

The Effect of Time on the Performance of Gait Biometrics

Darko S. Matovski, Mark S. Nixon, Sasan Mahmoodi, John N. Carter

Abstract—Many studies have shown that it is possible to recognize people by the way they walk. However, there are a number of covariate factors that affect recognition performance. The time between capturing the gallery and the probe has been reported to affect recognition the most. To date, no study has shown the isolated effect of time, irrespective of other covariates. Here we present the first principled study that examines the effect of elapsed time on gait recognition. Using empirical evidence we have shown for the first time that elapsed time does not affect recognition significantly in the short to medium term. By controlling clothing, a Correct Classification Rate (CCR) of 95% has been achieved over 9 months, on a dataset of nearly 2000 gait sequences/samples. We have created a new multimodal temporal database to enable the research community to investigate various gait and face covariates in a formal manner. Our results show that gait can be used as a reliable biometric over time and at a distance. We have demonstrated that clothing drastically affects performance regardless of elapsed time. A move towards developing appearance invariant recognition algorithms is essential.

I. INTRODUCTION

The way we walk, combined with our posture, is known as *gait*. Many studies have demonstrated that gait has the potential to become a powerful biometric for surveillance and access control. Recently, there has been a rapid growth in the number of surveillance systems in efforts to improve safety and security. These systems have yet to include recognition capability. The major advantages of gait are: non-invasive, hard to conceal and capable of being acquired at a distance. Gait can be detected at a very low resolution and therefore it has a unique advantage over all other biometrics for visual surveillance systems.

There are two main approaches to gait recognition: model-based and model free. Model based approaches use the human body structure [1] and model-free methods use the whole motion pattern of the human body [2]. In this paper we use a model-free approach called Gait Energy Image (GEI) [3] since it is one of the most popular gait representations used so far. There are a number of *covariates* (known as exploratory variables) that affect gait recognition. There are studies quantifying the effect of covariates on the performance of gait recognition algorithms [4] [5]. Covariate factors can be related either to the subject itself (e.g. different clothing) or to the environment (e.g. different walking surface). The research in [6] considers five covariates that

affect recognition: viewing angle, shoe type, walking surface, carrying objects and elapsed time between sequences being compared. Understanding these factors is crucial to developing robust and accurate gait recognition algorithms. The effect of a particular covariate on the recognition performance depends on the algorithm adopted. However, there is a consensus in the literature that the time taken between recording the gallery and the probe affects recognition performance the most [6]. Time as a covariate has not been considered explicitly in many studies mainly due to the lack of suitable database.

Since gait is a behavioural biometric, an important question arises: “Is it possible to recognise someone reliably after a certain period of time has elapsed?” Permanence is an important characteristic of any biometric. A biometric trait that changes significantly over time is unlikely to be useful for recognition.

In previous studies that have considered the effect of elapsed time on recognition by gait, one study [6] reported a substantial drop in recognition performance (at rank 1) from 78% to 3% when shoe and time (6 months) covariates are introduced, using the Gait Challenge database. In [7], a different recognition algorithm is used on the same dataset. Seventy silhouettes are chosen and manually subtracted from the background. The recognition performance over time dropped to 10%. In [1], two different recognition algorithms are employed on MIT database. Experiments are performed over two months in an indoor environment with different backgrounds and lighting. On a dataset of 24 subjects, the recognition rate varies between 30-60%. Another study [8] using the MIT database shows a performance drop to 45% from 100% for samples taken on the same day. Experiments on the UMD database reveals a drop in recognition performance to 30% over 3 months using 25 subjects in indoor environment [9]. In [10] a probe consisting of 10 subjects is matched to a gallery containing 115 subjects. The subjects in the probe are filmed wearing a variety of footwear, clothes and carrying conditions. After applying a predictive model, CCR of 26% is achieved. Clothing was not controlled in any of these studies. For more details of the databases see Table I.

Many studies on covariate factors, such as [4] [5], do not consider temporal data and focus solely on data acquired on the same day.

In this paper, we present empirical evidence to show that by controlling clothing the recognition performance is not affected drastically over 9 months. This is the longest time period yet considered for gait and we report, rather to our

This work was supported by *The National Physical Laboratory* (NPL UK). All of the authors are with the School of Electronics and Computer Science (ISIS research group), University of Southampton, SO17 1BJ United Kingdom. Contact: {dm08r, msn, sm3, jnc}@ecs.soton.ac.uk

surprise, a much higher recognition rate than any previous study.

Papers such as [11] treat time (over 6 months) as the most difficult covariate for gait recognition. They have proposed a new pattern classification method to solve the elapsed time problem in gait recognition. The work presented in this paper shows that elapsed time does not appear to affect the recognition performance considerably in the short to medium term. Our work aims to direct the research community towards solving other gait covariates that appear to affect performance more considerably than elapsed time. Consequently we provide evidence to show that clothing seems to be the major problem for model-free approaches. Studies such as [12] have recently started to approach this problem. Lack of a suitable database has been a major obstacle for understanding the effect of time. Our novel database enables researchers to examine each covariate separately and discover new algorithms invariant to each covariate. We also show that the ability to recognise someone from a large distance, at a low resolution is not affected by elapsed time.

The rest of this paper is organised as follows: section II describes existing temporal datasets and describes our new multimodal temporal dataset. Section III elaborates on the methodology used to obtain the results which are presented in section IV. Finally, the discussion and the future work are presented in section V.

II. DESCRIPTION OF THE DATASETS

A. Existing Temporal Datasets

Table I shows some of the most well known datasets that contain temporal data. Our database is shown in bold at the bottom of the table.

TABLE I
A SUMMARY OF EXISTING TEMPORAL GAIT DATASETS

Name	Subj.	Samples	Time Diff	Views	In/Out
MIT 2001	13	194	3 months	1	Y/N
HumanID	33	132	6 months	2	N/Y
SOTON 2002	10	≈100	0,7,15 months	1	Y/N
UMD	25	100	3 months	4	N/Y
SOTON Temporal	25	≈2000	0,1,3,4,5,8, 9 months	12	Y/N

Other databases add more than one covariate at a time and it is difficult to quantify the individual effect. It is not possible to compare two samples of the same subject taken at two distinctive time periods while keeping all other covariates the same. For example using the Gait Challenge dataset three covariates are added at the same time: clothes, shoes and time. Furthermore, none of the databases contains samples over a time period of subjects wearing the same clothes. This is a *key difference* between our new database and all the existing ones.

None of the existing databases can be used to investigate the effect of time in a formal manner. It is impossible to quantify whether the recognition performance varies due to changes of

gait over time or because other factors such as clothes, shoes or background change over time. Another advantage of our database is the availability of more than just two acquisition sessions. It is also possible to achieve analysis between 6 different time periods.

B. Our Multimodal Temporal Dataset

The data is captured using the University of Southampton Multi-Biometric Tunnel. The “tunnel” is a constrained environment that contains 12 synchronised cameras to capture subject’s gait, a camera to capture face video and an ear camera. The intended use of the “tunnel” is at airports and other high throughput environments. The data is acquired automatically in a non-invasive manner as the subject walks through it. Details of the arrangements are explained elsewhere [13].

A novel and unique database (*SOTON temporal*) has been created that enables analyses to be performed over time whilst controlling covariates that are known to affect gait recognition performance. Some of the factors that have been accounted for and remained constant over time are:

- *Environment*: Background, lighting, walking surface, position of cameras (automatically calibrated)
- *Subjects*: The subjects do not wear shoes. No major changes in any subject’s life style were reported.
- *Other*: 20 samples/subject/session are collected to enable samples of similar speeds to be compared.

Data is acquired in four sessions spanning nine months. There are 25 subjects (17 male and 8 female), between 20-55 years old. The same subjects were used in all sessions. Not all subjects were available for each session. The number of subjects available for each session is shown in brackets in Fig. 2.

The sessions taken in months 0, 1 and 4 consist of 20 samples per subject. Subjects wore white overalls over their normal clothes. The last session (month 9) consists of 10 samples of each subject wearing normal clothes in addition to 10 samples of each subject wearing overalls.

The only instruction to subjects is to “walk normally”. An example of a raw data sample is shown in Fig. 1.

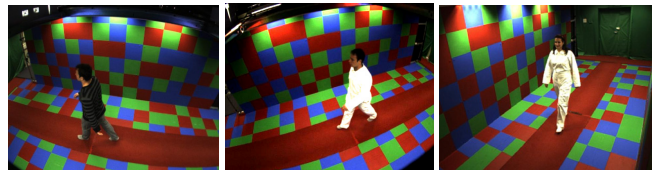


Fig. 1. - Raw data samples from SOTON multimodal temporal database

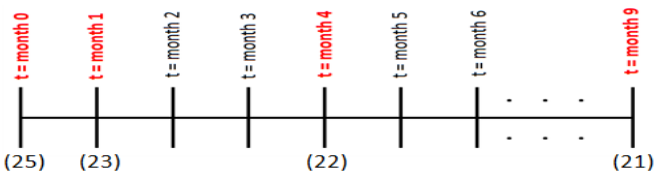


Fig. 2. - Data acquisition sessions

In hindsight, a session taken in month 7 would have provided data for difference of 2, 6 and 7 months.

Using our database, it is possible for the first time to concentrate on the effect of time and clothing on the performance of gait recognition, *ceteris paribus*.

III. METHODOLOGY

The data processing steps are shown in Fig. 3. There are at least two gait cycles for each sample as the subjects walk from the start to the end of the tunnel. The *gait cycle finder* identifies the most suitable cycle. Gait Energy Image is computed over a complete gait cycle. If $P(0)_{x,y}$ is the first image of the cycle and the next image is $P(1)_{x,y}$ the average silhouette can be computed using (1) [3].

$$AS(t)_{x,y} = \frac{1}{t} \sum_{i=0}^t P(i)_{x,y} \quad (1)$$

where t is the total number of images in the gait period.

Three views have been used in our experiment: *top* view, *side* view and *front* view. Two versions of each view have been computed: normalised (*Norm*) and non-normalised (*NN*). *N-fold* cross validation has been performed for all experiments.

The feature reduction module reduces the size of the signatures and enables us to perform recognition at different resolutions. The feature concatenation module creates a new gait signature by combining side, front and top view (*S+F+T*).

Euclidian distance is used for the k-Nearest Neighbour classifier (k-NN). Finally the CCR (Correct Classification Rate) and EER (Equal Error Rate) is computed for all experiments. A ROC (Receiver Operating Characteristic) and intra-inter class variation diagrams have also been produced for some experiments.

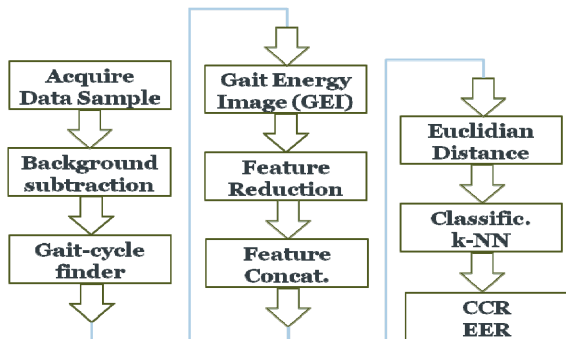


Fig. 3. Data processing steps

IV. RESULTS

A. The Effect of Time

An experiment is carried out to investigate the effect of time on the performance of gait recognition. The combinations of probe and gallery used in this experiment are shown in table II. Each probe and gallery consist of 10 samples per subject. The number of subjects for each month is shown in Fig. 2. In this experiment, we only use the overall

samples to provide consistent clothing over time. A 4-fold cross-validation is performed and the highest recognition rates are shown on the graph in Fig. 4. The error-bars for *S+F+T* indicate the lowest results of the 4-fold cross-validation. The error bars for the side and top view are not shown for clarity. Instead the standard deviation is shown in Table III.

TABLE II
THE COMBINATION OF PROBE AND GALLERY

Probe (month) <i>10 samples/subject</i>	Gallery (month) <i>10 samples/subject</i>	Time Difference (months)
0	0	0
1	0	1
4	1	3
4	0	4
9	4	5
9	1	8
9	0	9

We have shown that the CCR does not drop considerably over time for all three views (*side, front, top*) and the fusion of these three (*S+F+T*).

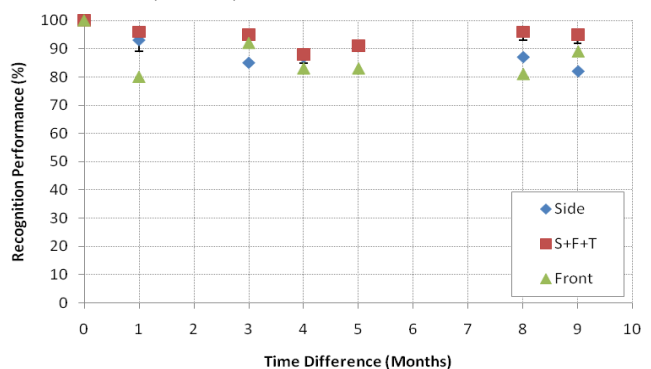


Fig. 4. - Recognition Performance over time for 2 different views and fusion of 3 different views. (*S+F+T* = Fusion of Side, Front and Top view)

TABLE III
THE STANDARD DEV. FOR THE RESULTS OF 4-FOLD CROSS VALIDATION

Time Difference (months)	Standard Dev. (Side)	Standard Dev. (Front)
0	0.01	0.01
1	0.03	0.02
3	0.02	0.06
4	0.03	0.01
5	0.01	0.03
8	0.01	0.01
9	0.00	0.01

The variation of performance over time is similar for the three different views considered and the fusion of all three views always delivers better results.

There is only a 5% drop in performance over 9 months when the combination of side, front and top is used. It is not clear whether the slight drop in performance over time is attributed to time (aging) itself or to covariate factor(s) that we have yet to consider. However, our hypothesis is that the variation of performance over time can be caused by change of clothing underneath the overall, as well as some change due to elapsed time. The best recognition is achieved when comparing similar temperature seasons (e.g. summer →

spring). Presumably, subjects wore similar type of clothes underneath the overall. Subjects wearing a body suit (without clothes underneath) was considered but discarded as it was seen as impractical.

The value of CCR for the front and side views is similar over time (except for time difference of 1 month). There is considerably more dynamic gait information available when the side view is used, compared to the frontal view. It is questionable how much of the dynamic information is used for recognition over time in the case of model-free approaches, such as the GEI. Nevertheless, the results show that the top view is not suitable for recognition over time.

Fig. 6 shows that the ROC curves for all combinations of gallery and probe shown in Table II are very similar and overlapping. The curve shown in green is for a gallery and probe taken on the same day and the blue one is for a time difference of 9 months when samples of different clothing are compared.

Fig. 7 shows that the intra/inter class variation is very similar for two different time periods.

The results presented in this section show that it is possible to recognize someone reliably by their gait after a certain time period has elapsed. Fig. 5 shows the significance of our results compared to results achieved so far.

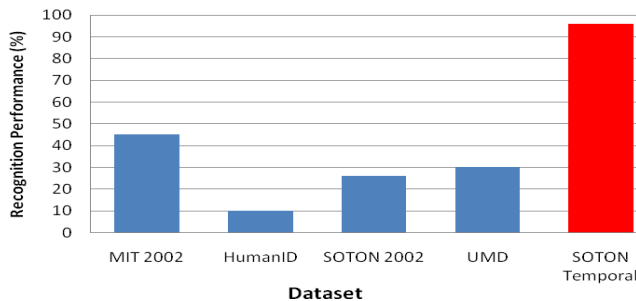


Fig. 5 – Highest CCR achieved by studies over time using various datasets. The results in red is the result achieved by our study using our new dataset.

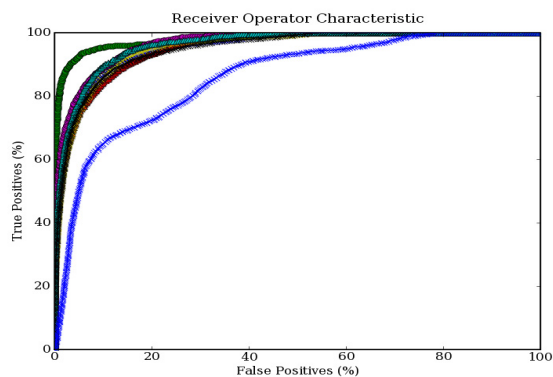


Fig. 6 – ROC curve for all combinations of gallery and probe as shown in Table II. In addition the last entry of Table IV is plotted.

TABLE IV
THE COMBINATION OF PROBE AND GALLERY

Probe (month) 10 samples/subject	Gallery (month) 10 samples/subject	Time Difference (months)
9 (Norm)	9 (Overall)	0
9 (Norm)	4 (Overall)	5
9 (Norm)	1 (Overall)	8
9 (Norm)	0 (Overall)	9

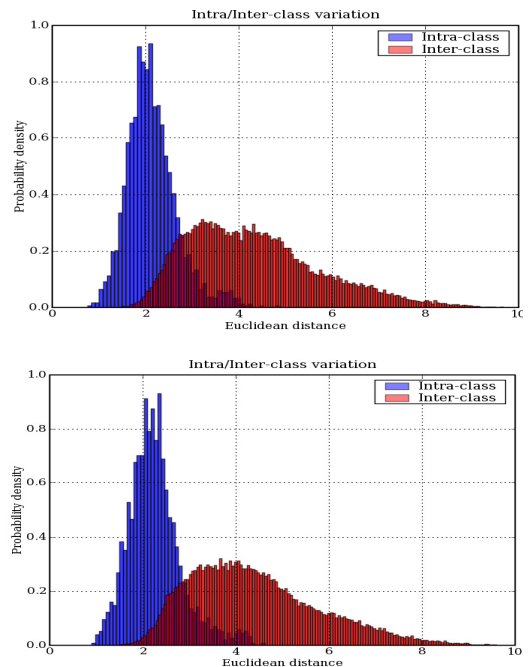


Fig. 7 – The Intra-inter class variation for time difference of 1 month (top) and time difference of 9 months (bottom)

B. The Effect of Clothes

The data acquired in month 9 differs from that in the other three. Ten samples were taken of each subject wearing normal clothes and another ten samples of each subject wearing overalls, only a few minutes apart. This data enables us to perform the analysis of different clothes over time and over few minutes. It enables us to quantify the effect of clothes, while keeping all other covariates constant.

In addition to the combinations of gallery and probe shown in Table II the combinations shown in Table IV are used for this experiment.

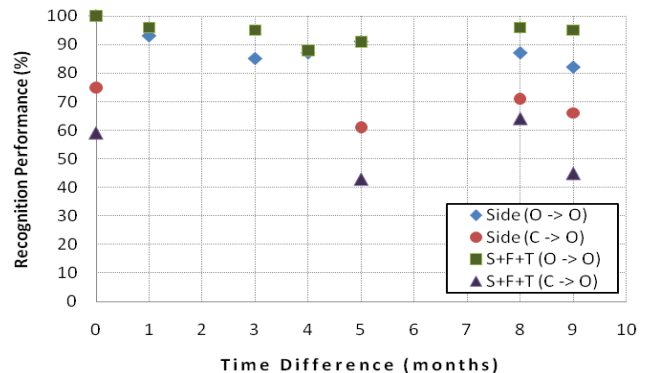


Fig. 8 - The recognition performance over time. (O -> O = Overall -> Overall, C -> O = Normal Clothes -> Overall)

2-fold cross validation is performed. Only the highest CCRs achieved are shown Fig. 8. The highest value for the standard deviation is 0.04.

The results presented in Fig. 8 clearly indicates that different type of clothes can lead to a large drop in performance. The finding is similar for all three views, but only the *side* and *S+F+T* are shown on the diagram for clarity. It is important to note that the drop in recognition is

consistent with time.

Fig. 9 further clarifies the effect of clothing. The probe and gallery are captured only few minutes apart. Five samples per subject are used for the probe and the gallery. The horizontal axis shows the combination of clothing considered. (N = Normal, O = Overall). Very high recognition results for all views are achieved if same types of clothes are matched. However, when overall is matched against normal clothes, the recognition rates can fall to 40%. Consequently, we have shown that recognition can be affected significantly regardless of elapsed time or any other covariates.

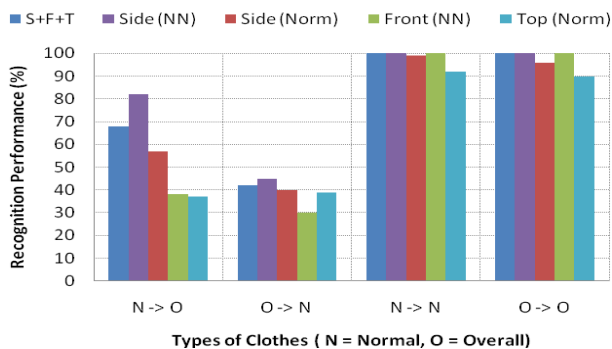


Fig. 9 – The effect of clothes on the recognition performance. The probe and the gallery have been taken few minutes apart.

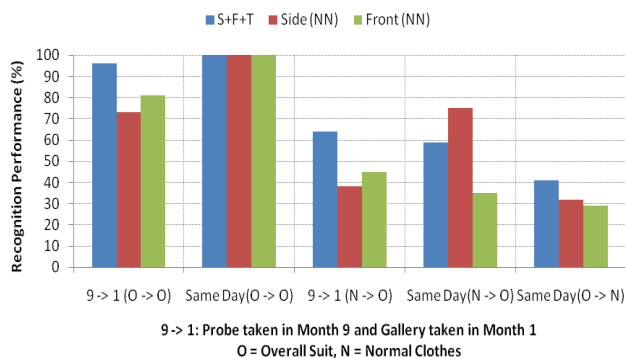


Fig. 10 – The Effect of clothes on recognition, over time period of 8 months

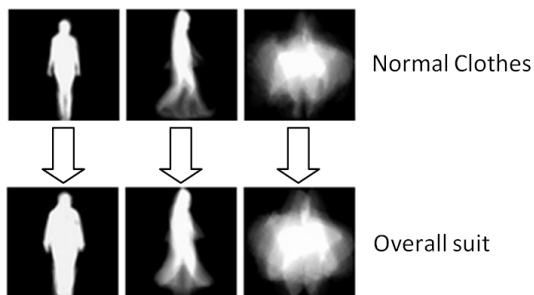


Fig. 11 - The effect of clothes on the Gait Signatures (samples taken few minutes apart) – Left = Front, Middle = Side and Right = Top

The results from the previous experiment (Fig. 9) clearly show that different clothing can have a major impact on recognition. An experiment is carried out to understand the effect of different clothing over time. Ten samples per subject are used for both the probe and the gallery. The results are presented in Fig. 10. Matching samples taken only few minutes apart with the same clothes leads to 100%

recognition for all three gait signatures. Very high recognition is achieved if samples of subjects wearing the same type of clothes are compared, regardless of the time difference. In this paper the results for time difference of 8 months is shown, but the results are consistent for all other time differences considered.

Fig. 10 shows that higher recognition rates can be achieved over a period of 8 months than over few minutes in the case of different clothing between the probe and the gallery. In the case of matching samples with normal clothes to samples with overall, the recognition for S+F+T and Front is higher over time than it is for samples taken few minutes apart. This is more evidence that model-free approaches for gait recognition such as GEI are affected heavily by different clothing and not by elapsed time.

The lowest recognition rate is obtained when samples are taken only few minutes apart, when the probe contains samples of subjects wearing overalls and the gallery consists of samples of subjects wearing normal clothes. This is the true for all the time differences we considered. In fact, this combination of probe and gallery always leads to a worse recognition, compared to matching samples of overall to normal clothes, taken over time longer time periods.

Fig. 11 shows the effect of clothing on the GEI signature. There is a considerable variation in both, the dynamic and static part for all three views. Clothing can change the body shape noticeably.

However, overalls can be considered ‘extreme’ clothing. People usually wear similar type of clothes. We believe that if ‘normal’ but different clothes are compared the recognition results would be much better.

C. Number of Features v.s. Performance over Time

A unique advantage of gait is the capability to do recognition at a distance, at a low resolution. Therefore, it is important to show how the performance changes with number of features. In this study, the number of features equates to the number of pixels in the GEI gait signature. If gait is captured at a distance there are fewer pixels in the image that can be used for recognition. Consequently a decrease in the number of features is equivalent to an increase in the distance at which the subject is observed. The rate at which the performance deteriorates with smaller number of features should be consistent over time. In other words, if the gallery and probe are acquired m month apart and a recognition performance of p is achieved with n number of features then a $p \pm q$ (where q is very small) performance for the same number of features n should be expected if the gallery and probe are acquired $m+o$ months apart. (o any value).

We have performed an experiment to show whether time has any effect on the ability to do gait recognition at a lower resolution. The findings are presented in Fig. 12. We have shown that irrespective of the time difference (except when the time difference is few minutes) between the probe and the gallery the degradation of performance as the number of

features decrease is fairly consistent. This is evidence that gait recognition can be used as a reliable biometrics at a distance over time. Fig. 12 shows that high performance can be achieved by using a small number of features. Using only 900 features (30x30) a CCR of 80%-90% can be achieved. However, a larger database is needed to correctly verify the number of features n needed to achieve a recognition p . The main aim of this experiment is to show that elapsed time does not play an important role to achieve recognition at various resolutions.

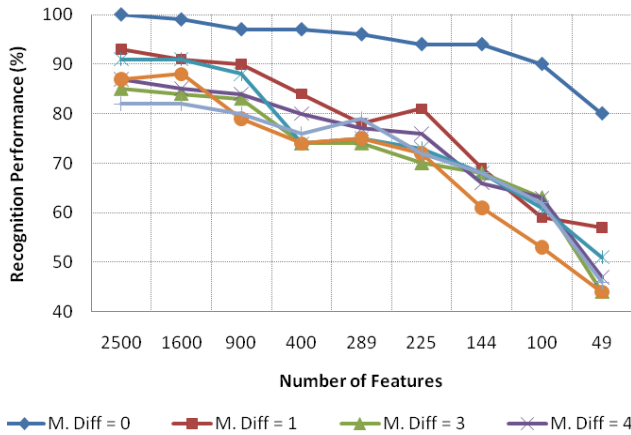


Fig. 12 – Number of features vs. Performance, over time

V. DISCUSSION AND FUTURE WORK

We have shown for the first time that gait can be used as a reliable biometric trait in the short and medium term if all covariate factors are properly accounted for. A CCR of 95% is achieved over period of 9 months using relatively simple techniques. We hypothesize that a CCR of nearly 100% could be achieved if the clothes underneath the overalls are controlled. As a comparison, the performance of face recognition can deteriorate linearly after a year has elapsed between capturing the gallery and the probe [14]. Further work could reveal that gait could be equally or more suitable biometrics for recognition over lengthy time periods. However, given that the problem of clothing can be overcome using techniques as in [15], the issue of gaining/losing weight would still pose a challenge in the case of model-free approaches, such as the GEI. Further research is needed to identify an age-invariant feature sub-space for gait.

In studies to date, there has been a lack of understanding of elapsed time on the performance of gait recognition. Time has been considered as a combination of other time-dependent covariate factors. The major obstacle for quantifying the effect of time and other factors independently is the lack of suitable databases. We have created a novel multimodal temporal database which could be used for developing and evaluating covariate invariant algorithms amongst many other things.

We have shown that the major problem for recognition over time can be attributed to change of clothes. We have

shown a dramatic fall in performance in case of ‘extreme’ changes in clothes. The CCR drops from 100% to 30% for *non-temporal* combination of gallery and probe thus achieving similar results to previous studies.

We have confirmed that recognition by gait at low resolution (at a distance) is not affected by elapsed time. This is a very significant finding because recognition at a distance is a major advantage of gait over all other biometrics.

More subjects, data over longer period and other recognition algorithms are needed to confirm the conclusions presented in this paper. The effect of other covariate factors such as speed and shoes will be investigated later.

The data we have collected for this study is also suitable for analyzing face from video over time in an unconstrained environment. Future work will include quality driven biometric fusion of gait and face. Learning what constitutes good gait and face samples with respect to covariates that change over time could lead to a revolutionary biometric system. We will be focusing on developing measures of quality that compensate for covariate factors.

REFERENCES

- [1] L. Lee and W. E. L. Grimson, "Gait analysis for recognition and classification," in *Proc. 5th IEEE Int. Conf. AFGR*, 2002, pp. 148-155.
- [2] J. Han and B. Bhanu, "Individual recognition using gait energy image," *IEEE Trans. PAMI*, **28**, pp. 316-322, 2006.
- [3] Z. Liu and S. Sarkar, "Simplest representation yet for gait recognition: Averaged silhouette," in *Proc. 17th IEEE Int. Conf. ICPR*, Cambridge, UK, 2004, pp. 211-214.
- [4] I. Bouchrika, M. Goffredo, J. Carter, and M. Nixon, "Covariate analysis for view-point independent gait recognition," in *The 3rd IAPR/IEEE Int. Conf. ICB Italy*, 2009.
- [5] I. Bouchrika and M. Nixon, "Exploratory factor analysis of gait recognition," in *The 8th IEEE Int. Conf. AFGR Amsterdam*, The Netherlands, 2008.
- [6] S. Sarkar, P. J. Phillips, L. Zongyi, V. Isidro Robledo, P. Grother, and K. W. Bowyer, "The humanID gait challenge problem: Data sets, performance, and analysis," *IEEE Trans. PAMI*, **27**, pp. 162-177, 2005.
- [7] Z. Liu, L. Malave, and S. Sarkar, "Studies on silhouette quality and gait recognition," in *IEEE Int. Conf. CVPR*, Washington, D.C., 2004, pp. 704-711.
- [8] R. T. Collins, R. Gross, and J. Shi, "Silhouette-based human identification from body shape and gait," in *Proc. 5th IEEE Int. Conf. AFGR*, 2002, pp. 366-371.
- [9] A. Kale, A. Rajagopalan, N. Cuntoor, and V. Kruger, "Gait-based recognition of humans using continuous HMMs," in *Proc. 5th IEEE Int. Conf. AFGR*, 2002, pp. 336-343.
- [10] G. Veres, M. Nixon, and J. Carter, "Model-based approaches for predicting gait changes over time," in *Proc. Int. Conf. ISSNIP*, 2005, pp. 325-330.
- [11] D. Tao, X. Li, X. Wu, and S. Maybank, "Elapsed time in human gait recognition: A new approach," in *Proc. IEEE Int. Conf. ICASSP*, 2006.
- [12] M. Altab Hossain, Y. Makihara, J. Wang, and Y. Yagi, "Clothing-invariant gait identification using part-based clothing categorization and adaptive weight control," *Pattern Recognition*, **43**, pp. 2281-2291, 2010.
- [13] R. Seely, S. Samangoei, M. Lee, J. Carter, and M. Nixon, "University of Southampton multi-biometric tunnel and introducing a novel 3D gait dataset," in *Proc. 2nd IEEE Int. Conf. BTAS*, Washington D.C., 2008.
- [14] Y. M. Lui, D. Bolme, B. Draper, R. Beveridge, G. Givens, and J. Phillips, "A meta-analysis of face recognition covariates," in *3rd IEEE Int. Conf. BTAS* Washington D.C., 2009.
- [15] A. Balan and M. Black, "The naked truth: Estimating body shape under clothing," in *Proc. ECCV*, 2008, pp. 15-29.