

ASM의 성능향상을 위한 형태 정렬 방식 제안

한 희 일[†]

요 약

본 논문에서는 ASM(active shape model)의 성능을 향상시키기 위하여 형태(shape) 정렬 방법과 이차원 특징벡터 추출 방법을 제안한다. 기존 알고리즘은 입력 이미지의 중간 검출 랜드마크와 기준 모델 간의 정렬을 위하여 스케일, 회전, 이동 정보만을 이용한다. 하지만 위의 평면적인 정보만으로는 얼굴과 같이 입체적인 물체의 포즈 변화나 삼차원적인 움직임 등을 제대로 반영할 수 없다. 이를 개선하기 위하여 자유도를 증가시킴으로써 형태의 복잡한 변화에 보다 강인한 형태정렬 방식을 제안한다. 또한, 멀티스케일로 이차원 프로파일을 구하고 이들의 공분산 행렬을 trimming하여 검출속도를 향상시키는 방법을 제안한다. 비교적 다양한 포즈로 촬영한 얼굴 이미지 데이터베이스를 이용하여 제안 알고리즘의 형태 검출 성능을 확인한다.

Proposing Shape Alignment for an Improved Active Shape Model

Hee-il Hahn[†]

ABSTRACT

In this paper an extension to an original active shape model(ASM) for facial feature extraction is presented. The original ASM suffers from poor shape alignment by aligning the shape model to a new instant of the object in a given image using a simple similarity transformation. It exploits only informations such as scale, rotation and shift in horizontal and vertical directions, which does not cope effectively with the complex pose variation. To solve the problem, new shape alignment with 6 degrees of freedom is derived, which corresponds to an affine transformation. Another extension is to speed up the calculation of the Mahalanobis distance for 2-D profiles by trimming the profile covariance matrices. Extensive experiment is conducted with several images of varying poses to check the performance of the proposed method to segment the human faces.

Key words: ASM, 형태정렬(Shape Alignment), 아파인 변환(Affine Transform), 주성분 분석(Principal Component Analysis)

1. 서 론

이미지에서 어떤 물체를 검출하고자 할 때, 그 물체의 형태가 고정적으로 일정한 형태를 띠면 쉽게 해결할 수 있지만, 사람의 얼굴이나 의료영상 속의 장기 같은 경우는 특정 범위 내에서 개인 별로 또는 시간에 따라 그 형태가 가변적이므로 이들을 적절히 모델링하는 것조차 쉬운 문제가 아니다. 예를 들어

사람의 얼굴은 개인 별로 모습이 다를 뿐만 아니라 모자나 안경을 쓸 수도 있고 수염을 기른 경우도 있으므로 이러한 가변적인 특성을 반영한 모델을 구성하는 것은 매우 어려운 문제에 속한다.

가변적인 형태를 검출하기 위하여 snake(active contour model)[1], ASM[2], AAM(active appearance model)[3] 등의 알고리즘이 제안되었다. snake 알고리즘은 형태 모델 정보를 포함하고 있지 않아서

※ 교신저자(Corresponding Author): 한희일, 주소: 경기도 용인시 모현면 왕산리 한국외국어대학교 정보통신공학과(449-791), 전화: 031-330-4356, FAX: 031)333-4256, E-mail: hihahn@hufs.ac.kr
접수일: 2011년 11월 22일, 수정일: 2011년 12월 22일

완료일: 2012년 1월 3일
[†] 정희원, 한국외국어대학교 정보통신공학과
※ 본 연구는 2011학년도 한국외국어대학교 학술연구비 지원으로 수행되었음.

그 형태와 관련 없는 윤곽선으로 수렴하는 경우가 발생하는 한계를 보이고 있으나 ASM과 AAM 등은 검출하고자 하는 물체의 다양한 형태에 대한 통계 정보를 이용하여 이러한 문제를 해결하고 있다. 이들은 검출하고자 하는 물체의 통계적 특성을 반영하여 충분한 양의 표본 이미지를 구한 다음, 각 이미지에서 물체의 특성을 나타내는 특징점들을 표시한다[4]. 이러한 특징점들을 랜드마크(landmark)라고 부른다. 여기서 충분한 양이란 표본 이미지의 수가 랜드마크의 수보다 충분히 커야 함을 의미한다. 그 점들은 경계선이거나 내부 또는 외부의 점일 수도 있다. 이러한 랜드마크들의 통계적 특성을 구하면 각 랜드마크마다 가변 정도를 파악할 수 있다. 이와 같이 모델을 구할 수 있으면, 검출하고자 하는 물체를 포함한 이미지에서 각 랜드마크에 해당되는 점들을 통계적 범위 내에서 찾아냄으로써 그 물체를 검출할 수 있다. 이를 위해서는 우선 그 물체의 대강 위치와 방향, 크기 등을 이용하여 물체의 초기 형태를 구하여야 한다. 초기 형태가 주어지면, 이로부터 입력 이미지에서 모델의 랜드마크에 해당되는 점들을 모델의 통계적 특성에 맞도록 찾아냄으로써 물체를 검출할 수 있다.

ASM이 훈련 이미지의 랜드마크 점들을 통해 구한 형태정보를 이용하여 랜드마크 인근 영역을 조사함으로써 보다 정확한 형태를 검출하고자 하는데 비하여 AAM은 ASM의 후속 모델로서 형태 정보뿐만 아니라 텍스처도 함께 이용하여 모델링함으로써 ASM의 성능을 개선시키기 위하여 제안되었다. 하지만 AAM은 모델을 구하고 검출하는데 계산시간이 너무 오래 걸리는 단점이 있다. 따라서 본 논문에서는 적절한 계산비용으로 우수한 성능을 보이는 ASM을 채택하였다. 현재까지 ASM의 성능을 개선시키기 위하여 많은 연구가 진행되고 있다. Xu, et al.[5]은 ASM의 성능을 개선하기 위하여 훈련 이미지를 이용하여 랜드마크를 정합하는 과정에서 procrustes 분석 방법을 제안하였다. Zheng, et al.[6]은 랜드마크가 에지 상에 있는지의 여부에 따라 특징벡터를 달리 정의하고, 프로파일 정합 과정에서도 Mahalanobis 거리에 이를 반영하는 방법과 함께 멀티 스케일로 검색할 때 스케일에 따라 프로파일의 길이를 가변적으로 조정하는 방법을 제안하였다. Huang, et al.[7]은 이미지의 밝기 변화에 의한 검색

성능 저하를 해결하기 위하여 2-D Gabor 필터를 이용한 특징벡터 추출 방안을 제안하였다. 지금까지 ASM의 성능을 개선시키기 위한 연구는 대부분 특징벡터나 프로파일 등을 구하는 방법을 통한 성능향상이 대부분이다. 그런데, Mohammad, et al.[8]은 자유도 4인 기존의 형태정렬 대신에 자유도 6의 어파인 변환을 이용한 정렬방식을 채택하여 성능을 개선시키고자 하였다. 이 알고리즘은 얼굴 이미지에서 두 눈동자와 입의 중간점 등, 세 개의 점을 찾은 다음, 이를 이용하여 6개의 미지수를 구하는 방식을 채택하고 있다. 하지만, 안경을 쓰거나, 머리카락이 한쪽 눈을 가리는 등, 이미지에 따라 세 점을 정확히 찾지 못하는 경우에는 적용하지 못하거나 정확도가 크게 저하되는 문제를 안고 있다.

본 논문에서는 위의 문제점을 해결하기 위하여 세 점을 이용하는 방식 대신에 이미지의 모든 랜드마크를 이용하여 변환 행렬을 구하는 방식을 채택함으로써 이미지의 특성에 구애 받지 않고 안정적으로 적용 가능한 정렬 방식을 제안한다. 또한, 1-D 프로파일과 함께 2-D 프로파일을 특징벡터로 이용하는데, 특히 2-D 프로파일 모델을 구하는 과정에서 이의 공분산 행렬의 크기가 매우 커서 랜드마크 검출 단계에서 엄청난 계산량이 요구된다. 이를 해결하기 위하여 공분산(covariance) 행렬을 trimming하는 방식을 채택하였다. 제안 알고리즘의 성능을 확인하기 위하여 케이프타운 대학에서 제작한 얼굴 이미지 데이터베이스를 이용하여 얼굴 모델을 구하고 이를 이용하여 얼굴을 검출하는 실험을 수행하였다.

본 논문의 구성은 다음과 같다. 2장에서는 ASM의 형태 모델을 리뷰하고 본 논문에서 제안한 형태 정렬 방식을 기술하며, 본 논문에서 채택한 프로파일 모델을 3장에서 간략히 소개한다. 제안 알고리즘을 수행하였을 때의 실험결과와 이에 대한 성능분석은 4장에서 설명한다. 마지막으로 5장에서는 결론을 맺고 향후 연구 진행방향에 대하여 논의한다.

2. ASM의 형태 모델

사람의 얼굴 형태를 표현하기 위한 모델을 구현하기 위해서는 얼굴의 가변 정도를 파악하는데 필요한 충분한 양의 훈련 이미지가 필요하다. 각 이미지 내의 얼굴 주위에 특정한 랜드마크를 표시하고 라벨링

한 다음, 동일한 라벨의 점들에 대한 통계적 특성을 구함으로써 얼굴형태에 대한 모델을 구할 수 있다. 여기서, 동일한 라벨의 랜드마크들은 훈련이미지 내에 있는 얼굴형태의 동일한 위치를 나타낼 수 있도록 하여야 한다.

2.1 형태 정렬

각 훈련 이미지에서 추출한 랜드마크들 간의 통계 정보를 추출하기 위해서는 우선 이 점들을 적절히 정렬시켜야 한다. \mathbf{x}_i 를 i 번째 형태에서 추출한 n 개의 랜드마크 점들의 벡터라고 정의하면, 이를 다음과 같이 나타낼 수 있다.

$$\mathbf{x}_i = \begin{pmatrix} x_{i0} & y_{i0} \\ x_{i1} & y_{i1} \\ \dots & \dots \\ x_{in-1} & y_{in-1} \end{pmatrix} \quad (1)$$

Cootes[2]는 두 개의 형태벡터 \mathbf{x}_i 와 \mathbf{x}_j 가 주어지면, \mathbf{x}_j 를 \mathbf{x}_i 의 포즈에 정합되도록 다음과 같이 similarity 변환하여 형태벡터 \mathbf{x}'_j 를 구하는 방법을 채택하였다.

$$\begin{bmatrix} x'_{jk} \\ y'_{jk} \\ 1 \end{bmatrix} = \begin{bmatrix} s_j \cos \theta_j & -s_j \sin \theta_j & t_x \\ s_j \sin \theta_j & s_j \cos \theta_j & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{jk} \\ y_{jk} \\ 1 \end{bmatrix} \quad (2)$$

여기서 파라미터 θ_j , s_j , t_x , t_y 는 다음과 같이 정의된 \mathbf{x}_i 와 \mathbf{x}'_j 간의 거리를 최소화하도록 정한다.

$$E = (\mathbf{x}_i - \mathbf{x}'_j)^T \mathbf{W} (\mathbf{x}_i - \mathbf{x}'_j) \quad (3)$$

여기서 \mathbf{W} 는 각 점의 비중을 별도로 제공하기 위한 대각 행렬로서 모든 랜드마크에 동일한 비중을 할당한다면 단위행렬이 된다. 대부분의 ASM과 AAM 관련 논문들은 이 방법을 채택하고 있다 [4,10,12]. 하지만, 이 방식은 스케일, 회전, 이동 정보만을 고려하는 자유도 4의 변환 방식이어서 형태의 다양한 변화를 제대로 반영할 수 없는 한계를 갖고 있다. 예를 들어, 정사각형의 형태는 변환 후에도 정사각형을 유지하므로 직사각형의 형태와 제대로 정렬이 되지 못한다. 이러한 단점을 개선시키기 위하여 본 논문에서는 다음과 같이 자유도가 6인 어파인 변환 방식으로 정렬시키는 알고리즘을 유도하였다.

$$\begin{bmatrix} x'_{jk} \\ y'_{jk} \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{jk} \\ y_{jk} \\ 1 \end{bmatrix} \quad (4)$$

\mathbf{x}_i 와 \mathbf{x}_j 를 어파인 변환한 \mathbf{x}'_j 간의 거리를 위의 식 (3)과 동일하게 정의하면 다음과 같이 나타낼 수 있고,

$$E = \sum_{k=0}^{n-1} w_k (x_{ik} - a_{11}x_{jk} - a_{12}y_{jk} - t_x)^2 + \sum_{k=0}^{n-1} w_k (y_{ik} - a_{21}x_{jk} - a_{22}y_{jk} - t_y)^2 \quad (5)$$

이를 최소화시키는 어파인 행렬은, least squares 알고리즘을 식 (5)에 적용하여 얻은 다음 행렬 식의 근을 구함으로써 얻을 수 있다.

$$\begin{bmatrix} A_1 & B & C_1 & 0 & 0 & 0 \\ B & A_2 & C_2 & 0 & 0 & 0 \\ C_1 & C_2 & W & 0 & 0 & 0 \\ 0 & 0 & 0 & A_1 & B & C_1 \\ 0 & 0 & 0 & B & A_2 & C_2 \\ 0 & 0 & 0 & C_1 & C_2 & W \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \\ t_x \\ a_{21} \\ a_{22} \\ t_y \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \\ \alpha_2 \\ \beta_2 \\ \gamma_2 \end{bmatrix} \quad (6)$$

여기서, 위 식의 여러 파라미터들은 다음과 같이 정의된다.

$$\begin{aligned} A_1 &= \sum_{k=0}^{n-1} w_k x_{jk}^2 & A_2 &= \sum_{k=0}^{n-1} w_k y_{jk}^2 \\ C_1 &= \sum_{k=0}^{n-1} w_k x_{jk} & C_2 &= \sum_{k=0}^{n-1} w_k y_{jk} \\ B &= \sum_{k=0}^{n-1} w_k x_{jk} y_{jk} & W &= \sum_{k=0}^{n-1} w_k \\ \alpha_1 &= \sum_{k=0}^{n-1} w_k x_{ik} x_{jk} & \alpha_2 &= \sum_{k=0}^{n-1} w_k y_{ik} y_{jk} \\ \beta_1 &= \sum_{k=0}^{n-1} w_k x_{ik} y_{jk} & \beta_2 &= \sum_{k=0}^{n-1} w_k x_{jk} y_{ik} \\ \gamma_1 &= \sum_{k=0}^{n-1} w_k x_{ik} & \gamma_2 &= \sum_{k=0}^{n-1} w_k y_{ik} \end{aligned} \quad (7)$$

본 논문에서는 N 개의 훈련 이미지에서 구한 형태 벡터를 정렬하기 위해 첫번째 형태 \mathbf{x}_0 를 기준 형태 벡터로 정하고, 나머지 각 형태벡터 ($\mathbf{x}_1 \sim \mathbf{x}_{N-1}$) 를 \mathbf{x}_0 와 정렬시킨 다음, 정렬된 형태벡터로부터 평균 형태벡터와 공분산 행렬 등의 통계정보를 구한다.

2.2 정렬된 형태벡터의 통계정보 추출

식 (1)에 주어진 형태 행렬은 n 개의 점으로 구성되어 있는데 연산의 편의를 위하여 다음과 같이 $2n$ 차원의 벡터로 표현한다.

$$\mathbf{x}_i = (x_{i0}, x_{i1}, \dots, x_{i(n-1)}, y_{i0}, y_{i1}, \dots, y_{i(n-1)})^T \quad (8)$$

N 개의 정렬된 형태벡터가 구해지면, 평균 형태벡터 $\bar{\mathbf{x}}$ 는 다음과 같이 구할 수 있다.

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=0}^{N-1} \mathbf{x}'_i \quad (9)$$

여기서 \mathbf{x}'_i 는 \mathbf{x}_i 를 \mathbf{x}_0 중심으로 정렬한 형태벡터이고 \mathbf{x}'_0 는 \mathbf{x}_0 와 동일하다. 이와 같이 구한 정렬된 형태벡터와 평균 형태벡터를 이용하여 PCA(principal component analysis)를 적용함으로써 이들의 통계정보를 추출한다. 즉, 다음과 같이 $2n \times 2n$ 공분산 행렬을 구한 다음,

$$\mathbf{S} = \frac{1}{N} \sum_{i=0}^{N-1} (\mathbf{x}'_i - \bar{\mathbf{x}})(\mathbf{x}'_i - \bar{\mathbf{x}})^T \quad (10)$$

이 행렬에서 고유값(eigen value)과 고유벡터(eigen vector)를 다음과 같이 구한다.

$$\mathbf{S}\mathbf{e}_k = \lambda_k \mathbf{e}_k, \quad k=0, \dots, 2n-1 \quad (11)$$

여기서, 고유벡터 \mathbf{e}_k 는 그 크기가 1로 정규화되도록 조정하고, λ_k 는 공분산 행렬 \mathbf{S} 의 k 번째 고유값으로서 내림차순으로 정렬되었다고 가정한다. 즉, $\lambda_k \geq \lambda_{k+1}$ 를 만족한다. 고유값 λ_k 는 N 개의 형태벡터에서 고유벡터 \mathbf{e}_k 방향으로 변하는 정도의 분산에 비례한다. 즉, 가장 큰 고유값 λ_0 에 해당되는 고유벡터 \mathbf{e}_0 은 N 개의 형태벡터에서 가장 크게 변하는 성분을 나타내고 그 분산은 λ_0 에 비례한다. 이러한 수학적 특성을 얼굴 검출에 적용하면, 충분한 양의 훈련 얼굴 이미지에서 형태벡터를 추출하여 공분산 행렬을 구한 다음 고유값과 고유벡터를 계산하면, 얼굴이 사람마다 어느 성분의 변화가 크고 작은지를 확인할 수 있다. 예를 들어, 그림 1에 제시한 바와 같이 임의의 형태벡터에서 \mathbf{e}_0 의 성분을 변화시키면 얼굴의 폭이 변화함을 알 수 있다. 이러한 사실은 얼굴의 폭이 그 특성 중에서 사람마다 가장 크게 변화하는 성분임을 나타낸다. 대부분의 경우 $2n$ 개의 고

유값 중에서 일부만 큰 값을 갖고 나머지는 매우 작은 값을 갖는다. 따라서, $2n$ 차원의 형태벡터 \mathbf{x} 를 $2n$ 개의 고유벡터 대신 그보다 훨씬 적은 수(편의상 t 개라고 정한다.)의 벡터로 표현할 수 있기 때문에 자유도를 크게 줄일 수 있어 문제를 단순화시킬 수 있는 효과를 얻을 수 있다. 본 논문에서는 다음과 같이 고유값의 합을 구한 다음, 그 합의 일정부분을 초과하는 최소의 고유벡터 수 t 를 구하였다.

$$\sum_{k=0}^{t-1} \lambda_k \geq Thr \quad (12)$$

일반적으로 Thr 은 응용에 따라 0.9~0.99의 값을 갖도록 정하는데, 본 논문에서는 0.98로 정하였다. 위에서 설명한 바와 같이 λ_k 는 고유벡터 \mathbf{e}_k 방향으로 변하는 정도의 분산을 나타내므로 Thr 은 검출능력을 에너지 비율로 나타낸 것으로 해석될 수 있다. 다시 말하면, 훈련 이미지 내에 존재하는 임의의 형태벡터 \mathbf{x} 는 다음과 같이 평균 형태벡터와 t 개의 고유벡터의 합으로 나타낼 수 있는데, 이는 대략 전체 에너지의 $100 \times Thr$ %의 정확도로 근사화시킬 수 있는 것으로 해석될 수 있다.

$$\mathbf{x} = \bar{\mathbf{x}} + \Phi \mathbf{b} \quad (13)$$

여기서, $\Phi = \{\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_{t-1}\}$ 는 t 개의 고유벡터로 구성된 행렬이고 $\mathbf{b} = (b_0, b_1, \dots, b_{t-1})^T$ 는 고유벡터의 비례상으로 구성된 열벡터이다. 위 식은 파라미터($b_0 \sim b_{t-1}$)를 적절히 변화시켜 줌으로써 임의의 형태벡터를 생성할 수 있음을 보여 준다. 앞에서 설명한 바와 같이 Thr 은 검출의 정확도를 에너지 비율로 표현한 것으로서 Thr 이 커질수록 검출의 정확도는 제고되지만 고유벡터의 수가 늘어나 계산량이 증가하고 파라미터 제어가 어려워지는 문제가 있다. 즉, Thr 이 1에 근사할수록, 고유벡터의 수 t 가 작을수록 적은 계산량으로 보다 정확한 검출결과를 얻을 수 있다. 본 논문에서 이용한 얼굴 이미지 데이터베이스로 실험한 결과 $Thr = 0.98$ 이고 $t = 20$ 일 때 적절한 난이도로 우수한 검출결과를 얻을 수 있음을 확인하였다. 또한, b_k 의 분산은 고유값 λ_k 과 일치하므로 b_k 가 가우시안 분포를 갖는다고 가정하면 대부분이 표준편차의 3배 범위 내에 존재하므로 b_k 의 범위를 다음과

같이 한정시키면,

$$-3\sqrt{\lambda_k} \leq b_k \leq 3\sqrt{\lambda_k} \quad (14)$$

대부분의 얼굴 형태에 대한 변화량을 감당할 수 있을 것으로 파악된다.

3. ASM의 프로파일 모델

형태벡터를 구성하는 각 랜드마크마다 그 주위의 이미지 특성이 서로 다르기 때문에 입력 이미지에서 이들을 효과적으로 검출하기 위해서는 각 점마다 별도의 특징벡터 모델이 필요하다. Cootes[2]가 처음 제안한 ASM은 1-D 프로파일을 특징벡터로 이용하였으나 그 이후로 본 논문을 포함한 대부분의 알고리즘은 2-D 프로파일도 함께 이용하기 시작하였다. 다음에서는 본 논문에서 구현한 두 프로파일 모델에 대하여 설명한다.

3.1 1-D 프로파일 모델

각 랜드마크마다 모든 훈련 이미지에서 윤곽선에 수직인 방향으로 1-D 프로파일 벡터 \mathbf{g} 를 추출한다. 여기서 윤곽선이란 인근 랜드마크들을 서로 연결하여 얻은 폐곡선을 말한다. \mathbf{g} 의 각 원소는 인근 원소 간의 차로 대치시킴으로써 이미지 밝기의 변화 등에 의한 영향을 줄인다. 각 랜드마크마다 N 개의 훈련 이미지에서 구한 프로파일에 대하여 평균 프로파일 벡터 $\bar{\mathbf{g}}$ 와 공분산 행렬 \mathbf{S}_g 를 구한다. 본 논문에서 이용한 이미지의 형태벡터는 68의 랜드마크로 구성되어 있으므로 그만큼의 $\bar{\mathbf{g}}$ 와 \mathbf{S}_g 를 구하여야 한다. 검출 단계에서는 각 중간 단계에서 추출한 프로파일 \mathbf{g} 와 모델의 평균 프로파일 벡터 $\bar{\mathbf{g}}$ 간의 Mahalanobis거리를 다음과 같이 구하여 최소값을 갖는 프로파일 \mathbf{g} 의 중심점을 후보 점으로 인식한다.

$$\text{Distance} = (\mathbf{g} - \bar{\mathbf{g}})^T \mathbf{S}_g^{-1} (\mathbf{g} - \bar{\mathbf{g}}) \quad (15)$$

3.2 2-D 프로파일 모델

각 랜드마크를 중심으로 13×13의 패치를 떼어내어 가로와 세로 방향의 차분을 각각 구한 다음 그 크기를 정규화시킨다. 형태모델을 구할 때와 유사한 방법으로 패치를 169×1의 벡터로 변환한 다음 1-D

프로파일에서와 같이 평균 프로파일 벡터 $\bar{\mathbf{g}}$ 와 공분산 행렬 \mathbf{S}_g 를 구한다. 그런데 \mathbf{S}_g 는 169×169의 정방행렬이어서 이를 이용하여 식 (15)와 같이 거리를 계산하는데 엄청난 계산량이 필요하다. 본 논문에서는 계산량을 줄이기 위하여 \mathbf{S}_g 를 trimming하는 방식을 채택하였다. 즉 \mathbf{S}_g 의 대각 원소는 매우 큰 값을 갖지만 이에서 멀어짐에 따라 그 값은 크게 감소한다. 이를 이용하여 본 논문에서는 대각선에서 4 이상의 거리를 갖는 원소에 대해서는 그 값을 0으로 지정하였다. 이 과정을 수행하면 169×169=28,561의 원소 중에서 $(3+1+3) \times 169 = 1183$ 개 (4.1%)의 원소만 남게 되므로 산술적으로 96% 정도의 계산량을 절약할 수 있는 이점이 있다. 원래 \mathbf{S}_g 는 positive definite 한 행렬이어서 식 (15)의 Mahalanobis거리를 구할 수 있지만, trimming을 하면 이를 보장할 수 없는 문제가 있다. 이를 해결하기 위하여 본 논문에서는 \mathbf{S}_g 를 trimming한 후에 강제로 positive definite하도록 조정하는 과정을 추가하였다[9,10].

4. 실험 결과

제안 알고리즘의 동작을 확인하고 성능을 분석, 평가하기 위해 본 논문에서는 케이프타운 대학교에서 제작한 MUCT 데이터베이스에서 1728개의 이미지(해상도: 640×480)를 선택하여 이용하였다. 본 논문에서는 1728개의 형태 정보를 이용하여 2.1과 2.2 절에서 설명한 방식대로 평균 형태벡터 $\bar{\mathbf{x}}$ 와 136×136의 공분산 행렬 \mathbf{S} 를 구하고 136개의 고유값과 고유벡터를 계산한 다음 이 중에서 고유값의 크기 순으로 식 (12)에 따라 $t = 20$ 개를 선택하였다.

그림 1은 평균 형태벡터 $\bar{\mathbf{x}}$ 에 3개의 고유벡터 ($\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2$)의 성분을 각각 변경시켰을 때 형태의 변화 과정을 보여 주고 있다. 그림 1(a)는 가장 큰 고유벡터 \mathbf{e}_0 의 비례상수 b_0 를 -3, -1, 0, 1, 3로 변경하면서 그린 형태이다. 위에서 설명한 바와 같이 사람의 얼굴 형태 정보 중에서 얼굴의 폭이 가장 크게 변화하는 성분임을 알 수 있다. 그림 1(b)에서는 \mathbf{e}_1 성분이 얼굴 형태에 미치는 현상을 나타내는 바, 이 성분은 얼굴 고유의 특성이라기 보다는 데이터 베이스의 이미지가 모두 정면에서 촬영한 것이 아니라 좌우 옆모습을 찍은 이미지가 적지 않아서 발생한 현상으

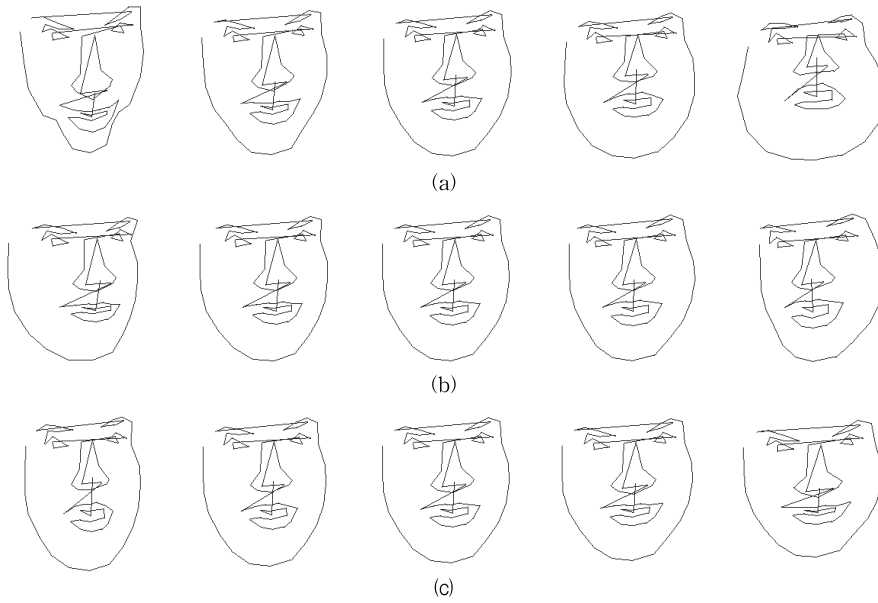


그림 1. 식 (13)에서 $b=-3, -1, 0, 1, 3$ (좌측부터)일 때 x 의 변화 (a) $k=0$, (b) $k=1$, (c) $k=2$

로 해석될 수 있다. 그림 1(c)를 보면 그 다음으로 많이 변하는 특성은 얼굴의 길이임을 유추할 수 있다. 이와 같이 고유값이 크면 클수록 정보량은 증가하지만, 반대로 작으면 해당 성분이 사람마다 거의 일정하므로 무시할 수 있다. 결국, 본 논문에서 구현한 ASM은 136개의 데이터로 구성된 형태벡터 \mathbf{x} 를 KL(Karhunen Loeve) 변환한 후 $t=20$ 의 데이터로 압축하는 알고리즘과 동일함을 알 수 있다.

검출하고자 하는 형태를 포함하는 이미지가 주어지면, 우선적으로 위치 정보와 함께 대강의 형태 정보를 추출하여야 한다. 초기의 형태 정보는 정확하지 않아도 별 문제는 없지만 원 형태 정보에서 크게 벗어나지 않아야 제대로 검출할 확률이 높아진다. 본 논문에서는 입력 이미지에서 얼굴과 눈의 위치를 찾고 대강의 형태정보를 찾기 위하여 Viola Jones 검출기[11]를 이용하였다.

입력이미지에서 얼굴을 검출하기 위하여 제일 먼저 Viola Jones 검출기로 추출한 얼굴의 위치, 크기, 방향 등의 정보를 이용하여 평균 형태벡터 $\bar{\mathbf{x}}$ 를 조정하여 이를 초기의 얼굴형태로 정한 다음 ASM 알고리즘을 반복적으로 적용하여 얼굴형태를 검색하였다. 이와 같이 구한 초기 형태벡터를 이용하여 이의 각 점(랜드마크)마다 윤곽선에 수직인 방향으로 후보 점을 찾는다. 여기서 윤곽선이란 형태벡터의 점

들을 서로 연결하여 얻은 폐곡선을 말한다.

그림 2는 1-D 프로파일을 형태로 모델링하고, 이를 기반으로 50개의 테스트 이미지를 이용하여 얼굴 검출을 수행하였을 때, 정렬 방법을 달리 하면서 각 랜드마크와 그에 해당되는 검출 점 간의 평균 자승오차 (mean square error)를 나타낸 것이다. 이 그림에서 알 수 있듯이, 기존의 similarity 변환 방식으로 정렬시키면 이미지에 따라 검출 정확도가 크게 변화하지만 본 논문에서 제안한 정렬방식을 이용하면 보다 정확한 검출 성능을 얻을 수 있다. 그림 4와 그림 5의 상단에 평균자승오차가 매우 큰 2번 이미지

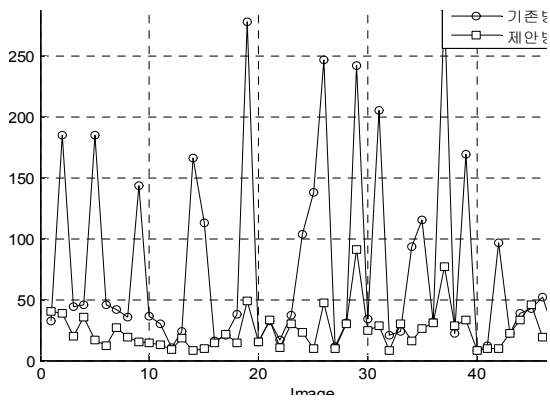


그림 2. 1-D 프로파일을 특징벡터로 이용 시 정렬 방식의 변화에 따른 검색 성능 비교

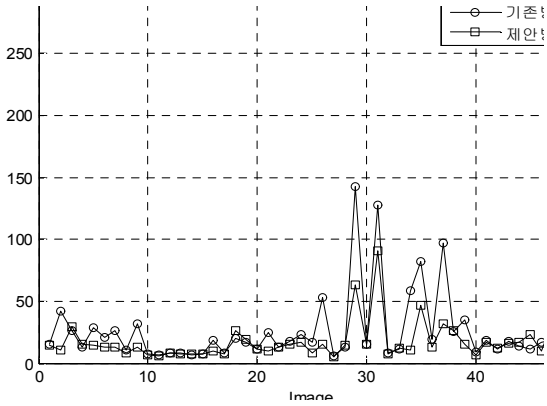


그림 3. 1-D와 2-D 프로파일을 특징벡터로 이용 시 정렬 방식의 변화에 따른 검색 성능 비교

와 9번 이미지에 대하여 1-D 프로파일만으로 얼굴을 검출한 결과를 제시하였는데, 이를 통해서도 제안 방식으로 형태를 정렬하면 보다 안정적이고 우수한 검출 성능을 얻을 수 있음을 확인할 수 있다.

그림 3은 1-D 프로파일과 2-D 프로파일을 함께 이용하여 구한 검출 결과를 보여 주고 있다. 예상한 바와 같이, 1-D 프로파일과 2-D 프로파일을 함께 이용하면 검출 성능이 크게 향상되며, 제안 방식을 이용하여 형태를 정렬하면 기존방식보다 우수한 결과를 얻을 수 있음을 확인하였다.

그림 4와 그림 5의 하단에는 평균자승오차가 매우 큰 2번 이미지와 9번 이미지에 대하여 1-D 프로파일과 2-D 프로파일을 순차적으로 이용하여 얼굴을 검출한 결과를 제시하였다. 표 1은 50개의 테스트 이미지에서 구한 mse에 대한 평균값을 나타내고 있다. 이를 통하여 본 논문에서 제안한 제안방식을 이용하여 형태를 정렬하면 보다 안정적이고 우수한 검출 성능을 얻을 수 있음을 확인하였다.

표 1. 기존 방식과 제안 방식의 평균 mse 비교

	기존방식	제안방식
1-D 프로파일	76.9414	25.1761
1-D + 2-D 프로파일	25.9493	17.2892

4. 결 론

본 논문에서는 ASM의 검출 성능을 향상시키기 위하여 형태정보의 정렬방식을 새로이 유도하고 실

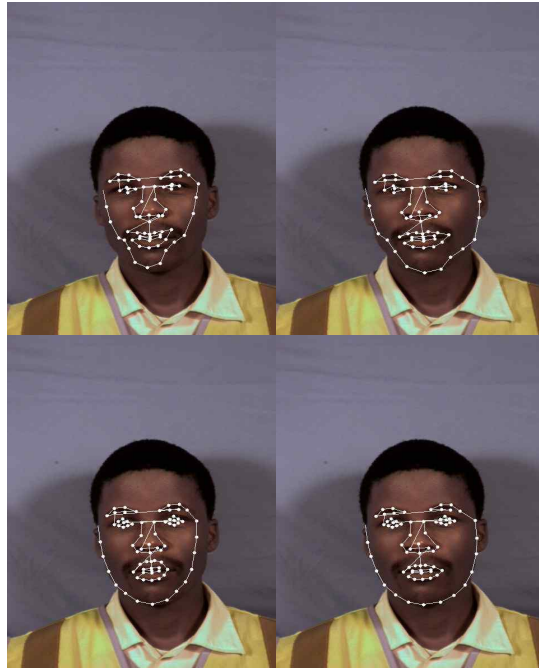


그림 4. 2번 테스트 이미지에서의 검색 성능 비교. 상단 : 1-D 프로파일 하단 : 2-D 프로파일 좌측 : 기존방식 우측 : 제안방식

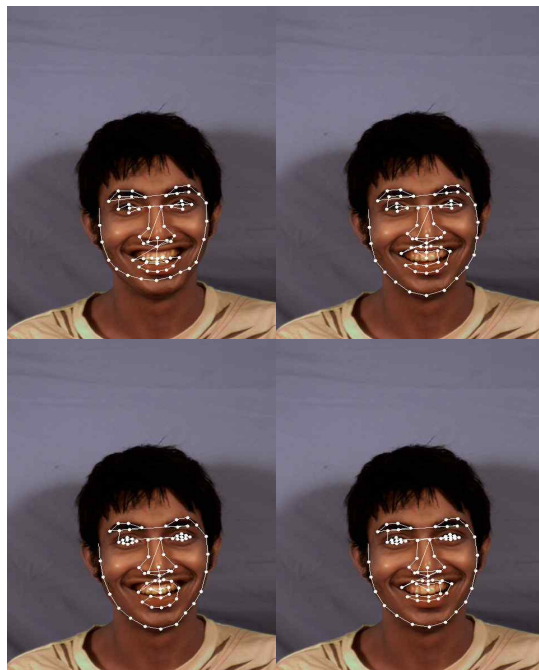


그림 5. 9번 테스트 이미지에서의 검색 성능 비교. 상단 : 1-D 프로파일, 하단 : 2-D 프로파일, 좌측 : 기존방식, 우측 : 제안방식

험을 통하여 제안 알고리즘의 동작을 확인하고 성능을 비교, 분석하였다. 제안 알고리즘은 자유도가 6인 정렬방식으로서, 기존의 정렬 방식인 자유도 4의 similarity 변환에 비해 보다 정확한 검출 성능을 보여주고 있음을 확인하였다. 이와 함께, 2-D 프로파일 모델을 구하는 과정에서 공분산 행렬을 trimming하는 방법을 채택함으로써 성능저하를 발생시키지 않으면서 계산량을 크게 줄일 수 있다. 제안 알고리즘은 AAM[3] 등에도 그대로 적용할 수 있으므로 이의 성능 향상에도 기여할 것으로 기대된다.

참 고 문 헌

[1] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active Contour Models," *Proceedings, First International Conference on Computer Vision*, pp. 259-268, IEEE Computer Society Press, 1987.

[2] T. F. Cootes, C. J. Taylor, and J. Graham, "Active Shape Model-Their Training and Application," *Computer Vision and Image Understanding* Vol.61, No.1, pp. 38-59, 1995.

[3] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active Appearance Model," *IEEE Trans. PAMI*, Vol.23, No.6, pp. 681-685, 2001.

[4] T. F. Cootes and C. J. Taylor, "Statistical Models of Appearance for Computer Vision," *Technical report, University of Manchester*, 2004.

[5] H. Xu and Z. Ma, "An Improved Active Shape Model for Facial Feature Location," *Second International Symposium on Intelligent Information Technology Application*, pp. 114-118, 2008.

[6] Z. Zheng, J. Jiong, D. Chunjiang, X. Liu, and J. Yang, "Facial Feature Localization Based on an Improved Active Shape Model," *Information Sciences*, Vol.178, No.9, pp. 2215-2223, 2008.

[7] H. Huang and S. Hsu, "Improved Active Shape Model for Facial Feature Localization," <http://repository.tudelft.nl>, 2011.

[8] M. H. Mohammad, A. M. Mohamed, and A. Ansari, "Improved Active Shape Model for Facial Feature Extraction in Color Images," *Journal of Multimedia*, Vol.1, No.4, pp. 21-28, 2006.

[9] James E. Gentle, *Numerical linear algebra for applications in statistics*, Springer, 1998.

[10] S. Milborrow and F. Nicolls, "Locating Facial Features with an Extended Active Shape Model," *ECCV'08 Proceedings of the 10th European Conference on Computer Vision: Part IV*, pp. 504-513, 2008.

[11] P. Viola and M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," *Computer Vision and Pattern Recognition Conference*, pp. 511-518, 2001.

[12] 유주환, 이종민, 김희율, "정규화된 형상 모델을 이용한 뼈 나이 측정 방법," *한국멀티미디어학회논문지*, 제12권, 제3호, pp. 383-396, 2009.



한 희 일

1980년 3월~1984년 2월 서울대학교 제어계측 공학과 공학사
 1984년 3월~1986년 2월 서울대학교 제어계측 공학과 공학석사

1992년 8월~1995년 12월 University of Arizona 전기및 컴퓨터 공학과 공학박사
 1987년 1월~1998년 3월 한국전자통신연구원, 선임연구원
 1998년 3월~현재 한국외국어대학교 정보통신공학과 교수
 관심분야: 신호처리, 영상처리, 컴퓨터비전, 패턴인식