

가상현실 방송 제작을 위한 모델 기반 카메라 보정

*오주현 **손광훈

* 연세대학교 디지털영상미디어연구실, KBS 방송기술연구소

** 연세대학교 디지털영상미디어연구실

*jhoh@diml.yonsei.ac.kr

Model-based Camera Calibration for Virtual Production

*Oh, Ju-Hyun **Sohn, Kwang-Hoon

* Yonsei University Digital Image Media Lab., KBS Broadcast Technical Research Institute

** Yonsei University Digital Image Media Lab.

요약

자연스러운 가상현실 제작을 위해서는 정확한 카메라 보정(camera calibration) 과정이 필수적인 선결 조건으로 요구된다. 그러나 기존의 영상처리에 의한 카메라 보정 방식은 특징점 추출에서의 에러 발생과 여러 장의 영상을 촬영해야 하는 등의 단점으로 줌렌즈 카메라 보정에는 사용되기 힘들었다. 본 논문에서는 카메라보정 객체의 모델에 기반하여 카메라 파라미터를 최적화하는 방법으로 카메라 보정을 구현하였다. 최적화 방법으로는 경사기반 방식에 비해 국부최적점에 갇힌 것으로 알려진 유전자알고리즘(genetic algorithm)을 사용하였다. 카메라 보정 객체에 낮은 공간주파수 성분을 보강하고, 목적함수에 영상의 밝기 정보를 포함하며, 유전자알고리즘을 사용함으로써 초기치가 최적점에서 멀리 떨어져있는 경우에도 수렴이 가능함을 실험적으로 확인하였다.

1. 서론

가상현실 제작(virtual production)은 이제 방송프로그램 전반에 걸쳐 보편적인 제작방식이 되었으며, 점차 그 활용도를 높여 가고 있다. 가상현실 제작은 일종의 혼합현실(mixed reality) 어플리케이션으로서, 콘텐츠의 완성도와 사실감에 있어서 그래픽의 품질 뿐 아니라 실사와 그래픽의 정확한 정합(registration)이 매우 중요하다. 정확한 정합은 카메라 보정과 카메라 트래킹의 정확도에 의해 결정된다. 따라서 기계적 카메라 트래킹 시스템을 사용하고 있는 대부분의 제작 현장에서 카메라 보정의 정확도가 제작된 콘텐츠의 품질을 크게 좌우한다고 말할 수 있다.

카메라 보정은 수작업에 의한 방식과 영상처리에 의한 방식으로 나눌 수 있다. 기존의 영상처리에 의한 카메라보정 방식은 영상에서 특징점(feature points)을 추출하고 이 특징점과 카메라보정 패턴(calibration pattern) 사이의 대응 관계(correspondence)를 구한 후 이들 사이의 선형방정식 관계를 이용하여 카메라 파라미터들을 계산하는 2단계 구조로 이루어진다 [1]. 이러한 방식은 특징점 추출 단계에서 필연적으로 발생하는 에러가 다음 단계로 전파되는 단점과 다단계 계산으로 인한 설정의 복잡성 등으로 인해 널리 사용되고 있지 않으며, 최근에 많이 사용되고 있는 호모그래피(homography)에 의한 카메라 보정 방식도 한 번의 카메라 보정을 위해 여러 장의 영상을 촬영해야 하는 단점이 있어 [2], 수많은 줌-포커스 쌍(zoom-focus pair)에 대해 보정을 실시해야 하는 줌렌즈 카메라에는 사실상 사용이 불가능하다 [3]. 이러한 이유로 실제 제작 현장에서는 수작업에 의한 카메라 보정

이 선호되고 있다.

본 논문에서는 복잡한 전처리과정 없이 한 장의 입력영상만으로 처리 가능한 모델 기반(model-based) 카메라보정 방법을 제안한다 [4, 5]. 그림 1은 카메라보정 객체의 모델과 카메라, 그리고 입력영상 간의 관계를 보여준다. 초기값으로 주어진 카메라 파라미터에 따라 투영(projection)된 모델과 입력영상 간의 유사도(similarity)를 목적함수(objective function)로 설정한 후, 이 목적함수를 최대화하는 카메라 파라미터를 구한다. 이 때 전체 영상을 합성하여 입력영상과 픽셀 단위로 비교하는 것은 시간이 많이 걸리므로, 패턴의 주요 측정점들(measure points)을 투영하여 입력영상의 해당 픽셀과 비교하였다. 각 측정점의 속성에 따라 조사된 유사도의 합을 에너지함수로 정의하고, 이를 최적화하기 위한 방법으로는 경사 기반(gradient-based) 방식에 비해 국부최적점(local optimum)에 갇힌 것으로 알려진 유전자알고리즘(genetic algorithm)을 사용하였다. 실험 결과에서는 가상의 입력영상을 사용한 컴퓨터 시뮬레이션 결과를 보인다.

2. 카메라와 카메라보정 객체 모델

가. 카메라 모델

가상스튜디오를 포함해서 대부분의 가상현실 제작을 위한 그래픽 소프트웨어에서는 단순한 핀홀(pinhole) 카메라 모델을 사용한다. 따라서 가상 카메라가 추적하는 실제 카메라 또한 보정 및 추적에 있어서 핀홀 모델을 사용하는 것으로 충분하다. 이는 더 복잡한 카메라 모델을

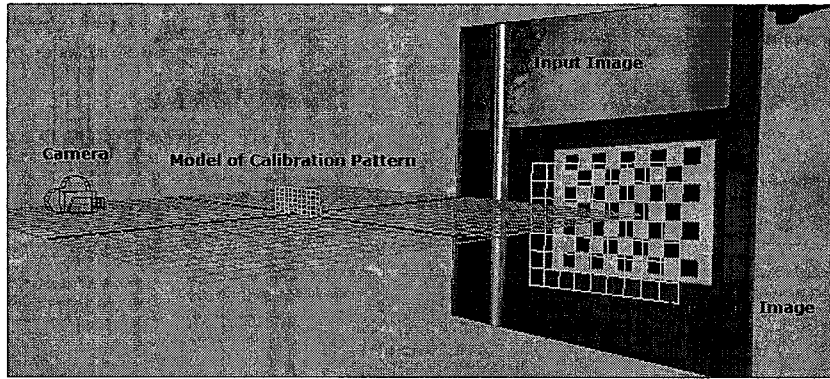


그림 1. 모델 기반 카메라보정의 개념

사용하더라도 다음 단의 그래픽 카메라에 적용할 때 편향 모델 이외의 파라미터들은 무시되기 때문이며, 본 논문에서도 보정할 카메라를 편향 카메라로 모델링하였다. 편향 카메라는 카메라의 위치(position)와 방향(orientation)을 나타내는 여섯 개의 파라미터들과 화각(field of view)으로 정의된다.

$$C = \{fov, r_x, r_y, r_z, t_x, t_y, t_z\} \quad (1)$$

식 (1)의 일곱 개의 파라미터들과 기 정의된 가로 세로 해상도 및 종횡비(aspect ratio)로부터 카메라의 투영 식을 구할 수 있다 [6, 7].

나. 카메라 보정을 위한 객체 모델

일반적으로 카메라 보정을 위해 사용되는 객체(calibration object)는 평면이나 입방체 위에 체스보드처럼 반복되는 패턴(pattern)을 표시하여 제작한다. 그러나 체스보드와 같이 일정한 주기성을 가지는 패턴은 특정한 공간 주파수(spatial frequency) 성분만을 가지므로, 최적화 과정에서 패턴의 한 주기 이상을 뛰어넘는 것이 불가능하여 국부 최적점에 수렴할 가능성이 크다. 따라서 본 논문에서는 그림2와 같이 일반적인 체스보드 패턴과 유사하면서 저주파 성분을 보강한 카메라 보정 객체를 사용하였다.

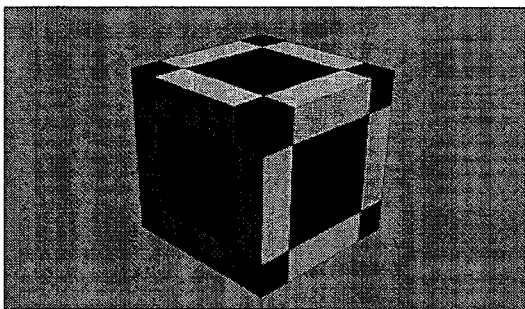


그림 2. 카메라보정 객체(calibration object)

실험에 있어서 카메라보정 객체는 그 중심점이 세계좌표계(world coordinates)의 원점에 있으며, 각 변의 크기가 1인 정육면체로 가정한다. 또한 일반적인 사용 환경을 고려하여 그림과 같이 3면이 보이도록 정육면체를 y축으로 45도 회전하였다.

3. 최적화(Optimization)

가. 전역최적점(Global Optimum) 탐색

특징점 추출이 없는 카메라보정을 제안한 L. Robert[4]는 흰 바탕에 검은 정사각형으로 이루어진 패턴을 사용하였으며, 사각형의 테두리에 위치한 측정점들의 gradient가 커야 한다는 점을 이용, 각 측정점에서의 gradient 제곱의 합을 목적함수로 설정하였다. 그러나 이 경우 초기치를 최적점에 거의 근사한 점으로 설정해야만 최적해를 얻을 수 있으며, 그렇지 않을 경우 국부최적점으로 수렴하고 만다.

본 논문에서는 전역최적점을 얻기 위해 다음과 같은 점들을 고려하였다:

- (1) 카메라보정 객체에 저주파 성분을 보강
- (2) 측정점의 gradient 뿐 아니라 밝기정보를 반영
- (3) gradient 이미지를 blurring하여 사용
- (4) 유전자 알고리즘의 사용

나열한 고려 사항들 중 (1),(2),(3)은 목적함수를 볼록한(convex) 형태에 가깝게 만들기 위한 방법이며, (4)는 볼록하지 않은(non-convex) 목적함수에서도 최적해를 찾기 위해 도입한 것이다.

나. 목적 함수(Objective Function) 정의

카메라 파라미터를 추출하기 위해, 임의의 상태로 주어진 카메라 파라미터가 얼마나 입력 영상과 부합하는지를 측정하는 목적함수 F 를 정의한다. 목적함수는 카메라 보정 객체의 표면에 배치된 측정점들에서 각각의 적합도(fitness)를 측정하여 더한 값으로 정의한다.

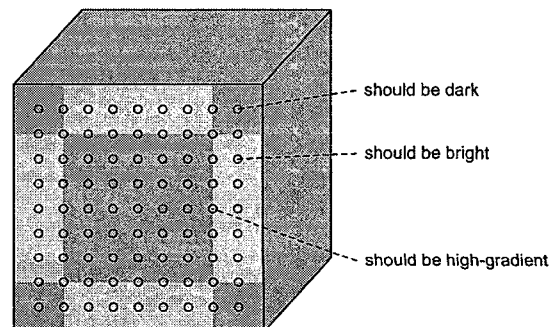


그림 3. 카메라 보정 객체와 측정점들

각각의 측정점들은 카메라보정 객체 모델의 표면 밝기에 따라서 그림 3과 같이 (Bright, Dark, High-gradient)의 세 가지 속성 중 하나를 가진다. 이와 같은 모델 포인트의 속성과, 실제 영상에서 측정된 픽셀 값을 비교하여 차이가 작으면 높은 적합도를, 차이가 크면 낮은 적합도를 가지도록 다음과 같이 카메라 파라미터 C 에 따른 적합도 함수 F 를 설정한다.

$$F(C) = \sum_i f(p_i) \quad (2)$$

where $f(p_i) =$

$I(Proj(C, p_i))$ if $Attrib(p_i) = Bright$

$1 - I(Proj(C, p_i))$ if $Attrib(p_i) = Dark$

$G(Proj(C, p_i))$ if $Attrib(p_i) = HG$

식에서 C 는 카메라 파라미터, p_i 는 i 번째의 측정점, I 는 입력 영상, G 는 gradient 영상을 각각 의미한다. $Proj(C, p_i)$ 는 점 p_i 를 C 에 의해 투영한 영상 위의 점을, $Attrib(p_i)$ 는 점 p_i 의 속성을 각각 나타낸다. 또한 영상은 $0 \leq I \leq 1$, $0 \leq G \leq 1$ 의 값 범위를 가진다.

나. 유전자 알고리즘 (Genetic Algorithm)

유전자 알고리즘은 최적화 문제에서 최적해를 찾기 위해 진화의 메커니즘을 도입한 방법이다. 유전자알고리즘에서 구하고자 하는 파라미터들은 일종의 유전자 스트링(genome)으로 간주되어, 자연계의 진화 과정에서 일어나는 상속(inheritance), 돌연변이(mutation), 자연선택(selection), 교차(crossover or recombination) 등을 거쳐서 최적의 상태로 진화한다 [8, 9]. 그림 4는 유전자에 해당하는 카메라파라미터에서 교차를 통해 다음 세대로 진화하는 과정의 예이다.

부	fov_1	r_{x1}	r_{y1}	r_{z1}	t_{x1}	t_{y1}	t_{z1}
모	fov_2	r_{x2}	r_{y2}	r_{z2}	t_{x2}	t_{y2}	t_{z2}
자1	fov_1	r_{x1}	r_{y1}	r_{z2}	t_{x2}	t_{y2}	t_{z2}
자2	fov_2	r_{x2}	r_{y2}	r_{z1}	t_{x1}	t_{y1}	t_{z1}

그림 4. 카메라 파라미터의 교차 예

교차 확률, 돌연변이 확률, 개체 수(population), 세대 수(generations) 등 실험을 위한 유전자알고리즘 변수들은 표 1과 같이 설정하였다.

표 1. 유전자 알고리즘 설정

설정변수	값
crossover	80 %
mutation	5 %
population size	50
generations	200
genome size	7

4. 실험 결과

가. 실험 영상

컴퓨터 그래픽으로 생성한 그림 5의 이미지를 실험 영상으로 사용하였으며, gradient는 매번 계산하지 않고 gradient 이미지를 만들어 사용하였으며, 빠른 수렴을 위해 Gaussian blurring하였다.

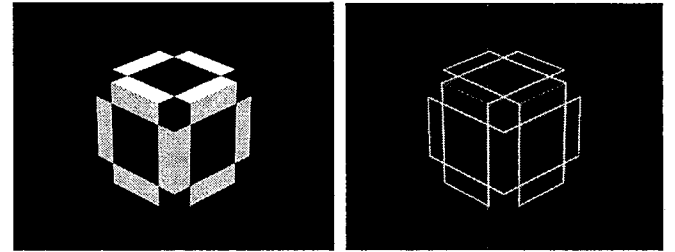


그림 5. 입력 이미지와 gradient 이미지

카메라보정 객체의 측정점들은 그림 3에 표시된 것처럼 카메라에서 보이는 3면에 각 81개씩, 총 243개를 배치하였다. 식 (2)에서 정의한 바와 같이 하나의 측정점에서 가질 수 있는 적합도 값은 0에서 1 사이이므로 F 가 가질 수 있는 이상적인 최대값은 243이다.

나. 실험 결과

표 2와 그림 6은 최적화에 의한 카메라 파라미터 추정 결과를 나타낸다. 그림 6에서 기존의 알고리즘으로는 추정이 불가능한 정도로 입력영상과 그 위에 투영된 모델 사이에 큰 차이를 보이고 있으나, 전역최적점을 찾기 위해 제안된 알고리즘으로 최적화를 수행한 결과, 모델이 입력영상에 정확하게 피팅(fitting)된 결과를 볼 수 있다. 모델의 측정점들은 정육면체 표면에 작은 원들로 표시되었으며, 원의 밝기는 해당 측정점에서 조사된 적합도 값을 나타낸다.

표 2. 카메라 파라미터 추정 결과

파라미터	초기값	결과값
fov	0.208	0.198
r_x	0.419	0.460
r_y	-0.015	0.027
r_z	0.071	0.006
t_x	-0.186	-0.271
t_y	4.711	5.011
t_z	-10.507	-10.110

제안된 알고리즘은 카메라보정을 위해 PC(2.2GHz dual CPU)에서 시각적 피드백을 포함하여 약 15초 정도의 수행시간을 요구하였다.

5. 결론

유전자 알고리즘을 사용하여 모델 기반 카메라 보정을 구현하였다. 저주파성분을 보강한 카메라보정 객체를 사용하고 목적함수에 blurring된 gradient 이미지를 사용하면서 측정점의 밝기 정보까지 함께 고려하여 최적화 과정이 안정적으로 수렴할 수 있게 하였다. 제안한 알고리즘은 복잡한 전처리나 특징추출 과정이 필요하지 않기 때문에 줌-포커스 쌍에 따라 수십에서 수백 번의 보정 과정을 필요로 하는 줌렌즈 카메라 보정에 효과적으로 사용될 수 있을 것으로 기대된다. 향후 제안된 알고리즘을 줌렌즈 카메라 보정에 실제로 적용하여 가상현실 제작에 활용할 계획이다.

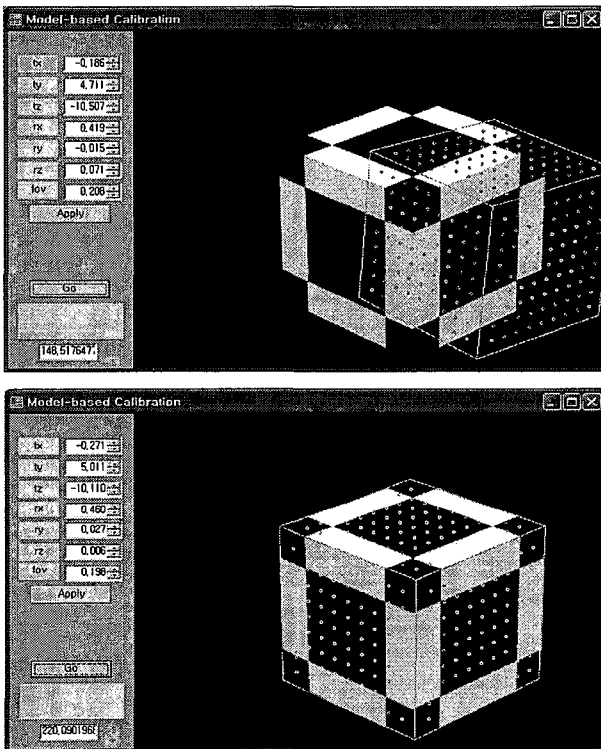


그림 6. 위: 초기 상태의 입력영상과 모델, 아래: 최적화 완료 상태

그림 7은 최적화에 따른 목적함수(적합도)와 FOV 파라미터의 추정 오류 변화를 보인다. 최적화가 진행됨에 따라 적합도는 증가하고 파라미터 추정 오류는 감소하는 것을 알 수 있다.

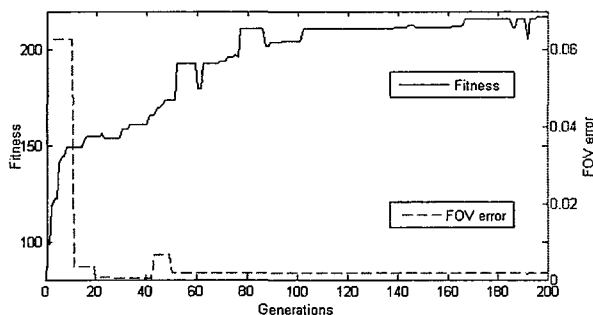


그림 7. 최적화 과정에 따른 적합도와 카메라보정 오류의 변화

참고 문헌

- [1] R.Y. Tsai, "Metrology Using Off-the-Shelf TV Cameras and Lenses", *IEEE Journal of Robotics and Automation*, Vol. 3, No. 4, pp. 323-344, August 1987.
- [2] Z. Zhang. "A flexible new technique for camera calibration". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330-1334, 2000.
- [3] Yong-Sheng Chen, Yi-Ping Hung, Chiou-Shann Fuh, Sheng-Wen Shih: "Camera Calibration with a Motorized Zoom Lens". *ICPR 2000*: 4495-4498
- [4] L. Robert. "Camera calibration without feature extraction", *IEEE Conference on Computer Vision & Image Processing: Pattern Recognition*, volume 1, pages 704--706, 1994.
- [5] Peter Eisert, "Model-Based Camera Calibration Using Analysis by Synthesis Techniques". *VMV 2002*: 307-314
- [6] Microsoft DirectX 9.0 SDK reference, August 2005.
- [7] Tom Miller, "Managed DirectX 9 kick start: graphics and game programming", Sams, 2003.
- [8] Holland, J.H., "Adaption in Natural and Artificial Systems", MIT Press, 1975.
- [9] Barry Laphorn, "A simple C# genetic algorithm", http://www.codeproject.com/cs/algorithms/btl_ga.asp.