Prédire l'âge de personnes à partir de photos du visage : une étude fondée sur la caractérisation et l'analyse de signes du vieillissement

Alex A. Nkengne

THESE DE DOCTORAT DE
L'UNIVERSITE PIERRE ET MARIE CURIE

Informatique Biomédicale
…………………………………………………….
ED 393 – SANTE PUBLIQUE : Epidémiologie et sciences de l'Information Biomédicale

Présentée par

M NKENGNE NGUIMEZONG Alex Albert

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE

Sujet de la thèse :

Prédire l'âge de personnes à partir de photos du visage : une étude fondée sur la caractérisation et l'analyse de signes du vieillissement

Predicting people's age from their facial image:
a study based on the characterization and the analysis of the signs of aging

Soutenue le 13 Juin 2008

Devant le jury composé de :

Philippe HUMBERT	Professeur, Université de Franche-Comté	Rapporteur
Alice CAPLIER	Professeur assistante, ENSERG	Rapporteur
Bernard FERTIL	Directeur de recherche, CNRS	Directeur de thèse
Nathalie ISSACHAR	Directeur de recherche, Johnson & Johnson	Co-encadrante
NIkiforos KOLLIAS	Directeur de recherche, Johnson & Johnson	Examinateur
Jean Gabriel GANASCIA	Professeur, Université de Paris VI	Examinateur
Maurice MILGRAM	Professeur, Université de Paris VI	Examinateur

Université Pierre & Marie Curie - Paris 6
Bureau d'accueil, inscription des doctorants et base de données
Esc G, 2ème étage
15 rue de l'école de médecine
75270-PARIS CEDEX 06

Tél. Secrétariat : 01 42 34 68 35
Fax : 01 42 34 68 40
Tél. pour les étudiants de A à EM : 01 42 34 69 54
Tél. pour les étudiants de EN à ME : 01 42 34 68 41
Tél. pour les étudiants de MF à Z : 01 42 34 68 51
E-mail : scolarite.doctorat@upmc.fr

Au Père Marie Elie,
A mes chers parents

Remerciements

Merci à vous qui avez conduit mes pas depuis la tendre enfance, à vous qui m'avez entouré de votre affection, choyé, protégé, encouragé.

Merci à vous ma famille, mes amis, mes nombreux pères et mères. Ce travail de thèse, est le fruit de votre Amour et de votre Sueur, soyez en fiers.

Merci à vous qui avez suscité en moi une vocation pour la science et m'avez montré les chemins de l'excellence.

Merci, à vous qui patiemment m'avez appris à décrypter les mystères du monde et à lier le bois au bois.

Merci à vous qui m'avez transmis votre flamme, votre passion de la science et de la vérité.

Merci à vous tous, enseignants et guides. Puissiez-vous retrouver dans cette thèse quelques unes des idées que vous vous êtes donnés tant de mal à m'inculquer.

Merci à vous qui avez accueilli à bras ouvert l'inconnu et m'avez donné une place à votre table.

Merci à vous qui m'avez donné l'opportunité de traverser les mers pour poursuivre un idéal.

Merci à vous, amis et mécènes qui avez cru en moi. Ce travail, je l'espère est à la mesure de la confiance que vous m'avez accordé.

La vie est faite de rencontres qui nous enrichissent et nous portent vers de nouveaux horizons. Maintenant que s'achève cette étape du voyage, je souhaiterais en particulier exprimer ma gratitude à ceux qui ont partagé ces années de thèse.

Mes premiers mots de remerciement vont à ceux qui ont rendu possible cette aventure, Nathalie Issachar et Bernard Fertil. Vous avez cru en moi et en mon projet et m'avez accueilli à bras ouvert dans vos équipes ; je vous en suis très reconnaissant.

J'exprime également toute ma reconnaissance aux membres de l'unité UMR 678 et en particulier à son directeur Alain Herment. Les moments passés avec vous ont été source d'émulation et d'inspiration.

J'ai une pensée particulière pour Alain Giron, véritable poumon de l'équipe IV ; ton dévouement m'aura redonné courage dans les moments d'incertitude.

L'équipe IV a été un véritable lieu de vie, ou j'ai eu le plaisir de me lier d'amitié à des étudiants tous plus généreux les uns que les autres. Arthur, Sylvain, Jean-François (mon super binôme), Caroline, Sahra et bien entendu Clara (par adoption), je n'oublierais pas les excellents moments passés en votre compagnie.

Une autre équipe qui aura su m'ouvrir ses bras est celle du Skin Care Research Institute. Grâce à vous, Christiane, Arlette, Ana, Valérie, Anne, Romain, Aline, Emilie, Estelle, Marion, Fanny, Maryline, Stéphanie, Delphine j'ai passé et continue de passer de très bons moments. Je vous exprime ma reconnaissance pour cette bonne humeur que vous savez si bien distiller. J'ai une attention particulière pour Christiane, Stéphanie, Delphine et Arlette qui ont largement contribué à la collecte et à l'analyse des données présentées dans ce travail. Merci de votre précieuse collaboration.

Je voudrais également remercier tous ceux dont les réflexions et l'énergie ont contribué à améliorer mon travail de thèse et ce document final. Christiane Bertin, Georgios Stamatas, l'increvable Alain et bien sur Bernard ; merci à vous pour le temps et l'énergie consacrée. Toute ma reconnaissance va également à Rachid Belaroussi qui s'est dévoué six mois durant pour me fournir une méthode adaptée de segmentation.

Enfin, j'ai l'honneur et le plaisir d'avoir dans mon jury les Professeurs Philippe Humbert, Alice Caplier, Maurice Milgram et Jean Gabriel Ganascia ainsi que les Docteurs Nikiforos Kollias, Nathalie Issachar et Bernard Fertil. Vos travaux à tous suscitent mon admiration. Je vous suis donc infiniment reconnaissant d'avoir accepté d'être membres de mon jury de thèse.

Résumé

L'âge a de tout temps constitué un attribut identitaire important. Nous avons développé au fil de l'évolution une aptitude innée à classer les individus en fonction de leur âge. Cette classification s'appuie en grande partie sur le visage et sur les transformations anatomiques qu'il subit au cours du temps. De plus en plus de traitements cosmétiques, dermatologiques et d'interventions chirurgicales s'attaquant à un signe ou un groupe de signes spécifiques du vieillissement sont mis en œuvre pour annuler, ou tout au moins masquer partiellement l'effet du temps sur le visage. On peut dès lors s'interroger sur l'influence de chacun des signes sur notre capacité à prédire l'âge d'un individu en observant son visage. Afin de construire un algorithme capable de déterminer l'âge d'individus à partir de leurs photos, nous nous sommes intéressés aux signes du vieillissement et à leur impact sur l'âge apparent. Dans un premier temps, nous avons déterminé et analysé les transformations anatomiques qui altèrent le visage à partir de l'âge adulte (au-delà de 20 ans). Puis nous avons étudié les signes sur lequel on se base pour prédire l'âge d'une personne. Enfin, nous avons construit et validé un modèle prédictif de l'âge en s'appuyant sur les observations précédentes.

Transformations anatomiques du visage avec l'âge : La prévalence d'un certain nombre de signes de vieillissement (rides, tâches brunes, forme du visage,…) a été mesurée sur un panel représentatif de femmes volontaires âgées de 20 à 74 ans. Ces données ont permis d'établir la cinétique d'apparition de ces signes.

Appréciation subjective de l'âge: Il s'agissait de déterminer les signes sur lesquels un observateur s'appuie lorsqu'il évalue l'âge d'un sujet. Pour ce faire, nous avons demandé à un panel constitué de 48 observateurs d'attribuer un âge aux volontaires sur lesquelles nous avions précédemment mesuré les signes du vieillissement. Nous avons confirmé avec ce groupe d'observateurs que la perception de l'âge est liée au sexe et à l'âge de l'observateur. De plus, à l'aide d'une régression PLS (Partial Least Square régression), nous avons établi des relations entre les signes du vieillissement et l'âge observé et démontré que selon que l'on soit jeune ou âgé, un homme ou une femme, on n'exploite pas les mêmes signes de vieillissement pour prédire l'âge.

Modèle de prédiction : Enfin, nous avons proposé un modèle s'appuyant sur la régression PLS pour prédire automatiquement l'âge à partir des photos du visage. Ce modèle présente la particularité d'associer, dans une approche unifiée, les signes relatifs à la couleur, à la forme et à la texture du visage à l'âge des sujets. A l'instar des Modèles Actifs D'apparence (AAM), le modèle construit vise à réduire fortement l'information portée par l'ensemble des pixels du visage. Toutefois, ce dernier est supervisé : Il est donc très approprié dans notre contexte puisque que l'on peut mettre en œuvre une procédure d'apprentissage pilotée par le but. Les performances sont de fait comparables à celles des humains.

Summary

Age has always been an important identity attribute. Mankind has developed, throw the evolutionary process, an innate ability to classify individuals according to their ages. This classification is partly driven by the face and the anatomical transformations that occur with time. Many dermatological, cosmetic or surgical procedures are developed to fight again signs of aging. Therefore, we can wonder how these procedures modify the perceived age of people who achieve them.

In order to build an algorithm that will predict someone age from his front face picture, we have study the signs of facial aging and their incidence on perceived age. Firstly we have analyzed the anatomical transformations that alter the adult face. Secondly, we have determined the signs of aging which mostly drive the human perception of age. Finally, we have built and validated a predictive model of age thanks to the results of the first two steps.

Anatomical transformation of the face with age: The importance of 21 signs of aging has been measured on a representative panel of Caucasian women aged between 20 and 74 years. This data have enabled to build the kinetic of facial aging.

Subjective judgment of age: The objective was to determine which signs drive observers' perception when evaluating people age. Forty height observers were therefore asked to give an age to the volunteers whose aging signs were previously measured. The perception of age was shown to be bias by the age and the gender of the observers. Moreover, the relation between the signs of aging and the perceived age was evaluated using Partial Least Square (PLS) regression. Particularly, it was shown that depending on the age or the gender of the graders, they do not similarly use the signs aging when predicting peoples' age.

Age prediction model: Finally a model was proposed to predict peoples' age from their front face image using PLS regressions. The proposed model combines and the signs of aging related with the color, the shape and the texture of the face. Like the Active Appearance Model (AAM), the proposed model allows to strongly reduce the dimensionality of the information carried by the pixels values. However, this model is supervised and is thus appropriate in our context of learning by sample. The ability of our model to predict peoples' age is finally comparable with human.

Table of contents

List of figures

List of tables

Main references

Cootes, T. F., G. J. Edwards, et al. (2001). "Active appearance models." IEEE Transactions on Pattern Analysis and Machine Intelligence **23**(6): 681-685.

Cootes, T. F., C. J. Taylor, et al. (1995). "Active shape models-their training and application." Computer Vision and Image Understanding **61**(1): 38-59.

CrimeLib (2002). Age Progression. Criminal Minds and Methods. C. T. s. C. Library.

Edwards, G. J., C. J. Taylor, et al. (1998). "Interpreting face images using active appearance models." 3 rdInternational Conference on Automatic Face and Gesture Recognition: 300–305.

Fu, Y., Y. Xu, et al. (2007). ESTIMATING HUMAN AGE BY MANIFOLD ANALYSIS OF FACE PICTURES AND REGRESSION ON AGING FEATURES. International Conference on Multimedia & Expo. Beijing, China.

Gandhi, M. R. (2004). A Method for automatic Synthesis of Aged Human Facial Images. Department of Electrical & Computer Engeneering. Montreal, Canada, McGill University. **Master Degree:** 107.

Geng, X. and K. Smith-Miles (2007). "Automatic Age Estimation Based on Facial Aging Patterns." IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE **29**(12): 2234-2240.

Geng, X., Z. H. Zhou, et al. (2006). "Learning from facial aging patterns for automatic age estimation." Proceedings of the 14th annual ACM international conference on Multimedia: 307-316.

Hayashi, J., M. Yasumoto, et al. (2002). "Age and gender estimation from facial image processing." SICE 2002. Proceedings of the 41st SICE Annual Conference **1**.

Horng, W. B., C. P. Lee, et al. (2001). "Classification of Age Groups Based on Facial Features." Tamkang Journal of Science and Engineering **4**(3): 183-192.

Hubball, D., M. Chen, et al. (2006). Evolutionary Morphing of Facial Images for Aging Simulation, University of Wales Swansea.

Kwon, Y. H. and N. da Vitoria Lobo (1994). "Age classification from facial images." Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on: 762-767.

Lanitis, A., C. Draganova, et al. (2004). "Comparing different classifiers for automatic age estimation." Systems, Man and Cybernetics, Part B, IEEE Transactions on **34**(1): 621-628.

Liu, J., N. Zheng, et al. (2007). "Estimating Aging Pattern by Aging Increment Distribution for Re-rendering of Facial Age Effects." LECTURE NOTES IN COMPUTER SCIENCE **4681**: 782.

Lobo, N. and Y. Kwon (1998). Automatic feature detection and age classification of human faces in digital images, Google Patents.

Ramanathan, N. and R. Chellappa (2006). "Face verification across age progression." IEEE transactions on image processing **15**(11): 3349-3361.

Ramanathan, N. and R. Chellappa (2006). "Modeling Age Progression in Young Faces." Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on **1**.

Ricanek Jr, K. and T. Tesafaye (2006). "MORPH: a longitudinal image database of normal adult age-progression." Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on: 341-345.

Seoul, K. (2004). Extraction and Manipulation of Wrinkles and Spots for Facial Image Synthesis. Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition.

Tiddeman, B. P., M. R. Stirrat, et al. (2005). "Towards Realism in Facial Image Transformation: Results of a Wavelet MRF Method." Computer Graphics Forum **24**(3): 449-456.

French long summary

I. Introduction

L'âge a toujours été un facteur important d'interaction sociale. La posture, le vocabulaire et l'intonation sont autant d'éléments qu'on adapte à l'âge de son interlocuteur. De manière plus large, on peut noter que la société propose des activités et des loisirs différents en fonction de l'âge des individus. Ainsi, pouvoir prédire l'âge demeure une aptitude importante qui se développe dès la tendre enfance. Cette aptitude s'appuie sur de nombreux signes qui peuvent aussi bien être liés à la tenue vestimentaire qu'à la posture ou aux traits du visage.

Dans cette thèse, nous nous intéressons en particulier au visage et aux transformations qui l'affectent avec l'âge, à la capacité d'un individu à s'appuyer sur ces transformations pour prédire l'âge et à la possibilité de construire un algorithme automatique de prédiction. Les objectifs de ce travail sont donc de:

1. Décrire les changements qui affectent le visage au cours du temps
2. Analyser l'influence de ces changements sur la perception de l'âge
3. Proposer un algorithme permettant de prédire l'âge d'individus à partir de leurs photos

Les effets du temps étant variables d'un sexe à un autre et d'une race à une autre, nous nous restreignons dans cette thèse à une population de femmes Caucasiennes.

Ce document est divisé en quatre chapitres. Le premier chapitre s'intéresse au vieillissement anatomique du visage et présente un résumé des connaissances acquises. Il décrit les phénomènes biologiques liés à l'âge et leur influence sur l'apparence du visage.

Dans le second chapitre, nous nous intéressons en particulier aux femmes Caucasiennes et présentons une étude menée afin de quantifier les transformations anatomiques du visage liées à l'âge. Nous évaluons l'évolution de 21 paramètres en fonction de l'âge sur une population constituée de 173 femmes âgées de 20 à 73 ans et établissons la cinétique du vieillissement.

Dans le troisième chapitre, nous mesurons l'influence relative des signes du vieillissement sur l'âge perçu. Nous demandons à 48 observateurs d'évaluer l'âge des 173 volontaires précédentes. Un modèle PLS permet ensuite d'établir une pondération des signes du vieillissement relativement à la perception d'âge.

Le quatrième chapitre s'appuie sur les résultats précédents pour proposer un modèle de prédiction d'âge à partir de photos. Ce modèle est utilisé de manière concluante sur les photos des 173 précédentes volontaires.

II. Vieillissement du visage

Le temps qui passe altère irrémédiablement le visage. De nombreuses publications décrivent l'ensemble des transformations qui modifient aussi bien la forme, que la texture et la couleur du visage. Le résumé de ces transformations est le suivant :

1. La forme du visage:

   La forme du visage est liée à la disposition des os, des tissus mous (muscles et graisse) et à la fermeté de la peau. Au cours du temps, chacun de ces éléments subit des changements importants. Quelques unes de ces transformations sont :

   - Au niveau du squelette :
     - La réduction de la hauteur du visage
     - La perte des dents et l'ostéoporose des os maxillaires et mandibulaires
     - L'augmentation da la largeur du visage
   - Au niveau des tissus mous
     - La myopathie de tous les muscles du corps
     - L'hypertrophie des muscles orbiculaires et frontaux entrainant l'apparition d'une ptose sur les sourcils et des poches sous les yeux
     - L'infiltration de graisse dans le nez provocant la ptose de la pointe de nez
     - La redistribution de graisses sur l'ensemble du visage le rendant plus concave
   - Au niveau de la peau
     - L'apparition de sillons marqués tels que le sillon nasogélien
     - La perte de fermeté favorisant le relâchement des paupières et sur la partie base du visage l'apparition de bajoues ou d'un double menton

2. La couleur du visage:

   La couleur du visage est principalement liée à l'épaisseur de l'épiderme et à la distribution des différents chromophores (mélanine et hémoglobine). Au fil du temps, la peau devient plus fine et les vaisseaux sous-jacents plus visibles, ce qui entraine un ternissement et un jaunissement du teint. Par ailleurs, l'exposition du visage au soleil favorise la non homogénéisation de la distribution de la mélanine provocant l'apparition de tâches pigmentaires.

3. La texture du visage :

   La texture du visage est altérée suite à l'apparition des rides et ridules. On peut distinguer les rides d'expressions qui apparaissent sur les parties du visage en mouvement telles que le front et des rides liés à l'âge. Toutes les rides se creusent avec le temps, suite à la dégradation du collagène et de l'élastine contenus dans le derme.

Bien que toutes les transformations décrites ci-dessus soient générales, elles s'opèrent plus ou moins vite en fonction de nombreux facteurs tels que l'origine raciale, le sexe, l'exposition au soleil, la pollution, la consommation de tabac, … La prochaine section s'intéresse spécifiquement au vieillissement de femmes Caucasiennes.

III. Vieillissement du visage des femmes Caucasiennes

Répondre aux attentes des femmes soucieuses de ralentir les effets du temps est l'une des principales missions des cosméticiens et chirurgiens plastiques. Dans ce contexte, la mesure quantitative des changements liés à l'âge permet de proposer de meilleurs traitements d'une part, et d'autre part de vérifier l'efficacité des procédures disponibles. La métrologie cutanée est la science qui s'est développée afin de mesurer les attributs cutanés et de suivre leur évolution. Johnson & Johnson dispose d'un laboratoire de métrologie au sein duquel une étude a été menée sur 173 femmes Caucasiennes âgées de 20 à 74 ans afin de mesurer les changements affectant leurs visages et liés à l'âge.

Pendant cette expérience 21 attributs choisis suite à une revue de la littérature ont été évalués par un expert à l'aide d'échelles visuelles analogiques. Le tableau ci-dessous fait le récapitulatif des attributs évalués et de leur corrélation avec l'âge.

Paramètre	Coeff corr (R)	p
Rides du front	0.61	0.00E+00
Sillon nasogélien	0.70	0.00E+00
Couleur de la peau	-0.66	0.00E+00
Paupières tombantes	0.61	0.00E+00
Rides de la patte d'oie	0.80	4.75E-40
Ride du lion	0.76	1.98E-33
Rides au dessus des lèvres	0.75	1.37E-28
Rides sous les yeux	0.57	1.11E-16
Ovale du visage	-0.57	1.11E-16
Ridules de la patte d'oie	0.50	3.80E-12
Epaisseur des lèvres	-0.48	1.40E-11
Tâches brunes	0.41	2.48E-08
Poches sous les yeux	0.40	4.00E-08
Définition du bord des lèvres	-0.39	1.26E-07
Ouverture des yeux	-0.41	1.43E-07
Uniformité du teint	-0.34	5.21E-06
Cernes	0.33	9.23E-06
Grain de peau	-0.08	2.86E-01
Rougeurs diffuses	0.07	3.51E-01
Radiance	-0.07	3.72E-01
Texture de la peau	0.04	5.82E-01

Table 0-1: Corrélation entre les attributs mesurés et l'âge des femmes

Comme le montre ce tableau de nombreux paramètres mesurés sont significativement liés à l'âge. En regardant la forme des courbes, nous mettons par ailleurs en évidence l'incidence de la ménopause sur 13 attributs. Sept attributs semble être négativement affectés par la ménopause (accélération de la dégradation), tandis que six d'entre eux sont positivement affectés (ralentissement).

Nous nous intéressons aussi à la reproductibilité de la méthode utilisée et montrons qu'elle convient pour mesurer la plupart des paramètres cutanés de manière reproductible. Deux experts sont invités à évaluer 31 volontaires et les résultats de leurs évaluations sont comparés en utilisant le coefficient de corrélation de Pearson. Nous obtenons une corrélation significative pour 70% des attributs comparés. Nous proposons néanmoins l'usage d'échelles construites à partir de photos afin de faciliter l'entrainement des évaluateurs et d'améliorer leur reproductibilité.

Les attributs évalués dans cette section sont utilisés par la suite pour comprendre les signes sur lesquels s'appuie un observateur pour prédire l'âge d'un sujet.

IV. Perception du vieillissement

Une revue de la littérature nous permets d'établir que :

- La capacité à prédire l'âge est acquise dans les premières années de l'enfance
- Cette capacité est sujette à un biais lié à la race, à l'âge et au sexe des observateurs ; chacun étant plus à même de prédire l'âge de personnes du même groupe que lui.
- La taille du visage, la couleur de la peau, la présence de rides, la vitalité du regard sont autant d'éléments qui permettent de juger de l'âge d'un individu. Néanmoins aucune hiérarchisation de ces paramètres ne semble faire le consensus, probablement à cause des protocoles expérimentaux utilisés qui ne permettent pas de directement comparer l'influence des différents attributs faciaux (approches unidimensionnelles).

Afin de déterminer les attributs influençant le plus la prédiction d'âge nous proposons une nouvelle méthodologie s'appuyant sur la régression PLS et permettant d'avoir une approche multidimensionnelle du problème.

Quarante huit observateurs ; 20 hommes et 28 femmes ; âgés de 20 à 64 ans évaluent l'âge des 173 précédentes volontaires à partir de photos de visage prises dans des conditions standardisées. L'âge apparent est définit comme l'âge moyen donné par les évaluateurs.

Dans un premier temps, la comparaison des évaluateurs masculins et féminins confirme que les femmes sont plus précises pour prédire l'âge d'autres femmes comme le laissait prévoir la littérature. Par ailleurs, la comparaison de trois groupes d'évaluateurs : les jeunes (20-35 ans), les personnes d'âge intermédiaire (36-55 ans), et les personnes âgées (56-65 ans) met en évidence des différences liées à l'âge des évaluateurs. Le groupe de jeunes montre une meilleure précision que les autres groupes. Néanmoins, nous ne confirmons pas sur ces données que les individus d'un groupe d'âge sont plus précis lorsqu'ils évaluent des gens de la même tranche d'âge.

Enfin, nous construisons sept modèles de régression PLS pour prédire l'âge à partir des 21 attributs mesurés. Chacun des modèles prédit respectivement l'âge réel, l'âge apparent donné par l'ensemble des évaluateurs, les âges apparents donnés par les groupes d'hommes et de femmes et finalement les âges apparents donnés par les trois groupes d'âge (Jeunes-âge intermédiaire, personnes âgées).

L'importance relative des pondérations du modèle PLS nous permet de connaître l'importance de chacun des attributs pour le modèle de prédiction d'âge. Nous établissons ainsi que :

- Les attributs les plus importants sont dans un ordre décroissant l'épaisseur des lèvres, la forme de l'ovale du visage, les rides autour de la bouche et l'uniformité du teint.

- Les hommes et les femmes ne présentent pas de différence significative concernant l'utilisation des différents attributs pour la prédiction d'âge.
- L'importance du sillon nasogélien, des cernes, des poches sous les yeux, de l'uniformité du teint décroit avec l'âge des observateurs. Au contraire, celle des rides de la patte d'oie, de l'ouverture des yeux, de l'épaisseur des lèvres augmente avec l'âge des volontaires.

Ces résultats qui vont au-delà de ce que décrit la littérature sont en cours de publication et gagneraient à être confirmés par d'autres expériences.

V. Prédiction de l'âge à partir de photos

Nous proposons un algorithme de prédiction de l'âge en s'appuyant sur les observations faites dans les parties précédentes. Cet algorithme qui s'inspire des Modèles Actifs d'Apparence (AAM) peut se décomposer en trois grandes étapes :

- D'abord les principaux traits du visage (contour, yeux, bouche, nez) sont localisés à l'aide de points d'encrages. Toutes les images sont alignées à l'aide de ces points de manières à ce que les traits principaux se retrouvent à la même position.
- Deux séries de modèles PLS sont construit pour prédire l'âge à partir des informations provenant des photos. Le modèle de forme s'appuie sur la position des points d'encrage et décrit les changements relatifs à l'âge dans la forme des visages. La deuxième série est constituée de trois modèles PLS décrivant les variations d'intensité des pixels sur chacun des canaux de couleur (RVB) en fonction de l'âge. Chacun de ces modèles constitue donc un descripteur de la texture du visage. Tous ces modèles fournissent un ensemble de variables latentes qui résument l'information relative à l'âge des individus de la base d'apprentissage.
- Les variables latentes des quatre précédents modèles sont normalisées avant de servir de variables d'entrée pour la construction d'un modèle global de prédiction d'âge.

Appliquée à notre base de données, cette approche nous permet de visualiser les attributs les plus importants pour la prédiction d'âge ainsi que leur transformation au cours du temps. Elle met ainsi en exergue l'importance de la forme de la bouche, du menton (ovale du visage) et de la taille des yeux. Elle permet aussi de rendre compte de l'importance de la région autour des yeux, ainsi que du sillon nasogélien et des bajoues.

Sur une base de test, le modèle proposé prédit l'âge avec une Erreur Moyenne Absolue (MAE) de 5.98 ans, qui est comparable à l'erreur humaine pour la même tâche. La capacité de prédiction s'améliore (MAE = 4.87 ans) lorsqu'on apprend à prédire l'âge perçu plutôt que l'âge réel. Ce résultat laisse entendre que le modèle s'appuie sur le même type d'informations que les observateurs humains.

VI. Conclusion et perspectives

Dans ce travail de thèse, nous nous sommes intéressés aux changements liés au vieillissement du visage et à leur perception. Dans un premier temps, nous avons quantifié sur une population de femmes caucasiennes l'évolution de signes liés à la couleur et à la texture de la peau, à la forme du visage et des lèvres ainsi qu'au contour des yeux. Les observations effectuées ont montré la corrélation de la plupart des signes choisis avec l'âge, et ceux malgré les variations interindividuelles. Par la suite nous avons analysé l'importance des signes du vieillissement sur la perception d'âge. Pour se faire nous avons proposé une approche originale, reposant sur la régression PLS. A notre surprise, nous avons trouvé que les personnes jeunes et les personnes âgées utilisent différemment les attributs du visage lorsqu'elles prédisent l'âge. Néanmoins, nous avons montré que les principaux attributs guidant la perception sont liés à l'épaisseur des lèvres, la forme de l'ovale du visage, les rides autour de la bouche et l'uniformité du teint. Tous ces enseignements ont alimentés notre approche pour la proposition d'un algorithme de prédiction automatique de l'âge à partir de photos.

Le modèle que nous avons proposé pour prédire l'âge s'appuie sur des informations aussi bien liées à la forme qu'à la couleur et à la texture du visage. Ce modèle utilise un nouvel algorithme de compression supervisée des visages, le Modèle Supervisé de Visage (SFM) reposant sur la régression PLS. Le SFM permet de réduire l'information portée par l'ensemble des pixels de l'image (plusieurs milliers) en un nombreux restreint de paramètres (quelques dizaines) portant l'information relative à l'âge. Cette approche nous permet de prédire l'âge avec une précision comparable avec celle d'un observateur humain.

Néanmoins, quelques améliorations pourraient encore augmenter les performances de cet algorithme. Dans un premier temps, il serait surprenant que la relation entre la valeur des pixels de l'image et l'âge des sujets n'ait pas une composante non-linéaire. Par conséquent, l'utilisation d'un modèle non-linéaire pour la compression des images du visage pourrait améliorer la qualité de la prédiction. Par ailleurs, le modèle de texture utilisé n'exploite que faiblement les informations contenues dans les hautes fréquences (rides et ridules). Des paramètres issus d'une analyse fréquentielle (Fourier ou ondelettes) permettraient de mieux encoder ces informations. Enfin, le modèle pourrait être étendu aux images en 3 dimensions (3D) afin non seulement de prédire l'âge à partir d'image 3D, mais surtout de simuler le vieillissement.

Introduction

Coming from an African country where people are respected for the graying of their hair and beard, one may be surprised in the western societies by the high number of products and procedures invented everyday to make people look younger. In fact, since Alexander the Great have started searching for the legendary Fountain of Youth, a great part of the humanity has struggled to stop or at least to slow down the effects of time on people's appearance.

Meanwhile, people's age remains an important attribute of social interactions. The way we act, the vocabulary we choose, our body attitude in front of someone else depends on his age. Moreover specific works, entertainment, lectures, movies, etc. are assigned to groups of age. Therefore people since childhood develop a capacity to estimate other people's age based on physical attributes. These attributes can be related to dress code or body gesture, but they are primarily linked to facial appearance. It is obvious that anyone can distinguish a baby face from an adult and from a senior one since some evident characteristics such as the size of the head and the number of wrinkles are affected dramatically by age. Consequently it would be useful to understand all the changes that occur on faces with age and their effect on people's appearance.

Furthermore, nowadays and even more so in the near future we will need to interact not only with human beings but also with computers and robots. A lot of research focusing on improving human to machine interactions is made in order to enable natural communication. One of the most accessible applications of these research results is the ability for the newest digital cameras to locate the faces in images and to capture pictures when people are smiling. Other applications include people identification at secured access points or faces tracking in video movies. However, to our knowledge, little work has been done on age prediction and there is no public accessible application that is commercialized which uses an age prediction algorithm. However many reasons have already been identified that justify the need for age prediction algorithms.

Objectives and motivation

This thesis focuses on changes that occur with age on Caucasian women's faces. Our three objectives are:

1. To describe the facial changes with age
2. To explore the incidence of these changes on the perceived age
3. To propose an automatic algorithm for age prediction based on facial images.

Various benefits could be derived from these three objectives in the fields of cosmetic dermatology, psychology and image processing. These benefits will be presented in details in the remaining of the document. Here are some examples:

- By understanding the contribution of each facial attribute to the overall aging appearance, one can design better procedures for facial rejuvenation.
- Understanding how people code and interpret faces is still an open problem in psychology. This work can contribute to better understanding of how signs of aging are read and interpreted.

Automatic algorithms for age prediction could be used for age-based indexing of facial images from a database. Furthermore, the work done for this algorithm could be extended for age simulation algorithms. Age simulation can be used in image processing of a facial image to show how someone will look after several years. Age simulation is useful to find missing children or to capture wanted fugitives.

Overview

Age or ages?

The concept of age measures the number of years that have passed since birth. When dealing with facial aging, one may be interested in three different kinds of aging: chronological aging, photoaging and apparent aging. Chronological aging or intrinsic aging refers to the natural physiological changes in the facial tissues. Those changes are related to the growth and the degeneration of the body over time and occur throughout the full body. Since the face is more exposed than the rest of the body to environmental factors such as sun, photoaging is also well noticeable on facial skin. The effects of photoaging underscore the ones produced by chronological aging leading to a gap between chronological age and apparent age. Apparent age is the average age given to a person by other peoples. It is a subjective judgment: a robust estimation can only be obtained by averaging multiple estimations. Some people look younger or older than their real age, depending on their genetic makeup, sun exposure, smoking habits and mood. Some events in life like the death of a parent or a long disease can accelerate the process of apparent aging. Make up, anti-aging creams or surgery can reduce the apparent age.

Does everyone age similarly?

We have already pointed out that apparent age depends on many internal and external factors. Changes with age are also tightly related to ethnicity. Within a same ethnic population, a wide variability is observed in physical appearance of individuals within the same age group. There is also a difference between apparent and chronological age. Thus, studies devoted to the measure of the anatomical transformations related to age should include a large number of subjects to account for individual variations. Since we are interested in facial changes with age, it is necessary to eliminate the characteristics that are related to the subject identity.

Thesis contributions

In this thesis we explore three different areas related to facial aging, confirm some recent findings and provide a unified approach of age prediction. The main contribution of this document includes:

- On facial aging:
  We quantitatively describe the age-related changes on Caucasian women
- On age perception:
  We confirm that the perception of age is related to the age and the gender of the observer. We also establish a ranking of the signs of aging and show that depending on the observer's age, he/she gives a different importance to each facial attribute
- On age prediction:
  We propose a new method to describe a face using a small set of. We also propose a new algorithm that enables to predict people's age with accuracy comparable to that of human.

Structure of the thesis

The next chapter describes the anatomical changes related to facial aging. It enables to understand the biological phenomena related to aging and their incidence on facial appearance.

In the chapter 2, we focus on aging of Caucasian women and present an experiment performed to record age-related changes in this population. We quantitatively describe the transformation of 21 facial attributes measured clinically on 173 volunteers from 20 to 74 years of age.

In the chapter 3, we focus on the influence of these 21 facial attributes on age perception. Firstly, we assign a perceived age to each of the 173 volunteers by a consensus among 48 assessors who were asked to give an age to these volunteers. Then we use a Partial Least Square (PLS) regression to build explanatory models of age. These models permit us to measure the relative contribution of each attribute to the perception of age.

The chapter 4 introduces a new age prediction algorithm. Firstly, we present a model that encodes facial images with a small set of parameters. We subsequently use it to predict the age of the 173 volunteers of our study from their facial images.

Chapter 1
Facial aging

The process of aging induces several changes on face. Despite of individual variations linked to genetic and living conditions, some general rules are found. These rules are gender and ethnic dependant. Differences between men's and women's aging are attributed to sex-specific characteristics such as skin thickness and hormonal activity, especially in relation with menopause (Broniarczyk-Dyla and Joss-Wichman 2001; Wines and Willsteed 2001; Raine-Fenning, Brincat et al. 2003; Sumino, Ichikawa et al. 2004; Quatresooz, Pierard-Franchimont et al. 2006).

Generally speaking, if faces are considered as 3D objects, the transformations that occur with age belong to three groups: shape, texture and color. Although unusual when looking at facial aging, this classification enables us to describe the aging process, keeping in mind that our objective is to extract signs of aging from facial images. Here is a short description of the three levels that we have considered:

- Changes on shape. They occur in the spatial dimensions with a centimetric range. They may be caused by the growth of bones or their osteoporosis, changes in muscle and fat distributions, skin sagging caused by the lost of elasticity and gravity. These changes can be recorded on low resolution images. They are affected by lighting conditions (orientations and shadows) and head positioning.

- Changes on texture. They happen in the spatial dimensions at a millemetric range. They are linked with skin aging and can involved wrinkles, pores and microrelief. They are not recorded on low resolution images.

- Changes on color. They are also linked with skin aging. They involve tone yellowish and unevenness with the apparition of brow spots.

This chapter depicts the anatomic transformations that influence the appearance of the face with age. Firstly, a brief description of the anatomy of the face is made in order to help better understanding these morphological transformations. Then the anatomic changes with age, reported by the literature are presented. Finally, the effects of some factors that influence the aging process are discussed.

I. Facial anatomy

The study of the facial anatomy is important for understanding the appearance of the face. Since anatomy is not the focal topic of this document, description is short: the objective is to highlight the main elements that contribute to the physical shape, color and texture of the face.

Facial anatomy is composed of three main elements: skin, soft tissue, and the underlying skeleton; each of them being affected by age differently.

I.1. Bones

Bones give the body a framework, maintain its shape, and protect vital organs. They also provide a place for muscles, supporting structures to attach. Bones are also sites for mineral storage and blood cells formation.

The face is composed of five main bones as shown on Figure 1-1 and support entry of the digestive and the respiratory systems. They also provide connection for muscles of facial expression and mastication, and they support organs for the main senses.

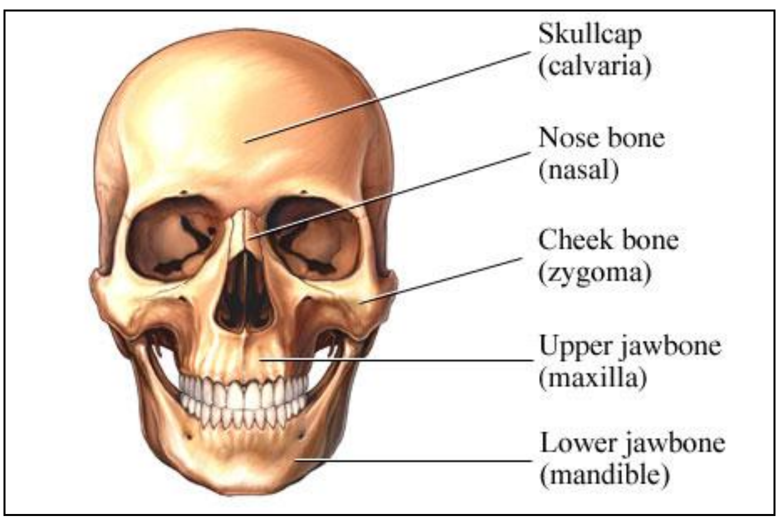


Figure 1-1: Main facial bones[1]

[1] Image copyright 2000 by Nucleus Communications, Inc. All rights reserved. http://www.nucleusinc.com

The relationship between facial bones and visual appearance of the face is not direct. For example, the shape of the eyes and eyelids, the tip of the nose and the lips cannot be predicted from the skull. However, the forehead, margins of the eyes, cheekbones, bridge of the nose, chin and overall facial shape are closely related with skull contours (Evison 2001).

I.2. Muscles and fat

Facial muscles and fat are known in the literature as soft tissues. Their distribution smoothes the bones shape and gives the face its identity. Simpson and Henneberg (Simpson and Henneberg 2002) have explored the relationship of soft tissues with craniometric dimension on Caucasian cadavers (17 males and 23 females). “Significant correlations between many soft-tissue depths and craniometric dimensions were found, suggesting a relationship between the amount of soft tissue present on the face and the size of the underlying bony skeleton. Soft-tissue depths were highly positively correlated with each other; craniometric dimensions were correlated but to a lesser extent.”

The specific contribution of muscle to facial shape is not extensively discussed in the literature, but a detailed description of their anatomy and their contribution to facial expressions and wrinkling is available (Freilinger, Gruber et al. 1987; Waters 1987; Waters and Terzopoulos 1990). Facial muscles can be divided into two groups: cutaneous and bones muscles.

The front face is made of more than sixteen major bones muscles (Figure 1-2) that give life to facial expressions and mouth movements (speech and chewing). The results of muscles movements arise on the skin and can lead to permanent wrinkles known as expression wrinkles.

On the upper face, the Frontalis (Occipitofrontalis) arises from the top of the skull to the orbital area. Contraction of this muscle is responsible for the rise of the brow, and may induce forehead wrinkles. Located in the intercilliary zone, the Procerus draws the skin between the eyebrows and causes horizontal wrinkles in the root of the nose. The Orbicularis oculi are circulars muscles around the orbital cavity. The internal portion of the muscles closes and open the eyelids while the low external portion downwards the eyebrows. The external portion also causes hard occlusion of the eyes, producing crow’s feet wrinkles on the temples.

The middle face is moved by the major and minor Zygomaticus, and the Levator labii. When contracted, the minor Zygomaticus elevates the portion of the lip where it is inserted while the major Zygomaticus lifts upward and laterally the corners of the mouth (smile). The Levator labii raises the later edge of the wing of the nose and a portion of the upper lip. It action leads to the apparition of the nasolabial folds.

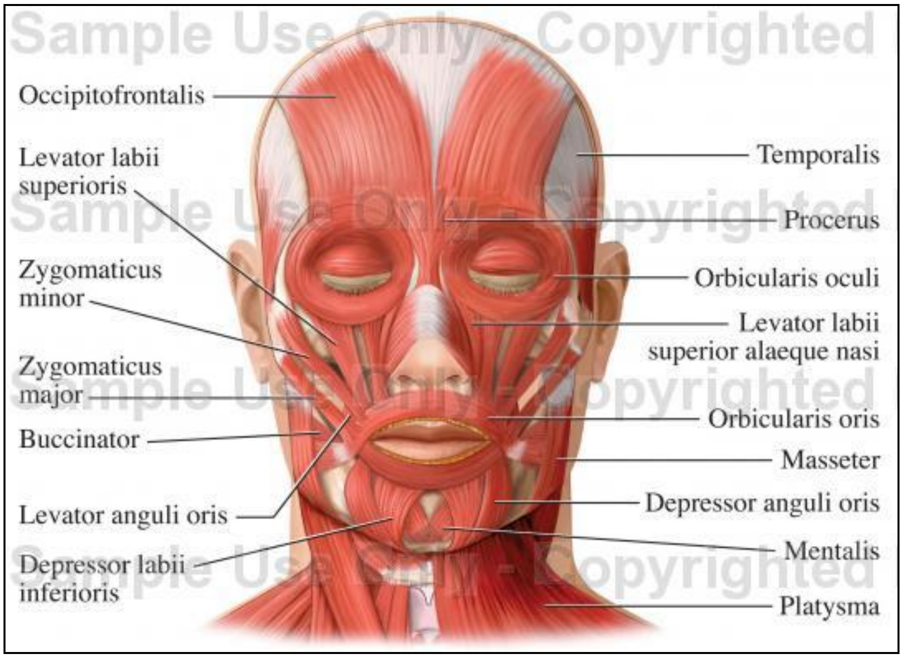


Figure 1-2: Facial muscles[2]

The movement of the mouth is also driven by the Buccinator, the Orbicularis oris, the masseter, the Depressor labii and the Levator anguli oris. Their coordination contributes to the mimics of the mouth and the mastication movements.

Some cutaneous muscles are part of the skin. The skin merged with the cutaneous muscles and fat - usually referred in the literature as the superficial musculoaponeurotic system (SMAS) – have an impact on apparent age and are manipulated through dermato-cosmetical procedures with a great benefit on apparent age (Barbara A. Gilchrest 2006).

I.3. Skin

The skin is the most voluminous body organ. Its main function is to be a protective barrier that interfaces against hostile environment including microbes and bacteria Facial skin in particular is exposed to many external aggressive factors such as sun and atmospheric pollution. Consequently, the facial skin aging will results from both chronological and extrinsic aging. To be able to understand changes occurring on the skin with age, one need to firstly describe its three layers organization (Figure 1-3).

[2] Image copyright 2000 by Nucleus Communications, Inc. All rights reserved. http://www.nucleusinc.com

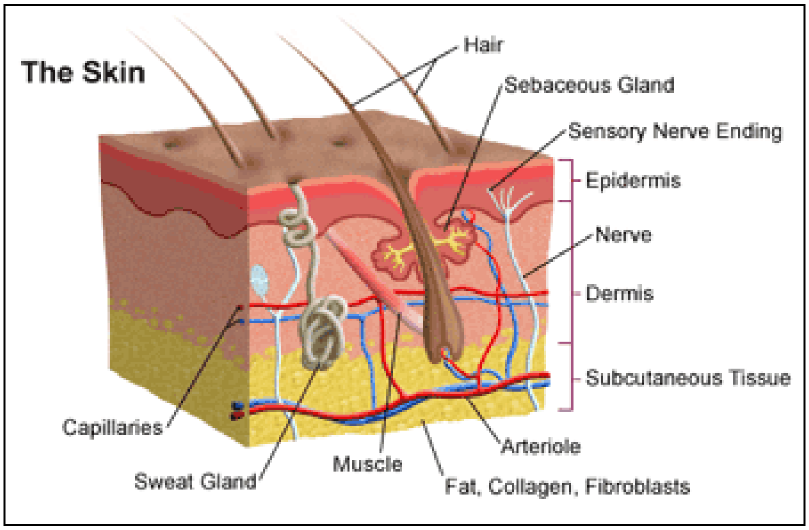


Figure 1-3: skin anatomy

The **epidermis** is the outer layer of the skin containing five sub-layers made with keratinocytes at different level of differentiation. Keratinocytes are produced in the basal or generative layer and start migrating to the upper layer while differentiating (Flattening and increasing surface). The top layer of the epidermis, the stratum corneum is made with flat non-nucleus and dead skin cells. The epidermis also contains melanocytes in its basal layer, which produce melanin. Melanin absorbs radiant energy from the sun and protects skin from ultraviolet radiations. The ratio of melanocytes and keratinocytes is approximately 1:4 in sun exposed area.

The primary function of the **dermis** is to support and sustain the epidermis. Apart from skin appendages and blood vessels the dermis is almost pure collagen. Collagen comprises 70% to 80% of the dry weight of human skin. Collagen fibers provide tensile strength and resistance to shear and other mechanical forces. Elastin fibers also contribute to the mechanical properties of the dermis. Even if they only represent 1% of the weight of the dermis, they resist to deformational forces and give the skin its elasticity. The cutaneous vessels located in the dermis are arranged in a horizontal network parallel to the surface of the skin.

The **subcutaneous tissue** is a layer made of fat and connective tissues containing larger blood vessels and nerves. This layer is important for the regulation of the skin and the body temperature.

Facial skin is clinically characterized mainly by its wrinkles and its color.

Skin wrinkles

Wrinkles are folds of the skin frequent on the face of aged people. They are caused by the changes in the thickness of the epidermis and structural transformation in the dermis related to aging. According to their aspect, there are three major groups of wrinkles (Kligman, Zheng et al. 1985):

- Crinkling-type wrinkles are fine wrinkles that formed from folded skin. They usually appear in old skin (person more than 75 years old) and may be caused by sun-damage on individuals with elastosis.
- The glyphic wrinkles have a crisscross pattern and are frequently seen on the cheek and the neck.
- The deep wrinkles form long and straight major lines or deep grooves. They usually appear on the forehead and the crow's feet area.

Skin color

Color is an essential component of physical properties of the skin (Andreassi 1995). Skin color arises from the interactions of light with the epidermis and the dermis (Stamatas and Kollias 2004). The perceived color also depends on the detector (eyes, camera, film…).

The interaction of light with the skin is related to the concentration of chromophores and the absorption coefficient of the constitutive layers of the skin. With an index of refraction of 1.55, the stratum corneum reflects approximately 5% of an incident perpendicular light (Kollias 1995). Most of the incident light crosses over the 10 µm of the stratum corneum and reaches the viable epidermis. In this area, the major chromophore is melanin. Melanin absorbs light stronger in the blue than in the red, resulting in a color that is rich in red and poor in blue (brown color) (Kollias 1995) .The absorbance spectrum of melanin shows a monotonic decrease towards longer wavelengths. Under the epidermis, the dermis provides the skin with nourishment through blood vessels. The hemoglobin found in the dermal blood is the chromophore responsible for the red appearance of the skin. The type of hemoglobin that carries oxygen molecules is call oxy-hemoglobin (oxy-Hb) while hemoglobin without oxygen is called deoxy-hemoglobin (deoxy-Hb). Each type of hemoglobin has its how absorbance spectrum.

It follows that the visual perception of skin color is the result of the contributions of three chromophores, melanin, oxy-Hb and deoxy-Hb. However, the skin color can hardly be linked with the concentration of one specific chromophore. As an example, Stamatas and Kollias (Stamatas and Kollias 2004)} have shown that skin darkening is not only linked with the concentration of melanin, but also depends on deoxy-Hb.

Another mode of interaction of light with the skin is scattering. Scattering occurs when light changes its trajectory after crossing over two different optical media. The principal component of scattering in the dermis is the collagen. Light scattering is also a function of wavelength. Blue light is more scattered than yellow, which penetrates deeper into the skin.

II. Changes with age

Several authors have described the overall changes on face with age. This section is based on the overall descriptions made by Friedman (Friedman 2005), Vacher (Vacher 2004), Taister & All (Taister, Holliday et al. 2000), Gilchrest & Krutmann (Barbara A. Gilchrest

2006) and Kanior & Kerth (Konior and Kerth 1990). The changes described involve those that take place from the twenties up to the eighties.

II.1. Changes on shape

Changes on facial shape with age are related to bones, muscles, fat and skin transformations and are tightly linked to body weight and gravity.

II.1.1. Bone transformation

Bartlett & All have explored age-related changes of the craniofacial skeleton of a Caucasian population (Bartlett, Grossman et al. 1992). They performed anthropometric measurements on 160 skulls selected randomly from 1500 specimens. Observed changes in craniofacial morphology included:

i. Appreciable reduction of facial height,
ii. Most marked in the maxilla and mandible, strongly correlated with loss of teeth,
iii. Modest increase in facial width,
iv. Modest increase in facial depth, except in those regions associated with tooth loss.

Vacher (Vacher 2004) considered the osteoporosis and the lost of teeth as the principal skull transformation linked with age. The osteoporosis is caused by the decalcification that occurs in all bones of the body and is even accentuated by the lost of teeth. Carlsson and Persson (Carlsson and Persson 1967) have shown that the front side of the mandible decreases by 2 mm two months after a tooth extraction, 4 mm after one year and approximately 7 mm after 5 years (Figure 1-4). The mandible resorption is more important for females than for males, meaning that it could be related to osteoporosis (Vacher 2004). The lost of teeth also affects the maxilla, but with a resorption of only 0.1 mm per year. Finally, bones changes mostly affect the lower part of the face.

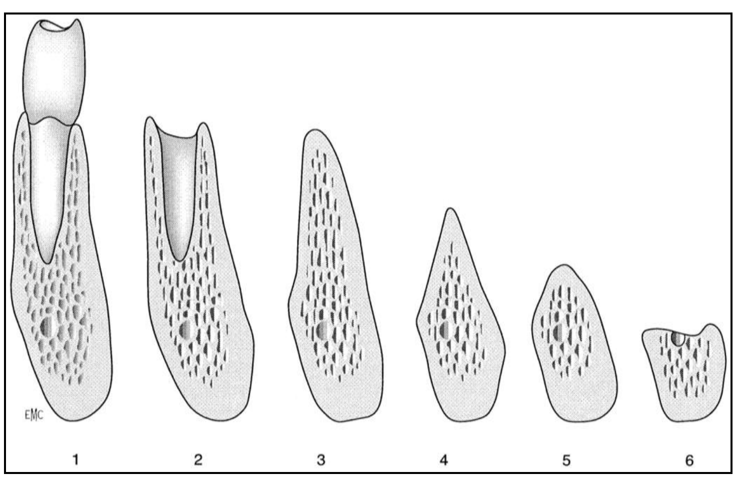


Figure 1-4: Resorption of the mandible after the extraction of a tooth (Atwood 1971)

II.1.2. Muscles and fat transformation

With increasing age, a degenerative myopathy of all the body muscles takes place due to the reduction of the number and the size of the muscular fibrous and vascular vessels. This leads to a reduction of about 30% of the muscle mass between 30 and 80 years of age (Vacher 2004). As an example, the frontalis loses its straight resulting in a ptosis of the eyebrow. Orbicularis muscles hypertrophy is responsible of the apparition of the orbital rims (Friedman 2005). Facial muscles are also affected by the modifications of the skeleton and facial fat. In most cases, a fat infiltration occurs in the skin muscles causing the nasal tip ptosis.

Fat also redistributes making the face looking more concave, with a hill and valley topography. Donofrio (Donofrio 2000) has pointed out both a fat atrophy in some areas of the face and a fat hypertrophy on other areas. Fat atrophy occurs in the forehead, periorbital, temporal, perioral and buccal areas while fat hypertrophy happens in the jowl, lateral nasolabial fold, lateral labiomental crease and lateral malar areas (Figure 1-5). These changes lead to old faces being more compartmented, with broken, wavy or concave shapes. Arcs, like the cheek arc in profile view or the mandible arc are converted to straight lines, living behind a relative excess of skin. Fat descents along the folds like nasolabial fold. Nasolabial fold can occur at all ages, but like the expression wrinkles, it increases with age (Yousif, Gosain et al. 1994).

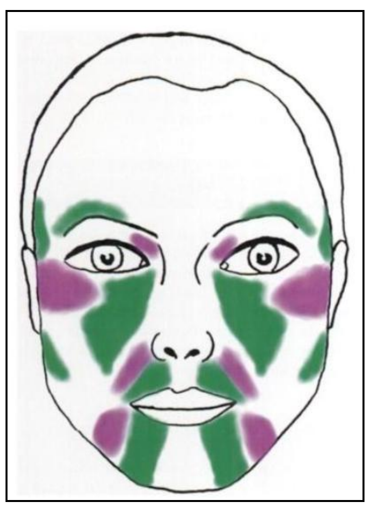


Figure 1-5: Fat atrophy (green) and accumulation (purple) (Sclafani and Romo 3rd 2000)

II.1.3. Skin transformation

Many skin shape transformations are the results of the changes in the deepest layers as previously described. The resorption of mandibular and maxillary bones and the fat absorption in the cheek and the chin area combined with the lost of the skin elasticity lead to chin ptosis. The nasolabial fold also becomes deeper with age as the combination result of facial movement and fat redistribution.

The most important change in skin properties that will affect facial shape appears in skin firmness and elasticity. Collagen and elastin are the main components of the connective tissues of the skin, responsible for its firmness and their degradation leads to stiffness, flacity and wrinkling skin.

Facial skin collagen is affected by both chronological aging and photoaging. With sun exposure, the amount of mature collagen (insoluble) has been reported to decrease, meanwhile the amount of soluble collagen increases (Kligman 1989). The overall density of the collagen in the dermis decreases and the fibrous network becomes coarser. This progressive disorder in the fibrous network may account in the loss of the skin tensile strength (Kligman and Balin 1989).

The elastic fibers are even more sensitive to photoaging. They degenerate and produce an amorphous mass that would never appear in unexposed skin. The fibers underlying reticular dermis become more numerous, thicker and disorganized (Kligman and Balin 1989). These transformations are clinically resulting in a thicker, more leathery, wrinkled and inelastic skin.

The jaw line and the eyes area are progressively affected by skin inelasticity. The accumulation of fat combined with the mandible resorption and the lost of skin elasticity leads to a chin ptosis and a weak mandibular line (Friedman 2005). On the eyes area, the effects of eyelids' sagging is combined with the hypertrophy of periorbital muscles leading to a reduction of eyes opening (Friedman 2005).

II.2. Changes on texture

Wrinkling is the principal change in skin texture with age. Facial wrinkles are related to two concomitants facts: skin aging and facial expression.

Both extrinsic and intrinsic aging lead to transformation in the thickness of the epidermis and the structure of the dermis as described in the previous section. The epidermal thickness is particularly small in the area of wrinkles and the dermoepidermal junction is flatter than in non-wrinkling regions. The reduction of the density of the dermis combined with the disorganization of the elastics fibers is also responsible of the appearance and the aggravation of skin wrinkles (Blog 1999).

Expression wrinkles arise from the repetitive movements of the skin caused by facial expressions. They tend to form perpendicular lines to the long axis of the muscles they appear on. When these muscles are circular, like around the eyes and the mouth, the wrinkle lines are perpendicular to a tangent drawn to the curve area (Friedman 2005).

II.3. Changes on color

The age related changes on skin color is caused by the modifications in the distribution of chromophores and by the transformation in the dermis structure. These changes are caused on sun exposed sites by both chronological aging and photo aging.

Ortonne has shown that the number of active melanocytes decreases about 10-20% per decade (Ortonne 1990) on un-exposed skin leading to an accentuation of pigmentation in older skin. The mechanisms behind these augmentation of melanocytes where explored by Okazaki et al. using ELISA analysis (Okazaki, Yoshimura et al. 2005).

Nishimori et al. (Nishimori, Pearse et al. 1998) measured the skin color parameters (L*,a*,b*) of 27 volunteers (aged between 18 and 61 years) using a spectrophotometer. They found that the coefficient [b*-5a*] was correlated both with the solar elastosis and with age. Their result suggests that solar elastosis accounts for the yellowish of the skin with age. The sun was also reported as the cause of many pigmentation disorders including lentigo and brown spots (Ortonne 1990). All these transformations lead to a dull uneven skin tone.

III. Factors influencing facial aging

III.1. Ethnicity

Some studies have been done to understand the similarities and differences in facial attributes as a function of age and ethnicity. It is a difficult task to evaluate the influence of ethnicity comparatively to climatic condition, pollution and lifestyle. Due to the wide variability within same age groups, a large number of subjects is required in order to record noteworthy changes in facial skin attributes with age.

Nouveau-Richard & al. (Nouveau-Richard, Yang et al. 2005) have compared a group of 160 Chinese women between 20 and 60 years of age with 160 French women within the same age range. Using clinical assessment, they found that wrinkles start later for the Chinese than for the French women. On the contrary, Chinese's women skins were more marked by brown spots comparatively to French women. The difference in terms of wrinkling was explained physiologically or climatically; an ethnic group-variation in melanosomes was reported as a possible explanation of the disparities in brown spot prevalence. Still, this study involves two groups of people leaving in different continents with different latitudes and lifestyle, the differences observed may also be linked to dissimilar exposure to the sun or diet.

Hillebrand & al. (Hillebrand, Levine et al. 2001) evaluated age-related changes from 3160 women spanning Caucasians, Asian Indians, African Americans, Latinos and East Asians;

44% of their study population living in Los Angeles and the rest living either in Europe (Rome and London) or in Japan (Akita). They measured wrinkling, pigmentation spots and pores from images acquired under controlled conditions. They also evaluated hydration with a corneometer, color with a chromameter, sebum excretion with a sebumeter and dryness with D-Squames analysis. When comparing people living in Los Angeles, independently of the age group, it was found that:

- Caucasian women are more wrinkled than all other ethnic groups, while East Asians were the least wrinkled
- African American women have more hyperpigmented spots than the other ethnic groups, followed by Caucasians.
- African Americans have two times more visible pores and more sebum excretion than the other ethnic groups
- Caucasian women are less hydrated in the face than the other groups, but not on the forearm and legs
- There was no significant difference between ethnic groups for skin flakiness.

This study involved a wide range of volunteers living in a same area. The measurements were done by trained experts to minimize acquisition bias. Unfortunately, it was not possible to collect all the data in a few period of time and the study span from October 1999 to March 2000 meaning big climatic differences between the two time points. Most of the methods used are well-known and have been validated and published. However to our knowledge, the image analysis algorithms used to measure wrinkling, pores and hyperpigmentation spots have not been published. Skin feature segmentation remains a difficult task, particularly when dealing with different skin types. The validity of the findings related to image analysis need to be confirmed.

III.2. Gender

The main differences between men and women aging is related with the specificity of their hormonal balances.

With the menopause, the production of ostrogen falls, accentuating women aging. At the bonny level, the demineralization is emphasized leading to the transformation of the maxilla and the mandibular (Friedman 2005). The skin is also affected and the production of collagen decreases, making the skin thinner, less elastic and more wrinkled. The reduction of fibroblast activity produces a dehydration of the dermis. There is also an epidermal atrophy caused by the reduction of the activity of keratinocytes. The secretion of the sebum also decreases, causing a less effective barrier of the stratum corneum.

III.3. Environment and lifestyle

Together with the hands, the face is the main exposed body part to environmental factors. Facial appearance is affected by sun exposure and atmospheric pollution but also by long-term moods and habits. Some old faces are described as happy and look younger while others are described as sad faces. It is a difficult task to quantify the influence of long-term happiness to the aging process since we cannot precisely assess ones' overall mood. However Rexbye & al. (Rexbye, Petersen et al. 2006) have explored the influence of lifestyle parameters such as social class, marital status, depression symptology score and number of children on the aging process. They included 1826 elderly twins who had a high-quality image taken for their study. They asked ten nurses to assess the visual age of their volunteers from the face images and the apparent age was defined as the average of the nurses' estimates. Using multivariate regression, they evaluated the influence of lifestyle and health parameters such as BMI (Body Mass Index), and cardiovascular and bronchitis diseases. They found that BMI, number of children for men, marital status and depression were significantly associated with facial aging. However, the main factors influencing facial aging where smoking and sun exposure.

The effect of smoking cigarettes on facial skin has been discussed in many papers (Lopez Hernandez, Tercedor et al. 1995; O'Hare, Fleischer et al. 1999; Leung and Harvey 2002; Trueb 2003; Freiman, Bird et al. 2004; Raitio, Kontinen et al. 2004; Just, Monso et al. 2005; Helfrich, Yu et al. 2007; Just, Ribera et al. 2007). It was known since 1971(Daniell 1971) that smoking increases the skin wrinkling, and the physiological mechanisms underlying this phenomenon were located in the dermis. Smoking increases the amount of amorphous elastotic material; this increase is caused by the degradation of elastic fibers as it also occurs in solar elastosis (Just, Ribera et al. 2007).

IV. Conclusion

From the twenties to the eighties, the facial anatomy is subjects to many changes both in the deepest structures like bones than in the superficial ones. These transformations visibly alter the shape, the texture and the color of faces as follow:

- In the lower part of the face, the jaw line shape move from oval to square due to the chin ptosis and the skin sagging. The lips become thinner and less contrasted
- In the middle part, the nasolabial folds become deeper, the cheek cave-in and a nasal tip ptosis often happens. Bags appear under the eyes and the upper eyelid start sagging, leading to smaller eyes.
- In the upper part of the faces, there are less noticeable shape changes. Wrinkles are the most visible sign of aging in this area of the face.
- Wrinkles progressively appear, both because of facial expression than because of alteration in the dermal layer of the skin.
- The distribution of chromophores in the face is affected by the sun exposure, leading to an uneven tone. The yellowish of the face with age is related to the elastosis increases on sun-exposed skin.

Even if age-related changes occur on everyone, there are some specific characteristics that are related to the gender or ethnic groups. In the next chapter, we specifically focus Caucasian women.

Chapter 2
Changes on Caucasian women's face with age

Women's aging is an important concern for cosmetic dermatologists, aesthetician surgeons and cosmetics company. In order to design products and procedures that better target the need for more efficient treatments, it is necessary to quantitatively evaluate the changes that happen on face with time. Particularly when these changes happen on the skin, they required adequate methods to be documented. Cutaneous metrology is the science that provides tools and methods for skin measurements.

In this chapter, we describe a study that we performed in the Skin Care Research Institute (SCRI), the cutaneous metrology laboratory of Johnson & Johnson. The study was done in order to better understand the changes that occur on healthy Caucasian women living in Paris and its suburbs.

The most efficient way to evaluate peoples' changes with age is to follow a cohort for decades and to record individual transformations as a function of time. Since this methodology is time and resource consuming, an alternative methodology was chosen: it was decided to include a large number of volunteers, with their age spanning over the whole range of interest (20-74 years). The size of the population was chosen so that a representative number of subjects for each group of ten years of age was available. Skin attributes were subsequently measured and regression models were built to link attributes with age. Data were collected over a short period of time (July 2004).

In this chapter, we describe the measured changes on Caucasian women. Firstly, we present the method used to collect and to analyze the data. Then we depict the transformation on faces, grouping them into shape, texture and color changes. Finally, we critically discuss the method we used during this experiment and propose some improvements.

I. Population and method

I.1. Population

The study followed the principles of the Helsinki convention. A random sample of 173 subjects from 20 to 74 years of age was taken from a database of 5000 healthy Caucasian women living in Paris and its suburbs and volunteers for evaluation of cosmetic products. They were aware of the goal of the study and signed a consent form. The main exclusion criteria included any surgical or non-surgical procedure on the face, long-term use of anti-aging cosmetic products or treatments, and previous facial exposure to sun tanning lamps. Intrinsic and extrinsic factors known to be linked with age such as heredity, life style, working condition, smoking and sun exposure were not considered as inclusion criteria, assuming that randomization reduced these biases.

The recruitment time limit does not allow having a perfectly uniform distribution. However, the number of volunteers per group of ten years was between 28 and 40 as shown in Figure 2-1.

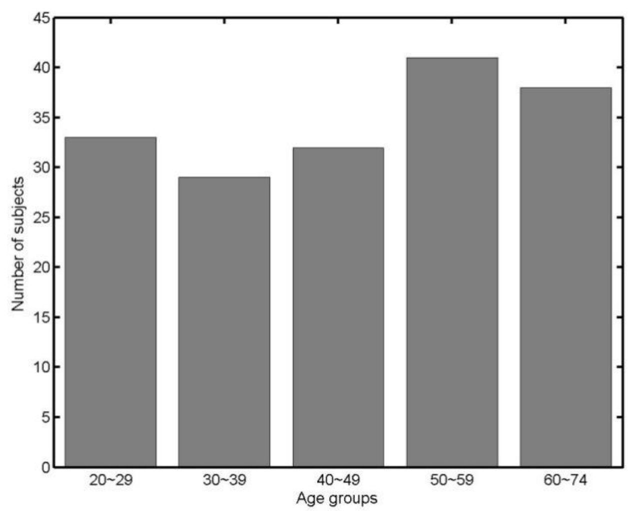


Figure 2-1: Distribution of subjects per age groups

I.2. Evaluation of skin attributes

Evaluation of skin attributes is commonly performed by dermatologists to assess diseased skin. With the development of skin metrology, new scientific methods are being developed to evaluate the properties and the aspects of healthy skins. These methods involved objective and subjective evaluations.

- Objective evaluation relates to procedures carried out by trained persons using either instrumental or sensory methods that produce reproducible and quantifiable results.

- Subjective evaluation relates to procedures involving sensory tests carried out by untrained people such as self-assessments.

In this study, skin attributes were objectively assessed under controlled conditions (room temperature, relative humidity, and light) after 15 minutes of acclimatization. Several attributes known as being linked to aging were measured using a Visual Analog Scale (Dexter and Chestnut 1995) (VAS). These attributes were chosen to evaluate the changes on skin that are influenced by both chronological aging and photoaging such as wrinkling (Fedok 1996), sagging (Fedok 1996), color (Ortonne 1990; Burt and Perrett 1995) and texture (Guinot, Malvy et al. 2002). Facial pictures were also taken under controlled positioning and the skin color was measured using a colorimeter.

I.2.1. Visual analog scale

Scaling and ranking methods are widely used in industry as quality control procedures, for product development, and in research. The methods, in their many varied forms, are almost certainly the most crucial step toward the practical application of sensory measurement techniques. The main advantages why scaling and ranking methods are preferred to even more accurate instrumental methods are:

1. They require much less time
2. They have wide range of applications
3. They can be used with large numbers of stimuli

"A scale has been defined as a graded arrangement, used in reporting assessments: it is divided into successive values, which may be graphic, descriptive or numerical. The successive values may be implicit, as in a graphic analogue scale, and the assessments are usually of magnitude (…) The scale is the instrument used by the assessors to make explicit their perceptions." (Land and Shepherd 1988) Scales are usually divided into four categories (Lawless and Malone 1986):

- Visual analog scale (VAS) also referred to as line or unstructured scales: The grader uses a continuous line with only two anchors at its extremities to run the measurement. Usually, the assessor marks a line at the point he or she feels appropriate and the distance along the line is measured. The response is assumed continuous and therefore the reported data, meaning that powerful statistics methods could be used for the analysis.
- Linear discrete scale (or machine scale): As the previous described scale, they are linear with only two anchors. But these scales are not continuous, most frequently because they are designed for digital devices that cannot handle analog input.
- Category scale: The grader is requested to classify the given subject into one of the proposed ordinal or nominal categories. For ordinal category scale, the categories are chosen to be equally separated. Usually adjectives such as 'none', 'weak', 'moderate' 'strong' and 'very strong' are used to build these scales. "The general principle involved in these category scales is that the adjectives are those commonly used in everyday life to describe perceived attributes, and the use of them has

therefore stabilized, i.e. people use them consistently." (Land and Shepherd 1988). If it is not the case, samples can be provided to the grader from each category so that he has a reference for comparison. The numbers of categories to choose when building the scale has been extensively discussed (Dunn-Rankin, Knezek et al. 2004) and there is not yet an agreement on the methodology to use to fix it.

- Magnitude estimation scale: A reference is shown to the grader who assigns a grade arbitrary. This reference will be used comparatively to evaluate the other subjects proportionally to the grade assigned to the reference.

The sensitivity and the reproducibility of these different types of scales have been extensively debated (Lawless and Malone 1986; Lawless and Malone 1986; Dunn-Rankin, Knezek et al. 2004; SeonYoung, O'Mahony et al. 2004), with no definitive answer. Category and line scales were reported slightly better than the other methods by Lawless and Malone (Lawless and Malone 1986). SeonYoung (SeonYoung, O'Mahony et al. 2004) did not see any difference between category and line scales in terms of data comparison. There may be a distinction when analyzing the data from these two types of scale. VAS are assumed to be continuous and linear, meaning that the data can be compared using parametric statistics. For category scale, non-parametric test would be required except if the number of subjects included or the number of grades in the scale is high enough. However in many publications, categorical data are analyzed as continuous.

Scales are commonly used in medicine. A review of literature in PUBMED shows an increase number of papers that refers to this methodology since early 70s. In dermatological researches particularly, scales have been used to measure skin wrinkles (Lemperle, Holmes et al. 2001; Day, Littler et al. 2004; Permissions and Surgery 2006), dryness (Serup 1995), color (Serup and Agner 1990), acne (Doshi, Zaheer et al. 1997), eczema (Charman and Williams 2000)…

In the present study, VAS were used and preferred to the other objective measurement methods for the following reasons:

1. VAS are widely used by dermatologists and researchers to assess the efficacy of treatments. Results from VAS are meaningful for these two communities
2. Clinical observations enable us to describe what is observable from a human point of view. This is particularly useful when examining which facial attributes drive the overall perception of aging (next chapter)
3. A trained grader can assess many parameters that are not measurable otherwise (Sloping upper eyelid, radiance, …)
4. Many measurements can be performed over a short period of time, meaning that we can follow many more parameters than when using devices
5. VAS enable more powerful statistics since the recorded data are assumed to be continuous.

Parameter assessed using VAS

We have assessed the severity of 21 attributes summarized in Table 2-1. The measurements were made by a trained technician who received a literate description of the extrema (minimum and maximum) of each VAS and used a slider to give a score during the assessments. Table 2-1: Parameters assessed using a Visual Analog Scale summarized the parameters measured, their meaning and their extrema values.

Attribute	Description	Minimum value (0)	Maximum value (12)
Crow's feet wrinkles	Number, length and depth of wrinkles	No wrinkles	Numerous and deep wrinkles at rest
Frown line	Number, length and depth of wrinkles	No wrinkles	Several and deep wrinkles at rest
Upper lip wrinkles	Number, length and depth of wrinkles	No wrinkles	numerous and deep wrinkles at rest
Cheek wrinkles	Number, length and depth of wrinkles	No wrinkles	numerous and deep wrinkles at rest
Nasolabial fold	Depth of the fold	Not visible at rest	Deep fold at rest
Brown spots	Number, size and color intensity of the spots	No brown spots	Numerous and large Dark brown spots all over the face
Skin color	The color of the skin in a scale going from pinky to yellowish	healthy glow (pinky) skin	Severe Sallowness (yellowish)
Forehead wrinkles	Number, length and depth of wrinkles	No wrinkles	Numerous and deep wrinkles at rest
Under the eyes wrinkles	Number, length and depth of wrinkles	No wrinkles	Numerous and Deep wrinkles at rest
Jaw line	Front shape of the jaw line	Oval shape	Severe sagging
Lips volume (volume)	The volume of the lips	Very Thin	Very Full
Crow's feet fine lines	Number, length and depth of fine lines	No fine lines	numerous fine lines covering the whole surface of crow's feet area
Bags under the eyes	Volume of the bags under the eyes	No bags	Very marked bags under eye
Eyes opening	Height of the eyes related with the upper eyelid sagging	Partly closed	Widely opened
Sloping upper eyelid	Covering of the upper eyelid by additional skin	No fold on the upper eyelid	Upper eyelid totally covered by additional skin
Lips definition	Color contrast between the vermilion and the surrounding skin	Blur border	Sharp border
Dark circles under the eyes	Color and surface of dark circles	No dark circles	Very marked circles under eye
Skin uniformity	Skin color evenness	Completely uniform	Non uniform all over the face
Skin texture	Smoothness of the skin related with pores and microrelief	Very Smooth	Very Rough
Radiance	Ability of the skin to scatter the light	Not radiant	Very radiant
Diffused redness	The color of the skin in a red scale	red	Not red

Table 2-1: Parameters assessed using a Visual Analog Scale

I.2.2. Facial pictures

In the past, documentation of skin conditions has been performed by dermatologists using imaging systems. Numerous atlases were built under the form of scientific material for educational purposes. With the development of digital cameras and the increase in digital computation power, storage and display efficiency, digital images have also been used as measurement tools to assess skin condition (Bittorf, Fartasch et al. 1997). In the last decades, digital images have also been used to build facial model of aging and to simulate age rejuvenation (Gandhi 2004).

In the present study, regular visible images are acquired in order to illustrate the changes measured clinically. The facial photographs are taken using a Nikon DS-505A with a resolution of 1280x1000 pixels. Photographs are taken under controlled lighting geometry and subject positioning. Two Lamps and a reflector panel are used to have diffused lighting falling from the top of the subject.

Pictures and clinical grading are referred as the JNJ database in the rest of the document.

I.2.3. Statistics

For each attribute measured using VAS, its relationship with age is studied. Subjects are divided in the following age groups: 20-29, 30-39, 40-49, 50-59 and 60-74. The mean and the standard deviation of attributes for each group are calculated and linear regressions are performed to measure the correlation between each attribute and age.

The facial images are used to calculate the average face for each group of age. Average faces are useful to capture the main features within the same age group. They are built by:

1. Manually selecting the main facial features (eyes, month, nose) from all the images in the group
2. Aligning the images using rigid registration so that the selected features overlap
3. Calculating the mean value of pixels at the same position in the whole image for each color channel.

Figure 2-2 represents the average face calculated for people between 20 and 29 years of age. The image is not as sharp as the original images since skin details among the different subjects do not match. However, the main features are sharp enough to preserve a correct estimation of the age of the "subject".

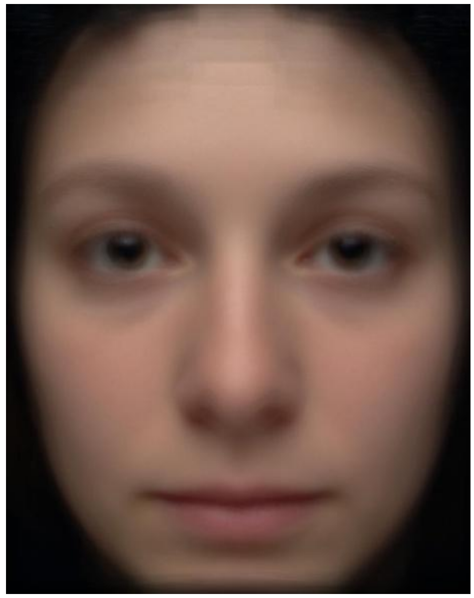


Figure 2-2: Example of average face

This face is calculated by averaging the pixels values of the 33 pictures of volunteers aged between 20 and 29. The main features (nose, eyes, mouth and jaw line) of the faces are aligned before the calculation of the average face.

II. Facial changes

The picture below (Figure 2-3) shows differences between a 20 year old woman and 60 year old woman. The attributes can be classified in terms of the characteristics (shape, texture and color) they described.

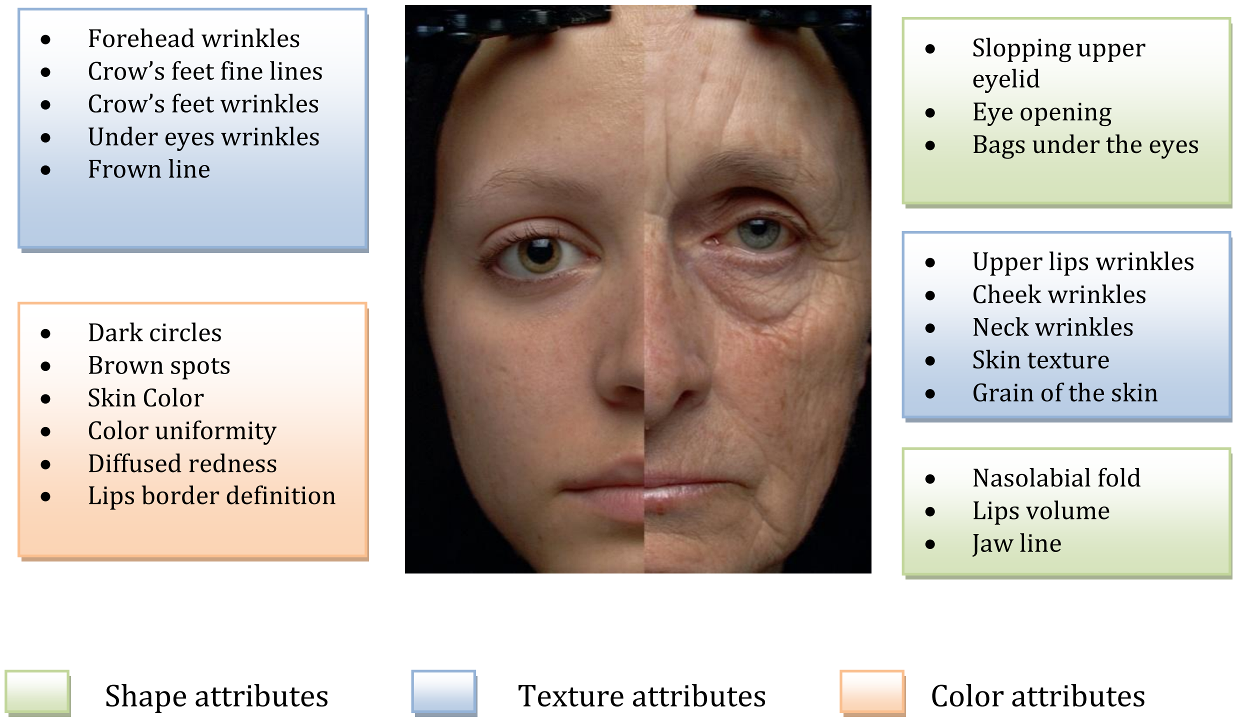


Figure 2-3: Difference between a young and an old woman face; attributes are assessed using VAS

II.1. Changes on shape

In order to capture overall changes on facial shape, average faces were made for each decade (Figure 2-4). Although small variations such as wrinkles or brown spots have disappeared with the averaging, shape changes are obvious. Therefore, the following can be noticed:

a) an increase of the nasolabial fold,
b) a reduction of lips volume,
c) a square shape of the jaw line,
d) a reduction of the size of the eyes,
e) an increase of bags under the eyes.

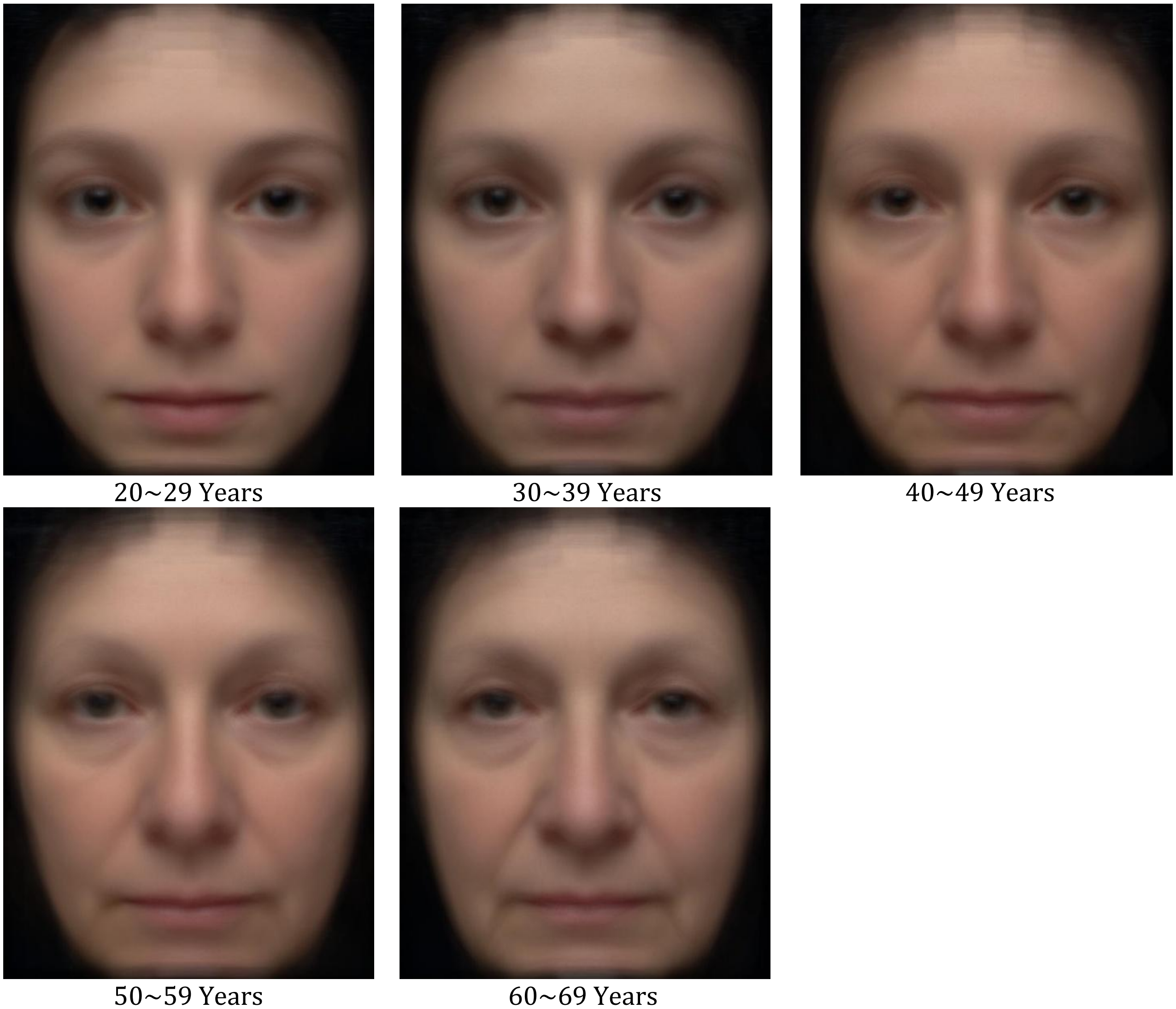
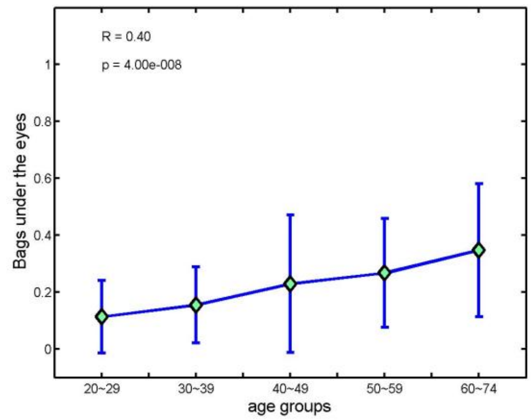


Figure 2-4: Average face per group of age

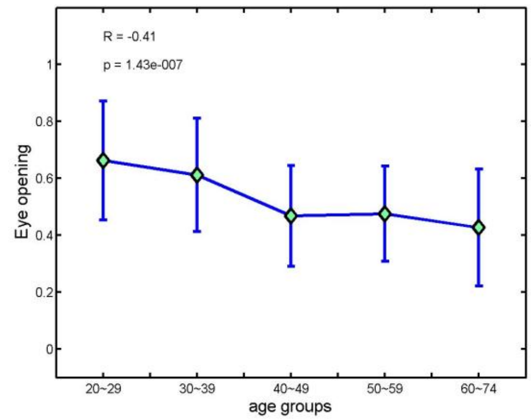
Each face is calculated by averaging the pixels values of the pictures of volunteers in the age range. The main features (nose, eyes, mouth and jaw line) of the faces are aligned before the calculation of the average face. These images capture the main changes related with age.

Bags under the eyes, eyes opening, slopping upper eyelids, nasolabial fold, lips volume and jaw line were assessed using VAS.

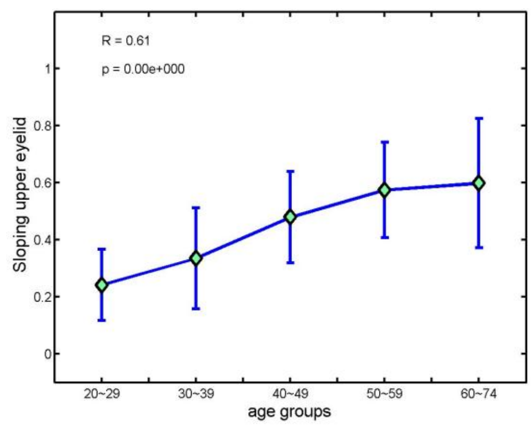
Figure 2-5 shows the relation between each attribute and age.



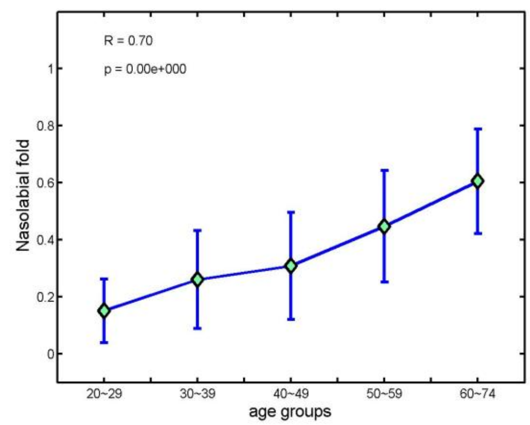
a) Bags under the eyes



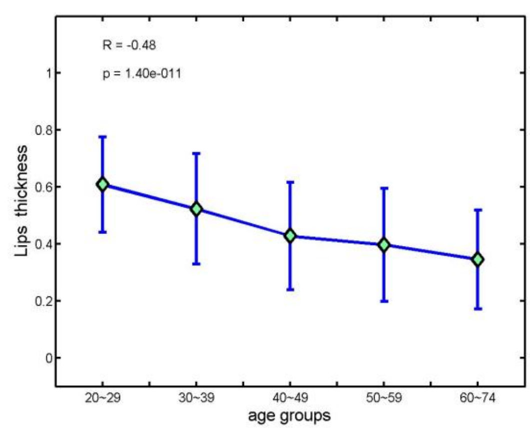
b) Eye opening



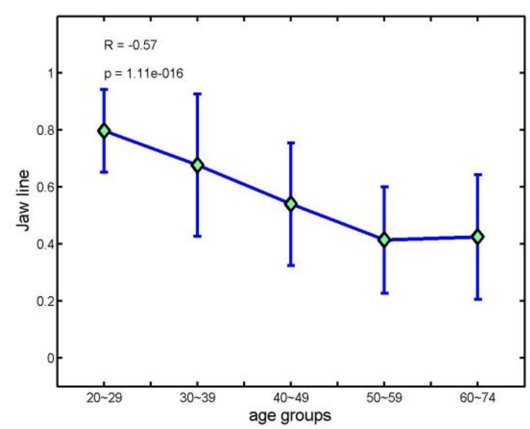
c) Sloping upper eyelid



d) Nasolabial fold



e) Lips volume



f) Jaw line

Figure 2-5: Changes on shape attributes with age

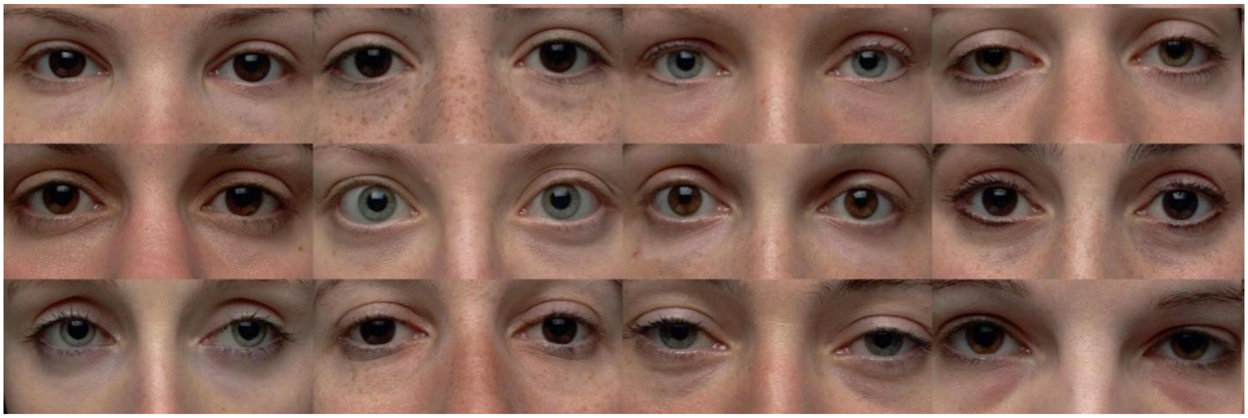
a) bags under the eyes, b) eye opening, c) sloping upper eyelid, d) nasolabial fold, e) lips volume, f) jaw line

The eyes area:

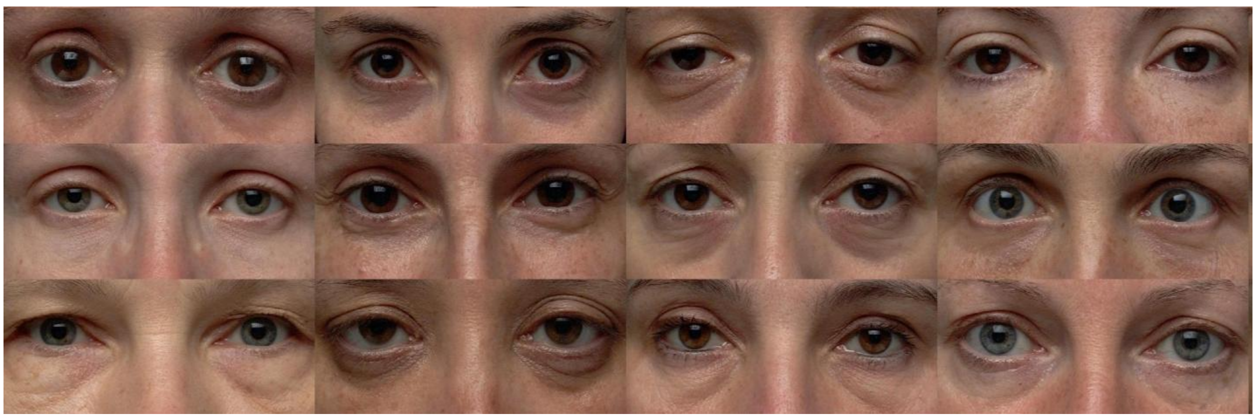
The area around the eyes ages as follows:

- The volume of the bags under the eyes linearly increases with age at least until 74 years old, but the slope of the raise is slow.
- The opening of the eyes significantly decreases until 40-50 years of age then reach a plateau.
- The sloping upper the eyelids linearly increases until 60 years of age.

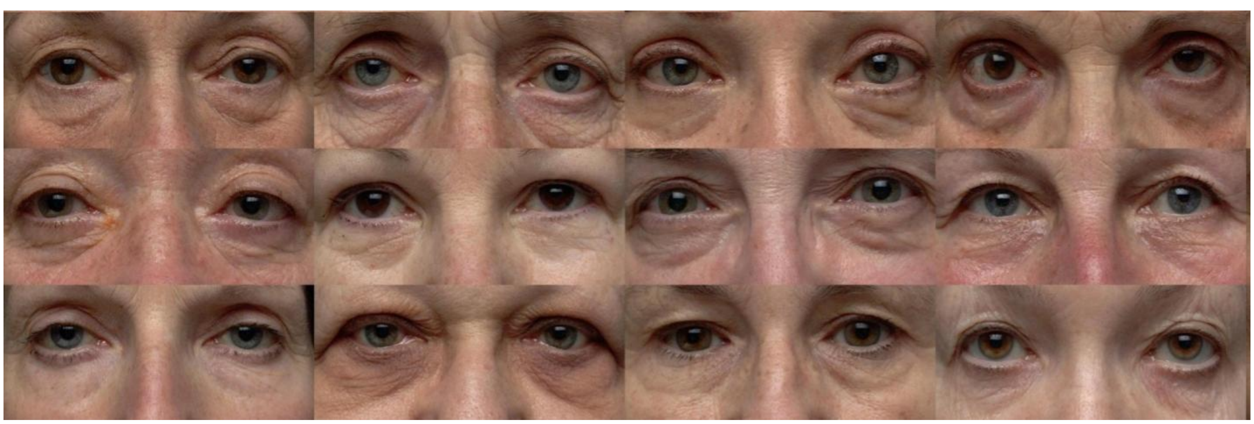
All these recorded changes are confirmed when observing the areas around the eyes of the facial images. Figure 2-6 presents the eyes of subjects randomly selected in three age groups: 20-29; 40-49; 60-74 years of age. The first group has their eyes widely open, with no sloping upper eyelids, and with small bags under the eyes for some people. In the second group, one can notice lost of firmness on the upper eyelid of most of the subjects, leading to smaller eyes opening. Bags under the eyes are more marked, in many more people. In the oldest groups, eyelid bags under the eyes are even more damaged.



a) Subjects between 20 and 29 years of age



b) Subjects between 40 and 49 years of age



c) Subjects between 60 and 74 years of age

Figure 2-6: Changes in the eyes area with age

The Lower face (Nasolabial fold, lips and jaw line)

The nasolabial fold becomes deeper with age. The increase particularly accelerates after 50 years of age. There is no significant difference between the groups 30-39 and 40-49 years.

The lip thickness slowly decreases with age, and the changes even become smaller after 50 years of age. In average, the changes are statistically visible after 20 years. The changes recorded by clinical grading are confirmed by imaging.

Figure 2-7 represents the average shape of faces per age group. Each shape is built using all the images from the age group. The images are manually landmarked and the average position of the landmarks is calculated. The resulted shapes show a reduction of the lips area, and particularly the lower lip. On the same figure, one can notice the transformation of the jaw line from an oval shape to a square one. Clinical grading shows that these transformations occur until 60 then reach a plateau.

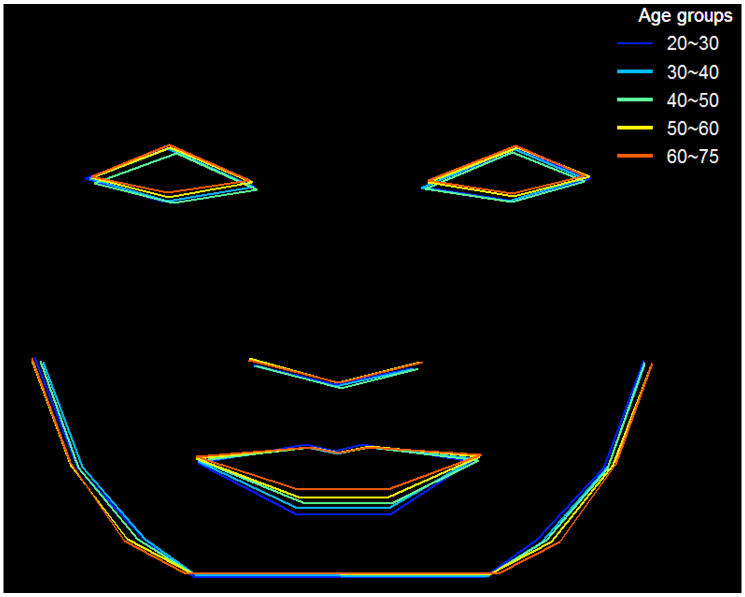
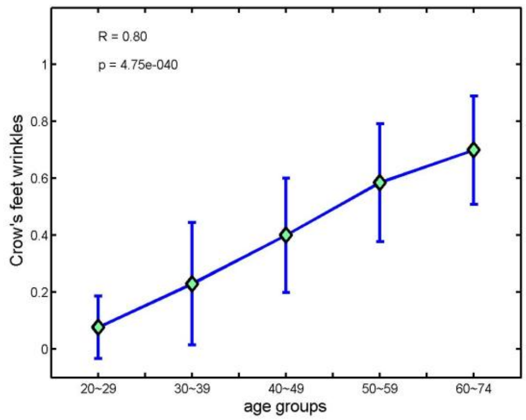


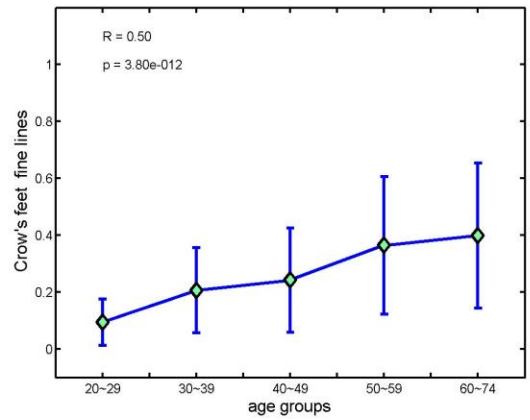
Figure 2-7: Changes of the mean facial shape with age

II.2. Changes on texture

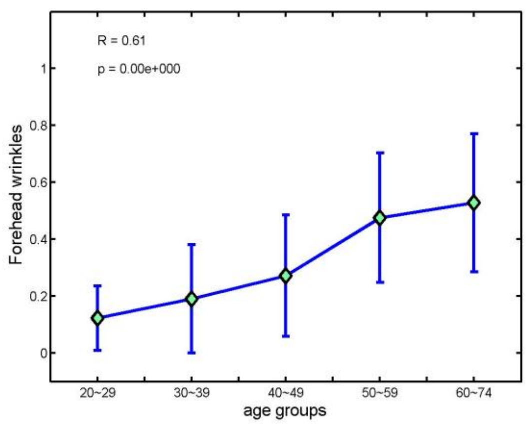
The textural parameters assessed are wrinkles, skin texture related to pores size, glyphics. Average faces do not capture these changes since they occur at a millimetric scale and cannot be aligned from one subject to another within a group of age. The Figure 2-8 displays the relation between each measured attribute and age.



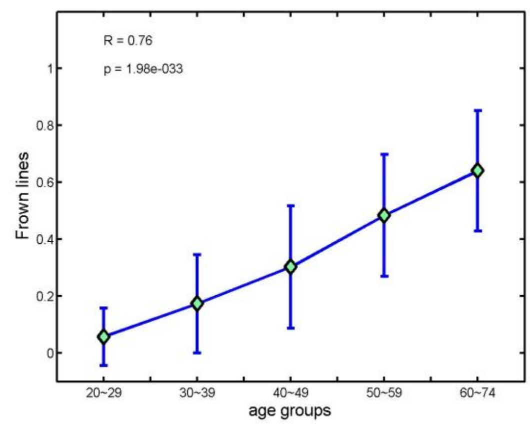
a) Crow's feet wrinkles



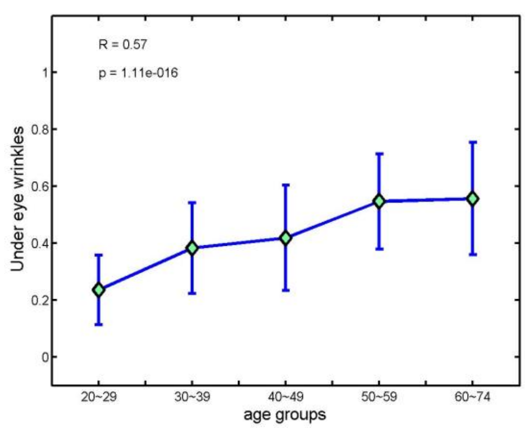
b) Crow's feet fine lines



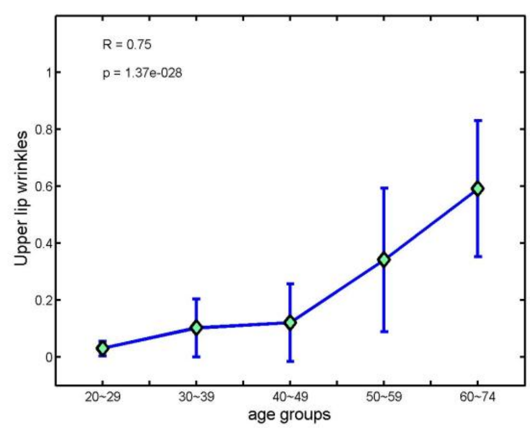
c) Forehead wrinkles



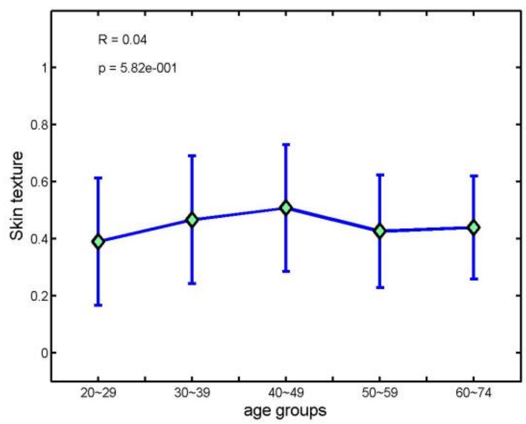
d) Frown line



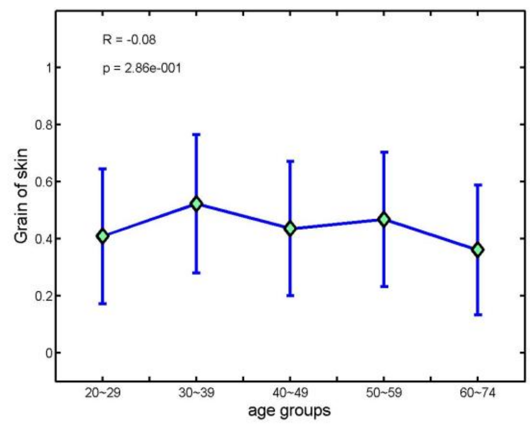
e) Under eye wrinkles



f) Upper lip wrinkles



g) Skin texture



h) Grain of the skin

Figure 2-8: Changes on skin textural attributes with age

a) Crow's feet wrinkles, b) crowfeet fine lines, c) forehead wrinkles, d) frown line, e) under eyes wrinkles, f) upper lip wrinkles, g) skin texture, h) grain of the skin.

Wrinkles:

Changes in the dermis structure, combined with facial expressions leads to wrinkles which become more numerous and deeper with age. VAS allows quantifying the following changes on facial wrinkles:

- The number and the depth of crow's feet wrinkles linearly increase with age.
- The crow's feet fine lines increase slower than the crow's feet wrinkle and reach a plateau at 60 years of age. Fine lines tend to increase fast in the early stage (30-40 years) and then after the menopause (50-60 years).
- The forehead wrinkles linearly increase with age.
- The depth of the frown lines linearly increases with age, each group being significantly different from the previous one.
- The under eye wrinkles slowly increases with two-plateau regions; the first one around 40 years and the second one above 60 years
- Upper lip wrinkles appear above 40 years and linearly increase with age

Texture:

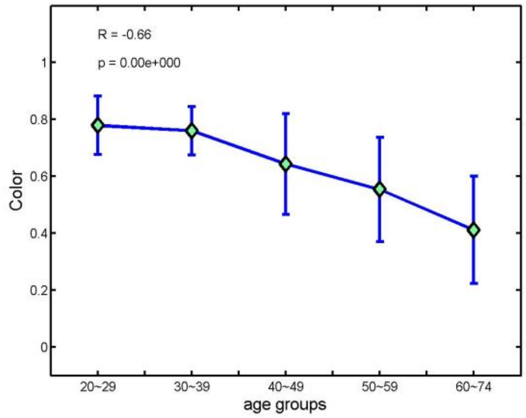
The correlation between facial skin texture (related to pores and glyphics) and age is not significant. The two measured attributes (skin texture and grain of the skin) are not related with age.

II.3. Changes in color

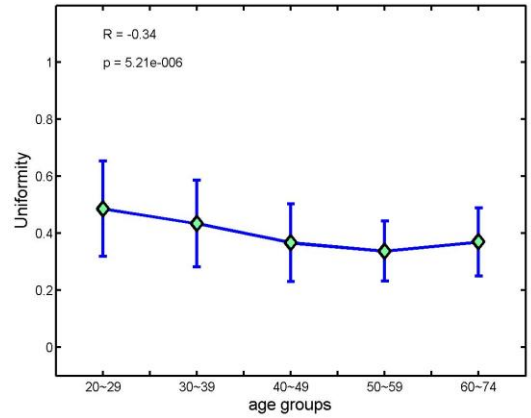
Old people's skin is described as uneven, yellowish and dull. These perceived changes are mainly related to photo-damage and pollution. The skin color attributes that were clinically measured for this study are:

- Color: A youthful skin is pinky while an older one looks yellowish. In this study, the color was graded between pink and yellow; the yellower the skin, the lower the given score.
- Color uniformity: The non-uniformity of skin in younger skin is mainly caused by acne. Older skins are affected by pigmentation disorders leading to spots.
- Brown spots: The number and the size of the brown spots should increase with age, as sun damage accumulates with time.
- Radiance: The radiance describes the skin ability to reflect light and is affected by pollution.
- Diffuse redness:
- Dark circles under the eyes: The skin under the eyes can be very thin, allowing the blood underneath the skin to become more visible.

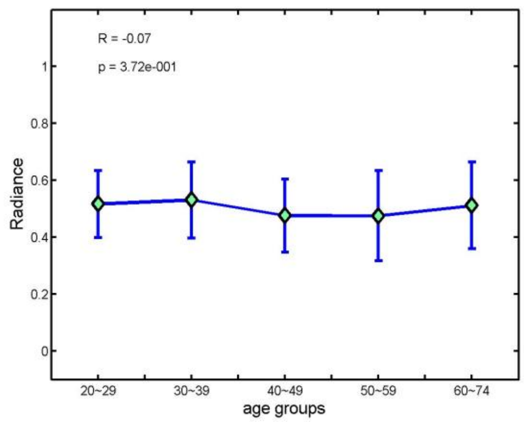
Figure 2-9 shows the changes on skin color attributes with age. As it can be seen, skin radiance and diffuse redness are not related to age. On the contrary, the skin color, the color uniformity and the brown spots are influenced by age. As expected, the color linearly changes from pink to yellow from 30 to 74 years of age. The skin uniformity slowly decreases between 20 and 50 years of age, while the number of the size of the brow spots significantly increases every 20 years. The dark circles slowly appear with age, but there is no significant difference between 20 and 50 years of age.



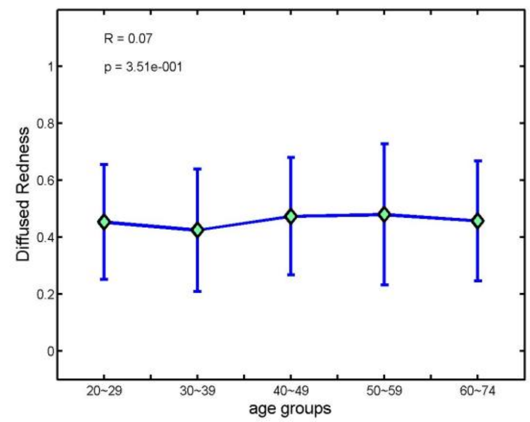
a) Color



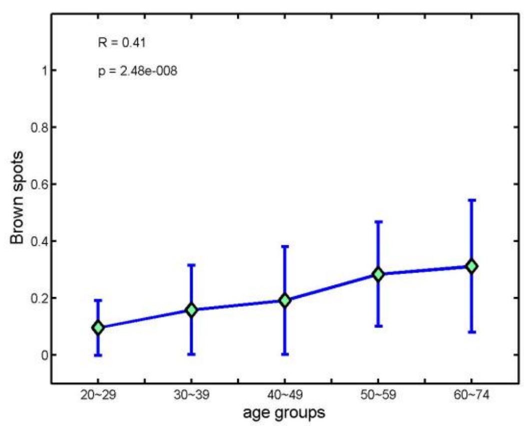
b) Uniformity



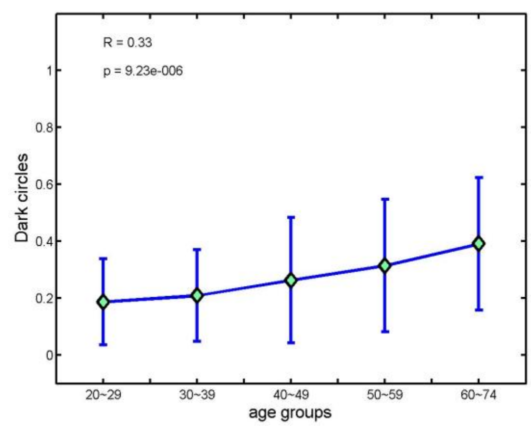
c) Radiance



d) Diffused redness



e) Brown spots



f) Dark circles

Figure 2-9: Changes on skin color attributes with age

a) color, b) skin tone uniformity, c) radiance, d) diffused redness, e) brown spots, f) dark circles

II.4. Discussion

As expected, both chronological and photo aging have a measurable effect on facial attributes. Despite the wide variability within each age group when assessing a single parameter, the size of our sample population enables us to detect significant differences among groups of age. Regarding the shape of the curve that links them to age, attributes can be classified into four different groups:

- *Attributes that are not affected by age*
    a. *Skin texture*
    b. *Grain of the skin*
    c. *Diffuse redness*
    d. *Radiance*

- *Attributes that linearly change with age*
    a. *Bags under the eyes*
    b. *Dark circles*
    c. *Brown spots*
    d. *Under eye wrinkles*
    e. *Frown lines*

- *Attributes that linearly change with age and reach a plateau around 40-50 years of age*
    a. *Jaw line*
    b. *Skin color uniformity*
    c. *Lips thickness*
    d. *Crow's feet fine lines*
    e. *Slopping upper eyelid*
    f. *Eye opening*

- *Attributes with change faster after 40-50 years of age*
    a. *Skin color*
    b. *Forehead wrinkles*
    c. *Nasolabial fold*
    d. *Upper lip wrinkles*
    e. *Lip drawing*
    f. *Cheek wrinkles*
    g. *Neck wrinkles*

In agreement with the literature, it is observed that only three of the measured attributes are not affected by age (grain of the skin, diffused redness and radiance). The changes we have quantified mainly correspond to those already qualitatively described. In most cases, the speed of deterioration of skin attributes changes around 40-50 years of age. These variations are probably related to hormonal imbalance disorder caused by the menopause

(Broniarczyk-Dyla and Joss-Wichman 2001; Wines and Willsteed 2001; Raine-Fenning, Brincat et al. 2003; Sumino, Ichikawa et al. 2004).

III. Accuracy of clinical grading

Visual scales are widely used for clinical evaluation of several anatomical or physiological parameters. They enable to directly measure attributes that cannot be simply evaluated with devices. They are easy to use since they only require a training of the grader. In some cases visual scales are more accurate than measurement with devices because human graders can combine many signs to give their judgment while devices usually focus on few of them. However, the accuracy and the reproducibility of each visual scale should be addressed before extensive use. In addition, the skin parameters assessed are multi-dimensional in contrast to a linear scale. As an example, the wrinkle severity rate is both related to the number, the length and the depth of wrinkles in the area of interest. Reproducibility of graders is therefore a main issue.

The grader who collected our data was well trained and had years of experiences in assessing clinical signs of aging. In order to use his results for other studies, we have decided to compare his grading with that of another expert.

The two graders were asked to rate the same skin attributes in-vivo on the same set of people. The assessments were done in-vivo by observing the subject's skin. Both graders were asked to grade the same parameters in a population of 31 subjects' age from 20 to 75 years of age. Their results were then compared using linear correlation. For each measured parameter, the table below (Table 2-2) gives the Pearson correlation coefficient between the two graders, its significance, the mean error and the standard deviation of error when comparing the two graders. The two graders are significantly correlated for most of the measured attributes, with a Pearson coefficient higher than 0.5 for 70% of the attributes.

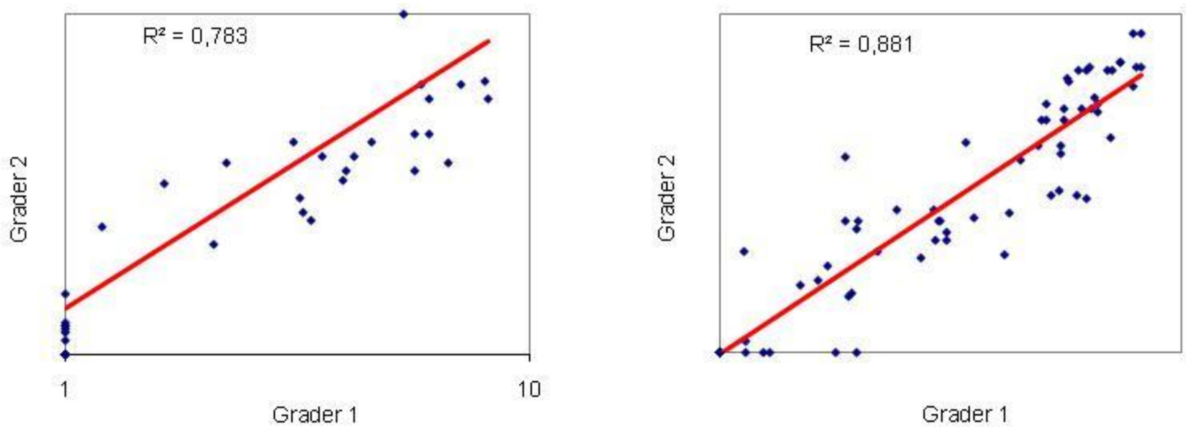
Parameter	Coeff Corr (R)	P value	Mean Error	Standard deviation of error
Jaw line	0.93	1.01E-16	-0.25	1.49
Frown line	0.89	1.82E-13	0.65	1.83
Forehead wrinkles	0.87	1.53E-12	1.51	1.73
Crow's feet wrinkles	0.81	1.89E-09	0.71	2.04
Brown spots	0.80	1.97E-09	-2.32	2.18
Sloping upper eyelid	0.80	2.11E-09	0.57	2.25
Eyes opening	0.79	7.93E-09	0.10	1.77
Lips color	0.76	4.98E-06	1.94	2.01
Upper lips wrinkles	0.75	8.37E-08	-0.65	2.22
Lips volume	0.74	8.47E-07	0.99	2.12
Nasolabial fold	0.67	6.92E-06	1.19	2.41
Dark circle	0.60	9.70E-05	-1.99	2.10
Under Eyes wrinkles	0.56	3.36E-04	0.61	2.32
Bags under the eyes	0.54	5.95E-04	0.09	2.65
Skin uniformity	0.54	7.65E-04	1.04	2.61
Skin texture	0.51	1.24E-03	-0.01	2.93
Cheek wrinkles	0.47	3.68E-03	-2.38	2.68
Diffuse redness	0.41	1.27E-02	-3.14	2.99
Grain of skin	0.40	1.30E-02	-0.58	2.48
Lips definition	0.37	4.70E-02	0.85	3.20
Radiance	0.36	3.03E-02	1.33	2.92
Skin color	0.33	5.12E-02	-0.34	3.52
crow's feet fine lines	0.28	9.29E-02	0.67	3.17

Table 2-2: Results of linear regressions made to compare the agreement among two graders

Some attributes were weakly correlated (Pearson $R < 0.5$ and $p < 10^{-2}$) meaning that the grading process could be improved to have a better agreement between graders. We have therefore proposed a new photonumeric scale based on calibrated digital parallel-polarized images of the face. Parallel-polarized images were preferred to regular images to increase specular reflections and to enhance the appearance of skin features such as wrinkles and furrows (Kollias 1997). One scale was built separately for each part of the face (forehead, crow's feet, under eyes). The detailed methodology, which is beyond the scope of this document have been presented at the American Academy of Dermatology (Washington 2007). The Figure 2-10 shows an example of nine grades scale that we proposed to assess the crow's feet wrinkles.

Figure 2-10: Nine grade scale built to assess the severity of crow's feet wrinkles

The improvement of the inter-observer agreement is illustrated on Figure 2-11 that compares the results of two graders on the crow's feet area, before and after the use of the picture scale.



a) Before using the scale

b) After using the scale

Figure 2-11: Improvement of graders agreement after the use of a proposed image scale, for the crow's feet area.

IV. Conclusion

In this chapter, we assessed the clinical changes in facial appearance of Caucasian women. VAS were used to clinically measure the facial attributes. A comparative study proved that VAS were reproducible and accurate tools to measure complex and multidimensional attributes such as wrinkles and color. However, new graders need to be intensively trained before providing fully reproducible evaluations. A wrinkle grading scale, made with high quality pictures was built to improve the training of the graders.

We particularly focused on changes in skin attributes and demonstrated that many visible transformations occur both on shape, textural and color aspects of the face as women aged. Some of these transformations where likely accelerated or decelerated along the life span.

Since many attributes have been reported in relation with facial aging, one may wonder which of them mostly drive the perception of facial aging. This question is addressed in the next chapter.

Chapter 3
Age perception

The facial appearance of a person does not always reflect the chronological age; some people look younger or older than they really are. In the previous chapter, we have described the main changes that occur with age and have seen the variability in facial attributes aspect within the same group of age. When talking about apparent age two questions could be addressed:

- Since apparent age is subjectively perceived by observers, is there a consensus on the apparent age that would be given to an individual? If not, will the observer being influenced by his own experience?
- What are the main facial attributes that drive our age perception?

Many studies have described the changes in facial attributes (skin color, wrinkles, sagging, micro relief, etc) with age, but few of them have analyzed their influence on the perceived age. The primary objective of this chapter is to analyze the contribution of individual skin attributes of the face on the perceived age of Caucasian women. Secondary objectives are to assess the influence of age and gender of graders with regards to the age perception. Firstly, we present the state of art, most of the findings coming from psychology literature; then we propose a ranking of all facial attributes as measured from our women database, relatively to age perception.

I. Previous work

It is generally claimed that "Humans can easily categorize a person's age group and are often precise in this estimation" (Gandhi 2004). The human capacity to estimate people's age has probably been developed throw the evolutionary process since age is a key point for social interaction. Shortly after birth, faces and material that look like faces already invoke particular interest to babies. Their attention is more attracted with faces than other visual stimuli (Macchi Cassia, Kuefner et al. 2006). To illustrate our ability to recognize faces, the reader will see plenty of faces in the image below although there is actually no face.

Figure 3-1: The thirteen faces image.

This image represents a man with two horses in a land wood. Depending on personal observation skills, one may find between 13 and 22 faces[3]

I.1. Biases in age perception

The ability to recognize and therefore to classify faces has been demonstrated to be plagued by several "own-group" biases. People are generally better at recognizing, remembering and classifying individuals belonging to their own group. The own-group bias effect has been found for such characteristics as age, race and gender; but the reasons underlying these biases are not clear yet (Rehnman 2007). We can assume that people are having more interactions with others within their own racial, age and gender group and are more trained (and efficient) with these categories of faces.

I.1.1. Own-race bias

The own-race bias is known as the capacity for people to better recognize faces from their own race (Dehon and Bredart 2001). A meta-analysis (Meissner and Brigham 2001) involving 35 articles with around 5,000 participants clearly highlights this bias, which is now taken into account for testimonies in the New Jersey Court. In fact it seems that face recognition in general, and age estimation in particular, involves learning processes through which people are trained by their environment.

Dehon & al. (Dehon and Bredart 2001) specifically studied the own-race bias in age estimation. They asked Caucasian and African participants to estimate the age of Caucasians and Africans from their facial pictures. Caucasian participants were better at estimating Caucasian faces than Africans'. However, Africans had the same performance for the two

[3] Image from : http://andfunforall.blogspot.com/2007/11/hidden-faces-illusion.html

groups. Since the study was run in Belgium, the authors suggest that African had been trained during their daily life to recognize and to classify Caucasian faces. This hypothesis is in line with Wright (Wright and Sladden 2003) who states that the ability to recognize other-group faces depends on the degree of exposure and contact with people from these groups.

I.1.2. Own-gender bias

Rehnman (Rehnman 2007) has recently published a state of art review about the own-gender bias in face recognition. He concluded that the women have greater accuracy in face recognition than men, and that women could more easily recognize female faces. On the contrary, men did not show any difference in recognizing women's or men's faces.

Looking at gender bias in age prediction, it has also been demonstrated that women and men do not estimate age the same way (Wright and Sladden 2003). A hypothesis to explain these differences is that depending on our age or gender, we focus on different facial attributes when estimating someone's age.

I.1.3. Own-age bias

The own-age bias has been documented by several authors, most of them trying to evaluate the confidence that could be given to eyewitness testimony. A review of literature made by Anastasi and Rhodes (Anastasi and Rhodes 2006) exhibits mixed and inconclusive results. Some authors have found an own-age bias in the oldest group, meaning that they better recognize oldest people. Others have found that young people better recognize young faces. Even if not yet well demonstrated, the own-age bias for face recognition is taken into account in the legal system with regard to eyewitness testimony.

Consequently, the influence of own-age in people's age estimation has also been explored. The underlying assumption is that people are judging other's age comparatively to themselves (Nagata and Inokuchi) and they should therefore be more accurate in their age range. An own age bias has been reported by George and Hole (George and Hole 1995). He ran two experiments including 25 young adults and 25 old adults who were asked to assess the age of faces aged between 5 and 70 years. In the second experiment, the volunteers were asked to judge three images per age group. It appears that young people overestimate the age of older people, older people overestimate the age of younger people, and finally that people are more accurate when estimating age within their own age group. Based on these results, Sörqvist (Sörqvist and Eriksson 2007) also tried to evaluate the own-age bias and to improve age estimation accuracy by training. He asked two groups of people: young (15-24 years) and middle age (34-46 years)) to evaluate 78 facial images from people in the three groups of age (young; middle age; old). He found that young participants were better with their age groups than with older people but the middle age group was not significantly better when judging people from their group of age. Thus, the own-age bias hypothesis was rejected.

I.2. Influence of facial attributes on perceived age

The estimation of age is based on facial attributes, but also on hairs, voice, body movements and posture (Rexbye and Povlsen 2007). However, some studies have investigated the facial attributes that influence age perception (Mark and Todd 1985; Henss 1991; Lanitis 2002). Most of them have used the well-described method of manipulations of facial images to assess which attributes influence the graders accuracy. These manipulations may include blurring the images to remove wrinkles, stretching to simulate the growth of bones, apply a round mask to remove information about jaw line drawing, and changes on skin pigmentation. George and Scaife (George 2000) focused on children's ability to predict age on unfamiliar faces. They presented four different versions of facial photographs to 134 children between 4 and 6 years old. The photographs were taken from volunteers between 1 and 80 years. The different versions of the face presented were: original image; internal facial features only (eyes-nose-mouth), skin-blur only, and overall blur (skin+features). Performances with the four sets of images were comparable, meaning that no conclusion could be drawn about the relative contribution of each facial feature.

However, it was claimed that adult observers are sensitive to many parameters including craniofacial shape (Mark and Todd 1985), skin surface information (Montepare and McArthur 1986; Burt and Perrett 1995) and color (Castanet and Ortonne 1997). Resbye and Povlsen (Rexbye and Povlsen 2007) also suggested that age estimation is more difficult on subjects with contradictory signs of aging such as a healthy skin but with an old-fashioned hair style. They asked 40 graders to evaluate 774 facial photographs of subjects older than 70 years old and to give them an age. Then the graders were asked to report the main features that have driven their perception. The main biological markers of aging were eyes and skin. In the eyes area, the graders would have focused on wrinkles, bags under the eyes, sunken and "watery" eyes and finally the vitality of the gaze. Concerning the skin, the graders would have focused on wrinkles on the face and the neck, and secondly on pigmentation, color and sagging.

II. Influence of facial features on perceived age of Caucasian women

In this section, we investigate the influence of the facial attributes we have measured in the previous chapter to age perception. To achieve this goal, we firstly define what is considered as the perceived age for each volunteer in the database.

II.1. Apparent age

In this section, we explore the ability of a panel of graders to give an age to the volunteers of our database. Consequently, we address the following questions:

- Is there a consensus in age estimation whatever the grader?

- How can we construct an apparent age based on the age given by a group of observers?

Firstly, we present the graders' population and the process used to give an apparent age. Then we measure the graders' agreement and propose a definition of the "apparent age".

II.1.1. Graders population

The estimation of age from facial pictures of the JNJ database is performed by a group of 48 "graders" (20 men and 28 women), from 20 to 64 years of age, recruited from a Caucasian population either at the University of Paris VI or within the volunteers of the Skin Care Research Institute. They were aware of the objectives of the experiment. The graders are divided into three groups of 15 years interval as defined: young (under 35 years), middle age (from 35 to 50 years) and senior (over 50 years). Figure 3-2 shows the age and gender distribution of graders. The total number of graders does not allow us to divide them in ten years groups of age as we previously did for the subjects' population. Therefore, they are divided in only three groups: young (20-35 years), middle aged (36-50 years) and senior(51-61) There are 41.7% of men in the population, 15% of them in the youngest group of age, 40% in the middle age group and 45% in the senior age group. There are 58.3% women in the population, 25% in the youngest group, 50% in the middle age group and 25% in the eldest group.

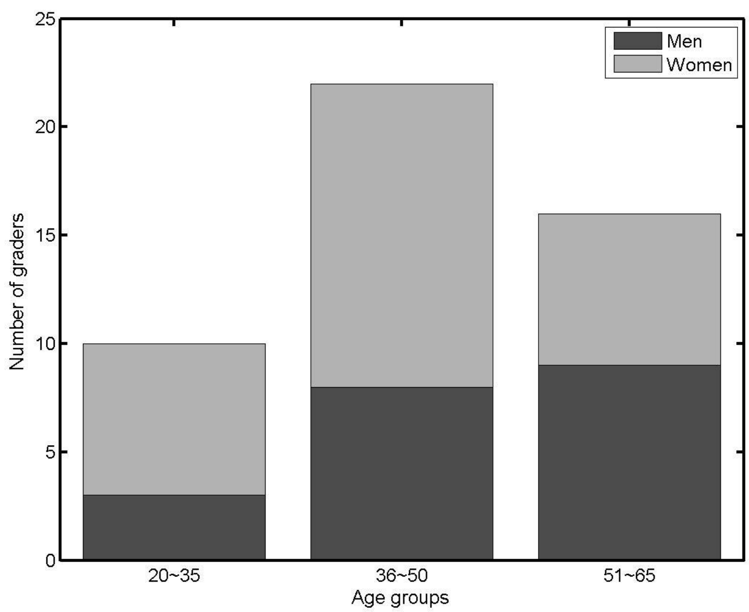


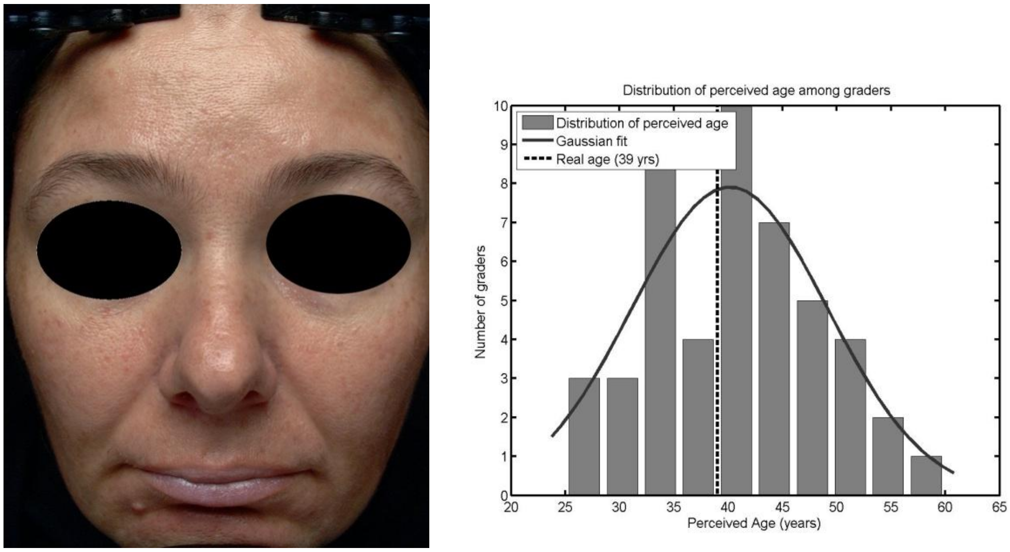
Figure 3-2: Distribution of graders among age and gender groups

II.1.2. Age estimation

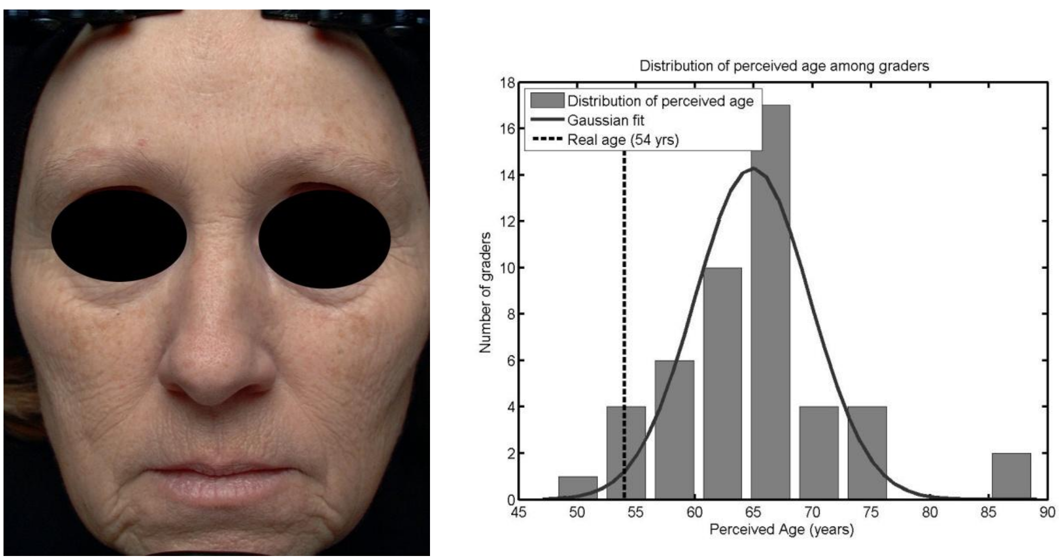
Photographs are sequentially and randomly presented to the graders on a computer screen. They are requested to give subjects' age using an interface that featured a slider bar spanning from 0 to 100 for age estimation. There is no time limit for the grading.

II.1.3. Agreement among the graders

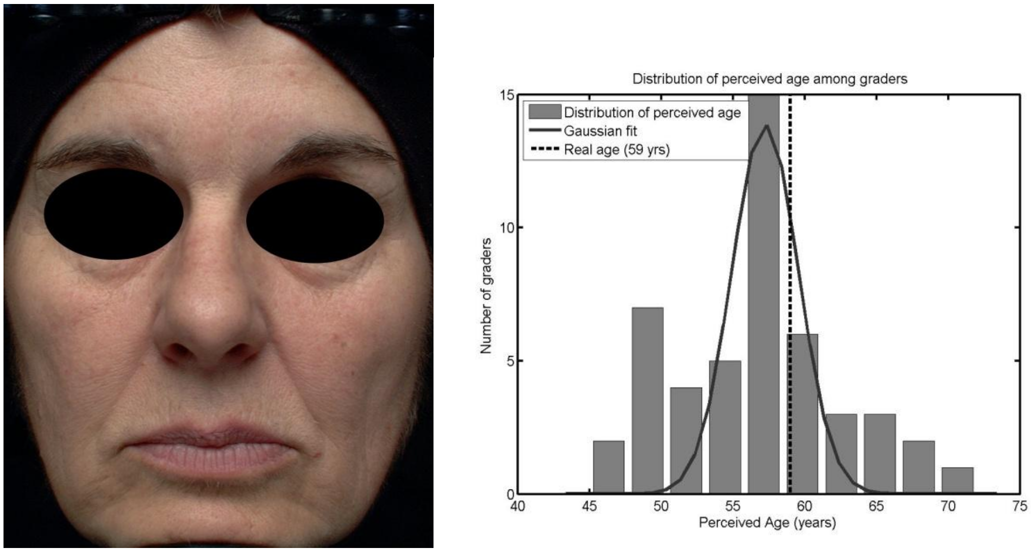
Each subject is given an age by each grader. The standard deviation in perceived age (from the distribution of ages given by the set of graders, for each subject) varies from 4.35 years to 9.36 years (mean= 6.50). Figure 3-3 displays four subjects a) with a small error (between real and perceived age) but a low agreement (among the graders), b) with a large error and a low agreement, c) with a small error and a high agreement, and d) with a large error and a high agreement.



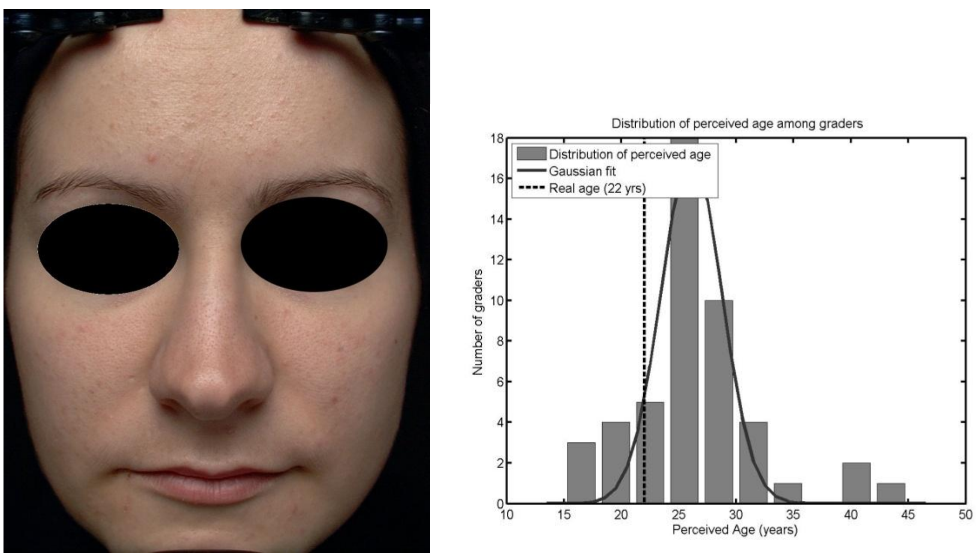
a) *Subject 1: Chronologic age=39 years; perceived age=41 years; standard deviation=8.2 years*



b) Subject 2: Chronologic age=54 years; perceived age=65; standard deviation=7.1 years



c) Subject 3: Chronologic age=59 years; perceived age=57 years; standard deviation=5.8 years



d) Subject 4: Chronologic age=22 years; perceived age=26 years; standard deviation=5.8 years

Figure 3-3: Agreement on perceived ages for four different subjects

In general, there is no relation between the agreement among the graders and their accuracy to predict the real age as it can be seen in the figure below (Figure 3-4). For each subject of the database, we plot the error in age prediction (real age minus predicted age) versus the standard deviation in perceived age. There is no correlation between the error in age estimation and the dispersion of the given age. On the contrary, we can notice that, as mentioned from the previous examples (Figure 3-3), some subjects are accurately judged with a high agreement among the graders; while for other subjects the graders strongly disagree.

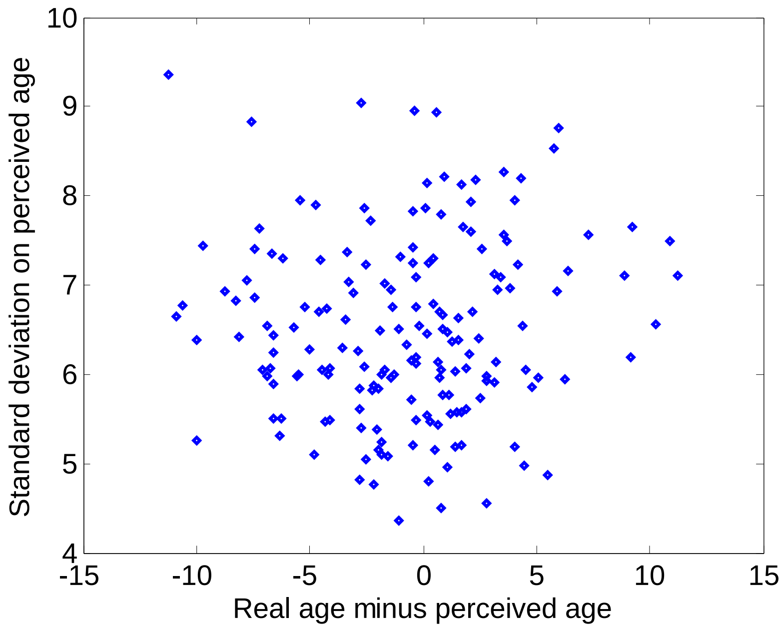
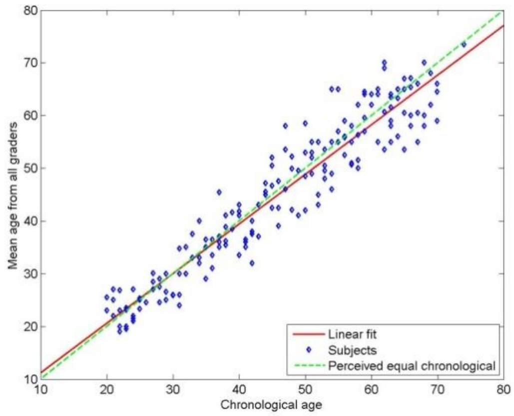


Figure 3-4: Correlation between accuracy and agreement among the graders.

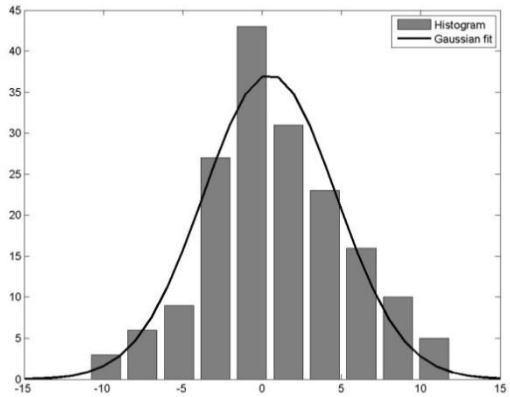
Each point represents the standard deviation on the perceived age versus the error in age estimation for each of the 173 subjects

II.1.4. Definition of apparent age

For 77.6% of the subjects, the given age follows a normal distribution (Jarque-Bera test of normality; significance =95%) with no outlier. Therefore for each subject, the apparent age could be defined as the mean value from age given by all graders. The Figure 3-5 displays the correlation between perceived age and real age. They are highly correlated (R= 0.95 $p<10^{-8}$ and the residual follows a Gaussian distribution with a mean of 0.9 years and a standard deviation of 4.49 years. Even if the difference between real age and perceived one is not high (0.9), a pair-wise t-test shows that the perceived age is significantly lower than the real age (p=0.009) meaning that in general, the women of the dataset were perceived as 1 year younger than they are in reality.



a) Correlation curve between real age and perceived age



b) Histogram of the differences between real age and perceived age(Residual)

Figure 3-5: Mean age versus perceived age

II.2. Age and gender biases

II.2.1. Own gender bias

Our volunteers' population is constituted with only females. If an own-gender bias exists, the female graders should be more accurate in age prediction than men. To test this hypothesis, we compare the perceived age given by the women with the one given by the men and with the real age using a pair wise t-test.

The mean perceived age within male graders (44.93) is significantly different from mean real age (mean= 45.85, p = 0.009) and from perceived age within female graders (mean=45.21, p = 0.038). There is no significant difference between perceived age by female graders and real age (p-value= 0.060). Male graders perceive that the subjects look younger than they really are and female graders are more accurate in age prediction than male graders.

II.2.2. Own-age-bias

The own-age-bias assumes that people are better with their own age group. Regarding the number of graders we have in our study, we have decided to divide the graders and the subjects' populations into three groups: the younger group under 35 years, the middle age group between 35 and 50 years and the eldest group with age over 50 years. This segmentation is relevant regarding what is done in the literature.

Table 3-1 shows the mean absolute error and the standard deviation for each age group of subjects and graders. The highest error and standard deviation is observed for the senior group of subjects. The youngest group of graders (less than 35 years) has smaller errors with a narrow distribution, being more accurate and having a better agreement than the other age groups. The smallest error and standard deviation are obtained when young graders judge young subjects.

		Subjects Age		
		Under 35	35-50	Over 50
Graders Age	Under 35	4.55 (3.50)	5.98 (4.82)	6.21 (4.78)
	35-50	4.98 (3.78)	6.43 (4.71)	7.52 (5.75)
	Over 50	5.40 (4.70)	6.65 (5.02)	6.42 (4.92)
	All graders	5.01 (4.06)	6.57 (4.95)	6.78 (5.33)

Table 3-1: Means and standard deviations of absolute values of errors in age prediction.

The means and standard deviations are calculated from the absolute values of the difference between real age and perceived age. Each cell corresponds to the mean and standard deviation of the errors from all the graders of an age group on subjects of an age group.

Table 3-2 shows the characteristics of the linear regression model built to explain the error in age perception based on the graders' and subjects' age. The coefficients associated with the subjects' and graders' ages are positive and significantly different from 0, meaning that the error in age perception significantly increases with subjects' and graders' age.

Term	Estimate	Std Error	Prob>|t|
Intercept	2.83	2.76E-01	<0.005
Subjects' age	0.05	3.64E-03	<0.005
Graders' age	0.02	4.88E-03	<0.005

Table 3-2: Parameters of the linear regression model between the absolute value of error in age estimation and subjects' and graders' ages

'Estimate' is the coefficient of the given parameter in the linear regression model and 'Std Error' is the standard error associated with this coefficient. 'Prob>|t|' is the value of the confidence test that the 'estimate' is significantly different from 0. The Estimate is significantly different from 0 for both subjects and graders age meaning that these two parameters influenced the error in age estimation.

II.3. Influence of facial attributes on the perceived age

In contrast with Rexbye(Rexbye and Povlsen 2007) who used a questionnaire to understand the influence of facial attributes on the perceived age and with George (George 2000) who modified the aspect of the attributes in facial images, we chose to build regression models to link facial attributes with perceived age. This approach enables to compare the influence of the different attributes and to rank them. Consequently, within the framework of regression methods, our problem belongs to the field of variable selection and classification. We build regression models that links age with the 22 facial attributes measured on people from the JNJ database.

II.3.1. Variables selection

Regression methods can be used to study the dependence of the perceived age Y on several facial attributes named variables $x_1, \dots, x_p$. Among regression methods, linear regressions are easier to interpret and consequently would better fit with our objective. The least square multiple linear regression is the most popular. It enables to describe the predicted variable Y as a linear function of $x_1, \dots, x_p$. The coefficients of the linear regression are chosen in order to minimize the mean square error between the real values Y and the predicted one. Even if least square regression models are easy to build, they are not interpretable when the variables $x_1, \dots, x_p$ are correlated. In this case, Partial Least Square (PLS) regression is a better tool to understand the contribution of each variable for the regression (Tenenhaus 1998) .

II.3.1.1. Partial least square (PLS) regression

The principle

Given a predicator $X(x_{n,p})$ made with n individuals described by p variables and $Y(y_n)$ the predicted variable, regression problems aimed at finding a function $\hat{f}: x_i \in X \to y_i \in Y$. When the dimension of X is high and particularly when X has a reduced rank $\alpha < m$, one may need to find a subset of variables t called latent variables, which carry most of the information from X.

To achieve this objective, PLS has been proposed by Wold et al. (Wold, Ruhe et al. 1984) The objective of the PLS is to find the set of variables t that will capture the maximum information from X while explaining the maximum of variance from Y. Therefore t is both related to X and Y as follows.

$$\boldsymbol{Y} = \boldsymbol{tQ} + \boldsymbol{f} \tag{3-1}$$

$$\boldsymbol{X} = \boldsymbol{tP} + \boldsymbol{e} \tag{3-2}$$

where t contains up to α latent variables $(t_1, t_2, \cdots, t_\alpha)$ defined as a basis for the reduced-rank component of X and Y; P and Q are matrices of coefficients relating t to X and Y respectively; and f and e are considered to be random errors. The latent variables t are built to be orthogonal; so that $cov(X, t_h) > cov(X, t_{h+1})$ and they maximized $cov(t_h, Y)$.

The idea behind the PLS is that since X has a reduced rank $\alpha < m$, there exists a subspace where less variables will be needed to describe the individuals while keeping the information related both to X and to Y. The objective of the method is to find this subspace made with latent variables t.

The following algorithms to construct the latent variables $t(t_i)$ and the final regression equation to link $X(x_i)$ and Y are well explained in the book "La regression PLS. Theorie et pratique" (Tenenhaus 1998).

Firstly, the first latent variable t_1 is built as follows:

$$t_1 = w_{11}x_1 + \cdots + w_{1p}x_p \tag{3-3}$$

where

$$w_{1j} = \frac{cov(x_j, y)}{\sqrt{\sum_{j=1}^{p} cov^2(x_j, y)}} \tag{3-4}$$

then, a regular regression of Y on t_1 is computed

$$Y = c_1 t_1 + y_1 \tag{3-5}$$

where c_1 is the regression coefficient and y_1 is the residual. Therefore, we obtain the following regression equation, where the coefficients are proportional to the importance of the predicator.

$$Y = c_1 w_{11} x_1 + \cdots + c_1 w_{1p} x_p + y_1 \tag{3-6}$$

If the model that was built is not powerful enough, one can construct a second component t_2 that will be a linear combination of the residuals x_{1j} of the regression of the variables x_j on the component t_1. t_2 is calculated as follows

$$t_2 = w_{21}x_{11} + \cdots + w_{2p}x_{1p} \tag{3-7}$$

where

$$w_{2j} = \frac{cov(x_{1j}, y_1)}{\sqrt{\sum_{j=1}^{p} cov^2(x_1, y_1)}} \tag{3-8}$$

Then the linear regression of Y on t_1 and t_2 is performed:

$$Y = c_1 t_1 + c_2 t_2 + y_2 \tag{3-9}$$

t_1 and t_2 are then substituted by their values and we obtain a linear relation between X and Y.

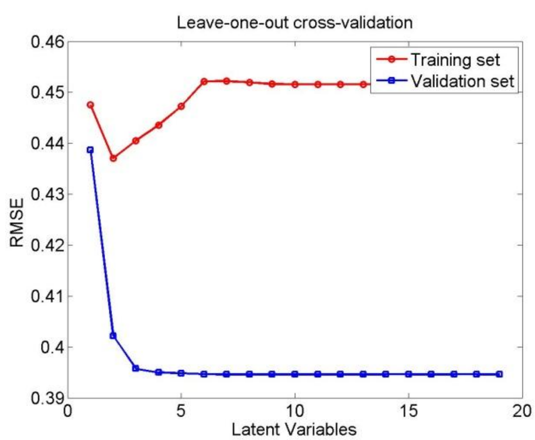
This procedure is iteratively applied until the "best" number of latent variables is reached. The best number of latent variable is defined using a cross-validation technique described in the annex 2. Another issue is also to define the confidence interval of the regression coefficients obtained using this algorithm. The bootstrap method has been shown to be a good tool to measure confidence intervals. The method is also debate in annex 2.

II.3.1. Estimation of influence of facial attributes using PLS

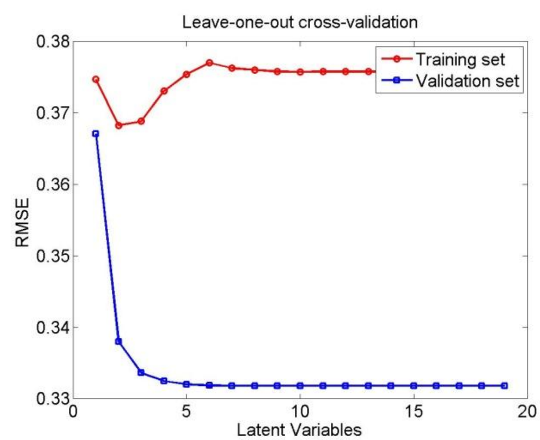
The clinical assessment (X) data and the ages (Y) are normalized in order to have their mean set to 0 and their standard deviation to 1. This procedure is highly recommended in order to have all the variables in the same scale.

Seven PLS models are built to predict age from the facial attributes. The first two models enable to predict the chronological and the perceived age. The three other models are used to predict the perceived age as given by the three groups of graders: young- middle age - seniors. The last two models are built to predict perceived age as estimated by men only and by women only.

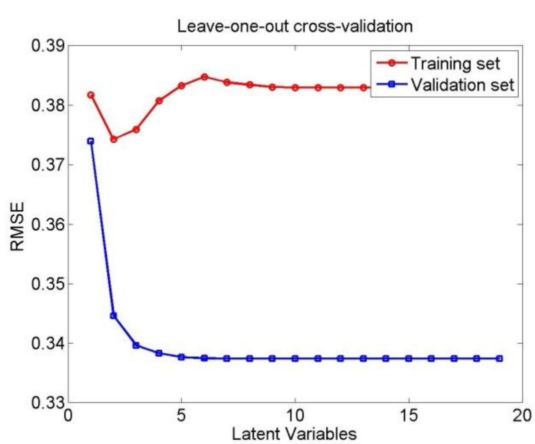
The number of latent variables is calculated for each model using the loo procedure. The figures below give the Root of Mean Square Error (RMSE – Refers to the annex 2 for its definition) in the training and the validation set for each model as a function of the number of latent variables.



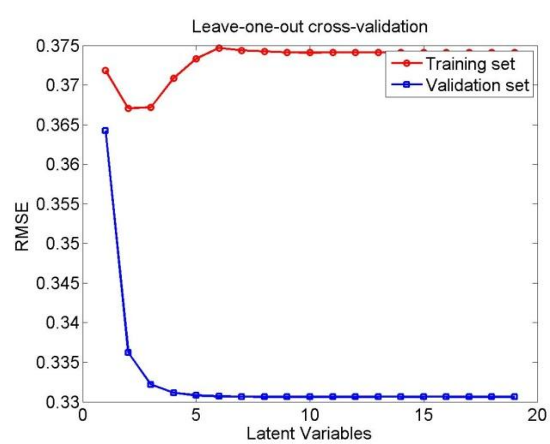
a) Chronological age



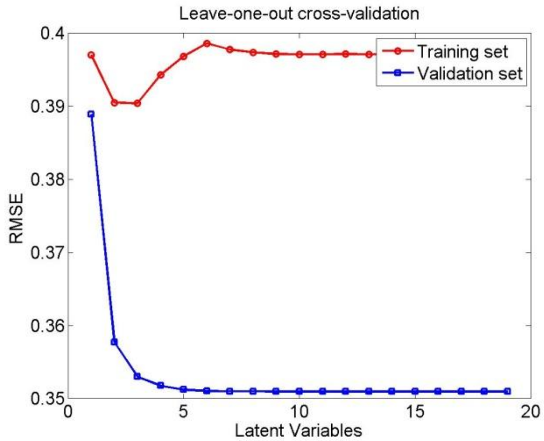
b) Perceived age (All graders)



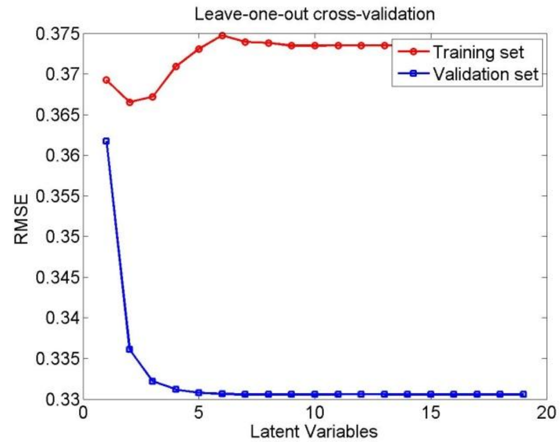
c) Perceived age (Men)



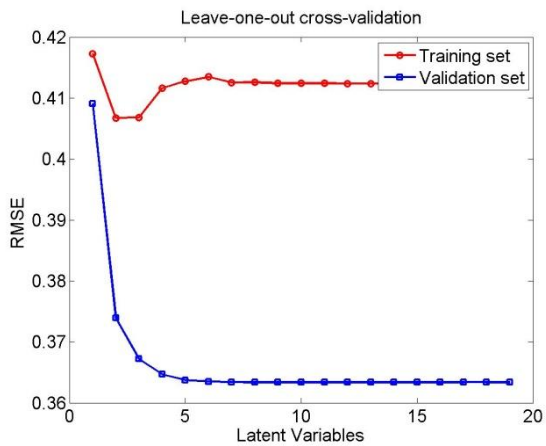
d) Perceived age (Women)



e) Perceived age (Graders under 35 years)



f) Perceived age (Graders between 35 and 55 years)



Perceived age (Graders over 55 years)

Figure 3-6: Validation of the five PLS models built for age prediction

Each graph represents the RMSE for the training and the validation set, as a function of the number of latent variables. The suitable number of variables is the one that minimize the RMSE for the validation set. These curves show that as the number of latent variables increase, the RMSE decrease for the training set, but not necessarily for the validation set.

For all models, two latent components enable to minimize the RMSE in the validation set of data. The confidence interval of each model coefficient is obtained using the bootstrap procedure with 250 resamplings. These confidence intervals allow comparing the contribution of variables among the different models.

Regarding the relative contribution of the different attributes in the perceived age and the real age models, some statistically significant differences are detected with some attributes (Figure 3-7) 'Eye opening' and 'Lips border definition' play an important role for chronological age while 'Under Eye wrinkles', 'Dark circle', 'Bags under the eyes', 'Skin Color uniformity' and 'Brown spots' play a more significant role in perceived age.

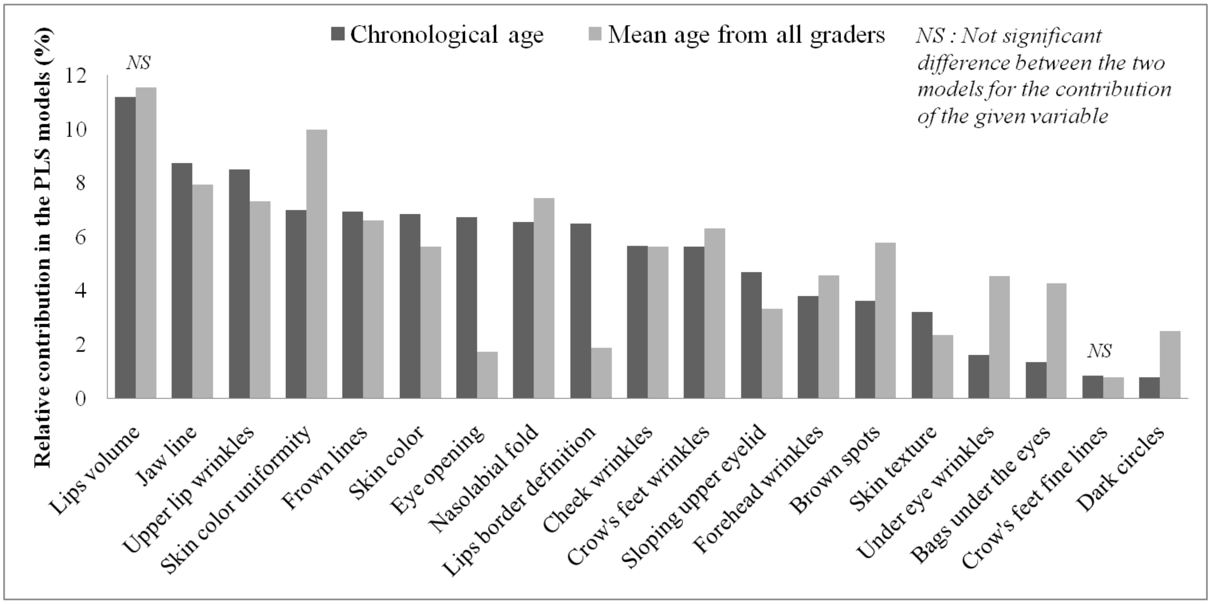


Figure 3-7: Comparison between PLS model of real age and perceived age

Each bar chart represents the relative contribution of the facial attributes while building the PLS model for age prediction. The highest the value, the more important is the attribute for the given model. The attributes are ranked from the most important to the least important for chronological age. The parameters whose contribution differs the most between the two models are skin color uniformity (p= 1.14E-59), lips border definition (p=6.46E-69), under eye wrinkles (p=7.43E-68) and bags under the eyes (p=2.12E-42)

The models of the three age groups' also present some differences as shown on Figure 3-8. Particularly, the oldest group overuses the attributes 'Eye opening', 'lips border definition' and 'lips volume'. In contrast, they disregard the attributes 'Nasolabial fold', 'Dark circles' and 'Brown spots'. Some cases may be used as illustration: Subject 1 (Figure 3-3) is 39 years old. The senior graders estimate her age to 40.5 while the young graders estimate it to 44 based on the nasolabial fold. Subject 4 (Figure 3-3) is 22 years old, her age is estimated to 25 by the young graders and to 28 by the senior graders, based on her lips.

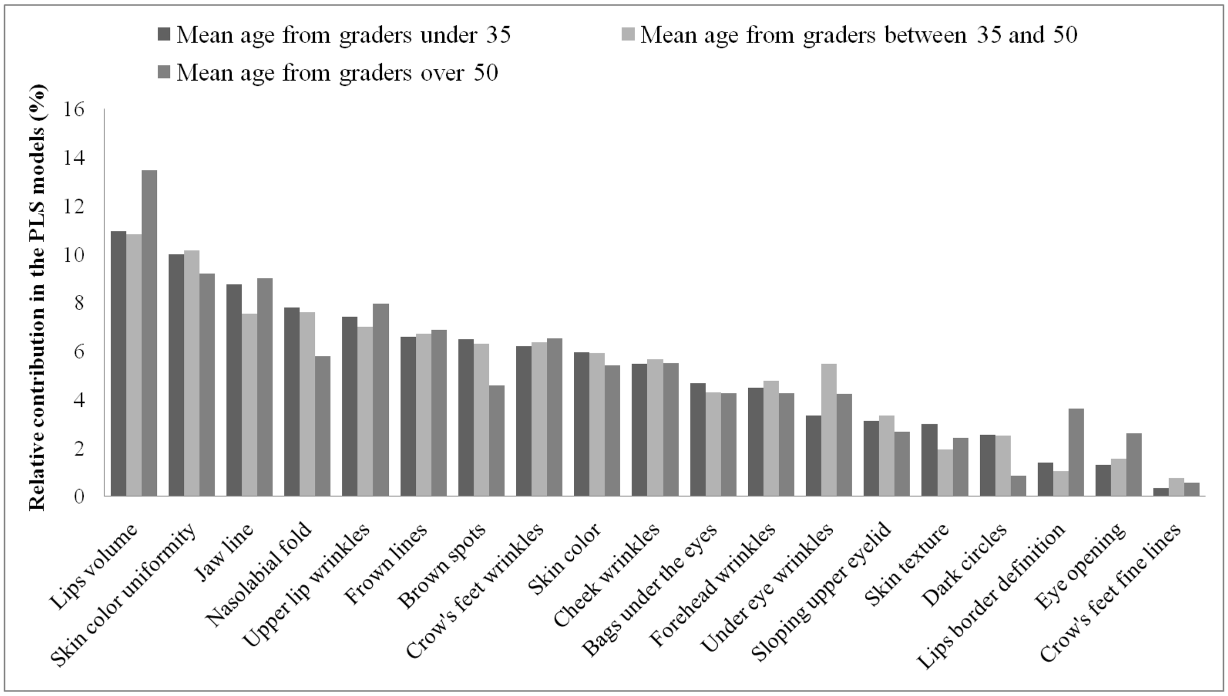


Figure 3-8: Comparison between PLS models of perceived ages for different age groups of graders

Each bar chart represents the relative contribution (in %) of the facial attributes while building the PLS model for age prediction. The higher is the value, the more important is the attribute for a given model. Attributes are ranked from the most important to the least important to predict perceived age by the youngest group of graders

The comparison between the PLS model of perceived ages by men does not show any statistical significant difference compared with the model built on perceived ages by women. The study does not highlight any difference in terms of the interpretation of attributes between men and women.

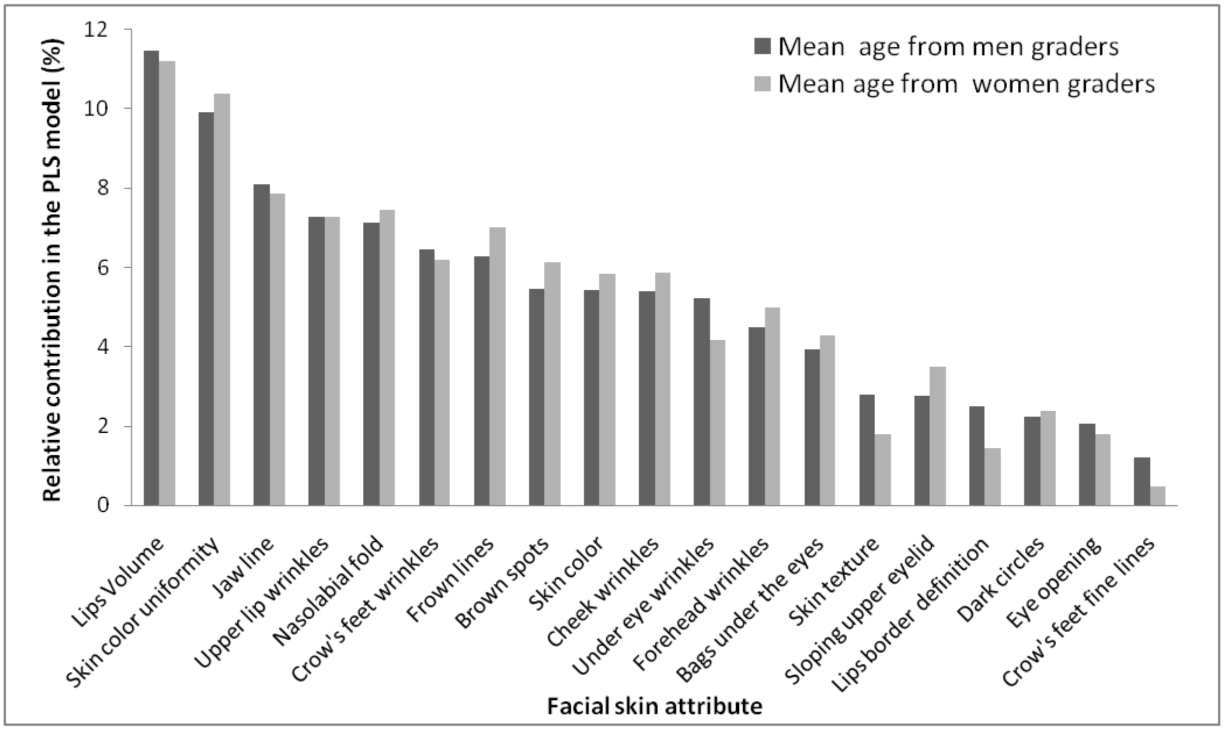


Figure 3-9: Comparison between PLS model of perceived ages from different gender groups of graders

Each bar chart represents the relative contribution (in %) of the facial attributes while building the PLS model for age prediction. The higher is the value, the more important is the attribute for a given model. Attributes are ranked from the most important to the least important to predict perceived age by the youngest group of graders

III. Discussion

In this chapter we have assessed the influence of different facial attributes on the perceived age. The perceived ages were given by a large set of Caucasian graders in order to reduce the own-group bias. The judgments were done on full face images with neutral expression. In real, faces are dynamic and expressions and emotions may also contribute to the perceived age. In addition, people also use hair, clothes and body posture information when available (Rexbye and Povlsen 2007). The way we set up our experiment does not enable to take into account the influence of facial expressions and hide all information but the face. Thus we may be able to focus on facial attributes. The pictures presented to the graders only shown faces, with no make-up and no visible hair.

Our main objective was to understand the influence of some skin facial attributes on perceived age. The choice of these attributes was driven by a review of literature on facial and skin changes with age (Fedok 1996; George 2000; Taister, Holliday et al. 2000; Guinot, Malvy et al. 2002).

To ascertain a causal link between two events, four conditions are required (Cohen 2003).

1. There should be a temporal precedence (The cause precede the result)
2. Some mechanisms explain how the causal effect lead to the result (Causal mechanism)
3. A change in the cause is followed by a change in the consequence (Association or correlation)
4. The effects of the specific cause can be isolated from the effects of other potential events that occur at the same time

In our case, the condition 1 and 2 are met since the attributes were suitably chosen. To evaluate the condition 3, several authors have artificially modified the aspect of each single facial attribute to measure the subsequent change in age perception. These works gave interesting insights about the importance of the attributes, but we agree with George et al (George 2000) who concluded that facial attributes are always used in combination. We have consequently chosen to measure the intensity of each facial attribute by building a multiple regression model to relate all attributes to age. PLS is an appropriate tool for that task since this regression model handles both highly correlated variables (facial attributes) and relatively small sample size (Wold, Sjostrom et al. 2001) (173 subjects). In addition, PLS enables to evaluate the relative contribution of each predicator (Burnham, MacGregor et al. 2001).

The accuracy of the graders about age estimation (RMSE=4.57) is similar to what has been observed by Lanitis (Lanitis 2002)(3.64). We may notice however, that the graders are more or less in agreement depending on the subject. The disagreement is often observed for subjects with contradictory signs of aging as suggested by Rexbye (Rexbye and Povlsen 2007). For example, subject 1 (Figure 3) presents few wrinkles for her age, but has a lot of brown spots and deeper nasolabial fold compared to her age group.

In order to have an overview of the influence of female facial attributes on perception, male and female graders were enrolled despite the fact that the graded subjects were only female. In this study, female graders were more accurate than male for age estimation. This may be explained by the fact that the evaluated population only included female. The own-gender bias is consistent with the results of Wright and Sladden (Wright and Sladden 2003) who found that females are better at identifying female faces and vice versa for male, although we do not address here whether males would be better at predicting age from male faces.

Our results show that the age of graders plays a role in their perception and confirm what has already been reported by George and Hole (George and Hole 1995): young people have an overall greater accuracy as compared to older people. One may argue on the contrary that older people should be more accurate since they have longer experience (and therefore longer training) than younger ones. On one hand, it appears from a literature review that training is not the only factor that influences people accuracy for age prediction. Both from our data and from the literature (Sörqvist and Eriksson 2007), the lack of accuracy from old people is mostly due to their overestimation of young people's age. On the other hand, Sörqvist and Eriksson (Sörqvist and Eriksson 2007) have shown that a training process does not increase the ability to predict the age of young people. Finally, old people may have also found it more difficult to estimate young people's age because the age-related changes occurring between 20 and 35 years of life are not as obvious as the changes occurring at a later age.

The importance of the nasolabial fold, dark circles, bags under the eyes, skin color and uniformity as well as brown spots decreases with graders' age. In contrast, the importance of the frown lines, crow's feet wrinkles and eye opening increases with graders' age.

Taken as a whole, these results suggest that our perception of aging is highly influenced by the appearance of the lip area (volume and upper lips wrinkles), the eye area (crow's feet and under-eye wrinkle, dark circles around eyes, bags under the eyes) and the skin tone uniformity (brown spots, skin color uniformity). The importance of the eye area confirms the finding of Lanitis (Lanitis 2002) and Rexbye (Rexbye and Povlsen 2007) who also found skin wrinkling and color as important features. Compared with the two previous authors, our results also drawn attention on the lips area and this finding might justify the recent popularity of fillers for lips.

IV. Conclusion

In this chapter, we have investigated the influence of facial attributes in perceived age and have found that the perceived age was highly correlated with the real age despite individual variability. We have tried to answer two questions that could help to better understand the key point to focus on when dealing with facial aging or rejuvenation. These two questions are:

1. Since apparent age is subjectively perceived by observers, is there a consensus on the apparent age that would be given to an individual? If not, will the observer being influenced by his own experience?
2. What are the main facial attributes that drive our age perception?

From a revue of the literature, we have found that age perception is biased by some own-groups effects like own-race, own-gender and own-age of graders. The gender and the age biases were explored and confirmed in our available dataset.

Regarding the second question, we have used an innovative method as compare to the literature and have found the main attributes that drive the perception of age from faces. These attributes are the lip area, the eye area and the skin tone. Regarding observer's age, young observers focus on color uniformity, nasolabial fold and dark circle, while older observers look at lips drawing and thickness and eyes opening. These surprising results need to be confirmed by other studies made on larger databases. One can also use eyes tracking systems as those used to explore face recognition, to track where people's eyes look at when observing a face to give an age.

Chapter 4
Age prediction

Age prediction from facial image belongs to the category of face analysis problems. These problems included but are not limited to: a) face localization, b) facial features segmentation, c) gender classification, d) facial expression recognition, e) age regression, f) age simulation.

Automatic age estimation can be seen as a multiple regression problem: the data are the subjects' facial images and the targets are the subjects' age.

Lanitis (A . Lanitis, C. Draganova et al. 2004) has pointed out four main motivations behind the works done for automatic age prediction:

1- *Age specific human interaction* could give computers and electronic devices the ability to interact with human being by taking into consideration peoples' age. Such a system could for example be used in combination with secure Internet access control to ensure that under-aged persons do not access to forbidden Internet pages. An Age Specific Human Computer Interaction (ASHCI) can also be implemented in daily life applications. As an example, a terminal could automatically adapt his vocabulary, interface and services to the customer's age. A vending machine could refuse to sell alcohol or cigarettes to the underage people (Geng, Zhou et al. 2006)
2- *Age-based indexing of face images* could permit an age-based retrieval of face images from databases. With the development of digital camera and Internet, image indexing is being an important field of research. A system that can recognize the age of a person would improve the performances of content-based web pictures browsers.
3- *Understanding the process of age perception by humans.* Work in the field of automatic age estimation could provide new insight or stimuli to help psychologists who study age perception by human.
4- *Development of age progression systems*. The methods used for age prediction provided numerical models of facial aging. These models can therefore be used to simulate aging on facial images. Aging simulation would be helpful for police

investigation to capture wanted fugitives (CrimeLib 2002) or for the finding of missing children.

We have previously study the signs of facial aging and have shown that despite several biases, human are good for age prediction. We have also identified the most important facial attributes for human age perception. In this chapter, we propose an automatic age prediction algorithm that uses the facial images for age prediction. Firstly, we present the methods available in the literature. Then, we introduce a new statistical model of face built for accounting all the age related changes in facial images. The facial model enables describing a face with a few set of parameters that are related to the regression/classification objective. Finally, we perform a linear regression on the extracted parameters in order to predict people's age.

I. Previous works

Despite face image processing is widely explored, age classification has received much less attention. Current approaches could be divided in three main categories: features based methods, statistical based methods and aging patterns based methods.

I.1. Feature based methods

In features based methods, a pre-processing enables to extract several facial cues known to be related with age and to use them for the classification task. This approach includes three main steps:

1- *Face segmentation* where the face is separated into anatomical parts based on the position of the main salient features (usually nose, mouth, eyes)
2- *Features measurement* where attributes related with facial aging (e.g. wrinkles, skin color,…) are quantify by a small set of parameters
3- *Age prediction* where the parameters extracted are used through a regression model to predict peoples' age.

To our knowledge, this approach was used for the first attempt of age prediction. Kwon and Lobo (Kwon and da Vitoria Lobo 1994; Lobo and Kwon 1998) have proposed an algorithm that combined facial wrinkles with craniofacial measurements.

Their face segmentation algorithm assumes that the position of the face is roughly known. Face contour and eyes iris are found using adapted contour snake algorithms. The position of the mouth is estimated using anthropometric assumption about its relative location in the face combined with an edge detector. The performances of the segmentation algorithm are not discussed in the publication, but one can expect poor performances on unseeing images since the edge of the mouth are not always well contrasted in the RGB color channels (Eveno 2003).

After the segmentation, Kwo and Lobo used the ratio of the distance between the eyes and the distance between the eyes and the nose to separate babies from adults while the skin wrinkles allow them to distinguish young adults from seniors. Kwo and Lobo successfully classify 47 faces into three groups: babies- young adult – seniors. The small size of the database and the absence of cross-validation makes it highly possible data overfitting.

Horng et al. (Horng, Lee et al. 2001) have proposed a more robust system using the same parameters (wrinkles and geometric ratios) than the previous one. A neural network performed the classification task. It achieved 81.6% of accuracy by separating 230 images in three age groups (babies – young adult – seniors).

An attempt made by Hayashi et al. to improve the accuracy in age prediction by increasing the number of age groups in their study failed (Hayashi, Yasumoto et al. 2002). They used a database of 300 individuals aged from 15 to 64 years of age. Pictures were taken under controlled conditions. The objective of the study was to predict people age and gender based on their facial wrinkles. Wrinkles were extracted using a Digital Template Hough Transform. The classification task was performed using a look-up table (LUT) that summarized changes for each wrinkle area. They obtained an accuracy of 27% on age estimation both because they have more age groups and because the LUT was not enough precise to handle multidimensional changes.

In conclusion, feature based methods are trying to measure some facial features known to be linked with age like wrinkles. Unfortunately up to now, the proposed approach only included few facial features. We have shown in the previous chapter that even if wrinkles are important, they are not the main factor driving aging perception. On the other hand, either measuring distances among features or estimating wrinkles are really sensitive to images acquisition conditions. Therefore, these parameters may bring a lot of noise to the final classifier. None of the recent attempts in age prediction have used this approach anymore. However this method has been extended for facial image synthesis (Seoul 2004; Ramanathan and Chellappa 2006).

I.2. Statistical based methods

Within statistical based methods, the objective is to perform the regression task from parameters statistically extracted from pixels values. The method could also be divided into three steps:

- Face registration where all the faces are aligned so that their main features are located at the same position.
- Face compression where the information handled by the pixels values is reduced using a dimension reduction approach and leads to a set small of eigenvectors characterizing each face.
- Age prediction using the eigenvectors from the previous step.

Gandhi (Gandhi 2004) used Support Vector Machines (SVM) to derive a function that can predict the age of a given frontal face image either from the pixels values or from the wavelet transform of the image. Using a database of 818 frontal images collected from Internet, he has been able to predict people age with an absolute error of 9.2 years. Although this result is not really accurate, we should keep in mind that the method only used spatial or frequency information from images acquired under various conditions of positioning and lighting.

Lanitis et al (A . Lanitis, C. Draganova et al. 2004) have reduced the error in age prediction up to 3.82 years. Before the classification, they built statistical face models by applying a PCA decomposition of facial shape and texture. These facial models have gained notoriety under the name of Active Appearance Models (Edwards, Taylor et al. 1998; Cootes, Edwards et al. 2001) (AAM). They enable to describe the main variation within a group of faces by a small set of parameters known as “face parameters” (22 parameters in the Lanitis experiment). “Each of the model parameters controls a systematic way in which face images vary within the training set” (A . Lanitis, C. Draganova et al. 2004). For example, the five first parameters in Lanitis work controlled changes in lighting, 3D pose, expression and individual appearance. All the parameters were used as input for classification, 4 classifiers and 3 classification strategies were compared. The best results were obtained with a combination of both appearance and age specific optimized quadratic functions that explained the relationship between the face parameters and the age of a person. Appearance specific classification improved the results whatever the classifier in use, meaning that the aging path is different from one type of face to another. This idea gave birth to the third category of algorithms for age estimation.

Lanitis work also suggested that PCA is not the most efficient way to reduce the number of parameters before the classification considering that the first parameters from PCA were not related with age (lighting, 3D pose, expression). Fu and Huang (Rehnman 2007) have comparatively explored different methods of projection from the pixel image space. They have compared PCA to Neighborhood Preserving Projections (NPP), Locality Preserving Projections (LPP) and orthogonal LPP (OLPP). After projecting the images in the different sub-spaces, they applied a multiple linear regression to determine peoples’ age. From two databases made with 4000 men and 4000 women aged from 0 to 93 years, they demonstrated that as compared to PCA, NPP,LPP and OLPP have better performance and required a lower number of variables in the sub-space. LP and OLPP achieved a Mean Absolute Error (MAE) of 8 years while the error varies from 22 years to 10 years with PCA. These results are worse than those obtain by Lanitis, but we may guess that it is because a multiple linear regression was used instead of a more sophisticated non-linear algorithm.

I.3. Aging patterns based method

Xin Geng and his team have recently proposed a new approach to predict peoples’ age from images (Geng, Zhou et al. 2006; Geng and Smith-Miles 2007). They stated that each person has a specific aging pattern, which is determined by his/her gene as well as external factors such as health, living style and weather. They defined an aging pattern as a sequence of personal face images sorted in time order. Therefore, they used images from each individual at different age.

In order to build the aging patterns, they firstly transformed the images into feature vectors using AAM. Each image was consequently converted into ***n*** parameters and the aging pattern vector of an individual included the ***p*** pictures of the given observer. Then they built a representative model for the aging patterns using PCA.

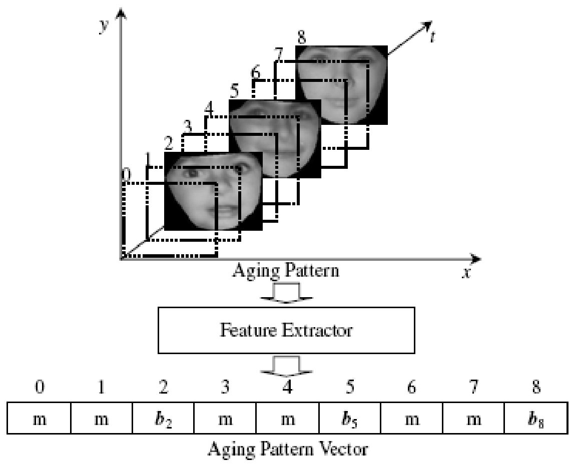


Figure 4-1: Vectorization of the aging pattern. (Geng and Smith-Miles 2007)

The 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'.

In order to predict the age from an unseen image, they firstly looked into the aging pattern subspace to find the suitable one. Then, they found the position of the given image within this pattern.

This algorithm was tested using the FG-NET Aging Database (Lanitis) made with 1002 face images from 82 different subjects. The authors reported a MAE of 6.77 years which was comparable to the accuracy of a human observer they involved in their experiment.

I.4. Comparative summary of the different approaches

Study	Number of subjects	Subject's age	Quality of the pictures	Compression algorithm	Classification algorithm	Validation strategy	Results
Kwon and Lobo (Kwon and da Vitoria Lobo 1994)	47 images	Babies-Young adults-Senior		Wrinkles extraction – face proportions	Threshold	Tested on 15 images	100% of accuracy for classification in 3 groups
Horng et al. (Horng, Lee et al. 2001)	230 images	Babies-Young adults-Senior		Wrinkles extraction – face proportions	Neural Network	230 images	81.6% of accuracy for classification in 3 groups
Hayashi et al. (Hayashi, Yasumoto et al. 2002)	300 images	Aged from 15 to 64 years and grouped by 10 years	Pictures acquired under controlled condition	Wrinkles extraction	Look Up Table	300 images	27% of accuracy for classification in 10 years age groups
Lanitis (A . Lanitis, C. Draganova et al. 2004)	330 images	Aged from 0 to 35 years	Variations in illumination, facial expression, positioning and make up	Facial model of appearance (AAM)	- Quadratic functions - Shortest Distance Classifier - Supervised Neural Networks - Kohonen Self-Organizing Map	80 images	MAE of 3.82 years with a combination of age and appearance specific optimized quadratic functions
Gandhi(Gandhi 2004)	818 images; 585 males and 233 females	Aged from 15 to 99 years	Variations in illumination, facial expression, positioning and make up. Resolution 160x160 pixels	None	Support Vector Regression	818 images in a 4-fold cross validation strategy	MAE of 9 years
Fu and Huang (Rehnman 2007)	4000 images; 2000 males and 2000 females	Aged from 0 to 93 years	Variations in illumination, facial expression, positioning and make up. Resolution 60x60 pixels	- PCA - NPP - LLE - OLPP	Multiple linear regression	2000 images	MAE around 10 years with PCA and NPP and 8 years with LLE and OLPP
Geng et al. (2007)	1002 images from 82 subjects (FG-net database) + 1724 images from 515 subjects (MORPH database)	Aged from 0 to 69 years	Variations in illumination, facial expression, positioning and make up. Resolution 5000 pixels	Facial model of appearance (AAM)	Linear Discriminant Analysis (LDA)	1002 images + 1724 images using a Leave-one-out strategy	MAE of 6.77 years with FG-net database and 8.07 with MORPH database

Table 4-1: Summary of the different works published on age prediction

II. Age prediction using Supervised Facial Model

II.1. Introduction

Our objective in this section is to build an algorithm for age prediction within the statistical based framework. A statistical based approach was preferred to a feature based one because it enables taking into account more age-related changes in faces. We couldn't test the aging patterns based approach since it require several images for each individual in the data set.

The statistical based approach necessitates a large dataset of images both to build the algorithm and to test its performances. The images in the training set should be chosen in order to include all possible factors of variability that could occur in the validation set. We believe that the images collected from Caucasian women during our aging study fits with these requirements to build a learning by sample algorithm for three reasons:

- There are acquired under standard lighting and positioning condition.
- The resolution of the images enables capturing small features.
- The population is uniformly distributed across age groups, with a minimum of 20 images per group of 10 years. That enables to take into account intra-groups variability.

When statistical models are built based on pixels values, it is required to align all the images in the data set so that the same facial features are roughly located at the same place. This step could be achieved by: (a) locating the main facial features in the images, (b) performing an image warping (or registration) to align all the images on a model. Facial features localization algorithms have been extensively discussed in previous works (Hammal and Caplier 2004; Vukadinovic and Pantic 2005; Belaroussi and Milgram 2006; Cristinacce and Cootes 2006). It appears that automatic localization is possible with a good accuracy. Since this problem was not an important point in the present work, we have not implemented these algorithms, assuming that facial features were already located using landmarks. We have subsequently used a rigid registration (rotation, translation and rescaling) to align the facial features on the average shape.

Faces encoding is the most critical step since it highly influences the outcome of the study. Based on images with thousands or millions of pixels, the objective is to find the best set of variables that capture information about facial aging. This task can be solved using data mining methods and/or image processing. Using data mining, each pixel is considered as a predicator variable and the problem is to predict the age from a large set of predicators. The most popular method available makes use of principal component analysis (PCA) to build a statistical model of the face. PCA allows finding a small subset of orthogonal variables that captures most of the variability within the facial images. However, PCA is not driven by the objective of the algorithm meaning that features needed for the regression task could be thrown away during the reduction of the number of variables. Consequently, we propose a Supervised Facial Model (SFM) based on the PLS regression.

The last step, age regression is done on a small subset of selected variables. Regarding the number of variables and their link with age, one can use linear or non-linear regression. Linear regression models are easier to build and to understand. Non-linear models are more powerful to describe complex relationship between the predicators and the predicted variables, but they should be used carefully since they could overfit the data.

This section presents our algorithm for age prediction. Firstly we describe the statistical model of facial appearance (AAM), which has been used by several authors to encode faces. Then we discuss the difference between PCA that is used for this model and PLS that we recommend for our own approach. Third, we introduce SFM and show how it can be used for age regression. Then we present the results obtain in our database. Finally, we discuss these results in the context of the state of art.

II.2. Statistical model of facial appearance

Cootes et al. have described the construction of statistical models of appearance in a seminal paper called “Active Appearance Model” (Cootes, Edwards et al. 2001). Given a set of images containing different occurrences of an object, the objective is to find a statistical model of the given object that will handle its variability in term of shape and texture. As an example, Cootes et al. build a model of human face, based on 400 images of faces.

These models combined a model of shape variation and a model of texture (also call appearance) variations. The models are built using a training set of labeled images, where main features are manually landmarked.

The statistical model of shape variation is built from the position of the landmarks on the training images. The principle is to align all the sets of landmarks into a common coordinate frame and to represent each individual by a vector x containing his landmarks positions. PCA is then applied to the landmarks position following the equation:

$$x = \bar{x} + b_s P_s \tag{4-1}$$

where $\bar{x}$ is the mean shape, P_s is a set of orthogonal modes of variation called eigenvectors and b_s is a set of shape parameters.

Before building the model of texture, all the faces are registered on the mean facial shape to reduce the impact of shape variations. The model of texture is subsequently built on pixel values following the same kind of projection than for the shape model:

$$g = \bar{g} + b_g P_g \tag{4-2}$$

where $\bar{g}$ is the mean vector of pixels values, P_g is a set of orthogonal modes of variations and b_g is a set of textural parameters.

The equation above is applied on each RGB channel for colored images.

The shape and the texture of the images can then be summarized by concatening vectors $\tilde{b}_s$ and $\tilde{b}_g$ obtained after the normalization of b_s and b_g respectively.

$$b = \begin{pmatrix} \tilde{b}_s \\ \tilde{b}_g \end{pmatrix} \tag{4-3}$$

Finally, a PCA is applied to vector b following the equation:

$$b = Qc \tag{4-4}$$

where Q are the eigenvectors and c is a vector of face parameters controlling both the shape and the texture models.

This model enables to summarize information related to the shape and the texture of the different occurrences of an object within a dataset. As an example, Cootes et al. summarized the variations in shape and texture in 400 images of faces containing 10,000 pixels and labeled with 122 landmarks with only 80 parameters. These parameters enable to explain 98% of the observed variations. The selected parameters enable to capture information related to head titling, facial expression, and identity. The Figure 4-2represents the first parameters and the information they capture.

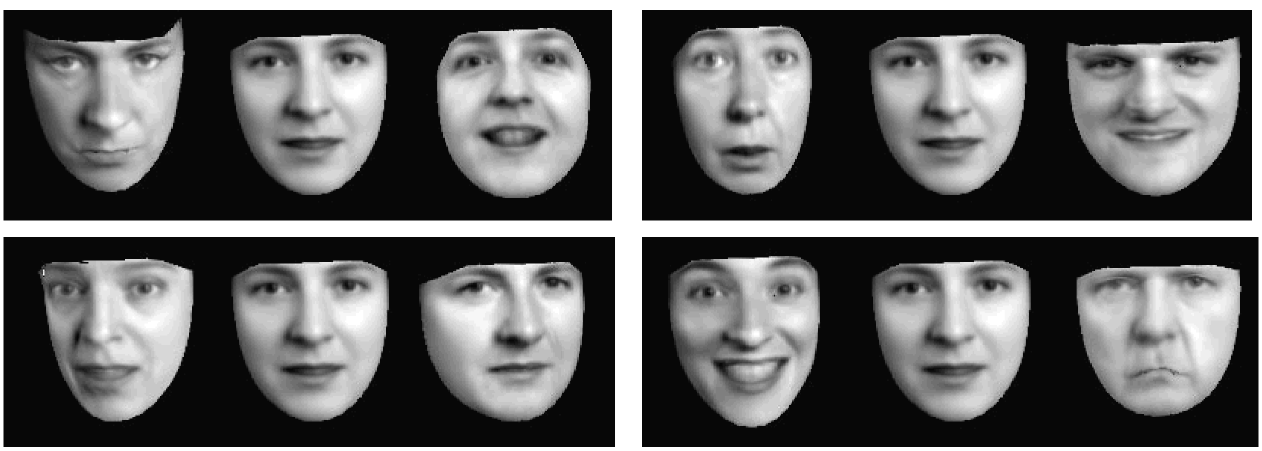


Figure 4-2: First four parameters of appearance and their variations (± 3 sd) (Cootes, Edwards et al. 2001)

Regarding the age prediction problem, one is especially concerned with parameters related to age. Since PCA is not supervised, this statistical model of appearance does not enable to make sure that it would be the case. Thus, we developed the SFM based on PLS regression which guarantee that the parameters that are obtained are the most relevant with respect to age. The comparison between PCA and PLS is provided in the annex 3.

II.3. Supervised Face Model

The supervised face model is derived from the shape and texture model, taken into advantage the fact that PLS is more powerful than PCA for rank reduction in regression problems. The main steps of the proposal algorithm can be summarized in Figure 4-3.

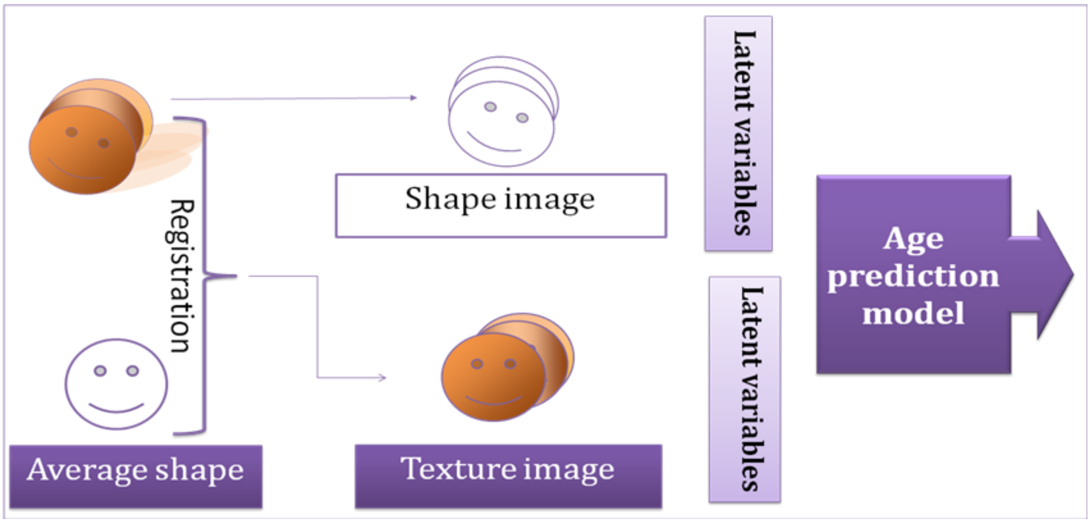


Figure 4-3: Diagram of the age prediction algorithm

First step: the images are registered so that the main facial features are aligned. Second step: two PLS models are built, one that captures information related with the shape of the face, and the other that encodes textural and color information. Third step: A final PLS regression enables to predict the age from the latent variables derived from the shape and textural latent variables

Given a set of facial images made with S pixels, the learning phase is made as followed:

1. The images of the training set are manually labeled, the main facial features being landmarked as shown in Figure 4-4. N landmarks points $P_i(x_i, y_i)$ provide a shape vector $X_{shape} = (x_{1..n}|y_{1..n})$ of $2 * N$ elements.

2. The mean shape is calculated from the position of the landmarks in the images of the training set, $\bar{X}_{shape} = (\bar{x}_{1..n}|\bar{y}_{1..n})$

3. The images are registered to the mean shape using linear warping transformations (Rotation, scaling, translation)

4. A PLS regression is applied to predict the age Y from the coordinates of the landmarks after alignment. The number of latent variables capturing information in the position of the landmarks related with age while preserving a good generalization power is fixed using the leave-one-out validation protocol: $t_{shape} = (t_i^{shape})_{i<2N}$. These latent variables represent the shape model of the face.

5. PLS regression models are also built to predict age from pixels values of each of the three-color channel (Red, Green, and Blue). After a cross validation with leave-one-

out, a small subset of latent variables $t_{texture} = \left(t_j^{red} \middle| t_k^{green} \middle| t_l^{blue}\right)_{(j,k,l)<S}$ are selected. These latent variables keep the main variation of pixel distribution related with age and therefore summarized the textural information. They represent the texture model of the face.

6. For each picture, a vector is made with latent variables from the shape and texture models. The vectors are normalized in order to make comparable shape and textural information.

7. Since shape and textural information may be correlated, an additional reduction of dimension is operated by means of a new PLS regression model aimed again at capturing information related to age from both the shape and texture latent variables. It gives rise to the overall face model. As for the other models, the number of latent variables to be selected is determined by leave-one-out.

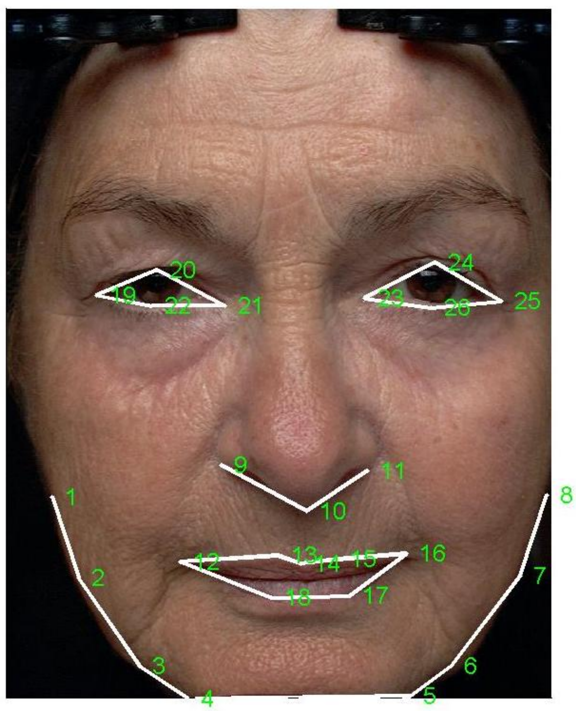


Figure 4-4: Manual landmarks on faces

The main facial features are localized by 26 points. Height landmarks are used for the jaw line, 7 landmarks for the mouth, 3 landmarks for the nose and 4 landmarks for each eye.

II.3.1. Image registration

Image registration or warping is the process of digitally manipulating an image so that any shapes portrayed is distorted to match a target. Image registration enables to reduce the variability related with the tilling of the head and to align the main facial features. It enables to give the same meaning to a pixel at a given position whatever the image. The images from our database were acquired with a precise positioning system so that the magnitude of the registration is expected to be small.

Image warping needs a source and a target shape, which can be described using landmarks on the image. In our case, the source shape is defined by 26 landmarks manually marked on each facial image and the target shape is the average position of the landmarks from the entire population. The number and the position of the landmarks were chosen according to the results of the previous chapter where we highlighted the influence of the jaw line, the size of the mouth and the eyes for age prediction.

Many warping techniques have been proposed by researchers, most of them consisting of three main steps:

(i) Finding the transformation function that enables to calculate the displacement of a pixel from the source image to the target or the contrary

(ii) Generate the target image using the transformation function

(iii) Re-sampling the output image to have a smooth output image

The warp function gives the correspondence between two sets of points, from the source and the target image respectively. Depending on the purpose of the warping, this function can produce a rigid or an elastic transformation.

In our case, we use the average position of landmarks in the whole learning set of images as a target. In order to keep the ratio in the face, a solid registration (rotation, translation and scaling) is applied following the equation:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha * \cos\theta & -\alpha * \sin\theta & t_x \\ \alpha * \sin\theta & \alpha * \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} . \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{4-5}$$

where x and y are the coordinates of a given point in the image we want to register while u and v are the coordinates of the same point after the warping. The warping is defined by four parameters: the rotation angle θ; the translation distances t_x and t_y; the scale factor α. These four parameters are calculated from the position of the landmarks in the labeled images using an optimization procedure.

After the position of each point from the initial image is calculated in the final one, a re-sampling of this image enables to smooth the output. A bilinear interpolation is used for image re-sampling.

II.3.2. PLS model of face

The SFM that we propose is based on the learning by sample approach. The main idea within this framework is to use a large set of data to establish general rules that would describe a population. As previously mentioned, the dataset should be large enough to take into account most of the variability factors expected from the population of the study. Furthermore, a validation procedure should be set up to make sure that the proposal model is not over fitting the learning dataset. Since a regression is done on the parameters extracted by SFM, two validation steps are required. Firstly, the SFM model is validated on 2/3 of the dataset. The regression model for age prediction is validated on the last third of the data.

PLS model of shape

The PLS model of shape describes the changes on the position of the landmarks which are related with age. After the faces have been aligned on the average face, a PLS regression is performed on a dataset X_{shape} containing 2/3 of the individuals in the database and N variables; each variable being the x or y position of a landmark in a given image. All the variables in X are normalized (set mean is set to zero and standard deviation to 1) so that they similarly contribute to the shape model. The adequate number of latent variables is fixed using the leave-one-out procedure, by defining the subset that minimizes the RMSE in age prediction for the testing set of data.

PLS model of texture

The PLS model of texture describes the changes on the texture of the face related with age. We call changes in the texture the variations in pixels distribution. These variations may carry textural information like wrinkles, but they may also carry shape information. As an example, the pixels of the lips are redder than those of the skin. Since the textural information is carried by the value of each pixel in the images, the shape of the lips will be visible on the textural data.

A PLS regression is performed for each color channel in the image. Given a color channel (Red for example), the dataset X_{red} is made of the red values of all the pixels. As for the shape model, the number of latent variables for each of the three texture models is selected using leave-one out.

Overall PLS model

The overall PLS model summarizes the information obtained from the shape and the texture models. The latent variables obtained from these 4 models (shape, red, green and blue textures) could have been directly used for age regression. However, the shape and the texture models are highly correlated as shown in Table 4-2. Particularly, the information derived from the different color channels is highly redundant. A PLS regression enables to build a smaller subset of variables that allows capturing most of the information contained in the shape and the texture models. In particular, the final model reduces redundancies in the texture models by combining the color information contained in the different color channels.

	Shape	Red	Green	Blue
Shape	1.00			
Red	0.80	1.00		
Green	0.79	0.97	1.00	
Blue	0.80	0.98	0.98	1.00

Table 4-2: Correlation coefficients between the first components of the shape and texture models.

The different models are built using 2/3 of images in our database. For each model, images are projected onto the first variable. Then a Pearson coefficient of correlation is calculated.

A PLS regression is performed on a vector $X_{overall} = \left(\tilde{t}_i^{shape} \middle| \tilde{t}_j^{red} \middle| \tilde{t}_k^{green} \middle| \tilde{t}_l^{blue}\right)_{i<2N,(j,k,l)<S}$ made with the normalized latent variables selected from the previous models.

Leave-one-out is used as usual to select the best number of latent variables that minimize the RMSE in age prediction.

II.4. Age prediction

II.4.1. PLS regression

The PLS done on $X_{overall}$ enables to build a subset of latent variables $t_{overall}$ that summarize information related with the shape, the color and the texture variations in the face dataset. This PLS model also enables to directly predict people age. From the equations equations

$$\boldsymbol{Y} = \boldsymbol{tQ} + \boldsymbol{f} \tag{4-6}$$

and

$$\boldsymbol{X} = \boldsymbol{tP} + \boldsymbol{e} \tag{4-7}$$

Y can be obtained, knowing X and t The predicted value of Y, that we call $\check{Y}$ is given by the formula

$$\check{Y} = B.X + f \tag{4-8}$$

where

$$B = W^{*}.Q' \tag{4-9}$$

,

$$W^{*} = W.(P'.W)^{-1} \tag{4-10}$$

and W is recursively calculated as explained in the previous chapter.

This regression is linear since Y is expressed as a linear combination of X. The matrix B gives the contribution of each X variable in the PLS regression model.

II.4.2. Non linear regression

The latent variables $t_{overall}$ can also be used as input vectors for a non linear regression which may improve the age prediction result. As an example, we test a Multilayer Perceptron neural network as Lanitis (A . Lanitis, C. Draganova et al. 2004) did. The network has as input the latent variables $t_{overall}$ and as an output the age Y. One hidden layer containing 1/3 of the number of neurons in the input layer is integrated in the network. The transfer function used for each neuron in hidden layers is a log-sigmoid since it response is non linear and positive. The transfer function of the output neuron is linear.

III. Results from JNJ database

JNJ database is made of 173 pictures with a definition of 1000x1280 pixels. Each color channel thus contains $128x10^4$pixels. In order to decrease the computation time, we have decided to reduce the size of the images to 250x320, (i.e. 80,000-pixels). Our images remain at least 16 times “heavier” than those used in the literature. It does not seem that the reduction of definition does not jeopardize the efficiency of observers to give an age.

Validation strategy: The proposal method builds a SFM and applies a regression to the parameters extracted. The SFM are constructed using the dataset and should be validated. Similarly, the overall regression model should be validated on an unseeing dataset. Regarding the size of our database, we have decided to apply the following methodology in order to have enough data both for the training and the validation. The 173 images are randomly divided into two sets: a training set containing 116 images and a validation set containing 57 images. The SFM are built on the training set and validated by means of the leave-one-out. Then the regression model is built on this same training set and tested on the validation set. The overall performance of the algorithm is thus calculated for the validation set.

III.1. Supervised face model

III.1.1. Model of shape

In order to build the shape model, we firstly determine the optimal number of latent variables using loo on the training set of data. The first 4 variables provide a good tradeoff between the minimization of the RMSE and the requirement to reduce the dimension as much as possible (Figure 4.5) These 4 variables capture 64.78% of the variance in the X_{shape} set and 82.88% of the variance in the Y set.

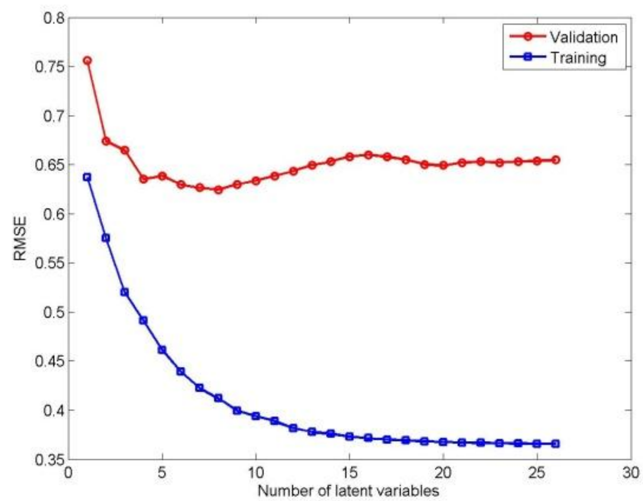


Figure 4-5: RMSE in age prediction based on the position of the landmarks as a function of the number of latent variables

Figure 4-6 shows the effects of varying the first two shape model parameters t_1^{shape} and t_2^{shape} through ±3 standard deviations as determined from the training set. It appears that the first component captures some of the changes linked with age such as the lips, the eyes and the jaw line shape. The second component gives the correction required to enhance the linear approximation brought by the first one. It particularly emphasizes some specific variations of the mouth.

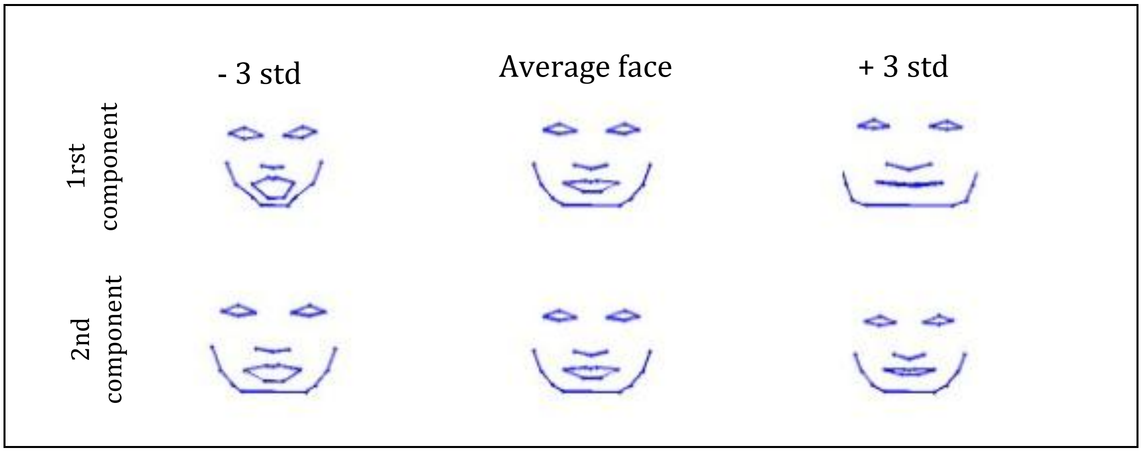
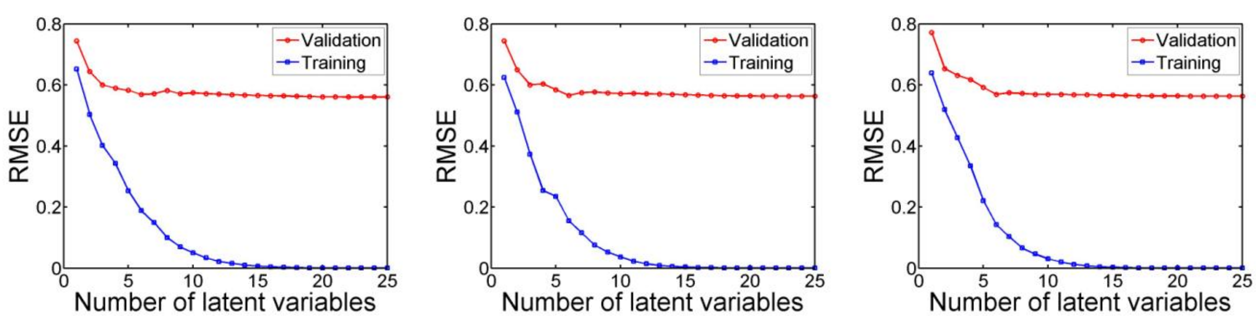


Figure 4-6: First two modes of shape variation

III.1.2. Model of texture

A PLS model of texture is built for each color channel. We also use loo to select the number of latent variables required for each model. The Figure 4-7 shows that six latent variables are optimal for each color channel. These variables capture 35.77%, 36.5% and 37.36% of the variance of the predicators, respectively for the red, green and blue channels. They capture 93.2%, 93.42% and 94.25% of the variance of the predicted variable.



a) red color channel b) Green color channel c) Blue color channel

Figure 4-7: RMSE in age prediction based on the pixels' values in each color channel as a function of the number of latent variables

The matrix of weights B enables displaying the contribution of each pixel to the prediction. We can thus see which part of the face mainly contributes to the texture model. Figure 4-8 shows the 4 first weights $b_{1..4}$ of the texture model for the red color channel. These weights are comparable to those of the other channels. It can be seen that the model mostly uses information coming from the eyes area and the jaw line.

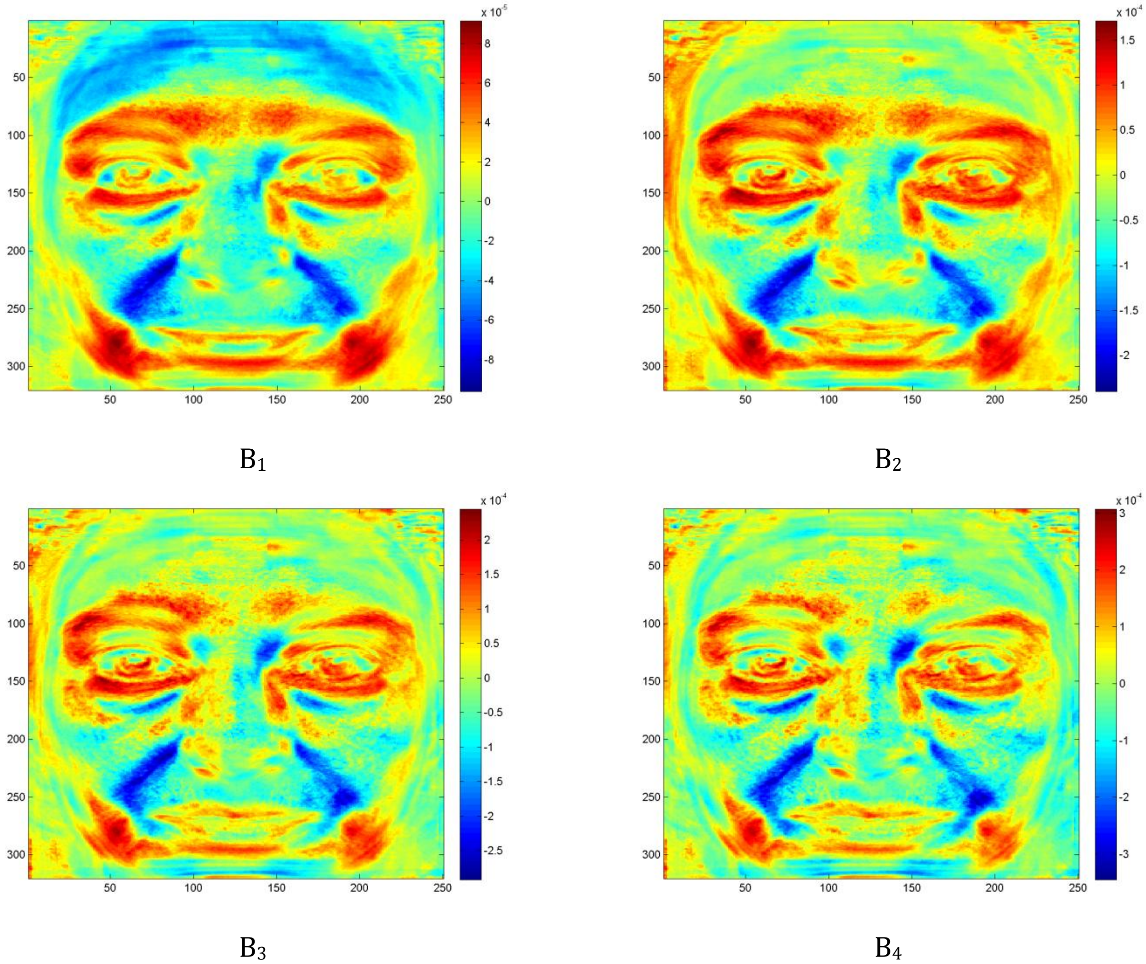


Figure 4-8: The four first weight b_1, b_2, b_3 and b_4 of the PLS model of texture

The areas in red and in blue are those who mainly contribute in the PLS model of texture (Red channel in the RGB color space)

III.1.3. Overall model

The overall face model is built using 22 variables; four from the model of shape and six from each model of texture. The number of latent variables to summarize this 22 variables set is selected using loo. It can be seen from the Figure 4-9 that 6 latent variables provide an optimal minimization of the RMSE in age prediction for the validation set.

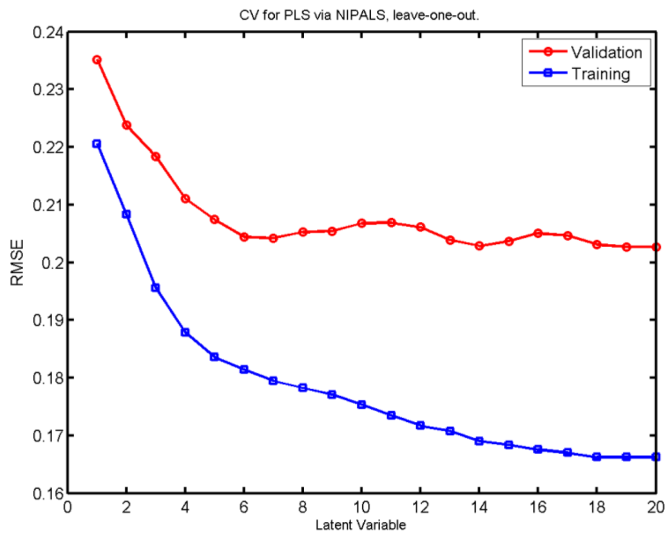


Figure 4-9: RMSE in age prediction based on the latent variables from the shape and texture models as a function of the number of latent variables

Figure 4-10 shows the effects of varying the first two face model parameters $t_1^{overall}$ and $t_2^{overall}$ through ±3 standard deviation as determined from the training set. It appears that the first component is positively correlated with aging. Particularly, we can notice the changes that occur on the upper eyelid, the bags under the eyes, the nasolabial fold, the lips thickness, the jaw line and the overall color. The second component is negatively correlated with age and enhances the linear approximation. It shows small changes in the cheeks and the nose and in the roundness of the face.

Figure 4-10: First two modes of face variation

The modes are calculated for each color channel and then combined to form a color image

Finally, the information contained in faces (80,000x3 pixels) is summarized with only six parameters. These parameters are used together for age prediction.

III.2. Linear regression using PLS

FSM sums up faces by a set of six variables that handle age related information about shape, color and texture. Using these variables, a PLS regression can be used to build a model for age prediction. A usual, validation is carried out using a different set of images.

The age for people in the training set is predicted with a high accuracy: the model exhibits a Mean Absolute Error (MAE) of 2.35 years and a Pearson R of 0.98 ($p<0.001$). These results show that FSM is an accurate representation of the age related attributes in the training dataset.

The predicted ages are less correlated with real ages in the validation set (Figure 4-11). The MAE is 5.98 years and the Pearson R of 0.85 ($p<0.001$). Seven subjects over 57 (12.3%) show an error of more than ten years.

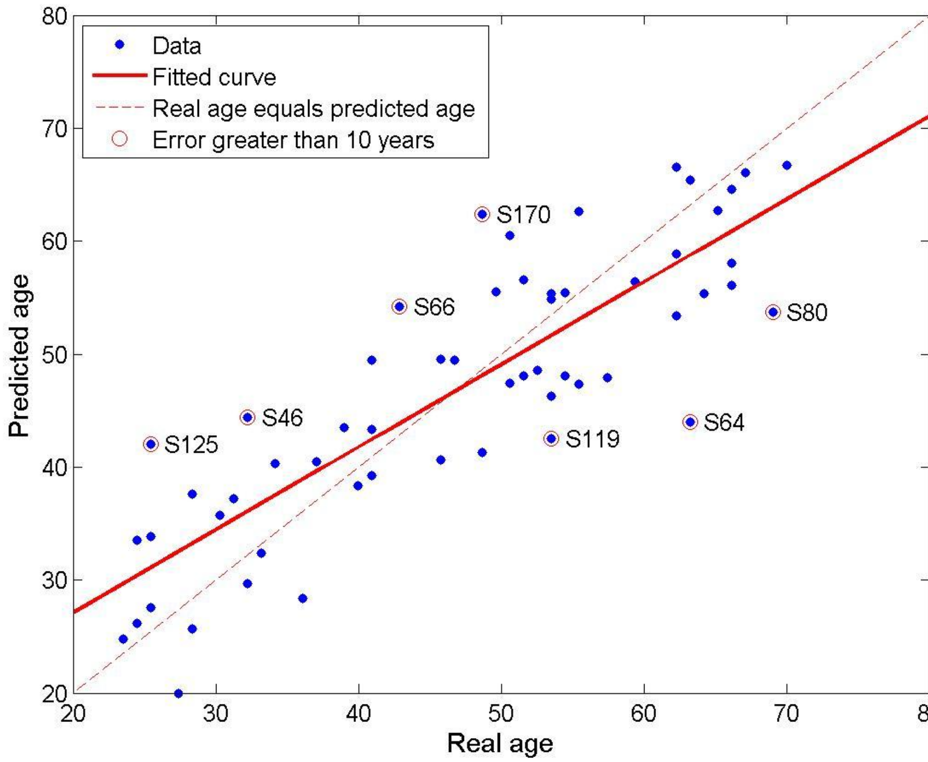


Figure 4-11: Correlation between real and predicted age for the validation set of data

The residual plot (Figure 4-12) shows that there is a systematic error (Figure 4-12). Older graders are over-aged while under graders are under aged.

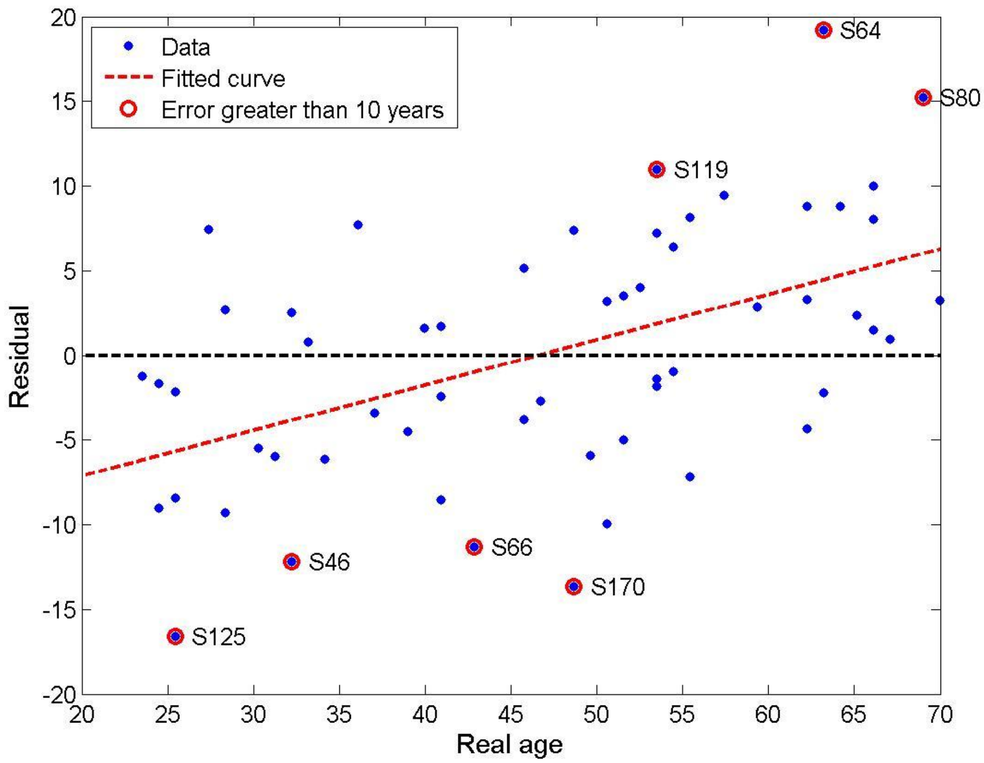


Figure 4-12: Residual plot of the age regression model

The mean of the residual is 0.12 and its standard deviation is 7.3. Younger people are over-aged while older graders are under aged.

Seven volunteers are given an error greater than ten years in age prediction (Figure 4-13). We can notice that our classifier has the same behavior than human graders for six out of these seven. The most unexpected error is achieved for subject S125. She is 24 years old and the system gives her 41 years. This large error is certainly caused by the presence of visible hairs on her picture.

Figure 4-13: Subjects with an error of more than ten years in age prediction

A pink background is used for subjects whose age is under estimated. A blue background is used for subjects whose age is over estimated.

III.2.1. Comparison with human graders

Comparison with human graders

The MAE of our classifier is 5.98 years, while for perceived age it is 4.1 years. Particularly, it can be seen from the graph below that our method tends to make more big errors than human. However, it must be pointed out that the perceived age from human graders is an average estimation of multiple graders. It is therefore more robust that individual estimations. The perceived age does not exactly reflect a single human behavior, particularly for big error in age prediction. In average the age of no subject is over or under estimate by more than 12 years by our group of graders. But when looking at individual values, we see that all the graders have over or under estimated the age of at least one person by more than fifteen years. The MAE of individual graders varies from 4.3 years to 9.3 years. The Figure 4-12 shows that classifier performs similarly than the best human grader.

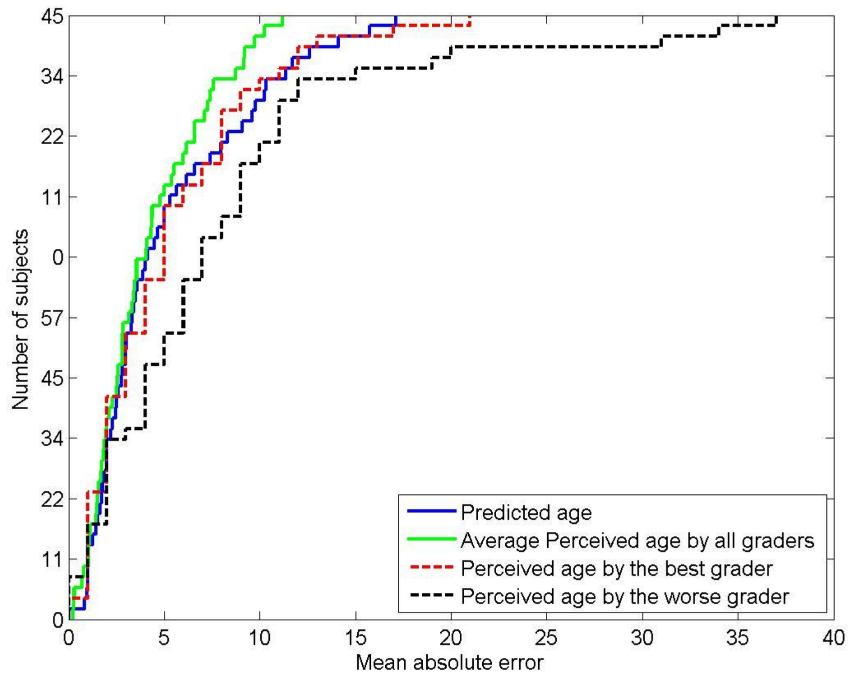


Figure 4-14: Cumulative distribution of the mean absolute error in age estimation

The blue curve is the cumulative distribution of the difference between the real and the predicted age using PLS regression. The green curve is the cumulative distribution of the difference between the real and the average perceived age given all human graders. The red and the black dotted line are respectively the cumulative distributions of the difference between real and perceived ages by the best and the worse human grader.

Predicting perceived age

We have also built a regression model to predict the perceived age instead of the chronological age. Fig 4-15 presents the results obtained for the validation set. RMSE is 0.8, MAE is 4.87 and the Pearson R is 0.85 ($p<0.001$). For this model, only four subjects are over or under estimated with more than ten years. This model is consequently better than the one built in the previous section.

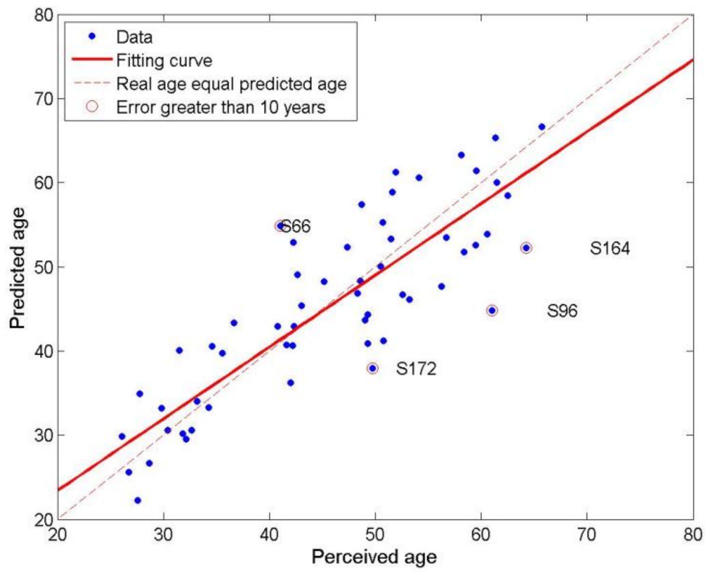


Figure 4-15: Correlation between perceived and predicted age on the validation set of data (PLS regression)

III.3. Comparison with AAM

The AAM model requires 32 components to summarized information contained in the facial images. Using these components, a linear regression enables to predict peoples' age with a MAE of 6.8 years on the validation set of data (Pearson R=0.8). Figure 4-15 shows the correlation between real and predicted age. Nine volunteers are given an error greater than ten years in age prediction. Surprisingly, the subjects S66 and S170 who were over-aged by our model are now under-aged. The contrary happens for the subject S119. In overall, the SFM based model is slightly better than the AAM based model.

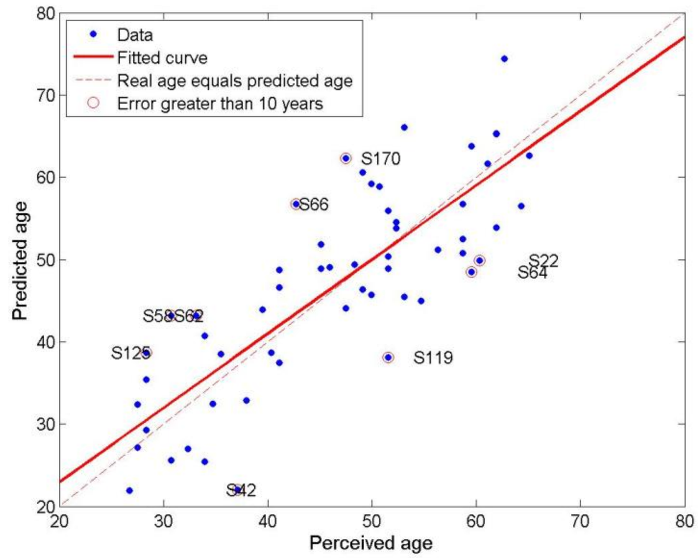


Figure 4-16: Correlation between perceived and predicted age on the validation set of data (AAM used for data image compression)

III.4. Non linear regression using neural network

The neural network enables to predict the age of people in the validation set with RMSE = 1, and MAE = 6.03. The Pearson correlation coefficient R is 0.85. Ten individuals are over or under estimated (error> 10 years) as it can be seen on figure 4-16. Surprisingly, these results are not better than those obtained using a linear model.

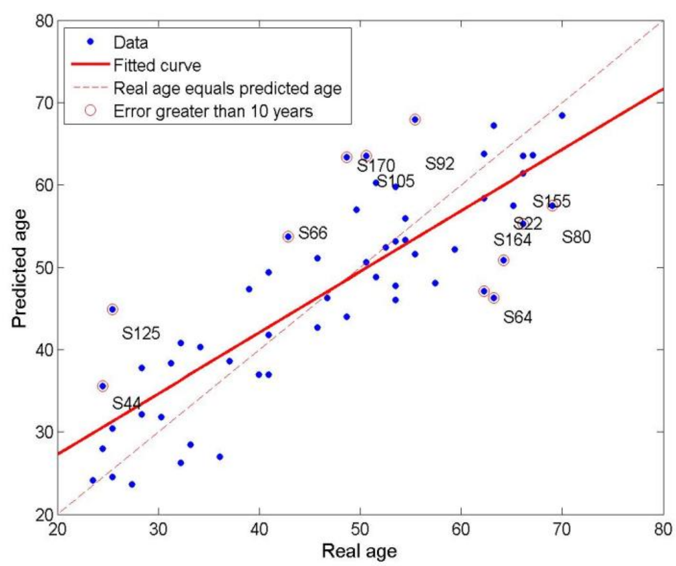


Figure 4-17: correlation between real and predicted age on the validation set of data (neural network model with PLS latent variables).

IV. Discussion

In this chapter, we have proposed an age prediction algorithm based on SFM, a new compression algorithm. SFM behave like AAM since it enables to summarize shape and textural information using a few set of components. These two models also allow visualizing the information carried by each component. They therefore help to better understand which part of the image is relevant regarding the classification task. However, SFM is driven by the goal and requires a fewer number of components than AAM. As an example, the final SFM model only used six components while the AAM needed 32 components. This difference can be explained by the criteria used to select the number of components in AAM. With this model, the required number of components to preserve most of the information in the face (usually 95% of the variance) is kept. Thus, information not related to the classification task remains in the components, leading to a bigger number of components. Moreover, the PLS algorithm computation requires a lot less memory than PCA algorithm computation (annex3).

The first component of the model of shape highlights the main attributes which are taken into account for the prediction. It can be seen that the model mainly uses the shape of the mouth, the eyes and the jaw line. The first component of the texture model highlights the areas around the eyes, the jaw line, the mouth and the nasolabial fold. These results are consistent with those obtained from human observers. The first component of the overall model of face confirms the importance of all these attributes and also shows a slight variation of the color from pink to yellow. This variation is consistent with graders observation. However, the proposed model does not capture information related to the small wrinkles and to the tone unevenness. Although these facial attributes are well correlated with age, they are not picked up by the model because they are not aligned from one image to another within the dataset. There are thus considered like details and the PLS models does not take them into account.

The age is predicted with a MAE of 5.98 years. This result is better than those obtained by Fu and Huang (Fu, Xu et al. 2007); it can be compared with the MAE of 6.22 years obtained by Geng et al. (Geng and Smith-Miles 2007) and is worse than those obtained by Lanitis (A . Lanitis, C. Draganova et al. 2004) using a combination of classifier. However, the accuracy of the proposal model is comparable with the best human grader in our population.

Not surprisingly, the model that predicts apparent age is better than the one which predict real age. It means that volunteers who look younger or older than their age introduce biases in the model which predict real age. This result suggests that perceived age should be used as the target when building age regression algorithms.

Using a non linear classifier does not improve the performance of the NN model. This may be due to the relative small number of individuals in the learning set, comparatively to what is usually required for NN.

V. Conclusion

The SFM is an adequate tool to summarize age-related changes. It also allows visualizing the main important attributes used for aged prediction. This model gives an interesting insight to the cosmetic dermatologist since it enables to build a picture summary of the changes that occur in facial appearance. While using this model, it clearly appears that facial aging is a multi-dimensional problem. Therefore efficient face rejuvenation should simultaneously address several signs of aging in order to produce an “aesthetically” consistent whole.

The proposed prediction model shows performances that matches human accuracy. These promising results could be improved by better handling wrinkles and tone unevenness information.

Conclusion and perspectives

This thesis work has allowed us to study facial aging under three different angles.

Firstly, we have explored the anatomical aging in order to better understand the changes occurring on face with age. We have set up and run a clinical study on women volunteers and have quantified the evolution of their facial signs of aging. This study provides a deeper knowledge of Caucasian women facial aging.

Secondly, based on the data collected, we have explored the perceptual aspect of facial aging. The importance of each facial attribute for age prediction was assessed using an original methodology. We have found that:

- Age perception is biased by own-gender and own-age group effects.
- The main attributes that drive the perception of age from faces are the lip area, the eye area and the skin tone.
- Regarding the age of the observer, young observers focus more on color uniformity, nasolabial fold and dark circle, while older observers look more at lips drawing and thickness and eyes opening.

Taken as a hole, the first part of the thesis (chapter 1, 2 and 3) has given us a better understanding of the human processing of age-related information. They have enabled us to verify the importance of the shape, the texture and the color of the face in age prediction.

In the last part of the thesis, we have proposed a model that learns how to predict people's age from a front face picture. This model is based on a new supervised compression algorithm of facial images named SFM. Thanks to the use of PLS, the SFM has demonstrated a higher power of compression comparatively to AAM described in the literature. Moreover, this model requires less machine memory and can therefore handle bigger images.

The final age prediction algorithm has shown performances comparable with Human. However, a non linear approach did not allow improving the results of the regression as expected.

An improvement to the proposed algorithm would be to use kernel PLS, which is non linear, when building the SFM.

Among the numerous perspectives that can arise from this work, we want to focus on three main ideas that can be implemented in the domain of image processing:

- Improving the age prediction model: The performance of the proposal model can be improved by:
    - Including more people in the dataset. In learning by sample approaches, the quality of the model often depends on the number of samples in the training set. In our case, our population was sampled in order to have a uniform distribution of subjects according to their age. However, the size of our training set was not as high as those reported in the literature. By including more volunteers, we could also extend the lower and the upper boundary of our age range from the childhood to older age.
    - Better encoding the high frequencies information. The overall model of face does not capture small changes in wrinkles and tone unevenness since these are high frequencies information. A preprocessing of the image in the spectral domain (Fourier our wavelet) could enable to better taken this information into account when building the texture models.
    - Using non linear model for encoding faces variation. PLS only encode linear relations between the position of the landmarks or the pixels values and the age of the volunteers. Consequently, non linear information is lost. This information could be captured using a non linear algorithm such as the kernel PLS when building the models of faces. However, models built using non linear algorithms will not enable to visualize its components as the PLS does.
- Extending the proposal method for age simulation: Given a new facial image, we can always calculate its projection into the face model space. By modifying its components into this space, we can also simulate changes related with age. Therefore the SFM approach could be extend to the simulation of rejuvenation or gain of age.
- Extending the SFM to other classification problems: SFM enables to summarize the information contain in the facial images regarding or regression objective. The same methodology can be applied for different regression/classification tasks such as gender classification or facial expression recognition.

Annexes

Annex 1: Colorimeter measurements

Assessing skin color has always been important in dermatologist investigation and clinical practices. A number of instrumental methods have been used since 1920s (Brunsting and Sheard 1929). Some of them focused on the perceived color while the others focused on its spectrum and the chromophores underlying it. Here are some methods to evaluate skin color either in terms of perceptual stimulus or wavelength spectrum (Kollias and Stamatas 2002):

- Visual observation
- The tristimulus system
- Diffuse Reflectance spectroscopy
- Spectral imaging
- Digital imaging

In the present study skin color is measured using a VAS and tristimulus colorimeter (Minolta chroma Meter CR200, Osaka, Japan). The Minolta Chroma Meters CR200 is a compact tristimulus colorimeter. The instrument consists of a control unit and a measuring head. The measuring head is featured with a pulsed xenon arc lamp that emits an intense white light covering the entire visible spectrum. To compensate for the variability of the xenon flash, the measuring head contains a double-beam feedback system, which measures both incident and reflected light. The color of the reflected light is analyzed by 3 high-sensitivity silicone photocells that are filtered to match the CIE standard observer curves for the blue (450 nm), green (550 nm) and red (610 nm) primary colors. Another 3 identically filtered photocells are used for the double-beam feedback system. The sensitivity of the photocells is adjusted to match the retinal response of CIE 1931 standard curves. The signal of the photodiode therefore simulates the human “eye-brain response”.

The measured skin color is expressed in the L*a*b* system(Weatherall and Coombs 1992). L*a*b* is designed to approximate human vision. It aspires to perceptual uniformity, and its L* component closely matches human perception of lightness. a* and b* are related to the color; a* varying from green to red and b* varying from blue to yellow. A representation of the L*a*b* color space is shown below (Figure 0-1)

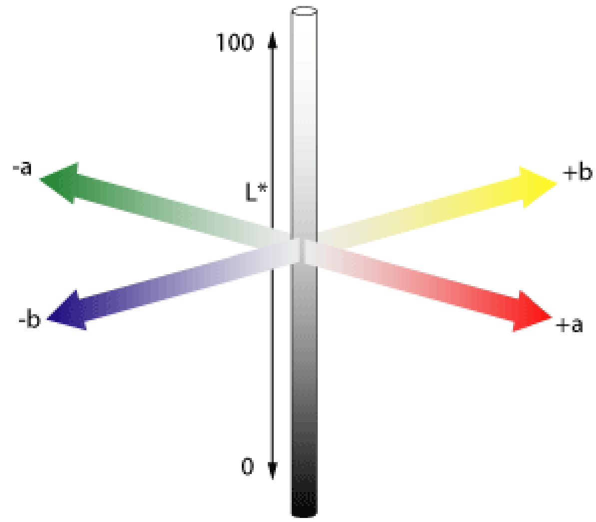


Figure 0-1: L*a*b* color space representation[4]

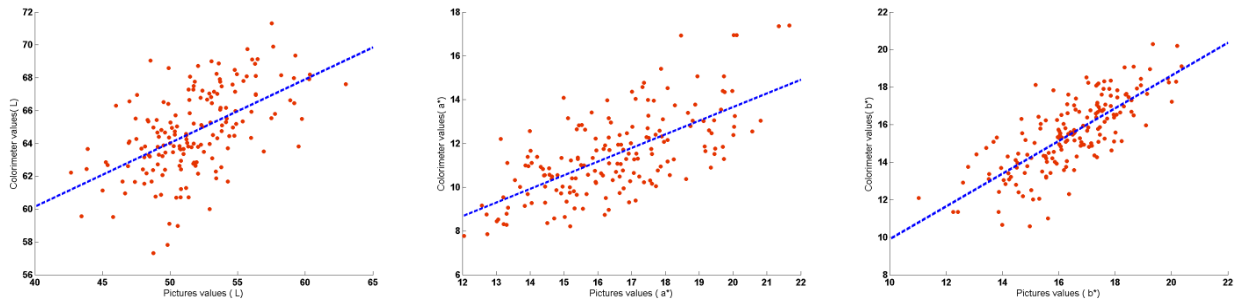
The colorimeter values L*a*b* are used to check for the consistency of the visual assessment of colors. The color VAS goes from pink to yellow. A correlation is made to link the given color from the scale to the measured one from the colorimeter.

Color perception

The above methodology for building picture scales was not found suitable for color measurements. We consequently decided to explore the link between color perception, measurements with colorimeters and digital images. Since the perceived color depends on the ambient light, its visual evaluation was done under a sunlight source. However, visual grading of skin color remains a difficult task because the color space is a 3-dimension space. In order to check the consistency of our expert to label colors, volunteers' skin color was also measured using a colorimeter. The colorimeter uses three colored LEDs as light source and measures the reflected light from the skin. The skin color was assessed in the L a*b* color system. The color was also evaluated from the facial images. A square of 268x268 pixels was cropped from the checked area and its average color was calculated in the RGB channel before being converted to the L a*b* channels for comparison.

The values from the colorimeter were highly correlated with those from the images ($p<0.005$) in both L, a*, b* as shown in Figure 0-2. Therefore we only use the values from the colorimeter for the remaining of the analysis.

[4] http://www.specialchem4coatings.com/tc/color/index.aspx?id=cielab



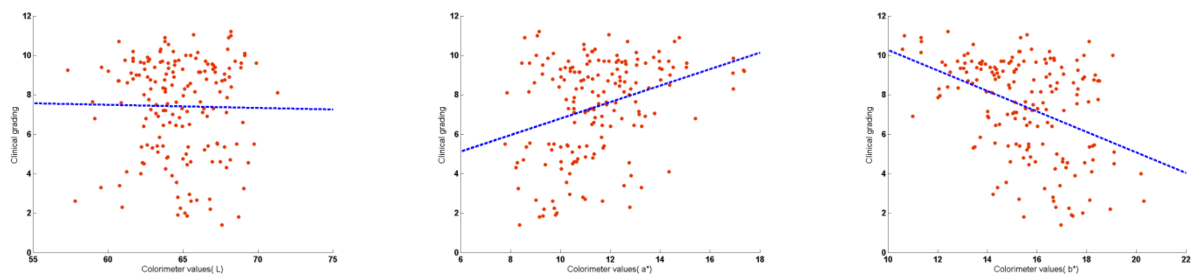
a) L* (Pearson R= 0.55, p<0.0001)

b) a* (Pearson R= 0.66, p<0.0001)

c) b* (Pearson R= 0.76, p<0.0001)

Figure 0-2: Correlation between colors measured from pictures and skin using colorimeters

The Figure 0-3 compares the measured L*a*b* color values with the clinical grading. It appears that perceived color measured with VAS is only related to a* and b*. This observation is not surprising since the proposal scale goes from pink to yellow, which are related to the chromaticity and not to the brightness in contrast with the L* value.



a) L* (Pearson R = 0.01, p=0.7892)

b) a* (Pearson R = 0.32, p<0.0001)

c) b* (Pearson R = -0.41, p<0.0001)

Figure 0-3: Correlation between the perceived color and the measurements

With the value of L* being fixed to its average within the population, one can look at the correspondence between a*, b* and the color from the VAS. Figure 19 represents the regression plane between the VAS and the a*,b* color channels, the color of the regression plane representing the given grade in an arbitrary color map. The VAS is highly significantly correlated (R=0.48, p<0.0001) with a* and b* values, meaning that the perceived color is linked with the measured one.

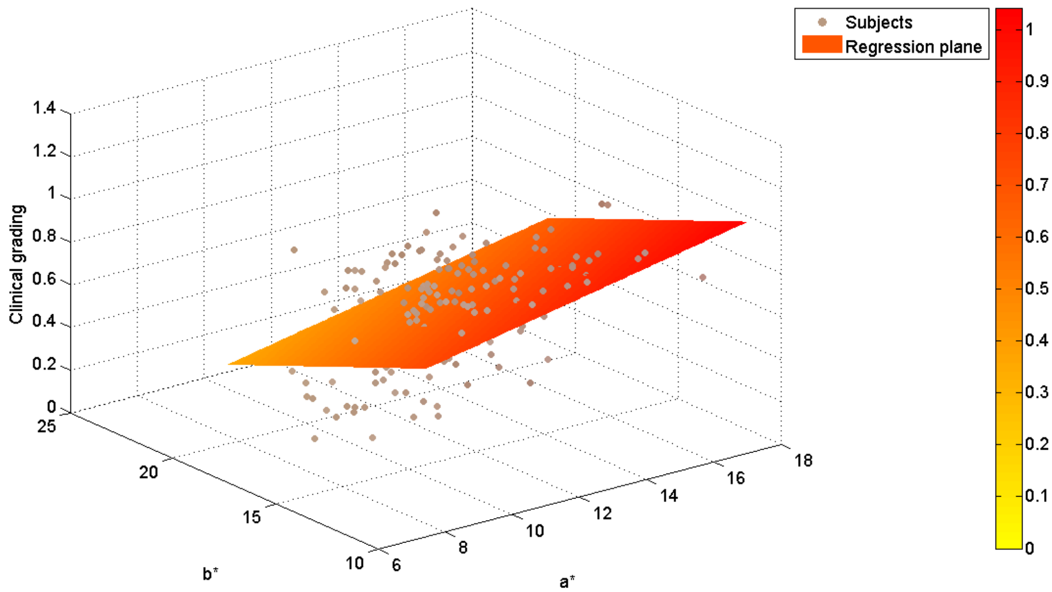


Figure 0-4: Multiple correlation between the a*,b* color values and the perceived color

Pearson R=0.48 p<0.0001

For a given value of the perceived age, one can represent a 2-dimension color space by displaying all the colors generated by varying a* and b* from their minimum to their maximum for the individuals with the given perceived value, L* being fixed to its average in the population. The color space of twelve grades evenly placed between the minimum given value (0.12) and the maximum (1) are represented in Figure 0-5. It can be seen that the graded colors effectively span from yellow to pink.

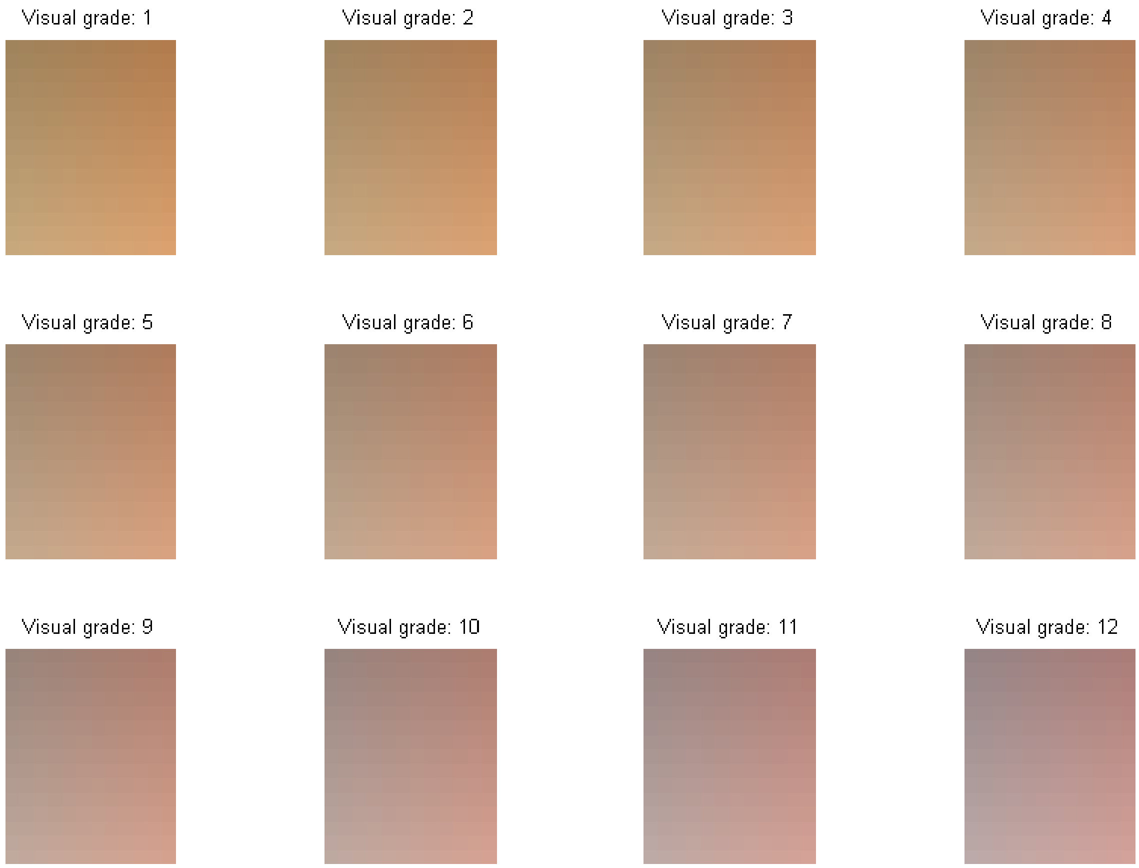


Figure 0-5: Representation of the perceived colors using colorimeters L*a*b* values

The grades are evenly over the [0.12 ,1] interval. For each grade, is plotted in the x-axis and y-axis correspond to a and b* respectively. a* and b* varies from their minimum to their maximum values within the subjects with a given grade. L* is fixed to its average within the whole population.*

Using VAS, we have found that skin color was significantly correlated with age ($R=0.66$, $p<0.0001$). This result is confirmed when doing a multiple correlation between VAS and a*,b* color values ($R=0.34$, $p<0.0001$), even if the correlation is lower.

Annex 2: Model validation

Quality of the model (Leave-one-out)

With latent variables models such as PLS, the best model is not necessarily the one that fits the data best. The generalization power of the model, which means its ability to fit an unknown individual, is also important. Since the number of variables is often high comparatively to the sample size, a model that over-fits data may not have a good predictive ability. Consequently, the quality of the observed data fit cannot be used to choose the number of latent variables to use. The number of latent variables must be chosen on the basis of how well the model fits observations not involved in the modeling procedure itself.

One method to choose the number of latent variables is to fit the model to only one part of the available data (the training set) and to measure how well models with different numbers of latent variables fit with the other part of the data (the validation set). This is called cross-validation.

When the number of individuals in the dataset is not high enough to split the data in two sets, one can use the leave-one-out (loo) procedure. Within the loo procedure, $n-1$ individuals are used to build the model (training set) and the last individual is used to validate it (validation set). The procedure is repeated n times with each individual being iteratively excluded from the training set. Consequently given h the number of latent variables, n models are built using $n-1$ individuals for the training set and one individual for the validation set. The quality of the model with h latent variables will be defined as the average value from the n test sets.

The loo procedure was used in our case to identify the best number of latent variables that enables to minimize the Root of Mean Square Error (RMSE) in the validation set. For a model with h latent variables, the RMSE is calculated as follows:

$$RMSE = \sqrt{\frac{1}{\tilde{n}}\sum_{i=1}^{\tilde{n}}(y_i - \hat{y}_{hi})^2}$$

Where $\tilde{n}$ is the number of individuals in the validation set and $\hat{y}_{hi}$ is the predicted value for the $i^{ème}$ individual in this set given by the model with h latent variables.

Confidence interval for coefficient values (Bootstrap)

To fully validate a regression model, particularly when we want to interpret the regression coefficient to understand the relative contribution of each variable to the model, it is also necessary to calculate the confidence interval of each model coefficient. Unfortunately, since PLS introduced the latent variables to link the predicted variable with the predicators, the analytical estimation of confidence intervals is not trivial (Burnham, MacGregor et al.

2001). One of the methods proposed in the literature to achieve this goal is the bootstrap (Lazraq, Cleroux et al. 2003).

The bootstrap enables to estimate properties (such as standard error, confidence interval, percentile points, correlation coefficients…) of regression coefficients by measuring them when sampling from an approximating distribution. One standard choice to approximate the distribution of regression coefficients is to build several models from randomly resampled sets of observed data.

In the bootstrap procedure, L regression models are built by randomly sampling with replacement the individuals in the dataset, each model leading to a set of regression coefficients $B^k\left(b_1^k, \dots, b_b^k\right)$ where $k \in \{1, \dots, L\}$. Therefore for each regression coefficient, we obtain L estimated values from which standard error, confidence interval, percentile, correlations… can be easily calculated.

The bootstrap procedure is used here to estimate the 95% confidence interval of each regression coefficient. There is not any consensus on a method to select the number of models to build (L). Thus L is usually chosen as big as possible taking into account the computation time. Two hundred is a classical value.

Annex 3: Comparison between PLS and PCA

Given a predicator $X(x_n)$ made with n individuals described by m variables and $Y(y_n)$ the predicted variable, regression problems aimed to find a function $\hat{f}: x_i \in X \rightarrow y_i \in Y$. When the dimension of X is high and particularly when X has a reduce rank $\alpha < m$, one may need to find a subset of variables t called latent variables, that will carry most of the information from X.

To achieve this objective, PCA has been widely used since it was introduced by Pearson(Pearson 1901). PCA is an orthogonal linear transformation that converts the data to a new coordinate system such that the projection on the data on each axis (also called latent variable) is a decreasing function of the rank of the given axis. In PCA, the first latent variable will therefore capture the maximum of the data variance; followed by the second latent variable, and so one. This property has enabled to extensively use PCA for dimensionality reduction in pattern recognition problems. Usually, a subset of latent variables is selected from the first variable up to a certain rank. The number of latent variables is usually chosen to have the maximum of variance captured with a minimum of variables.

The PCA algorithm requires calculating the covariance matrix

$$C = \frac{1}{n} X.X' \tag{0-1}$$

where X' is the transpose of X. Then C is diagonalized leading to a eigenvectors matrix V and a eigenvalues matrix D. D and V follow the equation:

$$V^{-1}.C.V = D \tag{0-2}$$

and D is a diagonal matrix of eigenvalues, each been associated with a eigenvector. The eignvalues are ordered and the eigenvector corresponding to the highest eigenvalue is chosen as the first axis. The second eigenvalue enable to choose the second axis and so one.

PLS is an alternative to PCA which has been proposed by Wold et al(Wold, Ruhe et al. 1984). The objective of the PLS is to find the set of variables t that will capture the maximum information from X while explaining the maximum of variance from Y. Therefore t is both related to X and Y as:

$$Y = tQ + f \tag{0-3}$$

$$X = tP + e \tag{0-4}$$

where t contains up to α latent variables $(t_1, t_2, \cdots, t_\alpha)$ defined as a basis for the reduced-rank component of X and Y; P and Q are matrices of coefficients relating t to X and Y respectively; and f and e are considered to be random errors. The latent variables t are built to be orthogonal; so that $cov(X, t_h) > cov(X, t_{h+1})$ and they maximized $cov(t_h, Y)$.

De Jong has demonstrated that the correlation coefficient of the regression of t^{PCA} on Y are always lower than those from the regression of t^{PLS} on Y, meaning that in a supervised framework, PLS pre-processing should be preferred to PCA. The Figure 0-6 shows a comparative example of projection both on PCA and PLS space. Data are distributed in a plane and are categorized in two groups. The first component in PCA captures the maximum of variance from data distribution while the first component from PLS capture the maximum of variance within the groups of subjects. Even if the projection of the data on the first PCA component or the first PLS component will not enable to separate the two categories of individuals in the population, the first PLS axis better take into account the difference between the two populations.

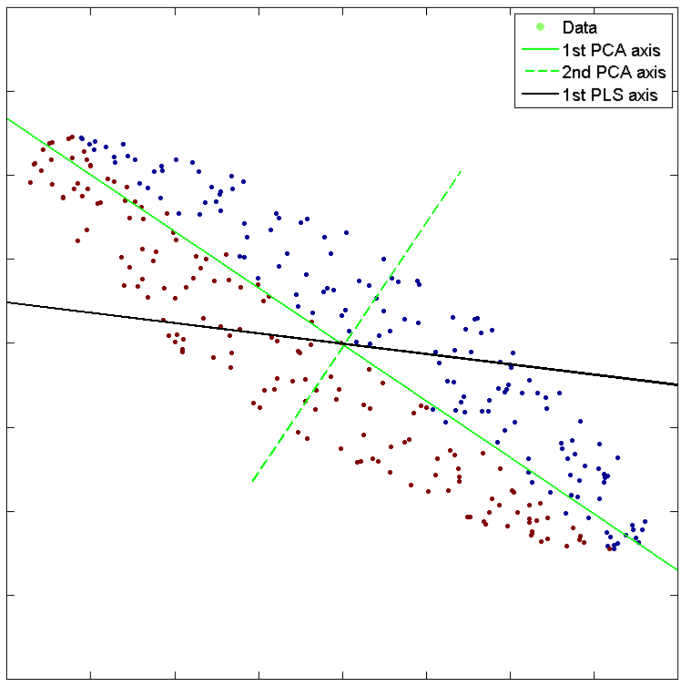


Figure 0-6: Comparison between PLS and PCA.

The individuals belong to two categories. By projecting the data on the first axis of the PCA, we capture most of the variability in the distribution of data. By projecting them on the first axis from the PLS, we have a better separation between the blue and the brown dots.

Another advantage of PLS versus PCA is the memory needed during computation. The PLS latent variables are built recursively as explained in the previous chapter by calculating a weight matrix W only using covariance of two columns of data. On the contrary with PCA, the covariance matrix C of all the variables in X is calculated. This matrix is of size $m * m$. Particularly when there are a lot of variables C requires a huge memory space. As an example, to calculate the latent variables of an image made with 80,000 pixels (250x320), it is required to handle a table of $64x10^8$ digits which represents 6103MB. With PLS, the biggest table would have a size of $m * a$ where $a \leq m$ is the reduce rank of the X table. Particularly when the number of individuals n is smaller than the number of variables m; $a < n$. With 200 individuals in the database, the maximum required size using PLS would be $16x10^5$ digits, which is required $4x10^3$ less space than PCA.

Finally, the PLS also enables to visualize the eigenfaces and therefore to understand which part of the image is mainly contributing to the prediction.

References

A . Lanitis, C. Draganova, et al. (2004). "Comparing different classifiers for automatic age estimation." Systems, Man and Cybernetics, Part B, IEEE Transactions on **34**(1): 621-628.

Anastasi, J. S. and M. G. Rhodes (2006). "Evidence for an own-age bias in face recognition." North American Journal of Psychology **8**(2): 237-252.

Atwood, D. A. (1971). "Reduction of residual ridges: a major oral disease entity." J Prosthet Dent **26**(3): 266-79.

Barbara A. Gilchrest, J. k. (2006). Skin aging. Skin aging, springer**:** 198.

Bartlett, S. P., R. Grossman, et al. (1992). "Age-related changes of the craniofacial skeleton: an anthropometric and histologic analysis." Plast Reconstr Surg **90**(4): 592-600.

Belaroussi, R. and M. Milgram (2006). "Face tracking and facial features detection with a webcam." Proceedings of the 3rd European Conference on Visual Media Production (paratre).

Bittorf, A., M. Fartasch, et al. (1997). "Resolution requirements for digital images in dermatology." Journal of the American Academy of Dermatology **37**(2): 195-198.

Blog, I. (1999). "A histological study of human wrinkle structures: comparison between sun-exposed areas of the face, with or without wrinkles, and sun-protected areas." British Journal of Dermatology **140**(6): 1038-1047.

Broniarczyk-Dyla, G. and E. Joss-Wichman (2001). "Ageing of the skin during menopause." J Eur Acad Dermatol Venereol **15**(5): 494-5.

Brunsting, L. A. and C. Sheard (1929). "The Color of the Skin as Analyzed by Spectrophotometric Methods: III. The Role of Superficial Blood." J. Clin. Invest **7**: 575–592.

Burnham, A. J., J. F. MacGregor, et al. (2001). "Interpretation of regression coefficients under a latent variable regression model." Journal of Chemometrics **15**(4): 265-284.

Burt, D. M. and D. I. Perrett (1995). "Perception of Age in Adult Caucasian Male Faces: Computer Graphic Manipulation of Shape and Colour Information." Proceedings: Biological Sciences **259**(1355): 137-143.

Carlsson, G. E. and G. Persson (1967). "Morphologic changes of the mandible after extraction and wearing of dentures. A longitudinal, clinical, and x-ray cephalometric study covering 5 years." Odontol Revy **18**(1): 27-54.

Castanet, J. and J. P. Ortonne (1997). "Pigmentary changes in aged and photoaged skin." Arch Dermatol **133**(10): 1296-9.

Charman, C. and H. Williams (2000). "Outcome Measures of Disease Severity in Atopic Eczema." Archives of Dermatology **136**(6): 763.

Cohen, J. (2003). Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences, Lawrence Erlbaum Associates.

Cootes, T. F., G. J. Edwards, et al. (2001). "Active appearance models." IEEE Transactions on Pattern Analysis and Machine Intelligence **23**(6): 681-685.

CrimeLib (2002). Age Progression. Criminal Minds and Methods. C. T. s. C. Library.

Cristinacce, D. and T. F. Cootes (2006). "Facial feature detection and tracking with automatic template selection." 7 thIEEE International Conference on Automatic Face and Gesture Recognition, Southampton, UK: 429–434.

Daniell, H. W. (1971). "Smoker's wrinkles. A study in the epidemiology of" crow's feet"." Ann Intern Med **75**(6): 873-80.

Day, D. J., C. M. Littler, et al. (2004). "The wrinkle severity rating scale: a validation study." Am J Clin Dermatol **5**(1): 49-52.

Dehon, H. and S. Bredart (2001). "An'other-race'effect in age estimation from faces." Perception **30**(9): 1107-13.

Dexter, F. and D. H. Chestnut (1995). "Analysis of statistical tests to compare visual analog scale measurements among groups." Anesthesiology **82**(4): 896-902.

Donofrio, L. M. (2000). "Fat Distribution: a Morphologic Study of the Aging Face." Dermatologic Surgery **26**(12): 1107-1112.

Doshi, A., A. Zaheer, et al. (1997). "A comparison of current acne grading systems and proposal of a novel system." International Journal of Dermatology **36**(6): 416-418.

Dunn-Rankin, P., G. A. Knezek, et al. (2004). Scaling Methods, Lawrence Erlbaum Associates.

Edwards, G. J., C. J. Taylor, et al. (1998). "Interpreting face images using active appearance models." 3 rdInternational Conference on Automatic Face and Gesture Recognition: 300–305.

Eveno, N. (2003). Segmentation des lèvres par un modèle déformable analytique, Thèse de l'INPG.

Evison, M. P. (2001). "Modeling Age, Obesity, and Ethnicity in a Computerized 3-D Facial Reconstruction." Forensic Science Communications **3**(2).

Fedok, F. G. (1996). "The aging face." Facial Plast Surg **12**(2): 107-15.

Freilinger, G., H. Gruber, et al. (1987). "Surgical anatomy of the mimic muscle system and the facial nerve: importance for reconstructive and aesthetic surgery." Plast Reconstr Surg **80**(5): 686-90.

Freiman, A., G. Bird, et al. (2004). "Cutaneous effects of smoking." Journal of Cutaneous Medicine and Surgery **8**(6): 415-423.

Friedman, O. (2005). "Changes associated with the aging face." Facial Plast Surg Clin North Am **13**(3): 371-80.

Fu, Y., Y. Xu, et al. (2007). ESTIMATING HUMAN AGE BY MANIFOLD ANALYSIS OF FACE PICTURES AND REGRESSION ON AGING FEATURES. International Conference on Multimedia & Expo. Beijing, China.

Gandhi, M. R. (2004). A Method for automatic Synthesis of Aged Human Facial Images. Department of Electrical & Computer Engeneering. Montreal, Canada, McGill University. **Master Degree:** 107.

Geng, X. and K. Smith-Miles (2007). "Automatic Age Estimation Based on Facial Aging Patterns." IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE **29**(12): 2234-2240.

Geng, X., Z. H. Zhou, et al. (2006). "Learning from facial aging patterns for automatic age estimation." Proceedings of the 14th annual ACM international conference on Multimedia: 307-316.

George, P. A. (2000). "The role of spatial and surface cues in the age-processing of unfamiliar faces." Visual Cognition **7**(4): 485-509.

George, P. A. and G. J. Hole (1995). "Factors influencing the accuracy of age estimates of unfamiliar faces." Perception **24**(9): 1059-73.

Guinot, C., D. J. M. Malvy, et al. (2002). "Relative Contribution of Intrinsic vs Extrinsic Factors to Skin Aging as Determined by a Validated Skin Age Score." Archives of Dermatology **138**(11): 1454-1460.

Hammal, Z. and A. Caplier (2004). "Eyes and eyebrows parametric models for automatic segmentation." Image Analysis and Interpretation, 2004. 6th IEEE Southwest Symposium on: 138-141.

Hayashi, J., M. Yasumoto, et al. (2002). "Age and gender estimation from facial image processing." SICE 2002. Proceedings of the 41st SICE Annual Conference **1**.

Helfrich, Y. R., L. Yu, et al. (2007). "Effect of smoking on aging of photoprotected skin - Evidence gathered using a new photonumeric scale." Archives of Dermatology **143**(3): 397-402.

Henss, R. (1991). "Perceiving age and attractiveness in facial photographs." Journal of Applied Social Psychology **21**(11): 933-946.

Hillebrand, G. G., M. J. Levine, et al. (2001). "The age dependent changes in skin condition in African-Americans, Asian Indians, Caucasians, East Asians & Latino's." IFSCC Mag. **4**: 259-266.

Horng, W. B., C. P. Lee, et al. (2001). "Classification of Age Groups Based on Facial Features." Tamkang Journal of Science and Engineering **4**(3): 183-192.

Just, M., E. Monso, et al. (2005). "Relationships between lung function, smoking and morphology of dermal elastic fibres." Experimental Dermatology **14**(10): 744-751.

Just, M., M. Ribera, et al. (2007). "Effect of smoking on skin elastic fibres: morphometric and immunohistochemical analysis." British Journal of Dermatology **156**(1): 85-91.

Kligman, A. M. and A. K. Balin (1989). Aging of human skin. Aging and the skin. A. M. Kligman and A. K. Balin. New York, Raven press: 1–42.

Kligman, A. M., P. Zheng, et al. (1985). "The anatomy and pathogenesis of wrinkles." British journal of dermatology(1951) **113**(1): 37-42.

Kligman, L. H. (1989). skin changes in photoaging. Aging and the skin. A. M. Kligman and A. K. Balin. New York, Raven press: 331–346.

Kollias, N. (1995). "The physical basis of skin color and its evaluation." Clinics in Dermatology **13**(4): 361-367.

Kollias, N. (1997). "‘Polarized light photography of human skin." Bioengineering of the Skin: Skin Surface Imaging and Analysis.

Kollias, N. and G. N. Stamatas (2002). "Optical Non-Invasive Approaches to Diagnosis of Skin Diseases." Journal of Investigative Dermatology **7**: 64-75.

Konior, R. J. and J. D. Kerth (1990). "Selected approaches to the aging face." Otolaryngol Clin North Am **23**(6): 1083-95.

Kwon, Y. H. and N. da Vitoria Lobo (1994). "Age classification from facial images." Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on: 762-767.

Land, D. G. and R. Shepherd (1988). "Scaling and ranking methods." Sensory Analysis of Foods: 155–185.

Lanitis, A. The FG-NET Aging Database: http://sting.cycollege.ac.cy/~alanitis/fgnetaging/index.htm.

Lanitis, A. (2002). "On the significance of different facial parts for automatic age estimation." Digital Signal Processing, 2002. DSP 2002. 2002 14th International Conference on **2**.

Lawless, H. T. and G. J. Malone (1986). "A comparison of rating scales: Sensitivity, replicates and relative measurement." J. Sensory Studies **1**(2): 229-252.

Lawless, H. T. and G. J. Malone (1986). "The discriminative efficiency of common scaling methods." J. Sensory Stud. **1**: 85-98.

Lazraq, A., R. Cleroux, et al. (2003). "Selecting both latent and explanatory variables in the PLS1 regression model." Chemometrics and Intelligent Laboratory Systems **66**(2): 117-126.

Lemperle, G., R. E. Holmes, et al. (2001). "A classification of facial wrinkles." Plast Reconstr Surg **108**(6): 1735-1750.

Leung, W. C. and I. Harvey (2002). "Is skin ageing in the elderly caused by sun exposure or smoking?" British Journal of Dermatology **147**(6): 1187-1191.

Lobo, N. and Y. Kwon (1998). Automatic feature detection and age classification of human faces in digital images, Google Patents.

Lopez Hernandez, B., J. Tercedor, et al. (1995). "[Skin aging and smoking]." Rev Clin Esp **195**(3): 147-9.

Macchi Cassia, V., D. Kuefner, et al. (2006). "A behavioural and ERP investigation of 3-month-olds' face preferences." Neuropsychologia **44**(11): 2113-2125.

Mark, L. S. and J. T. Todd (1985). "Describing perceptual information about human growth in terms of geometric invariants." Percept Psychophys **37**(3): 249-56.

Meissner, C. A. and J. C. Brigham (2001). "Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review." Psychology, Public Policy, and Law **7**(1): 3-35.

Montepare, J. M. and L. Z. McArthur (1986). "The influence of facial characteristics on children's age perceptions." Journal of Experimental Child Psychology **42**(3): p303-14.

Nagata, N. and S. Inokuchi "Subjective Age Obtained from Facial Images-How Old We Feel Compared to Others." V. Palade, RJH, Jain, L., eds.: Knowledge-Based Intelligent Information and Engineering Systems Part2, Lecture Notes in Artificial Intelligence **2774**: 877–991.

Nishimori, Y., A. D. Pearse, et al. (1998). "Elastotic degenerative change and yellowish discolouration in photoaged skin." Skin Research and Technology **4**(2): 79-82.

Nouveau-Richard, S., Z. Yang, et al. (2005). "Skin ageing: A comparison between Chinese and European populations A pilot study." Journal of Dermatological Science **40**(3): 187-193.

O'Hare, P. M., A. B. Fleischer, et al. (1999). "Tobacco smoking contributes little to facial wrinkling." Journal of the European Academy of Dermatology and Venereology **12**(2): 133-139.

Okazaki, M., K. Yoshimura, et al. (2005). "Correlation between age and the secretions of melanocyte-stimulating cytokines in cultured keratinocytes and fibroblasts." BRITISH JOURNAL OF DERMATOLOGY-SUPPLEMENT- **153**: 23.

Ortonne, J. P. (1990). "Pigmentary changes of the ageing skin." Br J Dermatol **122**(suppl 35): 21-8.

Pearson, K. (1901). "On lines and planes of closest fit to systems of points in space." Philosophical Magazine **2**(6): 559-572.

Permissions, O. and D. Surgery (2006). "Reproducibility of Two Four-Point Clinical Severity Scores for Lateral Canthal Lines (Crow's Feet)." Dermatologic Surgery **32**: 1256.

Quatresooz, P., C. Pierard-Franchimont, et al. (2006). "Skin climacteric aging and hormone replacement therapy." Journal of Cosmetic Dermatology **5**(1): 3-8.

Raine-Fenning, N. J., M. P. Brincat, et al. (2003). "Skin aging and menopause: implications for treatment." Am J Clin Dermatol **4**(6): 371-8.

Raitio, A., J. Kontinen, et al. (2004). "Comparison of clinical and computerized image analyses in the assessment of skin ageing in smokers and non-smokers." Acta Dermato-Venereologica **84**(6): 422-427.

Ramanathan, N. and R. Chellappa (2006). "Modeling Age Progression in Young Faces." Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on **1**.

Rehnman, J. (2007). The role of gender in face recognition, Department of Psychology, Stockholm University.

Rexbye, H., I. Petersen, et al. (2006). "Influence of environmental factors on facial ageing." Age Ageing **35**(2): 110-5.

Rexbye, H. and J. Povlsen (2007). "Visual Signs of Ageing: What are We Looking at?" International Journal of Ageing and Later Life **2**(1): 61-83.

Sclafani, A. P. and T. Romo 3rd (2000). "Injectable fillers for facial soft tissue enhancement." Facial Plast Surg **16**(1): 29-34.

SeonYoung, J., M. O'Mahony, et al. (2004). "A comparison of category and line scales under various experimental protocols." Journal of Sensory Studies **19**(1): 49-66.

Seoul, K. (2004). Extraction and Manipulation of Wrinkles and Spots for Facial Image Synthesis. Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition.

Serup, J. (1995). "EEMCO guidance for the assessment of dry skin (xerosis) and ichthyosis: clinical scoring systems." Skin Res Technol **1**: 109-114.

Serup, J. and T. Agner (1990). "Colorimetric quantification of erythema-a comparison of two colorimeters (Lange Micro Color and Minolta Chroma Meter CR-200) with a clinical scoring scheme and laser-Doppler flowmetry." Clin Exp Dermatol **15**(4): 267-272.

Simpson, E. and M. Henneberg (2002). "Variation in soft-tissue thicknesses on the human face and their relation to craniometric dimensions." American Journal of Physical Anthropology **118**(2): 121-133.

Sörqvist, P. and M. Eriksson (2007). "Effects of training on age estimation." Applied cognitive psychology **21**(1): 131-135.

Stamatas, G. N. and N. Kollias (2004). "Blood stasis contributions to the perception of skin pigmentation." Journal of Biomedical Optics **9**: 315.

Sumino, H., S. Ichikawa, et al. (2004). "Effects of Aging, Menopause, and Hormone Replacement Therapy on Forearm Skin Elasticity in Women." Journal of the American Geriatrics Society **52**(6): 945-949.

Taister, M. A., S. D. Holliday, et al. (2000). "Comments on Facial Aging in Law Enforcement Investigation." Masthead **2**(2).

Tenenhaus, M. (1998). La régression PLS: théorie et pratique, Ed. Technip.

Trueb, R. M. (2003). "Association between smoking and hair loss: Another opportunity for health education against smoking?" Dermatology **206**(3): 189-191.

Vacher, C. (2004). "Anatomie du vieillissement craniofacial." EMC-Dentisterie **1**(3): 201-213.

Vukadinovic, D. and M. Pantic (2005). "Fully Automatic Facial Feature Point Detection Using Gabor Feature Based Boosted Classifiers." Systems, Man and Cybernetics, 2005 IEEE International Conference on **2**.

Waters, K. (1987). "A muscle model for animation three-dimensional facial expression." Proceedings of the 14th annual conference on Computer graphics and interactive techniques: 17-24.

Waters, K. and D. Terzopoulos (1990). "A physical model of facial tissue and muscle articulation." Visualization in Biomedical Computing, 1990., Proceedings of the First Conference on: 77-82.

Weatherall, I. L. and B. D. Coombs (1992). "Skin Color Measurements in Terms of CIELAB Color Space Values." Journal of Investigative Dermatology **99**(4): 468-473.

Wines, N. and E. Willsteed (2001). "Menopause and the skin." Australasian Journal of Dermatology **42**(3): 149-160.

Wold, S., A. Ruhe, et al. (1984). "The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses." SIAM Journal on Scientific and Statistical Computing **5**: 735.

Wold, S., M. Sjostrom, et al. (2001). "PLS-regression: a basic tool of chemometrics." Chemometrics and Intelligent Laboratory Systems **58**(2): 109-130.

Wright, D. B. and B. Sladden (2003). "An own gender bias and the importance of hair in face recognition." Acta Psychologica **114**(1): 101-114.

Yousif, N. J., A. Gosain, et al. (1994). "The nasolabial fold: an anatomic and histologic reappraisal." Plastic and reconstructive surgery(1963) **93**(1): 60-69.