

투영된 DC히스토그램의 거리근사를 이용한 이미지 검색 기법

정희원 강응관*, 최종수*

Image Retrieval Using the Distance Approximation of the Projected DC Histogram

Eung-kwan Kang*, Jong-soo Choi* *Regular Members*

요 약

본 논문에서는 JPEG/MPEG과 같은 블록 기반의 영상 압축 기법에 사용되는 DCT변환 부호화를 이용하여 블록의 평균을 의미하는 DC계수를 추출하고, 이로부터 구성된 DC영상을 수평/수직 방향으로 투영시킴으로써 얻어지는 영상의 위치 정보를 히스토그램 분포 특성으로 변환시켜 데이터베이스의 색인 정보로 저장한 후, 입력되는 질의 영상에 대해 동일 처리를 수행한 결과의 히스토그램 그래프와의 거리 근사를 비교하여 유사도를 측정함으로써 영상을 검색하는 방법을 제시한다. 실험 결과 제안된 방법이 검색에 있어 우수한 성능을 갖추고 또한 상당한 양의 처리 시간과 메모리 공간을 줄일 수 있음을 확인하였다. 향후 제안한 방법은 색상과 같은 다른 색인 정보와 결합할 경우 보다 나은 영상 색인과 검색 수단을 제공할 것이며, 또한 비디오 장면 전환 검출에 유용하게 이용할 수 있다.

ABSTRACT

This paper presents a new indexing technique, for image retrieval, which calculates the similarity of distance between projected DC histogram in database and query histogram result from the same process. By DCT, which is essential procedure in block-based JPEG and/or MPEG standard, DC image can be constructed directly in compressed domain and this represents the average of block segmented image. We can obtain the positional information by projection with horizontal and vertical direction, and map this information into histogram that is used as database indices. Distance approximation of the projected DC histogram is used as a matching criterion. From the simulation results, the proposed method clearly show the validity and superiority in respect of computation time and memory space, and that in conjunction with other techniques for indexing, such as color, can provide a powerful framework for image indexing and retrieval. Furthermore, the proposed method can be applied to video scene change detection usefully.

I. 서론

최근 반도체 및 통신장비 기술의 발달과 더불어 사용자 세대의 변화에 따라 정보기기가 다루는 정보의 형태가 기존의 문자나 음성 위주에서 벗어나 영상이나 비디오, 오디오 등을 중심으로 한 종합적인

멀티미디어 형태로 전환되고 있으며, 또한 관련 응용 분야도 점차 확대되는 추세이다. 이에 따라 지금까지의 MPEG-1, -2, -4와 같은 디지털 동영상 및 오디오 압축 부호화의 개념으로부터 벗어나, 효율적인 데이터의 저장 및 관리, 검색 등을 위해 정보의 내용물을 갖고 대상을 표현하고자 하는 MPEG-7^[1]에 대한 표

* 중앙대학교 전자공학과(jschoi@candy.ee.cau.ac.kr)
논문번호 : 98380-0826, 접수일자 : 1998년 8월 26일

준 작업이 곧 진행될 예정이다.

한편, 인간의 정보 취득에 중요한 역할을 담당하는 시각 정보는 정보 전달 효과가 음성이나 문자에 비해 뛰어나기 때문에 영상 정보에 대한 수요는 날이 갈수록 급증하고 있으며, 참고 문헌^[2]에 따르면 현재 까지 제작된 비디오 필름의 양은 대략 6백만 시간 분량이 넘고, 해마다 약 10% 이상의 증가를 거듭하고 있으며 이는 디지털화 할 경우 1.8 million GB의 MPEG 부호화된 비디오 데이터량에 해당한다. 따라서 내용 기반의 대상 표현을 함에 있어, 효율적인 정보 처리 및 관리 체계 없이는 이러한 엄청난 양의 비디오 데이터 증가에 효과적으로 대처할 수가 없다.

과거의 비슷한 예로서 현재 일반적인 문서의 정형화된 표현 수단으로 색인 및 검색에 널리 이용되는 문자 위주의 데이터베이스 관리 시스템 (DBMS : Database Management System)의 필요성이 대두되었던 것과 일맥상통한다고 볼 수 있다. 그러나 기존 DBMS의 경우 대상 자체가 단순히 구조화된 텍스트 데이터 형태를 위해 설계되었는데, 영상 혹은 비디오 데이터의 경우 그 특성이 기본적으로 기존의 문자 위주와는 다르고, 데이터 모델링, 특징 추출 및 인덱싱 (indexing) 등 여러 가지 측면에서 새로운 처리 기법들을 요한다^[3]. 현재까지 보편적으로 사용되고 있는 멀티미디어정보 검색 방식으로는 텍스트 기반과 내용 기반의 두 경우로 크게 분류할 수 있으며, 특히 압축된 데이터의 검색에 사용되는 색인 정보를 DCT (Discrete Cosine Transform), 웨이블릿 변환 (Wavelet Transform), 벡터 양자화 (Vector Quantization), 움직임 벡터 (motion vector) 등을 이용하는 방식이 제안되었다^[4].

본 논문에서는 DCT의 직교 변환 특성을 이용하여 영상의 대표 평균값, 즉 DC 성분만으로 원영상에 대해 축소된 형태의 DC 영상을 재구성하여 이를 1차원으로 투영시킨 결과의 히스토그램 분포를 데이터베이스의 색인 정보로 저장하고, 입력되는 질의 영상에 대한 히스토그램 분포와의 거리 근사를 비교함으로써 영상을 검색하는 기법을 제시한다.

논문의 구성은 다음과 같다. 먼저 2장에서 DCT 및 투영 (projection)에 대한 간략한 이론적 배경을 고찰한 후, 3장에서는 투영된 DC히스토그램의 거리근사를 이용한 이미지 검색 기법에 대해 설명하고 DCT 방식을 본 논문에서 취하는 근본 이유 및 타당성에 대해 분석하며 또한 이를 투영과 접목시켰을 때의 효용성에 대해 살펴본다. 4장에서는 제안한 방식을 적용한 실험 결과 및 분석 내용을 검토하고, 5장에서

결론 및 향후 연구 방향에 대해 언급한다.

II. 이론적 배경

영상을 일정한 크기의 블록으로 나누어 각 블록에 대해 변환 처리를 행함으로써 블록내의 평균값 (DC 성분)으로부터 최고주파 성분에 이르기까지 여러 주파수 성분으로 분해되는 과정을 직교 변환이라 하는데, 이와 같이 DCT는 공간상의 영상 데이터를 주파수상의 에너지 분포로 변환시키는 직교 변환이며, 일반적인 영상 데이터는 변환 후 저주파 영역으로 집중되는 경향이 있으므로 시각 특성 (psycho-visual)의 효과적인 이용을 용이하게 한다는 이점과 구현이 비교적 쉽다는 장점이 있다^[5]. DCT의 일반 정의식은 다음과 같다.

$$F(u, v) = \frac{2(C(u)C(v))}{\sqrt{MN}} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \cos \left[\frac{(2i+1)u\pi}{2M} \right] \cos \left[\frac{(2j+1)v\pi}{2N} \right] \quad (1)$$

여기서, $\begin{cases} u=0, 1, \dots, M-1 \\ v=0, 1, \dots, N-1 \end{cases}$

그리고 $\begin{cases} C(u) = C(v) = 1/\sqrt{2}, & u=v=0 \\ 1, & u=v \neq 0 \end{cases}$

한편, 원형 (original)의 고차원 데이터를 낮은 차원의 부 공간으로 투영시키는 기법은 특히 영상 분석에 있어 유용하다. 즉, $n \times n$ 의 2차원 영상을 n 개의 1차원 데이터 요소로 투영시킴으로써 상당한 양의 처리 시간과 메모리 공간을 줄일 수 있다. 단순 2진 영상에 대한 투영의 예를 그림1에 나타내었다.

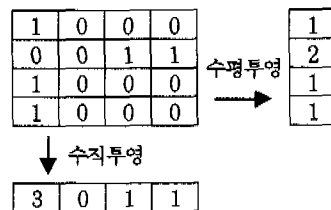


그림 1. 2진영상 투영의 예

III. 투영된 DC히스토그램의 거리근사를 이용한 이미지 검색 기법

본 장에서는 DCT변환 부호화를 이용하여 DC계수를 추출하고, 이로부터 구성된 DC영상의 투영된 히

스토그램 분포를 데이터베이스의 색인 정보로 저장한 후, 입력되는 질의 영상에 대해 동일 처리를 수행한 결과의 히스토그램 그래프와의 거리 근사를 비교하여 유사도를 측정 함으로써 영상을 검색하는 방법에 대해 기술하며, 처리되는 전체 과정은 그림2와 같다.

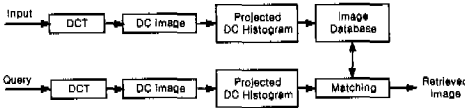


그림 2. 전체 블록도

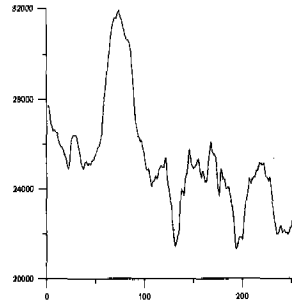
먼저 DC영상을 구성하는 방법으로 [6-8]은 영상을 8×8블록으로 나누어 블록의 단순 평균값을 취하는 방식을 제안하였다. 그러나 이들 방식은 DCT 수행에 따르는 계산량을 상당량 줄일 수 있는 반면, 실제 압축된 데이터 상태에서부터 DC영상을 구성하는 데는 제약이 따른다. 따라서 본 논문에서는 영상에 대해 2장에서 기술한 식(1)의 DCT를 수행한 후 $u=v=0$ 일 경우에 추출되는 DC계수들을 입력으로 받아들여 DC영상을 구성하는 방식을 취한다. 이는 실제 응용에 있어 압축된 데이터 즉, DCT 기반의 JPEG/MPEG 비트열로부터 별도의 처리나 복호화 과정을 거치지 않고도 바로 적용시킬 수 있다는 장점이 있다. 그림3은 256×256 크기의 lena영상에 대해 8×8 블록으로 DCT를 수행한 후 추출된 DC계수만으로 구성된 32×32 크기의 DC영상의 한 예를 나타낸 것이다.



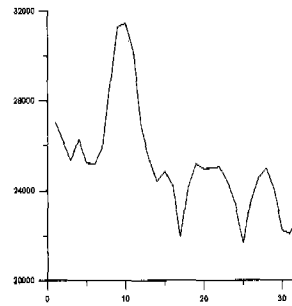
그림 3. (a) 원영상 (b) DC 영상

다음으로, 그림3에서 구성된 DC 영상으로부터 수평/수직 방향으로 투영시킨 결과의 히스토그램 분포를 그림4에 나타내었다. DCT를 수행한 결과의 DC값은 해당되는 블록을 대표할 수 있는 평균을 의미하므로 그림4에서 보듯이 DC 영상을 투영시킨 결과의 분포 특성은 원영상을 투영시킨 결과의 분포 특성과 거의 일치하는 것을 볼 수 있다. 이를 정리해 보면, DCT기반의 JPEG/MPEG 비트열로부터 DC계수를 바

로 추출할 수 있으므로 시스템 구현이 용이하고, DC 계수를 이용한 DC 영상은 원영상에 대해 공간적으로 DCT를 수행하는 블록 크기 N×N배로 축소된 형태이며, 이를 투영시킨 결과는 1차원적인 2N개의 데이터로 표현되므로 처리 시간이나 메모리 공간 측면에서 매우 효율적임을 알 수 있다.



(a)



(b)

그림 4. (a) 수평 투영된 원영상 히스토그램 (b) 수평 투영된 DC영상 히스토그램

한편, 히스토그램 분포에 대한 유사도를 비교하는 척도로 히스토그램 교차 함수 (histogram intersection measure) 방식[9]이 많이 이용되고, 이를 정규화 함수식은 다음과 같이 정의된다.

$$H(I, Q) = \frac{\sum_{i=1}^n \min(I_i, Q_i)}{\sum_{i=1}^n Q_i} \quad (2)$$

여기서, I_i , Q_i , 그리고 n 은 데이터베이스 영상의 히스토그램, 질의 영상의 히스토그램, 그리고 히스토그램 빈 (bin)의 개수를 나타낸다.

그러나, 이 수식에 있어서의 문제점은 히스토그램 분포가 서로 교차하지 않는 경우, 예를 들면 입력으

로 들어오는 질의 영상의 분포 특성이 데이터베이스의 히스토그램 분포에 비해 레벨이 아주 작을 경우에도 유사도는 1로 표현되므로, 같은 분포 특성으로 오인될 수 있다. 이러한 문제점을 해결하기 위해 본 논문에서는 수평/수직 방향으로 투영된 히스토그램 분포 특성을 거리 함수를 이용하여 서로 근사 시킨 후 산술 평균을 구함으로써 보다 정확한 정합 기준이 됨을 모의 실험을 통해 확인하였고, 수식은 다음과 같이 정의하였다.

$$H_{dis} = \frac{H_{hor} + H_{ver}}{2} = \frac{\sum_{i=1}^g \sqrt{(I_i - Q_i)^2}}{g \times M} + \frac{\sum_{j=1}^g \sqrt{(I_j - Q_j)^2}}{g \times N} \quad (3)$$

여기서, I_i, I_j, Q_i, Q_j, g 그리고 M, N 은 데이터베이스에 저장되어 있는 i 번째, j 번째 DC계수들의 투영 히스토그램, 입력으로 들어오는 질의 영상의 투영된 DC 히스토그램, 영상의 그레이 레벨, 그리고 투영축의 히스토그램 bin의 개수를 각각 나타낸다.

IV. 실험 및 고찰

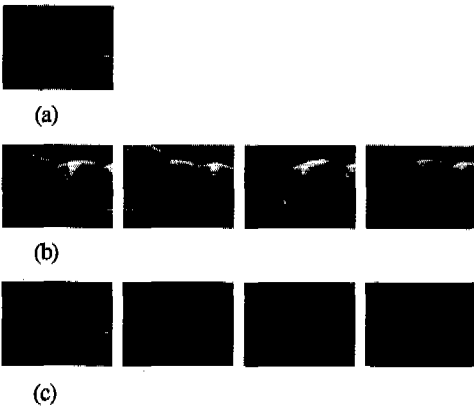


그림 5. 검색결과 (a) 질의영상 (b) [9]를 이용한 결과 (c) 거리근사를 이용한 결과

본 장에서는 제안한 영상 검색 기법의 성능을 평가하기 위해, 'Wallace and Gromit'이라는 240×176 영상 크기의 256 그레이 레벨 해상도를 갖는 200프레임의 영상을 갖고 실험한 결과에 대해 기술한다. 먼저, 그림5는 질의 영상이 입력되었을 때, DCT를 수행하여 DC 영상을 구성한 후 투영시킨 히스토그램과 데이터베이스에 저장된 투영된 DC 히스토그램

과의 유사도를 [9]의 평가 함수를 이용한 것과 본 논문에서 제안하는 거리 근사를 이용한 것을 가장 유사도가 높은 순으로 각각 나타낸 검색 결과이다.

그림6은 위의 검색 결과로 얻은, 그림5의 (b)와 (c) 영상 가운데 유사도가 높은 첫번째와 두번째 영상에 대한 수평 투영된 히스토그램 분포를 나타낸 것이고, 표1은 식(2)와 식(3)에 의해 계산된 유사도 결과이다.

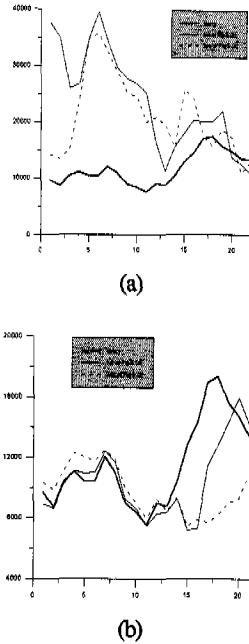


그림 6. 수평 투영된 DC 히스토그램 (a) 그림5의 (b)영상 (b) 그림5의 (c)영상

표 1. [9]와 거리 근사를 이용한 유사도

Query		Database				
		1 st	2 nd	3 rd	4 th	
수평 투영	[9]	그림5(b)	0.982067	0.981896	0.972355	0.966414
		그림5(c)	0.893603	0.807401	0.764364	0.652656
	H_{hor}	그림5(b)	49.124889	39.003796	43.836758	42.504105
		그림5(c)	5.891823	10.553089	11.114613	15.769242
수직 투영	[9]	그림5(b)	1.0	1.0	1.0	1.0
		그림5(c)	0.858841	0.830305	0.772442	0.652100
	H_{ver}	그림5(b)	47.499802	37.362328	41.391608	39.460660
		그림5(c)	6.630752	6.216829	7.613932	11.601026
산술 평균	[9]	그림5(b)	0.9910335	0.990943	0.9861775	0.983207
		그림5(c)	0.8762220	0.8188530	0.7684030	0.6523775
	H_{dis}	그림5(b)	48.3123455	38.1630620	42.5841630	40.9823325
		그림5(c)	6.2612925	6.3649590	9.3642725	13.6851340

위의 실험 결과를 토대로 내용을 분석해 보면, 그림6의 (a)와 같이 히스토그램의 레벨차가 클 경우, 즉 입력으로 들어오는 질의 영상의 투영 히스토그램 분포가 데이터베이스의 투영 히스토그램 분포에 비해 작을 경우에 식(2)에 근거한 [9]의 방식을 적용시킨 검색 결과는 실제로 사용자가 원하는 영상이 아님에도 유사도는 표1에서 보듯이 높게 나타나는 것을 볼 수 있다. 이에 비해 본 논문에서 제안하는 거리 근사를 이용한 검색 방법은 표1의 결과로부터 원하는 영상인 그림5의 (c)와 같은 영상을 얻을 수 있다.

그림7은 이에 대한 또 다른 실험으로 (a)와 같은 질의 영상이 입력되었을 때, 거리 근사를 이용하여 표2의 계산에 의해 가장 유사도가 높은 영상 순으로 검색되어 나타난 결과이다. 그 밖의 실험에서도 검색 결과가 뛰어나다는 것을 확인하였다. 따라서 이상의 실험 결과에 근거해 볼 때 히스토그램의 거리 근사를 이용한 유사도 평가 방법이 검색에 있어 보다 타당하고 우수한 성능을 갖는다는 것을 알 수 있다.

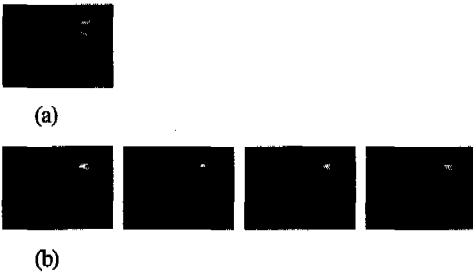


그림 7. 검색결과 (a) 질의영상 (b) 거리근사를 이용한 결과

표 2. 거리 근사를 이용한 유사도

DataBase Query	1 st	2 nd	3 rd	4 th
수평 투영	5.496604	8.650723	8.865479	9.530140
수직 투영	9.021506	7.487549	7.784424	9.154386
산술평균	7.2590550	8.0691360	8.3249515	9.3422630

V. 결 론

본 논문에서는 DCT를 이용하여 DC계수를 추출하고, 이로 구성된 DC영상을 수평/수직 방향으로 투영 시킴으로써 얻어지는 영상의 위치 정보를 히스토그램 분포 특성으로 변환시킨 후 거리 근사를 비교함으로써 영상을 검색하는 방법을 제시하였다. 모의 실험 결과 제안된 방법이 검색에 있어 우수한 성능을

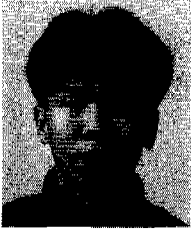
갖추고 있으며, 또한 투영된 DC 성분을 이용함으로써 상당한 양의 처리 시간과 메모리 공간을 줄일 수 있는 장점이 있다. 향후 제안한 방법은 색상과 같은 다른 색인 정보와 결합할 경우 보다 우수한 영상 색인 및 검색 수단을 제공할 것이며, 또한 비디오 장면 전환 검출에도 유용하게 이용할 수 있을 것이다.

참 고 문 헌

- [1] ISO/IEC JTC1/SC29/WG11 N2082 MPEG-7 Context and Objectives
- [2] Hampapur, A., "Design video data management systems," *Ph.D thesis*, The Univ. of Michigan, 1994.
- [3] Elmagarmid, A. K. et al., "Video database systems, issues, products, and applications," *Kluwer Academic Publishers*, 1997.
- [4] Mandal, M. K. et al., "Image and video indexing in the compressed domain," *SPIE Proceedings: Multimedia Storage and Archiving Systems II*, Vol. 3229, pp. 2 ~ 13, 1997.
- [5] Rao, K. R. and Yip, P., *Discrete Cosine Transform, Algorithms, Advantages, Applications*, Academic Press, 1990.
- [6] Yeung, M. M. and Liu, B., "Efficient matching and clustering of video shots," *IEEE ICIP-95*, Vol. I, pp. 388 ~ 341.
- [7] Yeo, B. L. and Liu, B., "On the extraction of dc sequence from MPEG compressed video," *IEEE ICIP-95*, Vol. II, pp. 260 ~ 263.
- [8] Yeo, B. L. and Liu, B., "Rapid scene analysis on compressed video," *IEEE Trans. Circuits and Systems for Video Technology*, Vol. 5, No. 6, pp. 533 ~ 544, 1995.
- [9] Swain, M. J. and Ballard, D. H., "Color indexing," *Int. J. Computer Vision*, Vol. 7, No. 1, pp. 11 ~ 32, 1991.

강 응 관(Eung-Kwan Kang)

정회원



1993년 2월 : 중앙대학교 전자공학과 졸업(공학사)

1995년 2월 : 중앙대학교 대학원 전자공학과 졸업(공학석사)

1995년 1월~1997년 5월 : 현대 전자 정보통신연구소 연구원

1997년 9월~현재 : 중앙대학교 대학원 전자공학과 박사과정 재학 중

<주관심 분야> 동영상 부호화, 멀티미디어 데이터베이스, 영상통신, 컴퓨터비전

최 종 수(Jong-Soo Choi)

정회원

1975년 2월 : 인하대학교 전기공학과 졸업(공학사)

1977년 2월 : 서울대학교 대학원 전자공학과 졸업(공학석사)

1981년 2월 : 일본 Keio대학 전기공학과 졸업(공학박사)

1981년 2월~8월 : Aloka(주) 연구소 연구원

1981년 9월~현재 : 중앙대학교 전자공학과 교수

<주관심 분야> 적외선 신호처리, 컴퓨터 시각, 영상 압축 및 영상처리 등임