

## 스테레오 카메라 추적을 통한 3차원 모바일 디스플레이에 증강된 객체의 입체감 향상

박정식 서병국 박종일\*

한양대학교

(nangsik, bkseo)@mr.hanyang.ac.kr, jipark@hanyang.ac.kr

## Consistent Augmentation on 3-D Mobile Display using Stereo Camera Tracking

Park, Jungsik Seo, Byung-Kuk Park, Jong-Il

Hanyang University

## 요약

본 논문에서는 스테레오 기반 카메라 추적 방법을 이용하여 3차원 모바일 디스플레이 상의 증강현실에서 증강된 가상 객체의 3차원 효과가 올바르게 나타나도록 하는 방법을 제안한다. 3차원 디스플레이에 가상 콘텐츠를 증강시키기 위해서는 카메라의 정확한 위치와 자세를 추정하는 것도 중요하지만, 증강된 가상 콘텐츠의 시각적 3차원 효과가 올바르게 나타나도록 하는 것도 중요하다. 이를 위해서는 좌우 카메라의 3차원 공간상의 관계가 유지되도록 양측 카메라의 위치 및 자세를 동시 추정(joint estimation)할 필요가 있다. 제안된 방법은 스테레오 카메라와 페릴렉스 베리어 방식의 3차원 디스플레이가 장착된 안드로이드 스마트폰에서 구현되고 실험되었다.

## 1. 서론

증강현실은 실제 세계에 대한 카메라 영상에 가상의 콘텐츠를 증강시켜 보여줌으로써, 가상의 객체가 실제 세계에 존재하는 것 같은 몰입감과 현실감을 제공하는 기술이다. 보다 실감나는 증강현실을 위해서는 현실 세계와 가상 세계를 정확하게 정합하는 것이 중요한데, 이는 카메라 추적을 통해 해결할 수 있다. 카메라 추적기술은 물리적인 센서 정보 혹은 영상 정보를 통해 3차원 공간상의 카메라 위치와 자세를 계산하는 기술이다. 최근 카메라, 다양한 센서가 탑재된 스마트폰의 보급이 확산되면서 모바일 증강현실 서비스에 대한 관심과 활용 가능성이 높아지고 있다.

비전기반 카메라 추적 기술은 센서기반의 방법과 달리 정확하고 강건하게 카메라 위치 및 자세를 계산할 수 있다. 초기의 비전기반 모바일 증강현실에는 시각적 마커를 이용한 카메라 추적 방법이 사용되었으나, 인위적인 마커를 사용한 부자연스러운 환경은 사용자의 몰입을 방해하는 단점이 있었다[1]. 따라서 자연스러운 환경에서 적용 가능한 특징점기반 추적 방법이 연구되어 왔다. 특징점기반 추적 방법은 추적 대상 환경에 존재하는 시각적 특징점을 이용한 방법으로, SIFT[2]와 FERNS[3]가 그 대표적인 방법이다. 특징점기반 추적 방법은 많은 연산량과 메모리 공간을 필요로 하기 때문에 낮은 성능을 가진 모바일 기기에서도 실시간으로 동작하도록 많은 연구가 진행되어 왔다[4-5].

일반적인 특징점기반 추적 방법은 추적 대상이 주로 평면으로 제한되기 때문에 3차원 환경에서 적용할 수 있는 방법이 연구되어 왔다. Klein 등은 3차원 특징점 지도를 생성하며 카메라 위치와 자세를 계산하는 PTAM을 제안하고, iPhone을 이용하여 구현, 검증하였다[6]. 하

지만, 특징점을 이용한 방법은 정확하고 강건한 카메라 추적을 위해 텍스처가 많은 환경을 필요로 한다. 3차원 모델기반 추적 방법은 추적 환경의 3차원 모델과 영상의 에지나 윤곽선을 이용하기 때문에 텍스처가 적은 환경에서도 추적이 가능한 방법으로서 컴퓨터 비전 및 로봇 분야에서 활발히 연구되어 왔고, 최근에는 모바일 플랫폼에서 구현되어 실시간으로 동작하였다[7-8].

한편, 디스플레이 기술의 발달로 여러 시점의 영상을 이용하여 입체감 있는 영상을 보여주는 3차원 디스플레이가 개발 및 상용화되었다. 최근에는 모바일 기기에도 3차원 디스플레이가 장착됨에 따라, 이러한 모바일 기기를 통해 더욱 실감나는 증강현실 서비스의 제공이 가능하게 되었다. 3차원 디스플레이를 이용하여 증강현실을 구현하기 위해서는 스테레오 혹은 다시점 영상에 대한 카메라의 정확한 위치와 자세를 추정하는 것이 필요하다[9]. 하지만, 실감 있는 3차원 콘텐츠를 제공하기 위해서 증강된 가상 콘텐츠의 시각적 3차원 효과가 올바르게 나타나도록 하는 것도 중요하다.

따라서 본 논문에서는 3차원 모바일 디스플레이를 이용한 증강현실에서 증강된 가상 객체의 입체감이 올바르게 나타나도록 하는 방법을 제안한다. 제안된 방법은 모바일 플랫폼 상에서 실시간 3차원 모델기반 추적방법[8]에 기초하여 두 카메라 간의 기하학적 관계를 통해 카메라 위치 및 자세를 동시 추정하고, 편위수정(rectification) 변환과 투영행렬의 보정을 통해 모바일 3차원 디스플레이에 증강된 가상 콘텐츠의 3차원 효과가 올바르게 나타나도록 한다.

## 2. 방법

\* 교신저자

### 가. 3차원 모델기반 카메라 추적

3차원 모델기반 카메라 추적 방법은 영상에서 검출된 예지와 영상 평면에 투영된 3차원 모델을 정합하여 카메라 위치 및 자세를 추정하는 방법이다. 먼저 이전 프레임의 카메라 위치 및 자세와 3차원 모델의 법선 벡터를 이용한 은면 제거법으로 후면의 선을 제거하고 두 면에 의해 공유되는 선을 제거하여 외곽선을 추출한다. 추출된 선은 영상 평면으로 투영되어 일정 간격으로 샘플링되고, 각 선의 법선 방향으로 가장 가까운 영상 예지를 찾아 정합한다.

카메라 위치 및 자세  $\mathbf{E}$ 는 카메라 움직임  $\Delta\mathbf{E}$ 에 의해 이전 프레임의 카메라 위치 및 자세로부터 갱신된다. 카메라 움직임  $\Delta\mathbf{E}$ 는  $x, y, z$ 축 방향에 대한 평행이동 및 회전 변환의 기저 행렬  $\mathbf{G}$ 와 각 변환 성분의 변화량  $\alpha$ 로 표현된다.

$$\Delta\mathbf{E} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} = \sum_{i=0}^5 e^{\alpha_i \mathbf{G}_i} \approx \mathbf{I} + \sum_{i=0}^5 \alpha_i \mathbf{G}_i \quad (1)$$

카메라 움직임이 작다는 가정 하에  $\Delta\mathbf{E}$ 은 식 (1)과 같이 근사되고, 반복적으로 대응점 간의 거리를 최소화하도록  $\alpha$ 를 최적화함으로써 구해진다.

$$\hat{\alpha} = \arg \min_{\alpha} \left\{ \sum_j^{N_e} \left\| d^j - \sum_{i=0}^5 \alpha_i (\mathbf{n}_i^j \cdot \mathbf{l}_i^j) \right\|^2 \right\} \quad (2)$$

### 나. 스테레오 카메라 추적

스테레오 카메라로 획득한 영상에서 카메라 추적을 각각 수행할 경우 두 카메라 간의 3차원 공간상의 기하관계  $\mathbf{E}_L \mathbf{E}_{L \rightarrow R} = \mathbf{E}_R$ 를 만족하지 않는 카메라 위치 및 자세가 추정될 수 있다. 이런 경우 증강된 두 영상을 3차원 디스플레이에 표시할 때, 3차원 효과가 올바르게 표현되지 않고 잔상처럼 표현된다.

그림 1은 각각 side-by-side 방식과 애너글리프(anaglyph) 방식으로 렌더링한 결과이다. 이 때, 지면상으로는 3차원 영상을 볼 수 없기 때문에 애너글리프를 사용하였다. 왼쪽 카메라의 위치 및 자세는 올바르게 추정되었으나 오른쪽 카메라의 위치 및 자세는 잘못 추정되었고, 좌우 영상이 3차원 영상으로 합쳐졌을 때 가상 객체가 어긋나게 되어 입체감이 올바르게 나타나지 않게 된다.

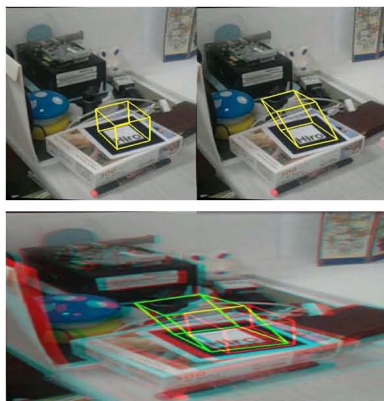


그림 1. 두 카메라 간의 기하학적 관계를 만족하지 않는 경우

이러한 문제는 두 카메라의 위치 및 자세를 동시 추정함으로써 해결할 수 있다. 마커의 경우 간단하게 다음과 같은 과정으로 카메라 위치 및 자세를 동시 추정할 수 있다. 오른쪽 영상에서 검출된 마커의 정점을 두 카메라 간의 기하학적 관계를 이용하여 왼쪽 영상으로 투영하고, 대응하는 왼쪽 영상에서 검출된 마커의 정점과의 중점으로부터 왼쪽 카메라의 위치 및 자세를 계산한다. 오른쪽 카메라의 위치 및 자세는 두 카메라 간의 기하학적 관계를 이용하여 왼쪽 카메라의 위치 및 자세로부터 계산된다. 마지막으로, 카메라 위치 및 자세 행렬에 편위수정 변환 행렬을 곱한 행렬과, 편위수정 후의 카메라 내부 행렬을 이용하여 가상 객체를 영상에 투영함으로써 편위수정된 영상에 올바르게 증강되도록 한다. 이를 위해서는 스테레오 캘리브레이션을 통해 각 카메라의 내부 파라미터와 두 카메라의 3차원 공간상의 관계를 구하는 과정과 편위수정 변환 행렬을 구하는 과정이 선행되어야 한다. 그림 2는 두 카메라의 위치 및 자세를 동시 추정한 결과로서, 양쪽 영상이 겹쳐졌을 때  $x$ 축 방향 시차 이외에는 어긋남이 없는 것을 볼 수 있다.

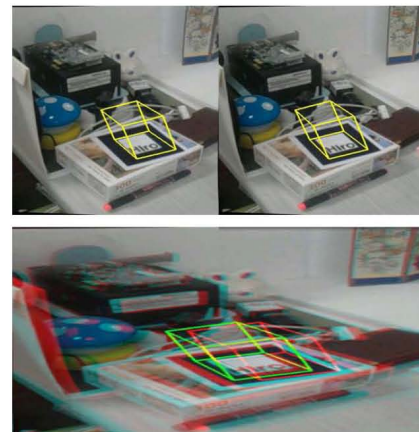


그림 2. 두 카메라 간의 기하학적 관계를 만족하는 경우

3차원 모델기반 카메라 추적에서는 카메라 위치 및 자세가 대응점으로부터 바로 계산되지 않고, 카메라 움직임을 추정하여 이전 프레임의 카메라 위치 및 자세로부터 갱신된다. 따라서 식 (3)과 같이 카메라 움직임을 동시 추정한 뒤 한쪽 카메라 위치 및 자세를 갱신하고, 두 카메라의 기하학적 관계를 이용하여 나머지 카메라의 위치 및 자세를 계산한다.

$$\begin{aligned} \Delta\mathbf{E}_L^* &= \arg \min_{\Delta\mathbf{E}_L} \left\{ \sum_i \|\mathbf{K}_L \mathbf{E}_L \Delta\mathbf{E}_L \mathbf{X}_i - \mathbf{x}_i\|^2 + \sum_j \|\mathbf{K}_R \mathbf{E}_R \Delta\mathbf{E}_L \mathbf{E}_{L \rightarrow R} \mathbf{X}_j - \mathbf{x}_j\|^2 \right\} \\ &= \arg \min_{\Delta\mathbf{E}_L} \left\{ \sum_i \|\mathbf{K}_L \mathbf{E}_L \Delta\mathbf{E}_L \mathbf{X}_i - \mathbf{x}_i\|^2 + \sum_j \|\mathbf{K}_R \mathbf{E}_R \Delta\mathbf{E}_L \mathbf{E}_{L \rightarrow R} \mathbf{X}_j - \mathbf{x}_j\|^2 \right\} \quad (3) \end{aligned}$$

여기서,  $\mathbf{K}$ 는 카메라 내부 파라미터를, 첨자 L과 R은 각각 왼쪽 카메라와 오른쪽 카메라를 나타내고  $\mathbf{E}_{L \rightarrow R}$ 은 왼쪽 카메라로부터 오른쪽 카메라로의 3차원 공간상의 변환 관계를 나타낸다. 마지막으로, 영상의 편위수정 변환 행렬을 카메라 행렬에 곱해준다.

### 3. 결과

실험에 사용된 모바일 기기는 LG Optimus 3D로서 ARM Cortex

A9 듀얼코어 1.2 GHz CPU, 512 MB RAM, 그리고 스테레오 카메라와 패럴랙스 배리어 방식의 무안경식 3차원 디스플레이가 탑재되었다.

Java의 느린 속도를 고려하여 카메라 입력 및 사용자 인터페이스 이외의 처리는 C/C++로 구현하고 JNI(Java Native Interface)를 통해 호출하였다. 3차원 디스플레이의 구동에는 LG Real3D SDK를 이용하였다. 카메라 입력은 좌우 카메라의 영상이 side-by-side 방식으로 결합된 720×480 YUV420 영상으로 추적 처리에는 Y 채널만 추출하여 사용하였다. 이 때, 카메라 내부 파라미터 및 두 카메라의 기하학적 관계가 변하지 않도록 초점거리를 고정시키고 auto convergence 옵션을 비활성화 시켰다.

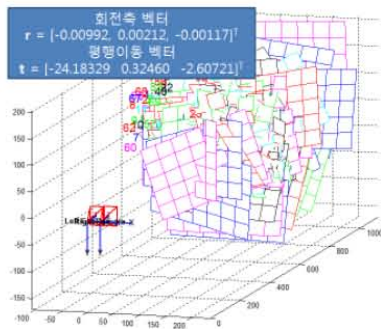


그림 3. Optimus 3D의 스테레오 카메라 캘리브레이션 결과

스테레오 카메라 캘리브레이션은 오프라인에서 수행하여 카메라 내부 파라미터, 왜곡 보정 파라미터, 두 카메라 간의 관계, 편위수정 정보 및 편위수정 후의 카메라 파라미터를 얻었다. 그림 3은 두 카메라의 관계를 보여준다. 두 카메라의 x축 방향 평행이동 성분은 24.18329 mm로, Optimus 3D의 스테레오 카메라의 실제 카메라간 거리 24 mm에 근접함을 볼 수 있다.

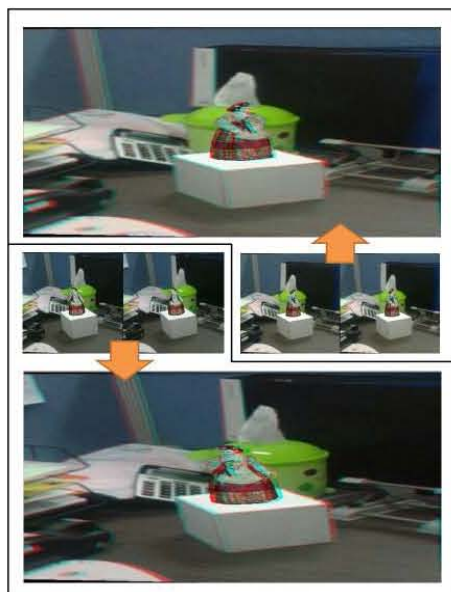


그림 4. 정육면체 모델에 대한 스테레오 추적 및 증강 결과

그림 4는 정육면체 모델에 추적 및 증강 결과를 나타낸다. 3차원 영상은 패럴랙스 배리어 모드로 렌더링하는 대신 애너글리프로 렌더링하였다. 스테레오 카메라의 양 카메라의 관계를 이용하여 카메라 위치 및 자세를 동시 추정하고 영상을 편위수정한 뒤, 편위수정 후의 카메라 파라미터를 이용하여 각 카메라 위치 및 자세를 보정한 결과이다. 애너글리프 영상을 통해 증강된 가상 객체의 3차원 효과가 올바르게 표현되는 것을 알 수 있다.

#### 4. 결론

본 논문에서는 스테레오 기반 카메라 추적 방법을 이용하여 3차원 모바일 디스플레이 상의 증강현실에서 증강된 가상 객체의 3차원 효과가 올바르게 나타나도록 하는 방법을 제안한다. 3차원 모델기반 추적 방법에 기초하여 두 카메라 간의 기하학적 관계를 유지하도록 카메라 위치 및 자세를 동시 추정하고, 영상의 편위수정에 따른 카메라 파라미터 변화를 보정한 결과를 증강현실에 반영하였다. 제안된 방법은 스테레오 카메라와 패럴랙스 배리어 방식의 3차원 디스플레이가 탑재된 스마트폰에서 실험되었다. 현재 모바일 플랫폼에서의 보다 정확하고 강력한 3차원 스테레오 카메라 추적 방법에 대한 연구를 진행 중에 있다.

#### 참고문헌

- [1] D. Wagner and D. Schmalstieg, "ARToolkitPlus for pose tracking on mobile devices," in *Computer Vision Winter Workshop*, 2007, pp. 139-146.
- [2] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.
- [3] M. Ozuysal, M. Calonder, V. Lepetit, and P. Fua, "Fast keypoint recognition using random ferns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, issues 3, pp. 448-461, 2010.
- [4] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, and D. Schmalstieg, "Pose tracking from natural features on mobile phones," in *ISMAR*, 2008, pp. 125-134.
- [5] S. Taylor and T. Drummond, "Multiple target localisation at over 100 FPS," in *BMVC*, 2009.
- [6] G. Klein and D. Murray, "Parallel tracking and mapping on a camera phone," in *ISMAR*, 2009, pp. 83-86.
- [7] J. P. Lima, V. Teichrieb, J. Kelner, and R. W. Lindeman, "Standalone edge-based markerless tracking of fully 3-dimensional objects for handheld augmented reality," in *VRST*, 2009.
- [8] B.-K. Seo, J. Park, and J.-I. Park, "3-D visual tracking for mobile augmented reality applications," in *ICME*, 2011, pp. 293-298.
- [9] F. Dionnet and E. Marchand, "Stereo tracking and servoing for space applications," *Advanced Robotics*, vol. 23, no. 5, pp. 579-599, 2009.