Localization of Facial Features using Pulse-Coupled Neural-Network and Active Contours

R. El Sayed* T. A. Yousef* A. M. Badawi**

*Teaching Assistant, Systems & Biomedical Engineering Dept., Faculty of Engineering, Cairo University.
**Ph.D., Systems & Biomedical Engineering Dept., Faculty of Engineering, Cairo University.

Abstract

Pulse-Coupled Neural-Network (PCNN) is a new promising image processing tool. Since the Pulse-Coupled Neural-Network firing scheme depends mainly on the shapes of the image, it is suitable for automated face segmentation because face images contains the same shape. In this paper, we present an algorithm for automatic facial features (eye, nose and mouth) detection in face images for different expressions based on PCNN-guided active contour models (snakes).

Keywords

Facial features detection, pulse-coupled neural-network, Active contours (snakes) and computer vision.

1. Introduction

Automatic face recognition has long been studied as it has a wide potential for applications. Several systems have been developed to identify faces from small face population via detailed face feature analysis, or by using neural nets or through model based approaches [1,3,5,21,23,25]. However, many of these systems suffer from robustness against changes in facial expressions. In addition, face is a rich source of information about human behavior. Facial expressions indicate emotion, pain, and brain function and regulate social behavior. In this paper, we aim to localize the facial features (eye, nose and mouth) in faces with different expressions. These localizations can then be the input for another module to run certain measurements to identify both the face and its expression.

Many researchers described the use of specialized techniques to locate facial features within a search area depending on the color, gray level, edges and time sequences of images [1-5]. For example, after skin color segmentation, Sobottka and Pitas extracted facial features from the detected face candidates. Facial features were detected from the rest of face by their low gray level. They applied preprocessing steps: grayscale erosion and an extreme sharpening operation, to enhance these facial features [3].

In this paper we consider using the PCNN, which is known for its good segmentation capabilities to segment the face image so that localization of facial feature is much easier. Active contour models (snakes) are used to capture the contour of each facial feature after being detected in the binary images generated by the PCNN. Kass et al. used active contour models (snakes) for tracking lips in image sequences [6]. A snake was initialized on the lips in a face image and it was able to deform and accurately track lip movements. Because the snakes used do not incorporate prior knowledge about expected shapes, this approach is easily confused by other structures present in the image and occlusion. Using the PCNN as a guide for the snakes provides both automatic detection as well as immunity against occlusion that is found in the original image but absent in the PCNN generated image sequence [10,11].

We have used the Pulse-Coupled Neural Network (PCNN) as the major component of our algorithm. PCNN was developed from studies of the visual cortex of small mammals made by e.g. Eckhorn [7,8]. The implementation of the PCNN was first carried out by Johnson [9,10] and the work has been extended by several others [11–14]. The stimulus for the PCNN is typically a 2D gray scale or color image. The PCNN will transform this input into a series of binary images. The internal feedback among the neurons allows similar pixels, i.e. pixels of similar intensity, to cluster into groups. The PCNN has the ability to map spatial input to temporal output (i.e. convert the input magnitude to a time domain pulses).

PCNN are different from classical neural network models, since no training is required. The properties of the PCNN can be adjusted by changing threshold levels and decay time constants. Other interesting features include the ability to detect edges, find texture information, etc. The PCNN has been shown to do segmentation and object isolation in a variety of applications [10 -12].

We noticed that PCNN firing scheme depends mainly on the shapes of the image, This is exactly what makes PCNN suitable for automated face segmentation, because face images contains the same shapes so the firing scheme is almost the same in all images (Fig. 3). With the aid of active contours, we can locate the facial features in most of the tested face expressions.

2. Pulse-Coupled Neural Network

Roughly a decade ago researchers were presenting models of the mammalian visual cortex [7,8]. Digital simulations revealed algorithms that were capable of extracting segments that were inherent in the original image [10 -14].

The original PCNN model considered in this paper is due to Eckhorn [1]. The typical PCNN neuron (fig. 1) has a feeding and linking input that is then combined in a second order fashion and then compared to a dynamic threshold [16]. The equations for a single iteration of the PCNN are:

$$F_{ij}[n] = e^{\alpha_F \delta_n} F_{ij}[n-1] + S_{ij} + V_F \sum_{kl} M_{ijkl} Y_{kl}[n-1] \qquad (1)$$

$$L_{ij}[n] = e^{\alpha_L \delta_n} L_{ij}[n-1] + V_L \sum_{kl} W_{ijkl} Y_{kl}[n-1] \quad (2)$$

$$U_{ij}[n] = F_{ij}[n](1 + \beta L_{ij}[n]) \quad (3)$$

$$Y_{ij}[n] = \begin{cases} 1 & \text{if } U_{ij}[n] > Q_{ij}[n], \\ 0 & \text{Otherwise} \end{cases} \quad (4)$$

$$Q_{ij}[n] = e^{\alpha_Q \delta_n} Q_{ij}[n-1] + V_Q Y_{ij}[n] \quad (5)$$

Where S is the stimulus, F is the feed, L is the link, U is the internal activity, Y is the pulse output and Q is the dynamic threshold. The local connections M and W are fixed (usually Gaussian). Through these local connections the activation of a neuron adds to the internal activity of the neighboring neurons. Groups of neurons receiving similar stimulus that are spatially close to each other tend to synchronize pulses. This is the foundation of the inherent segmentation ability of the system. The PCNN is a one layer neural network that requires no training. The input gray-level image is the stimulus S that activates the neurons and the output of the neuron is the matrix Y (called a Pulse image) that result by iteratively computing equations (1-5).

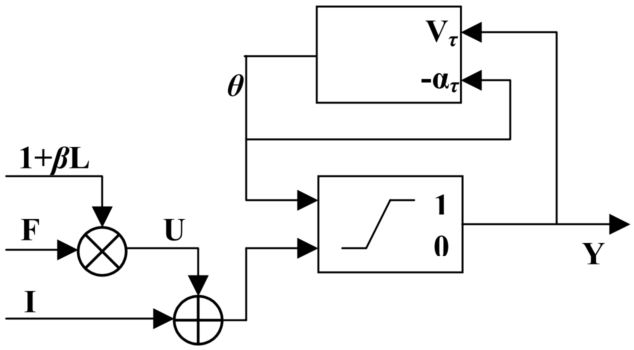


Fig. 1. The PCNN neuron

3. Active Contour Model

Deformable models have proved their potentiality in the last two decades and have given good results in computer vision research and many other applications.

The concept of the deformable model is based upon the existence of internal energy or some forces in the model trying to preserve some internal features in the model these features may be lengths, angular measures, curvature…etc. While the model is trying to be influenced by its internal energy or forces, another external energy or forces oppose it, which limits its response.

Based on this simple criterion many models have been proposed and found many applications in field of Computer Vision, Image Registration, Image Contouring and many other applications.

A model for representing image contours in a form that allows interaction with higher-level processes has been proposed by Kass et al. [6]. This active contour model is defined by an energy functional and a solution is found using techniques of variational calculus

The energy functional being minimized has continuity term and curvatures term which serve as smoothness constraint in addition to the image energy which is considered as external energy term which guides the active contour towards image features which minimize the contour total energy.

Total Snake energy

$$E^*_{snake} = \int_0^1 E_{\text{int}}(v(s)) + E_{image}(v(s))ds \quad (6)$$

Where

$$E_{\text{int}} v(s) = (\alpha(s)\,|v_s|^2 + \beta(s)\,|v_{ss}|^2)/2 \quad (7)$$

Where $|v_s(s)|^2$ is defined as follows:

$$|v_s(s)|^2 = (x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 \quad (8)$$

Where (x_i, y_i) and (x_{i-1}, y_{i-1}) are two successive points lying on the contour.

$$|v_{ss}(s)|^2 = (x_{i-1} - 2x_i + x_{i+1})^2 + (y_{i-1} - 2y_i + y_{i+1})^2 \quad (9)$$

Where α and β are the contour internal energy parameters which weights each term relative to the other usually α=1 and β=1.

$$E_{image} v(s) = w_{line} E_{line} v(s) + w_{ledge} E_{edge} v(s) \quad (10)$$

E_{line} v(s)= I (x, y) which is the gray level intensity at point (x, y).

This means that if w_{line} is positive the contour will possess minimum energy at dark areas, and if it takes negative values positive the contour will possess minimum energy at bright areas.

E_{edge}v(s)= | ∇I (x, y) |.

The previous equations have shown how to calculate the energy of the contour point at any location (x, y). the absolute values of the parameters x, y doesn't have any significant meaning yet their relative values plays important role in the contribution of each energy term.

4. Image Dataset and Preprocessing

Our dataset for human face images is from the "Biomedical Face Database" which contains a set of faces taken on March 2000 at the Biomedical Laboratory of Biomedical and Systems Engineering, Faculty of Engineering, Cairo University (Fig. 2). It contains seven different images of 100 distinct subjects. For all of the subjects, the images were taken with the same lighting conditions, facial expressions (neutral, happy, sad, surprise, angry, disgust and fear) and with no eyeglasses. All the images are taken against a dark homogeneous background and the subjects are in up-right, frontal position (with tolerance for some side movement and zooming effects). All the images are resized to a size of 256x256 and histogram equalized before being used in any processing. All pixels values are scaled to the range [0,1].

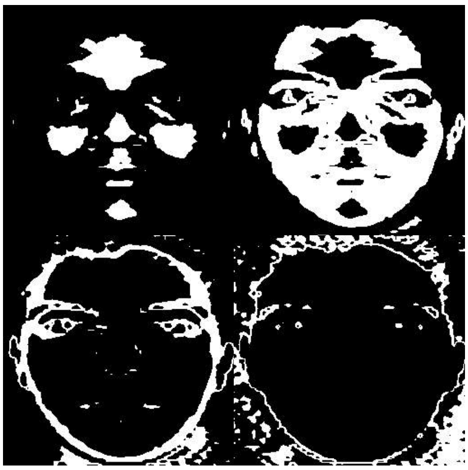
Fig. 2. Samples of the "Biomedical Face Database" showing the seven expressions for each subject.

5. Localization Technique

The PCNN was used to generate a sequence of four PCNN pulse images of the input gray-scale face image. The first image is almost black (due to an initial high threshold value) so no data can be extracted from it. The second image contains the brightest areas in face (forehead, cheeks and nose). The third image has the majority of the face points firing, this pulse image is the most interesting one because it cleanly isolates the face area excluding the hair and background so we initialize an active contour (snake) [6], near the image boundary and it is left to automatically adhere to the face contour in the image excluding the trouble-makers hair styles and background (fig. 3). We call this contour "The Face Contour".

5.1. Nose Localization

Having the Face Contour, we get the point slightly below (20 pixels down) the center of the smallest rectangle enclosing this contour; this point is the initial nose position.
Then we search in the second pulse image for the nearest white area near this point, this area should be the nose area. We scroll down this area to the lowest ON pixel, this point is the nose tip (fig. 3).



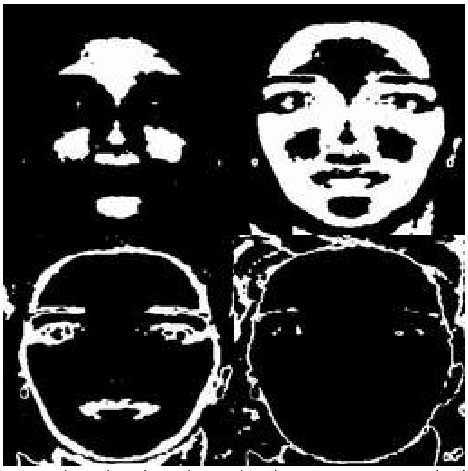


Fig. 3. Two Samples showing the pulse image sequence obtained by the PCNN on the face image. The images shown are the second through the fifth pulse images. Firing scheme is almost the same in the tow sequences

5.2. Mouth Localization

To get the mouth, we initiate an active contour below the nose tip obtained in the previous step and bounded by the snake obtained around the whole face. This new snake is left to adhere to the points representing the mouth in the fourth pulse image. The outer rectangle enclosing this snake is shown in figure (3).

5.3. Eye Localization

In the fourth pulse image, we initiate another two new active contours to catch the left and right eyes; these contours are initially bounded as follows:

- The nose tip from the bottom side.
- The Face Contour from the top and lateral side.
- The vertical axis of the nose tip from the axial side for each eye.

Due to the interference between the eye and the brow in the PCNN sequence images, the active contour could not in most of the cases isolate the eye from the its brow. So, we used a template-matching algorithm to identify exactly the eye depending on the elliptical nature of the eye [20]. Given the location of the eye (the bounded rectangle of its elliptic shape), A. H. Mohammed extracts all the eye segments (i.e. Iris, pupil) and iris texture [25].

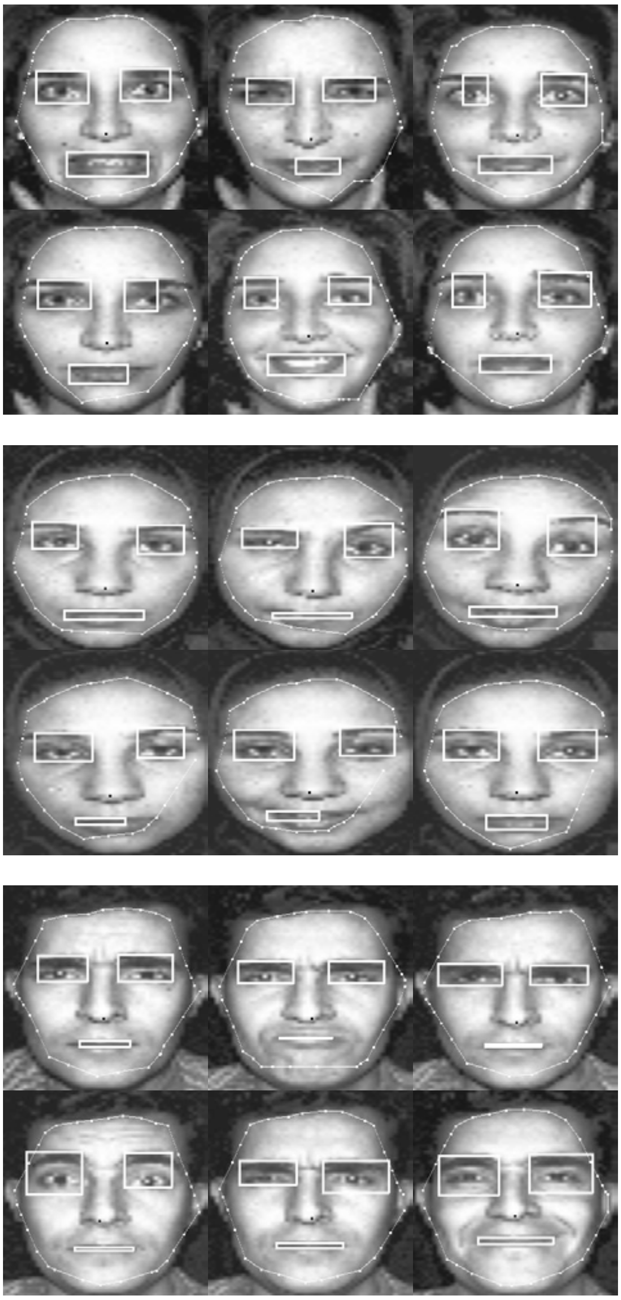


Fig. 4. Shows the results for 4 subjects (male, female, coated-hair female and the third author) in different 6 facial expressions. The black point is the nose tip

6. Results

From the 700 images in our data set (100 subject, 7 expressions per subject), nose and mouth are correctly localized in about 609 images (87%). Eyes were correctly localized in about 487 images (69.57%) raised to 530 (75.71%) after the use of template matching.

We have tried to apply this algorithm on another data sets, and it seems that the selected number of images from the PCNN sequence may differ from a data set to another. In our data set we used pulse images two and three, while in another data set (MIT data set), images four and six seem to give better results, This is due to differences in lightening conditions, background patterns and zoom factors iv different datasets. A normalizing and preprocessing step are necessary before applying the proposed algorithm to any other datasets.

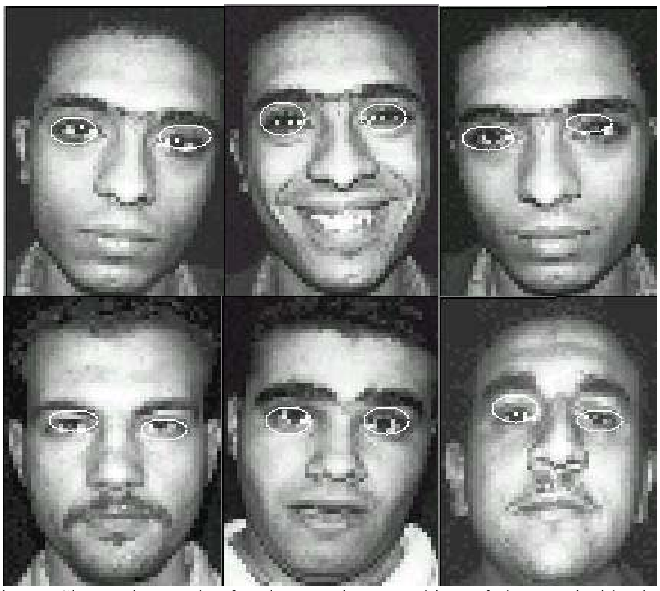


Fig. 5. Shows the results for the template matching of the eye inside the rectangle bounding the active contour around the eye. The top three results are for the same subject in different expression, the other from different subject in different expressions.

7. Conclusion

Using PCNN in localization of facial features seems to be promising. It is not computationally exhaustive or time consuming relative to many other algorithms [1,3,21]. It does not need training. It is robust against different facial expressions and to spatial shifts in the image (as long as the whole face is within the image) and can tolerate slight side movements and zooming effects. However, it still needs more enhancements to improve its robustness against changes in orientation of faces i.e. the rotation angle from upright position.

In the given here algorithm, the most traditional active contour model was employed [17]. Updated with better deformable models [18-19], we believe that, this method can improve the detectability of eye and mouth areas, however, our emphasis here was on the PCNN segmentation capabilities in facial images.

8. References

1. L. C. Desilva, K. Aizawa, M. Hatori, "Detection and Tracking of Facial Features by Using Edge Pixel Counting and Deformable Circular Template Matching," IEICE Trans. Inf. & Syst. vol. E78-D, no. 9, Sept. 1995, pp. 1195-1207.
2. Karin Sobottka, Ioannis Pitas, "Segmentation and Tracking of Faces in Color", *Proceedings of second International Conference on Automatic Face and Gesture Recognitio*n, Killington, Vermont, Oct. 1996, pp. 236-241.
3. Qing Song, B.Eng., "Robust face detection and pose estimation", A Proposal for Doctoral Research, (1998).
4. Choong Hwan Lee, Jun Sung Kim and Kyu Ho Park, "Automatic Human Face Location in a Complex Background Using Motion and Color Information," Pattern Recognition, Vol. 29, No. 11, pp.1877-1889, 1996.

5. A. Gee, R. Cipolla, “Fast Visual Tracking by Temporal Consensus,” Image and Vision Computing, Vol. 14, pp. 105-114, 1996.
6. M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: Active Contour Models”, International Journal of Computer Vision 321-331 (1988).
7. R. Eckhorn, H. J. Reitboeck, M. Ardnt, and P. Dicke, “Feature Linking via Synchronization Among Distributed Assemblies: Simulation of Results from Cat cortex”, Neural Computation 2, pp. 293-307 (1990).
8. I. A. Rybak, N. A. Shevtsova, V. A. Sandler, Neurocomputing 4, 93-102, (1994).
9. Johnson, J. L., “Pulse-Coupled Neural Networks”, Proceedings Adaptive Computing: Mathematics, Electronics and Optics, S. S. Chen and H. J. Caulfield, Orlando, Vol. CR55, pp. 47-76 (1994)
10. Johnson, J. L., “Pulse-Coupled Neural Nets: Translation, Rotation, Scale, Distortion and Intensity Signal Invariance for Images”, Ap. Opt. Vol. 33,No.26, pp. 6239-6253, (1994)
11. Raganath, H. S., Kuntimad, G., “Image Segmentation Using Pulse-Coupled Neural Networks”, Proceedings of IEEE International conference on Neural Networks, Orlando FL., pp. 1285-1290, (1994)
12. Raganath, H. S., Kuntimad, G., Johnson, J.L., “Pulse-Coupled Neural Networks for Image Processing.”, Proceedings of IEEE Southeastcon, Raeigh, NC, (1995).
13. T. Lindblad and J. Kinser, Image Processing using Pulsed Coupled Neural Networks, SpringerVerlag, London, (1998).
14. Raganath, H. S., Kuntimad, G., “Iterative Segmentation using Pulse-Coupled Neural Networks”, Proceedings Applications and science of Artificial Neural Networks II, Orlando, FL, Vol. 2760,pp.543-554, (1996).
15. J. M. Kinser, Proc. Of SPIE, 3728, Stockholm, 222-229, (1998).
16. J. M. Kinser, Foveation by a Pulse-Coupled Neural Network, IEEE-NN, Vol. 10,No. 3, (1999).
17. D. Terzopoulos and R. Szeliski. “Tracking with Kalman Snakes”. In A. Blake and A. Yuille, editors, Active Vision. The MIT Press, 1992.
18. T.F. Cootes and C.J. Taylor, "Active Shape Models- -Smart Snakes", Proc. British Machine Vision Conference, 1992, p266, Springer Verlag, 1992.
19. S. C. Zhu, T. S. Lee, and A. L. Yuille, "Region competition: Unifying snakes, region growing, energy /Bayes/MDL for multi-band image segmentation," in Int. Conf. on Computer Vision, (Boston, MA), pp. 416--423, 1995.
20. Xiaoguang Jia, “Extending the Feature Set for Automatic Face Recognition”, (1993).
21. Yingli Tian, T. Kanade and J. F. Cohn, " Recognizing Action Units for Facial Expression Analysis ", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 2, February, 2001.
22. Henry A. Rowley, Shumeet Baluja, Takeo Kanaade, “Rotation Invariant Neural Network-Based Face Detection”, IEEE Transactions on Pattern Analysis and Machine Intelligence, (1997).
23. Baback Moghaddam, Alex Pentland, “Probabilistic visual learning for object detection”, In Fifth International Conference on Computer Vision, pages 786–793, Cambridge, Massachusetts, June 1995. IEEE Computer Society Press, (1995).
24. Young Ho Kwon and Niels da Victoria Lobo, “Face Detection Using Templates,” Proceedings-International Conference on Pattern Recognition, pp. 764-767, (1994).
25. A. H. Mohammed, “Optimization of Iris Texture Image, Representing and Coding”, a thesis for Master of Science, System & Biomedical Eng. Dept., Faculty of Engineering, Cairo.