

입술 특징점에 기반한 입의 기하학적 왜곡 보정

Geometric Correction of Mouth Based Key Points of Lips

황동국, 박희정*, 전병민

충북대학교, 안동과학대학*

Hwang Dong-Guk, Park Hee-Jung*,

Jun Byoung-Min

Chungbuk National Univ., Andong Science Coll.*

요약

본 논문은 기하학적으로 왜곡된 입 모양을 보정하는 기법을 제안한다. 제안한 기법은 특징추출 단계와 보정 단계로 구성된다. 특징추출 단계에서는 원영상과 목적영상의 입술모델에 따라 각각의 특징과 특징점을 찾고 보정 단계에서는 부분 영상의 사상위치를 결정하고 어파인 변환을 적용하여 입의 왜곡을 보정한다. 여러 형태의 입모양을 실험한 결과, 많은 부분에 존재하는 왜곡이 보정된 것으로 나타났다.

Abstract

In this paper, we propose a method that corrects the geometric distortion of mouth in an image. the method is composed of two steps - detecting key points and correcting geometric distortion. First, key points of lips in source and destination images are found by using lips detection algorithm. Then, the two images are mapped by using affine transformation and information found in first step. In experiment result for various mouths with different geometric distortion, we found that the proposed method have satisfactory efficiency.

I. 서론

입술정보는 음성인식(voice recognition), 독순술(lip-reading) 및 립싱크(lip-sync) 같은 여러 분야에서 사용된다. 그러나 입술영상(lips image)에는 화자(speaker)의 습관 및 자세와 카메라의 위치, 거리 및 상태에 따라 다양하고 복잡한 기하학적 왜곡(geometric distortion)이 존재할 수 있다. 따라서, 입술영상에 존재하는 왜곡을 감소시키거나 제거하는 것은 입모양 연구에 있어 중요한 요소이다.

최근 몇 년 동안에 입술영상 처리 기술에 대한 많은 알고리즘이 제안되었다. Zhwang[1]은 입술영상을 추출하기 위해 색상과 윤곽선 정보를 사용하였고 Nicolas Eveno[2][3]는 영상에서 입술영역을 검출하기 위하여 입술과 피부색의 관계를 분석하였고 입술 윤곽선을 바탕으로 6개의 특징점(key points)을 검출하는 알고리즘을 제안하였다. 또한 S.L. Whang[4]은 파라메트릭 입

술모델과 점기반 입술모델을 결합한 실시간 입술 윤곽선 추출 알고리즘을 제안하였다.

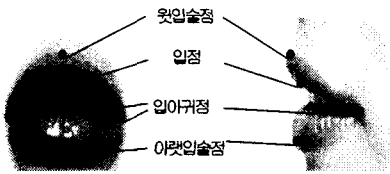
본 논문에서는 기하학적으로 왜곡된 입 모양을 보정하는 기법을 제안한다. 먼저, 원영상의 입술모델에 따라 입모양의 특징점과 특징을 추출한 후, 특징을 이용하여 목적영상의 입술모델을 생성한다. 다음 단계에서, 원영상과 목적영상의 사상영역을 4개로 나누어 어파인 변환(affine transformation)[5]을 적용함으로써 기하학적 왜곡을 보정한다.

제안한 알고리즘 실험은 한국어 '아' 발음에서 나타나는 입모양에 대한 14개 동영상 프레임으로 하였다. 입술영역 추출은 Eveno[3]가 제안한 방법에 의해 입술영상을 얻었다. 실험결과, 대부분의 실험결과영상에서 입술영상이 정규화되고 기하학적 왜곡이 보정된 것을 육안으로 확인할 수 있었다.

II. 입술 위상과 특징

1. 입술 위상

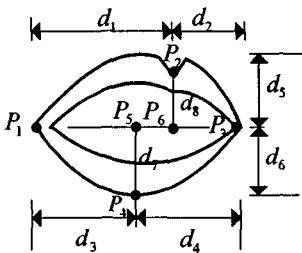
입술 위상은 발음할 때 자리 잡는 입술의 모양이다. 입술은 윗입술과 아랫입술로 구성되며 이 두 입술이 접하는 곳이 구각점이다. 그림 1과 같이 구각점은 입술의 좌우에 위치하며 정면으로 보았을 때 윗입술점-입점-아랫입술점은 일직선상에 위치한다.



▶▶ 그림 1. 입술 위상

2. 입술 특징점 및 특징

일반적으로 입모양은 윗입술점, 입점 및 아랫입술점을 지나는 직선에 좌우 대칭이다. 그러나 영상획득 할 때 발생하는 기하학적 왜곡 또는 화자의 의도적인 입모양 왜곡으로 인하여 비대칭이 될 수 있기 때문에 윗입술점-입점-아랫입술점이 일직선상에 위치하지 않을 수 있다. 따라서 그림 2와 같이 입술모델을 비대칭으로 모델링한다.



▶▶ 그림 2. 원영상의 입술모델

그림 2는 오른쪽으로 왜곡된 입 모델로 6개 특징점 $P_1 \sim P_6$ 과 8개 특징 $d_1 \sim d_8$ 을 나타낸다. P_3 과 P_1 는 오른쪽 구각점과 왼쪽 구각점, P_2 는 윗입술점, P_4 는 아랫입술점, P_5 와 P_6 는 아랫입술과 윗입술의 중심점이다. P_1 ,

P_5 , P_6 및 P_7 은 일직선에 존재한다.

특징점 $P_1 \sim P_4$ 는 영상에서 검출되거나 사전에 주어지는 정보이고 P_5 과 P_6 은 식 (1)로 얻어진다.

$$\sigma = \frac{a_{ry} - a_{ly}}{a_{rx} - a_{lx}}, \quad \sigma\sigma' = 1$$

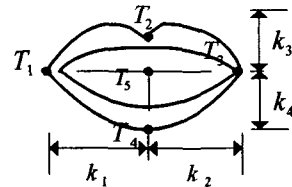
$$\begin{cases} c_x = \frac{(\sigma' l_x - \sigma a_{lx} + a_{ly} - l_y)}{(\sigma' - \sigma)} \\ c_y = \sigma(c_x - a_{lx}) + a_{ly} \end{cases} \quad (1)$$

여기서, (c_x, c_y) 는 중심점(P_5 또는 P_6)의 좌표이고 (l_x, l_y) 는 P_2 나 P_4 이다. 또한, (a_{lx}, a_{ly}) 와 (a_{rx}, c_{ry}) 은 왼쪽 구각점 P_1 이나 오른쪽 구각점 P_3 이 된다. 즉, 직교하는 두 직선의 교점을 구하는 방법이다.

특징은 특징점들 간의 거리를 나타내며 d_1 와 d_2 는 좌우 입술의 넓이이고 윗입술 대칭성을 나타내며, d_3 와 d_4 는 아랫입술의 넓이이고 아랫입술 대칭성을 판별하는데 사용될 수 있다. 또한 d_5 와 d_6 은 아랫입술의 높이와 윗입술의 높이를 나타내며 입의 전체 넓이 W 와 높이 H 는 다음과 같다.

$$W = d_1 + d_2 \quad \text{또는} \quad W = d_3 + d_4$$

$$H = d_5 + d_6$$



▶▶ 그림 3. 목적영상의 입술모델

III. 제안한 입모양의 보정 기법

본 논문에서 제안한 기하학적 보정 기법은 두 단계로 이루어진다. 먼저, i) 특징추출 단계(preprocessing step)는 입술모델에 따라 원영상의 특징점과 특징, 목적영상의 특징점과 특징을 구하는 단계이다. 다음단계는 ii) 보정 단계로 이전 단계의 정보를 이용하여 원영상과

목적 영상의 사상(mapping) 위치를 결정한 후 어파인 변환을 적용하여 왜곡을 보정한다.

특징추출 단계에서, 원영상에서 특징점 $P_1 \sim P_4$ 을 얻는다. 이 특징점들과 식(1)을 적용하여 특징점 P_3 와 P_6 을 구한 후 원영상의 특징인 $d_1 \sim d_8$ 을 구한다. 다음으로, 목적영상의 특징점 $T_1 \sim T_5$ 을 얻는다. 목적영상의 특징점 T_1, T_3 및 T_5 는 특정 위치로 고정되지만 T_2 와 T_4 의 수직 위치는 가변적으로 식(2)와 식(3)에 의해 결정된다.

$$I_{dh}/2 - d_7 \frac{k_1 + k_2}{d_3 + d_4} \quad (2)$$

$$I_{dh}/2 + d_8 \frac{k_1 + k_2}{d_3 + d_4} \quad (3)$$

여기서, I_{dh} 는 목적영상의 높이이다.

보정 단계에서, 사상위치를 결정하기 위하여 원영상의 특징점과 목적영상의 특징점을 기반으로 영상을 4개 부분 영역으로 분할한다. 각각의 부분 영상에 개별적인 어파인 변환을 적용하며 부분 영상 보간 방법으로는 양선형 보간(bilinear interpolation) 기법을 사용한다. 최종적으로 4개의 보정된 영상을 통합하여 최종 목표 영상을 얻는다.

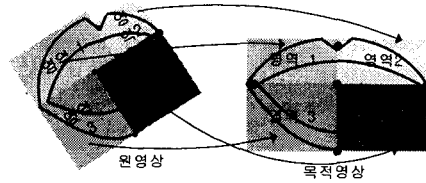
$$R1 : (P_1, P_2, P_6) \xrightarrow{f_1} (T_1, T_2, T_5)$$

$$R2 : (P_3, P_2, P_6) \xrightarrow{f_2} (T_3, T_2, T_5)$$

$$R3 : (P_1, P_4, P_5) \xrightarrow{f_3} (T_1, T_4, T_5)$$

$$R4 : (P_3, P_4, P_5) \xrightarrow{f_4} (T_3, T_4, T_5)$$

위 식에서 $R1, R2, R3, R4$ 는 사상 영역이고 f_1, f_2, f_3, f_4 는 어파인 변환이다. 어파인 계수는 대응하는 특징점들에 의해 결정된다.



▶▶ 그림 4. 원영상과 목적영상의 사상

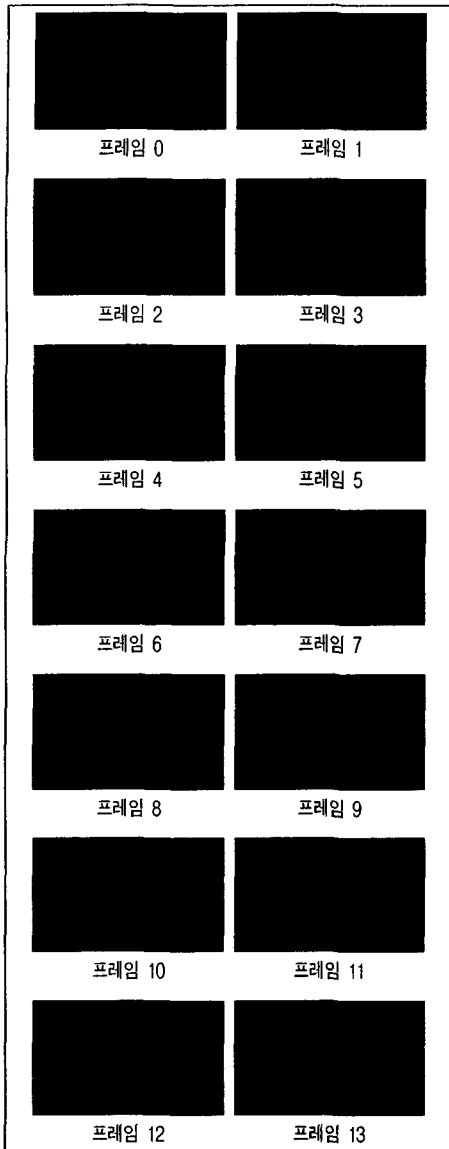
VI. 실험 및 실험결과

한국어 '아' 발음에 대한 14개의 프레임에 대해 제안한 알고리즘을 실험하였다. 실험영상(원영상)은 화상카메라 Logitech QuickCam Pro 4000에서 얻은 동영상 프레임을 입술 세그먼트 알고리즘을 사용하여 추출한 계조영상이다.

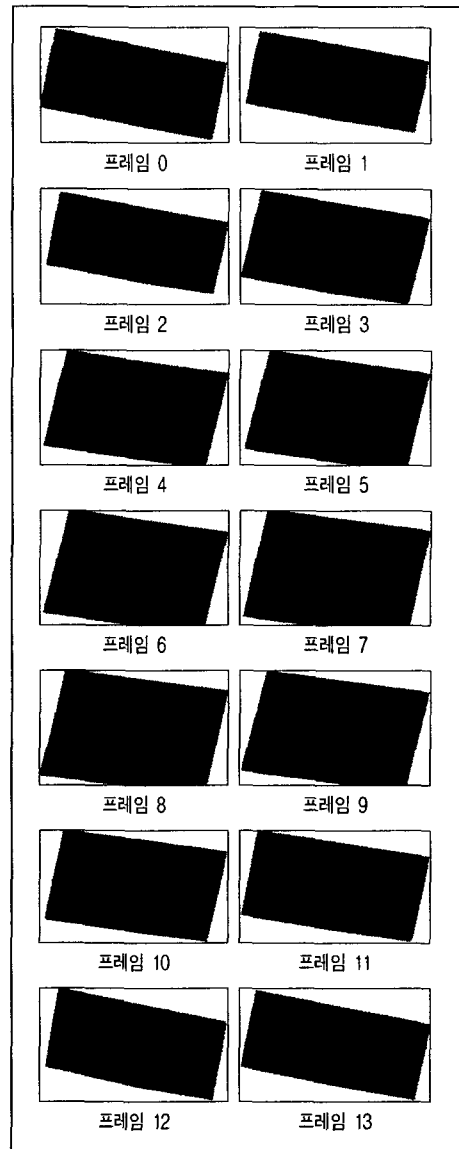
그림 9의 실험결과영상(목적영상)에서는 기하학적으로 보정된 결과영상을 보여준다. 한쪽으로 기울어진 입술영상을 보정한 후 입의 수평크기가 일정하도록 정규화한다. 실험결과영상에서 많은 부분의 기하학적 왜곡이 보정되었으나 특징점 추출 성능에 따라 많은 영향을 받는 것으로 나타났다. 또한, 실험영상의 특징점 P_3 와 P_6 의 불일치는 실험결과영상에서 불연속선으로 나타난다.

V. 결론

본 논문은 입술영상에 나타나는 기하학적 왜곡을 보정하는 기법을 제안하였다. 제안한 기법의 특징추출 단계와 보정 단계를 처리 후 왜곡이 보정된 것을 실험결과에서 볼 수 있었다. 그러나 실험결과영상에서 나타날 수 있는 불연속선을 해결하기 위해서는 정확한 특징점 추출이 요구된다. 따라서, 향후에 강건하게 특징점을 검출하는 알고리즘과 정확한 입술 윤곽선 추출 알고리즘에 대한 연구가 선행되어야 한다.



▶▶ 그림 5. '아' 발음의 입모양에 대한 실험영상
(14 프레임)



▶▶ 그림 6. 실험결과영상(14 프레임)

■ 참고문헌 ■

- [1]. X. Zhang and R.M. Mersereau, "Lip Feature Extraction Towards and Automatic Speechreading System," International Conference on Image Processing, 2002.
- [2]. N. Eveno, A. Caplier and P.-Y. Coulon, "Key Points Based Segmentation of Lips," ICME '02. Proceedings. Vol 2 , pp.125 -128, 2002.
- [3]. N. Eveno, A. Caplier and P.-Y. Coulon, "New Color Transformation for Lips Segmentation," 2001 IEEE Fourth Workshop on Multimedia Signal Processing, pp.3-8, 2001.
- [4]. S.L. Wang, W.H. Lau and S.H. Leung, "A New Real-time Lip Contour Extraction Algorithm", Speech, and Signal Processing, Vol 3, pp. III_217 -III_220, 2003.
- [5]. Y.Zhu, L.D. Seneviratne and S.W.E. Earles, "General Plane Curve Matching under Affine Transformations," Proceedings of the 1996 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vol. 3, pp.1425 -1432, 1996.