



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0055840
(43) 공개일자 2020년05월22일

(51) 국제특허분류(Int. Cl.)
G06K 9/32 (2006.01) G06K 9/00 (2006.01)
(52) CPC특허분류
G06K 9/3258 (2013.01)
G06K 9/00456 (2013.01)
(21) 출원번호 10-2018-0138457
(22) 출원일자 2018년11월12일
심사청구일자 2020년02월13일

(71) 출원인
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
박용주
경기도 고양시 일산동구 탄중로 416,908-403(중산동, 중산마을9단지아파트)
김상윤
서울특별시 은평구 연서로11길 6-6(구산동)
(74) 대리인
박종한

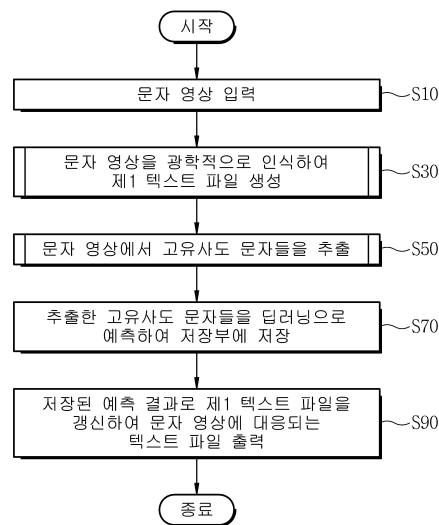
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 선별적 딥러닝을 이용한 문자 인식 장치 및 방법

(57) 요약

본 발명은 선별적 딥러닝을 이용한 문자 인식 장치 및 방법에 관한 것으로, 문자의 인식율을 높이면서 딥러닝에 소요되는 학습 시간과 하드웨어 비용을 절감하기 위한 것이다. 본 발명은 문자 인식을 수행할 문자 영상을 입력 받고, 입력받은 문자 영상에서 문자들을 광학적으로 인식하여 제1 텍스트 파일을 생성하고, 문자 영상에서 고유사도 문자들을 추출하고, 추출한 고유사도 문자들을 딥러닝으로 예측하여 저장부에 저장하고, 제1 텍스트 파일에 포함된 고유사도 문자들을 저장된 예측한 결과로 갱신하여 문자 영상에 대응하는 텍스트 파일을 출력하는 선별적 딥러닝을 이용한 문자 인식 장치 및 방법을 제공한다.

대표도 - 도3



(52) CPC특허분류

G06N 3/08 (2013.01)

G06K 2209/01 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 10063314

부처명 산업부

연구관리전문기관 한국산업기술평가관리원

연구사업명 산업핵심기술개발사업

연구과제명 (R)확장·유연 구조의 조합형 I/O를 지원하는 비전 블록 개발

기 여 율 1/1

주관기관 주식회사 크래비스

연구기간 2016.05.01 ~ 2019.04.30

명세서

청구범위

청구항 1

문자 인식을 수행할 문자 영상을 입력받는 문자 영상 입력부;

상기 문자 영상에서 문자들을 광학적으로 인식하여 제1 텍스트 파일을 생성하는 광학 문자 인식 엔진;

상기 문자 영상에서 고유사도 문자들을 추출하는 문자 추출부;

상기 추출한 고유사도 문자들을 딥러닝으로 예측한 결과를 저장부에 저장하는 딥러닝 엔진; 및

상기 제1 텍스트 파일에 포함된 고유사도 문자들을 상기 저장부에 저장된 예측한 결과로 갱신하여 상기 문자 영상에 대응하는 텍스트 파일을 출력하는 텍스트 파일 출력부;

를 포함하는 선별적 딥러닝을 이용한 문자 인식 장치.

청구항 2

제1항에 있어서, 상기 문자 추출부는,

상기 문자 영상에 포함된 문자들을 개별적으로 분리하여 조각 문자 영상을 생성하고, 생성한 조각 문자 영상들 중에서 고유사도 문자가 포함된 조각 문자 영상들을 추출하고, 추출한 조각 문자 영상들을 상기 딥러닝 엔진에 전달하는 것을 특징으로 하는 선별적 딥러닝을 이용한 문자 인식 장치.

청구항 3

제2항에 있어서, 상기 문자 추출부는,

상기 제1 텍스트 파일에서 고유사도 문자들의 위치를 파악하고,

상기 문자 영상에 포함된 문자들의 배열에 대응되게 상기 조각 문자 영상들을 순서대로 저장하고,

상기 제1 텍스트 파일에서 파악한 상기 고유사도 문자들의 위치 정보를 기반으로 순서대로 저장한 조각 문자 영상들 중 상기 고유사도 문자들의 위치에 해당되는 조각 문자 영상을 추출하고,

추출한 조각 문자 영상들을 상기 딥러닝 엔진에 전달하는 것을 특징으로 하는 선별적 딥러닝을 이용한 문자 인식 장치.

청구항 4

제3항에 있어서, 상기 문자 추출부는,

상기 분리한 조각 문자 영상의 시작점 또는 중심점의 XY 좌표값을 이용하여 상기 분리한 조각 문자 영상을 순서대로 저장하는 것을 특징으로 하는 선별적 딥러닝을 이용한 문자 인식 장치.

청구항 5

제3항에 있어서, 상기 문자 추출부는,

두 부분 이상으로 분리된 문자를 좌표값 유사도 측정 기법으로 하나의 문자로 인식하여 조각 문자 영상을 생성하는 것을 특징으로 하는 선별적 딥러닝을 이용한 문자 인식 장치.

청구항 6

제3항에 있어서, 상기 딥러닝 엔진은,

상기 추출한 조각 문자 영상들에 대한 딥러닝으로 문자를 예측한 결과를 상기 저장부에 저장하되, 상기 예측한 결과에는 상기 추출한 조각 문자 영상들의 위치에 대한 정보를 포함하는 것을 특징으로 하는 선별적 딥러닝을 이용한 문자 인식 장치.

청구항 7

제1항에 있어서, 상기 고유사도 문자는,

i(소문자 아이), I(대문자 아이), l(소문자 엘) 및 1(숫자)로 이루어진 제1 그룹에서 선택된 제1 문자; 및

o(소문자 오), O(대문자 오) 및 0(숫자)로 이루어진 제2 그룹에서 선택된 제2 문자;

중에 적어도 하나를 포함하는 것을 특징으로 하는 선별적 딥러닝을 이용한 문자 인식 장치.

청구항 8

문자 인식을 수행할 문자 영상을 입력받는 단계;

상기 문자 영상에서 문자들을 광학적으로 인식하여 제1 텍스트 파일을 생성하는 단계;

상기 문자 영상에서 고유사도 문자들을 추출하는 단계;

상기 추출한 고유사도 문자들을 딥러닝으로 예측한 결과를 저장하는 단계; 및

상기 제1 텍스트 파일에 포함된 고유사도 문자들을 상기 저장부에 저장된 예측한 결과로 갱신하여 상기 문자 영상에 대응하는 텍스트 파일을 출력하는 단계;

를 포함하는 선별적 딥러닝을 이용한 문자 인식 방법.

발명의 설명

기술 분야

- [0001] 본 발명은 문자 인식 장치 및 방법에 관한 것으로, 더욱 상세하게는 문자 인식 오류가 빈번한 문자에 대해서 선별적인 딥러닝을 수행하여 문자 인식율을 높일 수 있는 선별적 딥러닝을 이용한 문자 인식 장치 및 방법에 관한 것이다.

배경 기술

- [0002] 제조현장 특히, 4차 산업의 중심에 있는 스마트 팩토리 분야에서 생산되는 제품의 이력을 관리하고 추적하기 위해 제품상에 직접 제품에 대한 설명을 각인하고 있으며, 이러한 현상은 점점 증가하고 있고, 이로 인해 제품에 새겨진 문자 정보를 인식하여 관리해야할 필요성도 함께 커지고 있다.
- [0003] 제품에 새겨진 문자 정보를 사람들이 수작업으로 일일이 확인하고 관리할 경우, 번거로울 뿐만 아니라, 시간이 지남에 따라 집중력이 떨어져 정확도에 문제가 발생한다.
- [0004] 이러한 번거로움과 문제점을 줄이기 위해 각인된 문자를 스캔하고, 스캔한 이미지에서 광학 문자 인식 기술(OCR; Optical Character Recognition)을 이용해 문자를 추출하는 방식을 통해 관리하고 있다. 여기서 광학 문자 인식(OCR)이란 스캐너를 통해 입력된 문서 영상에서 문자에 해당하는 부분의 내용을 인식하는 기술을 말한다.
- [0005] 그러나 광학 문자 인식을 이용한 문자 인식에서도 문제점이 존재한다. 예컨대 광학 문자 인식 인식에 있어서, i(소문자 아이), I(대문자 아이), l(소문자 엘), 1(숫자) 그리고 o(소문자 오), O(대문자 오), 0(숫자)와 같이 서로 유사도가 높은 문자의 경우 문자 인식률이 다른 문자에 비해서 상대적으로 낮다.
- [0006] 이러한 문제점을 해결하기 위해서, 광학 문자 인식에 딥러닝 기법을 활용하는 방안이 소개되고 있다. 하지만 기존의 딥러닝 기법은 많은 학습용 데이터와 학습용 데이터에 대한 많은 학습 시간, 학습을 위한 높은 사양의 컴퓨터(하드웨어)가 요구되고 있다.

선행기술문헌

특허문헌

- [0007] (특허문헌 0001) 등록특허공보 제10-1801153호 (2017.11.27. 공고)

발명의 내용

해결하려는 과제

[0008] 따라서 본 발명의 목적은 문자의 인식율을 높이면서 딥러닝에 소요되는 학습 시간과 하드웨어 비용을 절감할 수 있는 선별적 딥러닝을 이용한 문자 인식 장치 및 방법을 제공하는데 있다.

과제의 해결 수단

[0009] 상기 목적을 달성하기 위하여, 본 발명은 문자 인식을 수행할 문자 영상을 입력받는 문자 영상 입력부; 상기 문자 영상에서 문자들을 광학적으로 인식하여 제1 텍스트 파일을 생성하는 광학 문자 인식 엔진; 상기 문자 영상에서 고유사도 문자들을 추출하는 문자 추출부; 상기 추출한 고유사도 문자들을 딥러닝으로 예측한 결과를 저장부에 저장하는 딥러닝 엔진; 및 상기 제1 텍스트 파일에 포함된 고유사도 문자들을 상기 저장부에 저장된 예측한 결과로 갱신하여 상기 문자 영상에 대응하는 텍스트 파일을 출력하는 텍스트 파일 출력부;를 포함하는 선별적 딥러닝을 이용한 문자 인식 장치를 제공한다.

[0010] 상기 문자 추출부는, 상기 문자 영상에 포함된 문자들을 개별적으로 분리하여 조각 문자 영상을 생성하고, 생성한 조각 문자 영상들 중에서 고유사도 문자가 포함된 조각 문자 영상들을 추출하고, 추출한 조각 문자 영상들을 상기 딥러닝 엔진에 전달할 수 있다.

[0011] 상기 문자 추출부는, 상기 제1 텍스트 파일에서 고유사도 문자들의 위치를 파악하고, 상기 문자 영상에 포함된 문자들의 배열에 대응되게 상기 조각 문자 영상들을 순서대로 저장하고, 상기 제1 텍스트 파일에서 파악한 상기 고유사도 문자들의 위치 정보를 기반으로 순서대로 저장한 조각 문자 영상들 중 상기 고유사도 문자들의 위치에 해당되는 조각 문자 영상을 추출하고, 추출한 조각 문자 영상들을 상기 딥러닝 엔진에 전달할 수 있다.

[0012] 상기 문자 추출부는, 상기 분리한 조각 문자 영상의 시작점 또는 중심점의 XY 좌표값을 이용하여 상기 분리한 조각 문자 영상을 순서대로 저장할 수 있다.

[0013] 상기 문자 추출부는, 두 부분 이상으로 분리된 문자를 좌표값 유사도 측정 기법으로 하나의 문자로 인식하여 조각 문자 영상을 생성할 수 있다.

[0014] 상기 딥러닝 엔진은, 상기 추출한 조각 문자 영상들에 대한 딥러닝으로 문자를 예측한 결과를 상기 저장부에 저장하되, 상기 예측한 결과에는 상기 추출한 조각 문자 영상들의 위치에 대한 정보를 포함할 수 있다.

[0015] 상기 고유사도 문자는 i(소문자 아이), I(대문자 아이), l(소문자 엘) 및 1(숫자)로 이루어진 제1 그룹에서 선택된 제1 문자; 및 o(소문자 오), O(대문자 오) 및 0(숫자)로 이루어진 제2 그룹에서 선택된 제2 문자; 중에 적어도 하나를 포함할 수 있다.

[0016] 본 발명은 또한, 문자 인식을 수행할 문자 영상을 입력받는 단계; 상기 문자 영상에서 문자들을 광학적으로 인식하여 제1 텍스트 파일을 생성하는 단계; 상기 문자 영상에서 문자들을 개별적으로 분리하여 조각 문자 영상을 생성하는 단계; 상기 문자 영상에서 고유사도 문자들을 추출하는 단계; 상기 추출한 고유사도 문자들을 딥러닝으로 예측한 결과를 저장하는 단계; 및 상기 제1 텍스트 파일에 포함된 고유사도 문자들을 상기 저장부에 저장된 예측한 결과로 갱신하여 상기 문자 영상에 대응하는 텍스트 파일을 출력하는 단계;를 포함하는 선별적 딥러닝을 이용한 문자 인식 방법을 제공한다.

발명의 효과

[0017] 본 발명에 따른 문자 인식 장치는 문자 영상에서 문자를 인식하는 데 있어서, 전적으로 딥러닝 기법에 의존하지 않고 유사도가 높은 문자들에 대해서만 선별적으로 딥러닝을 수행함으로써, 문자의 인식율을 높이면서 딥러닝에 소요되는 학습 시간과 하드웨어 비용을 절감할 수 있다.

[0018] 즉 문자 영상에서 유사도가 낮은 문자들은 광학 문자 인식 엔진으로 인식하고, 유사도가 높은 문자들은 딥러닝 엔진으로 예측한 후 병합하여 텍스트 파일을 출력함으로써, 문자 인식율을 높일 수 있다.

[0019] 그리고 딥러닝 엔진은 문자 영상에 포함된 전체 문자 중 유사도가 높은 문자들에 대해서만 선별적으로 딥러닝을 수행함으로써, 기존의 문자 영상 전체에 포함된 문자들에 대한 딥러닝을 수행하는 것과 비교해서 딥러닝에 소요되는 학습 시간과 하드웨어 비용을 절감할 수 있다.

도면의 간단한 설명

- [0020] 도 1은 본 발명의 실시예에 따른 선별적 딥러닝을 이용한 문자 인식 장치를 보여주는 블록도이다.
- 도 2는 도 1의 제어부의 세부 구성을 보여주는 블록도이다.
- 도 3은 본 발명의 실시예에 따른 선별적 딥러닝을 이용한 문자 인식 방법을 보여주는 흐름도이다.
- 도 4는 도 3의 제1 텍스트 파일을 생성하는 단계를 보여주는 상세 흐름도이다.
- 도 5는 조각 문자 영상을 생성하고, 도 3의 고유사도 문자들을 추출하는 단계를 보여주는 상세 흐름도이다.
- 도 6은 도 5의 조각 문자 영상을 생성하는 단계에서의 두 부분 이상으로 분리된 문자를 하나의 문자로 인식하기 위해 좌표값 유사도 측정 기법을 설명하기 위한 예시도이다.
- 도 7 및 도 8은 도 5의 생성한 조각 문자 영상들을 저장하는 방법을 설명하기 위한 예시도이다.

발명을 실시하기 위한 구체적인 내용

- [0021] 하기의 설명에서는 본 발명의 실시예를 이해하는데 필요한 부분만이 설명되며, 그 이외 부분의 설명은 본 발명의 요지를 흐트리지 않는 범위에서 생략될 것이라는 것을 유의하여야 한다.
- [0022] 이하에서 설명되는 본 명세서 및 청구범위에 사용된 용어나 단어는 통상적이거나 사전적인 의미로 한정해서 해석되어서는 아니되며, 발명자는 그 자신의 발명을 가장 최선의 방법으로 설명하기 위해 용어의 개념으로 적절하게 정의할 수 있다는 원칙에 입각하여 본 발명의 기술적 사상에 부합하는 의미와 개념으로 해석되어야만 한다. 따라서 본 명세서에 기재된 실시예와 도면에 도시된 구성은 본 발명의 바람직한 실시예에 불과할 뿐이고, 본 발명의 기술적 사상을 모두 대변하는 것은 아니므로, 본 출원시점에 있어서 이들을 대체할 수 있는 다양한 균등물과 변형예들이 있을 수 있음을 이해하여야 한다.
- [0023] 이하, 첨부된 도면을 참조하여 본 발명의 실시예를 보다 상세하게 설명하고자 한다.
- [0024] 도 1은 본 발명의 실시예에 따른 선별적 딥러닝을 이용한 문자 인식 장치를 보여주는 블록도이다. 도 2는 도 1의 제어부의 세부 구성을 보여주는 블록도이다.
- [0025] 도 1 및 도 2를 참조하면, 본 실시예에 따른 문자 인식 장치는 광학 문자 인식을 기반으로 하여 선별적인 딥러닝을 수행하는 장치이다. 문자 인식 장치는 문자 인식을 수행할 문자 영상에 대한 광학 문자 인식 엔진(73)을 적용한 후, 적용 결과에서 인식 오류가 빈번하거나 인식 오류가 높은 문자나 숫자(이하 '고유사도 문자'라 함)에 대해서 추가로 딥러닝 엔진(77)을 적용하여 최종적으로 문자 인식의 정확도를 높인다.
- [0026] 이러한 본 실시예에 따른 문자 인식 장치는 인터페이스부(10), 저장부(30), 표시부(50) 및 제어부(70)를 포함한다. 제어부(70)는 문자 영상 입력부(71), 광학 문자 인식 엔진(73), 문자 추출부(75), 딥러닝 엔진(77) 및 텍스트 파일 출력부(79)를 포함한다.
- [0027] 인터페이스부(10)는 문자 인식을 수행할 문자 영상을 제공하는 문자 영상 제공 장치와 통신을 수행하거나, 사용자로부터 선택 신호를 입력받아 제어부(70)로 전달할 수 있다. 즉 인터페이스부(10)는 통신부, 카메라부 또는 입력부를 포함할 수 있다.
- [0028] 인터페이스부(10)는 통신부를 통하여 문자 영상 제공 장치로부터 문자 영상을 수신하여 제어부(70)로 전달할 수 있다. 인터페이스부(10)는 제어부(70)의 제어에 따라 문자 영상 제공 장치로 문자 영상에 대응하는 텍스트 파일을 전송할 수 있다.
- [0029] 인터페이스부(10)는 카메라부를 통하여 문자 영상을 촬영하여 제어부(70)로 전달할 수 있다. 여기서 카메라부는 피사체의 촬영 기능을 수행하며, 제어부(70)의 제어에 따라서 문자 인식을 수행할 서류를 촬영하여 문자 영상을 생성한다. 카메라부는 문자 영상을 제어부(70)로 전송한다. 카메라부는 이미지 센서, 신호처리부 및 영상처리부를 포함하여 구성될 수 있다. 이미지 센서는 촬영한 영상의 광신호를 아날로그 신호로 변환한다. 신호처리부는 아날로그 신호를 디지털 신호로 변환한다. 그리고 영상처리부는 신호처리부를 통해 입력되는 영상신호를 처리하여 문자 영상을 획득하며, 획득한 문자 영상을 제어부(70)로 전송하거나 저장부(30)에 저장할 수 있다.
- [0030] 인터페이스부(10)는 입력부를 통하여 문자 인식을 수행할 문자 영상을 선택하고, 인식할 문자의 종류를 선택할 수 있다. 인식할 문자의 종류는 문자 인식 장치가 인식할 수 있는 문자의 종류로서, 예컨대 한글, 영어,

일본어, 중국어 등이 될 수 있다. 여기서 입력부는 문자 인식 장치의 조작을 위한 복수의 키를 제공하며, 사용자의 키선택에 따른 선택 신호를 발생하여 제어부(70)로 전달한다. 이러한 입력부로는 버튼, 키패드, 터치패드와 같은 포인팅 장치, 터치스크린 등의 입력장치가 사용될 수 있다.

- [0031] 저장부(30)는 문자 인식 장치의 동작 제어시 필요한 프로그램과, 그 프로그램 수행 중에 발생하는 정보를 저장한다. 저장부(30)는 문자 인식에 필요한 실행프로그램을 저장한다. 실행프로그램은 광학 문자 인식 프로그램과 딥러닝 프로그램을 포함한다. 광학 문자 인식 프로그램으로는 TesseractOCR이 사용될 수 있다.
- [0032] 저장부(30)는 문자 영상과, 문자 영상의 문자 인식을 통하여 생성한 텍스트 파일을 저장한다. 저장부(30)는 문자 영상의 문자 인식의 중간 단계에서 생성되는 제1 텍스트 파일 또는 딥러닝 예측 결과를 저장할 수도 있다. 저장부(30)는 고유사도 문자에 대한 정보를 저장할 수 있다.
- [0033] 여기서 고유사도 문자는 형상적으로 서로 유사하여 광학 문자 인식 시 인식율이 상대적으로 낮은 문자이다. 예컨대 고유사도 문자는 i(소문자 아이), I(대문자 아이), l(소문자 엘) 및 1(숫자)로 이루어진 제1 그룹에서 선택된 적어도 1개의 문자를 포함할 수 있다. 또는 고유사도 문자는 o(소문자 오), O(대문자 오) 및 0(숫자)로 이루어진 제2 그룹에서 선택된 적어도 1개의 문자를 포함할 수 있다.
- [0034] 표시부(50)는 문자 인식 장치에서 실행되는 각종 기능 메뉴를 비롯하여 저장부(30)에 저장된 정보를 표시할 수 있다. 표시부(50)는 문자 인식에 필요한 메뉴, 화면 또는 창을 표시할 수 있다. 표시부(50)로는 LCD, LED, AMOLED 또는 터치스크린이 사용될 수 있다. 터치스크린(touch screen)은 표시장치와 입력장치로서의 역할을 동시에 수행한다.
- [0035] 그리고 제어부(70)는 문자 인식 장치의 전반적인 제어 동작을 수행하는 마이크로프로세서(microprocessor)이다. 제어부(70)는 문자 영상에 대한 광학 문자 인식을 기반으로 하여 선별적인 딥러닝을 수행하여 문자 인식하고, 인식한 문자를 텍스트 파일로 출력한다.
- [0036] 이러한 제어부(70)는 문자 영상 입력부(71), 광학 문자 인식 엔진(73), 문자 추출부(75), 딥러닝 엔진(77) 및 텍스트 파일 출력부(79)를 포함한다. 문자 영상 입력부(71)는 문자 인식을 수행할 문자 영상을 입력받는다. 광학 문자 인식 엔진(73)은 문자 영상에서 문자들을 광학적으로 인식하여 제1 텍스트 파일을 생성한다. 문자 추출부(75)는 문자 영상에서 고유사도 문자들을 추출한다. 딥러닝 엔진(77)은 추출한 고유사도 문자들을 딥러닝으로 예측하고 예측한 결과를 저장부(30)에 저장한다. 그리고 텍스트 파일 출력부(79)는 제1 텍스트 파일에 포함된 고유사도 문자들을 저장된 예측 결과로 갱신하여 문자 영상에 대응하는 텍스트 파일을 출력한다.
- [0037] 여기서 문자 영상 입력부(71)는 문자 영상을 인터페이스부(10) 또는 저장부(30)로부터 입력받는다.
- [0038] 광학 문자 인식 엔진(73)으로는 TesseractOCR 엔진이 사용될 수 있다. TesseractOCR 엔진을 사용하여 제1 텍스트 파일을 생성하는 과정을 설명하면 다음과 같다.
- [0039] 먼저 광학 문자 인식 엔진(73)은 인터페이스부(10)를 통하여 문자 영상에서 인식을 문자의 종류를 선택받는다. 예컨대 문자의 종류는 영어 또는 한글일 수 있다.
- [0040] 다음으로 광학 문자 인식 엔진(73)은 tesseract 객체를 생성한 후 tesseract 객체를 초기화 한다.
- [0041] 다음으로 광학 문자 인식 엔진(73)은 문자 영상에서 문자 인식을 수행할 OCR 영역 지정한다.
- [0042] 다음으로 광학 문자 인식 엔진(73)은 지정된 OCR 영역에 포함된 문자에 대한 인식을 수행한 후, 인식 결과를 UTF-8 타입으로 변환할 수 있다.
- [0043] 여기서 UTF(Unicode-Transformation-Format)는 유니코드 형태의 문자를 변환하기 위한 규격이다. UTF는 가변 길이 문자 인코딩 방식으로, 부호 길이(비트수)에 따라 UTF-8, UTF-16, UTF-32 등이 있다. 이들 모두는 가변 길이 인코딩 방식이기 때문에, 아스키(ASCII) 코드의 길이에 맞춰 호환이 가능하다.
- [0044] 대부분의 텍스트는 UTF-8로 저장되고 있는데, UTF-8 인코딩은 문자열을 8bit 기반으로 저장한다. 즉 영문과 숫자 기호는 1byte로, 한글과 한자 등은 3byte로 가변적으로 저장하는 방식으로 전 세계 모든 글자들을 한꺼번에 표현할 수 있다.
- [0045] UTF-8이 8-bit를 기반으로 하듯이, UTF-16은 16-bit 기반으로 문자열을 저장한다. 예컨대 'A'라는 문자 표현의 경우 UTF-8은 1byte로 표현하고, UTF-16은 2byte로 표현하며, UTF-32는 4byte로 표현한다.
- [0046] 그리고 광학 문자 인식 엔진(73)은 변환 결과를 제1 텍스트 파일로 저장한다. 제1 텍스트 파일의 확장자는

".txt"일 수 있다.

- [0047] 문자 추출부(75)는 문자 영상에 포함된 문자들을 개별적으로 분리하여 조각 문자 영상을 생성하고, 생성한 조각 문자 영상들 중에서 고유사도 문자가 포함된 조각 문자 영상들을 추출하고, 추출한 조각 문자 영상들을 딥러닝 엔진(77)에 전달한다. 즉 문자 추출부(75)는 제1 텍스트 파일에서 고유사도 문자들의 위치를 파악한다. 문자 추출부(75)는 문자 영상에 포함된 문자들의 배열에 대응되게 조각 문자 영상들을 순서대로 저장한다. 문자 추출부(75)는 제1 텍스트 파일에서 파악한 고유사도 문자들의 위치 정보를 기반으로 순서대로 저장한 조각 문자 영상들 중 고유사도 문자들의 위치에 해당되는 조각 문자 영상을 추출한다. 그리고 문자 추출부(75)는 추출한 조각 문자 영상들을 상기 딥러닝 엔진(77)에 전달한다.
- [0048] 문자 추출부(75)는 두 부분 이상으로 분리된 문자를 좌표값 유사도 측정 기법으로 하나의 문자로 인식하여 조각 문자 영상을 생성할 수 있다. 예컨대 두 부분 이상으로 분리된 문자는 영문자의 경우, i, j가 해당될 수 있다. 이에 대한 설명은 도 6을 참조로 설명에서 후술하도록 하겠다.
- [0049] 문자 추출부(75)는 분리한 조각 문자 영상의 시작점 또는 중심점의 xy 좌표값을 이용하여 분리한 조각 문자 영상을 순서대로 저장할 수 있다. 이에 대한 설명은 도 7 및 도 8을 참조로 설명에서 후술하도록 하겠다.
- [0050] 딥러닝 엔진(77)은 추출한 조각 문자 영상들에 대한 딥러닝으로 문자를 예측하여 예측한 결과를 저장부(30)에 저장한다.
- [0051] 그리고 텍스트 파일 출력부(79)는 제1 텍스트 파일에 포함된 고유사도 문자들을 저장부(30)에 저장된 예측한 문자들로 갱신하여 문자 영상에 대응하는 텍스트 파일을 출력한다. 텍스트 파일 출력부(79)는 텍스트 파일을 표시부(50)를 통하여 표시하거나, 저장부(30)에 저장할 수 있다.
- [0052] 이와 같이 본 실시예에 따른 문자 인식 장치는 문자 영상에서 문자를 인식하는데 있어서, 전적으로 딥러닝 기법에 의존하지 않고 고유사도 문자들에 대해서만 선별적으로 딥러닝을 수행함으로써, 문자의 인식율을 높이면서 딥러닝에 소요되는 학습 시간과 하드웨어 비용을 절감할 수 있다.
- [0053] 즉 문서 영상에서 유사도가 낮은 저유사도 문자들은 광학 문자 인식 엔진(73)으로 인식하고, 고유사도 문자들은 딥러닝 엔진(77)으로 예측한 후 병합하여 텍스트 파일을 출력함으로써, 문자 인식율을 높일 수 있다.
- [0054] 그리고 딥러닝 엔진(77)은 문자 영상에 포함된 전체 문자 중 고유사도 문자들에 대해서만 선별적으로 딥러닝을 수행함으로써, 기존의 문자 영상 전체에 포함된 문자들에 대한 딥러닝을 수행하는 것과 비교해서 딥러닝에 소요되는 학습 시간과 하드웨어 비용을 절감할 수 있다.
- [0055] 이와 같은 본 실시예에 따른 문자 인식 장치를 이용한 문자 인식 방법에 대해서 도 3 내지 도 5를 참조하여 설명하면 다음과 같다. 여기서 도 3은 본 발명의 실시예에 따른 선별적 딥러닝을 이용한 문자 인식 방법을 보여주는 흐름도이다. 도 4는 도 3의 제1 텍스트 파일을 생성하는 단계를 보여주는 상세 흐름도이다. 그리고 도 5는 조각 문자 영상을 생성하고, 도 3의 고유사도 문자들을 추출하는 단계를 보여주는 상세 흐름도이다.
- [0056] 먼저 S10단계에서 문자 인식 장치는 문자 인식을 수행할 문자 영상을 입력받는다. 다음으로 S30단계에서 문자 인식 장치는 문자 영상에서 문자들을 광학적으로 인식하여 제1 텍스트 파일을 생성한다. 다음으로 S50단계에서 문자 인식 장치는 문자 영상에서 고유사도 문자들을 추출한다. 이어서 S70단계에서 문자 인식 장치는 추출한 고유사도 문자들을 딥러닝으로 예측하여 저장부에 저장한다. 그리고 S90단계에서 문자 인식 장치는 제1 텍스트 파일에 포함된 고유사도 문자들을 저장된 예측 결과로 갱신하여 문자 영상에 대응하는 텍스트 파일을 출력한다.
- [0057] 이때 S10단계에서 문자 인식 장치는 인터페이스부(10)를 통하여 문자 영상에서 인식을 문자의 종류를 선택받는다. 예컨대 문자의 종류는 영어 또는 한글일 수 있다.
- [0058] S30단계는 도 4에 도시된 바와 같이 수행될 수 있다. 먼저 S31단계에서 문자 인식 장치는 tesseract 객체를 생성한 후, S33단계에서 tesseract 객체를 초기화 한다.
- [0059] 다음으로 S35단계에서 문자 인식 장치는 문자 영상에서 문자 인식을 수행할 OCR 영역 지정한다.
- [0060] 다음으로 S37단계에서 문자 인식 장치는 지정된 OCR 영역에 포함된 문자에 대한 인식을 수행한 후, S39단계에서 인식 결과를 UTF-8 타입으로 변환할 수 있다. 이때 문자 인식 장치는 UTF-8 타입 이외에 UTF-16 또는 UTF-32 타입을 사용할 수 있다.
- [0061] 그리고 S41단계에서 문자 인식 장치는 변환 결과를 제1 텍스트 파일로 저장한다.

- [0062] S50단계는 도 5에 도시된 바와 같이 수행될 수 있다. 먼저 S51단계에서 문자 인식 장치는 제1 텍스트 파일에서 고유사도 문자들의 위치를 파악한다.
- [0063] 다음으로 S53단계에서 문자 인식 장치는 입력된 문자 영상을 전처리한다.
- [0064] 다음으로 S55단계에서 문자 인식 장치는 문자 영상에 포함된 문자들을 개별적으로 분리하여 조각 문자 영상을 생성한다.
- [0065] 다음으로 S57단계에서 문자 인식 장치는 생성한 조각 문자 영상들을 문자 영상에 포함된 문자들의 배열에 대응되게 순서대로 저장한다.
- [0066] 이어서 S59단계에서 문자 인식 장치는 파악한 고유사도 문자들의 위치에 대응되는 조각 문자 영상들을 추출한다. 즉 문자 인식 장치는 제1 텍스트 파일에서 파악한 고유사도 문자들의 위치 정보를 기반으로 순서대로 저장한 조각 문자 영상들 중 고유사도 문자들의 위치에 해당되는 조각 문자 영상을 추출한다.
- [0067] 그리고 S61단계에서 문자 인식 장치는 추출한 조각 문자 영상들을 딥러닝 엔진(77)에 전달한다.
- [0068] 이때 본 실시예에서는 S51단계 이후에 S53단계를 수행하는 예를 개시하였지만 이것에 한정되는 것은 아니다. 예컨대 S51단계와 S53단계는 병렬적으로 수행될 수도 있고, S53단계 이후에 S51단계를 수행할 수 있다. 물론 S51단계는 S59단계 이전에 수행한다.
- [0069] 한편 S55단계에서 문자 영상에 포함된 문자들을 개별적으로 분리하여 조각 문자 영상을 생성할 때, 도 6에 도시된 바와 같이, 두 부분 이상으로 분리된 문자를 좌표값 유사도 측정 기법으로 하나의 문자로 인식하여 조각 문자 영상을 생성할 수 있다. 여기서 도 6은 도 5의 조각 문자 영상을 생성하는 단계에서의 좌표값 유사도 측정 기법을 설명하기 위한 예시도이다.
- [0070] 도 6을 참조하면, 문자 영상을 조각 문자 영상으로 분리할 때, i 와 j 같이 영역이 분리되어 있는 문자의 경우, 영상 처리를 할 경우 하나의 문자로 바로 분리되지 않는다. 따라서 본 실시예에서는 별도로 제안하는 좌표값 유사도 측정 기법을 적용하여 두 부분 이상으로 분리된 문자를 하나의 문자로 인식하여 조각 문자 영상을 생성한다.
- [0071] 좌표값 유사도 측정 기법은 두 부분 이상으로 분리된 조각 문자 영상의 x 좌표값을 이용한다. 도 6의 (a)를 참조하면, I 영역과 II영역은 각각 직사각형 형태로, I 영역은 시작점(a,b)와 끝점(g,h)로 표시할 수 있고, II영역은 시작점(c,d)와 끝점(e,f)로 표시할 수 있다.
- [0072] I 영역과 II영역의 좌표값을 비교하여 x 좌표값이 비슷한 영역을 하나의 영역으로 판단하여 하나의 문자로 인식하도록 한다. 즉 i 의 경우, I 영역의 시작점(a,b)와 II영역의 시작점(c,d)에서 x 좌표값인 a 값과 c 값이 유사하다. 물론 I 영역의 끝점(g,h)와 II영역의 끝점(e,f)에서 x 좌표값인 g 값과 e 값이 유사하다.
- [0073] 이러한 문자에 대해서는, 도 6의 (b)에 도시된 바와 같이, 시작점이 (γ, ι)이고 끝점이 (ς, κ)인 새로운 III영역을 지정하여 하나의 조각 문자 영상으로 저장한다. 이때 III영역의 시작점은 (γ, ι)=(c,b)이고, III영역의 끝점은 (ς, κ)=(e,f)이다.
- [0074] S57단계에서 생성한 조각 문자 영상들을 문자 영상에 포함된 문자들의 배열에 대응되게 순서대로 저장할 때, 도 7 및 도 8에 도시된 바와 같이, 분리한 조각 문자 영상의 XY 좌표값을 이용하여 분리한 조각 문자 영상을 순서대로 저장할 수 있다. 여기서 도 7 및 도 8은 도 5의 생성한 조각 문자 영상들을 저장하는 방법을 설명하기 위한 예시도이다. 도 7 및 도 8에서 위는 문자 영상의 문자열을 보여주고 있고, 아래는 조각 문자 영상의 저장 순서를 나타낸다. 도 7 및 도 8에서 위는 순서대로 저장하기 전의 contours를 찾아 개별적으로 분리하여 저장된 조각 문자 영상을 보여주고 있고, 아래는 조각 문자 영상이 순서대로 저장된 배열을 나타낸다.
- [0075] 문자 영상에 대해서 contours를 찾아 분리하는 방법을 사용할 경우, 도 7에 도시된 바와 같이, 문자 영상의 문자열과 저장되는 조각 문자 영상의 순서가 상이하다. 이 경우 저장한 조각 문자 영상을 문자 영상의 문자열에 대응되게 재배열해야 한다.
- [0076] 도 7에 도시된 바와 같이, 문자 영상의 문자열들을 문자별로 조각 문자 영상으로 저장할 때, 순서대로 저장하기 위해서 각 조각 문자 영상에 대한 XY 좌표값을 이용하여 순서대로 저장한다.
- [0077] 이때 X축 방향으로 위치하는 문자열의 조각 문자 영상은 시작점 또는 중심점의 x 좌표값을 이용하여 순서대로 저장한다. Y축 방향으로 위치하는 문자열은 Y 좌표값을 이용하여 순서대로 저장한다. 이와 같이 조각 문자 영상

은 조각 문자 영상은 시작점 또는 중심점의 XY 좌표값을 이용하여 순서대로 저장한다.

[0078] 본 실시예에서는 문자 영상의 문자열에 대응되게 조각 문자 영상들을 순서대로 저장하기 위해서, 조각 문자 영상들의 시작점의 x 좌표값 또는 중심점의 x 좌표값을 비교하여 빠른 순서대로 배치하는 방법을 사용하였다.

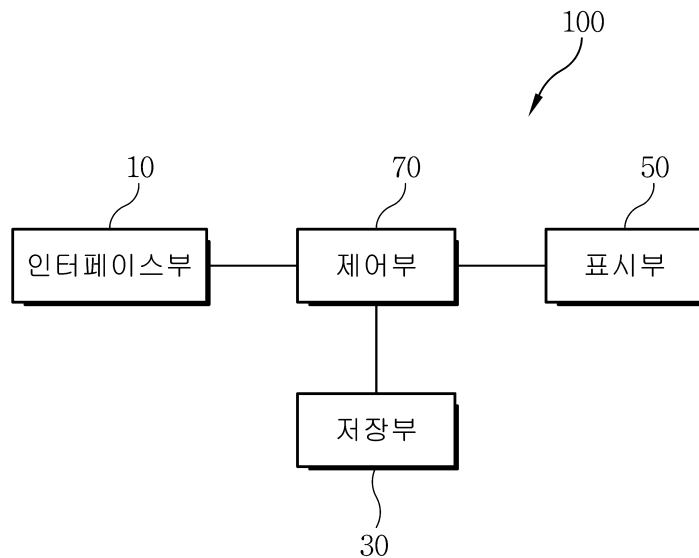
[0079] 한편, 본 명세서와 도면에 개시된 실시예들은 이해를 돕기 위해 특정 예를 제시한 것에 지나지 않으며, 본 발명의 범위를 한정하고자 하는 것은 아니다. 여기에 개시된 실시예들 이외에도 본 발명의 기술적 사상에 바탕을 둔 다른 변형예들이 실시 가능하다는 것은, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게는 자명한 것이다.

부호의 설명

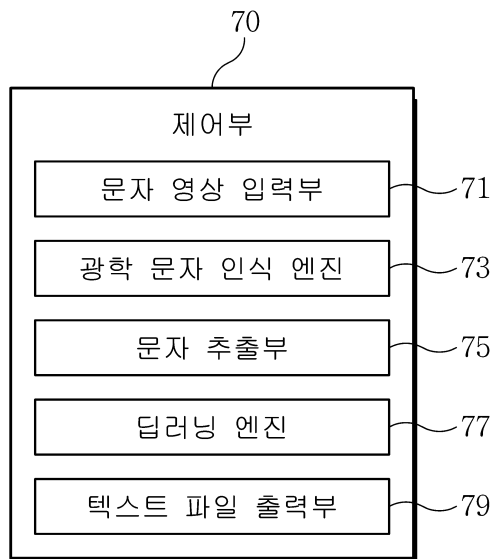
[0080] 10 : 인터페이스부
30 : 저장부
50 : 표시부
70 : 제어부
71 : 문자 영상 입력부
73 : 광학 문자 인식 엔진
75 : 문자 추출부
77 : 딥러닝 엔진
79 : 텍스트 파일 출력부
100 : 문자 인식 장치

도면

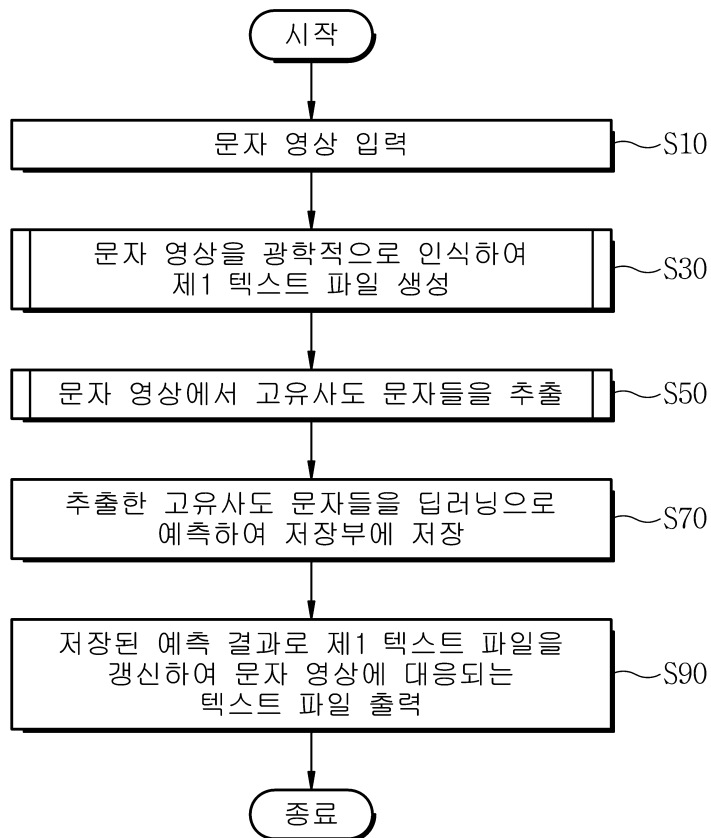
도면1



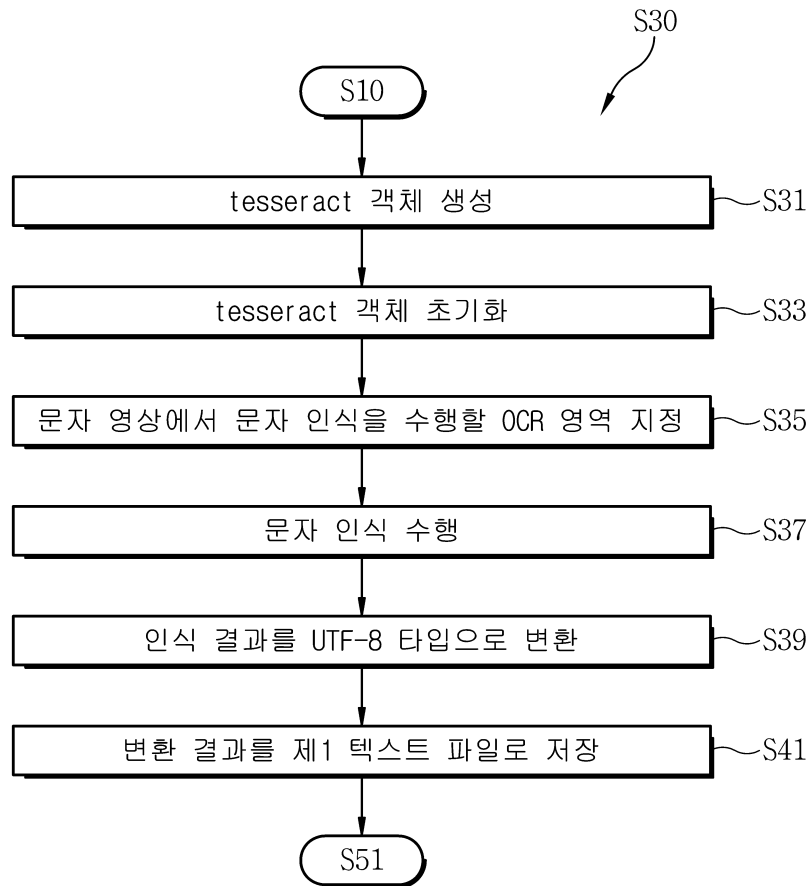
도면2



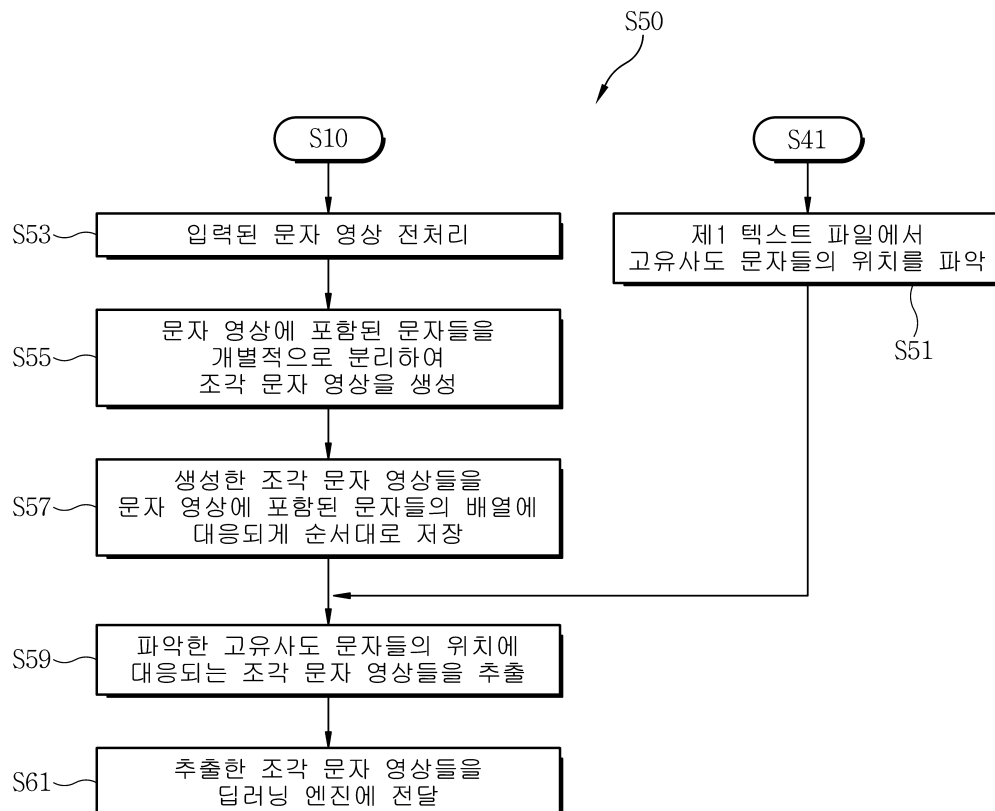
도면3



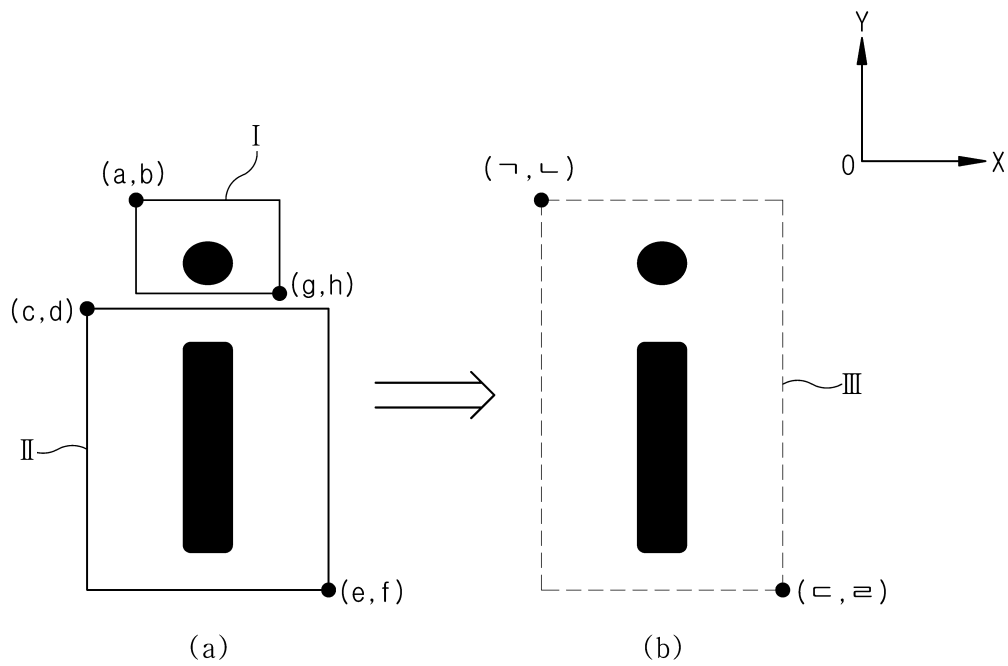
도면4



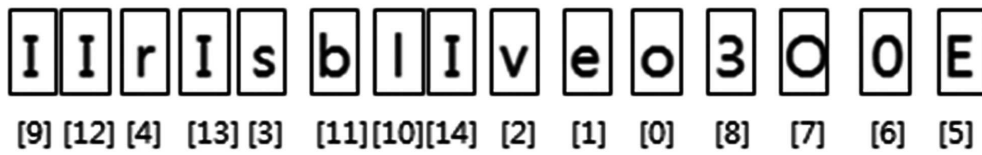
도면5



도면6



도면7



도면8

