

# 신경망을 이용한 내용 기반 이미지 분류

이재원, 김상균  
인제대학교 전산학과

## A Contents-Based Image Classification Using Neural Network

Jae-Won Lee, Sang-Kyoon Kim  
Dept. of Computer Science, Inje University

### 요약

본 논문에서는 신경망을 이용한 내용 기반 이미지 분류 방법을 제안한다. 분류 대상 이미지는 인터넷상의 다양한 이미지들 중 오브젝트 이미지이며 웹 에이전트를 통하여 획득하고 정규화 과정을 거친다. 획득한 이미지를 분류하기 위한 특징은 웨이블릿 변환 후 추출된 질감 특징이다. 추출된 질감 특징을 이용하여 학습패턴을 생성하고 신경망을 학습한다. 그리고 구성된 신경망 분류기로 이미지를 분류한다. 본 연구에서는 다양한 질감 특징들 중에서 대비(contrast), 에너지(energy), 엔트로피(entropy)를 이용하여 특징을 추출한다. 실험에 사용한 데이터는 30종류에 대하여 각각 10개씩, 300개의 이미지들을 학습 데이터, 테스트 데이터로 사용하여 구성된 분류기의 인식률을 실험하였다.

### 1. 서론

컴퓨팅 환경이 발달하고 인터넷이 보편화되면서 멀티미디어 데이터의 사용이 증가하고 있다. 그 중에서도 이미지 정보를 처리하고 검색하는 방법론에 대한 연구가 활발히 진행되고 있다.

기존의 이미지 검색은 텍스트 기반(text-based) 검색 방법을 이용한 결과를 제공하고 있다. 이러한 텍스트 기반의 검색 엔진들은 이미지 파일의 이름, 표제, 키워드와 같은 속성에 근거하여 이미지를 분류한다. 이는 바이트마다 숫자나 문자라는 의미가 부여되는 텍스트와는 달리, 단순한 비트열인 이미지 데이터를 텍스트처럼 정보화하기 위해서, 사람이 직접 이미지 데이터를 해석해서 그 내용을 문자로 표현하는 것과 유사한 방식이다.

하지만 텍스트 기반 검색은 사용자에게 많은 양의 검색 결과를 제공해줄 수는 있으나, 이미지 데이터가 가지는 방대함, 모호성, 그리고 불완전성으로 인하여 사용자가 원하는 정확한 검색 결과를 얻기에 부족함이 많다.

이러한 문제점을 해결하기 위해 최근에 활발히 연

구되고 있는 방법이 내용 기반(contents-based) 검색 방법이다. 이 방법은 이미지 데이터로부터 표현되는 컬러(color), 모양(shape), 질감(texture) 등의 특징(feature)들을 자동으로 추출하여 이 정보를 이용하여 검색하는 방법이다[1,2].

이미지의 내용 기반 검색 방법은 이미지 데이터 내의 다양한 정보를 검색 기준으로 사용하기 때문에 텍스트 기반 방법보다 정확한 검색 결과를 제공한다. 그리고 이러한 다양한 정보를 이용해 사용자 입장에서 시각적이고 편리한 사용자 질의 인터페이스를 제공할 수 있다.

본 연구에서는 이러한 내용 표현 요소들을 자동으로 추출하고, 추출된 특징치를 학습패턴으로 구성하여 역전파(BP; back propagation) 학습 알고리즘을 사용한 신경망(neural network)을 통하여 학습한다. 그리고 웹 에이전트를 통하여 자동으로 획득된 이미지 파일을 정규화하여 검색 엔진에 전송한다. 전송된 데이터에서 특징을 추출하고, 이 특징치를 이용하여 신경망 분류기가 이미지를 분류한다.

## 2. 검색 엔진

본 연구에서 제시하는 이미지 검색 엔진의 전체 구성과 흐름은 그림1과 같다.

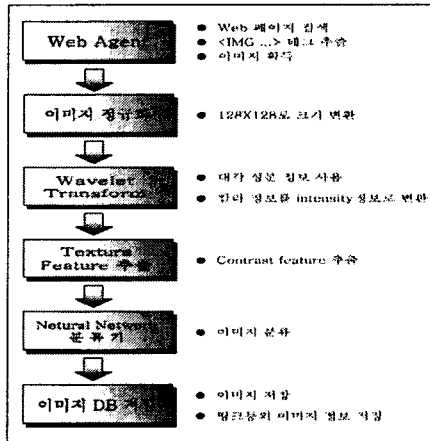


그림1. 검색 엔진의 전체 구성 및 흐름

### 2.1. 웨이블릿 변환(wavelet transform)

웹 에이전트에 의해 획득된 정규화된 이미지는 웨이블릿 변환을 사용하여 공간적인 특징을 추출한다. 웨이블릿 변환은 웨이블릿(wavelet)이라는 기저 함수(band function)를 이동(translating), 확장(dilating)함으로써 이미지 데이터를 주파수 영역으로 변환한다 [3]. 그리고 변환 후에 생성되는 부 밴드들을 분석함으로써 이미지에 대한 정보를 추출할 수 있다.

웨이블릿 변환을 이차원 이미지에 적용하였을 때 4개의 성분으로 나누어진다. 이 4개의 성분은 시각에 민감한 정보를 가지고 있으며, 저주파 성분과 수직 성분, 수평 성분, 그리고 대각 성분으로 나누어진다. 본 연구에서는 이미지의 공간적인 특징을 이용하기 위하여 대각 성분인 고주파(HH2) 영역을 사용하였다. 그림2는 2-채널 필터로 2단계 웨이블릿 변환한 후 이미지의 분할된 영역이다.

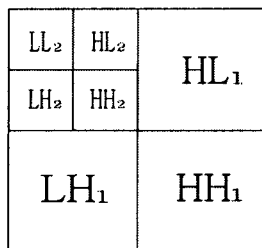


그림2. 웨이블릿 분할영역

### 2.2. 컬러 모델 변환

컬러 이미지 정보를 인지하는 방법은 RGB 컬러 모델과 HSI 컬러 모델이 있다. RGB 요소들은 상호 관계가 너무 크기 때문에 이미지 처리 알고리즘을 수행하기 어렵다[4,5]. 따라서 많은 이미지 관련 응용 프로그램에서 사용되고 있는 HSI 컬러 모델을 사용하여 그 중 명암(intensity) 값을 나타내는 I 값을 사용하였다.

### 2.3. 질감(texture) 특징 추출

일반적으로 이미지 검색에 사용되는 질감 특징은 통계적인 특징, 구조적인 특징과 스펙트럼 특징을 이용하여 표현된다. 통계적인 특징은 이미지의 거칠고 부드러운 정도를 나타낸다. 구조적인 특징은 이미지 내의 수평선의 반복과 같은 규칙적인 배열을 표현한다. 스펙트럼 특징은 높은 에너지 성분을 분석함으로써 이미지 내의 전체적인 주기성을 알아내는데 사용된다[6,7].

질감의 특성을 표현하는 요소에는 질감의 미세 정도를 나타내는 거침(coarseness), 명도의 높낮이 분포를 나타내는 대비(contrast), 질감이 가지고 있는 특정한 방향을 나타내는 방향성(directionality) 등이 있다.

본 연구에서는 이러한 질감 특징들 중, 대비(contrast)와 에너지(energy), 엔트로피(entropy)를 각각 사용하여 분류기의 인식률을 실험해 보았으며, 각 질감 특징들을 구하는 식은 식(1)과 같다.

$$Contrast = \sum_{n=0}^{N-1} n^2 \left( \sum_{i=0}^{N-1} \sum_{j=\frac{N}{2}}^{N-1} W \right) \quad n = |i-j|$$

$$Energy = \frac{1}{n} \sum_{i=0}^{N-1} W(i)^2$$

$$Entropy = - \sum_{i=0}^{N-1} W(i) \log(W(i)) \quad (1)$$

### 2.4. 학습패턴 생성

웨이블릿 변환을 통하여 획득된 원본 이미지의 대각성분 영상의 8×8 영역 단위로 질감 특징치를 구한다. 이 8×8 마스크를 2칸씩 왼쪽에서 오른쪽으로, 위에서 아래로 중복 이동하여 각 영역별 특징치를 구한다. 이를 통해 추출된 49개의 특징치를 신경망을 학습시키기 위한 학습패턴으로 구성한다. 그림3은 질감 특징들 중 대비(contrast)를 이용하여 특징 추출 및 학습패턴을 구성하는 예이다.

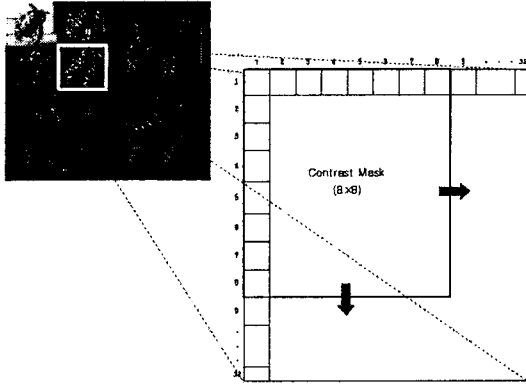


그림3. 학습패턴 생성의 예

2.5. 신경망 분류기(neural network classifier)

위의 과정으로 추출된 특징치를 신경망의 입력 값으로 사용하여 분류기를 구성한다. 신경망의 학습 알고리즘 중 역전파(BP; back propagation) 알고리즘은 LMS법칙을 일반화시킨 것으로 신경망의 실제 출력값과 목표값과의 평균 차이를 최소화하는 경사 탐색기법을 사용한다. 신경망의 구성은 그림4와 같으며, 학습 알고리즘은 아래와 같다[8].

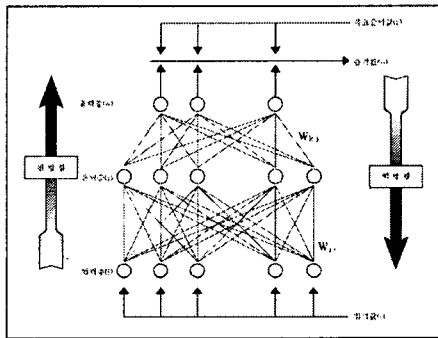


그림4. 신경망의 구성도

본 연구에서는 이미지 데이터에서 추출된 특징치 49개를 입력패턴으로 제시하였다. 신경망은 49개의 입력노드로 구성된 입력층과 49개의 은닉노드로 구성된 은닉층, 그리고 분류할 이미지 종류의 개수인 30개의 출력노드로 구성된 출력층으로 설계하였다. 신경망의 학습은 에러값(minimum error)이 0.1이 될 때까지 연결강도(weight)를 조정하여 학습하였다. 그림5는 위와 같이 구성된 신경망을 학습시키는 프로그램을 실행한 예이다.

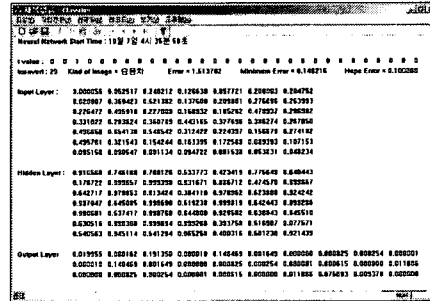


그림5. 신경망 학습 프로그램의 예

3. 실험 및 분석

실험을 위하여 펜티엄 III PC, 윈도우즈 환경에서 비주얼 C++을 사용하여 실험 및 분석을 위한 시스템을 구현하였다. 그림6은 실험 및 분석을 위해 구현한 프로그램을 실행한 예이다.

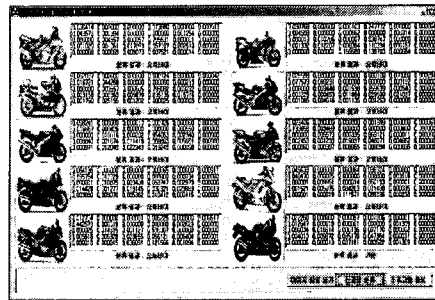


그림6. 신경망 분류 실험 프로그램의 예

각각의 질감 특징을 적용하였을 때의 인식률을 실험한 결과는 표1과 같다. 이미지 종류당 10개의 이미지 30종류, 300개의 이미지를 학습 데이터로 신경망을 학습시켜서 학습 데이터 300개와 테스트 데이터 300개, 총 600개의 이미지를 실험 데이터로 사용하였다.

표1. 인식률 실험 결과

질감 특징	학습데이터	테스트데이터	총 인식률
Contrast	185/300 (61.7%)	159/300 (53.0%)	344/600 (57.3%)
Energy	148/300 (49.0%)	126/300 (42.0%)	274/600 (45.7%)
Entropy	132/300 (44.0%)	115/300 (38.3%)	247/600 (41.2%)

종류별 인식 결과를 분석해 보면, 인식 결과가 높은 이미지 종류는 시계(90.0%), 개(79.0%), 의자(76.5%), 장갑(73.3%) 등이었다. 그리고 인식 결과가 낮은 이미지 종류는 장미(12.5%), 헬기(9.5%), 금낭화(6.4%) 등이었다.

다양한 이미지 종류들을 분석한 결과 각 특징별로 인식 결과가 높고, 낮은 이미지 종류가 상이하였다. 그리고 예를 들어 개와 말, 사과와 배, 상어와 고래 등의 서로 형태나 색상이 비슷한 종류의 이미지들은 오인하는 경우가 많았다. 또한 금낭화, 헬기, 전투함 등과 같은 종류의 이미지들은 형태가 복잡하거나, 이미지 내에 색상 정보가 다양하게 분포되어 있어 낮은 인식 결과를 보여주었다.

#### 4. 결론

본 연구에서는 웨이블릿 변환을 이용하여 이미지의 공간 정보를 획득하고, 여러 가지 특징들 중 질감 특징을 추출하여 이를 학습패턴으로 하는 신경망을 통해 이미지를 분류하는 방법을 제안하였다.

하지만 위의 실험 결과 분석에서와 같이 특징을 하나만 사용할 경우 이미지 종류별로 인식 결과에 대한 차이가 있었다. 이는 각 종류에 적합한 특징들을 모두 또는 다수를 사용하는 방법을 통하여 보다 높은 검색 결과를 얻을 수 있다.

따라서 본 연구는 차후 발전 방향으로 다수의 특징들을 적용할 수 있는 다중 특징(multiple feature)을 추출하고, 이를 다차원 입력노드를 가진 다층 퍼셉트론(multi-layer perceptron) 기반의 신경망을 구성하고자 한다.

그리고 인터넷상의 다양한 이미지들을 보다 효율적으로 분류하기 위하여 ART-2 신경망을 이용한 계층적 신경망(hierarchical neural network)을 구성하여 보다 많은 종류의 이미지들에 대해 높은 검색 결과를 얻을 수 있도록 개선하고자 한다.

#### [참고문헌]

- [1] Batber, W. and Equitz, C. F., "Query By Content for Large On-Line Image Collection", IEEE, 1995
- [2] Ang, Y. H., Li, Z., and Ong, S. H., "Image retrieval based on multidimensional feature properties", SPIE, 2420, 1995, pp.47-57
- [3] 안영아, 장주만, 김지균, 김태윤, "2차원 웨이블릿을 이용한 이미지 검색 시스템 설계", 1999년 한국

- 정보과학회 춘계 학술발표논문집, 6, 1999, pp.1390-1412
- [4] 김진숙, 김호성, "색과 공간정보를 이용한 영상 검색 기법", 한국정보과학회 추계 학술발표논문집, 24, 1997, pp.483-486
- [5] 강영미, "Wavelet과 신경망을 이용한 내용기반 열굴 검색 시스템", 창원대학교 석사논문, 1998
- [6] 김희승, "영상인식-영상처리, 컴퓨터비전, 패턴인식, 신경망", 생능출판사, 1994
- [7] Jain, R., Kasturi, R., and Shunck, B. G., "Machine vision", ISBN, 1995, pp.234-248
- [8] Fausett, L., "Fundamentals of Neural Networks", Prentice Hall, 1994, pp.289-333