

점진적 움직임 기반 구조를 위한 적응적인 영상 켄레 제거 방법

고재련 호요성

광주과학기술원 전기전자컴퓨터공학부

{jrko, hoyo}@gist.ac.kr

Adaptive Image Pair Resection for Incremental Structure from Motion

Jaeryun Ko and Yo-Sung Ho

Gwangju Institute of Science and Technology (GIST)

요약

점진적 움직임 기반 구조(Incremental Structure from Motion)는 다양한 시점에서 촬영한 영상들을 하나 씩 점진적으로 추가하여 3차원 장면을 복원하는 방법이다. 3차원 구조 복원에 사용되는 영상 켄레들 중에는 불필요한 켄레들도 충분히 포함되어 있으므로 복원된 구조의 불안정성과 불필요한 영상 켄레 처리로 인한 성능 손실이 발생할 수 있다. 이 논문은 상대적으로 불필요한 영상 켄레를 입력 영상 집합에 맞게 적응적으로 제거하는 방법을 제안한다. 대응점 탐색 단계에서 기하학적 검증 작업 전후로 총 두 번의 영상 켄레 제거가 실행되며, 통계적인 방법 및 기하학적으로 검증된 대응점 비율을 이용하여 문턱치를 결정한다. 실험 결과 3차원 복원 결과에 지장을 주지 않으면서 복원에 필요한 영상 켄레 개수를 효과적으로 줄일 수 있었다.

1. 서론

점진적 움직임 기반 구조(Incremental Structure from Motion) 방법은 다양한 시점에서 촬영한 영상을 하나씩 추가하여, 해당 영상을 촬영한 카메라의 위치를 추정하고 촬영된 장면을 3차원으로 복원하는 방법 중 하나이다. 이 논문은 3차원 복원을 위해 점진적 움직임 기반 구조에 사용될 다시점 영상들 중에서 중요성이 상대적으로 떨어지거나 큰 영향을 미치지 않는 영상 켄레들을 제거하는 방법을 제안한다. 대응점 탐색 단계에서 기하학적 검증 전후로 각각 다른 적응적 문턱치를 계산하여 두 번의 영상 켄레 제거를 수행했다. 첫 제거 과정에서는 영상의 특징점 정합 개수를 이용하였으며, 두 번째 제거 과정에서는 기하학적으로 검증된 정합 개수 비율을 계산하여 자동적으로 문턱치를 선정하여 제거를 실시했다. 이를 통해 생성된 3차원 점군 생성 실험을 진행하여 결과를 비교했다.

2. 점진적 움직임 기반 구조 파이프라인

점진적 움직임 기반 구조는 그림 1과 같이 하위 작업들이 순차적으로 처리되는 파이프라인 형태의 알고리즘이며 크게 대응점 탐색 단계 및 점진적 복원 단계 두 가지로 구분할 수 있다 [1].

대응점 탐색 단계는 입력된 다시점 영상 각각으로부터 크기와 회전 불변성을 가진 특징점을 추출하고 기술하는 작업으로 시작한다. 지역적 특징 기술자를 추출한 후 영상 간 대응되는 특징점 쌍을 획득하기 위한 정합을 실시한다. 이렇게 획득한 대응된 특징점 쌍은 실제 3차원 장면 상에서의 유효한 특징점 쌍이라는 것을 확인할 수 없으므로, 영상 간 평면 사영변환(Homography)를 추정하는 방법으로 각각의 특징점 쌍이 기하학적으로 정확하지 않음을 검증하는 과정을 거치게 된다 [2].

이후 획득한 영상 켄레들 중, 초기 3차원 복원 모델을 생성하기에 가장 적절한 켄레를 선택하여 복원한 후 점진적 복원 단계로 진입한다. 영상 정합 작업에서는 남은 영상들을 하나 씩 추가하여 상대적인 카메라 위치를 추정하고, 삼각화(Triangulation) 작업을 통해 추가된 영상으로 얻어진 특징점 쌍들을 이용하여 실제 3차원 구조에 대한 점군을 획득한다. 앞의 과정에서 누적된 오차들을 최소화하기 위한 방법으로 광속 조정(Bundle Adjustment) 작업을 실시하며, 이후 이상치 제거 등으로 잡음으로 취급되는 점군들을 제거한다. 모든 영상에 대하여 반복적으로 모든 과정이 수행되면 점진적 움직임 기반 구조 파이프라인은 종료된다.

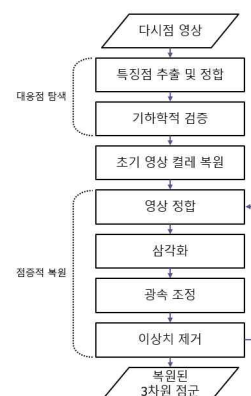


그림 1. 점진적 움직임 기반 구조 파이프라인의 순서도

3. 적응적인 영상 켄레 제거 방법

점진적 움직임 기반 구조 방법의 대응점 탐색 단계에서는 다시점 영상들로부터 충분히 많은 수의 영상 켄레들을 획득할 수 있다. 하지만

몇몇 영상 쉼체들은 3차원 복원에 큰 영향을 주지 않는 의미 없는 쉼체 인 경우들도 많으므로, 이러한 영상 쉼체들을 복원 과정에서 쓰지 않고 미리 제거하면 복원 속도 개선 및 복원된 3차원 구조의 안정성에 기여 할 수 있다.

그림 2는 제안한 적응적인 영상 쉼체 제거 방법의 흐름을 보여주며 그림 1에서의 대응점 탐색 단계를 정확히 대체할 수 있다. 먼저 통계적인 영상 쉼체 제거는 1차 제거 작업으로, 다시점 영상 쉼체들의 집합 $\mathbf{P} = \{p_1, p_2, \dots, p_N\}$ 에 대하여 각 영상 쉼체들이 가진 특징점 쌍의 개수 n_{p_i} 의 평균 $\mu_{\mathbf{P}}$ 와 표준편차 $\sigma_{\mathbf{P}}$ 을 구한다. 그리고 영상 쉼체 제거를 위한 문턱치 T_1 를 식 (1)와 같이 정의한다. 각각의 영상 쉼체들에 대한 n_{p_i} 가 문턱치 T_1 보다 작을 경우에 해당 영상 쉼체를 이 후 움직임 기반 구조 파이프라인에서 제외시킨다.

$$T_1 = \max(\mu_{\mathbf{P}} - 0.5\sigma_{\mathbf{P}}, 0) \quad (1)$$

1차 영상 쉼체 제거 이후 기하학적 검증을 통해 검증된 영상 쉼체 집합 $\mathbf{P}' = \{p'_1, p'_2, \dots, p'_G\}$ 를 획득한다. 각각의 영상 쉼체들은 기존의 특징점 쌍 개수 n_{p_i} , 뿐만 아니라 기하학적으로 검증된 특징점 쌍 개수 g_{p_i} 도 갖게 된다. 이를 이용하여 특징점 쌍으로부터 어느 정도 기하학적으로 검증되었는지에 대한 비율을 식 (2)과 같이 정의한다.

$$r_{p'_i} = g_{p'_i} / n_{p'_i} \quad (2)$$

또한 2차 영상 쉼체 제거 작업을 위한 문턱치 T_2 를 식 (3)에 따라 평균값과 중앙값 중 크기가 작은 값을 결정하여 각각의 영상 쉼체들에 대한 $r_{p'_i}$ 가 T_2 보다 작을 경우에 해당 영상 쉼체는 기하학적으로 중요성이 뒤떨어진다고 판단하여 제거한다.

$$T_2 = \min\left(1/G \sum_{i=1}^G r_{p'_i}, \text{median}(r_{p'_i})\right) \quad (3)$$

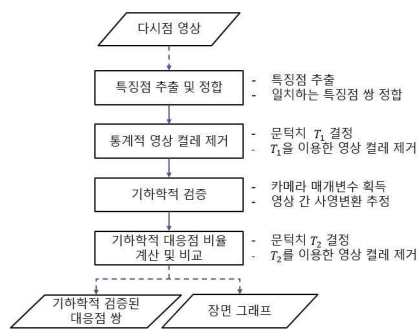


그림 2. 적응적인 영상 쉼체 제거 방법의 순서도

4. 3차원 복원 실험 결과

실험에 사용된 데이터로 99개의 다시점 영상으로 이루어진 영상 집합인 *Capitole*을 사용했다 [3]. 실험은 OpenMVG에서 제공되는 점진적 움직임 기반 구조 방법과 비교하여 진행되었다 [4]. 그림 3은 각

방법에서 사용된 영상 쉼체의 개수를 보여주는 삼각행렬로, 가로축과 세로축이 각각 영상의 지수(Index)를 의미한다. 제안한 방법은 영상 쉼체를 제거하기 때문에 기존의 기하학적 검증만 거친 방법보다 많은 수의 영상 쉼체가 제거된다. 하지만 그림 4와 같이 복원된 3차원 구조를 비교한 결과에서 유의미한 구조의 점군은 줄어든 영상 쉼체 개수에 비해 거의 차이가 없었다.

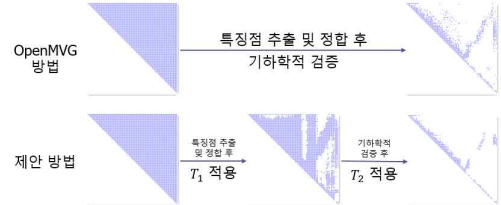


그림 3. 복원에 사용된 영상 쉼체 수의 시각적 비교

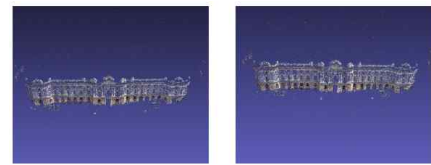


그림 4. 복원된 3차원 구조 결과 비교

5. 결론

이 논문은 점진적 움직임 기반 구조에서 상대적으로 복원에 불필요한 영상 쉼체들을 다수 제거하면서도 유의미한 3차원 복원 결과를 얻어낼 수 있는 적응적인 영상 쉼체 제거 방법을 제안했다. 첫 번째 쉼체 제거 과정에서는 해당 다시점 영상 집합으로부터 추출하여 대응된 특징점 쌍의 개수를 이용하여, 이 특징점 쌍 개수가 부족하다고 판단되는 영상 쉼체를 제거했다. 두 번째 쉼체 제거 과정에서는 특징점 쌍들 중 기하학적으로 검증된 쌍의 비율을 이용하여 문턱치 값을 결정하고 기준에 미달하는 영상 쉼체들을 한 번 더 제거했다. 이 결과 복원에 사용된 영상 쉼체의 수는 대폭 감소시키면서도 3차원 복원 구조의 결과 차이는 거의 없는 것을 확인했다.

감사의 글

이 논문은 2017년도 광주과학기술원의 재원으로 GRI(GIST연구원) 사업의 지원을 받아 수행된 연구임.

참고 문헌

- [1] J. Schonberger and J. Frahm, "Structure-from-motion Revisited," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4104-4113, 2016.
- [2] P. Moulon, P. Monasse and R. Marlet, "Adaptive structure from motion with a contrario model estimation," *Asian Conference on Computer Vision*, pp. 257-270, 2012.
- [3] A. Cohen, C. Zach, S. N. Sinha and M. Pollefeys, "Discovering and Exploiting 3D Symmetries in Structure from Motion," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1514-1521, 2012.
- [4] P. Moulon, P. Monasse and R. Marlet, "OpenMVG", <https://github.com/openMVG/openMVG>.