

실시간 처리를 위한 DCT 영역에서의 불연속 경계 검출

(Discontinuity Detection in the DCT Domain for Real-Time Processing)

김 태 용 [†] 한 준 희 ^{**}
(Tae Yong Kim)(Joon Hee Han)

요 약 DCT(Discrete Cosine Transform) 변환은 MPEG과 JPEG 표준에 의하여 영상이나 비디오 영상을 압축하는데 사용되어져 왔다. 본 연구에서는 이상적인 경계와 관련된 특성을 유도하고, 실시간 처리를 위하여 모델에 바탕을 둔 DCT 영역에서의 불연속 경계 평가 방법을 제안한다. 이 방법은 방향 검증과 위치 정렬 등의 평가로 구성된다. 두 가지의 평가 방법에 의하여 경계의 다양한 방향과 위치를 알 수 있으며, DCT 계수들을 표준화된 형식으로 정렬시킬 수 있고, 표준화된 DCT 계수에서 이상적인 계단 경계의 특성과 비교하여 경계의 크기를 산출할 수 있다. DCT 계수가 8x8의 블록 단위로 이루어져 있어 경계의 표현이 조밀하지는 않지만 처리 시간이 빠르고 잡음에 강한 특성을 가지고 있어 다양한 실시간 응용 분야에 사용될 수 있을 것이다.

Abstract The Discrete Cosine Transform (DCT) has been used for compressing videos or images with standards like MPEG and JPEG. In this paper, we derive DCT properties related to a ideal edge and propose a model-based edge evaluation technique in the DCT domain. This technique consists of a direction verification and a position alignment method with an evaluation criterion. The direction verification and position alignment causes the DCT coefficients to be of the centralized form, which enables an evaluation regardless of various positions and directions of edges. The evaluation criterion examines the standard position and evaluates the magnitude of an edge by using the properties of the ideal step model. Although the detected edges are rough in a low-resolution image for the size (8x8 pixels) of DCT blocks, experimental results show that this technique is fast in processing and robust against noise.

1. 서 론

DCT(Discrete Cosine Transform)는 영상압축 분야에서 광범위하게 사용되고 있는 변환으로 정지 영상 압축 표준인 JPEG, 영상 회의의 표준인 H.261 및 동영상 압축 표준인 MPEG (MPEG-1, MPEG-2) 등에 채택되어 있다. 이러한 표준에서 2차원의 DCT 변환을 8x8 화소 블록 단위로 수행하여 공간 영역에서 주파수 영역으로

변환하며, 64개의 변환 계수가 양자화 되어 실질적인 압축이 이루어진다.

공간 영역에서 디지털 영상과 비디오의 내용 인식을 위하여 물체의 경계선(boundary)은 물체의 구조적인 정보를 가지고 있고, 처리되어야할 정보량을 상당히 줄일 수 있으므로 매우 중요한 정보 중에 하나이다. 경계(edge)는 디지털 영상에서 가장 기본적인 불연속으로서 밝기값 및 색상의 상대적인 변화에 의하여 두 개의 영역을 구분할 수 있는 특성을 가지고 있다. 이러한 경계는 공간 영역에서 저주파 필터링과 기울기 변화 감지 방법에 의하여 검출되어져 왔지만[1][2], 만약 영상이 압축되어져 있을 경우에는 압축을 완전히 해제하고, 경계를 검출하고, 다시 압축해야 하는 처리상의 비효율성을 가지고 있다.

· 본 연구는 교육부 BK21 사업의 지원을 받아 수행되었음.

† 비 회 원 : 중앙대학교 영상공학과 교수
kimty@postech.ac.kr

** 중 신 회 원 : 포항공과대학교 컴퓨터공학과 교수
joonhan@postech.ac.kr

논문접수 : 2000년 8월 1일

실사완료 : 2000년 12월 4일

따라서, 압축된 자료에서의 영상 처리는 압축을 해제 하지 않고 직접 압축 상태의 자료를 처리하는 기술이 제안되어 왔다. 이러한 방법은 전송률을 상당히 낮출 수 있고, 압축 해제 및 재압축의 과정을 없앨 수 있기 때문에 처리 시간을 크게 줄일 수 있다. Chang과 Messerschmitt[3]는 DCT 영역에서 비디오 처리를 위한 특별한 함수를 지속적으로 발전시켰고, Smith와 Rowe[4]가 보다 일반화된 형태로서 영상의 대수적 처리 알고리즘을 정의하였다. 이러한 대수적 특성을 바탕으로 DCT 영역에서의 스칼라 덧셈, 스칼라 곱셈, 픽셀 덧셈, 픽셀 곱셈, 플립(flip), 전단 변형(shear), 회전, 재배열(rearrangement) 등의 처리 방법이 제안되었다[3][4][5][6][7]. 이러한 알고리즘은 압축 해제 후 처리 방법보다 시간적으로 50배 이상 빠른 성능을 나타낸다[4][7].

그러나, 영상 이해(image understanding)나 내용 기반(content-based)의 응용 분야에서는 DCT 영역 작업이 거의 진행되지 않았고, 단지 Shen과 Sethi가 잘 유도된 경계 검출 방법을 제안하였지만[8], 이 방법은 DCT 계수와 등방성(isotropic)의 연산자(Laplacian of Gaussian) 사이의 효율적인 컨볼루션 연산에 중점을 두었기 때문에 궁극적으로는 DCT 역변환을 필요로 하고, 따라서 DCT 영역의 직접 처리보다 상대적으로 많은 처리 시간을 필요로 하였으며, DCT 계수의 개수에 민감하고 블록 경계에서 왜곡을 발생시켜 경계의 위치를 비교하기 어려웠다.

본 연구에서는 DCT 영역에서 경계의 직접적인 검출을 제안한다. 널리 사용되는 8x8 블록의 DCT는 단순히 압축을 위하여 정의되었으므로 단순한 지역 주파수 특성만을 반영한다. 결과적으로 불연속 경계의 변화는 각각 DCT 계수의 변화를 초래하고 경계의 위치를 모르는 상황에서는 기존 방법의 위치 불변(shift-invariant) 특성을 사용할 수 없기 때문에 경계의 검출이 매우 어렵다. 본 연구는 DCT 영역에서 블록 내부의 경계를 회전시키고, DCT 계수 크기를 보상(compensation)하여 수직과 중앙 위치로 정렬시켜 모든 경계를 표준화시킨다. 표준화된 경계와 이상적인 경계의 모델을 비교하여 경계정도를 평가할 수 있고, 8x8 블록에서의 경계의 크기를 분석할 수 있다. 이러한 방법은 DCT 영역에서 직접 수행되므로 처리 속도가 빠르고, 잡음에 강하다는 것을 실험을 통하여 나타내었다.

본 논문은 2장에서는 블록 내부에서 경계의 정렬과 불연속 크기의 평가 방법을 수식적으로 전개하였으며, 3장에서는 DCT 영역에서 불연속 경계의 회전 및 검출 방법을 서술하였다. DCT 영역에서의 경계 검출 결과와

분석을 4장의 실험에 나타내었으며, 연구 결론은 5장에 서술하였다.

2. 위치 정렬 및 높이 평가 (Position alignment and height evaluation)

DCT 영역에서의 DCT 계수 자체는 공간적인 불연속의 크기나 위치와는 직접적인 이론적 연관관계를 나타낼 수 없으며, 불연속의 모양 및 블록 내부에서의 위치에 따라서 그 값의 의미가 달라지므로 실제적인 평가는 불가능하다. 본 논문에서는 불연속의 모델을 정의하고, 이에 대응하는 AC 성분만을 유도하는 계산을 사용하여 블록 내부에서 가장 큰 대표 불연속의 크기를 계산할 수 있도록 하였다.

DCT 영역에서의 경계 검출을 위하여 블록내부의 경계 위치를 표준화된 형태로 기술할 필요가 있고, 평가를 원활히 하기 위하여 가장 큰 경계의 위치를 표준화된 위치(블록의 중앙)로 이동시켜야 한다. 이러한 처리는 주파수 성분 계수의 보상(compensation)으로 이루어지며, 이를 위치의 정렬이라고 정의한다. 이렇게 정렬된 위치에서 대칭성 평가를 통하여 실제의 경계 위치를 분석하고, 경계의 크기를 계산할 수 있다. 본 장에서는 이러한 일련의 처리를 위한 DCT 계수의 특성을 유도한다.

2.1 DCT 변환 (Discrete Cosine Transform)

DCT는 JPEG과 MPEG의 표준으로서 공간 여분의 제거로 압축을 수행하는데, 8x8 공간 픽셀에 적용되어 압축의 첫 단계에서 수행되며, 압축 해제시 마지막으로 역 DCT(inverse DCT)가 수행된다. 8x8 픽셀에 적용되는 DCT와 IDCT의 수학적 정의는 다음과 같다.

$$F(u, v) = \frac{1}{4} C(u)C(v) \sum_{i=0}^7 \sum_{j=0}^7 f(i, j) \cos \frac{\pi u(2i+1)}{16} \cos \frac{\pi v(2j+1)}{16}$$

$$f(i, j) = \frac{1}{4} \sum_{u=0}^7 \sum_{v=0}^7 C(u)C(v)F(u, v) \cos \frac{\pi u(2i+1)}{16} \cos \frac{\pi v(2j+1)}{16}, \quad (1)$$

여기에서 $u, v=0$ 일 때 $C(u), C(v) = \frac{1}{\sqrt{2}}$, 그 이외에는 1이다. FDCT는 영상의 공간 영역 표현을 효율적으로 코딩하기 위하여 주파수 영역으로 변환하는 기능을 수행하며, 일원 직교 변환(unitary orthogonal transform)으로 매트릭스 곱셈과 덧셈에 배분적(distributive)인 특성을 갖는다.

2.2 불연속 위치 정렬(Discontinuity position alignment)

블록, 공간 영역에서 위치 불변(shift-invariant)의 가정에 의한 연산자로서 경계가 검출되지만, DCT 변환을 위하여 8x8 블록으로 나눌 경우 불연속의 위치가 DCT

블록 내부의 임의의 장소에 나타나기 때문에 같은 크기의 불연속 경계라도 DCT 계수는 많이 변하게 되므로, DCT 계수를 검색하는 것만으로는 고주파 성분의 존재는 확인할 수 있지만, 경계의 크기나 위치를 정확하게 파악할 수 없다. 따라서 경계의 위치 변화에 의한 DCT 계수의 변화를 극복하기 위하여 이상적인 경계를 정의하고, 그의 DCT 계수를 분석한다.

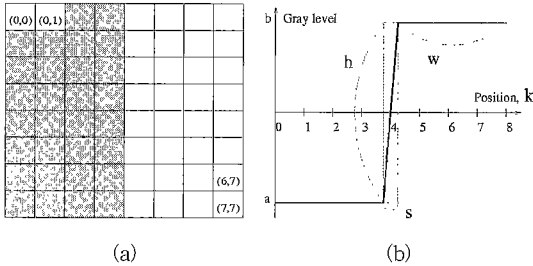


그림 1 이상적인 불연속 경계 모델 : (a) 공간 영역에서의 8x8 블록과 좌표, (b) 수평축에 따른 밝기값 변화와 불연속의 높이(h), 폭(w), 경사도(s) 및 위치(k).

(그림 1)의 (a)는 공간 영역에서의 이상적인 경계 불연속의 변화와 공간 및 주파수 영역의 좌표를 나타낸다. 8x8 블록의 밝기값은 식(1)의 j=3과 j=4 사이에서 변화하며, j=0, 1, 2, 3 위치에서 f(i,j)=a와 j=4, 5, 6, 7 위치에서 f(i,j)=b로 수식화되며 그 크기는 $-127 \leq f(i,j) \leq 127$ 이다. 이러한 경우 F(0,v)의 DCT 계수만 0이 아닌 값을 가지며 나머지 계수는 0으로 표현된다. 수평축으로 절단했을 경우의 밝기 변화를 (그림 1)의 (b)에 나타냈으며, 여기에서 이상적인 불연속은 높이, 폭, 경사도 및 위치로 모델링되며 각각은 다음과 같이 나타낸다.

$$\begin{aligned}
 h &= |b-a|, \\
 w &= \text{MIN}(\sum_j 1 \text{ if } f(i,j)=a, \sum_j 1 \text{ if } f(i,j)=b), \\
 s &= \sum_j 1 \text{ if } f(i,j) \neq a \text{ or } f(i,j) \neq b, \\
 k &= j \text{ if } f(i,j)=|b-a|/2.
 \end{aligned} \tag{2}$$

중심 위치(k=4)는 이미 알고 있는 이상적인 경계 모델과 비교하여 각종 모델 파라미터를 계산하기 쉬운 위치이므로, 임의의 불연속 위치($1 \leq k \leq 7$)를 중심 위치로 이동시킨 것을 정렬이라고 정의한다. 이러한 위치 정렬을 위하여 DCT 계수에 보상 주파수를 가감하여 주는데, 이는 공간 영역에서의 위치 이동과 같은 효과를 나타낸다. 보상 주파수는 $c_k(h, v) = F_4(0, v) - F_k(0, v)$ 에 의하여 얻어지는데, 여기에서 F_k 는 불연속의 위치가 k

일 경우의 DCT 계수를 나타낸다. F_3 의 경우를 유도하여 정리하면 다음과 같다.

$$\begin{aligned}
 c_3(h, v) &= F_4(0, v) - F_3(0, v) \\
 &= \frac{2}{\sqrt{2}} \left(\sum_{j=0}^7 f_4(0, j) \cos \frac{\pi v(2j+1)}{16} - \sum_{j=0}^7 f_3(0, j) \cos \frac{\pi v(2j+1)}{16} \right) \\
 &= \frac{2}{\sqrt{2}} (f_4(0,3) \cos \frac{7\pi v}{16} - f_3(0,3) \cos \frac{7\pi v}{16}) \\
 &= -\frac{2}{\sqrt{2}} h \left(\cos \frac{7\pi v}{16} \right).
 \end{aligned} \tag{3}$$

식 (3)의 결과를 응용하여 나머지 위치의 보상 주파수를 구하면 다음과 같이 표현된다.

$$\begin{aligned}
 c_1(h, v) &= -\frac{2}{\sqrt{2}} h \left(\cos \frac{3\pi v}{16} + \cos \frac{5\pi v}{16} + \cos \frac{7\pi v}{16} \right) \\
 c_2(h, v) &= -\frac{2}{\sqrt{2}} h \left(\cos \frac{5\pi v}{16} + \cos \frac{7\pi v}{16} \right) \\
 c_3(h, v) &= -\frac{2}{\sqrt{2}} h \left(\cos \frac{7\pi v}{16} \right) \\
 c_5(h, v) &= \frac{2}{\sqrt{2}} h \left(\cos \frac{9\pi v}{16} \right) \\
 c_6(h, v) &= \frac{2}{\sqrt{2}} h \left(\cos \frac{9\pi v}{16} + \cos \frac{11\pi v}{16} \right) \\
 c_7(h, v) &= \frac{2}{\sqrt{2}} h \left(\cos \frac{9\pi v}{16} + \cos \frac{11\pi v}{16} + \cos \frac{13\pi v}{16} \right).
 \end{aligned} \tag{4}$$

이 결과를 이용하여 보상된(위치가 중심으로 이동된) DCT 계수는 다음 식으로 얻을 수 있다.

$$\hat{F}_k(0, v) = F(0, v) + c_k(h, v),$$

여기에서 $0 \leq v \leq 7$, h는 (그림 1) (b)에 표시된 불연속의 높이이고, $\hat{F}_k(0, v)$ 의 k는 보상하기 전의 불연속 경계가 있는 위치를 나타낸다. 그렇지만 아직 불연속 경계의 위치 k는 알 수 없으므로 정렬의 정당성을 검증하여야 한다.

2.3 정렬 검증(Alignment verification)

임의의 불연속 경계의 위치 k를 중앙 위치로 이동시키는 보상 주파수를 전장에서 유도하였지만, DCT 계수만으로는 블록 내부에서의 실제의 위치를 알 수 없으므로 각각의 위치에서 보상되어 정렬된 DCT 계수 \hat{F}_k 를 평가하여 어떤 위치에서 불연속 경계가 존재하는지를 검증한다.

위치 검증을 위하여 블록의 수직 중간 축을 중심으로 플립하고 부호를 바꾼 원점 대칭의 특성을 사용한다. 공간 영역에서 입력 블록을 f(i,j)라 하고 출력 블록을 g(i,j)라 할 때, 원점 대칭은 다음과 같이 표현된다.

$$g(i, j) = -f(i, 7-j); \quad 0 \leq i, j \leq 7.$$

주파수 영역에서 입력 블록을 F(u,v)라 하고, 출력 블록의 계수를 G(u,v)로 나타낼 때 주파수 영역에서의 원점 대칭은 다음과 같이 유도된다.

$$\begin{aligned}
 G(u, v) &= \frac{C(u)C(v)}{4} \sum_{i=0}^7 \sum_{j=0}^7 g(i, j) \cos \frac{\pi u(2i+1)}{16} \cos \frac{\pi v(2j+1)}{16} \\
 &= -\cos(\pi v) \frac{C(u)C(v)}{4} \sum_{i=0}^7 \sum_{j=0}^7 f(i, j) \cos \frac{\pi u(2i+1)}{16} \cos \frac{\pi v(2j+1)}{16} \\
 &= -\cos(\pi v) F(u, v).
 \end{aligned} \tag{5}$$

여기에서 $j'=7-j$ 이다. 따라서 식 (5)와 같이 주파수 영역에서 입력 블록에 대한 원점 대칭 결과는 입력 블록에서 $G(u, v) = -\cos(\pi v)F(u, v)$ 와 같이 직접적으로 구해진다.

(그림 1) (b)와 식 (5)에 따라서 원점 대칭은 $G(u, v) = -\cos(\pi v)F(u, v) = F(u, v)$ 로 나타나므로, 만약 블록의 k 위치에 불연속 경계가 있었을 경우 $\hat{F}(0, 2)$, $\hat{F}(0, 4)$, $\hat{F}(0, 6)$ 의 DCT 계수들은 모두 0이 되어야 한다. 또한, 정렬된 DCT 계수에서의 불연속의 모양은 $\hat{F}(0, 1)$, $\hat{F}(0, 3)$, $\hat{F}(0, 5)$ 및 $\hat{F}(0, 7)$ 들의 값에 따라서 정의되어 진다. 따라서, 검증 평가식을 다음과 같이 정의할 수 있다.

$$D_k = \hat{F}_k(0, 2)^2 + \hat{F}_k(0, 4)^2 + \hat{F}_k(0, 6)^2, \quad (6)$$

여기에서 $k=1, 2, \dots, 7$ 이고, 원래의 불연속 위치 k 는 보상된 DCT 계수와 평가식 D_k 에 의하여 고정된 높이 h 에 의하여 평가받는다. 만약 불연속 경계가 존재할 경우 정확한 높이일 경우 해당 위치에서 식(6)의 평가식이 0이 되어야 하고, 임의의 높이를 사용했을 경우에는 해당 위치에서 평가식 D_k 가 최소가 되기 때문에 위치 k 를 찾을 수 있다.

2.4 높이의 평가(Height evaluation)

불연속 경계의 높이는 (그림 1) (b)와 같이 불연속을 가로지르는 밝기값의 차로 정의되었고, 이를 이용하여 많은 정보를 담고 있는 뚜렷한 경계를 구분할 수 있다. 전 장에서 유도된 불연속의 정렬 평가로 위치 정보를 얻은 후, 이 위치에서 높이를 구할 수 있다. 높이의 평가 함수는 전 장의 D_k 와 같으나 높이 값 h 가 변화된다.

Cos 함수는 주기함수이고, $f(0, j)$ 값은 (그림 1) (b)에서 a 혹은 b 로 정하였으므로, 파라미터 h 에 의한 특정 위치 k 에서의 D_k 는 다음과 같이 전개된다.

$$\begin{aligned} D_k(h) &= \hat{F}_k(0, 2)^2 + \hat{F}_k(0, 4)^2 + \hat{F}_k(0, 6)^2 \\ &= [c_k(\tilde{h}, 2) + c_k(h, 2)]^2 + [c_k(\tilde{h}, 4) + c_k(h, 4)]^2 + [c_k(\tilde{h}, 6) + c_k(h, 6)]^2 \\ &= (\tilde{h}\alpha_{k_2} - h\alpha_{k_2})^2 + (\tilde{h}\alpha_{k_4} - h\alpha_{k_4})^2 + (\tilde{h}\alpha_{k_6} - h\alpha_{k_6})^2 \\ &= (\tilde{h} - h)^2(\alpha_{k_2}^2 + \alpha_{k_4}^2 + \alpha_{k_6}^2). \end{aligned} \quad (7)$$

여기에서 \tilde{h} 는 실제 높이이고, $c_k(h, v)$ 는 k 위치에서 식 (2)에 의한 보상 주파수이며, α_{k_2} , α_{k_4} , 및 α_{k_6} 는 위치 k 에서 주파수로 전개된 고정된 상수이다. 높이 파라미터 h 는 밝기값 한계에서 변화할 수 있지만, 실제와 같은 값일 경우($\tilde{h}=h$)에 식(7)의 평가식이 최소가 되어서 실제 높이를 평가할 수 있다.

3. 회전과 검출(Rotation and detection)

본 장에서는 DCT 영역에서 회전에 대한 기존의 연

구를 살펴보고, 경계의 빠른 검출을 위한 구현 방법을 제시한다.

3.1 DCT 영역에서의 회전 (Rotation in the DCT domain)

이제까지 유도되었던 정렬 및 평가는 불연속 경계의 폭이 8x8블록의 반을 차지하고($w=4$), 경계의 방향이 수직하다는 가정을 가지고 있었다. 이 중 반을 차지한다는 가정은 압축을 위한 고정된 블록의 크기(8x8) 때문에 쉽게 해결하기는 어렵지만 영상의 해상도가 커지면 영향력이 현저히 줄어들게 된다.

수직 방향 가정은 주파수 영역에서의 회전과 전단 변형 연산(shearing)에 의하여 해결할 수 있다. 불연속 경계의 방향을 미분의 방향으로 $\tan^{-1}(\frac{\partial f}{\partial y} / \frac{\partial f}{\partial x})$ 와 같이 정의하면 90°의 배수의 회전은 DCT 계수의 부호를 바꿈에 의하여 쉽게 얻을 수 있다[7].

그 외의 방향은 회전보다는 전단 변형 연산에 의하여 얻는 것이 빠르고, 잘 정의되어 있다[3] [5] [7]. 이 전단 변형 연산은 열 추출과 열 이동 연산의 결합으로 가능하며, 다음의 식으로 정의된다.

$$\begin{aligned} f_s &= \sum_i w_i f s_i, \\ F_s &= \sum_i W_i F S_i, \end{aligned} \quad (8)$$

여기에서, f_s 는 공간 영역의 블록이고, w_i 매트릭스는 i 번째 열을 추출하고, s_i 는 1차원으로 이동시키는 매트릭스이다. F , W_i , S_i 는 각각 공간 영역에 대응하는 DCT 영역에서의 매트릭스이다[3]. 식 (8)에서와 같이 DCT 영역에서의 전단 변형 연산은 공간 영역에서와 같이 직교 변환의 배분 성질을 단순히 적용함으로써 얻을 수 있고, W_i 및 S_i 는 미리 계산하여 사용할 수 있다. 그러나, 다양한 각도에 적용하기 위해서는 너무 많은 계산 테이블과 복잡한 연산이 필요하여 기존의 공간 영역의 연산보다 더 나은 효율을 기대하기 어려우므로[3], 본 연구에서는 45°의 배수 각도에만 전단 변형 연산을 수행한다.

3.2 검출의 구현 (Detection implementation)

회전 및 전단 변형을 위한 $\tan^{-1}(\frac{\partial f}{\partial y} / \frac{\partial f}{\partial x})$ 의 계산은 $f(i, j) - f(i, j-1)$ 과 같이 디지털 차분으로 근사화하여 경계 방향을 유도하였다. 이를 이용하여 수평의 방향을 갖는 높이 h 의 경계는 $\sum_{j=1}^7 |\frac{\partial f}{\partial x}| \approx \sum_{j=1}^7 |f(i, j) - f(i, j-1)| = \sum_{j=1}^7 |f_j(i, j)| = h$ 와 같이 유도되고, 이러한 특성을 수직 및 수평으로 고려하면, 계단 경계의 미분으로부터 얻은 DCT 계수의 합은 다음과 같이 전개된다.

$$\begin{aligned}
 M_V &= \sum_{u=1}^7 |F(u, 0)|, \\
 M_H &= \sum_{v=1}^7 |F(0, v)| \\
 &= \sum_{v=1}^7 \left| \frac{2}{\sqrt{2}} \sum_{j=0}^7 f_j(0, j) \cos \frac{\pi v(2j+1)}{16} \right| \quad (9) \\
 &= \frac{2}{\sqrt{2}} h \sum_{v=1}^7 \left| \cos \frac{\pi v(2j+1)}{16} \right| \\
 &= \frac{2}{\sqrt{2}} 4.58h = 6.48h.
 \end{aligned}$$

이 값은 미분의 크기(h)와 비례하므로 기존의 $\tan^{-1}(\frac{\partial f}{\partial y} / \frac{\partial f}{\partial x})$ 를 대신하여 $\tan^{-1}(\frac{M_V}{M_H})$ 를 사용하여 근사화한다. 이것으로 DCT 계수 블록에 포함된 불연속의 방향을 판단하여 45° 및 135°는 전단 변형 연산을 수행하고, 90°의 배수는 부호 변경으로 회전시킨다.

전단 변형이나 회전에 의하여 경계의 방향을 수직으로 변환한 후, 정렬에 의하여 경계의 위치를 찾고, 높이를 평가한다. 식 (7)과 같이 높이는 파라미터에 따라 2차 함수 형식으로 변화하므로 기울기 강하(gradient descent) 방법[9]을 사용하여 빠르게 구현할 수 있다. 만약 임계값이 주어졌던 경우에는 높이(h) 대신 임계값을 파라미터로 사용하여 D_k 를 계산하면, 한번의 연산에 의하여 경계의 존재를 판단할 수 있다. 즉, 임계값을 높이로 사용하였을 경우, D_k 기울기 부호가 음이면 블록에 있는 경계의 높이가 임계값보다 높고, 기울기 부호가 양이면 경계의 높이가 임계값보다 적어서 검출되지 않는다. 이러한 원리를 이용하면 식(6)의 평가식에 의하여 정렬 평가시 높이 판별이 동시에 이루어 질 수 있다.

블록간의 불연속은 한 블록의 평균인 DC 성분 ($F(0,0)/8$)을 교차 기울기(cross-gradient)처럼 서로 미분함으로써 얻을 수 있다[10].

$$\|\nabla F_{DC}\| = [F_{DCx}^2 + F_{DCy}^2]^{\frac{1}{2}}, \quad (10)$$

$$\nabla F_{DC} = [F_{DCx}, F_{DCy}]^T = \left[-\frac{\partial F_{DC}}{\partial x}, \frac{\partial F_{DC}}{\partial y} \right]^T.$$

DCT 영역에서 불연속 경계를 찾는 방법을 요약하면 다음과 같다.

- ① JPEG이나 MPEG의 블록으로부터 밝기 요소의 DCT 계수를 습득.
- ② 식 (8)과 (9)를 이용하여 필요하면 회전 및 전단 변형 연산을 수행.
- ③ 식 (4)를 사용하여 DCT 계수를 정렬하고, 식 (6)을 이용하여 위치를 평가.
- ④ 식 (7)을 이용하여 불연속 경계의 높이를 계산하거나 임계값과 비교.
- ⑤ 영상의 모든 블록을 처리할 때까지 ①에서부터 반

복 수행.

⑥ 식 (10)을 이용하여 블록간 불연속 검출.

4. 실험 (Experiments)

영상에서의 자막(caption)은 자동적인 내용기반 영상 검색에 있어서 유력한 정보를 제공하여 주고, 디지털 비디오 저장에 위한 비디오 분석에 의하여 검출, 추출 및 인식되어 만들어진 색인 정보로도 사용될 수 있다[11][12]. 제안한 불연속 경계 검출을 자막 검출에 적용하기 위하여 방송에서 습득한 640x480크기의 영상을 사용하였으며, 45° 및 135°에 대한 전단 변형 연산과 90° 및 180°의 회전을 사용하였다. 이러한 4 방향에 대칭성을 이용하여 8방향의 방향성을 고려하였으며, W_x 와 S_x 는 미리 계산하여 테이블에 저장하여 사용하였다. 또한, 비교를 위하여 사용한 FDCT와 IDCT 구현은 Loeffler



(a)



(b)



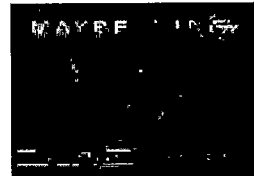
(c)



(d)



(e)



(f)

그림 2 DCT 영역에서 광고 영상의 불연속 경계 검출 :

- (a) 방송에서 습득한 광고 영상, (b) 임계값 40 일 때 검출된 경계 영상, (c) 임계값 40과 평활화를 사용하여 검출한 경계 영상, (d) 임계값 80 일 때 검출된 경계 영상, (e) 임계값 80일 때 전단 변형 연산을 사용하지 않은 경계 영상, (f) 임계값 80일 때 블록간 경계만 검출한 영상.

[13] 등이 제안한 방법을 사용하였으며, 공간 영역에서의 경계 검출은 Canny 연산자를 사용하였다.

(그림 2)의 (a)는 사람 얼굴과 광고 문안이 들어 있는 방송 영상을 보여준다. 이 영상에서 불연속 경계의 높이가 40 이상으로 검출된 영상을 (b) 그림에 나타내었으며, 고주파 성분을 평활화한 영상을 (c)에 나타내었다. 평활화는 높이를 구할 때 AC 고주파 성분을 적게 고려한 처리로써 얼굴과 자막사이의 경사경계가 검출됨을 나타낸다. (d)는 임계값을 80으로 하여 높이가 확실한 경계만을 검출한 영상을 나타낸다. 그림에서 보는바와 같이 자막에서의 불연속이 얼굴보다 크므로 임계값을 크게 하여도 자막의 대부분이 검출됨을 나타낸다. 또한, 그림 (e)와 같이 시간을 많이 소모하는 전단 변형을 수행하지 않아도, 자막의 경우 대부분 수평이나 수직으로 형성되므로, 5% 블록에 대한 전단 변형 연산을 수행한 그림 (d)와 비교할 때 자막의 정보가 대부분 유지됨을 알 수 있다. 제안한 블록 내부의 경계를 검출하지 않고 블록간의 경계만을 검출한 경계 영상을 그림(f)에 나타내었다. 블록 내부를 검사하지 않았기 때문에 얼굴 부분과 아래쪽 자막의 세부 내용이 검출되지 않았음을 보여준다.

(그림 3)의 (a)는 투명한 자막을 가진 뉴스 영상을 나타낸다. 공간 영역에서 Canny 연산자에 의하여 경계를 검출하였을 경우를 (b)에 나타내었고, 제안한 방법에 의하여 DCT 영역에서 검출된 경계 블록의 영상을 검출 위치의 정확성을 나타내기 위하여 그림 (c)에 나타내었으며, 큰 불연속을 갖는 경계의 위치가 거의 일치함을 알 수 있다.

블록 내부의 대표 불연속의 크기를 찾는 과정에서 연결 안된 백색 잡음이 AC 성분에 큰 영향을 미치므로 잡음의 영향을 분석하였다. 실험에서는 가우스 백색 잡음을 사용하였으며, Kitchen과 Rosenfeld[14]가 정의한 $SNR = (h/\sigma_N)^2$ 로 모델링하였다. 여기에서 h 는 불연속의 높이이고, σ_N 은 잡음의 표준편차를 나타낸다. 실험에서는 $h=25$ 를 사용하였다. 만약 원래 영상이 잡음(SNR=1)의 영향을 받았을 경우를 (d)에 나타내었다. Canny 연산자의 평활화 크기를 균일하게 사용하였기 때문에 검출 결과는 잡음의 영향을 매우 심하게 받았음을 그림 (e)에서와 같이 관찰할 수 있다. 그러나, DCT 영역에서 제안한 방법은 블록 단위의 검색을 수행하여 가장 큰 경계만을 검출 대상으로 간주하므로 파라미터의 변경 없이 잡음에 강한 특성을 가지고 있고, 잡음이 없는 영상 (c)와 경계 위치의 표시에서 큰 차이가 없음을 그림 (f)에서 살펴볼 수 있다.

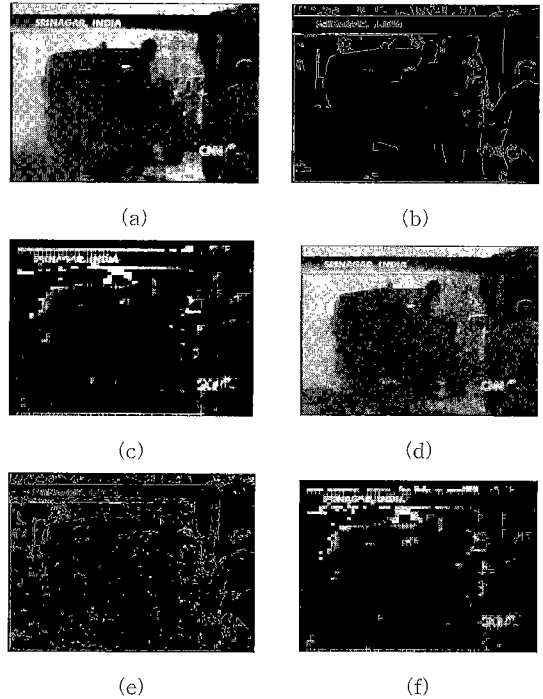


그림 3 공간 영역에서 Canny 연산자(임계값=0.9, $\sigma=1$)를 사용하여 얻은 경계 영상과 DCT 영역에서 제안한 방법(임계값=40)에 의하여 얻은 경계 영상 : (a) 방송에서 습득한 뉴스 영상, (b) Canny 연산자에 의하여 얻은 경계 영상, (c) 제안한 방법에 의하여 얻은 경계 영상, (d) 잡음(SNR=1)을 첨가한 뉴스 영상, (e) 잡음 영상에서 Canny 연산자로 검출한 경계 영상, (f) 잡음 영상에서 제안한 방법에 의하여 얻은 경계 영상.

표 1 영상의 크기(pixels)에 따른 처리 시간 비교 (msec): T_{d0} 는 전단 변형 연산이 없을 경우의 경계 검출 시간, T_{d5} 및 T_{d10} 는 각각 5%와 10%의 블록에 대하여 전단 변형 연산을 수행했을 경우의 경계 검출 시간, T_t 는 IDCT와 FDCT의 수행 시간, T_{td} 는 IDCT와 FDCT 수행 후 Canny 연산자로 경계를 검출했을 경우의 수행 시간.

Image size	T_{d0}	T_{d5}	T_{d10}	T_t	T_{td}
256 x 256	9	62	114	104	284
480 x 480	35	182	395	370	982
720 x 720	91	510	1057	843	2265
1024 x 1024	181	546	1136	1701	7949

<표 1>은 경계를 검출하기 위한 처리와 변환에 필요한 수행 시간을 1/1000초로 나타내고 있다. 표에서 T_{d1} , T_{d5} 와 T_{d10} 는 DCT 영역의 제안한 방법에서 각각 전단 변형 연산을 0%, 5% 및 10% 수행하여 경계를 검출했을 경우의 수행 시간을 나타낸다. T_s 는 다른 연산 없이 IDCT와 FDCT를 수행했을 때의 시간을 나타내며, T_{id} 는 압축된 영상을 공간 영역에서 검출한 후 다시 압축하는데 필요한 시간을 나타내는 것으로, IDCT, FDCT 및 Canny 연산 시간을 종합한 것이다. 표에서와 같이 DCT 영역에서 제안한 방법은 전단 변형을 수행하지 않을 경우 IDCT+FDCT 변환에 비하여 매우 빠름을 알 수 있고, 전단 변형 연산을 수행하는 경우에도 영상의 크기, 불연속의 개수 및 전단 변형 연산의 수 등에 따라서 다르지만, 전반적으로 공간 영역에서의 경계 검출보다 약 2배에서 40배까지 빠름을 보여준다. DCT 영역에서의 효율적인 컨볼루션을 수행하여 화소 단위의 경계를 검출했던 Shen과 Sethi[8]의 실험 결과 (2 ~ 10 배의 속도 향상)와 비교할 때 상당히 효과적임을 알 수 있다.

전단 변형 연산과 불연속 경계의 개수가 많은 경우에는 IDCT+FDCT 변환보다 많은 시간을 필요로 하지만, (그림 2)의 (d)에서와 같이 자막 검출 같은 특수한 응용 분야에서는 전단 변형 연산을 전혀 수행하지 않아도 많은 정보를 손실하지 않기 때문에 다양한 실시간 응용 분야에 사용될 수 있을 것이다.

5. 결 론

DCT 영역에서의 영상 조작에 대한 연구가 지속적으로 진행되어 왔음에도 불구하고, 영상이해나 특징추출에 대한 DCT 영역에서의 연구는 부족한 상태였다. 본 연구에서는 DCT 영역에서 위치 정렬 및 평가에 대한 특성을 유도하였고, 빠른 불연속 경계 검출 방법을 제안하였다. 본 연구에서 제시한 방법은 실험 결과에서와 같이 불연속 검출에서 매우 효율적이고, 잡음에 강한 특성을 가지고 있으므로, 실시간 처리가 필요한 응용 분야의 원천 기술로서 사용될 수 있겠다. 그러나, 전단 변형이나 회전을 보다 효율적으로 수행할 수 있는 알고리즘의 개발이 필요하고, 압축과 처리를 상호 연결할 수 있는 코딩 기법에 대한 연구가 지속되어야 하겠다.

참 고 문 헌

[1] John Canny, "A Computational Approach to Edge Detection," IEEE Trans. on PAMI, Vol. 8, pp.

679-698, 1986.
 [2] D. Marr and E. Hildreth, "Theory of Edge Detection," Proceeding of Royal Society of London, Vol. B-207, pp. 186-217, 1980.
 [3] Shin-Fu Chang and David G. Messerschmit, "Manipulation and Compositing of MC-DCT Compressed Video," IEEE JSAC Special Issue on Intelligent Signal Processing, 1994.
 [4] Brian C. Smith and Lawrence A. Rowe, "Algorithms for Manipulating Compressed Images," IEEE Computer Graphics and Applications, pp. 34-42, 1993.
 [5] Shin-Fu Chang, "New Algorithms for Processing Images in the Transform-Compressed Domain," SPIE Symposium on Visual Communication and Image Processing, 1995.
 [6] B. Chitprasert and K. R. Rao, "Discrete Cosine Transform Filtering," Signal Processing, Vol. 19, pp. 233-245, 1990.
 [7] Bo Shen and Ishwar K. Sethi, "Inner-Block Operations On Compressed Images," ACM Multimedia '95, pp. 489-498, 1995.
 [8] Bo Shen and Ishwar K. Sethi, "Convolution-Based Edge Detection for Image/Video in Block DCT Domain," Journal of Visual Communication and Image Representation, Vol. 7, pp. 411-423, 1996.
 [9] Robert M. Haralick and Linda G. Shapiro, 'Computer and Robot Vision', Addison Wesley, pp. 605-606, 1992.
 [10] Rafael C. Gonzalez and Richard E. Woods, 'Digital Image Processing,' Addison Wesley, pp. 197-201, 1993.
 [11] Rainer Lienhart, "Automatic Text Recognition for Video Indexing," Fourth ACM International Multimedia Conference, 1996.
 [12] Jun Ohya, Akio Shio and Shigeru Akamatsu, "Recognizing Characters in Scene Images," IEEE Trans. on PAMI, Vol. 16, pp. 214-220, 1994.
 [13] C. Loeffler, A. Ligtenberg and G. Moschytz, "Practical Fast 1-D DCT Algorithms with 11 Multiplications," Proc. Int'l. Conf. on Acoustics, Speech, and Signal Processing (ICASSP '89), pp. 988-991, 1989.
 [14] Les Kitchen and Azriel Rosenfeld, "Edge Evaluation Using Local Edge Coherence," IEEE Trans. on SMC, Vol. SMC-11, pp. 597-605, 1981.



김 태 용

1986년 2월 한양대학교 전기공학과 학사. 1988년 2월 한양대학교 전자통신공학 석사. 1998년 2월 포항공과대학교 컴퓨터공학과 박사. 1988년 3월 ~ 1999년 2월 한국통신운용연구단 연구원. 1999년 3월 ~ 1999년 12월 포항공과대학교 컴퓨터공학과 연구원. 2000년 1월 ~ 중앙대학교 첨단 영상대학원 영상공학과 교수. 관심분야는 영상통신, 영상처리, 컴퓨터비전, 컴퓨터보안.



한 준 회

1979년 한양대학교 전자공학과 졸업. 1981년 서울대학교 대학원 전자공학과(공학석사). 1988년 미시간 대학교 엔지니어, 컴퓨터 공학(공학박사). 1995년 캘리포니아 대학, 샌디에고, 방문연구원. 1988년 ~ 1999년 현재 포항공과대학교, 컴퓨터공학과 조교수, 부교수. 관심분야는 컴퓨터 비전, 패턴인식임.