

색상비율을 이용한 SIFT 성능향상

안보혁, 정종률, 최병욱
한양대학교 정보통신대학원 전자컴퓨터통신공학과
e-mail: abh3802@mlab.hanyang.ac.kr, bellaw@hanyang.ac.kr,
buchoi@hanyang.ac.kr

Improving Performance of SIFT Using Color Ratio

Bo Hyuck An, Jong Leul Chung, Byung-Uk Choi

Department of Electronics & Computer Engineering, Hanyang University.

요 약

효과적이고 정확한 물체인식은 컴퓨터 비전 연구 분야에 있어 매우 중요한 부분이다. 조명, 카메라 회전등의 외부 환경의 변화에 의해 서로 다르게 획득되는 영상에 대해서도 강인하도록 동일한 특징점을 추출하고 매칭할 수 있는 방법으로 SIFT(Scale Invariant Feature Transform)매칭이 많이 사용되어 왔다. 그러나 기존의 SIFT 기술자는 특징점 주변의 그레이만을 이용하여 기술하기 때문에 물체의 그레이정보가 유사하며 색상이 다르더라도 그레이 정보만 유사할 경우에도 매칭되는 단점이 있다. 이러한 문제점을 개선하기 위하여 본 연구에서는 기본영역과 확장영역의 색상 히스토그램에 기반 한 기술자를 추가하여 오매칭에 대한 인식 성능을 향상 시키는 방법을 제안한다.

1. 서론

컴퓨터 비전 연구는 영상의 정확한 이해를 목적으로 한다. 영상을 이해하는 방법 중 물체 인식은 중요한 부분을 차지한다. 이러한 물체 인식의 방법은 크게 형태 기반 인식 방법과 특징 기반 인식방법이 있다. 형태의 특징을 나타내는 방법은 데이터양이 방대하며 영상의 특징을 추출하기 전에 영상을 분할해야 한다는 단점이 있다. 이에 비해 특징 기반 인식 방법은 모델 특징들의 데이터양이 상대적으로 적은 장점과 시스템의 수행 속도가 빠르고 입력영상을 분할할 필요가 없다는 장점이 있다. 그러므로 본 논문에서는 영상에서의 특징을 기반으로 한 방법을 사용한다. 기존의 특징 기반 인식 방법 중 그레이의 Gradient를 사용하는 SIFT(Scale Invariant Feature Transform) 기술자를 많이 사용해 왔다. 이 방법은 물체내부의 특징점을 추출하여 특징점 주변의 국부 영역에 대한 그레이 Gradient 특징을 기술하고, 데이터베이스에 그레이 특징을 저장한다. 이때 입력영상이 들어올 경우 영상 특징점의 그레이 Gradient를 이용한 특징을 기술하며, 모델의 특징점 기술자와 입력영상의 특징점 기술자간의 비교를 통해 일치하는 특징점을 찾는다[1,2,3].

기존의 SIFT 기술자는 조명과 영상의 크기와 회전 변화를 고려한 기술방법이다. 특징에 기반 한 비전연구에서 정확한 영상의 이해를 하기 위해서는 점대점의 매칭에 있어서 정확한 매칭이 선행 되어야 된다. 하지만 SIFT매칭의 경우 특징점 주변의 그레이 Gradient만을 기술하기 때문에 색상이 다르더라도 매칭이 되는 단점이 있다. 이러한 오매칭을 줄이며, 정확도를 높이기 위한 방법으로 본 논문에서는 SIFT 매칭 결과를 판별하는 방법으로 Hue를 사용하여 색상 히스토그램을 이용하고자한다.

2. SIFT 특징점과 관심영역 검출 및 기술자 생성

2.1 SIFT 특징점 검출

SIFT는 영상에 대한 DoG(Difference of Gaussian)를 적용함으로써 후보점을 검출한다. 이후 찾아진 후보점들에 대하여 주변 화소들을 이용한 안정성을 검사하여 특징점을 검출한다[4].

2.2 관심영역 설정

관심영역은 스케일 공간상의 특징점이 해당하는 DoG 마스크영역을 선택하였기에 크기변화에 있어 변하지 않는 특징을 가지게 되며, 영역내부에서 주

변 상하좌우 화소를 이용하여 특징점의 방향과 크기를 가진 기준벡터를 생성한다. 이렇게 생성된 기준 방향은 영상의 회전이나 크기의 변화가 있더라도 항상 동일한 방향성을 갖기 때문에 영상이 회전에 대해 강인하다.



(a) 관심영역 및 기준방향 설정 (b) 크기와 회전변화가 있는 영상에서의 관심영역

그림 1. 특징점에서 기준벡터를 중심으로 한 관심 영역의 매칭

그림 1의 (a)는 특징점에서 기준벡터 방향으로 관심영역이 회전한 이미지이며, (b)는 (a)의 입력영상이 크기와 방향이 변하여 들어왔다 하더라도 동일한 영역과 기준방향을 갖는 것을 보여준다. SIFT 기술자는 선택된 관심 영역에 대해 기술한다.

3. 색상 변화도 기술자 생성

본 절에서는 SIFT 기술자를 이용한 매칭의 정확도를 높이기 위한 색상 기술자에 관하여 설명한다.

본 장에서 사용하는 변화도 기술자를 위한 2개의 영역 중 2장에서 설명한 관심영역을 기준영역이라 하며 기준영역을 중심으로 확장되는 영역을 확장 영역이라 하겠다. 색상정보를 표현하기 위해서 기준영역에서 확장영역으로 확장될 때 변하게 되는 색상의 관계를 색상 기술자로 사용한다.

기술자의 정보를 생성할 경우 하나의 관심영역에서만 색상정보를 확인하면 유사한 색상비율을 가지는 관심영역이 존재할 경우 동일한 정보를 갖는다고 판단하는 문제가 발생한다. 이러한 문제를 해결하기 위해 기준영역과 또 다른 하나의 영역인 확장영역을 사용하여 기술자를 생성하게 된다. 확장영역은 기준영역의 2배의 크기를 가지는 영역을 사용한다.

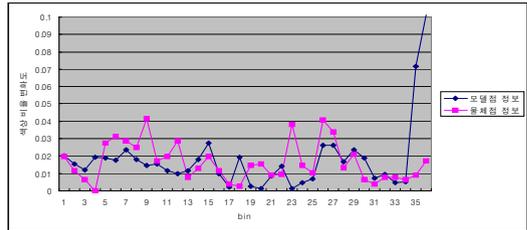
색상의 변화도를 나타내는 기술자를 생성하기 위해 각각의 영역에서 색상의 정규화 히스토그램을 생성하며, 색상의 비율이 영역이 확장되어가며 변하게 되는 색상의 관계를 나타내기 위해 기준영역과 확장영역과의 히스토그램의 차이를 사용해 색상 변화도

기술자를 기술하는 것이다.

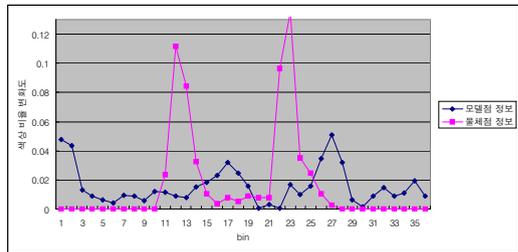
즉, 히스토그램간의 차이는 영역이 확장될 때 변하게 되는 색상의 변화도로 볼 수 있으므로 앞으로는 색상 변화도라 하겠다. 색상 변화도 기술자는 2장에서 검출한 크기와 회전에 강인한 영역 안에서 색상의 정보를 사용하기에 크기와 회전변화에 강인한 정보라 할 수 있다.

$$DH[i] = |EH[i] - H[i]| \quad (i = 1, 2, \dots, 36) \quad (1)$$

식 (1)에서 $H[i]$ 는 기준영역의 정규화 히스토그램을 나타내며, $EH[i]$ 는 확장영역의 정규화 히스토그램을 나타낸다. 영역간의 차분의 결과인 $DH[i]$ 는 영역이 확장되며 변하게 되는 색상 변화도 정보를 나타낸다.



(a) 매칭된 점의 주변 색상 변화도



(b) 오매칭된 점의 주변 색상 변화도

그림 2. 특징점 주변의 색상 변화도 예

그림 2의 (a)는 매칭된 점간의 DH 를 비교한 결과이며, (b)는 오 매칭된 점에 대한 DH 를 비교한 결과이다. 여기서, 모델점은 모델영상에서의 특징점에 해당하는 주변 색상 변화도정보 이며, 물체점은 입력영상에서의 특징점에 해당하는 주변 색상 변화도 정보이다. 그림 2의 (a),(b)에서 보는 것과 같이, 오 매칭된 점의 색상 변화도는 극명한 차이를 가지고 있음을 알 수 있다.

4. 유사도 비교

SIFT 특징점 매칭에 의해 매칭된 점이 오매칭된 점인지를 검증하기 위해서 3장에서 정의한 색상 변화도 기술자의 유사도를 계산한다.

$$\begin{aligned} DH_m &= \{\{r_{mi}\}, i = 1, 2, \dots, 36\} \\ DH_o &= \{\{r_{oi}\}, i = 1, 2, \dots, 36\} \end{aligned} \quad (2)$$

식 (2)의 DH_m , DH_o 는 각각 모델의 특징점과 입력영상에서 추출된 특징점에서의 색상 변화도 기술자이다. r 은 각각의 색상 요소들이다. DH_m 과 DH_o 의 유사도 계산은 Jia 등이 제안한 Gaussian Weighted Histogram Intersection (GWHI)을 사용한다. 이 방법은 동일한 bin에서만 값을 할당하는 것이 아니라 그 주변 bin에도 가우시안 값에 비례하는 가중치를 주면서 유사도를 계산하는 방법이다[5].

5. 실험

제안하는 방법의 성능을 평가하기 위하여, SIFT알고리즘의 실행 결과와 색상 변화도를 추가한 결과를 비교하였다.

성능평가의 척도로는 정확률(precision)과 재현률(recall)을 사용하였다.

$$\text{재현률} = \frac{\text{전체 매칭된 개수} - \text{오매칭된 개수}}{\text{모델의 특징점 개수}} \quad (3)$$

$$\text{정확률} = \frac{\text{전체 매칭된 개수} - \text{오매칭된 개수}}{\text{전체 매칭된 개수}} \quad (4)$$

식 (3)의 재현률은 모델의 특징 개수에 대한 정확한 매칭이 어떤 비율을 이루는가를 나타내며, 식 (4)의 정확률은 매칭된 개수 중 정확한 매칭이 어떠한 비율을 이루는가를 나타낸다.

제안하는 색상 변화도의 판별을 적용하면 오검출로 판별되는 부분이 제거되어 정확률은 상승하게 된다.

그림 3은 일반적인 환경의 물체에 대한 실험으로 컵과 축구화에 대한 SIFT결과와 색상 변화도 기술자를 이용하여 SIFT결과를 판별한 영상이다. 실험 물체에서의 오매칭 결과는 원형으로 표시하였다.

그림 3의 (a)는 SIFT 매칭을 수행한 결과로, 컵과 축구화로 매칭 되어야 될 부분이 주변 그레이 정보가 유사한 배경 부분의 특징점으로 오매칭됨을 확인할 수 있다. (b)는 색상의 변화도 기술자를 적용한 결과로서 주변의 색상 변화도 정보를 봄으로써 오매

칭이 제거되는 것을 볼 수 있으며, 그 결과를 흰색 선으로 표시하였다.



그림 3. (a) SIFT 매칭 결과 (b) 색상 변화도 기술자를 적용한 결과

표 1은 각 방법에 대한 정확률과 재현률을 수치적으로 비교한 결과이다.

성능평가를 위한 실험은 시작영상 28개, 회전영상 28개, 크기변화영상 28개로 총 84장의 이미지를 사용하였으며, 회전영상과 크기변화 영상은 시작영상에서의 크기의 변화와 회전을 적용한 영상을 의미한다.

표 1에서의 (a)는 SIFT결과에 대해 관심영역의 색상을 추가 하였을 경우 SIFT결과에 비해 변하게 되는 재현률과 정확률의 증가 감소를 확률값으로 나타낸다. (b)는 SIFT결과에 대해 기준영역과 확장영역에서의 색상 변화도를 추가 하였을 경우 SIFT결과에 비해 변하게 되는 재현률과 정확률의 증가 감소를 확률값으로 나타낸다.

표 1의 결과로서, 하나의 영역색상만을 사용하였을 경우 카메라의 회전등과 같은 외부 변화에 의해 서로 다르게 영상이 획득 되는 경우 색상은 많은 변화

가 발생할 수 있으므로 변화도 기술자보다 낮은 재현률을 가지게 된다.

표 1. SIFT결과와 색상정보 만을 비교한 그래프 & SIFT결과와 색상변화도 판별을 적용한 그래프

	(a) SIFT+색상 판별		(b) SIFT + 색상 변화도	
	recall	precision	recall	precision
시작 영상	-41%	32%	-28%	38%
회전 영상	-52%	66%	-38%	76%
크기 변화	-49%	65%	-32%	72%
평균	-47.33%	54.33%	-32.66%	62.00%

색상 변화도 기술자는 영역간의 색상의 비율이 변하는 값을 나타냄으로 영상의 크기변화와 회전의 변화가 있을 경우에도 영역의 변화도는 변하지 않기에 영역 색상 판별 방법보다 14.67% 높은 재현률을 가지게 된다. 정확률 또한 영역의 색상판별 방법보다 색상 변화도 기술자를 적용한 결과가 7.67% 높은 결과를 보여주며, 그 이유는 매칭점의 평가에 사용되는 색상 변화도 기술자방법이 영역의 색상만을 보는 방법보다 많은 정보를 표현할 수 있기에 높은 정확률을 값을 나타낸다.

6. 결론

SIFT 알고리즘은 모델 특징점 주변의 그레이 Gradient정보가 입력영상의 특징점 주변의 그레이 Gradient가 유사할 경우 오매칭이 발생한다. 이러한 경우 SIFT 매칭 성능을 향상시키기 위해 Hue를 이용한 색상 변화도 기술자를 사용 하였다. 또한 크기와 회전에 변하지 않는 특징 성능을 유지하기 위해 관심영역에서의 색상 변화도 정보를 사용하였다.

다양한 실험을 통해 살펴본 결과 재현률은 감소하게 되지만 정확률은 상승하는 결과를 볼 수 있었다. 또한 재현률이 줄어드는 크기에서도 색상 변화도를 볼 경우가 색상만을 볼 경우 보다 재현률의 변동이 적음을 볼 수 있다.

제안된 매칭 방법은 색상정보를 이용하여 SIFT의 오매칭을 제거하여 정확률이 향상 되었다. 이 방법은 색상의 정보를 구별해야 되는 물체인식 분야와 의료 등의 정확률을 목적으로 하는 분야에서 주목될

수 있을 것으로 기대된다.

7. Reference

- [1] T. Poggio and S. Edelman, "A Network That Learns to Recognize Three - Dimensional Object," Nature, Vol. 343, pp. 263-266, Jan 1990.
- [2] R. P. Rao, "Top - Down Gaze Targeting for Space-Variant Active Vision," In Proceedings of ARPA Image Understanding Workshop, pp. 1049-1058, Nov 1994.
- [3] C. Schmid and R. Mohr, "Combining Greyvalue Invariants with Local Constraints for Object Recognition," In Proceedings of Computer Vision and Pattern Recognition, pp. 872-877, jun 1996.
- [4] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, pp. 91-110, Jan 2004.
- [5] W. Jia, H. Zhang, X. He, and Q Wu, "Gaussian Weighted Histogram Intersection for License Plate Classification," In Proceedings of the 18th International Conference on Pattern Recognition, pp. 574-577, Aug 2006.