DESIGN OF A MORPHOLOGICAL MOVING OBJECT SIGNATURE AND APPLICATION TO HUMAN IDENTIFICATION

Olivier Barnich and Marc Van Droogenbroeck

University of Liège
Montefiore Institute, INTELSIG Group
Liège, Belgium

ABSTRACT

Many computer vision systems try to infer semantic information about a video scene content by looking at the time series of the silhouettes of the moving objects. This paper proposes a new inter-frame feature set (signature) based on piecewise surfacic descriptions of binary silhouettes. It captures the dynamics of moving objects and compacts it into a robust set of features suitable for classification. To assess its ability to represent motion information, we use it to build a complete gait recognition algorithm that we test on a database of 21 different subjects. To highlight the efficiency of our signature, we use *frontal* views instead of *side* views of persons, which is less discussed in literature and is considered to be harder as the movement of legs is not visible. In that context, the high recognition rates obtained (over 95% of correct identifications) proves that our signature is appropriate to describe moving objects.

Index Terms— Identification of persons, Surveillance, Pattern recognition, Signal analysis, Video signal processing

1. INTRODUCTION

Automatic human motion analysis is a very active research field in computer vision [1] and one of its most challenging goals is the automatic identification of humans. The first step of many motion analysis systems consists in the use of a background subtraction algorithm to extract time series of silhouettes of the moving objects that are in the field of view of the camera. If the silhouettes are tracked along multiple frames, semantic information about the scene content can be inferred by extracting and classifying a compact set of features, called *signature* hereafter, from the moving silhouettes of each object. Such a framework can be used for multiple purposes, like human detection [2] or lip reading [3] but its most famous application is automatic gait recognition. Since its introduction in [4], human identification through gait recognition has become popular. A major advantage of gait recognition towards other biometrics is that it can be achieved using

images taken from a large distance of the users. Gait recognition techniques might not be as precise as face recognition algorithms but, while most face recognition techniques need a constrained image of the person's face to function properly and thus require its cooperation, no user cooperation is necessary to perform gait recognition. All in all, because of its relative lack of precision, gait recognition is still best suited to reinforce a decision made in a multi-modal biometric system [5, 6].

Most gait recognition techniques use images captured from the side of the persons. The so-called "lateral views" are believed to be the most suitable views. But their use in real indoor conditions leads to serious operational issues. To capture long sequences of walking persons from their side, cameras have to be placed at a large distance. While it is manageable in outdoor environments, it is more difficult indoor. In hallways, it is often impossible to capture lateral view images. As a consequence, the use of a frontal view is mandatory in most hallways. Images captured from the front or the back of the users are thought to be less appropriate. They introduce geometric and scale transformations of the walkers that are absent from the side view. Furthermore, on these views, the pendular motion of the legs that is believed to contain a large amount of information about the gait is less visible [7]. However, the human ability to recognize walking people from behind supports the idea that building an automatic recognition scheme is possible.

Gait signatures can be extracted using many types of features: *e.g.* raw area of resized silhouettes [8], horizontal and vertical projections [9], raw contours [10], or their Fourier descriptors [11]. All these signatures have been used to recognize walkers from a side view. However in [12], Soriano *et al.* successfully recognized people from their gait using *frontal-view* cameras only. They used series of Freeman encoding of re-sized silhouette contours as a gait signature and obtained high recognition rates on a small database (4 persons). The use of the contour information is a source of noise in any practical environment where the contour of the extracted silhouettes are often of poor quality due to inaccuracies inherent to background subtraction algorithms. The use of a surfacic

---
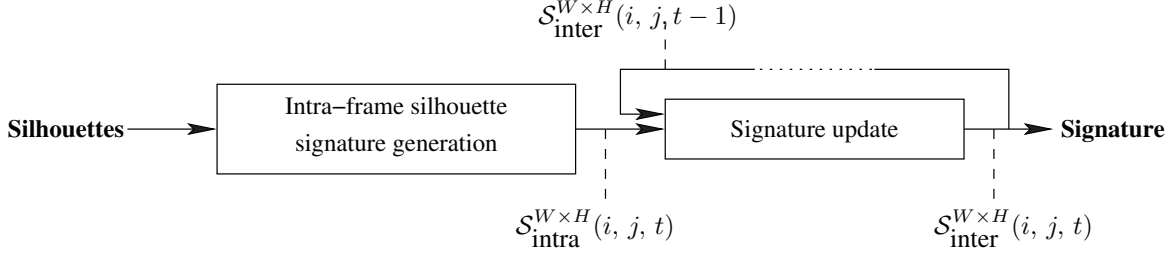
Fig. 1. Signature extraction from time series of silhouettes.

criterion is an option to increase the robustness. Furthermore, the database used in [12] was too small to definitively conclude that the used signature was relevant enough.

This paper describes a new moving object signature based on a morphological piecewise surfacic description of time series of silhouettes. To assess the relevance of our signature, we use it as a the key component of a complete frontal view gait recognition framework and test it on a database of 21 different persons.

The proposed new object signature is detailed in Section 2. In Section 3, we describe and evaluate our complete gait recognition algorithm. Section 4 concludes this paper.

## 2. DESIGN OF A MOVING OBJECT SIGNATURE

The inter-frame moving object signature is a temporal aggregation of silhouette signatures. The whole process, as shown on Figure 1, is described hereafter.

2.1. Intra-frame silhouette signature

Our intra-frame signature is based on a morphological operator called *cover by rectangle*s that provides a piecewise surfacic description of the shape of the silhouettes. We have used the cover by rectangles in [2] to extract information from single silhouettes. Intuitively, the cover by rectangles $C(S)$ of a silhouette S is the set of all the largest non-redundant rectangles that can be wedged inside of it (see Figure 2). Since the rectangles of $C(S)$ overlap each other, the description provided by $C(S)$ is redundant (*i.e.* robust). On the other hand, each rectangle of $C(S)$ covers at least one pixel that is not covered by any other rectangle of $C(S)$. Consequently, the use of $C(S)$ ensures a robust description of the shape S with a limited amount of redundancy.

Since the number of rectangles contained in $C(S)$ can be large (more than a thousand), it is impractical to use all the rectangles directly as a set of features. We believe that most of the information resides in the rectangle sizes distribution densities. These densities can be estimated using a discrete histogram whose bins correspond to the ratios of rectangles that fall within given size intervals.

More precisely, we denote by α the number of rectangles of $C(S)$ (so $\alpha = \sharp(C(S))$) and index the rectangles of $C(S)$ with
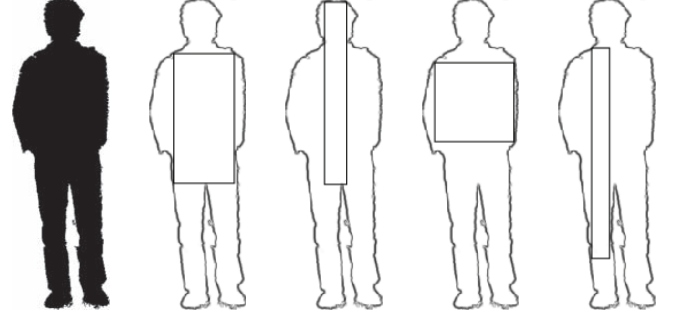
Fig. 2. The cover by rectangles of a silhouette is the union of all the largest rectangles that can be wedged inside of it [2].

a parameter d so that $R_d (d = 1, \ldots, \alpha)$ are the rectangles of $C(S)$. We also denote by w^{max} (resp. h^{max}) the width (resp. height) of the widest (resp. tallest) rectangle of $C(S)$. Next, we build size distribution histograms and partition the widths and the heights of the rectangles respectively into M bins $B^W(i)$ and N bins $B^H(j)$

$$B^W(i) = \left] i\frac{w^{max}}{M}, \ (i+1)\frac{w^{max}}{M} \right] \tag{1}$$

$$B^H(j) = \left] j\frac{h^{max}}{N}, \ (j+1)\frac{h^{max}}{N} \right] \tag{2}$$

where $i = 0, \ldots, M - 1$ and $j = 0, \ldots, N - 1$. Finally, the normalized two-dimensional histogram $\mathcal{S}_{intra}^{W \times H}(i, j)$ is defined as

$$\mathcal{S}_{intra}^{W \times H}(i, j) = \frac{1}{\alpha} \sharp \left\{ R_d | w_d \in B^W(i), \ h_d \in B^H(j) \right\} \tag{3}$$

where w_d and h_d denote the width and the height of the rectangle R_d respectively.

$\mathcal{S}_{intra}^{W \times H}(i, j)$ constitutes the set of features that we use as a signature to characterize the information contained in a single silhouette S.

2.2. Inter-frame signature update

To capture the dynamics of a moving object, we need an inter-frame signature that handles time series of silhouettes.

854
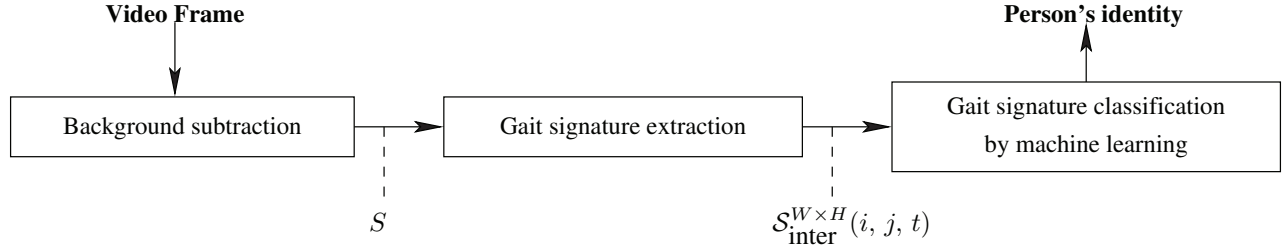
Fig. 3. Architecture of the gait recognition algorithm.

Let t be the time of the current frame and $\mathcal{S}_{\text{intra}}^{W \times H}(i, j, t)$ be the signature of S at time t. A new inter-frame signature $\mathcal{S}_{\text{inter}}^{W \times H}$ is defined as the concatenation of L consecutive histograms:

$$
\mathcal{S}_{\text{inter}}^{W \times H}(i, j, t) = \left\{ \mathcal{S}_{\text{intra}}^{W \times H}(i, j, t - (L-1)), \ldots, \right.
$$
$$
\left. \mathcal{S}_{\text{intra}}^{W \times H}(i, j, t-1), \mathcal{S}_{\text{intra}}^{W \times H}(i, j, t) \right\}
$$
$$(4)$$

The signature $\mathcal{S}_{\text{inter}}^{W \times H}$ is updated frame by frame with the silhouette histogram corresponding to the silhouette of the current frame (see Figure 1). Note that this signature characterizes the object dynamics of the L last frames.

In the following Section, we show how powerful $\mathcal{S}_{\text{inter}}^{W \times H}(i, j, t)$ is for frontal view gait recognition.

3. APPLICATION TO GAIT RECOGNITION

The purpose of the gait recognition algorithm described in this Section is to infer the identity of the walking person. For this application, we need to determine the appropriate values for the parameters (M, N, and L) of $\mathcal{S}_{\text{inter}}^{W \times H}(i, j, t)$ to capture and classify the gait dynamics of a walking human.

3.1. Gait recognition algorithm

The architecture of the gait recognition algorithm is drawn in Figure 3.

We extract the time series of binary silhouettes from the video stream using a background subtraction algorithm called *Extended Gaussian Mixture Model,* as proposed in [13]; this algorithm is said to be adequate to deal with noisy sensors and changing illumination. Then the signature of the object in the current frame is computed and used to update the inter-frame signature.

Finally $\mathcal{S}_{\text{inter}}^{W \times H}(i, j, t)$ is used as the input of a machine learning algorithm called extra-trees [14] whose purpose is to classify the signature and to provide the identity of the person. The extra-trees consist in a forest of independent decision trees. They are fast, accurate, and able to deal with the high dimensionality of the feature space we use.

For consistency over time, we refine the output and use a majority vote policy over the last V frames to smooth the re-

sults since the extra-trees might output a different class label for each new frame.

3.2. Results

We now discuss the performances of the recognition algorithm. It is worth noting that our signature is computationally efficient enough to be used in real-time applications: the whole recognition system runs real-time for standard resolution images (640×480 pixels) on a conventional computer.

During preliminary tests, we have determined appropriate values for the parameters of the method : the number of bins M and N, the number of frames L aggregated in a single gait signature, and the length V of the sliding temporal window used for the majority vote.

Depending on the sizes of the silhouettes and on the size of the training dataset, M and N should be comprised between 10 and 20. The use of values closer to 20 generally leads to better recognition rates. However, large numbers of large silhouettes (therefore with many rectangles) are needed to populate a 20×20 bins histogram with appropriate statistical significance.

The parameter L is adapted according to the framerate of the cameras in order for the gait signature to contain at least a complete gait cycle. Incorporating more gait cycles increases the precision of the classification but also lowers the reactivity of the system (note that the first estimate of the identity of the walking person is given by the system after a delay of $L + V$ frames).

A similar reasoning applies to V (the number of frames used for the majority vote). An increase of its value enhances the stability of the results but also increases the recognition system latency.

To assess the relevance of our gait signature and the performances of the whole system, we have performed tests on a database composed of series of 4 walking sequences of 21 different persons. This database contains frontal-view images exclusively and it was built using a consumer market webcam.

Identification rates are shown on Figure 4. For high values of V, the ratio of correct identifications can reach values as high as 97%. We also see on Figure 4 that the choice of $M = N = 15$ leads to better results that $M = N = 20$. The small sizes of some silhouettes of the dataset might explain
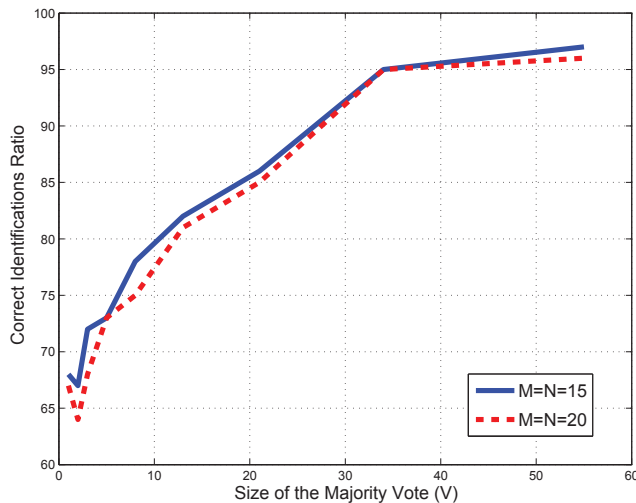
Fig. 4. Performances on a database of 21 persons.

this. When a person is far from the camera, his silhouettes are too small to contain enough rectangles to correctly populate a 20×20 bins histogram.

The high identification rates (over 95%) prove that our intra-frame gait signature succeeds in capturing relevant information about the gait of a person.

4. CONCLUSIONS

This paper presents an innovative signature to characterize time-series of binary silhouettes of a moving object. It is based on a morphological operator called "cover by rectangles", and it consists in a temporal aggregation of piecewise surfacic descriptions of silhouettes.

To evaluate the performances of our signature, we have chosen to incorporate it in a gait recognition strategy, where it is used to represent a moving object. During the evaluation, it has appeared that, despite the general belief that lateral views are better suited for gait recognition, it is possible to build a system capable to work with front views. Moreover several tests show that, with the help of our signature, the recognition rates are close to 95% of correct identifications on a database of 21 different humans.

5. REFERENCES

[1] J. Aggarwal and Q. Cai, "Human motion analysis: A review," *Computer Vision and Image Understanding*, vol. 73, pp. 90–102, 1999.

[2] O. Barnich, S. Jodogne, and M. Van Droogenbroeck, *Robust analysis of silhouettes by morphological size distributions*, vol. 4179 of *Lecture Notes on Computer Science*, pp. 734–745, Springer Verlag, 2006.

[3] H. Murase and R. Sakai, "Moving object recognition in eigenspace representation: gait analysis and lip reading," *Pattern Recogn. Lett.*, vol. 17, no. 2, pp. 155–162, 1996.

[4] S. Niyogi and E. Adelson, "Analyzing and recognizing walking figures in XYT," *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 469–474, June 1994.

[5] M. Nixon, T. Tan, and R. Chellappa, *Human identification based on gait*, Springer, 2006.

[6] X. Zhou and B. Bhanu, "Feature fusion of side face and gait for video-based human identification," *Pattern Recognition*, vol. 41, no. 3, pp. 778–795, 2008.

[7] C. Yam, M. Nixon, and J. Carter, "Automated person recognition by walking and running via model-based approaches," *Pattern Recognition Letters*, vol. 37, no. 5, pp. 1057–1072, 2004.

[8] J. Foster, M. Nixon, and A. Prügel-Bennett, "Automatic gait recognition using area-based metrics," *Pattern Recognition Letters*, vol. 24, no. 14, pp. 2489–2497, 2003.

[9] A. Kale, N. Cuntoor, B. Yegnanarayana, A. Rajagopalan, and R. Chellappa, "Gait analysis for human identification," in *Proceedings of the International Conference on Audio-and Video-Based Person Authentication*, Guildford, UK, 2003, pp. 706–714.

[10] L. Wang, T. Tan, H. Ning, and W. Hu, "Silhouette analysis-based gait recognition for human identification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1505–1518, December 2003.

[11] S. Mowbray and M. Nixon, "Automatic gait recognition via fourier descriptors of deformable objects," in *Audio Visual Biometric Person Authentication*, J. Kittler and M. Nixon, Eds. 2003, pp. 566–573, Springer.

[12] M. Soriano, A. Araullo, and C. Saloma, "Curve spreads: a biometric from front-view gait video," *Pattern Recognition Letters*, vol. 25, no. 14, pp. 1595–1602, 2004.

[13] Z. Zivkovic, "Improved adaptive gausian mixture model for background subtraction," in *Proceedings of the 17th International Conference on Pattern Recognition*, 2004, vol. 2, pp. 28–31.

[14] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine Learning*, vol. 36, no. 1, pp. 3–42, 2006.