

객체가 있는 칼라 영상에 대한 객체별 그룹핑 방법에 대한 연구

김성영*, 박창민**, 권규복***, 김민환***

* 창원전문대학 멀티미디어과

** 성심외국어대학 경영정보시스템전공

*** 부산대학교 컴퓨터공학과

A Study on Color Image Grouping Method based on Color Objects

Sung-Young Kim*, Chang-Min Park**, Kyu-Bok Kwon*** and Min-Hwan Kim***

* Dept. of Multimedia, Changwon College

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

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

요 약

본 논문에서는 객체가 포함되어 있는 칼라 영상들의 집합으로부터 유사한 시각적 특징을 가진 객체를 공통적으로 포함하는 영상들을 추출하여, 전체 영상들을 소수의 그룹으로 묶을 수 있는 방법을 제안한다. 먼저 영상에 포함된 객체는 주로 영상의 내부 위치에 나타난다는 사실에 착안하여, 영상의 경계와 접하지 않는 비경계 영역들을 병합하여 객체가 포함된 영역으로 간주한다. 그러나, 병합된 내부 영역에는 객체가 아닌 영역들도 포함하고 있을 수 있으므로, 칼라 히스토그램을 이용하여 내부 영역에 대한 특징으로 이용한다. 이러한 내부 영역의 칼라 특징들이 분포하는 형상을 분석하여 밀집도가 높은 클러스터를 그룹으로 추출한다. 이때, 밀집도는 일반적으로 사용하는 공간적인 분포 대신에 히스토그램 인터섹션에 의한 유사도를 이용하여 정의한다. 즉, 서로 유사도가 높은 것들이 집중되어 분포되어 있는 경우에 밀집도가 높은 클러스터로 간주하여 추출하는 방법을 사용한다. 클러스터의 형상 및 개수를 자동적으로 결정할 수 있는 방법도 제안한다. 실험에 의해, 추출한 클러스터의 칼라 영상들이 동일한 객체를 포함하고 있음을 알 수 있었으나, 향후 보다 안정화된 방법의 개발이 필요하다. 아울러, 클러스터별로 객체의 의미를 부여할 수 있는 방법론의 개발도 필요함을 알 수 있었다.

1. 서론

최근 영상 데이터의 급속한 팽창으로 인하여 영상 데이터를 효율적으로 관리하고 활용할 수 있는

방법이 부각되고 있다. 근래에는 각 영상에 대해 주석(annotation)을 달아 관리하는 방법의 문제점을 극복하기 위해, 영상의 시각적 특징(텍스처, 칼라 분포,

모양 정보 등)을 이용하여 내용기반으로 검색하는 방법들이 활발히 연구되고 있다. 그러나, 이러한 방법들은 각 영상을 개별적으로 관리하는 접근 방법으로서, 전체 영상을 그룹핑하여 분류 관리하며 계층적으로 서비스하는 접근 방법보다 효율적이지 못하다. 한편, 인터넷에서의 키워드 기반 검색을 통해 수집한 영상들을 분류없이 순차적으로 서비스하는 현재의 방법들보다 유형별로 분류하여 사용자가 원하는 것을 단계적으로 서비스해 주는 방법이 보다 효율적이다. 이러한 접근 방법은 최근의 텍스트 기반 검색 엔진에서도 주요 관심사가 되고 있다.

영상은 내용에 따라 동물 혹은 인물 등과 같은 특정 피사체 위주로 촬영되어 내부에 뚜렷한 객체가 포함되어 있는 경우와 석양이나 바다 풍경과 같이 뚜렷한 객체가 존재하지 않는 경우로 크게 나눌 수 있다. 객체가 포함되어 있는 영상에서의 사용자의 주된 관심 영역은 주로 객체 영역에 한정되며, 배경은 큰 관심의 대상이 되지 못한다. 따라서, 배경 영역을 제외한 객체 영역만을 추출하여 효과적으로 분류할 수 있는 방법론이 필요하다. 반면에, 뚜렷한 객체가 포함되어 있지 않은 영상에서는 영상의 전체적인 분위기나 구도에 따라 분류의 기준이 선정될 수 있으므로, 영상의 전체 영역을 분류에 이용하여야 한다.

본 논문에서는, 뚜렷한 객체가 포함되어 있는 영상을 대상으로 영상 분할을 통해 객체 단위로 영역을 추출하고, 추출된 영역의 시각적 특징을 이용하여 유사한 특징을 갖는 영상들을 그룹핑하는 방법을 제안한다(그림 1).

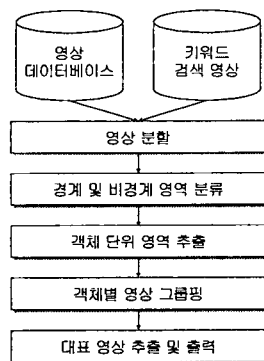


그림 1. 제안하는 방법의 작업 흐름도

논문의 2장에서는 영역 분할, 구분, 병합에 의해 객체를 포함하는 영역을 추출하는 방법에 대해 설명하고, 3장에서는 객체별 영상 그룹핑 방법에 대해 다룬다. 이어서, 4장에서는 실험 결과를 보이며, 5장에서 결론 및 향후 연구 방향에 대해 설명한다.

2. 객체를 포함하는 영역의 추출 방법

여기에서는 영상 분할 및 영역 병합에 의해 객체를 포함하는 영역을 추출하는 방법을 설명한다.

2.1. 영상 분할

일반적인 영상 분할에서는 픽셀 단위의 정교한 영역 경계 추출이 요구되지만, 영상 검색 및 그룹핑을 위한 영상 분할에서는 영역 경계는 다소 부정확하더라도 특징적인 영역들을 고속으로 추출하는 것을 목적으로 한다. 이때, 텍스처에 의해 영역이 너무 세분화되는 것은 바람직하지 못하다.

텍스처에 무관하게 영역 분할을 하는 방법으로 [5,8]에서 제안된 것을 이용할 수 있다. [5]에서는, 칼라 히스토그램에 기반한 새로운 영역 유사도 함수를 정의하고, 이 유사도를 이용하여 영상 분할하는 방법을 제안하고 있다. 이 방법은 고속의 영역 분할이 가능하나 영역의 개수를 한정하여 분할해야 하는 문제점이 있다. 내용기반 영상검색 시스템인 Netra에서 사용되고 있는 방법[8]에서는 색상과 텍스처 분포 유사도를 반영하는 J 값을 정의하여 영역을 분할하는데 사용하고 있다. 이 방법은 다소 처리 속도가 떨어지나 좋은 영역 분할 성능을 보여준다. 이에, 본 연구에서는, [8]에서의 방법을 이용하여 영역 분할을 수행하였다.

2.2. 영역 구분 및 병합

영상에 포함된 객체 영역은 단일 색상이나 텍스처로 구성되기 보다는 여러 가지 색상이나 텍스처로 구성되는 경우가 일반적이다. 따라서, 많은 경우에 영상에서의 한 객체가 여러 영역으로 분할되어 나타나게 된다. 그러나, 이러한 영역들을 병합하여 의미 있는 객체 영역을 추출하는 것은 매우 어렵다. 인간 처럼 객체 영역을 잘 구별하기 위해서는, 많은 사전

지식(pre-knowledge)을 지능적으로 처리할 수 있는 방법론이 필요하다. 그러나, 현재 기술로는 아직 적합한 방법을 찾기가 어렵다. 이에, 본 연구에서는 영상에서 관심있는 객체가 주로 중앙 내부에 나타난다는 사실에 착안하여, 영상 경계에 접하지 않는 영역들을 병합하여 객체 영역으로 가정하는 방법론을 사용한다.

내부 영역을 추출하기 위해서, 먼저 분할된 영역들을 대상으로 수식 (1)을 사용하여 경계 영역(boundary region, BR)과 비경계 영역(non-boundary region, NBR)으로 구분한다. 즉, 영역의 총 경계 길이(L_T)에 비해 영상의 경계에 해당하는 경계 길이(L_B)의 비율이 높은 영역을 BR로 분류하고, 나머지를 NBR로 분류한다.

$$R = \frac{L_B}{L_T} \text{ ----- (1)}$$

다음에는, NBR들 중에서 서로 인접해 있는 것끼리 병합하여 내부 영역을 형성한다. 이러한 내부 영역에는 객체와 관련이 없는 NBR 즉, 배경 영역도 포함될 수 있다. 또한, 병합된 내부 영역 중에는 배경 영역들로만 구성된 것도 발생할 수 있다. 따라서, 병합에 의한 내부 영역을 객체 영역이라고 판정할 수는 없다. 그러나, 어떤 병합된 내부 영역에는 객체의 특징(칼라, 텍스처 등)이 많이 반영되어 있다고 할 수는 있다.

한편, 한 영상에서 여러 개의 내부 영역이 추출될 수 있다. 이러한 현상은 작은 내부 영역을 제거함으로써 다소 완화할 수 있다. 즉, 어느 정도 큰 내부 영역만을 객체 영역으로 가정한다(그림 2).



그림 2. 병합에 의해 추출된 내부 영역

3. 객체별 영상 그룹핑

병합에 의해 형성한 내부 영역에는 객체 영역뿐만 아니라 배경 영역도 포함될 수 있다. 또한, 배경 영역만으로 구성된 내부 영역도 존재할 수 있다. 이와 같은 내부 영역들을 이용하여 동일한 객체가 있는 칼라 영상들을 그룹핑하기 위해 활용할 수 있는 특징으로는 칼라분포 정보와 텍스처 정보가 있다. 내부 영역의 모양(shape) 정보는 활용할 수가 없다. 본 연구에서는 칼라분포 특징만을 칼라 히스토그램으로 표현하여 영상을 객체별로 그룹핑하는데 활용한다.

유사한 칼라분포를 갖는 내부 영역들을 그룹핑하기 위해서 K-means, ISODATA 등과 같은 기존의 클러스터링 기법을 활용할 수 있다. K-means 방법에서는 K개의 클러스터를 형성하되, 각 클러스터로 분류되는 특징벡터와 그 클러스터의 대표 특징벡터와의 에러들을 총합하여 최저가 되도록 하는 접근 방법을 사용한다. 특징벡터의 분류 방법으로는 최소거리 분류 기법(minimum distance classifier)을 사용하고, 에러로는 특징공간에서의 거리 정보를 사용한다. ISODATA 방법은 K-means 방법에서의 클러스터 개수 설정의 어려움을 극복하기 위한 것으로서, K-means 방법과 유사하게 에러를 정의하고 최소화하는 접근 방법을 사용한다.

그러나, 이러한 거리 기반 에러의 정의와 최소거리 분류 기법을 사용하는 클러스터링 방법으로는 칼라분포 정보를 기반으로 칼라 영상을 그룹핑할 수가 없다. 먼저, 거리 기반의 에러는 특징공간에서의 공간적인 분포 정보를 고려하여 정의된 것으로서, 각 특징변수가 개체간의 상이성을 잘 표현할 수 있는 것일 때 의미를 갖는다. 칼라 히스토그램과 같이 전체가 하나의 정보 즉, 칼라분포 정보를 나타내는 경우에는 효용성이 매우 떨어진다. 따라서, 내용기반 영상검색 분야에서는 거리 정보 대신에 칼라 히스토그램 인터섹션 방법을 사용하여 유사도를 기반으로 분류하는 방법을 취한다. 한편, 최소거리 분류 기법은 추출할 클러스터의 형상 정보를 제대로 반영하지 못하는 문제점이 있다. 예를 들어, 그림 3과 같은 형태로 클러스터를 형성하는 특징벡터들이 인접하여 분포하고 있는 경우에, 바람직한 클러스터 경계는 (a)

로 설정되는 것이지만 클러스터 간의 중심위치에서 중간 지점에 클러스터 경계가 설정되므로 B 클러스터에 포함되어야 할 특징벡터가 A 클러스터에 포함되는 문제가 발생된다.

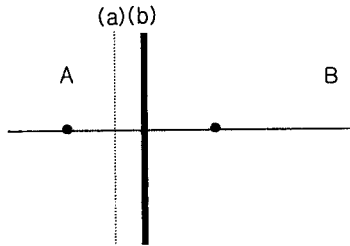


그림 3. 클러스터 경계 설정 문제

이에, 본 논문에서는 먼저 내부 영역에 대한 칼라 히스토그램을 특징벡터로 간주하며, 히스토그램 인터섹션 기법을 특징벡터간의 유사도를 측정하는 방법으로 사용한다. 각 내부 영역에 대한 특징벡터의 밀집도는 유사도가 임계값(T_i) 이상인 특징벡터들과의 유사도 값을 모두 더한 s 값으로 계산한다. 다음에는, 이러한 밀집도가 높은 특징벡터를 중심으로 클러스터를 구함으로써, 클러스터의 형상 정보가 잘 반영된 클러스터링 방법을 사용한다.

먼저 s 값이 큰 특징벡터일수록 밀집도가 높은 것이므로, 가장 s 값이 큰 특징벡터를 중심으로 첫 번째 클러스터를 형성한다. 이 클러스터에 포함되는 초기 특징벡터는 s 값을 구하는데 기여한 것들로 정한다. 다음에는, 유사도에 대한 임계값(T_i)을 낮추어가며 클러스터에 속하는 특징벡터들을 병합해 나간다. 새로 병합되는 특징벡터들의 유사도 합이 현저히 감소하면 클러스터의 확장을 멈추고, 그 때까지의 특징벡터들을 묶어 하나의 클러스터를 형성한다. 두 번째 클러스터의 중심은 아직 클러스터에 할당되지 않은 특징벡터 중에서 가장 큰 s 값을 갖는 것으로 결정하고, 첫번째 클러스터와 동일한 방법으로 클러스터의 크기를 결정한다. 모든 특징벡터가 어떤 한 클러스터에 할당될 때까지 새로운 클러스터를 생성하는 작업을 반복한다. 이때, s 가 일정값 이하로 낮은 값을 갖는 특징벡터는 밀집도가 낮아 클러스터를 형성하기

힘든 것이므로, 사전에 “기타 클러스터”로 분류하여 클러스터링 작업에서 배제한다. 마지막 작업으로서, 두 개 이상의 클러스터에 중복 포함되는 특징벡터는 유사도가 가장 높은 클러스터로 최종 할당하는 과정을 수행한다.

4. 실험 및 토의

본 논문에서 제안한 방법은 Windows98 환경의 Pentium III PC에서 Visual C++ 6.0으로 구현하였다. 실험을 위하여 먼저 ‘호랑이’를 키워드로 사용하여 현재 서비스되고 있는 영상 검색 엔진에서 검색된 60개의 영상을 입력으로 하였다. 검색된 영상에는 키워드와 정합되는 호랑이가 포함되어 있는 영상과 호랑이와는 무관한 영상들이 거의 유사한 비율로 함께 존재한다.

검색된 모든 영상들에 대한 분할 및 영역 병합을 통하여 생성된 각 내부 영역의 칼라분포 특징을 우선 DB에 저장하여 그룹핑이 이루어지도록 하였다. 표 1에는 임계값(T_i)을 변화시켜 가면서 객체별 그룹핑된 결과를 나타내었다.

표 1. 호랑이 영상 그룹핑 결과

| 임계값 (T_i) | 클러스터 개수 | 클러스터 내의 영상 분포율 | | 기타 클러스터 | |
|---------------|---------|--------------------------------------|-------------|---------|------|
| | | 호랑이 | 기타영상 | 호랑이 | 기타영상 |
| 0.2 | 1 | 39개 94.4% 백호 29.5% 황호 70.5% | 17개 5.6% | - | - |
| 0.3 | 2 | 34개 82.9% 백호 26.5% 황호 73.5% | 7개 17.1% | 7개 | 10개 |
| | | 1개 2.5% 백호 100% 황호 0% | 3개 7.5% | | |
| 0.4 | 2 | 31개 77.5% 백호 25.8% 황호 74.2% | 9개 22.5% | 7개 | 12개 |
| | | 3개 100% 백호 0% 황호 100% | 0개 0% | | |

먼저 기타 클러스터로 분류된 영상에는 호랑이 영상보다는 호랑이 이외의 영상이 많은 것을 볼 수 있다. 물론 호랑이 영상도 비교적 많이 포함되고 있는데, 이들 영상들은 객체 영역이 잘못 추출되었거나

일반 호랑이와는 상이한 칼라분포를 갖는 호랑이 영상들이다. 한편, 추출된 클러스터에는 호랑이 영상들이 높은 비율로 포함되어 있거나 혹은 호랑이 이외의 기타 영상들이 높은 비율로 포함됨으로써 영상에 대한 그룹핑이 적절히 수행되었음을 알 수 있다. 임계값을 0.3으로 하여 구한 결과에서, 두 번째 클러스터에는 기타 영상이 높은 비율로 포함되어 있는데, 이 클러스터의 구성 영상수가 적기 때문에 별로 문제가 되지 않으며 오히려 호랑이 영상과 기타 영상을 분리하는데 도움이 된다.

다음으로는 호랑이, 곰, 기린, 원숭이 영상이 섞여 있는 영상들의 집합을 대상으로 실험하였다. 그런데, 이들 영상들은 호랑이와 기린이 대체적으로 유사한 칼라 분포 특징을 가지고 있고, 원숭이의 경우에는 칼라 분포의 변화가 심하여 기대와는 차이가 나는 결과가 나타났다. 이것은 객체 영역 추출의 어려움과 추출된 객체 영역의 칼라 분포 특징만을 사용하기 때문에 발생하는 문제점으로 판단된다. 향후, 이에 대한 보완이 필요하다.

5. 결론 및 향후 연구

본 논문에서는 객체가 포함되어 있는 칼라 영상들의 집합으로부터 유사한 시각적 특징을 가진 객체를 공통적으로 포함하는 영상들을 추출하여 전체 영상들의 집합을 소수의 그룹으로 묶어 줄 수 있는 방법을 제안하였다. 제안한 방법에 대한 실험을 통해, 칼라분포가 어느 정도 차이가 나는 경우에는 그룹핑 효과가 있음을 확인하였다. 그러나, 향후에 객체 단위의 영역 추출에 대한 방법의 보완과 텍스처 정보 등과 같은 보다 많은 객체의 시각적 특징을 반영할 수 있는 방법론의 보완이 필요함을 확인할 수 있었다.

앞으로, 특정 객체가 포함되지 않은 자연 영상(석양, 바다 영상 등)에 대한 그룹핑 방법에 대해서도 연구할 것이다. 또한, 객체에 대한 사전 지식을 최대한 반영하여 보다 지능적인 그룹핑이 가능하도록 하는 방법론에 대해서도 연구할 것이다.

[참고 문헌]

- [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
- [5] 유윤주, 길연희, 김성영, 박창민, 김민환, "내용기반검색을 위한 자연영상의 칼라영역분할 방법", 한국 멀티미디어학회 '99추계학술발표논문집, pp. 724-729, 1999
- [6] 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
- [7] 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
- [8] Y. Deng, B. S. Manjunath, and H. Shin, "Color Image Segmentation", IEEE Conference on Computer Vision and Pattern Recognition, pp. 446-451, 1999
- [9] V. Ronda, M. H. Er., W. Ser, "Face Detection, Tracking and Recognition - A Study", The Fifth International Conf. on Control, Automation, Robotics, and Vision, Sinapore, 9-11 December, pp. 50-55, 1998
- [10] J. Yang, W. Lu, A. Waibel, "Skin-Color Modeling and Adaptation", CMU-CS-97-146, 1997
- [11] M. Swain and D. Ballard, "Color Indexing", Int'l Journal of Computer Vision, Vol 7, No. 1, pp. 11-32, 1991