

## 효과적인 뉴스 검색을 위한 얼굴 영역의 추출

윤 일 태, \*정 승 도, \*조 정 원, \*\*배 영 래, \*\*\*최 병 옥  
한양대학교 정보통신대학원, \*전자통신전파공학과,  
\*\*한국전자통신연구원, \*\*\*전자전기컴퓨터공학부  
전화 : 02-2290-0363 / 핸드폰 : 011-476-8513

### Detection of face region for an effective News Retrieval

Il-Tae Yoon, Seung-Do Jeong, Jung-Won Cho, Young-Rea Bae, Byung-Uk Choi  
The graduate school of information & communication, Hanyang University  
E-mail : President@mlab.hanyang.ac.kr

#### Abstract

The retrieval techniques of multimedia contents have been developed along with MPEG-7.

But as human being distinguishes objects with the eyesight, researches for high retrieval efficiency applied a high level computer vision technology are difficult because of the increase of processing time caused by complexity of algorithm, and the difficulty of an implementation.

In this paper, for an effective news retrieval using the human face information, we suggest a method which extracts face information like as location, size and number of the face in the news video, and then we prove the validity of method by experiment.

양한 물체(안경, 수염, 화장) 등이 원인이 되고 있다. 또한 얼굴 검출의 문제점과 함께, 얼굴 표정 변화와 화면상의 얼굴 크기 변화, 포즈 변화에 따라 인식에 있어서도 어려움이 따른다.

본 논문에서는 인물 위치와 크기 정보 등을 이용하여 효과적인 동영상 검색이 가능하도록, 관련 연구 결과와 비교하여 얼굴 검출의 정확도를 향상시키고, 검출 알고리즘을 인물 자체가 갖고 있는 정보량이 가장 크다고 할 수 있는 뉴스 비디오에 적용하여 본 논문의 타당성을 입증하고자 한다.

실제 뉴스 비디오에서 앵커 프레임의 경우는 인물의 등장 위치와 수가 상당히 정형화되어 있기 때문에, 본 논문에서 제시하는 방법으로 얼굴의 존재 유무를 판별하고 앵커 프레임을 검출할 수 있었다.

#### I. 서론

MPEG-7과 더불어 디지털 미디어 콘텐츠의 검색 기술은 날로 발전을 하고 있다. 멀티미디어 데이터에 포함된 인물에 대한 정보는 검색의 효율성을 높이는 데 크게 기여할 수 있으나, 그 기반 기술인 얼굴 검출과 인식 기술은 고도의 컴퓨터 비전 기술을 요구하기 때문에 검색 시스템으로의 응용에 어려움이 따랐다.

대표적인 얼굴 검출의 문제점으로는 인종에 따른 얼굴색의 변화, 3차원 얼굴 형상에 따른 조명의 변화, 다

#### II. 얼굴 검출

##### 2.1 얼굴 검출의 전체 시스템 구성

본 논문에서 제시하는 전체 시스템은 얼굴 검출 시스템 부분과 데이터베이스에서 검색하는 두 부분으로 구성되어 있다.

우선 얼굴 검출 시스템은 크게 3가지 처리 과정을 거친다. 처리의 순서는 전처리 과정과 일차 후보 영역의 검출 그리고 이차 영역의 검출 과정을 거치게 되

며, 결과적으로는 영상 내에서 얼굴의 존재 정보와, 얼굴의 개수, 얼굴의 위치 정보가 데이터베이스에 저장된다. 그림 1은 얼굴 검출 시스템의 구성을 보여주고 있다. 그림 2는 데이터베이스에 저장되는 얼굴 정보를 나타내고 있다.

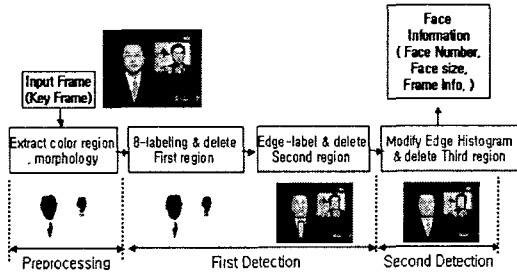


그림 1. 얼굴 검출 시스템의 구성

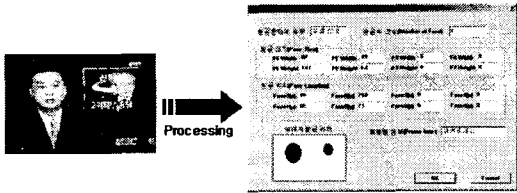


그림 2. 데이터베이스에 저장되는 얼굴 정보

각 프레임별로 얼굴에 대한 정보가 구축되기 때문에 사용자의 다양한 질의(예를 들면, 얼굴 개수가 2개인 프레임, 앵커 프레임 등)를 허용하는 검색이 가능하도록 하였다. 그림 3은 전체 시스템의 구성을 보여주고 있다.

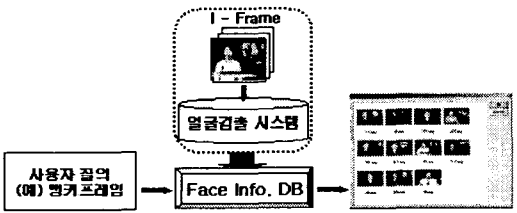


그림 3. 전체 시스템의 구성

## 2.2 얼굴 추출을 위한 전처리 과정

뉴스 영상에서 색상 정보를 이용하여 의미적으로 중요한 가치를 갖는 앵커 프레임의 얼굴 정보를 추출할 수 있다.

전처리 과정에서는 얼굴 색상 정보를 이용하여 일차적으로 얼굴 후보 영역을 검출한다. 먼저 얼굴색의 분포를 살펴보면 RGB 컬러 스페이스를 이용하는 것보다

는 YCbCr 컬러 스페이스나 HSI 컬러 스페이스를 이용하는 것이 효과적으로 얼굴 후보 영역을 검출할 수 있다[1]. 이것은 얼굴로만 구성된 Thumbnail 이미지에 대하여 각 컬러 성분의 분포를 조사하면 RGB 컬러 스페이스보다 YCbCr 컬러 스페이스에서 각 성분의 분산이 작음을 통해 알 수 있다. 그림 4는 컬러 스페이스의 분포를 보여주고 있다.

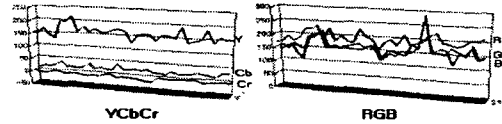


그림 4. 얼굴 색의 각 성분별 분포

본 논문에서는 실험 결과를 바탕으로 YCbCr 컬러 스페이스를 이용하여 일차적인 얼굴 후보 영역을 검출하였다. 그러나 일차적인 후보 영역의 이미지에는 많은 노이즈를 포함하고 있어 노이즈를 제거하고, 조명이나 그림자에 의하여 얼굴 영역이 분리될 수 있기 때문에 모폴로지(morphology) 확장(dilation) 연산을 적용하였다. 그림 5는 전처리 과정을 보여 주고 있다.

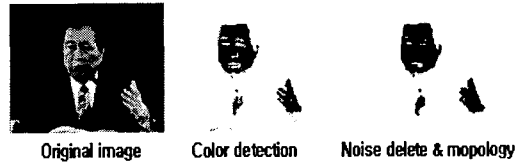


그림 5. 전처리 과정

## 2.3 일차 후보 영역 검출

이미지에는 여러 사람의 얼굴이 포함될 수 있다. 예를 들어 앵커 프레임의 경우는 앵커만의 얼굴이 있을 수도 있으며, 어깨 길이 창에 얼굴을 포함하는 이미지가 있을 수도 있다. 따라서 2개 이상의 얼굴 영역도 정확하게 검출해 낼 수 있어야 한다.

얼굴을 효과적으로 분리해 내는 방법으로 기존의 연구에서 누적 히스토그램을 이용한 방법[2], 레이블링을 이용하는 방법[3] 등이 제시되어 있으나, 누적 히스토그램의 경우는 얼굴이 같은 축 상에 연속으로 존재하는 경우에는 효과적으로 얼굴을 분리하지 못하며, 레이블링을 이용하는 방법은 얼굴의 효과적인 분리는 가능하지만 많은 계산 시간이 필요하다는 단점을 가지고 있다.

본 논문에서는 8방향 레이블링을 이용하여 얼굴 영역을 분리하였으며, 계산 시간을 줄이기 위하여 한번 레이블이 된 곳은 스탬프를 두어 건너뛰도록 하는 개

## 효과적인 뉴스 검색을 위한 얼굴 영역의 추출

선된 레이블링 알고리즘을 적용하였다. 이러한 과정을 거치게 되면 같은 숫자로 레이블링된 오브젝트의 MBA(Minimum Boundary Box)를 구할 수 있게 된다. 이 때 영상 내에 존재할 수 있는 얼굴의 크기 조건과 비례 조건을 MBA에 적용하여 간단하게 얼굴이 아닌 영역을 제거 할 수 있게 된다.

다음으로 이미지 전체가 아닌 MBA구역에서 수평 방향 에지(edge) 성분만 검출을 하여 다시 검출된 에지에 대하여 8방향 레이블링을 수행한 후 레이블된 에지 들에 대하여 새로운 Edge\_MBA를 생성한다. 이 때 얼굴의 에지 성분으로 나타나는 눈썹, 눈, 입의 크기와 비례 조건을 이용하여 눈썹, 눈, 입의 후보가 될 수 없는 에지 들을 포함하는 Edge\_MBA를 제거한다. 즉 최종적인 MBA의 후보는 MAB구역 내의 Edge\_MBA의 수가 5개 이상인 MBA들이다. 그림 6은 이러한 과정을 나타내고 있다.

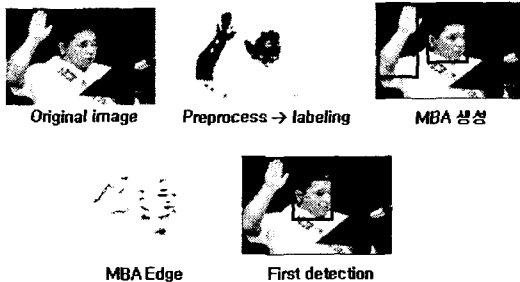


그림 6. 일차 후보 영역의 추출 과정

### 2.4 이차 영역 검출

일차 후보 영역의 검출 과정 후에도 얼굴 영역이 아닌 부분을 여전히 얼굴로 인식하는 경우가 있다. 따라서 이차 영역 검출 과정을 통하여 더욱 정확한 얼굴 영역을 검출한다.

이차 영역 검출 과정에서는 이미 전처리 과정에서 컬러 정보를 이용하여 얼굴의 후보 영역을 검출하였기 때문에 얼굴이 가지는 색상 특성을 제외한 고유한 특성을 이용하였다. 즉 얼굴이 가지는 에지의 특성 정보를 이용하여 최종적인 얼굴 영역을 검출하도록 하였다.

얼굴의 에지 성분 특성을 보면, 수직 성분의 에지 성분보다는 수평 성분의 에지의 개수가 더욱 많다. 그리고 정규화된 얼굴 MBA구역에서의 에지 성분 분포는 거의 비슷하다.

본 논문에서는 MPEG-7에서 제안하고 있는 Texture Descriptor 중에서 Edge Histogram[4]을 변형하여 최종 얼굴 영역을 검출하였다. MPEG-7의 Edge

Histogram은 총 125개의 특징 벡터를 추출하게 된다. 이 중에서 5개는 Global Edge Histogram이며, 80개는 이미지를 16등분한 각각의 부(sub) 이미지에 대한 5방향의 Local Edge Histogram이다. 또한 Semi Local Edge Histogram을 구하여 나머지 40개의 특징 벡터를 구성한다. 그림 7은 5방향의 에지 성분을 보여주고 있다.

1 -1	1 1	1.4 0	0 1.4	2 -2
1 -1	-1 -1	0 -1.4	-1.4 0	-2 2
수직방향	수평방향	45°방향	135°방향	전체방향

그림 7. 5방향 Edge 성분

본 논문에서는 MBA구역의 이미지만을 대상으로 특징 벡터를 구하기 때문에 특징 벡터의 수를 줄일 필요가 있다. 즉, MBA구역의 이미지를 64×64 이미지로 정규화 한 후 Global Edge Histogram과 Semi Local Edge Histogram의 특징 벡터만을 추출하였다. Local Edge Histogram의 특징 벡터를 제외한 이유는 정규화된 이미지의 크기가 작으므로 크게 중요한 요소로 작용을 하지 못하기 때문이다.

총 45개의 특징 벡터를 이용하여 미리 학습된 얼굴 데이터의 45개의 특징 벡터들과 비교하여 얼굴 영역을 최종적으로 판단한다. 또한 MPEG-7의 Edge Histogram의 경우는 5방향의 방향성을 조사한 후 가장 큰 값을 갖는 성분의 특징 벡터를 카운트하였는데, 본 논문에서는 5방향의 성분에 대한 에지의 수 진부를 카운트한다.

그림 8은 전처리 과정과 일차 후보 영역으로 검출한 뒤 마지막으로 후보 검증 처리를 한 결과 이미지를 보여주고 있다.

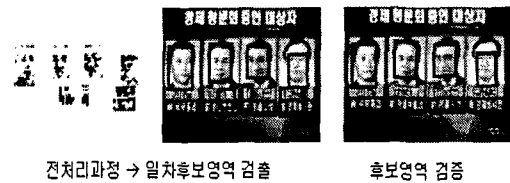


그림 8. 이차 영역 검출

### 2.5 Ground Truth Table

Ground Truth Table이란 MPEG-7에서 사용되는 정답 이미지들의 그룹이다. 즉, Ground Truth Table의 작성을 통해 실험 데이터에 대한 검출물을 객관적으로 증명해 보일 수 있다.

본 논문에서는 각각의 이미지에 대하여 얼굴에 대한 존재 여부와 얼굴의 개수, 얼굴의 위치, 그리고 프레임 정보를 Ground Truth Table로 작성하였다.

그러나 각각의 이미지에 대하여 개개인의 관점에 따라서는 얼굴 존재의 유무 판정을 내리기가 곤란한 경우가 있다. 본 논문에서는 주관적인 견해를 최대한 줄이기 위하여 얼굴을 통하여 그 사람이 누구인지 식별이 가능한 경우에만 얼굴이 존재하는 것으로 판정하였다.

그림 9에서 Ground Truth Table을 작성하는 예를 보여 주고 있다.

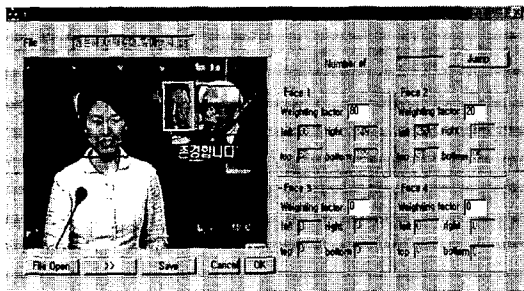


그림 9. Ground Truth Table의 작성

### III. 얼굴 검출 결과

본 논문에서는 시스템의 성능을 보다 객관적으로 판단하기 위하여 실험에 사용 될 데이터를 MPEG-1으로 압축된 KBS, MBC, SBS 뉴스 동영상의 I-Frame을 무작위 추출하여 634장의 실험 데이터 셋을 구성하였다. 또한 얼굴 검출 결과의 성공률을 Ground Truth Table과의 비교를 통해 산출하였다.

Ground Truth Table은 수작업으로 작성하였으며, 프레임 정보로는 뉴스 검색 시스템에서 중요한 분할 등의 중요한 정보로 사용되고 있는 앵커 프레임으로 우선 한정하였다.

얼굴 검출 알고리즘을 통하여 검출된 얼굴 정보는 얼굴 존재의 유무, 얼굴의 개수, 유추할 수 있는 프레임 정보를 포함하고 있다.

표 1에서 얼굴 유무의 성공률은 634장 전체 이미지에 대하여 작성하였으며, 얼굴의 개수는 얼굴이 있는 이미지에 대하여 작성을 하였다. 그리고 프레임 정보는 Ground Truth Table에서 앵커 프레임으로 판단한 이미지에 대하여 검출률을 작성하였다. 표에서 볼 수 있듯이 얼굴의 개수가 증가될수록 검출률이 떨어졌는데, 이는 얼굴의 개수가 증가함에 따라 이미지의 색상

복잡도가 증가하기 때문으로 볼 수 있다.

표 1. 실험 결과

Ground Truth Table			제안한 방법		
I-Frame 인물 정보		개수	검출 개수	검출율	
Face of Face (224개)	Number 1개	121	107	88%	
	2개	61	43	70%	
	3개	29	19	66%	
	4개 이상	13	8	62%	
앵커 프레임		53	48	91%	
Non-face I-Frame(410개)		410	373	91%	

전체적인 오검출의 요인은 일차적인 후보 영역의 선정이 색상 정보를 이용하기 때문에 마지막 이차 영역의 검출 과정으로도 제거되지 못한 이미지가 대부분이다.

### IV. 결론

본 논문에서는 얼굴 정보를 통하여 뉴스 검색의 효율성을 높일 수 있는 방법을 제시하였다. 결과에서 보듯이 얼굴 정보를 통해 앵커 프레임을 검색할 수 있었으며, 얼굴이 존재하는 프레임을 조건에 따라 검색하는데 이용할 수 있다.

향후 연구 과제로는 검출된 영역의 물리적인 위치와 크기 정보 이외에도 얼굴 인식 알고리즘을 적용하여 보다 구체화된 얼굴 정보를 얻을 수 있도록 하여야 할 것이다. 또한 뉴스뿐만 아니라 다양한 종류의 동영상 데이터에 대한 연구도 진행되어야 할 것이다.

### 참고 문헌

- [1] Douglas Chai and King N. Ngan, "Location Facial Region of Head and Shoulders Color Images", International Conference Automatic Face and Gesture Recognition, pp. 124-130, April 1998
- [2] A.Yuille, P.Hallinan, and D.Cohen, "Feature extraction from faces using deformable templates" Int.J. Computer Vision, Vol.8 No.2, pp. 99-111, 1992
- [3] 윤호섭 외 2인, "눈 영역 추출에 의한 얼굴 기술기 교정", 대한전자공학회 논문지, 제33권, 제12호, pp. 1886-1897, 1996. 12
- [4] ISO/IEC JTC1/SC29/WG11/N3321, "Mpeg-7 Visual part of eXperimentation Model Version 5.0", Nordwijkerhout, March 2000