Weight estimation from visual body appearance

Carmelo Velardo and Jean–Luc Dugelay
EURECOM
2229 route des Crêtes, Sophia Antipolis, France
http://image.eurecom.fr
{lastname}@eurecom.fr

***Abstract*— Weight is a biometric trait which has been already studied in both the forensic and medical domains. In many practical situations, such as videosurveillance, weight can provide useful information for re–identification purposes but needs to be estimated from visual appearance (images or video). In this work we study the feasibility of weight estimation from anthropometric data directly accessible from the available image material. A model is retrieved via multiple regression analysis on a set of anthropometric features. A large medical database is exploited for the model training, while its validation is performed both on ideal and realistic conditions. The performance analysis of our approach shows that under noisy data conditions the system provides accurate estimations, putting the basis for a future work towards an automatic weight estimation.**

I. Introduction

Over the centuries, two different methodologies 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 human body weight is measured 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 [1], [2], [3]. During medical emergencies, measuring the weight of a patient can be difficult, due to the impossibility to move him, 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 [1]. Obviously a visual estimation cannot provide the precision of a scale, and in some cases a more accurate estimation can make the difference (e.g. for doing an anesthesia before a surgery).

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 videosurveillance records.

Scientists from this domain have shown interest in height and weight estimation in crime scenes. Regarding height, several studies [4], [5] 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 [6], 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.

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.

The paper is organized as follows: in I-A the related works are presented. Section II shows the proposed method; Section III the database exploited for our experiments is described. Section IV 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 images are used to verify our estimation. Finally the conclusion about our algorithm will be presented along with the future expectations.

A. Related works

Researchers working on Biometrics put recently a spotlight on the so called Soft Biometrics [7], [8]. These are not to be confused with Hard Biometrics, the traits commonly used until now (e.g. fingerprints, iris, retina, face appearance, and so on). Hard Biometrics have some peculiar characteristics such as *robustness* and *distinctiveness* (i.e. the capacity to differentiate two persons) which explain the use of the adjective "hard". Comparatively, Soft Biometrics do not exhibit these characteristics; as a result, in order to perform identification one cannot exploit these as done until now with the Hard Biometrics. However, it has been recently demonstrated that Soft Biometrics are useful to ameliorate

This work was partly supported by the french national project VideoID.

TABLE I

CLASSIFICATION OF SOME SOFT BIOMETRICS TRAITS SHOWING THEIR DISTINCTIVENESS AND PERMANENCE VALUES

Soft biometrics	Distinctiveness	Permanence
Height	Medium	High
Weight	Medium	Medium
Skin color	Low	Medium
Age	Medium	Medium
Facial hairs	Low	Medium
Gender	Low	High
Glasses	Low	Medium
Ethnicity	Medium	High
Eye color	Medium	High
Hair color	Medium	Medium

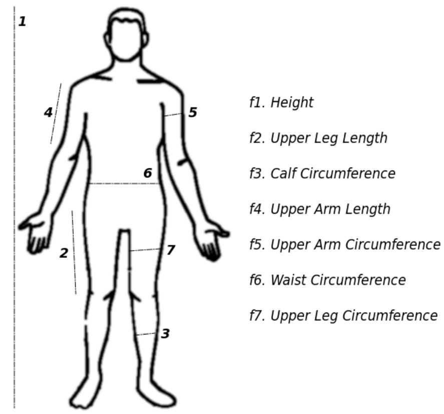

Fig. 1. Measures taken into account in our work on weight estimation.

the quality of the recognition and the speed of the biometric systems [7].

Soft Biometrics can be classified based on characteristics such as importance, stability, and reliability; an attempt of classification of such features can be observed in Table I. Some Soft Biometric traits are more *permanent* than others, while a lot of them lack of *distinctiveness*. This is indeed one of the main characteristics of these traits. For this reason 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.

Several works concentrated their efforts on the extraction of Soft Biometric traits [9], [10], [11], with the purpose of using them for recognition or tracking. Ethnicity and hair color were used in [12] to recognize and verify identities of clients in a biometric system. The authors have demonstrated that Soft Biometrics inscrease the recognition accuracy of the classic biometric system based only on face appearance. Moreover, in [7] 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. Similarly, in [13] the authors demonstrated that is possible to improve a gait recognition system by using “semantic biometrics” (read Soft Biometrics). Although among those traits weight was also present, its extraction was obtained via visual inspection led human operators.

These works have demonstrated the two major contributions of Soft Biometrics to recognition systems: increasing the speed of existing techniques (helping pruning the less likely solutions), and improving the recognition accuracy (by preventing rough errors).

Some of the Soft Biometrics traits are in relation to some measures of the human body, as is the case of height. Height estimation problem is a topic already mature in the literature and it has been exploited several times. One of the earliest approaches is presented in [9]. Some further improvements [10] led an increase of the precision using multiple measures and a statistical approach for removing the measures that more likely are outliers. Precise measurement of height has been already used in combination with other features to track people across multiple camera systems [14], and allows the identification of the same person in multiple video streams.

On the other hand, the estimation of weight has never been fully explored by the Biometrics community. Weight is still a challenging feature to be visually extracted, and to the best of our knowledge no baseline algorithms exist. In this paper a study about the feasibility of the weight estimation is proposed. For that reason an estimator has to be created in order to compute the weight starting from the visual appearance of the body and a representative set of features as to be defined on the human silhouette. We propose a methodology that, starting from these visual features, creates a regression model for weight estimation, and we suggest the inclusion of the weight in the already existing list of Soft Biometrics.

II. 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 mentioned in Section I-A. There, the authors exploited 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 chose 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 1.

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 this 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 problem fitting. This is formally described as following:

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

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 noise added to the linear relation among the variables:

$$y_i \;=\; \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \varepsilon_i \qquad i = 1, \ldots, n \tag{2}$$

To find the solution, the ordinary least squares method was 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 was conducted where the fitting was performed while varying the number of involved features from 1 to 7, so that all the possible combinations of features are explored (i.e. $2^7 - 1 = 127$).

Hereafter will follow a description of the database used for our analysis. Then the outcome of our experimental analysis will be presented.

III. Sources of data

As already introduced, weight is significant for the medical community; for this reason it is included in several databases used in such domain. NHANES [15] database fully satisfied each of our requirements. This dataset is unique because of its characteristics: size of the population, and time span analysis. This data set was collected from a large population of individuals (more than 28000 people), over a period of 6 years (from 1999 to 2005) by the Centers for Disease Control and Prevention during the National Health and Nutrition Examination Survey. The purpose of this survey was the monitoring of American population, and the assessment of health and nutritional conditions of adults and children in the United States. The database is a significant source of data regarding a wide range of different statistics: diseases diffusion, health conditions, physical measurements, and so on. We were interested in the section of the data set that collects some body measurements.

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 kept 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 1.

Some other features were present (i.e. subscapular and triceps skin fold), that were already reported, by the medical community, to be related to the quantity of fat, and then to the weight of persons. Nevertheless, we discarded them, since the 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 was incomplete, the data set was filtered out in order to keep only complete set of measurements. Additionally, we did 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 over 27000 subjects.

IV. Experimental results

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

The section is divided as follows: a first part will be 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 will be conducted where the measures are estimated from the images of a standard resolution camera commonly employed in videosurveillance.

A. Results under ideal conditions

We considered ideal conditions the use of measurements not affected by error. In this sense the original data present in the dataset where used. To perform the experiments we divided the database in training and testing set, respectively 70% and 30% of the available data. The training led 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:

$$\begin{aligned} \text{weight} \;=\; & -122.27 + 0.48 f_1 - 0.17 f_2 + 0.52 f_3 \\ & +0.16 f_4 + 0.77 f_5 + 0.49 f_6 + 0.58 f_7. \end{aligned} \tag{3}$$

For the list of features please refer to Figure 1. In this case the significancy 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 explored the full space of possible estimators (all the 127 possible combinations of 7 features). For the sake of clarity, we decided to report only the most representative results.

For assessing 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 models 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 was better than the others) we considered the cumulative value for the error range $\pm 10\%$ since this is the one considered acceptable in the medical community.

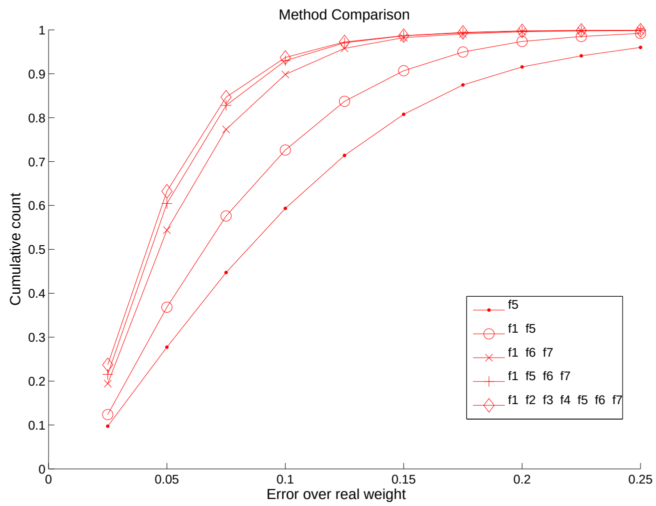

Fig. 2. Results for the multiple regression approach. It is evident that more information provided to estimator helps to increase the precision. (*For the meaning of each number please refer to Figure1*).

Figure 2 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 4).

TABLE II
PERFORMANCES OF HUMAN VISUAL WEIGHT ESTIMATION.

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

As a comparison we report in Table IV-A the results of an experiment [1] conducted at the Western Hospital of Melbourne. The experiment aimed at collecting visual estimates of the weight of 1137 patients. The study propose the statistics of the estimate conducted by the patients and the medical personnel. The weight of each patient was firstly estimated by himself, secondly the nurses and the physicians were asked to estimate it. The estimation was only performed visually, that makes this experiment comparable with the one we propose.

The precision of $\pm 5\%$ the weight of the patients is achieved by the patient and by most of the trained nurses. The physicians instead do not achieve the same results and their error is more spread (up to $\pm 20\%$ of the original weight).

As per Table IV-A, 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 (Fig. 2) can be observed 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 for more than 60% of our data set. An important result is also the consideration regarding the error at $\pm 10\%$ in Figure 2; at this level our system starts performing (93%) better than the patients (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.

B. *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. The analysis conducted in the first part of this paper is a bit optimistic. This level of precision can indeed be reached when the model is applied on data that are not affected by error, or in any case with an error that can be considered not significant for the estimation.

We obviously understand that such analysis is far from real cases, 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 a weight estimation application, 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 will describe in this section the results of such a work and we will 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.

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

$$\tilde{X} = [x_1, \cdots, x_n] + [\eta_1, \cdots, \eta_n]$$
$$|\eta_i| \leqslant x_i\,\epsilon, \quad \epsilon = \{0.05, 0.10, 0.15\}\,. \tag{4}$$

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.

As one can see from equation 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 one can observe the impact of 5% magnitude noise. The system still performs better than the professional result of [1]. The two other figures (4, 5) show the impact as

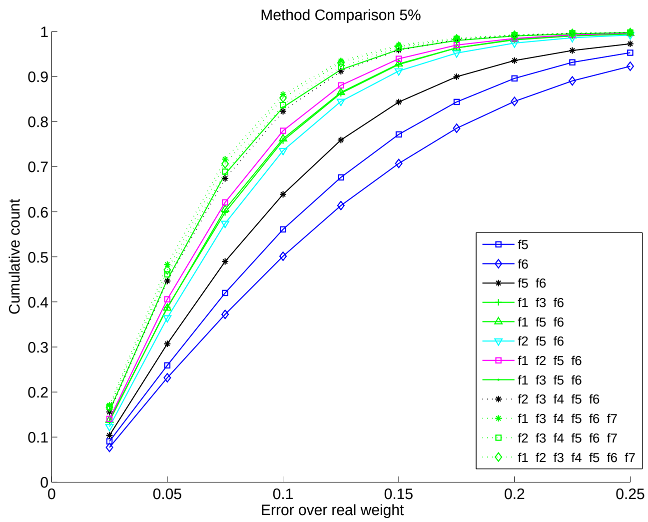

Fig. 3. Estimation response to data biased by 5% magnitude noise.

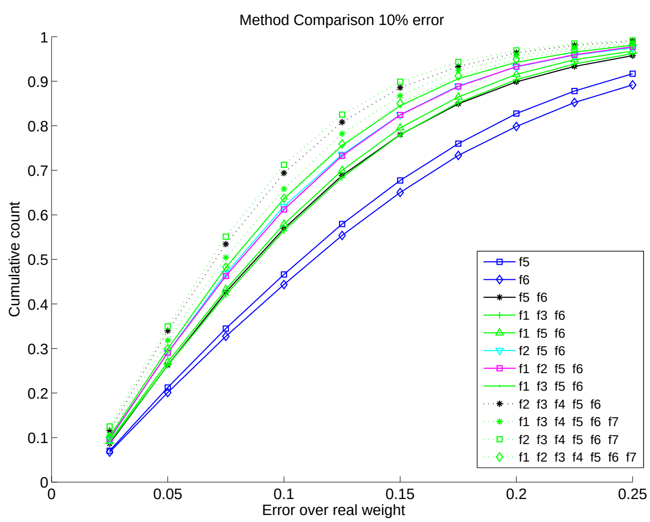

Fig. 4. Estimation response to data biased by 10% magnitude noise.

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 get 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 estimation (i.e. physicians in [1]).

C. Real case analysis

In order to measure the estimation performance in a real case scenario, and to confirm results obtained from the noise impact analysis, an experiment was 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 created a test set collecting some pictures of persons using a videosurveillance camera to recreate a possible application scenario.

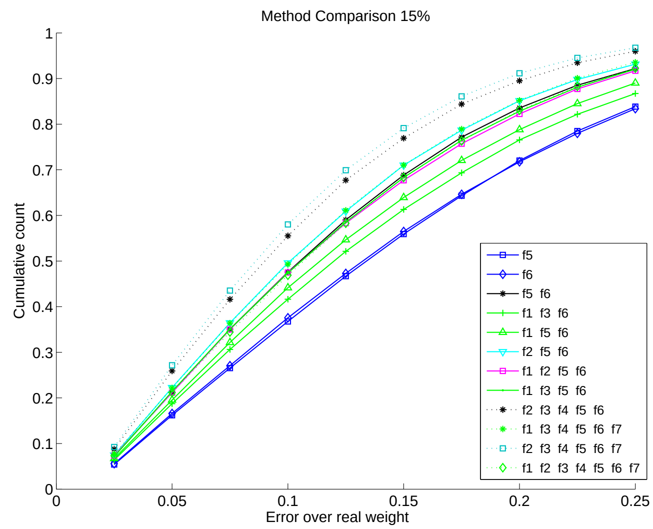

Fig. 5. Estimation response to data biased by 15% magnitude noise.

The database is composed by 20 different subjects (15 males and 5 females). The pictures were taken at fixed distance from the camera. Two different poses were experimented: a frontal shot and a profile one. A total of 40 pictures is available in the database. Example of such images are shown in Figure 6. In addition to the pictures, the weight was measured, using a common scale, to be considered as ground truth.

Fig. 6. An example of the two poses recorded during our database collection. A frontal and a profile pose are experimented. One should notices for the second subject how his clothes are hiding the real shape of the body.

In order to compute the weight using our model, an estimation of body parts size is needed. Since the important information about the 3D shape of the body is not available, an estimation 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, and calf) as the diameter of the cylinder approximating 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.

Several techniques are present in the mature literature of body parts detections, like [16], and each of them could serve for the purpose of this paper. However, being this a

TABLE III

THE ESTIMATION RESULTS OBTAINED FROM OUR DATABASE OF PEOPLE. THE ERROR IS SHOWN IN PERCENTAGE W.R.T. 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

preliminary analysis for this work, 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: for that we trivially compared the height of the subject with an object of known height inside the scene (eventually some other techniques could be exploited like proposed in [14], [9]); the second step involves the approximation of the other measures according to the height estimated in the previous step: for this a proportion computation was sufficient; the last part is the application of our statistical method that estimates the weight of the subject.

For this last step, all the estimators (127) computed in our theoretical study were experimented. The one presenting the best performance is the one that consider all the features but the calf information (*f3*).

What appears clearly visible from the example images is that subjects physical aspect varies a lot. For example, the loosen shirt and pants 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 estimation of the body parts measures. This statement explain why evidently the error for the calf measure has a significant negative impact on the estimation. Large pants increase the uncertainty about the part of the leg between the knee and the foot making poorer the performance of the system that considers its value.

The outcomes of our analysis are summarized in Table III. As we can see, the results vary in the whole output range. The average estimate error for this experiment is 4.3%, it is then confirmed that our system is able to estimate the weight from visual clues with an approximation of $\pm5\%$ of error over the real weight of the subject.

V. CONCLUSION

We presented a method to estimate the weight of a human body that exploits anthropometric features, known to be related to human appearance and correlated to the weight. 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 performace 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.

We believe that the proposed method will serve as a baseline for possible future approaches on the topic, and that the results presented in this work may contribute to stimulate new studies in many research sectors: from forensic analysis to video surveillance, from biometrics to medical research.

ACKNOWLEDGMENTS

The authors would like to thank Angela D'Angelo and Judith Redi for their precious comments. Additionally we thank the anonymous reviewers for their very insightful comments. Finally we acknowledge the projects that partly supported this work: the French national projects ANR VideoID and OSEO Biorafale.

REFERENCES

[1] S. Menon and A. M. Kelly, How Accurate is Weight Estimation in the Emergency Department?, *Emergency Medicine Australasia*, 2005

[2] T. R. Coe and M. Halkes and K. Houghton and D. Jefferson, The Accuracy of Visual Estimation of Weight and Weight and Height in Pre–operative Supine Patients, *Anaesthesia*, 1999

[3] W. L. Hall II and G. L. Larkin and M. J. Trujillo and J. L. Hinds and K. A. Delaney, Errors in Weight Estimation in the Emergency Department: Comparing Performance by Providers and Patients, *The Journal of Emergency Medicine*

[4] I. Duyar and C. Pelin, Body Height Estimation Based on Tibia Length in Different Stature Groups, *American Journal of Physical Anthropology*, 2003

[5] M. C. De Mendoca, Estimation of Height from the Length of Long Bones in a Portuguese Adult Population, *American Journal of Physical Anthropology*, 2000

[6] K. Krishan, Establishing correlation of footprints with body weight–Forensic aspects, *Forensic Science International*, 2008

[7] A. K. Jain and S. C. Dass and K. Nandakumar, Soft Biometric Traits for Personal Recognition Systems, *ICBA*, 2004

[8] A. Ross and K. Nandakumar and A. K. Jain, Handbook of Multibiometrics, *Verlag*, 2006

[9] A. Criminisi and I. Reid and A. Zisserman, Single View Metrology, *International Journal of Computer Vision*, 1999

[10] S. H. Lee and J. S. Choi, A Single-View Based Framework for Robust Estimation of Height and Position of Moving People, *PSIVT*, 2007

[11] H. C. Kim and D. Kim and Z. Ghahramani and S. Y. Bang, Appearance based gender classification with Gaussian processes, *Pattern Recognition Letters*, 2006

[12] G. L. Marcialis and F. Roli and D. Muntoni, Group–specific face verification using soft biometrics, *Elsevier*, 2009

[13] S. Samangooei and B. Guo and M. S. Nixon, The Use of Semantic Human Description as a Soft Biometric, *BTAS*, 2008

[14] C. Madden and M. Piccardi, Height Measurement as a Session-based Biometric for People Matching Across Disjoint Camera Views, *University of Technology Sydney*, 2005

[15] National Health and Nutrition Examination Survey, *Center for Disease Control and Prevention*, 1999–2005

[16] K. Mikolajczyk and C. Schmid and A. Zisserman, Human detection based on a probabilistic assembly of robust part detectors, *Computer Vision-ECCV 2004*, 2004