3D surface segmentation
3D scans
polygonal mesh approximation

Leszek LUCHOWSKI [*]

SEGMENTING 3D MESH IMAGES OF THE HUMAN FACE BY LOCAL QUADRIC PARAMETRIZATION

An approach to 3D surface (mesh) segmentation is proposed, consisting in clustering vertices by approximating quadric surfaces. Seed groups of adjacent points give rise to initial quadric patches, which then grow by absorbing unassigned mesh vertices and compete with neighbouring patches by seizing their vertices. While the algorithm is intended for general use, it has been applied to 3D representations of human faces and the quadric patches tend to settle on salient facial features such as cheeks, chin or the ridge of the nose. Early results are presented and further work outlined.

1. INTRODUCTION

Segmentation of 3D shapes is not in itself a new subject. The technique proposed in this paper was developed as a general-purpose framework in which to detect salient 3D features of arbitrary shapes and represent them as analytical constructs that can be influenced by information from other sources, and yield quantifiable parameters. The use of standardized equations can benefit relaxation-like strategies of multisensor fusion [2]. The technique was tried on 3D scans of human faces [3, 4] and the results show some potential, although more work is needed and full segmentation of faces will undoubtedly require the addition of dedicated, model-driven criteria [1].

2. OUTLINE OF THE ALGORITHM

The input 3D surface is represented by a polygonal mesh, which remains unchanged throughout the procedure. Arbitrary, non-overlapping fragments of the mesh are approximated by quadric patches. A patch with its associated vertices will be referred to as a *patch*. The patches are then allowed to grow, move and decay by gaining vertices (from the pool of unassigned vertices or from neighbouring patches) or losing them. Vertices change their affiliation according to which adjacent patch has a quadric equation that best matches the coordinates. Certain parameters describing the quality of the patch are also taken into account so as to promote quadrics that best match the data. If a patch has lost so many vertices to its neighbour that it is has become

---

- Institute for Theoretical and Applied Informatics of the Polish Academy of Sciences,
Bałtycka 5, 44-100 Gliwice, Poland, www.iitis.gliwice.pl/zksw e-mail: leszek.luchowski@iitis.gliwice.pl

degenerated, this patch is destructed and its remaining vertices become available for the creation of new patches or expansion of existing ones.

3. INPUT DATA

The input data is organized as a polygonal mesh where each vertex has a defined set of neighbours (Fig. 1, leftmost image), so that neighbourhoods of given topological radii can be defined.

4. FORMATION OF INITIAL FIEFDOMS

For each vertex in the mesh, a list is created of neighbouring vertices (including itself) within a radius of R_b mesh edges. If none of the vertices on this list has yet been assigned to any patch (initially, the set of patches is empty), then they are grouped together, and their positioned are approximated (in the least-squares sense) by a second-degree polynomial $P_i(x, y, z)$. The group of vertices with its associated polynomial form a *patch* F_i.

The resulting, non-overlapping patches cover most of the mesh, with narrow gaps depending on R_b and on local mesh topology (Fig. 1, second image).

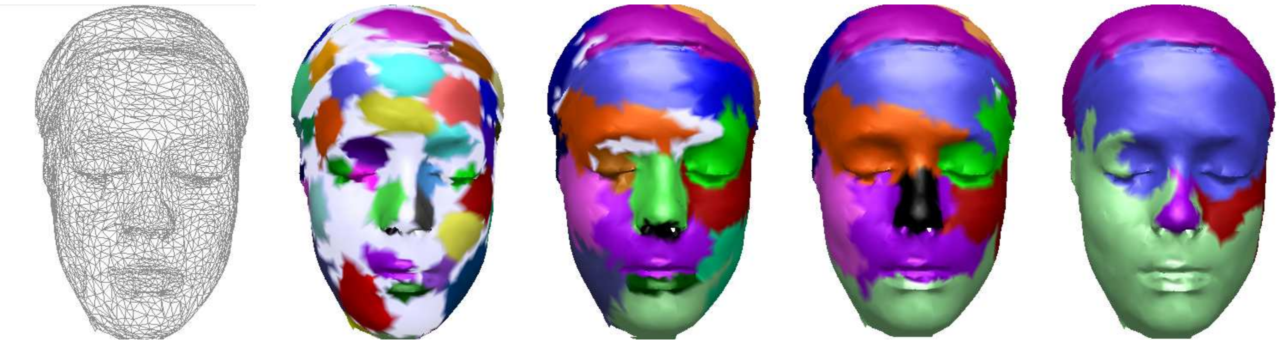

Fig.1 Left to right: input mesh, initial patches, iterations 129, 300, and 999.

5. EXPANSION, COMPETITION AND DRIFT

Each patch F_i in turn scans its immediate neighbourhood, i.e. mesh vertices which do not belong to F_i but are connected to at least one of the vertices of F_i by a mesh edge. Each such neighbouring vertex V_j is considered as a candidate to be seized into F_i. Net gain from such a seizure is defined as:

$$G_{F_iV_j} = \begin{cases} E_0 - E_{F_iV_j} & \text{, if } V_j \text{ was not assigned to any fiefdom} \\ E_{F_kV_j} - E_{F_iV_j} & \text{, if } V_j \text{ was assigned to an } F_k \text{ larger or equal to } F_i \text{ in n. of vertices} \\ E_{F_kV_j} - E_{F_iV_j} + T_s & \text{, if } V_j \text{ was assigned to an } F_k \text{ smaller than } F_i \end{cases} \quad (1)$$

where

E_0 is an arbitrary bonus for adopting an unassigned vertex,

$E_{F_iV_j}$ is the error with which vertex V_j fits the quadric equation of patch F_i.

T_S is a bias introduced to help overcome stalemates between two or more patches covering parts of the same continuous area.

If one or more candidates propose to bring net gains higher than zero, then the one which brings the highest is seized into F_i and removed from the patch F_k to which it previously belonged, if any.

A patch can also shed one of its edge vertices, if it matches its equation with an error above a rejection threshold T_R.

Patches that lose vertices in either way are re-examined to verify if they can still define a stable quadric surface. Those whose topology has degenerated into a nearly linear structure (does not contain a radius-R_b neighbourhood of any of its vertices), or that contain a total of fewer than 18 vertices) are deconstructed and their vertices are recycled, i.e. made available for others as unassigned. The number 18 was chosen as a compromise between, on one hand, allowing for small patches to represent relatively fine details, and on the other, preventing a patch from having so few vertices that they can be interpolated (rather than approximated) by a quadric, and thus having error zero and stagnating.

If loss of vertices cuts a gap across a patch splitting it into two or more unconnected parts, only the largest of them is kept while the vertices in others are made unassigned. Again, the truncated patch is checked for stability and possibly deconstructed.

All these steps: expansion and competition between existing patches, detection and deconstruction of degenerate ones, and creation of new ones out of unassigned vertices, are repeated in a cycle until change ceases or until a preset maximum number of iterations is reached.

6. INITIAL RESULTS

The algorithm was first tried on mesh representations of simple solids such as a cube and a cylinder and performed correctly, identifying the greatest parts of each body that can be represented by a quadric (simple planes, pairs of intersecting planes, and cylindrical surfaces).

Then an attempt was made to use the algorithm on 3D mesh images of human faces. The patches tend to cover identifiable structures such as the nose, chin, or eyes. Large areas such as cheeks and the hair cap worn for scanning tend to be split into two or more patches, which is an undesirable effect and will need to be solved. Over successive iterations, patches expanded and merged until most of the face is represented by two patches (and the cap by a third), and the nose was the only smaller detail that kept a patch of its own. While the overall potential of the technique to detect salient landmarks is confirmed, there is a need for more targeted choice of parameters and possibly a model-driven technique before this approach can reliably label parts of the human face.

7. PENDING WORK

In its present form, the algorithm has not been proven to converge, and is simply stopped after a fixed number of iterations. To guarantee convergence, criteria for vertex seizure need to be modified so that only vertices which decrease the mean quadric approximation error in a patch will be accepted. In this way, the process of patch growth and drift will only move the configuration of patches in one direction along the axis of total approximation error (decreasing), and thus will converge after reaching a minimum.

Other perspectives to be investigated include:

- A possibility should be introduced for a large patch to split spontaneously into two (even without being disrupted by loss of vertices) so as to better approximate the data.
- flexible dynamic tuning of algorithm parameters from initial rough fits to sub-vertex precision.
- using the quadric polynomial to derive geometric primitives. Depending on the type of quadric surface (plane, cone, ellipsoid, paraboloid etc) it defines different sets of planes, lines, points and directions. They will be computed for the purposes of measurement, labeling, and matching.

BIBLIOGRAPHY

[1] KOWALSKI P., TOMAKA A., LIŚNIEWSKA-MACHORORSKA B., Landmarks identification using active appearance models. Archiwum Informatyki Teoretycznej I Stosowanej, t. 17, , z 4/2005, pp 251-264

[2] CHAARI A., LUCHOWSKI L., ROUSSEAU J., Multiple-view stereo matching of points using epipolar constraints and relaxation, Archiwum Informatyki Teoretycznej i Stosowanej, Vol. 17, No 4, 2005, pp. 241-250

[3] TOMAKA A., LIŚNIEWSKA-MACHOROWSKA B., The application of the 3D surface scanning In the facial features analysis, Journal of Medical Informatics & Technologies, Vol. 9, October 2005, pp. 233-240

[4] TOMAKA A., LUCHOWSKI L., SKABEK K., TARNAWSKI M., 3D Head surface scanning techniques for orthodontics, Journal of Medical Informatics & Technologies, Vol. 9, October 2005, pp. 123-130

ACKNOWLEDGMENT

This work was partly financed under the Polish Government research grant No 3 T11C 028 27.