방향성 특징 기술자를 이용한 식물 잎 인식

강수명^{*}, 윤상민^{**}, 이준재^{***}

요 약

환경의 변화에 따라 급속도로 변화하는 생태계에 대한 체계적인 연구를 위해 식물의 정보를 수집 분석하기 위한 연구가 활발하게 진행되고 있다. 특히, 스마트 기기의 카메라를 이용하여 언제 어디서나 사용자가 원하는 식물의 종류를 검색할 수 있는 기술에 대한 관심이 증가하고 있다. 본 논문은 식물 인식 및 생태계 분석을 위해 다양한 식물의 잎을 종류별로 분석할 수 있는 방법에 대해 제안한다. 이를 위해, 카메라로부터 입력된 식물 잎 사진의 관심 영역을 GrabCut을 통해 배경과 분리한 후, 형태 기술자 추출 방법인 SIFT(Scale-Invariant Feature Transform), HOG(Histogram of Oriented Gradient)를 이용하여 형태 기술자를 추출하고, 이것을 부호화 기법 및 공간 피라미드 방법을 이용한 분류 특징 벡터를 만든다. SVM(Support Vector Machine)을 통한 식물 잎 분류 및 인식한다. 다양한 식물 잎에 대한 실험 결과를 통해 비슷한 색상이나 형태를 가지고 있더라도 방향성 특징 기술자를 활용한 식물 잎 분류 방법이 매우 효율적임을 알 수 있다.

Plant leaf Classification Using Orientation Feature Descriptions

Su Myung Gang^{*}, Sang Min Yoon^{**}, Joon Jae Lee^{***}

ABSTRACT

According to fast change of the environment, the structured study of the ecosystem by analyzing the plant leaves are needed. Expecially, the methodology that searches and classifies the leaves from captured from the smart device have received numerous concerns in the field of computer science and ecology. In this paper, we propose a plant leaf classification technique using shape descriptor by combining Scale Invariant Feature Transform (SIFT) and Histogram of Oriented Gradient (HOG) from the image segmented from the background via Graphcut algorithm. The shape descriptor is coded in the field of Locality-constrained Linear Coding to optimize the meaningful features from a high degree of freedom. It is connected to Support Vector Machines (SVM) for efficient classification. The experimental results show that our proposed approach is very efficient to classify the leaves which have similar color, and shape.

Key words: Plant Recognition(식물 인식), LLC(지역 제한적 선형 부호화), Feature Extraction(특징 추출)

※ 교신저자(Corresponding Author): 이준재, 주소: 대구광역시 달서구 달구벌대로 1095 계명대학교 성서캠퍼스 공과대학 게임모바일콘텐츠학과 (704-701), 전화: 053) 580-6682, E-mail: joonlee@kmu.ac.kr

접수일: 2013년 8월 7일, 수정일: 2014년 2월 4일
완료일: 2014년 2월 6일

^{*} (주)지오씨엔아이 공간정보기술연구소
(E-mail: myang.gang@gmail.com)

^{**} 국민대학교 컴퓨터공학부

(E-mail: smyoon@kookmin.ac.kr)

^{***} 계명대학교 게임모바일콘텐츠학과

※ 본 연구는 *한국연구재단의 기초연구사업(2010-0023438)과 **기상청 차세대도시농림융합스마트기상서비스개발(WSIE)사업(153-3100-3133-302-350)의 지원으로 수행되었음.

1. 서 론

최근 레저 및 여가 활동의 증가로 인해 야외에서 산책나 들꽃 등의 식물을 접할 수 있는 상황이 빈번해짐에 따라, 식물에 관한 관심도가 높아지고 있다. 다양한 식물들을 생활 주변 곳곳에서 쉽게 접할 수 있으나, 식물들의 이름이나 관련 정보를 획득하기는 쉽지 않을 뿐 아니라 잘못된 정보로 인해 많은 오류가 발생한다. 따라서, 사용자 중심의 식물 검색 기술은 일반인들의 식물에 관한 이해도를 높일 수 있을 수 있으며, 웰빙(well-being), 에코(ecology) 등의 자연주의와 유기농 식품에 대한 관심을 높일 수 있다. 또한, 교육적인 측면에서도 일반인이나 어린 학생들도 다양한 식물에 대한 이해를 높이는데 기여할 수 있다. 그림 1은 주변 환경에서 획득한 식물 잎 영상이다. 이처럼 여러 가지 식물들은 인간의 생활 영역 곳곳에서 쉽게 접할 수 있으나, 그 식물들의 형태가 매우 유사하기 때문에 이름이나 관련 정보를 정확하게 분류하기는 쉽지 않다. 특히 문자 기반의 검색으로는 정확한 정보를 찾아내기 힘들기 때문에 최근 영상 기반의 검색을 사용한 영상 분류 방법이 대두되고 있다[1-5].

최근 여러 검색 엔진 등에서 제공하는 영상 기반의 검색은, 영상에 들어있는 모든 특징을 사용할 수 있다는 장점이 있어 사용자가 원하는 정보를 더욱 손쉽게 정확하게 얻을 수 있다. 또한, 영상 기술의 발전으로 고화질 영상 획득이 가능한 스마트폰뿐 아니라 Wi-Fi(Wireless Fidelity) 등의 통신 기술이 접목된 디지털카메라 등이 보급됨에 따라 영상의 획득 및 처리가 쉬워져 영상 기반의 검색 응용이 쉽게 되었다.

그림 1. 주변 환경에서 획득한 식물 잎 영상

컴퓨터 비전 분야에서는 주어진 영상의 특성을 분석하여 영상 인식 및 검색할 수 있는 기술에 관한 연구인 내용 기반 영상 검색 (CBIR, Content-based Image Retrieval)에 관한 연구가 활발히 진행되고 있다[6].

본 논문은 사용자가 언제 어디서나 카메라를 통해 얻은 식물의 잎의 특성을 분석하여 사용자에게 식물에 대한 정확한 정보를 제공하는 것을 목표로 한다. 본 논문의 구성은 총 5장으로 구성되어 있으며, 2장에서는 식물 분류와 관련된 기존 연구에 대해 알아보고, 새롭게 적용할 수 있는 기존 영상 분류를 위한 지역적 제한 선형 부호화 기법에 대해 설명한다. 3장에서는 입력 영상에 대해 GrabCut 방법을 이용하여 관심영역을 추출한 후, SIFT 방법과 HOG 방법을 이용하여 특징을 추출한다. 다음 부호화 기법 및 공간 피라미드 방법을 이용하여 특징 벡터를 만들고 SVM을 통해 분류하는 방법에 대해 설명한다. 4장에서는 실험을 위한 데이터 구성과 분류 성능에 대해 알아보고, 마지막으로 5장에서는 본 실험으로 얻은 결론 및 연구되어야 할 앞으로 연구할 방향에 대해 언급한다.

2. 기존의 연구

식물 분류를 위한 모양의 특성 분석을 위해 식물의 형태학적 특성을 이용한다. 형태학적 특성은 어떠한 생물체의 모습이나, 내부 구조 등에서 외관 비율, 직사각형적 비율을 이용한다. 그림 2는 식물 잎을 통해 얻을 수 있는 여러 가지 형태 정보를 보여준다[2-4].

이러한 특성을 결합하여 분류한 Du 등 [2]의 제안에서는, 입력 영상을 회색 조 영역으로 변환하여 이진화 영상과 외곽선을 추출하고, 이 외곽선 영상을 통해 몇 가지 특성을 이용하여 특징들을 추출하였다. 이를 이용하여 MMC(Move Median Centers) 분류기를 통해 특징들을 군집화하였다. 이 실험은 속도 측면에서 이전 연구보다 장점을 보이고, 특성을 결합하는 것이 좋은 결과를 보이나, 인식률이 90% 정도라서, 형태학적 특징의 결합으로는 유사한 형태를 가지고 있는 다양한 식물의 분류에 어려움이 있다[2-4].

따라서, 본 논문에서는 형태 정보의 특성을 검출할 수 있는 특징점 검출 기술과 유사도 비교 방법에 대해 초점을 맞추어 연구한다. Bama 등 [5]은 영상에 관한 인식 및 검색 기술에 관한 연구인 CBIR을 식물

그림 2. 식물 잎을 통해 얻을 수 있는 여러 가지 영상

잎 분류 기법에도 적용하였다. Bama 등 [5]은 식물 잎 영상에서 SIFT 기술자를 추출하여 식물 잎 분류에도 적용 될 수 있음을 알 수 있다[5].

2.1 SIFT(Scale-Invariant Feature Transform)

Lowe [7]에 의해 제안된 SIFT 알고리즘은 불변성이 강한 지역적 특징이기 때문에, 크기가 다르거나 회전된 식물 잎이라도 특징은 동일하게 추출된다. 추출 과정은 다음과 같다. 우선 스케일 공간을 생성하여 그 공간에서 작업함으로 크기 불변성 특징 검출을 가능하게 한다. 스케일 팩터인 표준편차 σ 를 상수 k 배만큼 증가시키면서 가우시안 함수 $G(x, y, \sigma)$ 와 입력 영상 $I(x, y)$ 로 컨볼루션 된 스케일 공간의 영상 $L(x, y, \sigma)$ 을 얻는다. 그 식은 다음과 같다.

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

$$\text{where } G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$

이 스케일 상의 불변한 특징점을 얻기 위해 DoG(Difference of Gaussian)함수를 k 배 증가한 표준편차를 사용하여 얻은 L 과 증가하지 않은 L 의 차이로 정의한다. 그림 3은 이를 스케일 변화에 따라 도식적으로 표현한 것이다[7].

그림 3. 입력 영상으로부터 얻어진 DoG

이때, DoG의 모든 점에 대해서 극점(maxima 혹은 minima)의 조건을 검사하여 통과한 점은 후보 특징점이 된다.

$$m(x, y) = \quad (3)$$

$$\sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (2)$$

$$\theta(x, y) = \tan^{-1}(L(x+1, y) - L(x-1, y) / (L(x, y+1) - L(x, y-1)))$$

식 (2)에서 $m(x, y)$ 와 $\theta(x, y)$ 는 주변 화소의 기울기의 크기를 이용해서, 특징점에 관한 크기와 방향 정보를 의미한다. 이렇게 얻어진 특징점의 주위에서 그림 4와 같이 소구역에 대해 기울기를 계산해서 히스토그램을 구한다[7].

SIFT 특성은 계산 복잡도가 높아서 연산 속도가 느리지만, 영상 내부 관심 영역의 위치나 크기, 회전에 불변한 특징을 얻기 때문에 비교적 좋은 결과를 얻어낸다.

그림 4. 특징점을 중심으로 기울기 계산

2.2 HOG(Histogram of Oriented Gradient)

Dalal 등 [8]에 의해 제안된 HOG는 국소 영역의 밝기에 관한 기울기 분포를 영역마다 히스토그램으로 만들어 하나로 합친 후 특징으로 사용하는 기술자이다. 기존의 연구에서도 식물 잎의 외곽에서 얻을 수 있는 많은 특징으로 식물을 인식하고 분류하였다.

식물 잎 외관의 모양에 중요한 정보가 많으며, 식물 잎의 형상이 중요한 특징이 될 수 있다. 따라서 형상 정보를 이용하는 HOG 특징 기술자가 식물 잎의 분류에 유용하게 사용될 수 있다.

HOG 특징을 구하기 위해서, 먼저 입력 영상 I 를 $n \times n$ 크기의 국소 영역으로 자른다. 이렇게 잘린 국소 영역에 대해 기울기를 구할 수 있는 두 방향의 커널을 거쳐 각 기울기 영상을 구한다. 이 기울기 영상의 각 영역에서 기울기 히스토그램을 구하게 된다. 계산된 기울기를 X° 도 간격으로 기울기의 크기인 m 을 누적한 히스토그램을 구한다.

$$m(i, j) = \sqrt{G_x(i, j)^2 + G_y(i, j)^2} \quad (3)$$

$$\theta(i, j) = \tan^{-1} \frac{G_y(i, j)}{G_x(i, j)}$$

이때 구해진 각 국소 영역의 히스토그램을 하나의 특징 벡터 v 로 보고, 정규화하여 관심 영역의 크기에 대해 불변한 특성을 있도록 한다. 정규화된 각 국소 영역의 히스토그램을 하나로 이어 붙인 벡터로 만든다. HOG는 영상의 형상 표현에 대해 강인하다는 장점이 있다. 또한 연산량이 SIFT보다는 상대적으로 적기 때문에 훨씬 빠른 수행결과를 보이지만, 영상의 회전에 관계하여 불변하지 않다는 단점이 있다. 그림 5는 입력 영상에 대한 기울기의 크기 영상과, 소구역에 누적된 기울기 방향을 보여준다[8].

그림 5. 입력 영상에 관한 HOG 추출 영상

2.3 영상 분류를 위한 지역적 제한 선형 부호화 기법

최근 영상 분류 방법들은 BoF(Bag of Feature 혹은 Bag of Words) 기반의 공간 피라미드 매칭(SPM, Spatial Pyramid Matching) 방법이 좋은 결과를 보

여준다[9-15]. BoF 방법은 특징 기술자들을 재표현하기 위하여 코드 워드(code word)로 구성된 코드북(code book)이라는 지역 특징 집합을 사용한다. 입력된 모든 영상에서 추출된 모든 특징기술자에 대해 비슷한 기술자들끼리 묶어줄 수 있는 코드 워드를 만든 후, 이를 모아놓은 집합을 코드북이라고 한다. 그 후 코드북의 몇 가지 코드 워드로 분류할 영상을 재표현한다. 그림 6은 여러 특징 벡터를 코드 워드로 재표현한 집합을 보여준다. 그림 7에서는 미리 구성된 코드

그림 6. 입력 영상에 관한 코드북 구성

그림 7. 코드북을 이용한 입력 영상 인식

그림 8. 일반적인 부호화의 과정 모식도

그림 9. SPM을 통한 특징 히스토그램 생성

북으로 입력 영상을 재표현하는 모식도이다[9-15].

SPM은 코드북을 생성할 때 지역적인 정보를 고려해 영상을 소구역으로 나누고 난 후, 그곳에 위치한 특징 기술자로부터 특징 벡터를 그림 8과 같이 만든다[9-15].

이를 구현하는 방법 중, Wang 등 [12]에 의해 제안된 지역적 제한 선형 부호화(LLC, Locality-constrained Linear Coding)는 선형 분류기를 사용하여 계산 복잡도를 감소시키고, 특징 기술자들을 부호화할 때 코드북의 지역적 거리 가중치의 중요성을 강조한 지역 좌표 부호화(LCC, Local Coordinate Coding) [13]를 개선하여 만든 알고리즘이다. 그림 9는 한 영역 내, 각 코드 워드의 히스토그램을 생성하는 도식이다[12].

입력 영상에서 추출한 SIFT 같은 특징 정보를 부호화하는 기존 방법은 벡터 양자화(VQ, Vector Quantization), 희소 부호화(SC, Sparse Coding), LLC 등이 있다. 그림 10은 이것들을 비교하여 보여준다[12].

그중 가장 간단 한 것은 벡터 양자화 방법이며, 모든 부호화의 기본이 된다. 이렇게 부호화된 코드 워드를 SPM을 통해 하나의 벡터로 만들어 영상의 특징으로 사용하게 된다. VQ는 SIFT나 HOG 등을 통해 만들어진 특징 기술자들에 대해서 최소자승법을 이용해 가장 지역적으로 가까운 코드 워드 값을 대응하는 과정으로 그 식은 다음과 같다[11-14].

$$\arg \min_c \sum_{i=1}^N \|x_i - Bc_i\|^2 \quad (4)$$

$$s.t. \|c_i\|_{\ell^0} = 1, \|c_i\|_{\ell^1} = 1, c_i \geq 0, \forall i$$

이렇게 만들어진 코드는, 특징 값과 코드북의 코드 워드가 1 대 1 대응을 하므로 양자화 오류가 많을 가능성이 있다는 점에서 한계를 가진다. 벡터 양자화의 오류를 해결하기 위해, 오직 하나의 대응하는 코드를 여러 개 가지도록 하는 방법이 희소 부호화이다. 이 방법을 통해서 만들어지는 코드는 적은 개수의 코드를 가질 수 있으므로 한 개의 대응 코드를 가지는 벡터 양자화보다 오류가 줄어들게 된다. 그 식은 다음과 같다[15].

$$\arg \min_c \sum_{i=1}^N \|x_i - Bc_i\|^2 + \lambda \|c_i\|_{\ell^1} \quad (5)$$

하지만 이렇게 여러 개의 코드 값을 가진다 하더라도, 단지 개수만 고려 부호화를 한 것일 뿐 코드 간의 지역적인 연관성이 떨어져 성능의 제한이 있다.

Yu 등 [13]이 제안한 LCC의 방법에서는 코드북에서 최소화해 코드를 택하여 부호화하는 Yang 등 [15]이 제안한 기존의 ScSPM(Sparse Coding SPM) 방법을 개선하여 지역적 거리의 가중치를 덧붙여 고려하였다.

특징 기술자 집합을 $X = \{x_1, x_2, \dots, x_N\} \in \mathbb{R}^{D \times N}$ 라고 하고, 미리 구성된 코드북의 집합을 B, 기술자 집합 X와 코드북 집합 B의 대응으로 만들어진, 입력 영상의 코드북은 $C = \{c_1, c_2, \dots, c_N\} \in \mathbb{R}^{M \times N}$ 라고 할 때 비용함수는 식 (6)과 같다.

$$\arg \min_c \sum_{i=1}^N \|x_i - Bc_i\|^2 + \lambda \sum_j |c_{ij}| \|x_i - b_j\|^2 + \mu \sum_j \|b_j\|^2 \quad (6)$$

그림 10. 여러 가지 부호화 방법 비교

식 (6)의 마지막 항은 기존 최소 부호화의 방법에, 특징 기술자와 코드북의 코드 워드 간의 거리를 함께 고려하여 부호화에 정확성을 높였다. 식(5)는 계산적으로 비용이 높은 L1-norm의 최적화 문제를 풀어야 하는 데 비해, 이를 개선한 LLC의 경우 계산 복잡도가 줄어들기 때문에, 실행 속도가 빠르다. 따라서 사용자가 식물 잎에 관한 정보가 필요한 당시에 영상을 획득하고 처리하여 정보를 얻기 위해서는 처리의 속도가 중요한 고려 사항이 되므로 LLC 방법은 본 논문에서 지향하는 실제 응용을 위한 효과적인 방법이다.

LCC의 계산 비용 문제를 해결한 LLC는, LCC의 후반부 항을 단순히 특징점과 코드 워드 간의 거리만을 고려하도록 단순화한 것이 가장 큰 특징으로 이의 비용함수는 식 (7)과 같다.

$$\arg \min_c \sum_{i=1}^N \|x_i - Bc_i\|^2 + \lambda \sum_j^M \|d_j \odot c_j\|^2 \quad (7)$$

$$s.t. 1^T c_i = 1, \forall i$$

이 식에서 $\odot$ 는 원소 간의 곱을 뜻하고, $d_j \in \mathbb{R}^M$ 은 지역적인 특성에 기여하는 부분인데, 입력된 특징점과 코드 워드의 유사도를 계산한다. d_j 는 입력 x_i 와 코드북 B의 각 항과 유클리드 거리를 구하여 지역적인 영역에 관한 가중치 감쇄 속도를 조절하는 용도이다.

입력 영상의 부호화 된 집합은 풀링(pooling) 과정을 통해 하나의 대표될 수 있는 벡터로 만들어진다. 이것은 선형 SVM의 입력 벡터가 되어, 훈련 과정과 테스트 과정을 거쳐 입력 영상이 어떤 영상인지 알 수 있는 근거가 된다.

3. 제안하는 방법

이 장에서는 식물 잎을 분류하기 위하여, 식물 잎의 형태 분석 기술자만을 통해 인식률을 높이는 방법을 제안한다. 먼저 자연환경에서 잎 영상을 획득한 후에, GrabCut [16]을 이용한 관심영역을 추출한다. 추출된 관심 영역은 회색 조 영상 등으로 전처리를 통하여 SIFT나 HOG를 이용한 특징기술자를 얻는다[7-8]. 이 특징기술자들을 부호화하여 영상마다 대표할 수 있는 특징 벡터로 만들어 풀링을 통해 분류기에 입력할 수 있는 입력 벡터를 생성한 후, SVM 분류기를 통해 이를 분류하여 입력한 잎 영상을 인식한다[17].

3.1 영상 획득 및 GrabCut을 이용한 관심영역 분할

이 장에서는 자연환경에서 획득한 식물 잎 영상에 대해 필요한 관심영역 분할 및 전처리 과정에 대해 설명한다. 획득한 영상에 대해 전처리를 하는 과정은 잎의 특징 추출하기 전에 필수적인 과정으로, 관심영역을 제외한 배경 영역을 제거함으로써 다음에 이어질 영상 처리의 시간적 효율을 위해 필수적인 과정이라고 할 수 있다.

스마트 폰 등의 휴대기기로 그 정보가 필요한 장소에서 바로 획득된 영상은 관심영역으로 분할되어야 한다. 영상의 배경에 해당하는 비 관심영역을 제거하면 좀 더 정확한 정보를 추출할 수 있는데, 이때 Rother et al. [16]이 제안한, Graph cuts를 기반의 사용자 인터페이스를 개선한 GrabCut 방법을 사용한다. 그림 11은 Graph cuts와 Grabcut의 방법을 비교한 것을 보여준다.

기존의 연구에서는 여러 가지 방법으로 원하는 식물 잎의 부분을 분할하였지만, GrabCut이라는 영상 분할 방법은 사용자의 관심영역은 전경, 그 외의 배경으로 분류하여 영상을 생성한다. 이후, 가우시안 혼합 모델(GMM, Gaussian Mixture Model)으로 모델링하여 에너지 최소화를 반복적으로 수행하게 된다[16]. 이 방법은 조작하기 쉽고 성능이 높지만, 연산 비용이 높아서 수행 속도가 느리다. 그림 12는 사각형 모양의 사용자 인터페이스를 통한 처리과정을 보여준다.

이렇게 GrabCut 알고리즘을 수행하면 그림 13에서처럼 하나의 잎만 분할된 영역으로 추출할 수 있다. 식물 잎은 꽃과 같이 색깔 특징이 명확히 구분되

그림 11. 관심영역 추출 방법 비교

(a) 획득된 영상 (b) 관심영역 선택

그림 12. 자연환경에서 얻은 잎 영상의 처리

그림 13. GrabCut을 통해 획득된 관심영역

는 물체에 비해 색깔 특징은 중요하지 않다. 획득된 영상에 대해 다음 식을 통하여 회색 조 영상을 구한다.

$$Gray = I_r \times 0.2989 + I_g \times 0.5870 + I_b \times 0.1140 \quad (8) \quad (8)$$

이를 통해, 그림 14와 같이 회색 조 영상이나, 이진화 영상, 혹은 외곽선 등을 추출하여 컴퓨터가 인식할 수 있는 특징 정보를 획득하여 식물 잎의 종류를 분류할 수 있게 된다.

(a) 회색 조 영상 (b) 이진화 영상 (c) 외곽선 추출 영상

그림 14. 관심 영역에 대한 전처리

3.2 SIFT, HOG 특징을 이용한 지역적 제한 선형 부호화 기법

이 장에서는 어떻게 식물의 잎으로부터 특징을 추

출하여 부호화하는지에 대해 기술한다. 본 논문에서는 특징을 추출하기 위한 여러 가지 기술자 방법 중 가장 대표적인 SIFT와 HOG를 사용한다[7-8]. 관심 영역을 획득한 영상으로부터 얻어진 특징벡터를 미리 구성해놓은 코드북과의 대응을 통해 부호화하고, 부호화된 특징 중 특징점마다 위치를 가지는 SIFT는 SPM을 통해 분류기에 입력될 벡터를 만든다.

3.2.1 SIFT 알고리즘을 이용한 특징 벡터 추출

입력 영상의 DOG에서 얻어진 특징점의 주변 구역에 대해 각 8방향의 기울기를 하여 히스토그램을 누적한다. 이 과정으로 구한 벡터는 방향성에 대해 불변하며, 특징점을 중심으로 4×4 의 구역에서 8방향의 히스토그램을 누적하므로 128개의 특징을 가지게 된다. 그림 15는 입력 영상에 대해 SIFT를 추출한 결과를 보여준다.

(a) 회색 조 영상 (b) SIFT 특징점 영상

그림 15. 잎 영상에 대한 SIFT 처리

3.2.2 HOG 알고리즘을 이용한 특징 벡터 추출

SIFT와 마찬가지로, 입력된 식물 잎 영상을 회색 조로 만든 후 각각 기울기 영상과 기울기의 크기 영상을 구한다. 본 논문에서는 이 과정에서, 기존의 HOG에서 사용하던 $[-1, 0, 1]$ 과 $[-1, 0, 1]^T$ 커널보다 훨씬 외곽선이 뚜렷하게 나타나는 Sobel 방법을 사용하여 영상을 처리하였다. 이 실험에서는 영상을 3×3 개로 나눈 후, 각 구역에 대해 32방향의 기울기를 누적하였다. 따라서 한 영상에 대해 총 288개의 특징을 가진다고 할 수 있다. 그림 16은 입력 영상에 대해 기울기 영상과 기울기의 크기를 추출한 것을 보여준다.

그림 17은 잎 영상의 기울기 크기를 32방향으로 누적된 히스토그램을 나타낸다. 하지만 HOG는 영상의 위치나, 회전에 대해 불변하지 않은 특성이 있기 때문에 회전각 보정을 통해 최대한 같은 위치와 회전각을 맞추어서 실험을 진행하여야 한다. HOG의 경

그림 16. 입력 영상에 대한 HOG 처리

그림 17. 누적된 히스토그램

우, 영상을 국소 영역으로 분할하여 기울기 히스토그램을 누적하여 만든 벡터로 이루어진 특징이므로 SPM의 방법을 적용할 수 없다.

3.2.3 지역적 제한 선형 부호화 및 분류

앞서 얻은 특징벡터에서, 이를 대표할 수 있도록 미리 만들어놓은 코드북 데이터베이스와 대응시켜 각 영상의 특징벡터를 부호화하여야 한다. 모든 입력 영상에 대해 추출한 특징 벡터를 k-means 알고리즘을 통해 군집화한 후 중심 벡터를 대표로 두어 코드 워드로 정의하여 코드북을 만든다. 이를 입력 영상의 특징벡터와 대응시켜 부호화하는 과정을 거친다. 실험에서는 코드북 수가 512보다 높을 때 좋은 결과를 나타내었다. 이 코드북으로 특징벡터를 부호화하여 이를 풀링하는 과정을 거친다.

3.2.4 풀링 및 SVM 분류

풀링하는 과정에서 SPM의 방법에 따라, SIFT에서 얻은 특징은 특징점마다 위치 값이 반환되기 때문에, 전체 영상에서 피라미드 형태 즉, 1×1 , 2×2 , 4×4 의 각각 소구역에 관한 특징벡터들을 누적하게 된다. 이렇게 각 영상에서 획득한 특징벡터를 부호화하였다면, 이를 분류하는 동시에 검증과정을 거쳐야 한다.

이때 사용되는 것이 SVM(Support Vector Machine)이다. Cortes 등 [17]에 의해 제안된 SVM은 부류마다 존재하는 여유치(margin)을 최대화하여 일반화 능력을 극대화하는 방법을 통해 분류하는 분류기법이며, 여타 다른 분류기에 비해 우수한 성능을 자랑한다.

4. 실험 결과

4.1 실험 자료 구성

본 실험에서는 식물 잎과 배경과 분리하는 전처리 과정에 대해, 앞서 GrabCut을 수행한 뒤에 전처리 과정을 가진다고 언급하였다. 각 영상에 대해 GrabCut은 대부분 우수한 결과로 관심영역을 추출한다. 그림 18은 잎 영상에 대해 GrabCut 알고리즘을 적용했을 때 수행 결과를 보여준다.

따라서 모든 입력 영상에 대해 GrabCut을 수행했다는 가정 아래, Wu 등 [18]이 실험한 Flavia에서 제공하는 영상 자료 집합을 사용하였다. 그림 19는 제공된 영상 자료 집합 중 몇 개의 예시를 보여준다. 이 자료 집합은 각 32종에 대한 분류가 되어있지 않

그림 18. 잎 영상에 관한 GrabCut 수행 결과

그림 19. Flavia 데이터베이스의 식물 잎 영상 예제

표 1. SIFT 특징 기술자를 통한 식물 분류 실험 결과

카테고리	카테고리별 영상 수	SVM	코드북 샘플링 개수	코드북	SPM 공간	정확도(%)	오차범위
32종	50장	영상 30장 랜덤 선택 10번 수행	2000	1024	1	0.92437	0.01070
					2	0.98062	0.00277
					4	0.98703	0.00295
			10000		1	0.95812	0.00476
					2	0.98281	0.00392
					4	0.98968	0.00392

고, 개수가 맞지 않아 실험의 정확도를 높이기 위하여 모든 영상에 대해 같은 종류의 잎을 분류한 후, 한 종류마다 각 50장씩 자료 집합을 구성하였다. 이때 각 영상의 크기를 모두 가로 400, 세로 300의 크기로 일정하게 축소하여 실험을 진행했다.

실험은 각 특징 기술자에 관한 비교를 중점적으로 하였으므로 회전과 위치 변화 특성에 강인한 SIFT와는 달리, HOG는 회전 변화에 불변하지 않으므로, 미리 구성한 각 종류 자료 집합의 식물 잎은 먼저 일괄적으로 회전을 갖게 보정한 후 실험하였다.

4.2 분류 성능

4.2.1 SIFT 특징 기술자를 이용한 분류 방법

이 실험에서는 Flavia [18]에서 제공하는 잎 영상을 이용하여 먼저 SIFT 기술자를 추출한다. 주어진 영상의 각 특징점을 중심으로 4×4 크기 영역을 만들어 각 8방향의 기울기를 계산한 128개의 기술자를 히스토그램으로 누적한다. 이렇게 누적된 SIFT 단일 벡터를 코드북과 대응되는 코드 워드로 변환하여 SPM에 따라 1×1 , 2×2 , 4×4 로 구성하였다. 전체 1,600개의 자료 집합에서 k-means 클러스터링을 이용해 구성된 1024 크기의 코드북과 입력 인자를 대응시켜 최대 풀링 한 것을 입력 벡터로 사용하여 선형 SVM을 통해 분류한다.

HOG 특징기술자는 SPM에 의한 풀링이 불가능하므로 SIFT 기술자와의 정확한 비교를 위해 SPM 공간을 점차 증가시키는 방법으로 각 공간이 늘어날 때마다 정확도와 처리 및 분류 시간에 대해 중점적으로 확인하였다. 이 분류기에 입력되는 벡터를 각 입력 영상에서 얻어진 특징 벡터의 부호화된 벡터로 보고, 본 실험에서는 우연한 경우에 결과 값이 좋아질 수 있는 경우를 배제하기 위하여 모든 데이터 집

합에서 10번 무작위의 실험을 하여 평균적인 결과 값을 얻었다. 이 실험에서 사용한 PC 사양은 2010년형 iMac, Intel core i3 3.07GHz 프로세서와 4GB RAM, 32비트 Windows 7 운영체제를 사용하였다.

표 1의 실험결과를 보면, SIFT와 부호화만으로도 식물 잎에서 얻을 수 있는 형태학적 특징을 결합한 것 보다 우수한 결과를 보인다는 점을 알 수 있다. SPM의 공간이 늘어날 때마다 정확도가 향상하는 것으로 보아, 정확도에 SPM이 큰 기여를 한다는 것 역시 알 수 있다. 하지만 표 2에서 보면, SIFT 기술자 추출 시간이 늦고 데이터베이스를 읽는 시간 역시 오래 걸리는 것으로 보아 실제 응용에서 처리 시스템이 좋지 않은 경우에 사용하기 어려울 것으로 보인다.

표 2. 처리 영역별 수행 속도 (단위: sec)

구 분	SIFT 추출	분 류
속 도	2.071	2.291

4.2.2 HOG 특징 기술자를 이용한 분류 방법

이전 실험과 동일한 조건에서 HOG 특징 기술자를 이용한 방법으로, 32종의 식물 잎을 각 50장씩 자료 집합을 구성한 후 실험을 진행하였다. 모든 입력 영상에 대해 기울기를 구할 수 있는 두 방향의 $[-1, 0, 1]$ 과 $[-1, 0, 1]^T$ 커널을 대신하여, Sobel 커널을 사용한 기울기 영상을 $n \times n$ 의 국소 영역으로 나눈 이후에 각 영역에서 계산되어 반환된 기울기 $-\pi$ 부터 π 를, 11.25° 간격으로 하나의 국소 영역에 대해 32개의 방향성으로 기울기의 크기를 누적한 히스토그램을 구하였다.

국소 영역에 대해 방향성의 기울기를 누적할 때, 입력 영상의 관심 영역인 식물 잎 부분과 관계가 없는 흰색 배경까지 함께 누적하면 결과에 나쁜 영향을

표 3. HOG 특징 기술자를 통한 식물 분류 실험 결과

카테고리	카테고리별 영상 수	SVM	코드북 샘플링 개수	코드북	국소영역	정확도(%)	오차범위
32종	50장	영상 30장 랜덤 선택 10번 수행	2000	512	3×3	0.87156	0.01266
				1024		0.77578	0.00914

줄 수 있다는 것을 여러 번의 실험을 통해 알게 되었다. 따라서 기울기를 누적할 때는 흰색 배경은 제외하여 식물 잎의 기울기를 누적하였다.

1,600개의 자료 집합에서 k-means 클러스터링을 이용해 구성된 1024 크기의 코드북과 입력 인자를 대응시켜 최대 풀링 한 후, 선형 SVM을 통해 분류를 수행하였다. 이때, HOG 특징 기술자는 SIFT 기술자와는 달리 특징점의 위치가 없으므로 각 SPM 공간을 점차 증가시키는 방법에는 효과가 없다. 따라서 코드북 수와 국소 영역의 개수에 따라 정확도와 처리 및 분류 시간에 대해 중점적으로 비교하였다. 이 실험에도 10번의 무작위의 실험을 하여 평균적인 결과 값을 얻었으며, 같은 PC 사양 아래에서 진행되었다.

표 3을 보면, 실험결과 HOG는 SIFT에 비해 좋지 않은 결과를 보인다는 점을 알 수 있다. 그 이유는 HOG 알고리즘으로 얻을 수 있는 특징 개수가 SIFT와 비교하면 너무 적다는 것에 있다. 따라서 너무 적은 양의 특징벡터로 과도하게 큰 코드북을 생성할 시에도 정확도가 떨어지는 결과를 낳게 된다. 하지만 표 4를 보면 SIFT 알고리즘을 사용하는 것 보다, 빠른 속도로 기술자를 추출하므로 실시간으로 검색 결과가 필요한 모바일 시스템에 적용하였을 때는 SIFT보다 효율적으로 보인다.

표 4. 처리 영역별 수행 속도 (단위: sec)

구 분	HOG 추출	분 류
속 도	0.123	1.101

5. 결론 및 향후 연구 방향

본 연구에서는 여러 가지 특징을 결합한 분류기법이 아닌, 단일 특징 기술자를 통한 식물 잎의 분류기법을 통해 인식률을 향상 시키는 방법을 제안하였다.

SIFT는 입력 영상 전체의 특징점을 찾아 많은 수의 특징 벡터를 가지는 것뿐 아니라, 특징점의 위치를 고려한 공간적 분할을 통해 인식할 수 있는 특징

의 수를 늘림으로써 인식률을 높이고 있다. 반면 HOG는 형상의 정보를 이용하는 기술자로 인식률의 향상을 기대하였으나, SIFT보다 상대적으로 좋지 않은 결과를 나타내었다. 하지만 HOG는 SIFT보다 빠른 속도로 처리될 수 있다는 장점이 있기 때문에 실제 환경에서 휴대용 모바일 기계 등에 적용되었을 때 높은 효율성을 보일 것이라 기대된다.

제안된 방법은 기존에 있는 여러 가지 기술자들과 부호화 방법을 결합하여 이전의 잎 자체에서 얻을 수 있는 특성을 결합한 방법보다 좋은 결과를 보여주 고 있으나, 완벽한 결과를 기대할 수는 없다. 좋은 결과를 보이고 있는 SIFT는 속도가 너무 느리다는 단점이 있으므로, 병렬처리를 통해 속도를 개선할 수 있는 연구와 앱으로 개발하여 사용자가 언제 어디서나 식물에 대한 다양한 정보를 쉽게 얻을 수 있는 기술을 향후 개발하도록 한다.

참 고 문 헌

- [1] 이재원, 김상균, “신경망을 이용한 내용기반 영상분류,” 한국멀티미디어학회논문지, 제5권, 제5호, pp. 505- 514, 2002.
- [2] J.X. Du, X.F. Wang, and G.J. Zhang, “Leaf Shape Based Plant Species Recognition,” *Applied Mathematics and Computation*, Vol. 185, No. 2, pp. 883-893, 2007.
- [3] 남윤영, 황인준, “모양기반 식물 잎 영상 검색을 위한 표현 및 매칭 기법,” 정보과학회 논문지: 소프트웨어 및 응용, 제32권, 제11호, pp. 1013-1021, 2005.
- [4] S.J. Kim and D.P. Kim, “Performance Evaluations for Leaf Classification using Combined Features of Shape and Texture,” 한국지능정보시스템학회:지능정보연구, Vol. 18, No. 3, pp. 1-12, 2012.
- [5] B.S. Bama, S.M. Valli, S. Raju, and V. Abhai

- Kumar, "Content Based Leaf Image Retrieval (CBLIR) using Shape, Color and Texture Features," *Indian Journal of Computer Science and Engineering*, Vol. 2, No. 2, pp. 202-211, 2011.
- [6] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based Image Retrieval at the End of the Early Years," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 22, No. 12, pp. 1349-1380, 2000.
- [7] D.G. Lowe, "Distinctive Image Features from Scale-invariant Keypoints," *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110, 2004.
- [8] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," *Proc. the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 886-893, 2005.
- [9] Ales Leonardis, Bernt Schiele, and Sven J. Dickinson, *Object Categorization: Computer and Human Vision Perspectives*, Cambridge University Press, 2009.
- [10] Y. Cao, C. Wang, Z. Li, L. Zhang, and L. Zang, "Spatial-bag-of-features," *Proc. the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3352-3359, 2010.
- [11] 김동현, 박혜영, "돌출맵 정보를 이용한 코드북 기반 객체인식," 정보과학회 논문지: 소프트웨어 및 응용, 제39권, 제4호, pp. 306-314, 2012.
- [12] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, "Locality-constrained Linear Coding for Image Classification," *Proc. the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3360-3367, 2010.
- [13] K. Yu, T. Zhang, and Y. Gong, "Nonlinear Learning using Local Coordinate Coding," *Advances in Neural Information Processing Systems*, Vol. 9, pp. 2223-2231, 2009.
- [14] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories," *Proc. the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 2169-2178, 2006.
- [15] J. Yang, K. Yu, Y. Gong, and T. Huang, "Linear Spatial Pyramid Matching using Sparse Coding for Image Classification," *Proc. the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1794-1801, 2009.
- [16] C. Rother, V. Kolmogorov, and A. Blake, "GrabCut: Interactive Foreground Extraction using Iterated Graph Cuts," *ACM Transactions on Graphics*, Vol. 23, No. 3, pp. 309-314, 2004.
- [17] C. Cortes and V. Vapnik, "Support-vector Networks," *Machine learning*, Vol. 20, No. 3, pp. 273-297, 1995.
- [18] A Leaf Recognition Algorithm for Plant Classification using Probabilistic Neural Network, <http://flavia.sourceforge.net/>, 2009.

강 수 명

2007년 3월~2011년 2월 계명대
학교 게임모바일콘텐츠
학과 공학사
2011년 3월~2013년 8월 계명대
학교 미디어아트학과 계
입학석사

2014년 3월~(주)지오씨엔아이 공간정보기술연구소 선임연구원

윤 상 민

1994년 3월~2000년 2월 고려대
학교 전자공학과 공학사
2000년 3월~2002년 2월 고려대
학교 전자공학과 공학석사
2007년 5월~2010년 6월 독일 다
름슈타트 공대 컴퓨터공
학 공학박사

2002년 2월~2005년 2월 삼성종합기술원 연구원
2010년 7월~2011년 7월 일본 AIST 박사후 연구원
2011년 9월~2012년 8월 연세대학교 미래기술연구원 조
교수
2012년 9월~현재 국민대학교 컴퓨터공학부 조교수

이 준 재

1986년 경북대학교 전자공학과
졸업(학사)
1990년 경북대학교 전자공학과
졸업(석사)
1994년 경북대학교 전자공학과
졸업(박사)

1995년~2007년 동서대학교 컴퓨터 정보공학부 부교수
1998년~1999년 Georgia Institute of Technology 방문
교수
2000년~2001년 ㈜파미 연구소장
2007년~현재 계명대학교 게임모바일콘텐츠학과 교수
관심분야: 영상처리, 3차원 컴퓨터 비전, 게임