

투영 벡터의 첨두치 간 거리를 이용한 영상검색

정희원 권동현*, 이태홍*

Image Retrieval Considering Inter-Distance between Peaks of Projection Vector

Donghyun Kwon*, Taihong Yi* *Regular Members*

요 약

필터를 이용한 내용기반 영상검색의 대표적 방법에 해당하는 히스토그램 인터섹션은 영상에서 필터가 가지는 값의 빈도 수를 이용하여 간단하면서도 효율적으로 영상을 검색하면서 영상 내에서의 물체의 이동, 회전에 대해 대처하는 장점을 가지는 방법이다. 그러나 영상에서의 위치 정보가 누락되어 다른 영상을 동일 영상으로 인지하기 쉽고, 동일 영상에 대해 조명 변화가 있을 때 히스토그램이 영향을 많이 받아 검색이 어렵다는 단점을 지닌 방법이다. 본 논문에서는 조명 변화가 있을 경우에도 검색이 가능한 방법을 제안한 것으로, 영상에 1차원 투영 기법을 적용해 투영 값이 큰 벡터들 간의 상대 거리를 이용하여 검색에 이용하였다. 검색 성능의 확인을 위하여 5가지 조명에서 얻은 영상 데이터베이스를 이용하여 실험하였으며, 실험 결과 기존의 방법에 비해 적은 색인 데이터량으로 검색이 가능하면서 조명 변화에도 대응할 수 있는 검색 결과를 얻을 수 있었다.

ABSTRACT

As one of the color image retrieval methods, histogram intersection is easy and simple method that counts the occurrence of color pixels. The method is invariant to translation and rotation about objects within image. But, because of the absence of spatial information, it is difficult to discriminate images of the similar histogram. Also under the variation of illumination, the change of histogram results in poor retrieval rate. Thus this paper proposes a method that can retrieve reasonably under varying illuminant condition. This method utilizes a characteristic that is invariant to peak distance of the projected data. In experimentation, by using images acquired for the five illuminations, we could verify the method needed reduced index data and resulted better retrieval in comparison with conventional methods.

1. 서론

하루가 다르게 발전하는 컴퓨터 기술과 급속도로 보급되고 있는 네트워크의 확산과 더불어 전세계에 분산된 정보를 얼마나 빨리 찾을 수 있는가 하는 것이 경쟁에서의 우위를 확보하는 중요한 요소가 되었다. 웹을 이용하여 원하는 대상을 찾기 위해 지금까지는 텍스트 위주의 검색 방법이 주류를 이루었으나, 웹 사이트의 증가와 함께 데이터의 양이 천문학적으로 늘어남에 따라 검색에 걸리는 시간이

상대적으로 늘어나게 되었고, 또한 인간의 주관에 의해 작성된 주석(annotation)정보의 한계로 최근에는 검색 사이트마다 멀티미디어를 이용한 검색 방법도 동시에 사용되는 추세이다.

멀티미디어 검색 방법 중, 인간이 일상 생활에서 이용하는 방법인 의미론적인 방법은 가장 고차원적인 방법으로 현재 이를 구현하기에는 기술적으로 어려움이 많은 상황이라 일부 한정된 적용 분야에만 사용되고 있으며, 미디어 자체에 포함된 정보에서 검색에 필요한 특징을 추출해 사용하는 내용이

* 영남대학교 대학원 전자공학과 영상통신연구실(kdh@lily.cc.yeungnam.ac.kr),
논문번호 : 010001-0216, 접수일자 : 2001년 2월 16일

반 검색(Content based retrieval) 방법이 현재 많이 사용되고 있다.

내용에 기반을 둔 영상 검색에서 많이 사용되는 특징으로 색(color), 스케치(sketch), 형상(shape), 텍스처(texture), 공간적 관계(spatial relationship) 등이 있다^[1]. 그 중에서 특히 컬러는 영상의 전체 특성을 잘 나타내면서도 구현이 쉽다는 특징으로 인해 Swain^[2]의 히스토그램 인터섹션을 이용한 방법 이후 활발히 연구되고 있는 분야이다. 그러나 히스토그램을 이용할 경우 공간 정보가 없어지는 단점 때문에 서로 다른 영상이라도 히스토그램만 같으면 같은 영상으로 검색하는 오류가 일어나기 쉽고, 조명 변화가 있을 때에 히스토그램이 전체적으로 변화하여 원하지 않는 영상이 검출되기 쉽다는 문제점이 있다.

공간 정보 부재 문제 해결 방법 중의 하나인 투영(projection)은 영상의 1차원적인 가로 및 세로 방향의 정보를 개략적으로 나타내는 방법으로, 영상을 블록화하여 통계적 특성을 이용할 경우 영상 크기에 상관없이 이용 가능하여 히스토그램에서의 공간 정보 부재 문제를 해결하기 위한 방법으로 사용될 수 있는 방법이다^[3].

히스토그램에서의 또 다른 문제점인 조명 변화에 대처하기 위하여 본 논문에서는 투영 기법을 이용하여 영상의 투영 벡터에서 동일 영상에 대해 조명이 변하는 환경에서의 특징을 추출하여 검색에 이용하였다. 조명이 변하는 환경에서의 영상검색에 관해서는 Funt^[4]가 속한 Simon Fraser University에서 가장 많은 연구가 이루어졌다. Funt가 제안한 CCCI(color constant color indexing) 기법은 조명이 변화하는 환경에서도 컬러 비는 일정하게 유지된다는 성질을 이용하여 검색을 하고 있다. 본 논문에서 사용하는 투영 벡터를 이용하여 추출한 특징은 조명 변화가 있을 지라도 물체가 가지는 컬러 간의 상대적 거리는 일정하다는 성질을 이용하여 검색을 수행하고 있다.

본 논문의 전개를 위하여 2절에서는 컬러 및 투영 기법을 이용한 영상검색에 대해 소개하고, 3절에서는 본 논문에서 제안하는 방법인 투영 벡터의 척 두치 간 거리를 이용한 검색 방법을 설명한 후, 4절에서는 실험 환경 및 결과를 분석하고, 마지막으로 5절에서 결론을 맺는다.

II. 컬러 및 투영을 이용한 영상검색

컬러를 이용한 검색 중 최초로 소개된 히스토그램 인터섹션은 질의 영상에 대한 히스토그램을 Q 라 하고, 데이터베이스 영상에 대한 히스토그램을 I 라 할 때

$$HI = \sum_{i=0}^N \min(I_i, Q_i) \quad (1)$$

와 같이 질의 영상에 대해 데이터베이스 영상에서 구한 히스토그램과의 차이 값을 합산하여 비슷한 정도를 표현하고, 그 차이가 가장 작은 영상을 가장 유사한 영상으로 검색하였다. 식 (1)은 영상을 검색할 때 히스토그램만을 비교 대상으로 하기 때문에 영상의 위치가 이동하거나 회전이 일어난 영상에 대해서도 우수한 검색 성능을 발휘한다.

그러나, 서로 다른 영상이라도 히스토그램은 같을 수 있기 때문에, 영상 내에서의 컬러 빈도 수만을 비교 대상으로 하는 컬러 히스토그램을 이용할 경우 잘못된 영상을 검출할 가능성이 많다. 이러한 공간 정보의 부재 문제를 해결하기 위하여 Stricker^[5]는 영상을 임의의 영역으로 나눈 후, 모멘트 정보를 이용하여 검색을 시도하였다.

히스토그램의 공간 정보 부재 문제 해결을 위한 방법의 하나로 그림 1에서와 같은 투영이 사용될 수 있다.

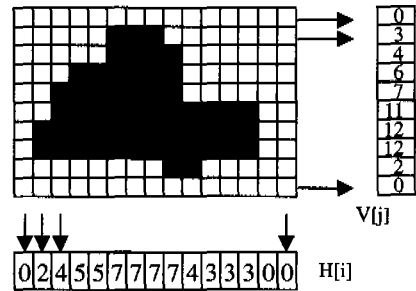


그림 1. 이진 영상에서의 투영 예

크기가 $N \times N$ 인 영상을 $I[r, c]$ 라 할 때 수평 및 수직 방향으로의 투영은 각각

$$H[r] = \sum_{c=0}^{N-1} I[r, c] \quad (2)$$

$$V[c] = \sum_{r=0}^{N-1} I[r, c] \quad (3)$$

와 같이 나타낼 수 있다.

영상에서의 1차 모멘트를 이용하여 물체의 위치를 알 수 있는 것과 마찬가지로, 영상에서의 1차 모멘트가 투영에서의 1차 모멘트로 나타난다는 성질을 이용하여 투영된 벡터에서 물체의 위치를 계산할 수 있다. 식

$$A = \sum_{c=0}^{N-1} V[c] = \sum_{r=0}^{N-1} H[r] \quad (4)$$

을 이용하여 수평 및 수직 방향의 위치를 계산하면 식

$$\bar{y} = \frac{\sum_{r=0}^{N-1} rH[r]}{A} \quad (5)$$

$$\bar{x} = \frac{\sum_{c=0}^{N-1} cV[c]}{A} \quad (6)$$

와 같이 행방향 y 및 열방향 x의 위치를 계산할 수 있다^[6].

그러나 투영은 영상의 크기에 따라 생성되는 행 및 열 방향의 투영 벡터 크기가 달라진다는 단점이 있다. 이러한 투영 벡터의 단점을 극복하고자 투영 벡터의 행 및 열 방향에 대해 블록을 구성한 후, 각 블록에 대해 각각 투영 벡터의 통계적 성질을 적용하면 적은 양의 색인 데이터로 영상 검색을 가능하게 할 수 있다^[3].

III. 투영 벡터의 침두치 간 거리를 이용한 영상검색

사진을 촬영할 때 주변 조명의 상태에 따라 인화 후의 사진은 많은 차이가 난다. 조명이 적절할 경우 색조가 뚜렷하고, 선명한 사진을 얻을 수 있지만 실내의 빈약한 불빛 환경에서 촬영했을 경우의 사진은 어둡고, 색조도 흐릿하게 나타난다.

컬러 히스토그램을 이용할 경우의 문제점으로 지적할 수 있는 것은 조명 변화에 대해 민감하다는 것이다. 예를 들어 그림 2의 (a), (b)에서와 같이 하나의 영상이 서로 다른 조명 환경에서 촬영되었다고 할 때 로그 단위로 나타낸 히스토그램은 그림 3의 (a), (b)에 나타낸 것과 같이 서로 다른 형태를 나타내어 히스토그램 인터섹션에 관한 식 (1)을 적용하면 차분 성분이 누적되어 유사도에서 차이가 많이 나는 서로 다른 영상으로 검색되게 된다. 이와 같이 같은 대상을 두고 영상을 얻는 과정에서 조명

이 변화할 경우 히스토그램을 이용한 방법으로 영상을 검색하면 조명의 변화에 따라 히스토그램의 변화가 많이 차이 나기 때문에 좋은 검색 결과를 기대하기 어렵다. Swain^[2]은 각 채널 별로 $r' = R/(R+G+B)$, $g' = G/(R+G+B)$, $b' = B/(R+G+B)$ 을 적용하여 조명 변화에 대처했지만 그 결과는 좋지 않았다. 또한 Funt^[4]는 CCCI 기법에서 센서 응답의 계수 모델을 이용할 경우 영상의 로그를 취한 값의 비는 조명 변화에 대해 일정하다는 성질을 수학적으로 밝히고 이를 실험하였다. 그러나 계수 모델이 성립하기 위한 전제 조건으로는 센서의 응답이 아주 좁은 대역에 대해 반응해야 한다는 문제점을 갖고 있기 때문에 일상적인 환경에서의 적용에는 무리가 있었다.

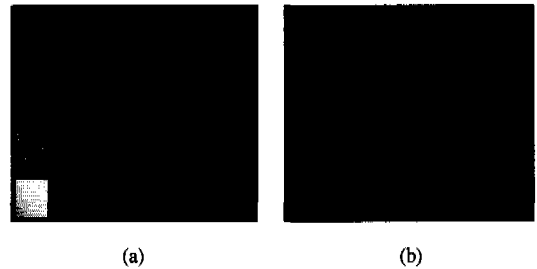


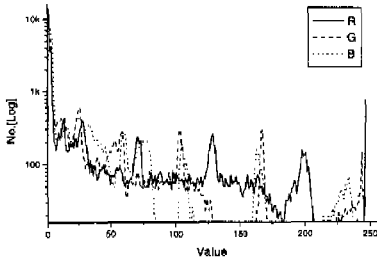
그림 2. 서로 다른 조명에서의 같은 물체

그림 4의 (a), (b)는 각각 그림 2의 (a), (b) 영상에 대한 수직 투영 벡터를 나타낸 것이다. 그림 3의 히스토그램 특성과는 달리 그림 4의 (a)영상과 (b)영상 사이에는 각 채널별로 투영 값의 크기는 달라지지만 그 형상은 변화가 거의 없다는 것을 알 수 있다. 투영 벡터의 이러한 성질을 이용하여 투영 벡터의 각 채널별로 임의의 개수의 침두치를 구한 후, 각 침두치 간의 상대 거리를 색인 데이터로 이용할 경우 조명이 변화하는 환경에서 동일 물체가 포함된 영상을 검색하는데 이용할 수 있다. 본 논문에서 침두치는 수평, 수직 투영 벡터에 대해 각각 3개씩 검출하였고, 그 결과로 구해진 상대 거리 색인 데이터는 수평, 수직에 대해 각각 2개씩 구해졌다.

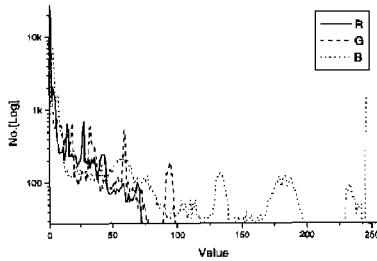
이를 3개의 채널에 대해 구하면 하나의 영상에 대한 총 색인 데이터는 12개이다. 침두치 간의 상대 거리를 구하기 위한 식은

$$d_{ij} = |p_i - p_j| \quad (7)$$

와 같이 각 침두치 간의 절대치를 이용하였다.

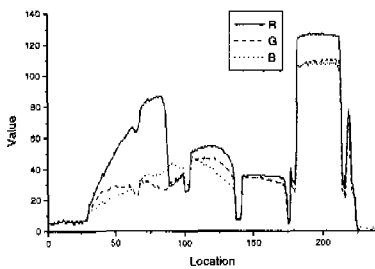


(a)

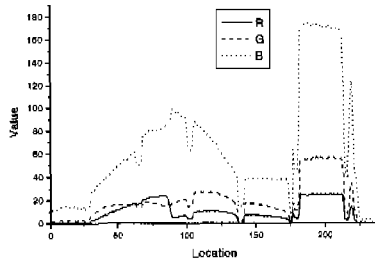


(b)

그림 3. 그림 2 영상의 히스토그램



(a)



(b)

그림 4. 그림 2 영상의 수직 투영 벡터

IV. 실험 및 실험 결과

조명이 변화하는 환경에서 검색 성능을 확인하기 위하여 본 논문에서는 Simon Fraser University에서 컬러의 연구를 위해 사용한 영상 데이터베이스를 사용하였다. 데이터 베이스를 구성하고 있는 영상은 하나의 물체 당 5가지의 조명을 이용하여 촬영한 11가지의 영상 55장을 이용하였다. 표 1에서는 실험에 사용된 조명을 나타내었다.

표 1. 실험에 사용된 조명 이름

순번	조명 이름
1	Macbeth 5000 K fluorescent tubes.
2	Macbeth 5000 tubes with Roscolux #3202 Full Blue filter.
3	Sylvania Cool White Fluorescent tube.
4	Philips Ultralume fluorescent tube.
5	Sylvania 75W halogen bulb

히스토그램 인터섹션을 이용한 실험을 위해 먼저 영상을 RGB 채널로 분리한 후, 각 채널에 대한 히스토그램을 구하여 색인 데이터로 저장하였다. 질의 영상이 Query by example로 주어졌을 때 질의 영상에 대한 히스토그램을 구한 후, 식 (1)을 사용하여 유사도를 비교하여 유사한 정도에 따라 순위를 매겼다. 유사도 비교 시의 척도(similarity measure)로는

$$S_{ij} = \sum_{i=1}^M \sum_{j=1}^N \sqrt{(I_{ij} - Q_{ij})^2} \quad (8)$$

의 L₂-norm을 사용하였다. CCCI를 위해서는 영상 데이터에 자연 로그를 취한 후, 라플라스 연산자를 구한 후, 히스토그램 인터섹션과 같은 형태로 처리하였다.

또한 제안된 방법의 구현을 위해서는 RGB 채널로 분리된 영상에 대해 각 채널별로 수직 및 수평 투영 벡터를 구해, 각 투영 벡터가 가지는 3개의 침투치를 투영 값으로부터 구하였다. 이 과정에서 침투치 간 최소 거리는 20 이상이 되도록 하였다. 이렇게 구한 침투치 간의 상대 거리는 식(7)을 사용하여 구했으며, 각 채널별로 4개씩 구하여 저장 후 색인 데이터로 사용하였다. 질의 영상이 주어졌을 때,

같은 과정으로 색인 데이터를 구한 후, 유사도 비교를 위해서는 히스토그램 인터섹션에서 사용한 식(8)을 이용하였다.

표 2에서는 히스토그램 인터섹션 및 CCCI를 이용하였을 때와 투영 벡터의 침투치 간 거리를 이용하였을 때의 검색 결과를 나타낸 것이다. 검색 결과는 하나의 영상을 다섯 가지 조명 상태에서 촬영한 영상을 사용했기 때문에 질의 영상 이외의 동일 물체가 포함된 영상 4개를 100 % 찾아 내는 최수로 나타내었다. 표 2의 결과에 나타낸 것과 같이 5위 이내에서 100 % 검출을 한 개수는 히스토그램 인터섹션이 2개, CCCI에서도 2개인 반면 투영 벡터의 거리를 사용하였을 때는 28 개, 10위 이내에 동일 영상 4개가 100% 검출된 개수는 히스토그램 인터섹션이 8개, CCCI에서는 11개인 반면 투영 벡터의 거리를 이용했을 경우에는 41개를 나타내었다. 또한 15위 이내에 동일 영상이 100% 검출된 개수는 히스토그램 인터섹션이 12개, CCCI에서는 19개, 투영벡터의 거리를 사용한 경우는 48개로 전체적으로 투영 벡터의 거리를 색인 데이터로 사용한 경우가 좋은 검색 결과를 나타낸다는 것을 알 수 있다. 또한 특기할 것은 투영 벡터의 거리를 이용한 경우 55개의 영상 중 100% 검색한 48개를 제외한 나머지 7개의 영상 중 5개의 영상에 대해서는 75% 검색 결과를 나타내었으며, 2개는 50%의 검색 결과를 나타내었다는 것이다. 이는 CCCI에서 75% 검출 20개, 50%검출 8개, 25% 검출 8개와 비교해도 좋은 결과인 것을 알 수 있고, 히스토그램 인터섹션에서 75% 검출 24개, 50% 검출 12개, 25 % 검출 4개, 하나도 검출하지 못한 경우 3개와 비교해도 우수한 결과이다.

표 2. 검색결과

검색방법	100% 검색 개수		
	5위 이내	10위 이내	15위 이내
HI	2개	8개	12개
CCCI	2개	11개	19개
Dis.of Prj.	28개	41개	48개

검색 결과의 내용면에서 보더라도 그림 2의 (a) ball 영상을 질의 영상으로 주었을 때 히스토그램 인터섹션과 CCCI에서는 (b) 영상을 15위 이내에서 검출해 내지 못한 반면 투영 벡터의 거리를 이용

했을 경우에는 4위로 검출하였다.

전체 검색 성공율에서 각 방법간의 차이를 보기 위해 15위 안에 하나라도 검색이 되면 1점을 주고 2개, 3개, 4개를 검색했을 때 2, 3, 4 점을 주어 전체 합을 계산한 점수는 히스토그램 인터섹션이 143 점을 기록한 반면, CCCI에서는 160점을 나타내었고, 투영 벡터의 상대거리를 이용한 경우에는 211점을 기록하여 두 방법에 비해 더 좋은 결과를 나타내었다. 각 방법에 따른 조명 별 검색 성공율을 비교했을 때 히스토그램 인터섹션 및 CCCI에서는 Sylvania Cool White Fluorescent tube 조명을 사용한 경우 검색이 제일 잘 되었으며, Macbeth 5000 tubes with Roscolux #3202 Full Blue filter 조명을 이용하였을 경우 검색 성능이 제일 떨어지게 나타났다. 반면에 투영 벡터의 거리를 이용한 경우에는 Macbeth 5000 K fluorescent tubes를 이용한 경우의 검색이 제일 좋았고, Sylvania 75W halogen bulb를 이용한 경우가 제일 처진 결과를 나타내었지만 최상과 최하 결과간의 차이가 거의 나지 않는 고른 분포를 나타내었다.

검색을 위해 사용되는 색인 데이터의 양을 비교해 보면 한 개의 영상 당 히스토그램 인터섹션 및 CCCI의 경우 256 레벨에 대해 세개의 컬러 채널만큼의 데이터가 필요하지만, 제안된 투영 벡터의 거리를 이용할 경우 한 채널당 4개, 도합 12개만 필요하기 때문에 데이터베이스의 용량이 커질수록 색인 데이터 차이가 많이 나게 된다.

V. 결론

본 논문은 히스토그램을 이용한 영상 검색에서의 단점으로 거론되는 공간 정보 부재와 조명 변화에 대처하지 못하는 문제점을 해결하기 위해 영상의 1차원 투영 기법을 이용하였으며, 투영된 데이터의 침투치 간 거리는 조명이 변화되는 환경에서도 일정하게 유지된다는 성질을 이용해 색인 데이터로 사용하여 영상 검색에 적용하였다.

다섯 가지 형태의 조명에서 얻은 영상 데이터베이스를 이용하여 실험을 한 결과 히스토그램 인터섹션 및 CCCI 기법에 비해 제안된 투영 벡터의 거리를 이용한 방법이 전체 검색에서 히스토그램 인터섹션의 143점과 CCCI 기법의 160점에 비해 211점으로 우수한 결과를 얻을 수 있었으며, 조명별로도 투영 벡터의 거리를 이용한 결과가 히스토그램 인터섹션에 비해 고른 우수한 검색 성능을 보여주

었다. 또한 검색을 위한 색인 데이터의 양에서도 투영 벡터의 거리를 이용한 방법은 하나의 영상 당 12개의 색인 데이터만으로 검색이 가능하였다.

참 고 문 헌

[1] F.Idris and S.Panchanathan, "Review of Image and Video Indexing Techniques," *Journal of Visual Communication and Image Representation*, vol.8, no.2, pp.146-166, June 1997

[2] Michael J. Swain and Dana H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol.7, no.1, pp. 11-32, 1991

[3] 권동현 외 3인, "투영벡터의 통계적 성질을 이용한 영상검색", *한국통신학회 논문지 제 25권 제7A호*, pp. 1044~1049, Jul. 2000

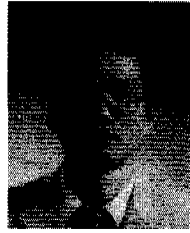
[4] Brian V. Funt and Graham D. Finlayson, "Color Constant Color Indexing," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.17, no.5, May 1995

[5] Markus Stricker and Alexander Dimai. "Color indexing with weak spatial constraint," *Storage and Retrieval for Image and Video Databases IV*, SPIE proceedings, vol. 2670, pp.1630-1639, 1996

[6] Ramesh Jain, *Machine Vision*, pp.35-38, McGraw-Hill 1995

권 동 현(Donghyun Kwon)

정회원



1990년 2월 : 영남대학교

전기공학과 졸업

1990년 8월 ~ '92년 8월 : 영남대

학교 전자공학과

석사 졸업

1992년 8월 ~ 1998년 6월 : 한국

전자통신연구원 연구원

1998년 9월 ~ 현재 : 영남대학교 전자공학과 박사과정
<주관심 분야> 영상 및 비디오 압축/신호처리, 영상/비디오 검색, HDTV 등

이 태 흥(Taihong Yi)

정회원

vol.25, no.1B 참조