마크로 블록타입에 대한 코드워드 할당이 적응적으로 가능한 효율적인 동영상 부호화 방법

*박시내 *심동규 **이하현 **임성창 **정세윤 **최진수

*광운대학교 컴퓨터공학과

**한국전자통신연구원 방통융합미디어연구부

*psea1118@kw.ac.kr

Efficient video coding method of adaptive allocated codeword for macroblock type

*Sea-Nae Park *Dong-gyu Sim **Hahyun Lee **Sung-Chang Lim **Seyoon Jeong **Jin Soo Choi

*Computer Engneering, Kwangwoon University

**Broadcasting and Telecommunications Media Research Department, ETRI

요약

최근 멀티미디어 기능을 가지는 다양한 기기가 보급되고, 이러한 기기들을 위해 다양한 화질과 해상도를 가지는 멀티미디어 서비스가 소비되고 있다. 이러한 서비스들은 목적 및 기기의 종류에 따라 전송환경이 다르기 때문에, 압축된 영상이 가지는 특성 또한 다르다. 일반적인 동영상 압축 표준에서 구문요소들에 코드워드를 할당하는 방법은 표준화 과정에서 얻어진 확률과 통계적인 수치에 의해 최적으로 결정된 것이다. 하지만 확률과 통계적인 수치는 영상의 특성에 따라 차이가 있기 때문에 최적의 코드워드는 항상 달라지는데, 현재 표준압축방법은 이러한 확률적인 특성이나 통계적인 수치를 반영하기 어려운 단점을 가진다.

이에 본 논문에서는 영상의 특징 및 부호화 조건에 따른 압축데이터의 특징을 반영하여 슬라이스 단위로 마크로 블록 타입에 대한 코드워드를 적응적으로 할당하는 방법을 제안한다. 제안하는 방법에서 부호화 조건에 따른 압축 데이터의 특징을 파악하기 위해 정해진 부호화 조건에 따라 한 번의 부호화를 수행하여 마크로 블록 타입에 대한 슬라이스 통계를 구한다. 그 다음 구해진 통계에 근거하여 마크로블록 타입에 대해 최적의 코드워드를 할당하고, 이 코드워드에 근거하여 다시 한 번 부호화를 수행하고 이렇게 발생된 비트스트림을 복호화기로 전송하게 된다. 본 논문에서 제안한 방법을 적용한 실험결과 BD-rate가 약 0.2~1.7% 정도 감소하는 결과를 얻을 수 있었다.

1. 서론

MPEG-1, MPEG-2, MPEG-4 및 H.264/AVC는 대표적인 동영상 압축 표준으로 그 표준이 정해진 목적에 따라 여러 가지 응용분야에서 사용되고 있다. 이러한 압축 표준들은 한 가지의 표준 방법 안에서도 여러 가지 프로파일로 나눠지는데, 프로파일은 동영상 부호화/복호화 과정에서 알고리즘 상 들어가는 기술 구성 요소들이 규격화 된 것을 의미한다. 이러한 프로파일을 통해 압축 표준 방법들은 다양한 응용에서 적합한 구조의 부호화/복호화 알고리즘을 가지게 된다.

일반적인 동영상 압축 방법에 대한 표준화는 부호화/복호화 알고리즘에 대한 표준과 비트스트림을 쓰는 방법인 구문에 대한 표준으로 나눌 수 있는데, 같은 알고리즘에 대해서도 어떠한 구문을 가지느냐에 따라 부호화 효율이 달라질 수 있다. 이러한 이유로 효율적인 구문을 결정하는 것도 동영상 압축 표준화에서 중요한 이슈가 된다. 대부분 구문요소들에 대한 효율적인 코드워드 할당은 확률 및 통계에 근거하여 결정이 되며, 동영상이 가지는 특징을 반영하여 최소의 비트로 압축된 데이터를 표현하는 방법을 결정하게 된다.

최근 멀티미디어 기능을 가지는 다양한 기기가 보급이 되고, 이러한 기기들을 통해IP-TV, DMB, 화상전화 등 다양한 멀티미디어 서비스(Multiemedia Serivce)가 소비되고 있다. 이러한 서비스를 위한 동영상들은 다양한 화질과 해상도를 가지며 각각의 전송 환경이 다르기 때문에, 압축된 영상이 가지는 특성 또한 다르다. 일반적인 동영상 압축 표준 방법은 구문요소들에 대하여 정해진 코드워드를 할당하는 방법으로 부호화 및 복호화가 진행되며, 정해진 구문요소에 대한 코드워드는 표준화 과정에서 얻어진 확률과 통계적인 수치에 의해 최적으로 결정된 것이다. 하지만 확률과 통계적인 수치는 영상의 특성에 따라 차이가 있으며, 현재 표준 압축 방법은 이러한 특성에 따라 변하게 되는 확률과 통계적인 수치를 반영하기 어려운 단점을 가진다. 이에 다양한 환경에서 이루어지는 동영상 서비스에서 구문 요소들에 대해 적응적으로 코드워드를 할당할 수 있는 효율적인 구조의 동영상 부호화 방법이 필요하다. 본 논문에서는 영상의 특징 및 부호화 조건에 따른 압축데이터의 특징을 반영하여 슬라이스 단위로 마크로블록 타입에 대한 코드워드를 적응적으로 할당하는 방법을 제안한다.

2008년 VCEG (Video Coding Experts Group)을 통해 현재 코딩

하는 마크로 블록에 대해 이전에 코딩된 마크로블록 타입의 통계적인 정보를 이용하여 코드워드를 할당하는 방법이 제안[1]되었으나, 이전 프레임의 정보를 사용하는 경우 화면전환이 일어나거나 이전 프레임이 IDR 프레임인 경우 이전 프레임의 통계를 사용할 수 없거나 부적절하다는 단점을 가진다.

본 논문의 구성은 다음과 같다. 논문의 2장에서는 H.264/AVC에서 마크로블록 타입을 기술하는 방법을 소개하고, 3장에서는 본 논문에서 제안하는 방법을 소개한다. 4장에서는 제안하는 방법에 대한 실험결과를 보여주고 마지막으로 5장에서 결론을 맺는 것으로 논문을 마친다.

2. H.264/AVC에서 마크로블로 타입을 기술하는 방법

H. 264/AVC에서 마크로블록 타입에 대한 기술은 통계와 확률에 근거한 코드워드 할당을 통해 이루어진다. 현재 부호화 하는 슬라이스의 타입이 결정되면 슬라이스 타입에 따라 슬라이스 내부의 마크로 블록이 가질 수 있는 마크로블록 타입의 종류들이 결정된다. 예를 들어 P슬라이스의 경우 intra_16x16, intra_4x4, skip, inter_16x16, inter_16x8, inter8x16, inter_8x8들을 마크로 블록 타입으로 가질 수 있으며, inter_8x8의 경우 서브 마크로블록 타입을 가지게 된다. 그리고 각 타입에 대한 코드워드는 다음과 같이 할당이 된다.

표 1. P 슬라이스 마크로 블록의 코드워드

Mb_type	Name of mb_type	Code word
1	P_L0_16x16	0
2	P_L0_L0_16x8	1
3	P_L0_L0_8x16	2
4	P_8x8	3
5	P_8x8ref0	4
. . .	intra	+α

표 1에서 마크로블록 타입별로 할당된 코드워드는 표준화가 진행되면서 통계적으로 많이 나오는 타입에 대해 짧은 길이의 코드워드를 할당하는 방법으로 정해진 것이다. P 슬라이스의 경우 이전에 부호화되고 복호화 된 프레임들과의 중복성 제거를 통해 압축을 하기 때문에 움직임이 적고 변화가 거의 없는 영상의 경우나 양자화 파라미터 값이 클 경우에는 skip 블록의 개수가 많이 나오지만, 움직임이 많고 고주파 성분이 많은 영상의 경우 inter_8x8로 예측되는 블록이 많아질 수 있고, 화면전환이 되거나 조명 조건이 짧은 시간 수시로 달라지는 경우 intra로 예측되는 블록이 많아지기도 한다. 표 2는 이러한 예를 알아보기 위하여 foreman qcif (176x144) 영상에 대해 한 프레임 내에서의 각 마크로 블록 타입별 빈도 수를 조사한 것이다. 표 2의 결과는 inter frame 예측에 대한 빈도수만을 표시한 것으로 skip과 intra 예측에 대한 결과가 제외된 것이다. 표2 에서 나타난 결과를 보면 일부 프레임에서 inter_8x8로 예측되는 블록의 숫자가 다른 마크로 블록 타입으로 예측되는 마크로블록의 숫자보다 많은 것을 알 수 있다. 이 경우 코드워드를 표 1과 같이 할당하기 보다 inter_8x8에 가장 짧은 코드워드를 할당하게 되면 효율적인 부호화가 가능하게 된다.

표 2. foreman에 대한 마크로블록 타입 빈도 수

Fame Number	P16x16	P16x8	p8x16	p8x8	P8x8 (ref0)
1	21	14	11	0	34
2	20	7	12	17	2
3	23	6	13	14	2

3. 제안하는 마크로 블록 타입에 대해 적응적인 코드워드를 할당 방법

본 논문에서는 동영상 압축 시 슬라이스 단위로 영상의 특징과 부호화 조건에 따른 압축 데이터의 특징을 분석하여 구문 요소에 대한 최적의 코드워드를 결정한 후, 이를 슬라이스 헤더에 표기하여 전송한다. 제안된 방법에서 동영상 부호화기는 영상의 특징 및 부호화 조건에 따라 결정된 구문요소에 대한 코드워드 변경 정보를 슬라이스 헤더에 명시하고 이에 따라 비트스트림을 작성하며, 복호화기는 슬라이스 헤더에 표기된 구문요소에 대한 코드워드 변경 정보에 맞추어 복호화를 진행한다.

가. 제안한 방법에서의 부호화 과정

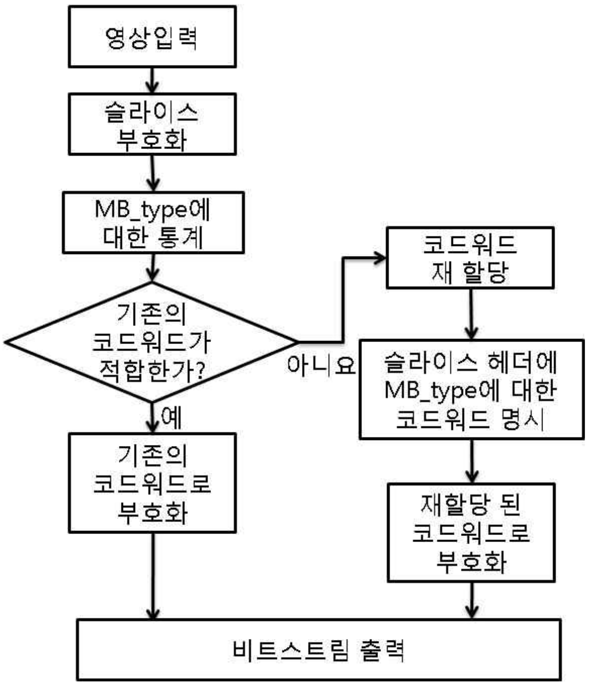

그림 1. 제안한 방법에 대한 부호화 과정

그림 1은 본 논문에서 제안하는 방법에 따라 마크로 블록의 발생 통계를 반영한 적응적인 코드워드 할당의 방법의 부호화 과정을 설명하기 위한 흐름도 이다. 제안된 방법은 입력된 영상에 대하여 기존의 부호화 방법을 통하여 슬라이스 단위로 부호화하고, 마크로 블록 타입에 대한 통계를 계산한다. 그리고 각 마크로블록 타입에 대하여 할당된 코드워드가 적합한지를 판단하게 되는데, 일반적인 엔트로피 부호화에서 효율적인 부호화를 위해서는 통계적으로 발생 빈도가 높은 요소에

대해서는 짧은 길이의 코드워드를 할당하고, 빈도가 낮은 요소에 대해서는 상대적으로 긴 길이의 코드워드를 할당한다. 이러한 규칙을 통해 각 마크로블록 타입들의 발생 빈도에 대한 통계가 할당된 코드워드에 일치하는지를 판단하는 단계가 기존의 코드워드에 대한 적합도를 판단하는 단계이다. 판단과정에서 마크로블록 타입에 대하여 기존의 코드워드가 적합하다고 판단이 되면 기존의 코드워드로 부호화하여 비트스트림을 생성하고, 적합하지 않다고 판단이 되면, 적합한 코드워드를 재할당하고 복호화기를 위해 이에 대한 정보를 슬라이스 헤더에 명시한다. 그리고 재할당된 코드워드를 이용하여 부호화하고 비트스트림을 생성한다.

표 3과 4는 슬라이스 헤더에 재할당 된 코드워드를 명시하기 위한 syntax 구조이다. 제안한 방법에서는 재할당 된 코드워드를 명시하기 위한 syntax 구조를 두 가지 제안하였다. 표 3은 해당슬라이스에서 기존의 코드워드에 대해 가장 비효율적으로 할당된 하나의 마크로블록 타입에 대한 코드워드만을 변경하는 방법이고, 표 4는 해당 슬라이스에 대하여 코드워드를 최적으로 할당하는 방법이다

표 3. 재 할당 코드워드를 위한 syntax_A

slice header ( )	**descriptor**
{	
syntax_reoder_flag	u(1)
if(syntax_reoder_flag) {	
reordering_value_num	ue(v)
reodering_size	ue(v)
}	
}	

표 3의 syntax_reorder_flag는 해당 슬라이스가 기존의 코드워드를 이용하여 부호화 되었는지 혹은 코드워드를 재할당하여 부호화 하였는지를 표시하기 위한 플래그이다. 플래그 값이 1이 되어 해당 슬라이스의 코드워드를 재할당한 경우 재 할당된 코드워드에 대한 세부 정보를 reordering_value_num과 reordering_size를 통해 표시한다. reordering_value_num은 가장 비효율적으로 코드워드가 할당된 마크로 블록타입의 번호를 뜻하고, reordering_size는 코드워드를 재 할당할 경우 재할당된 코드워드와 기존 코드워드와의 차이 값을 뜻한다.

표 4. 재 할당 코드워드를 위한 syntax_B

slice header ( )	**descriptor**
{	
syntax_reoder_flag	u(1)
if(syntax_reoder_flag) {	
sorting_order	ue(v)
}	
}	

표 4의 syntax_reorder_flag는 표 3과 같이 해당 슬라이스가 기존의 코드워드를 이용하여 부호화 되었는지 혹은 코드워드를 재할당하여 부호화 하였는지를 표시하기 위한 플래그이다. 플래그 값이 1이 되어 해당 슬라이스의 코드워드를 재할당한 경우 재할당 된 코드워드에 대한 세부 정보를 sorting_order를 통해 표시한다. B 방법에서는 마크로블록 타입에 대해 구해진 통계를 기반으로 마크로블록 타입을 정렬하고 이를 표기하였다.

나. 제안한 방법에서 복호화 과정

그림 2는 본 논문에서 제안한 방법에 대한 복호화 과정을 설명하기 위한 흐름도이다. 복호화기는 비트스트림을 입력받고, 슬라이스 특징 추출 단계에서 입력된 비트스트림의 슬라이스 헤더에서 마크로블록 타입에 대한 코드워드 재할당 관련 정보를 추출한다. 추출된 정보에서 전달된 비트스트림의 슬라이스가 기존의 코드워드로 부호화 되었는지를 판단하고, 기존의 코드워드로 부호화 된 비트스트림이면 표 1과 같이 마크로블록 타입에 대하여 기존의 코드워드를 이용하여 복호화한다. 현재 슬라이스의 마크로블록 타입이 기존의 코드워드가 아닌 새로운 코드워드로 부호화 되었다면, 그 정보를 추출하고 이를 이용하여 마크로블록 타입을 복호화하면 된다.

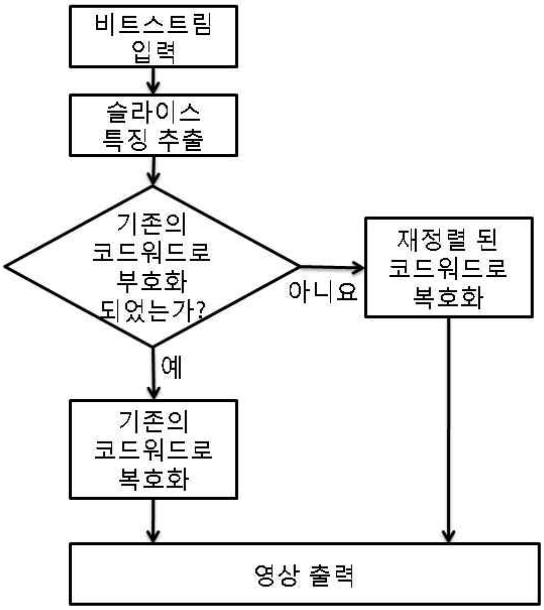

그림 2. 제안한 방법에 대한 복호화 과정

4. 실험결과

본 논문에서 제안한 방법에 대한 성능평가는 표 5의 실험환경을 통해 이루어졌다. 제안한 방법은 KTA 참조 소프트웨어 2.2 버전[2]에 구현하여 실험하였으며, qcif와 720p 영상에 대하여 실험을 진행하였다. 표 5에 나타나지 않은 실험환경은 VCEG (Video Coding Experts Group)에서 권장하는 실험환경에 근거하여 실험을 진행하였다.

표 5. 실험환경

Codec		KTA reference software (version 2.2)
qcif	**Sequences**	Foreman, News, Stephen
	Size	176 x 144
	Number of frames	300
	Frame rate	30Hz
	Intra period	1 (I, P, P, P...)
720p	**Sequences**	BigShips, City_corr, Crew
	Size	1280 x 720
	Number of frames	150
	Frame rate	60Hz
	Intra period	1 (I, P, P, P...)

표 5는 제안한 방법에 대하여 syntax_A를 사용한 경우의 qcif 영상에 대하여 실험한 결과이고, 표 6은 720p 영상에 대하여 실험한 결과이다. 실험결과 qcif 영상에 대해서는 0.2-1.0 %의 BD-bitrate[3] 이득이 있었고, 720p 영상에 대해서는 0.9-1.7%의 BD-bitrate 이득이 있었다. 실험결과에서 제안한 방법이 bitrate 측면에서 효과가 있었던 부분은 작은 크기의 영상에서는 서브 마크로블록 타입을 가지는 경우와 intra 블록으로 예측되는 경우이고, 큰 크기의 영상에서는 intra 마크로블록으로 예측이 되는 경우로 표 1에서 보듯이 intra로 예측되는 블록은 inter로 예측되는 경우에 비하여 상대적으로 긴 길이의 코드워드가 할당되기 때문에 intra 예측 블록이 많은 경우 기존의 방법에서는 비효율적으로 코드워드가 할당되게 되어 제안한 방법이 효과를 가지게 된다.

표 5. syntax_A를 이용한 실험결과 (qcif 영상)

Sequence	QP	Conventional		Proposed method		BDPSNR(dB)	BDBitrate(%)
		Bitrate	PSNR	Bitrate	PSNR		
Foreman. qcif	I=22, P=23	1368642	41.949	1367652	42.014	0.052	-0.664
	I=27, P=28	832578	37.913	832173	37.918		
	I=32, P=33	489755	34.127	488851	34.195		
	I=37, P=38	297851	30.764	295877	30.795		
News. qcif	I=22, P=23	1333132	43.237	1289564	43.316	0.102	-1.030
	I=27, P=28	897596	39.04	896658	39.093		
	I=32, P=33	596574	35.01	594787	35.024		
	I=37, P=38	384706	31.214	384512	31.296		
Stephen. qcif	I=22, P=23	2365898	42.393	2362458	42.42	0.026	-0.246
	I=27, P=28	1673394	37.55	1669875	37.556		
	I=32, P=33	1106017	32.691	1104985	32.698		
	I=37, P=38	669942	28.351	669098	28.369		

표 6. syntax_A를 이용한 실험결과 (720p)

Sequence	QP	Conventional		Proposed method		BDPSNR(dB)	BDBitrate(%)
		Bitrate	PSNR	Bitrate	PSNR		
BigShips	I=22, P=23	7787082	39.863	7755523	39.882	0.027	-0.929
	I=27, P=28	2333979	36.561	2328131	36.573		
	I=32, P=33	843787	33.821	841504	33.848		
	I=37, P=38	414792	31.481	412686	31.504		
City_corr	I=22, P=23	13037337	40.051	13025682	40.058	0.035	-1.120
	I=27, P=28	3073463	36.229	3068077	36.23		
	I=32, P=33	1089785	33.211	1079349	33.264		
	I=37, P=38	568773	30.449	564808	30.565		
Crew	I=22, P=23	6586560	41.565	6417232	41.66	0.058	-1.739
	I=27, P=28	2169314	39.227	2119848	39.228		
	I=32, P=33	992798	37.04	989158	37.007		
	I=37, P=38	546571	34.686	538242	34.668		

표 7과 8은 제안한 방법에 대하여 syntax_B를 이용한 경우에 대한 실험결과이다. 실험결과 qcif 영상에 대해서는 0.3 - 1.3 %의 BD-bitrate 이득이 있었고, 720p 영상에 대해서는 1.4 - 1.9 %의 BD-bitrate 이득이 있었다. syntax_B를 이용한 경우 최적의 한 슬라이스에 대하여 최적의 코드워드를 구하고, 이를 이용하여 마크로블록 타입에 대한 부호화를 진행하였기 때문에 syntax_A에 대하여 약간의 BD-bitrate 이득을 얻을 수 있었다.

5. 결론

본 논문에서는 마크로블록 타입에 대한 코드워드를 슬라이스 단위로 적응적으로 할당하는 방법을 통해 마크로블록 타입에 대한 부호화를 효율적으로 하는 방법을 제안하였다. 최적의 코드워드를 할당하기 위해 부호화는 2-pass로 진행하였으며, 각 슬라이스 별 최적의 코드워드는 슬라이스 헤더에 2가지 방법을 통해 표기하였다. 제안한 방법에 대한 실험결과 약 0.2-1.9%의 BD-bitrate 이득을 얻을 수 있었다. 현재 제안한 방법은 P슬라이스의 마크로블록 타입에 대해서만 진행을 하였으므로, 앞으로 B슬라이스에 대한 마크로 블록 타입에 대한 연구와 슬라이스 헤더에 코드워드를 표기하는 방법을 좀 더 효율적으로 표기할 수 있는 방법에 대한 연구를 진행할 예정이다.

표 7. syntax_B를 이용한 실험결과 (qcif 영상)

Sequence	QP	Conventional		Proposed method		BDPSNR(dB)	BDBitrate(%)
		Bitrate	PSNR	Bitrate	PSNR		
Foreman. qcif	I=22, P=23	1368642	41.949	1367342	42.015	0.06	-0.76
	I=27, P=28	832578	37.913	831813	37.921		
	I=32, P=33	489755	34.127	488501	34.199		
	I=37, P=38	297851	30.764	295532	30.796		
News. qcif	I=22, P=23	1333132	43.237	1289204	43.326	0.14	-1.37
	I=27, P=28	897596	39.04	896199	39.123		
	I=32, P=33	596574	35.01	594425	35.067		
	I=37, P=38	384706	31.214	384179	31.3		
Stephen. qcif	I=22, P=23	2365898	42.393	2362078	42.445	0.04	-0.36
	I=27, P=28	1673394	37.55	1669535	37.564		
	I=32, P=33	1106017	32.691	1104635	32.71		
	I=37, P=38	669942	28.351	668745	28.371		

표 8. syntax_B를 이용한 실험결과 (720p)

Sequence	QP	Conventional		Proposed method		BDPSNR(dB)	BDBitrate(%)
		Bitrate	PSNR	Bitrate	PSNR		
BigShips	I=22, P=23	7787082	39.863	7755203	39.91	0.03	-1.17
	I=27, P=28	2333979	36.561	2327780	36.573		
	I=32, P=33	843787	33.821	841175	33.856		
	I=37, P=38	414792	31.481	412323	31.51		
City_corr	I=22, P=23	13037337	40.051	13025312	40.105	0.05	-1.45
	I=27, P=28	3073463	36.229	3067731	36.23		
	I=32, P=33	1089785	33.211	1079001	33.272		
	I=37, P=38	568773	30.449	564467	30.577		
Crew	I=22, P=23	6586560	41.565	6416879	41.681	0.06	-1.97
	I=27, P=28	2169314	39.227	2119497	39.231		
	I=32, P=33	992798	37.04	988798	37.009		
	I=37, P=38	546571	34.686	537815	34.67		

감사의 글

본 연구는 방송통신위원회, 지식경제부 및 한국산업기술평가관리원의 IT 원천기술개발사업의 일환으로 수행한 연구로부터 도출된 것이다. [과제관리번호:2008-F-011, 과제명 : 차세대 DTV 핵심 기술 개발]

참고문헌

[1] Katsuyuki Nakamura, Masashi Takahashi, Toru Yokoyama, Kazushi Akie, Seiji Mochizuki, and Kenichi Iwata, "An Extended VLC for Coding Modes," ITU-T Q.6/SG16 VCEG Meeting, VCEG-AJ15, Oct. 2008

[2] KTA Software, http://iphome.hhi.de/suehring/tml/download/KTA/.

[3] Gisle Bjøntegaard, "Improvements of the BD-PSNR model," ITU-T Q.6/SG16 Document, VCEG-AI11, July 2008.