

적분 투영을 이용한 특징 기술 및 정합

이만희, 박인규

인하대학교 정보통신공학과

maninara@inha.edu, pik@inha.ac.kr

초록: 본 논문에서는 움직임에 의한 블러가 존재하는 영상에서 특징 정합 성능을 높이기 위하여 네 방향의 적분 투영을 이용한 새로운 특징 기술자를 제안한다. 제안하는 특징 기술자는 0 도, 45 도, 90 도, 그리고 135 도의 네 방향에 대한 적분 투영을 이용한다. 실험 결과 제안하는 기술자의 경우 움직임에 의한 블러 뿐만 아니라 카메라 초점에 의한 블러 및 영상 회전 과 밝기 변화에도 강인한 성능을 보여주는 것을 확인할 수 있다.

주제어: feature descriptor, integral projection, motion blur

I. 서론

특징 기술 및 정합은 다양한 컴퓨터 비전 분야에서 가장 기본적으로 수행되는 선행 알고리즘으로 알려져 있다. 이를 위하여 기존에 다양한 특징 기술 기법들이 제안되어 왔고 널리 활용되어 왔다. 이러한 기존의 기술자들의 경우 영상 회전 및 크기 변화에 강인하도록 설계되었다. 그에 따라 영상에 움직임에 의한 블러가 존재할 때 기존의 기술자로 특징을 정합할 경우 많은 어려움이 따른다. 따라서 본 논문에서는 움직임에 의한 블러가 존재하는 영상에서 특징 정합의 성능을 높이기 위한 새로운 특징 기술자를 제안한다.

II. 적분 투영

본 논문에서는 적분 투영을 이용한 새로운 특징 기술자를 제안한다. 일반적으로 적분 투영의 경우 영상에서 얼굴이나 눈의 경계를 찾거나 동영상에서 카메라의 전역적인 움직임을 찾는데 널리 사용되어 왔다. 본 논문에서는 0도, 45도, 90도 그리고 135도 네 방향의 적분 투영을 이용하여 특징을 기술한다. 이산 영상에서 가로 방향의 적분 투영 함수 IP_h 는 식 (1)과 같이 정의된다.

$$IP_h(y) = \sum_{x=x_1}^{x_2} f(x, y) \quad (1)$$

적분 투영은 식 (1)에서 보는 바와 같이 임의의 범위 x_1 과 x_2 사이의 모든 함수 값 $f(x, y)$ 의 합을 의미한다.

한편, 컴퓨터 비전 분야에서 일반적인 블러의 경우 원본영상과 PSF (point spread function)간의 합성곱으

로 표현된다. 이 때, 선형 움직임에 대한 PSF의 경우 직선으로 모델링 되고 이러한 PSF에 의해 생성된 블러 영상에서 움직임의 방향과 동일한 방향의 적분 투영은 블러가 없는 영상에서의 적분 투영과 일치하게 된다. 따라서, 이와 같은 적분 투영을 이용하여 특징을 기술할 경우 선형적인 움직임에 의한 블러에 강인한 특성을 갖게 된다.

III. 제안하는 기술자 생성 과정

본 논문에서 제안하는 기술자를 생성하기 위하여 우선 영상에서 찾아진 특징은 일정한 크기의 패치로 정규화 된다. 그림 1(a)는 본 논문에서 제안하는 패치의 크기를 보여준다. 그 후, 정규화 된 패치로부터 0도, 45도, 90도, 그리고 135도의 적분 투영을 계산한다. 이 때, 45도와 135도 대각선 방향의 적분 투영을 계산할 경우 연산의 효율을 위하여 그림 1(b)와 같은 샘플링 패턴을 이용하여 고속으로 적분 투영을 계산한다. 그림 1(b)에서 빨간색과 녹색은 각각 홀수와 짝수 번째 순서의 샘플 위치를 의미한다. 또한 그림 1(a)에서 보는 바와 같이 가로와 세로방향뿐만 아니라 대각선 방향에서도 적분 투영에 사용되는 샘플의 개수가 다르기 때문에 각각의 적분 투영 값은 해당 적분 투영의 샘플 개수를 이용하여 정규화한다. 결과적으로 하나의 방향 마다 32 차원의 적분 투영 값이 계산되고 최종적으로 네 방향의 128 차원을 갖는 기술자를 생성할 수 있다.

제안하는 기술자가 밝기 변화에 강인한 특성을 갖기 위하여 네 개의 적분 투영에서 각각 적분 투영의 중심 값을 기준으로 값의 차이를 이용하여 기술자를 생성한다. 또한 해당 차이의 최대와 최소를 이용하여 모든

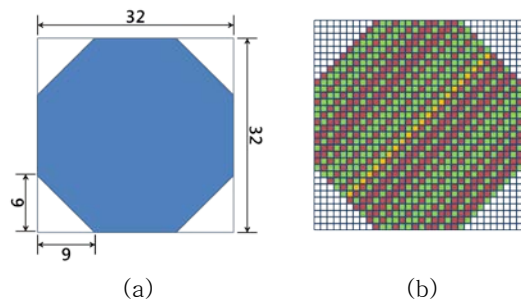


그림 1. 제안하는 특징 기술자 패치. (a) 패치 크기. (b) 대각선 방향의 샘플링 패턴.

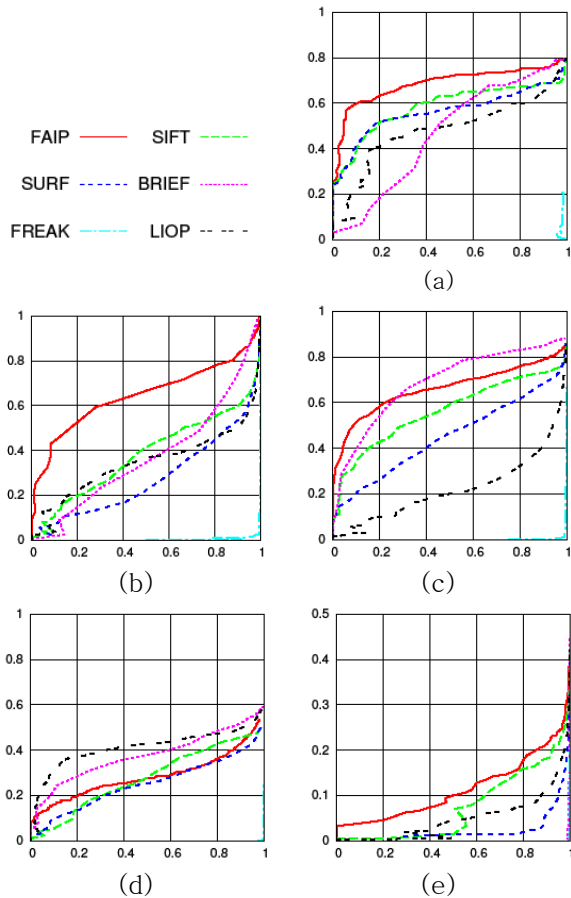


그림 2. 다양한 영상 환경에 대한 각각의 기술자들의 성능 비교. 가로축과 세로축은 1-precision과 recall의 의미. (a)(b) 움직임에 의한 블러. (c) 카메라 초점에 의한 블러. (d) 밝기 변화. (e) 영상 회전.

적분 투영 값을 정규화 함으로써 급격한 밝기 변화에서도 유사한 기술자를 생성할 수 있다.

한편 기존 기술자들의 경우 영상에 블러가 존재할 때 특징의 표준 방향을 정확하게 추정하지 못하여 정합 성능이 저하되는 경향이 있다. 따라서 제안하는 기술자의 경우 특징 패치에서 표준 방향을 추정하지 않고 네 개의 적분 투영 중 가장 편차가 큰 하나의 적분 투영을 선택하고 해당 적분 투영을 기준으로 나머지 적분 투영을 배열함으로써 영상의 회전에 강인한 특성을 가질 수 있다. 이 때, 각각의 적분 투영 값은 LO 기울기 최소화를 통하여 잡음에 민감하게 반응하는 것을 방지하였다.

표 1. 각각의 기술자에 대한 수행시간 비교.

기술자 종류 (차원)	기술자 생성 시간 (ms)
제안하는 기법 (128)	108
SIFT (128)	2,189
SURF (128)	388
BRIEF (256)	43
FREAK (512)	84
LIOP (255)	3,221

뿐만 아니라 제안하는 기술자의 경우 네 방향의 적분 투영을 이용하여 특징을 기술하기 때문에 영상의 잡음에도 강인한 특성을 갖게 된다.

IV. 실험 결과 및 결론

본 논문에서는 움직임에 의한 블러가 존재하는 영상뿐만 아니라 카메라 초점에 의한 블러 및 밝기 변화와 영상 회전에 대한 데이터 셋을 이용하여 실험을 진행하였다. 제안하는 기술자의 성능을 측정하기 위하여 SIFT[1], SURF[2], BRIEF[3], FREAK[4], LIOP[5] 기술자와 성능을 비교하였다. 그림 2에서 보는 바와 같이 제안하는 기술자는 움직임에 의한 블러가 있는 영상뿐만 아니라 카메라 초점에 의한 블러와 영상 밝기 변화, 그리고 영상 회전에 대해서도 좋은 성능을 나타내는 것을 확인할 수 있다.

표 1은 1,000개의 특징에 대한 기술자를 생성하기 위하여 각각의 기술자마다 필요한 수행 시간을 보여주고 있다. 표 1에서 보는 바와 같이 제안하는 기술자의 경우 바이너리 기술자들에 비하여 좀 더 많은 수행시간을 필요로 하지만 패치 기반의 기술자들에 비하여 많은 향상이 있는 것을 확인할 수 있다.

감사의 글

본 연구는 미래창조과학부 및 정보통신기술진흥센터의 정보통신·방송 연구개발사업의 일환으로 수행하였음. [2013(10047078), 다시점 블랙박스 영상을 이용한 교통 사고현장 3차원 재구성 기술개발]

참고문헌

- [1] D. G. Low, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, November 2004.
- [2] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-up robust features," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346-359, June 2008.
- [3] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "BRIEF: Binary robust independent elementary features," *Proc. of European Conference on Computer Vision*, pp. 778-792, September 2010.
- [4] A. Alahi, R. Ortiz, and P. Vandergheynst, "FREAK: Fast retina key-point," *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 510-517, June 2012.
- [5] Z. Wnag, B. Fan, and F. Wu, "Local intensity order pattern for feature description," *Proc. of IEEE International Conference on Computer Vision*, pp. 603-610, November 2011.