FACE DETECTION USING GRADIENT VECTOR FLOW

MAYANK VATSA, RICHA SINGH, P. GUPTA

Department of Computer Science & Engineering Indian Institute of Technology Kanpur
Kanpur INDIA, 208016
E-MAIL: {mayankv, richas, pg} @ cse.iitk.ac.in

Abstract:

Face detection is an important research area having wide applications in man-machine interface, visual surveillance and face recognition. We have proposed an algorithm for face detection using Gradient Vector Flow in gray level images which overcomes the problem for localization and initialization. The algorithm has been tested on various face databases and the result shows an accuracy of 97% with invariance to pose and orientation.

Keywords:

Face Detection; Active Contour Models; Gradient Vector Flow

1. Introduction

Face detection is an important task for various applications related to human face processing. It is to identify image regions that contain a face irrespective of the position, orientation and lightning condition. Various algorithms have been used for detecting faces like neural networks, template matching, skin color and principal component analysis. These algorithms generally have the problem related to initialization and localizing the face in the image. Another algorithm based on Active contour model, also referred as snakes are also being used to detect the edges and locate the face boundary. However there are some drawbacks with these models. For initialization of contour it depends on other mechanisms like interaction with users or some higher level image understanding processes to specify the approximate shape and starting position of the desired contour. Another problem with these models is that the initial contour should be closed to the boundary and should converge to concave boundaries.

In this paper Gradient Vector Flow (GVF) [8], which is a new class of forces for active contour models or snakes [1], is used for face detection from grey level images in controlled background and the accuracy is obtained to be 97%. Next section deals with the active contour models. In Section 3 face detection algorithm using Gradient Vector Flow is described. Section 4 presents the results of our algorithm on various face databases and finally conclusion and future work is given in Section 5.

2. Active Contour Models

An active shape model depicts the actual physical and higher-level appearance of features. Once released within a close proximity to a feature, active shape models interact with local image features (edges, brightness) and gradually deform to take the shape of the feature. There are generally three types of active shape models in the contemporary facial extraction research [1, 6, 7].

The first type uses a generic active contour called snakes, introduced by Kass et al. [1]. Active contours, or snakes, are commonly used to locate a head boundary [2, 3, 4, 5]. A snake is first initialized at the proximity around a head boundary. It then locks onto nearby edges and subsequently assumes the shape of the head. Gunn and Nixon [2] made the term sensitive to the image gradient so that the contour is convergent toward edge locations. They have also solved the problem of inefficiency in extracting non-convex features due to the tendency to attain minimum curvature by introducing a parameterized snake model for face and head boundary extraction. The parameterized model biases the contours toward the target shape and there by allows it to distinguish false image features and not be trapped by them. In addition to gradient information, the external energy term in [5] includes a skin color function which attracts the contour to the face region. In Huang and Chen [3] and Lam and Yan [6] fast iteration methods (greedy algorithms) are used for convergence.

The second type uses deformable templates, introduced by Yuille et al. [7], to take into account facial features and to improve the performance of snakes by incorporating global information of the eye. It also improves the reliability of extraction process.

Cootes et al. [8] have later proposed the use of a new generic flexible model which they have termed as smart snakes and Point distributed models (PDM) [2] to provide an efficient interpretation of the human face. PDM is a

0-7803-7865-2/03/$17.00 ©2003 IEEE

compact parameterized description of the shape based on statistics. The architecture and the fitting process of PDM are different from the other active shape models.

3. Face Detection

In this paper we have used Gradient Vector Flow (GVF) [9] for detecting the faces. GVF [9] is a modified form of the active contours [1]. The active contour model, i.e. snakes, introduced by Kass et al. [1], deforms a contour to lock onto the features of interest within an image. Usually the features are lines, edges and/or object boundaries. The snakes have been used for detecting desired and actual boundary if the approximate boundary of the object is given. A traditional snake is a curve $X(s) = [x(s), y(s)]$ for $s \in [0, 1]$, that moves through the spatial domain of an image $I(x, y)$ to minimize the energy function defined in [1] as

$$E_{energy} = \int_0^1 E_{energy}(X(s))ds \quad (1)$$

$$E_{energy} = \int_0^1 E_{in}(X(s)) + E_{ext}(X(s))ds \quad (2)$$

where E_{in} is the internal energy that provides smoothness to the contour and E_{ext} is the image energy calculated from an image data which determines the relationship between snake and the image.

The internal energy, E_{in}, of snake which is found to be the sum of the two terms such as elasticity and bending, is defined [1] as

$$E_{in} = [\alpha \left|\frac{dX(s)}{ds}\right|^2 + \beta \left|\frac{dX(s)^2}{ds^2}\right|^2]/2 \quad (3)$$

where α and β are the weighted parameters to control snake's tension and rigidity respectively. By controlling them we can control the limitation up to which snake is allowed to stretch and bend. Note that α represents the importance of the elasticity term in snake's internal energy while β describes the significance of stiffness term in snake's internal energy. These parameters are to be fixed for detecting the boundary of interested object in the image. The term $[\alpha|\frac{dX(s)}{ds}|^2]$ is a first-order derivative of the position of the snake and controls the tension in snake. Hence, it is called the elasticity energy $E_{elastic}$ of snake. Again the term $[\beta|\frac{dX(s)^2}{ds^2}|^2]$ is a second-order derivative of snake's position and controls the rigidity of snake, i.e. its ability to bend. Therefore, it is called the bending energy $E_{bending}$ of snake [1, 9].

Using Eq. (3) in Eq. (2) we get

$$E_{energy} = \int_0^1 \frac{1}{2}[\alpha|X'(s)|^2 + \beta|X''(s)|^2] + E_{ext}(X(s))ds \quad (4)$$

Again from [1], the external energies which lead the active contour towards step edge are found as

$$E^1_{ext}(x, y) = -|\nabla I(x, y)|^2 \quad \text{and}$$

$$E^2_{ext}(x, y) = -|\nabla(G_\sigma(x, y) * I(x, y))|^2 \quad (5)$$

where $G_\sigma(x, y)$ is a two-dimensional Gaussian function with standard deviation σ and ∇ as the gradient operator.

From the definition of snakes [1], minimization of Eq. (4) gives us the amount of energy that is required for snake to shrink to next position. Therefore, at the end of deformation, i.e. on the edges, this energy should be equal to zero. This implies that internal force of the snake is equal to the force by the image gradient. It can be expressed as [1, 9]

$$E_{enegy} = \alpha X''(s) - \beta X^{iv}(s) - \nabla E_{ext} = 0 \quad (6)$$

Note that $F_{in} = \alpha X''(s) - \beta X^{iv}(s)$ is the internal force [1, 9] that tries to stop the stretching and bending and $F^{(p)}_{ext} = \nabla E_{ext}$ is the external potential that tries to pull the snake towards the desired image edge. Then the force balance equation, defined in [1], becomes

$$F_{in} + F^{(p)}_{ext} = 0 \quad (7)$$

To solve Eq. (6), the snake is made dynamic i.e. X as the function of time t and s then Eq. (6) becomes

$$X_t(s,t) = \alpha X''(s,t) - \beta X^{iv}(s,t) - \nabla E_{ext} \quad (8)$$

This gives the solution of Eq. (6) which is the traditional snake [1]. But it has lots of limitations and is difficult to be used for face detection. The main problem is that its deformation is initialized normally and not done automatically. To produce a suitable set of parameters for snake initialization is another difficulty. Parameters/weights of energies have large impact in snake's behaviors and totally control the performance of deformation process. Automatically, choosing a set of parameters which can be used for a particular object in image of interest is a difficult task. Therefore, these constants are usually up for user to decide. When we initialize snake, the control points are normally put around the snake outside the object of interest. This becomes a problem if the object is too close to the border of the picture frame. Also, even if there is enough space to initialize snake, it may hit the picture frame during deformation process and produce an error which interrupts the process. Therefore, accurate edge tracing is not possible to complete. If object in the image has an edge which shows a concavity then snake may have a problem in tracing the edge at the concaved part. This is because the capture range at these parts of edges is too far from the snake. Also even if they are not far from snake, the gradient force on each wall of concavity effectively cancels each other and leaving the snake with no gradient force to work and hence no deformation can take place.

C. Xu. et. al. [9] then proposed a new algorithm for active contour models which was based on a new type of external force field, called Gradient Vector Flow [9]. GVF snake keeps the strong and useful part of the traditional snake: the internal force. But, instead of image gradient, it creates its own force field called the GVF force field [9] which was the dense vector field derived from image by minimizing energy function in a variational framework. The minimization is achieved by solving a pair of decoupled linear partial differential equations which defuses the gradient vectors of a gray level edge map computed from the image.

The new static external force field $F_{ext}^{(p)} = A(x, y)$ defined in [9], comes from the premise of mathematics of Helmholtz theorem. It states that most general static field can be decomposed into two components: irrotational i.e. curl-free component and solenoidal i.e. divergence-free component. Now (8) becomes

$$X_t(s,t) = \alpha X''(s,t) - \beta X^{iv}(s,t) + A \qquad (9)$$

where A is the external force field and which is independent of t. The solution of the above equation gives a GVF snake [9]. If we look at the traditional external force field of snake, the only useful forces are the ones near the edges. The rest of the image normally has very little variation in gradient and hence it is ignored by snake deformation. Therefore, if we can find an algorithm to further extend the gradient force near the edges into the rest of the image, we could increase snake's capture range and have a better deformation at concaved edges. The algorithm in [9] goes through a computational diffusion process. The concept is to create a new vector field called Gradient Vector field [9] from an image data. This force is located in a homogeneous region. Due to the diffusion process a force is assigned which points towards the edges of the objects. Now, the edge map $T(x, y)$ derived in [9] from the image $I(x, y)$ has the property that it is closest to the image boundaries. So, for i = 1 and 2

$$T(x, y) = -E^{i}_{ext}(x, y) \qquad (10)$$

Note that from [9], for E^{i}_{ext} the value of i varies from 1 to 4. But in our case as the face image is gray level we need only E^{1}_{ext} and E^{2}_{ext}, and the Gradient Vector Force field has been defined in [9] as

$$A(x, y) = (u(x, y), v(x, y))$$

which minimizes the energy function E in (4) to

$$E = \iint \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla T|^2 |A - \nabla T|^2 \, dxdy \quad (11)$$

Computing the GVF field [9] using variations of calculus in [10], we get the following Euler equations

$$\mu\nabla^2 u - (u - T_x)\cdot(T_x^2 + T_y^2) = 0 \qquad (12)$$

$$\mu\nabla^2 v - (v - T_x)\cdot(T_x^2 + T_y^2) = 0 \qquad (13)$$

where ∇^2 is the Laplacian operator. In [10] GVF is then generalized to three dimensions with u and v as functions of time and a GVF deformable surface is implemented.

$$u_t(x,y,t) = \mu\nabla^2 u(x,y,t) - (u(x,y,t) - T_x(x,y)\cdot(T_x(x,y)^2 + T_y(x,y)^2) \qquad (14)$$

$$v_t(x,y,t) = \mu\nabla^2 v(x,y,t) - (v(x,y,t) - T_x(x,y)\cdot(T_x(x,y)^2 + T_y(x,y)^2) \qquad (15)$$

The steady-state solution of the linear parabolic equations is the solution of these equations. These equations are decoupled and hence can be solved as the separate scalar partial differential equations in u and v. The above equations are known as generalized diffusion equations and are from the description of desirable properties of external fields for active contours.

The above solution of GVF is valid for gray level face images. To compute GVF for gray level face images, the edge-map function must first be calculated. $T(x, y) = -|\nabla(G_\sigma(x, y)*I(x, y))|^2$ (Eq. 5) is robust in case of noisy images. Any noise removal algorithm like median filtering, morphological filtering and anisotropic diffusion can be used to improve the edge maps. Now, GVF is computed in the usual way using Eq. (14) and Eq. (15). The range of α and β are set to be in the range of 2.0 to 3.0. Thus we get the desired boundary of the face image. The four corners of the boundary are then extracted to draw a bounding box of it and this part of the image is cropped to get the detected face image.

4. Results

The proposed algorithm is tested on the Pentium III [800 MHz with 64 MB RAM only] and on three databases – (i) AR database [16], (ii) the ORL – AT&T database [17] and (iii) IITK database [18] (the database maintained at Indian Institute of Technology Kanpur India). In this section we have discussed experimental results of our algorithms on three databases.

The sample results of the face detection module are shown in Fig. 1. GVF performs well in controlled background images taken from the three databases. The three sets of results are shown on IITK Face Database. Consider the images obtained in case A, B or C. Results of all three databases are shown in Table 1. In the figure showing the results, first image is the original image containing face which also shows the bounding box drawn by extracting the boundaries of the face. Second one gives the edge map obtained while third is the gradient of the edge map. Fourth image is showing the GVF field and fifth image is the output from the detection module i.e., the detected face. It contains the image within the bounding box.

The results show an accuracy of about 97% on the three databases in which the maximum accuracy is obtained on the AR database of 97.88%. The accuracy on ORL database is 96.50% and on the IITK database it is found to be 97.20%.

5. Conclusion

In this paper face detection using Gradient Vector Flow [9] has been implemented. The algorithm has been tested on three databases AR, ORL and IITK face database and the accuracy obtained is up to 97%. This overcomes the problem of face localization and initialization. Also the algorithm is robust to changes in pose and orientation.

The future work will be the real time implementation of this algorithm with considerations to its time complexity.

Acknowledgements

Authors would like to acknowledge Dr. Aleix Martinez for providing the AR Face Database.

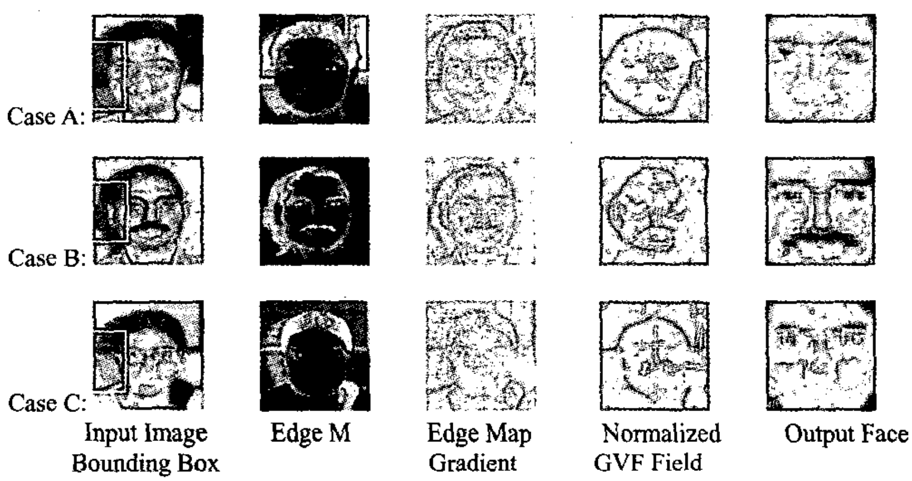

Figure 1 Results of the Face Detection Module

Table 1 Results of Face Detection on Three Databases

Database and the number of images in it	Accuracy
ORL - AT&T Database	96.50 %
AR Database	97.88 %
IITK Database	97.20 %

References

[1] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active Contour Models" International Journal of Computer Vision, Vol.1, No.4, pp.321-331, 1987.

[2] S. R. Gunn and M. S. Nixon, "A dual active contour for head and boundary extraction", In IEE Colloquium on Image Processing for Biometric Measurement, London, Apr. 1994.

[3] C. L. Huang and C. W. Chen, "Human facial feature extraction for face interpretation and recognition", Pattern Recognition, Vol. 25, 1992, 1435–1444.

[4] Nikolaidis and I. Pitas, "Facial feature extraction and pose determination", Pattern Recognition Vol. 33, 2000, 1783–1791.

[5] H. Wu, T. Yokoyama, D. Pramadihanto, and M. Yachida, "Face and facial feature extraction from color image", In IEEE Proceedings of 2nd International Conference on Automatic Face and Gesture Recognition, Vermont, Oct. 1996, pp. 345–349.

[6] K. M. Lam and H. Yan, "Fast greedy algorithm for locating head boundaries", Electronics. Letters. 30, 1994, pp. 21–22.

[7] L. Yuille, P. W. Hallinan, and D. S. Cohen, "Feature extraction from faces using deformable templates", International Journal of Computer Vision 8, 1992, pp. 99–111.

[8] T. F. Cootes and C. J. Taylor, "Active shape models—smart snakes", In Proceedings of British Machine Vision Conference, 1992, pp. 266–275.

[9] Chenyang Xu and Jerry L. Prince, "Snakes, Shapes and Gradient Vector Flow", IEEE Transaction on Image Processing, Vol.7, No.3, pp. 359-369, March 1998.

[10] Moritz Storring, Hans J.Andersen, Erik Granum, "Skin Color Detection under Changing Lightning Conditions", 7th Symposium on Intelligent Robotics Systems, July 1999, Coimbra, Portugal.

[11] R. Feraud, O. J. Bernier, J.-E. Villet, and M. Collobert. "A fast and accurate face detector based on neural networks". IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 1, pp.42-53, 2001.

[12] Buciu, C. Kotropoulos and I. Pitas, "Combining Support Vector Machines for Accurate Face Detection", IEEE International Confernce on Image Processing, 2001.

[13] O. Jesorsky, K. J. Kirchberg, R. W. Frischholz, "Robust Face Detection Using the Hausdorff Distance", International Conference on Audio-Video Biometric Personal Authentication 2001, pp. 90-95.

[14] I. Craw, H. Ellis, and J. Lishman, "Automatic extraction of face features" Pattern Recognition Letters, 5 pp. 183-187, 1987.

[15] I. Craw, D. Tock, and A. Bennett. "Finding face features". In Proceedings of the Second European Conference on Computer Vision, pp. 92-96, 1992.

[16] http://rvl1.ecn.purdue.edu/~aleix/aleix_face_DB.html

[17] www.uk.research.att.com/facedatabase.html

[18] www.cse.iitk.ac.in/users/pg