

정규논문 (Regular Paper)

방송공학회논문지 제19권 제2호, 2014년 3월 (JBE Vol. 19, No. 2, March 2014)

<http://dx.doi.org/10.5909/JBE.2014.19.2.205>

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

스테레오스코픽 3D 콘텐츠 제작의 효율성 향상을 위한 자동 영상정렬 및 모니터링 기법

김재인^{a)}, 김태정^{a)†}

Automated Image Alignment and Monitoring Method for Efficient Stereoscopic 3D Contents Production

Jae-In Kim^{a)} and Taejung Kim^{a)†}

요 약

고품질의 스테레오스코픽 3D 콘텐츠를 제작하기 위해서는 입체피로를 유발하는 문제요인들을 최소화하는 것이 중요하다. 촬영과정에서 스테레오 좌우 영상 간에 발생하는 수직시차는 입체피로의 주된 요인으로, 정확한 입체시를 유지하기 위해 필수적으로 제거될 필요가 있다. 본 논문에서는 콘텐츠 제작의 효율성 향상을 위하여 후처리 과정에서뿐만 아니라 촬영과 동시에 영상정렬이 수행된 결과를 모니터링할 수 있는 자동화된 방식의 영상처리 기법을 제안하였다. 제안방법은 대응점 추출과 기하구조 추정, 그리고 수직시차의 제거로 이어지는 편위수정 부분과 카메라 움직임 탐지 부분으로 구성되어 있으며, 각각에 대해 성능분석이 실시되었다. 실험결과, 제안방법은 비교분석을 위해 사용된 기존방법들에 비해 보다 향상된 성능을 나타냈으며, 카메라 움직임 탐지부분에서도 98.35% 성공률의 뛰어난 탐지성능을 보여주었다. 이러한 일련의 성능검증을 통해 실제 콘텐츠 제작현장에서 본 논문의 제안방법이 효과적으로 활용될 수 있음을 확인할 수 있었다.

Abstract

Minimization of visual fatigue is important for production of high quality stereoscopic 3D contents. Vertical disparity of stereo images occurred during contents production is considered as the main factor of visual fatigue. To ensure correct stereoscopy vertical disparity needs to be eliminated. In this paper, a method for automated image alignment was proposed for Stereoscopic 3D contents generation and post-processing steps. The proposed method consists of two parts: rectification for image alignment and camera motion detection. The proposed method showed that its rectification performance was the most superior among the existing methods tested and that camera motion detection had a success rate of 98.35%. Through these evaluations, we confirmed that the proposed method can be effectively applied to 3D contents production.

Keyword : Stereoscopic 3D contents, Image alignment, rectification, camera motion detection

a) 인하대학교 지리정보공학과 (Department of Geoinformatic Engineering, Inha University)

† Corresponding Author : 김태정(Taejung Kim)

E-mail: tejid@inha.ac.kr

Tel: +82-32-860-7606

※ 본 연구는 인하대학교의 지원과 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(2011-0009721).

· Manuscript received January 10, 2014 Revised February 21, 2014 Accepted February 21, 2014

I. 서론

스테레오스코픽 3D(Stereoscopic 3D, S3D) 콘텐츠의 제작방법은 크게 세 가지로 구분된다. 그 중 첫 번째는 컴퓨터 그래픽 기술을 이용하여 콘텐츠를 제작하는 컴퓨터그래픽 방식이다. 이 방식은 사전에 의도한 대로 입체감을 정확히 구현할 수 있어 고품질의 콘텐츠 제작이 가능하다는 장점이 있다. 그러나 제작비용 자체가 워낙 고비용일뿐더러 현실적 표현에 한계가 있어 애니메이션 제작이나 또는 실사 촬영의 보조적 수단으로 활용되는 경우가 대부분이다. 두 번째는 기존의 2D 콘텐츠를 3D로 변환하는 방식이다. 이는 말 그대로 기존의 콘텐츠를 재활용 할 수 있다는 점에서 큰 장점을 가질 수 있다. 그러나 전문화된 저작도구가 부재하며, 또한 많은 노동력이 요구될 뿐 아니라 애초에 3D 화면 출력을 고려하여 제작된 콘텐츠가 아니기 때문에 입체 품질이 좋지 못하다는 단점이 있다. 마지막으로 실사방식은 스테레오로 구성된 두 대의 카메라를 이용하여 실제 대상 장면을 촬영하는 방식이다. 이 방식은 입체시의 원리에 입각하여 실제 대상 장면을 촬영하기 때문에 앞서 두 방식에 비해 현실감 있는 고품질의 입체표현이 가능하다는 장점이 있다. 이 때문에 콘텐츠 제작방법으로는 가장 일반적으로 사용되고 있다. 그러나 이 방식 역시 전문 인력 및 기술의 부족과 값비싼 장비가 요구되어 제작에 많은 어려움이 따르고 있는 실정이다^{1,2)}.

실사방식으로 콘텐츠를 제작하는데 있어 기술적 주요 애로사항들은 대부분 시청자가 콘텐츠를 시청할 때 느끼게 되는 시각적인 피로, 즉 입체피로와 밀접한 관계를 가지고 있다. 입체피로는 크게 좌우 영상간의 색감, 밝기 및 선명도 차이로 인해 발생하는 광학적 문제와 카메라들의 자세 및 위치 등이 잘못 조정되어 발생하는 기하학적 문제로 나눌 수 있다. 이들 중에 특히 기하학적 문제의 경우는 3D 콘텐츠의 동작원리인 입체시 수립과 매우 밀접한 관계가 있기 때문에 입체피로를 최소화하여 고품질의 콘텐츠를 제작하기 위해서는 반드시 선결되어야 할 문제로 인식되고 있다^{2,3)}.

기하학적 문제는 카메라 간에 발생하는 수직시차를 최소화하고 사전에 의도한 대로 적합한 수평시차를 조정하는

것이 관건이다. 여기서 수평시차는 실제 시청자가 느끼게 될 입체감을 의미하기 때문에 콘텐츠 제작에 있어 매우 중요한 요소로 다뤄지지만 사실상 이는 수직시차가 완벽히 제거되었다는 가정하에 그 의미를 가질 수 있기 때문에 기하학적인 문제들 중에서 가장 선결되어야 할 것은 수직시차 제거를 위한 영상정렬이라 볼 수 있다. 이는 입체시 수립의 기본 전제조건으로, 만약 촬영된 영상에 정렬오류로 인한 수직시차가 발생되었을 경우, 결과적으로 어지러움과 두통과 같은 심각한 입체피로가 유발될 수 있다.

이런 이유로 실제 촬영 전에는 수직시차를 최소로 하기 위해 카메라 리그(Rig)라 불리는 고가의 장비를 이용하여 영상정렬 작업을 진행하고 있다. 그러나 이 과정은 전문 인력 및 기술이 요구되에도 불구하고 모든 과정이 수작업으로 진행되기 때문에 완벽하게 수직시차가 제거된 상태로 촬영이 이루어지기가 현실적으로 매우 어려우며, 또한 촬영 도중에 카메라의 자세 및 위치가 변경될 때마다 반복적으로 동일한 과정이 재수행되어야 한다는 점에서 콘텐츠의 제작을 더욱 더 어렵게 만들고 있다. 그뿐 아니라, 이런 이유로 제작 후반부 과정에서는 잔존하는 수직시차를 완전히 제거하기 위해 추가적인 후처리 작업이 이루어지지만 의도치 않게 과도한 시차가 발생된 경우에는 재촬영이 불가피하여 전체 작업진행에 큰 차질을 야기하게 될 소지가 있다. 그렇기 때문에 이러한 문제를 미연에 방지하기 위해서는 촬영과 동시에 최종결과를 확인해 볼 수 있는 신뢰적인 영상정렬 모니터링 기술이 필요하다.

따라서 본 논문에서는 이러한 현실정에서 실제 후처리 과정에서도 사용될 만큼 정확하고 신뢰적이며 또한 카메라 자세 및 위치가 변경될 때마다 자동적으로 영상의 수직시차 정렬을 가능케 할 수 있는 자동화된 방식의 영상정렬 및 결과 모니터링 기법을 제안하고자 한다.

II. 배경이론 및 연구사례

영상간의 수직시차 제거를 위한 영상정렬은 편위수정(Rectification)이라는 작업을 통해 이루어지게 된다. 이 편위수정은 크게 기하추정과 영상변환의 두 과정으로 구성되

며, 이 때 영상변환의 경우는 추정된 기하구조를 바탕으로 이루어지기 때문에 결과적으로 전체 편위수정의 정확도는 기하추정의 정확도에 종속적이다. 영상정렬을 통해 수직시차가 제거되면 두 영상사이의 모든 대응점들은 동일한 수직좌표를 가지게 된다. 이 때문에 편위수정은 컴퓨터비전과 사진측량분야 모두에서 공간정보 추출을 위한 영상정합(Image matching)의 전처리 과정으로 매우 중요하게 다루어져 왔다.

편위수정은 일반적으로 에피폴라 기하구조를 바탕으로 이루어진다. 에피폴라 제약조건은 이러한 기하구조를 충족시키기 위한 전제조건으로 다음의 그림 1과 같이, 두 영상의 투영중심 O, O' 과 공간상의 한 점 P , 그리고 각 영상에 투영된 지점 p, p' 이 공통의 한 평면(Epipolar plan)상에 존재해야 함을 의미한다.

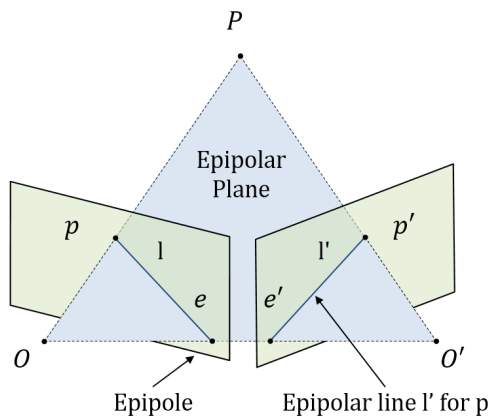


그림 1. 스테레오 영상의 에피폴라 기하
 Fig. 1. Epipolar geometry of stereo image pair

위 그림 1의 에피폴라 기하에서 에피폴(Epipole)은 두 투영중심을 연결한 기선이 각각의 영상평면과 만나는 지점으로서 영상의 모든 에피폴라선(Epipolar line)은 이 에피폴을 지나게 된다는 특징이 있다. 여기서 에피폴라선은 \overrightarrow{OP} (혹은 $\overrightarrow{O'P}$)가 다른 한 쪽 영상평면에 투영되었을 때 나타나는 직선 l' (혹은 l)을 의미한다. 이는 좌측영상의 특정 한 점에 대한 우측영상의 대응점은 에피폴라선상에 존재하게 됨을 의미한다.

전술한 에피폴라 제약조건을 수식으로 나타내면 다음과 같이 쓸 수 있다. 여기서 T, P_1, P_2 는 각각 $T = \overrightarrow{OO'}$, $P_1 = \overrightarrow{OP}$, $P_2 = \overrightarrow{O'P}$ 을 의미한다.

$$(P_1 - T)^T \cdot [T \times P_1] = 0 \quad (1)$$

이를 행렬로 표현하면 $T = [T_x \ T_y \ T_z]$ 라 했을 때, 다음과 같은 형태로 나타낼 수 있다. 여기서 R 은 기준좌표계에 대한 각 영상의 3×3 회전행렬을 의미한다.

$$P_2^T \begin{bmatrix} 0 & -T_z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{bmatrix} P_1 = P_2^T S P_1 = p_2^T R_2 S R_1 p_1 = 0 \quad (2)$$

그리고 다시 위 식은 영상좌표 q 와 카메라좌표 p 사이의 관계, 즉 $q = Mp$ 로부터 다음의 식 (3)과 같이 영상좌표계로 표현될 수 있다. 여기서 행렬 M 은 카메라행렬로서 초점거리 f 와 영상의 주점좌표 s_x, s_y 로 구성되어 있다.

$$q_2^T (M_2^{-1})^T R_2 S R_1^T M_1^{-1} q_1 = q_2^T F q_1 = 0 \quad (3)$$

$$q = \begin{bmatrix} 1 & 0 & s_x/f \\ 0 & 1 & s_y/f \\ 0 & 0 & 1/f \end{bmatrix} \begin{bmatrix} x \\ y \\ f \end{bmatrix} = Mp \quad (4)$$

따라서 스테레오 영상의 기하 추정 은 결론적으로 위의 식 (2)에서 회전행렬 R 과 행렬 S 의 구성요소인 기선벡터 T 를 복원하는 과정이라 할 수 있다. 여기서 컴퓨터비전방식과 사진측량방식의 가장 큰 차이점은 추정을 위한 미지수 정의와 추정에 사용되는 방법론에 있다.

기하추정이 수행되고 난 뒤에는 이를 바탕으로 수직시차 제거를 위한 영상변환 행렬이 계산된다. 이는 영상의 에피폴을 무한대로 사상(Mapping)시키는 변환행렬로서 결과적으로 영상의 에피폴라선을 수평으로 정렬시킴으로서 수직시차를 제거하게 된다. 이 때, 중요한 것은 변환에 따른 결과로 영상왜곡이 발생될 수 있다는 점이다. 이 때문에 영상변형을 최소화 하도록 변환행렬을 설계하는 것이 매우 중요하다.

현재 스테레오스코픽 3D 콘텐츠 제작기술과 관련된 연구는 컴퓨터비전분야에서 활발히 진행되고 있다. 컴퓨터비전의 방식에서는 일반적으로 식 (3)에서 F로 나타낸 3×3의 기본행렬(Fundamental matrix)을 미지수로 간주한다. 가장 대표적인 방법으로는 Hartley^[4]가 발표한 8점 알고리즘이 있다. 그러나 이 방법의 경우는 특이값 분해(Singular Value Decomposition, SVD)를 사용하는 선형해법이기에 때문에 기하추정에 사용되는 대응점 오차에 민감할 뿐만 아니라 영상의 비정렬도가 큰 경우에는 제대로 동작하지 못하는 경우가 발생하곤 한다. 이 때문에 실제로는 비선형 해법을 위한 초기값 계산에 자주 사용되고 있다. 이러한 비선형 해법들은 보통 비용함수(Cost function)를 정의하고 Levenberg-Marquardt 기법과 같은 비선형 최적화 기법 등을 통해 미지수들을 추정한다. 에피폴라선과 대응점간의 거리를 이용한 방법들이 연구된 바 있으며^[5,6], 이와 유사하게 특정 기하학적 관계를 이용한 방법들이 연구된 바 있다^[7,8]. 이와 다른 관점에서 수행된 연구로는 영상정렬 시 원영상 대비 변형을 최소화하기 위해 고안된 Quasi-Euclidean 방법이 있다^[9].

그러나 상기의 방식들은 비선형 최적화 기법을 통해 미지수들을 추정함에 따라 자칫 국소적 최소에 빠질 우려가 있으며, 또한 카메라 내부(초점거리, 주점의 위치, 렌즈왜곡 계수 등) 및 외부 파라미터(자세 및 위치) 모두를 미지수로 간주하기 때문에 보다 일반화된 방법일 수는 있으나 추정을 위한 정합점 분포 및 오차에 민감하게 반응한다는 문제점이 발생한다. 이를 해결하기 위한 방법으로 카메라 외부 파라미터만을 미지수로 간주하여 비선형 최소제곱추정 기법을 통해 영상 기하구조를 추정하고 수직시차를 제거한 사진측량의 접근방식이 제안된 바 있다^[10,11].

III. 제안 알고리즘

본 논문의 제안방법이 실제 스테레오스코픽 3D 콘텐츠 제작분야에서 활용되기 위해서는 편위수정의 정확도뿐만

아니라 촬영환경의 다양성에 대비하여 기하추정을 위한 대응점의 위치오차, 과도한 카메라 틸트(Tilt) 및 기선거리(Baseline) 등에 강인성이 보장되어야 한다. 또한 작업 소요 시간 단축을 위해서도 전과정의 고속화 및 자동화가 가능해야 할 것이다. 이를 위해 본 논문에서는 제안방법을 구성하는데 있어 일반적인 편위수정의 작업절차인 대응점 추출, 기하구조 추정, 수직시차 제거과정 외에도 카메라 움직임 탐지과정을 추가적으로 고려하여 각각의 단계에서 상호 유기적으로 상기의 요구사항들을 충실히 반영할 수 있도록 알고리즘을 설계하였다.

먼저 대응점 추출단계에서는 영상 상호간의 기하구조 추정을 위한 대응점 추출이 이루어진다. 이 단계에서는 전체 작업의 자동화 및 고속화 달성을 위해 특징점 기반 정합기법을 사용하여 자동화된 방식으로 대응점을 추출할 수 있도록 하였다. 이를 위해 빠른 처리속도가 장점인 SURF (Speed Up Robust Feature) 알고리즘^[12]을 이용하였으며, 보다 고속화를 위해 GPU 기반 병렬 컴퓨팅 기법을 적용하였다. 또한 이러한 방식으로 추출된 대응점에는 작은 위치오차 뿐만 아니라 오정합으로 인해 발생된 과대오차까지도 포함될 수 있으므로 이들을 걸러내기 위한 방법으로 RANSAC (RANdom Sample Consensus) 알고리즘^[13]을 적용하였다.

기하구조 추정단계에서는 기존 컴퓨터비전방식과는 달리 공면조건에 기반한 사진측량방식의 상대표정 알고리즘을 이용하여 스테레오 영상의 상호 기하구조를 추정하고자 하였으며 이 때, 대응점에 잔존하는 위치오차, 과도한 카메라 틸트 및 기선거리와 같은 오차요인들에 대해서 강인성을 확보하고자 사전 기하학적 제약조건을 도입한 베이지안(Bayesian) 접근방식을 적용하였다. 이러한 접근방식의 특징은 관측치로 사용되는 대응점까지도 상대표정요소들과 함께 최소제곱추정과정에서 미지수로 간주된다는 점이다. 또한 사전 제약조건으로 주어진 각 미지수들의 공분산행렬은 최소제곱추정과정에서 가중치의 개념으로 사용되어 추정의 안정성을 높일 수 있게 해준다^[14,15]. 이러한 접근방식을 본 논문의 목적에 맞게 변형하여 수식으로 표현하자면

1) 상대표정요소는 스테레오 영상 상호간의 자세 및 위치차이를 나타낸다. 일반적으로 좌측영상을 기준으로 우측영상의 X, Y, Z축 회전요소 ω , ϕ , κ 와 Y, Z축 방향 위치요소 T_Y , T_Z 가 이에 포함되며, X축 위치요소는 모델공간의 축척과 관계되어 보통 상수로 간주된다.

2장의 공면조건식 (2)를 모델방정식으로 하였을 때, 정규방정식과 미지수들의 보정값 행렬은 다음의 식 (5)와 같이 쓸 수 있다.

$$B\Delta y + \varepsilon = w, \Delta y = (B^T C^{-1} B)^{-1} B^T C^{-1} w \quad (5)$$

여기서 설계행렬 B 와 미지수 보정행렬 Δy , 사전 공분산행렬 C , 그리고 잔차행렬 w 는 다음과 같이 정의된다.

$$B = \begin{bmatrix} A_p A_x \\ I_p 0 \\ 0 I_x \end{bmatrix}, \Delta y = \begin{bmatrix} \Delta p \\ \Delta x \end{bmatrix}, w = \begin{bmatrix} w_F \\ w_p \\ w_x \end{bmatrix}, C = \begin{bmatrix} C_F 0 0 \\ 0 C_p 0 \\ 0 0 C_x \end{bmatrix} \quad (6)$$

행렬 B 에서 A_p 와 A_x 는 각각 표정요소와 대응점에 대한 설계행렬로서 n 개 대응점에 대해서 $n \times 5$, $n \times 2n$ 의 크기를 가진다. 이들 각각은 모델방정식인 식(2)를 F 라 했을 때, 다음과 같이 F 에 대한 각 요소들의 편미분 값으로 구성된다.

$$A_p = \begin{bmatrix} \frac{\partial F}{\partial \omega} & \frac{\partial F}{\partial \phi} & \frac{\partial F}{\partial \kappa} & \frac{\partial F}{\partial T_Y} & \frac{\partial F}{\partial T_Z} \end{bmatrix} \quad (7)$$

$$A_x = \begin{bmatrix} \frac{\partial F}{\partial x_2} & \frac{\partial F}{\partial y_2} & \dots & \frac{\partial F}{\partial x_2^n} & \frac{\partial F}{\partial y_2^n} \end{bmatrix} \quad (8)$$

잔차행렬 w 에서 w_F 는 모델방정식에 대한 $n \times 1$ 의 잔차행렬을 의미하며, w_p 와 w_x 는 각각 표정요소와 대응점에 대한 5×1 , $2n \times 1$ 의 잔차행렬을 의미한다. 마지막으로 행렬 C 에는 미지수들의 사전 제약조건이 정의된다. $C_F(n \times n)$ 는 각각의 대응점에 대해서 모델방정식의 잔차가 가질 수 있는 분산을 의미하며, $C_p(5 \times 5)$ 와 $C_x(2n \times 2n)$ 는 표정요소와 대응점의 추정값이 가질 수 있는 분산을 의미한다. 본 논문의 경우에는 스테레오스코픽 3D 콘텐츠 제작에 활용을 목적으로 하고 있기 때문에 이를 반영하여, 실제 촬영현장에서 좌우 카메라의 위치는 최대한 수평을 유지하도록 조정하고 있다는 점으로부터 T_Y , T_Z 에 대한 사전 제약조건은 T_X 를 100으로 설정하였을 때, 약 1 정도의 값으로 정의하였으며, 나머지 회전요소에 대해서는 약 30°로 정의하였다. 그리고

개별 대응점들은 상호 독립적으로 추출되었고 또한 각각의 표정요소 역시 상호 독립적인 관계를 가지기 때문에 행렬 C 의 대각행렬요소를 제외한 나머지는 모두 0으로 설정하였다.

기하추정이 완료된 다음에는 이를 바탕으로 수직시차 제거단계에서 영상변환이 이루어진다. 이 과정을 통해 변환된 좌우 영상으로부터 입체시가 가능해진다. 본 논문에서 사용한 영상변환 방법은 투영변환(Projection transformation)기법으로 중심투영의 기본원리를 바탕으로 영상왜곡을 최소화할 수 있는 방식이다^[11]. 변환과정은 모델공간좌표계에 대한 수직영상변환과 촬영기선에 대한 수직영상변환, 두 단계로 구분할 수 있다. 모델공간에 대한 수직영상변환의 경우에는 보통 모델좌표계의 기준을 좌측영상의 카메라 좌표계와 일치시키기 때문에 우측영상에 대해서만 이루어지게 되며, 여기에는 추정된 회전행렬 R 의 전치행렬이 적용된다. 그런 다음 두 영상을 촬영기선에 대해 수직영상으로 변환하게 되면 결과적으로 두 영상의 수직시차는 제거되어진다. 이 과정에서 요구되는 회전행렬 R_B 의 회전요소들은 추정된 기하구조로부터 다음과 같이 계산되며,

$$\Omega = \frac{\omega_1 + \omega_2}{2}, \Phi = -\tan^{-1} \frac{T_Z}{(T_X^2 + T_Y^2)}, K = \tan^{-1} \frac{T_Y}{T_X} \quad (9)$$

최종적으로 수직시차 제거를 위한 변환행렬 R_T 는 다음과 같이 계산된다.

$$R_T = R_B R^T \quad (R_B = R_\Omega R_\Phi R_K) \quad (10)$$

그리고 최종 계산된 변환행렬로부터 다음의 투영변환식이 수립되어 그림 2와 같이 영상변환이 이루어지게 된다.

$$\begin{aligned} x_r &= f_r \times \frac{r_{11}x_o + r_{12}y_o + r_{13}f_o}{r_{31}x_o + r_{32}y_o + r_{33}f_o} \\ y_r &= f_r \times \frac{r_{21}x_o + r_{22}y_o + r_{23}f_o}{r_{31}x_o + r_{32}y_o + r_{33}f_o} \end{aligned} \quad (11)$$

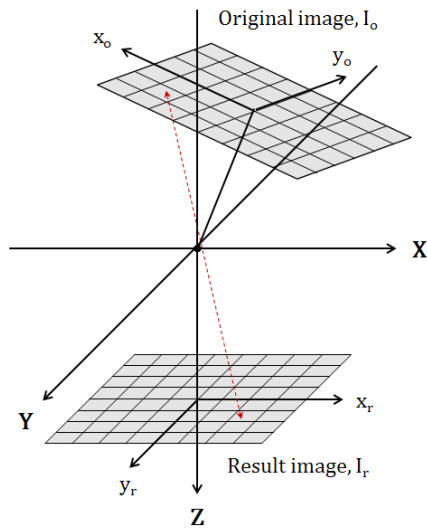


그림 2. 중심투영변환의 원리
Fig. 2. Principle of projection transformation

마지막으로 카메라 움직임 탐지단계에서는 이전 프레임

영상과 비교하여 카메라의 움직임을 자동으로 탐지하게 된다. 이는 카메라의 움직임이 없을 경우에 대응점 추출과 기하추정의 과정을 생략함으로써 처리속도를 향상시키기 위한 목적으로 고안되었다. 카메라의 움직임 여부는 이전 프레임에서 산출된 변환행렬을 이용하여 영상변환을 수행하고 결과영상에 대해 수직시차를 측정해봄으로써 판별될 수 있다. 본 논문에서는 그 판별 기준을 1.5 픽셀로 설정하여, 해당 수치 이상의 수직시차가 발생될 경우 카메라가 움직였다고 판단하였다.

제안방법의 전체적인 작업절차는 다음 그림 3과 같이 정리해볼 수 있다.

IV. 실험 및 결과분석

제안방법의 성능검증은 영상 편위수정부분과 카메라 움직임 탐지부분으로 나뉘어 평가가 이루어졌다. 먼저 편위

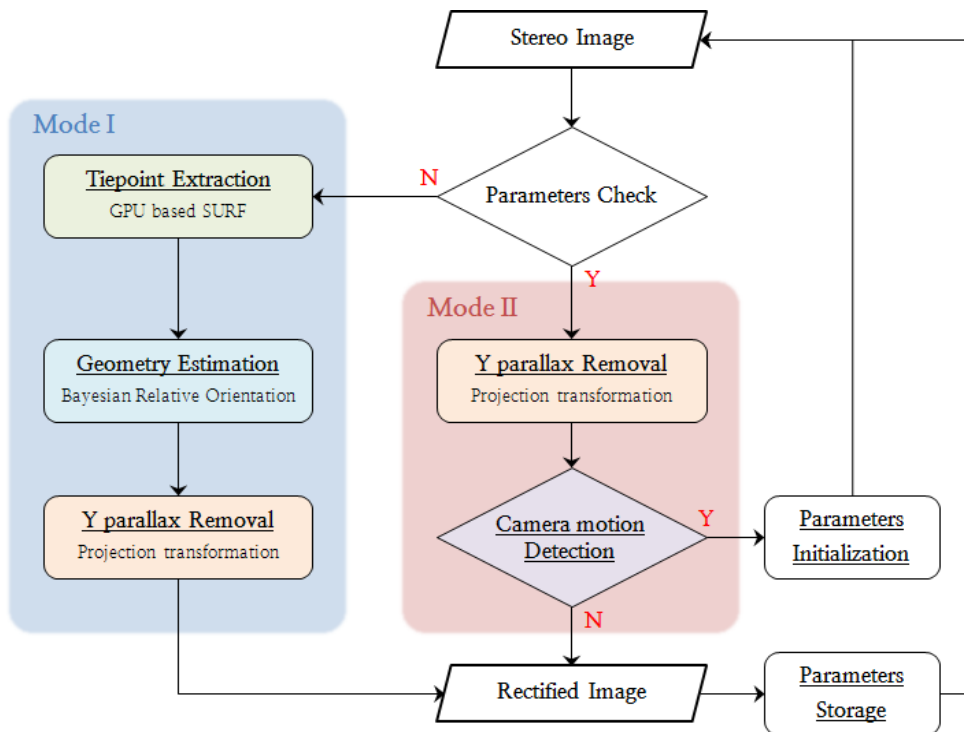


그림 3. 제안방법의 작업흐름도
Fig. 3. Workflow of the proposed method

수정 기법의 성능평가에서는 기하추정의 정확도 및 강인성, 그리고 영상변환에 따른 왜곡도 등이 분석되었다. 실험자료에는 이러한 측면들이 효과적으로 분석될 수 있도록 카메라 틸트 및 기선거리, 영상크기 등을 달리하여 총 3개 데이터셋이 제작되었다. 여기서 영상 시퀀스의 경우는 카메라의 움직임을 고정하여 100개 프레임 영상을 갖도록 촬영하였으며, 성능검증을 위한 대응점들은 수동으로 직접 10점씩 취득하였다(그림 4 참고). 표 1을 보면 Data 1에서 4로 갈수록 오차요인이 증가하는 것을 확인할 수 있다.

표 1. 성능검증용 데이터셋의 상세내용
 Table 1. Dataset details used for evaluation

Dataset	Data 1	Data 2	Data 3
Image size	320×240	320×240	640×480
Baseline	Narrow	Narrow	Wide
Tilt angle	Small	Medium	Large



그림 4. 성능검증에 사용된 영상시퀀스와 평가용 대응점
 Fig. 4. Image sequences and tie-points for evaluation (From top Data 1, 2, and 3)

평가를 위한 성능지표로는 앞서 언급한 바와 같이 기하추정의 정확도 및 강인성과 영상변환에 따른 왜곡도 분석을 위해 총 4가지 지표가 사용되었다. 먼저 정확도 측정에는 결과영상에 잔존하는 수직시차인 E_r 이 사용되었으며, 강인성 측정에는 에피폴라선의 변동을 의미하는 E_o 가 사용되었다. 여기서 E_r 는 편위수정된 검사점들의 위치좌표가 전체 영상 시퀀스에 대해서 가지게 되는 표준편차를 의미하며, 만약 기하추정이 이상적으로 이루어졌을 경우에는 0의 값을 가지게 된다. 즉, E_r 가 작을수록 모든 영상 프레임에 대해서 기하구조가 일관적이고 안정적으로 추정됨을 의미한다. 영상변환에 따른 왜곡도 측정에는 Mallon^[16]이 제안한 결과 영상의 직교성(Orthogonality) E_o 와 상대적인 축척변화 E_a 가 사용되었다. 여기서 E_o 와 E_a 는 각각 결과영상에서 X, Y축이 이루는 각도와 두 대각선 길이의 비를 의미하며, 이상적일 경우 E_o 은 90°의 값을 가지게 되고, E_a 의 경우는 1의 값을 가지게 된다.

성능검증의 신뢰도를 높이기 위하여 기존 방법과의 성능비교를 수행하였다. 여기에는 Sampson 비용함수에 기반한 기하학적 최소화 기법인 Hartley^[8] 방법과 영상왜곡을 최소화하기 위해 개발된 Fusiello^[5] 방법이 사용되었다. 다음 표 2는 방법별 측정된 성능지표 결과를 정리한 것이다.

표 2. 편위수정 성능분석 결과
 Table 2. Result of rectification evaluation

Dataset	Method	E_r (pixel)	E_o	E_a	E_v (pixel)
Data 1 (98 pts)	Hartley	0.96	92.02	1.03	18.72
	Fusiello	0.63	89.84	1.00	9.27
	Proposed	0.38	90.01	1.00	0.57
Data 2 (30 pts)	Hartley	0.87	94.81	1.06	8.19
	Fusiello	1.07	90.61	1.01	16.05
	Proposed	0.57	90.09	1.00	1.72
Data 3 (55 pts)	Hartley	1.56	90.78	1.02	8.27
	Fusiello	1.25	90.29	1.00	15.55
	Proposed	0.64	90.07	1.00	1.36

실험은 Data 1, Data 2, Data 3 각각에 대해서 평균적으로 98, 30, 55개 대응점이 추출되었다. 수직시차 부분에서 오차요인이 증가함에 따라 전반적으로 성능이 저감되는 양상이 나타났으나 제안방법은 서브픽셀의 정확도를 유지

하며 가장 우수한 성능을 보여주었다. 또한 E_o , E_e 로 표현되는 영상왜곡도 측면에서도 제안방법은 가장 최소의 변형만을 나타냈으며, 기하추정의 일관성 E_e 결과에서도 가장 안정적인 모습을 보여주었다. 이러한 결과들은 여러 오차요인들에 대해 강인성이 요구되는 콘텐츠 제작에 있어서 제안방법이 효과적으로 활용될 수 있음을 잘 보여주고 있다.

성능분석의 나머지 한 부분으로 카메라 움직임 탐지기법의 검증은 앞서 실험과는 달리 실제 카메라 움직임이 포함되어 있는 640×480 크기의 영상 시퀀스(245개 프레임 영상)를 제작하여 수행되었다. 참값과의 비교검증을 위해, 사용된 영상 시퀀스에서 실제 카메라 움직임이 발생한 부분들을 직접 육안으로 분석하여 제안방법을 통해 산출된 결과와 비교해보았다. 다음의 그림 5는 실험에 사용된 영상 시퀀스에 대해서 카메라 움직임이 발생한 부분을 나타낸 것이다. 그림에서 빈칸으로 표시된 부분은 카메라의 움직임이 탐지되지 않은 프레임을 의미한다.

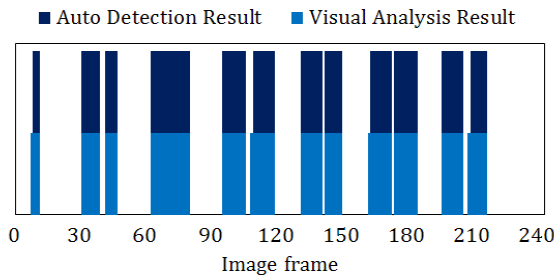


그림 5. 카메라 움직임 탐지결과
Fig. 5. Result of camera motion detection

결과에서 볼 수 있듯이 육안분석결과와 제안방법의 결과는 상당히 유사하게 산출되었다. 육안분석 결과와 비교하였을 때, 몇몇 부분에서 탐지시점이 다소 지체되는 현상이 확인됐지만 정량적으로 약 98.4%의 정확도 산출이 가능한 만큼 실제 현장에서도 충분히 활용 가능할 것으로 판단된다.

위에서 사용된 영상 시퀀스를 이용하여 추가적으로 처리 소요시간을 측정해 보았다. 대응점 추출, 기하구조 추정, 수직시차 제거, 그리고 카메라 움직임 탐지를 각각 단계 I에서 IV로 표기하여 각 단계별 평균소요시간을 표 3에 나타냈으며,

카메라 움직임이 발생한 경우와 발생되지 않은 경우를 각각 그림 3에서와 같이 모드(Mode) I과 II로 표기하여 전체 동작소요시간을 표 4로 정리하였다. 이 때, 실험은 Intel Core i5-2320 CPU, NVIDIA GeForce GTX 550 Ti 그래픽 카드 4.00GB 메모리 성능의 데스크탑 컴퓨터에서 수행되었다.

표 3. 단계별 요구되는 평균 처리시간
Table 3. Average time required for each step

Step	I	II	III	IV
Sec.	0.0584	0.0318	0.0217	0.0592
FPS	17.1312	31.4331	46.0402	16.8797

표 4. 촬영모드별 처리시간
Table 4. Total time required for each camera mode

Camera Mode	Combination	Sec.	FPS
Mode I	I + II + III	0.1119	8.94
Mode II	III + IV	0.0810	12.35

위의 결과는 단계 I에서 추출된 평균 44.8개의 대응점을 기준으로 측정된 것이다. 카메라가 움직이고 있는 경우 평균적으로 초당 8.9개 프레임의 처리가 가능하였으며 고정된 카메라에서는 초당 12.4개의 프레임 처리가 가능하였다. 이 결과는 실시간 처리의 기준을 약 30 fps라 보았을 때, 그 절반에 못 미치는 속도지만 콘텐츠 제작을 위한 결과 모니터링의 경우에는 충분히 활용될 수 있을 것이라 판단된다.

V. 결론 및 향후연구

본 논문에서는 스테레오스코픽 3D 콘텐츠 제작에 활용하기 위한 자동화된 방식의 영상정렬 및 결과 모니터링 방법을 제안하였다. 제안방법에서는 발생 가능한 여러 오차요인들에 대한 강인성 확보를 위해 베이지안 상대표정 알고리즘을 적용하여 기하구조를 추정하였으며, 이를 위한 대응점의 추출은 자동화 및 고속화를 목적으로 GPU 기반의 SURF 알고리즘을 적용하였다. 추정된 기하구조를 바탕

으로 영상의 수직시차를 제거하기 위해 중심투영변환 기법을 적용하였으며, 불필요한 반복과정을 생략하고 카메라 움직임 변화에 즉각 대응하기 위해 자동화된 탐지기법을 구현하였다.

제안방법의 성능분석 결과, 편위수정 부분은 기존 방식들과 비교하였을 때 기하 정확도, 오차 강인성, 영상 왜곡도 측면 모두에서 가장 우수한 성능을 나타내었으며, 카메라 움직임 탐지부분에서도 약 98.4%의 탐지성공률로 신뢰적인 성능을 보여주었다. 또한 처리속도 분석부분에서는 9~12 fps 정도의 처리속도 확보가 가능함에 따라 실제 콘텐츠 제작에 있어 실용적인 저작도구로서 활용 가능함을 확인할 수 있었다. 향후 연구계획으로는 제안방법의 성능향상과 활용성 극대화를 위한 방안으로 FPGA(Field Programmable Gate Array) 기반 하드웨어 시스템을 설계하고 구현할 계획이다.

참 고 문 헌 (References)

[1] Kim, H., Kim, H., Jang, H., Nam, S., Technical Trend of Stereoscopic Content Production, Electronics and Telecommunications Trends, vol. 26, no. 1, pp. 13-24, 2011.
 [2] Zilly, F., Kluger, J., Kauff, P., Production Rules for Stereo Acquisition. Proceedings of the IEEE, vol. 99, no. 4, pp. 590-606, 2011.
 [3] Kang, Y., Ho, Y., An efficient image rectification method for parallel multi-camera arrangement, IEEE transactions on Consumer Electronics, vol. 57, no. 3, pp. 1041-1048, 2011.
 [4] Hartley, R., In defense of the eight-point algorithm. IEEE transactions on pattern analysis and machine intelligence, vol. 19, no. 7, pp. 580 - 593, 1997.
 [5] Hartley, R., Zisserman, A., Multiple View Geometry in Computer

Vision, 2nd ed. Cambridge University Press, New York., 2003.
 [6] Wu, H., Yu, Y., Projective Rectification with Reduced Geometric Distortion for Stereo Vision and Stereoscopic Video. Journal of Intelligent and Robotic Systems, vol. 42, no. 1, pp. 71-94, 2005.
 [7] Banno, A., Ikeuchi, K., Estimation of F Matrix and image rectification by double quaternion, Information Sciences, vol. 183, no. 1, pp. 140-150, 2012.
 [8] Fathy, M. E., Hussein, A. S., Tolba, M. F., Fundamental matrix estimation: A study of error criteria, Pattern Recognition Letters, vol. 32, no. 2, pp. 383-391, 2011.
 [9] Fusiello, A., Irsara, L., Quasi-Euclidean epipolar rectification of uncalibrated images, Machine Vision and Applications, vol. 22, no. 4, pp. 663-670, 2011.
 [10] Kim, J., Kim, H., Lee, T., Kim, T., Photogrammetric approach for precise correction of camera misalignment for 3D image generation. Proceedings of the IEEE International Conference on Consumer Electronics, pp. 396-397, January 13-16, 2012.
 [11] Kim, J., Kim, T., Precise Rectification of Misaligned Stereo Images for 3D Image Generation, Journal of Broadcast Engineering, vol. 17, no. 2, pp. 411-421, 2012.
 [12] Bay, H., Ess, A., Tuytelaars, T., Gool, L. V., SURF: Speeded up robust features, Computer Vision and Image Understanding, vol. 110, no. 3, pp. 346-359, 2008.
 [13] Fischler, M. A., Bolles, R. C., Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, Communications of the ACM, vol. 24, no. 6, pp. 381 - 395, 1981.
 [14] Grodecki, J., Dial, G., Block adjustment of high-resolution satellite images described by rational polynomials, Photogrammetric Engineering & Remote Sensing, vol. 69, no. 1, pp. 59-68, 2003.
 [15] Kim, J., Kim, T., Development of photogrammetric rectification method applying Bayesian approach for high quality 3D contents production, Generation, Journal of Broadcast Engineering, vol. 18, no. 1, pp. 31-42, 2013.
 [16] Mallon, J., Whelan, P., Projective rectification from the fundamental matrix. Image and Vision Computing, vol. 23, no. 7, pp. 643-650, 2005.

저 자 소 개



김 재 인

- 2010년 8월 : 인하대학교 지리정보공학과 학사
- 2013년 2월 : 인하대학교 지리정보공학과 석사
- 2013년 3월 ~ 현재: 인하대학교 지리정보공학과 박사과정
- 주관심분야 : 3D 입체영상 생성, 위성영상 품질분석, 위성영상 활용

저 자 소 개



김 태 정

- 1991년 8월 : 한국과학기술원 전자공학과 학사
- 1992년 10월 : University College London 석사(위성공학)
- 1996년 2월 : University College London 박사(원격탐사)
- 1995년 8월 ~ 2001년 3월 : KAIST 인공위성연구센터 선임연구원
- 2001년 4월 ~ 2003년 8월 : KAIST 인공위성연구센터 연구교수
- 2003년 9월 ~ 현재: 인하대학교 지리정보공학과 정교수
- 주관심분야 : 영상정합, 3D 입체영상 생성, 센서모델링, 자동 DEM 생성, 수치도화기 개발