

어파인 변환에 불변하는 지역 검출기에 대한 특징 기술자의 성능 평가

이만희, 박인규
 인하대학교 정보통신공학과
 {maninara@inha.edu, pik@inha.ac.kr}

Performance Evaluation of Local Descriptors for Affine Invariant Region Detector

Man Hee Lee and In Kyu Park

Department of Information and Communication Engineering, Inha University

요 약

본 논문에서는 어파인(affine) 변환에 불변하는 지역 검출기에 대하여 다양한 기술자의 성능을 비교하였다. 지난 수 년간 다양한 특징 기술자들이 연구되어 왔고, 이러한 특징 기술자들은 각각의 목적에 따라 상이한 특성을 갖고 있기 때문에 동일한 조건에서 다양한 기술자들의 성능을 비교하는 연구가 필요하다. 그러나 어파인 변환에 불변하는 지역 검출기에 대해 최적의 특징 기술자를 찾는 연구는 미흡한 실정이다. 따라서 본 논문에서는 지역적인 패치 기반의 특징 기술자뿐만 아니라 바이너리 기술자와 최근에 제안된 기술자들의 성능을 비교하였다. 제안하는 실험에서는 MSER (maximally stable extremal regions) 검출기를 이용하여 어파인 변환에 불변하는 지역을 검출하였고, 영상 확대 및 축소, 회전, 시점 변환 및 변형 가능한 물체에 대하여 각각 기술자의 성능을 비교하였다.

1. 서론

특징 추출 및 기술은 vSLAM (visual simultaneous localization and mapping)이나 SfM (structure from motion)과 같은 대부분의 컴퓨터 비전 분야에서 널리 사용되고 있다. 이를 위하여 SIFT (scale invariant feature transform) 혹은 SURF (speeded-up robust features)와 같은 다양한 특징 검출기 및 기술자들이 꾸준히 개발되어 왔고, 이러한 기술자들의 성능을 비교하는 연구도 진행되었다[1][2]. 한편, Harris affine 이나 Hessian affine, 또는 MSER (maximally stable extremal regions)과 같은 어파인 변환에 불변하는 지역 검출기도 제안되었다. 이러한 지역 검출기들은 영상에 큰 시점변화가 있을 경우 기존의 특징 검출기에 비하여 특징의 재검출율이 높기 때문에 그 사용범위가 증가하는 추세이다. 그러나 이러한 어파인 변환에 불변하는 지역 검출기의 경우 해당 검출기를 위한 특징 기술자가 존재하지 않기 때문에 단순한 패치 정합이나 기존의 기술자를 이용하여 특징 정합을 수행한다.

그러나 이와 같은 어파인 변환에 불변하는 지역 검출기의 경우 기존 기술자들의 성능평가가 이루어지지 않았기 때문에 적합한 특징 기술자를 선택하는데 어려움이 있다. 따라서 본 논문에서는 이러한 어파인 변환에 불변하는 지역 검출기에 대하여 다양한 기술자의 성능을 비교한다. 본 논문에서는 MSER 을 이용하여 어파인 변환에 불변하는 지역을 검출하였으며, 지역적인 패치 기반의 특징 기술자뿐만 아니라 BRIEF (binary robust independent elementary features)와 FREAK (fast retina keypoint)와 같은 바이너리 특징 기술자,

그리고 LIOP (local intensity order pattern)과 같은 최근의 특징 기술자들의 성능을 비교하였다. 또한 본 논문에서는 영상 확대 및 축소, 회전, 시점 변환 및 변형 가능한 물체에 대하여 각각 기술자의 성능을 비교하였다.

2. 성능 평가 방법

본 논문에서는 다양한 영상 환경에서 총 여섯 가지의 특징 기술자의 성능을 비교하였다. SIFT 기술자는 특징 패치를 4x4 개의 그리드로 나눈 뒤 각각의 그리드에서 다시 4x4 개의 서브 영역에 대하여 방향성을 구하고 각각의 방향성을 8 방향으로 정량화 한 뒤 방향의 히스토그램을 이용하여 특징을 기술한다. SURF 도 SIFT 와 유사한 방법으로 특징을 기술한다. 이 때, 적분 영상과 Harr 웨이블릿 필터를 이용하여 빠른 속도로 기술자를 생성할 수 있다. BRIEF 는 바이너리 기술자로서 특정 위치 사이의 밝기 차이를 이용하여 0 과 1 의 바이너리를 생성한다. 이 때, Gaussian 함수를 이용하여 잡음에 민감하게 반응하는 것을 방지하고 Gaussian 분포의 표본 추출을 이용하여 기술자를 생성한다. 또한 고속의 기술자 비교를 위하여 Hamming 거리를 이용한다. FREAK 역시 바이너리 기술자로서 사람의 망막을 묘사한 표본 추출 방법을 이용하여 기술자를 생성한다. 이 때, 각각의 표본 위치는 각각 다른 Gaussian 함수를 이용하여 부드럽게 변환한다. Shape descriptor 는 MSER 을 위하여 제안된 기술자로서 MSER 로 찾아진 영역의 외형을 이용하여 기술자를 생성한다. 이 때, 기술자가 계산되는 방법은 SIFT 와 유사하다. LIOP 는 영상의 회전과 밝기 변화에 강인한 기술자로서 가장 최근에 제안된

기법이다. 우선 패치 내부의 모든 밝기 값을 순서대로 정렬한 뒤 몇 개의 집단으로 구분한다. 그 후 각각의 집단에 대하여 집단 내부의 모든 위치에서 주변에 대한 밝기의 상관관계를 이용하여 LIOP 값을 결정하고 이것의 히스토그램을 이용하여 기술자를 생성한다.

본 논문에서는 MSER 로 찾아진 타원 형태의 영역을 우선 고정된 크기의 원으로 정규화 한다. 그 후 해당 패치의 강인한 방향성을 이용하여 모든 방향을 일치시킨 뒤 위에서 언급한 각각의 기술자들을 생성한다. 특징의 정합을 위하여 본 논문에서는 nearest neighbor thresholding 기법을 이용하였다. 이 기법은 하나의 특징에 대하여 비교하고자 하는 모든 특징과의 기술자 거리를 구하고 가장 거리가 작은 기술자를 정합 쌍으로 결정하는 방법이다.

3. 실험 결과

위에서 언급한 모든 기술자들은 각 저자에 의해서 공개된 코드를 사용하거나 OpenCV 와 VLFeat 라이브러리를 이용하여 구현되었다. 각각의 패치는 16 픽셀의 가장자리를 포함하여 144x144 크기를 갖는다. 예외적으로 BRIEF 의 경우 같은 크기의 가장자리를 포함하여 82x82 의 패치 크기를 이용하였다. 또한 모든 실험은 Intel Core i7 2.7GHz 의 CPU 와 16GB 메모리의 환경에서 측정되었다.

그림 1 은 다양한 영상 환경에 대하여 각각의 기술자들의 성능을 비교한 그래프를 보여주고 있다. 그림 1 에서 가로축은 1-precision 을, 세로축은 recall 을 나타낸다. 그림 1 에서 보여주는 바와 같이 대부분의 경우 LIOP 가 좋은 성능을 나타내는 것을 확인할 수 있다.

표 1 은 각각의 기술자들의 수행시간을 나타낸다. 표 1 에서 확인할 수 있는 바와 같이 바이너리 기술자의 경우 고속의 기술자 생성과 정합을 수행할 수 있는 것을 알 수 있다. 그러나 가장 좋은 성능을 보여준 LIOP 의 경우 기술자 생성에 다른 기술자들에 비하여 20 배 이상의 수행시간을 필요로 하는 것을 알 수 있다.

4. 결론

본 논문에서는 어파인 변환에 불변하는 지역 검출기에 대하여 다양한 기술자의 성능을 비교하였다. MSER 지역 검출기를 이용하여 검출된 특징에 대하여 SIFT, SURF, BRIEF, FREAK, Shape descriptor, 그리고 LIOP 기술자의 성능을 비교하였다.

표 1. 500x500 정합의 수행 시간 비교 (ms).

| 기술자 종류 (차원) | 기술자 생성 | 기술자 정합 |
|------------------------|--------|--------|
| SIFT (128) | 2,189 | 142 |
| SURF (128) | 388 | 157 |
| BRIEF (256) | 43 | 84 |
| FREAK (512) | 84 | 163 |
| Shape descriptor (128) | 2,888 | 112 |
| LIOP (255) | 42,827 | 267 |

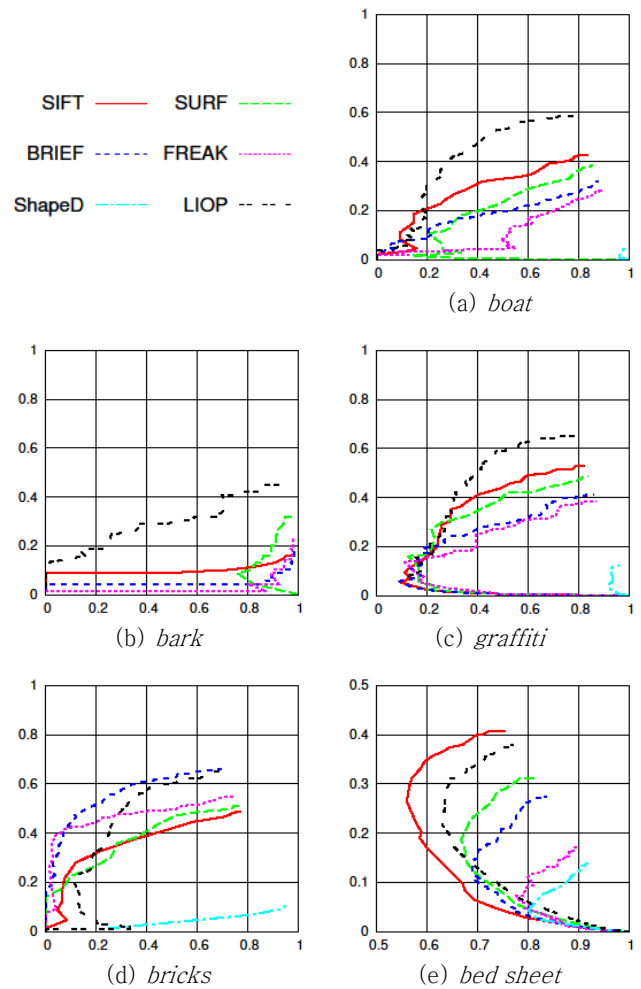


그림 1. 다양한 영상 환경에 대한 각각의 기술자들의 성능 비교. 가로축은 1-precision 을, 세로축은 recall 을 의미. (a)(b) 영상 확대 및 축소, 회전. (c)(d) 시점 변환. (e) 변형 가능한 물체.

다양한 영상 환경에서의 실험 결과 가장 최근에 제안된 기술자인 LIOP 가 대부분의 실험에서 좋은 성능을 보여주었다. 그러나 높은 연산량으로 인해 실시간 활용에는 어려움이 있어 다른 가속화 방법을 적용하거나 바이너리 기술자를 사용할 필요가 있다.

감사의 글

본 연구는 미래창조과학부 및 한국산업기술평가관리원의 산업융합원천기술 개발사업(정보통신)의 일환으로 수행하였음. [10047078, 다시점 블랙박스 영상을 이용한 교통 사고현장 3 차원 재구성 기술개발]

참고문헌

[1] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," IEEE TPAMI, vol. 27, no. 10, pp. 1615-1630, October 2005.
 [2] K. Mikolajczyk, et al., "A comparison of affine region detectors," IJCV, vol. 65, no. 1-2, pp. 43-72, November 2005.