

히스토그램 기반 영상 처리를 위한 압축영역에서의 고속 히스토그램 추출 기법

Fast Histogram Extraction Scheme for Histogram-based Image Processing

*박준형, **엄민영, ***최윤식

*Jun-Hyung Park, **Minyoung Eom, ***Yoonsik Choe

Abstract - Due to development of Internet network environments and data compression techniques, the size and amount of multimedia data has greatly increased. They are compressed before transmission or storage. Dealing with these compressed data such as video retrieval or indexing requires the decoding procedure most of the time. In video retrieval and indexing a color histogram is one of the most frequently used tools. We propose a novel scheme for extracting color histograms from images transformed into the compressed domain using 8x8 DCT(Discrete Cosine Transform). In this scheme an averaged version of original image is obtained by filtering DCT coefficients with a filter we designed.

Key Words : Histogram, Fast extraction, Compressed domain, DCT, Decoding

1. 서론

최근 들어 멀티미디어 데이터의 양과 수요가 급증하고 있고, 그 적용분야 또한 다양해짐에 따라, 데이터의 효율적인 조작과 처리의 중요성이 더욱 높아지고 있다. 이러한 멀티미디어 데이터는 저장과 전송의 효율성을 위해 압축영역으로 변환되고, 각종 화면을 통한 재생은 물론 편집 및 검색 등 대부분의 분야에 적용될 때 디코딩 과정을 거쳐서 공간 영역으로 전환되어 사용된다. 그러나 이 디코딩 과정은 많은 계산량과 긴 처리 시간을 요구하기 때문에 데이터 사용의 효율성을 떨어뜨린다고 할 수 있다. 멀티미디어 데이터의 처리에 쓰이는 대표적인 요소들 중 하나인 히스토그램을 얻어내기 위한 기존의 연구 중에 이러한 디코딩의 문제점을 해결하기 위해 8x8 DCT 계수로부터 디코딩 과정 없이 원영상의 averaged-version 영상을 복원해내는 기법과[1], 압축 영역의 계수들로부터 곧장 히스토그램을 얻어내는 기법[2], [3] 등이 있는데, 이들은 각각 복원 영상의 해상도가 원영상의 1/4에 한정된다는 한계와[1], 예지 히스토그램의 추출에 그 적용분야가 한정되어 있다는 한계가 있다[2], [3]. 본 논문에서는 히스토그램 추출을 위해 고속으로 공간 영역의 영상을 복원하는 기법을 제안한다. 이 기법은 8x8 DCT 계수의 선형 조합을 통해 averaged-version 영상을 복원해 내는 것으로서, 원래 해상도의 1/2인 영상의 복원까지도 가능하며 이는 다양한 히스토그램의 추출에 적용적으로 쓰일 수 있다. 2장에서는

제안 기법에 대해 자세히 설명하고 3장에서는 제안 기법을 통해 복원된 영상에 대한 실험 결과를 제시한다.

2. DCT 영역에서의 영상 추출

DCT는 공간 영역의 샘플을 주파수에 따라 분리하여 주파수 영역(압축 영역)의 계수로 변환시키기 때문에 DCT 계수는 quantization, VLC(Variable Length Coding) 등을 통해 압축하기에 효율적인 상태가 된다. 현재 주로 상용되고 있는 표준들, 이를테면 JPEG, MPEG 1, 2 등은 모두 이 DCT 과정을 포함하고 있다. 본 절에서는 8x8 DCT 계수로부터 디코딩 과정 없이 히스토그램 추출에 적합한 공간 영역의 영상을 복원하는 기법에 대해 설명한다.

2.1 원영상의 1/4 해상도 영상 추출

본 절에서는 8x8 DCT 계수의 선형 조합으로 원영상의 1/4 해상도(1/16 크기)에 해당하는 averaged-version 영상을 추출하는 기법에 대해 설명한다. 이 기법은 NxN DCT 계수 중 DC 계수가 공간 영역의 NxN 블록의 평균값에 상응한다는 사실에 근간을 둔다. 식 (1)에서 c_{00} , x_{ij} , m 은 각각 4x4 DCT의 DC 계수, (i, j)위치의 픽셀 값, 4x4 블록의 평균값을 나타낸다.

$$c_{00} = \frac{1}{4} \sum_{i=0}^3 \sum_{j=0}^3 x_{ij} \quad (1)$$

$$= \frac{1}{4} \cdot 16 \cdot m = 4m$$

8x8 DCT 계수로부터 m 을 얻어내는 기법을 살펴본다. X, D, C 를 각각 8x8 픽셀 블록, 8x8 DCT 매트릭스, 8x8

저자 소개

- * 박준형 : 연세大學 전기전자공학科 碩士課程
- ** 엄민영 : 연세大學 전기전자공학科 博士課程
- *** 최윤식 : 연세大學 전기전자공학科 正教授 · 工博
- ※ 본 연구는 한국 과학재단 목적 기초 연구(R01-2003-000-1 0412-0)지원으로 수행되었음.

DCT 계수라 하자.

$$DXD^T = C \Leftrightarrow D^T CD = X \quad (2)$$

식 (2)는 8x8 DCT와 IDCT 과정을 보여준다[4]. T는 transpose를 나타낸다. 한편, 매트릭스 F 를 그림 1(a)와 같이 정의하면,

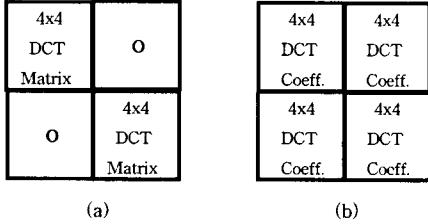


그림 1. (a) 매트릭스 F (b) 4x4 DCT 계수 E

식 (3)과 같이 X 에 F 를 취하여 E 를 얻을 수 있고, 식 (2)에 의해

$$FXF^T = E \quad (3)$$

$$FXF^T = F(D^T CD)F^T = (FD^T)C(DF^T) \quad (4)$$

$$= YCY^T = E$$

8x8 DCT 계수 C 로부터 $Y = [y_{ij}]$ 와의 연산을 통해 직접 E 를 얻을 수 있다. E 는 그림 1(b)에서 보듯이 4x4 DCT 계수 블록 네 개의 묶음이다. $E = [e_{ij}]$ 에서 네 개의 계수 $e_{00}, e_{04}, e_{40}, e_{44}$ 들은 각각 4x4 DCT DC 계수이므로 식 (1)과 같이 4x4 픽셀 블록의 평균값을 가지고 있게 되고, 결국 이것들을 취해 원영상의 1/4 해상도에 해당하는 averaged-version 영상을 얻어낼 수 있는 것이다.

여기서 매트릭스 곱 형태로 표현된 식 (4)를 식 (5)와 같이 고치면 8x8 DCT 계수 $C = [c_{ij}]$ 의 선형 조합으로 각 DC 계수 e_{nk} 를 표현할 수 있다.

$$e_{nk} = \sum_{i=0}^7 \sum_{j=0}^7 y_{kj} y_{ni} c_{ij}, \quad n, k \in \{0, 4\} \quad (5)$$

여기서 y_{ij} 와 f_{ij} 는 다음과 같다.

$$y_{ij} = \sum_{n=i}^{i+3} f_{in} d_{jn}, \quad F = [f_{ij}], \quad D = [d_{ij}] \quad (6)$$

$$f_{ij} = \sqrt{\frac{1}{2}} \cos\left(\frac{(2(j \bmod 4) + 1)(i \bmod 4)\pi}{8}\right) \quad (7)$$

2.2 원영상의 1/2 해상도 영상 추출

1/2 해상도(1/4 크기) 영상의 추출은 1/4 경우와 크게 다른 것이 없다. 매트릭스 F 가 그림 2. 와 같이 달라지면서 E 는

식 (8)과 같이 2x2 DCT 계수 블록 16개를 가지게 된다.

즉, 1/2의 경우 블록 하나당 16개의 DC 계수를 얻고 이것

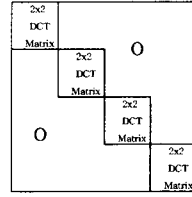


그림 2. 매트릭스 F

들이 가지고 있는 2x2 픽셀 블록의 평균값으로부터 1/2 해상도에 해당하는 영상을 복원하는 것이다.

$$e_{nk} = \sum_{i=0}^7 \sum_{j=0}^7 y_{kj} y_{ni} c_{ij}, \quad n, k \in \{0, 2, 4, 6\} \quad (8)$$

여기서 y_{ij} 와 f_{ij} 는 다음과 같다.

$$y_{ij} = \sum_{n=i}^{i+1} f_{in} d_{jn} \quad (9)$$

$$f_{ij} = \cos\left(\frac{(2(j \bmod 2) + 1)(i \bmod 2)\pi}{4}\right) \quad (10)$$

제안 기법에서 필요한 계산은 실제로 보이는 것보다 간단하다. 우선 식 (6)과 (9)처럼 $F = [f_{ij}]$ 가 포함되는 계산의 경우 F 는 각각 50%와 75% 만큼의 원소가 0이기 때문에 실제로 계산에 관여하는 부분은 8x8 중에서 50%와 25% 뿐이다. 그리고 E 에서 필요한 것은 4개 또는 16개의 DC 계수뿐이기 때문에 e_{nk} 의 계산 시 $Y = [y_{ij}]$ 은 8x8 중에 실제로는 $[y_{0*}], [y_{4*}]$ 2 행(1/2 해상도 경우 4 행) 만이 관여하는 것을 알 수 있다. 그리고 앞에서 언급했듯이 식 (5)와 (8)은 더 이상 2차원 매트릭스 계산이 아닌 64개 원소를 가진 1차원 벡터의 내적의 형태로 생각할 수 있고, 하나의 상수 ξ_{ij} 로 표현할 수 있는 $y_{kj} y_{ni}$ 부분은 미리 계산해서 가지고 있을 수 있기 때문에 한층 더 계산상의 복잡도가 낮아지게 된다. 그림 3은 미리 계산된 ξ_{ij} 중 하나씩을 8x8 DCT 계수 블록에 상응하는 매트릭스 형태로 나타낸 것이다. ξ_{ij} 의 성긴 구조 때문에 8x8 DCT 계수와의 내적은 최대 25개 또는 49개의 원소만을 필요로 함을 알 수 있다. 또한 경우에 따라 적용적으로 제한된 개수의 DCT 저주파 계수만을 사용하여 근사화 하는 것도 가능하다.

3. 실험 결과

본 절에서는 원영상과 복원 영상에서 얻어낸 컬러 히스토그램과 에지 히스토그램들을 비교한다. 히스토그램의 비교는 식 (11)의 히스토그램 유사도 측정법을 사용하였다.

$$S = \frac{H \cdot \hat{H}}{\sqrt{H \cdot H} \sqrt{\hat{H} \cdot \hat{H}}} \quad (11)$$

0.5	0.4731	0	-0.1931	0	0.2093	0	-0.0603
0.4574	0.4197	0	-0.1647	0	0.2063	0	-0.0817
0	0	0	0	0	0	0	0
-0.1551	-0.1442	0	0.0506	0	-0.0338	0	0.0287
0	0	0	0	0	0	0	0
0.1063	0.0883	0	-0.0338	0	0.0226	0	-0.0192
0	0	0	0	0	0	0	0
-0.0061	-0.0617	0	0.0207	0	-0.0102	0	0.0162

(a)

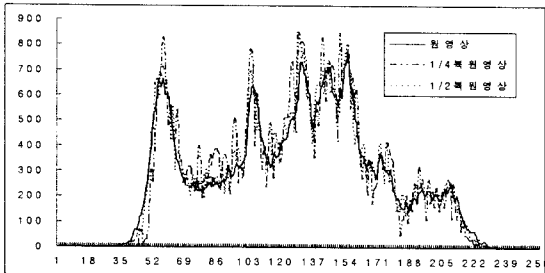
0.25	0.7204	0.2201	0.1135	0	-0.0732	-0.0217	-0.2827
0.3204	0.4105	0.2365	0.1442	0	-0.0463	-0.1225	-0.0817
0.0310	0.2465	0.2134	0.1034	0	-0.0630	-0.0364	-0.0589
0.1227	0.1642	0.1039	0.0706	0	-0.0318	-0.0481	-0.0287
0	0	0	0	0	0	0	0
-0.0732	-0.0963	-0.0626	-0.0328	0	0.0238	0.0264	0.0192
-0.0197	-0.1320	-0.0384	-0.0421	0	0.0209	0.0364	0.0584
-0.0967	-0.0917	-0.0590	0.0297	0	0.0102	0.0264	0.0162

(b)

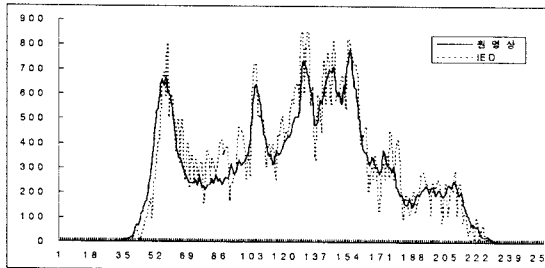
그림 3. DCT 계수와의 내적에 쓰이는 ξ_{ij}

(a) 1/4 해상도 (b) 1/2 해상도

여기서 H 와 \hat{H} 는 각각 원영상과 복원영상에서 추출한 히스토그램 bin 벡터이고, \cdot 는 벡터의 내적을 뜻한다. S 가 최대한 1에 가까울수록 두 히스토그램이 비슷하다고 할 수 있다. 그림 4의 (a)는 원영상과 제안 기법으로 복원된 영상으로부터 추출한 컬러 히스토그램을 나타낸다. 그림 4.(b)는 IED(Image Extraction in DCT domain)[1] 기법으로 복원된 영상의 컬러 히스토그램이다. 그림 5는 에지 히스토그램이다. 표 1은 원영상과 복원영상간의 S 를 나타낸다.

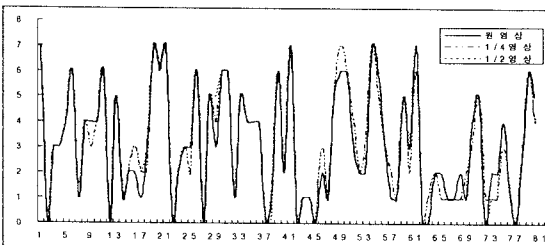


(a)

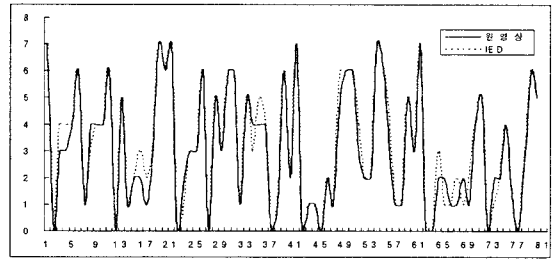


(b)

그림 4. lena 원영상과 복원영상의 컬러 히스토그램



(a)



(b)

그림 5. lena 원영상과 복원영상의 에지 히스토그램

표 1.에서 볼 수 있듯이 제안 기법을 통한 복원 영상들은 대부분 1에 매우 가까운 S 값을 갖는다. 그리고 모든 경우 S 는 1/2 경우가 1/4 경우보다 높은 것을 알 수 있다. 또한 IED를 통해 얻어진 1/4 해상도 영상보다 제안 기법을 통해 얻은 영상의 S 가 더 높음을 알 수 있다.

표 1. 원영상과 복원영상의 히스토그램 유사도

	lena			lori			boat		
	1/4	1/2	IED	1/4	1/2	IED	1/4	1/2	IED
S (컬러)	0.97630	0.9934	0.9759	0.96260	0.97940	0.96250	0.96210	0.98740	0.9613
S (에지)	0.98960	0.9948	0.9929	0.98680	0.99500	0.97650	0.99130	0.99450	0.9788

4. 결론

본 논문에서 우리는 히스토그램을 얻기 위해 8×8 DCT 계수의 선형 조합을 통해 곧장 averaged-version 영상을 복원하는 기법을 제안하였다. 제안 기법을 통해 원영상의 1/4, 1/2 해상도의 영상을 복원 가능하며, 이들로부터 추출한 히스토그램은 원영상의 히스토그램과 상당히 유사하다. 따라서 이 기법은 압축 영역에서의 히스토그램을 통한 영상 처리에 적용될 수 있다.

참고 문헌

- [1] G. C. Feng, J. Jiang, "Image extraction in DCT domain", IEE Proc. -Vis. Image Signal Process., Vol. 150, No. 1, February 2003
- [2] Minyoung Eom, Yoonsik Choe, "Fast Extraction of Edge Histogram Descriptor in DCT Domain Based on MPEG-7", Enformatika Systems Sciences and Engineering vol. 9, pp. 209-212, Nov. 2005
- [3] 송진호, 엄민영, 최윤석, "웨이블릿 영역에서의 웨이블릿 계수를 이용한 에지 히스토그램 고속 추출 기법", 전자공학회, 2005. 9. 제42권 SP편 제5호, pp. 137-144
- [4] K. R. RAO, "Techniques and Standards for Image, Video and Audio Coding", Prentice Hall PTR, Upper Saddle River, New Jersey 07458