

점증적으로 증가하는 타원형 군집화: 피부색 영역 검출에의 적용

(Elliptical Clustering with Incremental Growth and its Application to Skin Color Region Segmentation)

이 경 미 ^{*}

(Kyoung-Mi Lee)

요약 본 논문에서는 군집화 알고리즘을 사용하여 피부색 영역을 분할하는 방법을 제안한다. 기존의 군집화 알고리즘들의 대부분은 주로 구형의 군집을 검출하고, 배치형으로 수행되며, 군집의 개수를 미리 정해야 한다는 문제점을 가지고 있다. 본 논문에서는 대표적인 타원형 군집화 알고리즘인 EM 알고리즘을 변형하여, 온라인으로 학습가능하며, 군집의 개수를 자동적으로 찾아낼 수 있는 EAM 알고리즘을 사용하였다. EAM 알고리즘의 유효성은 피부색 영역 분할에 대해 증명되었다. 실험결과를 군집의 개수가 미리 주어지지 않더라도, EAM 알고리즘은 주어진 영상에 대해 자동적으로 옳은 군집의 개수를 찾아냈고, EM 알고리즘과 비교하여 더 좋은 분할 결과를 보여주고 있다. 영역에 대한 조건부 확률을 이용하여 성공적인 피부색 영역의 탐지 및 분할 결과를 얻었다. 또한 사람이 포함된 영상을 분류하는 문제에도 적용하여 좋은 분류 결과를 얻었다.

키워드 : 타원형 군집화, EM 알고리즘, 온라인 학습, 자원할당 알고리즘, 피부색 영역 검출

Abstract This paper proposes to segment skin color areas using a clustering algorithm. Most of previously proposed clustering algorithms have some difficulties, since they generally detect hyperspherical clusters, run in a batch mode, and predefine a number of clusters. In this paper, we use a well-known elliptical clustering algorithm, an EM algorithm, and modify it to learn on-line and find automatically the number of clusters, called to an EAM algorithm. The effectiveness of the EAM algorithm is demonstrated on a task of skin color region segmentation. Experimental results present the EAM algorithm automatically finds a right number of clusters in a given image without any information on the number. Comparing with the EM algorithm, we achieved better segmentation results with the EAM algorithm. Successful results were achieved to detect and segment skin color regions using a conditional probability on a region. Also, we applied to classify images with persons and got good classification results.

Key words : Elliptical clustering, EM algorithm, On-line learning, Resource-allocating algorithm, Skin color region detection

1. 서론

최근 많은 연구자들이 다양한 분야에서 사람과 기계 사이의 발달된 상호작용을 개발하기 위해 노력하고 있다. 특히, 사람과 기계사이의 통신에서 사람의 얼굴이나 자세는 중요한 정보의 요소가 될 것이다. 따라서 얼굴이나 인체 검출은 얼굴 및 자세 인식 시스템의 기본 단계

로, 범인 인식 시스템, 보안 시스템으로부터 각종 보안 시스템, 신용 카드, 디지털 영화 편집 시스템, 내용 기반 영상 검색 시스템 등의 응용 분야에 적용될 수 있다. 이러한 광범위한 얼굴 및 자세 인식의 적용 분야에서 피부색 영역을 검출하는 것은 얼굴이나 인체의 초기 위치를 추정할 수 있게 한다. 분할된 영역은 얼굴 검출 및 인식, 손 분할 및 손동작 분석, 자세 및 동작 인식 등을 위하여 처리될 것이다.

일반적으로 주어진 영상에서 피부색 영역을 검출하는 문제는 화소에 대해 피부색인지 아닌지를 분류할 수 있는 조건을 제시하는 것이다. 이런 화소 단계 방법은 피

· 본 연구는 한국과학재단 목적기초연구(R04-2003-000-10092-0) 지원으로 수행되었음

* 정 회 원 : 덕성여대 컴퓨터과학부 교수

kmlee@duksung.ac.kr

논문접수 : 2004년 2월 2일

심사완료 : 2004년 6월 17일

부색 모델을 설계하고, 각 화소와 모델사이의 거리를 계산하여 피부색 모델에 얼마나 가까운지를 판단하게 된다. 피부색 판단의 중요한 요소가 되는 피부색 모델링에 대해 여러 방법들이 제안되었다[1]. 피부색 모델의 경계선이 조건식 등을 이용하여 명시적으로 주어질 수 있는데[2], 구현은 간단하지만 적절한 명시적인 조건을 찾는 데 어려움이 있다. 그래서 이러한 조건들을 훈련 데이터의 통계적 정보에 의해 구하는 방법이 연구되었다. Jones과 Regh는 피부색 히스토그램의 통계적 정보를 제공하여 자동 피부색 분할을 시도하였다[3]. 많이 이용되는 통계적 도구는 Look-Up 테이블[4], 조건부 확률[3,5], 또는 SOM[6] 등이 있다. 이런 방법은 비교적 간단하고 빠르지만, 좋은 모델을 구하기 위해서 많은 양의 훈련 데이터가 필요하고, 통계 자료를 보관할 추가적인 저장 공간이 요구된다. 몇몇 연구자들은 가우시안(Gaussian)에 기반을 둔 피부색 분포 모델을 이용하여 대용량 훈련 데이터 요구 문제를 해결하였다[7,8]. 그러나 이런 방법들의 치명적인 문제는 이론적인 정당성을 제시하지 못하는 모델에 대한 유효성이다. 지금까지 살펴 본 피부색 모델과 이것을 이용한 화소 단계에서의 피부색 영역 검출 방법은 대체로 구현이 쉽고 빠르지만, 조명 변화와 같은 잡음에 민감하다는 단점을 가지고 있다. 이 문제를 해결하기 위해서 영역 단계에서 피부색 영역을 검출할 수 있는데, 이는 좋은 영역 분할 알고리즘을 사용하여야 한다는 전제를 가지고 있다.

본 논문에서는 군집화 알고리즘을 이용하여 화소, 특히 피부색 화소를 그룹화하고 영상을 분할하였다. 다양한 종류의 군집화 알고리즘이 문헌상으로 소개되어 영상 분할에 사용되었지만, 군집화 알고리즘을 실제적으로 적용할 때는 여러 가지 문제점들이 야기될 수 있다. 우선, 이들 군집화 알고리즘은 주어진 데이터 집합에 대한 군집의 개수 J 를 미리 알아야 한다. 일반적으로 이런 방법은 군집의 개수가 정확하게 선택되었을 때만 유용하다. 너무 적은 J 로는 군집화를 잘 하기 어렵고, 너무 큰 J 는 일반화 성능을 떨어지게 한다. 따라서 J 에 대한 사전 정보가 없는 경우에 자동적으로 적절한 군집의 개수를 찾아내는 알고리즘이 절대적으로 요구되어 진다. 기존 군집화 알고리즘의 또 다른 문제점은 군집화 결과에 영향을 미치는 측정 방법의 선택이다. 많은 군집화 알고리즘이 유클리디안(Euclidean) 거리를 사용하는데, 이 방법은 구형의 군집을 찾는 데만 적당하다. 하지만, 실제로 대부분의 군집들은 구형이 아니다. 그래서 유클리디안 거리를 사용하는 군집화 알고리즘은 크고 길게 늘어진 형태의 군집을 찾는 데 적절하지 않다. 세 번째는, 공분산의 역행렬을 구할 때 특이성(singular) 또는 특이성에 근사하게 되는 조정(regularization) 문제이다.

이 문제를 해결하기 위해 공분산을 적절하게 초기화하거나[9], 타원형과 구형의 측정치를 결합하는 시도[10]가 있었다. 본 논문에서는 확률적 타원형 군집화 알고리즘인 EM 알고리즘[11]을 사용하였다. EM 알고리즘에 감소요소와 증가기준을 결부시켜 군집의 개수를 점증적으로 증가시킴으로써 군집의 개수를 미리 알려야하는 문제를 해결할 수 있다. 또한, 조정문제는 점차적(step-wise) 방정식으로 역행렬을 직접 갱신함으로써 해결될 수 있다.

본 논문에서는 점증적으로 증가하는 타원형 군집화 알고리즘에 의해 화소 단계에서 영역 군집화를 수행한 후, 구해진 영역에 대한 통계적 확률을 이용하여 영역 단계에서 피부색 영역을 검출하는 알고리즘을 제안한다. 2장에서 배치형 EM 알고리즘, 온라인 EM 알고리즘, 그리고 점증적으로 증가하는 온라인 EM(EAM) 알고리즘 등을 차례로 소개한다. 3장에서는 EAM 알고리즘에 의한 영역 분할과 분할된 영역에 대한 피부색 영역 검출 과정을 설명한다. 피부색 영역 검출과 피부색 검출에 의한 영상 분류에 대한 실험 결과는 4장에 나타나고, 5장에서는 결론을 기술한다.

2. 점증적으로 증가하는 타원형 군집화

본 장에서는 영상 분할을 위해 사용되는 군집화 알고리즘에 대해 소개한다. 군집화 문제는 패턴 인식부터 컴퓨터 시각까지 다양한 영역에서 중요한 역할을 수행한다. d -차원의 공간에서 N 개의 데이터가 주어지면, $X = \{x_n | n=1, \dots, N\}$, 군집화 알고리즘은 같은 군집 내 데이터가 높은 유사성을 가지도록 하는 반면, 다른 군집에 속한 데이터들의 비유사성이 높도록 한다. 군집화 알고리즘의 목표는 주어진 데이터를 J 개의 군집 $\{c_j | j=1, \dots, J\}$ 중 하나에 할당하는 것으로, 비슷한 특성을 가진 데이터들이 같은 군집에 포함되도록 한다. 각 군집 c_j 은 N_j 개의 데이터를 포함하고, 중심 μ_j 과 공분산(covariance) 행렬 Σ_j 의 두 매개변수로 표현될 수 있다.

2.1 타원형 군집화: EM 알고리즘

대표적인 타원형 군집화 알고리즘인 EM 알고리즘은 기대(expectation: E-단계)와 최대화(Maximization: M-단계)의 두 단계로 구성된다. E-단계는 비관측 변수의 분포를 찾고, M-단계는 이들에 대한 최대 우도(maximum likelihood)를 측정한다. 두 단계는 반복되면서 정(true) 우도를 향상시키며, 이미 국부 최대에 도달했다면 변하지 않고 그대로 나뉜다.

EM 알고리즘은 군집의 최대 우도 매개변수 $\theta = \{\mu_j, \Sigma_j\}$ 에 대한 초기 가정으로 시작하고, E-단계와 M-단계를 반복적으로 적용함으로써 계속적으로 추정한다(알

고리즘 I). θ 을 이용하는 입력 데이터 x^n 와 군집 c_j 에 대한 결합(joint) 확률이 $P(x^n, j | \theta)$ 라고 하면, x^n 에 대한 주변(marginal) 확률은 $P(x^n | \theta) = \sum_j P(x^n, j | \theta)$ 이다. 만일 일반화된 지수함수를 사용한다면,

$$P(x^n, j | \theta) = \exp\left[-(x^n - \mu_j) \chi_{\Sigma_j}^{-1} (x^n - \mu_j)'\right]$$

$\bar{\theta}$ 이 현 매개변수의 추정치라 하자. E-단계는 아래 Bayes규칙에 따라 c_j 의 사후(posterior) 확률을 계산한다.

$$P(j | x^n, \bar{\theta}) = \frac{P(x^n, j | \bar{\theta})}{\sum_{i=1}^N P(x^n, i | \bar{\theta})} \quad (1)$$

식 (1)을 이용하여, M-단계는 전체 데이터에 대한 기대 로그-우도를 정의한다.

$$l(\theta | \bar{\theta}, X) = \sum_{n=1}^N \sum_{j=1}^J P(j | x^n, \bar{\theta}) \log P(x^n, j | \theta) \quad (2)$$

EM 알고리즘은 국부 최댓치에 도달할 때까지, $l(\theta | \bar{\theta}, X)$ 의 증가를 보장한다. 이러한 보장 방식은 EM 알고리즘이 다 천천히 수렴하도록 한다. M-단계는 [12]에 따라 매개변수를 갱신한다.

$$\mu_j = \frac{\sum_{n=1, x^n \in c_j} x^n P(j | x^n, \bar{\theta})}{\sum_{n=1, x^n \in c_j} P(j | x^n, \bar{\theta})} \quad \text{과} \quad \Sigma_j = \frac{\sum_{n=1, x^n \in c_j} m_j m_j' P(j | x^n, \bar{\theta})}{\sum_{n=1, x^n \in c_j} P(j | x^n, \bar{\theta})} \quad (3)$$

여기서 $m_j^n \equiv x^n - \mu_j$ 이고(')는 전치행렬을 의미한다.

μ_j 를 임의로 초기화 한다.
 Σ_j 를 단위행렬로 초기화 한다.
 $t = 1$.
 While $\exists j: c_j(t) \neq c_j(t-1)$
 E-단계: 모든 데이터에 대해 x^n ,
 식 (1)을 이용하여 최근접 군집 c_j 을 찾는다.
 군집의 데이터 개수를 증가시킨다. $N_j = N_j + 1$.
 M-단계: 모든 군집에 대해 c_j ,
 식 (2)을 최대화하고, 식 (3)을 이용하여 μ_j 과 Σ_j 을 구한다.
 $t = t + 1$.

알고리즘 I. EM 알고리즘

2.2 온라인 EM 알고리즘

전 장에서 살펴 본 EM 알고리즘은 배치(batch)형 알고리즘이다. 즉, E-단계와 M-단계는 완전한 데이터 집합에 대해 전체적으로 수행된다. 그러나 데이터가 계속 해서 도착하는 경우에는 점증적 학습을 이용한 적응적 학습이 적당할 것이다. 최근, EM에 대한 점증적 접근이 Nowlan[13]에 의해 연구되었고, Neal과 Hinton[14]에

의해 증명되었다. Backer 와 Scheuners[15], 그리고 Sato와 Ishii[16]은 타원형 군집화를 위한 온라인, 순차 알고리즘이 소개하였다. 이들 온라인 EM 알고리즘은 군집화를 주어진 군집 개수로 초기화하고, 데이터가 도착할 때마다 타원형 군집을 갱신한다. 배치 알고리즘은 잘 구분된 국부 최적으로 빠질 확률이 높은 반면에, 온라인 알고리즘은 한 번의 반복이 갱신과정을 보다 작게 만들어 최적해 근처로 빨리 접근하게 된다[17]. 다음은 [16]에 기반을 둔 온라인 EM 알고리즘이다.

$\theta^n \equiv \{\mu_j^n, \Sigma_j^n\}$ 이 n 번째 관측자 x^n 후의 추정치라 하자. 온라인 EM 알고리즘에서 step-wise 방정식을 사용하여, 식 (3)을 아래와 같이 고칠 수 있다.

$$\mu_j^n = \frac{f_j^{n-1}(x) + \eta^n [x^n P_j^n - f_j^{n-1}(x)]}{f_j^{n-1}(1) + \eta^n [P_j^n - f_j^{n-1}(1)]} \quad \text{과}$$

$$\Sigma_j^n = \frac{f_j^{n-1}(m_j m_j') + \eta^n [m_j^n m_j^n' P_j^n - f_j^{n-1}(m_j m_j')]}{f_j^{n-1}(1) + \eta^n [P_j^n - f_j^{n-1}(1)]} \quad (4)$$

여기서 $m_j^n \equiv x^n - \mu_j^n$ 이고, $P_j^n \equiv P(j | x^n, \theta^{n-1})$ 이다. 감소 요수 λ^n 는 n 에 대한 감소함수로 $0 \leq \lambda^n \leq 1$ 이므로, 학습 매개변수 $\eta^n = \left(1 + \frac{\lambda^n}{\eta^{n-1}}\right)^{-1}$ 는 $\frac{1}{n} \leq \eta^n \leq 1$ 이다. 또한, $f_j^{n-1}(x) = f_j^{n-2}(x) + \eta^{n-1} [x^{n-1} P_j^{n-1} - f_j^{n-2}(x)]$ 이다. 마찬가지로 $f_j^{n-1}(1)$ 와 $f_j^{n-1}(m_j m_j')$ 도 재귀적으로 정의된다.

모든 군집에 대해 j ,
 μ_j^0 을 임의로 초기화한다.
 Σ_j^0 을 단위행렬로 초기화한다.
 $n = 1$.
 각 데이터 x^n 에 대해,
 E-단계: 식 (1)을 이용하여 최근접 군집 c_j 을 찾는다.
 M-단계: 식 (2)을 최대화시키고, 식 (4)을 이용하여 μ_j^n 과 Σ_j^n 을 갱신한다.
 $N_j = N_j + 1$.
 $n = n + 1$.

알고리즘 II. 온라인 EM 알고리즘

λ^n 을 적절하게 선택하면, 식 (4)은 식 (3)과 동일하다 [5]. 같은 입력 데이터가 반복적으로 제시된다고 하자. 즉, $x^{n+(t-1)N} = x^n$ 이고, 전체 데이터 집합의 반복이 끝날 때마다, $n = N$, θ 가 갱신된다. 만일 λ^n 이 아래와 같이 정의된다면,

$$\lambda^n = \begin{cases} 0 & \text{if } n = (t-1)N + 1 \\ 1 & \text{otherwise,} \end{cases} \quad (5)$$

η^n 는 $\eta^n = \eta^{(t-1)N+n} = 1/n$ 으로 정의된다. 각 반복이 시작될 때마다, $f_j^{n-1}(x)$ 는 $f_j^0(x) + \eta^n[x^1 P_j^1 - f_j^0(x)] = x^1 P_j^1$ 으로 초기화된다. 전체 데이터의 반복이 끝날 때, $f_j^{(t-1)N}(x) = f_j^N(x)$ 는 $\bar{\theta} = \theta^{t-1}$ 에 대해 만족한다. $f_j^{n-1}(1)$ 과 $f_j^{n-1}(m_j; m_j)$ 도 마찬가지로 적용된다. 따라서 λ^n (식 (5))를 가진 온라인 EM 알고리즘은 배치 알고리즘과 동일하다.

2.3 점증적으로 증가하는 EM 알고리즘

군집화 알고리즘의 주요 목적은 데이터에 대한 정확한 적용이 아니라, 주어진 군집의 개수 J 에 데이터를 일반화할 수 있는 통계적 과정을 설계하는 것이다. 따라서 적절한 군집 개수의 선택은 군집화 결과에 상당히 영향을 끼친다. 그러나 많은 경우에 적절한 군집의 개수를 결정하는 것은 쉽지 않다. 그림 1 은 세 개의 타원형 군집의 예를 보여주고 있다. 만일 최적의 개수보다 적은 수가 주어지면($J=2$), 배치나 온라인 EM 알고리즘은 서로 다른 군집을 합쳐서 불충분 군집화를 야기한다. 또한, 큰 군집 개수($J=4$)는 하나의 군집을 여러 군집으로 나누어, 초과 군집화를 발생시킨다. 이런 문제를 해결하기 위해서, 고정되지 않은 과정의 온라인 설계에 적절한 자원 할당(Allocation)을 결합시킨 EAM 알고리즘이 소개됐다[5]. 새 데이터가 나타날 때마다, 데이터 증가 임계치에 기반을 두어 존재하는 군집의 매개변수를 갱신하거나 새로운 군집을 추가한다. 군집들은 이전에 군집된 데이터를 반영하고 현재의 상태를 포함한다.

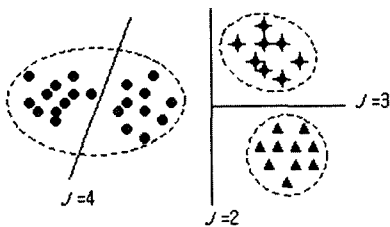


그림 1 EM 알고리즘에 의한 세 군집의 예($J=3$). 보다 적은 개수($J=2$)는 서로 다른 군집을 합치고, 보다 큰 개수($J=4$)는 하나의 군집을 나눈다.

EAM 알고리즘은 적은 군집의 개수로 시작하거나 ($J_0 \leq J$), 초기 군집 없이($J_0 \leq 0$) 시작할 수 있다. 초기 군집의 개수가 주어지면, 알고리즘은 EM 또는 온라인 EM 알고리즘과 같은 방법으로 J_0 개의 군집을 초기화한다. 중심 μ_j^0 은 임의로 정해지고, 공분산 Σ_j^0 은 단위 행렬로 초기화 된다. 데이터에 대한 정보가 없다면, 알고리즘은 군집을 포함하지 않는다. 새 데이터 x^n 가 올

때마다, EAM 알고리즘은 우선 존재하는 군집들과 비교하고, 식 (1)을 이용하여 최근접 군집을 찾는다. 증가기 준에 따라, 알고리즘은 임계치 ϵ^n 보다 큰 근접 군집에 대한 사후 확률을 계산한다.

$$P(\text{nearest} | x^n, \bar{\theta}) > \epsilon^n \tag{6}$$

임계치는 큰 값으로 시작되고($\epsilon^0 = \epsilon_{max}$), 최소값 ϵ_{min} 에 도달될 때까지 감소요소 γ 가 곱해진다($0 < \gamma < 1$). 임계치 ϵ^n 는 반복마다 달라져서, 큰 임계치 ϵ^n 를 가진 군집들이 처음부터 증가되어 진다. 이는 전체적으로 군집화를 안정되게 한다. 즉, 처음엔 더 많은 데이터가 보이며, 점점 ϵ^n 가 줄어들어 결국 전체 군집화는 세밀하게 된다.

식 (6)은 데이터 x^n 가 존재하는 어떤 군집과도 비슷하지 않고, 존재하는 군집들로 적절하게 표현되기 어렵다는 걸 입증한다. 따라서 식 (6)을 만족하면, 새 군집은 x^n 으로 만들어지고, 군집의 개수 J 가 증가된다. 새로 할당된 군집의 중심이 입력 데이터로 설정되고, 공분산은 존재하는 최근접 군집에서 새 데이터에 대한 사후 확률에 비례적으로 결정된다. 새 군집은 다음과 같이 결정된다[18,19]:

$$\begin{aligned} \mu_{j+1}^n &= x^n, \\ \Sigma_{j+1}^n &= K(x^n - \mu_{nearest}^n)(x^n - \mu_{nearest}^n) \end{aligned} \tag{7}$$

여기서 $\mu_{nearest}^n$ 은 x^n 로부터 존재하는 최근접 군집의 중심이고, K 는 함수의 평활도를 결정하는 겹침(overlap) 요소에 대한 상수이다. 이는 새 데이터가 새로 만들어진 군집에 보다 잘 대응되도록 한다. 따라서 데이터들을 군집화한 후에 각 군집은 데이터 공간에서 서로 인접한 데이터의 군집을 나타내게 되어, 마침내 EAM 알고리즘은 데이터 군집의 적절한 개수를 찾는다.

새 데이터 x^n 가 존재하는 군집 중 하나와 유사하다면, M-단계는 식(4)을 이용하여 중심과 공분산을 갱신한다. 군집의 중심은 존재하는 데이터를 향해 움직이는 반면, 공분산 행렬은 그 데이터에 의해 표현되는 공분산을 향해 움직인다. 그러나 이런 갱신은 $(\Sigma_j^n)^{-1}$ 이 존재하는 경우에만 가능하다. 불행하게도, 실제로 역행렬은 특이(singular) 행렬이 될 수 있다. d -차원 공간에서 $(\Sigma_j^n)^{-1}$ 의 비특이(non-singularity) 행렬을 보장하기 위해서, 조정된 공분산 행렬은 step-wise 방정식에 의해 정의될 수 있다.

$$\Sigma_{jR}^n = \Sigma_j^n + \beta \frac{\Sigma_j^n}{d} I_d \tag{8}$$

여기서 β 는 상수이고 ($0 < \beta < 1$), I_d 는 d -차원 단위행렬이다. 그리고 반복될 때 마다 $(\Sigma_{jR}^n)^{-1}$ 의 역행렬을 계산하는 것은 비효율적이므로, 역행렬은 직접 갱신되어

저야 한다. 표준이 되는 방법 [20]을 이용하여 역행렬은 직접 갱신되고, 이는 step-wise 방정식으로 구할 수 있다. $(\sum_{j \in R}^n)^{-1}$ (식 (11))을 직접 갱신하기 위한 자세한 유도는 본 논문의 끝에 있는 부록에서 소개되고 있다. EAM 알고리즘은 알고리즘 III에 요약되었다. 새로운 군집이 더해진 후, 일정 시간 동안 존재하는 매개변수를 갱신하고, 다른 할당을 허용하지 않는 것이 유용하다.

초기 군집에 대해 $j = 1, \dots, J_0$
 μ_j^0 을 임의로 초기화한다.
 \sum_j^0 을 단위행렬로 초기화 한다.
 $\epsilon^0 = \epsilon_{\max}$.
 $J = J_0$.
 $n = 0$.
 모든 데이터에 대해 x^n
 E-단계: 식 (1)을 이용하여 최근접 군집 c_j 을 찾는다.
 식 (6)을 만족하면, 식 (7)을 이용하여 새 군집을 할당한다. (A-단계)
 $N_{j+1} = N_{j_1} + 1$.
 $J = J + 1$.
 만족하지 않으면, 식 (2), (4)와 (8)을 이용하여 μ_j^n 와 \sum_j^n 을 갱신한다.(M-단계)
 $N_j = N_j + 1$.
 만일 $\epsilon^n > \epsilon_{\min}$ 이면, $\epsilon^{n+1} = \gamma \epsilon^n$.
 $n = n + 1$.

알고리즘 III. 점증적으로 증가하는 EAM 알고리즘

점증적으로 증가하는 EAM 알고리즘은 적절한 J 와 ϵ_{\max} 의 선택으로 온라인 EM 알고리즘과 동일함이 증명되었다 [5]. $\gamma \approx 1$ 이면, ϵ_{\max} 이 모든 군집에 대해 가장 큰 사후 행렬로 정의된다고 하자. 초기 군집의 개수가 최적($J_0 = J$)으로 주어진다면, 이는 제안된 EAM 알고리즘이 단순히 A-단계를 무시함으로써 온라인 EM 알고리즘과 동일함을 의미한다. 만일 J_0 이 더 적은 개수로 주어진다면($0 \leq J_0 < J$), 어떤 μ_j 이 군집 c_j 에 첫 번째 유사한 데이터로 초기화될 수 있고, 온라인 EM 알고리즘은 알고리즘의 처음에 임의로 모든 군집을 초기화한다. 결론적으로 EAM 알고리즘의 기대치는 최대 우도 기대치로 수렴된다.

3. 피부색 영역 분할에의 적용

본 장에서는 EAM 알고리즘을 피부영역 분할 문제에 적용하려 한다. 우선, 입력 영상의 잡음을 제거하기 위해 미디언 필터(median filter)를 이용하여 RGB 색공간에서 평활화시킨 후, EAM 알고리즘을 이용하여 주어진 영상을 화소 단계에서 군집화를 수행한다. 군집화로 얻어진 영역에 대해 피부색 확률을 구해 영역 단계에서의

피부색 영역을 검출하려 한다.

화소 단계 처리로 EAM 알고리즘과 같은 군집화 알고리즘을 적용하기 전에 영상 표현의 문제를 고려하여야 한다. 영상 표현이란 입력 영상에서 특징(feature) 영상의 집합으로의 변환으로 생각할 수 있다. 군집화에서 어떤 특징 표현법을 사용하느냐는 대부분 영상의 성질이나 분할의 대상 또는 목표에 의해 결정된다. 본 논문은 피부색 영역 분할을 목표로 하기 때문에, 피부색 검출을 위한 색공간인 RGB, YCbCr, HSV, CIE Lab, CIE Luv, 정규화된 RGB 등을 사용할 수 있는데, 광도를 분리할 수 있는 YCbCr와 HSV 색상 모델을 이용하였다. 사용된 YCbCr와 HSV 모델은 RGB 색상 영상의 밝기 값 보정을 통하여 외부 조명의 영향을 줄여주는 전처리 과정을 거친 보정된 RGB 영상을 변환시킨 YCbCr와 HSV 색상영상을 생성하였다. 그러면, EAM 알고리즘은 주어진 영상을 6개의 색상정보로 이루어진 벡터가 화소의 개수만큼 있는 데이터 집합으로 간주하고 군집화를 수행하게 된다.

화소 단계에서 EAM 알고리즘에 의해 주어진 영상을 분할한 후, 피부색 확률모델을 이용하여 영역 단계에서 피부색 영역을 찾아낸다. 피부색에 대한 통계학적 모델은 1000개의 영상을 포함하는 영상 데이터베이스로부터 피부색 화소의 히스토그램을 학습하여 만들어진다[3]. 이 때, 히스토그램은 HSV와 YCbCr 색공간을 k -평균(means) 알고리즘을 사용하여 각각 64개의 색군집으로 양자화하여 만든다. 색공간에서 피부색에 대한 조건부 확률은 피부색 히스토그램 내 각 색의 화소 수를 전체 피부색 히스토그램 화소 수로 나누어서 구할 수 있다. 따라서 각 화소 x 의 피부색에 대한 조건부 확률은 두 색공간 내 확률의 평균으로 한다.

$$P(x|skin) = \frac{1}{2} \left\{ \frac{Y(x)}{\sum_x Y(x)} + \frac{H(x)}{\sum_x H(x)} \right\} \quad (9)$$

여기서 $Y(x)$ 는 YCbCr 색공간에서의 피부색 히스토그램을, $H(x)$ 는 HSV 색공간에서의 피부색 히스토그램을 나타낸다. 그러면, 주어진 영상에서 각 화소의 확률이 임계치를 기준으로 피부색 화소와 그렇지 않은 화소로 분류함으로써 피부색 영역을 분할할 수 있다. 이와 같은 화소 단위의 분류는 주어진 영상에 내재된 분포의 복잡도에 관계없이, 쉽게 확률 모델을 구할 수 있다는 장점을 갖는다. 그러나 화소 단위의 분류는 피부색과 비슷한 배경을 포함할 수 있기 때문에 여전히 불충분하다.

이 문제를 해결하기 위해서 화소 단위 대신에 영역 단위로 접근할 수 있다. 분할된 영역 R 에 대한 피부색 조건부 확률은 R 의 각 화소의 확률을 평균하여 계산한다.

$$P(R|skin) = \frac{\sum_{z \in R} P(x|skin)}{Area(R)} \quad (10)$$

여기서 $Area(R)$ 은 영역 R 의 면적을 의미하고, 면적은 영역 내 전체 화소의 개수로 한다. 확률이 임계치 이상인 영역을 피부영역으로 한다.

4. 실험 결과

4.1 피부색 영역 검출

제안된 알고리즘은 운영체제 Windows 2000XP 상에서 JAVA를 이용하여 구현되었으며, 실험은 Pentium-IV 1.8GHz인 CPU와 512 MB 메모리 사양의 Windows 2000 XP 환경에서 수행하였다. 본 논문에서는 다양한 인종, 사람 수, 조명 환경을 포함하는 200개의 영

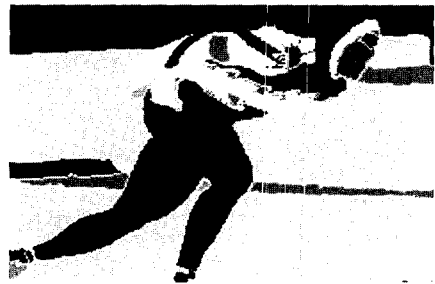
상을 인터넷에서 얻어 실험에 사용하였다. 화소 단계에서 EAM 군집화 알고리즘을 적용해서 영상을 분할하기 위해서 다음과 같은 값을 이용하여 EAM을 실행시켰다:

$d = 6, \epsilon_{max} = 0.7, \epsilon_{min} = 0.07, v = 0.999, \kappa = 0.9,$
그리고 $\beta = 0.3$.

그림 2는 비교실험을 위하여 주어진 영상(그림 2(a))에 대해 EAM 알고리즘(그림 2(b))과 함께 EM 알고리즘(그림 2(c), (d)와 (e))을 적용시킨 결과를 보여준다. 2장에서 언급했듯이, EM 알고리즘은 적절한 J 가 주어졌다고 가정한다. 즉, 그림 2(c), (d)와 (e)는 각각 군집의 개수 $J = 6, 8$ 과 10 으로 얻어진 결과 영상들이다. 반면, EAM 알고리즘은 초기 군집 없이($J_0 = 0$) 시작해서, 최적의 군집 개수($J = 8$)를 찾아 영역을 구분하였다. 여기서 의상과 같은 색을 가진 배경을 같은 영역으로 간주



(a) 원래 영상



(b) EAM 알고리즘 사용



(c) EM 알고리즘 사용: J = 6

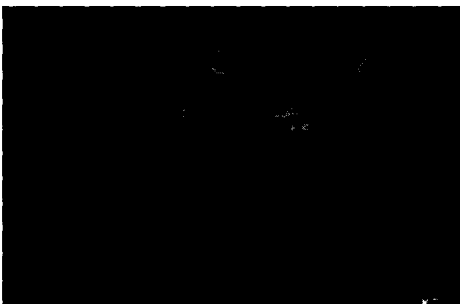


(d) EM: J = 8



(e) EM: J = 10

그림 2 영상 분할 결과



(a)



(b)

그림 3 피부색 영역 분할 결과: (a) 그림 2(a)에 대한 화소 단위 분할과 (b) 영역 단위 분할

하였다. 옳은 군집의 개수로 EM 알고리즘을 수행한 분할 결과(그림 2(d))와 비교하면 EAM 알고리즘은 잡음 조각이 비교적 적었고 적은 분할 오류를 획득했다.

두 군집화 알고리즘 모두가 데이터의 출현 순서에 의존하므로, 데이터의 출현 순서에 따라 영상 분할 결과가 달라질 수 있다. 따라서 본 장에서는 입력 영상에 대한 20개의 서로 다른 순서의 화소 집합을 두 알고리즘에 적용하였고, 그림 2(a)은 EAM 알고리즘을 이용하여 평균 92.35%로, $J = 8$ 인 EM 알고리즘을 이용하여 평균 86.25%로 옳게 분할되었다.

그림 3은 그림 2(a)에 대한 피부색 영역 분할 결과를 보여주고 있다. 군집화 알고리즘을 수행하지 않고, 식

(9)을 이용한 화소 단위로 분할한 결과(그림 3(a))는 옷이나 배경의 붉은 화소들을 함께 분할하고 그늘진 턱부분을 포함시키지 못하는 반면, EAM 군집화 알고리즘과 식 (10)을 이용한 영역 단위 분할 결과(그림 3(b))는 비교적 피부 영역을 잘 검출하였다.

그림 4는 EAM 군집화 알고리즘 적용 결과와 피부색 영역 분할 결과의 일부를 보여주고 있다. 200개의 실험 영상에 대해 20개의 서로 다른 순서의 화소 집합에 대한 평균 영상 분할률과 평균 피부색 영역 검출율이 표 1에 주어진다. 영상 분할률은 두 군집화 알고리즘을 적용하여 옳게 군집화된 화소의 비율을 나타내며, 피부색(화소 또는) 영역 검출율은 피부색 영역으로 옳게 검출



그림 4 피부영역 분할 결과: (a) 원래 영상, (b) EAM을 이용하여 분할된 영상, (c) 피부색 확률 모델을 이용하여 분할된 피부색 영역

표 1 실험 영상에 대한 평균 영상 분할률과 피부색 영역 검출률

	화소 단계				화소 + 영역 단계	
	식 (9)	[3]*	[6]*	[7]*	EM + 식(10)	EAM + 식(10)
평균 영상 분할률					84.65%	90.45%
평균 피부색 영역 검출률	85.75%	90%	90%	78%	86.15%	92.35%

* 저자에 의해 발표된 피부색 검출 결과

된 화소의 비율을 말한다. 표 1은 실험영상에 대해 EAM 알고리즘을 이용하여 화소와 영역을 모두 고려한 피부색 영역 검출이 식 (9)만을 이용한 화소 단계 방법과 EM 알고리즘을 이용한 검출 방법 보다 좋은 결과를 획득하였음을 보여주고 있다.

4.2 피부색 검출에 의한 영상 분류

EAM 군집화 알고리즘과 피부색에 대한 영역 단위 조건부 확률을 이용한 피부색 영역 검출 및 분할 알고리즘은 사람이 들어 있는 영상을 찾는 것과 같은 영상 분류에 사용될 수 있다. 사람을 검출하기 위한 여러 복잡한 방법을 사용하는 대신에, 여기서는 피부색 영역 검출만을 이용하여 사람을 포함하는 영상을 분류하려 한다. 우선, 각 영상에 대해 Jones & Rehg의 논문[3]에서 사용된 5가지 특징 중 아래 4개의 특징만을 구한다:

- 피부색 영역들의 면적 비율
- 피부색 영역의 평균 조건부 확률
- 가장 큰 피부색 영역의 크기

• 피부색 영역의 개수

피부영역의 평균 밀집도와 함께 5개의 특징은 3계층 RBF 신경망의 입력으로 주어지고, 이 신경망은 영상이 사람을 포함할 가능성을 출력한다. 피부색에 대한 통계학적 모델에 사용된 1000개의 영상과 WBIS[21]에 사용된 10,000개의 영상 중 사람을 포함하지 않은 500개의 영상을 선택하여 신경망을 학습시켰다. 10,000개의 영상 중 학습에 포함되는 않은 영상에 대해 사람을 포함한 200개의 영상과 그렇지 않은 105개의 영상으로 얻어진 분류 결과가 표 2에 보여진다. 결과는 Jones & Rehg 방법보다 좋은 분류 결과를 보여주고 있으며, 특히 사람이 포함된 영상들은 훨씬 잘 분류해 냈다(그림 5(a)). 그러나 여전히 남아있는 문제는 사람이 포함되지 않았지만, 피부색과 유사한 색을 가진 배경이나 물체를 가진 영상에 대한 분류이다(그림 6(a)). 사람이 포함되었어도 음영이나 의상 등에 의해 피부색 노출이 안 된 경우 역시 제대로 분류해 내지 못했다(그림 6(b)). 이는

표 2 사람을 포함한 영상에 대한 분류율: Jones & Rehg의 방법[3]와 비교

분류		제안된 방법	Jones & Rehg
실제	예측 결과		
사람 포함	사람 포함	87.5% (175/200)	83.2%
사람 없음	사람 없음	70.5% (74/105)	71.3%
평균 분류율		81.6% (249/305)	77.5%



(a) 사람이 포함되었다고 분류된 영상

(b) 사람이 포함되지 않았다고 분류된 영상

그림 5 옳은 분류 결과 영상 예



(a) 사람이 포함되었다고 분류된 영상

(b) 사람이 포함되지 않았다고 분류된 영상

그림 6 잘못된 분류 결과 영상 예

텍스처나 형태와 같은 보다 복잡한 접근을 통해 해결될 수 있다. 그림 5와 6은 각각 옳게 분류된 결과 영상과 잘못 분류된 결과 영상을 보여주고 있다.

5. 결론

본 논문에서는 군집화 알고리즘을 이용하여 화소 단계에서 영상을 분할하고, 분할된 영역을 피부색 영역으로 분류하는 피부색 영역 검출 알고리즘을 제안하였다. 일반적으로 화소 단계처리는 피부색과 비슷한 배경 등의 잡음에 약한 반면, 영역 단계처리는 좋은 영역 분할 알고리즘의 사용을 전제로 한다. 본 논문에서는 점층적으로 증가하는 타원형 알고리즘인 EAM 알고리즘을 이용하여 영역 분할을 수행하였다. 대부분의 군집화 알고리즘이 옳은 군집의 개수를 미리 요구하는 반면, EAM은 주어진 데이터(영상)에 대해 군집의 개수를 자동적으로 찾아내기 때문에, 데이터에 적합한 분할을 수행할 수 있다. 본 논문의 4장에서 EM 알고리즘을 사용한 군집화의 결과와 피부색 영역 검출 결과를 비교하여 보면, 적절한 군집화가 결과에 많은 영향을 미치고 있음을 실험적으로 보여주었다. 제안된 알고리즘은 사람을 포함한 영상 분류에도 적용하여 평균 81.6%의 결과를 얻었다.

제안된 방법은 몇 가지 방향으로 연구를 확장시키나 갈 수 있다. 첫째, EAM 알고리즘이 입력 공간에서 군집의 중심이 밖에 있는 군집이나 군집의 확률이 무시할 만큼 작은 군집을 잘라낼 수 있도록 한다. 이는 또한 초과 군집화 문제($J_0 > J$)를 해결할 수 있다. 둘째, 색상만을 이용하는 피부색 확률을 확장하여, 영역의(형태적인) 특징을 포함할 수 있는 측정 방법에 대한 연구가 필요하다. 본 논문에서 제안하는 알고리즘은 영상에서 사람 검출과 자세 인식에 관한 연구의 일부가 된다.

참고 문헌

- [1] V. Vezhnevets, V. Sazonov, and A. Andreeva, A Survey on Pixel-Based Skin Color Detection Techniques, In *Proc. Graphicon-2003*, pp. 85-92, 2003.
- [2] P. Peer, J. Kovac, and F. Solina, Human skin colour clustering for face detection. In *Proc. of International Conference on Computer as a Tool*, 2003.
- [3] M.J. Jones and J.M. Reng, Statistical color models with application to skin detection, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1:274-280, 1999.
- [4] L. Sigal, S. Sclaroff, and V. Athitsos, Estimation and prediction of evolving color distributions for skin segmentation under varying illumination. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 2, 152-159, 2000.
- [5] 이경미, 점층적으로 증가하는 타원형 군집화, 덕성여대 자연과학논문집, 10권, 135-147, 2004.
- [6] D. Brown, I. Craw, and J. Lewthwaiye, A SOM based approach to skin detection with application in real time systems. In *Proc. of the British Machine Vision Conference*, 2001.
- [7] J. Y. Lee and S. I. Yoo, An elliptical boundary model for skin color detection. In *Proc. of the 2002 International Conference on Imaging Science, Systems, and Technology*, 2002.
- [8] J.-C. Terrillon, M.N. Shirazi, H. Fukamachi, and S. Akamatsu, Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images. In *Proc. of the International Conference on Face and Gesture Recognition*, 54-61, 2000.
- [9] J.-M. Jolion, P. Meer, and S. Bataouche, Robust clustering with applications in computer vision, *IEEE transactions on pattern analysis and machine intelligence*, vol. 13, no. 8, pp. 791-802, 1991.
- [10] J. Mao and A.K. Jain, A self-organizing network for hyperellipsoidal clustering (HEC), *IEEE transactions on neural networks*, vol. 7, no. 1, pp. 16-29, 1996.
- [11] A.P. Dempster, N.M. Laird and D.B. Rubin, Maximum likelihood from incomplete data via the EM algorithm (with discussion), *Journal of the royal statistical society B*, vol. 39, no. 1, pp. 1-38, 1977.
- [12] L. Xu, M.I. Jordan, and G.E. Hinton, An alternative model for mixtures of experts, *Advances in neural information processing systems*, In G. Tesauro, D.S. Touretzky, and T.K. Leen (Eds.), 7:633-640, Cambridge, MA:MIT Press, 1995.
- [13] S.J. Nowlan, Soft competitive adaptation: neural network learning algorithms based on fitting statistical mixtures, PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA., 1991.
- [14] R.M. Neal and G.E. Hinton, A new view of the EM algorithm that justifies incremental and other variants, *Learning in graphical models*, pp. 355-368, 1998.
- [15] S.D. Backer and P. Scheunders, A competitive elliptical clustering algorithm, *Pattern recognition letters*, vol. 20, no. 11/13, pp. 1141-1147, 1999.
- [16] M. Sato and S. Ishii, On-line EM algorithm for the normalized Gaussian network, *Neural computation*, vol. 12, no. 2, pp. 407-432, 2000.
- [17] P. Sceunders, A comparison of clustering algorithms applied to color image quantization, *Pattern recognition letters*, vol. 18, no. 11/13, pp. 1379-1384, 1997.

[18] K.-M. Lee and W.N. Street, An adaptive resource-allocating network for automated detection, segmentation, and classification of breast cancer nuclei, *IEEE Transactions on neural networks*, vol. 14, no. 3, pp. 680-687, 2003.

[19] J. Platt, A resource-allocating network for function interpolation, *Neural computation*, vol. 3, no. 2, pp. 213-225, 1991.

[20] S. Jrgen, *Pattern classification: A unified view of statistical and neural approaches*, John Wiley & Sons, NY, 1996.

[21] J. Li and J.Z. Wang, Automatic linguistic indexing of pictures by a statistical modeling approach, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1075-1088, 2003.

로써, 공분산 행렬 $(\Sigma_{jR}^n)^{-1}$ 의 역을 직접 갱신할 수 있다.



이 경 미

1993년 덕성여대 전산학과(학사). 1996년 연세대 전산학과(석사). 2001년 미국 Univ.of Iowa 전산학과(박사). 현재 덕성여대 컴퓨터과학부 전임강사. 관심분야는 멀티미디어 처리 및 검색, 컴퓨터시각, 패턴인식, 신경망, 기계학습

부 록

식 (8)을 이용하여 직접 조정된 $(\Sigma_{jR}^n)^{-1}$ 의 역행렬을 구하기 위해서,

$$\begin{aligned}
 &= \frac{[(1-\eta^n)\Lambda_j^{n-1} + \eta^n P_j^n \varphi_j^{n-1}]^{-1}}{[(1-\eta^n)f_j^{n-1}(1) + \eta^n P_j^n]^{-1}} \text{ 이 때} \\
 \Lambda_j^{n-1} &= f_j^{n-1} \left(m_j, m_j' \right) + \beta \frac{f_j^{n-1}(|m_j^n|^2)}{d} I_d \text{ 과} \\
 \varphi_j^{n-1} &= m_j^n m_j'^n + \beta \frac{|m_j^n|^2}{d} I_d, \\
 &= \frac{[I + kP_j^n (\Lambda_j^{n-1})^{-1} \varphi_j^{n-1}]^{-1} (\Lambda_j^{n-1})^{-1}}{[I + kP_j^n (f_j^{n-1}(1))^{-1}]^{-1} (f_j^{n-1}(1))^{-1}} \text{ 이 때 } k = \frac{\eta^n}{1-\eta^n}, \\
 &= \frac{\left(I - kP_j^n (\Lambda_j^{n-1})^{-1} \varphi_j^{n-1} \right. \\
 &\quad \left. + k^2 P_j^n (\Lambda_j^{n-1})^{-1} \varphi_j^{n-1} P_j^n (\Lambda_j^{n-1})^{-1} \varphi_j^{n-1} - k^3 \dots \right) (\Lambda_j^{n-1})^{-1}}{\left(I - kP_j^n (f_j^{n-1}(1))^{-1} \right. \\
 &\quad \left. + k^2 P_j^n (f_j^{n-1}(1))^{-1} P_j^n (f_j^{n-1}(1))^{-1} - k^3 \dots \right) (f_j^{n-1}(1))^{-1}} \\
 &= \frac{\left((\Lambda_j^{n-1})^{-1} - kP_j^n (\Lambda_j^{n-1})^{-1} \varphi_j^{n-1} (\Lambda_j^{n-1})^{-1} \right. \\
 &\quad \left. \times (1 - k\alpha(m_j^n) + k^2 \alpha^2(m_j^n) - \dots) \right)}{\left((f_j^{n-1}(1))^{-1} - kP_j^n (f_j^{n-1}(1))^{-1} (f_j^{n-1}(1))^{-1} \right. \\
 &\quad \left. \times (1 - k\alpha(1) + k^2 \alpha^2(1) - \dots) \right)} \text{ 이 때} \\
 \alpha(m_j^n) &= P_j^n (\Lambda_j^{n-1})^{-1} \varphi_j^{n-1} \text{ 과 } \alpha(1) = P_j^n (f_j^{n-1}(1))^{-1}, \\
 &= \frac{(\Lambda_j^{n-1})^{-1} - \frac{kP_j^n}{1+k\alpha(m_j^n)} (\Lambda_j^{n-1})^{-1} \varphi_j^{n-1} (\Lambda_j^{n-1})^{-1}}{(f_j^{n-1}(1))^{-1} - \frac{kP_j^n}{1+k\alpha(1)} (f_j^{n-1}(1))^{-1} (f_j^{n-1}(1))^{-1}} \quad (11)
 \end{aligned}$$

$(f_j^{n-1}(1))^{-1}$ 과 $(\Lambda_j^{n-1})^{-1}$ 의 역행렬을 재귀적으로 갱신함