

MCT와 신경망을 이용한 얼굴 오검출 감소 알고리즘 개발

라승탁*, 이승호**

*국립 한밭대학교 전자·제어공학과

**국립 한밭대학교 전자·제어공학과

*e-mail:saint2347@naver.com, **e-mail:shlee@cad.hanbat.ac.kr

Development of Reduction Algorithm for Face Detection Error Using MCT and Neural Network

Seung-Tak Ra*, Seung-Ho Lee**

*Dept of Electronics&Control Engineering, Hanbat National University

**Dept of Electronics&Control Engineering, Hanbat National University

요 약

OpenCV(Open Computer Vision)에서 제공하는 얼굴 검출 알고리즘은 Haar-like feature와 Cascade 방식을 이용하여 얼굴의 패턴을 찾아내 얼굴을 검출한다. 그러나 우연히 얼굴이 아닌 곳이 얼굴과 유사한 패턴일 경우, 얼굴로 인식하는 오류를 범하게 된다. 따라서 본 논문은 MCT(Modified Census Transform)와 신경망을 이용하여 잘못된 얼굴 검출 영역을 감소시키는 알고리즘을 제안한다. MCT는 다양한 조명 조건에서도 강인한 얼굴 영상의 지역적 구조 특징을 추출하기 위하여 사용되고, 신경망 알고리즘은 Haar-Cascade 알고리즘의 얼굴 검출 방법으로 검출된 영역이 실제로 얼굴인지 아닌지를 판단하기 위하여 사용된다. 실험에서 사용된 6개의 데이터들은 인터넷에서 수집한 것으로서, Haar-Cascade 알고리즘의 얼굴 검출 방법으로 얼굴을 검출하였을 때 오검출된 영역이 1개 이상 존재한다. 본 논문에서 제안한 알고리즘으로 실험한 결과, Haar-Cascade 알고리즘의 얼굴 검출 방법에 비하여 오검출된 영역이 감소된 것을 확인할 수 있었다.

1. 서론

오래전부터 이용해왔던 사람의 신원 확인을 위한 수단으로는 주민등록증, 운전면허증 등과 같은 물리적 매체가 있다. 그러나 이러한 물리적 매체의 경우 사용자가 소지하지 않으면 신원 확인을 할 수 없고, 혹시라도 타인이 매체를 소지한다면 악용될 수 있는 단점이 있다.

이런 문제점을 보완하기 위한 지문 인식, 홍채 인식 등과 같은 생체 인식 기술은 오래전부터 다양하게 연구가 진행되어 왔다. 이러한 생체 인식 기술 중에서도 얼굴 인식 기술은 다른 생체 정보에 비해 인식 과정에서 사용자에게 강제성이 적고, 비 접촉식이라는 특징이 있어 거부감이 덜하다. 특히, 그 활용가치를 인정받고 상업적 응용분야로의 진출을 위해 독립적으로 활발한 연구가 진행 중이다.

OpenCV에서 제공하는 얼굴 검출 방법[1]은 Haar-like feature와 Cascade 방식을 이용하여 얼굴의 패턴을 찾아내어 얼굴을 검출한다. 하지만 우연히 얼굴이 아닌 곳이 얼굴과 유사한 패턴일 경우, 얼굴로 인식하는 오류를 범하게 된다. 이에 본 논문은 이러한 오검출된 얼굴 영역을 감소시키기 위하여 얼굴로 검출된 영역에 한하여 MCT와 신경망 알고리즘을 사용한다.

본 논문의 구성은 다음과 같다. 2장에서는 MCT에 대하여 설명을 하였다. MCT는 다양한 조명변화에 둔감하게

만드는 방법으로 영상의 지역적 구조 특징 정보만을 추출할 수 있다. 얼굴로 검출된 영역에 한해서 MCT를 수행하여 지역적 구조 특징을 추출한다. 3장에서는 신경망 알고리즘에 대하여 설명하였다. 신경망을 통해 얼굴로 검출된 영역이 진짜 얼굴인지 아닌지 판단하게 된다. 신경망에 입력으로 들어가는 얼굴 검출 영역은 MCT가 수행되었기 때문에 조명변화에 강인한 특징벡터로 이루어져 있다. 4장에서는 본 논문에서 제안하는 얼굴 오검출 감소 알고리즘에 대해 설명한다. 5장에서는 본 논문에서 제안한 알고리즘을 인터넷에서 수집한 데이터들을 실험하여 효용성을 입증한다.

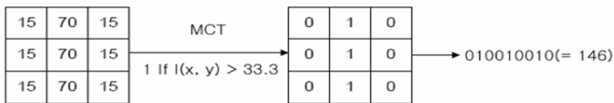
2. MCT

컴퓨터가 영상을 인식하는 과정에서 만약 그 영상이 조명에 의한 영향이 적어 고려하지 않아도 된다면, 인식하려는 객체의 반사성질에 의한 영향만이 고려되어 가장 이상적이라고 할 수 있다[2][3]. 그러나 영상에서 객체의 밝기 값 $I(X)$ 는 조명 값 $L(X)$ 와 객체가 빛을 반사하는 반사 값 $R(X)$ 의 곱으로 정의된다. 또한 카메라가 영상을 취득할 때 생기는 상수 값인 Gain(g)와 바이어스 값 b 도 밝기 값 $I(X)$ 에 영향을 준다. 따라서 밝기 값 $I(X)$ 는 아래와 같은 식으로 정의될 수 있다. X 는 각 픽셀의 위치 (x, y) 를 나타낸다.

$$I(X) = gL(X)R(X) + b \quad (1)$$

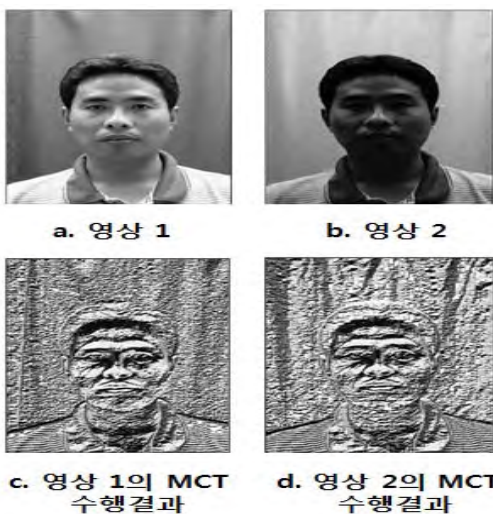
식 (1)에 의해 조명 값 $L(X)$ 에 대한 어떠한 가정이 없다면 반사 값 $R(X)$ 를 구하는 것은 불가능하다. 따라서 MCT는 영상의 특징으로 $R(X)$ 만을 이용하기 위해 매우 작은 크기의 윈도우내에서 $L(X)$ 의 값은 변하지 않는다는 가정을 한다. 이러한 가정을 통해 지역적 구조 특징은 $R(X)$ 에 의한 영향만을 받게 된다. 따라서 어떤 영상에 대하여 MCT를 수행하면 조명 값 $L(X)$ 에 의한 영향은 거의 없고, 오직 객체의 반사성질 $R(X)$ 만을 고려한 영상이 된다.

MCT를 수행하는 방법은 다음과 같다. 조명에 둔감하게 만들기 위한 기존의 대표적인 변환 방법 중의 하나인 LBP[4]와는 다르게 중심 픽셀의 밝기 값을 기준으로 하지 않는다. 기준이 되는 값은 그 윈도우 내의 픽셀들의 평균 밝기 값이다. 이렇게 기준을 정해지면 윈도우 내의 각 픽셀 밝기 값과 비교함으로써 중심 픽셀에도 값을 부여할 수 있고, 따라서 MCT는 윈도우 내의 모든 가능한 구조적 특징을 표현할 수 있다. 윈도우 내의 픽셀들의 평균 밝기 값보다 밝으면 1, 어두우면 0으로 인코딩하고 그 결과를 비트 스트링으로 연결한 값이 MCT를 수행한 값이 된다. 그림 1은 MCT를 수행하는 과정을 나타낸다.



(그림 1) MCT 수행 과정

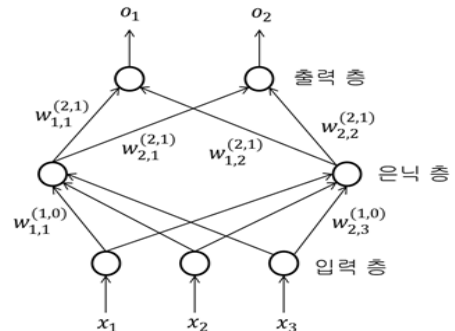
그림 1에서 수행한 방법처럼 3x3 크기의 윈도우를 적용시켜 동일한 조건에서 조명만 다르게 한 영상에 대해 MCT를 수행하면 그림 2와 같다. 그림 2의 a와 b는 조명의 영향 때문에 확연한 차이가 나지만 c와 d는 조명의 영향을 거의 받지 않은 것을 확인할 수 있다.



(그림 2) MCT 수행 결과 영상

3. 신경망

본 논문에서 사용한 신경망은 역전과 알고리즘을 적용시켰다. 역전과 알고리즘이란 신경망의 출력 값이 원하는 출력 값이 나오도록 오차 신호를 역방향으로 전파하여 노드 사이의 가중치를 수정하는 기법이다. 그림 3은 신경망의 구조를 나타낸다. 입력 층은 셋, 출력 층은 두 노드를 갖는다. 여기서 w 는 가중치(Weight)를 뜻한다. w 의 아래 첨자는 도착 노드와 시작 노드의 번호를 나타내고, w 의 위 첨자는 도착 층과 시작 층 번호를 의미한다.



(그림 3) 신경망의 구조

입력 벡터 x_p 가 있다고 할 때 p 는 $1 \leq p \leq P$ 의 범위를 갖고, x_p 에 대한 K 차원 희망 출력 d_p 는 식 (2)와 같이 정의된다. 그 다음 입력 벡터와 그 입력에 대한 희망 출력을 묶어 하나의 훈련 세트를 구성하고 그 식은 (3)과 같다. 이제 입력 벡터 x_p 를 신경망에 입력했을 때 신경망의 실제 출력 o_p 는 식 (4)와 같다.

$$d_p = (d_{p,1}, d_{p,2}, \dots, d_{p,K}) \quad (2)$$

$$(x_p, d_p) : p = 1, \dots, P \quad (3)$$

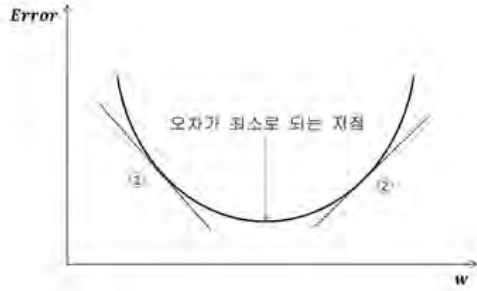
$$o_p = (o_{p,1}, o_{p,2}, \dots, o_{p,K}) \quad (4)$$

역전과 신경망의 목표는 실제 출력 벡터 o_p 가 희망 출력 벡터 d_p 와 거의 비슷하게 되도록 가중치 w 를 조절하는 것이다. 다시 말해 모든 입력 벡터들의 오차 합이 최소가 되도록 가중치 w 를 조절하는 것이다. 오차에 대한 함수식은 식 (5)와 같다. 오차 함수를 제곱 형태로 정하는 이유는 그 함수가 미분 가능하고 함수의 미분 값이 선형 일차식 꼴이므로 경사 강하 알고리즘에 적용하기 쉽기 때문이다.

$$\text{제곱 합 오차} = \sum_{p=1}^P \sum_{k=1}^K (d_{p,k} - o_{p,k})^2 \quad (5)$$

경사 강하 알고리즘에 따르면 w 의 가중치 변화 방향

은 $-dE/dw$ 방향과 같아야 한다. 그림 4에서 ①과 같이 기울기가 음수인 경우 w 는 증가하는 방향으로, ②와 같이 기울기가 양수인 경우 w 는 감소하는 방향으로 나아가야 한다. 위의 그림 3 신경망의 구조와 같이 한 은닉층을 갖는 인공신경망에서 w 의 변화 양은 식 (6)과 식 (7)과 같이 표현될 수 있다.



(그림 4) 2차 오차 함수 그래프

$$\Delta w_{k,j}^{(2,1)} \propto \left(\frac{-dE}{dw_{k,j}^{(2,1)}} \right) \quad (6)$$

$$\Delta w_{j,i}^{(1,0)} \propto \left(\frac{-dE}{dw_{j,i}^{(1,0)}} \right) \quad (7)$$

오차 함수 E 는 식 (8)과 같이 $w_{k,j}^{(2,1)}$ 를 변수로 갖는 함수이다. 따라서 식 (6)은 E 를 $w_{k,j}^{(2,1)}$ 에 대해 미분하여 계산될 수 있고 그 값은 식 (9)와 같다.

$$E = \sum_k (d_k - o_k)^2 = \sum_k (d_k - S(\sum_j w_{k,j}^{(2,1)} x_j))^2 \quad (8)$$

$$\frac{dE}{dw_{k,j}^{(2,1)}} = -2(d_k - o_k) S(net_k) x_j \quad (9)$$

또한 식 (8)의 마지막 항에서 $x_j = \sum_i (w_{j,i}^{(1,0)} x_i)$ 이므로 E 는 또한 $w_{j,i}^{(1,0)}$ 를 변수로 갖는 함수이다. 따라서 식 (7)은 E 를 $w_{j,i}^{(1,0)}$ 에 대해 미분하여 계산될 수 있고 그 값은 식 (10)과 같다.

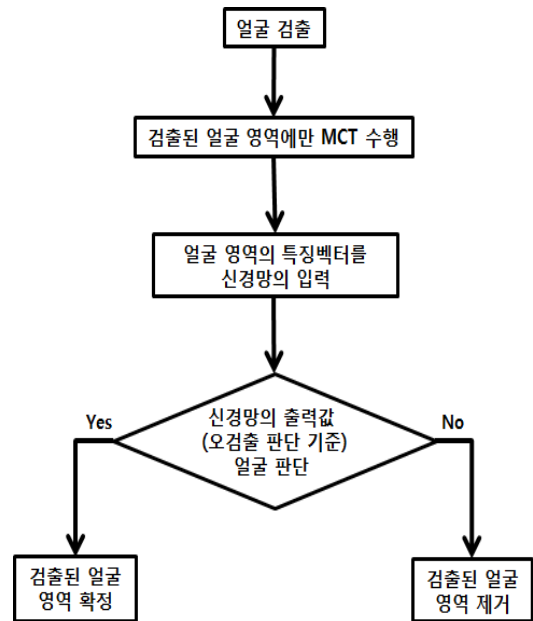
$$\frac{dE}{dw_{j,i}^{(1,0)}} = \sum_{k=1}^K -2(d_k - o_k) S(net_k) w_{k,j}^{(2,1)} S(net_j) x_i \quad (10)$$

여기서 net_k 는 k 층의 모든 노드들의 가중된 합을 뜻한다. 최종적으로 역전파 알고리즘을 이용하여 신경망을 학습시킨다는 의미는 위의 식들을 이용하여 오차 신호를 역방향으로 전파하여 노드 사이의 가중치 즉, w 를 원하는 에러율이 나올 때까지 계속 수정하여 최적의 가중치를 얻는 것을 의미한다.

4. 얼굴 오검출 감소 알고리즘

본 논문에서 제안하는 얼굴 오검출 감소 알고리즘은 다음과 같은 과정을 수행한다. 첫 번째, 입력 영상에 대해 Haar-Cascade 알고리즘의 얼굴 검출 방법을 통해 얼굴을 검출한다. 만약 입력 영상에 우연히 얼굴과 유사한 패턴이 존재한다면 얼굴이 아닌 곳을 얼굴로 검출하는 오검출이 존재할 수 있다. 두 번째, 얼굴로 검출된 얼굴영역에 대해서만 MCT를 수행한다. MCT를 수행함으로써 다양한 조명 조건에서도 강인하게 얼굴로 검출된 영역의 지역적 구조 특징을 추출할 수 있다. 세 번째, MCT를 수행한 얼굴영역의 특징 벡터를 신경망의 입력으로 넣는다. 신경망의 출력 값에 따라 얼굴인지 아닌지를 판단하여 최종적으로 신경망의 출력 값이 얼굴로 판별된 값으로 나왔다면 본래 얼굴영역으로 검출된 부분을 얼굴로 확정한다. 만약 신경망의 출력 값이 얼굴이 아닌 것으로 판별되었다면 본래 얼굴영역으로 검출된 부분은 제거한다.

신경망을 학습시킬 때 사용한 데이터들은 인터넷에서 수집하였고, 얼굴 이미지와 비얼굴 이미지를 학습시켰다 [5]. 이때 비얼굴 이미지는 오검출된 이미지를 뜻한다. 그림 5는 본 논문에서 제안하는 얼굴 오검출 감소 알고리즘의 흐름도를 나타낸다.



(그림 5) 얼굴 오검출 감소 알고리즘의 흐름도

5. 실험결과

5.1 실험에 사용한 데이터들

실험에서 사용된 6개의 데이터들은 인터넷에서 수집한 것으로서, Haar-Cascade 알고리즘의 얼굴 검출 방법으로 얼굴을 검출하였을 때 오검출된 영역이 1개 이상 존재한다. 그림 6은 실험에 사용된 데이터 중의 일부이다.



(그림 6) 오검출된 영역이 존재하는 데이터

5.2 실험결과 및 고찰

그림 7은 본 논문에서 제안한 알고리즘을 사용하여 그림 6의 얼굴이 아닌 영역이 제거된 영상을 나타내고 있다. 인터넷에서 수집한 6개의 데이터들에 대하여 표 1은 Haar-Cascade 알고리즘과 본 논문에서 제안하는 알고리즘의 실험 결과를 비교하여 나타낸 것이다. 표 1에서 나타난 바와 같이 본 논문에서 제안한 알고리즘의 얼굴 오검출률이 Haar-Cascade 알고리즘에 비하여 우수한 것이 입증되었다.



(그림 7) 오검출된 영역이 감소된 데이터

<표 1> 6개의 데이터들에 대한 실험 결과

	Haar-Cascade 알고리즘		제안한 알고리즘	
	얼굴 검출률	오검출 횟수	얼굴 검출률	오검출 횟수
데이터 1	100% (1/1)	1	100% (1/1)	0
데이터 2	100% (1/1)	1	100% (1/1)	0
데이터 3	100% (1/1)	1	100% (1/1)	0
데이터 4	46% (6/13)	1	46% (6/13)	0
데이터 5	100% (10/10)	1	100% (10/10)	0
데이터 6	92% (11/12)	1	92% (11/12)	0

5. 결론

OpenCV에서 제공하는 얼굴 검출 알고리즘은 Haar-like feature와 Cascade 방식을 이용하여 얼굴의 패턴을 찾아내 얼굴을 검출한다. 그러나 우연히 얼굴이 아닌 곳이 얼굴과 유사한 패턴일 경우, 얼굴로 인식하는 오류를 범하게 된다. 이러한 문제점을 보완하기 위하여 본 논문에서는 MCT와 신경망을 이용하여 잘못된 얼굴 검출 영역을 감소시키는 알고리즘을 제안하였다. 본 논문에서 제안한 알고리즘으로 실험한 결과, Haar-Cascade 알고리즘의 얼굴 검출 방법으로 얼굴을 검출하였을 때 보다 오검출된 영역이 감소된 것을 확인할 수 있었다.

향후 MCT를 사용해 조명의 영향에 강한 것뿐만 아니라 노이즈에도 강인하게 작용할 수 있는 변환법에 대한 연구가 필요하다. 얼굴이 검출된 영역에 한하여 조명과 더불어 노이즈에도 강인하게 특징벡터를 추출할 수 있다면 오검출을 더욱 감소시킬 수 있을 것으로 예상된다.

ACKNOWLEDGMENT

This research was financially supported by the Ministry of Education(MOE) and National Research Foundation of Korea(NRF) through the Human Resource Training Project for Regional Innovation (No.2015H1C1A1035818)

참고문헌

- [1] P. Viola and M. Jones, "Robust Real-Time Face Detection", International Journal of Computer Vision, pp. 137-154, 2004.
- [2] B. Froba and A. Ernst, "Face detection with the modified census transform", Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 91-96, May. 2004.
- [3] 조광신, 박수경, 심동규, 이수연, "조명변화에 강인한 MCT와 프레임 연관성 기반 실시간 얼굴인식 시스템", 전자공학회 논문지 45권 CI편, 제3호, 5월, pp. 123-134. 2008.
- [4] T. Pajdla and J. Matas, "Face Recognition with Local Binary Patterns", Computer vision-eccv 2004. Springer Berlin Heidelberg, 2004., pp. 469-481.
- [5] HA Rowley, S Baluja and T Kanade, "Neural network-based face detection.", IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 23-38, Jan. 1998.