

# 가상시점 영상 생성을 위한 경계 잡음 제거와 홀 채움 기법

고민수\*, 유지상<sup>o</sup>

## Boundary Noise Removal and Hole Filling Algorithm for Virtual Viewpoint Image Generation

Min Soo Ko\*, Jisang Yoo<sup>o</sup>

요약

본 논문에서는 3D 워핑(warping) 기법을 이용하여 가상시점 영상생성 시 화질 개선을 위하여 경계 잡음(boundary noise)을 제거하고 홀(hole) 채움을 하는 새로운 기법을 제안한다. 경계 잡음은 가상시점 영상 합성 시 기준 시점 영상과 깊이 영상 내 객체의 경계 불일치로 발생되며 홀은 기준시점 영상에서 보상할 수 없는 가려짐 영역(occlusion)으로 정의된다. 제안하는 기법에서는 경계 잡음 제거를 위해 먼저 배경 화소들의 평균과 절대 값 비교를 통해 경계 잡음에 해당되는 화소를 검출하고 검출된 화소를 홀 영역으로 확장한다. 경계 잡음 영역이 포함된 확장된 홀 영역은 나선형 가중 평균(spiral weighted average) 기법과 기울기 탐색(gradient searching) 기법을 혼용하여 채우게 된다. 나선형 가중 평균 기법은 깊이 정보를 사용함으로 객체 정보를 최소로 사용하지만 결과 영상이 번지는 단점이 있다. 기울기 탐색 기법은 영상의 기울기를 이용하여 세밀한 부분을 보존할 수 있는 장점이 있다. 따라서 각각의 결과를  $\alpha$  가중치로 조합하여 생성된 가상 시점은 두 기법의 장점을 동시에 적용하기 때문에 좋은 화질을 얻을 수 있다. 실험을 통해 제안하는 기법의 성능이 기존의 다른 기법보다 우수하다는 것을 확인하였다.

**Key Words** : 3D Warping, Virtual viewpoint image, Freeview-point TV, Multi-view image, Hole filling, 3차원 워핑, 가상 시점 영상, 자유시점 TV, 다시점 영상, 홀 채움

### ABSTRACT

In this paper, performance improved hole-filling algorithm including boundary noise removing pre-process which can be used for an arbitrary view synthesis with given two views is proposed. Boundary noise usually occurs because of the boundary mismatch between the reference image and depth map and common-hole is defined as the occluded region. These boundary noise and common-hole created while synthesizing a virtual view result in some defects and they are usually very difficult to be completely recovered by using only given two images as references. The spiral weighted average algorithm gives a clear boundary of each object by using depth information and the gradient searching algorithm is able to preserve details. In this paper, we combine these two algorithms by using a weighting factor  $\alpha$  to reflect the strong point of each algorithm effectively in the virtual view synthesis process. The experimental results show that the proposed algorithm performs much better than conventional algorithms.

※ 본 연구는 2012년도 지식경제부 및 한국산업기술평가위원회의 산업원천기술개발사업(정보통신)의 일환으로 수행한 연구[10039199, 인지품질 기반 스케일러블 3D 비디오 코덱 핵심 기술 연구]로서, 관계부처에 감사 드립니다.

◆ 주저자 : 광운대학교 전자공학과 디지털미디어 연구실, kmsqwet@kw.ac.kr, 정회원

◦ 교신저자 : 광운대학교 전자공학과 디지털미디어 연구실, jsyoo@kw.ac.kr, 종신회원

논문번호 : KICS2012-03-115, 접수일자 : 2012년 3월 7일, 최종논문접수일자 : 2012년 7월 31일

## I. 서 론

2009년 3D 아바타 영화의 흥행 성공으로 3차원(3D) 미디어에 대한 관심 증가하고, 3D 산업은 영화, 게임, 방송에 이르기 까지 다양한 분야에서 새로운 미디어 서비스로 자리 잡고 있다. 그러나 최근에는 안경착용에 따른 불편함, 시각적 피로, 다양한 콘텐츠의 부족 등으로 기대만큼 3D 산업이 활성화되고 있지 않다는 평가도 나오고 있는 상황이다. 따라서 3D 산업의 지속적인 활성화를 위해서는 2D 콘텐츠와 같이 쉽게 3D 콘텐츠를 획득할 수 있고, 안경 없이 고화질로 3D 영상을 즐길 수 있어야 하며, 피로가 최소화되어야 한다. 이러한 조건을 만족하기 위해서는 무안경식 다시점 3D 영상을 이용한 3차원 미디어 서비스 기술 개발이 필요하다<sup>[1]</sup>.

일반적으로 스테레오 3D 영상을 얻기 위해서는 두 대의 카메라를 이용한다. 하지만 스테레오의 양안 방식은 반드시 안경을 착용하여야 하며, 스테레오 무안경 방식의 경우에는 입체감을 느낄 수 있는 시점이 하나로 제한되어 시청자가 제한된 시역을 벗어나면 입체감을 느낄 수 없거나 어지러움을 느낄 수 있다. 다시점 무안경 디스플레이는 스테레오 디스플레이보다 시점의 개수가 증가하기 때문에 관찰자의 시역이 확대되어 보다 자연스러운 입체 디스플레이 구현이 가능하다. 그러나 시점의 개수가 증가함에 따라 데이터양도 증가하기 때문에 실시간 처리가 힘들고 초고속, 광대역의 전송채널이 요구된다. 이러한 문제를 해결하기 위해 수신단에서 스테레오 영상으로부터 원하는 시점 수만큼의 다시점 영상을 합성하기 위한 연구가 활발히 진행되고 있다. 무안경 다시점 3D 디스플레이는 시점 수에 맞는 다시점 영상을 만들어 정해진 포맷으로 다중화하여 입력하거나, 한 장의 영상과 깊이 영상을 입력하여 렌더링 엔진에 의해 다시점 영상을 생성하는 방식이 대부분이었다. 하지만 최근에는 보다 쉽게 다시점 3차원 영상을 제공하기 위해 2차원 영상으로부터 다시점 영상을 생성하거나, 좌우 스테레오 영상으로부터 다시점 영상을 생성하는 기능을 채택하는 경우가 늘어나고 있는 추세이다<sup>[2]</sup>.

최근에 MPEG에서도 새로운 3D 부호화 기술의 요구에 따라 다시점 비디오 기반의 MVC(multi-view video coding) 기술과 FTV(freeview-point TV) 기술에 대한 연구가 진행되었다. MVC 기술은 다시점 영상을 보다 효과적으로 부호화하기 위한 기법으로 개발되었다. FTV 기

술은 사용자가 기존의 다시점 영상 사이의 임의 시점을 선택하여 시청할 수 있는 기술로 논의되었으나, 2007년 4월에 열린 JVT 회의 이후에 그 범위를 3DTV 응용을 위한 FTV/3DV 표준 시스템 모델로 확장하여 개발되고 있다. 먼저, 수신단에서 다시점 색상 영상과 각 색상 영상에 대응하는 깊이 데이터(MVD: multi-view video plus depth)를 획득하고, 이를 부호화해 전송한다. 전송된 MVD 데이터는 복호화 과정을 거쳐 수신단에서 중간 영상 생성을 통해 다시점 3D 비디오를 구성하는 방식이다. 따라서 고화질의 중간 영상 생성을 위해 정확한 깊이 정보를 확보하고 중간 영상을 합성하는 방법에 대한 연구가 필요하게 되었다. 이를 위해 MPEG에서는 깊이 추정 소프트웨어(DERS; depth estimation reference software)와 가상 시점 영상 생성 소프트웨어(VSRS; view synthesis reference software)를 이용한 결과에 대한 평가를 실시하고 있다<sup>[3,4]</sup>.

가상시점(virtual viewpoint) 영상 합성을 위한 방법에는 크게 스테레오 영상의 변이정보를 이용하는 양방향 선형 보간법(bidirectional linear interpolation)과 깊이 정보를 이용하는 3D 워핑(3D warping) 기법 등이 있다. 양방향 선형 보간법은 변이(disparity) 정보를 이용하여 기준 시점(reference view) 영상 사이의 제한된 시점의 중간 시점 영상을 생성하는 방법이다. 3D 워핑은 기준 영상의 깊이 정보와 카메라 파라미터를 이용하여 가상 시점에 해당하는 가상 카메라의 파라미터를 정의하고, 영상의 실제 3D 좌표를 산출한 뒤 가상 카메라의 위치에 재투영하는 방법으로, 정의한 카메라 파라미터를 통해 기준 영상 사이의 중간시점 영상뿐만 아니라 임의 가상 시점 영상을 생성하는 등 보다 자유로운 시점을 제공할 수 있다.

본 논문에서는 고화질의 가상 시점 영상을 생성하기 위해 3D 워핑 기법을 사용한다. 먼저 기준 영상 화소들을 실세계 좌표(world coordinate)로 산출하고 원하는 가상 시점으로 재투영 하게 된다. 이때 가상시점 영상 생성 시 객체의 이동으로 생기는 가려짐 영역(occlusion)에 의해 홀(hole)이 발생한다. 그리고 기준 시점 영상과 깊이 영상 내 객체의 경계 위치가 일치하지 않아 경계 잡음(boundary noise)이 발생할 수 있다. 이러한 홀과 경계 잡음은 생성된 가상시점 영상의 화질을 저하시키는 요인이 되므로 본 논문에서는 이러한 요소를 제거하여 가상 시점 영상의 화질을 개선하기 위한 후처리 기법을 제한하고자 한다.

제안하는 기법에서는 먼저 홀에 근접하여 생기는 경계 잡음을 제거하기 위해 배경 화소들의 평균을 구하고 절대 값 비교를 통해 경계 잡음 화소를 검출한다. 경계 잡음으로 판단된 화소는 홀 영역으로 확장하여 홀 채움 기법을 통해 홀과 함께 채워지게 된다. 공통 홀(common-hole) 영역은 기준 시점 영상에는 정보가 없어서 참조 할 수 없는 영역이다. 따라서 홀 영역 주변의 화소를 통해 그 값을 예측하여야 한다. 기존에는 이러한 공통 홀(common-hole)을 채우기 위해 선형 보간법과 인페인팅(in-painting) 기법이 많이 사용 되었다. 선형 보간법은 공통 홀의 끝단에 존재하는 화소 값을 홀 영역의 구간 길이로 나누어 선형적으로 값을 더하거나 빼는 과정을 통해 선형 보간을 수행하는 방법이다. 이 방법은 수행속도가 빠르지만 채워진 공통 홀의 품질이 좋지 못한 단점이 있다. 인페인팅 기법은 영상의 손상된 영역을 복원하기 위해 고안된 기법으로 손상된 영역의 주위의 색상정보로부터 추측하여 최대한 자연스럽게 복원하는 방법이다. 현재 인페인팅은 화소 단위의 복원 방법 및 질감 합성 방법 등 다양한 방법들이 연구되고 있다. 하지만 인페인팅 기법은 주변 화소의 색상 정보만을 이용하기 때문에 객체와 배경 영역을 구분하지 못하고, 따라서 채워진 영역이 부자연스럽게 보이는 단점이 있다<sup>5,6)</sup>.

본 논문에서는 이러한 문제들을 해결하기 위해 깊이 정보를 이용하여 효율적으로 공통 홀을 채우는 나선형 가중 평균 기법(spiral weighted average algorithm)을 제안한다. 하지만 나선형 가중 평균 기법은 일종의 평균 필터이므로 홀 채움 후에 번짐(blur) 현상이 발생한다. 본 논문에서는 채워진 홀의 번짐 현상을 줄이고 영상의 연속성을 유지하기 위해 기울기 탐색 기법(gradient searching algorithm)을 적용한다. 결국 최종 홀 채움 기법은 나선형 가중 평균 기법과 기울기 탐색 기법을 최적의 가중치로 조합하여 두 기법의 장점을 모두 갖도록 설계한다. 실험을 통해 기존의 기법들과 비교를 통하여 제안하는 기법의 성능이 우수하다는 것을 확인하였다.

본 논문의 II장에서는 카메라의 기하학적 구조에 기반을 둔 3D 워핑 기법에 대해 설명하고, III장에서는 고화질 가상시점 영상을 생성하기 위한 후처리 기법을 제안한다. IV장에서는 실험을 통하여 제안한 기법의 성능을 평가하며, V장에서 결론을 맺는다.

## II. 3D 워핑을 이용한 가상시점 영상생성 기법

3D 워핑(3D warping)이란 카메라의 기하학적 구조에 기반을 두고 영상의 깊이 정보와 카메라의 내부 파라미터를 이용하여 영상의 실제 좌표(세계 좌표계; world coordinate)를 산출한 후, 가상 시점을 위한 카메라의 위치로 재 투영하여 원하는 가상 시점 영상을 생성하는 과정이다.

그림 1(a)는 핀홀 카메라(pin-hole camera) 모델을 3D 구조로, 그림 1(b)는 2D 구조로 각각 표현한 것이다. 일반적으로 핀홀 카메라의 상은 Z축의 -f 위치에 역상으로 생긴다. 하지만 이 경우 3D 좌표 상에서 해석하는 것이 쉽지 않기 때문에 영상이 맺히는 평면을 Z축 상의 카메라 초점 거리(focal length) f로 옮겨 해석한다. 실제로 카메라 좌표계(camera coordinate)에서 물체의 3D 좌표가 영상 평면(image plane)에 투영되는 관계는 그림 1(b)와 같이 삼각 비례 법으로 해석이 가능하다.

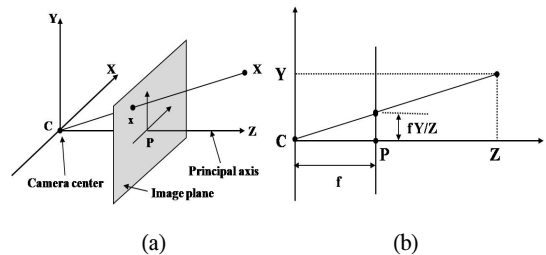


그림 1. 핀홀카메라의 기하학적 구조  
(a) 3D 구조 (b) 2D 구조  
Fig. 1. Geometrical structure of pin-hole camera  
(a) 3D structure (b) 2D structure

3D 실세계 좌표계(세계 좌표계; world coordinate)를 카메라 좌표계(camera coordinate)로 변환하기 위해서는 좌표계의 회전(rotation)과 이동(translation) 변환이 필요하다. 이 두 가지 변환은 카메라의 외부 파라미터(extrinsic parameter)로 정의된다. 카메라 좌표계로 변환된 물체의 좌표는 영상 평면(image plane)의 좌표로 다시 변환된다. 이 변환 관계는 그림 1(b)와 같이 초점 거리 f를 포함하는 카메라의 내부 파라미터(intrinsic parameter)로 정의될 수 있다. 결국 카메라의 외부 파라미터와 내부 파라미터를 이용한 변환 관계에 의하여 식 (1)과 같이 실세계 좌표계에 존재하는 객체의 좌표가 영상 평면상의 좌표로 변환되는 것이다.

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = K[R|T] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

여기서  $x, y$ 는 영상 평면에 투영된 물체의 2D 좌표를 나타내며,  $K$ 는 카메라 내부 파라미터(intrinsic parameter),  $R$ 은 카메라의 회전 행렬,  $T$ 는 카메라의 이동 벡터를 나타내며,  $X, Y, Z$ 는 실세계좌표계의 좌표를 의미한다.  $K[R|T]$ 를 투영 행렬(projection matrix)이라고 정의한다.

식 (1)의 행렬 연산 과정을 식 (2)와 같이 역으로 계산 하면 영상 좌표를 실세계 좌표로 역변환 할 수 있다. 이 때 실제 깊이 값  $Z$ 를 구하기 위해서는 식 (3)과 같이 시차 정보(disparity)  $D$ 를 이용하게 된다.

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = K[R|T] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \Rightarrow K^{-1} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = R \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} + T \quad (2)$$

$$\Rightarrow R^T K^{-1} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} - R^T T = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$Z(i, j) = \frac{1.0}{\left( \frac{D(i, j)}{255.0} \times \left( \frac{1.0}{MinZ} - \frac{1.0}{MaxZ} \right) + \frac{1.0}{MaxZ} \right)} \quad (3)$$

여기서  $Z(i, j)$ 와  $D(i, j)$ 는 각각 영상 내  $(i, j)$  좌표에서의 깊이값과 시차값이고,  $MinZ$ 와  $MaxZ$ 는  $Z$ 값이 가지는 최소, 최대값을 의미한다.

그리고 가상 시점을 생성하기 위해 가상 시점 카메라의 카메라 파라미터를 정의한다. 그리고 변환된 실세계 좌표와 가상 시점 카메라의 파라미터를 식 (1)을 적용하면 식 (4)와 같이 나타낼 수 있다. 식 (4)를 통해 실세계 좌표가 가상시점 영상의 좌표로 변환 되게 되고 이를 통해 가상 시점 영상을 생성 할 수 있다.

$$\begin{bmatrix} x_v \\ y_v \\ 1 \end{bmatrix} = K_v[R_v|T_v] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (4)$$

여기서  $x_v, y_v$ 는 생성된 가상 시점 영상 평면상의 좌표를 의미하며,  $K_v, R_v, T_v$ 는 각각 가상 시점 카메라의 내부 파라미터, 회전 행렬, 이동 벡터를 나타낸다. 따라서 가상 시점 영상을 생성하기 위해서는 해당 위치의 가상 카메라의 내, 외부 파라미터를 먼저 구해야 한다. 식 (5)는 좌, 우 시점 간의 카메라 거리를 1로 하고 가상 시점의 카메라와 두 기준시점 사이의 거리

비율인  $Ratio$ 에 따라 선형적으로 보간 하는 방법을 나타내었다.

$$\begin{aligned} K_v &= (1 - Ratio) \times K_L + Ratio \times K_R \\ R_v &= (1 - Ratio) \times R_L + Ratio \times R_R \\ T_v &= (1 - Ratio) \times T_L + Ratio \times T_R \end{aligned} \quad (5)$$

여기서  $V$ 는 가상 시점,  $L$ 은 좌 시점,  $R$ 은 우 시점 카메라를 각각 의미한다<sup>7,8)</sup>.

### III. 제안하는 후처리 기법

그림 2(c)와 같이 기준 시점 영상을 2절에서 설명한 3D 워핑 기법을 통해 가상 시점 영상을 생성하게 되면 가려짐 영역(occlusion)이 발생한다. 가려짐 영역은 기준 시점 영상에는 존재하지 않기 때문에 발생하며 홀(hole)의 형태로 나타나게 된다.

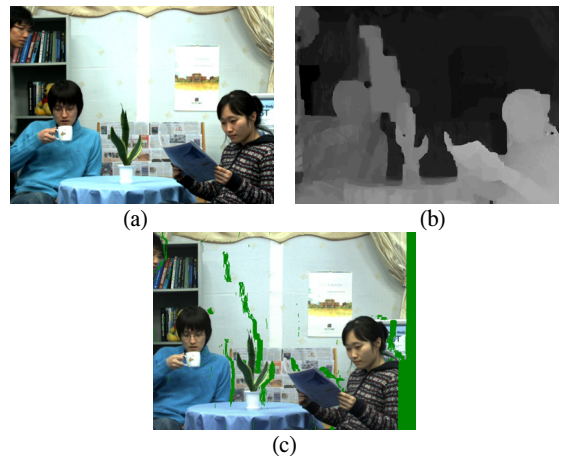


그림 2. 3D 워핑을 이용한 가상 시점 영상 생성  
(a) 기준시점 영상 (b) 깊이영상 (c) 가상 시점 영상  
Fig. 2. Virtual viewpoint image synthesis using 3D warping  
(a) reference image (b) depth image (c) virtual viewpoint image

좌, 우 영상을 기준시점으로 하고 가상시점 영상의 위치가 두 기준시점 영상의 사이에 위치할 경우에는 가려짐 영역은 기준시점 영상을 서로 참조하여 대부분 채울 수 있다. 하지만 가상시점 영상이 두 기준 시점 영상의 밖에 존재하거나 기준 시점 영상이 하나뿐 일 경우 가려짐 영역은 대부분 홀의 형태로 잔존하게 된다. 본 논문에서는 두 번째 경우와 같이 홀로 남는 영역이 많은 경우에 효과적으로 이 홀들을 채울 수 있는 후처리 기법을 제안하고자 한다. 그림 3은 제안하는 기법의 블록도이다.

제안하는 기법에서는 먼저 3D 워핑 기법을 통해 원하는 가상 시점 영상을 생성한다. 이때 발생하는 홀의 인접 지역에서 경계 잡음 영역을 검출한다. 경계 잡음으로 판단된 영역은 홀 영역으로 포함하여 확장하게 되고 확장된 홀 영역은 나선형 가중 평균(spiral weighted average) 기법과 기울기 탐색(gradient searching) 기법을 적절하게 조합하여 채운다<sup>9,10</sup>.

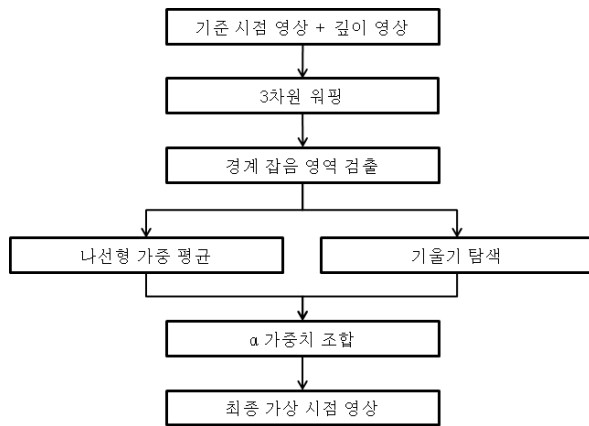


그림 3. 제안하는 기법의 블록도  
Fig. 3. Block diagram of the proposed algorithm

### 3.1. 경계 잡음 영역 검출

경계 잡음(boundary noise)은 기준시점 영상에 존재하는 객체의 경계와 이에 대응하는 깊이 영상 객체의 경계가 일치하지 않기 때문에 발생하는 잡음이다. 경계 잡음은 그림 4(a)와 같이 객체의 잔상과 유사한 형태로 나타나게 되는데 이 잡음을 제거하지 않고 홀을 채우게 되면 그림 4(b)와 같이 경계 잡음의 색상이 홀을 채울 때 사용하게 되어 채워진 홀의 품질이 저하된다. 본 논문에서는 먼저 그림 5와 같이 배경 영역에서 수평방향으로 일정 영역의 평균을 구하고 이를 홀 인접 지역 화소부터 비교하여 절대값 차이가 일정 임계치보다 크면 경계 잡음으로 판단한다. 경계 잡음이 연속으로 생길 경우를 생각하여 처음 화소가 경계 잡음으로 판단되면 다음 화소까지 비교를 하여 연속된 경계 잡음까지 검출한다. 검출된 경계 잡음은 홀 영역으로 포함시켜 제거한다. 그림 4(c)는 제안하는 기법을 통해 경계 잡음을 제거하고 홀을 채운 영상이다. 제거하지 않고 채운 영상과 비교하여 비교적 깨끗하게 홀이 채워진 것을 확인할 수 있다.

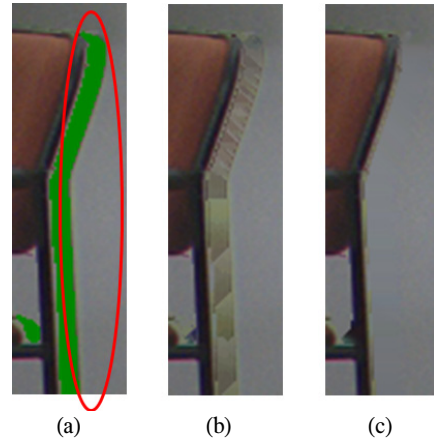


그림 4. (a) 경계 잡음 (b) 경계 잡음 제거 없이 홀 채움을 한 결과 (c) 제안하는 기법으로 경계 잡음을 제거하고 홀 채움을 한 결과

Fig. 4. (a) Boundary noise (b) Hole filling with boundary noise (c) Hole filling after removing the boundary noise with proposed algorithm



그림 5. 제안하는 경계 잡음 제거 기법  
Fig. 5. Proposed algorithm to remove boundary noise

### 3.2. 홀 채움 순서

그림 6은 기존의 홀 채움 방법으로 홀의 최 외곽 화소부터 시작하여 홀의 중심이 마지막에 채워지게 된다. 그림 6(a)와 같이 홀의 최 외곽끼리 맨 처음 순서로 채워지게 되며 중심으로 갈수록 나중에 채워지게 된다. 이 경우 홀 영역에 인접해 있는 객체의 색상 정보를 이용할 가능성이 커지게 되며 그림 6(b)와 같이 홀 채움 결과의 품질이 저하된다. 따라서 제안하는 기법에서는 가상시점 영상과 기준시점 영상과의 위치를 고려하여 배경에서 가까운 부분부터 채우는 방법을 사용한다. 그림 7은 가상 시점 영상이 기준 시점보다 우측에 있을 경우에 홀을 채우는 순서를 나타낸다. 이 경우 가상 시점이 기준 시점 보다 우측에 있으므로 객체의 우측에 홀이 생긴다. 따라서 홀의 우측이 배경과

가까운 영역이 되며 이 부분부터 홀을 채우게 된다. 그림 7(a)와 같이 배경에 인접한 홀 화소부터 먼저 채워지게 되고 객체에 가까운 홀 화소일수록 나중에 채워지게 된다. 그림 7(b)의 결과를 통해 제안하는 기법이 기존 방법보다 외부 객체의 영향을 덜 받는 것을 확인 할 수 있다.

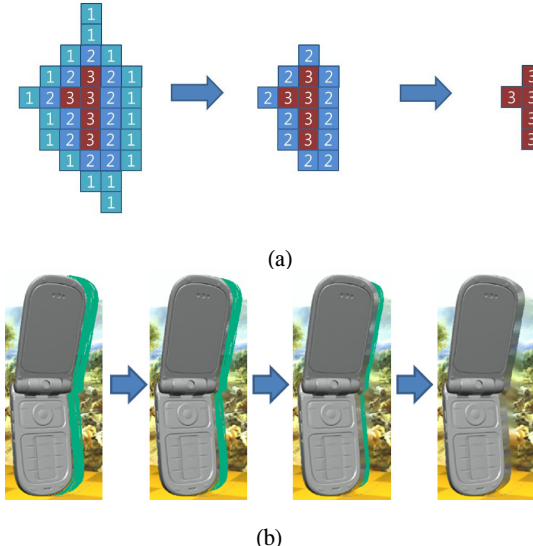


그림 6. (a) 홀을 외곽 화소부터 채우는 기법 (b) 결과 영상  
Fig. 6. (a) Hole filling from outside pixels (b) results

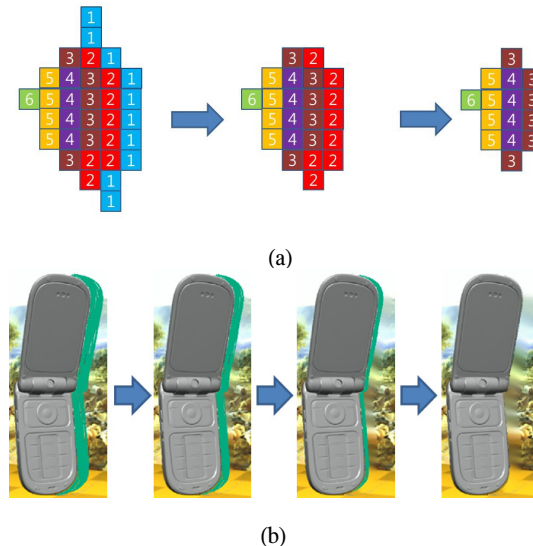


그림 7. (a) 제안하는 홀 채움 순서 (b) 결과 영상  
Fig. 7. (a) Proposed order of hole filling (b) results

### 3.3. 제안하는 홀 채움 기법

#### 3.3.1. 나선형 가중 평균 기법

가려짐 영역에서 채울 수 없는 영역은 공통 홀

(common-hole)의 형태로 잔존하는데 기준시점에서 참조할 수 있는 더 이상 정보가 없으므로, 생성된 가상시점 내에서 홀의 주변 화소정보를 이용하여 해당 화소의 값을 예측하여야 한다. 기존에는 이러한 공통 홀(common-hole)을 채우기 위해 인페인팅(in-painting) 기법이 많이 사용 되었다. 하지만 인페인팅 기법은 주변 화소의 색상 정보만을 이용하기 때문에 객체와 배경 영역을 구분하지 못하고, 따라서 채워진 영역이 부자연스럽게 보이는 단점이 있다. 본 논문에서는 이 문제를 해결하기 위해 깊이 정보를 이용하여 효율적으로 공통 홀을 채우는 나선형 가중 평균 기법(spiral weighted average algorithm)을 새로이 제안한다. 제안하는 기법에서는 3.2절의 그림 7에서 설명한 홀 채움 순서를 적용하여 성능을 개선하고 정확도를 높이도록 한다.

제안하는 기법에서는 나선형 탐색 기법을 사용하여 홀 화소와 가장 근접한 화소들을 후보로 사용 할 수 있도록 한다. 식 (6)은 거리의 가중치와 색상의 가중치를 사용하는 Bilateral filter와 유사하다. 하지만 색상 정보 대신 깊이 정보를 사용하여 홀 채움에 더 적합하다. 이는 거리의 가중치와 깊이 정보의 가중치를 사용하면 홀 화소에 근접하고 깊이 정보가 비슷한 화소일수록 높은 가중치 부여하여 배경과 비슷한 색상을 가지면서 객체의 색상 침범이 적어지도록 하기 때문이다.

그림 8은 제안하는 나선형 가중 평균 기법의 순서도이다. 먼저 홀의 한 화소를 기준 화소로 하여 이웃(8-neighbor) 화소가 갖는 가장 작은 깊이 값을 초기 깊이 값(initial depth)으로 결정한 후 그림 9와 같이 기준 화소를 중심으로 나선형으로 탐색을 수행하며 초기 깊이 값과 탐색 화소간의 깊이 값의 변화가 일정 임계치 이하인 화소들을 검출한다. 검출된 화소들의 색상과 깊이 값에 기준 화소와의 거리 차이에 따른 가중치를 곱하고 이 값들의 평균값 구하여 기준 화소의 색상 값과 깊이 값으로 할당하게 된다. 위와 같은 과정을 제안한 홀 채움 순서대로 모두 적용하여 공통 홀을 채우게 된다. 현재 순서의 홀이 모두 채워지면 차음부터 같은 과정을 반복한다. 식 (6)은 위의 과정에 설명한 공통 홀의 화소에 할당되는 새로운 색상 값과 깊이 값을 구하는 식이다.

$$ST(x, y) = \sum_{j=0}^m \left( \frac{e_j T_j(p, q)}{E} \right), SD(x, y) = \sum_{j=0}^m \left( \frac{e_j d_j(p, q)}{E} \right)$$

$$E = \sum_{i=0}^m e_i, \quad e = \frac{D(p, q)}{W(p, q)} \quad (6)$$

$$\text{where, } D(p,q) = \begin{cases} 1, & |Initail\_depth - d(p,q)| < th \\ 0, & \text{else} \end{cases}$$

여기서  $ST(x,y)$ 와  $SD(x,y)$ 는  $(x,y)$  좌표에서의 보상할 홀의 색상 값과 깊이 값이고,  $W(p,q)$ 는  $(p,q)$ 에서의 유클리디언 거리에 따른 가중치를 나타내며,  $D(p,q)$ 는  $(x,y)$ 좌표와  $(p,q)$ 좌표에서의 깊이 값 차이에 의한 가중치이다.  $T(p,q)$ 와  $d(p,q)$ 는  $(p,q)$ 에서의 색상 값과 깊이 값을 나타낸다<sup>[11,12]</sup>.

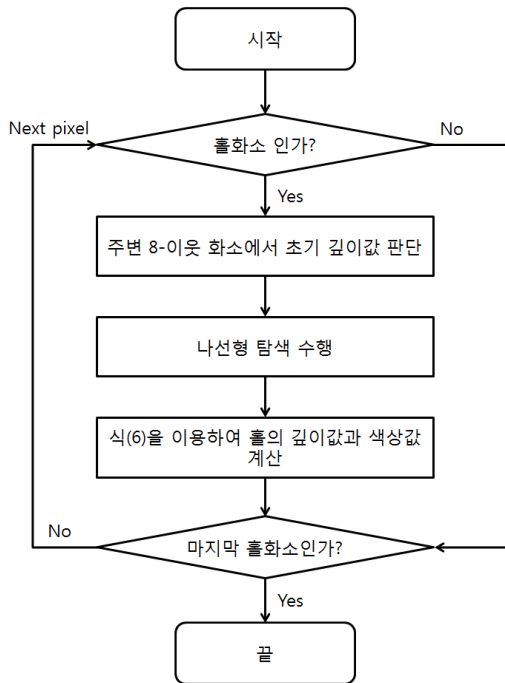


그림 8. 나선형 가중 평균 기법의 순서도  
Fig. 8. Flow chart of spiral weighted average algorithm

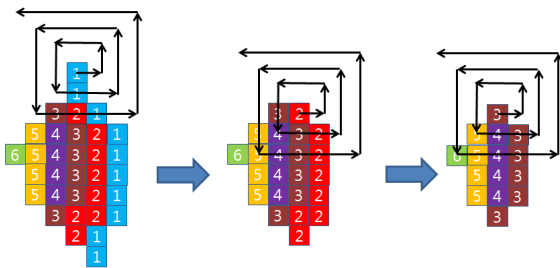


그림 9. 나선형 가중 평균 기법을 이용한 홀 채움 과정  
Fig. 9. Hole filling process by using spiral weighted average algorithm

### 3.3.2. 기울기 탐색 기법

나선형 가중 평균 기법은 일종의 평균 필터이므로 홀 채움 후 번짐(blur) 현상이 발생할 수 있다. 본 논문에서는 채워진 홀의 번짐 현상을 줄이면서 영상의 연속성을 유지하기 위해 기울기 탐색 기법(gradient

searching algorithm)도 적용한다. 기울기 탐색 기법은 홀 화소 주변의 화소 중 기울기가 가장 큰 화소를 찾는 기법이다. 따라서 홀이 기울기가 큰 방향의 색으로 채워지게 되어 영상 연속성을 유지 할 수 있다. 제안하는 기울기 탐색 기법은 다음과 같은 과정으로 구성 된다<sup>[9,10]</sup>.

- (1) 그림 10(a)와 같이 홀 화소 주변의 12 방향에 존재하는 화소들의 색상 값 차이로 기울기 값을 구하고, 최대 기울기를 갖는 방향의 화소를 찾는다.
- (2) 그림 10(b)와 같이 (1)에서 찾은 화소 주변의 최대 기울기 방향과 같은 세 방향 내에서 색상 값이 가장 비슷한 화소를 찾고 그 위치로 이동한다.
- (3) 탐색 범위 내에서 (2)의 과정을 반복한다.
- (4) (1)부터(3)의 과정에 찾은 모든 화소들의 평균을 구하여 공통 홀의 화소 값으로 할당한다.

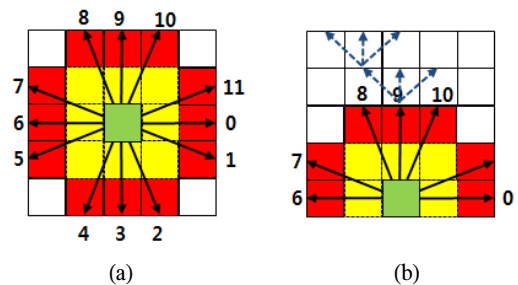


그림 10. 기울기 탐색 기법  
(a) 1 단계 (b) 2 단계  
Fig 10. Gradient searching algorithm  
(a) 1st step (b) 2nd step

### 3.3.3. 최종 홀 채움 과정

나선형 가중 평균 방법은 깊이 정보만을 이용하기 때문에 홀 채움 후에 번짐(blur) 현상이 나타나는 단점이 있다. 이러한 단점은 기울기 탐색 기법을 이용하여 영상의 연속성을 유지하도록 하여 보상할 수 있다. 따라서 나선형 가중 평균 방법과 기울기 탐색 기법으로 구한 각각의 화소 값을 식 (7)과 같이  $\alpha$  가중치를 적용하여 공통 홀의 최종 화소 값으로 결정하게 된다.

$$FT(x,y) = \alpha GT(x,y) + (1-\alpha)ST(x,y) \quad (7)$$

여기서  $GT(x,y)$ 와  $ST(x,y)$ 는 각각 기울기 탐색 기법과 나선형 가중 평균 기법으로 구한 홀의 색상값이고,  $F(x,y)$ 는 두 결과의 가중치 합을 취한 최종적으로 할당된 홀의 색상 값이다. 그림 11은 각  $\alpha$  가중치를 달리하여 공통 홀을 채운 결과 영상이다.  $\alpha$  가중치가 0.1

일 경우에 기울기 탐색 기법의 영향이 적어지므로 색이 번지는 현상이 나타나는 반면  $\alpha$  가중치가 0.6일 경우 기울기 탐색 기법의 결과와 나선형 가중 평균 기법의 결과가 적절히 조합되어 번짐 현상이 줄면서 객체의 연속성이 유지되는 것을 확인할 수 있다<sup>9,10</sup>.

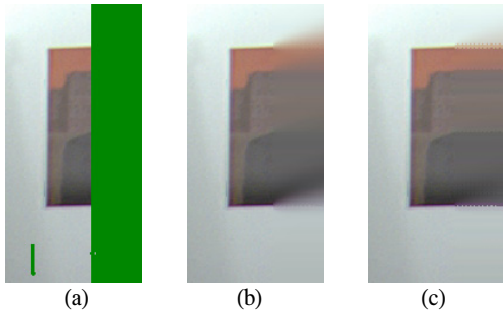


그림 11. 홀 채움 결과  
 (a) 홀 영상 (b)  $\alpha = 0.1$  (c)  $\alpha = 0.6$   
 Fig 11. Results of hole filling process  
 (a) Hole image (b)  $\alpha = 0.1$  (c)  $\alpha = 0.6$

#### IV. 실험 및 결과

본 논문에서는 MPEG에서 제공하는 다시점 영상 시퀀스인 Cafe, Book\_Arrival, Lovebird1, Mobile 시퀀스를 실험 영상으로 사용한다. 객관적인 화질평가를 위해서  $n$ 번 카메라의 영상을 사용하여  $(n+1)$ 번 카메라 영상을 생성하고 주어진 원본  $(n+1)$ 번째 영상과 PSNR을 비교하였다. 그리고 제안하는 기법의  $\alpha$  값을 실험을 통해 가장 좋은 결과가 나온 0.6을 사용하였다. 제안하는 기법의 성능을 기존의 선형보간법, Telea의 인페인팅 기법, VSRS 3.5 alpha에 구현된 2가지 방법의 성능과 비교하였다. 선형보간법은 홀의 수평방향의 양끝의 화소값을 이용하여 그 사이의 값들을 선형적으로 보간하는 방법이다. Telea의 인페인팅 기법은 영상의 기울기를 고려하여 평활화 예측을 통해 이를 채우는 방법이다. VSRS 1번 방법은 나선형 가중 평균 기법만을 사용한 기법이고 VSRS 2번 방법은 bilateral filter를 사용한 기법이다. 그림 12와 그림 13은 홀을 채운 후 생성된 가상시점 영상이다. 본 논문에서 제안하는 기법과 달리 결과 영상을 보면 기존 기법들은 경계 잡음을 제거하지 않고 홀 채움을 하므로 홀 주변에서 번지는 현상이 나타나는 것을 알 수 있다.

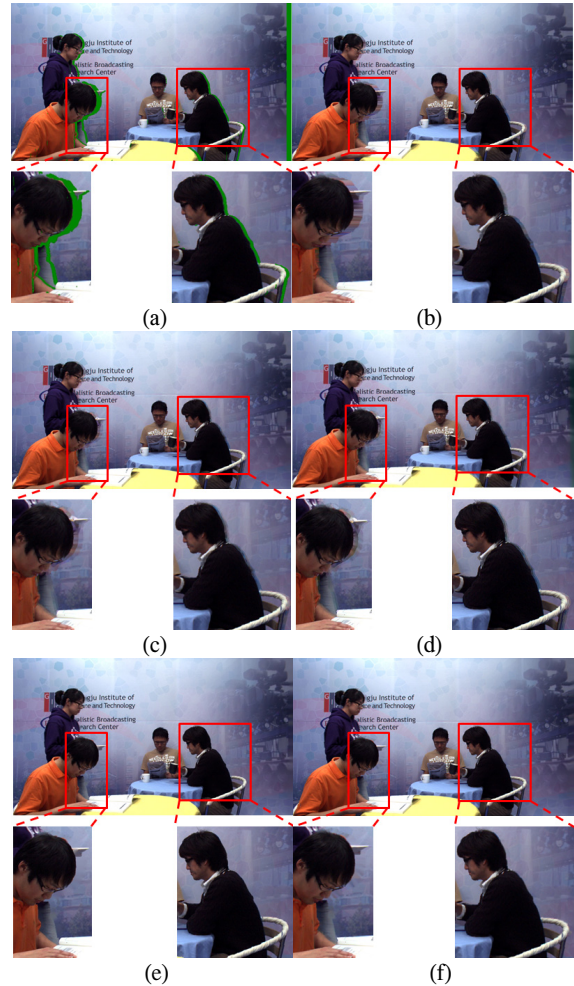


그림 12. 가상 시점 생성 결과  
 ("Cafe" 5번 시점, 0번째 프레임)  
 (a) 3D 워핑 결과 (b) 선형보간법 (c) 인페인팅 기법  
 (d) VSRS 1 (e) VSRS 2 (f) 제안한 기법  
 Fig 12. Generated virtual view  
 ("Cafe" 5th view, 0th frame)  
 (a) 3D warping result (b) Linear interpolation  
 (c) In-painting algorithm  
 (d) VSRS 1 (e) VSRS 2 (f) Proposed algorithm

표 1은 기존의 기법들과 제안한 기법을 이용하여 생성한 가상시점 영상을 주어진 원본 영상과 프레임별 PSNR을 구하여 평균을 구한 결과이다.



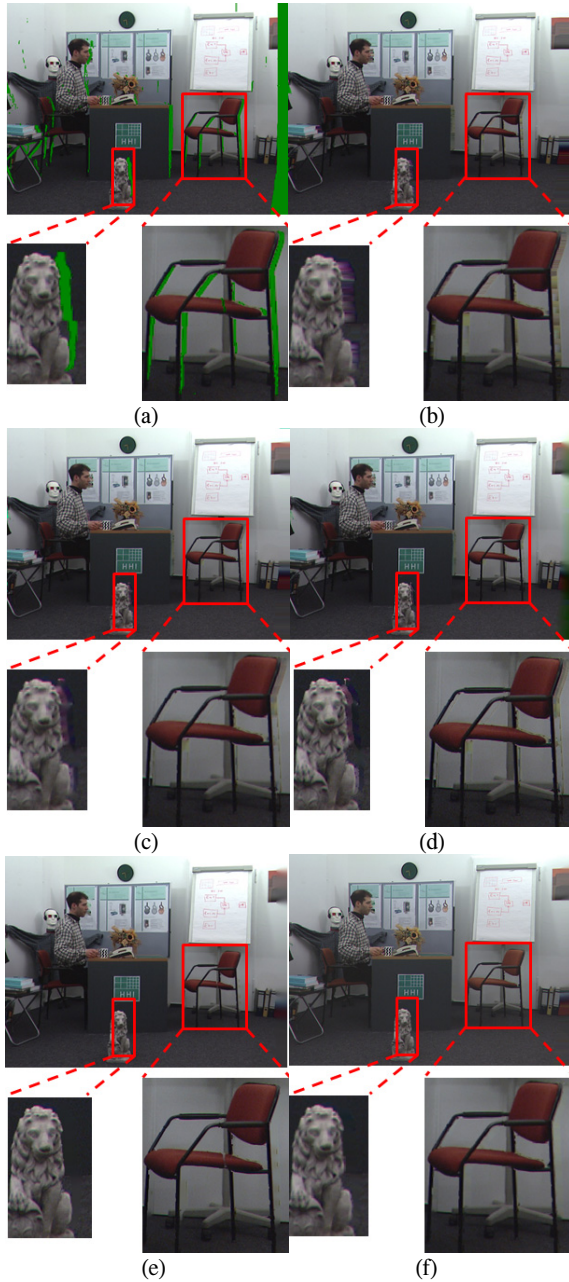


그림 13. 가상 시점 생성 결과  
 (“Book\_Arrival” 6번 시점, 0번 프레임)  
 (a) 3D 워핑 결과 (b) 선형보간법 (c) 인페인팅 기법  
 (d) VRSRS 1 (e) VRSRS 2 (f) 제안한 기법

Fig 13. Generated virtual view  
 (“Book\_Arrival” 6th view, 0th frame)  
 (a) 3D warping result (b) Linear interpolation algorithm  
 (c) In-painting algorithm (d) VRSRS 1 (e) VRSRS 2  
 (f) Proposed algorithm

표 1에서 보면 제안한 기법을 이용할 경우 기존의 기법을 이용하는 경우보다 대부분 PSNR이 향상 되는 것을 확인할 수 있다. 이러한 결과를 얻을 수 있는 이유는 제안하는 기법에서는 경계 잡음을 먼저 제거를 하고 홀 채움을 했고, 색상 정보 뿐 아니라 깊이 정보를

보까지 활용하여 홀 주변의 객체로 인한 영향을 효과적으로 제거 할 수 있었고, 또한 영상의 기울기 정보를 이용하여 영상의 연속성을 유지할 수 있었기 때문인 것으로 판단된다.

표 1. 제안한 기법의 성능 비교

Table 1. Performance comparison of the proposed algorithm

PSNR (dB)	선형 보간법	화소 단위 인페인팅 [5]	VRSRS1 [12]	VRSRS2 [13]	제안하는 기법
Cafe	34.187	34.374	34.273	35.400	35.279
Book Arrival	34.911	35.510	34.842	35.559	35.590
Lovebird1	30.077	30.070	30.068	30.159	30.125
Mobile	40.623	40.687	40.575	40.051	40.843
Average	34.950	35.160	34.940	35.292	35.459

## V. 결론

본 논문에서는 가상시점 영상을 생성하는 과정에서 발생하는 공통 홀을 효과적으로 채우기 위한 새로운 홀 채움 기법을 제안하였다. 제안한 기법에서는 먼저 기준 시점과 깊이지도 내의 객체 경계부분 불일치로 발생하는 경계 잡음을 검출하고 검출된 경계 잡음을 홀 영역으로 포함시킨다. 경계 잡음이 포함된 공통 홀 영역은 나선형 가중 평균(spiral weighted average) 기법과 기울기 탐색(gradient searching) 기법을 적절히 적용하여 채우게 된다. 나선형 가중 평균 기법은 깊이 정보를 사용함으로 객체 정보를 최소로 사용하지만 결과 영상이 번지는 단점이 있다. 기울기 탐색 기법은 영상의 기울기를 이용하여 세밀한 부분을 보존할 수 있는 장점이 있다. 두 기법의 장점을 최대한도로 반영하기 위하여 각 기법의 결과를  $\alpha$  가중치로 조합하여 홀 채움을 하였다.

실험을 통해 제안하는 기법을 통해 생성된 가상시점 영상이 기존의 다른 기법보다 우수한 품질을 보이는 것을 확인하였다. 그리고 원본 영상과의 PSNR 측정을 통해 평균 PSNR 또한 상승하는 것을 확인하였다. 이러한 결과를 얻을 수 있었던 이유는 제안하는 기법에서는 경계 잡음을 먼저 제거를 하고 홀 채움을 했고, 색상 정보 뿐 아니라 깊이 정보까지 활용하여 홀 주변의 객체로 인한 영향을 효과적으로 제거 하면서 동시에 영상의 기울기 정보를 이용하여 영상의 연

속성을 유지할 수 있었기 때문인 것으로 판단된다.

### References

[1] G.M. Um, G.H. Cheong, W.S. Cheong, N.H. Hur, "Technical Development and Standardization Trends of Multi-view 3D and Free-viewpoint Video", *The Magazine of the IEK*, vol.38, no.2, pp.18-23, Feb. 2011.

[2] G.M. Um, G. Bang, N.H. Hur, J.W. Kim, "Standardization Trends of 3D Video Coding in MPEG", *ETRI, Electronics and Telecommunications Trends*, vol.24, no.3, pp.61-68, Jun. 2009.

[3] ISO/IEC JTC1/SC29/WG11, "Introduction to 3D Video", M9784, May 2008.

[4] Y. Mori, N. Fukushima, T. Yendo, T. Fujii and M. Tanimoto, "View generation with 3D warping using depth information for FTV", *ELSEVIER, Signal Processing : Image Communication*, vol.24, issue.1-2, pp.65-72, Jan. 2009.

[5] A. Telea, "An image in-painting technique based on the fast marching method", *J. Graphics Tools*, vol.9, no.1, pp.25-36, 2004.

[6] A. Criminisi, P. Perez, K. Toyama, "Region Filling and Object Removal by Exemplar-Based Image In-painting", *IEEE Trans. Image Processing*, vol.13, no.9, pp.1200-1212, Sept. 2004.

[7] Se-hwan Park, Hyuk Song, Eun-young Jang, Namho Hur, Jinwoong Kim, Jin-soo Kim, Sang-hun Lee, Ji-sang Yoo, "Compensation Method for Occluded-region of Arbitrary-view Image Synthesized from Multi-view Video", *The Journal of Korea Information and Communications Society*, vol.33, no.12, pp.1029-1038, Dec. 2008.

[8] Tae-june Kim, Eun-young Chang, Nam-ho Hur, Jin-woong Kim, Ji-sang Yoo, "Virtual Viewpoint Image Synthesis Algorithm using Multi-view Geometry", *The Journal of Korea Information and Communications Society*, vol.34, no.12, pp.1154-1166, Dec. 2009.

[9] ISO/IEC JTC1/SC29/WG11, "Improved common-hole filling for extrapolation view synthesis", M21318, Jul. 2011.

[10] Min Soo Ko, Dong Wook Kim, Douglas L. Jones, Jisang Yoo, "A New Common-hole Filling Algorithm for Arbitrary View Synthesis", *4th International conference of 3D systems and applications*, pp.242-245, Jun. 2012.

[11] Min Soo Ko, Gun Bang, Won-Sik Cheong, Namho Hur, Jisang Yoo, "A New Post Processing Algorithm for Generated Arbitrary View", *3rd International conference of 3D systems and applications*, pp.183-187, Jun. 2011.

[12] ISO/IEC JTC1/SC29/WG11, "Boundary noise removal and hole filling for VSRS 3.5 alpha", M19992, Mar. 2011.

[13] ISO/IEC JTC1/SC29/WG11, "Implementation of Hole Filling Methods for VSRS 3.5.alpha", M20005, Mar. 2011.

#### 고 민 수 (Min Soo Ko)



2010년 2월 광운대학교 전자공학과 학사  
 2012년 2월 광운대학교 전자공학과 석사  
 2012년 3월~현재 광운대학교 전자공학과 박사과정  
 <관심분야> 3D 입체영상처리/압축, 스테레오 매칭, 영상 신호 처리

#### 유 지 상 (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> 3D 입체영상처리/압축, 웨이블릿 기반 영상처리, 비선형 디지털 신호처리