

# 다시점 비디오 영상 부호화기를 위한 시점 간 PSNR 분산 감소 비트율 제어 알고리즘

류승철, 서정동, Xingang Liu, Zheng Yan, 손광훈

연세대학교

khsohn@yonsei.ac.kr

## Reduced Inter-view PSNR variance Rate-control Algorithm for Multi-View Video coding

Seungchul Ryu, Jungdong Seo, Xingang Liu, Zheng Yan, Kwanghoon Sohn

School of Electrical and Electronic Engineering, Yonsei University

### 요약

다시점 영상 비디오를 부호화하기 위해 다시점 영상 부호화기(MVC)가 제안되었다. MVC가 다시점 영상 비디오를 효율적으로 압축함에도 불구하고 여전히 제한된 대역폭의 채널을 통해 다시점 영상을 전송하기에는 한계가 있다. 따라서 제한된 대역폭을 효과적으로 활용할 수 있도록 효율적인 비트율 제어 알고리즘이 요구된다. 2D 영상 부호화를 위한 비트율 제어 알고리즘은 활발한 연구가 이루어졌지만 다시점 영상 부호화를 위한 비트율 제어 알고리즘은 아직 충분히 연구되어지지 않았다. 본 논문에서는 시점 간 화질 분산과 계층적 B 픽처 구조를 고려한, 다시점 영상 부호화를 위한 비트율 제어 알고리즘을 제안한다. 제안된 알고리즘은 부호화 전 단계, 시점 레벨 비트율 제어, GOP 레벨 비트율 제어, 픽처 레벨 비트율 제어의 네 단계로 이루어진다. 부호화 전 단계에서 시점 간 비트율의 비중과 계층 레벨 간의 비중을 계산한다. 그리고 시점의 목표 비트율, GOP의 목표 비트율, 픽처의 목표 비트율이 각각 시점 레벨 비트율 제어단계, GOP 레벨 비트율 제어단계, 픽처 레벨 비트율 제어 단계에서 계산된다. 실험 결과는 제안된 방식이 시점 간 화질(PSNR)의 분산을 줄이면서 비트율을 제어하는데 효과적이라는 것을 보여준다.

### 1. 서론

최근 3DTV는 HDTV의 다음 세대 방송매체로 각광받고 있다. 3DTV의 비디오 신호를 부호화하기 위해서는 새로운 3D 비디오 부호화 방식이 필요하다. 이에 새로운 부호화 방식을 위한 많은 연구가 진행되었다 [1-2]. 그 결과로 여러 대의 카메라로부터 여러 시점의 영상을 동시에 획득한 영상을 부호화 하는 다시점 영상 부호화 방식이 제안되었다. 다시점 영상 부호화 방식은 각 시점 간의 데이터 중복성을 줄여 효율적으로 다시점 영상을 압축한다. 하지만 다시점 영상 부호화의 적용을 위해서는 여전히 해결되어야 할 문제점들이 남아있다. 많은 실질적 적용 분야에서 사용가능한 대역폭이 제한되어 있기 때문에 다시점 영상 부호화의 비트율 제어는 3D 비디오의 부호화에서 매우 높은 중요성을 차지한다.

2D 영상의 경우 각각의 표준(MPEG-2, MPEG-4, H.264 등)에 맞는 비트율 제어 알고리즘들이 활발하게 연구되었으나[3-5] 대부분의 2D 영상 부호화를 위한 비트율 제어 알고리즘은 다시점 영상 부호화의 비트율 제어에 적합하지 않다[6-9]. 이에 다시점 영상 부호화를 위한 비트율 제어 알고리즘의 개발이 요구된다.

다시점 영상 부호화 국제 표준에서는 인접한 시점 간의 데이터 중복성을 줄여 부호화 효율을 높이기 위해 다른 시점의 복원된 영상을 참조 영상으로 사용한다[10]. 따라서 다시점 영상 부호화의 비트율을 제어하기 위해서는 모든 시점을 위한 각각의 비트율을 제어하여야 한다. 이를 위한 몇 가지의 다시점 영상의 비트율 제어 알고리즘이 제안되었으나[11-12] 제안된 다시점 영상 부호화 비트율 제어 알고리즘에서는 다시점 영상 부호화의 특징 중 하나인 계층적 B 픽처 구조(Hierarchical B Picture Structure)를 고려하지 않거나[11] 큰 비중을 두지 않았다.[12] 또한, 다시점 영상의 경우 시점 간 화질을 제어해주지 않으면 각 시점 간의 화질 차이가 생겨 문제가 될 수 있다. 따라서 본 논문에서는 계층적 B 픽처 구조를 고려하고, 시점 간 화질차이를

줄이기 위한 다시점 영상 부호화에 적합한 비트율 제어 알고리즘을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 다시점 영상 부호화의 구조에 대해서 살펴보고 3장에서 계층적 B 픽처 구조에서 시점 간 화질 차이를 줄이기 위한 다시점 영상 부호화의 비트율 제어 알고리즘을 제안한다. 제안된 기법의 실험 결과를 4장에서 기술하고 마지막으로 5장에서 결론을 맺는다.

### 2. 다시점 영상 부호화 구조

다시점 영상은 여러 시점의 영상을 통해 사용자에게 3D 효과를 줄 수 있다. 따라서 사용자는 영상을 좀 더 사실적으로 감상할 수 있다. 다시점 영상을 압축하기 위해 H.264/AVC를 바탕으로 한 다시점 영상 부호화가 개발되었고 2006년 1월, 방콕에서 열린 MPEG 미팅에서 다시점 영상 부호화 구조가 결정되었다. 그림 1은 다시점 영상 부호화기(Multi-view video coding)의 부호화 구조를 보여준다. 다시점 영상 부호화에서는 계층적 B 픽처 구조가 사용된다[13]. 계층적 B 픽처 구조에서는 다른 GOP 구조들(IBBP, IPPP)과는 달리 B-픽처가 다른 B-픽처의 참조 영상으로 사용되어질 수 있다.

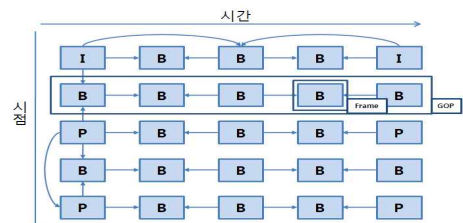


그림 1 다시점 영상 부호화 계층적 B 픽처 구조

이전의 다시점 영상 부호화 연구는 부호화의 효율을 높이는 데 초점이 맞춰져 있었다. 그럼에도 불구하고 제한된 대역폭의 채널을 통해 다시점 영상을 전송하기에는 한계가 있다. 특히, 시점의 수가 늘어남에 따라 더 많은 대역폭을 필요로 하게 된다[14].

## 2. 제안된 비트율 제어 알고리즘

본 논문에서는 다시점 영상 부호화기(MVC)에 적합한 비트율 제어 알고리즘을 제안한다. 이 알고리즘은 JM12.0의 H.264/AVC 비트율 제어 알고리즘을 바탕으로 다시점 영상 부호화의 구조를 고려하여 확장되었다. 제안된 알고리즘은 4단계로 이루어져 있다. 첫 번째 단계는 부호화 전 단계이고 두 번째 단계는 시점-레벨 비트율 제어, 세 번째 단계는 GOP-레벨 비트율 제어, 마지막으로 네 번째 단계는 픽처-레벨 비트율 제어이다. 각 단계의 과정을 아래에서 상세하게 기술한다.

### A. 부호화 전 단계

제안된 알고리즘에서는 시점 간 화질의 분산과 계층 레벨간의 비트율의 비중을 이용하여 비트율 제어를 수행한다. 이후의 단계에서 사용할 파라미터들을 부호화 전 단계에서 하나의 GOP를 미리 부호화하여 계산한다. 부호화 전 단계에서 계산하는 파라미터는 다음과 같다.

- $W(i)$  : 시점 간 비트율의 비중 (시점 간 PSNR 분산 고려 전)

$$W(i) = \frac{Bits(i)}{\sum_j Bits(j)} \quad (1)$$

$Bits(i)$ :  $i$  번째 시점의 부호화에 사용된 비트  
 $W(i)$ :  $i$  번째 시점 영상의 비트율 비중

- $H(i)$ : 계층 레벨간의 비트율의 비중

$$H(i) = \frac{Average\ Bits(i)}{\sum_j Average\ Bits(j)} \quad (2)$$

$Average\ Bits(i)$ :  $i$  번째 계층 레벨 픽처를 부호화하는데 사용된 평균 비트  
 $H(i)$ :  $i$  번째 계층 레벨 비중

- $W_i$  : 시점 간 비트율의 비중 (시점 간 PSNR 분산 고려 후)

$$W_i = W(i) \times (1 + \alpha(1 - \frac{PSNR(i)}{PSNR_{ave}})) \quad (3)$$

$PSNR(i)$ :  $i$  번째 시점의 PSNR  
 $PSNR_{ave}$ : 모든 시점의 평균 PSNR  
 $\alpha$ : PSNR 분산 파라미터 (12)

각 시점 영상을 부호화 하는데 필요한 비트의 비중을 하나의 GOP를 통해 계산한다( $W(i)$ ). 그리고 시점 간 화질의 분산을 고려하여 식-(3)에 의하여 시점 간 비트율의 비중을 계산한다.

계층적 B 픽처 구조의 특성 상 계층 레벨 간의 비트율의 비중이 다르다. 따라서 이 비중을 고려하여 비트율 제어를 수행하여야 한다. 부호화 전 단계에서 계산된 계층 레벨간의 비트율의 비중은 부호화 단계에서 첫 번째 GOP의 비트율 제어에 사용된다.

### B. 시점 레벨 비트율 제어

시점-레벨 비트율 제어 단계에서는 각 시점의 목표 비트율( $TV_i$ )을 계산한다. 다시점 영상 부호화에서 각 시점 영상을 부호화하는데 필요한 비트는 시점의 종류(I시점, P시점, B시점)에 따라 다른 비중을 가진다. 또한 영상의 특성에 따라 그 비율의 정도는 달라진다. 따라서 시점 간 목표 비트율 계산 시 이 점을 고려하여야 한다.

위에서 계산된 시점 간 비트율 비중( $W(i)$ )을 이용하여 각 시점의 목표 비트율( $TV_i$ )을 계산한다. 그 방법은 식-(4)와 같다.

$$TV_i = Total\ Bits\ rate \times W_i \quad (4)$$

$TV_i$ :  $i$  번째 시점 영상의 목표 비트율

Total Bits rate는 모든 시점의 영상의 비트율이며  $TV_i$ 는  $i$  번째 시점의 목표 비트율이다.

### C. GOP 레벨 비트율 제어

#### 1) GOP 레벨 비트율 제어

GOP 레벨 비트율 제어 단계에서는 각 GOP를 부호화하는데 사용될 목표 비트를 할당한다. 그 방법은 식-(5)와 같다.

$$T_{GOP,j} = \frac{TV_i}{frame\ rate} \times GOP_{size} + Remains - \frac{R_{1st\ anchor}}{GOP\ 개수} \quad (5)$$

$T_{GOP,j}$ :  $i$  번째 시점,  $j$  번째 GOP의 목표 비트율  
 $Remains$ : 이전 GOP를 부호화하고 남은 비트  
 $R_{1st\ anchor}$ : 첫 번째 픽처를 부호화하는데 사용된 비트

각 시점의 목표 비트율( $TV_i$ )을 프레임 레이트로 나눈 후 GOP의 크기를 곱하여 GOP를 부호화하는데 사용될 목표 비트를 계산한다. 여기에 이전 GOP를 부호화하고 남은 비트를 더해주고 첫 번째 앵커 픽처에서 사용한 비트를 각 GOP 목표 비트율에서 균등하게 감한다. 첫 번째 앵커 픽처를 부호화하는데 필요한 비트는 상대적으로 크기 때문에 첫 GOP의 목표 비트율에 포함시키지 않고 전체 영상에 균등하게 나누어 전체적인 화질 저하를 막을 수 있다.

#### 2) 첫 픽처 QP값 계산

각 GOP의 첫 픽처의 QP값은 이전 GOP의 높은 계층 레벨 픽처의 평균 QP값을 사용한다. 본 논문에서는 높은 계층 레벨 픽처로 앵커 픽처와 레벨1 픽처, 레벨2 픽처를 사용하였다.

#### 3) 계층 레벨 비중 계산

기존의 다시점 영상 부호화 방식에서는 계층적 B 픽처 구조가 사용된다. 각 계층 레벨간의 픽처들을 부호화하는데 사용되는 비트는 다른 비율을 가지게 된다. 따라서 제안된 알고리즘에서는 이전 GOP의

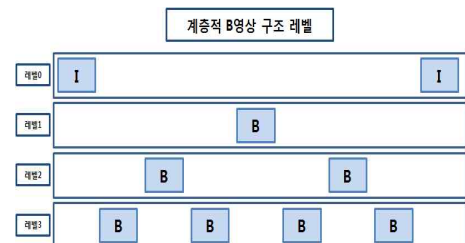


그림 2 계층적 B 픽처 구조의 레벨

표 1 다시점 부호하용 위한 비트율 제어 실험 결과

영상	목표비트율	JMVC 6.0			제한 된 알고리즘			RCE (%)	PSNR 감소 (dB)	PSNR 분산 감소
	(kbps)	비트율(kbps)	PSNR(dB)	PSNR 분산	비트율(kbps)	PSNR(dB)	PSNR 분산			
ballroom	724.6720	724.6720	32.3936	0.0078	734.0543	32.2453	0.0295	1.29	0.1482	-0.0217
	1149.7110	1149.7110	34.5326	0.0131	1156.4049	34.3166	0.0411	0.58	0.2160	-0.0280
	1880.1040	1880.1040	36.5910	0.0173	1889.7358	36.3495	0.0246	0.51	0.2415	-0.0073
	3153.8400	3153.8400	38.3554	0.0280	3153.1679	38.0867	0.0049	0.02	0.2686	0.0231
	5593.2396	5593.2396	39.9462	0.0530	5568.8543	39.6603	0.0066	0.44	0.2859	0.0464
rena	304.9900	304.9900	35.9473	0.6207	303.9382	35.6768	0.2805	0.34	0.2906	0.3402
	494.2640	494.2640	38.4255	0.4348	492.6050	38.1165	0.1130	0.34	0.3090	0.3218
	831.9980	831.9980	41.0237	0.2848	828.3259	40.5650	0.0637	0.44	0.4587	0.2211
	1438.7650	1438.7650	43.4472	0.2199	1431.3506	43.0330	0.0219	0.52	0.4142	0.1980
	2572.7610	2572.7610	45.6115	0.1682	2556.6716	45.2288	0.0098	0.63	0.3827	0.1584
exit	257.6790	257.6790	35.2938	0.0190	259.4050	35.0222	0.0098	0.67	0.2716	0.0091
	402.6030	402.6030	37.0424	0.0067	404.8888	36.8573	0.0006	0.57	0.1851	0.0061
	674.7160	674.7160	38.4794	0.0104	675.6469	38.2940	0.0000	0.14	0.1855	0.0104
	1288.8170	1288.8170	39.5577	0.0143	1286.5457	39.4680	0.0030	0.18	0.0897	0.0113
	2934.6000	2934.6000	40.6388	0.0176	2923.5679	40.5429	0.0020	0.38	0.0959	0.0156
akkokayo	494.5700	494.5700	34.1007	0.0159	495.5086	33.8186	0.0564	0.19	0.2821	-0.0405
	736.7830	736.7830	36.5324	0.0340	737.3703	36.2666	0.0417	0.08	0.2658	-0.0077
	1148.4790	1148.4790	39.0065	0.0674	1146.6913	38.7586	0.0379	0.16	0.2479	0.0295
	1870.9930	1870.9930	41.2999	0.0868	1864.6963	41.0550	0.0386	0.34	0.2449	0.0483
	3290.3780	3290.3780	43.3109	0.0856	3277.8495	43.0266	0.0501	0.38	0.2843	0.0355

계층 레벨간의 비트율의 비중( $H(i)$ )을 바탕으로 계층 레벨의 목표 비트를 할당한다. 단, 첫 GOP에서는 이전 GOP의 정보가 없으므로 부호화 전 단계에서 얻은  $H(i)$ 를 사용한다. 레벨간의 비트율 비중을 계산하는 방법은 식-(6)과 같다. Average Bits( $i$ )는 각 계층 레벨의 프레임들의 평균 복잡도이다. 계층 레벨 간의 비트율의 비중( $H(i)$ )는 픽처 레벨 비트율 제어 단계에서 각 픽처의 목표 비트율( $T_i$ )를 계산하는데 사용된다.

$$H(i) = \frac{\text{Average Bits}(i)}{\sum_j^{\text{Max level}} \text{Average Bits}(j)} \quad (6)$$

Average Bits( $i$ ): 이전 GOP의  $i$ 번째 계층 레벨 픽처들의 평균 복잡도  
 $H(i)$ :  $i$ 번째 계층 레벨의 비중

#### D. 픽처 레벨 비트율 제어

픽처-레벨 비트율 제어 단계에서는 각 픽처의 목표 비트율( $T_{i,j}$ )을 계산한다.

##### 1) 픽처 레벨 비트율 제어

계층 레벨간의 비트율의 비중( $H(i)$ )과 현재 GOP의 남은 목표 비트율( $T_{GOPi,j}$ )을 이용하여 현재 픽처의 목표 비트율을 계산한다. 그 방법은 식-(7)과 같다.

$$T_{i,j} = \frac{H(j)}{\sum_{k=\text{Hier level}} H(k)N(k)} \times T_{GOPi,j} \quad (7)$$

$N(j)$ : 각 계층 레벨의 남은 픽처의 수  
 $T_{i,j}$ :  $i$ 번째 시점,  $j$ 번째 픽처의 목표 비트율

##### 2) R-D model

제한된 알고리즘에서 사용된 R-D model은 식-(8)과 같다. 2차 모델에 의해 계산되어진 양자화 레벨  $Q_i$ 는 R-D 최적화 과정에 사용

된다[16]. 파라미터  $x_i, y_i$ 는 선형 회귀법을 이용하여 계산되고,  $MAD_i$ 는 이전 픽처( $i-1$ )의  $MAD$ 를 이용한 선형 모델에 의해 예측된다.

$$\frac{R_i - H_i}{MAD_i} = \frac{x_i}{Q_i} + \frac{y_i}{Q_i^2} \quad (8)$$

$x_i, y_i$ : 모델 파라미터  
 $Q_i$ : 양자화 파라미터  
 $MAD_i$ : 현재 픽처의 복잡도  
 $R_i$ : 현재 픽처를 부호화하는데 사용되는 비트  
 $H_i$ : 헤더 정보, 모션 벡터 비트

### 3. 실험 결과

제한된 다시점 영상 부호화 비트율 제어 알고리즘의 성능을 테스트하기 위해 4개의 영상(ballroom, rena, exit, akkokayo)을 사용하였다. 각 영상의 특성을 표 2에 나타내었다.

제한된 비트율 제어 알고리즘은 JMVC 6.0 [15] 참조 소프트웨어에 구현되었다. 제한된 비트율 제어 알고리즘이 구현된 JMVC와 제한된 알고리즘이 빠진 JMVC로 각각 실험하여 그 결과를 비교하였다. 실험 조건은 표 3와 같으며 비트율 제어 정확도를 측정하기 위해 RCE(Rate Control Error)가 사용되었다.

$$RCE = \frac{|R_{\text{target}} - R_{\text{actual}}|}{R_{\text{target}}} \times 100\% \quad (9)$$

$R_{\text{target}}$ : 목표 비트율  
 $R_{\text{actual}}$ : 실제 부호화 비트율

여기서  $R_{\text{target}}$ 과  $R_{\text{actual}}$ 은 각각 목표 비트율과 실제 부호화 비트율을 의미한다.

표 1에 제한된 비트율 제어 알고리즘의 실험 결과를 나타내었다. 평균 RCE는 0.41%로, 제한된 알고리즘을 사용하여 효과적으로 비트율을 제어할 수 있음을 알 수 있다. 또한 비트율 제어 알고리즘을 적용하지 않았을 때보다 시점 간 화질(PSNR)의 분산이 줄었다(평균

0.0685, 최대 0.3402). 따라서 제안된 비트율 제어 알고리즘을 이용하여 시점 간 화질의 분산을 줄이면서 비트율 제어를 효과적으로 수행할 수 있음을 확인 할 수 있다.

표 2 실험 영상

실험 영상		
영상	해상도	시점
ballroom	640x480	0~2
rena	640x480	38~40
exit	640x480	0~2
akkokayo	640x480	0~2

표 3 실험 조건

실험 조건	
프레임 레이트	25 frame/s
GOP 길이	16
탐색 범위	64 pixels
탐색 모드	Fast search
총 픽처 수	81
참조 영상 갯수	2

#### 4. 결론

본 논문에서 부호화 전 단계에서 얻은 파라미터를 바탕으로 다시 점 영상 부호화를 위한 시점 간 화질의 분산 감소 비트율 제어를 수행하는 알고리즘을 제안하였다. 실험 결과를 통해 제안된 알고리즘이 다시점 영상의 비트율을 효과적으로 제어함(평균 RCE : 0.41%)을 알 수 있다. 또한 제안된 방식을 적용하지 않았을 때보다 시점 간 화질의 분산이 감소함을 확인 하였다. 향후 시점 간 화질의 분산 감소 외에도 하나의 시점 내 화질의 변동을 줄여 전체적인 화질 저하를 방지할 수 있을 것을 기대한다.

#### 5. 참고문헌

[1] Ishfaq Ahmad, "Multi-view video : get ready for next-generation television", *IEEE Distributed Systems Online*, 2007.

[2] A. Kubota et al., "Multi-view imaging and 3DTV", *IEEE Signal Processing Magazine*, 2007.

[3] Z. G. LI, F. Pan, K. P. Lim and G. N. Feng, "Adaptive basic unit layer rate control for JVT", *JVT-G012*, Mar, 2003.

[4] MA S, GAO W, WU F. "Rate control for JVC coding scheme with HRD considerations", *Proceedings of International Conference on Image Processing[C]*, 2003.

[5] JINAG M Q, YI X. Q., LING N., "Improved frame-layer rate control for H.264 using MAD ratio", *Proceedings of IEEE International Symposium on Circuits and Systems*, 2004.

[6] X. Jing, L.P. Chau, W.C. Siu, "Frame complexity-based rate-quantization model for H.264/AVC intraframe rate control". *IEEE SIGNAL PROCESSING LETTER*, vol. 15, 2008.

[7] Y. Liu, Z. G. Lim Y. C. Soh, "A novel rate control scheme for low delay video communication of H.264/AVC

standard", *IEEE Transaction on Circuits and Systems for video technology*, vol 17, no. 1, January, 2007.

[8] W. Yuan, S. Lin, Y. Zhang, W. Yuan, H. Luo, "Optimum bit allocation and rate control for H.264/AVC", *IEEE Transaction on Circuits and Systems for video technology*, vol. 16, no. 6, June, 2006.

[9] D.K. Kwon, M.Y. Shen, C.C. J. Kuo, "Rate control for H.264 video with enhanced rate and distortion models", *IEEE Transaction on Circuits and Systems for video technology*, Vol. 17, no. 5, May, 2007.

[10] K. P. Lim, "Text Description of Joint Model Reference Encoding Methods and Decoding Concealment Methods", *JVT-K049*, Mar, 2004.

[11] T. Yan, L. Shen, P. An, H. Wang, Z. Zhang, "Frame-layer Rate control algorithm for Multi-view video coding", *Proceedings of the first ACM/SIGEVO Summit on Genetic and Evolutionary Computation*, 2009

[12] S. Park, D. Sim, "An Efficient Rate-control Algorithm for Multi-view Video Coding", *The 13th IEEE International Symposium on Consumer Electronics*, 2009.

[13] M. Flierl, B. Girod, "Multi-view Video Compression - Exploiting Inter-image similarities-", *IEEE SIGNAL PROCESSING MAGAZINE 1*, July, 2007.

[14] A. Leontaris, A.M. Tourapis, "Rate control reorganization in the JM (Joint Model) reference software" Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, *JVT-W042*, Apr, 2007.

[15] Joint Video Team of ITU-T VCEG and ISO/IEC MPEG. WD1 Reference software for MVC (JMVC) 6.0, Doc. *JVT-AF14*, November, 2009.

[16] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, G.J.Sullivan, "Rate-constrained coder control and comparison of video coding standards", *IEEE Trans. Circuits Systems Video Technol.* 13, 2003.