EFFICIENT GAIT-BASED GENDER CLASSIFICATION THROUGH FEATURE SELECTION*

Raúl Martín-Félez, Javier Ortells, Ramón A. Mollineda and J. Salvador Sánchez
Institute of New Imaging Technologies and Dept. Llenguatges i Sistemes Informàtics
Universitat Jaume I. Av. Sos Baynat s/n, 12071, Castelló de la Plana, Spain

Keywords: Gender classification, Gait, ANOVA, Feature selection.

Abstract: Apart from human recognition, gait has lately become a promising biometric feature also useful for prediction of gender. One of the most popular methods to represent gait is the well-known Gait Energy Image (GEI), which conducts to a high-dimensional Euclidean space where many features are irrelevant. In this paper, the problem of selecting the most relevant GEI features for gender classification is addressed. In particular, an ANOVA-based algorithm is used to measure the discriminative power of each GEI pixel. Then, a binary mask is built from the few most significant pixels in order to project a given GEI onto a reduced feature pattern. Experiments over two large gait databases show that this method leads to similar recognition rates to those of using the complete GEI, but with a drastic dimensionality reduction. As a result, a much more efficient gender classification model regarding both computing time and storage requirements is obtained.

1 INTRODUCTION

The last decades have witnessed the wide study of *gait* as a novel and appealing biometric feature. It mainly consists of recognizing people by their particular manner of walking, what is a human skill proved by early psychological studies (Johansson, 1975; Cutting and Kozlowski, 1977). But in addition, humans are also able to distinguish the gender of a person by their gait (Kozlowski and Cutting, 1977; Davis and Gao, 2004). In general, a person's gender is more accurately appreciated from a face or a voice, but gait allows to obtain this information at a distance, in a non-contact and non-invasive way, and without requiring the subject's willingness. These advantages have stirred up the interest of the computer vision community for conceiving gait-based gender recognition systems (Huang and Wang, 2007; Li et al., 2008; Yu et al., 2009) since a number of applications may benefit from the development of such systems: demographic analysis of a population, systems to analyse the customer's behaviour at supermarkets, advanced interaction of robots, etc. Nevertheless, there are also important weaknesses that hinder the use of gait. For instance, gait analysis is very sensitive to deficient segmentation of silhouettes or variations in clothing, footwear, walking speed, carrying conditions, etc.

In the computer vision literature, two main approaches to describe gait can be found. The model-based methods (Davis and Gao, 2004; Yoo et al., 2005; Huang and Wang, 2007) extract *dynamic features* from subject's movements by matching the joint locations with a kinematic model of the human body. However, the free-model techniques (Han and Bhanu, 2006; Li et al., 2008; Yu et al., 2009; Makihara et al., 2011) obtain *static attributes* related to the subject's appearance from a sequence of silhouettes, what implicitly might contain motion information.

In general, model-free methods have lower computational cost than model-based ones and allow to acquire features in a easier way. One of the most commonly used methods of this approach is *Gait Energy Image* (GEI) (Han and Bhanu, 2006). It has been proved to be a robust gait descriptor in different classification tasks, such as gender classification (Li et al., 2008; Yu et al., 2009) and human recognition (Han and Bhanu, 2006; Bashir et al., 2008). It consists of obtaining an average silhouette image to represent both body shape and movements over a gait cycle.

*This work has been partially supported by projects CSD2007-00018 and CICYT TIN2009-14205-C04-04 from the Spanish Ministry of Innovation and Science, P1-1B2009-04 from Fundació Bancaixa and PREDOC/2008/04 grant from Universitat Jaume I. The CASIA Gait Database collected by Institute of Automation, Chinese Academy of Sciences has been used in this paper.

Martín- Félez R., Ortells J., A. Mollineda R. and Salvador Sánchez J. (2012).
EFFICIENT GAIT-BASED GENDER CLASSIFICATION THROUGH FEATURE SELECTION.
In *Proceedings of the 1st International Conference on Pattern Recognition Applications and Methods*, pages 419-424
DOI: 10.5220/0003774404190424

Some studies have demonstrated that several body parts depicted in a GEI have a higher discriminative power than others for gender classification. For example, (Li et al., 2008) proposed to separate human silhouettes into seven components (head, arm, trunk, thigh, front leg, back leg, and feet). Results gave the arm (which includes chest) and front leg as the most discriminative parts. On the other hand, (Yu et al., 2009) described a better segmentation of body components based on results of a psychological study, and suggested the hairstyle, back and thigh regions as the most discriminative body parts. Another recent study (Makihara et al., 2011) uses frequency domain features to obtain similar results: the hair, back, breast and legs seem to be the most significant body parts.

(Bashir et al., 2008) presented an attempt to reduce the high dimensionality of GEI by means of feature selection methods to segment only the dynamic features, since static features are more affected by covariate factors such as clothing changes or carrying conditions. As the number of selected features was still high, they used a combination of Principal Component Analysis (PCA) and Multiple Discriminant Analysis (MDA) to reduce the dimensionality. Results prove that a better performance can be achieved by exploiting the discriminative information of GEI.

Based on conclusions of previous works, in this paper, a simple methodology to select and exploit the most discriminative features of GEI is proposed and applied to gender classification. Firstly, the discriminative power of each pixel (feature) in a GEI is obtained through an analysis of variance (ANOVA) for the whole gallery set. Then, a binary mask is obtained by retaining only those pixels having the highest ANOVA values. Given an unknown probe sample, it is projected onto the binary mask, and a lower dimensional representation is obtained. Experiments on two large gait databases prove that small discriminative feature subsets lead to similar classification accuracies than that obtained from the original GEI, but with a drastic reduction of the dimensionality. It produces an important improvement in terms of computational and storage costs.

2 BACKGROUND

In this section, the basic concepts of GEI representation and analysis of variance (ANOVA) are described.

2.1 Gait Energy Image (GEI)

This well-known free-model method was proposed by (Han and Bhanu, 2006). It basically creates an *average silhouette* for a gait sequence, which reflects the shape of the body parts and, to some extent, their changes over time (gait dynamics). In the resulting image (GEI), the higher the intensity of a pixel is, the more time that pixel belongs to subject silhouettes accross the gait sequence. The main advantages of this method are its robustness to silhouette noises and the reduction of storage and time requeriments since a unique image is used to represent a whole sequence.

In order to create a GEI given a gait sequence, its frames must be preprocessed as follows: 1) foreground (a silhouette) is segmented from background; 2) the bounding box enclosing all silhouette pixels is extracted as a new cropped silhouette image; 3) this image is scaled to a new one having a prefixed common height and a variable width to keep its particular aspect ratio; 4) all normalized silhouettes are horizontally centered from its upper-half horizontal centroid; 5) given the set of preprocessed silhouettes of a gait sequence $\{I_t(x,y)\}$ with $1 \leq t \leq N$, N being the number of silhouettes, and (x,y) referring to a specific position in the 2D image space, each gray-level pixel of a GEI is computed as in Eq. 1.

$$GEI(x,y) = \frac{1}{N}\sum_{t=1}^{N} I_t(x,y) \quad (1)$$

2.2 Analysis of Variance (ANOVA)

The ANOVA F-statistic (Brown and Forsythe, 1974) is a measure that assesses the discriminative capability of several features in an independent way. It is calculated as in Eq. 2, where x_{ij} is the j^{th} sample of class i, c is the number of classes, n_i is the number of samples of class i, $n = \sum_{i=1}^{c} n_i$, $\bar{x}_i$ is the mean of samples in class i, and $\bar{x}$ is the mean of $\bar{x}_i$. The greater the F-statistic value for a feature is, the better its discriminative capability.

$$F = \frac{\frac{1}{c-1}\sum_{i=1}^{c} n_i\left(\bar{x}_i - \bar{x}\right)^2}{\frac{1}{n-c}\sum_{i=1}^{c}\sum_{j=1}^{n_i}\left(x_{ij} - \bar{x}_i\right)^2} \quad (2)$$

A study to measure the relevance of each GEI pixel to discriminate gender was performed in other works (Makihara et al., 2011; Yu et al., 2009) by analysing their variance through ANOVA. They stated that the most discriminative pixels are those ones located in the head/neck (because of the hair style), the back region (due to the thinner body trunks of women in comparison with men), and the thigh region.

They made a particular interpretation of ANOVA that has been replicated in this work. For Eq. 2, $c = 2$ because there are two classes (men and women), x_{ij} is each particular GEI from a collection, $\bar{x}$ is the resulting image of averaging all GEIs, and $\overline{x_i}$ is a mean im-

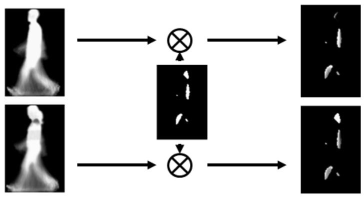

Figure 1: Example of a man (top) and a woman (bottom) projected onto the ANOVA-based binary mask.

age obtained from all GEIs of a particular class (men or women). The computation of the F value is made at pixel level since each GEI pixel is considered as an independent feature. As a result, an ANOVA F matrix is obtained, in which each value $F(x,y)$ corresponds to the relevance of the GEI pixel (x,y) to discriminate between men and women. This matrix is transformed into a gray-scale image by normalizing its values in the range $[0,255]$. In this image, the whitest pixels are those with a better discriminative capability, i.e., those pixels with a large variance between genders and a relatively small variance within each gender. On the other hand, the darkest values represent irrelevant pixels.

3 METHODOLOGY

In this section, a general overview of the method is introduced. The main goal of the proposal is to explore the potential of a novel ANOVA-based feature selection technique, rather than to provide a comprehensive study about feature reduction in GEI.

Given a gallery (training) set and a probe (test) set composed of two different collections of gait sequences, the *learning procedure* is as follows:

- For each gallery gait sequence, its corresponding GEI is computed as explained in Section 2.1.
- ANOVA is computed from all GEIs in gallery.
- A small number of the most discriminative pixels is selected from the ANOVA results. This subspace is represented by a binary mask image.
- Each gallery GEI is projected onto the binary mask to obtain a reduced description of the GEI, as depicts Figure 1. The resulting gallery is used to train a classifier.

In the *evaluation process*, given a probe gait sequence, its corresponding GEI is computed and projected onto the binary mask to obtain its reduced description. Finally, a gender decision is made for this reduced representation by the classifier.

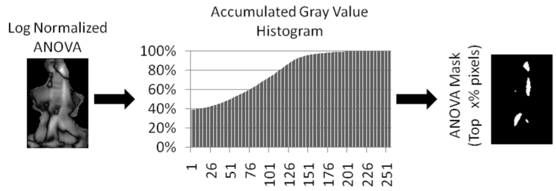

Figure 2: Generation of ANOVA-based binary mask.

3.1 ANOVA-based Feature Selection Method

The novel idea proposed in this paper is to create a binary mask image M, where the most discriminative GEI pixels for gender classification according to ANOVA are highlighted. Firstly, since ANOVA values are not normally distributed and a few very high values disturb the representation, a logarithmic normalization with basis 2 is proposed to overcome this problem. Then, the log normalized ANOVA matrix is transformed into a gray-scale image. M is defined as in Eq. 3, where χ denotes the percentage of most discriminative log normalized ANOVA pixels to be selected and $\theta(\chi)$ is the lowest gray value for all the selected pixels. A useful tool to obtain $\theta(\chi)$ is an accumulated gray-value histogram with the 255 gray possible levels in the x-axis, and the accumulated percentage of ANOVA pixels in the y-axis. An example can be seen in Figure 2.

$$M(x,y) = \begin{cases} 1 & \text{if } LogNorm\_ANOVA(x,y) \geq \theta(\chi), \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

4 EXPERIMENTS AND RESULTS

The experiments have been addressed to assess the effectiveness of the new method for a gender classification task in comparison to that provided by the plain use of the original GEI method. Besides, they also aim at measuring the influence of the parameter χ, which represents the percentage of more significant ANOVA pixels that are highlighted in the binary mask M, in terms of classification performance and reduction of time and storage costs.

4.1 Databases and Preprocessing

Experiments have been carried out on two public large gait databases. One of them is CASIA Gait Database (CASIA) (CASIA, 2005) - Dataset B, which consists of indoor videos of 124 subjects captured from different viewpoints, and some sequences include changes in clothing and carrying conditions. This database is unequally distributed as concerns gender and contains samples of 93 men and 31 women, which gives an imbalanced ratio of 3:1. For each subject, only their six side-view gait sequences without changes in clothing or carrying conditions have been used in the present experiments, what gives a total of 744 sequences.

The second gait database is the Southampton HID Database (SOTON) (Shutler et al., 2002) - Subset A. It is composed of indoor videos of 115 subjects filmed from a side view without any covariate condition. This database is also imbalanced regarding gender, containing samples of 91 men and 24 women (an imbalanced ratio of 4:1). Since a different number of sequences per subject ranging from 6 to 42 is available, a total number of 2162 sequences have been considered in experiments.

Both databases provide well-segmented foreground images that have been used as inputs to the silhouette extraction step (see Section 2.1). Then, these silhouettes have been scaled and horizontally aligned to fit an image template of 128×88 pixels, which have been the basis to compute the different GEIs.

4.2 Performance Evaluation Protocol

A stratified 5-fold cross validation scheme is repeated five times to estimate the recognition rates and to reduce the impact of subset singularities. In addition, all samples of a subject are put in the same fold, i.e., when a person's gait sequence is in the probe set, none of their sequences are used for training. Each gallery fold feeds a classifier that later performs a classification session on the corresponding probe fold. The performance of the classification is assessed through unbiased performance measures (see Section 4.4) in order to avoid misleading results.

4.3 Classifier Setting

Two different classifiers have been used to test the potential of the new method. One of them is a Support Vector Machine (SVM) classifier, which was selected because of its common high performance in two-class problems with few samples (Boser et al., 1992). Like in previous works (Li et al., 2008; Yu et al., 2009), a linear kernel with $C = 1$ has been used.

On the other hand, a Nearest Neighbour (1NN) rule has been used because it is the simplest and most commonly used supervised classifier. The Euclidean distance has been adopted to measure the similarity between probe and gallery gait sequences.

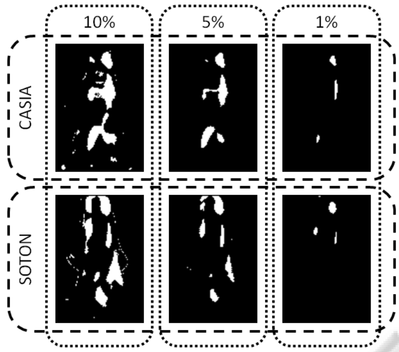

Figure 3: Samples of ANOVA-based masks for both databases.

4.4 Unbiased Performance Measures

Most of the performace measures for a two-class problem can be derived from a 2×2 confusion matrix, which is defined by the numbers of positive and negative samples correctly classified (TP and TN respectively), and the numbers of misclassified positive and negative samples (FN and FP respectively). For example, the overall accuracy is computed as $Acc = (TP + TN)/(TP + FN + TN + FP)$, but there are empirical evidence claiming that this measure can be strongly biased with respect to class imbalance and proportions of correct and incorrect classifications (Provost and Fawcett, 1997). Thus, some perfomance measures solving the shortcomings of accuracy are defined as follows:

- *True Positive Rate (TPr)* is the percentage of positive (women) samples that are correctly classified, $TPr = TP/(TP + FN)$.
- *True Negative Rate (TNr)* is the percentage of negative (men) samples that are correctly classified, $TNr = TN/(TN + FP)$.
- *Gmean* (or *Geometric Mean*) uses the accuracies separately measured on each class, $Gmean = \sqrt{TPr * TNr}$. The aim is to maximize the accuracy on both classes while keeping their accuracies balanced.

Unlike previous related works (Huang and Wang, 2007; Li et al., 2008; Yu et al., 2009) where only the overall accuracy has been used, in this work three measures are calculated to provide more reliable unbiased results. TPr and TNr give the individual class performances, and *Gmean* provides a global unbiased information about the system performance.

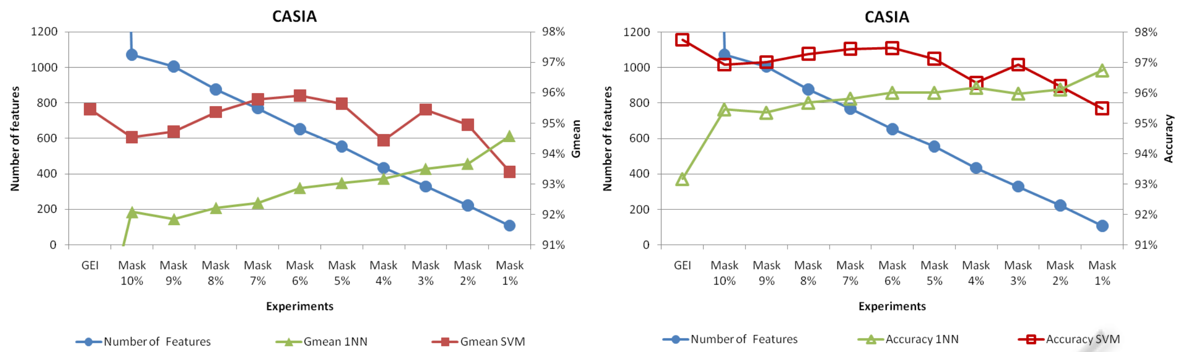

Figure 4: Results from CASIA database using Gmean (left) and Accuracy (right).

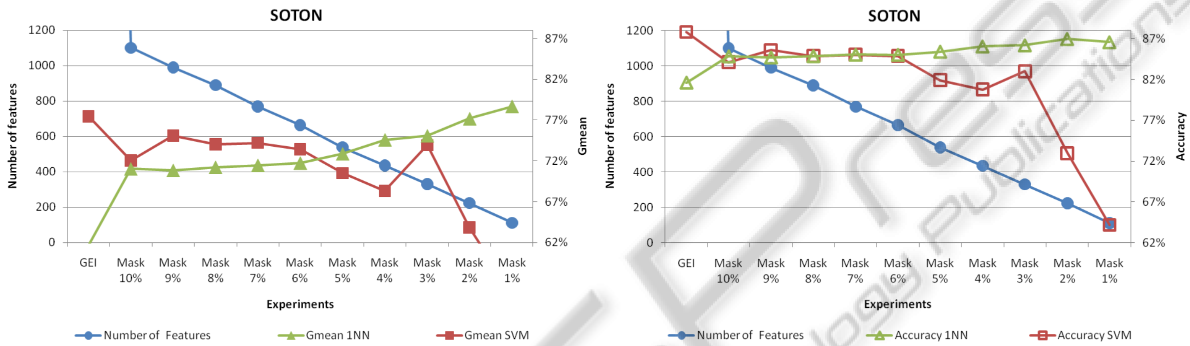

Figure 5: Results from SOTON database using Gmean (left) and Accuracy (right).

4.5 Discussion of Results

Several experiments to assess the best value for the parameter χ have been carried out on both databases. Its value has ranged from 10% to 1% with a decrement of a 1% in each experiment. The particular case of $\chi = 10\%$ means that the binary mask is highlighting the 10% of pixels with the highest log normalized ANOVA intensity values. In other words, given the accumulated gray value histogram of the log normalized ANOVA, the gray value $\theta(\chi)$ corresponding to $100\% - \chi$ is extracted, and then all log normalized ANOVA pixels with gray value equal or greater than $\theta(\chi)$ are highlighted in the corresponding place of the binary mask.

An example of the binary mask evolution depending on the value of the parameter χ is shown in Figure 3 for both databases. From the analysis of these images, some comments can be pointed out:

- The most discriminative GEI pixels for gender classification are in the hair style and the back and thigh regions, what matches with conclusions in (Makihara et al., 2011; Yu et al., 2009).
- The chest region appears as relevant for SOTON, but irrelevant for CASIA. A possible reason might be that Asian women (in CASIA) usually have an average breast cup size smaller than that for the European women (in SOTON).
- Some parts of the trunk are selected as discriminative features in CASIA. It might be due to its recording conditions, which produce some noisy silhouettes with holes in the trunk area that are erroneously taken as significant features.
- The area between the legs is highlighted as discriminative for SOTON but not for CASIA, what probably is due to the greater variety of clothing styles in SOTON.

Figures 4 and 5 show the results for the original GEI method and the approach here introduced as a function of the parameter χ. Each figure shows the different experiments in the x-axis, the number of features in the left y-axis, and the classification performance in terms of either accuracy or geometric mean in the right y-axis. Both performance measures have been despicted to demonstrate the biased results of accuracy in comparison with those reported of *Gmean*.

From the general analysis of these figures, some conclusions can be drawn:

- The number of features for the original GEI method is $128 \times 88 = 11264$ pixels. However, the dimensionality with the proposed method is drastically reduced ranging from approximately 1100 features with $\chi = 10\%$ to about 100 features with $\chi = 1\%$. It produces a strong improvement in terms of time and storage costs.

- *Gmean* results are more reliable than those of *Acc* because both databases have an imbalanced number of gait sequences corresponding to each gender. For example, for the particular case of SOTON and 1NN classifier, a relatively high accuracy of about 80% is obtained with the original GEI method, but it hides an inadmissible 40% of success on the women class with an almost perfect classification rate for the men class (98%). In this case, *Gmean* better represents this biased behaviour with a low value of about 60%.
- The proposed method obtains *Gmean* and *Acc* values similar to those of the original GEI, but with a drastic dimensionality reduction. For SVM classifier, the best balanced trade-off between performance and number of features might correspond to the 3% Mask. However, for 1NN classifier, the best solution is that with the highest dimensionality reduction (1% Mask). This classifier shows a tendency to improve its results as the number of features decreases. In fact, the worst results correspond to the original GEI, i.e., with all features. A possible reason is that the number of samples is too low in comparison with the number of features, what produces a very spread feature space.

By considering both performance and number of features, the best solution is probably that in which 1NN classifier is used with the 1% Mask, since differences with the best SVM results are not significant.

5 CONCLUSIONS

In this paper, an ANOVA-based algorithm has been used to select the most relevant GEI features for gender classification. The experiments carried out on two large databases with a SVM and a 1NN classifiers have shown that a similar performance to that of the original GEI can be achieved by using only its most discriminative information, what leads to an important reduction in computing time and storage requirements. In particular, the 1NN approach has obtained the highest success rates (comparable or better than those of SVM) with the lowest number of features.

With respect to future work, a more comprehensive study including other feature selection/extraction methods should be addressed. In addition, the parameter χ should be automatically estimated by using a validation set, since its value might depend on the singularities of the gallery set. In SOTON, the higher imbalanced ratio produces worse results, thus some techniques to deal with imbalance should be applied in order to improve the overall *Gmean* result.

REFERENCES

Bashir, K., Xiang, T., and Gong, S. (2008). Feature selection on gait energy image for human identification. In *Proc. IEEE Int'l Conf. on Acoustics, Speech and Signal Processing, ICASSP 2008*, pages 985–988.

Boser, B. E., Guyon, I. M., and Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *Proc. 5th Annual Workshop on Computational Learning Theory*, pages 144–152.

Brown, M. B. and Forsythe, A. B. (1974). The small sample behavior of some statistics which test the equality of several means. *Technometrics*, 16(1):129–132.

CASIA (2005). CASIA Gait Database. http://www.sinobiometrics.com.

Cutting, J. and Kozlowski, L. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 9:353–356.

Davis, J. and Gao, H. (2004). Gender recognition from walking movements using adaptive three-mode PCA. In *IEEE CVPR, Workshop on Articulated and Nonrigid Motion*, volume 1.

Han, J. and Bhanu, B. (2006). Individual recognition using gait energy image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(2):316–322.

Huang, G. and Wang, Y. (2007). Gender classification based on fusion of multi-view gait sequences. In *Proc. 8th Asian Conf. Computer Vision*, volume 1.

Johansson, G. (1975). Visual motion perception. *Scientific American*, 6(232):76–80.

Kozlowski, L. and Cutting, J. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, 21:575–580.

Li, X., Maybank, S., Yan, S., Tao, D., and Xu, D. (2008). Gait components and their application to gender recognition. *IEEE Trans. SMC-C*, 38(2):145–155.

Makihara, Y., Mannami, H., and Yagi, Y. (2011). Gait analysis of gender and age using a large-scale multi-view gait database. *LNCS. Computer Vision ACCV 2010*, 6493:440–451.

Provost, F. and Fawcett, T. (1997). Analysis and visualization of classifier performance: Comparison under imprecise class and cost distributions. In *Proc. of the 3rd ACM SIGKDD*, pages 43–48.

Shutler, J., Grant, M., Nixon, M. S., and Carter, J. N. (2002). On a large sequence-based human gait database. In *Proc. 4th Int'l Conf. on RASC*, pages 66–71.

Yoo, J., Hwang, D., and Nixon, M. (2005). Gender classification in human gait using support vector machine. In *Proc. ACIVS*, pages 138–145.

Yu, S., Tan, T., Huang, K., Jia, K., and Wu, X. (2009). A study on gait-based gender classification. *IEEE Transactions on Image Processing*, 18(8):1905–1910.