

간판영상에서 한글 인식 성능향상을 위한 가중치 기반 음소 단위 분할 교정

Weighted Disassemble-based Correction Method to Improve Recognition Rates of Korean Text in Signboard Images

이명훈*, 양형정**, 김수형**, 이귀상**, 김선희***

코난테크놀로지 미디어기술지원*, 전남대학교 전자 컴퓨터 공학부**, Carnegie Mellon University***

Myung-Hun Lee(kh2821@naver.com)*, Hyung-Jeong Yang(hjyang@jnu.ac.kr)**,

Soo-Hyung Kim(shkim@jnu.ac.kr)**, Guee-Sang Lee(gslee@jnu.ac.kr)**,

Sun-Hee Kim(wkdal749@cs.cmu.edu)***

요약

본 논문에서는 휴대폰 카메라를 통해 간판영상의 한글문자를 인식한 후 오인식 된 결과를 교정하는 방법으로 인식 후보를 음소단위 분할하고 연산 가중치를 적용한 weighted Disassemble Levenshtein Distance(wDLLD)를 제안한다. 제안된 방법은 인식된 문자열을 음소 단위로 분할한 후 입력 형태의 거리 값을 산출하여, 가장 유사한 상호명을 데이터베이스에서 검출 한다. 제안된 방법의 효율성을 검증하기 위해, 전국의 상호명 중 중복되는 상호명을 제거한 130만개의 상호명을 이용하여 데이터베이스 사전을 구축 하였다. 또한 대표적인 문자열 비교 알고리즘인 Levenshtein Distance와 음소를 분할하여 적용한 Disassemble Levenshtein Distance 방법, 그리고 본 논문에서 제안한 인식 후보의 음소 단위 분할 방법과 연산 가중치를 적용한 weighted Disassemble Levenshtein Distance의 교정율을 비교 분석 하였다. 그 결과 제안된 weighted Disassemble Levenshtein Distance(wDLLD)은 Levenshtein Distance와 Disassemble Levenshtein Distance방법에 비해 각각 평균 29.85%와 6%의 인식률의 향상을 보였다.

■ 중심어 : | Disassemble Levenshtein Distance | 문자인식 | 간판 영상 | Levenshtein Distance | 후처리 |

Abstract

In this paper, we propose a correction method using phoneme unit segmentation to solve misrecognition of Korean Texts in signboard images using weighted Disassemble Levenshtein Distance. The proposed method calculates distances of recognized texts which are segmented into phoneme units and detects the best matched texts from signboard text database. For verifying the efficiency of the proposed method, a database dictionary is built using 1.3 million words of nationwide signboard through removing duplicated words. We compared the proposed method to Levenshtein Distance and Disassemble Levenshtein Distance which are common representative text string comparison algorithms. As a result, the proposed method based on weighted Disassemble Levenshtein Distance represents an improvement in recognition rates 29.85% and 6% on average compared to that of conventional methods, respectively.

■ keyword : | Disassemble Levenshtein Distance|Text Recognition | Signborad images | Levenshtein Distance | Postprocessing |

* "본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음 (NIPA-2012-C1090-1211-0008)"

* "이 논문은 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임 (2011-0029429)"

접수번호 : #120120-007

접수일자 : 2012년 01월 20일

심사완료일 : 2012년 02월 02일

교신저자 : 양형정, e-mail : hjyang@jnu.ac.kr

I. 서론

최근 스마트폰 기반의 멀티미디어 응용 어플리케이션이 개발되어 사용자가 필요한 비주얼 혹은 오디오 형태의 정보를 스마트폰을 통하여 캡처하고, 이와 관련된 정보의 검색이 가능하게 되었다. 따라서 점점 더 많은 사용자들이 정보 검색에 스마트폰을 사용하게 됨으로써 검색을 위해 멀티미디어 정보를 인식할 수 있는 인식 시스템의 연구가 활발하게 진행되고 있다. 그 중 가장 활발하게 연구 및 개발되고 있는 인식 방법은 카메라를 통한 자동 인식 시스템이다. 예를 들어, 카메라를 이용하여 바코드를 인식하거나[1], 도서의 표지를 인식하여 도서 정보를 제공하는 등 인식을 이용한 다양한 형태의 어플리케이션들이 등장하였다[2]. 이러한 어플리케이션의 공통점은 시각정보를 인식하여 사용자가 필요로 하는 형태로 정보를 제공하는 것이다.

우리의 주변에는 다양한 형태로 구성된 정보전달을 목적으로 하는 매개체들이 많다. 길거리에서 필요한 정보를 얻기 위해 가장 먼저 찾고 흔히 볼 수 있는 매체로 간판을 들 수 있다. 간판의 특징은 직관적인 시각 정보를 가장 많이 포함하고 있다는 것이다. 간판에서 정보를 획득하기 위한 방법으로는 스마트폰 및 PDA 등의 고해상도 컬러 카메라를 장착한 모바일 단말기를 사용하는 것이다. 예를 들어 외국 관광객이 국내에 왔을 때 가이드가 없다면 한글을 모르는 외국인은 한글로 표시된 한글 간판을 통해 의사를 결정하는데 어려움이 있다. 이러한 불편함을 해결하기 위해 모바일 폰으로 촬영한 간판영상의 텍스트를 인식한 후 번역하여 제공한다면 의사결정에 도움을 줄 수 있다. 이러한 시스템을 구현하기 위한 가장 핵심적인 기술 연구 중의 하나가 문자 인식 기술이다.

문자인식은 오랫동안 연구되어온 분야로, 현재는 다양한 분야에서 응용되고 있다. 문자인식은 비교적 단순한 문서의 인쇄체를 인식하는 단계를 지나 복잡한 배경을 가지는 자연영상의 텍스트를 인식하는 연구까지 진행되고 있다[3][4]. 그러나 문서와 같이 정형화된 형태를 인식하는 것보다 같이 높은 인식률을 비정형화된 형태의 매체에서는 기대하기 어렵다. 인간이 문자를 인식할

때는 개별문자를 인식 한 후에 전체문자의 문맥이 맞는지를 파악하게 된다. 즉, 눈에 보이는 문자의 형태에만 의존하지 않기 때문에 정확하게 인식할 수 있다. 그러나 대부분의 인식기는 전처리 과정을 거쳐 획득한 개별 문자를 인식하여 그 결과를 출력한다. 따라서 인간이 문자를 인식할 때와 같이 유연성을 갖춘 시스템을 만드는 데는 개별 문자의 인식 기술로는 한계가 있다. 그러므로 이러한 한계성을 극복하기 위한 방법으로 문자 인식과정에서 발생하는 오인식률을 줄일 수 있는 효과적인 방법이 요구된다.

본 논문에서는 간판 인식 및 번역 시스템에서 간판영상으로부터 한글의 개별문자 인식을 향상을 위한 인식 후보 음소 단위 분할 방법과 연산 가중치를 적용한 오인식 교정 후 처리 방법을 제안한다. 제안한 방법은 휴대폰 카메라를 이용하여 촬영한 간판 영상을 문자인식 기에서 인식한 후 오인식 된 문자를 교정함으로써 기존의 방법보다 인식률을 향상 시킬 수 있다. 오인식 문자 교정으로 인한 인식률의 향상은 사용자에게 잘못된 정보를 제공하여 혼란을 초래할 수 있는 문제점을 줄일 수 있을 뿐만 아니라 원하는 정보를 좀 더 정확하게 제공할 수 있다. 본 논문에서는 제안된 방법의 효율성을 검증하기 위해 130만개의 상호명 데이터베이스를 이용하여 실험을 수행하였고, 대표적인 문자열 비교 알고리즘인 Levenshtein Distance와 음소를 분할하여 적용한 Disassemble Levenshtein Distance 방법과 비교 하였다. 실험 결과 제안된 가중치와 음소 분할에 기반한 weighted Disassemble Levenshtein Distance(wDL) 방법은 Levenshtein Distance와 Disassemble Levenshtein Distance방법에 비해 각각 평균 29.85%와 6%의 인식률의 향상을 보였다.

논문의 구성은 다음과 같다. 2장에서는 문자인식 및 오인식 교정에 관한 연구들을 살펴보고, 3장에서는 본 논문에서 제안하는 weighted Disassemble Levenshtein Distance에 대해 기술한다. 4장에서는 기존의 인식기로부터 산출된 인식결과를 이용하여 제안한 알고리즘을 적용한 오인식 교정 실험 결과를 보이고 5장에서는 결론 및 향후 연구에 대하여 기술한다.

II. 관련연구

기존의 문자 인식은 대부분 문서나 명함과 같이 정형화된 형태로부터 스캔된 이미지 영상에서 텍스트를 추출하여 인식하였다. 그러나 최근에는 자연 영상으로부터 문자를 추출하는 연구가 진행되고 있다. 이러한 자연 영상에서 문자인식은 정형화된 문서와는 다르게 다양한 배경들이 존재하기 때문에 어려움을 겪고 있다. 뿐만 아니라 문자 인식에서 발생하는 오인식 문제로 인하여 잘못된 정보를 제공하게 되는 문제점을 가지고 있다. 따라서 자연 영상에서 문자를 추출하여 인식하는 연구와 인식의 오류률을 줄이기 위한 연구가 다양한 형태로 이루어지고 있다.

Haritaoglu는 카메라가 부착된 스마트폰이나 PDA와 같은 개인 휴대 기기로 획득한 자연 영상으로부터 텍스트를 추출하고 변환하는 기술에 대한 연구를 진행하였다[5]. Yang은 자연 영상으로부터 의미 정보 추출 및 인식 그리고 변환을 수행하는 시스템을 개발하여 다양한 환경에서 서비스 성능을 시험하였다[6]. Ezaki는 디지털 카메라로 획득된 자연 영상으로부터 일본 문자를 검출하여 인식한 후 영어로 번역하는 방법을 연구하였으며[7], Li는 자연 영상으로부터 텍스트를 추출하고 인식하는 연구를 수행하였다[8]. 또한 [4]에서는 간편의 영문자를 인식하여 스페인어로 번역하여 제공해주는 연구를 수행하였다.

Wick는 오인식 에러율을 최소화 하기 위해 OCR(optical character recognition) 인식에서 문맥적 지식의 확률적 표현에 기초한 Viterbi알고리즘(VA)을 적용하였다[9]. [10]과 [11]은 멀티폰트(multifont) OCR 인식을 위해 문자열 정합(string matching) 알고리즘을 적용하였으며, [12]은 글자들의 위치 관계성을 고려하여 틀린 글자를 탐지하고 교정하는 Binary N-gram알고리즘과 Viterbi 알고리즘(VA)을 이용하여 오인식 문제에 접근하였다. Shinghal은 Dictionary Look up(DLA) 알고리즘을 이용하여 사전내에 같은 단어가 있는지를 검사하고, 문자 간의 혼동확률을 측정하여 사전내의 단어와 문자 정합을 시도하였다[13][14]. [15]은 부속 문자열을 처리하기 위한 연구를 수행하였다. 그러나 위와 같

은 오인식 처리 방법들은 오류 교정률이 떨어지며, 오류 교정을 위한 많은 비용이 요구된다.

Bansal은 위와 같은 문제점을 해결하기 위해 사전 검색을 제한하면서 마르코프 가정에 의한 사후 확률이 최대가 되도록 하는 방법에 의해 오인식된 단어를 수정하였다[16]. Shinha는 오인식 수정을 위해 VA와 DLA를 순차적으로 적용하는 예측 교정 알고리즘을 제안하였다[17]. 예측 교정 알고리즘은 예측 단계와 교정단계로 구성되며, 예측 단계는 VA를 이용하여 단어를 찾고 이 단어가 사전에 포함되어 있는지 확인한다. 만약 단어가 사전에 포함되어 있지 않다면 교정단계로 넘어간다. 즉, 예측 교정 알고리즘은 예측 단계에서 VA의 출력력을 확인하고 사전에 없는 경우만 교정단계에 들어간다. 예측 단계에서는 오인식 교정을 위해 많은 시간을 필요로 하지 않지만 교정단계를 거치게 되는 경우 교정을 위한 많은 시간이 소요된다.

Levenshtein 알고리즘[19]은 비교하는 두 문자열의 길이가 서로 달라도 계산이 가능하여, 높은 오인식 교정율을 보임으로써, 맞춤법 검사기, OCR 시스템, 자연어 번역 등 다양한 분야에 사용되고 있다. Takahashi는 단어를 구성하는 문자의 조합과 개별 문자의 후보들을 이용하는 후처리 방법을 제안하였다[18]. 개별 문자의 조합은 단어를 구성하는 문자들의 사용 빈도수를 조사하여 빈도수가 낮은 문자 4개 이하의 문서를 선택하여 문장을 만들 수 있는 조합을 고려하였다. 또한 유사도를 측정하고 후보 단어를 선택하였다. 유사도는 Levenshtein 거리를 이용하였으며, 후보들의 순위를 고려하여 유사도를 측정하였다.

본 논문에서는 휴대폰 카메라를 이용하여 촬영된 간편 영상 인식에서 발생하는 오인식 문제를 해결하기 위해 오인식을 내포하고 있는 인식결과를 교정하는 weighted Disassemble Levenshtein Distance 후처리 방법을 제안한다.

III. wDLD: weighted Disassemble Levenshtein Distance

본 논문에서는 오인식을 교정하기 위해 문자 교정 알

고리즘으로 많이 사용되는 Levenshtein Distance (LD)에 인식 후보 음소 단위 분할 방법과 연산 비용의 가중치를 적용한 weighted Disassemble Levenshtein Distance (wDL) 알고리즘을 제안한다.

1. 상호명 데이터 베이스

본 논문에서는 간판이라는 제한된 영역에서의 텍스트를 인식하기 위하여 전국 상호명 데이터베이스를 구축하였다. 전체 400여만개의 상호명 데이터베이스에서 중복되거나, 혹은 영문자, 숫자가 포함된 데이터를 제외한 130여만개의 상호명 데이터베이스를 구축하였다. 영문자는 A부터 Z까지 26자로 분류되어 인식되며, 한글은 초성 19개, 중성 21개, 종성 28개의 조합으로 이루어져 있다. 본 논문에서는 이중 인식된 한글을 대상으로 오인식 교정을 수행하였다.

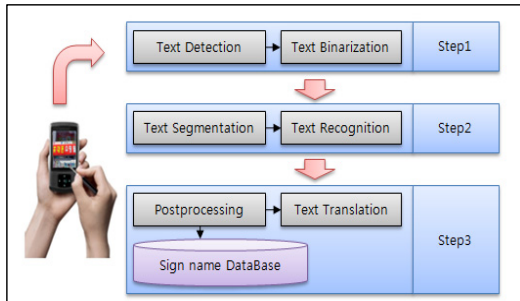


그림 1. 간판영상 문자 인식 시스템

한글의 조합 가능한 모든 글자 수는 11172자이며, 이중 실제 사용되지 않는 글자를 제외하면 글자 수는 2667자이다. 따라서 한글인식은 영문인식에 비해 분류 개수가 상대적으로 많아 인식에 대한 어려움이 있다. 따라서 본 논문에서는 인식 성능 향상을 위해 2667자 중 실제 간판 상호명에서 사용되는 글자로 분류 개수를 줄였다. 전국 상호명 데이터베이스를 분석한 결과 실제 사용된 글자의 개수는 1920자이며, 본 논문에서는 이중 사용 빈도가 낮은 글자를 제외한 총 사용빈도 98%에 해당되는 808자를 실험에 이용하였다.

휴대폰을 이용한 간판영상 문자 인식 시스템은 그림 1에서와 같이 3단계로 구성된다. 첫 번째 단계로는 입

력받은 간판영상에서 텍스트 영역을 찾고, 텍스트 영역에서 글자영역과 배경영역을 이진화 한다. 두 번째는 이진화된 문자영상을 각각의 문자 영상으로 분리하고 촬영 각도에 의해 왜곡된 문자영상을 보정하여 개별 문자 인식기에 입력하게 된다. 개별 문자 인식기에서는 신경망을 통해 분리된 글자 영상을 한 글자씩 인식하게 된다. 세 번째 단계는 인식된 문자를 후처리하여 사용자에게 정보를 제공하게 된다. [그림 2]는 개별 문자 인식기로부터 인식된 입력 값을 보여준다.



그림 2. 인식기로부터 입력된 데이터

력데이터 R 은 간판 영상으로부터 획득된 문자열로서 수식 (1)과 같이 정의 한다.

$$R = \langle R_1, \dots, R_m \rangle \quad (1)$$

여기서, R_1, R_2, \dots, R_m 은 m 개의 음절로 된 상호명을 의미하며, 각각의 음절마다 인식 후보 음절을 갖는다. 따라서 수식 (2)와 같이 후보를 고려한 R_i 을 정의 할 수 있다.

$$R_i = \langle r_{i1}, \dots, r_{in} \rangle \quad (2)$$

예를 들어, n 개의 후보 음절을 갖고 m 개의 음절을 갖는 상호명을 인식하였다면 입력데이터 R 은 다음 수식 (3)과 같은 행렬을 얻을 수 있다.

$$R = \begin{bmatrix} r_{11} & \dots & r_{1n} \\ \vdots & & \vdots \\ r_{m1} & \dots & r_{mn} \end{bmatrix} \quad (3)$$

입력된 R 로부터 가장 유사한 상호명을 찾기 위해 상

호명 데이터베이스를 검색하고 검색된 결과를 D 라고 정의한다. 본 논문에서는 입력 문자열과 가장 유사한 문자열을 찾기 위해 검색된 결과 D 와 입력 문자열 R 사이의 거리 값을 측정한다. 데이터베이스 검색 시, 입력 문자열 R 과 같은 길이의 문자열만을 검색하게 된다면, 문자 분할 과정에서 발생하는 분할 실패나 과 분할된 입력 문자열의 경우 검색된 결과 D 에 정답 문자열이 포함되지 않은 문제점이 발생하게 된다. 따라서 대상 문자열 검색 시 입력 문자열 R 의 길이는 ± 2 까지의 길이를 검색 대상으로 한다.



그림 4. Levenshtein Distance 연산과정

2. 제안된 후처리 방법

본 논문에서 제안된 후처리 방법인 wDL의 기반이 되는 Levenshtein Distance(LD)[19]는 문자열에 대한 거리 값을 산출하기 위해 사용되는 대표적인 거리 척도이다. LD거리 척도는 두 문자열이 주어졌을 때 한 문자열이 다른 문자열로 변환되는데 필요한 연산 비용을 산출한다. LD의 거리 계산은 수식 (4)와 같다.

$$LD(R, D) = \min \begin{cases} d(R_{m-1}, D_{k-1}), & R_m = D_k \\ d(R_{m-1}, D_k) + \alpha, & \text{Insertion} \\ d(R_m, D_{k-1}) + \beta, & \text{Deletion} \\ d(R_{m-1}, D_{k-1}) + \gamma, & \text{Substitution} \end{cases} \quad (4)$$

여기서 α, β, γ 는 삽입, 삭제, 대치 연산에 대한 연산 비용이며, LD에서는 $\alpha = \beta = \gamma = 1$ 로 주어진다. 따라서 그림 3에서와 같이 간판 영상으로부터 인식된 문자열의 거리 값을 계산한다. LD의 연산 과정은 [그림 4]와 같다. LD는 R 의 후보 문자 중에 신뢰도가 가장 높은 1순위의 음절만을 이용하며, 입력문자를 교정할 때 소요되는 연산은 각 삽입, 삭제, 대치 연산으로 보고 각 연산 비용을 모두 1로 계산한다.

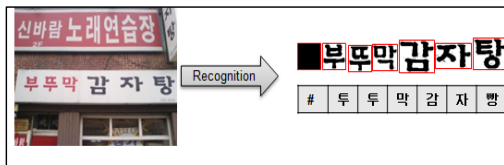


그림 3. 간판 문자 인식

본 논문에서는 한글의 특성을 고려하면서 연산 비용을 최소화 하기 위해 연산 가중치를 적용하고 인식 후보 순위에 음소 단위 분할을 적용한 향상된 wDL (weighted Disassemble Levenshtein Distance)를 제안한다. 한글은 영문과는 다르게 초성, 중성, 종성의 조합으로 이루어져 있으며, 각 음소별로 유사한 형태가 존재한다. 따라서 인식기는 이러한 한글의 구조적 형태로 인해 오인식이 발생하는 경우가 많다.

본 논문에서는 오인식의 형태를 고려하기 위해 입력 데이터의 인식 후보를 음소 단위로 분할하고 분할된 후보 음소의 연산 비용을 최소화 하기위해 연산 비용 가중치를 적용한다. 본 논문에서 제안된 wDL의 경우, 하나의 분할된 음절의 음소는 다른 후보 음절의 분할된 음소와 중복 되는 경우가 발생한다. 따라서 wDL에서는 후보순위가 높고 중복 횟수가 많을수록 신뢰도가 높다고 가정하여 대치 연산비용을 조절한다. 즉, 후보순위가 높고 중복 횟수가 많을수록 연산 비용은 적게 소요된다. 후보 순위에 따른 연산비용을 [표 1]과 같이 정의한다.

표 1. 분할된 음소의 대치 연산비용

입력	후보 음소					
	1	2	3	4	5	없음
R_1	r_{11}	r_{12}	r_{13}	r_{14}	r_{15}	
R_2	r_{21}	r_{22}	r_{23}	r_{24}	r_{25}	
...	
R_m	r_{m1}	r_{m2}	r_{m3}	r_{m4}	r_{m5}	
W	$W_1=1$	$W_2=0.8$	$W_3=0.6$	$W_4=0.4$	$W_5=0.2$	$W=0$

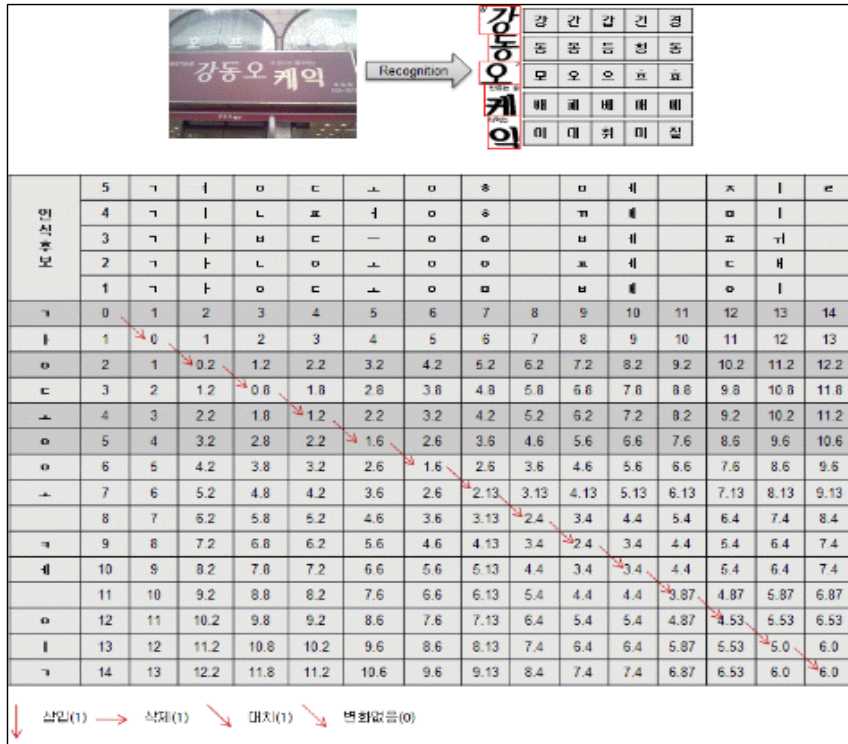


그림 5. 음소단위 분할 후 연산과정

음소를 분해하게 되면 각각의 후보끼리 일치하는 경우가 존재한다. 이런 경우 중복되는 것과 합하여 수식(5)와 같이 연산 비용을 산출한다.

$$W = 1 - \frac{\text{각 중복된 자음의 연산비용의 합}}{\text{전체 후보순위의 연산비용의 합}} \quad (5)$$

각 중복된 자음의 연산비용의 합은 중복된 후보들의 가중치를 더하여 계산되고, 위의 [표 1]에서의 모든 가중치 값을 더하면 전체 후보순위의 연산 비용의 합을 얻을 수 있다. 따라서 수식(5)를 계산하게 되면 음소 가중치가 전체 가중치에서 차지하는 비율을 얻을 수 있다. [그림 5]는 음소 단위로 분할 후 수행되는 wDLD의 연산 과정이다. 산출된 대치 연산 비용은 후보 순위와 중복 횟수가 고려되어 산정된다. 따라서 동일한 중복 횟수가 하더라도 후보 순위에 따라 연산 비용은 다르게 나타난다.

IV. 실험 결과

본 논문에서는 휴대폰 카메라로 촬영한 간판영상에서 텍스트 영역을 검출하여 인식한 후 인식된 결과 중 오인식 된 결과를 제안된 방법을 이용하여 교정하였다. 실험에 사용된 인식 결과는 3가지 인식을 통해 인식된 결과이다. 본 논문에서는 제안된 방법의 효율성을 검증하기 위해 LD와 1순위의 후보만을 음소 단위로 분할한

DLD(Disassemble Levenshtein Distance: DLD), 그리고 입력 데이터의 인식 후보를 음소 단위로 분할하고 가중치를 이용하여 교정하는 wDLD (weighted Disassemble Levenshtein Distance)를 비교 수행하였다. [표 2]는 3가지의 인식을 통해 후처리 없이 인식된 결과이다.

표 2. 3가지 인식기의 인식률

종류	전체	오인식	인식률
A	200	176	22%
B	200	99	50.5%
C	200	64	68%

본 논문에서는 인식기의 성능 별 오인식 교정 알고리즘 성능을 평가하기 위해, 서로 다른 3가지의 인식기를 이용하여 인식 결과를 도출하였고, 그 결과 인식기의 성능은 A<B<C 순으로 나타났다. 제안된 방법의 성능 평가를 위해, 오인식 된 결과를 어느 정도 교정이 이루어졌는지를 나타내는 교정률과, 오인식을 포함한 인식 결과에서 오인식 교정 후 얼마만큼 인식률이 향상되었는지를 나타내는 인식률을 비교하였다. 인식률은 수식 (6)을 통해 산출하였다.

$$\text{인식률} = \frac{\text{정답과 일치하는 간판의 수}}{\text{총 간판의 수}} \times 100 \quad (6)$$

교정률 다음과 같이 수식 (7)를 통해 산출하였다.

$$\text{교정률} = \frac{\text{바르게 교정된 수} - \text{올바른 간판을 잘못 변경한 수}}{\text{오인식된 수}} \times 100 \quad (7)$$

평균 52.7%로 교정 전 인식율에 비해 평균 9.2%의 향상된 결과를 볼 수 있었다 DLD는 1순위의 인식 결과만으로 음소 단위를 분할하여 교정을 수행한 결과, 교정 후 인식률은 평균 76.5%로 나타났다.

표 3. 오인식교정결과

알고리즘	인식기	오인식	교정 전 인식률	성공	실패	교정 후 인식률	향상률	교정률
LD	A	176	12%	34	142	29%	17%	19.32%
	B	99	50.5%	14	85	57.5%	7%	14.14%
	C	64	68%	7	57	71.5%	3.5%	10.93%
	평균	113	43.5%	18.3	94.7	52.7%	9.2%	14.79%
DLD	A	176	12%	67	109	45.5%	33.5%	38.07%
	B	99	50.5%	76	23	88.5%	38%	76.77%
	C	64	68%	55	9	95.5%	27.5%	85.94%
	평균	113	43.5%	66	47	76.5%	33%	66.93%
wDLD	A	176	12%	93	83	58.5%	46.5	52.84%
	B	99	50.5%	86	13	93.5%	43%	86.87%
	C	64	68%	59	5	97.5%	29.5%	92.19%
	평균	113	43.5%	78	35	82.5%	39%	76.54%

그러나 제안된 방법의 경우, 5순위까지의 인식 후보를 음소 단위로 분할하여 교정을 수행한 결과 LD와 DLD에 비해 각각 29.8%와 6% 향상된 결과를 보였다. 그림 6은 오인식 교정 후 인식률에 대한 결과를 그래프로 나타낸 것이다.

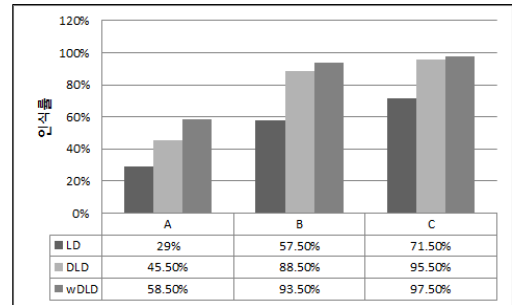


그림 6. 오인식 교정 후 인식률 그래프

wDLD의 경우, A인식기는 LD와 DLD에 비해 29.50%와 13% 더 좋은 성능을 보였으며, A에 비해 인식 성능이 더 좋은 B와 C는 LD의 경우, 36%와 26% 향상된 성능을 보였으며, DLD는 5%와 2% 더 향상된 인식 성능을 보였다. 즉, 전체적으로 wDLD가 LD와 DLD에 비해 더 높은 성능을 보였다. 그림 7은 오인식에 대한 교정률을 보여준다. DLD, wDLD는 인식기의 성능이 향상될수록 교정률도 향상되었지만, LD의 경우 인식기 성능이 향상 될수록 교정률은 감소되어 인식기의 성능 향상이 교정률 향상에는 어떠한 영향도 미치지 않는다는 것을 확인하였다.

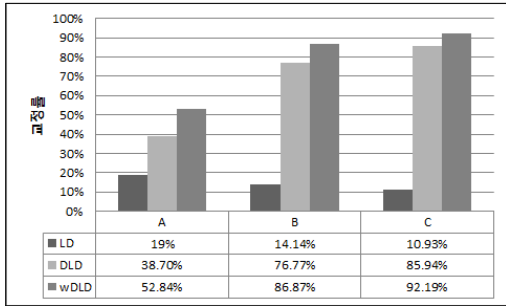


그림 7. 오인식 교정률 그래프

[표 4]는 각 알고리즘을 이용하여 교정한 결과이다. 음소 단위로 분할한 DLD와 wDLD의 경우는 교정 결과의 후보들이 구조적으로 서로 유사하게 나타난다. 후보 순위에 정답과 일치하는 글자가 존재하지 않고 그와 유사한 글자들이 많이 포함되는 경우는 음소 단위로 분할하는 것이 더 좋은 성능을 보였다.

표 4. 오인식 교정 상위 5순위

원본	인식결과	순위	LD	D-LD	iD-LD
삼화페인트	상 회 페 언 토 삼 함 패 인 심 활 할 때 연 살 회 화 개 안 회 회 매 안 돌	1	상남페인트	삼화페인트	삼화페인트
		2	상명페인트	삼호페인트	삼화페인트
		3	상미페인트	삼화페인트	삼호페인트
		4	상아페인트	삼미페인트	삼화페인트
		5	상아탑페인트	상아페인트	삼향페인트
광주상사	광 추 심 사 평 푸 사 썬 투 사 망 루 사 김 무 사 투 성 사 성 서 사 치	1	나주상사	광주상사	광주상사
		2	남광상사	광우상사	광우상사
		3	남주상사	경주상사	경주상사
		4	남주상사	공주상사	공주상사
		5	내광상사	광문상사	성주상사
만민장 의사	감 환 잠 의 사 깔 던 장 의 자 관 민 품 활 의 자 만 든 뜰 찰 역 사 민 든 공 공 먹 차	1	나동장 의사	남면장 의사	감천장 의사
		2	나사렛 의사	남산장 의사	감천장 의사
		3	나주장 의사	남원장 의사	만민장 의사
		4	낙동장 의사	남천장 의사	남면장 의사
		5	낙영장 의사	삼산장 의사	남천장 의사
영광굴비	앞 황 굴 비 양 방 굴 비 양 양 굴 비 항 판 굴 비 평 급 굴 비 색	1	남도굴비	영광굴비	영광굴비
		2	남선굴비	원광굴비	쌍용굴비
		3	남양굴비	이황굴비	중앙굴비
		4	남인굴비	남양굴비	평화굴비
		5	남현굴비	남현굴비	양지굴비
강동오케익	강 등 모 배 이 등 응 폐 대 청 오 호 채 등 등 후 취 호 호 호 매 효 효 효 개 매 매 매 갈	1	낙동강오리	강동오케익	강동오케익
		2	네오페이스	강오토바이	강오토바이
		3	네오페이스	강동오토바이	강동오토바이
		4	대강이앤비	강동꽃배달	강동꽃배달
		5	대구바이오	강동오치과	강서오케이
월드문구	월드문구 월드문구 월드문구 월드문구 월드문구	1	남도문구	월드문구	월드문구
		2	능도문구	미도문구	월드문구
		3	다운문구	랜드문구	미도문구
		4	다운월드	만도문구	월드문구
		5	닥트월드	월드문고	월드문화

V. 결론

본 논문에서는 휴대폰 카메라 기반 간판영상 텍스트 인식 시스템의 인식결과로부터 인식 후보의 음소 단위 분할을 적용하여 간판 상호명 데이터베이스의 후보군들과의 거리 값을 산출하고 가장 유사한 상호명으로 교정하는 방법을 제안하였다. 후보군 데이터베이스 구축을 위해 130만 여개의 전국 상호명 데이터베이스를 구축하였다. 또한 데이터베이스에서 사용되는 한글 문자를 조사하여 인식기의 인식 부류를 2,667자에서 808자로 낮추었다. 성능 평가를 위해 서로 다른 3가지 인식기를 통해 인식된 결과를 LD와 DLD, 그리고 제안된 wDLD를 이용하여 후처리 하였으며, 후처리 후 인식률과 교정률을 측정하였다.

제안된 wDLD는 기존의 LD보다 29.8% 향상된 인식률을 보였으며, 61.75%의 높은 교정률을 보였다. DLD

는 LD에 비해 평균 23.8%의 향상된 인식률과 52.14% 향상된 교정률을 보였지만 wDLD는 DLD에 비해 6% 더 높은 인식률과 9.61% 향상된 성능을 나타냈다. 전체적으로 wDLD가 기존의 방법들의 비해 좋은 오인식 교정 결과를 보였으며, 인식 후보의 음소 단위 분할을 적용 하였을 시 교정률 향상을 가져왔다. 그러나 wDLD는 DLD에 비해 연산과정에서 필요한 메모리 공간이나 연산량이 많이 요구되기 때문에 DLD에 비해 처리 속도는 떨어지게 된다. 따라서 향후 연구로는 이러한 단점을 극복하기 위한 추가적 연구가 필요하다.

참 고 문 헌

- [1] A. Wojciechowski and K. Siek, "Barcode Scanning from Mobile-Phone Camera Photos Delivered Via MMS: Case Study," ER Workshops, pp.218-227, 2008.
- [2] D. M. Chen, S. S. Tsai, R. Vedantham, R. Grzeszczuk, and B. Girod, "Streaming Mobile Augmented Reality on Mobile Phones," ISMAR, pp.181-182, 2009.
- [3] C. Thillou and B. Gosselin, "Natural scene text understanding," Vision Systems, Segmentation and Pattern Recognition, Ch.16, pp.307-333, 2007.
- [4] A. Canedo-Rodriguez, S. H. Kim, J. H. Kim, and Y. Blanco-Fernandez, "English to Spanish translation of signboard images from mobile phone camera," Southeastcon, 2009. SOUTHEASTCON. IEEE , pp.356-361, 2009.
- [5] I. Haritaoglu, "Scene text extraction and translation for handheld devices," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Vol.2, pp.408-413, 2001.
- [6] J. Yang, X. Chen, J. Zhang, Y. Zhang, and A. Waibel, "Automatic Detection and Translation of Text from Natural Scenes," Proceedings of the IEEE 2002 International Conference on Acoustics, Speech and Signal Processing (ICASSP '02), Vol.2, pp.2101-2104, 2002.
- [7] N. Ezaki, M. Bulacu, and L. Schomaker, "Text detection from natural scene images: towards a system for visually impaired persons," Pattern Recognition, ICPR 2004. Proceedings of the 17th International Conference on , Vol.2, pp.683-686, 2004.
- [8] 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.
- [9] M. L. Wick, M. G. Ross, and E. G. Learned-Miller, "Context-Sensitive Error Correction: Using Topic Models to Improve OCR," International Conference Document Analysis and Recognition, Vol.2, pp.1168-1172, 2007.
- [10] W. S. Rosenbaum and J. J. Hilliard, "Multifont OCR Postprocessing System," IBM Journal of Research and Development, Vol.19, No.4, pp.398-421, 1975.
- [11] S. Dobrisek, J. Zibert, N. Pavesic, and F. Mihelic, "An Edit-Distance Model for the Approximate Matching of Timed Strings," Pattern Analysis and Machine Intelligence, IEEE Transactions, Vol.31, No.4, pp.736-741, 2009.
- [12] J. J. Hell and S. N. Srihari, "Experiments in Text Recognition with Binary n-Gram and Viterbi Algorithms," Pattern Analysis and Machine Intelligence, IEEE Transactions, Vol.PAMI-4, No.5, pp.520-530, 1982.
- [13] R. Shinghal and G. T. Toussaint, "Experiments in Text Recognition with the Modified Viterbi Algorithm," Pattern Analysis and Machine Intelligence, IEEE Transactions, Vol.PAMI-1,

No.2, pp.184-193, 1979.

[14] S. W. Kim and Y. Aoki, "A Postprocessing of HANGUL Recognitions Using Dictionary Lookup," JTC-CSCC: Joint Technical Conference on Circuits Systems, Computers and Communications Vol.2, pp.1013-1017, 1993.

[15] R. S. Boyer and J. S. Moore, "A fast string searching algorithm," Comm. of ACM, Vol.20, No.10, pp.762-772, 1977.

[16] V. Bansal and R.M.K. Sinha, "Integrating knowledge sources in Devanagari text recognition system," Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions, Vol.30, No.4, pp.500-505, 2000.

[17] R. M. K. Shinha and B. Prasada, "Visual text recognition through contextual processing," Pattern Recognition, Vol.21, No.5, pp.463-479, 1988.

[18] H. Takashi, N. I. Amano, and A. Yamashita, "A spelling correction method and its application to and OCR system," Pattern Recognition, Vol.23, No.3/4, pp.363-377, 1990.

[19] T. Okuda, E. Tanaka, and T.Kasai, "A Method for the Correction of Garbled Words Based on the Levenshtein Metric," Computers, IEEE Transactions, Vol.C-25, No.2, pp.172-178, 1976.

저 자 소 개

이 명 훈(Myung-Hun Lee)

준회원



- 2008년 2월 : 광주대학교 컴퓨터 공학과(공학사)
- 2011년 2월 : 전남대학교 전자컴퓨터공학부(석사)
- 2011년 ~ 현재 : 코난테크놀로지 연구원

<관심분야> : 데이터마이닝, 정보검색

양 형 정(Hyung-Jung Yang)

정회원



- 1991년 : 전북대학교 전산통계학과(학사)
- 1993년 : 전북대학교 전산통계학과(석사)
- 1998년 : 전북대학교 전산통계학과(박사)

• 2003년 ~ 2005년 : 카네기멜런 대학교 연구원
 • 2005년 ~ 현재 : 전남대학교 전자컴퓨터공학부 부교수
 <관심분야> : e-Design, 데이터 마이닝, e-Learning

김 수 형(Soo-Hyung Kim)

정회원



- 1986년 2월 : 서울대학교 컴퓨터 공학과(공학사)
- 1988년 2월 : 한국과학기술원 전산학과(공학석사)
- 1993년 8월 : 한국과학기술원 전산학과(공학박사)

• 1990년 9월 ~ 1996년 12월 : 삼성전자 멀티미디어연구소(선임연구원)
 • 2000년 12월 ~ 2002년 1월 : 캐나다 Concordia 대학 CENPARMI 연구소 (방문교수)
 • 1997년 ~ 현재 : 전남대학교 전자컴퓨터정보통신공학부 교수
 <관심분야> : 인공지능, 패턴인식, 문서영상 정보검색, 유비쿼터스컴퓨팅

이 귀 상(Guee-Sang Lee)

정회원



- 1980년 : 서울대학교전기공학과 공학사
- 1982년 : 서울대학교대학원 전자계산기공학과(공학석사)
- 1991년 : Pennsylvania 주립대학 전산학(이학박사)

• 1984년 ~ 현재 : 전남대학교 전자컴퓨터공학부 교수
 <관심분야> : 멀티미디어통신, 영상처리및컴퓨터

김 선 희(Sun-Hee Kim)

정회원



- 2003년 : 한국교육개발원 멀티 미디어학(학사)
 - 2006년 : 동국대학교 컴퓨터공학(석사)
 - 2011년 : 전남대학교 전산학과 (박사)
 - 2011년 : 전남대학교 차세대 휴대폰 인터페이스 연구 센터 전임연구원
 - 2012년 ~ 현재 : Carnegie Mellon University 방문연구원
- <관심분야> : 데이터 마이닝, 센서 마이닝, 바이오 데이터 마이닝