

다시점 360도 영상을 사용한 자유시점 영상 생성 방법

조영광*, 안희준*

*서울과학기술대학교 전기정보공학과
choyg@seoultech.ac.kr, heejune@seoultech.ac.kr

Free view video synthesis using multi-view 360-degree videos

Young-Gwang Cho*, Heejune Ahn*

*Dept. of Electrical and Information Engineering,
Seoul National University of Science and Technology

요 약

360 영상은 시청자가 시야방향을 결정하는 3DoF(3 Degree of Freedom)를 지원한다. 본 연구에서는 다수의 360 영상에서 깊이 정보를 획득하고, 이를 DIBR (Depth-based Image Rendering) 기법을 사용하여 임의 시점 시청기능을 제공하는 6DoF(6 Degree of Freedom) 영상제작 기법을 제안한다. 이를 위하여 기존의 평면 다시점 영상기법을 확장하여 360 ERP 투영 영상으로부터 카메라의 파라미터 예측을 하는 방법과 깊이영상 추출 방법을 설계 및 구현하고 그 성능을 조사하였으며, OpenGL 그래픽스기반의 RVS(Reference View Synthesizer) 라이브러리를 사용하여 DIBR을 적용하였다.

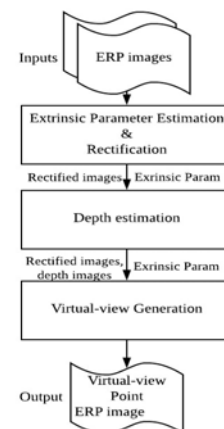
1. 서론

최근 HMD(Head Mounted Display)와 같은 장치와 360도 카메라가 보급되면서 가상현실을 위한 콘텐츠 개발이 활발히 진행되고 있다[1]. 그러나 렌더링을 통해 시점변화가 용이한 애니메이션기반 방식과 달리 실사영상을 바탕으로 한 가상시점 6DoF(6 Degree of Freedom) 제공은 어려움이 있다.

기존의 연구와 표준화 과정을 보았을 때 임의시점 영상을 생성하기 위해서는 깊이정보를 사용하는 것이 필수적일 것으로 보인다. 최근 표준화에서도 3DoF+(3 Degree of Freedom +), 6DoF 영상표준화에서 깊이정보를 사용하는 방향을 연구 중이다[1]. 본 논문은 기존의 다시점 평면영상의 다시점 영상 생성기법을 사용하였지만, 그 실현 과정에서 평면투영과 구형투영 영상 사이의 차이점으로 인하여 세부적인 알고리즘을 확장하였다. 두개이상의 ERP(Equi-Rectangular Projection)의 360도 영상을 입력으로 하여 카메라간의 외부 파라미터를 추정하여 위치관계를 파악한다. 이를 바탕으로 카메라를 정합(Rectification)하고 영상의 깊이를 추정한다. 추정된 깊이를 활용하여 각 픽셀들의 3차원 위치를 계산하여 가상시점에서의 영상을 생성한다.

2. 제안하는 ERP 가상시점 영상 생성 시스템

가상시점 생성을 위한 순서는 크게 3단계로 진행된다. 첫 번째로 사용자로부터 ERP 영상을 입력받아 각 영상의 카메라 간의 외부 파라미터를 추정하고 정합한다. 두 번째로 정합된 영상들을 이용하여 SAD(Sum of Absolute Difference) 방법을 통해 깊이 영상을 생성한다. 세 번째로 정합된 두 영상과 생성된 깊이 영상을 바탕으로 각 픽셀들의 가상 3차원 위치를 계산하여 구하고자하는 가상시점에서의 영상을 생성한다.



(그림 1) 가상시점 생성 절차

가. 카메라 파라미터 추정 및 정합

ERP 영상의 포맷은 위도 θ 와 경도 ϕ 및 반지름이 r (본 논문에서는 1로 고정)인 구좌표계를 따른다. 다수의 ERP 영상의 카메라 상대 파라미터를 추정하기 위해 화소위치를 구면 좌표계에서 직교좌표계(Cartesian coordinate)로 변환한다. 좌표계는 그림 2와 같이 MPEG의 OMAF(Omnidirectional Media Format)[2]의 좌표계를 따른다. 단위 구면에 투영된 구면 좌표와 직교좌표계로 변환하는 식은 다음과 같다.

$$\begin{aligned} x &= -\sin\theta\cos\phi & (1) \\ y &= \sin\theta\sin\phi \\ z &= \cos\theta \\ \theta &= \arccos(z) & (2) \\ \phi &= \arctan(-\frac{y}{x}) \end{aligned}$$

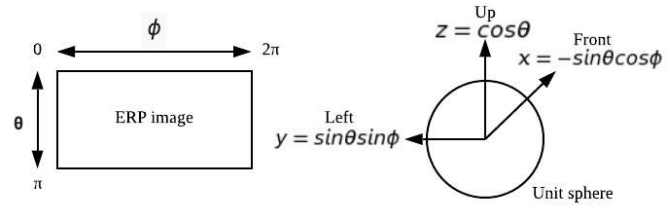
본 논문에서는 ERP 영상에서 SURF를 이용하여 특징점 추출하고 대응점으로부터 8-Point 카메라 파라미터 추정 알고리즘[3]을 사용한다. 평면 투영 이미지와 달리 ERP 투영에서는 래디컬 외곡이 발생하며(특히, 상단 및 하단), 기존의 SURF를 적용하기 위해서 Hajime Taira[4]가 제안한 회전을 통해 외곡이 적은 좌표영역으로 회전을 하는 방법을 사용하였다. 또한 신뢰성을 높이기 위하여 RANSAC(Random sample consensus)에 기반을 둔 알고리즘을 적용한다.

1. N개의 대응점에서 N/4개의 대응점을 무작위로 선택한다.
2. 8-Point 방법을 이용하여 R|t를 구한다.
3. 구해진 R 행렬을 XYZ-Eular 회전 벡터 R_vec 으로 변환한다.
4. 변환된 R_vec을 배열R_vec_arr에 저장한다. 이후 정확하게 측정된 회전 벡터 R_vec를 얻기 위해 다음과 같은 과정을 임의의 M회 수행한다.
 1. R_vec_arr중 m번째 R_vec과 다른 R_vec간 유클리드 놈(Euclidean norm)을 구한다.
 2. 구해진 놈들을 배열 norm_arr에 저장 후, 정렬한다.
 3. 정렬된 놈 중 하위 20% 및 상위 20%를 제외한 norm_arr의 평균을 구한다.
 4. 구해진 norm의 평균을 m번째 R_vec의 오차로 한다.

위와 같은 과정을 통해 구해진 XYZ-Eular 회전 벡터 R_vec 중 오차가 가장 적은 R_vec와 그와 대응되는 이동 벡터 t를 최종 외부 파라미터로 한다.

다음으로, 8-Point 외부 파라미터 추정 실험을 위해 두 개의 실사 ERP 영상을 이용하였다.

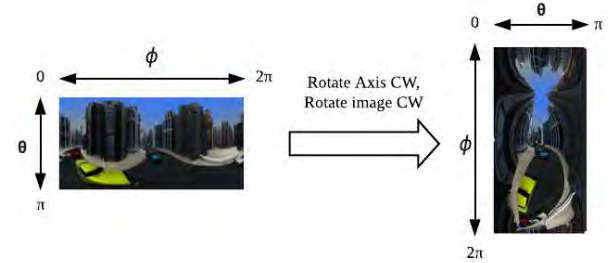
1. 좌우에 위치한 두 카메라 간 최초 회전 행렬 및 이동 벡터 $R_0|t_0$ 를 측정한다.
2. 우측 ERP 영상에 임의의 회전 행렬 R_i 을 적용한다.
3. 두 카메라 간 Essential 행렬을 구한 뒤 $R_e|t_e$ 를 추출한다.
4. 임의로 적용시킨 $R_i = R_e R_0^{-1}$ 이어야 할 것이다.
5. 이동 벡터는 $t_0 = t_e$ 이어야 할 것이다.



(그림 2) ERP 영상의 Spherical coordinate(왼쪽) 및 Cartesian coordinate와 Unit sphere(오른쪽)

나. 깊이추정

기존 원근 투영(Perspective projection)영상 기반의 디스패리티(disparity)계산 및 깊이 영상 생성 방법에 대한 많은 연구가 있다[5]. 본 논문에서는 실시간 처리에 중점을 두고 OpenCV 라이브러리에서 가속화지원이 되는 StereoBM[6]을 이용하여, 기존의 잘 정리되어있는 스테레오 매치(Stereo match) 방식을 ERP 영상에 변형하여 사용하였다. ERP 영상은 투영 방식 특성상 기존의 등극선 기하(Epipolar Geometry)를 그대로 적용할 수 없기 때문에 그림 3과 같이 영상의 좌표축을 시계 방향으로 90도 회전하는 H. Kim 등[7]의 방식을 사용하였다.



(그림 3) 원본 ERP 영상(왼쪽)과 StereoBM 적용을 위해 좌표축 회전 후 다시 영상을 회전한 입력 영상(오른쪽) 디스패리티(시차) 맵을 획득한 후 이를 이용해 각 픽셀들에 해당하는 거리를 계산한다. 그림 4와 같이 B 는 카메라 간 거리이며 h 는 회전된 영상의 가로길 이로 정의했을 때, 구면 좌표계의 각의 차이 d_θ 는

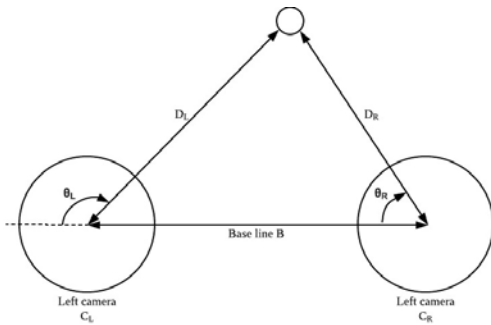
화소 디스패리티 d 로부터 식 3으로 얻어진다. 이로부터 구의 중심에서 투영된 물체까지의 거리 D_L 과 D_R 은 회전된 입력 영상의 구면 좌표계에서의 위도 θ_L, θ_R 에서부터 식 (4)으로 구해진다.

$$d = \theta_L - \theta_R \tag{3}$$

$$d_\theta = d \times \frac{\pi}{h}$$

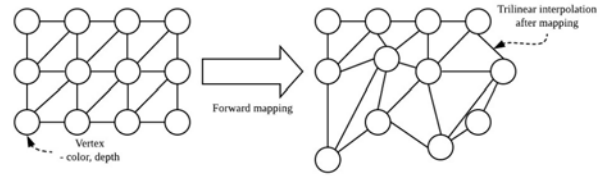
$$D_L = B / \left(\frac{\sin \theta_L}{\tan(\theta_L - d_\theta)} - \cos \theta_L \right) \tag{4}$$

$$D_R = B / \left(\cos \theta_R - \frac{\sin \theta_R}{\tan(\theta_R + d_\theta)} \right)$$

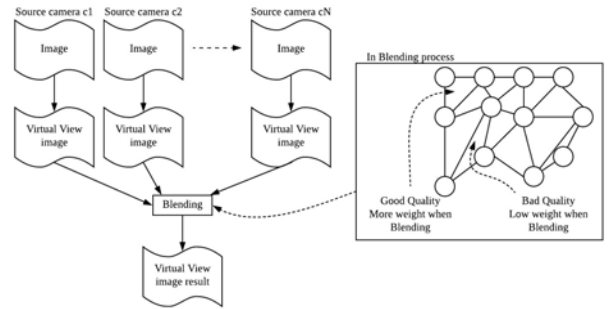


(그림 4) 정합된 두 단위 구 형태의 입력에서의 거리정보다. 가상 시점 생성

DIBR (depth-based image Rendering) 기법은 다시점 영상 생성을 위하여 꾸준히 연구 되어 왔다. 통상의 DIBR 기법은 가장시점의 모든 화소에서의 깊이정보를 필요로 하여 깊이 값을 추정하기 위한, 특히 깊이정보가 없는 occlusion영역을 제거하는 휴리스틱에 대한 연구가 주를 이루었다[8]. 본 연구에서는 OpenGL의 그래픽스 처리기법을 사용한 RVS(Reference View Synthesizer)[9]를 사용한다. RVS는 그림5와 같이 필터링이 필요한 역방향 사상 과정을 거치지 않고 각각의 픽셀을 메쉬 노드(vertex)로 정의하여 인접한 메쉬 삼각형들을 구성한다. 이를 순방향 사상을 진행하여 삼선형 보간법을 이용하여 빈 공간을 제거한다. 그리고 그림6과 같이 복수의 원본 영상으로부터 얻어진 가상 시점 영상들을 혼합함으로써 위치에 의해 발생하는 빈 영역을 제거하고 높은 품질의 결과물을 만들어내도록 한다. 이를 위해 순방향 사상 과정에서 발생하는 각 삼각형들의 변형 정도를 품질척도로 이용하여 삼각형의 변형이 적은 픽셀의 가중치를 크게 주어 최종 결과물의 품질을 향상시킬 수 있다.



(그림 5) RVS의 forward mapping과 삼선형 보간법을 이용한 가상시점 생성



(그림 6) RVS의 Blending 과정

3. 실험

8-Point 외부 파라미터 추정은 다음순서로 두 개의 실사 ERP 영상을 이용하였다. 실사 영상은 외부에서 카메라 간의 간격을 약 1m 로 하여 촬영한 것이다. 측정된 회전 행렬들은 XYZ-Euler 회전 벡터로 변환한 후 유클리드 놈을 구하여 오차를 측정하였다. 또한 추정된 외부 파라미터를 바탕으로 영상 정합을 진행하였다. 그 후 정합된 영상을 바탕으로 생성된 거리 영상은 그림8과 같다.

<표 1> 회전 행렬 추정 결과 및 오차, degree 표현

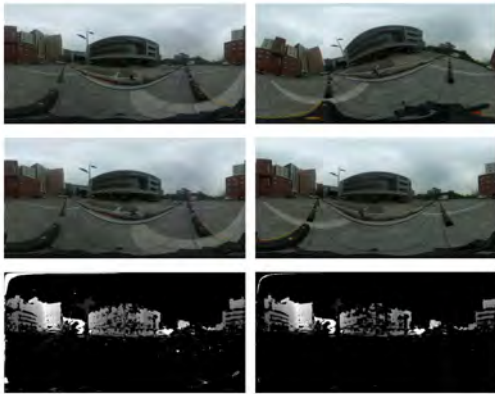
XYZ-Euler vector (x, y, z) of R_i	XYZ-Euler vector (x, y, z) of $R_e R_0^{-1}$	Error
(15, 0, 0)	(14.984, -0.016, 0.247)	0.248
(0, 15, 0)	(0.069, 14.938, -0.166)	0.190
(0, 0, 15)	(-0.304, 0.046, 14.961)	0.309
(15, 15, 0)	(14.884, 14.963, 0.184)	0.220
(15, 0, 15)	(14.782, 0.022, 15.255)	0.336
(0, 15, 15)	(-0.324, 15.029, 15.004)	0.325
(15, 15, 15)	(14.621, 15.189, 15.309)	0.524

<표 2> 이동 벡터 추정 결과 및 오차, 단위 벡터

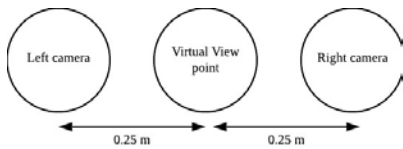
t_0	t_e	Error
(-0.001, -0.999, -0.011)	(0.06, -0.998, -0.002)	0.062
	(0.063, -0.997, -0.004)	0.064
	(0.046, -0.998, -0.019)	0.048
	(0.092, -0.995, 0.001)	0.094
	(0.073, -0.994, 0.069)	0.109
	(0.147, -0.989, 0.003)	0.149
	(-0.045, -0.998, -0.024)	0.046

가상시점 생성 실험을 위해 컴퓨터 그래픽스 도구인 Blender[10]로 만들어진 영상 및 거리 영상을 사용하였다. 이를 RVS 프로그램의 입력으로 하여 가상시점을 생성해보았다. 두 영상의 카메라 위치 및 가상시점의 위치는 그림9와 같으며 그림10과 그림11

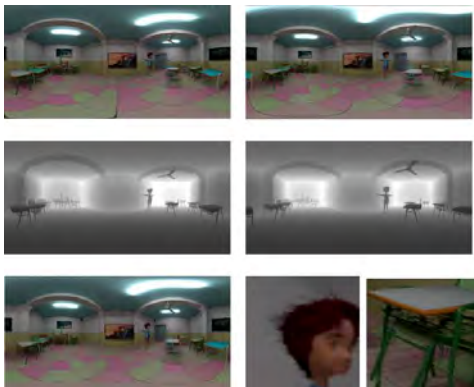
은 생성된 결과를 보여준다.



(그림 8) 정합 전 좌측 영상 및 우측영상(상단), 정합 후 좌측 영상 및 우측 영상(중간), 정합된 영상으로부터 구해진 좌측 및 우측 거리 영상(하단)



(그림 9) 입력 영상과 가상시점의 카메라 위치

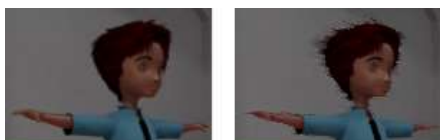


(그림 10) 입력 영상(상단), 입력 거리 영상(중간), 생성된 가상시점(왼쪽 하단), 가상시점의 부분 확대(오른쪽 하단) 생성된 가상시점과 같은 위치에서의 Ground Truth 영상과 PSNR을 구하였다. PSNR 계산은 MPEG의 WS-PSNR[11] 프로그램을 이용하였다.

본 연구에서 다중 360 ERP 영상으로부터 자유시점 영상생성을 위하여 필요한 절차들을 확장하고 실험하였다. 앞으로 효과적인 카메라 추정기법과 깊이 영상 생성에 대한 연구 추가적으로 요구된다.

<표 3> Ground Truth 영상과 생성된 가상시점 영상의 WS-PSNR 결과

YUV채널	Y	U	V
WS-PSNR	36.747	49.324	45.193



(그림 11) 가상 시점 위치에서의 Ground Truth 영상(왼쪽) 과 RVS로 생성된 영상(오른쪽)의 부분 확대

감사의 글

본 연구는 과학기술정보통신부, 정보통신기술진흥센터, 방송통신산업기술개발사업의 지원을 받았음. (시청자 이동형 자유시점 360VR 실감미디어 제공을 위한 시스템 설계 및 기반기술 연구(2016-0-00144).

참고문헌

- [1] 호요성, "MPEG-I 표준과 360도 비디오 콘텐츠 생성", 전자공학회지, vol. 8, pp. 52 - 57, 2017.
- [2] Miska M. Hannuksela, "An Overview of the OMAF Standard for 360° Video", Data Compression Conference, Snowbird, UT, USA, 2019, pp. 580-593.
- [3] H.C. Longuet-Higgins, "A Computer Algorithm for Reconstructing a Scene From Two Projections," Nature, vol. 293, pp. 133 - 135, 1981.
- [4] Hajime Taira, "Robust feature matching for distorted projection by spherical cameras", IPSJ Transactions on Computer Vision and Applications, Japan, 2015, pp. 84-88.
- [5] Heiko Hirschmuller, "Stereo Matching in the Presence of Sub-Pixel Calibration Errors", IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009, pp. 437-444
- [6] Bradski G, "The OpenCV Library", Dr. Dobb's Journal of Software Tools, 2000.
- [7] H. Kim, "3D Scene Reconstruction from Multiple Spherical Stereo Pairs", International Journal of Computer Vision, vol. 104, pp. 94-116, 2013.
- [8] S. Zinger, "Free-viewpoint depth image based rendering", Journal of Visual Communication and Image Representation, vol. 21, pp. 533-541, 2010.
- [9] MPEG, "Reference View Synthesizer (RVS) manual", The 124th MPEG meeting, Macau, China, 2018.
- [10] Community BO, Blender - a 3D modelling and rendering package, Stichting Blender Foundation, Amsterdam, 2018.
- [11] MPEG, "WS-PSNR Software Manual", The 124th MPEG meeting, Macau, China, 2018.