Facial Features Segmentation by Model-Based Snakes

Petia Radeva and Enric Martí
Departament d'Informatica, Facultat de Ciencies,
Universitat Autonoma de Barcelona, 08193 Bellaterra, Spain
e-mail petia@upisun1.uab.es

Abstract

Deformable models have recently been accepted as a standard technique to segment different features in facial images. Despite they give a good approximation of the salient features in a facial image, the resulting shapes of the segmentation process seem somewhat artificial with respect to the natural feature shapes. In this paper we show that active contour models (in particular, rubber snakes) give more close and natural representation of the detected feature shape. Besides, using snakes for facial segmentation frees us from the problem of determination of the numerous weigths of deformable models. Another advantage of rubber snakes is their reduced computational cost.

Our experiments using rubber snakes for segmentation of facial snapshots have shown a significant improvement compared to deformable models.

1 Introduction

The successful automatic extraction of facial features is strongly dependent on the existance of feature model that should be descriptive enough to embody common shape, but yet flexible to allow some degree of variation [4]. With this fact we can explain the popularity of deformable models [14] for the segmentation of facial features. Deformable models are collections of parameterized curves (normally, parabolae and circles) which, taken together, describe the expected shape of the feature to be detected in the image [2]. The template interacts dynamically with the original image and with the peaks, the valleys and the edges derived from the image. An energy function is determined that estimates and reduces the deviation of the template from the image characteristics. The minimum of the energy function corresponds to the best fit with the image.

Due to the special shape of the deformable template (consisting of parabolic curves and circles) the segmented facial features seem somewhat artificial. We should admit that almost never can we see such a regular and symmetric eye, that the eyelashes have exactly the shape of parabolic curves and the iris is seen as a perfect circle. Another problem of the deformable models technique is the previous determination of the parameters of the parabolae and the circle of the model and of the numerous weights for the constituents of the energy determined empirically [14, 2, 10, 13]. Xie et al. in [13] propose to normalize the different external energies in order to avoid the predetermination of the weights. This way the parameters to be assigned are reduced and the rest of them are derived from the grey level mean and standard deviation of the eye area. Another problem pointed out by many researchers working on the facial features segmentation by deformable models [10, 13] is the difficult segmentation of the lower eyelash and the lower lip. We show here, that the use of snakes is convenient for facial segmentation because they adjust themselves better to the facial image and help avoid the mentioned problems of deformable models.

The main difference between deformable models and rubber snakes is that the latter are not restricted to some predetermined family of shapes, and this way the segmented features have more natural representation. The classical snakes [6] do not take into account the specific *a priori* knowledge available, have more parameters to update and, hence, they are computationally slower [14]. For this reason we chose to segment the facial features by rubber snakes [7] whose deformation is model-guided similarly to the elastically deformable models proposed by Terzopoulos et al. [11]. The rubber snakes begin their deformation from a particular model of an individual feature and change their shape in accordance to the data in the image maintaining minimal difference with the eye model. This way they achieve to adjust themselves better to the edges and valleys of the image and do not need a previous determination of so many energies weights and parameters of parameterized curves due to the explicit use of a feature model. The only parameters necessary to determine previously are the parameters of elasticity and rigidity that are global for all the pixels of the snake. Similarly to the deformable models, the rubber snake deforms on edge and valley potentials but in a reduced computational process.

The deformable models have been prefered to the active contour models because of their global deformation [14], the shape changes are propagating along the parabolic curves and circles. In order to preserve this global nature of the deformation process we propose to use an additional component to the internal energy acting as springs between certain pairs of snake pixels and that take charge of the approximately symmetric changes in the snake shape. Finally, in this work the standard scheme for the feature localization and segmentation is modified to optimize the features localization, to solve some problems of the segmentation as for example this of the lower eyelash and of the lower lip, and to reduce the calculus of the deformation process.

The article is organized as follows: in section 2, we expose the fundamentals of rubber snakes. In section 3, we discuss the localization of the features of interest: the eyes, the eyebrows and the mouth. Section 4 is dedicated to features segmentation. Finally, conclusions are reported.

2 Rubber Snake Model: Fundamentals

A rubber snake is a continuous curve that, from an initial state, tries to position itself dynamically on image characteristics (f.e. edge or valley points).

External forces are designed to minimize the snake external energy so that it is pushed towards certain characteristics of the original image. These forces are associated to a potential which is defined in terms of a distance map of the image characteristics [3]. For the facial segmentation we use two potential fields: edge potential field and valley potential field. The construction of the edge potential field is done in terms of a signed distance potential that explicitly incorporates information about the gradient direction of the edges derived from the image [8].

An internal force of the rubber snake minimizes its internal energy keeping it near the feature model. The internal energy is as follows:

$$E_{int} = \alpha(u_s(s) - u_s^0(s))^2 + \beta(u_{ss}(s) - u_{ss}^0(s))^2$$

where $u(s) = (x(s), y(s))$ is the snake curve, s is the curve arc-length, $|u_s(s)|$ is the membrane energy of the snake, $|u_{ss}(s)|$ is its thin-plate energy, $|u_s^0(s)|$ and $|u_{ss}^0(s)|$ are the membrane energy and the thin-plate energy of the feature model [7]. The parameter of elasticity α and the parameter of rigidity β take charge of the control of the snake deviation from the model in each step of its movement. Thus, the rubber snake explicitly incorporates structural information about the desired final snake shape. The internal force of the rubber snake always attempts to compensate the changes caused by the external forces and to preserve the object model controlled by the parameters α and β.

The minimum of the snake energy corresponds to the best fit with the image. The total energy of the snake is represented by means of the energies addition functional:

$$E_{snake} = \int_0^1 E_{int}(u(s)) + E_{ext}(u(s))ds =$$

$$\int_0^1 \alpha(u_s(s) - u_s^0(s))^2 + \beta(u_{ss}(s) - u_{ss}^0(s))^2 +$$

$$P_{edg}(u(s)) + P_{vall}(u(s))\ ds$$

In order to extract different types of characteristics from the original image, we use the facet model. It provides a more exact edge points detector, the zero-crossing of the second directional derivative [5]. This is an edge detection method which has a good localization as well as a good estimation of the edge gradient. Another advantage is that it is very sensitive to weak image characteristics, this way it gives us more information about the facial features normally characterized by smooth contours. The facet model also allows to extract in the same computational process the edge points with their magnitude and direction, the valley points and the ridge points.

3 Features Localization

In order to segment a facial feature by rubber snakes, we need to locate an initial snake (a model) near the feature. Also, as the snake deformation is guided by a model, some normalization of the initial snake in scale and orientation should be done using the information derived in the stage of features localization.

A recommended technique for detection of facial features is by template matching [9]. The main disadvantage of this approach is the expansive computational cost. In order to reduce this cost we combine the template matching with the localization by integral projections [1]. Instead of matching the templates on all the image, it is examined for locations that fulfill some predetermined constraint (for example, bilateral symmetry) and thus the matching is done in a delimited region.

Given an image $I(x,y)$ we consider the vertical and the horizontal integral projections in the $[x_1, y_1] \times [x_2, y_2]$ rectangle defined as:

$$V(x) = \Sigma_{y=y_1}^{y_2} I(x,y), \quad H(y) = \Sigma_{x=x_1}^{x_2} I(x,y)$$

Brunelli and Poggio [1] use the integral projections on the edge map partitioned in terms of edge directions. We found out that more exact results can be obtained applying the projection analysis on the intensity image because of the smooth contours of most of the facial features. One example of vertical and horizontal projections of a face can be seen in Fig. 1 and Fig. 2.

In contrast with most features localizing techniques that begin from the eyes extraction, we firstly detect the nose crest. The nose localization has the advantage that we are looking for a homogeneous bright area, where we do not find external elements that cover it (as compared to the mouth or eyes, where a beard or sunglasses may occur, making difficult their localization). Therefore, the vertical projection is characterized by a well-differentiated peak corresponding to the nose. The nose position is important to locate the eyes symmetrically with respect to the nose. The eyes are localized by two symmetrical pits in the vertical projection and by one pit in the horizontal projection. Another pit in the horizontal projection near the eye location corresponds to the eyebrows. It can be more or less strong than the eye pit and hence we use the fact that the eyes are bellow the eyebrows to determine the features location. Other pits can appear in the vertical and horizontal projections due to shades and wrinkles around the eyes. Therefore, different hypotheses are constructed about the features location and a template matching is performed in a neighbourhood of the supposed location of the features. It is necessary, since the projections can change in some limits the horizontal and vertical position of the features in case of different facial orientations.

For the localization of the eyes we used a template matching because of the complex structure of the eyes. The template consists of one horizontal segment and three vertical segments looking for dark areas corresponding to the iris and the upper eyelash and for brigth areas corresponding to the eye whites. The template is normalized in accordance to the expected feature scale, obtained by the projection analysis. The localization of the eyebrows and the mouth is simpler with respect to the eye localization: the eyebrows are localized by looking for two valleys

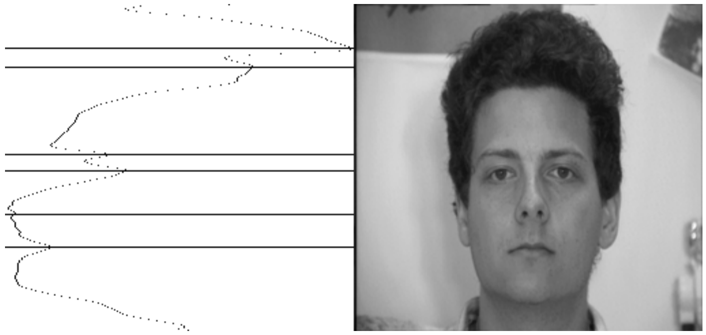

Figure 1: Vertical projections of a face

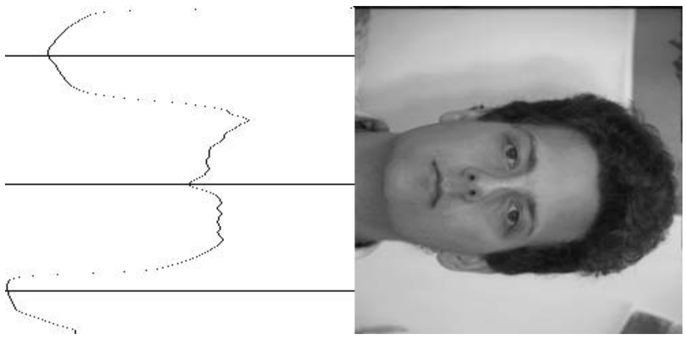

Figure 2: Horizontal projections of a face

bilaterally symmetrical with respect to the nose in the neighbourhood of a pit of the horizontal projection. The mouth is detected looking for a valley with its height near a pit in the horizontal projection and symmetrical with respect to the width position of the nose.

4 Features Segmentation

Using the information regarding the expected scale and orientation obtained from the distance between both irises, the iris models are scaled and used as initial snakes to adjust to the facial image.

4.1 Eye Segmentation

The segmentation of the eyes by rubber snakes is performed in a sequential manner analogously to the template deformation on different epochs used by Yuille et al. [14]. The reason is that it is easier to detect the irises because they appear as two circular spots symmetrically with respect to the nose. Once they are detected, their location is used to orient the models corresponding to the eyelashes. In order to keep the circular shape of the iris, the iris snakes try to maintain the thin-plate energy of the model. Nevertheless, in cases of occluded part of the iris the rubber snake obtains the exact contour, producing this way the realistic iris shape.

After the detection of the irises, the eye contours are considered. In contrast to the well-known works on eyes segmentation, our observations on different eyes images are that it is more convenient to look for the eye shape from the valley points than from the edge points, because of the nature of the eyelashes.

In order to increase the global control on the snake shape deformation we add a new term to the internal force of the rubber snake definition. This force takes charge of the symmetric changes of the rubber snakes, assuring this way certain symmetry between both eyelashes. The new internal energy for an eye is:

$$E_{int} = \alpha(s)(u_s(s) - u_s^0(s))^2 + \beta(u_{ss}(s) - u_{ss}^0(s))^2 + c(u(s) - u^{sym}(s))$$

where $u^{sym}(s))$ is the inverted and symmetric vector of the snake curve $u(s)$ with respect to both eye ends. Given that the number of pixels for each eyelash is equal and the first pixel and the middle pixel of the snake should correspond to both ends of the eye, it is easy to see that the snake for each eye deforms symmetrically with respect to both its ends. The coefficient c controls the strength of the symmetry constraint.

Another feature that helps eye segmentation is provided by the iris contours. Using the fact that the iris contour should lay inside the eye, we put the detected irises into the valley potential. If the eye model has some part near these valleys, and there are not any stronger valleys due to eyelashes that make the snake go out from the iris valleys, the snake position remains, so that the iris is always inside the final eye contour detected by the snake.

Once the eye contours have been detected, the results are tested considering for each eye contour its symmetrical one with respect to the nose crest, thus

Figure 3: An example of segmented eyes and eyebrows through rubber snakes

forming two pairs of eyes and repeating the deformation process. The best result of the deformation process is chosen as final result.

When the lower eyelash has large external energy (i.e. the snake fails to detect a valley corresponding to the lower eyelash), the snake is deformed on the edge potential. Both the ends of the snake corresponding to the lower eyelash are fixed. The snake curve is divided in three parts and they are deformed together but looking for edges with different gradient directions. The outside parts of the snake adjust to edge points with gradient directed towards the inside of the eye (due to the whites of the eye) and the central part of the snake is adjusting to edge points with gradient direction towards the outer part of the eye (due to the touch of the eyelash with the iris). Parallely, an eyelash snake is deformed adjusting itself to edge points with gradient direction towards the inside of the eye. This is the case when the eyelash does not touch the iris of the eye. The energy of the final snakes is estimated and the snake with the lower energy is considered as the result of the segmentation. An example of different segmented eyes can be seen in Fig. 3.

4.2 Eyebrow Segmentation

The contours of the eyebrows are difficult to detect in many facial images because of the texture of the eyebrows. It turned out that the valley points information derived from the original image gives us more consistent data about the position and shape of the eyebrows than that of the edge points. Following this idea, we firstly use an open snake to detect a valley line corresponding to the eyebrow skeleton. Afterwards, two initial models are put around the detected valley and the best deformation is considered as final result of the eyebrow segmentation. Using the results of two models is preferable because of the existance of two general shapes of eyebrows one can see. The first model corresponds to horizontal eyebrows, while the second one corresponds to convex eyebrow shapes. In order to help to the segmentation of the eyebrow contour, an internal energy similar to that of the eye is designed to add a new term. It provides symmetry to the contour with respect to the skeleton of the eyebrow. The left and right eyebrows are segmented independently and their results are tested for symmetry. An example of different segmented eyebrows is given in Fig. 3.

4.3 Mouth Segmentation

The segmentation of the mouth is similar to that of the eyebrows. A difference is that here, instead of the skeleton of the eyebrow, we have used the line between the lips detected on the valley potential by an open snake. Afterwards, the mouth model is scaled and oriented in accordance to the line between the lips and put onto the edge potential in order to detect the contours of the lips. The snake is deforming symmetrically with respect to the line between the lips. However, the constraint on the symmetry is lower than in the case of the eyebrows and the eyes, because the lips may be of different width and also in order to capture the existance of convexities in the upper lip that have no corresponding symmetrical convexities in the lower lip.

It has been already reported by other authors [2] that the lower lip often creates problems to be detected due to scarce edge points in the edge map. To solve the problem we use the gradient direction of the edge points. Depending on the illumination conditions, the lower lip may be seen as a dark area, as a bright area or, more frequently, as a smooth composition of dark-bright-dark areas. Since the edge potential contains explicit information about the gradient direction of the edge points, three snakes are constructed corresponding to the different possibilities for the edge points and the snake with best energy estimation after their deformation is considered as the result of the lip segmentation. In Fig. 4 an example of different segmented mouths is shown.

5 Results

We have applied the localization and segmentation technique by rubber snakes presented here on 26 facial images. The conditions on the facial images were to allow 15% deviation in scale, orientation and translation from the center of the image. The illumination when making the photos has been central. The background is supposed to be approximately homogeneous. As a result, in 23 images, the eyes were segmented correctly. The other three failed to detect correctly the irises because of their bright colour. In the mouth segmentation, some problems appeared

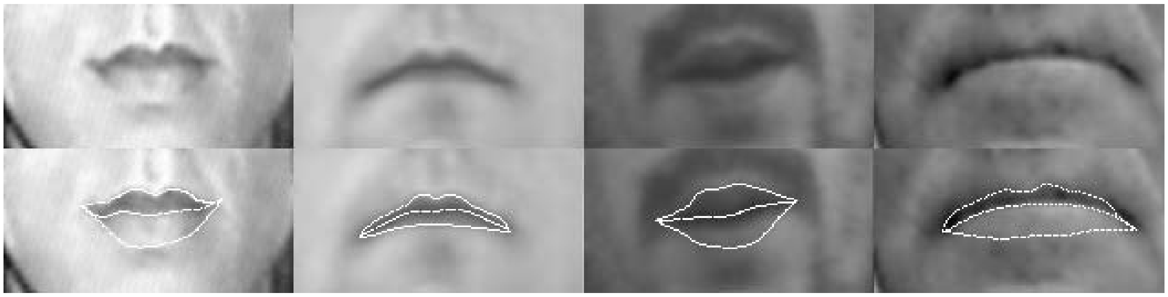

Figure 4: An example of segmented mouths through rubber snakes

when the line between the lips is shorter than the valley line in the image due to the presence of wrinkles next to the mouth. In three of the images, the snake could not detect exactly the end of the mouth. The problem we saw with the eyebrows detection is their eventual merge with shadows above the eyes as well as the occlusion of the eyebrow by the hair. We are working on incorporation of high-knowledge about the faces to guide and to test the snake deformation results. An extremely interesting issue to be considered for facial segmentation is the implementation of snakes for the different features that should deform simultaneously and dependently from each other.

6 Conclusions

The facial segmentation by rubber snakes can be considered as a more global technique than the one by deformable templates and less global than by original snakes. It has the advantage that the snakes capture different shapes, unrestricted in a certain family, according to the image characteristics and the imposed constraints on its "distance" to a given model. Besides its possibility to obtain approximately symmetrical shapes, its principal difference from the deformable templates is that it does not need to modelize the object it is looking for, since the model is used implicitly. We also show that the role of valley points is significant for the correct localization and segmentation of most of the facial features as f.e. the eye. This fact allowed us to simplify the scheme of facial segmentation by rubber snakes using only two external energies. The result is the reduced complexity of the segmentation process and no need of determination of a number of weights used in other facial segmentation techniques.

Since this segmentation technique is not depending on the features we have considered until now, it is easy transferred in segmentation modules for other objects. Our future work is related to the application of the segmentation technique by rubber snakes to detect other facial features and objects as wrinkles, spectacles, a moustache, etc. as part of our future work on facial expressions recognition.

References

[1] R. Brunelli and T. Poggio, "Face Recognition: Features versus Templates", *IEEE Transactions on Pattern Recognition and Image Analysis*, vol. 15, No. 10, October, 1993, pp. 1042-1052.

[2] Ch.-W. Chen and Ch.-L. Huang, "Human Face Recognition from a Single Front View", *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 6, No. 4, pp. 571-594, 1992.

[3] I. Cohen, L. D. Cohen and N. Ayache, "Using Deformable Surfaces to Segment 3-D Images and Infer Differential Structures", *CVGIP: IMAGE UNDERSTANDING*, Vol.56, pp.242-263, 1992.

[4] G. Chow and X. Li, "Towards a System for Automatic Facial Feature Detection", *Pattern Recognition*, Vol. 26, No. 12, 1993, pp. 1739-1755.

[5] R. M. Haralick and L. G. Shapiro, "The Facet Model," *Computer and Robot Vision, Vol.1,* Addison-Wesley, New York, 1992.

[6] M. Kass and A. Witkin and D. Terzopolous, "Snakes: Active Contour Models," *Proc. of Int. Conf. on CV*, pp.259-268, London, 1987.

[7] P. I. Radeva, "Segmentation of Radiographs by Rubber Snakes. Application to the Assessment of Skeletal Maturity", *Master Thesis*, UPPIA, UAB, September, 1993.

[8] P. I. Radeva and J. Serrat, "Rubber Snake: Implementation on Signed Distance Potential", In *Vision Conference SWISS'93*, pp. 187-194, September, 1993.

[9] A. Samal and P. A. Iyengar, "Automatic Recognition and Analysis of Human Faces and Facial Expressions: A Servey", *Pattern Recognition*, Vol. 25, No. 1, 1992, pp. 65-77.

[10] M. A. Shackleton and W. J. Welsh, "Classification of Facial Features for Recognition", *Proc. of Conf. on CVPR*, Hawaii, 1991, pp. 573-579.

[11] D. Terzopoulos, J. Platt, A. Barr and K. Fleisher, "Elastically Deformable Models", *Computer Graphics*, Vol. 21, July, 1987, pp. 205-214.

[12] D. Terzopoulos and K. Waters, "Analysis of Facial Images using Physical and Anatomical Models", *Proc. of Conf. on CVPR*, 1989, pp. 727-732.

[13] X. Xie, R. Sudhakar and H. Zhuang, "On Improving Eye Feature Extraction using Deformable Models", *Pattern Recognition*, Vol. 27, No. 6, 1994, pp. 791-799.

[14] A. L. Yuille, P. W. Hallinan and D. S. Cohen, "Feature Extraction from Faces using Deformable Templates", *International Journal of Computer Vision*, Vol. 8, no. 2, 1992, pp. 99-111.