이진 특징 기술자의 군집화를 이용한 특징점 고속 정합

박정식 박종일1)

한양대학교

nangsik@mr.hanyang.ac.kr, jipark@hanyang.ac.kr

Fast keypoint matching using clustering of binary descriptors

Park, Jungsik Park, Jong-Il

Hanyang University

요약

이진 특징 기술자는 실수 벡터 형태의 특징 기술자보다 빠르게 특징점 추출 및 정합이 가능하고 메모리 공간도 적게 차지하는 장점이 있다. 하지만, 특징점의 수가 많아질수록 정합에 많은 시간이 소요되므로 실시간 처리가 중요한 객체 추적에 적용하기 위해서는 정합의 고속화 방법에 대한 연구가 필요하다. 이에 본 논문에서는 이진 특징 기술자의 군집화를 통한 특징점의 고속 정합 방법을 제안한다. 제안된 방법은 k-means 군집화 알고리즘을 기반으로 정합을 위한 기술자 탐색을 효과적으로 수행함으로써 군집화를 사용하지 않는 기존의 정합 방법에 비해 빠르면서도 높은 정확도를 유지한다.

1. 서론

특징 기술자는 영상의 국소 특징점 주변의 텍스쳐 정보를 표현함으로써 각각의 특징점을 구별할 수 있도록 해주는 정보이다. 특징 기술자의 대표적인 방법으로 SIFT(Scale-Invariant Feature Transform)와 SURF(Speeded-Up Robust Features)가 있다[1,2]. SIFT와 SURF는 추적 영상의 다양한 변화에 매우 강건하게 수행되므로 영상 인식, 객체 추적, 3차원 재구성 등 다양한 분야에 활용된다. 하지만, 이와 같은 특징 기술자를 사용할 경우 실수 데이터에 대한 연산을 주로 하기 때문에 연산 시간과 메모리 공간이 많이 필요하다. 따라서 보다 작은 자료형으로 구성된 특징 기술자에 대한 연구가 활발히 진행되어왔으며, 최근에는 비트열로 구성된 이진 특징 기술자인 BRIEF(Binary Robust Independent Elementary Features), BRISK(Binary Robust Invariant Scalable Keypoints), ORB(Oriented FAST, Rotated BRIEF)가 제안되었다[3,4,5].

BRIEF에서는 특징점 주변 패치에서 임의로 선택된 두 샘플 픽셀의 밝기값을 비교하여 그 결과를 이진 특징 정보로 이용하고, 반복적으로 수행한 이진 특징을 나열하여 비트열로 기술한다. ORB와 BRISK는 특징점 주변 패치에서 선택된 샘플 픽셀의 밝기값 비교 결과를 비트열로 기술한다는 점에서 BRIEF와 유사하지만, BRIEF의 회전 불변성을 제공하고 샘플링 패턴을 개선한 방법이다. 이러한 이진 특징 기술자는 실수 벡터로 구성된 특징 기술자에 비해 연산 속도도 빠르고 매우 작은 메모리 공간을 차지한다. 또한, 특징 기술자의 정합을 위해 해밍 거리(Hamming distance)를 이용하는데, 이는 XOR 비트연산으로 빠르게 수행할 수 있으며 SIMD(Single Instruction Multiple Data) 명령어를 사용할 경우 한 번의 명령으로 128개의 요소에 대한 XOR 연산이 가능하므로 더욱 빠르게 정합을 수행할 수 있다.

하지만, 이진 특징 기술자의 경우에도 특징점의 수가 많아지는 경우에는 정합에 많은 시간이 소요되므로 실시간 처리가 중요한 객체 추적에 적용할 경우에 어려움이 있다. 이러한 문제를 해결하기 위해서 이진 특징 기술자의 고속 정합 방법에 대한 연구가 진행되고 왔으며[6], 본 논문에서는 k-means 군집화를 이용하여 이진 특징 기술자의 고속 정합 방법을 제안한다. 제안하는 방법은 군집화를 사용하지 않는 기존의 정합 방법에 비해 빠르면서도 높은 정확도를 유지하며, 실험 결과를 통해 이를 검증한다.

2. 방법

가. 이진 특징 기술자에 대한 k-means 군집화

이진 특징 기술자는 해밍 공간상의 벡터로 표현될 수 있다. 해밍 공간상에서 k개의 벡터를 선택하여 각 군집의 대표 벡터로 지정하고, 특징 기술자 벡터와 대표 벡터의 해밍 거리를 계산하여 가장 가까운 군집에 소속시킨다. 여기서 군집에 적어도 하나 이상의 특징 기술자 벡터가 소속되도록 보장하기 위해서 k개의 대표 벡터는 특징 기술자 벡터 중에서 랜덤하게 선택한다. 특징 기술자 벡터가 모두 군집에 소속되면, 각 군집에 소속된 벡터의 평균 벡터를 계산하여 대표 벡터로 갱신한다. 이후로 각 군집에 소속된 특징 기술자 벡터를 모두 제거하고, 다시 특징 기술자 벡터를 새로운 대표 벡터 중에서 가까운 군집에 소속시키고 각 군집의 대표 벡터를 갱신하는 과정을 반복한다. 만약 모든 군집의 대표 벡터가 새로운 대표 벡터와 동일한 경우에 반복을 종료한다. 이 때, 각 군집에 속한 특징 기술자 벡터와 대표 벡터의 해밍 거리의 최대값을 각 군집의 반경으로 정의한다. 위의 군집화 과정은 기준

1) 교신저자.

영상에서 추출된 특징 기술자에 대해서만 수행된다.

나. 군집을 이용한 정합

정확한 정합을 위해서는 기준 특징 기술자와 입력 특징 기술자가 1:1 대응이 되도록 할 필요가 있다. 하지만 기준 특징 기술자에 대해서만 군집화를 수행할 경우 입력 특징 기술자의 대응점 탐색은 빠르게 수행되지만, 기준 특징 기술자의 대응점 탐색에 시간이 많이 소요된다. 따라서 기준 영상 특징 기술자에 대한 군집과 동일한 대표 벡터를 갖는 입력 특징 기술자에 대한 군집을 생성하여, 동일한 군집에 속하는 특징 기술자에 한해서 대응점 탐색을 시도함으로써 빠르게 정합한다. 이 때, 대표 벡터와 가장 가까운 하나의 군집에만 소속시킬 경우 특징 기술자가 잘못된 군집에 소속될 가능성이 있으므로 대표 벡터와의 해밍 거리가 군집의 반경보다 작은 모든 군집에 입력 특징 기술자를 소속시킨다. 그리고, NNDR(Nearest Neighbor Distance Ratio)를 적용하여 실제로는 대응점이 아니지만 최근접 이웃으로 판정되는 것을 방지한다.

3. 실험결과

실시간 추적 응용을 가정하여 기준 영상을 다양한 시점에서 촬영한 비디오 시퀀스와 기준 영상을 정합하는 실험을 수행하였다. 실험에 사용된 기준 영상 및 비디오 시퀀스의 일부가 그림 1에 나타나 있다. 기준 영상은 640×512의 해상도를 가지고, 비디오 시퀀스는 640×480의 해상도를 갖는 총 708프레임의 영상으로 구성되었다. 실험은 2.83 GHz 쿼드코어 CPU와 6 GB RAM을 가진 PC에서 수행되었다.

특징점 검출기는 FAST 코너 검출기[7]를 사용하였고 특징 기술자는 BRIEF를 사용하였다. 하지만, BRIEF는 회전 불변성 및 스케일 불변성이 없기 때문에, 특징점을 중심으로 반경 3픽셀 이내에서 계산한 평균 그래디언트 벡터의 방향을 특징점의 주방향으로 사용하고, SIFT에서 사용하는 것과 유사한 스케일 공간을 사용함으로써 회전 불변성과 스케일 불변성을 가지도록 하였다.

군집화를 사용한 정합 방법과 사용하지 않은 정합 방법을 비교하기 위하여 각각의 방법을 동일한 비디오 시퀀스에 적용하여 정합 소요 시간과 정합 정확도를 측정하였다(표 1, 그림 2). 여기서 정확도는 군집화를 사용하지 않은 정합 방법으로 구한 대응점 중에서 군집화를 사용한 정합으로 구한 대응점과 일치하는 비율을 나타내고, k는 군집의 개수이다. 기준 영상에서는 3258개의 특징점을, 입력 영상에서는 평균 200개의 특징점을 각각 검출하여 사용하였다. 군집화를 사용하지 않는 경우에 비해 군집화를 사용한 경우 높은 정합 정확도를 가지면서도 약 3~5배의 속도를 나타내는 것을 확인할 수 있었다.

그림 1. 실험에 사용된 기준 영상(좌) 및 비디오 시퀀스의 일부(우)

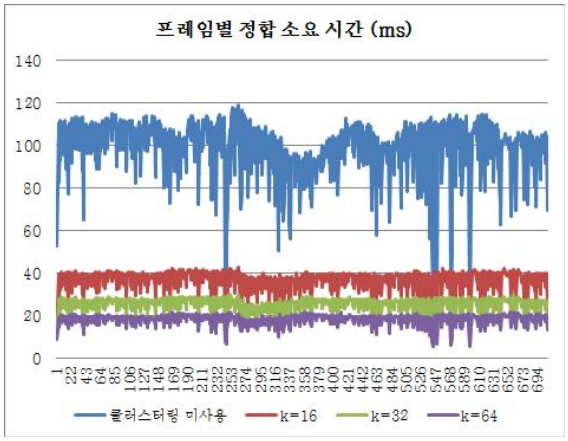

그림 2. 군집화를 사용하지 않은 정합 방법과 사용한 정합 방법의 프레임 별 정합 소요시간

표 1. 군집화를 사용하지 않은 정합 방법과 사용한 정합 방법의 평균 정합 소요 시간 및 정합 정확도

	정합 소요 시간	정합 정확도
군집화 미사용	98.69 ms	1.000
군집화 사용(k=16)	36.42 ms	0.987
군집화 사용(k=32)	25.32 ms	0.973
군집화 사용(k=64)	18.30 ms	0.949

4. 결론

본 논문에서는 이진 특징 기술자의 군집화를 통한 특징점의 고속 정합 방법을 제안하였다. 제안된 방법은 k-means 군집화 알고리즘에 기반한 정합 방법으로써 군집화를 사용하지 않는 기존의 정합 방법에 비해 빠르면서도 높은 정확도를 유지함을 확인하였다. 하지만 실제 응용에서의 특징점 추출과 특징 기술자 추출 등의 과정을 포함한 전체 과정의 처리 시간을 고려할 때 속도 개선이 더 이뤄져야 하며 동시에 정합 정확도가 떨어지지 않도록 최적화할 필요가 있다. 따라서 이진 특징 기술자의 고속 정합에 대해 더 많은 연구가 필요하다.

참고문헌

[1] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *IJCV*, vol. 60, no. 2, pp. 91-110, 2004.

[2] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Surf: Speeded-up robust features," *CVIU*, vol. 110, issues 3, pp. 346-359, 2008.

[3] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "BRIEF: Binary robust independent elementary features," *Proc. of ECCV*, 2010.

[4] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," *Proc. of ICCV*, pp. 2548-2555, 2011.

[5] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," *Proc. of ICCV*, pp. 2564-2571, 2011.

[6] M. Muja and D. G. Lowe, "Fast Matching of Binary Features," *Proc. of CRV*, pp.404-410, 2012

[7] E. Rosten and T. Drummond, "Machine Learning for High-Speed Corner Detection," *Proc. of ECCV*, 2006