Topological Features for Monitoring Human Activities at Distance

Javier Lamar Leon[1][✉], Raúl Alonso[1], Edel Garcia Reyes[1],
and Rocio Gonzalez Diaz[2]

[1] Patterns Recognition Department, Advanced Technologies Application Center,
7th A # 21406 e/ 214 y 216, Rpto. Siboney, CP 12200 Playa, La Habana, Cuba
{jlamar,rbaryolo,egarcia}@cenatav.co.cu
[2] Applied Math Department, School of Computer Engineering,
Campus Reina Mercedes, University of Seville, Seville, Spain
rogodi@us.es

Abstract. In this paper, a topological approach for monitoring human activities is presented. This approach makes possible to protect the person's privacy hiding details that are not essential for processing a security alarm. First, a stack of human silhouettes, extracted by background subtraction and thresholding, are glued through their gravity centers, forming a 3D digital binary image I. Secondly, different orders of the simplices are applied on a simplicial complex obtained from I, which capture relations among the parts of the human body when walking. Finally, a *topological signature* is extracted from the persistence diagrams according to each order. The measure cosine is used to give a similarity value between topological signatures. In this way, the powerful topological tool known as persistent homology is novelty adapted to deal with gender classification, person identification, carrying bag detection and simple action recognition. Four experiments show the strength of the topological feature used; three of they use the CASIA-B database, and the fourth use the KTH database to present the results in the case of simple actions recognition. In the first experiment the named topological signature is evaluated, obtaining 98.8 % (lateral view) of correct classification rates for gender identification. In the second one are shown results for person identification, obtaining an average of 98.5 %. In the third one the result obtained is 93.8 % for carrying bag detection. And in the last experiment the results were 97.7 % walking and 97.5 % running, which were the actions took from the KTH database.

Keywords: Gait-based recognition · Topology · Persistent homology · Gender classification · Carrying bag detection · Action recognition

1 Introduction

Objects detected by a video surveillance system are usually classified into different categories: human, vehicle, animal, etc. In the case of persons, it is useful another

level of categorization, which gives clues for the interpretation. After finding the class to which an object belongs, one may try to identify it and interpret its individual behavior in the scene, as well as its interaction with other objects. We consider that nonrigid objects and its actions should be described by the dynamic spatial relations among its different parts.

Methods based on geometric features extracted from silhouettes or its contours [1] have been widely used for gait recognition tasks. However, the stability of such features is affected by deformations in the shape of the silhouette. Even for the same individual, little changes on the walking direction, illumination variations and the way the clothes fit to the human body, may cause variability on the geometric features. We conjecture that topological descriptions based on the persistence of homology classes are more invariant to changes and noise in the silhouette shape than classical approaches. This kind of features have been previously used to match nonrigid shapes [2,3], because they are invariant under continuous deformations of the object.

Homology is a topological invariant frequently used in practice [4,5]. The ranks of the homology groups, also called Betti numbers, coincide in the first three dimensions with the number of connected components, tunnels and cavities of the object respectively, this as a consequence of the *Alexander duality* [6]. In particular, the homology could be a robust representation, because the shape of connected components and holes may change under geometric transformations, but their amount will be more stable. Given that is not enough to reach the invariance for the representation, but also needed a set of discriminating features, the approach called homological persistence, which is introduced from now on, will be used in order to elevate the discriminating power of the representation.

A *k-simplex* σ in $\mathbf{R^d}$ is the convex hull of a set S of $k+1$ affinely independent points, where $0 \leq k \leq d$. The dimension of σ is $dim(\sigma) = |S|-1 = k$. In (Fig. 1a) from left to right are shown k-simplices of dimensions 0, 1 and 2, which are the only used in this work. For every $U \subseteq S$ the simplex σ' defined by U is said to be a *face* of σ, if $U \neq S$ then σ' is a *proper face* of σ. A simplex σ is a *facet* or *coface* of a simplex σ' when σ' is a face of σ.

Let K be a collection of simplices, K is a *simplicial complex* if it satisfies two properties, namely (i) if σ' is a face of σ and $\sigma \in K$ then $\sigma' \in K$, and (ii) if
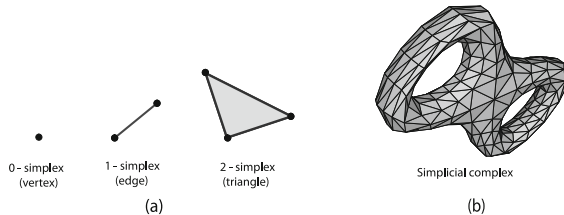

Fig. 1. (a) Simplices of dimensions 0, 1 and 2 (left to right), (b) Simplicial complex using 2-simplices (triangles) as maximal dimension simplices.

$\sigma_1, \sigma_2 \in K$ then $\sigma_1 \cap \sigma_2$ is empty or a face of both (Fig. 1b). The *dimension* of K is the largest dimension of any of its simplices. It is important to point out that a subset L of K is a *subcomplex* of K if it is a complex itself, i.e., L is a subcomplex of K iff it satisfy property (i), note that L inherit (ii) from K.

Let K be a simplicial complex with n simplices, a *filter function* $f : K \rightarrow \mathbf{R}$ is a function that assigns a real number to each simplex in K. A *filter* of K is an ordering of its simplices $[\sigma_1, \sigma_2, ..., \sigma_n]$, which satisfy that each prefix $K_i = [\sigma_1, \sigma_2, ..., \sigma_i]$ is a subcomplex of the next prefix K_{i+1}. Persistent homology concerns, given a filter, with how long persist homology classes (connected components, tunnels, cavities, etc.) after they are born; note that looking to a filter as a growing simplicial complex, we may see that homology classes are born and die. The homology classes that persist throughout all the filter i.e., that are born and never die, are the homology classes of K. An algorithm to compute persistent homology takes as input a filter, and gives out a collection of pairs representing the birth and death time of the homology classes. The difference between the birth and death time of a homology class is called its *persistence*.

The persistence diagrams and barcodes are two ways of representing the collection of pairs given up by the algorithm, they are used to study and visualize the persistent homology. In (Fig. 2 [7]) is shown an example of barcode representation, where the start of a horizontal bar represents the birth of a homology class and the end its death. An easy handling matlab implementation of an algorithm for computing persistent homology can be found[1].
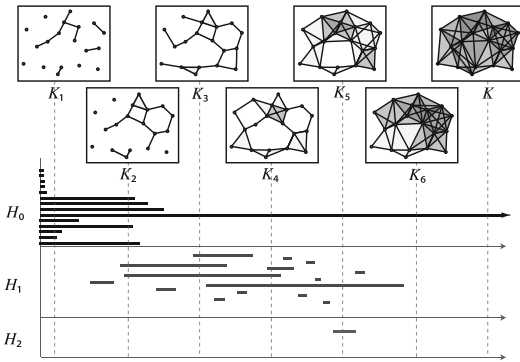

Fig. 2. Persistent barcodes. K_5 betti numbers are $H_0 = 1, H_1 = 1, H_2 = 1$ and K_2 betti numbers are $H_6 = 1, H_2 = 1, H_0 = 1$. Took from [7].

In this paper, a version of the topological features based on homological persistence given in [8,9] is presented, which is valid for gender classification, gait based person identification, carrying bag detection and simple action recognition. As aforementioned, monitoring these human activities makes possible to protect the person's privacy hiding details, such as faces, that are not essential for

---

[1] http://comptop.stanford.edu/programs/plex-2.0.1-windows.zip

processing a security alarm. The topological features can be used for monitoring human activities even when is possible to use images of high resolution.

The rest of the paper is organized as follows. Section 2 is devoted to describe the method for obtaining the topological signature. Experimental results are then reported in Sect. 3. We conclude this paper and discuss some future work in Sect. 4.

2 Topological Signature for Activities Monitoring

In this section, a topological signature is presented that is used for gender classification, person identification, carrying bag detection and simple action recognition at distance. A traditional approach for constructing a simplicial complex, departs from a points cloud, for which it is necessary to recover topological relations among its points. In our case there are structural relations among the points (pixels), and in fact there are temporal relation too, given that we work with videos. We take advantage of this relations to construct the simplicial complex.

2.1 The Simplicial Complex $\partial K(I)$

First, the foreground (person) is segmented from the background by applying background subtraction and thresholding. The sequence of resulting silhouettes is analyzed to extract one *subsequence of representation*, which includes at least a gait cycle [10].

The 3D digital binary image $I = (\mathbb{Z}^3, B)$ (where $B \subset \mathbb{Z}^3$ is the foreground), is built by stacking silhouettes of a subsequence of representation, aligned by their gravity centers (gc), see (Fig. 3a) and (Fig. 3b). The 3D cubical complex $Q(I)$ associated to I contains the unit cubes with vertices $V = \{(i,j,k), (i+1,j,k), (i,j+1,k), (i,j,k+1), (i+1,j+1,k), (i+1,j,k+1), (i,j+1, k+1), (i+1,j+1,k+1)\}$ and all its faces (vertices, edges and squares) iff $V \subseteq B$.
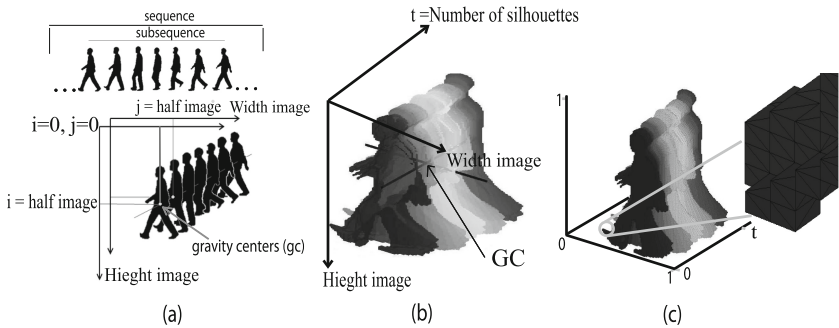

Fig. 3. (a) Silhouettes aligned by their gravity centers. (b) $I = (\mathbb{Z}^3, B)$ obtained from the silhouettes (GC is the gravity center of I). (c) The border simplicial complex $\partial K(I)$.

The squares of $Q(I)$ that are faces of exactly one cube in $Q(I)$ are subdivided in two triangles. The obtained triangles and their faces (vertices and edges) make up the *border simplicial complex* $\partial K(I)$ (see Fig. 3c). Finally, coordinates of the vertices of $\partial K(I)$ are *normalized* to coordinates (x, y, t), where $0 \leq x, y \leq 1$ and t is the number of silhouette of the subsequence of representation.

2.2   Filters for $\partial K(I)$

The topology of $\partial K(I)$ is, in general, very poor. However, in this and the next subsections we present how, using persistence diagrams, it is possible to get a topological signature from $\partial K(I)$ that captures relations among the parts of the human body when walking, and is robust against small input-data perturbations.

Up to medical researches [11,12], the natural human gait is defined as a succession of rhythmic and alternate movements of the limbs and the torso. Therefore, the difference in the gait is given by the relative position of the limbs and torso in each moment (structural features). A growing scheme of the simplicial complex $\partial K(I)$ respect to a useful selected filtration can encode the relations among the parts of the human body, which can be therefore very discriminating.

When a view direction d is chosen, two filters for $\partial K(I)$ are obtained as follows. All simplices belonging to $\partial K(I)$ are associated with two filter functions f_+ and f_-. For each vertex $v \in \partial K(I)$, $f_+(v)$ is the distance between v and the plane normal to d passing through the origin of the reference frame, while $f_-(v) = -f_+(v)$. Edges and triangles are associated to the smallest value that f_+ (resp. f_-) assumes on their vertices. Being the simplices of $\partial K(I)$ finite in number, we can determine a minimum value for f_+, say $f_{\min}$, and a maximum one, $f_{\max}$. It is now possible to induce two filters on $\partial K(I)$ by ordering its simplices according to increasing values of f_+ and f_- respectively, or according to increasing dimension of the simplices in case of tie, or arbitrarily otherwise. Denote these filters by $K_{[f_{\min}, f_{\max}]} = [\sigma_1, \ldots, \sigma_k]$ and $K_{[-f_{\max}, -f_{\min}]} = [\sigma'_1, \ldots, \sigma'_k]$.

2.3   Persistence Diagrams and Topological Signatures

Given a simplicial complex K, a filter function f, and the corresponding filter $[\sigma_1, \ldots, \sigma_k]$ for K, if σ_i completes a $p-$cycle (p is the dimension of σ_i) when σ_i is added to $K_{i-1} = [\sigma_1, \ldots, \sigma_{i-1}]$ then a p-homology class γ *is born at time i*, otherwise, a $(p-1)$-homology class *dies at time i*. The difference between the birth and death time of a homology class is known as its persistence, which quantifies the significance of a topological attribute. If γ never dies, we set its persistence to infinity. Drawing an horizontal segment $[i, j)$, in a 2D plane, for a p-homology class that is born at time i and dies at time j, we get the p-barcode diagram of the filtration. It represents a p-homology class by a segment whose length is the persistence of that class.

In this paper, barcodes are first computed for $K_{[f_{\min}, f_{\max}]}$ and $K_{[-f_{\max}, -f_{\min}]}$. Then, the barcodes are explored according to a uniform sampling. More precisely, given an integer $n > 0$, $n - 1$ cuts are performed homogeneously in the complex $K_{[f_{\min}, f_{\max}]}$ (resp. $K_{[-f_{\max}, -f_{\min}]}$) as follows: Let's suppose $K_{[f_{\min}, f_{\max}]} = \{\sigma_0, \ldots, \sigma_m\}$

and let's take $P_i = [\sigma_{\lfloor \frac{(i-1)m}{n} \rfloor + 1}, \ldots, \sigma_{\lfloor \frac{im}{n} \rfloor - 1}]$, $1 \leq i \leq n$, as the partitions given by the cuts. For a fixed i, we compute:

(a) Number of homology classes that were born or persist when the simplex $\sigma_{\lfloor \frac{(i-1)m}{n} \rfloor}$ is added, and, persist or die when the simplex $\sigma_{\lfloor \frac{im}{n} \rfloor}$ is added.
(b) Number of homology classes that were born in P_i.

An analogous process is done for $K_{[-f_{\max}, -f_{\min}]}$. A vector of $2n$ entries is then formed containing (a) in entry $2i$ and (b) in $2i + 1$; this way we obtain for a given dimension p, a vector for each filter of K.

The *topological signature for a gait subsequence considering a fixed direction of view* consists in four $2n$-dimensional vectors: (V_1, V_2, V_3, V_4) constructed as explained above. Consider that we take into account two dimensions ($p = 0$ and $p = 1$) and two filters ($K_{[f_{\min}, f_{\max}]}$ and $K_{[-f_{\max}, -f_{\min}]}$).
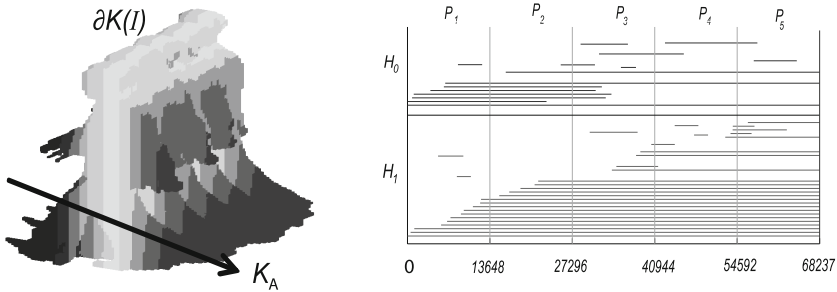

Fig. 4. An example of computation of the first element of a topological signature.

For example, consider $\partial K(I)$ given in (Fig. 4) and the direction of view K_A. We perform 4 uniform cuts on $\partial K(I)_{[f_{\min}, f_{\max}]}$, see the green lines in the persistence barcode representation. Let's fix $i = 2$ and dimension $d = 0$, according to Fig. 4, the number of homology classes that persist or were born in σ_{13648}, and, persist or die in σ_{27296} are $H_0 = 6$ in dimension 0 and the number of the homology classes that were born in P_2 are $H_0 = 2$ in dimension 0.

2.4   Comparing Topological Signatures

The topological signatures for two gait subsequences associated with a fixed view direction, say $V = \{V_1, \ldots, V_4\}$ and $W = \{W_1, \ldots, W_4\}$, can be compared according to the following procedure: for every $i = \{1, \ldots, 4\}$ compute:

$$S_i = \frac{V_i \cdot W_i}{\|V_i\| \cdot \|W_i\|}. \tag{1}$$

which is the cosine of the angle between the vectors V_i and W_i. Observe that $0 \leq S_i \leq 1$ since the entries of both vectors are always non-negative. Then, the

total similarity value for two gait subsequences, O_1 and O_2, considering a fixed view direction, is the sum of the 4 similarity measures computed before:

$$S(O_1, O_2) = S_1 + S_2 + S_3 + S_4. \tag{2}$$

3   Experimental Results

3.1   Human Gender Classification

Human gender classification can be obtained based on face [13], voice [14] or gait [15,16]. Dynamic features when the person walks give the possibility to classify gender at a distance, without any interaction with the subject [10,17,18]. This fact can improve the performance of intelligent surveillance systems and it can reduce the false positive rate during re-identification of an individual on a wide network camera. People not only observe the global motion properties while human walks, but they detect motion patterns of local body parts. For instance, women tent to swing their hips more than their shoulders. On the contrary, men tent to swing their shoulders more than their hips [19]. Moreover, men have in general wider shoulders than women [20].

Experiment 1. The performance of the proposed method is evaluated using the lateral view (90 degrees respect to the camera) in CASIA-B database. This database is composed of 124 subjects, 92 men and 32 women. For each person in the database there are 6 walking sequences, each one provided with background subtraction.

In order to avoid bias we selected 25 men and 25 women to perform the experiment. The 50 subjects were divided in 25 disjoint sets, each one containing two subjects (one man and one woman). Only one of these 25 sets was used for testing. The remaining 24 sets were used for training. The correct classification rate (CCR) is the average of the 25 combination of the cross validation.

The experimental protocol was made according to [15,16]. In this experiment, a subsequence of representation corresponds to the whole sequence, which has two gait cycles as average. We fixed $n = 24$ and used 3 view directions. The first one is vertical (i.e. parallel to axis y). The second one forms 45 degrees with axes x and y and 90 degrees with axis t. The third one is parallel to axis t, (see Fig. 5). In each experiment, the results of our method are compared with the methods presented in [15,16].

The aim of this experiment is to evaluate the topological signature for gender classification. Table 1 shows the 25-fold-cross-validation of CCR for the whole body using lateral view, as aforementioned. We can see that the topological signature provides better results.

3.2   Gait Recognition

Gait recognition is a challenging problem that gives the possibility to identify persons at a distance without any interaction with the subjects, which is very important in real surveillance scenarios [10,18].
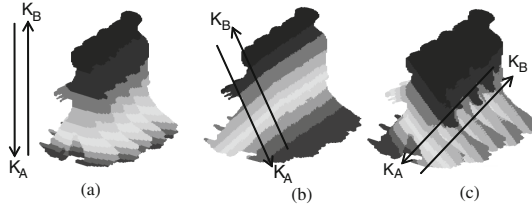
Fig. 5. View directions used in the experiments.

Table 1. Correct classification rates for gait based gender classification.

Method	Avg computer [15]	Avg human observers [15]	MCRF [16]	Our method
CCR	95.97	95.45	98.3	98.88

Experiment 2. We show the performance of the proposed method on the 11 views of the CASIA-B database, which contains 124 subjects. In this experiment the subsequence of representation consists of all the sequence, and the n parameter was set to 24.

We have used 4 view directions, they are shown in (Fig. 6) by black segments going through the simplicial complex and intercepting the GC point. The first one is parallel to the y axis and perpendicular to x and t, (Fig. 6a). The second one is parallel to the x axis and perpendicular to y and t, (Fig. 6b). The third one forms 45 degrees with the x and y axes and is perpendicular to t, (Fig. 6c). And the last is perpendicular to the previous one, perpendicular to the t axis and as aforementioned goes through the GC point. In all the cases in (Fig. 6) is shown the simplicial complex interleaved between two planes orthogonal to the view direction.

The experiment was carried out using 4 video sequences for training, and 2 for testing. The results are compared with the ones in [1]. In Table 2 is shown the cross validation average (15 combinations) of correct classification rates. It can be seen in Table 2 that the topological approach has a better performance for almost all the view angles, but the algorithm developed in [1] performs better for view angles close to 0 degrees. Therefore, these complementary behaviors conduce us to think in a combination of both approaches in future works.

3.3   Carrying Bag Detection

When the goal is to detect a simple behavior, namely a person has left a bag somewhere, it may be important to know that the person is carrying a bag. As carrying some object changes the normal body movements, topological features could be used to differentiate if the person left the bag [21–23].

Experiment 3. In this experiment we used, once again, the CAISA-B database. From the 124 persons, we selected 100 persons carrying bag and 100 walking
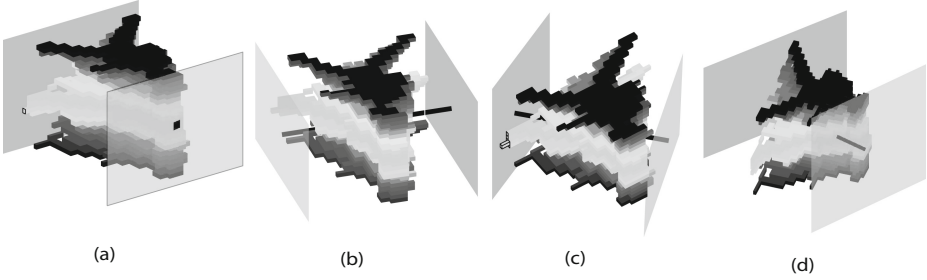
(a)   (b)   (c)   (d)

Fig. 6. View directions used in the experiments.

Table 2. Correct classification rates for gait based recognition.

Method	0	18	36	54	72	90	108	126	144	162	180	**Avg**
Wavelet(FD) [1]	100	100	100	93.4	81.1	90.3	90.3	83.3	91.9	92.7	97.6	92.9
Our method	99.3	99.1	98.8	98.3	97.6	98.0	98.3	98.3	98.2	98.2	99.0	98.5

normally. All the sequence was selected as the subsequence of representation, and the n parameter was set to 24.

Only lateral view was used in this case. For each person walking normally the 6 sequences provided by the database were used, while for parsons carrying bag the 2 sequences provided by the database were used. The same 4 view directions used for gait based recognition were used in this case.

It is important to point out that different kinds of bags carried by the persons, as well as the variation of position where they carry those bags, makes harder the classification, in (Fig. 7) are shown some details. Another hard situation emerges when the body occludes the bag, in this case even humans show difficulties to detect the bag.

Fig. 7. CAISA-B carrying bag images.

In our experimentation protocol, the 200 subjects to analyze were divided in 100 disjoint sets, each containing a person carrying a bag and a person without bag. One of these sets was used for testing, while the remaining 99 were used for training. The average of the 100 results obtained from the cross validation gives

up a correct classification rate of 92.5 % in the case of persons carrying bags, and 95.1 % in the case of persons walking in natural conditions.

3.4 Simple Action Recognition

In the case of simple action recognition, two action were taken from the KTH database, namely the action to run and the action to walk. Usually, to descry a person running gives the idea of the occurrence of an unusual event, which argues the importance of detecting this action in real surveillance scenarios.

Experiment 4. In order to test the performance of the proposed method, 20 persons were selected from the 25 persons in KTH database, 10 of them running and 10 walking. The subsequence of representation consists of all the sequence, and only lateral view was used.

The same four view directions used for gait recognition were selected in this case, while the n parameter took 24 as value. For each person four video sequences were used to carry out the experiment. The videos were filmed in an outdoor scenario as shown in (Fig. 8).
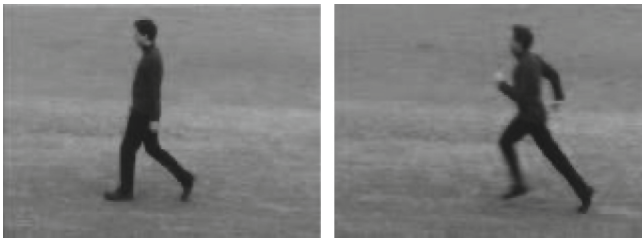

Fig. 8. Persons running and walking taken from the KTH database.

The 20 subjects selected were partitioned in 10 disjoint sets, each containing 2 subjects, one running and one walking. Only one of these 10 sets was used for testing, and the remaining 9 sets were used for training. The correct classification rate is the average of the 10 results obtained from the cross validation. In the case of persons running the result was 97.7 %, while for persons walking the result was 97.5 %.

4 Conclusions and Future Work

This paper shows that is possible to use the same topological feature, previously used for gait based human identification at a distance, in other tasks concerning to activities monitoring, including gender classification, carrying bag detection and simple action recognition. This kind of features have showed to be robust and discriminant. There are many issues to take into account in future works

in order to improve the promising results obtained in this work. Also, in future works, it is possible to model the crowd and its behavior as a whole and interpret the movements of the different parts characterizing the dynamic of holes and connected components.

References

1. Chen, C.H., Liang, J.M., Zhao, H., Hu, H.H., Tian, J.: Frame difference energy image for gait recognition with incomplete silhouettes. PRL **30**(11), 977–984 (2009)
2. Hilaga, M., Shinagawa, Y., Kohmura, T., Kunii, T.L.: Topology matching for fully automatic similarity estimation of 3d shapes. In: Proceedings of the 28th Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '01, pp. 203–212 (2001)
3. Bronstein, A.M., Bronstein, M.M., Kimmel, R.: Topology invariant similarity of nonrigid shapes. Int. J. Comput. Vis. **81**(3), 281–301 (2009)
4. Zomorodian, A., Carlsson, G.: Localized homology. Comput. Geom. **41**(3), 126–148 (2008)
5. Zomorodian, A.: Computational topology. In: Atallah, M., Blanton, M. (eds.) Algorithms and Theory of Computation Handbook, vol. 2, 2nd edn. Chapman & Hall/CRC Press, Boca Raton (2010)
6. Zomorodian, A.: Topology for Computing. Cambridge University Press, New York (2009)
7. Ghrist, R.: Barcodes, the persistent topology of data. BAMS. Bull. Am. Math. Soc. **45**, 61–75 (2008)
8. Lamar-León, J., García-Reyes, E.B., González-Díaz, R.: Human gait identification using persistent homology. In: Álvarez, L., Mejail, M., Gómez, L., Jacobo, J. (eds.) CIARP 2012. LNCS, vol. 7441, pp. 244–251. Springer, Heidelberg (2012)
9. Lamar, J., Garcia, E., Gonzalez-Diaz, R., Alonso, R.: An application for gait recognition using persistent homology. Electron. J. Image-A **3**(5) (2013)
10. Nixon, M.S., Carter, J.N.: Automatic recognition by gait. Proc. IEEE **94**(11), 2013–2024 (2006)
11. Murray, M.P.: Gait as a total pattern of movement: including a bibliography on gait. Am. J. Phys. Med. Rehabil. **46**(1), 290–333 (1967)
12. Winter, D.A.: Biomechanics and Motor Control of Human Gait: Normal, Elderly and Pathological. University of Waterloo Press, Waterloo (1991)
13. Golomb, B.A., Lawrence, D.T., Sejnowksi, T.J.: SEXNET: a neural network identifies sex from human faces. In: Lippmann, R.P., Moody, J.E., Touretzky, D.S. (eds.) Advances in Neural Information Processing Systems, vol. 3, pp. 572–579. Morgan Kaufmann Publishers Inc., San Mateo (1991)
14. Harb, H., Chen, L.: Gender identification using a general audio classifier. In: Proceedings of the 2003 International Conference on Multimedia and Expo, ICME '03, vol. 1, pp. 733–736. IEEE (2003)
15. Yu, S., Tan, T., Huang, K., Jia, K., Wu, X.: A study on gait-based gender classification. IEEE Trans. Image Process. **18**(8), 1905–1910 (2009)
16. Hu, M., Wang, Y., Zhang, Z., Zhang, D.: Gait-based gender classification using mixed conditional random field. IEEE Trans. Syst. Man Cybern. Part B **41**(5), 1429–1439 (2011)
17. Kale, A., Sundaresan, A., Rajagopalan, A.N., Cuntoor, N.P., Chowdhury, A.K.R., Kruger, V., Chellappa, R.: Identification of humans using gait. IEEE Trans. Image Process. **13**(9), 1163–1173 (2004)

18. Goffredo, M., Carter, J.N., Nixon, M.S.: Front-view gait recognition. In: Biometrics: Theory, Applications, and Systems, pp. 1–6, 29 September–1 October (2008)
19. Aggarwal, J.K., Cai, Q.: Human motion analysis: a review. Comput. Vis. Image Underst. **73**(3), 428–440 (1999)
20. Mather, G., Murdoch, L.: Gender discrimination in biological motion displays based on dynamic cues. Proc. Biol. Sci. **258**(1353), 273–279 (1994)
21. Dondera, R., Morariu, V.I., Davis, L.S.: Learning to detect carried objects with minimal supervision. In: CVPR Workshops, pp. 759–766. IEEE (2013)
22. Senst, T., Evangelio, R.H., Eiselein, V., Pätzold, M., Sikora, T.: Towards detecting people carrying objects - a periodicity dependency pattern approach. In: Richard, P., Braz, J. (eds.) VISAPP (2), pp. 524–529. INSTICC Press (2010)
23. BenAbdelkader, C., Davis, L.: Detection of people carrying objects: A motion-based recognition approach. In: Proceedings of Fifth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 378–383. IEEE (2002)