

객체 분할 및 주성분 분석 기반의 얼굴 추적 인식 알고리즘

성민영, 김대현, 이응주
동명정보대학교 정보통신공학과

Face Tracking and Recognition Algorithm Based On Object Segmentation and PCA

Min-Young Sung, Dae-hyun Kim and Eung-Joo Lee
Dept. of Information/Communication Eng., TongMyong Univ. of Information Technology.

E-mail : ejlee@tit.ac.kr

Abstract

본 논문에서는 실시간 출입통제시스템에 적용이 가능한 복잡한 배경에서의 다중 얼굴 영역 검출과 추적을 통한 얼굴 인식 알고리즘을 제안하였다. 제안된 알고리즘에서는 배경영상과 입력된 연속적인 프레임간의 차영상을 적용함으로써 물체의 움직임을 감지한 후, IISI컬러 좌표모델을 이용하여 얼굴의 1차 후보 영역을 검출하고, 잡음제거를 위해 모폴로지 연산을 수행하였다. 또한 Line Projection을 이용한 객체 분할법(Object Segmentation)으로 객체를 분할함으로써 다중 얼굴 영역을 추출하였다. 또한 추출된 얼굴영역에서 눈 영역 검출을 통해 각각의 얼굴 영역들을 검증하였으며 검증된 얼굴들의 최외각 4개의 좌표를 이용하여 얼굴 추적율을 높였다. 마지막으로 얼굴 인식은 추출된 얼굴 영역으로부터 주성분 분석(PCA: Principle Component Analysis)방법을 이용함으로써 97~98%의 높은 인식율을 보였다.

1.서론

최근 정보 통신 기술과 멀티미디어 기술의 급속한 발달로 인해 개인의 신원 확인 및 보안 시스템에 관한 관심이 급증하고 있다. 생체 인식과 관련된 연구로는 얼굴, 지문, 제스처, 서명, 음성, 홍채, 정맥 인식 등의 분야가 있으며, 그 중 얼굴 인식은 자신의 신체 일부를 인식장치에 접촉하지 않고 자연스러운 상태에서 인식할 수 있는 비접촉식으로 주목받고 있는 분야이다.

얼굴 인식을 위한 과정은 얼굴 영역을 검출한 후, 검출된 얼굴 영역을 추출하여 인식하는 과정으로 되어 있다. 얼굴 인식을 위한 첫 단계로서 얼굴 영역을 검출 하는 방법에는 색 정보를 이용하는 방법[1], 색 정보와 동영상을 이용하는 방법[2], 신경망을 이용하는 방법[3] 등 여러 가지 방법들이 있으나, 본 논문에서는 다중 얼굴 영역과 실시간 동영상에서 빠른 얼굴 검출을 위해 색 정보와 함께 객체 분할법(Object Segmentation)을 이용하는 방법을 제안한다. 또한, 기존의 얼굴 인식을 위한 방법으로는 크게 두 가지로 나뉘 수 있다.

얼굴 영상 전체를 이용하여 특징을 계산하는 방법인 영상 기반의 인식방법[4]과 눈, 코, 입과 같은 특징점을 추출하여 이들 간의 기하학적인 관계를 분석하거나 비교하여 얼굴 인식하는 특징점 기반의 인식방법[5]이 있다. 그러나 이러한 방법들은 실시간 처리에 문제점이 잘 일어난다. 본 논문에서는 수행 속도가 빠른 장점을 지닌 PCA방법[6]을 채택하였다. 본 논문의 전체적인 구성은 움직임 영역을 검출한 후, IISI 좌표계에 의한 색 분할로 1차 얼굴 후보 영역들을 검출하고 선투영(Line Projection)에 의한 객체 분할법으로 객체를 분할함으로써 다중 얼굴 영역들을 추출하고 그 얼굴 영역들에 대해 눈을 검출함으로써 얼굴을 검증하도록 하였다. 검증된 각각의 얼굴 영역들의 최외각 4개의 좌표만을 가지고 얼굴을 추적하고, 주성분 분석 기반 방법으로 얼굴을 인식하도록 하였다.

2. 물체 움직임 영역 검출

본 장에서는 실시간 프레임에서 움직임을 감지하기 위해 차영상을 이용하였으며, 이 방법은 두 영상간의 화소값 차를 이용하여 처리 시간이 단축되는 결과를 가져오는 장점이 있다. 물체 움직임 감지를 위해 우선, 배경 영상과 실시간으로 입력되는 영상간의 차영상을 구하고 현재 영상과 이전 영상 간의 차영상을 구한다. 차영상을 구하는 식은 다음 두 식과 같다.

$$BI_{0j}(x, y) = \begin{cases} 1 & \text{if } |B(x, y, t_0) - I(x, y, t_j)| > T \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

$$DI_{ij}(x, y) = \begin{cases} 1 & \text{if } |I(x, y, t_i) - I(x, y, t_j)| > T \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

여기서, T 는 조명에 의한 잡음을 제거하기 위한 문턱치값이다. 또, $B(x, y, t_0)$ 은 배경영상이고, $I(x, y, t_i)$ 과 $I(x, y, t_j)$ 는 시간 t_i 와 t_j 에서 획득된 영상들이고, $BI_{0j}(x, y)$ 와 $DI_{ij}(x, y)$ 는 T 에 의해 이진화된 차영상들이다. 그리고 최종적인 차영상 $Diff(x, y)$ 은 배경영상과 입력영상간의 논리곱으로 구하고 식(3)에 나타내었다.

$$Diff(x, y) = BI_{0j}(x, y) \cdot DI_{ij}(x, y) \quad (3)$$

구해진 차영상의 전체 픽셀에 대해 변화된 개수가 전체 영상에 대한 최소한의 움직임으로 판단되는 변화된 개수 이상일 때, 움직임이 있는 것으로 판단한다. 그림 1은 배경영상과 입력영상들 간의 차영상을 구해 움직임을 검출한 결과를 보여준다. 그림 1(d)에서 사각형은 움직임이 검출된 영역에 해당한다.

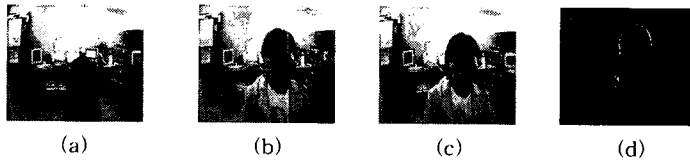


그림 1. 물체 움직임 영역 검출: (a) 배경영상, (b) 이전영상, (c) 현재영상, (d) 움직임 영역 검출.

3. 다중 얼굴 영역 검출과 추적

다중 얼굴 영역을 검출하기 위해 HSI 컬러 모델에 의한 색 분할로 1차 얼굴 후보 영역을 검출한다. 잡음 제거를 위해 모폴로지 연산을 수행한 후, 선 투영(Line Projection)에 의해 각각의 얼굴을 의미 있는 객체로 분할하여 얼굴을 추출한다. 마지막으로 얼굴 검증 단계로서 눈을 검출하여 얼굴을 판단하도록 하였다.

3.1 HSI 정보를 이용한 색 분할

연속된 컬러 영상에서 1차 후보 얼굴 영역을 검출하고자 본 논문에서는 명도와 색차를 독립적으로 다룰 수 있는 HSI 정보 중 색상과 명암도 정보를 이용하여 피부색 영역을 분할하였다. 이때, 조명의 영향을 받을 경우 얼굴 피부색에 해당하는 분포도가 명암도에 의해 변화하는 것을 볼 수 있다[7]. 따라서 조명의 영향을 줄이기 위해 앞장에서 생성한 배경 영상과 입력되는 영상간의 전체 평균 명암도 성분을 비교하여 일정한 임계치 α 에서 벗어날 때, Hue값 범위를 조정하여 검출하였다. 식(4)~식(6)은 RGB좌표계에서 HSI좌표계로의 변환 공식을 보여준다.

$$I = \frac{1}{3} (R + G + B) \quad (4)$$

$$S = 1 - \frac{3[\min(R, G, B)]}{(R + G + B)} \quad (5)$$

$$H = \cos \left[\frac{\frac{1}{2} |(R - G) + (R - B)|}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right] \quad (6)$$

움직임이 발생한 영역내의 밀집된 Hue성분 분포를 구함으로써 잡음을 제거하기 위한 열림(Opening)연산을 수행한다. 적어도 얼굴 하나의 크기값이 존재할 전체 영상에 대한 픽셀 수의 1/5정도의 값에 해당하는 임계값(Th_{face})에 만족되면 다음 단계를 수행한다. 그림 2는 HSI 칼라 모델에 의한 색 분할 및 열림 연산 수행 후의 결과 영상인 $M(x, y)$ 을 보여준다.

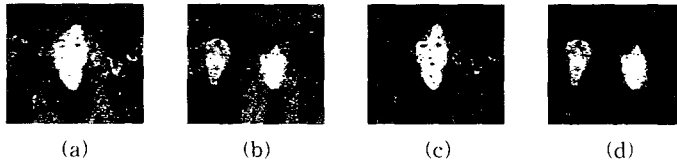


그림 2. HSI 정보를 이용한 색 분할 및 열림 연산; (a) 1차 열림 후보 영역(1인), (b) 1차 열림 후보 영역(2인), (c) (a)의 영상에 대한 열림 연산 결과, (d) (b)의 영상에 대한 열림 연산 결과.

3.2 객체 분할법에 의한 다중 얼굴 영역 추출

앞서 구한 1차 열림 후보 영역들 간의 객체들을 추출하고자 우선, $M \times N$ 크기인 $M(x, y)$ 에서 선 투영 처리를 수행한다. 선 투영이란 한 라인에 대한 픽셀의 누적값을 구하는 것이고 이를 식 (7)과 (8)에 나타내었다.

$$V(x) = \sum_{y=0}^N M(x, y) \quad \text{if } |M(x, y)| = 1 \quad (7)$$

$$H(y) = \sum_{x=0}^M M(x, y) \quad \text{if } |M(x, y)| = 1 \quad (8)$$

그 다음, 두 $V(x)$ 와 $H(y)$ 로 구한 프로파일(Profile)에서 얼굴 영역을 추출하기 위해 컨볼루션(Convolution)연산을 수행함으로써 객체의 모양을 좀 더 구체적으로 표현하고, 골(Valley)과 마루(Peak)점을 찾는데 사용하였다. 또한, 객체 주변의 잡음을 고려하여 큰 객체를 기준으로 두 개의 문턱치($T1, T2$)들을 구하여 프로파일 형태를 수정함으로써 잡음을 제거할 뿐 아니라, 1차적으로 가까이에 있는 객체를 분리 할 수 있다. 여기서, 문턱치($T1, T2$)들은 $V(x), H(y)$ 각각의 상위 20%에 해당하는 값들이다. 수정 후, 두 프로파일에 대해 각각 하나의 객체가 존재하면 다음 단계인 얼굴 검증 단계를 수행한다. 그렇지 않을 경우 객체 분할 과정을 수행한다. 그림 3에 선 투영에 의한 객체 분할 방법을 나타내었다.

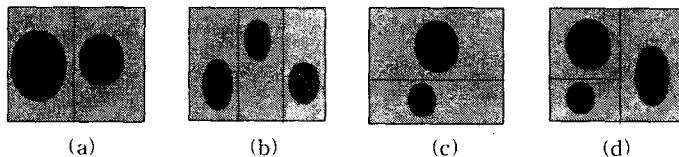


그림 3 선 투영에 의한 객체 분할 방법 모형; (a) 2개가 좌우로 있는 경우, (b) 3개가 좌우로 있는 경우, (c) 2개가 상하로 있는 경우, (d) 상하, 좌우로 총 3개가 있는 경우.

객체 분할 방법은 먼저 $V(x)$ 의 프로파일에서 각각의 객체에 대한 넓이를 계산한다. 최대 넓이에 해당하는 객체를 기준으로 상대 객체들의 넓이 비율이 임계치 k 보다 작으면 그 객체는 큰 객체를 기준으로 멀리 떨어져 있는 객체이거나 얼굴이 아닌 객체로 판단하여 추출하지 않는다. 인식을 위한 최대 거리를 고려하기 때문에 이와 같은 조건은 가능하다. 위의 과정을 수행한 후, $V(x)$ 의 프로파일에서 하나의 객체를 가지면 $H(y)$ 의 프로파일에도 위와 같은

과정을 반복 수행한다. $H(y)$ 도 객체 제거 후 하나의 객체가 나오면 분할을 중단한다. 그렇지 않고 $V(x)$ 에서 객체 제거 후에도 두개의 객체 또는 세 개의 객체가 나오면 그 객체들 간의 중심을 기준으로 $M(x, y)$ 을 분할한다. 분할된 $M(x, y)$ 에 대해 다시 위와 같은 과정을 각각 반복 수행한다. 즉, 분할된 영상에서 각각 하나의 객체만 가지도록 분할을 계속 수행한다. 그림 4는 객체 분할법에 의한 얼굴 영역을 검출한 결과 영상이다.

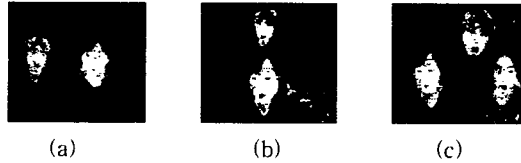


그림 4 객체 분할법에 의한 얼굴 영역 검출; (a) 얼굴이 상하에 위치할 경우, (b) 얼굴이 좌우로 위치할 경우, (c) 상하, 좌우로 위치할 경우.

3.3 다중 얼굴 영역 검증과 얼굴 추적

이 절에서는 인식을 위한 얼굴 검출의 마지막 단계로서 추출된 각각의 객체에 대한 눈의 존재 여부를 판단한다. 얼굴 영역들을 대상으로 HSI의 정보 중 I성분을 이용하여 명암도 영상으로 변환한다. 그리고 평균 이진화를 수행한 후, x축 방향으로 투영하여 프로파일 형태로 눈 특징자의 존재 여부를 파악한다. 눈이 얼굴의 윗부분에 존재한다는 형태학적 정보를 이용하여 프로파일의 윗부분만을 검색하여 골(valley)의 간격, 위치, 다음 골의 존재 유무, 모양으로 눈을 판단한다. 그림 5는 눈 영역 검출 과정 및 결과를 보여준다. 그림 5 (d)는 이진화된 얼굴의 수평 프로파일로서 원안에 있는 골 형태가 눈 영역이며 얼굴과 비슷한 형태를 가지고 있는 손의 경우는 이러한 골의 형태를 가지지 않으므로 얼굴과 손을 구별할 수 있다.

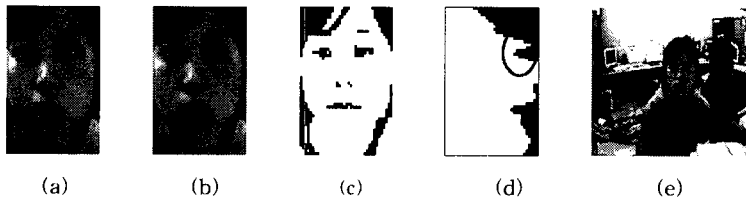


그림 5. 눈 검출 과정 및 결과; (a)추출된 얼굴 영역, (b) 명암도 변환된 영상, (c) 평균 이진화, (d) 프로파일 형태, (e) 검출된 눈 영역 결과 영상.

또한, 검증된 각각의 얼굴들은 최외각 4개의 좌표의 이동으로 추적(Tracking)하며, 다수의 얼굴 추적의 경우, 얼굴들의 4개의 좌표 이동점간의 가장 짧은 거리에 해당하는 점들로 이루어진 얼굴이 현재 영상에서의 그 얼굴에 대한 이동 객체가 된다. 그림 6에 연속 프레임에 대한 얼굴 추적 결과를 나타내었다.

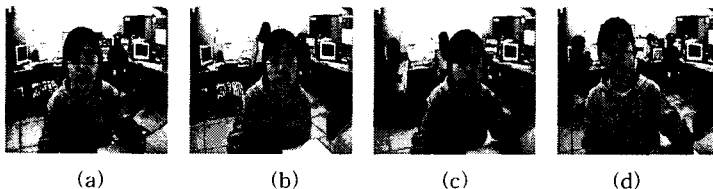


그림 6. 연속적인 프레임에 대한 얼굴 추적 결과.

4. 주성분 분석(Principle Component Analysis)를 이용한 얼굴 인식

추출된 얼굴 영역들 중 한명의 사용자가 카메라에 일정한 거리에 서있고, 임계값 ρ 의 명암값 분포를 가질 때 자동 캡처(Capture)하는 형식으로 된다. 캡처된 영상들은 DB에 저장되어 인식을 위한 참조 얼굴이 되거나 인식을 위한 영상이 된다.

$$C = \frac{1}{M} \sum_{n=1}^M \phi_n \phi_n^T \quad (9)$$

식(9)는 DB에 저장된 얼굴 영상과 평균 얼굴과의 차이를 공분산 행렬로 나타낸 공식이다. 이 공분산 행렬 중 가중치 값이 큰 것 순으로 정렬하여 일반적으로 얼굴 인식에서 전체 가중치 값의 92 ~ 95%정도에 해당하는 값까지 얼굴을 선정하여 입력되는 얼굴과 같이 재배열하는데 이를 고유 얼굴(Eigen face)이라 한다.

$$w_{ik} = u_k^T (\Gamma_i - \Psi) \quad (k = 1, 2, \dots, M') \quad (10)$$

$$\Omega_i^T = [u_{i1}, u_{i2}, \dots, u_{iM'}] \quad (11)$$

식(10)에서 입력 영상에 대한 가중치 값은 고유 얼굴과의 연산을 의미하여, 식(11)에서는 각 얼굴 영상의 성분벡터를 의미한다. 인식을 원하는 영상은 전체 얼굴 영상이 아닌 이러한 성분 벡터 값들만으로 얼굴을 표현하게 되고, 인식할 때 그 입력 영상에 대해서 동일한 과정으로 성분 벡터를 구한 후 Euclidian distance가 최소가 되는 얼굴 영상이 바로 가장 흡사한 얼굴로 인식된다. 그림 7은 PCA를 이용한 인식 과정을 보여준다.

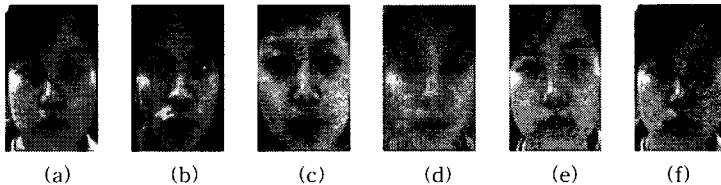


그림 7. PCA를 이용한 얼굴 인식; (a)(b)(c) 정규화된 얼굴 인식용 DB영상들, (d) DB 얼굴영상에 대한 평균, (e) 추출된 인식을 위한 얼굴 영상, (f) 인식 결과 DB에 저장된 동일한 인식.

5. 실험 및 결과

본 실험은 CCD 컬러 카메라와 Matrox Meteor-II board를 이용해서 30frame/sec의 640×480, 24bit 컬러 영상을 입력으로 하였다. 실험에서는 다중 얼굴 영역 추출과 인식을 위해 100인을 대상으로 실험하였다.

표 1. 다중 얼굴 영역 추출과 인식 결과.

전체 영상	다중 얼굴 영역 추출율			얼굴 인식율
	1인	2인	3인	
100	99.8%	98.7%	97.2%	96.1%

표 1은 다중 얼굴 영역 추출율과 인식율을 보여준다. 표 1에서 보는 것처럼 1인과 2인 추출율은 비교적 높게 나타났으나 3인 추출율은 1인과 2인 추출율에 비해 상대적으로 낮게 나타났다. 정확한 추출 시 인식율은 97~98.4%로 높은 인식율과 빠른 처리 속도를 보였다.

7. 결 론

본 논문에서는 움직임 감지 후, 색 분할법으로 얼굴 피부영역을 분할하여 객체 분할법에 의한 다중 얼굴 영역 추출 및 추적을 제안하여 추출율을 높였으며, 또한, 얼굴 인식에 적합한 얼굴 영상을 자동 캡처하여 인식율을 높였다. 인식은 PCA를 이용하여 실시간 인식 시스템에 적용이 가능하도록 하였다. 향후 과제로는 국부적인 조명의 영향을 줄이는 방향과 보다 정확한 얼굴 검출에 관한 연구 및 기울어진 얼굴을 보정하는 과제가 남아 있다.

[참고 문헌]

- [1] Demas Sanger, Yoichi Miyake, et al., "Algorithm for Face Extraction based on Lip Detection," Journal of Imaging Science and Technology, Vol.41, No.1, Jan./Feb., 1997.
- [2] Stan Birchfield, "Elliptical Head Tracking Using Intensity Gradients and Color Histogram," CVPR, 1998.
- [3] Henry A. Rowley, Shumeet Baluja and Takeo Kanaed, "Human Face Detection in Visual Scenes," IEEE Conference on Computer Vision and Pattern Recognition, 1998.
- [4] A.S.Geroghiades, P.N.Belhumeur and D.J.Kriegman, "From few to many: illumination cone models for face recognition under variable lighting and pose." IEEE Trans. Pattern Analysis and Machine Intelligence, vol.23, no.6, pp.643-660, June, 2001.
- [5] Robert Brunelli and Tomaso Poggio, "Face Recognition : Feature versus Templates," PAMI, Vol. 15, no. 10, pp.1042-1052, Oct 1993.
- [6] A.Pentland, B.Moghaddam, T.Starner, and M. Turk, "View-Based and Modular Eigenfaces for face Recognition", Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 84-91, 1994.
- [7] R. S. Berns, Principles of Color Technology, John Wiley & Sons Inc., pp. 22-23, 2000.
- [8] Jean-Christophe Terrillon, Martin David and Shigeru Akamatsu "Detection of Human Faces in Complex Scene Images by Use of a Skin Color Model and of Invariant Fourier-Mellin Moment." IEEE Trans 1998.