

인공객체영상 및 자연객체영상에 대한 자동 분류

구경모^{o*}, 박창민^{**}, 김민환*

* 부산대학교 컴퓨터공학과

** 영산대학교 컴퓨터정보공학부

Automatic classification for man-made object image and natural object image

Kyung-Mo Koo^{o*}, Chang-Min Park^{**}, Min-Hwan Kim*

* Dept. of Computer Engineering Pusan National University

** School of Computer & Information Engineering Youngsan University

Email : kmkool@pusan.ac.kr, cmpark@ysu.ac.kr, mhkim@pusan.ac.kr

요약

영상 분류는 내용기반 영상검색에서 부적절한 이미지를 효과적으로 걸러 낼 수 있게 하여 영상 검색의 성능을 향상 시키는데 큰 역할을 하게 된다. 따라서 최근 의미 있는 영상의 분류가 내용기반검색분야에서 중요한 과제로 대두되고 있다. 본 논문에서는 에지 정보를 이용해서 객체 영상을 인공객체영상과 자연객체영상으로 분류하는 방법을 제안한다. 직선형태의 에지를 많이 가지는 인공객체의 경우 에지 방향 히스토그램의 에너지가 자연객체에 비해 높은 값을 가지기 때문에 객체 분류에 유용한 정보로서 에지 정보를 활용하였다. 또한 에너지 값을 낮추는 원형의 에지가 인공객체영상에서 주로 발견되는 점을 이용하여, 제거에 의해 분류의 성능을 높이고자 하였다. 한편 가벼 필터를 이용한 분류 결과에 비해 에지 정보를 이용한 분류가 성능 면에서 보다 나은 결과를 얻을 수 있었다.

1. 서론

영상분류는 주어진 영상들을 의미 있는 영상 그룹으로 분류하는 것을 말한다. 이러한 영상분류는 사용자가 원하는 영상 그룹을 계층적으로 탐색할 수 있게 한다. 따라서 수 많은 영상들이 존재하는 영상 데이터베이스에서 영상들을 의미적으로 분류하는 것은 매우 중요한 일이다.

영상분류는 최근 몇 년 동안 내용기반영상검색(CBIR) 분야에서 많은 연구가 있어 왔다.[1-4] 하지만 이러한 연구는 실내/실외, 도시영상/풍경영상과 같이 영상의 전체영역이 영상의 내용을 표현하고 있는 전경영상에 대해서 주로 이루어져 왔을

뿐 특정 객체영역이 영상의 내용을 표현하는 객체 영상의 분류에 대한 연구는 미약한 실정이다. 그러나 실제로 영상검색에서, 사용자는 영상의 전체 영역이 아니라 일부 영역에 대해서만 관심을 가지고 질의를 하는 경우가 많다. 예를 들어 초원의 배경에 호랑이가 있는 영상이 질의로 주어졌을 경우, 사용자는 호랑이 영역에 관심을 가지며 배경에 관계없이 호랑이가 있는 영상이 검색 결과로 제공되길 원할 것이다. 이러한 경우 객체에 기반한 영상분류가 영상검색의 성능 향상에 중요한 역할을 하게 된다.

본 논문에서는 객체를 인위적으로 제조 되었거나

나 가공된 객체와 인위적인 조작이 없이 자연적으로 생성된 객체로 분류한다. 전자를 인공객체로 정의하며 도구, 운송수단, 건축물 등이 여기에 포함된다. 그리고 인공객체를 포함하는 영상을 인공객체영상이라 정의한다. 후자의 경우는 자연객체로 정의하며 동물, 식물 등이 여기에 속한다. 또한 자연객체를 포함하는 영상을 자연객체영상이라고 정의 한다. 그림 1 은 인공객체영상과 자연객체영상의 예를 보여준다.



인공객체영상의 예



그림 1 객체영상의 분류 예

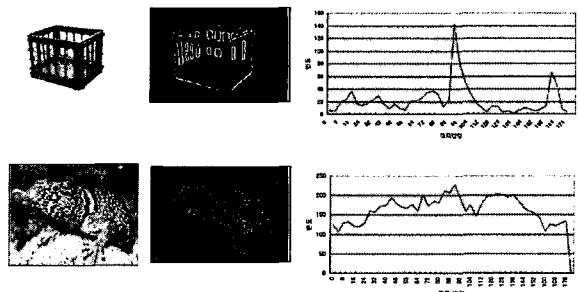
본 논문에서는, [5]에서 제안한 방법으로 객체영상에서 객체를 추출하고, 추출된 영역을 대상으로 에지 정보와 정보를 이용해서 인공객체영상과 자연객체영상으로 분류하는 방법을 제안한다. 분류된 결과는 영상 데이터베이스의 인덱싱에 유용하게 사용될 수 있을 것이다.

2. 객체영상 분류

2.1 에지 정보를 이용한 객체 영상의 분류

인공객체영상과 자연객체영상의 분류에 사용되는 특성으로서 객체영역의 에지 정보를 이용한다. 인공객체의 경우 객체 내부에서 직선 형태의 에지를 많이 가지고 있다. 이는 인위적인 제조 과정이나 가공의 흔적 때문이다. 반면에 인위적인 조작이 없는 자연객체의 경우 직선 형태의 에지보다는 자연스러운 곡선형태의 에지들이 많이 나타난다. 따라서 객체영역의 에지 방향 히스토그램을 보면, 인공객체의 경우 에지 화소가 특정 방향에 대해 높은 빈도를 가지는 반면, 자연객체의 경우에는 다양한 방향에 대해 고루 분포하는 것을 관찰 할 수 있다. 그림 2 는 대표적인 인공객체영상과 자연객체영상에 대해 객체영역의 에

지 영상과 에지 방향 히스토그램을 보여준다.



인공객체영상과 자연객체영상의 에지 영상과 에지 방향 히스토그램

그림 2 객체의 에지 영상과 에지 방향 히스토그램

객체의 경우 같은 종류일지라도 영상에서 보여지는 시점에 따라 에지의 방향이 달리 나타난다. 때문에 에지 방향 히스토그램 그 자체를 분류 특성으로 사용하는 것은 적절하지 않다. 따라서 에지 방향에 관계 없이 얼마나 많은 에지 화소가 밀집되는가를 측정하기 위해 히스토그램의 에너지 e 을 다음과 같이 계산하여 사용한다.

$$e = \sum_d f^2(d) \quad [1]$$

이때 d 는 에지의 방향을 의미하며 $f(d)$ 는 d 에서 빈도를 의미한다. 에지 추출을 위해서 Canny Operator를 사용했으며, 객체 추출의 어려요소를 감안해서 객체와 비객체간의 경계에 해당하는 에지는 배제했다.

그림 3 은 트레이닝 영상으로 사용된 인공객체들과 자연객체들에 대한 에지 방향 히스토그램 에너지의 분포이다.

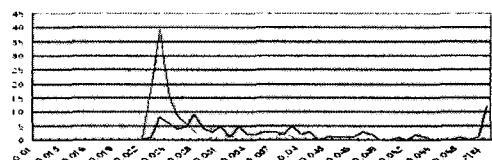


그림 3 에지 방향 에너지의 분포

그림 3에서 볼 수 있듯이 자연객체의 경우 전체의 80%정도가 0.02691을 기준으로 낮은 에너지를 가지고 있으며 인공객체의 경우도 전체의 80%정도가 높은 쪽에 분포하는 것을 볼 수 있다. 이것은 에지 방

향 히스토그램의 에너지가 인공객체영상과 자연객체 영상을 분류 할 수 있는 유용한 특성임을 보여준다.

인공객체영상과 자연객체영상에 대한 분류를 위하여 트레이닝 영상들에 대해 에러를 최소화 할 수 있는 에너지 σ 를 결정한다.

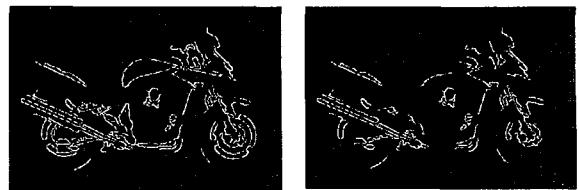
$$\sigma = \arg \min_{\sigma \in R} \{E_n^{(\sigma)} + E_m^{(\sigma)}\} \quad [2]$$

$E_n^{(\sigma)}$ 과 $E_m^{(\sigma)}$ 은 σ 에서 자연객체에서의 에러영상 개

수와 인공객체에서의 에러영상 개수를 나타낸다. 본 논문에서는 실험을 통해 σ 을 0.02691로 결정했다. 그리고 σ 를 기준으로 높은 에너지를 가지는 객체를 인공객체로 분류하며 낮은 에너지를 가지는 객체를 자연객체로 분류한다.

한편 인공객체 영상에서 원형을 띠는 에지가 에너지를 감소시키는 원인이 된다. 이러한 원형의 에지는 자연영상에서는 좀처럼 보기 힘든 특성이다. 때문에, 본 논문에서는 인공객체와 자연객체 간의 분류 정확성을 높이기 위해 원형의 에지를 제거 하는 작업을 선행한다.

원형의 에지를 찾기 위해서 에지 영상에 대해 Hough변환을 수행한다. 실험에서는, 추출된 원형을 구성하는 에지 화소의 개수가 실제 원주의 25% 이상일 때, 이를 원형 에지라고 판단하고 제거한다. 이 때 알고리즘의 특성으로 인하여 복잡한 에지를 가지는 객체의 경우 실제로 원형을 이루지 않는 에지들이 원형의 에지로 인식되어 제거 되는 경우가 있다. 이를 막기 위해서 일정 크기 이상으로 연결된 에지들에 대해서만 독립적으로 원형 에지를 추출하여 제거한다. 그림 4 는 원형의 에지를 포함하는 인공영상에서 원형 에지를 제거하고 난 후, 에너지가 증가하여 자연객체로 분류되었던 인공객체가 올바로 분류된 경우이다 이것은 원형 에지의 제거가 인공객체와 자연객체의 분류 정확성을 높이는데 기여한다는 것을 보여준다.



에너지 = 0.026

에너지 = 0.037

그림 4 원형 에지의 제거 후 증가된 에너지

2.2. 가버 필터를 이용한 객체 영상의 분류

가버 필터를 이용한 영상분류는 과거에도 있어왔다. [2]에서는 가버 필터를 이용한 파워 스펙트럼 템플릿으로 전경영상을 자연영상과 인공영상으로 분류했다. 예를 들어 고층 빌딩이 많이 나타나는 도시 영상의 경우 파워 스펙트럼의 형태가 특정 방향으로 나타난다. 이러한 특성은 영상을 도시영상과 자연영상으로 분류 할 수 있는 유용한 정보가 된다. 또한 이러한 점은 인공객체와 자연객체 분류에서도 유사하게 적용 될 수 있다. 그림 5 는 인공객체의 파워스펙트럼을 보여준다. 그림에서 볼 수 있듯이 인공객체 영상의 파워스펙트럼은 특정 방향으로 높은 값을 가진다. 본 논문에서는 이러한 특성을 이용해서 객체를 인공객체와 자연객체에 대해 분류해 보았다.

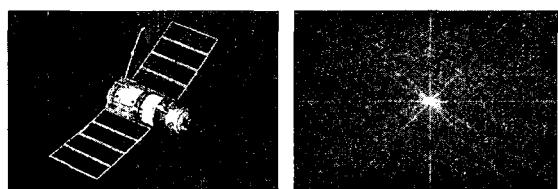


그림 5 인공객체의 파워스펙트럼

먼저 영상에 대해 비객체로 판단되는 영역의 화소를 $(R,G,B)=(0,0,0)$ 값으로, 객체영역의 경우 원영상의 밝기값으로 대체 한다. 이를 24개의 필터 뱅크 (scale =4, orientation = 6)로 구성된 가버 필터를 이용해서 24차원의 특성벡터를 만든다.

분류 방법은 테스트 영상에 대해 KNN(k-nearest neighbor : k=15)를 적용하여 다수결 원칙에 따른다.

3. 실험 및 토의

본 논문에서 제안한 방법은 PC에서 Visual C++로

구현하여 실험하였다. 트레이닝 영상은 Corel 영상에서 객체가 있는 영상 중에 임의로 인공객체영상 100개와 자연객체영상 100개를 선택하였으며, 테스트영상은 Corel 영상과 인터넷에서 얻어진 임의의 영상을 각각 100개씩 선택하여 실험 했다.

그림 6은 원형 예지 제거 전후의 트레이닝 영상 대한 예지 방향 히스토그램의 에너지 분포이다. 이것은 원형예지 제거가 분류 정확성 향상에 기여한다는 것을 보여준다.

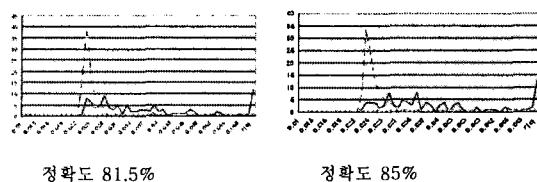


그림 6 원형 예지 제거 전,후의 에너지 분포

표 1은 테스트 영상에 대해 본문에서 언급한 두 가지 방법들을 이용해서 분류했을 때 나타나는 정확도를 나타낸다. 이는 정확성 측면에서 두 방법이 유사한 하다는 것을 보인다. 하지만 실제 연산 과정에서 소비되는 시간은 가버 필터가 더 많이 걸린다. 이는 정확성과 효율을 모두 고려했을 때, 예지 정보를 이용했을 때가 더 높은 성능을 가진다는 것을 보여준다.

표 1 분류 결과

	트레이닝 영상	테스트 영상
예지정보(원형예지제거)	85%	82.5%
가버필터	84%	83%

한편 원형 예지를 제거 할 경우, 자연객체에서도 예지의 일부분이 제거되어 자연객체의 예지 방향 히스토그램의 에너지가 전반적으로 소폭 상승되어 일부 자연객체가 인공객체로 잘못 분류되는 경우가 발생한다.

4. 결론 및 향후 연구

본 논문에서는 객체영상을 인공객체영상과 자연객체영상으로 자동 분류하는 방법을 제안하였다. 제안

된 방법을 객체영상에 적용해 본 결과, 인공객체영상과 자연객체영상을 분류하는데 어느 정도 만족할만한 결과를 얻을 수 있었다. 그러나 인공객체가 자연객체의 형태를 띠는 경우 또는 자연객체의 형태가 비교적 단순한 형태를 띠는 경우에는 예지 정보만을 이용해서 분류하기 어렵다. 따라서 향후 예지 정보 이외의 객체 내부의 질감 정보나 색상 정보와 같은 분류 능력이 있는 영상특성을 사용해서, 분류 성능을 더욱 높일 수 있는 방향으로 연구를 진행 할 것이다.

[참고 문헌]

- [1] A. Vailaya, M. Figueiredo, A. Jain, H. J. Zhang, "A Bayesian Framework for Semantic Classification of Outdoor Vacation Images," Proc. SPIE Storage Retrieval Image Video Database VII, vol. 3656, pp. 415-426, Jan. 1999
- [2] Aude Oliva, Antonio Torralba, Anne Guerin Dugue, Jeanny Herault, "Global Semantic Classification of Scenes using Power Spectrum Templates" CIR'99, Newcastle, UK, mars 1999
- [3] A. Vailaya, M. Figueiredo, A. Jain, H. J. Zhang, "Image classification for content-based indexing," Trans. On Image Processing, vol. 10, no. 1, pp. 117-130, Jan. 2001
- [4] M. Szummer, R.W. Picard, "Indoor-outdoor image classification," IEEE Int'l Workshop Content-Based Access Image Video Databases, pp. 42-51, Jan. 1989
- [5] S. Kim, S. Park, M. Kim, "Central object extraction for object-based image retrieval," Int'l conf, CIVR 2003, vol. 1, pp. 39-49, July, 2003.