

하다마드 트랜스폼 차이를 이용한 빠른 화면 내 모드 결정 기법

*황웅 **김주혁 ***김태호 ****정제창

한양대학교 전자컴퓨터통신공학과

*gogowinner@naver.com **kjh76363@gmail.com

crewx@hanyang.ac.kr *jeong@hanyang.ac.kr

Fast intra mode decision algorithm by using hadamard transform difference

*Hwang, Ung **Kim, Joo-Hyuk ***Kim, Tae-Ho ****Jeong, Je-Chang

Dept. Electronics and Computer Eng. Hanyang University

요약

H.264/AVC에서는 율-왜곡 최적화 기법을 사용하여 부호화 성능을 이전 보다 대폭 향상시켰다. 율-왜곡 최적화 기법은 선택할 수 있는 모든 코딩 모드들의 비용을 계산한 후, 최소의 비용을 가진 모드를 선택하여 부호화를 한다. 따라서 코딩 모드들이 많아질수록 복잡도 역시 증가하며, 기존 표준보다 다양한 코딩모드를 지원하는 H.264에서는 복잡도가 더 많이 증가한다. 본 논문에서는 SAHTD(Sum of Absolute Hadamard Transform Difference)를 이용하여 화면 내 모드 결정부에서 빠른 화면 내 모드 결정을 하면서도 압축 손실이 적게 발생하는 방법을 제시하였다. 휘도 신호 16x16블록, 휘도 신호 4x4블록, 색차 신호 8x8블록 모두 SAHTD를 기준으로 최적의 모드가 결정된다. 휘도 신호 16x16블록에서 SAHTD값에 따라 4x4블록 부호화 과정을 조기 종결이 가능하게 하는 것과, 휘도 신호 4x4에서는 SAHTD 값이 가장 낮은 3개의 모드와 MPM(Most Probable Mode)을 후보 모드로 선택하여 율-왜곡 최적화를 수행하는 방식이 사용되었다. MPM이 SAHTD값이 가장 낮을 경우 MPM이 휘도신호 4x4블록의 최적의 모드로 선택되고 종료된다. JM 참조 소프트웨어를 통한 실험결과 제안된 기법은 기존 JM의 방식에 비해 화면 내 프레임의 부호화 시간의 84.7% 감소와 0.049dB의 PSNR 감소율, 0.595%의 비트 증가율을 보여주었다.

1. 서론

H.264/AVC(Advanced Video Coding)은 가장 최신의 비디오 코덱 표준이다. 이것은 ITU-T(International Telecommunication Union), VCEG(Video Coding Experts Group), ISO/IEC(International Organization for Standardization/International Electrotechnical Commission) MPEG 표준 위원회로 구성된 JVT(Joint Video Team)에서 개발하였다. H.264/AVC에서는 이전 비디오 코덱 표준들에 비해 높은 부호화 압축 성능을 자랑한다. 이것은 H.264/AVC에서 다양하고 복잡도가 높은 부호화기술들을 사용했기 때문인데, 이 중에는 다중참조 프레임(multiple-reference frame)의 사용과 여러 가지 화면 내 모드(multiple intra prediction modes)의 사용, 다양한 블록 사이즈(variable block sizes)의 사용, RDO(Rate-Distortion Optimization) 등과 같은 방법들이 있다[1]. 이러한 비약적인 복잡도의 상승 이유로 인해, 복잡도를 낮추면서도 화질은 그대로 유지하기 위한 다양한 속도 향상 알고리즘들이 제안되어 왔다[2]-[3].

Wang[2]은 DES(dominant edge strength)을 이용해 사용하는 예측 모드의 수를 줄이는 방법을 사용해 복잡도를 낮추었다. Tsai[3]은 주어진 블록 내의 여덟 가지 적응 특징을 찾아내 기술기의 정도를 판단해 가장 가능성 있는 모드를 찾아내는 방식으로 복잡도를 낮추었다. Zeng[4]은 SAHTD(Sum of Absolute Hadamard Transform

Difference)의 값을 이용하는 조기 종결 알고리즘을 제안하였다. 이것은 Wang[2], Tsai[3]과 같은 다른 속도 향상 알고리즘들에 비해 앞선 속도향상 효과와 낮은 화질 감소가 나타났다. 하지만 실험 결과 Zeng[4]의 휘도신호 4x4블록 모드 결정 방식은 개선의 여지가 있는 것으로 나타났다.

본 논문에서는 Zeng[4]가 갖고 있는 장점을 이용하면서도 SAHTD값으로 모드를 판단하는 점을 개선하여 다음과 같은 화면 내 모드 선택방법을 제시하였다. 색차블록 8x8에는 모든 모드의 SAHTD값을 구하여 최소 SAHTD값을 가진 모드를 최적의 모드를 선택한다. 휘도 블록 16x16에서는 모든 모드의 SAHTD값을 구하여 최소 SAHTD값을 가진 모드를 최적의 모드를 선택한다. 또한 최적 모드의 SAHTD값이 정의된 임계값 보다 작을 경우 휘도블록 4x4모드에 대한 예측모드는 모두 생략하고 16x16에서 선택된 모드를 최종 모드로 결정한다. 마지막으로 휘도블록 4x4에서는 모든 모드의 SAHTD값을 구하고 가장 SAHTD 값이 작은 3개의 모드를 선택한다. 여기에 MPM(Most Probable Mode)를 추가한 총 4개의 모드가 후보 블록으로 선택된다. 이 4개의 모드를 가지고 율-왜곡 최적화 방식으로 4x4의 최적의 모드를 선택한다. 최종적으로 휘도블록 16x16과 4x4중 최종 블록 크기 선택은 4x4블록의 16개 율-왜곡 값의 합과 16x16블록의 값 중 적은 것으로 선택된다.

본 논문의 구성은 다음과 같다. 2장에서는 기존의 화면 내 모드 결정 방식을 설명하고, 3장에서는 제안한 기법을 실험 데이터와 함께 설명한다. 4장에서는 본 알고리즘의 실험결과와 함께 이전 속도 향상 알고리즘들과 비교 평가한 실험결과를 보여주고, 5장에서는 결론을 맺는다.

2. 기존의 알고리즘

기존의 화면 내 예측은 부호화하고자 하는 블록 주위의 인접한 픽셀들을 이용하는데 이는 픽셀간의 간격이 가까울수록 높아지는 픽셀간의 공간적 상관관계(Spatial Correlation)를 이용하기 위한 것이다. 공간적 상관관계가 높으면 픽셀 간 비슷한 화소 값을 가질 확률이 높아 정확한 예측이 가능해지고, 이로 인해 예측오차가 줄어들어 부호화 효율이 향상 된다. H.264/AVC 에서는 기본적으로 휘도신호를 16x16과 4x4의 블록 중에서 선택하여 부호화하는데, 블록의 크기 선택은 윌-왜곡 최적화를 통해 이루어진다. 크기가 가변적인 부호화 블록의 선택을 하는 이유는, 부호화 하는 매크로블록 안의 픽셀들이 높은 공간적 유사성을 갖지 않을 때에 보다 부호화 단위를 세분화하여 예측픽셀과 부호화대상 픽셀간의 공간적 상관 관계를 향상시킬 수 있도록 만든 것이다. 최종적으로 윌-왜곡 최적화를 통해 최적의 블록크기와, 블록에 따른 최적의 모드를 각각 판단할 수 있게 한다.

휘도 신호의 화면 내 예측은 그림 1과 같이 16x16 블록 단위에서는 Vertical, horizontal, DC, Plane의 총 4가지 모드로 이루어지고, 4x4블록 단위에서는 DC를 제외하고는 8가지 방향을 사용하여 총 9가지 모드로 이루어진다.

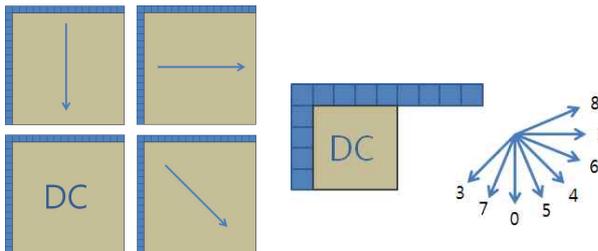


그림 1. 16x16의 4가지 모드(a), 4x4의 9가지 모드(b)

색차 신호의 8x8 블록 단위에서 화면 내 예측은 휘도 신호의 16x16 블록 예측과 비교했을 때 유사한 4가지의 모드로 이루어져 있다. 각각의 모드 방향에 따라 예측된 결과는 다음을 같이 윌-왜곡 최적화 공식에 의해 판단한다.

$$J = D + \lambda \cdot R \quad (1)$$

D는 부호화-복호화 이후의 재구성된 값과 원본과의 전체 왜곡이며, λ 는 라그랑주 승수이고 R은 복호화 이후의 비트율이다. JM에서는 복잡도를 줄이기 위하여 최적의 모드를 선택하는 다른 방안을 다음과 같이 지원한다. P는 MPM일 경우 1, 아닐때는 0이다.

$$J_{SATD} = SATD + 4 \cdot P \cdot \lambda_{MODE} \quad (2)$$

식(2)의 SATD는 다음과 같이 정의된다.

$$SATD = \sum_i^N \sum_j^N |c_{ij}| \quad (3)$$

식(3)의 c_{ij} 로 구성되는 C 행렬은 다음과 같이 정의된다.

$$C = H(A - B)H^T \quad (4)$$

여기서 H는 Hardmard Transform이며 내용은 다음과 같다.

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \quad (5)$$

3. 제안하는 알고리즘

제안하는 알고리즘은 고속으로 화면 내 예측을 하기 위하여, Zeng [4]이 제안한 방식을 기반으로 개선하였다. 휘도 신호 16x16 블록에서는 4가지 모드의 예측모드에 대한 각 SAHTD 값을 계산한 후 가장 작은 SAHTD 값을 갖는 모드를 최적의 예측 모드로 결정 한다. 이때 최소 SAHTD 값이 임계값 이하인 경우에는 변화가 적은 평탄한 영역으로 간주하여 16x16 블록에서 결정된 예측모드를 사용하여 부호화를 진행하고, 휘도 신호 4x4 블록의 예측 모드는 윌-왜곡 최적화 기법의 후보 모드로 선택되지 않는다. 임계값은 다음과 같이 계산할 수 있다.

$$T = 50 \cdot QP + 400 \quad (6)$$

임계값을 계산하는 공식은 실험적으로 결정되었다. 16x16블록이 4x4블록보다 부호화 효율이 높아 최종 부호화 블록 크기로 결정되었을 때 최적의 모드가 가지는 SAHTD 값이 위 임계값 이하일 확률이 90%정도가 되었다.

Zeng[4]의 휘도신호 화면 내 4x4 블록의 모드 결정 방식은 다음과 같다. 8개의 모드는 방향성에 따라 다음과 같은 4개의 그룹으로 나뉘어진다. {0,5,4}, {0,7,5}, {1,6,4}, {1,8,3}. 만약 0번 모드가 1번 모드보다 SAHTD 값이 작으면 1, 2번 그룹이 선택되고 그 다음의 5, 7번 모드 중 5번 모드가 SAHTD 값이 작으면 1번 그룹이 후보모드 그룹으로 선택된다. 이와 같이 계층적으로 SAHTD 값에 따라 3가지 모드를 가지는 하나의 그룹을 결정한다. DC모드는 그룹에 상관 없이 항상 포함되어 총 4가지의 후보 모드가 결정된다. 만약 MPM이 후보모드들과 비교해서 가장 작은 SAHTD 값을 가지는 경우 윌-왜곡 최적화 과정 없이 MPM을 최종 모드로 결정한다. 후보모드로 선택된 모드들만으로 윌-왜곡 최적화 방식을 이용하여 최적의 모드를 결정한다. 최종적으로 휘도신호의 4x4 블록 후보모드 중에서 윌-왜곡 최적화 공식을 통해 최적의 모드를 찾아 최종 모드로 결정한다. 이 과정을 매크로블록 내의 모든 휘도신호 4x4 블록 블록에 대해 수행한다. 이 후 최종 블록크기는 최적모드가 결정된 16x16 블록의 윌-왜곡 값과, 모든 4x4 블록들의 윌-왜곡 값을 비교해서 값이 더 낮은 블록으로 결정한다.

| 후보모드 | 예측성공률(%) |
|--------------|----------|
| 기존 Zeng(4후보) | 79.63 |
| SAHTD(1) | 53.65 |
| SAHTD(2) | 74.19 |
| SAHTD(3) | 84.04 |
| SAHTD(4) | 89.68 |
| MPM | 54.94 |
| SAHTD(2)+MPM | 90.84 |
| SAHTD(3)+MPM | 94.67 |

표 1 : 기존 Zeng[4] 방식의 후보모드 결정기법과 가장 작은 SAHTD 값을 갖는 모드부터 후보모드 결정기법과 MPM을 추가로 후보모드로 두는 결정기법의 율-왜곡 최적화 방식과의 결정모드 일치률

이 알고리즘의 휘도신호 4x4 블록의 화면 내 예측의 수행 방식에서는 몇 가지 문제가 발견되었다. 첫째, 최소 혹은 낮은 SAHTD값을 가지는 모드가 후보 모드에 들어가지 못할 가능성이 있었다. 표 1에서 확인해 볼 수 있는 것은 가장 작은 SAHTD 값을 가지는 모드가 최적의 모드일 확률은 53.65%로 MPM과 비슷할 정도로 높았다. 이를 후보 모드에 항상 넣지 않는 것은 옳지 않은 판단으로 볼 수 있었다. 표 1의 통계는 4장의 실험결과와 동일한 영상과 동일한 조건에서 얻은 결과이다. 둘째, MPM은 최종모드로 결정될 확률이 가장 높은 모드인데 후보 모드에 포함되지 않았다. 최적의 모드가 될 확률이 가장 높은 MPM이 후보모드에 들어가지 않은 것도 상당한 손해라 판단되었다. 표 1에서 MPM이 최적의 모드일 확률과 이를 낮은 SAHTD값을 가지는 모드 순으로 후보모드를 선택한 방식과 결합하여 적용한 것의 실험 결과를 볼 수 있다. SAHTD가 작은 순으로 2-3를 선택하고 MPM을 항상 선택하는 것이 기존 방식보다 개선된 방식으로 나타났다. 본 논문에서는 SAHTD가 작은 순으로 3개와 MPM을 후보모드로 채택하는 방식을 채택하였다. 자세한 과정은 다음과 같다.

1. 휘도신호의 모든 화면 내 4x4 모드에 대해서 SAHTD를 구한다.
2. MPM이 가장 낮은 SAHTD 값을 가질 경우 MPM을 최적의 모드로 결정하고 종료한다. 이 조건이 아닐 경우 다음 과정을 계속한다.
3. SAHTD가 낮은 순으로 3개의 모드를 후보모드로 결정한다.
4. MPM이 이 3개의 후보에 포함되었을 경우는 3개의 모드를 가지고 율-왜곡 최적화 방법을 사용하여 최종 모드를 결정하고, MPM이 이 3개의 후보에 들어있지 않은 경우 MPM을 후보모드에 추가하여 4개의 후보모드로 율-왜곡 최적화 방법을 사용하여 최종 모드를 결정한다.

색차신호의 화면 내 8x8 예측은 휘도 신호 16x16에서 사용했던 방법과 유사하게 4가지의 예측 모드중 가장 작은 SAHTD 값을 갖는 모드를 최종 모드로 결정한다.

4. 실험결과 및 분석

제안하는 알고리즘의 구현은 JM 13.2 참조 소프트웨어[5]에서 이루어졌다. 실험에 사용된 컴퓨터는 Intel(R) Core(R)2 Quad CPU Q9550 2.83 GHz와 램 4 GB로 구성되었다. 실험에는 QCIF(foreman,

news, container, silent, coastguard), CIF(forman, paris, mobile, tempete, stefan), SIF(blowingbubbles, racehorse, basketballpass, bqsquare)의 영상들이 이용되었다.

각 영상은 각각 영상의 최대 프레임 개수인 250, 300, 500, 600의 모든 프레임을 사용하여 인코딩하였다. 모든 영상의 프레임은 화면 내 프레임으로 인코딩하였다. QP는 22,27,32,37를 적용하였다. 또한 CABAC 엔트로피 부호화 기술을 사용하였다.

| | BD-Bitrate (%) | BD-PSNR (dB) | Time reduction (%) |
|----------|----------------|--------------|--------------------|
| Wang[2] | 3.371 | -0.260 | 0.490 |
| Tsai[3] | 1.983 | -0.153 | 0.597 |
| Zeng[4] | 2.171 | -0.170 | 0.877 |
| Proposed | 0.650 | -0.050 | 0.859 |

표 2 : QCIF(176x144)영상들에서의 실험 결과

| | BD-Bitrate (%) | BD-PSNR (dB) | Time reduction (%) |
|----------|----------------|--------------|--------------------|
| Wang[2] | 2.570 | -0.232 | 0.551 |
| Tsai[3] | 1.695 | -0.152 | 0.598 |
| Zeng[4] | 2.011 | -0.172 | 0.862 |
| Proposed | 0.609 | -0.057 | 0.847 |

표 3 : CIF(352x288)영상들에서의 실험 결과

| | BD-Bitrate (%) | BD-PSNR (dB) | Time reduction (%) |
|----------|----------------|--------------|--------------------|
| Wang[2] | 2.407 | -0.173 | 0.509 |
| Tsai[3] | 1.808 | -0.131 | 0.610 |
| Zeng[4] | 1.738 | -0.125 | 0.850 |
| Proposed | 0.526 | -0.039 | 0.835 |

표 4 : SIF(416x240)영상들에서의 실험 결과

실험결과 Wang[2]과 Tsai[3]는 전체 평균 51.7%, 60.2%의 시간 감소율로 86.3%의 시간 감소율을 보이는 Zeng[4] 과 비교했을 시에는 낮았으며 비트증가율은 비슷했다. 제안한 알고리즘은 0.595%의 비트증가율을 나타내어 1.97%의 비트 증가율을 보이는 Zeng[4]보다 개선이 많이 된 것을 볼 수 있었다. 제안한 알고리즘은 84.7%의 시간 감소로 86.3%의 시간감소율을 보이는 Zeng[4]보다 약간 낮지만 이것은 제안한 알고리즘을 구현한 코드의 최적화문제 이거나 컴퓨터 시간의 오차 범위로 보여진다. 이론상으로 휘도신호 4x4블록 부호화에서 3-4개의 후보군이 선택되는 제안된 알고리즘이 항상 4개의 후보모드를 갖는 Zeng[4]에 비해 나은 결과를 나타내는 것이 옳다고 여겨진다.

5. 결론

본 논문에서는 H.264 빠른 화면 내 프레임 부호화를 위한 향상된 화면 내 모드 결정 기법을 제안하였다. 휘도신호 화면 내 16x16의 빠른 화면 내 모드 결정을 위해서 SAHTD를 사용하여 최종 모드를 결정하

고 최종 모드의 SAHTD 값이 일정 경계 값 이상일 경우 16x16을 부호화에 사용할 최종 블록으로 결정하여 휘도신호의 4x4 화면 내 모드 결정 단계를 생략하였다. 색차신호의 8x8 블록 부호화는 각 모드의 SAHTD를 구하여 최소의 SAHTD 값을 가지는 모드가 색차신호 8x8 화면 내 모드의 최종 모드로 결정되었다.

4x4화면 내 모드 결정에 있어서는 SAHTD 값이 낮을수록 최적의 모드일 확률이 높아 낮은 순서대로 3개의 모드를 후보모드로 결정하였다. 또한 MPM이 가장 최종모드가 될 확률을 높으므로 이를 항상 후보 모드에 포함시켰다. 제안된 알고리즘을 통해 부호화를 수행한 결과, 화면 내 프레임의 부호화 시간 84.7% 감소, 0.049dB PSNR 손실, 0.595%의 비트증가율이 평균적으로 나타내는 것을 확인했다. 이 실험결과는 제안된 알고리즘이 빠른 부호화를 시도하면서도 최적의 모드를 선택할 확률을 증가시킴으로 빠른 부호화로 인한 손실을 최소화하였음을 보여주었다.

참조문헌

- [1] T. Wiegand, G. J. Sullivan, G. Bjonte, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560 - 576, Jul. 2003.
- [2] J. C. Wang, J. F. Wang, J. F. Yang, and J. T. Chen, "A fast mode decision algorithm and its VLSI design for H.264/AVC intra-prediction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 10, pp. 1414 - 1422, Oct. 2007.
- [3] A.-C. Tsai, A. Paul, J.-C. Wang, and J.-F. Wang, "Intensity gradient technique for efficient intra-prediction in H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 5, pp. 694 - 698, May 2008.
- [4] H. Q. Zeng, K. K. Ma, C. H. Cai, "Hierarchical Intra Mode Decision for H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 6, pp. 907-912, June 2010.
- [5] Joint Video Team. Reference Software JM13.2 [Online]. Available: http://iphome.hhi.de/suehring/tml/download/old_jm/

* 이 논문은 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(2011-0011313)