2012-ENST-015

EDITE - ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Traitement du signal et des images »

présentée et soutenue publiquement par

Carmelo VELARDO

le 23 Avril 2012

Estimation visuelle d'indices anthropométriques

Directeur de thèse : **Jean-Luc DUGELAY**

Jury

M. François BREMOND, Directeur de recherche, INRIA	Rapporteur
M. Frédéric DUFAUX, Directeur de recherche, CNRS	Examinateur
M. Jean-Marc ODOBEZ, Directeur de recherche, IDIAP	Examinateur
M. Fabio ROLI, Professeur, University of Cagliari	Rapporteur

T H È S E

TELECOM ParisTech

école de l'Institut Télécom - membre de ParisTech

Anthropometry and Soft Biometrics for Smart Monitoring

Carmelo VELARDO

A doctoral dissertation submitted to:

TELECOM ParisTech

in partial fulfillment of the requirements for the degree of:

DOCTOR OF PHILOSOPHY

Major subject: Image Processing, Biometrics

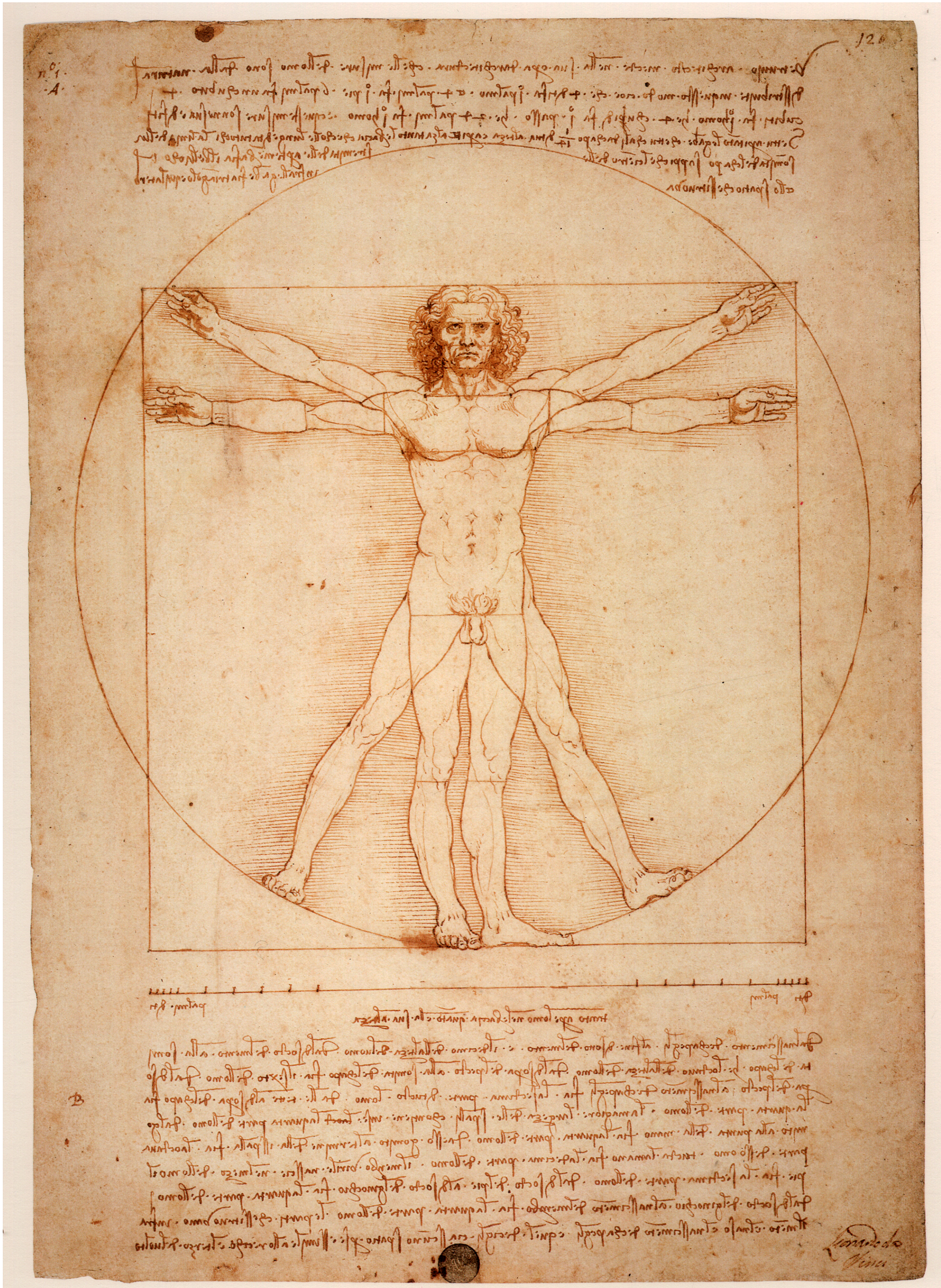
Approved by the following examining committee:

Dr. Jean-Marc Odobez	*President*
Prof. Fabio Roli	*Reviewer*
Dr. François Brémond	*Reviewer*
Dr. Frederic Dufaux	*Examiner*
Prof. Dugelay Jean-Luc	*Advisor*

A journey of a thousand miles
begins with a single step
— Sayings of 老子 (Lao Tzu)

Pluralitas non est ponenda sine necessitate
— William of Ockham

If we knew what it was we were doing,
it would not be called research, would it?
— Albert Einstein



The Vitruvian Man is a world-renowned drawing created by Leonardo da Vinci circa 1487. It is accompanied by notes based on the work of the famed architect Vitruvius (from which the name). It depicts a male figure simultaneously inscribed in a circle and square and it is based on the correlations of ideal human proportions with geometry described by the ancient Roman architect.

Abstract

The increasingly rapid convergence between machine and human languages has created a growing interest in the semantic analysis of multimedia contents. More recently, the creation of natural user interfaces has significantly reduced the gap between man and machine. Despite the recent advances in the technology, much remains to study in order to fill this semantic gap. That is to say the lack of coincidence between the information the machine extracts from the data and the way humans interpret that content.

Biometric research has advanced in these regards considering the recent focus on "Soft Biometrics". Those traits have attracted the attention of the research community in that they have some characteristics peculiar to "Hard Biometrics" while they improve the semantic content they carry.

In this dissertation we explore the Body Soft Biometrics concept both from a theoretical and a practical point of view. Being this field of study closely related with medicine, throughout the manuscript we refer to many studies conducted in the medical domain.

Taking advantage of datasets used for demographics, we analyze the relation between body parts and personal traits like anthropometric measures, weight, and gender. The ability to extract this information is tested by using both the 2D image analysis and by processing 3D video streams.

Subsequently, we use Soft Biometrics for pruning a biometric database, enabling faster and more accurate response of a face recognition system. Moreover, we build an application that demonstrates the feasibility of people re-identification for limited groups of users.

Furthermore, we present two medical applications for health conditions tracking. The former enables the interaction between the user and an automatic system that performs a medical check up of user's body, providing hints and lifestyle suggestions. The latter is intended to support the monitoring of cosmonauts' weight losses due to the lack of gravity in outer space.

Résumé

La convergence de plus en plus rapide entre la communication des machines et celle de l'homme a créé un intérêt croissant pour l'analyse sémantique des contenus multimédia. Plus récemment, la création d'interfaces naturelles qui permettent une communication plus aidée aux utilisateurs a considérablement réduit la distance entre l'homme et l'ordinateur.

Malgré les progrès technologiques récents, il reste encore beaucoup à faire pour combler le soi-disant "semantic gap" qui permettrait à la machine d'interpréter le contenu d'une manière plus humaine. Aussi en ce qui concerne les systèmes biométriques, des progrés ont été réalisés. Récemment, la "biométrie douce" a commencé à attirer l'attention de la communauté de recherche. La biometrie douce est un nouveau domaine de recherche qui vise à explorer les traits de l'homme, toute en permettant l'intercommunication entre l'homme et la machine, pour cette raison on parle également de "traits sémantiques".

Dans cette thèse, nous explorons le concept de la biométrie douce par rapport à la structure du corps humain, tant dans un cadre théorique, que dans un contexte pratique. Comme ce champ d'étude est fortement liés à la médecine, tout au long de la thèse nous ferons de nombreuses références à des études dans le domaine médical. En ce qui concerne les études et les bases des données utilisés normalement pour des statistiques sur la population, nous effectuons une analyse des différentes parties du corps humain afin d'être en mesure d'extraire de l'information biométrique comme le poids, et le sexe de la personne. La capacité d'extraire cette information a été testé en utilisant à la fois l'analyse d'images 2D ou de flux vidéo 3D.

En outre, nous avons utilisé l'ensemble des données anthropométriques pour effectuer une opération d'élagage permettant une exécution plus rapide des algorithmes de reconnaissance biométrique. Une autre application montre que dans le cas où l'on considère un groupe limité d'utilisateurs, la biométrie douce peut également être utilisée pour réidentifier les utilisateurs d'un système de vidéo surveillance. Enfin, nous montrons deux applications utilisables dans le domaine médical, le premier montre la possibilité d'intereagire avec l'utilisateur en analysant la santé de son corps. Le système peut ensuite distribuer automatiquement des conseils fondés sur l'utilisation de certaines formules qui font partie de l'état de l'art médical. La dernière application est basée sur un système qui cherche à contrôler le problème de perte de poids dans des conditions d'absence de gravité. Le système conçu surveille le poids des cosmonautes grâce à l'utilisation exclusive de la vision artificielle.

Riassunto

La sempre più veloce convergenza tra i linguaggi delle macchine e quello umano ha dato vita a un crescente interesse verso l'analisi semantica di contenuti multimediali. Più recentemente la creazione di interfacce naturali, che permettano la comunicazione diretta delle intenzioni dell'utente, ha ridotto sensibilmente la distanza tra uomo e computer. Nonostante questi ultimi avanzamenti in campo tecnologico, molto resta da fare per colmare il cosiddetto "gap semantico" che non permette alla macchina di interpretare in maniera direttamente comprensibile all'uomo i contenuti da lei analizzati.

Anche per quanto riguarda i sistemi biometrici sono stati fatti numerosi passi avanti. Recentemente i "Soft Biometrics" hanno cominciato ad attrarre l'attenzione della comunità di ricerca. Essi sono un nuovo campo di ricerca che intende esplorare dei tratti tipici dell'uomo, ma che permettano al tempo stesso la intercomunicabilità tra uomo e macchina, per questo motivo sono stati anche definiti *tratti semantici*.

In questa tesi esploriamo i Soft Biometrics relativi alla struttura del corpo umano sia da un punto di vista teorica, sia in ambito pratico. Essendo questa materia di studio strettamente legata alla medicina, si troveranno ampi riferimenti a studi condotti in questo ambito di ricerca. Approfittando di set di dati utilizzati per la demografia, analizziamo il rapporto tra le parti del corpo e tratti personali quali misure antropometriche, peso e sesso. La capacità di estrarre queste informazioni è testata utilizzando sia l'analisi di immagini 2D, sia sfruttando l'elaborazione di flussi video 3D.

Successivamente abbiamo usato l'insieme di dati antropometrici per effettuare una

operazione di *pruning* per permettere una esecuzione più veloce degli algoritmi di riconoscimento biometrico. Una ulteriore applicazione dimostra come in caso di insiemi limitati di persone i tratti soft biometrics possano essere utilizzati anche per re-identificare gli utenti di un sistema di videosorveglianza. Per ultimo mostriamo due applicazioni in ambito medico, la prima consente l'interazione tra l'utente e un sistema automatico che esegue un check up medico del corpo umano, fornendo suggerimenti per un corretto stile di vita e identificando il peso ideale della persona. La seconda applicazione è destinata ad aiutare i cosmonauti nel monitoraggio delle perdite di peso inevitabili in situazioni di microgravità.

Acknowledgments - Remerciements - Ringraziamenti

Since I was a child the passion of research have driven me into the unexplored. Everything evolves with time and experience, and much has changed since then, the path towards the PhD contributed to those changes, in good.
I think that challenging ourself to achieve always more is one of the keys that pushes us to go far beyond what we thought our limits lied. Sure this power is an internal source of energy, but it is the results of external affluents that come into our stream to make it bigger. Counting the number of people I encountered while running towards the end of this PhD is a difficult task. After the year of promo2007 at Eurecom almost 6 years passed.
Research would have been not so passionating without my advisor, prof. Jean-Luc Dugelay, he made me realize what really means doing research and for this I am grateful. Special thanks to the thesis committee: Dr. François Brémond, Dr. Frederic Dufaux, Dr. Jean-Marc Odobez, and prof. Fabio Roli; to have the patience to read and to comment this dissertation. Obviously, I cannot forget all Eurecom staff that was very kind in all these years, thank you: Gweanelle, Patrick, Carine, Sebastien, and all the others that made my stay at Eurecom very pleasant. Another thought to my colleagues that followed me through this path, Cucillo, Teo, Simon, Claudia and Miriam, Antonio and Mario, Alessandro, Corrado and Xiaolan, Francesco and Lei and Lorenzo, Nesli and Antitza; you are definitely too many to come into mind all at once.
A special thanks to those people that participated both to my professional and personal life: Angela and Judith, Marco and Dolores, Giuseppe and Giovanna; you should never forget you are all special people.
Other people left an indelible impression in my life experience: Rachele our roads ran in parallel and hopefully they will continue so, F.F. I mention you by your initials because the initials are the key, Raffaella you are special, continue being so; Antonio, Marco-SparaJury, Callippo, Teccio and all the friends from the *secondopiano* group, because of you a lot of things have never been done, but a lot of better ideas came to life.
And last, but first in this downward but ascending list of importance is my family that supported me again and again and again with no doubts, with no reconsiderations, just like every family should do, to you my biggest and thankful **Grazie!**

Abbreviations

AAM Active Appearance Model
ANN Artificial Neural Network
ANOVA Analysis of Variance
BMI Body Mass Index
BMMD Body Mass Measurement Device
BMR Basal Metabolic Rate
CAESAR Civilian American and European Surface Anthropometry Resource
CBR Content Based Retrieval
CDC Center for Disease Control and Prevention
CMC Cumulative Match Characteristic curve
DARPA Defense Advanced Research Projects Agency
DLT Direct Linear Transformation
FBI Federal Bureau of Investigation
FERET Face Recognition Technology
FOV Field of View
IBW Ideal Body Weight
ISS International Space Station
KNN K-Nearest neighbor
Lab Colorspace: Luminance and *a,b* color opponent dimensions
LDA Linear Discriminand Analysis
MLP Multilayer Perceptron
NASA National Aeronautics and Space Administration
NHANES National Health and Nutrition Examination Survey
NUI Natural User Interface
PCA Principal Component Analysis
PCH Probabilistic Color Histogram
RGBD Red, Green, Blue, and Depth camera
SLAMMD Space Linear Acceleration Mass Measurement Device
SVM Support Vector Machine
WHO World Health Organization

List of Figures

List of Tables

Contents

CHAPTER 1

Introduction

The identification process has always been of foremost importance for human beings. Identifying other people allows interactions and other social activities that are vital to us. However, those innate capabilities that characterize humans, are not available to computers and automatic systems. For that, computer scientists have spent last decades working on the Biometrics domain to embed such capabilities into machines.

For many years biometric information have been exploited to extract information under the form of features so as to provide identity insights to machines. Notwithstanding their identification purpose has been abundantly explored, a lot remains to investigate about identification when contact and subject cooperation cannot be guaranteed.

Moreover, while humans can easily understand and describe semantic concepts and use them to indicate actions or subjects, computers still encounter difficulties to abstract a plausible representation of a person from his/her physical attributes description.

For this reason, recently, research community felt the necessity to explore techniques that would allow machines to exploit this semantic representation in order to perform some actions, mainly with regards to the identification problem. Filling the semantic gap will enable a complete and satisfactory human/machine interaction[1].

[1] The semantic gap is the lack of coincidence between the information that one can extract from

In the last years Soft Biometrics, a new branch of Biometric studies, have started exploring the semantic representation of people and received particular attention. Soft biometrics' principal aim is to extract quasi-unique (or common) information about someone from images or videos. Those traits are especially oriented towards the description of semantic attributes and thus often referred to as *semantic traits* [5].

While in the biometric approach a template (e.g. face's image) is extracted from the biometric specimen (e.g. the face itself), and a digital version of it is compared against a gallery of probes, with soft biometrics we deal with semantic traits (e.g. one's height, or the color of his/her shirt) whose general validity does not enable robust identification approach per se. The first approach is generally valid for humans, as for example we are able to identify people from their pictures. Nevertheless, humans are able to exploit semantic level descriptions to perform similar actions (e.g. identifying someone in the crowd by its physical description). In the future employing soft biometrics will enable the use of characteristics to describe people to machine and vice-versa.

Thanks to their semantic content, being able to extract those traits will provide to the machine a deeper knowledge about the user with which it is interacting. All this could be exploited to enhance current human-machine interaction techniques facilitating many tasks and enabling a better control. An example could be represented by vending machine that could avoid providing alcohol or cigarettes to children; or that could find for a tall robust person wearing a black cap that stole my red purse minutes ago.

1.1 Motivations, objectives and contributions

The motivations that have driven this thesis work came from the lack of adequate exploration of *soft biometrics* and *anthropometry* within the Biometric community. For what regards technologies, algorithms, and ideas, many concepts remains unexplored and only recently the community started to scratch the surface of this interesting topic. To further motivate our work we explore different applications where soft biometric extraction can be employed to open wider the perspectives of this research topic.

Our contributions lay on the development and formulation of a clear definition

the sensory data and the interpretation that the same data has for a user in a given situation. [4].

of what are soft biometrics, and how to classify them in different areas (namely: body, face, and accessories classes). Moreover, to clearly distinguish the benefit that derive from the use of such features we leverage on the characteristics proper of the soft biometrics: human compliant, enrollment free, non-intrusive. Especially the first one is of fundamental importance as it underlines the inner capability that makes soft biometrics understandable both by the human and by the machine.

Thus, the objectives of our work are to provide a reliable way to extract body soft biometric traits, and to demonstrate their applicability in case of user re-identification, biometrics database pruning, and human machine interaction.

Our contribution focused especially on the study of *body soft biometric* traits. We explored different ways of extracting weight, height, and gender information employing different techniques and sensors, and we applied the result of the detection step in many different applications. As we will further read, while in the case of height estimation other algorithms existed to obtain this trait, in case of the weight our work provides one of the first applications of body anthropometry for weight estimation using video sensors and we have found several applicative domains that could benefit from such capabilities.

1.2 Thesis organization

The manuscript can be divided in two main parts. The first one present a state of the art about soft biometrics and some soft biometrics extraction techniques (Chapter 2 and Chapter 3). The second part is instead devoted to the presentation of some applications that exploit body soft biometrics (Chapter 4, Chapter 5, and Chapter 6).

In Chapter 2 we retrace the history of Soft Biometrics and we review their late introduction into Biometrics studies. We clarify the up to date definition that better describes those traits and we review the most important traits with a special attention devoted to Body Soft Biometrics. Moreover, we define the domain of application of Soft Biometrics and we introduce the main three areas that will be explored in this dissertation.

Chapter 3 explores different ways of extracting body soft biometrics. We provide some state of the art techniques description to extract height, and we explore deeper the part regarding the extraction of anthropometric measures thanks to new affordable 3D sensors. Thanks to the National Health and Nutrition Examination Survey (NHANES) dataset collected by the American Center for Disease Control and prevention (CDC) a theoretical study is described on a technique for weight estimation from anthropometric data.

In Chapter 4 we investigate body soft biometrics capabilities to perform pruning of a hard biometrics database. We demonstrate how a signature composed by body soft biometrics can improve both retrieval speed and accuracy. And we show in a practical experiment how we can improve at least of two times the performance of the recognition algorithm. Our pre-classification step based on anthropometric measures is tested in conjunction with a face recognition algorithm whose performance obtains benefit from the pruning phase. Finally, we clearly identify the trade off among pruning, accuracy, and mensuration error of an anthropometric based system.

In Chapter 5 we explore the possibility of performing re-identification in a video surveillance environment thanks to a body soft biometric signature composed of the height, a measure correlated with the weight of the person, and the clothes color of subjects. We show the different techniques used to extract such parameters and then we provide results on the people re-identification part. In this Chapter we see that exposing publicly Soft Biometrics characteristics still preserve the privacy of the users under surveillance.

In Chapter 7 we draw some conclusions and we explore plausible new perspectives and future works.

This dissertation has been partially based on the following materials.

- Velardo, Carmelo; Dugelay, Jean-Luc, *What can computer vision Tell you about your weight?*, Under submission/preparation

- Velardo, Carmelo; Dugelay, Jean-Luc; Paleari, Marco; Paolo, Ariano, *Building the space scale or how to weigh a person with no gravity*, 1st IEEE International Conference on Emerging Signal Processing Applications, January 12-14, 2012, Las Vegas, Nevada, USA

- Velardo, Carmelo; Dugelay, Jean-Luc, *Improving identification by pruning: a case study on face recognition and body soft biometric*, Research report RR-12-262, pp 1-20

- Velardo, Carmelo; Dugelay, Jean-Luc; Daniel, Lionel; Dantcheva, Antitza; Erdogmus, Nesli; Kose, Neslihan; Min, Rui; Zhao, Xuran, *Introduction to biome-*

try, Book chapter of “Multimedia Image and Video Processing” (2nd edition); CRC Press; 2011

- Velardo, Carmelo; Dugelay, Jean-Luc, *Real time extraction of body soft biometric from 3D videos*, ACM Multimedia 2011, 28 November-1 December, 2011, Scottsdale, Arizona, USA

- Velardo, Carmelo; Araimo, Claudia; Dugelay, Jean-Luc, *Synthetic and privacy-preserving visualization of camera network outputs*, 5th ACM/IEEE International Conference on Distributed Smart Cameras, August 22-25, 2011, Ghent, Belgium

- Dantcheva, Antitza; Velardo, Carmelo; D'angelo, Angela; Dugelay, Jean-Luc, *Bag of soft biometrics for person identification : New trends and challenges* Mutimedia Tools and Applications, Springer, October 2010, pp 1-39

- Velardo, Carmelo; Dugelay, Jean-Luc, *Weight estimation from visual body appearance*, IEEE 4th International Conference on Biometrics: Theory, Applications and Systems, September 27-29, 2010, Washington DC, USA, pp 1-6

- Velardo, Carmelo; Dugelay, Jean-Luc, *Face recognition with DAISY descriptors*, ACM SIGMM Multimedia and Security Workshop, September 9-10, Rome, Italy, pp 95-100

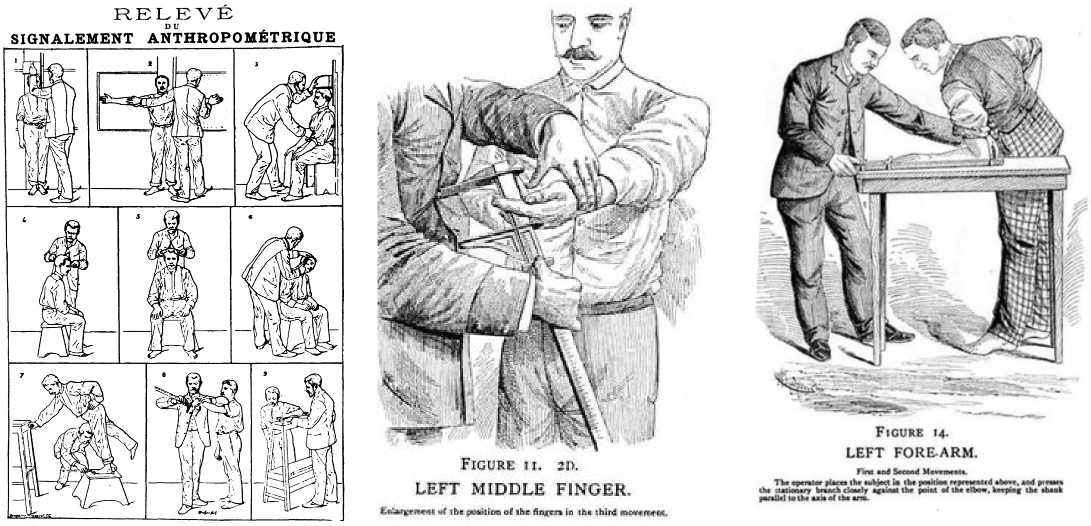
- Paleari, Marco;Velardo, Carmelo;Huet, Benoit;Dugelay, Jean-Luc, *Face dynamics for biometric people recognition*, IEEE International Workshop on Multimedia Signal Processing, October 5-7, 2009, Rio de Janeiro, Brazil, pp 1-5

CHAPTER 2

Soft Biometric: a state of the art

2.1 Introduction

The term “Biometry” comes from the fusion of two ancient Greek words: $\beta\iota o\varsigma$ (bios: life) and $\mu\varepsilon\tau\rho o\nu$ (metron: to measure, to count). Those two words indicate that there is something related to life (the human nature) that can be measured, or counted. Biometry is indeed the science that tries to understand how to measure characteristics which can be used to distinguish individuals. Humans have developed such skills during the evolution: the brain has specialized areas to recognize faces [6] and to link identities with specific patterns (behavioral or physical [7]). Researchers in the biometry field have always tried to automatize such processes making them suitable to be run on a computer or a device. The study of biometric patterns led to the definition of requirements which have to be respected to make a human trait feasible to be used in a recognition process. A biometric trait can be summarized as: a characteristic that each person should have (Universality), in which any two persons should present some differences (Distinctiveness), that should not drastically vary over a predefined period (Permanence), and that should be quantitatively measurable (Collectability). Furthermore, the biometric traits can be divided into the two following classes: physical and behavioral. To the former class belongs the face appearance, the pattern of the iris and the fingerprint, the structure of the retinal blood vessels as well as the shape of the ear. Each of these traits can be additionally subdivided into genotypic and randotypic, the former indi-



(a) (b) (c)

Figure 2.1: The images show the methodology applied to gather information from the suspect in the *bertillonage* system. The procedure was standardized by Bertillon in his book [8].

cates a correlation with some genetic factors (like hereditary similarities in twins), the latter describes traits that develop randomly during the fetal phase. Behavioral biometrics develop as we grow older and are not a priori defined. To those traits belong the gait, or even the keystroke pattern (the way of typing on a keyboard).

Lately a new biometric concept, Soft Biometric (also called *semantic* [5]), gained influence as it can increase the reliability of a biometric system and can provide substantial advantages: soft biometric features reveal biometric information, they can be partly derived from hard biometrics, they do not require the enrollment and can be acquired non intrusively without the consent and cooperation of an individual.

Those traits share a lot of similarities with the work of Alphonse Bertillon, a French police officer that first explored biometrics and set up the very first biometric identification system based on anthropometry classification. Before that criminals could be identified based on eyewitness accounts or by branding (which however was discontinued in France by 1809 leaving a gap in the methods for criminals identification). The modern method devised by Bertillon (for this reason dubbed *Bertillonage*) consisted in a series of anthropometric measures whose acquisition method was standardized in Bertillon manual [8] (see figure B.1 for some illustrations).

Data were accurately collected together on cards depicting identification pictures of the suspect, and then stored on special drawers. The cards were indexed according to a given combination of the anthropometric measures so as to guarantee the possibility of a fast access to subjects' file. An example of Bertillon file (the mugshots ancestor) is shown in figure B.2.

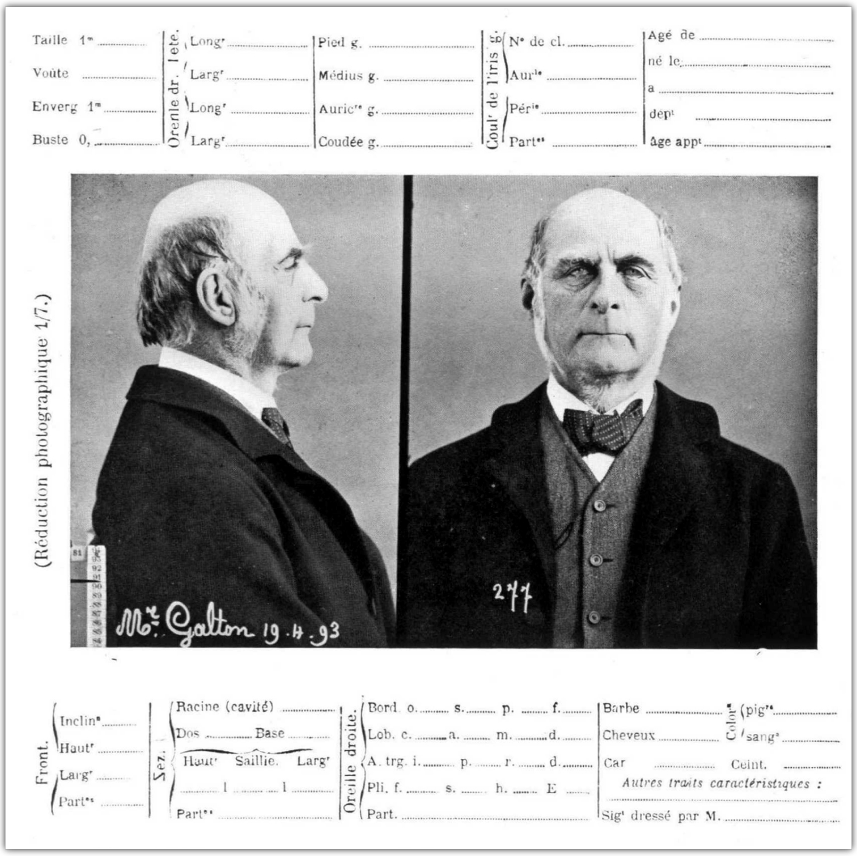


Figure 2.2: An example of Bertillon's file. It shows the picture of Mr. Francis Galton during a visit he conducted to Bertillon's laboratory in 1893. Galton is mostly known for his studies on the uniqueness of fingerprint pattern.

Notwithstanding its wide use across France, Great Britain and United States of America [9] the Bertillonage was found flawed as in the case of Will West vs. William West [10] that instead proved the higher reliability of fingerprint based systems. From that point on anthropometry use for identification has been discontinued and it was substituted by fingerprint matching and others more distinctive biometric techniques. However, lately big attention has been drawn on features called Soft Biometrics that recall for some aspects the Bertillon system.

2.2 New definition of Soft Biometrics

Despite the flourishing studies that explore the soft biometric topic, literature presents a lack of a formal definition that describes what can be considered belonging to this category. For this reason we derived a new definition of soft biometrics that generalizes this concept [11].

> “ *Soft biometrics are physical, behavioral or adhered human characteristics classifiable in predefined human compliant categories. These categories are, unlike in the classical biometric case, established and time-proven by humans with the aim of differentiating individuals. In other words the soft biometric trait instances are created in a natural way, used by humans to distinguish their peers.* ”

Traits accepting this definition include but are not limited to: gender, weight, height, age, eye color, ethnicity and so on. An increase in resources (such as an improved resolution of the sensors, or an increased computational capability) can lead to expanding of the traits amount and furthermore of the trait instances. The nature of soft biometrics features can be binary (for example presence of glasses), continuous (height) or discrete (ethnicity) [11] (an example of classification of soft biometric traits can be found in Table 2.1). In case of discrete values we can further divide each trait in trait instances. We refer to trait instances as the sub categories soft biometric traits can be classified into. Example for trait instances of the trait hair color could be: blond, red and black. It is important to notice that continuous value can be eventually discretized.

Like Biometrics, those characteristics can be differentiated according to their distinctiveness and permanence, whereby distinctiveness measures the power of a trait to distinguish subjects within a group, and permanence relates to the time invariability of a trait. Both of those characteristic are mostly in a lower range for soft biometrics than they are for classical biometrics (e.g. hair color, or presence of beard, can change during time and are shared among people). For this reason a distinction is traced between the *Hard* and the *Soft Biometrics*, in which the first are traits that can singularly provide a higher (harder) degree of distinctiveness, thus are more suitable for identification.

Furthermore it is of interest with which estimation reliability a trait can be extracted from an image or a video. By exploiting these three qualities, namely distinctiveness, permanence and estimation reliability, the importance of a soft biometric trait can be determined. We note that the classification of soft biometric

Table 2.1: Table of soft biometric traits

Soft Biometric trait	**Face, Body, Accessory**	**Nature of value**	**Permanence**	**Distinctiveness**	**Subjective perception**
Anthropometric measures	Body	Continuous	Medium/High	Medium/High	Medium
Height	Body	Continuous	Medium/High	Medium	Medium
Weight	Body	Continuous	Low/Medium	Medium	Medium
Body shapes	Body	Discrete	Medium	Medium	Medium
Clothes color	Accessory	Discrete	Low	Medium	Medium
Gait	Body	Continuous	Medium	Medium	High
Gender	Face/Body	Binary	High	Low	Low
Marks	Face/Body	Discrete	High	Medium/High	Low
Age	Face/Body	Continuous	Low/Medium	Medium	Medium
Skin color	Face	Continuous	Medium	Low	Medium
Hair color	Face	Continuous	Medium	Medium	Medium
Eye color	Face	Continuous	High	Medium	Medium
Beard	Face	Binary	Low/Medium	Low	Medium
Mustache	Face	Binary	Low/Medium	Low	Medium
Facial measures	Face	Continuous	High	Medium	Medium/High
Facial shapes	Face	Discrete	High	High	High
Facial feature measures	Face	Continuous	High	High	Medium/High
Facial feature shape	Face	Discrete	High	High	High
Make–up	Face	Discrete	Low	Low	Medium
Ethnicity	Face	Discrete	High	Medium	Medium
Glasses	Accessory	Binary	Low/Medium	Low	Low

traits can be expanded to evaluate or deduce aspects like accuracy and importance, depending on the cause for application.

To summarize all the characteristics that individuate soft biometrics we provide a description to some of them in the following list.

- *Human compliant*: Traits conform with natural human description labels.
- *Computational efficient*: Sensory and computational requirements are negligible.
- *Enrollment free*: Training of the system is performed off-line and without prior knowledge of the inspected individuals.
- *Deducible from classical biometrics*: Traits can be partly derived from images captured for primary (classical) biometric identifier (e.g. eye color from iris images).
- *Non intrusive*: Data acquisition is user friendly or can be fully imperceptible.
- *Identifiable from a distance*: Data acquisition is achievable at long range.
- *Not requiring the individual's cooperation*: Consent and contribution from the subject are not needed.
- *Preserving human privacy*: The stored signatures are visually available to everyone and serve in this sense privacy.

Recently, soft biometric traits have been employed to preliminary narrow down the search in a database, in order to decrease the computational time for the classical biometric trait. Jain et al. who first introduced the term *soft biometrics* performed related studies on using soft biometrics [12] for pre-filtering and fusion in combination with classical biometric traits.

A further application approach is the fusion of soft biometric and classical biometric traits to increase the system performance and reliability. Recently soft biometric systems have been employed also for person recognition [11, 13] and continuous user authentication [14].

Further studies evolve traits extraction algorithms concerning eye color [15], height [16], clothes color [17] or predictability of human metrology [18]. Ethnicity and hair color were used in [19] to recognize and verify identities of clients in a biometric system. The authors have demonstrated that Soft Biometrics increase the recognition accuracy of the classic biometric system based only on face appearance. In [20] the authors demonstrated that it is possible to deal with the real time extraction of Soft Biometrics, and that the combination of such traits can lead to an acceptable level of recognition accuracy.

All these works demonstrate the two major contributions of Soft Biometrics to recognition systems: increasing the speed of existing techniques (helping to prune the less likely solutions), and improving the recognition accuracy (by preventing rough errors). However, because the permanence of such traits can be very low compared to biometric traits, soft biometrics features are known to be better exploited in session based systems (e.g. video surveillance systems), where the lack of distinctiveness becomes less significant since the domain of the system is a known subspace of the entire population.

2.3 Traits

2.3.1 Body soft biometrics

Height, gait, body weight and color of clothes concern the body and are the main traits that can be extracted from a distance. The best distinctiveness is provided by gait detection, which is why gait is occasionally referred to as a classical biometric.

2.3.1.1 Anthropometric measures

The studies on anthropometric measures are not generally driven by biometric use. While at the beginning anthropometry was a technique used in physical anthropology to study the physical development of the human species; nowadays it is employed in industrial/clothing design, ergonomics, and architecture to optimize the products to the customers needs. Other interesting works regards the study of population statistics so as to monitor changes in lifestyle, and nutrition to track body dimensions (e.g. obesity epidemic) [21].

The first biometric application of anthropometry is due to Alphonse Bertillon. His anthropometry-based classification method was used to identify criminals, and it is indeed one of the few examples of anthropometric measure used as biometric identifier.

After the historical contribution of Bertillon, one of the first works that tried to estimate anthropometric measures from images is the one presented in [22]. The authors of this paper, use a priori statistical information about the human body, to establish the correspondence between a set of manually marked points and the segments that compose the body parts. In a second step a set of postures is considered and finally pose and anthropometric measurements are obtained.

Results are achieved by minimizing an appropriate cost function and according to a model inspired by human body statistics collected for medical research.

The recent commercialization of *millimeter wave scanners* and full 3D body scanners has raised the interest of the research community. Some works have suggested that the idea of anthropometry based people identification is feasible and proposed some approaches. An example is the work of [23] where the authors investigate the utility of 1D anthropometric measurements as a biometric trait for human identification. They analyze 27 measurements from 2144 subjects, by reducing those measures to a smaller set thanks to dimensionality reduction techniques they obtain rank-1 identification of 83% and 94% using just ten and fifteen features.

Other interesting works on anthropometric measures are shown in [24] where height, stride, and other measures are taken into account for people identification; and in [25] where anthropometric measures are estimated from calibrated monocular sequences. By tracking subjects across multiple cameras the authors estimate stature, shoulder breadth, and link them with specific features provided by gait to perform people identification.

2.3.1.2 Height

Even if height is part of the more general anthropometric measures, we dedicate to it a separate section because the computer vision community explored deeply its extraction and possible applications.

Height estimation is an already mature topic in the literature and it has been exploited several times. One of the earliest approaches is presented in [16], the authors use the content of the image to compute geometrical properties of objects that lie on the same plane, later they can compare objects dimensions. By knowing the height of given objects in the scene they are able to measure height of people in the camera field of view (FOV). Extending this last work, the authors of [26] propose further improvements using multiple measures and a statistical approach to remove outliers, using the proposed approach they arrive to a precision of 1cm for subjects walking in an unconstrained scenario.

Precise measurement of height has been already used in combination with other features so as to track people across multiple camera systems, and to allow the identification of the same person in multiple video streams [27]. The estimation is performed via the computation of height related to the real world coordinates estimated in camera images.

Height is possibly one of the most used in real cases and can become under certain circumstances a crime evidence. It is indeed one of the main factors used in photogrammetry. This technique is nowadays widely used to estimate anthropometric measures from images or video surveillance footages. The Netherlands Forensic Institute has performed a comparison [28] of two methods for obtaining body height measurements from images. One is based on projective geometry and the other one on 3D modeling of the crime scene. Keeping the same camera settings setup, the authors demonstrate that the predictions of both methods are accurate, but changing camera position makes the first algorithm less reliable.

Moreover, also the 3D reconstruction of the environment can be helpful as this kind of analysis greatly simplify the extraction of measurements. The possibility of using such a technique is explored in [29] where the authors use landmarks within the scene to enable the automatic collection of subjects' height measurements.

2.3.1.3 Weight

Since the beginning weight was introduced within the list of the soft biometric traits [12]. However it was not fully explored as soft biometric traits.

A field where weight is considered an important feature is represented by medical studies. Here the main interest is represented by the visual extraction capability and reliability of human operators in case of emergency situations where there is no possibility of using scales like in [2, 30, 31].

Other interests are represented by the use of weight as a foremost important feature that helps to monitor body health status [32]. Additionally, a branch of medical studies explore the forensic aspect of weight estimation [33] so as to recover information from latent traces that help to recognize victims or crime suspects.

To the best of our knowledge, the only paper which involves weight directly referring to it as a soft biometrics is [34], where the authors use a scale to weigh clients of a fingerprint recognition system. By exploiting weight and body fat measurements the authors reduce the total error rate of the system by 2.4%.

Another work [5] considered weight as a discrete value visually defined by subjects participating to a psycho-visual experiment. However, the values used (Very Thin, Thin, Average, Fat, Very Fat) show that rather than the weight itself, the description refers to the way fat is distributed on the inspected body. That is to say users described the body build of subjects rather than their body mass. A similar experiment is reported on [35] where the authors propose other features alongside with weight. Additionally, the work points out the relevance of this trait for eyewitness testimony.

Some experiments are performed by [36, 37] that involve fitting a 3D human model to a point cloud obtained in the first case by a RGBD camera, and in the second case by a series of stereo cameras. In this case extracting the weight is straightforward if we consider the average density of the human body. While in the first work the weight (and gender) estimation is an interesting side effect, in the second case the authors purposely try to extract such information.

2.3.1.4 Gender

Gender recognition has been already widely explored in social and cognitive psychology work in the context of both face and body analysis. From an image processing point of view, the topic offers a plethora of approaches. The latest efforts employ a selection of fused biometric traits to deduce gender information. For example in [38] gait energy images and facial features are fused to perform gender recognition. Similarly, the authors of [39] segment the human silhouette into seven components (head, arm, trunk, thigh, front-leg, back-leg, and feet). They to study the effectiveness of the seven human gait components in more than 500 different experiments on people identification and gender recognition. On average, good performance are achieved by discriminative gait features on a dataset composed of 1870 sequences from 122 subjects.

Another gender recognition approach is followed by [40]; in this case still images are analyzed to provide basic information about gender and pose. The authors propose a joint framework for rough pose estimation and gender classification. Both the systems are based on biologically-inspired features in combination with manifold learning techniques that achieve the maximum performance of about 80% on a public available pedestrian database.

The difference between the performance of these two approaches lies on the higher informativeness carried by gait sequences that better describe the human shape.

2.3.1.5 Gait

Gait is a complex pattern that involves not only some anthropometric parameters but especially behavioral information. Among all the soft biometric traits, it is one of the most explored, and because of its distinctiveness it is debated as being actually a hard biometric. One of the first experiments (1973) on gait analysis is presented in [41], where the author uses lights attached to the joints of the human body to record subjects' gait patterns. The author demonstrates how observers can recognize walking people familiar to them just by the light traces they leave while walking. Since 1970's many other authors were interested in the topic of automatic gait recognition. In [42] a spatiotemporal signature is extracted by the moving silhouette, later on a principal component analysis is employed to discard irrelevant information. Finally, supervised pattern classification techniques are performed in the lower–dimensional eigenspace. In order to provide more discriminative power, both the structural and behavioral characteristics of gait are captured.

Another interesting work is proposed in [5], where gait is chosen as primary biometric trait to be coupled with "semantic biometrics", that seems to be a very similar concept to soft biometrics. Using ANOVA they first outline the most important semantic traits. After they merge the results of the signature generated by gait with the one generated by the semantic information so as to identify users of the biometric system. Other ways of performing human identification via gait analysis are based on the human silhouette and on model based systems like [43, 44].

Gait analysis is not only used to identify people or gender, it is actively used also in the medical domain. In this case it is exploited to understand the patterns for pathologically abnormal patients like in [45]. Although in the medical field the use of markers has been widely exploited, lately some studies have started to involve new techniques like computer vision, or new sensors like accelerometers [46, 47] to analyze this trait.

2.3.2 Face soft biometrics

Former work on soft biometrics has been performed predominantly with the aim of preprocessing. In face recognition for person identification, for instance, beard detection and removal serves as an improvement of recognition results, disregarding the information of the presence of beard.

2.3.2.1 Color based

The color based facial soft biometric traits (eye, skin, and hair color) are the most obvious facial identifiers, mentioned primarily by humans, when describing people's face, and carrying important clues about ethnicity [48]. The main challenges for skin classification are on the one hand the low spread of different skin colors in color space, and as a consequence, on the other hand the high illumination dependence of classification. The latter is described in various skin locus papers, for example in [49].

Hair color is detected by techniques similar to the ones of skin color and often they are considered together, but it has more broadly scattered color categories. The work of [50] summarizes a method for human head detection based on hair-color, the authors propose the use of Gaussian mixture density models to describe the distribution of hair color. In [51] the fuzzy theory is used to detect faces in color images, here two fuzzy models describe the skin color and hair color, respectively.

Unlike the other color based facial soft biometrics, eye color detection is a relatively new research topic. Probably because 90% of humans possess brown eyes this color based trait seems not very explored. The advantage of eye color classification is the availability of all necessary information in images used for iris pattern analysis, in other words, iris color could be considered as a free side effect. Work on fusion between iris texture and color can be found in [52], where the authors fuse iris and iris color with fingerprint to provide performance improvement in respect with the unimodal systems. Leveraging on this side effect of iris analysis, in [53] the authors successfully use the iris color as an index of a color based pruning method that improves system's speed.

2.3.2.2 Gender

Although gender is a characteristic that principally affects the body appearance, many works exist that try to extract this information from face.
For example, in [54] the authors propose a combined system for gender and expression recognition by modeling the face using an Active Appearance Model (AAM), feature extraction and finally linear, polynomial and radial basis function based support vector machines for classification.
The work in [55] proposes the use of Adaboost on several weak classifiers, applied on low resolution gray scale images with good results.
The authors of [56] present a novel multi-modal gender recognition system, based on facial appearance, head and mouth motion, exploiting a unified probabilistic framework.

An interesting study [48] links together ethnicy and gender. While people are generally prone to obtain better performance when trying to estimate the gender from people of their same ethnicity, computers seems not to be affected by this effect. The authors conclude that automatic systems based on image features (as well as pixel based techniques) are able to generalize the gender recognition problem disregarding the ethnicity information.

2.3.2.3 Beard and Mustache detection

Another typical face trait is provided by the presence of beard or mustache, especially for men's categorization. Generally considered as an obstacle to the face recognition algorithms because of its high time variability, the study of beard and mustache detection is principally focused towards a detect-and-remove approach. As major example, in [57] authors show an algorithm for beard removal from images of people with facial hairs, the system is designed using the concept of structural similarity and coordinate transformations.

2.3.2.4 Age

Age is mainly related to the structure of the face and it plays an important role for long time employable systems based on face or body and is a challenging and relatively new field.
An interesting study on face changes over time is the one in [58], which includes a biometric, forensic, and anthropological review, and further discusses techniques to simulate aging in face images.
In [59] the authors distinguish children from adults based on the face/iris size ratio.
In [60] facial skin regions of Caucasian women are analyzed, a partial least square regression models is built to predict the chronological and the perceived age. They find out that the eye area and the skin color uniformity are the main attributes related to perceived age.

2.3.2.5 Ethnicity

Ethnicity recognition is very much debated because of its ethical and sociological implications. Nevertheless, this trait may be once again relevant for face recognition.
In the context of ethnicity a uniquely defined classification is a difficult and important task. For recognition of Asian and non-Asian faces in [61] machine learning framework applies a linear discriminant analysis (LDA) and multi scale analysis.

Another framework, integrating the LDA analysis for input face images at different scales, further improves the classification performance.
The authors of [62] devise an ethnicity recognition approach based on gabor wavelets transformation, combined with retina sampling for key facial features extraction. Finally a support vector machine classifier (SVM) is used for ethnicity classification providing very good results, even in the presence of varying lighting conditions.

2.3.2.6 Facial measurements

Differently for what happens with anthropometric measures, generally face measurements are considered as a hard biometric and then used for identification purposes [63]. Later studies continue employing facial measurements, extending the concept to 3D data analysis [64].
Recent work on facial soft biometrics is performed on scars, marks and tattoos by the authors in [65]. Other very recent works employ facial measurements and other information to study the correlation between this ensemble of features and a facial aesthetics index rated by a big population of individuals [66].

2.3.3 Accessory soft biometrics

The new soft biometrics definition allows the inclusion of accessories among the aforementioned traits. Accessories can indeed be related to personal characteristics (as sight problems in case of glasses), or personal choices (as adornment in case of jewelry and clothes). One of the first example is clothes color detection. According to the definition of soft biometrics these characteristics can be added to the list of already mentioned traits.

2.3.3.1 Clothes color and clothes classification

Several works focused on the use of color information from clothes to re-identify people in video surveillance scenarios. One of the common ways to discriminate among different targets is represented by histogram-based methods [67].
Content based retrieval systems have demonstrated how those techniques are well suited to retrieve similar images, however they are strongly affected by changes in appearance and illumination [68].
Another positive aspect of the histogram-like methods [69] is that they are straightforward and fast to compute. For this reason those works were taken as baseline

techniques in [70] where the authors present a set of methods that show promising performances considering both the size of the database and the simplicity of implementation of the techniques themselves. Later D'Angelo et al. [17] improved the performance proposing the probabilistic color histogram as new histogram-like descriptor.

An interesting yet preliminary result on clothes categorization is shown in [71] where authors are able to classify clothes in different categories as a side effect of 2D body silhouette analysis.

2.3.3.2 Eye Glasses detection

As for the beard and mustache, also for glasses the state of the art explore ways of removing this attribute so as to ameliorate the automatic face recognition results.
One of the earliest works for glasses detection was performed by [72], they exploit edge detection on a preprocessed gray level image to enhance some characteristics of glasses. Certain face areas are observed and an indicator for glasses presence. The best identifying part is found to be the nose region between the eyes, where the bridge is usually located.
A 3D method to detect glasses frames is presented in [73], where 3D features are obtained by a trinocular stereo vision system. Another approach for glasses extraction is employed in [74], where a face model is established based on the Delaunay triangulation.
Up to now, the best results on glasses detection are achieved on thermal images from [75] where in the data fusion process, eyeglasses, which block thermal energy, are detected from thermal images and replaced with an eye template.

2.3.4 Traits instances and direct measures

One of the most important aspects that characterizes the soft biometric traits is the possibility of using them both by value or as discrete, semantic terms; this is well summarized in Table 2.1 where the traits are classified as Discrete or Continuous value. However, for many of the continuous value establishing a direct correspondence to discrete terms it is straightforward, like pointed out in [11, 76]. One should notice that in those works the nomenclature is different, the authors of [11] refer to *trait's instances* while the authors of [76] to *trait's terms*; in both cases the authors refer to the value assigned to the soft biometric traits.

However, in case of discrete values, the earliest works that draw a definition of traits and propose their classification has been performed by the eyewitness testimony research community. A representative work of such studies is [77] where an analysis is conducted on archives of crime records. The authors draw conclusions on the traits that are often recalled during the eyewitness testimony. Traits like sex, height, ethnicity, and skin color are the most accurate ones. Other works in the same research community mention Hair Color and Length, and a quite interesting part related to clothes color and type.

A special attention was reserved to traits related to the body build, or generally the body appearance. Remarkable is the effort of MacLeod et al. [35] that conduct an exhaustive analysis on the whole body description from both a static and dynamic point of view. From a total of 1238 descriptors, they extract a series of 23 traits that are formulated as a five points, discrete scale. The authors validate their preliminary analysis conducted on the body build thanks to an experiment where annotators are asked to describe video footages. The statistical analysis of the annotations validates the author's choices.

A stronger validation of those results is included in [78] where a similar list of traits is proposed by Samangooei. The author analyzes traits correlation with ANOVA, and successively he uses them to create a content based retrieval (CBR) system based on gait sequences, and face images.

While it is straightforward that the discrete instances of soft biometrics can be generally compliant (if well defined) with a human semantic description, one might think that such assumption is not valid for continuous values. However, the discrete semantic description can be recovered if we consider the statistics of a trait's distribution (e.g. the distribution of stature among French population), or by exploiting experiments similar to the one conducted by [78] to link human semantic annotations and real measurements.
Similar analysis would grant the possibility to store the exact continuous value of a soft biometric trait, while keeping the human compliant characteristic (implicit for soft biometrics). The use of continuous values can enable properties that the semantic classification step can weaken or delete, this concept will be exploited in Chapter 4 where we explore the possibility of performing pruning thanks to our anthropometry based signature.

2.4 Techniques and applications of soft biometrics

We are already using soft biometrics every day without even knowing it. An example is shown in figure 2.3 where among the "hard biometric" information (the two photos) we find detailed description of physical characteristics we previously included within the list of soft biometrics (weight, height, tattoos, and so on). Each

Figure 2.3: The informative file for the missing person "Kristine Nicole Hamilton", among the others one can clearly identify the soft biometric traits. *[Image courtesy of FBI.]*

of the soft biometric traits previously presented can easily find application mainly in the three following areas.

2.4.1 Pruning

Pruning consists in reducing the search space of one biometric identification algorithm so that the template of the person to be identified lays in a subset of the original full *Biometric Database*.

This can be achieved by indexing the biometric database and then using indexes to pre-classify entries. The result of such indexing is usually a signature smaller than the original template that can be used to discard obvious non-matching templates.

In order to compute the performance of pruning algorithms, one has to measure the resulting accuracy and the *penetration rate*, that is to say the fraction of database we are able to isolate to perform the identification. The higher is the penetration rate, the larger is the portion of database considered in the identification phase. The lower this penetration rate, the fewer comparisons will be expected and the more efficient will be considered the pruning system.

Another measure of performance is the *binning error*, thus the error that is made when we mis-classify one trait's value. This usually happens when we choose to divide our feature set into different categories (e.g. the height as short, normal, tall).

One of the main purposes of soft biometrics could be to obtain descriptions of characteristics that are easy to use (e.g. the color of the eye, height, gender) that may help to index the biometric database thus reducing the penetration rate. We explore the use of soft biometrics for pruning in Chapter 4.

2.4.2 Identification

We already discussed how hard biometric traits are well suited for the identification task. On the other hand soft biometrics systems cannot rely on the same distinctiveness power that allows their use as templates for identification.

We can say that soft biometric traits do not provide enough entropy (in the Shannon sense) to have a reliable source of identity.

Nevertheless, gathering together several soft biometric traits we could increase (at will) this amount of information, thus providing a way to perform identification via soft biometrics.

We employ soft biometrics for identification in chapter 5 where we draw the theory of gathering together several traits for increasing the entropy of the ensemble of features, and we provide a real case scenario where we identify users of a video-surveillance system.

2.4.3 Semantic description

Another important aspect of soft biometrics is their human compliant characteristic which makes them conform with natural human description labels. It means that a computer that extracts soft biometrics identifiers can directly communicate those values to a human operator without any form of processing technique, and vice-versa.

Thanks to this property a better human machine interaction is possible as soft biometrics enables the extraction of semantic information from the image/video.

We investigate the possibility of extracting soft biometrics information with the purpose of people description in Chapter 6.1 and 6.2 sections.

CHAPTER 3

Body soft biometrics extraction

In this chapter we will examine the techniques for body soft biometrics extraction. We will start with anthropometric measures and then we will move through height, weight estimation and finally gender. Some of these techniques are directly related to work that we have published.

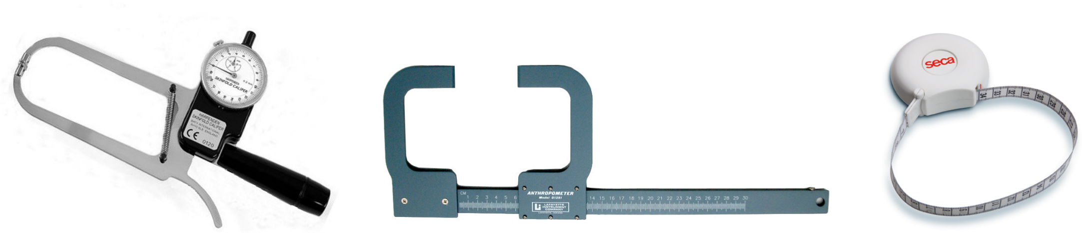
3.1 Anthropometry and biometrics

If we take into account both biometric and anthropometry terms' meaning, we can easily realize the strong bond that links the two concepts together. As seen in the previous chapter the first one means "measuring life" while the second one relates to the measurement of the human individual.
Widely used in the past for the purposes of understanding human physical variation (e.g. paleoanthropology and craniometry), nowadays anthropometry is better recognized as a crucial science for industrial design, clothing design, and ergonomics. In those fields the statistical distributions of body dimensions are exploited to optimize and to facilitate the use of products.

A part from the scientific studies, humans are proved to utilize anthropometric analysis in everyday life. We characterize body into three main phenotypes (Ectomorph, Endomorph, and Mesomorph) which help us during the recognition process. For example, a physical description is of utmost importance in many cases of missing persons or for crimes investigations.

As previously mentioned, the first to use and to introduce anthropometry as a recognition tool was Alphonse Bertillon. He is considered the father of all Biometrics because his technique gave birth to all the following studies on identification through the use of personal information provided by biometrics identifiers (e.g. Galton's fingerprint identification system).

In order to extract anthropometric information a straightforward way is by direct measure. It consist in measuring lengths and girths of the human body by using specific tools (calipers, or body tape meters, see figure 3.1).



(a) (b) (c)

Figure 3.1: Images of three instruments used in body anthropometry. Figure (a) represents a caliper used to measure skin-fold thickness. Figure (b) a caliper (also called anthropometer) used to measure lengths, and (c) a closed body tape meter used for girths mensurations.

Additionally to manual inspection, it is possible to adopt some computer vision techniques to obtain the automatic and semi-automatic extraction of body parts measures.

In [79] we employed a semi-automatic method to extract body measures.

In order to exploit the measures to estimate subjects' weight, we performed an estimation of body parts' sizes. Since the important information about the 3D shape of the body is not available in 2D pictures, an estimate of the anthropometric measurements was computed. For what concerns the circumferences involved in the computation, we considered the width of the body part (upper arm, leg, waist) as the diameter of the cylinder that approximates that particular body part. This is obviously a straightforward approximation of the real measure that introduce errors in the process, nevertheless the results of our estimation look promising.

Although several techniques are present in the mature literature of body parts detections, like [80], a manual tagging of the markers used for length information was performed directly on the images. The experiment we conducted can be mainly divided in three steps. The first step is represented by height estimation:

Figure 3.2: An example of the two poses recorded during our database collection. A frontal and a profile pose are experimented. One should notice for the second subject how his clothes are hiding the real shape of the body.

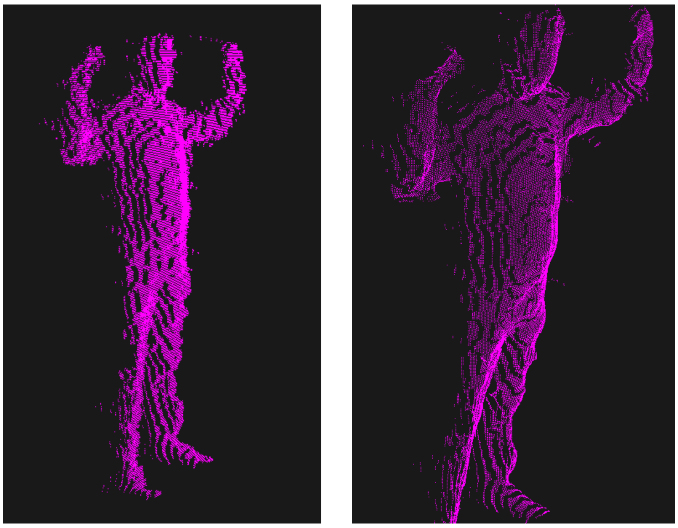
for that we exploited a similar method to the one provided by [16]. We compared the height of the subject with an object of known height inside the scene. The second phase involves the approximation of the other measures according to the height estimated in the previous step: in our case we studied the proportions of each different body part with respect to the height estimated measure.

Another way of extracting anthropometric measures is to exploit 3D information provided by special sensors. The market has lately seen dramatic improvements in the use of time of flight cameras and structured light cameras. Those devices enable automatic perception of depth, and thanks to calibration they are able to transform the scene into its 3D equivalent. The most evident example of such devices is the Microsoft Kinect sensor, originally developed for gaming and natural user interaction (NUI), it is more and more used by researchers for its 3D capabilities.

We exploited its 3D sensing functions in [81, 82] to provide our automatic soft biometric extraction system with the possibility to measure anthropometric traits of subjects in front of the RGBD sensor.

To attenuate both the error produced by the random noise of the Kinect sensor, and to minimize the effect of depth quantization, we employ a surface smoothing technique provided by [83]. In figure 3.3 we can observe the result of the smoothing algorithm.

In order to extract anthropometric measures we must obtain the location of each



(a) (b)

Figure 3.3: An example of two cloud points: (a) without any preprocessing, and (b) with surface smoothing applied. It is clear that the smoothing minimize the quantization that characterizes the first cloud point, and provide better understanding of the body shape.

body part. For this we resort to a person detection and body part tracking algorithm provided by [84]. Those two algorithms combined make possible to firstly detect the person has he/she moves, and then to track each of his/her body parts across time. A list of body parts is provided in figure 3.4.

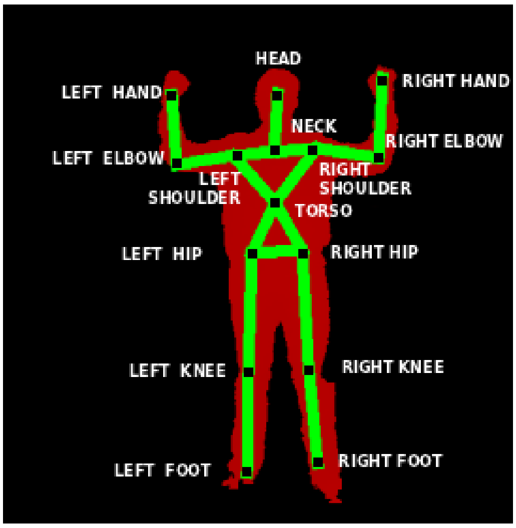


Figure 3.4: The image shows the joints for the skeleton provided by PrimeSense OpenNI framework.

Once the skeleton is provided, the body parts segmentation is performed. The following algorithm is used which is based on the distance of each element of the point cloud from the segment corresponding to the body part considered.

If we consider a generic point $C = (C_X, C_y, C_z)$, and a segment $\overline{AB}$ formed by the points A and B, we can define a conditional test to check if the point lies within the segment using the parameter r whose formula is the following

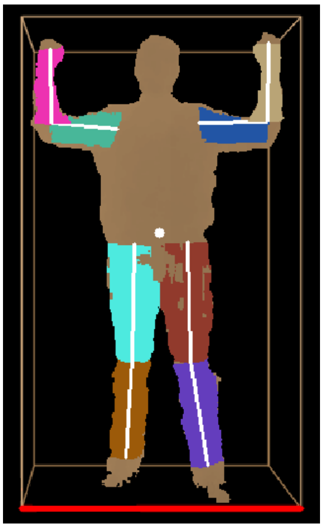
$$r = \frac{(C_x - A_x)(B_x - A_x) + (C_y - A_y)(B_y - A_y) + (C_z - A_z)(B_z - A_z)}{(B_x - A_x)(B_x - A_x) + (B_y - A_y)(B_y - A_y) + (B_z - A_z)(B_z - A_z)}.$$

We obtain that $P = A + r(B - A)$ is the perpendicular projection of C over $\overline{AB}$; in the equation the parameter r has the following properties

$$\begin{array}{cl} r = 0 & P \equiv A \\ r = 1 & P \equiv B \\ r < 0 & P \textit{ is on the backward extension of } \overline{AB} \\ r > 1 & P \textit{ is on the forward extension of } \overline{AB} \\ 0 < r < 1 & P \textit{ is interior to } \overline{AB}. \end{array}$$

Once we verify that P lies on the segment, we can compute the distance of C from its projection. We segment each body part by using different thresholds (to model the different sizes of the limbs), finally we obtain a result similar to the one shown in figure 3.5. As all the body parts are segmented we proceed with the extraction of 6 anthropometric measures. Thanks to the calibration parameters embedded in the Kinect camera we are able to transform each of the points sensed in the scene into a triplet of values in millimeters. This enables us to directly compute distances, and for this reason we just consider the skeleton previously tracked to compute the arm and leg length.

A different consideration has to be made for circumferences because the Kinect sensor is able to perceive only the frontal surface of the body, loosing any reference to the back side of the person. Because of this drawback, we have to find correspondences between the sensed and the real dimensions. To do so, a set of 5 candidates were drawn from our dataset to be used as “training” step. By computing multiplicative factors that can extrapolate measures from real values. The parameters are shown in Table 3.1. This is a straightforward approximation that worked in our case; however, increasing the complexity of the extraction algorithm



(a)

(b)

Figure 3.5: The image shows the outcome of our limb segmentation based on the skeleton provided by PrimeSense OpenNI framework.

Table 3.1: The estimation of extrapolation factors that link real with sensed circumferences.

Measure	Extrapolation factor
Arm circumference	2.7
Waist circumference	1.5
Leg circumference	2.8

(e.g. using multiple Kinect sensors, or allowing the person to turn around) would increase the precision for each measure.

In Table 3.2 we show the performances that our algorithm achieved in terms of mean absolute error. The results are computed on a self recorded database composed of 15 subjects. The measures will later on be used to extract weight estimation from 3D data.

3.2 Height estimation

Height estimation is a fundamental part of soft biometrics. This trait is by far the most commonly used when people describe each other [76]. Its extraction is however not free from some issues. First of all it depends on the mean used to collect images and videos used for the analysis. RGB images and videos are generally

Table 3.2: The estimation error for the automatic analysis of 15 people of the 3D dataset we recorded using the Kinect sensor (male/female subjects).

Measure	**Absolute error**
Height	1.9 cm
Arm length	3.6 cm
Arm circumference	3.2 cm
Waist circumference	8.4 cm
Leg length	2.9 cm
Leg circumference	1.7 cm

not capable of gathering 3D information unless some preprocessing steps are performed (e.g. camera and scene calibration), while 3D cameras or stereo-systems are embedded with those capabilities.

In the former case some techniques were studied that allow an automatic system to process images and videos in order to extract height estimations of objects in the scene. The forerunner of single view metrology is Criminisi [16, 85]; his approach provides ways of computing lengths of elements within an image thanks to the implicit information provided by the content of the scene. Provided the image's vanishing points and vanishing line we can compare the dimensions of objects sharing the same plane in the scene up to a scaling factor; the approach is well summarized in figure 3.6.

Improving this technique, the work in [26] describes an algorithm that is able to provide very low error rates for height estimation. The combination of a calibration step performed via a known pattern placed on the ground of the scene, plus an intensive use of statistical methods, helps to refine the height's measure. Thanks to those elements the precision of 1 cm is achieved.

If we consider sensing techniques more sophisticated like 3D scanners or Kinect-like devices the height estimation problem becomes easier to solve. Those systems are usually composed of special cameras that are already calibrated and that transform the sensed information into a 3D representation that preserves spatial information.

Thanks to such technologies is possible to measure distances between points that are *seen* by the 3D sensor. Provided the aforementioned characteristics, the height estimation problem is transformed into the detection of the head and feet of the subject. Once this input is provided, the height of a standing subject can be computed as the distance between the two points. In case of other poses

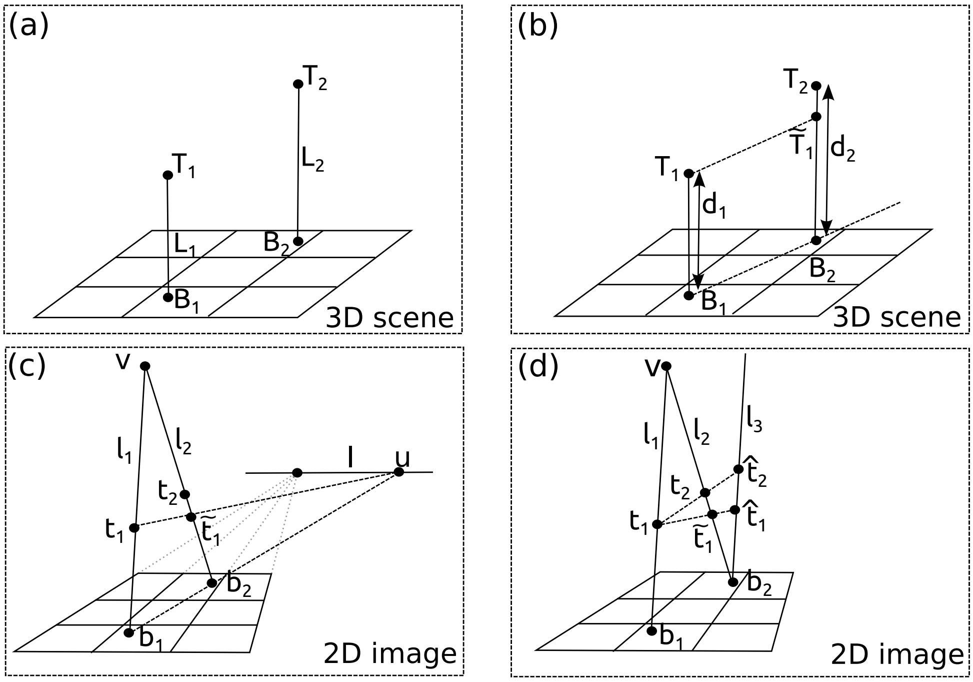


Figure 3.6: The image shows an illustration of how to compute length ratios of parallel scene lines. (a) Vertical lines segments L_1 and L_2 have length d_1 and d_2 respectively, their base points (B_1 and B_2) are on the ground plane. Thanks to the information within the scene we can compute the length ratio $d_1 : d_2$. (b) In the scene: L_1 segment's length may be transferred to L_2 by constructing a line parallel to the ground plane that generates the point $\tilde{T_1}$. (c) l is the ground plane vanishing line, and v the vertical vanishing point, and u is the horizontal vanishing point. Thanks to this information we can establish the correspondences of points in one segment onto the other one ($\tilde{t_1}$). (d) The line l_3 is parallel to l_1 in the image. After construction of the points $\hat{t}_1$ and $\hat{t}_2$ we obtain the distance ratio $d(b_2, \hat{t}_1) : d(b_2, \hat{t}_2)$ (that is the computed estimate of the ratio $d_1 : d_2$).

more complex computations have to be envisaged like estimating height from limbs measurements, or fitting a 3D model to the cloud point and then computing the height on the 3D adapted model.

3.3 Weight estimation

Over the centuries, two different methods have been adopted to measure the weight of an object: the spring scale and the balance. While the first one measures the local force of gravity that acts on the object, the second one is used to compare the weight of the unknown object with the one of a known standard mass. The way of measuring human body weight usually belongs to the first case. There exist a variety of practical situations where having an estimate of a person's weight is extremely useful. Unfortunately, in many of these situations none of the aforementioned techniques can be adopted.

The weight of a person is often considered as an indicator of his/her physical aspect and, in many cases, of his/her health conditions. For that reason, the medical community showed interest in the topic, and several works already analyzed the problem of weight estimation under difficult conditions [2, 30, 31]. During medical emergencies, measuring the weight of a patient can be difficult, due to the impossibility to move him/her, or also to some disabilities of the patient (e.g. some types of mental disorder). Often these situations are managed via the visual estimation of the weight of the patient performed by a trained personnel [2]. It is clear that a visual estimation cannot provide the precision of a scale, and in some cases a more accurate estimation can make the difference (e.g. in case of anesthesia administration).

Weight estimation could have a considerable value also in forensic science. Together with some others physical traits (e.g. height, color of the hairs, body-build) the weight is often part of the first description of a fugitive and could be considered helpful for an automatic search through video-surveillance records [78].

Scientists from the forensic domain have shown interest in height and weight estimation in crime scenes. Regarding height, several studies [86, 87] tried to estimate the height of deceased individuals from measurements of their long bones, exploiting the high correlation existing between this measure and the height. A regression model was then built that could estimate subjects height. In [33], instead, the authors utilize the evidences from crime scenes to collect footprints and link them with the weight of the suspects. The study concludes that the body weight has a strong correlation with the footprint.

Similarly to these last studies, our work explores a new way of estimating the weight of a person using images coming from a camera. The main idea has been extrapolated from the medical and forensic community that already used some anthropometric measurements to estimate the height of a subject. In our case using measurements of anthropometric traits we are able to estimate the weight. We will demonstrate that an acceptable level of precision can be reached by our method.

We present hereafter, the related works in Section 3.3.1. Section 3.3.2 and 3.3.3 show respectively the proposed method and the database exploited for our experiments. Section 3.3.4 presents the results of our estimation process, additionally a study is made to validate our framework under noisy input data. Successively an experiment is described where 2D images are used to verify our estimation.

3.3.1 Related works

As previously seen, the estimation of weight has never been fully explored by the biometric community. Weight is still a challenging feature to be visually extracted. In this part of our work we propose a study about the feasibility of the weight estimation through the use of anthropometric measures. For that reason we have to train an estimator so as to compute the weight starting from the body visual appearance. To describe one's body silhouette, a representative set of features has to be defined. We propose a method that, starting from these visual features, creates a regression model for weight estimation.

Although our first system for weight estimation was later improved by using non linear regression algorithms (more details in Chapter 6), in this section we explore our first approach based on multiple linear regression thus hypothesizing a linear dependency between measures and body weight.

3.3.2 Inferring the model

In order to study the feasibility of weight estimation, a model has to be deducted that could perform the estimate from anthropometric measurements. The main idea was inspired from the works of height estimation previously mentioned (Section 3.3). Those works exploit different anthropometric measures, highly correlated with the height of the human body, to estimate subjects height.

Our hypothesis is that this set of measures should be comprehensive of the whole body shape (upper and lower part) and then being reasonably correlated to the weight. For this reason we choose a set of anthropometric features related to various part of the body and that describe the full body (height), the torso (waist), and the limbs (arms and legs measures). This set of features is depicted in 3.7.

Our work can be easily justified by the inversion of the body mass index (BMI) equation. This value is indeed the one that mostly correlates with the amount of fat in the human body. Its formula $\mathit{BMI} = \frac{\mathit{Weight}[kg]}{\mathit{Height}.^2[m.^2]}$ can be reverted to show that $\mathit{Weight} = \mathit{BMI} \times \mathit{Height}.^2$. Since BMI model also the physical aspect of a person, we

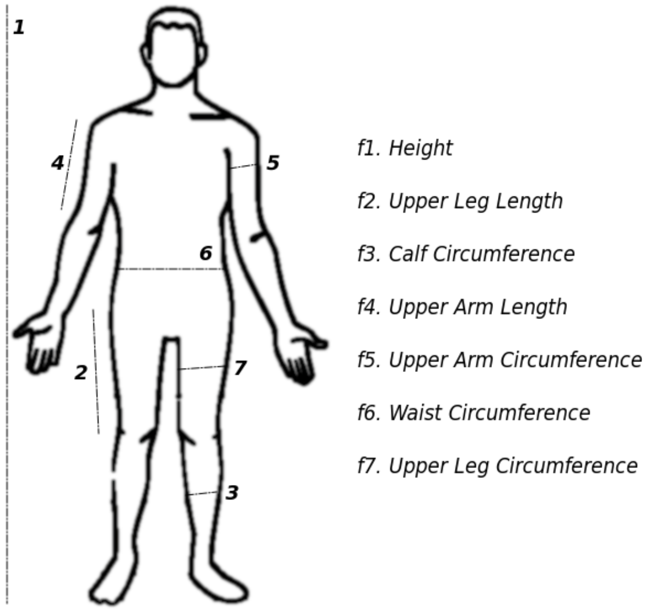


Figure 3.7: Measures taken into account in our work on weight estimation.

can derive that the physical description provided by the anthropometric measures should be enough to extrapolate a weight estimation.

Our purpose is to understand how each of these features, or each of their possible combinations, impacts the estimation phase. We propose to create a simple model, or multiple models that, starting from these measurements, can provide a good estimation of subjects' weight. As no precise relation could be found between the anthropometric measures and the weight, a linear dependency was assumed. To create such a model, we base our approach on multiple linear regression analysis, as this kind of analysis provides a powerful tool for data fitting. This is formally described as follows:

$$D = \{y_i,\, x_{i1}, \ldots, x_{ip}\}, \qquad i = 1, \ldots, n \tag{3.1}$$

where D is a given data set composed by a dependent variable (y being the weight) and several (p) input variables (x being the anthropometric features). A linear dependency can be inferred between the dependent variables and the input variables $x_{i1} \ldots x_{ip}$ up to an error term ε. This last element is a random variable that models the constant term of the linear relation among the variables:

$$y_i \;=\; \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \varepsilon_i \qquad i = 1, \ldots, n \tag{3.2}$$

To find the solution, the ordinary least squares method is adopted that minimizes the sum of squared residuals. By fitting the model to the data, one finds the optima β coefficients which shape the function in order to minimize the estimation error.

An analysis is conducted where the fitting is performed while varying the number of involved features from one to seven, and considering all the possible feature combinations (i.e. $2^7 - 1 = 127$).

Hereafter follows a description of the database used for our analysis. Then the outcome of our experimental analysis is going to be presented.

3.3.3 Sources of data

As already introduced, weight is significant for the medical community; for this reason it is included in several databases used in this domain. Creating such databases is generally an important investment and for this reason they are generally considered a precious resource. The Centers for Disease Control and Prevention (CDC), a United States federal agency under the Department of Health and Human Services is one of the organizations whose main purpose is to protect public health and safety in the United States. The CDC focus its attention on studying and applying disease prevention and control systems. It promotes yearly the National Health and Nutrition Examination Survey [21] (NHANES), a study devoted to understand health statistics behind a representative sample of the American citizens.

NHANES findings are the basis to determine the prevalence of major diseases and of risk factors, to define US standards for measurements like height, weight, and blood pressure. The importance of this database lies on its size, that makes it a suitable statistical source of information. If we consider the period taken into account for our studies that goes from 1999-2008, we benefit of more than 27000 records. This dataset is rich of heterogeneous information regarding everything related to health status of the subjects (medical analysis, demographics, and so on). A prevalent part of this dataset is composed of a set of anthropometric measures that is taken under controlled conditions.

These measurements are representative of the physical aspect of the subjects that participated to the survey. From all the measurements available in the database, we only keep the ones related to the physical aspect, while discarding the others. The collection of the data set was conducted by a trained group of people and by guaranteeing the same measurement conditions, reasonably excluding the possibility of errors. The list of features can be visually verified in figure 3.7.

Some other features are present (i.e. subscapular and triceps skin fold), that the medical community already reported as being related to the quantity of fat, and then to the weight of people. Nevertheless, we discarded them, since our analysis is intended to create a baseline for future techniques that will try to estimate the weight of a person only from its visual aspect.

Since for some individuals the database presents missing data, the data set is filtered in order to keep only complete set of measurements. Additionally, we do not consider in our study the range below 35 and beyond 130 kilograms since they are not significantly represented in the data set. We point out that even after filtering the non relevant observations, the population taken into account still counts a large number of subjects.

3.3.4 Experimental results

In this section the experimental results are presented. The aim of our work is to identify the best set of features for the problem of weight estimation in order to minimize the error and to assess its feasibility.

The section is divided as follows: the first part is related to the experiments performed in ideal conditions (i.e. the data as they are present in the NHANES dataset), the second part refers to the experiments assessing the robustness of our approach against the noise. An additional experiment is conducted where the measures are estimated from the images of a standard resolution camera commonly employed in video surveillance.

3.3.4.1 Results under ideal conditions

We consider ideal conditions the use of measurements not affected by error. In this sense the original data present in the dataset are used. To perform the experiments we divide the database in training and testing set, respectively 70% and 30% of the available data. The training lead to the values of the β coefficients used for the estimation. For example we report hereafter the β coefficients when all the features are considered:

$$\begin{aligned} \text{weight} &= -122.27 + 0.48f_1 - 0.17f_2 + 0.52f_3 \\ &\quad +0.16f_4 + 0.77f_5 + 0.49f_6 + 0.58f_7. \end{aligned} \tag{3.3}$$

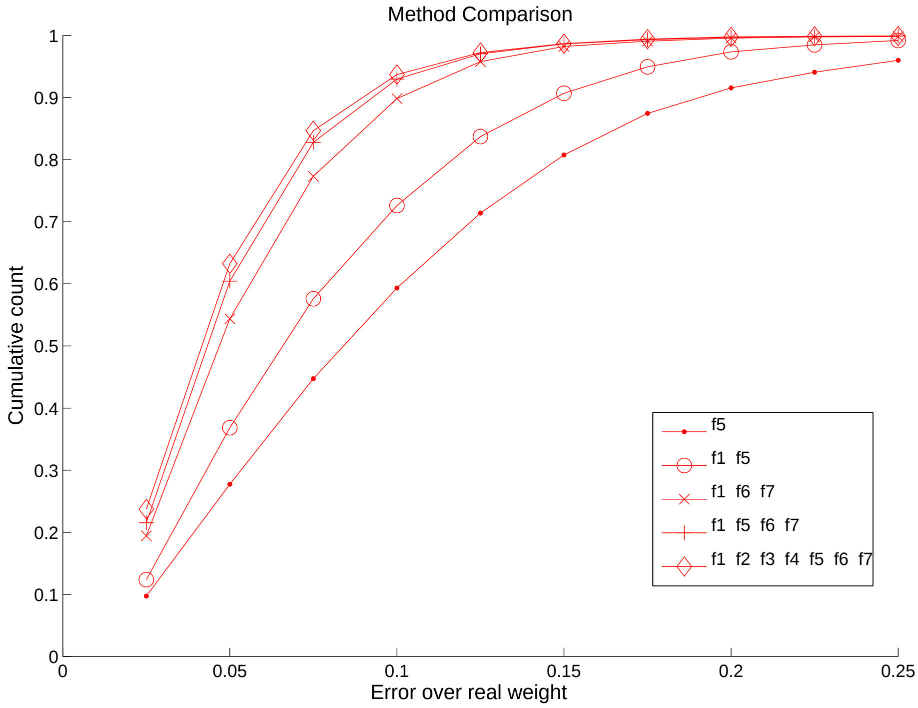


Figure 3.8: Results for the multiple regression approach. It is evident that more information provided to estimator helps to increase the precision. (*For the meaning of each number please refer to figure 3.7*).

For the list of features please refer to figure 3.7. In this case the significance of each β coefficient was confirmed through an analysis of the t-statistics, that is to say, being the P-values all close to 0 all the features are significant to the estimator.

In our analysis we explore the full space of possible estimators (all the 127 possible combinations of seven features). For the sake of clarity, we decide to report only the most representative results.

To assess the quality of a feature set we used the cumulative function of the relative mean squared error distribution. Such evaluation allows us to understand the real performance of each model as it shows the cumulative number of estimations under a certain value of mean squared error (i.e. $E = \frac{|\tilde{W}-W|}{W}$ where W is the real weight and $\tilde{W}$ is the estimated one). As quality measure for the curves (i.e. to decide which curve is better than the others) we consider the cumulative value for the error range $\pm 10\%$ since this is the one considered acceptable in the medical community.

Figure 3.8 presents the results of our estimation attempt. The first model is based on the circumference of the upper arm (*f5*). Among the available features, this is indeed the most correlated to weight, thus it is better at the estimate.

One can notice how the results are more precise as the number of features involved in the estimation process grows. As a result the method that better approximates the weight is the one that exploits all the variables (see equation 3.3).

Table 3.3: Performance comparison between our estimation system and the human visual estimation in [2].

	Error range	
	$\pm 5\%$	$\pm 10\%$
Patients	74%	91%
Nurses	44%	78%
Physicians	33%	59%

As a comparison we report in Table 3.3 the results of an experiment [2] conducted at the Western Hospital of Melbourne. The experiment aims at assessing the visual estimation performance when judging someone else's weight. A total of 1137 patients are involved in the study. The first task consists in measuring the precision of weight self assessment of the patients. For this task, without using a scale, each patient has to report his/her own weight. Later on the nurses and physicians are asked to estimate the patient's weight. The estimate is only performed visually, that makes this experiment comparable with the one we propose. Also in this case the error is computed as percentage of the original, real weight ($E = \frac{|\tilde{W} - W|}{W}$).

As expected the best precision is always achieved by the patients themselves. They are indeed the ones that has a better understanding and prior knowledge about their weight. As explained by the study, nurses obtain the second best result as they are used to assist patients constantly and for that reason they acquire with time the ability to deduce one's weight. On the other hand, the physicians do not achieve the same results and their error is more spread (up to $\pm 20\%$ of the original weight).

As per Table 3.3, the nurses result of $\pm 5\%$ of the actual weight is accomplished for 500 subjects, that over 1137 patients represents the 44% of the data set.

From our analysis (figure 3.8) we can observe that better performance are achieved by our method. Using the multiple regression approach we are able to deduce with the same error ($\pm 5\%$) the weight of more than 60% of our data set. An important result is also the consideration regarding the error at $\pm 10\%$ in figure 3.8; at this level our system starts to perform better than the patients (93% vs. 91%).

We consider this as a significant achievement since, in self evaluation, people are aware (most of the time) about their own weight; that also explains their precision in self-estimating.

3.3.4.2 Biased measures analysis

To the best of our knowledge, our analysis is the first attempt of making a comprehensive study on weight estimation exploiting anthropometric data. However, the analysis conducted in the first part of this section is a bit optimistic since the reported precision can be achieved when the model is applied on data that are not affected by error, or when this is negligible.

We obviously understand that such analysis is far from a real case, which not always include the possibility of precise measurements, or good measurement tools (e.g. video surveillance, forensic analysis of images, medical emergencies, and so on). We are also aware that an application to perform weight estimation, or a system exploiting this estimation to accomplish other tasks, could not have the possibility of having the same precision of the NHANES data set features.

Thus we conducted an additional study to evaluate the loss of performance and the magnitude of errors one can expect from real case scenarios.

For this purpose an analysis of the impact of data biased by a random noise is of interest. We describe in this section the results of such a work and we discuss the impact of noise at different levels of magnitude. We are going to see that, notwithstanding the loss in terms of precision, the system can still be considered good enough to be used as automatic system for the computation of an acceptable weight estimation where the working conditions do not allow to obtain precise measurements.

Experimental setup The error analysis was conducted by adding the data with a random error as follows:

$$\begin{aligned} \tilde{X} &= [x_1, \cdots, x_n] + [\eta_1, \cdots, \eta_n] \\ |\eta_i| &\leqslant x_i\,\epsilon, \quad \epsilon = \{0.05, 0.10, 0.15\}\,. \end{aligned} \tag{3.4}$$

where the x_i are the original features of the human body, and the η_i are the corresponding noise terms. The magnitude of such noise is computed as a random value between $-x_i\,\epsilon$ and $+x_i\,\epsilon$, where ϵ represents the factor that indicate the error range.

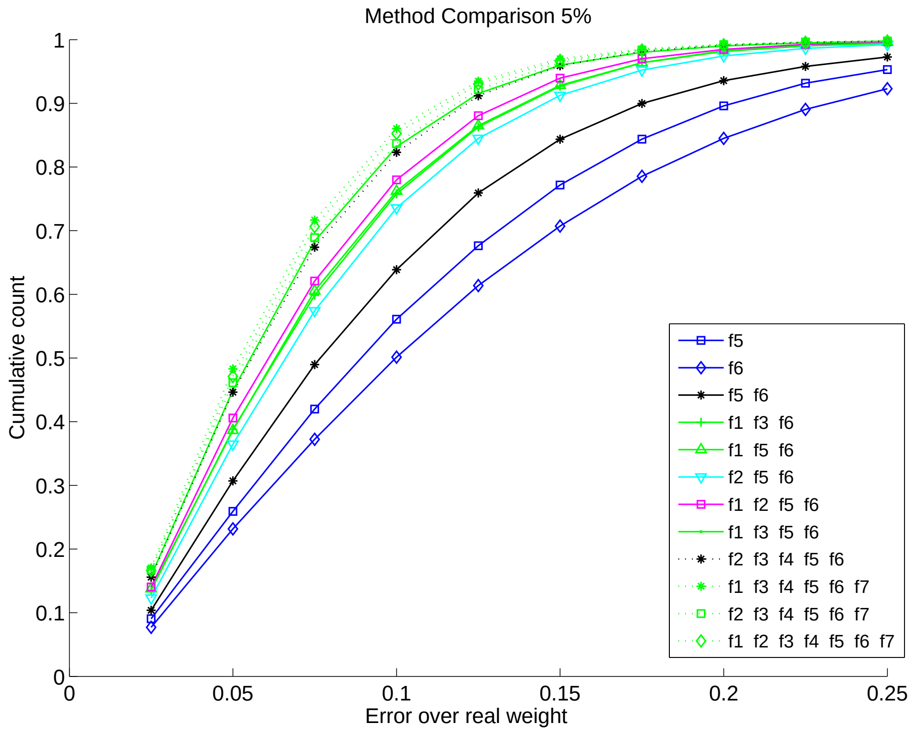


Figure 3.9: Estimation response to data biased by 5% magnitude noise.

As one can see from equation 3.4, for our case study we take into account three different values of ϵ. The choice of these values allows to track the response of the system as the error increases in magnitude.

As before, due to space constraints and clarity we do not present all results; for that reason we selected the ones we considered the most representatives.

In figure 3.9 one can observe the impact of 5% magnitude noise. The system still performs better than the professional result of [2]. The two other figures (3.10,3.11) show the impact as the magnitude of noise increases. They are included in order to understand how and how much the noise makes the performance decrease. What we can observe is that the estimate gets more linear as the noise increases of value. Nonetheless, the system achieves good results even in our extreme case (i.e. 15% of magnitude); our system still has similar performance as the non-specialized human estimate (i.e. physicians in [2]).

3.3.4.3 Real case analysis

In order to measure the estimate performance in a real case scenario, and to confirm the results obtained from the analysis on the impact of noise, an experiment

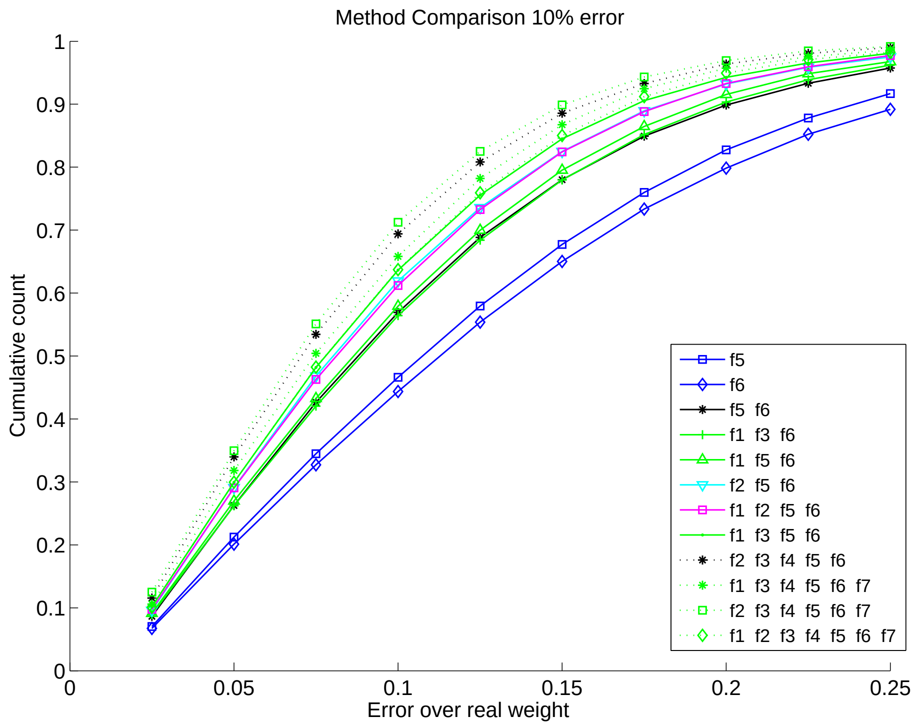


Figure 3.10: Estimation response to data biased by 10% magnitude noise.

is conducted on real images.

To the best of our knowledge, databases presenting weight as ground truth and that could fit our needs are not available to the community. For this reason we create a test set collecting some pictures of persons using a video surveillance camera to recreate a possible application scenario.

The database is composed by 20 different subjects (15 males and 5 females) whose pictures are taken at fixed distance from the camera. Two different poses are experimented: a frontal shot and a profile one, a total of 40 pictures is available in the database (see an example in figure 3.2). A common scale was employed to measure subjects' weight used as ground truth to compare the result of our estimate.

To obtain the measurements the method explained in the first part of section 3.1 is used. This method consists in extracting landmarks from the images manually and then estimate each anthropometric measure accordingly to a preliminary step that measures height.

In order to capture again the performance of our weight estimator, we consider all the 127 equations previously collected by our theoretical study. The outcome

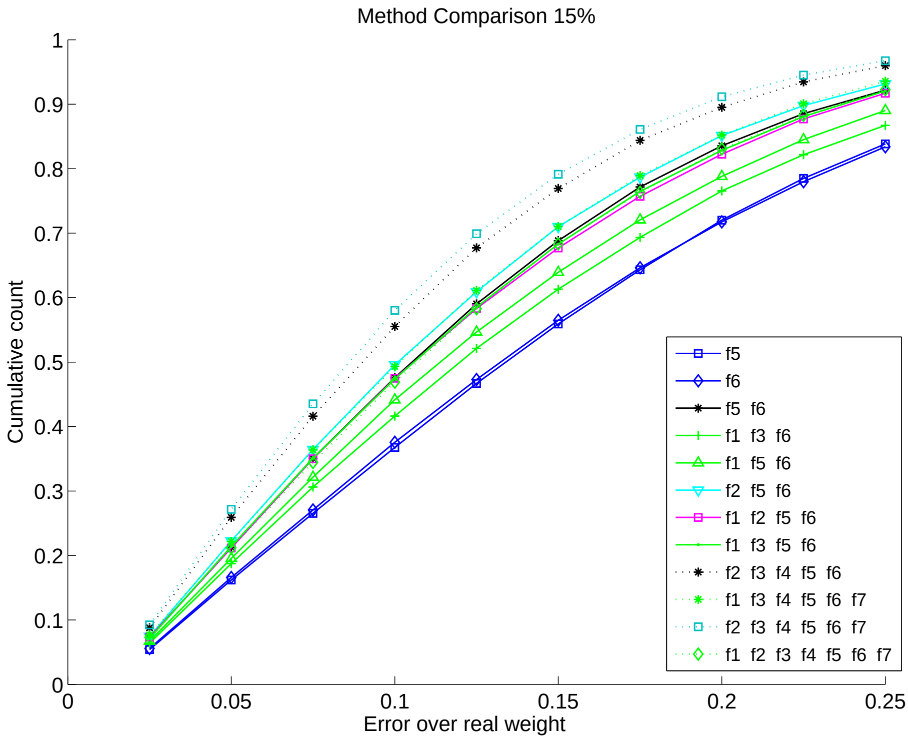


Figure 3.11: Estimation response to data biased by 15% magnitude noise.

of each formula is examined and the results ranked in terms of distance from the ground truth. We obtain that the best results are provided by the formula that does not include the calf measure (*f3*). This outcome justifies our previous observation that even if the best estimator is the one considering all features; as we consider noisy anthropometric measures, the contribution of each single measure changes.

In this case the unreliability of calf measure is due to the style of trousers used by subjects. Large trousers increase the uncertainty about the part of the leg between the knee and the foot making poorer (though manually marked) the position of the landmarks for calf estimation and, subsequently, the performance of the system that considers this value. As appears clearly visible from the example images 3.2, subjects' physical aspect varies a lot. For example, the loosen shirt and trousers of the second subject increases the error probability as they hide the real shape of the body parts. The clothes obviously represent a very challenging aspect for our algorithm as their contribution is an important source of error in the estimate of the body parts' measures.

The outcomes of our analysis are summarized in Table 3.4. As we can see, the results vary in the whole output range. The average estimate error for this

Table 3.4: The estimation results obtained from our database of people. The error is shown in percentage with respect to the real weight of the person.

Subject	Error %	Subject	Error %
1	1.23	11	1.38
2	1.79	12	4.94
3	6.19	13	3.15
4	8.97	14	2.17
5	6.48	15	8.51
6	0.15	16	2.11
7	8.66	17	2.16
8	2.01	18	3.44
9	1.23	19	0.21
10	4.47	20	16.25

experiment is 4.3%, it is then confirmed that our system is able to estimate the weight from visual clues with an approximation of $\pm 5\%$ of error over the real weight of the subject.

3.4 Gender classification

Gender classification includes the sum of techniques used to detect and classify a range of phenotypic characteristics typical of one of the two main sex classes (female, male). Phenotypes refers to the observable characteristics which are embedded into our DNA (Genotype) and that define our human physical aspect.

We have seen in Chapter 2 that gender recognition is a well established and explored topic in both face and body contexts. The usual approach for face based gender recognition systems is to exploit low level features directly extracted from the image of the face. Features are usually processed thanks to machine learning tools like Multilayer Perceptron (MLP) as in [88] or Support Vector Machines (SVM) as in [89, 90].

On the other side, the main approach used in body analysis focused more on gender recognition through gait analysis like in [39]. However, to exploit gait we need a full video sequence to be able to recognize the gender of a subject, while in the face approach we need to have a close up view of the person's face. By exploiting anthropometric measures instead, one could be able to perceive measures using only one picture of the person taken from distance.

Considering the better performances obtained in the weight estimation approach using anthropometric measures in combination with MLP, we considered the same technique for gender classification.

In this case we explore the possibility of classifying the gender of subjects according to their anthropometric set of measures. To this purpose we use the MLP as binary classification tool.

Also in this case measures are scaled to the range of $\{-1, +1\}$ and provided as input to the neural network properly dimensioned (in this case six neurons are used for the hidden layer). Finally we obtain the performance visible in Table 3.5.(a) where we can see that we overpass significantly the random choice of 50%. However those performances did not mapped successfully in the real case; indeed, by using the measures extrapolated by our automatic system, the performance were drastically worse. Unfortunately, the noisy extraction of body parts' circumferences affected the classification performance of the MLP. To solve this issue for our automatic gender extraction application, we resorted to the use of length measures, that demonstrated being more reliable. In Table 3.5.(b) we show the results of a similar approach. In this case the results applied to the anthropometric measures automatically extracted matched our expectations.

	Male	Female
Male	89.8%	10.2%
Female	11.9%	88.9%

(a)

	Male	Female
Male	85.5%	14.5%
Female	19.8%	80.2%

(b)

Table 3.5: The two tables show confusion matrices for two different experiments. (a) All measures from the NHANES database are considered, the single hidden layer of the MLP is composed of 6 neurons. (b) 3 Measures are used (Height, Arm length, Leg length), the single hidden layer of the MLP is composed of 3 neurons.

Even if the size of the NHANES is a precious characteristic that allows our results to be statistically valid, nevertheless its set of measures is quite limited. To test at which extent one could use anthropometric measures for the gender recognition problem, we explored the use of a secondary database collected by the authors of [3] at the U.S. Naval Postgraduate School in Monterey, California.

The database is composed of 507 individuals between 20 and 30 years old, all physically active. It shows an interesting extended set of measurements with respect to the one of the NHANES database used for weight. Unfortunately the dataset is not big enough for generalizing a system for weight estimation, neither a general framework for large scale data analysis; but it is sufficient to demonstrate the performance of gender recognition via anthropometric measures.

The set of measures included in the dataset are the following, for each of the features the corresponding distribution is shown in figure 3.12.

Diameters measures:

- Biacromial diameter, or "shoulder breadth"
- Biiliac diameter, or "pelvic breadth"
- Bitrochanteric diameter, or "hips breadth"
- Chest depth between spine and sternum at nipple level, mid-expiration
- Chest diameter at nipple level, mid-expiration
- Elbow diameter, sum of two elbows
- Wrist diameter, sum of two wrists
- Knee diameter, sum of two knees
- Ankle diameter, sum of two ankles

Girth Measurements:

- Shoulder girth over deltoid muscles
- Chest girth, nipple line in males and just above breast tissue in females, mid-expiration
- Waist girth, narrowest part of torso below the rib cage, average of contracted and relaxed position
- Navel (or "Abdominal") girth at umbilicus and iliac crest, iliac crest as a landmark
- Hip girth at level of bitrochanteric diameter
- Thigh girth below gluteal fold, average of right and left girths
- Bicep girth, flexed, average of right and left girths
- Forearm girth, extended, palm up, average of right and left girths
- Knee girth over patella, slightly flexed position, average of right and left girths
- Calf maximum girth, average of right and left girths
- Ankle minimum girth, average of right and left girths
- Wrist minimum girth, average of right and left girths

Other Measurements:

- Height (cm)

- Age (years)
- Weight (kg)
- Gender (1 - male, 0 - female)

We can clearly see that the list spans over the entire human silhouette and can be used to describe peculiar details of someone's body. However, we will demonstrate that such an amount of data is not completely relevant.

Considering the entire database, we divide it into train, validation, and test set respectively of 50%, 15%, and 35%, keeping the two classes evenly distributed along the different set.

We experimented with different configurations of parameters of which we provide in Table 3.6 the results of the testing set.

In case of Table 3.6.(a) all the measure were tested, the deep knowledge provided by the quantity of measures is reflected into the very high performance of the recognition rates (99.4% in average for the two different classes). However the quantity of measures used is not really necessary as can be seen in Table 3.6.(b) and (c). Here the features were empirically selected to map the canonical measures that humans learn when gender has to be distinguished by silhouette. The areas of the Chest, Waist and Hips are then explored, considering first the diameters (b) of these body parts, then the circumferences (c). In both cases we can see how the average classification rate drop of a negligible amount, keeping the performance above 90%.

	Male	Female
Male	99.0%	1.0%
Female	1.3%	98.7%

(a)

	Male	Female
Male	93.50%	6.5%
Female	5.9%	94.1%

(b)

	Male	Female
Male	93.3%	6.7%
Female	1.1%	98.9%

(c)

Table 3.6: The three tables show the confusion matrices for different experiments. (a) All measures are considered (without considering Age and Weight), a MLP with one hidden layer of 10 neurons is employed. (b) 3 Measures are used (Biacromial, Biiliac, Bitrochanteric diameters) and a MLP with one hidden layer with 2 neurons. (c) 3 Measures are used (Chest, Waist, Hip circumferences) and a MLP with one hidden layer with 2 neurons.

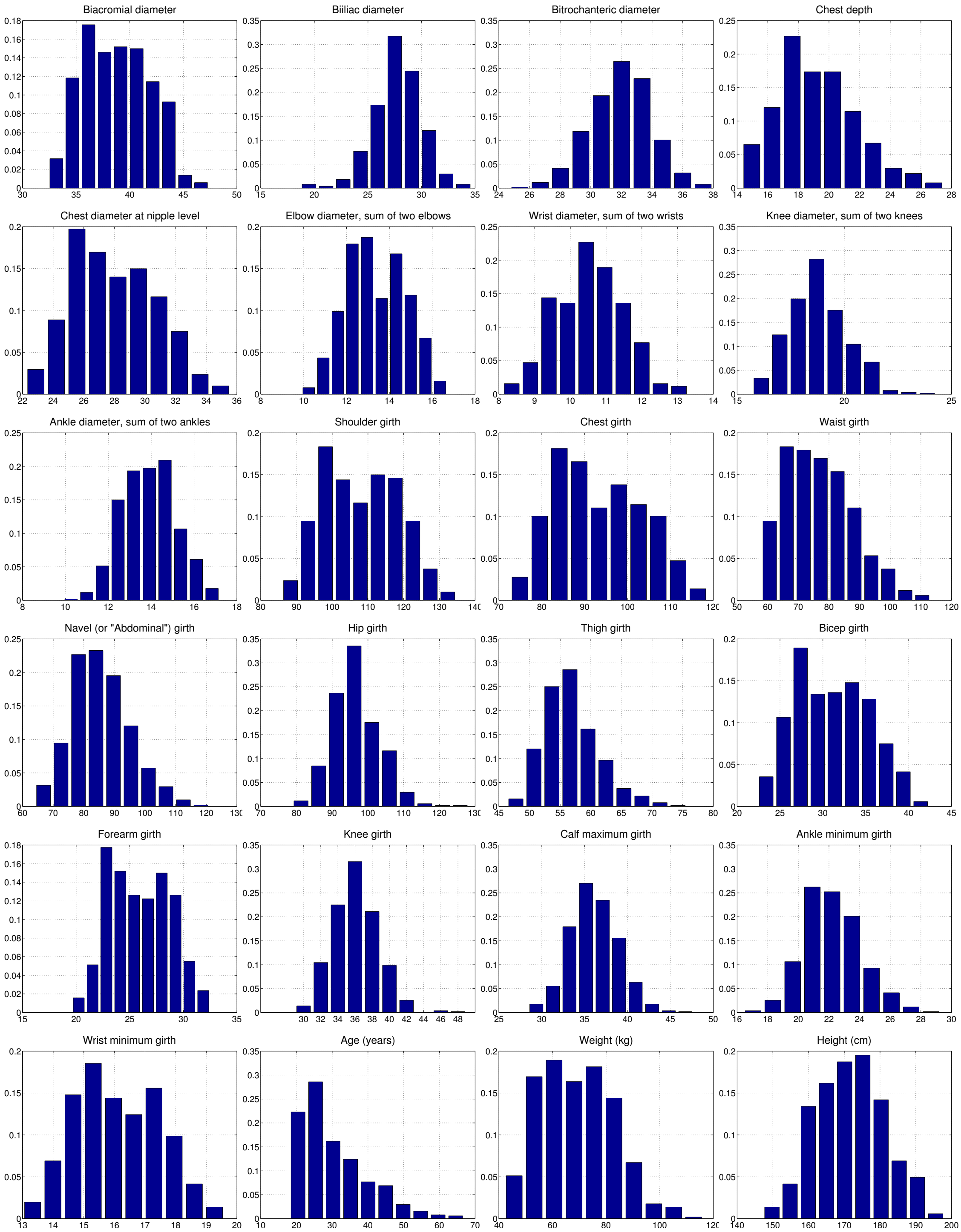


Figure 3.12: The figures depict the normalized distributions for each of the features belonging to the complete anthropometric dataset. We avoid showing the gender distribution as the classes are evenly distributed across the two classes (1-Male, 0-Female). *Y axis varies across the figures.*

3.5 Conclusion

In this Chapter we presented an overview on the techniques that we used within our research. State of the art techniques, like the ones presented for height estimation, were explored together with techniques that we devised, like anthropometric measures extraction and weight estimation from body measures.

Especially for this last technique a multiple regression analysis confirmed the suitability of these features to estimate the weight, when linearly combined. The proposed model was tested in various conditions, and its performance are validated testing the loss of precision in case the data are biased by a random noise. We demonstrated that the theoretical responses of our system outperform a human-based estimation experiment, and that the addition of noise does not invalidate our method; additionally a real case scenario was tested to validate our outcomes.

A similar technique was finally used to estimate the gender of people thanks to a subset of the anthropometric measures used in the weight estimation experiment. To fully validate the idea of gender estimation through body measurements we employ a richer database and we show that very good (theoretical) performance are achievable.

3.5.1 Critical analysis of the contributions

We have seen that Soft Biometric traits span a large amount of different aspects of the human body. It is clear that not all of them could be acquired at the same time or could exploit the same acquisition conditions, as they usually depend on the application scenario. Yet, the *human compliant* characteristic, that we previously stressed, underlines how the soft biometrics acquisition is in general *unobtrusive* for the user compared to the one of *hard* biometrics, where user's consent is generally crucial for the acquisition of the biometric template.

While for certain soft biometrics traits is easier to obtain an unobtrusive acquisition system (e.g. color of the clothes or height estimation), for some other traits this is still not completely possible. Considering this aspect the use of new techniques to acquire some of the soft biometric traits (e.g. estimating weight via images and videos) proposes new challenges that we started to tackle.
On the other side, the recent advancements in the technology of sensors and techniques, would increase the capabilities of extracting such features (e.g. millimeter body scanners for anthropometric measures, see fig. 3.13 for an example).

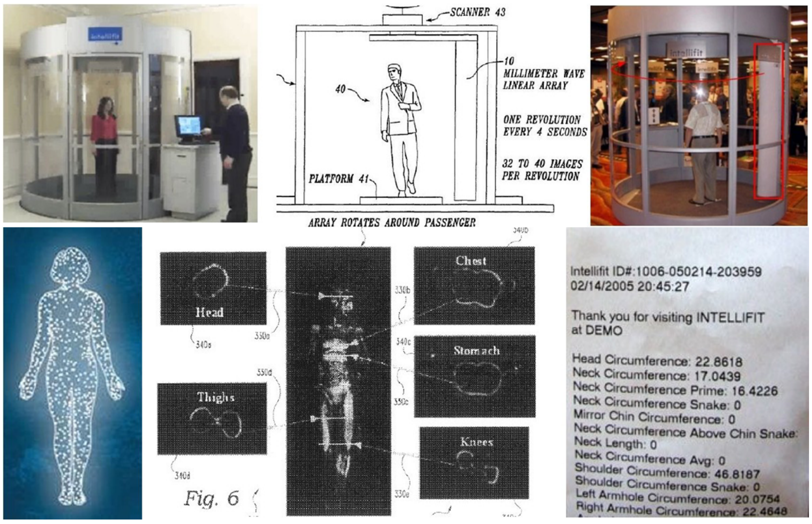


Figure 3.13: A rotating millimeter wave scanner that is capable of extracting measures of body parts and circumferences.

An interesting continuation of such work may rely on the implementation of an algorithm for automatic landmarks annotation. State of the art techniques are various and goes from the use of anthropometric models inferred from videos [91], to the use of 2D images [92]. The aforementioned systems are already able to provide the necessary accuracy to perform our body weight analysis.

Estimate the 3D appearance of the person from monocular images is also an interesting approach that is presented in [93]. Although the precision of the system seems to be promising, the authors experiments only using one subject and only using three poses. Unfortunately, as the authors clearly state in their article, the SCAPE eigen-shape representation does not provide any direct control parameters corresponding to intuitive attributes like gender, height or weight that can be specified by a user. Moreover, a similar experiment was performed in [94] but also in this case, although the experiments about creation of the 3D modelling system are convincing, the experiments about weight estimation take into consideration a very limited set of pictures (only 2 subjects).

CHAPTER 4

Pruning search in databases

In this chapter we investigate body soft biometrics capabilities to perform pruning of a hard biometrics database improving both retrieval speed and accuracy. Our pre-classification step based on anthropometric measures is elaborated on a medical dataset for large scale analysis to guarantee statistical meaning of the results. Our assumptions are verified by testing our system on a chimera dataset and tested in conjunction with an algorithm for face recognition. We clearly identify the trade off among pruning, accuracy, and mensuration error of an anthropometric based system. Even in the worst case of $\pm 10\%$ biased anthropometric measures, our approach improves the recognition accuracy guaranteeing that only half database has to be considered.

4.1 Introduction

As seen in Chapter 2, anthropometric measures fulfill the definition of soft biometrics [11]: they do not provide a specific pattern for identification (e.g. fingerprint and iris), they are human-compliant, and they can be extracted without cooperation from the user. Anthropometry is the science that uses human body measurements to study human variation and differences. This research field is particularly useful in case of medicine, industrial design, fashion, and other areas.

In this chapter we explore the use of anthropometric measurements to create a pre-classification scheme that is able to prune the search space of a subsequent hard biometrics (face recognition) module. We demonstrate that a hard biometric based face recognition system, properly pre-processed by our soft biometric anthropometry features, improves its performance in both accuracy and recognition speed. By using our anthropometric signature for pruning a face database, we exploit the independent complementary information provided by the body soft biometrics that makes this accuracy and speed gain possible.

However, the anthropometric mensuration process is inaccurate due to several factors. Sources of error are mainly due to the sensing device (tape meter, 3D scanner, ...), to the human operator, and to the inner variability of human measures [95]. Even if the sensing procedure could be eased in a near future by a new generation of specialized devices like body scanners or 3D video sensors like Microsoft Kinect, errors will eventually affect the recognition performance of an anthropometric system. Given this observation, we study how an increasing error during the mensuration step can affect the retrieval process and we show that even in presence of strong error magnitude (10% of the real measure) anthropometric measures could still be used for pruning the search space improving the performance of other biometrics (like face).

To evaluate our system and to verify that the results are statistically meaningful, we perform our analysis on the dataset NHANES presented in Chapter 3. The dataset includes more than 27000 subjects with the corresponding anthropometric measurements, and to the best of our knowledge this is the first time an anthropometry-based system includes this amount of users. Furthermore, in order to verify the performance of a complete system, we exploit the results of the pruning with a hard biometrics database. Therefore we check how the performance varies accordingly when we perform pruning with anthropometric measures and identification with an algorithm for face recognition.

The Chapter is structured as follows: we review works on anthropometric systems and search in biometric databases in Section 4.2. In Section 4.3 we present the case study and the methodology to couple the pruning system with an algorithm for face recognition. Finally in Section 4.4 we analyze the results obtained by our statistical analysis, and we show the performance gain achievable.

4.2 Previous work

4.2.1 Anthropometry-based recognition

After the historical example of the “Bertillonage” system, used in the 1882 for profiling and identifying prisoners (see figure B.1), the first example of anthropometric study that involves identification of people is the one proposed by Daniels in [96]. This study presents the quest for the “Average man”, i.e. a person with all the analyzed characteristics falling into average values (up to a given accuracy range). The author studies the possibility that such an individual could exist. The experiment involves a database of 4064 men of the Air Force flying personnel from which 131 measurements are extracted. Through an elimination process, the study shows how it is impossible that a person belongs to average classes in all his/her measures.

Later on, the study of Daniels was exploited by [97] to implement a people recognition system based on multiple biometrics. The article presents the result on fusing an anthropometric signature with the output of a system for gait analysis. They reach 90% of accuracy in a database consisting of 48 individuals.

To the best of our knowledge, the work presented by Ober et al.in [23] is the last one that explicitly involves people recognition based on anthropometric measures. The authors exploit the CAESAR 3D dataset that contains the 3D scans and 2D measures of approximately 4400 individuals (divided by classes of gender, age, and weight). Ober et al. analysis involves 27 different anthropometric measures. Their approach is based on the dimensionality reduction of these 27 measures through Linear Discriminant Analysis (LDA) analysis. They reach 97% accuracy in a subset composed of 2000 subjects.

Another interesting work [98] uses the 3D scans of the CAESAR dataset to analyze the effect of the error over the measures. They study the possible performance of an anthropometric recognition system as the reliability of the capture degrades. The original measures are altered adding error to the 3D landmarks to recreate different capturing conditions. Afterwards, they analyze the performance loss when looking for the altered data in the dataset.

All these methods start from the assumption that such a quantity of measures (27 in the case of [23]) should be easily available to the system in order to perform the identification. However, for some applications it would be challenging to obtain some of them (e.g. the foot breadth). Moreover, we believe that a real case scenario would be affected by errors similar to what studied by [98]. Thus making seriously challenging to create a system that could easily perform identification with noised input data.

4.2.2 Soft-biometrics based database pruning

A discussion on the possibility of using pruning via soft biometrics is already present in [99] where the author analyze the future possibility of pruning big public biometric databases using characteristics of the biometric templates. A later formalization of the pruning idea is already present in [11]. In [76] by Samangooei et al. semantic information are coupled with a gait signature to retrieve corresponding people from a database of individuals. Our case study differs from all these works as we do not consider the quantization step (*binning*) applied to each single feature. As clearly explained in the work of [100], although binning helps to improve pruning capabilities, it introduces another source of error too. At the time of pruning, if we miss the bin in which the user record exists, then a false reject error is generated irrespective of the matching accuracy, thus deteriorating the performance of the biometric system. Since in our case we consider the feature vector as composed of continuous values this statement does not generally apply to our system. Moreover, as the error span across different measures the effects of the overall error decrease.

4.3 Proposed case study

To better model a big population with real data we resorted to the already mentioned NHANES dataset. This dataset fits our need in that it contains anthropometric measures and its cardinality is very large. Here we recall some of the characteristics of the dataset reminding the measures that compose the entries for each subject. The measures are the following:

1. height
2. arm circumference
3. arm length
4. waist circumference
5. leg circumference

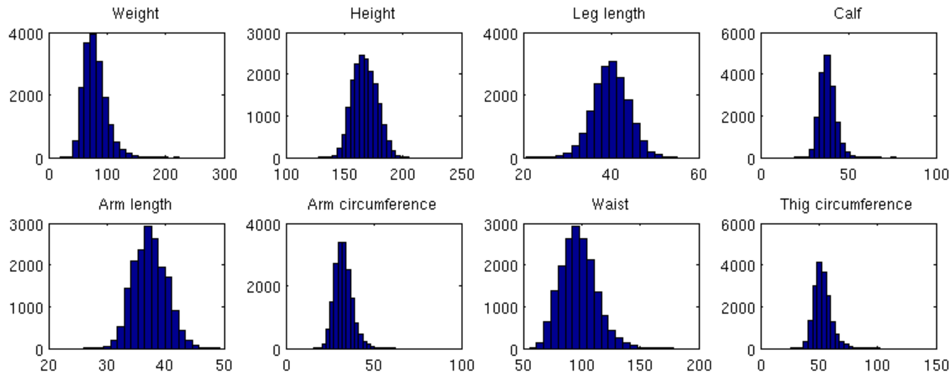


Figure 4.1: Image shows the histograms of values that characterize the measures included in the NHANES dataset.

6. leg length
7. calf circumference
8. body weight.

In figure 4.1 we show the distribution of values that belong to each class spanning the entire dataset.

Like all the databases of large dimension, NHANES dataset suffers of a considerable amount of missing data. Since in our case we cannot afford to have missing data in any of the previously mentioned measures, we discarded each subject that presented one of more missing measures. We discarded as well the subjects below 20 years old, as they had prevalence of missing data and became less represented in the dataset.

Afterward, we store the measures available in the dataset in a feature vector to be used in our pre-classification scheme. Prior to the normal identification module, we use those features to pre-classify the subject to be identified. In our case all the features are used as continuous value, and each of the features are considered of the same importance and then of the same discriminative power. One could argue that in order to select the features with most discriminative power, a study similar to [23, 79] should be conducted. In the first case LDA is used over a training set to extract the maximum information while discarding non useful information. In the second case, a brute force search is performed to check which feature combination performs best in estimating weight of a person. Nevertheless, to perform an analysis similar to the one of [23] we should consider a plethora of measures which

(a) (b) (c) (d)

Figure 4.2: The images show four face samples from the FERET *fa* gallery the images are warped so that eyes, nose, and mouth features are aligned across different samples.

is out of the scope of this study as we aim at easily define the physical shape of each user.

The pruning capabilities offered by our body soft biometric signature are used to prune a hard biometric database. We use face recognition against a complete face dataset to show the performance gain provided by our pruning scheme. In this case, the results are obtained using the well known FERET biometric face dataset that consists of 1195 subjects. We thus select a subset of people from the NHANES dataset so as to build a "chimera" database where each identity (face) is associated with a randomly chosen anthropometric feature vector [11].

A non verifiable cross correlation may affects face appearance and body dimensions, e.g. a large face would not likely belong to a slim person. As preventative measures, we normalized the face images and we iteratively randomized the choice of individuals from NHANES. The first action consists in warping all the images so that eyes, mouth, and nose positions will always fall at the same fixed locations (see figure 4.2); the second procedure consists in repeating several times the same experiment using always different subjects and then averaging the results.

Exploiting the code provided by [101] we create an algorithm for baseline recognition that exploits their eigenfaces implementation. Another famous face dataset (AR Face Database) is employed to create the eigenspace that will be used to recover the identity of the test subjects. The recognition results are compared as cumulative matching characteristic (CMC) curves. We demonstrate how the pruning can be used not only to speed up the hard biometric algorithm, but how it can be exploited to identify the best performing parameters that allow an increase of performance without any loss of accuracy.

4.4 Experimental results

Our experimental results can be divided in two sections. In the first part we analyze the pruning capabilities of an anthropometric based system as the mensuration error increases. Similarly to [79, 98] to each anthropometric measure is added a noise of increasing intensity to simulate a real mensuration system. The results discuss the amount of pruning acceptable as the mensuration error progressively increases. We will see that a trade off is possible so as to not interfere with the hard biometric recognition performance. The second part of this section is devoted to the use of the pruning capabilities in a full system where the identification is performed using an algorithm for face recognition. This algorithm is built and considered as baseline. This second part is also composed of two different analysis: firstly we discuss the achievable performance increase in accuracy that our anthropometric based pruning permit, and secondly, the trade off between accuracy, mensuration error, and penetration rate is shown.

4.4.1 Performance analysis of the anthropometry system

As biometric systems become more reliable, they can be applied to monitor or identify a bigger population. An example is the challenging Aadhaar Indian project [102] that aims at identifying the entire Indian population. Similar large scale systems will soon face problems due to the considerably high number of users to control. Exploiting the NHANES dataset we conduct our simulation over the entire dataset population, so as to guarantee statistical significance of the results. This provide us more than 17500 subjects from an original population size of almost 27000 individuals.

The literature of search and pruning of biometric databases proposes the *penetration rate* as the measure to compare the performance of a pruning algorithm. It consists on the fraction of database we are able to isolate to perform the identification. The higher is the penetration rate via the analysis of our feature vector, the larger is the portion of database considered in the identification rate [1]. However, the pre-classification step (binning) can also be affected by an error. The binning error can impair the performance of the recognition algorithm performed afterwards.

[1] A penetration rate of 1 means that the full database is analyzed (i.e. no pruning is done), while a value of 0.5 means that half database is considered.

Measure	Mean	Std	Min	Max
Weight	7.9	1.9	2.5	21.8
Height	16.7	1.0	13.0	20.3
Leg length	3.9	0.4	2.2	5.4
Calf	3.8	0.4	2.1	7.5
Arm Length	3.7	0.3	2.6	4.8
Arm circ	3.2	0.5	1.7	6.1
Waist	9.7	1.5	5.9	17.5
Leg circ	5.2	0.7	2.7	10.0

Table 4.1: We summarize the statistics of 10% error magnitude. The units are expressed in kilograms for body weight, and centimeters for all the other measures.

In our case we consider our feature vector as the ensemble of the anthropometric measures. To rank our database we employ a euclidean distance metric as suggested in [98]. To simulate a real mensuration system we consider a varying error, and in order to increase the randomness of this bias, an approach similar to the one proposed in Chapter 3 is used, where the error is applied varying randomly the error for each measure.

Considering our feature vector $F = [f_1, \dots, f_n]$ we add to each measure separately an error proportional to the magnitude of the original feature and with random sign, so as to obtain its biased version F_ϵ:

$$F_\epsilon \quad = \quad [f_1(1 + \alpha_1 w_1), \dots, f_n(1 + \alpha_n w_n)] \tag{4.1}$$

where α_i is a binary random variable that takes values in the set $\{-1, +1\}$, and the error is computed as function of the original value (f_i) times the power assigned to the noise at a given iteration (w_i).

To compare our error to the one observed in [23, 98] we summarize the error statistics in Table 4.1. Considering the average error, one can see that we are close to what one could expect from the worst case scenario.

For each given error magnitude we iterate over the entire dataset, at each iteration every biased vector is compared against the original dataset and a distance matrix is built. From the distance matrix we are able to compute a cumulative matching characteristic curve (CMC) that summarizes the performance of the pruning. The curve indicates the probability of observing the client in the first N-best candidates.

At the end of our analysis we obtain a CMC curve for each error magnitude

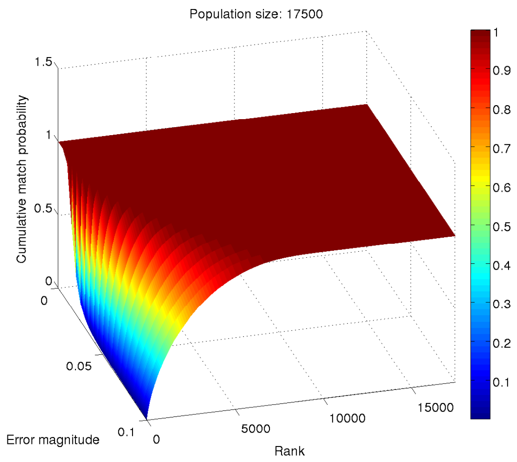


Figure 4.3: The plot shows the CMC curves as a 3D surface.

considered. We can hence plot the result as 3D plot (figure 4.3) in order to show the trend of the graph; or in 2D (figure 4.4) to better compare different experiments.

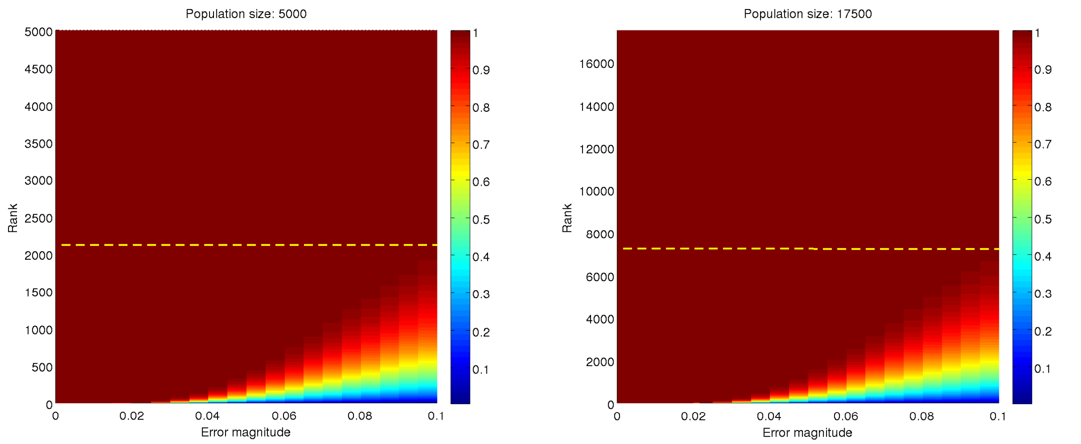
In figure 4.3 we can clearly see that if the error magnitude increases, we have to consider bigger portions of the results to guarantee that the client is among the first results.

We performed our analysis over different population sizes drawn randomly from the original dataset. For the sake of brevity, in figure 4.4 we show the experiments conducted with 5000 and 17500 subjects considered, respectively 30% and 100% of our dataset. In the first case the experiment was conducted 100 times and the results were averaged.

If we consider the maximum values admitted for the error magnitude (10% of the original value) we clearly see that a penetration rate of 50% can be obtained with no effort. In the first case (4.4.a) the client is always (probability $= 1$) within the first 2000 results. We remark that this penetration rate corresponds to 100% accuracy of the system, thus it cannot reduce the performance of the recognition algorithm applied afterwards. The second case (4.4.b), that considers the full original dataset, confirm our results since 50% of the dataset (8000 subjects in this case) is still a valid choice to have 100% accuracy of the pruning system.

4.4.2 Recognition accuracy increase by pruning

As previously mentioned, the performance of a recognition algorithm can be affected by the error introduced by the pre-classification algorithm or, contrarily, the



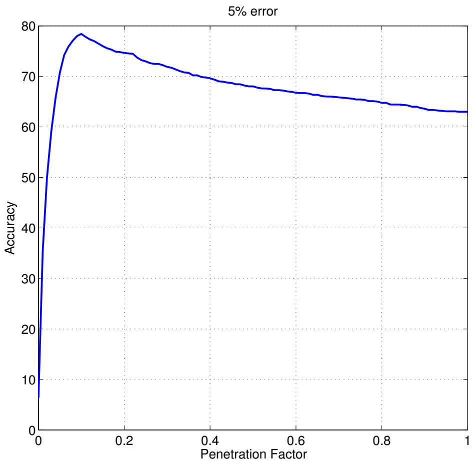
(a) (b)

Figure 4.4: The plot shows the CMC curves for two different population sizes: (a) 5000, (b) 18000 subjects. The intensity of the color indicates the probability of the CMC curve.

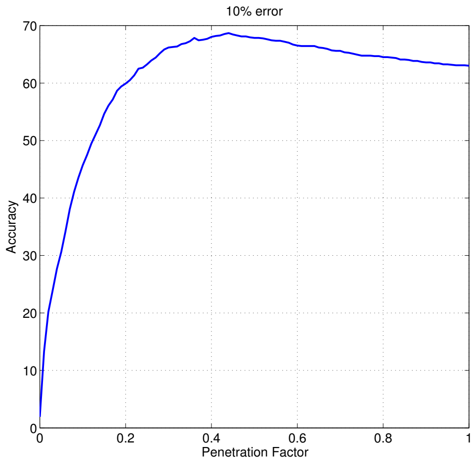
recognition algorithm can benefit from the pruning both in speed and accuracy. The analysis we conducted in the previous part, shows that a conservative choice could be the selection of just half database, which preserves the performance of the recognition algorithm even in the worst case (that we consider being the 10% magnitude error).

To verify our assumption we analyze in this section the performance of the cascade composed by our pre-classification scheme based on anthropometric measures, and a baseline system for face recognition. The algorithm produces at rank 1 a result of 63% accuracy. To analyze how the pruning affects this accuracy, we decided to consider two error magnitudes: 5% and 10%. Although a conservative choice could be made by selecting the penetration rate that provides us with the certainty of finding the client within the pruned results, for the sake of completeness we tested our full system for increasing values of penetration rate. This analysis shows us the best trade-off between pruning and accuracy and leads us to the best performing system. At each iteration we used the anthropometric feature vector associated with the identity to perform a fast search in the anthropometric dataset. After all the entries are ranked according to our pruning metric, the face recognition is performed on the subset defined by the penetration rate.

Figure 4.5 summarizes the results obtained by such analysis. Our assumptions are confirmed: as we consider the lowest penetration rate, the final recognition ac-



(a)

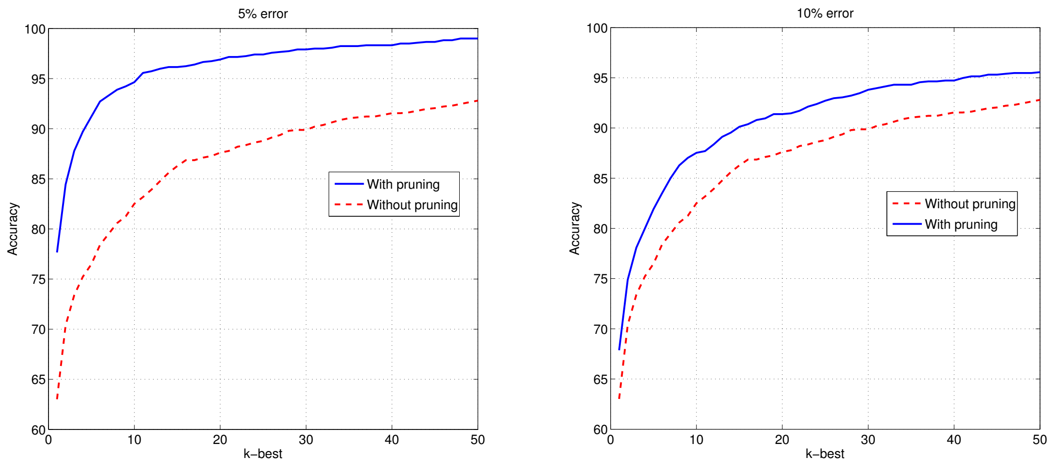


(b)

Figure 4.5: Rank one accuracy variation as the Penetration rate increase. Error at 5% (a) and 10% (b) magnitude are considered.

curacy suffers from the poor performances of the pruning algorithm, that is not able to provide the client with a given certainty (e.g. in the range $[0, 0.1]$ of figure 4.5.a). In both cases the curves behavior is very similar; the performance at rank 1 accuracy increases up to a global maximum, and then, asymptotically, falls back to the baseline accuracy result (63%) as the portion of pruned identities considered gets larger. Indeed, as we consider more subjects in the recognition process, we analyze more images that could include some face samples that increase the false acceptance rate. The maximum corresponds to the operating point that best put together the benefit of the pruning and the recognition algorithm. By selecting the two maxima we can define the optimum operating points for the two systems (see figure 4.5) that have to operate with 5% and 10% error magnitude. In the first system the maximum indicates the best penetration rate of 0.1, then 10% of the pruned results; while in the second case the range falls between 0.4 and 0.5.

Furthermore, to complete our analysis we show in figure 4.6 the full rank curves of the two systems. In the first case (figure 4.6.a) the chosen penetration rate consider the first 100-best while in the second case (figure 4.6.b) the top 600 results obtained by the pruning algorithm are further analyzed. The gain is both high in the sense of accuracy and speed performance (since the penetration rate is smaller) in the former, while in the latter both the gains are reduced, but still our system performs better than the baseline algorithm and we are able to prune half database



(a) (b)

Figure 4.6: The accuracy increase achievable after setting the penetration factor at its best in both (a) 5% and (b) 10% noise magnitude.

out from our recognition, speeding up the recognition phase by a factor of $2\times$.

In case of a recognition algorithm more expensive in terms of resources, one could be interested in reducing at the minimum the search space size. Therefore, a measure of the loss of performance caused by selecting a smaller penetration rate is needed.

Then, in order to analyze the full response of the system to both the penetration rate and the error in the anthropometry mensuration step, we present the result of figure 4.7. Here we show the performance of the global system in terms of rank-1 results as both the penetration factor and the error magnitude vary. One could exploit such graph to understand whether to leverage on the first parameter to obtain a faster system, or to invest into a better mensuration system to lower down the second factor, thus approaching the best possible results. In our case we can clearly see that by choosing small penetration rate values we can guarantee good performance only if we are able to lower down the error of the mensuration system. If we are not aware of the amount of error generated by our mensuration system, a good choice could be a penetration rate of 20% (i.e. the algorithm for face recognition would check only 1 image out of 5). That choice guarantees us very good performance in case of 5% error magnitude, and close to the baseline rank-1 accuracy in case of 10% error.

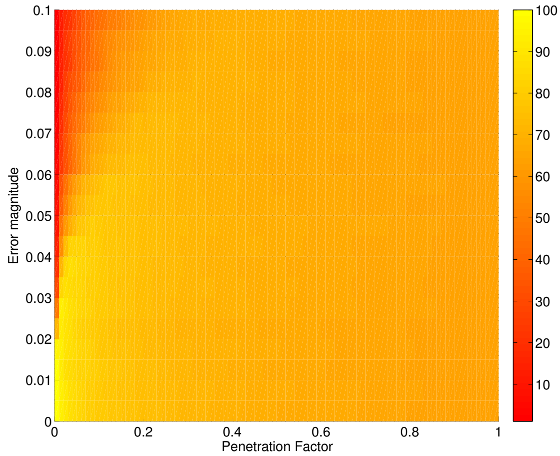


Figure 4.7: The graph shows the accuracy achievable by the cascade of our pruning module followed by the face recognition one. The plot helps identifying the best operational point. One can see that lowest penetration factors are effective only if mensuration error is reduced accordingly. By knowing the possible maximum error of our mensuration scheme, it is easy to understand which penetration factor to use in order to maximize the performance of our system.

4.5 Conclusion

We presented a work that exploits body soft biometric traits to pre-classify entries of a biometric database to improve both speed and accuracy of a hard biometric identification system. The anthropometric measures used as feature vector were extracted from a medical dataset for large scale analysis, thus guaranteeing statistical meaning of the results. Our work clearly identifies the trade off between penetration rate (i.e. the speed up achievable by pruning the dataset) and the accuracy of the identification performed by the system after pruning, and the response to the noise of the mensuration system. The behavior of a complete system is analyzed by using a well known face recognition technique and database in cascade to our pre-classification module. Our results distinctly show that even in the worst case of $\pm 10\%$ error magnitude in the anthropometric measures, the pruning system is able to speed up the search of $2\times$ factor while guaranteeing an increase of accuracy performance.

4.5.1 Critical analysis of the contributions

Chimera dataset are a tool that encountered many criticisms in the Biometric community, in [103] the authors demonstrate in theory and practice that the use of chimera dataset does not always imply that the performance observed in the simulated database will always be similar to the one of a real dataset.

However, their approach mainly considers the performance of fusion schemes, while we presented here a pruning algorithm that really considers in a separate way the two feature set. Nevertheless, we agree that, when possible, the use of real multimodal database is better than using simulated chimera datasets. Following this idea one may want to exploit the aforementioned CAESAR dataset [104] that contains 3D scans of the full body of more than 4000 subjects (see fig. 4.8 for an example). In this case the anthropometric measures might be extracted by the 3D body shape and the face recognition might be performed on the 3D face model.

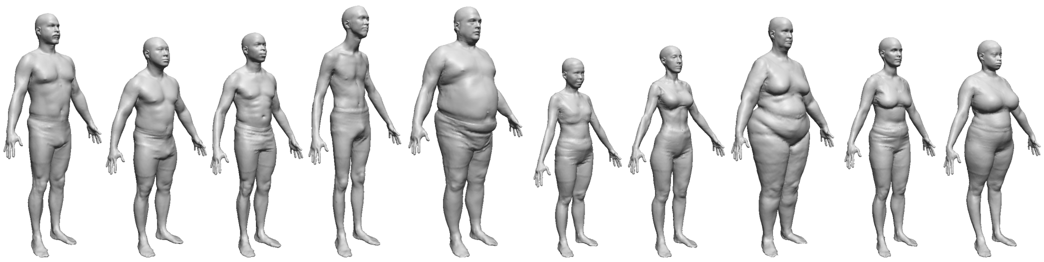


Figure 4.8: CAESAR commercial dataset that contains 3D full body scans of more than 4000 subjects.

CHAPTER 5

Identification

In the case of bag of soft biometrics [11], identification may also be possible for groups of varying size. In order to study the behavior of an identification system based on soft biometrics, we explore the binning of soft biometrics into different categories. We define three traits related to the body (weight, height, and clothes color) and we use them to re-identify subjects walking in a video surveillance scenario. Moreover, we propose a system for people tracking in video surveillance networks that preserves privacy by mapping people recorded by a video sensor network in the map of the corresponding surveilled areas. Thanks to this approach, it is possible (1) to synthesize multiple video outputs into a unique picture, and (2) to control privacy by offering the possibility to filter information to viewers. We demonstrate the feasibility of our approach by showing a real case application scenario.

5.1 Introduction

With the fast increase of the number of video sensors employed in surveillance applications, several crucial challenges must be solved. On the one side, problems usually arise due to the large amount of data that makes overwhelming and no longer feasible human supervision. On the other side, this quantity of data generates important concerns about the privacy of people under surveillance.

In the first case, in addition to tiredness and loss of focus of attention an important problem to overcome is the loss of the *situation awareness* of the surveillance agent. That is to say, the confusion generated from monitoring an activity presented as two different and non-correlated views, especially if the areas to supervise are unknown [105]. In the second case, we would like to enforce the privacy of surveilled users without waiving, at the same time, the possibility of tracking a person as he/she moves across the surveilled areas.

Although the straightforward solution to the first challenge would be to increase the number of video supervisors, it is clear that this is not scalable as it requires large human effort. Usually, security operators need accurate planning of time schedules to maximize their focus of attention, and site inspections are necessary to acquaint the agents with on site knowledge [106], each of these causes translate into a considerable economic drawback.

To face those limitations, many approaches have been explored that try to automate and to ease the work of surveillance operators. In [105] a 3D environment is simulated where the images from a distributed camera network are projected on a reconstruction of surveilled areas. The reconstructed space aims at increasing scene understanding and people localization inside the scene. A similar approach was envisaged by the authors of [107] that added audio cues and time analysis for the events in a surveillance network. This make possible a more precise localization of subjects within the video sensor network. Another approach employed in large-scale systems enforces situation awareness by linking cameras with the maps of the building. The system described in [108] displays camera locations within a map so that a spatial reference is always available for each camera. Yet the agent has to make a mental effort to link images and camera locations on the map.

Another problematic aspect of video sensor networks are privacy concerns. Typical approaches rely on obfuscating the images coming from the cameras involved in the recording. According to the level of obfuscation a person may be partially (e.g. only the face) [109], or totally [110] hidden. While the partial obfuscation leaves clues useful to perform visual tracking (i.e. body shape and color information), it is clear that, even if coupled with automatic tracking, the total obfuscation creates further difficulties to the surveillance agents.

We seek to solve both issues by using a map that summarizes all the ongoing activities in a surveillance network (see figure 5.1). Additionally, it provides a human compliant description of the physical appearance that makes feasible

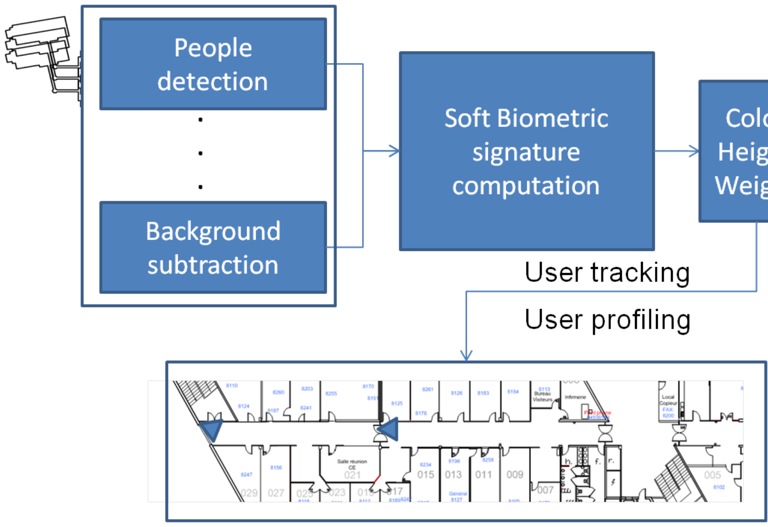


Figure 5.1: Scheme that summarizes the modules of our system. Images are analyzed in order to compute the soft biometric signature. Such signature is afterward used both to re-identify and to profile the surveilled person.

the re-identification task. The proposed framework provides an automatic solution for re-identifying people while respecting their privacy. We first establish a link between the real scene monitored by multiple video sensors with its map in blueprints. Secondly, we track and re-identify people moving around the building thanks to soft information on bodies. Finally, we display information concerning recorded people as a gradual function of existing authorizations and/or requirements.

We perform single camera tracking that enables us to extract soft biometrics traits [11] that describes physical attributes of the person. We later use those attributes to re-identify the person in another camera. We do not make use of temporal clues, neither our cameras are overlapping. Thus, to re-identify people, we only rely on the distance metric computed from soft biometric traits. Our design automatically deals with privacy issues: the representation of the user in the map is indeed completely anonymous. The tracking capability is maintained by the soft information provided as person's description.

In the remaining part of the Chapter, we introduce the mapping technique that creates the link between images and building plan in Section 5.2. Then, in Section 5.3 we present the soft biometric signature that is used to re-identify and to describe people. In the last Section, we provide examples of our application scenario and we draw our final considerations.

5.2 Linking video camera outputs and building map

The system we propose aims at generalizing and summarizing an entire video surveillance system into a global view that corresponds to the map of the covered areas (see 5.1).

Although the proposed approach best fits the scenario of indoor surveillance, it is generic enough to be applied to almost any building (e.g. undergrounds, airports). We require that a map of the area under surveillance is available (which is often the case for any building). Information from surveilled areas is exploited to create a correspondence between images and the building's map. Our method exploits the existing furnishings (e.g. doors), shared among different views as invariant features to create the mapping. Thanks to these elements, it is possible to inter-calibrate images from the camera with the map of the building.

5.2.1 Homography mapping

Obtaining the ground plane location of people in the scene means we should project the estimated positions for each frame on a single picture that represents the building map.

The invertible mapping from points in the image plane to points in the map plane is a planar projective transformation since it maps lines to lines (three collinear points in the image plan will still lay on the same line after projection and vice versa).

From an algebraic point of view, if we represent each two-coordinate point *(x, y)* as a homogeneous 3-vector ***p***=*(x, y, z)*, we can formulate a planar projective transformation as a linear mapping represented by a non-singular *3 x 3* matrix called homography matrix [111]:

$$[x'y'z']^t = H_{3\times3}\,[xyz]^t \tag{5.1}$$

Therefore, once we have the homography matrix we can compute the ground plane location of a person on the map as a projection of the corresponding location of his/her feet on the image.

We choose to solve the problem of estimating the 2D homography matrix by means of normalized Direct Linear Transformation (DLT) algorithm proposed by Hartley and Zisserman in [111]. The DLT algorithm exploits a set of corresponding points (***p'*** $\leftrightarrow$ ***p***) to compute an estimate of the homography mapping between two

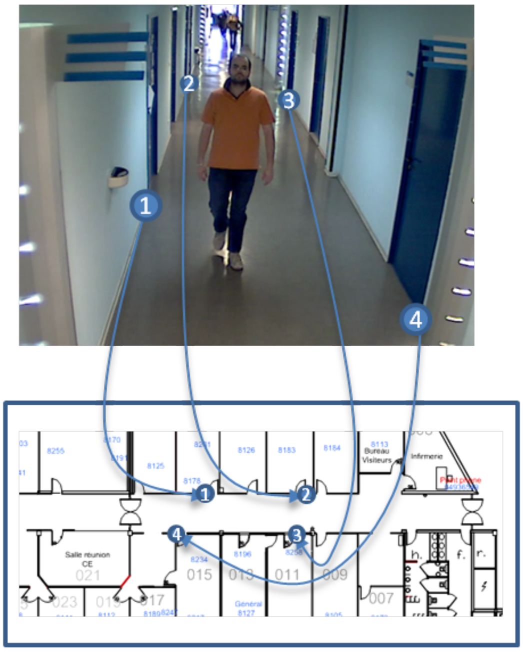


Figure 5.2: The image shows an example of point correspondences needed to compute the homography mapping which allows the transformation from the image space to the map plane. It is important to choose features that are shared between the two modalities.

planes. Since H is defined up to a scale factor, the transformation led by the 9 elements of H has 8 degrees of freedom. Since for each pair of corresponding points we have 2 linear equations in the H elements, it follows that if we have 4 corresponding points $\boldsymbol{p'} \leftrightarrow \boldsymbol{p}$ then it exists only one solution H (if no more than 2 collinear points exist).

In our case, point correspondences are established using standard elements present both in the map and in the images. For our system (see figure 5.2), we choose the position of the doors on the floor to compute the homography matrix H. This is just an example of how to exploit information shared among the recorded scenes and the map of the building.

5.2.2 Single camera tracking

A tracking algorithm keeps trace of the location of the surveilled subject on the map. We use this trace to compute the soft biometric signature. To achieve that we exploit state of the art algorithms for both detecting and tracking people. In order to detect persons we use the well-known histogram of gradients approach from [112].

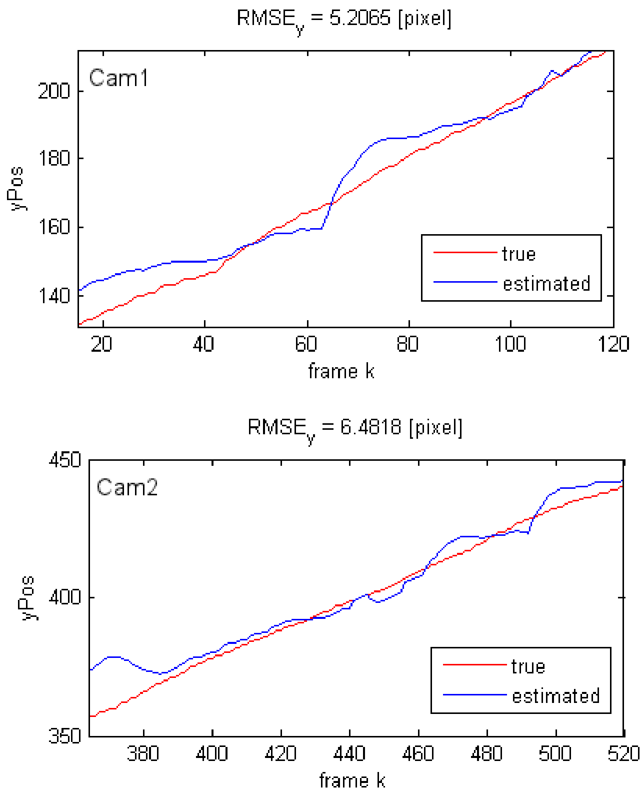


Figure 5.3: The two plots represent two examples of trajectories estimated from the ground plane location of the subject. A small value in *yPos* axis indicates that the subject is far away from the camera. The two tracks are compared to the ground truth (the straight line). It is noticeable how the uncertainty is generally bigger when the subject is far away from the camera.

The feet position is estimated as the middle point of the bounding box previously extracted; then it is projected onto the building map thanks to the homography mapping. Finally, we track the position of the user in the map domain by using a Kalman filter approach. The tracking performs better in this domain since the constant velocity hypothesis is more suitable for such a scenario. In figure 5.3 one can observe an example of tracked trajectories for two subjects walking in two different cameras.

5.3 Soft biometrics based people re-identification

Hard biometric traits are intrinsically good for the identification task thanks to the greater complexity of their patterns (e.g. fingerprint and iris patterns). However, hard biometric traits represent a great challenge in case of video surveillance sce-

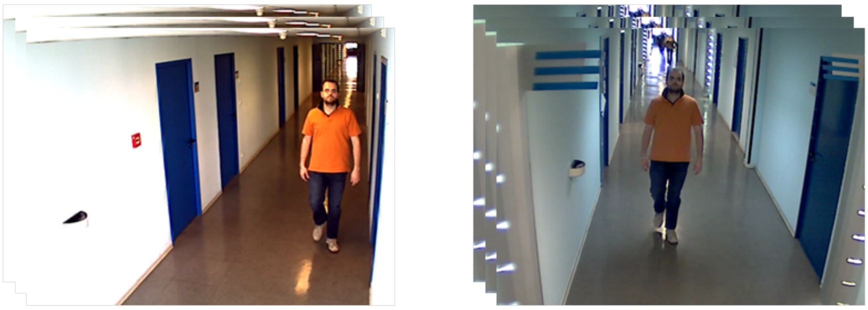


Figure 5.4: Images from the sequences taken by two different cameras.

narios. Their acquisition is indeed problematic through camera normally employed in such situations. For this reason, body soft biometric traits provide an additional source of information about the identity of subjects in cases where the quality of video cameras does not match the requirements of classical biometrics (see figure 5.4).

In this chapter we seek to provide a practical example of the possible uses of soft biometrics in a surveillance scenario. We propose the use of height, weight, and cloth colors to perform the quasi-immediate re-identification process when cameras do not overlap and when no temporal information is exploited.

In the following part, we explore the three different components of our soft biometric signature and we present our fusion scheme for the re-identification task.

5.3.1 Height

Stature is the first characteristic that determines the physical aspect of someone. Several methods exist to estimate the height of vertical objects in images the most known is presented in [16]. Some of these techniques were adapted to estimate people's height in videos or images [26]. Moreover, these algorithms can be divided in two classes depending on how they compute height. They can exploit information extracted from the captured scene [16], or they can use scene and camera knowledge, thus camera calibration [26].

Since the cameras employed in surveillance systems are generally not calibrated, in our system we exploited the former approach. In our case, the knowledge of the scene is limited to the size of corridors' doors. No limitations exist to the application of this technique: it is indeed sufficient to find within the scenes a common target of know height to provide stature estimates. An example of height values measured from the images of our surveillance system is show in 5.5.

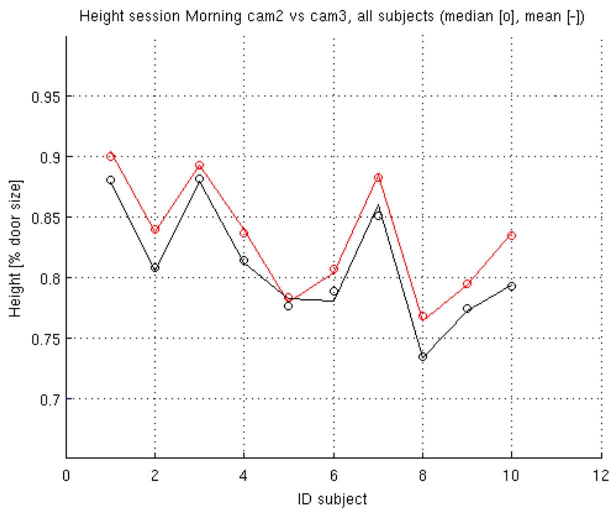


Figure 5.5: Values of height extracted from two different cameras. The values are reported as function of doors' height in the two different scenes. The graph shows the height estimation results for two different sessions.

In our case, doors' height does not change across Scene 1 and Scene 2. Therefore, we are able to compute the height of people as function of the doors' height (here used as measuring unit).

5.3.2 Weight

Weight is a recently exploited soft biometric trait. To the best of our knowledge, the only paper that involves weight for identification [34] uses a scale to weigh users of a system for fingerprint recognition so as to increase the recognition performance by including this trait in the classification scheme.

Unlike our experiment already presented in Section 3.3 where we analyzed the possibility of estimating people's weight from images by using anthropometric measures, in this case it is impossible to extract a set of anthropometric measures as our scenario involves uncalibrated cameras. Then, to provide a measure related to weight, we compute to the area occupied by the silhouette of the person with respect to the distance of the person from the camera. The extraction of the foreground is performed with the algorithm presented in [113].

The area's measure is computed as $WPR = \frac{W}{A}$ a ratio of foreground pixels (W) over the entire bounding box area (A) previously extracted by the people detector module.

Since this value have a tendency to vary according to the distance of the sub-

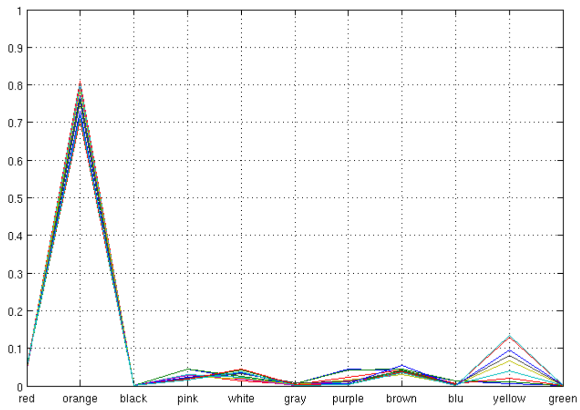


Figure 5.6: Considering pixel Lab components, the probabilistic color histogram measure the probability of the pixel to belong to a particular color class. In the figure the PCH of consecutive frames are shown. In this case the correlation among frames is evident.

ject from the camera we normalize it for each camera according to such distance (in the map plane) so as to make it a constant term. Moreover, by filtering the results obtained for several frames, we can guarantee enough independence of such result from the capturing camera.

5.3.3 Clothes color

Using colors to track people across a camera network can be challenging. As colored surfaces are exposed to different condition of illumination, the perceived color can vary substantially (see 5.4).

The probabilistic color histogram (PCH) presented in [17] is used to provide a color descriptor which is invariant to the illumination conditions. Images are converted into Lab color space and the pixel further elaborated using a fuzzy KNN into the 11 culture colors classes (*red, orange, black, pink, white, gray, purple, brown, blue, yellow, and green*).

Consequently, for each given pixel we obtain its soft classification score for each given color out of the 11. The descriptors are in the form of a PCH that represents the probability for a given pixel to be represented by one of the 11 classes previously mentioned.

An example of a probabilistic color histogram is shown in figure 5.6.

5.3.4 Bag of body soft biometrics for re-identification

To extract information that describes each user (height, weight, and color) a foreground and background segmentation algorithm is necessary. We employ the commonly available system provided by [113]. The silhouette obtained, is used both as weight correlated measure, and to separate the silhouette color from the background to compute the color descriptors.

After the extraction of the features, an approach similar to the one presented in [11] is exploited. In our case the bag of soft biometrics is limited to characteristics of physical body appearance.

Then, the results are fused together using a weighted sum fusion scheme:

$$D_{\mathit{total}} = \alpha D_{\mathit{color}} + \beta D_{\mathit{height}} + \gamma D_{\mathit{weight}} \tag{5.2}$$

where each *D* is a euclidean distance measure of the corresponding feature and α, β, and γ are weighting parameters chosen to provide more importance to the soft biometric trait which provides higher entropy than the others. The values chosen are 0.6, 0.2, and 0.2 respectively so that color (which is more discriminating) will weigh more.

To validate our idea we exploit a set up similar to the one presented in [114]. Then, we analyze recorded video sequences of ten subjects walking in a corridor environment. Challenges are presented by the varying illumination conditions and the compression applied to the videos (at the source). The videos recorded are analyzed with the processing steps already presented. Results of such fusion lead to the confusion matrix in figure 5.7. All the subjects are correctly re-identified as moving from one camera to the other.

5.4 Displaying appropriate information

Video surveillance is a technology that provides many benefits and, at the same time, generates many concerns because users' privacy is generally involved. For this reason, in the past years, a common solution was to record videos without employing surveillance agents. This guarantees that, if needed, the videos can still be available to law enforcement agencies. In other cases, the monitoring of qualified agents is required, thus guaranteeing their prompt intervention in case of emergency. Problems may arise as the number of camera grows because the number of required agents increases as well [106].

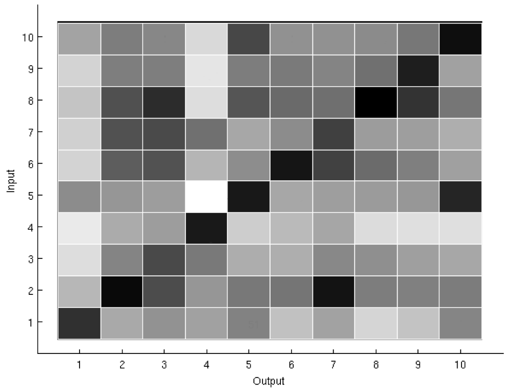


Figure 5.7: Confusion matrix representing the re-identification output of our system (darker color means smaller distance). The Input being the signature computed at Camera1, and Output being the outcome search in Camera2. The minimum distance subject along the rows gives the corresponding person among the two cameras.

The two presented are objective limitations to the use of video surveillance systems. In the first one, the limit is the possibility of having only passive surveillance, abandoning the possibility of a prompt security intervention. In the second case, the limitation lies in the large number of surveillance agents needed to cope with the number of surveilled cameras.

In order to tackle both these problems we propose a system that is able to summarize a surveillance network by providing to the agent the digest of the scenes recorded. An example of such system can be observed in figure 5.8.

Each person in the video surveillance system is represented as a two-color square. The colors of the square represent the upper and lower part of the clothes (i.e. torso/legs) obtained from the soft biometric signature computation. Although the specific design of the PCH descriptor provide us with slight intrinsic robustness to illumination variations, we compute the final color components' used in the user interface as follows:

$$C_{torso,legs} = \textit{NPCH}_{torso,legs} C_{matrix} \tag{5.3}$$

where *NPCH* is the normalized probabilistic color histogram (computed as $\frac{PCH}{||PCH||}$) and C_{matrix} is the color components matrix for the 11 culture colors chosen in our setup. Each of the RGB components of the culture colors where chosen from the values established by the color naming experiments performed by [115].

Our schema does not provide to the security agent images of the live record-

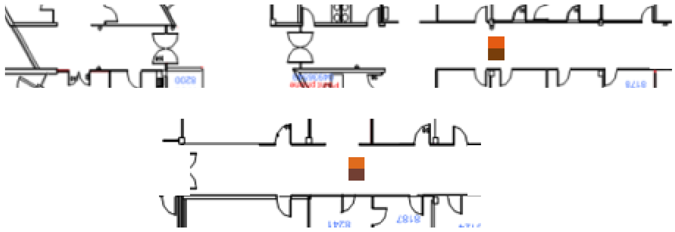


Figure 5.8: Example of the output of our system. It shows portion of the same corridor recorded by two cameras in two different moments and under different light conditions. The color of the top of the square matches the color of the torso, while the bottom part matches the color of the legs.

ings; this allows tracking people while ensuring their privacy. The description provided by our human compliant signature can be used to visually tracking a person, too. Our presentation framework do not alter in any manner the images and for this reason videos can still be recorded so that law enforcement agents could access them if needed.

5.5 Conclusion

We presented a new approach to video surveillance that aims at increasing situation awareness of surveillance operators while keeping an eye on both the re-identification and the privacy of the surveilled users.

Our system combines an algorithm for single camera tracking, a re-identification module based on soft-biometric signature, and a visualization interface that preserves privacy of users. We exploit geometrical correspondences between the scene and the map to compute a homography that links people with their current location in the scene. Small colored rectangles depict the users inside the map of the building. We exploit body soft biometrics (height, weight, and color) to create a signature that allows re-identification of people as they move through the camera network. Moreover, we use the soft biometric signature to enforce the situation awareness of the surveillance agent by creating a correspondence between the colors of persons' representations as they move in the map. We are able to summarize the information from the scene; it provides a good solution in terms of privacy and security management of buildings.

In order to provide an additional level of privacy preservation a similar system could provide ways of filtering displayed information to the user according to a given level of rights. In this way, users with higher permissions can easily access to the entire content of the surveillance system, while low ranks allow only the supervision of the essential information provided by the system. Another possible future upgrade, would link additional soft biometric traits (like hair color, gender, and ethnicity) with the shape of the person representation in order to provide additional clues for the visual tracking.

5.5.1 Critical analysis of the contributions

Our work confirms the results of our previous analysis presented in [116], where we provide an experimental formulation to compute group's size that minimize collision (or interference) probability in case of *identification* through *bag of soft biometrics*. In our case we analyzed the possibility that a smaller bag, composed uniquely of three soft biometric traits, might be used to perform re-identification in a non-cluttered small environment (i.e. a corridor).

The Probabilistic Color Histogram feature presented in [17] was tested against color histogram based techniques on the well known VIPER dataset. However, it would be interesting to provide results for more challenging database like the ones generally used in the video-surveillance [117] (i-LIDS, PETS, CAVIAR). However, this implicitly require the use of a bigger set of soft biometric traits as usually those scenes are cluttered and contains a lot of users.

Moreover, improvements in the entire system may be obtained combining the characteristics of state of the art re-identification systems (composed of higher dimensionality features set) and the privacy preserving characteristics of soft biometric traits.

CHAPTER 6

Semantic annotation and monitoring

In this chapter two different applications are going to be examined that exploit our capability of extracting soft biometrics information to perform analysis of contextually different problems related to weight.

In the first case an application is designed that allows people to self assess their body health conditions.

In the second case an application that monitors the weight of cosmonauts will be presented and compared with the applications used nowadays on board of the International Space Station (ISS).

6.1 Telemedicine

Extreme overweight is known as obesity and it is spreading widely like an epidemic. Notwithstanding the alarming situation, people do not respond properly to the warnings launched by the health institutions and the medical specialists feel the urge of preventive tools and methods that could increase people self-awareness about weight problems [32]. 3D body analysis can be employed to fill the gap of self diagnosis methods. We propose a system based on the Microsoft Kinect RGBD sensor to help people to detect weight problems and, possibly, to guide them through improvements in their lifestyle. Our application extrapolates anthropometric measures from the body silhouette and 3D information. The anthropometric measures are used to create a statistical model that estimates subjects weight using information from a medical dataset. Thanks to the knowledge acquired while analyzing the user, the system can provide, through its interface, healthiness measures like person's ideal weight, current BMI, estimated calories' intake and ideal one. All information that are necessary for a healthy lifestyle that would mitigate the effects of weight problems.

6.1.1 Introduction

Poor diet, lifestyle choices, and an under regulated food market are the main causes of the widespread obesity worldwide. Obesity is when a person is carrying too much body fat with respect to his/her height and sex. Generally, a person is considered obese if his/her body mass index (BMI) is 30 kg/m^2 or greater.

United States spends more than $\$300$ billions each year to treat obesity, diabetes, and cardiovascular diseases [118]. All those diseases, together with cancer, are the majority cause of mortality as can be seen in figure 6.1 extrapolated by the World Health Organization (WHO) website. Overweight is one of the main causes of this spending as it has been identified as one of the main factors that generates those diseases. United States are not alone since, according to the statistics of the WHO, many countries experience an increase of overweight people whose number has reached a size proper to epidemics. This is demonstrated by data of the WHO infobase website [1] that shows BMI trends that are irreversibly growing. For this reason, overweight has been identified as one of the main problems to face before arriving to a dead end.

Even worse, people develop the overweight status at younger age compared to the past, this situation provides them more time to develop a morbid condition. In figure 6.2 the data extrapolated from [119] shows the percentage of European children population (7–11 years) that falls between the overweight range. Since for these subjects the risk of becoming obese is very high, activities have been set up from nations and organizations to prevent this eventuality. For example, a vast number of European research projects and associations where lately created[2] that contribute to study the problem and to promote awareness campaigns.

Obesity can be treated by losing weight, which can be achieved through a healthy, calories-controlled diet, and increased exercise. However, the lifestyle

[1] https://apps.who.int/infobase/Indicators.aspx

[2] http://ec.europa.eu/research/health/medical-research/diabetes-and-obesity/index_en.html

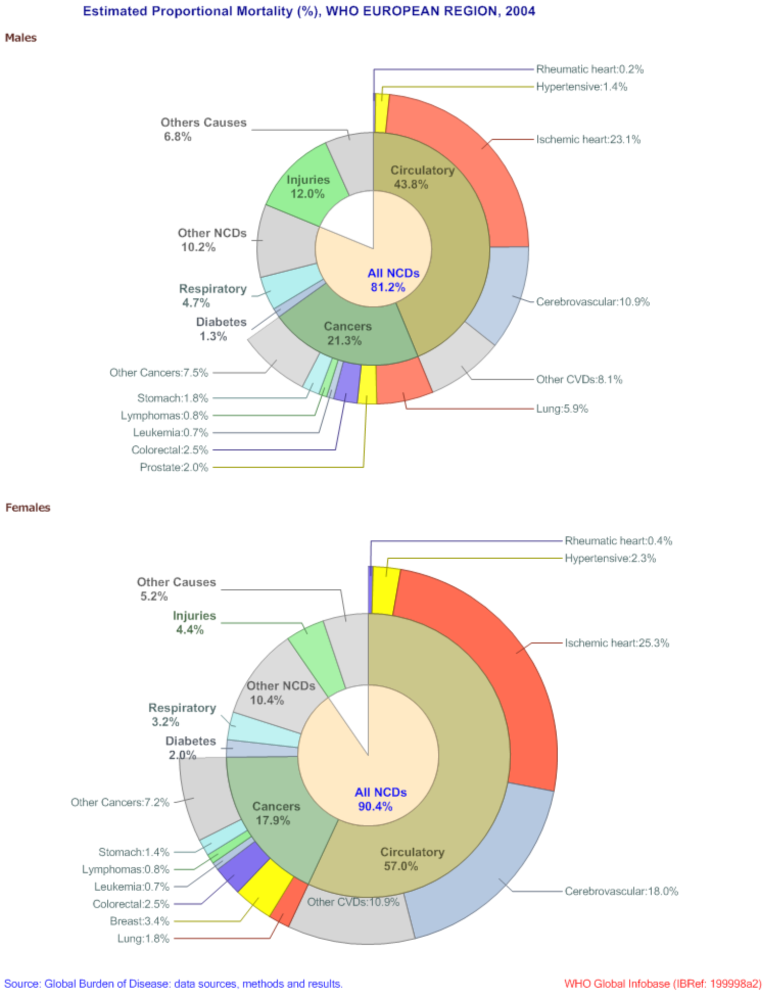


Figure 6.1: The two graphs show the rates of mortality factors for the European region. For both the two classes (male, female) the major contribution is given by Cardiovascular diseases. *[Image courtesy of WHO infobase]*.

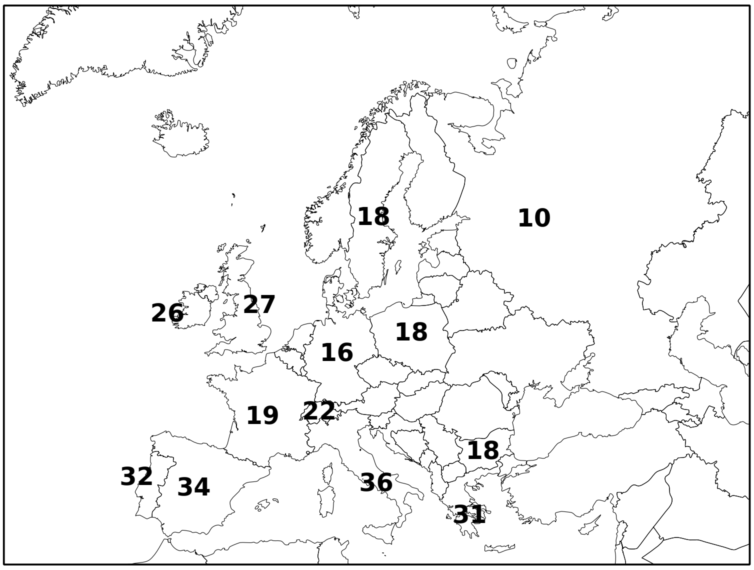


Figure 6.2: Prevalence of overweight children (aged 7–11 years) as a percentage in European countries.

changes necessary for weight loss can be challenging and not feasible to achieve without the proper support and control. For these reasons we need technologies that could help people to develop their self-awareness so as to achieve a better control of their body. Computer vision, by now entered in our daily life could be a favored mean for providing such new techniques. Algorithms like silhouette analysis [120], automatic extraction of the weight [79] and measures of the body [81], as well as the 3D model reconstruction of the human body [91], may be used as medical devices [82] or telemedicine equipments.

In this section we aim at complementing the information we are able to extract thanks to the Kinect camera and out weight estimation algorithm to provide to the user a better understanding and awareness of his/her own body condition. Through the use of computer vision techniques and resorting to commercial 3D equipments (e.g. the Microsoft Kinect sensor), we extract information about the user's body in order to estimate anthropometric measures, weight, and successively BMI and ideal/healthy weight. The result is an interface that shows to the subject its current position in a scale of *healthiness*. Additionally, considering the measures of the body, the system provides clues about the correct and healthy energy intake.

In Section 6.1.2 we explore the medical concepts of *ideal body weight* and *healthy weight* and we summarize some of the methods to obtain those two values. In Section 6.1.3 we present the techniques that allow the extraction of the measures and the weight estimate. In Section 6.1.4 the final application is de-

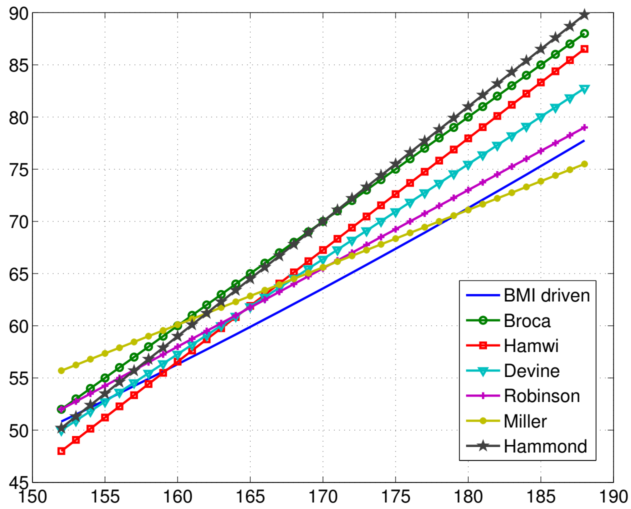


Figure 6.3: The figure shows the outcome of the different equations analyzed in [1].

scribed that provide to the user insights on *ideal weight*, BMI, and lifestyle choices to help reducing weight problems.

6.1.2 Subjective and Objective Ideal weight

The medical community agrees that considering the bones structure, the height, and the sex of a person, there is a weight range which minimizes the risk of contracting illnesses like cardiovascular diseases, and some form of cancer. For the purpose, several methods exist (e.g. bioelectrical impedance) that yield to compute with high precision the amount of fat in one's body; but those techniques require specific machines and trained technicians. For this reason specific equations where analyzed in the recent years to obtain an approximation of optimal body weight.

However, the medical literature contains two contrasting definitions and formulations of the ideal body weight (IBW) problem. IBW was introduced in 1959 when MetLife insurance company proposed the height-weight tables. By considering the sex and height of a subject, the tables assessed the weight able to provide the lowest mortality rate. Since its introduction, this term passed through many formulation and became something not clearly defined. During the last years several works tried to reformulate it so as to have a clear idea of what this value represents.

While on the one side a *healthy* weight can be scientifically defined as the

Table 6.1: Comparison of *ideal body weight* equations.

Source	**IBW equation**
Broca (1871)	(kg) = height (cm) - 100
Hamwi (1964)	(men) = 106 lb [+ 6 lb/in $>$ 5 ft]
	(women) = 100 lb [+ 5 lb/in $>$ 5 ft]
Devine (1974)	(men) = 50 kg [+ 2.3 kg/in $>$ 5 ft]
	(women) = 45.5 kg [+ 2.3 kg/in $>$ 5 ft]
Robinson (1983)	(men) = 52 kg [+ 1.9 kg/in $>$ 5 ft]
	(women) = 49 kg [+ 1.7 kg/in $>$ 5 ft]
Miller (1983)	(men) = 55.7 kg [+ 1.39 kg/in $>$ 5 ft]
	(women) = 53 kg [+ 1.33 kg/in $>$ 5 ft]
Hammond (2000)	(men) = 48 kg [+ 1.1 kg/cm $>$ 150 cm]
	(women) = 45 kg [+ 0.9 kg/cm $>$ 150 cm]

right combination of body cell mass, extracellular water, and nonfat connective tissue [121]; on the other hand ideal weight does not seems to have a clear and scientific definition.

We thus may refer to two different terms one of which is an objective evaluation, the other is a subjective one. One is a scientific measurable quantity, while the other is more subtle as it depends on a personal belief. To make a neat distinction between the two concepts, we will refer respectively to objective IBW (oIBW), and subjective IBW (sIBW).

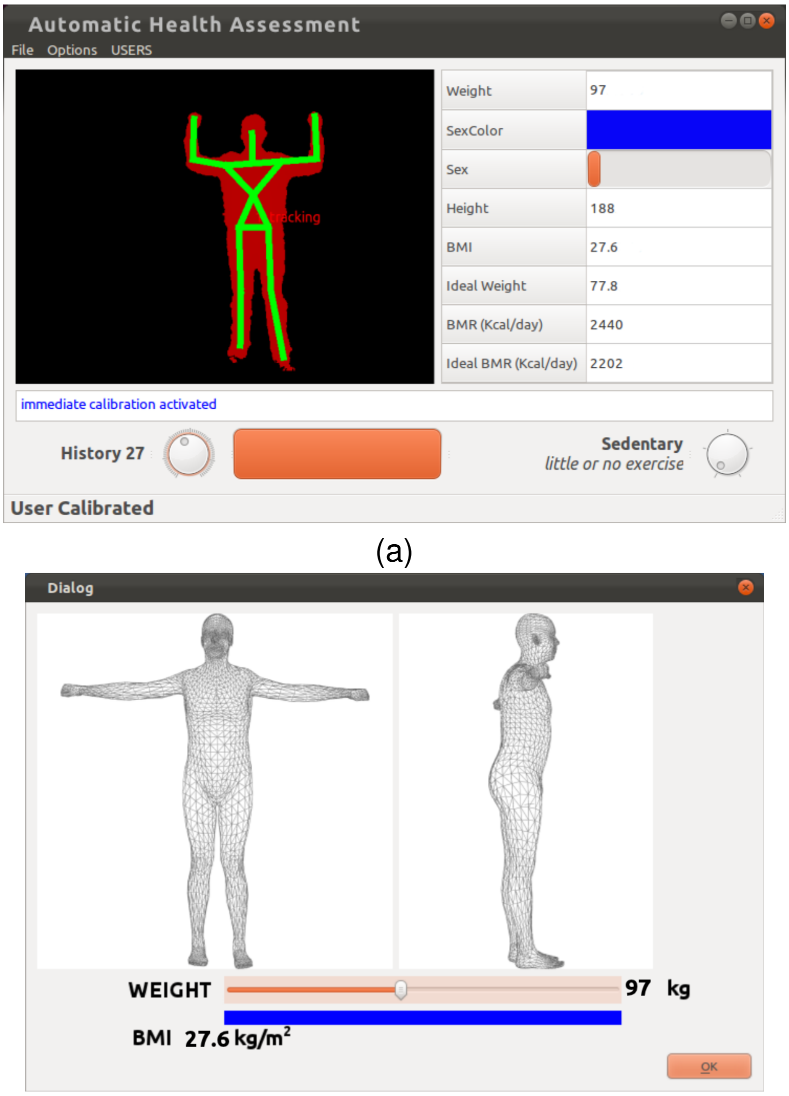
In [122] the authors identified sIBW as the value that replies to the question: "Ideally, how much would you like to weigh at the moment?". Crawford and Campbell demonstrate that people are poorly educated about the healthy weight range. This is especially true for men, whose definition of sIBW is always higher than the healthy weight.

Other works concentrated on finding formulas to express oIBW. An interesting summary of the state of the art is made in [1], the authors group all the known equations (see Figure 6.3 and Table 6.3) and compare them with the height-weight tables produced for MetLife. The results of the comparison is that the formulas are quite similar even if some of those are of unknown source like Broca and Hamwi, which makes difficult to clarify their validity.

The study concludes that Robinson's formula is the closest one to the results of the MetLife height-weight tables, and the one that fits closely to the values provided by considering 22 kg/m^2 BMI (the center of the normoweight range):

$$\textit{oIBW [kg]} = (52 - 3s) + [(h - 152) \times (0.75 - 0.08s)]$$

where $s = \{0 \rightarrow \textit{male}, 1 \rightarrow \textit{female}\}$ and h represents subject's height. Nevertheless, the article suggests to refer directly to the BMI range between 18.5 and 24.9 kg/m^2 that corresponds to the *normal* range. Considering this range the oIBW becomes 22 kg/m^2, that is to say the value associated with the lowest morbidity for both men and women. If we consider the example in Figure 6.4 where subject's height is 1.88 m, and we revert *BMI* formula (*weight[kg]* $=$ *BMI* $\times$ (*height[m]*)^2), we obtain a normal range of 65–88 kg, and an oIBW (at 22 kg/m^2) of 78 kg.



(b)

Figure 6.4: Interface of the application for the automatic self assessment of body health. Image (a) shows the main windows with all the elaborated data. Image (b) shows the interface to the bmivisualizer [123] images that makes clear the link between BMI and body shape.

6.1.3 Weight estimation

Our application exploits the work performed to extract anthropometric measures from an RGBD video and then estimate weight that presented in Chapter 3. While in our preliminary work we considered the weight as a linear combination of anthropometric measure (subsequently using multiple linear regression to train a weight estimator). In this section we try to exploit a possible non linear relation between the anthropometric measures and body weight using a regressor built from a multiple layer perceptron.

We take advantage again of the anthropometric measures provided by the publicly available medical database NHANES, that contains several anthropometric measurements for more than 27000 subjects.

NHANES dataset is split in training, validation, and test set respectively of 30%, 35%, 35%. We firstly test our system obtaining the results shown in figure 6.5 where the cumulative distribution of the relative error is shown. Considering the males population, the majority of the database (90%) falls in the interval of $\pm 5\%$ relative error. That is, if we consider a person weighing 90 kg the correspondent error is of 4.5 kg.

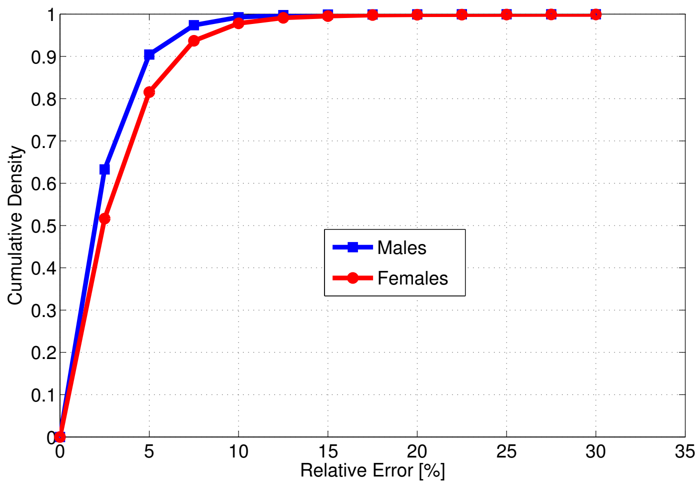


Figure 6.5: The figure depicts the cumulative density function of the relative error. The two groups (Males and Females) are trained and tested separately.

Exploiting the capabilities of the Microsoft Kinect RGBD sensor, our system is able to capture the 3D information of user's silhouette. The silhouette is analyzed and used to extrapolate the anthropometric measures needed to estimate user's weight.

A database of 15 subjects was recorded under minimum controlled conditions where the subjects were asked to wear without loose garments and without heels. As explained in Chapter 3 since the Kinect is able to perceive only the frontal surface, our system had been tuned to estimate what is on the back of the person. Table 6.1.3 summarizes again the absolute errors and reports as well the relative errors of the anthropometric measures gathered with our system. Moreover, we show the results on the weight estimation using the MLP trained on the NHANES dataset. The higher error for the arm circumference is due both to the poor resolution of the Kinect combined with the quantization of depth measurements. These are characteristics of the sensor that combined make of the arm a flat surface at the distance of 2–3 meters. The noisy 3D data provided are exploited over time, the measures are repeated several times and the median value is considered to prevent errors due to the sensor noise.

The precision of the height estimation is definitely the best results among the others measured values. Its precision is of foremost importance in our case as both the BMI and other values computed by our system depend on this term.

In the future the use of combined different sensors and techniques (e.g. face recognition, age and gender estimation) could improve the outcome of our combined mensuration and weight estimation scheme. At present we are considering that age is provided as manual input to our system. To top this, we can exploit the visual appearance of the subject to figure out the sex of user. By using a binary Neural Network classifier to recognize the gender of subjects from the limited set of anthropometric measures in NHANES dataset, we achieve a preliminary accuracy of over 80%. Further improvements may be introduced by increasing the set of measures, using diverse sensors, or by exploiting modalities like RGB image analysis.

6.1.4 Application description

The application we designed uses the anthropometric measures extracted thanks to the Kinect sensor to provide to the users a better vision of their health status.

Limiting ourself to just measure the current health status could not help the person to correctly tackle eventual weight problems. For this reason our system provides helpful insights to the current BMI and ideal weight. Moreover, we display the current (estimated) intake of calories, together with the amount that should be taken in case the ideal weight is assumed as target. The basal metabolic

Table 6.2: The table summarize the statistics about the measurements extracted from a 3D video.

Measure	Absolute error	Relative error
Height (measured)	1.9 cm	1.1%
Arm length (measured)	3.6 cm	12.3%
Arm circumference (measured)	3.2 cm	10.7%
Waist circumference (measured)	8.4 cm	10.0%
Leg length (measured)	2.9 cm	6.2%
Leg circumference (measured)	1.7 cm	3.4%
Weight (deducted)	**2.7 Kg**	**3.6%**

rate (BMR) is one of the measures that correlates most with our daily energy intake [124]. Similarly to the ideal weight, many formulas have been developed to help dietetics to compute this value without necessarily resort to more expensive and time consuming exams. However, in this case the literature shows disagreement inasmuch as some works establish a direct correlation between direct measurements and estimates [124] while others [125] emphasize the error magnitude that the predictions carries in some cases (ill and/or elderly patients).

Nowadays, for historical reasons, the most common equation is the one from Harris-Benedict. However, this has been demonstrated to overestimate in many cases the real value of BMR. To be in line with the last outcomes present in the literature, we resort to the estimation through the use of Mifflin-St. Jeor equation that has been validated as the most accurate [125] in general conditions. Table 6.3 summarizes the formulas used in our system.

Table 6.3: Comparison of *ideal body weight* equations.

Gender	BMR equation (kcal/day)
men	$10\times$ *weight (kg)* $+6.25\times$ *height (cm)* $-5\times$ *age (y)* $+5$
women	$10\times$ *weight (kg)* $+6.25\times$ *height (cm)* $-5\times$ *age (y)* -161

Nevertheless, the BMR indicates the expenditure that a body usually has while resting under given ambient conditions. To compute the real energy that a body need during the entire day, this value has to be multiplied by an appropriate activity factor:

- 1.200 = sedentary (little or no exercise)
- 1.375 = lightly active (light exercise/sports 1-3 days/week, ca. 590 Cal/day)

- 1.550 = moderately active (moderate exercise/sports 3-5 days/week, ca. 870 Cal/day)

- 1.725 = very active (hard exercise/sports 6-7 days/week, ca. 1150 Cal/day)

- 1.900 = extra active (very hard exercise/sports and physical job, ca. 1580 Cal/day)

Thanks to the BMR computation we are able to estimate the current daily calories' intake of a subject. Computing the difference between this value, and the BMR at ideal weight value, we can provide a range of possibilities for losing (or increasing) weight so as to reach the ideal goal.

In figure 6.4 we show the interface of our system. All the input and output parameters that are computed by our automatic algorithms are easily recognizable.

In the example depicted in the figure the estimated calories intake for the subject is 2440 kcal/day, the system suggest that the BMR at the oIBW should be circa 2202 kcal/day. By reducing of 200 kcal/day his/her BMR and by keeping the same activity level, the subject could reach his/her oIBW. Reporting a 5% error in the weight due to the computer vision estimation will make oscillate the value of ± 50 kcal/day, which still make our system a helpful information.

6.1.5 Conclusion

Obesity is an important health problem related to being extremely overweight. It is spreading widely like an epidemic and preventive tools and methods are needed to increase people self-awareness. We believe that computer vision, and especially 3D body analysis can be employed to fill the gap of self diagnosis methods. We propose a system based on the Microsoft Kinect RGBD sensor, our application extrapolates anthropometric measures from the body silhouette and 3D information. The data correspond to well defined anthropometric measures directly related to a statistical model that estimates subjects weight using data from the NHANES medical dataset. Thanks to the knowledge acquired while analyzing the user, the system can provide through its interface healthiness measures like person's ideal weight, current BMI, estimated calories' intake and ideal one. All those values are intended as support to a healthy lifestyle that would mitigate the effects of weight problems.

6.2 Applications to space

Thanks to our collaboration with the Center for Space Human Robotics of the Italian Institute of Technology we contributed to provide a possible solution to the problem of mass estimation of cosmonauts. Since the very beginning of space exploration, cosmonauts have suffered from weight losses which need to be particularly monitored during long term missions in space stations to insure their health and well being. In 1965-6 Thornton successfully built a device able to measure the body mass of cosmonauts in the micro-gravity space environment using passive linear spring-mass oscillators. Since then, space stations have been equipped with labs containing, among others, bulky devices like Thornton's. In this work we apply our recent advancements in computer vision algorithms allowing us to estimate the weight of a person within 4% error using 2D and 3D data extracted from a low-cost Kinect RGBD camera output.

6.2.1 Astronauts' weight

On April 12, 1961, the Soviet cosmonaut Yuri Alekseyevich Gagarin orbited the Earth, for the first time in the history of mankind; since then, in the last 40 years more than 500 cosmonauts [126] have been sent to the space.

One of the findings throughout the history of spaceflights concerns the loss of body mass that affects cosmonauts [127]. Typically, these losses are small (about 2%), but they can reach up to 10 to 15 percent of preflight body mass in particular cases. Most of the observed loss of body weight is accounted for by loss of muscle and adipose tissue [128]. As a matter of fact, micro-gravity severely changes the human physiology leading to loss of muscle mass and muscle volume, weakening muscle performance, especially in the legs.

To avoid some of these phenomena, since the 70's, cosmonauts are given precise exercises and feeding routines which were, in the following years, gradually adjusted and incremented to the current levels [128]. As a result, in a typical day on board of the International Space Station (IIS), crew members spend twelve hours working, two exercising, two preparing and eating meals, and eight hours sleeping. Weight, or more generally body mass, has been always considered as a good indicator of correct nutritional status and well being; however, a problem still exists on how to measure and track cosmonauts' weight in micro-gravity conditions.

Over the centuries, two main methodologies have been adopted to measure

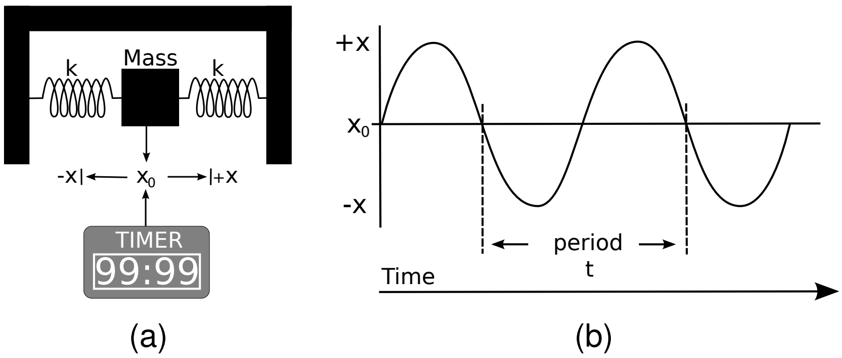


Figure 6.6: The images shows the principles of mass measuring devices: (a) the mass oscillates thanks to springs of known parameters. (b) The period measure is directly proportional to the mass. *[Images courtesy of the National Space Biomedical Research Institute]*.

the weight of an object: the spring scale and the balance. Unfortunately, both of these techniques make the assumption that a relatively strong and constant gravity acceleration field is affecting the object we want to weigh.

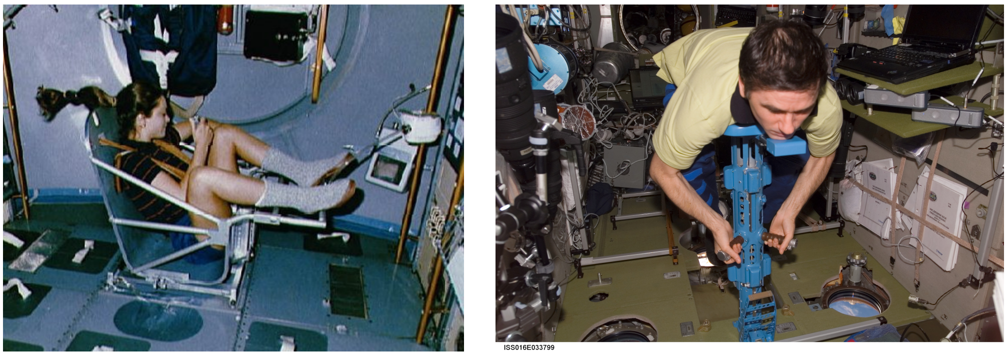
Because in space this assumption is not true, in order to measure mass during spaceflights, an acceleration field must be artificially generated.

The general principle is shown in figure 6.6, a sample mass (M) is made oscillating thanks to two springs of known constant (k) constrained to linear motion along a predefined axis. When the mass is displaced from the equilibrium position (x_0) and then released, it will undergo a round trip undamped oscillation in a period of time (t). By reversing the formula $t = \frac{2\pi M}{k}$, we obtain the mass measure as $M = \frac{kt}{2\pi}$.

In 1965/66 the US cosmonaut and researcher William Thornton, developed and tested a device (figure 6.7.a) that, making use of such a passive calibrated linear spring, was able to estimate mass of people by computing their characteristic frequency of vibration. Since then, similarly designed devices are sent into space to measure the mass of cosmonauts, garbage containers, and scientific specimens.

The current technology for weighing cosmonauts evolved since the Thornton device and involves a precisely calibrated spring attached to a support to which the cosmonaut is rigidly fastened and which is moved several times to estimate the kinematics of the structure and thus the mass of the cosmonaut (see figure 6.7.b).

Such a system assumes the cosmonaut to be a rigid body during accelerations and decelerations, and it requires both space and energy, two resources which are quite scarce in space explorations. The precision of such a system, has been re-



(a) (b)

Figure 6.7: The images shows the Thornton first experimental device (a), and the last and still used Russian BMMD device (b) *[Images courtesy of the Image Science and Analysis Laboratory, NASA Johnson Space Center]*.

cently debated [129, 130, 131], especially due to deviation from the baseline which can be in the order of 6 to 7 kilograms for the latest NASA device, the SLAMDD, and in the order of 2 kilograms for the latest Russian BMMD device currently serving the European ISS.

6.2.2 Vision Based Weight Estimation

Vision is a very interesting way to estimate the weight of a subject; in fact, as seen in Chapter 3, such approach is used in hospitals during emergencies. In space, cameras (possibly 3D) could easily be integrated into the walls and automatically record cosmonauts in their daily life autonomously tracking their weight several times a day unobtrusively, accurately, and even without the need for cooperation.

Considering our preliminary results whose conclusions were drawn in Chapter 3 the visual estimate differs from the true value with an average error of 4.3%. Supposing a 80 kilograms average weight of cosmonauts this translates into an error of ± 3.44 kilograms which is in-line with the results of the much more expensive, energy and space consuming, devices on board of the ISS [129, 130, 131]. The main limit of our first approach is that the features are computed off-line by manually tagging subjects' anthropometric landmarks.

Exploiting our framework already used for the ideal weight computation we present our plans for the development of next generation vision based weighing device.

In the future a biometric system such as face recognition could recognize the subject and give us the exact age, sex, expected anthropometric measures, and

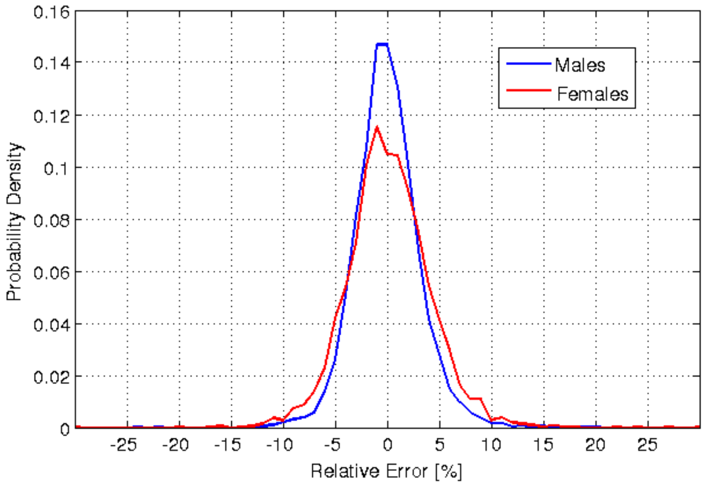


Figure 6.8: In our theoretical test a subset of people divides the database respectively by age (older than 20 years), and by gender. The figure shows the probability density function of the relative error for both males and females subjects.

many other a-priori knowledge that could help the weight estimate. At present we are constraining the database to represent as closely as possible the space environment. In particular, considering that about 65% of the cosmonauts are selected in United States and age between 26 and 46 years (the average age being 34 years) [3] we notice that the algorithm can be designed and targeted to work by exploiting such prior information.

Moreover, we can exploit the gender difference to apply different models to the two gender classes. To provide further automatism the visual appearance of the subject can be exploited to figure out the sex of the cosmonaut using computer vision algorithms like we have already seen in Chapter 3. Methods like the ones we presented can provide (in case of accurate measures) very high accuracy if we consider a full list of anthropometric measures. Also in this case possible improvements may be introduced by coupling other modalities (e.g. RGB image analysis).

We report here the weight estimation result obtained by selecting (from the 27000 subjects of the NHANES database) entries corresponding to the two separated male and female classes which only includes people older than 20 years. For the sake of brevity we will discuss only the results for the male class, the considerations are equal for the female class since the performance do not vary considerably.

The results shown in figure 6.8 are obtained using an Artificial Neural Network

[3] http://astronauts.nasa.gov/content/faq.htm

(ANN) with a single hidden layer consisting of 5-neurons and a hyperbolic tangent sigmoid transfer function. The experiment was performed dividing the dataset in training, validation, and testing set, respectively of 35%, 15%, and 50% ratios. We obtain that, over the entire male class, the mean absolute relative error is 0.023, thus in average the system error falls within the range of $\pm 2.3\%$.

In figure 6.8 the probability density function is Gaussian shaped with a standard deviation of 0.032 meaning that 95.5% of estimates fall into the range of $2\times\sigma = 6.4\%$ relative error. For the male class the entire dataset is covered considering a relative error of $0.1 = 10\%$ (almost $3 \times \sigma = 9.6\%$, thus 99.7% of the dataset). From the one side, it is clear that being the ANN approach a non linear regression technique, it is more capable of modeling non linear relationships among measures and weight of the subjects. On the other hand by selecting only a specific gender and age subset helps the regressor to find a more suitable way of estimating weight.

The result just presented in this section is interesting because it shows us what could be a theoretical limit to weight estimation through visual analysis. To our knowledge this limit could be further improved if additional anthropometric measures are considered, thus increasing the complexity of the computer vision algorithms and sensors.

Exploiting the framework we presented previously in Section 6.1.3 we could possibly provide a solution to cosmonauts whose interface is shown in figure 6.9.

6.2.3 Further Work and Discussion

We have introduced the issue of determining the mass of people in space microgravity environment and presented a solution to this challenge based solely on computer vision techniques. Our technology relies on RGBD cameras such as the Microsoft Kinect. The current development makes use of estimated anthropometric measures extracted from a frontal 3D view of the user. Thanks to our algorithms we are able to extract limbs measurements and to estimate subjects' weight up to 2.7 kg average absolute error. Future work could focus on further ameliorating the 3D reconstruction by both using multiple (and therefore closer) Kinect sensors and implementing algorithms specifically targeted at registering non rigid structures like human body. In this way one could be able to build a full body 3D model from different viewpoint and different postures.

Finally, as also acknowledged by John Charles, chief scientist on NASA's hu-

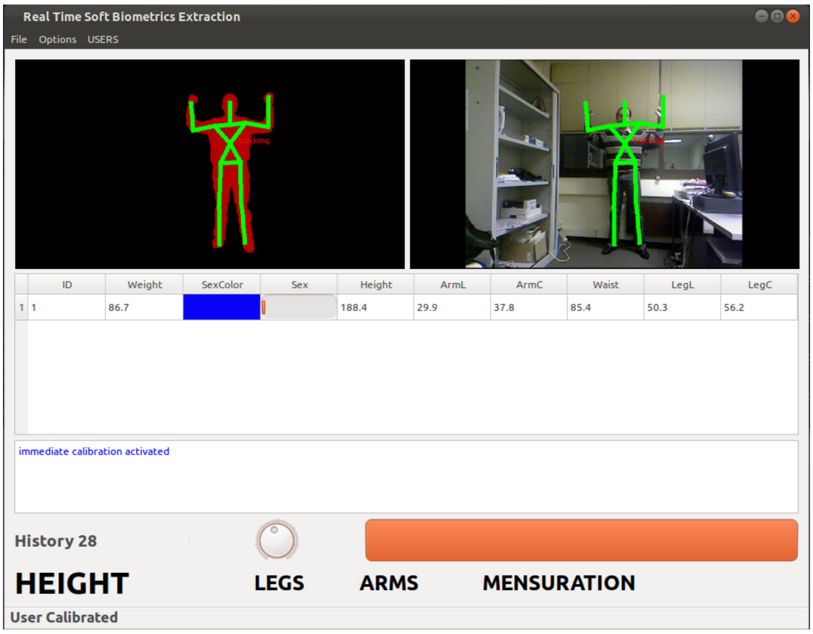


Figure 6.9: Interface of our automatic weighing module. The system analyzes user's silhouette and 3D shape to extract anthropometric measures and apply the formulas extrapolated from NHANES dataset analysis.

man research program in Houston, Texas [4], the possibility of combining current devices, our Space Scale, and other medical technologies could enable the possibility of assessing over time the quality of body mass lost in term of either bone, blood, muscular, or fat masses.

6.2.4 Critical analysis of the contributions

Notwithstanding the unusual size of NHANES dataset (27000 subjects) that provides enough data to obtain statistically valid results, it is not often the case for biometric database to have those dimensions. Moreover, the quite recent commercialization of the Kinect sensor impaired us to perform a bigger campaign of acquisition for our database, nor other databases for body analysis through Kinect were presented to the community.

A possible solution to that problem might be the simulation of a Kinect dataset by adding noise to the cloud points provided by dataset like the already mentioned CAESAR dataset [104].

[4] http://www.newscientist.com/article/mg21228443.700-kinect-weighs-astronauts-just-by-looking-at-them.html

CHAPTER 7

Conclusion

This dissertation explored the possibilities offered by body soft biometrics, a new set of biometric traits that share specific characteristics as seen in Chapter 2. In the same chapter we underlined an extensive, yet not exhaustive, list of soft biometrics traits and we clearly defined what can be included into this category. Moreover, we define the neat difference between traits instances and continuous values, and we explain why for automatic systems is more convenient to work with the latter, while managing the user communication with the former. Furthermore, we introduce the main three areas of application that may benefit from the use of Soft Biometrics.

In Chapter 3 aside some state of the art techniques for height estimation, we present our approach to anthropometric measure estimation from both 2D images and RGBD videos. Inspired by the work on forensic height estimation, we demonstrated that a similar approach can yield to weight estimation. A large medical dataset (27000 records) is considered for our initial tests about theoretical performance and about the impact of noisy measurements. Furthermore, we take advantage of our 2D anthropometric measuring system to provide experiments on real data. Finally we apply similar considerations to gender classification. However, we show that in case of classification, noisy measurements yield to limited results in the real case scenario.

The pruning capabilities offered by a body soft biometric signature are exploited in Chapter 4. We demonstrate that pre-classification based on anthropometry has beneficial effects on a large face recognition system, and it improves both speed and accuracy of the hard biometric system. Our results clearly identify the trade-off between noisy measurements, pruning capabilities, and the accuracy of the final system.

In Chapter 5 a camera network scenario is used to test the possibility of performing people re-identification. Our results show that, although the quantity of information is limited, a body soft biometric signature could be enough in case of small groups of subjects. Thus, we verify the theoretical outcomes of a previous work on Bag of Soft Biometrics [11]. Moreover, we show how revealing the soft biometrics is generally privacy safe.

In Chapter 6, body soft biometric are used to profile users of a medical system. A first application analyzes the user to automatically check up his/her body's condition. Moreover, the system provides lifestyle suggestions and identifies one's ideal weight. To obtain such results we apply state of the art medical formulas to parameters extracted from the body silhouette.
Since in case of microgravity dangerous weight losses usually occur, the second application is conceived to support cosmonauts' weight monitoring. Our work compares the performance of existing systems on board of the International Space Station with our computer vision based technology.
Anthropometry and RGBD sensors are in strong relation and in the very near future they are going to be used together to provide improved services in gaming, wellness, and automatic tailoring.

This thesis presented the opportunity of exploring and connecting different research domains, it is not always an easy task for many factors. Some of the reasons are related to different research patterns and different ways of conceiving research, these lead to increase the difficulty of reading, understanding, and experimenting techniques. On the other hand the fascination that those other worlds emit can yield to many open doors and as many different paths. We have found a balance between novelty of the approaches, application scenarios, and new technologies. In the next and last section we indicate some of the open doors that remain unexplored.

7.1 Future Works

In this work we have outlined the principles of Body Soft Biometrics; we have presented some applications that leverage the semantic understanding provided by those traits to explore different domains. The main three areas of interest we investigated are pruning, identification, and people description via body analysis. We leave the door open to other contributions in some aspects of our research.

7.1.1 New traits

The list of traits we provided, is not intended to be exhaustive, many other features may be visible and appreciated under different circumstances, and depending on the needs of the application in which they have to be used.
In the future, new sensors empowered with better resolutions may help to capture details that currently are difficult or impossible to gather at a distance. Furthermore, better technologies might help refining the mensuration process to obtain a better precision.
An example of additional trait can be represented by the description of the physical style of one subject, not limiting ourself to the identification of clothes colors but to the type of clothes, not only to the hair color but also to their style. This additional information may also help providing a quality factor for body part measures. First experiences in this field are found in [71].
Other traits may be found in other domain of biometrics like speech analysis where voice tone, accent, and intonation may reveal interesting soft biometric characteristics.

7.1.2 Database

A touchy aspect in every type of experimental research is the creation of adequate databases. In our case the NHANES dataset represented a rare source of information for the problem of weight estimation.
Nevertheless, although being a very large dataset (27000 individuals), it contains only 9 body traits (also considering weight and gender), this induces implicit limitations to studies that may be performed on the data.
For this reason other datasets have to be analyzed that could provide a richer set of variables. A possible, existing, solution may be the CAESAR [104] dataset that contains multiple 3D scans of more than 4.400 US, Canadian, and European citizens. Having a 3D scan instead of linear measures increases the range of possible studies that can be performed on body soft biometrics and anthropometry.

7.1.3 New sensors and technologies

Body scanners represent a good example of interesting new technology that might find application in exploring body soft biometrics. Those new devices are able to control the body silhouette through clothes without requiring a manual intervention. While today the inspection of full body scans is still performed visually by an operator, in the future the procedure could be automatized, and the scenario we described in Chapter 4 could become reality.
As we have already seen that commercial 3D equipments may be used for biometric extraction; other sensors like thermal cameras, embedded sensors, and sensors networks may represent an interesting source of information for extracting biometric traits.

7.1.4 Research studies

Many research paths are still unexplored. We believe that dynamic analysis can unveil other interesting areas where to investigate soft biometric applications.
The first section to analyze would be studying the interaction between weight and gait. For example to provide more insights about someone's identity, or to understand unusual patterns that may lead to discover malicious intentions or behaviors, or yet to establish the link between those two soft biometric traits.
The expression of movements can also be considered as soft biometric trait itself. Some studies already showed the feasibility of people identification from the way some actions are expressed [132].

During the noise impact analysis of Chapter 3 we experimented the decrease of performance of some formulas that initially provided the best results. This suggests that some of these features are more robust to measurement error than others, and it enforces the belief that a possible results improvement could be to provide a reliability measure to each of the extracted trait values.
A similar consideration can be done for the work of Chapter 4 applying such reliability measure will improve also the pruning phase, undoubtedly ameliorating the performance of the classification scheme.

A solution may come from some recent studies [18] that already demonstrated how different body parts show interesting correlation patterns. Those correlations may be explored to reduce measurements errors, or to provide solution to the missing data problem. Similar outcomes may be obtained by advances in 3D reconstruction that could improve anthropometric measures collection and body parts segmentation.

APPENDIX A

French Version : Introduction

Le processus d'identification a toujours été de premiere importance pour les êtres humains. L'identification d'autres personnes permet des interactions et d'autres activités sociales qui sont essentielles pour nous. Toutefois, ces capacités innées qui caractérisent les êtres humains, ne sont pas disponibles pour les ordinateurs et les systèmes automatiques. Pour cela, les informaticiens ont passé ces dernières décennies de travail sur le domaine de la biométrie pour donner ces capacités aux ordinateurs.

Pendant de nombreuses années des informations biométriques ont été exploitées pour extraire de l'information sous la forme de caractéristiques et pour fournir des indications d'identité aux ordinateurs. Malgré que l'utilisation pour l'identification a été abondamment exploré, il reste beaucoup à étudier sur l'identification lorsque le contact et la coopération des sujets ne peut pas être garantie.

En outre, même si les humains peuvent facilement comprendre et décrire des concepts sémantiques et de les utiliser pour indiquer des actions ou des sujets, les ordinateurs rencontrent encore des difficultés à extraire une représentation plausible d'une personne à partir de son/sa description physique. Pour cette raison, récemment, la communauté de recherche a commencé à explorer des techniques qui permettent aux machines d'exploiter cette représentation sémantique. Combler le fossé sémantique permettra une complète et satisfaisante interaction entre l'homme et la machine.

Dans les dernières années la biométrie douce, une nouvelle branche d'études biométriques, a commencé à explorer la représentation sémantique de personnes et a reçu une attention croissante. L'objectif principal de la biométrie douce est d'extraire des informations quasi-uniques (ou communes) sur quelqu'un, à partir d'images ou des vidéos. Ces traits sont surtout orientés vers la description des attributs sémantiques et donc souvent dénommée traits sémantiques [5].

Dans l'approche biométrique classique un modèle (par exemple l'image du visage) est extrait à partir de l'échantillon biométrique (par exemple, le visage lui-même), enfin une version numérique de celui-ci est comparée à une galerie de modèles. Avec la biométrie douce nous opérons avec des traits sémantiques (par exemple : la taille, ou la couleur des vêtements) dont la validité générale ne permet pas l'identification per se. La première approche est généralement valable pour les humains, comme par exemple nous sommes en mesure d'identifier les personnes à partir de leurs photos. Néanmoins, les humains sont capables d'exploiter les descriptions de niveau sémantique pour effectuer des actions similaires (par exemple nous pouvons identifier quelqu'un dans la foule à partir de sa description physique). Dans l'avenir employer la biométrie douce permettra l'utilisation des caractéristiques pour décrire les personnes à la machine, et vice-versa.

Grâce à leur contenu sémantique, être capable d'extraire les traits fournira à la machine une connaissance plus approfondie sur l'utilisateur avec lequel il est en interaction. Tout cela pourrait être exploité afin d'améliorer les techniques actuelles d'interaction homme-machine. Un exemple pourrait être représenté par un distributeur automatique qui pourrait éviter de fournir de l'alcool ou des cigarettes aux enfants; ou qui pourraient trouver pour une personne robuste avec un bonnet noir qui a volé mon sac à main rouge.

A.1 Motivations, objectifs et contributions

Les motivations qui ont conduit les travaux de cette thèse sont venu de l'absence de l'exploration adéquate de la biométrie douce et de l'anthropométrie au sein de la communauté biométrique. Pour ce qui concerne les technologies, les algorithmes, et des idées, des concepts nombreux sont encore inexplorés et seulement récemment la communauté a commencé à investiguer ce sujet intéressant. Pour motiver davantage notre travail, nous explorons différentes applications où l'extraction des traits de biométrie douce peuvent être employées pour ouvrir plus largement les perspectives de ce thème de recherche.

Nos contributions sont basées sur le développement et la formulation d'une définition claire des traits de biométrie douce, et la façon de les classer dans différents domaines (par exemple : du corps, du visage, et les classes d'accessoires). En outre, pour distinguer clairement l'avantage qui vient avec l'utilisation de ces éléments, nous exploitons des caractéristiques propres de la biométrie douce : compatibles avec les modalités des hommes, ne nécessitent pas de registration, non-intrusives. Surtout la première est d'une importance fondamentale en ce qu'elle souligne la capacité intrinsèque qui fait la biométrie douces compréhensible au même temps par l'être humain et par la machine.

Les objectifs de notre travail sont de fournir un moyen fiable pour extraire les traits biométriques douces, relatifs au corps, et à démontrer leur applicabilité dans le cas de ré-identification de l'utilisateur, de l'élagage des base de données biométriques, et de l'interaction homme-machine.

Notre contribution se focalise en particulier sur l'étude des traits biométriques douces relatifs au corps. Nous avons exploré différentes manières pour extraire le poids, la taille, et les information du genre de la personne employant des techniques et des capteurs différentes. Nous avons appliqué les résultats dans de nombreuses applications différentes.

Comme nous le verrons plus tard, même si dans le cas des l'estimation de taille existait d'autres algorithmes pour obtenir ce trait, dans le cas du poids, notre travail constitue l'une des premières applications de l'anthropométrie pour l'estimation du poids, surtout en utilisant des capteurs vidéo; en plus nous avons trouvé plusieurs domaines d'applications qui pourraient bénéficier de ces capacités.

A.2 Organisation de la thèse

Le manuscrit peut être divisé en deux parties principales. La première partie présente un état de l'art sur la biométrie douce et quelques techniques d'extraction de biométries douces (chapitres 2 et 3). La deuxième partie est plutôt consacrée à la présentation de certaines applications qui exploitent les traits de biométrie douce relatifs au corps (Chapitre 4, Chapitre 5, et le Chapitre 6).

Dans le chapitre 2, nous revenons sur l'histoire de la biométrie douce et nous passons en revue leur introduction tardive dans les études biométriques. Nous clarifions une nouvelle définition qui décrit mieux ces traits, en plus nous passons en revue les traits les plus importants avec une attention particulière consacrée à la biométrie douce relatif au corps. Par ailleurs, on définit le domaine d'application de la biométrie douces et nous présentons les trois principaux domaines qui seront explorés dans cette thèse.

Le chapitre 3 examine différentes façons d'extraire des données biométriques douces relatives au corps. Nous fournissons l'état de l'art de certaines techniques pour l'extraction de la taille, ensuite nous explorons plus la partie qui concerne l'extraction des mesures anthropométriques grâce à de nouveaux capteurs 3D abordables (comment la Kinect). Un étude théorique est décrite sur une technique d'estimation de poids à partir de données anthropométriques grâce à la base des données National Health and Nutrition Examination Survey (NHANES) enregistrée par l'American Center for Disease Control and Prevention (CDC).

Dans le chapitre 4, nous étudions les capacités des caractéristiques biométriques douce pour effectuer l'élagage d'une base de données biométriques conventionnelles. Nous montrons comment une signature composée par la biométrie douce du corps peut améliorer la vitesse de récupération des données et à la fois la précision du système. Ensuite nous montrons notre résultats dans une expérience pratique comment nous pouvons améliorer au moins de deux fois la performance de l'algorithme de reconnaissance. Le pré-classification sur la base de mesures anthropométriques est testée en conjonction avec un algorithme de reconnaissance de visage dont les performances obtient avantage de la phase d'élagage. Enfin, nous avons clairement identifier le compromis entre la taille, la précision et l'erreur d'un système de mensuration basé anthropométrique.

Dans le chapitre 5, nous explorons la possibilité d'effectuer de la ré-identification dans un environnement vidéo grâce à un organisme de surveillance à signature biométrique douce composé de la taille, une mesure en corrélation avec le poids de la personne, et la couleur des vêtements du sujets. Nous montrons les différentes techniques utilisées pour extraire ces paramètres, puis nous présentons des résultats sur la re-identification. Dans ce chapitre, nous voyons que l'exposition au public des caractéristiques biométriques douce préserve bien la vie privée des personnes sous surveillance.

Dans le chapitre 7, nous donnons quelques conclusions plausibles et nous explorons de nouvelles perspectives et des travaux futurs. Cette thèse a été rédigée en partie sur les publications suivantes.

- Velardo, Carmelo; Dugelay, Jean-Luc, What can computer vision Tell you about your weight?, Under submission/preparation

- Velardo, Carmelo; Dugelay, Jean-Luc; Paleari, Marco; Paolo, Ariano, Build-

ing the space scale or how to weigh a person with no gravity, 1st IEEE International Conference on Emerging Signal Processing Applications, January 12-14, 2012, Las Vegas, Nevada, USA

- Velardo, Carmelo; Dugelay, Jean-Luc, Improving identification by pruning : a case study on face recognition and body soft biometric, Research report RR12-262, pp 1-20
- Velardo, Carmelo; Dugelay, Jean-Luc; Daniel, Lionel; Dantcheva, Antitza; Erdogmus, Nesli; Kose, Neslihan; Min, Rui; Zhao, Xuran, Introduction to biometry, Book chapter of Multimedia Image and Video Processing (2nd edition); CRC Press; 2011
- Velardo, Carmelo; Dugelay, Jean-Luc, Real time extraction of body soft biometric from 3D videos, ACM Multimedia 2011, 28 November-1 December, 2011, Scottsdale, Arizona, USA
- Velardo, Carmelo; Araimo, Claudia; Dugelay, Jean-Luc, Synthetic and privacy-preserving visualization of camera network outputs, 5th ACM/IEEE International Conference on Distributed Smart Cameras, August 22-25, 2011, Ghent, Belgium
- Dantcheva, Antitza; Velardo, Carmelo; D'angelo, Angela; Dugelay, Jean-Luc, Bag of soft biometrics for person identification : New trends and challenges Multimedia Tools and Applications, Springer, October 2010, pp 1-39
- Velardo, Carmelo; Dugelay, Jean-Luc, Weight estimation from visual body appearance, IEEE 4th International Conference on Biometrics: Theory, Applications and Systems, September 27-29, 2010, Washington DC, USA, pp 1-6
- Velardo, Carmelo; Dugelay, Jean-Luc, Face recognition with DAISY descriptors, ACM SIGMM Multimedia and Security Workshop, September 9-10, Rome, Italy, pp 95-100
- Paleari, Marco; Velardo, Carmelo; Huet, Benoit; Dugelay, Jean-Luc, Face dynamics for biometric people recognition, IEEE International Workshop on Multimedia Signal Processing, October 5-7, 2009, Rio de Janeiro, Brazil, pp 1-5

APPENDIX B

Biométrie douce : état de l'art

B.1 Introduction

Le terme *biométrie* provient de la fusion de deux anciens mots grecs : bios (la vie) et metron (mesurer, compter). Ces deux mots indiquent qu'il y a quelque chose lié à la vie (la nature humaine) qui peut être mesuré ou compté. La biométrie est la science qui essaie de comprendre comment mesurer les caractéristiques qui peuvent être utilisés pour distinguer les individus. Les êtres humains ont développé ces compétences au cours de l'évolution : le cerveau a des domaines spécialisés afin de reconnaître les visages [6] et de relier les identités avec des modèles spécifiques (comportementales ou physiques [7]). Les chercheurs dans le domaine de la biométrie ont toujours essayé d'automatiser ces processus en les rendant aptes à être exécuté sur un ordinateur ou un périphérique. L'étude de modèles biométriques conduit à la définition des exigences qui doivent être respectés pour faire un trait humain possible d'être utilisé dans un processus de reconnaissance.

Une caractéristique biométrique peut se résumer ainsi : une caractéristique que chaque personne devrait avoir (universalité), dans laquelle deux personnes doivent présenter certaines différences (spécificité), qui ne devrait pas radicalement varier sur une période prédéfinie (permanence), et qui devrait être quantitativement mesurables (recouvrement). En outre, les traits biométriques peuvent être divisés en deux classes suivantes : physiques et comportementales. Pour la première catégorie appartient l'apparence du visage, le modèle de l'iris et les

(a) (b) (c)

Figure B.1: Les images montrent la méthodologie appliquée pour recueillir des informations sur le suspect dans le système du bertillonnage. La procédure a été normalisé par Bertillon, dans son livre [8].

empreintes digitales, la structure des vaisseaux sanguins de la rétine ainsi que la forme de l'oreille. Chacun de ces traits peut en outre être subdivisé en génotypique et randotypic, le première indique une corrélation avec des facteurs génétiques héréditaires (comme des similitudes chez les jumeaux), l'autre caractéristique décrit les traits qui se développent de façon aléatoire lors de la phase foetale. Les biométries comportementale se développent quand nous vieillissons et ils ne sont pas a priori défini. à ces traits appartiennent la démarche, et même la façon de taper sur un clavier.

Dernièrement, un nouveau concept biométrique, la biométrie douce (également appelé sémantique [5]), a acquis une influence, car il peut augmenter la fiabilité d'un système biométrique et peut fournir des avantages : les caractéristiques de biométrie douce révèlent des informations biométriques, ils peuvent être en partie issu de la biométrie classique, ils ne nécessitent pas de l'enregistrement de l'utilisateur et peuvent être acquises de manière non intrusive, sans le consentement ni la coopération d'un individu.

Ces traits partagent beaucoup de similitudes avec le travail d'Alphonse Bertillon, un agent de police française qui a d'abord exploré la biométrie et a mis en place un système d'identification biométrique basée sur la classification anthropométrique. Avant son système, les criminels ont pu être identifiés sur la base de récits de témoins oculaires ou par le marquage (qui a toutefois été abandonnée en France en 1809 laissant un vide dans les méthodes pour l'identification des criminels). La méthode moderne conçu par Bertillon (pour cette raison surnommé Bertillonnage) a consisté dans une série de mesures anthropométriques dont la méthode d'acquisition a été standardisée dans le manuel de Bertillon [8] (voir la figure B.1 pour quelques exemples).

Les données était rassemblées avec précision sur des fiches représentant des images d'identification du suspect, puis rangées sur des fichier. Les fiches était indexées en fonction d'une combinaison donnée des mesures anthropométriques de façon à garantir la possibilité d'un accès rapide au dossier de sujets. Un exemple de fichier Bertillon (l'ancêtre des photo d'identité judiciaire) est montré dans la figure B.2.

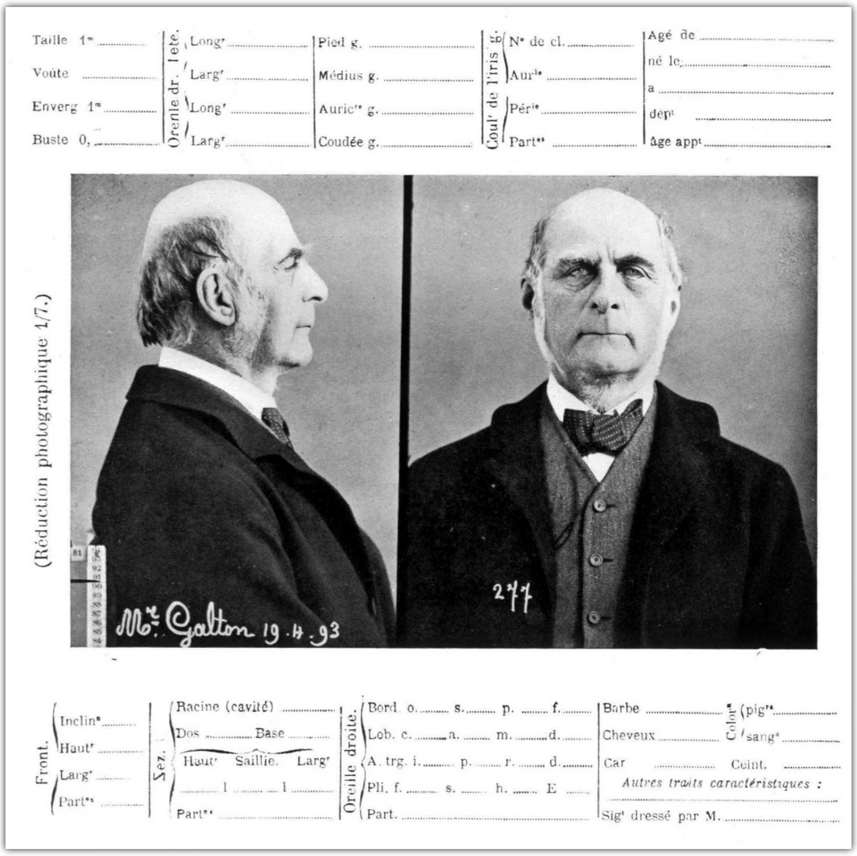


Figure B.2: Un exemple de fichier de Bertillon. Il montre la photo de M. Francis Galton au cours d'une visite faite au laboratoire de Bertillon en 1893. Galton est surtout connu pour ses études sur l'unicité du motif d'empreintes digitales.

En dépit de son utilisation à grande échelle à travers la France, la Grande-Bretagne, et les états-Unis d'Amérique [9], le Bertillonnage a été trouvé défectueux comme dans le cas de Will West vs William West [10] qui a prouvé la plus grande fiabilité des systèmes à base des empreintes digitales. à partir de là l'utilisation des mesures anthropométrique pour l'identification a été abandonnée et il a été remplacé par la correspondance des empreintes digitales et d'autres techniques biométriques plus distinctives. Cependant, dernièrement une grande attention a été tracée sur les caractéristiques appelées à biométrie douces qui rappellent pour certains aspects l'ancien système de Bertillon.

B.2 Nouvelle définition de la biométrie douces

Malgré les nombreuses études qui explorent le thème de la biométrie douce, la littérature présente une absence d'une définition formelle qui décrit ce que peuvent être considérés comme appartenant à cette catégorie. Pour cette raison, nous avons dérivé une nouvelle définition de la biométrie douce qui généralise ce concept [11].

> ❝ *Les traits de biométrie douce sont des traits physiques, comportementaux ou des caractéristiques que l'homme a déjà classées dans des catégories prédéfinies. Ces catégories sont, contrairement au cas de la biométrie classique, déjà présentes et bien établies par les humains dans le but de différencier les individus. C'est-à-dire, les instances des traits de biométrie douce sont créés d'une manière naturelle, utilisées par les humains pour distinguer leurs pairs.* ❞

Les traits qui acceptent cette définition comprennent, mais ne sont pas limités au genre, le poids, la taille, l'âge, la couleur des yeux, l'origine ethnique et ainsi de suite. Une augmentation des ressources (comme une amélioration de la résolution des capteurs, ou une grande capacité de calcul) peut conduire à l'expansion de la quantité des traits et en outre à l'expansion du nombre de valeurs qu'on est capable d'échantillonner. La nature des fonctions biométriques douce peut être binaire (par exemple la présence ou pas de lunettes), continu (taille) ou discrètes (ethnie) [11].

En cas de valeurs discrètes, nous pouvons diviser encore chaque trait dans ses échantillons. Des exemples pour les instances de la couleur des cheveux pourrait être : blonde, rouge et noir. Il est important de noter que les valeurs continues peuvent être éventuellement discrétisées.

Comme pour la biométrie classique, ces caractéristiques peuvent être différenciés en fonction de leur caractère distinctif et de la permanence, ou le caractère distinctif mesure la puissance d'un trait de distinguer les sujets au sein d'un groupe, et la permanence se rapporte à l'invariabilité d'un trait dans le temps. Ces deux caractéristiques sont pour la plupart dans une gamme inférieure pour la biométrie douce en les comparant à la biométrie classique (par exemple la couleur des cheveux, ou la présence de la barbe, peut changer au cours du temps et sont partagés entre les personnes). Pour cette raison, une distinction est tracée entre la biométrie classique et la biométrie douce; le premier sont des traits qui peuvent singulièrement fournir un degré plus élevé de reconnaissance, ce qui est plus approprié pour l'identification.

En outre, il est d'intérêt d'évaluer avec quelle fiabilité l'estimation d'un trait peut être extrait d'une image ou une vidéo. En exploitant ces trois qualités, c'est-à-dire le caractère distinctif, la permanence, et la fiabilité d'estimation, l'importance d'un trait de biométrique douce peut être déterminée. Nous notons que la classification des traits de biométrie douce peut être étendu afin d'évaluer ou de déduire des aspects tels que la précision et l'importance, en fonction de l'application.

Pour résumer parmis toutes les caractéristiques que caractérisent la biométrie douce nous fournissons une description de certains d'entre eux dans la liste cidessous .

- Conforme aux descriptions humaines : Traits conformes à des descriptions existantes.

- Efficacité en termes de calcul : exigences des capteurs et de calculs sont négligeables.

- Sans nécessité d'enregistrement de l'utilisateur : la création du système peut être effectuée hors-ligne et sans connaissance préalable des personnes inspectés.

- Déductible de la biométrie classique : Les traits peuvent être calculés à partir d'images capturées pour l'identification des traits biométrique classiques (par exemple la couleur des yeux à partir d'images de l'iris).

- Non-intrusif : L'acquisition des données est non intrusive ou peut être entièrement imperceptible.

- Possibilité d'identification à distance : L'acquisition des données est réalisable à distance.

- Ne nécessitant pas de la coopération du sujet : On n'a pas besoin strictement du consentement et de la contribution du sujet concerné par la surveillance.

- Préserver la vie privée : Les signatures stockées sont usuellement visuellement accessible à tous.

Récemment, les traits de biométrie douce ont été employées pour améliorer la recherche dans une base de données biométrique, afin de diminuer le temps de calcul nécessaire pour le trait classique biométrique. Jain et al. qui a introduit la biométrie douce, a effectué des études connexes sur l'utilisation de la biométrie douce [12] pour le pré-filtrage et la fusion en combinaison avec des traits biométriques classiques.

Une nouvelle approche est la fusion de la biométrie douce et celle classique pour augmenter les performances d'un système et sa fiabilité. Récemment les systèmes à biométrie douce ont été employées aussi pour [11, 13] la reconnaissance de la personne et pour l'authentification continue des utilisateurs [14].

D'autres études ont évalué les algorithmes d'extraction des traits comme la couleur des yeux [15], la taille [16], la couleur des vêtements [17] ou l'anthropométrie [18]. L'origine ethnique et la couleur de cheveux ont été utilisés dans [19] pour reconnaître et vérifier les identités des clients dans un système biométrique. Les auteurs ont démontré que la biométrie douce augmente la précision de la reconnaissance du système biométrique classique basé seulement sur l'apparence du visage. Dans [20] les auteurs ont démontré qu'il est possible de traiter avec l'extraction en temps réel de la biométrie douce, et que la combinaison de ces traits peut conduire à un niveau acceptable de précision au niveau de la reconnaissance.

Tous ces ouvrages témoignent de deux contributions majeures de la biométrie douce pour les systèmes de reconnaissance : ils augmentent la vitesse des techniques existantes (élagage des solutions les moins probables), et ils améliorent la précision de la reconnaissance (par la prévention des erreurs grossières). Cependant, parce que la permanence de ces traits peut être très faible par rapport à des traits biométriques, les caractéristiques de biométrie douce sont connues pour être mieux exploitée en des systèmes basés sur des sessions, c'est-à-dire sur des systèmes (par exemple de vidéo-surveillance) où l'absence de caractère distinctif devient moins important puisque le domaine du système est un sous-espace connu de toute la population.

B.3 Traits

B.3.1 Biométrie douce pour le corps

La démarche, la taille, le poids corporel et la couleur des vêtements concernent le corps et sont les principaux traits qui peuvent être extraits à distance. Le meilleur critère distinctif est fourni par la première, ce qui explique pourquoi la démarche est parfois considéré comme un trait de biométrie classique.

B.3.1.1 Les mesures anthropométriques

Les études sur les mesures anthropométriques ne sont généralement pas entraînées par l'utilisation biométrique. Tandis qu'au début l'anthropométrie était une technique utilisée dans l'anthropologie physique pour étudier le développement physique de l'espèce humaine; dans nos jours elle est employé dans le design industriel, l'habillement, l'ergonomie et l'architecture afin d'optimiser les produits aux besoins des clients. D'autres études intéressantes sont liées à l'étude des statistiques démographiques, à la façon de surveiller les changements de mode de vie, et de la nutrition, pour suivre les dimensions du corps (épidémie d'obésité, par exemple) [21].

La première application biométrique de l'anthropométrie est due à Alphonse Bertillon. Sa méthode basée sur l'anthropométrie avait comme but la classification pour identifier les criminels, c'est en effet l'un des rares exemples de mesure anthropométrique utilisé comme identifiant biométrique.

Après la contribution historique de Bertillon, l'un des premiers travaux qui a essayé d'estimer les mesures anthropométriques à partir d'images est celui présenté dans [22]. Les auteurs de cet article, utilisent des informations statistiques sur le corps humain, pour établir la correspondance entre un ensemble de points marqués manuellement et les segments qui composent les parties du corps.

Dans une deuxième étape, un ensemble de postures est considéré et enfin la pose et les mesures anthropométriques sont obtenus.

Les résultats sont obtenus en minimisant une fonction de coût appropriée et selon un modèle inspiré par les statistiques du corps humain prélevés pour la recherche médicale.

La commercialisation récente de scanners à ondes millimétriques et de scanners corporels 3D a suscité l'intérêt de la communauté de recherche. Certains travaux ont suggéré que l'idée d'identifier des gens grâce à l'anthropométrie est faisable et ont proposé des solutions. Un exemple est le travail décris dans [23] où les auteurs examinent l'utilité de mesures anthropométriques 1D comme une caractéristique biométrique pour l'identification humaine. Ils analysent 27 mesures de 2144 sujets, en réduisant ces mesures à un ensemble plus petits grâce à des techniques de réduction de dimensions, ils obtiennent une identification de 83% et 94% en utilisant seulement de dix et quinze dimensions.

D'autres travaux intéressants sur les mesures anthropométriques sont présentés dans [24] où la taille, la démarche, et d'autres mesures sont prises en compte pour l'identification des personnes, et dans [25] où des mesures anthropométriques sont estimées à partir des séquences monoculaires calibrées. Avec le suivi des sujets à travers plusieurs caméras, les auteurs estiment la taille, la largeur des épaules, et les rattache avec des caractéristiques spécifiques prévues par la démarche pour effectuer l'identification des personnes.

B.3.1.2 Taille

Même si la taille fait partie des mesures plus générales anthropométriques, nous y consacrons une part de cette section parce que la communauté en vision par ordinateur a exploré profondément son extraction et les applications possibles.

L' estimation de la taille est un sujet déjà mature dans la littérature et il a été exploité à plusieurs reprises. Une des premières approches est présentée dans [16], les auteurs utilisent le contenu de l'image pour calculer les propriétés géométriques des objets qui se trouvent sur le même plan, plus tard, ils peuvent comparer les dimensions des objets. En connaissant la taille des objets donnés dans la scène, ils sont capables de mesurer la taille des personnes dans le champ de la caméra de vision (FOV). L'extension de ce dernier ouvrage, les auteurs de [26] proposent de nouvelles améliorations en utilisant plusieurs mesures et une approche statistique pour éliminer les valeurs aberrantes, en utilisant l'approche proposée, ils arrivent avec une précision de 1 cm pour les sujets à pied dans un scénario sans contrainte.

La mesure précise de la taille a été déjà utilisée en combinaison avec d'autres fonctions afin de suivre les gens à travers des systèmes de caméras multiples, et de permettre l'identification de la même personne dans plusieurs flux vidéo [27]. L' estimation est effectuée via le calcul de la taille de rapport avec le monde réel et les coordonnées estimées en images de la caméra.

La taille est peut-être l'un des traites de biometrie douce les plus utilisés dans la vie réels et elle peut devenir dans certaines situations la preuve d'un crime. Elle est en effet l'un des principaux facteurs utilisés dans la photogrammétrie. Cette technique est aujourd'hui largement utilisée pour estimer les mesures anthropométriques à partir d'images ou de vidéo de surveillance. L' Institut médico-légal des Pays-Bas a effectué une comparaison [28] de deux méthodes pour obtenir des mesures de taille du corps à partir d'images. L'une est basée sur la géométrie projective et l'autre sur la modélisation 3D de la scène du crime. Avec la même caméra et la même configuration des ses paramètres, les auteurs démontrent que les prédictions des deux méthodes sont précis, mais si la position de la caméra change, le premier algorithme devient moins fiable.

Par ailleurs, la reconstruction 3D de l'environnement peut être utile à ce genre d'analyse pour simplifier considérablement l' extraction de mesures. La possibilité d'utiliser une telle technique est étudiée dans [29] où les auteurs utilisent repères au sein de la scène pour permettre la collecte automatique de les tailles des sujets.

B.3.1.3 Poids

Depuis le début, le poids a été introduite dans la liste des traits de biométrie douce [12]. Toutefois, il n'a pas été pleinement explorées autant que les autres traits biométriques.

Un secteur où le poids est considéré comme une caractéristique importante est représenté par des études médicales, où le principal intérêt est représenté par la capacité d'extraction visuelle et la fiabilité d'estimation des opérateurs en cas de situations d'urgence où il n'y a pas de possibilité d'utiliser des balances comme dans [2, 30, 31].

D'autres intérêts sont représentés par l'utilisation du poids en tant que élément important qui permet de suivre l'état de santé du corps [32]. En outre, une branche dés études médicales explore l'aspect médico-légal de l' estimation du poids de manière à récupérer des informations à partir de traces latentes qui aident à reconnaître les victimes ou les suspects d'un acte criminelle [33].

Le seul article qui utilise le poids faisant directement référence à un trait de biométrie douce est [34], où les auteurs utilisent une balance pour peser les clients d'un système de reconnaissance d' empreintes digitales. En exploitant le poids et les mesures de graisse corporelle, les auteurs réduisent le taux d'erreur totale du système de 2,4%.

Un autre travail [5] a considéré le poids comme une valeur discrète défini visuellement par les sujets participant à une expérience psycho-visuel. Toutefois, les valeurs utilisées (très mince, mince, moyen, gras, très gras) montrent que, plutôt que le poids lui-même, la description se réfère à la façon dont la graisse est répartie sur l'organisme. C'est-à-dire les utilisateurs décrit la structure du corps des sujets plutôt que de leur masse corporelle. Une expérience similaire est signalée sur [35] où les auteurs proposent d'autres fonctions aux côtés du poids. En outre, le travail indique l'importance de ce trait dans le cas de témoins oculaires.

Certains tests effectués par [36, 37] impliquent la créatione d'un modèle 3D humain à partir d'un nuage de points obtenu dans le premier cas par une caméra RGB-D, et dans le second cas par une série de caméras stéréo. Dans ce cas, l'extraction du poids est simple si l'on considère la densité moyenne du corps humain. Alors que dans le premier travail du poids (et le sexe) l'estimation est un effet secondaire intéressant, dans le second cas, les auteurs délibérément essaient d'extraire ces informations.

B.3.1.4 Genre

La reconnaissance du genre a déjà été largement explorée dans les travaux psychologie sociale et cognitive dans le contexte de l'analyse du visage et celle du corps. D'un point de vue de traitement d'image, le sujet offre une pléthore d'approches. Les derniers efforts emploient une sélection de traits biométriques fusionnés pour en déduire des informations sur le genre. Par exemple, dans [38] des "images d'énergie" de la démarche et les traits du visage sont fusionnés pour effectuer la reconnaissance du genre.

Les auteurs de [39] segmentent la silhouette humaine en sept composantes (tête, bras, tronc, cuisses, avant-jambe, dos-jambe, et pieds). Ils étudient l'efficacité des sept composants de la démarche dans plus de 500 expériences différentes sur l'identification des personnes et la reconnaissance du genre. En moyenne, de bonnes performances sont atteintes par des caractéristiques de la démarche discriminantes sur un ensemble de données composé de 1870 séquences de 122 sujets.

Une autre approche de la reconnaissance du genre est suivie par [40]; dans ce cas des images fixes sont analysées afin de fournir des informations de base sur le genre et sur la pose. Les auteurs proposent un cadre commun pour l'estimation de la pose et du genre. Les deux systèmes sont basés sur des caractéristiques inspirés par la biologie, en combinaison avec des techniques d'apprentissage "manifold” des performances maximales d’environ 80% sont obtenus sur une base de données publique.

La différence entre les performances de ces deux approches se trouve dans le contenu enrichi d’information qui appartient aux séquences de la démarche qui représentent au mieux la forme humaine.

B.3.1.5 Démarche

La démarche est un modèle complexe qui implique non seulement des paramètres anthropométriques, mais des informations comportementales. Parmi tous les traits de biométrie douce, il est l’un des plus explorée, et en raison de son caractère très distinctif, il est discuté comme étant en réalité un trait de biométrie classique. Une des premières expériences (1973) sur l’analyse de la démarche est présentée dans [41], où l’auteur utilise des réflecteurs attachés aux articulations du corps humain pour enregistrer un schéma de marche des sujets. L’auteur montre comment les observateurs peuvent reconnaître la marche des personnes qui leur sont familières seulement par les traces qu’ils laissent grâce aux réflecteurs tout en marchant. Depuis 1970, de nombreux autres auteurs se sont intéressés au thème de la reconnaissance automatique de la démarche. Dans [42] une signature spatio-temporelle est extraite par la silhouette en mouvement, et puis une analyse en composantes principales est utilisée pour éliminer l’information non pertinente. Enfin, des techniques de classification supervisée sont utilisées dans l’espace en dimensions réduites pour la classification. Afin de fournir plus de puissance discriminative, les caractéristiques structurelles et celles comportementales de la démarche sont capturés.

Un autre travail intéressant est proposé dans [5], où la démarche est choisie comme une caractéristique biométrique primaire à être couplée avec des “biométrie sémantiques”, qui semble être un concept très similaire à la biométrie douce. En utilisant ANOVA les auteurs décrivent d’abord les traits sémantiques les plus importants. Après, ils fusionnent les résultats de la signature générée par la marche avec celui généré par l’information sémantique de façon à identifier les utilisateurs du système biométrique. D’autres moyens de réaliser l’identification humaine avec l’analyse de la démarche sont basés sur la silhouette humaine et sur des systèmes basés sur des modèles comme dans [43, 44].

L’analyse de la démarche est non seulement utilisée pour identifier les personnes ou le genre, mais elle est activement utilisée également dans le domaine médical. Dans ce cas, il est exploité pour comprendre des modèles pour les patients pathologiquement anormales, comme dans [45]. Même si dans le domaine médical l'utilisation de marqueurs a été largement exploitée, ces derniers temps certaines études ont commencé à impliquer de nouvelles techniques comme la vision par ordinateur, ou des nouveaux capteurs comme des accéléromètres [46, 47] pour analyser ce trait.

B.3.2 Biométrie douce du visage

Des travaux sur la biométrie douce ont été effectuées essentiellement dans le but du pré-traitement des données. Dans le but de la reconnaissance de visage pour l'identification des personnes, la détection de barbe, par exemple, a été utilisée pour améliorer les résultats de la reconnaissance, en considérant pas la partie d'image avec la présence de la barbe.

B.3.2.1 Couleur

Les traits faciaux de biométrie douce basées sur les couleurs (comme la couleur des yeux, la peau et les cheveux) sont les parties les plus évidents du visage, mentionnés principalement par l'homme en décrivant le visage des gens, et qui apporte d'importants indices sur l'origine ethnique [48]. Les principaux défis pour la classification de la couleur de la peau sont d'une part la faiblesse de l'écart entre différentes couleurs dans l'espace de couleur, et en conséquence, d'autre part la dépendance élevé par rapport à l'éclairage. Cela est décrit dans divers documents comme par exemple [49].

La couleur des cheveux est détectée par des techniques similaires à celles de la couleur de la peau, mais pour ce trait il y a largement plus de catégories dispersées. Le travail de [50] résume une méthode de détection de la tête humaine basée sur les couleurs, les auteurs proposent l'utilisation de modèles de densité gaussien pour décrire la distribution de la couleur des cheveux. Dans [51] la logique floue est utilisée pour détecter les visages dans des images en couleur, ici deux modèles flous décrient la couleur de la peau et la couleur des cheveux, respectivement. Contrairement à la detection de la couleur de la peau et des cheveux, la detection de la couleur des yeux est un sujet de recherche relativement nouveau. Probablement parce que 90% des humains possèdent des yeux bruns ce trait de couleur ne semble pas très exploré. L'avantage de la classification par la couleur des yeux est la disponibilité de toutes les informations nécessaires dans les images utilisées pour l'analyse de l'iris, en d'autres termes, la couleur de l'iris pourrait être considéré comme un effet secondaire. Un travail sur la fusion entre la texture et la couleur de l'iris peut être trouvée dans [52], où les auteurs fusionne la couleur de l'iris avec les empreintes digitales pour assurer une amélioration des performances en ce qui concerne les systèmes uni-modales. S'appuyant sur cet effet secondaire de l'analyse de l'iris, dans [53] les auteurs utilisent avec succès la couleur de l'iris comme un indice d'une méthode d'élagage de couleur qui améliore la vitesse du système.

B.3.2.2 Sexe

Bien que le genre est une caractéristique qui affecte principalement l'apparence du corps, de nombreux travaux tentent d'extraire cette information du visage. Par exemple, dans [54] les auteurs proposent un système combiné pour le sexe et la reconnaissance des expressions par la modélisation du visage en utilisant un modèle d'apparence active (AAM), extraction de caractéristiques et de fonctions de base, enfin linéaire, polynomiale et radiale sur la base des machines à vecteurs de support pour la classification. Le travail décris dans [55] propose l'utilisation d'Adaboost sur plusieurs classificateurs faibles, appliqué sur des images à faible résolution d'échelle de gris avec de bons résultats. Les auteurs de [56] présente un système multi-modale de reconnaissance du sexe, basé sur l'image du visage et sur les mouvements de la tête et de la bouche.

Une étude intéressante [48] met en relation l'ethnie avec le genre. Alors que les gens sont généralement enclins à obtenir de meilleures performances lorsque vous essayez d'estimer le sexe des gens sachant leurs appartenances ethniques, les ordinateurs ne semble pas être affectée par cet effet. Les auteurs concluent que les systèmes automatiques basées sur des caractéristiques de l'image (ainsi que des techniques basées sur des pixels) sont en mesure de généraliser le problème de reconnaissance du genre sans tenir compte de l'information ethnique.

B.3.2.3 Détection de barbe et moustache

Un autre trait typique du visage est fourni par la présence de barbe ou de moustache, en particulier pour la catégorisation des hommes. Généralement considéré comme un obstacle par les algorithmes de reconnaissance de visage à cause de sa variabilité élevée dans le temps, l'étude de la détection de barbe et des moustaches est principalement axé vers une approche de détection et compensation.

Comme exemple principal on peut citer [57] où les auteurs montrent un algorithme pour retirer la barbe à partir d'images de personnes. Le système est conçu en utilisant la notion de similitudes structurelles et des transformations de coordonnées.

B.3.2.4 Age

L'age est principalement liée à la structure du visage et joue un rôle important pour des systèmes de longue durée aptes au travail sur le visage ou sur le corps. En plus, c'est un domaine difficile et relativement nouveau. Une étude intéressante sur les changements du visage au fil du temps est celui de [58], qui inclut un étude biométrique, médico-légales et anthropologiques, et discute encore des techniques pour simuler le vieillissement dans les images de visage. Dans [59] les auteurs distinguent les enfants des adultes sur la base du rapport de taille/iris. Dans [60] des régions de la peau du visage des femmes de race blanche sont analysés, un modèle de régression est construit pour prédire l'ordre chronologique et l'âge. Ils découvrent que le contour des yeux et de l'uniformité de la couleur de la peau sont les principaux attributs liés à l'âge.

B.3.2.5 Ethnie

La reconnaissance ethnique est très débattue en raison de ses implications éthiques et sociologiques. Néanmoins, ce caractère peut-être très pertinent pour la reconnaissance de visage. Dans le cadre de l'ethnie une classification définie de manière unique est une tâche difficile et importante. Pour la reconnaissance des visages asiatiques/non asiatiques, dans [61] des techniques d'apprentissage automatiques sont appliquées à une analyse LDA et à une analyse multi-échelle.

Les auteurs de [62] ont conçu une approche de reconnaissance de l'ethnie basée sur la transformation des ondelettes de Gabor, combinée avec l'échantillonnage à rétine pour l'extraction de caractéristiques clé du visage. Enfin un classificateur SVM est utilisé pour la classification de l'ethnie fournit de très bons résultats, même en présence de différentes conditions d'éclairage.

B.3.2.6 Mesures du visage

Différemment de ce qui se passe avec les mesures anthropométriques, de manière générale les mesures visage sont considérés comme un élément biométrique classique, plusieurs fois utilisées pour l'identification des personnes [63].

Des études ultérieures continuent à employer des mesures du visage, même en étendant le concept à l'analyse de données 3D [64]. Des travaux récents sur la biométrie faciale douce ont été effectués sur les cicatrices, sur les marques et sur les tatouages par les auteurs de [65].

D'autres travaux très récents emploient des mesures du visage et d'autres informations pour étudier la corrélation entre cet ensemble de caractéristiques, et un indice d'esthétique facial évalué sur une grande population de personnes [66].

B.3.3 Biométrie douce et accessoires

La nouvelle définition de biométrie douce qui nous avons introduite, permet l'inclusion d'accessoires aux traits déjà mentionnés. Les accessoires peuvent en effet être liés à des caractéristiques personnelles (comme des problèmes de vue en cas de lunettes), ou des choix personnels (comme ornement dans le cas des bijoux et des vêtements). Un premier exemple est la détection de la couleur des vêtements. Selon la définition de la biométrie douce ces caractéristiques peuvent être ajoutés à la liste des traits mentionnés.

B.3.3.1 La couleur et la classification des vêtements

Plusieurs travaux ont porté à l'utilisation d'informations de la couleur des vêtements pour ré-identifier les personnes dans des scénarios de vidéo-surveillance. Une des façons pour discriminer entre différentes cibles est représentée par des méthodes basées sur des histogrammes [67]. Les systèmes de récupération du contenu (surtout sur Internet) ont montré comment ces techniques sont bien adaptés pour récupérer des images similaires, mais en même temps elles sont fortement influencées par les changements dans l'apparence de l'objet et par l'illumination [68]. Un autre aspect positif des méthodes de type histogramme [69] c'est qu'elles sont simples et rapides à calculer. Pour ces raisons, ces publications ont été pris comme référence dans [70] où les auteurs présentent un ensemble de méthodes qui présentent des performances prometteuses compte tenu à la fois de la taille de la base de données, vet de la simplicité de mise en oeuvre des techniques elles-mêmes. Plus tard, D'Angelo et al. [17] ont amélioré la performance d'une méthode basée sur des histogrammes de couleur probabiliste comme nouveau type de descripteur.

Un résultat intéressant, encore préliminaire, sur la catégorisation des vêtements est montré dans [71] où les auteurs sont en mesure de classer les vêtements dans différentes catégories comme un effet secondaire de l'analyse de la silhouette du corps 2D.

2.3.3.2 Détection des lunettes

Comme pour la barbe et les moustaches, l'état de l'art sur la détection des lunettes explore les moyens d'éliminer cet attribut de manière à améliorer les résultats en reconnaissance de visage. L'une des premières publications pour la détection des lunettes a été réalisée par [72], ils exploitent la détection de contours sur une image en niveaux de gris pré-traitée pour améliorer certaines caractéristiques des lunettes. Certains zones du visage sont observées et un indicateur de la présence de verres est cherché. La partie qui indique la mieux la présence des lunettes se trouve dans la région du nez entre les yeux. Une méthode 3D pour détecter les lunettes est présenté dans [73], où des images 3D sont obtenues par un système de vision trinoculaire. Une autre approche pour l'extraction des lunettes est employée dans [74], où un modèle de visage est établi sur la base de la triangulation de Delaunay. Jusqu'à maintenant, les meilleurs résultats sur la détection des lunettes sont réalisés sur des images thermiques [75], dans le processus de fusion de données, les lunettes qui bloquent l'énergie thermique sont détectées à partir d'images thermiques et remplacées par un modèle des yeux.

Bibliography

[1] B. Shah, K. Sucher, and C. B. Hollenbeck. Comparison of ideal body weight equations and published height-weight tables with body mass index tables for healthy adults in the United States. *Nutrition in clinical practice*, 21(3):312–319, 2006.

[2] S. Menon and A. M. Kelly. How accurate is weight estimation in the emergency department? *Emergency Medicine Australasia*, 17(2):113–116, 2005.

[3] G. Heinz, L. J. Peterson, R. W. Johnson, and C. J. Kerk. Exploring Relationships in Body Dimensions. *Journal of Statistics Education*, 11(2), 2003.

[4] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(12):1349–1380, 2000.

[5] S. Samangooei, M. Nixon, and B. Guo. The Use of Semantic Human Description as a Soft Biometric. In *Biometrics: Theory, Applications, and Systems*, September 2008.

[6] C. A. Nelson. The development and neural bases of face recognition. *Infant and Child Development*, 10(1–2):3–18, 2001.

[7] A. J. O'Toole, D. A. Roark, and H. Abdi. Recognizing moving faces: A psychological and neural synthesis. *Trends in Cognitive Sciences*, 6(6):261–266, 2002.

[8] A. Bertillon. Instructions for taking descriptions for the identification of criminals and others by the means of anthropometric indications. *Chicago: American Bertillon Prison Bureau*, 1889.

[9] G. Pavlich. The subjects of criminal identification. *Punishment & Society*, 11(2):171, 2009.

[10] R. D. Olsen. A Fingerprint Fable: The Will and William West Case. Identification News, November 1987.

[11] A. Dantcheva, C. Velardo, A. D'Angelo, and J.-L. Dugelay. Bag of soft biometrics for person identification: new trends and challenges. *Multimedia Tools and Applications*, 2, 2010.

[12] A. K. Jain, S. C. Dass, and K. Nandakumar. Can soft biometric traits assist user recognition? *Proceedings of SPIE*, 5404:561–572, 2004.

[13] A. Dantcheva, J.-L. Dugelay, and P. Elia. Person recognition using a bag of facial soft biometrics. In *Proceedings of MMSP*, 2010.

[14] K. Niinuma, U. Park, and A. K. Jain. Soft biometric traits for continuous user authentication. *Transactions on information forensics and security*, 5(4), 2010.

[15] A. Ross, K. Nandakumar, and A. K. Jain. *Handbook of multibiometrics*. Springer-Verlag New York Inc, 2006.

[16] A. Criminisi, I.D. Reid, and A. Zisserman. Single view metrology. *International Journal of Computer Vision*, 40(2):123–148, 2000.

[17] A. D'Angelo and J.-L. Dugelay. People re-identification in camera networks based on probabilistic color histograms. In *Visual Information Processing and Communication, SPIE Electronic Imaging, Vol. 7882, 23-27 January, 2011, San Francisco, CA, USA*, 2011.

[18] D. Adjeroh, D. Cao, M. Piccirilli, and A. Ross. Predictability and correlation in human metrology. In *Proceedings of WIFS*, 2010.

[19] G. L. Marcialis, F. Roli, and D. Muntoni. Group-specific face verification using soft biometrics. *Journal of Visual Languages and Computing*, 20(2):101–109, 2009.

[20] A.K. Jain, S.C. Dass, and K. Nandakumar. Soft biometric traits for personal recognition systems. *Biometric Authentication*, pages 1–40, 2004.

[21] Center for Disease Control CDC and Prevention. National Health and Nutrition Examination Survey.

[22] C. Barron and I. A. Kakadiaris. Estimating anthropometry and pose from a single image. In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, volume 1, pages 669–676. IEEE, 2000.

[23] D. B. Ober, S. P. Neugebauer, and P. A. Sallee. Training and feature-reduction techniques for human identification using anthropometry. In *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*, pages 1–8. IEEE, 2010.

[24] B.A. Chiraz, R. Cutler, and L. Davis. View-invariant estimation of height and stride for gait recognition. In *Proceedings of ICBA*, volume 2359/2006, pages 155–167, 2002.

[25] C. BenAbdelkader and L. Davis. Estimation of anthropomeasures from a single calibrated camera. In *Proceedings of Automatic Face and Gesture Recognition*, pages 499–504, 2006.

[26] S.H. Lee, T.E. Kim, and J.S. Choi. A single-view based framework for robust estimation of heights and positions of moving people. *ICCE: International Conference on Consumer Electronics*, pages 503–504, 2010.

[27] C. Madden and M. Piccardi. Height measurement as a session-based biometric for people matching across disjoint camera views. In *Proceedings of Image and Vision Computing, New Zealand*, pages 282–286, 2005.

[28] G. Edelman, I. Alberink, and B. Hoogeboom. Comparison of the Performance of Two Methods for Height Estimation. *Journal of Forensic Sciences*, 55(2):358–365, 2010.

[29] D. De Angelis and R. Sala. New method for height estimation of subjects represented in photograms taken from video surveillance systems. *International Journal of Legal Medicine*, 121:489–492, 2007.

[30] T.R. Coe, M. Halkes, K. Houghton, and D. Jefferson. The accuracy of visual estimation of weight and height in pre-operative supine patients. *Anaesthesia*, 54(6):582–586, 2002.

[31] W. II L. Hall, G. L. Larkin, M. J. Trujillo, J. L. Hinds, and K. A. Delaney. Errors in weight estimation in the emergency department: comparing performance by providers and patients. *Journal of Emergency Medicine*, 27(3):219–224, 2004.

[32] I.I. Korhonen, J. Pärkkä, and M. Van Gils. Health monitoring in the home of the future. *IEEE Engineering in Medicine and Biology Magazine*, 22(3):66–73, 2003.

[33] K. Krishan. Establishing correlation of footprints with body weight–Forensic aspects. *Forensic science international*, 179(1):63–69, 2008.

[34] H. Ailisto, E. Vildjiounaite, M. Lindholm, S. M. Mäkelä, and J. Peltola. Soft biometrics–combining body weight and fat measurements with fingerprint biometrics. *Pattern Recognition Letters*, 27(5):325–334, 2006.

[35] M. D. MacLeod, J. N. Frowley, and J. W. Shepherd. Whole body information: Its relevance to eyewitnesses. *Adult eyewitness testimony: Current trends and developments*, pages 125–143, 1994.

[36] A. Weiss, D. Hirshberg, and M. J. Black. Home 3D Body Scans from Noisy Image and Range Data. In *ICCV: International Conference on Computer Vision 2011*, 2011.

[37] K. Pirker, M. Rüther, H. Bischof, F. Skrabal, and G. Pichler. Human Body Volume Estimation in a Clinical Environment. *AAPR/OAGM: Visual Learning*, 2009.

[38] S. Caifeng, G. Shaogang, and P. W. McOwand. Learning gender from human gaits and faces. In *Proceedings of AVSS*, pages 505–510, 2007.

[39] X. Li, S. J. Maybank, S. Yan, D. Tao, and D. Xu. Gait components and their application to gender recognition. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 38(2):145–155, 2008.

[40] G. Guo, G. Mu, and Y. Fu. Gender from body: A biologically-inspired approach with manifold learning. *Computer Vision–ACCV 2009*, pages 236–245, 2010.

[41] G. Johansson. Visual perception of biological motion and a model for its analysis. *Perceiving events and objects*, 3, 1973.

[42] L. Wang, T. Tan, H. Ning, and W. Hu. Silhouette analysis-based gait recognition for human identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(12):1505–1518, 2003.

[43] S. A. Niyogi and E. H. Adelson. Analyzing and recognizing walking figures in XYT. In *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on*, pages 469–474. IEEE, 1994.

[44] J. Han and B. Bhanu. Statistical feature fusion for gait-based human recognition. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 2, pages II—-842. IEEE, 2004.

[45] D. A. Winter. Human balance and posture control during standing and walking. *Gait & Posture*, 3(4):193–214, 1995.

[46] G. I. Zdragkas and J. N. Avaritsiotis. Gait analysis and automatic gait event identification using accelerometers. In *BioInformatics and BioEngineering, 2008. BIBE 2008. 8th IEEE International Conference on*, pages 1–6, 2008.

[47] C. Zong, M. Chetouani, and A. Tapus. Automatic gait characterization for a mobility assistance system. In *Control Automation Robotics Vision (ICARCV), 2010 11th International Conference on*, pages 473–478, 2010.

[48] G. Farinella and J.-L. Dugelay. Demographic Classification: Do Gender and Ethnicity Affect Each Other? In *ICIEV: International Conference on Informatics, Electronics and Vision*, 2012.

[49] A. Hadid, M. Pietikaeinen, and B. Martinkauppi. Color-based face detection u sing skin locus model and hierarchical filtering. In *Proceedings of ICPR*, pages 196–200, 2002.

[50] Z. Sun, G. Bebis, X. Yuan, and S. J. Louis. Genetic feature subset selection for gender classification: a comparison study. In *Proceedings of WACV*, page 165, 2002.

[51] S. Gutta, J. R. J. Huang, P. Jonathon, and H. Wechsler. Mixture of experts for classification of gender, ethnic origin, and pose of human faces. *Transactions of Neural Networks*, 11:948–960, 2000.

[52] R. Zewail, A. Elsafi, M. Saeb, and N. Hamdy. Soft and hard biometrics fusion for improved identity verification. In *Proceedings of MWSCAS*, 2004.

[53] N. B. Puhan and N. Sudha. A novel iris database indexing method using the iris color. In *Proceedings of ICIEA*, pages 1886–1891, 2008.

[54] Y. Saatci and C. Town. Cascaded classification of gender and facial expression. In *Proceedings of FG*, pages 393–400, 2006.

[55] S. Baluja and H. A. Rowley. Boosting sex identification performance. *International Journal of Computer Vision*, 71:111–119, 2006.

[56] F. Matta, U. Saeed, C. Mallauran, and J.-L. Dugelay. Facial gender recognition using multiple sources of visual information. In *Proceedings of MMSP*, pages 785–790, 2008.

[57] N. Kaushik and A. Mittal. A novel beard removal method based on structural similarity and coordinate transformations. In *Proceedings of INDICON*, pages 1–6, 2006.

[58] E. Patterson, A. Sethuram, M. Alber, K. Ricanek, and M. King. Aspects of age variation in facial morphology affecting biometrics. In *Proceedings of BTAS*, 2007.

[59] H. Weda and M. Barbieri. Automatic children detection in digital images. In *Proceedings of ICME*, 2007.

[60] A. Nkengne, C. Bertin, G. N. Stamatas, A. Giron, A. I. N. Rossi, and B. Fertil. Influence of facial skin attributes on the perceived age of Caucasian women. *Journal of the European Academy of Dermatology and Venerology*, 22:982–991, 2008.

[61] X. Lu and A. K. Jain. Ethnicity identification from face images. In *Proceedings of SPIE*, volume 5404, pages 114–123, 2004.

[62] S. Hosoi, E. Takikawa, and M. Kawade. Ethnicity estimations with facial images. In *Proceedings of FG*, pages 195–200, 2004.

[63] M. Nixon. Eye spacing measurement for facial recognition. In *Proceedings of SPIE*, volume 575, pages 279–285, 1985.

[64] J. Carnicky and J. D. Chorvat. Three-dimensional measurement of human face with structured-light illumination. *Measurement science review*, 6:1–4, 2006.

[65] J.-E. Lee, A. K. Jain, and R. Jin. Scars, marks and tattoos (SMT): Soft biometric for suspect and victim identification. In *Proceedings of BCC*, 2008.

[66] A. Dantcheva and J.-L. Dugelay. Female facial aesthetics based on soft biometrics and photo-quality. In *ICME 2011, IEEE International Conference for Multimedia and Expo, July 11-15, 2011, Barcelona, Spain*, Barcelona, SPAIN, 2011.

[67] O. J. Zeeshan, O. Javed, Z. Rasheed, K. Shafique, and M. Shah. Tracking across multiple cameras with disjoint views. *ICCV: International Conference on Computer Vision*, pages 952–957, 2003.

[68] J. Huang, S. R. Kumar, M. Mitra, R. Zabih, and W. J. Zhu. Image indexing using color correlograms. *CVPR: Computer Vision and Pattern Recognition*, pages 762–768, 1997.

[69] V. Kettnaker and R. Zabih. Counting people from multiple cameras. *Multimedia Computing and Systems*, pages 267 –271, 1999.

[70] D. Gray, S. Brennan, and H. Tao. Evaluating Appearance Models for Recognition, Reacquisition, and Tracking. In *PETS: Performance Evaluation for Tracking and Surveillance*, page 5, 2007.

[71] P. Guan, O. Freifeld, and M. Black. A 2d human body model dressed in eigen clothing. *ECCV: European Conference on Computer Vision*, pages 285–298, 2010.

[72] X. Jiang, M. Binkert, B. Achermann, and H. Bunke. Towards detection of glasses in facial images. *Pattern Analysis and Applications*, 3:9–18, 2000.

[73] H. Wu, G. Yoshikawa, T. Shioyama, S. Lao, and M. Kawade. Glasses frame detection with 3D Hough transform. In *Proceedings of ICPR*, volume 16, pages 346–349, 2002.

[74] Y. Xiao and H. Yan. Extraction of glasses in human face images. In *Proceedings of ICBA*, pages 1–23. Springer, 2004.

[75] J. Heo, S. G. Kong, B. R. Abidi, and M. A. Abidi. Fusion of visual and thermal signatures with eyeglass removal for robust face recognition. In *Proceedings of CVPRW*, page 122, 2004.

[76] S. Samangooei and M. Nixon. Performing content-based retrieval of humans using gait biometrics. *Semantic Multimedia*, pages 105–120, 2008.

[77] P. J. Van Koppen and S. K. Lochun. Portraying perpetrators; the validity of offender descriptions by witnesses. *Law and Human Behavior*, 21(6):661–685, 1997.

[78] S. Samangooei. *Semantic biometrics*. PhD thesis, University of Southampton, School of Electronics and Computer Science, 2010.

[79] C. Velardo and J.-L. Dugelay. Weight estimation from visual body appearance. In *BTAS: Biometrics Theory Applications and Systems*, pages 1–6. IEEE, 2010.

[80] K. Mikolajczyk, C. Schmid, and A. Zisserman. Human detection based on a probabilistic assembly of robust part detectors. *Computer VisionECCV 2004*, I:69–82, 2004.

[81] C. Velardo and J.-L. Dugelay. Real time extraction of body soft biometric from 3D videos. In *ACM MM: ACM International Conference on Multimedia*, pages 3–4, 2011.

[82] C. Velardo, J.-L. Dugelay, M. Paleari, and P. Ariano. Building the space scale or how to weigh a person with no gravity. In *ESPA: Emerging Signal Processing Applications*, volume 1, 2012.

[83] R. B. Rusu and S. Cousins. 3D is here: Point Cloud Library (PCL). In *IEEE International Conference on Robotics and Automation (ICRA)*, 2011.

[84] PrimeSense. OpenNI library. [Last seen: 9th March 2012] http://www.openni.org/.

[85] A. Criminisi, I. D. Reid, and A. Zisserman. Single View Metrology. In *ICCV*, pages 434–441, 1999.

[86] I. Duyar and C. Pelin. Body height estimation based on tibia length in different stature groups. *American journal of physical anthropology*, 122(1):23–27, 2003.

[87] M.C. De Mendoca. Estimation of height from the length of long bones in a Portuguese adult population. *American journal of physical anthropology*, 112(1):39–48, 2000.

[88] S. Gutta and H. Wechsler. Gender and ethnic classification of human faces using hybrid classifiers. In *Proceedings of IJCNN*, volume 6, pages 4084–4089, 1999.

[89] H. C. Lian and B. L. Lu. Multi-view gender classification using local binary patterns and support vector machines. *Advances in Neural Networks-ISNN 2006*, pages 202–209, 2006.

[90] J. Ylioinas, A. Hadid, and M. Pietikäinen. Combining contrast information and local binary patterns for gender classification. *Image Analysis*, pages 676–686, 2011.

[91] A. O. Balan, L. Sigal, M. J. Black, J. E. Davis, and H. W. Haussecker. Detailed human shape and pose from images. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007.

[92] P. Meunier and S. Yin. Performance of a 2d image-based anthropometric measurement and clothing sizing system. *Applied Ergonomics*, 31(5):445–451, 2000.

[93] P. Guan, A. Weiss, A.O. Balan, and M.J. Black. Estimating human shape and pose from a single image. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 1381–1388. IEEE, 2009.

[94] L. Sigal, A. Balan, and M. Black. Combined discriminative and generative articulated pose and non-rigid shape estimation. In *NIPS*, pages 1337–1344, 2008.

[95] V. Tillmann and P. E. Clayton. Diurnal variation in height and the reliability of height measurements using stretched and unstretched techniques in the evaluation of short-term growth. *Annals of Human Biology*, 28(2):195–206, 2001.

[96] G. S. Daniels. The “Average Man” ? Technical report, Air Force Aerospace Medical Research Lab Wright-Patterson, 1952.

[97] R. D. Green and L. Guan. Quantifying and recognizing human movement patterns from monocular video images-part ii: applications to biometrics. *Circuits and Systems for Video Technology, IEEE Transactions on*, 14(2):191–198, 2004.

[98] A. Godil, P. Grother, and S. Ressler. Human identification from body shape. In *3-D Digital Imaging and Modeling, 2003*, pages 386–392, 2003.

[99] J. L. Wayman. Large-scale Civilian Biometric Systems-Issues and Feasibility. In *Proceedings of Card Tech/Secur Tech ID*, volume 732. Card Tech / Secur Tech ID, 1997.

[100] A. Mhatre, S. Palla, S. Chikkerur, and V. Govindaraju. Efficient search and retrieval in biometric databases. In *SPIE Defence and Security Symposium*, volume 5779, pages 265–273. Citeseer, 2005.

[101] K. Delac, M. Grgic, and S. Grgic. Independent comparative study of PCA, ICA, and LDA on the FERET data set. *International Journal of Imaging Systems and Technology*, 15(5):252–260, 2005.

[102] R. Dass and S. Pal. Challenges of Identity Management: A Context in Rural India. In *Information Security and Digital Forensics*, volume 41 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pages 172–183. Springer Berlin Heidelberg, 2010.

[103] N. Poh and S. Bengio. Can chimeric persons be used in multimodal biometric authentication experiments? In Steve Renals and Samy Bengio, editors, *Machine Learning for Multimodal Interaction*, volume 3869 of *Lecture Notes in Computer Science*, pages 87–100. Springer, 2006.

[104] K. M. Robinette, H. Daanen, and E. Paquet. The CAESAR project: a 3-D surface anthropometry survey. In *3-D Digital Imaging and Modeling, 1999. Proceedings. Second International Conference on*, pages 380–386, 1999.

[105] S. Fleck, F. Busch, P. Biber, and W. Straber. 3d surveillance a distributed network of smart cameras for real-time tracking and its visualization in 3d. In *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW'06. Conference on*, page 118. IEEE, 2006.

[106] M. M. Trivedi, T. L. Gandhi, and K. S. Huang. Distributed interactive video arrays for event capture and enhanced situational awareness. *Intelligent Systems, IEEE*, 20(5):58–66, 2005.

[107] M. Baklouti, M. Chamfrault, M. Boufarguine, and V. Guitteny. Virtu4D: A dynamic audio-video virtual representation for surveillance systems. In *Signals, Circuits and Systems (SCS), 2009 3rd International Conference on*, pages 1–6. IEEE, 2009.

[108] C. F. Shu, A. Hampapur, M. Lu, L. Brown, J. Connell, A. Senior, and Y. Tian. IBM smart surveillance system (S3): a open and extensible framework for event based surveillance. In *Advanced Video and Signal Based Surveillance, 2005. AVSS 2005. IEEE Conference on*, pages 318–323. IEEE, 2005.

[109] F. Dufaux and T. Ebrahimi. Scrambling for video surveillance with privacy. In *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW'06. Conference on*, page 160. IEEE, 2006.

[110] M. Upmanyu, A. M. Namboodiri, K. Srinathan, and C. V. Jawahar. Efficient privacy preserving video surveillance. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 1639–1646. IEEE, 2009.

[111] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*, volume 2. Cambridge Univ Press, 2000.

[112] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893. Ieee, 2005.

[113] L. Li, W. Huang, I. Y. H. Gu, and Q. Tian. Foreground object detection from videos containing complex background. In *Proceedings of the eleventh ACM international conference on Multimedia*, pages 2–10. ACM, 2003.

[114] R. Chellappa and P. Turaga. Recent advances in age and height estimation from still images and video. In *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*, pages 91–96. IEEE, 2011.

[115] Experiments on color naming. [Last seen: 9th March 2012] http://xkcd.com/color/rgb/.

[116] A. Dantcheva, C. Velardo, and J.-L. Dugelay. Bag of soft biometrics for person identification. *Multimedia Tools and Applications*, pages 1–39, 2011.

[117] S. Bak, E. Corvee, F. Brémond, M. Thonnat, et al. Person re-identification using spatial covariance regions of human body parts. *Proc. Advanced Video and Signal based Surveillance*, pages 435–440, 2010.

[118] M. Pollan. Big Food vs. Big Insurance. The New York Times, September 2009.

[119] A. E. Matthews. Children and obesity: a pan-European project examining the role of food marketing. *European journal of public health*, 18(1):7–11, February 2008.

[120] J.M. Lu and M.J.J. Wang. Automated anthropometric data collection using 3D whole body scanners. *Expert Systems with Applications*, 35(1-2):407–414, 2008.

[121] M. P. Pai and F. P. Paloucek. The Origin of the "Ideal" Body Weight Equations. *The Annals of Pharmacotherapy*, 34:1066–1069, September 2000.

[122] D. Crawford and K. Campbell. Lay definitions of ideal weight and overweight. *International journal of obesity*, 23(7):738–745, 1999.

[123] Max Planck Gesellshaft. Bmi Visualizer. [Last seen: 9th March 2012] http://www.bmivisualizer.com.

[124] C. L. Kien and F. Ugrasbul. Prediction of daily energy expenditure during a feeding trial using measurements of resting energy expenditure, fat-free mass, or Harris-Benedict equations. *The American journal of clinical nutrition*, 80(4):876, 2004.

[125] D. Frankenfield, L. Roth-Yousey, and C. Compher. Comparison of predictive equations for resting metabolic rate in healthy nonobese and obese adults: a systematic review. *Journal of the American Dietetic Association*, 105(5):775–789, 2005.

[126] W. Harwood. CBS NEWS: Current Space Demographics, April 2010. [Last seen: 9th March 2012] www.cbsnews.com/network/news/space/democurrent.html.

[127] W. Thornton. Anthropometric changes in weightlessness. *Webb Assoc. Anthropometric Source Book,*, 1, 1978.

[128] H.W. Lane and S.M. Smith. *Nutrition in Space*, pages 783–788. Williams and Wilkins, Baltimore, 1998.

[129] K. Shimada and Y. Fujii. Reconsideration of body mass measurement on the International Space Station and beyond. *Key Engineering Materials*, 381:461–464, 2008.

[130] Y. Fujii, K. Shimada, and K. Maru. Instrument for measuring the body mass of astronauts under microgravity conditions. *Microgravity Science and Technology*, 22:115–121, 2010.

[131] K. Shimada and Y. Fujii. Issues with Body Mass Measurement on the International Space Station (ISS). *Applied Mechanics and Materials*, 36:9–20, 2010.

[132] European Project. ActiBio. [Last seen: 10th March 2012] http://www.actibio.eu.