3D Human Face Soft Tissues Landmarking Method: An Advanced Approach

Original
3D Human Face Soft Tissues Landmarking Method: An Advanced Approach / Vezzetti, E., Marcolin, F., Vincenzo, S.. - In: COMPUTERS IN INDUSTRY. - ISSN 0166-3615. - (2013). [10.1016/j.compind.2013.04.006]

Publisher:
Elsevier

Published
DOI:10.1016/j.compind.2013.04.006

3D Human Face Soft Tissues Landmarking: An Advanced Approach

Enrico Vezzetti[1]*, Federica Marcolin, Vincenzo Stola*
Dipartimento di Ingegneria Gestionale e della Produzione
Politecnico di Torino, Italy

Abstract

Face study and description through landmarks is a key activity in many fields, in particular the medical one for both corrective and aesthetic surgery purposes. In a previous work, we used a geometric approach to automatically extract landmarks working in the Differential Geometry domain, through the coefficients of the Fundamental Forms, the Principal Curvatures, Mean and Gaussian Curvatures and also with derivatives and Shape and Curvedness Indexes. In this paper we describe the improvements made to our algorithm previously developed by using a new parameterization of the mesh, new geometrical descriptors, and new conditions.

1. Introduction

Thanks to the strong development of 3D acquisition systems, three-dimensional face description has become more attractive, as the capture of 3D faces is faster and cheaper than in the past. One of the methods used for studying faces is the one involving the extraction of key points shared by all faces and with a particular biological meaning: they are called **landmarks**. In human face fifty-nine landmarks could be collected, but the most famous ones are nearly twenty. Some algorithms identify faces by extracting landmarks, or features, from an image of the subject's face. A possible algorithm may analyze the relative position, size and shape of the eyes, nose, cheekbones, and jaw. Then, these features are used for various applications, including maxillo-facial surgery. In fact,

---

[1] Corresponding Author.
E-mail address: enrico.vezzetti@polito.it (E. Vezzetti).

landmarks help surgeons in studying the patient's face, especially in the pre-operation phase for predicting the actions of the surgery and for better stating how to intervene.

Much research has been carried out on this topic. In their various publications, Alker *et al.*, Frantz *et al.*, and Wörz *et al.* proposed multi-step differential procedures for subvoxel localization of 3D point landmarks, addressing the problem of choosing an optimal size for a region-of-interest (ROI) around point landmarks [1] [2]. They introduced an approach for the localization of 3D anatomical point landmarks based on deformable models. To model the surface as a landmark, they used quadric surfaces combined with global deformations [3] [4]. They proposed a method based on 3D parametric intensity models, which are directly fitted to 3D images, and introduced an analytic intensity model based on the Gaussian error function in conjunction with 3D rigid transformations and deformations in order to efficiently model anatomical structures [5]. Finally these researchers introduced a new multi-step approach to improve detection of 3D anatomical point landmarks in tomographic images [6].

Romero *et al.* presented a comparison of several approaches that use graph matching and cascade filtering for landmark localization in 3D face data. For the first method, they apply the structural graph matching algorithm relaxation-by-elimination using a simple distance-to-local-plane node property and a Euclidean-distance arc property. After the graph matching process has eliminated unlikely candidates, the most likely triplet is selected, by exhaustive search, as the minimum Mahalanobis distance over a six dimensional space, which corresponds to three node variables and three arc variables. A second method uses state-of-the-art pose-invariant feature descriptors embedded into a cascade filter to localize the nose tip. After that, local graph matching is applied to localize the inner eye corners [7]. Then, it describes and evaluates their pose-invariant point-pair descriptors, which encode 3D shape between a pair of 3D points. Two variants of descriptor are introduced: the first one is the point-pair spin image, which is related to the classical spin image of Johnson and Hebert, and the second one is derived from an implicit radial basis function (RBF) model of the facial surface. These descriptors can really encode edges in graph based representations of 3D shapes. Here they show how the descriptors are able to identify the nose-tip and the eye-corner of a human face simultaneously in six promising landmark localisation systems [8].

Bevilacqua *et al.* [9] presented a method for landmarks detection in which landmarks are detected by the use of a "sliding vector" (an observation window that at every step “scrolls” along the particular direction of movement) on the polygonal mesh to search geometrical-statistic features typical of the face.

Ruiz *et al.* [10] presented an algorithm for automatic localization of landmarks on 3D faces. An Active Shape Model (ASM) is used as a statistical joint location model for configurations of facial features. The ASM is adapted to individual faces through a guided search, whereby landmark specific Shape Index models are matched to local surface patches. Similarly, Sang-Jun *et al.* [11] applied the Active Shape Models to extract the position of the eyes, the nose, and the mouth. Nair and Cavallaro [12] presented a method based on a statistical model called Point Distribution Model (PDM) to detect the landmarks. The PDM is similar to the Active Shape Model (ASM), but it uses the points coordinates instead of the colour of an image used by ASM. Firstly, a set of candidate points are detected through the use of the Curvedness Index; subsequently, a PDM is used to detect the landmarks; finally, the face registration is performed using the IPC algorithm.

Salah *et al.* [13] proposed a coarse-to-fme method for facial landmark localization: it relies on unsupervised modelling of landmark features obtained through different Gabor filter channels.

D’Hose *et al.* [14] presented a method for localization of landmarks on 3D faces: it uses Gabor wavelets to extract the curvature of the 3D faces, which is then used for performing a coarse detection of landmarks.

In our previous work [15] we presented a first method to detect nine landmarks (*pronasal*, *nasion*, *subnasal*, *alae*, *endocanthions*, *exocanthions*) using geometrical descriptors. In this paper we propose an improvement to the previous algorithm, in order to have more precise measurements for a greater number of faces. This work is based on a geometric formalization of the face which may be used not only in the medical field, but also in other applications such as face detection, face recognition, authentication, and facial expression recognition.

2. The proposed method

A facial landmark is a point which all faces share and has a particular biological meaning. In particular we may distinct two types of landmarks:

1. **hard-tissue** landmarks, which lie on the skeletal and may be identified only through lateral cephalometric radiographs;
2. **soft-tissue** landmarks, which are on the skin and can be identified on the point clouds generated by the scanning.

In this paper we considered only soft-tissue landmarks, since a radiograph is more invasive (and harmful) than a photogrammetric acquisition system. Although soft-tissue landmarks are nearly fifty-nine, in this paper we considered nine identifiable ones (*pronasal*, *nasion*, *subnasal*, *alae*, *endocanthions*, *exocanthions*) plus other assistant points which better define nose and eyes shapes, as shown in Figure 1. The landmarks close to the mouth are not taken into consideration due to their pose-dependency, while the ones close to the face boundaries have been ignored because in those zones the scan is not accurate.

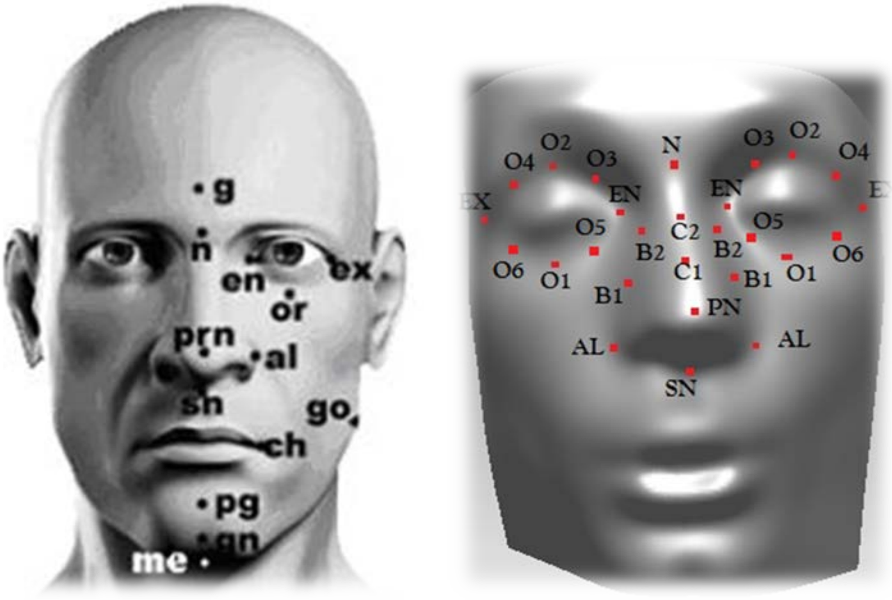


Figure 1. **(left)** The Anthropometric soft-tissue landmarks (g-*glabella*, n-*nasion*, en-*endocanthion*, ex-*exocanthion*, or-*orbital*, prn-*pronasal*, sn-*subnasal*, al-*alae*, ch-*cheilion*, pg-*pogonion*, gn-*gnathion*, go-*gonion*, me-*menton*) [16]; **(right)** landmarks detected by our proposed method (PN-*pronasal*, SN-*subnasal*, AL-*alae*, N-*nasion*, EN-*endocanthions*, EX-*exocanthions*, C1 and C2-*ridge of nose*, B1 and B2-*base of nose*, from O1 to O6-*boundaries of eye*).

The geometrical analysis is based on a mesh surface that approximates the point cloud, therefore a parameterization is necessary; in our previous work we used the following representation to describe the surface:

$$S = \begin{cases} x = X \\ y = Y \\ z = f(X, Y) \end{cases},$$

but with this form some analytical problems rose (e.g., with this form, the points which have the same x and y-coordinates are not taken). As a result, we interpolated the shell points with a parametric function, that has been fitted to the mesh using the parameterization:

$$S = \begin{cases} x = X(v, u) \\ y = Y(v, u) \\ z = Z(v, u) \end{cases}, \quad 1 \leq u, v \leq 150,$$

which solves the previous analytical problems. The reference system used has the following features, as shown in Figure 2:

1. the x-axis is horizontal, from left to right;
2. the y-axis is vertical, from bottom to top;
3. the z-axis is outgoing from xy-plane.

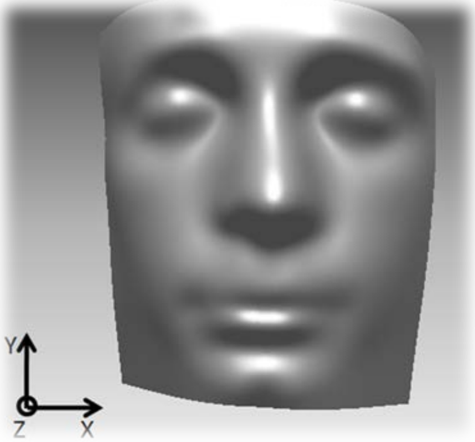


Figure 2. The reference system used in this work.

The u and v-parameters are the parametric directions of the surface; as shown in Figure 3, the u-direction is nearly horizontal, while the v-direction is nearly vertical. Keeping constant one of them, we obtained some curves on the surface; in particular, keeping constant the u-parameter we obtain the curves at u-constant, while keeping constant the v-parameter we obtain the curves at v-constant. These curves are will be used to detect some landmarks.

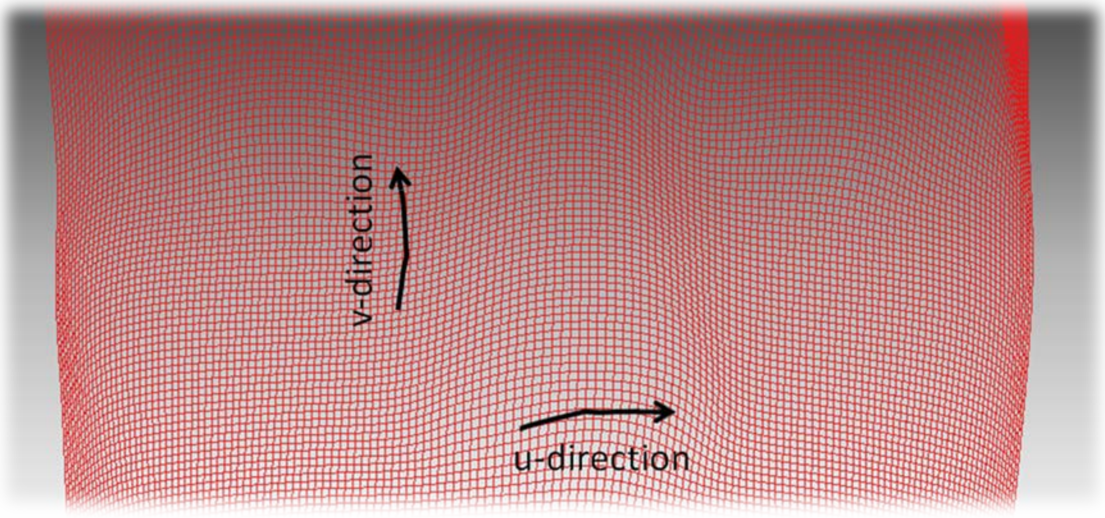


Figure 3. Detail of the face parametric surface (eyes region); the red grid indicates the parametric coordinates: the $\boldsymbol{u}$-direction is nearly horizontal, while the $\boldsymbol{v}$-direction is nearly vertical.

For studying the features of the surface the differential analysis may be used; in particular, the first, the second and the mixed derivates, the coefficients of the Fundamental Forms, the curvatures, Shape and Curvedness Indexes, and Tangent Map are used. The First and the Second Fundamental Forms are employed to measure the distance on surfaces and are defined by

$$Edu^2 + 2Fdudv + Gdv^2,$$

$$edu^2 + 2fdudv + gdv^2,$$

respectively, where E, F, G, e, f, and g are their coefficients and are calculated by the following formulas:

$$E = \|D_u\|^2,$$

$$F = \langle D_u, D_v \rangle,$$

$$G = \|D_v\|^2,$$

$$e = \langle N, D_{uu} \rangle,$$

$$f = \langle N, D_{uv} \rangle,$$

$$g = \langle N, D_{vv} \rangle,$$

where

$$D_u = \begin{cases} X_u = \frac{\partial X(v,u)}{\partial u} \\ Y_u = \frac{\partial Y(v,u)}{\partial u}, \\ Z_u = \frac{\partial Z(v,u)}{\partial u} \end{cases}$$

$$D_v = \begin{cases} X_v = \frac{\partial X(v,u)}{\partial v} \\ Y_v = \frac{\partial Y(v,u)}{\partial v}, \\ Z_v = \frac{\partial Z(v,u)}{\partial v} \end{cases}$$

$$D_{uu} = \frac{\partial Du}{\partial u}, D_{uv} = \frac{\partial Du}{\partial v}, D_{uv} = \frac{\partial Dv}{\partial v}, N = \frac{D_u \times D_v}{|D_u \times D_v|}.$$

Curvatures are used to measure how a regular surface bends in R^3. If D is the differential and N is the normal plane of a surface, then the determinant of DN will be the product of the Principal Curvatures ($det(DN) = (-k_1)(-k_2) = k_1k_2$), and the trace of DN will be the negative of the sum of Principal Curvature ($trace(DN) = -(k_1 + k_2)$). In the point P, the determinant of DN_P is the **Gaussian Curvature** (K) at P, while the negative of half of the DN trace is called the **Mean Curvature** (H) at P. In terms of the principal curvatures it can be written:

$K = k_1k_2,$

$H = \frac{k_1+k_2}{2},$

where k_1 and k_2 are the **Principal Curvatures**. Starting from the coefficients of the Fundamental Forms, we may calculate the Gaussian and Mean Curvatures with the following formulas:

$K = \frac{eg-f^2}{EG-F^2},$

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

Obtained the Gaussian and Mean Curvatures we may calculate the Principal Curvatures in this way:

$k_1 = H + \sqrt[2]{H^2 - K},$

$k_2 = H - \sqrt[2]{H^2 - K}.$

The most used descriptors are surely the Shape (S) and Curvedness (C) Indexes, introduced by Koenderink *et al.* [17]:

$$S = -\frac{2}{\pi}\tan^{-1}\frac{k_1+k_2}{k_1-k_2}, \qquad S \in [-1,1], \qquad k_1 \geq k_2,$$

$$C = \sqrt[2]{\frac{{k_1}^2+{k_2}^2}{2}}.$$

Their meaning is shown in Figure 4, 5, 6, and in Table 1.

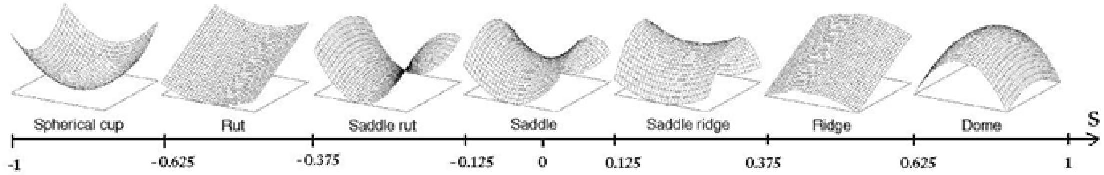


Figure 4. Illustration of Shape Index scale divided into seven categories. Different subintervals of its range [-1,1] correspond to seven geometric surfaces.

Class	S	Type	H	K
Cup/Pit	$[-1; -0.625)$	Elliptical convex	+	+
Rut/Valley	$[-0.625; -0.375)$	Cylindrical convex	+	0
Saddle rut/Saddle valley	$[-0.375; -0.125)$	Hyperbolic convex	+	-
Saddle	$[-0.125; 0.125)$	Hyperbolic symmetric	0	-
Saddle ridge	$[0.125; 0.375)$	Hyperbolic concave	-	-
Ridge	$[0.375; 0.625)$	Cylindrical concave	-	0
Dome/Peak	$[0.625; 1]$	Elliptical concave	-	+

Table 1. Topographic classes [16]. S is the Shape Index, H and K are the Mean and Gaussian Curvatures, respectively. The first line means that if a surface lies in the class of cup (namely its shape is comparable to a pit or a cup), its Shape Index will be in the range between -1 and -0.625; it will also have an elliptical convexity and both the Mean and Gaussian Curvatures will be positive. Similarly for the other lines of the table.

Figure 5. Curvedness Index scale, whose range is $(-\infty; +\infty)$.

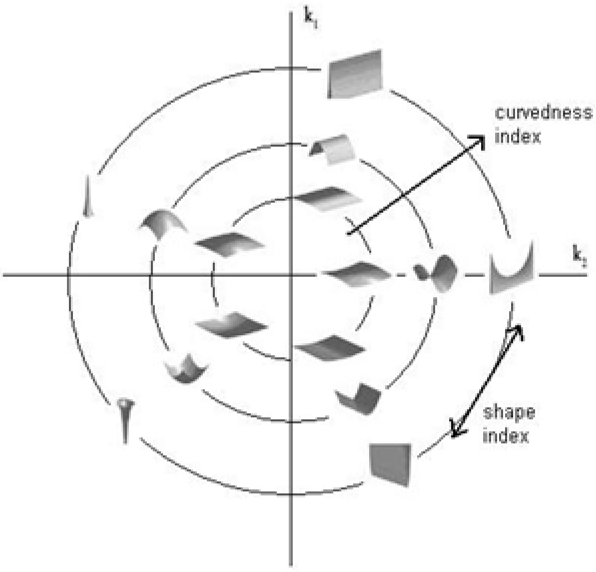


Figure 6. The effects on surface structure from variations in the Curvedness (radial coordinate) and Shape Index (angular coordinate) parameters of curvature, and the relation of these components to the Principal Curvatures ($\boldsymbol{k_1}$ and $\boldsymbol{k_2}$). Indexes ($\boldsymbol{S}$, $\boldsymbol{C}$) are viewed as polar coordinates in the $\boldsymbol{k_1 k_2}$-plane, with planar points mapped to the origin. The degree of curvature increases radially from the centre [16].

The Tangent Map is an index used by Perakis *et al.* [18] to detect the points which have the normal outward with respect to the centroid of the surface like nose and chin regions. The Tangent Map is calculated by the following formula:

$T(P) = \langle N(P), R(P) \rangle$,

where N is the normal of the surface at P and R is the straight line passing through the centroid of the surface and P.

The detection process consists in two phases: firstly, the algorithm defines the region of interest of a single landmark through geometric constrains; secondly, it detects the landmarks through a research of the critical points of an appropriate descriptor.

Indeed, the critical point positions depend by the face orientation, therefore it is necessary to rotate the face in a standard pose. It is represented by the following angles:

$$\theta_{x\_ref} = -0.467384745634034\, rad,$$

$$\theta_{y\_ref} = 0\, rad,$$

where θ_{x_ref} indicates the rotation angle around x-axis, so that the straight line through *pronasal* and *nasion* lies on xy-plane, while θ_{y_ref} indicates the rotation angle around y-axis so that the straight line through two forehead points equidistant from *nasion* lies on xy-plane, as shown in Figure 7.

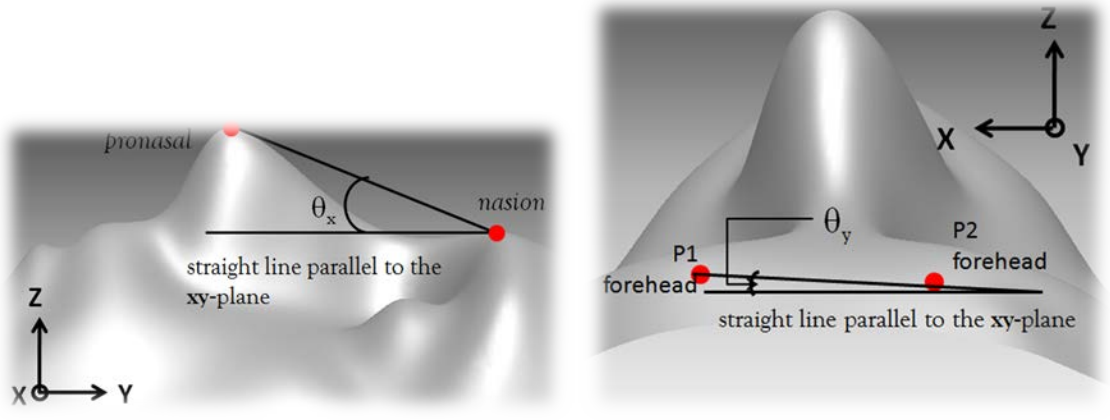


Figure 7. Representation of the angles $\boldsymbol{\theta_x}$ and $\boldsymbol{\theta_y}$.

Firstly, the function detects *pronasal*, *nasion*, and two forehead points equidistant from *nasion* in this way:

$$p1 = \begin{cases} u = u_n - 15 \\ v = v_n + 10 \end{cases},$$

$$p2 = \begin{cases} u = u_n + 15 \\ v = x_n + 10 \end{cases}.$$

Next, the function estimates the face angles θ_x and θ_y and finally it rotates the face so that the new angles θ_x and θ_y are respectively equal to θ_{x_ref} and θ_{y_ref}. Since θ_x and θ_y are estimated separately, the new angles θ_x and θ_y will not be equal to reference angles, therefore the function repeats the entire procedure five times. After five iterations the differences between the face angles and reference angles are very small, therefore negligible.

The rotation angle around z-axis is not considered for the standard pose because the estimate of this angle is not very precise, therefore the face after the rotation may be in a wrong pose. However, this angle is used to detect some landmarks, so its estimate is necessary. Firstly, we consider a point (C) that lies on the ridge of the nose, as shown in Figure 8.

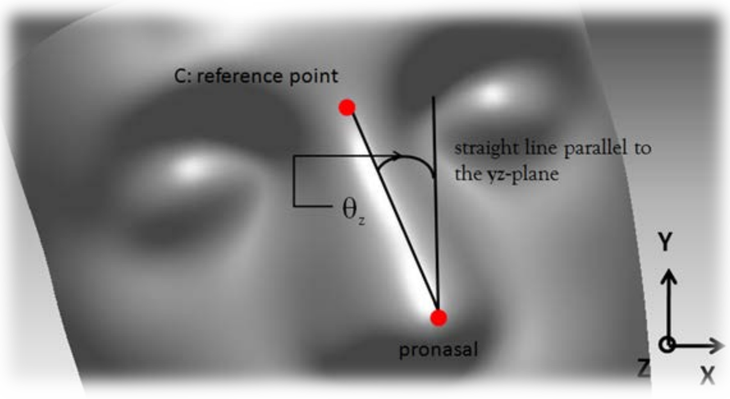


Figure 8. Representation of the angle $\boldsymbol{\theta_z}$.

The coordinates of this point are obtained in this way:

1. the v-coordinate (v_c) is calculated by the formula: $v_c = v_{PN} + 20$;
2. the u-coordinate (u_c) is found by searching for a local maximum along u-direction (with v-coordinate equal to v_c).

Secondly, we calculate the rotation angle by the following formula:

$$\theta_z = -\tan^{-1}\frac{x_{pn} - x_c}{y_{pn} - y_c},$$

where

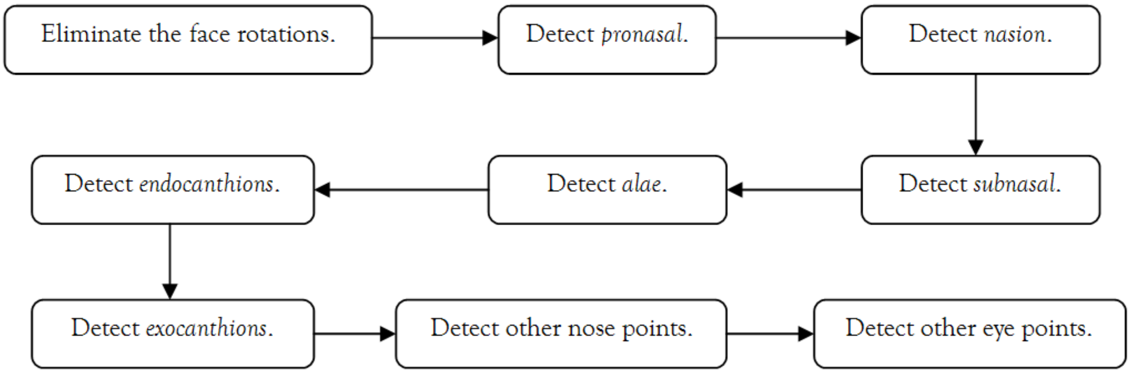
$x_{pn} = \boldsymbol{X}(v_{pn}, u_{pn})$,

$y_{pn} = \boldsymbol{Y}(v_{pn}, u_{pn})$,

$x_c = \boldsymbol{X}(v_c, u_c)$,

$y_c = \boldsymbol{Y}(v_c, u_c)$.

When the process of the face rotation ends, the algorithm starts the landmark detection phase, summarized in the following scheme.

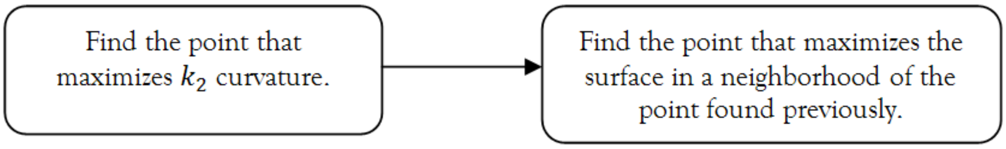


2.1 Pronasal

The *pronasal* (PN) is the point on the nose tip and surely is the point most easily identifiable by human eye, especially because it is the most salient when the face is well oriented. In a standard face, the region of nose is characterized by following features:

1. values of k_2 curvature is higher compared to the other regions;
2. in our reference system the nose tip is an local maximum, so it is a critical point.

The idea consists in maximizing k_2 curvature, so the system finds the *pronasal* region and then the algorithm searches the maximum of the surface in a neighborhood of the point found.



In most of the faces, the point that maximizes k_2 curvature is the nose tip, but in particular cases (for example hooked noses or noses with large nostrils) this is not true, therefore the research of the local maximum becomes necessary.

2.2 Nasion

The search of the *nasion* (N) is more complex than the *pronasal* one because the region surrounding the *nasion* has some geometric features present in other areas of the face, therefore the reduction of the region to be analyzed is necessary. We could analyze the face profile, shown in Figure 9, which is represented as a bi-dimensional function; in particular, the abscissa indicates v-parameter, while the ordinate indicates z-coordinate of the face. For each v-parameter, the value of the face profile is equal to the maximum of the face surface by varying the u-parameter.

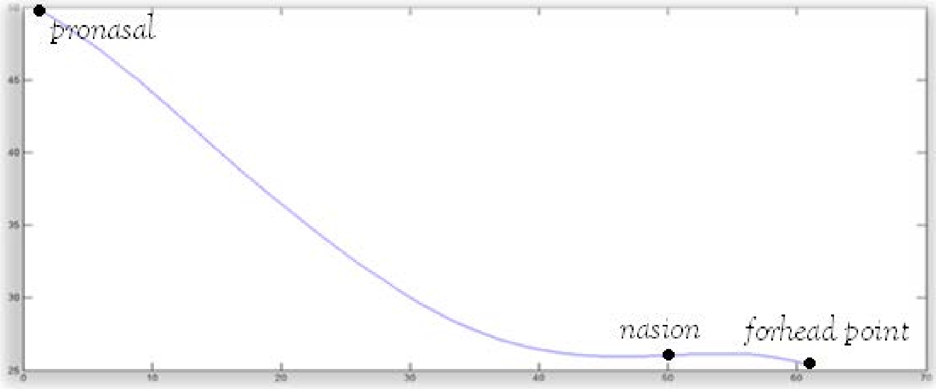


Figure 9. Profile of the region overlying the *pronasal*. The abscissa indicates $\boldsymbol{v}$-parameter, while the ordinate indicates the $\boldsymbol{z}$-coordinate of the face. In particular, the region subtended by the function is a portion of the nose. In $\boldsymbol{v} = \mathbf{0}$ we have the $\boldsymbol{z}$-coordinate of the *pronasal*, while the $\boldsymbol{v}$ maximum value corresponds to the $\boldsymbol{z}$-coordinate of the forehead point.

To facilitate the search of the *nasion*, the profile is extracted only in the region overlying the *pronasal*, therefore the value $v = 0$ is the nose tip, while the maximum v-value indicates a forehead point, as shown in Figure 9. In a standard face, the region of *nasion* is characterized by the following features:

1. high values of the face profile derivative, as in Figure 10;
2. the *nasion* is a critical point;
3. Shape Index lies in the range corresponding to the surface of saddle ridge, saddle point, saddle rut, or rut: $-0.625 \leq S < 0.375$;
4. the Gaussian Curvature is less than or equal to zero: $K \leq 0$;
5. the Mean Curvature lies in the range between -0.5 and 0.5: $-0.5 < H < 0.5$;
6. the coefficient *f* lies in the range between -0.1 and 0.1: $-0.1 < f < 0.1$;
7. low values of the coefficient g;
8. the *nasion* is a saddle point in the Tangent Map, as shown in Figure 11.

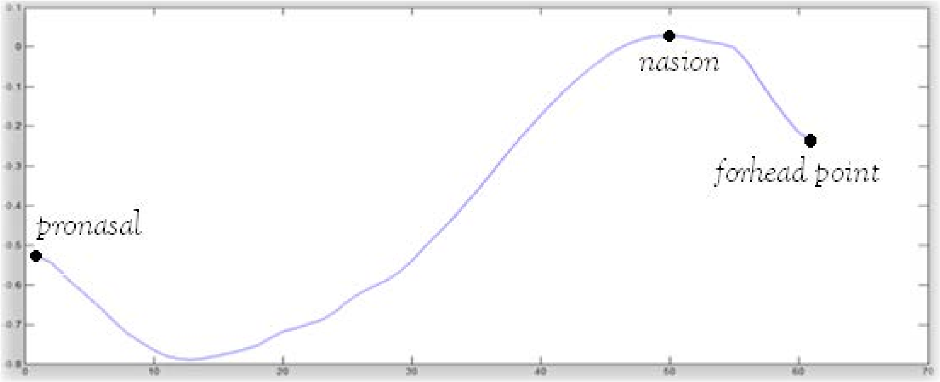


Figure 10. Derivative of the profile of the region overlying the *pronasal*.

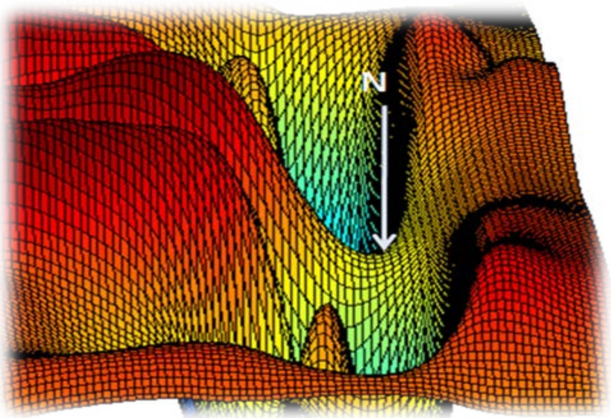


Figure 11. Particular of the tangent map in proximity of the *nasion* (N).

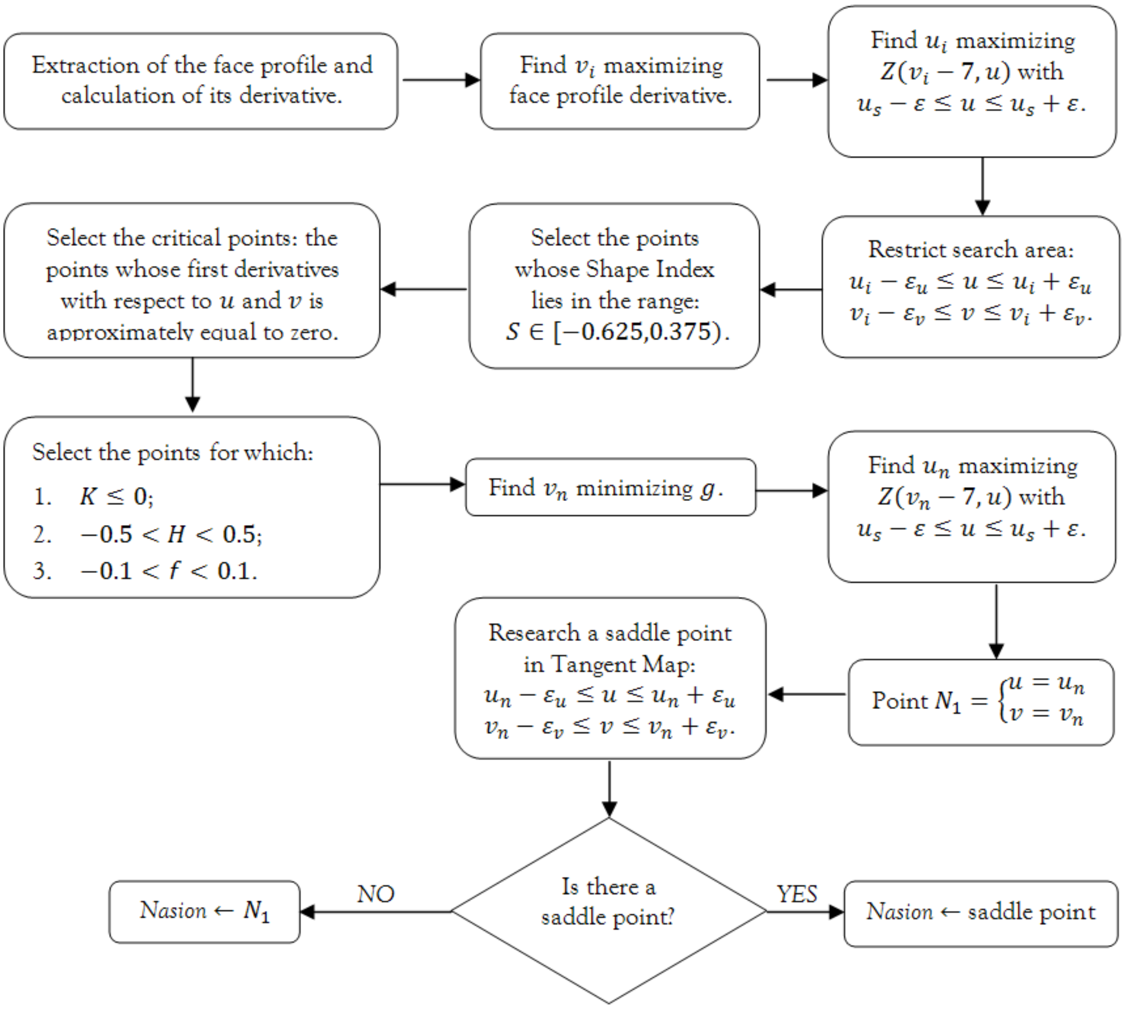
Firstly, the algorithm reduces the region of interest through the analysis of the face profile derivative: maximizing the derivative, the initial v-coordinate (v_i) is found; then the initial u-coordinate (u_i) is estimated maximizing $Z(v,u)$ keeping constant the initial v-coordinate and varying the u-coordinate in a neighbourhood of u_s coordinate calculated by the formula:

$$u_s = u_{pn} - tan\,\theta_z * (v_i - v_{pn}).$$

Since the forehead is a nearly flat region, the presence of noise and small rotations of the face may be due to an incorrect estimate of the u_i parameter; consequently, the maximization of $Z(v,u)$ is performed keeping constant v-parameter to the value $v_i - 7$ instead of v_i (we assume that the θ_z doesn't affect the u_i parameter because the distance between $v_i - 7$ and v_i is very small), since the curve $Z(v_i - 7, u)$ intersects the ridge of nose, which is a region more robust to noise than the forehead.

Secondly, the algorithm considers a set of points that are located in a neighbourhood of the point identified by parametric coordinates (u_i and v_i) previously found. These points are filtered by conditions 2, 3, 4, 5, and 6, while minimizing the coefficient g the algorithm finds a point (u_n and v_n) which could be the *nasion*. Indeed, this research has the same problem as previous research, so the estimate of the u_n parameter could be incorrect. As a result of this, the algorithm estimates the u_n parameter with the same method described above (in this case v_i is equal to v_n).

Finally, the algorithm searches a saddle point in the Tangent Map in a neighbourhood of the point identified by coordinates parametric (u_n and v_n) previously found. If the research produces no result then the algorithm considers as *nasion* the point obtained with the previous research. In fact, the Tangent Map depends on the head pose, therefore, if the face is not in a standard pose, the point found could be completely wrong; however, if the face is in a standard pose, the point found will be the *nasion*.

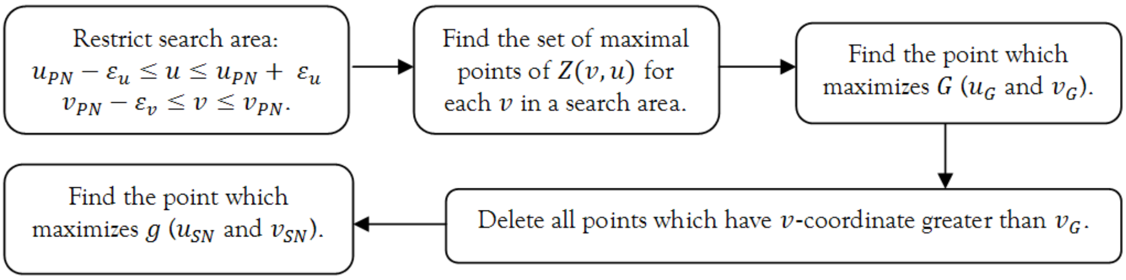


2.3 Subnasal

The *subnasal* (SN) is the point which lies exactly below the nose, in that dimple above the mouth. This point is characterized by:

1. high values in g and G;
2. maximum in $Z(v, u)$ in the u-direction.

Firstly, the algorithm finds the set of maximal points of $Z(v, u)$ for each v in a region underlying the *pronasal*; next, it considers the point which maximizes G (u_G and v_G) among points previously found and deletes from the set the points which have v-coordinate greater than v_G. Finally, among remaining points, the *subnasal* corresponds to the point that maximizes g.



2.4 Alae

The *alae* (AL) are the two points which lie on the left and the right of the widest part of the nose: therefore their distance is exactly the nose width. These points have a particular behaviour in Z; in fact, each point has a local maximum in the v-direction, as shown in Figure 12.

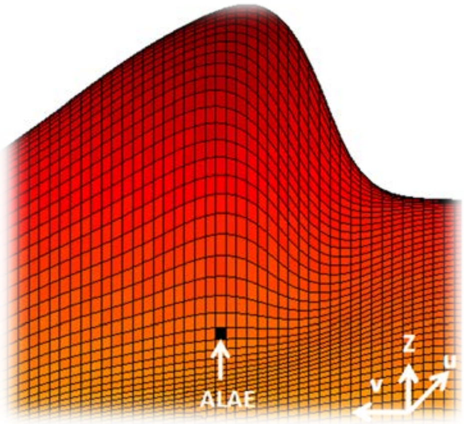
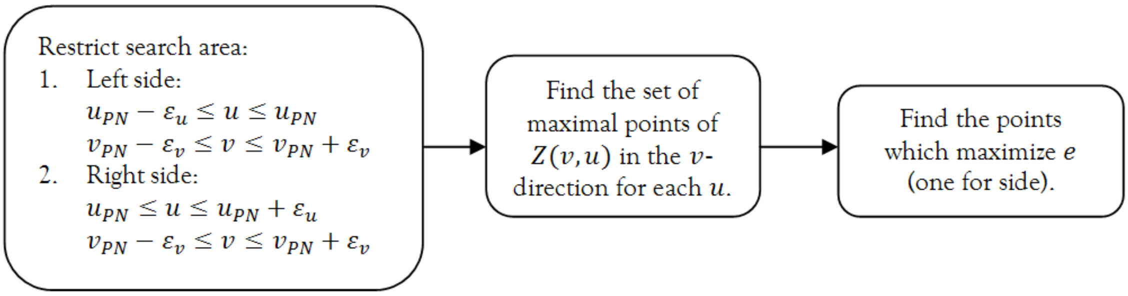


Figure 7. Detail of the left *alae*. The horizontal lines are the curves at $\boldsymbol{u}$-constant, while the vertical lines are the curves at $\boldsymbol{v}$-constant. We can note that the curve at $\boldsymbol{u}$-constant passing through the *alae* has a local maximum in it.

In summary:

1. the coefficient e has two local maximums in them;
2. $Z(v,u)$ has two local maximums in them in the v-direction.

Firstly, the algorithm finds the local maximums in the v-direction for each u-parameter smaller than u_{PN},for the left side, or greater than u_{PN}, for the right side. Then, the algorithm chooses as *alae* the points which maximise the coefficient e, one for side.



Restrict search area:
1. Left side:
   $u_{PN} - \varepsilon_u \leq u \leq u_{PN}$
   $v_{PN} - \varepsilon_v \leq v \leq v_{PN} + \varepsilon_v$
2. Right side:
   $u_{PN} \leq u \leq u_{PN} + \varepsilon_u$
   $v_{PN} - \varepsilon_v \leq v \leq v_{PN} + \varepsilon_v$

Find the set of maximal points of $Z(v,u)$ in the v-direction for each u.

Find the points which maximize e (one for side).

2.5 Endocanthions

The *endocanthions* (EN) are the two points at which the inner ends of the upper and lower eyelid meet. Since the upper and lower eyelid cannot be well defined in the mesh, the extraction of the *endocanthions* is not trivial. Therefore in this work the algorithm detects the two critical points, one for side, located between eye and nose, namely the inner hollows near the eyes. Since these points are very close to the real *endocanthions*, we call *endocanthions* the points detected. In a human face, these critical points are local minimums, but in a mesh this could be false (there is always a local minimum in the u-direction in Z but there could be a local minimum in the v-direction in Z) due to the noise introduced by the scanner or the excessive use of filter to reduce the noise. However, we may consider the local minimums in the v-direction in Z in the nose area, as shown in Figure 13.

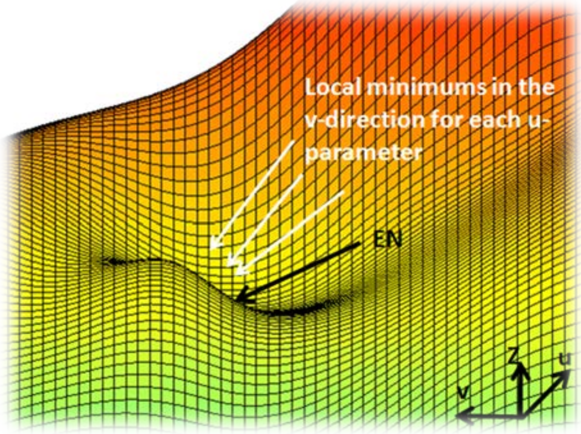
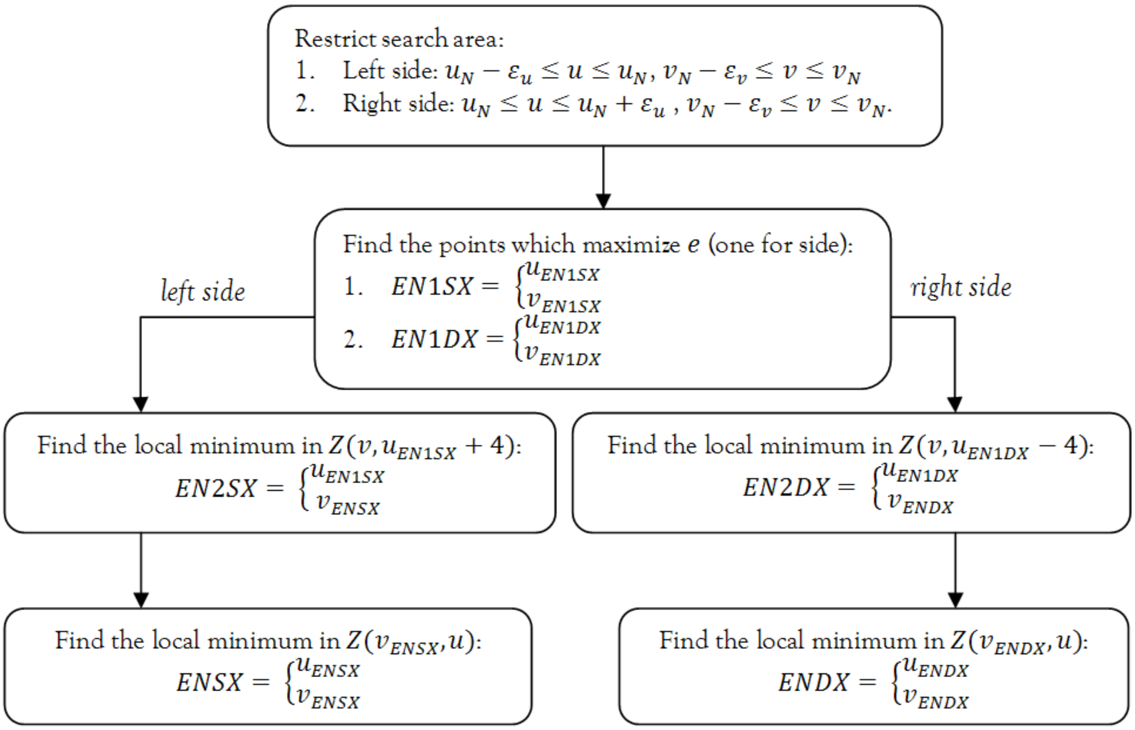


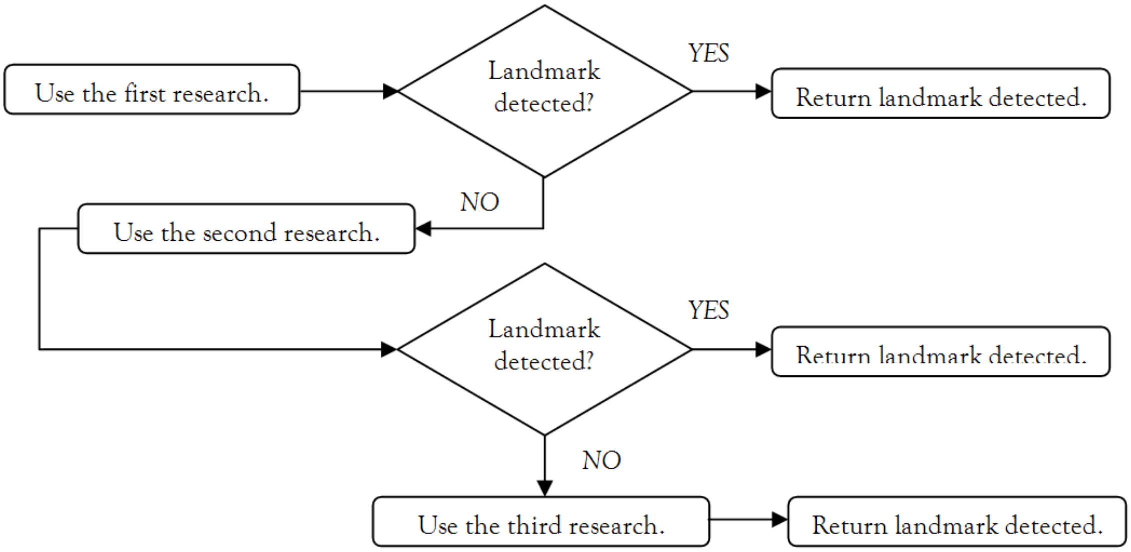
Figure 8. Detail of the *endocanthion* region. The horizontal lines are the curves at u-constant, while the vertical lines are the curves at v-constant. The *endocanthion* is behind the eyelid, but we can note that its v-coordinate corresponds to the local minimums of the curves at u-constant located on the side of the nose.

Firstly, the algorithm finds two local minimums of the coefficient e, one for side, in the nose area; in this way, two points ($EN1SX$ and $EN1DX$) lying on the nose bottom are found. Then, the algorithm searches the local minimums: in particular, the v-coordinates (v_{ENSX} and v_{ENDX}) are detected searching the local minimums in $Z(v, u_{EN1SX} + 4)$ and $Z(v, u_{EN1DX} - 4)$, while the u-coordinates are detected searching the local minimums in $Z(v_{ENSX}, u)$ and $Z(v_{ENDX}, u)$.



2.6 Exocanthions

The *exocanthions* (EX) are the two points at which the outer ends of the upper and lower eyelid meet. The search of these two points is the most complicated due to their position on the face. In fact, the points do not have well defined geometrical features; for instance they are not critical points. Consequently, our algorithm uses in sequence various search methods from the most accurate to the least accurate. The most accurate uses stronger conditions, thus the search can sometimes find no points; in this case, the algorithm uses a less accurate search method to detect these landmarks.



The first research method uses these conditions:

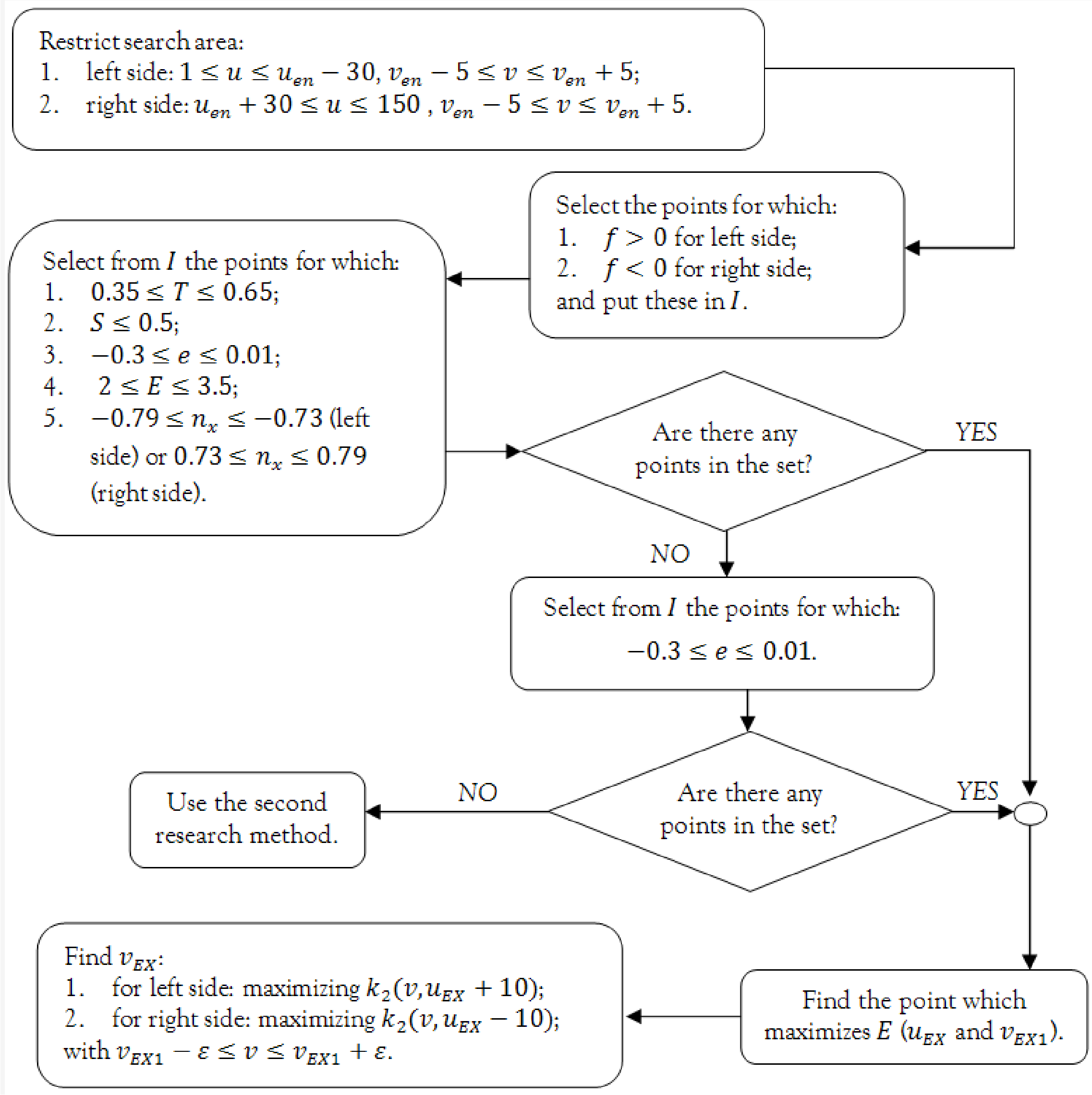
1. the Tangent Map must be in the range between 0.35 and 0.65 ($0.35 \leq T \leq 0.65$);
2. the Shape Index must be less than 0.5: $S \leq 0.5$;
3. the coefficient e must be in the range between -0.3 and 0.01: $-0.3 \leq e \leq 0.01$;
4. the coefficient E must be in the range between 2 and 3.5: $2 \leq E \leq 3.5$;
5. the x-component (n_x) of the normal deviation vector must be:
    a. left eye: $-0.79 \leq n_x \leq -0.73$;
    b. right eye: $0.73 \leq n_x \leq 0.79$;
6. the coefficient f must be positive for the left eye and negative for the right eye;
7. the point stays in this parametric region:
    a. left eye: $u_{ex} \leq u_{en} - 30, \quad v_{en} - 5 \leq v_{ex} \leq v_{en} + 5$
    b. right eye: $u_{ex} \geq u_{en} + 30, \quad v_{en} - 5 \leq v_{ex} \leq v_{en} + 5$

The normal deviation vector (n_d) is obtained by the formula:

$$n_d = n - \bar{n}$$

where n is the normal of a point, while $\bar{n}$ is the average of the normal of all points of the face. If the research finds no points, then the algorithm will perform a new research using only the conditions 7, 6, and 3; if the research finds no points again, then the algorithm will use the second research method. If the previous research finds a set of points, then the algorithm will consider up to 13 points (the points number obtained with the previous research could be less than 13) maximizing the derivative Z_{vv} and of these, it considers up to 8 point minimizing the derivative Z_{vv}. Subsequently, the algorithm finds the coordinates of landmark in this way:

1. finds the point which maximizes E (this point has coordinate u_{ex} and v_{ex1});
2. finds the point which maximizes k_2 curvature in a neighborhood of v_{ex1} (high value of the k_2 Principal Curvature).

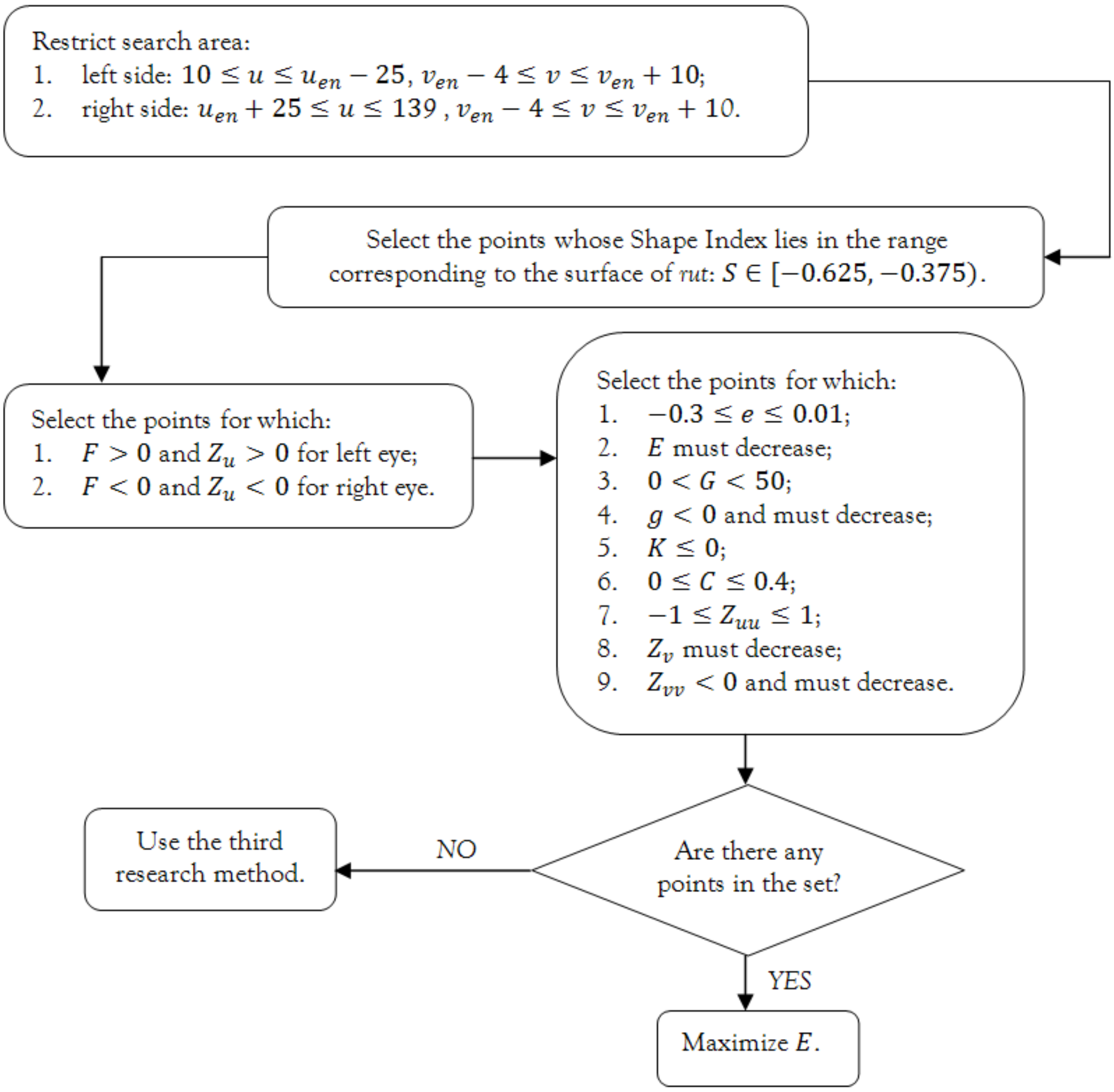


The second research method uses the same approach described in [15] ; therefore, the region of interest has these features:

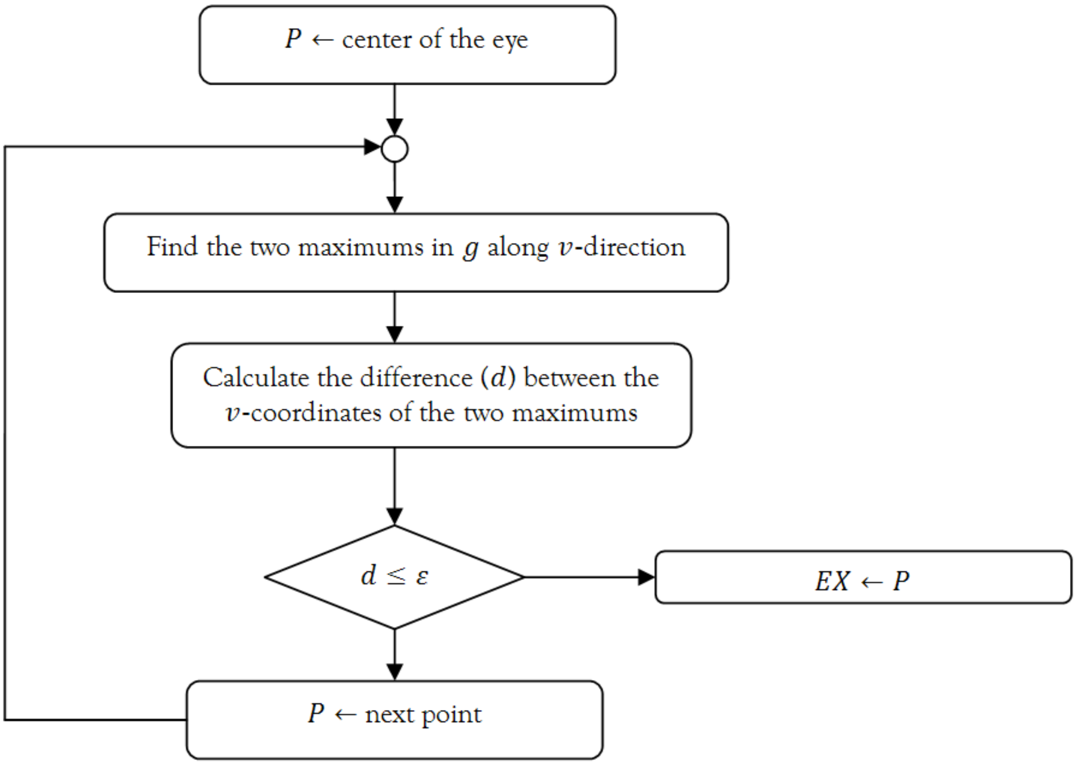
1. the Shape Index must be in the range between -0.625 and -0.375 corresponding to the surface of rut: $-0.625 \le S \le -0.375$;

2. the coefficient e must be in the range between -0.3 and 0.3: $-0.3 \leq e \leq 0.3$;
3. the coefficient E must decrease;
4. the coefficient G must be in the range between 0 and 50: $0 < G < 50$;
5. the coefficient g must decrease and must be negative;
6. the Gaussian Curvature must be positive: $K \leq 0$;
7. the Curvedness Index must be in the range between 0 and 0.4: $0 \leq C \leq 0.4$;
8. the derivative Z_u must decrease and:
    a. must be positive for left eye;
    b. must be negative for right eye;
9. the derivative Z_{uu} must be in the range between -1 and 1: $-1 \leq Z_{uu} \leq 1$;
10. the derivative Z_v must decrease;
11. the derivative Z_{vv} must decrease and must be negative;
12. the coefficient F must be positive for the left eye and negative for the right eye;
13. the point stays in this parametric region:
    a. left eye: $10 \leq u_{ex} \leq u_{en} - 25, \quad v_{en} - 4 \leq v_{ex} \leq v_{en} + 10$;
    b. right eye: $u_{en} + 25 \leq u_{ex} \leq 139, \quad v_{en} - 4 \leq v_{ex} \leq v_{en} + 10$.

Firstly, the algorithm finds all points which satisfy the previous conditions; secondly, it considers as *exocanthions* the two points which maximizes the coefficient E, one for side.



The last research method is based on the definition of the *exocanthions*: they are the two points at which the outer ends of the upper and lower eyelid meet. The problem of this approach is the eye edge detection without the use of texture. In this work, the eye edges are detected through the analysis of the coefficient g; in particular, the algorithm starts from center of eye and analyzes some points moving to the left (for left eye) or to the right (for right eye). For each point analyzed having coordinates u_P and v_P, it finds the two maximums in g along v-direction: one has a v-coordinate greater than v_P, while the other has v-coordinate lower than v_P. The algorithm stops when a point having a very small difference between the v-coordinates of the two maximums is found.



2.7 Other nose points

These points are not landmarks but are detected to obtain the shape of the nose with a good approximation; in particular, there are two points on the ridge of the nose and four points (two for side) on the base of the nose. The two points on the ridge of the nose have respectively an ordinate equal to $^1/_3$ and $^2/_3$ of the distance between *nasion* and *pronasal*, while the other points stay on the two straight lines perpendicular to the ridge of the nose, each of them passing through the points on the ridge.

The two points on the ridge of nose are detected this way:

1. calculates coordinates using the following formulas:
    a. first point: $y_{C1} = \frac{2*(y_{pn} - y_n)}{3} + y_n$, $x_{C1} = \frac{2*(x_{pn} - x_n)}{3} + x_n$;
    b. second point: $y_{C2} = \frac{(y_{pn} - y_n)}{3} + y_n$, $x_{C2} = \frac{(x_{pn} - x_n)}{3} + x_n$;
2. finds the u and v parameters (u_{C1}, v_{C1}, u_{C2}, v_{C2}) of the coordinates previously found;
3. finds the maximums in $Z(v_{C1}, u)$ and $Z(v_{C2}, u)$ in a neighborhood of the point u_{C1} and u_{C2}.

The four point on the base of nose are detected this way:

1. two points (one for side), both called $B1$, are detected maximizing the coefficient e along straight line perpendicular to the ridge of the nose passing through $C1$;
2. the other two points (one for side), both called $B2$, are detected maximizing the coefficient e along straight line perpendicular to the ridge of the nose passing through $C2$.

2.8 Other eye points

Similarly, these points are not landmarks and are detected to obtain the shape of the eye with a good approximation. In particular, three point on the eye upper edge ($O2$, $O3$, $O4$) and three point on the eye lower edge ($O1$, $O5$, $O6$) are detected. All points are researched through the maximization of the coefficient g keeping constant the u-parameter. The method:

1. calculates the u-coordinates with the following formulas:
   $u_{O1} = \frac{u_{ex}+u_{en}}{2}$,
   $u_{O3} = \frac{u_{o1}+u_{en}}{2}$,
   $u_{O4} = \frac{u_{ex}+u_{01}}{2}$,
   $u_{O2} = u_{O1}$,
   $u_{O5} = u_{O3}$,
   $u_{O6} = u_{O4}$;

2. calculates the initial v-coordinates with the following formulas:
   $v_{O1} = \frac{v_{ex}+v_{en}}{2} - 2$,
   $v_{O3} = \frac{v_{o1}+v_{en}}{2} + 2$,
   $v_{O4} = \frac{v_{ex}+v_{01}}{2} + 2$,
   $v_{O2} = v_{O1} + 4$,
   $v_{O5} = v_{O3} - 4$,
   $v_{O6} = v_{O4} - 4$;

3. finds the v-coordinates:
   a. for the upper the algorithm starts from the initial v-coordinates and searches the maximum of the coefficient g by moving upward;
   b. for the lower the algorithm starts from the initial v-coordinates and searches the maximum of the coefficient g by moving down.

3. Experimental Validation

The algorithm for landmark extraction was elaborated and implemented in Matlab®. Thirty-six faces of nine people with different facial expressions were scanned with a Minolta Vivid 910 and used for the experimentation. The scanned people were all Caucasian, male and female, from 20 to 40 years old. The computing time of the algorithm is less than a second for each face surface.

In the following paragraphs the results of the landmark extraction procedures will be presented (using one face as sample) in a graphical way, where the descriptor graphical behaviour is shown, while the corresponding landmark neighbourhood reduction is shown on the right one.

3.1 Pronasal

The zone-of-interest of the *pronasal* is found through a maximization process of the k_2 curvature. Then, the landmark is extracted finding a local maximum in Z in the area of the points-of-interest.

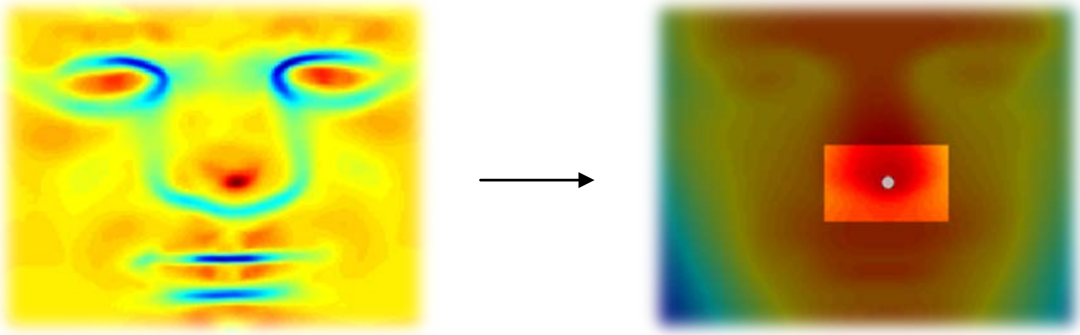


Figure 14. **(left)** the graph of the $\boldsymbol{k_2}$ curvature; the higher values are in the *pronasal* region, therefore maximizing $\boldsymbol{k_2}$ (white point on the **right image**) we can find the area-of-interest (the brightest region in the **right image**).

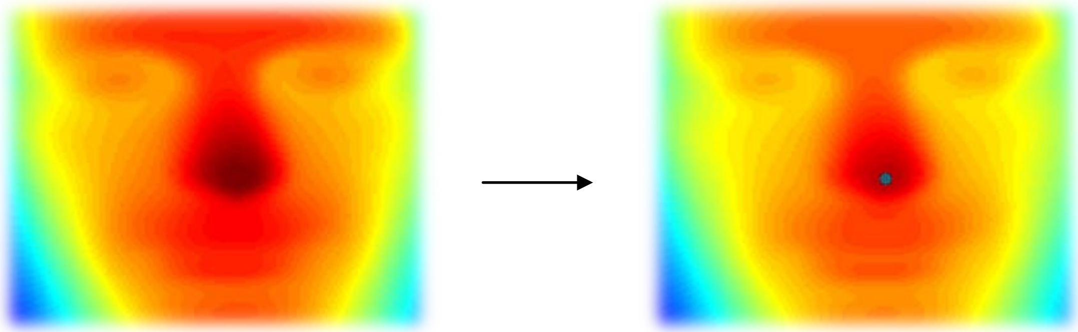


Figure 15. **(left)** representation of the *z*-coordinate (matrix $\boldsymbol{Z}$), namely of the face surface; **(right)** the landmark extracted by maximizing $\boldsymbol{Z}$ in the area-of-interest previously found.

In the previous example the absolute maximum in k_2 curvature corresponds to the *pronasal*, but in particular cases (for example hooked noses or noses with large nostrils) this is not true, as shown in Figure 16.

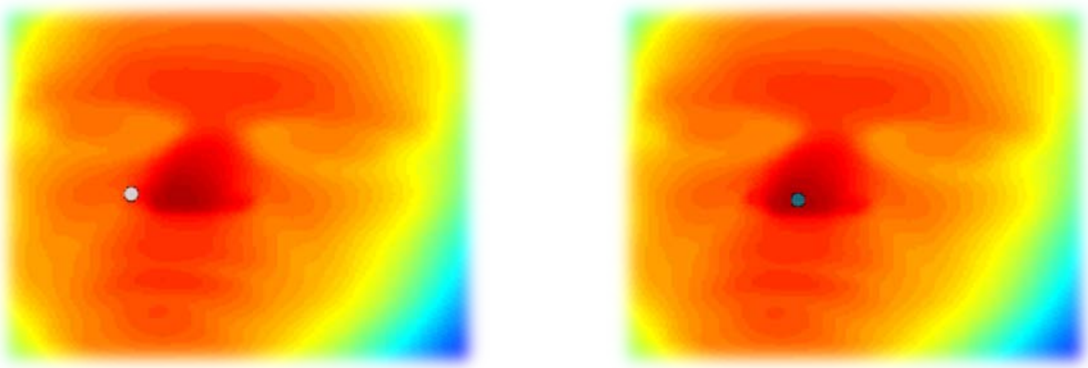


Figure 16. **(left)** in white the $\boldsymbol{k_2}$ curvature maximum which is very close to the left *alae*; **(right)** in blue the *pronasal* detected only by maximizing the *z*-coordinate.

3.2 Nasion

The search of the *nasion* is divided in three phases: firstly, the area-of-interest is found through a study of the face profile; secondly, a possible landmark is extracted by a search with geometric descriptors; finally, the landmark is detected through a search of a saddle point in the Tangent Map in a neighbourhood of the point previously found.

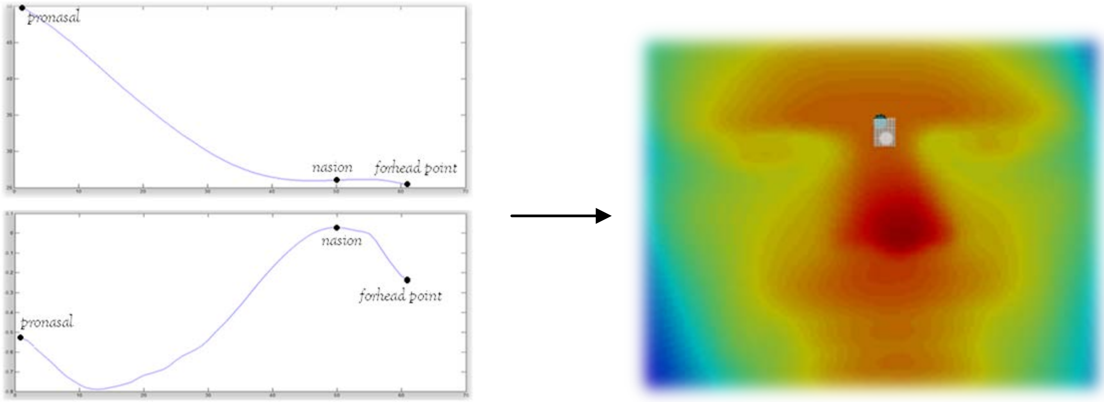


Figure 17. **(top-left)** the face profile of the region overlying the *pronasal*; **(bottom-left)** the face profile derivative (Figures 9 and 10 show a larger view); **(right)** in the bright area-of-interest, in blue the real *nasion*, and in white the point detected after the first phase: the $\boldsymbol{v}$-parameter of the nasion is computed maximizing the face profile derivative, while the $\boldsymbol{u}$-parameter is the maximum of $\boldsymbol{Z}$ in $\boldsymbol{u}$-direction.

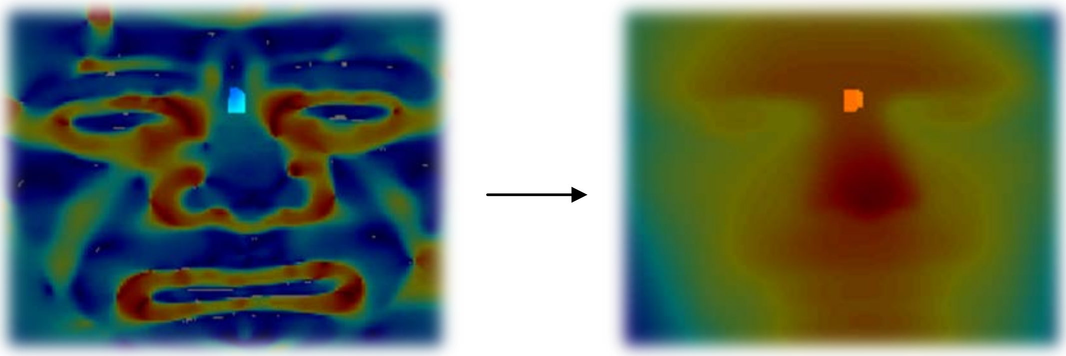


Figure 18. **(left)** in the brightest region there are the points whose Shape Index lies in the range corresponding to the surface of saddle ridge, saddle point, saddle rut, or rut in the zone-of-interest; **(right)** the reduction of the neighbourhood of interest on the face applying this restriction.

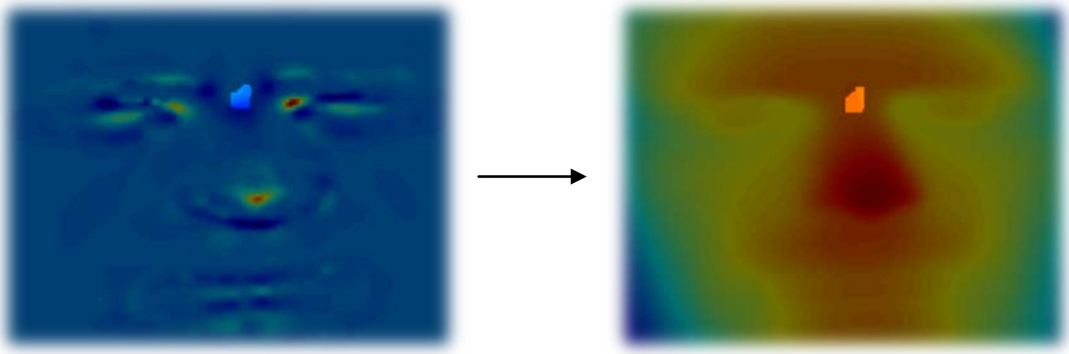


Figure 19. **(left)** in the brightest region there are the points which satisfy the condition on the Gaussian Curvature $(\boldsymbol{K} \leq \mathbf{0})$; **(right)** the reduction of the points-of-interest on the face applying this restriction.

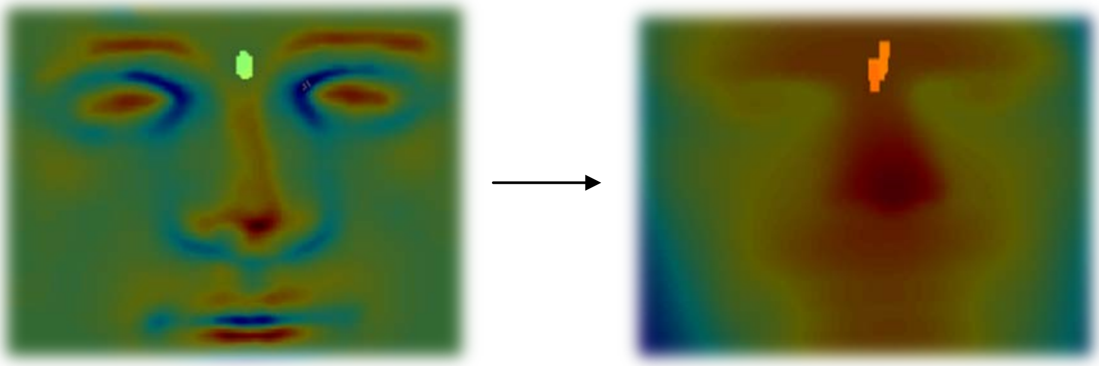


Figure 20. **(left)** in the brightest region there are the points which satisfy the condition on the Mean Curvature $(-\mathbf{0.5} < \boldsymbol{H} < 0.5)$; **(right)** the reduction of the neighbourhood on the face applying this restriction.

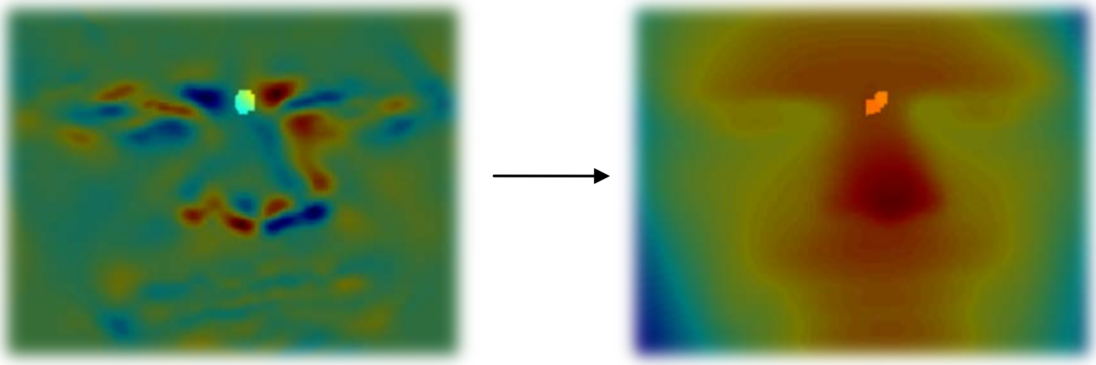


Figure 21. **(left)** in the brightest region there are the points which satisfy the condition on the coefficient f $(-\mathbf{0.1} < \boldsymbol{f} < 0.1)$; **(right)** the reduction of the points-of-interest on the face applying this restriction.

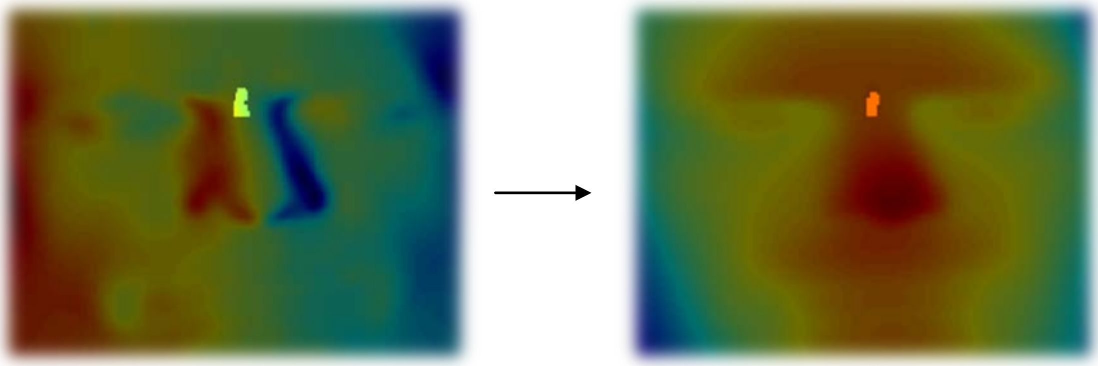


Figure 22. **(left)** in the brightest region there are the critical points in the $\boldsymbol{u}$-direction ($\boldsymbol{Z_u}$); **(right)** the reduction of the neighbourhood on the face selecting these critical points in the previous region-of-interest.

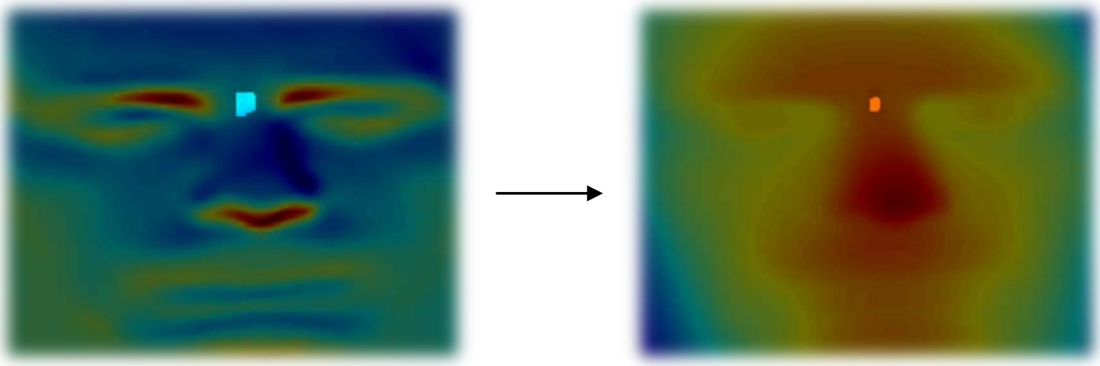


Figure 23. **(left)** in the brightest region there are the critical points in the $\boldsymbol{v}$-direction ($\boldsymbol{Z_v}$); **(right)** the reduction of the points-of-interest on the face selecting these critical points in the previous region-of-interest.

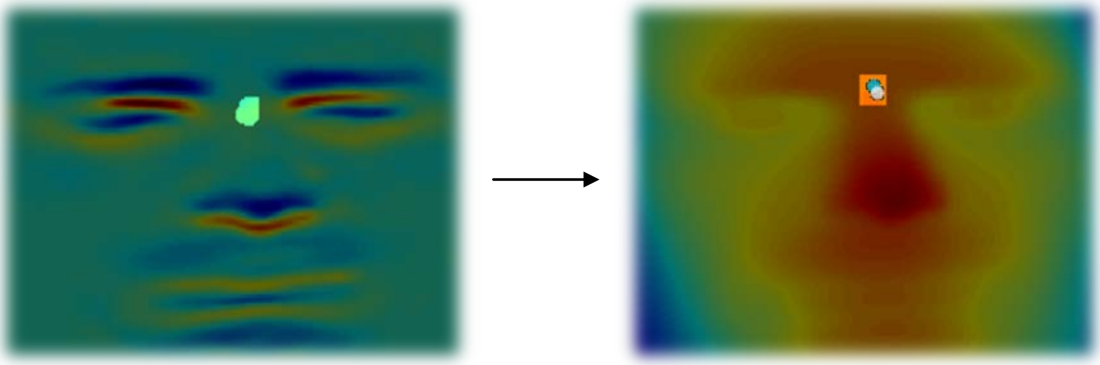


Figure 24. **(left)** the coefficient $\boldsymbol{g}$: the minimization is made in the brightest region; **(right)** in white the point extracted from the minimization of $\boldsymbol{g}$, in blue the real *nasion*; in bright the area-of-interest used in the next phase. This is the end of the second phase for the *nasion* extraction.

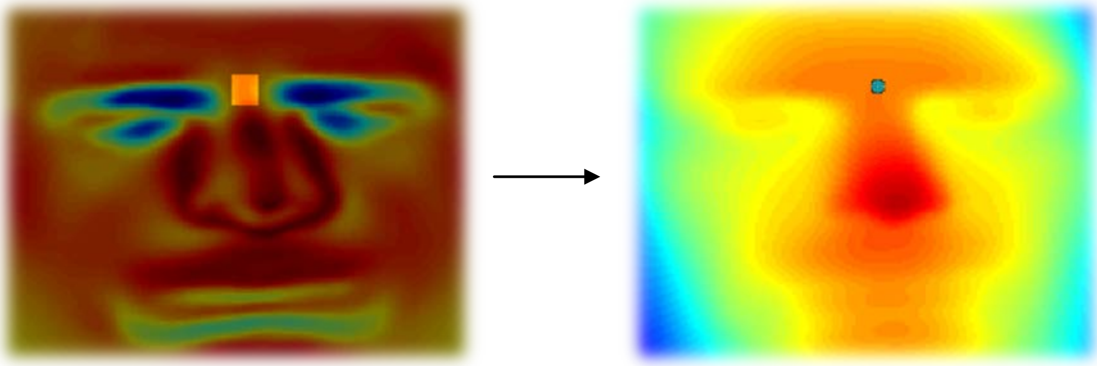


Figure 25. **(left)** the Tangent Map $\boldsymbol{T}$; **(right)** the detected *nasion*, which corresponds to a saddle point in the Tangent Map. This is the third and ending phase of the process.

3.3 Subnasal

The process of extraction of the *subnasal* consists in a maximization of the coefficients g and G. Firstly, the area-of-interest is found through a maximization of Z in u-direction; Secondly, the v-coordinate is found through a maximization of the coefficient G; finally, the *subnasal* is detected maximizing the coefficient g in a region underlying to the point extracted in the second phase.

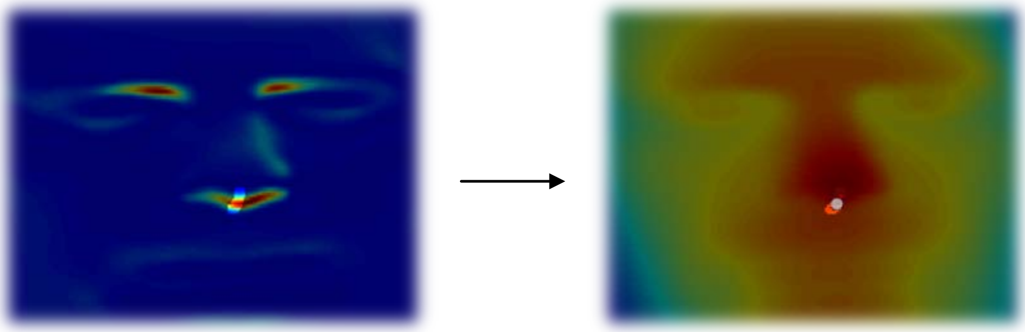


Figure 26. **(left)** the coefficient $\boldsymbol{G}$: the maximization is made in the brightest region; **(right)** in white the point which maximizes $\boldsymbol{G}$; the brightest region is the new area-of-interest.

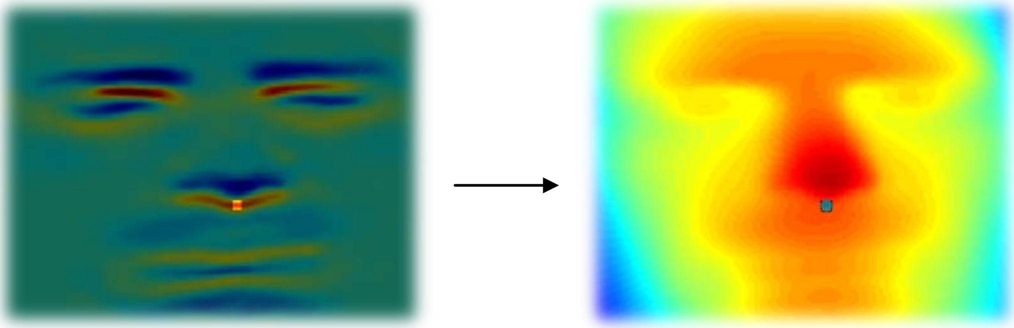


Figure 27. **(left)** the coefficient $\boldsymbol{g}$: the maximization is made in the brightest region; **(right)** in blue the *subnasal* detected by maximizing $\boldsymbol{g}$.

3.4 Alae

The extraction of the *alae* consists in two maximizations made in two zones-of-interest. These areas are found by searching the maximum of Z in v-direction in the left and right region of the *pronasal*. Then, the *alae* are the points which maximize the coefficient $\boldsymbol{e}$.

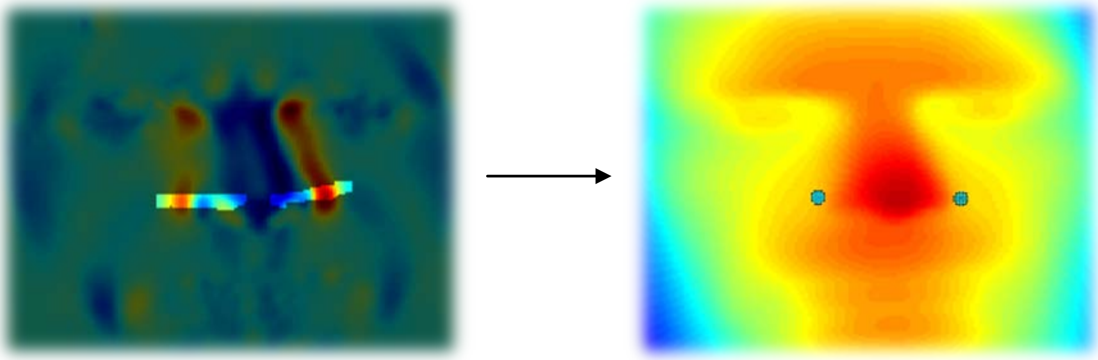


Figure 28. **(left)** the coefficient $\boldsymbol{e}$: the maximizations are made in the brightest regions; **(right)** in blue the *alae* detected by maximizing $\boldsymbol{e}$.

3.5 Endocanthions

The process of extraction of the *endocanthions* consists in two phases: firstly, the coefficient e is maximized in two areas-of-interest close to *nasion*; secondly, the *endocanthions* are detected through a search of two local minimums close to the two points previously extracted. The search of the local minimum is performed with the method previously described.

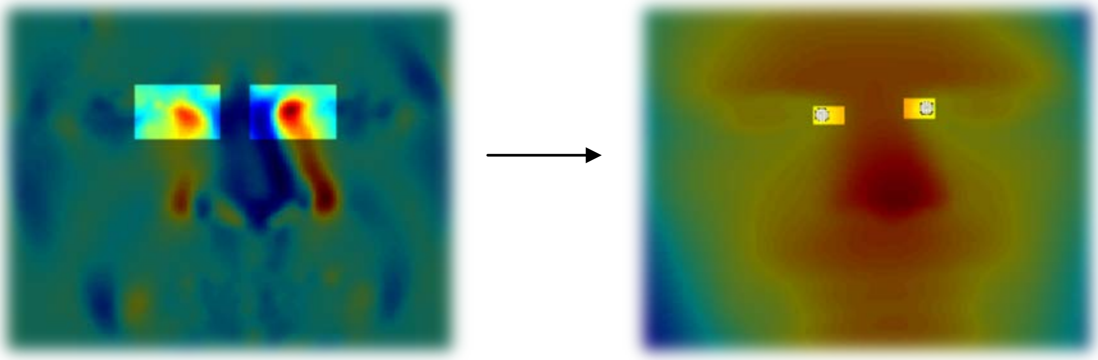


Figure 29. (left) the coefficient ***e***: the maximizations are made in the brightest regions; **(right)** in white the points detected by maximizing ***e***, while the brightest regions are the new areas-of-interest used in the next phase.

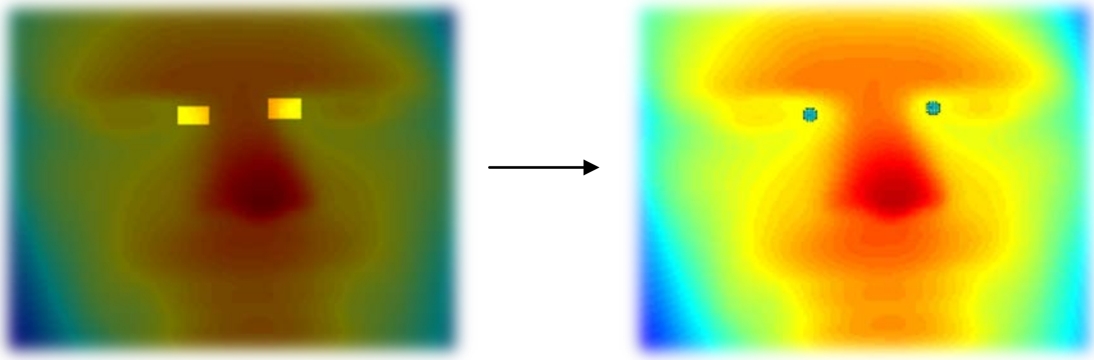


Figure 30. (left) representation of the *z*-coordinate (***Z***): the minimizations are made in the brightest regions; **(right)** in blue the detected *endocanthions*.

3.6 Exocanthions

The search of the *exocanthions* is performed through the use of the three search methods, from the most accurate to least accurate.

3.6.1 The first research method

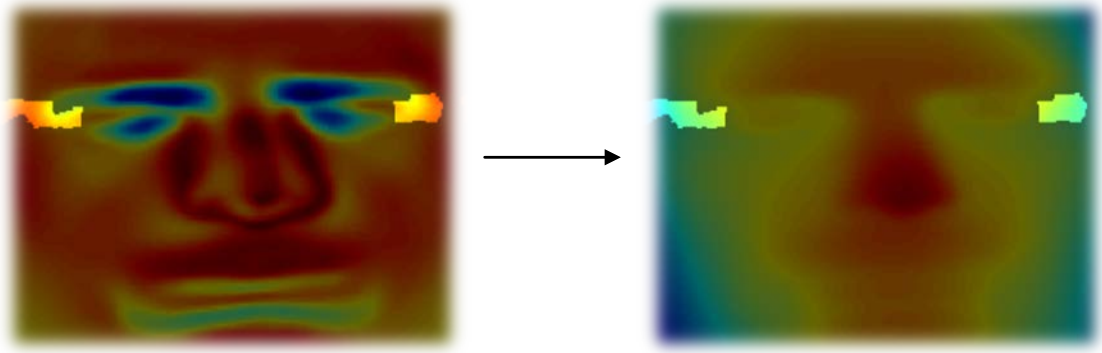


Figure 31. **(left)** in the brightest regions there are the points which satisfy the condition on the Tangent Map ($\boldsymbol{0.35 \leq T \leq 0.65}$); **(right)** the reduction of the points-of-interest on the face applying this restriction.

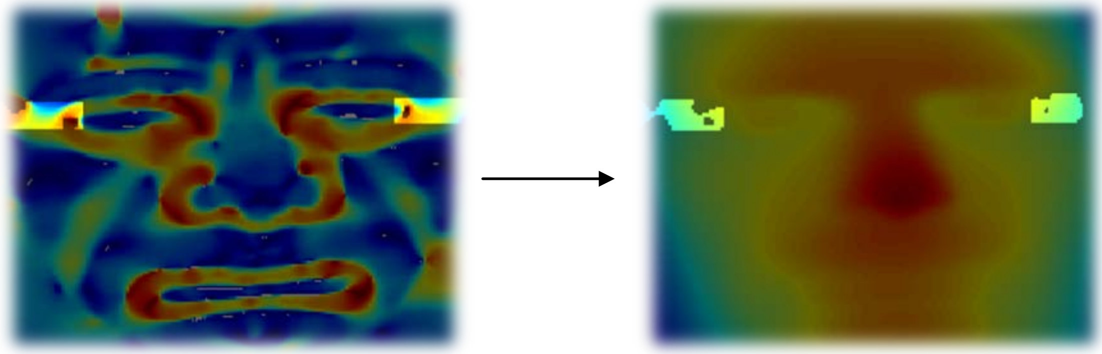


Figure 32. **(left)** in the brightest regions there are the points which satisfy the condition on the Shape Index ($\boldsymbol{S \leq 0.5}$); **(right)** the reduction of the area-of-interest on the face applying this restriction.

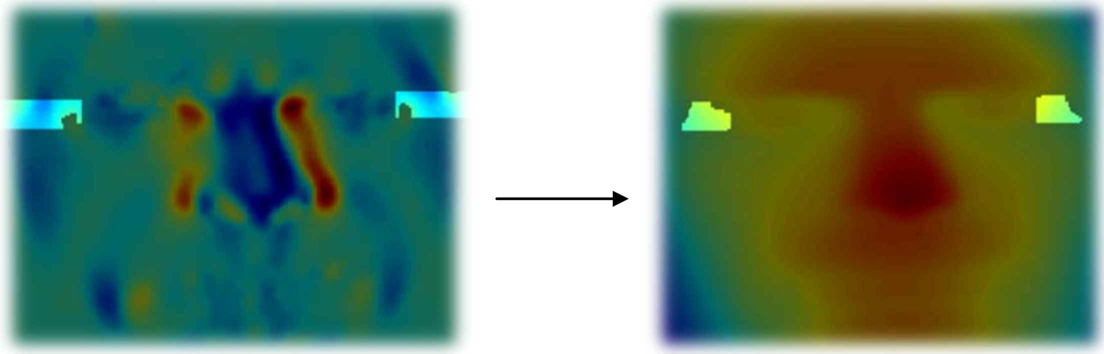


Figure 33. **(left)** in the brightest regions there are the points which satisfy the condition on the coefficient *e* ($\boldsymbol{-0.3 \leq e \leq 0.01}$); **(right)** the reduction of the neighbourhood on the face applying this restriction.

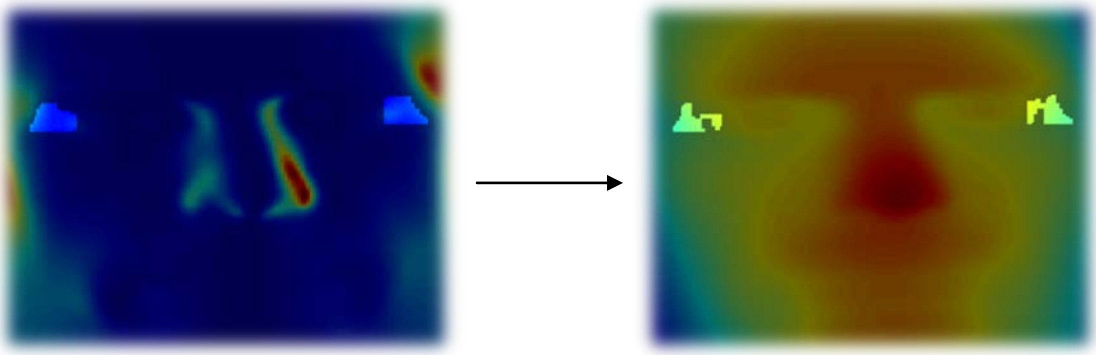


Figure 34. **(left)** in the brightest regions there are the points which satisfy the condition on the coefficient E ($\boldsymbol{2 \leq E \leq 3.5}$); **(right)** the reduction of the points-of-interest on the face applying this restriction.

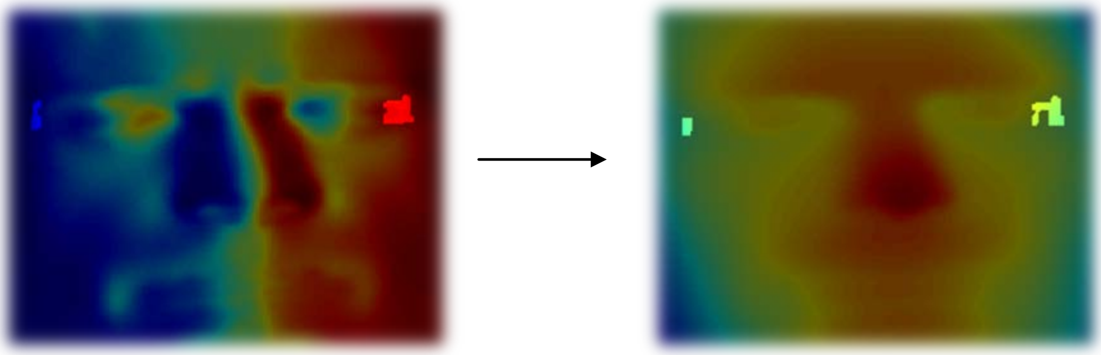


Figure 35. **(left)** in the brightest regions there are the points which satisfy the condition on the $\boldsymbol{x}$-component ($\boldsymbol{n_x}$) of the normal deviation vector (left eye: $\boldsymbol{-0.79 \leq n_x \leq -0.73}$; right eye: $\boldsymbol{0.73 \leq n_x \leq 0.79}$); **(right)** the reduction of the area-of-interest on the face applying this restriction.

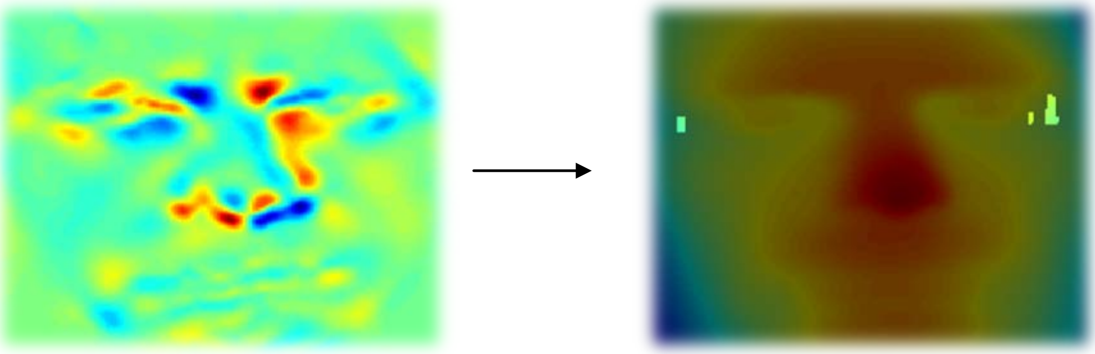


Figure 36. **(left)** in the brightest regions there are the points which satisfy the condition on the coefficient f (positive for the left eye and negative for the right eye, so this coefficient is used to choose the side); **(right)** the reduction of the neighbourhood on the face applying this restriction.

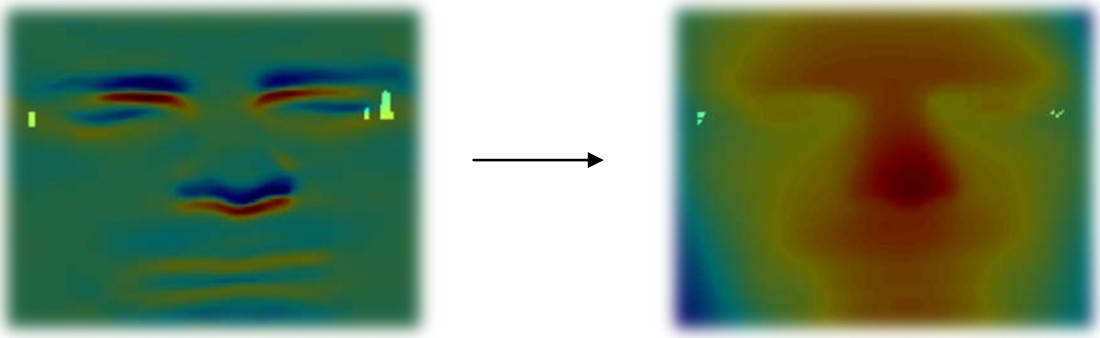


Figure 37. (left) in the brightest regions there are the points which are taken into consideration for maximizing $\boldsymbol{Z_{vv}}$; the method considers up to 13 points maximizing $\boldsymbol{Z_{vv}}$ and, among them, it considers up to 8 points minimizing $\boldsymbol{Z_{vv}}$; **(right)** the reduction of the points-of-interest on the face applying this restriction.

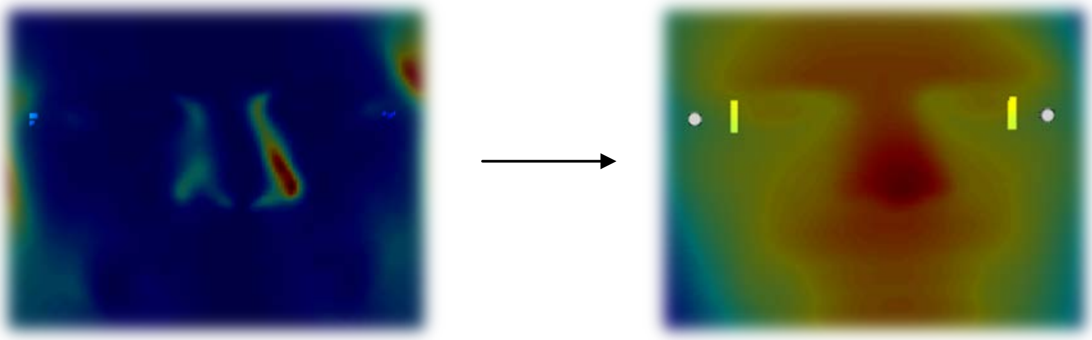


Figure 38. (left) in the brightest regions there are the points which are taken into consideration for maximizing the coefficient $\boldsymbol{E}$; **(right)** in white the points in which the coefficient $\boldsymbol{E}$ is maximum in the area-of-interest; in bright the zones-of-interest used in the next step.

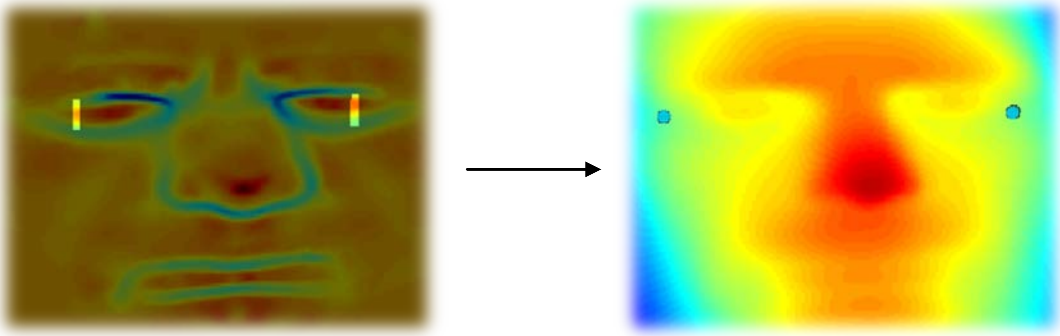


Figure 39. (left) in the brightest regions there are the points which are considered for the maximization of the $\boldsymbol{k_2}$ curvature; **(right)** in blue the detected *exocanthions*.

3.6.2 The second research method

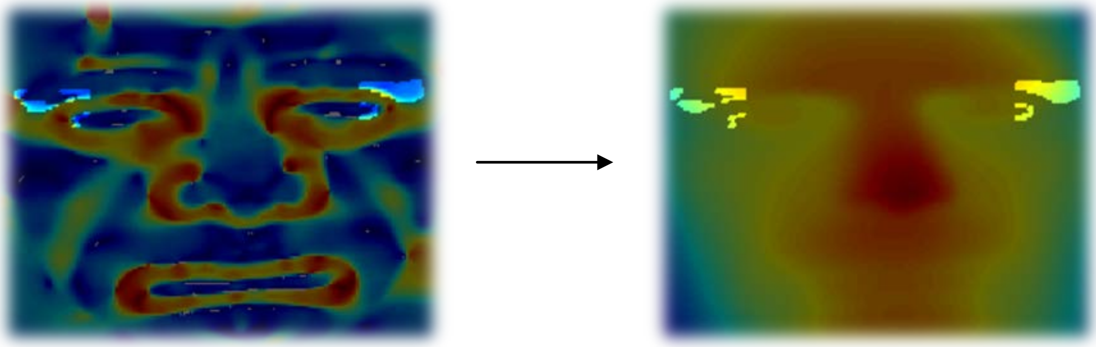


Figure 40. **(left)** in the brightest region there are the points whose Shape Index lies in the range corresponding to the surface of rut; **(right)** the reduction of the points-of-interest on the face applying this restriction.

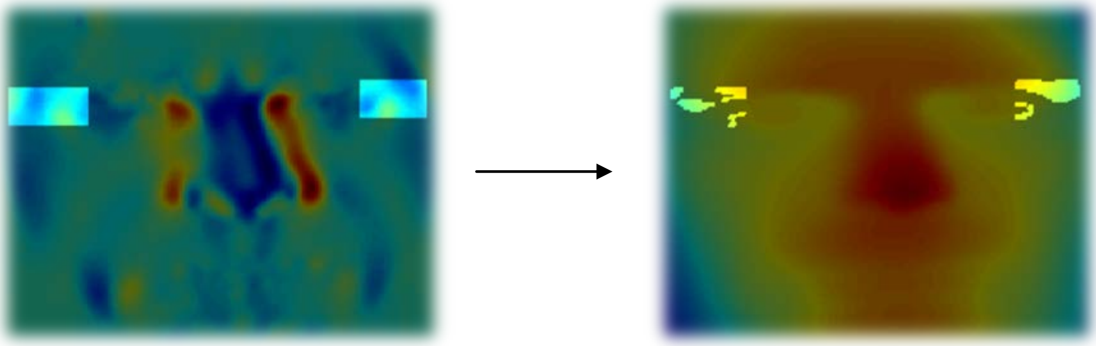


Figure 41. **(left)** in the brightest regions there are the points which satisfy the condition on the coefficient $\boldsymbol{e}$ ($\boldsymbol{-0.3 \le e \le 0.3}$); **(right)** the reduction of the neighbourhood on the face applying this restriction.

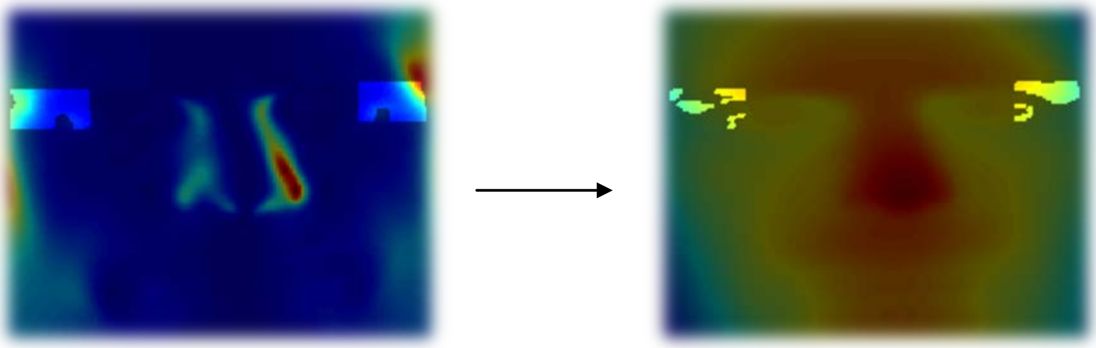


Figure 42. **(left)** in the brightest regions there are the points which satisfy the condition on the coefficient $\boldsymbol{E}$ (E decreases in the *exocanthions* zone); **(right)** the reduction of the area-of-interest on the face applying this restriction.

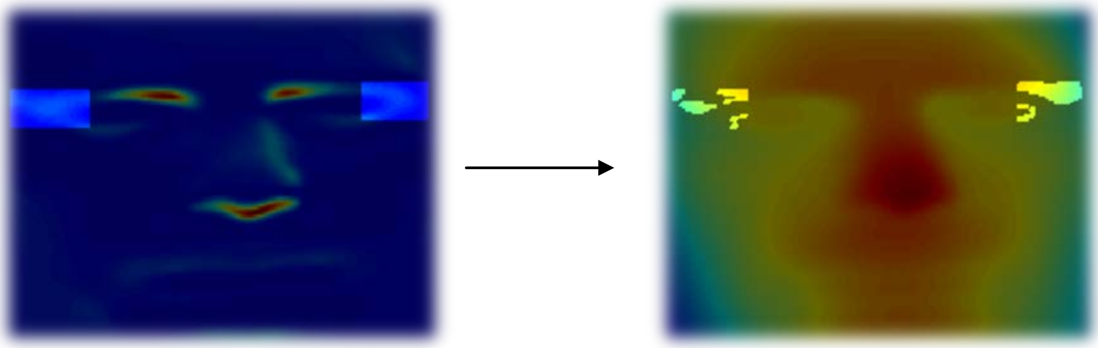


Figure 43. **(left)** in the brightest regions there are the points which satisfy the condition on the coefficient $\boldsymbol{G}$ ($\mathbf{0} < G < 50$); **(right)** the reduction of the points-of-interest on the face applying this restriction.

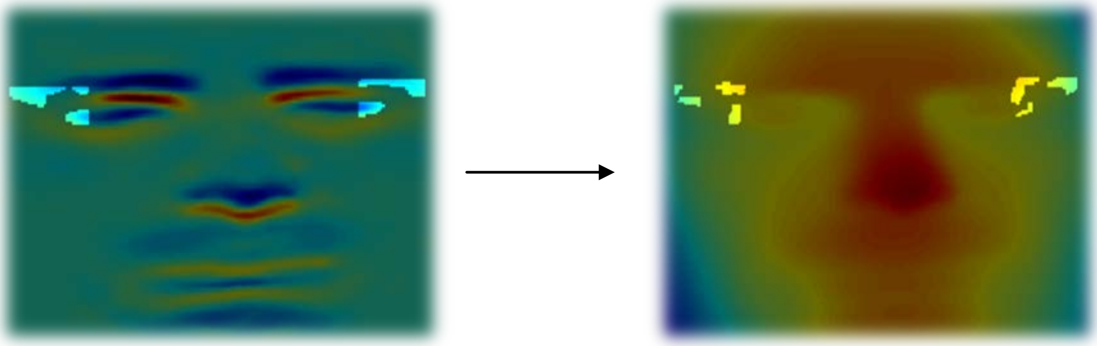


Figure 44. **(left)** in the brightest regions there are the points which satisfy the condition on the coefficient $\boldsymbol{g}$ (g decreases in the zone of the *exocanthions* and is negative); **(right)** the reduction of the area-of-interest on the face applying this restriction.

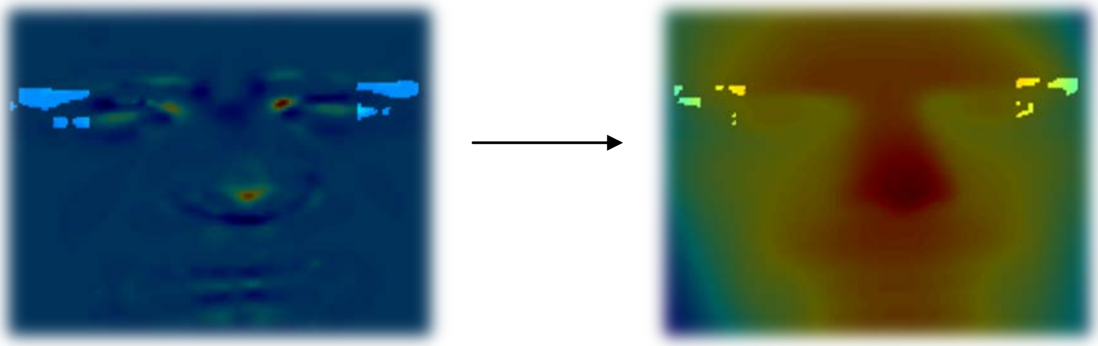


Figure 45. **(left)** in the brightest regions there are the points which satisfy the condition on the Gaussian Curvature ($\boldsymbol{K} \leq \mathbf{0}$); **(right)** the reduction of the neighbourhood on the face applying this restriction.

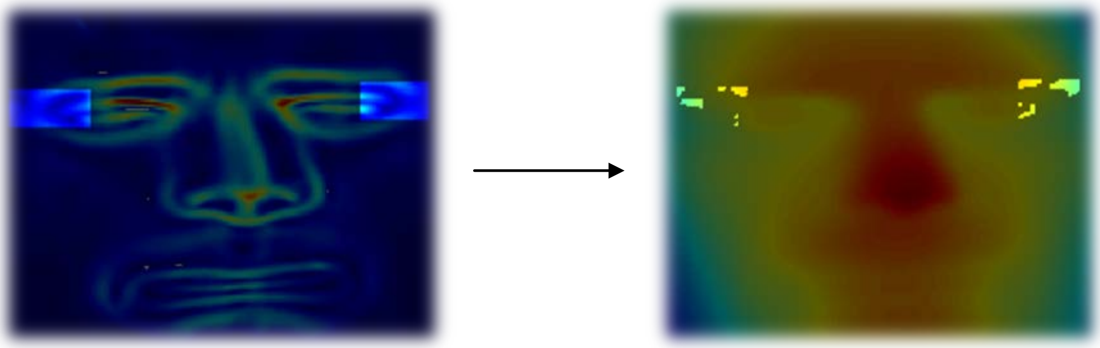


Figure 46. **(left)** in the brightest regions there are the points which satisfy the condition on the Curvedness Index $(\boldsymbol{0 \leq C \leq 0.4})$; **(right)** the reduction of the zone-of-interest on the face applying this restriction.

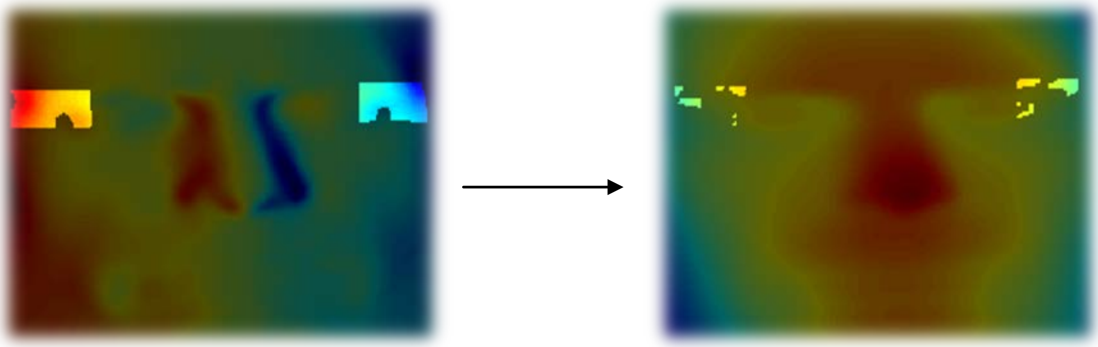


Figure 47. **(left)** in the brightest regions there are the points which satisfy the condition on the $\boldsymbol{Z_u}$ ($\boldsymbol{Z_u}$ must decrease and is positive for left eye, negative for the right); **(right)** the reduction of the zone-of-interest on the face applying this restriction.

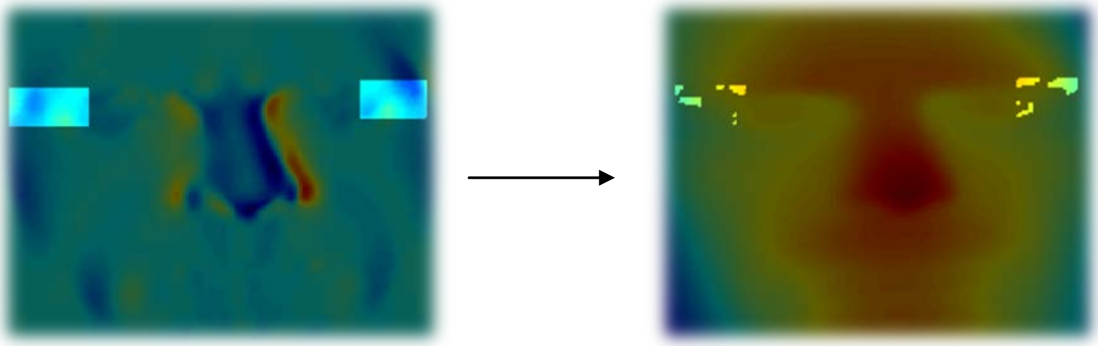


Figure 48. **(left)** in the brightest regions there are the points whose satisfy the condition on the $\boldsymbol{Z_{uu}}$ $(\boldsymbol{-1 \leq Z_{uu} \leq 1})$; **(right)** the reduction of the points-of-interest on the face applying this restriction.

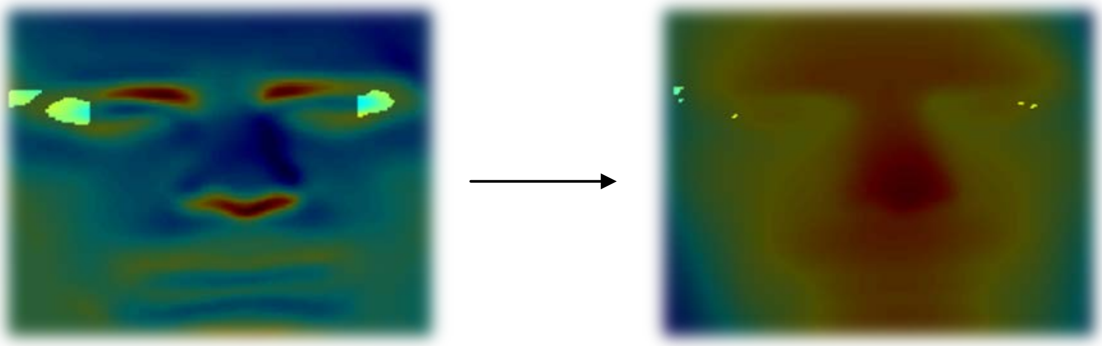


Figure 49. **(left)** in the brightest regions there are the points which satisfy the condition on the $\boldsymbol{Z_v}$ ($\boldsymbol{Z_v}$ decreases in the *exocanthions* zone); **(right)** the reduction of the neighbourhood on the face applying this restriction.

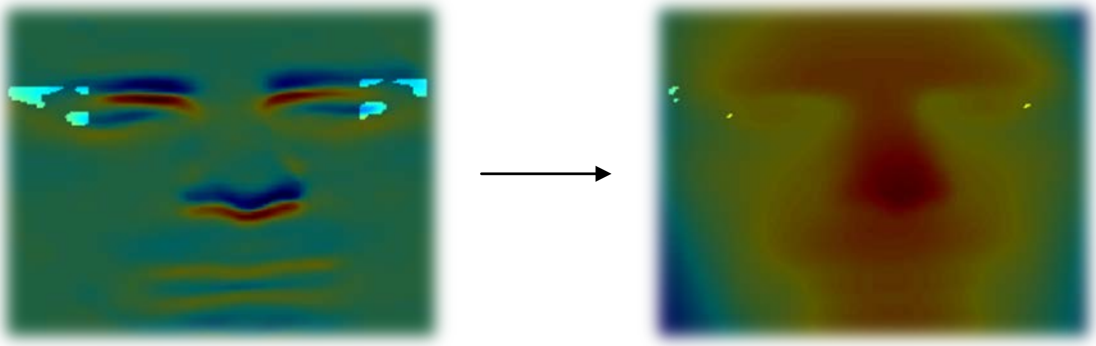


Figure 50. **(left)** in the brightest regions there are the points which satisfy the condition on the $\boldsymbol{Z_{vv}}$ ($\boldsymbol{Z_{vv}}$ decreases and is negative in the area); **(right)** the reduction of the points-of-interest on the face applying this restriction.

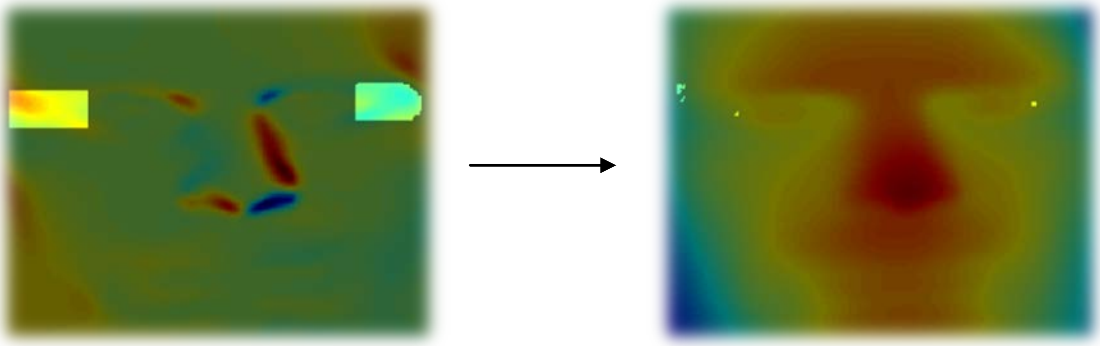


Figure 51. **(left)** in the brightest regions there are the points which satisfy the condition on the coefficient $\boldsymbol{F}$ (this coefficient is used to choose the side: $\boldsymbol{F}$ is positive for the left eye and negative for the right eye); **(right)** the reduction of the zone-of-interest on the face applying this restriction.

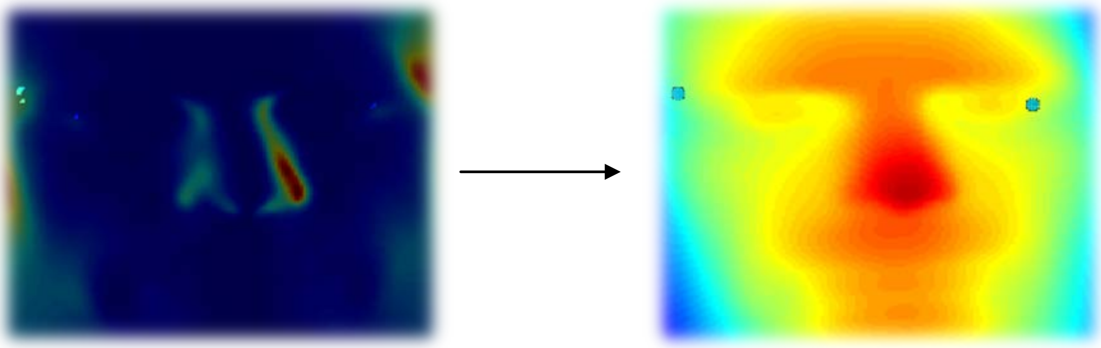


Figure 52. **(left)** in the brightest regions there are the points which are taken into consideration for maximizing the coefficient $\boldsymbol{E}$; **(right)** in blue the landmarks detected by maximizing $\boldsymbol{E}$.

Comparing the points extracted using the second research method with the ones extracted using the first one, we can note that the second method is less accurate; in fact, the left *exocanthion* is located on the right of the detected point, while the right *exocanthion* is very close to the extracted point.

3.6.3 The Third research method

The general idea is to detect the upper and the lower eyelid and to find the outer end where they meet: this search is performed using the coefficient g.

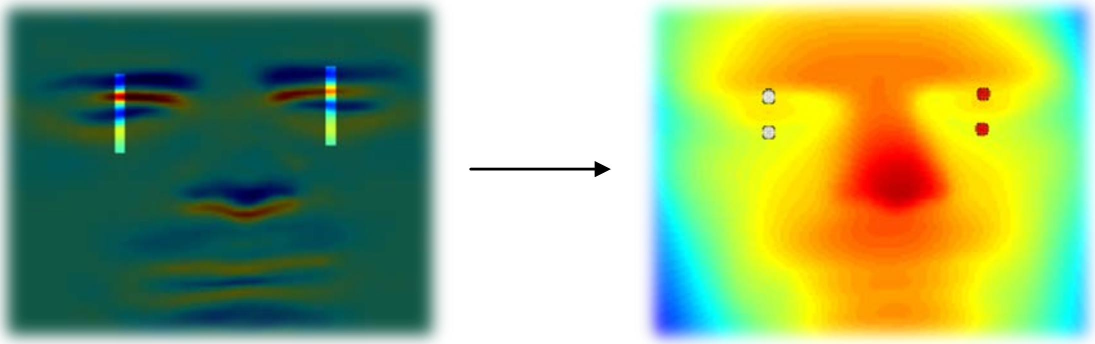


Figure 53. **(left)** the brightest regions are the areas-of-interest where the upper and the lower edges will be detected; **(right)** the two white points are the upper and the lower eyelid of the left eye, while the two red points are the upper and the lower eyelid of the right eye.

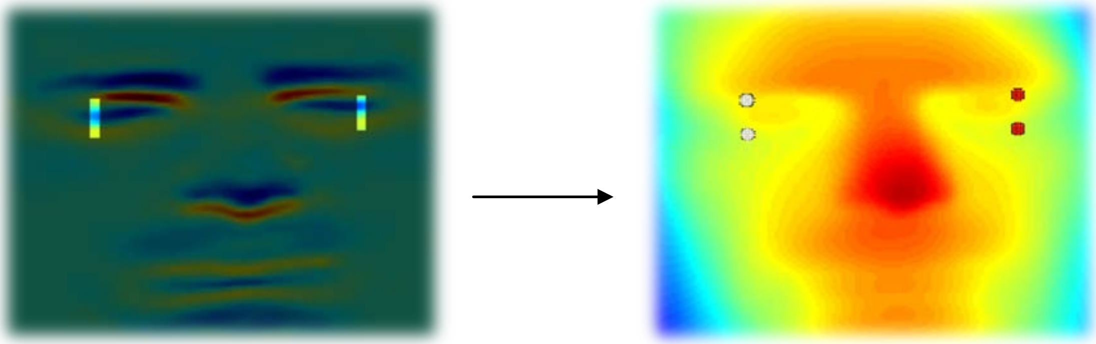


Figure 54. **(left)** the brightest regions are the new zones-of-interest where the upper and the lower edges will be detected: we can see how the new areas are more external than the previous ones; **(right)** the two white points are the upper and the lower eyelid of the left eye, while the two red points are the upper and the lower eyelid of the right eye.

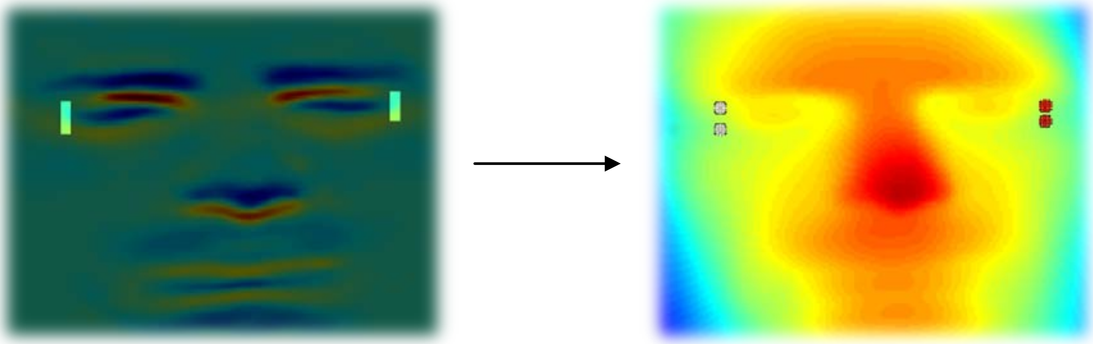


Figure 55. **(left)** the brightest regions are the new areas-of-interest where the upper and the lower edges will be detected; **(right)** the two white points are the upper and the lower eyelid of the left eye, while the two red points are the upper and the lower eyelid of the right eye.

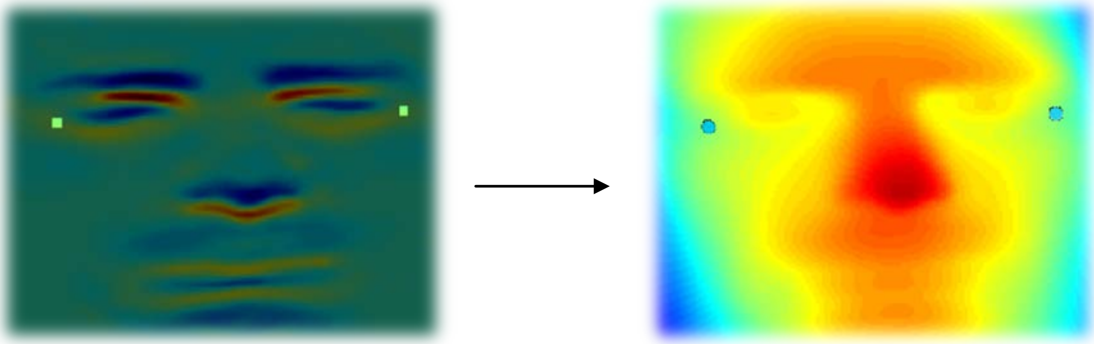


Figure 56. **(left)** the brightest regions are the new zones-of-interest where the *exocanthions* lie; **(right)** the two blue points are the detected *exocanthions*.

3.7 Other nose points

The search consists in two phases: firstly, the two points on the ridge of the nose are extracted through a search of maximums in the u-direction (the v-parameters are calculated by formulas previously described); secondly, the points on the nose base are detected maximizing the coefficient e along straight lines perpendicular to the ridge of the nose passing through the points previously detected.

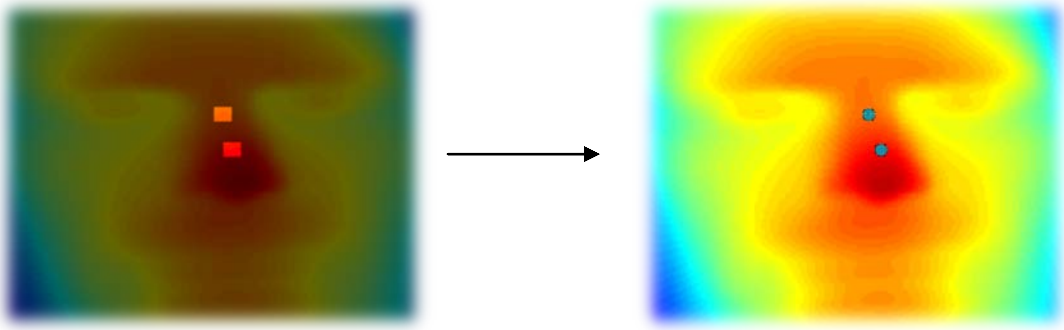


Figure 57. **(left)** the brightest regions are found by the formulas previously described; in these areas the maximums in the $\boldsymbol{u}$-direction are extracted, detecting two points on the ridge of the nose (as shown in the **right** figure).

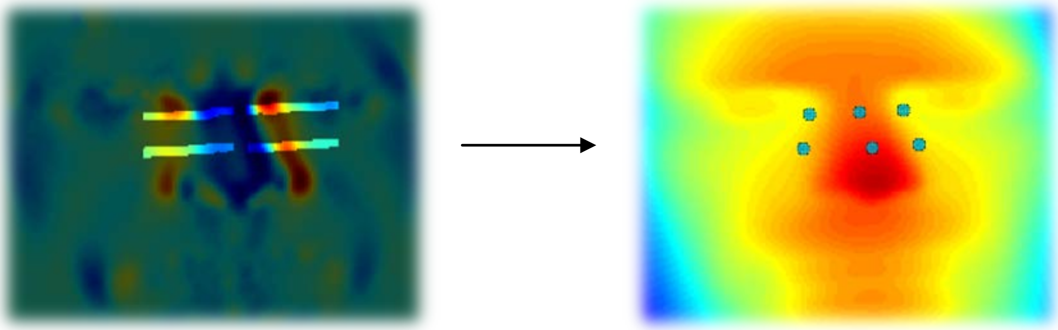


Figure 58. **(left)** the brightest regions lie along straight lines perpendicular to the ridge of the nose passing through the points previously detected; in these areas the maximums of the coefficient $\boldsymbol{e}$ are extracted, detecting 4 points on the base of the nose (as shown in the **right** figure).

3.8 Other eye points

These points are detected through the maximization of the coefficient g in some areas-of-interest; there are calculated by formulas previously described.

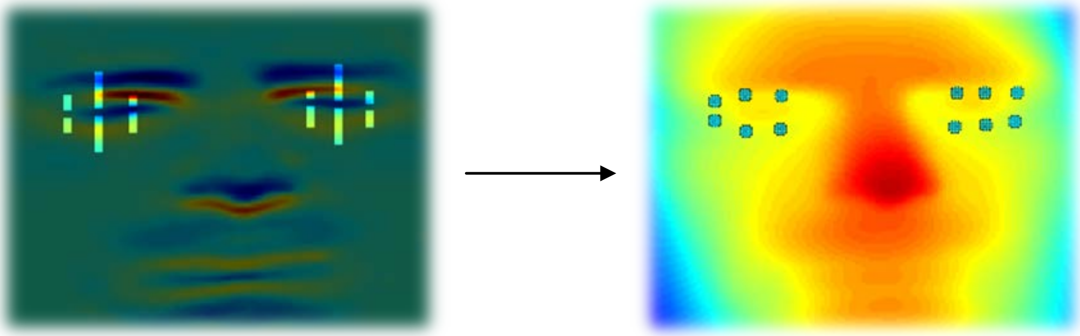
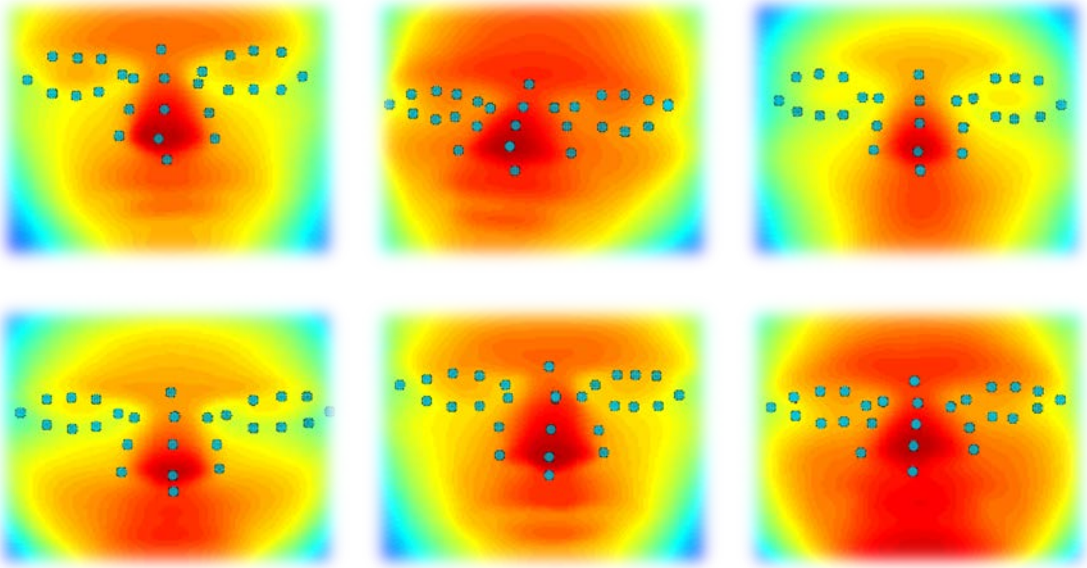


Figure 9. (left) the brightest regions lie in a neighborhood of the points obtained through the formulas previously described; in these areas the maximums of the coefficient ***g*** are extracted, detecting 6 points which delimit the eye (as shown in the **right** figure).

4. Results

The results of the algorithm, performed on nine faces belonging to different people, are shown in Figure 60, while in Figure 61 there are the results of the algorithm performed on three faces of the same person but with three different face expressions.



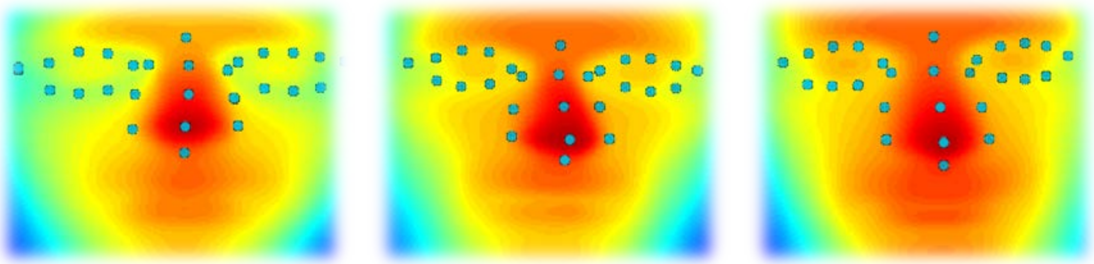


Figure 60. The extracted landmarks for nine of the thirty-six faces; the faces belong to different people.

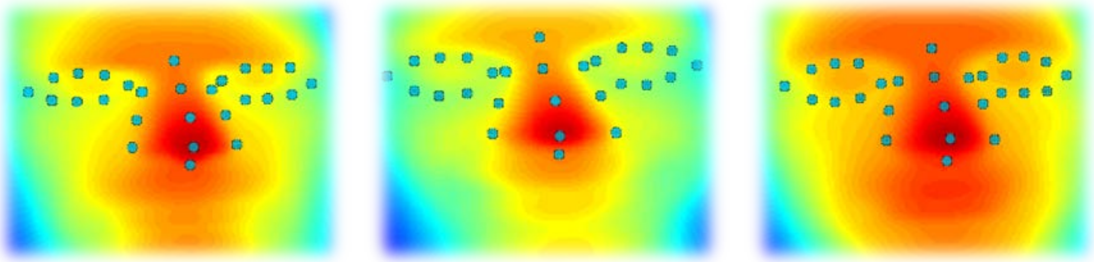


Figure 61. The extracted landmarks from faces belonging to the same person but with different expressions; **(left)** the face expression is disgust; **(center)** the face expression is enjoyment; **(right)** the face expression is sadness.

The computing times could be divided in two types:

1. the computing time of the algorithm, which is less than 1 seconds;
2. the processing time of the parametric surface fitting on the point cloud, which is about 10 seconds.

To verify the goodness of the extracted landmarks, a brief statistical study was performed. In this study we considered only the principal landmarks (*pronasal, subnasal, two alae, two endocanthions, two exocanthions, nasion*) since the others depend from the first ones and are not a biological meaning; furthermore, in our previous work [15] we extracted and detected only the principal landmarks.

Firstly, the landmarks of thirty-three faces were hand-detected from a plastic surgeon, so that we could compare these with the extracted ones. Only the principal landmarks were considered, omitting the ones which define the nose and eye shapes. Subsequently, Euclidean distances between the correct landmarks and the respective points given by algorithms were computed. However, in order to compare them, a normalizing operation is necessary; the idea was to normalize the distances by dividing them by the diagonal of the face bounding box. A bounding box is an invisible rectangular 3D box in which the face is somehow inscribed.

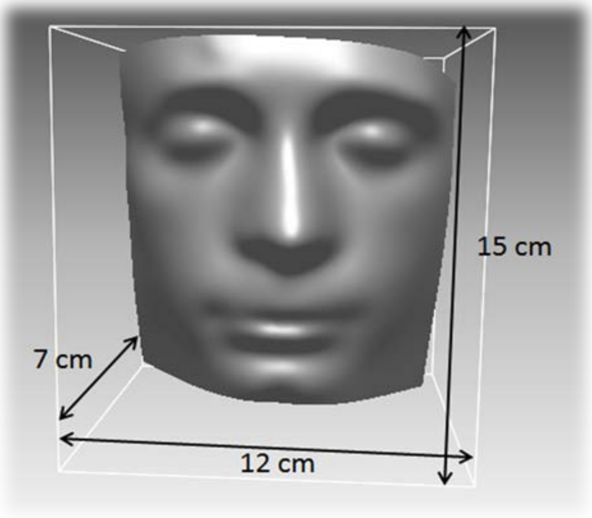


Figure 62. The measures of the bounding box of a face standard; we can note that the face width is about 12 cm, the face height is about 15 cm and the face depth is about 7 cm; applying Pitagora's theorem we can compute the diagonal of the bounding box ($\boldsymbol{D} \approx \mathbf{20}$ cm).

As shown in Figure 62, the sides of this box have a standard length, therefore, through a simple proportion, the normalization could be performed. The proportion is the following:

$$e : d_l = D : d_f,$$

where d_l is the distance between the correct landmark and the detected landmark, e is the normalized distance which must be computed (we can call it error), D is the diagonal of the bounding box of the face standard and d_f is the diagonal of the bounding box of the face where the landmarks are detected.

Once the normalized distances were computed, the sample mean μ and sample variance σ of these errors e_i were calculated:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} e_i, \qquad \sigma = \frac{1}{N-1}\sum_{i=1}^{N}(e_i - \mu)^2.$$

Since the diagonal of the bounding box of the standard face is given in centimetres, the errors and the mean will be in centimetres, while the variance will be in square centimetres.

Table 1 shows the errors on the detected landmarks with the geometric method, while table 2 shows the errors on the detected landmarks with the previous method.

Face	**PN**	**SN**	**Left *alae***	**Right *alae***	**N**	**Left EN**	**Left EX**	**Right EN**	**Right EX**
A disgust	0.069171	0.000000	0.102939	0.123921	0.071087	0.000000	0.000000	0.097312	0.122469
A enjoyment	0.066893	0.000000	0.125730	0.128060	0.094019	0.067379	0.195063	0.094766	0.226491
A fear	0.000000	0.259112	0.149469	0.167483	0.000000	0.000000	0.104483	0.000000	0.099717
A anger	0.000000	0.067840	0.127325	0.127756	0.068132	0.000000	0.108492	0.097135	0.000000
A straight	0.000000	0.142837	0.179085	0.000000	0.070954	0.070999	0.112884	0.000000	0.223418
A surprise	0.114984	0.108180	0.156197	0.000000	0.116447	0.093057	0.207247	0.116473	0.129299
A sadness	0.000000	0.127983	0.146393	0.156274	0.000000	0.000000	0.000000	0.104106	0.103056
B crooked mouth	0.074126	0.127731	0.000000	0.101559	0.000000	0.227978	0.000000	0.227402	0.306174
B puffy cheeks	0.101334	0.000000	0.000000	0.103273	0.000000	0.215730	0.207599	0.219302	0.000000
B open jaw	0.000000	0.121855	0.000000	0.105569	0.072183	0.220197	0.105198	0.107886	0.219968
B straight	0.000000	0.000000	0.000000	0.100350	0.097820	0.000000	0.123507	0.000000	0.096688
B smile	0.000000	0.000000	0.000000	0.000000	0.100900	0.075040	0.000000	0.000000	0.000000
C straight	0.000000	0.114335	0.106097	0.106188	0.080256	0.135223	0.107755	0.110031	0.107983
D crooked mouth	0.000000	0.000000	0.087976	0.000000	0.105260	0.000000	0.109156	0.107136	0.210427
D puffy cheeks	0.000000	0.000000	0.114842	0.130374	0.000000	0.000000	0.000000	0.109147	0.000000
D open jaw	0.071371	0.000000	0.000000	0.122839	0.000000	0.106740	0.113939	0.112331	0.000000
D straight	0.000000	0.123360	0.118684	0.000000	0.103075	0.110749	0.115518	0.112444	0.116453
D smile	0.071398	0.071987	0.000000	0.000000	0.000000	0.109750	0.113182	0.108528	0.112773
D crimped face	0.073178	0.131339	0.000000	0.093595	0.073183	0.104346	0.000000	0.000000	0.000000
E straight	0.073169	0.000000	0.098395	0.000000	0.000000	0.000000	0.000000	0.000000	0.117037
G straight	0.000000	0.000000	0.126099	0.000000	0.000000	0.000000	0.284262	0.000000	0.438821
H straight	0.000000	0.000000	0.000000	0.000000	0.000000	0.085962	0.125901	0.000000	0.202206
I disgust	0.000000	0.150083	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
I enjoyment	0.223603	0.123796	0.000000	0.000000	0.000000	0.000000	0.000000	0.084974	0.000000
I fear	0.000000	0.142259	0.000000	0.000000	0.000000	0.090407	0.000000	0.000000	0.000000
I anger	0.071645	0.142161	0.000000	0.000000	0.215278	0.085699	0.087036	0.085460	0.000000
I straight	0.000000	0.139943	0.000000	0.159684	0.087663	0.000000	0.089150	0.000000	0.000000
I surprise	0.000000	0.147411	0.000000	0.000000	0.070279	0.000000	0.000000	0.000000	0.000000
I sadness	0.072578	0.264788	0.145623	0.000000	0.000000	0.000000	0.120678	0.000000	0.106266
L disgust	0.000000	0.000000	0.000000	0.162794	0.000000	0.000000	0.000000	0.000000	0.097149
L enjoyment	0.000000	0.000000	0.000000	0.092274	0.000000	0.092863	0.000000	0.000000	0.000000

L fear	0.000000	0.120756	0.150874	0.156420	0.000000	0.000000	0.000000	0.083277	0.000000
L anger	0.081910	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.083499	0.000000
L straight	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
L surprise	0.000000	0.000000	0.113773	0.000000	0.000000	0.080680	0.000000	0.000000	0.000000
L sadness	0.000000	0.120888	0.000000	0.000000	0.000000	0.084069	0.000000	0.083885	0.000000
Mean	0.032371	0.076351	0.056931	0.059400	0.039626	0.057135	0.067529	0.059586	0.084344
Variance	0.002529	0.006033	0.004427	0.004366	0.002801	0.004667	0.006031	0.004112	0.011399

Table 1. The errors on the detected landmarks with the new method, their mean and variance.

Face	**PN**	**SN**	**Left *alae***	**Right *alae***	**N**	**Left EN**	**Left EX**	**Right EN**	**Right EX**
A disgust	0.000000	0.462384	1.377968	2.136248	4.781825	0.822180	2.965512	0.000000	3.198966
A enjoyment	0.092863	0.273609	0.800119	1.107198	4.995738	0.189145	2.773832	0.387509	3.526048
A fear	5.302633	6.037101	1.618077	2.046464	2.769072	nd	1.174162	2.095775	0.879138
A anger	0.000000	0.614204	0.831798	1.579049	5.792273	0.198926	3.301386	0.394830	3.251960
A straight	5.578411	6.032036	1.501161	1.220283	3.734194	nd	0.723541	nd	1.478767
A surprise	0.096215	0.690938	1.215083	1.814219	1.148803	0.000000	1.938825	0.812242	1.345096
A sadness	0.091621	0.660839	0.639964	1.342685	5.193155	0.854284	3.191624	0.000000	3.183644
B crooked mouth	0.000000	0.095451	1.222273	1.429403	5.487973	0.552146	0.000000	0.415520	3.545822
B puffy cheeks	0.000000	0.000000	1.562388	1.411442	1.870658	0.406317	0.000000	0.407503	0.000000
B open jaw	0.000000	0.093815	1.385636	1.694579	1.083590	0.966525	0.160395	0.898588	0.699424
B straight	0.000000	0.159025	1.513443	1.529998	1.320302	0.565691	3.827539	0.627688	3.646372
B smile	0.097797	0.000000	1.509481	1.420211	2.265855	0.837827	0.310812	0.655393	0.598336
C straight	3.404893	2.625729	1.084339	1.497613	1.197322	0.000000	0.836250	0.000000	0.523755
D crooked mouth	0.000000	0.000000	1.247381	1.552350	0.267925	0.000000	0.778690	0.000000	0.554704
D puffy cheeks	0.000000	0.000000	1.046686	1.115114	0.682151	0.135822	0.949677	0.000000	1.516129
D open jaw	0.000000	0.094580	1.325889	1.374958	0.678932	0.000000	0.894567	0.000000	1.296277
D straight	0.136921	0.094441	1.400497	1.567730	0.703805	0.688054	0.897806	0.000000	0.950379
D smile	0.000000	0.000000	1.116347	1.298311	1.250866	0.419763	1.201163	0.138033	1.021320
D crimped face	0.000000	0.000000	1.144163	1.439502	0.819671	0.000000	0.967131	0.000000	0.818975
E straight	0.192742	0.283780	1.142788	1.406316	1.128696	0.261565	1.390415	0.326844	0.763752
G straight	1.992700	1.088195	nd	Nd	5.804329	0.000000	3.331233	2.548691	3.158839
H straight	1.553049	0.573259	nd	Nd	5.060891	0.948022	4.051764	0.788119	2.902754
I disgust	0.000000	0.980362	2.150249	0.594056	nd	0.464803	0.827523	0.630289	3.688349
I enjoyment	5.644865	4.960352	2.043826	1.871001	2.727443	nd	1.717792	nd	1.196462
I fear	0.157299	0.572124	1.408674	2.035180	5.903441	0.764994	3.350541	0.264345	3.417757

I anger	0.000000	0.478551	2.066248	2.241016	nd	0.359521	3.467147	0.258844	3.494618
I straight	0.000000	0.429635	1.781803	2.035776	5.781285	0.615706	2.989192	0.498771	2.695455
I surprise	5.777090	6.386505	1.828581	1.971882	3.518407	nd	nd	Nd	nd
I sadness	0.000000	1.035113	2.079713	2.111924	nd	0.376088	3.420194	0.484326	2.318734
L disgust	0.000000	0.000000	3.268479	1.530134	2.397277	0.648415	2.235085	0.263401	2.435169
L enjoyment	0.102192	0.000000	0.628435	1.014100	1.336682	0.481138	1.094927	0.847771	0.801169
L fear	0.000000	0.493908	0.789340	1.640910	2.978528	0.254219	2.673390	0.463005	2.472777
L anger	0.000000	0.109172	0.977452	1.861438	3.010985	2.500448	2.857788	0.576667	2.422708
L straight	5.298176	5.910212	1.922673	3.373614	3.174130	1.431659	1.294068	nd	3.776546
L surprise	0.000000	0.857835	0.914187	1.821694	nd	0.240388	3.943875	0.612449	3.007881
L sadness	0.210671	0.326624	0.963366	1.458054	2.512631	0.603883	2.793119	0.255596	2.287509
Mean	0.992504	1.178327	1.397309	1.633660	2.855589	0.518360	1.952313	0.489131	2.082160
Variance	3.863761	3.919188	0.286411	0.224900	3.474020	0.252777	1.596996	0.313147	1.410543

Table 2. The errors on the detected landmarks with the previous method, their mean and variance; the word "nd" means not detected.

The comparison between the overall means of the two methods and the graphical representations of mean error and standard deviation ($\sqrt{\sigma}$) for each method are shown in Figures 63, 64, and 65.

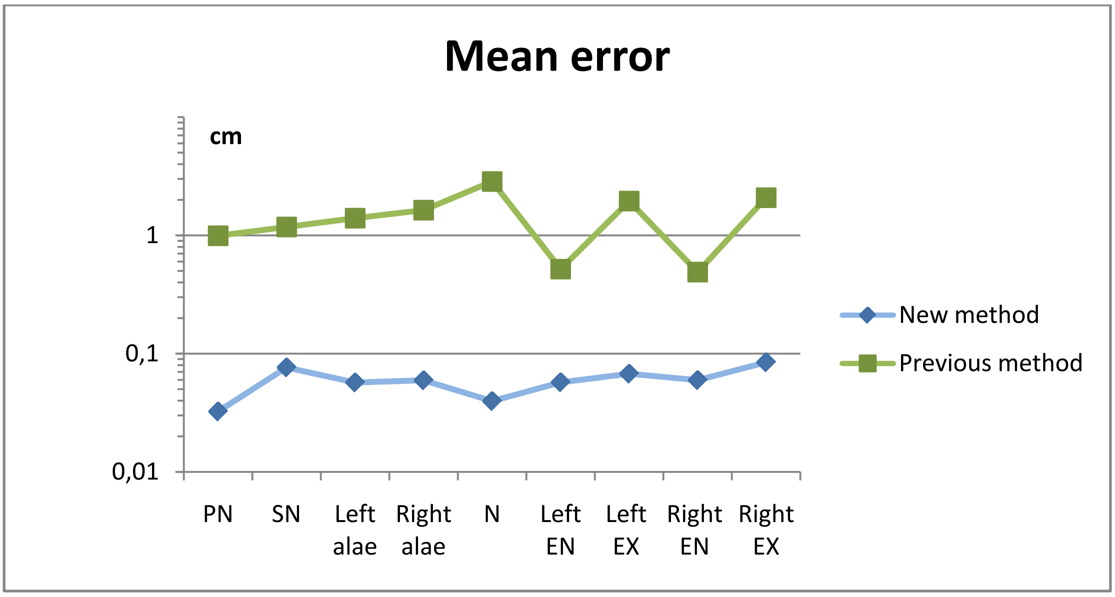


Figure 63. A graphical representation, in logarithmic scale, of mean error of the two methods; the mean error for each detected landmark in the new method is lower than the one in the previous method.

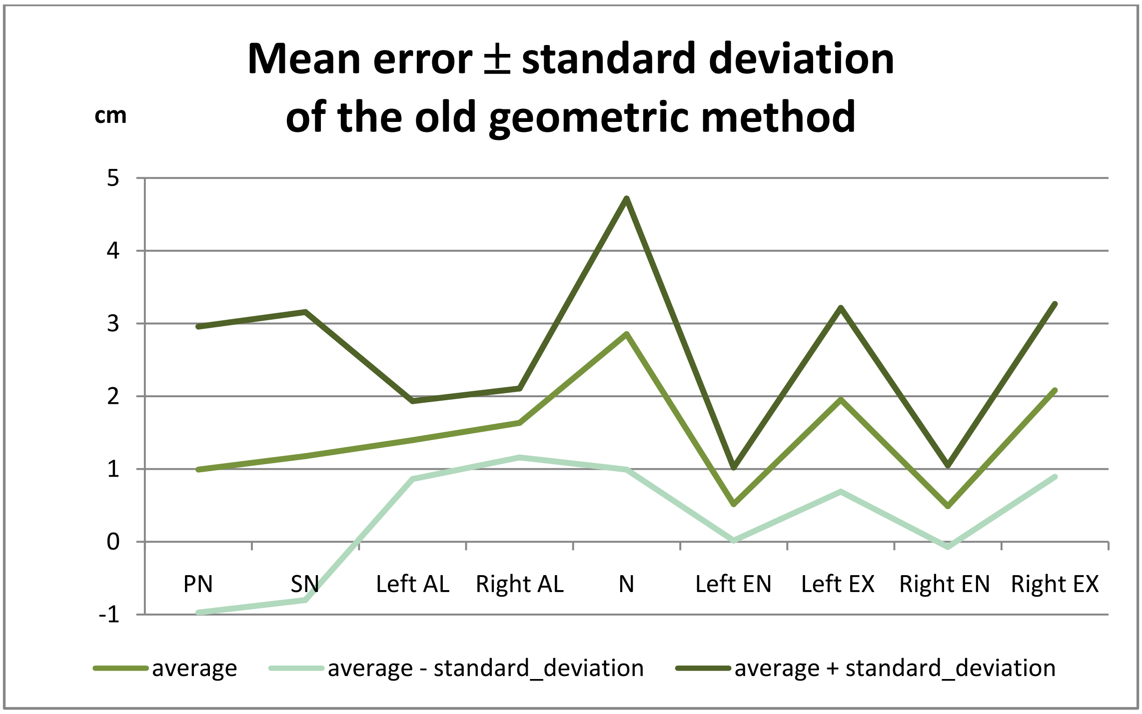


Figure 64. A graphical representation of mean error and standard deviation of the previous method.

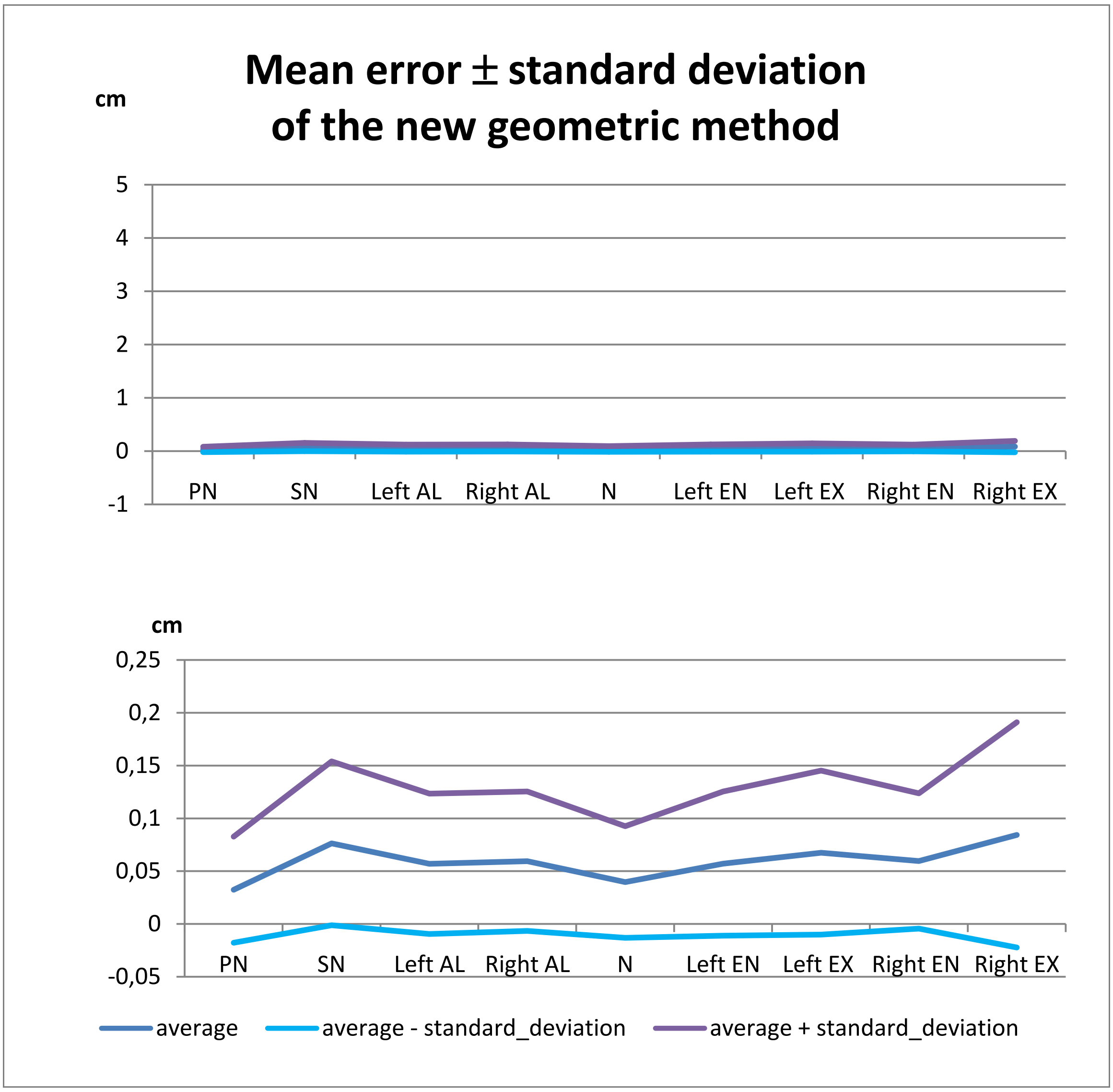


Figure 65. A graphical representation of mean error and standard deviation of the new method; (**top**) the axis scale is the same of the Figure 64, to better compare the two methods; (**bottom**) the same chart with a more suitable ordinate scale.

Figure 63 shows that, in the old method, the mean errors of the *exocanthions* are similar in the right and left parts of the face, while, in the new method, the values in the charts show a certain asymmetry between the two face sides. It may be due to the fact that, during the acquisition, the head was not perpendicular to the scanner, but slightly rotated to the left, which has led to a noisier acquisition of the right side; this has caused a bad detection of landmarks in the right side. The presence of noise in the right sides has somehow slightly altered descriptors behaviour. Furthermore, the *exocanthions* regions-of-interest and search regions are more narrow in the new method than in the old one. The alteration of descriptors behaviour together with this change of restriction conditions in landmark extraction give rise to the asymmetry in the results values of the right and left *exocanthions* of the new method. Another fact to point out is that, whereas the *exocanthions* are not critical points and so it is difficult to detect them, the *subnasal* may be occluded, since its region could be acquired incorrectly if the head is slightly tilted forward during the acquisition.

Concerning the quality of the results, all Figures 63, 64, and 65 point out that, in the new method, both the mean error and the variance error is generally lower than in the old one, with a deviation of one magnitude order. Numerically speaking, the mean error in the new method is lower than a millimeter for each landmark. In particular, the landmark that is detected more correctly is the *pronasal* (the mean error is about 0.03 cm), while the worst are the *subnasal* and the *exocanthions* (the mean errors are between 0.06 and 0.08 cm). Nevertheless, the mean error reaches similar values for all landmarks, keeping them in a range between 0.03 and 0.08 cm. So, it can be said that the mean error has been stabilized for all the landmarks. This stabilization means that, with this new method, the error has been "contained" with respect to the old one. In other words, all the landmarks may be considered as well-localized and more precise than in the previous method.

This behaviour could also be recorded with other graphical representations of the results, i.e. scatter plots and distribution functions for both the methods, shown in Figures 66, 67, 68, and 69. The scatter plot is a by-points representation on the Cartesian plane of the positions of the obtained landmarks. It is likely to put the found points in the same reference system, where the origin stands for the correct landmark position. The scattered points on the plane pretend to show the position of the points obtained by our algorithms, in particular the direction and the distance from the correct landmark.

The distribution function shows how many landmarks (n, on the ordinate axis) takes a particular distance value ($|e|$, on the abscissa axis). For simplicity sake, the distribution is not continuous but discrete. To obtain it, we discretized the set of error distances splitting them up into 50 short ranges for both the old and the new method results. A minimum and a maximum distance has been chosen: the minimum is equal to 0 and concerns the landmarks obtained by the algorithms which are in the same position of the correct landmarks; the maximum is equal to 5 cm and concerns the points whose distances from the correct landmark is equal or greater than 5 cm. Then, we subdivided the range between 0 and 5 in short intervals all equal to 0.1 cm, thus obtaining a discretization of the range of distances in 50 short ranges. Then, a graphical representation was done with the split-up distance range from 0 to 5 on the abscissa axis and the number of occurrences of the distance ranges on the ordinate.

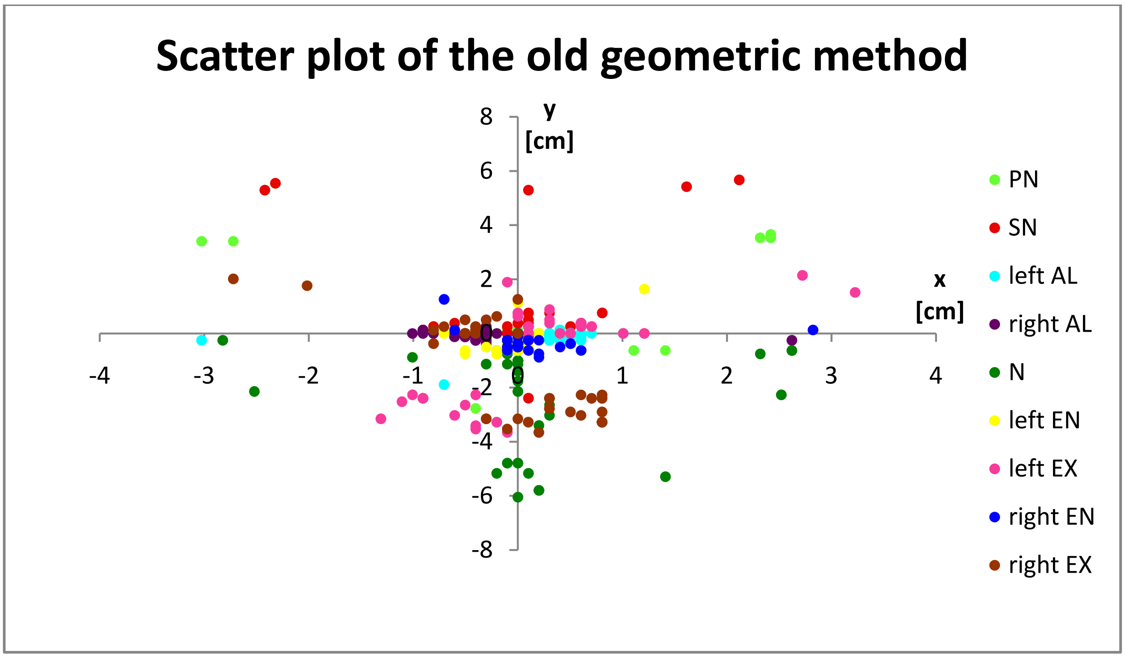


Figure 66. Scatter plot of the previous method results. The origin of the axis represents the correct location of every landmark. The scattered points are the positions of the landmarks obtained with our old algorithm. In this representation, the direction of the positioning of the obtained landmark is kept equal to the real one, while the absolute value of the distance the bi-dimensional distance on x and y axis, namely an approximation of the 3D distance.

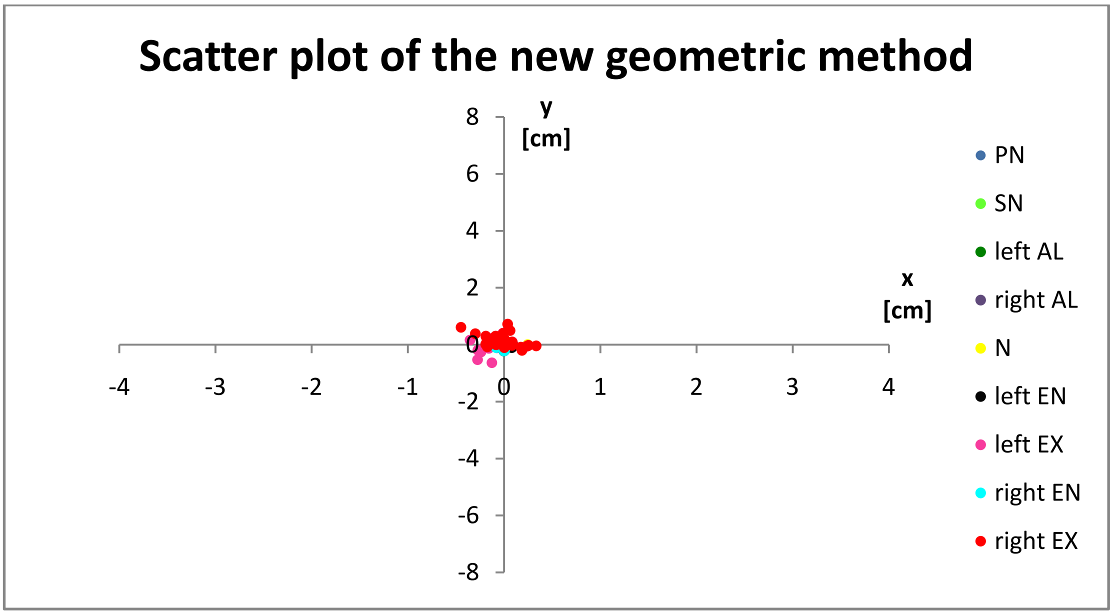


Figure 67. Scatter plot of the new method results. The origin of the axis represents the correct location of every landmark. The scattered points are the positions of the landmarks obtained with our new algorithm. In this representation, the direction of the positioning of the obtained landmark is kept equal to the real one while the absolute value of the distance the bi-dimensional distance on x and y axis, namely an approximation of the 3D distance.

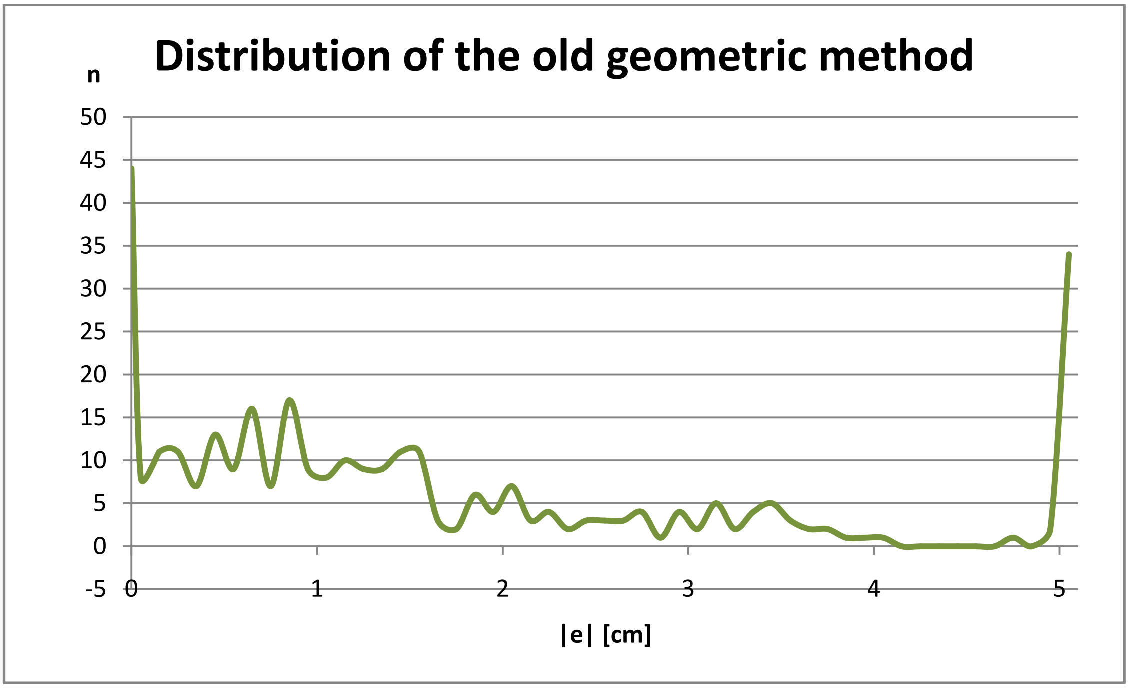


Figure 68. Distribution function for the results of the old method. The discrete distribution function shows how many landmarks (n, on the ordinate axis) takes a particular distance value ($|e|$, on the abscissa axis).

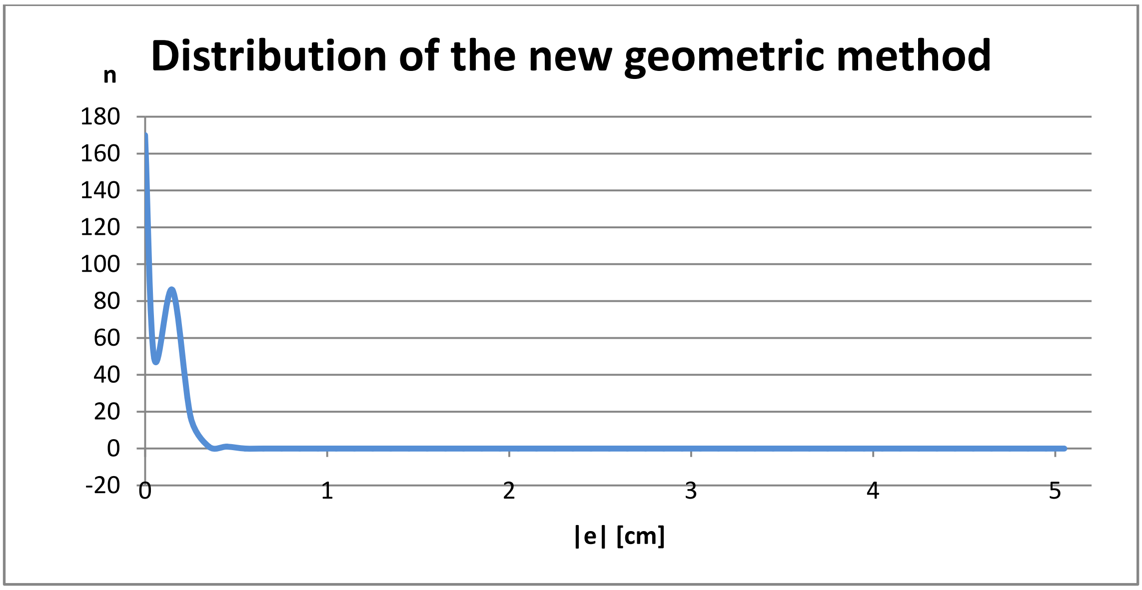


Figure 69. Distribution function for the results of the new method. The distribution function shows how many landmarks (n, on the ordinate axis) takes a particular distance value ($|e|$, on the abscissa axis).

As can be seen from Figures 66 and 67, while in the scatter plot of the previous algorithm the points are more sparse on the plane, with this new method the error has been kept below a certain limit and all the points could be considered as correctly localized. Figures 68 and 69 show the same respective behaviour of the previous ones. The distribution of the old method fluctuates around the value 10, namely there are roughly 10 landmarks for each distance range, meaning that many detected landmarks have a not negligible distance from the real one. On the contrary, the distribution of the new method approximately has an exponential decreasing; in fact, there are no landmarks whose distance between “true” and “obtained” is greater than 0.438821 cm. In particular, 170 distances are equal to zero. This confirms the extraction correctness and gives further confirmation onto goodness of the proposed method.

It was also asked to a plastic surgeon to judge our results. He stated that our obtained landmarks could be considered as correctly extracted in the faces. Furthermore, the value 0.438821 cm was considered as an acceptable and reliable threshold for a correct localization of the points of interest. Considering that most of the obtained landmarks have a distance from the real one which is much lower than this threshold, the surgeon confirmed the accurateness of the results.

5. Conclusion

In this paper we proposed an improvement to our previous geometric method. The general idea was to consider the face like a free-form surface and to use the Differential Geometry to detect and to extract the landmarks. Compared to our previous method, the algorithm has been refined with new geometrical descriptors and new conditions on the old ones. As a result, this new method is much more precise than the old one, in particular, in the worst case the error of the new method is about 0.5 cm, while, in the old method, some landmarks have not been even detected. As a further endorsement, a plastic surgeon confirmed not only the correction, but also the accuracy of the results.

6. References

[1] FRANTZ, S., ROHR, K. AND STIEHL, H. S. (1998) “Multi-Step Procedures for the Localization of 2D and 3D Point Landmarks and Automatic ROI Size Selection”, *Lecture Notes in Computer Science* 1406: 687-703.

[2] FRANTZ, S., ROHR, K. AND STIEHL, H. S. (1999) “Improving the Detection Performance in Semi-automatic Landmark Extraction”, *Lecture Notes in Computer Science* 1679: 253-262.

[3] FRANTZ, S., ROHR, K. AND STIEHL, H. S. (2000) “Localization of 3D Anatomical Point Landmarks in 3D Tomographic Images Using Deformable Models”, *Lecture Notes in Computer Science* 1935: 492-501.

[4] ALKER, M., FRANTZ, S., ROHR, K. AND STIEHL, H. S. (2001) “Improving the Robustness in Extracting 3D Point Landmarks from 3D Medical Images Using Parametric Deformable Models”, *Lecture Notes in Computer Science* 2208: 582-590.

[5] WÖRZ, S. AND ROHR, K. (2005) “Localization of anatomical point landmarks in 3D medical images by fitting 3D parametric intensity models”, *Medical Image Analysis* 10(1): 41-58.

[6] FRANTZ, S., ROHR, K. AND STIEHL, H. S. (2005) “Development and validation of a multi-step approach to improved detection of 3D point landmarks in tomographic images”, *Image and Vision Computing* 23(11): 956-971.

[7] ROMERO, M. AND PEARS, N. (2009) “Landmark Localisation in 3D Face Data”, *6th IEEE International Conference on Advanced Video and Signal Based Surveillance*, 73-78.

[8] ROMERO, M. AND PEARS, N. (2009) “Point-pair descriptors for 3D facial landmark localisation”, *IEEE 3rd International Conference on Biometrics: Theory, Applications, and System*,1-6.

[9] BEVILACQUA, V., MASTRONARDI, G., PIARULLI, R., SANTARCANGELO, V., SCARAMUZZI, R., AND ZACCAGLINO, P. (2009) “Experimental Comparison among 3D Innovative Face Recognition Frameworks”, *Lecture Notes in Computer Science* 5755: 1096–1105.

[10] RUIZ, M. C. AND ILLINGWORTH, J. (2008) “Automatic landmarking of faces in 3D-ALF3D”, *5th International Conference on Visual Information Engineering-IEEE Conferences*, 41-46.

[11] SANG-JUN, P. AND DONG-WON, S. (2008) “3D face recognition based on feature detection using active shape models”, *International Conference on Control, Automation and Systems-IEEE Conferences*, 1881-1886.

[12] NAIR, P. and CAVALLARO, A. (2009) “3D Face Detection, Landmark Localization, and Registration Using a Point Distribution Model”, *IEEE Transaction on Multimedia* 11(4): 611-623.

[13] SALAH, A. A. AND AKARUN, L. (2006) “Gabor Factor Analysis for 2D+3D Facial Landmark Localization”, *IEEE 14th Signal Processing and Communications Applications*, 1-4.

[14] D’HOSE, J., COLINEAU, J., BICHON, C., AND DORIZZI, B. (2007) “Precise Localization of Landmarks on 3D Faces using Gabor Wavelets”, *First IEEE International Conference on Biometrics: Theory, Applications, and Systems*, 1-6.

[15] VEZZETTI, F. AND MARCOLIN, F., “Three-dimensional Face Morphology Analysis: Soft-Tissue Landmark Formalization”, under review.

[16] CALIGNANO, F. (2009) “Morphometric methodologies for bio-engineering applications”, PhD Degree Thesis, Politecnico di Torino, Department of Production Systems and Business Economics.

[17] KOENDERINK, J. J. AND VAN DOORN, A. J. (1992) “Surface shape and curvature scales”, *Image and Vision Computing* 10(8): 557-564.

[18] PERAKIS. P., PASSALIS, G., THEOHARIS, T., AND KAKADIARIS, L. A. (2010) “Automatic 3D Facial Region Retrieval from Multi-pose Facial Datasets”, *Eurographics 2009 Workshop on 3D Object Retrieval*: 37-44.