

# 잎의 외형을 이용한 산나물 자동인식 시스템

김성중\*, 주휘윤\*, 김현정\*\*, 원일용\*  
\*서울호서전문학교 사이버해킹보안과  
\*\*건국대학교 컴퓨터공학과

e-mail:12-502722@daum.net, hush1231@hanmail.net, nygirl@konkuk.ac.kr,  
clccc@shoseo.ac.kr

## Automatic Recognition System for Wild edible greens using Leaf External Form

Seong-Jung Kim\*, Hwi-Yoon Ju\*, Hyun-Jung Kim\*\*, Il-Yong Won\*  
\*Cyber Hacking Security, Seoul Hoseo Technical College  
\*\*Dept of Computer Science and Engineering, Konkuk University

### 요 약

본 논문은 잎의 외형을 이용한 산나물 자동 인식 시스템을 제안한다. 더 나은 산나물 자동 인식 시스템을 위해 잎의 외형 추출뿐만 아니라 방향벡터를 이용하여 정확도를 향상시켰다. 이를 위해 BP와 HMM 알고리즘을 이용하여 개선하였고, 외형 특징점을 추출하여 표현하였다. 제안하는 시스템의 성능은 실험으로 증명하였으며, 어느 정도 의미 있는 결과를 얻을 수 있었다.

### 1. 서론

최근 등산 인구의 증가 및 웰빙에 대한 관심이 늘어가면서 산나물에 대한 관심도 늘어나고 있다[1]. 특히 웰빙에 대한 관심이 고조되면서 자연에서 얻을 수 있는 산나물이 인기를 끌고 있다. 그러나 산나물에는 산나물과 흡사한 모양의 야생식물 및 독초가 존재한다. 산나물인지 아닌지는 전문가가 아니면 일반인은 구별이 어렵고, 독초를 식용으로 오인하여 잘못 알고 먹으면 식중독 및 생명을 위협하는 등의 위험한 경우가 많다[2]. 따라서 수많은 산나물을 전문가의 도움이 없이 자동으로 인식하는 시스템에 대한 연구가 필요하다.

기존의 식물 자동인식 연구들은 일반적으로 식물 전체에 대한 연구들로 주로 식물의 종류를 판단하는 방법이다. 또한 산나물을 대상으로 하는 연구가 있으나 잎맥의 비율과 히스토그램을 이용한 연구방법[3]으로 산나물 자동인식 시스템의 초기 연구이므로 여러 가지 면에서 많은 제약이 있다.

따라서 본 논문에서는 산나물 잎의 외형을 대상으로 특징을 추출하는 기계학습 알고리즘과 방향 벡터를 사용하여 산나물 자동 인식하는 시스템을 제안한다.

본 논문의 구성은 다음과 같다. 2장은 관련 연구에 대해 설명하고, 3장에서는 본 논문에서 제안하는 잎의 외형을 이용한 산나물 자동인식 시스템을 설명한다. 4장에서는 실험에 대한 분석 및 결과 분석을 하고, 마지막 5장에서는 결론 및 향후 과제에 대해 언급한다.

### 2. 관련 연구

#### 2.1 Suzuki 알고리즘

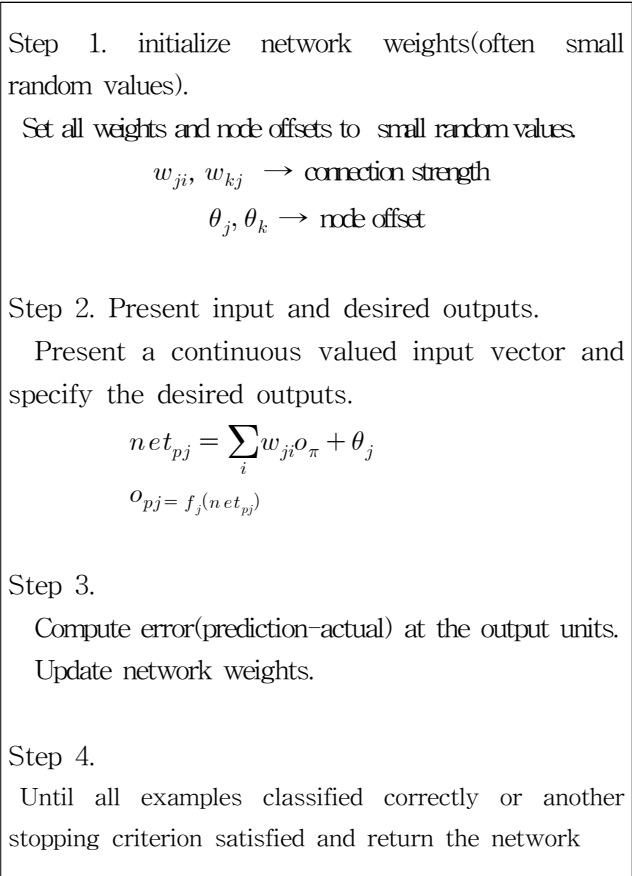
Border Following[4,5]은 객체의 외곽선을 추출하는데 사용하는 알고리즘으로, 효율적으로 추출된 윤곽선을 표현한다. 본 논문에서 사용된 Suznki 알고리즘은 border following을 기반으로 한 알고리즘 중 하나이다.

Suznki알고리즘[6]의 특징은 크게 두 가지 방향으로 나타낸다. 하나는 이진화된 이미지의 경계 중에서 둘러싸는 관계를 결정한다. 이때, 1-픽셀의 연결된 요소와 0-픽셀(background 또는 hole)은 외곽 테두리와 hole 경계 때문에 일대일로 대응한다. 이 방법은 이미지의 재구성 없이 어느 정도의 특징들을 추출할 수 있는 이미지의 대표를 산출한다. 다른 하나는 앞에서 제시한 방법의 변형된 버전으로써, 오직 바깥쪽 경계만을 따른다. 이 방법은 이진화된 이미지의 성분 계산, 수축, 구조적인 위상 분석에 효과적으로 사용된다.

#### 2.2 Back Propagation

BP(Back Propagation) 알고리즘[7]은 신경망을 훈련시키기 위해 사용하는 기술로, 다층 신경망이며, 감독학습 방법을 사용한다. 학습결과에 따른 에러가 출력노드로부터 내부 노드로 역으로 전파하는 것이다. 즉, 변경 가능한 가중치(weight)에 대해 네트워크 에러의 기울기를 계산하는데 사용된다.

이 알고리즘의 학습 과정은 크게 4단계로 다음과 같이 나타낸다.



(Fig 1) Back Propagation Learning Procedure

### 2.3 Hidden Markov Model

HMM(Hidden Markov Model)[8,9]은 각 상태간 전이 확률을 가지는 유한 상태 기계이며, 시간적 상태 또는 생성된 데이터는 은닉되고 관측되는 데이터는 출력으로 하여 상태를 파악하여 데이터를 인식하고자 하는 인식 과정이면서 관측 가능한 심벌로서 관측 불가능한 프로세서를 확률로서 추정하는 방식이다. 즉, 관찰이 불가능한 은닉의 확률론적 과정을 관찰이 가능한 기호를 발생시켜 다른 확률론적 과정을 통해 모형화하는 확률론적 가정이다.

HMM을 실제 응용 분야에 적용하기 위해서는 다음 3가지 기본 문제가 해결되어야 한다.

#### 2.3.1 확률 평가 문제

관측열  $O = \{O_1, O_2, O_3, \dots\}$  과 모델  $\lambda = (A, B, \Pi)$  이 주어져 있을 때, 이 관측열이 모델로부터 발생할 확률  $P(O|\lambda)$  을 효과적으로 계산할 수 있어야 한다. 이 문제는 전향(Forward)과 후향(Backward) 알고리즘[9]에 의하여 해결된다.

$$P(O|\lambda) = \sum_{i=1}^N \alpha_T(i) \quad (1)$$

$$\beta(i) = \sum_{j=1}^N a_{ij} b_j(o_{t+1}) \beta_{t+1}(j) \quad (2)$$

$$1 \leq i \leq N, t = T-1, \dots, 1$$

#### 2.3.2 최적 상태열을 찾는 문제

관측열  $O = \{O_1, O_2, O_3, \dots\}$  과 모델  $\lambda = (A, B, \Pi)$  이 주어져 있을 때, 관측열을 가장 잘 설명하는 최적 상태열  $q = \{q_1, q_2, q_3, \dots\}$  을 찾을 수 있어야 한다. 이 문제는 비터비(Viterbi) 알고리즘[9]에 의하여 해결된다.

$$\delta_t(i) = \max_{q_1, q_2, \dots, q_{t-1}} p(q_1, q_2, \dots, q_t = i, o_1, o_2, \dots, o_t | \lambda) \quad (3)$$

$$\delta_{t+1}(j) = \max_i [\delta_t(i) a_{ij}] \cdot b_j(o_{t+1}) \quad (4)$$

#### 2.3.3 파라미터 추정의 문제

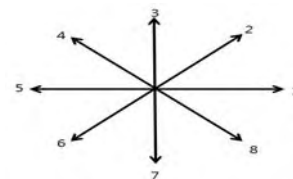
$P(O|\lambda)$  를 최대화하는 모델  $\lambda = (A, B, \Pi)$  의 파라미터를 추정하여 구할 수 있어야 한다. 이 문제는 바움-웰치(Baum-Welch) 재추정 알고리즘[9]에 의하여 해결된다. 이를 이용하여 새로 구성된 모델과 이전 모델이 생성하는 확률의 차이가 특정값 이상이 될 때까지 모델의 파라미터를 변경하면서 새로운 모델을 생성한다. 또한 계산된 변수값을 이용하여 모델의 파라미터를 재추정한다.

$$\xi(i, j) = \frac{\alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}{P(O|\lambda)} \quad (5)$$

$$\gamma_t(i) = \sum_{j=1}^N \xi_t(i, j) \quad (6)$$

### 3. 잎의 외형을 이용한 산나물 자동인식 시스템

본 논문에서 제안하는 방법은 크게 두 가지로 나뉜다. 하나는 잎의 외형을 기준으로 하여 외형을 추출한다. 이때, findContours 함수를 이용하여 산나물 잎의 외형을 추출한 후, 방향벡터를 적용하여 추출한 잎 외형의 이미지를 수치화한다. (Fig 2)는 본 논문에서 사용한 8방향 방향벡터이다.



(Fig 2) Vector Direction

두 번째는 첫 번째 과정에 의해 얻어진 잎의 외형을 수치화한 데이터를 각각 BP와 HMM 알고리즘을 이용하여 학습 및 추출한다. 이렇게 추출한 영상에 대한 근사값을 계산한다.

(Fig 3)은 제안하는 잎의 외형을 이용한 산나물 자동인식 시스템에서의 인지 방식을 간략하게 알고리즘으로 표현했다.

<Learning>

1. Leaf External Form extracts for each type of Plant.
2. Eight direction vector digitize the leaf external form image of step 1.
3. When it has continuously same values by step 2, it has to deal with as two overlapping values.
4. To learning the obtain values from step 3. At this time, we use to BP and HMM algorithm.

<Recognition>

1. The type of Leaf is normalization constant value between 0 and 1.
2. Apply eight direction vector to contour of leaf.
3. When they are classified as the normalization constant values in step 1. Of these look for analogous to the step 2's values.
4. The class value with the highest similarity, in other words, is selected an approximate value.

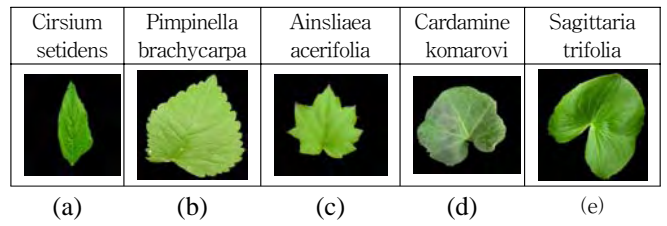
(Fig 3) Proposed Learning Algorithm

먼저 학습 단계에서는 각각의 종에 대하여 대표 샘플을 구하여 잎 외형을 방향벡터로 수치화한 학습 데이터를 저장한다. 인식 단계에서는 학습 단계에서 저장한 데이터를 이용하여 비교한 후, 가장 근사치에 있는 종을 최종 산나물의 종류로 결정한다.

#### 4. 실험 및 결과 분석

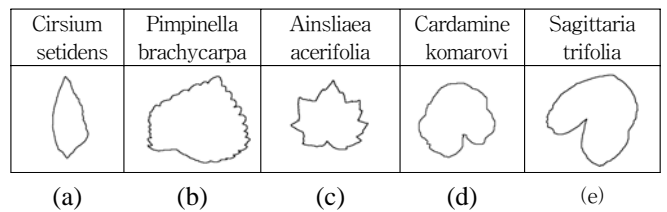
실험을 위해 준비한 산나물은 국내에서 수집할 수 있는 대표적인 산나물 중에서 20종을 선택하여, 각각에 대하여 크기나 관점이 다른 영상 10개씩을 각각 수집하여 사용했다.

(Fig 4)는 본 논문에서 실험을 위해 사용한 잎 이미지의 일부이다.



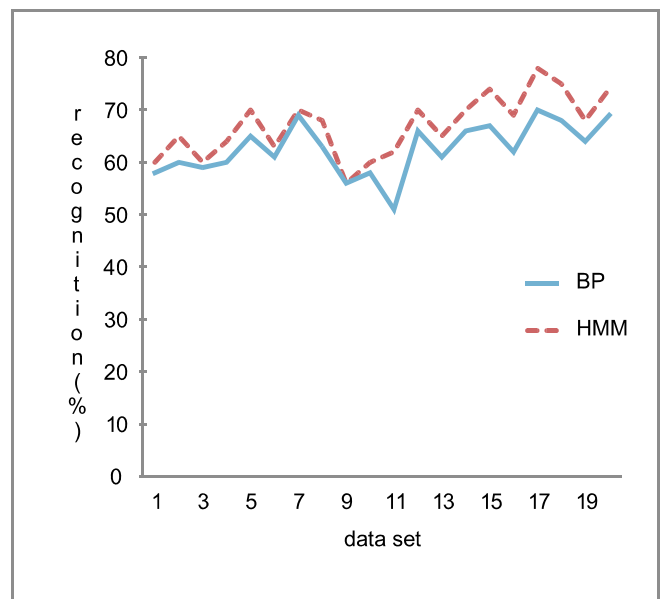
(Fig 4). Part of Experimental data (leaves)

실험은 각각의 산나물 이미지 중 50%를 랜덤으로 선택하여 학습하고, 나머지 50%로 테스트에 사용했다. (Fig 5)는 (Fig 4)에서 잎의 외형만을 추출한 이미지 일부이다.



(Fig 5) Obtained leaf shape images

학습과 인식은 Fig 1에서 제안한 알고리즘의 단계를 통해 실험하였다. 실험 결과는 (Fig 6)와 같다.



(Fig 6) Result of Experimental

(Fig 6)은 BP와 HMM을 이용한 잎의 종류별 인식률이다. 이 결과를 분석한 결과 BP를 통한 인식률은 평균 62.6% 정도이고, HMM을 통한 인식률은 평균 67.05% 정도로 잎의 외형을 방향벡터를 이용하여 수치화시킨 데이터를 학습시키는 알고리즘은 BP보다 HMM이 더 우수하다. 실험 결과를 분석한 결과, BP가 HMM보다 인식률이 낮은 이유를 분석한 결과 고려해야 할 요소가 많아 가장 적절한 수치를

알아내지 못한 것이라고 생각된다.

## 5. 결 론

본 논문에서는 산나물 검색을 위해 잎 외형의 모양을 추출하여 검색에 활용하였다. 방향벡터를 이용하여 잎의 외형을 수치화한 데이터를 BP와 HMM에 적용하여 학습시키고 분류하는 방법을 제안했다. 본 논문에서 제안한 방법의 유용성을 위해 실험으로 분석하였으며, HMM이 BP보다 더 우수했다.

향후 과제는 잎의 외형만이 아닌 잎의 내부의 잎맥까지 이용해 분류하는 방법을 결합하는 것과 더 나아가 휴대성을 위한 스마트폰 등 모바일 기기에서도 문제없이 실제 동작할 수 있는 실용성 연구가 추가로 필요하다.

## 참고문헌

- [1] 산림청, “스마트폰 증강현실을 이용한 식물정보 인식 타당성 분석연구”, 목원대학교 산학협력단 및 산림청, 2011
- [2] 강윤숙(식품기준과), “야생식물, 함부로 먹으면 큰일나요! -식용으로 오인하기 쉬운 꽃, 독초의 구별법 및 주의사항-”, 식품의약품안전처 보도자료, 2013.05.03
- [3] 김진서, 권수정, 김현정, 원일용, “산나물 자동 인식 시스템”, 한국정보처리학회 추계학술발표대회 논문집 제 20권, 제 2호, 2013
- [4] Image Processing-Laboratory 6: Border Tracing Algorithm, “[http://users.utcluj.ro/~raluca/ip\\_2014/ipl\\_06epdf](http://users.utcluj.ro/~raluca/ip_2014/ipl_06epdf)”
- [5] T.D. Haig, BE ,Y. Attikiouzel, BSc, PhD, M.D. Alder, MEngSc, PhD, “Border following : new definition gives improved borders”, IEE PROCEEDINGS-I, Vol. 139, No. 2, APRIL 1992
- [6] S. Suzuki, K. Abe, “Topological Structural Analysis of Digitized Binary Images by Border Following,” in CVGIP Vol.30, No.1, pp.32-46, 1985.
- [7] D. S. Kim, “Neural Network Theory and Application”, High-Tech Information, 1992
- [8] G. Y. Kim, “A Study of Behavior Analysis of Upper Part Body in the Wearable Environment,” M.E. dissertation, Wonkwang University, Jeollabuk-do, Korea.
- [9] 엄재홍, “은닉마르코프모델을 이용한 정보추출”, 서울대학교 대학원 컴퓨터공학석사 논문, 2001