

論文2000-37CI-1-5

## 문자기반 유사상표 검색을 위한 가중치 부여 근사매칭

## (Weighted Approximate Matching for Character-based Similar Trademark Retrieval)

徐昌德\*, 金會律\*

(Changduck Suh and Whoi-Yul Kim)

## 요 약

본 논문에서는 특허청에 등록된 상표들 중에서 90%를 차지하는 문자기반 상표를 대상으로 질의상표와 유사한 상표들을 보다 정확하게 검색하기 위한 방법을 제안한다. 제안한 방법은 상표 칭호에 대한 근사매칭으로 1차 유사도를 계산하며 동일 유사도 그룹에 대해 상표 외형에 대한 영상 처리로 2차 유사도를 계산하여 순위를 부여한다. 전체 시스템의 한 부분으로 기존 도형상표와 함께 구현하였으며, 성능평가지 기존 근사매칭보다 16.2%의 정확도 향상을 보였다.

## Abstract

Character-based trademarks constitute 90% of registered trademarks at the Korean Patent Office. This paper proposes a method to improve the precision rate when searching for similar trademarks in such systems. The proposed method first calculates the similarity measure by an approximate matching for the strings of trademarks followed by the similarity measure by an image processing. The method has been implemented and merged with the existing device-mark retrieval system to improve precision rate by 16.2% compared to other approximate matching methods.

## I. 서 론

멀티미디어 시대, 정보화 시대를 맞이하여 다양한 형태의 매체가 급속하게 증가하고 있다. 과거 텍스트 위주의 정보에서 현재는 영상, 동영상, 소리 등의 여러 매체가 함께 제공되고 있으며 특히 인터넷의 발달로 대용량의 영상 정보들을 손쉽게 접할 수 있게 되었다. 하지만 이들을 검색하기 위해서는 인간의 많은 인지적 노력이 필요하며, 이를 컴퓨터로 자동 검색하기 위한

연구가 계속되고 있다. 영상에 대한 특징을 자동으로 추출하여 검색에 이용하는 내용 기반 영상 정보 검색은 여러 분야에 적용될 수 있는데 그 중 하나가 상표(trademark) 관련 분야이다.

특허청에 따르면 최근 인터넷 관련 상표와 밀레니엄, Y2K 등 세기말과 관련한 상표 및 IMF 관련 상표의 출원이 꾸준히 증가하고 있으며<sup>[1]</sup>, '99 상반기 중 국내 상표출원은 모두 52,362건으로 전년 동기에 비해 58.2%나 급증한 것으로 나타났다<sup>[2]</sup>. 상표법상의 상표란 상품을 생산·가공·증명 또는 판매하는 것을 업으로 하는 자가 자기의 상품을 타인의 상품과 식별되도록 하기 위하여 사용하는 標章<sup>[2]</sup>을 말하는 것으로 상표심사기준 중 제19조에서는 先등록상표와 동일 또는 유사한 상표에 대한 기준이 명시되어 있다.

심사관들은 출원된 상표가 제19조에 의거한 유사성

\* 正會員, 漢陽大學校 電子電氣工學部

(Division of Electrical &amp; Computer Engineering, Hanyang University)

接受日字: 1999年10月2日, 수정완료일: 1999年12月3日

을 띄는지를 판단하기 위해 신 등록된 상표 원부를 찾아보거나 그와 유사한 사본을 유지 보관하여 그것을 수작업으로 찾아보는 형태로 처리하여 왔다<sup>13)</sup>. 새로 출원된 상표 중에서 약 25% 정도는 이미 출원된 상표와 비슷하다는 이유로 등록이 거절되므로 출원 이전에 또는 새로운 상표 제작 단계에서 이미 출원등록된 상표가 있는지를 알아볼 필요가 있다<sup>14)</sup>.

본 논문은 상표출원이 급증하고 있는 현 시대에서 이러한 많은 시간과 노력을 요하는 방법을 줄이고자 하는 시도로 기존의 부분매칭이 아닌 상표 특성을 고려한 근사매칭과 영상처리 기법을 결합한 가중치 부여 근사매칭 방법을 제안한다.

## II. 관련연구

### 1. 상표의 유형 및 유사성

상표의 유형분류는 STAR<sup>5)</sup>에서는 그림 1처럼 네 가지로 분류하고 있으며, 국내 특허청에서는 약 50만건의 상표에 대해 기호, 문자, 도형, 색채, 입체, 결합상표로 구분하기도 하지만 그 경계가 모호한 상표들도 많다.

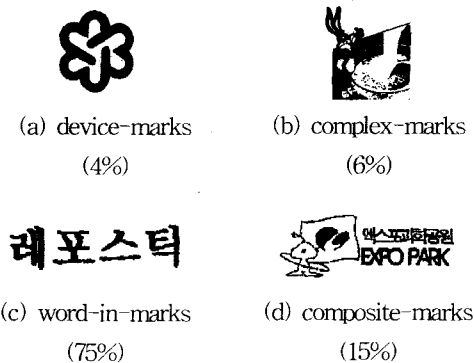


그림 1. 상표 유형 및 비율  
Fig. 1. Category and ratio of trademarks.

이들의 구성비율은 문자만으로 이루어진 상표(word-in-marks)가 75%, 문자가 포함된 상표(composite-marks)가 15%로 전체 상표의 90%가 문자관련 상표로 이루어져 있다<sup>6)</sup>. 지금까지는 평면적인 문자나 도형만을 사용했지만 추후 좀더 다양한 행태의 상표가 등장할 것이며 최근 국내에서도 색채상표와 입체상표를 인정하고 있고, 심지어는 소리, 냄새상표까지 영역이 넓어지고 있는 추세이다<sup>14)</sup>.

상표의 유사여부의 관찰 방법은 전체적, 객관적, 이격적(離隔的) 관찰을 원칙으로 하되 상표 구성 중 인상적인 부분에 대하여 비교하며 유사판단은 원칙적으로 상표의 칭호, 외관, 관념 중 어느 하나가 유사하여 거래상 상표출처의 오인, 혼동의 우려가 있는 상표는 유사한 것으로 본다<sup>7)</sup>. 이러한 사실들에 비추어 유사상표 검색 시스템이란 실제 상표법에 의한 모든 유사성 판단 기준을 적용할 수 없으며, 단순히 상표의 외관적 유사성이나 상표명에서 글자의 일치 정도와 기타 일부 기준만으로 유사도를 계산하게 된다.

출원상표가 아닌 상표 데이터베이스(DB)내 임의의 등록상표로 질의를 하는 경우라도 시스템은 기존 선등록된 상표DB에서 수십개 이상의 유사상표들을 출력할 수 있다. 이는 시스템이 상표DB 내에 유사상표가 존재하는 것으로 본다는 의미이다. 하지만 실제로 이들은 상표법상의 유사상표는 아니며 타업종이거나 같은 회사에서 출원한 경우에 해당된다. 따라서 시스템에서 언급하는 상표의 유사성은 상표법 상의 유사성과는 다를 수 있으며, 시스템은 출원상표에 대해 상표법상의 유사상표가 등록상표 내에 존재하는지를 심사관이 판단하기 위한 보조 수단으로서 사용되거나 신규 출원상표와 충돌 가능성이 있는 기존의 등록상표를 자동으로 검색하여 신규 출원자가 참조하는데 이용될 수 있다.

### 2. 기존 상표검색 시스템 및 문제점

상표검색 시스템의 질의 형태에 따른 구분은 크게 문자열과 영상 두 가지로 나뉘어지는데 90%를 차지하는 문자기반 상표의 경우 문자열로 질의하며 매칭 또한 문자열을 대상으로 한다. 반면 영상으로 질의하는 경우는 4%에 불과한 device-marks를 대상으로 하는 내용기반 검색 즉 영상 자체를 대상으로 한다.

IBM의 QBIC<sup>18)</sup>에서는 사진, 우표, 상표 검색이 가능한데 상표의 경우 1,000개의 상표를 대상으로 한 데모 시스템<sup>1)</sup>이 있다. 이 시스템은 하나의 원본 상표로부터 여러 모양(shape) 정보들을 추출하여 미리 정의된 Design Code들로 입력되며 검색시 하나라도 같은 디자인 코드를 갖는 상표들은 모두 무순위로 출력된다. 또한 문자열 검색방법으로 부분매칭도 함께 제공한다. 그 밖에 싱가포르의 STAR<sup>5)</sup>, ETL의 TRADEMARK System<sup>2)</sup>, CMC의 TRADEMARK REGISTRATION &

1) <http://www.qbic.almaden.ibm.com/tmdemo/>

2) <http://www.etl.go.jp:8080/etl/taiwa/TradeMark/>

표 1. 문자열 대상 매칭방법의 문제  
Table 1. Matching problems for strings.

문제점	비교 문자열				Sim
	상표1	문자열(Q)	상표2	문자열(S)	
① 의미적 으로 무관	대구병원	대구병원	구 병 산	구병산	2/5
	제안화장품	새한화장품	한화형광	한화형광	2/7
② 가중치 고려 안됨	주식회사 한컴 HANCOM INC	주식회사한컴	주식사정성	주식회사정성	6/10
			한컴홈	한컴홈	2/7
	주식회사우성산업	우성산업주식회사	우성산업	주식회사우성산업	12/14
			화성산업주식회사	화성산업주식회사	12/14
③ 글자형태 고려 안됨	한컴	한컴	한컴	한컴	2/2
			한컴 HANCOM	한컴	2/2

IMAGE RETRIEVAL SYSTEM” 등이 있으며 국내의 경우는 SEARCH ENGINE”, WIPS”, ZOOM” 등이 있다. 이 중 QBIC, WIPS, ZOOM은 문자열 검색 방법을 제공하지만 모두 순위부여가 불가능한 부분매칭만을 제공한다.

문자상표를 내용기반으로 검색하는 경우도 있지만 이는 부정확한 결과를 나타낸다. 글자 형태는 영상의 관점에서 서로 상당히 유사해 보이기 때문에 완전히 다른 글자로 된 상표라 할지라도 상대적으로 높은 유사도를 나타낼 수 있으며 반대로 같은 글자라도 크기, 굵꼴, 위치 등 여러 외형상의 요인에 따라 유사도에 큰 영향을 미쳐 전혀 다른 경우로 판단할 수 있다. 따라서 문자기반 상표는 문자열을 대상으로 하는 검색방법이 적합한데 이에는 정확매칭(Exact Matching), 부분매칭(Partial Matching), 근사매칭(Approximate Matching)이 있다. 원하는 특정상표를 찾는 것이 목적이 아니라 유사상표가 목적인 시스템에서 정확매칭은 부적합하다.

기존 상표검색에 사용하는 부분매칭 또한 정확매칭과 마찬가지로 매칭여부만을 나타내기 때문에 순위부여가 불가능하며 질의어가 길수록 매칭 가능성이 떨어지므로 2~3자의 질의어로 여러 번 질의해야만 원하는 유사상표들을 찾을 수 있다. 이는 질의자에 따라 검색 결과가 달라지는 문제가 발생할 수 있다. 실 예로 ‘주식회사대성’과 ‘대성’은 유사상표<sup>[7]</sup>이지만 ‘주식회사대성’으

로 질의시 ‘대성’을 검색하지 못하며 ‘피아스’로 질의시에도 ‘에스피아’, ‘아스파’, ‘텔피아’, ‘모피아’ 등을 검색하지 못한다.

이러한 문제점을 해결하기에는 근사매칭이 적합한데 해밍거리 기반 근사매칭은 다양한 길이를 갖는 상표명에 적용하기에는 부적합하다. 또한 질의 문자열을 비교 문자열로 변환하기 위해 입력, 삭제, 치환되는 기호의 수를 최소로 하는 k-근사매칭<sup>[6]</sup>이 있는데 이 방법은 알고리즘이 복잡하고 단어들의 위치가 서로 바뀐 상표는 유사도 매우 낮게되는 단점이 있다. 이는 n-gram 기반 근사매칭으로 해결할 수 있는데 두 글자씩 중첩되게 분절하면서 비교하는 방법은 간결하면서도 상표검색에 우수한 성능을 나타낸다<sup>[6]</sup>. 그러나 여전히 몇 가지 문제점이 발생하는데 표 1에 그러한 경우를 예시하였다. 이는 문자열을 대상으로 하는 기존 여러 부분매칭 및 근사매칭에서도 항상 발생하는 문제이다.

표 1의 첫 번째인 문제점 ①은 단어 사이의 무의미한 글자들(Q의 굵은 고딕체)로 인해 유사도가 높아지는 경우이며, 문제점 ②는 단어의 중요도를 고려하지 않고 일치된 글자 수로만 판단함으로써 발생하는 문제이고, ③은 같은 문자열이라도 외형을 고려하지 않아 발생하는 문제이다. 따라서 기존의 단순 기계적 매칭 방법을 그대로 상표검색에 도입할 경우 보다 정확한 순위부여를 할 수 없다.

### III. 상표특성을 고려한 근사매칭

제안하는 매칭방법은 상표의 특성을 고려한 근사매

3) <http://cmchome.cmc.stph.net/cmc/ip/trims/trims.htm>  
 4) <http://cherry.kist.re.kr/~kij/query.html>  
 5) <http://www.wips.co.kr/>  
 6) <http://www.zoomtech.co.kr/>

표 2. 문자기반 상표의 단어 구성별 유형분류

Table 2. Classification of character-based trademarks by the combination of words.

R/U	문법	예제	문자열	설명
(R)	<부사>		빙그레	등록어만으로 구성
	<Ng>+<Ng>		산장+녹차	
(U)	<Np>		새모스	미등록어로 구성
	<외래어>		폰타나	
(R, U)	<Np>+<Ng>		창덕+전기	등록어와 미등록어로 구성
	<Ng>+<Np>		원조+신승	

R : 어휘사전에 등록된 단어      예) 일반명사(<Ng>), 접두사, 접미사, 형용사, ...  
 U : 어휘사전에 등록되지 않은 단어      예) 고유명사(<Np>), 외국어, 의성어, 합성어, ...

표 3. 문자기반 상표의 모양별 유형분류

Table 3. Classification of character-based trademarks by shape.

줄																			
글꼴	산내들	산내들	누룽지	이탈이탈의	크린퍼그츠	코끼리	물價販	ALLIN	전남항공	한양	산내들	대양탱크	현호총명탕	현대건설	농우유	농우유	농우유	농우유	농우유
속성	포도랑	포도랑	누룽지	이탈이탈의	크린퍼그츠	코끼리	물價販	ALLIN	전남항공	한양	산내들	대양탱크	현호총명탕	현대건설	농우유	농우유	농우유	농우유	농우유
언어	삼성	삼성	호	BIOCAN	511	반다이	훈컴	농무로	PC출력	200대뉴스									
크기	한양	산내들	대양탱크	현호총명탕	현대건설	농우유	농우유	농우유	농우유	농우유	농우유	농우유	농우유	농우유	농우유	농우유	농우유	농우유	농우유
비율																			
위치																			
그림																			

칭으로 문자열을 대상으로 하는 가중치 부여 근사매칭과 상표 영상에 대한 내용기반 접근 방법을 결합한 방법이다. 1절에서는 적합한 상표검색 알고리즘을 찾는 데 단서가 되는 상표의 유형분류를 시도하며, 2절에서는 문자열 근사매칭을 위해 단어와 가중치와의 관계를 고려해 표 1의 문제 ①, ②를 해결한다. 3절에서는 유사도가 같은 상표 그룹단위로 영상 처리 기법을 적용하여 표 1의 문제 ③을 해결하고 4절에서는 두 방법을 결합한다.

1. 문자기반 상표의 유형분류

상표의 유사성을 계산하기 위한 방법을 찾기에 앞서 문자기반 상표의 유형을 두 가지 관점에서 분류하였다. 첫째는 단어 구성별 분류로 표 2에 예시하였고 둘째는

모양별 분류로 표 3에 예시하였다. 이들은 표 1의 문제를 해결하기 위한 단서를 제공한다.

표 2는 상표명을 이루는 글자 내용 즉, 단어의 통사적 특성에 따라 분류한 것으로 단어가 일반 어휘사전에 존재하는 경우와 그렇지 않은 경우가 있다. 여러 단어로 구성되어 있는 상표의 통사 정보를 형태소 분석에 이용할 수도 있지만 이는 본 논문의 범위를 벗어날 뿐더러 적용 범위가 매우 제한적이다. 따라서 이러한 통사 또는 의미 정보를 단어와 가중치와의 관계로 대신하여 표 1의 문제 ①, ②를 해결한다. 이에 대한 자세한 내용은 2절에서 언급한다.

표 3의 ①은 word-in-marks를 다섯 종류로, ②는 composite-marks를 세 종류로 세분화한 것이다. '줄'은 가로 한두 줄인 경우가 많으며 '글꼴'은 아직까지 명조

체나 고딕체가 주류지만 도형화된 글자체도 점차 증가하고 있다. '비율'은 글자와 도형이 차지하는 공간 비율을 말하는 것이며 '위치'는 글자와 도형이 차지하고 있는 공간관계에 따른 것으로 글자의 좌측 또는 위쪽에 위치한 경우가 많았다. '그림'은 복잡한 도형이나 자연 영상에 가까운 그림이 들어간 상표로, 문자가 차지하는 공간 비율은 작다. 이러한 정보들은 문자기반 상표의 외향적 특성을 파악하는데 도움을 주며 이는 표 1의 문제 ③을 해결하기 위해 어떠한 영상처리 기법을 적용해야 할지를 결정하는 단서를 제공한다. 보다 자세한 결정 과정 및 적용 기법은 3절에서 언급한다.

2. 문자기반 상표의 문자열 근사매칭

부분매칭과 달리 근사매칭은 여러 단어를 한 번에 질의하여 검색할 수 있지만 공백문자 없이 입력하여 매칭하게 되면 표 1의 문제 ①이 발생할 수 있다. 이는 질의 문자열을 단어 단위로 분리하기 위한 형태소 분석과 함께 수십만 어휘가 등록된 전자사전이 요구됨을 뜻한다. 또한 표 2에서 보듯이 다양한 형태의 미등록어가 존재하는 상표의 특성상 오분석 가능성이 높아진다. 검색엔진의 경우 두 개 이상의 키워드를 동시에 줄 때, 논리 연산자, 공백문자, 또는 쉼표가 오는데, 본 시스템에서는 공백문자만 오며 이를 단어 구분자(delimiter)로 삼았다. 질의한 단어들이 많이 포함될수록 또 포함되지 않더라도 일부 일치되는 글자가 많을수록 유사도는 높아진다.

표 1의 문제 ①, ②를 해결하기 위해서는 상표명을 구성하는 각 글자들이 어떠한 단어에 속하는지 또 그 단어가 상표명에서 어떠한 역할을 하는지 알아야 한다. 이를 위해 질의어를 그래프 단위로 분절하여 비교하되 문제 ①의 경우 특정 그래프가 어떤 단어내의 그래프인지 아니면 인접한 두 단어간의 그래프인지를 파악할 수 있어야 하며 문제 ②의 경우는 각 그래프들이 단어별로 서로 다른 가중치를 가질 수 있어야 한다.

그래프는 여러 형태로 나타낼 수 있는데 본 논문에서 언급하는 그래프에 대한 정의는 식 (1)과 같다. 여기서 G는 임의의 문자열을 분절한 그래프들을 원소로 하는 집합이다. 임의의 문자열을 S라 하면 이에 대한 그래프는 소문자 s로 나타낸다고 할 때 각 그래프 s는 p1 ~ p4의 특성을 갖는다.

$$G(S) = \{s_1, s_2, s_3, \dots, s_n\} \tag{1}$$

$$= \{s_i \mid p1(s) \wedge p2(s) \wedge p3(s) \wedge p4(s)\}$$

조건명제

$$p1(s) : s_i \subset G$$

$$p2(s) : s_i \text{의 문자 수} = 2(\text{그래프 크기})$$

$$p3(s) : s_i \cap s_{i+1} \text{의 문자 수} = 1$$

$$p4(s) : s_i = (s_{i-1} \cap s_i) + (s_i \cap s_{i+1})$$

제안한 근사매칭에 적용하기 위해 여러 단어로 이루어진 질의어를 대상으로 하여 식 (1)을 다시 살펴보고자 한다. 문자열 S는 k개의 단어로 이루어져 있으며 S내의 i번째 단어 W<sub>i</sub>는 m개의 그래프로 분절된다고 가정할 때 S와 W<sub>i</sub>에 대한 그래프의 집합 G(S), G(W<sub>i</sub>) 및 G(S)의 원소 수 n<sub>S</sub>는 식 (2)와 같이 나타낼 수 있으며 그 예를 함께 예시한다.

$$S = W_1 W_2 \dots W_k$$

$$G(W_i) = \{w_{i1}, w_{i2}, \dots, w_{im}\}$$

$$G(S) = \{s_1, s_2, s_3, \dots, s_n\} \tag{2}$$

$$= \{w_{11}, w_{12}, \dots, w_{12}, w_{21}, \dots, w_{23}, \dots, w_{im}, w_{ij}, w_{j1}, w_{j2}, \dots\}$$

$$n_S = n(G(S)) = \sum_{i=1}^k n(G(W_i)) + k - 1$$

S = "한국주식회사" W<sub>1</sub> = "한국", W<sub>2</sub> = "주식회사",  
 k = 2 G(W<sub>1</sub>) = {한국}, G(W<sub>2</sub>) = {주식, 식회, 회사}  
 G(S) = {한국, 국주, 주식, 식회, 회사}  
 s<sub>1</sub> = w<sub>11</sub>, s<sub>2</sub> = w<sub>12</sub>, s<sub>3</sub> = w<sub>21</sub>, s<sub>4</sub> = w<sub>22</sub>, s<sub>5</sub> = w<sub>23</sub>  
 n<sub>S</sub> = n(G(W<sub>1</sub>)) + n(G(W<sub>2</sub>)) + 1("국주") = 5

G의 원소 즉, 그래프들 s<sub>i</sub>는 단어내 그래프 s<sub>w</sub>와 단어간 그래프 s<sub>c</sub> 중의 하나로 나뉘어지는데 이들의 집합 G<sub>w</sub>와 G<sub>c</sub>를 식 (3)과 같이 정의한다. 앞서의 예 G에서 G<sub>c</sub>는 {국주}가 되며 G<sub>w</sub>는 G에서 G<sub>c</sub>를 뺀 집합이다.

$$G = G_w \cup G_c, \quad G_w \cap G_c = \emptyset$$

$$G_w = G_{w1} \cup G_{w2} \cup \dots \cup G_{wk} = \{s \mid p(s)\}$$

$$p(s) : s = w_{ij}, \quad w_{ij} \in G(W_i), \quad 1 \leq i \leq k, \quad 1 \leq j \leq m \tag{3}$$

$$G_c = \{s \mid q(s)\}$$

$$q(s) : s = wij, \quad wij \notin G(W_i), \quad wij \notin G(W_j),$$

$$wij = (w_{im} \cap w_{ij}) + (w_{ij} \cap w_{j1}),$$

$$1 \leq i \leq k-1, \quad 2 \leq j \leq k, \quad j = i+1$$

이제 s<sub>w</sub>와 s<sub>c</sub>를 구분하여 가중치(ω)를 주며 s<sub>w</sub>는 단어별로 가중치를 부여한다. 이때 그래프와 가중치를 쌍(s<sub>i</sub>, ω<sub>i</sub>)으로 하는 집합 G<sub>w</sub> 및 가중치 ω는 식 (4)와

같이 정의한다.

$$G_w = \{(w_{11}, \alpha_1), (w_{12}, \alpha_1), \dots, (w_{12}, \beta), (w_{21}, \alpha_2), \dots, (w_{km}, \alpha_k)\}$$

$$= \{(s_i, \omega_i) \mid s_i \in G, \omega_i = \omega(s_i), s_i = s_w \text{ or } s_c, s_i = s_c \rightarrow s_{i+1} = s_w\}$$

$$\omega = \{\omega_1, \omega_2, \omega_3, \dots, \omega_n\}$$

$$= \{\alpha_1, \alpha_1, \dots, \alpha_1, \beta, \alpha_2, \alpha_2, \dots, \beta, \alpha_i, \dots, \alpha_i\}$$

where  $\omega(s_w) = \alpha_i, \omega(s_c) = \beta, 0 < \beta < \alpha_i$  (4)

질의에 사용된 모든 단어의 가중치를 서로 다르게 주는 것은 비효율적이므로, 많이 사용되는 평이하면서도 고빈도 단어나 상표심사기준에 언급되어 있는 식별력이 떨어지는 단어에 대해서만 낮은 가중치를 주는 방법을 선택하였다.

상표DB 내에서 빈번히 나타나는 고빈도 단어가 질의어에 포함된 경우, 검색된 상표명들 상당수가 모두 같은 고빈도 단어를 포함하고 있으므로 이들의 유사도는 고빈도 단어보다는 질의어의 나머지 단어와의 유사성에 더 큰 비중을 두는 것이 타당하다. 또한 보통명칭을 포함 표 1의 ②와 같이 식별력이 떨어지는 단어는 유사성 비교에서 상표의 주요부분이 될 수 없다<sup>[7]</sup>. 특히 고빈도 단어나 식별력이 없는 단어이면서 글자 수가 많은 경우, 검색된 상표명들의 유사도는 이들에 의해 더 큰 영향을 받았기 때문에 질의어의 나머지 단어는 상대적으로 별 영향을 미치지 못한 상태가 된다. 이러한 이유로 가중치를 낮게 주는 것이며 이후 본 논문에서는 이러한 단어들을 모두 고빈도 단어라 칭한다.

이로서 각 그래프의 종류 및 가중치는 최종적으로 식 (5)과 같이 정의한다.  $\beta$ 를 작게 할수록 개별 글자 단위가 아닌 단어의 일치에 비중을 두는 것이며,  $\alpha_f$ 가 작을수록 상표에서 단어의 의미 역할을 중요시하는 것이다.

$$s = \begin{cases} s_w & : \text{단어내 그래프} \\ s_{wf} & : \text{고빈도 } s_w \\ s_c & : \text{단어간 그래프} \end{cases} \quad (5)$$

$$\omega(s) = \begin{cases} \alpha & \text{if } s = s_w \\ \alpha_f & \text{if } s = s_{wf} \\ \beta & \text{if } s = s_c \end{cases} \quad \text{where } 0 < \beta < \alpha_f < \alpha = 1$$

이 때 그래프의 집합  $G$ 는 식 (3)에서 식 (6)으로 수정되며  $p(s), q(s)$ 는 식 (3)의 경우와 같다. 앞서의 예  $G$ 에 대한  $G_w$ 는 {한국},  $G_{wf}$ 는 {주식, 식회, 회사},

$G_c$ 는 {국주}가 된다.

$$G = G_w \cup G_{wf} \cup G_c, \quad G_w \cap G_{wf} \cap G_c = \emptyset$$

$$G_w = \{s \mid p(s), W_i : \text{고빈도아님}\}$$

$$G_{wf} = \{s \mid p(s), W_i : \text{고빈도}\}$$

$$G_c = \{s \mid q(s)\} \quad (6)$$

최종적으로 질의 문자열  $Q$ 와 비교 문자열  $S$ 와의 유사도(Sim)는 식 (7)로 정의한다. 이는 유사계수 공식 중의 하나인 Dice 계수<sup>[18]</sup>를 기초로 변형한 것이다.

$$Sim(Q, S) = \frac{2 \times \sum_{i=1}^{n_q} \{ \omega_i \mid q_i \subset S \}}{\sum_{i=1}^{n_q} \omega_i + n_s} \quad (7)$$

where  $n_q = n(G(Q)), n_s = n(G(S))$ :  $Q, S$ 의 그래프 수

그래프 수 =  $\begin{pmatrix} l-1 & \text{if } l > 2 \\ 1 & \text{else} \end{pmatrix}$   $l$ : 문자열의 글자 수

$G(Q)$ :  $Q$ 에 대한 그래프의 집합으로 식(6)을 만족

$G(S)$ :  $S$ 에 대한 그래프의 집합으로 식 (1)을 만족

$q_i$ :  $Q$ 에 대한  $i$ 번째 그래프

$\omega_i$ :  $q_i$ 의 가중치

질의어를 어떻게 분절하여 비교하는 지는 근사매칭의 성능에 큰 영향을 미치는데, 두 글자 단위로 중첩하여 분절하여 비교하는 기존 방법( $M_g$ )에서, 단어 단위로 중첩하지 않고 분절하여 비교하는 방법( $M_w$ ),  $M_g$ 에서 단어간 그래프는 비교하지 않는 방법( $M_c$ ), 그리고  $M_g$ 에 가중치를 부여한 경우( $M_m$ )가 있다. 이에 대한 분절 예를 표 4에 나타내었으며  $M_w$ 와  $M_c$ 는 본 논문에서 제안한  $M_m$ 의 중간 단계라 할 수 있다.

### 3. 문자기반 상표의 영상 처리

영상 처리에 우선하여 문자상표를 처리하게 되면 그림 2에서 보듯이 넷째 줄을 제외하고는 전혀 다른 상표명임에도 유사성이 큰 것으로 계산되는 경우가 발생한다. 이들은 영상 처리의 관점에서는 유사할지 모르나 상표법이나 일반적 관점에서는 전혀 무관한 상표들이다.

창호가 같지만 외관이 다른 문자기반 상표가 그림 3에 예시되어 있는데 이는 같은 회사에서 등록한 경우이거나 타업종 간의 상표이다. 전체 상표DB를 대상으로 그림 2의 결과를 얻기보다는 그림 3처럼 글자가 같

표 4. 문자열의 분절 예

Table 4. Examples of divided strings.

방법	Q(결외어)	Q'(부분 문자열)	비 고
$M_g$	"가나다라마"	{가나, 나다, 다라, 라마}	기존 방법
$M_w$	"가나다 라마"	{가나다, 라마}	'다라'가 생성 안됨
$M_c$	"가나다 라마"	{가나, 나다, <u>다라</u> , 라마}	'다라'는 비교시 제외
$M_m$	"가나다 라마"	{가나, 나다, <u>다라</u> , 라마}	식 (5)에 의한 가중치



그림 2. 영상처리에 의한 유사상표

Fig. 2. Similar trademarks by the image's point of view.



그림 3. 칭호가 같지만 모양이 같지 않은 상표

Fig. 3. Trademarks with the same name but not by shape.

거나 거의 유사한 경우를 대상으로 영상처리를 행하는 것이 타당하다. 이는 문자열 근사매칭에 의해 검색된 같은 유사도를 갖는 그룹 단위의 내용기반 검색을 뜻한다.

영상처리 기법 중 도형상표의 경우는 주로 모양으로 유사도를 비교<sup>[10-13]</sup>하는데 문자기반 상표의 경우는 글자 모양의 특성을 반영할 수 있는 기법이 필요하다. 한글은 획으로 이루어져 있는데 자음 일부를 제외하고는 직선이 대부분이므로 이에 민감하게 반응하도록 프로잭션(Projection) 기법을 사용한다. 이러한 기법을 이용한 예로 전표영상에서 문자열 찾기<sup>[14]</sup>, 인쇄된 한글의 글꼴 인식<sup>[15]</sup> 등이 있다.

가로  $W$ , 세로  $H$ 인 이진영상으로 된 상표 영상을  $B[i, j]$ 라 할 때, 수직·수평방향의 프로젝션을 각각

$P_v[j], P_h[i]$ 라 하면 식 (8)과 같이 구할 수 있다<sup>[16]</sup>.

$$P_v[j] = \sum_{i=1}^H B[i, j] \quad \text{where } j=1, 2, \dots, W$$

$$P_h[i] = \sum_{j=1}^W B[i, j] \quad \text{where } i=1, 2, \dots, H$$
(8)

이렇게 하여 구해진  $P_v, P_h$ 를 프로파일(profile)이라고 하는데 이를 이용하여 영상의 특성을 파악할 수 있다. 그림 4에 프로파일에 의한 영상 예를 보인다.

상표 영상으로부터 문자 영상만을 추출하거나 어떤 종류의 글꼴인지 또 몇 줄로 구성되어 있는지 등의 문자열에 대한 특성을 파악하는 것은 매우 힘들다. 그림 4(d)나 표 3의 ②를 보면 글자와 도형의 경계가 모호하거나 글자의 위치가 다양한 형태로 존재하기 때문이다.

표 5. 같은 유사도 그룹에서의 2차 유사도에 의한 출력 결과 예

Table 5. Output examples by the 2nd similarity distance in the same group.

질의	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
하나로	하나로	하나로	HANARO 하나로	하나로	하나로	하나로	하나로	하나로	하나로	하나로
한컴프레스	한컴프레스	한컴프레스	한컴프레스	한컴프레스	한컴프레스	한컴프레스	한컴프레스	한컴프레스	한컴프레스	한컴

따라서 문자영상의 특성과 도형도 함께 비교할 수 있도록 분리하지 않고 두 상표의 프로파일로부터 거리를 구해 2차 유사도 계산에 이용한다.



그림 4.  $P_v, P_h$ 에 의한 영상 예  
Fig. 4. Example images by  $P_v, P_h$ .

그림 2의 셋째 줄같이 극히 일부의 세로줄로 구성된 상표를 제외하고는 프로파일은  $P_h$ 는 그림 4에서 보듯이 유용성이 떨어지므로  $P_v$ 만을 이용한다. 두 프로파일  $P1_v$ 와  $P2_v$ 에 대한 거리는 두 점의 거리 city block distance<sup>[16]</sup>를 이용하여 정규화한 식 (9)를 사용한다.

$$d(P1_v, P2_v) = \left( \sum_{n=1}^W \frac{|P1_v[n] - P2_v[n]|}{H} \right) / W \tag{9}$$

where  $W, H$ : width, height of image

4. 2단계 유사도 계산

적용하고자 하는 여러 방법들을 결합하는 것은 여러 가지가 있을 수 있는데, 두 방법 를 결합한 방법을  $M$ 이라 하고 대상  $DB$ 를  $D$ 라고 하면 일반적으로 식 (10a)나 (10b)와 같은 방법을 사용한다.

$$M = \omega M_1(D) + (1 - \omega)M_2(D), \quad 0 \leq \omega \leq 1 \tag{10a}$$

$$M = M_2(M_1(D)) \tag{10b}$$

$$M = \sum_{i=1}^R M_2(D_i), \quad \text{where } M_1(D) = \sum_{i=1}^R D_i \tag{10c}$$

$D_i$ : 유사도가 같은  $i$ 번째 그룹

$M_1$ 이 문자열 근사매칭이고  $M_2$ 가 영상처리일 때 식 (10a)는 그림 2와 같은 결과를 초래하며 식 (10b)의 경

우  $M_1$ 에 의해 걸러진(filtering) 것 즉, 유사도가 있는 것을 대상으로  $M_2$ 를 적용하기 때문에 전자에 비해 그 비율이 상당히 줄어들긴 하지만 역시 그림 2의 결과가 발생한다.

따라서 문자열 근사매칭으로 순위를 부여한 후 각 순위별로 2차 유사도를 평가하는 식 (10c)에 의한 방법을 사용한다. 이 방법은 그룹간 상대적 우선순위가 결정되기 때문에 식 (10a)나 식 (10b)에서와 달리  $D_i$ 에서 2차 유사도 계산으로 가장 순위가 낮은 경우라도  $D_{i-1}$ 의 가장 높은 순위보다 절대적 우위에 있다.

같은 1차 유사도를 갖는 그룹내 상표들 간에는 표 5와 같은 두 가지 경우가 존재하는데 첫 번째 줄은 그림 3과 같은 상표가 존재할 때이며 두 번째 줄은 문자열 길이가 서로 다르면서도 같은 유사도인 경우이다. 식 (7)에 의한 1차 유사도가 각각 1과 0.4인 경우로 같은 순위를 갖지만 식 (9)에 의한 2차 유사도 계산으로 이들은 서로 다른 순위가 부여되어 표 5와 같이 출력 된다.

표 5에서 알 수 있듯이 첫 번째 경우를 영상처리한 것은 의미가 있으나 이와 같은 경우 즉, 그림 3에 해당하는 비율이 매우 적었다. 두 번째의 경우는 질의 문자열과 비교 문자열이 다르므로 외형보다는 글자에 우선하게 되어 영상처리가 큰 의미를 갖지 않음을 알 수 있는데, 출력결과의 첫 번째인 ‘쌍용익스프레스’는 외형상으로 질의상표 ‘한컴프레스’의 영상 폭과 비슷하지만 글자 내용상 마지막의 ‘한컴’이 더 유사하기 때문이다.

문자기반 상표의 영상처리는 그림 3과 같은 상표 및 도형화된 문자상표가 증가할수록 좀 더 큰 비중을 차지하겠지만 현재로서는 그 비중이 상대적으로 낮기 때문에 두 방법의 결합형태로 식 (10c)를 선택하였으며, 시스템 구축은 문자열로 질의하면 문자열 근사매칭만으로 검색이 이루어지며 상표 영상으로 질의하면 식 (10c)의 결합된 방법으로 검색이 이루어지도록 하였다.



#### IV. 실험결과

##### 1. 실험방법

실험에 사용할 상표DB는 국내에 등록된 상표DB로부터 문자기반 상표 14만여 개를 임의로 발췌하여 실험하였다.

1차 유사도 계산에 의한 기존 방법과의 비교를 위해 표 6의 상표들을 질의어로 선정하였다. 단일어로 된 상표는 기존방법과 차이가 나지 않으므로 제외하고, 고빈도 단어와 두 단어에서 네 단어까지 선정하였다. 상표 영상으로 질의하는 경우는 상관없지만 문자열로 질의시 다섯 단어 이상을 하나의 문자열로 질의하는 경우는 드물기 때문이다. 표 6의 '2차 유사도' 항목은 영상처리를 위해 그림 3에 해당하는 상표들 중에서 선정하였다.

이러한 질의를 대상으로 표 4의  $M_g, M_c, M_m$  세 가지 방법을 각각 사용하여 1차 유사도에 의한 검색 결과 순위를 비교하였다.  $M_w$ 는 DB에 등록된 모든 상표명의 띄어쓰기가 객관적으로 정확하게 되어있어야 하는 부담이 있을뿐더러 외형이 완전히 무시되어 상표검색에 부적합하므로 비교 대상에서 제외하였다. 이 때 식 (5)에 따른 가중치  $\alpha, \alpha_f, \beta$ 는 각각 1.0, 0.8, 0.5로 설정하였다.  $\alpha_f$ 와  $\beta$ 를 작게 할수록 외형보다는 글자의 의미를 중요시하는 것이며  $\beta=0, \alpha_f=\alpha=1$ 로 하면  $M_c$ 와 같이,  $\beta=\alpha_f=\alpha=1$ 로 하면  $M_g$ 와 같게 된다.

실험 결과 같은 순위를 가지는 그룹의 원소 수가 수 개에서 100개 이상까지 매우 불규칙하게 나타났는데 유사상표들을 모두 포함시키기 위해서는 유사도가 0을 초과하는 모든 것을 대상으로 조사하여야 하지만 그 수가 수천 개 이상이며 유사도가 낮을수록 유사상표가

존재할 확률이 낮아지므로 1차 유사도 계산에 의한 순위 그룹 10개 내에 모두 존재하는 것으로 가정하였다.  $M_g, M_c$ 의 경우는 2차 유사도 계산을 하지 않으므로 동일 유사도 그룹별로 맨 앞쪽에 위치한 임의의 상표를 선정하였으나  $M_m$ 의 경우는 영상처리에 의한 2차 유사도 계산후 그룹별로 가장 높은 순위의 상표를 선정하였다. 이 밖에 각 방법의 순위차가 10을 초과하면서 유사도가 0.5이상인 것들도 선정하였다.

이로서 한 번의 질의로 최소 10개 이상의 상표들이 자동 추출되도록 하였는데 이를 수작업으로 적합상표들을 선택한 후 각 방법의 순위에 따라 출력하였다. 이때 적합상표들의 개수는 각 방법 모두 같지만 그들 사이에 부적합 상표들이 얼마나 포함되어 있느냐에 따라 정확도에서 차이가 난다.

1차 유사도에 의한 순위를 부여한 결과,  $M_m$ 에서 2차 유사도 계산에 의한 효과를 얻을 수 있는 그림 3과 같은 상표의 출현은 매우 적었다. 따라서 2차 유사도 계산에 의한 효과를 측정하기 위해 단일어로 구성된 표 6의 '2차 유사도' 항목에 해당하는 칭호를 갖는 상표들로만 따로 실험하였다.

##### 2. 평가척도

검색 성능을 평가하기 위한 척도로는 여러 가지가 있으며 상황에 따라 적절한 방법을 선택하게 되는데 일반적으로 재현율(RR: recall rate)과 정확도(PR: precision ratio)<sup>[17]</sup> 또는 Bull's Eye Performance(BEP)와 평균순위(AR: average rank)<sup>[18]</sup>등을 사용한다. 표의 '2차 유사도' 항목의 2차 유사도 계산 결과는  $RR=1, PR=1$ 이므로 그림 3의 인지적 순위와 표 5의 계산에 의해 출력된 순위와의 오차를 city block distance<sup>[16]</sup>로 평가하였다. 이들에 대한 식은 (11)~(14)이며 오차와 AR은 수치가 작을수록, RR, PR, BEP는 높을수록 좋은

표 6. 성능평가를 위한 질의 상표

Table 6. Examples of query trademarks for performance evaluation.

	고빈도	두 단어	세 단어	네 단어	2차 유사도
①	도서출판 한컴	계측기 정보	독도 해물 구이	월간 화학 공정 기술	하나로
②	대영 공업사	효원 조명	미원 다도 해물	신일 진공 청소기 싹쓸이	나드리
③	슈슈 시스템	싱싱 그린	에경 미스 바디	고려 수지 요법 혈도	산내들
④	녹수 산업	효도 농산	초등 전진 수학	농심 배추 김치 라면	참술
⑤	한주 식품	산동 원예	사조 인삼 광어	해태 갈아 만든 사과	홍길동

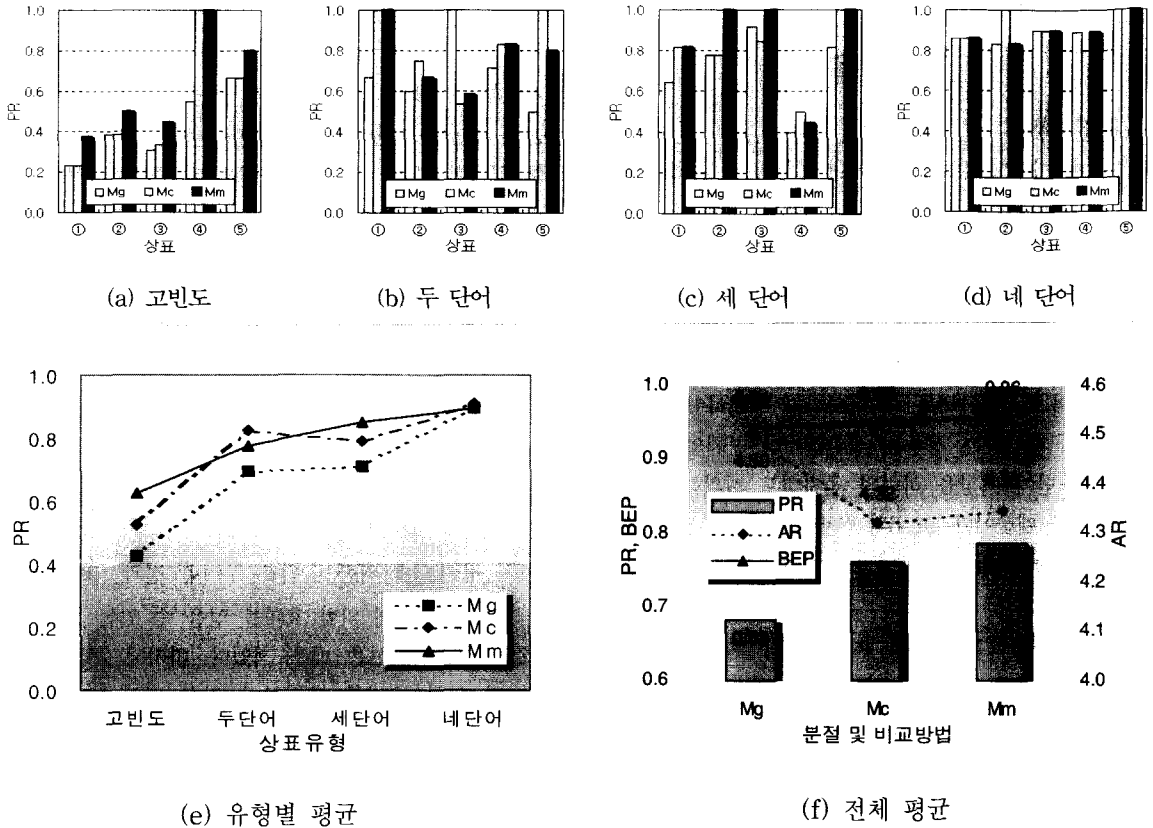


그림 5. 세 방법의 검색성능 평가  
 Fig. 5. Performance evaluation of 3 methods.

성능을 나타낸다.

$$RR = \frac{R_s}{R} \quad \text{where } R_s : \text{검색된 적합상표 개수}$$

$$PR = \frac{R_s}{S} \quad \text{where } R : \text{DB내 적합상표 개수}$$

$$S : \text{검색된 상표 개수}$$

$$\text{오차} = \left( \sum_{i=1}^n |R_{ih} - R_{ic}| \right) / n \quad (12)$$

where  $R_{ih}$  : 상표  $i$ 의 인지적 순위  
 $R_{ic}$  : 상표  $i$ 의 출력된 순위

$$BEP(\%) = \frac{R_{S_{2R}}}{R} \times 100 \quad (13)$$

where  $R_{S_{2R}}$  : 2R 순위내 적합상표의 개수  
 $R$  : DB내 적합상표 개수

$$AR = \sum_{i=1}^S R_i / S \quad (14)$$

where  $R_i$  : 검색된  $i$ 번째 상표의 순위  
 $S$  : 검색된 상표의 개수

본 논문에서는 표 1의 문제 해결이 목적이므로 PR에 의한 평가가 의미가 있으나 BEP와 AR도 참고적으로 함께 측정하였다. PR은 범위를 마지막 적합상표로 하여  $RR = 1$ 인 상태에서 계산하였다.

3. 실험 결과

그림 5는 세 방법에 대한 검색 성능을 평가한 것으로 PR에 대한 단어 유형별(a~d)과 그 평균(e), (f)는 질의 상표 20개 전체에 대한 평균을 나타낸 것으로 AR과 BEP(100을 곱하지 않았음)도 함께 표시하였다. (f)에서 보듯이 기존  $M_g$ 에 비해 제한한  $M_m$ 의 경우 정확도가 16.2% 향상되었으며 부차적으로 AR과 BEP도 약간 상승하였다.

유사도 비교시 상표명의 의미에 더 큰 비중을 두려면  $\alpha_f$ 와  $\beta$ 를 보다 낮게 주었으면 되었을 것이나 상표에서는 외형 또한 무시할 수 없으며, 의미와 외형 어느 쪽에 더 큰 비중을 두어야할지는 주관적인 문제이다.

그림 5의 제안한 방법  $M_m$ 에는 2차 유사도 계산을 위한 영상처리 기법이 포함되어 있는데 그에 대한 결과는 그림 6에 나와있다. 표 6의 '2차 유사도' 항목에 해당하는 5개 상표에 대한 2차 유사도 계산 결과, 식 (12)에 의한 평가로 각 상표의 순위 오차를 나타낸 것으로 이 중 상표 ②는 인지적 순위와 가장 큰 차이를 보이나 상표 ⑤는 인지적 순위와 동일한 결과를 산출하였다.

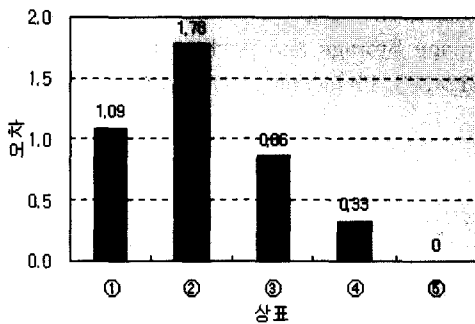


그림 6.  $M_m$ 의 2차 유사도  
Fig. 6. The 2nd similarity measure of  $M_m$ .

실험결과 그림 3에 해당하는 상표가 전체 등록상표에 비해 매우 적었기 때문에 2차 유사도 계산 효과는 전체 검색 성능에 비해 상대적으로 저조한 편으로 나타났다. 추후 실험 상표 수가 훨씬 더 많아지고 도형화된 문자상표가 증가하게 되면 문자기반 상표에서도 영상처리 기법이 점차 비중을 차지할 것으로 보인다.

### V. 결론

본 논문에서는 문자기반 유사상표 검색을 위한 가중치 부여 근사매칭을 제안하였다. 기존 부분매칭이나 근사매칭이 글자의 의미를 고려하지 않아 부정확한 결과를 초래하는 문제점을 해결하였으며 또한 상표의 특성을 고려한 방법이다. 전체 상표의 90%를 차지하는 문자기반 상표는 영상 처리 기법만으로는 한계가 있으므로 문자열 근사매칭에 우선하여 접근하였으며 프로젝트에 의한 영상처리 기법을 함께 결합하였다.

14만여 상표를 대상으로 실험하였으며 재현율을 1로 하여 부적합한 상표가 많이 포함되도록 한 상태에서 정확도를 측정된 결과 기존 방법에 비해 16.2%가 높아졌다. 이로서 제안한 방법은 기존 근사매칭의 문제점을

해결하여 정확도를 높일 수 있음이 입증되었으며 질의 상표의 단어간 그래프가 포함된 상표가 DB에 많이 존재할수록 우수한 성능을 나타낸다. 실험결과 2차 유사도 계산을 위한 영상처리 기법은 문자열 처리 기법에 비해 유용성이 떨어짐을 알 수 있었으나 보다 더 정확한 순위를 부여하고 추후 도형화된 문자상표의 증가에 대비하는데 의미를 부여할 수 있을 것이다.

제안한 방법은 PC상에서 Visual C++로 개발하여 기존 도형상표 검색 시스템과 결합하였으며, 웹상에서 검색이 가능한 데모 시스템<sup>1)</sup>은 약 3만여 상표를 대상으로 한다.

### 참고 문헌

- [1] "특허와상표", 대한변리사회, 제463호, 간별지, 1998.11.20
- [2] <http://www.kipo.go.kr/>, 국내 특허청 홈페이지 news
- [3] 특허업무전산처리운영지침서(도형상표검색), 특허청
- [4] <http://www.zoomtech.co.kr/>, ZOOM technology 홈페이지 상표이야기
- [5] C.P.Lam, J.K.Wu, B.Mehtre, "STAR - a System for Trademark Archival and Retrieval," *Proceedings 2nd ACCV'95*, vol. 3, pp. 214-217, 1995.
- [6] 서창덕, 김희율, "공유이중그래프를 이용한 유사상표 검색", 제11회 영상처리 및 이해에 관한 워크샵 발표 논문집, pp. 334-339, 1999.2
- [7] 의장·상표심사기준, 특허청, 1999.8
- [8] IBM Almaden Research Center, "Query by Image and Video Content: The QBIC System," *IEEE Computer*, pp. 23-31, Sep. 1995.
- [9] Andrew Binstock, John Rex, *Practical Algorithms for Programmers*, Addison-Wesley, 1995.
- [10] Y-S Kim, W-Y Kim, "Content-Based Trademark Retrieval System Using a Visually Salient Feature," *Journal of Image and Vision Computing*, vol. 16, pp. 931-939, Aug. 1998.
- [11] A. Vailaya, Yu Zhong and A. K. Jain, "Shape-based retrieval: A case study with

1) <http://trademark.hanyang.ac.kr>

- trademark image databases," *Proc. of IEEE Int. Conf. on Pattern Recognition*, no. 9, vol. 31, pp. 1369-1390, 1998.
- [12] J. P. Eakins, "Retrieval of trademark images by shape feature," *Proc. of ELVIRA Conf.*, May 1994.
- [13] G. Cortelazzo, G. A. Mian, G. Vezzi and P. Zamperoni, "Trademark shapes description by string-matching techniques," *Pattern Recognition*, vol. 27, no. 8, pp. 1005-1018, 1994.
- [14] 윤준성, 정현철, 최영우, 김재희, "잡영이 포함된 저해상도 전표 영상의 분석", 제11회 영상처리 및 이해에 관한 워크샵 발표 논문집, pp. 91-96, 1999.2
- [15] 박문호, "인쇄된 한글문서의 폰트형 인식에 관한 연구", 광운대학교 석사학위논문, 1996
- [16] Robert M. Haralick, Linda G. Shapiro, *COMPUTER AND ROBOT VISION*, Vol. II, Addison-Wesley, 1993.
- [17] 정영미, *정보검색론*, 구미무역, 1993.2
- [18] Asha Vellaikal, C.-C. Jay Kuo, "Content-Based Image Retrieval Using Multiresolution Histogram Representaion," *SPIE Digital Image Storage and Archiving Systems*, pp. 312-323, Philadelphia, Oct. 1995.

## 저 자 소 개

徐昌德(正會員) 第34卷 C編 第5號 參照  
현재 한양대학교 전자전기공학부 박사과정



金 齋 律(正會員)

1980년 한양대학교 전자공학과 학사. 1983년 Pennsylvania State University 전기공학과 석사. 1989년 Purdue University 전기공학과 박사. 1989.9~1994.2 University of Texas 조교수. 1994.3~현재 한양대학교 부교수로 재직 주관심분야는 영상처리, 컴퓨터비전, 패턴인식, 컴퓨터그래픽스