

MPEG-4 MAC(Multiple Auxiliary Component) 기반

스테레오스코픽 비디오 부호화

조숙희, 윤국진, 안충현
한국전자통신연구원 방송시스템연구부

Stereoscopic Video Coding Using MPEG-4 Multiple Auxiliary Component

Sukhee Cho, Kugjin Yun, and Chunghyun Ahan
Radio & Broadcasting Research Department
Electronics and Telecommunications Research Institute
E-mail : shee@etri.re.kr

Abstract

We propose stereoscopic video coding method using the syntax of MAC(Multiple auxiliary component) that was added to the MPEG-4 visual version 2 in order to describe the transparency of the video object. We also define the novel MAC's semantics in MPEG-4 that should support the proposed coding method. The major difference between the existing coding method and the proposed coding method is the addition of the residual texture coding. The proposed coding method assigns disparity map and residual texture to 3 components of MAC: one component for disparity map and the rest 2 components for the luminance and chrominance data of the residual texture, respectively. The performance of the proposed method is evaluated in terms of PSNR by computer simulations.

I. 서론

스테레오스코픽 비디오는 좌우 양안에 각각 대응되는 2 장의 좌우 영상으로 구성된다. 좌우 영상은 같은 장면에 대하여 일반적으로 65mm 정도의 간격을 두고 획득되므로, 좌우 영상은 서로 높은 상관정보를 가진다 [2]. 스테레오스코픽 비디오는 좌우영상의 상관정보를 이용하지 않고 부호화한다면, 일반 2 차원 비디오 부호화 방식에 의해 좌우 영상이 각각 독립적으로 부호화될

수 있다. 그러나, 좌우 영상의 상관 정보인 변이정보를 이용하지 않고 부호화하는 것은 부호화 효율이 저하된다. 따라서, 좌우 영상사이의 상관정보를 이용하여 공간 중복성(redundancies)의 최소화를 목표로 하는 스테레오스코픽 비디오 부호화 알고리즘에 대한 연구가 많이 제안되어 왔다. MPEG-2 MVP(Multi-View Profile)과 MPEG-4의 TS(Temporal Scalability)를 이용한 부호화가 대표적인 스테레오스코픽 비디오 부호화 방식이다[1][2].

TS 를 이용한 부호화 방식은 실제 H/W 로 구현하여 사용되는 경우가 매우 적다. 왜냐하면, TS 기반의 부호화 구조는 매우 복잡하여 구현이 어려우며, 2 개의 층(layer)에서 좌우 부호화 스트림 각각 출력되므로, 이들 2 개 부호화 스트림의 프레임 동기 맞추기 위해서는 다중화/역다중화 시스템이 별도로 필요하기 때문이다. 반면, 비디오 코덱으로부터 한 개의 부호화 스트림이 출력된다고 하면, 송수신을 포함하는 전체 시스템 구성 측면에서 보다 간략하게 스테레오스코픽 시스템 구축이 가능할 것이다. 따라서, 스테레오스코픽 비디오를 부호화하여 한 개의 부호화 스트림을 출력할 수 있는 구조를 갖는 MPEG-4 MAC(Multiple-Auxiliary Components)을 이용하여 부호화하는 방식을 제안한다.

II. TS 를 이용한 스테레오스코픽 비디오 부호화

TS 를 이용한 부호화 방식은 스테레오스코픽 비디오 부호화의 가장 대표적인 방식으로 기저층(base-layer)과 상위층(enhancement-layer)으로 구성하여, 좌영상은 기저층에서, 우영상은 상위층에서 부호화한다. 좌영상은 기존의 2 차원 비디오 부호화 방식과 동일한 방식으로 블록단위의 움직임 보상만을 이용하여 부호화하며, 우영상은 블록단위의 움직임 및 변이 보상을 이용하여 부호화한다. TS 를 이용한 부호화 방식에서 그림 1 은 여러 형태의 예측방식 중 전형적인 형태의 하나이다. 좌영상은 기저층에서 움직임 보상만을 수행하여 I-, P-, B-picture 로서 부호화하고, 우영상은 상위층에서 움직임 및 변이 보상을 수행하여 P-와 B-picture 로서 부호화한다.

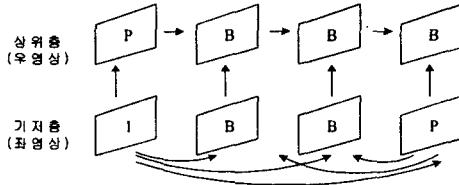


그림 1. 움직임 및 변이 예측 방식의 일 예.

III. MPEG-4 MAC

MAC 은 비디오 객체의 투명도(transparency)를 나타내기 위하여 MPEG-4 Visual Part 14496-2 의 버전 2 에 추가되었다[2]. 현재 MAC 신택스(Syntax)에 대한 시맨틱스(Semantics) 에는, VOP 로부터 다른 시점의 영상을 생성하는데 사용하기 위한 정보로서 변이맵(Disparity map)에 대한 정의가 포함되어 있다. 이러한 MAC 을 이용한 스테레오스코픽 비디오 부호화는 변이정보를 이용하여 부호화할 수 있으며, 한 개의 부호화 스트림으로 출력되는 구조를 갖는다. 그러나, 기존의 MAC 은 변이 보상된 영상과 원 영상과의 차영상을 처리할 수 있는 신택스와 시맨틱스가 정의되어 있지 않다. 반면, 실제 촬영된 스테레오스코픽 비디오는 차영상 데이터에 대한 정보 없이 변이보상만 이루어진 복원 영상에 대해서는 고화질을 보장할 수 없다는 문제점이 있다.

MAC 은 화소단위에서 VOP(Video Object Plane)에 대하여 정의되어 있으며, 알파정보, 변이정보, 깊이정보와 같이 비디오 객체와 관련된 데이터를 포함한다. 기

존 MAC 은 3 개의 보조요소(Auxiliary components)로 구성되어 있으며, 표 1 과 같은 형태의 정보들이 보조요소 정보로서 정의되어 있다.

표 1. VIDEO_OBJECT_LAYER_SHAPE_EXTENSION 의 시맨틱스 의미

video_object_layer_shape_ext.	aux_comp_type[0]	aux_comp_type[1]	aux_comp_type[2]	aux_comp_count
0000	ALPHA	NO	NO	1
0001	DISPARITY	NO	NO	1
0010	ALPHA	DISPARITY	NO	2
0011	DISPARITY	DISPARITY	NO	2
0100	ALPHA	DISPARITY	DISPARITY	3
0101	DEPTH	NO	NO	1
0110	ALPHA	DEPTH	NO	2
0111	TEXTURE	NO	NO	1
1000	USER DEFINED	NO	NO	1
1001	USER DEFINED	USER DEFINED	NO	2
1010	USER DEFINED	USER DEFINED	USER DEFINED	3
1011	ALPHA	USER DEFINED	NO	2
1100	ALPHA	USER DEFINED	USER DEFINED	3
1101-1111	t.b.d.	t.b.d.	t.b.d.	t.b.d.

3.1 MAC 을 이용한 기존 부호화 방식

표 1 에서 video_object_layer_shape_extension 변수가 '0001'~'0100'사이의 값일 때, 1 개 이상의 보조요소가 'Disparity' 정보를 가질 수 있다. 스테레오스코픽 비디오는 'Disparity'의 타입을 갖는 보조요소에 화소단위의 변이맵을 할당하여 우영상을 부호화할 수 있다. 그림 2 는 우영상에 대하여 변이맵을 갖는 MAC 을 이용한 기존 부호화 방식을 나타낸다.

좌영상은 VOP 부호화에 의하여 기존 2 차원 부호화 방식과 동일하게 그림 3 과 같이 움직임 보상의 DCT 기반 부호화를 수행한다. 화소단위로 구현된 변이맵은 MAC 의 한 보조요소에 할당하여, 그림 4 의 AC(Auxiliary Components) 부호화에 의해 움직임 보상의 DCT 기반 부호화를 수행한다. 이 때, AC 부호화는 VOP 부호화 과정과는 달리 움직임 벡터 예측을 수행하지 않고, VOP 부호화 과정에서 구현된 움직임 벡터를 그대로 이용하여 움직임 보상을 수행한다.

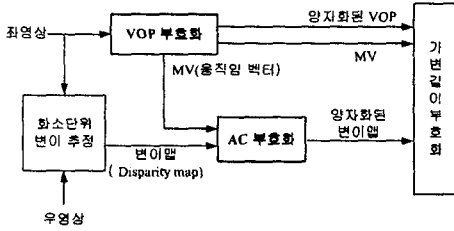


그림 2. 기존 MAC 을 이용한 부호화 구조

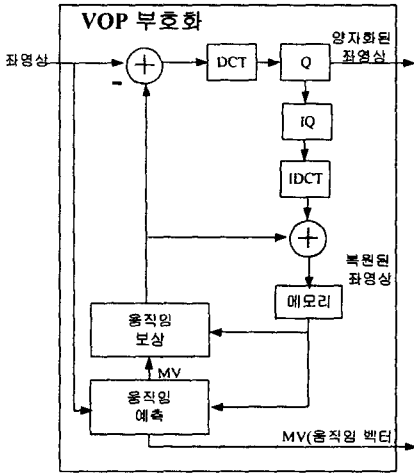


그림 3. VOP 부호화 과정.

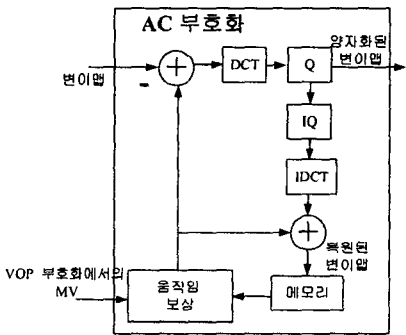


그림 4. AC 부호화 과정.

IV. MPEG-4 MAC 의 확장 및 제안 부호화 방식

기존의 MAC 이 변이정보를 수용하여 부호화할 수 있는 구조를 갖는다고 하지만, 실제 촬영에 의해 획득된 스테레오스코픽 비디오 부호화에 대해서는 한계가

있다. 특히, 변이 보상된 영상과 원영상과의 차영상 데이터를 수용할 수 있는 syntax 와 semantics 가 정의되어 있지 않다. 그러나, 차영상 데이터없이 변이맵만으로 복원된 우영상의 화질은 상당히 떨어진다.

따라서, 본 논문에서는 MPEG-4 MAC 을 이용하여 스테레오스코픽 비디오의 부호화를 효율적으로 수행하기 위하여 필요한 새로운 시맨틱스(semantics)를 정의하여, MAC 을 확장한다. 또한, 확장된 MAC 을 이용하여 스테레오스코픽 비디오를 부호화하는 방식을 제안한다.

표 2. VIDEO_OBJECT_LAYER_SHAPE_EXTENSION 의 새롭게 정의된 시맨틱스 의미.

video_object_layer_shape_ext.	aux_comp_type[0]	aux_comp_type[1]	aux_comp_type[2]	aux_comp_count
1101	DISPARITY	Luminance Residual Texture	NO	2
1110	DISPARITY	Luminance Residual Texture	Chrominance Residual Texture	3

4.1 MAC 의 확장

현재 MPEG-4 에서는 video_object_layer_shape_extension 변수가 '1101'~'1111'사이의 값일 때는 모든 보조요소에 대한 정보 종류를 정의하지 않고 있다. 본 논문에서는 스테레오스코픽 비디오를 효율적으로 부호화하기 위하여 표 2 와 같이 MAC 에 새로운 semantics 를 정의한다.

4.1 확장된 MAC 을 이용한 제안 부호화 방식

그림 5 는 확장된 MAC 을 이용하여 제안하는 스테레오스코픽 비디오 부호화 방식을 나타낸다. 제안방식과 기존방식과의 가장 큰 차이점은 변이보상된 영상과 원영상과의 차영상에 대한 부호화이다.

제안방식의 부호화는 우영상에 대한 데이터로서 변이맵과 차영상 데이터를 MAC 의 3 개 보조요소에 할당한다. 변이맵은 첫번째 AC 에, 휘도성분에 대한 차영상 데이터는 두번째 AC 에, 컬러성분에 대한 차영상 데이터는 세번째 AC 에 할당하여 그림 5 의 과정에 의해 부호화된다.

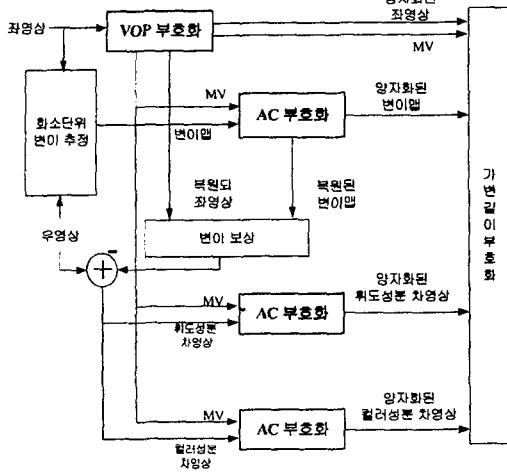


그림 5. 제안 부호화 방식의 구조.

V. 실험 결과

본 논문에서는 제안방식을 포함하여 아래의 5 가지의 부호화 방식에 대하여 시뮬레이션 결과를 비교 분석하였다. 각 부호화 방식에 대하여 좌영상에 대한 부호화 방식은 동일하며, 우영상에 대한 부호화 방식만 다르다. 방식 1은 기존 MAC의 이용한 부호화 방식으로 우영상에 대한 정보로서 변이맵을 첫번째 AC[0]에 할당하여 부호화하였다. 방식 2-1은 확장된 MAC을 이용한 제안 부호화 방식으로 변이맵을 첫번째 AC[0]에 할당하고, 휘도성분의 차영상 데이터를 두번째 AC[1]에 할당하여 부호화하였다. 방식 2-2는 확장된 MAC을 이용한 제안 부호화 방식으로 변이맵을 첫번째 AC[0]에 할당하고, 휘도성분 및 컬러성분의 차영상 데이터를 두번째 AC[1] 및 세번째 AC[2]에 각각 할당하여 부호화하였다. 방식 3은 좌우 영상을 각각 독립적으로 수행한 MPEG-4 부호화 방식을 적용하였다. 마지막으로, 방식 4는 TS를 이용한 MPEG-4 부호화를 수행하였다.

테스트 시퀀스는 720x480 해상도를 가지며 YcbCr 4:2:0 포맷을 갖는 100 프레임의 좌우 영상을 사용하였다. 'puppy' 시퀀스는 한 개의 객체에 움직임이 적으며, 'soccer2'는 움직임 많은 시퀀스이다. 위에서 언급한 5 가지의 모든 부호화에서 좌영상 부호화는 고정 QP 값, $QP(I,P,B)=(4,8,12)$ 로 설정하여 부호화하여 동일한 PSNR 값을 갖도록 하고, 우영상의 PSNR 값에 의하여 성능을 비교 평가하였다.

그림 6에 나타낸 바와 같이 2 가지 테스트 시퀀스에 대하여 제안 부호화 방식 2-1 및 방식 2-2는 기존 MAC을 이용한 부호화 방식인 방식 1보다는 예상대로 4dB~7dB 높은 PSNR 값을, TS를 이용한 부호화 방식보다는 약간 낮은 PSNR 값을 얻었다.

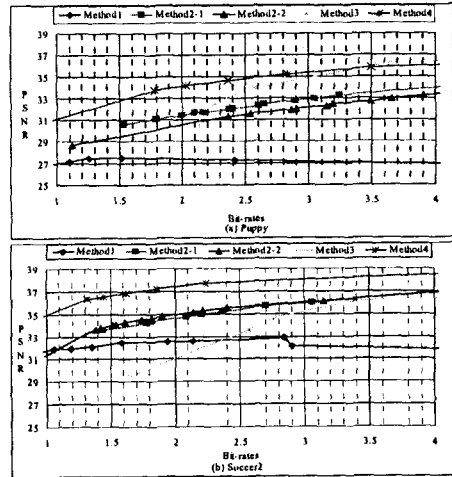


그림 6. 좌영상이 $QP(I,P,B)=(4,8,12)$ 로 설정되어 부호화된 경우, 복원된 우영상에 대한 PSNR

VI. 결론

본 논문에서의 제안 부호화 방식은 TS를 이용하여 부호화하는 전형적인 부호화 방식과는 달리 한 개의 부호화 스트림 형태로 출력하므로, 두 영상을 프레임단위로 다중화 및 역다중화하는 시스템이 필요없다는 장점을 가진다. 향후, MAC의 각 보조요소에 대한 최적의 비트율 할당 및 TS를 이용하는 부호화 방식보다 부호화 효율을 향상시킬 수 있는 알고리즘에 대한 연구가 수행될 것이다.

참고문헌

- [1] Naito S. and Matsumoto H., "Advanced rate control technologies for 3D-HDTV digital coding based on MPEG-2 Multi-View Profile," Proceedings 1999 International Conference on Image Processing, IEEE, vol.1, pp.281-285, Piscataway, USA.
- [2] Generic Coding of Audio-Visual Objects - Part 2 : Visual, ISO/IEC 14496-2 : 2001.