Role of linking parameters in Pulse-Coupled Neural Network for face detection

Young Wan Lim*, Jin Hee Na**, and Jin Young Choi**

* Mobile Multimedia Laboratory, LG Electronics Institute of Technology
(E-mail: ywlim@neuro.snu.ac.kr)
**School of Electrical and Computer Engineering, Seoul National University, Seoul, Korea
(Tel : +82-2-872-7283; E-mail: {jhna | jychoi}@neuro.snu.ac.kr)

Abstract: In this work, we have investigated a role of linking parameter in Pulse-Coupled Neural Network(PCNN) which is suggested to explain the synchronous activities among neurons in the cat cortex. Then we have found a method to determine the linking parameter for a satisfactory face detection performance in a given color image. Face detection algorithm which uses the color information is independent on pose, size and obstruction of a face. But the use of color information encounters some problems arising from skin-tone color in the background, intensity variation within faces, and presence of random noise and so on. Depending on these conditions, PCNN's linking parameters should be selected an appropriate values. First we obtained the mean and variance of the skin-tone colors by experiments. Then, we introduced a preprocess that the pixel with a mean value of skin-tone colors has the highest level value (255) and the other pixels have values between 0 and 255 according to normal distribution with a variance. This preprocessing leads to an easy decision of the linking parameter of the Pulse-Coupled Neural Network. Through experiments, it is verified that the proposed method can improve the face detection performance compared to the existing methods.

Keywords: face detection, skin-tone region, Pulse-Coupled Neural Network, linking parameter

1. INTRODUCTION

Face detection, a technique which searches the location of face from input images, is researched actively with face recognition. Especially face detection which uses the color information has a merit that it is independent on size, angle and obstruction of a face. But the use of color information encounters some problems arising from skin-tone color in the background, strong illuminate effect, intensity variation within faces, and presence of random noise and so on. To remove these problems, face detection algorithms in the color image which is independent to shape and pose of face are suggested. But the detection performance is not that good if skin-tone color model is inappropriate or the color of background is similar to skin-tone color.

In this paper, we suggested a preprocessing to improve the detection performance. We attempted to transform the color image into gray level image so that the pixel with a mean value of skin-tone colors has the highest value (255) and the other pixels have values between 0 and 255 according to a normal distribution. Then, we segment the skin-tone region using Pulse-Coupled Neural Network (PCNN), which is proposed by Eckhorn and Johnson. As a result, we can segment the face region effectively based on pixel value similarity and spatial proximity. The preprocessing also leads to an easy decision of the linking parameter of the PCNN. For robust face detection performance, we add a postprocessing.

The outline of the paper is as follows. Section 2 gives a brief overview of PCNN. In section 3, we describe the face detection using PCNN. The result of experiment is presented in section 4. Finally, conclusions and recommendations for future research are given in section 5.

2. PULSE-COUPLED NEURAL NETWORK

Eckhorn's neuron model has been proposed to explain the synchronous activity among neural assemblies in the cat cortex induced by feature dependent visual activity. Synchronization means that the neurons with similar inputs pulse together. Therefore, if an image is applied as input data, the network will group objects in the image based on pixel value similarity and spatial proximity. But this model has certain properties that diminish the neuron's utility in image processing applications [4]. PCNN is the modified Eckhorn's neuron model which is more suitable for image processing applications, and it has the image processing abilities such as segmentation and edge detection [3].

2.1 standard PCNN

Pulse coupled neuron (PCN) model which constructs PCNN is shown in Figure 1 below. When PCNN is applied to image processing, one neuron is correspond to one pixel. The four parts that form the basis of the neuron are: feeding receptive field, linking receptive field, internal state, and pulse generator.

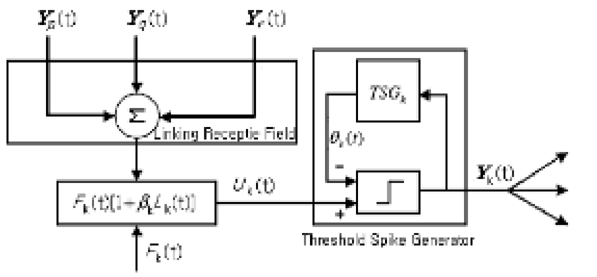

figure 1. pulse coupled neuron

1) Feeding receptive field : The feeding receptive field of the PCN N_k receives the feeding inputs. We use the pixel value of the image as an feeding input of PCNN.

2) Linking receptive field : The linking receptive field of the PCN N_k receives the linking inputs from the neighbor neurons. If one of the neighbor neuron N_l pulses, the output of N_l becomes the linking input of N_k.

$$L_k(t) = e^{\alpha F} L_k(t) + V_L \sum W_{kl} Y_l(t) \qquad (1)$$

where α is the time constant, V_L is normalized constant and W is weight matrix. Linking input is constructed by weighted sum of surrounding neuron's output and its decaying condition.

3) Internal state : The linking input modulates the feeding input in a nonlinear fashion to yield the internal state of the PCN N_k is

$$U_k(t) = F_k(t)(1 + \beta_k L_k(t)) \qquad (2)$$

where β_k is the linking parameter of N_k. It takes charge of an important role that controls the linking strength which influenced to the internal state of neuron.

4) Pulse generator : Pulse generator consists of a output pulse generator and a threshold spike generator. Output pulse generator determines output as follows.

$$Y_k(t) = \begin{cases} 1 & \text{if } U_k(t) > \theta_k(t) \\ 0 & \text{otherwise} \end{cases} \qquad (3)$$

where $\theta_k(t)$ is a threshold signal. When N_k pulses, $\theta_k(t)$ is set to initial value V_k which is greater than any possible value of U_k. If N_k pulses at time t_1 and t_2, the operation of threshold spike generator is described as follows.

$$Y_k(t) = \begin{cases} V_k & t = t_1 \\ V_k \exp(-(t - t_1)/\tau_k) & t_1 < t < t_2 \\ V_k & t = t_2 \end{cases} \qquad (4)$$

where τ_k is the decay time constant of threshold spike generator.

To add to this, pulse coupled neuron, which we described above, has the following characteristics.

- At t=0, all neurons pulse together.
- The threshold signal generators of all neurons are identical each other. In other words, All neurons have the same amplication factor and the decay time constant. $(V_k = V_l = V_\theta,\ \tau_k = \tau_l = \tau_\theta)$
- The value of linking parameter is independent of the neuron. $(\beta_k = \beta_l = \beta)$
- Each neuron receives linking inputs from the eight neighbors.

The characteristics of the pulse coupled neurons reflect the synchronous activities. By means of linking, feeding input is modulated when the internal state is calculated. And this modulation changes the internal state from F_k to $F_k(1 + \beta L_k)$. If the neuron has $\theta_k(t)$ value which lies between $[F_k,\ F_k(1 + \beta L_k)]$, then they fire together. Therefore, the neurons which have the similar inputs can be segmented together based on pulse period at this step. Note that whether neurons fire synchronously or not mainly depends on the values of linking parameter. In this paper, we suggested the effective method to determine the linking parameter of PCNN which is applied to the face detection.

2.2 Fast Linking Algorithm

If PCNN is applied to an image under various conditions of illuminations, we cannot get a face region effectively because a brightness of the image is distorted by illuminations [3]. To segment a face region robustly, we apply fast linking algorithm to standard PCNN. The principle of fast linking algorithm is to take advantage of the absent of reflection and other second order effects in the early iterations when waves propagate from intensity rich areas to those with lower intensity (cf. [Lindbald and Kinser, 1998]). Fast linking algorithm allows the generated linking wave to propagate all the way to the border of the image. Thus, it allows the linking wave to propagate faster than the feeding wave. In the standard PCNN, the linking wave was restricted to the size of linking and feeding field. The first neuron or group of neurons that fire will be the start point of a propagating linking wave.

The fast linking algorithm is achieved in the following way by using four neurons to explain it more clearly.

① Initialize the PCNN
② For the skin-tone region which is obtained by an experiment, if the internal state of neuron 1 satisfies $U_1 = F_1(1 + \beta L_1) > \theta_k$ when threshold signal reaches V_{t1}, neuron 1 fires and transports the output to the neighbor neurons, neuron 2 and neuron 3, as the linking input.
③ For the neuron 2 which receives the linking input at process ②, if the internal state of neuron 2 satisfies $U_2 = F_2(1 + \beta L_2) > \theta_k$, neuron 2 fires and transports the output to the neighbor neuron, neuron 4, as the linking input.
④ For the neuron 3 which receives the linking input at process ②, if the internal state of neuron 3 satisfies $U_3 = F_3(1 + \beta L_3) < \theta_k$, neuron 3 doesn't fire and reduces the value of θ_k. $(\theta_k = V_{t2})$
⑤ After processes ②¯④, For the neuron 4 which does not fire yet, the internal state of neuron 4 gets updated.
⑥ For the neuron 3 which does not fire yet, if $U_3 > \theta_k (= V_{t2})$, neuron 3 fires, otherwise neuron 3 does not fire and reduces the value of θ_k. $(\theta_k = V_{t2})$
⑦ Until no neurons are fired, repeat the processes ⑤¯⑥. And determine all firing neurons to one group.
⑧ Reduces the value of θ_k $(\theta_k = V_{t3},\ V_{t3} < V_{t2} < V_{t1})$, and repeat the processes ②¯⑦.

Fast linking algorithm presented above, neurons which fire at $\theta_k = V_{t1}$ and $\theta_k = V_{t2}$ form a group. And in order to reduce the processing time, we make fast linking algorithm to be used in the skin-tone region which is calculated in advance. While standard PCNN transmits linking wave at the next step within only receptive field, PCNN with fast linking algorithm transmits linking wave at the same step beyond the range of receptive field. Therefore, we expect that PCNN with fast linking algorithm can segment skin-tone region under various conditions of illuminations, because PCNN with fast linking algorithm can segment the neurons of which input is similar.

3. Face Detection Using PCNN

PCNN presented in the previous section has some problems to apply face detection. In this chapter, we present the problems and suggest the method to solve these problems.

3.1 Decision of linking parameter

The general approach to face detection using PCNN is to adjust the parameters of the network so that the neurons corresponding to the pixels of a skin-tone region pulse together and the neurons corresponding to the pixels of background regions do not pulse together. Therefore, the goal is to choose the network parameters such that the segmented region exactly corresponds to a face in the image. If such parameters exist, face region is easily segmented. However, most of input images include various objects that intensity distributions can be much different or overlapping.

Suppose input image contains two objects whose intensity distributions are as follows.

- Intensity distribution of object A : 200 ~ 250
- Intensity distribution of object B : 20 ~ 70

To segment each objects, β should satisfy to the following conditions.

$$X_{A,\min}(1+\beta L_A) \geq X_{A,\max}$$
$$\beta \geq (\frac{X_{A,\max}}{X_{A,\min}} - 1)/L_A = 0.25/L_A \quad (5)$$

$$X_{B,\min}(1+\beta L_B) \geq X_{B,\max}$$
$$\beta \geq (\frac{X_{B,\max}}{X_{B,\min}} - 1)/L_B = 0.25/L_B \quad (6)$$

If $L_A = L_B$, the minimum value of linking parameter of B is ten times greater than that of A. Therefore, if the minimum value of linking parameter is used to segment object A, much larger region than desired region is segmented. In this way the appropriate linking parameter is changed according to the range of pixels of each object, thus, we can not use the same linking parameter to segment the different object.

In this study, instead of getting an appropriate linking parameter in order to extract the skin-tone region, we transform the range of the skin-tone region into the range of region which is segmented effectively by the predetermined linking parameter. And if the PCNN did not produce satisfactory results, we extracted the face by postprocessing.

3.2 Improvement for face detection

In this study, we use the U, V images as inputs of the PCNN. The following equation shows how to convert from RGB image to YUV image.

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (7)$$

In order to segment the face robustly by the predetermined linking parameter, we need to know the color distribution of skin-tone region. First, we obtained the mean and variance of the skin-tone colors by experiments. Then, we perform a preprocess that the pixels with mean values of skin-tone colors have the highest level values (255) and the other pixels have values between 0 and 255 according to normal distribution with a variance. If we predetermine the value of linking parameter which is appropriate to segment the transformed skin-tone region, we can get a face region effectively. In addition, the objects which do not correspond to the face form a small group by necessity, because the value of linking parameter is inappropriate to segment those objects. Therefore, we can easily eliminate the objects which are not face by postprocessing.

We use 52 images which include facial regions only in order to check up the distribution of facial regions in UV space. Figure 2 shows the U,V distributions of the facial images.

fig 2.1 facial image

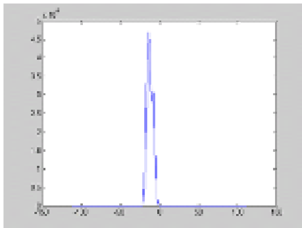

fig 2.2 U distribution

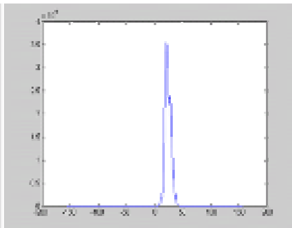

fig 2.3 V distribution

In this experiment, the distributions of the skin-tone regions are a little different because of various illumination conditions. Therefore, it requires that the skin-tone region has a wide range of pixel values. We transform the pixels with mean values of skin-tone colors have the highest level values (255) and the other pixels have values between 0 and 255 according to normal distribution.

$$F_u = 255 * \exp \frac{(X - U_{mean})^2}{U_{var}} \quad (8)$$

$$F_u = 255 * \exp \frac{(X - V_{mean})^2}{V_{var}} \quad (9)$$

where $U_{mean} = -14.2$, $U_{var} = 250$, $V_{mean} = 34.0$ and $V_{var} = 450$. Figure 3 represents the result of image which transformed to the desire region.

figure 3.1 input image

figure 3.2 U image

figure 3.3 V image

Equation (8), (9) represent that the distributions of U and V are different. Therefore, we use the different linking parameters for U and V images which can extract the skin-tone region effectively. Because the condition which linking parameter should be satisfied is $X_{\min}(1+\beta L) \geq X_{\max}$ for the maximum and minimum values of skin-tone region $X_{\max}$, $X_{\min}$, we can get minimum value of linking parameter. The proposed preprocessing makes the minimum value of linking parameter smaller because the difference between $X_{\max}$ and $X_{\min}$ become small. When we elevate the linking parameter from small value, we can find optimal linking parameter easier.

3.3 Pulse Inhibition Algorithm

In general, if there is a noise in the image, a certain pixel value can be much higher than other neighbor neurons. Especially, for the facial image, the luminosity of the nose or the forehead can be much higher than the other regions. In this case, the performance becomes deteriorate and we cannot help adjusting the linking parameter. Therefore, we modify the pulse coupled neuron by adding the pulse inhibition algorithm [4].

Suppose that the pixel value of a certain neuron is X, and it is going to be fired. If equation (10) is satisfied, the internal state of that neuron is modified.

$$X - \frac{1}{N}\sum X_{neighbor} > Threshold \qquad (10)$$

The modified internal state describes as follows.

$$U_k(t) = F_k(1+\beta L_k(t))(1-\gamma H_k) \qquad (11)$$

where, H_k is determined by the differences from the neighbor neuron's pixel values and γ is positive constant.

4. Experiments

In order to demonstrate the performance of the proposed method, we carry out the experiment. The size of images in this experiment is 400×300. We use the U and V image as the input of PCNN. After we perform a PCNN for each U and V images, then, we determine the final facial region as a common facial region of two cases.

We use the kinds of input images as follows.

- Type 1 : Image which contains skin-tone color in the background
- Type 2 : Image which does not contain skin-tone color in the background
- Type 3 : Image which contains various size of face
- Type 4 : Image which contains various angle of face
- Type 5 : Image which contains the distorted skin-tone color by illumination.

As described in table 1, the performance of the face detection using proposed model is improved considerably in compare with conventional methods. In this case, the conventional PCNN model also uses the fast linking algorithms in order to segment facial region robustly under various illumination conditions.

Type	Num of success		
	skin-tone color model[5]	PCNN with fast linking	Proposed PCNN
1 (20)	12	7	16
2 (20)	20	14	20
3 (20)	18	12	17
4 (20)	20	19	20
5 (20)	8	15	14
Total (100)	78	67	87

table 1. results of face detection

For the result of type 5, the performances of face detection with PCNN are much better than those of the case which uses the skin-tone color model. This result confirms that face detection with PCNN is more effective than conventional method even if the brightness of image is changed by illumination. In addition, the result of type 2 image shows that the performance of the proposed PCNN is better than those of the conventional PCNN and the skin-tone color model.

Because the conventional PCNN method extracts the facial region regardless of the linking parameter, we may not get the satisfactory performance of face detection. For the many values of the linking parameters, we make an experiment with the conventional PCNN for the type 1 image.

figure 4.1 input image

figure 4.2 beta = 0.01

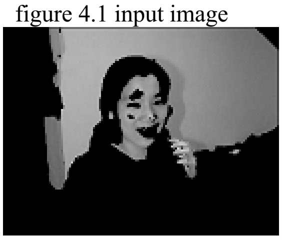

figure 4.3 beta = 0.1

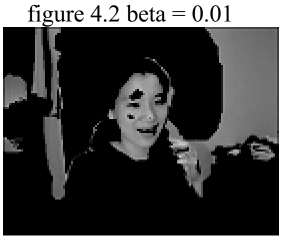

figure 4.4 beta = 0.055

As we can see in the above figure 4, we should find the appropriate value of linking parameter by trial and error. And it is difficult to find the appropriate one. When we use the proposed method, we can get the following result.

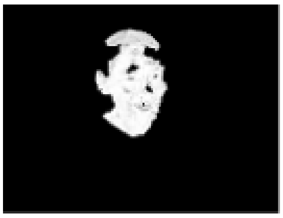

figure 5.1 result of U image

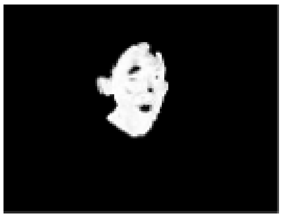

figure 5.2 result of V image

figure 5.3 final result

When intensity ranges of the complex background objects can be overlapped with those of the skin-tone region. And although the determined linking parameter is suitable for segmenting the skin-tone region, the other regions can be segmented together. However, the region which does not correspond to skin-tone region is a small size group as a necessity because of the inappropriate linking parameter. In these cases, we can get rid of these regions by the binary median filter.

figure 6.1 input image

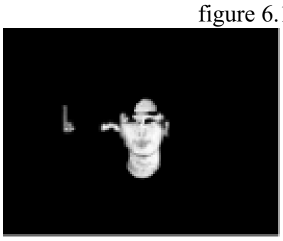

figure 6.2 before filtering

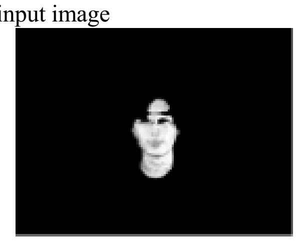

figure 6.2 after filtering

5. CONCLUSIONS

In this paper, we suggested the method that can effectively determines the linking parameter of PCNN. We perform a preprocessing that the pixel with a mean value of skin-tone colors has the highest level value (255) and the other pixels have values between 0 and 255 according to normal distribution with a variance. This preprocessing leads to an easy decision of the linking parameter of the PCNN and improves the performance of the face detection. That is, we predetermined the linking parameter appropriately which is suitable for segmenting the facial region. And we applied the fast linking algorithm and the pulse inhibition algorithm to improve the performance of the face detection.

However, in order to predetermine the linking coefficient, it still requires trial and error. And if the image contains the strong illumination, we can not get a desired region. If we use not only the color information but also the feature information of face, the performance of face detection can be improved.

ACKNOWLEDGMENTS

This research is supported by Korea Ministry of Commerce, Industry and Energy.

REFERENCES

[1] G. Yang and T. S. Huang, , “Human face detection in complex background,” *Pattern Recognition*, Vol. 27, No. 1, pp. 53-55, 1994.

[2] A. Lanitis, C. J. Taylor, and T. F. Cootes, “An automatic face identification system using flexible appearance models,” *Image and Vision Computing*, Vol. 13, No.5, pp. 393-401, 1995.

[3] Hitoshi Yamada, Yuuki Ogawa, Kosei Ishimura, Mitsuo Wada, “Face Detection Using Pulse-Coupled Neural Network,” *SICE Annual Conference in Fukui*, August 4-6, 2003.

[4] G. Kuntimad and H. S. Ranganath, “Perfect Image Segmentation Using Pulse-Coupled Neural Networks,” *IEEE Transaction on Neural Networks*, Vol. 10, No. 3, May 1999, pp. 591-598.

[5] Soo-Chang Pei, Ching-Long Tseng, “Robust face detection for different chromatic illuminations,”, *Image Processings 2002, International Conference on,* Vol. 2, 22-25 Sept. 2002.