

일반논문-07-12-3-06

스케일러블 비디오 부호화에서 통계적 가설 검증 기법을 이용한 고속 화면간 모드 결정

이범식^{a)}, 김문철^{a)†}, 함상진^{b)}, 김병선^{b)}, 이근식^{b)}, 박근수^{b)}

Fast Intermode Decision for Scalable Video Coding using Statistical Hypothesis Testing

Bumshik Lee^{a)}, Munchurl Kim^{a)†}, Sangjin Hahm^{b)}, ByungSun Kim^{b)}, Keunsik Lee^{b)},
and Keunsoo Park^{b)}

요 약

본 논문에서는 가변 움직임 블록을 이용한 움직임 예측과 계층적 B-화면 구조가 결합되어 많은 연산량이 요구되고 있는 SVC 구조에서 통계적 가설 검증 방법을 적용하여 화면간 움직임 모드 결정을 효율적으로 수행함으로써 부호화 과정의 연산량을 크게 줄일 수 있는 화면간 모드 고속 결정 알고리즘을 소개한다. 제안된 방법은 16x16 매크로 블록과 8x8 하위 매크로 블록에 통계적 가설 검증 기법을 적용하여 실행되며, 현재 블록과 복원된 참조 블록간의 화소값을 비교하여 율-왜곡 최적화 (RDO: Rate Distortion Optimization) 기반 움직임 모드 결정을 고속으로 수행함으로써 SVC의 부호화 과정에서 소요되는 부호화 시간을 전체 연산량 대비 최대 69%까지 감소시킨다. 그러나 연산량 감소에 따른 비트율의 증가나 화질의 열화는 무시할 수 있을 정도로 적음을 실험을 통해 확인하였다.

Abstract

In this paper, a fast intermode decision scheme is introduced by efficiently performing the mode decision using statistical hypothesis testing for hierarchical B-picture coding of SVC, in which much computational power is expended for combined variable block sizes and hierarchical B-pictures. The hypothesis testing in the proposed method is performed on 16x16 and 8x8 blocks to have early termination for RD computation of all possible modes. The early termination in intermode decision is performed by comparing the pixel values of current blocks and corresponding motion-compensated blocks. The proposed scheme exhibits effective early termination behavior in intermode decision and leads to a significant reduction up to 69% in computational complexity with slight increment in bit amounts. The degradation of visual quality turns out to be negligible in terms of PSNR values.

Keywords : Scalable Video Coding, H.264, AVC, hypothesis testing, intermode decision

1. 서 론

a) 한국정보통신대학교(ICU)
Information and Communications University
b) 한국방송공사(KBS)
Korea Broadcasting System
† 교신저자 : 김문철(mkim@icu.ac.kr)

스케일러블 비디오 부호화(SVC: Scalable Video Coding)^[1]
는 H.264/AVC의 확장된 부호화 방식으로 시간, 공간 및

화질적 계위성(Temporal, Spatial and Quality Scalability)을 지원한다. SVC는 ISO/IEC의 MPEG(Moving Picture Expert Group)과 ITU-T의 VCEG(Video Coding Expert Group)에 의해 구성된 JVT(Joint Video Team)에 의해 표준화가 거의 완료 상태에 있다. SVC는 시간 계위성(Temporal Scalability)을 지원하기 위해 H.264/AVC의 B-화면의 부호화 구조를 확장한 계층적 B-화면 구조를 채택하였다. 그림 1에서는 SVC에서 시간 계위성을 지원하기 위해 채택된 계층적 B-화면의 개념도를 보여주고 있다. 계층적 B-화면 기술은 다양한 시간 해상도를 지원하기 위한 시간 분해 구조로 폐루프(closed loop)구조를 기반으로 하며, 시간 계위에 따른 분해 레벨(decomposition level)을 갖는다^[2]. SVC의 계층적 B-화면은 주화면(key picture)인 I/P 화면 사이에 B-화면을 삽입함으로써 시간 계위성을 얻을 수 있다. 그림 1은 GOP의 크기가 $8(=2^N, N=3)$ 인 계층적 B-화면 GOP구조의 개념도를 보여준다. 여기서 N 은 정수로 나타낼 수 있으며 총 $(N+1)$ 개의 시간 해상도를 가지는 비트스트림으로 부호화할 수 있다. 그림 1에서 I와 P 화면은 주화면으로 부호화되며 시간 계위는 0이다. 이 두개의 화면만으로 가장 시간 해상도가 작은 비트스트림을 구성할 수 있다. 두 번째로 B_1 은 시간 계위가 1이며, B_2 는 시간 계위가 2, 마지막으로 B_3 은 시간 계위가 가장 낮은 3이 된다. 따라서 보다 다양한 시간 해상도를 구성하기 위해서는 주화면 사이에 보다 많

은 B-화면을 삽입해야 하며 이는 B-화면 부호화시 수행해야 할 양방향 움직임 벡터 추정, 양방향 다중 참조 화면등을 사용하기 때문에 연산량과 복잡도는 H.264/AVC에 비하여 크게 증가하게 된다. SVC의 기본 계층은 H.264/AVC와 호환적이므로, 움직임 예측과 모드 결정과정에서 $16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8, 8 \times 4, 4 \times 8$ 그리고 4×4 와 같은 7개의 서로 다른 크기를 갖는 블록을 사용한다. SVC에서 사용되고 있는 계층적 B-화면 구조는 다양한 시간 계위성을 얻기 위해 주화면인 I와 P 화면을 제외하고는 한 GOP(Group of Pictures)내에서 모두 B-화면을 사용하므로 H.264/AVC와 비교했을 때 연산량 증가와 함께 부호화 지연도 급격히 증가한다. B-화면은 양방향 움직임 벡터인 LIST0와 LIST1을 사용하고 다중 참조화면 사용시 양방향에서 모두 사용하므로 주화면 사이에서 많은 계층적 B-화면의 사용은 SVC 부호화기의 연산량을 현격하게 증가시키는 요인으로 작용한다.

현재까지는 H.264/AVC의 부호화 과정에서 연산량 감소를 위해 많은 연구들이 진행되어 왔다. Libo Yang은 무 움직임 검출(ZMD, Zero Motion Detection)에 기반하여 초기에 움직임 벡터 추정과 모드 결정을 완료하는 방법을 제안하였다^[3]. Wu는 모드 결정을 초기에 완료하기 위한 방법으로서 균일한 영역과 움직임이 적은 부분을 검출하는 텍스처 분석 방법을 사용하였다. 텍스처 분석방법으로 소벨연

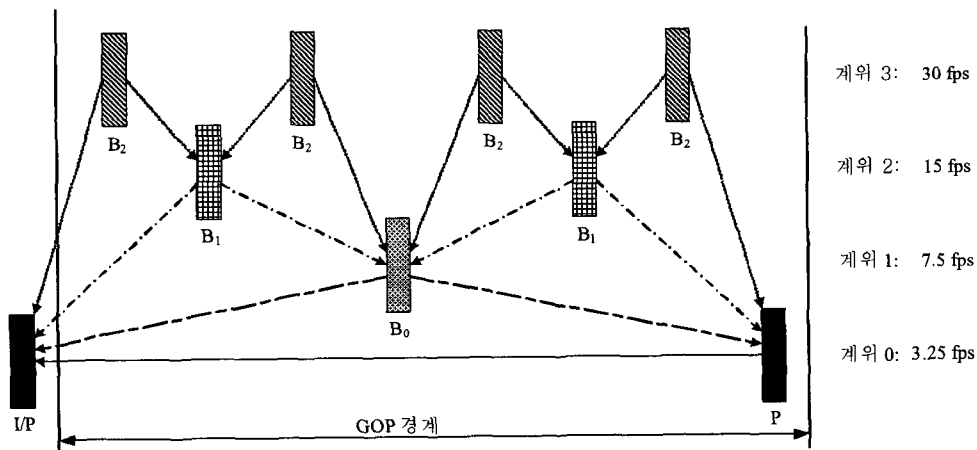


그림 1. 계층적 B-화면의 구조
Fig. 1. The structure of hierarchical B-pictures

산자를 사용하였고 SAD 값을 이용하여 현재 블록이 정적인 영역인지 아닌지를 판단한다. 이러한 분석방법으로 정적인 영역에 큰 블록을 할당하고 그렇지 않은 경우에 작은 블록을 할당하는 방법을 제안하였다^[4]. Jing은 전체 화면 대비 현재 매크로 블록이 얼마나 균일한 영역인지를 판단하여 블록 크기를 결정하는 방법을 사용하였다^[5]. 또한 [6]에서 AZCB(All Zero Coefficient Block)을 검출하여 고속 화면간 모드를 결정하는 방법이 제안되었다. 이것은 정수변환 후 얻어진 양자화 계수의 합이 임계치 이하이면 그 블록의 양자화 계수를 0으로 간주하는 방법이고 그 조건이 만족하면 초기에 모드 결정 과정을 완료함으로써 부호화 속도를 증가시키는 방법을 제안하였다. 그러나 기존의 제안된 방법들은 실험을 통한 임계치를 설정해야하며 임계치를 바르게 설정하지 않으면 율-왜곡(Rate-Distortion, 이하 RD) 성능이 나빠지는 단점이 존재한다.

본 논문에서는 현재 부호화되는 블록의 화소값과 움직임 보상된 참조 블록의 화소값을 통계적 가설 검증법으로 비교하여 움직임 벡터의 측정 및 모드 결정과정을 초기에 완료하여 고속 화면간 모드 결정을 달성하는 알고리즘에 대하여 소개한다. 본 논문은 다음과 같이 구성된다. 2장에서는 SVC에서의 화면 간 모드 결정 방법이 소개된다. 3장에서는 통계적 가설 검증기법에 대하여 소개하고, 통계적 가설 검증 기법을 이용하여 제안된 고속 모드 결정 알고리즘

을 소개한다. 5장에서는 제안된 알고리즘을 이용하여 복잡도 감소와 성능유지 측면에서 실험결과를 보인다. 6장에서 결론으로 본 논문을 마친다.

II. SVC에서 화면간 모드 결정 방법

1. RDO 기반의 모드 결정

그림 2는 H.264/AVC 기반의SVC 부호화기의 구조를 나타낸다. 화면간의 중복성을 제거하기 위하여, 화면간 예측이 연속된 화면 사이에서 수행된다. 이전의 비디오 부호화 표준에서는 움직임 벡터 추정이 16x16 매크로 블록 단위로 수행되었다. 그러나 16x16 매크로 블록 만으로는 다양한 비디오 움직임 및 텍스처의 특성을 반영하지 못하므로, H.264/AVC 비디오 표준에서는 16x16, 16x8, 8x16, 8x8, 8x4, 4x8, 4x4 등 총 7가지 다양한 크기를 갖는 블록을 사용하여 움직임 추정의 정밀도를 향상하였다. 다시 말해 16x16 블록과 같이 블록 사이지가 큰 경우는 움직임이 거의 없거나 균일한 영상 특성을 보이는 영역에 대한 움직임 추정에 사용되는 경향을 보이며 8x8~ 4x4 같이 작은 블록들은 움직임이 큰 영역이나 세밀한 특성을 보이는 영역에서의 움직임 추정에 이용된다. 이러한 블록 사이즈의 결정은 모든

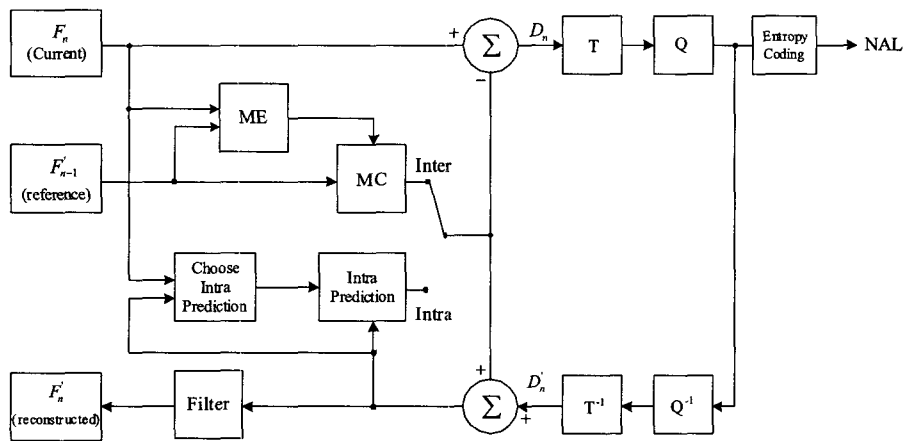


그림 2. H.264/AVC 기반 SVC 부호화기^[4]
Fig. 2. H.264/AVCbased SVC encoder

블록에 대하여 RD 비용을 구하여 최소의 RD 비용을 갖는 블록을 할당하도록 하는 원칙을 따른다.

그림 3은 계층적으로 구분되는 다양한 크기를 갖는 블록들을 보여준다. 첫 번째 단계에서는 16x16, 16x8, 8x16 등 비교적 큰 크기를 갖는 블록이 있으며, 하위 단계에서는 8x8, 8x4, 4x8, 4x4 와 같이 작은 크기를 갖는 블록이다. 이렇게 다양한 블록 크기를 갖는 모드 결정과 움직임 벡터 결정으로 연산량이 급격히 증가하였다. 왜냐하면, 각 블록 별로 각각의 움직임 벡터가 존재하며 하나의 매크로 블록에서 블록 사이즈를 결정하기 위하여 7개의 모든 블록에 대하여 RD 비용을 구하는 과정의 복잡도가 크기 때문이다. 이는 부호화 효율성을 높이기 위해 요구되는 연산량의 증가 비용이다.

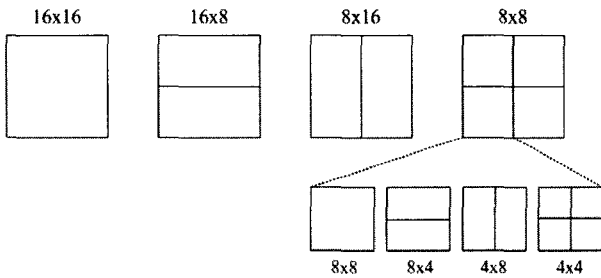


그림 3. 다양한 크기를 갖는 블록 들
Fig. 3. block mode with various sizes

SVC에서는 화면간 모드 결정이 GOP 단위로 수행되며 가장 적절한 블록 크기를 결정하는 것은 다음 식을 최소화 하는 블록을 최적의 블록으로 판단하는 것이다

$$J(s, c, Mode | QP, \lambda_{mode}) = SSD(s, c, Mode) + \lambda_{mode} \cdot R(s, c, Mode | QP) \quad (1)$$

여기서 s는 원영상 c는 복원 된 영상을 의미한다. QP는 화면간 모드 결정 시 사용되는 양자화 계수를 의미한다.

λ_{mode} 는 라그랑지 계수로서 $0.85 \times 2^{\frac{QP-12}{3}}$ 와 같이 계산된다

[1]. SSD(sum of Squared Differences)는 원영상과 복원된 영상간의 왜곡도를 나타내며, 값이 클수록 왜곡이 심하며 모드 결정이 잘못된 것을 의미한다. 그 것은 다음과 같이

계산된다.

$$SSD(s, c, Mode | QP) = \sum_{x=1}^{16} \sum_{y=1}^{16} (s_Y[x, y] - c_Y[x, y, Mode | QP])^2 + \sum_{x=1}^{16} \sum_{y=1}^{16} (s_U[x, y] - c_U[x, y, Mode | QP])^2 + \sum_{x=1}^{16} \sum_{y=1}^{16} (s_V[x, y] - c_V[x, y, Mode | QP])^2 \quad (2)$$

SVC 부호화기에서는 양자화 계수가 시간 계위에 따라 다르게 적용된다. 이것은 시간으로 다양한 비트율을 추출 할 수 있도록 계층적으로 시간 계위를 구성하지 않는 H.264/AVC 부호화기와는 다른 방식이다. 그림 1에서 보는 바와 같이 2^N 개의 화면으로 구성되는 GOP 구조에서 최대 시간적 계위 수는 (N+1)이다. QP_0 를 초기에 설정한 양자화 계수라고 하면 각각의 시간 계위에서 적용되는 양자화 계수는 다음과 같이 구할 수 있다[1].

$$QP(l) = \begin{cases} Round(QP_0 - 1.7 \cdot (N - 1 - l)), & \text{if } l = 0 \\ Round(QP_0 + 2 - 1.7 \cdot (N - 1 - l)), & \text{if } l \neq 0 \end{cases} \quad (3)$$

여기서 l은 시간 계위를 나타내며 l=0이면 한 GOP내에서 연속하여 참조할 수 있는 두 화면의 시간거리가 가장 멀고 주화면을 구성할수 있는 시간 계위이다. 그리고 QP_0 는 SVC 부호화기에 적용되는 양자화 계수의 초기값이다.

또한 Round()는 실수를 가장 가까운 정수로 나타내기 위한 연산자이다[1]. 이것은 서로 다른 시간 계위에 대하여 각각 다른 양자화 계수가 적용됨을 의미한다. 예를 들어, 3가지의 서로 다른 시간 해상도를 만들어 낼 수 있는 GOP 크기가 4인 GOP 구조에서 QP_0 가 초기값으로 28이 설정되어 있다면, 첫 번째 시간 계위(l=0) 즉, 주화면에 대해서는 양자화 계수가 25가 적용되고, 두번째 계위(l=1)에서는 27, 그리고 마지막 계위(l=3)에서는 28이 적용된다. 각 시간 계위 별로 다른 양자화 계수가 결정되는 이유는 각 화면이 서로 다른 중요도를 가지고 있기 때문이다. GOP를 구성하는 주

화면은 GOP 내의 모든 B 화면의 참조 화면이 될 수 있거나 직간접적으로 영향을 미칠 수 있으므로 가장 중요도가 높다^[8]. 그리고 계위가 높은 B 화면일수록 참조되는 수가 줄어들므로 중요도가 감소하게 된다. 따라서 계위가 높은 B 화면에 대해 양자화 계수값을 증가시켜 적용하게 된다. 다음 식은 GOP내의 각 화면의 중요도와 양자화 계수와의 관계를 보여준다.

$$\text{주화면(Key Pictures)} > B_1 > B_2 > B_3 \quad (4)$$

식 (4)는 화면의 중요도를 나타내며 이러한 중요성을 감안하여 양자화 계수를 다음 (5)와 같이 다르게 적용한다.

$$QP_{\text{key}} < QP_1 < QP_2 < QP_3 \quad (5)$$

여기서 QP_{key} 는 주화면에 적용되는 양자화 계수를 나타내며 QP_1, QP_2, QP_3 는 각각의 증가된 계위의 B화면의 부호화에 적용되는 양자화 계수이다.

2. 다양한 크기를 갖는 블록의 통계적 특성

통계적 가설 검증을 이용한 고속 화면간 모드 결정방법

을 적용하기에 앞서 각 시험 영상에 대한 모드 결정분포에 대하여 알아보기로 한다. 이것은 H.264/AVC 기반의 SVC 부호화기가 서로 다른 블록 크기를 사용함으로써 부호화 효율을 증가시키기는 하나 복잡도 또한 큰 폭으로 증가되어 영상의 특성과 시간 계위에 따라 화면간 모드 결정이 어떻게 분포하는지를 알아보기 위함이다. 그리고 본 통계적 특성을 이용하여 화면간 모드 결정시 복잡도를 개선하는 제안 알고리즘을 소개한다.

표 1은 서로 다른 7가지 블록 모드를 사용한 B-화면 부호화시 화면간 모드 결정에 대한 통계적 분포를 보여준다. 여기서 설정된 양자화 계수 QP는 28과 36을 사용하였으며, *Mother&Daughter*, *Foreman*, *Harbour* 영상이 사용되었다. 표 1에서 8x8 모드는 8x8 모드와 그 하위 모드인 8x4, 4x8, 4x4 모드를 포함한다. *Mother&Daughter*와 같이 많은 부분에서 고정된 배경과 정적인 움직임을 가지는 영상은 DIRECT와 16x16 모드와 같은 큰 블록들에 의해 화면간 모드가 결정되는 비율이 매우 크다. 반면 *Foreman*과 *Harbour*처럼 움직임이 크거나 미세한 영역을 포함하는 영상들은 양자화 계수 28을 적용하였을 경우 8x8 블록모드와 같은 작은 블록들로 선택되는 비율이 27%이상으로 나타난다. 이것은 빠른 움직임 또는 미세한 부분으로 구성된 영상 영역을 부호화할 때, 작은 블록들로 선택되어 움직임 예측 및 보상을 하는 것이 RD 비용 측면에서 유리하기 때문이

표 1. 시간 계위에 따른 다양한 블록 크기의 통계적 분포
Table 1. Statistical distribution of block modes according to the temporal scalability level

시험영상	QP	시간계위	Direct (%)	16x16 (%)	16x8 (%)	8x16 (%)	8x8 (%)
Mother&Daughter	28	1	52.6	22.8	8.2	7.2	9.3
		3	86.2	8.7	1.8	1.8	1.5
	36	1	71.2	17.7	4.0	4.7	2.4
		3	94.2	4.3	0.6	0.8	4.3
Foreman	28	1	15.7	24.0	13.2	19.3	27.8
		3	42.3	34.3	9.1	8.8	5.4
	36	1	26.8	34.3	15.4	13.0	10.6
		3	65.7	26.3	4.0	3.2	0.7
Harbour	28	1	24.8	25.5	6.0	16.6	27.1
		3	38.8	32.7	7.6	13.5	7.3
	36	1	43.2	16.2	4.5	8.9	16.2
		3	65.5	27.5	2.1	3.7	1.2

다. 선택된 블록 모드의 분포는 양자화 계수에 따라 다른 패턴을 갖는다. 같은 영상에 대하여 더 큰 양자화 계수를 적용하였을 때 보다 큰 블록들로 선택되는 비율이 커지게 된다. 이것은 큰 양자화 계수로 양자화 했을 경우, 영상의 미세한 부분을 잃게 되어 작은 블록으로 움직임 보상을 하는 블록의 수가 줄어 들기 때문이다. 표 1에서 흥미로운 것은, 같은 영상에 대하여, 블록 모드의 분포가 시간 계위에 따라 서로 다른 양상을 보인다는 점이다. 표 1에서 보는 것처럼 모든 시험 영상에 대하여 시간 계위가 낮을수록 보다 작은 블록으로 움직임 보상을 하는 것을 알 수 있다. 이것은 시간 계위가 낮은 영상은 GOP내의 다른 화면에 대해 움직임 예측 및 보상시 직간접적으로 영향을 주기 때문에 보다 높은 화질을 유지해야 하기 때문이다. 만일 주화면 또는 시간 계위가 낮은 영상이 좋은 화질을 유지하지 못한다면 GOP내의 전체 화면의 화질에 영향을 미치므로 시간 계위가 낮을수록 양자화 계수 값을 작게 적용하게 화질을 높게 유지해야 하는 제약이 따른다. 이것은 식 (3), (4), (5) 에서 보는 것처럼 시간 계위에 따라 서로 다른 양자화 계수를 적용함으로써 이루어진다. 표 1에서 보는 것처럼 화면간 블록 모드 결정의 통계적 특성은 양자화 계수와 영상의 특성 그리고 시간 계위에 따라 서로 다른 특성을 보인다. 한편 화면간 블록 모드를 결정하기 위해 실제로 많은 불필요한 연산이 존재하며 이를 줄이는 연구가 필요하다. 본 논문에서는 시간 계위에 따른 특성을 고려한 계층적 B-화면의 부호화에 적용되는 고속 화면간 모드 결정방법에 대하여 소개한다.

III. 통계적 가설 검증

통계적 가설 검증을 움직임 블록 고속 결정에 응용하고자 하는 이유는 다음과 같다. 현재 부호화 하고자 하는 16x16 매크로블록의 화소 값들과 움직임 추정 및 보상을 통해 얻은 16x16 참조 매크로블록의 화소값들이 같거나 매우 유사하다면 그 16x16 매크로블록을 현재 부호화 하고자 하는 화소값들의 대표 블록으로 결정 할 수 있다. 이 경우 16x8, 8x16, 8x8, 8x4, 4x8, 4x4 블록에 대한 움직임 추정 및 보상을 수행하지 않고 16x16 매크로블록을 최적의 움직

임 결정 모드로 결정함으로써 많은 연산량을 줄일 수 있다. 마찬가지로 8x8 블록에 대해서도, 현재 블록과 참조 블록간에 매우 큰 유사성이 존재하면 하위 블록, 8x4, 4x8, 4x4에 대한 부호화 과정없이 현재 블록을 8x8블록으로 결정한다. 그 다음 문제는 어떻게 이 두 블록간의 화소가 매우 유사한지 아닌지를 판정하는가이다. 일반적으로 통계학에서 두 모집단의 데이터의 유사성을 검정하기 위한 방법으로 평균과 분산을 이용한 가설검증 기법을 사용하고 있다⁹⁾. 본 논문에서도 이러한 가설 검증 이용하여 현재 부호화 하고자 하는 매크로블록의 화소값들과 움직임 추정 및 보상을 통해 찾은 참조 화면에서의 블록 화소값의 평균과 분산을 이용하여 두 매크로블록의 유사성을 판정한다.

1. 두 평균의 가설에 대한 검증

통계적 가설 검증법을 SVC의 고속 화면간 모드 결정 방법에 사용하기 위하여, X_C 와 X_R 를 각각 현재 블록과 참조 블록의 독립적인 화소값들의 랜덤 변수로 정의하며 정규분포를 따르는 것으로 가정하고 다음과 같이 정의한다.

$$X_C \sim N(\mu_C, \sigma_C^2), X_R \sim N(\mu_R, \sigma_R^2) \quad (6)$$

여기서 μ_C 와 μ_R 은 각각 현재 블록과 참조 블록의 화소값들의 참 평균(true means)이라고 가정하고, 두 평균을 비교하여 두 블록의 화소값들의 유사성을 검증하기 위한 가설을 다음과 같이 정한다.

$$\begin{aligned} H_0 : \mu_C &= \mu_R \\ H_1 : \mu_C &\neq \mu_R \end{aligned} \quad (7)$$

여기서 H_0 를 귀무가설(null hypothesis), H_1 을 대립가설(alternative hypothesis)이라고 하며 그 것은 두 블록의 표본 화소 그룹이 같은 모집단으로부터 온 것인지를 판단한다. 만일 설정된 귀무가설이 받아들여지면, 현재 블록과 복원된 참조 블록 내의 두 개의 화소 그룹은 같은 모집단으로부터 온 것으로 판단하고 현재 부호화 하고자 하는 블록을

부호화 모드로 결정하게 된다. $\bar{X}_C, \bar{X}_R, S_C^2, S_R^2$ 을 각각 현재 블록과 참조 블록 내 화소의 표본 평균과 표본 분산이라고 하자. 이 때 알려지지 않은 모집단의 분산이 같다고 가정하고 식 (7)에 대하여 가설 검증을 실시한다. 두 평균에 대하여 귀무가설을 테스트 하기 위해 식 (8)과 같이 자유도 $\nu = n_C + n_R - 2$ 를 가진 검증통계량(test statistics) t 를 사용한다^[9].

$$t_0 = \frac{\bar{X}_C - \bar{X}_R}{S_p \sqrt{\frac{1}{n_C} + \frac{1}{n_R}}} \quad (8)$$

여기서 n_C 와 n_R 은 각각 가설 검증을 하고자 하는 현재 블록과 참조 블록의 표본의 수를 나타낸다. 따라서 통계적 가설 검증을 수행하기 위해 선택한 표본의 수는 16x16블록인 경우, 16x16(=256), 8x8블록인 경우 8x8(=64)이 된다. 이것은 평균에 대한 통계적 가설 검증을 수행하기에 충분히 큰 표본의 수라고 할 수 있다. 추출된 표본의 크기가 30보다 큰 대표본^[10]이기 때문에 평균 검증을 위해 t 분포를 이용하는데 아무런 문제가 없다. 한편 공분산 S_p^2 은 다음과 같이 정의한다.

$$S_p^2 = \frac{(n_C - 1)S_C^2 + (n_R - 1)S_R^2}{n_C + n_R - 2} \quad (9)$$

S_p^2 은 두 개의 표본 집단을 대표하는 분산으로 공분산이라고 한다. 일단 t 검증통계량이 계산되면 α 를 갖는 신뢰 구간(confidence interval) $-t_{\alpha/2, \nu} \leq t_0 \leq t_{\alpha/2, \nu}$ 을 설정하여, 신뢰 구간 내에 t 검증통계량 값이 존재하면, 귀무가설을 받아 들이고 그렇지 않을 경우는 받아 들이지 않는다. 여기서 귀무가설을 받아들인다는 의미는 평균 비교에 기초하여 두 블록의 화소값들이 신뢰구간에 대해 같다고 판정하는 것이다. 물론 평균의 비교만으로는 두 집단 (현재 블록과 참조 블록)의 화소들이 일치한다고 판정하기 어려울 뿐만 아니라 평균에 대한 검증을 위해 가정했던 분산에 대한 조

건을 검증해야 하므로 분산 비교를 추가적으로 수행함으로써 신뢰도를 더 향상시킬 수 있다.

2. 두 분산의 가설에 대한 검증

두 블록간의 평균에 대한 가설 검증과 마찬가지로 두 블록간의 분산에 대한 가설 검증을 실행한다. 식 (7)과 같이 분산 비교에 대한 가설을 다음과 같이 설정한다.

$$\begin{aligned} H_0: & \sigma_C^2 = \sigma_R^2 \\ H_1: & \sigma_C^2 \neq \sigma_R^2 \end{aligned} \quad (10)$$

여기서 σ_C^2 와 σ_R^2 은 각각 현재 블록과 참조 블록 화소값들의 참 분산(true variances)이다. 위 가설을 검증하기 위한 검증 통계량 z 를 사용한다^[11].

$$Z_0 = \frac{S_C - S_R}{S_p \sqrt{\frac{1}{2n_C} + \frac{1}{2n_R}}} \quad (11)$$

여기서 공통 표준 편차의 예측치인 S_p 는 (9)와 같이 정의한다. 분산 비교에 의한 두 그룹간의 화소 유사성을 비교하기 위해 신뢰 구간을 설정하고 계산된 z 검증 통계량이 신뢰 구간 내에 존재하면 귀무가설을 받아 들이고 그렇지 않으면 귀무가설을 기각한다. 여기서 귀무가설을 받아 들인다는 의미는 분산 비교에 기초하여 두 블록의 화소값들이 신뢰구간에 대해 같다고 판정하는 것이다. 그림 4는 귀무가설을 받아들일지를 판단하는 신뢰 구간의 설정을 나타낸다. 하나의 매크로블록에 대한 통계적 가설 검증 테스트에서 z 통계 분포가 신뢰 구간내에 있으면 귀무가설을 받아들여지게 된다. 일반적으로 두 모집단의 분산에 대한 유사도를 조사하기 위한 방법으로 F 통계 분포를 이용한다. 이때 F 분포는 비교하는 두 개의 통계 분포가 정규 분포를 따라야 한다는 제한이 있다. 그러나 두개의 샘플의 크기가 제안된 알고리즘에서와 같이 16x16(=256)또는 8x8(=64)과 같이 충분히 크다면 정규 분포를 가정하지 않고도 충분히 분산에 대

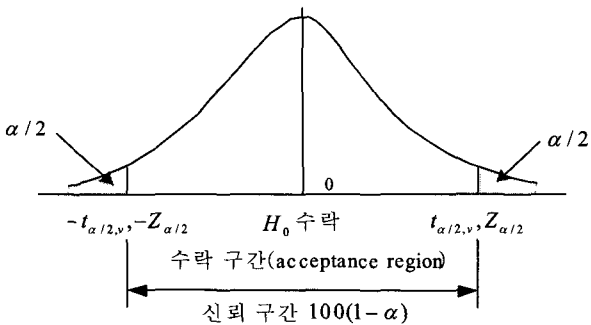


그림 4. 통계 분포에서의 신뢰 구간의 설정
 Fig. 4. Confidence interval in the test statistics

한 유사도 검증을 할 수가 있다. 이때 표본의 표준편차 S_R 과 S_C 가 평균 α_C 과 α_R 그리고 분산 $\sigma_C^2/2n_C$, $\sigma_R^2/2n_R$ 를 갖는 근사화된 정규 분포를 따른다는 사실에 근거한다^[9]. 따라서 본 논문에서는 분산을 테스트 하기 위한 통계 분포로 z 통계 분포를 이용하였다.

그림 4는 t 또는 z 통계 분포의 예를 보여준다. 일단 신뢰 구간을 설정하여 샘플의 평균과 분산을 이용한 검증 통계량의 통계 분포가 신뢰 구간 내로 들어오면 두 모집단의 평균이 같다고 판단하는 것이다. 이것은 현재 매크로블록과 참조 매크로블록의 영역이 충분히 정적이거나 균일한 영상 영역일 때 가능하므로 평균에 대한 통계적 가설 검증이 통과되면 이것은 큰 블록으로 모드가 결정될 가능성이 높다는 것을 의미한다.

IV. 제안하는 통계적 가설 검증을 이용한 고속 블록 모드 결정 과정

먼저 부호화 하고자 하는 16x16 매크로블록에서 평균과 분산에 대한 가설 검증을 실시하고 귀무가설이 모두 받아들여지면 현재 블록과 참조 블록은 같은 모집단으로부터 온 일치된 블록이라고 판단한다. 이 경우는 현재 블록과 참조 블록의 평균과 분산이 신뢰구간 내에서 일치하는 값을 가지므로 하위 블록에 대한 RDO과정은 불필요하게 된다. 따라서 하위 블록(16x8, 8x16, 8x8, 8x4, 4x8, 4x4)에 대한 움직임 벡터 추정과 모드 결정 없이 현재 블록을 16x16 모

드로 결정하여 부호화를 수행한다. 만약 두 가설 검증에서 하나라도 귀무가설이 받아들여지지 않으면 16x8, 8x16, 8x8 블록에 대해 일반적인 움직임을 예측을 수행한다. 가설 검증이 받아들여지지 않는 경우는 현재 블록과 참조 블록의 평균과 분산이 신뢰구간 내에서 일치하지 않아 두 블록이 같은 모집단으로부터 형성된 것이 아님을 의미한다. 따라서 이 경우는 하위 블록으로 세분하게 나뉘어진 영역에서 움직임 예측 및 보상을 할 필요가 있다. 8x8 블록에 대한 RDO 수행 시 8x8 블록에 대해 다시 16x16 매크로 블록에서와 같은 가설 검증을 수행한다. 마찬가지로 평균과 분산에 대한 가설 검증에서 귀무가설이 모두 받아들여지면 현재 블록과 참조 블록은 같은 모집단으로부터 온 일치된 블록이라고 판단하기 때문에 하위 블록(8x4, 4x8, 4x4)에 대한 움직임 벡터 추정과 모드 결정 없이 현재 블록을 8x8 모드로 결정하게 된다. 이와 같이 연산량 감소를 위한 고속 화면간 블록 모드 결정 방법은 3장에서 기술한 통계적 가설 검증에 기반한다. 제안하는 고속 화면간 블록 모드 결정의 구체적인 방법은 다음과 같은 과정을 따른다.

- 과정 1: 현재 슬라이스가 단방향(P) 예측이면 먼저 스킵(SKIP)모드 실행, 양방향(B) 예측이면 직접(DIRECT)모드를 실행한다.
 - 과정 2: 16x16 매크로 블록에 대하여 움직임 벡터를 추정한다.
 - 과정 3: 움직임 보상된 참조 매크로 블록과 현재 블록에 대하여 통계적 가설 검증을 실시한다.
 - 과정 4: 평균과 분산에 대한 가설 검증을 모두 통과하면 현재의 블록은 직접/스킵 모드와 16x16 블록의 RD 비용을 비교하여 최적의 모드를 결정하고 모드 결정과정을 완료한다. 하나라도 통과하지 않으면 과정 5로 진행한다.
 - 과정 5: RD 최적화 (RDO)에 기반하여 16x8, 8x16, 8x8 블록에 RDO를 수행한다.
- 8x8 블록인 경우
- 과정 5-1: 각 8x8 블록에 대하여 움직임 벡터 추정 및 RD 비용을 계산한다.
 - 과정 5-2: 현재 8x8블록과 움직임 보상된 참조 블록

- 에 대하여 통계적 가설검증을 실시한다
- 과정 5-3: 각각의 8x8 블록에 대하여 가설 검증이 통과 되면 현재 블록은 8x8 블록으로 결정되고 8x8 하위 블록인 8x4, 4x8, 4x4 블록에 대한 모드 결정 과정을 생략한다. 그렇지 않으면 과정 8로 간다
- 과정 5-4 : 8x4, 4x8, 4x4에 대한 RDO를 수행한다.
- 과정 5-5: 다음 8x8 블록으로 간다.
- 과정 8: RDO 기반 가장 최적의 모드를 결정한다.
- 과정 9: 다음 매크로 블록에 대한 모드 결정을 진행한다.

그림 5에서 보는 바와 같이 현재 부호화 되고 있는 블록이 16x16 크기의 1번 블록 일 때 최적의 참조 블록이 전방향에서 2번 블록이면 1번과 2번 블록에 통계적 가설 검증을 수행 하게 된다. 이 때 두 블록간의 평균과 분산에 대한 귀무가설이 모두 받아들여지면 현재의 매크로 블록은 16x16으로 확정된다. 만일 두개의 가설 중 하나라도 받아들여지지 않으면 하위 블록에 대한 움직임 추정 및 보상을 통하여 모드 결정과정을 거치게 된다. 마찬가지로 8x8 하위 매크로 블록에 대하여 동일한 방법으로 통계적 가설 검증을 수행 하게 되며 평균과 분산에 대한 귀무가설이 모두 받아들여지면 현재 블록은 8x8블록으로 확정되고 그렇지 않으면 8x4, 4x8, 4x4 하위 블록에 대해 움직임 추정 및 보상을 통하여 모드 결정 과정을 거치게 된다. 그림 6은 통계적 가설

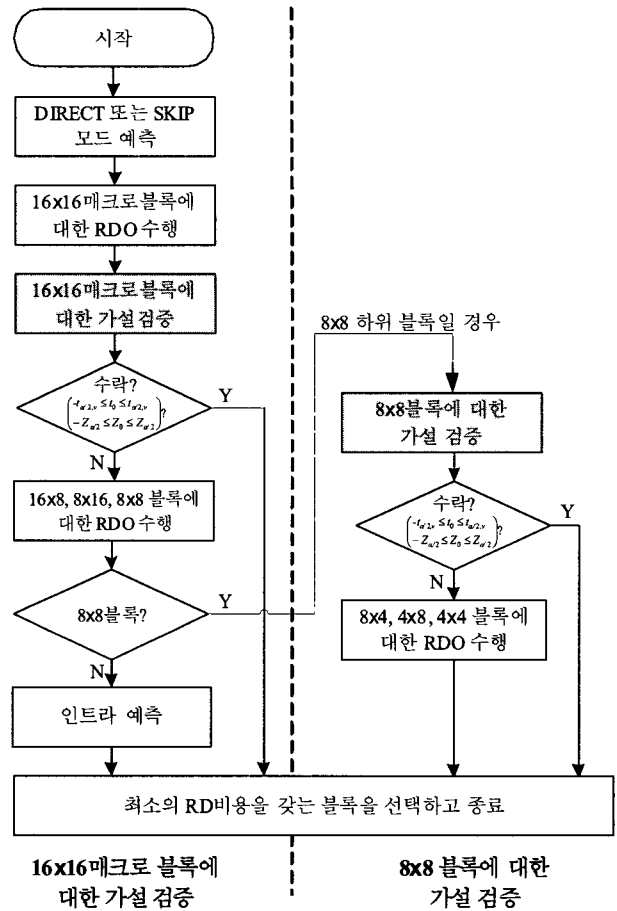


그림 6. 제안된 고속 화면간 모드 결정 알고리즘의 순서도
Fig 6. Block diagram for proposed fast intermode decision

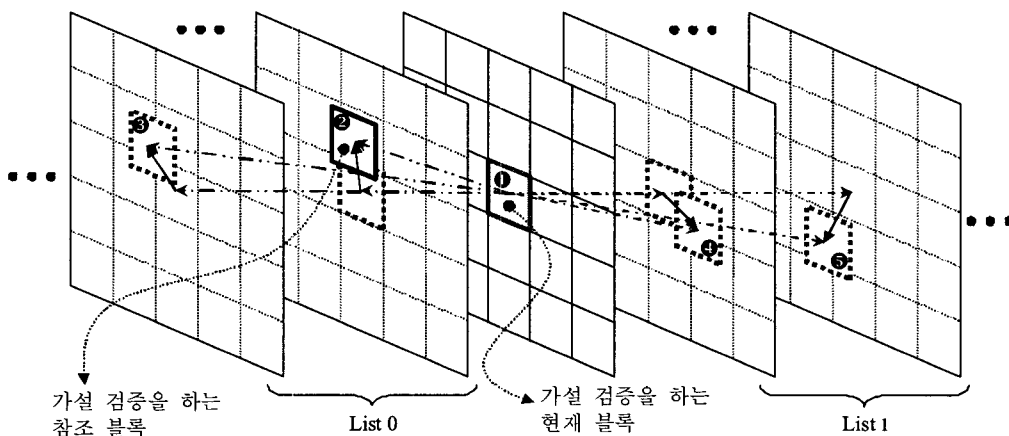


그림 5. 현재 블록과 참조 블록에 대한 가설 검증 테스트
Fig. 5. Statistical hypothesis testing for current and reference blocks

검증을 이용한 고속 화면간 모드 결정 방법에 대한 블록 다이어그램이다. 제안 알고리즘은 16x16 매크로 블록과 8x8 블록에서 가설 검증이 적용되며, 평균과 분산에 대한 귀무가설이 모두 받아들여 질 때는 하위 블록에 대한 움직임 추정 및 보상과정을 수행하지 않음으로써 그 만큼의 부호화 시간을 감소시킬 수 있다.

그림 7~10은 다양한 특성을 가지는 비디오 영상에 대한 t 검증통계량과 z 검증통계량의 분포를 나타낸다. 본 검증통계량의 분포를 얻기 위해 양자화 계수(QP), 28에 대하여 실험을 시행하였고 움직임이 느린 *Mother & Daughter* 영상 그리고 움직임이 빠른 *Foreman* 영상의 100 화면에 대해 각각 실험을 수행하였다.

그림 7~8에서 보는 것처럼 움직임이 빠른 영상일수록 t 검증통계량의 분포가 넓게 퍼져 있는 것을 확인 할 수 있다. 이것은 식(9)에서 보는 것처럼 느린 영상일수록 현재 영상

과 참조 영상간의 평균값의 유사도가 높기 때문이며, 빠른 영상일수록 느린 영상에 비해 원영상과 참조 영상간의 화소값의 평균의 유사도가 낮음을 알 수 있다.

그림 9과 10은 양자화 계수값이 28일 때 각 테스트 영상에 대한 z 검증통계량 분포를 나타낸다. 이것 역시 마찬가지로 움직임이 빠르거나 복잡한 영상일수록 넓게 퍼져 있고 느린 영상일수록 샘플들의 값이 중앙에 집중하는 것을 알 수 있다. 동일한 신뢰 구간을 적용할 때 움직임이 느린 *Mother & Daughter* 영상이 통계적 가설 검증을 통과할 확률이 높다는 것을 알 수 있다.

V. 실험 결과

실험을 위해 JSVM 6.0^[11] 참조 소프트웨어에서 설정한

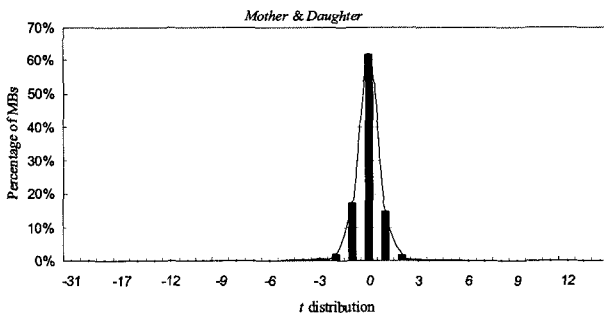


그림 7. t 통계 분포 (Mother & Daughter, QP 28)
Fig. 7. t distribution (Mother & Daughter, QP 28)

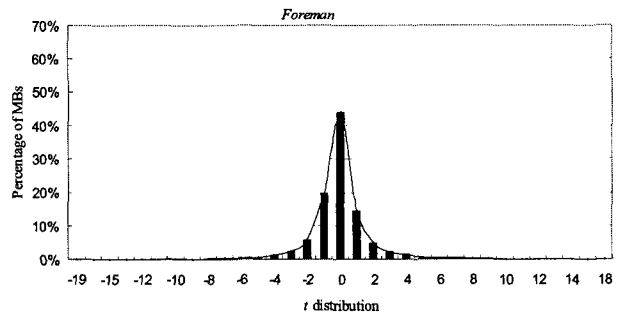


그림 8. t 통계분포 (Foreman, QP 28)
Fig. 8. t distribution (Foreman, QP 28)

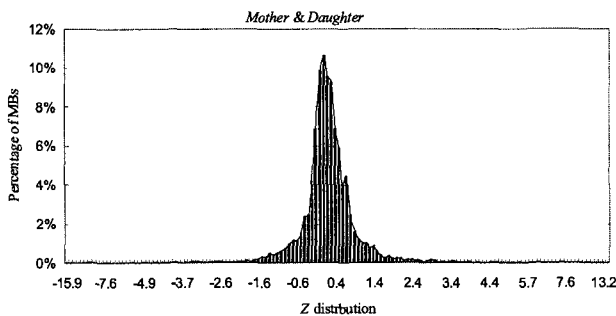


그림 9. z 통계 분포 (Mother & Daughter, QP 28)
Fig 9. z distribution (Mother & Daughter, QP 28)

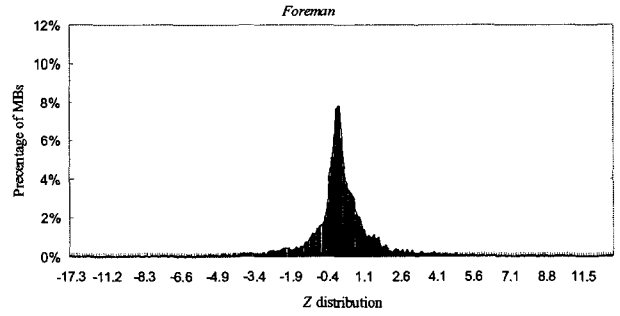


그림 10. z 통계 분포 (Foreman, QP 28)
Fig. 10. z distribution (Foreman, QP 28)

실험환경은 표 2에 나타낸다.

표 2. 부호화
Table 2. Encoding parameter

참조화면의 수	2
GOP 크기	4(1BBBP), 8(1BBBBBBP)
움직임벡터정밀도	1/4 화소
움직임벡터탐색범위	32화소
비용함수	정수화소 탐색시 : SAD 부화소 탐색시 : SATD
공간 및 화질 계위	사용안함
신뢰구간 100(1- α)	Option 1 : - 모든 시간 계위에 대하여 20% Option 2 : - $10i + 10, 0 \leq i \leq \log N_{GOP} / \log 2$, (N_{GOP} : GOP의 크기, i : 시간 계위)
컴퓨터 환경	Pentium IV- 3.2 GHz cpu , 1GB 주메모리

제안된 알고리즘의 목적을 부호화 과정에서 연산량을 감소시켜 전체 부호화 시간을 줄이는 것을 목표로 한다. 따라서 부호화 시간을 측정하여 원 참조 소프트웨어 (JSVM)보다 얼마만큼의 부호화 시간의 감소를 가져왔는지 측정하였다. 그리고 제안된 알고리즘이 적용되었을 때 화질 열화를 측정하기 위하여 각 화면 별로 PSNR 평균을 계산하여 제안 알고리즘 적용 전과 후의 PSNR 변동량을 조사하였다. 마지막으로 제안된 적용 전 후의 비트율 변화량을 측정하였다. 시험 영상으로 움직임이 적은 순서대로 *Mother&Daughter*, *Silent*, *Carphone*, *Foreman*, *Coast Guard QCIF(30Hz)* 영상을 사용하였고, 움직임은 적으나 텍스처가 미세한 특징을 보여주는 *Harbour* 영상을 사용하였다. 그리고 CIF 영상으로서 움직임이 다소 적은 *News CIF*와 움직임이 많은 *Mobile & Calendar* 영상을 사용하였다. 표 3은 제안하는 알고리즘을 적용해서 부호화 했을 때 PSNR 변화량과 비트율 증가량을 SVC 참조 소프트웨어인 JSVM 6.0과 비교하여 표시하였다.

1. 고정된 신뢰 구간을 이용한 고속 화면간 모드 결정

통계적 가설 검증 수행시 신뢰 구간을 다르게 적용하였

을 때 RD 성능과 부호화 시간의 변화 대하여 알아볼 필요가 있다. 왜냐하면 신뢰구간을 크게 하면 귀무가설이 받아들여질 확률이 높기 때문에 고속 화면간 모드 결정을 달성할 수가 있지만 그에 따른 RD 성능의 저하도 예상되기 때문이다. 먼저 실험 영상, *Foreman*에 대하여 신뢰 구간 20%, 40% 그리고 60%을 각각 모든 시간 계위에 동일하게 적용하여 실험을 수행하였다. 그림 11은 서로 다른 신뢰구간에 따른 RD 성능에 대한 그래프이다. 신뢰구간을 20%로 했을 경우 비트율 증가분이 JSVM6.0에 비하여 최대 1.74% 증가하는데 그쳤지만 신뢰 구간을 60%로 증가 시키면 7~10%가량 비트율이 증가하였다. 이것은 신뢰 구간을 증가시키면 현재 블록과 참조 블록이 일치한다고 판단하는 비율이 커지는 것을 의미한다. 즉 큰 블록으로 모드를 결정하는 비율이 커지게 되므로, 세밀한 영상 영역이나 움직임이 큰 부분에서는 효율적이 못해 부호화 효율이 떨어지기 때문이다. 그러나 PSNR 값은 신뢰 구간의 설정에 따라 영향을 덜 받게 된다. 그것은 PSNR 값은 양자화 계수에 의해 절대적으로 영향을 받기 때문에 신뢰구간의 설정의 영향이 덜하기 때문이다.

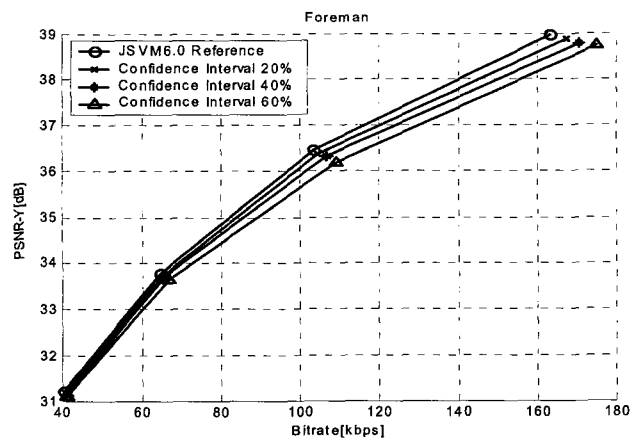


그림 11. 신뢰구간(Confidence Interval)에 따른 RD 성능, Foreman, QCIF, 30fps
Fig. 11. RD performance for different confidence intervals (Foreman, CIF, 30fps)

그림 12는 신뢰 구간을 다양하게 적용했을 경우 부호화 시간 감소 비율을 나타낸다. 테스트 영상은 *Foreman*을 사

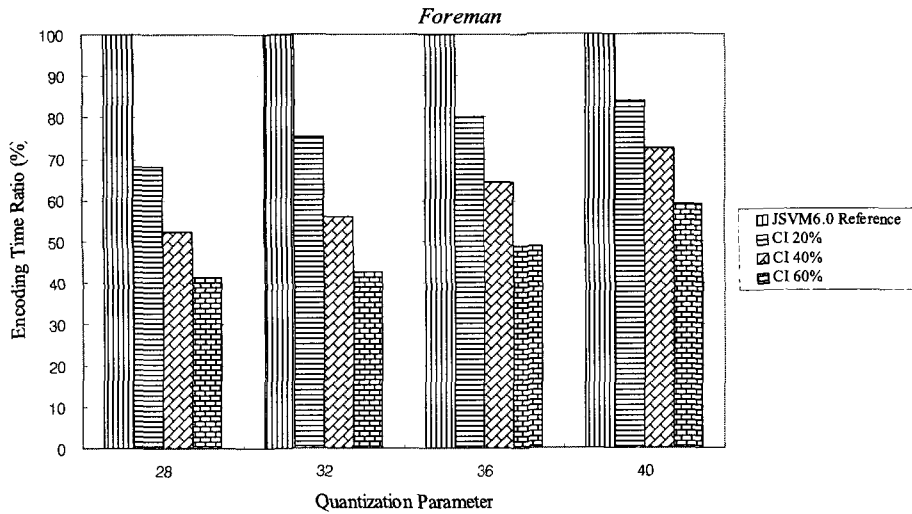


그림 12. 신뢰 구간에 따른 부호화 시간의 비교
 Fig. 12. The comparison of encoding time for different confidence intervals

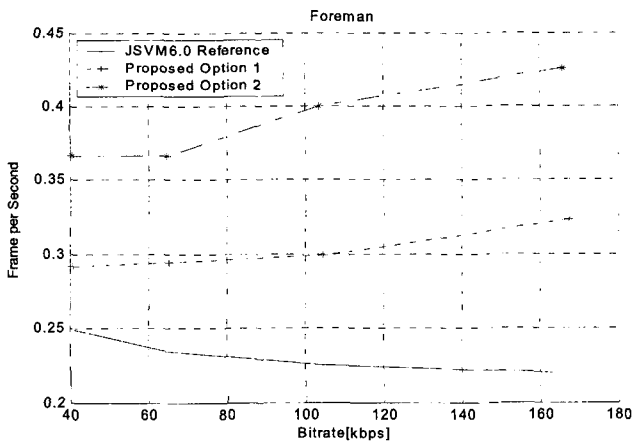


그림 13. Foreman의 부호화 시간 (GOP 크기=8)
 Fig 13. The encoding time for Foreman(GOP size =8)

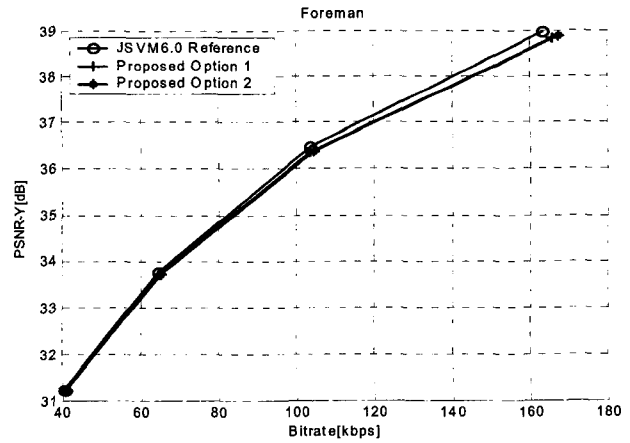


그림 14. Foreman의 RD 성능 그래프 (GOP 크기=8)
 Fig 14. RD-curves for Foreman (GOP size=8)

용하였고 양자화 계수는 28, 32, 36, 40을 사용하였다. 그림 13과 그림 14는 RD 성능과 부호화 감소 사이의 trade-off 관계를 보여준다. 즉 신뢰 구간을 크게 하면 그림 14에서 보는 것처럼 양자화 계수 28을 적용하였을 경우 최대 60% 정도의 부호화 시간 감소 비율을 보여주지만 비트율 증가도 7% 이상 발생하는 것을 알 수 있다. 따라서 신뢰 구간은 RD 성능과 복잡도 감소 비율을 적절히 고려하여 설정해야 한다.

2. 시간계위에 따라 다른 신뢰 구간을 적용한 고속 화면간 모드 결정

두 번째 방법은 표 1에서 보는 바와 같이 시간 계위에 따라 블록 모드의 분포가 매우 다르게 나타나므로 그에 따라 시간 계위별로 신뢰 구간을 다르게 적용하므로써 보다 효율적인 고속 화면간 모드 결정을 할 수가 있다. 예를 들어, GOP내에서 가장 중요도가 높은 주화면이나 시간 계위

표 3. 평균 부호화 시간의 단축

Table 3. Average encoding time saving

QP	Option 1						Option 2					
	GOP 크기= 4			GOP 크기= 8			GOP 크기= 4			GOP 크기= 8		
	28	32	36	28	32	36	28	32	36	28	32	36
Mthr &Dter	42.1	32.8	28.3	41.3	36.8	28.4	49.5	38.6	32.2	56.3	51.7	39.5
Silent	40.4	28.3	21.3	43.9	33.1	23.5	47.9	35.8	26.9	61.4	52.8	35.6
Carphone	36.3	30.1	22.0	38.6	30.0	25.2	44.5	36.5	27.0	54.5	48.8	42.1
Coast Guard	42.0	32.8	21.2	36.8	32.5	24.7	44.5	36.5	27.0	54.5	45.4	38.3
Foreman	34.7	27.4	23.0	31.8	24.8	20.1	41.7	34.7	29.1	48.2	43.9	35.6
Harbour	56.0	42.2	29.8	50.1	40.7	30.1	63.5	52.5	40.0	68.9	60.7	46.9
News	41.4	32.9	25.7	38.7	32.2	22.6	41.3	36.2	28.3	54.1	47.4	37.4
Mobile	37.0	31.8	20.2	27.7	26.0	21.1	40.0	32.0	20.6	50.5	44.1	37.2
Avg	41.2	32.3	23.9	38.6	32.0	24.5	46.6	37.9	28.9	56.1	49.4	39.1

unit : ((JSVM6.0sec - Proposedsec)/JSVM6.0sec)x100%

가 낮은 화면에 보다 작은 신뢰 구간을 적용하고 시간 계위가 높은 화면에 대해서는 보다 큰 신뢰 구간을 적용하여 보다 고속의 부호화를 수행할 수 있도록 한다. 이것은 표 1에서 보는 것처럼, 시간 계위가 높을 수록 큰 블록으로 결정되는 확률이 크기 때문에 보다 큰 신뢰 구간을 적용하여도 RD 성능의 저하에는 거의 영향이 없음을 예상할 수 있다. 즉 시간 계위가 높은 화면들은 신뢰 구간을 작게함으로써 그 구간내에서 현재 블록과 참조 블록이 일치한다고 판단하는 비율을 적게하여 화면의 화질을 일정 수준 유지할 수 있게 하고, 시간 계위가 낮은 화면들에 대해서는 표 1에서처럼 큰 블록들의 블록 모드 결정 비율이 크므로 신뢰 구간을 크게 하여 가설 검증을 통과하는 비율을 높인다.

표 3에서는 신뢰 구간을 모든 시간 계위에 대하여 20%로 설정하여 가설 검증을 실시한 Option1과 시간 계위에 따라 다르게 적용한 Option 2를 이용했을 때의 부호화 시간 단축 비를 나타낸 것이다. 표에서 보는 것처럼 같은 조건에서 Option 2가 Option 1에 비해 최대 5~18% 가량 부호화 시간 감소를 더 얻을 수 있음을 확인 할 수 있다. Option 2에서는 시간 계위가 높은 화면에 대해서 통계적 가설 검증 수행시 보다 큰 신뢰 구간을 설정하였기 때문에 시간 계위가 높은 화면은 가설 검증을 받아 들이는 비율이 고정된 신뢰구간을 적용하는 Option 1보다 크게 되므로 부호화 시간의 감소효

과를 더욱 크게 얻을 수 있다. 표 4는 제안된 알고리즘의 PSNR과 비트율 변화량을 측정된 것이고 [12]에서 기술한

표 4. BDBR과 BDPSNR

Table 4. BDBR and BDPSNR

	GOP Size	BDBR (%)		BDPSNR (dB)	
		Option1	Option2	Option1	Option2
Mother & Daughter	4	0.08	0.77	-0.01	-0.04
	8	1.05	1.21	-0.05	-0.06
Silent	4	0.75	0.96	-0.04	-0.05
	8	0.93	1.14	-0.05	-0.06
Carphone	4	1.31	0.64	-0.06	-0.03
	8	1.55	1.61	-0.08	-0.08
Coast Guard	4	1.78	0.87	-0.06	-0.03
	8	1.69	2.02	-0.06	-0.07
Foreman	4	1.38	0.64	-0.07	-0.03
	8	2.30	1.63	-0.12	-0.08
Harbour	4	1.14	1.18	-0.05	-0.05
	8	1.55	1.70	-0.07	-0.07
News (CIF)	4	1.28	0.88	-0.07	-0.04
	8	1.38	1.28	-0.08	0.07
Mobile (CIF)	4	1.73	1.44	-0.07	-0.06
	8	1.67	1.98	-0.07	-0.08
Avg.	4	1.18	0.92	-0.05	-0.04
	8	1.52	1.57	-0.07	-0.05

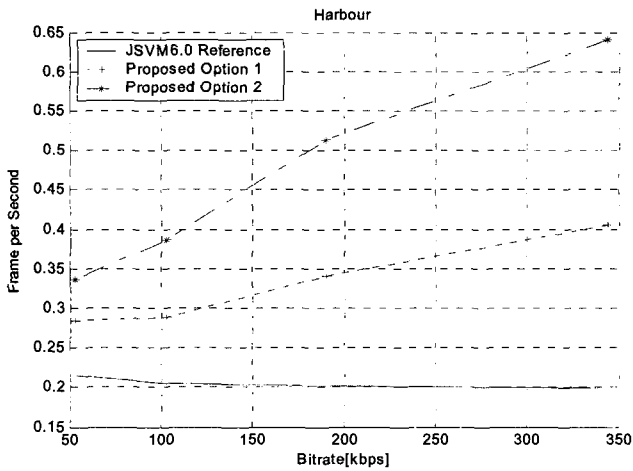


그림 15. Harbour의 부호화 시간 (GOP 크기 =8)
Fig. 15. The encoding time for Harbour(GOP size =8)

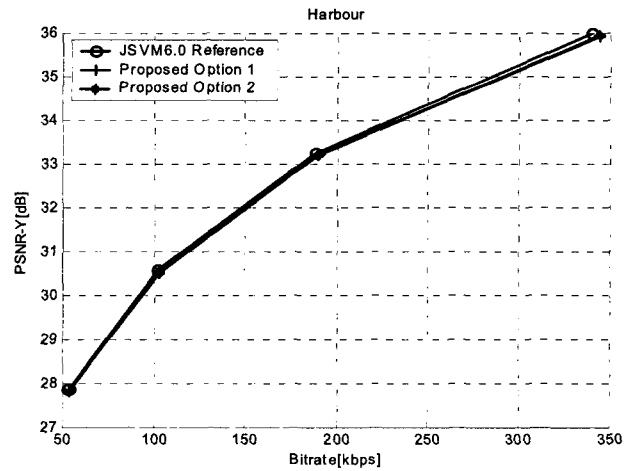


그림 16. Harbour의 RD 성능 그래프 (GOP 크기 =8)
Fig. 16. RD-curves for Harbour(GOP size=8)

BDPSRN과 BSBR을 이용하였다. 표에서 보는 것처럼 평균 PSNR의 감소는 최대 0.07dB이고 평균 비트율 증가는 최대 1.5%가 된다. 이것은 실험 영상에 대한 평균값으로 전체 RD 성능에 비하여 무시할 수 있는 정도로 작은 수준이다. 한편, Option1과 Option 2에 대한 RD 성능의 변화는 거의 비슷한 수준이며 시간 계위에 따라 서로 다른 신뢰 구간을 적용하여도 RD 성능에는 거의 영향이 없음을 확인할 수 있다. 표 1에서 확인하였듯 시간 계위가 높은 화면일수록 16x16 블록과 같은 큰 블록으로 화면간 모드 결정이 될 확률이 크므로 시간 계위가 높은 화면에 대하여 더 높은 신뢰 구간을 적용하여도 RD 성능에는 별로 영향을 주지 않기 때문이다.

시험영상 *Foreman*과 *Harbour* 영상에 대한 RD 성능 그래프를 그림 14와 16에 나타낸다. 두가지 옵션은 SVC 참조 소프트웨어를 이용한 RD 성능과 PSNR과 비트율 측면에서 거의 유사한 곡선을 보여주고 있다. 따라서 제안한 두 가지 방법 모두 참조 소프트웨어와 거의 동일한 RD 성능을 얻을 수 있음을 확인할 수 있다. 그러나 그림 13과 15의 부호화 시간에 대한 그래프에서 보는 것처럼 시간 계위에 대한 화면간 모드 결정의 통계적 특성을 고려하여 제안한 알고리즘을 적용하면 같은 RD 성능을 유지하면서도 고속으로 부호화 할 수 있음을 알 수 있다.

VI. 결론

SVC 부호화기의 연산량 감소를 위해 통계적 가설 검증을 이용한 화면 간 고속 모드 결정 알고리즘을 제안하였다. 현재 블록과 참조 블록간의 평균과 분산에 대한 통계적 가설을 설정하여 그 가설의 기각 여부를 검증하는 알고리즘이다. 현재 블록과 참조 블록간의 통계적 가설 검증이 받아들여지면 검증 영역은 균일한 영역이므로 하위 블록에 대한 부호화 과정없이 블록 모드 결정을 조기에 완료하게 된다. 기존의 방법들은 많은 실험을 통하여 최적의 임계치를 찾아내며 각기 다른 특성을 가지는 영상에 대하여 다른 임계치를 사용하지만 제안된 알고리즘은 통계적인 분포에 신뢰 구간을 적용하여 신뢰 구간만큼의 연산량 감소를 얻기 때문에 신뢰도 높은 고속 모드 결정을 할 수 있다. 또한 GOP내에서 시간 계위에 따른 서로 다른 블록 모드 분포를 이용하여 그에 따른 신뢰구간을 다르게 적용함으로써 제안 알고리즘은 최대 69%까지 부호화 시간 감소 효과를 얻을 수 있었으며, PSNR은 최대 0.07dB 감소, 비트율은 최대 1.57% 증가한데 그쳤다. 이는 RD 성능측면에서 무시할 수 있을 정도로 작은 수치이다. 따라서 RD 성능의 열화를 최소화 하면서 연산량을 크게 줄였다고 할 수 있다.

참고 문헌

- [1] J.Reichel, H.Schwarz and M.Wien, ISO/IEC JTC 1/SC 29/WG 11: 'Text of ISO/IEC 14496-10:2006/PDA M3 Scalable Video Coding,' MPEG76/N8015, Montreux, April 2006.
- [2] 박성호, 김원하, 한우진, "H.264/AVC-Scalable Extension의 표준화 연구동향과 알고리즘 분석", 방송공학회 논문지, 제 10권, 4호, 2005년 12월
- [3] L.Yang, K.Yu, J.Li and S.Li, " An Effective Variable Block-Size Early Termination Algorithm for H.264 Video Coding', IEEE Transaction on Circuits and Systems for Video Technology, vol 15, no 6, June 2005.
- [4] D.Wu, F.Pan, K.P.Lim, S.Wu, Z.G.Ki, X.Lin, S.Rahardja and C.C.Ko, "Fast Intermode Decision in H.264/AVC Video Coding," IEEE Transaction on Circuits and Systems for Video Technology, vol 15, no 6, July 2005.
- [5] X.Jing and L.P.Chau, "Fast approach for H.264 inter mode decision," ElectronicsLetters, vol.40, no.7, August 2004.
- [6] Y.H. Kim, J.W.Yoo, S.W. Lee, J.Shin, J.Paik and H.K. Jung, "Adaptive mode decision for H.264 encoder," Electronics Letters, vol.40, no.19, September 2004.
- [7] I.Richardson, "H.264 and MPEG-4 Video Compression - Video Coding for Next-generation Multimedia," Wiley, 2003.
- [8] H.Schwarz, D.Marpe and T.Wiegand, "Analysis of Hierarchical B Pictures and MCTF," IEEE International Conference on Multimedia & Expo (ICME 2006), Canada, July 9-12, 2006
- [9] W.Hines, D.Montgomery, D.Goldman and C.Borror, "Probability and Statistics in Engineering," Wiley, 4th Edition, 2003.
- [10] 김학범, "통계방법과 자료분석", 도서출판 대명, 2004년
- [11] ISO/IEC JTC 1/SC 29/WG 11/N7797: 'JSVM 6 Software,' MPEG76/N8016, Montreux, April 2006
- [12] G. Bjontegaard, "Calculation of Average PSNR Differences between RD-curves," Doc. VCEG-M33, April. 2001.

저자 소개



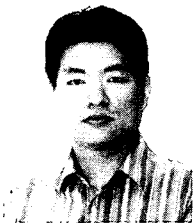
이 범 식

- 2000년 8월 : 고려대학교 전기전자전파공학부 학사
- 2006년 8월 : 한국정보통신대학교 공학부 석사
- 2006년 8월 ~ 현재 : 한국정보통신대학교 공학부 박사 과정
- 2000년 9월 ~ 2003년 9월 : (주)기산텔레콤
- 주관심분야 : Scalable Video Coding, H.264/AVC, 내용기반 영상 검색



김 문 철

- 1989년 : 경북대학교 학사
- 1992년 : University of Florida, Electrical and Computer Engineering, 석사
- 1996년 : University of Florida, Electrical and Computer Engineering, 박사
- 1997년 1월~2001년 2월 : 한국전자통신연구원 선임연구원, 방송미디어연구부 실감영상연구팀/영상미디어 연구팀 팀장
- 2001년 2월 ~ 현재 : 한국정보통신대학교 부교수
- 주관심분야 : MPEG-4/7/21/A/E, 멀티미디어 정보처리, 대화형 멀티미디어 방송, 지능형 에이전트 기반 멀티미디어 응용, 영상 및 신호처리



함 상 진

- 1996년 2월 : 연세대학교 전자공학과 학사
- 1998년 8월 : 연세대학교 전자공학과 석사
- 2001년 7월~현재 : 한국방송 방송기술연구팀, 현 U콘텐츠처리보호 선임연구원
- 주관심분야 : 고효율 영상압축, 콘텐츠 저작권 보호 기술, 영상 신호 처리

저 자 소 개



김 병 선

- 1994년 2월 : 경북대학교 전자공학과 학사
- 1996년 2월 : 경북대학교 전자공학과 석사
- 1996년 1월 ~ 현재 : 한국방송 방송기술연구팀 현 U 콘텐츠처리보호 선임연구원
- 주관심분야 : H.264/AVC, Scalable Video Coding, 방송콘텐츠 저작권보호 기술, 영상 및 신호처리



이 근 식

- 1985년 : 서울대학교 제어계측공학 학사
- 1987년 : 서울대학교 제어계측공학 석사
- 1988년 - 현재 : 한국방송 방송기술연구팀, 현 U콘텐츠처리보호 책임연구원
- 주관심분야 : 고효율 영상압축, 영상 신호처리, 디지털 방송제작 장비, 방송 콘텐츠 저작권보호



박 근 수

- 1982년 : 서울대학교 제어계측공학과, 학사
- 1984년 : KAIST 전기전자공학과, 석사
- 1984년 3월 ~ 현재 : 한국방송 방송기술연구소 근무, 현 U콘텐츠처리보호 수석연구원
- 주관심분야 : HD/SD 영상효과 및 처리 방송제작기술, DTV 방송콘텐츠 보호기술, 유비쿼터스 콘텐츠 처리기술