

# Three-Dimensional Features for Facial Gestures Simulation

J.Finat<sup>1</sup> and A.Hurtado<sup>2</sup>

<sup>1</sup>Junior Professor, Computer Science Engineering, IST Building, University of Valladolid

<sup>2</sup>Young Researcher, MoBiVAP R& D Group, Scientific Park, University of Valladolid

## Abstract

Image- and range-based information provide a support for three-dimensional modeling of human faces. Their main objective is the generation of a bottom-up model given by a collection of facets which, after matching, can be applied to recognition, medicine or gestures simulation in multimedia applications. A three-dimensional model contains several layers including discrete clouds of points or superimposed piecewise linear (triangular or quadrangular meshes) or piecewise smooth models (usually obtained by smoothing PL-models, after smoothing edges). The availability of a piecewise smooth model allow to apply deformation methods for comparing different configurations by taking geometric features (appropriate points or meaningful curves) as control elements. Intrinsic geometric features are linked to curvature maps, and more specifically to parabolic curves which separate concave and convex regions, characterized by the local constancy of signature of the total curvature. Discrete models arising from points clouds and their superimposed PL-structures pose specific problems due to structural gaps between eventually discontinuous PL-structures and ideal PS-models, which concern to discrete inputs (including irregular distributions), incomplete character of information (requiring restoration relative to propagation models on surfaces bounding solids), automatic identification of intrinsic features relevant for surface segmentation (distinction between parabolic curves and apparent contours, appearing as critical locus for projections), and relations between different kinds of meshes and related transformations (arising from lifting quadrangular meshes in views), between others. All of them are meaningful for generating gestures from image- or range-based information. In this work we develop some tools which are focused towards: a) Estimation of topological invariants linked to local and global properties which can be read in terms of triangular or quadrangular meshes; b) Intrinsic character of features (curvature maps), and c) Detection and estimation of geometric features. Nevertheless the intrinsic character of geometric properties (linked to curvature maps), results obtained from appearances arising from meshes depend in a crucial way of the resolution chosen and how sampling procedure is performed. The three above problems are meaningful for estimating intrinsic data in terms of an approximate location of critical points (Morse Theory), interpretation of appearances in terms of intrinsic data (a discrete approach to the Geometry of Discriminant Locus for the projection map) and controlling propagation phenomena (Poisson models, in our case) in order to restore, estimate and modify facial gestures. We provide some illustrations of our results concerning to each above mentioned problem in terms of a discrete version of eventually Stratified Morse Theory (for including real or apparent singularities), an explicit description of curvature maps arising from local Monge form (local equation of surface at each point obtained from the closest points), Poisson-based propagation phenomena for volumetric restoration, grouping of vector references corresponding to principal curvatures, detection of parabolic curves and their application to generate gestures for scanned human true faces.

## 1 Introduction

Following a functional approach, the human face is composed of skeletal, muscular and skin superimposed layers which can be considered as components of a 3D model; the description of interrelations between the three layers becomes more complex when one intends to model facial gestures. A facial gesture is given by a collection of typical configurations of the three above subsystems along time. Modeling the coupling between the three layers is a very hard task and several approaches have been developed for simplifying it from long time ago ([11]). From a mathematical viewpoint there is an artificial hierarchy based on appearances which can be structured in different ways; following an increasing order of complexity, levels

of this hierarchy can be described in terms of fiducial points (critical points for some height or depth scalar function), meaningful curves (separating concave and convex regions) or maps of curvatures (principal curvatures or their arithmetic and geometric means, typically). It is not elementary how to correlate the functional (human subsystems) and structural approaches (mathematical models). A top-down approach with a previous knowledge of facial muscles, starts with a set of fiducial points (very often with landmarks for making easier their tracking); "aligned" fiducial points provide a support for constructing non-linear models based on NURBS [15] or B-splines [14]. Control points are submitted to internal or external forces whose characteristics (direction, magnitude) depend on location of points and muscle shapes; in [5] one can find a description which provides a good support for this approach, with a large complexity model because it involves to a lot of muscles in the human face. Accuracy of models is strongly dependent of the application to be developed: very accurate models are necessary for facial surgical simulations [16] or for face recognition [3], whereas coarser models are enough for animation effects in entertainment industries [24].

Alternately and in view of troubles for modeling different layers according to the above functional framework, one can develop a bottom-up approach nearer to appearance-based reasoning which is based in morphological external "aspects". Some meaningful work for appearances-based facial animation has been developed from the mid nineties [17], and more recently for facial expression simulation [26]. For an appearance-based modeling the most important aspects are features linked to *meaningful curves* which "separate" regions bounding the human head or *fiducial points* lying on such meaningful curves. Surface segmentation is a decomposition of the face in regions with common characteristics from the geometric viewpoint: a coarse description of surface segmentation takes in account concave or convex regions which are separated by the above meaningful curves, whereas a finer description will give us a map of curvatures [25]. Separatrices given by parabolic curves, and regions bounded by them, are crucial curves for simulating facial expressions. Indeed, meaningful curves provide the support for propagation models and curvature maps imposes constraints on the allowed variations of regions for a human face. Geometric features for recognition and gestures simulation can be  $0D$  (fiducial points),  $1D$  (isocurves, parabolic curves, curves of mean-total curvatures) or  $2D$  (concave-convex regions, curvature maps); all of them have been used from the early nineties. Gaussian curvature has been applied by Moreno et al (2003) for generating a regional segmentation of a 3D model for human face; by using the signs of the mean and total curvatures, they detect a lot of regions (in fact 86) with a large computational effort, which is unpractical for simulating human gestures in applications requiring an interactive response.

In more advanced models from the mathematical viewpoint, the synthesis of human gestures can be posed in terms of elastic models given by Partial Differential Equations (PDE) with boundary constraints; in [9] boundary curves are extracted automatically using  $3D$  data of human faces obtained using a  $3D$  scanner. Nevertheless the mathematical elegance of this approach we have not still a "dictionary" between typical facial gestures and tensions at meaningful elements. In absence of this dictionary, for simulating facial gestures it is still necessary to perform a semi-automatic strategy based in manual interaction. In our more elementary approach, boundary constraints are supported by mobile curves which "separate" concave and convex regions. Convex (resp. concave) are characterized by a positive (resp. negative) sign for the total or Gaussian curvature  $\kappa_t$ . Hence, parabolic curves (given by the vanishing of  $\kappa_t$ ) are the most meaningful objects for separating regions following a bottom-up approach and the first goal to achieve is the semi-automatic generation of a map of parabolic curves. Simulation of gestures can be performed by discretizing parabolic curves and introducing like-gradient forces at a discrete collection of points belonging to parabolic curves.

Gradient fields provide a simple common framework for modeling, giving isocurves and isosurfaces (as in the Classical Morse Theory framework) linked to scalar functions as a by-product useful for recognition of static shapes. Similarly, deformations linked to the minimization of energy functionals defined on a superimposed mesh provide a gradient-like framework for kinematic modeling of facial gestures with bounded variation. The feedback between static and dynamic modeling is very common in Robotics, but its use has been comparatively less developed in facial gestures simulation. Intuitively, a normalization of the gradient field (unit normal vector at each point) and its first order variation allow to recover features of the facial surface including behavior at critical points and vector representation (the Weingarten matrix represents in fact the Hessian of a local representation of the surface). Intrinsic features are given locally by principal curvatures; their estimation can be performed from discrete or continuous information. Nevertheless the available discrete support, irregularities in the distribution of clouds of points and the consequent lack of

robustness for a pure discrete modeling, suggest a continuous approach based in our case on (i) a local fitting of a surface of low degree given by the Monge's form to the nearest points in a discrete neighborhood of each pivot point; (ii) an identification a local distribution of vector fields whose elements represent principal curvature lines or, alternately, total and mean curvature lines (for smoothing of such vector distributions see); (iii) an adjustment of deformable shapes to virtual displacement of control points located on parabolic curves.

A crucial issue for connecting static and dynamic modeling is the robust identification of curves supporting the distribution and propagation effects corresponding to simple expressions in human face. For us, this role is played by *parabolic curves*, i.e., curves along which the total curvature is null, which "separate" convex and concave regions in human face. A simplified map of the human face in terms of the sign of total curvature gives a collection of regions adjusting themselves to regions with constant sign for total curvature; the boundaries of each region are given by space parabolic curves. In our approach, most of these regions are of quadrangular type. The complexity of facial gestures requires to specify the combination of effects corresponding to different muscles; in other words, forces supported on parabolic curves are coupled between them, in such way that a contraction or an expansion of one of them generates similar effects in the nearest parabolic curves according to dynamical parameters (stiffness or damping matrices) linked to the muscle elasticity. Efforts from such curves are propagated following transversal directions (which are generically "almost" orthogonal) to parabolic curves; propagation is constrained by behavior of "opposite" side of quadrilateral following typical reactive biomechanical models. Furthermore, elasticity of skin covering muscle-skeletal subsystems is not the same as muscular elasticity, and this disparity is the responsible of folding effects in skin which are characteristic ones of facial gestures, and are perceived as wrinkles in the skin, more or less remarkable depending on (real or simulated) skin elasticity.

This paper is organized as follows: Section 2 is focused towards inputs and static modeling arising from laser scanning of human faces in our case. Next, section 3 introduces some variants of Digital Classical Morse Theory (CMT) which are adapted to pathologies of functionals defined on a PL-model of human face. The fourth section is devoted to pose a framework for synthesis of facial gestures, able of integrating geometric, kinematic and dynamic aspects linked to scalar, vector and tensor functionals which are typical in CMT corresponding to height, energy functional or curvatures, e.g. A natural feedback, including pathologies linked to singular points of functionals linked to deformations (and subsequent wrinkles), requires some extensions CMT for dynamic modeling which are sketched in section 5. Some technical details involving the shape estimation and a Poisson-like propagation model are displayed in the section 6. A lot of work remains to be done; some open issues are presented along the section 7, where some possible extensions are presented.

## 2 Data Capture and discrete static modeling

This section is focused towards the semiautomatic generation of models from range data arising from laser scan devices. A large amount of recent work is referred to 3D face representation which is focused towards recognition issues. The 3D representation can be constructed from a collection of simultaneously captured images (from stereo devices) or range information (arising from scanner devices). In this work, the input is given by a cloud of points for each face which has been taken with a Minolta 910 laser scan in a full frontal pose. The input is given as a cloud of several thousand of points; a geometric model is superimposed to the cloud of points by inserting edges connecting nearest points, following some variant of ICP algorithm, giving a 3D triangular morphable mesh  $M$  which is taken as a piecewise-smooth  $M_S$  can be associated. It is necessary to extract different kinds of features from  $M$  or  $M_S$ . Range-based information provides robust models which can be real-time updated [19].

Three dimensional data capture of human face can be performed by using laser scans or range image devices: (a) for laser scan devices inputs are given by a cloud of 3D colored points; (b) for range image, the  $z$ -coordinate is obtained from an interpolation linked to a stereo device with a large diversity of algorithms. Some recent contributions focused towards Recognition ([12]) are based on a combination of the above techniques with algebraic-statistical analysis (Linear Discriminant and Principal Component) giving good results but a higher computational effort. In this work, we restrict ourselves to 3D data arising from scanning a human face.

From the resulting cloud of 3D points  $\mathbf{P}_i$ , a triangular mesh is generated following different algorithms; between them Iterative Closest Points (ICP in the successive) is used also for aligning data and matching different scans. For each triangle  $\langle \mathbf{P}_i, \mathbf{P}_j, \mathbf{P}_k \rangle$  the unit normal vector  $\mathbf{u}_{ijk} = \mathbf{n}_{ijk}/\|\mathbf{n}_{ijk}\|$  stores the initial information for meaningful geometric features of human face; a measurement of local variation of normal vector provides the intrinsic invariants of human face in terms of principal curvatures  $\kappa_1, \kappa_2$  as eigenvalues of the curvature matrix  $K$ . Coefficients  $tr(K), det(K)$  of the characteristic polynomial  $|K - \lambda I|$  of the curvature matrix are the most common algebraic invariants  $2\kappa_m, \kappa_t$  corresponding to the mean and total or gaussian curvatures.

To start with, we are only interested in geometric properties of the resulting mesh to low resolution levels, with a special attention to morphable aspects. Geometric invariance for each volumetric object  $V$  is linked to intrinsic properties for the surface  $S$  bounding the solid, i.e.  $\partial V = S$ ; principal curvatures  $\kappa_1, \kappa_2$  (or alternately, mean  $\kappa_m = \frac{1}{2}(\kappa_1 + \kappa_2)$  and total  $\kappa_t = \kappa_1\kappa_2$  curvatures) are the main local invariants for a piecewise smooth surface  $\mathcal{M}$  at each point  $\mathbf{P}$ : they characterize the surface at each point up to a similarity transformation. This result has been used in several works for solving recognition issues from image- or range-based models of human faces from the mid nineties [INSERTAR REFERENCIAS!!]. In this work, we paid more attention to maps of curvatures for controlling human face deformations linked to the gestures synthesis.

If  $r_{ijk}$  is the line through barycenter with director vector  $\mathbf{n}_{ijk}$  and  $\mathbf{Q}_{ijk} := r_{ijk} \cap \mathcal{M}$  is the nearest point to  $\mathbf{B}_{ijk}$ , then  $\mathbf{n}_{ijk}$  approximates the "true" normal vector  $\mathbf{n}(\mathbf{Q}_{ijk})$  to the ideal PS-surface  $\mathcal{M}$  at the point  $\mathbf{Q}_{ijk}$  (which is usually unknown). Ideally in the PS-framework), the pencil of planes  $\mathcal{H} = (H^{(\alpha)})$  through the line  $r_{ijk}$  cuts out on the surface  $\mathcal{M}$  a pencil of plane curves  $C_{\mathbf{P}}^{(\alpha)}$ ; extremal values (maximum and minimum) of sectional curvatures correspond to the principal curvatures  $\kappa_1, \kappa_2$  which are the eigenvalues of the curvature matrix  $K_{\mathbf{Q}}$  at  $\mathbf{Q} \in \mathcal{M}$ ; consequently,  $det(K_{\mathbf{Q}}) = \kappa_1\kappa_2$  and the trace  $tr(K_{\mathbf{Q}}) = \frac{1}{2}(\kappa_1 + \kappa_2)$ , as coefficients of the polynomial characteristic  $|K_{\mathbf{Q}} - \lambda I|$ , are also invariants of  $K_{\mathbf{Q}}$  and are called total and mean curvatures at  $\mathbf{Q}$ , respectively. Let us remark that in the discrete approach, one replaces

- the true normal vector  $\mathbf{n}$  at  $\mathbf{Q}_{ijk} \in \mathcal{M}$  by an approximation given by the normal vector  $\mathbf{n}_{ijk}$  at the barycenter  $\mathbf{B}_{ijk}$  of a triangle generating 2-dimensional secant plane "near" to the true tangent plane  $T_{\mathbf{Q}}\mathcal{M}$ ;
- the pencil of planes  $\mathcal{H} = (H_{\mathbf{Q}}^{(\alpha)})$  through  $\mathbf{Q} \in \mathcal{M}$  by  $\mathcal{H}_{PL} = (H_{ijk}^{(\alpha)})$  throughout the almost normal directions  $\mathbf{n}_{ijk}$  which cuts to the PL-approach  $\mathcal{M}_{PL}$  along a collection of plane polygonals  $C_{ijk}^{(\alpha)}$  near to  $C^{(\alpha)}$ ;
- the true piecewise surface  $\mathcal{M}$  bounding the volumetric object by low-degree algebraic surface which can be constructed by using the Monge Form which is described in section 6.

### 3 Morse Digital Theory for volumetric description

This section is focused towards the extraction of "geometric features" from models constructed in the precedent section.

Deformation of meshes for artificial gestures synthesis preserving the global topology requires to identify control points and meaningful curves, which are "critical" for some functional. Traditional Morse theory (MT) reconstructs volumetric objects by adjoining cells at critical points of functionals; in the simplest case corresponding to a Morse function, the dimension of each cell to be adjoined depends on the signature of the Hessian at each non-degenerate critical point. Typical Morse functionals are given by a scalar function (height or depth, e.g.), a vector functional (energy, e.g.) or a tensor functional (curvature, e.g.). In this work we develop an appearance-based approach which is modeled following some general principles of Digital Morse Theory (DMT) which is a discrete version of traditional Morse Theory. DMT is given by a piecewise-linear approach to MT which is supported by a mesh constructed from a collection of points  $\mathbf{P}_i = (X_i, Y_i, Z_i)$  arising from range techniques, e.g. Most common scanning of faces are taken following a front-parallel orientation; in this case, the selection of critical points with respect to the relative depth (w.r.t. the scanner laser device) provides a collection of the most meaningful points which can be grouped in relative minima (nose, mouth, chin, e.g.) or relative maxima (eyes, ears, e.g.); these points are commonly chosen as fiducial points, including landmarks in vision-based approaches.

### 3.1 Some ideas of Stratified Morse Theory

Stratified Morse Theory (SMT in the successive) is well suited for reconstructing singular analytic varieties from a well specified stratification, where strata verify "good adjacency conditions" for limits of tangent and secant planes ([13]). In Classical Morse Theory (CMT), the smooth manifold  $M$  is reconstructed by adjoining *cells* (topologically equivalent to open disks) whose dimension is linked to the signature of the Hessian at each critical point; from the topological differential viewpoint, a compact manifold is given as a CW-complex by a finite collection of cells and matching procedures along critical points of Morse functions which are propagated by using transversal directions to level hypersurfaces. Instead, in SMT the singular stratified variety  $X$  is reconstructed by adjoining *pairs of pseudo-cells* representing Tangential and Normal Morse Data (TMD and NMD, in the successive) which are adjoined as "links" (see [Gor88] for details). An analytic variety as a stratified space is given by a finite collection of tangential and normal Morse Data, and a collection of links for local matching.

### 3.2 An application to dynamic modeling of facial gestures

For simulating facial gestures, if we forget holes (corresponding to mouth, nose, ears), the human head is a topological sphere as manifold. Hence, from the static viewpoint it can not be considered as a singular variety. However, from the dynamical viewpoint, facial gestures introduce "singularities" corresponding to wave fronts which are propagated from parabolic curves separating regions with positive and negative total curvature. Wrinkles are the most characteristic features for gestures recognition; wrinkles can be interpreted as apparent singularities linked to a discrete propagation wave from parabolic curves on the human face. After activating the gesture, its propagation along orthogonal directions to some parabolic curves follows a dissipative pattern.

From the support viewpoint, there is a feedback between geometric (tangential data linked to the support) and dynamic aspects (normal aspects linked to propagation phenomena) of facial gestures. The feedback between geometric and dynamic is usually formulated in Lagrangian terms, involving "generalized" coordinates  $(\mathbf{q}, \mathbf{p})$ , with the contact constraint  $\mathbf{p}dt - d\mathbf{q} = 0$ . Following this model, the geometry of stable configurations (minimal energy) do act as restoring forces for recovering neutral facial expressions. A simple model for recovering the original shape is given by a *spring-damper system* (Hooke) given as  $\mathbf{F}_\ell = -B_\ell(\mathbf{x}_\ell - \mathbf{x}_{\ell_0}) - D_\ell\dot{\mathbf{x}}_\ell$  where  $B_\ell$  is the *stiffness matrix*, and  $D_\ell$  is the *damping matrix* associated to each facial gesture.

First-order variation of  $\mathbf{q}$  along time giving  $\mathbf{p}$ , could be measured from a discretization of video sequence analysis or, alternately; in our case, it arises from integrating the behavior of forces  $\mathbf{F}$  acting on each control point along a direction (in the geometric framework is given as the "Lie contraction" of a 2D distribution along a direction). In practice, we have only information about appearances of facial gestures and no information about quantitative measures  $\mathbf{p}$ , still. Because of this limitation, we have chosen an even more classical approach based on the total work  $W$  performed by the system. Classical formulation gives the total *scalar work*  $W_s = \mathbf{F}d\mathbf{q}$ . In the extended Geometric Algebra or Clifford framework, a first approach to the total *geometric work* is defined as the "mixed" quantity  $W_g = \mathbf{F}d\mathbf{q} + \mathbf{F} \wedge d\mathbf{q}$  where  $\wedge$  represents the cross product. The scalar work as scalar product is the contraction of the force vector along the direction corresponding to the displacement. Due to the lack of space and to avoid additional mathematical details, we shall limit ourselves to consider the scalar component  $\mathbf{F}d\mathbf{q}$  as a "tangential" component along the space direction where the force  $\mathbf{F}$  is acting on, and the bivector component  $\mathbf{F} \wedge d\mathbf{q}$  as a "normal" component representing a "privileged direction" contained in the planar element along which deformation are holding on.

In discrete models and for applying a finite element method, it is necessary to identify the range where forces are acting on. It is clear that the variability range depends on personal character and morphological characteristics of human face; in absence of a psychological profile assigned to virtual actors and following our approach, morphological characteristics for facial gestures are bounded by the adjacent parabolic curves. Hence, the damping matrix  $D_\ell$  must be multiplied by an almost Gauss-shaped function following the direction of forces irregularly distributed along parabolic curves. Such forces are issue of a reference point given by the intersection of (at least two parabolic curves; in absence of a true tangent plane for dynamical singularities related to facial gestures, forces are propagated following a spray whose support is theoretically modeled as a tangent cone at each control point which is activated for gestures simulation of human face.



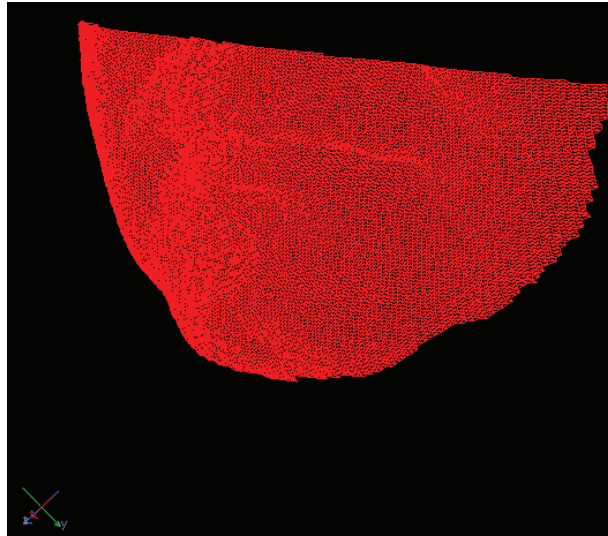


Figure 1: A mesh around commissures of lips where deformation takes place

## 4 A framework for synthesis of facial gestures

Our main goal is the synthesis of facial gestures from neutral expressions by using deformation theory. To avoid problems with self-occlusions and incomplete information, a common technique is based on warping which has been used in early stages of Computer Graphics [4]. This approach is suggested by the parametric description of deformations. Nevertheless accurate results for the static approach (quiet expressions in absence of external or internal forces), this approach does not take into account the effects linked to the different responses of muscle and skin components linked to the functional approach. To incorporate them, a more complete picture must take into account energy-based or forces-based models.

### 4.1 Energy-based models

Neutral expressions (passive, quiet, worried, sadness) are characterized by local minimal energy of the mesh superimposed to the model and, consequently; as local minima for energy functionals, they correspond to local stable configurations for superimposed meshes. Typical extreme gestures (smiling, happiness, laugh, astonishment/surprise, angry, enojo) are characterized by maximal energy at nearest nodes compatible with activation impulses bounded by maximal local deformations of the mesh; contrarily to local minima for energy functionals, they correspond to local unstable configurations for superimposed meshes.

In the classical approach, the integration of total work performed by forces acting on different control points, gives the energy functional defined on the mesh. The combination of repulsive and attractive forces between facial gestures by one side, and the dissipation effects linked to shape characteristics of human face and gestures, on the other hand, suggests the choice of a hybrid energy functional with components linked to external attractive forces and internal or elastic forces which tend to restore the initial displacements. In the following subsections we develop an alternate approach based on forces and displacements produced by them.

### 4.2 Dynamical modeling in terms of parabolic curves

Forces in charge of facial dynamics can be attractive or repulsive ones around a collection of control points or small pieces of curves corresponding in our case to pieces of parabolic curves of facial surface. Parabolic curves are defined by the vanishing of Gauss or total curvature. Hence they "separate" positive (convex) and negative curvature (concave) regions. Nevertheless small size modifications, the relative distribution

between concave and convex regions in human face is not modified along facial gestures. Values of curvatures are modified along facial gestures, but the sign of total curvature remains constant. This simple remark justifies the use of parabolic curves as "control curves". Hence, it suffices to identify, simulate and control deformation processes which are issue from parabolic curves for each facial gesture.

Usually, position changes of fiducial points generate small displacements at neighbor points in the mesh; these displacements are simulated in terms of forces which are propagated along edges connecting each fiducial point with the nearest levels in discrete models. Following this model, there is also a natural hierarchy for controlling facial expressions which can be performed in terms of fiducial points for the applied forces; this control must be translated to local propagation methods on discrete meshes. The discrete model for propagation is limited to a number  $\nu$  of the nearest points to control points located at meshes. The parameter  $\nu$  can be modified by the user, and imposes constraints on the energy distribution between neighbor levels following the opposite of a distribution quasi-normal function.

There are two computational approximations to 3D modeling of human faces which can be labeled as volumetric or surface elements: a) *Volumetric modeling* is obtained in an iterative way by carving decreasingly size cells (following octrees algorithms) and "sculpting" skew planes (identified and managed from marching cubes algorithms, e.g). b) Alternately, *surface-based modeling* starts with information arising from images or range information, and generates a 3D model by processing and analyzing image- and range-based information. The resulting model must be compatible with internal structure (muscle-skeletal system) and/or external (appearances) characteristics of human faces. In this work, we restrict ourselves to an appearance-based bottom-up approach with parabolic curves as control elements for dynamic effects.

## 5 Extending CMT for dynamic modeling

The crucial hypothesis of Classical Morse Theory (CMT) for reconstructing the static shape of a smooth compact connected manifold  $M$  is the non-degeneration of  $f$  at critical points  $\mathbf{P}_c \in f^{-1}(c)$ . This hypothesis allows to identify the nature of topological cells to be attached at each critical point for a Morse function and, hence, to recover the global shape from matching local data given by isosurfaces. Roughly speaking and if we forget holes corresponding to mouth and nose for shape modeling, human head is topologically equivalent to a sphere. Hence, CMT is relevant for reconstructing smooth models of human head, and most shape reconstruction could be ideally performed by small deformations of isocurves guided by control curves; in more formal terms, isosurfaces are isotopic and can be connected between them by means of integral curves of transversal fields to the boundaries to the boundaries. This approach is already well known, but it ignores peculiarities linked to the most meaningful elements for gestures simulation linked to parabolic curves which are responsible of human expressions. To overcome this difficulty there are (at least) two meaningful topological techniques:

- Introduction of *vector functionals* (energy) or *tensorial functionals* (curvatures) instead of scalar Morse functions (height, depth); two natural candidates are a *spatial distribution of energy* or *curvature tensor* defined locally by the curvature matrix. In our case, the invariants of the curvature matrix are given by mean and total curvature vector fields.
- *Analysis of extremals or singularities for functionals*: They are supported by parabolic curves which are confluent at their extremes (lips and eyes, for the human face); usual Morse functions (height or depth, typically) are regular ones on these locuses.

In both cases, parabolic curves are the crossing-zero locus for the total curvature understood as a vector field defined on the surface bounding the human head. Up to elementary non-degenerate singularities (one maximum and one minimum) for scalar Morse functions (height, depth), there are no other meaningful singularities for scalar functionals. However, the situation changes if we consider zero-crossing of curvature functionals (which in local terms would correspond to degenerate Hessian forms); in simulating facial gestures, it is possible to identify dynamic singularities of propagation phenomena whose temporal variations (wrinkles) allow to identify human gestures.

The discrete analogue of a smooth slice is given by a *volumetric interval* between two isolevel surfaces; as PL-object it can be decomposed in a union of adjacent tetrahedrals between two PL-isosurfaces. The

retraction of an "horizontal" volumetric interval of triangles to an intermediate connected polygonal provides a discrete interpolation between level sets corresponding to the upper and lower envelopes of solid stripes. Hence, the intermediate polygonal can be interpreted as a discrete deformation retract, which is called the "nerve" of the solid stripe. Inversely, a local solid stripe can be recovered from its nerve by attaching a handle with control points given by neighbor vertices. In this case, propagation method follows an nearest searching algorithm on a solid stripe of the original mesh with a similar pattern for the smooth case. Unfortunately, this analysis is not enough for dynamic singularities.

To start with, let us remember that at a singular point (typical in SMT) there is no a tangent plane but a *tangent cone*, and one needs to have an information about degeneracy conditions at singular points. A discrete approach to the simplest punctured tangential cones appearing in SMT is illustrated in the following picture: This information is contained in a retraction of the punctured tangential and normal Morse Data given by the boundary of small (product of) disks around the singularity. Roughly speaking, both data are given as pairs of topological pieces with a very simple product structure (not necessarily smooth). Hence, an important topological difference between the Classical Morse Theory (CMT) and Stratified Morse Theory (SMT) concerns to basic pieces to be patched: In CMT they are ordinary cells (balls or products of intervals), whereas in SMT furthermore ordinary cells, one has conical singularities. Conical singularities of human face do correspond to commissures of lips and eyes.

### 5.1 Data Estimation for cones

Singularities appearing in the adherence of strata are characterized by normal and tangent *cones*. Thus, data extraction of cones plays an important role for SMT. From a theoretical viewpoint, Normal and Tangential Morse Data are contained in sections of cones by disks outside the vertex of the cone. A discrete version of products of disks is given by products of volumetric intervals or "slabs". So, horizontal slabs can be considered as a discrete version of a cylinder (as "fat" PL-parallels) which is a product of disks; however, vertical slabs can have a piecewise linear conical structure (as "fat" PL-meridians). Labeling of different types of triangles of an alpha-shape provides the construction of parallel and meridian slabs:

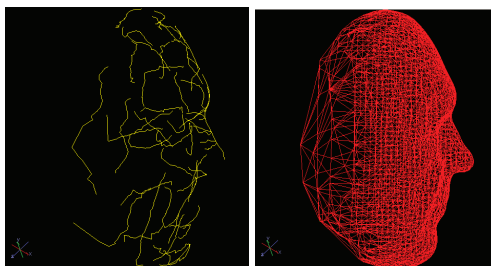


Figure 2: A mesh around commissures of lips where deformation takes place

## 6 Shape Estimation: Local Properties

Topological analysis arising from CMT provides a very useful information about global topology of a coarse model, including an estimation of topological genus. However, it does not allow to discriminate between appearances or metric information which are very useful for shape characteristics. The low-level meaningful appearances are concave and convex regions, but concavity estimation is a NP-hard problem and volumetric segmentation in terms of concave and convex regions. Metric information involves to the estimation of curvatures which is performed in terms of interpolations from local Monge forms at each triangle of the mesh.



## 6.1 Local Monge forms and appearance estimation

By taking local coordinates around a critical point and with the axis  $Oz$  away from tangent plane, the Taylor expansion of a local representation for a surface around a point is represented (up to infinitesimal transformations) by its  $k$ -jet which is infinitesimally equivalent to a polynomial of degree  $d \geq 2$  in 2 variables, i.e., the  $d$ -jet has a local form  $z(x, y) = J_{B,d}(x, y) + h.d.t.$  (h.d.t.: higher degree terms), being

$$J_{B,d}(x, y) = \sum_{k=0}^{k=d} \sum_{i=k}^{i=0} \frac{B_{k-i,i} x^{k-i} y^i}{i!(k-i)!}$$

with  $B_{00} = B_{10} = B_{01} = 0$  due to the choice of local coordinates. For a non-umbilic point, the quadratic part is diagonalizable and the 4-jet is given by the truncated Taylor polynomial of degree four:

$$z(x, y) = \frac{1}{2}(k_1 x^2 + k_2 y^2) + \frac{1}{6}(b_0 x^3 + 3b_1 x^2 y + 3b_2 x y^2 + b_3 y^3) + \frac{1}{24}(c_0 x^4 + 4c_1 x^3 y + 6c_2 x^2 y^2 + 4c_3 x y^3 + c_4 y^4) + h.d.t.$$

which is called the *standard Monge form*, being  $k_1, k_2$  the principal curvatures (distinct eigenvalues for the quadratic part). Degree four polynomials allow to estimate concavity and convexity directly from control points, and provide a first approach to the local behavior of curvatures around each point. Standard Monge form provides a locally homogeneous representation for the map of curvatures, without using regularization or smoothing operators. The restriction relative to the degree of polynomials is adopted from noise considerations. Higher degree polynomials are possible, but the presence of noise and some troubles for matching together surfaces of higher degree make very difficult to obtain a global curved model.

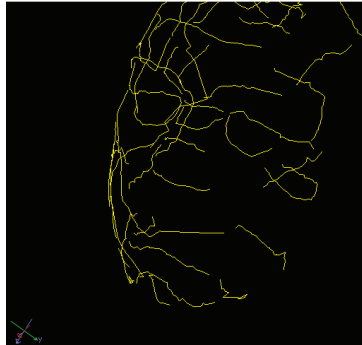


Figure 3: Local constancy of curvatures along a solid stripe

In particular, the Monge local form allows to verify the *local constancy* of curvatures along horizontal or vertical solid stripes of cones, and consequently detect pathologies corresponding to the (un)folding of cells or bifurcations of vector fields linked to propagation phenomena in presence of singularities. This example displays a locally symmetric behavior for the map of curvatures along cones which are the basic piece for reconstructing more complex objects following a discrete version adapted to SMT methodology. Far from being casual, the displayed example plays a fundamental role in modelling complex objects, not only from the topological viewpoint (SMT), but also from the differential geometric viewpoint.

## 6.2 Poisson propagation model

Gradient field manipulation of meshes superimposed to the human face can be performed with Poisson Propagation Model (PPM). Contrarily to the usual Laplacian-based methods, PPM is rotationally invariant which is an advantage for visualizing complicated meshes and, by construction it supports a natural propagation model. However, it displays some troubles regarding interpolation which require the introduction of additional optimization criteria. Simulation of facial gestures require to specify the "limits" which involve to

- The "weighted parabolic curves" bounding the adjacent regions to the current parabolic curve which would represent the "regions of influence" for activation/inhibition phenomena;
- the strength fields responsible of translations of fiducial points which are responsible of facial wrinkles

In our scheme, the connection between weights (geometric aspects) and strengths (dynamical aspects) will be given by an energy function (kinematic aspects) which can activate/inhibit one level or two levels of the adjacent regions to the parabolic curve. This kinematic feedback gives variable weights  $w_i$  corresponding to related gestures: the common regions for smiling and laughing are deformed with variable intensity depending on a sigmoid-shape energy functional which is defined on the nearest vertices to the mesh with deformations following orthogonal directions to the parabolic curve of reference. In practice, these variations can be simulated by means of the homotopy or continuation method involving to the behavior of curves associated to extremal gestures following discrete relaxation methods.

## 7 Discussion and possible extensions

Input data arise from once a scan of a human face captured with a Minolta 910. In our case, the cloud of points involves only to frontal aspects, and there is no information about the rear. Because of this, the application of Poisson process fills out the rear part by generating very large triangles which are not trustworthy. Local shape can be inferred in two quite different ways: from osculating spheres linked to curvature elements (Monge form) or from parameterization arising from lifting a regular grid to a surface (bicubic interpolation). We have followed the first approach, but for blending purposes, a bicubic interpolation could be more efficient: a local bicubic interpolation needs two transversal directions and four control points for each direction, which can be taken as the consecutive vertices of three consecutive triangles along a space coordinate curve. Such coordinate curves are given in our case by a discrete approach to principal curvature lines, corresponding to maximal and minimal variation of principal curvatures. In an experimental way, we have found that maximal expressiveness corresponds to opposite movements (shortening and widening) along both principal directions.

Some aspects which have not been considered in this work and which are currently in different stages of development concern to

- *Representation and management in meshes*: Usual meshes for image-based information is given by quadrangular meshes, whereas usual low-resolutions meshes for range-based information are given by triangles. Even if we apply re-meshing strategies, the comparison between quadrangular and triangular meshes requires some refinements of usual harmonic analysis (as solutions for propagation phenomena controlled by laplacians defined on smooth variations). Linked deformations needs a discrete version of *conformal mappings* (preserving angles) for the management of quadrangular meshes. Similarly to the methodology developed in 3D Reconstruction where objects are replaced by transformations, it is necessary to develop an efficient computational implementation for conformal transformations of  $2D$  and  $3D$  meshes, and relations between them. In a similar way to the  $2D$  case, laplacian deformation preserves volume, but it is not rotationally invariant, giving some errors for advanced visualization. Hence, it is necessary to introduce external constraints involving boundary surfaces in terms of offset meshes.
- *Optimization of  $2D$  quadrangular and  $3D$  triangular meshes* linked to propagation models on surfaces with a complicated shape to different resolutions. A first formulation concerns to the lifting to  $3D$  objects of integral curves corresponding to distributions linked to quadrangular meshes; in this work, we have considered only some vector fields corresponding to mean and total curvature lines, but it is necessary to develop a real-time feedback with information arising from range images. This goal implies to identify and characterize the behavior of vector fields linked to apparent contours and their plane projections in terms of the curvatures variation on the surface  $S$ .

## 8 Conclusions

A three-dimensional geometric decomposition of appearances for human face can be described in terms of the sign of Gauss or total curvature  $\kappa_T$ . A local estimation of principal curvatures is performed in terms of Monge forms. Parabolic curves separate elliptic and hyperbolic regions. The union of parabolic curves provides a map which is used for identifying facial regions and obtaining an appearances-based segmentation in an automatic way. The simulation of facial gestures is performed in terms of a feedback between geometric and dynamic aspects in terms of the geometric work  $W_G$ . The geometric work is given by the sum of a scalar and a bivector quantity representing tangential and normal data. This dynamic approach is restated in terms of Stratified Morse Theory. This framework allows to include pathologies linked to traditional geometric and dynamic aspects such as different types of singularities or the support of vector fields superimposed to a surface  $X$  bounding a rigid body  $Y$  which are bifurcating along certain curves. In our case, we take  $Y$  as the frontal part of human head,  $X$  as the human face, and the distribution of the total and mean curvatures as meaningful vector fields on the human face  $X$ . Vector fields are modified by the application of forces which are propagated along parabolic curves and their orthogonal directions following typical dissipative patterns. Bifurcations are visualized as wrinkles of the human face, which are characteristic ones of facial expressions. In the current state of our approach, wrinkles are generated in a manual way. In the next future, we hope to develop a semi-automatic generation of facial gestures by means the control of parameters linked to a discretization linked to an appropriate discretization of parabolic curves and the application and the application of forces at meaningful points of such discretization.

*Acknowledgments:* CGAL 3.3.1 has been used for several developments and applications regarding PL-constructions (alpha-shapes, e.g.) and computational differential geometric issues (including Monge form, e.g.). MeshLab provides an excellent visualization for maps of curvatures, and it is in the origin of own software developments regarding further computations which involve eventually singular objects. From financial viewpoint, this work is partially supported by the "Proyecto Singular Estratégico PATRAC (Patrimonio Accesible: I+D+i para una cultura sin barreras)", PS-380000-2009-2 of the Spanish Ministry of Science and Innovation which has been co-funded with FEDER grants.

## References

- [1] AIM & SHAPE: "Advanced and Innovative Models and tools for the development of semantic-based systems for handling, acquiring, and processing knowledge embedded in multidimensional digital objects". FP6, IST NoE 506766, <http://www.aimatshape.net/>
- [2] M.Attene, S.Biasotti, M.Moratara, G.Patanè, M.Spagnuolo and B.Falcidieno: "Computational methods for understanding 3D Shapes", *Computer and Graphics* 30, 323-333, 2006.
- [3] V. Blanz and T. Vetter: "Face recognition based on fitting a 3D morphable model", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, 1063-ss, 2003.
- [4] Bookstein, F. L.: "Principal warps: thin-plate splines and the decomposition of deformations", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 11, pp. 567-585, 1989.
- [5] V. Contreras: *Artanatomy. Anatomical Basis of Facial Expression Learning Tool*. Spain, 2005. Available at <http://www.artnatomia.net>.
- [6] Computational Geometry Algorithms Library, V.3.3.1
- [7] Edelsbrunner H, Harer J, Zomorodian A.: Hierarchical Morse-Smale complexes for piecewise linear 2-manifolds, *Discrete & Computational Geometry* 30(1):87-107, 2003.
- [8] P. Ekman and W. Friesen: *Manual for the Facial Action Coding System, Consulting Psychologists Press*, 1978. Palo Alto, CA, 1977
- [9] E.Elyan, H.Ugail: "Reconstruction of 3D Human Facial Images Using Partial Differential Equations", *J. of Computers*, Vol. 2, No. 8, 2007

- [10] F.Erol: *Modeling and Animating Personalized Faces*, M.Sc. thesis, 2002.
- [11] L. Farkas, *Anthropometry of the Head and Face*, Raven Press, New York, 1994.
- [12] B. Gokberk, A.A. Salah, L. Akarun: "Rank-based decision fusion for 3D shape-based face recognition", *Intl Conf on Audio and Video based Biometric Person Authentication (AVBPA)*, Springer-Verlag LNCS 3546, 2005, 1019-1028.
- [13] M.Goresky and R.MacPherson: "Stratified Morse Theory", *Erg. der Math. und ihre Grenzgebiete*, 3.Folge, Band 14, Springer-Verlag, 1988.
- [14] M. Hoch, G. Fleischmann, and B. Girod: "Modeling and animation of facial expressions based on B-splines", *Visual Computer*, vol. 11, pp. 87, 1994.
- [15] D. Huang and H. Yan: "Modeling and animation of human expressions using NURBS curves based on facial anatomy", *Signal Processing: Image Communication*, vol. 17, pp. 457, 2002
- [16] R. M. Koch, S. H. M. Roth, M. H. Gross, A. P. Zimmermann, and H. F. Sailer: "A Framework for Facial Surgery Simulation", Budmerice, Slovakia, 2002
- [17] Y. Lee, D. Terzopoulos, and K. Waters: "Realistic modeling for facial animation", Los Angeles, CA, USA, 1995.
- [18] N.Magnenat Thalman, P.Kalra, and M.Escher. "Face to Virtual Face", *Proc. of the IEEE*, Vol. 86, N. 5, 870-883, 1998.
- [19] S. Malassiotis and M. G. Strintzis: "Robust real-time 3D head pose estimation from range data", *Pattern Recognition*, vol. 38, 1153-ss, 2005
- [20] <http://meshlab.sourceforge.net/>
- [21] J.Milnor: "Morse Theory", *Ann. of Math.Studies* 51, Princeton Univ. Press, Princeton, NJ, 1963.
- [22] Mortara M, Patane G, Spagnuolo M, Falcidieno B, Rossignac J.: Plumber: a method for a multi-scale decomposition of 3D shapes into tubular primitives and bodies. In: *Proceedings of Solid Modeling and Applications (Poster Session)*, 2004. p. 339-44.
- [23] G.Rote and G.Vegter: "Computational Topology: An Introduction", in J.D.Boissonnat and M.Teillaud (eds): *Effective Computational Geometry for Curves and Surfaces*, Springer-Verlag, 2007.
- [24] Y. Song, L. Bai, and Y. Wang: "3D object modeling for entertainment applications", in *Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology* (Hollywood, California) ACM Press, 2006
- [25] H. Tanaka, M. Ikeda, and H. Chiaki: "Curvature-based face surface recognition using spherical correlation", *IEEE Conf. on FGR*, 372-377, 1998.
- [26] Y. Zhang, E. C. Prakash, and E. Sung, "Face alive", *Journal of Visual Languages and Computing*, vol. 15, 125, 2004.