

Research Article

Human Gait Recognition Based on Multiview Gait Sequences

Xiayi Huang and Nikolaos V. Boulgouris

Department of Electronic Engineering, Division of Engineering, King's College London WC2R2LS, UK

Correspondence should be addressed to Nikolaos V. Boulgouris, nikolaos.boulgouris@kcl.ac.uk

Received 6 June 2007; Revised 10 October 2007; Accepted 23 January 2008

Recommended by Juwei Lu

Most of the existing gait recognition methods rely on a single view, usually the side view, of the walking person. This paper investigates the case in which several views are available for gait recognition. It is shown that each view has unequal discrimination power and, therefore, should have unequal contribution in the recognition process. In order to exploit the availability of multiple views, several methods for the combination of the results that are obtained from the individual views are tested and evaluated. A novel approach for the combination of the results from several views is also proposed based on the relative importance of each view. The proposed approach generates superior results, compared to those obtained by using individual views or by using multiple views that are combined using other combination methods.

Copyright © 2008 X. Huang and N. V. Boulgouris. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

Gait recognition [1] aims at the identification of individuals based on their walking style. Recognition based on human gait has several advantages related to the unobtrusiveness and the ease with which gait information can be captured. Unlike other biometrics, gait can be captured from a distant camera, without drawing the attention of the observed subject. One of the earliest works studying human gait is that of Johansson [2], who showed that people are able to recognize human locomotion and to identify familiar persons, by presenting a series of video sequences of different patterns of motion to a group of participants. Later, Cutting and Kozlowski in [3] used movinglight displays (MLDs) to further show the human ability for person identification and gender classification.

Although several approaches have been presented for the recognition of human gait, most of them limit their attention to the case in which only the side view is available since this viewing angle is considered to provide the richest information of the gait of the walking person [4–7]. In [8], an experiment was carried out using two views, namely, the frontal-parallel view and the side view, from which the silhouettes of the subjects in two walking stances were extracted. This approach exhibited higher recognition accuracy for the frontal-parallel view than that of the side view. The side view was

also examined in [9] together with another view from a different angle, and the static parameters, such as the height of the walking person, as well as distances between body parts, were used in the template matching. Apart from the recognition rate, results were also reported based on a small sample set using a confusion metric which reflects the effectiveness of the approach in the situation of a large population of subjects. The authors in [10] synthesize the side view silhouettes from those captured by multiple cameras employing visual hull techniques. In [11], the perspective projection and optical flow-based structure of motion approach was taken instead. In [12], information from multiple cameras is gathered to construct a 3D gait model.

Among recent works, the authors in [13] use improved discriminant analysis for gait recognition, the authors in [14] use information of gait shape and gait dynamics, while the authors in [15] use a gait energy image (GEI). However, all above approaches are based only on side view sequences.

In this paper, we use the motion of body (MoBo) database from the Carnegie Mellon University (CMU) in order to investigate the contribution of each viewing direction to the recognition performance of a gait recognition system. In general, we try to answer the fundamental question: *if several views are available to a gait recognition system, what is the most appropriate way to combine them in order to enhance the performance and the reliability of the system?* We provide

a detailed analysis of the role and the contribution of each viewing direction by reporting recognition results of systems based on each one of the available views. We also propose a novel way to combine the results obtained from different single views. In the proposed approach, we set a weight for each view, based on its importance as it is calculated using statistical processing of the differences between views. The experimental results demonstrate the superior performance of the proposed weighted combination approach in comparison to the single-view approach and other combination methods for multiple views.

The paper is organized as follows. Section 2 presents the recognition performance of individual views in a multiview system. The proposed method for the combination of different views is presented in Section 3. Section 4 reports the detailed results using the proposed approach for the combination of several views. Finally, conclusions are drawn in Section 5.

2. GAIT RECOGNITION USING MULTIPLE VIEWS

The CMU MoBo database does not contain explicitly the reference set and the test sets as in [5]. In this work, we use the “fast walk” sequences as the reference set, and the “slow walk” sequences as the test set. As mentioned in the introduction, our goal is to find out which viewing directions have the greatest contribution in a multiview gait recognition system. To this end, we adopt a simple and straightforward way in order to determine the similarity between gait sequences in the reference and test databases. Specifically, from each gait sequence, taken from a specific viewpoint, we construct a simple template T by averaging all frames in the sequence

$$T = \frac{1}{N_T} \sum_{a=1}^{N_T} t_a, \quad (1)$$

where t_a , $a = 1, \dots, N_T$, are the silhouettes in a gait sequence and N_T is the number of silhouettes. This approach for template construction was also taken in [15–17].

Let T_i , R_j denote the templates corresponding to the i th and the j th subjects in the test database and the reference database, respectively. Their distance is calculated using the following distance metric:

$$d(T_i, R_j) = \|T_i - R_j\| = \left\| \frac{1}{N_{T_i}} \sum_{\alpha=1}^{N_{T_i}} t_{i\alpha} - \frac{1}{N_{R_j}} \sum_{\beta=1}^{N_{R_j}} r_{j\beta} \right\|, \quad (2)$$

where $\|\cdot\|$ is the l^2 -norm and $t_{i\alpha}$, $r_{j\beta}$ are the silhouettes belonging to the i th test subject and j th reference subject, respectively. The associated frame indices α and β run from 1 to the total number of silhouettes in a sequence (N_{T_i} and N_{R_j} , resp.). Essentially, a template is produced for each subject by averaging all silhouettes in the gait sequence. Specifically, the Euclidean distance between two templates is taken as a measure of their dissimilarity. In practice, this means that a smaller template distance corresponds to a closer match between two compared subjects.

In order to evaluate the contribution of various viewing directions in the human gait recognition, we choose MoBo

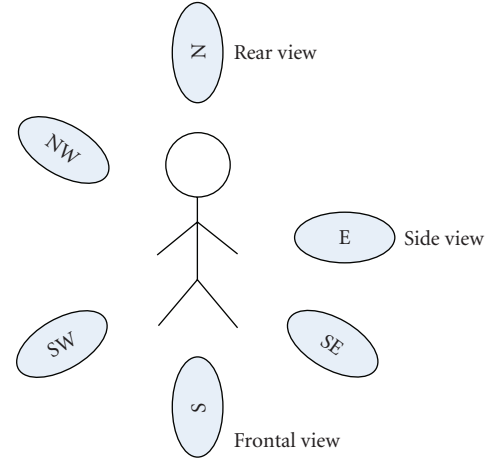


FIGURE 1: Camera arrangement in the CMU MoBo database. Six cameras are oriented clockwise in the east, southeast, south, southwest, northwest, north, with the walking subject facing toward the south.

TABLE 1: The recognition rates of the five viewing directions reported at rank 1 and rank 5.

| Camera location | Rank 1(%) | Rank 5(%) |
|-----------------|-----------|-----------|
| East | 84 | 92 |
| Southeast | 64 | 76 |
| South | 88 | 96 |
| Northwest | 76 | 92 |
| Southwest | 72 | 76 |

database [18] from the CMU which contains walking subjects captured from six cameras located in positions as shown in Figure 1. The database consists of walking sequences of 23 male and 2 female subjects, who were recorded performing four kinds of activities, that is, fast walk, slow walk, and so on. Before the application of our methodologies, we use bounding boxes of silhouettes, then align and normalize all silhouettes so that they have uniform dimensions, that is, 128 pixels tall and 80 pixels wide, in order to eliminate height differences of the walking subjects. We use five (see Figure 2) out of the six available viewing directions, omitting the north view, since it is practically identical to the south view (i.e., the frontal view). The cumulative match scores for each of these five viewing directions are shown in Figure 4, and the recognition rates at rank 1 and rank 5 are reported in Table 1.

One can see clearly from Table 1 that the results obtained using the south and the east viewing directions are the best, especially at rank 1. Results achieved using the rest of the viewing directions are worse. This is a clear indication that the south and the east viewing directions capture most of the gait information of the walking subjects and, therefore, are the most discriminant viewing directions. In the next section, we will show how to combine results from several viewing directions in order to achieve improved recognition performance.

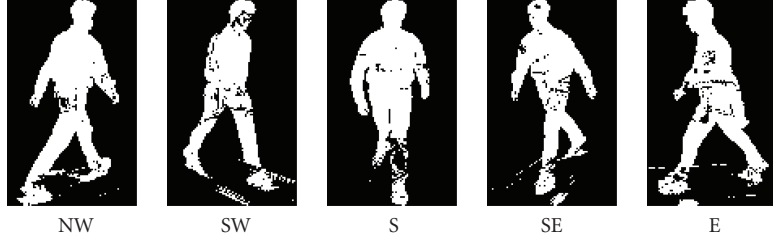


FIGURE 2: Available views for multiview gait recognition.

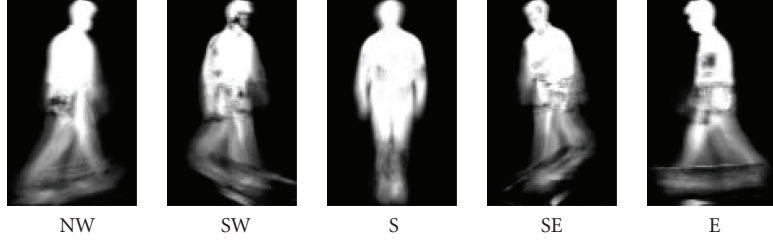


FIGURE 3: Templates constructed using the five available views.

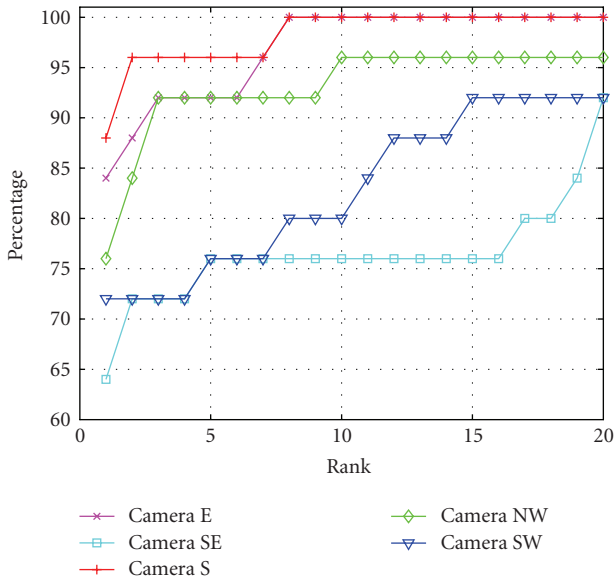


FIGURE 4: Cumulative match scores for five viewing directions, namely, the east, southeast, south, southwest, and the northwest.

3. COMBINATION OF DIFFERENT VIEWS USING A SINGLE DISTANCE METRIC

In this section, we propose a novel method for the combination of results from different views in order to improve the performance of a gait recognition system. In our approach, we use weights in order to reflect the importance of each view during the combination. This means that instead of using a single distance for the evaluation of similarity between walking persons i and j , we use multiple distances between the

respective views and combine them in a total distance which is given by

$$D(T_i, R_j) = \sum_{v=1}^V w_v d_v(T_i, R_j), \quad (3)$$

where V is the total number of available views. Therefore, our task is to determine the weights w_v , which yield smaller total distance when $i = j$, and larger when $i \neq j$.

Suppose that d_{f_v} , $v = 1, 2, \dots, V$, are random variables representing the distances between a test subject and its corresponding reference subjects (i.e., “within class” distance), and d_{b_v} , $v = 1, 2, \dots, V$, are random variables representing the distances between a test subject and a reference subject other than its corresponding subject (i.e., “between class” distance).

In order to maximize the efficiency of our system, we first define the distance D_f between corresponding subjects in the reference and test databases:

$$D_f = \sum_{v=1}^V w_v d_{f_v} = \mathbf{w}^T \cdot \mathbf{d}_f, \quad (4)$$

and the weighed distance between noncorresponding subjects:

$$D_b = \sum_{v=1}^V w_v d_{b_v} = \mathbf{w}^T \cdot \mathbf{d}_b. \quad (5)$$

In an ideal gait recognition system, D_f should always be smaller than D_b . In practice, a recognition error takes place whenever $D_b < D_f$. Therefore, the probability of error is

$$\begin{aligned} P_e &= P(D_b < D_f) \\ &= P(\mathbf{w}^T \cdot \mathbf{d}_b < \mathbf{w}^T \cdot \mathbf{d}_f) \\ &= P(\mathbf{w}^T \cdot (\mathbf{d}_b - \mathbf{d}_f) < 0). \end{aligned} \quad (6)$$

We define the random variable z as

$$z = \mathbf{w}^T \cdot (\mathbf{d}_b - \mathbf{d}_f) \quad (7)$$

if we assume that \mathbf{d}_b and \mathbf{d}_f are normal random vectors, then z is a normal random variable with probability density distribution

$$P(z) = \frac{1}{\sqrt{2\pi}\sigma_z} e^{-(1/2)((z-m_z)^2/\sigma_z^2)}, \quad (8)$$

where m_z is the mean value of z , σ_z is the variance of z .

Therefore, using (7) and (8), the probability of error in (6) is expressed as

$$P_e = P(z < 0) = \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}\sigma_z} e^{-(1/2)((z-m_z)^2/\sigma_z^2)} dz. \quad (9)$$

Furthermore, if $q = (z - m_z)/\sigma_z$, then the above expression is equivalent to

$$P_e = \int_{-\infty}^{-m_z/\sigma_z} \frac{1}{\sqrt{2\pi}} e^{-(1/2)q^2} dq. \quad (10)$$

The probability of error can therefore be minimized by minimizing $-m_z/\sigma_z$, or equivalently by maximizing m_z/σ_z . To this end, we have to calculate m_z and σ_z . If $E\{\cdot\}$ denotes statistical expectation, then the mean value of z is

$$\begin{aligned} m_z &= E\{z\} \\ &= E\{\mathbf{w}^T (\mathbf{d}_b - \mathbf{d}_f)\} \\ &= \mathbf{w}^T (E\{\mathbf{d}_b\} - E\{\mathbf{d}_f\}) \\ &= \mathbf{w}^T (\mathbf{m}_{d_b} - \mathbf{m}_{d_f}), \end{aligned} \quad (11)$$

where \mathbf{m}_{d_b} and \mathbf{m}_{d_f} are the mean vectors of \mathbf{d}_b and \mathbf{d}_f . The variance of z is

$$\begin{aligned} \sigma_z^2 &= E\{(z - m_z)^2\} \\ &= E\{(\mathbf{w}^T (\mathbf{d}_b - \mathbf{d}_f) - \mathbf{w}^T (\mathbf{m}_{d_b} - \mathbf{m}_{d_f}))^2\} \\ &= E\{(\mathbf{w}^T (\mathbf{d}_b - \mathbf{m}_{d_b}) - \mathbf{w}^T (\mathbf{d}_f - \mathbf{m}_{d_f}))^2\} \\ &= E\{(\mathbf{w}^T (\mathbf{d}_b - \mathbf{m}_{d_b}) - \mathbf{w}^T (\mathbf{d}_f - \mathbf{m}_{d_f})) \\ &\quad \times ((\mathbf{d}_b - \mathbf{m}_{d_b})^T \mathbf{w} - (\mathbf{d}_f - \mathbf{m}_{d_f})^T \mathbf{w})\} \\ &= E\{\mathbf{w}^T (\mathbf{d}_b - \mathbf{m}_{d_b}) (\mathbf{d}_b - \mathbf{m}_{d_b})^T \mathbf{w} \\ &\quad - \mathbf{w}^T (\mathbf{d}_b - \mathbf{m}_{d_b}) (\mathbf{d}_f - \mathbf{m}_{d_f})^T \mathbf{w} \\ &\quad - \mathbf{w}^T (\mathbf{d}_f - \mathbf{m}_{d_f}) (\mathbf{d}_b - \mathbf{m}_{d_b})^T \mathbf{w} \\ &\quad + \mathbf{w}^T (\mathbf{d}_f - \mathbf{m}_{d_f}) (\mathbf{d}_f - \mathbf{m}_{d_f})^T \mathbf{w}\}. \end{aligned} \quad (12)$$

If we assume that \mathbf{d}_b and \mathbf{d}_f are independent, then

$$\begin{aligned} \sigma_z^2 &= \mathbf{w}^T \cdot E\{(\mathbf{d}_b - \mathbf{m}_{d_b}) (\mathbf{d}_b - \mathbf{m}_{d_b})^T\} \cdot \mathbf{w} \\ &\quad + \mathbf{w}^T \cdot E\{(\mathbf{d}_f - \mathbf{m}_{d_f}) (\mathbf{d}_f - \mathbf{m}_{d_f})^T\} \cdot \mathbf{w} \\ &= \mathbf{w}^T \cdot \Sigma_{d_b} \cdot \mathbf{w} + \mathbf{w}^T \cdot \Sigma_{d_f} \cdot \mathbf{w}. \end{aligned} \quad (13)$$

Therefore, the optimization problem becomes equivalent to maximizing

$$\begin{aligned} \frac{m_z^2}{\sigma_z^2} &= \frac{\mathbf{w}^T \cdot (\mathbf{m}_{d_b} - \mathbf{m}_{d_f}) \cdot (\mathbf{m}_{d_b} - \mathbf{m}_{d_f})^T \cdot \mathbf{w}}{\mathbf{w}^T \cdot \Sigma_{d_b} \cdot \mathbf{w} + \mathbf{w}^T \cdot \Sigma_{d_f} \cdot \mathbf{w}} \\ &= \frac{\mathbf{w}^T \cdot \Sigma_{d_c} \cdot \mathbf{w}}{\mathbf{w}^T \cdot (\Sigma_{d_b} + \Sigma_{d_f}) \cdot \mathbf{w}}, \end{aligned} \quad (14)$$

where

$$\Sigma_{d_c} = (\mathbf{m}_{d_b} - \mathbf{m}_{d_f}) \cdot (\mathbf{m}_{d_b} - \mathbf{m}_{d_f})^T. \quad (15)$$

The maximization of the above quality is reminiscent of the optimization problem that appears in two-class linear discriminant analysis. Trivially, the ratio can be maximized by determining a vector \mathbf{w} that satisfies [19]

$$\Sigma_{d_c} \cdot \mathbf{w} = \Lambda (\Sigma_{d_b} + \Sigma_{d_f}) \cdot \mathbf{w} \quad (16)$$

for some Λ . In the case that we are considering, the optimal \mathbf{w} is given by

$$\mathbf{w} = (\Sigma_{d_b} + \Sigma_{d_f})^{-1} \cdot (\mathbf{m}_{d_b} - \mathbf{m}_{d_f}). \quad (17)$$

If we assume that the distances corresponding to different views are independent, then

$$(\Sigma_{d_b} + \Sigma_{d_f})^{-1} = \begin{pmatrix} \frac{1}{\sigma_{d_{b1}}^2 + \sigma_{d_{f1}}^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_{d_{b2}}^2 + \sigma_{d_{f2}}^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_{d_{bV}}^2 + \sigma_{d_{fV}}^2} \end{pmatrix}, \quad (18)$$

where V is the total number of available views. Therefore, the optimal weight vector is

$$\mathbf{w} = \left(\frac{\mathbf{m}_{d_{b1}} - \mathbf{m}_{d_{f1}}}{\sigma_{d_{b1}}^2 + \sigma_{d_{f1}}^2} \frac{\mathbf{m}_{d_{b2}} - \mathbf{m}_{d_{f2}}}{\sigma_{d_{b2}}^2 + \sigma_{d_{f2}}^2} \cdots \frac{\mathbf{m}_{d_{bV}} - \mathbf{m}_{d_{fV}}}{\sigma_{d_{bV}}^2 + \sigma_{d_{fV}}^2} \right)^T. \quad (19)$$

Of course, the practical application of the above theory requires the availability of a database (other than the test database) which will be used in conjunction with the reference database for the determination of \mathbf{m}_{d_b} , \mathbf{m}_{d_f} , σ_{d_b} , σ_{d_f} . In our experiments, we used the CMU database of individuals walking with a ball for this purpose.

In the ensuing section, we will use the weight vector in (19) for the combination of views and the evaluation of the resulting multiview gait recognition system.

4. EXPERIMENTAL RESULTS

For the experimental evaluation of our methods, we used the MoBo database from the CMU. The CMU database has 25

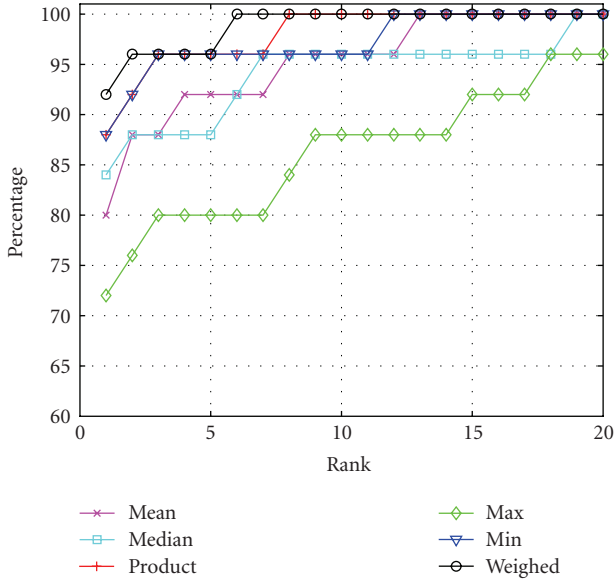


FIGURE 5: Cumulative match scores for the proposed and the other five combination methods.

subjects walking on a treadmill. Although this is an artificial setting that might affect the results, using this database was essentially our only option since this is the only database that provides five views. We used the “fast walk” sequences as reference and the “slow walk” as test sequences. We also used the “with a ball” sequences in conjunction with the reference sequences for the determination of the weights in (19). The comparisons of recognition performance are based on cumulative match scores at rank 1 and rank 5. Rank 1 results report the percentage of subjects in a test set that were identified exactly. Rank 5 results report the percentage of test subjects whose actual match in the reference database was in the top 5 matches. In this section, we present the results generated by the proposed view combination method. These results are compared to the results obtained using different single views and other combination methods.

Initially, we tried several simple methods for the combination of the results obtained using the available views. Specifically, the total distance between two subjects was taken to be equal to the *mean*, *max*, *min*, *median*, and *product* of the distances corresponding to each of the five viewing directions. Such combination approaches were originally explored in [20]. As shown in Figure 5 and Table 2, among all the above combination methods, the most satisfactory results were obtained by using the *Product* and *Min* rules.

In the sequel, we applied the proposed methodology for the determination of the weights in (3). Based on (19), the weights for the combination of the distances of the available views were calculated and are tabulated in Table 3. As seen, the most suitable views seem to be the frontal (east) and the side (south) views since these views are given the greater weights.

The above conclusion is experimentally verified by studying the recognition performance that corresponds to each of the views independently. The cumulative match scores and

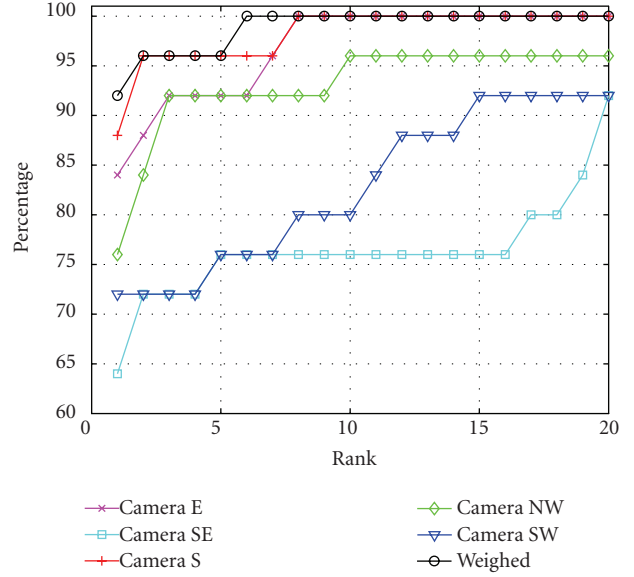


FIGURE 6: Cumulative match scores for five viewing directions and the proposed combination method.

TABLE 2: The recognition rates of the proposed and the other five combination methods.

| Combination method | Rank 1(%) | Rank 5(%) |
|---------------------------|-----------|-----------|
| Mean | 80 | 92 |
| Median | 84 | 88 |
| Product | 88 | 96 |
| Max | 72 | 80 |
| Min | 88 | 96 |
| <i>Weighed (proposed)</i> | 92 | 96 |

TABLE 3: The weights calculated by the proposed method.

| View | East | Southeast | South | Northwest | Southwest |
|--------|---------------|-----------|---------------|-----------|-----------|
| Weight | 0.3332 | 0.0603 | 0.4036 | 0.1188 | 0.0842 |

TABLE 4: The recognition rates of the five viewing directions and the proposed combination method.

| View | Rank 1(%) | Rank 5(%) |
|---------------------------|-----------|-----------|
| East | 84 | 92 |
| Southeast | 64 | 76 |
| South | 88 | 96 |
| Northwest | 76 | 92 |
| Southwest | 72 | 76 |
| <i>Weighed (proposed)</i> | 92 | 96 |

the recognition rates that are achieved using each view as well as those achieved by the proposed method are shown in Figure 6 and Table 4, respectively. As we can see, the south and the east views have the highest recognition rates, as well as the highest weights, which means that the weights calculated by the proposed method correctly reflect the

TABLE 5: The verification rates for the single-view and combined-views methods.

| Viewing direction | Verification rate (%) | | | Combination method | Verification rate (%) | | |
|---------------------------|-----------------------|------------|------------|---------------------------|-----------------------|------------|------------|
| | FAR 5% | FAR 10% | FAR 20% | | FAR 5% | FAR 10% | FAR 20% |
| East | 88 | 96 | 96 | Mean | 88 | 92 | 96 |
| Southeast | 68 | 72 | 76 | Median | 92 | 94 | 96 |
| South | 92 | 96 | 100 | Product | 92 | 96 | 96 |
| Northwest | 80 | 92 | 92 | Max | 72 | 76 | 84 |
| Southwest | 76 | 76 | 84 | Min | 92 | 96 | 100 |
| <i>Weighed (proposed)</i> | 96 | 100 | 100 | <i>Weighed (proposed)</i> | 96 | 100 | 100 |

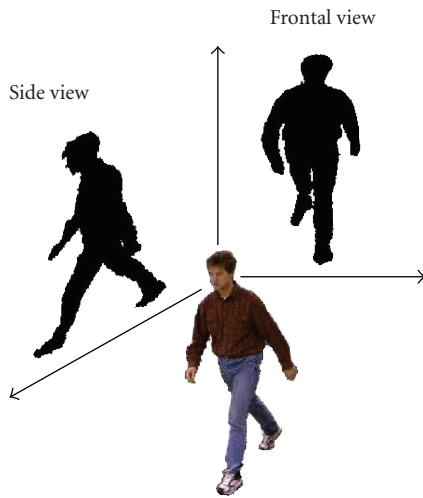


FIGURE 7: Frontal view and side view.

importance of the views. The results obtained by the proposed combination method are superior to those obtained from single views.

Since superior results are generally achieved using the frontal (east) and side (south) views (see Figure 7), the proposed method was also used to combine those two views. Figure 8 shows that the combination of the east and the south views using the proposed method has much better performance than using the views individually. It is interesting to note that, in theory, using two views should be sufficient for capturing the 3D information in a sequence. Although here we use silhouettes (so there is no texture that could be used for the estimation of 3D correspondence), it appears that the combination of these two views seems very efficient. By trying other combinations of the two views, we discovered that the optimal combination of the east and the south view is the only one which outperforms all single views.

The proposed system was also evaluated in terms of verification performance. The most widely used method for this task is to present *receiver operating characteristic* (ROC) curves. In an access control scenario, this means calculating the probability of positive recognition of an authorized subject versus the probability of granting access to an unauthorized subject. In order to calculate the above probabilities, different thresholds were set for examining the distances

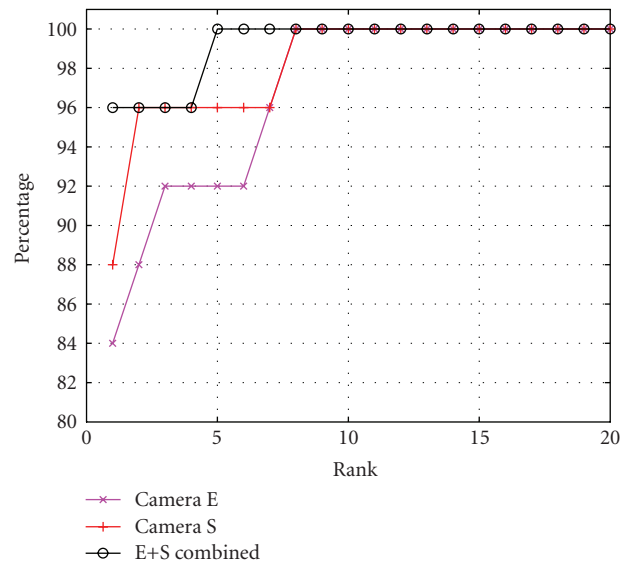


FIGURE 8: Cumulative match scores for the east and the south viewing directions and the proposed combination method.

between the test and reference sequences. We calculated the distances for the five intraviews, and combined them using weights and five other existing methods mentioned in the previous section. Figure 9 shows the ROC curves of the methods using single views and combined views. In Table 5, verification results are presented at 5%, 10%, and 20% false alarm rate for the proposed method and the existing methods. As seen, within the five viewing directions, the frontal (east) and side (south) views have the best performances; and among the five existing combination methods, the *Min* method obtains the best results. As expected, the proposed method has superior verification performance, in comparison to any of the single-view methods as well as in comparison to the other methods for multiview recognition.

5. CONCLUSION

In this paper, we investigated the exploitation of the availability of various views in a gait recognition system using the MoBo database. We showed that each view has unequal discrimination power and therefore has unequal contribution to the task of gait recognition. A novel approach was proposed for the combination of the results of different views

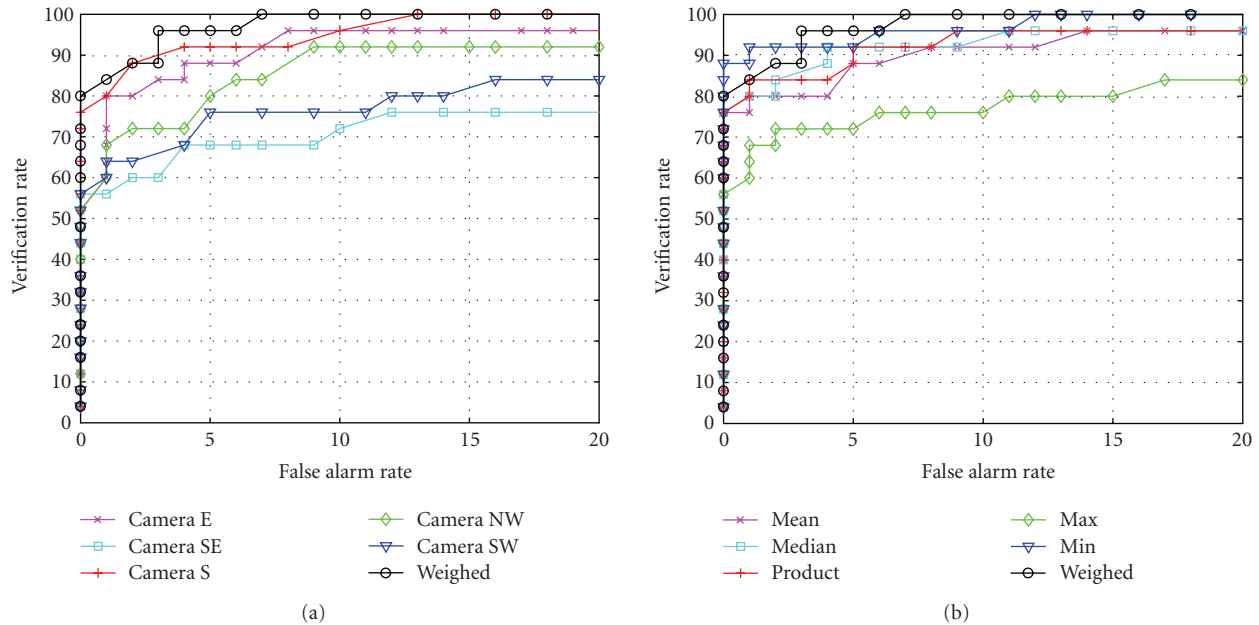


FIGURE 9: The ROC curves: (a) single-view methods and the proposed method, (b) the proposed and five existing combination methods.

into a common distance metric for the evaluation of similarity between gait sequences. By using the proposed method, which uses different weights in order to exploit the different importance of the views, improved recognition performance was achieved in comparison to the results obtained from individual views or by using other combination methods.

ACKNOWLEDGMENT

This work was supported by the European Commission funded FP7 ICT STREP Project ACTIBIO, under contract no. 215372

REFERENCES

- [1] N. V. Boulgouris, D. Hatzinakos, and K. N. Plataniotis, "Gait recognition: a challenging signal processing technology for biometric identification," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 78–90, 2005.
- [2] G. Johansson, "Visual motion perception," *Scientific American*, vol. 232, no. 6, pp. 76–88, 1975.
- [3] J. E. Cutting and L. T. Kozlowski, "Recognizing friends by their walk: gait perception without familiarity cues," *Bulletin Psychonomic Society*, vol. 9, no. 5, pp. 353–356, 1977.
- [4] L. Lee and W. E. L. Grimson, "Gait analysis for recognition and classification," in *Proceedings of the 5th IEEE International Conference on Automatic Face and Gesture Recognition (FGR '02)*, pp. 148–155, Washington, DC, USA, May 2002.
- [5] S. Sarkar, P. J. Phillips, Z. Liu, I. R. Vega, P. Grother, and K. W. Bowyer, "The humanID gait challenge problem: data sets, performance, and analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 162–177, 2005.
- [6] N. V. Boulgouris, K. N. Plataniotis, and D. Hatzinakos, "Gait recognition using linear time normalization," *Pattern Recognition*, vol. 39, no. 5, pp. 969–979, 2006.
- [7] M. Ekinici, "Gait recognition using multiple projections," in *Proceedings of the 7th IEEE International Conference on Automatic Face and Gesture Recognition (FGR '06)*, pp. 517–522, Southampton, UK, April 2006.
- [8] R. T. Collins, R. Gross, and J. Shi, "Silhouette-based human identification from body shape and gait," in *Proceedings of the 5th IEEE International Conference on Automatic Face and Gesture Recognition (FGR '02)*, pp. 351–356, Washington, DC, USA, May 2002.
- [9] A. Y. Johnson and A. F. Bobick, "A multi-view method for gait recognition using static body parameters," in *Proceedings of the 3rd International Conference on Audio and Video-Based Biometric Person Authentication (AVBPA '01)*, pp. 301–311, Halmstad, Sweden, June 2001.
- [10] G. Shakhnarovich, L. Lee, and T. Darrell, "Integrated face and gait recognition from multiple views," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '01)*, vol. 1, pp. 439–446, Kauai, Hawaii, USA, December 2001.
- [11] A. Kale, A. K. R. Chowdhury, and R. Chellappa, "Towards a view invariant gait recognition algorithm," in *Proceedings of IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS '03)*, pp. 143–150, Miami, Fla, USA, July 2003.
- [12] G. Zhao, G. Liu, H. Li, and M. Pietikainen, "3D gait recognition using multiple cameras," in *Proceedings of the 7th IEEE International Conference on Automatic Face and Gesture Recognition (FGR '06)*, pp. 529–534, Southampton, UK, April 2006.
- [13] D. Tao, X. Li, X. Wu, and S. J. Maybank, "General tensor discriminant analysis and gabor features for gait recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 10, pp. 1700–1715, 2007.
- [14] Z. Liu and S. Sarkar, "Improved gait recognition by gait dynamics normalization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 6, pp. 863–876, 2006.
- [15] J. Man and B. Bhanu, "Individual recognition using gait energy image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 2, pp. 316–322, 2006.

- [16] Z. Liu and S. Sarkar, "Simplest representation yet for gait recognition: averaged silhouette," in *Proceedings of the 17th International Conference on Pattern Recognition (ICPR '04)*, vol. 4, pp. 211–214, Cambridge, UK, August 2004.
- [17] G. V. Veres, L. Gordon, J. N. Carter, and M. S. Nixon, "What image information is important in silhouette-based gait recognition?" in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04)*, vol. 2, pp. 776–782, Washington, DC, USA, June–July 2004.
- [18] R. Gross and J. Shi, "The cmu motion of body (MoBo) database," Tech. Rep. CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon University, Pittsburgh, Pa, USA, 2001.
- [19] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, John Wiley & Sons, New York, NY, USA, 2001.
- [20] 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.