

신경망을 이용한 도로가 포함된 야외영상 인식

(Recognition of Outdoor Scenery Containing Roads using Neural Network)

이 호 중 [†]

(Hyo Jong Lee)

요약 야외에서 인지되는 자연 경치는 다양한 개체, 빛의 산란, 또는 변화를 주는 많은 요소들 때문에 컴퓨터 영상처리에서 인식하기가 쉽지 않다. 본 논문에서는 다층 인지 신경망을 이용하여 도로가 포함된 야외영상에 나타나는 개체들을 인식하는 방법을 연구하였다. 자연 영상을 영역화한 후, 각각의 영역들에 대하여 색상과 기하학적인 특성에 근거하여 특성벡터를 추출하고 이를 신경망에 입력하여 각 영역을 구분하는 2단계의 알고리즘을 제안한다. 먼저 야외 영상들을 개선된 영역 확장법과 병합과정에 의하여 개체별로 영역화하였다. 영역화된 영상은 자연 영상과 함께 영상 데이터베이스에 저장되고, 이 자료들을 이용하여 각 영역의 특성벡터를 계산하였다. 이 특성 벡터를 구성된 신경망의 입력층에 전달하면, 각 영역은 27개의 개체 중의 하나로 출력층에서 인식된다. 제안된 방법은 학습에 사용된 데이터, 학습에 사용되지 않은 새로운 데이터, 그리고 모두 합하여 놓은 데이터의 세가지 데이터 군에서 무작위로 선별하여 인식률을 측정하였다. 학습된 데이터에서는 99.4%까지의 인식률을 보여주었고, 학습되지 않은 데이터에 대해서도 최고 89.1%까지의 인식률을 나타내었다. 제안된 방법은 평균적으로 88.1%~97.7%의 인식률을 보여주어 자연 경치의 인식에 신뢰성 있는 방법으로 사용될 수 있음을 증명하였다.

Abstract It is not easy to recognize every object found in outdoor scenery. Even with the help of advanced image processing techniques, the complex conditions, such as various shapes, different light dispersion, and other unstable conditions keep these objects from being identified. In this paper the recognition method of objects from natural scenery containing roads has been studied with a neural network of a multi-layer perceptron architecture. A two-stage algorithm has been proposed to label each segmented region with feature vectors of color and geometric characteristics. They are fed into the neural network after the whole image has been segmented properly.

First, outdoor scenery has been segmented with the modified region growing method and merging process. The segmented regions are stored in an image database where the original image and feature vectors are computed utilizing the image database. As the feature vectors are fed into the input layers, each region is classified into one of 27 different object lists. The performance of the proposed method has been measured with three image groups of 'used for training', 'not used for training', and combined image data of both. The recognition rate for the 'used for training' data and 'not used for training' data have reached up to 99.4% and 89.1%, respectively. The proposed method shows the average recognition rate of 88.1%~97.7% for any data set and is considered as a reliable algorithm to recognize outdoor scenery.

1. 서론

컴퓨터를 통한 물체의 인식과 자동화는 컴퓨터비전 분야에서 오래 동안 연구되어 오고 있다. 무인자동차, 로봇의 시각장치, 자동지상운행 시스템, 또는 구미 선

· 이 논문은 1999년도 전북대학교의 지원 연구비에 의하여 연구되었음

† 중신희원 : 전북대학교 전자정보공학부 · 공업기술연구소 교수
hlee@moak.chonbuk.ac.kr

논문접수 : 2000년 7월 27일

심사완료 : 2000년 12월 20일

진국에서 계획되고 있는 교통정보관리 시스템 등에서 야외영상의 인식은 매우 중요하나, 그 복잡성으로 인하여 도로나 교차로, 차량, 또는 교통 표지판의 인식 등 국부적인 연구가 이루어지고 있을 뿐이다[1, 2]. 최근 Gotoh 등[3]은 도시 지역의 건물과 같이 정형적인 물체의 인식을 기하적으로 제구성하는 알고리즘을 개발했으나 야외에서 보여지는 초목이나 산들과 같은 비정형 물체에 대해서는 인식을 시도하지 못하였다. Gamba 등[4]도 야외 사진에서 자주 보이는 원근점과 검출된 윤곽선, 그리고 영역화의 결과로 얻어진 정보를 융합하여 영역을 인식하는 방법을 소개하였다. Iyatomi 등[5]은 자연 영상의 인식에 근접한 방법을 시도하면서, 영상의 해석을 위하여 신경망에 기반을 둔 퍼지 추론방법을 제안하기도 하였다.

국내에서는 아직까지 자연영상의 인식에 대한 연구가 활발히 전개되지 않고 있는 실정이다. 자동주행 시스템이나 교통량의 측정 목적으로 일부에서는 도로의 경계선, 도로 분리선의 기하학적인 특성을 이용한 연구[6]가 이루어지고 있거나, 야외에서 발견되는 인공물을 인식하는 시도[7]가 있었을 뿐이다. 자연영상이나 도로영상의 인식을 위해서는 세심한 처리가 요구되고 있다. 도로 영상과 같은 야외영상에서 나타나는 특징으로는 산재하는 대상, 다양한 형태, 독특한 질감(texture), 불 균일한 조명 상태 등을 들 수 있는데, 이러한 특성들을 고려하여 존재하는 사물들을 정확하게 인식할 수 있기 위해서는 영상의 특징을 이용하되 너무 그러한 특징에 의존하지 않는 조심스러운 접근 방법이 요구되는 것이다.

본 논문에서는 21세기의 첨단 과학기술을 예견하면서 앞서 기술한 무인 자동차나 로봇 공학에서 선행되어야 할 자연영상 인식의 처리방법을 제안한다. 단일 물체나 단순한 형태의 물체 등을 인식하는 특수 알고리즘이 아니고, 야외에서 발견되는 사물들을 인식할 수 있는 범용 알고리즘으로 개발하였다. 또한, 다양한 질감에 내구성을 지니며 자동차, 도로 및 도로 분리선을 주변에 있는 물체들과 명확히 구분하여 앞서 언급한 자동항법이나 로봇 시각장치 등에 적용될 수 있도록 한다.

그러나, 자연 영상에서 분포하는 모든 객체들을 인식하는 것은 여러 가지 제약점이 따르기 때문에, 대상 영상들을 소규모 도시에서 볼 수 있는 도로를 포함한 영상으로 국한하여 도로와 도로를 중심으로 발견되는 가로수, 자동차, 건물, 마을 등과 같은 객체를 위주로 인식하였다. 햇빛의 과도한 산란을 방지하기 위해서 영상의 촬영은 가급적 구름이 낀 날을 택하여 수집하였다. 실험 영상의 특성은 영상데이터베이스에서 다시 소개한다.

자연 영상의 효율적인 분할방법과 분할된 영상에서 특성 벡터를 추출한 후, 신경망을 통한 학습을 이용하여 효율적인 인식을 수행하도록 하는 고급 영상처리 방법을 제안한다.

본 논문의 구성방법은 다음과 같다. 제 2장에서는 실험에 사용된 영상 데이터베이스의 특성과 자연 영상 인식의 첫 번째 단계인 영역분할 방법을 기술한다. 제 3장에서는 인식을 위해서 각 영역별로 계산하는 특징 벡터와 이를 이용하여 영역에 관한 정보를 학습시키는 신경망의 구성에 대하여 설명한다. 제 4장에서는 수집된 자연 영상에 적용한 실험 결과를 기술하고 제 5장에서는 결론을 기술한다.

2. 영역분할

자연 영상의 인식은 결국 자연 영상에 존재하는 객체들을 인식하는 문제이므로 이 객체들을 영역 분할할 필요가 있다. 영역 분할이 정확히 수행되고 나서 이들의 특징 벡터가 계산될 수 있다. 영역 분할의 기본 원리는 영상의 불연속성을 일정 문턱값(thresholding)을 이용하여 결정하는 것이며[8] 자연 영상 처리 뿐만 아니라 의료 영상처리[9]와 같이 육안으로 관찰하기 어려운 부분에서도 활용되고 있다.

컬러 영상의 영역분할은 여러 논문에서 다루어져 왔으나, 그 내용을 크게 네 가지로 요약할 수 있다. 첫째, 문턱값을 정하여 영역과 영역간의 급격한 차이를 구분하는 방법[8, 10]인데, 이는 잡음에 민감하며 적절한 문턱값을 지역적으로 설정하지 못하기 때문에 합리적이지 못하다. 둘째는 K-mean 방법과 같이 통계적인 원리에 기반을 두어서 영역화를 시행하는 것이다[11]. 이 방법은 거리함수를 정하여 특징 벡터들의 성격을 그룹화하는 것으로 적절한 최소거리값을 선정할 경우 상당히 좋은 결과를 보여주고 있다. 셋째는 윤곽선을 추출하여 분석한 후 이를 바탕으로 영역화를 시도하는 방법이다[12]. 일반적으로 추출된 윤곽선은 연속성이 약하여 최근에는 잘 사용되지 않고 있다. 마지막으로는 최소 영역을 설정하고 그 영역에서부터 주위 방향으로 조건을 만족할 때에 확장해 나가는 영역 확장방법이다[13,14]. 이 방법은 영역화를 정확하게 하는 장점 이외에도 잡음 영역을 적절하게 주위에 포함시킬 수도 있어서 좋은 결과를 내고 있다.

복잡한 자연 영상에서 영역을 정확히 구분 짓는 것은 쉬운 일이 아니다. 본 절에서는 영역을 기존의 윤곽선 탐지 방법에 의해서 구분하는 경우와 영역의 확장 방법에 의한 분할을 비교 설명한다. 또한 영역의 분할과 인

식을 위해서는 자연영상의 수집이 요구되며 이를 체계적으로 정리하여야 하므로 영상 데이터베이스의 구축이 필요하다.

2.1 영상 데이터베이스

본 연구에서 다루는 자연 경치의 영상은 준 도시에서 전형적으로 보여지는 도로를 포함하거나 도로를 중심으로 하는 영상으로 구성하였다. 자동차를 주로 객체로 간주하였으며 너무 분주하지 않은 장소들을 선택하였다. 선정된 야외경치는 구름이 적절하게 분포하여 지나친 햇빛의 노출이 나타나지 않는 날씨를 택하여 35mm 컬러 슬라이드 필름에 촬영하였다. 카메라의 위치는 자동차 운전자의 평균 눈 높이를 기준으로 하였다.

이러한 자연 영상은 고해상도 스캐너로 채집되어 압축 손실을 피하기 위해서 24-비트 컬러 영상 형태 그대로 저장하였다. 약 120여장의 컬러 영상 이미지가 최종적으로 실험에 활용하였다. 이러한 자연 영상은 고해상도 스캐너로 채집되어 압축 손실을 피하기 위해서 24-비트 컬러 영상 형태 그대로 저장하였다. 약 120여장의 컬러 영상 이미지가 최종적으로 실험에 활용되었다.

각 자연 영상은 다음 절에서 설명하는 반자동 처리 방법으로 분할되었고 각 영역은 영상 전문가에 의하여 명명되었다. 이 과정은 영상 전문가의 주관에 의해서 구분될 수 있다는 단점은 있으나, 사람에 의하여 각 영역이 정확히 검사된 후, 영역별로 그 개체가 명명(label)되므로 인식 결과의 성패를 직접 비교할 수 있다는 장점을 제공 해준다.

영상 전문가는 영역을 명명하는 과정에서 표 1에서와 같이 총 27개의 하위 개체를 가정하였다. 이들 개체는 다시 9개의 상위 개체와 13개의 중위 개체로 구분하여 자연 영상 데이터베이스가 객체들의 상하관계를 유지하도록 하였다. 상위 개체는 개념적으로 인식되는 것이며 하위 개체는 상위 개체를 구성하는 또 다른 개체들을 의미한다. 개체 등급에 있어서 다분화는 개체의 정확한 색상 정보를 신경망에 입력할 수 있고, 인식하고자 하는 대상을 상세히 구분할 경우와 대략적으로 구분할 경우를 나누어 유용하게 사용될 수 있다.

예를 들어서 자연 영상에서 나타나는 자동차는 본체와 유리창 등으로 각기 달리 영역화가 되지만, 이들을 총체적으로 구분할 때는 자동차라는 상위 개체로 구분할 수 있다. 이러한 개념은 식생(vegetation)을 다년생 나무, 1년생 잡초류 및 논으로 구분하거나, 또는 지붕, 창 및 벽으로 구성되는 가옥에서처럼 문제의 복잡도에 따라 유연하게 분류할 수 있어서 편리하게 이용될 수 있다.

표 1 데이터베이스의 개체분류

	하위개체	중위개체	상위개체
1	하늘	1	하늘
2	아스팔트	2	도로
3	콘크리트		
4	시골길	3	황색선
5	흙길(주위)		
6	황색차선		
7	백색차선	4	백색선
8	도로마크	5	장애물
9	전봇대		
10	가로등	6	나무
11	녹색나무		
12	갈색나무	7	풀
13	녹색풀		
14	갈색풀		
15	녹색논	8	논
16	황색논		
17	본체	9	자동차
18	창문		
19	아파트마을	10	원거리마을
20	맑은산	11	산
21	원거리산		
22	아파트벽	12	건축물
23	아파트창문		
24	주택지붕		
25	주택벽		
26	주택창문	13	기타
27	기타		

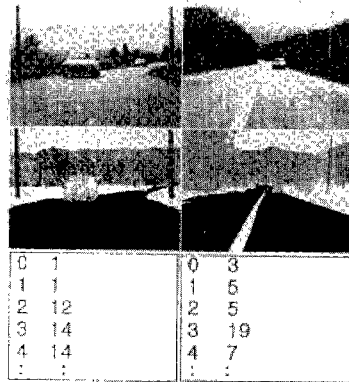


그림 1 영상 데이터베이스 레코드에 예

구체적인 개체의 목록과 이들의 관계를 표 1에 요약하였다. 본 연구의 결과는 27개의 개체를 편의에 따라 임의로 조절할 수 있으나 13개의 중위 개체를 근거로 실험 결과에 사용하였다. 실험 결과에서 언급하는 13개의 개체는 표 1에 나와 있는 바와 같이 하늘, 도로, 황색선, 백색선, 장애물, 나무, 풀, 논, 자동차, 원거리 마을, 산, 건축물 및 기타로 구분되었다.

이와 같은 과정을 모든 영상에 대하여 처리하여, 영상 데이터베이스의 자료들은 3쌍의 자료가 한 레코드를 형성하여 저장되었다. 즉, 일치되는 자연 영상, 분할 영상. 그리고 명명자료가 쌍을 이루며, 데이터베이스를 쌍의 형태로 구성시켰다. 그림 1은 영상 데이터베이스에서 전형적으로 나타나는 3쌍으로 구성된 자료파일 예를 보여 주고 있다. 명명자료는 문자 파일로 각 영역이 무엇을 나타내는지를 숫자로 표기하고 있다.

2.2 윤곽선 탐지

영역 분할에 사용되는 윤곽선은 1차 또는 2차 미분 연산자에 의해서 두 영역의 불연속성을 잘 나타내 주는 속성을 이용한다. Marr와 Hildreth[15]는 Gaussian 필터와 기울기를 고려하여 윤곽선을 탐지하는 방법으로 LoG(Laplacian of Gaussian)를 수식 (1)과 같이 도입하였다.

$$h(x,y) = \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$$

$$\nabla^2 h = \left(\frac{r^2-\sigma^2}{\sigma^4}\right)\exp\left(-\frac{r^2}{2\sigma^2}\right) \quad (1)$$

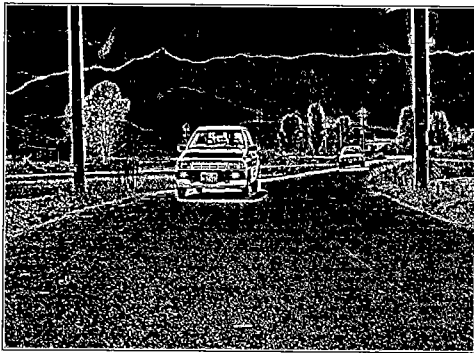


그림 2 자연 영상으로부터 LoG 윤곽선 검출

그러나, 자연 영상과 같이 복잡도가 높으며 인식하고자 하는 대상과 배경과의 뚜렷한 차이가 나타나지 않으며, 잡음이 많고, 색상의 변화가 급격하지 않은 영상에서는 윤곽선이 잘 탐지되지 않는다. 특히 윤곽선이 부분적으로 탐지되어서 같은 영역을 둘러싸고 있는 경계로서의 윤곽선들을 다시 연결하는 작업도 쉽지 않다. 그림 2는 수식 (1)과 같이 자연 영상에 LoG 윤곽선 검출 연산을 실행한 결과를 보여주고 있다. 나타난 바와 같이 원 영상의 경우에 도로는 균일하게 보이나, 무수한 잡음이 발생했다. 특히 도로와 양 옆 논과의 경계지역이나, 가까운 거리에 있는 산과 멀리 있는 산과의 경계들이 불분명한 것을 볼 수 있다. 이러한 결과는 자연 영상을

평탄화 연산을 하거나 Gaussian 필터링, 또는 평균값 필터링 연산과 조합하여 LoG 윤곽선 검출 연산을 실행하여도 비슷하게 잡음이 나오는 결과를 보여준다.

이 것은 다양한 비정형 개체들이 분포하는 자연 영상이 빛의 영향을 심하게 받아서 잡음을 유발시키기 때문이다. 따라서, 자연 영상의 영역 분할을 위해서는 다른 방법이 시도되어야 한다.

2.3 영역 확장 방법

기존의 영역화 방법은 컬러 색상 좌표계에서 적절한 문턱값을 기준으로 하는 것이다. 그러나, 자연 영상의 경우에는 영상의 복잡도가 커서 적절한 문턱값을 영역마다 찾는 것이 불가능하다. Fuh[16] 등은 컬러 공간에서 컬러 간의 거리를 측정하여 유사도를 측정한 후, 일정 문턱값 내에 들었을 경우에 영역으로 포함시키는 영역 확장 방법을 제시하였다.

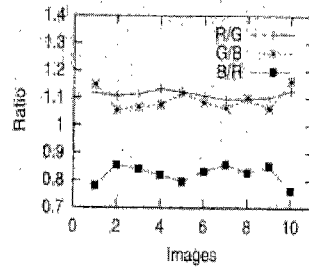


그림 3 자연영상에서 도로의 컬러 R, G, B간의 비율

본 논문에서는 고정된 문턱값을 영역마다 지정하는 대신 영역별로 나타나는 색상신호의 비례로 유사도를 판단하였다. 예로써, 도로 영역의 빨강신호(R)와 녹색신호(G)의 비(R/G)는 일정한 값의 범위를 나타낸다. 녹색신호(G)와 파란색 신호(B)의 비는 잔디나 나뭇잎 등의 식물분포 영역을 나타내는 중요한 기준값으로 사용될 수 있다. 그림 3은 자연 영상에서 촬영된 도로의 R/G, G/B 및 B/R의 비율을 그래프로 나타낸 것이다. 데이터베이스의 영상을 대상으로 실행한 결과 채널간의 비율의 차이는 날씨나 시간과 같은 환경의 영향을 받지 않는 것을 알 수 있었다. 실사 비율의 값은 변할 지라도 비율간의 차이는 계속 유지되기 때문에 영역화에 중요하게 사용될 수 있다는 장점이 있다. 그림 3에서처럼 컬러 채널 간의 비율은 항상 일정한 값의 범위에서 분포하는 것을 볼 수 있다. 다른 개체들도 각각 비슷한 값의 범주에서 컬러 채널 간의 비율 분포를 보여준다.

채널 간의 비율이 같은 영역에서는 일정하게 분포한다는 특성은 컬러값 속성을 사용하여 두 영역을 분할하

는 것이며, 자연 영상 전역에 걸쳐서 임의의 시작화소를 중심으로 영역을 확장시켜 나가는 것이다. 즉, (x, y) 에 위치한 화소의 색상 밝기 $I(x,y)$ 를 식 (2)와 같이 나타낼 수 있다.

$$I(x, y) = \sqrt{\frac{R(x, y)^2 + G(x, y)^2 + B(x, y)^2}{3}} \quad (2)$$

$$\approx \frac{R(x, y) + G(x, y) + B(x, y)}{3}$$

임의의 영역에서 색상 밝기의 변위(variance)는 검사 대상의 화소를 인접 영역에 포함되는지의 여부를 판단하는데 결정 요소로 사용할 수 있다. 식 (3)은 변위를 이용하여 해당 화소의 포함 여부를 결정하는 식을 보여 주고 있다.

$$var(R) = \frac{\sum_{i_n=0}^{\text{Number of pixel}} I^2(i_n)}{\text{Total number of Region}} - Mean^2(R) \quad (3)$$

여기에서 R 은 영역을 나타내며 $I(i_n)$ 은 영역 R 의 i 채널의 밝기를 의미한다. 따라서, 새로운 화소와 영역의 색상 밝기의 변위차가 주어진 문턱값보다 작으면 같은 영역으로 포함시킬 수 있다. 적용된 영역 확장 알고리즘의 단계는 다음과 같이 정리될 수 있다.

그림 4와 같은 자연 영상에 영역 확장 알고리즘을 적용하면 그림 5와 같은 결과 영상이 나온다. 일반적으로 영역

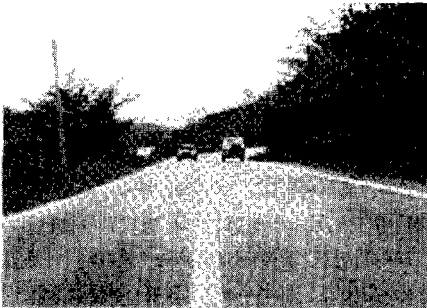


그림 4 자연 영상의 예



그림 5 그림 4에 영역 확장 알고리즘을 적용한 결과 영상

확장 알고리즘의 1차적인 결과 영상은 그림 5에서 나타나는 것처럼 영역의 수가 너무 과다하게 나타나게 된다.

영역 확장 알고리즘

1. 임의의 수 n 영역에서 컬러 채널 간의 비율을 계산하고, 영역 R_0 에서 R_{n-1} 까지의 씨앗값 r_i 를 임의로 설정한다.
2. 만약 영역 R_i 의 인접 화소의 r_i 값이 주어진 문턱값보다 작으면, 그 인접 화소를 영역 R_i 에 포함시킨다.
3. 단계 2를 모든 인접화소에 대하여 반복하여 영역 R_i 에 포함되는지 여부를 결정한다.
4. 포함되지 않는 인접화소는 새로운 영역 R_j 로 구분한다 ($j > n-1$).
5. 단계 1에서 단계 4를 영상의 모든 화소가 포함될 때까지 반복한다.

따라서, 영상에 인지된 전 영역에 걸쳐서 모든 영역들은 인접 영역과 유사성을 비교해서 같은 영역에 속하는지를 검사하는 불필요한 영역을 제거하는 병합 단계가 필요하다. 병합 단계는 두 영역간의 유사성을, 화소가 영역에 포함되는지를 판단할 때와 같이, 두 영역에서 색상의 밝기 변위를 이용하여 결정한다. 즉, 두 영역간의 색상의 변위차가 주어진 문턱값보다 작으면 동일한 영역으로 간주하는 것이다. 이와 같이 병합 과정은 과다하게 영역 분할된 영상을 적절한 영역으로 분할해주는 역할을 하게 된다. 그림 6은 그림 5에서 과다하게 영역 분할된 영상을 정리하여 최종 영역 분할한 것을 보여 준다.

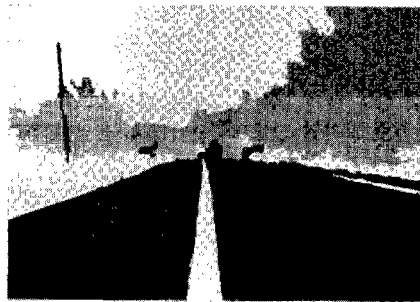


그림 6 병합 과정을 통한 그림 5의 최종 영역 분할

3. 신경망 설계

신경망에서 채택된 활성화 함수의 특성에 따라 선형과 비선형 모델로 분류할 수 있다. 비선형 모델은 각 처리요소의 활성화 함수의 비선형 특성 때문에 복잡한 경계 영역 문제를 효과적으로 처리할 수 있다. 특히 다층계 영역 문제를 효과적으로 처리할 수 있다. 특히 다

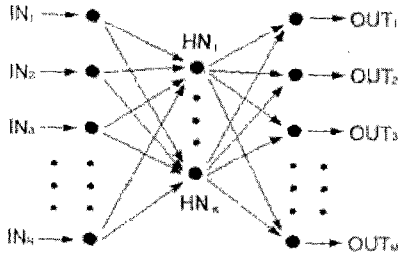


그림 7 3-층 구조의 신경망

층퍼셉트론 신경망 모델은 Rumelhart[17] 등에 의해 비선형 다층 구조 신경망에 대한 학습 알고리즘이 제안되어 계산 기능이 크게 향상되면서 충분한 처리요소만 보장되면 어떤 연속사상(continuous mapping)이라도 근사화 할 수 있어 음성 합성 및 인식 문제, 그리고 영상 패턴 인식 등의 분야에서 효율적으로 응용되고 있다. 본 연구에서도 $1/(1+e^x)$ 의 비선형 활성화 함수와 학습결과를 재귀적으로 나타내는 교차신호를 사용하여 실험을 수행하였다. 그림 7은 실험에 사용된 신경망의 구조를 보여주고 있다.

그림 7은 입력층(IN), 출력층(OUT), 그리고 은닉층(HN)의 3층 신경망 구조로 구성되었다. 입력층에서는 각 영역의 특성을 산출하여 계산된 입력 자료를 받아들인다. 입력 자료는 분할된 해당 영역에서 5개의 색상 정보와 4개의 기하 정보로 총 9개의 입력을 사용하였다.

이들은 영상처리에서 전통적으로 다루어 오는 특성[18]이며, 구체적으로 색상 정보는 휘도, 채도, 명도, R/G 및 G/B를 사용했으며, 기하 정보로는 영역의 크기, 위치, 가로와 세로의 비율 및 치밀도(compactness)를 사용하였다. 이들 입력에 사용된 특성 벡터들은 신경망을 구성하는 단계에서 우수한 결과를 내는 것으로서 선별되었다. 은닉층 노드 개수는 시뮬레이션 결과로부터 정확도와 수렴 시간을 고려하여 선택하였다. 일반적으로 은닉층의 노드수가 25개에서 48개 사이에 분포할 때 수렴속도와 정확도가 우수하였다. 본 연구에서는 45개의 은닉층 노드를 설정하여 최종적으로 실험하였다.

출력층의 노드 개수는 자연에 분포하는 개체들을 최종 27개로 구분하여 27단으로 정하였다. 인식대상은 2.1절에서 언급한 영상 데이터베이스에서 규정한 개체 상하관계를 정립하여 정의하였다. 따라서, 이들 출력 결과는 사용자의 요구 사항에 알맞게 최종 인식 대상의 개수로 변화시킬 수 있다.

4. 실험 결과

구성된 신경망을 이용하여 영상 데이터베이스에 입력된 영상들에 대한 인식을 실험하였다. 영상 데이터베이스에 저장된 총 120장의 자연 영상 중에서 80장을 무작위 선별하여 신경망의 학습에 사용하였다. 실험은 학습에 이용된 80장의 학습 데이터군, 실험에 이용되지 않은 40장의 비학습 데이터군, 그리고 전체 영상에서 무작위로 선별한 무작위 데이터군의 3종류의 데이터에 대하여 실시되었다. 제안된 자연 영상 인식 알고리즘의 효율을 측정하기 위해서 데이터군에 따른 실험별로 27개의 개체를 중심으로 영역 전체 영상의 화소 수와 정확하게 인식된 영역의 화소 수의 비율로 정확도를 측정하였다.

그림 8은 신경망 훈련에 사용한 학습 데이터군 영상 중의 하나이며, 그림 9는 그림 8의 영상을 인식한 결과를 보여주고 있다. 그림 9의 중앙부에 있는 두 개의 화살표는 잘못 인식된 영역을 나타내고 있다. 제안된 방법에 의한 학습 데이터군에 적용한 자연 영상의 정확한 인식율은 최대 99.4%까지 되었다. 학습에 사용된 영상 중 무작위로 이미지를 선택하여 실험한 결과는 영상 전체 화소에 대해서 0.6%~4% 정도의 오차율을 보여 주었다. 왼쪽 화살표가 가리키는 영역은 건축물로 명명되었으나, 원거리에 있는 산이 안개에 둘러싸인 부분이다.

오른쪽 화살표가 가리키는 영역에서는 전봇대에 부착된 물체를 자동차로 오인식하고 있다.



그림 8 학습 데이터군 영상의 원영상 예

'S'라고 표시된 부분은 움직이는 자동차의 그림자에 해당되는 부분이다. 그림자는 자동차의 일부분이 아니므로 자동차 영역과 분리되어야 한다. 그러나, 그림자를 제거하는 과정은 간단한 문제가 아닌 또 다른 연구 분야이므로 본 연구에서는 그림자를 자동차 영역에 포함시켜 영상 데이터베이스의 영역 분할 영상으로 등록하였다. 따라서, 그림자 영역은 자동차 영역으로 명명하여도 정확한 인식으로 취급하였다.



그림 9 분할 영역의 인식 영상(화살표는 오인식 지역)

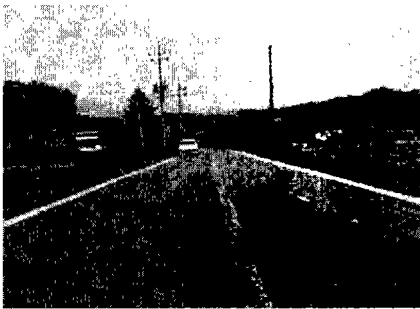


그림 10 인식에 사용한 비학습 데이터군 영상의 예

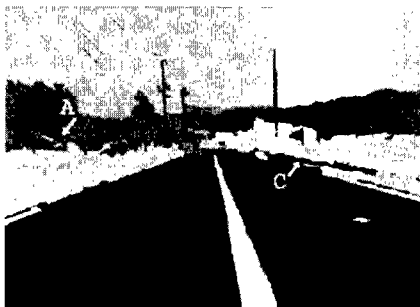


그림 11 그림 9의 인식 영상(화살표는 오인식 지역)

이러한 조건에서 전체적으로 그림 9의 정확한 인식률은 98.2%이었다. 그림 10은 비학습 데이터 집합의 한 예제 영상이며, 그림 11은 실험된 신경망의 결과 영상을 보여주고 있다. 비학습 데이터 집합 영상에서의 인식률은 학습 데이터 집합 영상의 인식률과 비교하여 일반적으로 약 10% 정도가 낮은 10%~13%의 오인식률을 나타내었다. 그림 11에서 보듯이 하늘, 도로, 또는 도로선 등은 대체로 정확하게 인식이 되었다.

오인식이 일어나는 대부분 영역은 원거리에 있어서 윤곽이 불분명한 것들로, 안개 낀 산이나, 원거리에 있는 스모그로 둘러싸인 마을들이 대표적인 예이다. 그림 11에서도 오인식된 영역은 영문 알파벳으로 표기하였다. 화살표 'A'로 표기된 영역은 나무에 둘러싸인 가옥이지만, 가옥을 놓치고 모두 나무로 인식하고 있다. 화살표 'B'로 표기된 영역은 스모그에 둘러싸인 산이지만, 그 영역은 자동차로 오인식되고 있다. 도로 옆에 방치된 인조물질(플라스틱류)은 역시 자동차로 오인식되고 있는 것을 화살표 'C'로 표시하고 있다.

학습에서 오인식이 발생하는 개체들의 관계를 파악하기 위해서 표 2와 표 3에 훈련영상과 비훈련 영상의 오인식 상관테이블을 각각 작성하였다. 세로축은 원 영상의 개체를 나타내고 가로축은 실험영상의 개체를 나타낸다. 가로축과 세로축의 순서는 표 1에 나와 있는 13개의 중위개체의 순서대로 배열한 것이다. 훈련영상의 오인식 상관 테이블에서 차선(3과 4)은 완전히 인식되고 있음을 보여주고, 나무(6)의 경우에 원거리에 있는 마을이나 산 등으로, 그리고 풀(7)이 자동차나 산으로 쉽게 오인식이 되고 있음을 알 수 있다. 13번 기타 영역은 도로 상에 있는 폐기물 또는 앞서 지정된 12개의 개체 이외의 개체로 크기 면에서 극히 작은 영역을 차지하여 상호간에 많은 오류를 범하고 있어서 크기가 작은 영역의 학습이 어려운 것을 보여주고 있다.

비훈련 영상의 오인식 상관 테이블은 훈련 영상에 비교하여 현저하게 오인식이 증가된 것을 보여주고 있다. 일부의 도로(2)는 백색선이나 자동차로 인식되고 있으며, 반대로 도로선(3과 4)이나 풀(7)이 도로로 인식되고 있기도 하다. 또한 산(11)이 장애물이나 나무로 오인식되거나, 건축물이 나무로 오인식되거나 반대로 나무가 건축물로 오인식되고 있는 것도 빈도가 높게 나타나고 있다.

마지막 실험은 학습 데이터군과 비학습 데이터군에서 입력 영상을 무작위로 선별하여 실행하였다. 대체로 무작위 데이터군에 대한 인식결과는 비학습 데이터군에 대한 실험 결과보다는 우수하나 학습 데이터군에 가지고 실행한 실험 결과보다는 우수하지 못했다. 학습 데이터군, 비학습 데이터군 및 무작위 데이터군의 인식률 결과를 표 4에 요약하였다.

최저 결과는 지정된 데이터군에서 오인식이 가장 심한 경우의 수치이며, 최고 결과는 지정된 데이터군에서 가장 양호한 인식을 보여주는 경우의 수치를 말한다. 또한, 표에서 언급하는 평균 결과는 실험에 이용된 모든 영상의 인식률을 취하여 평균을 낸 수치이다.

표 2 훈련 영상의 오인식 상관 테이블

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	47	0	0	0	0	0	0	0	0	0	0	1	0
2	0	50	0	3	0	0	3	0	0	0	0	0	4
3	0	0	6	0	0	0	0	0	0	0	0	0	0
4	0	0	0	20	0	0	0	0	0	0	0	0	0
5	0	0	0	0	48	1	0	0	0	0	0	0	0
6	2	0	0	1	1	130	3	0	1	5	3	1	2
7	0	1	1	4	0	8	148	6	3	0	3	2	1
8	0	0	0	0	0	0	0	36	0	0	0	1	0
9	0	0	0	0	0	0	0	0	54	0	0	1	1
10	0	0	0	0	0	0	0	0	0	2	0	0	1
11	0	0	0	1	0	1	1	0	0	1	56	1	1
12	0	0	0	0	0	0	0	0	0	0	0	25	0
13	2	1	0	1	0	2	0	0	1	2	2	0	48

표 3 비훈련 영상의 오인식 상관 테이블

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	20	0	0	0	2	2	0	1	1	1	12	0	0
2	0	57	1	13	0	0	4	0	4	0	1	0	5
3	0	2	17	0	1	0	3	0	0	0	0	0	0
4	0	2	0	7	0	1	0	0	0	2	0	0	3
5	4	0	0	0	49	6	1	0	2	8	5	4	2
6	1	0	0	2	13	91	17	12	1	0	8	20	2
7	0	28	3	1	5	34	148	57	1	3	1	3	13
8	0	0	0	0	0	1	36	51	0	0	0	1	1
9	1	0	0	0	1	6	1	0	10	1	6	3	0
10	4	1	0	0	3	9	0	1	0	12	25	3	3
11	0	0	0	0	10	24	7	1	1	0	12	3	0
12	0	0	0	0	1	11	6	1	0	2	0	2	0
13	0	3	0	2	4	9	8	1	4	0	15	1	6

표 4 각 학습 방법에 따른 인식률

데이터군	인식률(%)		
	최저	최고	평균
학습 데이터군	96.0	99.4	97.7
무작위 데이터군	92.8	99.2	96.0
비학습 데이터군	87.0	89.1	88.1

5. 결론

자연 영상의 인식은 다른 영상과 달리 복잡도가 높아서 인식이 쉽지 않았다. 자연 영상에 분포하는 모든 개체를 정확하게 인식하는 문제는 현재까지 도전을 필요로 하는 분야이다. 본 연구에서는 자연 영상의 인식을 위해서 먼저 각 개체별로 영역의 분할을 하였다. 영역의 분할은 씨앗 화소로부터 영역을 점차 확장해 나가는 방법을 통해서 1차적으로 실행한 후, 과밀하게 분할된 영

역들의 색상의 변위차를 이용하여 불필요한 영역을 없애고 통합하였다. 이들 분할 영상은 전문가가 반자동식으로 각 영역을 명명하고, 원래의 자연 영상과 분할 영상이 쌍을 이루어 영상 데이터베이스에 저장하였다.

영상 데이터베이스에 구축된 자료들을 기반으로 자연 영상의 개체들을 학습하기 위해서 3-층 신경망을 구성하였다. 각 영역에서 추출한 9개의 색상 정보와 기하 정보는 신경망의 입력층에 전달되고, 자연 영상에서 흔히 발견되는 27개의 개체를 출력층의 단자에 연결하였다. 이 27개의 출력 개체는 다시 미리 정의된 13개의 중분류 개체로 구분되었다.

영상 데이터베이스의 일부 영상을 구성된 신경망을 통하여 실험에 활용하였다. 학습이 이루어진 후, 학습에 사용된 영상, 학습에 사용되지 않은 비학습 영상, 그리고 이 두 집단을 포함한 영상의 세 종류의 데이터에 대하여 각각 인식률을 측정된 결과, 학습 데이터군에서는 최고 99.4%의 인식률을 나타냈으며 비학습 데이터군에서는 최고 89.1%의 인식률을 보여주었다. 평균적으로 학습 데이터군, 비학습 데이터군 및 무작위 선별 데이터의 영상에 대하여는 97.7%, 88.1% 및 96.0%의 인식률을 각각 보여 주었다.

도로가 포함된 자연 영상에서 존재하는 대부분 개체의 인식은 신경망을 통하여 양호하게 인식이 되었으며, 원거리에 있어서 안개나 스모그, 또는 영역이 너무 작은 부분을 차지하는 개체에서 종종 혼동스러운 인식 현상을 보여주었다. 차후에 해결되어야 할 과제로는 인식률을 향상시키는 것과 계산 시간을 단축시켜서 입력되는 비디오 영상에 대해서도 실시간에 인식을 할 수 있도록 개선하는 것이다. 이러한 연구가 효과적으로 수행될 때에, 무인 자동차나 원격 시스템에서 요구하는 것과 같이 인간의 시각 시스템과 유사한 성능을 지니는 시스템으로 활용될 수 있을 것이다.

본 연구에서 데이터로 사용한 영상들은 구름이 약간 끼어서 빛의 산란이 심하지 않은 상태에서 수집되었으며, 모든 영상들은 도로를 중심으로하여 자연에서 발견되는 개체들을 포함하고 있다. 이러한 영상의 수집 과정에서 제한 시켰던 기상조건이나 인식할 수 있는 대상도 개체 수적인 면에서 확장하여 다양한 사물들을 인식할 수 있는 연구가 향후에 수행되어야 할 것이다.

참고 문헌

[1] Foresti, G., Murino, V., Regazzoni, C.S., and Vernazza, G., "A Distributed Approach to 3D Road Scene Recognition," IEEE Trans on

- Vehicular Technology, pp 389-406, Vol 43, No 2, May 1994.
- [2] Gay, M., "Segmentation Using Region Merging With Edges," Procs of the 5th Alvey Vision Conference, pp 115-119, Sept., 1989.
- [3] Tomohiko Gotoh, Mineichi Kudo, Jun Toyama and Masaru Shimbo, "Geometry Reconstruction of Urban Scenes by Tracking Vertical Edges," 1999 Third International Conference on Knowledge-Based Intelligent Information Engineering Systems, 31st Aug. - 1st Sept. 1999, Australia.
- [4] Gamba P, Lodola R, Mecocci A, "Scene Interpretation by Fusion of Segment and Region Information," Image & Vision Computing, Vol. 15, No.7, pp , 1997.
- [5] Iyatomi, Hitoshi and Hagiwara, Masafumi, "Knowledge Extraction from Scenery Images and the Recognition Using Fuzzy Inference Neural Networks," Proceedings of the 1998 IEEE International Conference on Systems, Man, and Cybernetics, Vol. 5, pp. 4486-4491, 1998.
- [6] 진병태, "제한된 검색 영역을 이용한 효율적인 차량계수", 정보과학회논문지(B): Software and applications vol. 23, no. 6, pp. 611-624, June 1996.
- [7] 정홍, "무인 주행을 위한 도로 인식 및 핸들 제어", 정보과학회 가을 학술발표논문집(II), vol. 25, no. 2, pp.497-499, October 1998,
- [8] 김진형, "아외에 위치한 인공물을 인식하기 위한 지식기반형 방법", 정보과학회논문지, vol. 14, no. 3, pp. 239-246, August 1987.
- [9] Grennhill, D. and Davies, E.R., "A New Approach to the Determination of Unbiased Thresholds for Image Segmentation," Image Processing and Its Applications, pp 519-523, July 1995.
- [10] Herodotou, N, Plataniotis, K. and Venetsanopoulos, A., "A Color Segmentation and Classification Scheme for Facial Image and Video Retrieval," in IX European Signal Processing Conference, Rhodes, Greece, September 1998.
- [11] A. R. Weeks and G. E. Hague, "Color Segmentation in the HSI Color Space Using the K-means Algorithm", Proceedings of SPIE, Vol. 3026 , pp 143-154, Feb 1997.
- [12] Vikram Chalana, Wendy Costa, and Yongmin Kim. "Integrating region growing and edge detection using regularization," In Proceedings of the SPIE Conference on Medical Imaging. SPIE, 1995.
- [13] N. Ikonomakis, K. N. Plataniotis, and A. N. Ventsanopoulos, "Grey-scale and colour image segmentation via region growing and region merging," Canadian Journal of Electrical and Computer Engineering 28-1,2, pp. 43-47, 1998.
- [14] Tremreau, A., and Borel, N., "A Region Growing and Merging Algorithm to Color Segmentation" Pattern Recognition, Vol. 30, No. 7, pp. 1191-1203, 1997.
- [15] D. Marr and E. C. Hildreth, "Theory of Edge Detection," Proceedings of Royal Society, London Britain, 207:187-127, 1980.
- [16] Fuh, Chiou-Shann, Cho, Shun-Wen and Essig, Kai, "Hierarchical Color Image Region Segmentation for Content-Based Image Retrieval System," IEEE Transactions on Image Processing, Vol. 9, No. 1, January 2000.
- [17] Rumelhart, D. E., Hinton, G. and Williams, R. J. "Learning Internal Representations by Error Propagation," In Parallel Distributed Processing (Vol. 1, pp. 318-362), Cambridge, MA, MIT Press, 1997.
- [18] Rafael C. Gonzalez and Richard E. Woods, Digital Image Processing, pp. 716, Addison-Wesley, 1992.



이 호 중

1957년 6월 19일 생. 1982년 2월 전북대학교 지구과학과 졸업(이학사). 1985년 8월 University of Utah 기상학과 졸업(이학석사). 1986년 8월 University of Utah 컴퓨터과학과 졸업(공학사). 1988년 8월과 1991년 8월 University of Utah 대학원 졸업(각각 공학석사와 공학박사). 1986년 ~ 1989년 University of Utah 컴퓨터과학과 조교. 1989년 ~ 1991년 University of Utah 선임프로그램머. 1991년 University of Utah 대기연구소 연구원. 1991년 9월부터 현재 전북대학교 조교수. 관심분야는 병렬처리, 인공지능, 병렬인식 알고리즘 개발 및 소프트웨어 공학.