POLITECNICO DI TORINO
Repository ISTITUZIONALE

Landmarking-based unsupervised clustering of human faces manifesting labio-schisis dysmorphisms

Original
Landmarking-based unsupervised clustering of human faces manifesting labio-schisis dysmorphisms / Conti, D., Bonacina, L., Froio, A., Marcolin, F., Vezzetti, E., Speranza, D.. - In: INFORMATICA. - ISSN 0350-5596. - 41:4(2017), pp. 507-516.

Availability:
This version is available at: 11583/2697306 since: 2018-01-14T10:27:17Z

Publisher:
Slovene Society Informatika

Published
DOI:

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

26 November 2026

Landmarking-Based Unsupervised Clustering of Human Faces Manifesting Labio-Schisis Dysmorphisms

Daniele Conti
Department of Applied Science and Technology, Politecnico di Torino
corso Duca degli Abruzzi 24, 10129 Torino, Italy

Luca Bonacina
Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano
Piazza Leonardo da Vinci 32, 20133 Milano, Italy

Antonio Froio
Department of Energy, Politecnico di Torino, corso Duca degli Abruzzi 24, 10129 Torino, Italy

Federica Marcolin and Enrico Vezzetti
Department of Management and Production Engineering, Politecnico di Torino
corso Duca degli Abruzzi 24, 10129 Torino, Italy
E-mail: federica.marcolin@polito.it

Domenico Speranza
Dipartimento di Ingegneria Civile e Meccanica, Universit degli Studi di Cassino e del Lazio Meridionale
Viale dell'Universit, 03043 Cassino (Frosinone), Italy

Keywords: facial dysmorphism, labio-schisis, diagnosis, feature extraction, landmarking, clustering, D-MST, artificial intelligence, decision support

Received: July 20, 2016

Ultrasound scans, Computed Axial Tomography, Magnetic Resonance Imaging are only few examples of medical imaging tools boosting physicians in diagnosing a wide range of pathologies. Anyway, no standard methodology has been defined yet to extensively exploit them and current diagnoses procedures are still carried out mainly relying on physician's experience. Although the human contribution is always fundamental, it is self-evident that an automatic procedure for image analysis would allow a more rapid and effective identification of dysmorphisms. Moving toward this purpose, in this work we address the problem of feature extraction devoted to the detection of specific diseases involving facial dysmorphisms. In particular, a bounded Depth Minimum Steiner Trees (D-MST) clustering algorithm is presented for discriminating groups of individuals relying on the manifestation/absence of the labio-schisis pathology, commonly called cleft lip. The analysis of three-dimensional facial surfaces via Differential Geometry is adopted to extract landmarks. The extracted geometrical information is furthermore elaborated to feed the unsupervised clustering algorithm and produce the classification. The clustering returns the probability of being affected by the pathology, allowing physicians to focus their attention on risky individuals for further analysis.

Povzetek: Predstavljena je D-MST metoda za nenadzorovano grupiranje slik obrazov za diagnosticiranje.

1 Introduction

Medical imaging has seen an important enhancement in the past decades thanks to various technological achievements. Magnetic Resonance Imaging (MRI), Computed Axial Tomography (CAT), X-ray imaging, Ultrasound scans imaging (US) provide physicians with valuable information to diagnostic purpose. In particular, foetal diseases attracted attentions and efforts with the common aim to improve the current diagnosis techniques, fostered by the objective of defining a tailored therapy as early as possible. A crucial role in this activity is played by three-dimensional ultrasound scans, which could provide in-depth detailed images of foetal morphology in a safe and non-invasive way. Despite technological improvements, medical image-driven diagnosis suffers the deficiency/absence of automated computer science treatment, even for diseases such as Fetal Alcohol Syndrome (FAS) and labio-schisis [1–5]. This work aims to provide a methodology and a tool for supporting the diagnosis of labio-schisis pathology (cleft lip), which has been chosen due to its relatively large incidence in the population [6]. This task is conceived for prenatal diagnosis and stems from a recently developed work [7], in which an automatic procedure was designed to process a stack of 2D ultrasound scans of foetal faces

by transforming the standard DICOM images in a PLY 3D model. The core of the proposed method relies on the clustering technique. Common algorithms for unsupervised data clustering belong to two main categories: partitioning algorithms and hierarchical clusterings [8]. Algorithms in the first class, such as K-means or the recently proposed Affinity Propagation (AP) [9], define a subset of individuals called *centroids*, i.e. the class exemplars, to which any other node is compared. Hierarchical clustering algorithms, such as Single Linkage, compare couples of individuals and merge the closest in a class, thus creating a chain of hierarchical dependencies. In the first case, the expected number of classes, i.e. of centroids, should be *a priori* defined (except for Affinity Propagation), while in the second case the pruning of the hierarchical tree specifies how many clusters to be returned [10]. Among them, the bounded Depth Minimum Steiner Trees (D-MST) unsupervised clustering algorithm is chosen for this study [11], [12].

For privacy reasons, after the feasibility test on an ideal foetuses dataset, the public Bosphorus database was adopted, containing facial depth maps of 105 adult individuals showing the seven fundamental facial expressions, [13]. The defect was artificially simulated on the faces by modifying some Bosphorus point clouds. This way, seven artificial faces were generated with left-sided and right-sided labio-schisis. The algorithm is designed to be robust against different defect types.

The work is structured as follows. Firstly, an outline of geometrical face description formalization is presented together with related feature extraction aimed at landmarks localization. Then, information coming from geometrical descriptors are exploited to feed the unsupervised D-MST clustering algorithm for discriminating individuals according to the presence/absence of the pathology.

2 Methods

The algorithm is meant to detect the presence/absence of cleft lip in a query face. It is designed to work with three-dimensional foetal faces obtained through automatic elaboration of bidimensional ultrasound scan stacks. On the other hand, it has been extensively tested with a large size adult individuals dataset.

2.1 Mathematical background

Bosphorus database provides coordinates of facial point clouds, obtained through laser scans, as a binary file. A pre-built routine is provided together with data for reading binary files, extract cloud points data, and return information as a matrix containing the Cartesian coordinate of each point. The facial surface can be seen from the mathematical standpoint as the locus defined as

$$z \in \mathbb{R}^3 : z = f(u, v)$$

and it can be referred to as a *free-form surface*. A free-form surface is required to be smooth, with normal vector defined almost everywhere but edges, cusps, etc., but not belonging to a simple mathematical class of surfaces, like conics for example. Anyway, it can be divided in subdomains, each of them treatable as a linear combination of simple geometries. Thus, we define a surface patch divided in domains as an n-tuple of functions:

$$f(u, v) = (f_1(u, v), f_2(u, v), ..., f_n(u, v)). \quad (1)$$

Taking advantage of this definition, in order to objectively compare one face to another, the surface is point-by-point mapped-on with entities belonging to the Differential Geometry domain, here called *geometrical descriptors*. Twelve different geometrical descriptors, together with first and second derivatives, are chosen: three coefficients of the first fundamental form, i.e. E, G, F, three coefficients of the second fundamental form, i.e. e, f, g, the Gaussian curvature K, the mean curvature H, the principal curvatures k_1 and k_2, the shape index S and the curvedness index C. In the following section we go through the adopted geometrical descriptors definitions [14, 20].

2.2 Geometrical descriptors

A free-form surface is not an Euclidean geometry. Thus, distances on a face cannot be computed with the standard formula $s^2 = \sum_{i=0}^{d}(u_i - v_i)^2$. The first fundamental form, also called Riemann metric, allows to define equivalent concept of distance upon a non-Euclidean surface. For $d = 2$, the infinitesimal distance element ds can be defined as $ds^2 = Edu^2 + 2Fdudv + Gdv^2$. E, F, G are the first fundamental form coefficients. They can also be expressed in terms of partial derivatives as

$$E = \|f_u\|^2, \quad (2)$$

$$F = \langle f_u, f_v \rangle, \quad (3)$$

$$G = \|f_v\|^2, \quad (4)$$

where $f_u = \frac{\partial f}{\partial u}$. Moreover, by defining the normal unit vector in point (u, v) belonging to the face domain as

$$N(u, v) = \frac{f_u \times f_v}{|f_u \times f_v|}(u, v), \quad (5)$$

we can also introduce the second fundamental form as $ds^2 = edu^2 + 2fdudv + gdv^2$, with

$$e = \langle N, f_{uu} \rangle, \quad (6)$$

$$f = \langle N, f_{uv} \rangle, \quad (7)$$

$$g = \langle N, f_{vv} \rangle, \quad (8)$$

where $\langle \cdot \rangle$ denotes the scalar product. In order to introduce curvatures, let us consider the tangent plane $T_p(f)$ to f in point $p = f(u_0, v_0)$; it can be defined as the two-dimensional vector subspace $Df(u, v) \subset \mathbb{R}^3$, where D is the functional differential operator. For each point

p, there exists a set of orthonormal vectors $\{e_1, e_2\}$ for the tangent plane T_p, such that $DN_p(e_1) = -k_1 e_1$ and $DN_p(e_2) = -k_2 e_2$, where k_1 and k_2 are called the principal curvatures and e_1 and e_2 the principals directions at p. In terms of the principal curvatures, Gaussian curvature K and mean curvature H can be introduced:

$$K = k_1 k_2 = \frac{eg - f^2}{EG - F^2}, \tag{9}$$

$$H = \frac{k_1 + k_2}{2} = \frac{eG - 2fF + gE}{2(EG - F^2)}. \tag{10}$$

Thus, the principal curvatures can be obtained as the roots of the quadratic equation $k^2 - 2Hk + K = 0$, resulting in

$$k_1 = H + \sqrt{H^2 - K} \tag{11}$$

and

$$k_2 = H - \sqrt{H^2 - K}. \tag{12}$$

An insightful method for evaluating curvatures was introduced by Koenderink and van Doorn [15], who defined the shape index S and the curvedness index C. They can be expressed in terms of the principal curvatures:

$$S = -\frac{2}{\pi} arctan \frac{k_1 + k_2}{k_1 - k_2}, \; with \, S \in [-1, +1], \; k_1 > k_2, \tag{13}$$

$$C = \sqrt{\frac{k_1^2 + k_2^2}{2}}. \tag{14}$$

The range spanned by the shape index can be partitioned into nine different intervals, spanning from cup to dome, each of them representing a particular shape. The curvedness index provides information about how gently the surface bends. Differently from other geometrical descriptors such as the shape index, it is not independent on the unit length and has the dimension of a reciprocal length. These geometrical descriptors are computed for each point of the face and exploited for both landmarking and clustering phases.

2.3 Landmarking

Geometrical descriptors are suitable to be mapped point-by-point on facial surfaces. So, by computing their values for all individuals in the dataset, a distribution of their local behaviour is obtained. Such a statistics can be exploited as characteristic information of the facial region and used to automatically localize facial landmarks. Landmarks are typical facial points, such as the nose tip, i.e. the *pronasal*, the nose basis, i.e. the *subnasal*, the internal and external eye extrema, i.e. the *endocanthions* and *exocanthions*. Figure 1 shows the most renown landmarks.

Landmarks can be automatically detected by setting tailored thresholds, empirically defined, in specific facial areas (where each landmark is more likely to be) for all geometrical descriptor facial maps. Focusing on the specific problem addressed in this work, we developed a program for automatically extracting *pronasal* and *labrum superior* points. For further details about landmarking, please refer to [16–23].

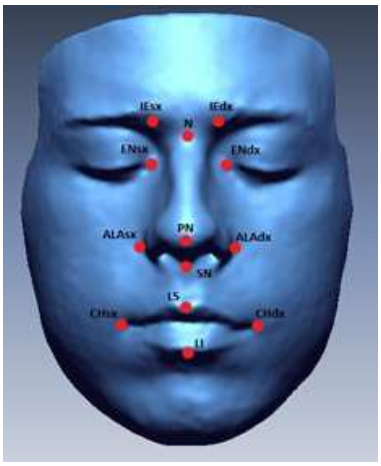

Figure 1: Main soft-tissue landmarks. From top to bottom: IEsx/IEdx, left and right *Inner Eyebrows*. N, *Nasion*. ENsx/ENdx, right and left *Endochantions*. ALAsx/ALArx, *Alae*. PN, *Pronasal*. SN, *Subnasal*. LS, *Labrum superior*. CHsx/CHdx, righty and left *Cheilions*. LI, *Labrum Inferior*.

2.4 Clustering

In order to perform the clustering, the input database is put in the form of a $N \times M$ matrix with a row for each individual to be classified and as many columns as the number of geometrical descriptors to be exploited for facial description purpose. For instance, considering all individuals available, if we report values of all seventeen geometrical descriptors expressed in the *labrum superior* landmark, a 105 x 17 matrix is created. In our case, values of a subset of geometrical descriptors expressed in an arbitrary surface portion, covering the central part of the upper lip and departing on a straight line from the identified LS landmark, are considered, in order to catch sufficient information about the possible presence of labio-schisis pathology. Clustering algorithms require the definition of a dissimilarity measure to be used for one-to-one face comparisons. It leads to the so called *Dissimilarity Matrix*, S, whose entries $s_{i,j}$ with $i, j = 1, ..., N$ are the dissimilarities between any couple of individuals i and j. $s_{i,j}$ can be defined in different ways, depending on the kind of data to be clustered [24]. Spearman's correlation rank distance $\rho_{i,j}$ is chosen as dissimilarity measure, in order to overcome problems related to descriptors, such as the shape and curvedness indexes, lying on different domains and with different measure scales. In particular:

$$\rho_{i,j} = \sum_{m=1}^{M} \frac{(x_{i,m} - \bar{x}_i) \cdot (x_{j,m} - \bar{x}_j)}{\sqrt{\sum_m (x_{i,m} - \bar{x}_i)^2 \cdot \sum_m (x_{j,m} - \bar{x}_j)^2}}, \tag{15}$$

$$s_{i,j} = 1 - \rho_{i,j}. \tag{16}$$

In 15, the usual definition of correlation is given, being ρ the Spearman's rank correlation coefficient. In particular, the variable x is the data rank, instead of the bare data itself as for Pearson correlation coefficient, i and j identify individuals and index m runs over geometrical descriptors values. In the second equation, s is the dissimilarity. The input dataset is treated as a fully connected weighted graph $G(n, e)$, with $n = 1, ..., N$, $e = \{(n_i, n_j)\}$, $i, j = 1, ..., N$, with N individuals and $\frac{N(N-1)}{2}$ weighted edges e, whose weight is the dissimilarity $s_{i,j}$ between the two individuals. A fictitious node, called *root*, is added to this graph and connected to all other nodes by a weight λ, empirically defined. From a Physical viewpoint, λ can be interpreted as the chemical potential of the system, i.e. the cost for adding an individual to the system itself, and it governs the most probable number of clusters returned. On the other hand, a depth parameter D is introduced to drive the final output. It is a constraint representing the maximum observable depth, namely the distance, in terms of nodes, from the root to the external leaves of clusters. Thanks to this additional parameter, D-MST interpolates between the Affinity Propagation algorithm, returning an arbitrary number of spherical clusters with $D = 2$, in which leaves are directly connected to the centroids via edge means with comparable dissimilarities, and the Single Linkage algorithm, in which $D > N$. In order to perform a classification, two variables (d_i, π_i) are assigned to any node and exploited to define an objective function to be minimized for detecting the optimal spanning tree T^* in the graph. The variable $d_i \in [2, N]$, $d_i \leq D$ is the distance, in terms of number of nodes, from the root, and assumes discrete values; variable $\pi_i = j$, $j \in [1, N]$, $j \neq i$ is a pointer tracking the ancestor of node i. Thus, the cost function is:

$$\begin{aligned} E(\{d_i, \pi_i\}_{i=1}^N) = \\ \sum_i s_{i,\pi_i} + \\ + \sum_{i,j \in \partial i} (h_{i,j}(\pi_i, \pi_j, d_i, d_j) + h_{j,i}(\pi_j, \pi_i, d_j, d_i)), \end{aligned} \tag{17}$$

where h_{ij} is defined as

$$h_{i,j} = \begin{cases} 0 & \{\pi_i = j \Rightarrow d_i = d_j + 1\} \\ -\infty & else, \end{cases} \tag{18}$$

and imposes an artificial constraint to the cost function that requires the returned optimum tree to be connected. In this terms, the probability of observing a configuration of variable for the optimum tree is given by the Boltzmann weight

$$P(\{\pi_i, d_i\}) \propto e^{-\beta E(\{\pi_i, d_i\})} \tag{19}$$

and it is maximized by a message passing algorithm described in [25] and [26].

3 Results

3.1 Preliminary analysis

Individuals' faces are reported in Bosphorus database as point clouds pre-ordered on a square grid and with the same orientation, in particular with nose oriented alongside with z-axis, x-axis aligned with chin-forehead line and, consequently, y-axis aligned from cheek to cheek. A first analysis of data contained in Bosphorus database is conducted by inspection, in order to examine the facial points suitable to our purpose. Indeed, referring only to faces with no expression, some facial point clouds showed degradation and low accuracy in shaping the face itself. In particular, it is not unusual to encounter data with a rough mouth surface that has no actual correspondence to the individual's picture accompanying data. Thus, all corrupted data were excluded from further analysis, resulting in an input matrix collecting 74 healthy individuals.
As a preliminary step, all facial point clouds are cropped in size, limiting the region of interest to a squared area. A four-pixels-side mean-filter is then applied in order to reduce noise and smooth surface peaks, where a pixel is intended as the squared surface area wrapping a point of the face. Moreover, the fact of having pre-oriented faces allowed us to avoid a pre-processing step aimed to provide data with a standard orientation. Most of all, it allowed to easily identify the central region of the face, by looking at those points with relatively higher z-coordinate. This way, the facial area containing nose and a mouth portion is identified and exploited for further analysis.

3.2 Computing geometrical descriptors and landmarking

Geometrical Descriptors are point-by-point computed, starting from derivatives. They are obtained by computing the surface gradients, along x and y directions, then by averaging values obtained in a ten-pixels-side window centred into the point of interest. All other geometrical descriptors can be obtained, as previously shown in Methods section, starting from first and second derivatives, and are easily computed for each point of the facial surface.

Once geometrical descriptors are obtained, they are exploited to define where the *pronasal* landmark and the *labrum superior* landmark are placed upon the surface. Our attention is focused most of all on the latter landmark, as it would affect the chosen area of investigation for the cleft

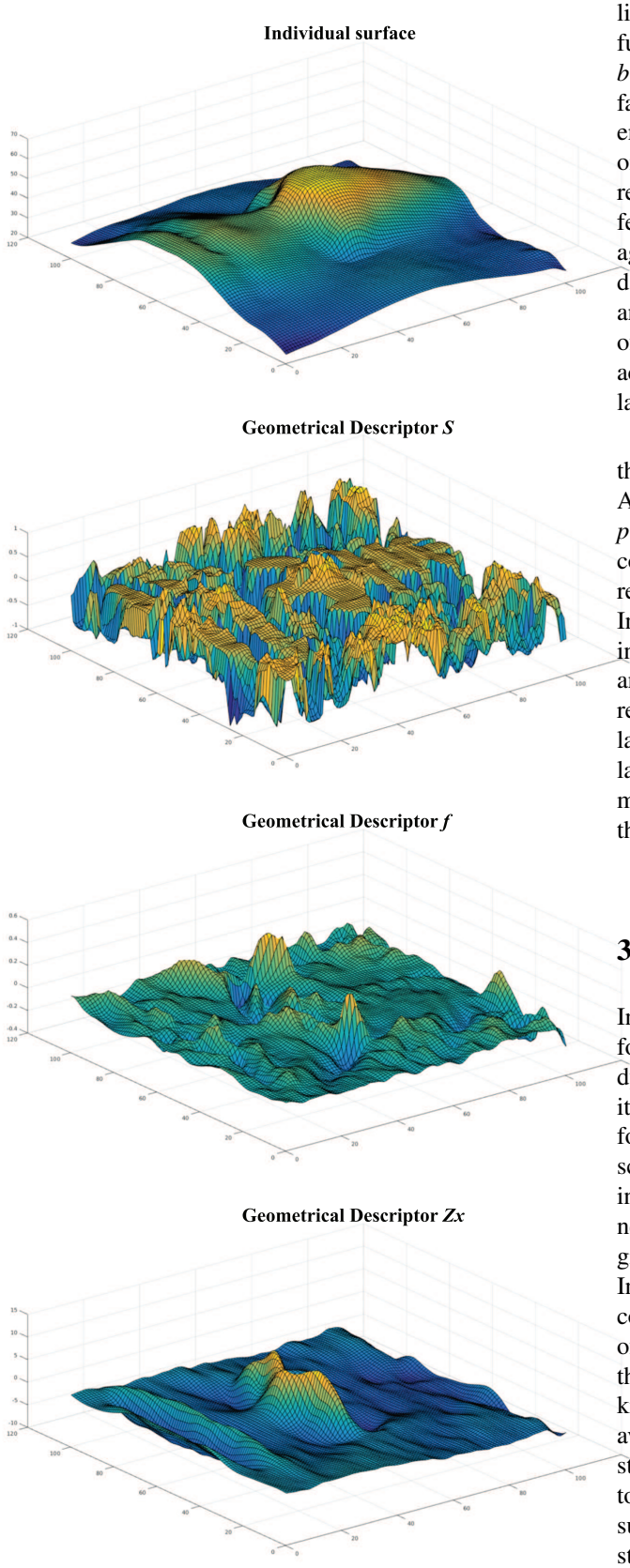

Figure 2: Geometrical descriptors mapped on a face. From top to bottom: bare facial surface, shape index S, second fundamental form f and derivative along x.

lip dysmorphism clustering. Indeed, the *pronasal* is helpful in proceeding with an accurate identification of the *labrum superior* itself. Starting from the central area of the face identified previously relying on z-coordinate, we set empirical constraints to values assumed by meaningful geometrical descriptors. They are the descriptors that, in the region of interest, present a characteristic behaviour. Referring to the previous work [19] and assessing the choice against this database, the shape index S, the second fundamental form coefficient f, and first derivatives along x and y directions have been chosen (figure 2). This subset of descriptors, conveniently constrained, leads to a 100% accuracy in the automatic determination of the *pronasal* landmark for the pre-processed database in exam.

Once the *pronasal* is identified, it is adopted to delimit the region of interest for the *labrum superior* detection. Approximatively half of the area going from chin to the *pronasal* itself is taken into consideration to detect this second landmark. In this region, the previous procedure is repeated changing only the geometrical descriptor adopted. In such a case, the chosen information relies upon the shape index S, the mean curvature H, the first derivative along x, and the second derivative along x. In this case, the accuracy reached by the algorithm is around 94%, but even when the landmark obtained by the algorithm and the ground truth landmark do not match perfectly, their relative distance remains around a few pixels. Thus, the error is not affecting the final output of the algorithm.

3.3 Prenatal applicability

In its original intention, the present work has been designed for pre-birth diagnosis of rare diseases manifesting facial dysmorphisms. Labio-schisis presents high incidence and it is clearly detectable through ultrasound-scans when the foetus is affected. Another connected pathology is palatoschisis. It is more difficult to be observed by US-scans inspection and its current diagnosis techniques would benefit of an automatic procedure for highlighting morphological differences that are symptomatic of the disease itself. In this perspective, the clustering procedure here proposed could be efficient only if it relies on an effective detection of the foetus's facial landmarks, which are clearly fuzzier than those of an adult. In our work, we tested the landmarking algorithm on the limited amount of real foetus data available and found that, after a tiny rearrangement of constraints imposed to the same subset of geometrical descriptors, the *pronasal* and the *labrum superior* landmarks were successfully detected (figure 3). Although this result is not statistically significant, it moves towards the application of such kind of procedure to pre-birth diagnosis. Indeed, once landmarks are identified correctly, the cleft lip manifesting face morphology reports similar differential geometry properties and, thinking toward a clustering perspective, it is totally equivalent to the case of adult individuals.

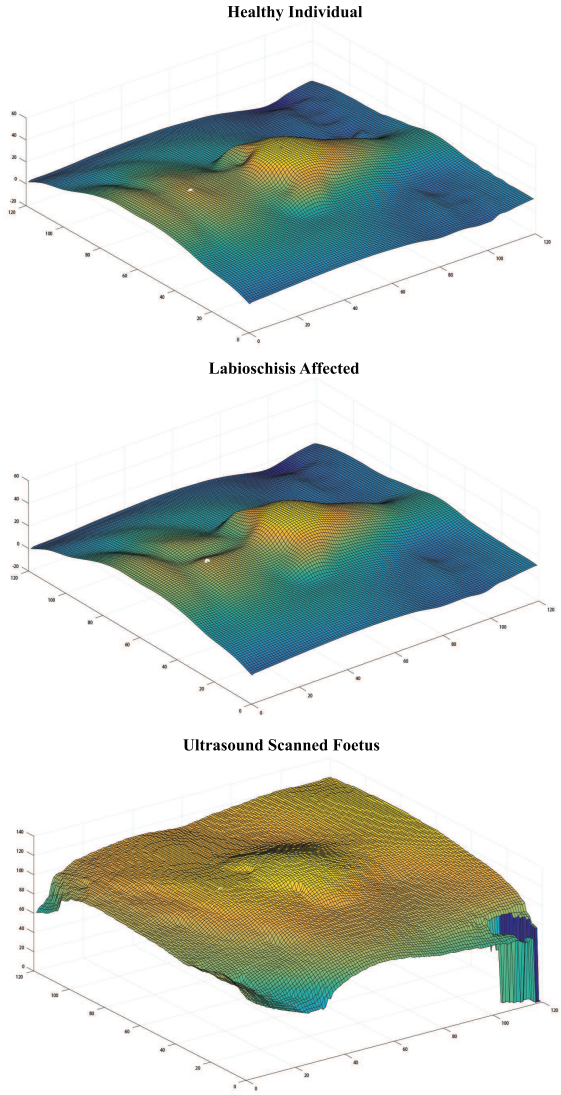

Figure 3: Surface Landmarking. For each surface, the white dot indicates the position of the *labrum superior* landmark. The comparison between faces not manifesting (top) and manifesting (centre) the labio-schisis is presented. Moreover, it is shown the case of landmarking on a foetal 3D model, (bottom). 0-valued points in the last figure indicate missing data.

3.4 Unsupervised clustering

Once the *labrum superior* landmark has been detected, the area of interest for the labio-schisis pathology is identified. In particular, a straight line laying on the upper lip is taken as a biometric information of the individual. The line length and width can be arbitrary, provided that it spans most of the lip itself, without invading other face regions. Therefore, if the cleft lip is present in the individual under investigation, it will also be in the region in exam. Moreover, the pre-orientation of individuals' face simplifies the identification of the lip line, avoiding a useless detection of its complete morphology.

A subset of meaningful geometrical descriptors expressed in any point composing this line is then stored into a row vector, building a matrix collecting all individuals in exam. The opposite in sign of the shape index S and of the coefficient of the first fundamental form G, the first and second derivatives along y of the free-form surface are sufficient to obtain the convergence of the clustering algorithm toward a successful classification. In our specific study, the chosen line is forty pixels-long and its width is three pixels, pinched at the *labrum superior* landmark. This choice is useful for the application of a median-filter on any three-pixels-side square of geometrical values spanning the line. Median filter smooths the descriptor behaviour along the surface, without adding artificial information to the one extracted from the geometrical analysis of the surface. In the end, the input matrix M presents a number of rows equal to 74 (number of healthy faces) plus 7 (number of artificially-induced cleft lip-affected faces) and a number of columns equal to 40 times the number of chosen geometrical descriptors.

Starting from M, the so called *similarity matrix* is computed. As previously mentioned, the similarity matrix S is a squared symmetric matrix which reports, with any of its entries s_{ij}, a measure of distance between any couple of individuals. In particular, the most suitable choice for the kind of data handled in this work is the Spearman's correlation rank distance, computed as $s_{ij} = 1 - \rho_{ij}$, where ρ is shown explicitly in equation 15. Once the similarity matrix is computed and the specific clustering depth D is set, the unsupervised clustering D-MST can be run.

As specified previously in the section Methods, D-MST is governed by an external parameter λ influencing the number of identified clusters. Lower values of λ would lead toward many clusters and non-linked individuals, while larger values of the parameter would return a single cluster collecting all nodes of the graph. In general, the maximum value of dissimilarities s_{ij} found in the S matrix is taken as upper bound for the parameter. In order to identify the most proper value of lambda to be chosen, stability regions of the clustering algorithm are investigated. They are intended as regions of the parameter space in which the algorithm converges toward a stable solution, in terms of number of clusters, despite the parameter change. So, the range between the minimum value of dissimilarity, exclu-

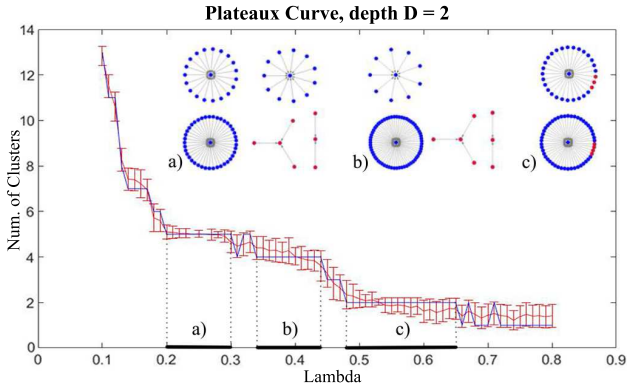

Figure 4: Number of clusters versus λ, for $D = 2$. Plateaux curve. The higher the λ value, the lower the number of clusters. For the plateaux associated to 5, 4 and 2 clusters, the relative clustering are overlaid. In particular, observing plateau a), two cleft lip clusters are detected correctly, even separating left- and right-sided cleft lips; on the contrary, healthy individuals population is divided in subpopulations as well. Increasing λ, blue clusters do not form a single class, while individuals affected by the pathology are merged to healthy clusters.

ded zero, and the maximum one, is linearly split up into a large number of bins. For each of them, the algorithm is run fifty times, in order to return results with a statistical significance. In such a way, what here is called *plateaux curve* is built and it plots the number of clusters returned as a function of the value of λ.

Such a procedure is repeated for the different depths investigated. Starting from depth $D = 2$, meaning a single link between centroid and leaves, the plateaux curve is obtained running the algorithm for values of lambda included into the range $\lambda \in [0.1, 0.8]$ and spaced 0.01. Figure 4 shows the passage from nearly no clusters, for λ close to 0, to a single cluster for λ large. The blue line shows the mode of number of clusters, while the red one is the average number of clusters with its standard deviation. An example of clustering obtained for lambdas falling in the specific plateaux range is shown; blue bullets represents healthy individuals, while red ones are individuals affected by the pathology.

The 5-clusters plauteaux, letter a) in figure 4, presents three healthy individuals clusters and two cleft lip clusters. Observing labels following bullets (not shown in the figure), one can analyze deeper the structure of the two red clusters and it is possible to appreciate how they are divided according to the presence of right- and left-sided cleft lip (see also figure 7). Proceeding to larger values of lambda, the range highlighted with letter b) indicates a region in which two of the healthy individuals clusters merge together and then merge again with both the cleft-lip clusters (letter c)). From this analysis it turns out how 2-MST unsupervised clustering is not able to unveil the inner structure of the proposed data and, trying to impose a spherical geometry, it returns more than one sub-population for the healthy individual class.

The same procedure is repeated moving to depth $D = 3$. Again the parameter λ is spanned from the minimum to the maximum of the dissimilarities, looking for the largest stability region of the algorithm. With higher clustering depths, it is found how the decay toward a single cluster plateaux is faster with respect to the spherical clustering. Thus, in order to build a reliable plateaux curve, a finer grating for lambda values is required, especially in the transition region. Moreover, at this depth it is quite common to encounter outliers, i.e. single nodes that are not assigned to any class. To this purpose, a green line is plotted as well, showing the mode of number of clusters composed by at least two nodes, in order to unveil the number of outliers present in the clustering.

Figure 5 shows the plateaux curve for depth $D = 3$. In such a case, two plateaux with no outliers are identified in the transition region between none and one cluster. Here lambda spans with a 0.002 step in order to track in detail the plateaux behaviour, while in the other regions of the curve a 0.005 step is kept. The most important parameter region, i.e. that with three clusters and indicated with letter b) in figure 5, discriminates well healthy individuals from those manifesting cleft lip. In particular, left- and right-sided cleft lips are clustered in two separated classes, while healthy individuals create a unique cluster. Then, the algorithm succeeds in identifying the investigated structure of the proposed data.

In order to understand the robustness of the clustering against the stochasticity of the algorithm, we compute the probability for any node of being assigned to its own class. Setting the value of lambda inside the plateaux, $\lambda = 0.18$, the algorithm is run fifty times, building a statistics of the

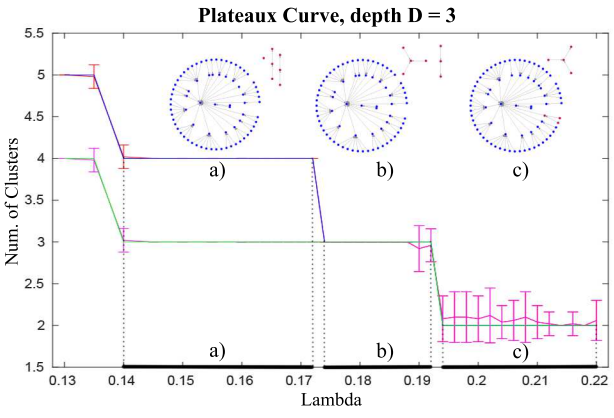

Figure 5: Number of clusters versus λ, for $D = 3$. Plateaux curve. The green solid line represents the obtained plateaux curve of the resulting clusters with at least 2 nodes. A single plateau, i.e. plateau b), is stable in converging toward a solution with no outliers and reports the discrimination of individuals with a pathology (see figure 7).

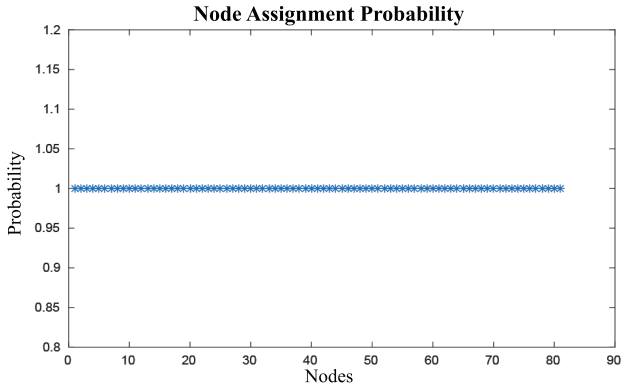

Figure 6: Assignment Probability Plot. The probability of a node to be assigned to its most frequent cluster throughout 50 runs of the clustering algorithm is plotted. No value below 1 is found.

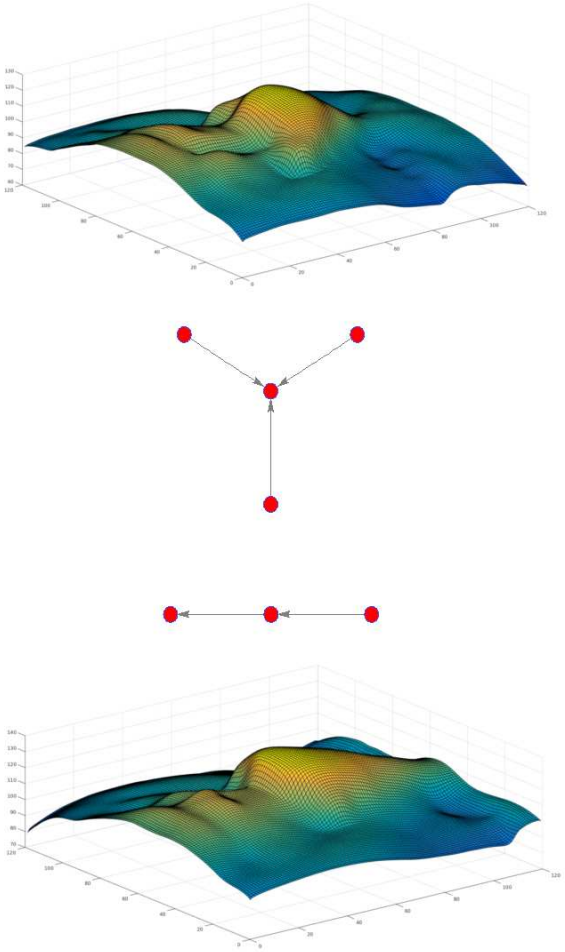

Figure 7: Resulting Clusters For $D = 3$. The 4-nodes cluster includes all the right-sided cleft lip affected individuals; an example of such individual surface is reported at the top of the figure. On the other hand, the 3-nodes cluster includes left-sided cleft lip affected individuals. An example of individual surface belonging to this class is reported at the bottom of the figure.

class assignment for any node. Plotting the node assignment probability, as shown in figure 6, it can be observed how all individuals show a probability of being member of their most frequent class equal to 1. Moreover, comparing each clustering returned in subsequent runs with the first one obtained, we can compute the overlap among clusterings, finding that the algorithm converges always to the same classifications, in terms of composition of the single clustering. In conclusion, such a clustering overlap and assignment probability allow us to state with high confidence the robustness of the classification returned.

Higher depths trees are also investigated, but results are not reported here, as they do not improve sensibly those obtained for $D = 3$. This means that, once the spherical constraint is relaxed, data do not show a longer range dependency and a three nodes correlation is able to catch the inner structure of clusters.

4 Conclusions

This work proposes an innovative automatic methodology for the diagnosis of cleft lip. The process is designed for the detection of diseases in the prenatal phase, thus its feasibility is tested upon the limitedly available ultrasound scans data and then deeply investigated for a large dataset of adult individuals. Bosphorus database is chosen for the testing, which seven cleft lip-affected individuals are added to, by artificially imposing the defect.

The algorithm maps each face with twelve differential geometry descriptors plus first and second derivatives with respect to x and y directions. It allows to determine facial landmarks, that enable face comparison. In this work, only the *pronasal* and the *labrum superior* landmarks are investigated. They are identified automatically by imposing thresholds on values expressed by a subset of geometrical descriptors. In particular, the *labrum superior* specifies the region of interest for the actual diagnosis.

In the second part of the algorithm, the geometrical descriptors expressed in the *labrum superior*'s neighbourhood are used to transform each face into a vector and create the input matrix for the unsupervised clustering algorithm. Any entrance of such built vector is used to perform the comparison between couples of individuals and compute their dissimilarity in terms of Spearman's correlation rank distance. In such a way, a squared symmetric matrix is computed and provided as input for the clustering itself.

Eventually, D-MST clustering algorithm allows to investigate regions of convergence stability to a certain number of clusters, the so called plateaux, imposing the maximum depth, i.e. the inner dependencies structure, of any cluster detected. It correctly separates left-sided and right-sided cleft lips, thus showing accurate diagnosis results.

This algorithm also opens the route for the definition of what is called *normotype*. The normotype can be considered as the representative face of a class of individuals,

collecting all the principal features distinguishing an individual as member of a particular category. The present algorithm is able to collect all healthy individuals in a single cluster, starting from features expressed in the lip region, in comparison to those manifesting a cleft lip. This allows a formalization of the normotype features. On the other hand, once the cleft lip population will account for a sufficient number of members, the labio-schisis normotype would be defined as well.

Other syndromes, like the Fetal Alchol Syndrome (FAS) or the palato-schisis syndrome, will also be investigated and geometrically formalized to be embedded in this algorithm, which could be a multi-syndrome diagnosing tool.

References

[1] T. S. Douglas, T. E. M. Mutsvangwa, "A Review of Facial Image Analysis for Delineation of the Facial Phenotype Associated With Fetal Alcohol Syndrome," *Am J Med Genet Part A*, vol. 152, no. 2, pp. 528536, 2010.

[2] S. Campbell, C. Lees, G. Moscoso, P. Hall, "Ultrasound antenatal diagnosis of cleft palate by a new technique: the 3D reverse face view," *Ultrasound Obstet Gynecol*, vol. 25, no. 1, pp. 1218, 2005.

[3] I. Aras, S. Olmez, S. Dogan, "Comparative Evaluation of Nasopharyngeal Airways of Unilateral Cleft Lip and Palate Patients Using Three-Dimensional and Two-Dimensional Methods," *The Cleft Palate-Craniofacial Journal*, vol. 49, no. 6, pp. e75e81, 2012.

[4] L. D. Platt, G. R. DeVore, D. H. Pretorius, "Improving Cleft Palate/Cleft Lip Antenatal Diagnosis by 3-Dimensional Sonography," *J Ultrasound Med*, vol. 25, no. 11, pp. 14231430, 2006.

[5] G. Tonni, M. Lituania, "OmniView Algorithm, A Novel 3-Dimensional Sonographic Technique in the Study of the Fetal Hard and Soft Palates," *J Ultrasound Med*, vol. 31, no. 2, pp. 313318, 2012.

[6] P. A. Mossey, J. Little, R. G. Munger, M. J. Dixon, W. C. Shaw, "Cleft lip and palate," *The Lancet*, vol. 374, no. 9703, pp. 1773-1785, 2009.

[7] L. Bonacina, D. Conti, A. Froio, F. Marcolin, E. Vezzetti, "Automatic 3D Fetus Face Model Extraction From Ultrasonography Through Histogram Processing", submitted to *IEEE Trans Biomed Eng*.

[8] A. K. Jain, M. N. Murty, P. J. Flinn, "Data Clustering: A Review", *ACM Computing Surveys*, vol. 31, no. 3, 1999.

[9] B. J. Frey, D. Dueck, "Clustering by Passing Messages Between Data Points," *Science*, vol. 315, no. 972, 2007.

[10] J.H. Jr. Ward, "Hierarchical grouping to optimize an objective function," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 236-244, 1963.

[11] M. Bailly-Bechet, S. Bradde, A. Braunstein, A. Flaxman, L. Foini, R. Zecchina, "Clustering with shallows trees," *Journal of Statistical Mechanics: Theory and Experiments*, vol. 2009, doi:10.1088/1742-5468/2009/12/P12010, 2009.

[12] M. Bailly-Bechet, C. Borgs, A. Braunstein, J. Chayes, A. Dagkessamanskaia, J.-M. Franois, and R. Zecchina. "Finding undetected protein associations in cell signaling by belief propagation," *Proceedings of the National Academy of Sciences*. 2010; 108(2):882-7, doi:10.1073/pnas.1004751108

[13] The Bosphorus Database Official Web Site: *http://bosphorus.ee.boun.edu.tr/default.aspx*

[14] E. Vezzetti, F. Marcolin, "Geometrical descriptors for human face morphological analysis and recognition," *Robotics and Autonomous Systems*, vol. 60, no. 6, pp. 928-939, 2012.

[15] J. J. Koenderick, A. J. van Doorn, "Surface shape and curvature scales", *Image and Vision Computing*, vol. 10, no. 8, pp. 557-564, 1992.

[16] E. Vezzetti, D. Speranza, F. Marcolin, G. Fracastoro, G. Buscicchio, "Exploiting 3D Ultrasound for Fetal Diagnostic Purpose Through Facial Landmarking," *Image Anal Stereol*, vol. 33, no. 3, pp. 167-188, 2014.

[17] E. Vezzetti, D. Speranza, F. Marcolin, G. Fracastoro, "Diagnosing cleft lip pathology in 3d ultrasound: A landmarking-based approach," *Image Anal Stereol*, vol. 35, no. 1, pp. 53-65, 2015.

[18] S. Moos, F. Marcolin, S. Tornincasa, E. Vezzetti, M. G. Violante, G. Fracastoro, D. Speranza, F. Padula, "Cleft lip pathology diagnosis and foetal landmark extraction via 3D geometrical analysis," *International Journal on Interactive Design and Manufacturing (IJIDeM)*, vol. 11, no. 1, pp. 1-18, 2017.

[19] E. Vezzetti, F. Marcolin, "Geometry-based 3D face morphology analysis: soft-tissue landmark formalization," *Multimed Tools Appl*, vol. 68, no. 3, pp. 895-929, 2012.

[20] E. Vezzetti, F. Marcolin, "3D human face description: landmarks measures and geometrical features," *Image and Vision Computing*, vol. 30, no. 10, pp. 698-712, 2012.

[21] E. Vezzetti, S. Moos, F. Marcolin, V. Stola, "A pose-independent method for 3D face landmark formalization," *Computer Methods and Programs in Biomedicine*, vol. 198, no. 3, pp. 1078-1096, 2012.

[22] E. Vezzetti, F. Marcolin, V. Stola, "3D Human Face Soft Tissues Landmarking Method: An Advanced Approach," *Computers in Industry*, vol. 64, no. 9, pp. 1326-1354, 2013.

[23] E. Vezzetti, F. Marcolin, "3D Landmarking in Multiexpression Face Analysis: A Preliminary Study on Eyebrows and Mouth," *Aesthetic Plastic Surgery*, vol. 38, pp. 796-811, 2014.

[24] R. Xu, C. Donald, *Clustering*, Wiley-IEEE Press, November 2008.

[25] M. Bayati, A. Braunstein, R. Zecchina, "A rigorous analysis of the cavity equations for the minimum spanning tree," *Journal of mathematical physics*, vol. 49, no. 12, 2008.

[26] M. Bayati *et al.*, "Statistical mechanics of Steiner trees," *Physical Review Letters*, vol 101, no. 3, 2008.