

## SURF(speed up robust feature)를 이용한 시점변화에 강인한 영상 매칭

손종인 강민성 손광훈

연세대학교 전기전자공학과

khsohn@yonsei.ac.kr

## View invariant image matching using SURF

Son, Jong-In Kang, Minsung Sohn, Kwanghoon

School of Electrical and Electronic Engineering, Yonsei University

## 요약

영상 매칭은 컴퓨터 비전에서 중요한 기초 기술 중에 하나이다. 하지만 스케일, 회전, 조명, 시점변화에 강인한 대응점을 찾는 것은 쉬운 작업이 아니다. 이러한 문제점을 보완하기 위해서 스케일 불변 특징 변환(Scale Invariant Feature Transform) 고속의 강인한 특징 추출(Speeded up robust features) 알고리즘들에 제안되었지만, 시점 변화에 있어서 취약한 문제점을 나타냈다. 본 논문에서는 이런 문제점을 해결하기 위해서 시점 변화에 강인한 알고리즘을 제안하였다. 시점 변화에 강인한 영상매칭을 위해서 원본 영상과 질의 영상간 유사도 높은 특징점들의 호모그래피 변환을 이용해서 질의 영상을 원본 영상과 유사하게 보정된 뒤에 매칭을 통해서 시점 변화에 강인한 알고리즘을 구현하였다. 시점이 변화된 여러 영상을 통해서 기존 SIFT,SURF와 성능과 수행 시간을 비교 함으로서, 본 논문에서 제안한 알고리즘의 우수성을 입증 하였다.

## 1. 서론

컴퓨터 비전 에서 서로 다른 두 영상에 대한 유사도를 측정 매칭하는 것은 3D reconstruction, 이미지 스티칭, 로봇 네비 게이션, 물체 인식, 얼굴 인식등에 응용될 수 있는 기초적인 기술이다.

영상 매칭을 위해서 2가지 과정이 필요하다. 첫째는, 영상에서 특징 점을 추출하는 것으로, 주요 방법으로는 헤리스 코너 검출기, 헤이시안 검출기, 고속 헤이시안 검출기 검출기 등이 있다. 둘째는, 특징점을 기준으로 표현자(descriptor)를 생성하는 것이다. 이러한 영상 매칭을 위한 특징점 추출과 표현자를 생성하는 알고리즘에 관한 연구<sup>[1][4]</sup>에 대표적인 것이 Lowe의 스케일 불변 특징 변환(Scale Invariant Feature Transform)<sup>[1]</sup> 알고리즘과 Bay의 고속의 강인한 특징 추출(Speeded up robust features) 알고리즘<sup>[5]</sup>이 있다.

헤리스 코너 검출기,헤이시안 검출기, 고속 헤이시안 검출기 검출기는 특징점을 찾는 방법으로 많이 이용되지만 크기 변화와 시점 변화에 취약하다. 이런 점을 보완한 SIFT,SURF 알고리즘은 영상의 크기, 회전 변환에 불변하는 특징점을 찾고 표현자를 구성하는 공통점을 갖는다. 하지만 시점 변화된 같은 물체나 장소에 대한 이미지에 대해서는 매칭율이 현저히 낮아지는 단점이 있다.

본 논문에서는 이러한 문제점을 해결하고 회전, 크기, 시점 변화에서도 강인한 이미지 매칭 방법을 제안한다.

2장에서는 기존 특징점 추출과 표현자를 사용하는 알고리즘에 대한 방법들과 호모그래피에 대한 소개 한다. 3장에서는 시점변화에 강인한 알고리즘을 제안한다. 4장에서는 기존 방법들(SIFT,SURF)과 신뢰도와 수행시간에 대해서 비교하고, 마지막으로 5장에서는 이 논문의 결론을 도출한다.

## 2. 영상의 특징점 추출 및 표현자 구성 방법 및 호모그래피 계산

## 2.1 특징점 추출 및 표현자 구성

영상 매칭을 위한 절차는 특징점 추출과 표현자 구성의 두 가지 단계로 구분된다. 특징점 추출을 위한 대표적인 방법들로는 헤리스 코너 검

출기<sup>[8]</sup>,헤이시안 검출기<sup>[4]</sup>, 고속 헤이시안 검출기<sup>[5]</sup>등이 있다.

헤이시안 검출기는 헤이시안 행렬 (1)을 이용하여 특징점을 추출하며 얼룩(Blob)과 능선(Ridge)검출에 강인하다.

$$H(x,y,\sigma) = \begin{bmatrix} LI_{xx}(x,y,\sigma)LI_{xy}(x,y,\sigma) \\ LI_{xy}(x,y,\sigma)LI_{yy}(x,y,\sigma) \end{bmatrix} \quad (1)$$

헤리스 코너 검출기는 코너점이 높은 양방향 곡률을 가지는 특성 이용하여 고유값과 코너응답함수를 통해 특징점을 찾는다. SURF에서 사용한 고속 헤이시안 검출기는 적분영상과 가우시안 2차 미분을 근사화 한 사각 필터를 이용하여 수행시간을 단축하였다

둘째로, 이렇게 찾은 특징점들을 표현자로 구성하는 방법으로는 SURF와 SIFT알고리즘이 있다. SIFT는 특징점을 중심으로  $4 \times 4$ 의 세부영역을 구성하고 그 세부 영역에서의 픽셀들의 기울기 방향과 히스토그램을 벡터로 구성하였다. SUFT는 특징점을 중심으로  $4 \times 4$ 의 세부영역을 구성하고 각 세부영역에서 Haar 웨이블릿을 사용하여 4,8개의 특징을 구함에 따라서 64, 128차원의 표현자 벡터를 구성하였고, 128차원 벡터를 사용할 경우 회전 영상 매칭에 강인하다.

## 2.2 Homography 계산 및 RANSAC

Homography는 왜곡된 영상을 보정하는 방법 중 하나<sup>[9]</sup>이며, 서로 대응되는 영상간에 2차원 Projective Mapping을 정의하는  $3 \times 3$  행렬이다. 어느 한 픽셀 $(x_i, y_i)$ 가 다른 영상에서  $(x_i, y_i)$  서 나타났고, 이 픽셀에 할당된 homography가 H라고 한다면,(2)의 관계가 성립한다.

$$H \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} wx_i \\ wy_i \\ w \end{bmatrix} \quad (2)$$

Homography를 통해서 영상을 보정하기 위해서, 8 DOF(Degree of Freedom)이 필요한데, 이 때 8 DOF는 서로 대응되는 4쌍의 점을 통해

구할 수 있다. 단, 이 4쌍의 점들 중에서 동일 선상에 최소 3개의 점이 있으면 안된다.

이러한 관계를 이용하여 본 논문에서는 원본 영상과 질의 영상 간에 SURF를 이용하여 가장 유사도가 높은 10쌍의 점을 찾아낸다.

RANSAC(Random Sample Consensus 이하 RANSAC),알고리즘은 거짓정보( outlier)를 포함하는 일련의 데이터 집합으로부터 수학적 모델의 인자들을 반복적인 작업으로 예측하는 기법<sup>[10]</sup>이다. 이 RANSAC 알고리즘을 통해 10쌍중에 거짓정보를 제거하고, Homography 행렬을 이용하여,질의 영상을 변환 하여 원본 영상과 유사한 시점으로 보정된 질의 영상을 구할 수 있다.

### 3. 제안하는 알고리즘

SURF는 안정적인 속도와 회전, 크기 변화에 강인한 이미지 매칭 알고리즘 이다. 하지만 시점변화에서는 낮은 매칭율을 보인다.

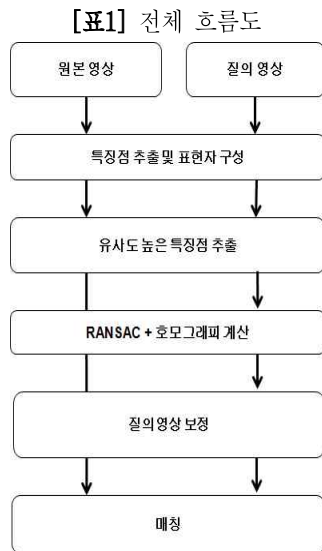
예를 들어서 그림1의 (a)와 (b) 경우 서로 같은 영상을 다른 시점에서 촬영한 영상이지만, SIFT와 SURF알고리즘을 사용할 경우 낮은 매칭율을 보인다.



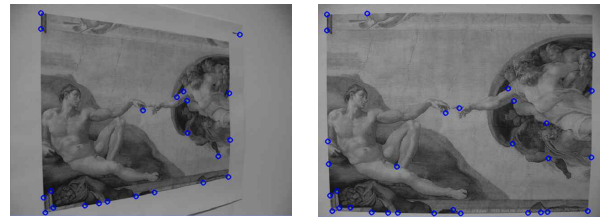
(a) 정면 영상 (b) 시점이 변화된 영상

[그림 1] 실험에 사용된 영상

이러한 단점을 보완하기 위해, 다음과 같은 알고리즘을 제안하였다.



원본 영상과 질의 영상을 헤리스 코너 검출기를 통해서 최대 100개의 특징점을 찾아낸다. 그림 2의 (a),(b)는 adam 원본과 질의 영상, (c),(d)는 graffiti 원본과 질의 영상에 대한 헤리스 코너 검출 결과를 보인다.



(a) adam 원본 영상 헤리스 코너 검출기 특징점 추출 (b) adam 질의 영상 헤리스 코너 검출기 특징점 추출

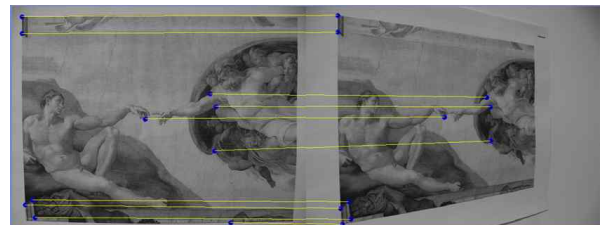


(c) graffiti 원본 영상 헤리스 코너 검출기 특징점 추출 (d) graffiti 질의 영상 헤리스 코너 검출기 특징점 추출

[그림 2] 특징점 추출

이렇게 얻은 헤리스 코너 검출기의 특징점들을 SURF와 동일한 방법으로 descriptor 이용하여 최소 4개에서 최대 10개의 유사도 높은 매칭쌍을 구해준다. 그림 3의 (a),(b)는 10개의 유사도 높은 특징점들을 매칭한 영상이다.

이 매칭된 특징점 간의  $3 \times 3$  Homography 행렬을 구해준다.



(a) adam영상 유사도 높은 특징점 추출



(b) graffiti영상 유사도 높은 특징점 추출

[그림 3] 유사도 높은 특징점 추출

원본 영상과 질의 영상간 유사도 높은 매칭점들을 RANSAC을 통해서 보다 정확한  $3 \times 3$  Homography Matrix를 구한다. 질의 영상을 계산된 Homography Matrix를 통해 변환하여 보정된 질의 영상을 구한다. 그림 4는 (a),(b)는 원본 영상을 (c),(d)는 질의 영상 (e),(f)는 질의 영상을 homography 변환된 보정된 영상을 보여준다.



[그림 4] 원본,질의,보정영상

원본 영상의 특징점과 표현자를 보정된 질의 영상의 특징점과 표현자를 통해서 다시 매칭시켜준다. 최종적으로 원본 영상과 질의 영상사이의 SURF매칭을 통해서 시점변화에 강인한 영상 매칭을 수행할 수 있다.

#### 4. 실험

제안하는 시점변화에 강인한 영상매칭 알고리즘은 C++언어로 구현하였으며, 인텔 i7-3기가 클럭 CPU와 2GB 메모리의 컴퓨터 환경에서 실험하였다.

실험에 사용된 영상은 K. Mikolajczyk가 공개한 database<sup>[11]</sup>의 adam 영상과 graffiti영상을 사용하였으며, 이 영상은 다 각도 시점변화 영상이 포함되어 있고, 영상 사이즈는 640×480 이다.

제안한 알고리즘의 성능<sup>[6]</sup>을 평가하기 위해 SIFT, SURF알고리즘과 수행시간을 측정하여 비교하였다.

또한 SIFT, SURF알고리즘과 영상간의 매칭개수를 비교하였다. 실험에 사용한 SIFT 알고리즘은 Rob Hess가 공개한 소스코드를 실험에 사용하였다.

수행시간	SIFT	SURF	제안한 알고리즘
graffiti	8.12(sec)	1.79(sec)	1.88(sec)
adam	3(sec)	0.31(sec)	0.43(sec)

[표2] 수행시간 비교

제안한 알고리즘은 SIFT에 비해서 graffiti영상은 약4.3배, adam영상은 약 7배 빠른 결과를 보이며, 기존 SURF에 비해서 크게 느리지 않는 것을 확인 할 수 있었다.

매칭개수	SIFT	SURF	제안한 알고리즘
graffiti	33	20	154
adam	1	2	50

[표3] 매칭개수 비교

제안한 알고리즘은 SIFT에 비해서 graffiti영상은 약 5배, adam영상은 약 50배 많은 매칭개수를, SURF에 비해서는 graffiti영상은 약 7.5배, adam영상은 약 25배 많은 매칭개수를 얻을 수 있었다.



(c) 제안한 알고리즘을 이용한 adam, graffiti 매칭 결과  
[그림 5] SIFT,SURF,제안한 방법 결과비교

## 5. 결론

본 논문에서는 기존 영상 매칭 알고리즘의 시점변화에 취약한 문제점을 해결하고, 시점변화에 강인한 영상 매칭 알고리즘을 제안하였다. 하지만 시점변화가  $90^\circ$  이상인 영상에서는 기존 방법과 비슷한 성능을 보인다. 향후에는 성능을 개선하기 위해 정밀한 매칭점 매칭 및 실시간 영상 매칭 알고리즘에 대한 연구를 진행 할 예정이다.

## 참고문헌

- [1] J.M. Morel and G.Yu, ASIFT : A New Framework for Fully Affine Invariant Image Comparison, SIAM Journal on Imaging Sciences, vol. 2, issue 2, 2009.
- [2] K. Mikolajczyk and C. Schmid, "Scale and Affine invariant interest point detectors". In IJC V 60(1):63-86, 2004. (Harris-Affine & Hessian Affine)
- [3] J.Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions". In BMVC p. 384-393, 2002 (MSER)
- [4] T. Tuytelaars and K. Mikolajczyk , "Local Invariant Feature Detectors" - Survey. In CVG, 3(1):1-110, 2008.
- [5] H. Bay, T. Tuytelaars, and L. V. Gool, "Surf: Speeded up robust features," European Conference on Computer Vision Vol. 3951.pp.404-417,2006
- [6] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.27, no. 10, pp. 1615-1630, 2005.
- [7] D. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," Int'l J. Computer Vision, Vol.60, no. 2, pp. 91-110, 2004.
- [8] C. Harris and M. Stephens, "A Combined Corner and Edge Detector," Proc. Alvey Vision Conf.pp.147-151, 1988.
- [9] Richard Hartley and Andrew Zisserman "Multiple View Geometry in computer vision". Cambridge University 2003.
- [10] M. A. Fischler, R. C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Comm. of the ACM, Vol 24, pp 381-395, 1981.
- [11] <http://www.robots.ox.ac.uk/~vgg/research/affine/>