

다양한 사용자 인터페이스를 고려한 내용기반 영상 검색시스템

방 상 배, 이 채 영, 남 재 열
계명대학교 컴퓨터공학과

Contents Based Image Retrieval System Considering Various User Interface

Sang-Bae Bang, Chae-Young Lee, Jae-Yeal Nam
Department of Computer Engineering, Keimyung University
skyhill@jinri.kmu.ac.kr, jynam@kmucc.kmu.ac.kr

요 약

내용 기반 영상 검색 시스템의 전체 구조를 살펴볼 때 질의영상을 입력하는 입력부, 입력된 질의영상의 색상 및 형태정보를 분석하여 DB내의 영상과 유사도를 측정하는 검색 엔진부, 그리고 유사도 측정 결과 검색된 영상을 사용자에게 보여주는 출력부로 나뉘볼 수 있다. 본 논문에서는 효율적인 검색 결과를 얻기 위해서 입력부에서는 사용자의 편의성을 고려한 인터페이스 설계를 논의하고, 검색엔진부에서는 질의영상의 한 빈에 대하여 비교영상의 인접한 빈 중 유사도가 높은 빈을 검색한 후 그 위치가 가까울수록 인접 빈에 높은 가중치를 부여하는 방식을 이용하여 히스토그램 인터섹션이 가지는 장점을 그대로 계승하면서 색상이나 명도, 채도에 약간의 차이가 존재하는 영상들 효율적으로 검색할 수 있는 검색 기법을 제안하였다. 또한 출력부에서는 검색 시스템의 검색 효율이 뛰어나더라도 그것을 사용자에게 효과적으로 제시되지 못할 경우 일어날 검색 시스템의 효율성 반감을 고려하여 영상 특징간의 관계성과 전체적인 일견의 효과를 제공할 수 있는 landscape 모델을 제안하였다.

1. 서 론

영상 검색 기법의 가장 큰 목적은 사용자가 찾고자하는 유사영상 혹은 사용자가 그린 질의영상을 입력하여 데이터베이스 내에서 가장 유사한 영상을 검색해 내는 것이다. 이런 질의영상에 의한 검색 기법은 질의영상의 내용을 기반(Content-based visual query)으로 유사도를 측정한다. 이러한 질의영상의 입력은 일반적으로 기존에 존재하고 있는 영상을 질의영상으로 사용하는 방법(Query by image)과 사용자가 생각하고 있는 영상의 특징을 그려서 질의영상을 생성하는 방법(Query by sketch)을 사용하고 있다. 검색엔진이 사용하는 영상의 특징은 크게 색상정보를 이용한 방법과 형태정보를 이용한 방법으로 나눌 수 있다. 색상정보를 이용하는 가장 범용적이고 대표적인 방법은 Swain등이 제안한 히스토그램 인터섹션(Histogram intersection) 기법이 있고 형태정보를 이용한 기법은 일반적으로 널리 알려진 경계선 검출 기법이나 이진영상을 이용한 기법 등을 들

수 있다. 그중, 영상 검색 기법 중에서 가장 많이 이용되는 히스토그램 인터섹션 기법은 그 동안 계산상의 용이성과 영상의 회전 및 이동 등에 강한 특성, 영상의 노이즈에 민감하지 않는 성질 등으로 색상을 기반으로 한 대부분의 영상 검색 기법에서 많이 이용되어 왔다. 하지만 형태정보를 판별할 수 없다는 단점을 제외하고도 컬러 히스토그램의 인터섹션을 이용하는 경우 인접한 히스토그램간의 유사도는 전혀 고려되지 않는 단점이 발생하게 된다. 히스토그램의 양자화를 통하여 어느 정도 보안이 가능하긴 하지만 영상에 따라 양자화의 단계가 달라질 경우 유사한 색상 분포를 가진 영상을 서로 다른 영상으로 판별하는 경우도 이 때문이다.

본 논문에서는 위에서 언급한 문제를 해결 혹은 보완할 수 있는 사용자 편의를 고려한 질의영상의 생성기법과 새로운 히스토그램 인터섹션 알고리즘을 제안한다.

2. 새로운 검색 시스템의 제안

2.1 사용자 편의성을 고려한 질의영상의 입력

기존의 검색 시스템의 질의영상 입력은 크게 사용자가 찾고자하는 영상의 특징들을 직접 그려서 질의영상을 만드는 방법(Query by sketch)과 사용자에게 질의영상의 예제를 제공하여 질의영상을 입력시키는 방법(Query by image)으로 나누어 볼 수 있다. 전자의 경우 사용자들의 시각적 기억력의 한계와 스케치 능력의 부족으로 찾고자하는 특징들을 정확하게 구분하기가 쉽지 않다는 단점이 있고 후자의 경우 하나의 질의영상이 찾고자하는 모든 특징요소를 가지고 있기는 힘들다는 단점이 있다. 또, 사용자가 생각하는 영상의 특징요소와 검색 시스템 내에서 추출되는 특징요소간의 이질감이 있기 때문에 검색 결과에 상당한 영향을 미친다.

본 논문에서는 이러한 이질감을 극복하고 사용자에게 보다 편리하고 효율적인 질의영상을 제공하기 위하여 개선된 인터페이스 모델을 제안한다. 제안된 인터페이스 모델은 크게 질의영상의 입력에서 질의영상이 포함하고 있는 영상특징의 추가, 삭제, 그리고 변경으로 나누어 볼 수 있다. 하나의 질의영상이 사용자가 원하는 모든 영상 특징들을 가지고 있기는 어렵기 때문에 질의영상의 입력부에서 두 개 이상의 여러 영상들과 합성을 통하여 이러한 영상특징의 부족분을 채워 나갈 수가 있다. 그리고 사용자가 원하지 않는 영상특징에 대해서는 사용자가 원하는 영상특징을 가지는 다른 영상의 일부를 복사와 붙이기를 통하여 원하지 않는 영상특징을 제거하고 새로운 영상특징으로 대체하여 검색효율을 향상시킬 수가 있다. 또, 사용자가 특정 질의영상과 데이터베이스 내의 영상이 비슷하다고 인지를 하더라도 검색시스템에 의한 검색순위가 상당히 뒤로 밀리는 경향을 종종 볼 수가 있다. 이는 검색시스템에서 형태정보는 비슷하지만 색상정보의 차이 때문에 그만큼 유사도가 떨어진다고 계산하기 때문이다. 여기서 질의영상 전체적인 명도, 채도, 색상조각 등의 편집 기능을 구현하여 사용함으로써 이러한 문제는 해결 가능하다.

사용자 인터페이스측면에서 사용자 직관적인 질의영상의 입력을 위하여 특정 영역내의 모든 영상을 한번에 많은 질의영상 후보로 보여주도록 Thumbnail browser 형태를 취하도록 구현하였고, 또 검색 결과로 나타나는 영상들을 다시 질의 영상으로 입력함으로써 검색 결과를 정렬하는 과정의 효과도 있도록 구현하였다.

2.2 인접 빈을 이용한 새로운 히스토그램 인터섹션 기법

영상이나 동영상의 검색에서 Swain등이 제안한 히스토그램 인터섹션 기법은 서로 유사도가 높은 영상은 컬러의 분포에 있어서도 서로 유사성이 높을 것이라는 가정 하에 히스토그램의 컬러 오브젝트를 비교함으로써

유사도가 결정된다. 유사도는 두 히스토그램의 교집합으로 나타내는데 다음 식 (1)과 같이 표현된다.

$$h_c(Q \cap T) = \frac{\Delta \min\{H_c(Q), H_c(T)\}}{|Q|} \quad (1)$$

여기서 Q 와 T 는 질의 질의영상과 데이터베이스 내의 영상이고, $H_c(Q)$ $H_c(T)$ 는 컬러에 대한 영상내의 픽셀 개수가 된다. 히스토그램 인터섹션 기법을 적용한 예를 보면 히스토그램 양자화 단계를 N 으로 하고, 각 양자화 단계의 컬러를 c_i 라 했을 경우 전체 영상에 대한 유사도는 다음 식 (2)로 표현될 수 있다.

$$h(Q \cap T) = \frac{\sum_{i=0}^N \min\{H_c(Q), H_c(T)\}}{\sum_{i=0}^N H_c(Q)} \quad (2)$$

히스토그램 인터섹션기법은 실제로 약간의 색상 차이만 존재하는 영상을 완전히 서로 다른 영상으로 인식하게 되는 문제가 발생한다. 이러한 문제를 해결하기 위해 본 논문에서는 히스토그램의 유사도를 계산할 때 인접한 빈 간의 유사성을 고려하여 전체적인 유사도를 계산하는 알고리즘을 제안한다. 우선, 히스토그램 인터섹션의 특성상 분모는 질의영상의 전체 픽셀 수의 총합으로 나타내어지고, 분자는 각각의 빈 내에서 질의영상과 비교 영상 중 작은 부분의 빈이 누적되게 된다. 이에 반해, 제안된 알고리즘에서는 전체 히스토그램 인터섹션이 크게 두 가지 방향으로 진행된다. 분모의 값은 변화가 없는 대신 분자의 값에서 주위의 빈들을 고려하게 된다. 이때, 분자의 값들을 임시로 저장하기 위한 변수로 U 를 둔다. 1) 먼저 질의영상의 히스토그램값이 비교영상의 값보다 작을 경우는 질의영상의 히스토그램값을 그대로 분자에 누적시켜 준다.

$$H_c(U) = H_c(Q)$$

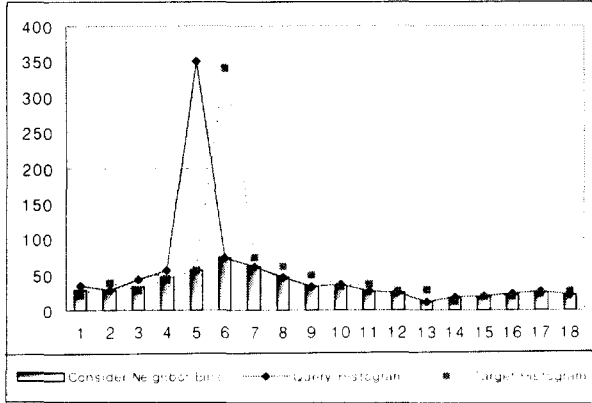
2) 반면에 비교영상의 히스토그램값이 질의영상의 히스토그램값보다 작을 경우는 질의영상의 빈과 비교영상의 이웃한 다른 빈과의 값을 비교하게 된다. 그중 가장 유사한 빈을 찾아내어 그 빈과 원래 비교대상인 빈과 가중치를 고려하여 그 결과를 분자로 반환하게 된다.

$$H_c(U) = \frac{w\{H_c(T)\} + M_c(T)}{w+1} \quad (3)$$

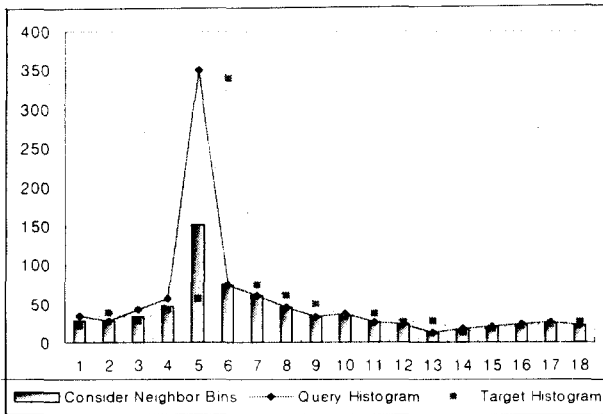
여기에서 $M_c(T)$ 는 $H_c(Q)$ 와 유사도가 가장 높은 빈의 값이고, w 는 유사 빈의 위치에 따른 가중치이다. 이렇게 구한 분자를 이용해서 전체 유사도를 구현하면 다음 식 (4)와 같다.

$$h_{min}(Q \cap T) = \frac{\sum_{i=0}^N H_c(U)}{\sum_{i=0}^N H_c(Q)} \quad (4)$$

그림 1.(a)를 보면 5번 bin의 경우 낮은 히스토그램 인터섹션 값을 갖는데 반하여 그림 1.(b)의 경우에는 이 값이 제안된 알고리즘에 의하여 개선되었음을 알 수 있다.



(a) 히스토그램 인터섹션

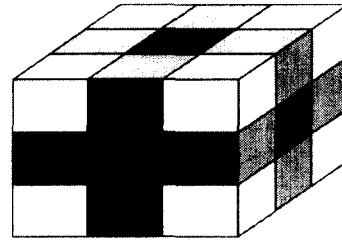


(b) 인접빈을 고려한 히스토그램 인터섹션
그림1. 인접빈을 고려한 유사도 향상

인접 빈을 이용한 알고리즘을 구현하기 위해서는 질의 영상의 bin과 가장 유사한 bin의 위치를 알아야 한다. 그 이유는 원래 비교대상이 되는 bin과의 거리가 가까우면 가까울수록 전체 유사도를 높게 책정해야 하기 때문이다. RGB 컬러 좌표는 3차원으로 표현되므로, bin 간의 거리 측정 시에도 x, y, z 세 축을 모두 고려한 형태로 나타낼 수 있다.

$$w = 2 \times (|x_q - x_t| + |y_q - y_t| + |z_q - z_t|) \quad (5)$$

본 논문에서는 RGB 좌표에서 인접 bin의 가중치를 식(5)에서처럼 각각의 축에 대한 절대치 오차를 이용하여 계산하였다.



■ $w = 2$ ■ $w = 4$ □ $w = 6$

그림 2. RGB 좌표에서의 가중치 w 의 값

그림 2는 RGB 컬러 좌표상에서의 가중치를 부여한 결과를 나타낸다. 그림에서 한 가운데의 bin이 현재 비교대상이 되는 bin이라고 하면 각각 이웃한 bin의 가중치는 그림 2와 같다.

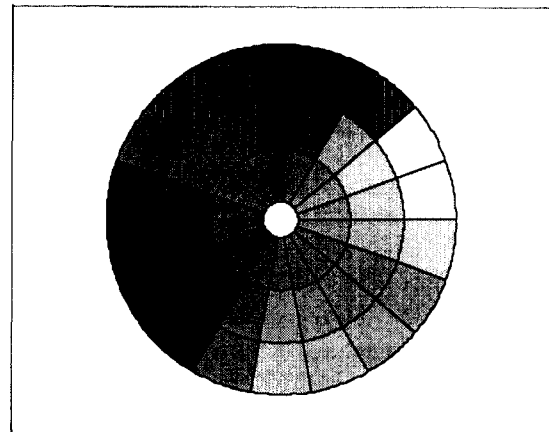


그림 3. Hue-Saturation 평면상에서의 인접bin 가중치

HSI 컬러 좌표의 경우에는 색상(hue) 성분의 양자화를 조밀하게 하므로, 인접bin도 2단계 떨어진 bin까지 고려 대상으로 한다. 그림 3의 예를 들면, 가운데 위에 있는 밝은 bin을 현재 비교대상이 되는 bin이라고 두면, 면이 인접한 경우에는 2의 가중치를 선이 인접한 경우나 Hue 성분의 면이 2단계 건너 인접한 경우에는 4의 가중치를 설정하였다.

2.3 효율적인 검색환경의 제공

검색결과가 너무 많거나 부정확해서 사용자가 쉽게 정보를 알아보지 못하거나 검색효율이 뛰어나더라도 그것이 사용자에게 효과적으로 제시되지 못하는 경우 검색시스템의 효율성을 반감시키는 결과를 초래하게 되는 것이다. 따라서, 검색결과와 효과적인 표현과 데이터의 전체적인 구조와 상호 관계를 사용자에게 효과적으로 제공할 필요가 있다. 이와 같이 영상 브라우저에 있어 기존의 Thumbnail, Zoom browser, Multiple window 등의 평면상의 제약된 출력 방법에 대해 비주얼 속성간의 관계성 파악의 어려움과 사용자가 원하는 결과와는 상

이한 검색결과와 출력에 대한 문제점의 해결 및 전체적인 일견의 효과를 제공함으로써 부가적 정보를 얻을 수 있게 되는 기회를 제공하며, 사용자와의 다양한 인터랙티브한 인터페이스의 제공 모델이 필요하다.

본 절에서는 영상 검색 시스템을 위한 영상 브라우징 기법으로 3차원 시각화 모델을 위한 패러다임을 제시한다. 3차원 기법은 효율적인 영상 브라우징과 전체적 영상의 분포를 알 수 있도록 하는 객체 공간을 제공하며 또한 사용자와 인터랙티브한 인터페이스를 제공하는 가상세계 공간을 제공한다. 아래에는 새로운 영상 브라우징 모델에 대한 프로토타입에 대해 기술한다.

첫째로는 전체적인 뷰를 제공함으로써 영상 검색 결과에 대한 사용자의 일견을 가능하게 하여 객체간 관계성의 파악과 경향성을 직관할 수 있는 메타포의 제공이다. 둘째는 애니메이션 기법을 사용하여 영상을 표현함으로써 사용자와의 인터랙티브한 인터페이스를 제공하는 것이다. 실제 크기의 영상들은 해당 객체를 클릭하거나 조작함으로써 전시하는 Zoom browser의 기능도 갖게 한다. 셋째로는 휴먼 인터페이스 측면에서 3차원공간을 가상공간처럼 사용하여 panning 및 zooming, navigation을 할 수 있게 하는 가상현실 개념의 지원이다.

이러한 프로토타입을 가지고 브라우징 모델로 Landscape 모델을 제안한다. 기존 영상검색 시스템의 검색 결과는 대부분 Thumbnail형식으로 표현되어 나타내게 되는데, Landscape 모델에서는 검색 키워드로 작용되어지는 영상의 컬러정보와 형태정보의 유사도를 기준으로 영상 브라우징 공간에서 좌측으로는 컬러정보의 유사도가 우세하며, 우측으로는 형태정보의 유사도가 우세함을 나타내게 하여 객체간 관계성의 파악을 가능케 하고 사용자가 전망을 보는 듯한 시선 제공의 기능을 가짐으로 전체적 뷰를 제공하게 된다. 3차원 공간에서 사용자가 직접 공간을 네비게이션하는 기능과 관찰자의 자유로운 시선처리의 기능, 해당 객체를 클릭 함으로써 상세정보의 출력 또는 전체영상을 출력함으로써 Zoom browser의 기능을 갖는 인터랙티브한 사용자 인터페이스를 제공하게 된다.

3. 실험 결과

본 논문에서 제안한 인접 빈을 고려한 히스토그램 인덱싱 알고리즘의 성능을 평가하기 위하여 MPEG-7 표준 제정을 위한 데이터 2300여 개의 영상을 대상으로 검색 효율을 측정하였다. 성능평가의 방법으로는 MPEG-7에서 제안된 NMRR(Normalized Modified Retrieval Rank)기법[-]을 이용한다. 정규화된 NMRR값은 [0.0, 1.0]사이의 값을 가지며 그 값이 작을수록 검색 효율성이 높음을 나타낸다. 실험은 위에서 설명한 영상 중 관련영상을 뚜렷이 구분할 수 있는 영상 15개를 선

택하여 그 값을 평균하여 계산하였고 그 결과는 그림 4에 나타나있다.

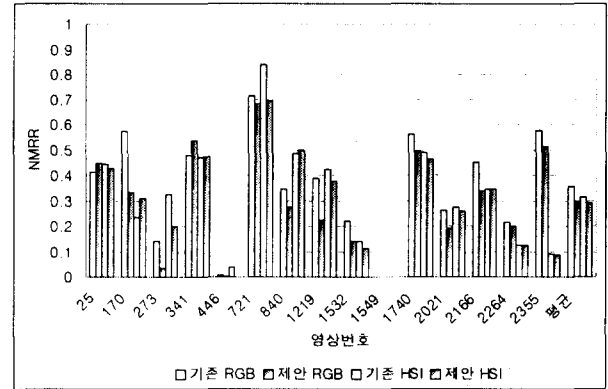


그림 4 15개 영상의 실험 결과

결과에서 보면 인접 빈을 고려한 알고리즘을 적용한 수치가 특정 몇몇을 제외하고는 전체적으로 검색효율이 뛰어난을 알 수 있다.

4. 결론

기존의 검색 시스템을 입력부, 검색엔진부, 출력부로 나누어 각각의 인터페이스 모델과 새로운 컬러 히스토그램 검색 알고리즘을 제시하였다 특히, 기존의 컬러 히스토그램에서 동일한 빈만을 고려했던 점을 보완하여 제안한 인접 빈간 유사도 고려를 통해 검색 효율의 개선을 가져올 수 있었다. 개선된 알고리즘을 통한 검색 효율향상뿐만 아니라 질의영상 입력부에서 간단한 전처리과정을 거쳐 검색효율의 향상이 이루어 질 수 있음을 보였고 또한, 검색 결과를 사용자에게 어떻게 효과적으로 보여 줄 것인지도 고려하여 사용자에게 좀더 손쉬운 검색이 이루어 질 수 있도록 하였다.

5. 참고문헌

- [1] J. Assfalg, Del Bimbo, P. Pala "Using positive and negative examples for precise image retrieval" Proc. of SPIE 24-26 January 2001
- [2] M. J. Swain and D. H. Ballard, "Color indexing," Int. J. of Computer Vision, vol. 7, no. 1, pp. 11-32, 1991.
- [3] Richard J. Qian, Peter J. L. van Beek and M. Ibrahim Sezan, "Image Retrieval Using Blob Histograms," ICME 2000, July/Aug. 2000.
- [4] Plaisant, C., Carr D. and Shneiderman B. "Image-browser taxonomy and guidelines for designers" IEEE Software 12, 2. Mar. 1995
- [5] ISO/IEC/JTC1/SC29/WG11 : "Licensing Agreement for the MPEG-7 Content Set," document no. N2466, Atlantic City, Oct. 1998.