Available online at www.sciencedirect.com
ScienceDirect
Image and Vision Computing xxx (2007) xxx–xxx
www.elsevier.com/locate/imavis

Automatic detection of abnormal gait

Christian Bauckhage [a,*], John K. Tsotsos [b], Frank E. Bunn [c]

[a] Deutsche Telekom Laboratories, Berlin, Germany
[b] Centre for Vision Research, York University, Toronto, Ont., Canada
[c] StressCam Operations and Systems Ltd., Toronto, Ont., Canada

Received 4 November 2005; received in revised form 26 September 2006; accepted 20 October 2006

Abstract

Analysing human gait has found considerable interest in recent computer vision research. So far, however, contributions to this topic exclusively dealt with the tasks of person identification or activity recognition. In this paper, we consider a different application for gait analysis and examine its use as a means of deducing the physical well-being of people. Understanding the detection of unusual movement patterns as a two-class problem suggests using support vector machines for classification. We present a homeomorphisms between 2D lattices and binary shapes that provides a robust vector space embedding of segmented body silhouettes. Experimental results demonstrate that feature vectors obtained from this scheme are well suited to detect *abnormal* gait. Wavering, faltering, and falling can be detected reliably across individuals without tracking or recognising limbs or body parts.

Keywords: Gait analysis; Shape encoding; Vector space embedding; SVM classification

1. Introduction

In times of increased interest in biometrics for authentication and access control, it comes as little surprise that gait analysis has become an active area of research. From a computer vision point of view, there are two major reasons why human gait is an appealing cue for person identification: in contrast to fingerprint- or retina scans, gait can be analysed from a distance. Yet, although no two body movements are ever the same, gait is as characteristic of an individual as other biometrics.

Scientific interest in this curiously familiar fact dates back to the French physician and photographer Etienne-Jules Marey (1830–1904). In the 1920s, Nikolai A. Bernstein conducted first clinical studies in Moscow, which revealed that gait is a unique personal trait. Unknown in the West, these findings were corroborated by medical studies from the 1960s [1]. Subsequent psychological experiments conducted in the 1970s revealed that humans effortlessly recognise people by the way they walk [2].[1]

Compared to its tradition in medicine, the history of human gait as a research topic in computer vision is rather brief. Nevertheless, the spectrum of methods that have been proposed for vision-based gait analysis is already vast. Recent and detailed surveys of different approaches can be found in [3,4]. From these surveys it becomes apparent that, although some contributions to the problem did not extract human silhouettes from image sequences [5,6], the majority of recent work relies on shape analysis. Little and Boyd [7] derive the shape of motion from optical flow and use this to compute characteristic features. Bobick and Johnson [8] subdivide the binary silhouette of walking

* Corresponding author. Tel.: +1 416 7362100x33972; fax: +1 416 736 5857.
E-mail address: bauckhag@cs.yorku.ca (C. Bauckhage).

[1] Although not verified experimentally, humankind knew this for a long time. For instance, in his 1599 tragedy *Julius Caesar*, Shakespeare has Cassius saying: "T' is Cinna; I do know him by his gait."

0262-8856/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2006.10.004

people into body parts and measure different relations among limbs. Similarly, Lee and Grimson [9] approximate silhouettes by ellipses and compute feature descriptors therefrom. Veres et al. [10] apply analysis of variance techniques to determine which parts of a binary shapes provide the most characteristic information for gait analysis. Collins et al. [11] propose template matching of body silhouettes as a baseline algorithm in gait recognition. Sarkar et al. [4] favour temporal shape statistics. Tolliver and Collins [12] store normalised silhouettes as vertices of similarity graphs and recover significant parts of the walking cycle from computing the eigenvectors of the Laplacian matrix of these graphs. BenAbdelkader et al. [13] compute self-similarity plots from silhouettes and apply subspace methods to recognise individuals. Other authors base their work on a statistical theory of shape developed by Kendal [14], Boyd [15], Veeraraghavan et al. [16] and Wang et al. [17] apply Procrustean distances to compare and classify human silhouettes.

Apart from the popularity of shape-based approaches, the papers considered in this brief survey reveal that gait analysis is almost exclusively being applied to identification tasks. However, gait can disclose more than identity. In psychological experiments, human subjects easily recognise different types of motion (e.g., walking, dancing, wavering, etc.) when asked to interpret moving light displays (obtained from filming subjects wearing light bulbs on their joints) or sequences of body silhouettes [18]. The reader may verify this claim by considering the two series of silhouettes shown in Fig. 1. The shapes in this figure were extracted from videos of walking people. While there is nothing unusual in the series in Fig. 1(a), which shows a subject slowly moving away from the camera, the subject in the series in Fig. 1(b) obviously has difficulties walking. The movements captured in this sequence suggest the person is in a dubious condition.

As there are numerous applications for the detection of *abnormal* gait, it seems worthwhile to explore techniques that can accomplish this. The work presented in the following is a first step in this direction. First, we describe a novel, fast and robust shape encoding scheme. We show that this encoding scheme provides a straightforward vector space embedding of shapes. Therefore, it leads to a rich choice of pattern recognition and machine learning techniques that deal with the classification of vectorial data. In Section 4, we present experimental results obtained from support vector classification of gait, which demonstrates that our approach allows for characterising types of movements across individuals. A summary and an outlook completes this contribution.

2. Robust shape encoding

Our approach to detecting abnormal gait from video was guided by the following considerations. Fig. 1 indicates that shape dynamics provide a strong cue for distinguishing normal from abnormal gait. Therefore, it appears natural to follow the general trend identified in the introduction and to base gait classification on shape analysis. In its most rudimentary form, the classification task of detecting abnormal gait can be treated as a two-class problem. Therefore, it might be solved using a support vector machine approach, because support vector machines are known to reliably accomplish two-class classification. Applying support vector machines, however, requires a vector space embedding of the silhouettes we are concerned with. In contrast to the identification task, abnormal gait detection should abstract from personal traits. The required vector space embedding should therefore capture general aspects of body silhouettes. Ideally, it also would be insensitive to noisy boundaries and would not require much computational effort. The shape encoding scheme presented in this section meets all these criteria.

Given binary images, as shown in Fig. 1, produced by some segmentation process, we understand a shape $\mathcal{S}$ to be a set of L pixels, $\mathcal{S} = \{\vec{p}_k \in \mathbb{R}^2 | k = 1, \ldots, L\}$. Fig. 2 visualises the following procedure that computes an $m \times n$ array $\mathbf{B}$ of boxes $\mathcal{B}_{ij}$, which can be thought of as a coarser representation of a shape:

(a) compute the bounding box $\mathcal{B}(\mathcal{S})$ of a pixel set $\mathcal{S}$
(b) subdivide it into n vertical slices of equal width;
(c) compute the bounding boxes $\mathcal{B}(\mathcal{S}_j)$ for each of the resulting pixel sets $\mathcal{S}_j$ where $j = 1, \ldots, n$;

Fig. 1. Silhouettes extracted from videos of walking people. Apparently, the information contained in sequences of shapes allows to distinguish between *normal* and *abnormal* gait. (a) Person slowly walking away from the camera; (b) Person wavering towards the camera.

Please cite this article in press as: C. Bauckhage et al., Automatic detection of abnormal gait, Image Vis. Comput. (2007), doi:10.1016/j.imavis.2006.10.004

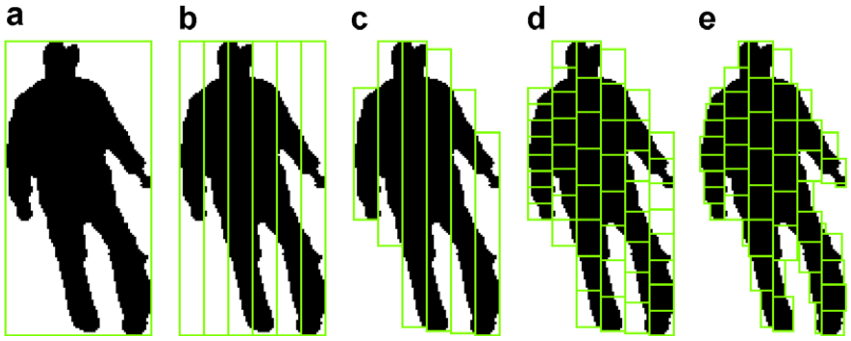

Fig. 2. Example of using bounding box splitting to map a 8×6 lattice onto a shape.

(d) subdivide each $\mathcal{B}(\mathcal{S}_j)$ into m horizontal slices of equal height;
(e) compute the bounding boxes $\mathcal{B}_{ij} = \mathcal{B}(\mathcal{S}_{ij})$ for each of the resulting pixel sets $\mathcal{S}_{ij}$ where $i = 1, \ldots, m$;

Obviously, this procedure is linear in mn. It does not require the computation of interpixel relations, and there is no need for repeated maximisation (or minimisation) steps. Therefore, the procedure's average computational complexity is $O(mn\hat{p})$, where $\hat{p}$ denotes the average number of pixels in a box $\mathcal{B}(\mathcal{S}_{ij})$.

Moreover, each box $\mathcal{B}(\mathcal{S}_{ij})$ as in Fig. 2(e) can be treated as a generalised pixel of height h_{ij} and width W_{ij} at location (x_{ij}, y_{ij}) as in Fig. 3. The storage requirement of a coarse shape representation is therefore a mere $4mn$. For small values of m and n, this iterative process of bounding box splitting yields a fast and storage efficient abstraction of shapes. In addition, Fig. 3 also exemplifies that already moderate array dimensions $m \times n$ can produce fairly accurate representations. This subjective impression is objectified by the graphs in Fig. 5. They summarise measurements obtained from 1178 silhouettes of an average size of 18878 pixels that were extracted from 6 video sequences of walking people. Given the height h and width w of a shape's initial bounding box, the array dimension m was computed as a function of n (see Fig. 4):

$$m(n) = \left\lfloor \frac{h}{w} n \right\rfloor \tag{1}$$

where $\lfloor x \rfloor$ indicates rounding $x \epsilon \mathbb{R}$ to the nearest lower integer, i.e., $\lfloor x \rfloor sup\{y \epsilon \mathbb{N} \mid y \leqslant x\}$.

While Fig. 5(a) shows the average compression rate $1 - 2mn/L$ obtained in shape encoding by bounding box splitting, Fig. 5(b) depicts the normalised Hamming distance $D = d_{\mathrm{H}}(\mathcal{S}, \mathbf{B})/L$ between a shape $\mathcal{S}$ and its coarse representation as a box array $\mathbf{B}$. Note that while the compression rate decreases slowly for growing values of n, the normalised reconstruction error decreases quickly. At a compression rate of about 93%, it already drops below 4%. Moreover, from the error bars in Fig. 5(b), we see that at the same level of compression the variance of the error already almost vanishes. Therefore, even box arrays of small sizes provide descriptions that capture the essential properties of a silhouette and are accurate across a wide variety of cases. The runtime required for bounding box splitting will be far from threatening real time constraints that are important in most video-based applications [19].

Furthermore, note that this simple method realizes a homeomorphism between 2D shapes and 2D lattices even though it only relies on basic computational geometry. If we begin counting the lattice coordinates (i,j) of individual boxes starting with the lower left of an array, then boxes $\mathcal{B}(\mathcal{S}_{ij})$ situated below other boxes will always have smaller lattice coordinates i. Also, lattice coordinates j of boxes to the left of other boxes are always smaller (see Fig. 6). Due to this topology preserving nature of the box array representation of shapes, it is straightforward to embed shapes in a vector space.

If the vector $\vec{v}$ denotes the location of the bottom left corner of the initial bounding box of $\mathcal{S}$, w and h denote its width and height and the vector $\vec{u}^{ij}$ denotes the centre of box $\mathcal{B}(\mathcal{S}_{ij})$, then the coordinates

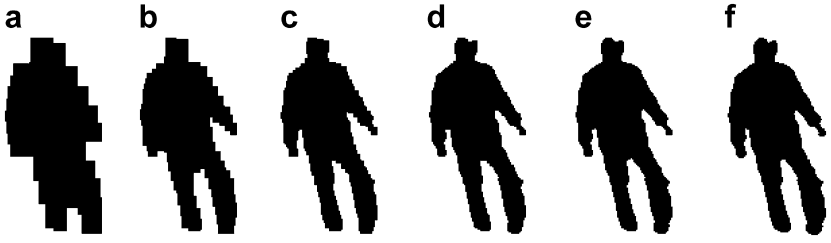

Fig. 3. Approximation of a binary shape by bounding box arrays of increasing dimensions $m \times n$. Given the number of columns n, the number of rows m in these examples was determined according to Eq. (1). (a) 8×4; (b) 16×8; (c) 32×16; (d) 64×32; (e) 128×64; (f) original.

Please cite this article in press as: C. Bauckhage et al., Automatic detection of abnormal gait, Image Vis. Comput. (2007), doi:10.1016/j.imavis.2006.10.004

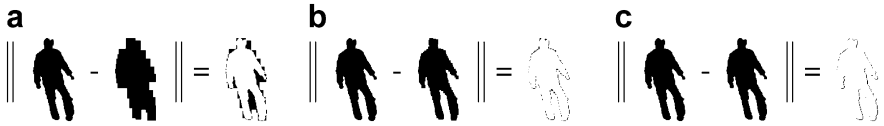

Fig. 4. Visualisation of shrinking Hamming distances d_H between a shape $\mathcal{S}$ and box array approximations $\mathbf{B}$ of growing array dimensions $m \times n$. (a) 8×4; (b) 32×16; (c) 128×64.

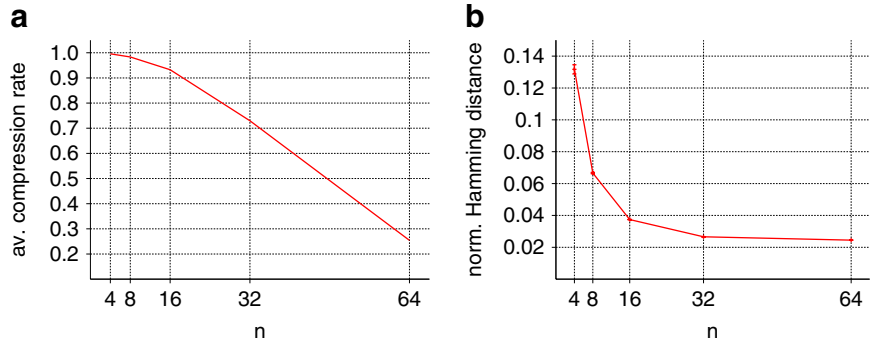

Fig. 5. Average compression rates and normalised Hamming distances obtained from shape encoding using box arrays of dimensions $m \times n$. The number of columns n was chosen as indicated; the corresponding number of rows m resulted from Eq. (1).

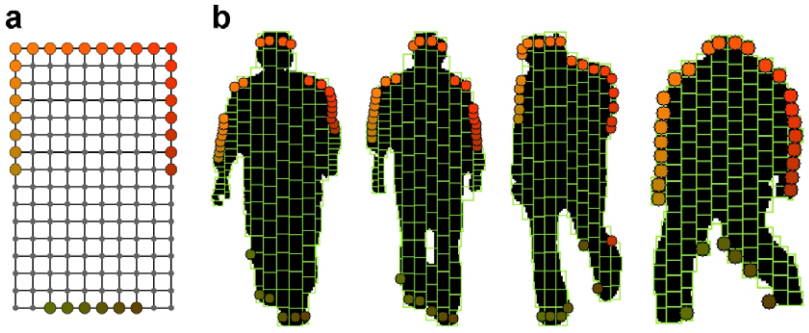

Fig. 6. A sample of $k = 30$ points on a 16×10 lattice and examples of how this sample is mapped onto different silhouettes.

$$\vec{\mu}^{ij} = \begin{pmatrix} (u_x^{ij} - v_x)/w \\ (u_y^{ij} - v_y)/h \end{pmatrix} \tag{2}$$

provide a scale invariant representation of $\mathcal{S}$. Sampling k points of an $m \times n$ lattice therefore allows to represent the shape $\mathcal{S}$ as a vector

$$\vec{r} = [\mu_x^{i_1 j_1}, \mu_y^{i_1 j_1}, \ldots, \mu_x^{i_k j_k}, \mu_y^{i_k j_k}] \epsilon \mathbb{R}^{2k} \tag{3}$$

where $i_\alpha < i_\beta$ if $\alpha < \beta$ and likewise for the index j.

Note that while this embedding in $\mathbb{R}^{2k}$ is scale invariant, it is not invariant against rotations. However, for real world applications of abnormal gait detection, it is reasonable to assume a camera setup where human silhouettes are within the image plane (as are the examples in Fig. 1). Therefore, since shapes of walking people will usually appear to be upright, rotation invariance is not of primary concern for our application. Rather, exceptional feature vectors that result from mapping a lattice onto a somewhat rotated silhouette will be a strong indicator of abnormal gait.

3. Features for gait classification

Choosing array dimensions $m \times n$ that cope with the requirements listed in Section 2 seems an arbitrary task. Obviously, too fine a grid would be more sensitive against distorted shapes or individual traits than a coarser one. A meshing that is too coarse, however, might not capture essential silhouette dynamics. From experimenting with a standard database of shapes, we found that arrays of sizes between a 140 and 240 cells allow for robust classification [19]. According to these empirical findings, a grid size of 16×10 was chosen for the experiments reported below.

Please cite this article in press as: C. Bauckhage et al., Automatic detection of abnormal gait, Image Vis. Comput. (2007), doi:10.1016/j.imavis.2006.10.004

This particular choice was further motivated by experiences from the Fine Arts: the ratio 3:5:8 is often considered to be a pleasing measure of the relative sizes of head, torso and legs of the human body. The two shapes on the left of Fig. 6(b) illustrate that this rule of thumb indeed is reasonable for upright silhouettes. Given this aesthetic basis for choosing $m = 16$ rows, the choice for the number of columns followed a similar path. Observe that 3:5 and 5:8 are approximations of the golden ratio $\phi = \frac{1}{2}(1 + \sqrt{5}) = 1.618\ldots$ So we chose $n = 10$ to approximate a golden rectangle - a rectangle with a side ratio of $1 : \phi$ and which is often considered a pleasing container of human silhouettes.

Having decided the lattice dimensions, it remains to determine which lattice points to sample in order to provide a suitable vector representation for gait classification. Experiments with a standard database of shapes revealed that in particular the points on the boundary of the lattice enable correct classification [19]. Moreover, from inspecting exemplary series of human shapes, it appeared that relations between head and shoulders and between the feet provide reliable cues to distinguish normal upright walking from any kind of wobbling. Fig. 6(b) exemplifies this. The two shapes on the left correspond to instances of the gait cycle of two subjects walking normally. In each case, the head is well above the shoulders, shoulders and arms nearly enclose a right angle and the feet are not too far apart. The silhouettes on the right, in contrast, were extracted from video sequences showing abnormal gait. Here, the relation between head and shoulders as well as between the feet varies arbitrarily. Moreover, as can be seen in Fig. 1, normal gait is also characterised by periodic motion of almost constant frequency of arms and feet, whereas abnormal movements are aperiodic and random. A sample of lattice points that correspond to those parts of a silhouette which depict the head and shoulders and, therefore, the arms and feet will provide valuable information for gait classification.

At this stage of our discussion, the simple mapping between 2D lattices and shapes introduced in Section 2 reveals its full potential. Due to the mapping's topology preserving properties, a sample of points with lattice coordinates on the upper and lower border of the lattice will usually, i.e., in the case of upright silhouettes, correspond to head, shoulders, arms, and feet. Fig. 6 displays the $k = 30$ lattice points we used in our experiments and where they are located on different silhouettes. Obviously, for normal gait, this scheme allows to roughly keep track of limbs. Neither feature tracking algorithms nor recognition procedures have to be applied to infer the location of significant body parts.

Since gait is a temporal phenomenon, information about a single instance of a gait cycle might not be sufficient to determine whether normal or abnormal gait is being observed. In order to incorporate temporal context into the classification process, at each time step t, we consider concatenated feature vectors

$$\vec{s}_t = \vec{r}_t \oplus \vec{r}_{t-1} \oplus \ldots \oplus \vec{r}_{t-\Delta} \tag{4}$$

where, for each $t_l \in \{t, \ldots, -\Delta\}$, we have

$$\vec{r}_{t_l} = [\mu_x^{i_k j_k}(t_l), \mu_y^{i_k j_k}(t_l)], \quad k = 1, \ldots, 30 \tag{5}$$

Therefore, a shape and its recent history will be characterised by means of a high dimensional feature vector $\vec{s}_t \in \mathbb{R}^{2k(\Delta+1)}$. For the experiments described in the next section, Δ was set to 0, 10 and 20, respectively, i.e., the 60 dimensional shape descriptors of the current frame and its predecessors were combined into feature vectors $\vec{s}_t$ of up to 1260 dimensions.

4. Experiments

In order to test the feasibility of our approach to abnormal gait detection, we experimented with a set of videos recorded in front of a green screen. Seven subjects were asked to walk in a normal fashion as well as in an unusual way, e.g., as if they were suffering a balance deficiency such as dizziness. In accordance with the setting of our application domain, the movement in these baseline sequences is either towards or away from the camera, and the angle to the camera varies between 0° and about 45°.

Motion segmentation was done using a simple statistical model where 50 frames of video without walking people were used to determine the colour distribution of the scene background. Subsequent filtering using a 5×5 median followed by a connected component analysis provides shapes, such as those seen throughout this paper.

Concerning our goal of distinguishing between normal and abnormal gait, many techniques developed in the field of pattern recognition are applicable. We opted for support vector machines (SVMs) with a radial basis kernel function because they are known to perform very reliably in two-class problems [20]. Also, the support vector property of these classifiers copes with possibly biased training data: in real world applications of gait analysis, examples of normal gait will outnumber those of abnormal gait. While some classifiers thus may favour the class with more examples, a trained SVM simply predicts according to those training examples which characterise the class boundaries. However, training an SVM requires solving a quadratic optimisation problem, which can be burdensome if the considered feature space is of high dimensionality and there are many training examples. We therefore applied the $\text{SVM}^{\text{light}}$ algorithm [21], which tackles these issues by decomposing the training process into a series of smaller tasks.

Table 1(a) summarises the training parameters of three series of baseline experiments. In each series, the same set of 7 videos showing 4 individuals was used for training. Four of these videos show normal gait. The remaining 3 contain examples of abnormal gait. Due to the varying length of temporal context, the number of frames available for training differs in each experiment. Choosing a temporal context of $\Delta = 20$, for instance, implies that the first twenty frames of each training sequence have to be skipped from

Please cite this article in press as: C. Bauckhage et al., Automatic detection of abnormal gait, Image Vis. Comput. (2007), doi:10.1016/j.imavis.2006.10.004

Table 1
Overview of parameters and results for training and testing abnormal gait detection

Δ	No of frames normal	No of frames abnormal	Leave one out error (%)
(a) Training Parameters and results			
0	1359	1128	26.5
10	1323	1101	11.6
20	1283	1071	6.1

Δ	Gait	No of frames	Accuracy (%)
(b) Test Parameters and results			
0	Normal	1227	73
	Abnormal	1413	61
10	Normal	1195	77
	Abnormal	1382	70
20	Normal	1157	72
	Abnormal	1350	82

training. The figures in the last column of the table are an estimate of the quality of the resulting SVM. Performing a 'leave one out' test on the individual frames of the training data produced the listed error rates. Since the error rates do not vanish, the feature space areas of normal and abnormal gait obviously overlap. However, if the temporal context increases, i.e., if the dimension of the feature space grows, the overlap decreases and thus the error rate decreases. This agrees with intuition and mathematical wisdom: the decision of whether or not a person moves abnormally is more reliable the longer the person is being observed. Also, in higher dimensional vector spaces, data will be more sparse, and there will be more space for partitioning. Further increasing the temporal context will further reduce the error rates. However, due to practical constraints imposed by our application scenario, we did not consider this option: first of all, due to the 'curse of dimensionality', linearly growing feature vectors require an exponentially growing amount of training data. Second of all, very high dimensional feature vectors considerably increase the computation effort of SVM-based classification (cf. e.g., [22]) and thus threaten the real time capability crucial for our system.

Parameters and results obtained on an independent test set are shown in Table 1(b). Here, we considered 7 videos of 5 individuals, 3 of which were not present in the training set. Out of the 7 test videos, 4 show instances of abnormal gait and 3 display usual walking behaviour. Again, due to the different temporal context, the number of frames varies among the experiments. What is noticeable is that, while a growing temporal context considerably improves the detection of abnormal gait, its effect on normal gait detection is inconsistent. Overall, especially for a temporal depth of $\Delta = 20$, the per-frame recognition accuracy can be considered acceptable. However, the rate of false positives seems to be too high to be used in real world applications of abnormal gait detection.

Our solution for this problem is depicted in Figs. 7 and 8. Note that, in order to deal with cluttered, possibly non-static scenes and inhomogeneous illumination conditions as in these examples, the motion segmentation procedure from our baseline experiments was improved using kernel density estimators for background modelling as proposed in [23]. Both figures cover a period of about 30 frames, taken from a sequence of a subject walking abnormally and normally, respectively. The idea of colouring the segmented shapes on the right of each panel was developed in collaboration with a partner from industry. It indicates the percentage of frames that were classified *abnormal* during the last 30 frames. The higher the number of dubious frames rises, the higher is the level of the red colouring. At the beginning of the part of the sequence that is shown in Fig. 7, only a single frame has been classified as abnormal gait. Corresponding to the wavering movement of the subject, however, the percentage of abnormal frames rises continuously throughout the sequence and reaches 100% in the panel at the bottom right.

In this manner, the colour coding illustrates the use of temporal context on a higher level of abstraction. It can

Fig. 7. Example of abnormal gait detection over a period of 32 frames; the temporal context for classification was set to $\Delta = 20$ frames. The colouring of the segmented shapes indicates the percentage of silhouettes in the past 30 frames that were classified to depict *abnormal* gait. Thus at the beginning of this sequence, the subject's gait appeared to be fairly normal whereas at the end each of his last 30 instances has been classified as abnormal gait.

Please cite this article in press as: C. Bauckhage et al., Automatic detection of abnormal gait, Image Vis. Comput. (2007), doi:10.1016/j.imavis.2006.10.004

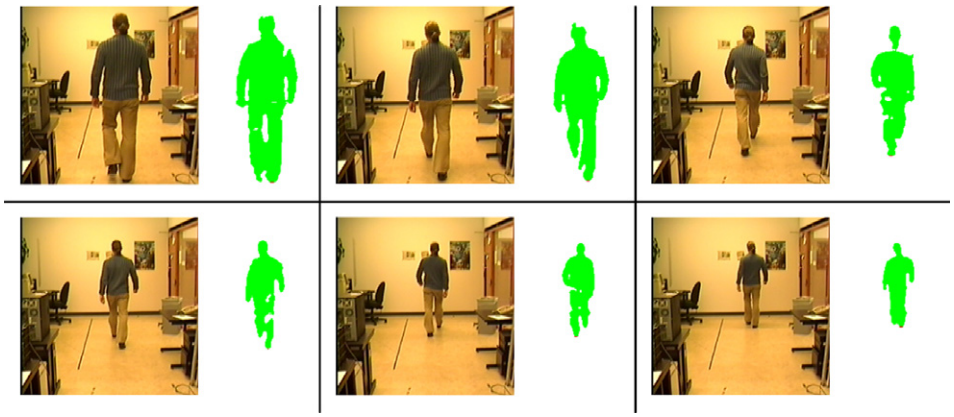

Fig. 8. Example of gait analysis over a period of 31 frames; again, the temporal context for classification was set to $\Delta = 20$ frames. The sequence shown here is an example of normal gait. None of the 31 frames covered in this figure was recognised as an instance of abnormal gait. Correspondingly, the segmented shapes on the right of each panel are homogeneously coloured in green

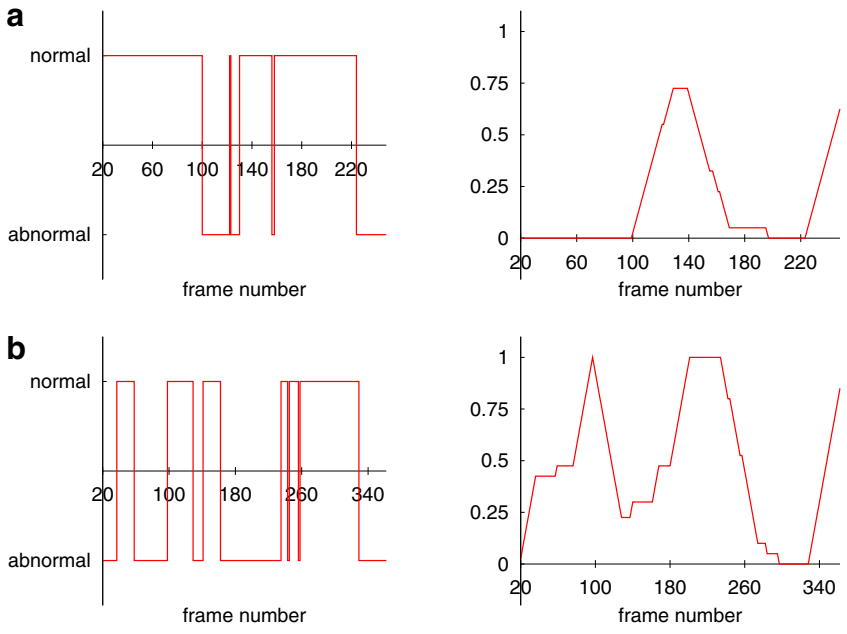

Fig. 9. Per frame classification results and corresponding percentage of dubious frames within the last 30 frames for two short sequences of normal and abnormal gait. (a) Example of a sequence of normal gait; (b) example of a sequence of abnormal gait.

be seen as a temporal filter that acts on the results of frame-wise classification. If one or several frames of a sequence are misclassified, it will have little effect on the general tendency or confidence level that becomes apparent from temporal filtering. This is further illustrated in Fig. 9 which depicts frame-wise classification results and the corresponding results of the temporal context filter for two of the test sequences considered in Table 1(b). In both cases, the temporal context for classification was set to $\Delta = 20$ frames and 30 frames were considered for subsequent temporal smoothing. For both sequences, there are considerably many misclassified frames, which, however, only occur over short periods of time, sometimes as short as a single frame. Consequently, for the sequence showing normal gait, the percentage of dubious frames within the last 30 frames never exceeds about 75%. For the sequence showing abnormal gait, on the other hand, there are several intervals where the percentage of dubious frames clearly exceeds this mark. For real world surveillance applications, it is thus possible to trigger an alarm once a certain level of abnormal gait has been detected over a given period of time.

Please cite this article in press as: C. Bauckhage et al., Automatic detection of abnormal gait, Image Vis. Comput. (2007), doi:10.1016/j.imavis.2006.10.004

5. Summary and outlook

This paper considered automatic gait analysis as a means to deduce if an observed walking pattern appears to be normal or not. In contrast to most contributions to visual gait analysis, the problem dealt with in this paper requires a representation that abstracts from individual gait characteristics but allows for the classification of gait across individuals. Addressing this requirement, we presented a homeomorphism between 2D lattices and shapes that enables a robust vector space embedding of silhouettes. Sampling suitable lattice points allows to roughly track the movement of limbs without requiring any limb recognition strategy.

Combining shape representations derived from several frames into lager feature vectors provides temporal context for the classification task. Experimental results underline that gait classification using support vector machines yields satisfiable accuracy. Temporal filtering of the classification results in further improvements of the reliability of the presented framework, because it lessens the effect of sporadic misclassifications.

Currently, we have been working to demonstrate the applicability and reliability of our gait classification scheme in different, unconstrained real world settings, using a motion segmentation technique that automatically acquires background models for non-rigid scenes. In addition, we are working to adopt our approach to situations with several people in such a way that they will occlude each other. To this end, we are experimenting with robust tracking techniques, as introduced in [24], that allow us to extract and analyse several silhouettes simultaneously.

References

[1] M. Murray, Gait as a total pattern of movement, American Journal of Physical Medicine 46 (1) (1967) 290–332.
[2] J. Cutting, L. Kozlowski, Recognizing friends by their walk: Gait perception without familiarity cues, Bulletin of the Psychonomic Society 9 (5) (1977) 353–356.
[3] T. Moeslund, E. Granum, A survey of computer vision-based human motion capture, Computer Vision and Image Understanding 81 (3) (2000) 221–268.
[4] S. Sarkar, P. Phillips, Z. Liu, I. Vega, P. Grother, K. Bowyer, The human id gait challange problem: Data sets, performance, and analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (2) (2005) 162–177.
[5] S. Niyogi, E. Adelson, Analyzing and recognizing walking figures in xyt, in: Proc. CVPR, 1994, pp. 469–474.
[6] R. Cutler, L. Davis, Robust real-time periodic motion detection, analysis, and applications, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (8) (2000) 781–796.
[7] J. Little, J. Boyd, Recognizing people by their gait: The shape of motion, Videre 1 (2) (1998) 2–32.
[8] A. Bobick, A. Johnson, Gait recognition using static, activity-specific parameters, in: Proc. CVPR, vol. I, 2001, pp. 423–430.
[9] L. Lee, W. Grimson, Gait analysis for recognition and classification, in: Proc. Int. Conf. on Automatic Face and Gesture Recognition, 2002, pp. 155–161.
[10] G. Veres, L. Gordon, J. Carter, M. Nixon, What image information is important in silhouette-based gait recognition? in: Proc. CVPR, vol. II, 2004, pp. 776–782.
[11] R. Collins, R. Gross, J. Shi, Silhouette-based human identification from body shape and gait, in: Proc. Int. Conf. on Automatic Face and Gesture Recognition, 2002, pp. 351–356.
[12] D. Tolliver, R. Collins, Gait shape estimation for identification, in: Proc. Int. Conf. on Audio and Video-Based Biometric Person Authentication, vol. 2688 of LNCS, Springer, 2003, pp. 734–742.
[13] C. BenAbdelkader, R. Cutler, L. Davis, Gait recognition using image self-similarity, EURASIP Journal on Applied Signal Processing 4 (2004) 572–585.
[14] D. Kendall, Shape manifolds, procrustean metrics and complex projective spaces, Bulletin of the London Mathematical Society 16 (1) (1984) 81–121.
[15] J. Boyd, Synchronization of oscillations for machine perception of gaits, Computer Vision and Image Understanding 96 (1) (2004) 35–59.
[16] A. Veeraraghavan, A. Chowdhury, R. Chellappa, Role of shape and kinematics in human movement analysis, in: Proc. CVPR, vol. I, 2004, pp. 730–737.
[17] L. Wang, H. Ning, W. Hu, T. Tan, Gait recognition based on procrustes shape analysis, in: Proc. ICIP, vol. III, 2002, pp. 433–436.
[18] G. Johansson, Visual perception of biological motion and a model for its analysis, Perception and Psychophysics 14 (2) (1973) 201–211.
[19] C. Bauckhage, J. Tsotsos, Bounding box splitting for robust shape classification, in: Proc. ICIP, vol. II, 2005, pp. 478–481.
[20] V. Vapnik, Universal learning technology: support vector machines, NEC Journal of Advanced Technology 2 (2) (2005) 137–144.
[21] T. Joachims, Making large-scale svm learning practical, in: B. Schölkopf, C. Burges, A. Smola (Eds.), Advances in Kernel Methods – Support Vector Learning, MIT Press, Cambridge, MA, 1999, pp. 169–184.
[22] W. Kienzle, G. Bakir, M. Franz, B. Schölkopf, Face detection – efficient and rank deficient, in: Proc. NIPS, 2005, pp. 673–680.
[23] A. Elgammal, D. Harwood, L. Davis, Non-parametric model for background subtraction, in: Proc. ECCV, vol. 1842 of LNCS, Springer, 2000, pp. 751–767.
[24] K. Okuma, A. Taleghani, N. de Freitas, J. Little, D. Lowe, A boosted particle filter: Multitarget detection and tracking, in: Proc. ECCV, vol. 3021 of LNCS, Springer, 2004, pp. 28–39.

Please cite this article in press as: C. Bauckhage et al., Automatic detection of abnormal gait, Image Vis. Comput. (2007), doi:10.1016/j.imavis.2006.10.004