제26회 한국정보처리학회 추계학술발표대회 논문집 제13권 제2호 (2006. 11)

MPEG-2/H.264 트랜스코더를 위한 DCT 기반 Intra 예측기법

이주경* 정기동**
*신라대학교 컴퓨터정보공학부
**부산대학교 컴퓨터공학과
e-mail:joklee22@silla.ac.kr

DCT-domain Intra Prediction Scheme for MPEG-2/H.264 Transcoder

Joo-Kyong Lee*, Ki-Dong Chung**
*Division of Computer and Information Engineering, Silla University
**Dept. of Computer Engineering, Pusan National University

요 약

H.264 코딩시 Intra 모드는 다양한 모드 즉, 매크로블록(MB) 당 16개의 4x4 블록 각각에 대한 9가지의 4x4 모드와 4가지의 16x16 모드의 오류 값을 계산하여 최상의 모드를 선택하게 된다. 이와 같은 픽셀 기반의 예측 기법을 DCT 기반의 MPEG-2/H.264 트랜스코더에 적용할 경우 DCT 변환의 특성으로 인하여 모드 예측을 위한 계산량이 높아지는 단점이 있다. 본 논문에서는 이러한 문제점을 해결하기 위해 MPEG2 DCT를 H.264의 정수형 DCT로 변환한 후, DCT 계수의 특징을 이용하여 Intra 모드를 결정하는 기법을 제안한다. 이때, 계수의 특징이 모호하여 모드를 결정하기 어려운 경우는 몇 가지 모드를 선택하여 오류 값을 계산하여 모드를 결정한다. 현재 정밀한 실험은 진행 중이며, 여러 동영상의 첫 Intra 프레임에 대한 실험을 수행한 결과 MB의 모드 결정의 정확도는 프레임 내 이미지의 특징에 영향을 많이 받았다. 예를 들면 Mobile과 같이 프레임 내의 픽셀 간 에지가 많이 존재하는 프레임은 추가적인 모드 결정을 사용하지 않아도 93%정도의 정확도를 보였으며 Akiyo, Foreman과 같이 이웃한 픽셀간 유사도가 상대적으로 높은 경우는 약 80% 정도의 순수 정확성을 보였다. 그러나 모드 판단이 모호한 경우의 모드도 결정한다면 90%이상의 정확도를 보일 것으로 예상된다. 향후 이미지의 특성에 따라 모드를 결정하는 값을 유동적으로 설정하는 기법을 연구하여 정확도를 높이는 연구를 수행하도록 하겠다.

1. 서론

유비쿼터스 컴퓨팅 시대를 맞이하여 시간과 공간의 제약을 받지 않고 다양한 멀티미디어 서비스를 제공받고자 하는 사용자들의 요구가 높아지고 있다. 이러한 사용자들의 요구를 만족시키기 위한 응용과 네트워크로 DMB(Digital Multimedia Broadcasting), Wibro(Wireless Broadband Internet)가 대표적이다. 그러나 진정한 유비쿼터스 서비스를 제공하기 위해서는 사용자의 위치, 단말기의 종류, 멀티미디어 코덱에 관계없이 언제든지 서비스를 받을 수 있어야 한다.

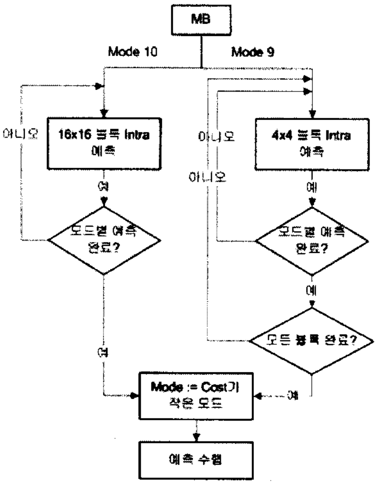
최근 ISO와 ITU-T에서는 차세대 동영상 압축 표준인 H.264(MPEG-4/AVC PART 10)를 공동으로 제정하였다[1]. H.264는 다양한 네트워크 환경에 쉽게 적응할 수 있는 유연성과 동영상 압축의 효율성이 기존 표준에 비해 월등하여[2] 유비쿼터스 서비스를 제공하기에 적합한 코덱이다. 본 논문은 유비쿼터스 환경을 위한 중요한 요소인 트랜스코더에 관한 것이며 그 중에서 특히, DCT 기반의 MPEG-2와 H.264의 트랜스코더를 효율적으로 구현하기 위한 기법에 대한 연구를 수행한 것이다. H.264에서는 기존의 압축 표준과 달리 Intra 프레임의 압축 효율을 높이기 위해 프레임 내 참조 기법을 사용한다. 따라서 MPEG-2와 H.264 간의 Intra 프레임 트랜스코딩을 위해서는 프레임 내의 연관성을 예측 및 보상하는 과정이 추가되어야 한다. 또한, Intra 프레임은 Inter

프레임의 참조 대상이 되므로 영상의 화질을 최대화 시키는 것도 중요하다. 그러나 DCT 기반의 Intra 예측은 MB당 16개의 4x4블록에 대하여 9번의 모드 예측, 4번의 16x16 모드 예측을 통해서 모드를 결정해야 하므로 DCT 상의 연산 복잡도를 높이는 원인이 된다. 본 논문에서는 이를 해결하기 위해 MPEG-2의 DCT 블록을 H.264의 4x4 정수형 블록으로 변환하고 각 블록의 DC 계수를 모아 4x4 하다마드 변환(Hadamard transform)을 수행한다. 이 결과를 이용하여 16x16 모드로 예측할 것인지, 4x4 블록으로 예측할 것인지 결정하게 된다. 이 과정은 MB 내의 4x4 블록의 평균값을 구하여 하다마드 변환을 통해 MB내의 에지를 찾는 것이다. 또한 각 블록의 예측기법도 DCT 데이터의 특징을 이용하여 선별적으로 적용한다. 이를 통하여 DCT 상에서의 예측을 수행하기 위한 계산 복잡도를 줄일 수 있다. 본 논문의 구성은 다음과 같다. 2장에서는 H.264 표준에서의 Intra 프레임 압축 기법 및 트랜스코더에서의 예측에 관한 연구를 살펴본다. 3장에서는 본 논문에서 제안하는 DCT상에서의 Intra 프레임 변환 기법을 알아보기로 한다. 4장에서는 실험결과를 살펴보고 5장에서 결론과 향후 연구에 대해 밝히도록 한다.

2. 관련 연구

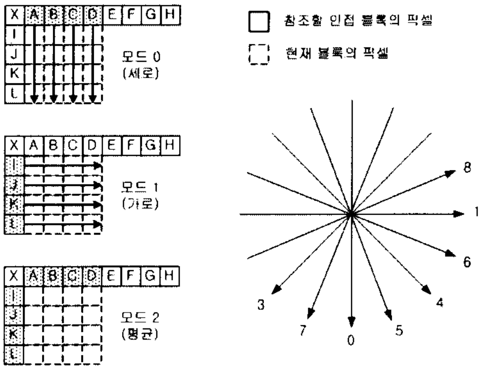
2.1 H.264의 Intra 예측 기법[1]

H.264 표준에서 Intra 모드의 MB은 (그림 1)과 같이 MB당 16x16 블록 크기의 예측과 4x4 블록의 예



(그림 1) Intra 모드의 예측 과정

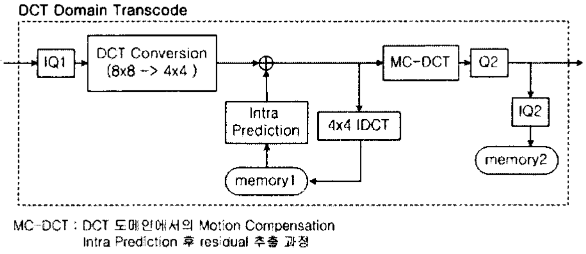
측을 수행하여 그 중 참조 값과의 차이가 적은 모드를 선택하게 된다. 참조 값은 16x16 블록의 경우 현재 MB과 인접한 MB가 압축된 후 복원된 경계선의 픽셀값이다. (그림 2)는 4x4 블록에 수행되는 9가지의 모드이며 이 중 최상의 블록을 선택한다. 16x16 블록의 경우 4가지 모드의 예측을 수행하고, 그 중 최상의 블록을 선택한다. 따라서 1개의 MB를 코딩하기 위해서는 16개의 4x4블록 각각에 대하여 9가지의 모드 연산 수행, 16x16 블록의 4가지 모드 연산을 수행하여 최종 모드를 결정하게 된다.



(그림 2) 4x4 블록의 9가지 모드

2.2 트랜스코더를 위한 Intra 예측 기법

[4]에서는 DCT 기반 트랜스코더의 장점을 유지하기 위해 픽셀 상에서 Intra 모드를 선택하고 실제 예측 연산은 DCT 상에서 수행되도록 하였다. 이 경우 모드 예측을 위한 계산량은 H.264의 연산과 동일하게 소요된다.

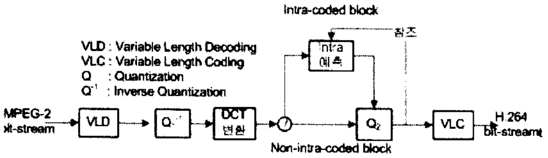


(그림 3) 최적 모드 선택을 위한 연산은 픽셀 상에서 수행하고 Intra 예측은 DCT 상에서 수행하는 트랜스코더 구조.

[5]에서는 H.264의 4x4 DCT 상에서 예측을 수행할 때 계수의 특징을 이용하여 선별적으로 예측할 수 있는 기법을 제안하였다. 그러나 16x16 블록 모드에 대한 연구가 없어 전체 Intra 예측을 위한 알고리즘이 요구된다.

본 논문에서는 DCT 상태에서의 트랜스코딩의 장점을 유지하면서 Intra MB의 빠른 모드 선택을 위한 기법을 제안한다.

3. DCT 기반의 빠른 Intra 모드 결정 기법



(그림 4) 열린 회로형의 DCT 기반의 트랜스코더

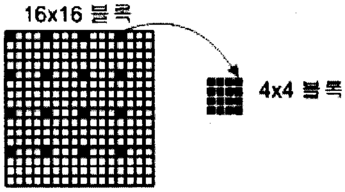
본 논문에서 가정하는 DCT 기반 트랜스코더 구조는 (그림 4)의 간단한 열린 회로형 트랜스코더이며 그 중에서 Intra MB를 위한 'Intra 예측' 블록의 모드 결정을 효율적으로 수행하기 위한 기법을 제안한다. 제안하는 기법은 크게 MB내의 모드 선택 과정과 선택된 모드의 하위모드 결정이다.

먼저 MB내의 모드 선택 과정을 살펴보자.

DCT 상에서의 Intra 모드 예측 연산은 각 모드에서 참조하는 값도 DCT 상태여야 하므로 계산복잡도가 픽셀 상의 연산보다 매우 높다. 따라서 DCT 특성을 이용하여 현재 MB 내의 픽셀 간의 연관성을 찾아내는 것이 중요하다. (그림 4)의 'DCT 변환' 블록에서는 8x8 DCT 블록을 4개의 4x4 정수형 DCT 블록으로 변환한다. 이때, intra MB 경우 'Intra 예측'을 수행하여 저장해야 할 비트의 값을 줄이게 되는데, 본 연구에서는 이 4x4 블록의 특징을 이용한다. 즉, 4x4 정수형 DCT 블록의 DC계수는 그 블록의 평균값을 저장하므로 MB내의 16개의 DC 계수는 그 MB를 1/16로 축소한 것과 같다. 이 특징을 이용하기 위해 MB내의 16의 DC 계수를 MB내의 위치 4x4 16개의 DC 계수에 다시 하다마드 변환을 수행하면 MB내의 에지 존재 여부를 확인할 수 있는 값들이 생성된다.

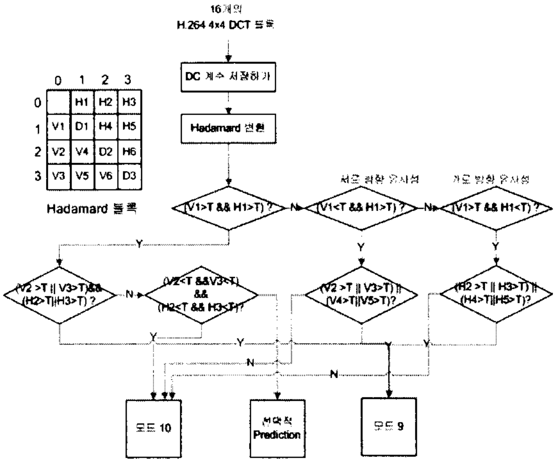
(그림 5)는 MB내에서 4x4 크기로 정수형 DCT된 블록에서 DC 계수를 추출하는 새로운 4x4 블록을 생성하는 과정을 보여준다. 이 블록을 DC 블록이라 하며 이 블록에 하다마드 변환을 수행한다.

하다마드 변환은 1과 -1로만 구성된 매트릭스를 이용하여 변환하는 것으로 H.264 정수형 변환에 비해 블록내 에지 여부를 좀 더 정확하게 판단할 수 있다. (그림 6)의 Hadamard 블록에서 V1, V2, V3 픽셀은 하다마드 변환 전의 블록에서 두 행을 더하여 부분합을 구하고 부분합의 차이값을 저장한 것이다.



(그림 5) DC 블록의 생성 과정

따라서 V1, V2, V3값이 작을수록 세로 방향으로 유사성이 높아진다. 동일한 방식으로 H1, H2, H3가 작을수록 가로 방향으로 유사성이 높아진다. MB 내에서 가로 또는 세로 방향으로의 유사성이 높다면 Intra 모드에서 16x16 예측(모드 10)을 수행할 가능성이 높으며 이와 반대인 경우(유사성이 낮은 경우) MB내에서 4x4 블록 단위로 예측(모드 9)을 수행할 가능성이 높아진다. (그림 6)은 이러한 특성을 이용하여 모드를 결정하는 알고리즘을 순서도로 나타낸 것이다. 이때, 두 모드의 특성을 포함하는 경우에는 각 모드의 예측을 직접 수행해야 한다.

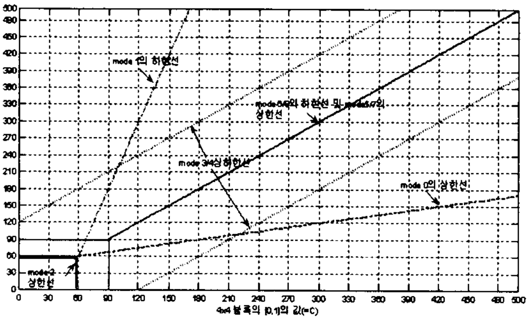


(그림 6) MB내의 예측 모드를 결정하는 알고리즘

MB의 예측 모드(모드10 또는 9)를 결정했다면 해당 모드의 하위모드를 선택해야 한다. 이때, 모드 10의 경우 가로 또는 세로 방향의 유사성이 존재하면 각각 해당 모드의 예측을 수행하게 된다. 그 외의 경우 나머지 모드 중 선택하게 된다.

모드 9(4x4 블록)의 경우 9가지의 모드를 계산해야 한다. 본 연구에서는 하다마드의 변환과 유사하게 AC 계수들의 값을 이용하여 가로 및 세로 방향 유사성을 판단하고 해당 모드만 연산을 수행한다. (그림 7)은 4x4 DCT 블록의 V1, H1의 값에 절대값을

취하여 각각 (|V1|, |H1|)을 좌표상에 나타내었을 때 실제 예측을 수행해야할 모드를 나타낸 것이다. 이는 다양한 실험을 통하여 경험적으로 계산한 것이다. 예를 들어 (V1, H1)이 (|50|, |-30|) 인 경우 mode0, 1, 2를 수행해야 한다. 그림에서 모드 3과 4



(그림 7) 4x4 블록에서의 모드 선택 기준

의 경우 동일한 그래프를 사용하고 있으나 모드 3의 경우 두 계수의 부호가 동일한 경우 모드 4의 경우는 다른 부호를 가질 때 사용한다. 또한 모드 5/7의 경우도 유사하게 각각 다른 부호, 동일부호를 가질 때 선택된다. 모드 6/8의 경우도 각각 다른 부호와 동일 부호일 때 선택된다. 이는 수직 또는 수평 방향으로의 유사성은 동일하나 오른쪽/왼쪽, 위/아래 방향으로의 유사성을 보이는 차이 때문에 발생하는 특성 때문이다.

4. 실험 및 성능 평가

현재 상세한 실험은 진행 중이며 Matlab을 이용하여 Akiyo, Foreman, Mobile 동영상의 첫 프레임에 대하여 실험한 결과는 아래와 같다. 표에서는 모드 9(4x4 블록단위의 예측 수행)과 모드 10(16x16 단위의 예측 수행)을 구분하는 순수 성공율/실패율을 나타낸 것이다. 순수 성공율/실패율은 판단이 모호하여 선택적인 모드를 계산하는 경우를 제외하고 (그림 6)의 알고리즘을 적용한 것을 의미한다. 실험에 적용된 T값은 18로 고정되어 있다.

<표 1> 제안한 알고리즘의 성공율

동영상 종류	Akiyo	Foreman	Mobile
프레임 당 MB 수	396	396	330
순수 성공수	306	307	310
순수 실패수	17	28	4

다양한 Threshold를 이용하여 실험한 결과 각 이미지마다 최상의 결과를 위한 Threshold가 달랐으며 Mobile과 같이 이미지내 인접 픽셀 간의 유사성이 낮을수록 T값도 낮은 것이 적합하다는 결론을 얻을 수 있었다.

5. 결론

본 논문은 DCT기반 MPEG-2/H.264의 Intra 예측을 수행하는 기법을 제안하였다. 먼저 MPEG-2 DCT를 H.264의 정수형 DCT로 변환한 후, DC 계수에 대하 하다마드 변환을 수행하여 MB내의 픽셀간 유사성을 판단한다. 이때, 계수의 특징이 모호하여 모드를 결정하기 어려운 경우는 몇 가지 모드를 선택하여 오류 값을 계산하여 모드를 결정한다. 현재 정밀한 실험은 진행 중이며, 여러 동영상의 첫 Intra 프레임에 대한 실험을 수행한 결과 MB의 모드 결정의 정확도는 프레임 내 이미지의 특징에 영향을 많이 받았다. 향후 이미지의 특성에 따라 모드를 결정하는 값을 유동적으로 설정하는 기법을 연구하여 정확도를 높이는 연구를 수행하도록 하겠다.

참고문헌

[1] Joint Video Team (JVT), "Final Draft Internationl Standard of Joint Video Specification",(ITU-T Rec. H.264| ISO/IEC 14496-10 AVC), 7th Meeting, Pattaya, Thailand, Mar. 2003

[2] Iain E. G. Richardson, "H.264 and MPEG-4 Video Compression", Wiley, England, 2003.

[3] T. Wiegand, G.J. Sullivan, "Overview of the H.264/AVC Video Coding Standard", IEEE Trans. on Circuits Syst, Video Technol., vol. 13, no. 7, pp. 560-576, Jul. 2003

[4] 강진미, 이주경, 정기동, "H.264 변환을 위한 DCT 도메인에서의 Intra 프레임 변환 기법", 한국정보과학회 2004 가을 학술발표논문집, 제31권, 제2호, pp.460-462, 2004년 10월

[5] Seong Soo Chun, Ja-Cheon Yoon, Sanghoon Sull, "Efficient Intra Prediction Mode Decision for H.264 Video", PCM 2005, Part I, LNCS 3767, pp.168-178, 2005