Persistent homology to analyse 3D faces and assess body weight gain

**D. Giorgi · M. A. Pascali ·
P. Henriquez · B. J. Matuszewski ·
S. Colantonio · O. Salvetti**

Abstract In this paper we analyse patterns in face shape variation due to weight gain. We propose the use of persistent homology descriptors to get geometric and topological information about the configuration of anthropometric 3D face landmarks. In this way, evaluating face changes boils down to comparing the descriptors computed on 3D face scans taken at different times. By applying dimensionality reduction techniques to the dissimilarity matrix of descriptors, we get a space in which each face is a point and face shape variations are encoded as trajectories in that space. Our results show that persistent homology is able to identify features which are well-related to overweight, and may help assessing individual weight trends. The research was carried out in the context of the European project SEMEOTICONS, which developed a multisensory platform which detects and monitors over time facial signs of cardio-metabolic risk.

1 Introduction

Overweight and obesity represent a major risk factor for a large spectrum of diseases, including cardiovascular; diabetes; musculoskeletal disorders; and some cancers [2]. According to the World Health Organization (WHO) [3],

D. Giorgi, M. A. Pascali, S. Colantonio, O. Salvetti
Institute of Information Science and Technologies, National Research Council of Italy, Via G. Moruzzi 1, 56124 Pisa, Italy
E-mail: maria.antonietta.pascali@isti.cnr.it

B. J. Matuszewski, P. Henriquez
Computer Vision and Machine Learning Research Group, School of Engineering, University of Central Lancashire, PR1 2HE Preston, UK

in 2014 around 39% of adults aged 18 and over were overweight, and 13% were obese. Moreover, overweight and obesity are increasing at a rapid rate in most of the European member states and in the Americas. On the other hand, individuals are increasingly motivated to play an active role in managing their own health [35]. Thus, a key issue is the development of personal health monitoring systems able to support people in estimating and tracking over time their health status, and offering tailored guidance towards lifestyle improvements [4,50]. These systems have to extract and analyse meaningful features connected with health status and disease risk. In this paper, we focus on weight gain as a risk factor for cardio-vascular disease, and study a method to quantify patterns in face shape variation due to weight gain.

This research was carried out in the context of the European project SE-MEOTICONS [1], which developed a multisensory platform in the form of a mirror. The platform detects and monitors over time facial signs of cardio-metabolic risk, cardio-vascular diseases being one of the leading causes of mortality worldwide. The guiding principle behind the design of the mirror is that it should easily fit into daily-life settings (the home, the gym, the pharmacy), by maximizing non-invasive and unobtrusive interaction with the users. Therefore, it requires contact-less data acquisition and non-invasive sign detection and analysis. According to a semeiotic model of the face for cardio-metabolic risk [14], the face signs include 3D morphological face descriptors of overweight and obesity, to be computed on a 3D face model reconstructed from range data acquired by a 3D scanner, possibly low-cost.

We propose to detect and track face changes connected with overweight and obesity using Persistent Homology [20], a computational topology technique for analysing shape, and shape changes in particular [7]. Given a 3D face scan labelled with a set of landmarks, we compute persistence intervals, a shape descriptor which gives information on the geometry and topology of the face. Then, evaluating face changes over time boils down to comparing the persistence intervals computed on 3D face scans taken at different times. Our algorithm develops in three steps:

- Represent faces trough the simplicial complex given by the Delaunay triangulation of 23 soft-tissue facial landmarks (Section 4.1);
- Filter the simplicial complexes by assigning to each edge the Euclidean or geodesic distance between the corresponding landmarks (Section 4.2);
- Compute and compare the persistence intervals on the filtered complexes to track face changes (Section 4.3).

Comparing the persistent intervals on a set of faces produces a dissimilarity matrix. By applying dimensionality reduction techniques to the dissimilarity matrix, we get a metric space in which each face is a point, and face shape variations are trajectories in that space (Section 5). We experimented on a dataset of synthetic 3D faces simulating weight changes, generated using a parametric morphable model [40], with promising results. By analysing the position of thin and fat people in this metric space, we show that persistent homology is able to identify features which are well-related to overweight. Also,

by analysing the shape patterns of single individuals as trajectories, we show that our technique helps assessing trends in weight change on individuals.

In addition, we performed a preliminary study on a set of real data. A set of proper methods for the automatic 3D reconstruction of 3D faces, registration and labelling of facial meshes were studied and implemented (Section 6.1 and 6.2). A key point is that the automation in the acquisition of the 3D scan and in the processing, preliminary to the computation of persistent homology, is a must for a system non-invasive and unobtrusive, able to trace over time the features related to overweight. Nine volunteers were enrolled, and an acquisition campaign was organized to collect for each volunteer the 3D scan of the face, the body weight, and the BMI, so as to compare our descriptors with standard physical parameters measuring overweight and obesity. The results demonstrate the feasibility of our approach in real settings, in terms of effectiveness and computational cost (Section 6.3).

2 State of the art

Back in 1942, D'Arcy Wentworth Thompson expressed the importance of investigating biological form in a fully quantitative manner [47]. We may say that D'Arcy Thompson's vision has come true: in the last century, *morphometrics* came of age, as the discipline dealing with the quantitative study of form [42]. This was mainly accomplished by applying univariate and multivariate statistics to measures such as linear distances, angles and ratios. In the 1980's, it became clear that a more complex approach to the study of shape was needed, which had to be able to capture the *geometry* of the morphological structures under study and retain its information through the analysis. It was the birth of *Geometric Morphometrics* [16], which quantifies the variation in the shape of anatomical objects using the Cartesian coordinates of anatomical landmarks, after the effects of non-shape variations (translation, rotation, scale) have been factored out. A rich statistical theory for shape analysis supported the analysis of shape variation [19,8].
Advances in statistics, computer vision and computer graphics provided large body of tools for shape representation and quantification, including formalisation of the statistical shape analysis, shape descriptors and similarity measures [54,44,7]. Among others, these include: deformable templates [51], geometric descriptors [52], geodesic representations [9], and level sets [55]. Due to a simplicity of the representation and the explicit form in which objects are measured, shapes are often represented by finite sets of points. For this type of representation, by far the most popular way to describe shapes and model shape variations is to use active shape models (ASM) [13]. Although the ASM has well known limitations, the method is still frequently used in real applications. The Kendalls shape analysis (KSA) [30], alleviate some of the ASM limitation, by constraining shapes to lie on a unit sphere rather than on a linear subspace. A different approach to shape representation is provided by the elastic shape analysis (ESA) framework [53,43], where shapes are represented

by functions rather than point sets. This approach removes a fundamental weakness of the ASM and KSA, i.e. their dependence on the adopted shape sampling scheme. Effective methods for applying ESA methodology for surface representation have been proposed in [31,32,48].

In the literature, computer vision tools have been applied to body analysis to estimate height, weight and other parameters enclosed in the body appearance, and most of the methods proposed are based on computing body measurements. Velardo and Dugelay proposed a model for the weight estimation based on geometric body measurements extracted from the 2D silhouette of the human body [49]. Giachetti and colleagues [24] used heterogeneous body scans to automatically extract geometrical parameters related to body fat. Their research was focused on parameters not depending on the precise location of anatomical landmarks, and robust against pose and mesh quality, so as to be used in healthcare applications. They found a high correlation between several parameters with total body less head (TBLH) fat and trunk fat (acquired by Dual-energy X-rays Absorptiometry).

Though it is well known that the face is involved in the process of fat accumulation, there is no consensus in the literature about which are the facial morphological correlates of body fat. An increase in some facial dimensions was observed in a study about the face morphology of obese adolescents [22]. Moreover, there are interesting studies showing that some geometrical facial features may be related to both attractiveness and Body Mass Index (BMI) [11]. Also, Lee et al. [34] found that facial geometric features measured on 2D images, include Euclidean distances, angles and face areas defined by selected soft-tissue landmarks are related to both BMI and Waist Circumference. We argue that shape changes cannot be attributed to single, local variations in the position of individual landmarks. Also, shape changes involve shifts in the position of landmarks relative to other landmarks. Therefore, we need techniques that enable one to globally analyse the landmark configuration on a face.

3 Basics in persistent homology

This section recalls the main mathematical concepts behind our method. Section 3.1 defines simplicial complexes and their filtration, which serve as an input to the computation of persistence intervals, summarized in Section 3.2.

3.1 Simplicial complexes and filtrations

A simplicial complex K is a collection of simplices (faces) such that:

- Any face of a simplex of K is in K;
- The intersection of any two simplices in K is either a face of both or empty.

One can think of K as a complex which grows from an initial state K^0 to a final state $K^n = K$: the complex growing is formally encoded in the definition

of a filtration of K.

A *filtration* of a simplicial complex K is a nested sequence of simplicial complexes

$$\emptyset \subset K^0 \subset \ldots \subset K^i \subset \ldots \subset K^n = K. \tag{1}$$

An example is the Rips filtration: if a space X is known through a finite number of samples, for a real number $\epsilon > 0$ the *Rips complex* $R_\epsilon(X)$ is the complex whose k-simplexes are the subsets $\{x_0, x_1, \ldots, x_k\}$ of X such that $d(x_i, x_j) \leq \epsilon$ for all pairs x_i, x_j with $0 \leq i, j \leq k$. Whenever $\epsilon < \epsilon'$, there is an inclusion $R_\epsilon(X) \to R_{\epsilon'}(X)$ that reveals a growing complex.

We compute the Delaunay complex of the configuration of the anthropometric landmarks on 3D faces, and filter the complex using the Rips filtration guided by the Euclidean and geodesic distances among the landmarks (Section 4.2).

3.2 Persistent homology

Simplicial homology plays a central role in topological data analysis: it accounts for the number of holes (of given dimension) of a topological space X and it can be readily computed on any abstract simplicial complex K associated to X, depending only on X. For example, on a 3D manifold 0-homology classes correspond to connected components, 1-homology classes to tunnels, and 2-homology classes to voids. Persistent homology essentially looks at homology at different resolutions, analysing how topological attributes change through the filtration: new connected components appear or connect to the old ones, tunnels are created and closed off, voids are enclosed and filled in, etc. For example, as for 0-homology, each homology class corresponds to a connected component, and a homology class is born when a point is added, forming a new connected component (a 0-cycle). A homology class dies when two points belonging to different connected components are connected by a 1-chain, thus becoming a boundary. The lifespan of topological attributes is encoded in a simple descriptor called persistence interval. The aim is to furnish a scale to assess the relevance of topological attributes, under the assumption that longevity is equivalent to significance [20]. Formally, the *j-persistent homology group* of K^i counts how many homology classes of K^i still survive in K^{i+j}. Persistence represents the life-time of cycles in the growing filtration. Given a filtered simplicial complex $\{K^i\}_{i=0,\ldots,n}$, the *j- persistent k-th homology group* of K^i is defined as a group isomorphic to the image of the homomorphism

$$\eta_k^{i,j} : H_k(K^i) \to H_k(K^{i+j}) \tag{2}$$

induced by the inclusion of K^i into K^{i+j}.

In order to evaluate and visualize the persistent homology of a filtered complex, persistence intervals are used: a *persistence interval* is a pair

$$(i, j), \text{ with } i, j \in \mathbf{Z} \cup \{+\infty\} \text{ and } 0 \leq i < j,$$

such that there exists a cycle that is completed at level i of the filtration and becomes a boundary at level j. Persistence intervals can be represented in a *barcode*, a collection of horizontal line segments in a plane whose vertical axis represents an (arbitrary) ordering of homology generators, and the horizontal axis corresponds to the filtration parameter. The length of each segment is the lifespan of the associated cycles in the growing filtration, that is, the length of the interval (i, j): the start-point and end-point of each segment correspond to the cycle birth and death, respectively. An equivalent representation is the persistent diagram. A persistent diagram $Dgm_d(K)$ of the persistent homology of the complex K in dimension d is given by a set of points in a plane: each point (i, j) is a generator of the d-homology with birth at i and death at j. Persistence intervals (diagrams) can be compared using the Bottleneck distance.

We use persistence intervals as descriptors for 3D faces, and the distance between persistence intervals as a measure of the difference between different faces (Section 4.3).

4 Face description and comparison

The proposed method for the representation and analysis of 3D faces develops in three steps:

a. Representation of faces using the Delaunay complex of 23 landmarks $\{l_1, \ldots, l_{23}\}$, $l_i \in \mathbf{R}^3$. The landmarks are a subset of Farkas' landmarks, picked up according to the findings in [34]. Figure 1 shows the set of landmarks on a template face model [40] (Section 4.1).
b. Filtering of the simplicial complexes by assigning to each edge the Euclidean or geodesic distance between the corresponding landmarks (Section 4.2).
c. Computation and comparison of persistence intervals (Section 4.3).

The algorithm for the computation of persistence intervals (with Euclidean filtration) of each 3D face is summarized in Table 1.

	Algorithm: PH computation on a 3D face, Euclidean filtration	
1.	$\Lambda = \{l_i, i = 1, .., 23\}$	extract landmarks from *face_mesh*
2.	$Dist = pdist(\Lambda)$	compute pairwise Euclidean distances on Λ
3.	$DTK = DelaunayTri(\Lambda, Dist)$	build simplicial complex
4.	$mSpace = expMSp(DTK, Dist)$	explicit the metric space
5.	$Kstream = createStream(mSpace)$	define the Vietoris Rips stream
6.	$I = computeIntervals(Kstream)$	compute persistence modules and intervals

Table 1 Short description of the computation of the persistence intervals of a face mesh, using the Euclidean filtration.
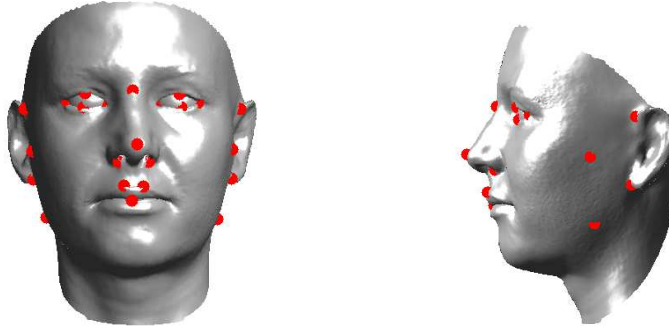
Fig. 1 The 23 landmarks used to represent faces.

4.1 The simplicial complex

A first choice for the representation of facial 3D data could be the flag complex obtained from the landmarks restricted to three dimensions, in which every pair of landmarks has an edge between them, and triangles and tetrahedra are consequently included. As suggested in [26], the drawback is that the flag complex would not generally have a geometric realization, whereas we do know that our data are inherently three-dimensional, since the landmarks come from human faces which live in the 3D Euclidean space. To preserve the Euclidean nature of data, we preferred a geometrically realizable subcomplex of the flag complex, namely the Delaunay triangulation of the 23 landmarks. We decided to compute the Delaunay triangulation on a template face, namely the Basel Face Model [40], rather than computing it on individual faces. We do this so that the abstract complex is fixed across all possible subjects, since we want to compare the absolute structure of the landmark configuration, common to all faces. Indeed, if the Delaunay triangulation was calculated separately for each face landmark configuration, the Delaunay triangulations could be slightly different, due to differences in the distances between landmarks for each subject.

4.2 The filtration

Once the complex is fixed, what varies on individual faces are the properties used for filtering the complex. We experimented with different filtrations. The first choice is a Rips filtration using the Euclidean distance between landmarks, similar to what has been done in [26] to study the outcome of clinical orthodontic procedures. All vertices enter at time t_0; the edge between landmarks i and j enters at time

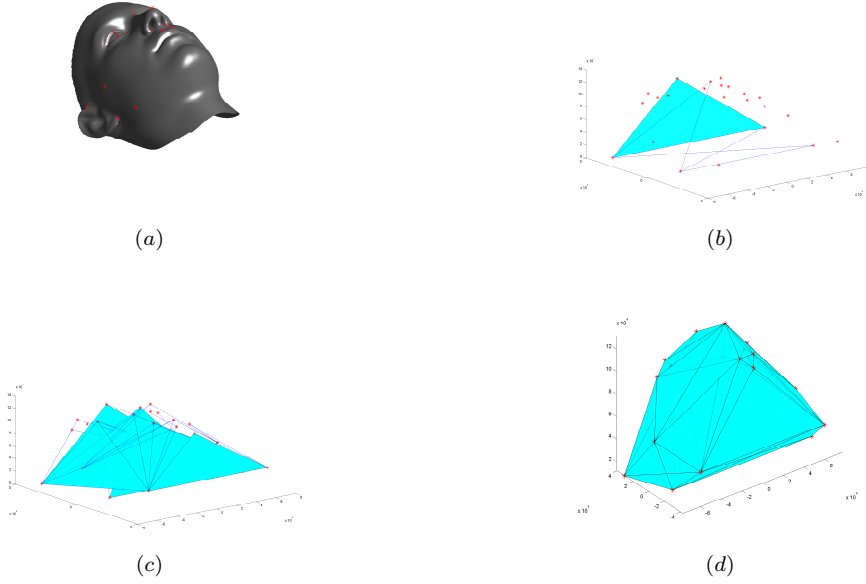$$t_{i,j} = M - d_E(l_i, l_j) \tag{3}$$

(a)

(b)

(c)

(d)

Fig. 2 The growth of the filtered complex associated to a face mesh: *(a)* single vertices associated with landmark points *(superimposed on a template face model)*; *(b)* some edges and the first face appear; *(c)* almost all the simplices appear; *(d)* the Delaunay complex of the face mesh.

where $d_E(l_i, l_j)$ is the Euclidean inter-landmark distance between landmarks i and j and $M = \max_{q,r} d_E(l_q, l_r)$; triangles and tetrahedra join the filtration when all of their faces have. Differently from [26], we change sign to the Euclidean distances. This is done so that landmarks that are far apart will have a smaller entry time. Figure 2 shows the process of growth of the complex. Also, differently from [26], we do not consider any normalization of the function across different individuals, as we are not interested in comparing inter-landmark distances within an individual with those same inter-landmark distances in other individuals, but rather on evaluating changes on the same individual in the process of gaining weight.

Moreover, we experimented with another filtration, which is similar to the one above but with geodesic distances instead of Euclidean ones:

$$t_{i,j} = M - d_G(l_i, l_j) \tag{4}$$

where $d_G(l_i, l_j)$ is the geodesic distance (computed using the Dijkstra algorithm, [15]) between landmarks i and j and $M = \max_{q,r} d_G(l_q, l_r)$. Geodesic distances take into account the intrinsic properties of faces, as they are bound to *walks* on surfaces. Geodesic distances encode different shape features than Euclidean distances: for example, the geodesic distance between the two land-
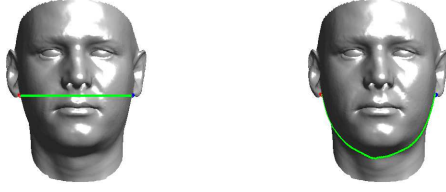
Fig. 3 Euclidean (left) and geodesic (right) distance between two landmarks.

marks in Figure 3 measures the length of the path passing below the chin, whereas the Euclidean distance measures the horizontal distance between the points. When facial soft tissues change with weight variations, the geodesics on the face can substantially change passing even through different parts of the face. Such a sudden change of region for some paths is exactly why the geodesic case is included in our study: this kind of phenomena produces a different pattern in the associated barcode, hence it could represent an interesting clue on the weight variation. We believe the filtered complexes above are a sensible choice to study the structure of a landmark configuration. Filtrations are defined in order to have descriptors invariant under translation and rotation, but purposely not invariant under scaling.

4.3 Computing and comparing persistence diagrams

We computed persistent homology for the complex of each face mesh, using the filtrations described in Equation 3 and 4, with homology of dimension 0, 1, and 2. That is, we computed 0-, 1-, 2-dimensional persistence intervals for each face. Figure 4 shows a face and its persistent intervals represented as barcodes. Persistence intervals of dimension d were compared via the Bottleneck distance, [17]. The computations were performed in MATLAB, with code adapted from the program JavaPlex [45].

Finally, the next Section describes how the shape dissimilarities are used to build a metric space in which each point represents a 3D face. Trajectories in this space are used to visualize patterns of 3D facial variations.

5 Analysis of patterns of shape variation

As a longitudinal study on real subjects to monitor weight and 3D face changes was not available, a dataset of synthetic 3D faces simulating weight changes was generated using a parametric morphable model [40] and used for the first experiments.
The Basel Face Model (BFM) is a morphable face model computed by applying PCA techniques to a large set of scans (200 real faces, 100 female and 100 male). The BFM can be deformed controlling the regression coefficients
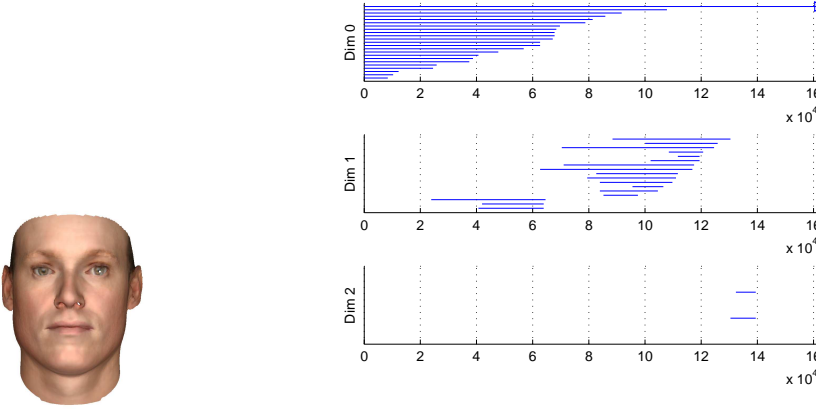
Fig. 4 Example of a face with its three barcodes. The barcode in the 0-dim (top) provides a visualization of persistence intervals representing connected components: all vertices enter at time 0, and only one connected component survives (being the face complex topologically equivalent to a ball in the 3D space). The barcode in the 1-dim (center) provides a visualization of persistence intervals representing loops: some loops appear while the complex growth, but all of them die before the face complex is completed. The barcode in the 2-dim (bottom) provides a visualization of persistence intervals representing voids: two voids appear while the complex growth, but both die soon after their birth.

extracted through PCA, accounting for face shape variations related to age, gender, height, and weight. Of course, such a model cannot be deformed in any real face, because it depends on the set of real faces used for its computation. Also, faces are labelled with different sets of anatomical landmarks (Farkas and MPEG4-FDP feature point coordinates and indices). These characteristics make the Basel Face Model a natural and effective choice for producing synthetic data to test the technique we developed.

We first generate 25 seeds from the mean face, varying randomly all the four parameters, in order to produce a synthetic study population of 25 subjects. Then, only the weight parameter of each seed is increased of a constant quantity in ten steps. An example of the sequence of fattening faces produced is given in Figure 5.

This gives a dataset of 250 faces, divided into 10 groups ordered according to increasing fatness. Computing persistent homology on this dataset gives 250×3 persistent intervals, one for each dimension (0-, 1-, 2-degree homology).

5.1 Metric space from shape dissimilarities

In statistical shape analysis, the analysis of shape variation is usually carried out in a feature space or in shape space (space of concatenated landmarks coordinates with imposed scale, translation and rotation equivalence class, i.e. similarity transformation invariance), i.e. a Kendall space [30]. Since persistent

Fig. 5 A seed face deformed in ten steps, by increasing the weight parameter of a constant quantity (20 units). The unit used in the BFM for the weight parameter cannot be related to standard measurement units (e.g. in kg).

diagrams are not vectors, we took a different route and worked in the metric space in which points are persistence diagrams, and the metric is given by the Bottleneck distance among them, as done in [23]. This approach is general and flexible, in that it can be adapted to arbitrary descriptors, other than feature-based (e.g., graphs).

Hence, the dissimilarity space built on the dataset yielded three different dissimilarity matrices, one for each homology dimension (0, 1, 2); each dissimilarity matrix is of dimension 250×250. Then, a dimensionality reduction technique was applied to these matrices. We experimented with classic Multidimensional Scaling (MDS, also known as principal coordinates analysis [25]) to explore and visualize our set of data, and got three $250 \times p$ matrices, with $p < 250$. In the analysis that follows, we set $p = 2$. Each row in the matrices represents the coordinates of a face in the lower-dimensional embedding space. Note that different embedding techniques could be used, possibly including non-linear dimensionality reduction techniques such as isometric feature mapping [46] and Laplacian eigenmaps [38].

As expected, given the intrinsic characteristic of faces, we found that homology of dimension 2 was not significant, whereas homology of degree 0 and 1 proved to be more informative. This can be seen from the observation of the dissimilarity matrices in Figure 6. Therefore, in what follows we will only analyse data pertaining to homology of degree 0 and 1. We first analyse visually the results separately, that is, for 0- and 1-homology, and for Euclidean- and geodesic-based filtration. This is done to study the different information they provide (Section 5.2). Then, we analyse quantitatively the results of integrated distances, given by the sums of matrices, in terms of classification rate (Section 5.3).

5.2 Qualitative analysis

The analysis of scatterplots in the embedding space seems to confirm that the proposed technique is able to identify 3D features which are well-related to overweight and obesity. Figure 7 (left) and 8 (left) show the first two embedded MDS coordinates, labeled by fatness level, from 1 to 10, for the filtrations based on Euclidean and geodesic distance, respectively, in dimension 0 and 1. It can be seen that in both cases the subjects are well distributed in the space according to their fatness level. In other words, our technique seems to
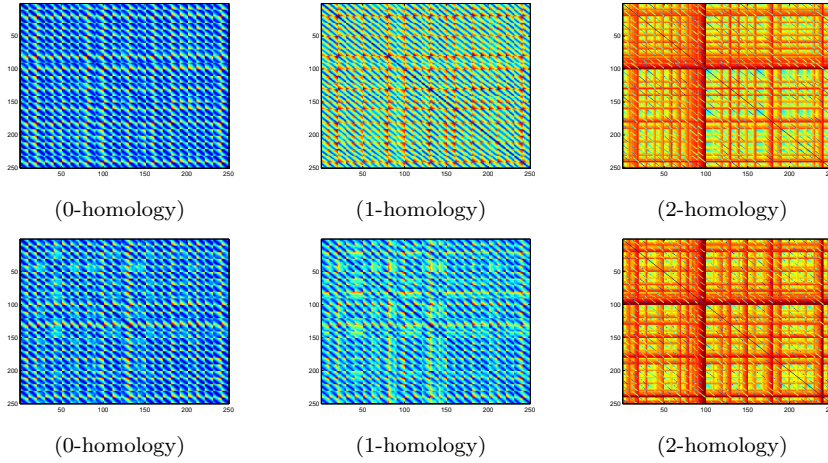
| (0-homology) | (1-homology) | (2-homology) |

| (0-homology) | (1-homology) | (2-homology) |

Fig. 6 Dissimilarity matrices between persistence intervals, for the Euclidean (top) and geodesic filtration (bottom). Colour map, ranging from blue to red, indicates increasing values of dissimilarity; e.g. the blue diagonal corresponding to entries (i, i) correctly indicates that the dissimilarity between the persistence diagram associated to the i-th face with itself is zero.

be able to separate faces of people in different groups. This can be better appreciated in Figure 7 (right) and 8 (right), which show the first two embedded MDS coordinates for a subset of faces, namely thinner people (red), medium people (green) and fatter people (blue) in our dataset. Since our essential task is the description of morphological change over time on a subject, we must check whether our technique enables us to discover a trend in a longitudinal study. A way to do this is visualizing the shape patterns of individuals as trajectories [12] in the dissimilarity space. Each individual has a trajectory which is made of ten consecutive points. For a given trajectory, we can analyse four attributes, namely *location* (the starting and ending points); *size* (the magnitude of the vector between the endpoints); *orientation* (the direction of the vector between the endpoints); and *shape*. In our context, the location depends on the specific, initial traits of each individual. The size is a measure of the difference in shape between the thinnest and the fattest morphing of the individual. The orientation is crucial: a consistent orientation would indicate that our technique is able to detect and encode the process of getting weight. Figure 9 shows the trajectories of three sample faces in our dataset in the embedding space given by the first two coordinates. It is clear that the orientation is consistent, from left to right in accordance with weight gain, but the shape of trajectories do differ, especially for 0-homology. Figure 10 shows the trajectories if only the first embedding coordinate is taken into account. As the trajectories are more homogeneous, it seems that the first coordinate alone is able to identify the trend in fat variation better than the first two coordinates.
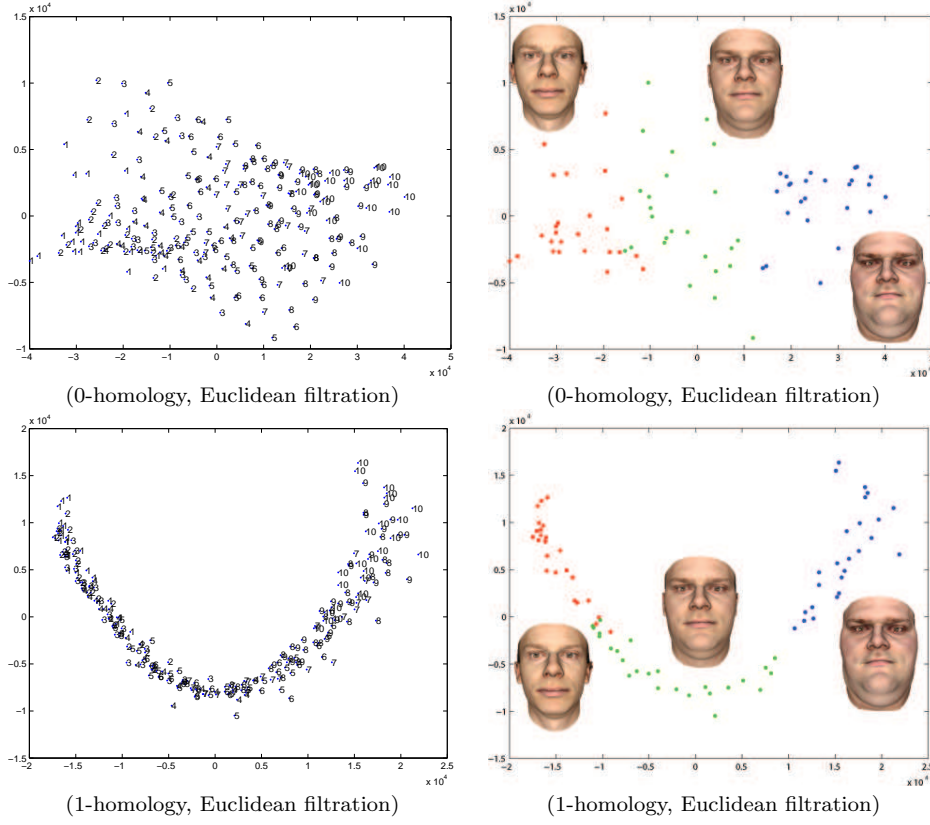
(0-homology, Euclidean filtration)

(0-homology, Euclidean filtration)

(1-homology, Euclidean filtration)

(1-homology, Euclidean filtration)

Fig. 7 Scatterplot of the first two embedded MDS coordinates, labeled by fatness level, from 1 to 10, for the Euclidean-based filtrations (left). The scatterplot with only a subset of faces shown (right).

5.3 Quantitative analysis

We have seen that persistence intervals based on both the Euclidean distance and the geodesic distance between landmarks are able to quantify shape variation, and that they take different properties into account. This suggests that summing up their information could be beneficial. We can define the distance between two faces as the sum of the Bottleneck distances between their persistence intervals in different dimensions, or with different filtrations.

Let us denote $d^0_{E,G}$ ($d^1_{E,G}$) the sum of distances in dimension 0 (dimension 1) obtained with the Euclidean and the geodesic filtration, and $d^{0,1}_E$ ($d^{0,1}_G$) the sum of distances with the Euclidean (geodesic) filtration in both dimensions 0 and 1.

The classification rate on our dataset for the integrated distances above defined was assessed and reported in Table 2. We also evaluated two popular
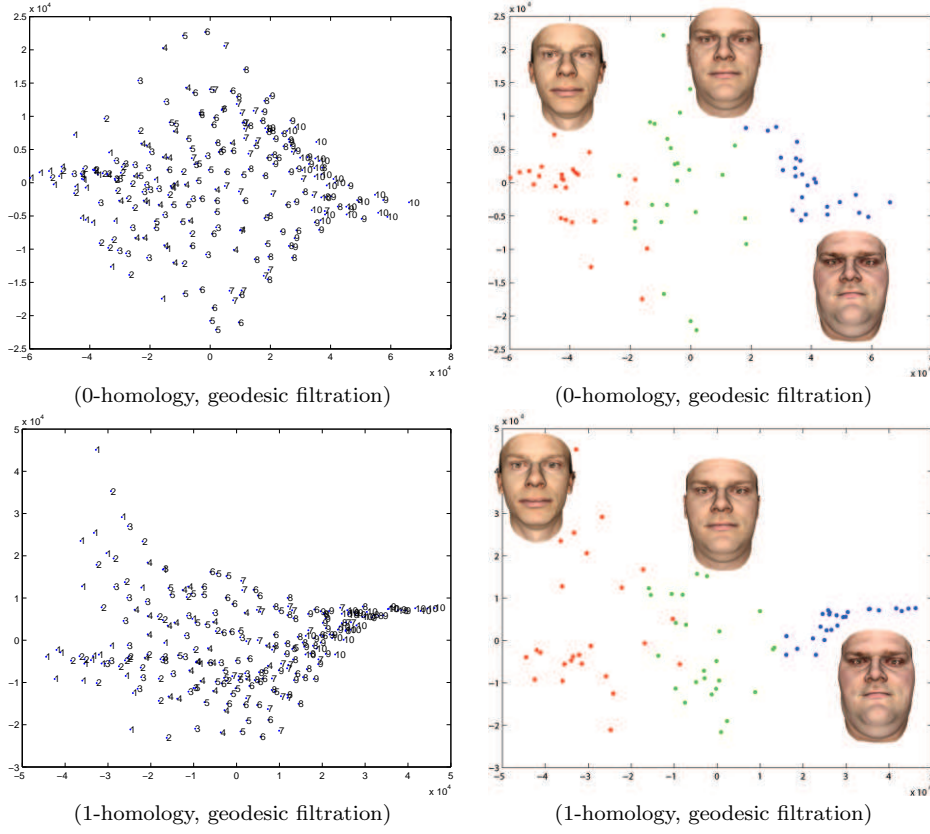
(0-homology, geodesic filtration)

(0-homology, geodesic filtration)

(1-homology, geodesic filtration)

(1-homology, geodesic filtration)

Fig. 8 Scatterplot of the first two embedded MDS coordinates, labeled by fatness level, from 1 to 10, for the geodesics-based filtrations (left). The scatterplot with only a subset of faces shown (right).

state-of-the-art shape descriptors, namely Shape Distributions (SD) [39] and Spherical Harmonics (SH) [29].

The classification rate refers to the number of subjects correctly attributed to their group, out of the 10 groups in the dataset, in a leave-one-out experiment. Notice that the classification task is very challenging, since the variation among consecutive groups, in terms of fat gain, is rather small. Therefore, we considered three classification rates, for correct prediction within the first, second and third choice. It can be seen that integrated distances perform better than or comparably to Spherical Harmonics, and significantly better than Shape Distributions.

The best performance seems to be provided by the filtration based on geodesic distances, with distances summed over homology dimensions.
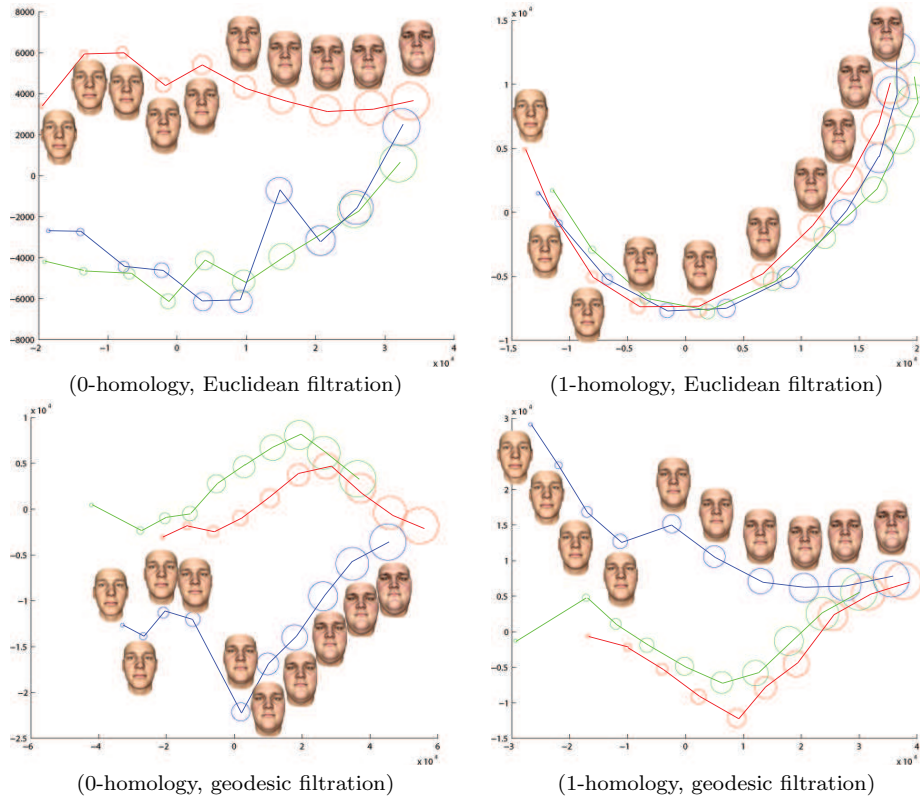
(0-homology, Euclidean filtration)

(1-homology, Euclidean filtration)

(0-homology, geodesic filtration)

(1-homology, geodesic filtration)

Fig. 9 Trajectories of three sample faces in the space given by the first two embedding coordinates.

	1st	2nd	3rd
$d_{E,G}^0$	42%	63%	**88.4%**
$d_{E,G}^1$	41.2%	64%	83.2%
$d_E^{0,1}$	30.8%	54.4%	77.2%
$d_G^{0,1}$	**42.4%**	**72.2%**	83.6%
SD	22.4%	44.8%	65.6%
SH	38.4%	65.6%	88.0%

Table 2 Classification rates for integrated distances and competitors. The best rates are marked as bold.

6 A preliminary study on real data

The method was applied on a small set of real data for extracting from the morphological analysis of the subject's face the information about body weight. The sample used is made of 9 adult subjects: 3 women and 6 men. The characteristics of the sample are listed in Table 3. The 3D facial data are acquired
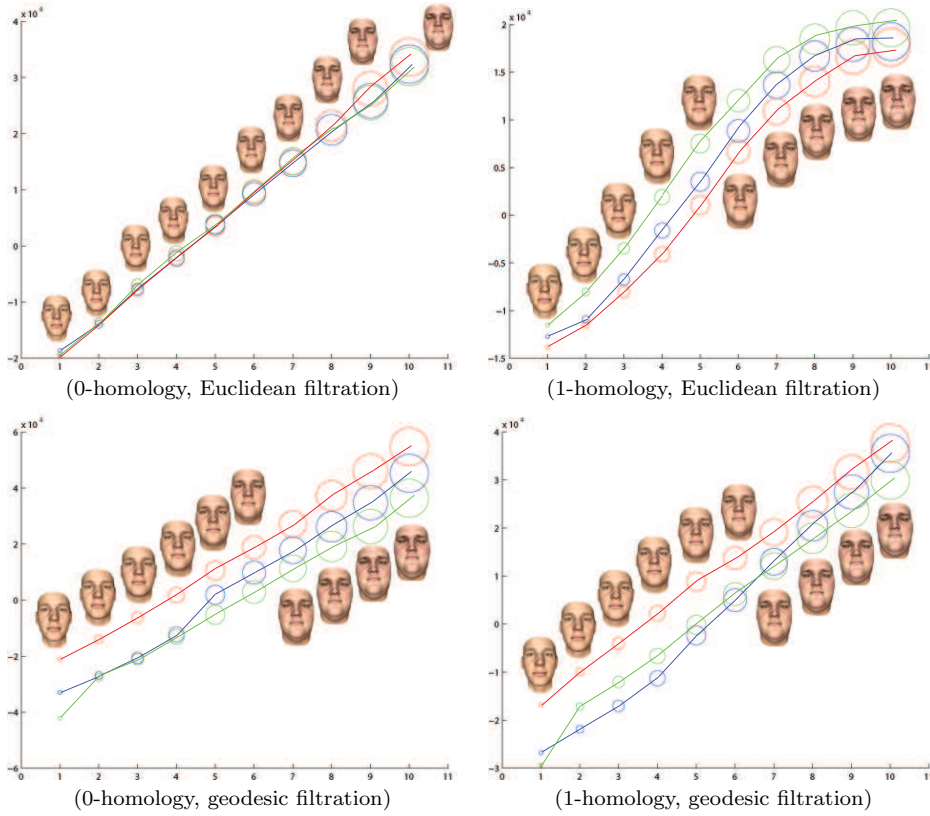
(0-homology, Euclidean filtration)   (1-homology, Euclidean filtration)

(0-homology, geodesic filtration)   (1-homology, geodesic filtration)

Fig. 10 Trajectories of three sample faces in the space given by the first embedding coordinate.

with a low cost depth sensor; then the 3D faces are automatically reconstructed and labelled, as described in Section 6.1 and Section 6.2. Finally, the results obtained with real data are described in Section 6.3. Of course, the analysis carried out for the large, synthetic dataset cannot be applied as it is to the set of real data, due to the fact that powerful tools, such as the multidimensional scaling, are no more appropriate with a so different sample size.

6.1 3D face reconstruction

This section describes a system for a 3D geometric face reconstruction based on inexpensive, readily available depth sensors such as Kinect or Asus Xtion. The requirement is to produce a 3D manifold mesh, representing face surface, accurate enough to enable an extraction of informative facial features.
The 3D reconstruction method recovers points cloud from a sequence of depth frames and merges them into a single consistent global object. This is achieved

Id	gender	BMI	weight (kg)
1	F	20.3	57.8
2	F	20.8	57.2
3	F	20.9	48.9
4	M	21.9	64.1
5	M	22.2	68.8
6	M	24	78.1
7	M	24.6	71.2
8	M	27.4	78.15
9	M	28.5	79.6

Table 3 Characteristics of the studied sample, subjects sorted by BMI ascending.

using a structure from motion algorithm [37]. The data fusion is performed by tracking the global model using a coarse-to-fine iterative closest point (ICP) algorithm with the reconstructed surface estimated using a truncated signed distance function (TSDF) and a Poisson meshing algorithm [28].

The first, preprocessing, step before applying the reconstruction procedure is the face segmentation. This step prevents background objects and body parts from interfering in the reconstruction process. The segmentation technique used to extract the person's face from the depth image is based on a face pose estimation method [21] and application of a bounding box segmentation around the estimated head centre. The face pose estimation [21] calculates, based on a random forest, the head position in a 3D space (x, y, z). The 3D head centre is then projected into the 2D depth frame and used as the centre of a rectangular bounding box of 180x180 pixels. All the depth points inside the defined bounding box will remain in the segmented face, the rest will be discarded as well as the pixels with bigger z values than $z_c + 15$cm, where z_c is the estimated distance of the head centre from the sensor.

The segmentation process is applied to a depth frame sequence where the user has rotated the head to the left and then to the right. Subsequently, the segmented depth data is passed to the reconstruction system in order to obtain the fused 3D model and the corresponding mesh.

The average area of the triangles forming the 3D face reconstruction meshes is $2.8mm^2$. The volume which is captured with the depth sensor is a cube of $512 \times 512 \times 512$ voxels which represents a volume of $1.5 \times 1.5 \times 1.5$ metres $(3.375m^3)$ on average. The reconstructed faces occupy $0.3 \times 0.3 \times 0.2$ metres volume, which corresponds to approximately $715,827$ voxels from the captured volume, with each voxel representing around $25mm^3$ volume.

6.2 3D face labelling

Different landmark detection methods in depth or 3D data can be found in the literature, such as [5],[18],[27] and [36]. Numerous approaches have been used to tackle this problem, for example the work described in [5] proposed a

heuristic method combined with a SIFT detector applied to local search windows to locate 9 landmarks automatically. In the approach presented in [27], 11 fiducial points are automatically located on a pair of range and portrait images using a search over an area centred at the average location of the fiducial point location in the training data. The technique explained in[36] selects 3 feature points by determining the local shape index at each point within the 2.5D scan.

The landmarking problem presented in this paper requires the automatic location of 23 landmarks (Farka's model [33], Figure 11), and some of them are very difficult to detect using the methods described above. In order to detect this challenging landmarks arrangement a method similar to the one described in [18] is needed. This approach enables detection of landmarks in the areas with no significant facial features.

The method used in this work to produce the approximate positions of the predefined 23 landmarks is based on a non-rigid registration of a deformable model [41]. Such 3D facial representation can model wide variety of faces. The 3D faces are represented by a low dimensional shape space vector (SSV) of the statistical shape model (SSM). As explained in [41], the SSM is trained using different faces with 83 landmarks selected manually in each face of the training dataset.

The process of labelling a 3D face reconstruction starts with the fitting of the deformable model to the reconstructed point cloud. This algorithm involves the estimation of the pose and shape parameters, which allow adapting the model to a new personalised model. The shape and the pose are iteratively estimated in turn. The iterative closest point algorithm is used to estimate the pose, and the projections on the estimated SSM are used to estimate the shape parameters encoded in the SSV.

Once the model has been matched to the 3D reconstruction, the previously selected 23 landmarks of interest are transferred to their corresponding closest point on the reconstructed 3D surface (see Figure 11). The results obtained with the proposed labelling method were compared to a manually obtained ground truth of 22 reconstructions. The manual labelling was only based on shape data, no texture information was used in the process. The average from three manual annotations performed by two observers was used as ground truth. The average error of the landmark identification is 6 mm. This should be compared with the corresponding calculated annotation error between the two observers of 2 mm.

The face deformable model used for this work contains 5,832 points (See Figure 11 on the top left), covering a volume similar to the reconstructed faces. Variations in resolutions and number of points used for the deformable registration could significantly influence landmarking accuracy. In previous works, it has been shown that using a smaller number of landmarks located around eyes, nose and mouth with a 3D deformable model representing only central part of the face could provide better results for landmarks on the mouth. Additionally, it has been also demonstrated that using 3D reconstruction with significantly higher number of vertices, when compared with the number of the model
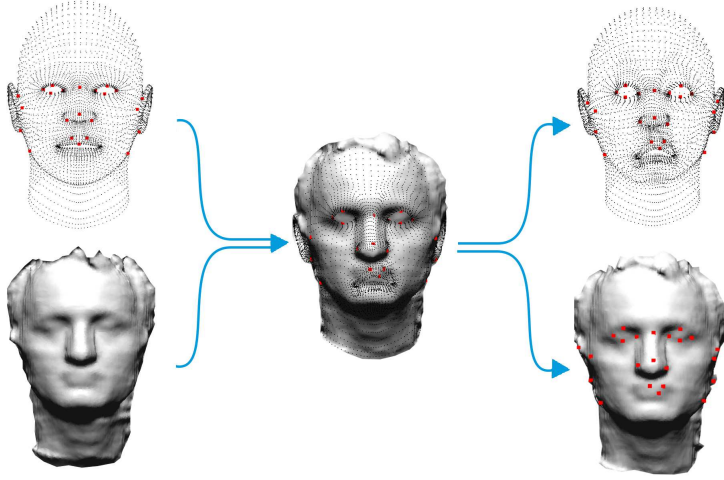
Fig. 11 Labelling process. Inputs: Manually labelled deformable model (top left) and 3D reconstruction (bottom left). The model registered to the reconstruction (centre). Outputs: Deformed model (top right) and labels transferred from the model to the 3D reconstruction (bottom right).

vertices, can lead to the model converging to a local minima resulting in an incorrect labelling. This was solved by down-sampling the reconstructed surfaces. For the described problem, the resolution of the adopted low cost depth sensor is suitable for the 3D reconstructions being labelled by the deformable model from [41]. Although position accuracy of the landmarks around mouth, cheeks and ears could be improved, they are the most challenging landmarks for detection as the mouth is the part of the face with the biggest deformations and the Farkas' landmarks on the cheeks and close to the ears do not have significant distinctive features to be robustly identified.

6.3 Results on real data

3D facial data acquired through low-cost sensors, such as in Figure 11, makes clear that the aim claimed in the paper's title is quite hard to reach. The very low requirements about the acquiring conditions, the number of points used for the shape description (only 23 landmarks), and the strong automation of the whole system, desirable for a day-life implementation (e.g. at home, by not skilled people) imply having data with low resolution and low accuracy, thus any subsequent processing may be negatively affected by them.

Nevertheless, the persistent homology keeps carrying a certain amount of information about body weight. As in the synthetic dataset, we found that homology of dimension 2 was not significant, whereas homology of degree 0 and 1 proved to be more informative. Then, barcodes in dimension 0 and 1 of each
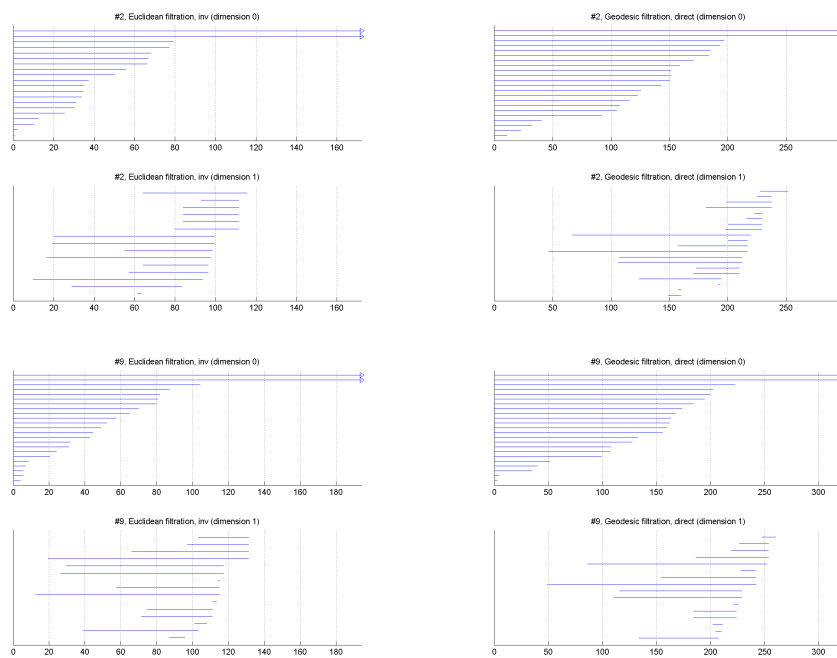
Fig. 12 Barcodes: Subject 2 (first row) and 9 (second row); both Euclidean (first column) and geodesic filtration (second column).
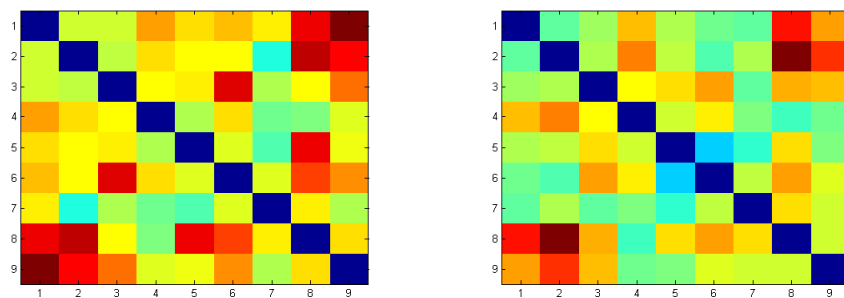

Fig. 13 Dissimilarity matrices computed on the set of real data. Colour map ranging from blue to red, according to increasing dissimilarity.

Stage	Time (seconds)
Segmentation	12
Reconstruction	14
Meshing	3
Labelling	1
PH computation	1
Total	**31**

Table 4 Average processing time of the different stages of the proposed method when it is run using a 3.6GHz i7 processor. 23 subjects from the dataset were used for these measurements.

subject has been analysed; barcodes of two subjects, with quite different BMI values, are shown in Figure 12).

A further analysis focused on the persistent intervals computed in dimension 0, both Euclidean and geodesic filtration, and on the persistent intervals computed in both dimension 0 and 1, with the geodesic filtration; this choice is motivated by the results obtained in the larger dataset of synthetic faces, already presented in Table 2. Let us denote $d^0_{E,G}$ the sum of distances in dimension 0 obtained with the Euclidean and the geodesic filtration, and $d^{0,1}_E$ the sum of distances with the Euclidean filtration in both dimensions 0 and 1. Then the dissimilarity matrices between the persistent intervals corresponding to $d^0_{E,G}$ and to $d^{0,1}_E$ are computed to assess if any discriminative value is maintained in the case of real data.

The sample has been sorted with increasing BMI. Thus, we expect that a dissimilarity matrix, able to encode differences among the subject in the set, should be a matrix with a blue diagonal, with the two corners up-right and down-left being red, and with the other cells in the middle going from blue to red accordingly to the distance from the diagonal. And in Figure 13 we discover dissimilarity matrices showing a colour pattern similar to the ideal good one.

Table 4 shows average processing time for different stages of the algorithm, including 3D reconstruction, labelling, and the persistent homology computations (reading the data, building the complex, computing persistence barcodes in dim 0 and 1, for both Euclidean and geodesic filtration, computing the dissimilarity matrix).

7 Conclusions

We described an ongoing work in the European project SEMEOTICONS, which is developing a multisensory platform which detects and monitors over time facial signs correlated with cardio-metabolic risk, and gives personalized guidance to individuals to improve their habits. Our contribution to the project is the automatic assessment of weight gain via 3D shape analysis, being obesity and overweight one of the main factors of cardio-metabolic risk. We used persistent homology, which offered a tunable framework for face description

and comparison. We qualitatively and quantitatively described the behaviour of our descriptors, and the results seem to be promising.

In literature, different choices of simplicial complex and filtration are possible, such as the Clique, the Witness and Lazy Witness, the Rips-Vietoris, the Delaunay, the Tidy: each of them has specific features and an optimal setting of application. A first direction of future research is the usage of a different filtration, defined on the whole face rather than on the landmark configuration, for example using Morse filtrations based on curvature or distances from reference points. A landmark-free approach would also give the advantage of not having to pre-compute face landmarks. Although landmarks can be identified with different strategies [10], locating landmarks with the desired accuracy could be difficult on bad-quality face scans.

The future work also include the use of hybrid descriptors which analyse both shape and texture, as for example the hybrid geodesic distance in [6]. Finally, an advantage of our framework is that it is flexible, meaning that it can be adapted to the study of face properties other than weight accumulation, by defining a different, *ad hoc* filtered complex. Therefore, we plan to investigate on the study of other signs, including facial asymmetry.

Concerning the application of the morphological analysis based on the persistent homology in day-life settings, the experimentation described in Section 6 proves that the method implemented provides a valid tool to analyse very peculiar 3D objects, such as human faces even in real settings. We plan to enlarge the sample size and to better manage the inaccuracy in the landmarks' location, in order to take advantage of statistical tools for the analysis of the results, and improve the discriminative power of the proposed method.

References

1. http://www.semeoticons.eu/
2. Global health observatory data. http://www.who.int/gho/ncd/risk_factors/overweight/en/
3. Obesity and overweight: fact sheet n 311; 2015. http://www.who.int/mediacentre/factsheets/fs311/en/
4. Andreu-Cabedo, Y., Henriquez, P., Colantonio, S., Coppini, G., Favilla, R., Germanese, D., Giannakakis, G., Giorgi, D., Larsson, M., Marraccini, P., Martinelli, M., Matuszewski, B., Milanic, M., Pascali, M., Pediaditis, M., Raccichini, G., Randeberg, L., Salvetti, O., Stromberg, T.: Mirror mirror on the wall...; an intelligent multisensory mirror for well-being self-assessment. In: Multimedia and Expo (ICME), 2015 IEEE International Conference on, pp. 1–6 (2015). DOI 10.1109/ICME.2015.7177468
5. Beretti, S., Ben Amor, B., Daoudi, M., del Bimbo, A.: 3d facial expression recognition using sift descriptors of automatically detected keypoints. In: The Visual Computer Journal (2011)
6. Biasotti, S., Cerri, A., Giorgi, D., Spagnuolo, M.: Phog: Photometric and geometric functions for textured shape retrieval. Computer Graphics Forum **32**(5), 13–22 (2013). DOI 10.1111/cgf.12168. URL http://dx.doi.org/10.1111/cgf.12168

7. Biasotti, S., Falcidieno, B., Giorgi, D., Spagnuolo, M.: Mathematical tools for shape analysis and description. Synthesis Lectures on Computer Graphics and Animation **6**(2), 1–138 (2014)
8. Bookstein, F.L.: Biometrics, biomathematics and the morphometric synthesis. Bulletin of mathematical biology **58**(2), 313–365 (1996)
9. Bronstein, A., Bronstein, M., Kimmel, R.: Numerical geometry of non-rigid shapes. Springer (2008)
10. Celikutan, O., Ulukaya, S., Sankur, B.: A comparative study of face landmarking techniques. EURASIP Journal on Image and Video Processing **13** (2013)
11. Coetzee, V., Chen, J., Perrett, D.I., Stephen, I.D.: Deciphering faces: Quantifiable visual cues to weight. Perception **39**(1), 51–61 (2010)
12. Collyer, M.L., Adams, D.C.: Phenotypic trajectory analysis: comparison of shape change patterns in evolution and ecology. Hystrix **24**(1), 75–83 (2013)
13. Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape models: their training and application. Computer Vision and Image Understanding **61**(1), 38–59 (1995)
14. Coppini, G., Favilla, R., Gastaldelli, M., Colantonio, S., Marraccini, P.: Moving medical semeiotics to the digitalrealm. semeoticons approach to face signs of cardiometabolic risk. In: HEALTHINF 2014 (2014)
15. Cormen, T.H., Leiserson, C.E., L.Rivest, R.: Introduction to algorithms. The MIT Press, Cambridge, MA (1994)
16. Corti, M.: Geometric morphometrics: An extension of the revolution. Trends in Ecology & Evolution **8**(8) (1993). DOI http://dx.doi.org/10.1016/0169-5347(93)90261-M. URL http://www.sciencedirect.com/science/article/pii/016953479390261M
17. d'Amico, M., Frosini, P., Landi, C.: Using matching distance in size theory: A survey. International Journal of Imaging Systems and Technology **16**(5), 154–161 (2006)
18. Dibeklioglu, H., Salah, A., Akarun, L.: 3d facial landmarking under expression, pose, and occlusion variations. In: 2nd IEEE International Conference on Biometrics: Theory, Applications and Systems, pp. 1–6 (2008)
19. Dryden, I.L., Mardia, K.V.: Statistical shape analysis, vol. 4. Wiley Chichester (1998)
20. Edelsbrunner, H., Letscher, D., Zomorodian, A.: Topological persistence and simplification. Discrete and Computational Geometry **28**(4), 511–533 (2002)
21. Fanelli, G., Weise, T., Gall, J., Van Gool, L.: Real time head pose estimation from consumer depth cameras. In: Annual Symposium of the German Association for Pattern Recognition, vol. 6835, pp. 101–110 (2011)
22. Ferrario, V.F., Dellavia, C., Tartaglia, G.M., Turci, M., Sforza, C.: Soft tissue facial morphology in obese adolescents: a three-dimensional noninvasive assessment. The Angle orthodontist **74**(1), 37–42 (2004)
23. Gamble, J., Heo, G.: Exploring uses of persistent homology for statistical analysis of landmark-based shape data. J. Multivar. Anal. **101**(9), 2184–2199 (2010). DOI 10.1016/j.jmva.2010.04.016
24. Giachetti, A., Lovato, C., Piscitelli, F., Milanese, C., Zancanaro, C.: Robust automatic measurement of 3d scanned models for human body fat estimation. IEEE Journal of Biomedical and Health Informatics **19**(2), 660–667 (2015)
25. Gower, J.: Principal coordinates analysis, In: Armitage, P. and Coulton, Theodore eds. Encyclopedia of Biostatistics. John Wiley and Sons, Inc. (1998)
26. Heo, G., Gamble, J., Kim, P.T.: Topological analysis of variance and the maxillary complex. Journal of the American Statistical Association **107**(498), 477–492 (2012). DOI 10.1080/01621459.2011.641430. URL http://dx.doi.org/10.1080/01621459.2011.641430
27. Jahanbin, S., Choi, H., Bovik, A.: Passive multimodal 2-d+3-d face recognition using gabor features and landmark distances. In: IEEE Transactions on Information Forensics and Security, pp. 1287–1304 (2011)
28. Kazhdan, M., Bolitho, M., Hoppe, H.: Poisson surface reconstruction. In: Eurographics symposium on geometry processing, pp. 61–70. Aire-la-Ville, Switzerland (2006)
29. Kazhdan, M., Funkhouser, T., Rusinkiewicz, S.: Rotation invariant spherical harmonic representation of 3d shape descriptors. In: Symposium on Geometry Processing, vol. 6 (2003)
30. Kendall, D.: Shape manifolds, procrustean metrics and complex projective spaces. Bulletin of the London Mathematical Society **16**(2), 81–121 (1984)

31. Kurtek, S., Klassen, E., Ding, Z., Srivastava, A.: A novel riemannian framework for shape analysis of 3d objects. In: Proc. IEEE Conf. Comput. Vis. Pattern Recog., pp. 1625–1632 (2010)
32. Kurtek, S., Klassen, E., Gore, J.C., Ding, Z., Srivastava, A.: Elastic geodesic paths in shape space of parameterized surfaces. IEEE Trans. Pattern Anal. Mach. Intell. **34**(9), 1717–1730 (2012)
33. Lee, B., Kim, J.Y.: Predicting visceral obesity based on facial characteristics. In: BMC complementary and alternative medicine (2014)
34. Lee, B.J., Kim, J.Y.: Predicting visceral obesity based on facial characteristics. BMC complementary and alternative medicine **14**(1), 248 (2014)
35. Loucks, E., Schuman-Olivier, Z., Britton, W., Fresco, D., Desbordes, G., Brewer, J., Fulwiler, C.: Mindfulness and cardiovascular disease risk: state of the evidence, plausible mechanisms, and theoretical framework. Curr Cardiol Rep. **17**(12) (Dec. 2015). DOI 10.1007/s11886-015-0668-7
36. Lu, X., Jain, A., Colbry, D.: Matching 2.5d face scans to 3d models. In: IEEE Trans. on Pattern Analysis and Machine Intelligence (2006)
37. Newcombe, R., Izadi, S., Hilliges, O., Molyneaux, D., Kim, D., Davison, A., Kohli, P., Shotton, J., Hodges, S., Fitzgibbon, A.: Kinectfusion: Real-time dense surface mapping and tracking. In: IEEE International Symposium on Mixed and Augmented Reality, pp. 127–136 (2011)
38. Niyogi, P., Smale, S., Weinberger, S.: Finding the homology of submanifolds with high confidence from random samples. Discrete & Computational Geometry **39**(1-3) (2008)
39. Osada, R., Funkhouser, T., Chazelle, B., Dobkin, D.: Shape distributions. ACM Transactions on Graphics (TOG) **21**(4), 807–832 (2002)
40. Paysan, P., Knothe, R., Amberg, B., Romdhani, S., Vetter, T.: A 3d face model for pose and illumination invariant face recognition. In: Proc. IEEE AVSS, Genova, Italy. IEEE (2009)
41. Quan, W., Matuszewski, B., Shark, L.K.: Improved 3-d facial representation through statistical shape model. In: IEEE International Conference on Image Processing, pp. 2433–2436 (2010)
42. Reyment, R.A.: An idiosyncratic history of early morphometrics. In: Advances in morphometrics, pp. 15–22. Springer (1996)
43. Srivastava, A., Klassen, E., Joshi, S.H., Jermyn, I.H.: Shape analysis of elastic curves in euclidean spaces. IEEE Trans. Pattern Anal. Mach. Intell. **33**(7), 1415–1428 (2011)
44. Srivastava, A., Turaga, P., Kurtek, S.: On advances in differential-geometric approaches for 2d and 3d analyses and activity recognition. Image and Vision Computing **30**, 398–416 (2012)
45. Tausz, A., Vejdemo-Johansson, M., Adams, H.: Javaplex: A research software package for persistent (co)homology. Software available at http://code.google.com/javaplex (2011)
46. Tenenbaum, J.B., De Silva, V., Langford, J.C.: A global geometric framework for nonlinear dimensionality reduction. Science **290**(5500), 2319–2323 (2000)
47. Thompson, Wentworth, D., et al.: On growth and form. On growth and form. (1942)
48. Tumpach, A., Drira, H., Daoudi, M., Srivastava, A.: Gauge invariant framework for shape analysis of surfaces. IEEE Trans. Pattern Anal. Mach. Intell. **39**(1), 46–59 (2016)
49. Velardo, C., Dugelay, J.L.: Weight estimation from visual body appearance. In: BTAS 2010, 4th IEEE International Conference on Biometrics: Theory, Applications and Systems, September 27-29, 2010, Washington DC, USA (2010)
50. Velardo, C., Dugelay, J.L.: What can computer vision tell you about your weight? In: Proc. of 20th European Signal Procesing Conference (EUSIPCO 2012), Bucharest (Romania) - August 27-31, 2012 (2012)
51. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Diffeomorphic demons: efficient non-parametric image registration. Neuroimage **45**, 61–72 (2009)
52. Windheuser, T., Schlickewei, U., Schmidt, F., Cremers, D.: Geometrically consistent elastic matching of 3d shapes: a linear programming solution. In: International Conference on Computer Vision (ICCV) (2011)
53. Younes, L.: Computable elastic distances between shapes. SIAM J. Appl. Math **58**, 565–586 (1998)

54. Younes, L.: Space and manifolds of shapes in computer vision: An overview. Image and Vision Computing **30**, 389–397 (2012)
55. Zhang, Y., Matuszewski, B., Histace, A., Precioso, F.: Statistical model of shape moments with active contour evolution for shape detection and segmentation. Journal of Mathematical Imaging and Vision **47**(1), 35–47 (2013)