

## CNN 모델을 이용한 얼굴 추출 및 보정 기법

조일국<sup>0</sup> 김호준

한동대학교 전산전자공학부

ilgook@gmail.com<sup>0</sup>, hjkim@handong.edu

### Face Detection and Region Refinement using a CNN Model.

Il Gook Cho<sup>0</sup>, Ho Joon Kim

School of Computer Science and Electronic Engineering, Handong Global University

#### 요 약

본 연구에서는 실내에서 입력 받은 영상의 조명과 크기 변화 등에 강인한 얼굴 검출 기법을 소개한다. 제안된 얼굴 검출 기법은 후보 영역 선정 과정과 얼굴패턴 검출 과정, 얼굴 영역 보정 과정으로 이루어진다. 후보 영역 선정 과정에서는 조명보정과 색상 필터, 움직임 필터를 이용하여 얼굴패턴의 후보 영역을 선정한다. 얼굴패턴 검출 과정에서는 CNN을 이용하여 특징을 추출하고, WFMM신경망을 이용하여 얼굴 패턴을 검증한다. 얼굴 영역 보정 과정은 형태학적 연산 등의 영상 처리를 이용하여 눈 영역과 입술 영역의 위치를 판별한 후 최종적인 얼굴 영역을 결정한다.

#### 1. 서 론

얼굴 영역 검출은 정지 영상이나 동영상으로부터 얼굴의 유무를 판단하고, 얼굴이 있다면 그 영역과 위치 등을 찾아내는 작업이다. 얼굴 검출은 얼굴 인식이나 인증, 표정 인식 등에 기반이 되는 기술로서 얼굴 영역의 검출 성능에 따라 후의 과정들의 성능이 많이 좌우된다. 이를 위한 많은 방법론과 연구 결과들이 국내외의 연구들을 통하여 발표되고 있는데, 컬러 이미지에서 피부색의 분포를 분석하는 방법[1], 얼굴의 기하학적 정보를 이용하는 방법[2], 다단계 필터링과 복합형 신경망을 사용하는 방법[3], 지역적 샘플링과 자기 구조화 지도를 상호 결합한 형태의 신경망을 사용한 방법[4] 등이 그 예이다.

얼굴패턴 검출은 특징 추출 과정과 패턴 분류 과정으로 구성된다. 특징 추출을 위하여 CNN(Convolutional Neural Network)모델을 사용하였는데 이는 생물학적인 시각신경계의 구조로부터 유추된 인공 신경망이다[4]. 첫 번째 계층에서 가보(Gabor) 변환을 사용하여 입력 영상에서 기본 특징을 추출하고, 연속되는 계층을 통하여 조합, 확장함으로써 최종적인 특징지도 생성하게 된다. 패턴 분류 과정에서 사용한 가중치를 가진 퍼지 최대 최소(Weighted Fuzzy Min-Max(WFMM))신경망은 일종의 뉴로퍼지 신경망으로 하이퍼 박스 기반의 패턴 분류 모델이다[5]. 이는 신경망의 빈도와 분포, 가중치를 고려하는 모델로서, 학습을 통하여 가중치가 조정되어 비정상적인 데이터에 의해 하이퍼 박스가 왜곡 되는 것을 개선하게 한다. 또한 특징과 클래스간의 연관도 요소를 고려하여 특징의 상대적 중요도를 평가 하면 선별적으로 입력특징을 사용할 수 있다.

후보 영역을 결정하기 위해 사용한 필터링 과정은 시스템의 실시간 응답특성을 고려한 전처리 단계로서 조명보정과 색상 필터, 움직임 필터를 적용하였다. 이 때 사용하는 색상 필터는 총 11가지의 색상 공간에서 추출해낸 피부색 값을 WFMM신경망에 학습 시킴으로써 피부색을 분류한다.

얼굴 영역 보정 과정은 최종적으로 선별된 얼굴 영역에서 눈의 위치와 입술의 위치를 판별하여 결정한다. 이는 얼굴 영역을 결정할 때 턱 밑의 목 부분까지 크게 결정되는 것을 방지하기 위한 방법으로 정확한 얼굴 영역을 검출 하기 위해 사용한다.

#### 2. 대상 시스템 개요

본 연구는 그림 1과 같은 구조의 얼굴 검출 시스템을 대상으로 한다. 그림에서 보인 바와 같이 시스템은 크게 전처리 단계와 얼굴 영역 검출 단계, 얼굴 영역 보정 단계의 3단계의 모듈로 이루어진다. 첫 단계인 필터링 과정에서는 영상에서의 움직임 요소, 색상 정보를 이용하여 후보 영역을 선별한다. 두 번째 단계인 검출 모듈은 특징 추출과 패턴 분류로 이루어진다. 첫 번째 계층에 가보 필터를 적용한 CNN을 사용하여 특징 지도를 생성하고, 이 결과를 WFMM신경망의 입력으로 사용하여 최종적으로 얼굴 패턴을 분류해낸다. 분류된 얼굴 패턴 영역은 실제 얼굴 영역에 비해 목 등을 포함하여 크게 나타나게 된다. 따라서 본 연구에서는 세 번째 단계로 얼굴 영역 보정 작업을 수행하게 되는데, 추출된 얼굴 영역 내의 눈과 입술을 찾아 내어 이에 따른 통계적 얼굴 크기를 이용하여 최종적인 얼굴 영역을 확정하게 된다.

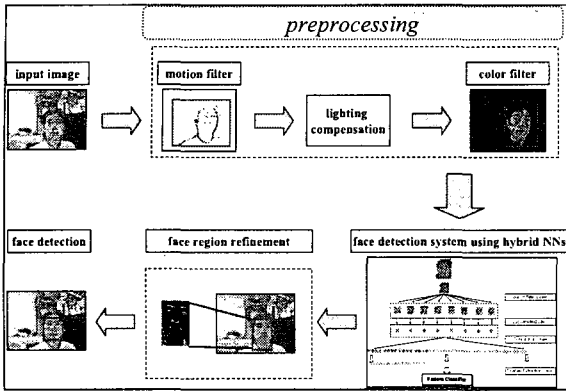


그림 1 얼굴 검출 시스템의 개요

3. 필터링

주어진 입력 영상에 대하여 모든 영역을 대상으로 하는 영역 추출 과정은 많은 시간이 소모되고, 복잡한 배경에서의 영역 추출의 성능을 저하시키며, 조명 변화에도 민감하게 반응한다. 이에 본 연구에서는 조명보정과 다단계 필터링을 통해 후보 영역을 선정하여 상기의 문제점들을 개선하고자 한다.

첫 번째 단계에서 사용한 움직임 검출에서는 본 시스템의 실시간 응답특성을 고려하여 단순한 방식으로 하였는데, 이는 연속 영상에서 각 픽셀의 색상의 변화를 구하고 이에 임계치를 적용함으로써 움직임 정보를 추출한다.

입력되는 영상에서 피부색은 조명에 매우 민감하게 반응한다. 따라서 본 연구에서는 조명 보정과정을 수행하는데 이를 위하여 "Reference White" [1]기법을 사용하였다. 이는 입력되는 전체 영상에서 luma값의 범위를 구한 후, 이 값의 상위 일정 비율의 픽셀을 특정 그룹으로 정의한다. 이에 속한 픽셀 수가 임계치보다 많으면 전체 영상에 대해 보정작업을 수행한다.

피부색 검출은 실제 환경에서 획득한 다량의 피부색 데이터로부터 총 11가지의 색상 공간에서 특징 값들을 얻어낸 후 WFMM신경망에 학습하여 수행한다. 이는 피부색 패턴 분류 기능뿐만 아니라, 학습된 신경망으로부터 특징과 하이퍼박스간의 상호 연관도 요소를 분석하여 피부색 검출을 위한 색상공간을 선별적으로 사용할 수 있게 한다.

4. 얼굴 검출 신경망

전처리 단계를 거쳐 결정된 후보 영역은 특징 지도 생성 및 패턴 분류 작업을 통하여 최종적인 얼굴 영역을 결정하게 된다. 특징 지도 생성을 위해 사용된 CNN은 계층적 다중 신경망이다. 첫 번째 계층에서는 38 X 42의 입력 영상으로부터 가보 변환을 적용하여 8방향의 28 X 32크기인 기본 특징들을 추출한다. 두 번째 계층에서는 2 X 2영역으로 표본 추출하여 14 X 16의 크기로 조정되고, 총 8개의 특징 영역 중 2개의 서로 다른 조합을 선택하여 총

44개의 특징 지도를 생성한다. 이는 패턴 분류 신경망의 입력으로 사용되게 된다.

제안된 시스템에서는 패턴 분류 신경망으로 WFMM신경망을 사용한다. 신경망의 활성화 특성은 다음과 같이 정의되는 하이퍼박스의 소속 함수로서 결정된다.

$$B_j = \{X, U_j, V_j, C_j, F_j, f(X, U_j, V_j, C_j, F_j)\} \quad (1)$$

$$\forall X \in I^n$$

식 (1)에서  $U_j$  및  $V_j$ 는 각각 하이퍼박스  $j$ 의 특징 범위에 대한 최대값 및 최소값 벡터를 의미하며,  $C_j$  및  $F_j$ 는 특징 집합의 평균값 및 특징의 유형을 의미한다. 이와 같이 정의된 하이퍼박스는 다음과 같은 식으로 그 활성화 특성이 정의된다.

$$b_j(A_h) = \frac{1}{\sum_{j=1}^n w_{ji}} \cdot \sum_{j=1}^n w_{ji} [\max(0, 1 - \max(0, \gamma_{ju} \min(1, a_{hi} - v_{ji}))) + \max(0, 1 - \max(0, \gamma_{jv} \min(1, u_{ji} - a_{hi}))) - 1.0] \quad (2)$$

식 (2)에서  $w_{ji}$ 는 특징  $i$ 와 하이퍼박스  $j$ 와의 가중치 요소를 의미한다. 또한 매개변수  $\gamma_{ju}$ 와  $\gamma_{jv}$ 는 소속 함수에서 특정범위의 가장자리의 기울기를 제어하는 상수이다.

본 연구에서 학습된 신경망으로부터 특징과 클래스간의 관계를 분석하기 위하여 다음과 같은 4종류의 연관도 요소를 정의할 수 있다.

- $RF1(x_i, B_j)$ : 특징값  $x_i$ 와 하이퍼박스  $j$ 와의 연관도 요소
- $RF2(x_i, C_k)$ : 특징값  $x_i$ 와 클래스  $k$ 와의 연관도 요소
- $RF3(X_i, C_k)$ : 임의의 특징타입과 클래스간의 연관도 요소
- $RF4(X_i)$ : 주어진 패턴 분류 문제에서 임의의 특징에 대한 상대적 중요도

이는 각각 다음과 같은 식으로 정의할 수 있다.

$$RF1(x_i, B_j) = w_{ij} \quad (3)$$

$$RF2(x_i, C_k) = \left( \frac{1}{N_k} \sum_{B_j \in C_k} S(x_i, (u_{ji}, v_{ji})) \right) \cdot w_{ij} - \frac{1}{(N_B - N_k)} \sum_{B_j \in C_k} S(x_i, (u_{ji}, v_{ji})) \cdot w_{ij} / \sum_{B_j \in C_k} w_{ij} \quad (4)$$

$$RF3(X_i, C_k) = \frac{1}{L_i} \sum_{x_j \in X_i} RF2(x_j, C_k) \quad (5)$$

$$RF4(X_i) = \frac{1}{M} \sum_{j=1}^M RF3(X_i, C_j) \quad (6)$$

위와 같은 연관도 요소를 적용하면 입력되는 특징과 클래스와의 관계를 정의하여 각 특징의 상대적 중요도를 평가

할 수 있게 된다. 이는 색상 필터에서 피부색 검출을 위한 주요한 특징을 선별하여 사용할 수 있게 하며, 얼굴 검출 신경망에서 44개의 특징지도를 통해 생성된 1848개의 특징 중 주요 특징만을 선별하여 사용하게 해 신경망의 규모를 줄이고, 연산량을 줄인다.

5. 얼굴 영역 보정

최종적으로 검출된 얼굴 영역은 실제 얼굴 영역보다 크게 영역이 선정되게 된다. 이는 색상 필터를 사용하여 후보영역을 선정 할 경우 목 등의 영역까지 얼굴의 후보 영역으로 선정하기 때문인데, 이를 개선하기 위해 눈과 입술 영역을 검출하여 통계 얼굴 크기를 고려하여 얼굴 영역을 보정한다. 얼굴 영역 보정 작업은 그림 2와 같이 진행된다.

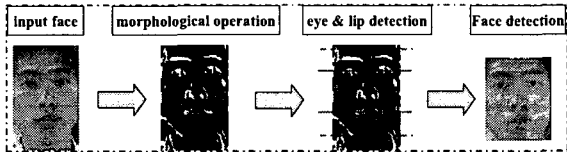


그림 2 얼굴 영역 보정 과정

입력된 얼굴 영역에 대하여 형태학적 닫힘 연산을 수행하여 원본 영상과의 차 영상을 구하면 얼굴 내의 눈과 입술의 영역이 추출된다. 눈과 입술의 영역으로부터 얼굴의 기하학적 정보의 통계적 크기[2,6]를 이용하여 최종적인 얼굴 영역을 추출해 내게 된다.

6. 실험 결과

본 연구는 실내 영상에서 실시간 얼굴 검출을 대상으로 한다. 그림 3과 그림 4는 제안된 이론을 적용한 얼굴 패턴 검출 시스템의 예이다.

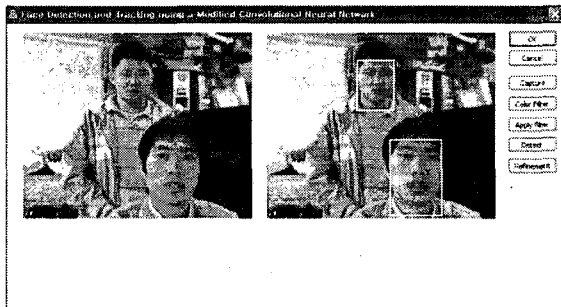


그림 3 얼굴 검출 시스템의 예

7. 결론

본 연구에서는 정확한 얼굴 검출을 위한 다단계 필터와 특화된 신경망 및 얼굴 보정 기법에 대해 소개 하였다. 이는 실내 영상을 대상으로 한 실시간 인식을 목표로 한다. 후보영역 선정을 위해 사용한 조명 보정과 필터링은 검

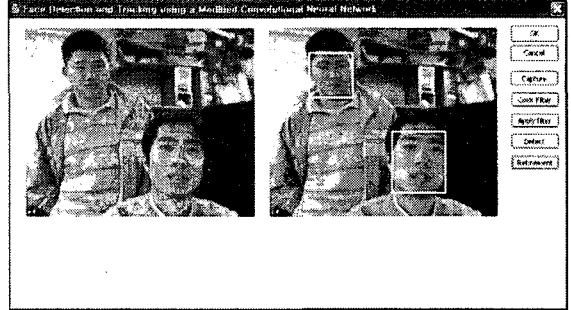


그림 4 얼굴 보정 후의 예

출 대상을 줄임으로써 시스템의 연산량과 속도를 개선한다. 얼굴 특징 지도 생성을 위해 사용한 CNN은 움직임이나 회전에 강인한 특징을 생성하게 하고, 얼굴패턴 검출을 위해 사용한 WFMM신경망은 특징과 클래스간의 연관도 요소를 고려하여 1848개의 특징들을 선별하여 사용하도록 한다. 정확한 얼굴 영역 선정을 위해 사용한 얼굴 보정 기법은 눈과 입술의 위치를 통해 정확한 얼굴의 영역을 추출하게 한다.

\* 이 연구는 산업자원부 뇌 신경 정보학연구 프로그램의 지원으로 수행되었음

8. 참고 문헌

- [1] Rein-Lien Hsu, Mohamed Abdel-Mottaleb and Anil K. Jain, " Face Detection in Color Images," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.24, No.5, pp. 696-706, 2002.
- [2] Frank Y.Shin, Chao-Fa Chuang, " Automatic extraction of head and face boundaries and facial features" , Information Sciences 158, pp. 117-130, 2004.
- [3] Raphael Feraud, Olivier J. Bernier, Jean-Emmanuel, and Michel Collobert " A Fast and Accurate Face Detector Based on Neural Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 1 2001
- [4] Christophe Garcia, Manolis Delakis, "Convolutional Face Finder: A Neural Architecture for Fast and Robust Face Detection," IEEE Transactions on Neural Networks, Vol. 26, No. 11, 2004
- [5] Ho Joon Kim, Hyun Seung Yang, A Modified FMM Neural Networks for Pattern Classification and Feature Analysis, Proceeding of IASTED Conference on Artificial Intelligence and Application(AIA-2005), 2005, 774~778
- [6] Kwok-Wai Wong, Kin-Man Lam, Wan-Chi Siu, "An Efficient Algorithm for Human Face Detection and Facial Feature Extraction under Different Condition", Pattern Recognition, Vol.34, 1993-2004 2001