(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년12월15일
(11) 등록번호 10-2613984
(24) 등록일자 2023년12월11일

(51) 국제특허분류(Int. Cl.)
G06V 30/26 (2022.01) *G06F 40/237* (2020.01)
G06F 40/295 (2020.01) *G06N 3/08* (2023.01)
G06V 20/17 (2022.01) *G06V 30/14* (2022.01)
G06V 30/164 (2022.01)
(52) CPC특허분류
G06V 30/26 (2023.08)
G06F 40/237 (2022.01)
(21) 출원번호 10-2023-0035717
(22) 출원일자 2023년03월20일
심사청구일자 2023년03월20일
(30) 우선권주장
1020220184082 2022년12월26일 대한민국(KR)
(56) 선행기술조사문헌
KR1020220086336 A*
KR102472447 B1*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
고려대학교산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
김현우
서울시 성북구 북악산로 813 (정릉동, 정릉우성아파트) 101동 809호
이소진
서울특별시 도봉구 덕릉로53길 10 (창동 대봉아트빌) 302호
(뒷면에 계속)
(74) 대리인
특허법인위솔

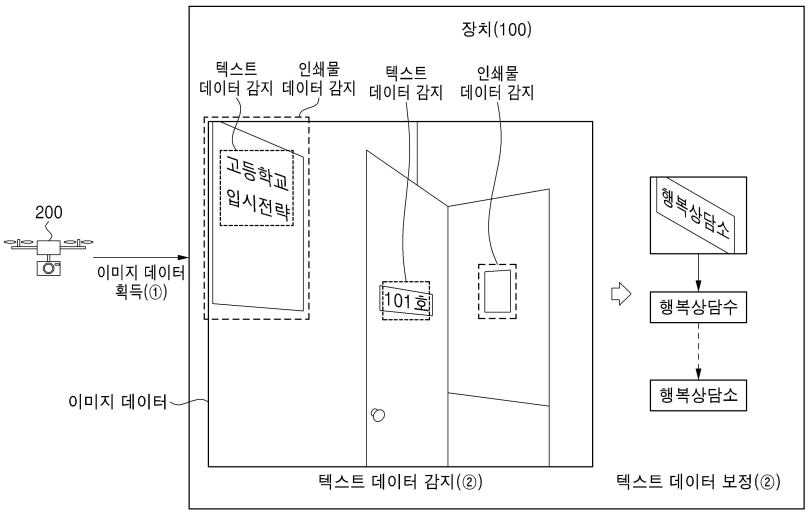
전체 청구항 수 : 총 8 항 심사관 : 강현일

(54) 발명의 명칭 **실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법 및 이를 위한 장치**

(57) 요 약

본 발명의 일 실시 예에 따른 프로세서 및 메모리를 포함하는 장치가 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식하는 방법은 (a) 무인 이동체로부터 이미지 데이터를 획득하는 단계, (b) 기 획득한 이미지 데이터에 포함된 텍스트 데이터를 인식하는 단계, (c) 상기 인식한 텍스트 데이터 내에 오탈자가 있는 경우, 상기 텍스트 데이터를 보정하는 단계 및 (d) 상기 보정한 텍스트 데이터를 기반으로 상기 이미지 데이터에 대한 위치 후보군을 선정하는 단계를 포함한다.

대 표 도



(52) CPC특허분류
G06F 40/295 (2020.01)
G06N 3/08 (2023.01)
G06V 20/17 (2023.08)
G06V 30/1444 (2023.08)
G06V 30/164 (2023.08)

(72) 발명자
주민석
서울특별시 강동구 고덕로80길 134 고덕숲아이파크 109동 1501호

이승훈
경기도 의정부시 평화로202번길 20 신일유토빌플러스아파트 110동 2203호

이 발명을 지원한 국가연구개발사업

과제고유번호	1711193663
과제번호	2020-0-01819-004
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성(R&D)
연구과제명	ICT명품인재양성(고려대학교)
기 여 율	1/1
과제수행기관명	고려대학교산학협력단
연구기간	2023.01.01 ~ 2023.12.31

명 세 서

청구범위

청구항 1

프로세서 및 메모리를 포함하는 장치가 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법에 있어서,

(a) 무인 이동체로부터 이미지 데이터를 획득하는 단계;

(b) 상기 획득한 이미지 데이터에 포함된 텍스트 데이터를 인식하는 단계;

(c) 상기 인식한 텍스트 데이터 내에 오탈자가 있는 경우, 상기 텍스트 데이터를 보정하는 단계; 및

(d) 상기 보정한 텍스트 데이터를 기반으로 상기 이미지 데이터에 대한 위치 후보군을 선정하는 단계;

를 포함하는,

실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법에 있어서,

상기 (b) 단계는,

(b-1) 상기 이미지 데이터에 포함된 모든 텍스트 데이터를 검출하여 상기 검출한 텍스트 데이터 각각을 둘러싼 텍스트 영역을 지정 및 산출하는 단계;

(b-2) 상기 검출한 텍스트 영역들 중 적어도 둘 이상의 텍스트 영역들 간의 거리가 기 설정된 거리 이하로 근접한 경우, 상기 거리가 기 설정된 거리 이하로 근접한 텍스트 영역들이 인쇄물 데이터 내에 포함되어 있는 텍스트 영역이라고 판단하는 단계;

(b-3) 상기 인쇄물 데이터에 포함되어 있는 텍스트 영역 내의 텍스트 데이터를 필터링 하는 단계; 및

(b-4) 상기 이미지 데이터에 포함된 텍스트 영역의 텍스트 데이터들을 인식하는 단계;

를 포함하는,

실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법.

청구항 2

제1 항에 있어서,

상기 (a) 단계 이후에,

(b′) 상기 이미지 데이터 내에 텍스트 데이터가 포함되어 있지 않는 경우, 상기 무인 이동체로부터 후속적인 이미지 데이터를 획득하는 단계;

를 포함하는,

실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법.

청구항 3

삭제

청구항 4

제1항에 있어서,

상기 (b-4) 단계는,

상기 이미지 데이터를 광학 문자 인식 인공지능 모델에 입력하여 상기 이미지 데이터에 포함된 텍스트 영역의 텍스트 데이터를 인식하는 단계인,

실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법.

청구항 5

제1 항에 있어서,

상기 (b) 단계 이후에,

(c′) 상기 인식한 텍스트 데이터 내에 오탈자가 없는 경우, 상기 텍스트 데이터를 기반으로 상기 이미지 데이터에 대한 위치 후보군을 선정하는 단계;

를 포함하는,

실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법.

청구항 6

제1 항에 있어서,

상기 (c) 단계는,

상기 텍스트 데이터에 포함된 단어들을 Symspell 알고리즘을 기반으로 보정하되, 상기 텍스트 데이터에 포함된 단어와 기 설정된 편집 거리 내에 해당하는 적어도 하나 이상의 보정 후보 단어- 상기 보정 후보 단어는, 상기 장치 내의 말뭉치 데이터에 저장되어 있는 데이터임-를 매칭하고, 상기 보정 후보 단어들 중 실내 장소명과 근접한 단어를 추출하여 보정하는 단계인,

실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법.

청구항 7

제1항에 있어서,

상기 (d) 단계 이후에,

(e) 상기 선정된 위치 후보군 중 가장 많이 선정된 위치 후보군을 상기 이미지 데이터에 대한 최종 위치 후보군으로 출력하는 단계;

를 더 포함하는,

실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법.

청구항 8

하나 이상의 프로세서;

네트워크 인터페이스;

상기 프로세서에 의해 수행되는 컴퓨터 프로그램을 로드(Load)하는 메모리; 및

대용량 네트워크 데이터 및 상기 컴퓨터 프로그램을 저장하는 스토리지를 포함하되,

상기 컴퓨터 프로그램은 상기 하나 이상의 프로세서에 의해,

(A) 사용자 단말로부터 단서 정보 및 무인 이동체로부터 실시간 탐지 정보 중 어느 하나 이상을 획득하는 프로세스;

(B) 상기 획득한 이미지 데이터에 포함된 텍스트 데이터를 인식하는 프로세스;

(C) 상기 인식한 텍스트 데이터 내에 오탈자가 있는 경우, 상기 텍스트 데이터를 보정하는 프로세스; 및

(D) 상기 보정한 텍스트 데이터를 기반으로 상기 이미지 데이터에 대한 위치 후보군을 선정하는 프로세스;

를 실행하는 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 장치에 있어서,

상기 (B) 프로세스는,

(B-1) 상기 이미지 데이터에 포함된 모든 텍스트 데이터를 검출하여 상기 검출한 텍스트 데이터 각각을 둘러싼 텍스트 영역을 지정 및 산출하는 프로세스;

(B-2) 상기 검출한 텍스트 영역들 중 적어도 둘 이상의 텍스트 영역들 간의 거리가 기 설정된 거리 이하로 근접한 경우, 상기 거리가 기 설정된 거리 이하로 근접한 텍스트 영역들이 인쇄물 데이터 내에 포함되어 있는 텍스트 영역이라고 판단하는 프로세스;

(B-3) 상기 인쇄물 데이터에 포함되어 있는 텍스트 영역 내의 텍스트 데이터를 필터링 하는 프로세스; 및

(B-4) 상기 이미지 데이터에 포함된 텍스트 영역의 텍스트 데이터들을 인식하는 프로세스;

를 포함하는,

실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 장치.

청구항 9

컴퓨팅 장치와 결합하여,

(AA) 무인 이동체로부터 이미지 데이터를 획득하는 단계;

(BB) 상기 획득한 이미지 데이터에 포함된 텍스트 데이터를 인식하는 단계;

(CC) 상기 인식한 텍스트 데이터 내에 오탈자가 있는 경우, 상기 텍스트 데이터를 보정하는 단계; 및

(DD) 상기 보정한 텍스트 데이터를 기반으로 상기 이미지 데이터에 대한 위치 후보군을 선정하는 단계;

를 실행하는 매체에 저장된 컴퓨터 프로그램에 있어서,

상기 (BB) 단계는,

(BB-1) 상기 이미지 데이터에 포함된 모든 텍스트 데이터를 검출하여 상기 검출한 텍스트 데이터 각각을 둘러싼 텍스트 영역을 지정 및 산출하는 단계;

(BB-2) 상기 검출한 텍스트 영역들 중 적어도 둘 이상의 텍스트 영역들 간의 거리가 기 설정된 거리 이하로 근접한 경우, 상기 거리가 기 설정된 거리 이하로 근접한 텍스트 영역들이 인쇄물 데이터 내에 포함되어 있는 텍스트 영역이라고 판단하는 단계;

(BB-3) 상기 인쇄물 데이터에 포함되어 있는 텍스트 영역 내의 텍스트 데이터를 필터링 하는 단계; 및

(BB-4) 상기 이미지 데이터에 포함된 텍스트 영역의 텍스트 데이터들을 인식하는 단계;

를 포함하는,

매체에 저장된 컴퓨터 프로그램.

발명의 설명

기 술 분 야

[0001] 본 발명은 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법 및 이를 위한 장치에 관한 것이다. 보다 자세하게는 무인 이동체로부터 획득한 이미지 데이터에 포함된 텍스트 데이터를 인식하고, 인식한 텍스트 데이터에 오탈자가 있는 경우, 텍스트 데이터를 보정하여 보정한 텍스트 데이터를 기반으로 실내 장소명을 예측하는 방법 및 이를 위한 장치에 관한 것이다.

배 경 기 술

[0003] 텍스트 인식 기술은 문서 또는 이미지 내에 포함된 텍스트를 단순히 검출할 뿐만 아니라 검출한 텍스트의 내용이 무엇인지 인식할 수 있는 기술이다.

[0004] 이러한 기술은 사용자들이 제출한 문서의 오류(예. 오탈자, 미기재 등)를 감지하여 수작업으로 문서의 오류를 판단하는 업무의 효율을 향상시켰을 뿐만 아니라, 구조를 요청한 사용자가 촬영한 이미지를 통하여 구조 현장의 단서(텍스트)를 확보하는 것에도 기여하였다.

[0005] 다만, 텍스트 인식 기술은 밝기, 진동에 의한 노이즈 등 여러 방해 요인에 의하여 아직까지 텍스트를 정확히 인식하지 못하고 있으며, 정확하지 않은 텍스트를 그대로 출력하기에 오히려 업무의 혼선을 가져오는 문제점을 발생하기도 하였다.

[0006] 예를 들어, 이미지에 포함된 한글 문자를 초성, 중성, 종성별로 인식하여 생성한 각 후보군에 대한 정답 확률을 산출한 후 최대 정답 확률을 갖는 초성, 중성, 종성 후보를 조합하는 시스템이 개시되었지만(대한민국 등록특허공보 제 10-2215593 호), 문자 인식 모델을 통한 문자 인식에서 발생한 오류를 보정하기 위해 확률을 이용하되 입력으로 사용된 이미지에 포함된 문자가 촬영 각도나 밝기에 의해 훼손되어 있을 경우 스스로 보정할 수 없다는 문제점이 발생하였다.

[0007] 본 발명은 이와 같은 문제점들에 착안하여 도출된 것으로써, 단순히 이미지에 포함되어 있는 텍스트를 인식하는 것에 그치지 않고, 인식한 텍스트의 오탈자를 감지하고, 오탈자가 포함된 텍스트를 자진하여 보정해주는 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법 및 이를 위한 장치를 제안하고자 한다.

선행기술문헌

특허문헌

[0009] (특허문헌 0001) 대한민국 등록특허공보 제 10-2215593 호(2021.02.05)

발명의 내용

해결하려는 과제

[0010] 본 발명이 해결하고자 하는 기술적 과제는 무인 이동체로부터 획득한 이미지 데이터 내에 포함된 텍스트 데이터를 인식할 뿐만 아니라 인식한 텍스트 데이터를 자진하여 보정해주는 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법 및 이를 위한 장치를 제공하는 것이다.

[0011] 본 발명이 해결하고자 하는 또 다른 기술적 과제는 무인 이동체로부터 획득한 이미지 데이터를 획득하였을 시, 이미지 데이터 내에서 실내 장소명 예측에 불필요한 텍스트 데이터(예. 인쇄물 데이터 내의 텍스트 데이터)를 필터링하여 정확한 실내 장소 예측명을 선정하는 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법 및 이를 위한 장치를 제공하는 것이다.

[0012] 본 발명의 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0013] 상기 기술적 과제를 달성하기 위한 본 발명의 일 실시 예에 따른 프로세서 및 메모리를 포함하는 장치가 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식하는 (a) 무인 이동체로부터 이미지 데이터를 획득하는 단계, (b) 상기 획득한 이미지 데이터에 포함된 텍스트 데이터를 인식하는 단계, (c) 상기 인식한 텍스트 데이터 내에 오탈자가 있는 경우, 상기 텍스트 데이터를 보정하는 단계 및 (d) 상기 보정한 텍스트 데이터를 기반으로 상기 이미지 데이터에 대한 위치 후보군을 선정하는 단계를 포함할 수 있다.

[0014] 일 실시 예에 따르면, 상기 (a) 단계 이후에, (b′) 상기 이미지 데이터 내에 텍스트 데이터가 포함되어 있지 않는 경우, 상기 무인 이동체로부터 후속적인 이미지 데이터를 획득하는 단계를 포함할 수 있다.

[0015] 일 실시 예에 따르면, 상기 (b) 단계는, (b-1) 상기 이미지 데이터에 포함된 모든 텍스트 데이터를 감지하여 상

기 감지한 텍스트 데이터 각각을 둘러싼 텍스트 영역을 지정 및 산출하는 단계, (b-2) 상기 이미지 데이터에 포함된 인쇄물 데이터를 감지하여 상기 인쇄물 데이터 내의 텍스트 영역들 간의 거리가 기 설정된 거리 이하로 근접한 경우, 상기 인쇄물 데이터에 포함된 텍스트 영역을 필터링 하는 단계 및 (b-3) 상기 이미지 데이터에 포함된 텍스트 영역의 텍스트 데이터들을 인식하는 단계를 포함할 수 있다.

[0016] 일 실시 예에 따르면, 상기 (b-2) 단계 이후에, (b-3′) 상기 이미지 데이터 내에 텍스트 영역이 존재하지 않는 경우, 상기 무인 이동체로부터 후속적인 이미지 데이터를 획득하는 단계를 포함할 수 있다.

[0017] 일 실시 예에 따르면, 상기 (b-3) 단계는, 상기 이미지 데이터를 광학 문자 인식 인공지능 모델에 입력하여 상기 이미지 데이터에 포함된 텍스트 영역의 텍스트 데이터를 인식하는 단계일 수 있다.

[0018] 일 실시 예에 따르면, 상기 (b) 단계 이후에, (c′) 상기 인식한 텍스트 데이터 내에 오탈자가 없는 경우, 상기 텍스트 데이터를 기반으로 상기 이미지 데이터에 대한 위치 후보군을 선정하는 단계를 포함할 수 있다.

[0019] 일 실시 예에 따르면, 상기 (c) 단계는, 상기 텍스트 데이터에 포함된 단어들을 Symspell 알고리즘을 기반으로 보정하되, 상기 텍스트 데이터에 포함된 단어와 기 설정된 편집 거리 내에 해당하는 적어도 하나 이상의 보정 후보 단어- 상기 보정 후보군 단어는, 상기 장치 내의 말뭉치 데이터에 저장되어 있는 데이터임-를 매칭하고, 상기 보정 후보 단어들 중 실내 장소명과 근접한 단어를 추출하여 보정하는 단계일 수 있다.

[0020] 일 실시 예에 따르면, 상기 (d) 단계 이후에, (e) 상기 선정된 위치 후보군 중 가장 많이 선정된 위치 후보군을 상기 이미지 데이터에 대한 최종 위치 후보군으로 출력하는 단계를 더 포함할 수 있다.

[0021] 상기 기술적 과제를 달성하기 위한 본 발명의 또 다른 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 장치는 하나 이상의 프로세서, 네트워크 인터페이스, 상기 프로세서에 의해 수행되는 컴퓨터 프로그램을 로드(Load)하는 메모리 및 대용량 네트워크 데이터 및 상기 컴퓨터 프로그램을 저장하는 스토리지를 포함하되, 상기 컴퓨터 프로그램은 상기 하나 이상의 프로세서에 의해, (A) 사용자 단말로부터 단서 정보 및 무인 이동체로부터 실시간 탐지 정보 중 어느 하나 이상을 획득하는 프로세스, (B) 상기 획득한 이미지 데이터에 포함된 텍스트 데이터를 인식하는 프로세스, (C) 상기 인식한 텍스트 데이터 내에 오탈자가 있는 경우, 상기 텍스트 데이터를 보정하는 프로세스 및 (D) 상기 보정한 텍스트 데이터를 기반으로 상기 이미지 데이터에 대한 위치 후보군을 선정하는 프로세스를 실행한다.

[0022] 상기 기술적 과제를 달성하기 위한 본 발명의 또 다른 실시 예에 따른 매체에 저장된 컴퓨터 프로그램은 컴퓨팅 장치와 결합하여, (AA) 무인 이동체로부터 이미지 데이터를 획득하는 단계, (BB) 상기 획득한 이미지 데이터에 포함된 텍스트 데이터를 인식하는 단계, (CC) 상기 인식한 텍스트 데이터 내에 오탈자가 있는 경우, 상기 텍스트 데이터를 보정하는 단계 및 (DD) 상기 보정한 텍스트 데이터를 기반으로 상기 이미지 데이터에 대한 위치 후보군을 선정하는 단계를 실행한다.

발명의 효과

[0024] 상기와 같은 본 발명에 따르면, 무인 이동체로부터 획득한 이미지 데이터 내에 포함된 텍스트 데이터를 인식할 뿐만 아니라 인식한 텍스트 데이터를 자진하여 보정함으로써, 오탈자나 촬영 환경(예. 조도가 높음 등)에 의해 발생하는 식별 불가능한 텍스트도 정확하고 신속하게 인식할 수 있는 효과가 있다.

[0025] 또한, 무인 이동체로부터 획득한 이미지 데이터를 획득하였을 시, 이미지 데이터 내에서 실내 장소명 예측에 불필요한 텍스트 데이터를 필터링함으로써, 실제 활용 가능한 데이터를 기반으로만 실내 장소명 예측이 실시되기에 실내 장소명 예측 정확도를 비약적으로 향상시킬 수 있다.

[0026] 본 발명의 효과들은 이상에서 언급한 효과들로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해 될 수 있을 것이다.

도면의 간단한 설명

[0028] 도 1은 본 발명의 제1 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 장치의 기본적인 구성을 나타낸 도면이다.

도2는 본 발명의 제1 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 장치의 기능적인 구성을 나타낸 도면이다.

도 3은 본 발명의 제2 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법을 개념

적으로 이해하기 위하여 나타낸 도면이다.

도4는 본 발명의 제2 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식의 대표적인 단계를 도시한 순서도이다.

도5은 본 발명의 제2실시 예에 따른 이미지 데이터에 포함된 텍스트 데이터를 인식하는 단계를 구체적으로 나타낸 도면이다.

도6은 본 발명의 제1 실시 예에 따른 문자 검출부가 텍스트 데이터 및 인쇄물 데이터를 검출하는 모습을 간단한 그림을 통하여 나타낸 도면이다.

도7은 본 발명의 제1 실시 예에 따른 필터링부가 인쇄물 데이터 내의 텍스트 데이터를 필터링하는 과정을 그림을 통하여 나타낸 도면이다.

도8은 본 발명의 제1 실시 예에 따른 후처리부가 텍스트 데이터 내의 오탈자를 보정하는 과정을 간단한 모식도로 나타낸 도면이다.

도9는 본 발명의 제3실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법을 구체적으로 나타낸 도면이다.

도10은 본 발명의 제4 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법을 구체적으로 나타낸 도면이다.

발명을 실시하기 위한 구체적인 내용

[0029] 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시 예를 상세히 설명한다. 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시 예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 게시되는 실시 예에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시 예들은 본 발명의 게시가 완전하도록 하고, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭한다.

[0030] 다른 정의가 없다면, 본 명세서에서 사용되는 모든 용어(기술 및 과학적 용어를 포함)는 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 공통적으로 이해될 수 있는 의미로 사용될 수 있을 것이다.

[0031] 또한, 일반적으로 사용되는 사전에 정의되어 있는 용어들은 명백하게 특별히 정의되어 있지 않는 한, 이상적으로 또는 과도하게 해석되지 않을 것이다.

[0032] 본 명세서에서 사용된 용어는 실시 예들을 설명하기 위한 것이며 본 발명을 제한하고자 하는 것은 아니다. 본 명세서에서, 단수형은 문구에서 특별히 언급하지 않는 한 복수형도 포함한다.

[0033] 명세서에서 사용되는 '포함한다 (comprises)' 및/또는 '포함하는 (comprising)'은 언급된 구성 요소, 단계, 동작 및/또는 소자는 하나 이상의 다른 구성 요소, 단계, 동작 및/또는 소자의 존재 또는 추가를 배제하지 않는다.

[0034] 도 1은 본 발명의 제1 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 장치(100)가 포함하는 기본적인 구성을 나타낸 도면이다.

[0035] 그러나 이는 본 발명의 목적을 달성하기 위한 바람직한 실시 예일 뿐이며, 필요에 따라 일부 구성이 추가되거나 삭제될 수 있고, 어느 한 구성이 수행하는 역할을 다른 구성이 함께 수행할 수도 있음은 물론이다.

[0036] 참고로, 본 발명의 명세서에서는 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 장치(100)의 구성을 기본적인 구성과 기능적인 구성으로 나뉘어 설명하며, 도1을 통하여 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 장치(100)가 주어진 역할을 수행하기 위하여 필수적으로 포함되어있어야 하는 기본 구성에 대하여 설명하고, 도2를 통하여 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 장치(100)가 실내 장소명을 예측하기 위하여 필요한 기능적인 구성을 설명하도록 한다.

[0037] 본 발명의 제1 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 장치(100)는 프로세서(10), 네트워크 인터페이스(20), 메모리(30), 스토리지(40) 및 이들을 연결하는 데이터 버스(50)를 포함할 수 있다.

[0038] 프로세서(10)는 각 구성의 전반적인 동작을 제어한다. 프로세서(10)는 CPU(Central Processing Unit), MPU(Micro Processer Unit), MCU(Micro Controller Unit) 또는 본 발명이 속하는 기술 분야에서 널리 알려져 있는 형태의 프로세서 중 어느 하나일 수 있다. 아울러, 프로세서(10)는 본 발명의 제2 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법을 수행하기 위한 적어도 하나의 애플리케이션 또는 프로그램에 대한 연산을 수행할 수 있으며, 인공지능 프로세서로 구현함이 바람직하다 할 것이다.

[0039] 네트워크 인터페이스(20)는 본 발명의 제1 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 장치(100)의 유무선 인터넷 통신을 지원하며, 그 밖의 공지의 통신 방식을 지원할 수도 있다. 따라서 네트워크 인터페이스(20)는 그에 따른 통신 모듈을 포함하여 구성될 수 있다.

[0040] 메모리(30)는 각종 데이터, 명령 및/또는 정보를 저장하며, 본 발명의 제2 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법을 수행하기 위해 스토리지(40)로부터 하나 이상의 컴퓨터 프로그램(41)을 로드할 수 있다. 도 1에서는 메모리(30)의 하나로 RAM을 도시하였으나 이와 더불어 다양한 저장 매체를 메모리(30)로 이용할 수 있음은 물론이다.

[0041] 스토리지(40)는 하나 이상의 컴퓨터 프로그램(41) 및 대용량 네트워크 데이터(42)를 비임시적으로 저장할 수 있다. 이러한 스토리지(40)는 ROM(Read Only Memory), EPROM(Erasable Programmable ROM), EEPROM(Electrically Erasable Programmable ROM), 플래시 메모리 등과 같은 비휘발성 메모리, 하드 디스크, 착탈형 디스크, 또는 본 발명이 속하는 기술 분야에서 널리 알려져 있는 임의의 형태의 컴퓨터로 읽을 수 있는 기록 매체 중 어느 하나일 수 있다.

[0042] 컴퓨터 프로그램(41)은 메모리(30)에 로드되어, 하나 이상의 프로세서(10)에 의해 (A) 무인 이동체로부터 이미지 데이터를 획득하는 프로세스, (B) 상기 획득한 이미지 데이터에 포함된 텍스트 데이터를 인식하는 프로세스, (C) 상기 인식한 텍스트 데이터 내에 오탈자가 있는 경우, 상기 텍스트 데이터를 보정하는 프로세스 및 (D) 상기 보정한 텍스트 데이터를 기반으로 상기 이미지 데이터에 대한 위치 후보군을 선정하는 프로세스를 실행할 수 있다.

[0043] 지금까지 간단하게 언급한 컴퓨터 프로그램(41)이 수행하는 오퍼레이션은 컴퓨터 프로그램(41)의 일 기능으로 볼 수 있으며, 보다 자세한 설명은 본 발명의 제2 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법에 대한 설명에서 후술하도록 한다.

[0044] 데이터 버스(50)는 이상 설명한 프로세서(10), 네트워크 인터페이스(20), 메모리(30) 및 스토리지(40) 사이의 명령 및/또는 정보의 이동 경로가 된다.

[0045] 이상 설명한 본 발명의 제1 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 장치(100)는 네트워크 기능을 보유한 물리적으로 독립된 전자 장치일 수 있으나, GPS 위치 추적 서버(미도시), 말뭉치 데이터 관제 서버(미도시)등으로부터 실내 장소명 예측에 필요한 데이터들을 수신해야 할 수 있으므로 해당 장치의 일 기능으로 구현할 수도 있으며, 이 경우 해당 장치는 유형의 물리적인 장치일 수도 있고, 가상의 클라우드(Cloud) 장치일 수도 있음은 물론이라 할 것이다.

[0046] 이이서, 도2는 본 발명의 제1 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 장치(이하 장치(100)이라 약칭함)의 기능적인 구성을 나타낸 도면이다.

[0047] 도2를 참조하면, 본 발명의 제1 실시 예에 따른 장치(100)는 이미지 데이터 입력부(110), 문자 검출부(120), 필터링부(130), 문자 인식부(140), 후처리부(150) 및 후보선정부(160)을 포함할 수 있다.

[0048] 참고로, 장치(100)의 기능적인 구성에 의한 구체적인 프로세스는 본 발명의 제2 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법을 설명할 때에 구체적으로 설명할 것이기에, 도2의 설명에서는 장치(100)의 기능적인 구성들의 역할에 대하여만 간략히 설명하도록 한다.

[0049] 먼저, 이미지 데이터 입력부(110)는 무인 이동체(200)로부터 획득한 이미지 데이터를 장치(100) 내에 입력하거나 읽어 들이는 구성이다.

[0050] 구체적으로, 이미지 데이터 입력부(110)는 장치(100)가 네트워크 인터페이스(20)를 통하여 실내외를 주행하는 무인 이동체(200)로부터 정지 또는 실시간 이미지 데이터를 획득하면, 이미지 데이터 내의 텍스트 데이터들을 단서로 해당 이미지 데이터의 실내 장소명을 예측을 실시하기 위하여 이미지 데이터를 장치 내에 읽어 들이거나 입력하는 구성이라고 말할 수 있다.

[0051] 또한, 본 발명에서 말하는 무인 이동체(200)는 사용자의 제어 신호에 의하여 원격으로 작동이 제어되는 이동체 또는 사용자의 입력 데이터를 기반으로 자율주행을 실시하는 이동체라고 말할 수 있다.

[0052] 무인 이동체(200)는 도3에 도시된 바와 같이 적어도 하나 이상의 프로펠러가 구비되어 있어 상공을 비행하거나 자율비행할 수 있는 항공 무인 이동체(200)로 활용될 수 있으나, 본 발명의 무인 이동체(200)는 이에 한정되지 않으며 육로를 주행하는 육로형 무인 이동체(미도시), 수면을 따라 이동하는 수상 무인 이동체(미도시)로서도 활용될 수 있는바, 다양한 환경, 공간에 따라 육/해/공을 이동할 수 있는 무인 이동체(200)라면 본 발명의 무인 이동체(200)가 될 수 있다

[0053] 무인 이동체(200)는 카메라가 구비되어 있어, 상기 카메라를 통하여 실내외 공간을 적어도 하나 이상의 프레임 단위로 실시간으로 촬영하여, 정지 이미지 또는 실시간의 영상 이미지 데이터를 획득할 수 있다.

[0054] 문자 검출부(120)는 이미지 데이터 입력부(110)로부터 입력된 이미지 데이터에 포함된 텍스트 데이터를 검출하는 구성이다.

[0055] 문자 검출부(120)는 단순히 이미지 데이터 내에 포함된 텍스트 데이터를 검출하는 것이 아니라 텍스트 데이터를 둘러싼 텍스트 영역을 생성하고, 이미지 데이터 내의 객체(예. 사람, 사물, 인쇄물 등)를 식별하여 인쇄물 데이터(예. 현수막, 포스터, 간판 등)를 감지할 수 있다.

[0056] 참고로, 앞서 언급한 텍스트 영역 및 인쇄물 데이터는 문자 검출(120)가 텍스트 데이터를 검출하는 모습이 도시된 도6을 설명할 때에 다시 한번 설명하도록 한다.

[0057] 필터링부(130)는 장치(100)가 이미지 데이터 내에 포함된 텍스트 데이터를 기반으로 실내 장소명을 예측할 때에 불필요한 텍스트 데이터를 필터링하는 구성이다.

[0058] 예를 들어, 장치(100)가 무인 이동체(200)로부터 학교 전체 공간 중 일부 공간에 대한 이미지 데이터를 획득하였다고 가정하였을 때, 상기 이미지 데이터 내의 '2-1반', '미술실', '급식실' 등은 실내 장소를 예측하는 것에 도움이 될 만한 텍스트 데이터가 될 수 있지만, '급식 메뉴표', '대학 입시 포스터' 에 포함된 텍스트 데이터는 실내 장소명을 예측하는 것에 불필요한 텍스트가 될 수 있는바, 본 발명의 필터링부(130)는 이미지 데이터 내의 인쇄물 데이터(예. 현수막, 포스터, 간판)에 포함된 텍스트 데이터를 제거하여 필터링할 수 있다.

[0059] 이렇게 필터링부(130)에 의한 불필요한 텍스트 데이터의 필터링 과정이 완료되면, 해당 이미지 데이터 내에 실내 장소를 예측하는 것에 필요한 텍스트 데이터만 남겨져 있기에, 문자 인식부(140)는 해당 이미지 데이터 내의 모든 텍스트 데이터를 OCR(Optical Character Recognition: 광학 문자 인식) 기술을 활용하여 인식하여 실내 장소명 예측에 필요한 정확한 결과 값을 산출할 수 있다.

[0060] 후처리부(150)는 딥러닝 모델이 인식한 텍스트의 보정을 위해 Symspell 알고리즘을 활용해 텍스트 데이터를 보정을 하는 구성으로서, Symspell 알고리즘을 통한 텍스트 데이터를 실시할 때에 문자열 메트릭(Metric)인 편집 거리(예. 레벤슈타인 거리)와 미리 정의해둔 말뭉치 데이터가 활용될 수 있다.

[0061] 후처리부(150)가 Symspell알고리즘, 자모 단위의 편집 거리, 말뭉치 데이터를 활용하여 텍스트 데이터를 보정하면, 후보 선정부(160)는 보정한 텍스트 데이터를 위치 후보군으로 선정한다.

[0062] 이후에, 후처리부(150) 및 후보 선정부(160)가 텍스트 보정과 위치 보군 선정이 무인 이동체(200)로부터 획득한 이미지 데이터를 통하여 연속적으로 실시되면, 후보 선정부(160)는 선정된 복수 개의 위치 후보군 중 가장 많은 위치 후보군을 최종 위치 후보군으로 출력하여, 해당 이미지 데이터의 실내 장소명이 최종 위치 후보군인 것으로 판단할 수 있도록 한다.

[0063] 이상 본 발명의 제1 실시 예에 따른 장치(100)의 기본적인 구성과 기능적인 구성에 대하여 간략히 살펴보았다.

[0064] 이하 본 발명의 제2 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법에 대하여 살펴보기 이전에, 본 발명의 제2실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법에 대한 개념을 간략히 살펴보도록 한다.

[0065] 도3은 본 발명의 제2 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법을 개념적으로 이해하기 위하여 나타낸 도면이다.

[0066] 본 발명의 제1 실시 예에 따른 장치(100)는 앞서 설명하였듯이 실내외의 어느 한 공간을 주행하는 무인 이동체(200)가 구비된 카메라를 통하여 상기 어느 한 공간을 촬영하여 생성된 이미지 데이터를 획득하면, 이미지 데이

터 내의 텍스트 데이터를 기반으로 상기 어느 한 공간의 실내 장소명이 무엇인지 예측해주는 장치이다.

[0067] 특히, 본 발명의 장치(100)는 단순히 획득한 이미지 데이터 내의 텍스트 데이터를 인식하고, 인식한 텍스트 데이터를 기반으로 실내 장소명을 분석 및 예측하는 것이 아니라, 텍스트 데이터 인식 과정에서 불필요한 텍스트 데이터를 필터링하고, 인식한 텍스트 데이터에 오탈자 또는 노이즈가 있을 경우, 텍스트 데이터를 보정하여 정확하고 효율적으로 실내 장소명을 예측할 수 있는 장치이다.

[0068] 이러한 장치(100)는 무인 이동체(200)로부터 획득한 실시간의 이미지 데이터에 해당하는 장소명을 정확하고 정밀하게 추론함으로써 복잡한 건물구조(예. 동일한 방이 복수 개로 설계됨)속에서 이미지 데이터에 해당하는 위치를 정확하고 신속하게 파악할 수 있다.

[0069] 도3을 참조하면, 이러한 기술적 특징을 보유한 장치(100)는 획득한 이미지 데이터를 기반으로 실내 장소명을 예측하기 위하여 먼저, 무인 이동체(200)로부터 이미지 데이터를 획득한다(①).

[0070] 이후에, 장치(100)는 획득한 이미지 데이터 내의 텍스트 데이터 및 인쇄물 데이터를 감지하고(②), 실내 장소명 예측에 불필요한 텍스트 데이터를 필터링 한다.

[0071] 다음으로, 장치(100)는 감지한 모든 텍스트 데이터를 OCR 기술을 활용하여 인식하고, 인식한 텍스트 데이터 중 어느 하나의 텍스트 데이터에 오탈자가 있는 경우, 오탈자가 포함된 텍스트 데이터를 보정한다(③).

[0072] 예를 들어, 무인 이동체(200)가 카메라를 통하여 행복상담소라는 표지판을 촬영하였으나, 촬영과정 중 무인 이동체(200)가 소정의 방해 요인에 의하여 흔들렸다는 것을 전제로, 장치(100)가 상기 무인 이동체(200)로부터 획득한 이미지 데이터 내의 텍스트 데이터를 인식하였을 때에 이미지 데이터 내의 노이즈(예. 블러 현상)'행복상담소'가 아닌'행복상담수'로 텍스트 데이터를 인식하였다면, 장치(100)는 앞서 설명한 후처리부(140)를 통하여 '행복상담수'를 오탈자로 인식하고, '행복상담수'를 '행복상담소'로 보정할 수 있다.

[0073] 이상 본 발명의 제2실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법에 대한 개념을 간략히 살펴보았다.

[0074] 다음으로, 앞서 언급한 텍스트 인식 과정, 불필요한 텍스트 필터링 과정, 오탈자 보정 과정이 구체적으로 어떻게 실시되는 지에 대하여 살펴보기 위하여 본 발명의 제2 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법을 구체적으로 설명하도록 한다.

[0075] 도 4는 본 발명의 제2 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법의 대표적인 단계를 도시한 순서도이다.

[0076] 이는 본 발명의 목적을 달성함에 있어서 바람직한 실시 예일 뿐이며, 필요에 따라 일부 단계가 추가되거나 삭제될 수 있고, 더 나아가 어느 한 단계가 다른 단계에 포함될 수도 있음은 물론이다.

[0077] 우선, 장치(100)는 무인 이동체(200)로부터 이미지 데이터를 획득한다(S101).

[0078] S101 단계는 실시간으로 주행 중인 무인 이동체(200)가 촬영한 정지 이미지 데이터 또는 연속적인 영상 이미지 데이터를 획득하는 단계일 수 있지만, 이에 한정되지 않고, 무인 이동체(200) 또는 사용자 단말(여기서 말하는 사용자 단말은, 이미지 데이터를 통하여 실내 장소명을 획득하고자 하는 사용자의 단말임)이 과거에 수집 획득한 정지 이미지 데이터 또는 연속적인 영상 이미지 데이터를 획득하는 단계일수도 있다.

[0079] 이후에, 장치(100)는 이미지 데이터에 포함된 텍스트 데이터를 인식한다(S102).

[0080] S102 단계는 전술한 문자 검출부(120)가 이미지 데이터에 포함된 텍스트 데이터와 인쇄물 데이터를 구분하고 검출하는 단계라고 볼 수 있으며, 이 때의 문자 검출부(120)는 텍스트만을 감지하는 텍스트 감지 알고리즘 또는 이미지에 포함된 객체를 구분하는 객체 탐지 알고리즘(예. YOLOv7)을 실행할 수 있다.

[0081] 관련하여, 무인 이동체(200)가 촬영한 모든 이미지 데이터에는 텍스트 데이터가 포함되어 있지 않을 수도 있는 바, S102 단계는 장치(100)가 획득한 이미지 데이터에 텍스트 데이터가 포함되어 있는지 판단하고, 판단 결과에 따라 다른 프로세스가 진행될 수 있으며, 이러한 프로세스에 대한 설명은 도5을 통하여 구체적으로 설명하도록 한다.

[0082] 도5는 본 발명의 제2실시 예에 따른 이미지 데이터에 포함된 텍스트 데이터를 인식하는 단계를 구체적으로 나타낸 도면이다.

[0083] 도5를 참조하면, S102 단계는 이미지 데이터에 텍스트 데이터가 포함되어 있는지에 대하여 판단하는 단계(S102-

1)를 포함한다.

[0084] S102-1 단계에서 장치(100)가 이미지 데이터 내의 텍스트 데이터가 포함되어 있다고 판단하지 않은 경우, 무인 이동체(200)로부터 후속적인 이미지 데이터를 획득한다(S102-2a). 쉽게 말해, 무인 이동체(200)로부터 실내 장소명 예측의 단서가 될만한 이미지 데이터를 재획득한다.

[0085] 한편, S102-1 단계에서 장치(100)가 이미지 데이터 내의 텍스트 데이터가 포함되어 있다고 판단한 경우, 이미지 데이터에 포함된 모든 텍스트 데이터를 검출하여 검출한 텍스트 데이터 각각을 둘러싼 텍스트 영역을 지정 및 산출한다(S102-2b).

[0086] 이후에, 장치(100)는 검출한 텍스트 영역들 중 적어도 둘 이상의 텍스트 영역들 간의 거리가 기 설정된 거리 이하로 근접한 경우, 상기 거리가 기 설정된 거리 이하로 근접한 텍스트 영역들이 인쇄물 데이터 내에 포함되어 있는 텍스트 영역이라고 판단한다(S102-3).

[0087] 참고로, S102-2b단계 및 S102-3단계에서 설명한 '텍스트 영역', '인쇄물 데이터'에 대한 설명은 도6 및 도7을 설명할 때에 다시 한번 설명하도록 한다.

[0088] 다음으로, 장치(100)가 인쇄물 데이터에 포함되어 있는 텍스트 영역 내의 텍스트 데이터를 필터링하며(S102-4), 이미지 데이터에 포함된 텍스트 영역의 텍스트 데이터들을 인식한다(S102-5).

[0089] 도6은 본 발명의 제1 실시 예에 따른 문자 검출부(120)가 텍스트 데이터 및 인쇄물 데이터를 검출하는 모습을 간단한 그림을 통하여 나타낸 도면이다.

[0090] 도6을 참조하면, 문자 검출부(120)는 도6에 도시된 이미지 데이터를 획득하면, 이미지 데이터 내의 텍스트 데이터를 검출하여, 텍스트 데이터를 둘러싼 텍스트 영역을 지정 및 산출한다.

[0091] 또한, 문자 검출부(120)는 텍스트 데이터를 검출할 뿐만 아니라 이미지 데이터 내의 객체가 포스터, 현수막, 간판인지 식별하여 인쇄물 데이터를 검출한다.

[0092] 예를 들어, 문자 검출부(120)가 도6에 도시된 이미지 데이터에 대한 문자 검출을 실시하였다면, 문 앞 표지판에 기재된 '101호'의 텍스트 데이터를 검출하고, 벽면에 부착된 인쇄물 데이터들 검출하되, 인쇄물 데이터 내에 포함된 '고등학교 입시전략'이라고 기재된 텍스트 데이터를 검출할 수 있다.

[0093] 문자 검출부(120)는 인쇄물 데이터에 포함된 텍스트 간의 배치관계가 일반적으로 서로 인접해 있다는 특징을 고려하여, 문자 검출부(120)가 최초로 검출한 텍스트 영역 또는 텍스트 데이터 간의 배치가 기 설정된 거리 이하인 경우, 인쇄물 데이터라고 식별할 수 있다.

[0094] 또한, 문자 검출부(120)는 과거의 학습 데이터(예. 과거의 학습된 포스터, 현수막, 간판)를 통하여 이미지 데이터에 포함된 객체가 인쇄물 데이터인지를 검출할 수 있다. 다시 말해, 문자 검출부(120)는 기 학습된 인쇄물 데이터의 특징(예. 형상, 크기/규격, 텍스트 배치 관계 등)을 기반으로 획득한 이비지 데이터에 포함된 객체가 포스터인지 실내 장소를 예측할 수 있는 표지판 또는 방팻말인지를 구분할 수 있다.

[0095] 이렇게 문자 검출부(120)가 텍스트 영역을 지정 및 산출하고, 인쇄물 데이터를 식별하는 이유는, 일반적으로 인쇄물 데이터는 소정의 이벤트를 홍보하거나 광고하는 용도 또는 정보를 제공하는 용도로 활용되기에 실내 장소명을 예측을 예측하는 것에 불필요하기 때문이다.

[0096] 다만, 인쇄물 데이터 내에도 인쇄물 데이터가 부착 또는 게시된 위치를 예측할 수 있는 텍스트가 포함되어 있을 수 있는바, 문자 검출부(120)는 인쇄물 데이터 내의 텍스트 데이터가 필요 데이터인지 불필요 데이터인지, 인쇄물 데이터 내에 포함된 텍스트에 직접적인 장소 관한 텍스트가 포함되어 있는지에 대한 여부를 판단하여 이후에 필터링부가 상기 텍스트를 제거하지 않도록 표시해줄 수 있다.

[0097] 도7은 본 발명의 제1 실시 예에 따른 필터링부(130)가 인쇄물 데이터 내의 텍스트 데이터를 필터링하는 과정을 그림을 통하여 나타낸 도면이다.

[0098] 앞서 도6에 대한 설명에서는 문자 검출부(120)에서 텍스트 영역의 근접 거리를 판단하여 상기 텍스트 영역이 인쇄물 데이터 내에 포함되어 있는지 판단한다고 설명하였지만, 이러한 인쇄물 데이터를 식별하는 프로세스는 문자 검출부(120)에 의하여 실시될 수 있지만 텍스트 데이터를 제거하는 필터링부(130)에서도 실시될 수 있다.

[0099] 이렇게 문자 검출부(120) 또는 필터링부(130)에서 식별된 인쇄물 데이터 내의 텍스트 영역은 필터링부(130)에 의하여 제거된다.

[0100] 예를 들어, 문자 검출부(120)가 '어린이', '영어 학원', '100일의 기적'이라는 텍스트 데이터를 검출하고, 검출한 텍스트 데이터를 둘러싼 각각의 텍스트 영역을 지정 및 산출하였으며, 상기 텍스트 영역들 간의 거리가 기설정된 거리 이하로 근접하다고 가정하였을 때, 문자 검출부(120) 또는 필터링부(130)는 '어린이', '영어 학원', '100일의 기적'이라는 텍스트 데이터를 포함하는 텍스트 영역이 인쇄물 데이터에 포함된 텍스트 영역이라고 판단하고, 필터링부(130)가 인쇄물 데이터에 포함된 텍스트 영역 내의 텍스트 데이터를 모두 제거할 수 있다는 것이다.

[0101] 이렇게, 장치(100)가 문자 검출부(120)을 통하여 텍스트 데이터를 검출하고, 필터링부(130)를 통하여 실내 장소명 예측에 불필요한 텍스트 데이터를 제거하고, 문자 인식부(140)를 통하여 검출한 텍스트 데이터가 무슨 내용인지 인식하였다면, 장치(100)는 인식한 텍스트 데이터에 오탈자가 있는지 판단하고, 텍스트 데이터 내에 오탈자가 있는 경우, 텍스트 데이터를 보정한다(S103).

[0102] S103 단계에서 이루어지는 텍스트 데이터 보정은, 장치(100)의 후처리부(150)에서 실시되는바, 도8에 도시된 후처리부(150)가 텍스트 데이터를 보정하는 과정을 설명할 때에 같이 설명하도록 한다.

[0103] 도 8은 본 발명의 본 발명의 제1 실시 예에 따른 후처리부(150)가 텍스트 데이터 내의 오탈자를 보정하는 과정을 간단한 모식도로 나타낸 도면이다.

[0104] 후처리부(150)는 딥러닝 모델이 인식한 텍스트의 보정을 위해 Symspell 알고리즘을 활용해 텍스트 데이터를 보정을 하는 구성으로서, Symspell 알고리즘을 통한 텍스트 데이터를 실시할 때에 문자열 메트릭(Metric)인 편집거리(예. 레벤슈타인 거리)와 미리 정의해둔 말뭉치 데이터가 활용될 수 있다.

[0105] 이러한 후처리부(150)가 텍스트 데이터 내의 오탈자를 보정하는 과정을 구체적으로 살펴보기 이전에 편집 거리, 말뭉치 데이터 및 Symspell 알고리즘의 개념을 간략히 살펴보도록 한다.

[0106] 먼저, 여기서 말하는 편집 거리는 한 단어를 다른 단어로 변경하는데 필요한 단일 문자 편집의 최소 수인 편집거리를 의미할 수 있으며, 본 발명에서의 편집거리는 글자 단위가 아닌 자모 단위로 편집 거리를 측정한다.

[0107] 한국어를 예를 들면, '행ㅞ'과 '행복'은 자모가 하나 다르므로 레벤슈타인 거리가 1이다. 영어를 예로 들면, 'happiness'와 'heppiness'는 문자 ‘a’와 ‘e’가 다르므로 이 경우도 레벤슈타인 거리가 1이다.

[0108] 이와 같이, 본 발명에서는 보정 가능한 텍스트 데이터의 언어는 한국어뿐만 아니라 영어, 불어, 독일어, 중국어 등이 포함될 수 있다.

[0109] 한편, 본 발명에서 말하는 말뭉치 데이터는 문헌 바탕의 단어와 각 단어의 빈도수가 기록된 데이터로서, 혼동하기 쉬운 문자에 대하여 교정된 문장의 확률을 부과할 뿐만 아니라 단어들 간의 조합이 실내 장소명일 확률을 부과해주는 역할을 수행한다.

[0110] 참고로, 본 발명의 말뭉치 데이터는 연세 말뭉치 용례 검색 시스템에서 제공하는 20세기 문헌의 약 1억 5천만 어절을 분석한 ‘연세 20세기 한국어 말뭉치’를 사용할 수 있으며, 이에 한정되지 않고, 단어의 편집 거리에 따라 실내 장소명이 예측 가능한 단어들이 데이터셋 형태로 매칭되어 있는 말뭉치 데이터라면 본 발명의 말뭉치 데이터가 될 수 있다.

[0111] 말뭉치 데이터는 총 4가지 형태소인 일반 명사(NNG), 체언 접두사(XPN), 명사 파생 접미사(XSN), 의존 명사(NNB)만을 사용하여 단어를 각 형태소로 정확히 나눌 수 있도록 하였다.

[0112] 또한, 말뭉치 데이터에 정의된 각 단어의 빈도는 로그 스케일(log scale)로 연산될 수 있어, 각 단어가 포함된 텍스트의 확률(예. 실내 장소명일 확률) 계산에 용이하게 활용될 수 있도록 하며, 타 검색에 쓰이는 각 단어를 연속한 개수로 묶어서 처리하는 방식인 N-gram을 사용하지 않는 대신, 베이즈 확률론에 따라 각 단어별 확률의 곱으로 문장의 확률을 표현할 수 있다.

[0113] Symspell 알고리즘은 앞서 설명한 편집거리 및 말뭉치 데이터를 기반으로 단어를 기준으로 주어진 단어에서 자음 혹은 모음에 대한 삭제(deletion) 과정을 통해 미리 일정 편집거리 이내의 단어를 해시 테이블에 저장하여 빠르게 보정할 단어를 탐색할 수 있도록 한다

[0114] 예를 들어, '공리'는 편집거리 1 이내의 단어로 삭제과정 중 'ㅇ'를 삭제한 '고리'와 'ㅗ'를 삭제한 'ㄱㅇ리'를 후보로 가지는 '궁리'와 'ㅣ'를 삭제한 '공ㄹ'를 후보로 가지는 '공로'가 후보에 추가된다. '공리'라는 단어에 대한 해시테이블에는 편집거리가 1인 단어로 '고리', '궁리', '공로' 등이 저장된다. 해시 테이블을 사용하는

이유는 같은 편집 거리를 가지는 단어들을 저장해서 탐색 과정을 빠르게 하기 위함이다.

[0115] Sympell 알고리즘은 각 단어의 형태소와 빈도를 바탕으로 실내 장소명에 들어갈만한 형태소로 먼저 말뭉치 사전을 전처리(preprocessing) 한 이후에 실시될 수 있다.

[0116] symspell 알고리즘은 삼각 행렬(triangular matrix) 방법을 통해 말뭉치 데이터의 편집 거리를 측정하여 문장을 보정할 수 있으며. 이 때, 문장의 앞부터 탐색하면서 각 편집거리마다 가장 높은 확률을 가진 문장을 저장한다. 나눈 글자가 말뭉치 사전에 없는 경우 편집 거리를 증가시키면서 해당되는 단어들 중 가장 높은 빈도를 가진 단어(예. 실내 장소명일 확률인 높은 단어)로 교체한다.

[0117] 이후에, Symspell 알고리즘은 텍스트의 끝까지 탐색하게 되었을 때, 각 편집거리에서 가장 확률이 높은 문장을 탐색하여 최종 결과를 출력한다.

[0118] 이상 본 발명의 편집 거리, 말뭉치 데이터 및 Symspell 알고리즘의 개념에 대하여 간략히 살펴보았다.

[0119] 다음으로, 도8을 참조하여 후처리부(150)가 텍스트 데이터를 보정하는 과정을 구체적인 예를 기반으로 설명하도록 한다.

[0120] 도8을 참조하면, 문자 인식부(140)가 '그림꿈나무반'을 소정의 방해 요인에 의하여 '그린끔나무반'으로 잘못 인식하였다고 가정하였을 때, 후처리부(150)의 Symspell 알고리즘은 편집거리 및 말뭉치 데이터를 기반으로 단어를 기준으로 주어진 단어에서 자음 혹은 모음에 대한 삭제(deletion) 과정을 걸쳐 '그린끔나무반' 중 '끔나무'가 오탈자인 것을 감지한다(①).

[0121] 이후에, 후처리부(150)는 말뭉치 데이터를 통하여 '끔나무'가 상기 설정된 편집 거리(예. 2이하의 거리) 내에 해당하는 적어도 하나 이상의 보정 후보 단어(예. 음나무, 꿈나무, 끝나무 등)를 매칭하고(②), 보정 후보 단어들 중 실내 장소명과 근접한 단어를 추출하여 보정한다(③).

[0122] 다시 말해, 장치(100)가 후처리부(150)를 통하여 텍스트 데이터를 보정하는 단계는, 텍스트 데이터에 포함된 단어들을 Symspell 알고리즘을 기반으로 보정하되, 상기 텍스트 데이터에 포함된 단어와 기 설정된 편집 거리 내에 해당하는 적어도 하나 이상의 보정 후보 단어를 매칭하고, 상기 보정 후보 단어들 중 실내 장소명과 근접한 단어를 추출하여 보정하는 단계이다.

[0123] 참고로, 여기서 말하는 보정 후보군 단어란 앞서 언급한 말뭉치 데이터의 해시 테이블에 저장되어 있는 데이터를 말할 수 있다.

[0124] 이렇게 후처리부(140)에 의한 텍스트 데이터 보정이 완료된 경우, 장치(100)는 후보 선정부(160)를 통하여 S103 단계에서 보정한 텍스트 데이터가 이미지 데이터에 대한 위치 후보군이라고 선정한다(S104).

[0125] 여기서 말하는 위치 후보군은 실내 장소명이라고 예측할 수 있는 최종 위치 후보군을 결정하기 위한 예비 위치 후보군이라고 말할 수 있다.

[0126] 이렇게, 장치(100)가 단일의 보정된 텍스트 데이터를 통하여 상기 텍스트 데이터가 실내 장소명이라고 판단하지 않는 이유는, 이미지 데이터 내에 실내 장소명을 판단할 수 있는 텍스트 데이터가 복수 개로 인식될 수 있고, 복수 개의 보정된 텍스트 데이터 중에서 가장 정확한 결과값을 선택하기 위함이다.

[0127] 이상 본 발명의 제2 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법을 자세히 살펴보았다.

[0128] 다음으로, 본 발명의 제3 및 제4 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법을 구체적으로 살펴보도록 한다.

[0129] 본 발명의 제3 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법은 장치(100)가 이미지 데이터에 포함된 텍스트 데이터를 인식하였을 때에, 인식한 텍스트 데이터 내에 오탈자가 없다고 판단한 경우에 대한 실시 예이다.

[0130] 도 9를 참조하면, 이렇게 장치(100)가 이미지 데이터에 포함된 텍스트 데이터를 인식하였을 때에 인식한 텍스트 데이터에 오탈자가 없는 경우, 인식한 텍스트 데이터를 이미지 데이터에 대한 위치 후보군으로 선정한다(S103′).

[0131] 예를 들어, 무인 이동체(200)가 어느 한 교실을 촬영한 이미지 데이터를 장치(100)에 전달하였을 때에, 장치(100)가 전달 받은 이미지 데이터에 포함된 '2-1반', '2학년 1반'및 '문과반' 이라는 텍스트 데이터를 인식하였

다고 가정하였을 때에, 장치(100)의 후처리부(150)는 상기 나열된 텍스트 데이터의 오탈자가 없다고 판단하고, 장치(100)의 후보선정부(160)가 '2-1반', '2학년 1반'및 '문과반'모두를 위치 후보군으로 선정할 수 있다는 것이다.

[0132] 한편, 본 발명의 제4 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법은 장치(100)의 후보선정부(160)가 위치 후보군을 선정한 이후에 복수 개의 위치 후보군 중 가장 실내 장소명으로 적합할 것 같은 최종 후보군을 출력하는 실시 예이다.

[0133] 도10을 참조하면, 장치(100)의 후보선정부(160)가 보정한 텍스트 데이터를 기반으로 이미지 데이터에 대한 위치 후보군을 선정하였다면, 후보 선정부(160)가 선정된 후보군 중 가장 많이 선정된 위치 후보군을 이미지에 대한 최종 후보군으로 출력한다(S105).

[0134] 다시 말해, 후보선정부(160)는 S104 단계에서 보정한 텍스트 데이터를 기반으로 위치 후보군을 선정할 뿐만 아니라 선정한 위치 후보군 중 동일한 위치 후보군이 있을 시 이를 카운팅하며, 카운팅한 위치 후보군 중 가장 많이 선정된 위치 후보군을 이미지 데이터에 대한 최종 후보군으로 출력할 수 있다는 것이다.

[0135] 참고로, 여러 공간을 주행하는 무인 이동체(100)가 현재 촬영 중인 특정 공간에 대한 실내 장소명을 예측 하기 위해서는 장치(100)에 의하여 선정된 후보군이 언제부터 카운팅되고, 카운팅이 언제 완료되는 지가 관건이 될 수 있는바, 무인 이동체(100)가 특정 공간을 진입한 시점부터 상기 특정 공간을 나가는 시점까지의 시간을 장치(100)가 측정하여 상기 특정 공간에 대한 최종 후보군을 출력할 수 있으며, 사용자(예. 장치 관리자)가 특정 시간 범위를 설정하도록 하여, 장치(100)가 기 설정된 시간 범위 동안 위치 후보군을 카운팅하고, 기 설정된 시간이 종료되면 최종 후보군을 출력하도록 할 수 있다.

[0136] 한편, 중복 서술을 방지하기 위해 자세히 설명하지는 않았지만, 본 발명의 제1 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 장치(100)와 본 발명의 제2 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법은 동일한 기술적 특징을 포함하는 본 발명의 제5 실시 예에 따른 매체에 저장된 컴퓨터 프로그램으로 구현할 수 있다. 이 경우 매체에 저장된 컴퓨터 프로그램은 컴퓨팅 장치와 결합하여, (AA) 무인 이동체로부터 이미지 데이터를 획득하는 단계, (BB) 상기 획득한 이미지 데이터에 포함된 텍스트 데이터를 인식하는 단계, (CC) 상기 인식한 텍스트 데이터 내에 오탈자가 있는 경우, 상기 텍스트 데이터를 보정하는 단계 및 (DD) 상기 보정한 텍스트 데이터를 기반으로 상기 이미지 데이터에 대한 위치 후보군을 선정하는 단계를 실행하되, 중복 서술을 방지하기 위해 자세한 설명은 생략하지만 본 발명의 제1 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 장치(100)와 본 발명의 제2 실시 예에 따른 실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법과 동일한 효과를 도출할 수 있을 것이다.

부호의 설명

[0138] 10: 프로세서

20: 네트워크 인터페이스

30: 메모리

40: 스토리지

41: 컴퓨터 프로그램

50: 데이터 버스

100: 장치

110: 이미지 데이터 입력부

120: 문자 검출부

130: 필터링부

140: 문자 인식부

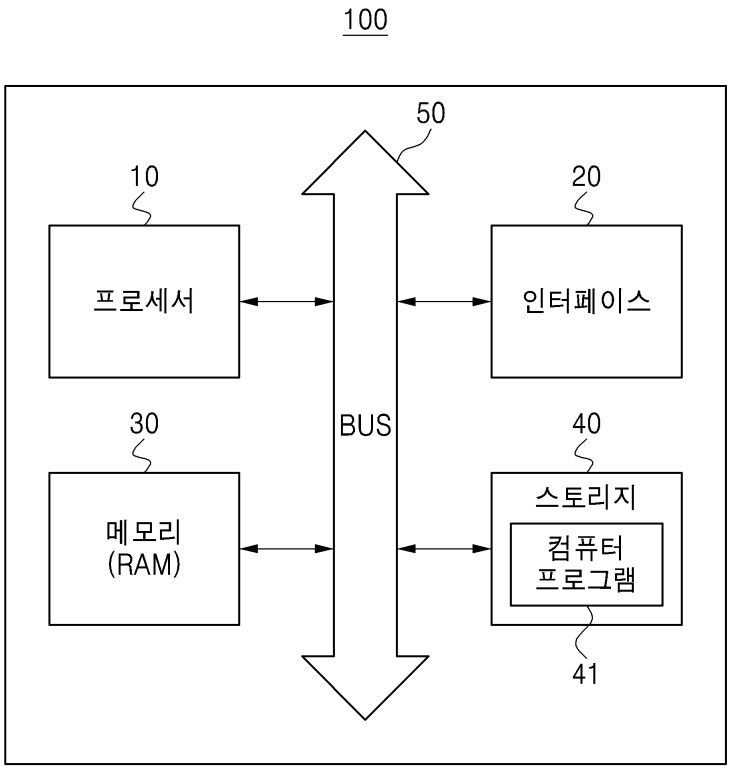
150: 후처리부

160: 후보선정부

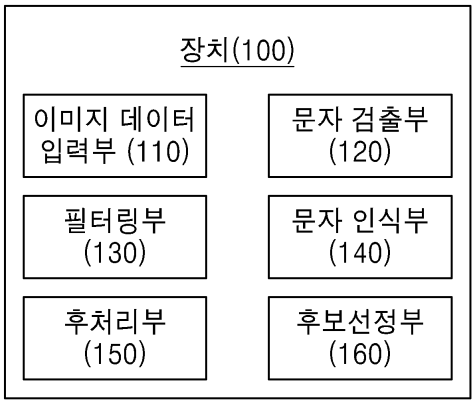
200: 무인 이동체

도면

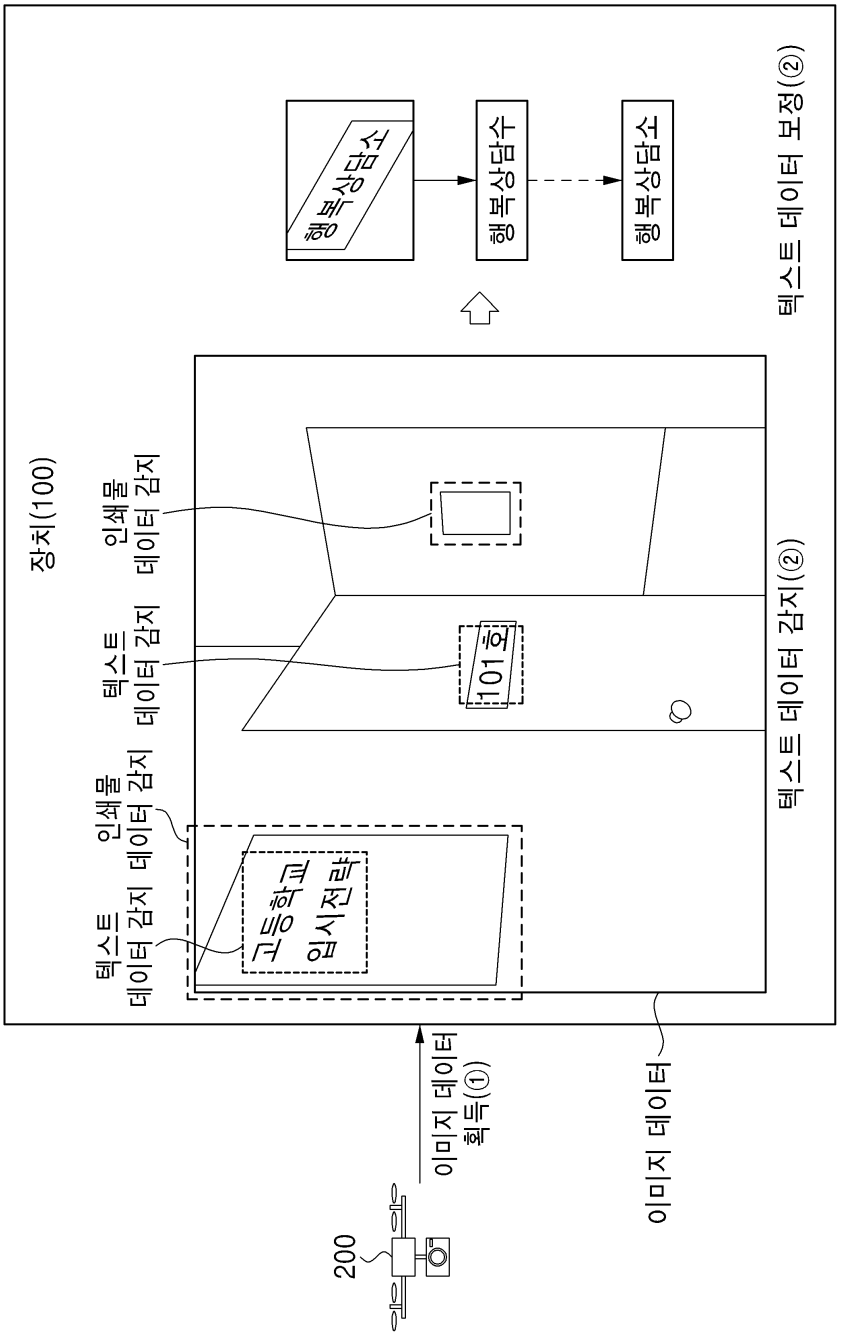
도면1



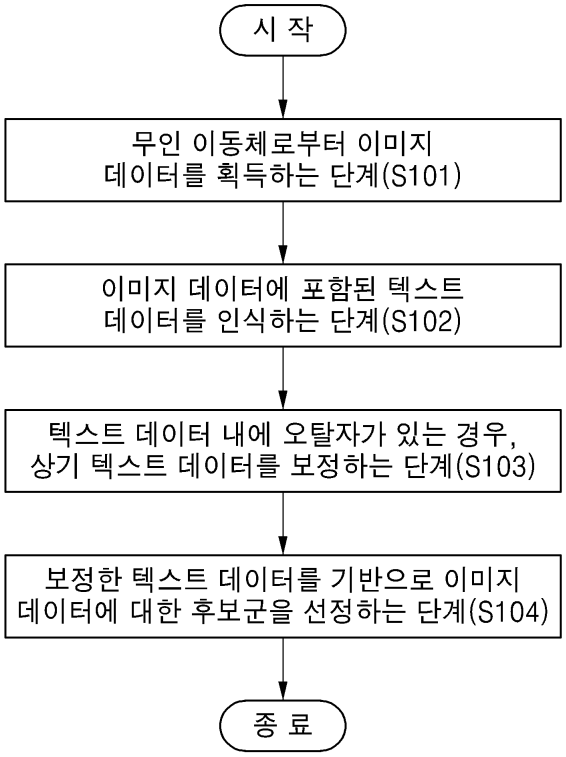
도면2



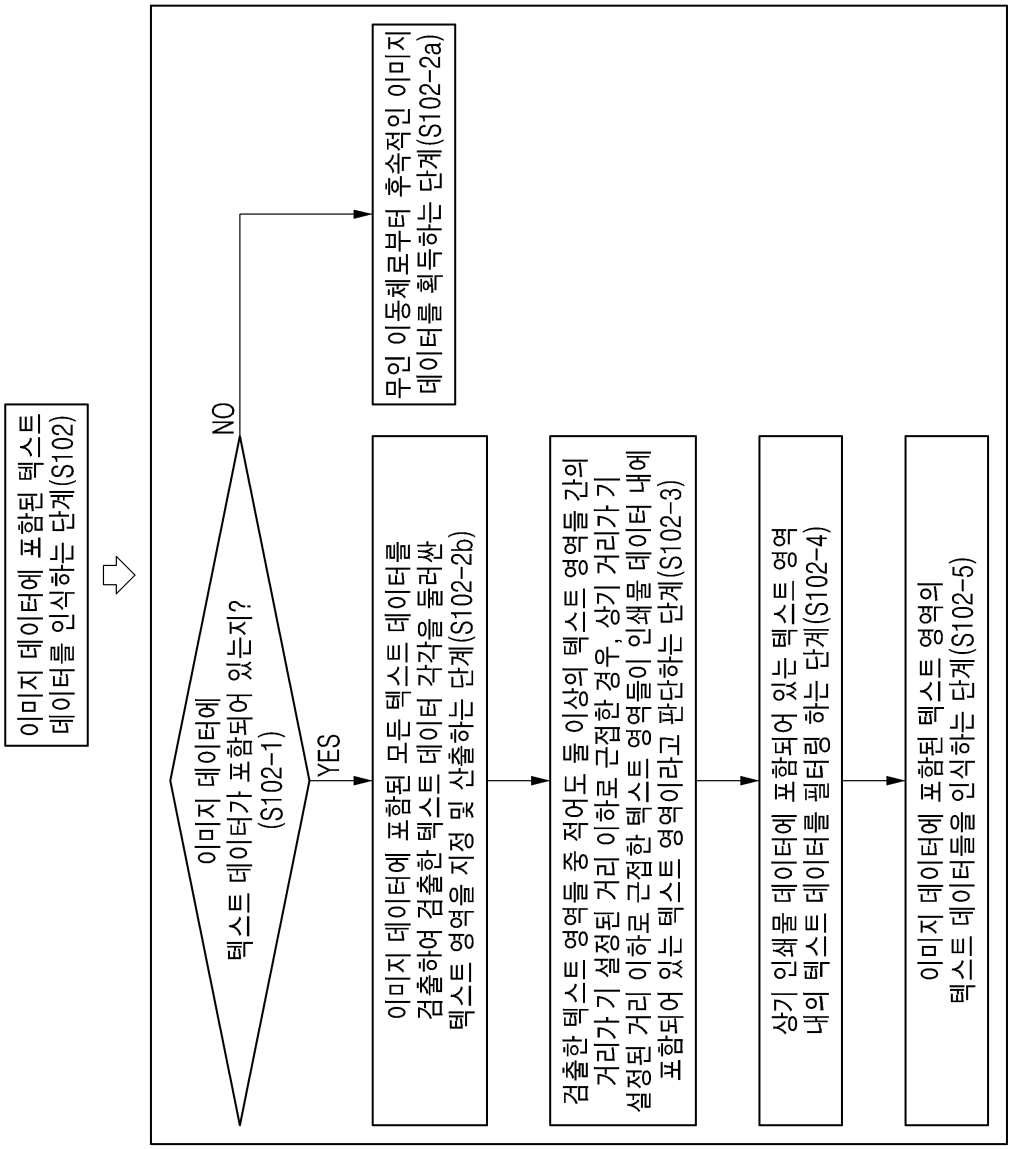
도면3


장치(100)
인쇄물 데이터 감지
텍스트 데이터 감지
인쇄물 데이터 감지
텍스트 데이터 감지
101호
텍스트 데이터 감지(②)
텍스트 데이터 보정(②)
행복상담소
행복상담수
행복상담서
200
이미지 데이터 획득(①)
이미지 데이터

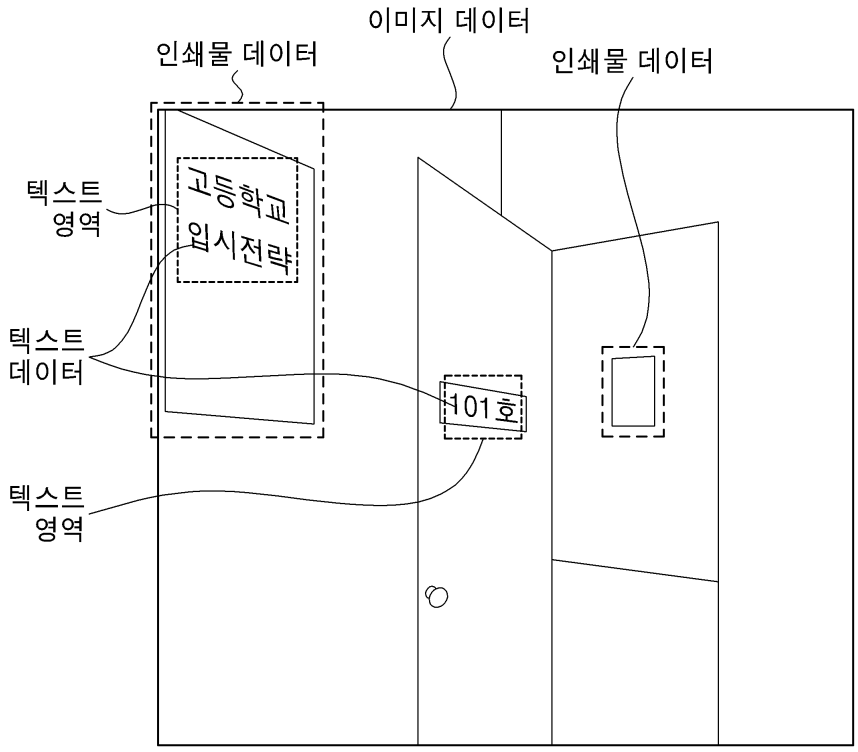
도면4


시 작
무인 이동체로부터 이미지
데이터를 획득하는 단계(S101)
이미지 데이터에 포함된 텍스트
데이터를 인식하는 단계(S102)
텍스트 데이터 내에 오탈자가 있는 경우,
상기 텍스트 데이터를 보정하는 단계(S103)
보정한 텍스트 데이터를 기반으로 이미지
데이터에 대한 후보군을 선정하는 단계(S104)
종 료

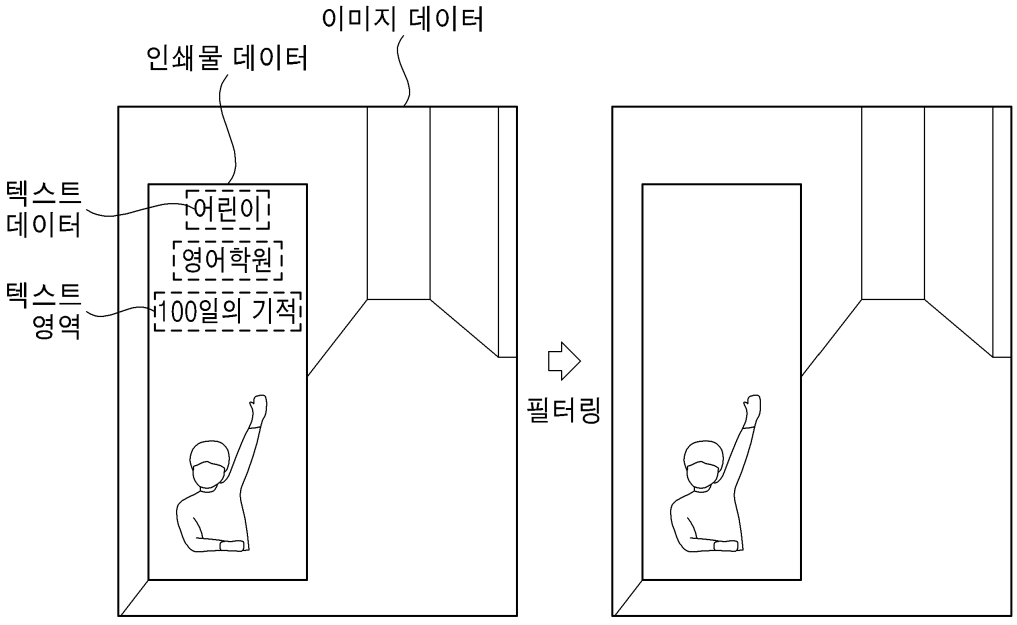
도면5


이미지 데이터에 포함된 텍스트 데이터를 인식하는 단계(S102)
이미지 데이터에 텍스트 데이터가 포함되어 있는지? (S102-1)
NO
YES
무인 이동체로부터 후속적인 이미지 데이터를 획득하는 단계(S102-2a)
이미지 데이터에 포함된 모든 텍스트 데이터를 검출하여 검출한 텍스트 데이터 각각을 둘러싼 텍스트 영역을 지정 및 산출하는 단계(S102-2b)
검출한 텍스트 영역들 중 적어도 둘 이상의 텍스트 영역들 간의 거리가 기 설정된 거리 이하로 근접한 경우, 상기 거리가 기 설정된 거리 이하로 근접한 텍스트 영역들이 인쇄물 데이터 내에 포함되어 있는 텍스트 영역이라고 판단하는 단계(S102-3)
상기 인쇄물 데이터에 포함되어 있는 텍스트 영역 내의 텍스트 데이터를 필터링 하는 단계(S102-4)
이미지 데이터에 포함된 텍스트 영역의 텍스트 데이터들을 인식하는 단계(S102-5)

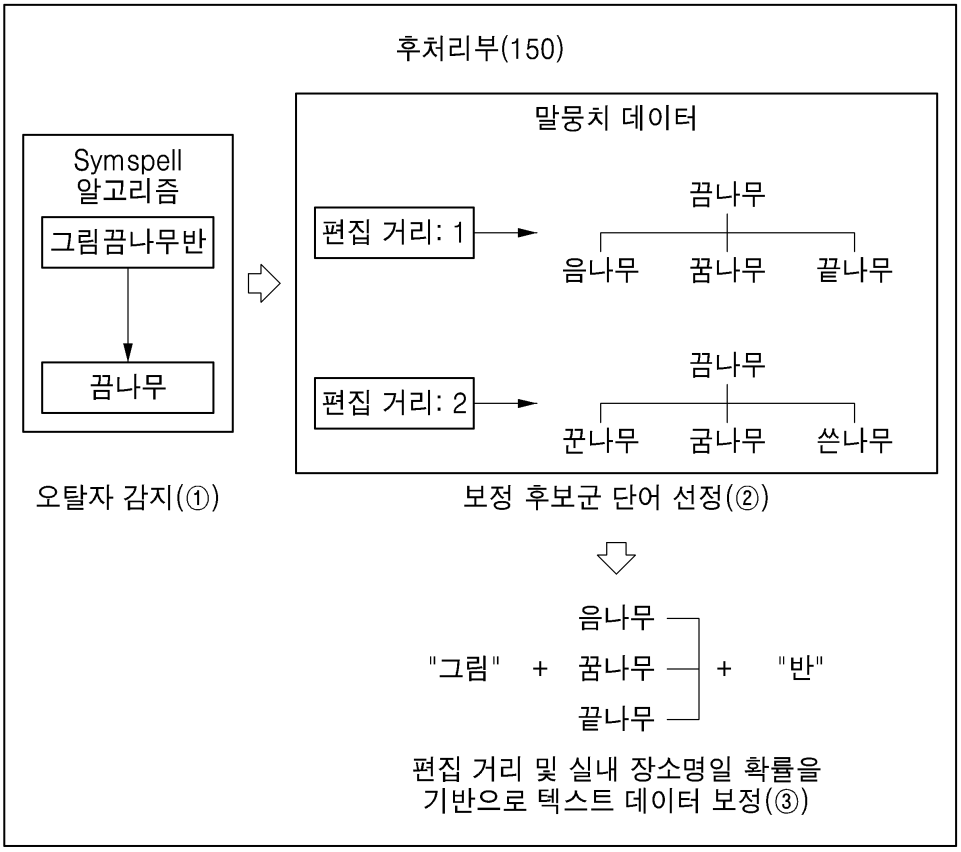
도면6


이미지 데이터
인쇄물 데이터
인쇄물 데이터
텍스트
영역
고등학교
입시전략
텍스트
데이터
101호
텍스트
영역

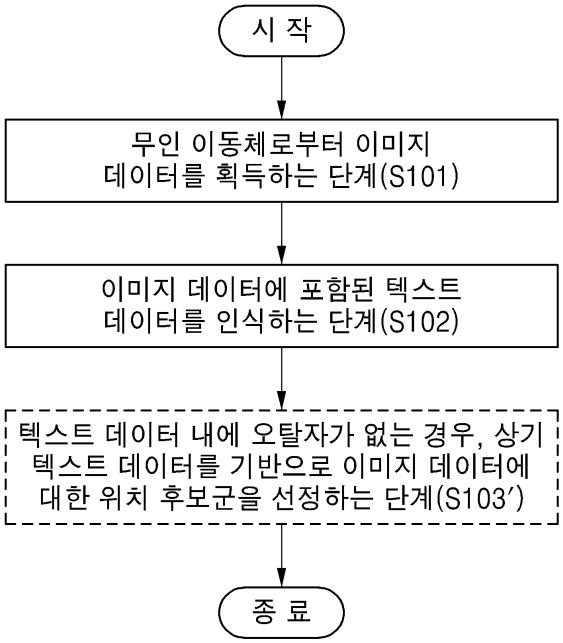
도면7


이미지 데이터
인쇄물 데이터
텍스트
데이터
어린이
영어학원
텍스트
영역
100일의 기적
필터링

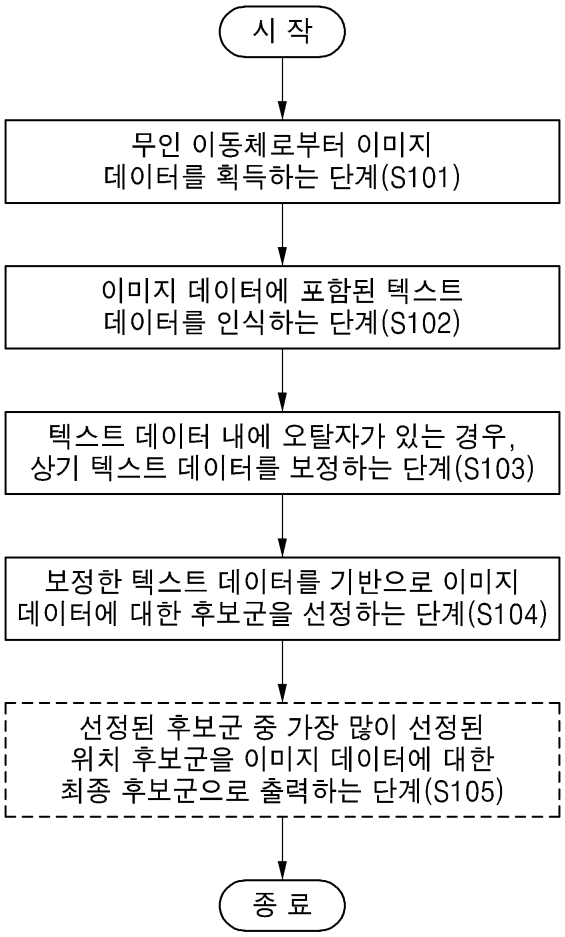
도면8


후처리부(150)
Symspell
알고리즘
그림끔나무반
끔나무
오탈자 감지(①)
말뭉치 데이터
편집 거리: 1
끔나무
음나무
꿈나무
끝나무
편집 거리: 2
끔나무
꾼나무
굼나무
쓴나무
보정 후보군 단어 선정(②)
"그림"
+
음나무
꿈나무
끝나무
+
"반"
편집 거리 및 실내 장소명일 확률을
기반으로 텍스트 데이터 보정(③)

도면9


시 작
무인 이동체로부터 이미지
데이터를 획득하는 단계(S101)
이미지 데이터에 포함된 텍스트
데이터를 인식하는 단계(S102)
텍스트 데이터 내에 오탈자가 없는 경우, 상기
텍스트 데이터를 기반으로 이미지 데이터에
대한 위치 후보군을 선정하는 단계(S103′)
종 료

도면10



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 6

【변경전】

제1 항에 있어서,

상기 (c) 단계는,

상기 텍스트 데이터에 포함된 단어들을 Symspell 알고리즘을 기반으로 보정하되, 상기 텍스트 데이터에 포함된 단어와 기 설정된 편집 거리 내에 해당하는 적어도 하나 이상의 보정 후보 단어- 상기 보정 후보군 단어는, 상기 장치 내의 말뭉치 데이터에 저장되어 있는 데이터임-를 매칭하고, 상기 보정 후보 단어들 중 실내 장소명과 근접한 단어를 추출하여 보정하는 단계인,

실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법.

【변경후】

제1 항에 있어서,

상기 (c) 단계는,

상기 텍스트 데이터에 포함된 단어들을 Symspell 알고리즘을 기반으로 보정하되, 상기 텍스트 데이터에 포함된 단어와 기 설정된 편집 거리 내에 해당하는 적어도 하나 이상의 보정 후보 단어- 상기 보정 후보 단어는, 상기 장치 내의 말뭉치 데이터에 저장되어 있는 데이터임-를 매칭하고, 상기 보정 후보 단어들 중 실내 장소명과 근

접한 단어를 추출하여 보정하는 단계인,

실내 장소명 예측을 위한 딥러닝 모델 기반 광학 문자 인식 방법.