

## 3차원 기하를 이용한 스테레오스코픽 비디오의 흔들림 측정

\*김종유 이현호 박상욱 김정환 김학섭 강지우 \*\*이상훈

연세대학교 전기전자공학과

\*jongky@yonsei.ac.kr \*\*slee@yonsei.ac.kr

## Camera Shake Estimation of Stereoscopic Video Using 3D Geometry

\*Kim Jongyoo Lee Hyunho Park Sanguk Kim Junghwan Kang Jiwoo \*\*Lee Sanghoon

Department of Electrical and Electronic Engineering, Yonsei University

## 요약

최근 3D 비디오의 흥행과 더불어 사람들이 스테레오스코픽 비디오를 쉽게 접할 수 있게 되었다. 하지만 부적절한 스테레오스코픽 비디오는 시청자로 하여금 어지럼증, 눈의 피로 등을 느끼게 할 수 있다. 따라서 비디오 콘텐츠를 분석하여 피로도를 예측할 수 있는 지표의 개발이 필요하다. 스테레오스코픽 비디오에서 피로도를 유발시키는 원인은 여러 가지가 있으며, 본 논문에서는 비디오의 카메라 움직임 정보에 집중한다. 카메라의 흔들림은 심각한 멀미를 유발할 수 있다. 이는 스테레오스코픽 비디오에서 더욱 강렬하게 나타나게 된다. 우리는 스테레오스코픽 비디오의 흔들림을 측정하기 위해 3D 복원 기술을 사용한다. 특히, 좌안 비디오와 우안 비디오의 정보를 모두 활용하는 최적화된 알고리즘을 제안한다. 이를 통해 우리는 각 프레임의 카메라 위치와 방향 정보를 획득할 수 있다. 마지막으로 획득한 3차원 카메라 움직임 정보로부터 카메라의 흔들림을 잘 표현할 수 있는 점수를 계산한다.

## 1. 서론

3D 영화가 크게 성공을 이루게 되면서 3차원 스테레오스코픽 영상이 대중적으로 다가오게 되었다. 많은 디스플레이 제조업체들이 3D TV를 출시하였고, 일반 사람들이 가정에서 쉽게 3D 영화를 접할 수 있게 되었다. 3D 영화란 소위 스테레오스코픽 비디오를 말한다. 좌안을 위한 영상과 우안을 위한 영상이 따로 존재하여 각 눈에 서로 다른 영상을 보여준다. 좌안 영상과 우안 영상에 존재하는 디스패리티에 의해 사람은 깊이 정보를 느끼게 된다. 깊이 정보는 기존의 2D 비디오에서는 느낄 수 없었던 것으로 더욱 깊은 생동감과 몰입감을 시청자에게 전해준다. 3차원 스테레오스코픽 비디오는 장점과 동시에 해결해야 할 문제를 안고 있다. 스테레오스코픽 비디오는 사람 눈의 감각을 속여 일반적인 것과는 다른 상황을 통해 입체감을 만들어낸다. 이에 따라 3D 영화를 시청하는 사람이 어지럼증, 구토, 눈의 피로 등을 느끼기도 한다. 따라서 3차원 스테레오스코픽 비디오가 인체에 잠재적으로 미칠 수 있는 영향을 연구하여 이를 미연에 방지할 수 있도록 하는 연구가 필요하다.

스테레오스코픽 비디오가 피로를 유발하는 원인은 여러 가지가 있다. 과도한 디스패리티에 의한 피로는 많은 연구가 이루어졌다. 이를 방지하기 위해 많은 실험을 통해 사람 편안한 영역을 정의하는 연구도 있었다 [1, 2]. 하지만 스테레오스코픽 비디오에서의 움직임 특징에 따른 피로도 연구는 부족하다. 몇몇 연구에서 깊이 방향의 움직임에 의해 피로감을 느낄 수 있다고 하였다 [3, 4]. 본 논문에서는 움직임 성분 중에서도 특히 카메라의 움직임에 중점을 둔다.

카메라가 부적절하게 움직이면 시청자가 어지럼증을 느끼게 된다.

기존의 2D 비디오에서도 이러한 현상이 존재한다. 사람이 직접 손에 들고 촬영한 영상을 볼 때 멀미와 같은 증상을 느낀다. 카메라의 부적절한 움직임이 3차원 스테레오스코픽 비디오에서 발생한다면 더욱 심각한 증상을 초래하게 된다 [5]. 더욱 큰 몰입감이 있기 때문에 짧게 시청하더라도 많은 어지럼증을 호소하게 된다.

본 연구에서는 스테레오스코픽 비디오의 흔들림을 측정할 수 있는 알고리즘을 제안한다. 기존의 연구들이 카메라의 움직임을 측정하기 위해 2차원 모션 흐름 정보를 활용하였던 것에 반해 본 연구에서는 3차원 기하 정보를 활용한 시스템을 제안한다. 여러장의 이미지로부터 3차원 모델을 복원해내는 SfM (Structure from Motion) 기술을 활용한다. 특히 3차원 스테레오스코픽 영상은 베이스라인이 고정되어 일정한 x축 거리를 유지한다는 특징을 활용한다. 획득한 카메라의 3차원 위치를 이용하여 카메라의 흔들림을 표현할 수 있는 특징을 추출해낸다.

## 2. 스테레오스코픽 카메라의 3차원 움직임 측정

SfM 기술은 3차원 구조를 복원하기 위해 사용되는 컴퓨터 비전 알고리즘이다. 여러 장의 이미지들을 입력으로 받아 카메라의 3차원 포즈 및 간략한 3차원 구조를 복원해낸다 [6]. 일반적인 다중 시점에 대한 SfM 기술은 대한 단일 카메라로 촬영된 영상을 대상으로 한다. 이미지들은 비디오 스트림이 아니라 비구적으로 조합되어 있다. 따라서 SfM을 수행하기 위해서는 모든 이미지에 대해 특징 추출 및 매칭이 선행된 후에 적절한 초기 이미지를 선택한 후 가능하다.

스테레오스코픽 비디오는 좌안영상과 우안 영상을 갖는다. 일반적

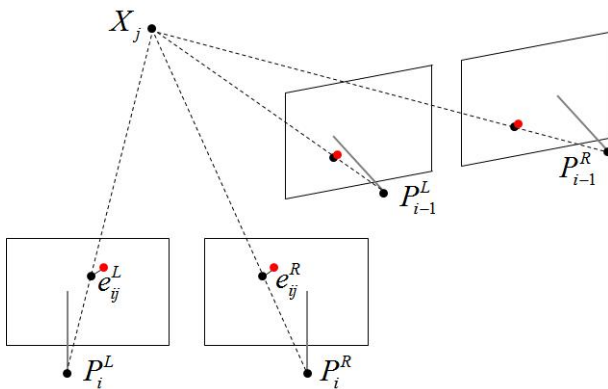


그림 1 스테레오 카메라의 기하

으로 사용하는 평행 방식의 스테레오 카메라는 좌안용 카메라와 우안용 카메라가 평행한 z축을 갖으며 카메라 리그에 의해서 고정된다. 따라서 이를 이용해 촬영한 스테레오 영상은 x축에 대한 디스페리티만 존재하게 된다. 카메라의 초점거리, 주점의 위치정보인 캘리브레이션 정보를 안다면 디스페리티로부터 실제 깊이를 측정하는 것이 가능해진다. 이 정보를 활용함으로써 모든 스테레오스코픽 비디오를 첫 프레임일부터 순차적으로 한 프레임씩 고려하여 카메라 위치 추정이 가능하다.

초기 구조는 첫 프레임을 이용하여 획득한다. 좌안 영상과 우안 영상의 특징점 매칭을 수행한다. 캘리브레이션 정보를 활용하여 매칭이 이루어진 특징점들은 3차원 공간상에 역으로 투영하여 간소한 형태의 초기 포인트 클라우드를 얻는다.

추가적인 매 프레임에 대하여 최적화 알고리즘을 수행하여 카메라 위치를 획득해낼 수 있다. 우선 새로운 프레임의 이미지에 대하여 특징점 추출을 수행한 뒤 3차원으로 복원된 점들과 매칭을 수행한다. 그림 1은 풀고자 하는 문제의 3차원 기하를 보여준다. 3차원 상에 복원되어 있는 점을  $X$ 라 할 때, 이 3차원 점은 카메라 매트릭스  $P$ 에 의하여 이미지에 투영되어 2차원 점  $x$ 로 나타난다. 이 관계는 간단히  $x = PX$ 로 표현된다. 하지만 실제로 측정을 통해 얻은 카메라 매트릭스는 이상적이지 않기 때문에 오차가 발생한다. 이 오차는 좌안 영상과 우안 영상에 대하여 각각 나타나게 된다.

$$e_{ij}^L = D(P_i^L(r_i, c_i)X_j - x_{ij}^L)$$

$$e_{ij}^R = D(P_i^R(r_i, c_i)X_j - x_{ij}^R)$$

여기서  $D(\dots)$ 는 두 점 사이의 유클리디언 거리를 나타내며,  $j$ 는 3차원 포인트의 인덱스,  $i$ 는 프레임의 인덱스를 나타낸다. 카메라 매트릭스는 카메라의 회전정보를 나타내는  $r$ 과 이동 정보를 나타내는  $c$ 에 의해 구성된다. 따라서 새로운 카메라 매트릭스를 획득하기 위해 다음의 코스트를 최소화시키는 최적화를 수행한다.

$$\min_{r_i, c_i} \left[ \sum_j (e_{ij}^L + e_{ij}^R) \right]$$

LM 알고리즘을 활용하여 투영 오차를 최소화시키는  $r_i$ 와  $c_i$ 를 얻을 수 있다. 이는 새로운 카메라 매트릭스  $P_i^L$ 와  $P_i^R$ 를 만들어낸

다. 이후 새롭게 추가된 이미지와 기존의 이미지와 특징 매칭을 수행하여 새로운 3차원 포인트를 추가한다. 일련의 과정이 반복되면서 새로운 프레임의 카메라 위치를 획득할 수 있다.

### 3. 흔들림 측정

앞서 구한 스테레오스코픽 카메라의 위치는 매 프레임의 카메라 매트릭스로 나타내어진다. 각각의 카메라 매트릭스는 해당하는 프레임의 카메라의 위치 및 회전 정보를 포함한다. 카메라 매트릭스의 스트림으로부터 카메라의 흔들림을 수치화할 수 있는 수치를 측정한다.

일반적으로 스테레오 카메라를 이용하여 촬영 시 주요한 객체에 영점을 맞추게 된다. 영점이란 좌안 영상과 우안 영상에서 디스페리티가 0이 되는 z축 거리를 말한다. 즉, 시청자가 일반적으로 주시하는 객체에서의 흔들림이 가장 잘 표현될 수 있다. 3차원 상에서 영점의 위치는 카메라 매트릭스로부터 쉽게 획득할 수 있다.

$$Z_0 = \begin{bmatrix} R^T & -t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ z_0 \\ 1 \end{bmatrix}$$

여기서  $z_0$ 는 영점의 z축 거리를 말한다. 위의 수식을 이용하여 모든 프레임에서의 영점위치를 알 수 있으며 이를 이용하여 카메라의 흔들림을 표현할 수 있는 지표로 만든다. 우리는 이를 점의 이동 점수 PMS라고 정의하고 다음과 같이 구한다.

$$PMS_i = \sqrt{\sum_{k=i-t_0}^i (\Delta Z_k - \mu_{Z_i})^2}$$

본 논문에서는  $t_0$ 를 1초로 두어 실험하였다. PMS는 영점 포인트의 흔들림을 표현하였으며 이는 카메라 움직임 중 롤 움직임에 대한 것은 포함하지 못하고 있다. 따라서 카메라의 z축에 대한 회전 움직임을 반영하는 다른 지표가 필요하다. 우리는 롤 움직임 점수 RMS라 정의한다. 카메라가 z축에 대하여 얼마나 회전하였는가를 구하기 위해서 우리는 회전 매트릭스로부터 오일러 각도로 변환을 한다.

우선 y축과 z축에 대한 참조 벡터를 표현하기 위해서 간단하게  $a_x = (0, 1, 0)^T$ ,  $a_z = (0, 0, 1)^T$ 으로 같이 정의한다. 이 때 i번째 카메라의 y축과 z축은 다음과 같이 나타난다.

$$a'_x = R_i a_x, \quad a'_z = R_i a_z$$

우리가 알아야 하는 것은  $r_y$ 와  $r'_z$ 에 의해 생기는 평면과  $r'_y$ 사이의 각도이다. 우리는 평면의 법선 벡터  $n = [a_y]_{\times} a'_z$ 를 이용할 수 있다. 이 때 우리가 롤에 해당하는 각도는 다음과 같이 구한다.

$$\theta = \frac{\pi}{2} - \cos^{-1} \frac{a'_y \cdot n}{|a'_y| |n|}$$

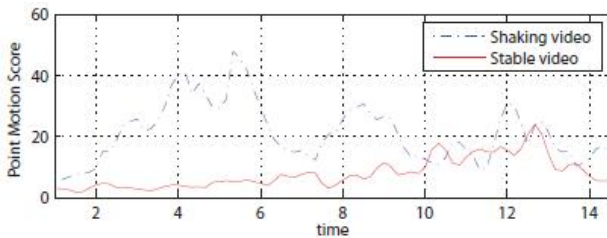
최종적으로  $\theta$ 의 표준편차를 이용하여 롤 움직임 점수 RMS를 구한다.

$$RMS_i = \sqrt{\sum_{k=i-t_0}^i (\Delta\theta_k - \mu_{\theta_i})^2}$$

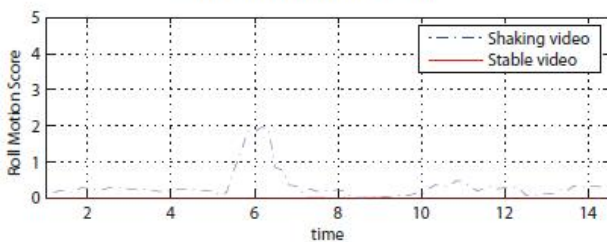
#### 4. 실험 결과

제안한 알고리즘을 사용하여 두 가지 종류의 비디오에 대하여 실험을 진행하였다. 첫 번째 비디오는 흔들림이 거의 없이 안정적인 상황에서 찍혔다. 두 번째 비디오는 손에 들고 촬영하여 흔들림이 존재한다. 두 테스트 비디오 모두 15초의 길이로 편집하였으며 매 5프레임마다 골라서 총 90 프레임을 사용하였다.

각각의 비디오는 앞서 제안한 스테레오스코픽 비디오에 최적화된 카메라 움직임 측정 알고리즘을 사용하였다. 모든 프레임에 대하여 카메라 매트릭스를 획득한 후 이를 이용하여 PMS와 RMS를 획득하였다. 실험 결과는 그림 2에 나타나 있다. 점수가 높을수록 흔들림이 더욱 심한 것을 표현한다. 직관적으로 흔들리는 영상의 PMS 및 RMS가 크게 나타난다는 것을 확인할 수 있다. 실험을 통해 복원된 카메라의 움직임과 3차원 구조는 그림 3에 나타나 있다. 좌쪽으로 파란색 점과 빨간색 선으로 연결된 부분이 복원된 카메라를 나타내며, 중앙에 녹색 점은 복원된 3차원 구조를 나타낸다. 또한 3차원 구조의 중앙에 파란색 점과 붉은색 선으로 연결도니 것은 영점의 이동을 나타낸다. 복원된 카메라의 움직임을 통해서 (b)의 영상에 흔들림이 존재함을 알 수 있다.



(a) Point motion score



(b) Roll motion score

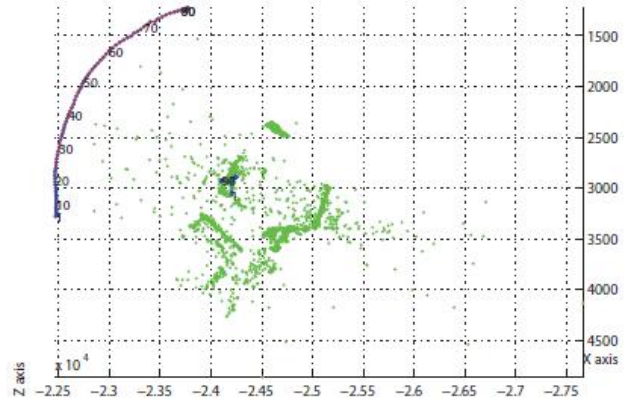
그림 2 테스트 비디오의 점수

#### 감사의 글

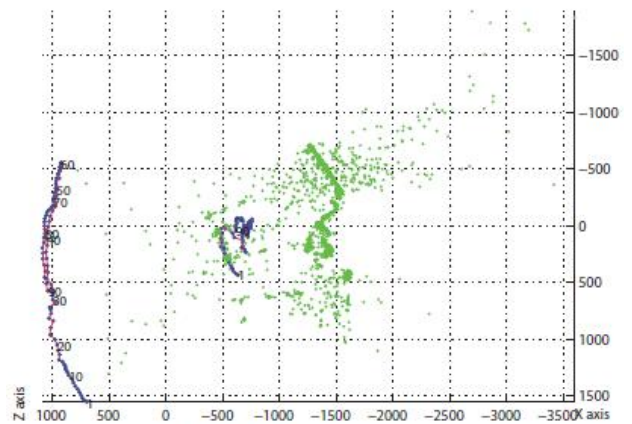
본 연구는 미래창조과학부 및 정보통신산업진흥원의 IT융합 고급인력 과정 지원사업의 연구결과로 수행되었음 (NIPA-2013-H0401-13-1003)

#### 참고문헌

[1] M. Wopking, "Viewing comfort with stereoscopic pictures: An experimental study on the subjective effects of disparity magnitude and depth of focus," *Journal of the SID*, vol. 3, pp. 101-103 (1995)



(a) Stable movement



(b) Shaking movement

그림 3 복원된 3차원 구조와 카메라 움직임

[2] M. Lambooj, W. Ijsselsteijn, M. Fortuin, and I. Heynderickx, "Visual discomfort and visual fatigue of stereoscopic displays: A Review," *J. Imag. Sci. Technol.*, vol. 53, pp. 0302011-03020114, (2009)

[3] S. Yano, M. Emoto, and T. Mitsuhashi, "Two factors in visual fatigue caused by stereoscopic HDTV images," *Displays*, vol. 25, no. 4, pp. 141-150, 2004.

[4] S. I. Lee, Y. J. Jung, H. Sohn, Y. M. Ro, and H. W. Park, "Visual discomfort induced by fast salient object motion in stereoscopic video."

[5] H. Ujike, and H. Watanabe, "Effects of stereoscopic presentation on visually induced motion sickness."

[6] R. Hartley and A. Zisserman, "Multiple view geometry in computer vision," 2000.