

Pulse-Coupled Neural Network를 이용한 얼굴추출 알고리즘

Face Detection Algorithm Using Pulse-Coupled Neural Network

임 영 완*, 나 진 희*, 최 진 영**

(Young Wan Lim, Jin Hee Na, Jin Young Choi)

Abstract - In this work, we suggested the method which improves the efficiency of the face detection algorithm using Pulse-Coupled Neural Network. Face detection algorithm which uses the color information 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, intensity variation within faces, and presence of random noise, and so on. Depending on these conditions, we obtained the mean and variance of skin-tone colors by experiments. Then we introduce a preprocess that the pixel with a mean value of skin-tone colors has highest level value(255) and the other pixels in the skin-tone region have values between 0 and 255 according to a normal distribution with a variance. This preprocess leads to an easy decision of the linking parameters.

Key Words : face detection, skin-tone region, Pulse-Coupled Neural Network, linking coefficient

1. 서 론

얼굴 추출(face detection)은 영상으로부터 얼굴의 위치를 자동으로 찾아내는 기술로서 얼굴 인식과 더불어 활발히 연구되고 있는 분야이다. 그러나, 동일한 인물이라도 얼굴의 기울어짐과 크기, 다른 물체에 의한 가려짐 등에 의해 얼굴의 형태가 달라질 수 있기 때문에 색상정보를 이용한 얼굴 추출 알고리즘이 형태정보를 이용한 방법에 비해 장점을 갖는다. 그러나, 이 경우에는 조명이나 주위 배경의 색상에 영향을 받을 수 있다. 즉, 피부색 모델의 영역이 부적절하거나, 배경의 색상분포가 피부색과 유사할 경우 얼굴을 제대로 찾아내기 힘들다. 이러한 문제를 해결하기 위해 본 논문에서는 입력으로 들어오는 컬러영상에서 피부색 영역에 해당하는 픽셀들이 다른 영역에 해당하는 픽셀들에 비해 더 큰 값을 갖고 분포되도록 한 후, Pulse-Coupled Neural Network(PCNN)을 이용하여 피부색 영역을 분리하였다. 그 결과, 픽셀값의 유사성과 공간적 근접성을 기반으로 하면서, 기존의 PCNN 알고리즘에 비해 linking coefficient의 결정에 덜 민감하도록 추출 성능을 개선하였다. 그리고 후처리과정(post processing)을 통해 보다 안정적으로 얼굴을 추출할 수 있도록 하였다.

2. Pulse-Coupled Neural Network

Eckhorn의 뉴런 모델은 고양이 시신경을 모델로 하여 인

저자 소개

* : 서울대학교 전기컴퓨터공학부 석사과정

** 正會員 : 서울대학교 전기컴퓨터공학부 부교수

공신경망을 구축하고자 하는 연구의 결과로 탄생하였다[1]. 이 모델은 입력 데이터 값의 유사성과 공간적 근접성을 기반으로 뉴런들이 동기화되어 동시에 출력을 낼 수 있게 하는 기능을 가지고 있다. 그러나 Eckhorn의 모델은 디지털 이미지 처리에 있어서 신경망의 효용성을 저하시키는 특징을 가지고 있기 때문에[2], 이를 디지털 영상처리에 보다 적합한 형태로 개선한 모델이 PCNN이다.

2.1 구조와 원리

PCNN을 구성하는 PCN(Pulse coupled neuron)은 그림 1과 같은 구조를 가지며 총 4개의 파트로 구분할 수 있다.

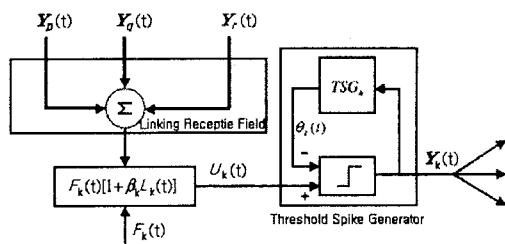


그림 1. PCNN의 구조

(1) **Feeding receptive field** : PCN N_k 는 외부 입력값

$F_k(t)$ 를 feeding input으로 받는다. 정지영상이 PCNN의 feeding input으로 들어올 경우, 각 뉴런의 입력은 입력영상의 해당 픽셀값이 된다.

(2) **Linking receptive field** : 각각의 뉴런은 이 linking receptive field를 통해 공간적으로 인접한 뉴런들의 출력값을 linking input으로 받는다. 즉, 이웃한 뉴런 N_l 이 펄스를 낼 경우, N_l 의 출력은 N_k 의 linking input으로 들어가게 된다.

$$L_k(t) = e^{-\alpha t} L_k(0) + V_L \sum W_{ki} Y_i(t) \quad (1)$$

여기서 α 는 시정수(time constant), V_L 은 정규화된 상수값을 나타낸다.

(3) Internal state :

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

여기서 β_k 는 뉴런 N_k 의 linking coefficient로서, 뉴런들을 적절하게 연결시키고 동기화하는데 중요한 역할을 한다.

(4) Pulse generator : threshold spike generator와 output spike generator로 나눈다. Output pulse generator에서는 식(3)을 통해 출력값을 결정한다.

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

여기서, $\Theta_k(t)$ 는 threshold signal로서 해당 뉴런 N_k 가 $Y_k=1$ 로 firing하면, 초기값 V_k 로 초기화된다. 만약, 시간 t_1 과 t_2 에서 뉴런 N_k 가 firing 한다면, $\Theta_k(t)$ 값은 다음과 같이 결정된다.

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

여기서, τ_k 는 감쇠 시정수(decay time constant)이다.

그 밖에도 위의 뉴런은 다음과 같은 특징을 갖는다.

- 하나의 뉴런은 하나의 픽셀에 해당한다.
- 모든 뉴런은 동일한 threshold spike generator를 사용한다. ($V_k = V_i = V_\theta$, $\tau_k = \tau_i = \tau_\theta$)
- 각각의 뉴런은 인접한 8개 뉴런으로부터 linking input을 받는다.

본 논문에서는 얼굴 영역을 추출하는데 있어 주로 색상 정보를 이용하기 때문에 U,V 값을 PCNN의 입력으로 사용하였다. 그리고, U,V 이미지 상에서 각각 얼굴 영역을 추출한 후, 공통영역을 최종 얼굴 영역으로 선택하였다.

2.2 PCNN의 개선

다양한 조명효과에 강인하고, 뉴런들의 동기화가 빠르게 일어나도록 하기 위해 앞에서 설명한 PCNN의 기본구조에 fast linking 알고리즘을 사용하였다[3]. fast linking은 threshold spike generator의 threshold 값의 변화를 기준으로 특정 threshold 값에서 출력이 1이 되는 뉴런들이 존재할 경우 그 뉴런들로부터 출력값이 1이 되는 새로운 뉴런이 없을 때까지 linking wave를 전달하는 방법이다. 구체적인 내용은 지면 관계상 생략한다.

노이즈에 의한 영향을 최소화하기 위해서 pulse 억제 알고리즘을 사용하였다[4]. 이미지 내부에 노이즈가 생길 경우, 특정 픽셀이 주위의 픽셀에 비해 큰 값을 갖게 되는 경우가 있다. 같은 얼굴 영역 내에서 소수의 뉴런이 주위 다른 뉴런에 비해 훨씬 큰 값을 가지게 되면 결국 얼굴 영역에 해당하는 픽셀값의 범위가 넓어지게 되고, 그 결과 β 값의 수정이 불가피해지며 성능도 저하된다. 이러한 문제를 해결하기 위해, 해당 뉴런의 픽셀값이 주위 뉴런들의 픽셀값에 비해 일정값 이상 를 경우, 그 뉴런의 내부 상태값을 조절하였다.

$$U_k(t) = F_k(1 + \beta L_k(t))(1 - \gamma H_k) \quad (5)$$

이 때, H_k 값은 주위 픽셀값과의 차이에 비례하게 결정해 주며, γ 값은 양의 실수값이다.

3. 성능향상을 위한 전처리 과정

PCNN을 이용해 얼굴을 추출하고자 할 경우, linking parameter인 β 값의 결정이 추출 성능에 많은 영향을 미친다. 그러나, 영상 안의 모든 추출 대상에 대해 적절한 β 값을 결정하는 것은 어려운 작업이다. 그 이유는 추출하고자 하는 영역의 픽셀값 범위에 따라 이를 추출하기 위한 적절한 β 값이 달라지기 때문이다. 본 논문에서는 얼굴을 효과적으로 추출하기 위해 피부색 영역을 미리 정한 β 값이 적절한 분리성능을 낼 수 있게 하는 영역으로 옮기는 작업을 먼저 수행하였다.

미리 정한 β 값이 적절한 분리성능을 낼 수 있도록 하기 위해서는 피부색 영역의 색상분포를 미리 알아야 한다. 그 후에, 피부색이 1바이트 채널에 들어가는 최고값인 255를 기준으로 분포되도록 한다면, β 값을 미리 결정해 두고도 얼굴을 배경과 분리할 수 있다. 즉, 큰 픽셀값을 갖는 피부색 영역에 알맞도록 선택된 β 는 상대적으로 작은 픽셀값을 갖는 배경이 피부색 영역과 함께 분리되지 않도록 한다. 또한, 이러한 경우 얼굴에 해당하지 않는 영역들은 이 영역들을 집단화(grouping) 하기에 부적절한 β 값을 갖게 되므로, 필연적으로 작은 영역들로 나타나게 된다. 그러므로 이 영역들은 후처리 과정을 통해 얼굴 후보영역에서 쉽게 제외시킬 수 있다.

우선, 실제 이미지에 피부색이 존재할 때, 그 분포가 어떤지 확인해 보기 위해, 피부색만 포함한 총 52장의 얼굴 이미지를 이용하여 U,V 값들의 평균값과 분산값을 구한 후, 이를 이용해 다음과 같은 가우시안 분포를 갖도록 하여 기준에 비해 피부색 영역이 좀 더 조밀한 분포를 갖도록 하였다.

$$F_u = 255 * \exp \frac{(X - U_{mean})^2}{U_{var}} \quad (6)$$

$$F_v = 255 * \exp \frac{(Y - V_{mean})^2}{V_{var}} \quad (7)$$

단, $U_{mean} = -14.2$, $U_{var} = 250$, $V_{mean} = 34.0$, $V_{var} = 450$ 값을 갖는다. 식 (6),(7)에서 알 수 있듯이 실제 이미지에서 피부색 영역의 U,V 값들은 서로 다른 분포 특성을 갖는다. 따라서 U와 V 이미지에 대해 각기 다른 β 값을 사용하여야 한다. β 가 만족해야 할 조건은 피부색 영역에 해당하는 픽셀의 최대값과 최소값에 해당하는 X_{max} , X_{min} 에 대하여 $X_{min}(1 + \beta L) \geq X_{max}$ 의 식에서 구할 수 있다. 그러므로 β 의 최소값을 결정할 수 있다. 그런데, 피부색 영역의 픽셀값을 255를 중심으로 가우시안 분포를 갖도록 하였으므로, X_{max} , X_{min} 값의 차이가 줄어들어 피부색 영역을 분리할 수 있는 β 의 최소값이 더 작아진다. β 를 작은 값에서부터 증가 시켜 가면서 255를 기준으로 분포하는 피부색 영역을 추출하기에 적절한 β 값을 찾으면 된다.

4. 실험 결과

실험에서는 크기가 400×300 인 이미지 100장을 사용하였다. 이때 사용되는 얼굴 영상 데이터는 실제 환경 속에서 얻어진 것으로 피부색과 비슷한 색상을 가진 배경을 포함하는 이미지(유형1), 피부색과 확연히 구분되는 배경을 포함한 이미지(유형2), 다양한 얼굴 크기(유형3) 및 방향(유형4)을 가진 이미지, 조명에 의해 얼굴 색상에 변화가 있는 이미지(유형5) 등을 사용하였다. 위의 특성을 지닌 이미지를 가지고 실험한 결과, 표1에서와 같이 기존의 컬러 영상을 이용한 얼굴 추출 방법[5]이나 기존의 PCNN을 사용한 방법[3]보다 좋은 성능을 보였다.

이미지 유형	추출 성공 개수		
	피부색을 이용한 추출	기존 PCNN	제시한 PCNN
1 (20개)	12	7	16
2 (20개)	20	14	20
3 (20개)	18	12	17
4 (20개)	20	19	20
5 (20개)	8	15	14
TOTAL (100개)	78	67	87

표 1. 성능 비교 결과

위의 결과를 살펴보면, PCNN을 이용한 방법이 기존의 피부색을 이용한 방법에 비해 조명효과에 더 강인함을 확인할 수 있으며, 전처리 과정을 포함한 제시한 방법이 기존의 PCNN을 이용한 방법에 비해 배경과의 분리에 더 좋은 성능을 보임을 확인할 수 있다.



그림1. 입력이미지

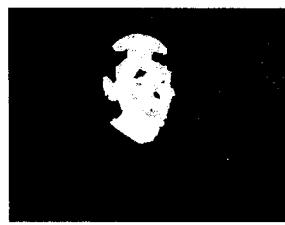


그림2. U 이미지



그림3. V 이미지



그림4. 최종 결과



그림5. 입력이미지



그림6. 기존의 피부색을
이용한 추출



그림7. 제시한 모델에
의한 추출

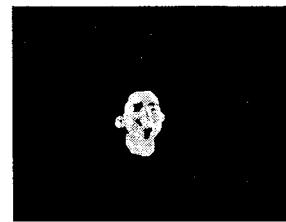


그림8. 이진 중간필터
적용

위의 그림1~그림8은 피부색과 비슷한 배경이 포함된 경우와 조명의 효과가 있을 경우에 실험해 본 결과이다. 3장에서 설명했듯이 설정한 β 값이 얼굴에 해당하지 않은 부분이 얼굴과 함께 집단화 되게 할 수도 있다. 이러한 부분들은 이진 중간 필터(binary median filter)를 사용하여 집단화된 영역에서 배제되도록 하였다.

5. 결 론

본 논문에서는 PCNN을 이용한 얼굴 추출 시, 전처리 과정을 통해 피부색에 해당하는 픽셀값의 분포가 255를 중심으로 가우시안 분포를 갖게 한 후, fast linking 알고리즘을 적용하였다. 그 결과, 기존의 알고리즘에 비해 β 값에 덜 민감하고 환경변화에 더 강인 얼굴추출 방법을 제시하였다. 그러나 여전히 β 값을 결정하는 방법이나, 강한 조명으로 인해 색상 왜곡이 생길 경우에 형태 정보를 이용하는 방법 등에는 여러 가지 개선의 여지가 남아있다.

참 고 문 헌

- [1] G.Yang and T.S.Huang, "Human face detection in complex background", Pattern Recognition, vol. 27, no. 1, pp.53-53, 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. Kuntmad 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 Processing. 2002. Proceedings. 2002 International Conference on , Volume: 2, 22-25 Sept. 2002.