

다시점 비디오 부호화 기술 동향

호요성 · 오관정(광주과학기술원 실감방송연구센터)

1. 서론

우리가 일상생활에서 즐기고 있는 디지털 시대가 도래된 이후, 다양한 멀티미디어 기술들이 매우 빨리 발전하고 있으며, 이를 바탕으로 디지털 콘텐츠 시장도 해마다 그 규모가 급속도로 성장하고 있다. 또한 인터넷을 통한 데이터 전달 기술이 급격히 향상되어 전송 속도도 현격하게 증가하여 다양한 멀티미디어 콘텐츠를 손쉽게 이용할 수 있는 기폭제 역할을 하고 있다. 더구나 최근에는 사용자들의 요구가 단순한 멀티미디어 콘텐츠의 소비 형태에서 대화형 콘텐츠와 실감 콘텐츠의 이용을 요구하는 수준에 이르고 있다. 이러한 흐름 속에 디지털 영상 분야에서는 최근 실감 영상에 대한 관심이 크게 높아지고 있다. 실감영상이란 3차원 입체감이 좀더 사실적으로 느껴지는 영상을 일컫는데, 최근에는 이러한 실감영상을 위한 기술로 양안시점 입체영상 (stereoscopic image), 다시점 비디오 (multi-view video), 홀로그래피 (holography) 등과 같은 3차원 영상

과 비디오 기술에 대한 연구가 활발히 진행되고 있다. 여기서 다시점 비디오란 두 대 이상의 카메라를 이용하여 촬영한 영상물로 여러 방식을 통해 단일 시점에 비해 다양하고 입체감있는 영상을 사용자에게 제공하는 3차원 영상처리의 중요한 주제이다¹⁾.

다시점 영상의 한 예인 파노라믹 (panoramic) 영상은 우주 항공 사진학, 컴퓨터 비전, 영상처리, 컴퓨터 그래픽스 분야에서 매우 흥미로운 연구 주제로 자리잡고 있으며, 이러한 파노라믹 영상 기술은 항공사진의 해석, 영상변화감지, 비디오 압축, 비디오 인덱싱, 카메라 해상도 및 시계 (field of view, FOV) 확대에서 간단한 영상 편집에 이르기까지 매우 다양한 분야에서 응용되고 있다. 또한 컴퓨터 비전에서는 다시점에서 획득된 영상을 이용하여 영상 내의 물체의 깊이(depth)와 시차(disparity) 정보를 파악하기 위한 연구를 수행하고 있으며, 컴퓨터 그래픽스에서는 영상기반 렌더링 (image based rendering, IBR)이란 기술로 획득된 다시점 영상들을 바탕으로 사실적인 가상

의 시점 영상을 생성하기 위해 노력하고 있다. 이러한 연구의 성과로, 현재 다시점 비디오 처리 기술은 전방향 카메라를 이용한 감시 시스템이나 게임에서 이용되는 3차원 가상 시점과 N개의 카메라 영상으로부터 입력된 임의의 영상을 선택할 수 있도록 하는 시점 스위칭 등으로 적용되고 있다. 또한 영화나 광고 영상의 경우에도 다양하게 응용되고 있다^[2, 3].

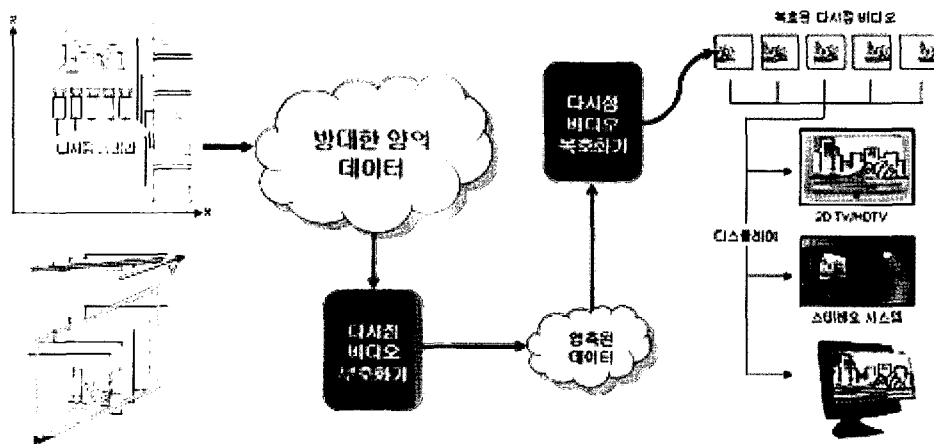
그러나 다시점 비디오는 사용자들에게 자유로운 시점과 넓은 화면을 통한 입체감을 제공하는 장점을 지니고 있지만, 영상 획득시에 필수적으로 존재하는 중복 조건, 순간적으로 처리해야 할 데이터 양의 증가, 고가의 장비가 요구되는 문제점으로 인해 다양한 서비스 개발이 제한되었다. 따라서 다시점 비디오를 효율적으로 처리하는 다양한 기술들을 개발하는 것이 매우 필요하다. 이에 발맞추어 멀티미디어 콘텐츠에 대한 표준 규격을 제정하는 국제 기구인 MPEG (moving picture experts group)에서는 2002년부터 3차원 오디오 비주얼 (3-dimensional audio-visual, 3DAV)이란 이름으로 3차원 영상처리에 관한 기술을 표준화

하는 작업을 진행하고 있다. 이후, 3DAV에서 다시점 비디오 부호화에 좀더 집중하여 MVC(multi-view video coding)란 이름으로 변경되었고, 2006년 7월에 열린 제77차 MPEG 회의에서부터 MVC에 관한 표준화 작업을 JVT(joint video team)에서 진행하게 되었다. 최근에는 MPEG에서 FTV(free-viewpoint television)에 대한 AHG(ad hoc group)을 만들어 다시점 비디오와 깊이 영상을 기반으로 FTV에 대한 표준화 작업을 준비하고 있다. 본 논문에서는 다시점 비디오 부호화에 대한 독자들의 이해를 돕기 위해 다시점 비디오 부호화에 대한 일반적인 내용을 소개하고, 최근에 연구되고 있는 관련 기술들을 살펴본다.

II. 다시점 비디오에 대한 소개

1. 다시점 비디오 시스템

다시점 비디오 시스템은 여러 대의 카메라를 이용하여 촬영된 다시점 영상을 획득하고



〈그림 1〉 다시점 비디오 시스템

전송하고, 다양한 디스플레이 기술을 이용하여 재생하는 시스템이다. 다시점 영상은 사용자에게 임의시점의 영상을 시청할 수 있게 하며, 여러 시점의 영상을 합성하여 보다 넓은 범위에서의 화면을 제공할 수 있다. 또한 인접한 시점의 두 영상을 스테레오 모니터나 다시점 영상 디스플레이 장치를 통해 디스플레이 하여 사용자에게 입체감있는 3차원 영상을 시청할 수 있게 한다. 이러한 다시점 비디오 시스템이 상용화되기 위해서는 방대한 양의 다시점 비디오 데이터를 효율적으로 압축하는 다시점 비디오 부호화 기술의 개발이 필요하다. <그림 1>은 일반적인 다시점 비디오 시스템에 대한 개념을 설명하고 있다.

2. 다시점 비디오 응용 분야

가. 자유시점 비디오와 자유시점 TV

자유시점 TV는 조밀한 간격으로 배치된 여러 대의 카메라로 3차원 장면을 촬영하고, 이를 바탕으로 광선 공간(ray space)를 정의한다. 이렇게 정의한 광선 공간을 이용하면 사용자의 요청에 따라 임의의 시점 영상을 자유로이 생성할 수 있다. 최근 일본에서는 자유 시점 TV에 대한 집중적인 연구를 진행하고 있는데, 이러한 다시점 비디오 시스템은 다음과 같은 분야에서 활용될 수 있다.

- 엔터테인먼트 : 콘서트, 스포츠 방송, 다중 사용자 게임, 영화, 드라마, 뉴스
- 교육 : 문화, 운동 종목의 교육, 의학 수술 과정
- 관광 : 동물원, 수족관, 식물원, 박물관
- 감시 시스템 : 교통 통제, 지하 주차장, 은행

- 기록 보관 : 공간 기록, 천연 보물의 기록, 전통 문화

나. 3차원 TV

3차원 TV는 스테레오로 제공되는 입체 영상의 확장이라고 쉽게 생각할 수 있다. 현재의 입체 영상은 스테레오를 기반으로 하나의 시점에 대한 입체 영상을 제공하지만, 우리가 연구하고 있는 3차원 TV는 다시점 비디오를 이용하여 다양한 위치에서의 입체 영상을 제공할 수 있어야 한다. 3차원 TV의 입력 영상인 다시점 비디오를 다시점 비디오 부호화기를 이용하여 압축하여 전송하고, 채널을 통해 전송된 비트스트림은 수신단에서 다시점 비디오 복호화기로 복원된다. 재구성된 다시점 비디오는 재생 장치의 특성에 따라 다양하게 변형되거나 합성되어 재생된다. 재생 장치의 예로는 2차원 디지털 TV, 스테레오 입체 TV, 다시점 3차원 TV 등이 있다. 3차원 TV는 다양한 시점의 입체영상을 지원하므로 생성해야 할 시점의 수가 촬영된 시점의 영상 수보다 많은 경우를 대비하여 중간 시점의 영상 보간(intermediate view reconstruction) 기술이 필요하다^[4].

다. 몰입형 원격회의

몰입형 원격회의란 단일 시점 비디오를 기반으로 한 현재의 원격회의 시스템에 다시점 비디오를 적용하여 보다 입체감있는 영상을 이용하는 원격회의를 일컫는다. 몰입형 원격회의를 통해 보다 현실감 있는 원격회의가 가능해진다^[4].

III. 다시점 비디오 부호화 표준화 작업

1. 실험 영상과 부호화 조건

현재 다시점 비디오 부호화 표준화 작업에서 권고하고 있는 실험 영상은 사용된 카메라의 개수, 카메라 배치, 초당 프레임수, 영상의 해상도, 장면의 복잡도, 카메라의 움직임 등을 고려하여 선정되었다. 또한, 선택된 실험 영상은 모두 카메라 변수를 포함하고 있으며, 다시점 비디오 표준화 그룹 내에서 이에 대한 검증 작업을 완료했다. 현재 표준화 그룹에서 제공하는 다양한 형태의 다시점 비디오 실험 영상의 특성을 <표 1>에 정리했다^[5].

<표 1>에서 보는 바와 같이, 카메라의 개수는 5대, 8대, 100대가 사용되었으며, 영상은 VGA (640×480) 크기와 XVGA (1024×768) 크기가, 초당 프레임 수는 15, 25, 30이 사용되었다. 또한, 카메라 배치는 1차원 평행, 1차원 수렴, 1차원 원호, 2차원 평행, 2차원 배열이 사용되었다. 모든 실험 영상은 YUV 4:2:0 형식으로 표현된다. 이때 영상 교정(rectification)이란 여러 대의 카메라를 이용하여 획득된 다

시점 영상들이 3차원 기하학 변환을 통해 정확하지 않은 카메라 구성으로 인해 발생하는 영상의 왜곡을 보정하는 작업을 의미한다.

이러한 실험 영상들은 다음 장에서 소개될 참조 소프트웨어와 부호화 조건^[5]을 통해 해당 부호화 알고리즘의 성능을 평가하는데 사용된다. 각 실험 영상들은 이를 촬영한 기관을 통해 인터넷으로 다운받을 수 있으며, 'Breakdancers' 영상의 경우에는 깊이 영상도 함께 제공된다^[5].

2. 다시점 비디오 부호화의 참조 소프트웨어

2005년 10월에 다시점 비디오 부호화에 대한 참조 소프트웨어로 Fraunhofer-HHI의 알고리즘이 선정되었으며, JSVM3.5 (joint scalable video model)을 기반으로 한 참조 소프트웨어가 2006년 2월에 배포되었다. 이어서 2006년 7월에 다시점 비디오의 병렬처리를 위하여 각 시점별로 독립적으로 부호화 하는 JMVM1.0 (joint multiview video model)이 배포된 이래, 2007년 5월에는 JMVM4.0이 배포되었다. 기본적으로 시점 방향에 대해서는

<표 1> 다시점 비디오 실험 영상의 특성

실험 영상	영상 특성	교정 (rectification)	카메라 개수	카메라 구성
Ballroom	VGA, 25fps	교정함	8	1차원 평행, 20cm 간격
Exit	VGA, 25fps	교정함	8	1차원 평행, 20cm 간격
Race1	VGA, 30fps	교정하지 않음	8	1차원 평행, 20cm 간격
Flamenco2	VGA, 30fps	교정하지 않음	5	2차원 평행, 직교, 20cm 간격
Uli	XVGA, 25fps	교정하지 않음	8	1차원 평행, 수렴 (convergent), 20cm 간격
Breakdancers	XVGA, 25fps	교정하지 않음	8	1차원 원호, 20cm 간격
Rena	VGA, 30fps	교정함	100	1차원 평행, 5cm 간격
Akko&Kayo	VGA, 30fps	교정하지 않음	100	2차원 배열, 수평 5cm 간격, 수직 20cm 간격

IBPBP... 구조를 시간 방향에 대해서는 계층적 B화면 구조를 사용하고 있다. <그림 2>는 다시점 비디오 부호화 참조 소프트웨어의 시공간 예측 구조를 보여주고 있다.

<그림 2>에서 S_n 은 n 번째 시점의 카메라를 의미하고, T_n 은 시간적으로 n 번째 화면을 의미한다. 화살표는 해당 화면을 예측하는데 참조되는 화면들 사이의 관계를 나타낸다. 시점 방향으로는 IBPBP 구조를 이용하고, 시간 방향으로는 계층적 B화면 구조를 이용한다. 계층적 B화면 구조란 <그림 2>에서와 같이 하나의 GOP 구조에서 GOP의 양 끝 화면(B_n)을 제외하고 중간 화면들은 모두 B화면으로 이루어진 GOP이고, 해당 화면들에 대해 먼저 해당 GOP의 중간에 위치한 화면(B_{n+1})이 양 끝 화면을 통해 부호화되고, 다시 해당 중간 화면과 양 끝 화면 사이의 화면들(B_{n+2})이 전후의 부호화된 화면들로부터 계층적으로 부

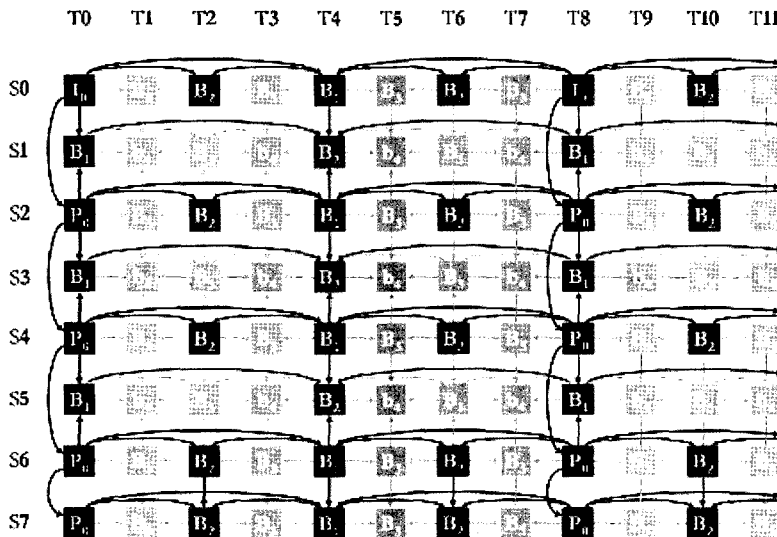
호화되는 GOP 구조를 말한다. 현재 시공간적 예측 구조에 대한 부호화 순서는 각 시점을 따로 따로 부호화한다. <그림 2>의 경우 I시점인 S_0 를 처음 부호화하고, S_0 의 복원 영상을 이용하여 P시점인 S_2 를 부호화한다. 마찬가지로, S_0 와 S_2 의 복원 영상을 이용하여 B시점인 S_1 을 부호화한다.

IV. 다시점 비디오 부호화 기술 소개

1. 조명 보상 기술

(Illumination Compensation)

다시점 비디오는 여러 대의 카메라를 이용하여 촬영한 영상이기 때문에 동일한 장면을 촬영하더라도 카메라의 위치 변이, 카메라의 특성 차이와 조리개 조절 변이 등의 이유로 조



(그림 2) 다시점 비디오 참조 소프트웨어의 시·공간 예측 구조

명값의 변화가 생길 수 있다. 이는 사람의 눈으로 그 차이를 크게 느낄 수 없으나 작은 화소값의 변화에도 민감한 비디오 부호화 과정에서는 문제가 될 수 있다. 따라서 이에 따른 다양한 조명 보상 기술이 제안되었고, 현재 매크로블록 단위의 조명 보상 기술이 JMVM 2.0에 포함되었다. 이 방법은 움직임 예측/보상 과정에서 매크로블록 단위로 조명을 보상하는 방법(illumination change-adaptive motion compensation)이다[7].

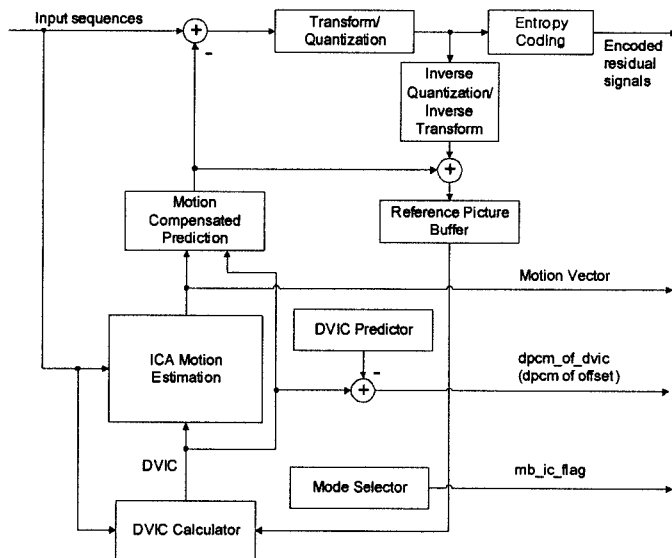
<그림 3>에 보인 것처럼, 기존의 다양한 매크로블록 모드에 조명을 보상한 16×16 모드를 추가하여 움직임 예측을 수행하고, 비트율-왜곡 (rate-distortion) 관점에서 조명을 보상한 16×16 모드가 선택될 경우 조명을 보상하며, 이 경우에 DVIC (difference value of illumination change) 값을 보낸다. 이 값은 dpcm_of_dvic라는 플래그(flag) 비트가 사용되며, 조명 보상이

이루어진 경우에 대해서는 기존의 SAD(sum of absolute differences)가 아닌 식 (1)과 같은 MR_SAD(mean-removed SAD)가 사용된다 [7].

$$MR_SAD(x, y) = \sum_{i=m}^{m+S-1} \sum_{j=n}^{n+T-1} \left\{ f(i, j) - M_{cur} - \left\{ r(i+x, j+y) - M_{ref}(m+x, n+y) \right\} \right\} \quad (1)$$

여기서 M_{cur}과 M_{ref}는 현재 블록과 참조 블록의 평균값을 나타내고, f(i,j)와 r(i,j)는 현재 화면과 참조화면의 화소값을 각각 나타낸다.

2007년 5월에 배포된 JMVM4.0에는 제안한 B_Skip/Direct_16×16 모드에 대해 조명 보상에 관련된 정보를 유추하여 사용하는 방법과 블록 단위의 조명 보상으로 인해 발생하



<그림 3> 매크로블록 기반의 ICA ME/MC 부호화 구조

는 블로킹 현상 (blocking artifact)를 줄이는 방법이 추가로 구현되었다⁸⁾. <그림 4>와 같이 조명 보상이 발생하는 인터 모드에 대해 디블로킹 필터를 적용하는 과정에서, 기존의 방법과 달리, 경계 부분에 대한 필터의 강도 (boundary filtering strength, bS)를 조정하는 방법을 조명 보상을 고려하여 결정함으로써 조명 보상으로 인한 블로킹 현상을 효율적으로 줄일 수 있다⁸⁾.

2. 움직임 정보 생략 모드 (Motion Skip Mode)

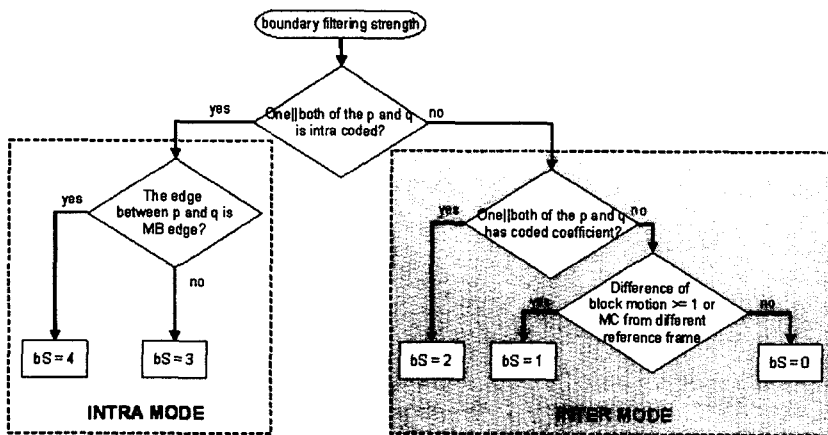
다시점 비디오에서는 인접하는 시점의 영상들 사이에는 시간적 상관도가 상당히 유사하다. 움직임 생략 모드는 이러한 특징을 고려하여 이전에 부호화된 인접 시점의 움직임 정보를 공유하는 방법을 통해 부호화 효율을 증진시킨다. 매크로블록에 대해 Motion Skip Mode가 선택되면 매크로블록 모드, 참조 인덱스, 움직임 벡터와 같은 움직임 관련 정보는 부호

화되지 않고, 인접 시점의 상응 매크로블록의 움직임 정보로 대체된다. 해당 기술은 2007년 5월에 JMVM4.0에 구현되었다⁸⁾.

움직임 정보 생략 모드는 크게 인접 시점에서 상응 매크로블록을 찾는 단계와 상응 매크로블록으로부터 움직임 정보를 유도하는 단계로 나뉜다. 상응 매크로블록은 시점간의 예측만을 이용해 부호화되는 기준 화면 (anchor frame)에서 매크로블록 단위로 측정된 전역 변이 값(global disparity value, GDV)을 이용하여 찾고, 해당 GDV 값은 기준 화면의 헤더에 부호화된다. 비기준 화면(non-anchor frame)에 대한 GDV 값은 식 (2)와 같이 해당 화면의 POC(picture order count)와 해당 비기준 화면이 속한 GOP의 양 끝단에 위치한 기준 화면의 GDV 값으로부터 유도된다.

$$GDV_{cur} = GDV_A + \left[\frac{POC_{cur} - POC_A}{POC_B - POC_A} \times (GDV_B - GDV_A) \right]$$

(2)



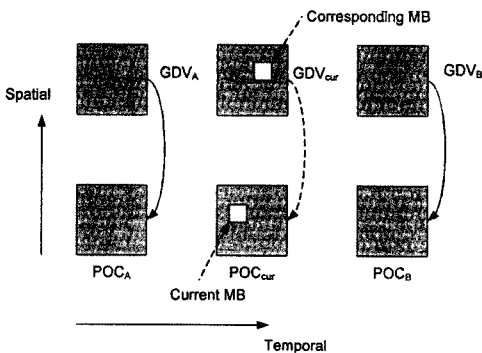
<그림 4> 조명 보상 디블로킹 필터링 조건

각 화면에 대한 GDV 값을 계산한 후 해당 화면에 속하는 매크로블록은 GDV 값을 통해 인접 시점에서 상응 매크로블록을 찾고, 상응 매크로블록의 움직임 관련 정보를 공유한다. 최근에는 GDV 계산의 복잡도를 줄이기 위해 GDV 값 계산을 식 (3)과 같이 단순화하여 계산하기도 한다^[9].

$$GDV_{cur} = GDV_A \quad (3)$$

움직임 정보 생략 모드는 기존의 다른 모드와 비트율-왜곡 (rate-distortion) 관점에서 비교되고, 움직임 정보 생략 모드가 최적으로 판단되는 경우 매크로블록 레벨에서 정의된 motion_skip_flag를 전송하여 움직임 관련 정보의 부호화를 대신한다. <그림 5>는 Motion Skip Mode에서 상응 매크로블록에 대한 개념과 GDV의 유도를 설명한다.

3. 적응적인 참조화면 필터링 (Adaptive Reference Filtering)



<그림 5> 움직임 정보 생략 모드에 대한 개념도

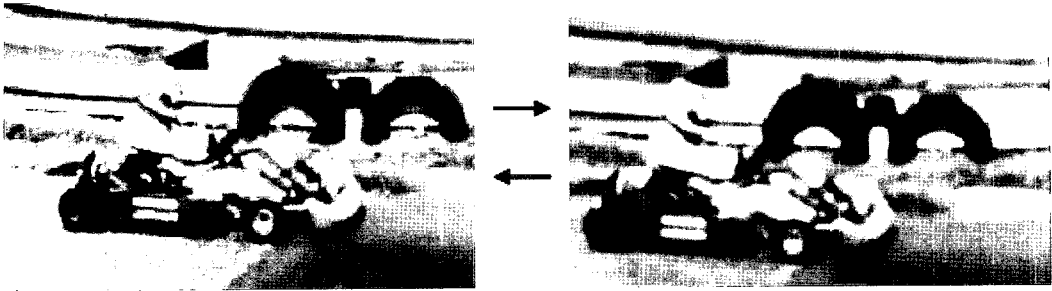
다시점 비디오에서는 여러 대의 카메라를 이용하여 동일한 3차원 장면을 촬영하기 때문에 인접한 화면에서 화면의 초점이 맞지 않는 문제가 발생할 수 있다. <그림 6>은 Race1 영상에 대해 인접한 두 화면에서의 초점 불일치로 인한 화면의 차이를 보여주고 있다.

이처럼 두 영상에 대해 한 시점의 영상은 전경에 초점이 맞춰 촬영되고, 다른 한 시점의 영상은 배경에 초점이 맞춰져 촬영이 되었다면 둘 사이의 상관도는 낮아지기 때문에, 인접 시점을 참조하여 부호화하는 시점 방향의 예측 효율이 떨어지게 된다. 최근 이를 보정하기 위해 참조 화면의 전경과 배경에 대해 적절한 필터링을 통해 마치 참조 화면이 부호화 할 화면과 동일한 초점에서 촬영된 것처럼 보정하는 방법이 제안되었다. 이를 통해 초점의 불일치로 인해 발생하는 참조 화면과 부호화 할 화면간의 상관도의 저하를 보정할 수 있어 부호화 효율의 증진이 기대된다^[10].

4. 합성 영상 예측 (View Synthesis Prediction)

다시점 비디오에서는 인접한 시점의 영상끼리는 상관도가 높다. 어떤 시점의 영상에 관한 정보는 이와 인접한 시점에도 대부분 존재하는 것을 알 수 있다. 따라서 만약 좌우 영상으로부터 그 사이에 존재하는 가상 시점에 대한 영상 생성이 가능하다. 이를 위해 현재 JVT 내에서는 합성 영상 예측 기술에 대한 연구가 진행되고 있다. 이러한 연구의 큰 흐름은 깊이 영상 생성과 이를 이용한 합성 영상 생성에 초점을 맞추고 있다.

깊이 영상이란 영상 속에 존재하는 객체들



〈그림 6〉 Motion Skip Mode에 대한 개념도

이 카메라와 상대적으로 얼마나 떨어져 있는지를 나타내는 상대적인 정보로 컬러 영상과 달리 4:0:0 컬러 포맷의 정보로 이루어져 있으며, 이 영상의 화소 값은 해당 화소의 깊이 정보를 알려준다. 현재 다시점 영상에서는 이러한 깊이 영상을 제공해 주지 않기 때문에 깊이 영상을 올바르게 만들어 내는 작업도 중요한 이슈이다^[11].

깊이 맵 생성에는 블록 기반 방법과 영역 기반의 방법이 제안되었다. 블록 기반의 방법은 복잡도는 낮으나, 한 블록내에 다른 깊이를 갖는 객체가 동시에 존재할 때 이를 효과적으로 처리하기가 쉽지 않다. 반면, 영역 기반 깊이 영상 생성의 경우 경계와 같은 부분에서 효과적으로 깊이를 얻어낼 수 있지만 영역 구분(image segmentation) 알고리즘의 결과에 의존적이다. 블록 기반이든 영역 기반이든 인접 시점에서 자기와 일치하는 영역을 찾아서 이 정보를 이용하여 깊이값을 계산한다. 깊이 값을 얻어내는 방법은 변이를 찾은 후 깊이 정보로 변환하는 방법과 3D 워핑 기술 기반으로 깊이 정보를 직접 찾는 방법이 제안되었다.

이렇게 생성된 깊이 영상을 부호화하는 방

법에는 깊이 영상을 비디오 데이터로 간주하고 이를 H.264나 JMVM을 이용하여 부호화하는 방법이나 동일 위치의 매크로블록을 부호화할 때 블록 단위로 부호화하는 방법이 제안되었다. 합성 영상 예측에 이용되는 깊이 영상은 다시점 영상을 부호화한 데이터와 함께 전송되어야 하는 부가정보 성격이므로 부호화 효율을 올릴 수 있을 정도의 화질이면서 최대한 압축되어야 한다. 합성된 영상을 이용한 다시점 비디오 부호화 방법으로는 합성된 영상을 추가적인 참조 화면으로 이용하는 방법과 부호화할 화면과의 차이 영상을 부호화하는 방법이 제안되었다. 현재 다양한 깊이 영상 생성 방법과 깊이 영상 부호화 방법, 그리고 이를 이용한 다시점 비디오 부호화 방법이 연구되고 있다^[11].

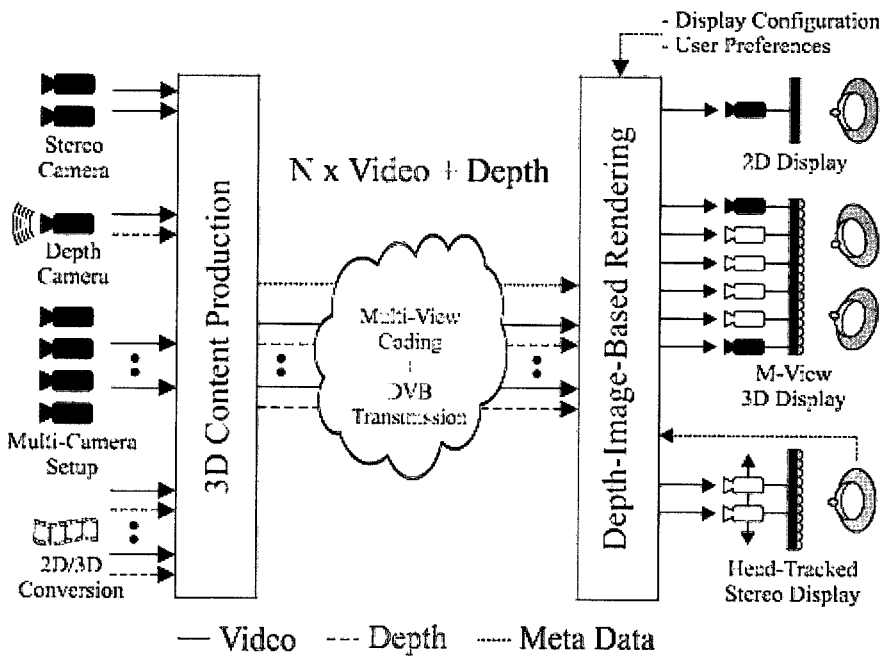
5. 가상시점 영상 생성 (Virtual View Synthesis) 및 자유시점 TV (Free Viewpoint TV)

다시점 비디오는 다양한 방법으로 디스플레이 될 수 있다. 우선 시청자가 다양한 시점의

영상 중에서 원하는 한 영상을 골라서 보는 자유시점 디스플레이를 지원하며, 다양한 입체 모니터를 통해 시청자에게 입체 영상을 제공할 수도 있다. 현재 스테레오 모니터와 다시점 모니터를 통한 입체 영상 재생이 가능하지만, 이러한 모니터는 입력 영상 사이의 간격이 특정 조건을 만족해야 입체감있는 영상을 제공하는 한계점이 있다. 따라서 일반적인 다시점 영상을 이러한 입체 모니터에 이용했을 경우에 입체감을 제대로 제공하지 못하는 경우도 있다. 이러한 문제점을 해결하기 위한 방법이 바로 가상시점 영상 생성이다. 여기서 가상시점이란 부호화 단에서 부호화를 위해 입력받은 시점의 영상 이외의 시점을 말한다. 이를 위해서는 일반적으로 깊이(depth) 영상이 필

요하다. <그림 7>은 최근 MPEG에서 논의되고 있는 FTV (free viewpoint TV)의 구조도를 보여준다. FTV의 경우에는 기존의 다시점 영상에 깊이 영상을 더하여 다양한 방식의 디스플레이 시나리오를 고려하고 있다^[2].

FTV에서는 입체 영상 생성에 필요한 여러 시점 혹은 가상 시점까지 촬영하거나 부호화할 수 없으므로 특정 N 시점의 영상만을 입력으로 받아 M 시점 ($M \geq N$) 영상을 디스플레이하는 시스템을 연구하고 있다. 이러한 가상시점 영상 생성에는 현재 깊이 영상과 카메라 변수가 이용된다. 합성 영상 예측 기술에서 사용되는 깊이 영상과는 달리, 가상시점 영상 생성을 위해 사용되는 깊이 영상을 고화질로 부호화되어야 한다. N 시점으로부터 M 시점



<그림 7> 자유시점 TV에 대한 구조도

을 만들어 내는 위치에 따라 다양한 시나리오가 가능하지만, 각 시나리오의 차이는 부호화기와 복호화기의 복잡도나 전송되는 데이터의 양이 다르다. 예를 들어, 부호화기 측에서 가상 시점 영상을 생성하여 M 시점의 영상을 생성하는 경우 부호화의 복잡도가 증가되고 전송되는 데이터 양도 많아지게 된다. 만약 부호화 측에서는 깊이 영상만 전송하고 복호화기 측에서 가상시점 영상을 생성하는 경우에는 전송되는 데이터 양은 줄일 수 있지만 복호화기의 복잡도가 증가하게 된다. 현재는 다양한 시나리오를 가정하여 이에 관한 연구를 진행하고 있다^[12].

IV. 맺음말

이 논문에서는 다시점 비디오 부호화(MVC)에 관한 일반적인 내용을 소개하고, 관련되는 기술의 표준화 작업에 대해 살펴보았다. 다시점 비디오는 사용자에게 자유 시점 비디오나 3차원 TV와 같은 보다 입체감있는 영상을 사용자에게 제공할 수 있지만, 처리해야 할 데이터 양이 방대하여 효율적인 부호화 기술이 요구되는 분야이다. 현재 JVT를 중심으로 MVC에 대한 국제 표준화 작업에 발맞추어 다양한 부호화 기술과 관련 분야의 연구가 활발히 진행 중이다. MVC에 대한 표준화 작업은 2008년 4월 최종 국제규격안(Final Draft International Standard, FDIS)이 마련될 예정이다. 이와 더불어 최근에는 다시점 비디오와 깊이 영상을 이용한 자유 시점 TV에 관한 연구가 시작되어 이에 관한 기술 개발을 선도할 예정이다.

감사의 글

본 논문은 광주과학기술원(GIST) 실감방송 연구센터(RBRC)를 통한 정보통신부 대학 IT 연구센터(ITRC)와 교육인적자원부 두뇌한국 21(BK21) 정보기술사업의 지원으로 수행되었다.

참고문헌

- [1] A. Smolic, K. Mueller, T. Rein, P. Eisert, and T. Wiegand, "Free Viewpoint Video Extraction, Representation, Coding, and Rendering," IEEE International Conference on Image Processing. Vol. 5, pp. 3287-3290, 2004.
- [2] R.S. Wand and Y. Wang, "Multiview Video Sequence Analysis, Compression, and Virtual Viewpoint Synthesis," IEEE Transactions on Circuit and System for Video Technology, Vol. 10, pp. 397-410, 2000.
- [3] A. Redert, M. Beeck, C. Fehn, W. IJsselsteijn, M. Pollefeys, L. Gool, E. Ofek, I. Sexton, and P. Surman, "Advanced three-dimensional television system technologies," Proc. 3D Data Processing Visualization and Transmission, 2002.
- [4] ISO/IEC JTC1/SC29/WG11, "Requirements on Multi-view Video Coding v.7," N8218, July 2006.
- [5] ISO/IEC MPEG & ITU-T VCEG, "Common Test Condition for Multiview Video Coding," JVT-U211, Oct. 2006.
- [6] ISO/IEC MPEG & ITU-T VCEG, "Joint Multiview Video Model (JMVM) 1.0," JVT-T208, July 2006.
- [7] ISO/IEC MPEG & ITU-T VCEG, "Joint Multiview Video Model (JMVM) 2.0," JVT-U207, Oct. 2006.

- [8] ISO/IEC MPEG & ITU-T VCEG, "Joint Multiview Video Model (JMVM) 4.0," JVT-W207, April 2007.
- [9] ISO/IEC MPEG & ITU-T VCEG, "MVC Decoder Simplification of Motion Skip Mode," JVT-X024, July 2007.
- [10] ISO/IEC MPEG & ITU-T VCEG, "Adaptive Reference Filtering for MVC with decoder complexity analysis," JVT-X060, July 2007.
- [11] ISO/IEC MPEG & ITU-T VCEG, "CE3: View Interpolation Prediction for MVC," JVT-W303, April 2007.
- [12] ISO/IEC MPEG & ITU-T VCEG, "Multi-view video plus depth (MVD) format for advanced 3D video systems," JVT-W100, April 2007.

저자소개



호요성

1981년 2월 서울대학교 전자공학과 학사
 1983년 2월 서울대학교 전자공학과 석사
 1989년 12월 Univ. of California, Santa Barbara, Department of Electrical and Computer Engineering, 박사
 1983년 3월-1995년 9월 한국전자통신연구소 선임연구원
 1990년 1월-1993년 5월 미국 Philips 연구소, Senior Research Member
 1995년 9월-현재 광주과학기술원 정보통신공학과 교수
 주관심 분야 : 디지털 신호처리, 영상신호 처리 및 압축, 멀티미디어 시스템, 디지털 TV와 고선명 TV, MPEG 표준, 3차원 TV, 실감방송



오관정

2002년 8월 전남대학교 정보통신공학부 학사
 2005년 2월 광주과학기술원 정보통신공학과 석사
 2005년 3월-현재 광주과학기술원 정보통신공학과 박사과정
 주관심 분야 : 다시점 비디오 부호화(MVC), H.264/AVC, 3차원 TV, 디지털 신호처리, 영상신호 처리 및 압축, 실감방송