JKIICE Journal of the Korea Institute of Information and Communication Engineering

한국정보통신학회논문지(J. Korea Inst. Inf. Commun. Eng.) Vol. 20, No. 2 : 375~384 Feb. 2016

# 부복호화기 효율을 고려한 다시점 영상 합성 기법

최인규[1] · 정원식[2] · 이광순[2] · 유지상[3*]

# Multi-view Synthesis Algorithm for the Better Efficiency of Codec

**In-kyu Choi[1] · Won-sik Cheong[2] · Gwangsoon Lee[2] · Jisang Yoo[3*]**

[1]Department of Electronic Engineering, Kwangwoon University, Seoul 01897, Korea
[2]Electronics and Telecommunications Research Institute, Daejeon 34129, Korea
[3*]Department of Electronic Engineering, Kwangwoon University, Seoul 01897, Korea

## 요 약

본 논문에서는 스테레오 영상과 스테레오 영상 바깥 시점의 위성 영상 그리고 상응하는 깊이 영상이 입력으로 주어질 때 이 입력 영상들을 압축에 적합한 포맷으로 변환하고 활용하여 중간시점 영상을 합성하는 새로운 기법을 제안한다. 송출 단에서 깊이 영상들은 하나의 global 영상으로, 위성영상은 시점 이동 시에 발생하는 프레임 밖의 영역 및 가려짐 영역과 같은 홀 영역에만 해당하는 residual 영상으로 변환 후에 데이터 량을 줄이기 위하여 다운샘플링하고 주 시점의 스테레오 영상과 더불어 HEVC 코덱으로 부호화 한 후 전송한다. 수신 단에서 복호화된 각 입력 영상들을 이용하여 스테레오 영상 간 그리고 스테레오 영상과 위성 영상간의 중간시점 영상을 합성한다. 실험을 통하여 제안한 포맷을 이용하여 합성한 중간 시점 영상을 MVD(Multiview Video plus Depth) 포맷을 이용하여 합성한 중간 시점 영상과 비교했을 때 전송하는 데이터의 비트율(bit-rate) 대비 객관적 그리고 주관적 화질이 뛰어남을 확인하였다.

## ABSTRACT

In this paper, when stereo image, satellite view and corresponding depth maps were used as the input data, we propose a new method that convert these data to data format suitable for compressing, and then by using these format, intermediate view is synthesized. In the transmitter depth maps are merged to a global depth map and satellite view are converted to residual image corresponding hole region as out of frame area and occlusion region. And these images subsampled to reduce a mount of data and stereo image of main view are encoded by HEVC codec and transmitted. In the receiver intermediate views between stereo image and between stereo image and bit-rate are synthesized using decoded global depth map, residual images and stereo image. Through experiments, we confirm good quality of intermediate views synthesized by proposed format subjectively and objectively in comparison to intermediate views synthesized by MVD format versus total bit-rate.

**키워드** : 포맷 변환, Global 깊이 영상, residual 영상, 영상 합성

Key word : Format conversion, global depth map, residual view, view synthesis

**Received** 20 October 2015, **Revised** 17 November 2015, **Accepted** 01 December 2015

*** Corresponding Author** Jisang Yoo(E-mail:jsyoo@kw.ac.kr, Tel:+82-2-940-5112)
Department of Electronic Engineering, Kwangwoon University, Seoul 01897, Korea

Open Access http://dx.doi.org/10.6109/jkiice.2016.20.2.375 print ISSN: 2234-4772 online ISSN: 2288-4165

ⓒThis is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(http://creativecommons.org/li-censes/by-nc/3.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright Ⓒ The Korea Institute of Information and Communication Engineering.

## Ⅰ. 서 론

3DTV 방송 서비스는 특수 안경을 통하여 좌안과 우안에 각각 하나의 영상을 제공하는 스테레오스코픽(stereoscopic) 방법이 보편적이다. 최근에 3DTV 방송 서비스의 활성화와 안경 착용의 불편함을 해소하기 위해서 무안경 방식(auto-stereoscopic)의 기술이 개발되고 있는데 어떤 시점에서도 양안으로 영상을 보기 위해서는 많은 시점이 존재해야 한다. 이러한 다시점 3D 기술은 기존 스테레오스코픽 방법과 달리 안경착용의 불편함이 없을 뿐만 아니라 시청자의 움직임에 따라 대응되는 시점의 영상을 보여주는 운동 시차(motion parallax)를 제공함으로써 입체감을 향상시킬 수 있는 장점을 지닌다. 하지만 무안경식 3D 디스플레이는 시점 수만큼의 많은 데이터를 저장하고 전송해야하기 때문에 매우 효율적인 부호화 방법과  여기에 적합한 데이터 포맷(format)이 필요하다.

이러한 목적으로 깊이 영상을 사용하여 다시점 영상을 합성하는 방법이 연구되고 있는데 그 이유는 몇 개의 입력 영상과 그에 상응하는 깊이 영상으로 원하는 목표시점으로 투영하여 많은 수의 시점 영상을 합성할 수 있기 때문이다. 관련하여 깊이영상 촬영 시스템 또는 스테레오 매칭(stereo matching)을 통한 깊이영상 합성에 대한 연구[1-3]와 획득된 깊이영상에서 오류를 제거하여 정교한 깊이 영상을 생성하는 연구[4,5] 등이 진행되고 있다. 깊이 영상을 이용한 워핑(warping)과 시점 이동시에 발생하는 홀 영역을 보간하는 연구[6-8] 또한 활발하게 진행되고 있다. 기존에 존재하는 2D 코덱과 호환이 가능하면서도 적은 데이터를 이용하는 다시점 부호화 방법[9]도 제안되었다. 이 방법은 부호화 전에 입력 영상과 깊이 영상 간의 불필요한 정보를 제거하여 하나의 global 영상과 깊이 영상을 생성하는 전처리 과정을 거친다.

본 논문에서는 중간영상 합성을 위해 필요한 스테레오 영상과 스테레오 영상 바깥 시점의 위성 영상 그리고 상응하는 깊이 영상이 입력으로 제공된다. 먼저 각각의 깊이 영상을 가지고 하나의 global 깊이 영상을 생성한다. 위성영상은 중복 정보(redundancy)를 제거하여 residual 영상으로 생성된다. 결국 스테레오 영상과 하나의 global 깊이 영상, 그리고 위성 영상의 residual 영상이 부호화되어 전송된다. 수신기에서는 부호화된 각각의 영상들을 복호화기를 통해 복원하고, 복원된 스테레오 영상과 깊이영상, 그리고 위성영상의 residual 영상을 가지고 중간 영상을 합성하게 된다. 여기서 부호화기는 HEVC 코덱을 사용하며 각각의 입력 영상에 적용하는 QP(quantization parameter)에 따라 합성되는 중간 영상의 결과가 다른데 이는 압축정도에 따라서 성능이 달라지기 때문이다. 입력 데이터의 비트율(bit-rate) 대비 합성된 중간 영상의 PNSR을 측정하여 깊이 영상 기반 3D 비디오 포맷 중의 하나인 MVD(Multiview Video plus Depth)을 통해 합성된 중간 영상과 그 성능을 비교한다.

본 논문은 다음과 같이 구성된다. 2장에서 송출단에서 데이터를 줄이기 위해 global 깊이 영상 및 residual 영상을 생성하는 방법을 먼저 소개한다. 이어서 3장에서는 복호화 된 스테레오 영상, global 깊이 영상 및 residual 영상을 이용하여 다시점 영상의 중간영상을 합성하는 기법에 대해 설명한다. 4장에서는 실험 결과로써 입력 데이터의 비트율 대비 합성된 중간 영상의 PSNR을 측정하여 MVD 포맷과 성능을 비교한 결과를 보여준다. 마지막으로 5장에서 본 논문의 결론을 맺는다.

## Ⅱ. 본 론

그림 1은 부호화기의 데이터 효율을 높이기 위해 송출 단에서 처리되는 과정을 나타낸다. 그림 1은 입력 영상이 4시점인 경우로 V2, V3는 주시점을 구성하는 기본 스테레오 영상(base view)이고 V1, V4는 주시점 영상의 바깥쪽에 존재하는 시점 영상으로 위성 영상(satellite view)이라고 부른다.

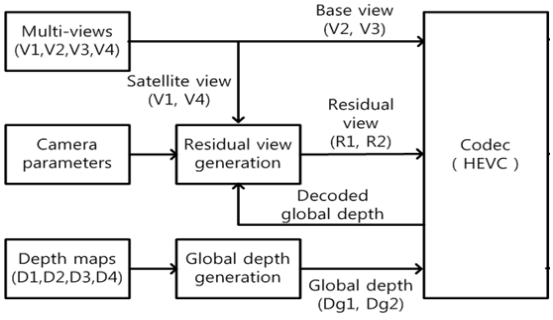



**Fig. 1** Processes in the transmitter

주시점의 스테레오 영상을 기준으로 V1 영상을 가장 왼쪽, V4 영상을 가장 오른쪽 시점의 영상으로 생각하면 된다. 또한 이 영상들에 대응하는 깊이 영상(D1, D2, D3, D4)과 카메라 파라미터 정보가 부호화기의 입력 데이터에 포함된다.

송출 단에서는 global 깊이 영상 생성과 위성 영상의 residual 영상을 생성하게 된다. 먼저 주어진 4개의 깊이 영상과 카메라 파라미터를 활용하여 하나의 global 깊이 영상을 생성한다. 생성된 global 깊이 영상과 카메라 파라미터를 활용하여 위성 영상(V1, V4) 시점의 residual 영상(R1, R2)을 생성한다. 입력으로 주어진 기본 스테레오 영상, 생성된 global 깊이 영상과 residual 영상이 각각 HEVC 부호화기를 거쳐 전송 된다.

### 2.1. Global 깊이 영상 생성

Global 깊이 영상을 생성하기 위해서 그림 2와 같이 깊이 영상 $D_1, D_2, D_3, D_4$를 중심 시점으로 투영(projection)한다. 이때 중심시점은 D2와 D3 시점의 중간시점이 된다. 깊이 영상 $D_1, D_2, D_3, D_4$는 입력 영상 V1, V2, V3, V4에 각각 대응한다. 시차가 $x$축 방향으로만 존재한다고 가정하면 투영 전의 $D_i$ 깊이 영상의 한 화소의 좌표 $(x_i, y)$와 투영 후의 좌표$(x_i^g, y)$간의 관계는 식 (1)과 같다.

$$\triangle x_i = \frac{fB_i}{255}\left(\frac{1}{Z_{near}} - \frac{1}{Z_{far}}\right)D_i(x_i, y) + \frac{fB_i}{Z_{far}} \quad (i = 1,2,3,4) \tag{1}$$

$$x_1^g = x_1 - \triangle x_1,\ x_2^g = x_2 - \triangle x_2,$$
$$x_3^g = x_3 + \triangle x_3,\ x_4^g = x_4 + \triangle x_4$$

여기서 f는 카메라의 초점 거리이고 $B_i$는 $D_i$ 카메라와 앞에서 정의한 중심 시점 $D_g$ 카메라 간의 거리(baseline length)이고 $Z_{near}$와 $Z_{far}$는 각각 가장 멀리 떨어진 물체와의 거리, 가장 가까운 물체와의 거리를 나타낸다. 이 거리는 모든 깊이 영상에서 같으며 카메라 파라미터에 포함되어 있다. $D_i(x_i, y)$는 화소 $(x_i, y)$에서의 깊이값을 나타내고 $\triangle x_i$은 $D_i(x_i, y)$와 $D^g(x_i^g, y)$ 간의 $x$축 방향 시차(disparity)를 의미한다.

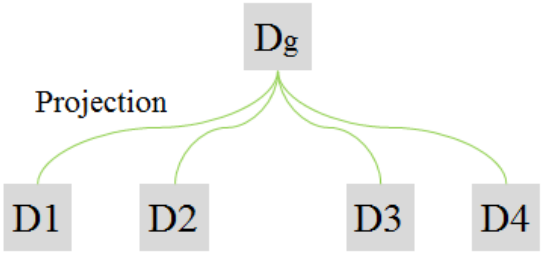


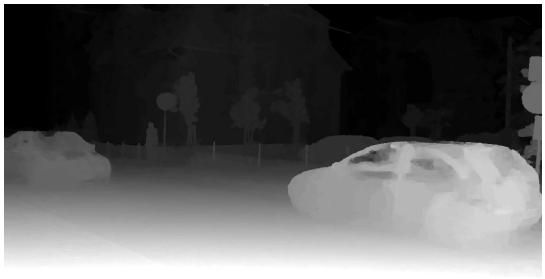

**Fig. 2** Generation of the global depth image

중심 시점으로 투영된 깊이 영상에는 홀이 발생할 수 있다. 이때 발생하는 홀(hole) 영역은 가장 인접한 좌, 우 화소 중에 홀이 아닌 화소의 값을 비교하여 작은 값으로 채운다. 그 이유는 투영 시에는 전경에 의해 가려져 있던 배경이 드러나는 경향이 있기 때문이다. 중심 시점으로 투영하고 홀 영역을 채운 4개의 깊이 영상은 식 (2)를 이용하여 평균값을 구하고 중심 시점에 해당하는 하나의 global 깊이 영상 $D_g(x, y)$를 생성하게 된다.

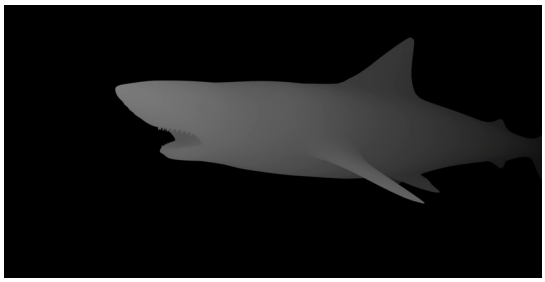
$$D^g(x,y) = \frac{D_1^g(x,y) + 2D_2^g(x,y) + 2D_3^g(x,y) + D_4^g(x,y)}{6} \tag{2}$$

여기서 $D_i^g(i = 1,2,3,4)$는 깊이 영상 $D_1, D_2, D_3, D_4$가 중심 시점으로 투영된 후, 홀 영역을 채운 영상을 각각 의미한다. 이때 중심 시점에서 가까운 깊이 영상은 투영된 깊이 영상과 상관도가 높을 것이라고 판단하여 큰 가중치를 부여함으로써 시점 이동으로 인한 오류를 줄이고자 하였다. 그림 3은 $D_2$와 생성된 global 깊이 영상을 비교한 결과이다.

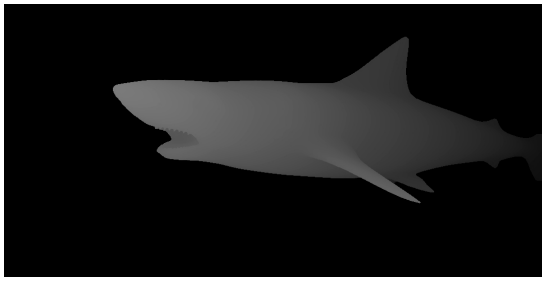
(a)



(b)



(c)



(d)

**Fig. 3** Given depth image and global depth image (a) Input depth image D2(Street) (b) Global depth image (Street) (c) Input depth image D2(Shark) (d) Global depth image(Shark)

이렇게 생성된 global 깊이 영상은 부호화 시 데이터의 양을 줄이기 위하여 수평, 수직 방향으로 1/2 만큼 다운샘플링(down sampling) 하게 된다. 수신 단에서 다시점 영상을 합성할 때는 global 깊이 영상을 원래 크기로 복원해야 하는 데 이 때 식 (3)의 보간 식을 이용하게 된다.

$$D^g(x,y)=\begin{cases}(D^g(x-1,y)-D^g(x+1,y))/2 \\ \text{if}\,|D^g(x-1,y)-D^g(x+1,y)|<th \\ \max(D^g(x-1,y),D^g(x+1,y) \\ \text{if}\,|D^g(x-1,y)-D^g(x+1,y)|<th\end{cases} \quad (3)$$

즉, 보간해야 할 화소의 좌, 우 화소의 차이값이 임계값보다 작으면 두 화소의 중간값으로, 임계값보다 크면 큰 값으로 보간하게 된다.

### 2.2. Residual 영상 생성

Residual 영상은 4시점의 입력영상 중에 좌측과 우측 끝에 존재하는 2개의 위성 영상을 가지고 생성하게 된다. 그림 4는 residual 영상을 생성하는 과정을 보여준다.

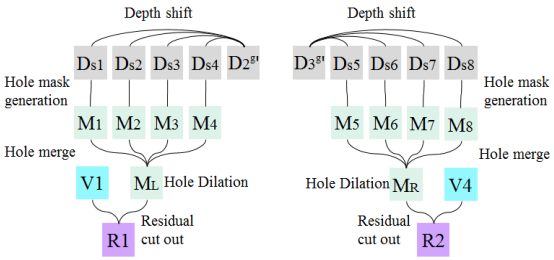



**Fig. 4** Generation of the residual image

Residual 영상은 수신 단에서 다시점 영상을 생성할 때 스테레오 영상과 위성 영상 간(V1과 V2, V3과 V4)의 중간 시점을 외삽으로 생성할 시에 발생하는 홀 영역을 채우기 위한 영상이다. 따라서 수신 단에서 생성하고자 하는 시점에 대한 홀 영역 정보를 가지고 있어야 한다. 홀 영역 검출은 복호화된 global 깊이 영상 $D^{g\prime}$을 이용하게 되는데 이는 수신 단에서 외삽에 의한 중간 시점 합성 시 일관성을 유지하기 위함이다.

Residual 영상을 생성하기 위하여 먼저 복호화된 global 깊이 영상 $D^{g\prime}$을 스테레오 영상 시점(V2와 V3)으로 투영하여 $D_2^{g\prime}$와 $D_3^{g\prime}$을 각각 생성한다. 이어서 $D_2^{g\prime}$와 $D_3^{g\prime}$을 각각 수신 단에서 합성하고자 하는 목표 시점으로 투영하여 $D_{si}(i=1,2,3,4,5,6,7,8)$을 생성한다. 여기서 $i$는 수신 단에서 합성하고자 하는 다시점 영상 수에 따라 다르게 되며 이 절에서는 스테레오 영상과 위성 영상 간의 중간 시점의 수를 좌우 각각 4개씩으로 설정하였다. 그리고 식 (4)를 이용하여 홀 영역을 판별한다.

$$(D_{si}(x+\delta,y)-D_{si}(x,y)) \geq th, \quad (x,y)\ is\ in\ the\ occlusion\ region \tag{4}$$

$$th = \frac{255}{fB_i(\frac{1}{Z_{near}}-\frac{1}{Z_{far}})}\delta,$$

$$(0 \leq \delta \leq fB_i(\frac{1}{Z_{near}}-\frac{1}{Z_{far}})\ \text{단, } \delta\text{는 정수})$$

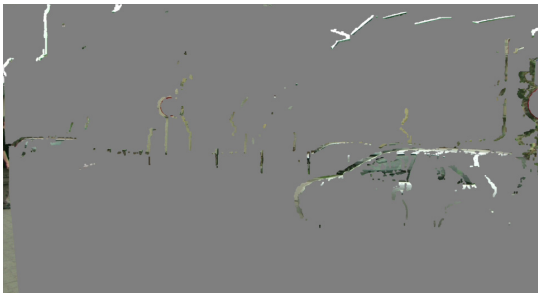
판별하고자 하는 화소의 깊이값 $D_{si}(x,y)$과 인접한 화소의 깊이값 $D_{si}(x+\delta,y)$의 차이가 임계값보다 클 경우 화소 $(x,y)$는 가려짐 영역(occlusion)으로 판별한다. 이 때 $B_i$는 $D_2^{g\prime}$ 또는 $D_3^{g\prime}$와 투영하는 목표 시점간의 기선장(baseline length)이며 $\delta$는 정수값을 가지는 시차(disparity)로 정의되고 임계값은 카메라의 초점거리 $f$와 기선장, $B_i$ 그리고 시차 $\delta$를 가지고 계산할 수 있다. 이렇게 홀 영역으로 판별된 화소와 시점 이동시에 깊이영상의 좌우측 경계부분이 사라지는 경우, 화소를 255의 값으로 할당하고 그렇지 않은 경우는 0으로 할당하여 홀 마스크(hole mask), $M_i(i=1,2,3,4,5,6,7,8)$를 생성하게 된다.

생성된 홀 마스크 중 $M_i(i=1,2,3,4)$는 모두 V1 위성 영상의 시점으로 시점 이동 후에 OR 연산을 통해 그림 4에 보인 마스크 $M_L$를 생성하게 된다. 마찬가지로 $M_i(i=5,6,7,8)$는 V4 위성 영상 시점으로 시점 이동 후 같은 연산을 하여 그림4와 같이 $M_R$을 만들게 된다. 합병된 마스크는 팽창(dilation) 연산을 적용하여 $M_L^{'}$, $M_R^{'}$을 만들고, 다시점 영상 합성 시에 경계를 채우지 못하는 오류를 줄인다[10]. Residual 영상은 홀 마스크 $M_L^{'}$, $M_R^{'}$에서 255에 해당되는 화소 값만 위성 영상 V1, V4에서 가져온다. 나머지 영역은 앞에서 가져온 화소의 평균값으로 채운 후 최종적으로 residual 영상 R1과 R2을 생성한다.

그림 5는 생성된 홀 마스크 및 residual 영상의 결과를 보여준다. Residual 영상도 데이터량을 줄이기 위해 가로, 세로 모두 1/2 크기로 다운샘플링 하여 부호화한다. 이때 저역필터인 (0.25, 0.5, 0.25)의 3-tap 필터를 적용하여 다운샘플링으로 인한 에일리어싱(aliasing) 현상을 줄일 수 있다.

(a)

(b)



(c)

**Fig. 5** (a) V4 bit-rate of 'street' (b) generated hole mask (MR') (c) residual image

## Ⅲ. 수신 단에서의 다시점 영상 합성기법

수신 단에서는 그림 6과 같이 복호화기를 거쳐 원래의 스테레오 영상, V2와 V3, global 깊이 영상 $D^{g\prime}$, residual 영상 R1과 R2을 각각 복원한 후 다시점 영상의 중간 영상을 합성하는 이용하게 된다. 이때 스테레오 영상 V2와 V3 사이의 중간 시점 영상은 내삽으로, 스테레오 영상과 위성 영상 사이(V1과 V2, V3와 V4)의 중간 시점 영상은 외삽 기법을 통해 합성하게 된다. 내삽은 global 깊이 영상과 주시점 영상인 스테레오 영상

만 가지고 합성하는 방법이고 외삽은 스테레오 영상과 global 깊이 영상 그리고 residual 영상을 이용하게 된다.

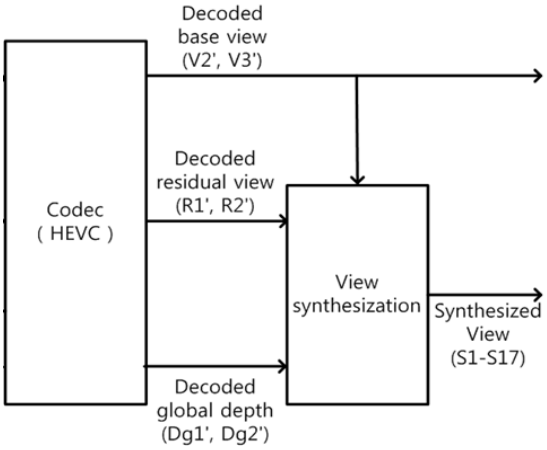



**Fig. 6** Multi-view synthesis process in the receiver

### 3.1. 내삽(Interpolation)

먼저 복호화된 global 깊이 영상($D^{g'}$)를 목표 시점으로 투영하고 2.1절에서와 같이 홀 영역을 채운다. 투영된 깊이 영상을 이용하여 좌 영상(V2')도 목표 시점으로 투영하여 가상시점 영상($V_{Li}$, $i=1,2,3,4,5,6,7$) 7장을 합성하게 된다.

이 때 홀 영역은 0, 그렇지 않은 영역은 0으로 하는 이진 마스크 영상($M_{Li}$, $i=1,2,3,4,5,6,7$)을 만들게 된다. 동일한 방법으로 우 영상(V3')을 이용하여 가상시점 영상($V_{Ri}$, $i=1,2,3,4,5,6,7$) 7장과 이진 마스크 영상($M_{Ri}$, $i=1,2,3,4,5,6,7$)도 생성하게 된다. 각각의 중간 시점 영상은 식 (5)를 이용하여 합성한다.

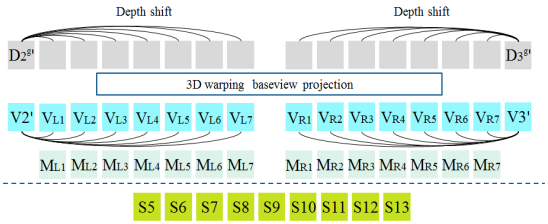



**Fig. 7** Synthesis of intermediate view by interpolation

$$S_{(i+5)}(x,y)=\begin{cases}(1-\alpha)\,V_{Li}(x,y)+\alpha\,V_{Ri}(x,y)\\(M_{Li}(x,y)=1\ \text{and}\ M_{Ri}(x,y)=1)\\ \\V_{Li}(x,y)\\(M_{Li}(x,y)=1\ \text{and}\ M_{Ri}(x,y)=0)\\ \\V_{Ri}(x,y)\\(M_{Li}(x,y)=0\ \text{and}\ M_{Ri}(x,y)=1)\\ \\common\ hole\ inpainting\\(M_{Li}(x,y)=0\ \text{and}\ M_{Ri}(x,y)=0)\end{cases}$$

$$(i=1,2,3,4,5,6,7\ \ S_5=V2',S_{13}=V3' \tag{5}$$

여기서 $V_{Li}(x,y)$와 $V_{Ri}(x,y)$는 각각 스테레오 영상의 좌, 우 영상을 목표 시점으로 투영한 가상시점 영상의 $x$, $y$ 위치에서의 화소 값, $M_{Li}(x,y)$과 $M_{Ri}(x,y)$은 이진 마스크 영상의 $(x,\ y)$ 위치에서의 화소 값 그리고 $S_{(i+5)}(x,y)$는 합성된 영상의 $(x,\ y)$ 위치에서의 화소 값을 각각 의미한다.

즉, 같은 목표 시점에 대해 좌, 우 마스크를 비교해서 해당 화소 값이 모두 1인 경우는 좌 영상(V2')을 투영한 가상 시점 영상(VLi)과 우 영상(V3')을 투영한 가상 시점 영상(VRi)의 화소를 모두 이용하여 보간한다. 이 때 가까운 시점 영상의 상관도가 크다고 판단하고 그 영상의 화소는 가중치를 크게 한다. 하나의 마스크 영상에서만 화소의 값이 1이라면, 한 시점에서 보이는 영역이 다른 시점에서는 보이지 않는 가려짐 영역으로 판단할 수 있기 때문에 이 마스크에 대응하는 가상 시점 영상의 화소 값으로 보간한다. 좌, 우 마스크 영상이 모두 0의 값을 갖는 경우에는 공통 홀로 판단하고 추후에 인페인팅(Inpainting) 기법을 통해 보간한다[11]. 하지만 좌, 우 가상 시점을 모두 이용하는 경우 복잡도가 높은 영역에서 불일치의 문제가 발생할 수 있다. 예를 들면 그림 8(a)의 영상에서 빨간색 원안의 갈대밭과 같이 복잡도가 높은 영역은 좌 영상과 우 영상을 같은 목표 시점으로 투영했을 때 완벽하게 일치하지 않아 오류가 발생하게 된다. 이 문제를 해결하기 위해서 좌, 우 마스크의 값이 모두 1인 경우에 식 (6)과 같은 부가적인 조건을 추가한다.

$$S_{(i+5)}(x,y)=\begin{cases}(1-\alpha)\,V_{Li}(x,y)+\alpha\,V_{Ri}(x,y) \\ \left|\,V_{Li}(x,y)-V_{Ri}(x,y)\,\right|_{(R,G,B)}<th \\ \\ V_{Ri}(x,y) \\ \left|\,V_{Li}(x,y)-V_{Ri}(x,y)\,\right|_{(R,G,B)}>th \\ \text{and } \alpha>0.5 \\ \\ V_{Li}(x,y) \\ \left|\,V_{Li}(x,y)-V_{Ri}(x,y)\,\right|_{(R,G,B)}>th \\ \text{and } \alpha\le 0.5\end{cases} \tag{6}$$

여기서 $V_{Li}(x,y)$와 $V_{Ri}(x,y)$는 각각 스테레오 영상의 좌, 우 영상을 목표 시점으로 투영한 가상시점 영상의 $(x,y)$ 위치에서의 화소 값, $\alpha$는 가중치 그리고 $th$는 임계값이다. 투영된 좌, 우 가상 시점 영상의 화소 값 차이를 비교하여 임계값보다 작은 경우에는 식 (5)와 동일하게 시점에 따른 가중치를 활용하고 화소 값 차이가 임계값보다 큰 경우에는 불일치로 판단하고 가까운 시점의 화소 값만 이용하여 보간 한다.

(a)

(b)

**Fig. 8** (a) Mismatching problem (b) Result of solving the mismatching problem

그림 8(b)은 그림 8(a)는 화소 값 차이 조건을 적용하여 가까운 가상 시점 영상만을 이용하여 불일치 문제를 해결한 결과를 보여준다.

### 3.2. 외삽(Extrapolation)

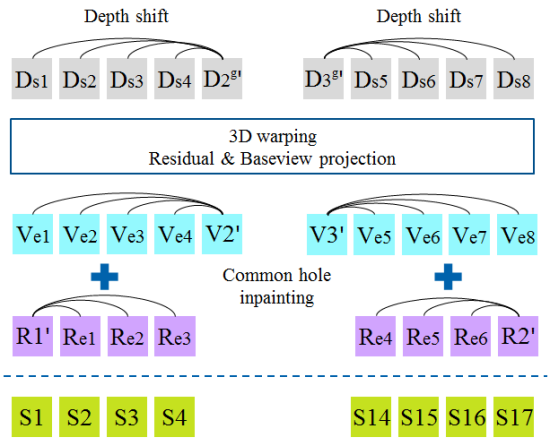



**Fig. 9** Synthesis of intermediate view by extrapolation

외삽 기법에서는 그림 9와 같이 복호화된 global 깊이 영상($D_2^{g'}$, $D_3^{g'}$)을 각 목표 시점으로 투영하고 2.1절에서와 같이 홀 영역을 채운다. 그리고 이 투영된 깊이 영상을 이용하여 복호화된 residual 영상 R1', R2' 및 스테레오 영상 V2', V3'을 목표 시점으로 투영하여 $R_{ei}(i=1\sim6)$과 $V_{ei}(i=1\sim8)$을 생성한다. 이때 시점 이동된 스테레오 영상에서 2.2절에서 언급했던 것과 같이 좌우측 경계 부분과 가려짐 영역에 해당하는 홀 영역이 발생할 수 있다. 이 경우 같은 목표 시점으로 투영된 residual 영상을 이용하여 비워있는 영역을 채운다. 이 후에도 남은 홀 영역은 공통 홀(common hole)로 판단하고 인근 화소들을 이용한 인페인팅(Inpainting) 기법을 이용하여 보간하게 되고 최종적으로 합성영상 $S_i(i=1\sim4,14\sim17)$을 생성하게 된다.

## Ⅳ. 실험결과

본 논문에서 제안하는 다시점 영상 합성 기법의 성능 비교를 위해서 실사 시퀀스인 MPEG 3DV CfP의

'Poznan_street', NICT에서 제공하는 CG 시퀀스, 'Shark'을 실험 영상으로 사용한다. 'Poznan_street' 시퀀스의 경우에는 PNSR을 측정할 기준 영상이 위성 영상(V1, V4) 밖에 없기 때문에 이 경우에만 성능비교를 하고 'Shark' 시퀀스의 경우에는 모든 중간 시점 영상에 대하여 측정을 한다. HEVC 압축 모드는 시퀀스의 프레임 율(frame-rate)과 동일한 주기로 I-프레임을 삽입하는 Random-access 방식을 이용한다.

### 4.1. 객관적 화질 비교

입력되는 4 시점의 영상과 이에 상응하는 4 시점의 깊이 영상을 모두 전송하는 MVD 포맷과 본 논문에서 제안하는 스테레오 영상, global 깊이 영상 및 residual 영상을 전송하는 포맷의 성능을 비교한다. 즉, 각 포맷의 데이터 비트율 대비 합성된 중간 시점 영상과 기준 영상 간의 PNSR을 측정하여 성능을 비교하였다.

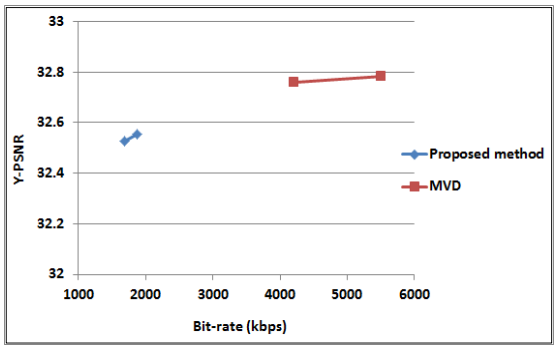



**Fig. 10** PSNR comparisonof synthesized intermediate view ('Poznan_street')

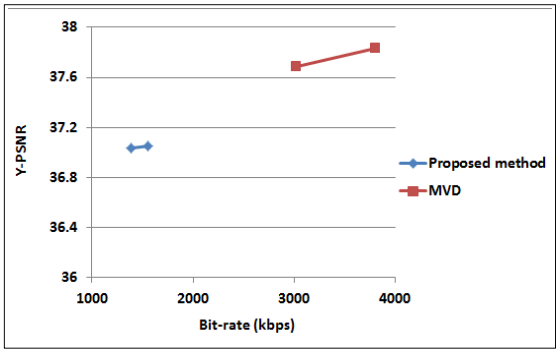



**Fig. 11** PSNR comparisonof synthesized intermediate view ('Shark')

그림 10과 11은 MVD 포맷과 제안한 포맷을 데이터 비트율 대비 PSNR을 측정한 결과를 각각 보여준다. MVD 포맷을 이용한 경우 객관적 수치가 조금 높지만 전송하는 데이터 량이 훨씬 많다는 것을 알 수 있다. 이것은 제안하는 기법이 훨씬 적은 데이터를 가지고 비슷한 화질의 중간 시점의 영상을 합성할 수 있다는 것을 의미한다.

### 4.2. 주관적 화질 비교

다음은 두 가지 기법을 이용하여 생성된 중간 시점 영상의 화질을 주관적으로 판단하고자 한다. 그림 12와 13은 각각 'Poznan_street', 'Shark' 시퀀스에 대하여 MVD 포맷과 제안한 포맷을 통해 합성한 중간 시점 영상이다. 본 논문에서 제안하는 포맷을 이용하여 합성한 경우 스테레오 영상을 목표 시점으로 투영 시에 발생하는 홀 영역을 residual 영상으로 채우는 방법이므로 깊이 영상이 정교하지 않으면 객체의 경계 부분에 오류가 발생할 수 있다. 이것은 4 시점의 깊이 영상을 한 시점의 global 깊이 영상으로 전송하는 포맷의 문제점이다. 특히 CG 시퀀스인 'Shark'의 경우 이런 현상이 좀 더 두드러지게 나타나는 것을 알 수 있다. 하지만 적은 데이터 량을 이용함에도 MVD 포맷을 이용한 결과와 비교했을 때 전체적으로 비슷하거나 더 좋은 화질의 영상을 합성할 수 있었다.

(a)

(b)

**Fig. 12** Synthesized intermediate view (a) the proposed method (b) MVD

(a)

(b)

**Fig. 13** Synthesized intermediate view (a) the proposed method (b) MVD

## Ⅴ. 결 론

본 논문에서 제안하는 다시점 영상 합성 기법은 무안경식 다시점 디스플레이를 위한 것이다. 송출 단에서 입력받은 스테레오 영상 및 위성영상, 그리고 대응하는 깊이 영상으로부터 부호화 효율을 높이기 위해 데이터 량을 줄인 global 깊이 영상과 residual 영상으로 변환하여 부호화한 후 전송하였다. 수신 단에서는 복호화 과정을 거쳐 전송된 영상으로부터 다시점 영상을 합성하게 된다. 이때 원래의 스테레오 영상 사이의 중간시점 영상은 내삽(Interpolation) 기법으로, 부가적으로 주어지는 위성 영상과 스테레오 영상 사이의 중간 시점 영상은 외삽(Extrapolation) 기법으로 합성하였다. 제안한 기법에서는 일반적인 중간 시점 영상 합성 기법에 추가 조건을 부여하여 영상내의 복잡한 영역에서 발생하는 좌, 우 영상의 화소 불일치 오류를 줄였다. 본 논문에서 제안하는 기법의 성능을 MVD 포맷과 비교하였을 때 부호화 시 적은 데이터 량을 이용함에도 불구하고 합성된 다시점 영상의 화질은 비슷하거나 더 좋다는 결과를 실험을 통해 확인할 수 있었다.


## ACKNOWLEDGMENTS

This work was supported by Institute for Information &communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (No. B0101-15-295, Development of UHD Realistic Broadcasting, Digital Cinema, and Digital Signage Convergence Service Technology)


## REFERENCES


[ 1 ] Z.-F. Wang and Z.-G. Zheng, “A region based stereo matching algorithm using cooperative optimization,” *in proceeding of the IEEE Conference on Computer Vision and Pattern Recognition,* Anchorage, pp. 1-8, 2008.

[ 2 ] X. Mei, X. Sun, M. Zhou, S. Jiao, H. Wang and X. Zhang, “On building an accurate stereo matching system on graphics hardware,” *Computer Vision Workshops (ICCV Workshops),* 2011 IEEE International Conference on, pp. 467-474, Nov. 2011.

[ 3 ] J. Liu, C. Li, F. Mei and Z. Wang, “3D entity-based stereo matching with ground control points and joint second-order smoothness prior,” *The Visual Computer,* July 2014

[ 4 ] P.-J. Lee and Effendi, “Nongeometric Distortion Smoothing Approach for Depth Map Preprocessing,” *IEEE TRANSACTIONS ON MULTIMEDIA,* vol. 13, no. 2, pp. 246-254, April 2011.

[ 5 ] C.-H. Wei, C.-K. Chiang and S.-H. Lai, “Iterative depth recovery for multi-view video synthesis from stereo videos,” *Asia-Pacific Signal and Information Processing Association,* 2014 Annual Summit and Conference (APSIPA), pp. 1-8, Dec. 2014.

[ 6 ] I. Daribo, and H. Saito, “A Novel Inpainting-Based Layered Depth Video for 3DTV,” *IEEE TRANSACTIONS ON BROADCASTING,* vol. 57, no. 2, pp. 533-541, June 2011.

[ 7 ] A. Criminisi, P. Perez and K. Toyama, “Region filling and object removal by exemplar-based image inpainting,” *IEEE TRANSACTIONS ON IMAGE PROCESSING,* vol. 13, no. 9, pp. 1200-1212, Sept. 2004

[ 8 ] B. Bartczak, P. Vandewalle, O. Grau, G. Briand, J. Fournier, P. Kerbiriou, M. Murdoch, M. Muller, R. Goris, R. Koch and R. van der Vleuten, “Display-Independent 3D-TV

Production and Delivery Using the Layered Depth Video Format," *IEEE TRANSACTIONS ON BROADCASTING,* vol. 57, no. 2, pp. 477-490, June 2011.

[ 9 ] T. Senoh, Y. Ichihashi, H. Sasaki and K. Yamamoto, "Simple Multi-View Coding with Depth Map", *3D Research,* vol. 4, no.2, June 2013.

[10] R. C. Gonzalez and R. E. Woods, "Digital Image Processing," *Pearson Education,* Inc, 2010.

[11] M. Bertalmio, A. L. Bertozzi and G. Sapiro, "Navier-Stokes, Fluid Dynamics, and Image and Video Inpainting," *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on,* vol. 1, pp. 355-362, 2001.

**최인규(In-Kyu Choi)**

2014년 2월 광운대학교 전자공학과 학사
2014년 3월 광운대학교 전자공학과 석사
※관심분야 : 컴퓨터 비전, 스테레오 매칭, 영상 신호 처리

**정원식(Won-sik Cheong)**

1994년 경북대학교 전자공학과 석사
2000년 경북대학교 전자공학과 박사
2000년 ~ 현재 한국전자통신연구원 입체방송연구실 책임연구원
※관심분야 : 3DTV 방송시스템, 모바일 방송, 영상처리 및 압축, MPEG 등

**이광순(Gwangsoon Lee)**

1993년 : 경북대학교 전자공학과 졸업 (학사)
1995년 : 경북대학교 전자공학과 졸업 (석사)
2004년 : 경북대학교 전자공학과 졸업 (박사)
2001년 ~ 현재 : 한국전자통신연구원 입체방송연구실 실장
※관심분야 : 3DTV 방송시스템, 무안경 3D, 라이트필드 영상처리 등

**유지상(Jisang Yoo)**

1985년 2월 서울대학교 전자공학과 학사
1987년 2월 서울대학교 전자공학과 석사
1993년 5월 Purdue Univ. EE, ph.D
1997년 9월 ~ 현재 광운대학교 전자공학과 교수
※관심분야 : 웨이블릿 기반 영상처리, 영상압축, 영상인식, 비선형 신호처리