On 3D Face Feature Segmentation Using Implicit Surface Active Contours

M. Krueger, P. Delmas, G. Gimel’farb
Dept. of Computer Science, Tamaki Campus, The University of Auckland, New Zealand.
Email: mkru007@aucklanduni.ac.nz

Abstract
Segmentation and extraction of face features from 2D images has been researched extensively in the past. Recent research shows that algorithms incorporating 3D information often perform better than purely 2D approaches in many areas of computer vision including face biometrics and recognition. Cheaper and readily available 3D data acquisition devices combined with improved 3D reconstruction algorithms delivering highly accurate data are leading a shift from 2D to 3D based data. In this paper we review our recently introduced algorithm for segmentation of 3D surfaces using active contours. Then we apply this framework to propose a novel method for feature segmentation from 3D faces. The performance of our model is illustrated on a set of 3D faces.

Keywords: Active contour, face features, 3D, lips, partial differential equation, surface segmentation

1 Introduction
Over the last decade, the field of 3D data processing has become immensely popular in the Computer Vision community. This has several reasons: acquisition tools such as 3D scanners have become more affordable. Algorithms for 3D reconstruction, e.g. from stereo images, provide an easily accessible and inexpensive means to generate three dimensional data. Furthermore, 2D problems are ill-posed as they often offer a restricted-partial representation of 3D reality and 2D approaches have shown to be error-prone in presence of challenging illumination conditions. Accordingly, recent research [1, 2, 3] shows that multi-modal approaches combining 2D (texture) and 3D (shape) information perform better than either 2D or 3D techniques alone in face biometrics.

Methods that are well-known and understood for planar images, e.g. isotropic and anisotropic diffusion [4, 5] and scale space analysis [5] are now often generalised to images painted on curved surfaces. Active contour models, a popular tool for object segmentation in 2D images, are currently enjoying a regain of attention for curved surfaces segmentation. Spira and Kimmel [6] applied the concept of geodesic active contours (GAC), to images on parametric surfaces. They did not incorporate a balloon term in their model thus limiting the scope of the method to ideal data (involving no or very little noise). Krueger et al. introduced in [7] a GAC model for the segmentation of implicit surfaces. An optimised algorithm converging in linear time was introduced in [8].

In [8] accurate lip contours extraction results were presented applying the proposed GAC model to 3D face data. However, essential details of the segmentation process such as the design of suitable stopping terms were omitted. In this paper we fill this gap thus describing in detail the complete process of 3D face feature (such as the lips) segmentation using GACs and the geometry of the face. To the best of our knowledge this has not been done before. This integrated approach is the main contribution of the paper.

The paper is organized as follows. Section 2 gives a brief overview of related work. Section 3 succinctly reviews the framework for GAC on implicit surfaces, the backbone of our algorithm for 3D face feature segmentation, as introduced in [8]. The details of the segmentation process will be presented in section 4. Section 5 presents results of 3D face features segmentation while section 6 concludes the paper.

2 Related Work
As there is a huge body of published work on the topic of active contours which is related to the theoretical background of our work, we only cite here the most important papers: Active contours (also known as “snakes”) were introduced in the
pioneering work by Kass et al. [9]. Derived from the gradient flow of an energy functional, a snake is a curve that deforms itself to minimise its total internal and external energy. The internal energy imposes regularity (first and second order) on the curve while the external energy attracts the curve to desired image features (e.g., edges, lines, regions, etc.). Later, Caselles et al. [10] introduced a general GAC following the curve to change its topology during the evolution. They further proposed a general GAC model including a balloon force which shrinks or expands the curve to desired image features (e.g., edges, lines, regions, etc.). Later, Caselles et al. [10] introduced the GACs, based on an energy functional \( E \) that integrates a positive edge indicator function \( f \) along the contour \( C \):

\[
E(C) = \int_C f(C) \, ds .
\]

The function \( f \) originally had small values for large image gradients and higher values elsewhere. In contrast to the classic snakes, the GACs are independent of a particular parameterisation. Caselles et al. [10] applied the level sets method [11] allowing the curve to change its topology during the evolution. They further proposed a general GAC model including a balloon force which shrinks or inflates the curve by a constant speed:

\[
\partial_t C = (f \cdot (\kappa - c) - \langle \nabla f, \nu \rangle) \nu .
\]

where \( \partial_t C := \frac{\partial C}{\partial t} \), \( \kappa \), and \( \nu \) are the time derivative, the curvature, and the unit normal field, respectively. The GAC converges towards a steady state if a sufficiently stable minimum of the respective total energy is reached.

Spira and Kimmel [6] applied the original GAC model to segment images painted on parametric surfaces. This was achieved by back-projecting the GAC equation from the surface onto the parameterisation plane and solving the resulting partial differential equation (PDE) with the level sets framework [11]. No balloon term was incorporated in their model. Subsequently, Krueger et al. showed in [8] that integrating a balloon term in such parametric framework leads to an ill-posed equation thus generating unstable results. To counter this, Krueger et al. [8] proposed to evolve GAC on surfaces using the implicit representation. In this framework a balloon force could be included, making it suitable for noisy real data. We will review this approach in more detail in Sect. 3. Please see [8] for further references on the theoretical aspects.

3D face feature segmentation is a relatively young and unexplored field of research: the bibliography on it is comparably small. Existing approaches usually focus on the detection of landmark points, see [1] for a list of references. In [3] the ear shape contour was segmented using depth and color information. To the best of our knowledge we are the first to publish work on face feature segmentation using only 3D shape data.

3 GAC on Implicit Surfaces

3.1 The Equation

After embedding a curve \( C \) as an isocontour in a 2D scalar function \( \phi \), the evolution of the curve is controlled by solving a suitable PDE for \( \phi \) (see e.g. [11, 12]). The GAC in (2) corresponds to the following level set PDE:

\[
\partial_t \phi = cf|\nabla \phi| + \langle \nabla f, \nabla \phi \rangle + f\kappa|\nabla \phi| ,
\]

where the terms \( I, II, \) and \( III \) denote a motion in normal direction, an advection flow, and a weighted mean curvature flow, respectively.

Applying Cheng et al.’s [13] framework for geometric curve evolution on implicit surfaces, Krueger et al. [7, 8] generalized the GAC model (3) to implicit surfaces. With a surface \( M = \{ x \in \mathbb{R}^3 : \psi(x) = 0 \} \) implicitly embedded in a cuboid \( \Omega \subset \mathbb{R}^3 \), a grayscale image \( I \) is a mapping \( I : M \rightarrow \mathbb{Z}^+ \). As in the classical GAC model, the image \( I \) should be smoothed before calculating an edge indicator function, e.g. by a low pass filter. If \( I_\sigma \) is the smoothed image, the edge indicator function can be defined as \( f = \frac{1}{1 + |\nabla I_\sigma|^2} \), \( \sigma \). Note, that \( f \) can be defined in terms of geometric quantities such as the surface curvature as well. As in [13] the projection operator \( P_P \) is defined by \( P_\varphi(v) = v - \frac{1}{|w|} \langle v, w \rangle w \), where \( v \) and \( w \) are vectors in \( \mathbb{R}^3 \).

Then the evolution for GACs on implicit surfaces is

\[
\partial_t \phi = cf|P_{\varphi} \nabla \phi| + \langle P_{\varphi} \nabla f, \nabla \phi \rangle
\]

\[
+ \mu f|\kappa| P_{\varphi} \nabla \phi .
\]

where the parameters \( c \) and \( \mu \) determine the influence of the balloon force and the regularity terms, respectively. The image gradient term has a normalised weight. Then the evolving contour on \( M \) is given by the zero level set \( C(t) = \{ x \in \Omega : \phi(x, t) = \psi(x) = 0 \} \).

Figure 1: Evolving narrowband on a face surface.

3.2 Numerics

The PDE in question (4) is partly second order parabolic (term \( II \)) and partly of Hamilton-Jacobi type (term \( I \)). Appropriate numerical schemes (central differences resp. essentially non oscillating
(ENO/WENO) schemes) were applied to accurately compute the spatial derivatives. The hyperbolic term \( I \) was approximated by a Lax-Friedrichs (LF) scheme with global flux and the time derivative by forward differencing. In [8] a narrowband algorithm was proposed that reduced the computational complexity of the algorithm to linear (see figure 1). We refer to this paper for more details on the numerical implementation.

4 3D Face Feature Segmentation

Figure 2 depicts the two possible sources of information when dealing with 3D face data, the texture and intrinsic geometric information such as the curvature. Note that face features like the lips and the eyes can be characterised by both texture and curvature information.

Figure 2: Human face with texture (left) and mean curvature colorisation (right).

Traditionally, 3D scanners emphasize depth resolution and accuracy over texture resolution (when they provide one). We therefore decided to use only curvature data for the algorithm in this paper. Note that we smoothed all the surfaces by iterating the mean curvature flow a few times, before further analysing them.

4.1 The Lips

Looking closer at the mouth region (figure 3), one can observe ridges delineating the lips from surrounding areas. However, these ridges tend to weaken close to the mouth corners. Thus, the segmentation of the lips cannot be achieved by a simple computation of ridge-lines. We therefore intend to segment the lips by applying our GAC model with a suitable stopping term \( f \), and without the need of the image gradient.

A naive approach is to define the stopping function as

\[
f = \frac{1}{1 + |H|^p}, \quad p \in \{1, 2\},
\]

where \( \hat{H} \) is a rescaled version of \( H \) with values between 0 and \( \hat{H}_{max} \). We separately rescaled both the negative and the positive sign range so that

\[
\hat{H}(x) \in \begin{cases} [0, \hat{H}_{max}) & \text{if } H(x) < 0 \\ [\hat{H}_{min}, \hat{H}_{max}] & \text{if } H(x) \geq 0. \end{cases}
\]

The graph of the function in (6) is shown in figure 5. Note, how it decreases in the direction of positive mean curvature values. Thus, the advection term in the GAC equation (4) pushes the contour into regions of high positive curvature. Optimally, the contour converges to the upper and lower ridges while assuming an equilibrium of the advection and the other two terms near the mouth corners.

Figure 3: Two examples for the mouth region with mean curvature colorisation. Regions of positive curvature are colored blue/pink, negative curvature regions have green/yellow/red color (on a grayscale printout these regions appear dark and bright respectively).

Figure 4: GAC with stopping term (5): initial contour (left) and contour after passing the mouth corners (right).

Figure 5: GAC with stopping term (6): initial contour (left) and contour after passing the mouth corners (right).
In our tests we noticed that stopping function $f_1$ has an undesired effect: since $f_1$ has a steep negative gradient for negative curvature values (remember equation (7)) the contour is repelled from the mouth corners too strongly (see results in section 5). To counter this, we designed a further stopping function $f_2$ with a more uniform negative slope (see figure 5):

$$f_2 = \frac{0.8}{\pi} \cdot \arccos \left( \frac{\tilde{H} - \tilde{H}_{\text{max}}/2}{\tilde{H}_{\text{max}}/2} \right) + 0.2. \quad (8)$$

4.2 The Eyes

In general, segmenting the eyes from 3D face scans is more problematic than segmenting the lips. The first and most obvious reason is that in most 3D face reconstructions the eyes are closed. In this case there is no obvious contour to be segmented. In the following we assume that the eyes are open.

Figure 6 shows two examples of eyes regions with superposed mean curvature. The eye pupil and iris boundaries are highlighted by circular ridges of high positive mean curvature. Again, the eye corners lack consistent negative curvature. While in some faces the curvature gap is rather small (like in 6(a)) it is prominent in others (see 6(b)). Unless the contour is initialised close to the eye contour, a balloon force is necessary to reach satisfactory convergence. In many cases this causes the contour to miss the outer eye corner features. We will provide further examples in section 5. More sophisticated models, e.g. including a higher order term penalising curvature or a shape prior, might be more promising for this application.

5 Experimental Results

We have tested our model on a set of 3D face surface using an Intel Pentium M 1.6 GHz Notebook endowed with MATLAB®. Firstly, we tested the freely available content of the Face Database of the MPI for Biological Cybernetics in Tübingen, Germany [14], which contains four triangulated 3D face datasets including texture. The triangulated surfaces have a resolution of between 70,000 to 90,000 vertices. Patches of the face surfaces were converted to implicit representation using a test version of the FastRBF toolbox [15]. They were embedded implicitly in cuboids with side lengths between 70 and 150 pixels.

Figure 7: Mean curvature of the mouth region before (left) and after (right) cutting the outliers (< 1% and > 99% percentiles). Bright areas have positive mean curvature, dark areas have negative mean curvature.

Overall, removing outliers from the curvature function $H$ prior to the GAC evolution improves the segmentation results. Figure 7 illustrates how the contrast of the curvature distribution is increased by removing lower and higher ends of the curvature $H$ spectrum (e.g. at 1% and 99% percentile thresholds). Next we rescaled the curvature function $\tilde{H}$ as described in section 4. However, removing parts of the curvature range has to be considered carefully. Indeed cutting too much often causes the contour to get trapped in undesired local minima in small areas of comparably high curvature. Extensive tests demonstrated that good results could be consistently achieved for $\tilde{H}_{\text{max}} = 2.5$ and the threshold values stated above.

5.1 The Lips

We first used the stopping function $f_1$. As expected the results were not optimal. Figure 8 (a) illustrates a segmentation result using $f_1$ with the parameters $\mu = 0.2$ and $c = 0.04$. Figure 8 (b) displays the GACs final form on the same face using stopping function $f_2$ (with parameters $\mu = 0.15$ and $c = 0.03$). The mouth corner region is better segmented. Considering this, all available datasets of the MPI database were processed with this stopping term and these parameters. As a
uniform time-step we chose $\tau = 1.0$. Figure 9 illustrates results of our algorithm on the datasets of the MPI Tuebingen face database. Except for one face (figures 9(g) and (h)) where lips are barely visible in both curvature and texture images, the segmentation results are close to subjective contours. Note as well that in the other three examples the computed ‘anatomical’ lip contour – based on the lip geometry – is very close to the actual – texture defined – lip contour. Despite the minor deviations, a geometry based lip contour can be very interesting for biometrics or face recognition. An expected refinement for future work would be to combine or sequentially process GACs on texture and curvature information. In figure 10 we show two further results of our algorithm on data acquired with a REXSCAN halogen light scanner (with sub-mm depth precision and 0.5 mm planar resolution). The data is visibly noisier while ridges are less prominent. Not surprisingly, we had to adapt the model parameters ($\mu = 0.32$ and $c = 0.0125$). The regularity term weight was increased and the balloon term weight lowered to obtain good results (see figure 10). Overall we applied our method to six 3D face datasets, delivering good results for five of them. The robustness of the algorithm with respect to parameter values is yet to determine. To answer this question, we are currently building a sufficiently large database of faces, encompassing various facial deformations, acquired under constant illumination conditions.

Processing times of the method are between 4-5 minutes and 10-12 minutes for one face, depending on the data quality. Keeping in mind, that our method is fully three dimensional, one cannot expect it to deliver real-time results. However, if the contour is initialised close to the desired result, processing times can be reduced considerably.

5.2 The Eyes

As indicated in section 4 our results on 3D eye segmentation so far are not as promising as the lip
results. For completeness sake they are mentioned below. We applied our model with stopping term $f_1$ to the four faces of the MPI Tübingen database [14]. For two of them we were able to obtain a good result. However, we had to adapt the parameters manually for each one of them. For the other two faces, our model failed due to the problems mentioned in section 4. Respective examples are illustrated in figure 11.

6 Conclusion

In this paper we have proposed a new method for 3D face feature segmentation. It is based on a new framework for the segmentation of surfaces using active contours [8] which takes due account of the 3D surface geometry. We have proved experimentally that our method can be applied to 3D face segmentation using only the three dimensional geometry of the face without the need for texture information thus freeing us from illumination conditions. Our preliminary results for lip contours extraction are promising. Yet, we are aware that there is still much work to do. The method has to be extensively tested on a large database of 3D faces acquired under constant conditions. It has to be determined how the model parameters can be estimated from the database parameters using for example the data noise distribution and general data quality on the one hand, and the face geometry of the particular face on the other hand. Moreover, it will be interesting to research how geometry and texture can be combined to achieve a multimodal method.

Having said that, there are some limitations of our method: the computational load of the fully 3D based approach makes it unsuitable for online resp. real-time applications. Furthermore, it can only work if the surface data have a sufficiently good quality. For blurry, very noisy or data with bad depth resolution it will probably not work.

However, we believe that we have proposed an innovative approach that can open up new avenues in applications such as biometrics or face recognition where 3D data can be readily obtained.

References


