

# 로컬영역에서 다중 특징을 이용한 물체인식

최경영<sup>0</sup>

디지털 미디어 연구소, 정보통신대학교

kyungyung@icu.ac.kr

## Object Recognition using Multiple Local Features

Kyung-Yung Choi<sup>0</sup>

Digital Media Laboratory, Information and Communication University

### 요 약

본 논문은 향상된 Scale Invariant Feature Transform (SIFT) 기법과 이로부터 얻어진 로컬 특징 영역에서 다중특징을 이용한 물체인식 방법에 대하여 논하였다. SIFT 기법 [1]은 물체의 크기, 회전, 3차원 좌표변환에 강인한 특성을 갖는다. 이 기법에서는 크기가 다른 가우시안 (Gaussian) 함수를 적용한 영상들의 차이에서의 최대 및 최소값이 특징점으로 결정된다. 하지만 SIFT 알고리즘의 특성상, 인식되어야 될 물체의 비교적 큰 크기 변화, 중요도가 낮은 특징점들의 추출, 그리고 서로 다른 물체에서 추출된 유사한 특징벡터들이 인식 시스템의 신뢰도를 저하 시킬 수 있다. 이에 대응방안으로, 본 논문에서는 상대적으로 낮은 인식정보를 갖는 추출된 특징점을 제거하기 위한 기법과 서로 다른 물체에서 생성된 유사 특징벡터의 구분을 위한 특징점에서의 방위 (orientation) 비교법 및 색차 (chrominance) 정보를 사용에 대하여 기술하였다.

### 1. 서 론

제한되지 않는 환경에서의 물체인식 시스템 구현에서, 임의의 거리와 각도에 의해 변하는 물체의 형태와 인식되어야 할 물체가 부분적으로 가려진 경우 (occlusion) 에 대한 대책이 필요하다. 인식 대상의 임의의 스케일 변화 및 3차원 변환에 강인한 기법으로 알려진 로컬영역의 특성해석을 이용하는 인식방법은 시각에 불변한 (view-invariant) 특징점들을 추출하고 이에 대응하는 로컬영역을 3차원 변환에 대하여 강인한 모델로 설계함으로써, 언급된 문제들에 대한 해결책으로 제시 되어왔다. 이런 기법은, 물체의 엣지 또는 실루엣 등 글로벌한 특징에 대한 기하학적 변환으로부터 물체를 찾는 방법들 [5,8] 과는 다르게, 물체가 부분적으로 가려진 문제에 대해서도 강인성을 유지 할 수 있다.

로컬 영역의 특징 해석기법은 1) 여러 가지 스케일에서 가우시안 (Gaussian) 미분 함수를 사용하여 얻어진 영상의 특성을 통해 코너, 원형, 접합점들의 특징점을 추출하는 방법 [2, 3, 9, 11], 2) 가우시안 필터링된 여러 영상들의 차이에서 얻어지는 로컬 극한값으로 특징을 추출하는 방법 [1, 12], 3) PCA (Principal component Analysis) 기반으로 물체의 로컬영역에 대한 주요한 정보를 저차원의 공간에 투영시키는 기법 [4, 7] 등으로 크게 구분되어 진다.

Lowe[2]에 의해 제안된 Scale Invariant Feature Transform (SIFT) 기법은 크기가 다른 가우시안 필터를 단계적으로 사용하여 블러링 된 영상을 얻는다. 얻어진 영상들의 차이인 DoG (Difference of Gaussian) 에 대한 로컬 극한값 측정으로 주어진 정지영상의 특징점들을 추출한다. 이 특징점들에 대응하는 영역에 대한 기울기 및

방위 (orientation) 해석으로 대응영역의 특성을 구분짓는 특징벡터를 생성한다. DoG를 이용하는 SIFT는, 물체의 크기, 회전, 3차원 좌표변환에 대하여 안정되고 지속적인 특징점을 찾는다 알려진 LoG (Laplace of Gaussian) 함수를 사용하여 얻어 질 수 있는 결과를 적은 계산량으로 추출한다 [13]. 그러나 DoG를 사용하는 SIFT 알고리즘의 특성상, 대비 (contrast) 가 약한 영역과 엣지 영역에서의 추출될 수 있는 다수의 중요도가 낮은 특징점들의 추출, 그리고 서로 다른 물체에서 추출된 유사한 특징벡터들이 인식 시스템의 신뢰도를 저하 시킬 수 있다. 이에 대응방안으로, 본 논문에서는 상대적으로 낮은 인식정보를 갖는 추출된 특징점을 제거하기 위한 기법, 그리고 유사한 특징벡터의 구분을 위한 로컬영역에서의 방위 및 색차 정보의 사용에 대하여 기술하였다

### 2. SIFT 특징 추출

#### 2.1 Scale Invariant Feature Transform

SIFT 기법은 크기가 다른 가우시안 함수를 적용한 영상,  $L(x,y,\sigma)$ , 들의 차에 대한 최대 최소값을 특징점으로 결정한다.

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y), \quad (1)$$

$$G(x,y,\sigma) = 1/(2\pi\sigma^2) \cdot e^{-(x^2+y^2)/2\sigma^2} \quad (2)$$

그림 1은 다중 스케일에서 극한값을 찾는 방법을 보인다. 크기,  $\sigma$ ,가 다른 가우시안 함수를 사용하여 얻어진 인접한 영상들의 차이로 DoG가 얻어진다. DoG들 중, 그림 1에서와 같이, DoG<sub>2</sub>의 c가 인접픽셀들 중 최소 또는 최대값을 가지면 c는 특징점으로 결정된다. 특징점의 기울기 크기와 방위는 근처 픽셀들의 값에 의해 결정되고, 스케일과 3차원 변환에 둔감하게 하는 특징 영역에 대한 해석은, 정해진 크기의 원형 가우시안 필터를 특징점에 대

음하는 영역의 기울기 크기와 방위에 적용하여 이루어진다. 이때 특징영역은 16개의 4x4의 세부영역으로 나누어지고 각 세부영역에서 얻어진 방위정보가 특징벡터를 생성한다.

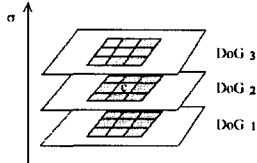


그림 1. 근접픽셀 검사로 얻어지는 SIFT 특징점 결정.

### 2.2 SIFT 기법의 보완

DoG는 엣지 및 그 주변에 강하게 반응하며, 이 부근에서 추출된 서로 다른 물체의 특징점들의 특성해석 결과는 비슷할 수 있다. 이는 매칭 절차에서 시스템의 의사 결정을 어렵게 한다. 또한, 그림 2에서 처럼 모델 물체의 특징 추출하기 위해 DoG의 극한값을 찾을 때, 배경 부분과 대비가 작은 영역에 특징점들이 추출되는 것을 볼 수 있다. 이렇게 검출된 특징점들은 배경과 대상 모델의 구분을 어렵게 할 뿐 아니라, 의미있는 정보를 제공하지 못한다. 또한 다음 인식 단계에서 불필요하게 많은 점들의 매칭은 인식시스템의 속도와 인식율을 저하시킨다. 이런 문제들에 대한 해결방안으로, 먼저, SIFT 입력단에 가우시안 함수를 적용하였다. 이 방법은 추출점의 빈도가 높은 영역과 더불어 효과적으로 엣지주변의 특징점들을 제거한다. 하지만 이 방법으로 엣지 주변의 특징점은 제거될 수 있지만, 대비가 약하여 엣지 크기가 약한 부분에서 검출된 인식률에 큰 영향을 주지 못하는 특징점들이 함께 검출되고 있다. 이런 특징점들은 정량화된 (normalized) DoG에서 검출된 특징점에 대하여 라플라시안 연산자를 적용한 후 임계값의 사용으로 제거될 수 있다. 마지막으로 Hessian 행렬의 곡률 (principle curvature)을 사용한 방법[13]으로 엣지영향을 감소시킨다.

### 2.3 방위정보와 색차를 이용한 유사 특징 벡터의 구분

SIFT 기법은 서로 다른 물체의 로컬영역에서 유사한 특성을 가진 특징 벡터가 존재할 때 인식률의 저하를 가져온다. 이러한 유사 특징점들은 첫째, 로컬 영역에서의 기울기의 특성이 유사하거나, 둘째, 추출된 특징벡터가 인식하고자 하는 물체의 특징 벡터군에 포함되지 않거나, 또는 인식하려 하지 않는 물체의 특징 벡터에 더 근접할 때 발생한다. 이에 대한 대응책으로 본 논문에서는 방위 비교법과 색차 정보의 활용이 사용되었다. 특징영역에서 생성된 특징벡터와 시스템이 가지고 있는 물체들 간의 Euclidean 거리를 비교하여 특징벡터에 대응하는 영역을 구분짓는 기존의 매칭기법만으로는 유사벡터의 구분에 강인함을 유지하기 어렵다. 특징점에서의 방위,  $\theta$ , 는 스케일, 회전, 3차원 변환등에 지속적인 특성을 보이며, 이 특성은 기존의 거리측정을 사용한 기법과 더불어 특성벡터를 구분 짓는데 사용되었다.

$$\theta(x,y) = \text{atan2}(L(x+1,y) - L(x-1,y), L(x,y+1) - L(x,y-1)) \quad (3)$$
 특징벡터 간의 거리측정 후, 인식되어야 할 물체로 구분된 특징영역에 대한 방위 정보를 시스템이 가지고 있는

모델에 가장 근접한 특징영역의 방위정보와 비교한다. 만약 두 방위의 차이가 정해진 임계오차 범위를 넘으면, 그 특징영역을 인식되어야 할 물체의 특징영역 중에서 제거한다. 이런 절차를 거친 특징점들에 대한 색차 검사로 최종적으로 특징영역을 구분한다. 일반적으로 주어진 영상 전체에 대하여 색차를 이용한 인식방법은 색감의 편차가 작게 분포된 물체의 인식을 위해 제한적으로 사용될 수 있지만, 물체와 배경의 색이 비슷한 경우, 여러가지 색으로 표면이 구성된 경우등이 색차기법의 일반적인 적용을 어렵게 한다. 이런 문제는 YCbCr 컬러영역을 특징점을 중심으로 지름 5 픽셀의 원형 영역에 적용하고, 조영조건에 둔감한 Cb와 Cr 성분을 측정함으로써 해결할 수 있다. 중심점으로부터 차등 가중치를 적용한 특징영역에서 생성된 색정보 벡터, [Cb, Cr], 와 시스템이 가진 근접벡터의 색정보 벡터와의 거리의 임계값으로 검출된 특징영역의 구분이 결정된다.

### 3. 실험 및 평가

제안된 기법들을 Desk-top 환경에서 물체들에 대하여 적용하였다. 실험에 사용된 물체는 형태가 비슷한 두가지 책이며 인식시스템은 이 두가지 책들 중에서 인식하려는 책을 특징영역에 대한 해석으로부터 구분해야 한다. 두가지 책에 대한 특징 벡터는 모델링과정에서 미리 생성되었고, 제안된 기법의 평가를 위해 임의의 조건에서 두 물체가 촬영되었다. 먼저 특징점 추출을 위해 2.2절에 기술된 향상된 SIFT 알고리즘이 적용되었다. 기존의 SIFT의 특성상 엣지부근에서 다수의 특징점들이 추출되었으며 유사한 형태의 물체들에서 나온 이러한 특징영역의 해석결과는 시스템의 의사 결정을 어렵게 한다. 그림 2 오른쪽에서 처럼 가우시안 함수를 시스템의 입력단에 적용하여 강한 엣지부근에서 검출될 수 있는 특징점들을 제거했으며, 가우시안 이미지의 차에 추출된 특징점에 대한 기울기 검사와 Hessian 곡률을 통해 중요도가 낮은 특징점들과 엣지 영향이 제거 되었음을 볼 수 있다.

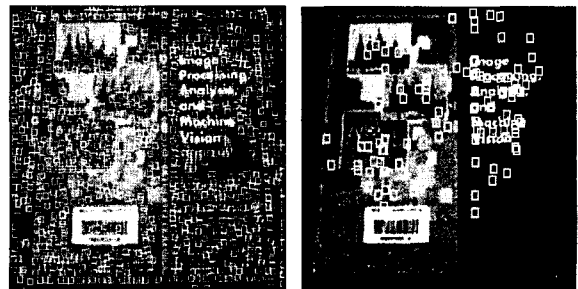


그림 2. 기존의 방법을 이용한 SIFT 특징점 추출 및 보완된 방법으로 추출된 SIFT 특징점 (오른쪽).

방위와 색차를 이용한 특징 구분 기법을 적용한 결과와 그렇지 않은 결과는 그림 3과 그림 4에서 비교 검토되었다. 그림 3에서 사각형 표시는 특징벡터에 대한 Euclidean 거리 측정의 결과가 왼쪽에 있는 책의 영역이라고 결정된 영역에 대응된다. 일부 추출된 특징점 중 인

식하려는 물체가 아닌 영역에서 검출된 false positive 가 있다. 방위 정보와 색차 정보를 사용한 기법은 매칭결과에서 검출된 특징 영역에 대하여 적용하였다. 이 방법은 2.3절에 기술한 후자의 문제점을 근본적으로 해결하진 못하지만, 서로 다른 색깔을 갖는 물체가 가진 방위와 색차 정보는 시스템이 잘못된 의사결정 확률을 줄인다. 그림 4에서 처럼 Euclidean 거리비교만으로 잘못 분류된 False positive 영역들이 2.3절에 기술된 방법을 사용한 그림 4에서 제거되었음을 볼 수 있다.

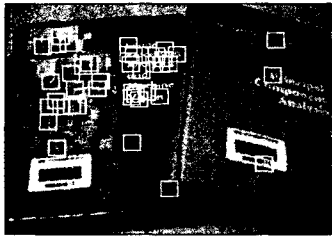


그림 3. 영상에서 추출된 특징벡터 중, Euclidean 거리 측정을 통해서 왼쪽의 책이라고 인식된 특징점의 위치.

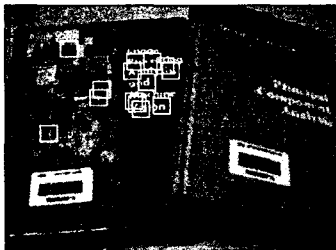


그림 4. 방위 및 색차정보를 사용하여 재분류 한 특징점의 위치.

#### 4. 결론 및 향후 연구 방향

본 연구에서는 [1] 에서 제안된 SIFT 기법상의 문제점인 불필요하게 많이 추출된 특징점 그리고 서로 다른 물체에서 생성된 유사 벡터문제 등에 대한 해결방안을 제안했다. SIFT 기법으로 추출된 특징점들 중, 대비가 작은 영역과 모서리와 같은 엣지가 강한 영역에서 상대적으로 불필요한 특징점들은 낮은 인식정보를 가지며, 특징 매칭 절차에서 의사결정을 복잡하게 하고 시스템 신뢰도를 떨어뜨린다. 상대적으로 높은 인식정보를 가진 특징점들을 그대로 유지하면서 그렇지 않은 특징점들의 제거는 로컬 영역에서의 기울기 해석으로 제거되었다. 향상된 SIFT 기법으로부터 추출된 특징점과 특징 영역에 대한 방위 및 색차 정보는 서로 다른 물체에서 생성 될 수 있는 유사 특징벡터를 효과적으로 구분함을 보였다.

추출된 로컬 영역을 다중의 세분화된 영역으로 구분하고 각 세부영역에서 생성된 특징벡터는 3차원 좌표변환에 강한 특성을 가진다 [10]. 그러나 추출된 특징 영역에 대응하는 특징 벡터를 생성하기 위해서 사용되고 있는 고정된 크기의 원형 윈도우 (fixed size circular window) 는 스케일의 현저한 변화 환경 아래서는 대상 영역에 포함되는 기울기 크기 및 방위정보의 차이로 그 특성이 유

지되기 어렵다. [2, 3, 11] 등에서 제안된 다중 스케일 기법과 해상도 변화 기법을 통한 특징점의 스케일 검출과 이에 따른 적응 원형윈도를 통한 특징영역의 해석이 해결 방법으로 고려 될 수 있다.

#### 5. 참고 문헌

- [1] D. G. Lowe, "Object Recognition from Local Scale-Invariant Features", *Proc. of the Int. Conf. on Computer Vision*, pp. 1150-1157, Corfu, Greece, 1999
- [2] T. Lindeberg, "Feature Detection with Automatic Scale Selection", *Int. J. of Computer Vision*, Vol.30, No.2, pp. 77-116, 1998
- [3] K. Mikolajczyk and C. Schmid, "Indexing based on scale invariant interest points", *Proc. of the 8th Int. Conf. on Computer Vision*, pp. 525-531, Vancouver, Canada, 2001
- [4] D. Jugessur and G. Dudek, "Local Appearance for Robust Object Recognition", *IEEE Computer Society Conf. On Computer Vision and Pattern Recognition*, Vol. II, pp. 1834-1839, S. Calolina, 2000
- [5] U. Hillenbrand and G. Hirzinger, "Object Recognition and Pose Estimation from 3D-Geometric Relations", *Fourth Int. Conf. On Knowledge-Based Intelligent Engineering Systems & Allied Technologies*, pp. 113-116, Brington, UK, 2000
- [6] D. Hall, J. L. Crowley, and C. de Verdiere, "View Invariant Object Recognition using Coloured Receptive Fields", *Machine Graphics and Vision Journal*, Special Issue, Vol.9, No.2, 2000
- [7] K. Ohba and K. Ikeuchi, "Detectability, Uniqueness, and Reliability of Eigen Windows for Stable Verification of Partially Occluded Objects", *IEEE Trans. on PAMI*, Vol. 19, No. 9, pp. 1043-1048, 1997
- [8] M. A. Magnor, "Geometry-based Automatic Object Localization and 3-D Pose Detection", *Proc. of IEEE Southwest Symposium on Image Analysis and Interpretation*, pp. 144-147, Santa Fe, USA, 2002
- [9] R. Ganger, S. Lazebnik, C. Schmid, and J. Ponce, "3D Object Modeling and Recognition Using Affine-Invariant Patches and Multi-View Spatial Constraints", *IEEE Computer Society Conf. On Computer Vision and Pattern Recognition*, Vol. II, pp. 272-280, June 18-20, Wisconsin, USA, 2003
- [10] S. Edelman, N. Intrator, and T. Poggio, "Complex Cells and Object Recognition", <http://www.ai.mit.edu/~edelman/mirror/nips97.ps.z>
- [11] Y. Dufournaud, C. Schmid, and R. Horaud, "Matching Images with Different Resolutions", *Proc. of the Conf. on Computer Vision and Pattern Recognition*, S. Carolina, USA, pp. 612-618, 2000
- [12] D. G. Lowe, "Local Feature View Clustering for 3D Object Recognition", *Proc. of the Int. Conf. on Computer Vision and Pattern Recognition*, pp. 682-688, Hawaii, USA, 2001
- [13] D. G. Lowe, "Distinctive image features from scale-invariant keypoints", preprint, 2003