IMPROVING IDENTIFICATION BY PRUNING: A CASE STUDY ON FACE RECOGNITION AND BODY SOFT BIOMETRIC

Carmelo Velardo, Jean-Luc Dugelay

Eurecom
2229 route des Crêtes, Sophia Antipolis, France
{velardo,dugelay}@eurecom.fr

ABSTRACT

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 large scale medical dataset to guarantee statistical meaning of the results, and tested in conjunction with a face recognition algorithm. Our assumptions are verified by testing our system on a chimera dataset. We clearly identify the trade off among pruning, accuracy, and mensuration error of an anthropomeasure 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.

Index Terms— Soft biometrics, pruning, face recognition.

1. INTRODUCTION

Soft biometrics are a new trend in biometric studies which exploits the information coming from non-reliable, non-discriminative human traits (height, eye color, ...). They were firstly investigated by Jain et al. to improve multi-modal fusion [1] and later they were exploited to perform identification [2, 3], or to simply extract information related to the user [4]. Anthropometric measures fulfill the definition of soft biometrics [2]: they do not provide a specific pattern for identification (e.g. fingerprint and iris), they are human-compliant (i.e. people can use them to describe someone's appearance), and they are available without user cooperation.

Anthropometry science uses human body measurements to study variation and differences of the human body. Particularly useful in case of medicine and industrial design, anthropometric studies were used at the beginning of biometric people identification: Alphonse Bertillon first derived a method that involved several body measures to classify and to identify criminals. However, anthropometry was replaced by a more reliable method based on fingerprints. Indeed, while soft biometrics have intrinsic problems with identification (e.g. lack of discriminative power above all), hard biometrics [2] provide higher distinctiveness, although their elaboration is often quite expensive in terms of resources. In this paper we use body soft biometric to prune the search space so as to ease the task of a subsequent hard biometrics module. We demonstrate that a face recognition system, priorly processed by our anthropometric features, improves its performance in both accuracy and recognition speed. The two independent complementary information provided by the anthropometric signature and the face appearance makes this accuracy and speed gain possible.

A possible application of anthropometric based pruning would be border control passport. For example, capturing both facial appearance and body measures at Paris airport will make easier, faster, and more robust the recognition of the passenger at New York, by filtering out thousands of other passengers that do not match the anthropometric signature. Even if the recent introduction of body scanners will definitely improve the mensuration capabilities of automatic systems, since they can perceive the shape under clothes fabric; the mensuration process will always be considered error prone. Sources of noise are mainly due to the sensing device (tape meter, 3D scanner, ...), to the human operator, and to the inner variability of human measures. Even if the sensing procedure will improve in a near future thanks to specialized devices, noise will eventually affect the recognition performance of an anthropometric system. For this reason we study how an increasing noise during the mensuration step can affect the retrieval process and we show that even in presence of strong noise magnitude (10% of the real measure) anthropometric measures are still useful to reduce the search space and, at the same time, to improve the performance of other biometric traits.

The article is structured as follows: we review works on anthropometric systems and search in biometric databases in Section 2. In Section 3 we present the anthropometric dataset used and we introduce the methodology of the study proposed to couple the pruning system with a face recognition algorithm. Finally in Section 4 we analyze the results obtained by our statistical analysis, and we show the gain in performance.

The authors thank Miriam Redi for her insightful comments. They also acknowledge the VideoID project for having supported their work.

2. PREVIOUS WORK

2.1. Anthropometry-based recognition

The first modern work on identification via anthropometric measures is that of Daniels in [5]. The author seeks to find the "Average man", that is to say a person whose traits fall into average ranges. Through an elimination process, the author demonstrates that 10 measures provide enough distinctiveness to exclude the existence of such a subject.

To the best of our knowledge, the latest work involving anthropometric-based people recognition is the one presented by Ober et al. in [3]. The authors exploit the CAESAR 3D dataset that contains the scans of 4400 individuals to analyze 27 different measures. They apply dimensionality reduction through Linear Discriminant Analysis (LDA) analysis to obtain 97% accuracy in a population of 2000 subjects. A similar work [6] degrades 3D measures to study the performance variation of an anthropometric recognition system.

All these methods start from the assumption that such a quantity of measures (27 in the case of [3]) 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 noise similar to what has been reported in [6]. Thus making seriously challenging to create a system that performs identification with noisy data.

2.2. Soft-biometrics based database pruning

A discussion on the possibility of using pruning via soft biometrics is already present in [7] and a later formalization is present in [2]. In [8] semantic information are coupled with gait signatures to identify people from a database of individuals. Using predefined classes to perform pruning increases the problem of misclassification and the result of pruning can be prejudiced. In our case, we consider continuous values which preserves us from such a drawback. Errors in one of the measure can be compensated by others.

3. PROPOSED CASE STUDY

We propose to use a body soft biometric signature (composed of several body measures) as a pre-classification module to be used prior to a recognition algorithm based on face recognition. The scope is to preselect a subset of users so that the recognition based on face analysis will be faster and more accurate. For this reason two different database were used: one for the anthropometric measures, the other one for the face appearance. For the first one we resorted to the large NHANES dataset, for the second to the well known FERET database.

NHANES [1] is a study promoted yearly by U.S. CDC office devoted to monitor health statistics of a representative sample of the American population. The size of this database makes it an important source of information. Indeed, we benefit of more than 28000 individuals recorded over a period that goes from 1999–2008. An important section includes a set of body measures taken under controlled conditions: height, arm circumference, arm length, waist circumference, leg circumference, leg length, calf circumference, body weight.

Like other databases of large dimension, NHANES suffers of a considerable amount of missing data. Since dealing with missing data is out of the scope of our study, we discarded subjects presenting missing measures. We consider the features as continuous value, and each of them is considered of the same importance, thus of the same discriminative power. One could argue that in order to select the features with most discriminative power, a study similar to [3, 4] should be conducted. Nevertheless, to perform a similar analysis, we should consider a plethora of measures which is out of the scope of this paper, indeed our aim is to use those measures as a simplified body shape description.

We used the well known FERET as face dataset. It contains, in its largest gallery, 1195 pairs of images. The algorithm for face recognition is extrapolated by the code for the Eigenfaces approach provided by the work of [9]. It was used for the face recognition algorithm that we consider as baseline.

4. EXPERIMENTAL RESULTS

In the first part of our analysis we explore the pruning capabilities of our anthropometric signature under noisy conditions. Similarly to [4, 6] each anthropometric measure is added with a normally distributed random noise of increasing intensity to simulate a real mensuration system. Using only the anthropometric measures we derive the best *penetration rate* (i.e. how much we prune the database) as the noise progressively increases. We show that a trade off is possible to not interfere with the hard biometric recognition performance.

In the second part we apply the result of the pruning to the recognition algorithm: in a first analysis we discuss the achievable performance gain in accuracy that our pruning enables, in a second paragraph we show the trade off between accuracy, mensuration error, and penetration rate. Here a *chimera dataset* has to be created; that is to say, fusing the data from both datasets we create virtual users composed by the anthropometric measures from NHANES and the facial appearance of FERET users.

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 iteratively randomized the choice

[1]The data we analyzed are freely available at: `http://www.cdc.gov/nchs/nhanes.htm`.

	Mean	Min / Max		Mean	Min / Max
Weight	7.9	2.5 / 21.8	Arm c.	3.2	1.7 / 6.1
Leg l.	3.9	2.2 / 5.4	Calf	3.8	2.1 / 7.5
Arm l.	3.7	2.6 / 4.8	Height	16.7	13.0 / 20.3
Waist	9.7	5.9 / 17.5	Leg c.	5.2	2.7 / 10.0

Table 1. Statistics of 10% noise magnitude; units are expressed in kilograms for body weight, and centimeters for all the other measures.

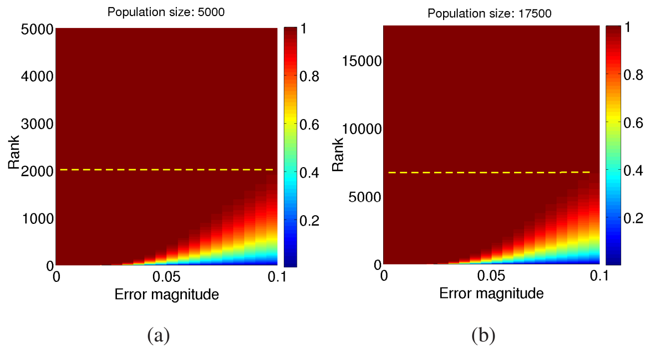


Fig. 1. CMC curves for population sizes of (a) 5000 and (b) 18000 subjects. The color indicates the probability value of the CMC curve. The dashed lines indicate the certainty (probability $= 1$) that the client is within the corresponding rank. For both the curves this value is around 50% of the database for 10% error magnitude.

of individuals from NHANES. The first measure consists in warping all the images so that eyes, mouth, and nose positions will always fall at the same fixed locations; the second procedure consists in repeating several times the same experiment using always different subjects and then averaging the results.

4.1. Anthropometry system performance analysis

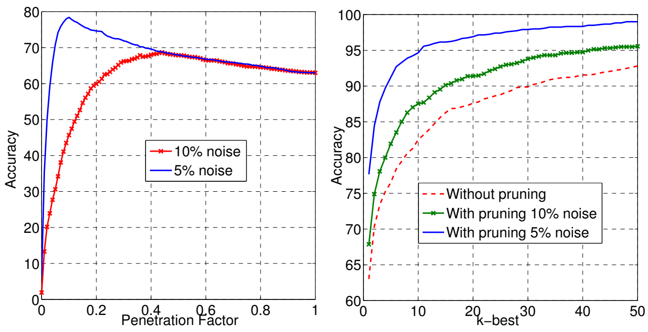
In order to evaluate the performance of the pruning phase, we use the penetration rate value. The penetration rate is inversely proportional to the pruning factor. The higher the penetration rate is, the larger is the portion of database passed to the recognition algorithm.

We consider our feature vector as the full set of anthropometric measures. To rank the results we employ an euclidean distance metric as suggested in [6]. To simulate a real mensuration system we consider a varying noise, and in order to increase the randomness of this bias, an approach similar to the one proposed by [4] is used, where the noise is assumed as follows: considering our feature vector $F = [f_1, \dots, f_n]$ we add to each measure separately a noise proportional to the magnitude of the original feature and with random sign. The result of the operation is the biased version of the feature $F_\epsilon = [f_1(1+\alpha_1 w_1), \dots, f_n(1+\alpha_n w_n)]$ where f_i is the real value of the feature, α_i is a binary random variable in the set $\{-1, +1\}$, and w_i is the magnitude assigned to the noise. w_i ranges from 0 to 0.10 (i.e. 10%). To compare our noise to the one observed in [3, 6] we summarize the noise statistics in Table 1. Considering the average noise, one can notice how close we are to a worst case scenario.

To evaluate how the anthropometric-based ranking performs at each different noise magnitude, we employ a cumulative matching characteristic curve (CMC) which summarizes the system performance. The curve indicates the probability of observing the matching subject (*client*) in the first N-best candidates. Once obtained the CMC curves for each noise magnitude, we stack them up to show them as a heatmap (fig. 1) so as to compare different experiments. We performed our analysis over different population sizes drawn randomly from the original dataset. For the sake of brevity, in Figure 1 we show the experiments conducted with 5000 and 17500 subjects considered (respectively 30% and 100% of NHANES dataset). The first experiment was conducted multiple times with random populations drawn from NHANES and the results were averaged; the second one was performed with the totality of the users. If we consider the maximum values admitted for the noise magnitude (10%) we clearly see that a penetration rate of 50% can be obtained, that is to say our client is always (probability $= 1$) within the first half of ranked results. As from Figure 1, in case of a population of 5000 subjects the client lays within the first 2000 results; while in the second case within the first 8000. 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.

4.2. Recognition accuracy gain by pruning

We analyze in this section the performance of the cascade composed by our pruning module and a face recognition algorithm. The baseline produces at rank-1 a result of 63% accuracy. To analyze how the pruning affects this result, we analyze 5% and 10% noise magnitudes. For the sake of completeness we tested our 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, i.e. the one that better exploits the capabilities of the two modules. At each iteration a set of 1195 virtual users is created associating to each FERET identity a randomly selected (and biased) feature vector from NAHNES. The feature vector is used to perform a fast search in the anthropometric dataset. Once all euclidean distances are computed and the identities are ranked accordingly, the face recognition is performed on the subset defined by the penetration rate chosen. Figure 2.(a) shows that, as we consider lower penetration rate values, the final rank-1 accuracy suffers from the poor performance of the pruning that is unable to provide the client



(a) (b)

Fig. 2. (a) Rank-1 accuracy variation as the penetration rate gets larger. 5% and 10% noise magnitude are considered. (b) Accuracy gain achievable with best penetration factor for 5% and 10%.

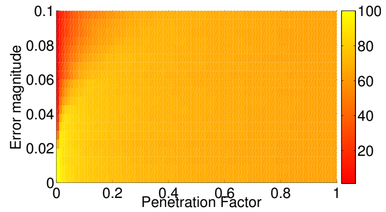


Fig. 3. The rank-1 accuracy as function of penetration rate and mensuration error. By fixing one parameter we can see which second choice give us the best performing system.

within a given certainty. In both cases the curves behavior is very similar; the rank-1 accuracy raises up to a global maximum, and then, asymptotically, falls back to the baseline accuracy result (63%) as the portion of pruned identities considered gets closer to the original database size. The maximum corresponds to the penetration rate value that best put together the benefit of the pruning and the recognition algorithm. For the first system the maximum indicates the best penetration rate of 0.1 (10% of the original set); while in the second case is 0.45. Furthermore, to complete our analysis we show in Figure 2.(b) the full rank curves of the two systems. In the case of 5% noise the gain is both high in the sense of accuracy and speed performance (since the penetration rate is smaller), while in the second case both the gains are reduced, but still the pruning provides improvements to the baseline algorithm while reducing the search space by a factor of $2\times$.

In case the recognition is 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. In order to analyze the full response of the system to both the penetration rate and the mensuration error, we present Figure 3. Here the performance of the global system are summarized in terms of rank-1 results as function of penetration rate and mensuration error. We can exploit this 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.

5. CONCLUSION

We demonstrated that even noisy body soft biometrics measurements can help to simultaneously reduce the search space and to increase the performance of a face recognition system. A medical database is used to analyze large scale system performance. We clearly identify the trade off between penetration rate and accuracy of the identification performed by the system, and the response to the noise of the mensuration system. Our results distinctly show that in the worst case of $\pm 10\%$ noise added to the anthropometric measures, the pruning system is still able to speed up the search of $2\times$ factor while guaranteeing an increase of accuracy performance. As normal continuation we will apply our results to a real case scenario where both face and body are considered (e.g. a passport control at state border). Moreover, identifying a confidence interval for the body measures or increasing the number of anthropometric measurements will undoubtedly ameliorate the performance of the pruning step.

6. REFERENCES

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

[2] A. Dantcheva, C. Velardo, A. D'angelo, and J.-L. Dugelay, "Bag of soft biometrics for person identification : New trends and challenges," *Mutimedia Tools and Applications, Springer*, October 2010.

[3] D.B. Ober, S.P. Neugebauer, and P.A. Sallee, "Training and feature-reduction techniques for human identification using anthropometry," in *4th IEEE BTAS*, September 2010, pp. 1–8.

[4] C. Velardo and J.-L. Dugelay, "Weight estimation from visual body appearance," in *4th IEEE BTAS*, September 2010.

[5] G.S. Daniels, "The "average man" ?," Tech. Rep., Air Force Aerospace Medical Research Lab Wright-Patterson, 1952.

[6] A. Godil, P. Grother, and S. Ressler, "Human identification from body shape," in *3-D Digital Imaging and Modeling*, 2003, pp. 386–392.

[7] J.L. Wayman, "Large-scale civilian biometric systems - issues and feasibility," Card Tech / Secur Tech ID, 1997.

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

[9] 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*, pp. 252–260, 2005.