

스케일러블 비디오 부호화에서 통계적 가설 검증 기법을 이용한 프레임 간 모드 결정

*이범식 *김문철 **함상진 **이근식 **박근수

*한국정보통신대학교(ICU) **한국방송공사(KBS)

{bslee, mkim}@icu.ac.kr, {cashy, kslee22, keunspark}@kbs.co.kr

Fast Intermode Decision of Scalable Video Coding using Statistical Hypothesis Testing

*Lee, Bumshik *Kim, Munchurl **Hahm, Sangjin **Lee, Keunsik **Park, Keunsoo

*Information and Communications University **Korea Broadcasting System

요약

스케일러블 비디오 코딩(SVC, Scalable Video Coding)은 MPEG(Moving Picture Expert Group)과 VCEG (Video Coding Expert Group)의 JVT(Joint Video Team)에 의해 현재 표준화 되고 있는 새로운 압축 표준 기술이며 시간, 공간 및 화질의 스케일러블리티를 지원하기 위해 계층 구조를 가지고 있다. 특히 시간적 스케일러블리티를 위해 계층적 B-픽처 구조를 채택하고 있다. 스케일러블 비디오 코딩의 기본 계층은 H.264/AVC와 호환적이므로, 모션 예측과 모드 결정과정에서 16x16, 16x8, 8x16, 8x8, 8x4, 4x8 그리고 4x4와 같은 7개의 서로 다른 크기를 갖는 블록을 사용한다. 스케일러블 비디오 코딩에서 사용되고 있는 계층적 B-픽처 구조는 키 픽처인 I와 P 픽처를 제외하고는 한 GOP (Group of Picture)내에서 모두 B-픽처를 사용하므로 H.264/AVC와 비교했을 때 연산량 증가와 함께 부호화 지연도 급격히 증가한다. B-픽처는 양방향 모션 벡터인 LIST0와 LIST1을 사용하고 양방향 모두에서 다중 참조 픽처를 사용하기 때문이다.

본 논문에서는 통계적 가설 검증을 이용하여 스케일러블 비디오 부호화에 적용 가능한 고속 프레임간 모드 결정 알고리즘 대해 소개한다. 제안된 방법은 16x16 매크로 블록과 8x8 서브 매크로 블록에 통계적 가설 검증 기법을 적용하여 실행되며, 현재 블록과 복원된 참조 블록간의 픽셀 값을 비교하여 RD(Rate Distortion) 최적화 기반 모드 결정을 빨리 완료함으로써 고속 프레임간 모드 결정을 가능하게 한다. 제안된 방법은 프레임 간 모드 결정을 고속화함으로써 스케일러블 비디오 부호화 기의 연산량과 복잡도를 최대 57% 감소시킨다. 그러나 연산량 감소에 따른 비트율의 증가나 화질의 열화는 최대 1.74% 비트율 증가 및 0.08dB PSNR 감소로 무시할 정도로 작다.

1. 서론

스케일러블 비디오 부호화 (SVC, Scalable Video Coding) [1]은 H.264/AVC의 확장 버전으로 시간, 공간 및 화질적 스케일러블리티를 지원한다. SVC는 MPEG (Moving Picture Expert Group)과 VCEG (Video Coding Expert Group)에 의해 구성된 JVT (Joint Video Team)에 의해 현재 표준화가 진행 중이다. SVC는 시간적 스케일러블리티를 지원하기 위하여 H.264/AVC의 B-picture의 유사 개념인 계층적 B-픽처를 채택하였다. 그림 1에서는 SVC에서 시간적 스케일러블리티를 지원하기 위해 채택된 계층적 B-픽처의 개념도를 보여주고 있다. 따라서 모션 벡터 추정 및 모드 결정시에 16x16, 16x8, 8x16, 8x8, 8x4, 4x8, 4x4와 같은 총 7개의 서로 다른 모드를 사용한다. SVC의 계층적 B-픽처는 키 픽처인 I/P 픽처를 제외하고 모두 B-픽처를 사용한다. 따라서 B-픽처 부호화시 사용되는 양방향 모션 벡터 추정, 양방향 다중 참조 픽처 등을 사용하기 때문에 연산량과 복잡도는 H.264/AVC에 비하면 상당량 증가하게 된다.

H.264/AVC의 부호화 부분에서 연산량 감소를 위해 많은 연구들이 진행되어 왔다. Libo Yang은 무 움직임 검출(zero motion detection)에 기반하여 조기에 움직임 벡터와 모드 결정을 완료하는 방법을 제안하였다 [2]. Wu는 모드 결정을 조기에 완료하기 위한 방법으로 homogeneous 텍스처와 움직임이 적은 부분을 검출하는 텍스처 분석 방법을 사용하였다 [3].

본 논문에서는, 현재 부호화되는 블록의 픽셀값과 움직임 보상된 참조 블록의 픽셀값을 통계적 가설 검증법으로 비교하여 움직임 벡터의 추정 및 모드 결정과정을 조기에 완료함으로써 고속 프레임 간 모드 결정을 달성하는 알고리즘에 대하여 소개한다. 본 논문은 다음과 같이 구성된다. 2장에서는 SVC에서의 프레임 간 모드 결정 방법이 소개된다. 3장과 4장에서는 통계적 가설 검증기법에 대하여 소개하고, SVC의 참조 소프트웨어인 JSVM 5.0 [4]에 어떻게 삽입되었는지와 통계적 가설 검증 기법을 이용한 제안된 알고리즘을 소개한다. 5장에서는 제안된 알고리즘을 이용하여 복잡도 감소와 성능유지 측면에서 실

사용하기 위하여, X_C 와 X_R 을 현재 블록과 참조 블록의 독립적인 정규 모집단이라고 가정하고 다음과 같이 정의한다.

$$X_C \sim N(\mu_C, \sigma_C^2), X_R \sim N(\mu_R, \sigma_R^2) \quad [4]$$

여기서 μ_C 와 μ_R 은 알려져 있다고 가정하고, 다음과 같은 가설 검증을 실시한다.

$$\begin{aligned} H_0 : \mu_C &= \mu_R \\ H_1 : \mu_C &\neq \mu_R \end{aligned} \quad [5]$$

여기서 H_0 를 귀무가설(null hypothesis)이라고 하며 그 것은 두 개의 픽셀 그룹이 같은 모집단으로부터 온 것인지를 판단한다. 만일 설정된 귀무가설이 받아들여지면, 현재 블록과 복원된 참조 블록내의 두 개의 픽셀 그룹은 같은 모집단으로부터 온 것으로 판단하고 테스트하는 블록의 크기로 모드를 결정하게 된다. $\overline{X_C}, \overline{X_R}, S_C^2, S_R^2$ 을 현재 블록과 참조 블록 내 픽셀들의 평균과 분산이라고 하자. 이때 알려지지 않은 모집단의 분산이 같다고 가정하고 식 (5)에 대하여 가설 검증을 실시한다. 두 평균에 대하여 귀무가설을 테스트하기 위해 다음과 같은 t 통계 분포를 사용한다. 이때 자유도 $\nu = 2n_1n_2 - 2$ 를 가진 t 통계 분포는 다음과 같다.

$$t_0 = \frac{\overline{X_C} - \overline{X_R}}{S_p \sqrt{2/(n_1n_2)}} \quad [6]$$

그리고

$$S_p^2 = \frac{(n_1n_2 - 1)S_C^2 + (n_1n_2 - 1)S_R^2}{2n_1n_2 - 2} \quad [7]$$

S_p^2 는 두 개의 샘플 집단을 대표하는 분산으로 공분산이라고 한다. 일단 t 통계 분포가 생성되면 α 를 갖는 신뢰 구간 (Confidence Interval)을 설정하여, 구간 내에 t 통계 분포가 존재하면, 즉 $-t_{\alpha/2, \nu} \leq t_0 \leq t_{\alpha/2, \nu}$ 이면 귀무가설을 받아들이고 그렇지 않을 경우 받아들이지 않는다.

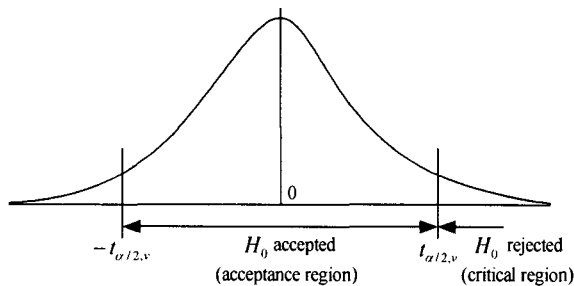


그림 4 t 통계 분포와 신뢰 구간

그림 4는 t 통계 분포의 예를 보여준다. 일단 신뢰 구간을 설정하여 샘플의 평균과 분산을 이용한 t 통계 분포가 신뢰 구간 내로 들어오면 두 모집단의 평균이 같다고 판단하는 것이다.

나. 두 분산의 가설에 대한 검증

두 개의 평균에 대한 가설 검증과 마찬가지로 두 개의 분산에 대한 가설 검증을 실행한다. 식 (5)와 마찬가지로 다음과 같이 분산에 대한 귀무가설을 설정한다.

$$\begin{aligned} H_0 : \sigma_R^2 &= \sigma_C^2 \\ H_1 : \sigma_R^2 &\neq \sigma_C^2 \end{aligned} \quad [8]$$

이 두 개의 분산에 대한 가설을 검증하기 위하여 z 통계 분포를 이용한다. z 통계 분포는 다음과 같이 정의한다.

$$Z_0 = \frac{S_C - S_R}{S_p / \sqrt{n_1n_2}} \quad [9]$$

여기서 공통 표준 편차의 예측치인 S_p 는 다음과 같이 정의한다.

$$S_p^2 = \frac{(n_1n_2 - 1)S_C^2 + (n_1n_2 - 1)S_R^2}{2n_1n_2 - 2} \quad [10]$$

z 통계 분포는 일반적으로 정규 분포를 따른다. 이 것 역시 일단 신뢰 구간을 설정하고 현재 테스트되고 있는 z 통계 분포가 신뢰 구간에 존재하면 귀무가설을 받아들이고 그렇지 않으면 받아들이지 않는다.

여기서 평균과 분산에 대한 귀무가설이 모두 받아들여지지만 현재 블록과 참조 블록은 같은 모집단으로부터 온 일치된 블록이라고 판단하기 때문에 현재 블록과 하위 블록에 대한 움직임 벡터 추정 및 모드 결정 없이 현재 블록을 16x16 블록이나 8x8 블록으로 결정하여 모드 결정 과정을 조기에 완료 한다.

4. 제안하는 프레임 간 고속 블록 모드 결정 과정

연산량 감소를 위한 고속 블록 모드 결정 방법은 3장에서 기술한 통계적 가설 검증 테스트에 기반 한다. 제안하는 블록 모드 결정 방법은 다음과 같은 과정을 따른다.

과정 1 : 현재 블록이 P-픽처 예측이면 먼저 스킵모드 실행, B 예측이면 직접모드를 실행한다.

과정 2 : 16x16 매크로 블록에 대하여 움직임 벡터를 추정한다.

과정 3 : 움직임 벡터를 이용하여 이동한 최적의 복원된 참조 매크로 블록과 현재 블록에 대하여 평균과 분산의 통계적 가설 검증을 실시한다.

과정 4 : 평균과 분산에 대한 가설 검증을 모두 통과하면 현재의 블록은 직접 또는 스킵 모드와 16x16 블록을 비교하여 최적의 모드를 결정하고 모드 결정과정을 완료한다. 통과하지 않으면 과정 5로 진행한다.

과정 5 : RD 최적화 방법에 기반하여 16x8, 8x16, 8x8 블록에 대하여 움직임 벡터를 추정한다.

과정 6 : 각 8x8 블록에 대하여 평균과 분산의 통계적 가설 검증 테스트를 한다.

과정 7 : 각각의 8x8 블록에 대하여, 가설 검증이 통과되면 현재 블록은 8x8 블록으로 결정되고, 8x8 서브 블록에 대한 모드 결정과정을 완료한다. 그렇지 않으면 과정 8로 간다.

과정 8 : RD 최적화 방법에 기반하여 8x, 8x4,4x8, 4x4 블록에 대하여 움직임 벡터 추정 및 모드 결정을 한다.

과정 9 : 다음 매크로 블록에 대한 모드 결정을 진행한다.

과정 2에서 RD cost가 최소가 되는 움직임 벡터를 추정하여 가설 검증 테스트를 할 참조 블록을 찾는다. 즉 다중 참조 프레임 중에서 몇 번째 프레임의 매크로 블록인지 찾아 현재 블록과 참조 블록간의 가설 검증 테스트를 한다. 그림 5는 통계적 가설 검증이 적용되는 현재 블록과 복원된 블록을 보여준다.

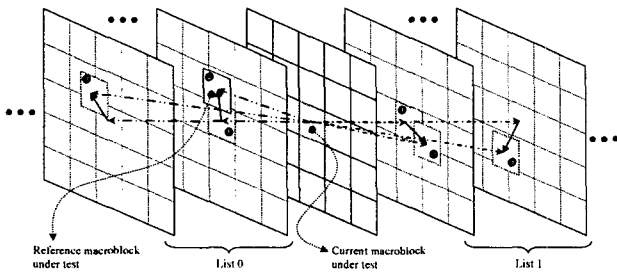


그림 5 본 알고리즘이 적용된 프레임

그림 6과 그림 7은 3장에서 기술한 t 와 z 통계 분포를 보여준다. 이것은 Mother&Daughter 샘플 비디오의 처음 100프레임을 사용하였고 양자화 파라미터는 32를 사용하였다.

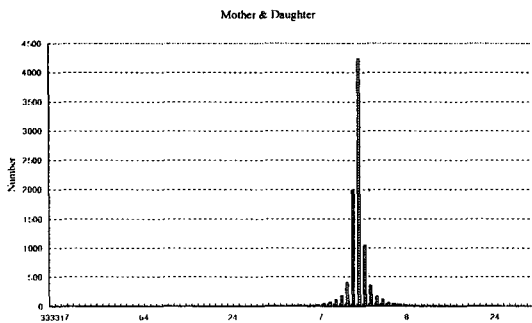


그림 6 Mother&Daughter 영상의 t 통계 분포

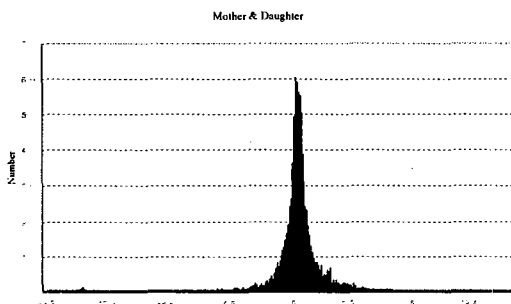


그림 7 Mother&Daughter 영상의 z 통계 분포

였다. Mother & Daughter 비디오는 움직임이 적은 비디오 샘플 중 하나이므로 현재 블록과 참조 블록간의 유사도가 높기 때문에 중앙에 많은 샘플이 집중하는 것을 알 수 있다.

5. 실험 결과

실험을 위해 SVC의 참조 소프트웨어인 JSVM5.0을 사용하였

다. 참조 프레임의 개수는 5개, GOP의 크기를 4로 하였으면 움직임 벡터의 정밀도는 1/4로 설정하였다. 움직임 벡터 추정 시 움직임 추정 범위는 32 픽셀로 설정하였고 움직임 벡터 추정을 위해 다이아몬드 움직임 벡터 추정법을 이용하였다. 움직임벡터 추정 시 왜곡도 추정을 위해 정수 픽셀단위에서는 SAD(Sum of Absolute Difference)를 서브 픽셀단위에서는 하다마드 변환을 사용하였다. 시간적 스케일러비티의 연산량 감소 추정을 위해 화질적 스케일러비티와 공간적 스케일러비티는 제외하고 실험을 하였다. 다양한 비디오 샘플을 사용하기 위하여 움직임이 적은 비디오 샘플부터 움직임이 많은 것까지 다양한 특성을 가지는 샘플 비디오를 사용하였으며 Clare QCIF, Mother &Daughter QCIF, Carphone QCIF, Foreman QCIF, Harbour QCIF를 사용하였다. 통계적 가설검증을 위해 설정해야하는 신뢰구간 (Confidence Interval)은 30%로 설정하였다. 본 실험을 위해 사용한 컴퓨터 환경은 3.2GHz 클럭 속도와 1GB 메모리를 가지는 펜티엄 4 데스크탑 PC를 사용하였다. 실험결과는 복잡도 감소로 인한 부호화 시간의 감소와 이에 따른 PSNR과 bitrate의 변화량을 측정하여 성능을 비교하였다. 본 알고리즘의 목적은 부호화 과정의 복잡도를 감소시켜 전체 부호화 시간을 감소시키는 것으로 한다. 따라서 부호화 시간을 측정하여 원 참조 소프트웨어보다 얼마만큼의 부호화 시간의 감소를 가져왔는지 측정한다. 그리고 본 알고리즘 적용 전후의 PSNR 변동량과 bit rate의 변화량을 dB와 %로 각각 측정하였다. 알고리즘의 적용 전후의 RD 성능의 변화량을 측정함으로써 알고리즘의 적용 후 복잡도 감소에도 RD 성능을 유지하는지 보기 위함이다. 테스트 영상으로는 움직임이 적은 순서대로 Clare(C), Mother & Daughter(M&D), Carphone(CA), Foreman(F) QCIF(30Hz) 영상을 사용하였고 다음으로 움직임은 적으나 텍스처가 미세한 특징을 보여주는 Harbour(H) QCIF 영상을 사용하였다.

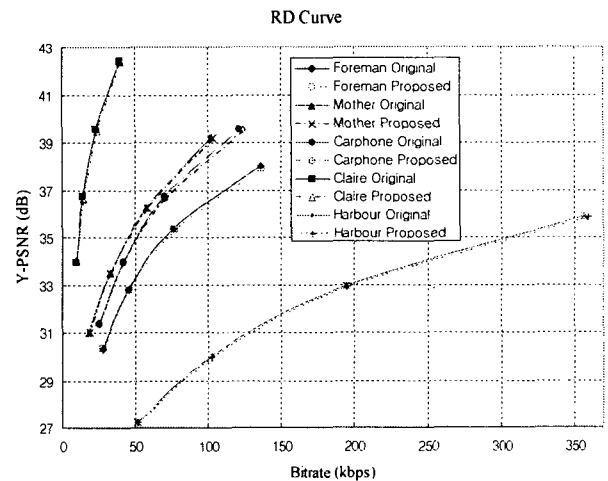


그림 8 제안된 알고리즘 적용 전후의 RD 성능

그림 8은 원 참조소프트웨어와 알고리즘 적용 후의 RD 성능을 그래프로 나타내었다. 알고리즘 적용 전 후의 PSNR과 bitrate는 거의 일치하므로 Rate-Distortion 측면에서 제안하는 알고리즘은 우수한 성능을 보여 줄 수 있다. 연산량 감소를 통한 부호화 감소는 표 1에 나타내었다. Qp 28에 대하여 32%부터 최대 57%까지 연산량 감소를 보여준다.

Qp	Δ	C	M&D	CA	F	H
28	PSNR(dB)	-0.05	-0.02	-0.06	-0.08	-0.05
	Bitrate(%)	1.74	0.87	1.63	0.71	0.75
	Complexity(%)	-48	-44	-34	-32	-57
32	PSNR(dB)	-0.07	0.00	-0.05	-0.03	-0.04
	Bitrate(%)	1.40	0.87	0.80	1.35	0.47
	Complexity(%)	-40	-34	-28	-27	-44
36	PSNR(dB)	-0.08	-0.02	-0.02	-0.04	-0.01
	Bitrate(%)	0.84	0.85	-0.15	0.57	-0.15
	Complexity(%)	-29	-26	-21	-22	-21
40	PSNR(dB)	0.04	-0.01	-0.01	0.04	-0.01
	Bitrate(%)	1.04	0.52	-0.08	-0.34	-0.07
	Complexity(%)	-23	-17	-21	-20	-21

표 1 RD 성능의 변화량과 복잡도 감소 비율

표 1에서 보는 것처럼 움직임이 많을수록 연산량 감소 비율이 줄어든다. 이것은 움직임이 많을수록 t 및 z 통계가 보다 고르게 분포하기 때문에 고정된 신뢰구간 하에서 가설 검증을 통과할 비율이 작아지기 때문이다. 그리고 같은 영상에 대하여 Qp가 증가할수록 연산량 감소 비율이 줄어든다. 이것은 높은 Qp를 가지고 양자화를 하면 복원된 영상에서는 미세한 부분이 높은 양자화 파라미터에 의하여 없어지는 효과를 발생하므로 원영상과 복원된 참조 영상간의 유사도는 떨어지게 지게 된다. 따라서 가설검증을 통과할 비율역시 낮아지게 때문이다.

표 1에서 보는 것처럼 본 알고리즘 적용 시 복잡도가 최대 57%까지 감소하며 이에 따른 RD 성능 저하는 PSNR의 경우 최대 0.08dB 감소, 비트율 증가분은 최대 1.74%이므로 이는 RD 성능 측면에서 무시할 수 있을 정도로 작은 수치이다.

6. 결론

SVC 부호화기의 연산량 감소를 위해 통계적 가설 검증을 이용한 프레임 간 고속 모드 결정 알고리즘을 제안하였다. 이것은 현재 블록과 참조 블록간의 통계적인 특성을 이용하기 때문에 기존의 SAD(Sum of Absolute Difference)나 SSD(Sum of Squared Difference)를 이용하여 임계값을 설정함으로써 모드 결정을 조기에 완료하는 방법보다 신뢰도가 높다. 왜냐하면 기존의 방법들은 많은 실험을 통하여 최적의 임계값을 찾아내며 각기 다른 특성을 가지는 콘텐츠에 대하여 다른 임계값을 사용하지만 제안된 알고리즘은 통계적인 분포에 신뢰 구간을 적용하여 신뢰 구간만큼의 연산량 감소를 얻기 때문에 콘텐츠의 특성과 무관하게 알고리즘을 적용할 수가 있다.

본 알고리즘은 최대 57% 만큼의 연산량 감소효과를 얻을 수 있으며 그럼에도 불구하고 PSNR가 최대 0.08,dB감소, 비트율이 최대 1.74% 밖에 증가하는데 그쳤다. 이는 RD 성능 측면에서 무시할 수 있을 정도로 작은 수치이다.

참고 문헌

- [1] ISO/IEC JTC 1/SC 29/WG 11 : 'Text of ISO/IEC 14496-10:2006/PDA M3 Scalable Video Coding', MPEG75/N7796, Thailand, January 2006
- [2] Libo Yang, Keman Yu, Jiang Li and Shipeng Li, " An

- Effective Variable Block Size Early Termination Algorithm for H.264 Video Coding", IEEE Transaction on Circuit and Systems for Video Technology, vol 15, no 6, June 2005.
- [3] D.Wu, F.Pan, K.P.Lim, S.Wu, Z.G.Ki, S.Rahardja and C.C. "Fast Intermode Decision in H.264/AVC Video Coding", IEEE Transaction on Circuits and Systems for Video Technology, vol15, no 6, July 2005.
- [4] ISO/IEC JTC 1/SC 29/WG 11/N7797: ' JSVM5 Software', MPEG 75/N7797, Thailand, January 2006
- [5] Iain E.G. Richardson, "H.264 and MPEG-4 Video Compression - Video Coding for Next-generation Multimedia", Wiley, 2003.
- [6] 박내리, 전동산, 김재근, 한종기, "Scalable Video Codec을 위한 효율적인 비트율 제어 기법", 방송공학회 논문지, 제10권, 제4호, 2005년 12월
- [7] William W.Hines, C.Montgomery, David M.Goldsman and Connie M.Borror, "Probability and Statistics in Engineering.", Wiley 4th Edition, 2003.