

객체 추출 및 객체별 그룹핑을 이용한 영상검색 결과의 단계적 서비스 방안

박창민*, 김성영**, 김민환***

* 성심외국어대학 경영정보시스템전공

** 창원전문대학 멀티미디어과

*** 부산대학교 컴퓨터공학과

A Scheme for Progressive Service of Retrieved Images based on Object Extraction and Grouping

Chang-Min Park*, Sung-Young Kim**, MinHwan Kim***

* Dept. of Management Information, System Sung Sim Foreign College

** Dept. of Multimedia, Changwon College

*** Dept. of Computer Engineering, Pusan National University

E-mail : cmpark@sungsim.ac.kr, sykim@changwon-c.ac.kr, mhkim@pusan.ac.kr

요 약

본 논문에서는 키워드를 입력해 검색된 영상들을 유사한 특징을 갖는 소수의 그룹으로 그룹핑하고 각 그룹을 대표하는 대표영상을 추출하여 우선적으로 사용자에게 보여주고 필요에 따라 나머지 영상들을 단계적으로 서비스할 수 있는 방안을 제시한다. 영상 그룹핑을 위한 각 영상의 특징은 영상에 포함된 중심 객체를 사용하여 추출한다. 이를 위해 검색 키워드는 객체와 연관성이 있는 단어로 제한하여 영상을 검색하며 검색된 영상으로부터 중심 객체를 추출할 수 있는 객체 추출 방법을 활용하였다. 각 영상으로부터 추출된 중심 객체에 대한 특징 벡터는 칼라 분포를 이용한다. 영상 그룹핑은 칼라분포로 표현되는 특징공간에서의 밀집도를 조사하여 높은 밀도로 모여있는 영역별로 추출하여 동일한 그룹으로 분류하였다. 대표 영상은 분류된 그룹에서 가장 밀집도가 높은 영상으로 선택된다. 한편, 얼굴이 포함된 영상은 사전에 따로 분류하고 얼굴 크기 및 얼굴 수에 따라 영상을 그룹핑하여 각 그룹에 대한 대표 영상을 선정한다. 본 연구에서 제안한 방법은 사용자에게 모든 검색 결과를 일괄적으로 보여주는 것에 비해 보다 빠른 시간 내에 사용자가 원하는 영상을 편리하면서도 효과적으로 확인할 수 있는 방법을 제공해 줄 수 있을 것으로 기대한다.

1. 서론

영상 검색 기술은 텍스트 기반 검색과 내용 기반

검색으로 구분할 수 있는데, 기술력 및 웹의 구조적인 특성으로 인해 현재는 텍스트 기반의 영상 검색

이 상용 서비스되고 있다. 텍스트 기반 영상 검색은 사용자가 키워드를 입력하여 관련된 내용을 포함하는 영상을 검색하는 방식이다. 이를 위해서는 웹 상에 존재하는 영상들에 대한 분류 작업이 선행된다. 영상 분류 작업은 아직까지는 영상의 내용에 대한 정보보다는 HTML 문서의 내용을 분석하여 이루어진다[4,5]. HTML 문서에서는 영상이 포함된 태그 주변의 텍스트 정보를 활용하여 영상 내용에 대한 실마리를 어느 정도 얻을 수 있으며, 이들 정보를 활용하여 영상을 분류하게 된다. 이와 같이 영상 분류가 주변 텍스트 정보에 따라 이루어지므로 검색된 결과 영상에는 키워드와 관련성이 적은 영상들도 상당수 포함된다. 이로 인해, 사용자는 검색된 많은 영상으로부터 원하는 영상을 확인하기 위해 시간과 노력을 소비하게 된다. 그런데 입력된 키워드와 관련된 영상들을 검색한 후 화면에 출력할 때 유사한 내용을 담고 있는 영상들을 분류하여 몇 개의 그룹으로 나누고 각 그룹의 대표 영상만을 우선적으로 보여준다면 검색 결과의 확인 시간을 줄이면서 사용자의 편의성을 상당히 증대시킬 수 있을 것이다.

이에 본 연구에서는 키워드로 검색된 영상을 유사한 시각적 특성을 가진 객체를 공통적으로 포함하는 영상들로 그룹핑[8]하고, 각 그룹에서 대표적인 영상을 추출하여 우선적으로 사용자에게 제공하고 나머지 영상은 필요에 따라 단계적으로 서비스할 수 있는 방법을 개발하였다. 영상에 포함된 중심 객체를 추출하기 위해서는 중심 객체 추출 방법[7]을 활용하였다. 영상의 중심 객체는 영상의 중심 부근에서 비교적 많은 영역을 차지하는 객체로 정의된다.

영상에서 얼굴의 포함 여부와 얼굴의 개수 및 크기 정보는 영상 분류에 유용하게 사용될 수 있는 정보이므로 얼굴 영역 추출 방법[6]을 활용하여 얼굴 영상만을 사전에 구분하고 얼굴 크기 및 얼굴 수에 따라 영상을 분류하여 사용자에게 따로 서비스할 수 있는 방법도 개발하였다.

본 논문의 2장에서는 얼굴 영역 및 객체 영역 추출에 대해 설명하고, 3장에서는 영상의 그룹핑 방법에 대해 설명한다. 4장에서는 그룹별 대표 영상의 선택방법과 검색 영상의 단계적 서비스 방안

에 대해 설명하고, 5장에서는 실험 결과 그리고 6장에서는 결론에 대해 설명하고자 한다.

2. 얼굴 객체 영역 추출

영상 그룹핑을 통한 대표 영상을 추출하기 위해 얼굴 검출 과정이 선행된다. 키워드로 검색된 영상에서 얼굴의 포함 여부는 사용자 관심의 대상이 된다. 먼저 얼굴을 포함하고 있는 영상을 추출하고 얼굴 수와 크기에 따라 분류하여 대표영상을 추출 한다.

얼굴이 포함되지 않은 영상은 객체별 그룹핑이 가능하도록 각 영상의 객체 영역 추출 과정이 필요하다. 객체 추출은 영상 분할을 수행하고 인접 영역을 병합한 후 객체를 대표하는 핵심객체영역과 배경을 대표하는 핵심배경영역을 선택하고 각 분할 영역에 대해 이들 영역과의 유사도 비교를 수행함으로써 이루어진다. 각 영상으로부터 객체 영역이 추출되면 객체에 대한 특징 벡터를 생성하고 특징 공간에서 밀집도를 조사하여 객체별 그룹핑이 이루어지고 각 그룹을 대표하는 대표 영상이 선정되어 사용자에게 우선적으로 제공된다. 대표영상 추출 과정은 그림 1과 같다.

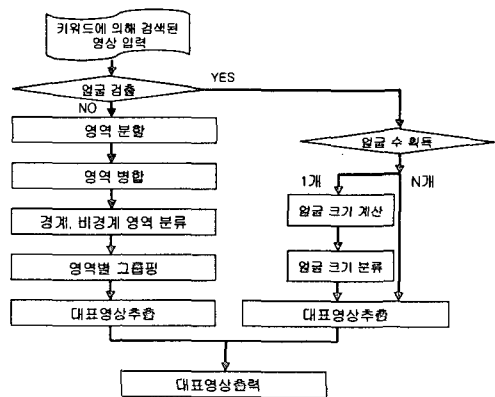
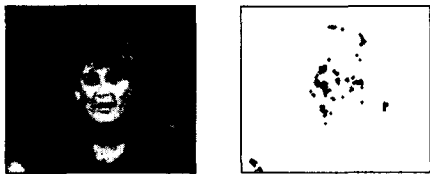


그림 1. 대표영상 추출 흐름도

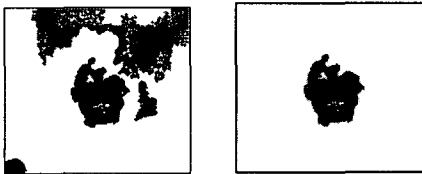
2.1 얼굴 영상 추출 및 분류

사람 얼굴 영역은 사용자 관심의 대상이 될 가능

성이 높은 영역으로써 이들을 분류 시켜줌으로써 검색에 유용하게 활용할 수 있다. 얼굴 영역은 얼굴 수와 크기가 중요한 정보가 되므로 이들에 따라 얼굴을 분류하였다. 우선 얼굴의 포함 유무를 결정하고 얼굴 포함 영상에 대해 얼굴 수를 파악하여 얼굴이 하나만 포함된 영상과 둘 이상 포함된 영상으로 분류한다. 얼굴이 하나만 포함된 영상은 다시 얼굴 크기에 따라 세가지 그룹으로 분류하였다. 얼굴이 전체 영상의 50% 이상을 차지하는 클로즈업된 얼굴을 포함하는 영상과 50-25%를 차지하는 영상, 그리고 25% 이하를 차지하는 영상으로 분류하였다. 얼굴 영역을 추출하기 위해서는 [6]에서 개발한 방법을 이용하였다. 얼굴 추출 단계는 그림 2와 같이 분석된 피부색 분포 특성을 이용하여 입력 영상으로부터 초기얼굴 후보 영역을 선택하고 초기 피부색 범위를 점진적으로 넓혀가면서, 초기 얼굴 후보 영역에 이웃하고 있는 픽셀들을 병합해 나가면서 피부색 영역을 점진적으로 확장한다. 최종적으로 얼굴 모양 정보 및 얼굴 기관(눈과 입)의 공간적인 위치 정보를 이용하여 얼굴 영역을 결정하게 된다.



(a) 원시영상 (b) 초기얼굴후보영역



(c) 피부색영역확장결과 (d) 최종추출된 얼굴영역

그림 2. 얼굴 영역 추출 과정

2.2 객체 추출

객체 추출을 위해서 우선 영상을 분할한다. 객체 추출을 위한 영역 분할은 많은 연산을 통한 정확

한 영역 경계 추출보다는 고속이면서 대표적인 영역 단위의 추출이 가능하면 유리하다. 객체별 그룹핑을 위해서는 많은 영상이 빠른 시간내에 처리되어야 되고 정확한 객체 영역의 경계 추출은 영상의 그룹핑 성능에 크게 영향을 미치지 못하기 때문이다. 분할된 영상으로부터 객체 영역을 추출하기 위해 [7]에서 제안한 방법을 사용하였다.

중심 객체를 추출하기 위해 [7]에서는 입력 영상에 대해 해상도를 줄여가며 영상 분할하여 계층적인 영상 집합을 생성하고 그 결과를 조합함으로써 영상 내부의 사소한 변화는 무시하며 상대적으로 중요한 변화가 발생하는 영역 단위의 추출이 가능하다. 저해상도에서 영상 분할된 결과에는 세부적인 변화 및 그림자 등이 어느 정도 무시되고 포괄적인 객체의 형태를 나타내며 영역이 분할되는 장점이 있다. 분할된 영역들은 영상 경계와의 교차정도에 따라 경계영역과 비경계영역으로 구분하여 정의된다. 비경계영역 및 경계 영역은 각각 객체를 구성하는 영역과 배경에 해당될 가능성이 높은 영역이다.

비경계영역 중에서 객체 영역을 대표하는 핵심 객체영역이 선택되고 경계영역 중에서 배경을 대표하는 핵심배경영역이 선택되어 최종 객체 추출에 이용된다. 최종 객체 영역은 각 비경계영역을 핵심객체영역 및 핵심배경영역과의 칼라 분포에 대한 유사도를 비교하여 핵심객체영역과 더욱 유사한 영역들의 합으로 결정된다.



그림 3. 객체 추출 결과

3. 객체별 영상 그룹핑

추출된 객체 영역은 영역 내부의 칼라분포 특징을 이용하여 유사한 칼라분포를 갖는 영역들을 동일 그룹으로 지정하여 전체 영상들이 소수의 그룹으로

분류되도록 한다. 최근 많이 사용되는 클러스터링 기법으로는 K-means, ISODATA 등이 있다. 이 방법들은 영역들이 그룹으로 할당될 때 할당되는 전체 데이터의 총 에러값을 최소화 시킬 수 있는 방향으로 분류된다. 이때 그룹의 초기 개수와 그룹의 병합 및 분할을 위한 임계값 등이 주어져야 된다. 이 값들은 전체적인 분류 성능을 좌우하는 중요한 요소지만 결정하기 어렵다는 문제가 있다.

본 논문에서는 [8]에서 제안된 방법을 이용하여 특징 공간(feature space)에서의 각 객체 영역의 밀집도를 파악하여 높은 밀도로 모여있는 영역들을 추출하여 동일한 그룹으로 분류할 수 있도록 한다. 특징 공간은 각 칼라 빈을 축으로 구성된다. 영상 그룹핑은 우선 밀집도가 높은 핵심영역을 중심으로 첫 번째 그룹이 선택한다. 밀집도는 객체 영역에 대해 나머지 영상의 객체 영역과의 히스토그램 인터섹션 값을 구하여 임계값 이상인 경우 이 값을 모두 더하여 계산된다. 첫번째 그룹에 포함되는 모든 영역들을 결정하기 위해 핵심영역을 중심으로 범위를 넓혀가며 주위 영역들을 포함시킨다. 이때, 포함된 영역들의 밀집도가 임계값을 넘지 않을 때까지 확장을 계속한다(그림 4).

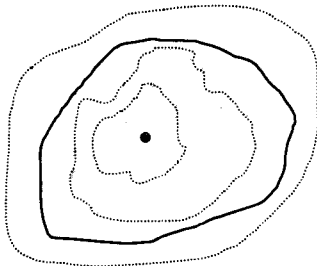


그림 4. 밀집도 기반 그룹핑

첫번째 그룹이 결정된 경우 다음 그룹의 중심을 결정하기 위하여 아직 그룹에 할당되지 않은 영역 중에서 가장 큰 밀집도를 가진 영역을 새로운 그룹의 중심으로 설정하고 첫번째 그룹핑과 동일한 방법으로 그룹의 크기를 결정한다. 또한 동일 영역이 두 개이상의 그룹에 중복 포함되는 경우가 있는데, 이들 중복 포함된 영역에 대해서는 히스토그램 인터섹

션[9]을 통해 소속 그룹 중에서 가장 높은 유사도를 갖는 그룹으로 재분류 하였다. 그런데 그룹을 형성하는 구성 벡터의 개수가 적은 경우에는 의미 있는 그룹으로 취급할 수가 없으므로 이런 그룹들은 기타 항목으로 분류된다.

4. 대표 영상 추출 및 단계적 서비스 방안

키워드를 입력하여 검색된 결과 영상에는 키워드와 관련이 적은 영상도 존재하지만 대체적으로 관련이 높은 영상들이 비교적 높은 비율로 존재한다. 또한 대부분의 경우 동일한 객체는 유사한 칼라 분포를 갖게 된다. 따라서 영상으로부터 추출된 객체를 기반으로 그룹핑을 수행하면 이들 객체를 포함하는 영상들이 비교적 높은 밀집도를 형성하게 될 것이다. 검색된 전체 영상에 대한 대표 영상은 각 그룹별로 하나씩 선택된다. 대표 영상은 밀집도가 가장 높아 각 그룹의 중심에 해당되는 영역으로 선택된다. 따라서 형성된 그룹의 개수만큼 대표영상이 생성된다. 각 그룹에서 선택된 대표 영상은 유사한 칼라 분포를 갖는 객체 영상을 대표하는 영상이므로 전체 검색 영상들을 대표하는 중요한 의미를 가진다고 할 수 있다.

현재 서비스되는 영상 검색 엔진은 검색된 모든 영상을 순차적으로 표시함으로써 사용자는 검색 결과를 확인하기 위해 많은 시간을 다시 소비하게 된다. 본 연구에서는 검색된 영상들로부터 선택된 대표 영상부터 사용자에게 출력하도록 한다. 사용자는 모든 영상을 확인하지 않더라도 검색된 결과 영상에 대한 대략적인 분석이 가능해질 수 있게 된다. 또한 대표 영상중에 사용자가 원하는 영상이 존재하지 않을 경우에는 각 그룹별로 세부 영상을 확인함으로써 영상 확인 시간을 단축시킬 수 있다. 세부 영상 중에도 원하는 영상이 존재하지 않는 경우에는 영상중의 하나를 선택하고 영상의 객체에 대한 칼라 분포 특징을 기준으로 재검색이 가능하도록 기능을 제공함으로써 검색 효율도 향상시킬 수 있을 것이다.

관련성이 적은 영상들을 기타로 분류함으로써 불필요한 영상을 확인하기 위해 소요되는 시간도 단축

할 수 있다. 이와 같이 검색된 영상들을 단계적으로 제공함으로써 다양한 검색 효율에 대한 이득을 얻을 수 있다.

5. 실험 및 토의

본 논문에서 제안한 방법은 Windows2000 환경의 Pentium IV PC에서 Visual C++ 6.0으로 구현하였다. 실험을 위하여 '호랑이'를 키워드로 사용하여 웹에서 검색된 94개의 영상을 입력 영상으로 사용하였다. 검색된 영상에는 키워드와 정합 되는 호랑이가 포함되어 있는 영상과 호랑이와는 무관한 영상들이 함께 존재하였다. 검색된 모든 영상들로부터 객체 영역을 추출하여 객체별로 칼라분포 특징들을 우선 DB에 저장한다. 그룹핑을 수행하여 분류된 그룹의 세부 영상과 각 그룹의 대표 영상은 그림 5와 같다. 3개의 그룹으로 형성된 영상을 보면 비교적 유사성을 가진 영역들을 가진 영상들이 각 그룹에 포함되는 것을 확인할 수 있었다. 임계값을 0.6으로 하여 나온 결과에서 첫번째 그룹과 두 번째 그룹을 보면 호랑이와 무관한 영상들도 함께 포함되어 있었다. 이들 영상들은 객체 영역이 잘못 추출되었거나 일반 호랑이와는 상이한 칼라분포를 갖는 영상들이다. 하지만 호랑이와 무관한 영상이라고 하더라도 영역별로 호랑이의 대표적인 칼라분포와 유사한 영역이 존재하여 호랑이 영상과 동일한 그룹으로 분류된 것이다. 이것은 객체 영역 추출의 어려움과 추출된 객체 영역의 칼라 분포 특징만을 사용하기 때문에 발생하는 문제점으로 판단된다. 따라서 영역에 대한 칼라 정보 외에 다른 지식 정보를 선택적으로 적용할 수 있는 방법이 절실히 요구된다.



(a) 첫번째 그룹 대표영상 및 구성 영상들



(b) 두 번째 그룹 대표영상 및 구성 영상들



(c) 세 번째 그룹 대표영상 및 구성 영상들

그림 5. '호랑이' 키워드 검색영상에 대한 그룹핑결과

얼굴을 키워드로 사용하여 검색된 영상들은 따로 분류하여 그룹을 형성하였다. 검색된 얼굴 영상들은 얼굴이 하나만 포함된 영상들과 둘 이상을 포함하는 영상들로 분류하였다. 얼굴이 하나만 포함된 영상은 전체영상에서 얼굴이 차지하는 비율에 따라 다시 세 가지 그룹(50%이상, 50~25%, 25%이하)으로 분류하였다. 그림 6은 얼굴분류에 대한 결과를 보여주고 있다. 분류된 영상들은 얼굴 영상 검색시에 유용하게 사용될 수 있을 것으로 예상된다.



(a) 얼굴크기가 50%이상을 차지하는 영상



(b) 얼굴크기가 25-50%를 차지하는 영상



(c) 얼굴크기가 25%이하를 차지하는 영상



(d) 얼굴이 두개 이상이 포함된 영상

그림 6. 얼굴 영상에 대한 분류 결과

6. 결론 및 향후 연구

본 연구에서는 키워드에 의해 검색된 객체가 포함되어 있는 칼라 영상들의 집합으로부터 유사한 시각적 특징을 가진 객체를 공통적으로 포함하는 영상들을 추출하고, 전체 영상들의 집합을 소수의 그룹으로 묶어 각 그룹의 대표영상을 추출할 수 있는 방법을 제시하였다. 추출된 대표 영상을 활용하여 키워드 검색된 모든 영상 중에서 우선적으로 대표 영상만을 디스플레이하고 사용자가 원할 경우 각 대표 영상과 동일한 그룹을 구성하는 나머지 영상들을 단계적으로 디스플레이 할 수 있도록 하였다. 본 연구에서 제안한 방법은 그룹을 이루지 않은 상태에서의 결과에 비해 보다 빠른 시간 내에 정확한 검색 결과를 얻을 수 있는 효율적인 방법이 될 수 있을 것으로 예상된다.

그룹핑된 각 영역에 대한 대표적인 특징값을 추출하여 검색 키워드와 연계 시켜 둠으로써 차후 키워드 검색에서 좀더 정확한 검색 결과를 빠르게 얻을 수 있을 것이다. 또한 객체별 저장된 특징값을 내용기반검색에 활용함으로써 단지 영역의 색상이나 텍스처 정보만을 이용하는 것이 아니라 영상에 포함된 의미 정보를 일부 포함시켜 영상을 검색할 수 있을 것으로 기대된다.

현재는 객체를 포함하는 영상들을 대상으로 대표 영상을 추출하지만 향후 객체가 포함되지 않은 자연 영상(석양, 바다 영상 등)에 대한 구분 및 분류 방법을 연구할 것이다.

[참고문헌]

- [1] W.Y. Ma, NETRA: A Toolbox for Navigating Large Image Databases, Ph.D. Dissertation, Dept. of Electrical and Computer Engineering, University of California at Santa Barbara, 1997
- [2] C. Carson, M. Thomas, S. Belongie, J. M. Hellerstein, and J. Malik, Blobworld: A System for region-based image indexing and retrieval, the Third Int. Conf. On Visual Information Systems, 1999
- [3] Gonzalez, Wood, Digital Image Processing, Addison Wesley, 1992
- [4] J. R. Smith and S. F. Chang, Searching for Images and Videos on the World-Wide-Web, Columbia University CTR Technical Report #459-96-25, August, 1996
- [5] C. Frankel, M. J. Swain and V. Athitsos, WebSeer: An Image Search Engine for the World Wide Web, Univ. of Chicago Technical Report TR-96-14, 1996
- [6] 문대성, 김성영, 김민환, 피부색 영역 확장에 의한 얼굴 영역 추출 방법, 한국 멀티미디어학회, 2000추계학술발표논문집, 2000
- [7] 김성영, 권규복, 박창민, 김민환, 칼라 영상에서의 객체영역 추출에 관한 연구, 한국 멀티미디어학회, 2001추계학술발표논문집, pp. 121-125, 2001
- [8] 박창민, 김성영, 권규복, 김민환, 객체가 있는 칼라 영상에 대한 객체별 그룹핑 방법에 대한 연구, 한국 멀티미디어학회, 2001춘계학술발표논문집, pp. 128-132, 2001
- [9] M. Swain and D. Ballard, Color Indexing, Int'l Journal of Computer Vision, Vol 7, No 1, pp. 11-32, 1991