

일반논문 (Regular Paper)

방송공학회논문지 제25권 제6호, 2020년 11월 (JBE Vol. 25, No. 6, November 2020)

<https://doi.org/10.5909/JBE.2020.25.6.954>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

## 다중시점 환경에서의 슈퍼픽셀 세그멘테이션 기반 깊이 영상 개선 알고리즘

정 우 경<sup>a)</sup>, 한 중 기<sup>a)†</sup>

### Depth Map Correction Algorithm based on Segmentation in Multi-view Systems

Woo-Kyung Jung<sup>a)</sup> and Jong-Ki Han<sup>a)†</sup>

#### 요 약

실감형 미디어에서 현실감을 느끼게 하는 가장 중요한 요소는 깊이 정보이다. 따라서 고품질의 실감형 미디어를 제작하기 위해서는 고품질의 깊이 정보를 획득하는 것이 필수적이다. 본 논문에서는 고품질의 깊이 정보를 획득하기 위하여 다중 시점 환경에서 깊이 지도를 개선하기 위하여 깊이 지도를 여러 개의 세그먼트로 분할 및 다중 시점간의 관계를 고려하는 알고리즘을 제안한다. 제안된 알고리즘은 슈퍼픽셀 세그멘테이션 기법을 사용하여 기준 시점의 깊이 지도를 여러 세그먼트로 나누고, 각 세그먼트를 인접 시점으로 투영한다. 이후 투영된 세그먼트의 정보를 이용하여 인접 시점의 깊이 지도를 평면 추정을 이용하여 개선한 후, 기준 시점으로 역투영된다. 여러 개의 인접 시점에 대해 이 과정을 반복하여 개선된 인접 시점들의 값들과 기준 시점의 초기 깊이 지도를 가중치 합으로 갱신하여 깊이 지도를 개선한다. 기존 다중 시점 스테레오 비전 알고리즘에 제안된 알고리즘을 적용한 시뮬레이션을 통해 제안된 알고리즘의 결과가 주관적 및 객관적으로 기존 알고리즘을 능가하는 것을 보인다.

#### Abstract

In immersive media, the most important factor that provides immersion is depth information. Therefore, it is essential to obtain high quality depth information in order to produce high quality immersive media. In this paper we propose an algorithm to improve depth map, considering the segmentation of images and the relationship between multiple views in multi-view systems. The proposed algorithm uses a super-pixel segmentation technique to divide the depth map of the reference view into several segments, and project each segment into adjacent view. Subsequently, the depth map of the adjacent view is improved using plane estimation using the information of the projected segment, and then reversed to the reference view. This process is repeated for several adjacent views to improve the reference depth map by updating the values of the improved adjacent views and the initial depth map of the reference view. Through simulation, the proposed algorithm is shown to surpass the conventional algorithm subjectively and objectively.

Keyword : depth map, multi-view system, super-pixel based segmentation

## I. 서론

VR(Virtual Reality) 기기 제작 벤처회사였던 오кул러스가 페이스북에 인수된 2014년 이래 VR 및 AR 하드웨어는 꾸준한 발전을 지속하였다. 이를 통해 VR 및 AR(Augmented Reality) 하드웨어가 보급되기 시작했으며, 비교적 고가로 인한 장벽을 낮추기 위하여 최근 저가형 하드웨어도 출시되고 있다. 이를 통해 VR 및 AR이 대중에게 보급되기 시작하였고 실감형 미디어가 주요 콘텐츠로 주목받고 있다. 이러한 추세에 따라 3-DoF(Depth of Field), 6-DoF 등의 시청자에게 다양한 시야를 제공함에 관련된 기술들이 자유 시점 영상을 구현하기 위한 중요한 기술들로 취급되고 있다. 실감형 미디어는 이름 그대로 사용자가 실감할 수 있는 미디어로, 사용자가 미디어를 통해 현실감을 느껴야 한다. 따라서 고품질의 실감형 미디어는 최대한 현실감 있게 제작하는 것이 필수 불가결한 요소이다. 이러한 현실감은 주로 깊이 정보에 의존하므로, 고품질의 깊이 정보를 획득하는 것은 고품질의 실감형 미디어를 제작하기 위한 가장 중요한 요소라고 할 수 있다.

지난 수십 년간 많은 연구자들에 의해 깊이 지도를 효율적으로 생성하기 위한 알고리즘들이 발명되었다<sup>[1][2]</sup>. 이러한 알고리즘들은 2개의 시점에 대해서 시차를 계산하는 스테레오 비전으로 시작하여, 최근에는 보다 많은 시점에 대해서 깊이 지도를 추정하기 위한 다중 시점 스테레오 비전 알고리즘으로 발전하였다. 그러나 카메라 캘리브레이션 과정 및 윈도우 기반 매칭 과정에서 발생한 노이즈로 인하여 깊이 정보의 품질이 저하하는 문제가 있었다. 이를 개선하기 위하여 최초 깊이 지도 추정 이후 후처리하는 방식의 깊이 지도 개선 알고리즘이<sup>[2][3][4]</sup> 제안되었었다. 그러나 이

러한 알고리즘들은 다중 시점 스테레오 비전이 아닌 2개의 카메라를 이용하는 스테레오 비전 환경에서 고려된 알고리즘이었고, 이를 다중 시점 환경에서 이용할 경우 시점 간 깊이 정보의 불일치로 인한 품질 저하가 발생할 수 있는 문제가 있었다. 따라서 다중 시점 환경에서 깊이 지도를 개선하기 위해서는 기준이 되는 시점의 깊이 지도뿐만 아니라, 이웃한 시점의 깊이 지도를 이용하여 각각의 깊이 지도를 개선하고, 이를 정합하는 과정이 필요하다.

본 논문은 다중 시점 환경에서 인접 시점의 데이터들을 이용하여 평면모델을 구성하여 기준 시점의 깊이 정보를 개선하기 위한 알고리즘을 제안한다. 제안된 알고리즘은 다중 시점 사이의 관계를 고려하여 영상 분할 기법을 이용하여 깊이 지도를 개선한다. 먼저 슈퍼-픽셀 영상분할 기법을 이용하여 기준 시점의 깊이 지도를 여러 개의 세그먼트로 나누고, 각각의 세그먼트를 인접 시점들로 투영한다. 이후 인접 시점의 깊이 정보를 이용하여 평면 모델을 추정하여 인접 시점의 깊이 정보를 개선한다. 개선된 인접 시점의 깊이 정보들을 다시 기준 시점으로 투영되고, 이를 참조 시점의 초기 깊이 정보와의 가중치 합을 계산하여 이를 이용해 기준 시점의 깊이 지도를 갱신한다.

이 논문은 다음과 같이 구성되어 있다. 2절에서는 기존 알고리즘의 제약사항들에 관해서 설명하며, 다중 시점 환경에서 깊이 지도를 향상시키기 위해 해결할 문제를 공식화한다. 3절에서는 깊이 지도에 있는 폐색 및 홀 영역과 관련된 문제를 개선하기 위한 알고리즘을 제안한다. 4절에서는 제안된 알고리즘의 성능을 입증하기 위해 시뮬레이션 결과를 제시한다. 5절에서는 이 논문을 간략히 결론 짓는다.

## II. 기존 연구 결과 분석

깊이 지도는 스테레오 비전 알고리즘인 패치-매칭을 이용해 추정하는 것이 일반적이다. 이러한 매칭 방법은 (1) 매칭 비용 계산, (2) 매칭 비용 집계 (3) 디스패리티 계산 및 최적화 (4) 후처리 등 4단계의 과정<sup>[5]</sup>으로 구성된다. 과정 (1)의 경우, 일반적으로 윈도우 기반으로 SAD, 혹은 NCC와 같은 단위를 이용해 매칭 정확도, 즉 매칭 비용을 계산한다. 과정 (2)에서는 이전 과정에서 계산한 매칭 비용을 모든

a) 세종대학교 전자정보통신공학과(Sejong University, Dept. of Electrical Engineering)

✉ Corresponding Author : 한종기(Jong-Ki Han)

E-mail: hjk@sejong.edu

Tel: +82-2-3408-3739

ORCID: <https://orcid.org/0000-0002-5036-7199>

※ This work was partly supported by the National Research Foundation of Korea (NRF) under Grant NRF-2018R1A2A2A05023117 and partly by Institute for Information & communications Technology Promotion (IITP) under Grant 2017-0-00486 funded by the Korea government (MSIT).

• Manuscript received September 1, 2020; Revised October 30, 2020; Accepted October 30, 2020.

픽셀에 대해서 집계한다. 과정 (3)에서는 집계한 매칭 비용을 이용해 각 픽셀의 최적 시차 값을 계산하고, 이를 이용해 깊이 지도를 추정한다. 과정 (4)에서는 추정한 깊이 지도를 개선하기 위하여 후처리 알고리즘을 이용한다.

다중 시점 환경에서의 깊이 지도는 2개 시점 환경에서 고려된 스테레오 비전 알고리즘을 다중 시점으로 확장한 다중 시점 스테레오 비전 알고리즘을 이용해 추정할 수 있다. 시점의 개수가 2개 이상으로 증가하였으므로, 매칭 비용 계산은 한 개의 시점에 대해서 계산하는 것이 아닌 모든 시점에 대해서 실행될 필요가 있다. 따라서 연산량을 줄이기 위하여 카메라의 내부 및 외부 파라미터를 이용하여 각 카메라 간의 관계를 파악한 뒤, 인접한 시점에 대해서만 계산하여 연산량을 줄이는 방법이 주로 이용된다.

일반적인 스테레오 비전의 깊이 지도 개선을 위한 후처리 알고리즘은 시점 차이로 인해 발생하는 폐색 영역 및 홀 영역을 채우는 것을 목표로 하고 있다. 폐색 영역의 경우, 다른 시점에서 정보가 존재하지 않으므로, 폐색 영역 주변 픽셀의 깊이 값을 이용하여 해당 영역을 채우는 방식이 일반적이다.

다중 시점 환경은 시점의 수가 2개 이상이므로, 다중 시점 스테레오 비전의 깊이 지도는 이중 시점 스테레오 비전의 깊이 지도와 다른 특성을 가진다. 한 시점에 대해서 매칭을 시도하며 발생한 폐색 영역은 다른 시점에서는 발생하지 않을 수 있다. 따라서 다중 시점 스테레오 비전을 위한 깊이 지도 후처리 알고리즘은 이중 시점 스테레오 비전의 것과 달리 다중 시점간의 깊이 정보 간의 관계를 고려할 필요가 있다. 만약 스테레오 비전 환경에서 사용하는 방법을 다중 시점 환경에 대한 고려 없이 적용하면 다중 시점의 깊이 영상에 불일치 문제가 발생할 가능성이 존재한다.

다중 시점 환경에서 깊이 지도를 추정하기 위한 기존 연구<sup>[6][7]</sup>에서는 깊이 정보를 개선하기 위하여 시점간의 깊이의 일관성을 고려한다. 이러한 연구에서는 인접 시점의 깊이 지도가 기준 시점의 깊이 영상으로 투영된다. 그 이후 기준 시점의 깊이 영상과의 깊이 값 차이를 평가한다. 기준 시점의 깊이 영상의 값은 인접 시점에서 해당하는 픽셀의 깊이 값을 투영하고, 일반적으로 미디안 필터 등을 이용하여 갱신된다.

기존 알고리즘<sup>[6][7]</sup>들을 다중 시점 환경의 다양한 테스트

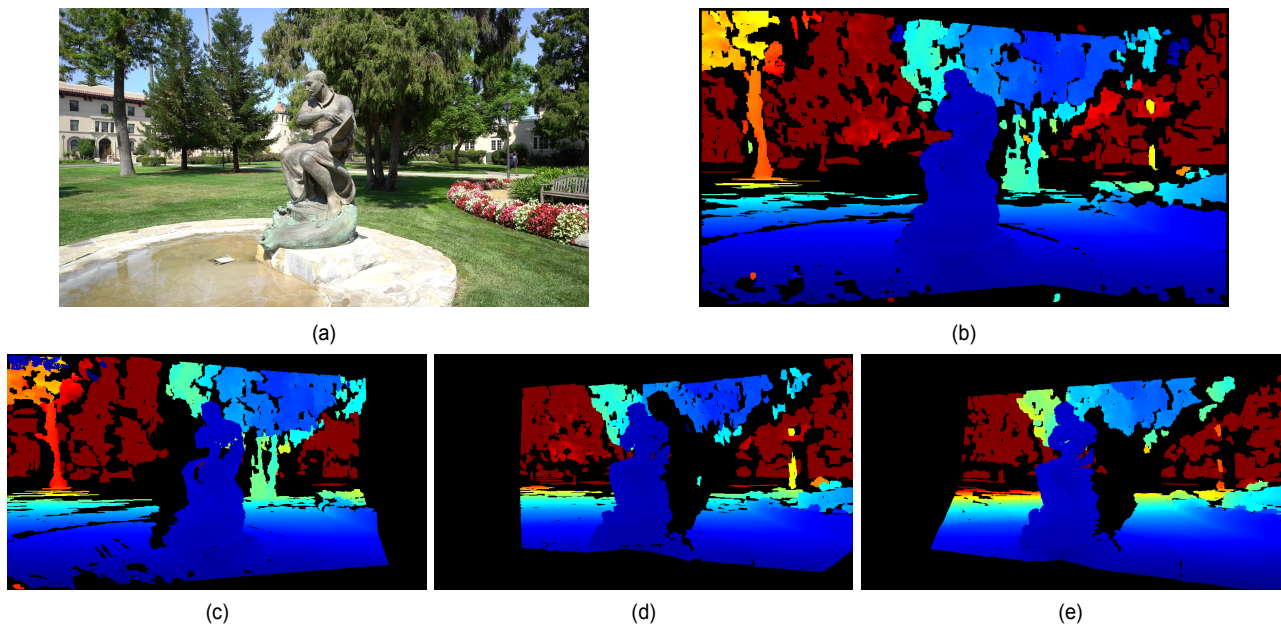


그림 1. [11] 내 Ignatius 데이터셋을 이용한 시점 간 투영으로 인한 폐색 및 홀의 발생 예. (a) 기준 시점의 텍스처 이미지, (b) 기준 시점의 깊이 맵, (c)~(e) 기준 시점 위치에 투영된 인접 깊이 지도들

Fig. 1. Examples of occlusion and holes resulted from projection between views using Ignatius dataset from [11] (a) texture image of reference view, (b) depth map of the reference view, (c)~(e) neighbor depth maps projected onto reference view position

영상에 적용한 결과, 투영된 깊이 영상에서 여러 개의 폐색 및 홀 영역을 발견할 수 있었다. 이러한 영역은 특정 시야 방향을 따라 개선을 시도하는 깊이 영상의 품질 저하를 초래할 수 있다. 그림 1은 투영된 깊이 영상에서 발생하는 폐색 및 홀 영역의 예시를 보인다. 그림 1 (a) 및 (b)는 기준 시점의 텍스처 영상 및 추정된 깊이 영상을 보인다. 그림 1 (c)~(e)는 인접 시점의 깊이 영상들을 기준 시점으로 투영한 결과들을 보인다. 그림 1 (b)~(e)에서 각 색상은 깊이 값을 의미하며, 청색은 가까운 깊이 값을, 적색은 멀리 있는 깊이 값을 나타낸다. 검은색 영역은 깊이 값이 존재하지 않는 영역이다. 그림 1 (c)에서 보인 것처럼, 투영된 인접 시점의 깊이 지도는 홀 영역이 관측되고, 폐색으로 인하여 일부 깊이 값이 사라지는 것을 확인할 수 있다. 이러한 영역들로 인하여 기준시점의 깊이 영상 개선이 효율적으로 이뤄지지 않게 된다.

따라서 본 논문에서는 기준 시점의 텍스처 영상을 이용하여 영상을 여러 개의 세그먼트로 분할한 후, 세그먼트 맵을 깊이 영상에 적용한다. 이후 이를 인접 시점으로 투영하여 인접 시점의 깊이 영상을 개선한 후, 개선된 인접 시점의 깊이 영상을 다시 역투영 하여 기준 시점의 깊이 영상을 개선하는 알고리즘을 제안한다.

### III. 알고리즘 제안

3절에선 4단계로 구성된 기준 시점의 깊이 지도 개선을 위한 알고리즘을 제안한다. ; (a) 텍스처 영상을 이용한 깊이 영상 분할, (b) 인접 시점의 깊이 영상에 대한 기준 시점

의 깊이 영상 투영 및 평면 추정, (c) 기준 시점의 깊이 지도에 수정된 인접 시점의 깊이 지도 투영, (d) 기준 시점의 깊이 지도의 개선. 제안된 개선 알고리즘의 전체 블록 다이어그램은 그림 2에 나타나 있으며, 다음과 같은 가정을 한다. 가정 1은 “다중 시점 환경에 채택된 모든 카메라의 내부, 외부 파라미터가 알려져 있다”는 것이고, 가정 2는 “각 픽셀의 매칭 정확도를 이용한 신뢰도가 계산되어 있다”는 것이다. 가정 1의 경우, 다중 시점 환경에서 계산되어 있어야 최초 깊이 지도를 추정할 수 있다<sup>[8]</sup>. 따라서 가정 1은 자동으로 충족된다. 가정 2의 경우, 패치-매칭 기법을 이용하는 등 각 픽셀의 매칭 정확도를 통해 신뢰도를 계산할 수 있으므로 충족된다<sup>[9]</sup>.

#### 1. 텍스처 영상을 이용한 깊이 영상 분할

제안된 알고리즘의 첫 번째 단계에서는 기준 시점의 깊이 영상에 슈퍼-픽셀 세그먼테이션<sup>[10]</sup>을 시행하여 분할한다. 그림 3은 슈퍼-픽셀 세그먼테이션을 통해 분할된 텍스처 이미지와 깊이 영상의 예를 보여준다. 그림 3에서 보인 바와 같이, 세그먼트 내 픽셀들은 유사한 성질과 값을 가진다. 제안된 알고리즘의 모든 과정은 그림 3 (c)와 같이 그림의 각 패치에 적용된다.

#### 2. 인접 시점의 깊이 영상에 대한 기준 시점의 깊이 영상 투영 및 평면 추정

이 단계에서는 기준 시점의 깊이 영상에 존재하는 각 세그먼트를 인접 시점들로 투영한다. 이를 위해 다중 시점 환

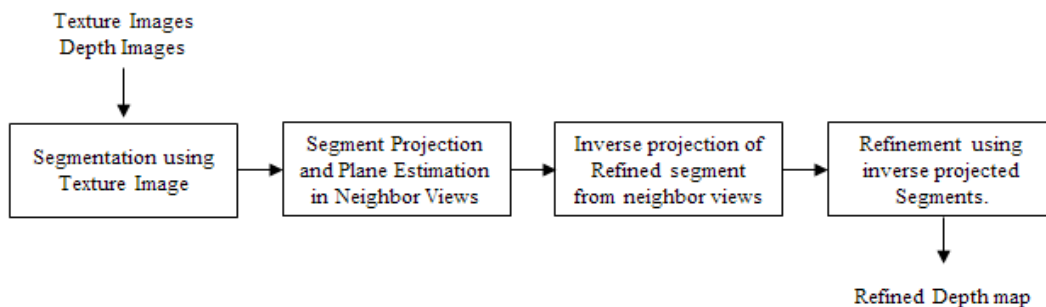


그림 2. 제안된 알고리즘의 전체 블록 다이어그램  
Fig. 2. Overall block diagram of the proposed refinement algorithm



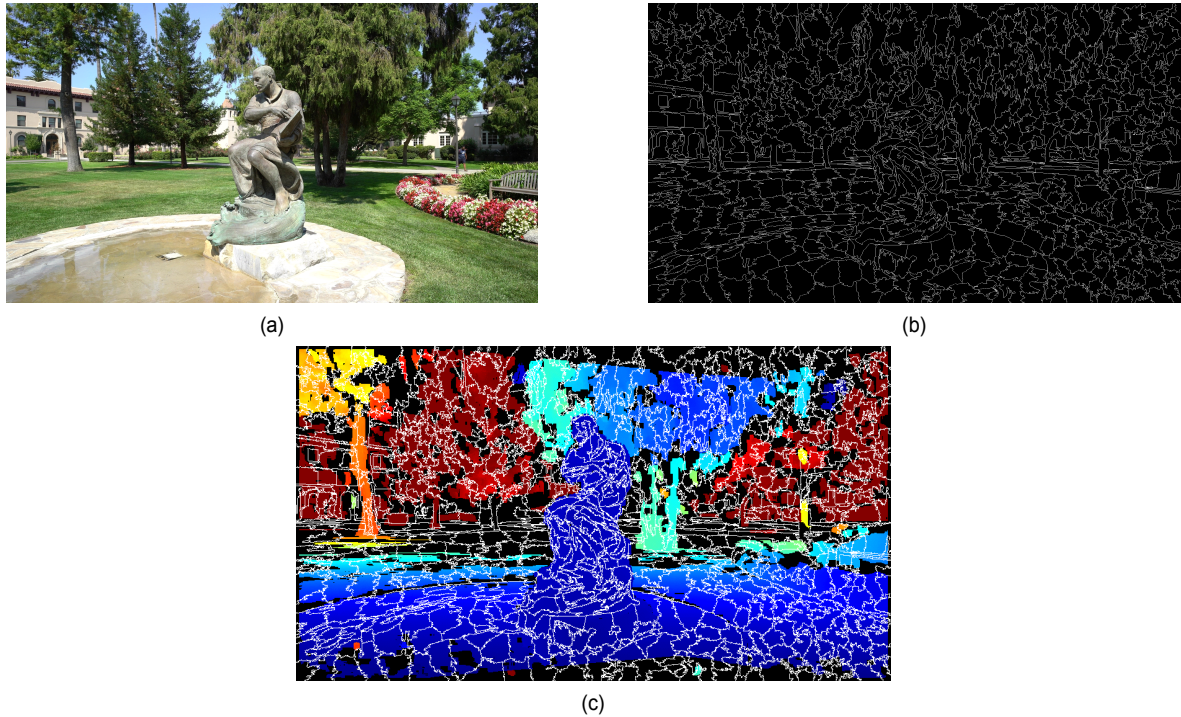


그림 3. [11] 내 Ignatius 데이터셋을 이용한 슈퍼-픽셀 세그멘테이션 예. (a) 기준 시점의 텍스처 영상, (b) 슈퍼픽셀 분할을 위한 마스크, (c) 슈퍼-픽셀 분할에 따라 분할된 깊이 지도

Fig. 3. Example for segmentation process using Ignatius dataset from [11] (a) texture image of reference view, (b) mask for super-pixel segmentation, (c) depth map divided according to super-pixel segmentation

경에 이용된 카메라들의 내부 및 외부 파라미터를 알 필요가 있다. 이는 가정 1에 의해 이미 알고 있는 상태이므로, 다음과 같은 공식을 이용하여 세그먼트를 기준 시점 이미지 좌표계, 기준 시점 카메라 좌표계, 월드 좌표계, 인접 시점 카메라 좌표계, 인접 시점 이미지 좌표계로 변환할 수 있다.

$$\begin{bmatrix} X_{Ref} \\ Y_{Ref} \\ Z_{Ref} \end{bmatrix} = \begin{bmatrix} \frac{x_{Ref}d(P_{Ref})}{f} \\ \frac{y_{Ref}d(P_{Ref})}{f} \\ \frac{f}{d(P_{Ref})} \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} x_n \\ y_n \\ z_n \end{bmatrix} = K_n \begin{bmatrix} X_n \\ Y_n \\ Z_n \\ 1 \end{bmatrix} = \begin{bmatrix} f_n & 0 & 0 & 0 \\ 0 & f_n & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_n \\ Y_n \\ Z_n \\ 1 \end{bmatrix} \quad (4)$$

여기서  $P_{Ref}$ 는 기준 시점의 세그먼트 내 픽셀을 의미한다.  $[x_{Ref}, y_{Ref}]^T$ 는 픽셀  $P_{Ref}$ 의 좌표를 의미한다.  $d(P_{Ref})$ 는 픽셀  $P_{Ref}$ 의 깊이 값을 의미한다.  $K_{Ref}, R_{Ref}, t_{Ref}$ 는 각각 기준 시점의 카메라의 내부, 외부, 이동 행렬을 의미한다.  $[X_{Ref}, Y_{Ref}, Z_{Ref}]^T$ 는 카메라 좌표계에서의  $P_{Ref}$ 의 좌표를

$$\begin{bmatrix} U_{world} \\ V_{world} \\ W_{world} \\ 1 \end{bmatrix} = \begin{bmatrix} X_{Ref} \\ Y_{Ref} \\ Z_{Ref} \\ 1 \end{bmatrix} R_{Ref}^T + t_{Ref} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 0 & 0 & c_x \\ 0 & 1 & 0 & c_y \\ 0 & 0 & 1 & c_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_{Ref} \\ Y_{Ref} \\ Z_{Ref} \\ 1 \end{bmatrix} \quad (2)$$

$$\begin{bmatrix} X_n \\ Y_n \\ Z_n \\ 1 \end{bmatrix} = \begin{bmatrix} U_{world} \\ V_{world} \\ W_{world} \\ 1 \end{bmatrix} R_n - t_n = \begin{bmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -c_x \\ 0 & 1 & 0 & -c_y \\ 0 & 0 & 1 & -c_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} U_{world} \\ V_{world} \\ W_{world} \\ 1 \end{bmatrix} \quad (3)$$

의미한다. (1)과 (2)를 적용함으로써 각각 기준 시점 카메라 좌표계에서의 좌표와 월드 좌표계에서의 좌표를 계산할 수 있다.  $[X_n, Y_n, Z_n]^T$  는 n번째 인접 시점의 카메라 좌표계에 존재하는  $P_n$ 의 좌표를 의미한다.  $K_n, R_n, t_n$ 는 각각 n번째 인접 시점의 카메라의 내부, 외부, translation 행렬을 의미한다. (3)과 (4)를 적용함으로써 각각 n번째 인접 시점 카메라 좌표계에서의 좌표와 인접 시점 이미지 좌표계에서의 좌표를 계산할 수 있다. 이때, n번째 인접 시점 이미지에서

의 깊이 값  $d(P_n)$ 는  $Z_n$ 과 동일하다.

그림 4에서 픽셀  $P_{Ref}$ 가 인접 시점의 패치 내  $P_n$ 로 투영된 것을 확인 할 수 있다. 제안된 알고리즘에서는 파란색 패치 영역의 픽셀의 위치 정보가 인접 시점으로 투영된 후, 투영된 위치 픽셀의 깊이 값에 따라 투영된 패치의 영역을 비-폐색 영역과 폐색 영역으로 나눈다. 만약 픽셀에 의미 있는 깊이 값이 있으면 비-폐색 영역에 할당된다. 만약 영역이 앞쪽 물체에 의해 가려져 깊이 정보를 가지고 있지 않거

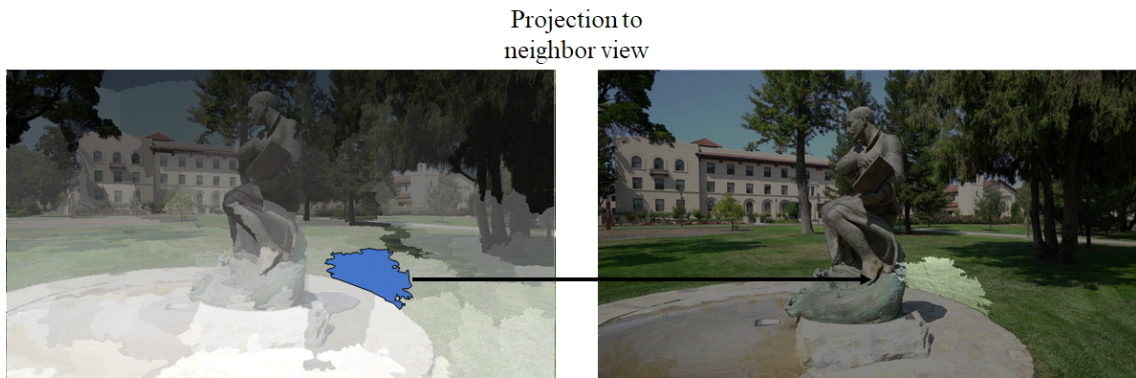


그림 4. [11] 내 Ignatius 데이터셋을 이용한 슈퍼-픽셀 세그먼트와 투영된 세그먼트의 예  
Fig. 4. Examples of a super pixel segment and the projected segment using Ignatius dataset from [11]



그림 5. [11] 내 Ignatius 데이터셋을 이용하여 투영된 세그먼트의 폐색 및 비-폐색 영역 예  
Fig. 5. Example of a occluded and non-occluded area of a projected segment using the Ignatius dataset in [11]

나 인접 시점의 깊이 값보다 커서 가려져 있다고 판단될 경우 해당 영역을 즉 폐색 영역으로 설정한다.

그림 5는 폐색 영역과 비-폐색 영역의 예시를 보이고 있다. 적색 패치 영역은 깊이 정보가 존재하지 않는 폐색 영역이다. 따라서 녹색 패치, 즉 비-폐색 영역의 깊이 값을 이용하여 평면 모델을 구성하고, 구성한 평면 모델을 적색 패치 영역에 적용하여 깊이 값을 조정한다. 평면 모델을 추정할 때, 평면 모델이 평면 함수를 구성하기 위해 선택한 깊이 값에 따라 달라진다. 따라서 RANSAC<sup>[12]</sup> 알고리즘을 사용하여 평면 모델을 최적화한다. 깊이 값에 대하여 RANSAC을 적용할 때, 초기 깊이 영상을 생성하는 과정에서 계산된 각 픽셀의 깊이 값의 신뢰도를 고려하여 가중치를 계산한다. 따라서 세그먼트 내 픽셀의 초기 깊이 값에 오류가 존재하더라도 신뢰도가 높은 픽셀을 이용한 평면이 추정되므로, 이를 통해 전체 오차를 줄인다.

### 3. 기준 시점의 깊이 지도에 수정된 인접 시점의 깊이 지도 투영

이전 과정에서 제안된 알고리즘을 적용하여 모든 인접 시점의 깊이 영상들을 개선한 후, 모든 인접 시점의 녹색

및 빨간색 영역의 모든 픽셀 및 깊이 값을 기준 시점으로 투영한다. 이 과정을 통해 다음 단계의 깊이 값 개선을 위한 입력 값을 제공한다.

### 4. 기준 시점의 깊이 지도의 개선

기준 시점의 깊이 영상에 있는 각 픽셀의 깊이 값은 초기 깊이 값과 다중 인접 시점에서 투영된 깊이 값의 가중치 합으로 대체된다. 이 갱신 과정은 평균, 미디안 필터링, 가우시안 필터링, 바이래터럴 필터링과 같은 다양한 필터링 함수들 중 하나로 구현될 수 있다. 이 논문에서는 인접 시점에서 투영된 깊이 값의 신뢰도가 가중치 요인으로 사용하여 가중 평균화 함수를 선택하였다.

$$new P_{ref} = \frac{d(P_{ref}) + \sum_{n=1}^N confidence(P_n) * d(P_{n \rightarrow ref})}{confidence(P_{ref}) + \sum_{n=1}^N confidence(P_n)} \quad (5)$$

식(5)는 본 논문에서 이용한 픽셀의 깊이 값 공식이다.  $P_{ref}$ 는 기준 시점 내 픽셀을 의미한다.  $d(P_{ref})$ 는  $P_{ref}$ 의 깊이 값을 의미한다.  $P_n$ 는 n번째 이웃 시점으로 투영된  $P_{ref}$

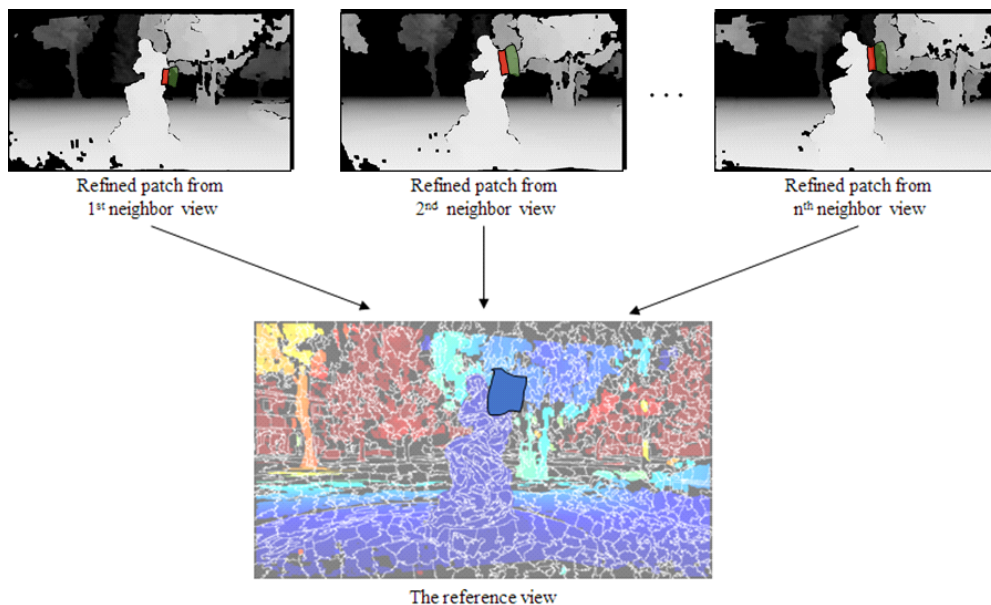


그림 6. 제안된 알고리즘의 3번째 과정 예  
Fig. 6. Example of the third phase of the proposed algorithm



를 의미한다.  $confidence(P_n)$ 는  $n$ 번째 이웃 시점으로 투영된  $P_{ref}$ 의 신뢰도를 의미한다.  $P_{n \rightarrow ref}$ 는  $P_n$ 을 기준 시점으로 역투영한 픽셀을 의미한다.

#### IV. 실험 결과

이 절에서 제안된 알고리즘은 [13]에서 설명하는 기준인 F1 점수를 기반으로 평가된다. F1 점수는 복원된 포인트 클라운드를 평가하기 위하여 완전성과 정확성, 이 두 가지의 평가 요소를 통해 계산할 수 있다. 각각의 평가 요소들은 ground truth 포인트 클라운드와 복원된 포인트 클라운드를 비교하여 결과를 얻을 수 있다. 완전성의 경우 ground truth 포인트 클라우드의 3D 포인트와 매칭되는 가장 가까운 복원된 포인트와의 거리를 이용하여, 정확성은 ground truth 포인트로부터 일정 임계 값 내에 있는 복원된 포인트의 비율을 이용하여 계산된다. 이후 완전성과 정확성의 조화 평

균값을 계산하여 이를 해당 복원된 포인트 클라우드의 F1 점수로 한다.

제안된 알고리즘의 성능은 openMVS<sup>[14]</sup>와 비교되며, 모든 시뮬레이션은 [13]에 제공된 테스트 데이터셋으로 수행되었다. 해당 데이터셋은 고해상도-다 시점, 저해상도-다 시점, 저해상도-이중 시점의 3가지 데이터셋이 존재하며 각각의 데이터셋 내에는 training 데이터셋과 test 데이터셋이 존재한다. 고해상도-다 시점 training 데이터셋은 총 13개의 영상 세트로 구성되어 있으며, 이 중 6개 영상 세트가 outdoor, 7개 영상 세트가 indoor 영상 세트로 분류된다. 본 실험에서는 고 해상도-다 시점 training 데이터셋 내 13개의 영상세트를 전부 이용하여 F1 score를 측정하였다.

그림 7을 통해 openMVS에 대해 제안된 알고리즘을 적용하여 계산된 gain을 보여주며, F1 점수가 알고리즘의 성능을 평가하는 기준으로 사용된다. 이 그래프에서 관찰한 바와 같이, tolerance가 증가함에 따라 이득이 감소하지만, 제안된 알고리즘은 tolerance의 전체 범위에 대해 이득이

표 1. Tolerance별 openMVS와 제안된 알고리즘의 평균 F1 점수 결과  
Table 1. Average of F1 score of openMVS and proposed method by tolerance

Tolerance	1	2	5	10	20	100	200	500
openMVS	0.147	0.326	0.962	2.180	5.060	13.218	25.303	42.917
openMVS+ Proposed Method	0.171	0.426	1.044	2.389	5.391	13.655	25.719	42.963

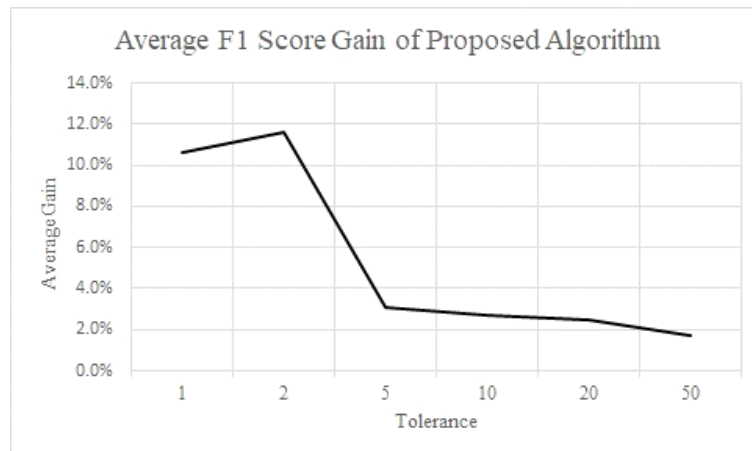


그림 7. 제안한 알고리즘의 평균 F1 점수 이득  
Fig. 7. Average F1 score gain of the proposed algorithm

존재한다. 이 결과로부터 제안된 알고리즘이 openMVS를 통해 계산된 깊이 영상보다 제안된 알고리즘을 통해 개선

한 깊이 영상이 더 정확하다는 것을 객관적으로 알 수 있다. 그림 8과 9는 ETH3D의 테스트 데이터셋을 openMVS와



그림 8. [13]의 facade 데이터셋을 이용한 포인트 클라우드 복원 결과 비교  
Fig. 8. Comparing point cloud reconstruction results using facade dataset from [13]

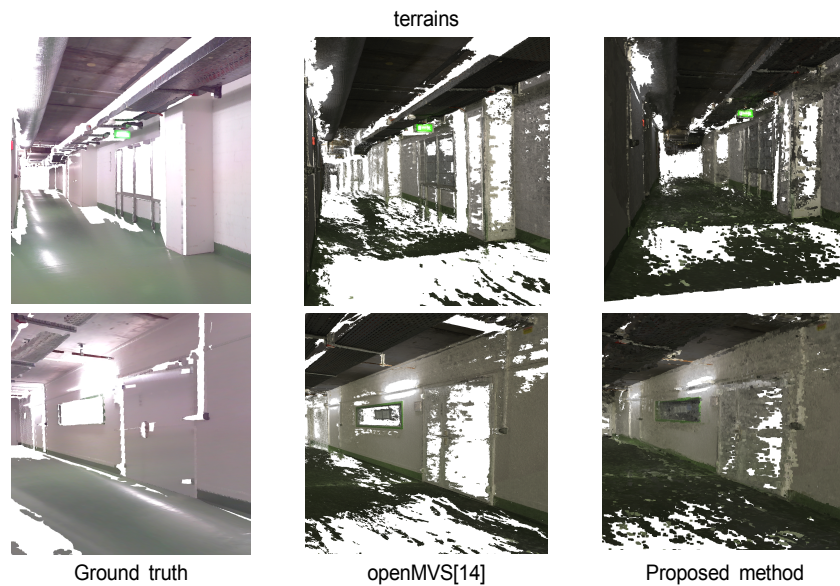


그림 9. [13]의 terrains 데이터셋을 이용한 포인트 클라우드 복원 결과 비교  
Fig. 9. Comparing point cloud reconstruction results using terrains dataset from [13]

제안된 알고리즘을 이용해 복원한 포인트 클라우드를 보인다. 이 그림에서 관찰된 바와 같이 제안된 알고리즘을 이용하여 복원한 포인트 클라우드는 OpenMVS를 이용하여 복원한 포인트 클라우드보다 ground truth와 더 유사한 것을 확인할 수 있다. 따라서 그림 7을 통해 제안된 알고리즘이 OpenMVS를 능가하는 것을 주관적으로 확인할 수 있다.

## V. 결 론

본 논문에서는 영상 분할 기법과 RANSAC을 이용한 평면 모델 추정을 활용하는 다중 시점 환경에서의 깊이 지도를 향상시키기 위한 알고리즘을 제안하였다. 기존 알고리즘은 인접 시점의 깊이 값을 향상시키지 않고 그대로 채택하여 기준 시점의 깊이 값을 갱신하는 반면, 제안된 알고리즘은 인접 시점의 향상된 깊이 값을 이용한다. 인접 시점의 깊이 영상을 향상시킴으로써 알고리즘이 기준 시점의 깊이 지도의 폐색 및 홀 영역과 관련된 문제를 극복할 수 있다.

## 참 고 문 헌 (References)

- [1] H. Hirschmuller, "Stereo Processing by Semiglobal Matching and Mutual Information," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 328-341, Feb. 2008.
- [2] A. Dziembowski, A. Grzelka, D. Mieloch, O. Stankiewicz and M. Domański, "Depth map upsampling and refinement for FTV systems", 2016 International Conference on Signals and Electronic Systems (ICSES), Krakow, pp. 89-92, 2016.
- [3] T. Chang, J. Kuo and J. Yang, "Efficient hole filling and depth enhancement based on texture image and depth map consistency," 2016 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), Jeju, pp. 192-195, 2016.
- [4] J. Lu, D. Min, R. S. Pahwa and M. N. Do, "A revisit to MRF-based depth map super-resolution and enhancement," 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, pp. 985-988, 2011.
- [5] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms" *Int. Jour. Computer Vision*, 47(1/2/3):7 - 42, 2002
- [6] T. Schöps, J. L. Schönberger, S. Galliani, T. Sattler, K. Schindler, M. Pollefeys, A. Geiger, "A Multi-View Stereo Benchmark with High-Resolution Images and Multi-Camera Videos", *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [7] H. Shih and H. Hsiao, "A depth refinement algorithm for multi-view video synthesis," 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, Dallas, TX, pp. 742-745, 2010
- [8] M. Kurc, O. Stankiewicz and M. Domański, "Depth map inter-view consistency refinement for multiview video," 2012 Picture Coding Symposium, Krakow, pp. 137-140, 2012.
- [9] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein and R. Szeliski, "A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms," 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 2006, pp. 519-528, doi: 10.1109/CVPR.2006.19.
- [10] R. Xiaofeng and J. Malik. "Learning a classification model for segmentation." *Proceedings Ninth IEEE International Conference on Computer Vision*, pp. 10-17 vol.1, 2003.
- [11] Arno Knapitsch, , Jaesik Park, Qian-Yi Zhou, and Vladlen Koltun. "Tanks and Temples: Benchmarking Large-Scale Scene Reconstruction". *ACM Transactions on Graphics* 36, no.4 (2017).
- [12] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography". *Commun. ACM* 24, pp. 381 - 395, 1981.
- [13] T. Schöps, J. L. Schönberger, S. Galliani, T. Sattler, K. Schindler, M. Pollefeys, A. Geiger, "A Multi-View Stereo Benchmark with High-Resolution Images and Multi-Camera Videos", *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [14] <https://github.com/cdcseacave/openMVS>



---

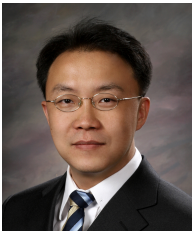
저 자 소 개

---



정 우 경

- 2015년 3월 ~ 현재 : 세종대학교 전자정보통신공학과 재학중
- ORCID : <https://orcid.org/0000-0003-2704-4436>
- 주관심분야 : 비디오 코덱, 영상 신호처리, VR



한 종 기

- 1992년 : KAIST 전기및전자공학과 공학사
- 1994년 : KAIST 전기및전자공학과 공학석사
- 1999년 : KAIST 전기및전자공학과 공학박사
- 1999년 3월 ~ 2001년 8월 : 삼성전자 DM연구소 책임연구원
- 2001년 9월 ~ 현재 : 세종대학교 전자정보통신공학과 교수
- 2008년 9월 ~ 2009년 8월 : University California San Diego (UCSD) Visiting Scholar
- ORCID : <https://orcid.org/0000-0002-5036-7199>
- 주관심분야 : 비디오 코덱, 영상 신호처리, 정보 압축, 방송 시스템