

이미지의 영역별 상대성을 이용한 영상 검색

신광성^o 서덕원 유기형 김대중 김미화 유강수* 곽훈성

전북대학교 컴퓨터공학과, 영상공학과*

{Waver^o duck0419, ghyoo, kdjid, kmh8475 ,mickey*, hskwak}@chonbuk.ac.kr

Image retrieval using of relatively similarity by region of image

Kwangsung Shin^o Duckwon Seo Gihyoung Yoo Daejung Kim Mihwa Kim Kangsoo Yoo*

Kwak Hoonsung

Department of { Computer, Image* } Engineering, Chonbuk University

요 약

MPEG-7은 형태나 질감, 컬러 등의 속성 및 분포 등을 나타내기 위한 여러 가지 기술자(descriptor)들을 제공함으로써 검색의 효율을 높일 수 있는 방법을 포함한 표준화를 제시하고 있다.

본 논문에서는 영상의 특징을 MPEG-7의 여러 기술자중 에지 히스토그램 기술자(Edge Histogram Descriptor)를 통해 얻는다. 추출된 에지 특징을 이용하여 영상의 정해진 각 영역의 전체 후보영상과의 상대적 유사도를 비교하여 검색에 반영함으로써 검색의 효율을 향상시키는 방법을 제안한다.

1. 서 론

최근 급속한 정보화의 발전과 WWW으로 대변되는 컴퓨터 네트워크의 발전에 따라 디지털로 표현된 영상 데이터의 양이 기하급수적으로 증가하고 있으며, 그에 따라 이러한 영상 데이터의 생성, 전송, 가공이 매우 용이해졌다. 따라서 다양한 형태의 영상 데이터로부터 사용자가 원하는 내용의 정보를 효율적으로 찾아내고 그 정보를 사용자가 원하는 형태로 변형, 또는 전송할 수 있도록 하는 기술의 필요성이 대두되고 있다.

영상을 분류하고 검색하는 방법은 현재까지 많이 연구되어져 왔으며, 크게 다음과 같은 두 가지 방법으로 분류할 수 있다. 첫 번째, 텍스트 기반 검색 기법을 들 수 있다. 이 방법은 대상이 되는 모든 영상 데이터에 대하여 사람이 문자로 직접 색인을 첨가하고, 사용자는 주제어(Keyword)를 통하여 원하는 정보를 검색하는 기법이다. 텍스트 기반 검색 기법은 제한된 범위 내에서는 정확도가 높다는 장점을 가지고 있지만 방대한 영상 데이터에 대해 사람이 일일이 주제를 부여하기 어렵고 색인에 따른 비용이 너무 많아지며, 또한 사용자와 색인을 첨가하는 사람과의 관점이 일치하지 않는 경우 검색의 효율성이 크게 떨어지고 영상 데이터가 복잡한 속성을 가지고 있으면 텍스트만으로는 정확하게 표현할 수 없다는 여러 가지 단점을 가지고 있다.

두 번째, 내용 기반 검색 기법은 텍스트 기반 검색 기법의 단점을 극복하기 위하여 영상 데이터의 내용을 대표할 수 있는 특징을 추출하여, 이를 기반으로 색인과 검색하는 방법이다[1]. 이 방법은 영상 데이터의 내용으로부터 표현되는 색상(color), 무늬, 질감, 윤곽선, 형태, 공간 관계 등과 같은 각각의 영상에 고유한 특징을 추출하여 색인 과정에 사용함으로써 데이터 베이스 구축에 필요한 시간 및 인력의 소모를 줄인다는 장점을 가지고

있으나, 영상 데이터로부터 정확한 내용을 추출하기가 어렵다는 단점을 가지고 있다.

이러한 내용기반 검색의 단점을 극복하기 위한 노력으로 최근에는 멀티미디어 데이터의 효율적 관리 및 검색을 위한 MPEG-7의 표준화가 진행되었다[2].

MPEG-7은 형태나 질감, 컬러 등의 속성 및 분포 등을 나타내기 위한 여러 가지 기술자(descriptor)들을 제공함으로써 검색의 효율을 높일 수 있는 방법을 포함한 표준화를 제시하고 있다.

본 논문에서는 영상의 특징을 MPEG-7의 여러 기술자중 에지 히스토그램 기술자(Edge Histogram Descriptor)[3]를 통해 얻는다. 추출된 에지 특징을 이용하여 영상의 정해진 각 영역의 전체 후보영상과의 상대적 유사도를 비교하여 검색에 반영함으로써 검색의 효율을 향상시키는 방법을 제안한다.

2. MPEG-7 개요

2.1 MPEG-7 정지 영상을 위한 기술자

MPEG-7 Visual Part에서는 영상을 기술하기 위한 방법으로 기술자(descriptor)를 제공한다. 색상을 기술하기 위한 기술자로는 Color space, Color structure, Dominant color, Color quantization, GoF/GoP color, Color layout, Scalable color가 있다. 질감을 기술하기 위한 것으로는 Edge histogram, Texture browsing, Homogeneous Texture의 3가지가 있다. 이중 edge histogram은 본 논문에서 영상 검색을 위해 사용된 것으로 방향성 에지 4종류와 비방향성 에지 하나를 통해 영상이 갖는 에지 공간적인 분포를 표현한다. 에지 히스토그램 기술자는 영상이 가지고 있는 에지 분포를 추출하기 때문에 사용자가 생각하는 주관적인 물체에 입각한 자연 영상의 내용기반 이미지 검색이 가능하다.

2.2 에지 히스토그램 기술자(Edge Histogram Descriptor)

에지 히스토그램(Edge Histogram) 기술자는 이미지가 가지고 있는 에지 정보를 추출하는 방법을 이용한 MPEG-7 기술자의 한 종류로서 블록 단위의 에지 특징 추출과 분할된 영상의 위치적 특성을 고려한 에지 히스토그램을 생성한다. 에지 히스토그램 기술자는 다음과 같이 3가지 의미를 갖는 에지 분포의 히스토그램 빈(bin)으로 히스토그램의 특징 정보를 구성한다. 첫 번째는 영상공간 내의 전체적인(global) 에지의 분포를 나타내는 빈이고, 두 번째는 일정 크기의 부분 영상(sub-image) 내 에지 분포를 나타내는 국부(local)특성의 빈, 세 번째는 에지 혹은 객체의 윤곽선의 수평 및 수직 이동에 대해 강인성을 주기 위해 일정 열(column)과 행(row) 그리고 특정 부분을 포함하는 부분 영역(semi-global) 에지 분포를 나타내는 빈으로 구성되어 있다. 이 세 가지 특징 정보의 추출은 픽셀단위가 아닌 블록을 단위로 하여 에지를 검출한다.[4]

여기서 영상의 에지 정보는 아래와 같이 다섯 가지의 에지 타입, 즉 수직 에지, 수평 에지, 45° 에지, 135° 에지 그리고 방향성 없는 에지 성분으로 구분할 수 있다.

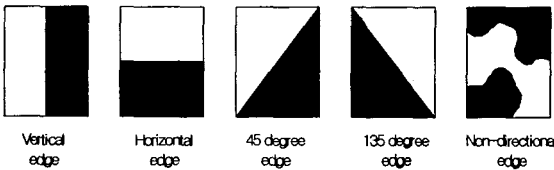


그림 1-1. 에지의 5종류

3. 제안한 영상 검색 방법

내용기반 이미지 검색은 텍스트 기반의 이미지 검색과 달리 식별 가능한 이미지에 대해 자동으로 특징 정보를 추출해야하고 추출된 특징 정보는 질의 처리과정에서 효율적인 검색이 이루어지도록 색인되어야 한다.

본 논문에서는 에지 히스토그램 디스크립터를 이용하여 영상 전체에 대해 로컬, 세미 글로벌, 글로벌 영역에 대하여 상세한 에지 성분을 검출하는 것으로 영상의 특성을 대표하는 특징 정보로써 규정했다. 영상의 특징 추출을 위해 에지 히스토그램을 이용한 이유는, 영상에서 에지 정보는 해당 영상의 전체적인 형태를 잘 나타내어 주는 요소이기 때문이다. 또한, 에지 히스토그램 디스크립터는 로컬 영역의 에지만을 검출하여 글로벌, 세미글로벌 영역의 에지를 계산해 낼 수 있는 장점이 있으므로, 영상을 기술하여 데이터베이스화하는 시간을 크게 줄일 수 있다는 이점이 있다. 영상에서 검출된 에지 정보를 검색에 사용하기 위해서 거쳐야 할 몇 가지 절차에 관해 알아본다.

3. 1. 히스토그램 정규화

히스토그램을 정규화 하는 방법은 여러 가지가 있을 수 있지만, 본 논문에서는 다음과 같은 정규화 과정을 거침으로써 영상 전체의 에지 량과 서브 이미지내의 에지 량을 상대적(relatively)으로 평가할 수 있다. 첫 번째로 로컬 에지 히스토그램의 인덱스(i)가 0~79까지라고 할 때 각 히스토그램 빈($local_edge(i)$)의 값을 각 영역(local, global, semi-global)의 각각의 빈이 가지고 있는 상세한 에지로 판정되어진 에지 인덱스 값을 합산한 값($normalized_sum$)을 이용해 정규화 하여 $feature_local(i)$ 의 값을 얻는다. 두 번째, 글로벌 에지 히스토그램의 정규화 과정도 동일하며 인덱스(i)가 0~4라 할 때 정규화 식은 (3-1)과 같다.

$$feature_global(i) = \frac{global_edge(i)}{normalized_sum}, 0 \leq i < 5 \quad (3-1)$$

세 번째, 세미글로벌 에지 히스토그램도 식 (3-2)에 의해 정규화 한다. 인덱스(i)는 0~64이다.

$$feature_semi_global(i) = \frac{semi_global_edge(i)}{normalized_sum}, 0 \leq i < 65 \quad (3-2)$$

3. 2. 상관도

영상의 상관도를 이용해서 사용자가 알고 있는 영상을 이용해 알지 못하는 영상들 중에서 가장 유사한 영상을 찾을 수 있다. 가장 유사한 영상은 상관도의 값이 가장 큰 영상을 선택함으로써 얻을 수 있다.

3. 3. 유사도 판정

3. 3. 1. 유클리디안 거리(Euclidean Distance)

예제 영상과 후보 영상의 각 픽셀간의 차분을 계산하는 Euclidean Distance는 식 (3-3)에서와 같이 정의된다.

$$D = \sum_x \sum_y |f(x,y) - w(x,y)| \quad (3-3)$$

3. 3. 2. 히스토그램 유사도 판정

본 논문에서는 비교 대상이 픽셀이 아닌 로컬, 글로벌, 세미글로벌 에지 히스토그램의 정규화 된, 총 150개의 빈을 서로 비교해야할 대상으로 정했으므로 유사도의 판정은 식 (3-4)과 같이 볼 수 있다[6][9]. 질의 영상의 정규화된 빈의 값과 후보 영상의 빈의 값을 비교하여 차이 값을 누적시키는 방법으로 두 영상 사이의 거리 값을 얻을 수 있다.

$$D = \sum_{i=0}^{149} |f(i) - w(i)| \quad (3-4)$$

3. 4. 영역별 상대적 거리 값 산출

식 3-4에 의해 거리 값을 비교한 유사도는 각 영역의 특성을 고려하지 않은 모든 영역에 대해 동일하게 차분 값을 누적하는 방법이기 때문에 각 영역의 상대적인 (relatively) 유사성을 반영하지 못한다는 단점을 가진다.

전체 영상의 각 영역인 로컬, 글로벌, 세미글로벌 에지 히스토그램의 유사도가 각각 질의 영상과 후보 영상의 차이를 비교하는 방법에 의해 거리 값이 결정되므로, 데이터베이스 내의 전체 영상의 각 영역별 상대적인 거리 값과는 각각 다를 수 있다. 각 영역마다의 거리 값을 따로 구한 후 데이터베이스 내의 모든 영상(w)에 대해 영역별 가장 큰 거리 값을 가진 영역을 취해 그 값으로 각각의 빈을 나눔으로써 모든 영상의 각 영역별 거리 값을 재조정한다.

3. 5. 최종적인 유사도 판정

결과 영상을 검색하기 위한 유사도의 판정은 식 (3-5)와 같다.

$$D(f,w) = \frac{D_{local}(w) + D_{global}(w) + D_{semi_global}(w)}{3} \quad ,w=1, \dots, n$$

$D_{local}(w)$, $D_{global}(w)$, $D_{semi_global}(w)$ 은 각각 식 (3-5)에서 얻어진 후보영상 w의 재조정된 각 영역의 거리 값이다.

결과 영상은 $D(f,w)$ 가 낮은 것부터가 가장 질의 영상과 유사한 영상이므로 가장 낮은 영상으로부터 M개만큼의 영상을 결과 영상으로 보여준다.

4. 결과

본 논문에서는 영역별로 특징이 에지 성분이 추출된 특징 정보를 이용해 영역별 유사성을 고려한 거리 비교 방법을 제안하였다. 에지 히스토그램에 의한 특징 추출과 검색은 매우 간단하면서도 작은 특징 벡터를 통해서 영상을 효율적으로 검색할 수 있다는 장점이 있다. 영상의 각 영역을 고루 이용함으로써 검색의 정확성을 높일 수 있으나, 본 논문에서 제안한 방법을 통해 각 영상의 영역별 상대적 유사성을 고려하여 검색한다면 더 좋은 검색 효율을 얻을 수 있다. 특히, 다양한 변화가 있는 사물의 영상이나 연속된 영상의 경우 각 영역의 상대적 유사성을 비교함으로써 영역에서의 변화에 좀더 민감하게 반응할 수 있다.

실험을 통하여 본 논문에서 제안한 방법이 기존의 방법에 비해 우수한 검색 효율을 나타냄을 증명하였다. 에지 추출을 통해 얻어진 특징 정보를 에지 분포 정보가 거리 비교에 적합하도록 정규화를 거치고, 1차 비교로 얻어진 거리 값을 가지고 다시 한번 영역별 거리 값을 정규화 하여 비교하는 방식으로 검색 효율을 높인 결과 배경과 구분된 유사 영상의 경우 약 15%, 연속된 유사 영상에 대하여는 16%, 복잡한 질의 영상에 대한 유사 영상은 약 9% 이상의 검색 효율의 향상이 있었다. 특히, 상위에서 검색된 유사 영상의 개수에서도 향상을 보이고 있지만, 전체 순위에 대해서도 큰 폭의 향상을 보임을 성능 평가 비교를 통하여 확인하였다.

향후에는 전경(관심 객체)과 배경이 잘 분리되어있지 않은 복잡한 내용의 영상의 경우라도 전체가 아닌 관심 부분만을 따로 정의하여 기술함으로써 주위 배경과는 무관하게 관심 객체를 검색할 수 있는 방법이 연구되어야 할 것이다.

참고 문헌

[1] 반가운, 유기형, 최재호, 곽훈성 "웨이브렛 변환을 이용한 내용기반 검색 시스템," 전자공학회 학계학술대회 논문집, Vol.21, No.1, pp.733-736, 1998.6

[2] MPEG-7 Homepage
"http://ipsi.fraunhofer.de/delite/Projects/MPEG7"

[3] José M. Martínez. "MPEG-7 Overview", ISO/IEC/JTC1/SC29/WG11 N5525, Pattaya, March 2003

[4] Park, D.K., Jeon, Y.S., Won, C.S., Park, S.-J. "Efficient use of local edge histogram descriptor", Proc. of ACM Multimedia 2000 Workshops, Marina del Rey, 2000

[5] Vaisey, J., Gersho, A "Image compression with variable block size segmentation," IEEE Tr. Signal Process., vol 40, no 8 (1992) 2040-2060

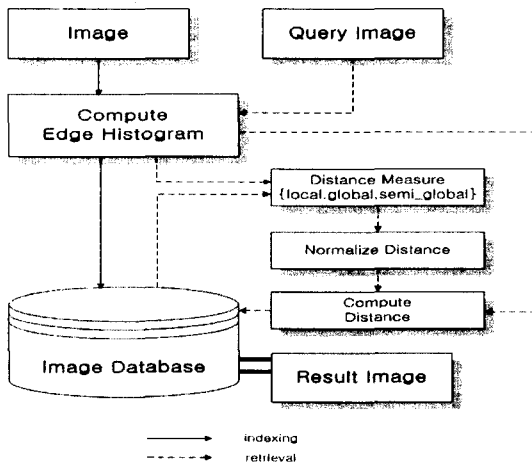


그림 3-1. 제안한 검색 시스템