

얼굴 영역 추출에 의한 장면 전환 검출에 관한 연구

최경애[°], 최기호
광운대학교 컴퓨터공학과

A Study on Scene Change Detection Using Facial Regions Extraction

KyungAe Choi, KiHo Choi
Dept. of Computer Engineering, Kwangwoon Univ.
E-mail : kyung-ae@hanmail.net, khchoi@gwu.ac.kr

요 약

본 논문에서는 효과적인 비디오 인덱싱을 위해 얼굴 영역 추출을 통한 장면 전환 검출 방법을 제시하였다. 히스토그램과 사람의 피부색 검출을 통해 사람의 얼굴을 포함하는 후보 프레임을 찾고, 얼굴 영역과 특징 추출을 통해 사람을 포함하는 키 프레임을 검출하여 비디오의 장면 전환 프레임을 검출하고, 실험을 통해 제안된 방법의 우수성을 보였다.

1. 서론

초고속 통신망의 보급과 멀티미디어 관련기술의 발달로 인하여 오디오, 동영상 등과 같은 대용량의 멀티미디어 데이터가 증가하고 영상을 효과적으로 검색하는 방법이 중요한 관심사로 대두되어 연구가 활발히 진행 중이다.

최근에는 비디오 검색에 대한 연구가 활발하게 이루어 지면서, 영상에 의한 질의(Query-by-image)와 얼굴의 의한 질의(Query-by-face)와 같은 다양한 형태의 질의를 만족시켜주는 시스템들이 필요하게 되었고 이에 따른 연구들이 활발히 진행되고 있다[1]. 이러한 비디오 검색을 위해서 최근에는 뉴스 동영상에서 앵커 장면을 추출하는 방법과 축구 동영상에서 장면전환 검출을 하는 방법들에 대한 연구가 진행되고 있다.

일반적으로 동영상에서 인접한 프레임은 시간적으로 많은 차이가 나지 않기 때문에 서로 비슷한 특성을 띠게 된다. 그러나 장면 전환이 일어났을 때에는 프레임간의 상관성이 현저히 줄게 되어 프레임의 특성간에 차이가 많이 나게 된다. 따라서 장면 전환 검출을 하기 위해서는 이런 프레임간의 차이를 이용한다. 프레임의 차이를 나타내기 위해 히스토그램, 픽셀간의 차이, 애지 변화 등을 이용하여 장면 전환 검출을 한다[2, 3].

컴퓨터 시스템을 이용하여 인간의 얼굴을 분석하는 방법은 매우 복잡하고 변수가 많은 기술이지만, 별도의 장비나 추가적인 제약을 가하지 않고도 분석이 가능하다. 그러므로, 인간의 얼굴을 분석하기 위해서 얼굴의 세부 특정 요소를 분석하여 얼굴 영역을 처리하는 방법과 전체 얼굴 영역을 하나의 특징 영역으로 추출하는 방법 등의 연구가 활발하게 진행

중이다.[4, 5]

얼굴 영역을 추출하기 위한 방법으로는 에지를 이용한 방법, 템플릿 매칭(Template matching), 주성분 분석법(Principal component analysis), 신경망을 이용한 방법들이 있다. 이런 방법들은 그레이 스케일(Gray scale) 영상을 이용하여 얼굴 영역 추출이 이루어 진다. 그러나 조명등에 영향을 받는 경우 얼굴 영역 추출이 정확히 되지 못하는 경우가 발생한다.[4, 6, 7]

최근에는 주로 칼라 영상을 이용하여 얼굴 영역 추출이 이루어진다. 칼라 영상은 밝기 정보 이외에도 색상 정보를 포함하고 있으므로 얼굴과 같이 고유한 색상을 지닌 영역을 분할하는데 유리하다. 이러한 점을 이용하여 얼굴의 피부색을 이용하여 얼굴 영역을 추출하는 방법이 많이 이용되고 있다.

본 논문에서는 효과적인 비디오 인덱싱(indexing)과 검색(retrieval)을 위하여 얼굴 영역 검출을 통한 장면 전환 검출을 한다. 우선 히스토그램방법을 통해 비디오 샷들의 키 프레임을 추출하고, 추출된 키 프레임을 중심으로 피부색 검출을 통해 얼굴 후보 프레임을 추출한다. 이렇게 추출된 후보 프레임에서 사람의 얼굴을 포함하는 프레임을 추출하기 위해 얼굴 영역과 특징을 검출하며, 실험을 통해 그 유효성을 입증하였다.

2. 장면 전환 및 얼굴 영역 추출

본 논문에서는 사람을 포함하는 프레임의 추출을 추출하기 위하여 사람의 얼굴을 검출하여 정확히 사람으로 인식된 프레임을 장면 전환이 이루어지는 키 프레임으로 설정한다. 사람을 포함하는 프레임으로 장면 전환 검출하기 위해서 다음과 같은 과정을 거치게 된다.

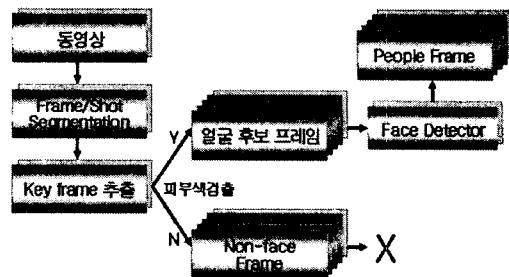


그림 1. 사람을 포함하는 프레임 추출 과정

2.1 장면분할

장면을 분할하기 위해서 가장 많이 사용하는 방법으로 히스토그램 비교(Histogram difference)방법들을 이용한다. 이 방법은 장면으로 분류해야 할 프레임의 색상분포가 거의 비슷하다는 성질을 이용하는데, 카메라의 움직임이나 회전, 물체의 움직임에 대해 영향을 적게 받는다. 그러나, 연속되는 두 샷이 비슷한 칼라 혹은 밝기 분포를 가질 때는 검출이 어렵다.[4]

이 방법은 조명 변화, 크기 변화에 민감하지 않기 때문에 많이 사용된다. 히스토그램의 차이를 구하는 식은 다음과 같이 표현된다.

$$H_i = \sum_{j=1}^n h_j \quad \text{식 (3.1)}$$

$$d(I_i, I_{i+1}) = \sum_{j=1}^n |H_i(j) - H_{i+1}(j)| \quad \text{식 (3.2)}$$

여기서 H_i 는 화상 안에서 칼라 i 의 전체 칼라 히스토그램을 나타내며, n 은 칼라 버켓의 수이다. 그리고, 식 3.2에서 나타난 I 는 동영상의 프레임을 나타내며 $d(I_i, I_{i+1})$ 은 연속된 프레임간의 히스토그램의 차이를 나타낸다.[4]

이러한 과정을 통해서 장면 전환을 검출하고, 장면 전환이 이루어진 대표 프레임으로 키 프레임(key-frame)을 추출한다. 이렇게 추출된 키 프레임들은 사람을 포함하는 프레임과 그렇지 않은 프레임으로 구분하기 때문에 사람을 포함하는 프레임들을 추출하기 위해 별도의 과정이 필요하다.

2.3 얼굴 추출

2.2 얼굴 후보 프레임 추출

얼굴을 포함하는 영상을 효율적으로 추출하기 위하여 사람의 얼굴이 갖는 특징을 바탕으로 키 프레임에서 얼굴 후보 프레임들을 추출하도록 한다. 본 논문에서는 사람의 피부색을 이용하여 얼굴 후보 프레임을 추출하도록 할 것이다.

칼라 영상에서 영역을 분할할 때는 칼라는 아주 유용한 정보이기 때문에 얼굴 영역 추출 및 인식 분야에서는 얼굴 후보 프레임을 추출해 내기 위해 칼라 정보를 많이 사용한다.[7,8] 본 논문에서 얼굴 후보 프레임을 추출하기 위해 먼저 RGB 영상을 인간의 칼라 인식과 밀접한 관련이 있는 YCbCr 칼라 공간으로 변환하여 처리한다.

$$\begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.5 \\ 0.5 & -0.419 & -0.081 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad \text{식 (3.3)}$$

YCbCr 영상으로 변환한 후 얼굴 후보 프레임을 추출하기 위해 Cb와 Cr값을 이용하여 처리한다. 영상에서 얼굴 후보 영역을 분리하기 위해서 피부색 정보를 많이 사용하는데, 이것은 사람의 피부색이 다른 객체들과 구별이 잘 되며, 연상이 빠르고 단순할 뿐만 아니라 사정 가정을 최소화 하기 때문에 최근에 많이 사용된다.

본 논문에서는 Chai의 논문에 나타난 피부색 참조 맵[9]을 이용하여 인종에 관계없이 피부색을 검출하여 얼굴 후보 프레임을 추출할 수 있도록 하였다.

$$B(x, y) = \begin{cases} 1 & \text{if } (77 \leq Cb \leq 127) \cap (133 \leq Cr \leq 173) \\ 0 & \text{otherwise} \end{cases} \quad \text{식 (3.4)}$$

식 (3.4)를 통해서 영상에서 피부색에 해당하는 영역을 모두 검출 할 수 있도록 하여, 이렇게 추출된 피부색을 통해서 얼굴 후보 영역을 추출한다.

본 논문에서는 기존의 얼굴 검출 방법을 개선해서 효과적으로 얼굴의 영역을 검출하도록 한다. 본 논문에서 제안된 방법은 다음과 같다.

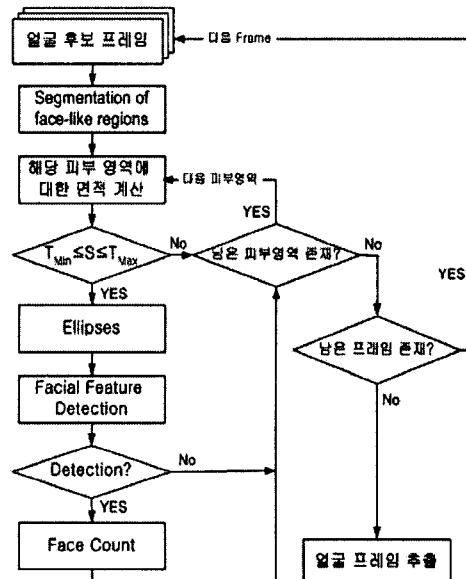


그림 2. 제안된 얼굴 추출 순서도

2.3.1 Segmentation of face-like regions

얼굴 영역을 검출하기 위해서는 먼저 피부 영역을 복잡한 배경에서 분리시켜야 한다. 장면에서 피부색 영역을 분리해 내기 위해 가로, 세로 scanning을 통해서 얼굴 영역을 검출해 낸다. 본 논문에서는 이러한 과정을 거친 영역을 가지고 얼굴에 대한 면적을 계산하여 작은 영역을 처리에서 제외시키도록 할 것이다.

본 논문에서는 면적을 계산하기 이전에, 영역들 각각에 인접해있는 영역이 존재하는지 검사하게 될 것이다. 인접 영역을 계산 함으로써 안경이나 마스크를 쓰고 있는 경우에도 얼굴 영역을 추출할 수 있도록 한다. 안경이나 마스크를 쓰고 있을 경우에는 얼

굴 영역이 인접하게 존재하기 때문에 이러한 과정은 수행한다.

인접한 두 개의 영역 A_1, A_2 가 존재한다면, 우선은 영역 사이의 거리 계산을 통해서 인접 여부를 결정한다. 두 영역은 동일한 x좌표상에 존재한다면, y좌표의 거리를 계산하여 인접 여부를 결정한다. 위에 있는 영역을 A_1 이라 하고 아래에 있는 영역을 A_2 라 할 때, A_1 영역의 최소 y좌표에서 A_2 영역의 최대 y좌표를 비교하여 임계값을 만족하면 두 개의 영역은 인접해 있다고 설정한다.

위의 조건을 만족하였다면 다음으로는 면적에 대한 연산을 통해서 영역에 대한 처리를 한다. 영역을 처리하기 위해서 A_1, A_2 의 면적을 구하는데 이때 구해진 면적을 a_1 과 a_2 라 한다. 두 영역의 면적을 비교하여 큰 영역에서 작은 영역의 면적을 나누어서 두 영역의 관계를 구한다. 이때 나눈 값을 S_1 이라 한다. a_1 영역이 a_2 영역보다 크다고 할 때 다음과 같은 식으로 나타난다.

$$S_1 = a_1 / a_2, \quad 0.3 \leq S_1 \leq 0.5 \quad \text{식 (3.5)}$$

식 (3.5)을 만족하게 되면 두 영역은 하나의 얼굴 일 가능성성이 높은 영역이므로, 두 영역을 합하여 얼굴 추출이 이루어 진다. 인접해 있는 두 영역의 면적이 같을 때는 서로 다른 얼굴일 가능성이 높으므로, 각각에 대한 얼굴 추출이 이루어 진다.

2.3.2 얼굴 영역 결정

위와 같은 과정을 수행한 영역에서 해당 영역에 대한 면적을 계산하여 너무 작은 영역은 얼굴 추출에서 제외를 시킬 것이다. 영역의 면적을 계산하기 위해 영역을 사각형으로 추출해 낸다.

$$S_2 = H/W, \quad 1 \leq S_2 \leq 2.6 \quad \text{식 (3.6)}$$

식 (3.6)은 얼굴의 비율을 계산하기 위한 것으로, 여기서 H 는 영역의 가로이고 W 는 영역의 세로이다. 본 식을 통해서 얼굴의 정면과 측면을 추출할 것이

다.

2.3.3 얼굴 영역 추출

얼굴 면적에 대한 연산을 수행 후 얼굴 영역을 타원으로 표현하여 얼굴 영역과 비얼굴 영역을 결정할 것이다. 본 논문에서 기본적인 타원 연산을 통해서 불필요한 피부 영역을 제거하고, 이 영역 내에서 애지 검출을 통해 얼굴의 특징을 추출할 것이다.

$$\frac{(x - \alpha)^2}{a^2} + \frac{(y - \beta)^2}{b^2} = 1 \quad \text{식 (3.7)}$$

이와 같은 과정을 프레임 내에서 검출된 영역들 사이에서 반복 수행하면서 추출된 얼굴 영역의 수를 계산한다. 이 과정을 통해서 프레임 내에 사람수를 파악할 수 있다.

이런 얼굴 영역 추출을 통해서 최종적으로 사람이 있다고 판단된 프레임들은 비디오의 키프레임으로 설정하여 장면이 전환되는 부분을 사람이 나오는 장면으로 재구성 한다

3. 실험 및 고찰

본 논문에서는 펜티엄 864MHz PC에 Microsoft Windows XP환경에서 Visual C++ 6.0을 이용하여 구현하였고, AVI 형식의 동영상 파일을 초당 30프레임, 320X240 픽셀로 입력 받아 실험에 사용하였다.

동영상의 각 프레임은 히스토그램을 이용하여 45% 이상 변화된 프레임을 장면 전환 된 키 프레임으로 하였다. 그림 3은 이러한 과정을 거쳐 추출된 키 프레임들을 보여준다.



그림 3. 히스토그램을 이용한 키 프레임 추출

여기서 추출된 키 프레임 중 사람의 피부색 검출을 통해 얼굴 후보 프레임이 결정된다. 그림 4는 추출된 얼굴 후보 프레임들을 보여준다.



그림 4. 얼굴 후보 프레임 추출

얼굴 후보 프레임들에서 얼굴 영역 추출과 특징들을 추출함으로써 사람을 포함하는 프레임들을 검출하였다

본 논문에서는 10분짜리 드라마 영상을 실험에 사용하여 표 1과 같은 결과를 얻었다. Ft는 전체 장면 전환 수, Fs는 후보 프레임이 검출된 수, Ff는 얼굴 영역이 추출된 프레임 수, Fn은 얼굴 영역이 아닌데 추출된 수, Fu는 얼굴 영역으로 추출되어야 하는데 추출되지 못한 횟수를 나타낸다.

Ft	Fs	Ff	Fn	Fu
319	248	186	18	7

표 1. 실험 결과

4. 결론

본 논문에서는 효과적인 비디오 인덱싱을 위해 얼굴 영역 추출을 통한 장면 전환 검출 방법을 제시하였다. 히스토그램과 사람의 피부색 검출을 통해 사람의 얼굴을 포함하는 후보 프레임을 찾고, 얼굴 영역과 특징 추출을 통해 사람을 포함하는 키 프레임을 검출하였으며, 실험을 통해 제안된 방법의 우수성을 보여주었다..

[참고문헌]

- [1] Kim Jong Bae, "Content-based Retrieval System Design for Image and Video using Multiple Feature" Journal of KISS Vol.26, no.12, 1999
- [2] 이미숙, 황본우, 이성환, "내용기반 영상 및 비디오 검색 기술의 연구 현황", 정보 과학회지 제 15권 제 9호, pp 10-19, 1997
- [3] 이동섭, 이양원, "뉴스 비디오의 방면 분할을 위한 키 프레임 추출 기법", 한국 정보과학회 학술발표논문집 Vol. 25, No.1 pp 613-615, 1998
- [4] M.H.Yang and David Kriegmap, "Detecting Face in Image: A survey" IEEE Trans. On Pattern Anal. & Machine Intelligence, Vol. 24, pp 34-58, 2002
- [5] Jianping Fan, David.K.Y.Yau, "Automatic Image Segmentation by Integrating Color-Edge Extraction and Seed Region Growing" IEEE Trans. On image Processing, Vol. 10, pp 1454-1466, 2001
- [6] V.Starovoitov and D.Samal, "Matching of Faces in Camera Image and Document Photographs," IEEE International Conference Acoustics: Speech and Signal Processing, Vol. IV, pp 2349-2352, June, 2000
- [7] K.Sobottka and I.Pitas, "Looking for Faces and Facial Features in Color Images", Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications, Russian Academy of Sciences, Vol.7, No.1 , 1997
- [8] C.W. Jung, J.S.Kim, S.G.Jahng, J.S.Chi, "Face Detection Algorithm Using Skin-Color Distribution and Facial Features in Color Images", ITC-CSAC2001, Tokushima Japan, July, 2001
- [9] Douglas Chai "Face Segmentation Using Skin-Color Map in Videophone Applications", IEEE Transactions on Circuits and System for Video Technology, Vol.9, No.4, pp 551-564, 1999