

형태 군집화를 이용한 입술 형태 모델과 입술 추출

장 경 식^{*}

요 약

이 논문은 입술의 모양을 효과적으로 추출하는 방법을 제안하였다. 입술은 Point Distribution Model에 근거하여 점들의 집합으로 표현하였다. Isodata 군집 알고리즘을 이용하여 전체 학습 영상을 입술 형태별로 군집화하고 주성분 분석법을 사용하여 각 군집에 대한 입술의 형태 모델을 구하였다. 추출 결과가 입력 영상의 실제 입술 위치를 올바르게 찾았는지 판정하기 위하여 입술 경계선 주변의 화소값들을 이용한 입술의 경계선 모델을 구하고 이를 반영하는 평가함수를 구성하였다. 형태 차이를 반영하여 얻은 입술 형태 모델을 사용하여 입술을 추출하기 때문에 전체 학습 영상의 평균 모양과 많은 차이를 보이는 입술을 정확하게 추출할 수 있었다. 여러 영상을 대상으로 실험하여 약 92%의 성공률을 얻었다.

Lip Shape Model and Lip Localization using Shape Clustering

Kyung Shik Jang^{*}

ABSTRACT

In this paper, we propose an efficient method for locating lip. The lip shape is represented as a set of points based on Point Distribution Model. We use the Isodata clustering algorithm to find clusters for all training data. For each cluster, a lip shape model is calculated using principle component analysis. For all training data, a lip boundary model is calculated based on the pixel values around the lip boundary. To decide whether a recognition result is correct, we use a cost function based on the lip boundary model. Because of using different models according to the lip shapes, our method can localize correctly the lip far from the mean shape. The experiments have been performed for many images, and show correct recognition rate of 92%.

Key words: lip localization, principal component analysis, Isodata clustering algorithm

1. 서 론

최근에 변화하는 입 모양을 추출하거나 추적하는 연구는 많은 관심을 끌고 있다. 말할 때 변화하는 입 모양 변화를 음성 데이터와 같이 사용하여 음성을 인식하면 인식을 개선할 수 있음이 알려져 있으며[1, 2], 그래픽 애니메이션(animation) 분야에서 화자의 입 모양을 쉽게 실시간으로 생성할 수 있다[3]. 또한, 변화하는 입 모양은 각 개인을 식별하는 방법으로 이용될 수 있다[4].

입술 형태를 추출하는 연구로서 Mlrhosseinl[5]은

접수일 : 2003년 3월 31일, 완료일 : 2003년 6월 12일

^{*} 정회원, 동의대학교 멀티미디어공학과 조교수

입술을 구성하는 에지(edge)의 각도, 방향, 곡률(curvature) 등 입의 형태를 반영하는 조건들을 만족하는 에지들을 찾음으로써 입술의 경계선을 추출하였다. Oliver[3]는 색 정보에 대한 입술 모델을 설정하고 입의 형태적 특성을 반영하는 특징들과 시간에 따른 입술의 형태 변화를 HMM(Hidden Markov Model)을 이용하여 모델링하고 추출하였다. Wark[6]는 에지 정보와 색 정보를 이용하여 위 입술과 아래 입술의 경계선을 각각 4차 방정식과 2차 방정식으로 근사화하여 입술의 경계선을 추출하였다. Basu[7]는 얼굴과 입술을 구성하는 화소들의 색과 공간적 배치의 연관성을 이용한 특징을 선정하여 입술 영역을 찾고 칼만 필터(Kalman filter)와 입술 색 모델을 사용하

여 입술을 추출하였다. Delmas[8]와 Lievin[9,10]은 색 정보와 MRF(Markov Random Process)에 기반한 Bayesian segmentation을 이용하여 입술 영역을 찾고 snake를 이용하여 입술의 경계선을 추출하였다. Luettin[11]과 Iain[4]는 ASM(Active Shape Model)을 기반으로 PDM(Point Distribution Model)과 주성분 분석법(principle component analysis)을 이용하여 입술 형태 모델을 구성하고 입술을 구성하는 각 점 인근 영역을 탐색하여 형태계수(shape parameter)와 자세계수(pose parameter)를 변화시키며 입술을 추출하였다.

색 정보를 이용하는 방법은 입술 색의 개인별 차이, 조명의 영향을 많이 받으며, 에지 및 투영을 이용하는 방법은 입술 주변에 많은 에지가 존재하기 때문에 입술의 정확한 경계선을 찾기 어렵다. snake를 이용하는 방법[8,9]의 경우 계수설정 및 수렴성(convergence)의 기준 등이 잘못되는 경우 오류를 발생시킬 수 있다. 또한, 많은 연구들이 입술이 다물어진 상태에서 입술의 바깥쪽 경계선을 찾았으며 이는 입을 벌린 경우 입 안쪽의 치아와 혀로 인하여 입술 안쪽의 경계선을 찾는 것이 어렵기 때문이다. Delmas의 방법[8]에서는 입술 바깥쪽 경계선의 추출 성공률은 높았지만 입을 벌린 경우에 안쪽 경계선의 추출 성공률은 낮았다. 이에 반해 ASM을 이용한 방법[11]은 입술을 벌린 경우에도 안쪽 경계선의 추출에 높은 추출 성공률을 보였다. 그러나 이 방법은 학습 영상을 대표하는 하나의 입술 모델만을 사용하기 때문에 학습 데이터의 평균 모양에서 많이 벗어난 형태의 입술 추출에는 실패하는 경우가 많았다.

본 논문은 입술의 형태를 효과적으로 추출하는 방법을 제안하였다. ASM을 이용하여 입술을 추출하는 기존의 방법에서는 전체 학습 데이터에서 하나의 입술 모델을 구하여 사용하기 때문에 평균모양에서 많이 벗어난 형태의 입술은 올바르게 추출하지 못한다. 이 논문에서는 전체 학습 영상을 Isodata 군집 알고리즘을 이용하여 입술을 형태별로 군집화하고 각 군집에 따른 입술 모델을 구하였으며, 입술을 구성하는 각 점 주변의 화소값 분포를 이용하여 전체 학습 영상에 대한 하나의 입술 경계선 모델을 구성하였다. 각 군집의 입술 형태 모델을 이용하여 추출한 결과 중에서 실제 입술 위치를 가장 올바르게 찾은 결과를 얻기 위하여 입술 경계선 모델을 이용한 평가

함수를 사용하였다. 형태 차이가 있는 입술을 각기 다른 입술 형태 모델로 표현하기 때문에 전체 학습 영상의 평균 모양과 많은 차이가 있는 입술을 정확하게 추출할 수 있다. 여러 입술 영상에 대해 실험한 결과 기존의 방법으로 추출하지 못하는 패턴의 입술을 정확히 추출하였으며 전체적으로 약 4.4%의 성공률 향상을 얻었다.

2. 입술 분석

2.1 입술 형태 군집화

PDM[12]을 이용하여 점들의 집합으로 입술의 형태를 표현하였다. 입술과 주변 피부와의 경계선을 따라 그림 1과 같이 바깥쪽 입술선 위에 21개, 안쪽 입술선 위에 20개의 점을 위치시켰다. 이때 각 점은 입술의 정해진 위치에 있게 하였다.

학습 데이터의 i 번째 입술은 식(1)과 같이 41개 점들의 좌표로 구성되는 벡터로 표현하였다. 학습 영상에 있는 입술의 크기, 형태가 다양하기 때문에 동일 크기, 방향을 갖고 무게중심 좌표가 동일하도록 이들을 스케일(scale), 평행이동(translation), 회전(rotation)시키는 선형 기하학적 변환을 통하여 정규화하였다[4]. 기존의 방법에서는 전체 학습 데이터를 하나의 군집으로 간주하기 때문에 평균 입술 형태와 많은 차이가 있는 입술을 추출하지 못하는 경우가 발생한다. 그림 2는 200개 학습 영상의 입술을 식(1)과 같이 벡터로 표현하고 주성분 분석법을 적용하여 2개 주성분(V_1, V_2)을 취하고 이들로 이루어지는 공간에 벡터 데이터를 투영하여 나타낸 것이다.

$$X_i = (x_{1i}, y_{1i}, x_{2i}, y_{2i}, \dots, x_{41i}, y_{41i})^T \quad (1)$$

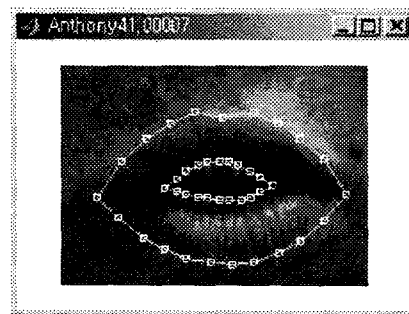


그림 1. 입술 표현

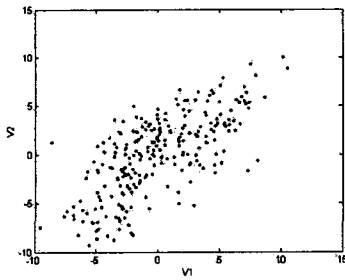


그림 2. 입술 형태 분포

패턴은 하나의 개념, 물체 등을 나타내는 개체이며, 어떤 패턴의 feature는 그 패턴의 특징을 나타내는 구분 요소들이다. 임의의 패턴 P_i 의 feature들을 $f_{ik}(k=1, \dots, N; N$ 은 총 feature의 수)로 나타낼 때, P_i 는 N 차원의 공간에서의 한 점으로 표현될 수 있다. 군집 알고리즘은 N -feature 공간에 놓인 패턴들을 유사한 것들끼리 묶어주는 알고리즘으로서, Maximum-Distance 알고리즘, K-means 알고리즘, Iso-data 알고리즘 등이 있다[13]. 이 논문에서는 식(1)의 입술을 구성하는 점들의 좌표로 구성되는 벡터를 하나의 패턴으로 사용하여 전체 학습영상들에 대해 군집화를 수행하였다. 데이터에 대한 군집의 수를 알지 못하기 때문에 Isodata 군집 알고리즘을 사용하였으며, 200개의 데이터를 이용하여 군집화한 결과 전체 데이터는 4개의 군집으로 나타났다. 그림 3은 군집화 결과의 예이다.

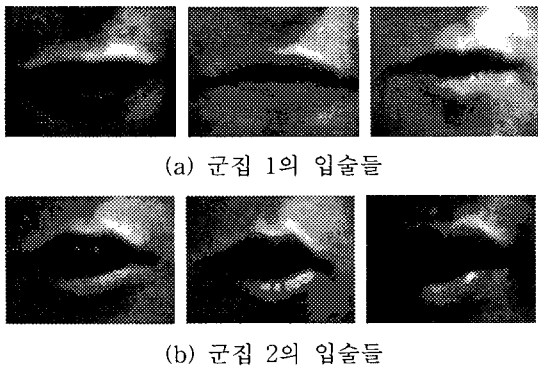


그림 3. 각 군집별 입술 예

2.2 군집별 입술 형태 모델

j 번째 군집에 속하는 벡터들의 평균 벡터 \bar{X}_j , 공분산 행렬(covariance matrix) Σ_j 을 구한 후 식(2)에

따라 공분산 행렬의 고유 벡터(eigen vector)와 고유 값(eigen value)을 구하였다. 이 식에서 p_{ij} 는 j 번째 군집에 대한 i 번째 고유 벡터로서 $p_{ij}^T p_{ij} = E$ 를 만족하는 직교행렬이고 λ_{ij} 는 p_{ij} 에 대응되는 고유값이며, 이때 E 는 단위행렬이다. j 번째 군집의 입술 모델은 식(3)에 의하여 표현된다. 이 식에서 P_j 는 고유값을 크기 순으로 정렬하고 가장 큰 값을 갖는 t 개의 고유값에 대응하는 고유벡터를 취하여 구성하였다.

t 는 $\sum_{i=1}^t \lambda_{ij} \geq V \sum_{i=1}^{\infty} \lambda_{ij}$ 에 의하여 결정하며 V 는 상수이다. b_j 는 고유 벡터들에 대한 가중치(weight)들로 구성되는 벡터로서 형태 계수(shape parameter)이다. 그림 4의 (a), (b)는 각각 군집 1과 군집 2의 입술 형태 모델에 의하여 생성한 입술 인스턴스(instance)로 가운데는 평균 입술 모양이고 좌우는 각각 $-2\sqrt{\lambda_{1j}}$, $2\sqrt{\lambda_{1j}}$ 만큼 변형된 입술 모양이다.

$$\Sigma_j p_{ij} = \lambda_{ij} p_{ij} \tag{2}$$

$$X_j = \bar{X}_j + P_j \cdot b_j \quad (j=1, \dots, 4)$$

where, $P_j = \begin{bmatrix} p_{1j} & p_{2j} & p_{3j} & \dots & p_{tj} \end{bmatrix}, \lambda_{ij} \geq \lambda_{(i+1)j}$

$$b_j = (b_{1j}, b_{2j}, \dots, b_{tj})^T \tag{3}$$

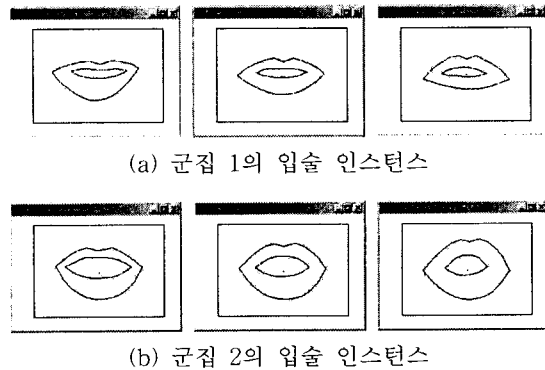


그림 4. 각 군집별 평균 입술과 변형된 입술

2.3 입술 경계선 모델

입술을 구성하는 41개 각 점에 인접한 영역의 화소값 분포를 이용하여 입술 경계선 모델을 설정하였다. 그림 5와 같이 학습 영상에서 입술을 구성하는 j 번째 점을 지나며 입술 경계선에 수직인 방향으로 현재 점을 중심으로 양쪽에서 n 개씩과 현재 점을 포

함하여 $(2n+1)$ 개 점을 선택하고 식(4)와 같이 그 점들의 화소값들로 구성되는 벡터(g_j)를 구하였다. 화소값을 직접 사용하는 경우에 조명 밝기 변화에 따른 영향을 받기 때문에 이를 최소화하기 위하여 식(4)에 따라 정규화 벡터(g_j^n)를 구하였다. 학습에 사용된 모든 입술 영상에 대해 j 번째 점에서의 경계선에 대한 평균벡터($\overline{g_j^n}$)와 공분산 행렬(S_j)을 식(5)에 따라 구하였다. 이 과정은 입술을 구성하는 모든 점에 대해 수행되어 입술 전체에 대한 경계선 모델을 구성하며 전체 학습 영상에 대해 하나의 경계선 모델을 구성하였다.

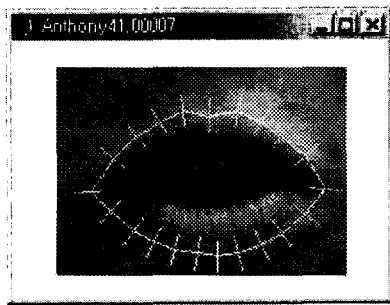


그림 5. 입술 경계선 모델

$$g_j = (g_{1j}, g_{2j}, \dots, g_{(2n+1)j})^T \text{ where, } g_{kj} \text{ 는 화소값}$$

$$g_j^n = \frac{g_j}{\sum |g_{kj}|} \text{ where, } g_{kj} = g_{(k+2)j} - g_{kj} \text{ (} k=1, \dots, 2n-1 \text{)}$$

(4)

$$S_j = \frac{1}{m} \sum_{i=1}^m (g_j^i - \overline{g_j^n}) (g_j^i - \overline{g_j^n})^T \text{ where, } m \text{ 은 학습 영상의 수}$$

(5)

입술 경계선 모델을 이용하여 입력 영상에서 입술을 탐색하는 과정은 입술 형태 모델을 이용하여 생성된 인스턴스를 구성하는 각 점이 자신이 가진 경계선 모델과 가장 유사한 화소값 분포를 갖는 영상 위의 점으로 이동함으로써 수행된다. 이를 위하여 인스턴스의 j 번째 점이 있는 영상위치에서 정규화된 화소값의 분포 H_j^n 를 구한다. 이 과정은 각 점의 경계선 모델을 구하는 과정과 같으며, H_j^n 는 현재 점을 중심으로 양쪽에서 $k(k > n)$ 개의 화소 값을 취하여 구성하였다. 이후 H_j^n 에서 차례로 $(2n+1)$ 개 값을 취하여 h_j^n 를 구성하고 이들 중에서 j 번째 점에서의 경계선 모델인 $\overline{g_j^n}$ 와 가장 비슷한 부분을 찾았다. 이를 위하여

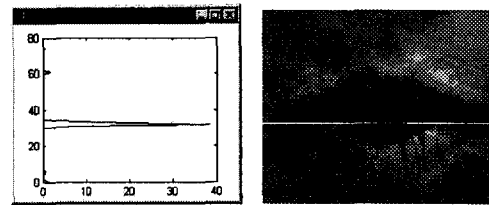
식(6)의 마할라노비스 거리(mahalanobis distance) 척도를 최소로 하는 h_j^n 의 가운데 부분에 해당하는 점을 이동할 점으로 정하였다. 이 과정은 인스턴스를 구성하는 모든 점에 대해 수행된다.

$$D_j = (h_j^n - \overline{g_j^n})^T S_j^{-1} (h_j^n - \overline{g_j^n}) \tag{6}$$

3. 입술 추출

3.1 입술 위치 설정

입술을 추출하는 첫 단계로서 위 입술과 아래 입술 사이의 경계선을 찾고 그 위치를 중심으로 입술을 탐색하였다. 영상의 각 열에서 식(7)을 이용하여 가장 어두운 화소들을 찾고 식(8)을 이용하여 입술 영상의 높이에 해당하는 크기를 갖는 벡터인 $L(y)$ 를 구하면 가장 큰 값을 갖는 위치에서 두 입술 사이의 경계선을 얻는다. 이 식들에서 $I(x, y)$ 는 (x, y) 위치에서의 화소값을 나타내고 K_1, K_2 는 상수이며, H_{Lip}, W_{Lip} 는 각각 영상의 높이와 폭이다. 그림 6의 (a)는 $L(y)$ 를, (b)는 두 입술 사이의 경계선을 나타낸 것이다. 그림 6의 (a)에서 x축은 $L(y)$ 값, y축은 영상 높이이다.



(a) (b)

그림 6. 입술 위치 검출

$$M(x) = \arg \min_{y \in H_{Lip}} I(x, y) \tag{7}$$

$$L(y) = \sum_{x \in W_{Lip}} \frac{(K_1 - \cosh \left(\frac{M(x) - \frac{H_{Lip}}{2}}{\frac{H_{Lip}}{K_2}} \right))}{K_1} \tag{8}$$

3.2 입술 추출

입술을 구성하는 각 점 주변의 정해진 영역들을 탐색하여 각 점은 자신이 갖는 입술 경계선 모델과

가장 비슷한 화소 분포를 갖는 점으로 이동하는 과정을 반복 수행하여 입술을 추출한다. j번째 군집의 입술 모델을 이용하여 입력 영상에서 입술을 찾는 과정은 다음과 같다.

```

bj = b = 0;      t = (W/2, YLip);
while(1) {
    Xj =  $\overline{X}_j + P_j \cdot b_j$ ;
    Xj = T(θ,t,s)(Xj);
    search Y;
    E = | T(θ,t,s)( $\overline{X}_j$ ) - Y |2를 최소화하는
        (θ,t,s)를 찾는다.;
    y = T(θ,t,s)-1(Y);
    y =  $\frac{y}{y \cdot \overline{X}_j}$ ;
    bj = PjT(y -  $\overline{X}_j$ );
    if( (Dmj =  $\sum_{i=1}^t (\frac{b_{ij}^2}{\lambda_{ij}})$ ) > Dmax)
    then bj = bj ·  $\frac{D_{max}}{D_{mj}}$ ;
    if( bj 와 (θ,t,s)가 수렴 ) then break;

```

이 과정에서 W는 영상의 폭, Y_{Lip}은 3.1절에서 구한 두 입술 사이 경계선의 y 좌표이며 D_{max}는 상수이다. T_(θ,t,s)(X_j)는 평행 이동량 t, 회전각 θ, 스케일 비 s를 사용하여 X_j를 선형 기하학적 변환한다. 인스턴스 X_j의 각 점에서 입술 경계선 모델을 이용하여 입술의 경계선에 해당하는 새로운 점을 찾으며, Y는 이들 점으로 구성되는 입술 모양이다. 각 점이 독립적으로 이동하기 때문에 이들 점으로 구성되는 전체적인 모양이 전형적인 입술 형태와 다르게 되므로 주어진 조건을 만족하지 못하는 경우 학습 데이터에서 얻어진 범위내의 형태계수를 찾고 적용하여 모양을 조정한다. 그림 7은 입술 추출과정 예로서 (b)는 (a)에서 각 점이 이동하여 Y를 구성한 것이고 (c)는 (b)

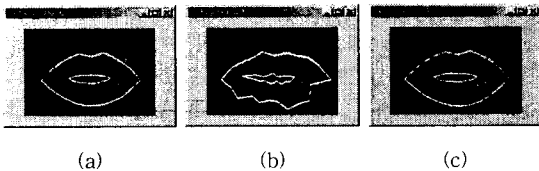


그림 7. 입술 추출 과정

에 대한 b값을 조정하여 입술 모양을 조정한 것이다.

3.3 평가 함수에 의한 인식 결과 판정

4개 군집의 입술 형태 모델을 이용하여 입력 영상에서 입술을 찾은 결과들에서 실제 입술 위치를 가장 잘 찾은 결과를 최종 인식결과로 판정하였다. 이를 위하여 2.3 절에서 각 점에 대한 경계선 모델을 구하는 것과 같은 방법으로 식(4)를 이용하여 추출 결과의 입술을 구성하는 각 점에서 화소값 분포를 구한다. 이때 경계선 모델을 구할 때와 같이 (2n+1)개의 점을 선택한다.

이 논문에서 사용한 평가함수는 식(9)와 같다. 이 식에서 g_{jk}ⁿ은 k번째 결과의 j번째 점에서의 화소값 분포이며 \overline{g}_j^n 은 j번째 점에서의 경계선 모델에 대한 평균이다. S_j는 j번째 점에서의 경계선 모델에 대한 공분산 행렬이다. 이러한 평가함수는 입술 모델을 구성하는 j번째 점에서의 화소값 분포와 추출된 입술을 구성하는 j번째 점에서의 화소값 분포와의 마할라노비스 거리를 구하고 이 과정을 모든 점들에 대해 수행하여 얻은 마할라노비스 거리들의 합을 나타낸 것으로 이 평가함수의 값이 가장 작은 결과를 최종 결과로 판정하였다.

$$D_k = \sum_{j=1}^4 (g_{jk}^n - \overline{g}_j^n)^T S_j^{-1} (g_{jk}^n - \overline{g}_j^n) \quad (k=1, \dots, 4) \tag{9}$$

4. 실험 및 고찰

실험을 위하여 Tulip 1 데이터베이스[14]에 있는 입술 영상들을 사용하였다. 이들은 다양한 조명상태에서 총 12명(남자 9명과 여자 3명)이 4개의 숫자를 두 번씩 발음할 때 입력된 연속적인 명암 영상(grey image)으로서 첫 번째 발음할 때의 영상들을 집합 1, 두 번째 발음할 때의 영상들을 집합 2라고 지칭한다. 이들은 입술 영역만을 포함하고 있으며 100*75 크기를 갖는다. 이 논문에서는 학습 영상으로 집합 1에서 200개의 입술 영상을 사용하였으며 실험 영상으로 집합 2에서 250개의 영상을 사용하였다. 펜티엄(Pentium) 4 PC(1.6 GHz)에서 실험하였으며 Matlab으로 구현하였다.

그림 8은 두 입술 사이의 경계선 추출 결과이다.

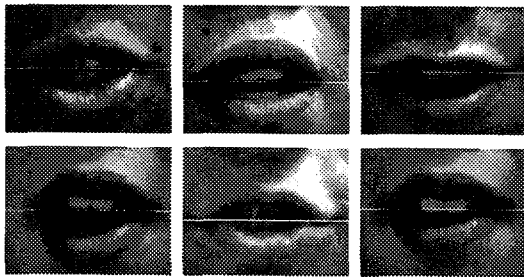
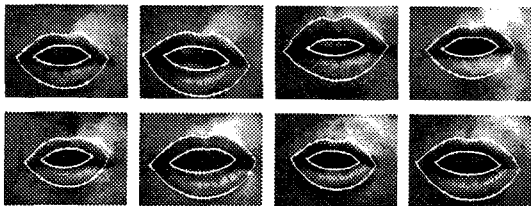


그림 8. 입술 경계선 검출

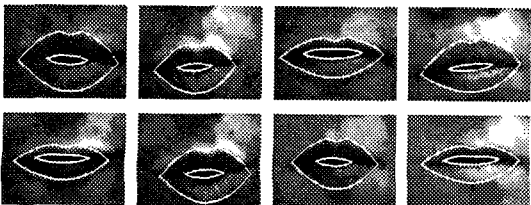
입술이 다물어진 경우뿐만 아니라 치아와 혀가 보이는 경우에도 정확하게 경계선을 추출하였다. 그림 9는 입술 추출 결과의 예이다. 입을 열어서 치아와 혀가 보이는 경우, 입을 다물어서 치아와 혀가 보이지 않는 경우 등 여러 상황에서 입술을 정확하게 추출하였다. 결과에서 보듯이 전반적으로 어두운 경우와 밝은 경우 등 조명 영향은 없었다. 실험 결과 모든 경우에 추출결과의 위치가 입술의 실제 위치와 많이 차이가 나는 결과는 없었으며, 오류인 경우는 실제 입술 경계선에서 조금 떨어지거나 정확한 입술 폭의 추출 실패로 인하여 전체적인 입술 모양이 실제 모양과



(a) 치아가 보이는 경우



(b) 치아 혹은 혀가 보이는 경우



(c) 치아와 혀가 모두 안 보이는 경우

그림 9. 입술 추출 결과

다르게 추출된 경우가 발생하였다. 입술 폭 추출 오류는 안쪽 경계선 폭이나 바깥쪽 경계선 폭의 추출 오류를 나타낸다. 실험결과를 분석하기 위하여 추출한 입술의 경계선이 실제 위치와 일정 거리 이상 떨어진 경우를 오류로 판정하였다. 이때 거리의 기준으로 윗입술 높이의 1/2을 사용하였다.

바깥쪽 경계선의 위 부분에 대한 추출 오류는 없었으며, 아래 부분에서 오류가 발생하였다. 이는 일반적으로 아래 입술과 인접 피부와의 경계가 윗입술과 인접 피부와의 경계보다 뚜렷하지 않기 때문이다. 오히려 입술의 폭을 잘못 추출하여 추출 결과가 실제 입술보다 옆으로 긴 형태의 오류가 많았다. 이는 입술의 위와 아래 부분에서 인접 피부와의 경계보다 입술 양 끝점에서 인접 피부와의 경계가 매우 불확실하기 때문이다. 안쪽 경계선의 추출 오류도 발생하였는데 이는 입을 벌린 경우 입술 사이에 보이는 치아와 혀로 인하여 경계가 많이 나타나기 때문이다.

이 논문에서는 입술을 추출하는 방법 중에서 가장 효과적이라 판단되는 Luettin의 방법[11]에서 본 논문과 같이 입술을 구성하는 점들이 독립적으로 이동하며 입술을 찾는 방법과 본 논문에서 제안한 방법의 결과를 비교한 것이다. Luettin의 방법에서는 입술을 다문 경우와 입술이 열린 경우 각각에 대한 입술 모델을 설정하였으나 이들을 구별하여 입력 영상에 적용하는 방법을 제시하지 못했기 때문에 이 논문에서는 입술을 다문 경우와 입술이 열린 경우 모두 하나의 입술 모델을 설정하였다. 그림 10은 Luettin의 방법과 제안한 방법의 결과 비교 예이다. (a)는 제안한 방법으로 추출한 결과이고 (b)는 기존의 방법으로 추출한 결과이다. 제안한 방법이 보다 실제 형태에 가깝게 추출하였음을 알 수 있다.

그림 11은 추출 오류의 예이다. (a)는 아래 입술을



(a) 제안한 방법



(b) 기존 방법

그림 10. 결과 비교

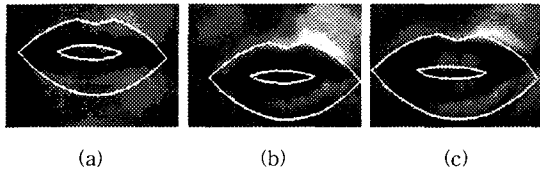


그림 11. 추출 오류

입 안쪽으로 넣은 상태로 입술의 바깥쪽 경계선 아래 부분을 추출하지 못한 경우이고 (b)는 입술의 양 끝 점을 추출하지 못한 경우이다. (c)는 벌린 입 사이로 보이는 치아로 인하여 안쪽 경계선을 잘못 추출한 결과이다.

표 1은 실험 결과를 나타낸 것이다. 경계선과 입술 폭에서의 오류가 동시에 발생한 경우는 입술 폭 추출 오류로 판단하였다. 표에서 보듯이 제안한 방법이 기존 방법에 비하여 약 4.4%의 성공률 향상을 얻었다.

표 1. 실험 결과

		추출 성공	추출실패			전체
			입술 바깥쪽 경계선	입술 안쪽 경계선	입술폭	
제안한 방법	영상 개수(개)	230	2	7	11	250
	비율(%)	92.0	0.8	2.8	4.4	100
기존 방법	영상 개수(개)	219	3	8	20	250
	비율(%)	87.6	1.2	3.2	8.0	100

6. 결 론

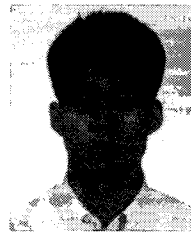
본 논문은 입술의 형태를 효과적으로 추출하는 방법을 제안하였다. 전체 학습 데이터에서 한 입술 모델을 구하여 사용하면 입술 형태 변화가 심한 경우에는 추출을 올바르게 하지 못하는 경우가 생긴다. 이 논문에서는 전체 학습 영상을 Isodata 군집 알고리즘을 이용하여 입술의 형태별로 군집화하고 각 군집의 입술 형태 모델을 각각 구하였다. 입술 경계선에서의 화소값 분포에 근거한 입술 경계선 모델을 구성하고 이를 반영하는 평가함수를 구성하여 각 군집의 입술 형태 모델을 이용하여 추출한 결과 중에서 입력 영상의 실제 입술 위치를 가장 올바르게 찾은 결과를 얻기 위한 평가 함수로 사용하였다. 많은 형태 차이가

있는 입술을 각기 다른 입술 모델로 표현하여 추출하기 때문에 입술 형태가 전체 학습 영상의 평균 모양과 차이를 보이는 입술을 정확하게 추출할 수 있었다. 여러 입술 영상에 대해 실험한 결과 기존의 방법으로 추출하지 못하는 패턴의 입술을 정확히 추출하였다. 입술을 구성하는 점들의 특징을 이용하여 탐색하는 방법을 개선하고 입을 다문 경우와 열린 경우를 구별하여 입술 모델을 설정하면 보다 높은 추출 성공률을 얻을 것으로 생각된다.

참 고 문 헌

- [1] Yang J., R. Stiefelhagen, U. Meier and A. Waibel, "Real-time Face and Facial Feature Tracking and Application," *Proceedings of Auditory-Visual Speech Processing*, pp. 79-84, 1998.
- [2] Kaucic R., A. Blake, "Accurate, Real-Time, Unadorned Lip Tracking," *Proceedings of the 6th International Conf. on Computer Vision*, pp. 370-375, 1998.
- [3] Oliver N., A. Pentland, "LAFTER: Lips and Face Real Time Tracker," *Proceedings of the 1997 Conf. on Computer Vision and Pattern Recognition*, (1997), pp. 123-129.
- [4] Iain Matthews, Timothy F. Cootes, J. Andrew Banghan, Stephen Cox and Richard Marvey, "Extraction of Visual Features for Lipreading," *IEEE Tans. on Pattern Recognition and Machine Analysis*, Vol 24, No. 2, , pp. 198-213, Feb. 2002.
- [5] Mirhosseinl A. R., H. Yan and K. M. Lam, "Adaptive Deformable Model for Mouse Boundary Detection," *Optical Engineering*, Vol. 37 No. 3(1998), pp. 869-875.
- [6] Wark T., Sridharan and V. Chandran, "An Approach to Statistical Lip Modelling for Speaker Identification via Chromatic Feature Extraction," *Proceedings of the 14th International Conf. on Pattern Recognition*, Vol. 1, pp. 123-125, 1998.
- [7] Basu S., N. Oliver and A. Pentlan, "3D

- Modeling and Tracking of Human Lip Motions," *Proceedings of the 6th International Conf. on Computer Vision*, pp. 337-343, 1998.
- [8] Delmas P., Y. Coulon and V. Fristot, "Automatic Snakes for Robust Lip Boundaries Extraction," *IEEE International Conf. on Acoustics, Speech and Signal Processing*, Vol. 6, pp. 3069-3072, 1999.
- [9] Lievin M., F. Luthon, "Unsupervised Lip Segmentation under Natural Conditions," *IEEE International Conf. on Acoustics, Speech and Signal Processing*, Vol. 6, pp. 3065-3068, 1999.
- [10] Lievin M., P. Delmas, Y. Coulon, F. Luthon and V. Fristot, "Automatic Lip Tracking : Bayesian Segmentation and Active Contours in a Cooperative Scheme," *IEEE Conf. on Multimedia, Computing and System*, pp. 691-696, 1999.
- [11] Luetttin J., N. A. Thacker and S. W. Beet, "Locating and Tracking Facial Speech Features," *Proceedings of The International Conf. on Pattern Recognition*, pp. 652-656, 1996.
- [12] M. B. Stegmann, R. Fisker, "On Properties of Active Shape Models," *Informatics and Mathematical Modelling*, Technical University of Denmark, 2000.
- [13] Morton Nadler and Eric P. Smith, *Pattern Recognition Engineering*, A Wiley-Interscience Publication, 1993.
- [14] Movellan J. R., "Visual Speech Recognition with Stochastic Networks," *Advances in Neural Information Processing System*. Vol. 7, MIT Press Cambridge, 1995.



장 경 식

1989년 연세대학교 전자공학과 졸업(학사)
 1991년 연세대학교 대학원 전자공학과 졸업(석사)
 1996년 연세대학교 대학원 전자공학과 졸업(박사)

1994년~1998년 2월 대우전자 전
 략기술 제 1 연구소 선임연구원
 1998년~현재 동의대학교 멀티미디어공학과 조교수
 관심분야: 컴퓨터 비전, 인공지능, Digital Contents
 Watermaking 등

E-mail : ksjang@dongeui.ac.kr

교 신 저 자

장 경 식 614-714 부산광역시 부산진구 가야동 산24 동의
 대학교 멀티미디어공학과