

가변 블록 DCT 를 이용한 영상 검색 시스템

김동우*, 서은주*, 장인동*, 한재혁* 안재형*
*충북대학교 정보통신공학과

e-mail : dubssi@chollian.net

Image Retrieving System Using Variable Block DCT

Dongwoo Kim*, EunJu Seo*, UnDong Jang*, JeaHyeck Han*, JaeHyeong Ahn*
*Dept. of Computer & Communication Engineering, Chungbuk Univ.

요 약

영상과 같은 다양하고 복잡한 데이터 검색은 기존의 키워드를 이용한 검색이 아닌 내용 기반 검색 방법이 요구된다. 본문은 JPEG/MPEG 과 같은 영상 압축 기법에서 사용되는 DCT 변환 부호화를 이용하여 각 블록의 평균인 DC 성분을 얻어 이 정보를 이용하여 영상을 검색하는 방법을 제시한다. 기존의 방법은 단순히 DCT 전체 성분을 이용하는 방법으로, 영상 전체의 DCT 정보를 처리하여야 하므로 시간과 메모리의 낭비가 있을 수 있다. 본문에서는 이러한 처리시간과 메모리의 낭비를 막기 위해 DC 성분만을 이용하여 영상 검색을 하며, 영상이 실제로 있는 부분에서는 블록의 크기를 가변적으로 변경하여 검색의 오류를 줄일 수 있다. 또한 검색을 전체 영상에 대한 빠른 1 차 검색과 동일 패턴을 가진 영상의 2 차 세부 검색으로 나누어 실행함으로써 보다 빠른 검색을 할 수 있다. 한편, 접쳐진 영상에 대해서도 검색이 가능하다.

1. 서론

현재 우리가 접하는 멀티미디어 정보 중 영상과 같이 다양한 형태의 데이터들은 그것과 연관된 텍스트나 키워드를 사용하는 기존의 검색방법으로는 그 영상을 제대로 묘사해 주지 못한다. 이들 검색 방법은 각 영상에 입력된 초기 텍스트 정보에 전적으로 의존하기 때문에 만약 질의어가 초기에 묘사되지 않은 영상속성을 참조하게 되면 검색은 대부분 실패하게 된다. 따라서 색상, 모양, 질감 등의 성분을 이용한 내용기반 검색방법이 필요하게 되었다.

현재까지 보편적으로 사용되고 있는 멀티미디어 정보의 내용기반 검색방법은 DCT(Discrete cosine transform), 웨이블릿 변환(Wavelet transform), 벡터 양자화(Vector quantization), 움직임 벡터

(Motion vector) 등을 이용하는 방식이 제안되었다.^[1]

제안 방법은 먼저 전체 영상에 대하여 영상이 있는 부분과 없는 부분을 나누어서 가변적으로 영상을 블록화 한 후 DCT 의 직교 변환 특성을 이용하여 영상의 대표 평균값인 DC(Direct current) 성분을 추출하여 데이터베이스의 색인 정보로 저장하고, 질의 영상에 대하여서 같은 처리과정을 거쳐서 1 차 전체 영상에 대한 빠른 검색을 시도하고, 검색이 안된 영상을 세부적으로 2 차 검색한다.

구성은 다음과 같다. 2 장에서 DCT 에 대한 간략한 이론적 배경을 살펴보고, 3 장에서는 가변블록 DCT 를 이용한 내용기반 검색 기법에 대하여 설명한다. 4 장에서 결론 및 향후 연구 방향에 대하여 언급하겠다.

2. DCT(Discrete cosine transform)

영상을 일정한 크기의 블록으로 나누어서 각 블록에 대해 변환 처리를 함으로써 블록내의 평균값(DC)으로부터 최고주파 성분에 이르기까지 여러 주파수 성분으로 분해하는 과정을 직교변환이라고 한다.

DCT는 직교변환을 이용하여 데이터를 특정한 성분으로 집중하여 데이터를 압축하며, JPEG과 MPEG 등 영상 압축 표준에 사용되는 기초 알고리즘으로 DFT(Discrete Fourier transform)와 같은 종류의 주파수 변환 방법이다. DCT의 수행과정은 영상데이터를 MxN의 블록으로 나누고 이 블록내의 화소에 대해서 DCT 연산을 행하여 원 영상을 2차원 주파수 평면에 나타내는 것이다.^[2]

$$B_{pq} = \frac{2C(p)C(q)}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos\left[\frac{(2m+1)p\pi}{2M}\right] \cos\left[\frac{(2n+1)q\pi}{2N}\right] \quad (2.1)$$

$$A_{mn} = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} \frac{2C(p)C(q)}{\sqrt{MN}} B_{pq} \cos\left[\frac{(2m+1)p\pi}{2M}\right] \cos\left[\frac{(2n+1)q\pi}{2N}\right] \quad (2.2)$$

여기서 $\begin{cases} p=0,1,\dots,M-1 \\ q=0,1,\dots,N-1 \end{cases} \begin{cases} C(p)=C(q)=1/\sqrt{2}, p=q=0 \\ C(p)=C(q)=1, p=q \neq 0 \end{cases}$

식 (2.1)은 일반적인 2차원 MxN DCT의 수식으로 원 영상의 특정 화소값 A_{mn} 에 대한 함수값 B_{pq} 를 구하는 공식이다. 식 (2.2)는 주파수 공간으로 변환된 함수값 B_{pq} 에서 원 영상 함수값 A_{mn} 을 구하는 역 DCT 수식이다.

그림 2.1의 4x4 DCT 변환 예에서처럼 영상에 대해 DCT를 수행할 경우 대부분의 중요한 정보는 일반적으로 저주파 영역(굵은실선, 특히 DC)에 집중 분포하기 때문에 저주파 영역의 정보만 저장하는 방식으로 영상을 압축할 수 있다. 또한 DFT와 달리 복소수 계산을 하지 않는 계산상의 장점과 구현이 비교적 쉽다는 장점이 있다.^[3]

한편, DCT는 그 계수의 분포 특성이 점진적이므로 특징 추출 과정에서 많이 사용되고 있다. 특히 DC부분의 정보만으로 전체 영상의 특징을 표현할 수 있으므로 전체 영상을 DCT 처리 후 DC 성분만을 데이터 베이스에 저장하여 영상 검색에 사용할 경우 우수한 성능을 보인다.^[4]

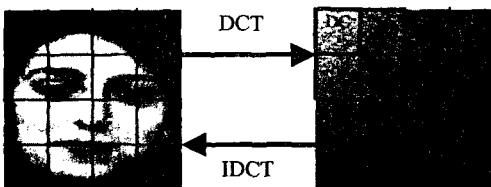


그림 2.1 4x4 DCT 변환 예

3. 가변블록 DCT를 이용한 내용기반검색

3.1 전체 시스템 구성

그림 3.1은 가변블록 DCT를 이용한 내용기반 검색 시스템의 전체 구성도이다.

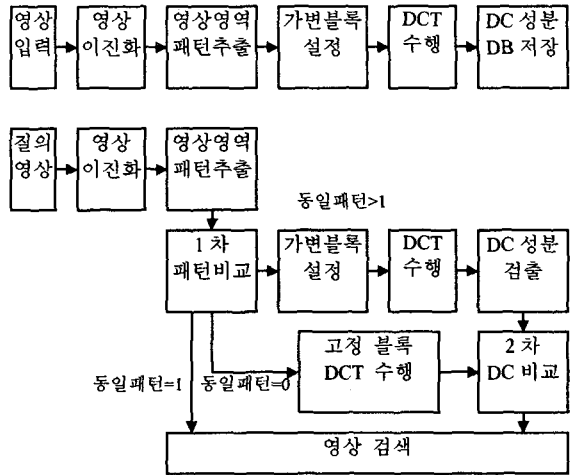


그림 3.1 가변블록 DCT 이용 내용기반검색 구성도

3.2 영상 정보 데이터베이스 생성

영상 정보 데이터베이스를 생성하기 위해 우선 입력된 영상의 전체 영상에 대하여 전처리 과정으로 이진화를 한다. 이진화는 그림 3.2처럼 256의 그레이 스케일의 영상을 0과 1의 값으로 이루어진 영상으로 만드는 과정이다.^[5]

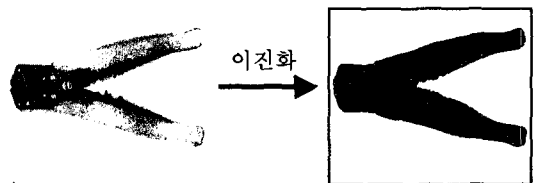


그림 3.2 영상의 이진화 결과

입력 영상이 이진화 되었으면 이진화 된 영상을 8x8 크기의 블록으로 나누어 각 블록에서 실제 영상 데이터(IMG_{exist} , 값이 1)의 합이 임계값($TH=32$) 이상이면 그 블록은 영상이 있는 영역(AR_{exist})으로 설정하고 IMG_{exist} 합이 임계값 이하이면 영상이 없는 영역($AR_{non-exist}$)으로 설정하게 된다.

128x128 영상에서는 16x16 개의 블록을 얻을 수 있으며 영상이 있는 영역 AR_{exist} 과 영상이 적게 들어 있거나 없는 영역 $AR_{non-exist}$ 을 첫번째 정보인 영상영역패턴 테이블에 저장한다. 또한 이 정보는 2차 검색에서 DCT를 가변블록 하기 위한 기준자료로 사용

하게 된다. 식 (3.1)은 영상영역패턴을 추출하는 수식이다.

$$\begin{cases} \text{if } \sum \text{IMG}_{\text{exit}} > \text{TH}, & AR_{\text{exist}} \\ \text{otherwise,} & AR_{\text{non-exist}} \end{cases} \quad (3.1)$$

다음 단계로 영상영역패턴 테이블이 완성 되었으면 이를 바탕으로 각 블록별 DCT를 수행하여야 한다. 기존의 방법은 영상영역패턴 테이블이 없으므로 전체 영상에 대하여 동일한 크기의 블록으로 DCT를 수행하여 블록크기에 따른 정보의 정확성과 계산량의 비례를 일정하게 정할 수 밖에 없었다.^[6] 그러나 본 문에서는 영상영역패턴 테이블을 미리 작성하였으므로 이것을 이용하여 원 영상에서 영상이 적거나 없는 영역인 $AR_{\text{non-exist}}$ 에서는 그대로 8x8 크기의 DCT를 수행하고, DCT를 진행하다가 영상이 많은 영역인 AR_{exist} 을 만나면 그 부분은 세부적으로 2x2 DCT를 수행하여 영상이 없는 부분에서는 빠른 처리를 수행하고, 영상이 실제로 있는 부분에서는 보다 세밀한 정보를 얻을 수 있다.

한편, DCT 후의 영상의 정보는 저주파 영역 특히 평균값인 DC에 집중되므로 모든 데이터를 정보로 얻는 것 보단 DC 값만으로 영상의 정보를 취하는 것이 보다 효과적인 이미지 검색을 할 수 있다.^[6] 이렇게 검출된 DC 정보를 영상 검색 데이터베이스에 저장하게 되면 영상 정보 데이터베이스의 생성이 완료 된다. 그림 3.3은 원 영상에서 AR_{exist} , $AR_{\text{non-exist}}$ 에 따른 가변블록 DCT를 수행한 후 정보의 양을 표현한 것이다.

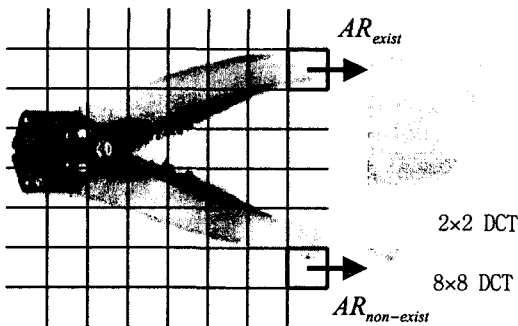


그림 3.3 가변블록 DCT의 적용

3.3 질의 영상 검색

질의 영상 검색 기법은 기본적으로 영상 정보 데이터베이스생성의 과정을 질의 영상을 통해 다시 거치면서 검색을 하게 된다.

우선 질의 영상이 들어오면 영상을 이진화 시킨 후 데이터베이스 생성과 동일하게 AR_{exist} , $AR_{\text{non-exist}}$ 정보를 16x16 개의 블록으로 갖는 영상영역패턴 테이블을 작성 할 수 있다.

질의 영상의 영상영역패턴 테이블이 얻어지면 그 시점에서 그림 3.1과 같이 1차 패턴 비교를 한다. 즉 영상 정보 데이터베이스의 영상영역패턴 테이블과 질의 영상의 영상영역패턴 테이블만을 비교하여 검색을 할 수 있다. 1차 패턴 비교는 3 가지 경우가 발생할 수 있다.

첫 번째는 검색을 통하여 동일 패턴 테이블을 가진 영상이 데이터베이스 상에 1개만 존재하는 경우로 원하는 영상을 1차 패턴비교에서 빠르게 동일 영상으로 검색할 수 있게 된다.

두 번째는 데이터베이스 상에 영상영역패턴 테이블의 값이 질의 영상과 동일한 정보가 여러 개 있는 경우로 질의 영상에 대해서도 영상 정보 데이터베이스 생성 과정과 같이 $AR_{\text{non-exist}}$ 에서는 8x8 DCT를 수행하며, AR_{exist} 에서는 보다 많은 정보를 얻기 위해 2x2 DCT를 수행하여 질의 영상에 대한 DC 정보를 얻은 후 동일 패턴 영상들에 대해서만 다시 그림 3.1과 같이 2차 검색을 시도하게 된다.

세 번째는 동일 패턴이 없는 경우로 동일한 영상은 없는 것으로 인식하고, 유사 영상을 검색하게 된다. 이때는 영상영역패턴 테이블을 통해 가변 블록을 이용할 수 없으므로 기존의 고정 블록 DCT를 이용하여 유사 영상을 검색하게 된다.

한편, 검색은 히스토그램 분포에 대한 유사도를 비교하는 척도로 히스토그램 교차 함수(histogram intersection measure) 방식^[7]이 많이 사용되고, 식 (3.2)은 이를 정규화 한 수식이다.

$$H(I, Q) = \frac{\sum_{i=1}^n \min(I_i, Q_i)}{\sum_{i=1}^n Q_i} \quad (3.2)$$

여기서, I_i , Q_i 그리고 n 은 데이터베이스 영상의 히스토그램, 질의 영상의 히스토그램, 그리고 히스토그램 bin의 개수를 나타낸다. 그러나 이 수식은 히스토그램 분포가 서로 교차하지 않는 경우, 예를 들면 입력으로 들어오는 질의 영상의 분포 특성이 데이터베이스의 히스토그램 분포에 비해 레벨이 아주 작을 경우 유사도가 1로 표현 되므로, 다른 영상을 같은 영상으로 오인할 수 있다.

이러한 문제점을 해결하기 위해서 수평/수직 방향으로 투영된 히스토그램 분포 특성을 거리 함수를 이용하여 서로 근사 시킨 후 산술 평균을 구함으로써

보다 정확한 검색 방법^[6]을 사용할 수 있고, 식 (3.3)과 같이 표현된다.

$$H_{dis} = \frac{H_{hor} + H_{ver}}{2} = \frac{\sum_{i=1}^M \sqrt{(I_i - Q_i)^2}}{g \times M} + \frac{\sum_{j=1}^N \sqrt{(I_j - Q_j)^2}}{g \times N} \quad (3.3)$$

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

4. 결론

본 문은 전체 영상에 대하여 영상이 있는 AR_{exist} 부분과 없는 $AR_{non-exist}$ 부분을 영상영역패턴 테이블을 만든 후 가변적으로 영상을 블록화 한 후 DCT의 직교 변환 특성을 이용하여 영상의 대표 평균값인 DC 성분을 추출하여 데이터베이스의 색인 정보로 저장하고, 질의 영상에 대하여서 같은 처리과정을 거쳐서 영상영역패턴 테이블을 이용하여 전체 영상에 대한 빠른 검색을 1차로 수행하고, 1차 과정에서 검색이 안된 영상을 가변 DCT 처리 후 DC 성분만을 가지고 수평/수직 방향으로 투영된 히스토그램 분포 특성을 거리함수를 이용하여 서로 근사 시켜 2차 검색하는 기법에 대하여 기술 했다.

제안 방법은 영상영역패턴 테이블만을 직접 비교하여 빠른 검색이 가능하며, 1차 검색이 실패 시에도 2차 검색에서 전체 데이터베이스가 아닌 동일한 패턴을 가진 영상에 대해서만 검색을 하므로 검색 속도가 빠르다. 또한 2차 검색도 DCT의 DC 성분만을 이용하므로 검색 속도와 메모리 효율성이 좋다. 한편, 영상이 없는 영역에서는 블록 크기를 크게 하여 속도를 높이고, 영상이 있는 영역에서는 블록 크기를 작게 가변적으로 하여 검색의 정확성을 높이는 효율적인 구조를 가진다.

앞으로 칼라 정보에 대한 연구가 추가 된다면 좀 더 복잡형태의 객체도 분류할 수 있는 시스템 구현이 가능하고, 이를 통해 여러 분야의 검색 시스템에서 다양한 활용이 기대 된다.

참고문헌

[1] Yun Q. Shi and Huifang Sun, "image and Video Compression for Multimedia Engineer," CRC Press, 2000
 [2] Jerry D. Gibson, Toby Berger, Tom Lookabaugh, Dave Lindbergh and Richard L. Baker, "Digital Compression

for Multimedia," Morgan Kaufmann Publishers, pp.228-243, 1998
 [3] Rao, K.R. and ip,P., "Discrete Cosine Transform, Algorithms, Advantages, Applications," Academic Press, 1990
 [4] Yasuo ARIKI and Yoshiaki SUGIYAMA, "Classification of TV Sports News by DCT Features using Multiple Subspace Method," IEEE TRANSACTIONS, vol.2 pp.1488-1491, 1998
 [5] Gregory A. Baxes, "Digital Image Processing," John Wiley & Sons Inc, pp.70-79, 1994
 [6] 강웅관, 최종수, "투영된 DC 히스토그램의 거리근사를 이용한 이미지 검색 기법," 한국통신학회논문지, vol.24, no.4B, pp741-745, 1999
 [7] Swain, M. J. and Ballard, D. H., "Color indexing," Int. J. Computer Vision, Vol.7, No.1, pp.11-32, 1991