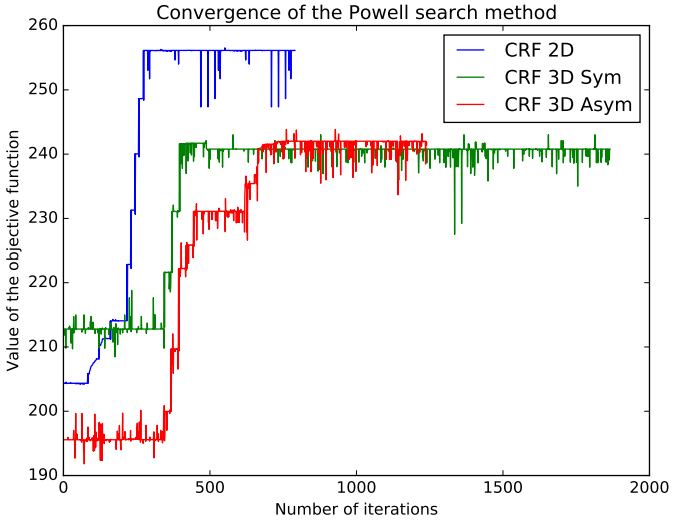


Figure 5.4: Convergence of the Powell search method.

	Sensitivity			Specificity		
	E	U	D	E	U	D
Before Optimization	0.96	0.66	0.85	0.83	0.93	0.85
After Optimization	0.91	0.70	0.93	0.96	0.92	0.96

Table 5.7: $CRF_{3D\ ASym}$ classification results with and without the parameter estimation using Powell's method.

The resulting segmentation with CRF_{2D} contains 17% of incoherent transitions. $CRF_{3D\ Sym}$ still contains 4% of incoherent transition while $CRF_{3D\ ASym}$ contains none. The incoherent transitions percentage are presented in Tab. 5.8.

Table 5.8: Percentage of incoherent transitions between the skin layers. Epidermis: E, Uncertain: U, Dermis: D.

	U → E	D → U	D → E	Total
Random Forest	0.06	0.02	0.03	0.11
RF + CRF_{2D}	0.07	0.03	0.06	0.17
RF + $CRF_{3D\ Sym}$	0	0.04	0	0.04
RF + $CRF_{3D\ Asym}$	0	0	0	0

The global accuracy, presented in Tab. 5.9, for the three experiments are similar. We also achieve a high specificity for the three classes (above 0.90%), see Tab 5.10. The $CRF_{3D\ Asym}$ parametrization had the effect of increasing the sensitivity of the uncertain area classification, compared to CRF_{2D} and $CRF_{3D\ Sym}$, while maintaining the epidermal and dermal sensitivity above 0.90%, see Tab. 5.11.

Table 5.9: Global accuracy percentage for the three regularization schemes.

	Accuracy
RF	0.79
RF + CRF_{2D}	0.87
RF + CRF_{3DSym}	0.85
RF + CRF_{3DAsym}	0.87

Table 5.10: Specificity of the three labeling in the regularized cases.

	Specificity		
	Epidermis	Uncertain	Dermis
CRF_{2D}	0.95	0.92	0.92
$CRF_{3D\ Sym}$	0.94	0.92	0.96
$CRF_{3D\ Asym}$	0.96	0.92	0.96

Table 5.11: Sensitivity of the three labeling in the regularized cases.

	Sensitivity		
	Epidermis	Uncertain	Dermis
CRF_{2D}	0.92	0.63	0.92
$CRF_{3D\ Sym}$	0.91	0.56	0.90
$CRF_{3D\ Asym}$	0.91	0.70	0.93

Examples of resulting segmentations at two following depths are presented in Fig. 5.5 and Fig. 5.6. A direct classification after features calculation leads to misclassifications and incoherent transitions between classes. The positive impact of our CRF model is noticeable.

5.4.4 Comparison to state-of-the-art methods

We compare our results to state-of-the-art methods. The results presented below should be considered with caution, because of the differences in classification level and dataset size, especially with the methods that perform an image classification rather than a pixel one.

Global accuracy of our model is comparable to state-of-the-art methods. The sensitivity and specificity results of the regularized CRF model are presented in Tab. 5.12 and Tab. 5.13. The specificity results obtained by Kurugol et al. [21] are not available.

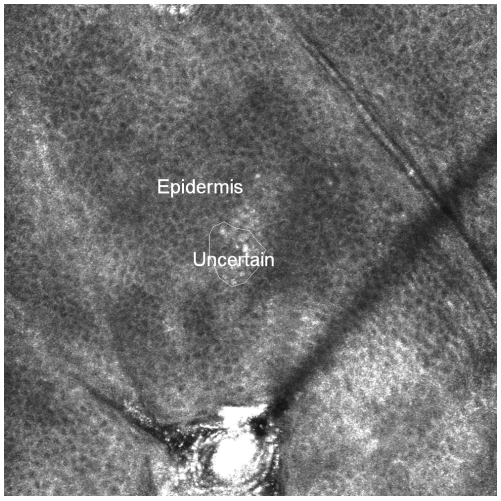
Deep learning methods seem to take into accounts the dependencies between images to perform the classification, but the regression might lack of interpretability.

Table 5.12: Sensitivity results of $CRF_{3D\ Asym}$ compared to state of the art methods.

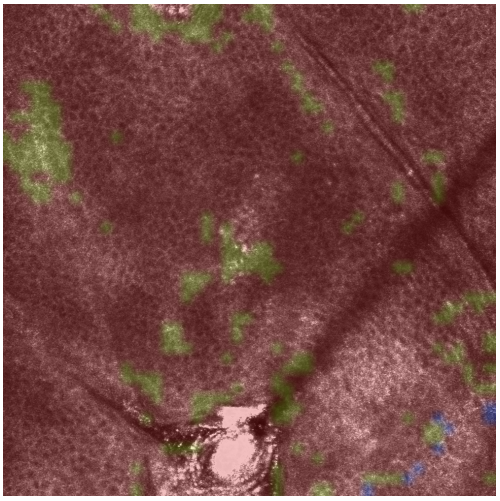
	Sensitivity			Goundtruth level	Number of RCM stacks
	Epidermis	Uncertain	Dermis		
$CRF_{3D\ Asym}$	0.91	0.70	**0.93**	Pixel	23
Kurugol et al. [21]	0.64	0.41	0.75	Pixel	15
Hames et al. [23]	0.87	0.79	**0.88**	Image	308
Bozkurt et al. [89]	**0.93**	0.83	0.84	Image	504
Kaur et al. [88]	0.74	0.51	0.68	Image	15

Table 5.13: Specificity results of $CRF_{3D\ Asym}$ compared to state of the art methods

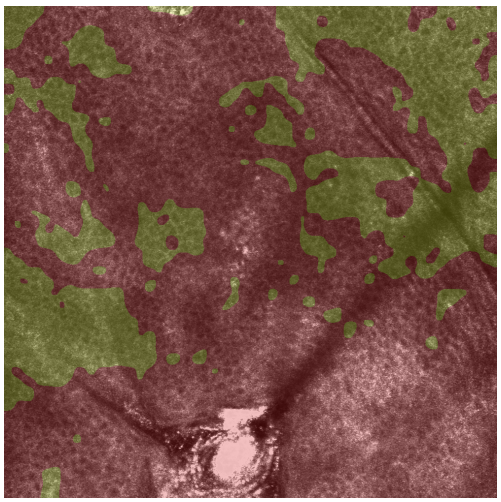
	Specificity			Goundtruth level	Number of RCM stacks
	Epidermis	Uncertain	Dermis		
$CRF_{3D\ Asym}$	**0.96**	**0.92**	**0.96**	Pixel	23
Hames et al. [23]	0.94	0.87	0.94	Image	308
Bozkurt et al. [89]	0.95	0.90	0.95	Image	504
Kaur et al. [88]	0.86	0.75	0.85	Image	15



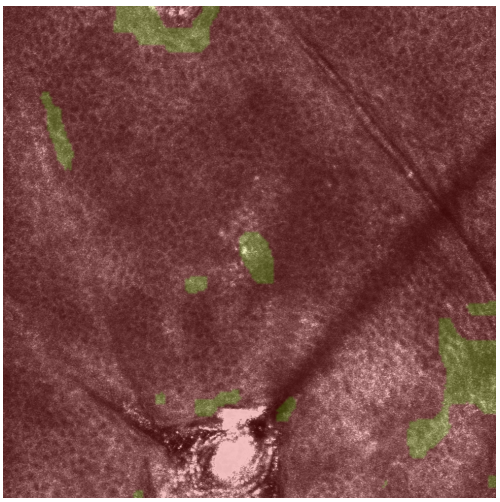
(a) Annotated image at depth d



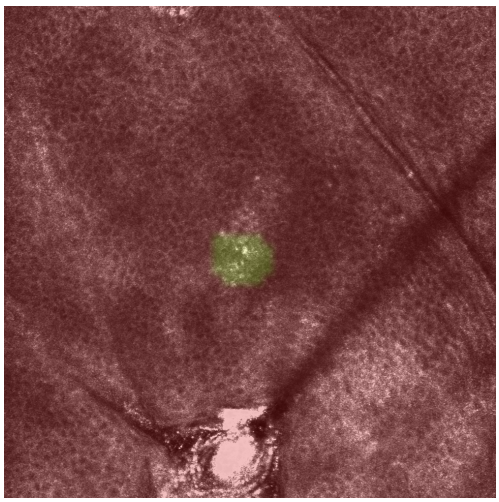
(b) RF at depth d



(c) CRF_{2D} at depth d

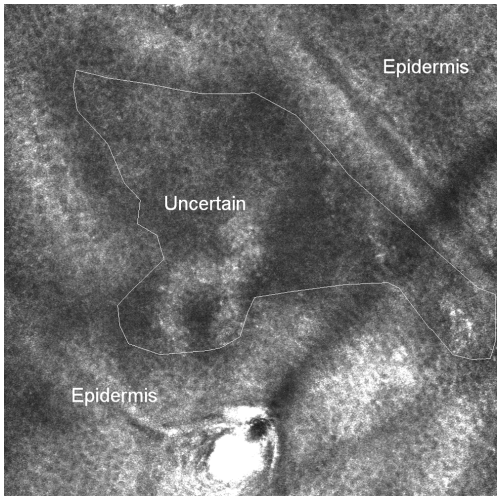


(d) $\text{CRF}_{\text{3D Sym}}$ at depth d

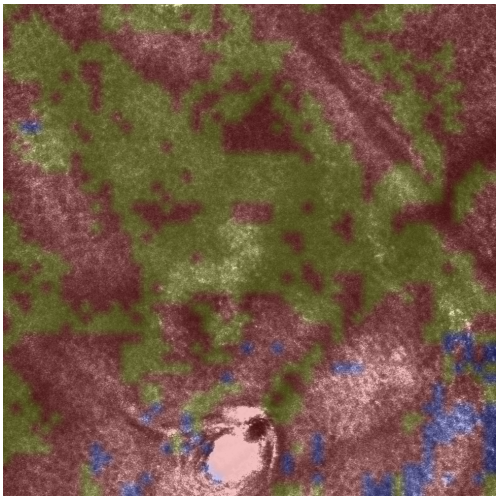


(e) $\text{CRF}_{\text{3D Asym}}$ at depth d

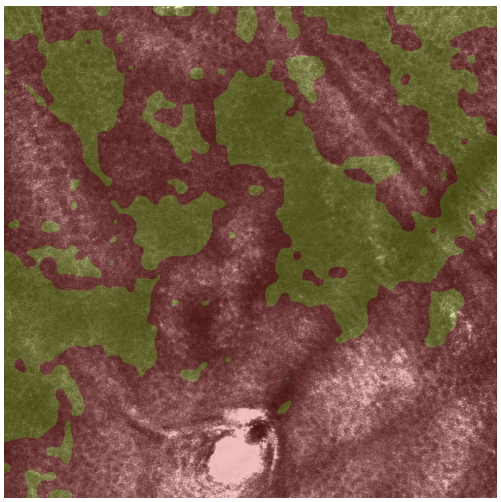
Figure 5.5: Segmentations at depth d: Epidermis (red), Uncertain (yellow), Dermis (blue). The addition of constrains into the CRF model improves the accuracy of the segmentation.



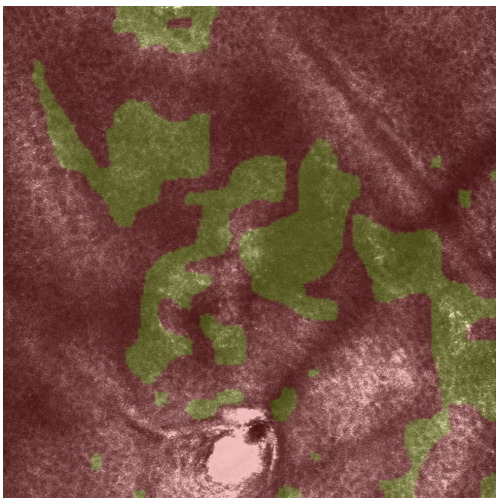
(a) Annotated image at depth $d+1$



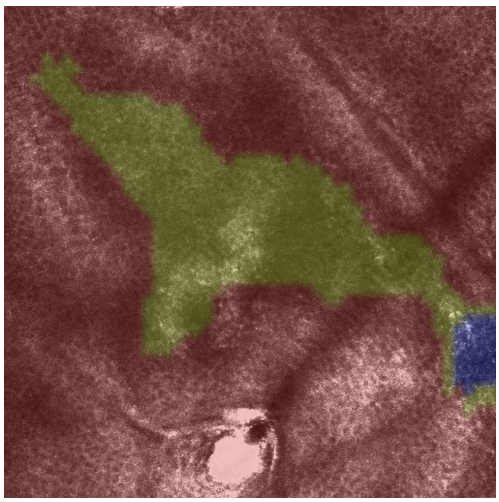
(b) RF classification at depth $d+1$



(c) CRF_{2D} at depth $d+1$



(d) $\text{CRF}_{\text{3D Sym}}$ at depth $d+1$



(e) $\text{CRF}_{\text{3D Asym}}$ at depth $d+1$

Figure 5.6: Segmentations at depth $d+1$. Incoherent transitions exists between depth d and $d+1$. One can notice the misclassification obtained by RF and CRF_{2D}. The use of $\text{CRF}_{\text{3D Asym}}$ provides a coherent segmentation.

5.5 DEJ representation as 3D Surface

Our resulting DEJ segmentation is a thick 3D area.

We propose to extract three relevant 3D surfaces from the DEJ segmentation:

- the transition from the epidermis to the uncertain area, which is the epidermal lower boundary and will be called the *epidermal surface*;
- the median surface of the uncertain area, which corresponds for each pixel to the median distance between the two borders of the uncertain area, and will further be called the *median junction surface*; and
- the transition from the uncertain area to the dermis, which is the dermal upper boundary and will be called *the dermal surface*.

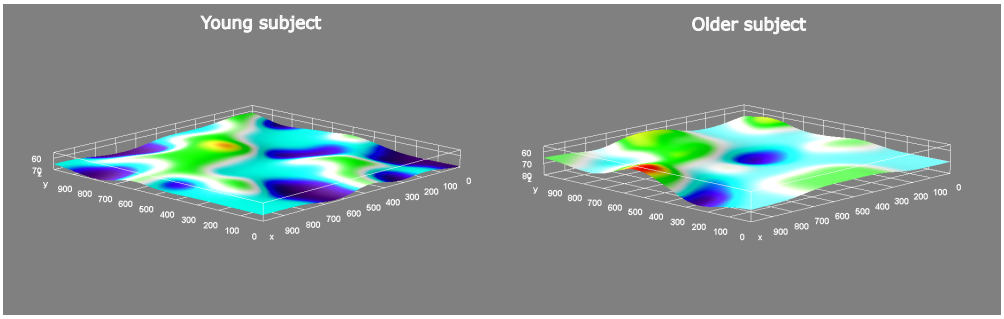
Examples of the 3D surfaces are presented in Fig. 5.7.
Each 3D surface will be used in the following to characterize the DEJ shape.

5.6 Conclusion

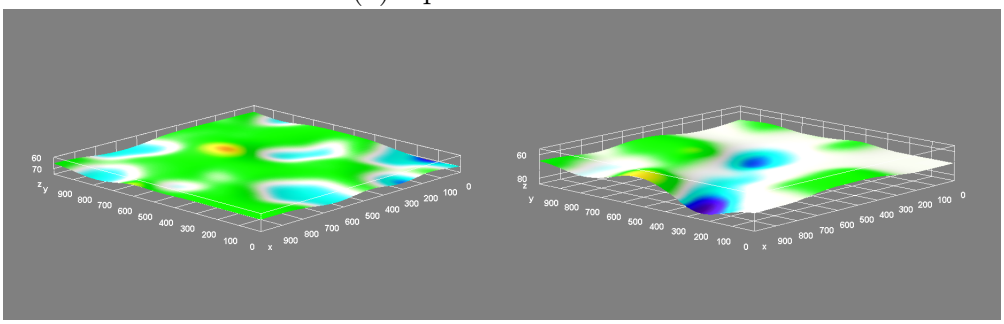
To summarize, our proposed regularization $\text{CRF}_{\text{3D Asym}}$ promotes spatial consistency and forbids incoherent layer transitions which allow us to increase the global accuracy of the segmentation from 79% to 87% while suppressing incoherent labellings. We have shown the positive impact of the use of a 3D CRF model. The 3D regularization also improves the classification results. We have also shown that the inclusion of asymmetrical potentials along the z-axis to model adjacency relations between the skin layers improves the performance of the model. The epidermal and dermal sensitivity are increased. The additional constrains ensure the coherence of the segmentation.

3D visualizations of the DEJ segmentations of young and older subjects, obtained with $\text{CRF}_{\text{3D Asym}}$, are presented in Fig. 5.8. The additional constrains proposed in our model allow to obtain a thick DEJ surface which is biologically coherent.

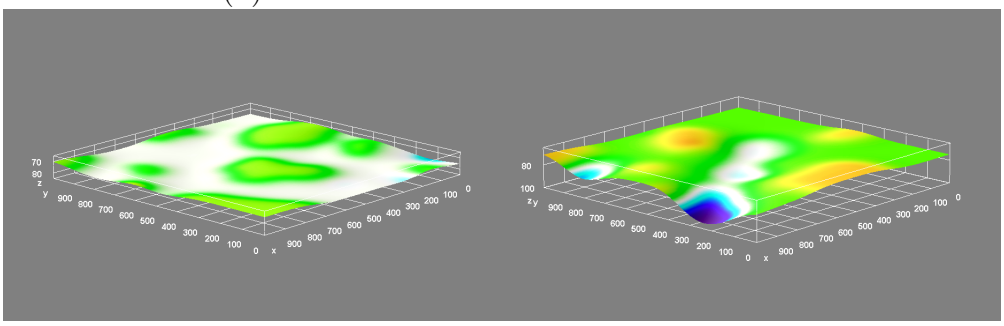
The DEJ segmentation produces a topological surface composed of level-lines which enables us to uses the characterization methods proposed in the next section. In the next chapter, we present our method to characterize the shape of our 3D segmentation.



(a) Epidermal Surface

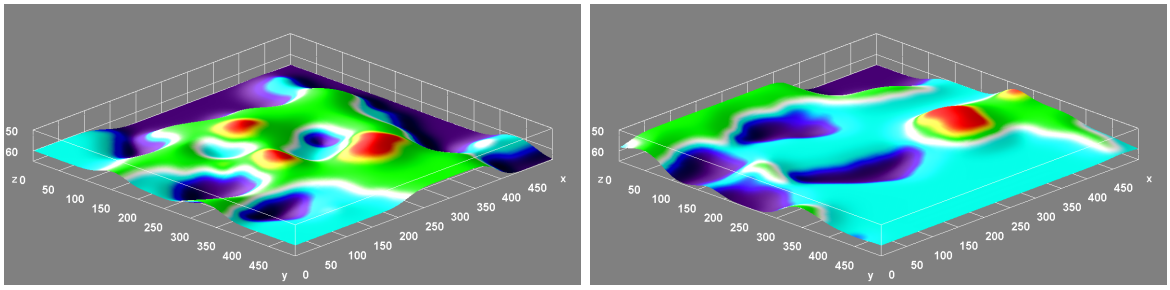


(b) Median surface of the Uncertain area



(c) Dermal Surface

Figure 5.7: Examples of 3D DEJ surfaces: (a) The Epidermis-to-Uncertain surface, which corresponds to the epidermal lower boundary, (b) The Median surface of the Uncertain area and (c) The Uncertain-to-Dermis surface, which corresponds to the dermis upper boundary (c).



(a) Young epidermis (b) Aged epidermis

Figure 5.8: Visual appearance of the lower border of the epidermal layer. Notice that the aged epidermis appears flatter than the young epidermis. This is something to be expected.

6
DEJ shape analysis

Contents

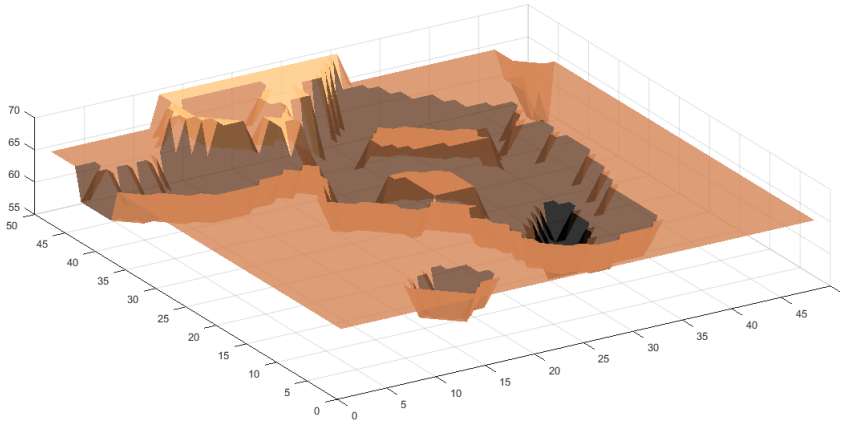
Visual methods to analyze the DEJ shape rely on the characterization of its peaks and valleys. Visual analysis focus on the number of dermal papillae per area [47, 50] and the shape of the dermal papillae [14, 15]. To reproduce visual expertise, Newton et al. [111] classified the skin layers into dermal and epidermal compartments using a Graph-Cut approach. They measured the height and volume of the peaks and valleys starting from the median plane of the DEJ surface. They showed no correlation in these measurements with skin aging. This highlights the need for additional methods to characterize the DEJ shape. Also, the question arises on which DEJ surface should be extracted from the thick segmentation to perform its shape characterization.

In Chap. 5, we have presented the method to segment the skin into three different layers : one containing the epidermis with high confidence, an uncertain area containing the DEJ and a last one containing the dermis with high confidence.

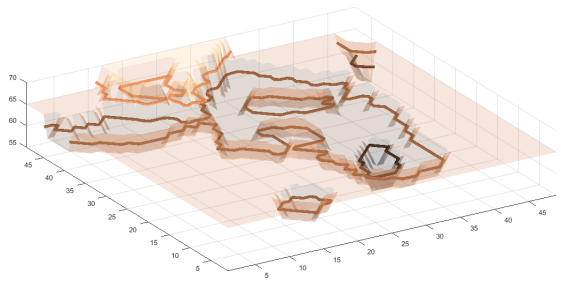
Our resulting segmentation of the 3D DEJ can be viewed as a topographic surface. Each level line is the contour of a region of the surface. The family of level lines of the

image is given a tree structure since they are ordered by inclusion, this representation is called the level-lines tree. An exemple of the DEJ topographic surface and its level lines are presented in Fig. 6.1.

(a) 2D representation of the DEJ segmentation



(b) 3D representation of the DEJ segmentation



(c) Family of level-lines

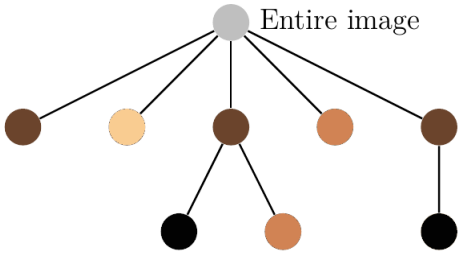


Figure 6.1: Example of the DEJ segmentation as a topographic surface. In (a), the DEJ segmentation is presented in 3D. Each level line is the contour of a level set of the surface (b). In (c), inclusion relationship between the level-lines is presented as a tree structure.

This representation as a tree structure enables a region-based analysis, rather than a pixel-based one, through the notion of connected operators which have been discussed in Chap. 4. They are region-based filtering tools that act by merging flat zones. Tree based connected operators consist of filtering the tree, i.e. removing some of its nodes, based on an attribute function characterizing the regions of the tree. The tree-based image representations popularized by the connected operators have been proved to be useful for many applications, such as image segmentation [63], image simplification [58, 112–114] or image classification [115, 116].

For instance, Salembier and Wilkinson [56] have proposed a set of filtering strategies that simplify images using a tree representation. The nodes of the tree that do not meet an attribute criterion are removed from the tree, the corresponding flat zones are filtered from the original image. A filtering strategy has also been used in [113] with some shape

attributes to simplify images so that the object of interest is enhanced. The node selection can also be performed by energy minimization to find the optimal cut [117].

In [116], the authors have proposed a multiscale morphological method for pattern-based analysis and classification of gray-scale image using connected operators relying on a tree representation. Pattern spectrum [118] are commonly used for image analysis and classification. They can be computed using a technique known as granulometry. Intuitively, a granulometry can be considered as a set of filters with an increasing criterion. Morphological and attribute profiles [115, 119] are widely used for the classification of high-resolution images. Attribute profiles [120] are multilevel decomposition of images obtained with a sequence of transformations performed by connected operators. They are obtained by the sequential application of a morphological attribute filter with an increasing criterion $T = \{T_\lambda : \lambda = 0, \ldots, n\}$. The main idea is to extract information on the tree structure during the filtering of the nodes.

The objective of this section is to provide a method able to characterize the DEJ shape and to discriminate between age groups. Our contributions lie in the representation of the DEJ as a level line tree and in its processing using two approaches:

- the extraction of topological descriptors of the tree and local features on relevant nodes;
- an attribute profile of the tree. The choice of the attribute function involved in the tree filtering is driven by the dermatologists' analysis. We propose the definition of a new measure of the tree structure: the surface area attribute.

Our proposed methods are able to discriminate the two age groups. They also agree with the dermatologists' analyses.

In this chapter, we first review the tree based image representation as a level-line tree. Then, we present our results with respect to skin aging skin aging when analyzing descriptors of the level-lines tree. And finally, we present the attribute profile of the DEJ surface area according to a compactness criterion.

6.1 Level line tree

The consideration of levels lines, i.e. topological boundaries of the connected components given by the upper or lower level sets, leads to a tree representation describing images. This representation is unique with respect to any image. This tree is called level-line tree [121], or tree of shapes [122].

An example of the level line tree representation is presented in Fig. 6.2.

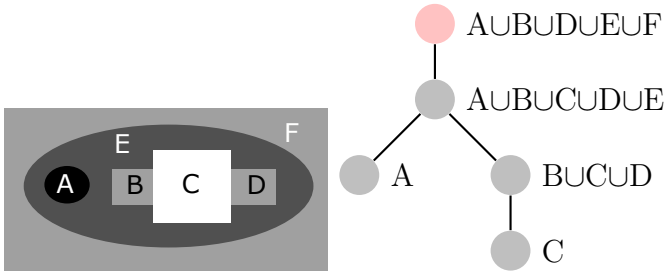


Figure 6.2: Tree-based image representations. The root of the tree is at the top represented by a red circle.

The two sets of connected components of upper and lower level sets can be merged using the notion of shape. A *shape* is a connected component of a upper or lower level with its holes filled.

Each node of the graph represent a connected component of the image f with its holes filled and each edge models the inclusion relationship between the connected component of f.

The level line tree is a unique representation of the image modeling the inclusion relations of its level lines. The leaves of the tree represent the regional maxima and minima. The bright and dark objects are treated in the same way. It is invariant to change in contrast and in scale.

The first efficient algorithm to compute the level line, or tree of shapes, was proposed by Monasse and Guichard in [122]. This algorithm first builds the min and max-trees. For each hole in a connected component in one tree, it finds the unique corresponding connected component in the other tree. The connected component is then put as a child of the one containing the hole. Song [121] proposed a top-down approach to compute the topographic map. Recently, Géraud et al. [123] proposed a quasi-linear algorithm to compute the topographic map using the Union-Find algorithm for efficiently finding equivalence classes.

6.2 Level line tree attributes

The aim of the section is the calculation of level-line tree node descriptors, in order to assess if they are capable of measuring the skin aging process.

The DEJ shape can be characterized by extracting topological descriptors of its corresponding level tree and local features on relevant nodes.

When performing a visual analysis on the DEJ shape, dermatologists focus on the dermal papillae appearance. The dermal papillae correspond to the peaks and valleys of the DEJ, i.e. the leaves in the level-line tree representation of the DEJ. We assume that the leaves of the level line tree are relevant nodes for the DEJ analysis.

For each node of the tree, we can assign an attribute value A that characterizes the region, i.e. shape represented by that node.

6. DEJ shape analysis

We define the *contrast* of a node as the absolute difference between its level and the level of its parent.

Topological descriptors of the tree and local features on the leaves of the tree are computed.

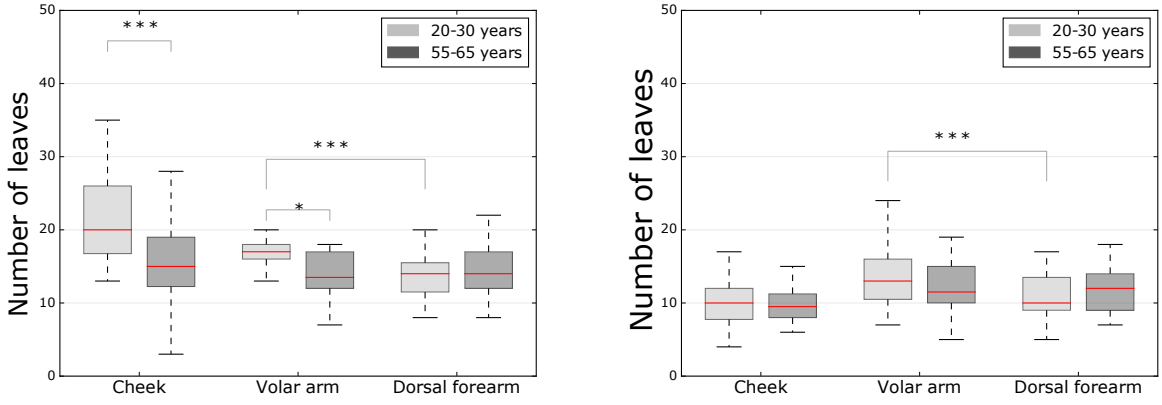
- Number of leaves, corresponding to the number of peaks and valleys of the DEJ.
- Depth of the leaves, corresponding to the number of nodes between the leaves and the root of the tree.
- Volume of the leaves, corresponding to the value: area $\times$ contrast with the area being the number of pixels composing a shape with its holes filled.
- Compactness of the leaves, corresponding to the value: $4\pi \times (\text{area}/\text{perimeter}^2)$.
- Complexity of the leaves, corresponding to the value: (perimeter/area).
- Elongation of the leaves, corresponding to the value: λ_1/λ_1 with λ_1 and λ_2 the eigenvalues of the inertia matrix of the region.

Except for the number of leaves, each tree attribute is the average of the value of all the leaves.

For each attribute, box-and-whisker plots are presented for each of the surfaces (epidermal, median junction and dermal surface). The box-and-whisker plots represent the data distributions through their quartiles. The vertical lines indicate the variability outside the upper and lower quartiles. Means of distributions are represented as horizontal lines inside the boxes. We conduct measurements on the same database described in Chap. 4 consisting in 15 subjects assigned to two age groups: a 7-persons group aged from 18 to 25 and another 8-persons group aged from 55 to 65. The clinical validation results will be further presented in Chap. 7.

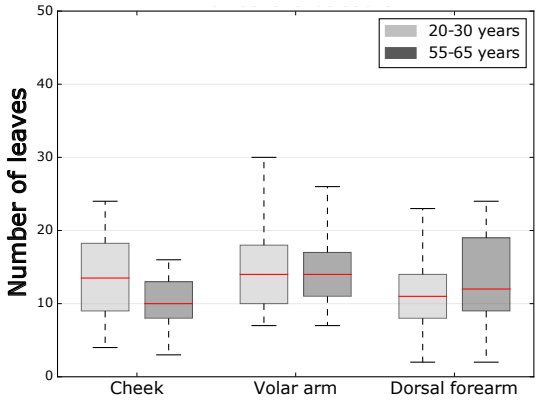
P-values lower than 0.05 are considered significant. The statistics significance is defined as follow:

- * : $0.01 < \textit{P-values} \leq 0.05$;
- ** : $0.001 < \textit{P-values} \leq 0.01$;
- *** : $\textit{P-values} \leq 0.001$.



(a) Epidermal Surface

(b) Dermal Surface



(c) Median Surface of the uncertain area

Figure 6.3: Number of leaves of (a) the Epidermis-to-Uncertain surface, (b) the Uncertain-to-Dermis surface and (c) the Median surface of the Uncertain area.

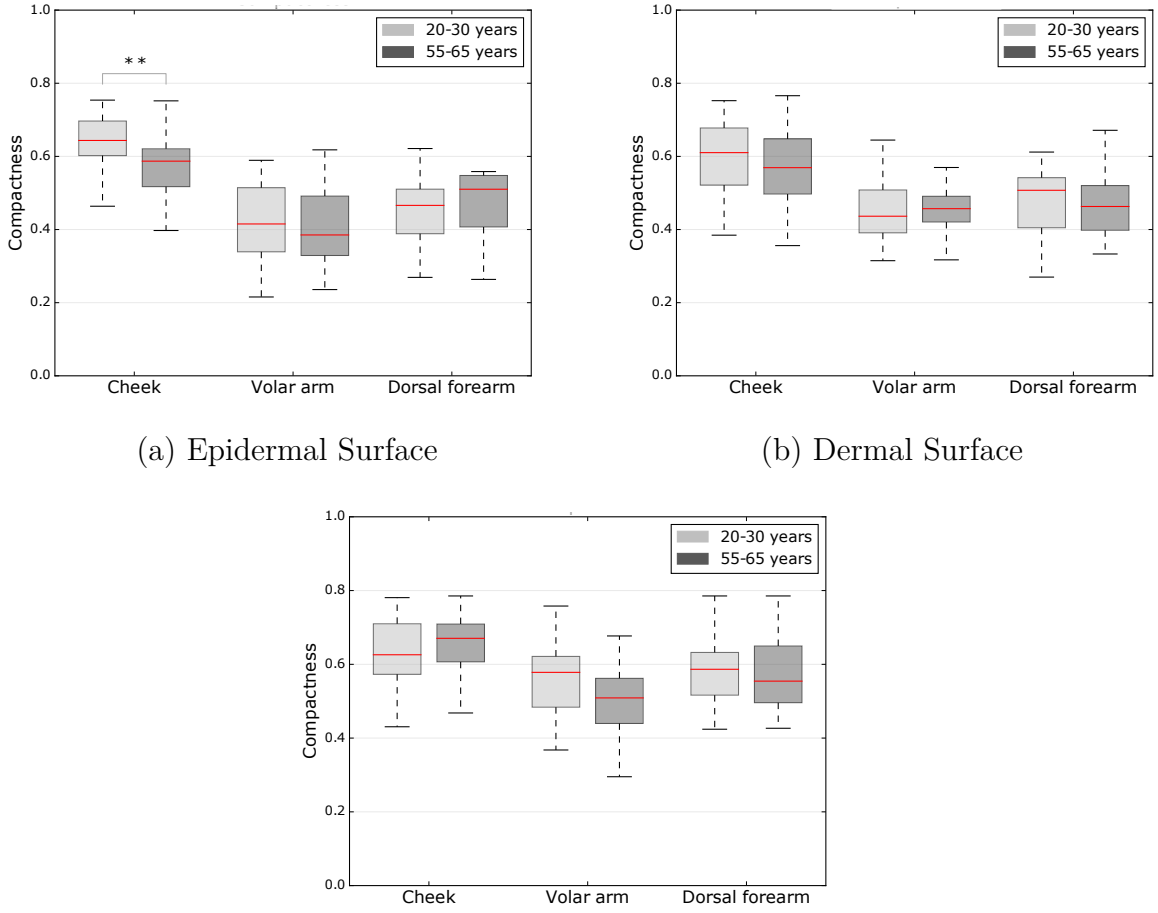
The elongation attribute and mean depth of the tree show no difference between age groups. Several attributes (number of leaves, volume and compactness) show significant differences toward chronological aging and photo-aging.

The number of leaves is significantly higher in the younger group (21.9 ± 6.8) than in the older group (15.2 ± 5.4) on the cheek and on the volar arm (16.6 ± 2.7 vs. 14.0 ± 3.1) for the epidermal surface (Fig. 6.3 (a)).

The compactness attribute is higher in the younger group (0.64 ± 0.07) than in the older group (0.56 ± 0.1) on the cheek also for the epidermal surface analysis (Fig. 6.4 (a)).

The volume attribute is higher in the younger group than in the older group on the cheek for the epidermal surface analysis (Fig. 6.5 (a)).

The complexity attribute is significantly higher in the older group than in the younger group on the Volar arm (photo-protected) for the median junction surface (Fig. 6.6 (c)).



(a) Epidermal Surface

(b) Dermal Surface

(c) Median Surface of the uncertain area

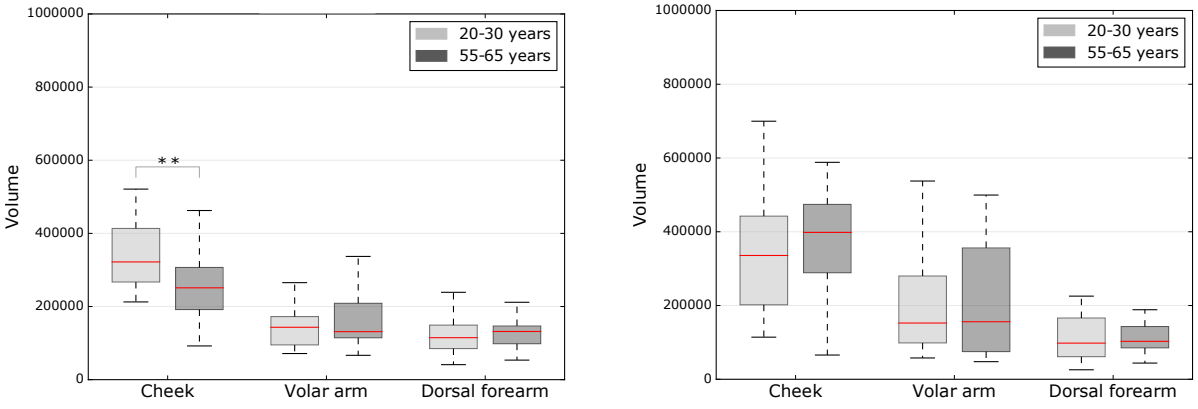
Figure 6.4: Compactness of the peaks and valleys of (a) the Epidermis-to-Uncertain surface, (b) the Uncertain-to-Dermis surface and (c) the Median surface of the Uncertain area.

We do not expect such difference. Even if the statistical analysis shows the significance of this result, we can observe a large standard deviation for both populations.

When comparing the Volar arm (photo-protected) and the dorsal foreram (photo-exposed), the photo-aging effect is quantified among the young subject with the decrease of the number of leaves from 16.4 ± 2.7 to 13.7 ± 3.6 for the epidermal surface and from 13.1 ± 3.7 to 10.9 ± 3.3 for the dermal surface (Fig. 6.3 (a,b)). The complexity is significantly lower on the photo-protected area than the photo-exposed area for both populations for the epidermal surface.

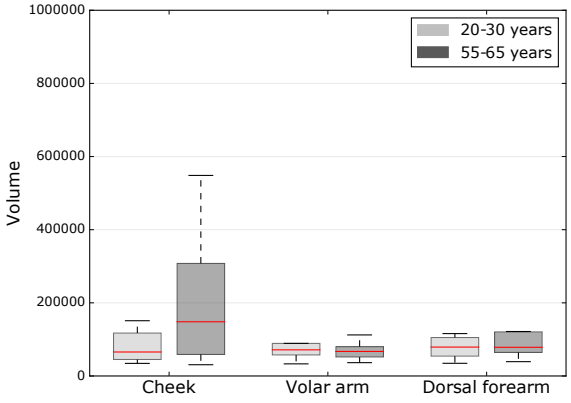
6.2.1 Choice of the DEJ surface

The dermal surface does not enable the measurement of chronological aging.



(a) Epidermal Surface

(b) Dermal Surface



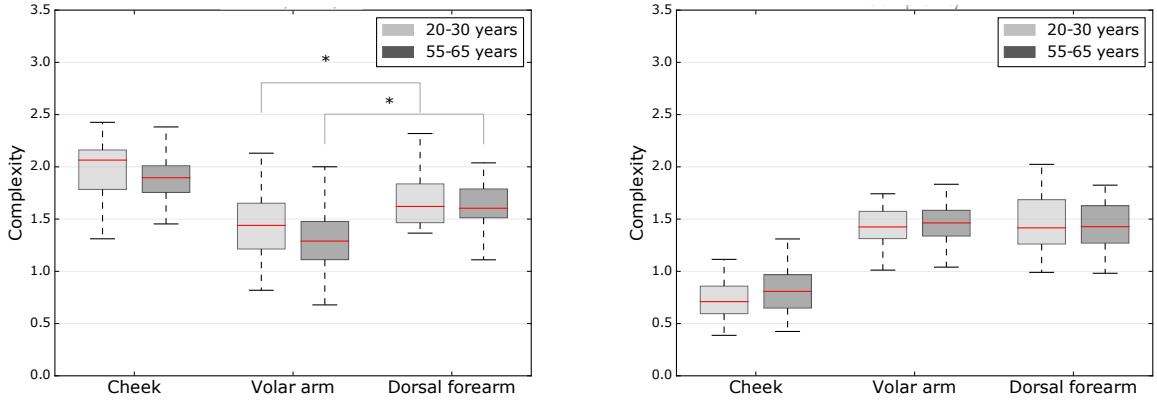
(c) Median Surface of the uncertain area

Figure 6.5: Volume of the peaks and valleys of (a) the Epidermis-to-Uncertain surface, (b) the Uncertain-to-Dermis surface and (c) the Median surface of the Uncertain area.

There is no significant results regarding photo-aging for the median junction surface. The median junction surface may have been smoothed due to thickness of the 3D segmentation, which present small undulations.

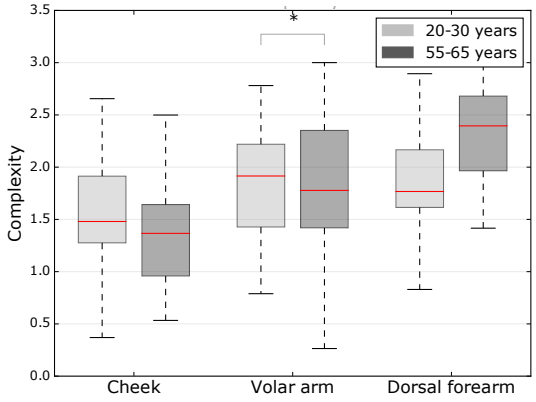
It appears that the epidermal surface is the most suited for the DEJ shape analysis. The transition from the epidermis to the uncertain area (containing the DEJ) is characterized by the transition from the epidermal honeycomb pattern to a blurry pattern. The dermis and the uncertain area may share some common blurry pattern, as the dermal collagen fibers may not appear as contrasted objects, which could bias the shape of the transition surface. In the following, we will focus on the analysis of the DEJ surface obtained from the transition from the epidermis to the uncertain area, i.e. the epidermal surface.

To summarize, we are able to quantify the flattening of the DEJ on the cheek and on the volar arm with age, and among the younger group regarding sun-exposure. An



(a) Epidermal Surface

(b) Dermal Surface



(c) Median Surface of the uncertain area

Figure 6.6: Complexity of the peaks and valleys of (a) the Epidermis-to-Uncertain surface, (b) the Uncertain-to-Dermis surface and (c) the Median surface of the Uncertain area.

interesting result is the loss of compactness of the leaves on the cheek with age. With age, the peaks and valleys of the DEJ become less compact. This result is in line with the visual analysis performed by dermatologists as they focus on whether the dermal papillae of the DEJ have a ringed or poly-cyclic shape.

The use of these attributes allows us to get some information on the DEJ shape.

In the next section, we propose to study the DEJ attribute total surface area using a filtering process with a compactness criterion.

6.3 Attribute Profile

In the last section, we have presented some attributes obtained from the DEJ representation as a level line tree and their evolution with skin aging. We aim to obtain a multilevel measure on the DEJ shape.

A way of obtaining a multilevel characterization of an image is through the use of morphological profiles. Traditional profiles consist of extracting the variations of the gray values during the successive filtering of the image by morphological operators at varying scales [119].

Attribute profiles are obtained by the sequential application of a morphological attribute filter with a criterion $T = \{T_\lambda : \lambda = 0, \ldots, n\}$. The main idea is to quantify the measurement variations (such as those of contrast, perimeter, scale, etc) from the tree during the node filtering to obtain a feature vector.

6.3.1 Choice of the attribute criterion

Two parameters need to be set to perform an attribute profile of the DEJ surface area. First, the criterion involved in the attribute filter needs to be determined. Second, the measurements extracted during the attribute profile needs to be chosen.

Concerning the choice of the filtering criterion, we have shown that the compactness of the peaks and valleys of the DEJ decreases with age. We hypothesize that this result bears some resemblance with with the visual analysis by dermatologists. Thus, the compactness attribute appears to be well-adapted for the filtering of the tree.

Then, we need to determine the measure to be extracted at each filtering step. As the DEJ is a 3D surface, separating the epidermis from the dermis, we aim to compute an attribute in connection with the DEJ anatomy. Among all possible criteria, the surface area is a good candidate to obtain global information on the complexity of the surface as it is directly linked to the number, depth and shape of peaks and valleys in the DEJ.

In the following, we first define the surface area attribute. Then, we describe the tree filtering rules and we finally present the surface area profile of the DEJ.

6.3.2 Definition of the surface area attribute

Here, we define the surface area attribute of the DEJ. We have considered the area of a shape as its number of pixels with its holes filled. We define the surface area S_s of a shape s and $S_{children}$ the surface area of its children. The surface area of s is defined as follow:

$$S_s = S_{children} + area_s - area_{children} + perimeter_s \times contrast_s \tag{6.1}$$

An example of the surface area attribute is presented in Fig. 6.7. The surface area attribute differs from the area attribute in the way that it takes into account the level of a node which, in our case, encodes the location in depth of the corresponding region in the 3D segmentation. The surface area attribute can be computed using Algorithm 1. The surface area attribute is calculated from the leaves up to the root of the tree. First, the product of the perimeter and contrast is calculated from the leaves to the root of the tree. Second, the area of each node (with its holes filled) is added to their corresponding surface area attribute.

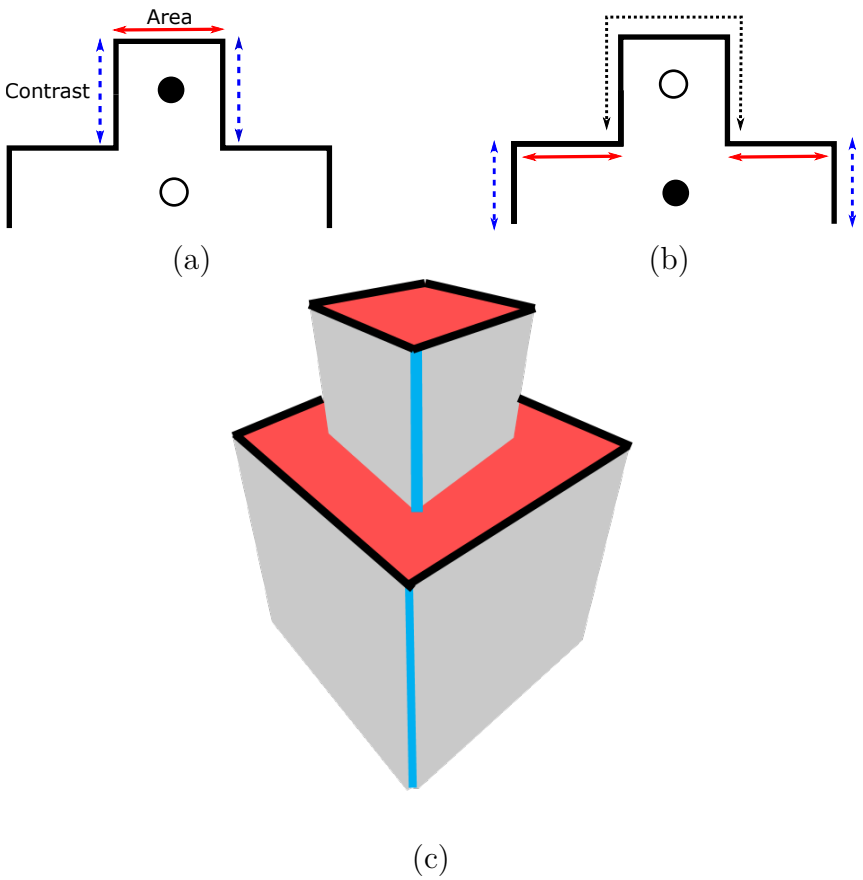


Figure 6.7: Computation of the surface area attribute. The surface area attribute is the sum of the area (red arrows) and the contrast × perimeter (blue arrows). In (a) the attribute is calculated on a leaf of the tree. In (b) it is calculated on its parent. A 3D representation of the surface area attribute is shown in (c). The red areas correspond to the area of the corresponding nodes. The contrast (blue lines) × perimeter (black lines) is calculated to obtain the grey areas which are added to the red areas to finally get the surface area attribute.

6.3.3 Tree Filtering rule

The simplification is done by pruning because the idea is to eliminate the connected components which do not meet the attribute criterion, i.e. by removing the nodes of the tree which do not meet the attribute criterion.

When filtering a tree, the question arises whether to prune the descendant of a filtered node or not.

Two approaches exist: the pruning and non-pruning strategies [124]. When a node is filtered with a pruning strategy, all of its descendants are collapsed onto the highest surviving ancestor. In the second one, a descendant of a filtered node may be preserved. The use of a pruning or a non-pruning strategy depends on the property of the attribute function A, whether it is increasing or not. An attribute function is said to be increasing if for any nodes N_1 and N_2, $N_1 \subseteq N_2 \Rightarrow A(N_1) \leq A(N_2)$.

Algorithm 1: Surface area attribute computation.

```
function Surface Area Attribute (tree, perimeter, area, level);
  Input  : A level line tree tree, attributes functions characterizing the area,
           perimeter and level of each node of the tree.
  Output : The surface area of the root of the tree S[root of the tree], corresponding
           to the total surface area of the tree.
for all nodes i of tree do
    S[i] = 0 ;
end
for all nodes i of the tree from the leaves to the root do
    if i ≠ root then
        p = parent[i];
        contrast[i] = | level[i] − level[p] |;
        S[i] = contrast[i] × perimeter[i] ;
        S[p]+ = S[i] ;
    end
end
for all nodes i of tree do
    S[i]+ = area[i] ;
end
return S[root of the tree] ;
```

If an attribute function is increasing, the pruning strategy is straightforward. Examples of increasing attributes are the area and the volume attribute.

For a non-increasing attribute A, tree filtering is not straightforward. Several strategies for non-increasing attributes have been proposed [56, 113, 116, 125].

They are illustrated in Fig. 6.8.

For a given attribute threshold t the strategies are defined as:

- *Min*: a node N is removed from the tree if $A(N) \leq t$ or if one of its ancestors is removed itself.
- *Max*: a node N is removed if $A(N) \leq t$ and if all of its descendants are removed as well.
- *Direct*: a node N is removed if $A(N) \leq t$ with no consequences for its descendants.
- *Subtractive*: a node N is removed if $A(N) \leq t$ and the gray levels of its surviving descendants are lowered.

The compactness attribute is a non-increasing attribute. We choose the subtractive rule because it removes the nodes which do not meet the compactness criterion while retaining the others, and it also preserves the contrast with the local background.

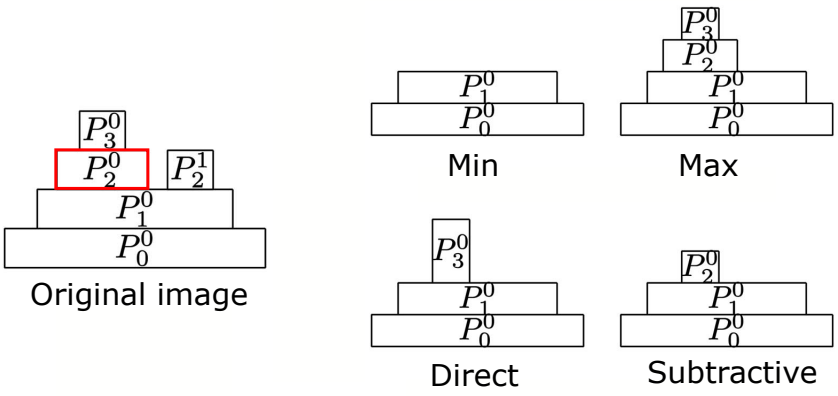
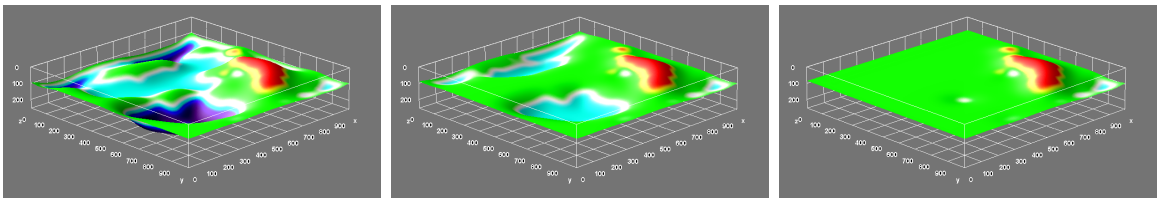


Figure 6.8: A gray-level image f is represented in 1D through its peaks components P. The peak component P_2^0 (in red) will be removed. The four filtering stratedy are presented. The figures are obtained from [116].

6.3.4 Surface area profile

We are interested in the total surface area attribute, i.e. the surface area attribute of the root of the tree.

An illustration of the DEJ surface evolution when undergoing a filtering process with an increasing compactness criterion is presented in Fig. 6.9.



(a) (b) (b)

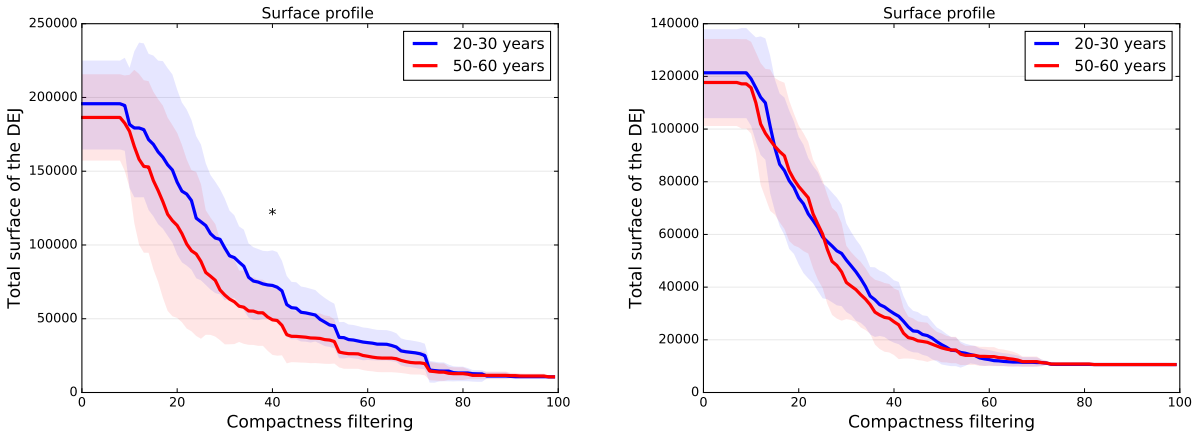
Figure 6.9: DEJ Surface filtering with an increasing compactness criterion. In (a) the compactness criterion is 0.20, in (b) 0.21 and in (c) 0.22. At each step, the total surface of the DEJ is calculated and added to the response profile.

Distributions are compared using a two-sample Kolmogorov–Smirnov test. The surface area profiles for the epidermal surface are presented in Fig. 6.10. The lines represent the mean values and the shadowed areas the standard deviations for each population (20-30 years old and 55-65 years old). Only the distributions on the cheek are statistically different.

All the distributions on the arms are identical.

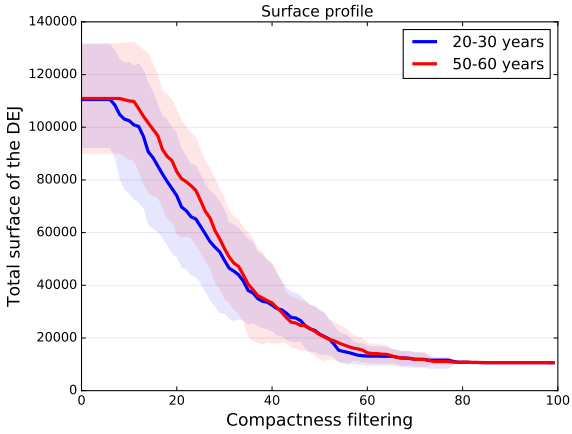
In order to reduce the inter-subject variability, we look at the *differential surface area profile*. This profile correspond to the loss of DEJ surface area attribute, expressed as a percentage, while applying compactness filters with various values.

The differential profiles are presented in Fig. 6.11. The differential profiles for the epidermal surface on the cheek are statistically different. We do not find statistical differences on the other areas.



(a) Cheek

(b) Volar arm

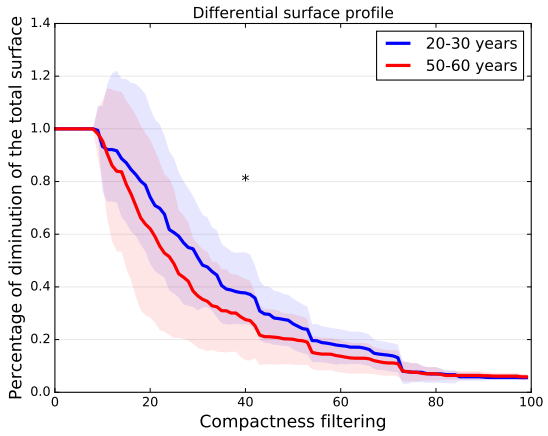


(c) Dorsal forearm

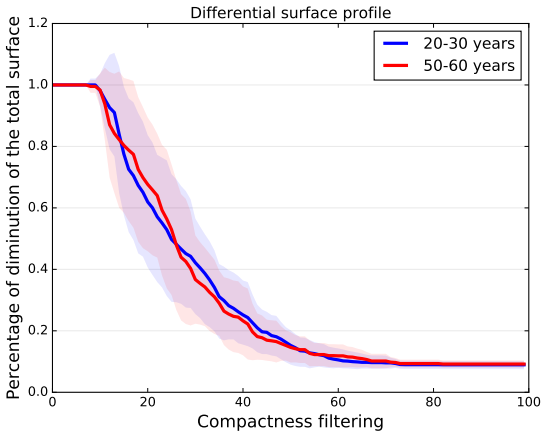
Figure 6.10: Surface area profiles with an increasing compactness criterion.

Finally, in order to obtain a descriptor of the DEJ shape we calculate the area under the curve, designated as AUC, of the differential profiles. We present box-and-whisker plots for the three acquisition areas (cheek, volar arm and dorsal forearm) in Fig 6.12.

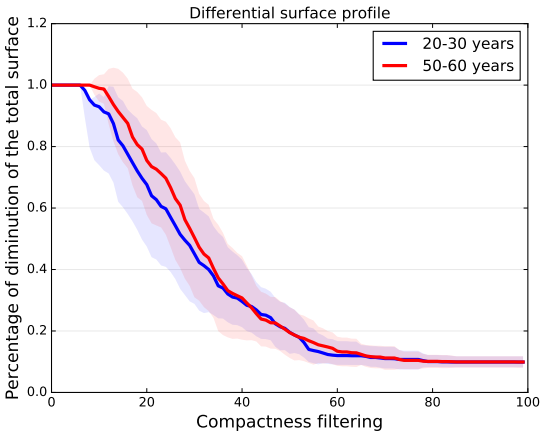
The AUC differential profile is higher in the younger group than in the older group on the cheek. When comparing the volar arm (photo-protected) and the dorsal foreram (photo-exposed), the photo-aging effect is quantified among both populations with the decrease of the AUC differential profile. We are able to assess the chronological aging on the cheek and the photo-aging for our two populations using the AUC measure.



(a) Cheek



(b) Volar arm



(c) Dorsal forearm

Figure 6.11: Differential surface area profiles.

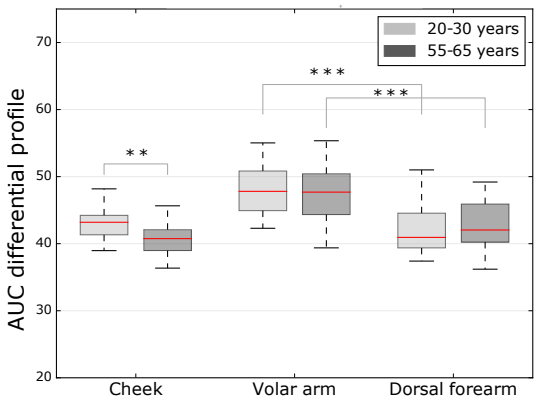


Figure 6.12: Area under curve of the differential surface area profiles.

These results supports our previous findings using the topological descriptors or the tree regarding the chronological aging on the cheek and photo-aging among the young

population. The use of the surface area profile brings additional information regarding the photo-aging among the older population compared to the topological descriptors of the level line tree. Nevertheless, both methods are relevant to assess chronological and photo aging through the characterization of the DEJ shape.

6.4 Conclusion

The DEJ shape has been described through its representation as a level line tree. Two analysis strategies have been developed: 1) the development of local and topological tree descriptors, 2) using the response profile of the DEJ surface area evolution when undergoing a filtering process.

On the one hand, topological descriptors and local features computed on relevant nodes of the tree allow the quantification of both chronological and photo-aging. On the other hand, the surface area profile brings additional information on the photo-aging process among the older group, while supporting the first results on chronological aging. The proposed method permit the discrimination between age groups.

7 Validation

Contents

In this chapter, we present the validation of the developped methods for the characterization of the epidermal honeycomb pattern and the analysis of the dermal-epidermal junction. Because of the importance of the characterization of collagen fibers for the cosmetic field, we also add our preliminary method and result for the characterization of the collagen fibers.

Two validation studies are presented. First, we present our clinical results on a large database annotated by experienced dermatologists. Then, we detail our results regarding the assessment of cosmetic products efficacy. In each section, we recall the databases involved in the analysis and we detail our guideline for the analysis before presenting our results.

Detailed description of the databases and image acquisition protocol are presented in Chap. 3.

7.1 Clinical validation

In this section, we present the clinical validation of our methods. The aim is to confront our proposed method to clinical annotations in order to evaluate their agreement. First, we describe the clinical database. Then, we detail our analysis strategy. Finally, we present our results for the epidermis, DEJ and dermis characterization.

7.1.1 Database

The study enrolls 160 subjects from 4 different ethnics backgrounds (Caucasian, Hispanic, African and Asian). Image acquisition are carried out on the cheek. The database is composed of 480 z-stacks.

The clinical database has been reviewed by experienced dermatologists. On all the 480 z-stacks, the following aging descriptors have been annotated:

- the regularity of the epidermal honeycomb pattern (0: regular or 1: irregular);
- the shape of the dermal papillae (0: ringed and 1: poly-cyclic);
- the collagen fibers dominant types (0: reticulated, 1: coarse).

The clinical study has been further described in Chap. 3.

7.1.2 Analysis guideline

We propose two analysis types:

1. the statistical analysis of our measures according to the clinical annotations. Significant differences are sought between ground-truth annotation sets;
2. some metrics (confusion matrix metrics and ROC analysis) of a predictive model build from the annotated data are evaluated.

The statistical analysis strategy is detailed in Chap. 3.

We only show the ROC for 5 out of the 10-folds cross validation for clarity. Taking all of these curves, it is possible to calculate the mean ROC, and see the variance of the curve when the training set is split into different subsets. This shows how the classifier output is affected by changes in the training data.

7.1.3 Epidermis

We recall here our two proposed measures to assess the regularity of the honeycomb pattern:

- WDC: the percentage of regular cells;
- WDR: the average size of the regular regions.

Fig. 7.1 shows box-and-whisker plots of all the measurements at all depths. No statistical difference is evident in the percentage of regular cells between honeycomb patterns annotated as regular and irregular. In our previous results on a smaller database (see Chap. 4), the percentage of regular cells, even if statistically different, remained stable with age.

The average size of regular regions is significantly higher among regular honeycomb patterns (6.6 ± 3.3) than among irregular honeycomb patterns (5.2 ± 3.1).

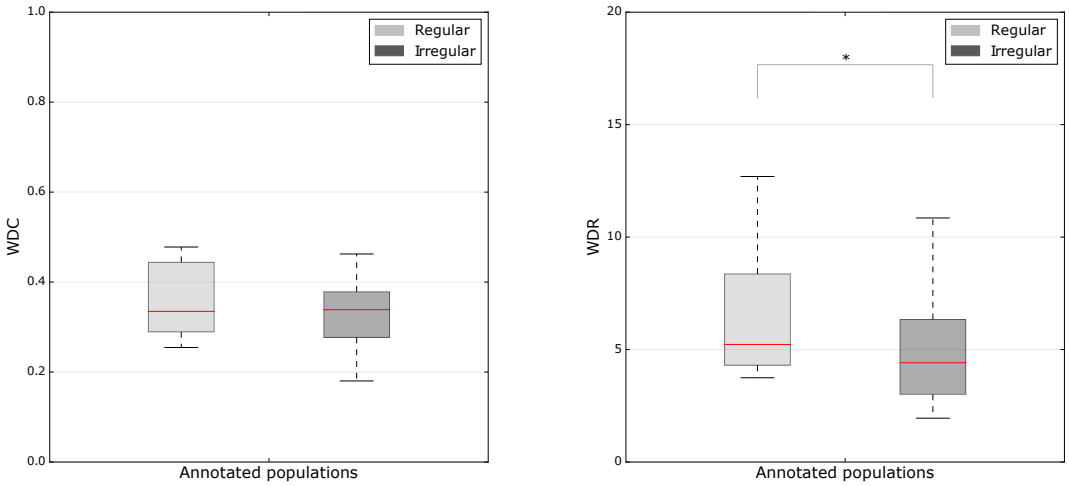


Figure 7.1: Clinical results of the regularity of the honeycomb pattern. The box-and-whisker plots represent the data distributions through their quartiles. The vertical lines indicate the variability outside the upper and lower quartiles. Means of distributions are represented as horizontal lines inside the boxes. Statistical results are represented as : * : $0.01 < P\text{-values} \leq 0.05$; ** : $0.001 < P\text{-values} \leq 0.01$; *** : $P\text{-values} \leq 0.001$.

In order to test if our method is able to predict the class (regular of irregular honeycomb pattern) of a subject, a Random Forest classifier is trained on the database. Each subject is characterized by its mean score value for all depth for both WDR and WDC. A 10-folds cross-validation gives an average precision score of 80%. We achieve sensitivity and specificity scores slightly above 80%, see Tab. 7.1.

The ROC calculated at each fold are presented in Fig. 7.2. The ROC obtained from the different splits in the data-set show different behaviors (see the ROC fold 2 and 4 on

Table 7.1: Honeycomb pattern classification.

	Scores
Accuracy	80
Sensibility	81
Specificity	81

Fig. 7.2). The best point on the mean ROC achieves a high true positive rate (~ 0.8) for a false positive rate around 0.2, but the standard deviation remains large. As the ROC show different behaviors, we obtain different AUROC values at each fold. The mean AUROC value is 0.81 ± 0.12, which can be considered as a good AUROC score even with a high standard deviation.

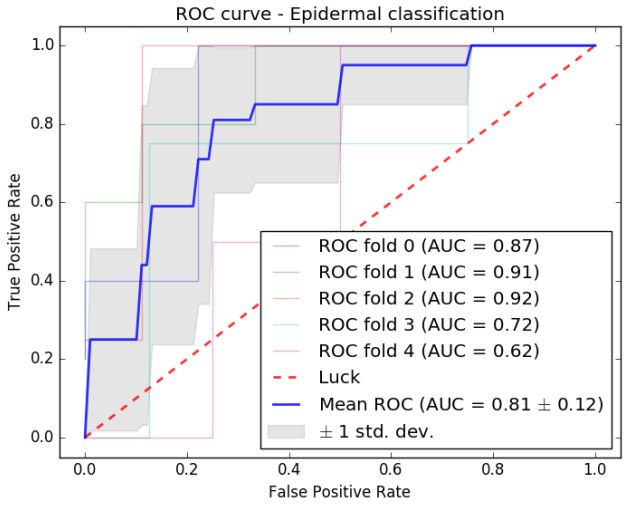


Figure 7.2: ROC curves for 5 folds of the cross-validation. The mean ROC curve is represented in blue with the standard deviation as a grey zone. At each fold, the AUROC is calculated. The dashed red line represent the ROC curve of a binary classifier with no prediction ability, i.e. no better than pure luck.

As discussed in Chap. 4, cell classification is dependent of cell shape (compactness and elongation). The shape parameters are influenced by neighboring cells. This assumption enlightens us on the significance of our results. The Random Forest classification tends to quickly classify a cell as irregular unless it is surrounded by regular cells.

Larger regular regions are consistent with visual assessment as dermatologists evaluate the extend of irregular honeycomb pattern in an image.

Even if our measures (WDC, WDR) show few statistical differences between annotated data, they allow us to gather sufficient information to achieve good predictability results.

One way to improve our results would be to work on the cells classification as regular or irregular (classification from which we calculate our measures). The introduction

of a range of regularity with several stages, rather than cruder, binary classification, could allow us to measure more subtle changes in the epidermal state. We could take inspiration from the work in [126], in which the authors propose a grading of epidermal lesions with several stages of epidermal cells regularity.

7.1.4 DEJ

The following set of measures to characterize the DEJ shape is analyzed:

1. the attributes on the leaves of the level line tree:
   - Number of leaves;
   - Compactness;
   - Complexity;
   - Elongation;
   - Volume.
2. the DEJ surface evolution during its filtering with a compactness criterion.

The number of leaves is significantly higher among ringed dermal papillae (24.9 ± 4.9) than among poly-cyclic dermal papillae (22 ± 4.7), see Fig 7.3.

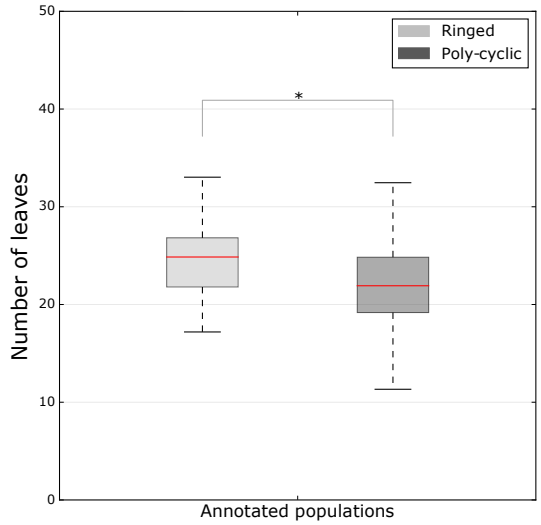


Figure 7.3: Number of leaves of DEJ annotated as ringed or poly-cyclic.

The compactness of the leaves of the DEJ is higher among ringed dermal papillea (0.45 ± 0.01) than among poly-cyclic dermal papillae (0.43 ± 0.02). The elongation of the leaves of the DEJ is higher among ringed dermal papillea (0.58 ± 0.02) than among poly-cyclic dermal papillae (0.56 ± 0.02). They are presented in Fig. 7.4.

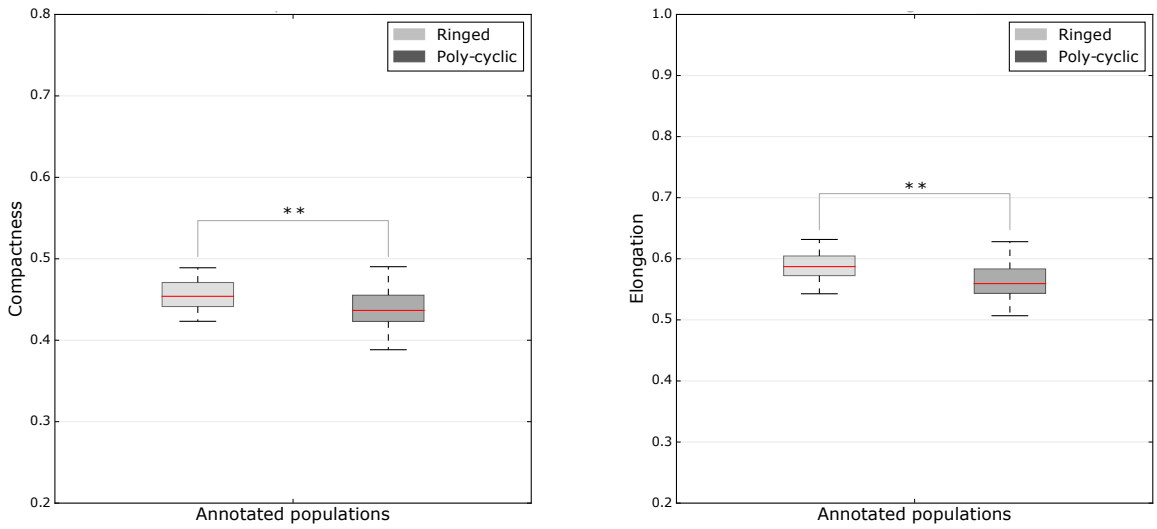


Figure 7.4: Compactness and elongation of the leaves of DEJ annotated as ringed or poly-cyclic.

We calculate the differential surface profile of the ringed and poly-cyclic DEJ with a compactness criterion. The profiles and their respective AUC are presented in Fig. 7.5. The two profiles are not statistically different, we can observe large standard deviation among both populations. On the other hand, for the calculated AUC, we do find a significant difference between the two populations.

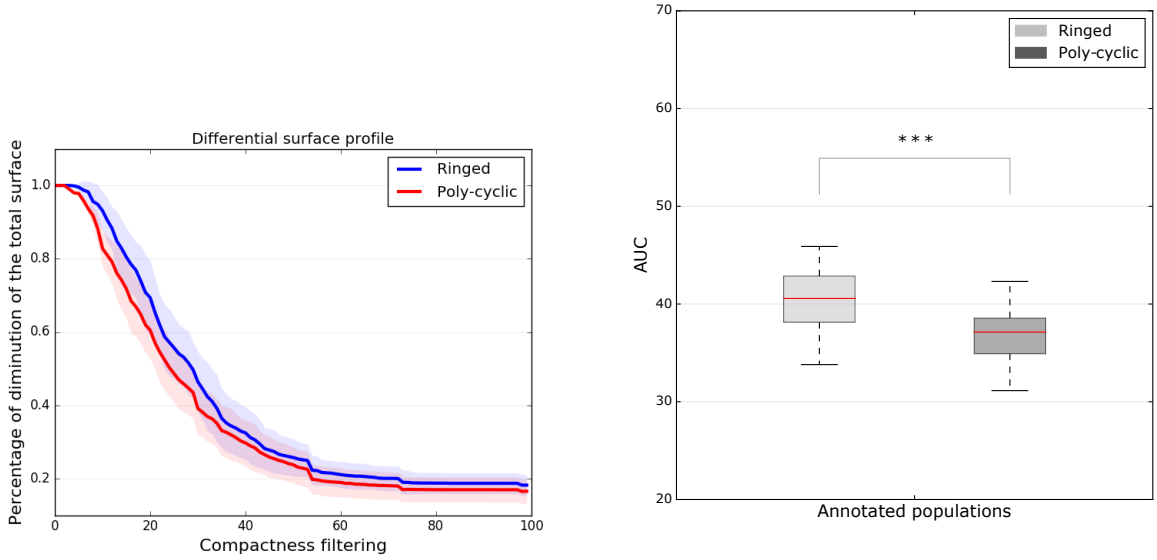


Figure 7.5: Surface profiles and AUC of DEJ annotated as ringed or poly-cyclic.

A Random Forest classifier is trained to predict the clinical annotations. Accuracy, sensitivity and specificity obtained from a 10-folds cross-validation are presented in Tab. 7.2. We

achieve 83%, 76% and 81% for respectively the global accuracy, sensitivity and specificity.

The ROC at each fold are presented in Fig. 7.6. At each fold, we achieve an AUROC value above 0.75 leading to a mean AUROC of 0.84 ±0.07.

Even if statistically different, we can not link the number of leaves to the clinical annotations as dermatologists focus on shape rather than number of dermal papillae. The number of leaves could be linked to a flattening of the DEJ.

The shape parameters (compactness and elongation) are more relevant to discriminate ringed from poly-cyclic dermal papillae. They mimic the visual analysis by dermatologists as they describe the shape of the leaves of the DEJ, i.e. the shape of the dermal papillae.

The DEJ annotated as ringed are composed of significantly more compact objects than the poly-cyclic ones.

We only incorporate images annotated as ringed or poly-cyclic which correspond to 67 subjects. As the visibility of dermal papillae is highly dependent on the skin phototype and acquisition zone, most confocal images cannot be annotated for the DEJ aging pattern. One step further would be the automated discrimination of visible and non-visible dermal papillae. We could rely on textural features as visible papillae appear with strong contrast while non-visible dermal papillae show an amorphous pattern. The question arises on how to interpret results obtained from non-visible dermal papillae. Further work in cooperation with dermatologist is needed to understand the structure of blurry DEJ.

Table 7.2: Clinical DEJ classification.

	Scores
Accuracy	83
Sensibility	76
Specificity	81

7.1.5 Dermis

We present in this section our preliminary method to characterize the collagen fibers.

The characterization of the collagen fibers is a key research indicator for the cosmetic field as it suffers from severe damage with aging. Within the clinical database, the collagen fibers are classified into two categories:

- thin reticulated collagen: bright thin fibrillar structures forming a delicate web-like pattern;
- coarse collagen structure: coarse filamentous structures with a tendency to be packed. A web-like pattern is still observable but with larger and irregularly spaced meshes;

Examples of the collagen fiber types are presented in Fig. 7.7.

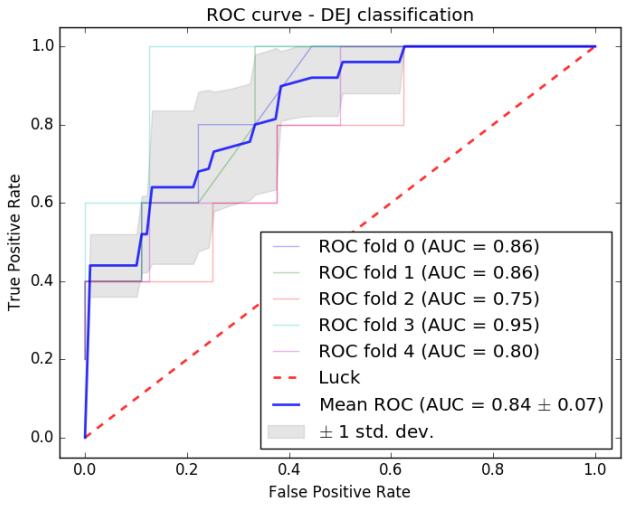
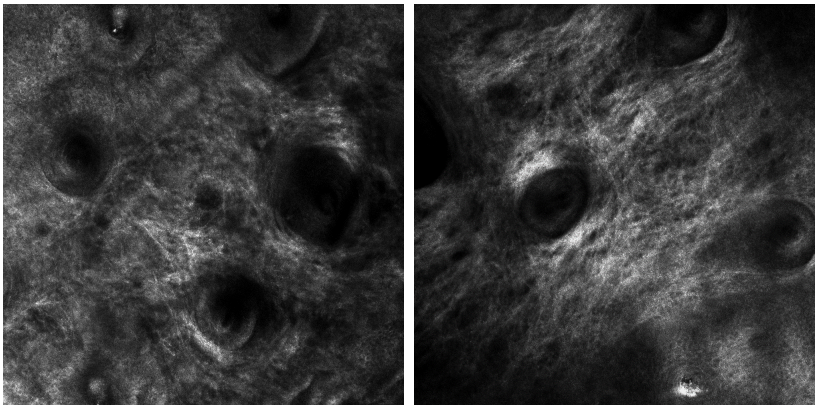
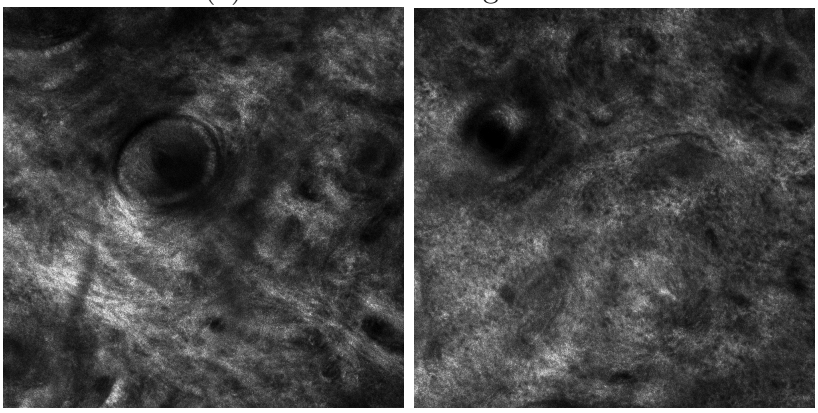


Figure 7.6: Roc curves at each fold of the cross-validation for the DEJ classification.



(a) Reticulated collagen fibers



((b) Coarse collagen fibers

Figure 7.7: Collagen fiber types: (a) reticulated fibers characterized by tiny reflective fibers orderly disposed to form web-like structure, (b) coarse fibers composed by thick fibers roughly arranged to form an irregular net.

Collagen fibers classification

First, a pixel-level classification as reticulated or coarse is performed, using a training set composed of 37 subsets of images containing only reticulated fibers and 45 subsets of images containing only coarse fibers. A Random Forest classifier is trained using Gabor features. Gabor filters are used to obtain textural information from the confocal images. When a Gabor filter is applied to an image, it gives the highest response at edges and at points where texture changes. A bank of Gabor filters can form a wavelet frame suitable for texture analysis [107]. Our Gabor filter bank is computed with 4 levels of frequency and 8 orientation with a 32×32 Gaussian kernel. As features, we use the Local Energy, i.e. the sum of the squared values of the Gabor feature image and the Mean Amplitude of the response, i.e. the sum of the absolute value. These textural features are calculated for every thirty pixels in the image.

Each pixel of the training data is classified as reticulated or coarse using a Random Forest classifier on a 10-fold cross-validation. Accuracy, sensitivity and specificity of the pixel classification are presented in Tab. 7.3. We achieve a global accuracy of 76%.

From the classification results, we calculate the percentage of an image classified as reticulated to measure the extend of reticulated fibers, as performed by dermatologists. We consider that if an image is composed of more than 50% of reticulated fibers, it can be entirely classified as reticulated. New confusion matrix metrics are calculated and presented in Tab. 7.4. The specificity of our classification increases from 70% to 82%.

The threshold of 50% has been empirically chosen and does not necessarily correspond to what dermatologists choose. Finding the optimum threshold, i.e. the threshold that mimics the visual analysis, could improve our results further.

Table 7.3: Pixel classification as reticulated or coarse.

	Scores
Accuracy	76
Sensibility	78
Specificity	70

Table 7.4: Image classification as reticulated or coarse.

	Scores
Accuracy	78
Sensibility	73
Specificity	82

Clinical validation for the dermis characterization

We now present our clinical results for the characterization of the collagen fibers. The percentage of reticulated fibers among both annotated populations are presented in Fig. 7.8. Images annotated as reticulated by the dermatologists have a significantly higher percentage of reticulated fibers than images annotated as coarse. This result agrees with the visual analysis by dermatologists who assess the extend of reticulated fibers within images. From the percentage of reticulated fibers, we can classify each image as reticulated or coarse. The accuracy, sensitivity and specificity of the classification are presented in Tab. 7.5. We achieve a good specificity of 89%. The sensitivity score could be further improved by the choice of an optimum threshold as discussed above.

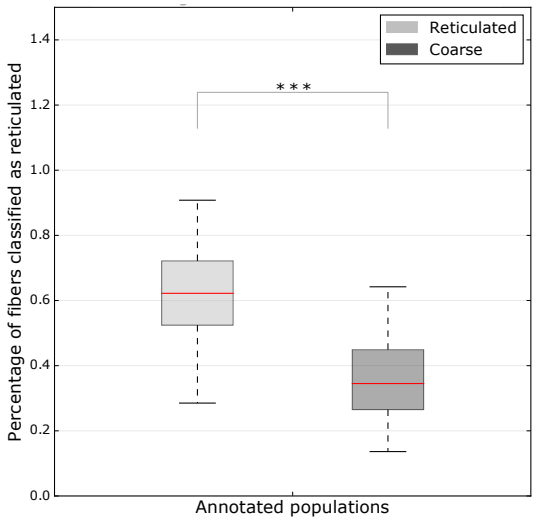


Figure 7.8: Percentage of reticulated collagen fibers among populations annotated as reticulated and coarse.

Table 7.5: Collagen fibers classification.

	Scores
Accuracy	80
Sensibility	63
Specificity	89

7.1.6 Conclusion

In this section, we have presented the clinical validation of our methods on a large database. To our knowledge, these are the first results comparing a computer-based approach to

dermatologists' approach for the assessment of skin aging using in vivo confocal microscopy. The significance of our proposed measurements and the prediction scores are summarized in Tab. 7.6 and Tab. 7.7. Our proposed measurements provide a consistent agreement with the clinical annotations. Since 4 ethnics backgrounds are enrolled in the study, our proposed methods can be considered independent of ethnicity. We have discussed several ways to improve the robustness of the proposed methods. One way to improve our results would be to take advantage of the ROC analysis to find an optimum threshold which could better recover the clinical annotations.

In the next section, we present our results regarding a the cosmetic product efficacy.

Table 7.6: Summary of the significance of our measurements methods when comparing clinical annotated sets. Statistical significances are consistent with dermatologists' annotations. Statistical results are represented as : NS : $0.05 < P\text{-}values$; * : $0.01 < P\text{-}values \leq 0.05$; ** : $0.001 < P\text{-}values \leq 0.01$; *** : $P\text{-}values \leq 0.001$.

Skin layer	Measurements	Statistical differences consistent with clinical annotations
Epidermis	Percentage of regular cells	NS
	Average size of regular regions	*
DEJ	Number of leaves	*
	Leaves compactness	**
	Leaves elongation	**
	Differential surface area profile	NS
	AUC Differential surface area profile	***
Dermis	Percentage of reticulated collagen fibers	***

Table 7.7: Summary of the clinical annotations prediction for each skin layer.

	Epidermis	DEJ	Dermis
Accuracy	76	78	80
Sensibility	78	73	63
Specificity	70	82	89

7.2 Cosmetic product efficacy

7.2.1 Database

Our database is composed of 17 subjects between 45 to 60 years old. Four different conditions are tested:

1. P1: A controlled condition with no product application;
2. P2: Application of a moisturizing cosmetic product;

3. P3: Application of an anti-aging cosmetic product containing Retinol;

4. P4: Application of an anti-aging cosmetic product containing AHA.

The two anti-aging products have different cellular action pathways. One contains 0.3% of Retinol, a well-known anti-aging substance that stimulates the collagen production. And the other 10 % of AHA agents that stimulate the skin epidermal renewal. Our goal is to measure the product efficacy and to further understand the different action pathways of active ingredients in vivo.

Acquisitions are carried out at three times: initial time $T0$, one-week of application $T1$ and two-weeks of application $T2$. At each time, three z-stacks are acquired.

The study protocol is described in Chap. 3.

7.2.2 Statistical analysis

The statistical analysis strategy was described in Chap. 3.

Statistical analysis is performed to assess the variations of our proposed measures over time on each of the acquired area. Our first concern is to evaluate if there are any significant variations over time for the controlled condition (condition 1 with no product application). If the controlled condition remains stable over time, we assume that if there is any change in the other conditions, these changes are due to the product action. Therefore, for each product, we compare variations with respect to the $T0$ acquisition.

If there is a significant variation over time in the controlled condition, we suppose that the skin intrinsic structures have been modified due to external factors, such as the environment. In that case, we analyze at each time the differences between the products and the controlled condition, i.e. we discard the variations due to external factors.

The statistics significance is defined as follow:

- LS : $0.05 < \textit{P-values} \leq 0.1$;
- * : $0.01 < \textit{P-values} \leq 0.05$;
- ** : $0.001 < \textit{P-values} \leq 0.01$;
- *** : $\textit{P-values} \leq 0.001$.

We do not use our prediction model, which has been described in the previous section, as the two analysis has been performed simultaneously.

7.2.3 Results

Epidermis

The two proposed measures are analyzed: the averaged percentage of regular cells (WDC) and averaged average size of regular regions (WDR) at all depths. There is no significant evolution neither in the percentage of regular cells (WDC) nor in the average size of regular regions (WDR). The controlled condition remains stable over time. No other condition presents significant changes. As the product P4 stimulates the epidermis renewal, an improvement of its epidermal measures over time is expected.

Then, the two measures at the deeper layer of the epidermis are analyzed. The epidermis renews himself in 4 to 6 weeks. Thus, an early sign of the product action might be observed at the bottom of the epidermis, where new cells should be produced. Even there, we find no significant evolution of our proposed measures over time for any conditions.

Two reasons might explain these results. First, 2 weeks of application might be too short to assess early signs of products action on the epidermis. Second, as discussed in the previous section, our cell classification as regular or irregular might not be sensitive enough to detect subtle changes in the honeycomb pattern conformation.

DEJ

We analyze our proposed set of measures to characterize the DEJ shape:

1. the attributes on the leaves leaves of the level line tree:
    - Number of leaves,
    - Compactness,
    - Complexity,
    - Elongation,
    - Volume;
2. the DEJ surface profile.

The number of leaves significantly increases for the controlled condition over time, see Fig. 7.9. Even if the P4 condition shows an increase of its number of leaves over time, when comparing the differences with the controlled condition, we find no significant evolution. We cannot conclude to a product action for the P4 condition. No other condition show significant difference compared to the control.

The complexity of the leaves remains stable over time for the controlled condition. Thus, we can analyze other conditions over time itself. The complexity of the leaves

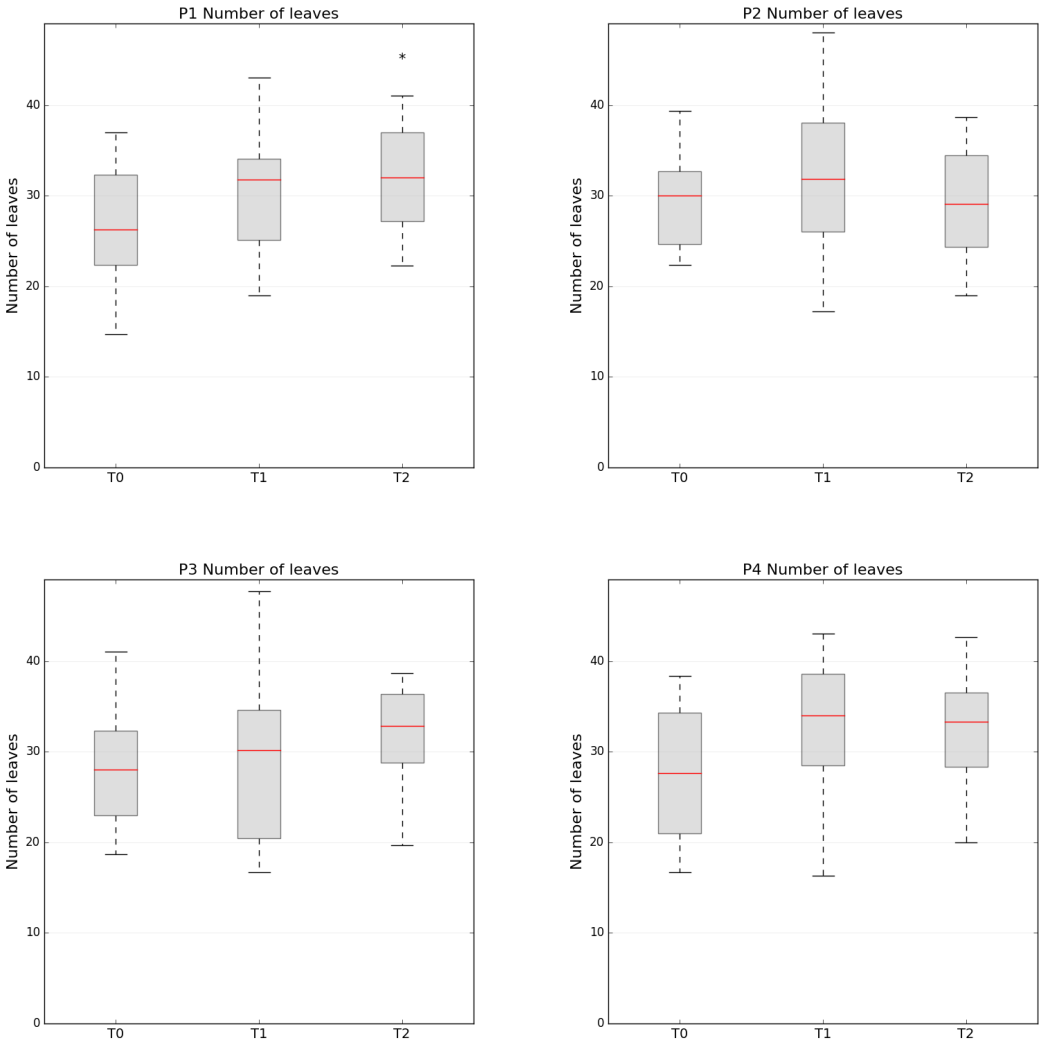


Figure 7.9: Number of leaves of the DEJ for the four condition. P1: controlled condition. P2: moisturizing product. P3: product containing Retinol. P4: product containing AHA.

decreases at $T2$ for the P3 condition, i.e. the leaves shape gains in regularity. In addition, the moisturizing product (condition P2) shows no significant evolution over time. Therefore, we can conclude that the results for the condition P3 must be due to the anti-aging action of the product and not to a moisturizing effect. The complexity results for all products are presented in Fig. 7.10.

The elongation of the leaves remains stable over time for the controlled condition. The elongation of the leaves decreases at $T2$ for both the P3 and P4 conditions, i.e. the peaks and valleys shape gains in circularity. The elongation results for all products are presented in Fig. 7.11.

Finally, there is a significant increase in the leaves volume for the P3 condition at $T2$, see Fig. 7.12. All other conditions remain stable over time. We can conclude to an

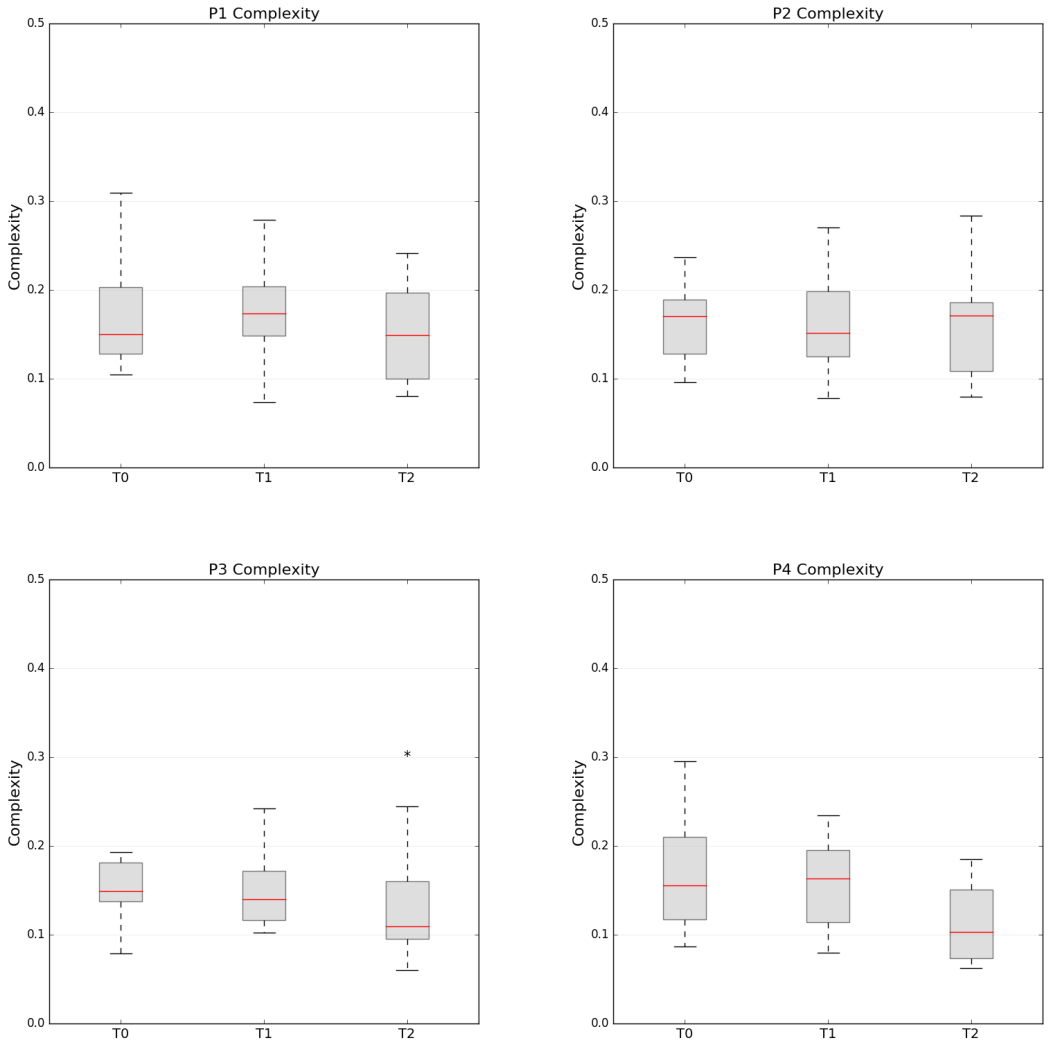


Figure 7.10: Complexity of the leaves of the DEJ for the four condition. P1: controlled condition. P2: moisturizing product. P3: product containing Retinol. P4: product containing AHA.

anti-aging action of the product P3 on the volume of the peaks and valleys of the DEJ.

There is no significant evolution for the compactness measure nor for the surface profile of the DEJ for any conditions.

Dermis

Last but not least, we present our preliminary results of the collagen fibers characterization for our cosmetic product efficacy study. We find a limit significant (p-value ≤ 0.1) increase in the percentage of reticulated fibers at $T2$ for the P3 condition and no evolution for any other condition. The results are presented in Fig. 7.13. Even if we obtain a low significance, the fact that product P3 is the only tested product which stimulates collagen

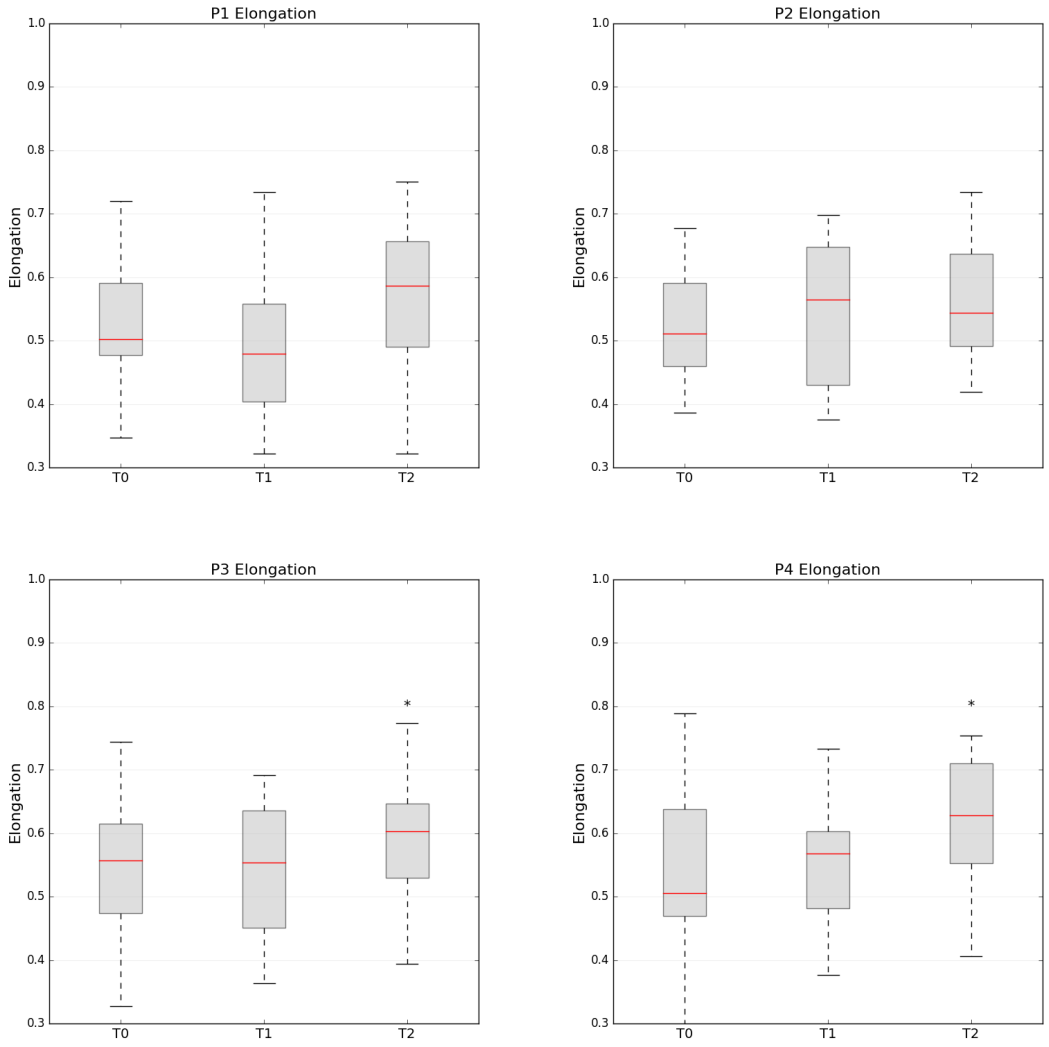


Figure 7.11: Elongation of the leaves of the DEJ for the four condition. P1: controlled condition. P2: moisturizing product. P3: product containing Retinol. P4: product containing AHA.

production supports the interpretation of our result. It could well be that, with one or two more weeks of application, we could obtain more significant results.

7.2.4 Conclusion

The results are consistent with the cosmetic products action except on the epidermis. We can surmise that the stimulation of the epidermal renewal (with product 4) may have a positive effect on the DEJ, as one of its roles is the support of the epidermis. Moreover, we have highlighted a trend in the stimulation of collagen production with the product

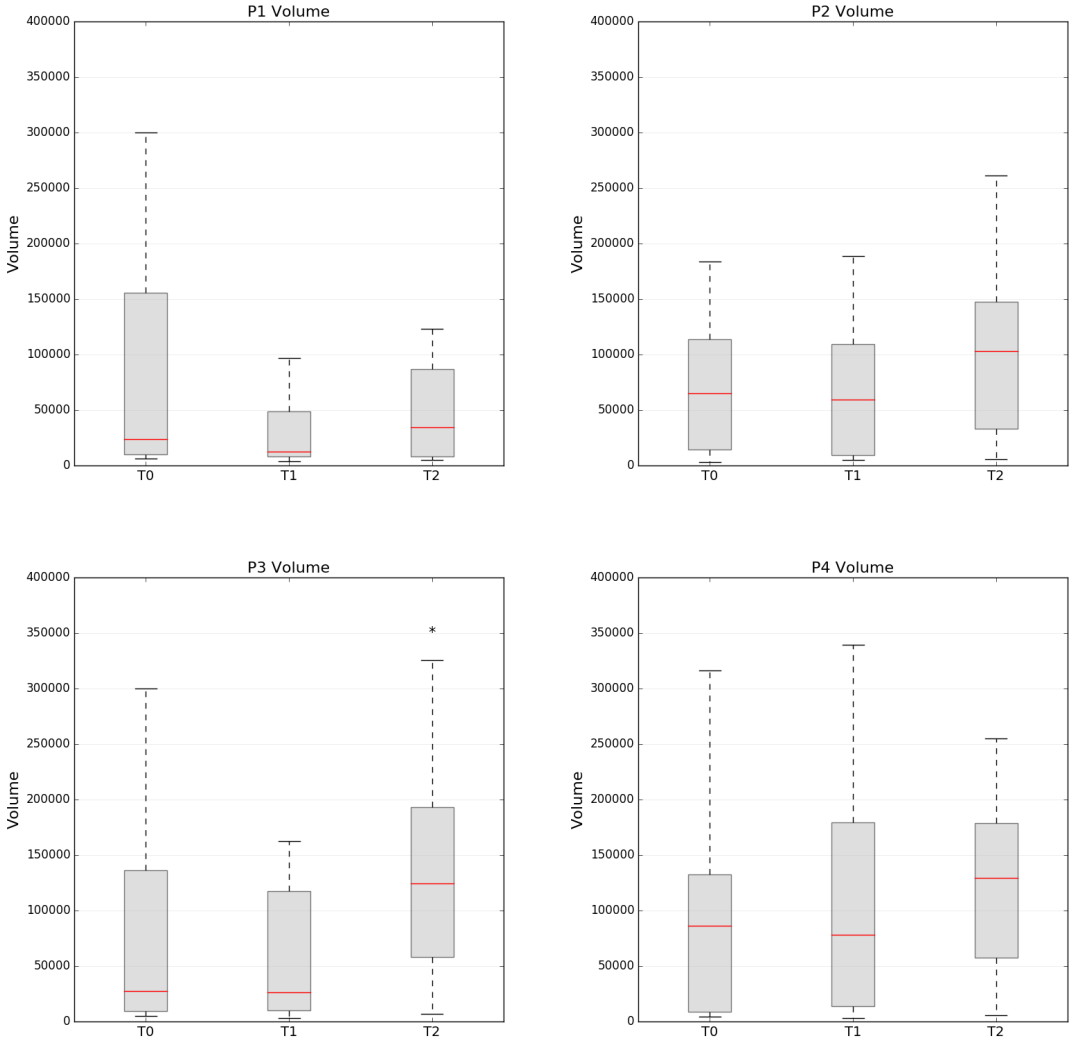


Figure 7.12: Volume of the leaves of the DEJ for the four condition. P1: controlled condition. P2: moisturizing product. P3: product containing Retinol. P4: product containing AHA.

3, which also can have a positive effect on the DEJ by improving its foundations. We are able to detect early sign of these products' anti-aging actions.

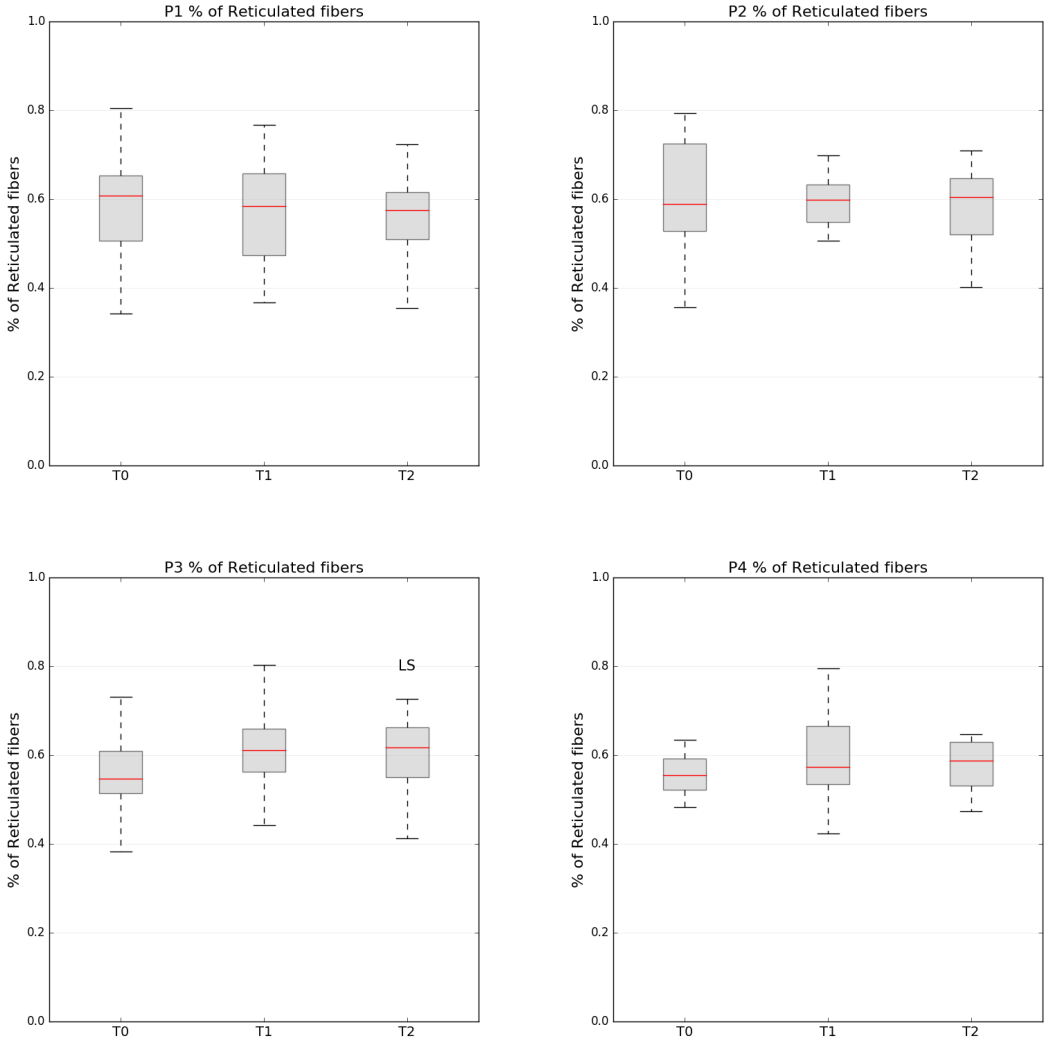


Figure 7.13: Percentage of reticulated fibers for the four condition. P1: controlled condition. P2: moisturizing product. P3: product containing Retinol. P4: product containing AHA.

8

Conclusion and perspectives

Contents

8.1 Conclusion and perspectives

This research work has allowed us to study the automation of the skin aging process using in-vivo confocal microscopy. We proposed several methods for the computational quantification of skin aging on confocal images.

Chapters 2 and 3 have provided a better understanding of the RCM images and the skin aging patterns. They detailed the skin inner structure and the skin appearance on confocal images. We have set up and run several studies on women volunteers to first perform a clinical validation of our methods using clinical annotations and to assess cosmetic product efficacy.

The chapter 4 is dedicated to the automatic characterization of the epidermal state through the study of the regularity of its honeycomb pattern. We proposed a segmentation method using prior biological information and two regularity measures which have enabled us to quantify differences between age groups.

In chapter 5, we have detailed our method to segment the skin layers. We took advantage of a biological information on the skin structure, i.e. its layers organization, to provide a robust segmentation method.

In chapter 6, we have proposed a hierarchical representation of the DEJ surface as a level line tree, in order to characterize its shape. The features used in chapter 6 to analyze the DEJ shape were chosen for their close connection with visual expertise, so that they should be relevant to reveal differences in the DEJ shape with age.

The chapter 7 summarizes the validations of our proposed methods compared to clinical annotations and describes a cosmetic product efficacy study. Our clinical analyses have revealed several ways we could improve our results. Among the numerous perspectives that arise from this work, we plan to focus on three main ideas:

- first, the cell classification involved in the characterization of the epidermal honeycomb pattern is influenced by neighboring cells. We want to refine this classification to detect more subtle changes in the epidermal state. Our first goal will be to study the regularity classification of the epidermal cells on a database of subjects with an healthy epidermis versus subjects suffering from Actinic Keratosis (AK). Actinic Keratosis is a pathological condition which can evolve to a cancerous condition. In AK lesions, epidermal keratinocytes are clearly irregular, they form an atypical honeycomb pattern. This study will allow us to better understand the regularity range between a healthy keratinocyte and a clearly irregular one. This should help us to improve our cell classifier by incorporating other features such as shape variance and size variance of epidermal cells;

- second, when analyzing annotated DEJ as ringed or poly-cyclic, we excluded the non-visible DEJ pattern. We intend to include these DEJ annotations which are, to date, considered as flat by dermatologists. Non-visible dermal papillae could be detected using textural features. The volume and volume variance of the peaks and valleys could be a meaningful indicator to evaluate the non visible dermal papillae and their suspected absence of relief. The question arises of the meaning of other parameters, such as compactness and elongation, on non-visible DEJ. Even if the dermal papillae are not visible, the DEJ still exists. A first start to study this question and open discussions would be to review 3D segmentations of non-visible DEJ with experienced dermatologists and to perform histological comparisons;

- third, the classification of the collagen fibers needs to be improved, especially on the relevance of the features used in the classification and also on the finding of an optimum threshold which agree with clinical annotations.

In the course of this work, we have developed methods inspired by the scoring system used by dermatologists to assess skin aging patterns on in-vivo confocal images. We have shown that our methods are not only capable of assessing changes due to skin aging, but that they also agree with the clinical annotations provided by experienced dermatologists. Moreover, we have shown that RCM has the power to quantify skin aging markers. By using this quantification over time, we can provide quantitative measurements of the efficacy of skin care products.

8.2 Perspectives in the use of the confocal microscope

In this section, we present several perspectives for the use of reflectance confocal microscopy in the context of skin studies. First, we present a study regarding the possibility of using the confocal microscope as an ex-vivo device. Then, we describe our skin phantom for confocal image restoration. And finally, we discuss the contributions that the detection of early signs of anti-aging products action could bring to the understanding of clinical events of skin aging and anti-aging cosmetic products actions.

8.2.1 Ex-vivo validation

Our first intent is to perform confocal acquisitions of the skin explant, followed by histological analysis in order to validate our methods. Therefore, we built a device allowing ex-vivo acquisitions using our confocal microscope. The use of skin explant could allow us to study the skin aging process under controlled conditions such as ultraviolet irradiation or application of Glucocorticoids (GC) which can induce accelerated aging.

We used a specific device designed to attach the confocal probe to the skin explant. The device consists of a fixed stainless steel part and a movable part in the center of which is housed a polymer cylinder connected to a spring, see Fig. 8.1. This arrangement allows the device both to adapt to the position of the probe and to cope with the varying sample thickness.

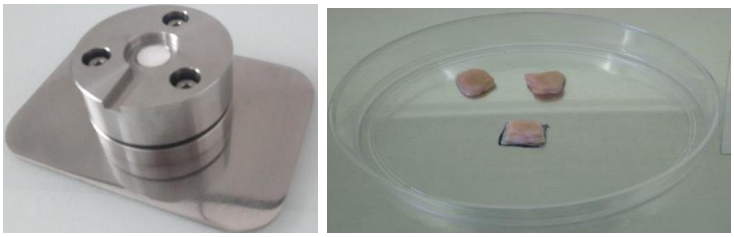
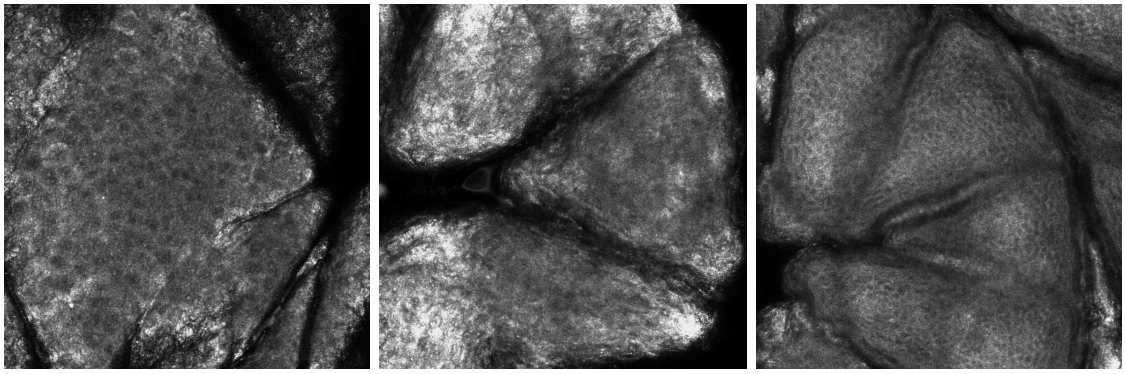


Figure 8.1: Ex-vivo imaging device.

We were not able to visualize either the epidermal honeycomb pattern or any other skin patterns on ex-vivo images, see Fig. 8.2(b).

To keep them alive, ex-vivo explants are put in an incubator at 37°C with a high humidity level. We suspect the explant preservation protocol to interfere with the image formation. Acquisitions on skin explant before any preservation treatment reveals skin patterns, see Fig. 8.2(c). Further reflection and discussion are needed to understand what are the processes that happens to explants as they are incubated.



(a) (b) (c)

Figure 8.2: Comparison of in vivo and ex vivo acquisitions. (a) In vivo image of the epidermal layer showing a honeycomb pattern. (b) Ex vivo image of the epidermal layer after it has been kept in a stove for preservation, the honeycomb pattern is not visible. (c) Ex vivo image of the epidermal layer before any treatment, the honeycomb pattern appears undamaged.

The ex-vivo device may be useful beyond histological validation. The skin explant can be imaged from the bottom, revealing deep collagen fibres and deeper skin tissue such as adipocytes. In-vivo acquisitions do not allow the visualization of such structures.

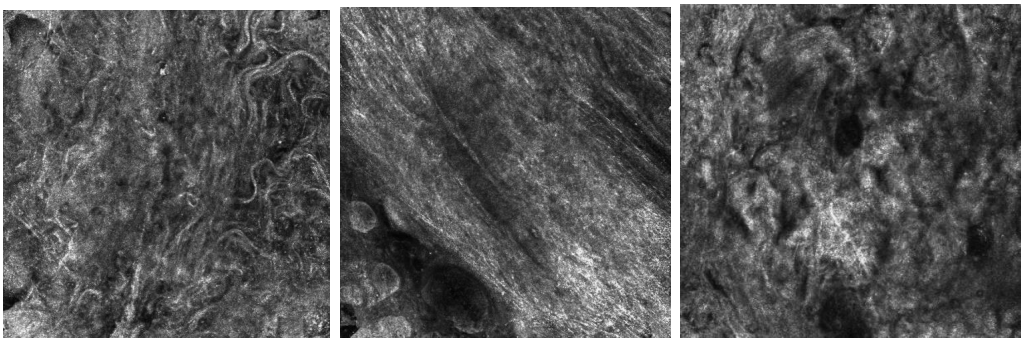
Example of deep collagen fibers are presented in Fig. 8.3. Three aging pathways have been tested: chronological aging, Glucocorticoids treatment and ultraviolet irradiation. We can observe that the fibers show differing patterns.

In Fig. 8.4, we present our first visualization of adipocytes (fat cells). Each round structure is an adipocyte containing a lipid droplet which appear as an oily structure.

We still need to standardize the acquisitions protocol, especially to set up the acquisition depth, but the use of this device could help us understand additional skin structures such as the deeper dermis and subcutaneous tissue.

8.2.2 Image restoration

During the acquisition, confocal microscopy images are corrupted by Poisson noise [127]. Unlike Gaussian noise, Poisson noise is signal-dependent, making it difficult to separate



(a) (b) (c)

Figure 8.3: Ex vivo collagen fibers undergoing (a) chronological aging, (b) aging due to Glucocorticoides treatment and (c) aging due to ultraviolet irradiation.

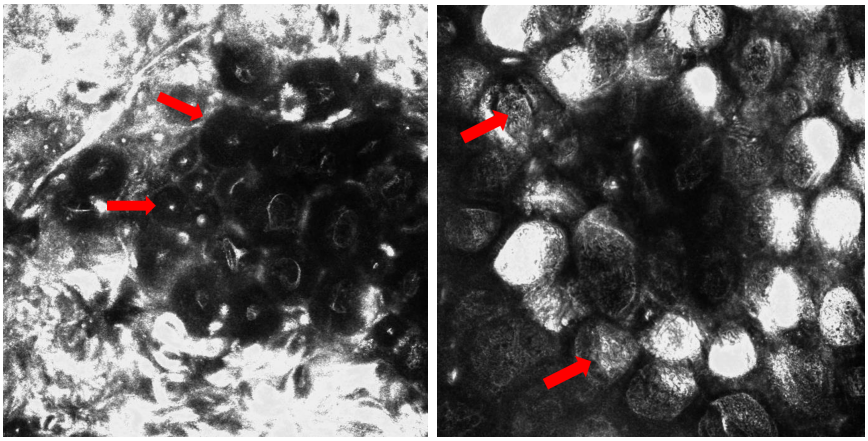


Figure 8.4: Adipocytes visualization using the ex-vivo imaging device. Red arrows indicate adipocytes containing lipid droplet which appear as oily structures.

from the signal itself. Poisson noise is intrinsically linked to the inherent natural variation of the incident photon flux. The images are also corrupted by spherical aberrations due to the optical path. To improve image quality, methods of image deconvolution can be used [128]. Even if one resorts to blind deconvolution schemes, a priori knowledge related to the Point Spread Function (PSF) is highly desirable. The PSF describes the response of an imaging system to a point source or point object. This knowledge can be acquired either by studying PSF theoretical properties or experimental data. Experimental PSFs may be measured using calibration beads or directly from the image by extracting small point-like objects [129]. Such PSFs can be used for instance to validate theoretical parametric PSF models or to assess the aberration of point spread function in given imaging systems.

One goal related to confocal microscopy images would be to assess experimentally the PSF depth variation by imaging sample with point sources (beads) placed at varying depths.

We aim to develop a skin phantom that includes the epidermis, the stratum corneum and the dermis [130]. It should mimics the optical properties of the skin tissue [131].

We developed an optical skin phantom in collaboration with the National Metrology Centre in Singapour. The design of our optical phantom is shown in Fig. 8.5.

The skin phantom is divided into three components which correspond to:

- the stratum corneum with a refractive index of 1,51 and a thickness of 15 μm;
- the epidermis with a refractive index of 1,34 and a thickness of 100 μm;
- the dermis with a refractive index of 1,40 and a thickness of 1mm.

each component layer is made of silicon material with specific refractive index. The refractive index for the three layers has been tested at 589 nm under 20 – 23 C°.

Each component should be filled with beads with a refractive index similar to the refractive index approaching the one of melanin (higher than 1.51). The beads size should be equal to the spinous cells size in the epidermis, i.e. between 15 and 25 μm. Soda lime microspheres were chosen as refractive beads. They have a refractive index of about 1.5. The medium diameter of the soda lime spheres is however higher than the theoretical one, about 35 μm to 45 μm.

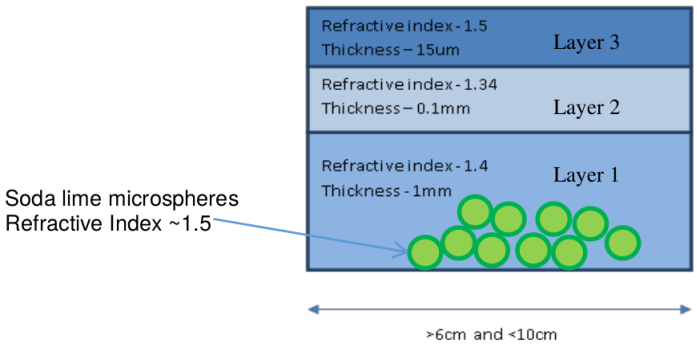


Figure 8.5: Scheme of the designed optical phantom.

Our optical phantom has not yet been tested. We hope it will help us improve confocal images quality.

8.2.3 Early signs of skin anti-aging products action

During our research work, we have proposed several methods to study the skin aging descriptors, which have been previously identified by dermatologists.

We aim to further understand the skin aging process and cosmetic products actions, by connecting clinical signs of aging to microscopic events.

The cosmetic study that has been performed produced encouraging results given the small size of our population and the short period of product application.

To date, targets of anti-aging cosmetic active ingredients have been mostly studied by histological means or ex-vivo experimentation. In-vitro models are reliable but they lack a more direct correlation to the in-vivo aging process. Knowledge obtained from in-vitro models is still hardly portable to in-vivo applications, results often show large differences.

The study of anti-aging cosmetic ingredients by means of RCM that is followed by a clinical study after several months of applications could help to understand their early actions and above all their clinical consequences.

A step further takes us to the prediction of clinical anti-aging product actions. By knowing the active ingredient pathway of action and its impact on skin cellular structures in-vivo after a short period of application, one could glimpse the clinical consequences to come and so, highlight better active ingredient in a much faster and accurate way.

References

[1] Gary J Fisher, Sewon Kang, James Varani, Zsuzsanna Bata-Csorgo, Yinsheng Wan, Subhash Datta, and John J Voorhees. "Mechanisms of photoaging and chronological skin aging". In: *Archives of dermatology* 138.11 (2002), pp. 1462–1470.

[2] Andrea Vierkötter, Ulrich Ranft, Ursula Krämer, Dorothea Sugiri, Verena Reimann, and Jean Krutmann. "The SCINEXA: a novel, validated score to simultaneously assess and differentiate between intrinsic and extrinsic skin ageing". In: *Journal of dermatological science* 53.3 (2009), pp. 207–211.

[3] R Rox Anderson. "Polarized light examination and photography of the skin". In: *Archives of dermatology* 127.7 (1991), pp. 1000–1005.

[4] Klaus-P Wilhelm, Anastasia B Cua, and Howard I Maibach. "Skin aging: effect on transepidermal water loss, stratum corneum hydration, skin surface pH, and casual sebum content". In: *Archives of dermatology* 127.12 (1991), pp. 1806–1809.

[5] Hachiro Tagami, Masatoshi Ohi, Keiji Iwatsuki, Yuko Kanamaru, Mizuho Yamada, and Bunjiro Ichijo. "Evaluation of the skin surface hydration in vivo by electrical measurement". In: *Journal of Investigative Dermatology* 75.6 (1980), pp. 500–507.

[6] Ronald R Warner, Mark C Myers, and Dennis A Taylor. "Electron probe analysis of human skin: determination of the water concentration profile". In: *Journal of Investigative Dermatology* 90.2 (1988), pp. 218–224.

[7] Hyo Sub Ryu, Young Hyun Joo, Sun Ok Kim, Kyoung Chan Park, and Sang Woong Youn. "Influence of age and regional differences on skin elasticity as measured by the Cutometer®". In: *Skin Research and Technology* 14.3 (2008), pp. 354–358.

[8] Dominique Batisse, Roland Bazin, and Thérèse Baldeweck. "Influence of age on the wrinkling capacities of skin". In: *Skin Research and Technology* 8.3 (2002), pp. 148–154.

[9] Shulian Wu, Hui Li, Hongqin Yang, Xiaoman Zhang, Zhifang Li, and Shufei Xu. "Quantitative analysis on collagen morphology in aging skin based on multiphoton microscopy". In: *Journal of biomedical optics* 16.4 (2011), pp. 040502–040502.

[10] J Lademann, A Knüttel, H Richter, N Otberg, RV Pelchrzim, H Audring, H Meffert, W Sterry, and K Hoffmann. "Application of optical coherent tomography for skin diagnostics". In: *Laser physics* 15.2 (2005), pp. 288–294.

[11] Yicong Wu and Jianan Y Qu. "Autofluorescence spectroscopy of epithelial tissues". In: *Journal of biomedical optics* 11.5 (2006), pp. 054023–054023.

[12] JB Pawley and Barry R Masters. "Handbook of biological confocal microscopy". In: *Optical Engineering* 35.9 (1996), pp. 2765–2766.

[13] Giovanni Pellacani, Pascale Guitera, Caterina Longo, Michelle Avramidis, Stefania Seidenari, and Scott Menzies. "The impact of in vivo reflectance confocal microscopy for the diagnostic accuracy of melanoma and equivocal melanocytic lesions". In: *Journal of investigative dermatology* 127.12 (2007), pp. 2759–2765.

[14] Caterina Longo, Alice Casari, Francesca Beretti, Anna Maria Cesinaro, and Giovanni Pellacani. "Skin aging: in vivo microscopic assessment of epidermal and dermal changes by means of confocal microscopy". In: *Journal of the American Academy of Dermatology* 68.3 (2013), e73–e82.

[15] Caterina Longo, Alice Casari, Barbara Pace, Silvia Simonazzi, Giovanna Mazzaglia, and Giovanni Pellacani. "Proposal for an in vivo histopathologic scoring system for skin aging by means of confocal microscopy". In: *Skin Research and Technology* 19.1 (2013).

[16] EMT Wurm, C Longo, C Curchin, HP Soyer, TW Prow, and G Pellacani. "In vivo assessment of chronological ageing and photoageing in forearm skin using reflectance confocal microscopy". In: *British Journal of Dermatology* 167.2 (2012), pp. 270–279.

[17] Daniel S Gareau. "Automated identification of epidermal keratinocytes in reflectance confocal microscopy". In: *Journal of biomedical optics* 16.3 (2011), p. 030502.

[18] Julie Robic, A Nkengne, Benjamin Perret, Michel Couprie, and Hugues Talbot. "Automated quantification of the epidermal aging process using in-vivo confocal microscopy". In: *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*. IEEE. 2016, pp. 1221–1224.

[19] Kivanc Kose, Christi Alessi-Fox, Melissa Gill, Jennifer G Dy, Dana H Brooks, and Milind Rajadhyaksha. "A machine learning method for identifying morphological patterns in reflectance confocal microscopy mosaics of melanocytic skin lesions in-vivo". In: *SPIE BiOS. International Society for Optics and Photonics* (2016), pp. 968908–968908.

[20] Marco Wiltgen and Marcus Bloice. "Automatic Interpretation of Melanocytic Images in Confocal Laser Scanning Microscopy". In: *Microscopy and Analysis*. InTech, 2016.

[21] Sila Kurugol, Kivanc Kose, Brian Park, Jennifer G Dy, Dana H Brooks, and Milind Rajadhyaksha. "Automated delineation of dermal–epidermal junction in reflectance confocal microscopy image stacks of human skin". In: *Journal of Investigative Dermatology* 135.3 (2015), pp. 710–717.

[22] Julie Robic, Benjamin Perret, Alex Nkengne, Michel Couprie, and Hugues Talbot. "Classification of the dermal-epidermal junction using in-vivo confocal microscopy". In: *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on*. IEEE. 2017, pp. 252–255.

[23] Samuel C Hames, Marco Ardigò, H Peter Soyer, Andrew P Bradley, and Tarl W Prow. "Automated segmentation of skin strata in reflectance confocal microscopy depth stacks". In: *PloS one* 11.4 (2016), e0153208.

[24] M David Egger and Mojmir Petran. "New reflected-light microscope for viewing unstained brain and ganglion cells". In: *Science* 157.3786 (1967), pp. 305–307.

[25] Peter M Andrews, W Matthew Petroll, H Dwight Cavanagh, and James V Jester. "Tandem scanning confocal microscopy (TSCM) of normal and ischemic living kidneys". In: *Developmental Dynamics* 191.1 (1991), pp. 95–102.

[26] H Dwight Cavanagh, James V Jester, John Essepian, William Shields, and Michael A Lemp. "Confocal microscopy of the living eye." In: *Eye & Contact Lens* 16.1 (1990), pp. 65–73.

[27] P Corcuff, C Bertrand, and JL Leveque. "Morphometry of human epidermis in vivo by real-time confocal microscopy". In: *Archives of dermatological research* 285.8 (1993), pp. 475–481.

[28] P Corcuff and J-L Lévêque. "In vivo vision of the human skin with the tandem scanning microscope". In: *Dermatology* 186.1 (1993), pp. 50–54.

[29] Milind Rajadhyaksha, Melanie Grossman, Dina Esterowitz, Robert H Webb, and R Rox Anderson. "In vivo confocal scanning laser microscopy of human skin: melanin provides strong contrast". In: *Journal of Investigative Dermatology* 104.6 (1995), pp. 946–952.

[30] Giovanni Pellacani, Anna Maria Cesinaro, and Stefania Seidenari. "In vivo assessment of melanocytic nests in nevi and melanomas by reflectance confocal microscopy". In: *Modern pathology* 18.4 (2005), p. 469.

[31] Giovanni Pellacani, Anna Maria Cesinaro, and Stefania Seidenari. "Reflectance-mode confocal microscopy of pigmented skin lesions–improvement in melanoma diagnostic specificity". In: *Journal of the American Academy of Dermatology* 53.6 (2005), pp. 979–985.

[32] Zeina S Tannousa, Martin C Mihma, Thomas J Flottea, and Salvador Gonzálezb. "In vivo examination of lentigo maligna and malignant melanoma in situ, lentigo maligna type by near-infrared reflectance confocal microscopy: comparison of in vivo confocal images with histologic sections". In: *Journal of the American Academy of Dermatology* 46.2 (2002), pp. 260–263.

[33] G Pellacani, P Pepe, A Casari, and C Longo. "Reflectance confocal microscopy as a second-level examination in skin oncology improves diagnostic accuracy and saves unnecessary excisions: a longitudinal prospective study". In: *British Journal of Dermatology* 171.5 (2014), pp. 1044–1051.

[34] P Guitera, SW Menzies, G Argenziano, C Longo, A Losi, M Drummond, RA Scolyer, and G Pellacani. "Dermoscopy and in vivo confocal microscopy are complementary techniques for diagnosis of difficult amelanotic and light-coloured skin lesions". In: *British Journal of Dermatology* 175.6 (2016), pp. 1311–1319.

[35] H Dwight Cavanagh, W Matthew Petroll, Hassan Alizadeh, Yu-Guang He, James P McCulley, and James V Jester. "Clinical and diagnostic use of in vivo confocal microscopy in patients with corneal disease". In: *Ophthalmology* 100.10 (1993), pp. 1444–1454.

[36] Salvador Gonzalez, Robert Sackstein, R Rox Anderson, and Milind Rajadhyaksha. "Real-time evidence of in vivo leukocyte trafficking in human skin by reflectance confocal microscopy". In: *Journal of Investigative Dermatology* 117.2 (2001), pp. 384–386.

[37] Pierre Corcuff, G Gonnord, GE Pierard, and JL Lévêque. "In vivo confocal microscopy of human skin: a new design for cosmetology and dermatology". In: *Scanning* 18.5 (1996), pp. 351–355.

[38] Misbah Huzaira, Francisca Rius, Milind Rajadhyaksha, R Rox Anderson, and Salvador González. "Topographic variations in normal skin, as viewed by in vivo reflectance confocal microscopy". In: *Journal of Investigative Dermatology* 116.6 (2001), pp. 846–852.

[39] Sieglinde Neerken, Gerald W Lucassen, Marielle A Bisschop, Egbert Lenderink, and Tom AM Nuijs. "Characterization of age-related effects in human skin: a comparative study that applies confocal laser scanning microscopy and optical coherence tomography". In: *Journal of biomedical optics* 9.2 (2004), pp. 274–281.

[40] Fabia Cristina Rossetti, Livia Vieira Depieri, and Maria Vitoria Lopes Badra Bentley. "Confocal laser scanning microscopy as a tool for the investigation of skin drug delivery systems and diagnosis of skin disorders". In: *Confocal laser Microscopy-Principles and Applications in Medicine, Biology, and the Food Sciences*. InTech, 2013.

[41] Pierre Corcuff, Olivier de Lacharrière, and Jean-Luc Lévêque. "Extension-induced changes in the microrelief of the human volar forearm: variations with age". In: *Journal of gerontology* 46.6 (1991), pp. M223–M227.

[42] C Trojahn, G Dobos, M Schario, L Ludriksone, U Blume-Peytavi, and J Kottner. "Relation between skin micro-topography, roughness, and skin age". In: *Skin Research and Technology* 21.1 (2015), pp. 69–75.

[43] JM Lagarde, C Rouvrais, and D Black. "Topography and anisotropy of the skin surface with ageing". In: *Skin Research and Technology* 11.2 (2005), pp. 110–119.

[44] WDAHI Manuskiatti, DA Schwindt, and HI Maibach. "Influence of age, anatomic site and race on skin roughness and scaliness". In: *Dermatology* 196.4 (1998), pp. 401–407.

[45] CO Barland, PM Elias, and R Ghadially. "The aged epidermal permeability barrier: basis for functional abnormalities". In: *Skin Barrier. New York, Taylor & Francis* 53552 (2005).

[46] Miranda A Farage, Kenneth W Miller, and Howard I Maibach. "Degenerative changes in aging skin". In: *Textbook of aging skin*. Springer, 2015, pp. 1–18.

[47] Sophie Garrido Lagarrigue, Jerome George, Emmanuel Questel, Christophe Lauze, Nicolas Meyer, Jean-Michel Lagarde, Michel Simon, Anne-Marie Schmitt, Guy Serre, and Carle Paul. "In vivo quantification of epidermis pigmentation and dermis papilla density with reflectance confocal microscopy: variations with age and skin phototype". In: *Experimental dermatology* 21.4 (2012), pp. 281–286.

[48] Robert M Lavker, Peishu Zheng, and Gang Dong. "Aged skin: a study by light, transmission electron, and scanning electron microscopy." In: *Journal of investigative dermatology* 88 (1987).

[49] Laure Rittié and Gary J Fisher. "Natural and sun-induced aging of human skin". In: *Cold spring harbor perspectives in medicine* 5.1 (2015), a015370.

[50] Kirsten Sauermann, Sven Clemann, Sören Jaspers, Thilo Gambichler, Peter Altmeyer, Klaus Hoffmann, and Joackim Ennen. "Age related changes of human skin investigated with histometric measurements by confocal laser scanning microscopy in vivo". In: *Skin research and Technology* 8.1 (2002), pp. 52–56.

[51] Richard G.B. Langley, Elizabeth Burton, Noreen Walsh, Iva Propperova, and Scott J. Murray. "In vivo confocal scanning laser microscopy of benign lentigines: Comparison to conventional histology and in vivo characteristics of lentigo maligna". In: *Journal of the American Academy of Dermatology* 55.1 (2006), pp. 88 –97. ISSN: 0190-9622. DOI: https://doi.org/10.1016/j.jaad.2006.03.009. URL: http://www.sciencedirect.com/science/article/pii/S0190962206007997.

[52] K Mizukoshi, K Yonekura, M Futagawa, T Nakamura, K Hirayama, and K Takahashi. "Changes in dermal papilla structures due to aging in the facial cheek region". In: *Skin Research and Technology* 21.2 (2015), pp. 224–231.

[53] James Varani, Michael K Dame, Laure Rittie, Suzanne EG Fligiel, Sewon Kang, Gary J Fisher, and John J Voorhees. "Decreased collagen production in chronologically aged skin: roles of age-dependent alteration in fibroblast function and defective mechanical stimulation". In: *The American journal of pathology* 168.6 (2006), pp. 1861–1868.

[54] Jean C Serra and Philippe Salembier. "Connected operators and pyramids". In: *SPIE's 1993 International Symposium on Optics, Imaging, and Instrumentation*. International Society for Optics and Photonics. 1993, pp. 65–76.

[55] Henk JAM Heijmans. "Connected morphological operators for binary images". In: *Computer Vision and Image Understanding* 73.1 (1999), pp. 99–120.

[56] Philippe Salembier and Michael HF Wilkinson. "Connected operators". In: *IEEE Signal Processing Magazine* 26.6 (2009).

[57] Philippe Salembier and Jean Serra. "Flat zones filtering, connected operators, and filters by reconstruction". In: *IEEE Transactions on image processing* 4.8 (1995), pp. 1153–1160.

[58] Philippe Salembier, Albert Oliveras, and Luis Garrido. "Antiextensive connected operators for image and sequence processing". In: *IEEE Transactions on Image Processing* 7.4 (1998), pp. 555–570.

[59] Jean Serra. "A lattice approach to image segmentation". In: *Journal of Mathematical Imaging and Vision* 24.1 (2006), pp. 83–130.

[60] Christian Ronse. "Partial partitions, partial connections and connective segmentation". In: *Journal of Mathematical Imaging and Vision* 32.2 (2008), pp. 97–125.

[61] Luc Vincent. "Morphological area openings and closings for grey-scale images". In: *Shape in Picture*. Springer, 1994, pp. 197–208.

[62] Ronald Jones. "Connected filtering and segmentation using component trees". In: *Computer Vision and Image Understanding* 75.3 (1999), pp. 215–228.

[63] Benjamin Perret, Jean Cousty, Olena Tankyevych, Hugues Talbot, and Nicolas Passat. "Directed connected operators: Asymmetric hierarchies for image filtering and segmentation". In: *IEEE transactions on pattern analysis and machine intelligence* 37.6 (2015), pp. 1162–1176.

[64] Benjamin Perret and Ch Collet. "Connected image processing with multivariate attributes: An unsupervised Markovian classification approach". In: *Computer Vision and Image Understanding* 133 (2015), pp. 1–14.

[65] Carlo Tomasi and Roberto Manduchi. "Bilateral filtering for gray and color images". In: *Computer Vision, 1998. Sixth International Conference on*. IEEE. 1998, pp. 839–846.

[66] Nobuyuki Otsu. "A threshold selection method from gray-level histograms". In: *IEEE transactions on systems, man, and cybernetics* 9.1 (1979), pp. 62–66.

[67] Luc Vincent and Barry Masters. "Morphological image processing and network analysis of cornea endothelial cell images". In: *Image Algebra and Morphological Image Processing III* 1769 (1992), pp. 212–226.

[68] Yann Gavet and Jean-Charles Pinoli. "Visual perception based automatic recognition of cell mosaics in human corneal endotheliummicroscopy images". In: *Image Analysis & Stereology* 27.1 (2011), pp. 53–61.

[69] Bettina Selig, Koenraad A Vermeer, Bernd Rieger, Toine Hillenaar, and Cris L Luengo Hendriks. "Fully automatic evaluation of the corneal endothelium from in vivo confocal microscopy". In: *BMC medical imaging* 15.1 (2015), p. 13.

[70] Luc Vincent and Pierre Soille. "Watersheds in digital spaces: an efficient algorithm based on immersion simulations". In: *IEEE Transactions on Pattern Analysis & Machine Intelligence* 6 (1991), pp. 583–598.

[71] Fernand Meyer. "Un algorithme optimal de ligne de partage des eaux". In: *Actes du* 2 (1991), pp. 847–859.

[72] Serge Beucher and Fernand Meyer. "The morphological approach to segmentation: the watershed transformation". In: *Optical Engineering-New York-Marcel Dekker Incorporated-* 34 (1992), pp. 433–433.

[73] Jean Cousty, Gilles Bertrand, Laurent Najman, and Michel Couprie. "Watershed cuts: Minimum spanning forests and the drop of water principle". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31.8 (2009), pp. 1362–1374.

[74] Jesus Angulo and Sabine Matou. "Automatic quantification of in vitro endothelial cell networks using mathematical morphology". In: *Proceedings of the 5th IASTED International Conference on Visualization, Imaging, and Image Processing (VIIP'05).* 2005, pp. 51–56.

[75] Per-Erik Danielsson. "Euclidean distance mapping". In: *Computer Graphics and image processing* 14.3 (1980), pp. 227–248.

[76] Jesús Angulo and Dominique Jeulin. "Stochastic watershed segmentation". In: *PROC. of the 8th International Symposium on Mathematical Morphology.* 2007, pp. 265–276.

[77] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. "Snakes: Active contour models". In: *International journal of computer vision* 1.4 (1988), pp. 321–331.

[78] Kamil Charłampowicz, Daniel Reska, and C Bołdak. "Automatic segmentation of corneal endothelial cells using active contours". In: *Advances in Computer Science Research* (2014).

[79] Marco Foracchia and Alfredo Ruggeri. "Corneal endothelium cell field analysis by means of interacting bayesian shape models". In: *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE.* IEEE. 2007, pp. 6035–6038.

[80] Thomas Dyhre Nielsen and Finn Verner Jensen. *Bayesian networks and decision graphs.* Springer Science & Business Media, 2009.

[81] Gilles Bertrand. "On topological watersheds". In: *Journal of Mathematical Imaging and Vision* 22.2-3 (2005), pp. 217–230.

[82] Dipika V Patel and Charles N McGhee. "Quantitative analysis of in vivo confocal microscopy images: a review". In: *Survey of ophthalmology* 58.5 (2013), pp. 466–475.

[83] Leo Breiman. "Random Forests". In: *Machine Learning* 45.1 (2001), pp. 5–32. ISSN: 1573-0565. DOI: 10.1023/A:1010933404324.

[84] Sindhu Ghanta, Michael I Jordan, Kivanc Kose, Dana H Brooks, Milind Rajadhyaksha, and Jennifer G Dy. "A Marked Poisson Process Driven Latent Shape Model for 3D Segmentation of Reflectance Confocal Microscopy Image Stacks of Human Skin". In: *IEEE Transactions on Image Processing* 26.1 (2017), pp. 172–184.

[85] S. C Hames, M. Ardigò, H P. Soyer, A. P Bradley, and T. W Prow. "Automated segmentation of skin strata in reflectance confocal microscopy depth stacks". In: *PloS one* 11.4 (2016), e0153208.

[86] Abdelghafour Halimi, Hadj Batatia, Jimmy Le Digabel, Gwendal Josse, and Jean-Yves Tourneret. "An unsupervised bayesian approach for the joint reconstruction and classification of cutaneous reflectance confocal microscopy images". In: *arXiv preprint arXiv:1703.01444* (2017).

[87] Eduardo Somoza, Gabriela Oana Cula, Catherine Correa, and Julie B Hirsch. "Automatic localization of skin layers in reflectance confocal microscopy". In: *International Conference Image Analysis and Recognition*. Springer. 2014, pp. 141–150.

[88] Parneet Kaur, Kristin J Dana, Gabriela Oana Cula, and M Catherine Mack. "Hybrid deep learning for Reflectance Confocal Microscopy skin images". In: *Pattern Recognition (ICPR), 2016 23rd International Conference on*. IEEE. 2016, pp. 1466–1471.

[89] Alican Bozkurt, Trevor Gale, Kivanc Kose, Christi Alessi-Fox, Dana H Brooks, Milind Rajadhyaksha, and Jennifer Dy. "Delineation of Skin Strata in Reflectance Confocal Microscopy Images With Recurrent Convolutional Networks". In: *Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on*. IEEE. 2017, pp. 777–785.

[90] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. "Imagenet classification with deep convolutional neural networks". In: *Advances in neural information processing systems*. 2012, pp. 1097–1105.

[91] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 770–778.

[92] Daphne Koller and Nir Friedman. *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

[93] Charles Sutton, Andrew McCallum, et al. "An introduction to conditional random fields". In: *Foundations and Trends® in Machine Learning* 4.4 (2012), pp. 267–373.

[94] John Lafferty, Andrew McCallum, Fernando Pereira, et al. "Conditional random fields: Probabilistic models for segmenting and labeling sequence data". In: (2001).

[95] Fuchun Peng and Andrew McCallum. "Information extraction from research papers using conditional random fields". In: *Information processing & management* 42.4 (2006), pp. 963–979.

[96] Fei Sha and Fernando Pereira. "Shallow parsing with conditional random fields". In: *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1*. Association for Computational Linguistics. 2003, pp. 134–141.

[97] Yan Liu, Jaime Carbonell, Peter Weigele, and Vanathi Gopalakrishnan. "Protein fold recognition using segmentation conditional random fields (SCRFs)". In: *Journal of Computational Biology* 13.2 (2006), pp. 394–406.

[98] Xuming He, Richard S Zemel, and Miguel Á Carreira-Perpiñán. "Multiscale conditional random fields for image labeling". In: *Computer vision and pattern recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE computer society conference on*. Vol. 2. IEEE. 2004, pp. II–II.

[99] Andreas C Müller and Sven Behnke. "Learning depth-sensitive conditional random fields for semantic segmentation of rgb-d images". In: *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE. 2014, pp. 6232–6237.

[100] Stefan Bauer, Lutz-P Nolte, and Mauricio Reyes. "Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2011, pp. 354–361.

[101] Patrick Ferdinand Christ, Mohamed Ezzeldin A Elshaer, Florian Ettlinger, Sunil Tatavarty, Marc Bickel, Patrick Bilic, Markus Rempfler, Marco Armbruster, Felix Hofmann, Melvin D'Anastasi, et al. "Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2016, pp. 415–423.

[102] S. Kumar and M. Hebert. "Discriminative random fields". In: *International Journal of Computer Vision* 68.2 (2006), pp. 179–201.

[103] van A Van der Schaaf and JH van van Hateren. "Modelling the power spectra of natural images: statistics and information". In: *Vision research* 36.17 (1996), pp. 2759–2770.

[104] Robert M Haralick. "Statistical and structural approaches to texture". In: *Proceedings of the IEEE* 67.5 (1979), pp. 786–804.

[105] Thomas P Weldon, William E Higgins, and Dennis F Dunn. "Efficient Gabor filter design for texture segmentation". In: *Pattern recognition* 29.12 (1996), pp. 2005–2015.

[106] Anil K Jain, Nalini K Ratha, and Sridhar Lakshmanan. "Object detection using Gabor filters". In: *Pattern recognition* 30.2 (1997), pp. 295–309.

[107] A Ahmadian and A Mostafa. "An efficient texture classification algorithm using Gabor wavelet". In: *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*. Vol. 1. IEEE. 2003, pp. 930–933.

[108] David Marr and Ellen Hildreth. "Theory of edge detection". In: *Proc. R. Soc. Lond. B* 207.1167 (1980), pp. 187–217.

[109] Oliver Kramer. "Iterated local search with Powell's method: a memetic algorithm for continuous global optimization". In: *Memetic Computing* 2.1 (2010), pp. 69–83.

[110] S. Kosov. *Direct Graphical Models C++ library*. http://research.project-10.de/dgm/. 2013.

[111] VL Newton, RS Bradley, P Seroul, M Cherel, CEM Griffiths, AV Rawlings, R Voegeli, REB Watson, and MJ Sherratt. "Novel approaches to characterize age-related remodelling of the dermal-epidermal junction in 2D, 3D and in vivo". In: *Skin Research and Technology* (2016).

[112] Philippe Salembier and Luis Garrido. "Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval". In: *IEEE transactions on Image Processing* 9.4 (2000), pp. 561–576.

[113] Michael HF Wilkinson and Michel A Westenberg. "Shape preserving filament enhancement filtering". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2001, pp. 770–777.

[114] Michel A Westenberg, Jos BTM Roerdink, and Michael HF Wilkinson. "Volumetric attribute filtering and interactive visualization using the max-tree representation". In: *IEEE Transactions on Image Processing* 16.12 (2007), pp. 2943–2952.

[115] Bin Luo and Liangpei Zhang. "Robust autodual morphological profiles for the classification of high-resolution satellite images". In: *IEEE Transactions on Geoscience and Remote Sensing* 52.2 (2014), pp. 1451–1462.

[116] Erik R Urbach, Jos BTM Roerdink, and Michael HF Wilkinson. "Connected shape-size pattern spectra for rotation and scale-invariant classification of gray-scale images". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29.2 (2007), pp. 272–285.

[117] B Ravi Kiran and Jean Serra. "Global–local optimizations by hierarchical cuts and climbing energies". In: *Pattern Recognition* 47.1 (2014), pp. 12–24.

[118] Petros Maragos. "Pattern spectrum and multiscale shape representation". In: *IEEE Transactions on pattern analysis and machine intelligence* 11.7 (1989), pp. 701–716.

[119] Mauro Dalla Mura, Jón Atli Benediktsson, Björn Waske, and Lorenzo Bruzzone. "Morphological attribute profiles for the analysis of very high resolution images". In: *IEEE Transactions on Geoscience and Remote Sensing* 48.10 (2010), pp. 3747–3762.

[120] Mauro Dalla Mura, Jon Atli Benediktsson, and Lorenzo Bruzzone. "Self-dual attribute profiles for the analysis of remote sensing images". In: *International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing*. Springer. 2011, pp. 320–330.

[121] Yuqing Song. "A topdown algorithm for computation of level line trees". In: *IEEE Transactions on Image Processing* 16.8 (2007), pp. 2107–2116.

[122] Pascal Monasse and Frederic Guichard. "Fast computation of a contrast-invariant image representation". In: *IEEE Transactions on Image Processing* 9.5 (2000), pp. 860–872.

[123] Thierry Géraud, Edwin Carlinet, Sébastien Crozet, and Laurent Najman. "A quasi-linear algorithm to compute the tree of shapes of nD images". In: *International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing*. Springer. 2013, pp. 98–110.

[124] Philippe Salembier. "Connected operators based on region-trees". In: *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*. IEEE. 2008, pp. 2176–2179.

[125] Benjamin Perret, Jean Cousty, Olena Tankyevych, Hugues Talbot, and Nicolas Passat. "Directed connected operators: asymmetric hierarchies for image filtering and segmentation". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 37.6 (2015), pp. 1162–1176. DOI: 10.1109/TPAMI.2014.2366145.

[126] G Pellacani, M Ulrich, A Casari, TW Prow, F Cannillo, E Benati, A Losi, AM Cesinaro, C Longo, G Argenziano, et al. "Grading keratinocyte atypia in actinic keratosis: a correlation of reflectance confocal microscopy and histopathology". In: *Journal of the European Academy of Dermatology and Venereology* 29.11 (2015), pp. 2216–2221.

[127] Charles Kervrann and Alain Trubuil. "An adaptive window approach for poisson noise reduction and structure preserving in confocal microscopy". In: *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on*. IEEE. 2004, pp. 788–791.

[128] Martin Laasmaa, Marko Vendelin, and Pearu Peterson. "Application of regularized Richardson–Lucy algorithm for deconvolution of confocal microscopy images". In: *Journal of microscopy* 243.2 (2011), pp. 124–140.

[129] M Von Tiedemann, Anders Fridberger, M Ulfendahl, and J Boutet De Monvel. "Image adaptive point-spread function estimation and deconvolution for in vivo confocal microscopy". In: *Microscopy research and technique* 69.1 (2006), pp. 10–20.

[130] Thorsten Bergmann, Sebastian Beer, Ulf Maeder, Jan M Burg, Peggy Schlupp, Thomas Schmidts, Frank Runkel, and Martin Fiebich. "Development of a skin phantom of the epidermis and evaluation by using fluorescence techniques". In: *Proc. SPIE*. Vol. 7906. 2011, 79060T.

[131] Giovanni Zaccanti, Adriana Taddeucci, Marco Barilli, Piero Bruscaglioni, and Fabrizio Martelli. "Optical properties of biological tissues". In: *Optical Tomography, Photon Migration, and Spectroscopy of Tissue and Model Media: Theory, Human Studies, and Instrumentation*. Vol. 2389. International Society for Optics and Photonics. 1995, pp. 513–522.