

고유벡터 분석을 이용한 조명 불변 영상 검색

김 용 훈, 이 태 흥

영남대학교 전자공학과 영상통신연구실

Illumination Invariant Image Retrieval using Eigenvector Analysis

Yonghun Kim, Taihong Yi

Visual Communications Lab., Department of Electronic Engineering, Yeungnam University

yhkim@lily.ee.yeungnam.ac.kr taehong@yu.ac.kr

요 약

본 논문에서는 조명의 변화에 의해 컬러 영상의 컬러 성분이 달라지더라도 영상 내 컬러간의 편차값을 나타내는 공분산 행렬(covariance matrix)의 고유벡터(eigenvector)와 영상 내 화소들의 컬러 성분과의 상관관계는 거의 변화하지 않는 특징을 이용한 조명 변화에 강인한 영상 검색 방법을 제안한다. 제안된 방법은 영상에서 컬러 성분들의 공분산 행렬과 공분산 행렬의 고유치(eigenvalue), 고유벡터를 계산한 후, 가장 큰 고유치에 관계된 고유벡터로 화소를 투영시키고, 투영된 벡터의 크기 성분으로 영상을 재구성한다. 재구성된 영상으로부터 7개의 불변 모멘트(moment)를 계산하고, 공분산의 가장 큰 고유치를 가중치로 부과하여 특징벡터를 추출한다. 7개의 불변 모멘트로부터 구한 특징벡터는 영상 내 물체의 이동, 영상의 회전, 크기 변화뿐만 아니라, 조명의 변화에 의해 컬러가 변화할 경우에도 유사한 영상을 잘 검색한다. 제안된 방법의 성능 확인을 위하여 5가지 조명에서 얻은 영상 데이터베이스를 이용하여 실험하였으며, 실험 결과 히스토그램 인터섹션에 비해 적은 특징량으로 검색이 가능하면서 조명 변화에도 대응할 수 있는 검색 결과를 얻을 수 있었다.

I. 서 론

멀티미디어 기술의 발달과 인터넷의 대중화로 인해 정보 검색에 대한 수요가 증가함에 따라 개인이 다물

수 있는 정보량이 기하 급수적으로 늘어났고, 그 형태도 텍스트 위주에서 멀티미디어 형태로 바뀌어 가고 있다. 이에 대용량 데이터에서 원하는 정보를 보다 빠르고 쉽게 검색할 수 있는 검색 기법이 필요하게 되었고, 그 방법 중의 하나로 멀티미디어 데이터 내의 내용으로부터 직접 정보를 이끌어 내어 사용할 수 있는 내용 기반 검색 기법이 현실적으로 많이 연구되고 있다.

내용 기반 영상 검색 기법 중, 특히 컬러를 이용한 기법은 Swain[1]등에 의해 히스토그램 인터섹션(histogram intersection) 기법이 발표된 이후, 구현의 용이함 등으로 인해 Mehmet[2]의 RTS 불변 컬러 인덱싱 등의 많은 연구가 되어 왔고, 현재 IBM의 QBIC[3]과 UC Berkeley의 Chabot[4]등 다양한 시스템에 사용되고 있다. 컬러 히스토그램을 이용할 경우, 계산이 간단하고, 카메라 위치의 미세한 변화에 강하다는 장점이 있다. 그러나, 이 방법은 동일 영상에 대해 조명 변화가 있을 경우, 히스토그램이 영향을 많이 받아 검색이 어렵다는 단점을 지니고 있다.

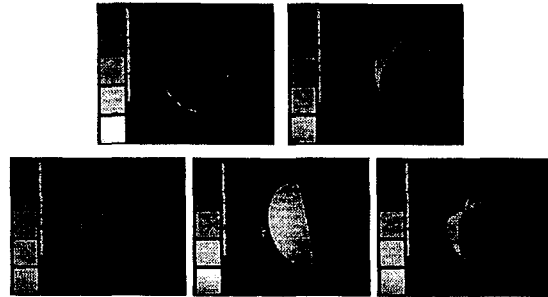


그림 1. 다양한 조명하에서의 컬러 영상

히스토그램에 기반한 방법들은 그림 1과 같이 조명 등 외부 요인에 의해 영상의 컬러 속성이 변형되는 경우에는 검색이 제대로 되지 않는 단점이 있다. 본 논문에서는 위와 같이 조명의 변화로 인해 컬러의 속성이 변형된 영상을 질의 영상으로 사용하는 경우에도 유사한 컬러 영상을 검색함을 목적으로 한다.

특징벡터의 추출을 위해 RGB 컬러 모델을 사용하여 컬러 성분들의 공분산 행렬을 구하고, 공분산 행렬로부터 고유치와 고유벡터를 계산한다. 그리고, 가장 큰 고유치에 대응하는 고유벡터를 화소를 투영시키고, 투영된 벡터의 크기 성분으로 영상을 재구성한다. 재구성된 영상으로부터 7개의 불변 모멘트를 계산하고, 공분산의 가장 큰 고유치를 기준으로 부과하여 특징벡터를 추출한다.

논문의 구성은 다음과 같다. 2장에서는 본 논문에서 제안한 방법의 특징벡터의 추출 방법과 유사도의 정의에 대해 설명하고, 3장에서는 실험 및 결과에 대해 알아본다. 마지막으로 4장에서 결론을 맺는다.

II. 고유벡터 분석을 이용한 조명 불변 영상 검색

영상에서 화소의 컬러 성분은 화소의 위치에 대한 벡터 함수 $X(i, j) = [R(i, j) G(i, j) B(i, j)]^T$ 로 표현된다. 여기서, T 는 전치 행렬이다. 영상 내 화소들의 컬러 성분의 평균을 나타내는 평균 컬러 벡터 m_X 와 컬러 성분에 대한 공분산 행렬 C_X 는 다음과 같이 정의된다.

$$m_X = E\{X\} \quad (1)$$

$$C_X = E\{(X - m)(X - m)^T\} \quad (2)$$

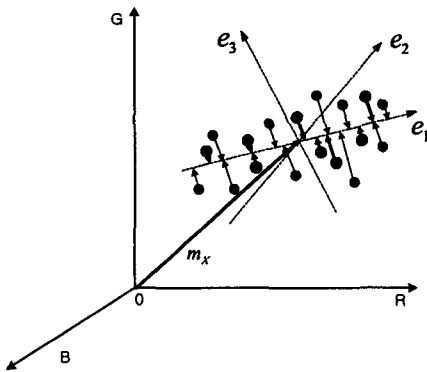


그림 2. 컬러 모델상의 고유벡터

영상 내 화소들의 컬러 성분에 대한 공분산 행렬 C_X 는 실수값을 가지는 3×3 대칭 행렬이기 때문에, 3개의 정규 직교인 고유벡터를 구하는 것은 항상 가능하다[5]. 공분산 행렬 C_X 의 고유치와 고유벡터는 다음의 수식으로 정의된다. 여기서, I 는 3×3 단위 행렬이다.

$$|C_X - \lambda I| = 0 \quad (3)$$

$$|C_X - \lambda I| e = 0 \quad (4)$$

식(3)에서 구한 $i=1, 2, 3$ 인 고유치 λ_i 를 가장 큰 값에서 작은 순으로 $\lambda_1 > \lambda_2 > \lambda_3$ 가 되도록 내림차순으로 정렬한다. 영상에서의 컬러 분포는 가장 큰 고유치 λ_1 에 대응하는 고유벡터 e_1 방향으로 집중된다. 이것은 가장 큰 고유치 λ_1 에 대응하는 고유벡터 e_1 은 컬러 영상의 컬러 정보를 가장 많이 포함하고 있다는 것을 의미하고 가장 작은 고유치에 대한 고유벡터는 가장 작은 컬러 정보를 담고 있다는 것을 의미한다. 그러므로, 그림 2와 같이 컬러 정보를 가장 많이 포함하고 있는 고유벡터 e_1 으로 화소의 컬러를 투영함으로써 데이터량을 효과적으로 줄일 수 있다.

영상에서 화소의 컬러 $X(i, j)$ 를 가장 큰 고유치 λ_1 에 대응하는 고유벡터 e_1 상으로 투영한 투영벡터 P_X 는 식(5)와 같이 표현되고, 투영된 벡터의 크기 성분으로 재구성된 영상은 식(6)과 같이 표현된다.

$$P_X(i, j) = e_1^T(X(i, j))e_1 \quad (5)$$

$$f_X(i, j) = \text{norm}\{P_X(i, j)\} \quad (6)$$

2차원 연속함수 $f(i, j)$ 에 대해, 차수 $(p+q)$ 를 가지는 모멘트는 다음과 같이 정의된다.

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} i^p j^q f(i, j) di dj \quad (7)$$

여기서, $p, q=0, 1, 2, \dots$ 이다.

유일성 정리에 의하여 만약 $f(i, j)$ 가 부분적으로 연속적이고, ij 평면의 유한한 구간 내에서만 0이 아닌 값을 가지면, 모든 차수의 모멘트가 존재하고 그 모멘트 열 m_{pq} 는 $f(i, j)$ 에 의해 유일하게 결정되어진다. 이와 반대로, m_{pq} 는 유일하게 $f(i, j)$ 를 결정한다[5].

디지털 영상에서 중앙 모멘트는 다음 수식으로 표현된다.

$$\mu_{pq} = \sum_i \sum_j (i - \bar{i})^p (j - \bar{j})^q f_X(i, j) \quad (8)$$

정규화된 중앙 모멘트는 η_{pq} 로 표기하며, 다음과 같이 정의된다.

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\frac{p+q}{2}}} \quad (9)$$

여기서, $\gamma = ((p+q)/2) + 1$ 이고, $p+q=2, 3, \dots$ 이다.

7개의 불변 모멘트의 집합은 두 번째와 세 번째 모멘트로부터 식(10)과 같이 유도된다[5]. 이 모멘트의 집합은 영상 내 물체의 이동, 영상의 회전, 크기 변화에 대해 불변이다.

$$\phi_1 = \eta_{20} + \eta_{02} \quad (10-a)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (10-b)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (10-c)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (10-d)$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (10-e)$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (10-f)$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (10-g)$$

식(10)에서 구한 7개의 불변 모멘트들의 생동폭을 감소시키기 위해 식(11)과 같이 로그를 취하고, 가장 큰 고유치 λ_1 을 로그를 취해 가중치로 적용한다. 본 논문에서 제안하는 영상의 특징벡터 F 는 식(12)와 같이 정의한다.

$$\phi_i^* = \log(\lambda_1) \times \log(\phi_i) \quad i=1, 2, \dots, 7 \quad (11)$$

$$F = [\phi_1^*, \phi_2^*, \phi_3^*, \phi_4^*, \phi_5^*, \phi_6^*, \phi_7^*] \quad (12)$$

컬러 영상의 검색은 질의 영상이나 데이터베이스내의 영상에 대해 특징벡터 F 를 추출하여 그 특징벡터들을 가지고 유사도를 측정하게 된다. 두 영상 사이의 유사도 S 를 구하기 위해 제안하는 방법은 특징벡터 F 의 성분들 사이의 상대적인 차이를 나타내는 d_1 거리 함수를 사용하여 유사도를 측정한다.

$$S(F, \bar{F}) = \sum_{i=1}^7 \frac{|\phi_i^* - \bar{\phi}_i^*|}{1 + \phi_i^* + \bar{\phi}_i^*} \quad (13)$$

식(13)에 나타낸 유사도 S 는 질의 영상 F 와 데이터베이스 내의 영상 \bar{F} 가 비슷할수록 0에 가까운 값을 가지게 되고, 다른 영상일 경우 상대적으로 큰 값을 가지게 된다.

III. 실험 및 결과

조명의 변화 등에 의해 영상 내 컬러의 속성이 변형되는 환경에서 제안된 방법의 검색 성능을 확인하기 위하여 본 실험에서는 Simon Fraser University에서 컬러 연구를 위해 사용한 영상 데이터베이스를 사용하였다. 데이터베이스는 그림 1과 같이 하나의 물체 당 5가지의 조명을 이용하여 촬영한 영상을 하나의 그룹으로 하여 3 그룹(A, B, C 그룹), 총 15장의 영상으로 구성하였다. 각 그룹에 속해 있는 1~5까지 5장의 영상은 모두 다른 5가지의 조명 아래에서 각각 촬영된 것으로 각 번호별로 실험에 사용된 조명의 이름을 표 1에 나타내었다.

표 1. 실험에 사용된 조명 이름

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

제안된 특징벡터와 d_1 거리합수를 이용하여 측정된 각 영상들 사이의 유사도를 표 2에 나타내었다. 표 2에서 좌측의 영상과 위쪽의 영상이 수직으로 마주치는 위치의 값이 두 영상 사이의 유사도이다. II장에서 언급한 바와 같이, 두 영상 사이의 유사도는 두 영상이 동일할 경우 0이 되고, 두 영상이 다를 경우에는 상대적으로 큰 값을 갖는다. 표 2에서 동일 그룹간의 유사도를 살펴보면 대부분의 값이 0.3 이하로 매우 작음을 알 수 있다. A 그룹간의 유사도는 최대 0.437이고, B 그룹은 최대 0.152, 그리고, C 그룹은 최대 0.403이다. 반면에 다른 그룹과의 유사도는 상대적으로 매우 크다는 것을 알 수 있다. A 그룹과 B, C 그룹과의 유사도는 거의 1 이상이고, B 그룹과 C 그룹과의 유사도 또한 동일 그룹간의 유사도에 비해 큰 값을 가진다. 이러한 사실은 제안된 특징벡터가 조명의 변화로 인해 컬러의 속성이 변형된 영상의 경우에도 조명의 변화에 큰 영향을 받지 않고 유사한 영상을 잘 검색해 낸다는 것을 의미한다.

IV. 결론

본 논문에서는 조명 변화 등 외부 요인에 의해 컬러의 속성이 변형된 경우에도 유사한 영상을 검색할 수 있는 새로운 방법을 제안하였다.

제안된 방법은 영상 내 컬러 성분들의 공분산 행렬을 구한 후, 고유치와 고유벡터를 계산하고, 가장 큰 고유치에 대응하는 고유벡터를 화소를 투영시킨 투영 벡터의 크기 성분으로부터 7개의 불변 모멘트를 계산하고, 공분산의 가장 큰 고유치를 가중치로 부과하여 특징벡터를 추출하였다.

다양한 조명을 이용하여 촬영한 영상을 대상으로 실험을 진행하여 제안된 특징벡터가 조명 변화에 강한 특성을 가짐을 나타내었다.

참고 문헌

- [1] Michael J. Swain and Dana H. Ballard, "Color Indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11-32, 1991.
- [2] Mehmet Celenk and Yuan Shao. "Rotation, translation, and scaling invariant color image indexing," *IS&T/SPIE Conference on Storage and Retrieval for Image and Video Databases VII*, pp. 623-630, January, 1999.
- [3] Myron Flickner *et al*, "Query by Image and Video Content : The QBIC System," *IEEE Computer*, vol. 28, no. 9, pp. 23-32, September, 1995.
- [4] Virginia E. Ogle and Michael Stonebraker, "Chabot : Retrieval from a Relational Database of Images," *IEEE Computer*, vol. 28, no. 9, 1995.
- [5] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison-Wesley, 1992.

표 2. 영상들간의 유사도 측정

		A 그룹					B 그룹					C 그룹				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
A 그 룹	1	0.000	0.276	0.162	0.390	0.226	1.216	1.201	1.070	1.066	1.066	0.856	0.708	0.496	0.451	0.639
	2		0.000	0.437	0.364	0.201	0.856	0.931	0.799	0.795	0.795	0.583	0.434	0.296	0.245	0.365
	3			0.000	0.229	0.064	1.283	1.357	1.227	1.224	1.224	1.015	0.868	0.656	0.612	0.799
	4				0.000	0.173	1.502	1.575	1.446	1.443	1.444	1.238	1.092	0.882	0.839	1.024
	5					0.000	1.345	1.418	1.289	1.285	1.286	1.078	0.931	0.719	0.676	0.862
B 그 룹	1						0.000	0.089	0.067	0.061	0.067	0.281	0.424	0.636	0.680	0.494
	2							0.000	0.135	0.138	0.152	0.366	0.501	0.712	0.755	0.570
	3								0.000	0.035	0.094	0.279	0.367	0.579	0.622	0.436
	4									0.000	0.058	0.259	0.363	0.575	0.619	0.433
	5										0.000	0.215	0.363	0.576	0.619	0.433
C 그 룹	1											0.000	0.150	0.363	0.407	0.220
	2												0.000	0.270	0.325	0.126
	3													0.000	0.138	0.187
	4														0.000	0.205
	5															0.000