Gait Recognition Using Geometric Features and Soft Biometrics

Konstantinos Moustakas, *Member, IEEE*, Dimitrios Tzovaras, *Member, IEEE*, and Georgios Stavropoulos

***Abstract*—This letter presents a novel framework for gait recognition augmented with soft biometric information. Geometric gait analysis is based on Radon transforms and on gait energy images. User height and stride length information is extracted and utilized in a probabilistic framework for the detection of soft biometric features of substantial discrimination power. Experimental validation illustrates that the proposed approach for integrating soft biometric features in gait recognition advances significantly the identification and authentication performance.**

***Index Terms*—Authentication, gait recognition, height, identification, soft biometrics, stride.**

I. Introduction

THE area of biometrics for access control and security has been extensively researched during the last four decades. Biometrics measure the unique physical or behavioral characteristics of individuals as a means to recognize or authenticate their identity. Common physical biometrics include fingerprints, hand or palm geometry, and retina, iris, or facial characteristics. Behavioral characteristics include among others signature, voice (which also has a physical component), keystroke pattern, and gait.

During the latest years, there has been a growing interest in the identification of the humans based on their way of walking that is motivated by the unobtrusiveness of the trait and the promising reported recognition rates. The results of this research field can be directly applied for surveillance, identity verification, and in medical applications as well.

Recently the term soft biometrics has been introduced in the literature to describe the biometric traits that are easily extracted but lack the distinctiveness and discriminating power of typical biometrics. It has been also reported that soft biometrics could increase the recognition or verification rates of typical biometric systems [1].

A. Previous Work

Most of the recent gait analysis methods can be divided into two main categories; model-based and feature-based methods. Model-based approaches, study static and dynamic body parameters of the human locomotion [2], like stride length, stride speed and cadence [3]. On the other hand, feature based techniques do not rely on the assumption of any specific model of the human body for gait analysis. They employ simple temporal correlation, linear time normalization [4], full volumetric correlation on partitioned silhouette frames [5], and dynamic time warping [6]. In general, feature-based methods are seen to perform better even if they exhibit limitations, e.g. dependency on the walking direction or image noise.

The use of soft biometrics for recognition has been only recently studied [1], [7]. Due to their low discriminating power, they are rarely used independently for recognition or verification but are used to verify hypotheses or in general to reduce the search space in typical biometric systems [8].

B. Contribution

The present paper proposes a novel scheme for the integration of two or even more soft biometric traits in a biometric recognition system using a stochastic framework. In particular, the "height" soft biometric trait along with the "stride length" are utilized to augment the information obtained by a gait recognition system and to ultimately advance its performance. It should be emphasized that the proposed approach goes beyond explicit weighted fusion of different traits [9] at the score level, including also soft biometrics, as performed in [1], [7] since it exhibits some disadvantages like the need of the computation of a soft biometric score or weighting functions for fusion at the score level based usually on posterior probabilities. The proposed stochastic framework includes soft biometrics directly in the estimation of the biometric score.

II. Augmenting Gait Recognition With Soft Biometrics

Fig. 1 illustrates the architecture of the proposed framework. The initial gait sequence is processed so as to extract the geometric gait features and the soft biometric features. Then using the proposed probabilistic framework the geometric gait score and the soft biometric measurements are combined so as to generate the final score.

Let f_g be the biometric geometric feature of a gait sequence. Let also Ω be the set of all identities $x_i, i \in \{1,\ldots,N\}$, where N is the number of elements in Ω. Now let f_{soft} be the soft biometric feature that can be both scalar or vectorial. In our case, f_{soft} is instantiated by f_h and f_s that correspond to the soft biometric features of height and stride length.

Without loss of generality we will provide the subsequent analysis for the soft biometric feature for height f_h. Let us also

Manuscript received October 14, 2009; revised December 23, 2009. First published January 22, 2010; current version published February 19, 2010. This work was supported by the European Union (EU) funded ACTIBIO IST STREP (FP7-215372). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Nikolaos V. Boulgouris.

The authors are with the Informatics and Telematics Institute/Centre for Research and Technology Hellas, Thermi-Thessaloniki 57001, Greece (e-mail: moustak@iti.gr; tzovaras@iti.gr; stavrop@iti.gr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2010.2040927

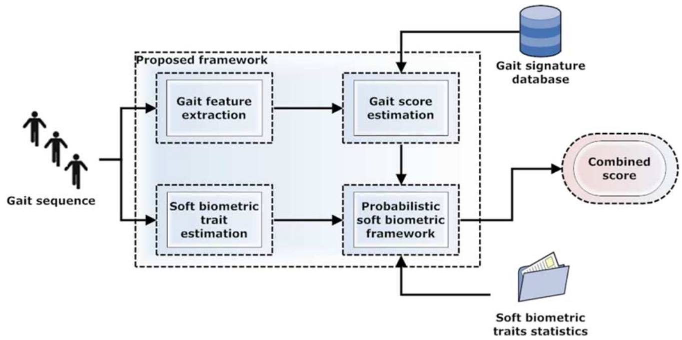

Fig. 1. Architecture of the proposed framework.

define the notion of minor cluster M_i as a subset of the soft biometric feature space F that exhibits notable variation in terms of f_h and has the following characteristics.

The *a priori* probability of a soft biometric feature to belong to a minor cluster is very low

$$p(f_h \in M_i) = p_i \ll 1. \tag{1}$$

There exists a subset S_i of Ω such as

$$\exists S_i \subset \Omega \begin{cases} \forall x \in S_i, & p(f_h \in M_i|x) > a \\ \forall x \notin S_i, & p\left(f_h \in M_i|x\right) \simeq p_i \end{cases}. \tag{2}$$

Moreover, M is the union of all minor clusters whose number N_M should be significantly lower than the size $|\Omega|$ of Ω

$$\begin{aligned} M = \cup M_i, \, \forall i = 1, \ldots, N_M \\ N_M \ll |\Omega|. \end{aligned} \tag{3}$$

Definition: A cluster M_i of the feature space, associated to a subset S_i of the set of identities Ω is characterized as minor cluster iff (1)–(3) hold.

The idea under the aforementioned analysis is that the soft biometric information could be highly discriminative for special cases of identities or groups of identities exhibiting soft biometric features that vary from the average features of the population [9].

An issue of high importance is to link the soft biometric information to the geometric gait feature f_g when the soft biometric feature is available. Let us define the recognition probability (or score) after measuring f_g and link it with the soft biometric height f_h

$$\begin{aligned} p(x|f_g) = & \, p\left(x|f_g, f_h \in M_i^h\right) p\left(f_h \in M_i^h\right) \\ & + p\left(x|f_g, f_h \notin M_i^h\right) p\left(f_h \notin M_i^h\right) \end{aligned} \tag{4}$$

where M_i^h denotes the minor cluster related to the "height" soft biometric. Moreover

$$p\left(x|f_g, f_h \in M_i^h\right) = \frac{p\left(f_h \in M_i^h|f_g, x\right) p(x)}{p\left(f_h \in M_i^h|f_g\right)}. \tag{5}$$

Thus

$$\frac{p\left(x|f_g, f_h \in M_i^h\right)}{p\left(x|f_g, f_h \notin M_i^h\right)} = \frac{p\left(f_h \in M_i^h|f_g, x\right) p\left(f_h \notin M_i^h|f_g\right)}{p\left(f_h \notin M_i^h|f_g, x\right) p\left(f_h \in M_i^h|f_g\right)}. \tag{6}$$

Since f_g and $f_h \in M_i^h$ are independent events then

$$\begin{aligned} p\left(f_h \in M_i^h|f_g, x\right) &= \frac{p\left(f_h \in M_i^h, f_g|x\right)}{p(f_g|x)} \\ &= \frac{p\left(f_h \in M_i^h|x\right) p(f_g|x)}{p(f_g|x)} \\ &= p\left(f_h \in M_i^h|x\right) \end{aligned} \tag{7}$$

Thus, (6) turns into

$$\frac{p\left(x|f_g, f_h \in M_i^h\right)}{p\left(x|f_g, f_h \notin M_i^h\right)} = \frac{p_x^h(1-p^h)}{(1-p_x^h)\,p^h} = k \tag{8}$$

where $p_x^h = p(f_h \in M_i^h|, x)$ and $p^h = p(f_h \in M_i^h)$. By these means (4) transforms to

$$p\left(x|f_g, f_h \in M_i^h\right) = \frac{k}{1 - p^h + kp^h} p(x|f_g). \tag{9}$$

Now if $x \notin S_i$ then $k = 1$. On contrary if $x \in S_i$ then assuming that the feature $f_h \in M_i$ then the following relationship should hold: $p(x|f_g, f_h \in M_i^h) \geq p(x|f_g)$, which is valid only for $k \geq 1$ in (9) that is ultimately reduced to the condition $p_x^h > p^h$ that is valid following (1) and (2). The variable α can be tuned according to the application and represents the strictness of the definition of the minor cluster in the soft biometric feature space.

Using a similar analysis the above concept can be extended to two or more auxiliary soft biometric modalities. For the present case, it can be easily shown that

$$\begin{aligned} & p\left(x|f_g, f_h \in M_i^h, f_s \in M_i^s\right) \\ & \quad = \frac{k}{1 - p^h + kp^h} \frac{l}{1 - p^s + lp^s} p(x|f_g) \end{aligned} \tag{10}$$

where f_s is the stride soft biometric feature, $p_x^s = p(f_s \in M_i^s|, x)$, $p^s = p(f_s \in M_i^s)$, M_i^s are the stride soft biometric minor clusters, and $l = p_x^s(1 - p^s)/(1 - p_x^s)p^s$.

III. Implementation Details

The application of the framework of Section II in a gait recognition scenario requires first of all the development of a geometric gait recognition algorithm. Moreover, the height and stride length soft biometric features should be extracted. Finally, the probabilities $p(f_h \in M_i^h)$, $p(f_s \in M_i^s)$ and $p(f_h \in M_i^h|x)$, $p(f_s \in M_i^s|x)$ have to be modelled. In the following subsections these implementation details are briefly described.

A. Geometric Gait Recognition Synopsis

The geometric gait recognition feature vector f_g is extracted using two different algorithms so as to illustrate the invariance of the proposed scheme on the geometric gait recognition algorithm utilized. The first algorithm that is presented in [10] and [11] is based on the radial integration transform (RIT) that is applied on gait sequence silhouettes. The algorithm results in a single geometric gait feature vector f_g of proven efficiency [10]. The second algorithm is based on gait energy images as described in [12] and is based on matching spatiotemporal images of human gait.

B. Height and Stride Length Estimation

A comprehensive analysis on the height and stride length estimation is out of the scope of this paper. However, in order to make this paper self contained a brief outline of the algorithms follow.

The height and the stride length soft biometric features are estimated utilizing the calibrated stereoscopic sequences that were obtained by capturing the HUMABIO [10] and ACTIBIO databases. Since real world coordinates and absolute distances can be extracted through calibrated stereoscopic sequences, the problem of the estimation of the height and stride length features is trivially reduced in the selection of the features that correspond to the highest-lowest part of the subject, concerning height, and to the largest distance between the legs within a gait cycle.

C. Modelling Minor Clusters and A Priori Probabilities

A significant task for applying the framework of Section II is to define the minor clusters and to model the probabilities $p(f_h \in M_i^h)$, $p(f_s \in M_i^s)$ and $p(f_h \in M_i^h|x)$, $p(f_s \in M_i^s|x)$.

Assuming a uniform partition of the soft biometric feature space F into N_F clusters, then the minor clusters are defined through (1), where the value of p_i is experimentally set as $p_i = 0.1$. The number of clusters L is set to $L = 10$ for both height and stride length soft biometric feature spaces.

As soon as information of a population is available then the probabilities $p(f_h \in M_i^h)$, $p(f_s \in M_i^s)$ can be directly estimated as long as the minor clusters are defined. Concerning probabilities $p(f_h \in M_i^h|x)$, $p(f_s \in M_i^s|x)$ their definition is flexible and could also depend on the application, e.g. whether the soft biometric feature space is discrete or continuous.

In the present case, both the soft biometric feature spaces for gait and stride length, are continuous. $p(f_h \in M_i^h|x)$ and $p(f_s \in M_i^s|x)$ are modelled using a Gaussian model

$$p\left(f_h \in M_i^h|x\right) = \begin{cases} \frac{1}{\sigma_i^h\sqrt{2\pi}} e^{-\frac{\left(f_h-\mu_i^h\right)^2}{2\sigma^2}}, & \text{if } f_h \in M_i^h \\ 0, & \text{if } f_h \notin M_i^h \end{cases} \quad (11)$$

where μ_i^h and σ_i^h are the mean and the standard deviation of the population of M_i^h. The same model is also applied for the estimation of $p(f_s \in M_i^s|x)$.

IV. Experimental Results

Since benchmarking databases like the USF gait challenge do not include depth information, the proposed algorithms have been tested in both the HUMABIO and ACTIBIO databases that include gait sequences captured with stereoscopic cameras. The HUMABIO database, extensively described in [10] was captured in an indoor environment and consists of 75 subjects in the first and 51 in the second capture session. The collection protocol had each person walk multiple times naturally along a predefined path, so that the view is approximately fronto-parallel. The ACTIBIO database is a proprietary activity recognition dataset that also includes two sessions of gait sequences from 28 subjects that were captured in two months difference. The subjects were asked to walk several times following patterns of increased complexity (e.g. fronto-parallel, progressive deviation in the walking direction).

It should be also mentioned that even if benchmarking of gait recognition algorithms can depend on the benchmarking database, this in not the case for the proposed framework. Assuming a correctly structured database that includes subjects with different soft biometric features close to the distribution observed in the real world, e.g. with different heights, the proposed framework is seen to augment and improve performance of any gait recognition algorithm implemented and does not depend on the database from a qualitative point of view. It is obvious that for the case of a database where all subjects would have the same height, the proposed algorithm would not offer any added value.

The proposed framework has been tested both in terms of identification and authentication rates. It should be mentioned that sequences from different recording sessions are used for enrolment and identification/authentication that is generally assumed as a challenge in gait recognition. The sessions have a time gap of six and two months for the HUMABIO and ACTIBIO databases, respectively.

A. Identification

Fig. 2(a) and 2(b) illustrate comparative results in the ACTIBIO database, using as geometric gait feature extraction the algorithms of [10] and [12] respectively. Fig. 2(c), illustrates comparative results in the HUMABIO database using for gait feature extraction the method described in [10].

All aforementioned diagrams illustrate the efficiency of the proposed approach using different algorithms for gait feature extraction and different databases. Four curves are displayed in each figure that correspond to the cumulative matching scores (CMS) using only the gait feature (gait), combined with height $(gait + height)$, stride length $(gait + stride)$, both soft biometrics $(gait + height + stride)$. As expected, augmenting the gait feature with additional soft biometric information significantly increases the gait recognition efficiency. It should be also emphasized that from a theoretical point of view the proposed framework is expected to advance the recognition rate for incorrect identification cases for subjects that exhibit soft biometric features of substantial discrimination power, i.e., for subjects that lie within the minor clusters.

A direct quantitative comparison of the proposed scheme with other existing algorithms for soft biometric fusion with classical biometric traits [7] should be performed only taking several considerations into account. First of all, the proposed framework provides direct means for exploiting soft-biometric information in classical biometric systems. The only precondition is the modelling of the conditional and a priori probabilities as described in Section III-C, which is a relatively trivial task. The approach in [7] provides a general score-level fusion framework that however needs to explicitly estimate weighting parameters that regulate the effect of the soft biometric traits in the final score. The estimation of these weights is a non-trivial task as also mentioned by the authors of [7]. Moreover, the proposed approach from a general point of view does improve the a posteriori probability of a candidate subject taking into account its

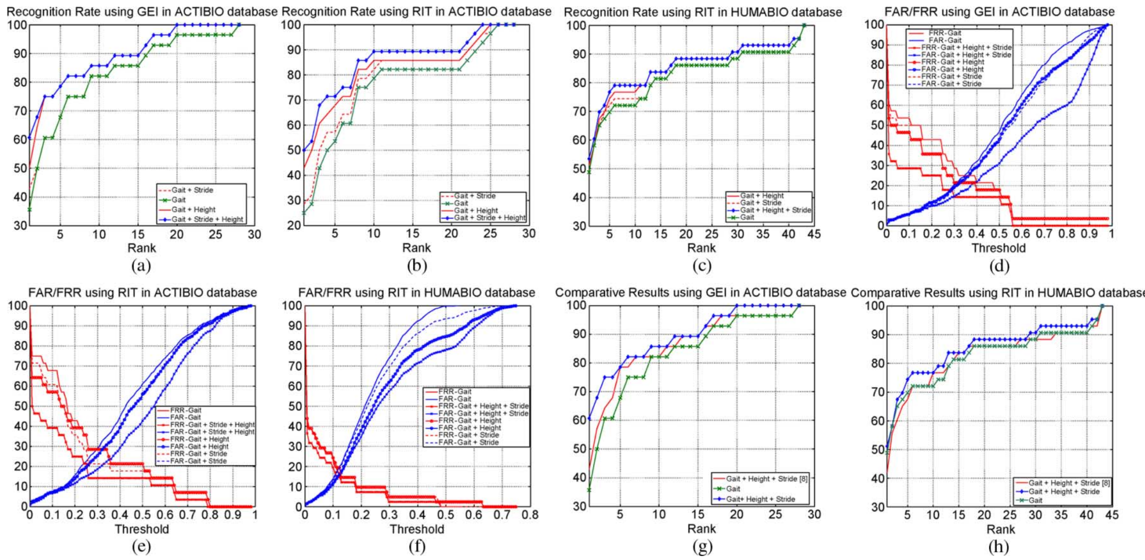

Fig. 2. (a) Cumulative matching scores (CMS) for the ACTIBIO database using as baseline the algorithm in [12], (b) CMS for the ACTIBIO database using as baseline the algorithm in [10], (c) CMS for the HUMABIO database using as baseline the algorithm in [10], (d) FAR-FRR diagrams for the ACTIBIO database using as baseline the algorithm in [12], (e) FAR-FRR diagrams for the ACTIBIO database using as baseline the algorithm in [10], (f) FAR-FRR diagrams for the HUMABIO database using as baseline the algorithm in [10], (g) comparative results of the proposed approach with the approach of [7] in the ACTIBIO and in the (h) HUMABIO database.

soft biometric feature, but does not diminish it if the soft-biometric measurement is not valid, which illustrates also resistance to soft biometric feature estimation noise.

However, even if the aforementioned algorithms are different in essence, Fig. 2(g) and 2(h) provide comparative results of the proposed framework and the approach in [7]. It should be mentioned that the weights needed in [7] are optimally selected by exhaustively searching in the parameter space for the optimal results. The improved efficiency of the proposed scheme is considered as a reflection of the resistance of the proposed approach to soft-biometric features estimation noise and the fact that only soft biometric features with high discrimination efficiency, i.e. minor clusters, are considered.

B. Authentication

Concerning authentication the false acceptance (FAR) and false rejection ratios (FRR) are extracted and illustrated in Fig. 2(d) and 2(e) for the ACTIBIO database using for gait feature extraction the algorithms of [10] and [12] respectively, while Fig. 2(f) depicts the FAR-FRR curves in the HUMABIO database. It should be mentioned that the proposed framework manages to decrease the FAR and FRR in the equal error rate EER point (4% to 15% decrement) in the ACTIBIO database depending on the soft biometrics used. Moreover, the increment in performance becomes more notable for difficult application scenarios where the state-of-the-art gait recognition and authentication scenarios cannot achieve very high recognition and authentication rates.

V. Conclusions and Future Work

An efficient framework for augmenting gait recognition algorithms with the height and stride length soft biometrics has been proposed. The advantage of this framework is that no fusion at the score level is needed, while it can also be directly applied to any gait recognition algorithm. Experimental validation illustrates the efficiency of the framework in different databases and using different algorithms for gait feature extraction.

References

[1] A. Jain, S. Dass, and K. Nandakumar, “Can soft biometric traits assist user recognition?,” *Proc. SPIE*, vol. 5404, pp. 561–572, 2004, Proc. SPIE, 2004.

[2] D. K. Wagg and M. S. Nixon, “On automated model-based extraction and analysis of Gait,” in *Proc. 6th IEEE Int. Conf. Automatic Face and Gesture Recognition*, Seoul, Korea, May 2004, pp. 11–16.

[3] L. Wang, H. Ning, T. Tan, and W. Hu, “Fusion of static and dynamic body biometrics for gait recognition,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 2, pp. 149–158, Feb. 2004.

[4] N. Boulgouris, K. Plataniotis, and D. Hatzinakos, “Gait recognition using linear time normalization,” *Pattern Recognit.*, vol. 39, pp. 969–979, 2006.

[5] S. Sarkar, P. Phillips, Z. Liu, I. Robledo-Vega, P. Grother, and K. Bowyer, “The human id gait challenge problem: Data sets, performance, and analysis,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 2, pp. 162–177, Feb. 2005.

[6] N. Boulgouris, K. Plataniotis, and D. Hatzinakos, “Gait recognition using dynamic time warping,” in *Proc. IEEE 6th Workshop on Multimedia Signal Processing*, Sep. 2004, pp. 263–266.

[7] A. Jain, S. Dass, and K. Nandakumar, “Soft biometric traits for personal recognition systems,” in *Proc. Int. Conf. Biometric Authentication*, Jul. 2004, pp. 731–738.

[8] J. Wayman, “Large scale civilian biometric systems,” in *Proc. CardTech/SecurTech Government*, 1997 [Online]. Available: http://www.engr.sjsu.edu/biometrics/nbtccw.pdf.

[9] G. Marcialis, F. Roli, and D. Muntoni, “Group-specific face verification using soft biometrics,” *J. Vis. Lang. Comput.*, vol. 20, pp. 101–109, 2009.

[10] D. Ioannidis, D. Tzovaras, I. Damousis, S. Argyropoulos, and K. Moustakas, “Gait recognition using compact feature extraction transforms and depth information,” *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 3, pp. 623–630, Sep. 2007.

[11] N. Boulgouris and Z. Chi, “Gait recognition using radon transform and linear discriminant analysis,” *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 731–740, Mar. 2007.

[12] J. Han and B. Bhanu, “Statistical Feature Fusion for Gait-Based Human Recognition,” in *IEEE Computer Soc. Conf. Computer Vision and Pattern Recognition*, 2004, pp. 842–847.