

시점 변화에 강인한 특징점 정합 기법

정 현 조*, 유 지 상^o

Feature Matching Algorithm Robust To Viewpoint Change

Hyun-jo Jung*, Ji-sang Yoo^o

요 약

본 논문에서는 FAST(Features from Accelerated Segment Test) 특징점 검출기와 SIFT(Scale Invariant Feature Transform) 특징점 서술자(descriptor)를 사용하여 시점 변화에 강인한 특징점 정합 기법을 제안한다. 기존의 FAST 기법은 영상의 에지 부분을 따라서 불필요하게 특징점을 많이 추출하게 되는데 이러한 단점을 주곡률(principal curvatures)을 적용하여 개선한다. 추출된 특징점을 SIFT 서술자를 통해 기술하고 시점이 다른 두 영상으로부터 구해진 정합쌍에 RANSAC(RANdom SAMple Consensus) 기법을 통하여 호모그래피(homography)를 계산한다. 시점 변화에 강인한 특징점 정합을 위해서 기준 영상의 특징점들을 호모그래피 변환을 통해 변경된 좌표와 시점이 다른 영상의 특징점 좌표간의 유클리드(Euclidean) 거리를 통해 정합쌍을 분류한다. 같은 물체나 장소에 대해 시점이 변화된 여러 영상에 대한 실험을 통해서 제안하는 정합 기법이 적은 계산량으로 기존의 특징점 정합 기법보다 우수한 성능을 보여주는 것을 확인하였다.

Key Words : viewpoint change, homography transform, Mashing

ABSTRACT

In this paper, we propose a new feature matching algorithm which is robust to the viewpoint change by using the FAST(Features from Accelerated Segment Test) feature detector and the SIFT(Scale Invariant Feature Transform) feature descriptor. The original FAST algorithm unnecessarily results in many feature points along the edges in the image. To solve this problem, we apply the principal curvatures for refining it. We use the SIFT descriptor to describe the extracted feature points and calculate the homography matrix through the RANSAC(RANdom SAMple Consensus) with the matching pairs obtained from the two different viewpoint images. To make feature matching robust to the viewpoint change, we classify the matching pairs by calculating the Euclidean distance between the transformed coordinates by the homography transformation with feature points in the reference image and the coordinates of the feature points in the different viewpoint image. Through the experimental results, it is shown that the proposed algorithm has better performance than the conventional feature matching algorithms even though it has much less computational load.

I. 서 론

컴퓨터 비전 분야에서 영상의 특징점 정합 기법은

움직임 검출, 얼굴 인식, 3D 이미지 복원, 파노라마 스티칭, 물체 인식, 스테레오 유사도 측정 등 여러 분야에서 널리 사용되는 기술이다. 특징점 정합 기법은

※ 이 논문은 2015년도 정부(미래창조과학부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No.R0126-15-1034, 채널 객체 융합형 하이브리드 오디오 콘텐츠 제작 및 재생기술 개발)

♦ First Author : Kwangwoon University, guswh7905@kw.ac.kr, 학생회원

° Corresponding Author : Kwangwoon University, jsyoo@kw.ac.kr, 종신회원

논문번호 : KICS2015-09-313, Received September 17, 2015; Revised November 17, 2015; Accepted December 7, 2015

영상에서 특징점 추출 알고리즘을 이용하여 특징점을 추출하는 과정, 추출된 특징점을 서술자(descriptor)를 이용하여 특징점마다 서술하는 과정, 마지막으로 서로 다른 영상에서 서술된 특징점을 비교하여 유사한 특징점들을 연결하고 잘못 연결된 특징점을 제거하여 올바르게 연결된 특징점을 분리하는 과정으로 이루어진다. 특징점 정합 기법의 성능은 영상에서 독특한 특징점을 검출하는 것과 각 특징점마다 구별되는 서술자를 생성하는 방식에 따라 좌우된다.

초기의 특징점 검출 방법인 해리스(Harris) 코너 검출기는 헤이시안(Hessian) 행렬의 고유값을 이용하여 영상의 특징점을 찾는 방법으로 영상의 스케일 변환에 취약한 약점을 가지고 있다⁴⁾. Lindeberg가 제안한 가우시안 미분 검출기(Laplacian of Gaussian)는 스케일 공간에서 라플라시안의 극대점을 얻어내는 방식으로 크기 변화에 강인한 특징을 가지고 있다¹⁾. Lowe는 가우시안 미분 검출기(Laplacian of Gaussian)를 가우시안 차분 검출기(Difference of Gaussian)로 근사하여 영상의 특징점을 검출하고, 특징점 주변의 기울기(gradient) 방향의 히스토그램으로 특징점을 표현하여 영상 정합을 수행하는 SIFT(Scale Invariant Feature Transform) 기법을 제안하였다¹⁾. Mikolajczyk는 해리스 코너 검출기를 수정하여 라플라시안 기반의 스케일 변화를 고려한 특징점을 검출하고, SIFT 특징점 서술자를 수정한 GLOH(Gradient Location and Orientation Histogram)로 서술자를 생성하여 영상 정합을 수행하는 기법을 제안하였다^{2,3)}. Bay는 계산량을 감소시키면서 우수한 정합 성능을 나타내는 SURF(Speeded Up Robust Feature) 기법을 제안하였다⁶⁾. SURF는 헤이시안 행렬의 극대점을 사용해 특징점을 검출하고, Haar 응답으로 특징점 서술자를 생성하여 영상 정합을 수행하는데, 적분 영상 기법을 사용하여 계산량을 크게 감소시켰다⁷⁾. 하지만 SIFT기법은 영상의 조명 변화나 영상의 크기, 회전 변화에 강인하지만 많은 계산량으로 인해 실시간으로 사용될 수 없는 문제점을 가지고 있다. SURF 기법은 같은 물체나 장소에 대해 시점이 변화된 영상에 대해서는 정합율이 낮아지는 단점이 있다. 따라서 본 논문에서는 수행 시간을 단축하면서 동시에 시점 변화에 강인한 특징점 정합 기법을 제안하고자 한다.

제안하는 기법에서는 특징점을 빠르게 추출하기 위해 단순히 화소 값의 비교만으로 계산량을 최소화하여 특징점을 추출하는 FAST(Features from Accelerated Segment Test) 기법을 적용한다⁵⁾. FAST 기법은 SIFT의 DOG 방법이나 SURF의 Haar

wavelet 방식보다 특징점을 추출하는 속도 측면에서 성능이 월등히 좋다. 하지만 기존의 FAST 기법은 영상의 에지 부분을 따라서 특징점을 불필요하게 많이 추출하는 단점이 존재하는데 이러한 단점을 주곡률(principal curvatures)을 적용하여 개선한다. 추출된 특징점에 대해서는 더 많은 정보(128차원)를 기술할 수 있는 SIFT 서술자(descriptor)를 사용한다. 시점이 다른 두 영상으로부터 구해진 정합쌍에는 올바른 정합쌍뿐만 아니라 오차가 있는 정합쌍이 발생할 수 있으며, 잘못된 정합쌍의 분포가 전체 데이터에서 많은 부분을 차지할 경우에는 특징점 정합의 정확성이 떨어진다. 따라서 구해진 정합쌍에서 오차가 있는 정합쌍들을 제거하기 위해 Fischler and Bolles이 제안한 RANSAC(RANdom SAMple Consensus)을 사용하며 정확하게 정합된 4개의 특징점 쌍을 가지고 호모그래피 계산을 수행한다^{8,9)}. 시점 변화에 강인한 특징점 정합을 위해서 RANSAC을 통해 구한 호모그래피 행렬과 기준 영상의 특징점들을 이용하여 호모그래피 변환을 통해 변경된 좌표와 시점이 다른 영상의 특징점 좌표간의 유클리디언(Euclidean) 거리를 계산하고 이 거리를 이용하여 임계값(threshold)과 비교를 통해 정합쌍을 분류한다.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문에서 제안하는 시점 변화에 강인한 특징점 정합 기법을 소개한다. 3장은 시점이 다른 여러 영상을 가지고 실험을 통해서 제안하는 기법과 기존 기법간의 성능 비교를 보여주며 4장에서 결론을 맺는다.

II. 시점 변화에 강인한 특징점 정합 기법

본 논문에서 제안하는 기법의 흐름도는 그림 1과 같다.

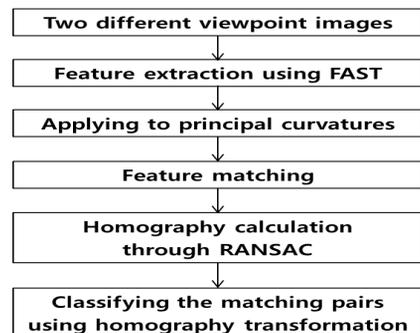


그림 1. 제안하는 특징점 정합 기법의 흐름도
Fig. 1. Flowchart of the proposed feature matching algorithm

2.1 FAST 기법을 이용한 특징점 추출

FAST(Features from Accelerated Segment Test) 기법은 실시간 특징점 추출에 적합한 기법으로 기존의 특징점 추출 기법인 Harris 코너 검출기, SIFT(Scale Invariant Feature Transform)의 DOG(Difference Of Gaussian), SURF(Speeded Up Robust Feature)의 Haar wavelet과 SUSAN(Smallest Uni-Value Segment Assimilating Nucleus Test)보다 빠른 속도로 특징점 추출이 가능하다.

그림 2는 특징점을 결정하기 위한 중심 화소와 주변 화소들과의 관계를 보여준다. FAST 기법은 영상에서 한 화소 p를 중심 화소라고 할 때 이로부터 거리가 3인 원을 형성한다. 형성된 원에 걸쳐있는 16개의 주변 화소를 $p \rightarrow x_k$ 라고 하면 밝기 값($I_{p \rightarrow x_k}$)과 중심 화소 p의 밝기 값(I_p)에서 임계값을 더하거나 뺀 값을 비교하여 중심 화소 p의 특징점 여부를 판별한다. 만약 16개의 화소들의 밝기 값 중 중심 화소 p에 임계값을 더한 밝기 값보다 더 크거나 중심 화소 p에 임계값을 뺀 밝기 값보다 더 작은 화소가 N개 이상 연속으로 있을 때 중심 화소 p를 특징점으로 한다.^[5]에 의하면 N의 값이 9일 경우 다음 프레임에서의 반복성(repeatability)이 가장 높고 임계값은 사용자가 임의로 설정하여 특징점을 검출할 수 있다.

이 과정에서 특징점 후보를 더 빠르게 추출하기 위해서 원을 구성하는 16개 화소들의 밝기 값을 비교하는 방법 대신에 그림 2에서 1,5,9,13번에 위치한 화소들의 밝기 값만을 비교한다. 4개의 화소들 중 3개 이상이 중심 화소의 밝기 값 I_p 에 임계값 t를 더한 $I_p + t$ 보다 밝거나 I_p 에 임계값 t를 뺀 $I_p - t$ 보다 어두울 경우에 해당 화소가 특징점일 가능성이 있다고 판단한

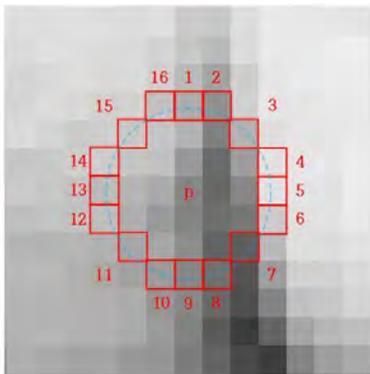


그림 2. 중심 화소 p를 기준으로 원 위에 존재하는 16개 화소
Fig. 2. 16 pixels on a circle centered at p

다. 일단 특징점일 가능성이 있다고 판단되면 16개 화소들의 밝기 값을 비교해서 중심 화소 p의 특징점 여부를 최종 판별한다. 특징점일 가능성을 만족하지 않으면 16개의 화소들의 밝기 값을 비교하지 않아도 되기 때문에 특징점 추출의 속도를 크게 증가시킬 수 있다.

위 과정은 영상의 모든 화소들에 대해 수행한다. 따라서 입력된 영상의 크기가 커질수록 계산량이 증가하며 수행시간이 늘어나는 문제가 있다. 이 문제를 해결하고자 FAST 기법에서는 결정 트리 구조를 사용한다. 트리 구조를 만들기 위해 우선 식(1)과 같이 각각의 화소들의 관계를 4가지 상태로 구분한다.

$$S_{p \rightarrow x_k} = \begin{cases} d, & I_{p \rightarrow x_k} \leq I_p - t \quad (dark) \\ s, & I_p - t < I_{p \rightarrow x_k} < I_p + t \quad (similar) \\ b, & I_p + t \leq I_{p \rightarrow x_k} \quad (bright) \\ u, & \quad (unknown) \end{cases} \quad (1)$$

$k = 1, \dots, 16$

여기서 $S_{p \rightarrow x_k}$ 는 원을 구성하는 16개 화소 중 k번째 화소의 상태를 나타낸다. 중심 화소 p의 밝기 값(I_p)에서 임계값을 뺀 값보다 원에 걸쳐 있는 주변 화소의 밝기 값($I_{p \rightarrow x_k}$)이 작으면 그 화소를 d로, 중심 화소 밝기 값과 임계값을 더한 값보다 크면 그 화소를 b로, 임계값을 뺀 값과 더한 값의 사이 값이면 s로 마지막으로 비교하지 않은 상태일 경우 u로 정의한다. 그림 3은 식 (1)을 통해 트리를 구성하여 깊이 탐색 방식(depth first search)으로 전체 화소에서 특징점을 추출하는 과정을 보여준다. 블록 안에 있는 알파벳은 원을 구성하는 16개 화소들의 상태를 나타낸다. 먼저 모든 화소들의 상태를 u로 초기화한다. 이후 중심 화소 p에 대해 식 (1)을 이용하여 16개 화소들의 상태를 결정한다. 깊이 탐색 방식에 의해서 그림 2에서 1번에 위치한 화소가 다시 중심 화소 p가 되어 다시 주변 16개의 화소들과 비교하는 과정을 반복한다. 이 과정에서 화소들 간의 상태 비교가 앞선 과정에서 먼저 결정되어 있는 경우에는 식 (1)의 과정을 진행하지 않는다. 위와 같은 상황이 모든 화소에 적용되기 때문에 식

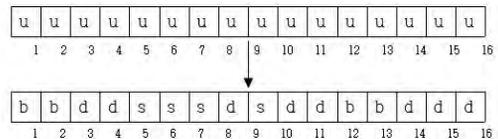


그림 3. 결정 트리 구조
Fig. 3. Decision tree structure

(1)의 과정을 생략하는 상황이 발생하여 수행 시간이 단축될 수 있다.

이 과정을 통해 추출된 특징점들은 주변 화소 밝기 값과의 차이를 이용하여 추출되었기 때문에 SIFT의 DOG 특징점 검출 기법과 비슷한 성능을 보이면서 동시에 처리 속도는 매우 빠르다. 하지만 FAST의 문제점인 SIFT의 DOG나 SURF의 Haar wavelet과는 다르게 스케일의 변화에 강인하기 위한 과정이 없기 때문에 추출된 특징점들이 군집되어 있는 경우가 많이 발생한다. 군집되어 있는 영역의 특징점은 각기 비슷한 서술자(descriptor)를 갖고 있기 때문에 특징점 정합을 할 경우 오류가 발생할 확률이 높아진다. 이를 해결하고자 NMS(Non Maximum Suppression) 방법을 이용하여 군집되어 있는 특징점 중에서 다음에 정의되는 값이 가장 큰 하나의 특징점을 선택한다. 특징점에 값을 부여하는 과정은 식 (2)와 같다.

$$V = \max \left(\sum_{x \in S_{bright}} |I_{p \rightarrow x} - I_p| - t, \sum_{x \in S_{dark}} |I_p - I_{p \rightarrow x}| - t \right) \quad (2)$$

여기서 $I_{p \rightarrow x}$ 와 I_p 는 각각 주변 화소의 밝기값과 중심 화소의 밝기 값을 나타낸다. 먼저 식 (1)에 따라 주변 16개의 화소의 상태를 결정한다. 이 후 결정 트리 구조에서 b상태($x \in S_{bright}$)에 해당하는 화소와 d 상태($x \in S_{dark}$)에 해당하는 화소에 대해 식 (2)를 적용하여 더 큰 값을 해당 특징점의 값으로 부여한다. 위의 과정을 인접한 특징점들에 대하여 모두 적용하여 가장 큰 값을 가지는 특징점을 군집 영역의 특징점으로 추출한다. 그림 4는 N이 9이고 임계값 t 가 20일 때 FAST 기법을 적용하여 특징점을 추출한 영상으로 (a)는 NMS 기법을 적용하지 않은 경우이고 (b)는 NMS



그림 4. FAST를 이용하여 특징점을 추출한 영상 그림 4(a) NMS기법을 적용하지 않은 경우 그림 4(b) NMS기법을 적용한 경우
Fig. 4. Feature points extracted by the FAST algorithm Fig. 4(a) With NMS Fig. 4(b) Without NMS

기법을 적용한 경우이다. NMS 기법을 적용했을 때 영상에서 불필요하게 많이 뽑힌 특징점의 개수가 상당수 줄어든 모습을 볼 수 있다.

2.2 주곡률(principal curvatures)을 이용한 특징점 추출 개선

NMS 기법을 적용했음에도 불구하고 FAST 기법은 영상의 에지(edge) 부분을 따라서 많은 특징점들이 추출된다는 단점이 있다. 이를 개선하기 위해서 곡면이 각 방향에 따라 굽은 정도를 나타내는 값인 주곡률을 이용한다. 앞에서 추출한 특징점마다 식 (3)과 같이 각의 방향별로 2차 미분값으로 이루어진 2x2 Hessian 행렬을 구한다.

$$H = \begin{pmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{pmatrix} \quad (3)$$

여기서 I_{xx} 는 해당 화소의 x 방향 2차 미분값, I_{yy} 는 y 방향 2차 미분값, I_{xy} 는 x ,와 y 방향으로 각각 미분한 값을 의미한다. 주곡률은 식 (3)의 행렬 H로부터 계산된다. 행렬 H의 고유값(eigenvalue)은 해당 화소의 주곡률과 비례하기 때문에 고유값을 직접 계산하는 것을 피하고 고유값의 비율만을 고려하여 주곡률을 구한다. 이 때, 고유값 중에 큰 값을 α 로 작은 값을 β 로 하고 식 (4)를 계산한다.

$$\begin{aligned} Tr(H) &= I_{xx} + I_{yy} = \alpha + \beta \\ Det(H) &= I_{xx}I_{yy} - (I_{xy})^2 = \alpha\beta \end{aligned} \quad (4)$$

여기서 $Tr(H)$ 은 행렬 H의 주대각선(main diagonal) 상에 있는 원소들의 합을 의미하고 $Det(H)$ 는 행렬 H의 행렬식(determinant)을 의미한다. 이 때 $Det(H)$ 값이 음수인 경우와 아래 식 (5)를 만족하지는 못하는 경우에는 해당 화소를 평탄한(flat) 지역이나 에지로 판단하기 때문에^[2] 에지 부분에 많은 특징점이 검출된다는 단점을 개선하기 위해서 이 경우는 특징점에서 제외한다. 이 때 r 은 임계값이 된다.

$$\begin{aligned} \frac{Tr(H)^2}{Det(H)} &= \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(r+1)^2}{r} \quad (r = \alpha/\beta) \\ \frac{Tr(H)^2}{Det(H)} &< \frac{(r+1)^2}{r} \end{aligned} \quad (5)$$

그림 5는 FAST 기법과 NMS 기법을 적용하여 특



그림 5. FAST 기법과 NMS 기법을 적용하여 특징점을 추출한 영상
 그림 5(a) NMS만 적용한 경우 그림 5(b) NMS기법과 주곡률을 모두 적용한 경우
 Fig. 5. Feature points extracted by applying FAST algorithm and NMS
 Fig 5(a) Only with NMS Fig 5(b) With NMS and the principal curvatures

징점을 추출한 영상으로 (a)는 NMS 기법만 적용한 경우이고 (b)는 NMS 기법과 주곡률을 모두 적용한 경우이다. (b)영상에서는 (a)영상과는 달리 에지 부분을 따라 존재하는 특징점이 상당수 줄어든 것을 알 수 있다.

2.3 SIFT 특징점 서술자 생성 및 정합

제안한 방법으로 영상에서 특징점을 추출한 후에는 각각의 특징점마다 정보를 포함하는 서술자(descriptor)를 생성한다. 본 논문에서는 시점 변화(viewpoint change)에 강인한 특징점 정합을 위해서 SURF의 서술자보다 특징점에 대해 더 많은 정보를 기술할 수 있는 SIFT의 서술자를 사용하여 각 특징점을 기술한다. 특징점 서술은 그림 6과 같이 각 특징점을 중심으로 지역적으로 서술한다. 4x4 크기의 블록 16개를 각각 8개의 방향으로 표현하며 각 화살표의 길이는 방향 히스토그램 값을 의미한다.

이렇게 모든 특징점들은 방향 히스토그램의 값들로

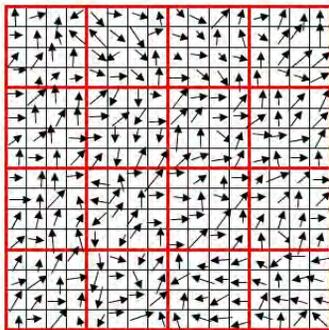


그림 6. SIFT의 서술자 블록 형태와 영상 기울기
 Fig. 6. Block of SIFT descriptors and image gradients

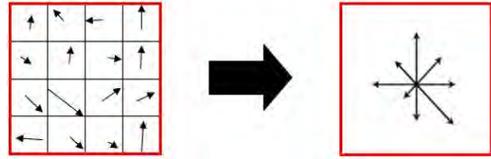


그림 7. SIFT의 특징점 서술자
 Fig. 7. Keypoint descriptor of SIFT

이루어진 벡터의 형태로 형성된다. 그림 7은 그림 6의 4x4 크기의 16개의 블록 중 하나로 방향 히스토그램의 8방향 성분을 SIFT 서술자 벡터로 표현한 예이다. 이와 같이 각 특징점을 16개 블록에 대해 각각 8개의 방향으로 표현하며 모두 128(4x4x8)차원으로 기술한다. 정합할 두 영상에서 추출된 특징점들 간의 128차원 벡터로 유클리디언 거리를 계산하고 그 거리 값이 임계값보다 작은 경우에 정합쌍으로 선정한다.

2.4 호모그래피(homography)를 이용한 정합쌍 분류

호모그래피는 왜곡된 영상을 보정하는 방법 중 하나로 서로 대응되는 영상 간의 2차원 사영 사상(projective mapping)을 정의하는 3x3 행렬이다. 시점이 다른 두 영상을 정합하기 위해서는 하나의 카메라 좌표계를 중심으로 호모그래피 변환을 적용해야 한다. 호모그래피 행렬을 구하기 위해서는 정확하게 정합된 특징점 4쌍이 필요하다. 특징점 4쌍을 선택하는 과정에서 이상점(outlier)이 사용될 경우 잘못된 호모그래피 행렬을 구할 수 있다. 호모그래피 행렬을 구하는 과정에서 이상점들을 제거하기 위하여 오류와 잡음이 섞여있는 데이터들로부터 적절한 모델을 예측하여 오차를 최소화하는 RANSAC(RANdom SAMple Consensus) 기법을 사용한다. 시점이 다른 두 영상에서 추출된 특징점들 간의 128차원 벡터 연산과 임계값 비교를 통해 추출된 정합쌍을 가지고 RANSAC 기법을 활용하여 호모그래피 행렬을 구한다. 식 (6)은 호모그래피 좌표 변환 식이다.

$$\begin{bmatrix} \hat{x}_1 w_1 \\ \hat{y}_1 w_1 \\ w_1 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \quad (6)$$

여기서 (x_1, y_1) 은 기준 영상에서 특징점의 좌표이고 (\hat{x}_1, \hat{y}_1) 는 호모그래피 변환에 의해 변환된 좌표이며 w_1 는 스케일 상수를 의미한다. 시점 변화에 강인한 특징점 정합을 위해서 기준 영상의 특징점들을 식

(6)처럼 호모그래피 변환을 통해 변경된 (\hat{x}_1, \hat{y}_1) 좌표와 시점이 다른 영상의 특징점 (x_2, y_2) 좌표를 이용하여 식(7)과 같이 두 좌표간의 유클리드 거리 연산과 임계값(T) 비교를 통해 정합쌍을 구한다.

$$\sqrt{(x_2 - \hat{x}_1)^2 + (y_2 - \hat{y}_1)^2} < T \quad (7)$$

그림 8은 시점이 다른 두 영상과 호모그래피 변환을 통해 보정된 영상을 보여준다. 그림 8(c)의 보정된 영상과 그림 8(b)의 시점이 다른 영상과 서로 유사한 형태를 갖는 것을 볼 수 있다.



그림 8. 시점이 다른 두 영상과 보정된 영상 그림 8(a) 기준 영상 그림 8(b) 시점이 다른 영상 그림 8(c) 호모그래피를 이용해 보정된 영상
Fig. 8. Two different viewpoint images and the corrected image. Fig 8(a) Reference image Fig 8(b) Query image Fig 8(c) Corrected image by homography

III. 실험 결과

본 논문에서는 Mikolajczyk가 제공하는 database 영상^[10]중에서 시점 변화가 존재하는 800x640 해상도의 Graffiti 영상 4장(정합을 위해 2장씩 선택해서 각각 A,와 B라 칭함), Caltech에서 제공하는 database 영상^[11]중에서 800x600 해상도의 시점 변화가 존재하는 세 종류의 Home Object 영상(A,B,C)을 실험영상으로 사용하였다. 실험 환경은 인텔 i5 쿼드코어 CPU, 8GB 램이며 Microsoft 사의 Visual Studio 2013으로 구현하였고 OpenCV 2.4.9버전을 사용하였다. 제안하는 기법의 성능을 평가하기 위해 기존의 특징점 정합 기법 중에서 성능이 우수한 SIFT, SURF 등의 기법과 성능 및 수행시간을 측정하여 비교하였다. 그림 9부터 그림 13까지는 제안한 기법과 SIFT, SURF 기법의 정합 결과를 보여준다. 제안한 기법이 각 영상에서 동일한 개수의 정합쌍을 추출할 때 기존의 정합 기법들보다 정합이 정확하게 되는 것을 볼 수 있다.

표 1과 표 2는 SIFT, SURF 기법과 제안하는 기법을 800x640 크기의 영상에 적용하여 각 과정별 수행 시간을 측정된 것이다. 제안하는 기법은 FAST 기법



그림 9. 'Graffiti A' 영상의 정합 결과, 제안한 기법(상)과 SURF(중), SIFT(하)
Fig. 9. Result of matching by the proposed algorithm(top), SURF(middle) and SIFT(bottom) for 'Graffiti A'



그림 10. 'Graffiti B' 영상의 정합 결과, 제안한 기법(상)과 SURF(중), SIFT(하)
Fig. 10. Result of matching by the proposed algorithm(top), SURF(middle) and SIFT(bottom) for 'Graffiti B'

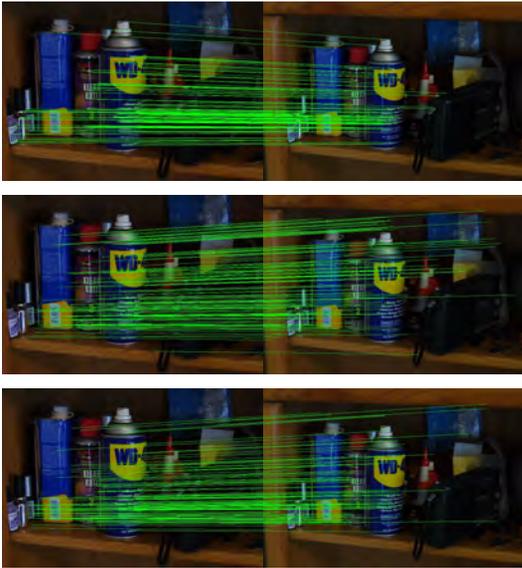


그림 11. 'Home Object A' 영상의 정합 결과, 제안한 기법(상)과 SURF(중), SIFT(하)
 Fig. 11. Result of matching by the proposed algorithm(top), SURF(middle) and SIFT(bottom) for 'Home Object A'



그림 13. 'Home Object C' 영상의 정합 결과, 제안한 기법(상)과 SURF(중), SIFT(하)
 Fig. 13. Result of matching by the proposed algorithm(top), SURF(middle) and SIFT(bottom) for 'Home Object C'

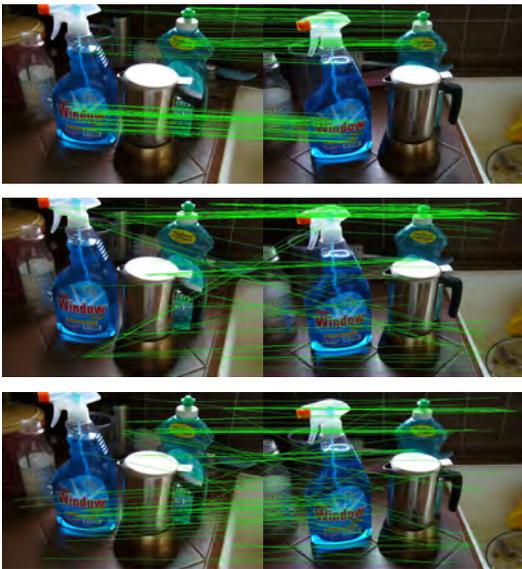


그림 12. 'Home Object B' 영상의 정합 결과, 제안한 기법(상)과 SURF(중), SIFT(하)
 Fig. 12. Result of matching by the proposed algorithm(top), SURF(middle) and SIFT(bottom) for 'Home Object B'

을 사용하였기 때문에 특징점을 추출하는 과정에서 기존의 추출 기법들보다 수행 시간이 많이 줄어든 것을 볼 수 있다. 특징점을 정합하는 과정에서도 기존의 정합 기법과 달리 제안하는 기법은 RANSAC을 통한

호모그래피 계산이나 호모그래피 변환을 이용한 정합 쌍 분류와 같은 과정이 더 추가되었지만 불필요한 특징점을 제거하였기 때문에 수행 시간이 더 줄어든 것을 볼 수 있다. 전체 수행 시간을 비교해보면 제안하

표 1. 'Graffiti A'의 단계별 수행 시간 비교
 Table 1. Comparison of runtime per each process for 'Graffiti A'

Process	Proposed Algorithm (ms)	Process	SIFT (ms)	SURF (ms)
Feature extraction using FAST	3	Feature extraction	362	147
Applying to principal curvatures	1			
Feature matching	178			
Homography calculation through RANSAC	4	Feature matching	322	312
Classifying the matching pairs using homography transformation	1			
Total	187	Total	684	459

는 기법이 SURF 기법보다는 수행 속도가 2배 정도 빠르고 SIFT 기법보다는 3.5배 정도 빠른 것을 확인할 수 있다.

표 3은 시점이 변하는 여러 영상에 대해 SIFT, SURF 기법과 제안하는 기법의 수행 시간을 측정된 결과이다. 각 영상에서 정합 기법마다 동일한 개수의 정합쌍을 추출할 때 제안하는 기법이 모든 영상에서 기존의 정합 기법보다 수행 시간이 더 적게 걸린 것을 알 수 있다.

표 2. 'Graffiti B'의 단계별 수행 시간 비교
Table 2. Comparison of runtime per each process for 'Graffiti B'

Process	Proposed Algorithm (ms)	Process	SIFT (ms)	SURF (ms)
Feature extraction using FAST	3	Feature extraction	375	149
Applying to principal curvatures	1			
Feature matching	196			
Homography calculation through RANSAC	3	Feature matching	327	389
Classifying the matching pairs using homography transformation	1			
Total	204	Total	702	538

표 3. 영상별 전체 수행 시간 비교
Table 3. Comparison of total runtime for each image

Image	Number of matching pairs	Proposed Algorithm (ms)	SIFT (ms)	SURF (ms)
'Graffiti A'	158	187	684	459
'Graffiti B'	178	204	702	538
'Home Object A'	127	92	635	224
'Home Object B'	89	210	682	466
'Home Object C'	140	170	760	300

IV. 결 론

본 논문에서는 FAST(Features from Accelerated Segment Test) 특징점 검출기와 SIFT(Scale Invariant Feature Transform) 특징점 서술자를 사용하여 시점 변화에 강인한 특징점 정합 기법을 제안하였다. 기존의 FAST 기법은 영상의 에지 부분에 불필요하게 특징점이 많이 검출되는데 이러한 단점을 주목률을 이용하여 개선하였다. 추출된 특징점마다 정보를 더 기술하기 위해서 SIFT 특징점 서술자를 이용해서 정합쌍을 구하였다. 시점 변화에 강인한 특징점 정합을 위해서 RANSAC(Random Sample Consensus)을 이용하여 호모그래피(homography) 행렬을 획득한 후 기준 영상의 특징점들을 호모그래피 변환을 통해 변경된 좌표와 시점이 다른 영상의 특징점 좌표간의 유클리디언(Euclidean) 거리를 통해 정합쌍을 분류하였다. 제안하는 영상 정합 기법은 SIFT나 SURF보다 같은 물체나 장소에 대해 시점이 변화된 여러 영상에서 전체 수행 시간을 비교해보았을 때, 제안하는 기법이 SURF 기법보다는 수행 속도가 2배 정도 빠르고 SIFT 기법보다는 3.5배 정도 빠른 것을 확인할 수 있고 동일한 개수의 정합쌍을 추출할 때 기존의 정합 기법들보다 정합이 정확하게 되는 것을 볼 수 있다.

References

- [1] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91-110, Nov. 2004.
- [2] K. Mikolajczyk, "Scale & affine invariant interest point detectors," *Int. J. Comput. Vision*, vol. 60, no. 1, pp. 63-86, Oct. 2004.
- [3] K. Mikolajczyk, "A performance evaluation of local descriptors," *Pattern Anal. and Machine Intell.*, vol. 27, no. 10, pp. 1615-1630, Oct. 2005.
- [4] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool. "A comparison of affine region detectors," *Int. J. Comput. Vision*, vol. 65, no. 1-2, pp. 43-72, Nov. 2005.
- [5] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," *9th Eur. Conf. Comput. Vision*, Graz, Austria, pp. 430-443, May 2006.

[6] E. Rosten, "Faster and better : A machine learning approach to corner detection," *Pattern Anal. and Machine Intell.*, vol. 32, no. 1, pp. 105-119, Jan. 2010.

[7] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-up robust feature," *Computer Vision and Image Understanding*, vol. 10, no. 3, pp. 346-359, Jun. 2008.

[8] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381-395, Jun. 1981.

[9] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," in *Proc. 2000 IEEE Conf. Comput. Vision and Pattern Recognition*, vol. 2, pp. 142-149, Jun. 2000.

[10] <http://www.robots.ox.ac.uk/~vgg/research/affine/>

[11] <http://www.vision.caltech.edu/pmreels/Datasets/>

[12] M. M. Hossain, H. J. Lee, and J. S. Lee, "Fast image stitching for video stabilization using sift feature points," *J. KICS*, vol. 39, no. 10, pp. 957-966, Oct. 2014.

[13] B. W. Chung, K. Y. Park, and S. Y. Hwang, "A fast and efficient haar-like feature selection algorithm for object detection," *J. KICS*, vol. 38, no. 6, pp. 486-497, Jun. 2013.

[14] H. K. Jang, "The more environmentally robust edge detection of moving objects using improved Canny edge detector and Freeman chain code," *J. KICS*, vol. 37, no. 2, pp. 37-42, Apr. 2012.

정 현 조 (Hyun-jo Jung)



2015년 2월 : 광운대학교 전자공학과 졸업
 2015년 3월~현재 : 광운대학교 전자공학과 석사과정
 <관심분야> 영상 신호처리, 기계 학습, 객체 인식

유 지 상 (Ji-sang Yoo)



1985년 2월 : 서울대학교 전자공학과 졸업
 1987년 2월 : 서울대학교 전자공학과 석사
 1993년 5월 : Purdue Univ. EE, Ph.D.
 1997년 9월~현재 : 광운대학교 전자공학과 교수

<관심분야> 비선형 디지털 신호처리, 웨이블릿 기반 영상처리