

스마트폰에서의 3 차원 모델 검색을 위한 전처리 기법

손기원, 이승재, 이만희, 박인규
 인하대학교 정보통신공학부

{kiwonshon86, creative.sjlee}@gmail.com, maninara@hotmail.com, pik@inha.ac.kr

Preprocessing for 3D Model Retrieval on Smartphone

Ki Won Shon, Seung Jae Lee, Man Hee Lee, In Kyu Park
 School of Information and Communication Engineering, Inha University

요 약

본 논문에서는 스마트폰에서 한 장의 깊이 영상을 이용하여 내용기반 3 차원 모델 검색 기법을 범용적으로 수행하기 위한 깊이 영상 전처리 기법을 제안한다. 우선, 모바일 증강현실 플랫폼 중 하나인 Vuforia 를 이용하여 다수의 영상 취득과 동시에 스마트폰 카메라의 외부 인자들을 추정한다. 그 후 취득된 영상과 외부 인자들을 이용하여 Multi-View Stereo 기법을 통해 깊이 영상을 추정한다. 이렇게 추정된 깊이 영상은 3 차원 모델 검색 기법의 질의 깊이 영상으로 활용되어 3 차원 모델 검색 기법의 입력으로 활용할 수 있다. 실험 결과 본 논문에서 제안하는 전처리 기법으로 추정된 깊이 영상을 이용하였을 경우 3 차원 모델 검색의 범용성이 확장되고 3 차원 모델 검색이 올바르게 수행된 것을 확인하였다.

1. 서론

최근 애니메이션, 영화, 게임 및 웹 서비스 등의 발전으로 3 차원 모델 데이터의 수요가 급격히 증가하고 있고, 이에 따라 멀티미디어 및 엔터테인먼트를 비롯한 다양한 분야의 어플리케이션에서 3 차원 모델의 필요성이 점차 증가하고 있다.

한편 다양한 형태의 질의(query) 입력 기반 서비스와 제공되고 있는 2 차원 영상 검색과는 대조적으로, 3 차원 모델 검색 기법은 3 차원 모델의 시각적 형태 및 특징에 따른 검색이 불가능하기 때문에 Google 3D Warehouse [1] 와 같은 텍스트 기반의 검색 서비스만이 제공되고 있다. 이와 같은 문제를 해결하기 위하여 다양한 방법 [2] 이 연구 되었지만, 입력으로 사용되는 질의 깊이 영상의 취득을 위해 Kinect 를 사용해야 하므로 일반적인 환경에서 3 차원 모델을 검색할 수 있는 활용성이 감소하게 된다.

따라서 본 논문에서는 이러한 문제점을 해결하기 위해 최근 널리 사용되고 있는 스마트폰을 이용하여 깊이 영상을 취득할 수 있는 기법을 제안한다. 본 논문에서는 스마트폰 환경에 최적화된 증강현실 플랫폼인 Vuforia [3] 를 이용하여 다수의 영상을 카메라의 외부 인자들과 함께 취득하고, 이를 이용하여 Multi-View Stereo [4] 기법을 통해 깊이 영상을 추정하는 방법을 제안한다.

본 논문의 구성은 다음과 같다. 제 2 절에서는 증강현실을 사용하여 카메라 외부 인자를 구하는 법을 살펴본 후, 제 3 절에서 깊이 영상 생성 과정을 알아본다. 제 4 절에서는 앞에서 생성한 깊이 영상을 입력 질의 영상으로 하여 3 차원 검색을 하며, 마지막으로 제 5 절에서 본 논문에 대한 결론을 맺는다.

2. 증강현실을 이용한 카메라 외부인자 추정

일반적인 증강현실 기법을 이용할 경우 카메라에서 사용자가 지정한 패딘의 위치는 카메라에서의 model-view 행렬로 나타나게 된다. 본 논문에서는 이를 이용하여 카메라의 외부 인자 행렬을 추정하였다.

$$E = [R | t] = M' \quad (1)$$

즉, 증강현실 기법으로 구해진 model view 행렬 M 의 역행렬 M' 이 카메라 외부 인자 행렬 E 가 되는 것이다.

기존의 Zhang [5] 의 방법의 경우 격자영상을 이용하여 카메라 행렬을 추정하기 때문에 검색하고자 하는 물체에 의해 격자가 가려질 경우 정확한 카메라 행렬의 추정이 어려운 단점이 있다. 또한 5 점 및, 8 점 추정 방법을 사용할 경우



그림 1. 취득된 영상과 추정된 카메라 외부 인자

사용자가 지정한 패턴이 없어도 카메라의 외부 인자를 추정할 수 있지만, 이러한 방법의 경우 특징 검출 및 정합을 통하여 3 차원 정보를 추정하기 때문에 연산량이 증가하게 된다. 이는 하드웨어 성능이 상대적으로 부족한 스마트폰 환경에는 적합하지 않다. 반면 증강현실을 이용하여 사용자가 지정한 패턴을 찾는 방식은 카메라의 외부 인자를 좀더 정확하고 효율적으로 찾는 데 도움을 준다.

본 논문에서는 Qualcomm 에서 개발한 모바일 증강현실 플랫폼인 Vuforia 를 사용하였다. Vuforia 는 다양한 형태의 패턴을 추적할 수 있고, 일부분(40%)이 가려져도 증강현실 기능을 수행하는 강인함이 장점이다. 그림 1 은 스마트폰으로 취득한 5 장의 영상과 각 영상에 해당하는 카메라 외부 인자를 3 차원 공간에 나타낸 것을 보여주고 있다.

3. 깊이 영상 생성

2 절에서 취득한 4~5 장의 영상과 해당 영상에 해당하는 카메라 외부 인자 행렬을 이용하여, Multi-View Stereo[4] 기법을 이용하여 깊이 영상을 추정하였다. 본 논문에서 사용하는 3 차원 검색 기법은 사용자가 검색하고자 하는 물체만의 깊이 영상이 필요하므로, 배경은 사용자 입력 단계에서 GrabCut 알고리즘을 이용하여 분할하였다. 그림 2 는 5 장의 영상으로부터 추정된 깊이 영상을 보여준다.

4. 3 차원 모델 검색

3 절에서 생성된 깊이 영상은 한 장의 깊이 영상을 이용한 내용기반 3 차원 모델 검색 기법[2] 질의 영상으로 사용될 수 있다. 생성된 깊이 지도는 64x64 의 해상도로 축소되며 대부분의 잡음이 사라진다. 그림 3 은 자동차에 대한

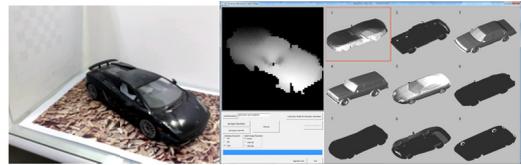


그림 3. 자동차에 대한 3 차원 모델 검색 결과

3D 모델 검색 결과를 보여준다.

5. 결론

본 논문에서는 3 차원 모델 검색 기법의 범용성 확장을 위해 널리 사용되고 있는 스마트폰 환경에서의 질의 깊이 영상을 취득하기 위한 전처리 기법에 대해 제안하였다. 본 논문에서는 증강현실을 이용하여 각 영상의 카메라 외부 인자를 추정하였고, 이를 이용하여 Multi-View Stereo 기법을 이용하여 깊이 영상을 취득하였다. 배경의 깊이 정보는 GrabCut 알고리즘을 이용하여 제거하였고, 생성된 깊이 영상의 잡음은 해상도가 축소되면서 대부분 사라지며 양질의 3 차원 모델 검색 기법의 입력 질의 깊이 영상으로 사용될 수 있게 된다. 실험 결과 본 논문에서 제안하는 전처리 기법으로 추정된 깊이 영상을 이용하였을 경우 3 차원 모델 검색의 범용성이 확장되고 3 차원 모델 검색이 올바르게 수행된 것을 확인하였다.

감사의 글

본 연구는 정보통신산업진흥원의 IT/SW 창의연구과정의 연구결과로 지식경제부와 (주)NHN 에 의해 지원된 과제로 수행되었음(NIPA-2012-(H0505-12-1003)).

참고문헌

- [1] Google 3D Warehouse, <http://sketchup.google.com/3dwarehouse>
- [2] M. S. Bae and I. K. Park, "Content-based 3D model retrieval using a single depth image from a low-cost 3D camera," *The Visual Computer*, vol. 29, no. 6-8, pp. 555-564, June 2013.
- [3] Qualcomm®, Vuforia™ <https://www.vuforia.com/>
- [4] J. Y. Chang, H. Park, I. K. Park, K. M. Lee, and S. U. Lee, "GPU-friendly multi-view stereo reconstruction using surfel representation and graph cuts," *Computer Vision and Image Understanding*, vol. 115, no. 5, pp. 620-634, May 2011.
- [5] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330-1334, 2000.

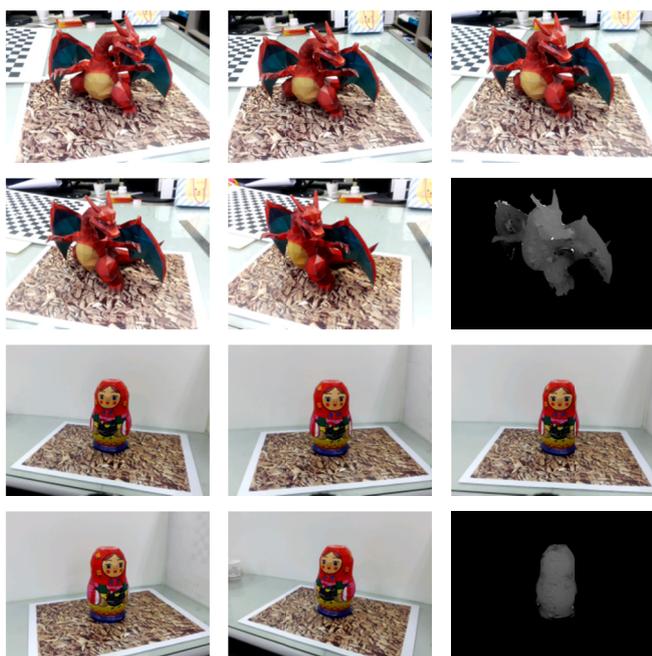


그림 2. 취득된 영상과 추정된 깊이 영상