

## 1. Reviewer

neuralix

## 2. Article

David G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, 60, 2 (2004), pp. 91-110.

(개인적으로 ICCV1999 제출본을 보조적으로 봄. 이 논문을 필수적으로 함께 보아야 할 것임.)

## 3. Stage 설명 및 Comment

### Local extrema detection

#### 알고리즘 설명

1. 입력 이미지를 두배로 키움 (bilinear interpolation)

2.  $\sigma = \sqrt{2}$ 로 Gaussian smoothing 적용. 이후  $\sqrt{2}$ 를  $\sigma$ 에 곱해 Gaussian 적용. 서로 인접한 smoothed 결과끼리 빼주어 DoG 영상 산출

(이상을 수 회 반복해 다수의 DoG를 산출하며 이를 octave라 함)

Gaussian smoothing은 1D convolution을 가로 세로 방향으로 각각 수행하여 연산 시간을 절약.

3. 최초 octave의 첫번째 smoothed된 영상을 1.5배 spacing(인접 4 pixel의 보간)에 따라 sampling하여 level이 다른 영상을 만듦. 이에 대해 2를 반복.

4. 각 픽셀에 대해, 한 level 내에서 주변 8 pixel을 검사하여 extrema인지를 검사. extrema일 경우 상하 각 레벨(서로 다른 레벨에 대해서는 각 레벨 당 9 pixel 모두와 비교)에 대해 검사를 수행. 이를 모두 만족하면 extrema로 인정.

#### Critique with Questions and Issues

합리적인 방법이라 생각됨. 단, 각 level에서 몇개의 DoG를 산출할 것인지는 fix시키지 못 했음.

### Accurate keypoint localization

#### 알고리즘 설명

Quadratic function에 의한 sub-sampling으로 key point의 위치를 보다 정교히 찾음. 이때  $D(x)$ 의 절대값이 0.03보다 작으면 위치보정 무시.

#### Critique with Questions and Issues

이와 같이 SIFT는 local image descriptor 전까지는 최대한 정교하게 찾으려는 노력을 하고 있음. 그러나 이 단계에 들 계산 cost를 고려해 볼 때 과연 이 단계가 필요한지 의문임.

### Eliminating edge responses

#### 알고리즘 설명

Hessian matrix을 구하고 이것이 반영한다고 가정하는 ellipse의 장단축 길이비를 이용해 edge response를 test. Ellipse의 길이비는 두 eigen value 값 간의 비가 반영하게 되는데 이를 직접적으로 구하는 것이 계산 cost가 많이 드므로 두 eigen value간의 비를 나타내는 변수  $r$ 과 증감성질을 공유하는 함수인  $(r+1)^2/r$ 의 최대화를 통해 구함.

#### Critique with Questions and Issues

알려져 온 내용에 계산효율성을 위한 아이디어를 제안함. 증감성질을 공유하는 함수의 사용은 훌륭함. 단, 어떤 point가 edge의 끊어진 end point일 가능성도 있음. 이는 뒤 단계의 지식을 통한 supervision으로 filtering되어야 할 것이나 현 단계에서는 이에 대한 연구가 별로 진행된 바 없음.

### Orientation assignment

#### 알고리즘 설명

각 픽셀의 magnitude(local extrema detection에서 이미 계산되었음)와 orientation을 구함. orientation은 36개의 bin으로 나뉘며, 크기  $m$ 을 가진 벡터는 각 bin이 나타내는 방향벡터와의 내적을 통해 decompose됨. 이는 Circular하게 배치된 상태에서 keypoint scale에 1.5배를 곱한 sigma 계수에 의해 smoothing됨.

#### Critique with Questions and Issues

orientation 산출 시 Lowe가 적용한 것처럼 adjacent pixel간의 차분을 사용할지 Sobel처럼 중심픽셀 양쪽간에 차분을 할지는 논란의 여지가 있겠음. 특히 SNR이 낮은 경우는 경험상, Sobel과 같이 서로 멀리 떨어진 위치간에 차분이 수행됨. Noise가 큰 경우에 대한 concerning은 사실 SIFT에서 빠져있는 것으로 보임.

Histogram을 구축하고 이에 대해 smoothing을 수행하는 것은 크게 보아 합리적인 framework이기는 하나 역시 object의 boundary 부분에서 나타나는 양상변화(;Lowe도 말하고 있는 artifact적인..)에 영향받을 소지가 있음. 이 역시 개선이 필요한 과제임.

#### Local image descriptor

##### 알고리즘 설명

각 픽셀은 이를 포함하는 8X8 grid를 가지는데 이때 grid의 각 픽셀에 대해 magnitude와 orientation vector 정보를 알아야 함. (이때 orientation은 중앙값이 일률적으로 빠짐) 이는 다시 2X2 grid로 변환되는데 grid의 각 칸은 4X4 pixel들의 orientation-magnitude 벡터들을 단일의 feature 벡터로 만든 것임. 이 feature vector들은 normalize됨.

#### Critique with Questions and Issues

Edelman의 논문을 본 적이 있는데 이를 잘 해결한 것으로 보임. out-of-plane motion의 작은 variation을 suppress하는 방법으로 아주 reasonable함. 그러나 역시 boundary에서의 양상변화에 대한 대처를 고려할 필요는 있겠음.

#### Nearest neighbor indexing

##### 알고리즘 설명

BBF(Best Bin First) 알고리즘을 사용

#### Critique with Questions and Issues

BBF 알고리즘을 모름

#### Feature Clustering

##### 알고리즘 설명

Hough 변환을 사용. 변수는 orientation, scale, location 세가지 임.

#### Critique with Questions and Issues

Hough 변환의 일반적 성질에 대해 생각해보는 기회가 되었음. 구체적인 구현에 대해서는 아직 이해하지 못 함. 단, 보통 Hough 변환이 cost가 비싸다는 점에서 다른 대안이 없을까 하는 생각을 해봄.

그리고 Feature clustering시, 서로 다른 level의 DoG에서 얻은 feature들도 clustering되는게 상관없는지? 아무튼 supervised indexing이 이뤄지는 것이니 상관은 없다고 보임.

#### Affine parameter 추정

##### 알고리즘 설명

Least square method를 이용하여 Affine 파라미터를 추정. 3개의 특징점이 찾아지면 그 cluster가 찾아진 것으로 판단.

#### Critique with Questions and Issues

단순한 방법을 이용했음. 구체적인 구현은 좀 더 이해가 필요.

### 전체 Critique

Feature detection, construction의 가장 핵심적인 문제인 파라미터의 축소에 큰 진전을 준 연구.  
AI에 대한 연구자의 concern이 motivation 및 방법 채택에 많은 영향을 끼친 것이 느껴짐.  
지금까지 나왔던 많은 연구를 집대성하고.. 자체적으로 개량한 연구  
연구자 및 연구실의 축적된 연구성과가 있었기에 가능했을 것임..