

2D 텍스처맵과 3D 모델을 이용한 2.5D 물체 추적 및 카메라 캘리브레이션 알고리즘

홍현석, 정명진
한국과학기술원 전기 및 전자공학부

Robust 2D Texture Map and 3D Model Based 2.5D Object Tracking and Camera Calibration

Hyun Seok Hong, Myung Jin Chung
Department of Electrical Engineering & Computer Science, Korea Advanced Institute of Science and Technology

Abstract - 기존 2D 추적기들은 영상에서 특정 평면 영역을 원근 투영하에서 만족할 만한 추적결과를 보여주었다. 하지만 2D 추적기는 2D 영역들로 이루어진 3D 물체를 영상에서 추적하는 경우, 물체자신의 회전에 의해 가려지거나 새로 나타나는 영역에 대해 대응하지 못하여 추적이 실패하게 되지만, 3D 정보를 이용한다면 이러한 사라짐과 나타나는 영역을 예측하고 완벽하게 추적할 수 있게 된다. 본 연구에서는 일련의 영상으로부터 3D 모델과 2D 텍스처맵을 추출하고, 이를 이용하여 3D 물체의 회전과 평행이동 움직임을 추적한다. 또한 카메라의 줌 파라미터를 모델링하고 추적기 알고리즘에 추가하여, 물체의 3차원 파라미터의 추적과 동시에 카메라 줌 파라미터를 추적하였다.

1. 서 론

컴퓨터비전이나 이미지프로세싱 분야에서 관심영역의 추적은 가장 많이 사용되는 기술 중 하나이며, 그 중요성 못지않게 많은 추적기법이 개발되어 왔다. 많은 추적기법 중에서 정밀도가 가장 높다고 알려진 방법은 gradient descent 방법이며, 이는 오차함수를 최적화시키는 최적해 파라미터를 찾음으로써 얻을 수 있다. Gradient descent 방법을 통해 영상추적을 한 최초의 알고리즘은 Lucas와 Kanade의 알고리즘[1]이며, 이는 주어진 두 영상을 정렬하기 위해, 첫 번째 영상을 어파인 혹은 원근 움직임을 통해 두 번째 영상과 일치될 수 있도록 오차함수를 정하고, 이 오차함수를 최소화하는 해당 움직임을 파라미터를 구하는 방법이었다. 최근 Luscas-Kanade 알고리즘은 Hager와 Belhumeur, 그리고 Baker와 Matthews에 의해 효율성을 획기적으로 개선시킨 알고리즘[3,4]이 개발되었다.

하지만 이러한 2차원 영상내에서 영상의 움직임을 추적하는 것은 근본적인 한계를 지닌다. 2차원 영상은 3차원 물체에 대한 사영된 영상이므로, 3차원 물체의 회전이나 이동에 의해 가려지는 부분이 생기거나, 가려졌던 부분이 다시 화면에 나타나는 경우에 전혀 대처할 수 없게 된다. 따라서 3차원 물체에 관한 정보를 활용하여 2차원 영상에서 물체를 추적한다면, 가려짐에 대해 훨씬 더 강한 알고리즘이 될 수 있다.

본 연구에서는 2차원 영상으로 구성된 텍스처맵과 3차원 물체의 CAD 모델을 이용하여 3차원 움직임과 카메라 파라미터를 추적하는 방법을 수식화 전개하고 그 실험결과를 보여준다.

2. 3차원 물체의 CAD모델 및 2차원 텍스처 맵 획득하기

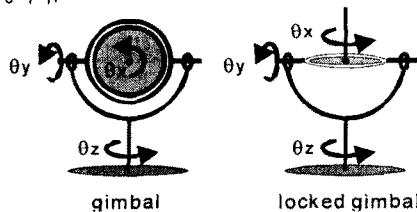
2차원영상을 사용하여 2차원 움직임을 추적하는 2D 추적기, MRI 영상처럼 3차원 영상에서 3차원 움직임을 추적하는 방법을 3D 추적기라고 명명하고, 2차원 텍스처맵 영상과 3차원 CAD모델을 사용하는 추적기를 2.5D 추적기라는 이름으로 사용하도록 한다.

3차원 CAD모델과 2차원 텍스처맵은 공장에서 출시되는 물체인 경우, 쉽게 얻을 수 있지만, 일반적인 물체의 경우 규격화 되지 않았으므로 직접 구해야 하는 필요성이 생긴다. 이는 SFM(structure from motion) 알고리즘을 통해 구성할 수 있다. 본 연구에서는 SFM 알고리즘 중에서 factorization method를 사용하여, 일반 3차원 물체가 사영된 일련의 영상들로부터, 물체의 3차원 CAD모델을 얻고, 2차원 텍스처맵을 구성하였다.

3. 2.5D 물체추적기

3.1 3차원 움직임 모델링

3차원 움직임은 고정된 물체에 대해, 움직이는 카메라의 3차원 회전행렬 (R)과 3차원 위치변위벡터(T)를 가리킨다. 이 회전과 위치변위는 물체에 대한 상대적인 값이며, 움직이는 물체와 고정된 카메라 등에 대해서도 유사하게 사용가능하다.



〈그림 1〉 Euler 회전표현을 구현하는 gimbal의 lock상태

3차원 회전을 모델링할 때, 가장 많이 사용되는 방식은 Euler xyz 방법이며, 대부분의 3차원 회전이 세 번의 연속적인 회전을 통해서 표현할 수 있다는 방식에 기초한 방법이며 3개의 파라미터를 갖는다. 3차원 회전의 자유도는 3개를 가지므로, 이 Euler xyz 방법은 파라미터수가 자유도와 동일하다는 장점이 있지만, 특정 회전위치에서는 임의방향 회전을 할수 없는 짐벌락(gimbal lock) 문제가 생기므로, 임의의 회전을 표현하는 방식으로는 사용할 수 없다. 컴퓨터 그래픽스에서 부드러운 움직임을 표현하기위해 많이 사용되는 quaternion 방식은 크기에 제한을 갖는 4개의 파라미터로 표현되며, 임의의 회전을 표현할 수 있다는 장점이 있다. quaternion 방식을 이용한 회전행렬은 다음과 같다.

$$R = \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix} \quad (1)$$

단, $q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$ 이며, 식(1)로 표현되는 행렬은 특정 회전 축 $u = (q_1, q_2, q_3)^T$ 을 기준으로 특정 회전량 $\theta = 2\cos^{-1}(q_0)$ 만큼 회전시키는 회전행렬이 된다. 3차원 위치변위벡터는 $T = (t_x, t_y, t_z)^T$ 이다.

3차원 움직임 파라미터 벡터를 $p_{RT} = (1 - q_0, q_1, q_2, q_3, t_x, t_y, t_z)^T$ 로 정의하면, 3차원 공간상의 점 $x_i = (x, y, z)^T$ 를 이동시키는 워프함수는 다음과 같이 표현할 수 있다.

$$W(x_i; p_{RT}) = R(p_{RT})x_i + T(p_{RT}) \quad (2)$$

3.2 Full perspective camera

3차원에서 평행한 두 직선은 2차원으로 사영되면 일반적으로 더 이상 평행하지 않게 된다. 즉 동일한 3차원 두 직선에 대해, 평행하게 사영된 영상과 평행하지 않게 사영된 영상사이에서는 perspective distortion이 나타나게 된다. 따라서 2D 추적기에서는 어파인 움직임으로는 추적이 불가능하게 되며, 2.5D 추적기에서는 weak perspective camera로는 추적이 불가능하게 된다. 본 연구에서는 full perspective camera는 행렬 K로서 다음과 같이 표현할 수 있다.

$$K = \begin{pmatrix} 1 + f_u s(1 + f_u) u_0 & & \\ 0 & 1 + f_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3)$$

카메라 초점거리 파라미터를 $p_K = (f_u, f_v)^T$ 로 정의하면, 사영함수 homogeneous coordinate 형태로 정의하면 다음과 같다.

$$P(x_i; p_K) \sim K(p_K)x_i \quad (4)$$

3.2 오차함수 정의 및 2.5D 추적알고리즘

두 파라미터 벡터 p_K 와 p_{RT} 의 초기값을 각각 p_{K0} 와 p_{RT0} 라 하고, 하나의 파라미터벡터 $p = (p_{RT}^T, p_K^T)^T$ 로 구성된 후, 워프함수를 다음처럼 하나의 함수로 구성 하자.

$$PW(x_i; p) = P(W(x_i; p_{RT0} + p_{RT}); p_{K0} + p_K) \quad (5)$$

3차원공간에서 N개의 점을 추적하는 2.5D 추적기를 위한 오차함수는 식 (2, 4, 5)를 이용하여, 다음과 같이 정의할 수 있다. (단, 0는 영벡터)

$$\sum_{i=1}^N [I(PW(x_i; p + \Delta p)) - T(PW(x_i; 0))]^2 \quad (6)$$

구하는 해는 Gauss-Newton 최적화를 통하면 다음과 같다.

$$\Delta p = H^{-1} \sum_{i=1}^N \left[\frac{\partial I(PW(x_i; p))}{\partial p} \right]^T [I(PW(x_i; p)) - T(PW(x_i; 0))] \quad (7)$$

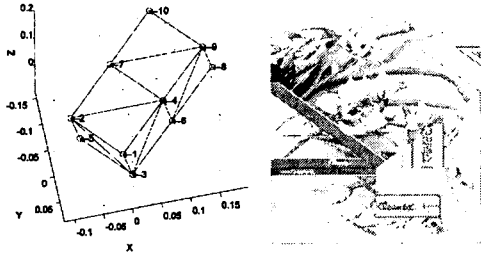
$$H = \sum_{i=1}^N \left[\frac{\partial I(PW(x_i; p))}{\partial p} \right]^T \left[\frac{\partial I(PW(x_i; p))}{\partial p} \right] \quad (8)$$

식(7)을 통해 계산되는 움직임 파라미터 벡터의 증분값을 통한 최종 파라미터의 업데이트는 다음식과 같다:

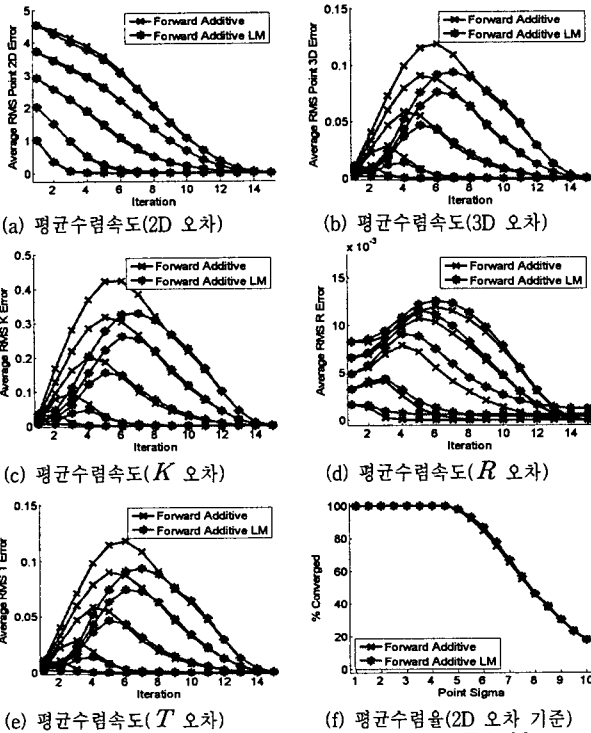
$$p \leftarrow p + \Delta p \quad (9)$$

4. 실험결과

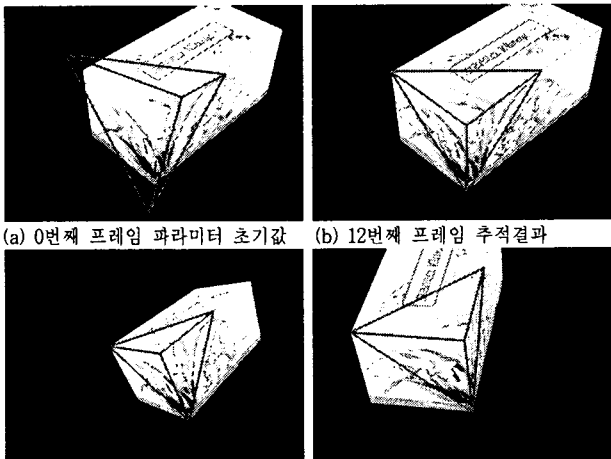
3차원 물체추적을 위해 그림2와 같이 3차원 삼각형 메쉬(mesh)모델로 이루어진 박스에 2차원의 텍스처맵을 구성하였다. 텍스처맵과 3차원모델의 사영된이미지 사이의 매핑은 homogeneous coordinate 표현과, 3x3행렬의 호모그래피(homography)를 통해 이루어진다.



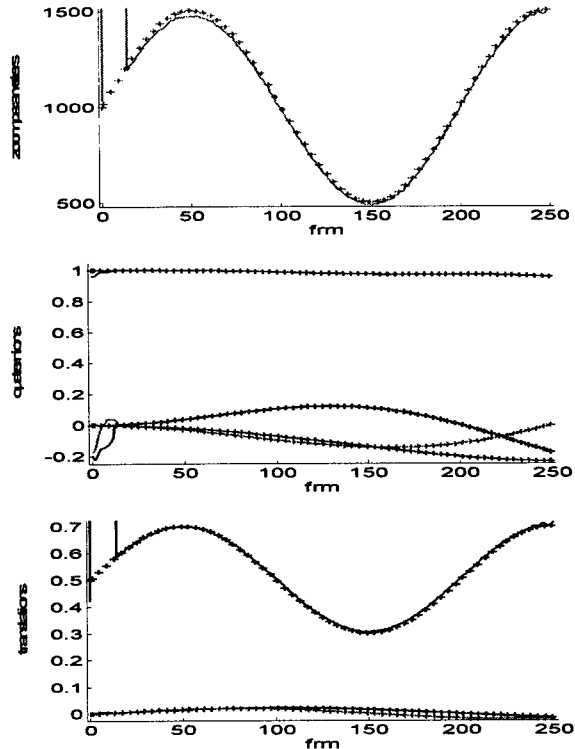
〈그림 2〉 3D 모델 및 2D 텍스처 맵



〈그림 3〉 2.5D추적기의 평균수렴속도 및 평균수렴율
(×:Gauss-Newton방법, * :Levenberg-Marquadt방법)



〈그림 4〉 합성된 일련의 이미지에 대해 2.5D 추적결과 (추적영역은 녹색 삼각형 메쉬로 표현)



〈그림 5〉 합성된 일련의 이미지에 대해 K, R, T 파라미터 추적결과 (10프레임단위의 +마크:ground truth, 실선:추적결과)

그림3은 5000개의 임의 변화시킨 영상에 대해, 2.5D추적결과를 보여준다. 2D 오차와 3D 오차는 각각 추적영역의 네 꼭지점의 2D와 3D 위치오차의 RMS(root mean square)값을 사용하였으며, 수렴 여부는 15번의 반복되는 최적화 계산 후에 2D RMS오차가 1픽셀오차보다 작은 경우에 수렴한 것으로 간주하였다. K, R 및 T 오차는 해당 파라미터 오차벡터의 놈(norm)을 사용하였다. 식(6)에서 보듯이 최적화하는 대상은 2차원 오차이므로 2차원 오차에 맞추어 추적을 초기오차보다 적어지게 진행하는 것에 비해, 3차원 관련 오차는 초기오차가 큰 경우 증가후 감소추세를 보여준다. 그림4와 그림5는 251장의 합성된 일련의 동일영상에 대해, 2.5D 추적결과를 보여준다. 그림4에서 보듯 2차원 오차는 점진적으로 감소하지만, 그림5에서 보면 3차원 오차들은 안정화에 들어서기까지 활발하게 변화를 진행한다. 특히, 유사한 물리적 역할 때문에, 카메라 줌 파라미터 f_u, f_v 와 3차원 변위벡터의 t_x 값 사이의 파라미터 조정이 활발하다. 그림4-(a)에서 보듯 초기 파라미터가 상당한 에러를 포함함에도 불구하고, 약 12프레임만에 전체 파라미터가 안정화에 들어서고 정밀한 추적을 진행한다 (파라미터 초기오차가 작을수록, 프레임간 많은 반복계산을 허용할수록 더 적은 프레임이 소요).

5. 결 론

기존 2D 추적기에서는 3차원 회전이나 변위에 의한 가려짐에 대응할 수 없는 반면, 3D 추적기는 가려짐과 나타남에 대해 예측하고 대응할 수 있다는 장점이 있다. 본 연구에서는 가장 정밀하다고 알려진 gradient descent방법에 기반하여 2차원 텍스처맵과 3차원 모델을 이용하여, quaternion을 통해 임의의 3차원 회전과 변위를 추적하고, 동시에 카메라의 줌까지 추적하는 2.5D 추적기를 개발하고 분석하였다.

Acknowledgement

본 연구는 정보통신부 대학 IT연구센터 육성, 지원사업의 연구결과로 수행되었습니다.

참고 문헌

- [1] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," In Proceedings of the International Conference on Artificial Intelligence, pp.674-679, 1981.
- [2] R. Dutter and P. Huber, "Numerical Methods for the Nonlinear Robust Regression Problem," J. Statistical Computer Simulation, vol. 13, no. 2, pp. 79 - 113, 1981.
- [3] G. D. Hager and P. N. Belhumeur, "Efficient Region Tracking with Parametric Models of Geometry and Illumination," IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(10):1025 - 1039, 1998.
- [4] S. Baker and I. Matthews, "Equivalence and efficiency of image alignment algorithms," In Proceedings of the IEEE Conference of Computer Vision and Pattern Recognition, vol. 1, pp. 1090-1097, 2001.