

# 히스토리 버퍼를 사용하여 떨림 현상을 줄이는 마커 추적

윤중현<sup>0</sup> 이범중, 박종승

인천대학교 컴퓨터공학과

{jhyoon<sup>0</sup>, leeyanga, jong}@incheon.ac.kr

## Stabilizing Camera Poses in Marker Tracking Using History Buffer

Jong-Hyun Yoon<sup>0</sup> Bum-Jong Lee Jong-Seung Park

Department of Computer Science & Engineering, University of Incheon

### 요 약

본 논문에서는 특정 마커를 사용하는 실감형 증강현실 시스템 상에서 카메라가 비정형적인 움직임을 하는 경우에 대하여, 다중 마커를 사용한 떨림 현상을 줄인 실시간 움직임 추적 기법을 제안하고자 한다. 카메라의 움직임을 추정하기 위하여 카메라와 마커 사이의 변환을 계산해야 한다. 이미지로부터 검출된 각 마커의 네 모서리 점들을 이용하여, 각 마커에 대한 변환을 계산한다. 마커는 서로 다른 로컬 좌표계를 가지고 있고, 마커에 대한 변환은 해당 마커의 좌표계에 의해 정의된다. 다중 마커의 로컬 좌표계로부터 최적의 카메라 움직임을 추정하기 위한 정합 알고리즘을 제안한다. 정합을 위한 방법으로 레퍼런스 마커를 사용한다. 레퍼런스 마커는 정합 과정에서 자동적으로 선택된다. 레퍼런스 마커를 기준으로 각 마커의 변환에 대해 신뢰성(confidence rate)을 기반으로 가중치를 적용함으로써 최적의 카메라 움직임을 추정할 수 있다. 또한 추정된 카메라의 움직임의 최적화를 위하여 히스토리 버퍼를 사용하여 떨림 현상을 제거하는 방법을 제안한다. 추정된 카메라의 위치에 대한 평균 필터 및 중간 필터의 개념과 유사한 보정 방법을 통해 떨림 현상을 제거한다. 실험을 통해 다른 방법들과 비교한 우리가 제안한 방법의 정확성을 확인할 수 있다.

### 1. 서 론

최근 실감형 증강현실분야에 대한 연구가 활발히 진행되고 있다. 증강 현실이란 실세계와 가상객체의 융합에 대한 연구 분야이다. 증강현실 시스템에서는 실세계뿐만 아니라 가상으로 생성된 이미지가 투영된 증강 환경의 모습을 볼 수 있다. 또한 게임 플레이어의 위치를 추적함으로써 증강 환경에서 플레이어와 가상의 객체 사이의 대화형 인터페이스 구성이 가능해진다. 사용자는 HMD나 FMD와 같은 디스플레이 장치를 사용하여 공간상에서 자유로운 움직임을 할 수 있고, 카메라를 부착하여 카메라로부터 들어오는 입력 이미지를 통해 실세계를 포함한 증강 환경을 경험할 수 있다. 위치 추적을 위한 여러 가지 방법들이 있으나 본 논문에서는 멀티플마커를 사용한 카메라의 움직임의 추적을 최적화하는 방법을 제안한다.

마커를 사용한 카메라 자세 추정 방법은 빠르고 안정적이다. 최근 마커 기반의 카메라 추적기법에 대한 다양한 연구가 진행되었다. Piekarski는 혼합현실(Mixed Reality) 기법에 대한 연구를 하였고, 이를 모바일 어플리케이션에 대해 적용하였다[1]. Siltanen은 증강현실기법을 사용하여 내부 인터리어를 할 수 있는 응용시스템을 개발하였다[2]. 이들 모두 마커 검출의 정확성을 높이기 위한 연구를 중점적으로 진행하였다. 또한 카메라의 위치 추정 오차를 줄이기 위한 많은 연구들이 진행되었다[3][4]. 대표적인 예로 사전에 카메라 캘리브레이션 과정을 수행함으로써 카메라에 의한 하드웨어적인 오차를 줄이

기 위한 연구가 많이 진행되고 있다[5]. 카메라 캘리브레이션을 할 때 왜곡 요소(distortion factor)를 사용함으로써, 카메라 위치 복원의 정확도가 높일 수 있다. 효과적인 카메라 캘리브레이션을 위한 연구 외에도, 카메라 캘리브레이션 과정 없이 카메라의 움직임을 복원하는 방법에 대한 연구도 활발히 진행되고 있다[6][7].

본 논문에서는 마커를 사용해 카메라의 움직임을 추정하고 추정된 결과의 보정을 통해 떨림 현상을 줄이는 방법을 제안한다. 이미지로부터 검출된 다수의 마커들에 대해서도 효과적인 융합을 통해 좀 더 정확한 플레이어 움직임의 복원이 가능하다. 또한 추정된 카메라의 움직임의 보정을 통해 떨림 현상을 줄인 좀 더 최적화된 움직임 추정 결과를 얻을 수 있다.

2장에서는 다중 마커를 사용한 본 논문에서 제안하는 카메라 움직임 추정 방법에 대하여 소개한다. 3장에서는 추정된 카메라의 움직임에 대하여 떨림 현상을 보정하는 방법에 대하여 설명한다. 4장에서는 실제 비디오 프레임에 대하여 우리의 방법을 사용한 실험 결과를 분석하고, 마지막으로 5장에서는 결론을 내린다.

### 2. 마커를 통한 카메라 움직임 추정

다중 마커를 사용한 카메라의 움직임 추정 기법을 제안한다. 제안한 알고리즘은 다음과 같다.

- Step 1. 이미지상의 N 개의 마커 검출
- Step 2. 검출된 i 번째 마커의 변환  $T_i$  계산
- Step 3. 이미지상의 마커의 센터와 이미지 센터와의 거리가 최소값을 갖는 마커를 레퍼런스 마커로 지정
- Step 4. 검출된 마커의 모든 변환들을 레퍼런스 마커의 좌표계로 변환.  $T_i \rightarrow T^i$
- Step 5. 검출된 모든 마커에 대한 오류율(error rate) 계산
- Step 6. 추정된 마커 변환에 대해 가중치 합(weighted sum) 계산.  $\rho$  : error rate,  $\sum(\rho T_i)$
- Step 7. 히스토리 버퍼를 사용한 추정된 변환의 보정

검출된 마커 중에서 레퍼런스 마커를 선택한다. 이미지에서 마커의 중심과 이미지 중심의 거리가 가장 작은 값을 갖는 마커를 레퍼런스 마커로 지정한다. 가운데 위치한 마커는 마커 검출에 대하여 가장 큰 정확도를 갖기 때문에 이미지 센터로부터 가장 가까운 마커를 레퍼런스 마커로 사용한다. 레퍼런스 마커에 가장 큰 가중치를 적용함으로써 카메라 움직임의 추정에 대한 신뢰성을 높일 수 있다.

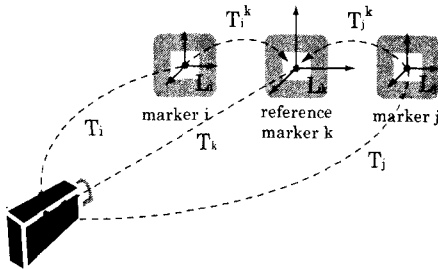


그림 1. 다중마커의 기하학적 관계

그림 1은 마커의 로컬 좌표계 사이의 관계를 나타내고 있다. 추정된 카메라의 움직임은 레퍼런스 마커의 오리엔테이션을 기준으로 해야 한다. 검출된 각 마커는 서로 다른 로컬 좌표계를 갖는다. 그러므로 각 로컬 좌표계를 레퍼런스 마커를 기준으로 하는 좌표계로 정합하는 과정이 필요하다. 정합과정에서 실제세계에서의 두 마커 사이의 기하적인 관계를 이미 알고 있다고 가정한다. 정합을 통해서 검출된 마커들의 모든 변환은 레퍼런스 마커의 오리엔테이션을 기준으로 식 (1)과 같이 정렬된다.

$$T_i^j = T_i(T_{i,k})^{-1} = T_i T_i^{-1} T_k = T_k \quad (1)$$

멀티플 마커의 정확한 가중치 합을 위하여 오차율(error rate)을 계산한다. 오차율은 타겟 마커에 대한 카메라 시선 방향과의 관계에 대한 값을 의미한다. 우리는 입력 이미지에서 마커의 코너 점을 사용하여 오차율을 정의한다.

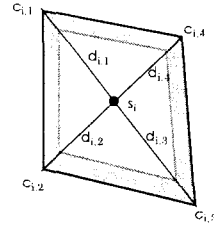


그림 2. 이미지 상에서 검출된 마커

그림 2는 이미지 상에서 검출된 마커를 보여준다.  $c_{i,1}, c_{i,2}, c_{i,3}, c_{i,4}$ 는 마커의 코너 포인트를 나타내며,  $s_i$ 는 두 대각 직선의 교차점을 나타낸다. 마커의 각 모서리 점으로부터 식 (2)와 같이 유도할 수 있다.

$$v_i = \frac{1}{4} \sum_{n=1}^4 (d_{i,n} - \bar{d}_i)^2 \quad (2)$$

$$\text{where } \bar{d}_i = \frac{1}{4} \sum_{n=1}^4 d_{i,n}^2 \text{ and } d_{i,k} = \|s_i - c_{i,k}\|$$

마커는 정사각형 모양을 가지고 있다.  $v_i$ 는 이미지로부터 검출된 마커의 모양이 정사각형 모양에서 얼마나 왜곡되었는지를 의미한다. 왜곡이 되지 않을수록  $v_i$ 는 0에 가까운 값을 갖게 된다.

또한 오차율 계산을 위해 마커가 이미지에서 차지하고 있는 영역을 고려한다. i 번째 마커의 이미지영역  $a_i$ 는 검출된 마커의 안쪽 픽셀 수로써 계산된다.  $v_i$ 와  $a_i$ 를 통해 계산된 오차율  $\rho$ 의 계산은 식 (3)과 같다.

$$\rho_i = w_i / \sum_{i=1}^N w_i, \quad (3)$$

$$\text{where } w_i = a_i / v_i, a_i = p_i / \sum_{i=1}^N p_i, \text{ and } v_i = v_i / \sum_{i=1}^N v_i$$

$w_i$ 는 오차율  $\rho$ 를 계산하기 위한 가중치이다. 계산된 오차율  $\rho$ 는 0에서 1사이의 값을 갖고, 1에 가까운 값을 가질수록 카메라가 마커를 정면에서 보고 있음을 의미한다. 계산된 오차율  $\rho$ 가 0에 가까운 값을 갖는다면 카메라는 마커를 정면에서 보고 있는 경우이다. 이런 경우 오차가 거의 없는 마커 검출 결과를 얻을 수 있다.  $\rho$  값은 카메라의 움직임 추정값을 개선하는데 사용된다. 다중 마커는 단일 마커보다 다양한 정보를 제공한다. 다중 마커가 제공하는 정보를 융합하는 과정은 식 (4)와 같다.

$$T = \sum_{i=1}^N \rho_i T_i \text{ where } \sum_{i=1}^N \rho_i = 1 \quad (4)$$

추정된 모든 카메라 변환들은 계산된 오차율에 의해서 새로운 값을 갖게 된다. 계산된  $T$ 은 싱글 마커나 일반적인 멀티플 마커의 방법보다 신뢰할 수 있는 카메라 움직임 정보를 제공한다.

### 3. 추정된 카메라 움직임 정보의 떨림 현상 보정

이미지에서 추정된 카메라의 움직임 변환 추정은 여러 환경 변수의 영향을 받는다. 가장 민감한 환경변수로 실

험환경에서의 빛을 들 수 있다. 다양한 환경 변수의 변화에 의한 영향으로 마커의 변환 추정은 안정성을 보장할 수 없다. 이번 장에서는 추정된 카메라 움직임의 떨림 현상 제거를 위해 히스토리 버퍼를 사용한 카메라 움직임의 보정 방법을 설명한다.

$n$  번째 카메라의 움직임을  $\theta(t_x^n, t_y^n, t_z^n, r_x^n, r_y^n, r_z^n)$ 라 할 때,  $\theta(t_x^n, t_y^n, t_z^n)$ 는 이동벡터를,  $\theta(r_x^n, r_y^n, r_z^n)$ 는 카메라의 방향 벡터를 의미한다. 떨림 현상 보정을 위하여 히스토리 버퍼를 사용한다. 히스토리 버퍼에 들어있는 이전 세 개의 움직임을 함께 고려하였을 경우

$$\theta(t_x^{n-3}, t_y^{n-3}, t_z^{n-3}) \theta(t_x^{n-2}, t_y^{n-2}, t_z^{n-2}) \theta(t_x^{n-1}, t_y^{n-1}, t_z^{n-1}) \theta(t_x^n, t_y^n, t_z^n) \quad (5)$$

식 (5)와 같은 인자를 고려할 수 있다.  $n$  번째 카메라의 움직임에 대하여  $n-3$  번째에서  $n-1$  번째 까지 히스토리를 함께 고려한다. 현재 카메라 위치의  $x$  좌표를 계산할 때,  $(t_x^{n-3}, t_x^{n-2}, t_x^{n-1}, t_x^n)$  중  $t_x^{n-3}$ 을 최소값,  $t_x^n$ 을 최대값이라 하면,  $n$  번째 마커의 이동벡터의  $x$  좌표는  $t_x^n = (t_x^{n-2} + t_x^{n-1})/2$ 의 값을 갖게 된다. 최소값과 최대값을 제외한 나머지 요소들의 평균값을 취함으로써, 떨림 현상의 원인이 되는 갑작스런 증가나 감소를 한 값을 배제한 값을 가질 수 있다. 이동 벡터의  $y, z$ 도 같은 과정을 통해 좌표값을 계산한다. 그림 3에서  $y$  좌표에 대한 보정의 한 예를 볼 수 있다. 7 번째에서 갑자기 증가된 값이 있고, 15 번째에서 갑자기 감소된 값이 나오는 것을 볼 수 있다. 이때, 그림에서 맨 아래쪽의 보정된 결과를 보면 심한 떨림이 있을 수 있는 값에 대하여 이전 값과 비교하였을 때 적절한 값으로 보정되는 결과를 확인할 수 있다.

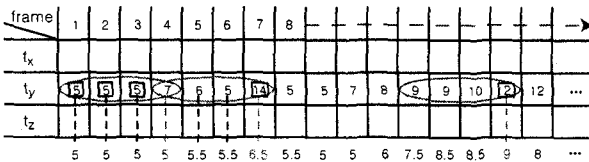


그림 3. 히스토리 버퍼를 사용한 좌표 보정

위와 같은 과정을 반복적 수행을 함으로써 좀 더 최적화된 결과를 얻을 수 있다. 하지만 마커를 사용한 카메라 추적기법은 실시간 시스템에 응용되는 경우가 많으므로 단일 수행을 통해 추정 결과를 보정하는 것이 시간적 측면을 고려하였을 때 올바른 것으로 보인다.

4. 실험 결과

다수의 실 데이터 셋을 사용하여 카메라의 움직임을 추정하는 실험을 하였다. 본 실험에서는 제안한 방법의 정확도 측정을 위하여 각 이미지마다 카메라의 움직임을 추정하고 미리 정의한 카메라의 움직임과 수치적으로 비교하였다.

그림 4와 그림 5는 실험에 사용된 연속되는 이미지 프레임들을 보여준다. 실험에서 사용된 마커는 ARToolkit의

마커와 ARTag의 마커를 혼합하여 사용하였다. 모든 마커들은 5cm의 길이의 정사각형 모양을 갖도록 하였다. 다중 마커는 5개의 단일 마커를 수평하게 정렬하였다. 그림 4는 카메라가 직선의 움직임을 갖는 실험에 사용된 이미지 프레임들을 나타낸다. 이때 카메라의 위치는 마커로부터 70cm 정도 떨어져 있었으며, 카메라의 시선 방향은 마커를 기준으로 정 면을 바라보도록 하였다. 그림 5는 카메라가 원형의 움직임을 갖는 실험에 사용된 이미지 프레임이다. 카메라는 50cm 정도의 반경을 갖는 원 테 위에서 움직였으며, 약 5도 정도의 간격을 갖고 움직이도록 하였다.

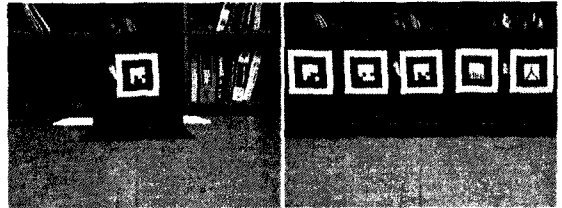


그림 4. 직선 움직임을 갖는 카메라에 대한 실험 환경



그림 5. 원형 움직임을 갖는 카메라에 대한 실험 환경

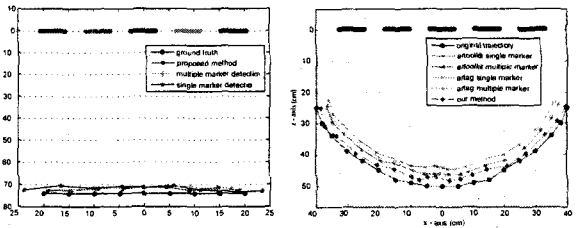


그림 6. 카메라의 움직임 복원

그림 6은 제안된 방법과 기존의 단일 마커, 다중 마커를 사용한 카메라의 움직임을 복원한 결과를 나타내고 있다. 모든 실험에서 카메라의 실제 움직임과 비교를 하였다. 마커의 위치를 5개의 굵은 직선으로 표시하였고, 복원된 카메라의 위치를 표시하고, 궤적을 선으로 연결하였다. 왼쪽 그림은 9 프레임에 대하여 직선 움직임을 갖는 카메라에 대한 실험 결과이고, 오른쪽 그림은 19 프레임에 대하여 원형 움직임을 갖는 카메라에 대한 실험 결과이다. 제안된 방법이 실제와 가장 근접한 결과를 얻는 것을 확인할 수 있다.

그림 7은 그림 6의 결과의 오차를 분석한 결과이다. 왼쪽 두 개의 그래프가 직선 움직임을 갖는 카메라의 실

험결과에 나타나고, 오른쪽 두 개의 그래프가 원형 움직임을 갖는 카메라의 실험 결과를 나타낸다. 모든 오차는 실제 측정된 카메라의 위치와 비교하였다. 그래프는 위에서부터 카메라 움직임 간격 오차, 카메라 위치의 x좌표 오차를 계산한 것이다. 카메라의 직선 움직임 실험에 대하여 단일 마커의 경우는 2cm 정도의 오차, 기존 다중 마커의 경우 1cm 정도의 오차, 제안된 방법의 경우 0.5cm 정도의 오차의 갖는 결과를 확인할 수 있다. 카메라의 원형 움직임 실험에 대하여 단일 마커의 경우는 3.3cm 정도의 오차, 기존 다중 마커의 경우 2.7cm 정도의 오차, 제안된 방법의 경우 1.8cm 정도의 오차의 갖는 결과를 확인할 수 있다. 그래프를 통한 오차를 확인한 결과도 제안된 방법이 가장 적은 오차를 갖는 것을 알 수 있다.

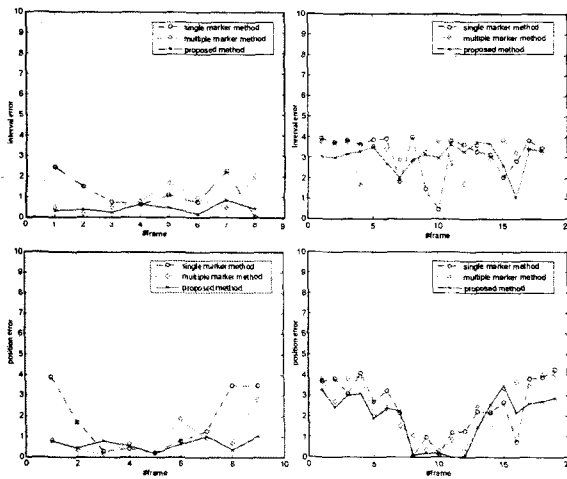


그림 7. 계산된 카메라 움직임의 추정 오차

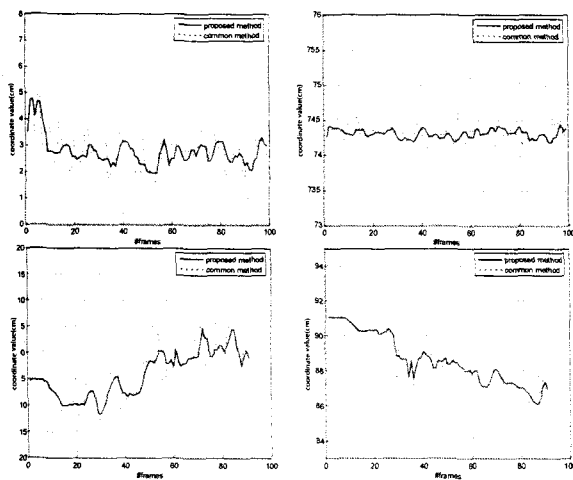


그림 8. 떨림 현상을 제거한 카메라좌표와 일반적으로 추정된 카메라좌표의 떨림 현상 비교

그림 8은 떨림 현상을 제거한 카메라의 움직임과 떨림 현상을 제거하지 않은 카메라의 움직임을 비교한 그래프이다. 위쪽 두 개의 그래프는 움직이지 않고 고정되어 있는 카메라에 대하여 약 100프레임에 걸쳐 추정된 카메라 위치의 좌표를 비교한 것이고 아래 두 개의 그래프는 직선 움직임을 하는 카메라의 위치에 대한 좌표를 비교한 것이다. 위쪽과 아래쪽 모두 왼쪽의 그래프가 x좌표의 보정된 결과를 비교한 것이고, 오른쪽 그래프가 z좌표를 비교한 보정 결과이다. 붉은 점선으로 표시되어 있는 것이 떨림 현상을 제거하지 않은 것으로써 파란색으로 표시된 떨림 현상을 제거한 것보다 편차가 심한 결과를 확인할 수 있다.

5. 결론

본 논문에서는 다중 마커를 사용하여 넓은 공간에서 실감형 증강현실 시스템에서 카메라의 움직임을 추정하고 떨림 현상을 보정할 수 있는 새로운 방법을 제안하였다. 우리가 제안한 알고리즘은 임의의 마커의 개수에도 적용할 수 있는 다중 마커 방법을 사용함으로써 이전 방법과 차별성을 갖고 있다. 하나의 마커가 검출 되었을 경우에도 본 알고리즘을 적용해서 카메라의 움직임을 추정할 수 있다. 이미지 프레임에서 마커를 검출하고 각 이미지 마다 레퍼런스 마커를 결정하였다. 그리고 각 마커의 로컬 좌표계를 레퍼런스 마커를 기준으로 하는 마커계로 변환함으로써 최적화된 카메라의 움직임을 계산할 수 있었다. 또한 빛과 같은 환경적 영향 때문에 불안정한 결과가 추정되는 경우에 대하여 보정을 통해서 떨림 현상을 줄일 수 있었다. 실험 결과를 통해서 기존의 방법과 비교하여 제안한 방법이 정확한 결과를 확인할 수 있었다.

본 논문에서 제안한 기법은 대화형 증강현실 게임 시스템의 구현에 상당히 유용할 것이다. 가상현실 분야에서도 네비게이션 기능을 제공함으로써 교육용 시뮬레이터 개발에도 활용될 수 있을 것으로 보인다.

향후 마커 사이의 기하적인 관계를 알지 못하는 경우에 유효한 카메라의 움직임을 추정하고 카메라의 떨림 현상을 좀 더 정확한 보정 방법에 대해 연구할 계획이다.

참고문헌

- [1] Piekarski, W., Avery, B., Thomas, B.H., "Hybrid indoor and outdoor tracking for mobile 3d mixed reality," I SMAR 2003, pp. 266-267, 2003.
- [2] Siltanen, S., Woodward, C., "Augmented interiors with digital camera images," Australasian User Interface Conference, Vol. 50, pp. 33-36, 2006.
- [3] Abdullah, J., Martinez, K., "Camera self-calibration for the ARToolkit," Proceedings of First International Augmented Reality Toolkit Workshop, pp. 84-88, 2002.

- [4] Malbezin, P., Piekarski, W., Thomas, B., "Measuring ARToolkit accuracy in long distance tracking experiments," First International AR Toolkit Workshop, 2002.
- [5] eddine Ababsa, F., Mallem, M., "Robust camera pose estimation using 2d fiducials tracking for real-time augmented reality systems," VRCAI 2004, pp. 431-435, 2004.
- [6] Bianchi, G., Wengert, C., Harders, M., Cattin, P., Szekely, G., "Camera-marker alignment framework and comparison with hand-eye calibration for augmented reality applications," ISMAR 2005, pp. 188-189, 2005.
- [7] Micilotta, A.S., Ong, E.J., Bowden, R., "Real-time upper body 3d pose estimation from a single uncalibrated camera," Eurographics 2005, pp. 41-44, 2005.