

LDA와 인덱싱 테이블을 이용한 피부영역 검출방법

양희성, 강호진, 이준호

성균관대학교 전기전자 및 컴퓨터공학부

Skin Color Segmentation Using LDA and Indexing Table

Hee-Sung Yang, Ho-jin Kang, June-Ho Yi

School of Electrical and Computer Engineering, Sungkyunkwan University

요약

본 논문에서는 복잡한 배경이나 조명 변화가 심한 영상에서도 피부영역을 정확하게 검출할 수 있는 피부영역 검출방법을 제안한다. 제안된 방법은 오프라인(off-line) 훈련과정과 온라인(on-line) 검출과정의 두 단계로 나누어진다. 훈련단계에서는 다양한 조명하에서 얻은 피부영상과 배경영상으로 구성된 훈련영상을 다차원의 열벡터로 표현하고 열벡터에 LDA(linear discriminant analysis)를 적용하여 선형변환된 특징벡터를 가지고 인덱싱 테이블을 생성한다. 검출단계에서는 카메라로 들어온 칼라영상을 여러 개의 조각영상으로 나누고 각각의 조각영상에 대하여 LDA를 적용하여 선형변환된 특징벡터를 구한다. 구해진 특징벡터를 미리 생성한 LDA 인덱싱 테이블에서 찾아 피부영역을 검출한다. 제안된 방법을 조명을 변화시킨 다양한 영상에 적용하여 실험한 결과 검출률이 상당히 우수함을 알 수 있었다.

1. 서론

카메라로 입력된 영상을 가지고 제스처를 인식하거나 얼굴을 인식하는 문제는 컴퓨터 비전의 중요한 응용분야이다. 하지만 오랜 연구에도 불구하고 제스처 인식이나 얼굴 인식은 제한된 환경에서만 이루어져 왔다. 이를 해결하기 위해서는 복잡한 배경이나 조명 변화에서 움직이는 손이나 얼굴 등의 피부 영역을 정확하게 추출하는 방법이 선행되어야 한다. 입력영상에서 피부영역을 획득하는 기존의 방법은 HSI 칼라공간이나 YIQ 칼라공간에서 히스토그램 인터섹션(histogram intersection) 방법[2-4]이나 화소단위에 기반한 룩업테이블(lookup table) 방법[5, 6]을 이용하는 것이 대표적이거나 조명변화가 있는 칼라영상이나 피부와 비슷한 색에 대해서는 검출률이 상당히 저조하여 제한된 환경에서만 실험되어

왔다.

본 논문에서는 조각영상을 다차원의 열벡터로 표현하여 LDA(linear discriminant analysis)를 적용하여 피부영역을 검출하는 방법을 제안한다. 제안된 방법은 H값과 S값으로 각각 구성된 조각영상은 LDA(linear discriminant analysis)를 이용하여 새로운 특징벡터공간의 벡터로 선형변환된다. 줄여진 차원으로 선형변환된 특징벡터를 가지고 인덱싱 테이블을 생성하며 LDA 인덱싱 테이블(indexing table)이라 칭한다. 구해진 LDA 인덱싱 테이블은 오프라인(off-line)에서 다양한 피부색으로 미리 구해진 특징벡터들과 비교되어 피부색인지 아닌지를 판단하게 된다.

제안된 방법을 조명을 변화시킨 다양한 영상에 적용하여 실험하였고 화소단위에 기반한 룩업 테이블 방법을 적용한 실험 결과와 비교 분석하였다. 기존의 방법들은 피부 영역을 화소단위로 처리하여 검출하기 때문에 조명 변화에서 검출률이 상당히 저조한 반면에 제안된 방법은 조각영상을 패턴(pattern)처리하여 피부색을 검출하므로 복잡한 배경이나 조명 변화에서도 보다 정확하게 피부영역을 검출할 수 있다.

2. 기존의 방법

룩업테이블 방법은 입력된 칼라영상을 화소단위로 처리한다. 오프라인(off-line)에서 피부영상인 훈련영상을 적절한 칼라공간으로 변환하여 크로미넌스의 2차원의 룩업테이블(look up table)을 생성한다. 온라인(on-line)에서 입력된 화소가 미리 구해진 룩업테이블의 피부영역에 속하면 피부로 정의하고 속하지 않으면 배경으로 정의하는 간단한 방법이다. 따라서, 계산속도가 상당히 빠르다는 장점이 있으나 조명변화가 있는 피부영역이나 피부색과 비슷한 칼라에는 검출률이 상당히 저조하다는 단점이 있다.

히스토그램 인터섹션 방법은 입력영상을 화소단위로 입력하지 않고 조각영상(patch image) 단위로 입력하여 피부영역에 속하는지 아닌지를 판단하게 된다. 히스토그램 인터섹션 방법은 오프라인에서 피부영상인 훈련영상을 이용하여 칼라 히스토그램을 생성한다. 온라인(on-line)에서 입력된 조각영상도 똑같이 칼라 히스토그램으로 생성한다. 생성된 칼라 히스토그램을 오프라인에서 구한 칼라 히스토그램을 가지고 정합(matching)되는 정도를 판단하여 임계값이상이면 피부영역으로 판단한다. 생성된 두 개의 칼라 히스토그램을 비교하기 위하여 대체로 식 (1)과 같은 정합점수(matching score)가 사용되었다[2-4].

$$M = \frac{\sum_{i,j} \min(H_{current}(i, j), H_{training}(i, j))}{\sum_{i,j} H_{current}(i, j)} \quad (1)$$

여기서, $H_{current}(i, j)$ 는 온라인에서 입력된 조각영상의 칼라 히스토그램을 의미하며, $H_{training}(i, j)$ 는 오프라인에서 생성한 훈련영상의 칼라 히스토그램을 의미한다. 입력된 조각영상은 식 (1)에 의한 정합점수가 임계값이상이면 피부영역으로 판단하고 이하이면 배경으로 판단한다. 히스토그램 인터섹션은 조각영상을 입력으로 한다는 점에서 화소단위의 룩업테이블 방법과 차이가 있다. 그러나 식 (1)의 정합점수는 결과적으로 피부색으로 판단된 각 색의 화소를 색상별로 단지 합한 것을 의미하므로 조각영상에 속한 화소들의 전체 상관관계를 고려한 것으로 볼 수는 없다. 히스토그램 인터섹션 방법은 룩업테이블 방법에 비하여 비교적 높은 검출률을 보이나 조명변화가 있는 칼라영상에는 여전히 검출률이 낮다.

그림 1의 (a)는 일정한 조명하에서 얻은 피부색을 HSI 칼라공간을 사용하여 칼라 히스토그램 분포를 나타낸 것이고, (b)는 조명 변화하에서 얻은 피부색에 대한 칼라 히스토그램 분포를 나타낸 것이다. 그림에서 보듯이 같은 피부색이라도 조명이 변화함에 따라 H와 S의 값이 다양한 분포를 갖게 되며 결과적으로 피부가 아닌 다른 색을 피부로 오검출을 하게 되는 이유가 된다[1].

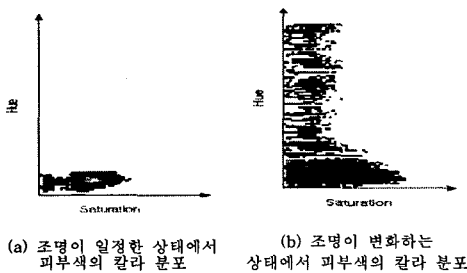


그림 1. HSI 칼라공간에서의 피부색의 칼라 히스토그램 분포

3. 제안된 방법

본 논문에서는 복잡한 배경이나 조명 변화가 심한 영상에서도 피부영역을 정확하게 검출할 수 있는 피부 영역 추출 방법을 제안한다. 제안된 방법은 그림 2에서 보는 바와 같이 오프라인(off-line) 훈련과정과 온라인(on-line) 검출과정의 두 단계로 나누어진다. 훈련단계에서는 다양한 환경에서 얻은 피부영상과 배경영상으로 구성된 훈련영상을 HSI 칼라공간으로 변환하고 훈련영상을 다차원의 열벡터로 표현하고 열벡터에 LDA(linear discriminant analysis)를 적용하여 새로운 특징벡터공간으로 선형변환(linear transformation)한다. 출어진 차원으로 선형변환된 특징벡터를 가지고 인덱싱 테이블을 생성하며 LDA 인덱싱 테이블(indexing table)이라 칭한다. 검출단계에서는 카메라로 들어온 칼라영상을 여러 개의 조각영상으로 나누고 각각의 조각영상에 대하여 LDA를 적용하여 선형변환된 특징벡터를 구한다. 구해진 특징벡터를 미리 생성한 LDA 인덱싱 테이블에서 검색하여 피부 영역을 검출한다.

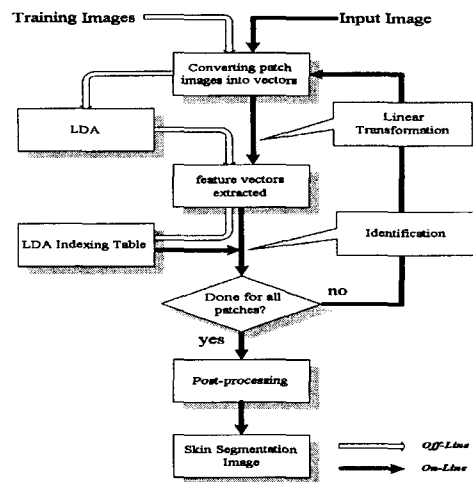


그림 2. 전체 시스템 순서도

3.1 입력영상

제안된 방법은 먼저 RGB 칼라공간의 입력영상을 크기가 $m \times n$ 인 여러 개의 조각영상으로 나눈다. 하나의 조각영상은 HSI 칼라공간으로 변환하면 HSI의 3개의 프레임(frame)으로 구성된다. 이 중에서 I 프레임은 조명 변화에 민감하므로 사용하지 않고 H와 S 프레임만을 사용한다. H와 S 프레임을 $2 \times m \times n$ 크기의 열벡터로 생성한다. 결국, 하나의 조각영상은 $2 \times m \times n$ 개의 원소를 갖는 열벡터로 변화되어 LDA의 입력으로 사용하게 된다. 4×4 크기의 하나의 조각영상을 열벡터화하는 과정을 그림 3에 나타내었다.

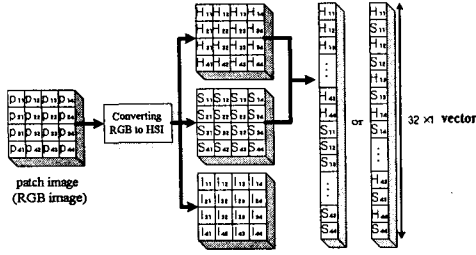


그림 3. 조각영상(4×4 크기)을 벡터로 변환하는 과정

3.2 LDA(Linear Discriminant Analysis)

LDA는 2개 이상의 클래스가 존재할 때 클래스간 분산(between-class covariance)과 클래스내 분산(within-class covariance)의 비(ratio)를 최대화하는 투영행렬(projection matrix)을 구하여 데이터를 선형변환을 취하면 클래스간의 분리가 최대로 일어나게 된다는 사실을 이용하는 패턴인식방법이다. LDA를 자세히 살펴보면 다음과 같다.

m 차원 벡터 공간의 열벡터 \vec{X}_{ij} 를 $m(m < n)$ 차원 특징 공간으로 매핑(mapping)하는 선형변환에 의한 새로운 특징 벡터 \vec{Y}_{ij} 는 식 (2)과 같이 정의된다.

$$\vec{Y}_{ij} = W^T \vec{X}_{ij} \quad (2)$$

여기서, \vec{X}_{ij} 와 \vec{Y}_{ij} 는 W 에 의한 선형 변환 이전 및 이후의 i 번째 클래스에 속한 j 번째 열벡터를 나타낸다. 클래스간 분산행렬을 식 (3)과 같이 정의하고

$$S_b = \sum_{i=1}^C (\vec{M}_i - \vec{M}) (\vec{M}_i - \vec{M})^T \quad (3)$$

클래스내 분산행렬을 식 (4)와 같이 정의한다.

$$S_w = \sum_{i=1}^C \sum_{j=1}^{N_i} (\vec{X}_{ij} - \vec{M}_i) (\vec{X}_{ij} - \vec{M}_i)^T \quad (4)$$

여기서, \vec{M}_i 는 i 번째 클래스의 평균, \vec{M} 은 전체 평균, C 는 클래스의 개수이며 N_i 는 i 번째 클래스의 데이터 개수이다. 클래스간 분산 행렬의 determinant와 클래스내 분산 행렬의 determinant의 비를 최대화하는 최적의 선형변환행렬 W_{opt} 를 식 (5)을 통해 계산하게 된다.

$$W_{opt} = \arg \max_W \frac{|W^T S_b W|}{|W^T S_w W|} \quad (5)$$

$$= [w_1 \ w_2 \ \dots \ w_m]$$

$\{w_i | i=1, 2, \dots, m\}$ 는 식 (6)에 구해지는 m 개의 가장 큰 고유치에 해당하는 고유벡터의 집합으로 구성된다. 여기서, m 은 대부분 $(C-1)$ 개의 0이 아닌 고유치를 가지므로 $(C-1)$ 이다.

$$S_b w_i = \lambda_i S_w w_i, \quad i=1, 2, \dots, m \quad (6)$$

3.3 피부 영역 검출

하나의 조각영상을 나타내는 열벡터는 LDA의 입력이 되며 LDA의 출력은 $(C-1)$ 개의 차원을 갖는 특징벡터가 된다. 훈련단계에서는 다양한 환경에서 얻은 피부 영상과 배경영상으로 구성된 훈련영상을 열벡터로 생성하여 LDA를 적용한다. 구해진 W_{opt} 에 의해 모든 훈련영상을 선형 변환된 특징벡터인 $(C-1)$ 차원의 열벡터들로 표현한다. 검출단계에서는 입력영상을 여러 개의 조각영상으로 나누고 조각영상에서 생성된 열벡터를

W_{opt} 에 의해 선형 변환된 특징벡터는 훈련단계에서 미리 구해진 특징벡터공간상에서 가장 가까운 특징벡터가 속하는 클래스로 검출하게 된다.

본 논문에서는 모든 훈련영상들로부터 계산된 특징벡터들과의 비교를 피하기 위하여 훈련단계에서 생성되는 특징벡터들을 LDA 인덱싱 테이블로 구성함으로써 획기적으로 계산시간을 단축하였다. LDA 인덱싱 테이블은 LDA를 통하여 선형 변환된 특징벡터의 첫 번째 원소와 두 번째 원소를 가지고 생성한다. 그 이유는 다음과 같다. 대부분의 패턴 인식에서는 식 (7)의 유클리디안(Euclidean) 거리를 사용하여 어느 클래스에 속하는지를 검증하게 된다.

$$d_{Euclidean}(C_i, \vec{y}) = \sqrt{(\Omega_1 - \gamma_1)^2 + (\Omega_2 - \gamma_2)^2 + (\Omega_3 - \gamma_3)^2 + \dots + (\Omega_{C-1} - \gamma_{C-1})^2} \quad (7)$$

여기서 C_i 는 i 번째 클래스 벡터이며 \vec{y} 는 온라인에서의 특징벡터이다. 그림 4에서 보듯이 선형 변환된 특징벡터는 클래스의 수에 무관하게 세 번째 원소부터는 거의 0으로 수렴하게 됨을 알 수 있다. 따라서, 식 (7)의 세 번째 항부터는 유클리디안 거리에 영향을 미치지 못하며 첫 번째와 두 번째의 원소만을 가지고 LDA 인덱싱 테이블을 구성해도 검출률에는 거의 변함이 없다.

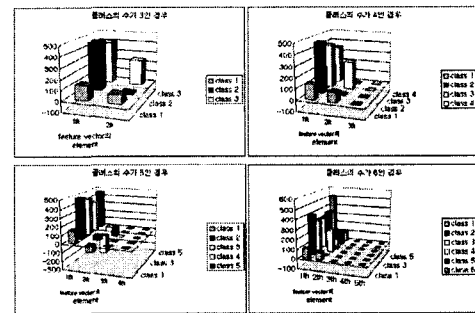


그림 4. 클래스 수와 특징벡터의 관계

검출과정에서는 유클리디안 거리를 이용하지 않고 특징벡터의 첫 번째와 두 번째 원소만을 가지고 훈련단계

에서 미리 생성된 LDA 인덱싱 테이블을 참조로 하여 입력된 조각영상이 어느 클래스에 속하는 지, 즉, 피부색에 해당하는 지 배경에 해당하는 지를 판단하게 된다.

4. 실험

본 논문에서 사용된 클래스의 수는 피부색을 나타내는 하나의 클래스와 배경을 나타내는 세 개의 클래스로 구성된다. 사용된 조각영상의 크기는 4×4 화소이며 각 클래스는 17424개의 훈련영상을 갖는다. 훈련영상으로 사용된 피부색은 여러 조명하에서 얻은 조각영상이다. 생성된 훈련영상인 조각영상은 32×1 열벡터로 변환되어 LDA를 적용한다. 선형변환된 특징벡터는 3×1 벡터가 되며 첫 번째와 두 번째 원소만을 가지고 LDA 인덱싱 테이블을 구성하며 그림 5에 나타내었다.

검출과정에서 입력영상은 320×240 화소이며 4800개의 조각영상으로 나누어진다. 하나의 조각영상은 선형변환되어 3×1 의 특징벡터로 표현된다. 훈련단계에서와 마찬가지로 첫 번째와 두 번째 원소를 가지고 미리 계산되어 있는 그림 5의 LDA 인덱싱 테이블과 비교하여 피부영역인지 아닌지를 구별하게 된다. 전체 입력 영상에 대하여 똑같은 방법을 적용하면 개략적인 피부영역을 검출하게 된다. 그러나 이러한 결과영상은 잡음 및 오검출로 인한 에러를 포함하고 있다. 1차적인 결과영상에 모폴로지(morphology) 연산을 취하고 피부영역의 8-이웃을 비교해 전부 배경이면 오검출(false detection)로 간주해 제거하였다. 그림 6은 입의의 조명하에서 얻은 영상에 대하여 본 논문에서 제안하는 방법으로 피부영역을 검출한 결과영상을 나타내었다.

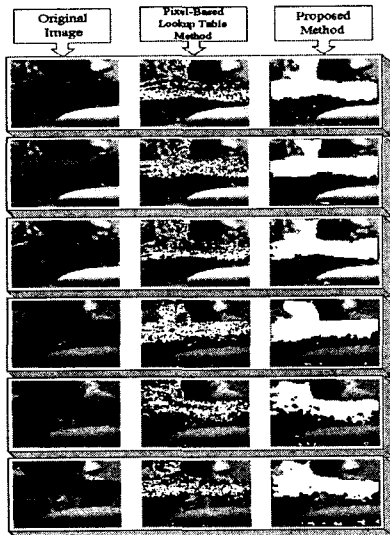


그림 6. 실험 결과

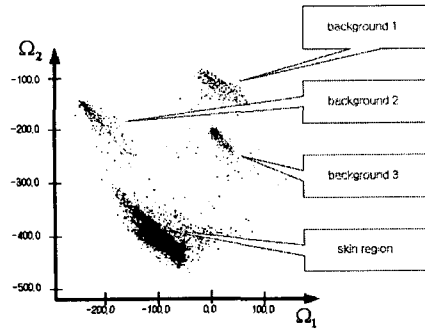


그림 5. 생성된 LDA 인덱싱 테이블

5. 결론

본 논문에서는 조명변화가 있는 영상에 강인한 피부영역 검출을 위하여 LDA 인덱싱 테이블을 이용한 피부영역 검출방법을 제안하였다. 제안된 방법의 장점은 다음과 같다. 첫째, 피부영역검출을 조각영상 단위로 처리하여 조각영상내의 화소값들의 상호관계를 고려하였다. 둘째, 훈련영상의 조각영상들로부터 계산되고 특징벡터를 LDA 인덱싱 테이블로 구성함으로써 모든 특징벡터들과의 비교를 피함으로써 획기적으로 계산시간을 단축할 수 있었다.

감사의 글

본 연구는 한국과학재단 '98 핵심전문연구(과제번호 981-0926-142-2) 지원으로 수행되었음.

참고문헌

- [1] M. J. Jones and J. M. Rehg, "Statistical Color Models with Application to Skin Detection," IEEE International Conference on Computer Vision and Pattern Recognition, Vol. 1, pp. 274~280, 1999.
- [2] M. J. Swain and D. H. Ballard, "Color Indexing," International Journal of Computer Vision. Vol 7, pp. 11~32, 1991.
- [3] S. Ahmad, "A Usable Real-Time 3D Hand Tracker," Conference Record of the Asilomar Conference on Signals, Systems and Computers, pp. 1257~1261, 1994.
- [4] D. Saxe and R. Foulds, "Toward Robust Skin Identification in Video Images," Proceedings of the 2th International Conference on Automatic Face and Gesture Recognition, pp. 379~384, 1994.
- [5] R. Kjellden and J. Kender, "Finding Skin in Color Images," International Conference on Automatic Face and Gesture Recognition, 1995.
- [6] J. Segen and S. Kumar, "Fast and Accurate 3D Gesture Recognition Interface," Proceedings of the 14th International Conference on Pattern Recognition. Vol. 1, pp. 86~91, 1998.