

방향성분 특징과 Fisher Measure를 이용한 간판영상 한글인식

임 준 식[†] · 김 수 형^{**} · 이 귀 상^{**} · 양 형 정^{**} · 이 명 은^{***}

요 약

본 논문은 모바일 폰 기반의 간판 영상내 한글 문자인식에 관한 연구로써 인식 대상은 간판영상에서 추출된 상호명으로 하였고 인식대상 문자 수는 상호명 빈도수 기반 808자로 한정하였다. 인식과정은 특징 추출, 대분류, 상세 분류로 구성되고 특징 추출과정에서는 문자영상의 크기, 잡음 및 왜곡에 강건한 비선형 방향성분 특징을 이용하였고 대분류 과정에서는 추출된 특징과 인식 대상문자에 대하여 최소거리 분류를 수행하고 10순위까지의 후보 문자를 추출하였다. 상세 분류 과정에서는 Fisher discriminant measure 이용하여 대분류에서 발생 할 수 있는 오인식 결과를 보완하였다. 실험결과 1순위 인식률은 80.45%이고 5순위까지의 누적 인식률은 93.51%를 보였다.

키워드 : 간판인식, 비선형 방향성분, 특징, 문자인식

Recognition of Korean Text in Outdoor Signboard Images Using Directional Feature and Fisher Measure

Lim Jun-Sik[†] · Kim Soo-Hyung^{**} · Lee Guee-Sang^{**} · Yang Hyung-Jung^{**} · Lee Myung-Eun^{***}

ABSTRACT

In this paper, we propose a Korean character recognition method from outboard signboard images. We have chosen 808 classes of Korean characters by an analysis of frequencies of appearance in a dictionary of signboard names. The proposed method mainly consists of three steps: feature extraction, rough classification, and coarse classification. The first step is to extract a nonlinear directional segments feature, which is immune to the distortion of character shapes. The second step computes an ordered set of 10 recognition candidates using a minimum distance classifier. The last step reorders the recognition candidates using a Fisher discriminant measure. As experimental results, the recognition accuracy is 80.45% for the first choice, and 93.51% for the top five choices.

Keywords : Signboard Recognition, Nonlinear Directional Segment, Feature, Character Recognition

1. 서 론

최근의 모바일 기기는 저전력, 고효율, 저렴한 가격의 하드웨어들이 개발됨에 따라 단순 음성 통화에서 영상 저장, 재생, 데이터 통신 및 영상 처리 기술 등을 적용한 응용분야로 사용 영역이 확대되고 있다. 이러한 하드웨어적 기술의 발전을 바탕으로 모바일 기기 사용자에게 다양한 정보를 제공하기 위한 목적으로 고해상도 카메라를 장착한 모바일 기기를 사용한 다양한 응용분야의 연구가 진행되고 있다[1-8].

본 논문에서는 간판영상을 대상으로 사용자에게 보다 효율적인 정보를 제공하기 위한 응용을 논하고자 한다. 예를 들어 간판 영상에서 획득된 텍스트를 인식 함으로써 외국인 여행자들에게 상점이나 기관 이름의 영문 번역 등의 내용을 제공 할 수 있다.

현재 자연영상 텍스트인식에 대하여 많은 연구가 진행 중인데, 미국 IBM Almaden 연구소에서는 카메라가 부착된 핸드폰이나 PDA등과 같은 개인 디지털 입력장치를 통해 현장에서 획득한 캡춰 이미지로부터 텍스트를 추출하고 자국의 언어로 번역하여 제공하는 기술에 대한 연구를 수행하였고[1], Carnegie Mellon 대학의 Interactive 시스템 연구소에서는 자연영상으로부터 추출한 텍스트를 인식하여 예제 기반 번역을 수행하는 연구를 진행하였다.[2-5]

류코쿠 대학과 시가대학에서는 모바일폰 기반 카메라를 이용하여 영상에서 텍스트를 인식에 관한 연구를 수행하였

※ 이 논문은 2007년도 전남대학교 연구년 교수연구비 지원에 의하여 연구되었음.

† 준 회 원 : 전남대학교 전자컴퓨터공학부 박사과정

** 중 신 회 원 : 전남대학교 전자컴퓨터공학부 교수

*** 정 회 원 : 전남대학교 전자컴퓨터공학부 연구원

논문접수 : 2008년 12월 29일

수정일 : 1차 2009년 2월 5일, 2차 2009년 3월 12일

심사완료 : 2009년 3월 13일

고 류코쿠 대학의 NHK 과학 기술 연구소에서는 디지털 카메라로 획득된 자연 영상으로부터 일본 문자를 검출하여 영어로 변환하는 연구를 수행하였다.[9] 중국 Tshinghua 대학의 OCR연구소에서는 자연 영상으로부터 텍스트의 위치를 찾아 인식하는 연구를 진행하였다.[10]

국내에서는 ETRI 영상인식연구팀에서 정보통신부의 “차세대 통합 단말 기술 개발”사업의 일환으로, 디지털 카메라로 간판, 책자 등을 촬영하면 내용을 즉시 문자 정보로 변환시키는 “휴대용 디지털 카메라 문자인식 핵심 기술”을 개발하였다.[11] 또한 전북대학교의 허기수 등은 간판 영상 인식을 목적으로 간판 영상 데이터베이스를 제작하였고[12] 전남대학교 박현일 등은 모바일 기기로 획득된 간판영상에서 전화번호를 인식함으로써 간판 영상에 대한 정보를 사용자에게 전달하고자 하였다.[13]

한글 문자 인식에 관련된 기존 연구를 살펴 보면 크게 구조적 방법과 통계적 방법으로 구분 할 수 있다.[14] 구조적 특징은 입력 문자의 기하학적인 모양에 기반을 두고 사람이 직관적으로 이해 할 수 있는 형태학적 분석을 수행한다.[15, 16] 통계적 특징은 입력 문자의 형태적 변환을 수행하여 수치적인 결과로 변환하는 과정을 수행하게 된다.[17-19]

본 논문의 연구 목적은 모바일 단말기에서 간판영상의 문자를 인식함으로써 간판영상의 정보를 사용자에게 전달하는데 있다. 이를 위하여 영상의 크기 및 왜곡에 강건한 방향 성분 특징을 이용하여 간판영상의 촬영 시 발생하는 문제점들을 해결하였고 대분류 및 상세 분류로 구성되는 2단계 인식 방법을 적용하여 1순위 인식률을 80.45%까지 향상 시켰다.

2. 한글 간판 인식

2.1 모바일 환경에서 간판 인식의 어려움

모바일 환경에서 문자의 인식은 처리 시간 및 자원의 한계로 인한 많은 문제점 들이 발생한다. 이러한 문제점은 글자영역을 추출하는 이진화 과정 및 글자의 분할을 수행하는 과정에서 발생하게 되며 크게 3가지 유형으로 구분 할 수 있다.

첫 번째는 그림자 및 조명의 영향으로 인해 글자 영역에 발생하는 잡음이다. 본 논문에서는 문자 인식의 상세 분류 과정에서 이러한 문제를 해결하고자 한다. (그림 1.a)은 조명의 영향이 강하게 나타나는 간판 영상의 예이다.

두 번째는 장애물로 인한 문자 영역의 손실로써 실제 간판영상을 촬영해 보면 (그림 1.b)와 같이 전선과 같은 장애



(a) 조명에 의한 손상 (b) 장애물에 의한 손상
(그림 1) 조명 및 장애물로 인한 문자 영역 손상

물이 간판영상에 포함되어 촬영된다. 이러한 영상을 이진화하게 되면 글자 영역에 많은 손실을 가져 오게 된다. 이러한 문제점은 자소 단위 인식 및 형태학적 인식 알고리즘을 적용하였을 경우 오인식의 주요 원인이 되게 된다.

마지막으로 간판 영상에는 촬영 거리 및 각도에 따른 글자의 크기의 다양성, 글자의 기울어짐, 글자 형태의 왜곡 등이 발생한다. 이러한 문제점은 글자영역 대한 크기 변환 작업 및 글자영역의 기울기 보정 작업이 필요하지만, 본 논문에서는 특징 추출 방법으로 비선형 분할을 수행함으로써 글자의 크기에 따른 영향을 최소화 하였고 방향 성분 특징을 이용하여 글자영역의 왜곡에 따른 문제점을 최소화 하였다.

2.2 이진화 및 글자 분할

문자 단위 인식을 위해서는 간판 영상에서 문자 영역과 배경영역을 구분하는 이진화 과정과 구분된 문자 영역에서 문자 단위로 분할하는 과정이 필요하다. 본 논문에서는 이진화 과정을 위하여 박종현 등이 제안한 퍼지 클러스터링과 히스토그램 분포를 이용한 이진화 방법을 사용하였으며[20] 문자 분리를 위하여 김의철 등이 제안한 연결 요소 분석에 의한 간판 영상의 문자 분할 방법을 사용하였다.[21] 박종현 등이 제안한 이진화 방법은 텍스트 영역의 검출 및 이진화를 위하여 모바일 폰에 장착된 카메라로부터 획득된 RGB 영상을 gray 영상으로 변환하고 수평 및 수직 방향의 에지 히스토그램을 추출하여 텍스트 영역을 검출한다. 검출된 텍스트 영역의 이진화를 위하여 칼라 클러스터링 방법을 적용하여 배경과 텍스트 영역을 구분하였다.

김의철 등이 제안한 문자 분할 방법은 글자 영역의 연결 요소를 분석 후 한글의 초, 중, 종성의 위치 관계를 바탕으로 병합하는 과정을 수행하여 문자 단위 분리를 수행하였다. (그림 2)는 모바일 시스템에 응용하기 위해 제안된 간판 영상에서의 텍스트 검출 및 이진화를 보여주고 있으며 (그림 3)은 이진화된 영상에서 문자 단위 분할 결과를 보여주고 있다.



(그림 2) 텍스트 검출 및 이진화

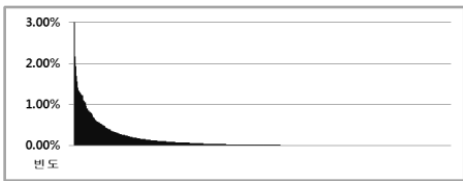


(그림 3) 문자 단위 분할

2.3 인식 대상 문자

한글은 총 2350자로 구성되는데 이중 광고 및 홍보를 위

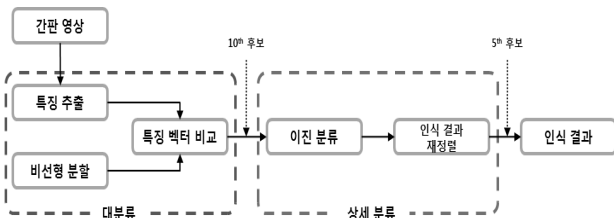
해 사용되는 문자는 한정되어 있다. 본 논문에서는 인식 대상 문자를 축소하기 위하여 진화 번호부를 활용하였고 총 400만개의 상호명으로부터 명사를 분할하고 분할된 명사 내 글자들의 빈도수를 측정하였다. (그림 4)는 진화 번호부 상호명에 사용되는 문자들을 빈도수 별로 보여 주고 있으며, 빈도수 순으로 808자의 문자들이 상호명에 사용되는 전체 문자의 95%이상을 차지 하는 것을 알 수 있었다. 이러한 결과로 인식 성능 향상과 처리 시간 단축을 위하여 인식 대상 문자를 2,350자가 아닌 808자로 한정하였다.



(그림 4) 상호명 기반 빈도수

2.4 한글 간판 인식 과정 개요

(그림 5)는 모바일 기기에서 한글 간판 영상내의 문자 인식 과정의 구성도이며 한글 간판 인식 과정은 대분류와 상세분류로 구성된다. 대분류 단계에서는 비선형 분할을 통해 방향 성분 특징(Directional Segment Feature : DSF) 을 추출하고 추출된 특징과 인식 대상 문자별로 훈련된 특징 사이의 Manhattan distance를 구하여 10순위까지의 후보를 추출하게 된다. 상세 분류는 Fisher discriminant measure를 이용하여 추출된 10순위 결과를 재정렬하고 5순위까지의 결과를 최종 인식 결과로 출력하게 된다.



(그림 5) 인식 과정 구성도

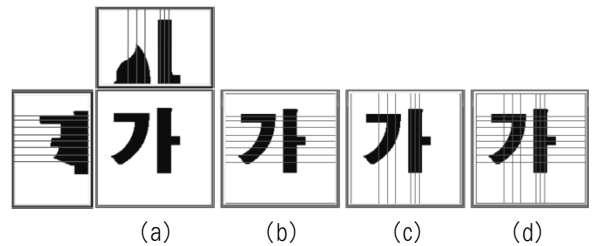
3. 특징 추출 및 대분류

대분류 단계에서는 입력 영상의 특징을 추출하고 각각의 인식 대상 문자의 특징과 입력 영상의 특징을 비교하여 정렬한 후 10순위까지의 결과를 추출하게 된다. 일반적인 문자 인식에서는 영상에서 문자들의 위치 및 배치 정보를 활용하여 문자들의 왜곡 보정을 수행하지만 간판영상은 문자들 사이의 간격이 넓고 문자의 수가 많지 않아서 영상의 보정이 어렵고 인식에 많은 어려움이 따른다. 또한 모바일 기기에서는 영상 변환과 같은 많은 처리시간을 요구하는 작업을 수행하기 어렵다. 이러한 문제들을 해결하고자 영상의 왜곡에 강건한 방향 성분 특징과 비선형 분할을 이용하여 특징을 추출하였다.

3.1 비선형 분할

특징 추출 과정에서 입력 문자영상의 형태 및 크기 변화를 주지 않고 특징을 추출하기 위하여 영상의 크기에 독립적인 성질을 갖는 비선형 분할을 수행 하였다. 비선형 분할은 선형 분할에 비해 처리과정이 복잡하지만 영상을 변환하는 과정을 생략 할 수 있고 영상의 크기 변화에 따른 왜곡도 극복 할 수 있다. 비선형 분할 과정은 다음과 같이 수행 된다.

- ① 문자 영상의 수평방향 히스토그램과 수직 방향 히스토그램을 계산한다.
- ② 각각의 히스토그램에 대하여 총 합을 구하고 나누고자 하는 mesh의 크기(N)로 나누어 각각의 임계값을 구한다. (그림 6.a)
- ③ 각각의 히스토그램에 대하여 임계값의 크기에 따라 분할한다. (그림 6.b, 6.c)
- ④ 수평 분할 값과 수직 분할 값으로 원영상을 분할한다. (그림 6.d)



(그림 6) non-linear mesh (a)히스토그램 (b)수평분할 (c)수직 분할 (d)분할결과

비선형 분할 과정을 수행한 후 각 cell의 특징 값을 계산한다. (식 1)은 각 cell의 특징 값을 구하는 과정으로 cell 내부의 black pixel들의 특징 값의 평균을 계산하게 된다. 여기서 $Cell_i(F)$ 는 i 번째 cell의 특징 벡터이고, N 은 cell 내부의 검은 화소의 수, F_k 는 k 번째 검은 화소의 특징 벡터이다.

$$Cell_i(F) = \sum_{k=1}^N F_k / N \tag{식1}$$

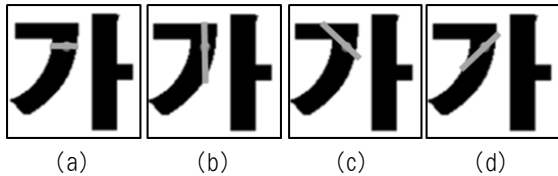
3.2 방향 성분 특징

방향 성분 특징은 각 화소가 문자 영역에서 수평 및 수직 성분과 대각 성분을 고려하여 특징을 추출하므로 문자의 왜곡에 강건한 특징을 추출 할 수 있고 처리 과정이 간단하여 처리 시간을 최소화 할 수 있다. 그리고 방향성분 특징 추출과정에서 대각 성분을 측정하기 위하여 영상을 45도 회전하여 측정하게 되는데 영상 회전은 처리 시간 증가와 영상의 변형을 가져오므로 본 논문에서는 영상을 회전하지 않고 삼각 함수를 이용하여 대각 성분의 특징 값을 근사하였다. 방향 성분 특징의 처리 단계는 다음과 같이 나눌 수 있다.

- ① 각 점(x,y)에 대하여 수평 방향 성분(RLH_{xy})을 계산한다.

(그림 7.a)

- ② 각 점(x,y)에 대하여 수직 성분(RLV_{xy})을 계산한다. (그림 7.b)
- ③ 수평 방향 성분(RLH_{xy})과 수직 성분(RLV_{xy})을 가지고 기여도 (DCH_{xy})와 (DCV_{xy})를 계산한다. (수식 2)
- ④ 대각 성분(RLR_{xy}, RLL_{xy})에 대하여 ①,②,③번 과정을 반복 수행한다. (그림 7.c) (그림 7.d)



(그림 7) 방향성분 특징 (a)수평성분 (b)수직성분 (c)대각성분 (d)대각성분

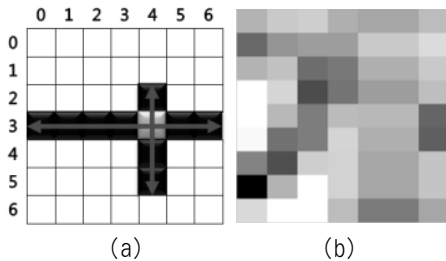
$$DCH_{x,y} = \frac{RLH_{x,y}}{(RLH_{x,y} + RLV_{x,y})}$$

$$DCV_{x,y} = \frac{RLV_{x,y}}{(RLH_{x,y} + RLV_{x,y})}, \quad (식2)$$

where $0 \leq x < W, 0 \leq y < H$

(식 2)에서 점(x,y)의 수평성분 기여도(DCH_{xy})는 수평 방향 성분(RLH_{xy})과 수직 성분(RLV_{xy})의 합으로 수평 방향 성분(RLH_{xy})을 나눈 값이다. 이와 같은 과정을 반복 수행하면 각 점(x,y)에 대하여 4가지 방향에 대한 기여도 DCH_{xy}, DCV_{xy}, DCR_{xy}, DCL_{xy}의 값을 구하게 된다.

(그림 8.a)는 방향성분의 이해를 돕기 위한 것이다. 점(4.3)에 대한 수평 방향 성분은 점(4.3)의 X축 방향의 black pixel의 수를 측정하여 구할 수 있고, 수직 방향 성분은 Y축 방향의 black pixel의 수를 측정하여 구할 수 있다.



(그림 8) 방향 성분 특징 (a)수직 수평 성분 추출 방법 (b) 인식 대상글자 “가”의 수평 성분

비선형 분할과정의 수행 후 각 화소의 방향성분 특징과 비선형 분할 정보를 바탕으로 각각의 cell에 대한 특징을 구하게 된다. 이때 각cell의 특징 값은 cell에 해당하는 검은 화소들에 방향 성분 값의 평균으로 한다. (그림 8.b)는 인식 대상 문자 “가”의 수평성분 특징 값이다. 비선형 분할 수행 시 분할 크기를 다양하게 추출하여 비교 분석하였고 가장 높은 성능은 보인 7x9의 mesh를 사용하였고, 따라서 하나

의 문자영상으로부터 추출되는 특징 벡터 차원의 수는 7x9x4(252)개 이다.

3.3 Manhattan distance를 이용한 대분류

입력 문자 영상에서 획득된 특징 벡터와 인식 대상 문자들의 특징 벡터 집합을 비교하는 과정으로 (식 3)을 수행하게 된다. 여기서 F 는 입력영상의 특징 벡터이고 O_i 은 i 번째 인식 대상 문자의 특징 벡터, D 는 입력 문자 영상과 인식대상 문자 집합 간의 특징 벡터의 비교 거리이고 d_i 은 i 번째 인식 대상 문자와 입력 영상간의 특징 벡터의 거리이다. 두 벡터간 거리 d_i 가 작은 값을 가질수록 유사한 특징을 갖는 문자가 된다.

$$F = \langle f_0, f_1, \dots, f_{250}, f_{251} \rangle$$

$$O_i = \langle o_{i0}, o_{i1}, \dots, o_{i250}, o_{i251} \rangle, i = 0 \sim 807$$

$$D = \langle d_0, d_1, \dots, d_{806}, d_{807} \rangle \quad (식3)$$

여기서 $d_i = \sum_{j=0}^{251} |f_j - o_{ij}|$

입력 문자와 인식대상 문자간의 특징 벡터의 거리 D 를 구하게 되면 오름차순 삽입 정렬을 수행하여 10순위까지의 인식 후보 그룹을 추출하게 된다.<표 1>

<표 1> 입력 문자 영상 “태”와 “래”의 대분류 결과

문자 영상	대분류 결과										
	순위	1	2	3	4	5	6	7	8	9	10
태	후보 문자	태	태	테	쇄	패	제	체	채	대	
래	후보 문자	래	빼	레	렉	배	깨	째	베	패	

본 장에서는 특징 벡터 추출 및 비교, 후보그룹 추출과정에 대하여 설명하였다. Manhattan distance를 이용한 대분류를 수행한 후 획득된 후보 그룹은 정확도가 높지 않지만 특징 비교까지의 처리 시간이 빠르며 처리 과정도 단순하다. 다음 장에 설명할 Fisher차원은 많은 양의 저장 공간 및 처리 시간을 갖기 때문에 후보 그룹의 수에 많은 영향을 받게 된다. 즉 대분류 후보그룹의 수에 따른 인식률과 처리 시간의 trade-off가 발생하고 후보 그룹의 수를 10으로 했을 경우 적당하다는 결과를 실험적으로 찾았다.

4. 상세 분류

상세 분류는 대분류 과정을 수행 후 추출된 후보그룹의 정확도를 향상 시키기 위한 과정으로서 후보그룹의 문자들과 입력 문자 영상의 특징 벡터를 비교하게 된다. 3장의 특징 추출 과정을 수행하였을 경우 서로 다른 글자 임에도 불구하고 비슷한 특징 벡터를 보이는 경우가 발생하고 이는 전체적인 인식 과정에서 인식률을 떨어뜨리는 가장 큰 요인

이다. 이러한 문제점을 해결하는 방법으로 Fisher 차원을 이용한 이진 분류기를 사용하였고 이진 분류기의 결과를 합산한 후 상위 5순위를 최종 인식 결과로 추출하게 된다.

4.1 Fisher 차원

Fisher 차원은 대분류의 특징 추출 단계에서 획득된 7x9x4차원의 특징 벡터 252개의 값 중 (식 4)로 표현되는 Fisher discriminant measure가 큰 32개를 문자 쌍 별로 선정할 것이다.

$$F_{ij}(k) = \partial_{ij}(k) / \{ \partial_i(k) + \partial_j(k) \} \quad (식4)$$

여기서 $\partial_i(k)$ 는 i -th 문자의 모든 샘플의 k -th 특징 값의 분산, $\partial_j(k)$ 는 j -th 문자의 모든 샘플의 k -th 특징 값의 분산, $\partial_{ij}(k)$ 는 i -th 문자의 모든 샘플과 j -th 문자의 모든 샘플의 k -th 특징 값의 분산이다. (식 4)는 i -th 문자와 j -th 문자의 특징 벡터들의 유사성을 측정하는 것이 목적으로 두 문자를 구분 하는데 필요한 특징 벡터를 선택하기 위하여 사용된다.

인식 대상 문자에 대한 Fisher 차원을 사전에 작성 하였고 정의된 Fisher 차원 순서쌍은 ${}_{808}C_2=326,028$ 개로 구성된다. 여기서 326,028개가 각각32차원이므로 모바일 기기 환경에서는 자원의 한계로 인해 사용하기 어려우며 각 특징값은 실수형이므로 많은 저장 공간을 필요로 한다. 이러한 문제점을 해결하기 위하여 326,028개의 순서쌍을 문자의 ASCII 코드 순으로 오름차순 정렬 후 index를 작성하였고 특징 값의 순번만을 short형으로 저장하여 실제 처리 과정에서는 필요한 순서쌍만 입력 받으므로 자원의 낭비 및 파일 입출력에 대한 처리 시간을 최소화 하였다.

4.2 인식 결과 재구성

Fisher 차원에 의한 이진 인식은 앞 절에서 설명한 252차원의 벡터를 기반으로 하여 32차원의 특징 공간 상에서의 거리를 이용한 이진 분류 함수를 구현하는 것이다. 즉 입력 문자 영상으로부터 추출한 252차원 방향성분 특징 벡터 F 와 인식 대상 문자 a, b 의 대표 벡터와의 유사도를 측정함에 있어 a 문자에 대해 정의된 Fisher차원(32차원의 부분 벡터) $Fisher_a$ 와의 거리 d_a 와 b 문자에 대해 정의된 Fisher차원(32차원의 부분 벡터) $Fisher_b$ 와의 거리 d_b 중 작은 값을 갖는 문자를 이진 분류 함수의 출력으로 간주 한다.(식 5)

$$F = \langle f_0, f_1, \dots, f_{250}, f_{251} \rangle$$

$$Fisher_a = \langle df_{a0}, df_{a1}, \dots, df_{a30}, df_{a31} \rangle$$

$$Fisher_b = \langle df_{b0}, df_{b1}, \dots, df_{b30}, df_{b31} \rangle \quad (식 5)$$

if $d_a < d_b$
 than output a
 else output b

이진분류 함수의 결과 재구성 과정은 대분류 과정에서 추출된 인식 결과의 순서쌍들에 대한 Fisher 차원에 대한 벡터를 불러오는 과정부터 시작된다. 순서쌍들의 Fisher 차원은 각 특징 벡터의 순번이므로 입력문자의 특징 벡터와 순서쌍 각각의 특징 벡터의 Fisher 차원에 해당하는 순번만을 이진 분류 함수의 입력으로 삼게 된다. 본 논문에서는 모든 순서쌍에(대분류에서 후보 그룹의 수가 10이므로 총 45개의 순서쌍) 대한 이진 분류를 수행하고 순서쌍 각각의 Fisher 차원 거리 비교 값을 이진 분류의 결과로 출력하게 되고 출력 결과를 각각의 후보 문자별로 합산하였다. 마지막으로 합산된 결과 값에 대하여 오름차순 정렬을 수행하고 정렬결과 중 상위 5개에 대하여 상세분류의 인식 결과로 출력하게 된다. <표 2>와 <표 3>은 대분류와 상세 분류의 출력 결과를 보여 준다.

<표 2>입력 문자 영상 “래”의 인식 결과

문자영상	대분류 결과										
	순위	1	2	3	4	5	6	7	8	9	10
래	문자	래	배	레	랙	배	개	제	베	패	
	상세 분류 결과										
	순위	1	2	3	4	5					
	문자	래	레	레	랙	배					

<표 3> 입력 문자 영상 “태”의 인식 결과

문자영상	대분류 결과										
	순위	1	2	3	4	5	6	7	8	9	10
태	문자	태	테	테	태	쇄	패	제	채	채	대
	상세 분류 결과										
	순위	1	2	3	4	5					
	문자	태	태	테	테	패					

5. 실험 환경 및 결과

5.1 실험 환경

실험을 위해 삼성 전자 SPH-M4650을 사용하였다. 실험 영상은 실험에 사용된 모바일 기기로 직접 촬영하여 수집하였고 수집된 영상에 대하여 모바일 기기 상에서 실험 하였다. (그림 9)는 모바일 기기상의 인식 프로그램 결과 영상이다. 좌측 상단은 인식 결과를 5순위 까지 보여주고 있으며 중간 부분은 이진화된 입력 영상이다. 그리고 하단에는 인식과정에서 소요된 시간을 보여주고 있으며 인식과정에 소



(그림 9) 인식 프로그램

요된 시간은 DB 파일을 읽어 오는 시간 및 인식과정에서 사용되는 모든 과정을 포함하게 된다. 인식 소요 시간은 문자당 평균 0.57초 간판당 3.1초의 처리 시간을 보였다.

5.2 인식 성능 분석

<표 4>는 대분류와 상세분류의 5순위까지의 인식 결과를 보여 주고 있다. 대분류는 10순위의 결과를 추출하지만 상세 분류와 정확한 비교를 위하여 5순위까지 측정된 결과를 기술한다. (표 4)에서 보는 것과 같이 대분류 시 1순위 인식이 37.83%, 5순위까지의 누적 인식이 84.55%이고, 상세 분류는 1순위 인식이 80.45%, 5순위까지의 누적 인식이 93.51%이며 이러한 결과는 상세 분류를 통해 정확율이 향상 된다는 것을 보여 주고 있다.

간판 문자 인식은 각각의 문자를 인식하는 성능도 중요하지만 간판에 있는 문자 전체를 인식 해야만 정확한 정보를 획득 할 수 있다. 그래서 본 논문은 간판 단위의 인식을 측정하였고 <표 5>는 그 결과이다. 간판 단위 인식이란 간판에 포함된 문자 모두가 5순위 후보군에 속할 때만 성공했다고 평가 했다. 인식 성공에서 각 순위는 간판 영상의 인식 결과 중 가장 낮은 순위를 표기한 것이다. 간판 단위 인식 결과에서 각 순위별 인식 결과는 간판 영상의 모든 문자에 대한 인식 결과가 해당 순위 안에 존재 한다는 것이고 이 결과는 간판 문자 데이터 베이스 검색 서비스 또는 간판 영상 번역 시스템의 중요한 지표로 사용된다. 예로 간판 영상 번역 시스템에서 사전기반 인식 및 번역을 수행하게 될 때 5순위까지의 인식 결과를 사용하여 작업을 수행하게 되면 간판의 글자수를 n이라고 할 때 5ⁿ 개의 글자 조합에 대하여 사전기반 유사도 검사를 수행하면 된다.

<표 4> 대분류와 상세분류의 인식 결과 분석

글자단위	대분류		상세 분류		
	5th인식률	누적인식률	5th인식률	누적인식률	
인식 성공	1th	37.83%	37.83%	80.45%	80.45%
	2th	36.59%	74.42%	7.34%	87.79%
	3th	4.09%	78.51%	2.32%	90.11%
	4th	3.37%	81.88%	0.46%	90.57%
	5th	2.66%	84.55%	2.94%	93.51%

<표 5> 간판 단위 인식 결과

간판 단위	인식수	인식률	누적 인식률	
인식 성공	1th	112	46.09%	46.09%
	2th	36	14.81%	60.91%
	3th	12	4.94%	65.84%
	4th	2	0.82%	66.67%
	5th	30	12.35%	79.01%
인식 실패	51	20.99%	20.99%	
합계	243			

5.3 상용 인식 프로그램과 비교 평가

본 논문에서 제안하는 한글인식 방법의 객관적 평가를 위해 최근에 판매되고 있는 상용 OCR소프트웨어와 비교 분석

하였다. 비교 대상으로는 아르미 6.5 pro와 HI 글눈 2003을 대상으로 하였고 인식 성능 비교 과정은 각각의 인식 프로그램이 최적의 성능을 나타낼 수 있도록 최적의 환경을 구성하였다. 아르미 6.5 pro과 글눈 2003의 입력으로 이진화된 글자 영역에 대하여 간판 단위로 입력하였고 분할과정에서 발생 할 수 있는 오류를 최소화 하기 위하여 분할에 실패할 경우 임의로 분할하여 실험 하였다. 두 프로그램 모두 3순위까지의 인식 결과를 제공하므로 제안한 알고리즘의 3순위까지의 누적 인식 결과를 비교 대상으로 하였다. <표 6>은 성능 비교 분석 결과이며 상용 인식기에 비해 제안된 알고리즘이 문자단위 인식성능과 간판단위 인식성능 모두에서 높은 수치를 보였다.

<표 6> 간판 단위 인식 결과

	제안 방법	아르미6.5	글눈 2003
3 th 누적인식률	90.1%	73.2%	69.7%
간판단위 인식률	65.8%	33.7%	31.6%

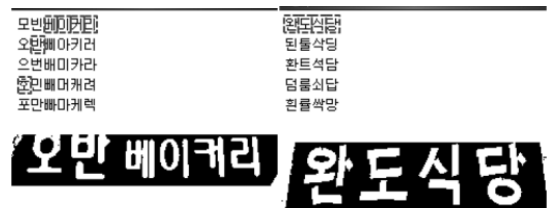
5.4 실험 결과

(그림 10)은 이진화된 간판 영상과 인식 결과이다. “현대 남선 비뇨기과”에서 “현”자에 해당하는 글자 영역에 조명의 영향으로 손실이 발생하였으나 1순위에서 인식 성공하였다.” 김형재 치과의원” 영상은 촬영 당시 간판과 촬영자 사이에 전선이 존재하여 글자영역 전반에 걸쳐 손실이 발생하였지만 모든 글자에 대하여 1순위 인식에 성공하였다.

(그림 11)은 모바일의 환경에서 입력 영상, 인식 결과, 이진화된 영상 그리고 소요시간을 보여주고 있다.



(그림 10) 인식 결과(1)



(그림 11) 인식 결과(2)

6. 결 론

본 논문에서는 간판 영상 내의 한글 인식 과정에 대하여 제안하였다. 간판 영상 인식 과정은 특징 추출 및 대분류, 상세분류로 구성된다. 특징 추출 방법은 영상의 왜곡에 강건한 방향 성분 특징을 사용하였고 대분류 과정에서 Manhattan distance를 이용하여 특징 비교를 수행하였다. 대분류 과정에서 10순위까지의 후보 그룹을 생성하고 상세 분류에서 Fisher 차원을 이용하여 재정렬하였다. 본 논문에서 제안한 인식 시스템은 간판 영상에 존재 하는 잡음, 손실, 기울어짐에 강건하고 문자 각각의 인식 결과 및 간판 영상 단위의 인식 결과가 우수함을 보였다. 하지만 간판 영상에서는 문자 이외의 다양한 정보가 존재하고 있다. 이를 위하여 간판 영상에서 문자 외에 획득 할 수 있는 심볼, 기호 등의 인식에 관한 연구를 진행할 것이다. 또한 최근 간판 영상에 많이 사용되고 있는 디자인 형태의 글꼴 인식에 관한 연구도 수행할 것이다.

참 고 문 헌

- [1] I. Haritaoglu, "Scene text extraction and translation for handheld devices," Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, Vol.2, pp.408-413, 2001.
- [2] C. Thillou, B. Gosselin, "Character Segmentation-by-Recognition Using Log-Gabor Filters", Proc. of IAPR International conference on Pattern Recognition, pp.901-904, 2006.
- [3] C. Thillou, B. Gosselin, "Color binarization for complex camera-based images," Proc. Electronic Imaging, pp.301-308, 2005.
- [4] C. Thillou, B. Gosselin, "Color Text Extraction from Camera-based Images the Impact of the Choice of the Clustering Distance," Proc. of IEEE International Conference on Documents Analysis and Recognition, pp.312-316, 2005.
- [5] C. THILLOU, "Natural Scene Text Understanding", Vision Systems, Segmentation and Pattern Recognition, Ch.16, pp.307-333, 2007.
- [6] 정인숙, 함대성, 오일석, "웹 이미지의 텍스트 추출을 위한 색 분산 방법의 실험적 평가", 영상처리 및 이해에 관한 워크샵, p.36, 2008.
- [7] 김지훈, 이택현, 김기웅, 김진형, "상향식 접근 방법을 사용한 강인한 간판 인식", 한국컴퓨터종합 학술대회 논문집, 제34권, 제1호, pp.234-235, 2007
- [8] J. Zhang, X. Chen, J. Yang, and A. Waibel, "A PDA-based Sign Translator," Proceedings of the Fourth IEEE International Conference on Pattern Recognition, pp.216-219, Oct., 2002.
- [9] Y. Watanabe, Y. Okada, Y. B. Kim, and T. Takeda, "Translation camera," Proceedings of International Conference on Pattern Recognition, Vol.1, pp.613-617, 1998.
- [10] C. Li, X. Ding, and Y. Wu, "Automatic text location in natural scene images," Proceedings of International Conference on Document Analysis and Recognition, pp.1069-1073, 2001.
- [11] <http://www.kosef.re.kr/bbs/>
- [12] 허기수, 오일석, "간판인식 연구를 위한 간판영상 데이터베이스," 한국 정보과학회 춘계 학술발표회, 2005.7.6
- [13] 박현일, 김수형, "칼라 군집화 및 비등방성확산필터를 이용한 저해상도 영상에서의 숫자열 인식," 제22회 한국정보처리학회 추계학술발표대회 논문집, 제11권 제2호, pp.839-842, 2004년 11월.
- [14] 오일석, Ching Y. Suen, "광학 문자 인식을 위한 거리 특징," 한국정보과학회논문지(B), 제25권 제7호, pp.1028-1043, 1998년 7월.
- [15] M. Shridhar, A. Badreldin. "A high-accuracy syntactic recognition algorithm for handwritten numerals," IEEE Tr. SMC 15(1), pp.152-158, 1985.
- [16] T. Pavlidis, F. Ali. "Computer recognition of handwritten numerals by polygonal approximations," IEEE Tr. SMC 5(6), pp.610-614, 1975.
- [17] Y. Le Cun, O. Matan, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, L.D. Jackel, H.S. Baird, "Handwritten Zip code recognition with multilayer networks," Proc. ICPR'90, pp.35-40, 1990.
- [18] J. Cao, M. Ahmadi, M. Shridhar, "A hierarchical neural network architecture for handwritten numeral recognition," Pattern Recognition, 30(2), pp.289-294, 1997
- [19] N.W. Strathy, C.Y. Suen, "A new system for reading handwritten ZIP codes," Proc. ICDAR, pp.74-77, Montreal, Canada, 1995
- [20] J.H. Park, T. N. Dinh, G.S. Lee, "Binarization of Text Region based on Fuzzy Clustering and Histogram Distribution in Signboards", Proc. WORLD ACADEMY OF SCIENCE, ENGINEERING AND TECHNOLOGY, Vol.33, pp.85-90, 2008.
- [21] 김의철, 김수형, 양형정 오상욱, "연결 요소 분석에 의한 간판 영상의 문자 분할," 한국멀티미디어학회 춘계 학술발표회, 2008년 5월.



임 준 식

e-mail : goethe0708@iip.chonnam.ac.kr

2004년 전남대학교 전산학과(이학사)

2007년 전남대학교 전산학과(공학석사)

2007년~현재 전남대학교 전자컴퓨터공학부 박사과정

관심분야 : 인공지능, 패턴인식, 문자 인식,

객체 추적 등



김수형

e-mail : shkim@chonnam.ac.kr
1986년 서울대학교 컴퓨터공학과(공학사)
1988년 한국과학기술원 전산학과(공학석사)
1993년 한국과학기술원 전산학과(공학박사)
1990년~1996년 삼성전자 멀티미디어연구소
선임연구원

2000년~2002년 캐나다 Concordia 대학 CENPARMI 연구소 방문교수

1997년 1월~현 재 전남대학교 전자컴퓨터공학부 교수
관심분야 : 인공지능, 패턴인식, 문서영상 정보검색, 유비쿼터스 컴퓨팅 등



양형정

e-mail : hjyang@chonnam.ac.kr
1991년 전북대학교 전산통계학과(학사)
1993년 전북대학교 전산통계학과(석사)
1998년 전북대학교 전산통계학과(박사)
2003년~2005년 카네기멜런 대학교 연구원
2005년~현 재 전남대학교 전자컴퓨터공학부 교수

관심분야 : e-Design, 데이터 마이닝, e-Learning 등



이귀상

e-mail : gslee@chonnam.ac.kr
1980년 서울대학교전기공학과(공학사)
1982년 서울대학교 전자계산기공학과(공학석사)
1991년 Pennsylvania 주립대학 전산학
(이학박사)

1984년~현 재 전남대학교전자컴퓨터공학부 교수
관심분야 : 멀티미디어통신, 영상처리및컴퓨터 등



이명은

e-mail : melee@chonnam.ac.kr
1998년 목포대학교 전자공학과(학사)
2001년 목포대학교 전자공학과(공학석사)
2007년 목포대학교 전자공학과(공학박사)
2006년~2007년 목포대학교 정보공학부
초빙교수

2007년~현 재 전남대학교 전자컴퓨터공학부 연구원
관심분야 : 의료영상처리, 3차원 시각화, 패턴인식, 인공지능 등