

# Fusion of dynamic and static features for gait recognition over time

Galina V. Veres, Mark S. Nixon, Lee Middleton and John N. Carter  
School of Electronics and Computer Science  
University of Southampton, UK.  
gvv,msn,ljm,jnc@ecs.soton.ac.uk

**Abstract** – *Gait recognition aims to identify people at a distance by the way they walk. This paper deals with a problem of recognition by gait when time-dependent covariates are added. Properties of gait can be categorized as static and dynamic features which we derived from sequences of images of walking subjects. We show that recognition rates fall significantly when gait data is captured over a lengthy time interval. A new fusion algorithm is suggested in the paper wherein the static and dynamic features are fused to obtain optimal performance. The new fusion algorithm divides decision situations into three categories. The first case is when more than two thirds of the classifiers agreed to assign identity to the same class. The second case is when the two different classes are selected by each half of classifiers. The rest falls into the third case. The suggested fusion rule was compared with the most popular fusion rules for biometrics. It is shown that the new fusion rule over-performs the established techniques.*

**Keywords:** Gait recognition, static and dynamic features, time-dependent covariates, fusion.

## 1 Introduction

The demand for automated person identification systems is growing in many important applications such as visual surveillance, access control, and smart interfaces, where reliable individual verification is essential. Recently much attention has been devoted to the use of human gait patterns as a biometric. Gait recognition discriminates individuals by the way they walk and has the advantage of being non-invasive, being readily captured without a walker's attention, and is less likely to be obscured than other biometric features.

Approaches to gait recognition can be broadly classified as being model-based and model-free. Model-based methods [1, 2, 3] model the human body structure and extract image features to map them into structural components of models or to derive motion trajectories of body parts. Model-free methods [4, 5, 6] generally characterise the whole motion pattern of the human body by a compact representation regardless of the underlying structure. In this paper we employ the model-based (dynamic) method of Wagg and Nixon [7] and the model-free (static) method of Veres et al [8].

For obtaining optimal performance, an automatic person identification system should integrate as many informative clues as available. In [9] decision fusion of face and gait

cues for the single camera case is reported based on two fusion scenarios: hierarchical and holistic. The first involves using the gait recognition algorithm as a filter to pass on a smaller set of candidates to the face recognition algorithm. The second involves combining the similarity scores obtained individually from the face and gait recognition algorithms. Simple rules like the SUM, MIN and PRODUCT are used for combination. Decision fusion is a special case of data fusion [10] and will be employed here. Further we concentrate only on gait recognition. There are various properties of gait that might serve as recognition features. They can be categorised as static features and dynamic features. Based on the idea that body biometrics includes both the appearance of human body and the dynamics of gait motion measured during walking, some efforts were made to fuse two completely different sources of information available from walking video for person recognition [11, 12].

However, in these works only databases recorded in a short time interval evaluated. Some studies over a lengthy time interval were reported for face recognition. In [13] images of 240 distinct subjects were acquired under controlled conditions, over a period of ten weeks. They showed that there was not a clearly decreasing performance trend over a period of ten weeks and concluded that degradation line is small enough as to be nearly flat over this time period. Other studies have shown that with data recorded over a period of years face recognition performance degrades linearly with time [14]. Some studies were done to show affects of aging on face recognition [15, 16, 17]. In [15] a systematic method for modelling appearance variation due to aging is presented. It was shown that aging variation is specific to a given individual, it occurs slowly and it is affected significantly by other factors, such as the health, gender and the lifestyle of the individual. Taking this into consideration then reasonably accurate estimates of age can be made for unseen images. In [16, 17] face identification experiments are presented, where the age of individuals in the gallery is significantly different than the age of individuals in the probe. It was demonstrated that automatic age simulation techniques can be used for designing face recognition systems, robust to aging variation. In this context, the perceived age of subjects in the gallery and probe is modified before the training and classification procedures, so that aging variation is eliminated. Some recent efforts [18]

were made to improve age estimation by devoting part of the classification procedure to choosing the most appropriate classifier for the subject/age range in question, so that more accurate age estimates can be obtained.

In this paper we consider a gait recognition problem when two databases (the gallery and probe) were recorded with a time interval of 6 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 imagery where subjects wore different shoes, clothes, or carried different bags. In real life the need to analyse such databases arises in security of access to a company or an embassy for example. It is possible to record people walking normally as a gallery, but later it will be necessary to recognize these people in different clothes, shoes, possibly carrying some bags and when time passes. It is shown that in this case correct classification rates fall significantly and recognition becomes unreliable. Similar results were obtained for the HumanID Gait Challenge Problem [19], where recognition fell from 82% to 6% after 6 months. Some other recent works reported a significant fall in recognition capability over lengthy time interval [20, 21, 22]. One way to handle reduction in recognition capability with time is to try to fuse different information.

In this paper a new fusion algorithm is suggested to be used in a two-stage data fusion framework. Before fusing static and dynamic features of gait we suggest reduction in dimensionality using different approaches such as feature extraction and feature selection to produce different sets of features representing both dynamic and static parts of gait. Then at the first stage the fusion of different classifiers for each dataset is performed. Some work was done in this area [23, 24, 25, 26] and in this paper we use only the most popular fusion rules in biometrics to compare them with the new approach. The second stage deals with the fusion of outputs from the first stage representing static and dynamic parts of gait. A suggested fusion algorithm divides all decision situations into three cases. The first case is when more than two thirds of classifiers agree on the same class. In this case this class is selected as an output of fusion process. The second case is when exactly half of the classifiers agree about the same class and the second half can all agree on a different class. In this case the class is chosen according to maximum sum of classifier weights. All the remaining decision combinations belong to the third case where local accuracy of each classifier is taken into consideration to make the final assignment. The suggested fusion rule was compared with the most popular fusion rules for biometrics. It is shown in the paper that the new fusion rule performs better than the established techniques, such as majority voting, min, max, sum, product and median rule. It happens since when the new fusion rule is used, then either majority vote is selected, or when there is no majority vote either the output of the most competent classifiers is chosen or the most popular subject among all classifiers is selected.

The remainder of this paper is organised as follows. Section 2 presents the suggested data fusion algorithm. The methodology of constructing feature sets is presented in

Section 3. Experimental results are presented and described in Section 4. Section 5 concludes this paper.

## 2 Data fusion algorithm

In this paper a new fusion rule applied to two-stage data fusion framework is suggested. The new fusion rule is presented first and then two-stage data fusion framework is described. We assume that a small set of classifiers is available and we are interested in combining their output to get the best possible result.

Consider a gait recognition problem where a subject  $\mathbf{x}$  is to be assigned to one of the  $c$  possible classes  $\Omega = \omega_1, \dots, \omega_c$  using a set of  $N$  classifiers  $R_1, \dots, R_N$ . Each classifier gets as its input a feature vector  $\mathbf{x} \in \mathbb{R}^n$  and assigns it to a class from  $\Omega$ , i.e.  $R_i : \mathbb{R}^n \rightarrow \Omega$ , or equivalently,  $R_i(\mathbf{x}) \in \Omega, i = 1, \dots, N$ . The classifier output is usually a  $c$ -dimensional vector with supports to the classes. In this paper the probe (subjects with unknown classes) is compared against the gallery (subjects with known classes), and the output of classifier  $i$  is distance matrix  $\mathbf{D}_i$  of size  $(M \times C)$ , where  $C = \sum_{i=1}^c n_g(i)$  is a number of subjects in the gallery,  $n_g(i)$  is a number of subjects in each class and  $M$  is a number of subjects in the probe. Then a subject  $\mathbf{x}$  is assigned to class  $\omega_s$  for a given classifier iff

$$\mathbf{D}_{is}(\mathbf{x}) \leq \mathbf{D}_{it}(\mathbf{x}), \quad \forall t = 1, \dots, C. \quad (1)$$

In case of more than one classifier the fusion rule is needed to make a final decision. The new fusion rule is presented in the next subsection.

### 2.1 New fusion rule

An algorithm describing a new fusion rule is presented below

- Start with  $N$  distance matrices  $\mathbf{D}_i$  and form a decision matrix  $\mathbf{Q}$  and a label matrix  $\mathbf{L}$ , whose sizes are a number of subjects  $M$  in probe by number of distance matrices  $N$  or voters/classifiers. The label matrix  $\mathbf{L}$  consists of the labels assigned to each subject in the probe. The label matrix  $\mathbf{L}$  is constructed as

$$\mathbf{L}_{ki} = \omega_{t(p)} \quad \text{such as} \quad D_i(\mathbf{x}_k) = \min_{p=1}^C D_i^p(\mathbf{x}_k), \quad (2)$$

where  $k = 1, \dots, M, i = 1, \dots, N$  and  $t(p)$  means the class from the gallery nearest to  $k$ th subject from the probe. The decision matrix  $\mathbf{Q}$  is a matrix of zeros at this stage. Define weights  $w_i$  for each classifier.

- For each subject in probe
  1. If more than two thirds of voters agree on the same subject, then choose this subject as the output.

assign  $\mathbf{x} \rightarrow \mathbf{L}_{ki}$  if

$$\mathbf{L}_{ki} = \mathbf{L}_{kj}, \quad i \neq j \quad \text{and} \quad \cup j \geq 2N/3, \quad (3)$$

where  $\cup j$  means a number of  $j$  for which condition (3) holds,  $j = 1, \dots, N$ . If for  $k$ th subject condition (3) is true, then assign  $\mathbf{Q}_{kj} = 1, \forall j$ .

2. If exactly half of the voters select one subject and the other half of voters selects another subject, calculate weights of voters for each subject, and select the subject with the maximum sum of voters weights.

$$\text{assign } \mathbf{x} \rightarrow \mathbf{L}_{ki} \quad \text{if} \\ \mathbf{L}_{ki} : \sum_{j=1}^{N/2} w_j(i) > \sum_{k=1}^{N/2} w_k(r) \quad , \quad (4)$$

where  $i$  is a position of classifier which output is considered as true;  $w_j(i)$  are weights of classifiers voting for one subject;  $w_j(r)$  are weights of classifiers voting for another subject. For all classifiers which selected the chosen subject assign 1 in the decision matrix  $\mathbf{Q}$ .

3. In the remaining cases
  - a) Choose the number of the best nearest neighbours  $K$  ( $K = 20$  is recommended, but this might change accordingly to application).
  - b) Sort distance matrices  $\mathbf{D}_i$  in ascending order with index matrix  $\mathbf{I}$  which records the positions of the sorted distances in the distance matrix.
  - c) Form the matrix  $\mathbf{K}_{nn}$  of  $K$  smallest distances and note their positions in the matrices  $\mathbf{D}_i$ , i.e. form matrix  $\mathbf{I}_{nn}$ .
  - d) Find the most popular subject in the matrix  $\mathbf{I}_{nn}$  and select it as the output. If several subjects have the same popularity, select the output from these subjects at random.
  - e) For all classifiers which selected the most popular subject assign 1 in the decision matrix  $\mathbf{Q}$ .

This algorithm takes into consideration the local accuracy of classifiers when there is no preferred subject. It is suggested to use this fusion rule together with the data fusion framework described below.

## 2.2 Data fusion framework

In this section a data fusion framework at a decision level is described. Databases are divided into gallery and probe. It is suggested before applying fusion algorithms to reduce the dimensionality of both static and dynamic data sets using selection and extraction methods reported in [8, 29, 30] in order to obtain usually different feature sets representing the same dataset and reduce computational load without loss or insignificant loss in gait recognition. The main stages of the suggested data fusion framework are

**Stage 1** Fusion of different classifiers for the same dataset

**Stage 2** Fusion of new datasets representing dynamic and static parts of gait.

Firstly, after reducing dimensionality, several classifiers are run on new data sets and their results are fused to improve recognition for a given set. A number of classifiers and which kind of classifiers to use may differ from one dataset to the next. A decision matrix and a fused distance matrix are formed for each data set. After that the fused distance matrices for static and dynamic parts of gait are combined using the same fusion rule to obtain a final fused distance matrix and a final decision matrix. Such fusion rules as majority voting, min, max, sum, product, median and the new fusion rule were tried and compared.

The algorithmic description of the data fusion framework is presented below.

1. Start with initial sets of features  $\Omega_s$  and  $\Omega_d$  representing static and dynamic parts of gait respectively. Select and/or develop a set of classifiers. Select and/or develop a set of data fusion rules
2. Run feature selection and feature extraction algorithms on each feature set  $\Omega_s$  and  $\Omega_d$ . Form four new features sets  $\Omega_{se}$  and  $\Omega_{de}$ , where features obtained by feature extraction algorithm are used, and  $\Omega_{ss}$  and  $\Omega_{ds}$  containing features obtained by feature selection algorithm.
3. For each new data set
  - Run selected classifiers
  - For each classifier form a distance/probability matrix
  - Normalise each distance/probability matrix (further referred to as the distance matrix for convenience) in such a way that values in the matrix were normalised to be between 0 and 1.
  - Choose a common rule of denoting the nearest neighbour and adjust distance matrices respectively, i.e. for example in all matrices 0 will mean the nearest and 1 will mean the furthest or vice versa.
  - Using a chosen fusion rule, form a label matrix, a decision matrix and a fused distance matrix
4. Combined the fused distance matrices corresponding to static and dynamic part of gait and form a final label matrix, a final decision matrix and a final fused distance matrix

The decision matrix is a matrix of binary values, where 1 means that subject was recognised as belonging to a class and 0 means subject was not recognised as belonging to a class. A fused distance matrix is constructed according to the chosen fusion rule and the corresponding decision matrix. In this paper an algorithm for forming the fused distance matrix using the new fusion rule is presented. This can easily be adapted to different fusion rules. The algorithm of fused distance matrix forming is as follows

1. Start with  $N$  distance matrices  $\mathbf{D}_i$  and decision matrix  $\mathbf{Q}$ , whose size is a number of subjects in probe  $M$  by number of voters or distance matrices  $N$ .

2. Define  $\mathbf{D}_f = \emptyset$  as a fused distance matrix.
3. For each subject in probe
  - Check how many votes assign the subject  $j$  in probe to a given class
  - If the number of voters is more or equal  $2 * N/3$ , then find a position of first 1 in decision matrix, note it as  $(j, i)$  and assign  $\mathbf{D}_f(j, :) = \mathbf{D}_i(j, :)$ .
  - If exactly half of the voters selects one subject and another half of voters selects another subject, calculate weights of votes for each subject, then find a position of first 1 in decision matrix among votes with highest weight, note it as  $(j, i)$  and assign  $\mathbf{D}_f(j, :) = \mathbf{D}_i(j, :)$ .
  - In all other cases find the most popular subject among votes following the new fusion rule and note the position of this subject as  $(s, i)$  in distance matrix. Assign  $\mathbf{D}_f(j, s) = \mathbf{D}_i(j, s)$ ,  $\mathbf{D}_f(j, 1 : s) = \mathbf{D}_i(j, 1 : s) + \mathbf{D}_i(j, s)$  and  $\mathbf{D}_f(j, s + 1 : M) = \mathbf{D}_i(j, s + 1 : M) + \mathbf{D}_i(j, s)$

This algorithm ensures that the fused distance matrix contains the best result after fusion from a recognition point of view, because either the majority vote is selected, or when there is no majority vote either the most competent classifier or the most popular subject among classifiers is selected.

### 3 Databases description and methods

Two databases were analysed in the paper, both comprising indoor (studio, controlled) data since the purpose of this paper is to investigate a gait recognition problem. The first database, called the large database (LDB), consists of 115 subjects performing a normal walk. The database arrangements are described elsewhere [27]. LDB can help to determine which image information remains unchanged for a subject in normal conditions and which changes significantly from subject to subject, i.e. it represents subject-dependent covariates. The small database (SDB) consists of a 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. In this case the covariate factor is more complicated, since it does not solely depend on the subjects, but on the other factors mentioned above. Each subject's data was captured during one continuous filming session and 14 different experiments for each person was recorded. The 14 experiments are designed to investigate the effect of five factors on performance. The five factors are studied individually. The first factor describes the influence of different footwear and includes the following experiments: subject walking in flip flops; subject walking with bare feet; subject walking in socks; subject walking in boots; subject walking in trainers; subject walking in own shoes. The second factor studies the affects of clothing on the performance such as subject wearing raincoat and subject wearing trench coat. The third factor includes experiments with subject carrying

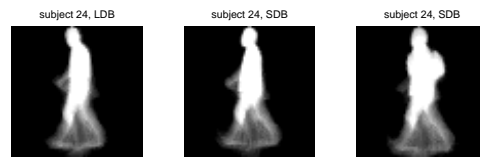


Fig. 1: Average silhouette representing subject 24: left picture is LDB, middle picture is SDB, walking normally, right picture is SDB, carrying a bag

different objects such as subject carrying hand bag; subject carrying barrel bag (with slung over shoulder, bag at hip height); subject carrying barrel bag (carried by hand on shoulder); subject carrying rucksack. Factor four investigate the affects of speed on recognition and includes such experiments as subject walking slowly and subject walking fast. And the last fifth factor looks at influence of short-time interval on the performance, using experiment where subject walks normally in his own shoes, but recorded 2 hours later than the first record. The SDB is intended to investigate the robustness of biometric techniques to imagery of the same subject in various common conditions (carrying items, wearing different clothing or footwear). An example of features for LDB and SDB is presented in Fig. 1. The figure shows subject 24 recorded in LDB and in SDB walking normally and carrying a bag. The SDB is intended to investigate the robustness of biometric techniques to imagery of the same subject in various common conditions (carrying items, wearing different clothing or footwear). It worth noticing that sequence of filming data in time was LDB first and then SDB with approximately 6 months difference between recording LDB and SDB, i.e. time-dependent covariates are added. Each subject ID consists of the following fields: session number; camera ID (the same for this experiment); subject number; data type (subject or background, subject in this case); experiment number (only for SDB); sequence number and direction of walk (right or left).

To produce binary subject silhouettes (the static part of gait) from the video the following steps are performed as reported in [8]. First, the video is calibrated to correct for radial distortion errors. Second, chromakey analysis is done to mark some known background pixels to assist with extracting the subject. The data was recorded in a laboratory with a fixed green background cloth and the track, on which the subject walks, was painted the same shade of green. The chromakeyed frames from video are then cropped with fixed values adjusted to the gait lab parameters to exclude laboratory equipment. The final silhouette is obtained by connected component analysis and morphological operators. The output is then a series of binary images with the subject in white on a black background. Period detection is used to determine the beginning and the end of a gait period. This is achieved by analysis of the variation of the distance between the legs. A gait cycle was defined as a sequence of silhouettes taken from one minimum of this parameter to the next minimum. The gait cycle closest to the centre of the field of view was chosen for further analysis. The signature is calculated by averaging the binary silhouettes over a complete gait cycle and as a result a 4096

dimensional feature vector obtained. This is known as the average silhouette.

The dynamic part of gait was obtained following the methods described by Wagg and Nixon [7]. 73-dimensional feature vectors were extracted, using model-based techniques. This signature derives from bulk motion and shape characteristics of the subject, articulated motion estimation using an adaptive model and motion estimation using deformable contours. After pre-processing to remove noise and background the sequence is edge detected. A motion compensated temporal accumulation algorithm is used to extract the bulk motion of the subject in the horizontal plane. This is then filtered using template matching, leaving only motion due to the subject. Shape estimation is then performed using a more accurate model of the subject's shape. Articulated motion is estimated by using sinusoidal models of hip, knee, ankle and pelvic rotation. These provide a starting point for model adaptation of the subject's limb movements. An adaptive process for joint location is then applied to the sequence to form a more accurate and robust model of limb movement. This adaptive process is based on an iterative gradient descent method repeated until no changes occur over the entire sequence. By feature selection, the processes described in [28] yield 45 parameters based on joint rotation models for the hip, knee and ankle and 18 parameters describing the subject's speed, gait frequency and body proportions. A further 10 parameters are extracted from the processes described in [7]. All of these parameters are normalised to make them size invariant.

## 4 Results

Before presenting the results some notations are introduced. In the text and tables the following notations will be used: *stat* is a feature set describing static part of gait; *dyn* is a feature set describing the dynamic part of gait; *stat<sub>s</sub>* and *stat<sub>e</sub>* are sets of features obtained after applying feature selection and feature extraction algorithms to the static part of gait respectively; *dyn<sub>s</sub>* and *dyn<sub>e</sub>* are sets of features obtained after applying feature selection and feature extraction algorithms to dynamic part of gait respectively. Different feature selection algorithms were chosen for dynamic and static parts of gait due to their different performance on different datasets. In case of static feature set the modification of sequential floating selection algorithm; backward feature selection algorithm [29] was used for dynamic feature set. The correct classification rate CCR is understood as a correct classification rate obtained by comparing SDB (the probe data) via LDB (gallery), if not mentioned otherwise. To show the influence of time on recognition, the CCR is calculated for each database separately, i.e. LDB is analysed via LDB (L/L) and SDB via SDB (S/S). Cross-validation using the 'leave-one-out' rule is performed for both datasets, using Euclidean distance as the most popular distance in gait recognition literature and the 1-nearest neighbour rule. The results are presented in Table 1 and show acceptable CCRs for both dynamic and static feature sets.

CCRs for original and reduced datasets when SDB via LDB analysed are presented in Table 2, i.e. when LDB

Table 1: Analysis of databases without time-dependent covariates

	L/L	S/S
<i>stat</i>	98.47%	99.90%
<i>dyn</i>	72.32%	90.24%

is considered as the gallery and SDB is considered as the probe. A subject in the probe can wear normal shoes, clothes and walk normally or can walk slower/faster than normal, wear different shoes, raincoat or even carry a bag/rucksack. We try to match this subject in probe to a subject in the gallery who walks normally, wears normal shoes and normal clothes (no raincoat) and does not carry any bags. Analysis is done when all features are taken into consideration in the probe and the gallery and when a reduced set of features is considered. It can be seen that as soon as time-dependent covariates are added to the analysis, the fall in CCR is very noticeable. The significant reduction in feature space is achieved without significant loss in CCR, and in case of dynamic part of gait the visible increase in CCR can be seen for *dyn<sub>s</sub>*. At the same time the CCRs are very low even in the best case and something should be done to improve CCR. One of the ways is to use the suggested data fusion framework.

Table 2: CCR for different datasets before fusion

Dataset	number of features	CCR
<i>stat</i>	4096	24.78%
<i>stat<sub>s</sub></i>	139	21.50%
<i>stat<sub>e</sub></i>	97	20.05%
<i>dyn</i>	73	5.30%
<i>dyn<sub>s</sub></i>	34	13.50%
<i>dyn<sub>e</sub></i>	23	6.00%

Five classifiers were chosen for Stage 1 fusion: Euclidean distance, City Block distance, Cosine, Dice coefficient and Distance infinity. Different classifiers will be the best from the recognition point of view for different datasets. Table 3 shows CCRs achieved by the best classifier for different datasets.

Table 3: CCRs achieved by the best classifier

	<i>stat<sub>s</sub></i>	<i>stat<sub>e</sub></i>	<i>dyn<sub>s</sub></i>	<i>dyn<sub>e</sub></i>
CCRs	23.48%	19.05%	14.16%	5.47%

Several fusion rules, such as majority voting (maj in further references), min rule, max rule, sum rule, product rule (prod in further references), median rule (med in further references) and new fusion rule (new in further references) were used. CCRs for different datasets after fusion of classifiers are presented in Table 4. Table 4 shows that fusion of classifiers allows some increase in the CCR for some datasets such as *stat<sub>e</sub>*, *dyn<sub>s</sub>* and *dyn<sub>e</sub>* with new fusion rule giving the best results in comparison with the best classifier for a given dataset.

Results after Stage 2 are presented in Table 5. A new fusion rule produces the best results at this stage. A 18% increase in CCR was achieved in comparison to the best re-

Table 5: Stage 2 of data fusion framework

	new	min	max	sum	prod	med	maj
CCR	27.70%	20.91%	22.30%	19.70%	17.93%	19.70%	6.48%

Table 4: CCR after fusion of classifiers

	$stat_s$	$stat_e$	$dyn_s$	$dyn_e$
new	23.10%	19.77%	14.75%	7.90%
min	22.51%	19.43%	13.33%	5.06%
max	13.89%	17.14%	14.02%	4.78%
sum	19.15%	19.60%	13.26%	6.06%
prod	22.02%	19.29%	12.92%	7.13%
med	19.15%	19.60%	13.26%	6.06%
maj	21.68%	19.25%	13.12%	5.30%

sult before fusion. Further work will be needed to increase CCR further. It was investigated in [31] that when local OR is used CCR of 61% was achieved, i.e. it was shown that it is possible to handle time-dependent covariates. However, local OR can be only used when identity of a subject is known on decision level. In many application as verification of subjects such knowledge is not available, therefore there is a necessity of developing fusion rules which do not require information about subject identity on decision level and can produce CCRs as near as possible to results obtained by local OR.

## 5 Conclusions

This paper deals with a problem of increasing correct classification rate when time-dependent covariates (6 months passed between the finish of recording the gallery and the start of recording the probe) together with some other covariates such as variety of footwear, clothes and carrying different bags are added to an analysed database for gait recognition. It was shown that CCRs are very low in this case. One of the ways to handle with this problem is to apply fusion of dynamic and static parts of gait. In this paper a new fusion rule applied to two-stage data fusion framework is suggested. Before moving to fusion it is suggested to reduce the dimensionality of datasets and obtain different sets of features by using feature selection and feature extraction algorithms. The first stage consists of fusing of different classifiers for each datasets of both dynamic and static features. All datasets representing dynamic and static parts of gait fused together on the second stage. It was shown that using this fusion framework and the suggested fusion rule helps to increase the CCR when both static and dynamic parts of gait used.

## Acknowledgment

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

## References

- [1] 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, Vol. 90, No. 1, pp. 1-41, 2003.
- [2] L. Lee and W.E.L. Grimson. *Gait analysis for recognition and classification*, Proceedings of the IEEE International Conference on Face and Gesture Recognition, pp. 155-161, 2002.
- [3] A. Kale, N. Cuntoor and R. Chellappa. *A framework for activity-specific human recognition*, Proceedings of the International Conference on Acoustics, Speech and Signal Processing, Orlando, FL, 2002.
- [4] A.F. Bobick and A. Johnson. *Gait extraction and description by evidence-gathering*, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 423-430, 2001.
- [5] 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, pp. 366-371, Washington, DC, 2002.
- [6] L. Wang, H.Z. Ning, T.N. Tan and W.M. Hu. *Fusion of static and dynamic body biometrics for gait recognition*, Proceedings of the Ninth International Conference on Computer Vision, Vol. II, pp.1449-1454, 2003.
- [7] D.K. Wagg and M.S. Nixon. *Automated markless extraction of walking people using deformable contour models* Computer Animation and Virtual Worlds, Vol. 15. No. 3, pp 399-406, 2004.
- [8] 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, Vol. II, pp. 776-782 2004.
- [9] A. Kale, A.K. RoyChowdhury and R. Chellappa. *Fusion of gait and face for human identification*, Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, Vol 5, pp. V- 901-4, 2004.
- [10] M.M. Kokar and J.A. Tomasik. *Data vs decision fusion in the category theory framework*, FUSION 2001, 2001.
- [11] A.I. Bazin and M.S. Nixon. *Probabilistic combination of static and dynamic gait features for verification*, accepted to SPIE Defense and Security Symposium, April 2005.
- [12] L. Wang, H. Ning, T. Tan and W. Hu. *Fusion of Static and Dynamic Body Biometrics for Gait Recognition*, Proceedings of the International Conference on Pattern Recognition, pp. 115-118, 2004.

- [13] D.A. Socolinsky and A. Selinger. *Thermal face recognition over time*, Proceedings of Proceedings of the 17th International Conference on Pattern Recognition, Vol. 4, pp. 187-190, 2004.
- [14] P. Phillips, D. Blackburn, M. Bone, P. Grother, R. Micheals and E. Tabassi. *Face recognition vendor test 2002 (FRVT 2002)*, Available at <http://www.frvt.org/FRVT2002/default.htm>, 2002.
- [15] A. Lanitis, C.J. Taylor and T.F. Cootes. *Modeling the process of ageing in face images*, Proceedings of the Seventh IEEE International Conference on Computer Vision, Vol. 1, pp. 131-136, 1999.
- [16] A. Lanitis and C.J. Taylor. *Towards automatic face identification robust to ageing variation*, Proceedings of Fourth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 391-396, 2000.
- [17] A. Lanitis, C.J. Taylor and T.F. Cootes. *Toward automatic simulation of aging effects on face images*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 4, pp. 442-455, 2002.
- [18] A. Lanitis, C. Draganova and C. Christodoulou. *Comparing different classifiers for automatic age estimation*, IEEE Transactions on Systems, Man, and Cybernetics- PartB: Cybernetics, Vol. 34, No. 1, pp. 621-628, 2004.
- [19] 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, Vol. 27, No. 2, 162-177, 2005.
- [20] Z. Liu, and S. Sarkar. *Simplest Representation Yet for Gait Recognition: Averaged Silhouette*, Proceedings of the 17th International Conference on Pattern Recognition, Vol. 2, 704-711, 2004.
- [21] Z. Liu, L. Malave and S. Sarkar. *Studies on Silhouette Quality and Gait Recognition*, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 4, 211-214, 2004.
- [22] Z. Liu, and S. Sarkar. *Effect of Silhouette Quality on Hard Problems in Gait Recognition*, Accepted for future publication: IEEE Transactions on Systems, Man, and Cybernetics- PartB: Cybernetics, 2005.
- [23] J. Kittler, M. Hatef, R.P.W. Duin and J. Matas. *On combining classifiers*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 3, pp. 226-239, 1998.
- [24] C. Tremblay and P. Valin. *Experiments on individual classifiers and on fusion of set of classifiers* Proceedings of the Fifth International Conference on Information Fusion, Vol. 1, pp. 272-277, 2002.
- [25] X. Lu, Y. Wang and A.K. Jain. *Combining classifiers for face recognition*, Proceedings of the IEEE International Conference on Multimedia and Expo, Baltimore, MD, Vol. 3, pp. 13-16, 2003.
- [26] O. Melnik, Y. Vardi and C. Zhang. *Mixed group ranks: preferences and confidence in classifier combination*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, No. 8, pp. 973-981, 2004.
- [27] 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, pp.66-72, 2002.
- [28] D.K. Wagg and M.S. Nixon. *On automated model-based extraction and analysis of gait* Proceedings of 6th IEEE International Conference on Automatic Face and Gesture Recognition, pp. 11-16, 2004.
- [29] G.V. Veres, L. Gordon, J.N. Carter, M.S. Nixon. *Feature extraction or feature selection in silhouette-based gait recognition?* Submitted to International Journal of Pattern Recognition and Artificial Intelligence.
- [30] P. Somol, P. Pudil, J. Novovicova and P. Paclil. *Adaptive floating search methods in feature selection* Pattern Recognition Letters, Vol. 20, pp. 1157-1163, 1999.
- [31] G.V. Veres, J.N. Carter and M.S. Nixon. *Towards automatic gait recognition* Submitted to IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics.