SIFT를 이용한 내시경 영상에서의 특징점 추출

Feature Extraction for Endoscopic Image by using the Scale Invariant Feature Transform (SIFT)

오장석*, 김호철**, 김형률*, 구자민*, 김민기***
J. S. Oh*, H. C. Kim**, H. R. Kim*, J. M. Koo*, M. G. Kim***

Abstract - Study that uses geometrical information in computer vision is lively. Problem that should be preceded is matching problem before studying. Feature point should be extracted for well matching. There are a lot of methods that extract feature point from former days are studied. Because problem does not exist algorithm that is applied for all images, it is a hot water. Specially, it is not easy to find feature point in endoscope image. The big problem can not decide easily a point that is predicted feature point as can know even if see endoscope image as eyes. Also, accuracy of matching problem can be decided after number of feature points is enough and also distributed on whole image. In this paper studied algorithm that can apply to endoscope image. SIFT method displayed excellent performance when compared with alternative way (Affine invariant point detector etc.) in general image but SIFT parameter that used in general image can't apply to endoscope image. The goal of this paper is abstraction of feature point on endoscope image that controlled by contrast threshold and curvature threshold among the parameters for applying SIFT method on endoscope image. Studied about method that feature points can have good distribution and control number of feature point than traditional alternative way by controlling the parameters on experiment result.

Key Words : SIFT, Feature Extraction, Endoscopic Image

1. 장 서론

이미지에서 매칭할 수 있는 특징점을 추출하는게 컴퓨터비전에서 중요한 문제이다.

특징점은 이미지 크기변화, 회전변환에 불변해야 한다. 또한 부분적으로 조명과 카메라의 포즈변화에 부분적으로 불변해야 한다.

이러한 연구는 대표적으로 [1],[2]에서 말하고 있다.

하지만 내시경 이미지는 일반 영상과는 다르다. 일반 영상에 비해 특징점으로 생각할 수 있는 점이 거의 없기 때문이다.

[1],[2]에서 말하고 있는 방법은 일반 영상에서 잘 적용이 되지만 내시경 영상에서는 그대로 적용하기에는 적합하지 않다.

실험을 통해 [2]방법의 파라미터 값을 변화시키면서 매칭하기에 좋은 특징점을 추출할 수 있다는 것을 알았다.

또한 간단한 알고리즘을 이용하여 적당한 파라미터 값을 찾을 수 있었다. 이러한 알고리즘을 통하여 얻은 값은 내시경 영상뿐만 아니라 일반 영상에도 적용할 수 있어 효율적인 방법이다.

*高麗大學 電子情報工學科 碩士課程
**高麗大學 醫學科 博士課程
***高麗大學 電子情報工學科 教授 · 工博

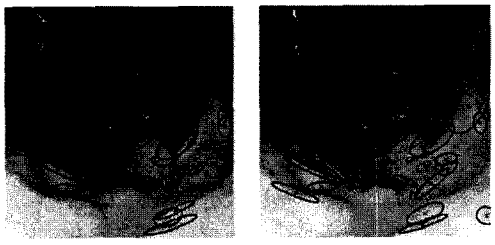
• 본 연구는 산업자원부의 지원을 받아서 이루어진 것임.

2. 장 본론

2.1 절 Affine invariant point detector

(a) (b)

그림1. 일반 영상에서의 affine invariant points



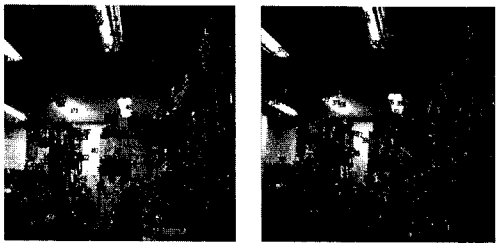
(a) (b)

그림2. 내시경 영상에서의 affine invariant points

그림1에서와 같이 일반영상에서는 (a)는 97개, (b)는 83개의 특징점을 추출할 수 있었다. 그러나 그림2에서와 같이 내

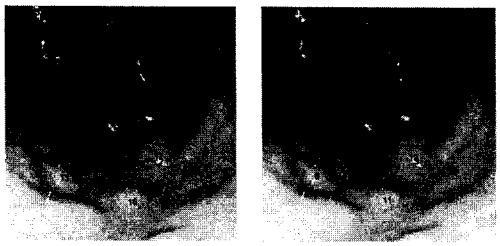
시경 영상에서는 (a)는 39개, (b)는 40개의 특징점을 추출할 수 있었다. 심각한 문제는 내시경을 찍을 때에 반사되는 조명부분에 민감하다는 것이다. 그리고 분포와 수가 적당하지 않아 매칭할 때 문제가 된다.

2.2 절 Scale Invariant Feature Transform (SIFT)



(a) (b)

그림3. 일반 영상에서의 SIFT적용한 특징점



(a) (b)

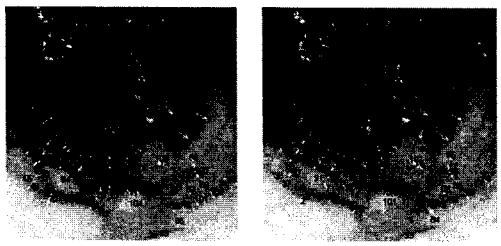
그림4. 내시경 영상에서의 SIFT적용한 특징점

그림3에서 얻은 결과를 보면 (a)는 485개, (b)는 414개의 특징점을 추출해 그림1보다 많은 매칭점을 가질 수 있다.

그러나 여전히 그림4에서 (a)는 18개, (b)는 13개로 내시경 영상에 적용하는 문제가 있음을 알 수 있다.

그림3과 그림4는 contrast threshold를 0.02로 하고, curvature threshold를 10으로 했을 때의 결과이다.

먼저 contrast threshold만 0.01로 하면 그림5와 같이 (a)는 111개, (b)는 110개의 특징점을 추출하는 결과를 얻을 수 있다.

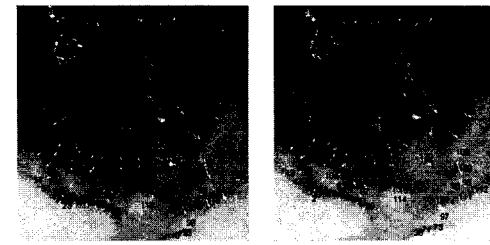


(a) (b)

그림5. 내시경 영상에서의 SIFT적용한 특징점

결과에서 볼 수 있듯이 영상에 따라 contrast threshold를 조정할 필요가 있다.

다음으로 contrast threshold를 0.01로 두고, curvature threshold만 50으로 조정하면 그림6의 결과를 얻을 수 있다.



(a) (b)

그림6. 내시경 영상에서의 SIFT적용한 특징점

그림6에서 (a)는 126개, (b)는 123개로 그림5와 비교할 때 별 차이가 없다. 다만 경계로 보이는 곳에 특징점이 조금 늘었다.

따라서 curvature threshold보다는 contrast threshold를 잘 조정하는 것이 필요하다.

SIFT의 순서를 보면 다음과 같다.

1. Scale-space extrema detection
2. Keypoint localization
3. Orientation assignment
4. Keypoint descriptor

위의 과정 중에 contrast threshold와 관련된 것은 두 번째 과정인 Keypoint localization이다.

$$D(X) = D + \frac{\partial D^T}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X \quad (1)$$

$$\hat{X} = -\frac{\partial^2 D^{-1}}{\partial X^2} \frac{\partial D}{\partial X} \quad (2)$$

$$D(\hat{X}) = D + \frac{1}{2} \frac{\partial D^T}{\partial X} \hat{X} \quad (3)$$

$$|D(\hat{X})| < contrast\ threshold \quad (4)$$

식(4)를 만족하는 특징점을 제거하게 되므로 contrast threshold값을 작게 할수록 많은 점을 추출하게 된다.

영상마다 $|D(\hat{X})|$의 값의 분포가 다 다르다. 실제로 각 그림에서 (a)와 (b)의 영상이 거의 똑같이 보이지만 알고리즘에 의해 얻은 contrast threshold값을 보면 다르다.

알고리즘은 각 영상에서 $|D(\hat{X})|$의 히스토그램을 분석하여 최고치에 이르는 값을 contrast threshold로 쓰는 것이다.

또한 curvature threshold, γ과 관련된 과정도 두 번째이다.

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (5)$$

$$Tr(H) = D_{xx} + D_{yy} = \alpha + \beta, \quad Det(H) = D_{xx} D_{yy} - (D_{xy})^2 = \alpha\beta \quad (6)$$

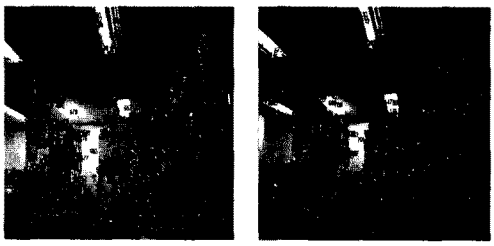
$$\alpha = r\beta, \quad \frac{Tr(H)^2}{Det(H)} = \frac{(\alpha+\beta)^2}{\alpha\beta} = \frac{(r\beta+\beta)^2}{r\beta^2} = \frac{(r+1)^2}{r} \quad (7)$$

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r} \quad (8)$$

식(8)을 만족하는 특징점을 제거하게 되므로 γ값을 크게 할수록 경계로 보이는 부분에서 조금 더 특징점을 추출할 수 있다.

3. 장 결론

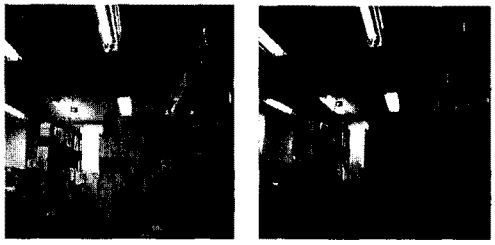
실험은 curvature threshold가 중요하게 작용하지 않으므로 10이라는 값으로 고정하였다.



(a) (b)

그림7. 일반 영상에서의 SIFT적용한 특징점

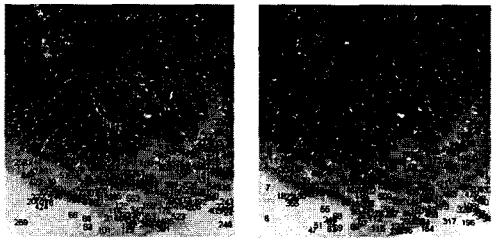
그림7에서 (a)는 687개, (b)는 772개의 특징점을 추출할 수 있었다. contrast threshold값은 (a)가 0.0077이고, (b)는 0.0056이었다.



(a) (b)

그림8. 그림7의 특정점을 가지고 매칭한 점

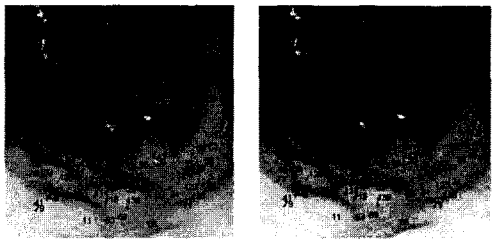
그림8에서는 실제로 매칭알고리즘을 통하여 얻은 점인데 135개의 매칭점을 얻었다.



(a) (b)

그림9. 내시경 영상에서의 SIFT적용한 특징점

그림8에서 (a)는 560개, (b)는 509개의 특징점을 추출할 수 있었다. contrast threshold값은 (a)가 0.0031이고, (b)는 0.0032이었다.



(a) (b)

그림10. 그림9의 특정점을 가지고 매칭한 점

그림10에서도 실제로 매칭알고리즘을 통하여 얻은 점인데 145개의 매칭점을 얻었다.

그리고 그림3은 91개, 그림4는 9개, 그림5는 54개, 그림6은 61개의 매칭점을 찾을 수 있었다.

이러한 매칭된 점들은 Fundamental Matrix를 구하는데 이용된다.

참 고 문 헌

[1] K. Mikolajczyk and C. Schmid, "An affine invariant interest point detector", ECCV'02, vol I, pp. 128-142.

[2] David G. Lowe, "Distinctive Image Features from Sacle-Invariant Keypoints", accepted for publicatoin in the International Journal of Computer Vision, 2004.

[3] David G. Lowe, "Object Recognition from Local Scale-Invariant Features", Proc. of the International Conference on Computer Vision, Corfu, September 1999, pp. 1150-1157.

[4] David G. Lowe, "Local feature view clustering for 3D object recognition", IEEE Conference on Computer Vision and Pattern Recognition, Kauai, Hawaii (December 2001), pp. 682-688.