

論文99-36T-9-5

## 화상전송 서비스를 위한 객체 분할 및 지식 기반 얼굴 특징 검출

### (Object Segmentation for Image Transmission Services and Facial Characteristic Detection based on Knowledge)

林春煥\*, 梁洪榮\*

(Chun-Hwan Lim and Hong-Young Yang)

#### 요 약

이 논문에서는 영상통신을 위한 객체 분할 알고리즘과 지식기반 얼굴 특징 검출 알고리즘을 제안한다. 이를 위해 일정한 조도 상태에서 비디오 카메라로부터 동일거리에서  $256 \times 256$  크기의 그레이 스케일 256 레벨로 두장의 영상을 취득한 후 가우시안 필터를 이용하여 영상 내에 존재하는 잡음을 제거한다. 취득된 두 입력영상은 비디오가 가리키고 있는 배경영상과 배경영상 내에 사람이 포함된 경우의 영상이다. 그리고 배경영상과 얼굴이 포함된 입력 영상의 차를 구한다. 다음으로 차영상에 대해 축소와 팽창 과정을 통해 얼굴영역 내의 잡음을 제거한 후 좌, 우에서 픽셀 값을 조사하여 마스크를 생성한다. 그리고 마스크를 얼굴이 존재하는 원 영상에 투영하여 배경에서 얼굴영역을 분할한 후 얼굴영역의 에지성분을 조사하여 눈, 코, 귀, 입을 분할한다. 시뮬레이션 결과를 통하여 제안된 알고리즘의 효율성을 입증한다.

#### Abstract

In this paper, we propose a facial characteristic detection algorithm based on knowledge and object segmentation method for image communication. In this algorithm, under the condition of the same lumination and distance from the fixed video camera to human face, we capture input images of  $256 \times 256$  of gray scale 256 level and then remove the noise using the Gaussian filter.

Two images are captured with a video camera, One contains the human face; the other contains only background region without including a face. And then we get a differential image between two images. After removing noise of the differential image by eroding and dilating, divide background image into a facial image. We separate eyes, ears, a nose and a mouth after searching the edge component in the facial image.

From simulation results, we have verified the efficiency of the proposed algorithm.

#### I. 서 론

영상통신 환경에서 정지 및 동영상 데이터를 압축하여 전송하고자 할 때에는 주로 픽셀 혹은 고정된 크기

의 블록단위로 압축하여 전송한다. 이것은 파형기반 압축 기술로 제 1세대 영상압축기술로 분류되며 JPEG, MPEG-1, 2, H.261, H.263 등이 파형기반 압축 기술을 바탕으로 제정되었다. 그러나 8-64kbps 급의 초저속 영상통신용이나 멀티미디어 응용을 위해서는 보다 고효율의 압축 기술이 요구된다. 고효율 압축기술로 최근 각광받고 있는 기술은 내용기반 영상압축기술로 영상을 픽셀 혹은 블록단위로 압축하는 것이 아니라 영상의 내용을 인식하여 의미있는 객체단위로 압축함으로써

\* 正會員, 東岡大學 電子科

(Dept. of Electronics Eng., dongkang College)

接受日字 : 1999年7月16日, 수정완료일 : 1999年8月29日

써 압축 효율을 높이고자 하는 기술이다. 즉, 객체기반 영상 압축 기술은 압축하고자 하는 영상을 의미있는 동질의 영역(객체)으로 분할한 다음 분할한 객체의 컬러, 움직임, 그리고 형태 정보를 압축한다.

이러한 객체기반 영상압축 기법들은 해석단계와 합성단계의 두단계로 이루어져 있다. MPEG-1, 2와 같은 영상압축 기법에서는 압축하고자 하는 영상에 대해 해석을 하지않고 영상의 내용에 관계없이 일정 크기의 블록으로 나누어 압축한다. 그러나 MPEG-4와 같이 객체단위로 압축하는 경우에는 필수적으로 영상을 해석하여 객체분할을 하는 해석단계가 필요하며 고효율의 압축을 달성하기 위해서는 해석단계와 합성단계가 적절히 잘 수행되어야 한다.

그러나 MPEG-4 표준안은 합성단계는 규정하고 있으나 해석단계는 아직 표준화하지 않고 있다. 따라서 해석 단계에서 분할된 영상의 결과에 따라 MPEG-4와 MPEG-7의 성능에 커다란 영향을 미치게 되고 합성단계의 입력으로 제공된 분할된 객체를 실시간 내에 만들어줄 영상 분할 알고리즘이 존재하지 않는 한 MPEG-4에 적용될 동영상은 극히 제한적일 수 밖에 없다.<sup>[1]</sup>

따라서 본 연구에서는 배경이 일정하고 전경만 변하는 영상통신 환경하에서 얼굴을 분할하여 효율적인 통신환경을 제공할 수 있는 객체 분할 알고리즘을 제안한다. 또한 분할된 얼굴 영상에서 지식에 근거한 해석 방법에 의해 세부적인 특징을 검출한다. 이를 위하여 II장에서는 얼굴영역 분할 연구 동향에 대해 살펴보고 III장에서는 객체 분할 알고리즘을 살펴본다. 그리고 제IV장에서는 시뮬레이션을 수행하고 제V장에서 결론을 맺는다.

## II. 연구 동향

컴퓨터로 얼굴을 분할하고 인식하기 위한 연구는 많은 방법을 통해 연구되고 있다. 일반적인 연구 흐름은 얼굴 부위별 상대적인 위치 및 특징적인 구조를 예상하고 이러한 예상이 실제 영상에서의 구조와 일치하는가를 확인하는 과정을 통해 얼굴인식이 이루어진다. 또한 인간의 얼굴은 대부분 좌우 대칭이고 눈썹아래 눈이 있고 그 아래 부분에 코가 존재하고 코 아래에 입술이 위치한다는 등의 일반적인 지식을 이용하여 얼굴인식을 하게 된다.

주어진 영상에서 얼굴의 존재 유무를 가려내는 초기의 방법에서는 주로 에지(Edge) 영상을 사용하였다. Sakai는 입력 영상에서 추출된 에지 지도에 타원형의 형판을 적용하여 대략적인 머리 영역을 설정하고, 설정된 머리 영역 내에서 눈, 입의 에지 영상을 검사함으로써 최종적인 머리 영역을 추출하는 방법은 제시하였다<sup>[2]</sup>. 이 방법은 조명의 방향에 크게 영향을 받는 단점을 갖는다. 또한 Kelly는 입력 영상으로부터 자동적으로 머리와 몸의 윤곽을 추출하고, 계속해서 눈, 코, 입의 위치를 추출하는 하향식 영상 해석 방법을 소개하였다<sup>[3]</sup>. 이 방법에서는 에지 영상을 사용하여 머리의 탐색 영역을 설정한 후, 탐색 영역 내의 에지 영상을 원영상을 투시하여 보다 정확한 머리의 윤곽을 찾아내었다. 일단 머리가 추출되고 나면, 같은 방법으로 눈, 코, 입 등의 특징이 추출된다.

영상 내에서 얼굴을 분할하기 위한 또 다른 방법으로는 얼굴에 대한 형판을 사용하는 방법이 있다. Craw 등은 주어진 영상에서 계층적인 크기의 영상과 형판을 사용하여 머리 영역을 추출하는 방법을 제시하였다<sup>[4]</sup>. 이 방법에서는 다양한 해상도의 영상이 사용되었으며, 머리 형판은 가장 낮은 해상도에서 구성되었다. 에지의 크기와 방향은 명도 영상에서 Sobel 연산자를 사용하여 검출되었으며, 머리의 윤곽선을 연결하기 위해서 방향 추적자가 사용되었다. 눈과 눈썹, 그리고 입술과 같은 특징들 또한 이와 유사한 방법으로 추출된다. Govindaraju는 복잡한 배경의 영상에서 얼굴의 위치를 찾기 위해 에지 영상으로 이루어진 머리 윤곽을 형판으로 사용하였다<sup>[5]</sup>. 머리 형판은 오른쪽 선, 왼쪽 선, 그리고 머리선의 3개의 부분으로 구성되며, 각각은 곡선의 길이, 벡터 형식의 현, 현과 곡선 사이에 둘러싸인 영역과 영역의 무게 중심으로 이루어진 4차원의 값으로 구성된다. 머리를 구성하기 위해서는 이 세 개의 성분이 모두 일정한 방향에 위치하여야 하며, 이 때 세 성분의 중심은 얼굴의 중심 위치가 된다. 이 방법에서 사용하는 형판은 특정한 스프링 기반의 모델에 따라서 크기 변형과 회전이 가능하다.

위의 두 방법과는 조금 다른 방법으로, Sirohey는 배경이 있는 영상으로부터 Canny의 에지 탐색기로 추출된 에지 영상과 명도 영상을 함께 사용하여 얼굴을 분할하였다<sup>[6]</sup>. 얼굴을 분할하기 위해서 우선 전치리를 통하여 에지 성분들 간의 교차점을 구하고, 선을 구성하는 에지 성분에 레이블을 할당한 후, 교차점에서 비슷

한 모양의 선을 형성하고 있는 에지 성분들을 서로 연결시킨다. 사람의 얼굴은 타원에 가깝기 때문에, 레이블된 에지 성분의 짝은 타원 방정식에 의해서 표현되어지며, 방정식에 사용되는 인자값들은 얼굴의 형태에 따라서 조정된다. 이 방법은 제약을 가하지 않은 48개의 영상에 대해 약 80%의 정확도를 보여 주었다.

### Ⅲ. 객체 분할 알고리즘

본 연구에서는 동일한 환경에서 자동적으로 입력영상에서 얼굴영역을 분할하고 분할된 얼굴영역에서 특징을 검출하는 알고리즘을 제안한다.

#### 1. 얼굴영역 분할

객체 기반 영상압축을 위해서는 배경영역과 얼굴영역을 정확히 분할해야 한다. 본 연구에서는 일정한 조도 상태에서 취득된 두 입력영상에 대해 차영상을 적용하여 배경영상에서 얼굴을 분할한다. 이를 위해 비디오로부터 동일거리에 있는 영상을 256×256 크기의 그레이 스케일(Gray Scale) 256 레벨로 취득하여 가우시안(Gaussian) 필터를 이용하여 영상 내에 존재하는 잡음을 제거한다. 취득된 두 입력영상은 비디오가 항상 가리키고 있는 배경영상과 배경영상 내에 사람이 포함된 경우의 영상이다.

다음으로 배경 영상과 얼굴이 포함된 입력 영상의 차를 구한다. 차영상을 적용할 때 실제 환경에서는 동일한 카메라로 동일한 영역을 취득하였다 하더라도 같은 영역에 대해 일정한 픽셀값을 얻기가 어렵고, 조도 상태도 미소하게 변한다. 따라서, 얼굴이 포함된 입력영상에서 얼굴영역 밖의 픽셀 값과 얼굴이 포함되지 않는 배경영상의 픽셀값이 정확히 일치하지 않으므로, 차영상을 구할 때에 다음과 같이 임계값을 주어 동일 카메라로 얻어진 영상의 픽셀 값이 미소하게 변화해도 원하는 영상을 구할 수 있도록 한다.

```

if
  |Image1(x, y) - Image2(x, y)| < threshold
  then Differ_Imgae(x, y) = 0
else
  Differ_Imgae(x, y)
  = |Image1(x, y) - Image2(x, y)|
  
```

(1)

단, Image1 : 배경 영상.

Image2 : 얼굴이 포함된 입력영상.

Differ\_Image: 차영상

식 (1)에서 배경의 픽셀 값과 얼굴의 픽셀 값이 비슷한 경우, 차영상을 적용하면 얼굴 영역의 일부가 배경 픽셀 값에 의해 제거된다. 또한 배경 영역에서 임계값보다 크게 변화되는 부분은 차영상 내에서 잡음으로 발생하게 되므로 이를 직접 마스크로 사용할 수가 없다. 하지만 약간의 조도 변화와 빛의 반사에 의하여 생기는 잡음은 발생 확률이 적어 영상 내에서 미소 면적을 차지하게 된다. 따라서, 차영상을 이진화 하면서 영상을 축소시킨다. 얼굴영역이 지나치게 축소되면 마스크가 얼굴영상보다 작아질 우려가 있다. 따라서 얼굴영상의 크기에 맞게 영상을 팽창시켜야 한다. 즉, 영상의 축소와 팽창을 이용하여 얼굴영역 내의 잡음을 완전히 제거한 후 좌, 우에서 픽셀 값을 조사하여 마스크를 생성한다. 그리고 마스크를 얼굴이 존재하는 원 영상에 투영하여 배경과 얼굴영역을 분할한다. 그림 1은 얼굴영역을 분할하기 위한 순서도를 나타낸다.

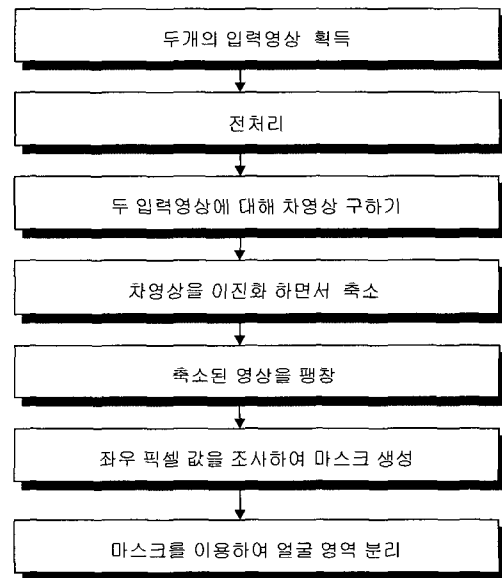


그림 1. 얼굴 분할을 위한 순서도  
Fig. 1. Flowchart for face segmentation.

#### 2. 특징 검출

사람의 얼굴에서 대부분의 특징정보는 눈썹, 눈, 코,

입, 뺨에 존재한다. 특히 뉴스를 중계한다거나 스포츠를 중계할때는 일반적으로 배경영상이 일정하고 아나운서의 입과 눈이 대부분의 변화정보가 된다. 따라서 효과적인 비디오 코딩을 위해서는 얼굴영역을 검출하는 작업 외에 얼굴의 특징검출이 중요한 과제가 된다.

분할된 얼굴영상에서 특징검출을 위해 Sobel 연산자를 이용하여 에지를 검출한다. 검출된 에지 성분들의 수평성분과 수직 방향 성분을 이용하여 눈의 영역을 찾는다.

또한 비디오로부터 일정한 거리에 있는 사람의 얼굴의 크기는 비슷하므로 픽셀의 거리를 지정하여 눈썹, 눈, 코 그리고 입을 각각 분할하였다. 그림 2는 특징검출을 위한 순서도이다.

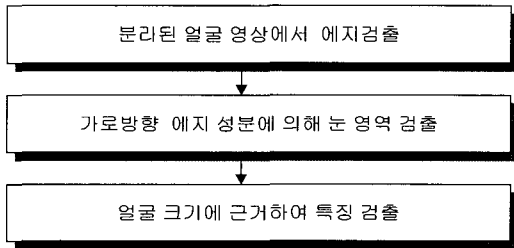


그림 2. 특징검출을 위한 순서도  
Fig. 2. Flow chart for feature detection.

#### IV. 시뮬레이션

배경이 일정하고 전경이 변하는 상황에서 영상데이터의 효율적인 전송을 위해서는 배경은 미리 전송하고 나중에 움직이는 객체만을 분할하여 전송하는 방법이 필요하다. 특히 토크 쇼와 뉴스 중계 등에서는 배경이 잘 변하지 않는다. 이와같은 상황에서 움직이는 물체만을 분할하는 작업은 중요한 문제가 된다.

본 연구에서는 일정한 조도상태에서 비디오 카메라로부터 동일거리에 있는 영상을 256×256 크기의 그레이 스케일 256 레벨로 취득하여 PC(266MHz)상에서 C언어를 이용하여 실험을 수행하였다.

이를위해 입력 영상을 가우시안(Gaussian) 필터를 이용하여 영상 내에 존재하는 잡음을 제거하고 전처리를 수행한 후 차영상을 이용하여 배경영상에서 얼굴을 분할하였다.

특징 추출을 위해 30인을 대상으로 입력영상으로 배경 및 얼굴이 포함된 영상을 각각 2장씩 취득하였다.

이 때 얼굴을 쉽게 분할하기 위하여 배경영상을 먼저 취득하고 다음으로 얼굴이 포함된 영상을 취득한다. 그림 4와 5는 이와 같은 입력영상을 나타낸다.

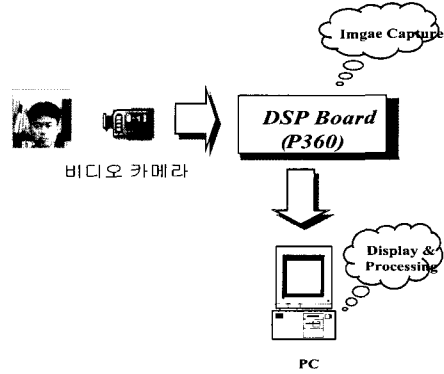


그림 3. 얼굴 검출 시스템  
Fig. 3. A system for Face detection.

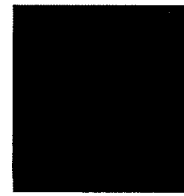


그림 4. 입력 영상(I)  
Fig. 4. Input image(I).



그림 5. 입력 영상(II)  
Fig. 5. Input image(II).

다음으로 취득된 두 입력영상에 대해 차영상을 구했다. 그림 6은 구해진 차영상을 나타낸다. 그런데 이 차



그림 6. 차영상  
Fig. 6. Differential image.

영상을 바로 마스크로 이용할 수는 없다. 따라서 차영상의 화소값을 0과 1로 변환하여 영상을 축소하고 확장하는 작업을 수행하였다.

그림 7은 확장된 영상을 이용하여 좌우에서 픽셀값을 조사하여 생성한 마스크 영상이다.



그림 7. 마스크 영상  
Fig. 7. Mask image.

마스크 영상을 얼굴이 존재하는 원영상에 투영하여 그림 8과 같이 배경에서 얼굴을 분할하였다.



그림 8. 분할 영상  
Fig. 8. Segmentation image.

분할된 얼굴영상에서 지식기반 특징 검출을 위해 그림 9와 같이 에지를 검출하여 가로 방향으로 에지 성분 분함을 구했다. 그림 10을 보면 에지성분의 가로방향 합 중에서 가장 큰 값을 갖는 영역이 눈 주위임을 알 수 있는데 여기에서 눈의 위치를 검출하였다.



그림 9. 에지 영상  
Fig. 9. Edge image.

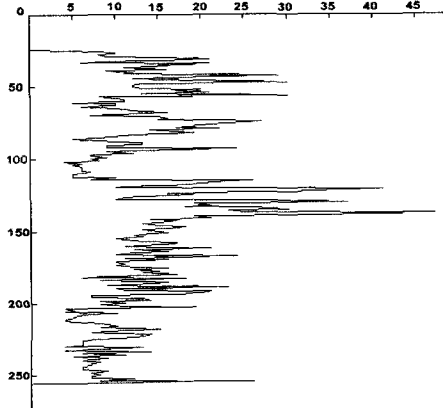


그림 10. 에지의 합  
Fig. 10. Sum of Edge values.

그리고 카메라로부터 일정한 거리에 있는 사람의 얼굴의 크기는 비슷하므로 그림 11과 같이 얼굴영역에 대해 유클리디안 거리를 설정하여 눈썹, 눈, 코, 입을 각각 검출하였다.

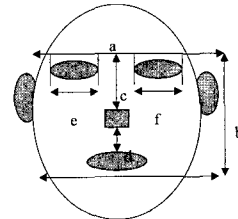
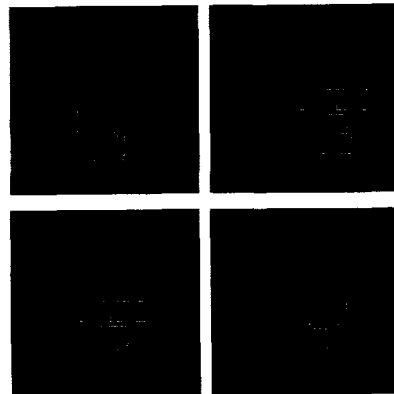


그림 11. 특징 검출을 위한 정규화된 거리  
Fig. 11. Thenormalized distance for feature detectioin.

그림 12는 얼굴의 특징검출의 예이다. 제안된 방법을 이용하여 전체의 입력영상에서 영상 내의 얼굴 위치에 무관하게 특징영역을 100% 분할할 수 있었다.



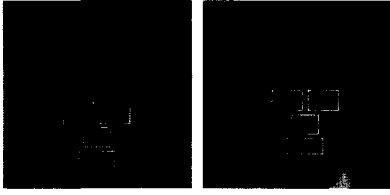


그림 12. 지식기반 특징검출  
Fig. 12. Characteristic detection.

### V. 결 론

본 연구에서는 영상통신을 위한 객체 분할 알고리즘과 지식기반 특징검출 알고리즘을 제안하였다. 이를 위해 일정한 조도 상태에서 배경영상을 먼저 취득한 후 사람이 포함된 영상을 취득하였다. 다음으로 두 영상에 대해 차영상을 구한 후 이진화 하여 축소와 팽창을 한 후 마스크를 생성하였고 이를 원 영상에 투영하여 얼굴을 분할하였다.

시뮬레이션 결과 획득된 영상에 대해 얼굴영역이 100% 분할되었고 또한 얼굴영상의 에지 특성과 인간 얼굴에 대한 지식을 이용하여 눈, 코, 귀, 입을 검출할 수 있었다.

따라서 제안된 알고리즘을 적용하면 배경이 미리 갖 추어지고 전경만 변화는 상황에서 변화는 전경만을 검출하여 전송함으로써 데이터 전송효율을 향상시킬 수 있다.

향후 연구방향은 다양한 조도 환경에서 얼굴을 분할 하는 문제이다.

### 참 고 문 헌

- [ 1 ]. 원치선 “내용기반 영상 압축을 위한 영상분할”, 한국통신학회지, 제14권 9호, pp. 96-106, 1997년 9월.
- [ 2 ] T. Sakai, M. Nagao, and Fukibayashi : Line extraction and pattern recognition in a photograph: *Pattern Recognition*, Vol. 1, pp. 233~248, 1969.
- [ 3 ] M.D. Kelly: Visual identification of people by computer: Tech. Rep. AI-130, Stanford AI Proj., Stanford, CA, 1970.
- [ 4 ] I. Craw, H. Ellis, and J. Lishman: Automatic extraction of face feature: *Pattern Recognition Lett.*, Vol. 5, pp. 183~187, 1987.
- [ 5 ] V. Govindaraju, S. N. Srihari, and D. B. Sher : A computational model for face location: *Proc. 3rd Int. Conf. on Computer Vision*, pp. 718~721, 1990.
- [ 6 ] S. A. Sirohey: Human face segmentation and identification: Tech. R.CAR-TR-695 : Center for Automation Research, Univ. of Maryland, 1993.

### 저 자 소 개



林 春 煥(正會員)  
1991年 朝鮮大學校 電子工學科 卒業 (工學士). 1993年 8月 朝鮮大學校 電子工學科 卒業(工學碩士). 1999年 2月 朝鮮大學校 大學院 博士課程 修了. 1999年 3月 東岡大學 電子科 兼任教授. 주관심분야는 영상통신, 컴퓨터 비전, 신경망, 자율이동로봇등임



梁 洪 榮(正會員)  
現在 東岡大學 電子科 教授. 주관심 분야는 영상통신, 컴퓨터 비전, 신경망, 자율이동로봇등임