

# 컬러영상에서 Pulse-Coupled Neural Network를 이용한 얼굴 추출 알고리즘

## Face Detection Algorithm Using Pulse-Coupled Neural Network in Color Images

임영완, 나진희, 최진영  
서울대학교 전기컴퓨터공학부

Young Wan Lim, Jin Hee Na, Jin Young Choi

Dept. of Electrical Engineering and Computer Sciences/Electrical Engineering  
Seoul National University  
E-mail : jhna@neuro.snu.ac.kr

### 요 약

본 논문에서는 컬러영상에서 Pulse-Coupled Neural Network를 이용한 얼굴 추출 알고리즘의 성능을 향상시키는 방법에 대하여 논의하였다. 색상정보를 이용한 얼굴 추출 알고리즘은 얼굴의 기울어진 정도나 크기 등에 영향을 받지 않으므로, 형태정보를 이용한 얼굴 추출 알고리즘에 비해 비교우위를 가진다. 그러나 조명의 변화가 심하거나, 피부색과 유사한 배경이 포함되어 있을 경우 적절한 성능을 내기 어렵다. 이러한 문제점들을 해결하기 위해 본 연구에서는 넓은 피부색 영역을 추출하고, Pulse-Coupled Neural Network를 통해 공간적으로 근접한 피부색 동종영역을 분리해 내는 방법을 사용하였다. 그리고 피부색 영역에 해당하는 픽셀들이 다른 영역들에 비해 큰 값을 갖도록 하여, Pulse-Coupled Neural Network의 linking coefficient를 보다 쉽게 결정하도록 하였다.

### 1. 서론

얼굴 추출(face detection)은 정지영상이나 동영상으로부터 얼굴의 위치를 자동으로 찾아내는 기술로서 얼굴 인식과 더불어 활발히 연구되고 있는 분야이다. 그러나 동일한 인물이라도 얼굴의 기울어짐과 크기, 다른 물체에 의한 가려짐 등에 의해 생김새가 달라질 수 있으며, 조명이나 주위 배경의 색상에도 영향을 받을 수 있다. 그러므로 얼굴 추출은 쉽지 않은 문제이다. 이러한 외적 요인들에 의한 효과를 제거하기 위해, 얼굴 형태와 자세에 대한 사전지식을 필요로 하는 흑백영상에서의 얼굴 추출 방식을 배제하고, 조명 효과나 배경에 강인함을 갖는 컬러영상에서의 얼굴 추출 알고리즘에 대한 연구가 많이 이루어졌다. 그러나 피부색 모델의 영역이 부적절하거나, 배경의 색상분포가 피부색과 유사할 경우 얼굴을

제대로 찾아내기 힘들다. 이러한 문제를 해결하기 위해 본 논문에서는 입력으로 들어오는 컬러영상 중 피부색 영역에 해당하는 픽셀들이 다른 영역에 해당하는 픽셀들에 비해 더 큰 값을 갖고 조밀하게 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 연결계수(linking coefficient)에 덜 민감하도록 추출성능을 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 인공신경망을 구축하고자 하는 연구의

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

2.1 PCNN의 구조와 원리

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

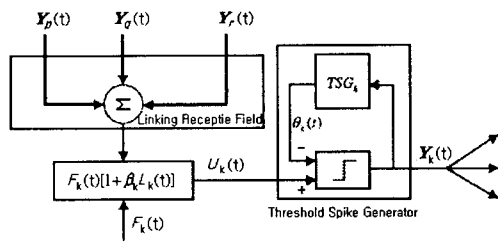


그림 1. PCNN의 구조

(1) Feeding receptive field : PCN \$N\_k\$는 외부

입력값 \$F\_k(t)\$를 공급입력(feeding input)으로 받는다. 정지영상이 PCNN의 외부입력으로 들어올 경우, 각 뉴런의 입력은 입력영상의 해당 픽셀값이 된다.

(2) Linking receptive field : 각각의 뉴런은 이 linking receptive field를 통해 공간적으로 인접한 뉴런들의 출력값을 연결입력(linking input)으로 받는다. 즉, 이웃한 뉴런 \$N\_l\$이 펄스를 낼 경우, \$N\_l\$의 출력은 \$N\_k\$의 연결입력(linking input)으로 들어가게 된다.

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

여기서 \$\alpha\$는 시정수(time constant), \$V\_L\$은 정규화된 상수값을 나타낸다.

(3) Internal state :

$$U_k(t) = F_k(t)(1 + \beta_k L_k(t)) \tag{2}$$

여기서 \$\beta\_k\$는 뉴런 \$N\_k\$의 연결계수로서, 뉴런들을 적절하게 연결시키고 동기화 하는데 중요한 역할을 한다.

(4) Pulse generator : Pulse generator는 output pulse generator와 threshold spike generator로 나뉜다. output pulse generator에서는 식(2)을 통해 출력값을 결정한다.

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

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.

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

여기서, \$\tau\_k\$는 감쇠 시정수(decay time constant)이다.

그 밖에도 위에서 제시한 뉴런은 다음과 같은 특징을 갖는다.

- 하나의 뉴런은 입력 영상의 한 픽셀에 해당한다. 즉, 뉴런의 외부입력 값 \$F\_k(t)\$는 픽셀 \$P\_k\$의 값인 \$X\_k\$과 같다.
- 연결계수 값은 뉴런과 관계없이 일정한 값을 갖는다.(즉, 모든 \$k\$와 \$l\$에 대해 \$\beta\_k = \beta\_l = \beta\$)
- 모든 뉴런은 동일한 threshold spike generator를 사용한다.(\$V\_k = V\_l = V\_\theta\$, \$\tau\_k = \tau\_l = \tau\_\theta\$)
- 각각의 뉴런은 인접한 8개 뉴런으로부터 연결입력을 받는다.

2.2 PCNN의 개선

위에서 제시한 기존의 PCNN 모델을 다양한 조명 아래서 사용하였을 경우, 만족할 만한 결과를 얻지 못하는 경우가 많다. 따라서 다양한 조명 효과에 강인하도록 기존의 PCNN 알고리즘에 fast linking 개념을 적용하였다[3]. fast linking은 threshold값의 변화를 기준으로 현재의 threshold 값에서 출력값이 1이 되는 뉴런들이 존재할 경우, 그 뉴런들로부터 시작하여 출력값이 1이 되는 뉴런이 없을 때까지 linking wave를 전달하는 방법이다. 네 종류의 뉴런 1,2,3,4를 예로 구체적인 알고리즘을 설명하면 다음과 같다.

- (1) PCNN을 초기화한다.
- (2) \$\theta\_k\$가 감소하다가 \$\theta\_k = V\_l\$ 일 경우, 실험적으로 구한 피부색 영역에 대해 뉴런 1이 \$U\_1 = F\_1(1 + \beta L\_1) > \theta\_k\$ 를 만족하면 이 뉴런의 출력값은 1이 되고 주위 뉴런 2,3에 연결입력을 전달한다.
- (3) (2)에 의해 연결입력을 받은 뉴런 2에 대해 \$U\_2 = F\_2(1 + \beta L\_2) > \theta\_k\$이면, 뉴런 2는 1을 출력으로 내고, 주위 뉴런 4에 연결입력을 전달한다.
- (4) (2)에 의해 연결입력을 받은 뉴런 3에 대해 \$U\_3 = F\_3(1 + \beta L\_3) < \theta\_k\$이면, 뉴런 3은 0을 출력으로 내고 \$\theta\_k\$값을 줄인다.(\$\theta\_k = V\_l\$)

(5) (2)~(4)과정 후에 실험적으로 구한 피부색 영역에 해당하는 픽셀들에 대해 아직까지 출력이 0인 뉴런 3,4 중에 뉴런 4의  $U_4$ 값을 갱신한다.

(6) 뉴런 3에 대해  $U_3 > \theta_k (= V_i)$ 이면, 1을 출력으로 내고, 아니면 0을 출력으로 낸다. 뉴런 4에 대해  $U_4 > \theta_k (= V_i)$ 이면, 1을 출력으로 내고, 아니면 0을 출력으로 낸 후,  $\theta$ 값을 줄인다. ( $\theta_k = V_i$ )

(7) 출력이 1인 새로운 뉴런이 없을 때까지 (5)~(6)을 반복한다. 그리고 출력이 1인 모든 뉴런들을 하나의 그룹으로 분리한다.

(8)  $\theta_k$ 값을 줄이고 ( $\theta_k = V_{i+1}, V_{i+1} < V_i < V_i$ ) (2)~(7)을 반복한다.

위의 알고리즘은  $\theta_k = V_i$ 를 갖는 뉴런들이  $\theta_k = V_i$ 를 갖는 뉴런들과 같은 그룹으로 분리될 수 있게 하였다. 그러므로 같은 얼굴에 해당하는 픽셀이 조명이나 다른 노이즈에 의해 주위의 픽셀과 다른 값을 갖더라도 얼굴 영역이 나누어지지 않고, 같은 그룹으로 분리될 가능성이 높아진다. 그러나, fast linking을 사용할 경우, 기존의 PCNN에 비해 연산량이 많아져 계산시간이 오래 걸리는 단점이 있다. 이를 해결하기 위해, 실험적으로 피부색 영역에 해당하는 픽셀들을 미리 구하여 탐색 범위를 줄인 후, 그 픽셀들에 대해서만 fast linking 알고리즘을 적용하도록 하였다. 자세한 내용은 다음 장에서 설명하기로 한다.

### 3. PCNN을 이용한 얼굴 추출

PCNN을 이용해 얼굴을 추출하고자 할 경우,  $\beta$  값을 적절하게 설정해야 하는데, 이것은 어려운 작업이다. 그 이유는 얼굴추출에 사용되는 입력 이미지에는 일반적으로 다양한 객체(object)들이 포함되어 있는데, 객체들 간의 intensity 분포 차이가 크게 나타날 수도 있고, 또한 intensity overlapping이 나타날 수도 있기 때문이다. 본 논문에서는 얼굴을 추출하기 위해 입력 이미지에 알맞은  $\beta$ 값을 구하는 대신, 피부색 영역을 미리 정한  $\beta$ 값이 적절한 분리성능을 낼 수 있게 하는 영역으로 옮기는 작업을 먼저 수행하였다. 그리고 배경이 얼굴과 비슷한 색상정보를 가지고 있어서, PCNN을 사용하여 효과적인 분리가 어려운 경우 후처리 과정(post processing)을 통해 얼굴을 추출하였다.

#### 3.1 PCNN을 위한 피부색 영역 추출

본 논문에서는 얼굴 영역을 추출하는데 있어 주

로 색상정보를 이용하기 때문에 U,V 값을 PCNN의 입력으로 사용하였다. RGB 이미지를 YUV 이미지로 변환하는 식은 다음과 같다.

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (5)$$

앞에서 언급했듯이 미리 정한  $\beta$ 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,  $\beta$ 값을 미리 결정해 두고도 얼굴을 배경과 분리할 수 있다. 즉, 큰 값을 갖는 피부색 영역에 알맞도록 선택된  $\beta$ 는 상대적으로 작은 값을 갖는 배경이 피부색 영역과 함께 분리되지 않도록 한다. 또한, 이러한 경우에 얼굴에 해당하지 않는 영역들은 이 영역들을 집산화(grouping)하기에 부적절한  $\beta$ 값을 갖게 되므로, 필연적으로 작은 영역들로 나타나게 된다. 이러한 경우 이 영역들은 후처리 과정을 통해 얼굴 후보영역에서 제외되어야 한다.

우선, 실제 이미지에 피부색이 존재할 때, 그 분포가 어떠한지 확인해 보기 위해, 피부색만 포함한 총 52장의 얼굴 이미지를 이용하여 U,V 값들의 분포를 살펴보았다. 그 결과, 조명 등 외부 효과에 의해 이미지마다 피부색 분포가 조금씩 다르게 나타났다. 그러므로 피부색 영역에 해당하는 픽셀값이 가장 크도록 하되, 그 분포를 보다 넓게 잡아야 할 필요가 있다. 실험으로 얻어진 U,V의 평균값과 분산값을 이용해 피부색 영역이 다음과 같은 가우시안 분포를 갖도록 하여 기존의 이미지에 비해 피부색 영역이 좀 더 조밀한 분포를 갖도록 하였다.

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

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

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  $\beta$ 값을 사용하여야 하는데, 이  $\beta$ 값은 255를 기준으로 분포하는 피부색 영역에 대해 최적의 분리성능을 낼 수 있도록 하였다.  $\beta$ 가 만족해야 할 조건은 피부색 영역에 해당하는 픽셀의 최대값, 최소값에 해당하는  $X_{max}$ ,  $X_{min}$ 에 대하여  $X_{min}(1 + \beta L) \geq X_{max}$ 이다. 그러므로  $\beta$ 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  $\beta$ 의 최소값이 더 작아진다. 그러므로  $\beta$ 를 작은 값에서부터 증가시킬 때, 더 쉽게  $\beta$  값을 찾을 수 있다.

### 3.3 Pulse 억제 알고리즘

일반적으로, 이미지 내부에 노이즈가 생길 경우, 주위 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  $\beta$ 값의 수정이 불가피해지며 성능도 저하된다. 이러한 문제를 해결하기 위해, 해당 뉴런의 픽셀값이 주위 뉴런들의 픽셀값에 비해 일정값 이상 클 경우, 그 뉴런의 내부 상태값을 조절하였다[4].

$$U_k(t) = F_k(1 + \beta L_k(t))(1 - \gamma H_k) \quad (8)$$

이 때,  $H_k$ 값은 주위 픽셀값과의 차이에 비례하게 결정해 주며,  $\gamma$ 값은 양의 실수값이다.

## 4. 실험 및 결과

본 논문에서 제안한 알고리즘을 실제 이미지에 적용하여 얼굴을 추출해 보았다. 실험에서는 크기가  $400 \times 300$ 인 이미지 100장을 사용하였다. 이 때 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. 그리고 적절한 얼굴 추출의 성공여부는 다음과 같은 조건을 만족시킬 경우와 아닌 경우로 구분하였다.

- 얼굴영역 내부의 눈,코,입 등이 제대로 나타날 경우(원래 이미지에 해당 특징들이 포함되어 있을 경우)
- 얼굴영역 내부의 픽셀들 중 약 90%이상이 얼굴로 추출될 경우
- 얼굴이 피부색과 유사한 배경과 인접해 있을 경우, 얼굴과 완전히 분리되지 않은 배경이 얼굴영역 전체의 pixel수의 10%를 넘지 않을 경우
- 얼굴과 분리된 group이 최종적으로 얼굴영역으로 추출되지 않을 경우

위와 같은 특성을 지닌 이미지와 평가 기준을 가지고 실험한 결과, 표1에서와 같이 기존의 컬러 영상을 이용한 얼굴 추출 방법[5]이나 기존의 PCNN을 사용한 방법[3]보다 좋은 성능을 보였다.

이때, 기존의 PCNN을 이용한 알고리즘에서는 본 논문에서 사용한 연결계수와 동일한  $\beta = 0.3$ 을 사용하였다.

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

먼저, 위의 얼굴 데이터 유형 중 피부색과 비슷한 색상을 가진 배경을 포함하는 경우를 살펴보면, 그림 2.1과 같다. 이 경우, 얼굴과 배경이 모두 피부색 영역에 속해 있으므로, 기존의 색상 정보를 이용한 방법으로는 얼굴 추출이 힘들다. 그리고 PCNN을 위한 기존 연구의 경우, 연결계수에 대한 고려 없이 주로 펄스 억제 방법에 의존하여 얼굴 영역을 추출한다[3]. 기존의 PCNN을 사용하여 연결계수를 변화시켜 가면서 얼굴을 추출해 보면 다음과 같다.



그림 2.1 입력이미지



그림 2.4  $\beta = 0.01$  경우



그림 2.3  $\beta = 0.1$  일 경우

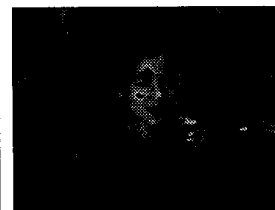


그림 2.4  $\beta = 0.055$  경우

위의 그림에서 알 수 있듯이 기존의 PCNN을 이용한 방법에서는  $\beta$ 를 미세하게 변화시키면서, 적절한 값을 찾아야 한다. 즉,  $\beta$  결정에 대한 적절한 고려가 이루어지지 않으면, 제대로 된 결과를 얻을 수가 없다. 그러나 본 논문에서 제시한

색상변환 방법을 사용하면 다음과 같은 결과를 얻을 수 있다.



그림 3.1 U image의 segment결과



그림 3.2 V image의 segment결과

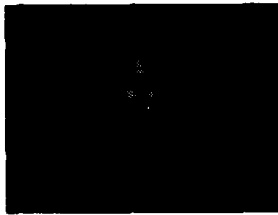


그림 3.3 최종 추출 결과

배경이 여러 객체들을 포함하여 복잡한 경우에는, 피부색 영역에 해당하는 영역들이 배경 속에 군데군데 나타날 수 있다. 또한 3장에서 설명했듯이  $\beta$ 값이 큰 픽셀값을 갖는 피부색 영역을 분리하기에 알맞도록 선택되었으나, 얼굴에 해당하지 않은 부분이 얼굴과 함께 집단화 될 수도 있다. 이러한 부분들은 이진 중간 필터(binary median filter)를 사용하여 집단화된 영역에서 배제되도록 하였다.



그림 4.1 입력 이미지



그림 4.2 이진 중간 필터 적용 전



그림 4.3 이진 중간 필터 적용 후

## 5. 결론 및 향후과제

본 논문에서는 PCNN을 이용한 얼굴 추출 시, 전처리 과정을 통해 피부색에 해당하는 픽셀값의 분포가 255를 중심으로 가우시안 분포를 갖게 하였다. 그리고 fast linking 알고리즘을 수

정하여, 기존의 알고리즘에 비해  $\beta$ 값에 덜 민감하게 하였으며, 환경변화에 더 강인한 얼굴추출 방법을 제시하였다. 그러나 여전히  $\beta$ 값을 결정하는 방법이나, 강한 조명으로 인해 색상 왜곡이 생길 경우에 형태 정보를 이용하는 방법 등에는 여러 가지 개선의 여지가 남아있다.

## 6. 참고문헌

- [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. 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 Processing. 2002. Proceedings. 2002 International Conference on , Volume: 2, 22-25 Sept. 2002.