Computing methods for facial aging prevention and prediction

Victor Martin

THESE DE DOCTORAT DE

CENTRALESUPELEC
COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 601
Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité : Signal, Image, Vision

Par
Victor MARTIN

Computing methods for facial aging prevention and prediction

Méthodes informatiques pour la prévention et la prédiction du vieillissement du visage

Thèse présentée et soutenue à « Neuilly-sur-Seine », le « lundi 11 février 2019 »
Unité de recherche : IETR UMR CNRS 6164 – Equipe FAST CentraleSupélec Rennes
Thèse N° : 2019-03-TH

Rapporteurs avant soutenance :

Alice Caplier	Professeur des Universités Grenoble-INP
Jean-Luc Dugelay	Professeur des Universités EURECOM

Composition du Jury :

Président :	Luce Morin	Professeur des Universités INSA Rennes
Examinateur :	Lionel Prévost	Professeur des Universités ESIEA
Dir. de thèse :	Renaud Séguier	Professeur des Universités CentraleSupélec Rennes
Co-dir. de thèse :	Aurélie Porcheron	Chargé de projet CHANEL Parfums Beauté

Titre : Méthodes informatiques pour la prévention et la prédiction du vieillissement du visage

Mots clés : Modèle actif d'apparence ; Réseau neuronal convolutionnel ; Vieillissement du visage ; Progression de l'âge ; Modélisation des rides ; Perception de la santé

Résumé : L'utilisation de la simulation informatique pour comprendre comment les visages humains vieillissent est un domaine de recherche en pleine croissance depuis des décennies. Cela a été appliqué à la recherche d'enfants disparus ainsi qu'aux domaines du divertissement, des cosmétiques et de la recherche en dermatologie. Notre objectif est de modéliser les changements des traits du visage liés à l'âge, afin de mieux les prédire.
Dans ce travail, une nouvelle perspective pour faire vieillir un visage est proposée : un modèle actif d'apparence axé sur les rides. Tout d'abord, les visages sont décomposés en termes d'apparence et de forme à l'aide d'un modèle actif d'apparence. Ensuite, les rides de chaque visage sont transformées en paramètres d'apparence et de forme. Une façon nouvelle et efficace de modéliser la distribution des paramètres des rides dans un visage est introduite.
Il est démontré que les visages artificiellement vieillis produits par le système influencent mieux la perception de l'âge que ceux produits par deux autres systèmes. Cet outil est une première étape dans la construction d'un système de vieillissement du visage plus précis.
En outre, une nouvelle méthode d'estimation de la santé utilisant un réseau neuronal convolutionnel est proposée. Ce système est capable de reproduire le jugement humain dans l'évaluation de la santé perçue. Il est présenté comment cet outil utilise les mêmes traits du visage que l'humain pour effectuer sa prédiction. Enfin, l'impact de caractéristiques faciales spécifiques jamais étudié auparavant sur la perception de la santé est établi.

Title: Computing methods for facial aging prevention and prediction

Keywords: Active Appearance Model; Convolutional Neural Network; Facial Aging; Age Progression; Modeling of wrinkles; Health Perception

Abstract: The use of computer simulation to understand how human faces age has been a growing area of research since decades. It has been applied to the search for missing children as well as to the fields of entertainment, cosmetics and dermatology research. Our objective is to elaborate a model for the age-related changes of facial cues which affect the perception of age, so that we may better predict them.
In this work, a new framework to make a face age is proposed: Wrinkle Oriented Active Appearance Model. First, faces are decomposed in terms of appearance and shape using Active Appearance Model. In addition, wrinkles in each face are transformed in appearance and shape parameters.
A new effective way to model the distribution of wrinkle parameters in a face is introduced. Finally, it is shown that artificially aged faces produced by the system better influence age perception than those produced by two other systems. This framework is a first step in the construction of a more accurate facial aging system.
In addition, a new health estimation system using a convolutional neural network is introduced. This system is able to estimate how a face is perceived in terms of health by humans. It is shown how this tool reacts in the same way as health perception by humans. Finally, the impact of specific facial features on health perception that have never been studied before is established.

Remerciements

Je tiens à exprimer ma profonde gratitude à mon directeur de thèse Renaud Séguier pour sa gestion exemplaire de cette thèse, son dynamisme, son pragmatisme et son optimisme débordant.

Je tiens à remercier également ma directrice de thèse Aurélie Porcheron pour sa bienveillance, sa disponibilité, son optimisme, et sa capacité à faire confiance.

Merci à mes rapporteurs Alice Caplier et Jean-Luc Dugelay ainsi qu'à mes examinateurs Luce Morin et Lionel Prévost d'avoir acceptés de juger cette thèse et pour leurs commentaires pertinents sur ce travail de recherche.

Je souhaite remercier toutes les personnes avec qui j'ai travaillé pendant mes 3 ans chez CHANEL à Pantin. Je tiens à remercier plus particulièrement mes collègues Randa, Sandra, Gabriel, Guylaine, Astrid, Jonathan, Andreia, Catherine, Caroline, François, Emma, Gaëlle, Sandra, Irina et Kim qui ont contribué à créer un environnement très stimulant et une bonne ambiance au travail.
Un remerciement tout particulier pour l'unité de statistiques Julie, Manue et Frédérique avec qui j'ai partagé de bons moments, amicaux et sportifs, ainsi que des discussions très instructives sur l'art de la statistique.
Merci à Frédérique Morizot pour m'avoir permis d'effectuer cette thèse dans son équipe et d'avoir su détecter le potentiel de l'analyse d'images appliquée au domaine de la cosmétique.

Je remercie Catherine Soladié de l'équipe FAST ainsi que tous les doctorants avec qui j'ai pu avoir de nombreux échanges dynamisants pendant ces trois ans.

Enfin, je ne pourrais terminer ces remerciements sans penser à ma famille, et plus particulièrement à mes parents dont le soutien, les encouragements et l'éducation que j'ai reçu ont contribué à l'aboutissement de ce travail. Merci à France pour son soutien inconditionnel et son aide précieuse pendant cette thèse.

Résumé

Dans ce travail, nous avons étudié la problématique de la simulation du vieillissement du visage. Une application de vieillissement artificiel du visage doit tenir compte des changements morphologiques liés à l'âge ainsi que des modifications de l'apparence de la peau afin d'obtenir des résultats réalistes. Le changement le plus important du visage avec l'âge est morphologique et résulte de la croissance du visage ; il se produit de la naissance jusqu'au début de l'âge adulte. Une autre modification morphologique liée à l'âge concerne les volumes du visage sous la peau dus aux variations de répartition des graisses ; ces volumes varient tout au long de la vie, de la naissance à la fin de l'âge adulte. Au cours de l'âge adulte, la peau du visage subit également des changements avec l'âge : l'apparition de rides, de l'affaissement et une augmentation des irrégularités pigmentaires. Toutes ces caractéristiques sont liées à l'âge de la peau et impactent la perception de l'âge du visage chez les adultes. Notre objectif est de modéliser les changements des traits du visage liés à l'âge, afin de mieux les prédire.

Dans la section suivant l'introduction, nous avons effectué un état de l'art des différentes méthodes de vieillissement du visage. Nous montrons que la simulation du vieillissement du visage peut être réalisée en utilisant une grande variété de techniques, mélangeant traitement d'images, statistiques, optimisation et réseaux neuronaux profonds. Les premières méthodes ont introduit des déformations de formes de visages directement inspirées de l'anthropométrie.
Des approches plus récentes ont proposé d'intégrer l'apparence à la forme. L'application de méthodes statistiques à l'image (analyses en composantes principales et régressions pour le modèle actif d'apparence) ont permis de ne pas avoir à intégrer beaucoup de connaissances *a priori* pour vieillir artificiellement un visage, ce qui est souhaitable pour comprendre le processus de vieillissement.
La plupart des méthodes développées par la suite se sont concentrées sur l'intégration des détails de texture dans la modélisation, en prenant comme base un modèle actif d'apparence, un prototype ou un réseau neuronal profond.

Dans cette thèse, nous avons proposé *WOAAM*, *Wrinkle Oriented Active Appearance Model*, un modèle actif d'apparence qui intègre explicitement les rides du visage (Fig. 1).
Tout d'abord, les visages sont décomposés en termes d'apparence et de forme à l'aide d'un modèle actif d'apparence. Les formes sont alignées sur la forme moyenne par transformation affine (analyse Procrustéenne). Les pixels du visage sont quant à eux déformés (*warped*) dans la forme moyenne par une transforma-

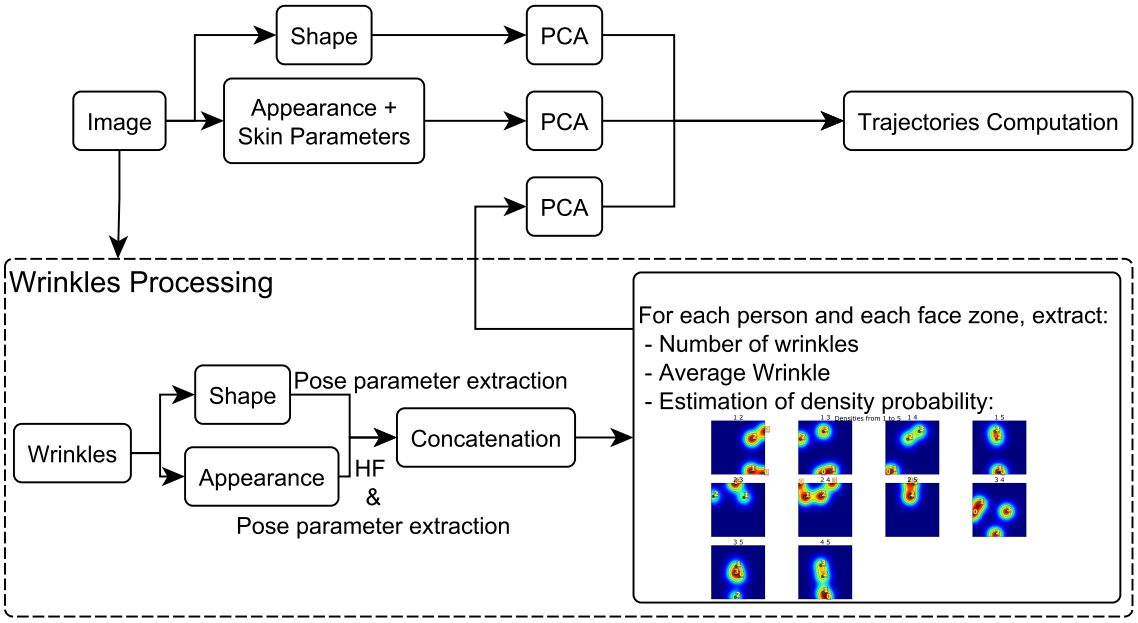

FIGURE 1 – Schéma du *Wrinkle Oriented AAM*.

tion affine par morceau (*Piecewise Affine Transform*). Une analyse en composantes principales (ACP) est ensuite effectuée pour la forme, et une autre pour l'apparence. Les poids de sortie de la forme et de l'apparence sont concaténés.
Dans *WOAAM*, en plus de la forme et de l'apparence, un troisième canal est rajouté pour les rides. Les rides de chaque visage sont transformées en paramètres d'apparence et de forme. Nous proposons 5 paramètres de forme définis tel que :

— les deux paramètres (c_x, c_y) représentent le centre de la ride

— la taille ℓ qui est égale à la distance géodésique entre le premier point et le dernier point de l'annotation de la ride

— l'angle a en degrés

— la courbure $\mathcal{C}$ calculée comme une minimisation aux moindres carrées de :

$$\min \parallel Y - \mathcal{C}X^2 \parallel_2^2 \tag{1}$$

avec Y (resp. X) l'ordonnée (resp. l'abscisse) de la ride centrée sur l'origine, et dont le premier et le dernier point sont alignés horizontalement.

Pour modéliser l'apparence d'une ride, un filtre à différence de Gaussiennes est utilisé pour ne conserver que les hautes fréquences, puis le résultat est converti du RGB vers des niveaux de gris. Ensuite, la ride est déformée (*warped*) dans une forme de ride de référence. Enfin, une fonction de Lorentz dérivée seconde définit dans l'équation 2 est adaptée (*fitted*) sur chaque colonne de la ride.

$$A * \frac{2\sigma\left(3\left(x-\mu\right)^2-\sigma^2\right)}{\left(\left(x-\mu\right)^2+\sigma^2\right)^3} + o \tag{2}$$

Les paramètres μ et o sont des paramètres de translations horizontale et verticale de la courbe utile à l'optimisation ; ils ne seront pas conservés à l'issue celle-ci. Les paramètres A et σ représentent respectivement l'intensité et la largeur de la ride. Les paramètres obtenus sur chaque colonne sont moyennés pour obtenir deux paramètres A et σ représentant l'intensité et la largeur moyennes de la ride.
Ainsi, une ride sera codée par 7 paramètres : $(c_x, c_y, \ell, a, \mathcal{C}, A, \sigma)$.

La répartition de ces paramètres de rides dans un visage et dans chaque zone du visage est codée par plusieurs informations :

— le nombre de rides

— la ride moyenne

— un ensemble de densités de probabilité jointes

Cet ensemble contient toutes les densités de probabilité jointe des dimensions prises deux à deux $\{P(c_x, c_y), P(c_x, \ell), \ldots, P(A, \sigma)\}$, déterminées par estimation par noyau (*Kernel Density Estimation*). Cet ensemble ainsi qui le nombre de rides et la ride moyenne définissent une représentation efficace des rides dans une zone du visage. Le même procédé est réalisé pour toutes les zones du visage et les représentations produites sont concaténées pour obtenir une représentation des rides au sein d'un visage. Une analyse en composantes principales est ensuite effectuée sur ces représentations (une par visage) pour ensuite être "reliée" au modèle actif d'apparence. Ainsi, en plus des deux canaux de forme et d'apparence de l'AAM, ce travail intègre un troisième canal dédié à l'intégration des rides. L'espace WOAAM résultant modélise les variations du visage en forme, apparence, et ride.
Nous avons introduit également une légère modification du canal d'apparence. Le visage est découpé en blocs et en bandes de fréquences avant l'analyse en composantes principales. Cette altération permet d'intégrer la micro texture et les pores de la peau dans le système de vieillissement.
Ainsi à chaque visage de la base correspond un point dans l'espace final. Pour réaliser un réel système de vieillissement, une régression cubique f est effectuée entre les poids de l'analyse en composantes principales, et les âges perçus correspondant. Pour faire vieillir un visage X_i de la base, il est projeté dans l'espace ACP final :

$$W_i = (X_i - \bar{X})C^T \tag{3}$$

Les poids subissent une translation Δ_{age} sur la trajectoire de vieillissement f :

$$\hat{W}_i = W_i + (f^{-1}(age + \Delta_{age}) - f^{-1}(age)) \tag{4}$$

Dans cette équation f^{-1} est calculée par simulation de Monte Carlo. Enfin, le point synthétique de l'espace final est reconstruit en forme, apparence, et rides.

$$\hat{X}_i = \hat{W}_i C + \bar{X} \tag{5}$$

Les résultats expérimentaux montrent que le système peut générer des visages réalistes pour le vieillissement et le rajeunissement, et que les visages vieillis et

rajeunis avec notre système influencent mieux la perception de l'âge que ceux produits avec deux autres systèmes.

Enfin, nous nous sommes intéressés à la façon dont un visage est perçu, non plus en termes d'âge, mais en termes de santé. Nous avons développé un système pour prédire comment un visage est perçu par d'autres personnes en termes de santé (Fig. 2). Le système est composé d'un réseau de neurones convolutionnel pré entrainé pour l'estimation de l'âge biologique qui va générer une représentation peu sensible aux variations de pose et d'illumination à partir d'une photo de visage. Ensuite une régression de type Ridge est utilisée pour prédire un *score de santé perçue* à partir d'une représentation générée à l'étape précédente. Les résultats montrent comment un tel système peut être utilisé pour comprendre plus en détail la perception humaine de la santé. De plus, il est montré en conclusion comment le système peut être utilisé pour évaluer des effets de maquillages sur la perception humaine de l'âge et de la santé.

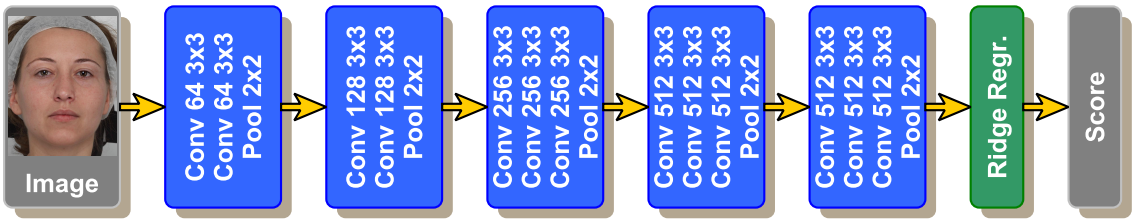

FIGURE 2 – Chaîne de calcul pour estimer un score de santé.

Publications

Journal Article

V. Martin, R. Séguier, A. Porcheron, and F. Morizot
Face aging simulation with a new wrinkle oriented active appearance model.
Multimedia Tools and Applications, **Published**, Pp. 1–19. 2018

Patents

V. Martin, A. Porcheron, and R. Séguier
European Patent *EP 17 306 901.4*. **Filed**, 2017

V. Martin, R. Séguier, and A. Porcheron
European Patent *EP 18 305 211.7*. **Filed**, 2018

V. Martin
European Patent *EP 18 306 600.0*. **Filed**, 2018

International Conference Articles

V. Martin, R. Seguier, A. Porcheron, and F. Morizot
Towards continuous health diagnosis from faces with deep learning.
In *Workshop PRIME @ MICCAI*. **Published**, Pp. 120–128. 2018

V. Martin, R. Seguier, A. Porcheron, and F. Morizot
Understanding the visual effects of cosmetic products on beauty via deep learning.
In *IFSCC*. **Published**, 2018

National Conference

V. Martin
Perception des visages et intelligence artificielle : application à la cosmétique.
In *Journée Cosmétique Augmentée par la Société Francophone d'ingénierie et d'imagerie Cutanée*. **Oral**, 2018

Contents

Introduction

Context & Motivation

The entity *Chanel Recherche & Technologie* from the *Chanel Parfums Beauté* fuels the scientific credibility of the brand and innovates in many fields. Within the entity, the *Biology & Women's Beauty* (BWB) department has three goals:

— conduct innovative research projects to develop high-level and original knowledge on skin and women's beauty,

— offer innovative cosmetic solutions (products, services, accessories) that meet women's needs,

— contribute to *Chanel Parfums Beauté* scientific credibility.

Among its main research themes, the department conducts multidisciplinary works on aging, to support skincare and cosmetic products development and communication. This work brings together scientists from different fields: biophysics, epidemiology, genetics, biology, and psychology. It is performed in multiple countries across the globe and focus on studying the aging of women's face, and more specifically the aging of facial skin. This works on aging spans from the understanding of the biological mechanisms involved to its clinical description, from the understanding of the influence of ethnics and environmental factors, to its consequences for face perception.

Aging of the skin and face in general is subject to genetic, morphological, and environmental factors.

Genetic Multiple studies conducted on several countries have shown differences due to ethnicities with regard to the age at which signs of aging appear (e.g. wrinkles, dark spots, facial sagging), and the severity of these signs. Nevertheless, the signs of skin aging are the same for all women, no matter their origins or environments (Guinot et al., 2006). Genome Wide Association Studies (GWAS) conducted by the BWB have shown the impact of the gene receptor of the melanocortine-1 (MC1R) on the severity of skin aging (Latreille et al., 2009).

Morphological Also, it appears that facial morphology – which differs between ethnic groups – partly explains the differences in skin aging (Farkas et al., 2004). For example, Asian female faces retain a so-called "babyface" appearance during aging unlike Caucasian female faces (Shirakabe et al., 2003). The skeleton of the face of Asian women helps to support the upper lip and avoids the sagging of this area with aging. Hence, the lips of these women maintain a relatively constant thickness with age; which is not the case concerning Caucasian women's

lips which lengthen and narrow with age.
Environmental Studies carried out by Chanel R&T, as others in the literature have highlighted the influence of extrinsic factors such as access to health care, depression, marital status, education and social status, on the appearance and evolution of facial cues with aging (Mayes et al., 2010; Rexbye et al., 2006) Among the most studied factors, there are sun exposure over the lifetime and the use of UV protection, but also smoking (Malvy et al., 2000; Rexbye et al., 2006). These studies show that aging and environmental factors influence the appearance of signs of aging, but also the change of apparent age (or perceived age) of a face.

Indeed, in addition to the biological age, the age perceived from a face by humans is measured. Looking old or young for one's age – perceived age - is an accurate measurement of the global biological aging process and is correlated with health condition (Christensen et al., 2004; Gunn et al., 2008, 2009). The perceived age, controlled for chronological age and sex, is also correlated with a molecular biomarker of aging, the leucocyte telomere length; and is a good predictor of longevity (Christensen et al., 2004, 2009). A methodology has been published and widely used in the literature to generate perceived age as a biomarker of facial aging. The methodology uses facial images of subjects to be presented to a group of "naive" age assessors, and show that the generated perceived age can be a reproducible measure to investigate facial aging.
Linked to age perception, the BWB is involved since years in researches on human perception of health from faces. Recent studies show that health perception from faces by humans is a good predictor of good health and healthy behaviors (Whitehead et al., 2012; Zebrowitz et al., 2014). People who are judged healthy are more attractive and receive more positive judgments (Feingold, 1992; Stephen et al., 2012). In addition, naive people would acutely detect signs of sickness from the face in an early phase after exposure to infectious stimuli and potentially contagious people (Axelsson et al., 2017). As one would expect, perceived health decreases with age (Fink et al., 2011b).
As a company selling cosmetics and skincare products, we are particularly interested in the perception of a face in terms of age and health, and of course how to alter this perception by applying products.

In this context, the BWB has teamed up with the *Facial Analysis Synthesis and Tracking* (FAST) team from CentraleSupelec Rennes led by *Renaud Séguier*. The ambition of this thesis is to deepen the understanding of the aging process. Our ultimate objective is to develop a model to make an input face age in a plausible manner, and to gather knowledge from such a model. More specifically, the integration of wrinkles and studying its interaction with aging was identified as a promising first step. Moreover, we thought it could be interesting to evaluate an aging path. A face that ages while maintaining a healthy appearance could be a "positive" aging. Hence, we expected that perceived health estimation could represent a good indicator of a positive or negative aging path. An additional goal for the BWB was to know if a computer vision profile might help the team. Indeed, the BWB uses a lot of image processing techniques in their process, and was interested in being advised on this subject.

Constraints & Insights

Our objective is to develop a system able to make a face age and to gather knowledge on aging from such a system. This objective gives rise to several specific constraints:

— Alterations applied to faces due to aging might be learned from a dataset, integrating the less prior knowledge as possible. Hence, shape, color and high-frequency details might be added or removed following statistic rules learned from data.

— Linked with the precedent item, the objective is to generate, not the most photo-realistic, but the most plausible aged face from an input face image. Thus, adding arbitrary texture to give the produced face a natural look as in some approaches is not something desirable.

— Ideally, the system would use human interpretable parameters to help understand the facial aging process.

— Understanding the facial aging process means that using low resolution images ($< 256 * 256$) as it's often used in face aging applications is not possible, as lot of details are lost at that resolution.

Contributions

In view of our objectives and constraints, we base our work on the Active Appearance Model (AAM) to simulate facial aging, to which we incorporate a specific channel to fully integrate wrinkles; computed aging trajectories will take into account shape, appearance and wrinkles, differing from other methods which use classic AAM and add a post-processing step to include wrinkles.

Concerning the face aging simulation, we propose 3 contributions.

1. The first contribution is the parametrization of each wrinkle where shape and texture are represented altogether by a very understandable 7-length vector. Conversely, such a vector can be used to produce a wrinkle in shape and texture just from parameters.
2. To represent a group of wrinkles in one facial zone, we propose an approximation of an arbitrary joint probability of n random variables, as the set of every joint probability for every random variable taken two at a time; that is our second contribution. A new method of sampling for our approximated density is introduced.
3. The parametrization of skin micro-texture and its integration is our third contribution.

In addition, we worked on health estimation from faces and its link with aging. The way that humans perceive health from faces has been a growing area of psychology research the last decades. Researchers gather health ratings from humans before pointing differences in faces that could explain differences of ratings. We aimed to automatize this task by training a Convolutional Neural Network on a related task (age estimation) combined with a Ridge Regression to rate faces. The first system able to estimate health scores from faces is introduced. We propose 3 other contributions.

4. As often in psychology research, the database we have at our disposal is quite scarce. We experiment to show whether a Convolutional Neural Network trained on a similar task with larger datasets available like biological age estimation combined with a simple estimator allows us to achieve great performance on our task.
5. We propose a new methodology where we show that our system is not only able to imitate judgments by humans, but more importantly that it uses the same main facial cues as humans.
6. In our last contribution, we highlight new links between facial features and heath perception.

Thesis Organization

In the next chapter, we describe existing age progression systems and classify them according to the assumptions used in their paradigm and highlight their limitations according to our context.
In the chapter 3, a novel age progression method is presented. Active Appearance Model is taken as a basis to model faces to which wrinkles are explicitly integrated. Results are presented and compared to results from existing solutions. We showed that the system can generate realistic faces for aging and rejuvenating, and such age-progressed faces better influence age perception than other methods.
In the chapter 4, a system able to imitate human judgments of health from faces is provided. A Convolutional Neural Network trained on a related task (biological age estimation) is used in conjunction with a Ridge Regression trained on health estimation. We evaluate the performance of our system and compare it with human performance on the same dataset. More importantly, we highlight the fact that our system is influenced by the same facial cues as humans to judge health from a face. Finally, we show how such a system can be employed to study the influence of several facial cues on health perception.

Chapter 1

Related Works

1.1 Introduction

Age progression has been an ever-growing field for several decades. It has been applied to the search for missing children (Scherbaum et al., 2007; Simonite, 2006), entertainment (Sydell, 2009), cosmetics (Aarabi, 2013; Boissieux et al., 2000) and dermatology research (Aarabi, 2013; Restylane, 2012). In this kind of applications, facial aging simulation must consider age-related morphological changes as well as skin appearance modifications in order to provide realistic results. The most dramatic change of the face with age is morphological and results from facial growth; it occurs from birth to early adulthood (Farkas et al., 2004). Another age-related morphological modification concerns the facial volumes due to fat distribution variations; they vary all along life, from birth to late adulthood (Donofrio, 2000). During adulthood facial skin also undergoes dramatic changes with age, including wrinkling and sagging, increases of pigmented irregularities (Yaar and Gilchrest, 2007). All these skin age-related features are keys in the perception of facial age in adults (Burt and Perrett, 1995; Fink et al., 2006; Mark et al., 1980; Samson et al., 2010).

Given the diversity of potential applications of facial aging and the growing variety of computer vision techniques, many methods have been developed in recent decades (Fu et al., 2010; Ramanathan et al., 2009). The figure 1.1 can give us an intuition about the state of the art of age progression.
One of the pitfalls of developing age progression applications lies in the results evaluation part. Indeed, it is necessary to evaluate the quality of the predictions, i.e. to evaluate if the faces aged by the algorithm are close to the real faces aged by time, which is not always possible, depending of the dataset used. Various other evaluation methods can be used: assessing the photorealism of produced faces, quantifying the preservation of identity, and assessing the difference between the expected age and the perceived age (e.g. if a face has been artificially aged 10 years, will this face be perceived 10 years older by humans?). The different evaluation methods are accounted for each paper in the Table 1.1. In this table, evaluation methods were abbreviated as follows:

CRAF Close to "real" aged faces.
PHOTO Photorealism of produced faces.
IP Preservation of identity.
DEPA Difference between expected age and perceived age.

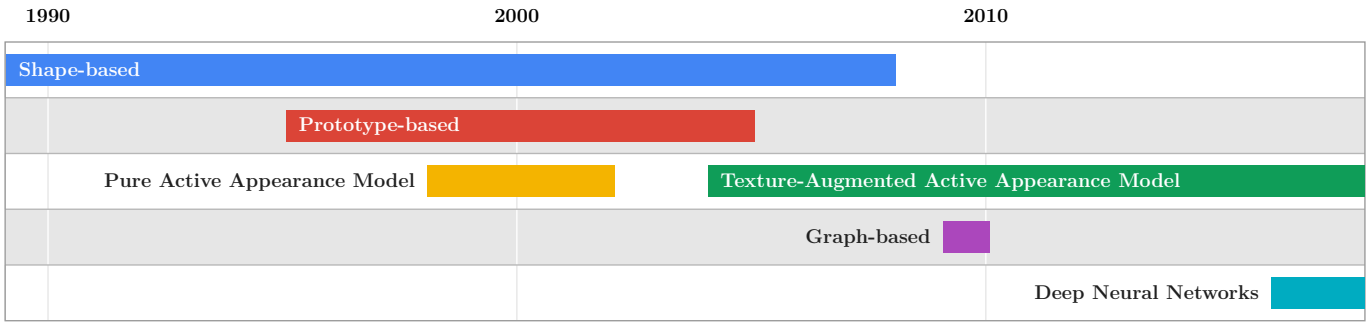

Figure 1.1 – A chronological overview of different families of methods applied to age progression.

Table 1.1 – State-of-the-art on age progression.

Authors	Main Methods	Dataset	Evaluation Methods			
			CRAF	PHOTO	IP	DEPA
Ramanathan and Chellappa (2006, 2008)	Revised cardioidal strain	FG-NET, Passport	X		X	
Lanitis et al. (1999)	AAM	Custom	X			X
Lanitis et al. (2002)	AAM	Custom	X		X	X
Geng et al. (2007)	AAM + Subspace	FG-NET, MORPH	X		X	
Tsai et al. (2014)	AAM + Patches Improv	FG-NET	X			X
Bukar et al. (2017)	AAM + Patches Improv	HQFaces, Dartmouth faces	X			
Burt and Perrett (1995)	Prototype	Custom				X
Tiddeman et al. (2001)	Wavelet Prototype	Custom				X
Tiddeman et al. (2005)	MRF Wavelet Prototype	Custom		X		X
Jinli Suo et al. (2009, 2010)	AAM + Graph	Custom, MORPH			X	X
Antipov et al. (2017)	acGAN	IMDb + Wiki			X	
Wang et al. (2016, 2018)	PCA + RNN	LFW, MORPH, CACD, Custom	X		X	X
Zhang et al. (2017)	CAAE	MORPH, CACD Custom	X			

1.2 Modeling based on shape

In the book "On Growth and Form", Thompson (1942) argues that biologists rely too much on evolution to explain the determination of shapes and structures in living organisms. Rather than evolution, Thompson reminds the importance of the physics and mechanics laws in shape determination. He analyses the physical constraints applied to biological forms as an explanation of their shape. Based on Thomson's work, Shaw et al. (1974) propose to model the growth of facial shape as mathematical operations. They identify two types of variations : the cardioidal strain and the affine shear, two types of affine transformations applied to the

contour of profile faces. Todd et al. (1980) revisit this approach and introduce the "revised" cardioidal strain transformation. This approach models human head growth as a fluid-filled spherical object with pressure. Each facial feature in polar coordinates (R_i, θ_i) is transformed with an internal pressure k using:

$$\begin{aligned} \hat{R}_i &= R_i(1 + k(1 - \cos(\theta_i))) \\ \hat{\theta}_i &= \theta_i \end{aligned} \tag{1.1}$$

Leveraging the "revised" cardioidal strain transformation model, Ramanathan and Chellappa (2006, 2008) propose an age progression application to analyze shape variations due to age for children under 18 years of age. Shapes are defined by a set of facial landmarks, and a model of facial deformation for aging during childhood is introduced. Then, faces are warped according to the deformation model to rejuvenate or age. This model permits them to estimate an age based on a face and to mock-up the face aging process for children. This model only takes on board shape variations because that is considered the principal source of variations from birth to adolescence. We can see on Figure 1.3 the transformation model (Fig. 1.2) being applied to two children's faces to make them look older.

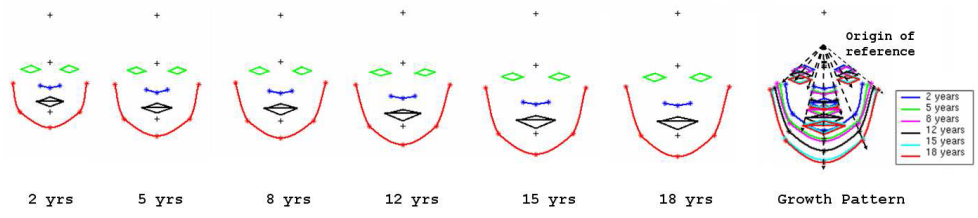

Figure 1.2 – The identified growth pattern from childhood to adolescence.

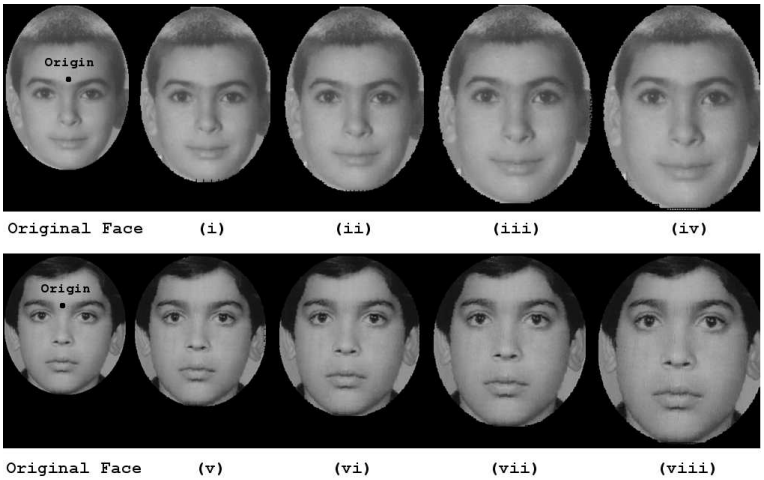

Figure 1.3 – The identified growth pattern is applied to two children's faces.

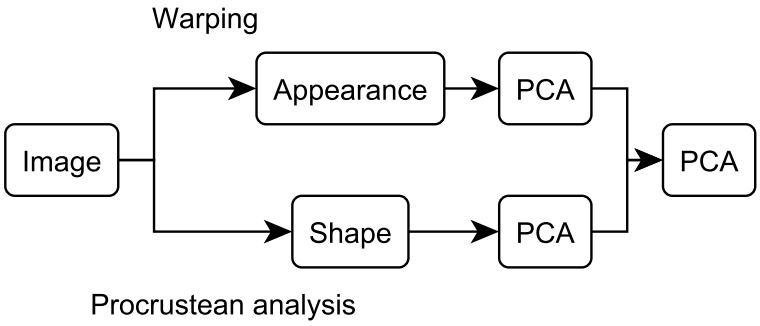

Figure 1.4 – Active Appearance Model Scheme

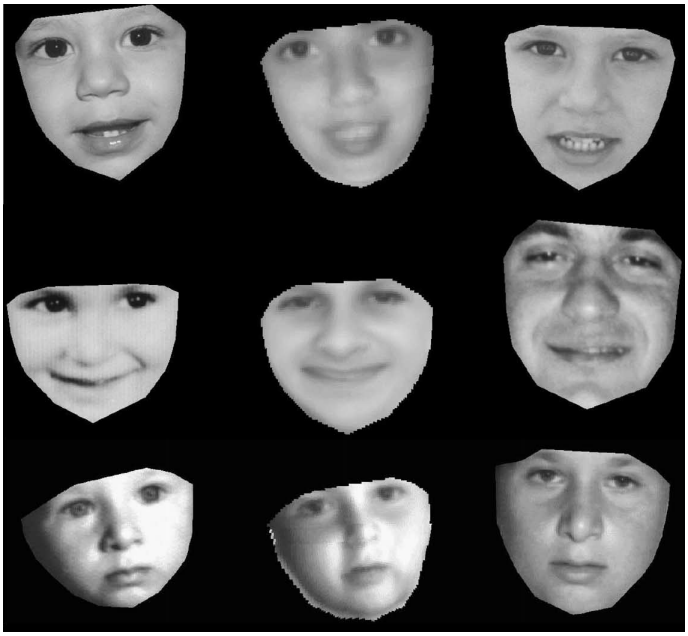

Figure 1.5 – Results from Lanitis et al. (2002). First and second column are original and age-progressed images, respectively. The third column shows the same subject at the target age.

1.3 Modeling based on appearance and shape

When elaborating a model for facial aging during adulthood, in addition to shape, appearance changes also need to be considered. The work of Lanitis et al. (1999, 2002) is the first to use Active Appearance Model on age progression (Fig. 1.4). We will now briefly present the functioning of Active Appearance Model (more explanation can be found in Sec 2.2 p. 46).
Active Appearance Model (Cootes et al. (1998)) is a statistical model which creates a subspace modeling appearance and shape variations from an annotated dataset of faces.
Firstly, landmarks are placed on facial features as eyes, mouth, nose and facial outline, and the coordinates of these points define the shape. Afterwards, shapes are *rigidly* aligned in translation, scaling and rotation with a procrustean analysis

using the mean shape as a reference.
Appearance information is then computed by warping every image into the mean shape, using each individual annotation. At last, only the pixels inside the convex hull of the facial outline defined by annotation are kept, and we have now the appearance feature.
After that, Principal Component Analysis (PCA) is carried out separately for shape and appearance, and a final PCA is made on the concatenation of shape weights and appearance weights. This creates a subspace which models variations of shape and appearance present in the dataset (Fig. 1.4). Regression of coordinates from this newly created space on age indicates the direction of facial aging. Finally, they can project a new face in this subspace, translate it in the face aging direction and reconstruct a shape and appearance to obtain an aged face. For this regression, they assess different strategies in their article:

Global Aging Function They test linear, quadratic and cubic functions to model the relation between all PCA weights and ages.

Person-Specific Aging Function As they use the FG-NET aging database which contains for each subject several photos at different ages, they compute a quadratic aging function for each subject. Assuming that similar faces age in a similar way, when a new face has to be aged, the aging function is a mixture of aging functions of closest faces in parameters space.

Lifestyle In addition to the previous method, they concatenate the raw PCA parameters with a vector quantifying lifestyle like sun exposition, health, economic situation or stress level, thus, the aging function is not only selected based on appearance but also on a lifestyle profile.

Qualitative results of this method can be seen in Fig. 1.5.

1.3.1 Aging Pattern

Geng et al. (2007) introduce the subspace named AGES (AGing pattErn Subspace) where each point represents the aging pattern of each subject. Taking advantage of having images of the same subject taken at different times in the FG-NET aging database (Fig. 1.6), they encode the whole aging pattern from a subject into a vector: they compute AAM parameters as Lanitis et al. (2002) for each image from the same subject before concatenating them in a sparse vector as in Fig. 1.7. If face images are available for certain ages, they are filled into the positions of the sparse vector corresponding to these ages. If not, the positions are zero-filled. After that, they wanted to create a subspace modeling these aging patterns, however, they couldn't directly use PCA as aging pattern vectors are largely incomplete/sparse. Thus, they propose to learn this subspace: with x_i an aging pattern, x_i^a and x_i^m respectively the available and missing components, they find the transformation matrix W which minimizes:

$$\frac{1}{N}\sum_{i=1}^{N}(x_i^a - \hat{x}_i^a)^T(x_i^a - \hat{x}_i^a) \text{ with } \hat{x}^a = \bar{x}^a + W^a(W^{aT}(x^a - \bar{x}^a)) \tag{1.2}$$

using an Expectation-Maximization algorithm combined with PCA. In simpler words, they find a W to project aging pattern vectors while minimizing the mean reconstruction error of available features.

Figure 1.6 – Two subjects taken at different times in the FGNET database. As we can see, the amount of photos for each subject varies from one individual to another.

Thus, with an incomplete aging pattern x_{test}, they can generate the fulfilled pattern $\hat{x}_{test}$ by projecting x_{test}, and back-projecting it as:

$$\hat{x}_{test} = \bar{x} + W(W^T(x_{test} - \bar{x})) \tag{1.3}$$

We can see examples of fulfilled patterns in Figure 1.8.

1.3.2 Texture Enhanced AAM

Nevertheless, AAM-based age progression is known to produce a blurry texture because wrinkles and spots are never perfectly aligned between people.
Facing this problem, more recent approaches (Bukar et al. (2017); Gandhi (2004); Tsai et al. (2014)) use AAM to produce appearance and shape, and add a post-processing step on appearance to superimpose patches of high-frequency details. Bukar et al. (2017) propose to use AAM with an addition, they linearly convert pixel intensities from the RGB colorspace to i1i2i3 colorspace (Ohta et al., 1980) to decorrelate each color channel from each other such as:

$$\begin{aligned} i1 &= \tfrac{1}{3}(R+G+B) \\ i2 &= \tfrac{1}{2}(R-B) \\ i3 &= \tfrac{1}{4}(2G-R-B) \end{aligned} \tag{1.4}$$

After that, they carry out PCA on each of the 3 new channels. Like the classic AAM, the concatenated vectors representing shape and appearance – here, 1 for shape, and 3 for appearance – before carrying out a final PCA creating a subspace modeling both shape and appearance variations.
The final PCA weights are regressed with sPLS, sparse Partial Least Squares regression. Given a $n * m$ matrix $C = \{c_1, c_2, \ldots, c_n\}^T$ of facial features, and a $n * 1$ vector a of ages, Partial Least Squares will find a decomposition describing both the covariate space of C and a. Adding the L_1 penalty, thus making the PLS sparse, improves variable selection performance, because the optimization process will try to explain as much information as possible with the lowest number of variables.

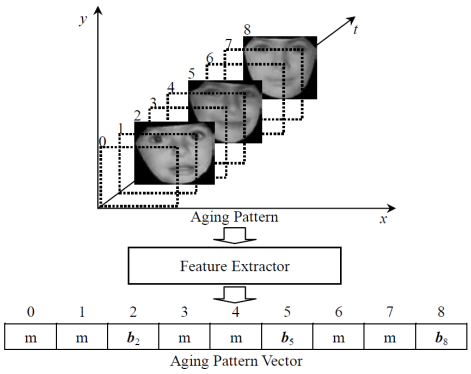

Figure 1.7 – Vectorization of the aging pattern. Ages (0-8) are marked at the top-left to the corresponding positions and above the corresponding feature vectors. The missing parts in the aging pattern vector are marked by 'm'.

Figure 1.8 – Results from the AGES method of Geng et al. (2007). Original images with the corresponding ages are on the first row. Generated images and 'real' aged images with the corresponding ages are on the second and third row, respectively.

At this step, they can simulate aging on faces, but the texture produced will suffer from the same burden defined earlier, i.e a blurry texture. Therefore, after the blurry face has been generated from AAM, they add an extra step to enhance texture. The generated face is divided into 72 overlapping patches and each patch will be replaced by a patch emanating from a real face from the database. With F_{init} and F_{aged} , the input face reconstructed and aged by AAM, respectively, the patch selection is done following 3 rules.

1. Stating that similar faces may have similar high-frequency details, authors propose to replace the patch P_k^{aged} from F_{aged} by the most similar patch P_k from the database.
2. When dealing with faces recomposed in patches, we have to be careful to choose matching adjacent patches. The overlapping region of the patch P_k has to be similar with the overlapping region of the patch above P_i^{aged} and to its left P_j^{aged}, both on the face F_{aged}.
3. Finally, the chosen patch P_k has to match the corresponding patch P_k^{init} on the original face F_{init}.

To compare patches between them, authors choose to use the Image Euclidean Distance from Li and Lu (2009). Indeed, simple euclidean distance is very sensitive to small translations, small rotations, noise or illumination. With two patches P_a and P_b with fixed size $M * N$, P_a^k the kth pixel of P_a, the euclidean distance can be written as follows:

$$ED(P_a, P_b) = \sum_{k=1}^{MN} (P_a^k - P_b^k)^2 \tag{1.5}$$

Authors propose to use Image Euclidean Distance defined as:

$$IMED(P_a, P_b) = \sum_{i=1}^{MN} \sum_{j=1}^{MN} g_{ij}(P_a^i - P_b^i)(P_a^j - P_b^j) \tag{1.6}$$

With g_{ij} a Gaussian distance function between two pixels position:

$$g_{ij} = \frac{1}{2\pi\sigma^2} exp\left(-\frac{{d_{ij}^S}^2}{2\sigma^2}\right) \tag{1.7}$$

d_{ij}^S being the euclidean distance between the spatial positions of pixels i and j, and σ, the standard deviation of the Gaussian function.
In simpler words, this distance is an euclidean distance where the correlation of pixels intensities in a neighborhood is integrated. The width of this neighborhood is defined by the parameter σ.

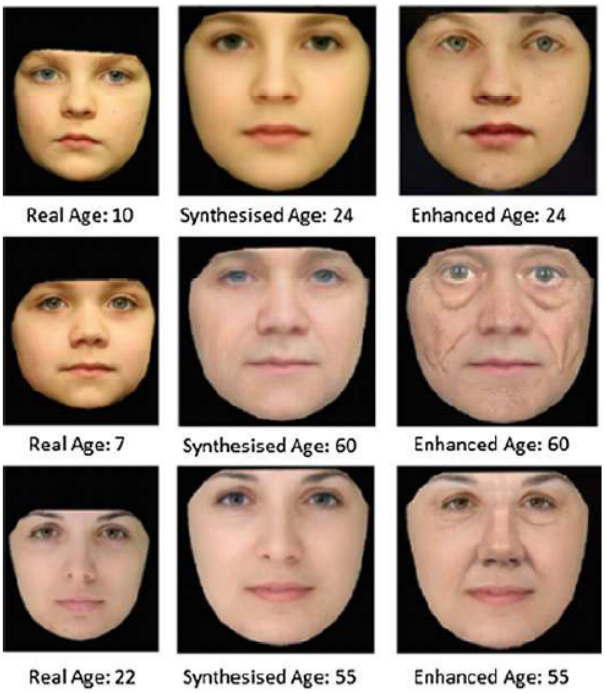

Figure 1.9 – Results from Bukar et al. (2017). Each row represents a different identity. Images on the left column are the original images prior to age synthesis. Middle images are AAM-synthesized and those on the right are enhanced outputs with the patch enhancement method.

We can see a subsample of their results in Figure 1.9.

Very similarly to the just described work from Bukar et al. (2017), Tsai et al. (2014) employ a patch enhancement method to counterbalance the blurriness of PCA generated appearance. More specifically, they base their work on the AGES method defined earlier, at which they add patch-based texture synthesis inspired

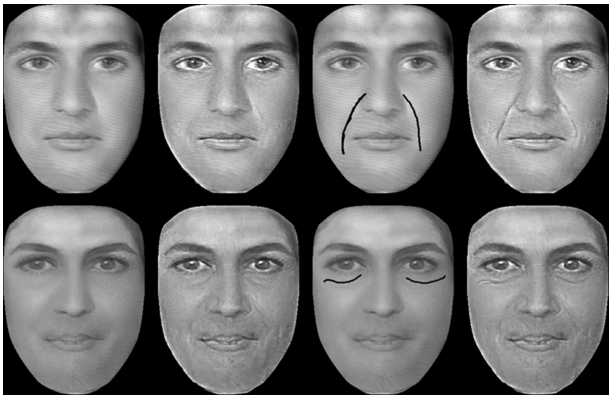

Figure 1.10 – Two examples of interactive aging enhancement from Tsai et al. (2014). The 1st column shows the original reference or predicted faces; the 2nd column shows the detailed face by patch-based transfer; the 3rd column shows indication curves assigned by users; the 4th column shows the results by interactive enhancement.

by Mohammed et al. (2009). The distance between two patches is defined as the sum between the euclidean distance between the two patches, and the euclidean distance between the gradient of the two patches:

$$D(P_a, P_b) = \sum_i \| P_a^i - P_b^i \|^2 + \| \nabla P_a^i - \nabla P_b^i \|^2 \tag{1.8}$$

In addition to this metric, several constraints are added to improve the synthesized face. To keep smooth stitching, the authors require that the distance D at boundary of adjacent patches should also be under a user-defined threshold. To keep the symmetry of faces, they restrict the two symmetric patches in the left and right parts of a face to be selected from the same image. However, even with this kind of constraints, authors argues that faces composed with selected patches cannot be smooth, and thus boundaries between patches appear in synthesized faces. That's why authors propose to use the famous Poisson image editing algorithm from Pérez et al. (2003) to remove color differences between patches while keeping gradient.
An interesting part of their method is the addition of an optional interactive step to synthesize wrinkles. As synthesized faces don't have gradient information, the selected patches from the database won't have a lot of gradient either. Authors propose to draw wrinkles directly on the blurry synthesized face, thus adding high frequency details, and only after to enhance the face. As a consequence, the patch selection step will take into account the drawn lines. We can assess the results of interactive synthesizing and aging in Figure 1.10 and 1.11, respectively.

1.3.3 Conclusion

However, about the two described method, while faces produced are plausible and better than AAM in terms of photo-realism, details added are not statistically learned for age progression, as texture patches that contain details are chosen with a similarity measure, and not with respect to a precise age. Thus, the

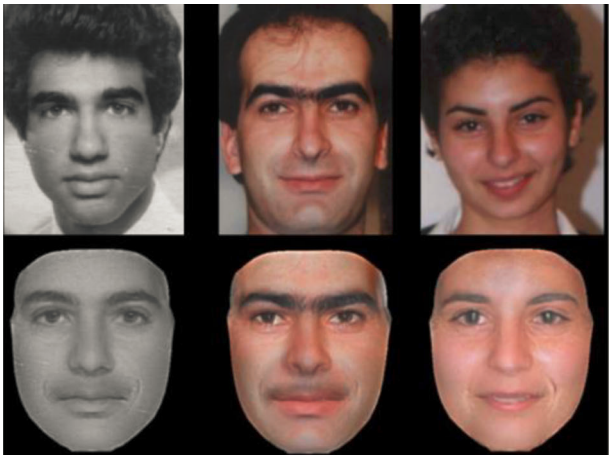

Figure 1.11 – Results from Tsai et al. (2014)'s work. The 1st rows: input images from the FG-NET. The 2nd rows: aged images of the 1st rows by their method.

key idea can be defined in simple terms: the high-frequency details of close faces in the database are transferred onto the generated blurry face. There is room for improvement.

1.4 Modeling based on prototype

Another approach creates a prototype (Burt and Perrett (1995); Rowland and Perrett (1995)), an average face from faces within a constrained age group, meant to represent typical features from this group. A younger face can be then warped in the mean shape, and the prototype blended on the texture of the younger face to make it look older. As for AAM-based methods, prototype-based methods suffer from the same problem; making an average face will blur out every non-aligned high frequency detail. Tiddeman et al. (2001, 2005) propose to add a post-processing step to enhance high frequency information on the average face. Tiddeman et al. (2001) extract fine details with wavelet decomposition for every face to add them on the final average face, with a parameter σ controlling the level of details to transfer. They choose to decompose faces using an over-complete wavelet basis. Images can be decomposed using a critical-basis of wavelets, decomposition has the same size as the original image, and reconstruction is guaranteed to be strictly equal to the original image. However, authors note that with a critical-basis decomposition, alteration of frequency bands can lead to artifacts after reconstruction. Thus, they choose to use an over-complete basis of 2D Gabor wavelets. A 2D Gabor function g can be formulated as a Gaussian function modulated by a cosine function:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \psi\right) \tag{1.9}$$

with λ the wavelength of the cosine function, θ the orientation, ψ the phase, σ the standard deviation of the Gaussian envelope, and γ the spatial aspect ratio.

As shown in Figure 1.12, authors choose to build a wavelet pyramid in two directions at three scales.

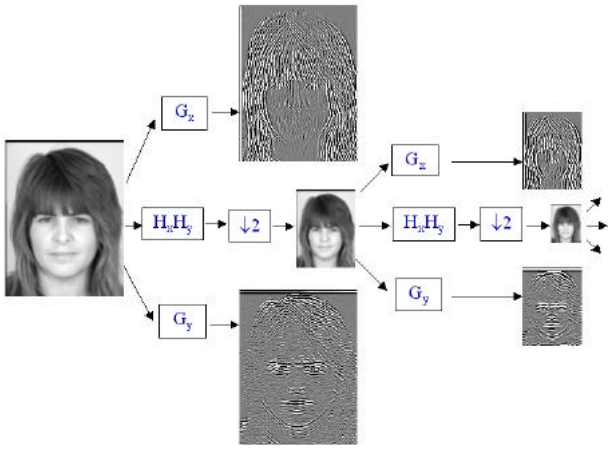

Figure 1.12 – A typical image (left) and the corresponding wavelet decomposition pyramid (right). The coefficients of the filters H and G are (1, 4, 6, 4, 1)/16 and (1, -4, 6, -4, 1)/16 respectively.

After that, they measure an edge strength σ_w by smoothing with two 1D cubic B-spline B_x, B_y, the magnitude in a frequency band w:

$$\sigma_w = B_x * B_y * |w| \tag{1.10}$$

Hence, the mean values of σ in a specific frequency band across the set of N facial images gives a measure of the average edge size in each region. Finally, we can enhance the texture in each frequency band $\bar{w}$ of the average face by multiplying it by the ratio between the average smoothed magnitude $\bar{\sigma}$ and the smoothed magnitude of the average face $\sigma_{\bar{w}}$. We can see an example in Figure 1.13 where this method is used to transfer a texture from a face to another.

$$\bar{w}_{new} = \bar{w}\frac{\bar{\sigma}}{\sigma_{\bar{w}}} \tag{1.11}$$

In Tiddeman et al. (2005), they combine wavelet decomposition with Markov Random Field to regenerate fine details on the average face, which produces more realistic results. Given an input face Z belonging to a young person, several young persons faces A_i, and several older persons faces B_i. The average face B is superimposed on the face Z, to produce an aged face $\hat{Z}$. However, the newly aged face won't have any wrinkle as an average face is often blurry. Hence, high frequency details of the face $\hat{Z}$ are synthesized pixel by pixel on the different scales and orientations, starting from the lowest scale to the highest. With a pixel $p_{\hat{Z}(x,y)}$ to synthesize and its known neighborhood $N(p_{\hat{Z}}(x,y))$, the goal is to maximize the following probability:

$$\arg\max_{p_{\hat{Z}(x,y)}} P(p_{\hat{Z}(x,y)}|N(p_{\hat{Z}}(x,y)), N(p_{A_{1..N}}(x,y)), N(p_{B_{1..N}}(x,y)), p_{A_{1..N}}(x,y), p_{B_{1..N}}(x,y))) \tag{1.12}$$

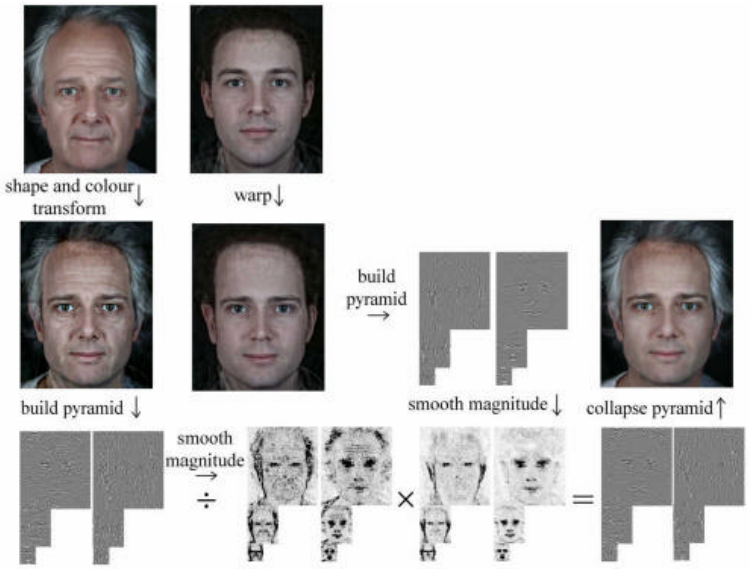

Figure 1.13 – One example of the texture enhanced transformation process from Tiddeman et al. (2001) where a texture from an target face (top right) is transferred to the original image (top left). The original image is transformed using the shape and color method and the target prototype is warped into the new shape. Wavelet pyramids are then built from these two images and their magnitudes are calculated. After rescaling, the subject's pyramid is collapsed to give the new image.

As specified by the authors, the neighborhood in $\hat{Z}$ only contains pixels already synthesized, at the same scale, and at other scales. The neighborhood is composed of the 12 closest pixels already synthesized, plus the 9 closest pixels at the precedent scale, and one pixel in each precedent scale remaining. To find the value of pixel $p_{\hat{Z}(x,y)}$, they empirically construct the conditional probability density with the faces used to make the average face, for the younger group A and for the older group B, using the Parzen-Rosenblatt window method. Given two 1D histograms H_A, H_B initialized with zeros, they can be computed for a precise pixel as:

$$\begin{aligned} H_A(p_{A_i}(x,y)) &+= \mathbb{G}(N(p_{A_i}(x,y)), N(p_{\hat{Z}}(x,y)), \Sigma) \\ H_B(p_{B_i}(x,y)) &+= \mathbb{G}(N(p_{B_i}(x,y)), N(p_{\hat{Z}}(x,y)), \Sigma) \end{aligned} \tag{1.13}$$

with $\mathbb{G}$ the multi-dimensional Gaussian function evaluated at $N(p_{A_i}(x,y))$, centered on $N(p_{\hat{Z}}(x,y))$, and with a uniform standard deviation matrix Σ. In simpler terms, the probability to synthesize the value p_{A_i} is equal to the sum of the Gaussian distance function between the neighborhoods of images $N(p_{A_i}(x,y))$ in the database having this exact pixel value p_{A_i}, and the neighborhood of the face to enhance $N(p_{\hat{Z}}(x,y))$. Thus, the choice of the pixel value to synthesize in image $\hat{Z}$ is influenced by the closer neighborhoods in the database from its own neighborhood. A 1D Gaussian smoothing is applied on the two histograms H_A, H_B to turn them into probability densities. In addition, the two probability density functions are converted in cumulative density functions F_A, F_B. At last, the value of pixel

$p_{\hat{Z}}(x,y)$ is found with the following equation:

$$p_{\hat{Z}}(x,y) = F_B^{-1}(F_A(p_{\hat{Z}}(x,y))) \tag{1.14}$$

In other words, the pixel value $p_{\hat{Z}}$ is transformed into a cumulative probability with F_A, before being remapped to a pixel value with F_B^{-1}. We note that, as F_B is a 1D cumulative density function, strictly increasing, we can trivially inverse it. The described process is done for every scale at every orientation, then, the whole pyramid can be collapsed to produce a real face. Results from this method are shown in Figure 1.14.

Figure 1.14 – Results from Tiddeman et al. (2005). The original East-Asian face images (left) are rejuvenated using European faces as examples (center). Clamping the output values to within 3 s.d. of the conditional mean (right) improves the stability of the synthesis.

1.4.1 Conclusion

Despite the dates of publication of these papers (2001-2005), we can notice interesting ideas.
First, this approach based on prototype alter both shape and appearance of faces with aging.
Secondly, the two methods to improve the high frequency details of faces are very different. The wavelet smoothing approach superimposes high frequency details shared by several faces in a specific age group. Thus, superimposed high frequency details are in a way representative of the target age group. On the other hand, the MRF method developed after uses the neighborhood of the original image and the neighborhood of the target age group to creates sharper and more realistic faces. However, details generated with this method are not very representative of the target age group, they can be linked to a specific identity.

1.5 Modeling based on graph

Jinli Suo et al. (2009, 2010) decompose faces in three levels: global face and hair as the first level, the different face parts (eyes, nose, mouth, ...) composed the second level, and the third level represents wrinkles. Their database is split into five age intervals, spaced over ten years. The face and facial parts are modeled by several AAMs. Thus, an AAM is used for the face, and others AAM are made for every local parts (eyes, mouth, ...). For the wrinkle level, they divided the face in 6 zones, and each wrinkle in each zone are annotated. Wrinkles are annotated and their properties (numbers, lengths, positions, orientations) are modeled by a Poisson distribution, for each property.

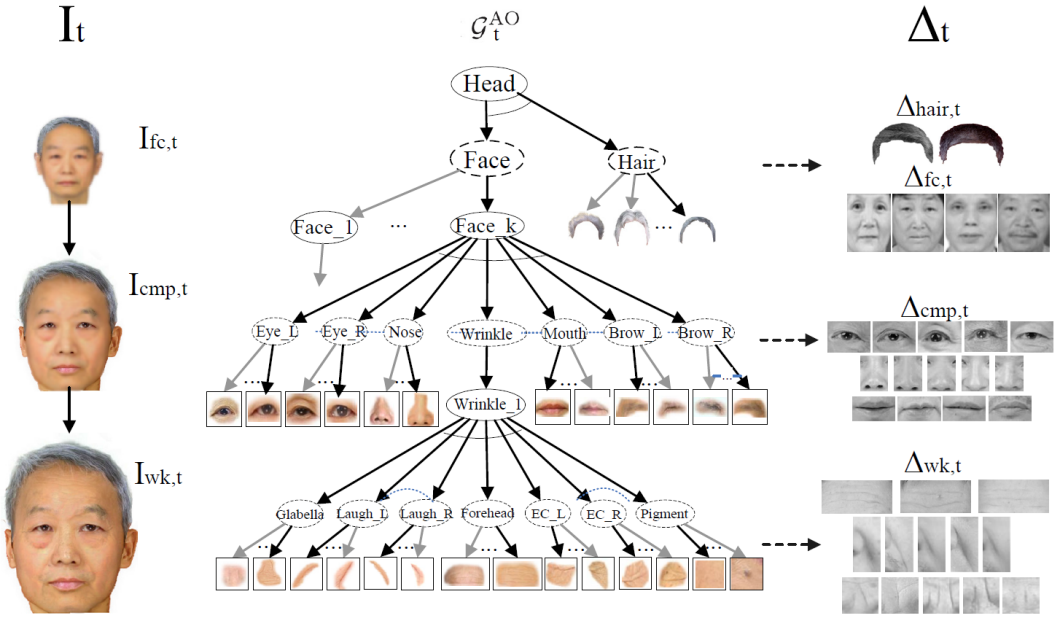

Figure 1.15 – Left: A high resolution face image I_t at age group t is represented at three resolutions – $I_{face,t}$, $I_{cmp,t}$ and $I_{wkl,t}$. Middle: All face images at age group t are represented collectively by a hierarchic And-Or graph $\mathcal{G}_t^{AO}$. The And nodes (in solid ellipses) in the graph $\mathcal{G}_t^{AO}$ represent coarse-to-fine decomposition of a face image into its parts and components. The Or-nodes (in dashed ellipses) represent alternative configurations. By choosing the Or-nodes, we obtain a parse graph G_t for a specific face instance. Right: Dictionary Δ_t includes $\Delta_{hair,t}$, $\Delta_{face,t}$, $\Delta_{cmp,t}$ and $\Delta_{wkl,t}$ at three levels from coarse to fine.

The three face levels are distributed into an And-Or Graph (Fig. 1.15). And nodes represent different parts of the face, whereas Or nodes represent the different realizations of these parts for the population in every age group. They use a first order Markov chain to model aging of parts of the face. Artificial aging can be created by decomposing a face, present in age group t, in a And-Or graph G_t, and to sample the probability $p(G_{t+1}|G_t)$ with Gibbs sampling algorithm; the graph G_{t+1} can be collapsed to generate a new face (Fig. 1.16). During synthesis, the low-frequency face is generated. Generated face parts are pasted on the gener-

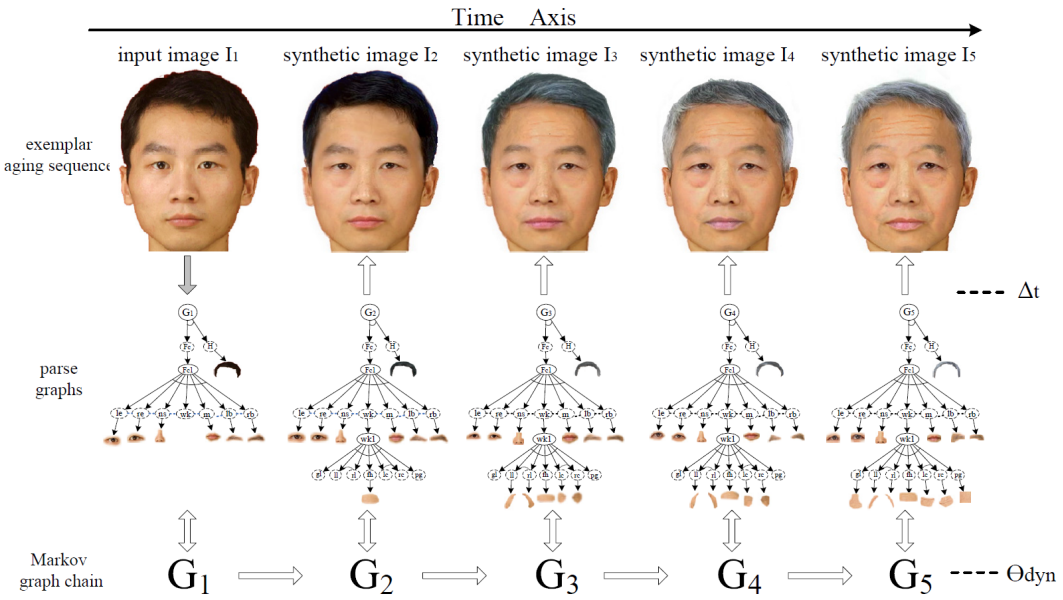

Figure 1.16 – Modeling the aging process as a Markov chain on parse graphs. Top row is a face image sequence at different ages, with the leftmost one being the input image and the other four being synthetic aged images. The second row is the parse graphs of the image sequence. The third row shows the Markov chain and θdyn includes the parameters for Markov chain dynamics.

ated face and blended with Poisson Image Editing from Pérez et al. (2003). About wrinkles, their number in a zone is randomly sampled from the prior distribution. After that, wrinkle shapes and appearances are picked randomly from wrinkles in the corresponding age group. Finally, the probability density for each wrinkle property (length, position, orientation) is randomly sampled and wrinkle shapes are modified to follow these properties. Thus, the wrinkle appearances can be warped in the wrinkle shapes, and blended on the synthesized face, again with the help of Poisson Image Editing. We can see some results in Figure 1.17.

1.5.1 Conclusion

We can note that's the first work to explicitly integrate hair and wrinkles in a face aging model. Hence, produced images are more realistic and more useful for a real-world application. However, aging of each face part is done independently of others, despite the known correlation between face parts (e.g eyebrows and hair color are correlated, eye lightness and face color as well). As a consequence, the combination of several facial parts can produce some unrealistic results, even if the realism of faces has not been addressed in these articles.

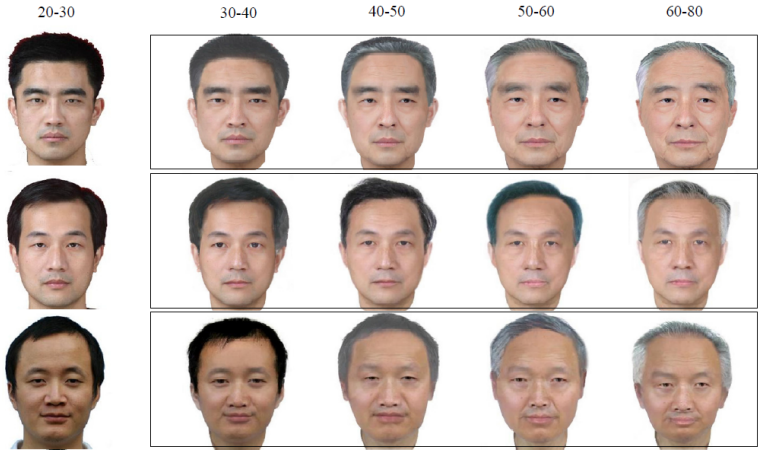

Figure 1.17 – Some aging simulation results from Jinli Suo et al. (2010)'s work. The leftmost column is the original images of the individuals in group 1. The 2nd to 5th columns are synthetic aged images at 4 consecutive age groups.

1.6 Modeling based on deep neural networks

Promising approaches (Antipov et al., 2017; Liu et al., 2017; Wang et al., 2016, 2018; Zhang et al., 2017) propose to use Deep Neural Networks to produce aged faces.

1.6.1 Generative Adversarial Network

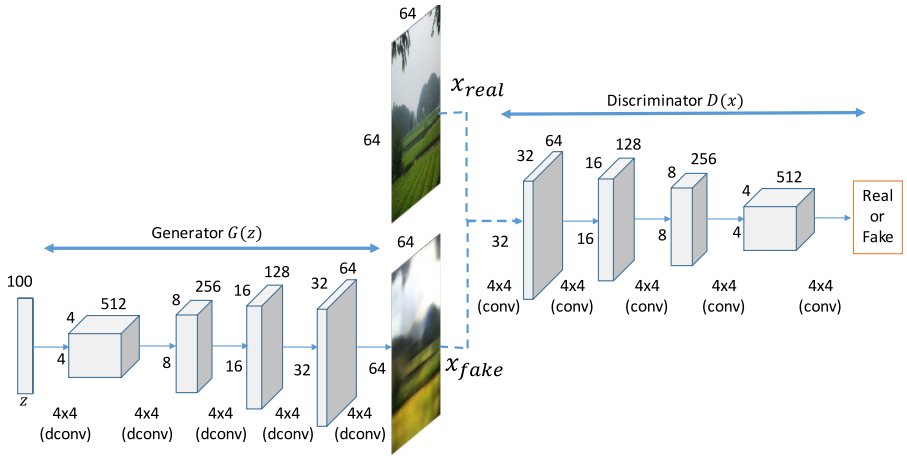

Figure 1.18 – Typical GAN architecture.

Goodfellow et al. (2014) introduced the Generative Adversarial Network (GAN) architecture (Fig. 1.18). This architecture is composed by a generator G and a discriminator D. The goal of G is to generate a realistic picture given a latent vector z. At the same time, the discriminator D has to discriminate between real pictures originating from the database, and images generated from G. The clever

trick of this architecture lies in the fact that the 2 networks are trained together and improve at the same time. The generator uses the classification from the discriminator as a loss to improve. The discriminator improves more and more to discriminate between real and generated images, while the generator gets better and better at fooling the discriminator. Thus, as the loss function for the generator G is not defined in the pixels space like PCA or Autoencoder, generated images are sharper than in previously methods like Section 1.3 and Section 1.4.

1.6.2 Age Conditional Generative Adversarial Network

To extend GAN to age progression, Antipov et al. (2017) propose age conditional Generative Adversarial Network (acGAN).
As we can see in Figure 1.19, a face passes through an encoder E to project the face in a latent vector z_0. The vector y_0 defines the person's age, one-hot encoded with six age categories: 0-18, 19-29, 30-39, 40-49, 50-59, 60+ years old. A reconstruction of the face $\bar{x_0}$ can be obtained by feeding the generator the two vectors: $\bar{x_0} = G(z_0, y_0)$. After that, to age a face, we can modify the vector y_0 with the desired target age and keep z_0 vector unmodified, and feed again the generator with the two vectors. We will explain with more detail, the training process of such a system.

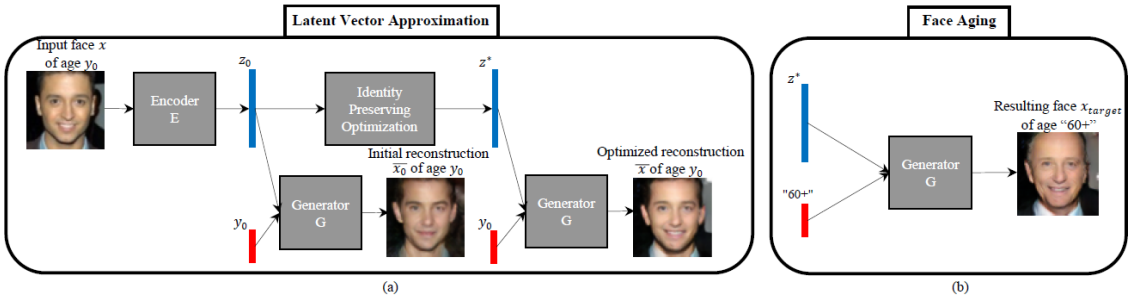

Figure 1.19 – acGAN scheme. (a) approximation of the latent vector to reconstruct the input image; (b) switching the age condition at the input of the generator G to perform face aging.

First, the conditional GAN (cGAN) is trained in a classic manner: cGAN training can be expressed as an optimization of the function $v(\theta_G, \theta_D)$, where θ_G and θ_D are parameters of G and D, respectively:

$$\min_{\theta_D} \max_{\theta_G} v(\theta_G; \theta_D) = \mathbb{E}_{x,y\sim p_{data}}[logD(x,y)] + \mathbb{E}_{z\sim p_z(z),\tilde{y}\sim p_y}[log(1 - D(G(z,\tilde{y}),\tilde{y}))] \tag{1.15}$$

Thus, the discriminator will learn to associate pictures with age classes, and the generator, thanks to the discriminator will be able to generate faces corresponding to a specific age class.
Secondly, they introduce an encoder E: they generated 100K pairs $(\bar{x_i}, G(z_i, y_i))$

where $z_i \sim N(0, I)$ are random latent vectors, $y_i \sim U$ are random age conditions uniformly distributed between six age categories, $G(z, y)$ is the generator of the priorly trained cGAN, and $\bar{x}_i = G(z_i, y_i)$ are the synthetic face images. Afterwards, E is trained to minimize the Euclidean distances between estimated latent vectors $E(x_i)$ and the ground truth latent vectors z_i. Thirdly, they introduce an identity-preserving latent vector optimization. Using a pretrained face recognition neural network embedding FR, and z_0 as an initial latent vector, they optimize z^*_{IP} such that:

$$z^*_{IP} = \arg\min_z \| FR(x) - FR(\bar{x}) \|_{L2} \tag{1.16}$$

In simpler terms, starting from z_0, they find a latent vector z^*_{IP} that retains identity better after reconstruction, than the initial reconstruction from z_0.

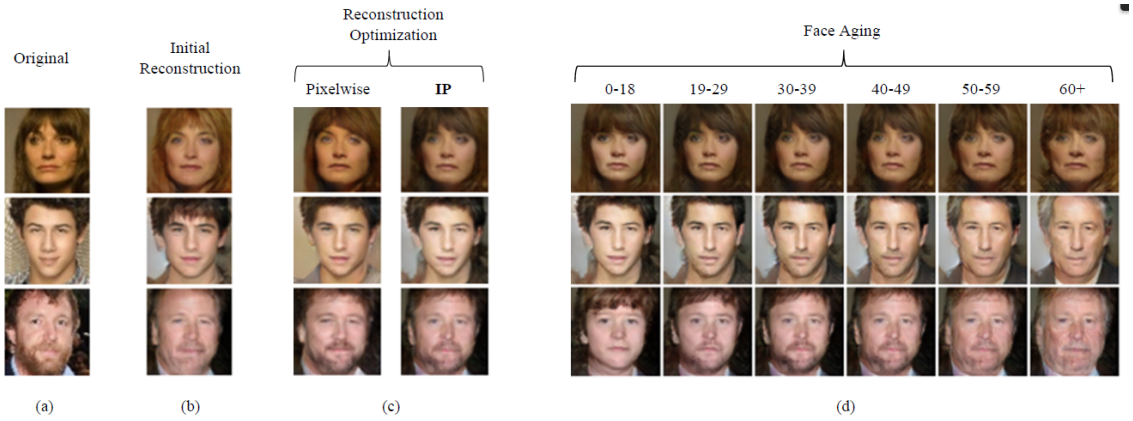

Figure 1.20 – Examples of face reconstruction and aging from acGAN. (a) original test images, (b) reconstructed images generated using the initial latent approximations: z_0, (c) reconstructed images generated using the "Pixelwise" and "Identity-Preserving" optimized latent approximations, and (d) aging of the reconstructed images generated using the identity-preserving latent approximations and conditioned on the respective age categories y (one per column).

Results show that faces can be aged/rejuvenated for long-term period, taking into account facial features like hair or beard (Fig. 1.20).

1.6.3 Recurrent Neural Network

Wang et al. (2016, 2018) introduce a Recurrent Face Aging (RFA) framework using a Recurrent Neural Network (RNN) which takes as input a single image and automatically outputs a series of aged faces. They divide the faces of each gender into 9 age groups, as shown in Figure 1.21. For training, their model requires the faces of the same person covering two adjacent groups. The RNN transforms a face across different age groups by decomposing the complex image generation process into a sequence of intermediate states with smaller and subtle changes. To be specific, RNN is applied to construct complex images iteratively where the rough

outlines are gradually replaced by precise forms, and lines are sharpened.Their method includes two main contributions: Face Normalization and RNN-based Face Aging Model.

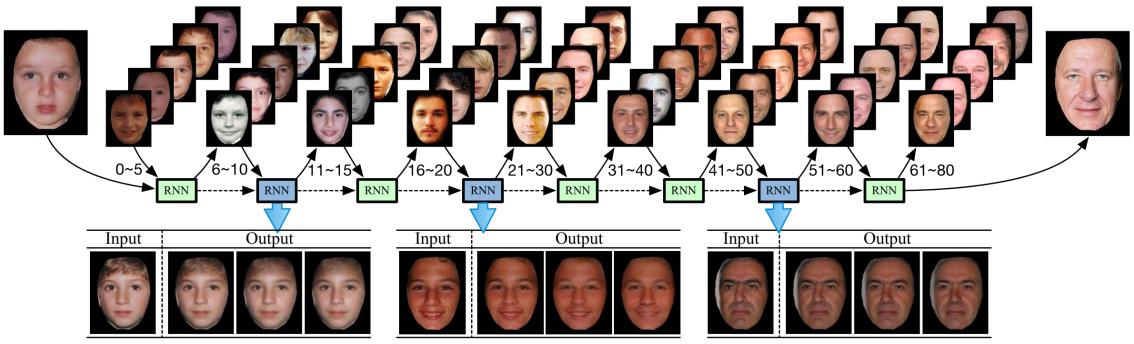

Figure 1.21 – The recurrent face aging (RFA) framework exploits a RNN to model the aging pattern. The aged face is synthesized by referring to the autoregressive memory of the previous faces. The intermediate transitional faces can also be synthesized.

Face normalization is an important preprocessing step for face aging. To normalize the faces, many works rely on facial landmarks and warping. To avoid distortions with landmarks-based warping, authors prefered to use optical flow. As shown in Figure 1.22, they mask images to keep only the face and, for each age group, they compute a PCA to define an eigenface space. Keeping the k first components, they project each image I in the low rank eigenface space and reconstruct it, to get I'. Afterwards, the optical flow can be computed from I' to I, and we can use it to obtain $\hat{I}$, the warped version of I. As $\hat{I}$ can have ghost artifacts due to complex optical flows, authors decided to make this process iterative. Starting with $k = 4$, and with n the number of faces in an age group, the images $[I_1, \ldots, I_n]$ are warped to get $[\hat{I}_1, \ldots, \hat{I}_n]$. After that, a new PCA is computed with the newly warped faces $[\hat{I}_1, \ldots, \hat{I}_n]$, and with k incremented by 1. The same process described above repeats until $k = 80$.
After that, for each couple of adjacent age groups, a RNN is trained to learn the transformation between the weights from the younger age group x_y and the weights from the older one x_o. With x'_o being the generated weights from the RNN, the loss is defined as a least square minimization of a difference between ground truth and predicted weights:

$$J = \| \ (x_o - x'_o) \ \|_2^2 \tag{1.17}$$

The authors used Whitening PCA - eigenvectors are divided by their corresponding eigenvalues, thus each eigenvector has a unit variance - to give each PCA dimensions the same importance.

To make a face age in a nutshell, an input face is aligned with optical flow and projected in the eigenface space corresponding to its age group. Its weights w_y

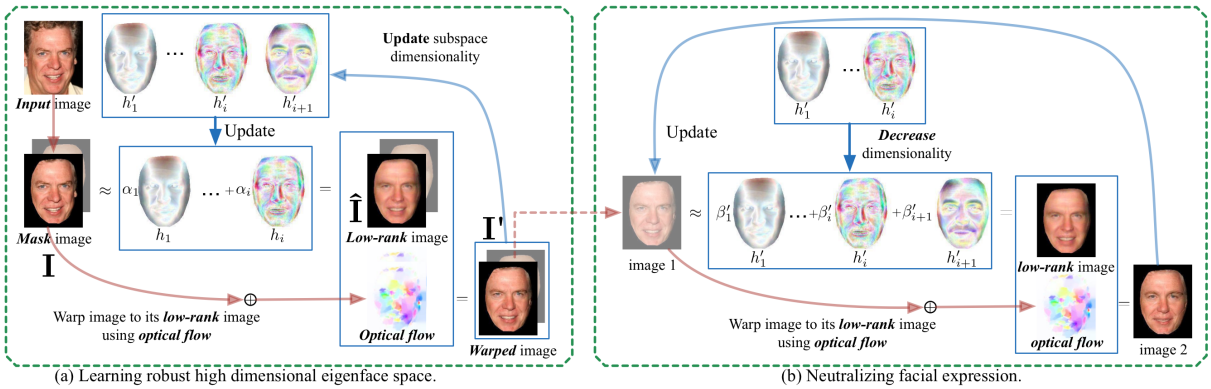

Figure 1.22 – In Wang et al' work, face normalization process consists of two steps. Step 1, shown in (a), is to learn a robust eigenface space incrementally which is insensitive to the errors brought by the optical flow. Step 2, shown in (b), is to neutralize the facial expressions progressively by decreasing the dimensionality of the eigenface space.

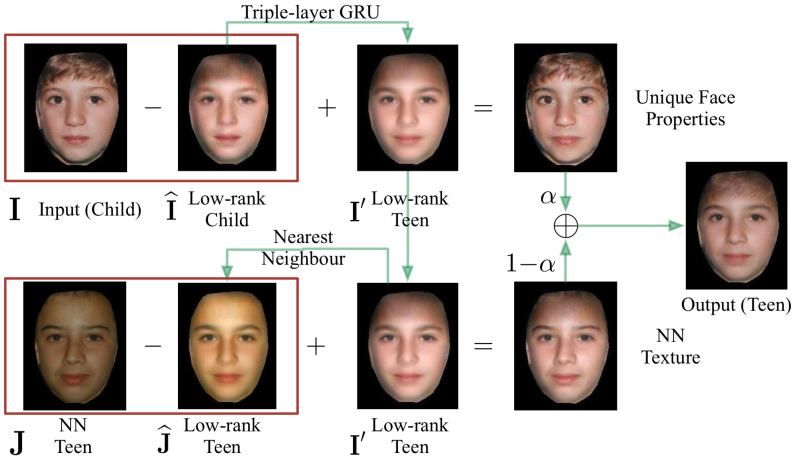

Figure 1.23 – Texture transfer with the nearest neighbor from Wang et al' work.

are fed to the RNN to predict the weights $\hat{w_o}$ corresponding to an older face. The nearest neighbours weights $w_{nearest\_\hat{w_o}}$ from the database are found. An older face can be synthesized by reconstructing a face from $\hat{w_o}$, and by adding to it the high frequency details from the face in the database corresponding to $w_{nearest\_\hat{w_o}}$ (Fig. 1.23).
We can see aging results in Figure 1.24.

1.6.4 Conditional Adversarial Autoencoder

Zhang et al. (2017) propose Conditional Adversarial Autoencoder (CAAE). They use an Autoencoder combined with 2 discriminators working on latent variables and output images to impose photo-realistic results (Fig. 1.25). The first discriminator D_z imposes latent variables z to be uniformly distributed to avoid "holes"

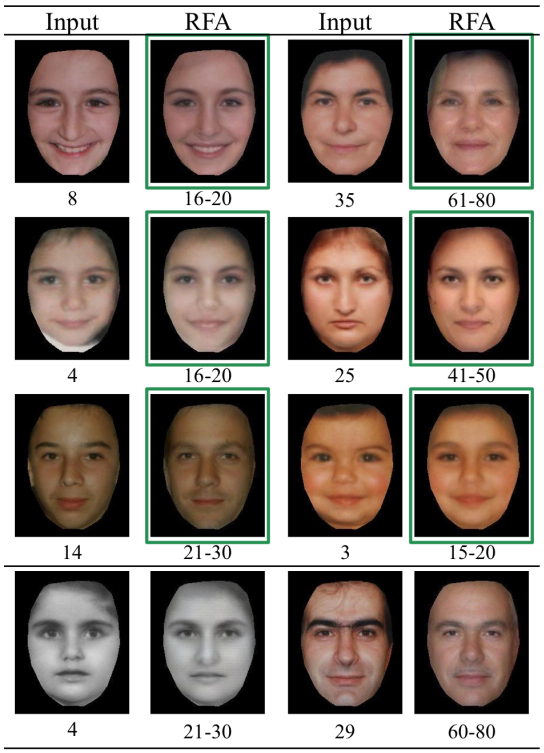

Figure 1.24 – Aging results from Wang et al' work.

in the latent space, and thus to produce a smooth age progression. The second discriminator D_{img}, inspired by the GAN architecture, discriminates between real images and generated images, and its loss is used to improve the photo-realism of pictures. Age progression is achieved by regressing the latent variables with respect to age.

First, authors propose to use a Conditional Autoencoder. An autoencoder is a neural network taking a face as input, and compressing it to a low dimension manifold of latent variables, that's the encoder E part. The second part of the neural network - the decoder, or in this case the generator G - takes the latent variables and reconstructs the input faces. The encoder and generator parts are trained as one neural network taking an input image and outputting the same image, with a L_2 loss. Age is divided into ten categories: 0–5, 6–10, 11–15, 16–20, 21–30, 31–40, 41–50, 51–60, 61–70, and 71–80. Each age is one-hot vector encoded, and concatenated to the latent variables during training, making the autoencoder conditional (Fig. 1.25). With l the one hot encoded age label, and x a face image, such an autoencoder can be trained with the following loss:

$$\min_{E,G} \parallel x - G(E(x), l) \parallel_{L2} \tag{1.18}$$

In addition to that, authors propose 3 types of regularization.
First, they introduce a multi layer perceptron D_z with the purpose of imposing a uniform distribution on latent variables. As encoder E learns a highly non-linear mapping of the input, authors argue that forcing a uniform distribution

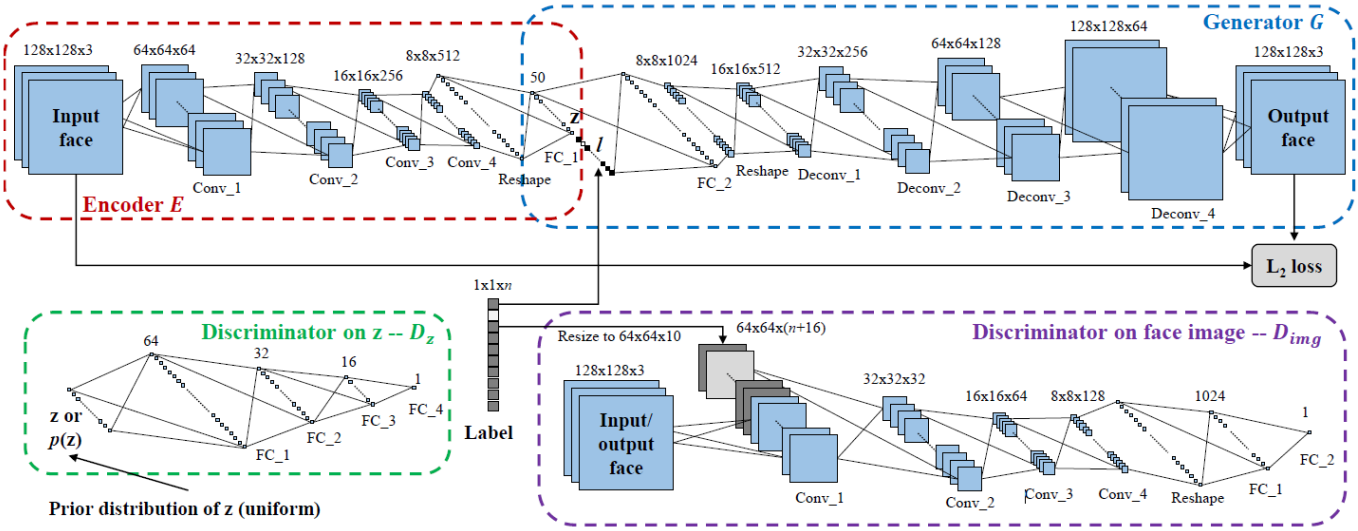

Figure 1.25 – Structure of the proposed CAAE network for age progression/regression. The encoder E maps the input face to a vector z (personality). Concatenating the label l (age) to z, the new latent vector $[z, l]$ is fed to the generator G. Both the encoder and the generator are updated based on the L_2 loss between the input and output faces. The discriminator D_z imposes the uniform distribution on z, and the discriminator D_{img} forces the output face to be photo-realistic and plausible for a given age label.

will reduce holes in this space. Thus, a displacement in this space will unlikely create faces far from faces from the database. Therefore, at each training step of encoder E, latent variables z and generated z^* following a uniform distribution $p(z)$ are fed to D_z to improve E and D_z:

$$\min_E \max_{D_z} \mathbb{E}_{z^* \sim p(z)}[logD_z(z^*)] + \mathbb{E}_{x \sim p_{data}(x)}[1 - log(D_z(E(x)))] \tag{1.19}$$

with p_{data} being the distribution of the training data, and $z^* \sim p(z)$ being the random sampling from the prior $p(z)$. Encoder E learns to output a uniform distribution while D_z learns to discriminate between samples following a uniform distribution and those that don't, that's an adversarial configuration.
Secondly, authors introduce a second adversarial regularization, the Convolutional Neural Network D_{img}. This ConvNet corresponds to the Discriminator in a typical GAN architecture (Fig. 1.18), as explained in Section 1.6.1 page 34. The discriminator D_{img} takes an input face, and a age label, making it Conditional actually, and outputs the probability that the picture is close to faces in the database. Similarly, the Generator G tries to fool the Discriminator D_{img} by generating better faces, and the Discriminator tries to discriminate between real and generated images. The loss to minimize is:

$$\min_G \max_{D_{img}} \mathbb{E}_{x,l \sim p_{data}(x,l)}[logD_{img}(x,l)] + \mathbb{E}_{x,l \sim p_{data}(x,l)}[1 - log(D_{img}(E(x),l))] \tag{1.20}$$

Thirdly, to avoid ghosting artifacts in faces generated by GANs, authors argue to minimize the total variance of each batch images. Such a regularization forces the generated faces to be less diverse, more average.
Finally, with $TV(\cdot)$ being the total variance calculation, the function to optimize is:

$$\min_{E,G}\min_{D_z,D_{img}} \lambda \parallel x - G(E(x),l) \parallel_{L2} + \gamma TV(G(E(x),l)) \\ \mathbb{E}_{z^*\sim p(z)}[logD_z(z^*)] + \mathbb{E}_{x\sim p_{data}(x)}[1 - log(D_z(E(x)))] + \\ \mathbb{E}_{x,l\sim p_{data}(x,l)}[logD_{img}(x,l)] + \mathbb{E}_{x,l\sim p_{data}(x,l)}[1 - log(D_{img}(E(x),l))] \tag{1.21}$$

The coefficients λ and γ balance the smoothness and high resolution.

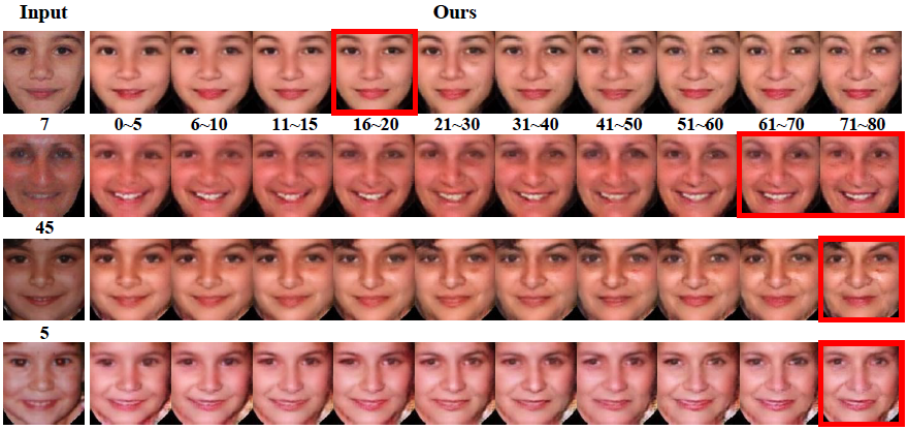

Figure 1.26 – Results from CAAE. The first column shows input faces, and the rest columns are their results from both age progression and regression.

We can see the results of their methods in Figure 1.26.

1.6.5 Conclusion

These Deep Learning methods, by their specificities, present interesting advantages compared to other techniques. First of all, a precise annotation of facial features is not needed for acGAN and CAAE. A face detector is used to crop faces, and the cropped results are fed to the network without attenuating variations of pose and expression. Above all, these systems can handle a learning with millions of samples as they are trained with the Stochastic Gradient Descent families of methods. In addition, as their methods use the adversarial loss introduced by Goodfellow et al. (2014), where high frequency details are taken into account by the discriminator, theoretically, fine facial details can be generated by the network.
However, age progression algorithms based on neural networks can produce in some cases unrealistic faces (e.g the 2 eyes of a reconstructed face can have different shapes). In addition, lots of these algorithms work on low resolution faces, at most 128x128 (Antipov et al., 2017; Liu et al., 2017; Zhang et al., 2017). Thus, as

the used faces are too small to show fine details, these face aging systems cannot generate faces with fine wrinkles.
Concerning the work from Wang et al. (2016, 2018) introducing the Recurrent Face Aging (RFA) framework, Recurrent Neural Networks are only used to make the transition from an eigenface from one age group to another. Hence, as faces are represented by PCA, this system suffers from the same burden as PCA-based methods. This work does not take advantage of the representation power of Convolutional Neural Networks as the previously mentioned methods.

1.7 Discussion

As the reader will have noticed, the simulation of facial aging involves a wide variety of techniques, mixing image processing, statistics, optimization and deep neural networks. The first methods introduced "carefully" chosen shape deformations of faces directly inspired by anthropometry. Approaches based on prototype and Active Appearance Model both proposed two enhancements. First, to embed the appearance along with the shape. Secondly, the use of statistics - PCA and regressions for AAM, average in an age group for prototype-based methods - allows us not to have to integrate much prior knowledge to make a face age, which is desirable to understand the face aging process. Active Appearance Model offers the advantage of continuous modeling. Most of the methods developed later focused on the integration of texture details in the produced results, taking as a basis Active Appearance Model, prototype, or Deep Neural Networks.

In the light of our objectives, we can draw up a number of remarks on the state of the art to better understand which research direction to follow to achieve it.
First of all, we can observe that a large majority of these methods models the face aging process as a discrete process. Indeed, in all methods, except Active Appearance Model, authors split their dataset in several age groups of approximatively 10 years, and their aging system model the transition from one age group to another. More specifically, Texture-Enhanced AAM methods model facial aging as continuous thanks to AAM, but split their dataset in age groups too when it comes to enhance produced faces. In view of the fact that face aging is a continuous process, this splitting, despite being convenient for algorithmic purpose, does not reflect the reality of face aging. Furthermore, it doesn't allow short-term aging and smooth transitions between faces at different ages.
Secondly, we can notice some weaknesses concerning approaches trying to include texture details. The key idea in most approaches is to assign to the aged face, the texture of the closest face in the database. Recent works decompose faces in patches, and the same assignation is made patch after patch. This process is made on the assumption that faces with close low frequency information will have the same high frequency details, which is quite reasonable but imperfect. It would be desirable to dispose of a system able to model the texture for the shape and the appearance, in statistically learning its distribution in faces of the

database.
Thirdly, as one of our objectives is to understand the process of facial aging, the way of superimposing facial details on faces is not optimal. Ideally, we would have a representation of facial details which would be human understandable, and in the same time, able to regenerate facial details on a face from this representation.

Chapter 2

Face Aging

2.1 Introduction

As defined in the introduction, our objective is to develop a system able to make a face age and to gather knowledge on aging from such a system. This objective gives rise to several specific constraints:

— Alterations applied to faces due to aging might be learned from a dataset, integrating the less prior knowledge as possible. Hence, shape, color and high-frequency details might be added or removed following statistic rules learned from data.

— Linked with the precedent item, the objective is to generate, not the most photo-realistic, but the most plausible aged face from an input face image. Thus, adding arbitrary texture to give the produced face a natural look as in some approaches is not something desirable.

— Ideally, the system would use human interpretable parameters to help understand the facial aging process.

— As shown in the previous chapter, low resolution images ($< 256 * 256$) are often used in face aging applications. Understanding the facial aging process means that using low resolution images is not possible, because lot of details are lost at that resolution.

In view of the current state of art and our constraints, we based our work on the Active Appearance Model to simulate facial aging. Active Appearance Model has the advantage to learn the modifications due to aging without integrating much prior knowledge. We will explain in detail the functioning of Active Appearance Model in Section 2.2. After showing some lacks of the model and its usage in facial aging simulation (Sec. 2.2.4.2), we will introduce our main contribution, the Wrinkle Oriented Active Appearance Model in Section 2.3. Results from our new model are shown in Section 2.4 and discussed in Section 2.5. The whole scheme of our system is shown in Figure 2.1.

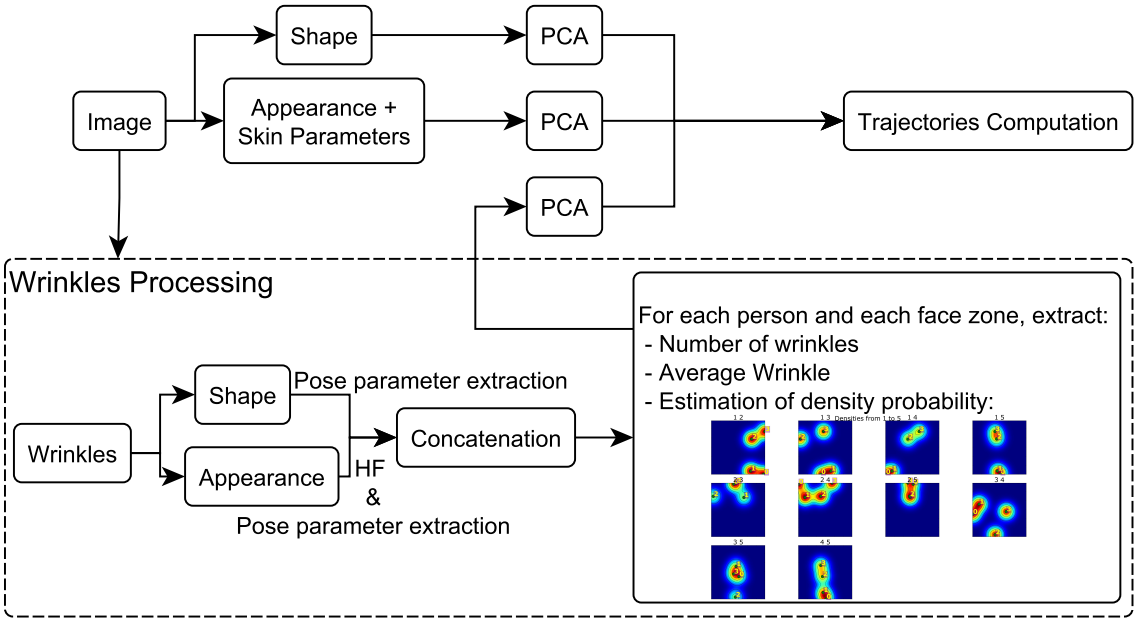

Figure 2.1 – Wrinkle Oriented AAM Scheme.

To analyze faces in the light of facial aging, we propose 4 contributions.

1. The first contribution is the parametrization of each wrinkle where shape and texture are represented altogether by a very understandable 7-length vector. Conversely, such a vector can be used to produce a wrinkle in shape and texture just from parameters (Sec. 2.3.1).
2. To represent a group of wrinkles in one facial zone, we propose an approximation of an arbitrary joint probability of n random variables, as the set of every joint probability for every random variable taken two at a time; that is our second contribution (Sec. 2.3.2). To sample from such a set, a new method of sampling is proposed (Sec. 2.3.4).
3. The parametrization of skin micro-texture and its integration is our forth contribution (Sec. 2.3.5.3).

2.2 Active Appearance Model

In this section, we will briefly present the functioning of the AAM and its usage in face aging simulation. More importantly, the limitations of AAM and propositions to circumvent them will be pointed out at the end.

Active Appearance Model (Cootes et al., 1998) is a statistical model which creates a subspace modelling appearance and shape variations in an annotated dataset of faces (Fig. 2.2).

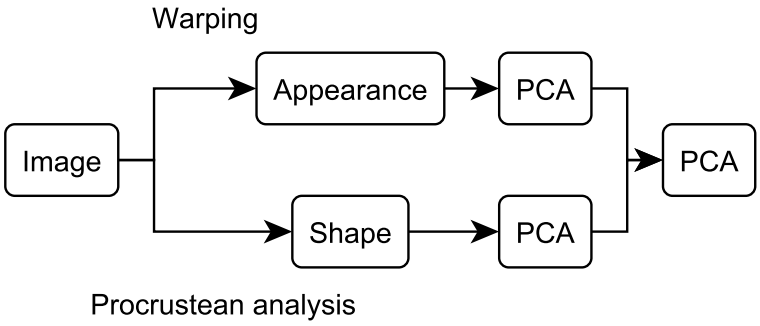

Figure 2.2 – AAM Scheme

2.2.1 Shape Normalization

Firstly, landmarks are placed on facial features as eyes, the mouth, the nose and the facial outline, and the coordinates of these points define the shape. Afterwards, shapes are *rigidly* aligned in translation, scaling and rotation with a procrustean analysis using the mean shape as a reference. For this analysis, each shape S_i is aligned to the mean shape $\bar{S}$ as a least square minimization of Eq. 2.1.

$$\arg\min_{A,\theta,b} \| \bar{S} - \left(A \times S_i \cdot \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} + b \right) \|_2^2 \tag{2.1}$$

with A, θ, and b respectively the scaling, rotation and translation parameters.

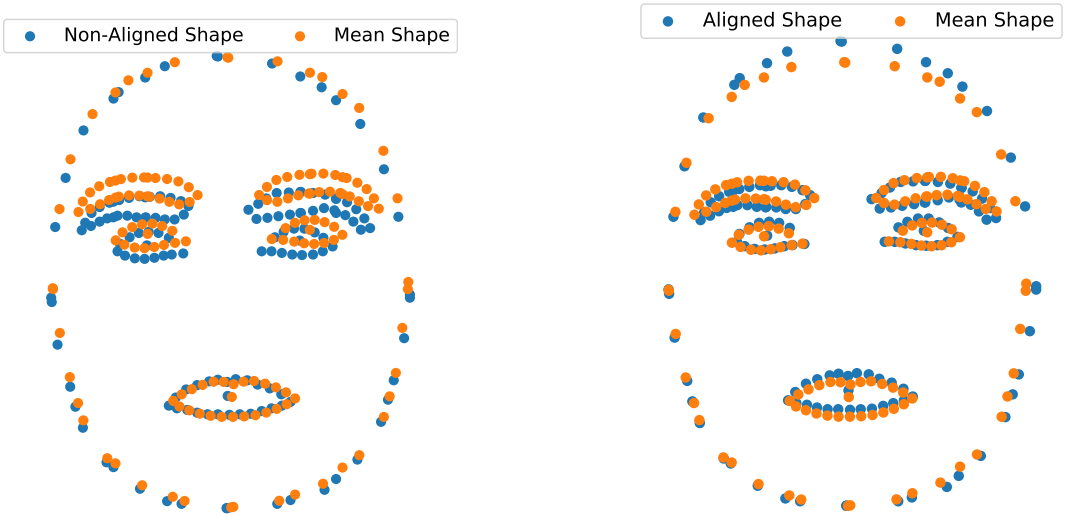

Figure 2.3 – Left: the shape of a face from the base (blue) and the mean shape (orange). Right: the two shapes are aligned using procrustean analysis.

2.2.2 Appearance Normalization

Secondly, face images are *warped* in the mean shape. The face is paved with triangles using a Delaunay Triangulation maximizing the minimum angle of all triangles. The same triangulation is made for the mean shape. After that, the inverse warping algorithm starts: for each triangle T_i' in the mean shape and each corresponding triangle T_i in the face to be warped, an affine function f_i^{-1} is determined so that $f_i^{-1}(T_i') = T_i$. Finally, to find the value of the pixel $I'_{x',y'}$ in the warped image I', we identify the triangle this pixel belongs to, we express it as a barycenter of the 3 vertices (A', B', C') of this triangle as

$$(x', y') = \left(w_1 A' + w_2 B' + w_3 C'\right) \text{ with } \sum_{j=1}^{3} w_j = 1 \tag{2.2}$$

with w_1, w_2, and w_3 the 3 new relative coordinates of the pixel $I'_{x',y'}$.
Afterwards, we compute the corresponding position in the original image I as

$$(x, y) := \left(w_1 f_i^{-1}(A') + w_2 f_i^{-1}(B') + w_3 f_i^{-1}(C')\right) \tag{2.3}$$

Now, with the bicubic interpolation, we can assign $I'_{x',y'} := B(I, (x, y))$ with $B(I, (x, y))$ a cubic combination of the 16 closest pixels to (x, y) in I. The process is repeated for each pixel in I^w (Fig. 2.4). At last, we take only the pixels inside the convex hull of the facial outline defined by annotation, and we have now the appearance feature (Fig. 2.5).

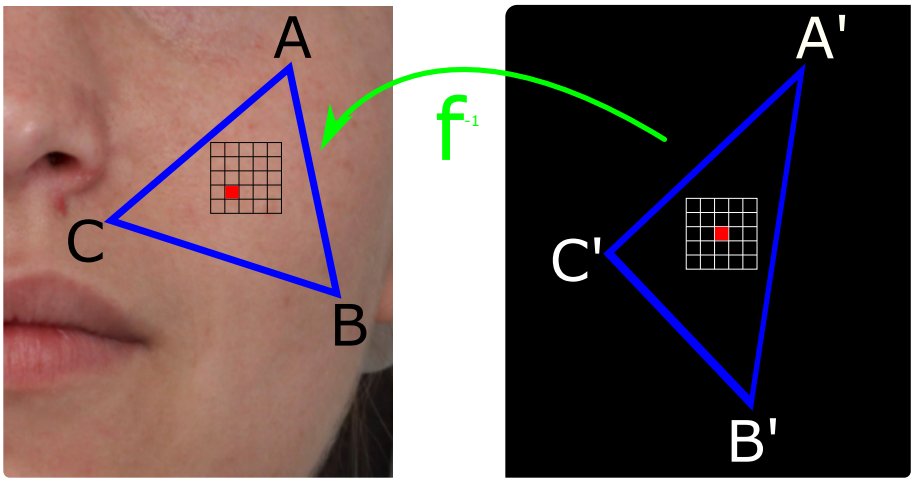

Figure 2.4 – Illustration of inverse warping.

2.2.3 Principal Components Analysis

Two Principal Component Analysis (PCA) are carried out, one for appearance, and one for shape. Principal Component Analysis transforms observations with correlated variables into values with new linearly uncorrelated variables. These

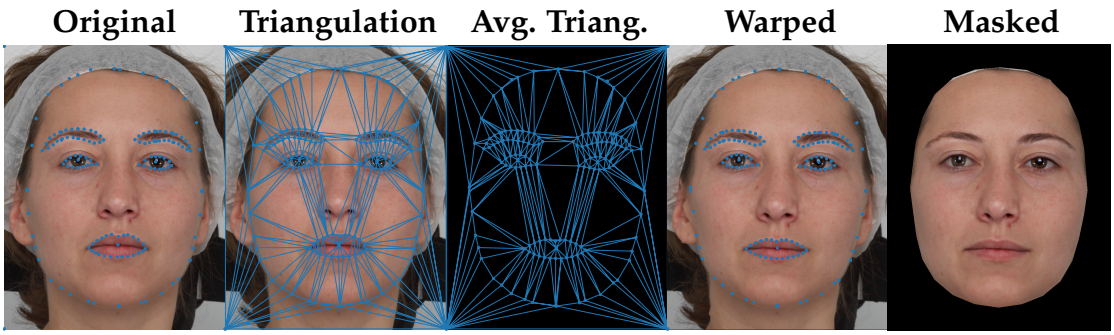

Figure 2.5 – The process to compute the appearance feature from an annotated face.

new variables are called the principal components. The n observations with d variables are projected into n vectors of dimensionality $min(d, n-1)$. This transformation is made such as a principal component accounts for as much variance as possible while being linearly uncorrelated from other components. In addition, each principal component explains more variance from the original observations than the following components. As a consequence, it often takes a few first components to explain much of the variance in the original set of observations.

2.2.3.1 PCA Calculation

Mathematically speaking, PCA on a set of observations X can be computed using a Singular Value Decomposition (SVD) of the covariance matrix of X. We will detail the calculation of a classical Principal Component Analysis.

Mean Subtraction First, the mean vector $\bar{X}$ is subtracted from X of dimensions (n, p):

$$\hat{X} = X - \bar{X} \tag{2.4}$$

Covariance Matrix Second, $\hat{X}$ having a mean of zero, $\hat{X}^T \otimes \hat{X}$ defines the sum of the product of the differences from the mean for every pair. Dividing by the number of products $n-1$ gives us an average, thus a covariance matrix C of dimensions (p, p):

$$C = \frac{\hat{X}^T \otimes \hat{X}}{n-1} \tag{2.5}$$

Eigenvalue Calculation Third, eigenvalues λ of the covariance matrix C have to be computed. For that, we can use a property of eigenvalues:

$$det(C - \lambda I) = 0 \tag{2.6}$$

with det the determinant operator, I the identity matrix of dimension (p, p), and λ a scalar.

Eigenvector Calculation Forth, now that eigenvalues are found, eigenvectors V have to be computed. For that, we can use a property of eigenvectors and eigenvalues:

$$(C - \lambda I)V = 0 \tag{2.7}$$

PCA Projection Finally, the projection of X in the PCA space is made by:

$$W = (X - \bar{X})V^T \tag{2.8}$$

and the perfect reconstruction of X from W can be achieved by:

$$X = WV + \bar{X} \tag{2.9}$$

2.2.3.2 Usage in the Active Appearance Model

PCA is used in Active Appearance Model to encode shape and appearance. Each shape are vectorized - e.g. 250 landmarks represented by a matrix of dimensions $(250, 2)$ are "flattened" to obtain a vector of dimensions $(500,)$. For appearance, as defined earlier, only the pixels inside the convex hull of the facial outline are kept. Once again, the matrices of dimensions $(P, 3)$ are flattened to vectors of dimensions $(P * 3,)$. A PCA is made for each features independently - one for shape and one for appearance. After that, shape weights W_{sha} are normalized to have the same global standard deviation as appearance weights W_{app}. In order to do so, shape weights are divided by the sum of eigenvalues of shapes, and multiplied by the sum of eigenvalues of appearances.

$$W_{sha} := \frac{W_{sha} \sum_{i=1}^{m} \lambda_{app,i}}{\sum_{i=1}^{n} \lambda_{sha,i}} \tag{2.10}$$

Afterwards, the shape and appearance weights are concatenated to produce a matrix $W_{app\&shape}$. Finally, a third PCA is carried out on $W_{app\&shape}$ to generate a subspace modeling variations in faces both in shape and appearance. The normalization made in Equation 2.10 prevents the final PCA to favor appearance weights, as these weights have larger variance in comparison to shape weights. The final weights W_{all} produced by the third PCA will be used by the next step to make a face age in terms of shape and appearance.

At this state, two remarks can be made.
First, as defined earlier PCA is perfectly reconstructible. Hence, a face used to construct the PCA can be projected in the final PCA subspace, and back-projected to recreate the original shape and appearance. The appearance can be warped to the reconstructed shape to get the original face.
Second, PCA is generative. A point can be synthetically created in the final subspace - at a random position for example - and back-projected to generate a new face.

2.2.4 Usage of Active Appearance Model in Face Aging Simulation

2.2.4.1 Trajectories & Simulation

To each image in the database corresponds a point in the final subspace. Our goal in this section is to identify a face aging direction. For each face, we have the age of the corresponding person. Hence, an arbitrary regression f can be made between faces in PCA space W_{all}, and ages. For example, a cubic polynomial can be chosen for f:

$$f(W_{all}) = A^T W_{all}^3 + B^T W_{all}^2 + C^T W_{all} + D \tag{2.11}$$

Therefore, function f is able to estimate an age given a weight from a face. To make a face with a perceived age a look older/younger of k years, we have to project it on the final subspace to obtain weights W_i , apply the Equation 2.12, and reconstruct an aged face from modified W_i.

$$W_i := W_i + (f^{-1}(a+k) - f^{-1}(a)) \tag{2.12}$$

In Equation 2.12, the function f^{-1} outputs PCA weights given an age. As multiple different faces can match the same age, $f(a)$ will return the average PCA weights $W_{mean,a}$ of this specific age.
To inverse f, a Monte Carlo simulation is made as in Lanitis et al. (2002). A lot of plausible weights W are generated; the corresponding age for each weight $w \in W$ is found by applying $f(w)$, and f^{-1} is a lookup table where for a given age a, $f^{-1}(a)$ is an average of all weights $W_a \subset W$ such as $f(W_a) = a$.

To sum up, the final PCA models variations of faces in shape and texture. In this PCA space, a regression on age is made to identify an aging trajectory. A face projected in this subspace is translated in the direction identified as an aging direction thanks to the regression. The translated point is reconstructed to a face, keeping the same identity, but looking older/younger.
We can note that multiple trajectories can be made for different aging trajectories. For example, after AAM calculation, the population could be divided into 2 groups, one with people who have been very sun-exposed, and another group with people who have not. Thus, two regressions could be made, one for each group. Finally, a new face could be projected in the subspace and translated using the most suitable trajectory to achieve the most precise aging simulation.

2.2.4.2 Problems encountered

When applying the previous method to make a face age, we obtain the result in Figure 2.6. Aging in terms of shape and color variations is quite well modeled by AAM. However, new wrinkles don't appear and existing wrinkles aren't amplified. That's the same for freckles.

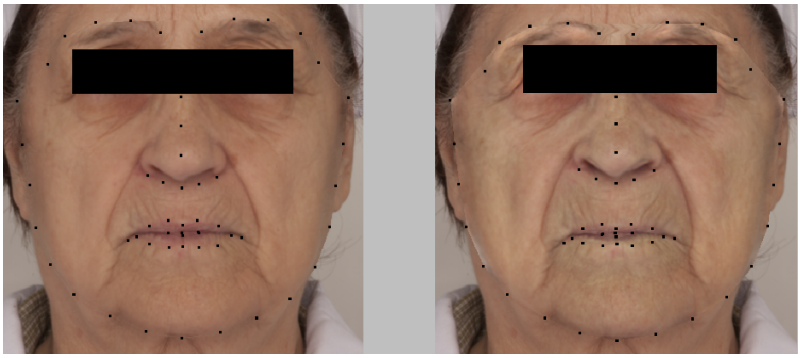

Figure 2.6 – Left: Original image. Right: After an aging of 15 years.

Explanation PCA computes an average appearance and creates its dimensions using deviations of appearances from this average. As wrinkles/freckles aren't perfectly aligned between people, PCA puts in different dimensions similar signs of facial aging.
This problem of misalignment can be seen from another angle in Figure 2.7. An average face made of 15 faces of aged people with a large amount of wrinkles. This average face, supposed to represent the group of 15 faces, is less wrinkled than the least wrinkled face of the group, and is perceived younger.

Proposition A representation of wrinkles less sensible to slight misalignments has to be found. With this representation, AAM would produce rejuvenated faces without wrinkles and aged faces with wrinkles. Similarly, an average face with this representation would present a level of wrinkles representative of the group used to make the average; wrinkles on the average face would have average parameters: an average number of wrinkles, average intensities, average lengths, ...

Figure 2.7 – An average face made of 15 old women faces. Faces are warped in the mean shape and an average is computed for each pixel.

2.3 Wrinkle Oriented Active Appearance Model

In view of the current state of art and our constraints, we base our work on the Active Appearance Model to simulate facial aging, to which we incorporate a specific channel to fully integrate wrinkles (Sec 2.3.1 & 2.3.2) in this subspace, computed aging trajectories will take into account shape, appearance and wrinkles, differing from other methods which use classic AAM and add a post-processing step to include wrinkles.
Afterwards, we detail how to synthesize aged faces from our new wrinkle oriented AAM (Sec. 2.3.4).
Finally, we propose to study the quality of our aging system by presenting images resulting from the aging and rejuvenating of faces (Sec. 2.4.2). Then, we show that this approach increases/decreases perceived age more precisely than the unmodified Active Appearance Model with an age estimation convolutional neural network (Sec. 2.4.3).

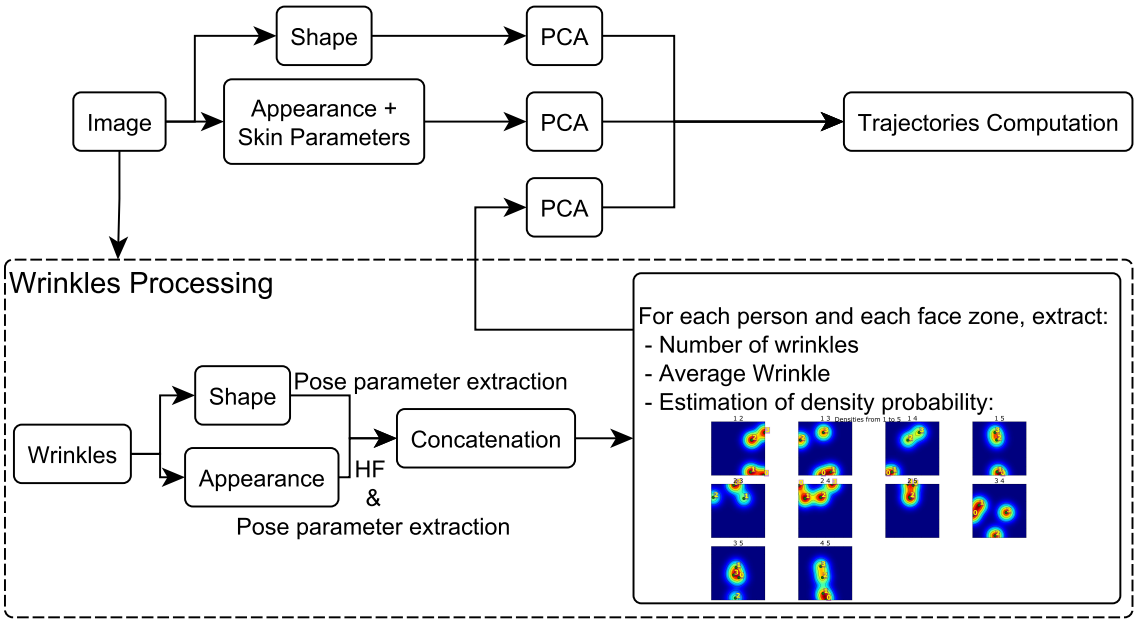

Figure 2.8 – Wrinkle Oriented AAM Scheme.

To analyze faces in the light of facial aging, we propose 3 contributions.

1. The first contribution is the parametrization of each wrinkle where shape and texture are represented altogether by a very understandable 7-length vector. Inversely, a wrinkle can be produced in shape and texture just from parameters from such a vector (Sec. 2.3.1).
2. To represent a group of wrinkles in one facial zone, we propose an approximation of an arbitrary joint probability of n random variables, as the set of every joint probability for every random variable taken two at a time; that

is our second contribution (Sec. 2.3.2). A new method of sampling for our approximated density mentioned above (Sec. 2.3.4).

3. Our third contribution is the parametrization of skin micro-texture and its integration within the Active Appearance Model (Sec. 2.3.5.3).

2.3.1 Wrinkle Model

We propose a separate model to analyze the shape and appearance variations of wrinkles.
First, wrinkles are annotated with 5 points for each wrinkle as shown in Figure 2.9. Afterwards, these 5 points are transformed into more explainable pose parameters containing:

— center (c_x, c_y) of wrinkle
— length ℓ which is equal to the geodesic distance between the first point and the last point of annotation
— angle a in degrees
— curvature $\mathcal{C}$ computed as least squares minimization of

$$\min \parallel Y - \mathcal{C}X^2 \parallel_2^2 \tag{2.13}$$

with Y (resp. X) the ordinates (resp. abscissa) of the wrinkle centered with the origin, and with first and last points horizontally aligned.

Here we just transformed the shape of a wrinkle in a 5-length vector $(c_x, c_y, \ell, a, \mathcal{C})$.

In addition, each texture wrinkle is extracted by making a bounding box around annotation and only keeping high frequency information by Difference of Gaussians (see Fig. 2.10). Here we blur texture with parameter $\sigma_b = 6$ and subtract blurring result with the untouched texture to make a high-pass filter and extract wrinkles. This filter has the advantage of being able to reconstruct perfectly the original image by simply summing the low and high frequency versions of the image. Here, as the wrinkle is high frequency information, we only keep the high frequency image and drop the low frequency version which contains skin color.
After that, wrinkle appearance is warped in the mean shape and then transformed in pose parameters. A second derivative Lorentzian function (Eq. 2.14) is fitted on each column and the average of every parameter found by fitting is kept (Fig. 2.11).

$$A * \frac{2\sigma\left(3\left(x-\mu\right)^2 - \sigma^2\right)}{\left(\left(x-\mu\right)^2 + \sigma^2\right)^3} + o \tag{2.14}$$

where μ and σ are respectively location and scale of the second derivative Lorentzian function, and, A and o are tweaking parameters to adjust the curve. Only A and σ are kept to characterize respectively depth and width of wrinkles.
Thus, we constructed a model able to transform a wrinkle in a set of 7 understandable parameters $(c_x, c_y, \ell, a, \mathcal{C}, A, \sigma)$, 5 for shape and 2 for appearance. On a side note, we can say that other pose parameters could have been computed.

Figure 2.9 – Example of wrinkles annotation in a face. Each wrinkle is annotated with 5 points, and a spline approximates its curve for visualization purpose. Two wrinkles of different colors are belonging to two different groups. The different groups will be listed in the next section (Sec. 2.3.2 p 56).

Taking the curvature parameter $\mathcal{C}$ as minimization of Eq. 2.13 is implicitly modeling wrinkle shapes as second order polynomials. For more accurate but more complex modeling, third or fourth order polynomials, or any parametric curve, could be used. Also, concerning appearance pose parameters, our modeling implicitly defines wrinkles as having uniform intensity and width. Instead of taking the average parameters (A, σ), several parameters (A_i, σ_i) could have been taken at different locations for each wrinkle appearance.
As wrinkles are analyzed in a separate channel (Fig. 2.8), wrinkles are removed from the appearance.

2.3.2 Robust Feature

The objective remains to obtain a representation of wrinkles for each face and to analyze them by applying PCA. As people have different numbers of wrinkles, we cannot just compute parameters for each wrinkle in a face and concatenate them to create a fixed-length representation usable with PCA. We have to find a fixed-length representation vector of wrinkles for each face.

We propose to estimate the probability density modeling the structure of wrinkles for each face and each zone.

Figure 2.10 – High frequencies extraction. Left: original image. Middle: Image Gaussian blurred with $\sigma_b = 6$. Right: Difference of Left and Middle Image to extract high frequencies. The parameter σ_b is relative to image resolution (i.e higher resolution implies higher σ_b), and can be found empirically.

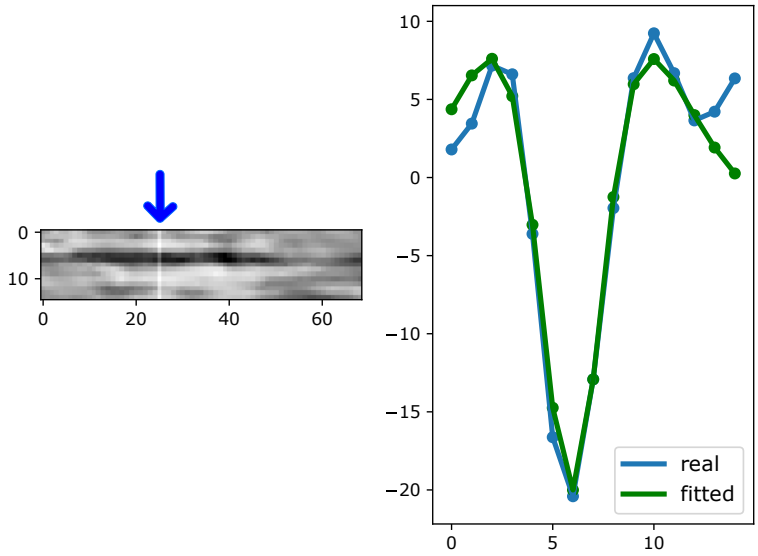

Figure 2.11 – Texture Fitting Example. Left: warped wrinkle; fitted column is highlighted. Right: in blue the pixels intensity variations and in green the fitting result.

Using the system introduced in Section 2.3.1, each wrinkle is represented by a 7-length vector. Faces are divided into 15 zones:

- 2 nasolabial folds
- 2 cheeks
- 2 crow's-feet
- 2 zones at the corner of the mouth
- 2 zones: below and above the mouth
- 2 zones below each eye
- forehead
- chin
- frown lines

We aim to compute a joint probability $P(d1, \ldots, d7)$ of wrinkles from each zone and each face. Unfortunately, such joint probabilities can have a very large memory footprint because of dimensionality, as the memory size of densities grows exponentially with dimensionality. To circumvent this problem, we propose an approximation of an arbitrary joint probability of n random variables by computing every joint probability for every random variable taken two at a time (Fig. 2.12). More precisely, we propose to approximate $P(d_1, \ldots, d_n)$ by the set

$\{P(d_1, d_2), P(d_1, d_3), \ldots, P(d_{n-1}, d_n)\}$. From now, when the number of dimensions n grows linearly, work memory no longer grows exponentially but quadratically $\Theta(\frac{n(n-1)}{2})$.
Joint probabilities are computed by Kernel Density Estimation (KDE) with a Gaussian kernel of standard deviation $\sigma_{kde} = 1.5$ for 60x60 densities; σ_{kde} parameter controlling the tradeoff between accuracy of wrinkles representation with a low σ_{kde}, and generalization with a higher σ_{kde}.

To better understand this approach, Figure 2.12 shows a forehead and the corresponding computed densities. The top left density $p(c_x, c_y)$ encodes the relationship between the position of wrinkles in abscissa and their positions in ordinate. Looking at this density, we see two Gaussians with high c_x values and two others with lower c_x values. It corresponds to the 4 wrinkles on the forehead, two on the left (low c_x values), and two on the right (high c_x values). Similarly, the next density $p(c_x, \ell)$ codes the relationship between the position in abscissa of wrinkles and their lengths. Looking at the picture, there are 3 wrinkles with a similar length, and a wrinkle shorter than others. As a consequence, the density $p(c_x, \ell)$ contains one Gaussian with a low ℓ value, and others with higher ℓ values.

Thus, for one face, we propose to extract a vector containing, for each of the 15 zones:

— number of wrinkles n_w in current zone,
— average wrinkle,
— densities computed with KDE on wrinkles where the average wrinkle was subtracted,

and to concatenate all 15 vectors to create the representation of wrinkles in one face. The number of elements in such a representation is: $15 * (1 + 7 + \frac{7(7-1)}{2} * 60 * 60) = 1134120$.

2.3.3 Linking with AAM

To sum up, we introduced three chained levels of representation.
1. The wrinkle level transforms each wrinkle on a picture of a face into 7 parameters $(c_x, c_y, \ell, a, \mathcal{C}, A, \sigma)$.
2. The zone level uses wrinkles in a zone, already converted to parameters at the previous level. At this stage, the number of wrinkles, the average wrinkles, and the ensemble of densities are computed and flattened to a vector to encode the variations of wrinkles in this zone. Such a vector will be named a zone wrinkles representation vector.
3. The face level concatenates all 15 vectors generated at the previous level to create the representation of wrinkles in one face. Such a representation will be named a wrinkles representation vector.
A third PCA is made for the new feature representing wrinkles. As for shape (Sec. 2.2.3.2 p 50), the PCA weights of wrinkles W_{wri} are normalized to have the

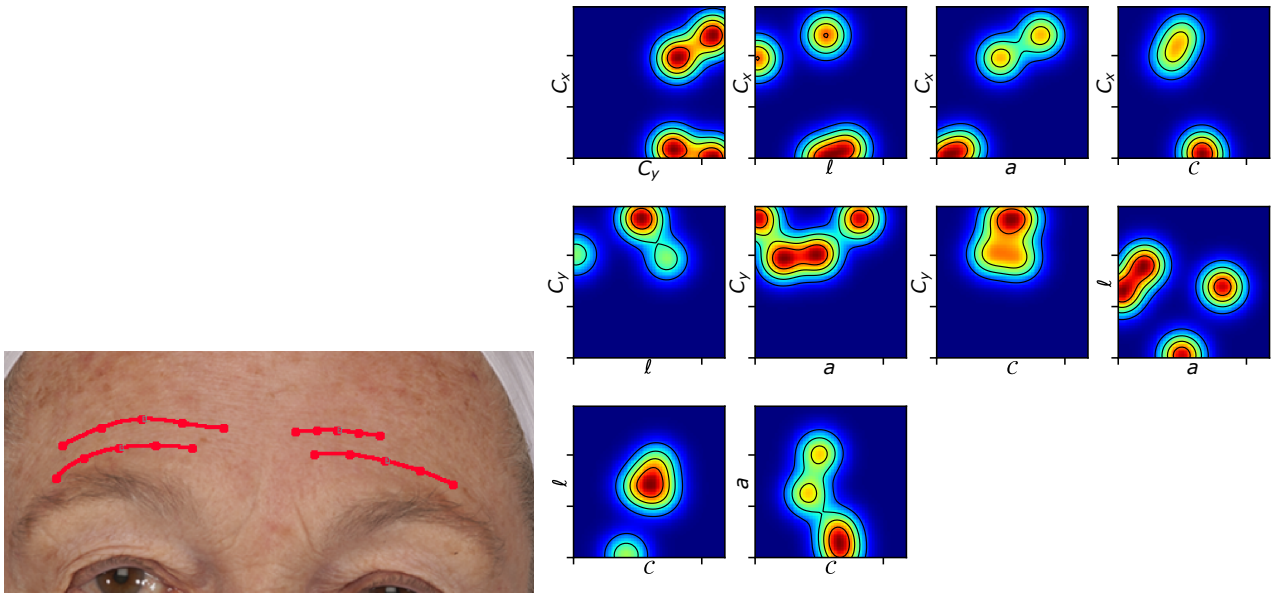

Figure 2.12 – Left: a forehead with 4 wrinkles. Right: ensemble of joint probabilities for the 4 wrinkles. With $n = 7$, there are $\frac{n(n-1)}{2} = 21$ densities; however, we only show 10 densities for convenient purpose (with $n = 5$ corresponding to $(c_x, c_y, \ell, a, \mathcal{C})$).

same global standard deviation as appearance weights W_{app} with Equation 2.15.

$$W_{wri} := \frac{W_{wri} \sum_{i=1}^{m} \lambda_{app,i}}{\sum_{i=1}^{n} \lambda_{wri,i}} \tag{2.15}$$

with λ_{app} (λ_{wri}), the eigenvalues of appearance (wrinkles) weights.
The final weights used for trajectories calculation are a concatenation of appearance weights, shape weights, and wrinkles weights.

2.3.4 Synthesizing Wrinkles

We now have a representation of wrinkles that we are able to incorporate in the classic AAM as seen on Fig. 2.8. PCA being perfectly invertible, we can reconstruct a shape, an appearance and a wrinkles representation vector from any point in the final PCA space.
The Sections 2.3.1 and 2.3.2 defined how to create a wrinkles representation vector from annotated wrinkles in one face. In this section, the inverse path is presented: how to generate wrinkles on a face from a wrinkles representation vector. First, the wrinkles representation vector is split into 15 zone wrinkles representation vectors encoding the variations of wrinkles in each zone. However we must define how to generate wrinkles from each of our zone wrinkles representation vectors.

We propose a new sampling method able to extract plausible wrinkles from a zone wrinkles representation vector, which contains joint probabilities. Algorithm's main point is finding a point iteratively, dimension after dimension, whose projections in each density is above a probability threshold; the threshold is decreased from 0.9 to 0.1 progressively to find the best candidate (Alg. 1 p 65).
First of all, peaks are found in $P(c_x, c_y)$ and SAMPLE function is called for each peak (p_x, p_y) found, from the peak with highest probability to the lowest.
We will present a step-by-step running of the function SAMPLE for a given peak $(39, 41)$. A vector $p = (39, 41, 0, 0, 0, 0, 0)$ is created which will contain the point's coordinate created by the function (Fig. 2.13).

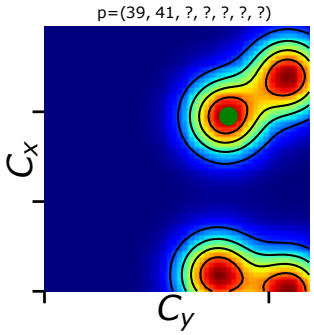

Figure 2.13 – The first two values of p are found by peak detection (the green point of coordinates (39, 41)).

After that, function GET_ARGMAX_MIN will extract two 1-D densities, $P(c_x = 39, \ell)$ and $P(c_y = 41, \ell)$, apply the minimum operator element-wise on them, and finally find the coordinate with highest probability ii such as $ii = \text{argmax}(\min(P(c_x = 39, \ell), P(c_y$ (Fig. 2.14 and 2.15). With $p_3 = ii$, if $P(p_3)$ is below the reference $P_{ref} = 0.9$, P_{ref} is decreased at 0.8; otherwise the search for p_4 begins with P_{ref} still equals to 0.9 and $p = (39, 41, p_3, 0, 0, 0, 0)$.
Here, $p_3 = 1$ and $P(p_3) = 0.52$, so P_{ref} is sequentially decreased from 0.9 to 0.8, then 0.7, then 0.6, and finally 0.5, where the value of $P(p_3)$ is accepted and the search for p_4 begins with P_{ref} equals to 0.5 and $p = (39, 41, 1, 0, 0, 0, 0)$.

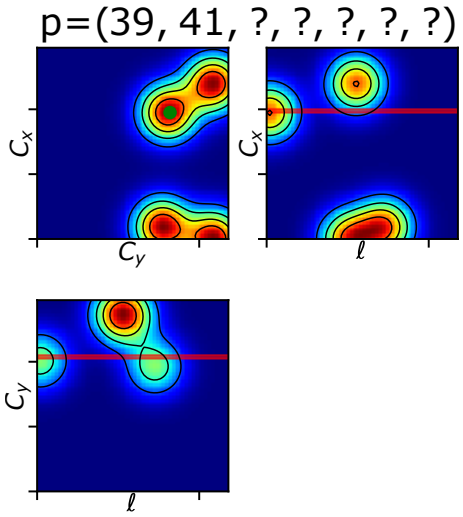

Figure 2.14 – The algorithm has to find $\ell = p_3$ a value that maximizes the probability in $P(c_x = 39, \ell)$ and $P(c_y = 41, \ell)$.

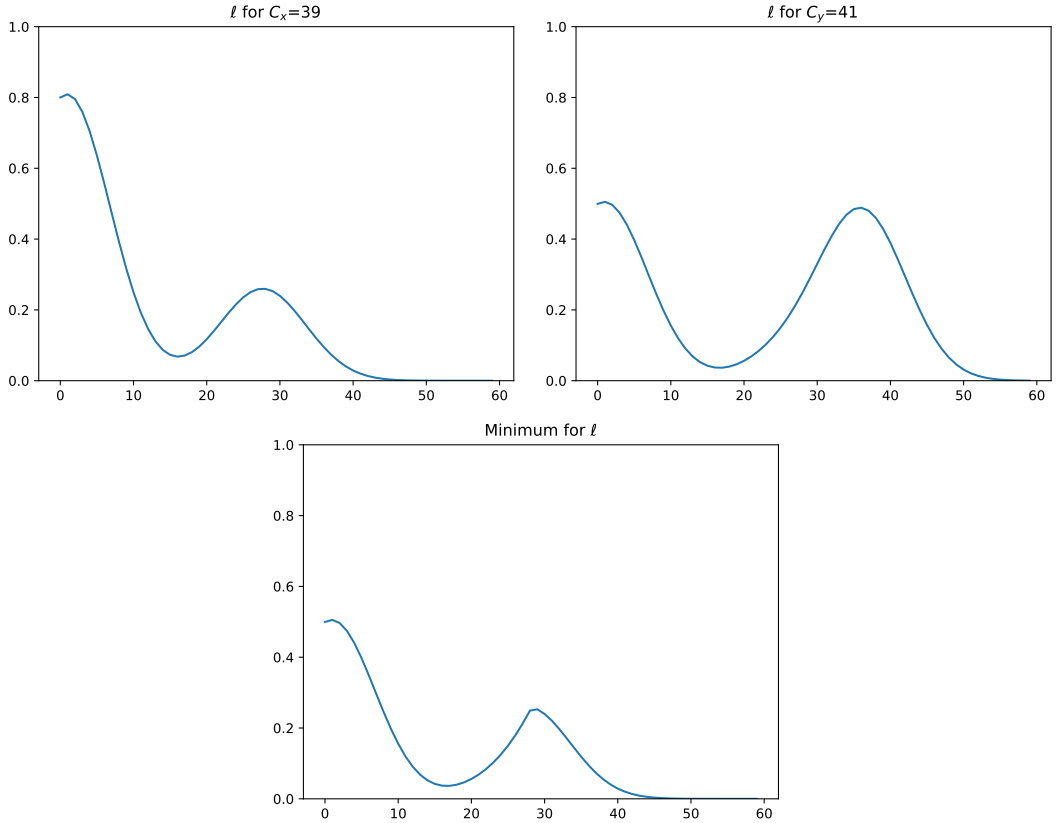

Figure 2.15 – The two extracted red lines on Figure 2.14 are the first two curves at the top, the third curve is the result of the element-wise minimum operator. We find that the maximum is obtained for $\ell = 1$.

For p_4, the same processing is made with the three 1-D densities $P(c_x = 39, a)$, $P(c_y = 41, a)$ and $P(\ell = 1, a)$. With $a = p_4$ found (Fig. 2.16 and 2.17), if $P(p_4)$ is

below the reference $P_{ref} = 0.5$, then backtracking starts: $P(c_x = 39, \ell = 1)$ and $P(c_y = 41, \ell = 1)$ are set to 0 and a new p_3 has to be found; otherwise the search for p_5 begins with $p = (39, 41, 1, p_4, 0, 0, 0)$.

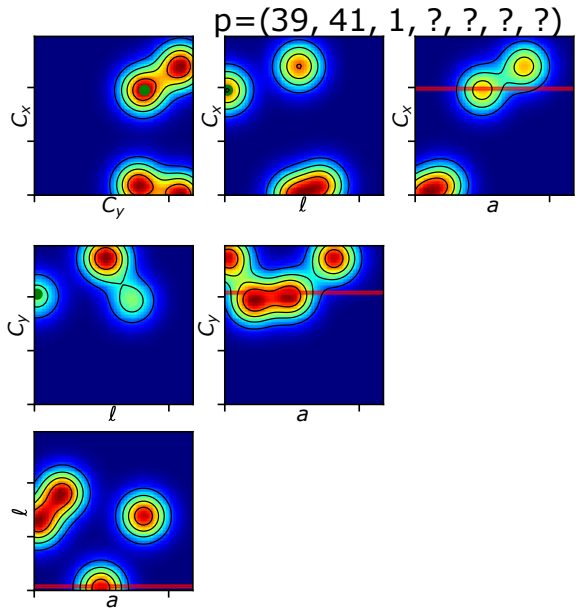

Figure 2.16 – The algorithm has to assign p_4 a value that maximizes the probability in $P(c_x = 39, a)$, $P(c_y = 41, a)$ and $P(\ell = 1, a)$.

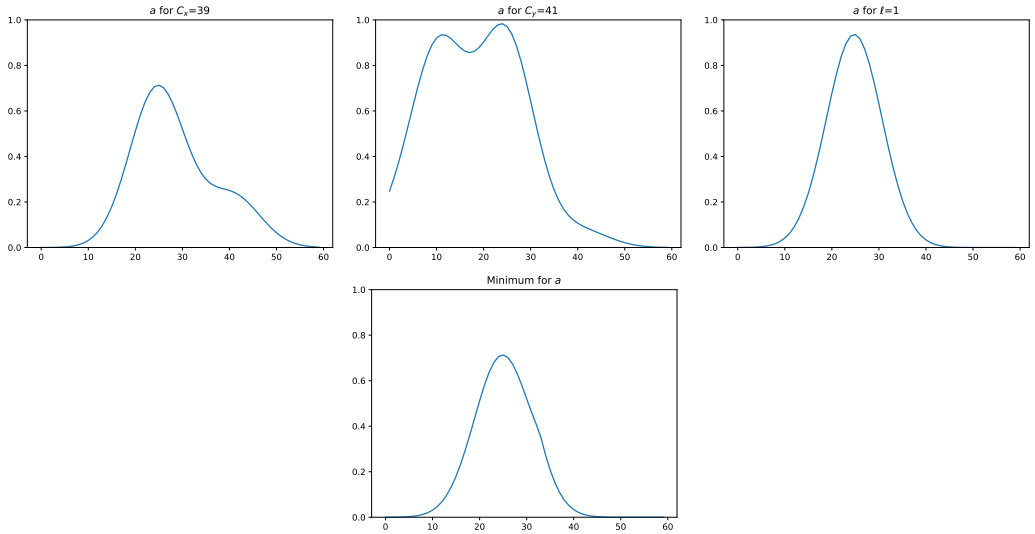

Figure 2.17 – The three extracted red lines on Figure 2.16 are the first three curves at the top, the fourth curve is the result of the element-wise minimum operator. We find that the maximum is obtained for $a = 25$.

As the algorithm keeps running, more and more cases are explored to finally get a point p which maximizes probabilities in densities given the starting peak

(p_x, p_y), and thus corresponds to a plausible wrinkle.
The zone wrinkles representation vector contains the number of wrinkles n_w to generate, the average wrinkle and the densities. We can create the n_w wrinkles parameters by running this algorithm n_w times and adding them to the average wrinkle.

Afterwards, we trivially have to produce wrinkles shape and appearance from parameters (see Sec.2.3.1 p.55 for definition of these parameters).
Shape is created from $(c_x, c_y, \ell, a, \mathcal{C})$ by sampling the polynomial defined by the curvature $\mathcal{C}$ until the specified geodesic length ℓ is reached. After that, points composing the shape are rotated according to angle a and finally center (c_x, c_y) is added to shape.
Appearance is produced by creating an empty image and variations of a second derivative Lorentzian function (see Eq. 2.14) of parameters (A, σ) are affected to each column (Fig. 2.18).
Finally, texture is warped in the newly created shape, for every wrinkle, and wrinkles are subsequently blended by adding the gradient of wrinkles with gradient of the underlying face.
To better understand the texture generation part, we will detail the synthesis of a wrinkle of parameters $(c_x, c_y, \ell, a, \mathcal{C}, A, \sigma) = (376, 757, 40, 91, 6.4e^{-3}, 15, 6)$. Figure 2.19 shows an example of a wrinkle warped and merged under a mouth. The input face is decomposed into a low frequency and a high frequency components, and the wrinkle texture in Figure 2.18 is warped into the shape defined at the previous step. An add operation is made between the warped wrinkle and the high frequency component. Finally, the face is reconstructed by adding the low frequency component and the modified high frequency component. We can appreciate how the added wrinkle integrates well with the other wrinkles and the underlying micro-texture, despite having only 7 parameters to encode wrinkles.

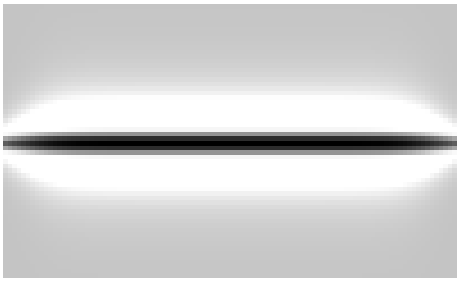

Figure 2.18 – Synthesized wrinkle using a second derivative Lorentzian function of parameters $(A, \sigma) = (15, 6)$. To generate a smooth wrinkle that will fit well with its environment on the face, a fade-in and a fade-out are applied at the start and at the end of the wrinkle on the parameters A and σ.

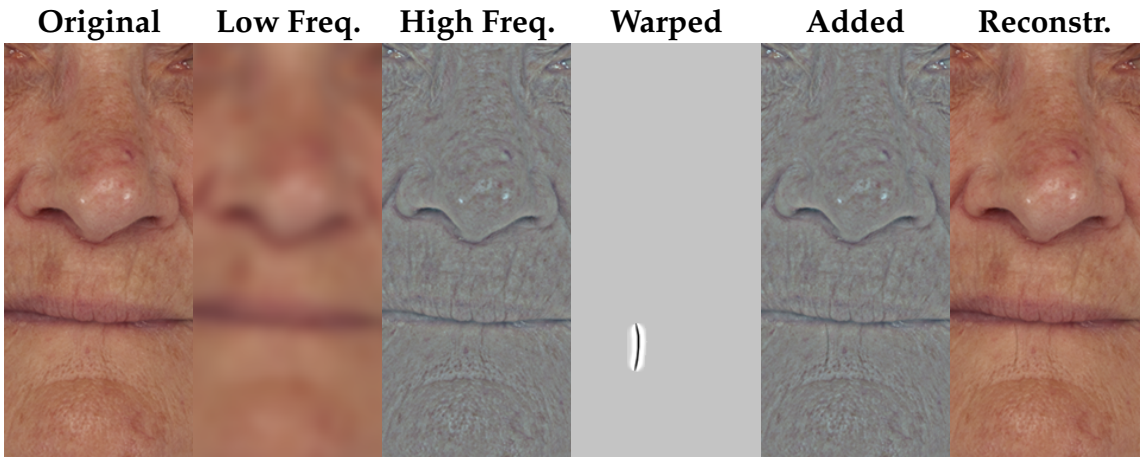

Figure 2.19 – Synthesis of a wrinkle of parameters $(c_x, c_y, \ell, a, \mathcal{C}, A, \sigma) = (376, 757, 40, 91, 6.4e^{-3}, 15, 6)$. Parameters have a simple interpretation. As $a = 91$, the wrinkle is rotated by 91°, that's why it's vertical. As $\ell = 40$, the wrinkle generated has a length of exactly 40 pixels.

Algorithm 1 Sample one point from an ensemble of joint probabilities

Require: D, list of every 2D joint probabilities

function GET_ARGMAX_MIN(D, n, p, i)
  $index \leftarrow i$
  $lines \leftarrow zeros(i, D_{index}.ncols)$
  for $j = 1$ `to` i **do**
    $lines[j, :] = D_{index}[p_j, :]$
    $index = index + (n - 2) - j$
  $min\_line = \min(lines)$ ▷ Min Pooling between lines
  $ii \leftarrow \text{argmax}\, min\_line$
  return $ii, min\_line[ii]$
function SAMPLE(D, n, p_x, p_y) ▷ densities, n dimensions and peak detected
  $p = zeros(n)$
  $p_1, p_2 \leftarrow p_x, p_y$
  for $P_{ref} = 0.9$ `to` 0.1 **do**
    $i \leftarrow 3$
    while $3 <= i <= n$ **do**
      $p_i, P(p_i) = Get\_argmax\_min(D, n, p, i)$
      if $P(p_i) \geqslant P_{ref}$ **then** ▷ Move on
        $i = i + 1$
      else ▷ Backtracking and removing path
        $i = i - 1$
        $index \leftarrow i$
        for $j = 1$ `to` i **do**
          $D_{index}[p_j, p_i] = 0$
          $index = index + (n - 2) - j$
  return p

2.3.5 AAM Modifications

2.3.5.1 Removing the final PCA

In addition to adding a new channel dedicated to wrinkle analysis as mentioned earlier, we decided to remove the final PCA from the classical AAM.
Historically, Active Appearance Model were first introduced to automate landmarking of images. For example, let's state that we want to landmark a face, given an AAM trained on faces. Using algorithms of optimization as gradient descent or metaheuristics, the goal is to find, iteratively, the best vector in the final subspace which reconstruction minimizes the square error with the input face, in pixels space. Thus, the final PCA is used as a mean of reducing the number of parameters in the search space, to speed up and facilitate the optimization process.
In the use we make of AAM, we don't need this additional reduction. Worse than that, the final PCA - in an unsupervised manner - could put on a same dimension variations linked with age and others which are not, perturbing the following

trajectories computation. Therefore, we removed the final PCA computation and used the concatenated weights $W_{app,sha,wri}$.

2.3.5.2 Aging Trajectory

The regression f such as $f(W_{app,sha,wri}) = age$ is a polynomial of degree 3. To reduce computational cost only interactions between the 10 first principal components of each of the 3 channels are kept.
We are going to explain our approach on a simple case: polynomial regression of degree 2 with 2 variables (x_1, x_2). A polynomial regression can be seen as a linear regression of polynomial variables.

Polynomial Variables First, a new set of 5 variables is computed from (x_1, x_2): $x_1^2, x_2^2, \mathbf{x_1 * x_2}, x_1, x_2$ (the interaction term is bolded).

Linear Regression Second, a linear regression is made on the new variables as in Equation 2.16.

$$f(x_1^2, x_2^2, x_1 * x_2, x_1, x_2) = a_1 x_1^2 + a_2 x_2^2 + a_3 x_1 * x_2 + a_4 x_1 + a_5 x_2 + a_6 \quad (2.16)$$

Hence, interactions are new variables which are products of different variables. For example, for a polynomial regression of degree 2 and 200 variables, a large number of variables has to be defined: 200 simple variables, 200 square variables, and $200(200-1)/2 = 19900$ interactions variables. The number of interactions variables grows quadratically with the number of variables and exponentially with the degree. Therefore, we decided to make a polynomial regression of degree 3, and to keep only interactions between the 10 first principal components of each of the 3 channels.

2.3.5.3 Skin Parameters

In addition to wrinkles, the micro-texture is an important cues to take into account to obtain a precise aging of faces, and to generate photo-realist images (Fig. 2.20).

Figure 2.20 – Two persons at different ages. Besides wrinkles, micro-texture is an important cue which help to distinguish these two faces.

The micro-texture in faces suffers from the same lack of modeling as wrinkles, due to the same problem as wrinkles: misalignment. As an attempt to integrate micro-texture, we developed a representation of the face accounting for texture in different spatial locations and frequencies. As shown in Figure 2.21, the face is divided in several blocks. In addition, the same face is split into 4 bands of frequencies as Figure 2.22. The split in frequencies is made using Differences of Gaussians, which is fast and perfectly reconstructible. Moreover, modification of one band or more does not produce many artifacts on the final reconstruction compared to critical-basis wavelet decomposition (Tiddeman et al., 2005; Zhang and Blum, 1999). Only the 3 first bands are kept as the last band with low frequency information does not contains any texture information.
Afterwards, for each block in each band of frequencies, the mean RGB (3 parameters) and the standard deviation RGB (3 parameters) are taken. With 600 blocks and 3 bands of frequencies, $600 * 3 * 6 = 10800$ skin parameters are obtained.

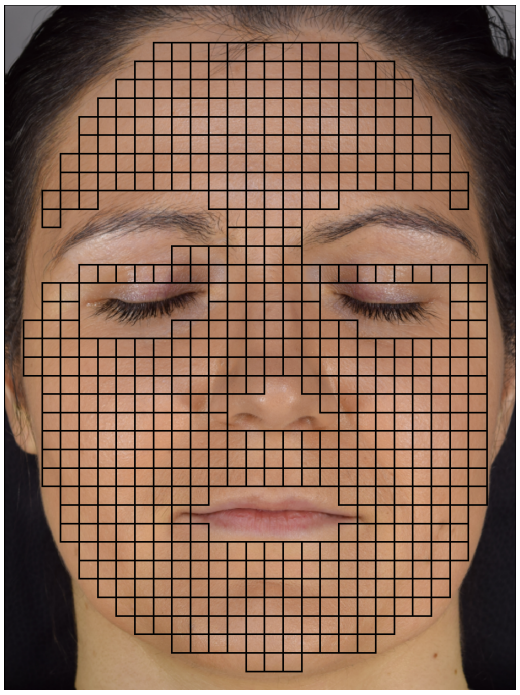

Figure 2.21 – The face is divided in several blocks.

Figure 2.22 – The face is split into 4 bands of frequencies.

The parameters are concatenated with the appearance vector containing the warped face. During aging, these parameters are likely to change. For reconstruction, we dispose of the warped face and the skin parameters, both modified by the AAM. From the new face, the same splitting in blocks and frequencies is made. Finally, mean and standard deviation of a block in a specific band of frequencies are altered to match the corresponding parameters.
This method helps the system learn correlation between micro-texture variations and aging without incorporating much knowledge on aging as for shape, appearance, and wrinkles.

2.4 Experiments

In this section, we present the database used in the experiments (Sec. 2.4.1). In the next section, we first present examples of aged and rejuvenated faces resulting from our model (Sec. 2.4.2), and after that we quantify the correlation between age progressed faces and the perception of these faces by an independent age estimation algorithm (Sec. 2.4.3). We show that our system is better correlated with the perception of age than the classic AAM (Sec. 2.4.3.2) and a state-of-the-art deep learning method.

2.4.1 Database

Our database consists of 400 Caucasian women taken in 2014, in frontal pose with a neutral expression and with the same lightning (Fig. 2.23). All faces are resized to 667x1000 resolution and annotated with 270 landmarks to locate eyebrows, eyes, mouth, nose, and facial contour (Fig. 2.24). Each face has been rated by 30 untrained raters to obtain a precise perceived age; perceived ages in the dataset range from 43 to 85 years with an average of 69 years. From this set of 400 faces, a subset with the same age range of 70 faces are chosen and 5 landmarks are placed on each wrinkle. Wrinkles annotation wasn't done on the whole database because it's a time-consuming task. Hence, the 400 faces are used to create the shape and appearance PCA subspaces, and the subset of 70 faces is used to create the wrinkle subspace and for the trajectory computation part.

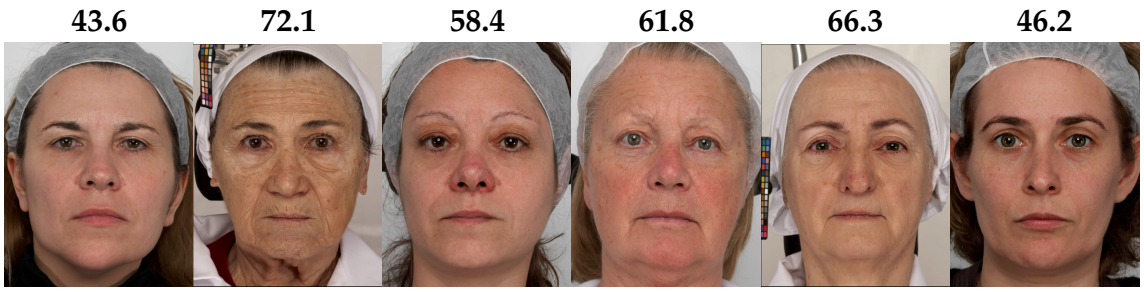

Figure 2.23 – A subsample of our database with their corresponding perceived ages.

Figure 2.24 – One face with an annotated shape (left) and annotated wrinkles (right). Each wrinkle is annotated with 5 points, and a spline approximates its curve for visualization purpose.

2.4.2 Qualitative Results

As seen on Figure 2.25, aging changes several known cues on a face (Donofrio, 2000; Farkas et al., 2004; Yaar and Gilchrest, 2007).
Concerning shape, the size of the mouth is reduced, especially the height of the lower mouth; eyebrows and eyes are both reduced as well, and we can see facial sagging at the lower ends of the jaw. Of course, the inverse phenomenon is observed with rejuvenating.
Concerning appearance, the face globally becomes whiter and yellowish, eyebrows and eyelashes are less present, and the mouth loses its red color as aging progresses.
With aging, more wrinkles appear and existing wrinkles are deeper, wider and longer. As we can see, new wrinkles created by our system are plausibly located with realistic texture (Figs. 2.26 and 2.27).
Furthermore, young skins are globally smoother than older ones. As shown in Figure 2.28, the performance of the system is improved by the addition of skin parameters.
The 3 faces shown here were not cherry-picked. The alterations made in terms of shape, appearance and wrinkles are representative of alterations made with aging/rejuvenating on the whole database.

In Figure 2.29, we check that our representation of wrinkles can be useful to generate more accurate average faces. With this representation, the average face contains wrinkles and their number, length and intensity are average in each area. As shown in Figure 2.30, the average operation made in the space of the wrinkles

Figure 2.25 – Face aging results. Left: Rejuvenating of 20 years. Middle: Original. Right: Aging of 20 years.

representation vectors, did not dampen the signal to the point where it disappears, in comparison to the same operation made in the pixels space.

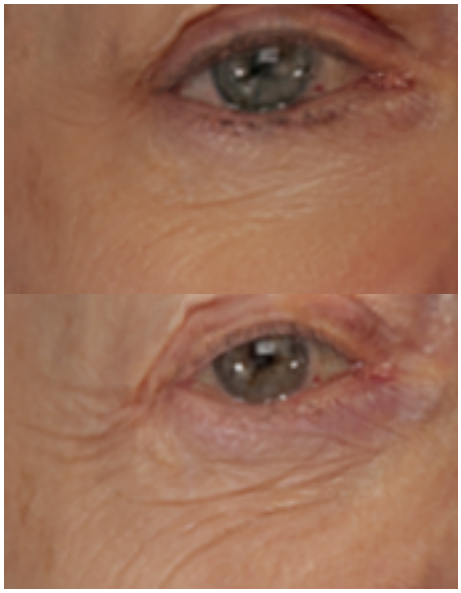

Figure 2.26 – Before and after aging wrinkles under the left eye. As we can see, the method doesn't produce any artifact nor suppress underlying micro-texture.

Figure 2.27 – Before and after aging wrinkles around the mouth. As we can see, the method amplifies existing wrinkles and create new ones.

Figure 2.28 – Rejuvenating with classic AAM (left) and with WOAAM (right). These pictures shows the effect of skin parameters. The right picture is slightly blurred in comparison to the left one due to a decrease of standard deviations in this location, thus, producing a face perceived younger.

Figure 2.29 – Average face made of 40 faces in the classic way (left), and using the representation of wrinkles (right). The right picture contains wrinkles and their number, length and intensity are average in each area.

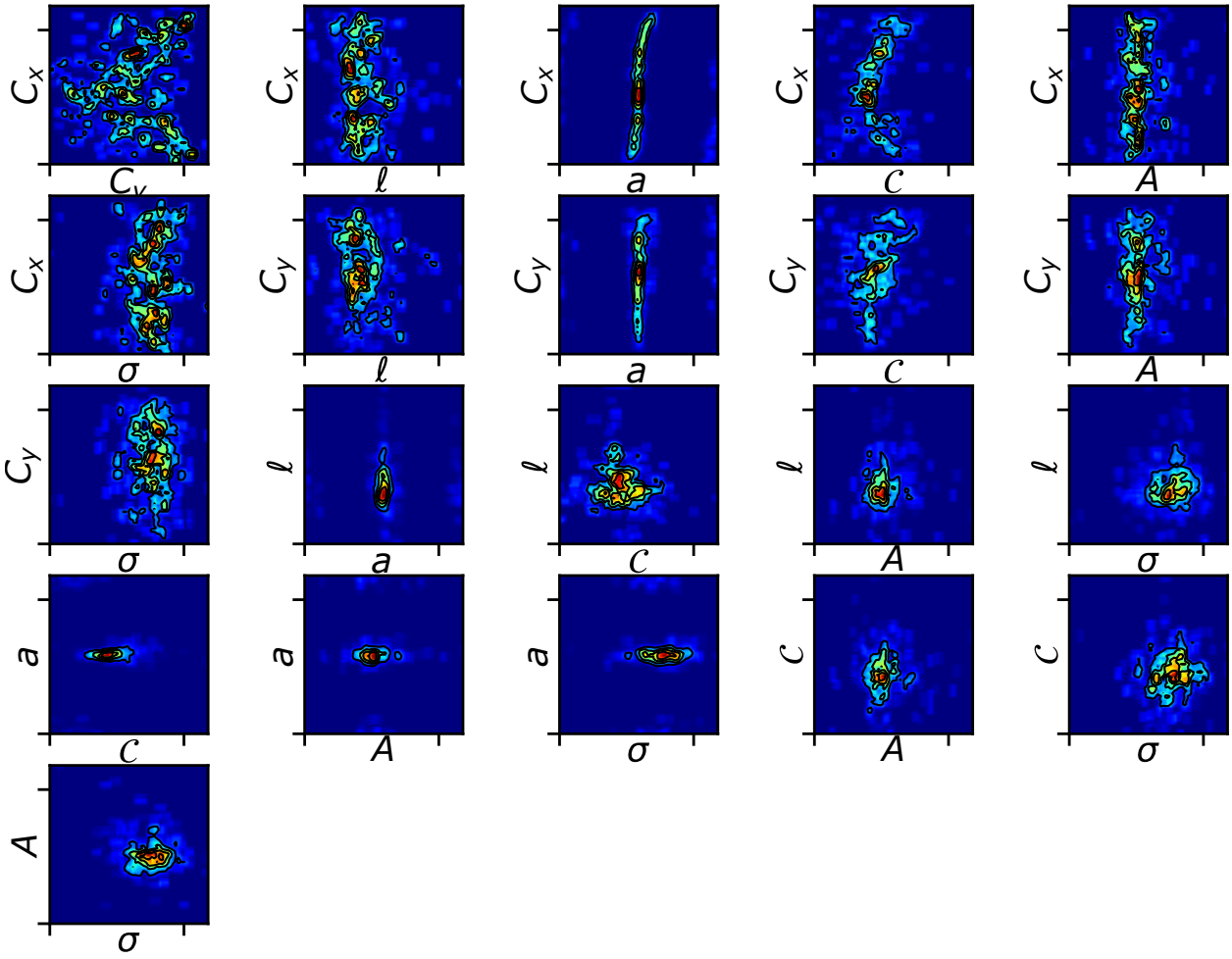

Figure 2.30 – The densities corresponding to wrinkles on the forehead for an average face made of 40 faces (Fig. 2.29).

2.4.3 Quantitative Results

2.4.3.1 Age Estimation

As in Rothe et al. (2015), we employed a pre-trained VGG-16 CNN (Simonyan and Zisserman, 2014) to create a face representation less sensitive to pose and illumination: we feed a picture as input where the face has been cropped and the representation produced is the output from *block5_pool*, the last pooling output. Afterwards, a Ridge regression is made in a 40-fold manner. As seen on Fig. 2.31, we obtain a R^2 score of 0.92 and an average absolute error and maximum absolute error of respectively, 2.8 years and 13.7 years. On the very same database, the average human estimates perceived age with an average absolute error and maximum absolute error of respectively, 5.5 years and 17.1 years. Here, the ground truth is the average of perceived ages by human, hence, that is not surprising that the regression performed better as it was trained to predict averages of perceived ages.

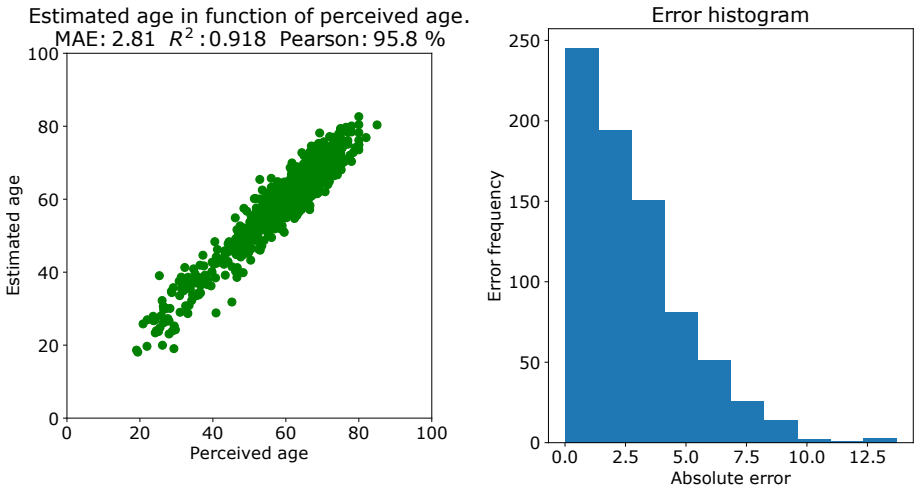

Figure 2.31 – Performance of our age estimation algorithm.

2.4.3.2 Comparison with prior works

For this experiment, we compare the perception of aged faces and the perception of rejuvenated faces for Active Appearance Model (AAM) (Lanitis et al., 2002), Conditional Adversarial Autoencoder (CAAE) (Zhang et al., 2017) and our method Wrinkle Oriented Active Appearance Model (WOAAM). For that, we used the age estimation system presented in the previous section. To test facial aging, we use faces with a perceived age of less than 60 years, and, for rejuvenating faces, a perceived age of 70 years and more. For AAM and WOAAM, each face is aged/rejuvenated 2 years at a time, and we compare, on average, the difference between estimated and expected age. For CAAE, each face is aged/rejuvenated 10 years at a time because this method use 10 discrete labels, and each label account for a 10-year interval.

As we can see on Figure 2.32 and 2.33, our method produces faces that are perceived as older than classic AAM and CAAE for aging, and younger for rejuvenating. In other words, a facial aging with WOAAM of y years better reduces the gap between the expected age and the age estimated by the age estimation system than a classic AAM or CAAE. For a 10-year aging period, the estimation of age has increased by 4.9 years for WOAAM, by 3.4 years for AAM, and by 2.9 years for CAAE. Also, for a 10-year rejuvenating period, the estimation of age has decreased by 4 years for WOAAM, by 2.3 years for AAM, and by 1.5 years for CAAE. On average, we improved performance by a factor of 1.5 over AAM, and by a factor of 2.5 over CAAE.
However, we can note that for a 10-year period of aging and rejuvenating, the estimation of age has been altered too slightly: respectively, by only 4 years and -3.4 years, which is low. This can be explained by the fact that we used only one aging trajectory, and because our model does not consider age spots.

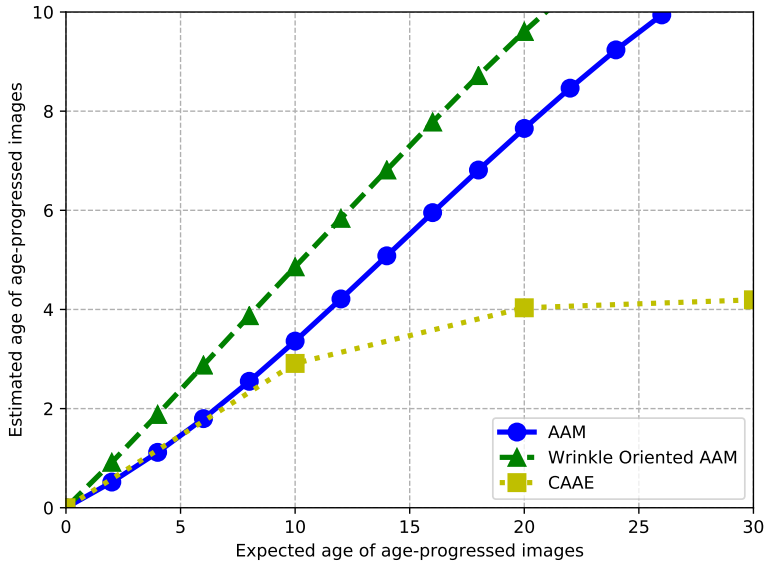

Figure 2.32 – Perception of faces aged of y years, in function of y going from 0 to 30 years, for the classic AAM, CAAE, and our Wrinkle Oriented AAM.

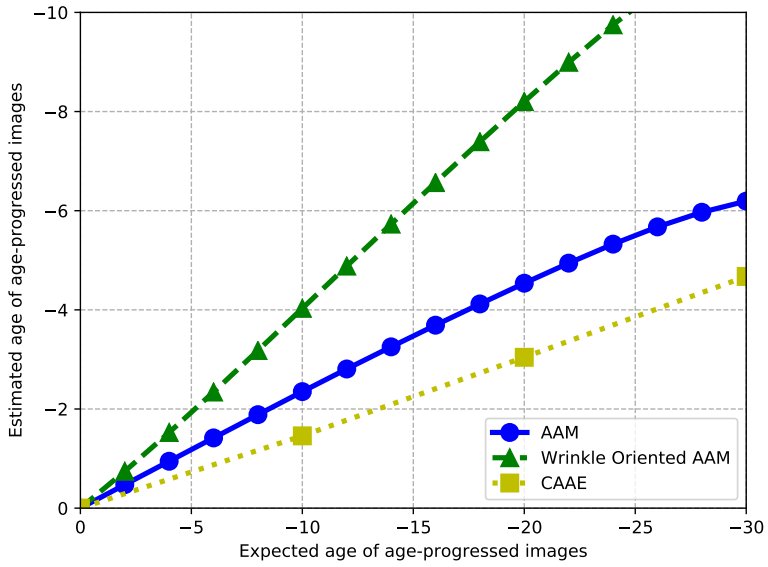

Figure 2.33 – Perception of faces rejuvenated of y years, in function of y going from 0 to -30 years, for the classic AAM, CAAE, and our Wrinkle Oriented AAM. Our rejuvenating system decreases the perception of age by only 4 years despite trying to rejuvenate by 10 years.

2.4.3.3 Understanding the effect of Aging on Wrinkles

Using our representation of wrinkles, age effects can be measured more accurately; we can quantify the effect of aging on wrinkles parameters. For each zone, various correlations between each wrinkle parameter and biological age are com-

puted. The same process is made using perceived age instead of biological age. The large tables containing the results are located in the annex of the document, Section A.1.1 p 103 for biological age and Section A.1.2 p 105 for perceived age. We use the following parameters in each zone: length, curvature, width, and intensity. For each parameter, average and maximum values are taken. Taking only the average could be confusing. As aging goes, new fine wrinkles appear, and the average length decreases. Hence, we could conclude that wrinkles length decreases with aging, which is known as false, despite being desirable. Therefore, the maximum value is taken in addition to the average. Besides these parameters, the number of wrinkles in each zone is taken too. We note that the results brought in this study are only valid for the population in our database i.e Caucasian women from 43 to 85 years with an average of 69 years.

Results shows as expected that the number of wrinkles increases with aging in most facial zones. In addition, the number of wrinkles is more correlated with perceived age than biological age. In terms of number of wrinkles, zones around the mouth are the most impacted by aging (Right & Left corner of the mouth, Top of & Below the mouth). Low correlation scores for forehead and frown lines, respectively 9% and 8%, can be explained by the fact that wrinkles usually appear at a younger age in these areas and are largely due to facial expressions.
Looking at the tables, the intensity of wrinkles are the most affected by aging (80% correlated with perceived age) – more than any other parameters. The most affected zones are the nasolabial folds and the corners of the mouth. Also, wrinkles width and length are 63% and 57% correlated, respectively, to perceived age. The area around the mouth is the most impacted zone (Top of & Below the mouth, Right & Left corner of the mouth). We can link this results with the work of Ezure et al. (2011): facial sagging is the most important visual cue affected by facial aging, and wrinkles at the corners of the mouth and nasolabial folds are mostly consequences of facial sagging; hence, the modeling of wrinkles seems to be consistent with clinical evaluation.

2.5 Discussion

We presented a new framework to analyze facial aging taking into account shape, appearance and wrinkles. We showed that the system can generate realistic faces for aging and rejuvenating, and such age-progressed faces better influence age perception than with Active Appearance Model or Conditional Adversarial Autoencoder. On average, we demonstrated an improvement factor of 2.0 over prior works.
Usually, image analysis software computes a global score from the face to assess the effects of skincare products or for dermatological analysis. The accuracy of the wrinkles modeling allows to go further in this kind of analysis.

Nevertheless, the model can be improved in several ways.
Firstly, the realism of the faces produced by the model has not been rated in this study.
Moreover, we know that facial aging is influenced by environmental factors like sun exposure, alcohol consumption or eating practices (Guyuron et al., 2009; Latreille et al., 2012). A potential next step could be to compute multiple trajectories in function of those factors.
Skin parameters as defined here are an easy way to integrate micro-texture. Skin parameters cannot create new pores or skin relief from a smooth skin, unlike the more complex wrinkles integration. The method can only amplify or dampen existing micro-texture. Thus, if the system learns that micro-texture increases with aging and we try to age a face with a smooth skin, change will be seen in terms of shape, appearance, and wrinkles, but nothing about skin relief and micro-texture. To circumvent this problem, a more complex representation of micro-texture has to be added to AAM, as we did for wrinkles.
In addition, dark spots must be included in the model to increase the accuracy of facial aging. Age spots could be incorporated in our model by creating a dedicated channel in our system, as we did for wrinkles. Afterwards, pose parameters of each age spots shape could be computed by fitting an ellipse to shapes and taking parameters of the fitted ellipses. Also, pose parameters of each age spots appearance could be computed by taking their mean RGB color. After that, we can carry out the same processing that we made for wrinkles. Firstly, to estimate the probability density modeling the structure of age spots for each face and each zone. Secondly, we can compute a PCA on our age spots representation vectors and connect the output to the final weights. Thus, aging trajectories would take into account age spots, in addition to shape, appearance and wrinkles.
Using only 2D images is not the best way to model a precise aging. Indeed, modeling in 2D does not fully catch 3D variations with ages as facial sagging, cheekbones alterations, or other variations of the distribution of fat mass in the face. An improvement could be to acquire a 3D database of faces, and to use WOAAM on it; shape would be composed of multiple 3D points, and appearance would be the texture flattened.

Chapter 3

Health Perception & Aging

3.1 Introduction

Judgments of a person's health based on facial appearance are a daily occurrence in social interactions. Understanding how we perceive health from a face is important because this judgment drives a wide array of social behaviors. Looking healthy has many positive real-life outcomes such as preferential treatment in the professional context, in the justice system or in dating interactions (Efran, 1974; Marlowe et al., 1996; Ritts et al., 1992; Spisak et al., 2014). Inversely, looking unhealthy is associated to lower self-esteem (Feingold, 1992) and may lead to a risk of social stigmatization and isolation (Henderson et al., 2016). A better understanding of how health is perceived and which facial cues alter this perception is likely to help reducing the negative social consequences which can follow.
Scientific recent evidences also show that facial healthy appearance is a good predictor of healthy behaviors (Whitehead et al., 2012) and good health (Re et al., 2011; Stephen et al., 2009a; Zebrowitz et al., 2014). Faces with an increase of oxygenated blood skin coloration are perceived healthier, and blood oxygenation level is known to be associated with cardiovascular fitness (Re et al., 2011). People with a healthy diet, such as daily consumption of fruits and vegetables, have a more attractive skin color and are perceived healthier (Whitehead et al., 2012). Sleep deprived people appear less healthy compared with when they are well rested (Axelsson et al., 2010). And people would acutely detect signs of sickness from the face in an early phase after exposure to infectious stimuli and potentially contagious people (Axelsson et al., 2017).
Previous works (Fink et al., 2011a,b, 2012) showed that perceived age and perceived health are correlated, however, facial cues used for the two perceptions aren't the same; health perception is influenced by facial adiposity and skin color while age perception is mostly predicted by skin aging signs.
Although several facial cues have been already identified as relevant cues to judge apparent health or to detect a general sickness state; more research are still necessary to better understand the mechanisms behind the perception of health.

To this end, we aimed to automatize human health perception by training a Convolutional Neural Network on a related task (age estimation) combined with a Ridge Regression to rate faces. Indeed, contrary to health ratings, large datasets with labels of biological age exist. The results show that our system outperforms average human judgments for health. Moreover, we highlight the fact that our system is influenced by the same facial cues as humans to judge health from a face. In this work, we introduce the first system able to estimate health scores from faces.

That's why we propose 3 contributions.
Firstly, as often in psychology research, the database we have at our disposal is quite scarce. We experiment to show whether a Convolutional Neural Network trained on a similar task with larger datasets available like biological age estimation combined with a simple estimator allows us to achieve great performance on our task.
Our second contribution is a new methodology where we show that our system is not only able to imitate judgments by humans, but more importantly that it uses the same main facial cues as humans.
In our last and third contribution, we highlight new links between facial features and heath perception.

3.2 Related Works

To the best of our knowledge, we introduce in this chapter the first work on automatic health estimation from face. Thus, the related works rely on the automatic age estimation from face.

Early works have been made by Kwon and Lobo (1999, 1994), they computed several distance ratios between landmarks at specific locations on faces to distinguish between the 3 classes babies, young adults, and seniors.
Lanitis et al. (2002, 2004) proposed to obtain a compact parametric description of face images using Active Appearance Model and to use this description to estimate ages. The authors tested a range of classifiers and regressions like linear regression, quadratic regression, cubic regression, and artificial neural network. This method presents the advantages of using both shape and appearance information from the face, and also to be more interpretable as we can display the most age correlated principal components.
Guo et al. (2009) proposed the Biological Inspired Features. Face images are firstly convoluted with several Gabor kernels extracting specific details in terms of scales and orientations. Secondly, the result undergoes a max pooling compensating for small translations and small rotations. Finally, the pooled feature is used with Support Vector Machines to estimate age with a low Mean Absolute Error.
Recent uses of deep convolutional neural networks have demonstrated great per-

Figure 3.1 – An excerpt of the Internet Movie Database with their corresponding biological age. As we see it above, the database contains faces with large variations in pose, illumination and color distribution. Pictures are resized to 224x224 before training.

formance and robustness on big datasets with large variations in pose and illumination. Rothe et al. (2015) proposed to use the ConvNet VGG-16 (Simonyan and Zisserman, 2014) pretrained on the ImageNet database for image classification. Thereafter, they finetuned it with a database of 500k celebrity faces to estimate biological age. Finally, they finetuned it again on the database of the ChaLearn LAP 2015 challenge which they won.

More recently, some researchers have begun to study whether it is possible to estimate less common attributes from the face such as intelligence (Qin et al., 2016), attractiveness (Chen et al., 2017; Chen and Zhang, 2016; Fan et al., 2017; Liu et al., 2016) or social relation traits (Zhang et al., 2015).

In view of the current state of art and our constraints, we use a Convolutional Neural Network trained on biological age combined with a Ridge Regression to assess health perception from faces (Sec. 3.3.1). Thereafter, we evaluate the system performance on our database and we compare it with human performance on the same database (Sec. 3.3.2). To validate our approach, we alter several facial cues known to influence human perception and we demonstrate that our health estimation system reacts in the same way as health perception by humans (Sec. 3.4.1). Finally, we establish the impact of specific facial features on health perception that have never been studied before (Sec. 3.4.2).

3.3 Health Estimation

3.3.1 Method description

Based on the age estimation method of Rothe et al. (2015), we employ the Convolutional Neural Network VGG-16 pre-trained on the ImageNet database (Simonyan and Zisserman, 2014) to detect 1,000 classes of objects, and trained it on the Internet Movie Database (IMDb) of celebrities (Fig. 3.1). We filtered the $\approx 500K$ images to keep only those containing faces with resolution greater than 120x120 pixels, no more than one face detected in each image, and only picture

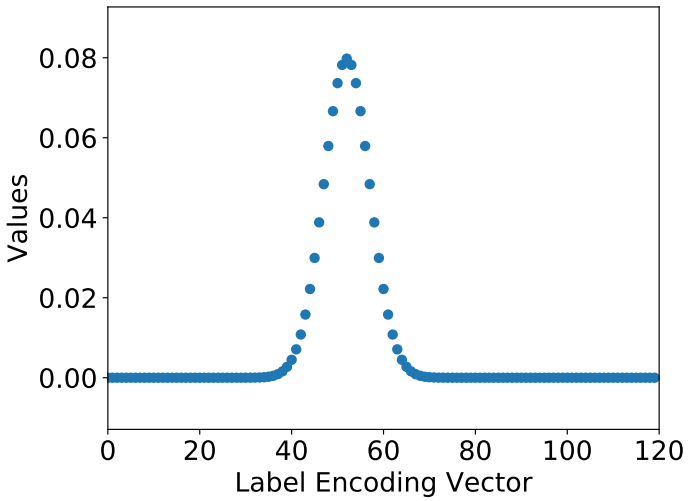

Figure 3.2 – Label distribution encoding with $N = 120, x = 32, o = 20$ and $\sigma = 5$. A vector of size 120 containing a discretized Gaussian centered on $32 + 20$ of standard deviation 5.

depicting people from 11 to 85 years old. Indeed, we want our application to work mainly on adult faces – the youngest faces being very different from the others, they were excluded. For each picture, we have the date of birth of the celebrity pictured and the date of the photo acquisition, thus we can deduce the biological age of the depicted person.
In addition, from the original VGG-16 architecture, we replace the final Multi Layer Perceptron containing a large part of the parameters, by a lighter one with one layer of 1024 units (Fig. 3.4) and an output layer of 120 units. The objective of doing so is to shift the learning effort onto the convolutional layers because the final Multi Layer Perceptron will be dropped as we want to estimate health and not biological age – thus, having the fastest training with the lowest score is not the main goal here. Using the same method as Antipov et al. (2016); Rothe et al. (2015), biological age estimation is seen as a classification method using *label distribution encoding* (Geng et al., 2013). An age $x \in \mathbb{R}$, is encoded by a vector of size N containing a discretized Gaussian centered on $x + o$ of standard deviation σ, with o an arbitrary offset. With $N = 120, x = 32, o = 20$ and $\sigma = 5$, Figure 3.2 display an encoded vector. With y the output vector of size 120 from the neural network, the age can be computed as: $\tilde{x} = \sum_1^{120} y_i * i - 20$.

Thus, the last 3 convolutional blocks and the fully connected layers has been trained on IMDb with Stochastic Gradient Descent with a Learning Rate of 10^{-4} on 1000 epochs with 10 steps per epoch and a batch size of 16. The decrease of the Mean Absolute Error for the training set and validation set can be seen in Figure 3.3.

After that, we have to develop our system of health estimation with only 140 images annotated with health scores (Fig. 3.5). Our database contains 140 photos of Caucasian women faces with a neutral expression in a controlled environment.

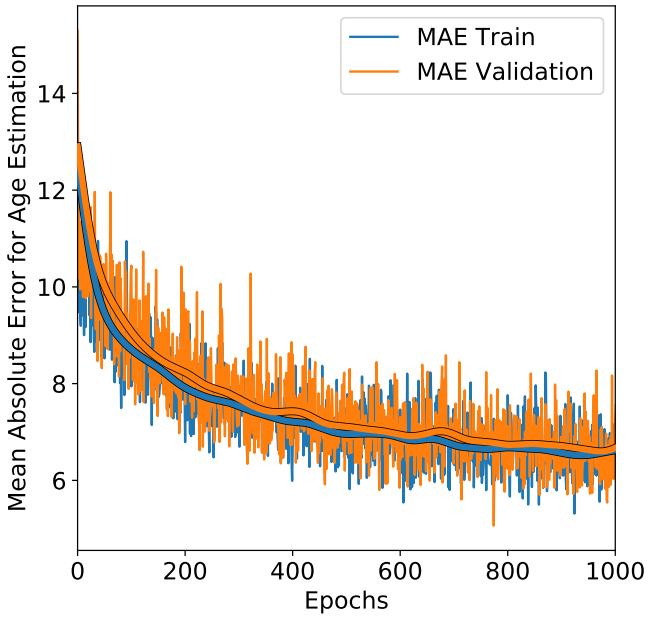

Figure 3.3 – Decrease of the Mean Absolute Error during the training for the train set and the validation set.

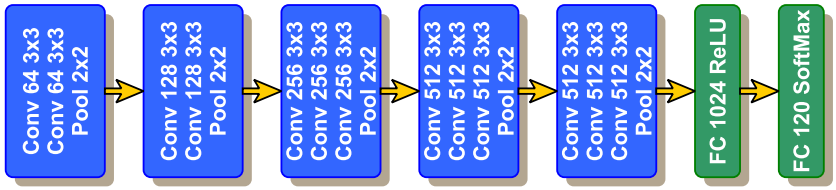

Figure 3.4 – Our architecture takes a 224x224 image and produces a probability distribution over all possible ages. The blue part has not been modified from the original VGG-16 architecture.

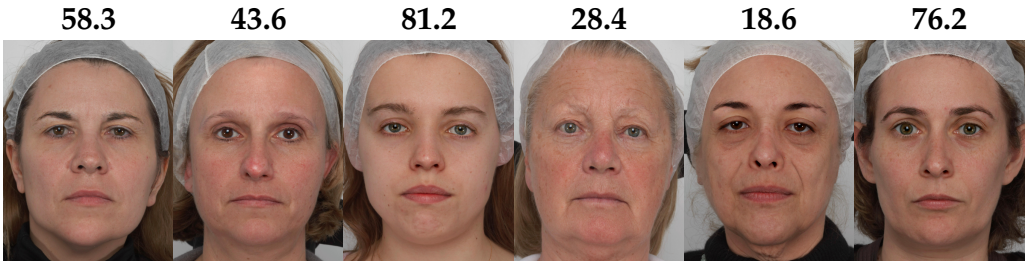

Figure 3.5 – An excerpt of our database with their corresponding perceived health scores. Faces were anonymized in this figure.

For every face, 74 raters were asked to evaluate health and to give a score from 0 to 100; 0 being perceived in "very bad health" and 100 being perceived in "very good health". Finally, for each image, we took the average of the 74 ratings to determine a reliable perceived health score.
We want to compute a representation of our faces from the newly trained ConvNet using only the convolutions and pooling blocks, and use a regression to estimate health scores from representations. The blue part in Figure 3.4 is kept to compute representations (feature extraction) and the green part is replaced by a regression. From an input image of size 224x224, the blue part outputs a vector of dimension d=512x7x7=25088 (512 features maps of dimension 7x7). The image and produced feature maps will pass through 5 successive 2x2 pooling layers - reducing its dimensions by two, five times ($224/2^5 = 7$).
The question remains, at which epoch can we stop the training for health estimation? If we take the weights at an early epoch, the system will be underfitted. In the same way, as biological age prediction is not the final goal, taking the weights corresponding to an advanced epoch with a low MAE is not the go-to choice to make.
We evaluate the suitability of ConvNet weights at each epoch for Health Estimation with a simple Linear Regression trained with a 40-fold configuration. We can see in Fig. 3.6 how the training on a different, but related, task can increase performance on our health estimation problem. At epoch 0, learning for biological age hasn't started yet and we get a relatively high MAE (9.0) for health estimation. In a second stage, learning for biological age greatly decreases Mean Absolute Error from 9.0 to 6.2. Finally, as learning progresses and the model specializes in biological age estimation, the error for health estimation increases. An optimal period is found around epoch 60 to take the weights for health estimation.

Now that we found the ConvNet weights to compute representations from faces, we test several estimators to asses health scores from representations. For each estimator, we evaluate a broad range of parameters and report those producing the best performance in Table 3.1. We used several estimators: linear regression without penalty, with ℓ_1 penalty (lasso), and with ℓ_2 penalty (ridge), support vector regression with a linear kernel and a Gaussian kernel, partial least squares regression, extremely randomized trees (Geurts et al., 2006), and K-nearest-neighbors. In the table, the Multi Layer Perceptron is composed of two layers containing n neurons for the first layer and 120 for the output layer.

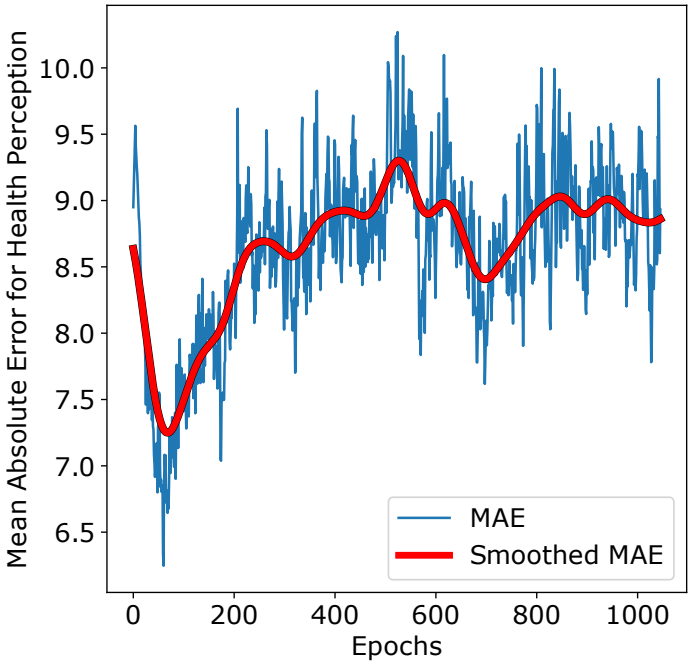

Figure 3.6 – Variation of the Mean Absolute Error in function of the epoch at which we choose the weights. Epoch 0 corresponds to VGG-16 just trained on ImageNet. The red curve has been Gaussian smoothed with $\sigma = 25$.

Table 3.1 – List of tested estimators. The estimator with the lowest Mean Absolute Error is bolded.

Estimator	MAE
Linear Regression	6.240
Ridge $\alpha = 10^{-3}$	**6.232**
Lasso $\alpha = 10^{-2}$	6.437
Linear SVR $C = 10^{3}$	6.355
RBF Gaussian SVR $C = 10^{4}, \gamma = 10^{-4}$	6.269
PLS Regression $n = 100$	14.64
Multi Layer Perceptron $n = 2048$	8.543
Extremely Randomized Trees $n = 200$	8.446
K-NN $K = 15$	8.778

As we can see on Table 3.1, simple estimators as a Linear Regression or a Linear Regression regularized with a low ℓ_2 penalty (Ridge Regression) can achieve the best performance given our dataset and the feature extraction method we chose earlier. We can explain the fact that simpler estimators perform better than more complex estimators as Random Forests or Multi Layer Perceptron by the scarce number of samples $n = 140$ in regard of the dimensionality of our features $d = 512 * 7 * 7 = 25088$.
The final architecture of our system is described in Figure 3.7. An input image pass through the network trained for biological age estimation freezed at epoch 60 and the output of the last pooling layer (feature extraction) is fed to a Ridge

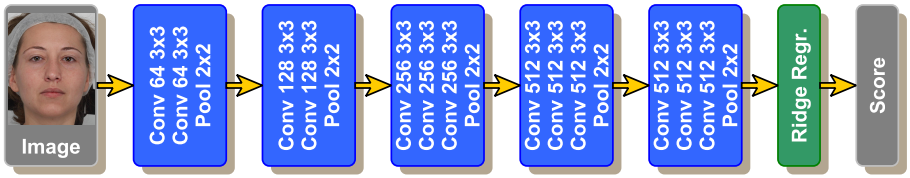

Figure 3.7 – The whole computation chain.

regression that will be trained for health estimation.

3.3.2 Experiment: System versus Human performance

We have 2 databases at our disposal:

A. 140 pictures of faces from Caucasian women , with perceived health scores given by 74 Caucasian women from the same age range (Fig. 3.5). Ages ranging from 20 to 60.

B. 140 pictures of faces from Chinese women with perceived health scores given by 74 Chinese women from the same age range. Ages ranging from 20 to 60.

Exploiting the previously described system, we trained the Ridge Regression in a 140-fold manner to assess its performance for each ethnicity.

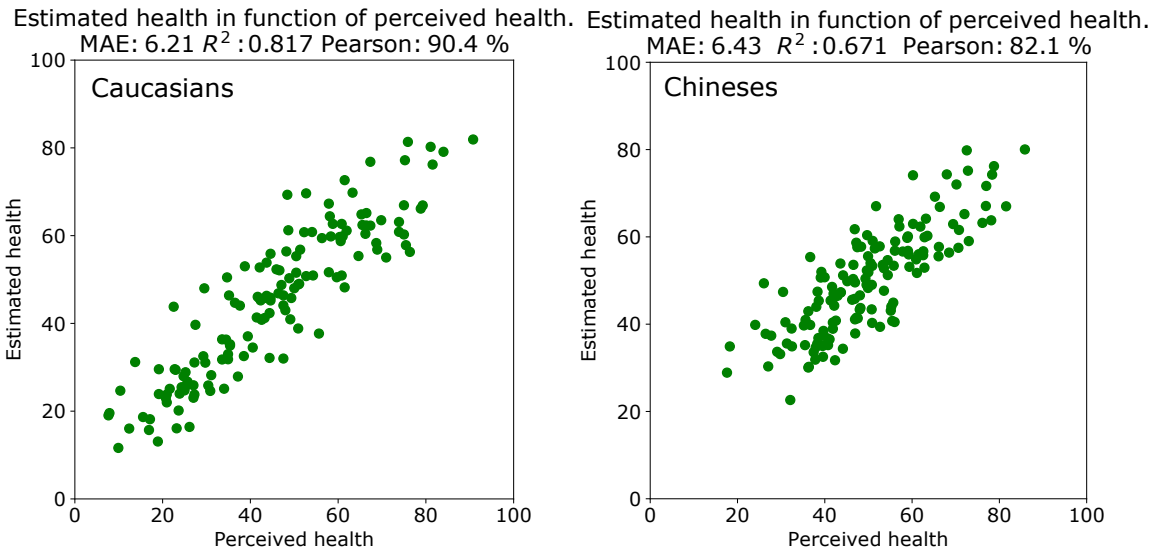

Figure 3.8 – Left graph: The predictions of the system compared to the perceived health scores for Caucasian faces. Right graph: The same has been done for Chinese faces.

As we can see on Figure 3.8, we can achieve good performance on our dataset with a scarce amount of data. The estimated health scores are 90.4% and 82.1%

correlated, for Caucasians and Chineses respectively, to the perceived health ratings. Using mean absolute error MAE, coefficient of determination R^2 and Pearson correlation PC, Table 3.2 shows that our system estimates health more accurately than an average human working on the same dataset.
In addition, among the 74 judges, one judge with the lowest MAE (i.e smallest difference in average between his ratings and the average ratings) is selected and placed in the table below under the name Best Human.

Table 3.2 – Performance of our health estimation system compared to human performance.

	Caucasians			Chineses		
	MAE	PC	R^2	MAE	PC	R^2
System	**6.21**	**90.4%**	**0.817**	**6.43**	**82.1%**	**0.671**
Average Human	18.4	64.7%	-0.387	20.23	46.6%	-2.60
Best Human	9.37	81.3%	0.637	9.04	77.1%	0.384

3.4 Understanding Health Perception and Facial Cues

The newly created system is able to estimate health scores from faces with a good performance on our datasets. We want now to employ our tool to understand how facial features influence the perception of health.
To this end, we alter several facial features known to influence health perception and assess the difference of health scores between the modified and the unmodified faces to highlight how our system use the same criteria as human to evaluate health (Sec. 3.4.1). Finally, we carry out the same process with several facial features which have never been studied before (Sec. 3.4.2).

3.4.1 Study on known criteria

As defined in the literature (Henderson et al., 2016), facial features like skin luminosity (Fink and Matts, 2008; Stephen et al., 2009a,b), red color of lips (Russell et al., 2016) or smile expression of the lips (Jones et al., 2018; Ostir et al., 2004) are known to greatly influence health perception in humans. Here, we test how these features influence the estimation of health of our system.
First, we convert our face images from RGB to the La*b* color space largely used in face perception research because it corresponds roughly to the color channel of the human visual system (see Appendix A.2 page 106 for more information on this color space). Second, we automatically modify them to amplify or soften each of the features (Fig. 3.9). Afterwards, we estimate health scores of newly modified faces, and we assess the average difference of health scores between the modified and the unmodified faces; we called this average the *Delta Score* in Table 3.3. In addition to compute the effect of each facial feature on health estimation, we compute a p-value with a paired t-test to quantify the statistical significance of the effect. Alterations displayed in Figure 3.9 were chosen empirically to be high enough to produce a visual effects while keeping a realistic rendering.

Figure 3.9 – Example of manipulation for facial features (”known features”). Four manipulations are shown: increase of lip redness a* (+10 units), decrease of lip redness a* (-10 units), modification of lip curvature (”smiling lips”) and increase of facial skin luminosity L* (+10 units).

In this study, we modified the smiling aspects of the face by manipulating the curvature of the lips only. With this simple manipulation we did not produce an authentic smile on the faces (Korb et al., 2014), so the effect of smiling on health estimation may have been underestimated.

Table 3.3 – Influence of ”known features” modifications on health estimation.

Modification	Delta Score	p-value
Lip: a* +10	0.61	**7.7e-10**
Lip: a* −10	-0.60	**1.3e-07**
Lip: Smiling	1.38	**2.1e-08**
Face: L* +10	5.56	**9.0e-32**
Face: a* +03	0.71	**8.1e-13**
Face: b* +03	0.84	**5.8e-18**

The results in Table 3.3 show positive and significant effects of increasing skin luminosity, redness and yellowness, and increasing lips redness and smiling on health estimation; but a negative and significant effect of reducing lips redness. In other words, female faces with redder and smiling lips, more luminous, yellow and red skin look healthier. These results are consistent with previous findings in psychology research performed with human judgments (Jones et al., 2018; Russell et al., 2016; Stephen et al., 2009b).

Figure 3.10 – Example of morphological manipulation for the eyebrows (width & height), the mouth (height), and the eyes (size).

3.4.2 Study on new criteria

Here, we repeat the previous process but, this time, by altering specific facial features with unknown relationship to health perception. In literature, barely any work have considered the link between morphological changes and perception of health, only the link between facial adiposity and perceived health (Coetzee et al., 2009; Henderson et al., 2016). That's why we propose to apply morphological changes to facial features such as the lips, eyes, and eyebrows, to quantify their impact on health estimation, and to discuss those results in light of literature. Figure 3.10 presents 4 examples of facial features manipulation in terms of their shape.

Table 3.4 – Influence of "unknown features" modifications on health estimation.

Modification			Delta Score	p-value
Eyebrow:	Width	+20%	1.77	**2.0e-14**
Eyebrow:	Height	+20%	0.06	0.51
Mouth:	Width	+20%	0.42	**5.8e-05**
Mouth:	Height	+20%	-0.001	0.97
Eye:	Size	-15%	-0.79	**1.9e-4**
Eyebrow:	Distance with Eyes	-50%	-1.20	**1.1e-3**

Using our system we determined the influence of facial features shape on health estimation. The relative size of the eyebrows, the mouth and the eyes is known to be sexually dimorphic (Bruce and Young, 1998; Re et al., 2013) and changes with age (Burt and Perrett, 1995; George and Hole, 1995; Samson et al., 2010). Here we show that the width of the eyebrows and the mouth, the size of the eyes and the

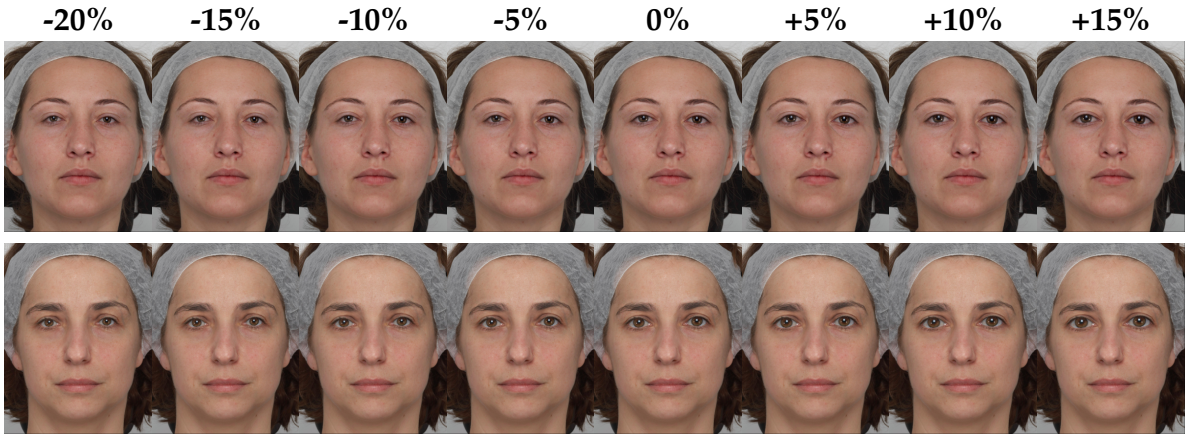

Figure 3.11 – Incremental modifications of eye size.

distance between the eyes and eyebrows influence the estimation of health. We found no correlation between eyebrow thickness and health perception, as the reported p-value is not significant. Female faces with bigger eyes, a larger distance eyes-eyebrows, larger lips and eyebrows look healthier.

In addition to determining the influence of facial features on health estimation, we can use our system to characterize this influence. We decide to modify eye size of faces from our dataset incrementally from -20% to +15% with a step of 5% (Fig. 3.11). Figure 3.12 shows the effect of this alteration on health estimates.

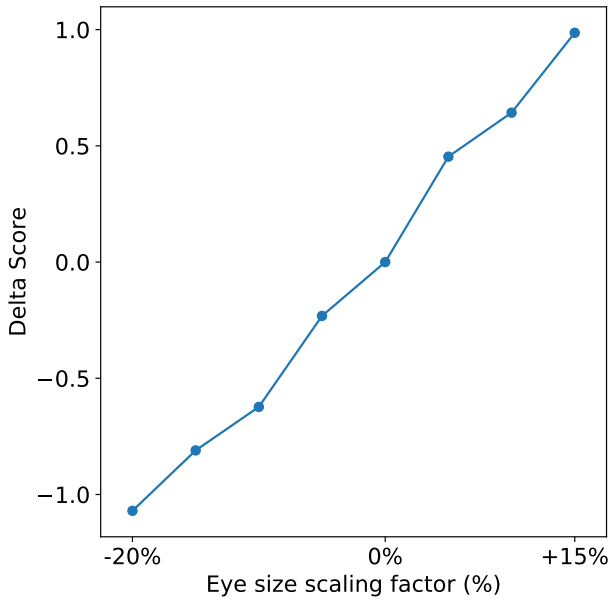

Figure 3.12 – Impact of eye size on health estimates. At each step, the variation of health scores is significant (with p being the p-value: $p < 0.01$).

Based on Figure 3.12, the modification of eye size influences linearly the estimation of health: the bigger the size of the eyes, the higher the health estimation.

Using this tool, the same process could be done with other facial cues to analyze their links with health perception.

3.5 Age & Health

In this section, we deepen the link between age and health perception. As shown in Figure 3.13, age and health perception are negatively correlated (-68%). With aging, health estimates decrease. Figure 3.14 shows the decrease of health estimates along aging using WOAAM. The 10 first individuals aged using WOAAM were taken from Section 2.4.3.2 p 75. Each color is an individual at a specific age. There are 31 points per individual as each face of age y has been rejuvenated/aged from $y-30$ to $y+30$ years, 2 years at a time. As we can see, during aging, our Wrinkle Active Appearance Model alters facial cues not only linked with age, but also cues linked with perceived health, consistently with a "real" aging.

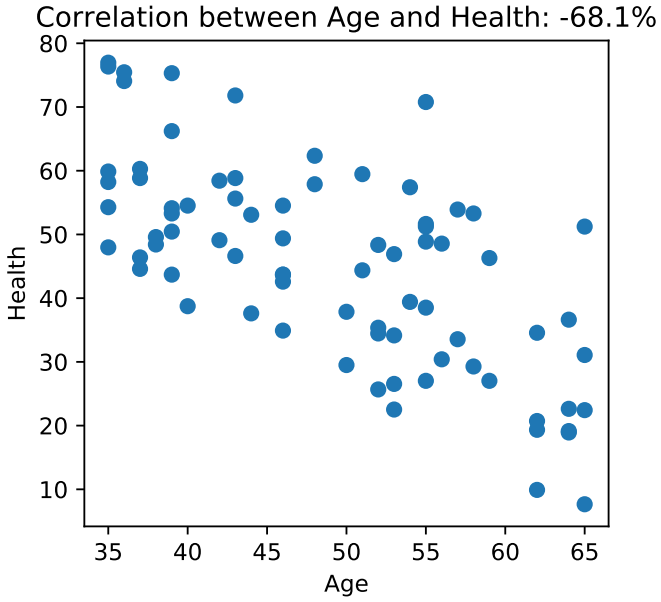

Figure 3.13 – Correlation between age and health perception of Caucasians.

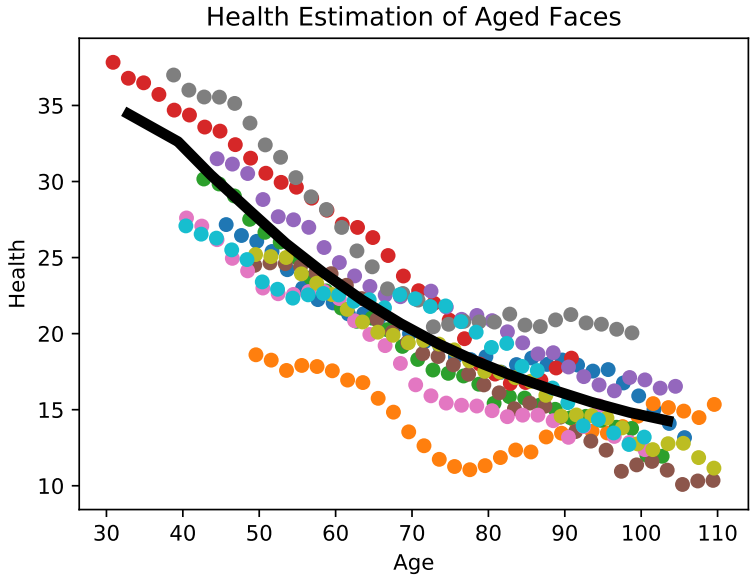

Figure 3.14 – Age in function of health estimated from rejuvenated/aged faces.

To give the reader an insight about distinctions between age and health perception, we present average faces made using health ratings decorrelated from age ratings.
Here, the database used have never been rated in terms of health by humans, the tool presented earlier was in charge of giving health ratings. As health perception and age are known to be correlated (Fink et al., 2011b), health ratings are linearly decorrelated with perceived age. To make this decorrelation, first we define a linear regression to predict health h from perceived age a:

$$\underset{\beta_0,\beta_1}{\operatorname{argmin}} \sum_{i=1}^{n} (h_i - f(a_i))^2 \quad \text{with} \quad f(a_i) = \beta_1 a_i + \beta_0 \tag{3.1}$$

Afterwards, decorrelated health ratings are defined as the residuals of the function f:

$$h'_i = h_i - f(a_i) \tag{3.2}$$

In doing so, we remove the variability in health ratings due to age ratings.
Finally, the ten faces with the greatest decorrelated estimated health are warped in a mean shape and averaged pixel-by-pixel into one average face; the left face in Figure 3.15. The same has been done for the ten faces with the lowest estimated health.
As we can see, two different faces are produced, highlighting facial cues that influence the estimation of health without influencing the estimation of age: lightness, colors, and facial expressions (smiling face at left vs. non-smiling face at right) seems to be important cues for health estimation decorrelated with perceived age.

Figure 3.15 – Average face of the 10 faces with the greatest estimated health to the left, and with the lowest estimated health to the right. Health ratings have been decorrelated from perceived age.

3.6 Discussion

This chapter described how we managed to develop an automatic system able to imitate human judgments of health. We trained a Convolutional Neural Network to estimate biological age and we used representations produced by the network of our scarce database to train a simpler estimator. We observed a very good performance of the system when we compared it to human judgments of health. More importantly we found that the system bases its health estimation on the same facial information than humans. We took advantage of that finding to identify new links between facial features and health perception. Similarly, we presented links between health estimation by our system and aging using faces from different individuals, and faces from the same individuals rejuvenated and aged with WOAAM.

Nevertheless, we identified several areas of improvement.
First, the use of a Linear Regression to rank the different ConvNet weights (Fig. 3.6) tends to favor this type of estimators in the next step where we compare the performance of different estimators (Tab. 3.1). We could have ranked the different weights using a multitude of estimators. In the same way, the use of a dataset to rank the weights can be seen as a form of learning on this specific dataset, introducing a risk of overfitting. However, the use of the never used before dataset with Chineses faces shows that this ranking generalizes well with other datasets and ethnic groups.
Moreover, by using more images annotated with health ratings, we could improve the performance of our system and make it more robust to variations in

pose and illumination.
To augment our database, we tested to generate new faces from our scare dataset by making specific average faces. One's facial cues are "pasted" onto the head of another face and the shape of the facial cues are warped into an average of the facial cues from both faces (Fig. 3.16). The Poisson blending (Pérez et al., 2003) is used to smoothly remove potential skin color differences. Afterwards, this new faces were annotated by humans in terms of health. Unfortunately, adding this new database in our training set has only increased our performance by a few percent.

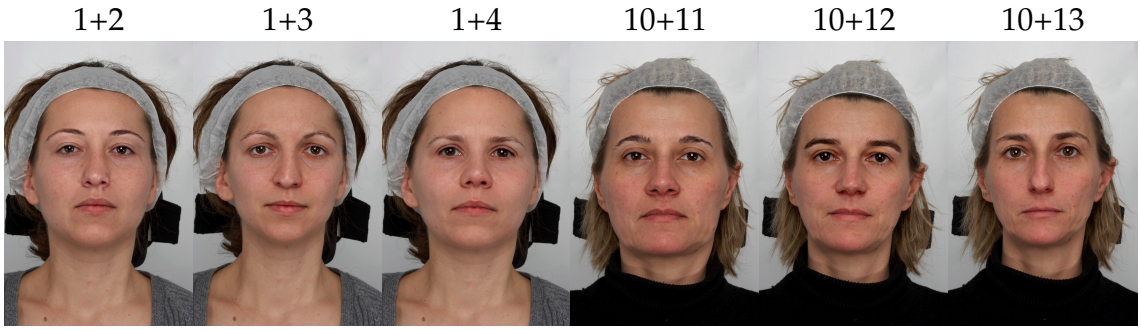

Figure 3.16 – The first face is the first picture of our dataset mixed with the second one. The second face is the first picture of our dataset mixed with the third one. The third face is the first picture of our dataset mixed with the forth one. For the next three faces, the same process was done starting not from the first face but from the tenth one.

Additional work will be necessary to test its performance on other demographic groups such as other ethnicities and men.

To conclude, we developed the first automatic health estimation system able to reproduce human judgments and show how it can contribute to better understand the perception of health. Such a system will help to further explore the use of facial cues and, more importantly, their validity in health judgments. Moreover, this tool will improve the evaluation of the properties of different makeup conditions, as we can see in the *Additional Works* section in the next chapter.

Conclusion

Preamble

We presented a framework able to age an input face in terms of appearance, shape, wrinkles, and microtexture. This work is the first to directly integrate wrinkles in the Active Appearance Model. First by introducing a new model of wrinkles, second by representing wrinkles distribution in a facial zone in an innovative manner. We showed that this approach outperforms classical AAM and a GAN-based approach. We believe that this method could be extended to integrate dark spots in AAM. The accuracy of our representation of wrinkles can help us, first to understand more precisely the effects of cosmetic and skincare products on the face, second, to build a better skincare evaluation process in analyzing wrinkles evolution over time more finely.

After that, we introduced a new system able to estimate health from faces to study the link between perceived health and aging. We showed how using a similar task with a large amount of annotated data available (biological age estimation) allows us to obtain a good performance with a scarce amount of data on our task (perceived health estimation). More importantly we found that the system bases its health estimation on the same facial information than humans. We took advantage of that finding to identify new morphological links between facial features and health perception. As perceived health is linked to the "real" biological health state, the system could be used on a daily basis to detect early signs of sickness. We suggest that a quick decrease of this index over time for a face could be an alert on a potential declining health condition, but this has to be demonstrated.

Future Work

During our investigation, some points have not been deepened sufficiently. Sometimes, they are related to the model, sometimes to the methodology. We consider that treating these points would enhance efficiency, we report them as follows:

— Using 3D faces could help the precision of face aging. Indeed, the facial aging process is largely influenced by 3D fat displacement under the skin and alterations of the facial skeleton (Bellity and Bellity, 2017; Mendelson and Wong, 2012). Facial sagging can cause wrinkle formations on the skin

surface. Thus, in using only 2D images, ones cannot integrate all the complexity of facial aging. Using 3D images would allow to model 3D shape deformations and to understand the causal effects between them and alterations of the skin surface.

— We focused our latest work on health estimation, but we could have followed the work on face aging making multiple aging trajectories. Indeed, we know that facial aging is influenced by environmental factors like sun exposure, alcohol consumption or eating practices (Guyuron et al., 2009; Latreille et al., 2012). A potential improvement could be to compute multiple trajectories in function of those factors. With such multiple trajectories, we could more precisely make a face ages. In addition, we could make a face ages in different aging directions and to show the results for prevention purpose, and to motivate healthy behaviors. However, computing multiple trajectories instead of one is trivial, and the focus on perceived health seemed more interesting from the company's point of view, thus we preferred to explore a new path in choosing to develop a health estimation system.

— As the datasets used contains only Caucasians, or Chineses concerning the work on health estimation, the obtained results are only valid for these specific ethnic groups. Despite being confident in the intrinsic qualities of our systems to generalize to other ethnicities as well, we cannot ensure it. Hence, to use databases with faces from other ethnic groups would be desirable to reinforce our systems robustness. Similarly, our work is focused only on women's faces, men could be added to our datasets.

— Similarly, this work could be extended to less constrained databases. We are currently working on extending it to "selfie" images shot with various illumination conditions; allowing us to deploy it on mobile devices.

Additional Work

When making a thesis in the industry, it's rare to work on only one subject during the 3 years. Thus, we had the opportunity to work on small but interesting projects.

Automatic clinical evaluation

In the department, dermatologists are asked to rate the severity of clinical aging signs. For example, when they are shown a picture of a face zoomed onto a nasolabial fold (Fig. 3.17), they have to rate on a scale from 1 to 7 the severity of aging of this zone. However, this clinical evaluation is very time-consuming: to evaluate all signs in a face, it takes a dermatologist between 10 and 15 minutes. We introduced a Proof-of-Concept able to give clinical score to zoomed pictures of nasolabial folds, with the same accuracy than dermatologists.

First, pictures are convoluted with several Gabor kernels followed by a Max Pooling operation for the feature extraction part, and Extremely Randomized Trees are used to make the decision. The system is trained with 300 images with ratings of a dermatologist, in a 10-fold configuration. In the first line of Table 3.5, we can see that ratings estimated with our system are 90% correlated with the ratings produced by the dermatologist on new input images. To give the reader an insight about this accuracy, we propose to measure intervariability and intravariability. The same dataset is rated by two dermatologists, and we see that ratings made by the first dermatologist is 85% correlated with the ratings produced by the second one, that's the intervariability. In addition, we asked the first dermatologist to rate the same dataset a second time, one month after the first scoring session, that's the intravariability. We see that ratings made by the first dermatologist is 91% correlated with the ratings produced one month later. We can observe that our system seems to have an accuracy close to dermatologists. However, more experiments would be needed to reinforce the robustness of this system.

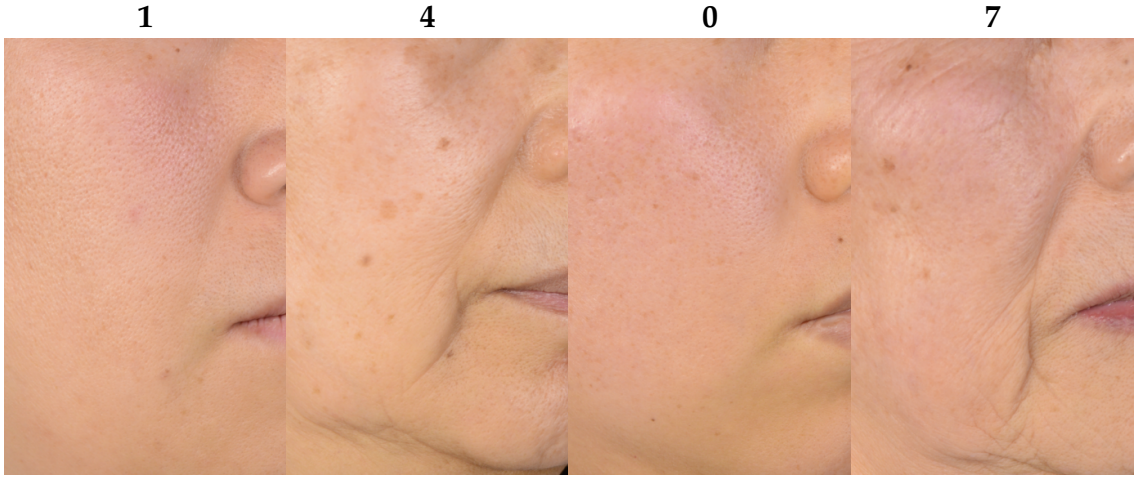

Figure 3.17 – Nasolabial folds and corresponding ratings by a dermatologist.

Table 3.5 – System Performance versus Human Performance.

	MAE	R^2	PC
Human1 vs System	**0.71**	**0.78**	**90%**
Human1 vs Human2	0.81	0.68	85%
Human1 T1 vs Human1 T2	0.65	0.81	91%

Microscopic Image Analysis

Biologists in the department have to assess the hydration power of active molecules on living cells on a regular basis. They apply on cells a solution containing the active molecule to test, before analyzing these cells with a microscope. The higher the fluorescence seen in the image, the better the hydration. We proposed a software to automatically extract the level of fluorescence from a cell image, thus

allowing to rank multiple molecules by hydration power.

The fluorescence is determined by the intensity of each cell. As the first image A in Figure 3.18, the intensity of the light is not the same in every zone of the image, avoiding an accurate measure of fluorescence. Using difference of Gaussians, we achieve to greatly dampen this variation of lightness. After that a simple thresholding is made to extract the average intensity value from cells. In Figure 3.19, the variations in the hydration power of vitamin E with time were computed with this tool.

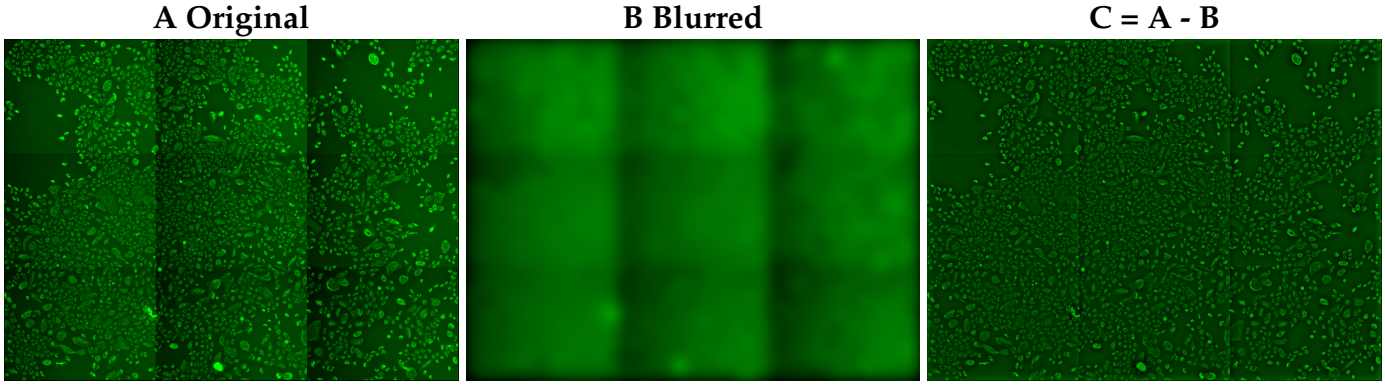

Figure 3.18 – A: Microscopic image. B: Gaussian Blurred. C: The difference between the original image and its blurred version.

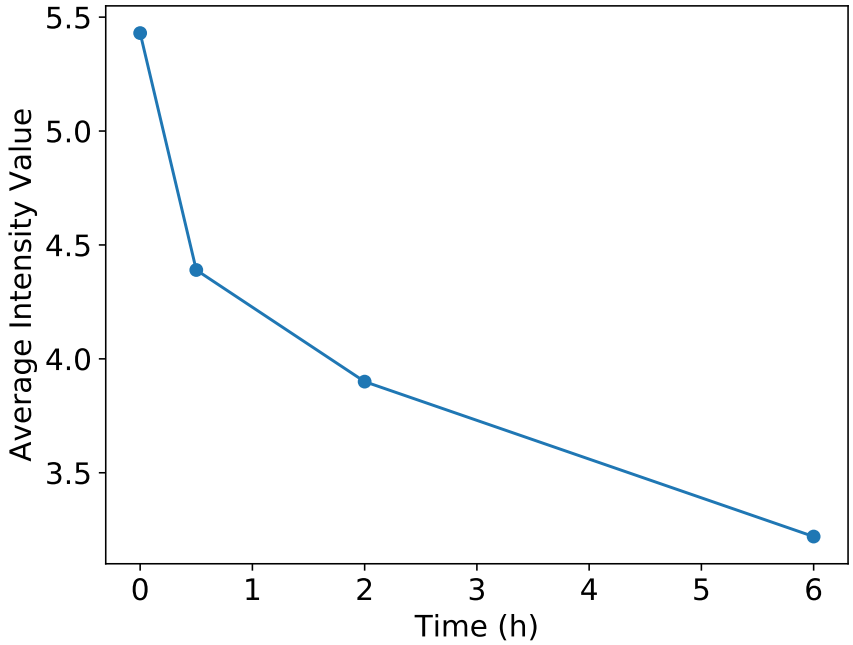

Figure 3.19 – Decrease in the hydration power of vitamin E on human cells with time (T+0, T+30min, T+2h, T+6h).

Virtual Makeup

Finally, we came up the idea of proposing the simple skin parametrization obtained in Section 2.3.5.3 page 66 as a virtual makeup simulation method. The split in space and frequency is well adapted to model a foundation or lipstick application in terms of color alterations and smoothing effect.
In Figure 3.20, the skin parameters from a face with foundation is applied to a face without make-up. We can see how the lightness and colors change (low frequency information) and how the skin texture (high frequency information) is realistically smoothed by the virtual make-up - pores and spots are greatly softened. Several artifacts due to the patch decomposition appear near the eyes and eyebrows. They could be mitigated using a blending operation at the edge of each patch, using for example the Poisson blending method (Pérez et al., 2003).

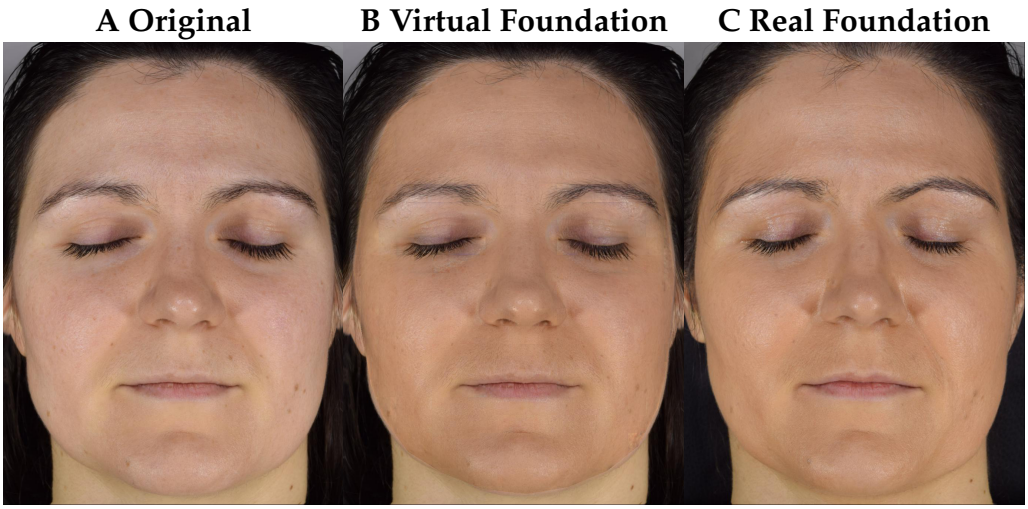

Figure 3.20 – Skin parameters computed from a face with foundation is applied to image A, to give image B. In image C, the real foundation is applied.

Makeup evaluation

Usually, makeups are evaluated using distinct specific metrics. For example, foundations are compared using their coverage effect, brightness effect, homogeneity effect, and many more. We propose to assess makeup in terms of age and health perception, more meaningful for consumers. We use the age estimation system presented in Section 3.3.2 p 86 and the health estimation system presented in Section 2.4.3.1 p 74. Both systems have been trained on faces without any makeup.
We disposed of facial photographs of 32 women taken under controlled illumination with a neutral pose. Each woman was made up by a makeup artist, to obtain 7 independent conditions:

1. without makeup
2. with concealer and foundation
3. the precedent condition where blush was added

4. the precedent condition where lipstick and lip liner were added
5. the condition 3 where eye liner, eye shadow, mascara, and eyebrow pencil were added
6. with full face makeup where the makeup artist where asked to give a natural look
7. with full face makeup where the makeup artist where asked to give a intense look

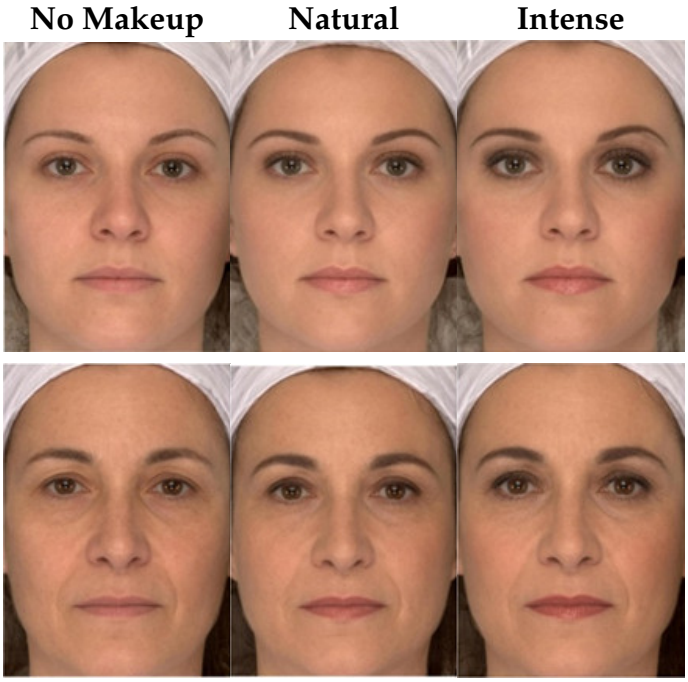

Figure 3.21 – Exemple of 2 faces under 3 conditions: no makeup, with makeup to give a natural look, and with makeup to give an intense look. These faces are average faces of women in their 30's (first line) and in their 50's (second line).

The Figure 3.21 shows 3 conditions for two average faces for 2 age ranges, 30 years and 50 years.
Leveraging the previously described systems, first we estimated the ages and health scores from faces without makeup (condition 1). After that, we estimated the ages and health scores from faces in one of the 6 other conditions. Finally, we computed an average difference between estimates to account the effect of a specific condition.
The influence of makeup conditions on estimated age and health is presented in Table 3.6. The effect of makeup appears to be low on estimated age but stronger on estimated health.
Only condition 3 shows a significant and positive influence on estimated age, while all conditions except condition 3 impact significantly and positively the estimation of health.
In the light of the literature we might explain the result on age estimation by recent findings showing differential effects of makeup on perceived age according to the age of the women (Dayan et al., 2015; Russell et al., 2018). Russell et al.

(2018) found that 40 - and especially 50-year-old-women did appear significantly younger when wearing makeup, 30-year-old women looked no different in age with or without makeup, while 20-year-old women looked older with makeup. Based on these results, we conducted a supplementary analysis by splitting our set of faces in four age groups, women in their 20s, 30s, 40s and 50s, with 8 women by age groups.
The results are shown on Figure 3.22. The graphs show similar pattern of results with our system compared to the results observed with human perception in Russell et al. (2018). We tested whether differential effects of makeup also exist on the estimation of health according to age, these results are shown on Figure 3.23. The pattern of results shows positive effects of makeup on health estimation that vary according to age groups; however the effect was significant only for women in their 30s. The small number of faces by age group might explain the non significant effects of makeup on the other age groups.

Table 3.6 – Influence of different makeup conditions on the estimation of age and health. Non significant values are grayed out.

	Age		**Health**	
	Delta Score	p-value	Delta Score	p-value
2. Foundation	1.02	0.07	1.97	0.03
3. Blush	1.98	0.003	0.85	0.41
4. Lipstick	1.13	0.05	3.72	0.002
5. Eye shadow	-0.40	0.52	2.96	0.003
6. Natural	-0.088	0.86	3.85	0.002
7. Intense	0.13	0.84	3.23	0.002

This new method aims to improve modeling of cosmetics effects on beauty. It should contribute to better understand the visual effects of the products on facial appearance and better predict consumer preferences.

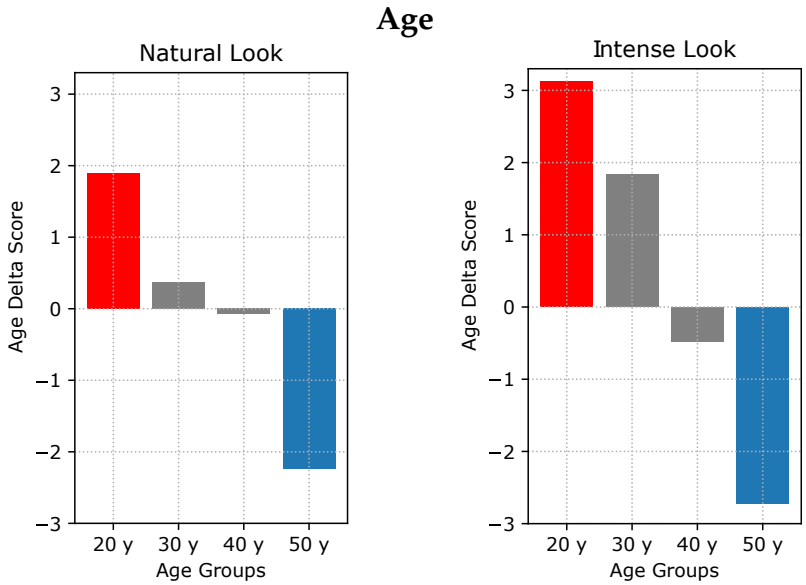

Figure 3.22 – We measure the average difference of scores for makeup giving a natural look (left) and makeup giving an intense look (right) for four age groups. Non significant variations are grayed out.

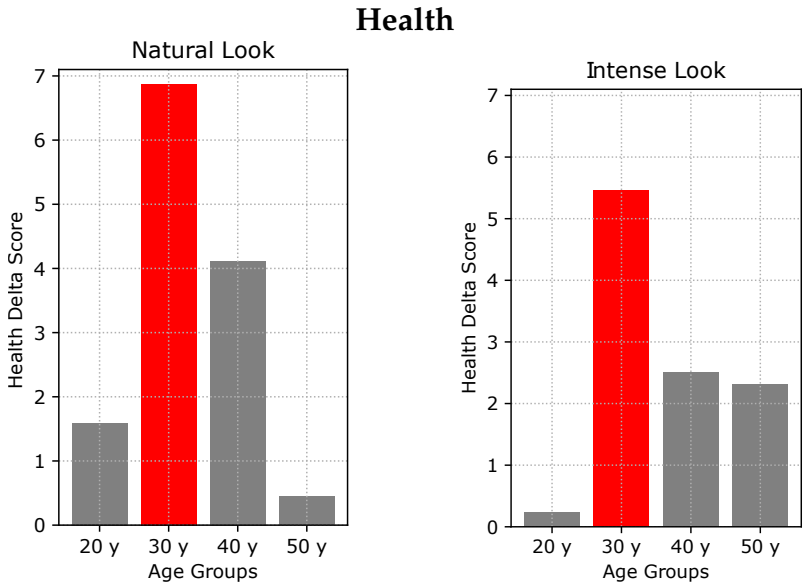

Figure 3.23 – We did an analysis similar to the one in Figure 3.22, but for health estimation. We measure the average difference of health scores for makeup giving a natural look (left) and makeup giving an intense look (right) for four age groups. Non significant variations are grayed out.

Appendix A

Appendices

A.1 Tables: Aging & Wrinkles

These tables are the results of the analysis introduced on Section 2.4.3.3 page 76. Using our representation of wrinkles, we can quantify the effect of aging on wrinkles parameters. For each zone, we computed various correlation between each wrinkle parameter and biological age (Sec. A.1.1). We did the same using perceived age instead of biological age (Sec. A.1.2).

A.1.1 Biological Age

Table A.1 – Correlation between number of wrinkles and biological age for each zone.

Zone	**Number**	
	PC	p-value
Right Corner of the Mouth	39%	**6.7e-03**
Left Corner of the Mouth	59%	**1.4e-05**
Forehead	12%	4.1e-01
Top of the mouth	44%	**1.8e-03**
Right Cheek	34%	**1.8e-02**
Left Cheek	28%	5.5e-02
Chin	29%	**4.6e-02**
Right Crow's-feet	37%	**1.1e-02**
Left Crow's-feet	41%	**4.1e-03**
Frown lines	8%	5.8e-01
Right Nasolabial Fold	40%	**4.9e-03**
Left Nasolabial Fold	40%	**4.9e-03**
Below the Mouth	49%	**4.1e-04**
Below Right Eye	18%	2.2e-01
Below Left Eye	13%	3.7e-01
All Zones	46%	**1.0e-03**

Table A.2 – Correlation between length and curvature of wrinkles and biological age for each zone. The mean and max values were taken for each parameter. The curvature is in fact the absolute value of the curvature parameter.

Zone	**Mean Length**		**Mean Curvature**		**Max Length**		**Max Curvature**	
	PC	p-value	PC	p-value	PC	p-value	PC	p-value
Right Corner of the Mouth	22%	1.4e-01	39%	**6.4e-03**	31%	**3.1e-02**	45%	**1.4e-03**
Left Corner of the Mouth	2%	9.0e-01	7%	6.4e-01	23%	1.2e-01	-4%	7.9e-01
Forehead	5%	7.3e-01	1%	9.7e-01	16%	2.7e-01	20%	1.7e-01
Top of the mouth	25%	9.1e-02	7%	6.2e-01	36%	**1.3e-02**	25%	9.2e-02
Right Cheek	-9%	5.3e-01	-14%	3.5e-01	3%	8.6e-01	25%	8.6e-02
Left Cheek	-0%	9.8e-01	-5%	7.5e-01	18%	2.2e-01	25%	8.5e-02
Chin	10%	5.1e-01	14%	3.6e-01	17%	2.4e-01	32%	**3.1e-02**
Right Crow's-feet	21%	1.6e-01	5%	7.4e-01	28%	5.3e-02	18%	2.3e-01
Left Crow's-feet	4%	8.0e-01	2%	9.0e-01	23%	1.2e-01	5%	7.4e-01
Frown lines	16%	3.0e-01	-4%	7.7e-01	20%	1.7e-01	29%	5.1e-02
Right Nasolabial Fold	38%	**7.9e-03**	-1%	9.6e-01	38%	**7.9e-03**	-1%	9.6e-01
Left Nasolabial Fold	31%	**3.1e-02**	7%	6.3e-01	31%	**3.1e-02**	7%	6.3e-01
Below the Mouth	34%	**1.8e-02**	20%	1.8e-01	40%	**5.4e-03**	40%	**5.7e-03**
Below Right Eye	9%	5.3e-01	20%	1.8e-01	31%	**3.2e-02**	23%	1.2e-01
Below Left Eye	15%	3.3e-01	1%	9.3e-01	21%	1.6e-01	20%	1.7e-01
All Zones	37%	**1.1e-02**	0%	9.8e-01	48%	**7.1e-04**	41%	**4.0e-03**

Table A.3 – Correlation between intensity and width of wrinkles and biological age for each zone. The mean and max values were taken for each parameter.

Zone	**Mean Intensity**		**Mean Width**		**Max Intensity**		**Max Width**	
	PC	p-value	PC	p-value	PC	p-value	PC	p-value
Right Corner of the Mouth	39%	**7.1e-03**	45%	**1.4e-03**	54%	**9.0e-05**	46%	**1.0e-03**
Left Corner of the Mouth	49%	**5.0e-04**	46%	**1.3e-03**	62%	**3.2e-06**	49%	**4.4e-04**
Forehead	16%	3.0e-01	-20%	1.9e-01	28%	5.8e-02	-10%	5.0e-01
Top of the mouth	37%	**9.7e-03**	34%	**1.9e-02**	23%	1.1e-01	32%	**2.7e-02**
Right Cheek	10%	5.0e-01	-1%	9.6e-01	27%	6.9e-02	4%	7.9e-01
Left Cheek	9%	5.6e-01	9%	5.5e-01	20%	1.8e-01	14%	3.3e-01
Chin	38%	**9.0e-03**	24%	1.1e-01	43%	**2.6e-03**	27%	6.6e-02
Right Crow's-feet	24%	1.0e-01	29%	**4.8e-02**	28%	6.0e-02	33%	**2.2e-02**
Left Crow's-feet	29%	**4.7e-02**	19%	2.1e-01	42%	**3.4e-03**	25%	9.1e-02
Frown lines	33%	**2.2e-02**	16%	2.7e-01	40%	**5.7e-03**	13%	3.7e-01
Right Nasolabial Fold	57%	**3.0e-05**	43%	**2.5e-03**	57%	**3.0e-05**	43%	**2.5e-03**
Left Nasolabial Fold	66%	**4.1e-07**	44%	**2.2e-03**	66%	**4.1e-07**	44%	**2.2e-03**
Below the Mouth	44%	**2.0e-03**	47%	**8.8e-04**	42%	**3.1e-03**	48%	**7.0e-04**
Below Right Eye	25%	8.7e-02	-17%	2.6e-01	34%	**2.1e-02**	-1%	9.7e-01
Below Left Eye	22%	1.3e-01	-24%	1.1e-01	36%	**1.3e-02**	9%	5.3e-01
All Zones	69%	**1.1e-07**	50%	**3.8e-04**	69%	**9.1e-08**	52%	**1.5e-04**

A.1.2 Perceived Age

Table A.4 – Correlation between number of wrinkles and perceived age for each zone.

Zone	**Number**	
	PC	p-value
Right Corner of the Mouth	60%	**7.6e-06**
Left Corner of the Mouth	62%	**3.3e-06**
Forehead	9%	5.5e-01
Top of the mouth	56%	**5.1e-05**
Right Cheek	38%	**8.5e-03**
Left Cheek	40%	**5.6e-03**
Chin	31%	**3.3e-02**
Right Crow's-feet	35%	**1.7e-02**
Left Crow's-feet	41%	**4.5e-03**
Frown lines	8%	5.9e-01
Right Nasolabial Fold	38%	**7.9e-03**
Left Nasolabial Fold	38%	**7.9e-03**
Below the Mouth	51%	**2.5e-04**
Below Right Eye	35%	**1.6e-02**
Below Left Eye	18%	2.2e-01
All Zones	55%	**5.5e-05**

Table A.5 – Correlation between length and curvature of wrinkles and perceived age for each zone. The mean and max values were taken for each parameter. The curvature is in fact the absolute value of the curvature parameter.

Zone	**Mean Length**		**Mean Curvature**		**Max Length**		**Max Curvature**	
	PC	p-value	PC	p-value	PC	p-value	PC	p-value
Right Corner of the Mouth	22%	1.3e-01	29%	**4.9e-02**	39%	**6.2e-03**	41%	**4.3e-03**
Left Corner of the Mouth	15%	3.1e-01	9%	5.3e-01	38%	**8.5e-03**	7%	6.3e-01
Forehead	-2%	8.7e-01	7%	6.5e-01	9%	5.3e-01	24%	1.1e-01
Top of the mouth	42%	**3.0e-03**	18%	2.2e-01	50%	**3.4e-04**	44%	**2.2e-03**
Right Cheek	11%	4.6e-01	-1%	9.6e-01	27%	6.4e-02	30%	**4.1e-02**
Left Cheek	20%	1.7e-01	3%	8.6e-01	39%	**6.3e-03**	32%	**3.0e-02**
Chin	-1%	9.4e-01	7%	6.3e-01	7%	6.5e-01	23%	1.2e-01
Right Crow's-feet	21%	1.5e-01	4%	8.0e-01	26%	8.0e-02	11%	4.8e-01
Left Crow's-feet	11%	4.6e-01	-4%	7.9e-01	25%	8.9e-02	-1%	9.5e-01
Frown lines	24%	1.0e-01	1%	9.2e-01	34%	**2.0e-02**	19%	2.1e-01
Right Nasolabial Fold	41%	**4.5e-03**	-18%	2.1e-01	41%	**4.5e-03**	-18%	2.1e-01
Left Nasolabial Fold	31%	**3.5e-02**	-6%	7.1e-01	31%	**3.5e-02**	-6%	7.1e-01
Below the Mouth	46%	**1.2e-03**	31%	**3.7e-02**	47%	**9.5e-04**	42%	**3.6e-03**
Below Right Eye	11%	4.7e-01	41%	**4.3e-03**	42%	**3.5e-03**	29%	**4.7e-02**
Below Left Eye	20%	1.9e-01	12%	4.3e-01	26%	7.2e-02	25%	9.3e-02
All Zones	43%	**2.6e-03**	-17%	2.6e-01	57%	**2.5e-05**	46%	**1.3e-03**

Table A.6 – Correlation between intensity and width of wrinkles and perceived age for each zone. The mean and max values were taken for each parameter.

Zone	Mean Intensity		Mean Width		Max Intensity		Max Width	
	PC	p-value	PC	p-value	PC	p-value	PC	p-value
Right Corner of the Mouth	40%	**5.0e-03**	50%	**3.0e-04**	56%	**4.6e-05**	53%	**1.1e-04**
Left Corner of the Mouth	60%	**9.2e-06**	50%	**3.1e-04**	70%	**3.4e-08**	57%	**3.3e-05**
Forehead	14%	3.6e-01	-12%	4.0e-01	19%	2.0e-01	-14%	3.6e-01
Top of the mouth	43%	**2.3e-03**	51%	**2.2e-04**	29%	**4.5e-02**	52%	**1.6e-04**
Right Cheek	29%	**4.9e-02**	16%	2.9e-01	42%	**3.0e-03**	19%	1.9e-01
Left Cheek	19%	1.9e-01	27%	6.6e-02	31%	**3.1e-02**	35%	**1.6e-02**
Chin	31%	**3.4e-02**	19%	2.1e-01	38%	**8.6e-03**	23%	1.2e-01
Right Crow's-feet	31%	**3.5e-02**	25%	9.5e-02	34%	**2.1e-02**	26%	7.9e-02
Left Crow's-feet	26%	7.8e-02	1%	9.3e-01	47%	**9.1e-04**	17%	2.6e-01
Frown lines	42%	**3.2e-03**	31%	**3.6e-02**	42%	**3.1e-03**	22%	1.3e-01
Right Nasolabial Fold	69%	**6.4e-08**	40%	**5.4e-03**	69%	**6.4e-08**	40%	**5.4e-03**
Left Nasolabial Fold	73%	**7.2e-09**	41%	**3.8e-03**	73%	**7.2e-09**	41%	**3.8e-03**
Below the Mouth	57%	**2.4e-05**	60%	**1.0e-05**	51%	**2.8e-04**	59%	**1.2e-05**
Below Right Eye	48%	**6.7e-04**	-4%	7.9e-01	51%	**2.8e-04**	14%	3.6e-01
Below Left Eye	37%	**9.8e-03**	-7%	6.4e-01	46%	**1.2e-03**	14%	3.6e-01
All Zones	80%	**2.1e-11**	61%	**4.7e-06**	80%	**2.4e-11**	63%	**2.2e-06**

A.2 La*b* color space

The goal of La*b* color space is to better represent colors in terms of human visual perception. Uniform changes in the L*a*b* color space aim to correspond to uniform changes in perceived color by humans. The three components are:

— **L** is the lightness value, ranging from 0 (black) to 100 (white)

— **a*** is the green-red component, usually ranging from -128 (green) to $+127$ (red)

— **b*** is the blue-yellow component, usually ranging from -128 (blue) to $+127$ (yellow)

To convert from RGB to La*b*, first RGB is converted to XYZ, then XYZ is converted to La*b*.

A.2.1 RGB to XYZ

RGB is first converted to XYZ (XYZ, 2018):

$$\begin{pmatrix} 2{,}768\,9 & 1{,}751\,7 & 1{,}130\,2 \\ 1{,}000\,0 & 4{,}590\,7 & 0{,}060\,100 \\ 0{,}000\,0 & 0{,}056\,508 & 5{,}594\,3 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{A.1}$$

A.2.2 XYZ to La*b*

Then, from XYZ to La*b* (LAB, 2018):

$$L^\star = 116\, f\left(\frac{Y}{Y_\mathrm{n}}\right) - 16 \tag{A.2}$$

$$a^\star = 500\left(f\left(\frac{X}{X_\mathrm{n}}\right) - f\left(\frac{Y}{Y_\mathrm{n}}\right)\right) \tag{A.3}$$

$$b^\star = 200\left(f\left(\frac{Y}{Y_\mathrm{n}}\right) - f\left(\frac{Z}{Z_\mathrm{n}}\right)\right) \tag{A.4}$$

where:

$$f(t) = \begin{cases} \sqrt[3]{t} & \text{if } t > \delta^3 \\ \frac{t}{3\delta^2} + \frac{4}{29} & \text{otherwise} \end{cases} \tag{A.5}$$

$$\delta = \frac{6}{29} \tag{A.6}$$

Here, (X_n, Y_n, Z_n) represents the standard illuminant.

A.3 Makeup Mirror

Our age estimation algorithm was implemented on a smart mirror to give instant age estimates, and observe how they vary according to different kinds of makeup.

Figure A.1 – The color and the brightness of the lights are controlled by the mirror. The red number is the instant estimate and the two green numbers are the mean and standard deviation over the 40 last estimates.

Bibliography

Parham Aarabi. How brands are using facial recognition to transform marketing, April 2013. URL https://venturebeat.com/2013/04/13/marketing-facial-recognition/.

Grigory Antipov, Moez Baccouche, Sid-Ahmed Berrani, and Jean-Luc Dugelay. Apparent age estimation from face images combining general and children-specialized deep learning models. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 96–104, 2016. URL http://www.cv-foundation.org/openaccess/content_cvpr_2016_workshops/w18/html/Antipov_Apparent_Age_Estimation_CVPR_2016_paper.html. bibtex: antipov_apparent_2016.

Grigory Antipov, Moez Baccouche, and Jean-Luc Dugelay. Face Aging With Conditional Generative Adversarial Networks. *arXiv:1702.01983 [cs]*, February 2017. URL http://arxiv.org/abs/1702.01983. arXiv: 1702.01983 bibtex: antipov_face_2017.

J. Axelsson, T. Sundelin, C. Axelsson, J. Lasselin, and M. Lekander. Identification of acutely sick people and facial cues of sickness. *Brain, Behavior, and Immunity*, 66:e38, November 2017. ISSN 0889-1591. doi: 10.1016/j.bbi.2017.07.138. URL http://www.sciencedirect.com/science/article/pii/S0889159117303653.

John Axelsson, Tina Sundelin, Michael Ingre, Eus J. W. Van Someren, Andreas Olsson, and Mats Lekander. Beauty sleep: experimental study on the perceived health and attractiveness of sleep deprived people. *BMJ*, 341:c6614, December 2010. ISSN 0959-8138, 1468-5833. doi: 10.1136/bmj.c6614. URL https://www.bmj.com/content/341/bmj.c6614.

Philippe Bellity and Jonathan Bellity. Facial Rejuvenation Enhancing Cheek Lift , Facial Rejuvenation Enhancing Cheek Lift. *Archives of Plastic Surgery, Archives of Plastic Surgery*, 44(6):559–563, October 2017. ISSN 2234-6163, 2234-6171. doi: 10.5999/aps.2016.01858. URL http://www.e-aps.org/journal/view.php?doi=10.5999/aps.2016.01858.

Laurence Boissieux, Gergo Kiss, Nadia Magnenat Thalmann, and Prem Kalra. Simulation of skin aging and wrinkles with cosmetics insight. In *Computer Animation and Simulation 2000*, pages 15–27. Springer, 2000. URL http://link.springer.com/chapter/10.1007/978-3-7091-6344-3_2. bibtex: boissieux_simulation_2000.

Vicki Bruce and Andy Young. In the Eye of the Beholder: The Science of Face Perception. January 1998.

Ali Maina Bukar, Hassan Ugail, and Nosheen Hussain. On Facial Age Progression Based on Modified Active Appearance Models with Face Texture. In Plamen Angelov, Alexander Gegov, Chrisina Jayne, and Qiang Shen, editors, *Advances in Computational Intelligence Systems*, volume 513, pages 465–479. Springer International Publishing, Cham, 2017. ISBN 978-3-319-46561-6 978-3-319-46562-3. doi: 10.1007/978-3-319-46562-3_30. URL http://link.springer.com/10.1007/978-3-319-46562-3_30. bibtex: angelov_facial_2017.

D. M. Burt and D. I. Perrett. Perception of Age in Adult Caucasian Male Faces: Computer Graphic Manipulation of Shape and Colour Information. *Proceedings of the Royal Society B: Biological Sciences*, 259(1355):137–143, February 1995. ISSN 0962-8452, 1471-2954. doi: 10.1098/rspb.1995.0021. URL http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.1995.0021. bibtex: burt_perception_1995.

F. Chen, X. Xiao, and D. Zhang. Data-Driven Facial Beauty Analysis: Prediction, Retrieval and Manipulation. *IEEE Transactions on Affective Computing*, PP(99):1–1, 2017. ISSN 1949-3045. doi: 10.1109/TAFFC.2016.2599534. bibtex: chen_data-driven_2017.

Fangmei Chen and David Zhang. Combining a causal effect criterion for evaluation of facial attractiveness models. *Neurocomputing*, 177:98–109, February 2016. ISSN 0925-2312. doi: 10.1016/j.neucom.2015.11.010. URL http://www.sciencedirect.com/science/article/pii/S0925231215016641. bibtex: chen_combining_2016.

Kaare Christensen, Maria Iachina, Helle Rexbye, Cecilia Tomassini, Henrik Frederiksen, Matt McGue, and James W. Vaupel. "Looking Old for Your Age": Genetics and Mortality. *Epidemiology*, 15(2):251, March 2004. ISSN 1044-3983. doi: 10.1097/01.ede.0000112211.11416.a6. URL https://journals.lww.com/epidem/fulltext/2004/03000/_looking_old_for_your_age___genetics_and_mortality.21.aspx.

Kaare Christensen, Mikael Thinggaard, Matt McGue, Helle Rexbye, Jacob v B. Hjelmborg, Abraham Aviv, David Gunn, Frans van der Ouderaa, and James W. Vaupel. Perceived age as clinically useful biomarker of ageing: cohort study. *BMJ*, 339:b5262, December 2009. ISSN 0959-8138, 1468-5833. doi: 10.1136/bmj.b5262. URL https://www.bmj.com/content/339/bmj.b5262.

V. Coetzee, D. I. Perrett, and I. D. Stephen. *Facial adiposity: A cue to health? Perception, 38(11)*. 1700–1711, 2009. URL http://doi.org/10.1068/p6423. bibtex: coetzee_facial_2009.

T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. In Hans Burkhardt and Bernd Neumann, editors, *Computer Vision — ECCV'98*,

Lecture Notes in Computer Science, pages 484–498. Springer Berlin Heidelberg, June 1998. ISBN 978-3-540-64613-6 978-3-540-69235-5. doi: 10.1007/BFb0054760. URL http://link.springer.com/chapter/10.1007/BFb0054760. bibtex: cootes_active_1998.

Steven H. Dayan, Katherine Cho, Mary Siracusa, and Selika Gutierrez-Borst. Quantifying the impact cosmetic make-up has on age perception and the first impression projected. *Journal of drugs in dermatology: JDD*, 14(4):366–374, April 2015. ISSN 1545-9616.

L. M. Donofrio. Fat distribution: a morphologic study of the aging face. *Dermatologic Surgery: Official Publication for American Society for Dermatologic Surgery [et Al.]*, 26(12):1107–1112, December 2000. ISSN 1076-0512. bibtex: donofrio_fat_2000.

M. G. Efran. The effect of physical appearance on the judgment of guilt, interpersonal attraction, and severity of recommended punishment in a simulated jury task. *Journal of Research in Personality*, 8(1), 1974. URL http://doi.org/10.1016/0092-6566(74). bibtex: efran_effect_1974.

Tomonobu Ezure, Eiichiro Yagi, Naomi Kunizawa, Tetsuji Hirao, and Satoshi Amano. Comparison of sagging at the cheek and lower eyelid between male and female faces. *Skin research and technology: official journal of International Society for Bioengineering and the Skin (ISBS) [and] International Society for Digital Imaging of Skin (ISDIS) [and] International Society for Skin Imaging (ISSI)*, 17(4):510–515, November 2011. ISSN 1600-0846. doi: 10.1111/j.1600-0846.2011.00526.x.

Y. Y. Fan, S. Liu, B. Li, Z. Guo, A. Samal, J. Wan, and S. Z. Li. Label Distribution-Based Facial Attractiveness Computation by Deep Residual Learning. *IEEE Transactions on Multimedia*, PP(99):1–1, 2017. ISSN 1520-9210. doi: 10.1109/TMM.2017.2780762. bibtex: fan_label_2017.

Leslie G. Farkas, Otto G. Eiben, Stefan Sivkov, Bryan Tompson, Marko J. Katic, and Christopher R. Forrest. Anthropometric measurements of the facial framework in adulthood: age-related changes in eight age categories in 600 healthy white North Americans of European ancestry from 16 to 90 years of age. *The Journal of Craniofacial Surgery*, 15(2):288–298, March 2004. ISSN 1049-2275. bibtex: farkas_anthropometric_2004.

A. Feingold. Good-looking people are not what we think. *Psychological Bulletin*, 111(2), 1992. URL http://doi.org/10.1037/0033-2909.111.2.304. bibtex: feingold_good-looking_1992.

B Fink and Pj Matts. The effects of skin colour distribution and topography cues on the perception of female facial age and health. *Journal of the European Academy of Dermatology and Venereology*, 22(4):493–498, April 2008. ISSN 0926-9959, 1468-3083. doi: 10.1111/j.1468-3083.2007.02512.x. URL http://doi.wiley.com/10.1111/j.1468-3083.2007.02512.x. bibtex: fink_effects_2008.

B. Fink, K. Grammer, and P. Matts. Visible skin color distribution plays a role in the perception of age, attractiveness, and health in female faces. *Evolution and Human Behavior*, 27(6), 2006. URL http://doi.org/10.1016/j.evolhumbehav.2006.08.007. bibtex: fink_visible_2006.

B. Fink, P. Matts, S. Röder, R. Johnson, and M. Burquest. Differences in visual perception of age and attractiveness of female facial and body skin: Perception of face and body skin. *International Journal of Cosmetic Science*, 33(2), 2011a. URL http://doi.org/10.1111/j.1468-2494.2010.00594.x. bibtex: fink_differences_2011.

B. Fink, P.J. Matts, D. D'Emiliano, L. Bunse, B. Weege, and S. Röder. Colour homogeneity and visual perception of age, health and attractiveness of male facial skin: Perception of male skin colour. *Journal of the European Academy of Dermatology and Venereology*, pages no–no, November 2011b. ISSN 09269959. doi: 10.1111/j.1468-3083.2011.04316.x. URL http://doi.wiley.com/10.1111/j.1468-3083.2011.04316.x. bibtex: fink_colour_2011.

B. Fink, P. Matts, D. D'Emiliano, L. Bunse, B. Weege, and S. Röder. Colour homogeneity and visual perception of age, health and attractiveness of male facial skin. *Journal of the European Academy of Dermatology and Venereology*, 26 (12), 2012. URL http://doi.org/10.1111/j.1468-3083.2011.04316.x. bibtex: fink_colour_2012.

Y. Fu, G. Guo, and T. S. Huang. Age Synthesis and Estimation via Faces: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(11):1955–1976, November 2010. ISSN 0162-8828. doi: 10.1109/TPAMI.2010.36. bibtex: fu_age_2010.

Maulin R. Gandhi. *A method for automatic synthesis of aged human facial images*. PhD thesis, McGill University, Montreal, Canada, 2004. URL http://www.cim.mcgill.ca/~levine/thesis-maulin.gandhi-lowres.pdf. bibtex: gandhi_method_2004.

X. Geng, C. Yin, and Z. Zhou. Facial Age Estimation by Learning from Label Distributions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35 (10):2401–2412, October 2013. ISSN 0162-8828. doi: 10.1109/TPAMI.2013.51.

Xin Geng, Zhi-Hua Zhou, and Kate Smith-Miles. Automatic age estimation based on facial aging patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 29(12):2234–2240, 2007.

Patricia A. George and Graham J. Hole. Factors Influencing the Accuracy of Age Estimates of Unfamiliar Faces. *Perception*, 24(9):1059–1073, September 1995. ISSN 0301-0066. doi: 10.1068/p241059. URL https://doi.org/10.1068/p241059.

Pierre Geurts, Damien Ernst, and Louis Wehenkel. Extremely randomized trees. *Machine Learning*, 63(1):3–42, April 2006. ISSN 1573-0565. doi: 10.1007/s10994-006-6226-1. URL https://doi.org/10.1007/s10994-006-6226-1.

Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative Adversarial Networks. *arXiv:1406.2661 [cs, stat]*, June 2014. URL http://arxiv.org/abs/1406.2661. arXiv: 1406.2661.

C. Guinot, J. Latreille, E. Mauger, L. Ambroisine, S. Gardinier, H. Zahouani, S. Guéhenneux, and E. Tschachler. Reference ranges of skin micro-relief according to age in French Caucasian and Japanese women. *Skin Research and Technology*, 12(4):268–278, November 2006. ISSN 1600-0846. doi: 10.1111/j.0909-752X.2006.00164.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0909-752X.2006.00164.x.

David A. Gunn, Peter G. Murray, Cyrena C. Tomlin, Helle Rexbye, Kaare Christensen, and Andrew E. Mayes. Perceived age as a biomarker of ageing: a clinical methodology. *Biogerontology*, 9(5):357, April 2008. ISSN 1573-6768. doi: 10.1007/s10522-008-9141-y. URL https://doi.org/10.1007/s10522-008-9141-y.

David A. Gunn, Helle Rexbye, Christopher E. M. Griffiths, Peter G. Murray, Amelia Fereday, Sharon D. Catt, Cyrena C. Tomlin, Barbara H. Strongitharm, Dave I. Perrett, Michael Catt, Andrew E. Mayes, Andrew G. Messenger, Martin R. Green, Frans van der Ouderaa, James W. Vaupel, and Kaare Christensen. Why Some Women Look Young for Their Age. *PLOS ONE*, 4(12):e8021, December 2009. ISSN 1932-6203. doi: 10.1371/journal.pone.0008021. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0008021.

Guodong Guo, Guowang Mu, Yun Fu, and Thomas S. Huang. Human age estimation using bio-inspired features. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 112–119. IEEE, 2009. bibtex: guo_human_2009.

Bahman Guyuron, David J. Rowe, Adam Bryce Weinfeld, Yashar Eshraghi, Amir Fathi, and Seree Iamphongsai. Factors contributing to the facial aging of identical twins. *Plastic and Reconstructive Surgery*, 123(4):1321–1331, April 2009. ISSN 1529-4242. doi: 10.1097/PRS.0b013e31819c4d42. bibtex: guyuron_factors_2009.

A. J. Henderson, I. J. Holzleitner, S. N. Talamas, and D. I. Perrett. Perception of health from facial cues. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1693):20150380., 2016. URL http://doi.org/10.1098/rstb.2015.0380. bibtex: henderson_perception_2016.

Jinli Suo, Xilin Chen, Shiguang Shan, and Wen Gao. Learning long term face aging patterns from partially dense aging databases. pages 622–629. IEEE, September 2009. ISBN 978-1-4244-4420-5. doi: 10.1109/ICCV.2009.5459181. URL http://ieeexplore.ieee.org/document/5459181/. bibtex: jinli_suo_learning_2009.

Jinli Suo, Song-Chun Zhu, Shiguang Shan, and Xilin Chen. A Compositional and Dynamic Model for Face Aging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3):385–401, March 2010. ISSN 0162-8828. doi: 10.1109/

TPAMI.2009.39. URL http://ieeexplore.ieee.org/document/4782970/. bibtex: jinli_suo_compositional_2010.

Alex L. Jones, Carlota Batres, Aurélie Porcheron, Jennifer R. Sweda, Frédérique Morizot, and Richard Russell. Positive facial affect looks healthy. *Visual Cognition*, 26(1):1–12, January 2018. ISSN 1350-6285. doi: 10.1080/13506285.2017.1369202. URL https://doi.org/10.1080/13506285.2017.1369202.

Sebastian Korb, Stéphane With, Paula Niedenthal, Susanne Kaiser, and Didier Grandjean. The Perception and Mimicry of Facial Movements Predict Judgments of Smile Authenticity. *PLOS ONE*, 9(6):e99194, June 2014. ISSN 1932-6203. doi: 10.1371/journal.pone.0099194. URL http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0099194.

Young H Kwon and Niels da Vitoria Lobo. Age Classification from Facial Images. *Computer Vision and Image Understanding*, 74(1):1–21, April 1999. ISSN 1077-3142. doi: 10.1006/cviu.1997.0549. URL http://www.sciencedirect.com/science/article/pii/S107731429790549X. bibtex: kwon_age_1999.

Young Ho Kwon and N. da Vitoria Lobo. Age classification from facial images. In *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 762–767, June 1994. doi: 10.1109/CVPR.1994.323894. bibtex: kwon_age_1994.

CIE LAB. CIELAB color space, September 2018. URL https://en.wikipedia.org/w/index.php?title=CIELAB_color_space&oldid=859056425. Page Version ID: 859056425.

A. Lanitis, C.J. Taylor, and T.F. Cootes. Modeling the process of ageing in face images. pages 131–136 vol.1. IEEE, 1999. ISBN 978-0-7695-0164-2. doi: 10.1109/ICCV.1999.791208. URL http://ieeexplore.ieee.org/document/791208/. bibtex: lanitis_modeling_1999.

Andreas Lanitis, Christopher J. Taylor, and Timothy F. Cootes. Toward automatic simulation of aging effects on face images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4):442–455, 2002. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=993553. bibtex: lanitis_toward_2002.

Andreas Lanitis, Chrisina Draganova, and Chris Christodoulou. Comparing Different Classifiers for Automatic Age Estimation. *IEEE transactions on systems, man, and cybernetics. Part B, Cybernetics : a publication of the IEEE Systems, Man, and Cybernetics Society*, 34:621–8, March 2004. doi: 10.1109/TSMCB.2003.817091. bibtex: lanitis_comparing_2004.

Julie Latreille, Khaled Ezzedine, Anissa Elfakir, Laurence Ambroisine, Sophie Gardinier, Pilar Galan, Serge Hercberg, Florian Gruber, Jonathan Rees, Erwin Tschachler, and Christiane Guinot. MC1r Gene Polymorphism Affects Skin Color and Phenotypic Features Related to Sun Sensitivity in a Population of French Adult Women. *Photochemistry and Photobiology*, 85(6):1451–1458, November 2009. ISSN 1751-1097. doi: 10.1111/j.1751-1097.2009.00594.

x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1751-1097.2009.00594.x.

Julie Latreille, Emmanuelle Kesse-Guyot, Denis Malvy, Valentina Andreeva, Pilar Galan, Erwin Tschachler, Serge Hercberg, Christiane Guinot, and Khaled Ezzedine. Dietary Monounsaturated Fatty Acids Intake and Risk of Skin Photoaging. *PLoS ONE*, 7(9), September 2012. ISSN 1932-6203. doi: 10.1371/journal.pone.0044490. URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3435270/. bibtex: latreille_dietary_2012.

Jing Li and Bao-Liang Lu. An adaptive image Euclidean distance. *Pattern Recognition*, 42(3):349–357, March 2009. ISSN 0031-3203. doi: 10.1016/j.patcog.2008.07.017. URL http://www.sciencedirect.com/science/article/pii/S0031320308003130.

Shu Liu, Yang-Yu Fan, Ashok Samal, and Zhe Guo. Advances in computational facial attractiveness methods. *Multimedia Tools and Applications*, 75(23):16633–16663, December 2016. ISSN 1380-7501, 1573-7721. doi: 10.1007/s11042-016-3830-3. URL https://link.springer.com/article/10.1007/s11042-016-3830-3. bibtex: liu_advances_2016.

Si Liu, Yao Sun, Defa Zhu, Renda Bao, Wei Wang, Xiangbo Shu, and Shuicheng Yan. Face Aging with Contextual Generative Adversarial Nets. In *Proceedings of the 2017 ACM on Multimedia Conference*, MM '17, pages 82–90, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4906-2. doi: 10.1145/3123266.3123431. URL http://doi.acm.org/10.1145/3123266.3123431. bibtex: liu_face_2017.

J-M. Denis Malvy, Christiane Guinot, Paul Preziosi, Loïc Vaillant, Michel Tenenhaus, Pilar Galan, Serge Hercberg, and Erwin Tschachler. Epidemiologic determinants of skin photoaging: Baseline data of the SU.VI.MAX. cohort. *Journal of the American Academy of Dermatology*, 42(1, Part 1):47–55, January 2000. ISSN 0190-9622. doi: 10.1016/S0190-9622(00)90008-2. URL http://www.sciencedirect.com/science/article/pii/S0190962200900082.

Leonard S. Mark, John B. Pittenger, Helen Hines, Claudia Carello, Robert E. Shaw, and James T. Todd. Wrinkling and head shape as coordinated sources of age-level information. *Perception & Psychophysics*, 27(2):117–124, March 1980. ISSN 0031-5117, 1532-5962. doi: 10.3758/BF03204298. URL https://link.springer.com/article/10.3758/BF03204298. bibtex: mark_wrinkling_1980.

C. M. Marlowe, S. L. Schneider, and C. E. Nelson. *Gender and attractiveness biases in hiring decisions: Are more experienced managers less biased? Journal of Applied Psychology, 81(1)*. 11–21, 1996. URL http://doi.org/10.1037/0021-9010.81.1.11. bibtex: marlowe_gender_1996.

Victor Martin, Renaud Séguier, Aurélie Porcheron, and Frédérique Morizot. Face aging simulation with a new wrinkle oriented active appearance model. *Multimedia Tools and Applications*, pages 1–19, July 2018a. ISSN 1380-7501, 1573-7721. doi: 10.1007/s11042-018-6311-z. URL https://link.springer.com/article/10.1007/s11042-018-6311-z.

Victor Martin, Renaud Séguier, Aurélie Porcheron, and Frédérique Morizot. Towards Continuous Health Diagnosis from Faces with Deep Learning. In Islem Rekik, Gozde Unal, Ehsan Adeli, and Sang Hyun Park, editors, *International Workshop on PRedictive Intelligence In MEdicine @ MICCAI 2018*, Lecture Notes in Computer Science, pages 120–128. Springer International Publishing, 2018b. ISBN 978-3-030-00320-3. doi: 10.1007/978-3-030-00320-3_15. URL https://link.springer.com/chapter/10.1007/978-3-030-00320-3_15.

Victor Martin, Renaud Séguier, Aurélie Porcheron, and Frédérique Morizot. Understanding the visual effects of cosmetic products on beauty via Deep Learning. In *IFSCC*, Munich, Germany, 2018c.

Andrew E. Mayes, Peter G. Murray, David A. Gunn, Cyrena C. Tomlin, Sharon D. Catt, Yi B. Wen, Li P. Zhou, Hong Q. Wang, Michael Catt, and Stewart P. Granger. Environmental and Lifestyle Factors Associated with Perceived Facial Age in Chinese Women. *PLOS ONE*, 5(12):e15270, December 2010. ISSN 1932-6203. doi: 10.1371/journal.pone.0015270. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0015270.

Bryan Mendelson and Chin-Ho Wong. Changes in the Facial Skeleton With Aging: Implications and Clinical Applications in Facial Rejuvenation. *Aesthetic Plastic Surgery*, 36(4):753–760, August 2012. ISSN 0364-216X. doi: 10.1007/s00266-012-9904-3. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3404279/.

Umar Mohammed, Simon J. D. Prince, and Jan Kautz. Visio-lization: Generating Novel Facial Images. In *ACM SIGGRAPH 2009 Papers*, SIGGRAPH '09, pages 57:1–57:8, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-726-4. doi: 10.1145/1576246.1531363. URL http://doi.acm.org/10.1145/1576246.1531363. bibtex: mohammed_visio-lization_2009.

Yu-Ichi Ohta, Takeo Kanade, and Toshiyuki Sakai. Color information for region segmentation. *Computer Graphics and Image Processing*, 13(3):222–241, July 1980. ISSN 0146-664X. doi: 10.1016/0146-664X(80)90047-7. URL http://www.sciencedirect.com/science/article/pii/0146664X80900477.

G. V. Ostir, K. J. Ottenbacher, and K. S. Markides. Onset of frailty in older adults and the protective role of positive affect. *Psychology and Aging*, 19(3), 2004. URL http://doi.org/10.1037/0882-7974.19.3.402. bibtex: ostir_onset_2004.

Patrick Pérez, Michel Gangnet, and Andrew Blake. Poisson Image Editing. In *ACM SIGGRAPH 2003 Papers*, SIGGRAPH '03, pages 313–318, New York, NY, USA, 2003. ACM. ISBN 978-1-58113-709-5. doi: 10.1145/1201775.882269. URL http://doi.acm.org/10.1145/1201775.882269.

Rizhen Qin, Wei Gao, Huarong Xu, and Zhanyi Hu. Modern Physiognomy: An Investigation on Predicting Personality Traits and Intelligence from the Human Face. *arXiv:1604.07499 [cs]*, April 2016. URL http://arxiv.org/abs/1604.07499. arXiv: 1604.07499 bibtex: qin_modern_2016.

Narayanan Ramanathan and Rama Chellappa. Modeling age progression in young faces. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, volume 1, pages 387–394. IEEE, 2006. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1640784. bibtex: ramanathan_modeling_2006.

Narayanan Ramanathan and Rama Chellappa. Modeling shape and textural variations in aging faces. pages 1–8. IEEE, September 2008. ISBN 978-1-4244-2153-4. doi: 10.1109/AFGR.2008.4813337. URL http://ieeexplore.ieee.org/document/4813337/.

Narayanan Ramanathan, Rama Chellappa, and Soma Biswas. Computational methods for modeling facial aging: A survey. *Journal of Visual Languages & Computing*, 20(3):131–144, June 2009. ISSN 1045-926X. doi: 10.1016/j.jvlc.2009.01.011. URL http://www.sciencedirect.com/science/article/pii/S1045926X09000032.

Daniel E. Re, Ross D. Whitehead, Dengke Xiao, and David I. Perrett. Oxygenated-Blood Colour Change Thresholds for Perceived Facial Redness, Health, and Attractiveness. *PLoS ONE*, 6(3), March 2011. ISSN 1932-6203. doi: 10.1371/journal.pone.0017859. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3063159/.

Daniel E. Re, David W. Hunter, Vinet Coetzee, Bernard P. Tiddeman, Dengke Xiao, Lisa M. DeBruine, Benedict C. Jones, and David I. Perrett. Looking Like a Leader–Facial Shape Predicts Perceived Height and Leadership Ability. *PLOS ONE*, 8(12):e80957, December 2013. ISSN 1932-6203. doi: 10.1371/journal.pone.0080957. URL http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0080957.

Restylane. THE RESTYLANE IMAGINE TOOL, 2012. URL http://www.restylane.ca/rest-imagine-tool/.

Helle Rexbye, Inge Petersen, Mette Johansens, Louise Klitkou, Bernard Jeune, and Kaare Christensen. Influence of environmental factors on facial ageing. *Age and Ageing*, 35(2):110–115, March 2006. ISSN 0002-0729. doi: 10.1093/ageing/afj031. URL https://academic.oup.com/ageing/article/35/2/110/28144.

V. Ritts, M. L. Patterson, and M. E. Tubbs. Expectations, impressions, and judgments of physically attractive students: A review. *Review of Educational Research*, 62(4), 1992. URL http://doi.org/10.3102/00346543062004413. bibtex: ritts_expectations_1992.

Rasmus Rothe, Radu Timofte, and Luc Van Gool. Dex: Deep expectation of apparent age from a single image. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pages 10–15, 2015. URL http://www.cv-foundation.org/openaccess/content_iccv_2015_workshops/w11/html/Rothe_DEX_Deep_EXpectation_ICCV_2015_paper.html. bibtex: rothe_dex_2015.

D.A. Rowland and D.I. Perrett. Manipulating facial appearance through shape and color. *IEEE Computer Graphics and Applications*, 15(5):70–76, September 1995. ISSN 02721716. doi: 10.1109/38.403830. URL http://ieeexplore.ieee.org/document/403830/. bibtex: rowland_manipulating_1995.

R. Russell, A. Porcheron, J. R. Sweda, A. L. Jones, E. Mauger, and F. Morizot. *Facial contrast is a cue for perceiving health from the face*. Journal of Experimental Psychology: Human Perception and Performance, Advance online publication, 2016. URL http://doi.org/10.1037/xhp0000219. bibtex: russell_facial_2016.

Richard Russell, Carlota Batres, Sandra Courrèges, Gwenaël Kaminski, Frédérique Soppelsa, Frédérique Morizot, and Aurélie Porcheron. Differential effects of makeup on perceived age. *British Journal of Psychology*, 0(0), 2018. ISSN 2044-8295. doi: 10.1111/bjop.12337. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/bjop.12337.

Nadine Samson, Bernhard Fink, Paul J. Matts, Nancy C. Dawes, and Shannon Weitz. Visible changes of female facial skin surface topography in relation to age and attractiveness perception. *Journal of Cosmetic Dermatology*, 9(2):79–88, June 2010. ISSN 1473-2165. doi: 10.1111/j.1473-2165.2010.00489.x. bibtex: samson_visible_2010.

Kristina Scherbaum, Martin Sunkel, H.-P. Seidel, and Volker Blanz. Prediction of Individual Non-Linear Aging Trajectories of Faces. In *Computer Graphics Forum*, volume 26, pages 285–294. Wiley Online Library, 2007. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1467-8659.2007.01050.x/full. bibtex: scherbaum_prediction_2007.

Robert Shaw, Michael McIntyre, and William Mace. The role of symmetry in event perception. In *Perception: Essays in honor of James J. Gibson.*, pages 317–317. Cornell University Press, Ithaca, NY, US, 1974. ISBN 0-8014-0835-0.

Yukio Shirakabe, Yoshiro Suzuki, and Samuel M. Lam. A New Paradigm for the Aging Asian Face. *Aesthetic Plastic Surgery*, 27(5):397–402, October 2003. ISSN 1432-5241. doi: 10.1007/s00266-003-2099-x. URL https://doi.org/10.1007/s00266-003-2099-x.

Tom Simonite. Virtual face-ageing may help find missing persons, 2006. URL https://www.newscientist.com/article/dn10164-virtual-face-ageing-may-help-find-missing-persons/.

Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. URL https://arxiv.org/abs/1409.1556. bibtex: simonyan_very_2014.

B. R. Spisak, N. M. Blaker, C. E. Lefevre, F. R. Moore, and K. F. B. Krebbers. A face for all seasons: Searching for context-specific leadership traits and discovering a general preference for perceived health. *Frontiers in Human Neuroscience*, 8(792.), 2014. URL http://doi.org/10.3389/fnhum.2014.00792. bibtex: spisak_face_2014.

I. D. Stephen, V. Coetzee, Law Smith, M., and D. I. Perrett. Skin blood perfusion and oxygenation colour affect perceived human health. *PLoS ONE*, 4(4), 2009a. URL http://doi.org/10.1371/journal.pone.0005083. bibtex: stephen_skin_2009.

I. D. Stephen, Law Smith, M. J., M. R. Stirrat, and D. I. Perrett. Facial skin coloration affects perceived health of human faces. *International Journal of Primatology*, 30(6), 2009b. URL http://doi.org/10.1007/s10764-009-9380-z. bibtex: stephen_facial_2009.

I. D. Stephen, I. M. L. Scott, V. Coetzee, N. Pound, D. I. Perrett, and I. S. Penton-Voak. Cross-cultural effects of color, but not morphological masculinity. *on perceived attractiveness of men's faces. Evolution and Human Behavior*, 33(4), 2012. URL http://doi.org/10.1016/j.evolhumbehav.2011.10.003. bibtex: stephen_cross-cultural_2012.

Laura Sydell. Building The Curious Faces Of 'Benjamin Button', 2009. URL https://www.npr.org/templates/story/story.php?storyId=100668766.

D. W. Thompson. On growth and form. *On growth and form.*, 1942. URL https://www.cabdirect.org/cabdirect/abstract/19431401837.

B. Tiddeman, D. M. Burt, and D. Perrett. Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications*, 21, 2001. URL http://doi.org/10.1109/38.946630. bibtex: tiddeman_prototyping_2001.

B. P. Tiddeman, M. R. Stirrat, and David I. Perrett. Towards realism in facial image transformation: Results of a wavelet mrf method. In *Computer Graphics Forum*, volume 24, pages 449–456. Wiley Online Library, 2005. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1467-8659.2005.00870.x/full. bibtex: tiddeman_towards_2005.

J. T. Todd, L. S. Mark, R. E. Shaw, and J. B. Pittenger. The perception of human growth. *Scientific American*, 242(2):132–134, 139A, 140 passim, February 1980. ISSN 0036-8733.

Ming-Han Tsai, Yen-Kai Liao, and I-Chen Lin. Human face aging with guided prediction and detail synthesis. *Multimedia Tools and Applications*, 72(1):801–824, September 2014. ISSN 1380-7501, 1573-7721. doi: 10.1007/s11042-013-1399-7. URL http://link.springer.com/10.1007/s11042-013-1399-7. bibtex: tsai_human_2014.

W. Wang, Z. Cui, Y. Yan, J. Feng, S. Yan, X. Shu, and N. Sebe. Recurrent Face Aging. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2378–2386, June 2016. doi: 10.1109/CVPR.2016.261.

W. Wang, Y. Yan, Z. Cui, J. Feng, S. Yan, and N. Sebe. Recurrent Face Aging with Hierarchical AutoRegressive Memory. *IEEE Transactions on Pattern Analysis and*

Machine Intelligence, PP(99):1–1, 2018. ISSN 0162-8828. doi: 10.1109/TPAMI.2018.2803166.

Ross D. Whitehead, Daniel Re, Dengke Xiao, Gozde Ozakinci, and David I. Perrett. You Are What You Eat: Within-Subject Increases in Fruit and Vegetable Consumption Confer Beneficial Skin-Color Changes. *PLOS ONE*, 7(3):e32988, March 2012. ISSN 1932-6203. doi: 10.1371/journal.pone.0032988. URL http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0032988.

CIE XYZ. CIE XYZ, January 2018. URL https://fr.wikipedia.org/w/index.php?title=CIE_XYZ&oldid=144673866. Page Version ID: 144673866.

M. Yaar and B. A. Gilchrest. Photoageing: mechanism, prevention and therapy. *The British Journal of Dermatology*, 157(5):874–887, November 2007. ISSN 0007-0963. doi: 10.1111/j.1365-2133.2007.08108.x. bibtex: yaar_photoageing_2007.

L. A. Zebrowitz, R. G. Franklin, J. Boshyan, V. Luevano, S. Agrigoroaei, B. Milosavljevic, and M. E. Lachman. Older and younger adults' accuracy in discerning health and competence in older and younger faces. *Psychology and Aging*, 29(3), 2014. URL http://doi.org/10.1037/a0036255. bibtex: zebrowitz_older_2014.

Zhanpeng Zhang, Ping Luo, Chen Change Loy, and Xiaoou Tang. Learning Social Relation Traits from Face Images. *arXiv:1509.03936 [cs]*, September 2015. URL http://arxiv.org/abs/1509.03936. arXiv: 1509.03936 bibtex: zhang_learning_2015.

Zhifei Zhang, Yang Song, and Hairong Qi. Age Progression/Regression by Conditional Adversarial Autoencoder. *arXiv:1702.08423 [cs]*, February 2017. URL http://arxiv.org/abs/1702.08423. arXiv: 1702.08423.

Zhong Zhang and R. S. Blum. A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. *Proceedings of the IEEE*, 87(8):1315–1326, August 1999. ISSN 0018-9219. doi: 10.1109/5.775414.

List of Figures

List of Tables