

능동적 형태 모델과 가중치 벡터를 이용한 입술 인식*

장경식

동의대학교 멀티미디어공학과
(ksjang@hyomin.donggeui.ac.kr)

이 논문은 입술의 형태와 입술 외곽선 부근의 화소값을 이용하여 입술을 효과적으로 인식하는 방법을 제안하였다. 입술의 형태는 학습 영상을 통계적으로 분석하는 능동적 형태 모델을 기반으로 구성하였다. 이 방법은 탐색시 초기 위치의 영향을 받기 때문에 이 논문에서는 입술의 형태에 기반한 가중치 벡터를 이용하여 두 입술 사이의 경계선을 찾고 탐색의 초기 위치로 사용하였다. 다양한 입술 영상들을 대상으로 실험하여 좋은 결과를 얻었다.

1. 서론

최근 휴먼 컴퓨터 인터페이스(human computer interface)를 개발하기 위한 분야로서 얼굴, 눈, 입의 인식 기술에 대한 연구가 많이 진행되고 있다. 이 분야에서 이전에 행해진 많은 연구들은 얼굴 영역의 추출에 중점을 두었기 때문에 추출된 얼굴 영역의 확인을 위하여 눈이나 입의 위치만을 찾았다[Samal and Iyengar, 1992; Sung and Poggio, 1998]. 그러나 근래 들어 시간에 따른 눈동자나 입술의 형태변화 및 움직임에 대한 정보를 추출하고 여러 분야에 응용하려는 연구가 활발히 진행되고 있으며, 이 중에서도 입 모양을 인식하거나 추적하는 연구는 많은 관심을 끌고 있다[Mirhosseinl et al., 1998; Oliver and Pentland, 1997].

입 모양 추출 결과는 여러 분야에 응용 가능하

다. 그래픽 애니메이션(animation)이나 가상 현실(virtual reality) 분야에서 화자의 입 모양을 보다 쉽게 실시간으로 생성할 수 있으며, 음성 데이터와 화자의 입 모양 변화를 같이 사용하여 음성을 인식하는 경우에 인식률에 많은 개선을 얻을 수 있음이 알려져 있기 때문에 이에 대한 연구가 많은 관심을 끌고 있다[Yang et al., 1998; Kaucic and Blake, 1998]. 또한, 사람이 말할 때 변화하는 입 모양을 각 개인을 식별하는 방법으로 사용하는 연구가 진행되고 있다[Wark et al., 1998].

Mirhosseinl[1998]은 입술을 구성하는 에지(edge)의 각도, 방향, 곡률(curvature) 등에 대하여 입의 형태를 반영하는 조건들을 정하고 이를 만족하는 에지들을 찾음으로써 입술의 외곽선을 추출하였다. Oliver[1997]는 색 정보를 이용하여 입술에 대한 모델을 설정하고 입의 형태적 특성을 반영하는 특징(feature)들을 사용하여 입술을

* 이 논문은 2001년 동의대학교 교내학술연구 지원사업에 의하여 연구되었음.

찾았으며, 시간에 따른 입술의 변화형태를 HMM (Hidden Markov Model)을 이용하여 모델링하고 인식하였다. Wark[1998]는 에지정보를 이용하여 위 입술과 아래 입술의 외곽선을 각각 4차 방정식과 2차 방정식으로 근사화하여 입술의 외곽선을 추출하였다. Basu[1998]는 얼굴과 입술을 구성하는 각 화소들의 색 정보와 공간적 배치의 연관성을 이용한 특징을 선정하여 얼굴과 입술 영역을 찾고 칼만 필터(kalman filter)와 입술 색 모델을 사용하여 입술을 추출하였다. Delmas [1999]와 Lievin[1999]은 색 정보와 HMM에 기반한 bayesian segmentation을 이용하여 입술 영역을 찾고 snake를 이용하여 입술의 외곽선을 추출하였다. Yang[1998]은 입술 색에 대한 모델을 구성하고 입력에 따라 모델을 변화시켜 입술 영역을 추출하였다.

색 정보를 이용하여 입술을 찾는 방법[Oliver, 1997; Basu, 1998; Lievin, 1999]의 경우에 입술 색은 개인별로 많은 차이를 보이며, 조명의 영향을 많이 받기 때문에 인식 오류의 가능성을 내포한다. 또한 에지 및 투영을 이용하는 방법[Mirhosseinl et al., 1998; Wark et al., 1998]에서는 입술과 주변의 피부와의 경계가 뚜렷하지 않기 때문에 입 모양의 정확한 외곽선을 찾기 어렵다. 또한, 조명이 변화하는 경우에 에지 추출이 어렵거나 사용하는 임계값을 경험적으로 구해야 하고 입술 크기 등이 변하는 경우에는 그 값을 변화시켜야 하는 어려움이 있다. snake를 이용하는 방법[Delmas, 1999; Lievin 1999]의 경우 초기화(initialization)와 계수(parameter) 설정 및 수렴성(convergence)의 기준 설정 등이 잘못되는 경우 오류를 발생시킬 수 있다.

이 논문은 입술의 형태와 입술 외곽선 부근의 화소값 분포를 이용하여 입술을 효과적으로 인식

하는 방법을 제안하였다. 입술의 형태는 학습 영상을 통계적으로 분석하는 능동적 형태 모델을 기반으로 표현하였다. 능동적 형태 모델은 탐색시 초기 위치의 설정에 따라 잘못된 결과를 얻을 수 있기 때문에 이 논문에서는 입술의 형태에 기반한 가중치 벡터를 이용하여 두 입술 사이의 경계선을 찾고 탐색의 초기 위치로 사용하였다. 입술을 인식하는 많은 방법들이 에지 및 색 정보를 기반으로 영상을 탐색하는 반면에 이 논문에서는 입술 외곽선 주위의 화소분포를 특징으로 정하고 이를 이용하여 영상을 탐색하였으며, 이 과정에서 특징의 정규화 과정을 거쳐 조명변화의 영향을 최소화시켰다. 다양한 입술 영상들을 대상으로 실험하여 좋은 결과를 얻었다.

2. 입술 분석

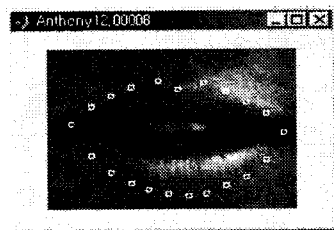
2.1 전체과정

이 논문에서 입술을 인식하는 과정은 입술 모델을 구성하는 단계와 입술을 인식하는 단계로 구성된다. 입술의 외곽선에 위치하는 점들의 좌표로 입술의 형태(shape)를 표현하였으며, 모든 학습 영상에 걸쳐 이들 점의 좌표들에 대한 평균(mean)과 공분산 행렬(covariance matrix)을 구하고 주성분 분석법(principal component analysis)을 사용하여 평균과 고유 벡터의 선형 함수로 입술 모델을 표현하였다. 또한 입술 모델을 구성하는 각 점에 인접한 영역의 화소값 분포를 이용하여 입술의 외곽선에 대한 특징을 설정하였다. 입술의 형태에 기반한 가중치 벡터를 이용하여 두 입술 사이의 경계선을 찾아 탐색의 초기 위치로 정하고 입술 모델을 구성하는 각 점에서의 특징을

이용하여 입술을 찾았다. 이 과정에서 평가함수로써 마할라노비스 거리(mahalanobis distance) 척도를 이용하였다.

2.2 입술 형태 모델링

능동적 형태 모델(active shape model)[Cootes et al., 1994; Luetin et al., 1996]을 이용하여 점들의 집합으로 입술의 형태를 표현하였다. 이 방법은 학습 영상들을 통계적으로 분석하고 이용하는 방법으로 사용자가 경험적 지식을 토대로 부여하는 여러 가지 제약의 사용을 배제할 수 있다. 입술과 주변 피부와의 경계선을 따라 <그림 1>과 같이 21개의 점을 위치시켰다. 이때 각 점은 입술의 정해진 위치에 있게 하였다. 예를 들면, 가장 왼쪽에 있는 점을 점 1로 정하고 시계방향으로 진행하며 번호를 부여하면, 점 1과 11은 항상 입술의 양쪽 끝에 놓이고 점 5, 6, 7은 입술의 위쪽에 위치한다.



<그림 1> 입술 모델링

$$X = (x_1, y_1, x_2, y_2, \dots, x_{21}, y_{21})^T \quad \text{식(1)}$$

입술은 식(1)과 같이 21개 점들의 좌표로 구성되는 벡터로 표현하였다. 학습 영상에 있는 입술의 크기 및 형태가 다양하기 때문에 입술을 구성하는 점들의 좌표 값은 영상마다 다르다. 따라서 이들을 대표하는 하나의 모델을 구성하기 위하여

동일 크기, 방향을 갖고 무게중심의 좌표가 동일하도록 이들을 스케일(scale), 평행이동(translation), 회전(rotation)시키는 선형 기하학적 변환(linear geometrical transform)을 통하여 정규화하는 과정이 필요하다. 방향은 입술의 양 끝점을 연결하는 직선과 x 축과의 각도를 의미한다. 이는 다음과 같은 과정을 거쳐 수행된다.

1. 각 벡터의 무게중심(\bar{x}, \bar{y})이 원점에 오도록 평행 이동한다.
2. 첫 번째 벡터를 크기가 1 되도록 스케일하고 ($X_1 = \frac{X_1}{|X_1|}$) 기준 벡터로 정한다.
3. 나머지 벡터들을 기준 벡터와 같은 크기, 방향을 갖고 무게 중심이 동일하도록 스케일, 회전, 평행 이동하고 이들의 평균 벡터(mean vector)를 구한다.
4. 평균 벡터를 크기가 1 되도록 스케일하고 기준 벡터와 같은 방향을 갖고 무게 중심이 동일하도록 변환하기 위한 회전각, 평행이동 양을 구한다.
5. 이 값들이 수렴하면 중지한다. 그렇지 않으면 평균 벡터를 기준 벡터로 정하고 단계 3으로 간다.

기준 벡터와 같은 크기, 방향을 갖고 무게중심이 동일하도록 나머지 벡터들을 정렬하는 과정은 다음과 같다. X' 이 기준 벡터이고 X 가 정렬시킬 벡터인 경우에 X 를 X' 에 대하여 정렬하는 것은 이들이 이차원 벡터인 경우에 식(2)의 E 를 최소화하는 $T_{(\theta, t, s)}(X)$ 를 구하는 것과 같다. 이때 s 는 스케일 비, θ 는 회전각이고, t 는 (t_x, t_y) 로서 X 와 X' 의 무게 중심간 차이이다. 우선 무게 중심이 동일하도록 이동하면 식(2)의 E 를 최소화하는 s, θ

는 식(3)과 같이 된다. 이 식들에서 $(x, y) \in X$, $(x', y') \in X'$ 이다.

$$E = |T_{(\theta, t, s)}(X) - X'|^2$$

$$\text{where, } T_{(\theta, t, s)} \begin{pmatrix} x \\ y \end{pmatrix} = s \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad \text{식(2)}$$

$$s \cos \theta = \frac{X \cdot X'}{|X|^2} \quad s \sin \theta = \frac{\sum_{i=1}^n (x_i y'_i - y_i x'_i)}{|x|^2} \quad \text{식(3)}$$

정규화된 벡터들의 평균 벡터를 구하고 식(4)에 따라 공분산 행렬을 구한 후 공분산 행렬의 고유 벡터(eigen vector)와 고유 값(eigen value)을 구하였다. 이 식에서 P_i 는 고유 벡터이고 λ_i 는 고유값이며 $P_i^T P_i = 1$ 을 만족한다. 입술 모델은 식(5)에 의하여 표현된다. 이 식에서 P는 고유값을 크기 순으로 정렬하고 가장 큰 값을 갖는 t 개의 고유값에 대응하는 고유벡터를 취하여 구성하였다. t 는 $\sum_{i=1}^t \lambda_i \geq V \sum_{i=1}^{2n} \lambda_i$ 식에 의하여 결정하며 V는 상수이다. b는 고유벡터들의 가중치(weight) 들로 구성되는 벡터이며, 식(6)에 의하여 결정된다. 식(5)를 이용하면 입술의 한 인스턴스(instance)는 평균벡터와 고유 벡터들의 선형조합으로 얻을 수 있다. 즉, 고유벡터들의 선형 조합(linear combination)만큼 평균 입술 모양을 변형 시킴으로써 입술의 인스턴스를 생성한다. <그림 2>는 모델에 의하여 생성한 평균 입술 모양과 변형된 입술 모양의 예이다. (a)는 평균 입술 모양이고 (b), (c)는 $\pm 3\sqrt{\lambda_1}$ 만큼 변형된 입술 모양을 나타낸다.

$$\Sigma = \frac{1}{m} \sum_{i=1}^m (X_i - \bar{X})(X_i - \bar{X})^T \quad \text{식(4)}$$

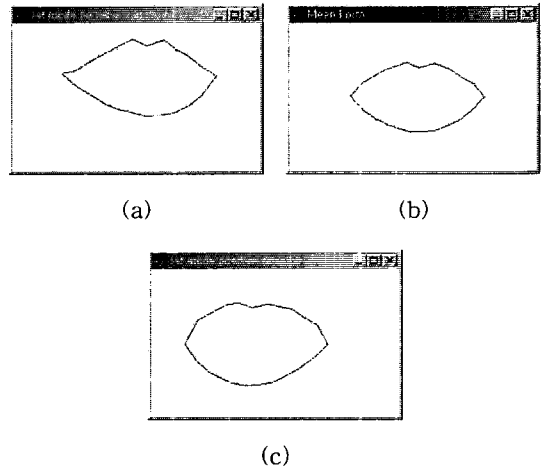
$$\Sigma P_i = \lambda_i P_i$$

$$X = \bar{X} + P \cdot b$$

$$\text{where, } P = \begin{bmatrix} P_1 & P_2 & P_3 & \dots & P_t \end{bmatrix}, \quad \lambda_i \geq \lambda_{i+1}$$

$$b = (b_1, b_2, \dots, b_t)^T \quad \text{식(5)}$$

$$b = P^T(X - \bar{X}) \quad \text{where, } -3\sqrt{\lambda_1} \leq b_i \leq 3\sqrt{\lambda_1} \quad \text{식(6)}$$



<그림 2> 입술 모델

2.3 화소값 분포를 이용한 입술선 특징 선정

입술과 피부와의 경계선에 대한 특징을 설정하고 이전 과정에서 구성한 입술 모델과 함께 이용하여 입력 영상에서 입술을 찾는다. 입술 모델을 구성하는 각 점이 입술의 외곽선에 정확히 존재하지 않고 입술 외곽선 부근에 다양한 예지가 존재하여 예지 정보를 사용하기 어렵기 때문에 각

점에 인접한 영역의 화소값 분포를 이용하여 입술 외곽선에 대한 특징을 설정하였다. <그림 3>과 같이 학습 영상에서 입술을 구성하는 각 점을 지나며 입술 외곽선에 수직인 방향으로 점의 양쪽으로 n 개씩과 자신을 합하여 (2n+1) 개 점을 선택하고 그 점들의 화소값들로 구성되는 벡터 g를 구하였다. 화소값만을 사용하는 경우에 조명 등의 밝기 변화에 따른 영향을 많이 받기 때문에 이를 최소화하기 위하여 식(7)에 따라 정규화하고 특징 벡터(g')를 구하였다. 학습에 사용된 입술의 동일 위치에 있는 점들에 대한 특징 벡터들이 정규분포를 따른다고 가정하고 식(8)을 이용하여 각 점에서의 특징들에 대한 평균벡터(\bar{g})와 공분산행렬(S_g)을 구하였다.

$$g = (g_1, g_2, \dots, g_{(2n+1)})^T$$

$$g = \frac{g}{g_{\max}} \quad \text{where, } g_{\max} = \max_{1 \leq i \leq (2n+1)} g_i$$

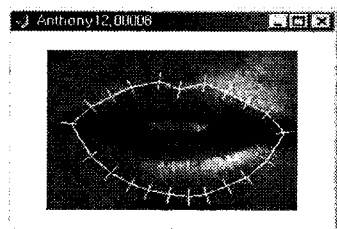
$$g' = \frac{g'}{\sum g'_i} \quad \text{where, } g'_i = g_{i+2} - g_i \quad (i=1, \dots, 2n-1)$$

식(7)

$$dg' = g' - \bar{g}$$

$$S_g = \frac{1}{m} \sum_{i=1}^m dg'_i dg'_i{}^T \quad \text{where, } m \text{은 학습 영상의 수}$$

식(8)



<그림 3> 화소값 분포 획득 구간

3. 입술 인식

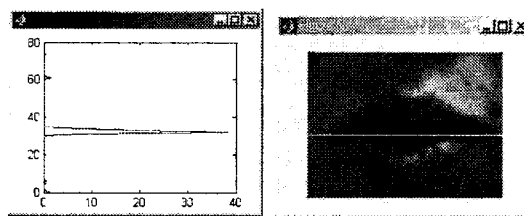
3.1 두 입술 사이의 경계선 설정

능동적 형태 모델을 이용하여 탐색하는 경우에 초기 위치의 선정에 따라 잘못된 결과가 나올 수 있다. 이 논문에서는 입술을 인식하는 첫 단계로서 두 입술 사이의 경계선을 찾고 그 위치를 중심으로 입술 외곽선을 탐색하였다. 영상의 각 열에서 식(9)를 이용하여 가장 어두운 화소들을 찾고 식(10)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 K₁, K₂는 상수이며, H_{Lip}, W_{Lip}은 각각 영상의 높이와 폭을 나타낸다. <그림 4>의 (a)는 L(y)를 나타낸 것이고 (b)는 두 입술 사이의 경계선을 나타낸 것이다.

$$M(x) = \arg \min_{y \in H_{Lip}} I(x, y) \quad \text{식(9)}$$

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

$$Y_{LipCenterLine} = \arg \max_{y \in H_{Lip}} L(y) \quad \text{식(10)}$$



(a) (b)

<그림 4> 입술선 검출

3.2 입술 인식

두 입술 사이의 경계선 위치가 정해지면 입술 모델을 그 위치로 평행 이동한 후 입술의 외곽선을 찾는 과정을 수행하였다. 이 과정은 모델을 구성하는 각 점의 특징을 이용하여 입술 경계선을 탐색하는 과정과 점들의 전체적인 모양을 허용범위 내로 조정하는 과정의 두 단계로 구성된다. 입술 모델을 구성하는 각 점의 특징을 사용하여 점 주변의 정해진 영역들을 탐색하고 각 점은 자신이 갖고 있는 입술선 특징과 가장 비슷한 특징을 갖는 점으로 이동한다. 이 결과 각 점이 독립적으로 이동하기 때문에 전체적인 모양이 입술과 다른 형태를 갖게 되기 때문에 식(5)의 입술 모델에서 고유벡터에 대한 가중치를 조정하여 입술 모델에서 생성할 수 있는 입술의 형태로 조정하였다. 입력 영상에서 입술을 찾는 과정은 다음과 같다.

1. $b=0$, (θ, t, s) 는 초기값을 설정한다. t 는 (t_x, t_y) 로서 평행 이동량을 나타내고 θ 는 회전각, s 는 스케일 비를 나타낸다. 이 논문에서는 t_y 는 이전 단계에서 구한 두 입술 사이의 경계선의 y 좌표를 사용하였으며, t_x 는 영상 폭의 1/2을 사용하였다. 또한 θ 는 0을 사용하였으며, s 는 임의의 값을 사용하였다.
2. 입술 모델을 표현하는 $X = \bar{X} + P \cdot b$ 식에 의하여 입술의 인스턴스를 생성하고 $T_{(\theta, t, s)}(X)$ 식에 의하여 스케일, 평행이동, 회전한다.
3. 인스턴스 각 점에서의 특징을 이용하여 입술의 경계선에 해당하는 새로운 점을 찾는다. 이들 점으로 구성되는 모양을 Y 로 정하고 Y 의 무게 중심을 원점으로 이동한다.
4. \bar{X} 에서 Y 로 되기 위한 (θ, t, s) 를 찾아서 이 값

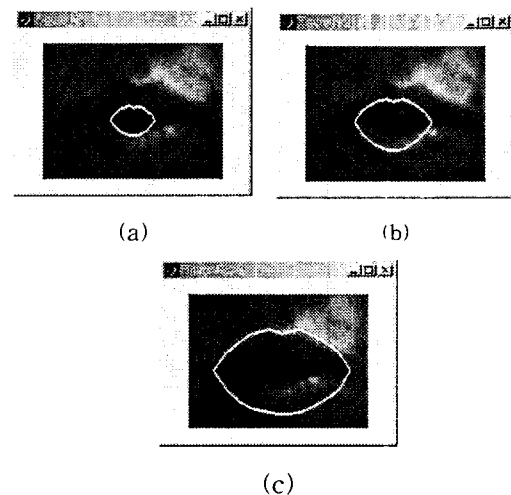
을 새로운 (θ, t, s) 로 정하고 $y = T_{(\theta, t, s)}^{-1}(Y)$ 식을 이용하여 역변환한 후 $y = \frac{y}{y \cdot \bar{X}}$ 식에 의하여 정규화한다.

6. $b = P^T(y - \bar{X})$ 식에 의하여 b 를 구하고 b 값이 식(11)을 만족하는지 조사한다. 만족하지 않으면 식(12)에 의하여 b 값을 재 조정한다. 이 식들에서 D_{max} 는 상수이다.
7. b 와 (θ, t, s) 가 수렴하면 중지한다. 그렇지 않으면 단계 2로 간다.

$$D_m = \sum_{i=1}^L \left(\frac{b_i^2}{\lambda_i} \right) \leq D_{max} \quad \text{식(11)}$$

$$b_i = b_i \cdot \frac{D_{max}}{D_m} \quad \text{식(12)}$$

<그림 5>는 입력 영상에서 입술을 찾는 과정이다. (a)는 모델이 초기 위치에 있는 상태이고 (b)는 각 점이 독립적으로 인근 영역을 탐색하여 새로운 위치로 이동한 후 전체적인 모양을 조정 한 결과이다. (c)는 여러 번의 반복적인 과정을 거쳐 최종적인 입술을 인식한 결과이다.



<그림 5> 입술 인식 과정

입술을 탐색하는 과정은 입력 영상에서 입술의 인스턴스를 구성하는 각 점에 인접 점들에서 특징을 추출하고 모델의 대응하는 점의 특징과 비교하여 가장 비슷한 특징을 갖는 위치로 각 점이 독립적으로 이동한다. 현재 입술의 인스턴스를 구성하는 각 점의 위치에서 화소값의 분포 G_s' 를 구하였으며, 이 과정은 입술 모델을 구성하는 각 점의 특징을 구하는 과정과 같다. 이때 G_s' 는 현재 점을 중심으로 양쪽에서 $m(m>n)$ 개의 화소 값을 취하여 구성하였다. 이후 G_s' 에서 차례로 $(2n+1)$ 개 값을 취하여 g_s' 를 구성하고 이들 g_s' 중에서 \bar{g} 와 가장 비슷한 부분을 찾았다. 이를 위하여 식(13)의 마할라노비스 거리(mahalanobis distance) 척도를 최소로 하는 g_s' 를 선택하고 g_s' 의 중간 부분에 해당하는 점을 이동할 점으로 정하였다. g_s' 는 G_s' 의 처음부터 $(2n+1)$ 개 값을 취한 벡터로 구성하고 이후에는 하나씩 점차 이동하며 $(2n+1)$ 개의 값을 취하여 구성하였다.

$$f(g_s') = (g_s' - \bar{g})^T S_g^{-1} (g_s' - \bar{g}) \quad \text{식(13)}$$

4. 실험 및 고찰

실험을 위하여 Tulip 1 데이터베이스 [Movellan, 1995]에 있는 입술 영상들을 사용하였다. 이들은 다양한 조명상태에서 12명이 4개의 숫자를 발음할 때 입력된 연속적인 명암 영상(grey image)으로서 입술 영역만을 포함하고 있으며 다양한 입술 형태를 가지고 100×75 크기를 갖는다. 이들 영상에서 입술의 위치는 대부분 중앙에 있으나 고정된 위치는 아니다. 이 논문에서는 이들 영상에서 학습 영상으로 27개의 입술 영

상을 사용하였으며 실험 영상으로 150개의 영상을 사용하였다. 펜티엄(Pentium) II PC(350 Mhz)에서 실험하였으며 matlab으로 구현하였다.

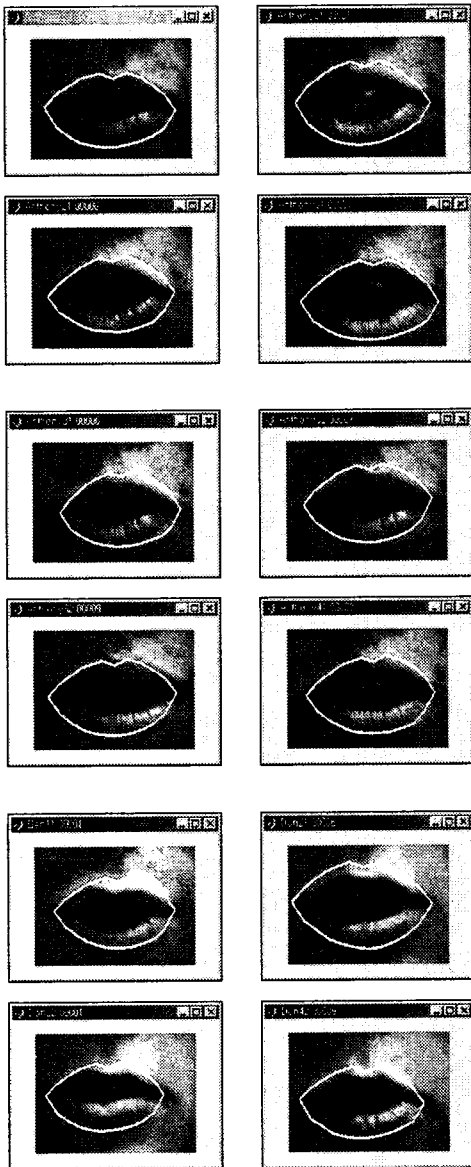
<표 1> 실험 결과

	인식 성공	인식 실패			전체
		윗 입술 외곽선 인식 실패	아래 입술 외곽선 인식 실패	위, 아래 입술 외곽선 모두 인식 실패	
영상 개수(개)	136	4	7	3	150
비율(%)	90.7	2.7	4.6	2	100

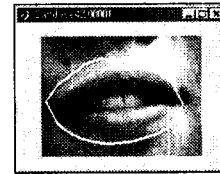
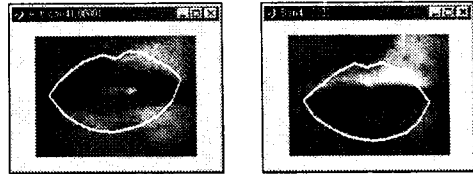
<표 1>은 실험 결과이다. 모든 경우에서 인식 결과의 위치가 입술의 실제 위치와 전혀 다른 결과는 없었으며, 오류인 경우는 실제 입술 외곽선에서 조금씩 떨어진 경우가 발생하였다. 실험결과를 분석하기 위하여 인식한 입술의 외곽선이 실제 위치와 일정 거리 이상 떨어진 경우를 오인식으로 판정하였고 일정 거리 이내에 있는 경우를 인식한 것으로 판정하였다. 이때 거리의 기준으로 윗 입술 높이의 1/3을 사용하였다. 입술의 아래쪽 외곽선 인식과정에서의 오류가 4.6%가 되었으며, 윗쪽 외곽선 인식 오류보다 높았다. 이는 일반적으로 아래 입술과 인접 피부와의 경계가 윗 입술과 인접 피부와의 경계보다 뚜렷하지 않기 때문이다.

<그림 6>은 실험에 사용한 영상들과 인식결과 의 예이다. 입을 다문 경우, 입을 조금 벌린 경우와 많이 벌린 경우, 입술을 내민 경우 등 여러 상황에서 입술을 정확히 인식하였다. <그림 7>은 오인식된 결과들이다. (a)는 입술을 입 안쪽으로 넣은 상태로 입술의 아래 외곽선을 인식하지 못한 경우이고 (b)는 입술의 윗쪽 외곽선을 인식하지 못한 경우이다. (c)는 위, 아래의 입술 외곽선 모두를 인식하지 못한 경우이다. 대부분 입술을 옆으로 길게 늘인 경우에 오인식이 발생하였고

모델 설정시 사용한 학습 영상의 입술과 형태, 조명 등이 많이 다른 경우에 오인식이 발생하였다. 이는 보다 다양한 학습 영상을 사용하여 입술 모델을 구성하면 개선될 것으로 생각된다.



<그림 6> 인식 결과



<그림 7> 오인식 결과

5. 결론

이 논문은 입술의 형태와 입술 외곽선 부근의 화소값 분포를 이용하여 입술을 효과적으로 인식하는 방법을 제안하였다. 입술의 형태에 기반한 가중치 벡터를 이용하여 두 입술 사이의 경계선을 찾고 탐색의 초기 위치로 사용함으로써 능동적 형태 모델에서 초기 위치에 따라 잘못된 결과가 나올 수 있는 문제를 보완하였다. 또한, 에지 및 색 정보를 사용하지 않고 입술 외곽선 주위의 화소분포를 특징으로 정하고 이를 이용하여 영상을 탐색하였으며, 이 과정에서 특징의 정규화 과정을 거쳐 조명변화의 영향을 최소화시켰다. 실험 결과 학습시 사용된 입술의 형태와 환경 등이 다른 경우에 오인식이 발생하였으나 보다 다양한 학습 영상을 사용하면 개선될 수 있을 것이라 생각된다.

참고문헌

- Basu S., N. Oliver and A. Pentlan, "3D Modeling and Tracking of Human Lip Motions", *Proceedings of the 6th International Conf. on Computer Vision*, (1998), 337-343.
- Cootes T. F., A. Hill, C. J. Taylor and J. Haslam, "Use of active shape models for locationg structures in medical images", *Image and Vision Computing*, Vol. 12, No. 6(1994), 355-365.
- 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(1999), 3069-3072.
- Kaucic R., A. Blake, "Accurate, Real-Time, Unadorned Lip Tracking", *Proceedings of the 6th International Conf. on Computer Vision*, (1998), 370-375.
- Lievin M., F. Luthon, "Unsupervised Lip Segmentation under Natural Conditions", *IEEE International Conf. on Acoustics, Speech and Signal Processing*, Vol. 6(1999), 3065-3068.
- 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*, (1999), 691-696.
- Luetin J., N. A. Thacker and S. W. Beet, "Locating and Tracking Facial Speech Features," *Proceedings of The International Conf. on Pattern Recognition*, (1996), 652-656.
- Mirhosseini A. R., H. Yan and K. M. Lam, "Adaptive Deformable Model for Mouse Boundary Detection", *Optical Engineering*, Vol. 37 No. 3(1998), 869-875.
- Movellan J. R., "Visual Speech Recognition with Stochastic Networks", *Advances in Neural Information Processing System*. Vol. 7, MIT Press Cambridge, 1995.
- Oliver N., A. Pentland, "LAFTER: Lips and Face Real Time Tracker", *Proceedings of the 1997 Conf. on Computer Vision and Pattern Recognition*, (1997), 123-129.
- Samal, A., P. A. Iyengar, "Automatic recognition and analysis of human faces and facial expressions: a survey", *Pattern Recognition*, Vol. 25(1992), 65-77.
- Sung K. K., T. Poggio, "Example-Based Learning for View-Based Human Face Detection", *PAMI*, Vol. 20, No. 1(1998), 39-51.
- 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(1998), 123-125.
- Yang J., R. Stiefelagen, U. Meier and A. Waibel, "Real-time Face and Facial Feature Tracking and Application", *Proceedings of Auditory-Visual Speech Processing*, (1998), 79-84.

Abstract

Lip Recognition Using Active Shape Model and Shape-Based Weighted Vector

Kyungshik Jang*

In this paper, we propose an efficient method for recognizing lip. Lip is localized by using the shape of lip and the pixel values around lip contour. The shape of lip is represented by a statistically based active shape model which learns typical lip shape from a training set. Because this model is affected by the initial position, we use a boundary between upper and lower lip as initial position for searching lip. The boundary is localized by using a weighted vector based on lip's shape. The experiments have been performed for many images, and show very encouraging result.

Key words : lip recognition, active shape model, weighted vector

* Department of Multimedia Engineering, Dongeui University, Busan, Korea