

모델기반 카메라 추적에서의 모델오차의 영향

이은주 김강수 서병국 *박종일

한양대학교

{heckism, vistavision, nwseoweb}@mr.hanyang.ac.kr *jipark@hanyang.ac.kr

Influence of Model Errors in Model-based Camera Tracking

Eun Joo Rhee Kangsoo Kim Byung-Kuk Seo *Jong-Il Park

Hanyang University

요약

본 논문에서는 모델 기반 카메라 추적 시 필요한, 사전 정의된 실측 모델의 정확성이 카메라 추적의 정확성에 미치는 영향에 대하여 논의한다. 이를 위하여 모델 기반과 특징점 기반의 카메라 추적에 가중치를 두고 혼합시켜 이용하는 하이브리드 카메라 추적 방법을 사용하고, 이 방법을 활용한 상호작용형 모델링(interactive modeling)을 이용하여 실측 모델을 제작한다. 또한, 상호작용형 모델링 과정에서 생기는 실측 모델의 오차로 인해 발생하는 카메라 추적 오차를 Cramer-Rao 하한(lower bound)을 이용하여 정의하고, 이 둘의 상관관계를 실험적으로 도출한다. 이를 통해 사전 정의된 실측 모델이 가질 수 있는 오차의 하한을 실험적으로 검증한다.

1. 서론

카메라 추적 기술은 영상으로부터 획득한 정보를 기반으로 카메라의 자세를 추정하는 기술로서, 로봇 비전(robot vision), HCI(human-computer interaction), 증강현실(augmented reality), 영상 감시 체계(visual surveillance) 등과 같은 다양한 컴퓨터 비전 응용 분야에 있어서 핵심 기술 중 하나이다. 특히, 증강현실에서는 가상 객체를 현실 공간에 정확하게 증강시키기 위하여 카메라와 대상물 간의 기하학적인 관계를 취득하는데, 이 때, 정교한 카메라 추적 방법이 필수적으로 요구된다. 정교한 카메라 추적을 위해서는 영상으로부터 지속적으로 추출이 가능한 대상물이 필요하며, 이를 위해 ARToolkit[1]과 ARTag[2] 등은 사전 정의된 형태의 마커를 이용한 카메라 추적을 제안했다. 마커를 이용한 카메라 추적은 제작이 간편하고 구현이 빠르다는 장점이 있지만, 굵은 정사각형 테두리를 사용하는 마커 형태는 주변 환경에 어울리지 못하고 사용자의 시야를 방해하는 문제점이 있다. 최근에는 이러한 마커 형태의 문제점을 극복하고자 자연 영상 내에서 특징점을 추출하거나 사전 제작된 3차원 모델과의 정합을 통해 카메라의 자세를 추정하는 다양한 방법들이 제안되고 있다. 특징점 기반 카메라 추적은 자연 영상에서 영상 강도의 변화에 따라 두드러진 특징점을 검출하고 이를 이용하는 방법으로, SIFT[3]나 SURF[4]와 같은 지역 서술자(local descriptor)로 특징점을 정의해 추적하거나 FERNS[5]와 같이 확률 모델을 이용한 패턴 분류 기법으로 추적한다. 이러한 방법들

은 마커를 이용하지 않기 때문에 사용자의 시야를 방해하지 않지만, 변화량이 다양하지 않은 영상에 대해서는 특징점을 추출 및 정의하기가 어려운 단점이 있다. 모델 기반 카메라 추적[6]은 사전 정의된 실측 모델을 기준 정보로 활용하는 방법으로, 카메라 입력 영상의 윤곽(edge)과 실측 모델의 윤곽을 정합시켜 카메라 자세를 추정한다. 그러나 이 방법은 사전 정의된 실측 모델의 윤곽과 입력 영상에서 검출한 대상물의 윤곽을 정합시키기 위하여 카메라의 초기 자세가 주어져야 하며, 영상의 변화량이 심해 윤곽 검출이 용이하지 않을 경우 활용에 어려움이 있다. [7]에서는 특징점 기반과 모델 기반 카메라 추적에 가중치를 두고 혼합시킴으로써 상황에 따라 적합한 영상 정보를 이용하는 하이브리드 카메라 추적 방법이 제안되었다. 이 방법은 다양한 시점으로부터 추출 및 3차원 복원된 특징점 정보를 이용하는 동시에, 사전 정의된 3차원 실측 모델을 이용하여, 3차원 물체 추적에 최적화 시킨 방법이다.

본 논문에서는 사전 정의된 실측 모델의 정확성이 카메라 추적의 정확성에 미치는 영향을 살펴보기 위하여 앞서 언급한 하이브리드 카메라 추적 방법을 이용하였으며, 이를 활용한 상호작용형 모델링 방식(interactive modeling)을 소개한다. 상호작용형 모델링을 이용하면 실측 모델을 보다 유연하게 제작할 수 있다. 하지만 상호작용형 모델링을 통해 제작된 실측 모델은 레이저 측정 후 제작된 모델에 비해 상대적으로 그 정확도가 떨어지며, 모델링 과정에서 생기는 실측 모델의 오차는 카메라 추적의 정확성에 큰 영향을 미칠 수 있다. 본 논문에서는 Cramer-Rao 하한을 이용하여 사전 정의된 실측 모델의 오차에 따른 카메라 추적 오차의 하한을 정의하고, 이를 통해 사전 정의된 실측 모델이 가질 수 있는 오차의 하한을 실험적으로 검증한다.

*교신저자: 박종일(jipark@hanyang.ac.kr)

감사의 글: 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2009년도 문화 콘텐츠산업기술지원사업의 연구결과로 수행되었음(3차년도).

2. 하이브리드 카메라 추적

특징점 기반과 모델 기반의 카메라 추적을 함께 이용한 하이브리드 카메라 추적을 위해서는 사전 정의 된 3차원 실측 모델과 특징점 데이터베이스가 준비되어야 한다. 특징점 데이터베이스를 구축하기 위해서 SURF[4]를 사용하여 다양한 시점으로부터 특징점을 검출하고, bundle adjustment[8]를 이용하여 특징점의 3차원 좌표를 복원한다. 이렇게 얻어진 3차원 특징점 정보들을 데이터베이스로 구축한다. 3차원 실측 모델과 특징점 데이터베이스가 주어지면, SURF[4]를 이용하여 입력 영상의 특징점을 검출하고, 검출된 특징점과 특징점 데이터베이스를 비교하여 초기 카메라 자세를 계산한다. 또한, Canny operator[9]를 이용하여 입력 영상의 윤곽을 검출하고, 이를 실측 모델의 윤곽과 정합시킴으로써 카메라 추적을 가능케 한다. 입력 영상과 실측 모델, 특징점 데이터베이스를 각각 비교하여 얻은 카메라 자세에는 재투영(re-projection) 오차가 생기는데, 보다 정확한 자세 갱신을 위하여 각각의 추적 방법에서 발생하는 오차의 의존도를 서로 다르게 융합시킨 후, 이를 최소화하는 가중치를 이용하여 카메라 자세를 계산 및 갱신한다. 하이브리드 카메라 추적의 전체적인 흐름은 그림1과 같다.

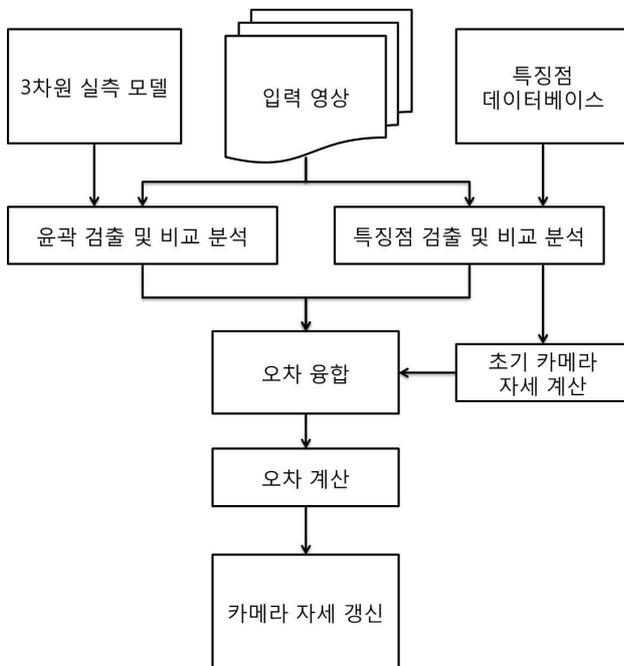


그림 1. 하이브리드 카메라 추적의 전체적인 흐름

카메라 자세는 리 군(Lie group)의 수학에 기반 하여 계산된다. 만약 카메라의 내부 변수 및 데이터베이스와 입력 영상의 특징점 사이의 상관도를 알고 있다면, 카메라와 추적 대상물 사이의 카메라 자세를 계산할 수 있다. 카메라 자세 갱신은 식1과 같다.

$$E^{(k+1)} = E^{(k)} M \quad (1)$$

이 때, M 은 4x4 행렬로 x, y, z 방향의 이동과 x, y, z 축의 회전을 나타내고, 다음 여섯 개 행렬의 가중치 합(weighted sum)으로써 표현된다.

$$M = \sum_{i=0}^5 \exp(\alpha_i G_i) \quad (2)$$

$$G_0 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, G_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, G_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (3)$$

$$G_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, G_4 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, G_5 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

가중치(α_i)는 특징점 기반 추적과 모델 기반 추적에서 발생하는 각각의 오차를 융합한 최종 오차의 값을 최소화하며, 그 값은 식4와 같다.

$$\alpha = \min_{\alpha} \{ \eta_e \| \mathbf{q} - P\alpha \|^2 + \eta_f \| \mathbf{s} - R\alpha \|^2 \} \quad (4)$$

여기서 $\eta_{e,f}$ 는 시각 신호(visual cue)의 의존성(reliability)을 나타 내주는 변수이고,

$$\eta_e = \frac{N_e}{S_e} - \frac{N_f}{D_f} \quad (5)$$

$$\eta_f = 1 - \eta_e \quad (6)$$

$P, R, \mathbf{q}, \mathbf{s}$ 는 각각의 시각 신호와 연관되는 행렬 및 벡터이다. 가 중치는 식4를 편미분하여 구할 수 있고, 그 결과는 다음과 같다.

$$\{ \eta_e (P^T P) + \eta_f (R^T R) \} \alpha = P^T \mathbf{q} + R^T \mathbf{s} \quad (7)$$

3. 상호작용형 모델링

하이브리드 방식의 카메라 추적을 효과적으로 하기 위하여 상호 작용형 모델링을 이용한다. 상호작용형 모델링을 이용하여 3차원 실측 모델을 제작할 수 있는데 이 때, 실측 모델은 데이터베이스에 저장되어 있는 3차원 특징점 정보를 이용하여 제작된다. 본 논문에서 이용된 상호작용형 모델링에서는 사용자가 제작하고자 하는 모델의 꼭지점을 직접 선택 및 입력하도록 하여 3차원 다면체 모델을 제작할 수 있도록 한다. 사용자의 상호작용을 통해 선택된 꼭지점들은 면을 이루게 되며, 면 내부에 존재하는 특징점과 데이터베이스에 저장되어 있는 3차원 특징점 정보 사이의 대응 관계를 통하여 3차원 실측 모델의 제작이 가능하다. 이렇게 제작된 3차원 실측 모델은 특징점 기반과 모델 기반 카메라 추적을 더불어 활용한 카메라 추적을 하는데 있어 필수적으로 요구 되는, 사전 정의 된 실측 모델로써 이용될 수 있다(그림 3). 본 논문에서는, 간단하고 직관적인 비교 및 분석을 위하여 상호작용형 모델링으로 제작하는 3차원 실측 모델을 직육면체 형태로 제한한다. 다음 장에서는, 이렇게 제작된 직육면체 형태의 3차원 실측 모델을 이용하여 사전 정의된 실측 모델의 오차와 모델 기반 카메라 추적의 오차 사이의 민감성을 알아본다.

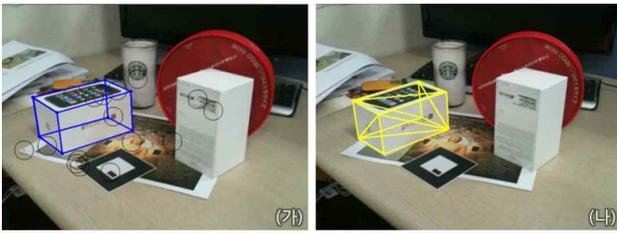


그림 3. 상호 작용형 모델링을 이용한 카메라 추적. (가) 상호 작용형 모델링으로 제작된 3차원 실측 모델. (나) 상호 작용형 모델링을 이용한 카메라 추적 결과

4. Cramer-Rao 하한을 이용한 추적 오차 분석

상호작용형 모델링을 이용하여 3차원 실측 모델을 보다 유연하게 제작하고, 이를 모델 기반 카메라 추적에 활용할 수 있다. 하지만, 상호 작용형 모델링을 통해 제작된 실측 모델은 레이저 측정과 같은 실측 과정을 통해 제작된 모델에 비해 상대적으로 그 정확도가 떨어지는 단점이 있다. 모델 기반 카메라 추적에서 사진 정의된 실측 모델의 정확성은 카메라 추적의 정확성을 결정하는 중요한 요인이며, 부정확한 실측 모델은 그림4와 같이, 카메라 추적의 정확성을 떨어뜨릴 수 있다. 사진 정의된 실측 모델의 정확성과 카메라 추적의 정확성 간의 관계를 살펴보기 위해 Cramer-Rao 하한(CRLB)을 이용하여 추적 오차의 하한을 정의한다.

Cramer-Rao 하한은 확률적으로 분포하는 데이터 추정량이 가질 수 있는 최소 분산을 의미한다. 간단하게 말하면, Cramer-Rao 하한 비편향추정량(unbiased estimator)으로부터 계산한 분산의 최소값은 피셔 정보(fisher information)의 역수가 된다. 스칼라 비편향추정량(scalar unbiased estimator)일 때 Cramer-Rao 하한은 다음과 같다.

$$var(\theta^*) \geq \frac{1}{I(\theta)} \quad (8)$$

이 때, 피셔 정보 $I(\theta)$ 는 우도함수(likelihood function)에 자연로그를 씌우고, 편미분한 값의 제곱의 조건부 확률을 나타낸다. 피셔 정보는 다음 식과 같이 정의되고,

$$[I(\theta)] = E \left[\left(\frac{\partial \ell(x; \theta)}{\partial \theta} \right)^2 \right] = -E \left[\frac{\partial^2 \ell(x; \theta)}{\partial \theta^2} \right] \quad (9)$$

$\ell(x; \theta) = \log f(x; \theta)$ 는 우도함수의 자연로그 값을, E 는 기대값을 나타낸다.

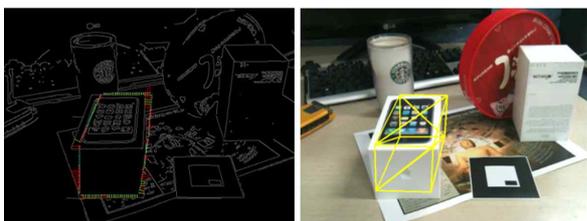


그림 4. 부정확한 3차원 실측 모델로 인한 카메라 추적 부정확성

Cramer-Rao 하한은 이용하여 추적 오차의 분산 최소값을 정의하기에 앞서, 실험에 사용 될 직육면체 형태의 3차원 실측 모델의 실측 정보는 그림5와 표1과 같다.

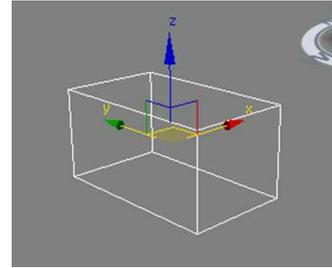


그림 5. 3차원 사진 모델의 실측 정보

표 1. 3차원 사진 모델의 실측 좌표

3차원 사진 모델의 실측 좌표 (단위:mm)								
꼭짓점	1	2	3	4	5	6	7	8
x축	-39.5	39.5	-39.5	39.5	-39.5	39.5	-39.5	39.5
y축	-67.5	-67.5	-67.5	-67.5	67.5	67.5	67.5	67.5
z축	-69.5	-69.5	0	0	-69.5	-69.5	0	0

실험을 위해서 3차원 사진 모델에 가우시안 노이즈(Gaussian noise)의 편차를 변화시켜 가며 추적 오차를 계산한다. 그림6에 보여진 실험 모델을 이용하여 가우시안 노이즈가 1에서 5로 변화할 때 추적 정확성의 변화를 알아본다. 가우시안 노이즈가 변화함에 따라 달라지는 추적의 추이는 그림 7에 보여 진다.

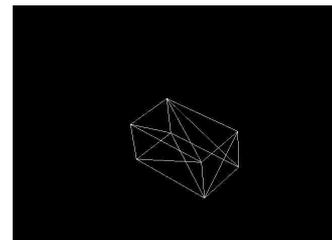


그림 6. 실험 모델

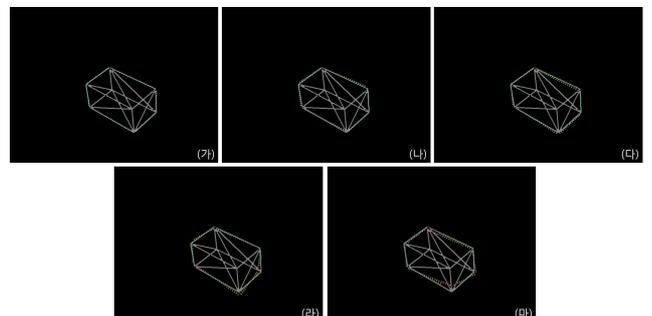


그림 7. 가우시안 노이즈 변화에 따른 추적 변화. (가) $\sigma = 1$, (나) $\sigma = 2$, (다) $\sigma = 3$, (라) $\sigma = 4$, (마) $\sigma = 5$.

상호작용형 모델링에서는 추적 대상물의 꼭짓점과 대응되는 좌표를 사용자가 직접 선택하기 때문에, 회전 오차는 없고 이동 오차만 있다고 가정한다. 따라서 Cramer-Rao 하한을 이용하여 추적 오차의 분

산 하한을 정의할 때, 가중치(α)의 분산 중, 이동 변수를 나타내는 $\alpha_0, \alpha_1, \alpha_2$ 에 초점을 두어 계산한다. 그림8-10은 가우시안 노이즈 편차의 변화에 따라, $\alpha_0, \alpha_1, \alpha_2$ 의 값이 변화하는 추이를 나타낸다. Cramer-Rao 하한 값과 실제 추적 과정에서 나타난 오차의 분산 값이 비슷한 추이를 나타내는 것을 확인할 수 있다.

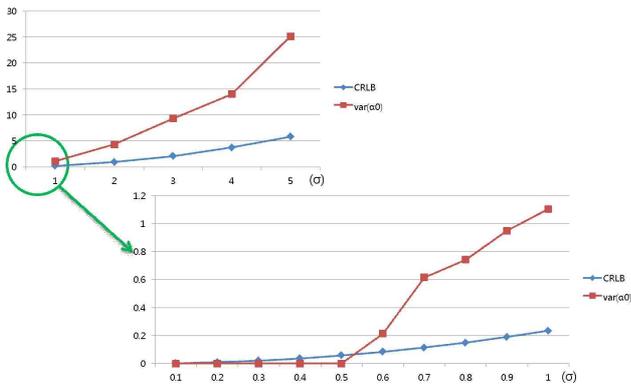


그림 8. α_0 Cramer-Rao 하한과 실제 오차 분산값

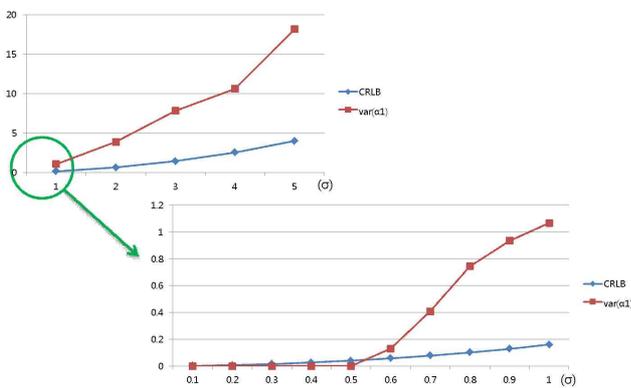


그림 9. α_1 Cramer-Rao 하한과 실제 오차 분산값

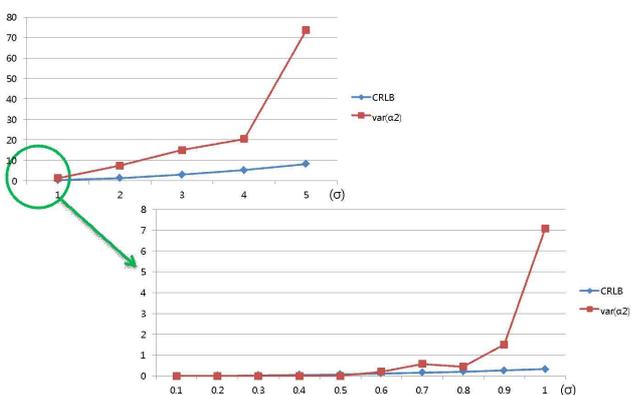


그림 10. α_2 Cramer-Rao 하한과 실제 오차 분산값

5. 결론

본 논문에서는 모델 기반 카메라 추적 시 필요한, 사전 정의 된 실측 모델의 정확성이 카메라 추적의 정확성에 미치는 영향에 대하여 논

의하였다. 이를 위하여 모델 기반과 특징점 기반의 카메라 추적에 가중치를 두고 혼합시켜 이용하는 하이브리드 방식의 카메라 추적 방법을 사용하고, 이를 활용한 상호작용형 모델링을 이용하여 3차원 실측 모델을 제작하였다. 또한, 모델 오차에 따른 추적 오차를 Cramer-Rao 하한을 이용해 정의하고 그 결과를 실험적으로 도출하였다. 하지만 본 논문에서 제안된 상호작용형 모델링은 실측 모델의 오차에 따른 카메라 추적 오차의 민감성을 보이기 위해, 간단하고 원시적인 직육면체 형태로 제한되었으며, 추적 오차를 분석함에 있어서도 Cramer-Rao 하한을 이용하여 오차의 분산 최소값을 정의하는 데 그쳤다. 향후, 더 정교하고 직관적인 모델링을 할 수 있는 방향으로 상호작용 모델링을 발전시킬 예정이고, 보다 정확하고 직관적으로 실측 모델의 오차가 카메라 추적의 정확성에 미치는 영향을 분석할 수 있는 오차 분석 방법을 연구 중에 있다.

6. 참고 문헌

- [1] H. Kato and M. Billinghurst. 1999. "Marker tracking and hmd calibration for a video-based augmented reality conferencing system." In *proceedings of IWAR'99*.
- [2] M. Fiala. 2005. "ARTag, a fiducial marker system using digital techniques." In *proceedings of IEEE CVPR'05*.
- [3] D.G. Lowe. 2004. "Distinctive image features from scale-invariant keypoints." *International Journal of Computer Vision (IJCV)*, vol. 60, no. 2, pp. 91-110.
- [4] H. Bay, A. Ess, T. Tuytelaars, and L.V. Gool. 2008. "Speeded-up robust features (surf)." *Computer Vision and Image Understanding*, vol. 110, issues 3, pp. 346-359.
- [5] M. Ozuysal, M. Calonder, V. Lepetit, and P. Fua. 2010. "Fast keypoint recognition using random ferns." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, issues 3, pp. 448-461.
- [6] T. Drummond and R. Cipolla. 2002. "Real-time tracking of complex structures with on-line camera calibration." *Image and Vision Computing*, vol. 20, issues 5-6, pp. 427-433.
- [7] H. Park, J. Oh, B. -K. Seo, and J.-I. Park. 2010. "Automatic confidence adjustment of visual cues in model-based camera tracking." *Computer Animation and Virtual Worlds*, vol. 21, no. 2, pp. 69-79.
- [8] M.I.A. Lourakis and A.A. Argyros. 2009. "SBA: A software package for generic sparse bundle adjustment." *ACM Trans. Mathematical Software*, vol. 36, no. 1, pp. 1-30.
- [9] J. Canny. 1986. "A computational approach to edge detection." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, issues 6, pp. 679-714.