

# 내용기반영상검색을 위한 유용한 웹 영상의 자동 검출 방법

임경은\*, 김성영\*\*, 김민환\*

\* 부산대학교 컴퓨터공학과

\*\* 금오공과대학교 컴퓨터공학부

## Automatic Detection of Useful Web Images For CBIR System

KyungEun Lim\*, SungYoung Kim\*\*, MinHan Kim\*

\* Dept. of Computer Engineering, Pusan National University

\*\* School of Computer Engineering, Kumoh National Institute of Technology

### 요 약

본 논문에서는 내용기반영상검색을 위해 웹으로부터 수집된 영상들 중에서 검색 요구가 높은 유용한 영상들을 검출하는 방법을 제안한다. 검색 요구가 높은 영상들 이외의 불용 영상들은 배너, 화면 캡처와 같이 그래픽 성격이 강한 특징이 있다. 이에 본 논문에서는 영상에서 가장 빈번하게 나타나는 우세 칼라(dominant color)에 대한 정보와 영상 내부에 위치하는 수평 및 수직 성분을 추출하여 유용한 영상에 대한 검출 기준으로 사용한다. 또한 기존에 사용되는 인접한 픽셀간의 칼라 거리 및 영상의 가로 및 세로 비율을 활용한다. 이들 네 가지 특징을 순차적으로 조합하고 다수결 원칙을 적용하여 유용한 대상을 검출할 수 있도록 한다. 제안된 방법은 웹에서 직접 수집한 2047장의 영상에 적용하여 약 90%의 정확도를 얻었다.

### 1. 서론

웹 기반 영상 검색 서비스를 제공하기 위해서는 웹 에이전트(web agent)를 통한 웹 영상의 수집 과정이 선행된다. 수집된 영상들에 대해서는 영상 검색에 활용 가능한 특징들을 추출하여 영상 데이터베이스를 생성한다. 그런데 웹에는 방대한 영상 데이터들이 포함되어 있으며 이들 중에는 영상 검색의 대상이 되지 않는 불필요한 영상들도 상당수 존재한다. 예를 들어, 광고, 배너, 슬라이드 자료 및 화면 캡처된 영상들은 유용한 내용을 가진다고 보기 어려우며, 또한 이들 영상에 대해서는 영상 검색 요구가 거의 발생하지 않는다. 사용자의 영상 검색 요구는 주로 자연

경관이나 특정 대상을 촬영한 사진 영상에 한정된다. 영상 검색 서비스를 제공하기 위해서는 방대한 데이터에 대한 수집 및 가공이 요구되므로 영상 데이터베이스 구축에 많은 시간이 소요된다. 검색 요구가 낮은 영상들을 사전에 검출함으로써 영상 데이터베이스를 신속하고 효과적으로 구축할 수 있으며 또한 검색 성능을 향상시킬 수 있다.

이에 본 논문에서는 웹에서 수집된 영상에 대한 분류 과정을 통해 영상 검색에 유용한 영상들을 자동으로 추출하는 방법을 제안하였다. 내용 기반 영상 검색에서 검색 요구가 낮은 영상들의 특징을 분석하여 보면 대부분 컴퓨터를 사용하여 인위적으로 생성된 그래픽 요소를 포함하고 있다. 그래픽을 포함하는

영상들은 자연 영상에 비해 제한된 칼라 수를 가지며 특정한 칼라가 높은 빈도수를 갖는 경향이 있다. 또한 수평 및 수직 성분이 강한 직선을 포함하는 비율이 높다. 본 논문에서는 그래픽 영상의 이와 같은 특징을 사용하여 검색 요구가 낮은 영상들을 검출할 수 있도록 한다. 또한 웹에서 영상의 자동 수집 시에 작성한 키워드 테이블을 활용하여 인명사진 리스트를 작성함으로써 인명사진에 등록된 인물 영상은 최대한 영상 데이터베이스에 포함될 수 있도록 한다.

## 2. 관련연구

Athitsos와 Swain[1]은 사용된 칼라 수, 최대 빈도 칼라 등의 칼라 정보를 사용하여 사진 영상과 그래픽 영상을 구분하는 연구를 수행하였다. 이들은 구분 대상을 사진과 그래픽에 한정하고 사진과 그래픽이 혼재된 영상이나 손으로 직접 그린 그림을 스캔한 경우와 같이 모호한 영상은 대상에서 배제하였다. Prabhakar et al.[3]는 분류 대상을 비즈니스 그래픽과 자연 사진으로 구분하였는데, 실험 대상은 209장의 영상들로 구성된 소규모 DB로 한정하고 있다. 분류 기준은 칼라, 텍스처, 에지 정보를 함께 활용하였다. 그러나 이들이 제시한 기준은 사진을 기본으로 편집한 영상이나 사진과 그래픽 혹은 텍스트가 혼합되어 있는 영상에서는 성능을 발휘하기 어렵다.

통신 환경이 현저히 개선되고, 그래픽 기법이 발달함에 따라 기존과는 다른 새로운 형태의 영상들이 웹 상에 속출하고 있다. 따라서 새로운 영역에 적용할 수 있는 영상의 특징을 추출하는 것이 필요하다.

## 3. 유용한 영상 검출을 위한 특징 정보

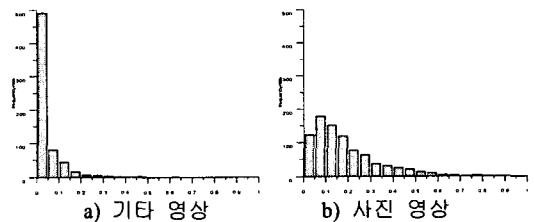
### 3.1 칼라 분포 특성

일반적으로 사진 영상은 광고, 배너 및 컴퓨터 그래픽 영상에 비해 풍부한 칼라 분포를 가지는 특징이 있다. 반면 그래픽 영상은 제한된 칼라 혹은 단일 칼라를 사용하여 넓은 면적이 채워지는 경우가 빈번하다. 따라서 영상에 포함된 우세 칼라(dominant color)의 분포 특성을 파악함으로써 사진 영상과 기타 영상들을 분류하는 것이 가능하다. 영상에서 우세 칼라가 넓은 영역을 갖는 경우에는 그래픽 영상일 가

능성이 크다. 우세 칼라를 선택하기 위해 본 논문에서는 우선 입력 영상에 대해 RGB 공간을 16×16×16 단계로 균일 양자화하여 4096개의 bin을 가지는 히스토그램을 생성하였다. 그런데 양자화된 칼라를 통해 우세 칼라의 분포 특성만을 고려하는 경우에는 사진 영상 중에서도 하늘, 사막 및 초원 영상과 같이 비교적 단순한 패턴으로 표현된 영상들은 기타 영상과 구분하기 어려운 경우가 발생한다. 이에 본 논문에서는 가장 빈번하게 발생하는 우세 칼라가 양자화 이전의 트루칼라 영상에서는 몇 개의 칼라로 구성되는지 조사하는 과정이 필요하다. 우세 칼라를 갖는 픽셀 및 구성 칼라의 개수의 상관 관계를 분류 특징으로 사용하기 위해 수식 (1)을 정의한다.

$$CN = \frac{NC_{DC}}{NP_{DC}} \quad \text{수식 (1)}$$

$NP_{DC}$ 는 영상에서 우세 칼라를 갖는 픽셀의 개수이고  $NC_{DC}$ 는 양자화 이전에 우세 칼라를 구성하는 칼라 수이다. 다양한 칼라 분포를 갖는 영상일수록  $CN$ 의 값은 증가하며 최대 '1'의 값을 갖는다. 사진 영상은 우세 칼라가 비교적 다양한 칼라로 구성(그림 1(b))되는 반면 기타 영상에서는 제한된 칼라로 구성(그림 1(a))되는 특징이 있다.



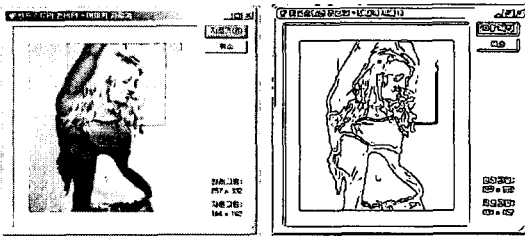
[그림 1] 우세 칼라에 대한 칼라 분포 특성

### 3.2 수평 및 수직 성분

그래픽 틀을 통해 사진 영상들을 변형하는 경우에는 두 장 이상의 사진 영상을 조합하거나 관심 있는 영역을 강조하기 위하여 배경을 조작하는 경우가 많다. 웹에서 무작위로 획득한 영상들의 구성을 살펴 보더라도 상당 수의 영상에서 인위적으로 배경이 조작 혹은 제거되었거나 액자 모양과 같은 테두리를

포함하고 있었다. 이런 영상들은 대부분 제품 판매를 위한 제품 사진이거나 화면 캡처된 경우로써 역시 검색의 요구 빈도는 매우 낮다.

배경이 제거되어 단색으로 처리된 경우는 칼라 분포 특성을 이용하여 쉽게 구분할 수 있지만 특정 객체의 외곽에 테두리 처리가 된 경우나 컴퓨터 화면 캡션 영상의 경우는 칼라 분포나 칼라 거리만을 이용해서는 구분되지 않는다. 그런데 이런 영상의 경우에는 일반 자연 영상에서는 거의 존재하지 않는 수평 및 수직 성분이 매우 강하게 나타나게 된다. 따라서 이와 같은 수평 및 수직 성분을 추출함으로써 인위적인 요소를 포함하는 영상들을 검출할 수 있다.



a) 화면캡처 영상      b) 에지 검출 결과  
[그림 2] 에지 검출을 통한 수평 및 수직 성분 추출

이에 본 논문에서는 Canny 연산자를 사용하여 에지를 검출하고 수평 및 수직 성분을 갖는 선분이 존재하는 지를 조사하였다. 이때 선분이 영상의 경계에 위치하는 경우와 영상 내부에 위치하는 경우를 별도로 구분하였다. 영상의 내부에 수평 및 수직 성분이 발견되면 이는 인위적인 조작에 의해 생성된 것으로 판단하여 해당 영상을 제거 대상으로 분류하도록 한다. 반면 영상의 경계에만 수평 및 수직 성분이 나타나는 경우에는 주로 인물 영상에서 단지 해당 인물을 강조하기 위한 테두리일 가능성이 높기 때문에 인명사전을 참고하여 선택적으로 제거하도록 하였다. 인명사전은 로봇을 통해 웹 영상을 수집할 당시 영상 주변의 텍스트를 분석하여 선정한 키워드 중 사용자의 선호도가 높은 인물에 대한 것을 등록한 테이블이다. 따라서 영상에 대한 키워드가 인명사전에 등록된 단어일 경우 유명인의 사진일 가능성이 높다고 판단한다. [그림 2]는 화면 캡처된 영상에서 수평 및 수직 성분을 갖는 선분들이 다량으로 검출

된 경우를 나타낸다. 총 17개의 수평 및 수직 성분의 선분들이 존재하는데 이중에 9개의 선분이 영상의 경계에 위치하고 8개는 영상 내부에 위치한다.

### 3.3 기타 특징 정보

#### ① 인접 픽셀간의 칼라 거리[1]

영상의 총 픽셀 수에서 인접 픽셀의 칼라가 자신과 동일한 픽셀의 개수를 뺀 값을  $f1$ 이라 한다. 그래픽 영상의 경우 넓은 영역이 단일 칼라로 표현되는 경향이 있으므로 낮은  $f1$ 을 갖는다. 인접 픽셀간의 칼라 거리가 임계치 보다 높은 픽셀의 개수를  $f2$ 라 한다. 사진 영상은 칼라가 부드럽게 변하는 반면 그래픽 영상은 칼라가 급격하게 변하는 경향이 있으므로 그래픽 영상에서  $f2$ 는 큰 값으로 나타난다 이에  $f=f2/f1$ 의 비율을 계산하여 분류 기준으로 사용한다.  $f$ 가 높을수록 그래픽 영상일 가능성이 커진다.

#### ② 영상의 가로 및 세로 길이 비율[1]

영상의 가로 및 세로 길이 중에서 큰 값을 작은 값으로 나눈 비율을 특징 정보로 사용한다. 사진 영상의 경우 인화지나 디지털 카메라의 해상도와 관련하여 그 비례가 일정한 범위 내에 존재하는 경향이 있지만 광고 영상이나 배너와 같이 가치 없는 영상들은 가로 및 세로의 비율이 매우 큰 값을 갖는다.

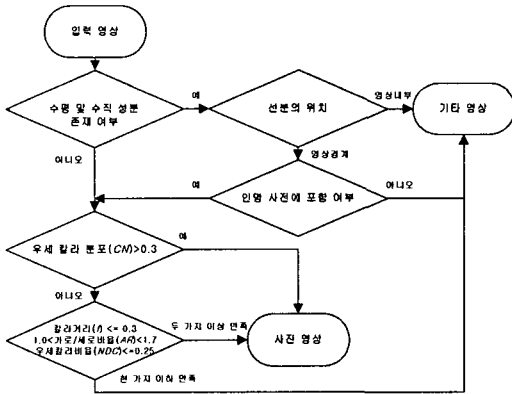
## 4. 구현 및 성능평가

### 4.1 분류방법

제안된 분류 기준들은 서로 상보적인 관계가 있는 것으로 어느 한가지만을 사용하여서는 좀을 성능을 내기 어렵다. 따라서 이들 기준들을 병합 적용하여 최종 분류 결과를 얻도록 한다. [그림 3]은 분류 시스템의 흐름도를 나타낸다.

입력된 영상에 대해 수평 및 수직 성분을 추출하여 영상 내부에 이들 성분이 존재하면 제거 대상인 기타 영상으로 분류한다. 영상 내부가 아니고 영상의 경계에 이들 성분들이 존재한다면 인명 사전을 참조하여 기타 영상으로의 분류 여부를 결정한다. 그래픽이 아닌 영상에 대해서는 다시 CN 값을 참조하여 0.3이상인 경우에는 사진 영상으로 최종 분류한다

CNO이 0.3미만인 경우에는 칼라 거리( $f$ ), 영상의 가로 및 세로 길이의 비율( $AR$ ) 및 영상 내에서 우세 칼라를 갖는 픽셀의 비율( $NDC$ )에 다수결 원칙(majority rule)을 적용하여 최종 분류한다.  $f$ 가 0.3이하,  $AR$ 이  $1.0 < AR < 1.7$ ,  $NDC$ 가 25%이하인 경우 각각 score1, score2, score3을 활성화하고 두 개 이상의 score가 활성화되면 사진 영상으로 분류한다.



[그림 3] 영상 분류 시스템의 흐름도

#### 4.2 성능평가

제안된 방법의 성능을 평가하기 위해 웹에서 무작위로 수집된 2047장의 영상 중에서 장축의 길이가 100 픽셀 미만인 영상을 제외한 1505장의 영상을 대상으로 실험을 수행하였다. 기타 영상의 검출 정확도는 84.41%(563/667)이었으며, 사진 영상의 검출 정확도는 93.67%(785/838)로 나타났다. [그림 4]는 기타 영상 및 사진 영상으로 분류된 결과를 나타낸다. 분류 결과를 살펴보면 기타 영상의 검출 정확도가 사진 영상에 비해 낮게 나타나고 있는데, 오류 영상 중 영화포스터나 잡지기사 등이 높은 비율을 차지하였다.

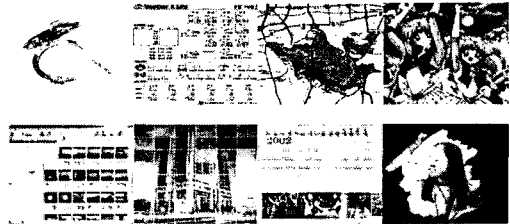
#### 5. 결론

본 논문에서는 웹에서 수집된 영상에 대해 영상 검색의 요구가 높은 영상들을 검출하여 제거할 수 있는 방법을 제안하였다. 검색 요구가 현저하게 낮은 불용 영상들은 배너, 화면 캡처와 같이 대부분 그래픽 영상으로써 그래픽 영상의 특징을 분석하여 분류 기준을 선정하였다. 제안된 방법을 적용함으로써 영

상 데이터베이스 구축 시간을 단축시킬 수 있으며 또한 검색 성능의 향상도 기대할 수 있다. 그런데 기타 영상의 검출 정확도가 사진 영상에 비해 상대적으로 저조하게 나타난다. 이를 해결하기 위해서는 의미 있는 영상에 대한 보다 구체적인 분류 기준을 제시하여 영상을 보다 세분화시켜 분류할 수 있는 방법에 대한 연구가 필요하다.



(a) 사진 영상



(b) 기타 영상

[그림 4] 영상 분류 결과

#### [참고문헌]

- [1] V. Athistos and M.J. Swain, "Distinguishing photographs and graphics on the world wide web," IEEE Workshop on Content-Based Access of Image and Video Libraries, Aug 1997, pp.10-17
- [2] C. Frankel, M. J. Swain, and V. Athistos, "WebSeer: An image search engine for World Wide Web," University of Chicago Department of Computer Science Technical Report Tr-96-14, Aug. 1996
- [3] S. Prabhakar, H. Cheng and J.C. Handley, "Picture-Graphics color image classification," Int'l Conference on Image Processing, 2002
- [4] R. Lienhart and A. Hartmann, "Classifying images on the web automatically," Journal of Electronic Imaging, Oct. 2002