Multiple Genetic Snakes for People Segmentation in Video Sequences

Lucia Ballerini

Dept. of Technology, Örebro University
70182 Örebro, Sweden
`lucia@aass.oru.se`

Abstract. In this paper we propose a method for finding people and segmenting their body parts in video image sequences. We propose the use of Genetic Snakes, that are active contour models, also known as snakes, with an energy minimization procedure based on Genetic Algorithms (GA). Genetic Snakes have been proposed to overcome some limits of the classical snakes, as initialization and existence of multiple minima, and have been successfully applied to images from different domains. We extend the formulation of Genetic Snakes in two ways, by adding an elastic force that couples multiple contours together and by applying them to color images. Experimental results, carried out on images acquired in our lab, are described.

1 Introduction

The analysis of image involving humans is currently a very active application domain in computer vision. This is motivated by the desire for improved man-machine interaction and therefore design machines which act more intelligently or human-like. A number of promising applications of "looking at people" are emerging. One of them is smart surveillance. In this case, an important information is to determine where and who are the humans in the environment.

The literature in this field is quite extensive and some special workshops devoted to this topic confirm the general interest in it. Due to space limitation we only mention few relevant works. See [1, 2] for surveys on the work on visual analysis for human movement and [3] for a review on techniques for detecting faces in images. Applications of active shape mmodel to human faces can be found in [4]. Pfinder [5] is a real-time system for tracking people which uses a multi-class statistical model of color and shape to segment a person. W^4 [6] is a real-time visual surveillance system for tracking people and their body parts in monochromatic imagery.

Several approaches for recognizing people have been proposed. Among them, the feature-based methods can roughly be subdivided in a segmentation step, where the person and its body parts are identified, a feature extraction step, where features like shape, size, color, texture are extracted in each body region, and a classification stage which produces the system output and hopefully the recognition. In this work, we focus on the first stage, the segmentation procedure.

We propose the use of Genetic Snakes [7], that are active contour models, also known as snakes [8], with an energy minimization procedure based on Genetic Algorithms (GA) [9]. Snakes optimization through GA proved to be particularly useful in order to overcome problems of the classical snakes related to initialization, parameter selection and local minima. New internal and external energy functionals have been proposed in our previous works and they have been successfully applied to a variety of images from different domains.

The purpose of this paper is to extend the Genetic Snakes model to handle complex contours, composed of distinct regions, allowing introduction of external knowledge expressed by additional energy terms. Each person is modeled with three contours that segment three main body parts (head, torso, legs) and using the knowledge of their relative position and size the contours are coupled together.

The organization of the paper is as follow: in Section 2 we briefly review active contours, the basic notions, their limitations and some improvements proposed in literature. In Section 3 we describe the Genetic Snakes model and we extend its formulation by the introduction of the multiple snake structure and of the binding energy. Experimental results are reported in Section 4.

2   Active Contours

Snakes are planar deformable contours that are useful in several image analysis tasks. They are often used to approximate the locations and shapes of object boundaries on the basis of the reasonable assumption that boundaries are piecewise continuous or smooth.

Representing the position of a *snake* parametrically by $\mathbf{v}(s) = (x(s), y(s))$ with $s \in [0, 1]$, its energy can be written as:

$$E_{snake} = \int_0^1 E_{int} \left[\mathbf{v}(s) \right] ds + \int_0^1 E_{ext} \left[\mathbf{v}(s) \right] ds \qquad (1)$$

where E_{int} represents the internal energy of the snake due to bending and it is associated with *a priori* constraints, E_{ext} is an external potential energy which depends on the image and accounts for *a posteriori* information. The final shape of the contour corresponds to the minimum of this energy.

In the original technique of Kass et al. [8] the internal energy is defined as:

$$E_{int} \left[\mathbf{v}(s) \right] = \frac{1}{2} \left[\alpha(s) \left| \frac{\partial \mathbf{v}(s)}{\partial s} \right|^2 + \beta(s) \left| \frac{\partial^2 \mathbf{v}(s)}{\partial s^2} \right|^2 \right]. \qquad (2)$$

This energy is composed of a first order term controlled by $\alpha(s)$ and a second order term controlled by $\beta(s)$. The two parameters $\alpha(s)$ and $\beta(s)$ dictate the simulated physical characteristics of the contour: $\alpha(s)$ controls the *tension* of the contour while $\beta(s)$ controls its *rigidity*.

The external energy couples the snake to the image. It is defined as a scalar potential function whose local minima coincide with intensity extrema, edges,

and other image features of interest. The external energy, which is commonly used to attract the snake towards edges, is defined as:

$$E_{ext}\left[\mathbf{v}(s)\right] = -\gamma|\nabla G_\sigma * I(x,y)|^2 \tag{3}$$

where $I(x,y)$ is the image intensity, G_σ is a Gaussian of standard deviation σ, ∇ is the gradient operator and γ a weight associated with image energies.

Due to the wide and successful application of deformable models, there exist survey papers focusing on different aspects of the model and its variants proposed in the literature [10–13].

The application of snakes and other similar deformable contour models to segment structures is, however, not without limitations. For example, snakes were designed as interactive models. In non-interactive applications, they must be initialized close to the structure of interest to guarantee good performance. The internal energy constraints of snakes can limit their geometric flexibility and prevent a snake from representing long tube-like shapes or shapes with significant protrusions or bifurcations. Furthermore, the topology of the structure of interest must be known in advance since classical deformable contour models are parametric and are incapable of topological transformations without additional machinery. Due to its own internal energy, the snake tends to shrink in case of lack of image forces, i.e. constant image backgrounds or disconnected object boundaries, and not to move towards the object.

Various methods have been proposed to improve and further automate the deformable contour segmentation process. See the above mentioned surveys for a review of some of them.

As concerns the energy minimization, the original model employs the variational calculus to iteratively minimize the energy. There may be a number of problems associated with this approach such as algorithm initialization, existence of local minima, and selection of model parameters. Simulated annealing [14], dynamic programming [15] and greedy algorithm [16] have been also proposed for minimization. However they are restricted either by the exhaustive searches of the admissible solutions either by the required accurate initialization.

Some authors propose the application of GA to active contours. Among those, MacEachern and Manku [17] introduce the concept of active contour state and encode the variants of the state in the chromosome of the genetic algorithm. Tanatipanond and Covavisaruch [18] apply GA to contour optimization with a multiscale approach. The fitness function is "trained" from a previously segmented contour. Ooi and Liatsis [19] propose the use of co-evolutionary genetic algorithms. They decompose the contour into subcontours and optimize each subcontour by separate GA working in parallel and co-operating. However, in these approaches the optimization is done in the neighborhood of the snake control points.

3   Multiple Genetic Snakes

Genetic Snakes are active contour models with an energy minimization procedure based on GA. An accurate description of implementation details along with a

discussion on the choice of the model coefficients can be found in [7]. Here we briefly summarize them.

The parameters that undergo genetic optimization are the positions of the snake in the image plane $\mathbf{v}_i = (x_i, y_i)$, for $i = 0, ...N$ where N is the total number of snake points. To simplify the implementation we use polar coordinates $\mathbf{v}_i = (r_i, \theta_i)$ with the origin in the centre of the contour. Actually, this point must lie inside the object, but its position may be arbitrary. The magnitudes r_i are codified in the chromosomes, while $\theta_i = 2\pi i/N$. The polar representation introduces ordering of the contour points and prevents the snake elements from crossing each other during evolution. The genetic operators can be implemented straightforward on this representation, no additional check is required to ensure that mutation and crossover produce valid individuals.

The fitness function is the total snake energy as previously defined in (1), where E_{int} and E_{ext} are defined in (2) and (3).

The initial population is randomly chosen in a region of interest defined by the user, and each solution lies in this region. This replaces the original initialization with a region-based version, enabling a robust solution to be found by searching the region for a global solution. The region of interest can be the image itself, so the solution can be searched in the whole image, making the initialization fully automatic.

Due to the genetic minimization procedure that does not require any restrictive assumption (continuity, existence of derivatives, etc.) on the fitness function, multiple cues (edge, color, motion, etc.) can easily integrated in the model.
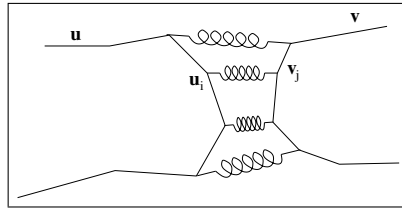

Fig. 1. Binding between adjacent snakes

In order to extend the model from a single contour to multiple contours we add an additional term to the internal energy, which we call binding energy. This is done by introducing an elastic force that connects together appropriate points of adjacent snakes, as shown in Fig. 1. The energy associated to the elastic force is assumed to be represented by the relation:

$$E_{bind}[\mathbf{v}(s), \mathbf{u}(t)] = \mu|\mathbf{v}(s) - \mathbf{u}(t)|^2 \tag{4}$$

where $\mathbf{v}(s)$ is one contour with parameter s, $\mathbf{u}(t)$ is another contour with parameter t and μ represents the elastic constant of the spring. The application points of the elastic forces are selected accordingly to the a priori knowledge of the structure to be segmented.

We also employ a modified version of the image energy to apply our model to color images. When the goal is to fit a snake to a boundary within an image, it is useful to preprocess the image with an edge detector so that the points of maximum gradient are emphasized. The edge detector most commonly employed uses the gradient of the image convolved with a Gaussian smoothing function.

We consider the gradient of the three components in the color space. Thus, the proposed image energy functional is composed of three terms and can be expressed as:

$$E_{ext}[\mathbf{v}(s)] = -\gamma_1|\nabla G_\sigma * I_1(x,y)|^2 - \gamma_2|\nabla G_\sigma * I_2(x,y)|^2 - \gamma_3|\nabla G_\sigma * I_3(x,y)|^2 \quad (5)$$

where $I_1(x,y), I_2(x,y), I_3(x,y)$ are the three color components of the image. In this work we did experiments using two color models: the RGB (red, green, blue) model and the HSI (hue, saturation, intensity) model. The weights are negatives so that local minima of E_{ext} correspond to maxima of the gradient, i.e. strong edges.

4   Application

We applied the model to segment people in images. The images were acquired with a webcam in the robotic lab at our institution. We have sequences with 9 different persons walking in the lab. The resolution is pretty low (160×120), but acceptable at this stage as we used these images only to test the method.

In this work we evaluated the application of multiple genetic snakes to the segmentation of the three body parts: head, torso, legs. Each person is modeled using three snakes, which represent the three body parts considered, chained together by means of the binding force. A region of interest is defined for each snake by setting its centre and the minimum and maximum magnitude for each $\mathbf{v}(s)$. Each snake is composed of 18 points, and the binding energy acts on three couples of consecutive points in each junction. In order to simulate the anatomical relationships, binding energy is assumed to act between the lower part ($\theta \simeq \pi/2$) of the first snake (head) and the upper ($\theta \simeq -\pi/2$) part of the second snake (torso). In the same way we connect the lower part of the second snake to the upper one of the third (legs).

The internal energy of the model is given by a weighted sum of E_{int} and E_{bind}, defined respectively in (2) and (4). In our implementation it is possible to use different α and β for each snake. The image energy E_{ext} is computed as defined in (5), by computing the Gradient of Gaussian ($\sigma = 0.3$) of the three color components (see Fig. 2). The fitness function is the total energy of the model i.e. the sum of the internal and image energy:

$$E = E_{int} + E_{bind} + E_{ext} \quad (6)$$

The model initialization is done automatically. We considered only walking people, therefore the three snake centres are aligned along the vertical middle axis. The a priori knowledge about standing people is used for initialization and for connecting together the snakes.

Fig. 2. An example of image and the Gradient of Gaussian of the three color components (RGB)

We assumed that our model is generally applied to a scene that consists of a relatively static situation such as an office, and a single moving person. We used the difference between two consecutive images to identify the moving edges and then we computed the centre of mass to find the initial position of snake centres. Background subtraction algorithms are commonly used to extract motion information. We decided not to use them to be able to generalize our method to images taken from a moving camera, for example mounted on a mobile robot. In this case the background will not be easily available.

The GA implementation adopted in this work is GAucsd-1.4 [20]. We used most of the default options proposed by the GAucsd package, i.e. Gray-code, fitness sigma scaling, two point crossover, roulette wheel selection. The parameters of the GA were: length of the genome = 162, population size = 3200, maximum number of generations = 162, crossover rate = 0.6, mutation rate = 0.00004. Each run takes about 1 min.

We performed several experiments varying the snake energy weighting coefficients. As more internal and external energy terms are considered in our model than in the classical snake, it is hard to determine the appropriate ratio between different forces in the total energy function of the model, due to lack of understanding on the effect of each force on the energy function.

The results showed in Fig. 3 have been obtained with following weights: $\alpha_1 = \alpha_2 = \alpha_3 = 1$, $\beta_1 = \beta_2 = \beta_3 = 1$, $\mu = 5$, $\gamma_1 = \gamma_2 = \gamma_3 = 1$. For this setting of weights we did 25 GA runs for each image. Equation (5) has been applied by separately using two color model (RGB and HSI). The left image of Figure 3 is obtained by using the RBG color model, the right image by applying (5) to the three components of the HSI model. We did not notice big differences related to the choice of the model. Results were not perfect. The segmentation accuracy can be easily improved by increasing the population size and the number of generation, but we decided to keep these parameters quite small in order to cope with the well known drawback of the GA, i.e. the time consuming.

We are also currently carrying out a study on the optimization of the snake weighting coefficients. This optimization lets the system himself learn which information are more important. Even the system is not presently working in real-time, mainly due to the use of GA, the optimization of the weights could be very useful for a possible future real-time application. For example, if we

Fig. 3. Examples of segmentation results

observe that some weights are very small, we can deduce that the relative info are irrelevant and avoid to compute them later.

5   Conclusions

In this paper a method for automatic segmentation of people in image sequences is described along with some results obtained with the proposed approach.

The proposed method allows to combine the a priori knowledge on the shape of people in the upright pose with the adaptative behaviour of active contours and GA. Binding energy allows to introduce adequate constraints on the geometry of the snake model. It is known that the snake model requires either a local minimizer with good initialization or otherwise a global minimizer. Genetic Snakes confront and overcome at the same time the two primary problems of initialization and optimization, and provide a global optimization with an automatic initialization.

Shape-tracking techniques relying on active contours have received a large amount of attention since the first work of Kass [8]. Several implementations have been developed and already applied to object tracking in video sequences. Most of them concentrates its effort on reducing the computational complexity to suite real-time contour tracking. It is not our purpose to compare our implementation with them, as our research is mainly directed to body part segmentation for people recognition than tracking.

The encouraging results reported prompt us that the method could be extended and applied to other images. Other extensions could consider the study of the parameters and the functionals governing the snake behaviour.

Acknowledgments

References

1. Gavrila, D.M.: The visual analysis of human movement: A survey. Computer Vision and Image Understanding **73** (1999) 82–98

 2. Aggarwal, J.K., Cai, Q.: Human motion analysis: A review. Computer Vision and Image Understanding **73** (1999) 428–440
 3. Yang, M.H., Kriegman, D.J., Ahuja, N.: Detecting faces in images: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence **24** (2002) 34–58
 4. Cootes, T., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models - their training and application. Computer Vision and Image Understanding **61** (1995) 38–59
 5. Wren, C.R., Azarbayejani, A., Darrel, T., Pentland, A.P.: Pfinder: Real-time tracking of the human body. IEEE Transactions on Pattern Analysis and Machine Intelligence **19** (1997) 780–785
 6. Haritaoglu, I., Harwood, D., Davis, S.L.: W^4: Real-time surveillance of people and their activities. IEEE Transactions on Pattern Analysis and Machine Intelligence **22** (2000) 809–830
 7. Ballerini, L.: Genetic snakes for medical images segmentation. In: Evolutionary Image Analysis, Signal Processing and Telecommunications. Volume 1596 of Lectures Notes in Computer Science., Springer (1999) 59–73
 8. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. International Journal of Computer Vision **1** (1988) 321–331
 9. Goldberg, D.E.: Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, Reading, MA (1989)
10. McInerney, T., Terzopoulos, D.: Deformable models in medical image analysis: A survey. Medical Image Analysis **1** (1996) 91–108
11. Jain, A.K., Zhong, Y., Dubuisson-Jolly, M.P.: Deformable template models: A review. Signal Processing **71** (1998) 109–129
12. Xu, C., Pham, D.L., Prince, J.L.: Image segmentation using deformable models. In Sonka, M., Fitzpatrick, J.M., eds.: Handbook of Medical Imaging. Volume 2. SPIE Press (2000) 129–174
13. Cheung, K.W., Yeung, D.Y., Chin, R.T.: On deformable models for visual patter recognition. Patter Recognition **35** (2002) 1507–1526
14. Storvik, G.: A bayesian approach to dynamic contours through stochastic sampling and simulated annealing. IEEE Transactions on Pattern Analysis and Machine Intelligence **16** (1994) 976 –986
15. Geiger, D., Gupta, A., Costa, L., Vlontzos, J.: Dynamic programming for detecting, tracking and matching deformable contours. IEEE Transactions on Pattern Analysis and Machine Intelligence **17** (1995) 294–302
16. Williams, D.J., Shah, M.: A fast algorithms for active contours and curvature estimation. CVGIP: Image Understanding **55** (1992) 14–26
17. MacEachern, L.A., Manku, T.: Genetic algorithms for active contour optimization. In: Proc. IEEE International Symposium on Circuits and Systems. Volume 4. (1998) 229–232
18. Tanatipanond, T., Covavisaruch, N.: An improvement of multiscale approach to deformable contour for brain MR images by genetic algorithms. In: Proc. IEEE International Symposium on Intelligent Signal Processing and Communication Systems, Phucket, Thailand (1999) 677–680
19. Ooi, C., Liatsis, P.: Co-evolutionary-based active contour models in tracking of moving obstacles. In: Proc. International Conference on Advanced Driver Assistance Systems. (2001) 58–62
20. Schraudolph, N.N., Grefenstette, J.J.: A user's guide to GAucsd 1.4. Technical Report CS92-249, Computer Science and Engineering Department, University of California, San Diego, La Jolla, CA (1992)