

Model-based approaches for predicting gait changes over time

Galina V. Veres, Mark S. Nixon and John N. Carter

Department of Electronics and Computer Science
University of Southampton, Southampton, SO17 1BJ

Abstract. Interest in automated biometrics continues to increase, but has little consideration of time. This paper deals with a problem of recognition by gait when time-dependent and time-invariant covariates are added. We have shown previously how recognition rates fall significantly for data captured over lengthy time intervals. We suggest predictive models of changes in gait due both to time and now to time-invariant covariates. A considerable improvement in recognition capability is demonstrated, with potential generic biometric application.

1 Introduction

Recently much attention has been devoted to use of human gait patterns as biometric. Gait recognition aims to discriminate individuals by the way they walk. Approaches to gait recognition can be broadly classified as being model-based and model-free. Model-based methods [2, 3] model the human body structure. Model-free methods [1, 8] generally characterise the whole motion pattern of the human body. In this paper we employ the model-free (static) method of Veres et al [6]. However, in these works only databases recorded over a short time interval were evaluated.

In this paper we consider a gait recognition problem when two databases (the gallery and probe) were recorded with a time interval of 6 and 12 months between the finish of recording the first database (gallery) and the start of recording the second database (probe), i.e. time-dependent covariates are added. Moreover, some extra covariates were added in the second database such as different shoes, clothes, carrying different bags. It is shown that correct classification rates fall significantly and recognition becomes unreliable over time. Similar results have been obtained for the HumanID Gait Challenge Problem [4], where recognition collapsed from 82% to 6% for data acquiring at an interval of 6 months.

Under the assumptions that similar people have similar changes in gait, several predictive models of gait changes are suggested in this paper. Three databases were analysed in the paper. To estimate the effects of time on gait recognition a new time-dependent predictive model is suggested. Time-dependent predictive matrices are estimated for every subject in the training set. Then these matrices are used to predict the possible changes in the gallery of the test set. If a subject in the gallery of the test set does not belong to the gallery in the training set, then the nearest neighbour from the training set is found. The

predictive matrix corresponding to the nearest neighbour is used and the difference is included in prediction of the test gallery. Time-invariant covariates are now included in the time-invariant predictive model. Our earlier approach predicted the time-dependent covariate [7]. However we did not take into account time-invariant covariates. In this paper we try to predict changes in gait due both to time-variant covariates and to time-invariant covariates. In this case the time-invariant predictive matrices are calculated which represent dependency between normal walk of a subject and different possible extra covariates of the same subject. The combined predictive model now takes into consideration both time-dependent and time-invariant covariates of gait. The models suggested in the paper are linear, since it is the first approach to handling changes in gait over a lengthy time interval and it is a reasonable initial assumption. We show that CCRs can be increased by several times when using the new predictive models and in some cases more than 90% CCRs were achieved.

Section 2 describes the suggested predictive models for changes in gait over time and the new combined model. The database description is presented in Section 3. Experimental evaluation is presented and described in Section 4. Section 5 concludes this paper.

2 Prediction of Gait Changes

We assume that it is possible to predict the gallery over the given time interval and achieve good recognition results by analysing the probe via the predicted gallery. In this case the training set consists of a set of subjects from the gallery and the same set of subjects from the probe. The test set consists of different set of the same subjects, possible extra subjects and possible extra experiments from the gallery and the probe.

2.1 Prediction of Time-dependent Covariates

In general case the predicted subject from the gallery can be defined as

$$\hat{\mathbf{g}}_{t+1}(i) = f(\mathbf{g}_t(i)), \quad (1)$$

where $\hat{\mathbf{g}}_{t+1}(i)$ is the i th predicted subject from a gallery, $\mathbf{g}_t(i)$ is the i th subject in the gallery, f is a prediction function for i th subject. We use a prediction function f in the form

$$\hat{\mathbf{g}}_{t+1}(i) = \mathbf{g}_t(i)\mathbf{W}_t(i), \quad (2)$$

where $\mathbf{W}_t(i)$ is $N_f \times N_f$ time-dependent predictive matrix for k th subject, N_f is a number of features for each subject.

Let the gallery and the probe be divided into groups (subjects). Let us consider at first the case when the number of groups in the gallery equals the number of groups in the probe and the groups are the same, i.e. training set. At first for each group in the probe and gallery the mean of the group is calculated

$$\bar{\mathbf{x}}_t^j(p) = \frac{1}{n_p^j} \sum_{i=1}^{n_p^j} \mathbf{x}_t^j(pi) \quad \text{and} \quad \bar{\mathbf{x}}_t^j(g) = \frac{1}{n_g^j} \sum_{i=1}^{n_g^j} \mathbf{x}_t^j(gi), \quad (3)$$

where $\bar{\mathbf{x}}_t^j(p)$ and $\bar{\mathbf{x}}_t^j(g)$ are the means of group j in the probe and the gallery respectively, $j = 1, \dots, n_g$, where n_g is a number of groups, n_p^j and n_g^j is a number of records in the j th group of the probe and of the gallery, respectively, $\mathbf{x}_t^j(pi)$ and $\mathbf{x}_t^j(gi)$ are records for i th subject in j th group in the probe and gallery respectively.

Then the time-dependent predictive matrix for each group j is calculated as

$$\mathbf{W}_t(j) = (\bar{\mathbf{x}}_t^j(g))^{-1} \bar{\mathbf{x}}_t^j(p), \quad (4)$$

where the inverse is calculated by pseudoinverse. Formula (4) allows to find changes in the j th subject's gait in two different time instances and can be derived from the equation

$$\bar{\mathbf{x}}_t^j(p) = \bar{\mathbf{x}}_t^j(g) \mathbf{W}_t(j), \quad (5)$$

where the gallery $\bar{\mathbf{x}}_t^j(g)$ was recorded in one time instance; the probe $\bar{\mathbf{x}}_t^j(p)$ in the next (future) time instance; the predictive matrix $\mathbf{W}_t(j)$ represents changes from the subject in the gallery to the subject in the probe due to time.

The i th predicted gallery subject from group j is

$$\hat{\mathbf{g}}_{t+1}(i) = (\mathbf{g}_t(i) + |\mathbf{g}_t(i) - \bar{\mathbf{x}}_t^j(g)|) \mathbf{W}_t(j), \quad (6)$$

where $|\mathbf{g}_t(i) - \bar{\mathbf{x}}_t^j(g)|$ describes deviation of the subject in the gallery from the mean of the group to which this subject belongs; $\mathbf{g}_t(i) + |\mathbf{g}_t(i) - \bar{\mathbf{x}}_t^j(g)|$ is a correction due to deviation from the group mean.

With obtaining $\mathbf{W}_t(j)$ for each group in the training set the training stage is finished and the time-dependent predictive matrices $\mathbf{W}_t(j)$ are saved for future use on the test set.

For the more general case (test set) when a number of groups in the gallery is not the same as number of groups in the probe and/or subjects are not the same both in the gallery and in the probe, the predictive matrix is constructed as follows. Here we present a case when a number of groups in the gallery is more than a number of groups in the probe. Two assumptions are made in this case

1. Every subject in the probe exists in the gallery.
2. The gait of the similar subjects will change in a similar manner with time.

We are looking forward to gathering more data to provide a theoretical analysis or statistical observation to support the second assumption.

If the same subject is presented both in the gallery of the training set and the gallery of the testing set, then group j which the subject from the test set belongs to is recorded and this subject is predicted. If the subject in the test set $\mathbf{g}_t^{test}(i)$ is not in the training set, then the nearest neighbour from the training set is found by finding r such as

$$r = j : \{\min_j |\mathbf{g}_t(j) - \mathbf{g}_t^{test}(i)|, j = 1, \dots, N_g\}, \quad (7)$$

where N_g is a number of records in the gallery of the training set. Then i th predicted subject is calculated as

$$\hat{\mathbf{g}}_{t+1}^{test}(i) = (\mathbf{g}_t^{test}(i) + |\mathbf{g}_t^{test}(i) - \bar{\mathbf{x}}_t^{j(r)}(g)|) \mathbf{W}_t(j(r)), \quad (8)$$

where $j(r)$ means that group j is chosen according to r th record of the training set.

After the predicted gallery is calculated as (6) or (8), the probe is classified via the predicted gallery. Prediction of time-invariant covariates is presented in the next subsection.

2.2 Prediction of Other Covariates

In this subsection we consider the dataset recorded over a short time interval when time will have little effect on recognition capability, and the effect of other covariates is much more likely. However, the subject cannot only walk normally but have some extra covariates like different clothes, shoes, wearing bags etc. In this case the gallery consists of subjects walking normally and the probe consists of records when some changes to normal walk are added. Then dependency between normal walk and any extra covariate can be expressed through the prediction of the gallery as

$$\hat{\mathbf{g}}_t^e(k) = f^e(\mathbf{g}_t(k)), \text{ or } \hat{\mathbf{g}}_t^e(k) = \mathbf{g}_t(k) \mathbf{W}^e, \quad (9)$$

where f^e is a function of changes according to e th experiment, \mathbf{W}^e is a square predictive matrix defining differences between normal walk and e th experiment.

The set of time-invariant predictive matrices \mathbf{W}^e can be calculated using training set, and used when there is a possibility of extra covariates. In this case the training set is divided into groups according to experiments and the means of each experiment for each subject is calculated as

$$\bar{\mathbf{x}}_t^e(k) = \frac{1}{n^{ek}} \sum_{i=1}^{n^{ek}} \mathbf{x}_t^{ek}(i), \quad (10)$$

where $\bar{\mathbf{x}}_t^e(k)$ is mean of e th experiment for k th subject and $\mathbf{x}_t^{ek}(i)$ is i th record in e th group for k th subject. Then the time-invariant predictive matrices \mathbf{W}^e can be estimated as

$$\mathbf{W}^e(k) = (\bar{\mathbf{x}}_t^0(k))^{-1} \bar{\mathbf{x}}_t^e(k), \quad (11)$$

where $\bar{\mathbf{x}}_t^0(k)$ represents mean of normal walk for k th subject and $\bar{\mathbf{x}}_t^e(k)$ are means of all other experiments for k th walk, since the dependency between the normal walk and extra time-invariant covariates can be presented as

$$\bar{\mathbf{x}}_t^e(k) = \bar{\mathbf{x}}_t^0(k) \mathbf{W}^e(k). \quad (12)$$

Then the predicted gallery for k subject in the e th experiment will be

$$\hat{\mathbf{g}}_t^e(k) = \mathbf{g}_t(k) \mathbf{W}^e(k). \quad (13)$$

After the predictive matrices are obtained, the predicted gallery is calculated as (13) and the probe is classified via the predicted gallery. The combination of time-variant and non-time dependent covariates of gait is presented in the next subsection.

2.3 Combination of Different Gait Changes

So far we consider changes of gait either due to time or extra covariates separately. In this subsection we consider the case when both time and extra covariates are presented. Then the predicted gallery can be presented as

$$\hat{\mathbf{g}}_{t+1}^e(k) = f^e(f(\mathbf{g}_t(k))), \quad (14)$$

where f is function defining time-dependent changes and f^e defines changes due to e th experiment for k th subject. Combining formulas (6), (8) and (13) the combined predicted gallery is

$$\hat{\mathbf{g}}_{t+1}^e(k) = (\mathbf{g}_t(k) + |\mathbf{g}_t(k) - \bar{\mathbf{x}}_t^j(g)|) \mathbf{W}_t(j) \mathbf{W}^e(k), \quad (15)$$

if k th subject exists both in the training and test sets or

$$\hat{\mathbf{g}}_{t+1}^{test,e}(k) = (\mathbf{g}_t^{test}(i) + |\mathbf{g}_t^{test}(i) - \bar{\mathbf{x}}_t^{j(r)}(g)|) \times \mathbf{W}_t(j(r)) \mathbf{W}^e(k), \quad (16)$$

if k th subject exists only in the test set.

Using formulas (15) and (16) we can predict the gallery taking into consideration both time-dependent and non time-dependent covariates.

3 Databases Description

Three databases were analysed in the paper, comprising indoor (studio) data. The first database, called the large database (LDB), consists of 115 subjects performing a normal walk. The database arrangements are described elsewhere [5]. The small database (SDB) consists of the sample of 10 subjects from the LDB. Each subject was filmed wearing a variety of footwear, clothes and carrying various bags. They were also filmed walking at different speeds. Each subject's data was captured during one continuous filming session. The temporal database (TDB) also consists of 10 subjects from the LDB in normal conditions, but each subject was re-filmed on different dates. One sequence for each subject consists of 50-60 frames describing gait. Examples of features from the LDB and the SDB are presented in Fig. 1. The figure shows subject 46's average silhouette computed for a sequence in the LDB and in the SDB walking normally, carrying a bag and wearing a trench coat. It worth noticing that sequence of filming data in time was LDB, SDB and TDB with approximately 6 months difference between LDB and SDB, and SDB and TDB.

For brevity we shall not fully describe the extraction methods, complete explanations of the techniques used can be found in [6]. This technique yields a 4096 dimensional feature vector derived from the subject's silhouette accumulated over a sequence of images.

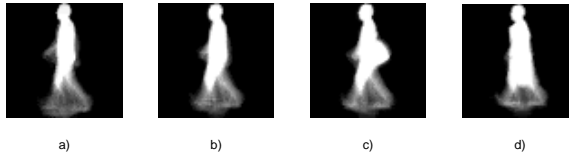


Fig. 1. Average silhouette representing subject 46: a) LDB, b) SDB, walking normally, c) SDB, carrying a bag, d) SDB, wearing trench coat

4 Experimental Evaluation

In this section we present the performance of the suggested predictive models in the tasks of recognizing people over lengthy time interval by gait, i.e when SDB was analysed via LDB S/L , TDB was analysed via SDB T/S and LDB T/L . All three data sets were divided between training set and testing set. The training set consists of 10 subjects belonging to all three databases walking normally for prediction of time-dependent covariates. However only 50% of records corresponding to a given subject in the database go to the training set, the remaining data belong to the testing set together with extra subjects for LDB and extra experiments (time-invariant covariates) for SDB. The system was tested on four different conditions. For the first we performed leave-one-out experiments using all training and test silhouettes walking normally, i.e. S_{10}/L_{10} , T/S_{10} and T/L_{10} . This experiment tested the accuracy of our approach in prediction of gait changes over time in subjects who had already provided training silhouettes. For the second experiment we added 105 new subjects to the gallery in the test set, i.e. S_{10}/L_{all} and T/L_{all} . In third experiment we investigate how adding extra time-invariant covariates will affect the time-dependent predictive model, i.e. S_{all}/L_{10} and T/S_{all} . The last experiment investigates the performance of the different predictive models when extra subjects and extra experiments are added, i.e. S_{all}/L_{all} and T/S_{all} .

The reduced datasets were used for representing all three databases, since it was shown [6] that recognition rates change little if a subset of the features is used. PCA was run on combination of three databases, and 500 features which represent 98.79% variation explained were selected for further analysis. To show the influence of time and not a degradation of data quality on recognition, the CCR is calculated for each database separately both for original and reduce datasets. The results are presented in Table 1 and show acceptable CCRs for three datasets. Moreover, CCRs change little when a number of features reduces from 4096 to 500.

CCRs for analysis of different datasets before training are presented in Table 2. It can be seen that as soon as time-dependent covariates are added to the analysis the fall in CCR is very noticeable, especially when all subjects and all experiments are considered.

CCRs after applying the time-dependent predictive model on training and testing sets are presented in Table 3. It can be seen from Table 3, that the time-

features	L_{all}/L_{all}	S_{all}/S_{all}	T/T
4096	98.47%	99.90%	99.30%
500	97.70%	99.84%	99.62%

Table 1. Analysis of databases without time-dependent covariates

S_{10}/L_{10}	T/S_{10}	T/L_{10}	S_{10}/L_{all}	T/L_{all}	S_{all}/L_{10}	T/S_{all}	S_{all}/L_{all}
69.54%	73.88%	64.67%	44.20%	36.56%	50.88%	63.30%	24.12%

Table 2. CCRs for reduced feature sets before training

dependent predictive model copes very well when either only time-dependent covariates present or extra subjects are added from LDB both for training and testing sets. In this case more than 98% CCRs are achieved for training sets and more than 91% CCRs are achieved for testing sets. Moreover, adding extra covariates to SDB when it is considered as a gallery does not change the recognition capability. However when extra covariates are added to SDB treated as a probe, CCRs decrease significantly.

set	S_{10}/L_{10}	T/S_{10}	T/L_{10}	S_{10}/L_{all}	T/L_{all}	S_{all}/L_{10}	T/S_{all}	S_{all}/L_{all}
training	99.53%	98.74%	98.59%	99.53%	98.59%	99.53%	98.74%	99.53%
testing	91.16%	91.32%	92.13%	91.16%	92.11%	26.46%	91.32%	26.46%

Table 3. CCRs after applying the time-dependent predictive model

The results of using time-invariant predictive model for SDB is presented in Table 4. Table 4 shows that it is possible to calculate the predictive matrices at the training stage and then apply them successfully on the testing stage, thus achieving good recognition performance.

Table 5 shows the results of applying the combined predictive model. The results presented in Table 5 are significantly better than results reported in Table 2.

5 Summary and Conclusions

This paper deals with a problem of increasing correct classification rate when time-dependent covariates together with time-invariant covariates are added to an analysed database for gait recognition. We have shown that CCRs are very low in this case. In this paper we suggest use of the prediction of gait over the given time interval and prediction of possible changes in gait due to time-invariant covariates. The time-variant and time-invariant predictive models were developed. The experimental results showed that good results can be achieved both on the training set and the testing set on four different conditions.

dataset	training	testing
SDB	97.79%	96.57%

Table 4. Prediction of time-invariant covariates

set	S_{all}/L_{10}	S_{all}/L_{all}
training	87.16%	87.16%
testing	85.76%	72.03%

Table 5. CCRs using the combined predictive model

Acknowledgment

We gratefully acknowledge partial support by the Defence Technology Centre 8 – 10 supported by General Dynamics.

References

1. R. Collins, R. Gross, and J. Shi “Silhouette-based human identification from body shape and gait” *Proceedings of the International conference on Automatic Face and Gesture Recognition*, Washington,DC, 2002
2. D. Cunado, M.S. Nixon, and J. N. Carter “Automatic extraction and description of human gait models for recognition purposes” *Computer Vision and Image Understanding*, **90**(1): 1-41, 2003
3. A. Kale, N. Cuntoor, and R. Chellapa “A framework for activity-specific human recognition” *Proceedings of IEEE International conference on Acoustics, Speech and Signal Processing*, Orlando, Fl., 2002
4. S. Sarkar, P.J. Phillips, Z. Liu, I.R. Vega, P. Grother, and K.V. Bowyer “The HumanID Gait Challenge Problem: Data sets, Performances, and Analysis ” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **27**(2): 162-177, 2005.
5. J. D. Shutler, M. G. Grant, M. S. Nixon, and J. N. Carter “On a Large Sequence-Based Human Gait Database” *Proc. 4th International Conf. on Recent Advances in Soft Computing*, Nottingham (UK), 2002
6. G.V. Veres, L. Gordon, J.N. Carter and M.S. Nixon “What image information is important in silhouette-based gait recognition?” *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Washington, D.C., USA, **II**: 776-782 2004.
7. G.V. Veres, M.S. Nixon and J.N. Carter “Modelling the time-variant covariates for gait recognition” *Proceedings of 5th International conference on Audio- and Video-Based Biometric Person Authentication*, New York, USA, 2005.
8. L. Wang , W.M. Hu and T.N. Tan “A new attempt to gait-based human identification” *IEEE Transactions on Circuits and Systems for Video Technology*, **14**(2): 149-158, 2002.