

# SOM 기반 웹 이미지 분류에서 고수준 텍스트 특징들의 효과

조 수 선<sup>†</sup>

요 약

본 논문에서는 웹 이미지의 분류 효과를 높이기 위해 이미지 자체에서 추출된 저수준의 비주얼 특징뿐만 아니라 이미지와 관련된 텍스트 정보로부터 나온 고수준 시맨틱 특징들을 이용하는 분류 방법을 제안한다. 이 고수준의 텍스트 특징들은 이미지 URL, 파일명, 페이지 타이틀, 하이퍼링크 및 이미지 주변 텍스트로부터 얻어진다. 분류 엔진으로는 Kohonen의 SOM(Self Organizing Map)을 사용한다. 고수준의 텍스트 특징들과 저수준의 비주얼 특징들을 동시에 사용하는 SOM 기반의 이미지 분류에서는 10개의 카테고리로부터 수집된 200개의 테스트 이미지들이 사용되었다. 분류 성능을 평가하기 위해 간단하면서도 새로운 두 가지 척도, 즉 통일 카테고리 이미지들의 산포 정도와 집적 정도를 나타내는 각각의 척도를 정의하고 사용하였다. 실험결과, SOM기반의 웹 이미지 분류에서는 고수준의 텍스트 특징들이 보다 유용한 것임이 밝혀졌다.

키워드 : 웹 이미지 분류, 고수준 텍스트 특징, 자기조직화지도

## The Effectiveness of High-level Text Features in SOM-based Web Image Clustering

Soosun Cho<sup>†</sup>

Abstract

In this paper, we propose an approach to increase the power of clustering Web images by using high-level semantic features from text information relevant to Web images as well as low-level visual features of image itself. These high-level text features can be obtained from image URLs and file names, page titles, hyperlinks, and surrounding text. As a clustering engine, self-organizing map (SOM) proposed by Kohonen is used. In the SOM-based clustering using high-level text features and low-level visual features, the 200 images from 10 categories are divided in some suitable clusters effectively. For the evaluation of clustering powers, we propose simple but novel measures indicating the degrees of scattering images from the same category, and degrees of accumulation of the same category images. From the experiment results, we find that the high-level text features are more useful in SOM-based Web image clustering.

Key Words : Web Image Clustering, High-level Text Features, Self-Organizing Map

### 1. 서 론

웹에서 디지털 이미지의 사용이 급격히 증가함에 따라, 웹 이미지의 분석, 분류, 그리고 검색 기술 등이 주요한 연구 주제가 되고 있다. 특히 웹을 거대한 분산 멀티미디어 데이터베이스로 보고, 그것을 다루는 방법론과 마이닝 알고리즘들이 이미지 마이닝의 한 분야로 자리잡고 있다. 연구[1]에서 제공된 웹 마이닝의 분류에서, 웹 콘텐츠 마이닝은 '웹 페이지 콘텐츠 마이닝', '텍스트 마이닝', '이미지 마이닝' 등을 포함한다.

한편, 이미지 마이닝의 주된 작업인 이미지 검색과 분류에서는 콘텐츠 기반 기술에 관심이 집중되고 있다. 주석 처리

되지 않은 대부분의 이미지를 검색하기 위해 다양한 콘텐츠 기반 이미지 검색 (CBIR: Content-Based Image Retrieval) 시스템들이 개발되어 왔다[2]. 이 CBIR 분야는 지난 10년간 획기적인 발전을 이루었다[3, 4]. CBIR은 이미지로부터 비주얼 특징들을 뽑아내고, 멀티미디어 데이터베이스 상의 이미지들을 인덱싱하기 위해 이 비주얼 특징들을 이용하는 아이디어에 기초를 두고 있다. 이미지들 간의 유사도를 결정하는 비교 방법은 이미지 특징들의 표현과 적당한 거리 함수(distance function)에 달려있다. 따라서 CBIR은 쿼리 이미지의 비주얼 특징들, 즉, 칼라 히스토그램, 질감, 모양 등에서 비슷한 값을 갖는 이미지를 찾도록 만들어진다.

그러나 CBIR은 단지 비슷한 특징들을 가진 이미지를 검색함으로써 전반적으로 관련이 없는 이미지를 리턴하는 취약점이 있다. 실제로 CBIR의 성능은 대부분 낮은 편이다. 주된

<sup>†</sup> 종신회원 : 국립중앙대학교 컴퓨터학과 조교수  
논문접수 : 2005년 9월 5일, 심사완료 : 2006년 2월 23일

이유는 저수준의 비주얼 특징들만으로는 이미지의 고수준 시맨틱 컨텐트를 표현할 수 없기 때문이다. 따라서 몇몇 연구들[5, 6]은 이미지 검색을 위해 저수준 특징들과 고수준 특징들을 결합하는 방법에 초점을 두고 있다.

본 논문에서는 웹 이미지 분류에서 저수준의 비주얼 특징들뿐만 아니라 웹 이미지와 관련된 텍스트 정보로부터 나온 고수준의 시맨틱 특징들을 이용함으로써 분류 성능을 얼마나 높일 수 있는지 실증적으로 확인해 보고자 한다. 이들 고수준 텍스트 특징들은 이미지 URL, 파일명, 페이지 타이틀, 그리고 주변 텍스트들로부터 얻어진다. 분류 엔진으로는 Kohonen의 자기조직화지도(SOM: Self Organizing Map)[7]를 이용한다.

논문의 나머지는 다음과 같이 구성된다. 제2절에서는 관련 연구들을 소개하고, 제3절에서는 특징 추출기와 SOM 기반 분류기를 포함하는 웹 이미지 분류 시스템 프로토타입의 구조와 기능들을 설명한다. 이미지 분류 실험 및 평가는 제4절에서 논의되고, 마지막으로 제5절에서 결론을 맺는다.

## 2. 관련 연구

본 연구에서 우리의 시도는 SOM에 기반한 웹 이미지 클러스터링이며, 이미 기존의 연구에서 SOM을 이용한 웹 이미지의 분류 또는 검색에 관한 것을 찾아볼 수 있다. 연구[8]에서는 SOM을 이용한 컨텐트 민감형의 풍부한 속성을 가진 이미지 분류기를 발표하였다. 이 시스템은 3가지 기능을 내세우는데 이는 다양한 저수준 이미지 특징들에 기반한 이미지 특징 추출, 유사성 분석 및 SOM을 이용한 이미지 분류 등이다. 이 연구는 칼라 기반, 예지 기반, 영역 기반 및 질감 기반의 4가지 이미지 특징들을 사용하였고, 100개의 이미지를 대상으로 그들의 방법을 평가하였다. SOM에서의 분류 평가 결과, 유사 이미지들은 동일 영역에 배치됨을 보였다. 이 연구에서는 다양한 비주얼 특징들을 채택하여 이미지 분류의 약점을 극복하고자 하였으며 본 연구에서 다루고자 하는 고수준 텍스트 특징들은 고려되지 않았다.

SOM을 이용한 대표적인 CBIR 시스템은 PicSOM[9, 10]이다. PicSOM은 트리구조 SOM (Tree structured SOM)에 기반을 두고 있다. 참조 이미지 집합이 주어지면, PicSOM은 그것과 유사한 이미지 집합을 검색할 수 있다. 유사성의 척도로서 트리구조 SOM을 이용한다. 그들은 이미지 특징으로 'average color', 'color moments', 'texture neighborhood', 'shape histogram', 'shape FFT' 등을 사용하였다. PicSOM의 최근 버전[11]에서는 유사성 척도로 다수의 병렬 트리구조 SOM을 이용하고, 이미지 특징으로는 MPEG-7 컨텐트 디스크립터와 같은 통계적 비주얼 데이터를 사용한다. 역시 고수준의 텍스트 특징은 고려되지 않았다.

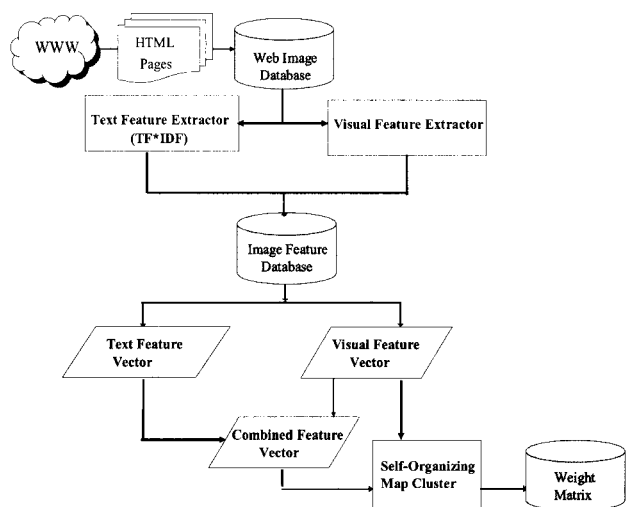
한편, 웹 이미지 검색에서 텍스트 특징들을 이용한 대표적인 연구로서, Zheng Chen 등[5]은 웹 마이닝을 이용한 웹 이미지 검색 엔진을 발표하였다. 이 연구는 이미지를 묘사하는 웹 페이지 상의 텍스트 정보를 추출하고, 이를 다른 저수준

의 이미지 특징들과 결합함으로써 검색의 효과를 높일 수 있었다. 또한 이들은 이미지 검색 성능을 높이기 위해 사용자의 피드백 로그 정보에 데이터 마이닝을 적용하였다. 이 연구에서 쿼리 이미지와 검색된 후보 이미지 사이의 유사도는 두 이미지의 특징 벡터들 간의 내적(dot product)으로 계산되었으며 SOM을 적용한 분류와는 거리가 있다.

웹 이미지 검색에서 SOM을 기반으로 하거나, 텍스트 특징들을 이용한 여러 관련 연구와 달리, 본 연구의 관심은 SOM 기반의 분류기에서 텍스트 특징을 이용할 때 분류의 성능이 얼마나 좋아지는가를 밝히는 데 있다. 따라서 우리는 두 가지 종류의 이미지 특징 벡터인 저수준 비주얼 특징 벡터와 고수준 텍스트 특징 벡터를 따로 구성하고, 200개의 웹 이미지로부터 이들 특징들을 추출하여 분류에 이용하였다. 분류 성능을 측정하기 위해 서로 다른 특징 벡터를 이용하여 SOM 분류기에서 각각 실험하였으며, 평가를 위해 동일 카테고리 이미지들의 산포 정도와 집적 정도를 나타내는 두 가지 척도를 정의하여 사용하였다.

## 3. 제안된 웹 이미지 분류 시스템

본 논문에서 제안되는 웹 이미지 분류 시스템의 구조는 (그림 1)과 같다. 이것은 3가지 주요 컴포넌트인 텍스트 특징 추출기, 비주얼 특징 추출기 및 SOM 기반 분류기로 구성된다. 특징 추출에 앞서 첫 번째 단계로, 이미지 분류를 위해 필요한 웹 이미지와 관련 텍스트 정보가 수집되며, 수집된 소스 이미지와 관련 텍스트 정보가 웹 이미지 데이터베이스에 저장된다. 특징 추출기의 프로세싱을 거친 후, 추출된 특징들은 이미지 특징 데이터베이스에 저장된다. SOM 기반의 분류기는 분류 시스템의 두뇌 역할을 하는 지능적인 컴포넌트이다. SOM 기반 분류기의 훈련과정에서, 소스 이미지를 수집한 각 오브젝트 카테고리의 가중치 행렬이 훈련의 결과물로 저장되며 이것은 SOM 분류기의 테스트를 위해 사용될 수 있다.



(그림 1) 웹 이미지 분류 시스템의 구조

### 3.1 이미지 수집

SOM 기반 이미지 분류 시스템에 적용하기 위해 10개의 카테고리로부터 200개의 이미지가 수집되었다. 10개의 카테고리는 'monkey', 'tomato', 'building', 'tree', 'boat', 'vehicle', 'bird', 'fish', 'bridge', 'rose'를 포함한다. 후보 이미지들의 수집은 모바일 에이전트의 일종인 웹 스파이더에 의해 수행될 수 있다. 웹 스파이더는 다수의 지정된 웹 페이지뿐만 아니라 그들과 하이퍼링크로 연결된 웹 페이지들을 방문하여 분석하고 관심있는 이미지들을 로컬의 웹 이미지 데이터베이스로 내려 받는다. 후보 이미지들로부터 각 카테고리마다 20개의 이미지와 관련 텍스트 정보를 선택한다. 따라서 수집된 전체 이미지의 수는 200개가 된다.

### 3.2 저수준 비주얼 특징들의 추출

인간의 감각은 주관적이기 때문에 하나만으로 충분한 비주얼 특징은 없다[3]. 더구나 하나의 특징에 대한 한가지 표현 또한 불충분하다. 따라서 특징에 대한 다수의 표현과 그 조합이 요구된다. 일반적으로 이미지는 칼라, 질감, 모양, 예지, 그림자 등의 특징들을 갖는다. 이 중 가장 유용한 특징은 칼라와 질감이다. 우리의 웹 이미지 분류 시스템에서는 5가지 칼라 히스토그램 특징들과 5가지 질감 특징들을 사용한다.

- **칼라 히스토그램 특징들** - 칼라 히스토그램 특징은 서로 다른 칼라 밴드(예를 들면 red, green, blue와 같은)에서 각 칼라 값의 레벨에 얼마나 많은 픽셀들이 분포하고 있는지를 나타내는 확률분포에 기초를 둔다[12]. 8비트 이미지라면 각 칼라 밴드 별로 256개의 레벨을 가지게 된다. 사용되는 5가지 칼라 히스토그램 특징은 평균, 표준편차, 비대칭도, 에너지, 그리고 엔트로피이다. 평균은 이미지 전체의 칼라 레벨 평균값을 나타내고, 표준편차와 비대칭도는 칼라의 대비 정도를 나타낸다. 에너지는 히스토그램 확률의 제곱합으로 계산되는데 최대값인 1인 경우, 이미지는 오직 한가지의 칼라 값을 가지는 것을 의미하고, 에너지 값이 작을수록 여러 가지 칼라 값으로 구성됨을 의미한다. 엔트로피는 이미지 칼라 분포의 복잡도를 에너지와는 반대로 나타내는 척도이다. 단순한 칼라 분포의 이미지는 낮은 엔트로피 값을 갖고, 복잡한 칼라 분포의 이미지는 높은 엔트로피 값을 가진다. 저수준 이미지 특징 추출에서는  $5 \times 3 = 15$  차원의 칼라 히스토그램 특징 벡터가 생성된다.
- **질감 특징들** - 질감은 영역 또는 블록 형태의 근접 특징으로 정의된다. 근접한 픽셀에 대한 각 픽셀의 변화량이 질감을 결정하는 것이다. 따라서 질감 특징은 픽셀 거리 값을 요구하며 다양한 질감 특징들을 계산하기 위해 co-occurrence 행렬,  $C_{a,d}(i, j)$ 를 구성하여 사용한다.  $C_{a,d}(i, j)$ 는 칼라 레벨  $i$ 와  $j$ 를 가지며 방향  $a$ , 거리  $d$ 만큼 떨어져 있는 픽셀들의 co-occurrence 행렬이다. 픽셀 쌍들의 거리가 길수록 거친 질감을 나타내고, 거리가 짧을수록 가는 질감을 나타낸다[13]. 질감분석에서 가장 일반적으로 사용되는 특징은 서브 밴드 계수의 에너지

이며  $\sum_i \sum_j C(i, j)^2$ 로 계산된다[12]. 비주얼 특징 추출에서 뽑아내는 5가지 질감 특징들은 에너지, 관성, 상관관계, 역 디퍼런스, 그리고 엔트로피이다. 각 질감 특징들은 범위 값과 평균 값을 가지므로 생성되는 특징 벡터는  $5 \times 2 = 10$ 차원이 된다.

### 3.3 고수준 텍스트 특징들의 추출

특징 추출의 결과 형태는 SOM 분류기의 성능에 큰 영향을 미친다. 사용되는 텍스트 특징들은 해당 이미지를 포함하고 있는 웹 페이지의 여러 텍스트 소스들로부터 경험적인 규칙에 따라 추출된다.

- **이미지 파일명과 URL** - 대체로 웹 이미지에는 의미있는 이름이 지정된다. 그러나 파일명과 URL의 유용성을 높이기 위해 다음과 같은 몇 가지 처리가 필요하다. (1) 파일명 필터링: 파일명에서 몇몇 구분자, 예를 들면 숫자, 하이픈, 밑줄, 파일명 확장자와 같은 것은 미리 제거되어야 한다. (2) URL 파싱: URL은 대부분 웹 페이지의 이미지에 대한 계층 정보를 나타낸다. URL로부터 유용한 정보를 얻기 위해서는 URL 파싱이 필요하다.
- **주변 텍스트** - 많은 웹 페이지에서 이미지는 저자가 나타내고자 하는 내용을 강화하기 위해 사용된다. 따라서 이미지 주변의 텍스트들은 그 이미지의 내용과 의미론적으로 관련이 있다고 여겨진다. 우리는 이미지의 상, 하, 좌, 우, 4가지 가능한 지역으로부터 텍스트를 추출하여 사용한다.
- **웹 페이지 제목 및 하이퍼링크** - 웹 페이지 제목과 이미지의 하이퍼링크들은 이미지 내용과 관련이 있을 수 있다. 따라서 이들 또한 이미지의 텍스트 특징 소스로 이용될 수 있다.

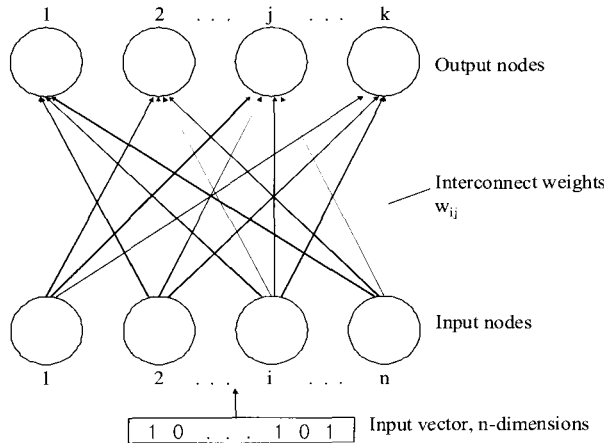
텍스트 특징들을 나타내기 위해 텍스트 특징 벡터가 구성되는데, 텍스트 특징 벡터의 각 키워드에 가중치를 주기 위해 TF\*IDF(term frequency & inverse document frequency) 방법[14]이 사용된다. 앞에서 언급한 것처럼, 우리는 SOM기반의 분류를 위해 10개 카테고리의 오브젝트 이미지들을 사용한다. 따라서 키워드는 'monkey', 'tomato' 등과 같은 10개의 카테고리 이름으로 주어진다. 텍스트 특징 벡터를 구성하는 규칙은 다음 식과 같다.

$$D_i = TF_i * IDF_i = (t_{i1} * \log(\frac{N}{n_1}), \dots, t_{ij} * \log(\frac{N}{n_j}), \dots, t_{im} * \log(\frac{N}{n_m})) \quad (1)$$

여기서  $D_i$ 는 이미지  $i$ 의 특징 벡터이고,  $t_{ij}$ 는 이미지  $i$ 의 텍스트 디스크립션에 나타나는 키워드  $j$ 의 빈도수이다.  $n_j$ 는 키워드  $j$ 를 포함하는 이미지들의 수를 나타내고,  $N$ 은 전체 이미지의 수이다. 구성되는 텍스트 특징 벡터는 키워드, 즉 카테고리 수에 따라 10차원이 된다.

### 3.4 자기조직화지도(Self-Organizing Map)

자기조직화지도는 클러스터링 작업에 사용될 수 있는 신경망의 일종이며 시트 모양의 신경망 배열로 시각화될 수 있



(그림 2) SOM의 구조

다. 다양한 입력 신호들에 대해 맵의 셀들은 정렬된 모양으로 조정된다. 학습과정은 경쟁적이고 무감독이므로 어떤 입력에 대한 올바른 출력을 정의할 필요가 없다. 또한 입력층과 출력층으로 구성되는 네트워크의 구조상 여러 단계의 피드백이 아닌 단 하나의 전방 패스만 사용하므로 수행 속도가 상당히 빠른 모델로서, 잠재적으로 실시간 학습처리를 할 수 있는 모델이다. 이러한 특징은 본 연구에서 응용하고자 하는 웹 이미지 분류 또는 검색 시스템에 적용될 수 있는 유용한 장점이 된다.

SOM 신경망 알고리즘은 인접한 출력 뉴런들이 비슷한 기능을 수행할 것이라는 가정하에 기존의 경쟁학습을 개선하여 입력 벡터와 가장 가까운 출력 뉴런과 그 이웃 뉴런들을 함께 학습시키는 알고리즘이다. (그림 2)는 SOM의 구조를 나타낸다.  $n$ 차원의 입력 데이터를 표현하는  $n$ 개의 입력 노드들과  $k$ 개의 분류영역을 표현하기 위한  $k$ 개의 출력 노드들로 구성되어 있다. 모든 입력 노드들은 모든 출력 노드들과 연결되어 있고 연결 가중치를 가진다.

SOM 알고리즘에서 입력 샘플 집합은 실수 벡터  $\mathbf{x}(t) \in \mathbf{R}^n$ 로 표현된다. 여기서  $t$ 는 샘플의 인덱스를 나타낸다. 출력 맵에서 각 노드  $i$ 는 가중치 벡터  $\mathbf{w}_i(t) \in \mathbf{R}^n$ 를 가지며 그 엘리먼트 수는 입력 벡터  $\mathbf{x}(t)$ 의 수와 같다.  $\mathbf{w}_i(t)$ 에서도  $t$ 는 샘플의 인덱스를 나타낸다. 가중치 벡터  $\mathbf{w}_i(t)$ 의 초기값은 무작위로 선택될 수 있다. 그러나 실제 응용에서 가중치 벡터는 어떤 정렬된 방법으로 적절하게 초기화된다. 각 입력 아이템은 어떤 메트릭에서 입력 벡터  $\mathbf{x}(t)$ 와 가장 잘 매치되는 가중치 벡터  $\mathbf{w}_i(t)$ 로 매핑된다. 자기조직화 알고리즘은 다음과 같은 기본 태스크들의 반복으로 정렬된 매핑을 생성한다.

- (1) 입력 벡터  $\mathbf{x}(t)$ 는 모든 가중치 벡터  $\mathbf{w}_i(t)$ 와 비교된다. 맵 상에서 최상의 매칭 유닛(노드), 즉, 어떤 메트릭(예를 들면 유클리드 거리)에서 가중치 벡터가 입력 벡터와 가장 유사한 노드가 결정된다. 이 최상 매칭 유닛(BMU: Best Matching Unit)은 승자(winner)로 불린다.
- (2) 승자의 가중치 벡터와 배열 상의 그 이웃 노드들은 학습 원칙에 따라 입력 벡터를 향해서 바뀌어진다. SOM 학습 프로세스에서 이상적인 경우는 각 입력 벡터  $\mathbf{x}(t)$

에 대해 승자와 그 이웃 노드들이 입력 데이터 공간의  $\mathbf{x}(t)$ 로 근접하여 변하는 것이다. 학습 프로세스 동안 하나하나의 변화는 서로 모순될 수도 있지만, 프로세스에서 결과는 가중치 벡터  $\mathbf{w}_i(t)$ 를 위한 정렬된 값들이 배열상에 나타나는 것이다. 사용 가능한 입력 샘플의 수가 제한적이라면 샘플들은 SOM 알고리즘에 반복적으로 제공되어야 한다. 학습 프로세스에서 가중치 벡터의 적합화는 다음 식들에 따라 이루어진다.

$$\begin{aligned} \mathbf{w}_i(t+1) &= \mathbf{w}_i(t) + \alpha(t)[\mathbf{x}(t) - \mathbf{w}_i(t)] && \text{for each } i \in N_c(t), \\ \mathbf{w}_i(t+1) &= \mathbf{w}_i(t) && \text{otherwise} \end{aligned} \quad (2)$$

여기서  $t$ 는 변수들의 이산 시간 인덱스이고,  $\alpha(t) \in [0,1]$ 는 학습 단계의 상대적 크기를 나타내는 스칼라 값이다.  $N_c(t)$ 는 맵 배열에서 승자 주변의 이웃 노드들을 나타낸다. 학습 프로세스의 초반에는 이웃의 반지름이 매우 크지만 학습과정에서 점점 좁아진다.  $\alpha(t)$  또한 학습과정에서 감소한다.

#### 4. 실험 및 평가

SOM 기반 분류기는 'monkey', 'tomato', 'vehicles' 등, 10개 카테고리로부터 나온 오브젝트 이미지들로 구성된 훈련 데이터 시리즈로 훈련된다. 이미지들은 현재 가동 중인 웹 사이트에서 수집되었으며, 대표 이미지 특징 벡터를 구성하기 위해 이들로부터 앞에서 언급된 비주얼 특징 및 텍스트 특징이 추출되었다. 비주얼 특징과 텍스트 특징을 사용함으로써 얻어지는 분류기의 성능을 조사하기 위해, 이 두 가지 종류의 이미지 특징들을 서로 다른 조합으로 묶어서 몇 가지 분류 실험에 적용하였다. 3가지의 특징 벡터들의 조합이 가능한데 그 것은 25차원의 칼라 및 질감 특징 벡터, 25차원의 칼라 및 텍스트 특징 벡터, 그리고 20차원의 질감 및 텍스트 특징 벡터이다.

##### 4.1 실험 방법 및 분류 맵

훈련을 위해 10 x 10 격자를 가지는 100개 노드의 SOM이 사용되었다. 훈련 사이클은 10회, 학습 파라미터  $\alpha$ 는 초기값 0.5에서 종료값 0.1까지 기하급수적으로 감소시켰다. 이웃함수는 가우시안을 사용하였고 이웃함수의 크기는 맵 전체 폭의 50%부터 1%까지 역시 기하급수적으로 감소시키며 실험하였다. 여기서 사용된 노드의 수 및 기타 파라미터 값들은 훈련 집합에서 유사한 클러스터들의 수를 고려하여 여러 번의 비교 실험을 거쳐 경험적으로 결정되었다.

(그림 3)은 저수준의 비주얼 특징인 칼라히스토그램 특징들과 질감 특징들을 사용했을 때 SOM 분류의 결과로 나타난 맵이다. 각 노드에는 굵은 글씨체로 클러스터의 사이즈가 표시되어 있고 각 카테고리 별로 가장 많은 데이터가 분류되는 노드에 카테고리 명이 표시되어 있다. 괄호안의 수는 중복 수이다. 그림에서 보듯이 10개 이상의 데이터를 포함하는 노드 수는 모두 7개이고, 'vehicle'와 'boat', 'tree'와 'bird', 'monkey'와 'building' 카테고리들은 동일 노드에 포함되어

28 Monkey(10) Building(7)		6	17 Fish(10)				6
1	1		1	4			
1			2				
1							
4	1						
	4						16 Bride(13)
4			11 Tomato(10)	6			36 Vehicle(14) Boat(11)
4			16 Tree(7) Bird(9)	4			
1							5
1		2					19 Rose(14)

(그림 3) 비주얼 특징 벡터를 사용하였을 때의 SOM 분류 맵

있음을 알 수 있다. 또한 유사한 비주얼 특징을 가지는 'vehicle'와 'boat', 'bridge' 등은 같은 노드나 이웃 노드에 분포되어 있음을 알 수 있다. 나머지 2가지 특징 벡터들의 조합, 즉 25차원의 칼라 및 텍스트 특징 벡터, 그리고 20차원의 질감 및 텍스트 특징 벡터를 사용하였을 때 각각의 SOM 분류 맵 형태는 (그림 3)과 뚜렷한 차이가 없고 정량적인 비교가 어려웠다. 따라서 좀 더 구체적인 분석을 위해 분류의 성능을 나타내는 정량적인 척도를 도입하여 비교해 보고자 한다.

4.2 분류 성능 척도

본 실험에서는 향상된 SOM 분류의 결과를 얻기 위해 각종 파라미터를 조절하는 것보다 입력 데이터를 구성하는 특징 벡터들의 종류에 따라 어떤 분류 결과가 나타나는지 알아보는 것이 더 중요하게 여겨진다. 고수준 텍스트 특징들이 얼마나 효과적인지 밝혀내고자 하기 때문이다. 이를 위해서는 분류 결과를 분석하고 성능을 측정할 수 있는 도구가 필요하다. 본 논문에서는 서로 다른 이미지 특징 벡터를 사용하였을 때 나타나는 분류기의 성능을 측정하기 위한 도구로서, 동일 카테고리로부터 나온 이미지들의 산포 정도와, 역시 동일 카테고리로부터 나온 이미지들이 하나의 결과 분류 노드에 집적되는 정도를 나타내는 척도를 제안한다. 먼저 어떤 카테고리  $i$  에 대하여 산포 정도를 나타내는 척도인  $S_i$  는 다음과 같이 정의된다.

$$S_i = n_i + N_i, \quad \text{if } N_i \text{ is larger than } 1, \\ = n_i, \quad \text{if } N_i \text{ equals } 1, \quad (3)$$

여기서  $n_i$  는 1개 또는 2개의  $i$  카테고리 이미지를 포함하는 분류 노드들의 수이고,  $N_i$  는 3개 이상의  $i$  카테고리 이미지를 포함하는 분류 노드들의 수이다. 하나의 카테고리로부터 나온 이미지들이 여러 개의 다른 분류 노드에 흩어져있을 때  $S_i$  값은 커진다. 이상적인 경우는 동일 카테고리로부터 나온 대부분의 이미지들이 하나의 분류 노드에 포함되는 것이다. 이 경우에  $N_i$  는 1이 되고 이것은  $S_i$  값에 추가되지 않는다.

동일 카테고리 이미지들이 하나의 결과 분류 노드로 집적되는 정도를 나타내는 척도,  $A_i$  는 다음과 같이 정의된다.

$$A_i = \text{Max}_{j=1}^{c_i} M_{ij} \quad (4)$$

여기서  $c_i$  는  $i$  카테고리 이미지들이 포함되는 분류 노드들의 수이고,  $M_{ij}$  는 분류 노드  $j$  에 포함되는  $i$  카테고리 이미지들의 수이다. 따라서  $A_i$  는  $i$  카테고리 이미지들이 가장 많이 포함된 분류 노드에서 그 이미지들의 수이다.

4.3 분류 성능 비교

제안된 2가지 척도를 사용하여 SOM 분류 결과를 분석한다. <표 1>은 25차원의 칼라 및 질감 특징 벡터, 25차원의 칼라 및 텍스트 특징 벡터, 그리고 20차원의 질감 및 텍스트 특징 벡터를 적용한 경우, 각각의 성능을 비교한 것이다. 식 (3), (4)에 정의된 바에 따르면  $S_i$  는 작을수록,  $A_i$  는 클수록 더 좋은 성능이라고 할 수 있다. <표 1>에서 보듯이 칼라 및 질감 특징 벡터를 사용하였을 때,  $A_i$  의 합은 105,  $S_i$  의 합은 43이지만, 칼라 및 텍스트 특징 벡터를 사용하였을 때,  $A_i$  의 합은 119로 증가하였고,  $S_i$  의 합은 38로 감소하였다. 질감 및 텍스트 특징 벡터를 사용한 경우는 성능 향상의 정도가 더 큰 것을 알 수 있다. 다양한 특징 벡터 조합을 사용했을 때 각각의 SOM 분류 맵만으로는 그 성능을 비교하기 어려웠으나 본 논문에서 제안하는 척도를 사용함으로써 텍스트 특징 벡터의 사용이 분류 성능에 더 좋은 영향을 미친다는 결과를 얻을 수 있다.

<표 1> 사용된 특징들의 조합에 따른 분류기 성능 비교

Category $i$	Degree of Accumulation, $A_i$			Degree of Scattering, $S_i$		
	color & texture	color & text	texture & text	color & texture	color & text	texture & text
1	10	7	10	4	4	4
2	10	15	12	5	3	3
3	7	10	13	5	4	4
4	7	10	11	5	5	3
5	11	10	16	3	5	2
6	14	14	15	3	3	3
7	9	11	15	5	2	3
8	10	14	12	4	4	4
9	13	15	12	5	4	4
10	14	13	10	4	4	5
SUM	105	119	126	43	38	35

5. 결 론

본 논문에서는 SOM기반의 웹 이미지 분류에 대한 새로운 시도를 소개하였다. 이 시도는 분류하고자 하는 이미지의 특

정으로 저수준의 비주얼 특징뿐만 아니라 고수준의 텍스트 특징을 동시에 이용하는 것이며, 그 결과 얼마만큼의 분류기 성능의 향상이 있었는지 측정해 보고자 하는 것이었다. 이를 위해 두 가지 종류의 이미지 특징들을 서로 다른 조합으로 묶어서 SOM 기반 분류기에 적용하여 그 성능을 평가하였다. 성능 평가를 위해 우리는 간단하면서도 새로운 2가지 척도, 즉, 동일 카테고리 이미지들의 산포 정도와 집적 정도를 나타내는 각각의 척도를 정의하고 사용하였다. 실험 및 평가에서는 칼라 히스토그램 특징들과 질감 특징들만 사용하였을 때보다 텍스트 특징들이 각각 추가되었을 때 더 큰 분류 성능이 나타남을 알 수 있었고 따라서 SOM 기반의 이미지 분류에서 텍스트 특징의 유효성을 확인할 수 있었다. 이는 현재의 CBIR 시스템에서 검색 또는 분류의 취약성을 극복하기 위해 이미지의 분류 또는 인덱싱에 적절한 고수준 텍스트 특징들을 사용하는 것이 하나의 해결책이 될 수 있음을 시사한다.

### 참 고 문 헌

[1] D. Arotaritei and S. Mitra, "Web mining: a survey in the fuzzy framework," *Fuzzy Sets and Systems*, Vol.148, pp.5-19, 2004.

[2] N. Gudivada and V. Raghavan, "Content-based image retrieval systems," *IEEE Computer*, Vol.28, No.9, pp.18-22, 1995.

[3] Y. Rui, T. Huang and S. Chang, "Image retrieval: current techniques, promising directions and open issues," *Journal of Visual Communication and Image Representation*, Vol.10, No.4, pp.39-62, 1999.

[4] 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 Transactions on Pattern Analysis and Machine Intelligence*, Vol.22, pp.1349-1379, 2000.

[5] Z. Chen, L. Wenyan, F. Zhang, M. Li and H. Zhang, "Web Mining for Web Image Retrieval," *Journal of the American Society for Information Science and Technology*, Vol.52, No.10, pp.831-839, 2001.

[6] Y. Lu, C. Hu, X. Zhu, H. Zhang and Q. Yang, "A unified framework for semantics and feature based relevance feedback in image retrieval systems," In proceedings of the 8th ACM international Conference on Multimedia, pp.31-38, 2000.

[7] T. Kohonen, 'Self-Organizing Maps', New York : Springer-Verlag, 1997.

[8] S.W.K. Chan and M.W.C. Chong, "Unsupervised clustering for nontextual web document classification," *Decision Support Systems*, Vol.37, pp.377-396, 2004.

[9] J. Laaksonen, M. Koskela, S. Laakso and E. Oja, "PicSOM-content-based image retrieval with self-organizing maps," *Pattern Recognition Letters*, Vol.21, pp.1199-1207, 2000.

[10] J. Laaksonen, M. Koskela, S. Laakso and E. Oja, "Self-organizing maps as a relevance feedback technique in content-based image retrieval," *Pattern analysis & Applications*, Vol.4, pp.140-152, 2001.

[11] J. Laaksonen, M. Koskela and E. Oja, "PicSOM-Self-organizing image retrieval with MPEG-7 content descriptions," *IEEE Transactions on Neural Networks*, Special Issue on Intelligent Multimedia Processing, Vol.13, pp.841-853, 2002.

[12] S. E. Umbaugh, Y. Wei, and M. Zuke, "Feature Extraction in Image Analysis," *IEEE Engineering in Medicine and Biology*, Vol.16, pp.62-73, 1997.

[13] T. Chang and C.-C.J. Kuo, "Texture analysis and classification with tree-structured wavelet transform," *IEEE Tr. IP*, Vol.2, pp.429-441, 1993.

[14] G. Salton, 'Automatic Text Processing', Addison-Wesley, pp.279-281, 1989.

### 조 수 선



e-mail : sscho@chungju.ac.kr

1987년 서울대학교 계산통계학과(이학사)

1989년 서울대학교 계산통계학과(이학석사)

2004년 충남대학교 대학원 컴퓨터과학과

(이학박사)

1989년~1994년 ㈜웅진미디어 CBE개발부

연구원

1994년~2004년 한국전자통신연구원 디지털홈연구단 선임연구원

2004년~현재 국립충주대학교 컴퓨터과학과 조교수

관심분야: 웹 데이터 마이닝, 기계학습