

# Lip Contour Detection by Multi-Threshold

Kim Jeong Yeop<sup>†</sup>

## ABSTRACT

In this paper, the method to extract lip contour by multiple threshold is proposed. Spyridonos et. al. proposed a method to extract lip contour. First step is get Q image from transform of RGB into YIQ. Second step is to find lip corner points by change point detection and split Q image into upper and lower part by corner points. The candidate lip contour can be obtained by apply threshold to Q image. From the candidate contour, feature variance is calculated and the contour with maximum variance is adopted as final contour. The feature variance 'D' is based on the absolute difference near the contour points. The conventional method has 3 problems. The first one is related to lip corner point. Calculation of variance depends on much skin pixels and therefore the accuracy decreases and have effect on the split for Q image. Second, there is no analysis for color systems except YIQ. YIQ is a good however, other color systems such as HVS, CIELUV, YCrCb would be considered. Final problem is related to selection of optimal contour. In selection process, they used maximum of average feature variance for the pixels near the contour points. The maximum of variance causes reduction of extracted contour compared to ground contours. To solve the first problem, the proposed method excludes some of skin pixels and got 30% performance increase. For the second problem, HSV, CIELUV, YCrCb coordinate systems are tested and found there is no relation between the conventional method and dependency to color systems. For the final problem, maximum of total sum for the feature variance is adopted rather than the maximum of average feature variance and got 46% performance increase. By combine all the solutions, the proposed method gives 2 times in accuracy and stability than conventional method.

Keywords : Lip Contour, Lip Detection, Lip Segmentation

# 다중 문턱치를 이용한 입술 윤곽 검출 방법

김 정 엽<sup>†</sup>

## 요 약

본 논문에서는 입술 윤곽선을 검출하기 위한 다중 문턱치 기반의 검출방법을 제안하였다. 기존의 연구 중 Spyridonos 등이 제안한 방법은 입력영상을 RGB로부터 YIQ 좌표계로 변환하여 Q 성분만을 이용하여 Q 영상을 얻는다. Q 영상으로부터 변화 점 검출을 통하여 입술 모양의 좌우 끝점을 얻어낸다. 좌우 끝점에 대한 수직 좌표의 평균값을 이용하여 Q 영상을 상하로 분리하고, 상하 영역 각각에 대하여 별도로 Q값을 대상으로 문턱치를 적용하여 후보 윤곽선을 추출한다. 추출된 후보 윤곽선에 특징치 거리를 이용하여 최적의 문턱치를 찾고, 해당하는 윤곽선을 최종 입술 윤곽선으로 결정한다. 이 때 사용되는 특징치 거리 D는 후보 윤곽선 상의 점들을 기준으로 주변 영역에 대한 차이의 절대값을 이용하여 계산한다. 기존연구의 문제점은 세 가지인데, 첫째는 입술 끝점 추출 과정에서 피부영역의 과도한 참여로 입술 끝점의 추출의 정확도가 감소하고, 따라서 후속되는 상/하 영역 분리에도 영향을 미친다. 둘째는 YIQ 칼라 좌표계를 사용하였는데, 다양한 칼라 좌표계에 대한 분석이 미비하므로 추가적인 분석이 필요하다. 셋째, 최적 윤곽선의 선택 시 적용하는 거리 값 파라미터의 계산 과정에서, 문턱치를 적용하여 구한 해당 윤곽선 주변의 데이터들에 의한 변화분을 계산하여 변화가 가장 큰 윤곽선을 입술 후보로 채택하는데, 변화분의 최대치를 기준으로 하기 때문에 검출된 입술영역이 기존보다 축소되는 문제점이 있다. 첫 번째 문제점을 해결하기 위하여 피부영역의 계산과정 참여를 줄여서 성능을 30%정도 향상시켰다. 두 번째는 YIQ 외에 HSV, CIELUV, YCrCb 등의 칼라 좌표계에 대한 성능테스트를 거쳐 기존연구 방법이 칼라좌표계에 대한 의존성이 없음을 확인하였다. 세 번째는 윤곽선 주변의 변화분 검토 시, 윤곽선 포인트 당 변화분의 평균값 대신에 변화분의 총량을 적용하여 46% 성능개선 효과를 얻었다. 이상의 내용을 모두 적용하여 제안한 통합방법은 기존연구 대비 2배의 성능향상과 안정성을 확보할 수 있었다.

키워드 : 입술 윤곽, 입술 검출, 입술 영역분할

## 1. 서 론

립 리딩(Lip-reading)을 위한 구성에서 영상 또는 비디오에서 입술영역 또는 윤곽선을 추출하는 것은 중요하다. 대상이 되는 입력 영상이나 비디오에서 활용할 수 있는 유효한 특

※ 이 논문은 2020학년도 영산대학교 교내연구비의 지원에 의하여 이루어진 것임.

† 정 회 원 : 영산대학교 성심교양대학 부교수

Manuscript Received : June 15, 2020

First Revision : August 25, 2020

Accepted : September 9, 2020

\* Corresponding Author : Kim Jeong Yeop(neocopy@ysu.ac.kr)

징을 얻을 수 있는 초기단계의 처리방법이기 때문이다. 얼굴이 포함된 영상으로부터 입술 모양 또는 윤곽선 정보를 추출하는 것은 중요하다. 입술의 모양은 장면 내에서 말하는 사람의 발음을 유추할 수 있는 중요한 정보이다[1]. 입술 영역으로부터 추출되는 정보를 이용하면 장면의 대화에서 일어나는 언어 소통과 관련된 해석의 정확도를 높이기 위하여 음성 신호에 더하여 부가적으로 이용된다. 특히, 소음이 강하게 발생하여 대화의 음성신호를 해석하기 어려운 경우, 화자에 의해 발음된 음소의 해석 등에서 효과적이다.

### 1.1 관련 연구

흑백 영상에서 입술영역을 추출하는 대표적인 방법은 템플릿(template)과 능동 모델(active model)을 이용하는 방식이다. 템플릿을 사용하는 방법[2]은 입을 중심으로 하는 영역에 대하여 일정한 형식으로 미리 템플릿을 정의하고, 찾으려고 하는 영역을 검출하기 위하여 영상으로부터 변환된 특징치 공간에서 저장된 템플릿에 가장 근접하는 입술 형태를 검색한다. 템플릿의 형태는 입술모양을 다수의 수학적 함수를 이용하여 표현되는 모양이다. 템플릿 방법은 영상 내부에서 찾으려고 하는 입술 모양이, 미리 저장된 데이터베이스 내부의 가장 유사한 템플릿과 근접하도록 결합된 수학 함수 파라미터를 최적화하는 것이 필요하다. 최적화에는 오차에 대한 에너지 함수를 최소화하는 방식을 사용한다.

능동 모델 방법[3]은 미리 정의된 모델의 실질적인 입술 영역과 투영된 대상 영역의 차이를 줄이기 위하여 파라미터를 조정하는 확률 모델이다. 검사하는 영상에서 입술영역에 대한 특정한 색상을 물리적으로 표시하는 방식도 이용되었다[4]. 이 방법은 입술 영상을 미리 기록하여 만들어진 데이터베이스와 함께 주로 사용되었다[5].

흑백과 달리 칼라 영상의 경우는 더욱 많은 정보를 포함하고 있다. 밝기 정보뿐만 아니라 색상정보를 추가로 활용할 수 있기 때문이다. 대부분의 방법들은 칼라 특징을 사용하기 위해 여러 가지 형태의 색상 값을 단일 또는 결합한 형태로 사용하였다. Wimmer 등[6]은 정규화 된 RGB 좌표계로  $R/(R+G+B)$ 와 같은 형태를 이용하였고, Liew[7]과 Wu[8] 등은 CIELab와 CIELuv 칼라 좌표계의 색상 성분을 결합하여 CIELabuv의 밝기 값 1개, 4가지 색상정보를 5차원 벡터로 구성하였다. Sridharan 등[9]은 다른 방식으로 정규화 된 RGB 좌표계를 사용하였는데, 파라미터는 R/G 또는 (r,g) 형태로 구성된 벡터와 그 값들에 대한 평균, 분산 값을 이용하여 특징을 구성하였다. Erol 등[10]은 원형의 RGB, 정규화 된 RGB, 색상 정보(hue)를 결합하여 특징 벡터를 구성하였다. Rohani 등[11]은  $R/(R+G)$ 와 같이 Wimmer와는 다른 형태의 정규화 된 RGB 좌표계를 특징으로 사용하였으며, 이 특징은 입술과 주변의 피부색 영역의 구분에 효과적이었다. Kim[12]은 Liew의 방법에 기반하고 효과적인 칼라 특징을 입술 영역 추출에 이용하였다. Spyridonos 등[13]은 다중 문턱치의 접근방식을 이용한 입술 윤곽선 검출 방법을 제안하

였다. 최근, 단일 알고리즘이 아닌 인공지능 이론을 이용한 딥 러닝(deep-learning) 기법들도 제안되고 있다. Guan 등[14]은 퍼지 딥 신경회로망(fuzzy deep neural network)을 이용하여 입술영역을 검출하는 방법을 제안하였다. NadeemHashmi 등[15]과 같은 딥 신경회로망 기반의 방법들은 입술영역이나 윤곽선 검출 보다는 학습대상을 립 리딩 전체 과정으로 설정한 경우가 대부분이다.

### 1.2 제안한 방법의 연구범위

앞에서 언급한 여러 방법들은 기본적인 단일 칼라 좌표계의 특징 값 또는 두 개 이상의 칼라 좌표계 성분 요소들을 결합하여 벡터 형태로 활용하였다. 다수의 칼라 성분을 결합하는 이유는 입술과 피부영역 색상의 유사도가 높아서 단일 칼라 특징 값만을 이용하는 경우, 효과적인 입술영역 검출이 어렵기 때문이다. 칼라 특징치의 선택에 따라서 입술과 피부 영역의 구분에 대한 성능도 다양하게 나타난다. Spyridonos 등은 YIQ 칼라 좌표계에서 Q성분만을 이용하였고, 최적의 문턱치를 결정하기 위하여 Q 값에 대하여 다수의 문턱치를 설정하고, 추출된 후보 윤곽선들의 포인트들을 기준으로 상하의 일정 거리에 있는 화소의 Q값 차이의 합이 최대인 윤곽선을 선택하는 입술영역 검출에 유리한 제약조건을 추가하였다. 입술영역 검출의 효율성을 위하여 후보 입술영역을 상하로 나누어 각각 추출하고 최종적으로 결합하였다. 본 연구에서는 상하 영역을 나누지 않는 접근방법과 입술영역 검출을 위한 새로운 제약조건을 제안하였다. 제안한 방법을 적용한 결과, 기존 방법보다 하우스도르프 거리[16] 기준으로 전체 평균 100% 정도의 개선 효과가 있었다. 2장에서는 기존의 입술 윤곽선 검출 방법과 문제점을 기술하고, 입술 끝점 추출 방법의 개선과 특징치 거리 제약조건을 포함하는 새로운 입술 윤곽선 검출방법을 제안한다. 3장에서는 실험결과와 고찰을, 4장에서는 결론을 맺는다.

## 2. 제안한 입술 윤곽선 검출 방법

### 2.1 기존의 연구방법 및 문제점

Spyridonos 등은 다수의 문턱치를 기반으로 얻어진 각 후보 윤곽선에 Q값의 편차를 이용한 제약조건을 적용하여 입술 윤곽선을 추출하였다. 전체적인 처리방법은 Fig. 1에 나타내었다.

기존의 방법은 입력영상을 RGB로부터 YIQ 좌표계로 변환하여 Q 성분만을 이용하여 Q 영상을 얻는다. Q 영상으로부터 열(column) 단위 표준 편차 데이터에 Killick 등[17]의 방법을 적용하여 입술 모양의 좌우 끝점을 얻어낸다.

좌우 끝점에 대한 수직 좌표의 평균값을 이용하여 Q 영상을 상하로 분리하고, 각각 상반부-영상, 하반부-영상이라 한다. 해당 알고리즘의 처리 과정을 단계별로 나타내면 다음과 같다.

- ① 입력영상을 RGB 좌표계에서 YIQ 좌표계로 변환함
- ② 입력영상으로부터 입술의 양 끝점을 추출하고 전체영상을 상하로 분할함
- ③ 상반부-영상에 대하여 일정한 Q 값으로 문턱치를 적용한 이진영상을 생성함
- ④ 이진영상에 잡음제거를 위하여 형태 필터링(morphological filtering)을 적용함
- ⑤ 잡음이 제거된 이진영상에서 시계방향으로 윗입술에 해당하는 윤곽선을 추출함
- ⑥ 추출된 윤곽선 정보를 기준으로 하여 특징치 거리 D를 계산함(최초로 구해진 값:  $D_{in}$ )

$$D = \sum_{i=1}^l \sum_{j=1}^l \frac{|q_i^{up} - q_j^{down}|}{l^2} \quad (1)$$

여기서,  $q$ 는 Q영상의 특정 화소의 값을,  $l$ 은 윤곽선 내부의 기준점 개수,  $q_i^{up}$ 은  $i$ 번째 윤곽선 기준점으로부터 위쪽으로 일정한 거리에 있는 한 개 화소의  $q$ 값을 의미한다. 실험에서는 7화소 거리에 있는 동일한 열(Column)의 화소 값을 사용하였다. 마찬가지로,  $q_j^{down}$ 은  $j$ 번째 윤곽선 기준점으로부터 아래쪽으로 일정한 거리에 있는 한 개 화소의  $q$ 값을 나타낸다.

- ⑦ 문턱치 값을 0.01 단위로 계속 증가시킨 다음 ③~⑥을 반복 처리하여 새롭게 구해진 D값과  $D_{in}$ 을 비교하여 D의 최댓값을 갱신함
- ⑧ 최대의 D를 가지는 윤곽선을 상반부-영상의 결과로 저장함

이후, 하반부-영상에 대해서도 동일하게 ③~⑧을 반복 처리하여 아랫입술에 대한 윤곽선을 추출한 다음 결합하여 최종 입술 윤곽선을 검출한다. 여기서 사용되는 문턱치의 최솟값과 최댓값은 Q 영상의 전체 화소에 대한 최소와 최대의 값을 이용하고, 매 반복 시 적용하는 증가분은 0.01로 하였다.

기존연구의 문제점은 다음과 같다.

1) 입술 끝점 추출 방법

Fig. 1의 3번째 블록에 해당하며, 입술을 포함하는 영상에서 세로방향의 분산(Variance) 값  $\sigma^2$ 를 각각 기록한 후, 변화점 검출방법[17]을 적용하여 좌우 끝점을 추출한다. 세로방향의 분산 값을 계산할 때 입술 주위의 피부영역 값이 많이 포함될수록 변화정도의 추정 정확성이 저하되는 경향이 있다.

2) 효율적인 칼라좌표계의 파라미터

기존연구에서는 YIQ 좌표계의 요소 중 Q를 이용하였다. 일반적인 연구들로부터 HSV의 색상(hue) 정보 등이 입술영역추출에 유리한 것으로 알려져 있다. 성능향상을 위하여 다른 칼라좌표계의 특징 값들에 대한 분석과정이 필요하다. 특

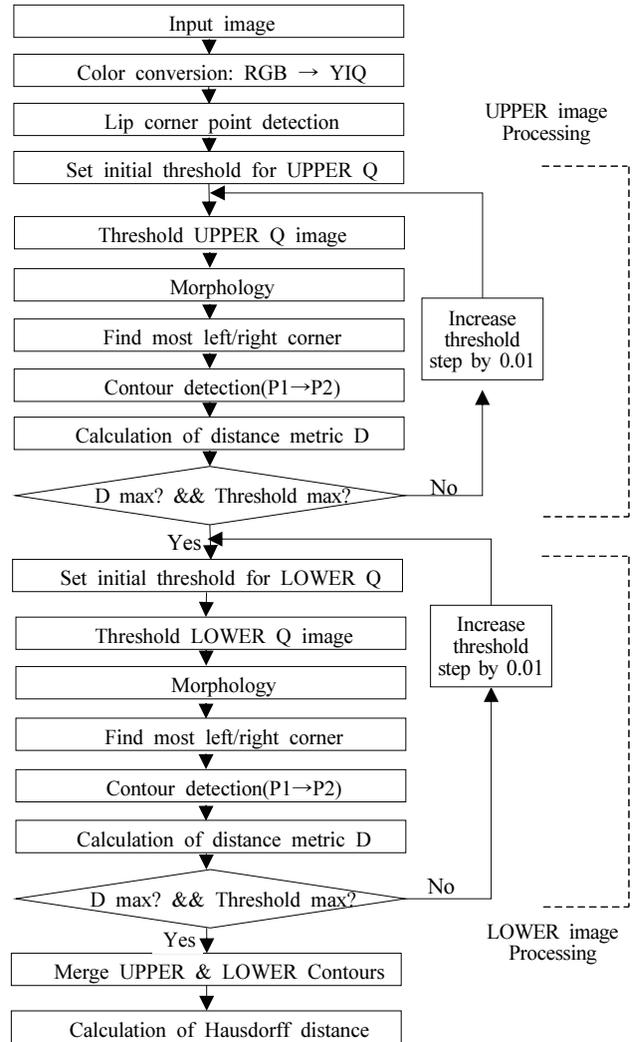


Fig. 1. The Algorithm Flowchart of Conventional Method

히, 입을 포함하는 영상을 상/하로 나누었을 때, 통상적으로 윗부분의 검출은 어려운데, 아랫입술에 비하여 면적이 작고, 피부영역과의 유사도가 높기 때문이다. 그러므로 효율적인 칼라 특징에 대한 검토가 필요하다.

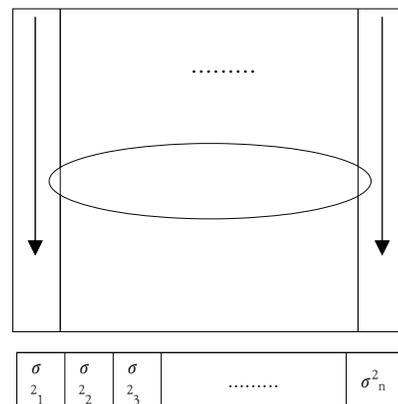


Fig. 2. The Estimation of Variances in Mouth Image

3) 최적 윤곽선의 선택 시 적용하는 거리 값

기존연구에서 제시한 문턱치 결정은 Q 영상을 기준으로 최솟값부터 최댓값까지 매 슬라이스(Slice) 당 최적 윤곽선 여부를 테스트한다. 이 때 사용되는 기준은 해당 윤곽선 주변의 데이터들에 의한 변화 분을 계산한다. 변화가 가장 크게 나타나는 윤곽선을 입술 후보로 채택하는데, 변화분에 해당하는 거리 값의 계산에서 윤곽선 포인트 당 평균 거리를 적용한다. 기존연구 결과를 분석했을 때, 입술 후보 영역은 주로 Q값의 변화가 가장 강하게 검출되는 슬라이스를 채택하였다. 실제로 입술 기준 영역을 수동으로 추출해보면, 변화가 가장 강하게 검출되는 부분 보다는 입술과 피부영역이 맞닿는 부분을 기준으로 하게 된다. 이 지점은 최대 변화 지점과는 차이가 있으므로 기존연구의 거리 값 계산에서 최대 변화 지점을 검출하는 것은 성능저하의 요인이 된다.

2.2 효율적인 칼라좌표계 특징치 분석

기존연구의 문제점 중 2)에 해당하며, 효율적으로 알려진 몇 가지 칼라 특징 치에 대하여 데이터 대비(Contrast)값을 분석하였다. 통상적으로 대비 값이 작은 경우보다는 큰 경우 성능향상을 기대할 수 있기 때문이다. 대비 값은 Equation (2)와 같고 영상내의 매 화소 당 5×5 윈도우 내의 지역 대비(Local contrast) 값  $L_c$ 를 전체 평균하여 비교하였다.

$$L_c = \frac{\max - \min}{\max + \min} \quad (2)$$

분석한 칼라 좌표 특징은 YIQ, HSV, CIELUV, YCrCb의 4종이며 결과는 다음과 같다. 분석에 사용한 실험영상은 Helen Dataset[18]의 1919장 중 1~100번까지를 대상으로 하였다.

Table 1. Result of Average Contrast for YIQ-HSV

| Images               | $Q_U$ | $Q_L$ | $H_U$ | $H_L$ | $Q_{U_{sup}}$ | $Q_{L_{sup}}$ |
|----------------------|-------|-------|-------|-------|---------------|---------------|
| 100 sample (#1~100)  | 0.142 | 0.140 | 0.235 | 0.239 | 30            | 18            |
| ALL sample (#1~1919) | 0.121 | 0.121 | 0.225 | 0.219 | 424           | 319           |

Table 1에 YIQ의 Q와 HSV의 H에 대하여 기존연구의 기준으로 상/하 영역으로 나눈 다음 대비 값을 나타내었다.  $Q_{U_{sup}}$ 은 상부영역에서,  $Q_{L_{sup}}$ 은 하부영역에서 Q가 H보다 대비 값이 큰 샘플의 수를 나타낸다. 결과적으로 Q보다는 H의 대비가 우세함을 확인할 수 있다. 100장의 샘플과 전체 데이터 셋의 결과에 경향성이 유지되므로 다른 칼라 특징 분석에는 100장의 샘플을 사용한다.

기존연구 방법에 칼라특징치만 변경한 후, 실험을 하여 최종 하우스도르프 거리를 계산한 결과는 다음과 같다.

하우스도르프 거리가 작을수록 입술영역 추출이 잘 된 것으로 판단할 수 있다. 실험결과를 검토해보면, HSV와 CIELUV

Table 2. Result of Average Contrast for YIQ-CIELUV(|u-v|)

| Color Features      | $Q_U$ | $Q_L$ | $L_U$ | $L_L$ | $Q_{U_{sup}}$ | $Q_{L_{sup}}$ |
|---------------------|-------|-------|-------|-------|---------------|---------------|
| 100 sample (#1~100) | 0.142 | 0.140 | 0.281 | 0.270 | 7             | 11            |

Table 3. Result of Average Contrast for YIQ-YCrCb(Cr/Cb-Cr<sup>2</sup>)

| Color Features      | $Q_U$ | $Q_L$ | $Y_U$ | $Y_L$ | $Q_{U_{sup}}$ | $Q_{L_{sup}}$ |
|---------------------|-------|-------|-------|-------|---------------|---------------|
| 100 sample (#1~100) | 0.142 | 0.140 | 0.109 | 0.103 | 79            | 90            |

Table 4. Result of Average Hausdorff Distance for Color Coordinate System Changes

| Color Features     | YIQ    | HSV    | CIELUV | YCrCb  |
|--------------------|--------|--------|--------|--------|
| Hausdorff distance | 30.657 | 33.861 | 23.986 | 43.117 |

칼라 특징의 대비 값은 0.2 근처로 유사하지만 입술 윤곽 검출 성능은 차이를 보이고 있다. YIQ와 HSV는 검출성능이 비슷하지만 대비 값은 큰 차이를 보인다.

문턱치를 기반으로 하는 Spyridonos 등의 기존연구 방법에서는 칼라 좌표계에 다른 특징치의 대비 값이 큰 영향이 없는 것으로 판단되므로, HSV나 CIELUV, YCrCb 등의 도입은 필요하지 않다. 즉, 칼라 특징치의 대비 값과 검출성능의 연관성은 발견되지 않았다.

2.3 제안하는 입술 윤곽선 검출방법

본 논문에서는 2.1절에서 제시한 기존의 Spyridonos 등의 입술영역 윤곽선 검출방법을 분석하여, 세 가지 개선 방향을 제시하였다. 칼라 좌표계의 선택은 2.2절의 분석을 통하여 YIQ를 그대로 유지하도록 하므로, 문제점 1)과 3)에 대한 개선을 제안하고자 한다.

1) 입술 끝점 추출 방법의 개선

Fig. 2에서 나타나는 입술 주변의 피부영역 참여를 최대한 배제하는 것이 성능향상에 유리하다고 판단된다. 그러므로 기존연구의 방법을 응용하여 Fig. 3 (a)의 점선으로 표시되는 최대한 피부영역을 배제한 영역을 대상으로 Fig. 3 (b)에서와 같이 새로운 분산 값을 계산한다. 변화 점 검출방법을 세로방향으로 먼저 적용하여 점선으로 표시되는 입술 상단과 하단의 후보 영역을 대상으로 Fig. 2와 같은 세로방향의 분산 값을 이용한 변화 점 검출을 실행한다.

제안한 입술 끝점 추출 방법만을 기존연구에 적용하여 실험한 결과, Helen Dataset 100장의 샘플에 적용한 평균 하우스도르프 거리는 21.231로 기존연구 방법에 의한 거리 30.657에 비하여 30% 정도의 개선효과가 있음을 확인하였

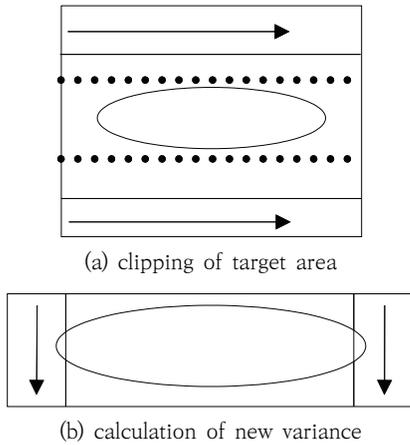


Fig. 3. The Estimation of New Variances in Mouth Image

다. 다음으로 기존연구의 문제점 중 마지막인 최적 윤곽선의 선택 시 적용하는 거리 값에 대하여 검토한다.

2) 최적 윤곽선의 선택 시 적용하는 거리 값

2.1절의 3)에 언급한 대로, Spyridonos 등의 기존연구는 Equation (1)에 의해 윤곽선 주변의 데이터 변화분이 가장 큰 슬라이스를 최적 윤곽선으로 채택하기 때문에 실질적인 윤곽선보다는 어느 정도 축소된 형태로 결과가 나오게 된다. Equation (1)에서는 윤곽선을 구성하는 기준 포인트를 중심으로 하여 상하로 일정한 거리 떨어진 화소의 Q 값의 차이를 누적하여 평균적인 거리를 계산하여 윤곽선의 신뢰도를 결정하였다. Equation (1)은 고려 대상이 영역의 형태를 가지기는 하지만 포인트를 중심으로 조합적인 거리의 평균치를 계산하여 이용한다. Equation (1)에서 윤곽선 포인트 당 평균치를 구하여 사용하기 때문에 윤곽선 포인트 주변의 데이터 변화량이 크기만 하면 윤곽선 후보 중에서 면적이 작은 윤곽선도 채택될 가능성이 높다. 실험을 통하여 관찰한 결과, Q 값이 낮은 부분 보다는 대체로 Q 값이 높은 부분의 슬라이스가 채택되는 경향이 있었다. 즉, 문제점 분석에서 언급한 윤곽선 포인트 당 평균적인 변화량은 Q 값이 높을수록 큰 경향을 가지므로, 선택된 최적의 윤곽선은 크기가 작은 것일 가능성이 높고, 하부 영역보다 윗입술을 포함하는 상부 영역에서 자주 발생하였다. 이 문제점을 해결하기 위하여 윤곽선을 구성하는 포인트 당 평균 데이터 변화량을 이용하는 것 보다는 전체 윤곽선 포인트에 대한 데이터 변화량의 총량을 선택하는 것이 유리하다. Equation (3)과 같이 거리 값  $D_p$ 를 평균이 아닌 총량을 기준으로 계산하면 이와 같은 문제점을 완화하는 효과가 있을 것이다.

$$D_p = \sum_{i=1}^l \sum_{j=1}^l |q_i^{up} - q_j^{down}| \quad (3)$$

제안한 거리 값  $D_p$ 의 효과를 확인해 보기 위해 기존연구 방법에 거리 값 계산 부분만 Equation (1)의 D에서 총량인

Equation (3)의  $D_p$ 로 바꾸고, Helen Dataset의 샘플 100장에 적용한 결과 평균 하우스도르프 거리는 16.530으로 기존연구의 30.657에 비하여 46% 정도의 개선효과가 있음을 확인하였다.

이상에서 제안한 문제점 1)과 3)의 개선방법을 통합하여 최종적인 입술 윤곽선 검출 알고리즘을 구성하였다. 두 가지 제안 점을 동시에 기존연구 방법에 적용한 결과 평균 하우스도르프 거리는 16.705로 확인되었다. 문제점 3)만 적용한 경우보다 0.175 높은 결과가 나와서 분석해본 결과 100장의 영상 중에서 #40번 영상은 수염이 많은 사람의 경우이며, 하우스도르프 거리가 각각 175, 217로 나왔고, 이를 제외한 나머지 99장 영상에 대한 평균 하우스도르프 거리는 14.772, 14.529로 통합 방법이 약간 우수함을 확인할 수 있었다.

이상의 제안한 통합방법의 전체 처리과정을 Fig. 4에 나타내고 설명하였다.

- ① 입력영상을 RGB 좌표계에서 YIQ 좌표계로 변환함
- ② 입력영상으로부터 제안한 방법으로 입술의 양 끝점을 추출하고 전체영상을 상하로 분할함
- ③ 상반부-영상에 대하여 일정한 Q 값으로 문턱치를 적용한 이진영상을 생성함
- ④ 이진영상에 잡음제거를 위하여 형태 필터링(morphological filtering)을 적용함
- ⑤ 잡음이 제거된 이진영상에서 시계방향으로 윗입술에 해당하는 윤곽선을 추출함
- ⑥ 추출된 윤곽선 정보를 기준으로 하여 제안한 특징치 거리  $D_p$ 를 계산함(최초로 구해진 값:  $D_{in}$ )
- ⑦ 문턱치 값을 0.01 단위로 계속 증가시킨 다음 ③~⑥을 반복 처리하여 새롭게 구해진  $D_p$ 값과  $D_{in}$ 을 비교하여  $D_p$ 의 최댓값을 갱신함
- ⑧ 최대의  $D_p$ 를 가지는 윤곽선을 최종 결과로 저장함

여기서 사용되는 문턱치의 최솟값과 최댓값은 Q 영상의 전체 화소에 대한 최소와 최대의 값을 이용하고, 매 반복 시 적용하는 증가분은 0.01로 적용하였다.

3. 실험결과 및 고찰

본 논문에서는 Helen Dataset[18]을 사용하였는데, 이 데이터베이스는 사진의 크기와 장면의 조명이 보정되지 않은 2300여장으로 구성되어 있다. 이 영상들은 인터넷으로부터 수집된 것이므로, 주변조명, 실내/실외, 인물 등의 영상 촬영 조건이 다양하고 전혀 보정되지 않았다. 본 논문에서 사용한 영상은 이 데이터베이스에서 그래픽으로 작성된 경우를 제외한 1919장이 사용되었고, 약 380장은 흑백사진이거나 사진이 아닌 그래픽 데이터여서 제외하였다. 제공되는 기본정보(ground truth information)는 전체 영상을 기준으로 각 영상마다 얼굴영역의 외곽선(contour), 눈, 코, 입의 외곽선 정보를 포함한다. 테스트 영상의 일부를 Fig. 5에 나타내었다.

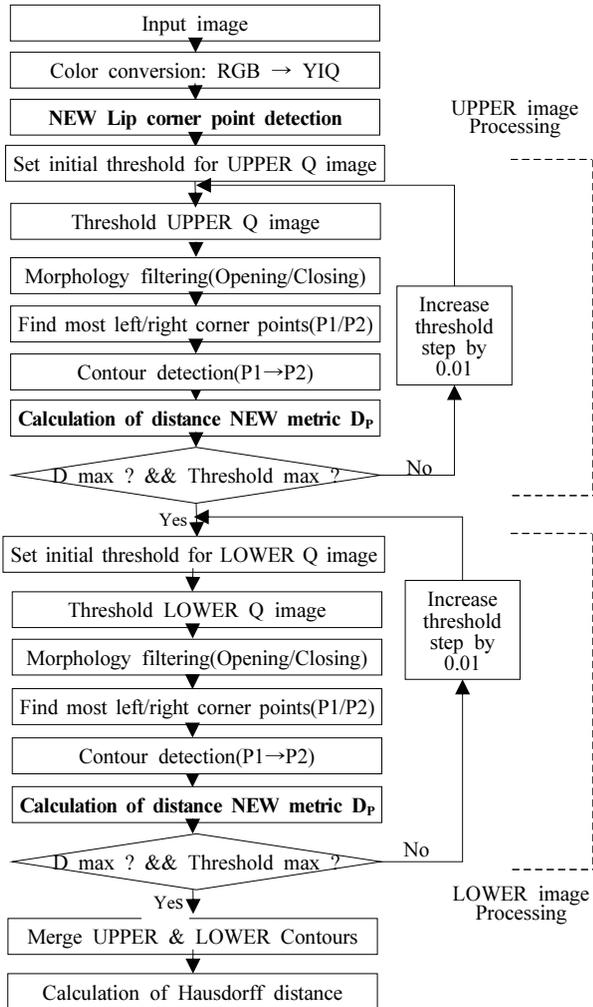


Fig. 4. The Algorithm Flowchart of Proposed Method

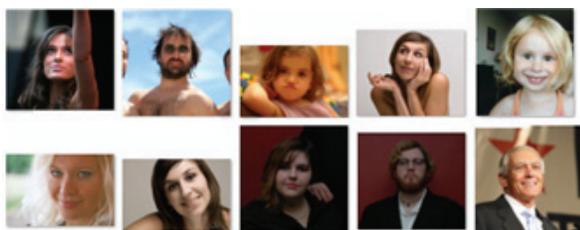
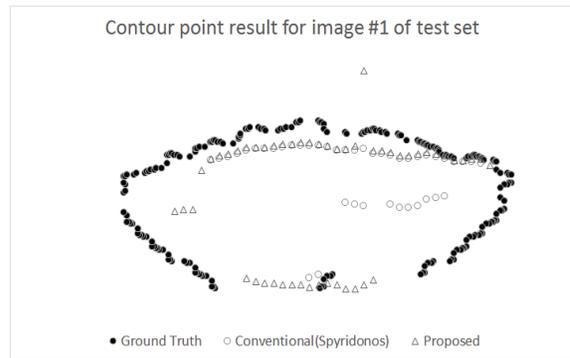


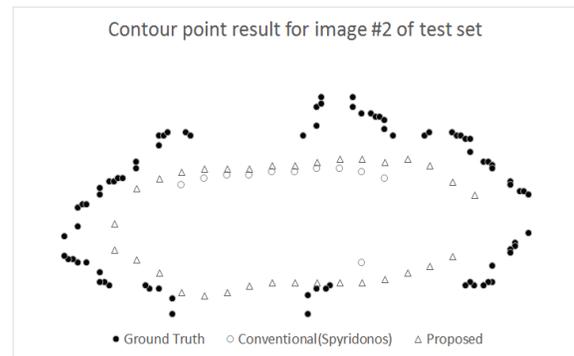
Fig. 5. Example Test Images from Helen Dataset

테스트 영상으로부터 얼굴영역을 추출하기 위해서 Nilsson[19]의 방법을 이용하였다. 추출된 얼굴영역에서 높이 기준 하위 1/3 범위와 너비를 기준으로 하여 중앙 1/2 을 입술 후보영역으로 추출하였다. 추출된 입술후보 영역에 대한 결과와 데이터 베이스에서 제공하는 기준 정보(Ground truth information)를 이용하여 하우스도르프 거리를 계산하였다.

실험에 의한 입술 윤곽선 추출 결과를 Fig. 6에 나타내었다. Fig. 6 (a)는 Fig. 5의 첫 번째 줄 첫 번째 영상에 대한 결과이고, (b)는 두 번째 영상에 대한 추출 결과를 나타내었다. Fig에서 짙게 채워진 원형 기호는 기준 정보를 나타내고,



(a) Contour result for image #1 of Fig. 5



(b) Contour result for image #2 of Fig. 5

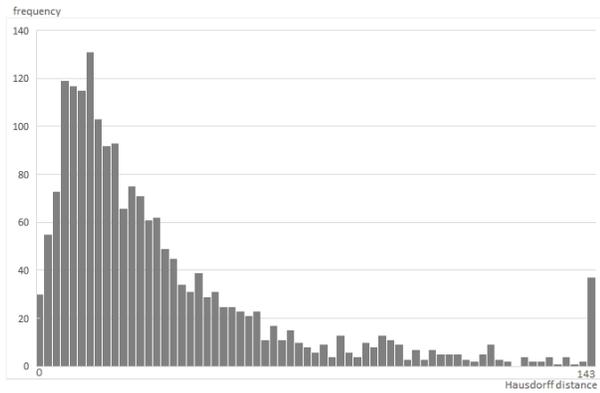
Fig. 6. Contour Points by Conventional and Proposed Methods

속이 비워진 원형 기호는 기존연구의 결과이고, 빈 삼각형 기호는 제안한 방법에 의한 결과이다. 시각적으로 보더라도 제안한 방법이 기존연구에 비하여 우수함을 확인할 수 있다.

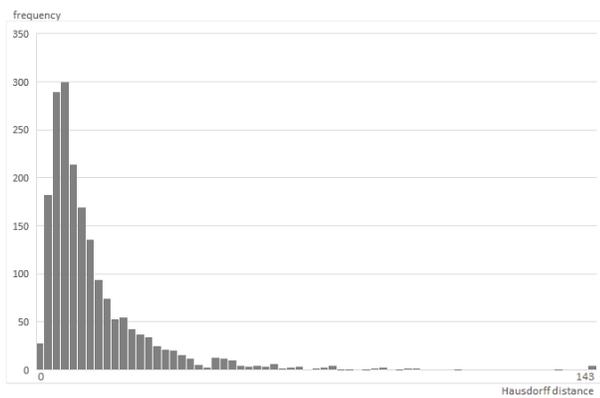
입술 윤곽선의 추출결과에 대한 성능평가는 시각적인 형태로 판단하기 어렵기 때문에 윤곽선 간의 유사도를 수치로 판별해 주는 하우스도르프 거리를 이용한다. 하우스도르프 거리가 "0"이라면 비교하는 윤곽선이 정확하게 일치한다는 것이며 거리 값이 작을수록 유사도가 높다. 성능평가에는 기존의 방법과 제안한 방법의 비교를 중심으로 하고, 성능 평가 결과는 Fig. 7과 Table 5에 나타내었다.

Fig. 7의 가로축은 하우스도르프 거리 값을 나타내고, 세로축은 해당 거리 값에 대한 발생 빈도이다. Fig. 7 (a)는 기존연구에 대한 총 1919장의 데이터 셋에 대한 하우스도르프 거리 값의 분포도이다. Fig. 7 (b)는 새로운 입술 끝점 추출과 수정된 특징치 거리  $D_p$ 를 적용한 제안 방법의 결과를 나타낸다. 그림에서 보는 바와 같이 제안한 방법이 기존연구에 의한 방법보다 성능이 우수함을 알 수 있다.

실질적인 하우스도르프 거리의 계산에서는 순방향거리(FHD: Forward Hausdorff Distance)와 역방향거리(RHD: Reverse Hausdorff Distance)의 두 가지를 각각 계산하고 더 큰 값을 최종 하우스도르프 거리(HDIS: Hausdorff Distance)로 결정한다. 윤곽선의 형태에 따라 순방향거리와 역방향거리의 값이 서로 다르게 나타날 수 있고, 성능평가에



(a) Histogram of Hausdorff distances (conventional)



(b) Histogram of Hausdorff distances (proposed)

Fig. 7. Performance Evaluation for Conventional and Proposed Methods

영향을 주게 되므로 대부분의 경우에는 두 값을 모두 계산한 다음 큰 값을 채택한다. 본 실험에 사용된 영상은 총 1919장이며, 각 방식별로 평균값과 분산 값을 Table 5에 나타내었다. Table 5에서  $F_{AVG}$ 는 순방향거리의 평균값을,  $F_{VAR}$ 는 순방향거리의 분산 값,  $R_{AVG}$ 는 역방향거리의 평균값,  $R_{VAR}$ 는 역방향거리의 분산 값,  $H_{AVG}$ 는 하우스도르프 거리의 평균값,  $H_{VAR}$ 는 하우스도르프 거리의 분산 값을 의미한다.

Table 5. Average and Variance Result of Hausdorff Distances for the Methods

| Methods      | $F_{AVG}$ | $F_{VAR}$ | $R_{AVG}$ | $R_{VAR}$ | $H_{AVG}$     | $H_{VAR}$ |
|--------------|-----------|-----------|-----------|-----------|---------------|-----------|
| Conventional | 38.261    | 33.147    | 23.514    | 25.323    | <b>39.570</b> | 34.368    |
| Proposed     | 14.661    | 8.947     | 18.846    | 17.422    | <b>19.884</b> | 17.198    |

제안한 방법은 기존의 방법에 비하여 평균 하우스도르프 거리가 50% 정도이고, 분산 값도 약 50% 정도로 나타났다. 평균 하우스도르프 거리는 정확성을 의미하며, 분산 값은 추출 성능의 안정성 척도라 할 수 있고, 제안한 방법은 기존의 방법 대비 2배의 정확성과 2배의 안정성을 가지는 것으로 볼 수 있다.

#### 4. 결 론

본 논문에서는 입술 영역의 윤곽선 검출을 위하여 다중 문턱치를 이용한 검출 방법을 제안하였다. 검출 성능을 향상시키기 위하여 피부영역의 참여를 감소시키는 입술 끝점 추출 방법과, 최적의 문턱치를 결정하기 위한 개선된 특징치 거리 계산 방법을 도입하였다. 또한 다양한 칼라 좌표계에 대한 분석을 통하여 특징치의 대비와 추출 성능은 무관함을 확인하였다. 새로운 특징치 거리 계산 방식을 통하여 윤곽선 검출 성능을 향상시킬 수 있었으며, 제안한 방법의 성능을 평가하기 위하여 윤곽선에 대한 하우스도르프 거리를 도구로 이용하였다. 기존 대비 정확성 측면과 검출 안정성 측면에서 각각 2배로 성능이 우수함을 확인하였다. 기존의 방법과 제안한 방법 모두 단일 윤곽선을 대상으로 검출 결과를 도출하므로, 향후 연구에서 다수의 윤곽선을 다룰 수 있는 방식으로 개선되어야 할 것이다.

#### References

- [1] G. Potamianos, C. Neti, J. Luetin, and I. Matthews, "Audio-Visual Automatic Speech Recognition: An Overview," In: *G. Bailly, E. Vatikiotis-Bateson and P. Perrier(Ed.), Issues in Audio-Visual Speech Processing*, MIT Press, 2004.
- [2] A. L. Yuille, P. W. Hallinan, and D. S. Cohen, "Feature extraction from faces using deformable templates," *International Journal of Computer Vision*, Vol.8, No.2, pp.99-111, 1992.
- [3] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models - their training and application," *Computer Vision and Image Understanding*, Vol.61, No.2, pp.38-59, 1995.
- [4] C. Bregler and Y. Konig, "Eigenlips for Robust Speech Recognition," *Proc. of IEEE International Conf. on Acoustics, Speech and Signal Processing*, pp.669-672, 1994.
- [5] G. I. Chiou and J. N. Hwang, "Lip Reading from Color Video," *IEEE Transactions on Image Processing*, Vol.6, No.8, pp.1192-1195, Aug. 1997.
- [6] M. Wimmer and B. Radig, "Adaptive skin color classifier," In *Proc. of the first ICGST International Conference on Graphics, Vision and Image Processing GVIP-05*, Vol.1, pp.324-327, 2005.
- [7] A. W. Liew, S. H. Leung, and W. H. Lau, "Segmentation of Color Lip images by Spatial Fuzzy Clustering," *IEEE Trans. on Fuzzy Systems*, Vol.11, No.4, pp.542-549, 2003.
- [8] Y. Wu, R. Ma, W. Hu, T. Wang, Y. Zhang, J. Cheng, and H. Lu, "Robust Lip Localization on Multi-View Faces In Video," *IEEE International Conf. on Image Processing*, pp.481-484, San Antonio, Texas, USA. 2007.

- [9] S. Sridharan, S. Lucey, and V. Chandran, "Adaptive Mouth Segmentation Using Chromatic Features," *Pattern Recognition Letters*, Vol.23, No.11, pp.1293-1302, 2002.
- [10] O. Erol, Y. Berkay, K. Harun, E. Hakan, and U. Mustafa, "Lip Segmentation Using Adaptive Color Space Training," *Proc. of the International Conference on Auditory-Visual Speech Processing*, Sept. 26-29, Tangalooma Wild Dolphin Resort, Moreton Island, Queensland, Australia, pp. 219-222, 2008.
- [11] R. Rohani, F. Sobhanmanesh, S. Alizadeh, and R. Boostani, "Lip Processing and Modeling based on Spatial Fuzzy Clustering in Color Images," *International Journal of Fuzzy Systems*, Vol.13, No.2, pp.65-73, Jun. 2011.
- [12] J. Kim, "Extraction of Lip Region using Chromaticity Transformation and Fuzzy Clustering," *Journal of Korea Multimedia Society*, Vol.17, No.7, pp.806-817, Jul. 2014.
- [13] P. Spyridonosa, A. F. Saintb, A. Likasb, G. Gaitanisc, and I. Bassukasc, "Multi-Threshold LIP Contour Detection," *2018 25th IEEE International Conference on Image Processing(ICIP)*, Athens, 2018, pp.1912-1916, 2018.
- [14] C. Guan, S. Wang, and A. W. Liew, "Lip image segmentation based on a fuzzy convolutional neural network," *IEEE Transactions on Fuzzy Systems*, 2019.
- [15] S. NadeemHashmi, H. Gupta, D. Mittal, K. Kumar, A. Nanda, and S. Gupta, "A Lip Reading Model Using CNN with Batch Normalization," *2018 Eleventh International Conference on Contemporary Computing (IC3)*, Noida, pp.1-6, 2018.
- [16] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.15, No.9, pp.850-863, 1993.
- [17] R. Killick, P. Fearnhead, and I. A. Eckley, "Optimal Detection of Changepoints With a Linear Computational Cost," *Journal of the American Statistical Association*, Vol.107, No.500, pp.1590-1598, Dec. 2012.
- [18] V. Le, J. Brandt, Z. Lin, L. Bourdev, and T. S. Huang, "Interactive facial feature localization," *European Conference on Computer Vision (ECCV)*, Berlin, pp.679-692, 2012.
- [19] M. Nilsson, J. Nordberg, and I. Claesson, "Face detection using local SMQT features and split up snow classifier," *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference*, Vol.2. IEEE, 2007.



**김 정 업**

<https://orcid.org/0000-0001-5807-1002>

e-mail : neocopy@ysu.ac.kr

1990년 경북대학교 전자공학과(학사)

1992년 경북대학교 전자공학과(석사)

2001년 경북대학교 전자공학과(박사)

1994년 ~ 2001년 삼성종합기술원(과장)

2001년 ~ 2013년 영산대학교 게임콘텐츠학과 부교수

2013년 ~ 현 재 영산대학교 성심교양대학 부교수

관심분야 : Color Image Processing & Computer Vision