내용기반 영상정보 검색기술에 관한 이론적 고찰

A Study on Content-based Image Information Retrieval Technique

노 진 구 (Jin-Goo Noh)*

〈목 차〉

Ⅰ. 서론
Ⅱ. 영상정보 검색기술의 연구현황
1. 내용기반 이미지 정보검색 연구 현황
2. 내용기반 비디오 검색연구 현황
3. 웹 이미지 검색연구 현황
4. 내용기반 영상정보 검색을 위한 질의
Ⅲ. 내용기반 영상정보 검색을 위한 인덱싱 기술
1. 내용기반 이미지 정보의 색인기법
2. 내용기반 검색을 위한 색인구조의 설계 요건
Ⅳ. 내용기반 영상정보 검색시스템 현황
1. 내용기반 이미지 검색시스템
2. 내용기반 동영상 검색시스템
Ⅴ. 결론 및 제언

초 록

최근의 디지털 도서관의 출현과 급속한 발전은 멀티미디어 데이터의 효과적인 탐색을 위한 수단의 필요성을 증가시키고 있다. 이에 따라 기존의 텍스트 기반의 정보검색에 관한 연구자들의 관심은 이미지 기반의 정보검색으로 모아지고 있다. 이에 본 연구는 최근까지의 멀티미디어 영상정보의 검색에 관한 여러 연구들 중 내용기반의 영상정보 검색방법에 관한 연구들을 살펴보고, 이들 연구에서 채용하고 있는 영상정보에 대한 다양한 특징정보의 추출과 인덱싱 기술에 대해 고찰해 보고자 한다. 아울러 영상정보의 효과적인 검색을 위해 채용되고 있는 색인구조의 설계 요건에 대해서도 살펴본다. 마지막으로 이러한 내용기반 영상정보 검색기술을 적용한 다양한 실제 시스템을 소개하고자 한다. 특히 디지털 도서관의 급속한 발전은 정보서비스에 이러한 영상정보 검색시스템의 적극적인 활용을 필요로 한다.

Abstract

The growth of digital image and video archives is increasing the need for tools that efficiently search through large amount of visual data. Retrieval of visual data is important issue in multimedia database. We are using contented-based visual data retrieval method for efficient retrieval of visual data. In this paper, we introduced fundamental techniques using characteristic values of image data and indexing techniques required for content-based visual retrieval. In addition, we introduced content -based visual retrieval system for use of digital library.

* 경북대학교 문헌정보학과 강사

Ⅰ. 서 론

최근 컴퓨터와 통신기술의 급속한 발달은 멀티미디어 정보서비스에 대한 요구를 증가시키고 있다. 또한 이미지 처리기술과 이미지 저장기술의 발전은 멀티미디어 데이터베이스에 관련된 기술 중에서도 이미지 데이터베이스 기술을 멀티미디어 정보서비스의 핵심으로 인식시키는데 기여하고 있다. 이미지 데이터베이스 기술은 시각화된 다양한 정보를 제공하는 정보서비스의 하나로서 디지털 도서관을 비롯하여 의료정보시스템, 지리정보시스템, 예술 화랑과 박물관 관리, 실내 디자인, 일기예보, 섬유 및 의상 디자인, 상표 데이터베이스 관리, 법률 및 범죄조사 시스템, VOD 등 여러 응용분야에 활용되고 있다. 특히, 디지털 도서관은 컴퓨팅과 커뮤니케이션 기술의 진보로 점차 그 성장속도를 가속화시키고 있다..

최근의 디지털 도서관의 출현과 급속한 발전은 멀티미디어 데이터의 효과적인 탐색을 위한 수단의 필요성을 증가시키고 있다. 이에 따라 기존의 텍스트 기반의 정보검색에 관한 연구자들의 관심은 이미지 기반의 정보검색으로 모아지고 있다. 또한 최근의 한 조사에 따르면 향후 가장 필요한 온라인 데이터베이스의 형식은 문자 데이터베이스보다 화상(그래프, 이미지, 동영상) 데이터베이스를 선호하고 있는 것으로 조사되고 있다.[1] 이러한 조사결과에서 보듯이 이제 방대한 이미지 데이터베이스의 지속적인 생산과 이러한 이미지 데이터의 색인과 검색은 중요한 문제로 대두되고 있다.

일반적으로 멀티미디어 정보검색 방법은 세 가지로 분류될 수 있다.

첫 번째는 원시적인 관찰에 의한 검색으로, 이 방법은 가장 초보적이고 원시적인 방법으로 처음부터 끝까지 관찰에 의해서 검색하는 방법으로 정보의 검색은 순차적인 검색(재생, 빨리감기, 되감기)만을 제공하고 있어 대용량의 멀티미디어 정보를 검색하기 위해서는 많은 시간을 소비해야 한다.[2]

두 번째는 텍스트 기반(text-based) 정보검색 방법으로, 모든 멀티미디어 검색 대상에 사람이 직접 색인을 부여하고 탐색자는 그 색인어에 부합되는 탐색어를 사용하여 원하는 정보를 검색하는 방법이다. 이 방법은 비정형의 멀티미디어 데이터에 대해서 제한된 범위 내에서 효율적인 정보검색은 가능하지만 대용량의 멀티미디어 데이터에 대한 수작업에 의한 색인작업이 시간과 비용면에서 너무나 방대하고, 멀티미디어 데이터의 성격상 데이터의 속성들을 소수의 키워드로 표현하기가 어렵다. 그리고 데이터 내용에 대해 색인자와 탐색자가 주관적 관점으로 텍스트로 표현하게 됨으로써 객관성이 결여되어 검색의 불일치를 발생시켜 검색의 효율

1) 한국데이터베이스진흥센터, 『'97 데이터베이스백서』, 서울; 同 센터, 1997. pp. 118-119.
2) 이종구, 장옥배, "비디오 인덱싱을 위한 텍스트 추출과 인식", 《한국정보과학회 가을 학술발표 논문집》, Vol. 24, No. 2(1997), pp. 351-354.

성이 크게 떨어진다.

세 번째는 내용기반 정보검색 방법으로, 멀티미디어 데이터의 특징정보를 추출하여 이를 기반으로 색인과 검색을 수행하는 방법이다.[3] 이 방법은 자동으로 멀티미디어 데이터의 특징을 추출하여 인덱싱함으로써 데이터베이스의 구축에 시간과 비용을 줄일 수 있지만, 멀티미디어 데이터로부터 정확한 이미지의 내용정보를 추출하기가 어렵다는 단점이 있다. 그러나 최근에는 컴퓨터 비젼이나 영상처리, 컴퓨터 그래픽스와 같은 분야의 연구결과들을 이용하여 멀티미디어 데이터로부터 정확한 정보를 추출하려는 연구들이 시도되고 있다. 이러한 연구들은 이미지 자체가 지닌 특징 데이터를 그대로 검색에 이용하려는 시도로, 이미지에 대한 특징 데이터는 영상처리나 컴퓨터 비젼 분야에서 연구되고 있는 이미지 특징 추출 시스템의 출력결과를 이용하는 것을 전제로 한다. 이와 같은 내용기반 멀티미디어 영상정보 검색기법은 다양한 응용분야에서 효율적이고 중요한 기술로 적용될 수 있다. 그럼에도 불구하고 현재까지는 이에 대한 기술력이 국내외적으로 아직 초보적인 단계에 머물러 있다.

이에, 본 연구는 최근까지의 멀티미디어 영상정보의 검색에 관한 여러 연구들 중 내용기반의 영상정보 검색방법에 관한 연구들을 살펴보고, 이들 연구에서 채용하고 있는 영상정보에 대한 다양한 특징정보의 추출과 인덱싱 기술에 대해 고찰해 보고자 한다. 아울러 영상정보의 효과적인 검색을 위해 채용되고 있는 색인구조의 설계 요건에 대해서도 살펴본다. 마지막으로 이러한 내용기반 영상정보 검색기술을 적용한 다양한 실제 시스템을 소개한다.

Ⅱ. 영상정보 검색기술의 연구현황

현재까지 주로 연구되고 있는 내용기반 영상정보 검색기술은 대상 데이터의 종류에 따라 이미지 검색과 비디오 검색의 두 가지로 나누어진다. 이미지 검색기술은 정지영상을 대상으로 하는 기술이며, 비디오 검색은 동영상을 대상으로 하는 기술이라 할 수 있다.[4]

1. 내용기반 이미지 정보검색 연구 현황

3) IBM Almaden Research Center, "Query by Image and Video Content: The QBIC System", *IEEE Multimedia*(Sept. 1995), pp. 23-32.

4) 이미숙, 황본우, 이성환, "내용 기반 영상 및 비디오 검색기술의 연구 현황", 《정보과학회지》, 제15권, 제9호(1997. 9), pp. 10-19.

이미지 정보검색에 관한 연구들은 아직 초보적인 단계로서 자동 단순 속성(attribute) 추출이나 수동식 속성 기술, 또는 캡션자료를 이용한 텍스트의 매칭에 의한 경우가 대부분이다. 특히, 이미지 정보는 색, 모양, 질감, 공간관계 등의 여러 가지 속성을 지니고 있기 때문에 이미지 데이터 자체의 내용을 자동으로 추출 분석하고 인덱싱하여 이를 검색에 사용함으로써 관련 정보를 검색하는 내용기반(content-based)의 검색기술에 대한 연구가 필요하다. 내용기반 이미지 검색기법은 이미지처럼 복잡한 구조를 가진 데이터에 효율적으로 접근하기 위해 제시된 대표적 방법이다. 내용기반 이미지 검색이란 이미지에서 이미지의 내용으로 표현되었던 특징들을 추출한 후, 추출된 특징 데이터들을 기반으로 이미지를 검색하는 방법을 말한다.[5] 내용기반 이미지 검색에 대한 연구로는 이미지 처리과정을 통해 이미지 데이터를 나타내는 특징값(characteristic value)을 추출하는 방법에 관한 연구와 추출된 특징값들을 효율적으로 인덱싱하는 방법에 관한 연구, 그리고 사용자 질의방법과 질의 처리방법에 대한 연구가 주를 이루고 있다.

이미지 정보를 효과적으로 검색하기 위해서는 먼저 이미지 정보의 속성이 추출되어야 한다. 내용기반 이미지 정보검색을 위한 속성으로는 색상(color), 질감(texture), 형태(shape), 스케치(sketch), 양(volume), 공간 유사성(spatial similarity), 브라우징(browsing), 객관적 속성(objective attribute), 주관적 속성(subjective attribute), 연속성(sequences), 키워드(key words), 자연어 문장(natural language text), 도메인 개념(domain Concepts) 등이 있다.[6] 이 중에서 특히 색상, 질감, 형태 그리고 공간관계에 의한 이미지 정보검색이 주로 연구되고 있다.

1) 색상에 의한 이미지 검색방법

색상에 의한 이미지 검색방법은 내용기반 이미지 검색의 초기부터 사용된 방법으로, 이미지 내의 조명, 관측위치, 크기 변화에 어느 정도 무관하다는 장점이 있다. 이 방법은 주로 색상 히스토그램 중복도, 색상 히스토그램 거리 가중치, 평균색상 거리, 색상 인접정보 등을 이용한 색상 유사도(color similarity)를 측정함으로써 이루어진다. 이러한 유사성은 이미지의 전체적인 색상 분포만을 계산하는 전역적 방법과 이미지를 일정한 영역으로 나누어 각각의 색상 분포를 계산하는 지역적 방법으로 구분할 수 있다.

색상 히스토그램을 이용한 이미지 검색은 이미지를 구성하는 지역적인 색상분포를 고려하지 않아서 시각적으로 유사한 색상 히스토그램을 가지는 영상이더라도 실제로는 전혀 다른

5) B. Furht, S.W. Smoliar, H. Zhang,, *Video and Image Processing in Multimedia Systems*, Kluwer Academic Publishers, 1994.

6) Venkat N. Gudivada, Gwang S. Jung, "An Algorithm for Content-based Retrieval in Multimedia Databases", *Proceedings of the International Conference on Multimedia Computing and Systems*, Hiroshima, Japan, June, 1996.

이미지를 검색할 가능성이 크다. 이와는 반대로 시각적으로는 전혀 다른 영상이더라도 색상 히스토그램에 대한 유클리드 거리(Euclidean distance)를 계산하였을 때 동일한 영상으로 판단하는 경우가 발생한다. 이 문제는 이미지의 여러 부영역에 대해 지역 색상 히스토그램을 유지하면 해결될 수 있지만,[7] 이를 유지하기 위한 저장공간의 비용상승이 문제가 된다.

색상 시그니쳐(color signature)를 이용한 이미지 검색방법은 색상 히스토그램 방식이 색상의 지역적인 분포를 고려하지 못한다는 단점을 개선하기 위해 제안된 방법이다.[8] 이 방법은 이미지의 색상 특징이 이미지 전체 화소(pixel)의 대부분을 차지하는 몇 개의 주요 색상(dominent color)에 의해 결정된다는 점에 착안하여, 주요 색상에 대하여 색상분포를 시그니쳐 형식으로 간단하게 표현한 후, 질의 이미지의 시그니쳐와 데이터베이스 이미지의 색상 시그니쳐를 AND연산하고 그 결과 비트열에서 비트값이 1인 비트의 수를 세어 두 이미지간의 유사도를 계산한다. 그러나 이 방법은 이미지의 전역적인 색상분포 특징을 고려하지 않아서 카메라 시점이 이동된 동일한 이미지나 이미지를 구성하는 물체가 이동, 회전, 변형된 경우에는 효과적으로 검색할 수 없다는 단점이 있다.

색상 정보를 사용한 이미지 검색에 있어서 색상공간 상에서의 거리 측정을 정확하게 하기 위한 색상 정보의 추출방법은 RGB 색상모델을 비롯하여 HSV, YIQ, YUV 색상모델 등을 이용할 수 있다.[9]

(1) RGB 색상모델

RGB 색상모델은 이미지 화소값을 R(Red), G(Green), B(Blue)의 세 값의 조합으로 나타내는데, 이 모델은 컴퓨터에서 색상을 표현할 때 일반적으로 사용되지만 시각적으로 균등하게 분포되지 않는다는 단점이 있다. 그래서 색상 히스토그램은 대개 RGB 색상모델을 사람의 인식시스템과 유사한 다른 색상 모델로 변환시켜, 정량화 과정을 거쳐 추출한다.

(2) YIQ, YUV 색상모델

YIQ, YUB 색상모델은 컬러 영상시스템에서 채용된 선형 색상변환 모델로, 두 방식 모두 RGB 색상모델을 한 개의 밝기(luminance)채널과 두 개의 색채(chrominance)채널로 변환시킨다. YIQ는 NTSC방식에, YUV는 PAL, SECAM 방식의 컬러 텔레비젼에 사용된다.

---

7) Y. Gong, H. Chau, and X. Guo, "Image Indexing and Retrieval Based on Color Histogram", *Proc. of the 2nd International Conference on multimedia modeling*, Singapore, Nov. 1995, pp. 115-126.

8) Tat Seng. Chua, Kian-Lee Tan, and Beng Chin. Ooi, "Fast Signature-based Color-Spatial Image Retrieval", *Proc. of the 4th ACM Multimedia Computing and Systems*, Ottawa, Ontario, Canada, June 3-6, 1997, pp. 362-369.

9) 임형준, 『결러내기 방식의 컬러 이미지 검색시스템 설계』, 미간본 석사학위 논문, 부산대학교 대학원, 1998.

(3) HSV 색상모델

HSV 색상모델은 인간의 시각능력에 유사한 색상모델로서, 색도(hue), 채도(saturation), 명도(value) 등의 각 구성요소는 인간의 시각시스템에 직관적으로 영향을 미치도록 설계되었다. HSV와 유사한 색상모델로 HSI(Hue, Saturation, Intensity), HSL(Hue, Saturation, Lightness) 색상모델 등이 있다. 이들 모델들은 RGV 채널값을 변환하여 색도, 채도, 명도값으로 변환한다.

그 외의 색상모델로는 Munsell 색상 순서시스템, CIE 색상모델 등이 있다.

이미지 검색을 위한 또 다른 방법으로 히스토그램의 두드러진 특징을 이용하는 방법이 있다. 이것은 색상 히스토그램 방법의 문제점이 인덱싱을 위한 추가적인 저장공간의 필요성과 계산의 복잡성에 있다는 점에 착안하여, 이에 대한 대안적인 방법으로 이미지의 색상분포가 확률분포로 해석 될 수 있는 모멘트를 이용하는 방법이다. 이 방법을 이용하여 인덱스의 수를 줄임으로써 저장공간을 절약할 수 있고, 검색속도의 향상 및 강건성을 지니는 장점을 지니게 된다. 하지만 이 방법은 low-order의 모멘트가 빛의 변화에 민감하다는 단점을 가지고 있다.[10]

2) 질감에 의한 이미지 검색방법

의류나 직물디자인, 벽지의 디자인 같은 응용영역의 특정한 질감 특성을 갖는 이미지의 검색에서는 이미지의 색상정보만으로는 정확한 검색결과를 기대할 수 없다. 따라서, 이미지의 질감정보를 추출하여 사용하는 질감정보 기반의 검색방법이 필요하다. 일반적으로 이미지검색에 사용되는 질감은 통계적인 특징과 구조적인 특징, 그리고 스펙트럼 특징으로 표현된다. 통계적인 특징은 이미지의 거칠고 부드러운 정도를 나타내며, 구조적인 특징은 이미지 내의 수평선의 반복과 같은 규칙적인 배열을 표현한다. 스펙트럼 특징은 주파수 스펙트럼의 형태를 기반으로 스펙트럼의 높은 에너지 성분을 분석함으로써 이미지 내의 전체적인 주기성을 알아내는데 사용된다. 질감 특징을 표현하기 위한 초기의 연구들에서는 대부분 단일 해상도에서 분석을 하였지만 최근에는 다중해상도와 다채널 질감분석에 초점을 두고 있다.

이와 같이 이미지 정보를 질감에 의해 모델링 하는 기법들은 다양하게 제안되고 있으며 각각의 특징은 다음과 같다.[11]

---

10) F. Idris and S. Panchanathan, "Review of Image and Video Indexing Techniques", *Journal of Visual Communication and Image Representation*, Vol. 8, No. 2(June 1997), pp. 146-166.

11) 이희영, 양형정, "내용 기반 이미지 정보 검색 기술 연구", 《전주공업전문대논문집》, 제27집(1997. 10), pp. 110-118.

(1) **Reaction-diffusion 모델:** 결정적 방법으로 점이나 줄무늬 등의 자연적 무늬를 나타내는데 효과적이다.

(2) **Macro random field 모델:** 잔디밭이나 모래사장, 모피와 같이 임의적이면서, 규칙적이고, 구조화된 패턴을 모델링 하는데 유용하다.

(3) **Cluster-based probability 모델:** 고차 확률(a higher-order probabilistic)모델로서 보다 복잡한 구조를 모델링 하는데 효과적이다.

(4) **Wavelet[12]-based 모델:** Wavelet에 기반한 시간-주파수 분석은 질감의 다양한 타입의 분석에 있어서 아주 효율적이고 강력하다.

(5) **A New Wold 모델:** 이미지의 질감특성을 주기성, 방향성, 임의성에 의한 유사성을 계산하여 모델링 하는데 유용하다.

(6) **Stochastic 모델:** 시간차에 따른 질감 변화를 표현하기 위한 모델로서 동영상 이미지에서 흐르는 강물, 바람에 날리는 나뭇잎 등을 나타내기 위한 지역 집약속성(local collective property)을 지원한다.

(7) **Synthetic flames 모델:** 다각형 조각을 이용하여 불꽃을 모델링 한다.

3) 형태에 의한 이미지 검색방법

형태에 의한 이미지 정보검색은 이미지나 이미지내의 객체의 모양을 일정한 수치 데이터로 변환하여 이를 기반으로 검색한다. 여기서 먼저 고려하여야 할 것은 이미지가 가지는 형태 특성이다. 첫 번째는 이미지내의 객체들의 윤곽선이 나타내는 형태로서, 형태정보의 추출이 간단한 반면, 동일한 물체에 대해서 보는 각도에 따라 서로 다른 형태의 윤곽선이 나타날 수 있다. 두 번째는 방향에 상관없이 물체들이 지니는 고유한 형태로서, 보는 각도, 크기, 위치 변화 등에 무관한 형태정보를 갖는 반면, 형태 추출과 검색과정이 복잡하다는 단점이 있다.

이렇게 추출된 형태정보는 여러 가지 방법에 의해 검색이 이루어진다. 첫 번째는 사용자로부터 입력된 스케치 이미지를 이미지 데이터베이스로부터 추출된 에지(edge)와 세선화된 이미지와의 관련성을 비교하여 검색하는 방법이다. 두 번째는 객체의 윤곽선을 체인으로 구성하고 체인코드의 매칭으로 이미지 정보를 검색하는 방법이며, 세 번째는 객체의 윤곽을 모델링한 1D 템플리트를 주로 사용하고, 2D 템플리트를 이진 레벨패턴을 매칭시키는데 이용하는 방법이 있다.

---

12) 웨이브릿 변환(Wavelet transform)은 신호분석(signal analysis)이나 이미지 압축 등의 여러 분야에서 최근에 많이 사용되고 있는 기술로 기존의 코사인 변환(cosine transform)이나 푸리에 변환(Fourier transform)에 비하여 공간에 대하여 지역적인 특성(localized in space)을 가지고 있으며, 이미지에 대하여 다해상도 분석(multi resolution analysis)이 가능한 특징을 가지고 있어서 질감이나 모양 등의 이미지 데이터의 특징 추출에 유용하게 사용될 수 있다.

이러한 형태 특징 데이터는 내용기반 이미지 검색에서 가장 응용이 어려운 부분 중의 하나이다. 형태정보를 사용한 검색방법은 다른 검색방법에 비하여 사용자에게 고수준의 검색환경을 제공할 수 있으나, 실세계에 존재하는 다양한 물체의 형태를 표현할 수 있는 수학적 모델이 없으며, 형태간 유사도를 측정할 수 있는 기준이 부족하다는 문제점을 가지고 있다. 실제로 이러한 형태에 의한 유사성 검사는 모델기반 비젼 응용(model-based vision application)과 같은 전문분야에서도 아직 어려운 연구 중의 하나로, 최근에는 많은 통계학적인 방법을 이용하여 보다 효율적인 방법들을 연구하고 있다.

4) 공간관계에 의한 이미지 검색방법

내용기반 이미지 검색에서 색상, 질감, 형태 등의 가시적 정보와 함께 이미지 내의 객체들 사이의 공간관계 정보는 이미지 정보를 검색하는데 가장 중요한 속성중의 하나로 인식되고 있다.[13] 이러한 공간관계 정보는 특히 공간관계에 의해 이미지를 표현할 경우 명시적으로 주어지지 않은 공간관계를 연역할 수 있기 때문에 지리정보시스템 같은 공간관계가 중요시되는 응용분야에는 특히 유용하다.

공간관계에 의한 이미지 검색은 결국 내용기반 이미지 정보검색을 위한 논리적 구조표현에 의해 이루어진다고 할 수 있다. 이러한 이미지 내의 객체들의 공간관계의 논리적 구조표현에 관한 연구로는 2D 스트링, 순서 트리플(ordered triple), E-R 다이아그램, 기호배열(symbolic array) 등이 있다.

(1) 2D 스트링[14]

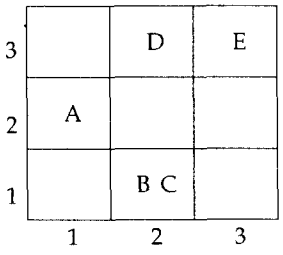
이미지 데이터 표현에 있어서의 2D 스트링은 이미지 처리나 패턴인식 등의 선행처리에 의해 심볼화된 객체를 X축과 Y축에 병렬된 최소의 사각형(Minimum Bounding Rectangles: MBR)으로 나타내고, 이들 사이의 공간관계성을 표현함으로써 심볼화된 이미지를 구축한다.

즉, S={O1,O2,.....,On}을 n개의 객체 집합이라 할 때, R={"=", "<", ":"}을 이용하여 X축과 Y축에 따라 객체들 사이의 관계를 나타낸다. 여기서, "="는 X나 Y 좌표값이 같음을 의미하고, "<"는 왼쪽(west)이나 아래쪽(south)을, ":"는 하나의 사각형에 동시에 두 개의 객체가 위치함을 나타낸다. 예를 들어, 객체 S={A, B, C, D, E}인 <그림 1>에 대한 2D 스트링은 (A<B: C=D<E, B: C<A<D=E)로 표현된다. (A<B: C=D<E)는 X축에 대해 A는 B의 왼쪽에 존재하고, B와 C는 동

13) C. C. Chang and S. Y. Lee, 'Retrieval of Similar Pictures on Pictorial Databases," *Pattern Recognition*, Vol. 24(1991), pp. 675-680.

14) S. K. Chang, Q. Y. Shi and C. W. Yan, "Iconic indexing by 2D string," *IEEE Transaction, Pattern Analysis*(Mar. 1987), pp. 413-428.

일한 사각형에 위치하고 D는 B, C와 같은 X축 좌표에, 그리고 D는 E의 왼쪽에 위치함을 의미한다.



〈그림 1〉 심볼화된 이미지

이렇게 표현된 이미지들에 대해서 2D 스트링의 매칭에 의한 유사 이미지 검색이 이루어진다. 2D 스트링에 의한 이미지 데이터의 표현은 이미지의 빠른 검색을 지원하기 위한 아이콘 인덱스의 구축을 간단하고 효율적으로 수행할 수 있게 한다.

(2) 순서 트리플

Chang과 Lee[15]는 아이콘 인덱스 구축을 위해 2D 스트링으로 표현된 이미지 데이터를 두 개의 객체와 공간관계를 명시하는 2D 스트링의 변형인 순서 트리플로 표현하고 있다. 두 개의 객체를 Oi, Oj라고 하고 두 객체 사이의 공간관계를 rij라 할 때, 이들을 위한 순서 트리플은 (Oi, Oj, rij)가 된다. 그리고 이러한 공간관계는 8개의 방위표시를 기준으로 하여 북쪽에서부터 시작하여 시계 반대방향으로 1~8까지의 정수를 할당하고, 같은 위치에 있을 경우에는 9를 할당하는 방법으로 순서 트리플을 구한다. <그림 1>을 예로 든다면, 객체 A의 북동쪽에 D가 위치하므로 이 정보는 (A, D, 8)로 표현된다. rij'는 rij의 반대 방향을 의미하고 (Oi, Oj, rij)는 (Oi, Oj, rij') 와 같은 트리플이다. 즉, 위의 예에서 (A, D, 8)은 (D, A, 4)와 같은 트리플을 의미한다. 따라서 하나의 이미지 내의 n개의 객체는 $n*(n-1)/2$개의 트리플로 표현된다. 이러한 순서 트리플들은 해쉬 함수를 이용하여 도치화일로 구성되고 , 사용자로부터 주어진 질의는 역시 트리플의 형태로 변환되어 이 도치화일에서 각 이미지를 표현하고 있는 트리플 집합의 부분집합이 되는가에 따라 검색결과가 도출된다.

이 방법은 이미지 데이터를 표현하는데 있어서 트리플 구조를 이용함으로써 자료구조가 간단하면서도 객체들 사이의 공간관계를 적절히 표현할 수 있으며, 순서 트리플로 이미지 데이

15) C. C. Chang and S. Y. Lee, "Retrieval of Similar Pictures on Pictorial Databases", *Pattern Recognition*, Vol. 24(1991), pp. 675-680.

터를 색인하여 도치화일을 구성함으로써 빠른 검색을 지원한다는 것이 장점이다. 그러나 검색 대상이 단순객체로 한정되며 객체의 의미를 고려하지 못한다는 단점도 있다.[16]

(3) E-R 다이아그램[17]

이미지 데이터 내의 객체와 속성, 그리고 관계성을 표현하는 방법으로 E-R 다이아그램을 이용한다. 여기에서 전처리 과정을 통해 인식된 객체들은 각각 엔티티로, 객체의 색이나 크기 등의 특정한 속성은 애트리뷰트로, 객체들 간의 관계성은 위치정보를 나타내는 공간관계와 악수나 마주 보는 행동 등을 나타내는 행위관계로 나타난다.

사용자 질의에 대한 검색은 E-R 다이아그램에 의해 표현된 이미지 데이터와 엔티티의 애트리뷰트 간의 유사성과, 관계간의 유사성에 의해 이미지에 대한 유사성을 계산함으로써 이루어진다. E-R 다이아그램에 의한 이미지 표현은 객체들 사이의 공간관계 외에 객체 자체의 속성을 표현할 수 있기 때문에 이미지 정보의 검색시 가시적 속성뿐 아니라 논리적 검색도 가능하게 하며, E-R 다이아그램을 위해 개발된 데이터베이스의 많은 기능을 이미지 데이터베이스에 쉽게 적용할 수 있다는 장점이 있다. 그러나, E-R 다이아그램을 이용하여 이미지 데이터를 표현하기 위해서 객체와 속성을 추출하는데 어려움이 있으며, 여전히 검색대상의 객체들이 단순 객체의 형태로서 개념에 기반한 검색을 지원하지 못한다. 또한 이미지 구조 표현을 위한 E-R 다이아그램과 질의의 E-R 다이아그램 간의 매칭이 복잡한 것도 하나의 단점이다.

(4) 기호배열[18]

이미지 정보 내의 공간 관계성을 표현하는 또 다른 방법으로 기호배열을 이용하는 방법이 있다. 기호배열에 나타나는 심볼은 상대적인 크기나 거리정보를 배제한 공간 위치정보만을 표현한다. 이 기호배열은 하나의 셀이 다른 기호배열로 매핑되는 다단계 계층구조를 갖는다. 예를 들어 <그림 2>에서 유럽을 나타내는 정수 4는 두 개의 셀에 위치함으로써 그 위치가 아프리카의 북쪽 내지는 복동쪽이라고 할 수 있다는 정보를 표현하고 있다.

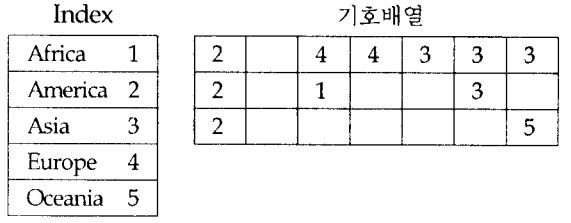
이 기호배열은 <그림 3>과 같이 다시 유럽을 나타내는 셀이 유럽 내의 나라들에 대한 공간정보를 표현한 다른 기호배열로 매핑되는 다단계 계층구조를 갖는다. 기호배열에서 제공하는

16) 양형정 외, "트리플 기반 이미지 색인을 이용한 개념 기반 이미지 검색", 《정보과학회논문지(B)》, 제25권, 제3호(1998. 3), pp. 478=489.

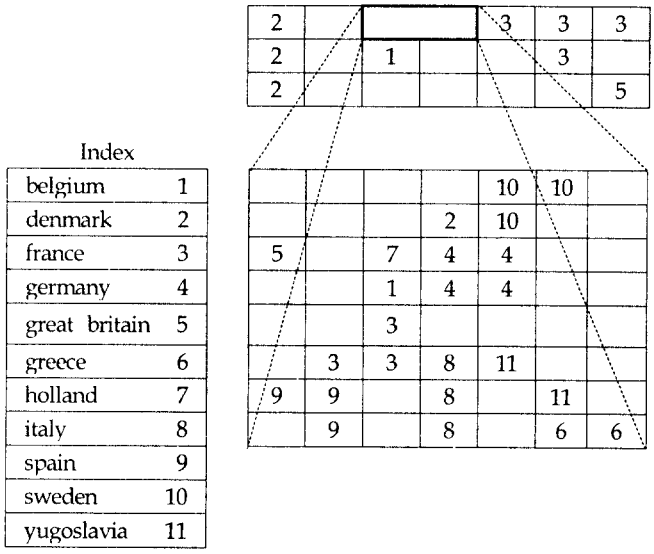
17) Y. Alp Aslandogan et al. "Design, Implementation and Evaluation of SCORE(a System for Content based Retrieval of pictures)", *Proceedings of 11th International Conference of Data Engineering(1995)*, pp. 280-287.

18) D. Papadias and T. Sellis, "Spatial Reasoning Using Symbolic Arrays", *Proceeding of International Conference GIS-From Space to territory Theories and Method of Spatio-temporal Reasoning in Geographic Space*, Pisa. Italy: Springer Verlag, 1992

계층구조는 여러 단계의 추상화 레벨에서 복잡한 공간관계를 표현하는데 사용될 수 있지만, 단순히 객체들 사이의 공간관계성만을 이용하여 복합객체를 형성할 뿐 사용자가 원하는 객체들의 의미를 고려하여 하나의 새로운 개념으로 유추되는 개념객체를 의미하지는 않는다.



〈그림 2〉 기호배열에 의한 이미지 표현



〈그림 3〉 기호배열을 이용한 이미지의 다단계 계층 표현

이상과 같이, 이미지 정보를 효과적으로 검색하기 위한 연구로서 색상, 질감, 형태, 공간관계에 의한 이미지 검색방법에 관해서 살펴보았다. 이러한 검색방법들은 영상의 일부 속성만을 다루고 있기 때문에 한 가지의 방법만으로는 사용자의 다양한 요구에 부응할 수 없다는 단점이 있으며, 특히 대용량의 영상 데이터베이스에 적용시킬 때에는 검색효율의 저하가 문제점으로 대두된다. 그러므로 검색효율의 향상을 위해서 여러 가지 속성들이 복합적으로 사용되고 있으며, 아울러 텍스트기반의 검색방법이 병행되고 있기도 하다. 그렇지만 진정한 의미의 내

용기반 영상검색시스템의 개발을 위해서는 영상 내의 객체들에 대한 인식기술과 고수준의 사용자의 질의로부터 적용이 가능한 적절한 검색조건을 생성할 수 잇는 기술이 필요하다. 아울러 대용량 데이터베이스에 대한 효율적인 검색을 위한 색인구조의 개발이 병행되어야 한다.

2. 내용기반 비디오 검색연구 현황

기존의 비디오 정보(텔레비전 방송이나 영화 또는 비디오)는 그 성격상 순차적인 검색방법을 제공하며, 검색에 의해 정보를 찾아도 그 자료를 직접 이용할 수가 없게 된다. 그래서 보다 빠르고 효과적인 비디오 정보의 활용을 위해서는 기존의 아날로그 방식의 비디오를 내용기반으로 검색할 수 있게 하는 것이 필요하다. 내용기반 비디오 검색은 비디오의 내용으로 정보를 검색하는 것으로, 영상, 음향, 텍스트로 검색하는 기술을 말한다.

내용기반 동영상 검색연구는 동영상의 내용을 대표할 수 있는 특징을 추출하여 이를 색인하기 위한 동영상 파싱, 대용량의 동영상 데이터와 이와 관련된 메타 데이터를 효율적으로 저장하기 위한 동영상 데이터 압축 및 저장, 그리고 사용자가 동영상 검색을 용이하게 할 수 있는 환경을 제공하는 검색 및 브라우징 등을 핵심기술로 하고 있다. 그러나 여기에서는 동영상 정보의 색인을 위한 동영상 파싱에 있어서의 비디오 분할과 비디오 색인에 대하여 살펴본다.

비디오 파싱[19]은 대용량의 비디오 데이터를 효율적으로 저장, 관리하고 사용자가 쉽게 검색할 수 있는 환경의 제공을 목적으로 하고 있다. 이러한 비디오 파싱은 대용량 비디오를 보다 적은 용량의 요약형 비디오로 만들기 위한 비디오 분할과 비디오의 내용을 효과적으로 표현하기 위한 비디오 색인과정으로 구성된다.

1) 비디오 분할

일반적으로 비디오는 프레임(frame), 샷(shot), 장면(scene)의 세 가지 구성요소로 이루어진다. 프레임이란 영상을 구성하는 최소 단위로서, 필름 한 장에 해당하는 하나의 영상을 나타낸다. 샷은 하나의 카메라로 촬영한 영상들을 나타내며, 샷 내에서는 필름이 끊기지 않고 연속적으로 연결되어 있다. 샷은 비디오 분할의 기본단위로 사용된다. 장면은 연속된 일련의 샷으로 구성되며, 주인공이나 특정한 장소와 같이 하나의 대상을 연속하여 촬영한 영상을 나타낸다. 비디오 분할은 컷(cut)이나 점진적 장면전환(gradual transition) 효과에 의해서 발생되는 샷 또는 장면 사이의 경계를 검출하는 기술이다. 비디오 분할기술은 분할 대상 비디오의 종류에 따라서 비압축 비디오에 대한 분할기술과 압축 비디오에 대한 분할기술로 나누어진다.[20]

19) 이미숙, 황본우, 이성환, op. cit.

(1) 비압축 비디오에 대한 분할기술

비압축 비디오에서의 비디오 분할은 압축되지 않은 비디오 스트림을 대상으로 인접한 프레임간의 특징차를 비교하여 샷의 경계를 찾아내는 기술로 사용하는 특징의 크기에 따라 화소 단위 분할방법, 부분영역 단위 분할방법, 프레임 단위 분할방법 등이 있다.

① 화소 단위 분할방법

화소 단위 분할방법은 연속된 두 프레임에서 상응하는 화소 사이의 특징차를 구하여, 그 차이가 임계값 이상이 되면 컷으로 검출하는 방법[21]으로 주로 화소의 농도값을 특징으로 사용한다. 화소 비교법은 한 쌍의 연속 프레임에서 대응하는 화소의 세기를 비교하여 얼마나 많은 변화가 일어났는지를 측정하여 특정 파라미터를 비교하여 절대차의 합이 임계값을 초과하면 프레임 사이의 장면변화가 존재하는 컷으로 인정하여 검출하는 방법이다. 그러나 이 방법은 카메라의 움직임과 물체의 움직임에 아주 민감하여 컷 검출에 어려움이 크다는 단점이 있다. 이 외의 방법으로는 프레임간의 화소값 차이를 구해 어떤 특정 임계값을 두고 이 값을 초과하는 픽셀들의 수를 셈으로써, 초과된 픽셀의 수가 컷을 검출하기 위한 임계값을 초과하는 지의 유무를 조사하는 방법이 있다. 그러나 이 방법은 줌(zooming), 패닝(panning)과 같은 카메라의 연산과 수평선회 움직임에 민감한 반응을 일으킨다는 단점이 있다. 이러한 단점을 보완하기 위한 방법으로, 카메라의 연산과 잡음을 최소화하기 위하여 비교 전에 3x3의 평균필터를 추가적으로 사용함으로써 보다 나은 결과를 산출할 수 있는 방법이 제안되고 있다. 그러나 이 방법은 또한 프레임 내의 객체의 움직임 영향을 근본적으로 해결할 수 없다는 단점이 있다.

② 부분영역 단위 분할방법

한 장의 프레임을 몇 개의 부분 영역으로 나누고 부분 영역별로 특징을 구한 후, 연속된 두 프레임에서 상응하는 부분 영역간의 특징차가 임계값 이상이 되는 프레임을 컷으로 검출하는 방법이다. 이 방법에는 주로 부분 영역별 히스토그램을 특징으로 사용한다.[22] 히스토그램 비교법은 화소 비교법보다 카메라 또는 물체의 움직임에 덜 민감하여 컷 검출에 정확성과 속도면에서 좋은 결과를 가져올 수 있는 방법이다.

③ 프레임 단위 분할방법

---

20) 김영민 외, "내용 기반 비디오 검색 기술", 《정보과학회지》, Vol. 16, No. 8(1998. 8), pp. 39-49.

21) H. Zhang, A. Kankanhalli and S. W. Smoliar, "Automatic partitioning of Full-motion Video," *Multimedia Systems*, Vol. 1, No. 1(1993), pp. 10-28.

22) D. Swanberg, C. F. Shu, and R. Jain, "Knowledge Guided Parsing in Video Databases," *Proc. of SPIE'93 - Storage and Retrieval for Image and Video Database*, San Jose, CA, Vol. 1908(1993), pp. 13-24.

프레임 단위 분할방법은 가장 일반적으로 사용되는 방법으로 한 장의 프레임에 대하여 대표적인 특징값을 구하고, 각 프레임간의 특징값의 차이가 임계값 이상이 되는 프레임을 컷으로 검출하는 방법이다. 프레임 단위의 검출방법에는 다음과 같은 방법이 있다.[23]

- 농도 히스토그램을 사용한 검출방법
- 색상 히스토그램을 사용한 검출방법
- 차영상(difference image)의 히스토그램을 사용한 검출방법
- 부분 영역별 히스토그램을 사용한 검출 방법
- 에지 영상을 사용한 검출방법
- 화소간 농도차를 사용한 검출방법

(2) 압축 비디오에 대한 분할기술

압축 비디오의 분할 방법은 사용되는 특징에 따라서 DCT 계수를 사용한 방법, 주파수 대역을 사용한 방법, VQ 레이블을 사용한 방법, 그리고 이동벡터를 사용한 방법으로 나누어진다.

① DCT 계수를 사용한 검출방법

DCT(Discrete Cosine Transform) 계수를 사용한 방법은 연속된 프레임의 DCT 계수의 차가 임계값 이상이 되는 프레임을 컷으로 검출하는 방법으로써, 대부분의 압축 비디오 분할방법이 DCT 계수를 사용하고 있다.[24] 이 방법은 M-JPEG나 MPEG 형식으로 압축된 비디오에 적용할 경우 효과적이고 처리시간이 빠른 반면, 점진적인 장면전환은 검출하지 못한다는 단점을 가지고 있다.

② 주파수 대역의 계수를 사용한 검출방법

Wavelet 방법에 의해 압축된 비디오 데이터를 분할할 때 효과적이다.

③ VQ 레이블을 사용한 검출방법

VQ(Vector Quantization) 방법에 의해 압축된 비디오 데이터를 분할할 때 효과적이다.

④ 이동 벡터를 사용한 검출방법

이동 벡터는 카메라 또는 물체의 상대적인 움직임에 의해 발생하는 것이다. 비디오의 인접

23) 이종구 외, "뉴스 비디오 검색을 위한 자동 인덱스 모델의 설계 및 구현," 《한국정보과학회 인공지능연구회 97년도 춘계 학술발표 논문집》(1997. 3), pp. 103-112.

24) F. Arman, A. Hsu and M. Y. Chiu, "Feature Management for Large Video Databases," *Proc. of SPIE'93 - Storage and Retrieval for Image and Video Databases*, San Jose, CA, Vol. 1908(1993), pp. 2-12

프레임들의 이동 벡터를 검출하면 카메라 또는 물체의 움직임에 대한 방향과 크기를 알 수 있다. 이동 벡터를 이용한 검출방법은 이동 벡터의 개수를 특징값으로 사용하여 컷을 검출하는 기술로 장면이 급변하는 변화는 대부분의 이동 벡터의 값이 0이 된다는 사실에 기반하여 비디오를 분할한다. 이동 벡터의 방향과 크기의 갑작스러운 변화는 장면변화가 있음을 의미한다. 이 방법은 이동 벡터의 방향을 분석함으로써 카메라 동작에 의한 장면전환도 분석이 가능하다는 장점이 있다. 그러니 이동 벡터를 계산하는데 많은 시간이 걸린다는 단점도 있다.

2) 비디오 색인

비디오 색인은 분할된 비디오의 대표 특성을 추출하여 특징정보로 사용하기 위한 기술을 의미한다. 그러나 비디오 데이터는 시간과 공간이 결합된 복잡한 형태를 띠고 있으므로 대표 특성을 찾아 색인하는 것은 매우 어려운 기술의 하나이다.

비압축 비디오나 압축 비디오 상에서의 비디오 색인 방법은 다음과 같은 방법들이 있다.

(1) 텍스트기반 색인방법

수작업에 의해 비디오의 의미를 사람이 파악하여 텍스트로 색인하는 방법이다. 그러나 이 방법은 많은 시간과 노력이 필요하며, 색인자와 사용자의 관점에 따라 색인어를 다르게 부여하거나 해석할 여지가 있으며, 특히 추상적인 내용의 색인어 부여와 해석은 더욱 그러하다. 따라서 일관성을 가지기 어렵다는 단점이 있다. 그래서 텍스트 이외의 비디오가 가진 고유의 내용적인 특징을 사용하여 비디오를 표현하려는 연구가 진행 중이다.

(2) 대표 프레임의 영상을 사용하는 방법

샷의 내용은 대표 프레임으로 표현될 수 있으며, 이들 영상들은 색상, 형태 등을 특징정보로 사용한다.[25] 그러나 비디오는 정지영상과는 달리 시간적 흐름에 대한 정보를 가지고 있기 때문에 정지영상과 같은 방식으로 색인하기에는 미흡한 점이 많다.

(3) 카메라 동작 분석 및 이동물체 분석을 이용한 색인방법

비디오 정보의 특성상 카메라의 동작 분석 및 이동물체의 분석을 통해 비디오의 시간적 특성을 표현하고자 하는 연구가 진행되고 있다.[26] 이러한 카메라 동작(fixed, panning, tilting,

25) 황본우, 노형기, 이성환, "색상 및 형태 정보를 이용한 내용 기반 영상 검색 시스템의 Web상에서의 구현," 《한국정보과학회 봄학술발표논문집》, 제25권, 제1호(1998, 4), pp607-609.

26) M. S. Lee et al., "Automotive Video Parsing Using Shot Boundary Detection and Camera Operation Analysis," *Proc. of 14th ICPR, Brisbane,* Australia, August 1998.

zooming, booming, dollying, tracking)은 크게 수평 이동, 상하 이동, 그리고 확대·축소의 세 종류로 분류할 수 있는데[27], 이에 대한 분석은 현재 수학적 모델을 사용한 방법과 패턴을 사용한 방법으로 나뉘어 연구가 진행되고 있다.

비디오 내의 이동물체에 대한 분석에 있어서, 배경이 고정되어 있을 때는 이동물체에 대한 이동 벡터만이 양수값을 갖는다는 사실에 기반하여 이동물체 영역과 이동방향을 분석할 수 있으나, 배경이 함께 이동하는 경우에는 이동물체 뿐만 아니라 배경의 움직임도 고려해야 한다.

(4) 비디오 내의 문자분할 및 인식과 객체분할 및 인식, 음성인식을 이용한 색인방법

비디오는 영상 정보와 더불어 텍스트 정보도 포함하고 있기 때문에 비디오 내의 텍스트 정보는 비디오의 주제를 나타내는 중요한 정보가 될 수 있다. 따라서 이것을 찾아내어 추출하고 인식한 후 결과를 특징정보로 사용하여 인덱싱 한다면 고수준의 검색을 가능하게 할 수 있다. 또한 비디오 내에서 중심이 되는 객체는 비디오의 내용을 대표할 수 있기 때문에, 배경과 구분이 되는 객체를 추출하여 인식하는 비디오 내 객체분할 및 인식기술은 데이터베이스 구축 시 자동화된 색인과정을 위해 매우 중요하다.

정확한 비디오 색인을 위해서는 이상과 같은 텍스트 정보와 내용기반 특징정보들을 종합적으로 고려하여 색인하는 것이 바람직하다.

3. 웹 이미지 검색연구 현황

인터넷의 발전으로 WWW에 존재하는 문서의 수는 급속히 증가하고 있다. 웹에는 텍스트 이외에 이미지, 비디오, 오디오 등의 다양한 형태의 정보가 존재하고 있으며, 최근에는 이러한 웹 상에서의 이미지를 효율적으로 검색하기 위한 연구가 시도되고 있다.

웹 이미지 검색에 관한 연구는 일반적으로 이미지의 내용에 기반을 둔 검색과 이미지 주위의 텍스트를 이용한 검색을 지원하고 있다. 웹 이미지 검색엔진은 웹 이미지의 내용기반 검색을 위하여 이미지 검색에 관한 기존 연구내용을 응용하고 있으며, 텍스트를 이용한 이미지 검색을 위해서는 웹 문서 내의 HTML 문서에서 특정 위치의 텍스트를 활용하고 있다. 그러나 텍스트는 문맥에 따라 의미가 달라지는 문제점이 있으므로 텍스트를 이용하여 웹 이미지를 분류하고 질의시 대상 이미지의 분류를 명시함으로써 검색효율을 높이는 방안도 연구되고 있다.[28]

---

27) A. Akutsu et al., "Video Indexing Using Motion Vectors," *Proc. of SPIE-Visual Communications and Image Processing* '92, San Jose, CA, USA, Vol, 1818(1992), pp. 1522-1530.

4. 내용기반 영상정보 검색을 위한 질의

내용기반 영상정보 검색을 위한 일반적인 질의의 종류는 다음과 같이 다양하게 구성되어 있다.

1) 색상(color)

색상에 의한 질의는 사용자가 이미지를 구성하는 색상의 정도를 명세하여 그와 유사한 이미지를 검색한다.

2) 질감(texture)

질감에 의한 질의는 사용자가 이미지를 구성하는 질감의 정도를 명세하여 그와 유사한 이미지를 검색한다.

3) 형태(shape)

형태에 의한 질의는 이미지들의 형태에 대한 유사성을 측정하여 질의 이미지와 유사한 형태를 가지고 있는 이미지들을 검색하는 방법이다. 이 때의 사용자 질의는 두 가지 방법에 의해서 입력이 된다. 첫 번째는 검색시스템이 제공하는 그래픽 도구를 이용하여 검색하고자 하는 이미지와 비슷한 이미지를 스케치하여 입력하는 방법이다. 이 방법은 사용자가 원하는 검색대상의 형태를 직접 그려서 입력하는 방식이므로 질의의 융통성이 크다. 또한 이 방식의 장점은 사용자가 원하는 검색형태를 시각적으로 입력하므로 원하지 않는 형태의 검색결과가 나오는 비율이 감소한다는 장점이 있다. 반면에 이 방법은 사용자가 입력한 스케치는 수학적인 정의가 어렵고, 기존의 잘 알려진 수학적 방법으로는 사용자의 입력정보와 실제 영상의 매칭이 어렵다는 문제점이 있다. 사용자 질의의 두 번째 방법은 기존의 이미지 데이터베이스에 저장된 적당한 예제 이미지를 선택하여 입력하는 방법이다.

4) 볼륨(volume)

볼륨에 의한 질의는 3차원 이미지(3D image)의 모양에 의한 질의 방법이다.

5) 공간적인 제약(spatial constraints)

공간적 제약에 의한 질의는 한 이미지 안에 있는 객체들간의 공간적(spatial), 위상적(topological) 관계에 의한 질의방법으로 객체들간의 인접(adjacency), 겹침(overlap), 포함

28) Shian-Hua Lin et al. "Extracting Classification Knowledge of Internet Documents with Mining Therm Association: A Semantic Approach", *ACM SIGIR '98*, 1998.

(containment) 등과 같은 관계에 대한 질의가 가능하다.

6) 브라우징(browsing)

브라우징에 의한 질의는 사용자가 이미지 데이터베이스에 대하여 이용 가능한 정보의 종류를 전혀 모를 때 사용하는 질의의 종류이다. 브라우징에 의한 검색행위는 전체검색의 초기 검색영역에 대한 정보를 주어 필요가 없는 정보공간을 사전에 배제시킬 수 있게 한다.

특히, 비디오 데이터의 경우 스키마가 복잡하고 내용기반 검색능력의 부족으로 기존의 질의 방법으로는 원하는 비디오 데이터를 검색하기 어려우므로 이를 보완하는 방법으로 비디오 브라우징 기술이 쓰인다.

비디오 브라우징은 효과적인 검색을 가능하게 할 뿐 아니라 비디오 파싱이나 편집 시 비디오 응용에 공통적으로 사용되기도 하는데, 브라우징의 방법에는 대표 프레임을 나열하는 방식, 계층적 장면전환 그래프를 이용한 스토리 보드 구성방식, 계층적 비디오 브라우징 방식 등이 있다.

7) 객관적 속성(objective attributes)

객관적 속성에 의한 질의는 이미지를 획득한 날짜나 이미지 형식과 같은 이미지의 객관적인 정보에 의한 질의로 일반적인 상업용 데이터베이스 질의 언어와 연결이 가능한 질의방법이다.

8) 주관적 속성(subjective attributes)

주관적 속성에 의한 질의는 객관적 속성과는 반대로 각각의 사용자에 의해 다르게 해석될 수 있는 이미지의 주관적 속성에 의한 질의 방법이다.

9) 동작(motion)

동작에 의한 질의는 시공간(spatiotemporal)상에서 변하는 이미지의 주기(sequence)에 의한 질의 방법이다.

10) 텍스트(text)

텍스트에 의한 질의는 텍스트에 기반한 이미지 검색을 요구할 때 필요하다. 이 때의 텍스트는 이미지의 객관적 속성을 표시하는 짧은 종류의 텍스트가 아니라 이미지에 관한 여러 정보를 묘사하는 일종의 풀 텍스트로 자연언어 처리기법과 일반적인 정보검색기법을 사용하여 질의를 처리한다.

11) 도메인 개념(domain concept)

도메인 개념에 의한 질의는 복합 질의(complex queries)로, 앞에서 설명된 질의의 다양한 종류들은 복합 질의를 형성하기 위한 기본적 연산자로 사용될 수 있다.

이상에서 언급한 다양한 유형의 질의가 모든 응용에 모두 필요한 것은 아니다. 각각의 응용에 적합한 질의의 종류를 선택하거나 혼합하여 이용하는 것이 바람직하다.

Ⅲ. 내용기반 영상정보 검색을 위한 인덱싱 기술

내용기반 이미지 검색시스템은 크게 해당 이미지를 가장 정확하게 표현할 수 있는 특징을 추출하는 단계와 추출된 특징을 사용하여 효율적인 검색을 지원하는 인덱싱과 색인구조를 구축하는 단계로 구성된다.[29] 이 장에서는 추출된 멀티미디어 데이터의 검색을 지원하는 색인기법과 색인구조에 관해 살펴본다.

1. 내용기반 이미지 정보의 색인기법

대용량성과 비정형성을 특징으로 하는 방대한 양의 영상정보 데이터베이스에서의 내용기반 검색의 중요한 문제는 효율적이고 신속한 검색을 지원하기 위해서는 적절한 이미지 색인기법이 필수적이다. 내용기반 이미지 검색을 위한 전형적인 색인구조는 이미지 데이터로부터 고차원의 특징 데이터를 추출한 후, 추출된 특징 데이터를 하나의 차원으로 간주하여 이미지 데이터를 색인하는 방법을 주로 사용하고 있다. 이러한 이미지 색인구조에서는 수십 또는 수백 개의 특징들로 표현되는 이미지를 색인해야 하므로 구현상의 어려운 문제점들이 아직까지 존재하고 있는 실정이다.[30]

내용기반 이미지 검색을 위해 기존의 다차원 색인구조를 사용하는 방법에 대한 많은 연구들이 진행되어 왔다. 대표적인 예로는 KD-트리, KDB-트리, hB-트리, LSD트리, 다차원 선형 해싱, 다차원 신장 해싱, 그리드 파일, BANG-화일, 확장형 셀 기법, R-트리, R+-트리, R*-트리

29) Tzi-cker Chiueh, "Content-Based Image Indexing," *Proceedings of the 20th VLDB Conference*(1994), pp. 582-593.

30) 최길성, 『내용기반 이미지 검색을 위한 다수의 사용자를 지원하는 효율적인 고차원 색인구조』, 미간본박사학위논문, 충북대학교 대학원, 1999.

등이 있다. 그러나 이러한 다차원 색인기법들은 특징 데이터의 수가 증가함에 따라 검색시간이나 기억공간이 지수적으로 증가되어 고차원의 데이터에 대해서는 순차적으로 탐색하는 것보다 못한 결과를 초래하여 색인구조로서의 기능을 상실하게 된다. 따라서 기존에 제시된 다차원 색인기법들은 고차원의 이미지 데이터를 색인하기에는 적합한 색인구조라고 할 수 없다.

최근 이러한 제약을 극복하기 위해 고차원의 이미지 데이터를 저차원으로 변환하는 기법[31]을 사용하거나, 고차원 이미지 데이터를 그대로 수용할 수 있는 색인구조들이 제안[32]되고 있다. 저차원 변환 기법은 거리관계를 유지하는 변환을 이용하여 고차원 특징 벡터를 낮은 차원으로 사상한다. 이 기법의 문제점은 대량의 데이터를 대상으로 할 경우에는 저차원 변환의 한계가 발생한다는 것이다. 고차원 데이터 색인구조들의 대표적인 예로는, TV-트리, SS-트리, VAM-KD트리, VAM Split R-트리, X-트리, SR-트리 등이 제안되었다.

이러한 고차원 데이터의 색인구조들은 각기 다른 접근방법과 특징을 가지고 있기 때문에 특정 응용환경에 적용할 때는 그 특성을 잘 파악하여 적절한 색인구조를 선택하는 것이 바람직하다. 아울러 특정한 고차원 데이터의 색인구조를 내용기반 검색시스템에 적용하기 위해서는 다음과 같은 사항들을 고려해야 한다.[33]

1) 특징 벡터의 성질

특징 추출과정을 통해 생성된 특징 벡터를 사용할 경우, 특정 색인기법이 특징 벡터의 성질에 잘 맞는지 여부를 판단하는 것이 중요하다.

2) 유사성 측정 함수의 성질

내용기반 검색에서 두 이미지간의 유사성을 측정하기 위해 사용하는 함수의 성질을 잘 알아야 한다. 예를 들어, 유클리드 거리 함수는 대부분 유용하지만 모든 경우에 다 적용될 수는 없다. 또한 색상 히스토그램처럼 특징 벡터들 사이에 간섭이 존재한다면 이것을 반영하는 2의 지수승배 형식의 거리 함수(quadratic distance bounding)를 이용해야 한다. 이와 같이 내용기반 검색시스템에 색인기법을 적용할 때는 사용하는 유사성 측정 함수의 성질을 잘 파악해야 한다.

3) 사용하는 질의의 종류

---

31) C. Faloutsos et al., "Efficient and Effective Querying by image content," *Journal of Intelligent Information System*, Vol. 3, No. 3(July 1994), pp. 231-262.

32) D. A. White and R. Jain, "Similarity Indexing: Algorithms and Performance," *Proc. SPIE*, Vol. 2670(1996), San Diego, USA, pp. 62-73.

33) 이동호 외, "대규모 이미지 데이터베이스에서 고차원 색인구조를 이용한 효율적인 내용기반 검색시스템," 《정보과학회논문지(B)》, 제26권, 제1호(1999. 1), pp. 52-65.

내용기반 이미지 검색에서 사용하는 질의의 종류를 잘 파악하여 이에 적합한 색인구조를 이용해야 한다. 어떤 색인구조들은 특정 질의에 대한 성능을 향상시키는데 초점을 두어 개발하였기 때문에 다른 종류의 질의에 대해서는 비효율적일 수 있다.

한편, 온라인 상에서 다수의 사용자들이 동시에 색인구조를 통해서 데이터를 조작하는 경우, 이들에 대한 동시성을 제공해주는 동시성 제어기능이 필요하다. 트리구조에 대한 동시성 제어는 데이터의 삽입이나 삭제를 하기 위해 트리를 변경하거나, 데이터 검색을 위해 트리에 접근하는 다수의 프로세스들이 존재할 때 간섭이나 충돌을 최소화시킨다. 또한 삽입이나 삭제 연산에 의해 트리가 변경 중이거나 변경된 상태가 트리를 탐색하고 있는 다른 프로세스에게 항상 일관된 정보에 접근하도록 보장해준다. 기존의 동시성 제어기법으로는 B-트리를 기반으로 한 것들이 있으며, 이 중에서 Blink-트리 기법은 링크를 적용함으로써 동시성 제공 수준을 향상시킨 기법이다. 또 다른 동시성 제어기법으로는 R-트리 계열의 기법들이 있는데 그 중에서 Rlink-트리는 Blink-트리에서처럼 삽입연산이 진행되는 동안에도 링크를 이용하여 탐색연산이 가능하도록 하는 LSN(Logical Sequence Number)이라는 새로운 개념을 도입하여 R-트리에 대한 기존의 동시성 제어기법들 중에서는 가장 우수한 것으로 알려졌다.

다음은 내용기반 이미지 검색을 위한 색인구조가 가져야 할 설계 요건에 대해 살펴보고자 한다.

2. 내용기반 검색을 위한 색인구조의 설계 요건

내용기반 이미지 검색에서 방대한 양의 이미지 데이터를 내용기반으로 검색하기 위해서는 빠른 검색을 지원하는 새로운 방식의 저장구조 및 색인구조가 필요하다. 그런데 이러한 색인구조들은 내용기반 이미지 검색을 효율적으로 지원할 수 있도록 다음과 같은 설계 요건들을 갖추어야 한다.[34]

1) 고차원 특징의 효과적인 수용

내용기반 이미지 검색을 위한 색인구조는 이미지 자체가 포함하고 있는 고차원의 특징들을 효과적으로 수용할 수 있어야 한다. 기존의 다차원 색인구조는 색인에 사용되는 데이터의 차원이 증가함에 따라 노드 수가 지수적으로 증가되어 트리의 성능이 지수적으로 감소하여 검색의 효율이 떨어지는 단점이 있기 때문에, 내용기반 이미지 검색을 위해 설계되는 색인구조는 고차원의 특징을 효과적으로 수용할 수 있는 기능을 갖춤으로써 이러한 단점을 해결할 수

34) 최길성, op. cit.

있어야 한다.

2) 겹침의 최소화

겹침(overlap)은 하나 이상의 특징 벡터가 차지하는 영역을 두 개 이상의 중간노드에 중복시켜 정의하는 것을 말한다. 이것은 두 개 이상의 트리 경로를 따라 탐색을 행하게 하여 색인구조의 성능을 떨어뜨린다. 문제는 이러한 겹침영역의 발생이 고차원으로 갈수록 데이터 수가 증가되거나 트리의 높이가 높아짐으로써 많아진다는 것이다. 따라서 고차원 색인구조는 그 성능이 현저하게 감소된다. 그러므로 내용기반 이미지 검색을 지원하는 고차원 색인구조의 설계는 이러한 겹침을 최소화하는 형태로 트리를 구성하는 것이 필수적이다.

3) 유사성 검색의 지원

일반적으로 내용기반 이미지 검색에서는 하나의 이미지는 고차원 공간상의 하나의 점으로 대응시키고 이에 대한 특징 데이터들이 행렬이나 벡터 형태로 표현되기 때문에 완전일치 검색이 아닌 유사성 검색을 수행하여 사용자가 원하는 이미지와 유사한 이미지를 검색한다. 따라서 내용기반 이미지 검색에 사용되는 추출된 특징 데이터들은 유사성 검색을 지원하는 색인구조로 구성할 수 있도록 적절한 벡터 형식과 유사성 척도로 정규화(nomalize)하는 사전처리(preprocessing) 과정이 필요하다. 또한 고차원 색인구조의 성능은 데이터들의 분포에 많은 영향을 받으므로 유사한 데이터들끼리의 군집화(clustering)가 필요하다.

4) 유사성 척도 함수의 선택

하나의 이미지를 고차원 공간상의 하나의 점으로 대응시키고 서로 상이한 독립된 차원들에 대해서 일관된 거리 척도만을 적용하여 두 이미지간의 유사성 정도를 판단할 경우에는 정확성에 한계가 발생한다. 그러므로 유클리디안 거리 척도에 특징 데이터들의 상관관계나 가중치를 부여하는 함수를 적용시킬 필요가 있다. 또한 각 차원의 중요도에 따라 색인트리를 구축하는 경우에는 특징 데이터의 성질과 사용된 함수가 색인구조에 부합하는지를 판단할 수 있어야 한다.

5) 다양한 질의 형태에 대한 효율적인 검색의 지원

특정의 색인구조는 특정 질의 형태에 맞도록 설계될 수도 있으므로 내용기반 이미지 검색에 사용하는 질의 종류를 잘 파악하여 이에 적합한 색인구조로 설계되어야 한다. 사용자는 시각적 도구나 질의어 등의 다양한 형태로 질의하고, 색인된 모든 특징들을 질의 이미지로 제시

하거나 특징의 일부 또는 범위형태로 질의하는 경우가 있기 때문에, 이러한 상황에 대해 검색 결과의 정확성을 보장하고 신속한 검색을 지원하기 위한 효율적인 이미지 색인구조의 설계가 필요하다.

6) 동적인 데이터 조작의 용이

색인구조는 내용기반 이미지 검색의 목적에 적합하도록 동적인 환경을 제공할 수 있어야 한다. 즉, 색인될 데이터들의 분포형태나 사전정보와 무관하게 색인구조가 생성될 수 있어야 하며, 특징 벡터들의 삽입이나 삭제가 자유롭게 수행될 수 있어야 한다. 특히, 색인구조는 데이터의 삽입이나 삭제연산을 수행하는 경우에는 연산의 영향을 덜 받는 범위 내에서 조작될 수 있도록 설계하는 것이 바람직하다.

Ⅳ. 내용기반 영상정보 검색시스템 현황

1. 내용기반 이미지 검색시스템

정지 영상을 대상으로 하는 내용기반 이미지 검색시스템은 이미 상용화되어 인터넷을 통하여 서비스가 제공되고 있다. 다음은 현재 WWW 등을 통하여 서비스가 제공되고 있는 대표적인 내용기반 이미지 검색시스템들이다.

(1) QBIC[35](http://wwwqbic.almaden.ibm.com/~qbic/qbic. html)

IBM의 Almaden 연구소에서 개발된 이미지 검색시스템으로 내용기반 이미지 검색뿐만 아니라 제한된 범위 내에서의 비디오 검색이 가능한 시스템으로 현재 여러 면에서 가장 뛰어난 검색시스템으로 평가받고 있다. 이 시스템은 사용자에게 색상이나 질감, 형태 등과 같은 기본적인 속성에 기반한 시각적 질의를 제공하여 검색의 폭을 확대시켰다. 아울러, 데이터베이스에 저장된 이미지에 대하여 사람이 의미정보를 부여할 수 있도록 허용함으로써 제한적이나마 의미정보에 의한 검색도 가능하다. 그러나, 복잡한 질의 인터페이스가 체계적으로 통합되어 있지 않아서 사용자 측면에서는 이용하기가 어려울 수 있다는 단점이 있다. 또한 특징 데

35) W. Niblack et al., 'The QBIC Project: Querying Image by Content using Color, Texture, and Shape," *Proc. SPIE Storage and Retrieval for Image and Video Databases*(Feb. 1993), pp. 173-187.

이터를 완전 자동으로 추출하지 못하고 부분적으로 사람의 수작업을 이용해야 한다는 단점도 있다.

(2) Virage(http://www.virage.com)

Virage사의 Virage 시스템은 시각정보에 대해 원시 이미지, 처리 이미지, 이미지 특징, 비디오 이벤트 등의 4가지 계층구조로 구성하고 색상의 배열이나 윤곽선, 질감, 형태 등의 특징을 추출하여 검색에 이용한다. 또한 검색을 위해 특징의 일부에 가중치를 설정하여 유사도를 계산하고 있으며, 다양한 형태의 이미지 검색을 할 수 있도록 지원하는 특징이 있다.

(3) MetaSEEK(http://www.ctr.columbia.edu/MetaSEEK)

탐색엔진은 급속히 발전하는 WWW상에서 정보를 찾기 위한 매우 강력한 자원이다. 그런데, 원하는 탐색엔진을 찾고 그것들을 어떻게 이용하는지를 학습하는 것은 아주 많은 시간낭비일 수 있다. 그러한 탐색도구들의 통합은 이용자들이 명백하고 효율적인 방법으로 전세계의 정보에 접근할 수 있게 한다. 이러한 시스템들을 메타 탐색엔진이라고 부른다. 최근 웹상의 시각적 정보 탐색엔진들의 출현은 동일한 효율 문제를 이끌어내고 있다.

MetaSEEK는 시각적 정보에 기반한 웹상의 이미지들을 찾기 위해 사용되는 내용기반 메타 탐색엔진이다. 이 시스템은 상이한 계층들의 이용자 질의들의 성능을 순위매기는 것에 의해 다수의 온라인 이미지 탐색엔진들을 지능적으로 선택하고 인터페이스하기 위해 설계되었다. 이용자 피드백이 순위매김 개선에 또한 통합되었다.

(4) Image Surfer[36](http://isurf.interpix.com)

Interpix Software사의 이미지 검색시스템으로 키워드 기반의 분류 트리(category tree)를 생성하고 이를 이용하여 웹 이미지를 분류하고 검색한다. Yahoo의 이미지 검색시스템으로 사용되고 있다.

(5) WebSEEK[37](http://www.ctr.columbia.edu/webseek)

Columbia 대학에서 개발한 WebSEEK는 Image Surfer와 마찬가지로 이 시스템은 WWW에서의 이미지와 비디오를 분류하고 탐색할 수 있다. 이미지와 관련된 텍스트를 이용하여 반자동으로 이미지를 분류한다. 이 시스템을 통한 탐색은 이미지 토픽으로부터의 선택에 의해서,

36) Yahoo's Image Surfer. http://ipix.yahoo.co.

37) J. R. smith, S. -F. Chang, "An Image and Video Search Engine for the World-Wide-Web", *Symposium on Electronic Imaging; Science and Technology - Storage & Retrieval for Image and Video Databases V*, San jose, CA.,(Fef. 1997).

혹은 텍스트나 이미지 특징을 사용하는 것에 의해서 이루어진다. 이 시스템을 이용하면 특정 분류 안에 들어있는 WebSEEK내의 이미지, 웹 상의 다른 이미지, 색상 선택, 색상 히스토그램, 스케치, 이미지 대중성, 이미지 속성, 해제, 카테고리 기반 검색을 할 수 있다. 현재 이 시스템에는 웹으로부터 추출된 650,000건의 이미지와 10,000건 이상의 비디오가 분류되어 있다.

(6) WebSeer[38](http://infolab.cs.uchicago.edu/webseer)

Chicago 대학에서 개발한 WebSeer 시스템은 이미지가 포함된 웹 문서에서 웹 이미지 내용을 분석하여 얻어진 이미지 헤더, 파일 종류, 크기, 날짜 등과 이미지 주위의 텍스트 정보를 이용하여 웹 이미지를 검색한다. 또한 이 시스템은 사람의 얼굴과 수평선 등의 객체를 이미지에서 자동으로 인식하고 색상과 질감 등의 정보를 사용하여 하나의 이미지를 여러 조각으로 분할한다.

(7) ImageRover[39](http://cs-www.bu.edu/groups/ivc/ImageRover/Home.html)

ImageRover는 WWW의 이미지 데이터베이스의 내용기반 탐색을 위한 단일 인덱스로 텍스트와 시각적 통계를 결합하여 사용하고 있다. 텍스트적 통계는 HTML 문서에 포함된 텍스트에 기반한 잠재적 의미 색인(latent semantic indexing)을 사용하여 벡터형태로 포착된다. 시각적 통계는 색상 및 정위(orientation) 히스토그램을 사용하여 벡터형태로 포착된다. ImageRover 색인에 대한 탐색을 시작하기 위해서는, 이용자는 기대하는 이미지들을 기술하는 소수의 키워드들을 명기한다. 그 다음에는 이용자는 적합성 피드백을 통하여 이 초기의 질의를 개선한다. 이 시스템은 WWW 상에서의 이미지의 수집, 해석을 담당하는 이미지 수집 서브시스템과 질의 서버와 사용자 인터페이스로 이루어진 이미지 질의 서브시스템으로 구성되어 있다.

(8) Excalibur Visual Retrieval Ware (http://www.eccalib.com/rev2/demos/vrw/cstdemo - intro.html)

Excalibur사에서 개발한 상용 이미지 검색 시스템으로 지문, 얼굴, 그리고 문자 인식에 활용할 수 있다. 이외에 비디오, 사진, 그리고 다른 유형의 이미지 데이터를 위한 구성요소도 다루고 있다.

---

38) Charles frankel, Michael J. Swain, Vassilis Athisos, "WebSeer: An Image Search Engine for the World Wide Web", TR 96-14(1996), U. Chicago.

39) Stan Sclaroff, Leonid Taycher, Macro La cascia, "ImageRover: A Content-Based Image Browser for the World Wide Web", Proc. *IEEE Workshop on Content-based Access of Image and Video Libraries*(1997).

(9) Photobook(http://www-white.media.edu/~tpminka/photobook/)

이 시스템은 MIT에서 개발한 내용기반 이미지 검색시스템으로 이미지 자체가 아니라 이미지들과 연관된 특징들을 비교하는 것에 의해 작업을 수행한다. Photobook은 질감 모델링, 얼굴 인식, 형태 매칭, 브레인 매칭, 상호작용적 분할 및 주석에 이용된다. 이 중에서 Faces는 얼굴 이미지 검색시스템으로 이 시스템은 약3,000여명의 7,562개의 얼굴 이미지 데이터베이스에 대하여 90% 이상의 정확한 검색율을 보여주고 있다.

(10) QVE[40]

Hirata와 Kato가 만든 시스템으로, 처음으로 시각적 질의를 통한 내용기반 검색을 제안한 시스템이다. 이미지의 외곽선 데이터를 그대로를 특징 데이터로 사용하기 때문에 실제 질의 처리 및 검색과정에서 유사성을 검사하기 위해서 데이터베이스에 저장된 각각의 이미지에 대해 이동(translation)이나 스케일링(scaling), 회전(rotating)과 같은 기하학적 변환을 일일이 고려해야 한다는 단점이 있다.

(11) VisualSEEK[41](http://disney.ctr.columbia.edu/VisualSEEK/)

WWW를 이용한 내용기반 이미지 검색시스템으로 Columbia 대학에서 개발되었다. 이 시스템의 가장 큰 특징은 시각적 특징에 의한 내용기반 질의(content-based querying)와 공간정보에 의한 공간질의(spatial querying)가 동시에 가능하다는 것이다. 이 시스템에서는 이미지에 대한 색상 영역에 관한 정보를 추출하고 이를 이용하여 각 색상 영역들간의 공간적인 관계를 추론함으로써 공간관계에 대한 특징을 추출하여 질의에 사용한다. 그러나 이러한 추론과정은 시간이 많이 소요되며, 색상정보에 의존한 검색만 가능하다는 단점을 가진다.

(12) Chabot(http://http.cs.berkeley.edu/~ginger/chabot.html)

UC Berkeley에서 개발한 Chabot[42]는 POSTGRES라는 자료 저장 시스템 위에 B-트리와 다차원 색인기법 중의 하나인 R-트리를 이용하여 단순 속성, 객체에 대한 설명정보(캡션 정보)를 이용한 텍스트기반 검색과 히스토그램을 구한 다음 색상에 대해 적절한 의미정보를 기술하여 검색을 수행한다. 이 시스템은 의미정보를 이용하므로 기존의 질의 처리 언어와 각종 응용 프로그램을 사용할 수 있다는 장점은 있으나 색상에 의한 검색만 가능하다는 단점이 있다.

---

40) K. Hirata and T. Kato, "Query by Visual Example-Content Based Image Retrieval," *Advances in Database Technology(EDBT '92)*(1992), pp. 56-71.

41) J. R. Smith and S.-F. Chang, "VisualSEEK: A Fully Automated Content-based Image Query System," *ACM Multimedia 96*, Boston, MA(1996).

42) V. E. Ogle, M. Stonebreaker, "Chabot: Retrieval from a Relation Database of Images," *IEEE Computer*, Vol. 28, No. 9(Sept. 1995), pp. 25-41.

이 시스템은 나중에 온라인 브라우징을 지원하고 내용으로 이미지를 검색할 수 있도록 확장되었다.

(13) SaFe(http://disney.ctr.columbia.edu/SaFe)

SaFe는 공간 및 특징 이미지 탐색을 위한 일반적인 시스템이다. 이 시스템은 영역이나 객체들의 공간적 배열에 의해 탐색과 이미지 비교를 위한 프레임웍을 제공한다. 이 시스템 내에서 질의, 객체 혹은 영역들은 이용자들에 의해 할당된다. 이러한 것들은 공간적 위치, 크기 그리고 색상 같은 시각적 특징에 대한 주어진 성질들이다. SaFe는 질의에 가장 정확히 매치되는 이미지를 찾아내며, 영역/특징 추출과 인덱싱을 위한 완벽히 자동화된 도구들을 사용한다.

(14) CORE(Content-based Retrieval Engine)

싱가폴 국립대학에서 개발한 내용기반 검색엔진 CORE[43]는 내용기반 검색기법과 이미지의 특징에 근거한 색인기법을 이용하여 멀티미디어 자료를 검색할 수 있으며, 다양한 응용 프로그램 개발을 위해 다수의 API(Application Program Interface)를 제공한다. 또한 가장 큰 특징은 사용자 질의의 재생성을 위한 사용자 질의 피드백을 지원하며, 키워드, 색상, 형태, 질감 등 이미지의 특징 벡터를 통한 시각 브라우징, 유사성 검색, 퍼지 검색, 시소러스 기반 텍스트 검색과 같은 다양한 검색기법을 제공한다.

(15) CEDAR(The Center of Excellence for Document Analysis and Recognition)

뉴욕 주립대학에서 개발한 CEDAR[44] 시스템은 부제목이 첨가된 인물 사진들에 대해서 인덱싱과 데이터 검색을 필요로 하는 응용분야를 위해 개발되었다.

2. 내용기반 동영상 검색시스템

(1) AVIS(Advanced Video Information System)

Maryland 대학에서 개발된 비디오 데이터베이스 프로토타입인 AVIS 시스템은 텍스트 정보 검색의 역 리스트와 공간 데이터 검색의 세그먼트 트리(Segment tree)를 변형하여 주인공의 이름이나 사건의 이름을 검색조건으로 하여 관련 비디오 프레임을 탐색하도록 해준다.

---

43) J. K. Wu et al., "CORE: a Content-based Retrieval Engine for Multimedia Information Systems," *ACM Multimedia System*, Vol. 3, No. 1(1995), pp. 25-41.

44) R. K. Srihari, "Automatic Indexing and Content-based Retrieval of Captioned Images," *IEEE Computer*, Vol. 1, No. 9(Sept. 1995), pp. 49-56.

(2) QBIC(Query BY Image Content, http://wwwqbic.almaden.ibm.com)

IBM사의 QBIC는 이미지 검색기능을 비디오 데이터의 검색에도 활용하고 있다. QBIC에서는 비디오 프레임간의 급격한 차이를 이용하여 샷을 자동으로 검출하고, 각 샷에 대해 대표 프레임을 선정해, 이 대표 프레임을 이미지 검색기술로 검색해 요청된 장면을 검색한다.

(3) Informedia(http://informedia.cs.cmu.edu/html/body-whitepage.html)

Carnegie Mellon 대학에서 개발한 Informedia는 비디오 내의 텍스트 정보와 객체인식, 오디오의 내용인식 등 고수준의 특징값을 사용한 비디오 색인기법을 사용함으로써 비디오 데이터베이스 구축 및 검색의 효율성을 높였다.

(4) SWIM System(http://www.iss.nus.sg/RND/MS/Projects/vc/project1.html)

싱가폴의 ISS(Institute of Systems Science)에서 개발하였다.

(5) WebClip(http://www.ctr.columbia.edu/webclip)

Columbia 대학에서 개발한 WWW 상에서의 압축된 비디오를 편집하고 브라우징하기 위한 분산 시스템이다. WebClip은 Columbia 대학에서 개발한 압축 비디오 편집, 파싱, 및 탐색(Compressed Video Editing, Parsing, and Search: CVEPS) 기술위에 구축된 웹 응용이다. 이 시스템의 특징은 압축된 도메인 비디오 편집, 내용기반의 비디오 검색, 다중 해상도 접근, 그리고 분산 네트워크 구조 등이다. 편집엔진과 탐색엔진은 특수효과를 번역하고 이용자들에 의해 표출된 시각적 질의를 처리하는 프로그램을 포함하고 있다.

(6) OVID(Object-oriented Video Information Database)

Kyoto Sangyo 대학의 Oomoto와 Kobe 대학의 Tanaka가 개발한 비디오 데이터베이스로 단일 연속 비디오 프레임 시이퀀스로 구성된 비디오 객체들, 하나 이상의 연속 비디오 프레임 시이퀀스로 구성된 비디오 객체들, 그리고 연속 여부에 상관없이 모든 유형의 비디오 객체들로 구분하여 검색한다.

(7) VideoQ(http://www.ctr.columbia.edu/VideoQ)

Columbia 대학에서 개발한 자동 객체지향 내용 기반의 비디오 탐색시스템이다. VideoQ는 풍부한 시각적 특징과 시공간적인 환경에 기반한 비디오 객체를 탐색할 수 있게 하는 새로운 탐색기술로서 전통적인 탐색방법(예, 키워드 혹은 주제별 탐색)을 확장시키고 있다. 이 시스템의 특징으로는 첫째 자동적인 독단적 형태 비디오 객체 분할과 트래킹, 둘째 색상, 질감, 형

태, 그리고 동작을 포함한 풍부한 시각적 특징 라이브러리, 셋째 다중 객체 시공간적 비디오 질의(절대 및 상대적 위치) 등이다. 이 시스템은 현재 대규모의 디지털 데이터베이스를 지원하고 있으며, 현재 약 2,000건의 샷이 축적되어 있다. 스케치에 의한 질의에 추가하여, 이용자는 텍스트에 의해 비디오 샷들을 브라우즈하거나 비디오를 탐색할 수 있다.

Ⅴ. 결론 및 제언

이상에서 최근까지의 멀티미디어 영상정보의 검색에 관한 여러 연구들 중 내용기반의 영상정보 검색방법에 관한 연구들을 살펴보고, 이들 연구에서 채용하고 있는 영상정보에 대한 다양한 특징정보의 추출과 인덱싱 기술에 대해 고찰해 보았다. 아울러 영상정보의 효과적인 검색을 위해 채용되고 있는 색인구조의 설계요건에 대해서도 살펴보았으며. 마지막으로 이러한 내용기반 영상정보 검색기술을 적용한 다양한 실제 시스템을 소개하였다.

내용기반 이미지 정보검색을 위한 특징정보를 추출하기 위해서는 여러 가지의 속성이 사용되고 있으며, 이 중에서 특히 색상, 질감, 형태 그리고 공간관계에 의한 이미지 정보검색이 주로 연구되고 있다. 그러나 이러한 검색방법들은 영상의 일부 속성만을 다루고 있기 때문에 한 가지의 방법만으로는 사용자의 다양한 요구에 부응할 수 없다는 단점이 있다. 특히 대용량의 영상 데이터베이스에 적용시킬 때에는 검색효율의 저하가 문제점으로 대두된다. 그러므로 검색효율의 향상을 위해서 여러 가지 속성들이 복합적으로 사용되고 있으며, 아울러 텍스트기반의 검색방법이 병행되고 있기도 하다. 그렇지만 진정한 의미의 내용기반 영상검색시스템의 개발을 위해서는 영상 내의 객체들에 대한 인식기술과 고수준의 사용자의 질의로부터 적용이 가능한 적절한 검색조건을 생성할 수 있는 기술이 필요하다.

내용기반 동영상 검색연구는 동영상의 내용을 대표할 수 있는 특징을 추출하여 이를 색인하기 위한 동영상 파싱, 대용량의 동영상 데이터와 이와 관련된 메타 데이터를 효율적으로 저장하기 위한 동영상 데이터 압축 및 저장, 그리고 사용자가 동영상 검색을 용이하게 할 수 있는 환경을 제공하는 검색 및 브라우징 등을 핵심기술로 하고 있다. 여기에서 비디오 파싱은 대용량의 비디오 데이터를 효율적으로 저장, 관리하고 사용자가 쉽게 검색할 수 있는 환경의 제공을 목적으로 하고 있다. 이러한 비디오 파싱은 대용량 비디오를 보다 적은 용량의 요약형 비디오로 만들기 위한 비디오 분할과 비디오의 내용을 효과적으로 표현하기 위한 비디오 색인과정으로 구성된다.

내용기반 이미지 검색시스템은 크게 해당 이미지를 가장 정확하게 표현할 수 있는 특징을 추출하는 단계와 추출된 특징을 사용하여 효율적인 검색을 지원하는 인덱싱과 색인구조를 구축하는 단계로 구성된다. 대용량성과 비정형성을 특징으로 하는 방대한 양의 영상정보 데이터베이스에서의 내용기반 검색의 중요한 문제는 효율적이고 신속한 검색을 지원하기 위해서는 적절한 이미지 색인기법이 필수적이다. 또한 방대한 양의 이미지 데이터를 내용기반으로 검색하기 위해서는 빠른 검색을 지원하는 새로운 방식의 저장구조 및 색인구조가 필요하다. 이러한 색인구조들은 내용기반 이미지 검색을 효율적으로 지원할 수 있도록 고차원 특징의 효과적인 수용과 겹침의 최소화 할 수 있어야 하며, 유사성 검색의 지원, 유사성 척도 함수의 선택, 다양한 질의 형태에 대한 효율적인 검색의 지원, 동적인 데이터 조작의 용이성 등의 설계 요건이 필요하다.

이러한 영상정보 검색 기술들은 다양한 영상정보 검색시스템에 적용되어 특정한 영상정보를 검색하는데 활용되고 있다. 특히 디지털 도서관의 급속한 발전은 정보서비스에 이러한 영상정보 검색시스템의 적극적인 활용을 필요로 한다.

아울러 보다 효과적인 영상정보 검색을 위해서는 사용자의 직관에 적합한 개념객체의 도입이 필요하며, 이를 위해서 이미지 정보를 위한 시소러스에 대한 연구가 필요하다. 또한 이용자를 고려한 감성 공학 측면의 영상정보 검색시스템에 대한 연구와 객체지향 기법을 이용한 시스템의 개발도 요구된다.

〈참고문헌은 각주로 대신함〉