(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년02월02일
(11) 등록번호 10-2210404
(24) 등록일자 2021년01월26일

(51) 국제특허분류(Int. Cl.)
*G08G 5/00* (2006.01) *G06T 7/60* (2017.01)
(52) CPC특허분류
*G08G 5/0078* (2013.01)
*G06T 7/60* (2013.01)
(21) 출원번호 10-2019-0127130
(22) 출원일자 2019년10월14일
심사청구일자 2019년10월14일
(56) 선행기술조사문헌
JP2009015406 A*
JP2019132664 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
**국방과학연구소**
대전광역시 유성구 북유성대로488번길 160 (수남동)
**고려대학교 산학협력단**
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
**강재웅**
대전광역시 유성구 북유성대로488번길 160 (수남동)
**송현승**
대전광역시 유성구 북유성대로488번길 160 (수남동)
*(뒷면에 계속)*
(74) 대리인
**제일특허법인(유)**

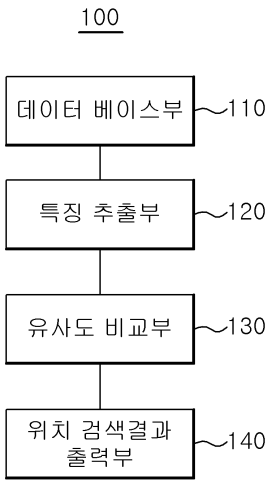
전체 청구항 수 : 총 6 항 심사관 : 김희주

(54) 발명의 명칭 **위치 정보 추출 장치 및 방법**

## (57) 요 약

본 발명은 촬영된 영상의 GPS 좌표나 DSM 등의 정보 없이 영상의 내용을 이용하여 위치를 추정함으로써 촬영된 영상의 위치를 정확하게 찾을 수 있도록 한 위치 정보 추출 장치 및 방법에 관한 것으로서, 위치 정보를 포함하는 지역 이미지를 복수 개 포함하는 데이터베이스부와, 타겟 이미지 및 상기 복수 개의 지역 이미지 각각으로부터 특징을 추출하도록 기 학습된 특징 추출부와, 상기 특징 추출부가 추출한 상기 복수 개의 지역 이미지 각각에 대한 특징과 상기 특징 추출부가 추출한 상기 타겟 이미지에 대한 특징 간의 유사도를 비교하는 유사도 비교부와, 상기 유사도 비교부가 비교한 결과에 기초해서 상기 복수 개의 지역 이미지 중 어느 하나를 선별하고, 상기 선별된 지역 이미지에 포함된 위치 정보를 이용해서 상기 타겟 이미지에 대한 위치 정보를 추출하는 위치 검색결과 출력부를 포함한다.

*대 표 도* - 도1



(52) CPC특허분류
*G06T 2207/20081* (2013.01)
*G06T 2207/30236* (2013.01)
(72) 발명자
**이성환**
서울특별시 성북구 안암로 145, 고려대학교 뇌공학과 (안암동5가)
**남우정**
서울특별시 성북구 안암로 145, 고려대학교 뇌공학과 (안암동5가)
**박성준**
서울특별시 성북구 안암로 145, 고려대학교 뇌공학과 (안암동5가)
**윤민섭**
서울특별시 성북구 안암로 145, 고려대학교 뇌공학과 (안암동5가)

# 명 세 서

## 청구범위

### 청구항 1

위치 정보를 포함하는 지역 이미지를 복수 개 포함하는 데이터베이스부와,

타겟 이미지 및 상기 복수 개의 지역 이미지 각각으로부터 특징을 추출하도록 기 학습된 특징 추출부와,

상기 특징 추출부가 추출한 상기 복수 개의 지역 이미지 각각에 대한 특징과 상기 특징 추출부가 추출한 상기 타겟 이미지에 대한 특징 간의 유사도를 비교하는 유사도 비교부와,

상기 유사도 비교부가 비교한 결과에 기초해서 상기 복수 개의 지역 이미지 중 어느 하나를 선별하고, 상기 선별된 지역 이미지에 포함된 위치 정보를 이용해서 상기 타겟 이미지에 대한 위치 정보를 추출하는 위치 검색결과 출력부를 포함하고,

상기 특징 추출부의 학습에 이용된 입력 데이터는, 소정의 지역을 제1 시점에 촬영한 기준 이미지, 상기 지역을 상기 제1 시점과 상이한 제2 시점에 촬영한 긍정 이미지 및 상기 지역과 상이한 지역을 촬영한 부정 이미지를 포함하고, 정답 데이터는, 상기 기준 이미지에 대한 기준 특징 벡터, 상기 긍정 이미지에 대한 긍정 특징 벡터 및 상기 부정 이미지에 대한 부정 특징 벡터를 포함하며,

상기 특징 추출부는,

상기 기준 특징 벡터와 상기 긍정 특징 벡터 간의 차이는 최소화시키고 상기 기준 특징 벡터와 상기 부정 특징 벡터 간의 차이는 최대화시키는 손실 함수를 이용해서 학습된 것인

위치 정보 추출 장치.

### 청구항 2

제 1 항에 있어서,

상기 타겟 이미지 및 상기 복수 개의 지역 이미지 각각은, 항공기에서 촬영된 영상인

위치 정보 추출 장치.

### 청구항 3

삭제

### 청구항 4

제1항에 있어서,

상기 특징 추출부가 추출한 특징은,

상기 특징 추출부에 포함된 장단기 메모리(long-short term memory, LSTM)를 통과함으로써, 가중치가 부여된 특징인

위치 정보 추출 장치.

### 청구항 5

위치 정보를 포함하는 지역 이미지를 복수 개 마련하는 단계와,

기 학습된 머신 러닝 모델이 상기 복수 개의 지역 이미지 각각으로부터 추출한 특징을 확보하는 단계와,

소정의 타겟 이미지를 입력받으면, 상기 기 학습된 머신 러닝 모델이 상기 타겟 이미지로부터 추출한 특징을 확보하는 단계와,

상기 복수 개의 지역 이미지 각각에 대한 특징과 상기 타겟 이미지에 대한 특징 간의 유사도를 비교하는

단계와,

상기 유사도에 대한 비교 결과에 기초해서 상기 복수 개의 지역 이미지 중 어느 하나를 선별하는 단계와,

상기 선별된 지역 이미지에 포함된 위치 정보를 이용해서 상기 타겟 이미지에 대한 위치 정보를 추출하는 단계를 포함하고,

상기 머신 러닝 모델의 학습에 이용된 입력 데이터는, 소정의 지역을 제1 시점에 촬영한 기준 이미지, 상기 지역을 상기 제1 시점과 상이한 제2 시점에 촬영한 긍정 이미지 및 상기 지역과 상이한 지역을 촬영한 부정 이미지를 포함하고, 정답 데이터는, 상기 기준 이미지에 대한 기준 특징 벡터, 상기 긍정 이미지에 대한 긍정 특징 벡터 및 상기 부정 이미지에 대한 부정 특징 벡터를 포함하며,

상기 머신 러닝 모델은,

상기 기준 특징 벡터와 상기 긍정 특징 벡터 간의 차이는 최소화시키고 상기 기준 특징 벡터와 상기 부정 특징 벡터 간의 차이는 최대화시키는 손실 함수를 이용해서 학습된 것인

위치 정보 추출 방법.

**청구항 6**

컴퓨터 프로그램을 저장하고 있는 컴퓨터 판독 가능 기록매체로서,

상기 컴퓨터 프로그램은, 프로세서에 의해 실행되면,

위치 정보를 포함하는 지역 이미지를 복수 개 마련하는 단계와,

기 학습된 머신 러닝 모델이 상기 복수 개의 지역 이미지 각각으로부터 추출한 특징을 확보하는 단계와,

소정의 타겟 이미지를 입력받으면, 상기 기 학습된 머신 러닝 모델이 상기 타겟 이미지로부터 추출한 특징을 확보하는 단계와,

상기 복수 개의 지역 이미지 각각에 대한 특징과 상기 타겟 이미지에 대한 특징 간의 유사도를 비교하는 단계와,

상기 유사도에 대한 비교 결과에 기초해서 상기 복수 개의 지역 이미지 중 어느 하나를 선별하는 단계와,

상기 선별된 지역 이미지에 포함된 위치 정보를 이용해서 상기 타겟 이미지에 대한 위치 정보를 추출하는 단계를 포함하는 방법을 상기 프로세서가 수행하도록 하기 위한 명령어를 포함하고,

상기 머신 러닝 모델의 학습에 이용된 입력 데이터는, 소정의 지역을 제1 시점에 촬영한 기준 이미지, 상기 지역을 상기 제1 시점과 상이한 제2 시점에 촬영한 긍정 이미지 및 상기 지역과 상이한 지역을 촬영한 부정 이미지를 포함하고, 정답 데이터는, 상기 기준 이미지에 대한 기준 특징 벡터, 상기 긍정 이미지에 대한 긍정 특징 벡터 및 상기 부정 이미지에 대한 부정 특징 벡터를 포함하며,

상기 머신 러닝 모델은,

상기 기준 특징 벡터와 상기 긍정 특징 벡터 간의 차이는 최소화시키고 상기 기준 특징 벡터와 상기 부정 특징 벡터 간의 차이는 최대화시키는 손실 함수를 이용해서 학습된 것인

컴퓨터 판독 가능한 기록매체.

**청구항 7**

컴퓨터 판독 가능한 기록매체에 저장되어 있는 컴퓨터 프로그램으로서,

상기 컴퓨터 프로그램은, 프로세서에 의해 실행되면,

위치 정보를 포함하는 지역 이미지를 복수 개 마련하는 단계와,

기 학습된 머신 러닝 모델이 상기 복수 개의 지역 이미지 각각으로부터 추출한 특징을 확보하는 단계와,

소정의 타겟 이미지를 입력받으면, 상기 기 학습된 머신 러닝 모델이 상기 타겟 이미지로부터 추출한 특징을 확보하는 단계와,

상기 복수 개의 지역 이미지 각각에 대한 특징과 상기 타겟 이미지에 대한 특징 간의 유사도를 비교하는 단계와,

상기 유사도에 대한 비교 결과에 기초해서 상기 복수 개의 지역 이미지 중 어느 하나를 선별하는 단계와,

상기 선별된 지역 이미지에 포함된 위치 정보를 이용해서 상기 타겟 이미지에 대한 위치 정보를 추출하는 단계를 포함하는 방법을 상기 프로세서가 수행하도록 하기 위한 명령어를 포함하고,

상기 머신 러닝 모델의 학습에 이용된 입력 데이터는, 소정의 지역을 제1 시점에 촬영한 기준 이미지, 상기 지역을 상기 제1 시점과 상이한 제2 시점에 촬영한 긍정 이미지 및 상기 지역과 상이한 지역을 촬영한 부정 이미지를 포함하고, 정답 데이터는, 상기 기준 이미지에 대한 기준 특징 벡터, 상기 긍정 이미지에 대한 긍정 특징 벡터 및 상기 부정 이미지에 대한 부정 특징 벡터를 포함하며,

상기 머신 러닝 모델은,

상기 기준 특징 벡터와 상기 긍정 특징 벡터 간의 차이는 최소화시키고 상기 기준 특징 벡터와 상기 부정 특징 벡터 간의 차이는 최대화시키는 손실 함수를 이용해서 학습된 것인

컴퓨터 프로그램.

## 발명의 설명

### 기 술 분 야

[0001] 본 발명은 항공 뷰(Air View)에서 촬영된 영상의 위치를 보다 정확하게 추정하기 위한 위치 정보 추출 장치 및 방법에 관한 것이다.

### 배 경 기 술

[0002] 최근 유무인 항공기의 발달 및 사진 기술의 발달로 인해, 고화질 항공 영상을 대량으로 쉽게 접할 수 있게 되었다. 이러한 대량의 항공 영상을 활용하기 위해서는, 각각의 영상이 어느 지역을 촬영했는지를 알아내는 것이 중요하다. 촬영 위치를 알기 위해서는 보통 항공 영상 이외에 GPS(Global Positioning System) 좌표나 DSM(Digital Surface Model) 등의 정보를 활용한다.

[0003] 그러나 GPS 좌표와 DSM 등의 정보는 기상과 지형지물의 상황, 전파 교란(JAMMING) 현상 등에 따라 정보가 부정확해지거나 사용할 수 없는 상황이 오기도 한다. 따라서 영상의 내용을 가지고 영상의 촬영 위치를 추정하면, 이와 같은 문제를 해결하여 성공적인 영상의 위치를 추정할 수 있다.

[0004] 한편, 내용 기반의 항공 영상 위치검색을 어렵게 하는 요소는 다음과 같다.

[0005] 첫 번째로, 항공 영상의 특징이다. 항공 영상은 일반적인 영상에 비해 크기가 크고 많은 객체가 존재하기 때문에 복잡한 계산이 필요하다. 따라서 기존의 전통적인 컴퓨터 비전 방법들은 좋은 성능을 기대하기 어렵다.

[0006] 두 번째로 대응하는 영상이 서로 다른 시간에 촬영되었다는 점이다. 따라서 영상의 밝기와 그림자가 달라지고, 사람이나 차량 따위의 유동적인 물체들의 정보 차이가 심하게 나타난다. 이러한 정보들의 차이는 내용 기반의 항공 영상 위치검색을 더욱 어렵게 한다.

## 발명의 내용

### 해결하려는 과제

[0007] 본 발명의 목적은 촬영된 영상의 GPS 좌표나 DSM 등의 정보 없이 영상의 내용을 이용하여 위치를 추정함으로써 촬영된 영상의 위치를 정확하게 찾을 수 있도록 한 위치 정보 추출 장치 및 방법을 제공하는 데에 있다.

### 과제의 해결 수단

[0008] 상술한 목적을 달성하기 위한 본 발명의 일 실시 예에 따른 위치 정보 추출 장치는 위치 정보를 포함하는 지역 이미지를 복수 개 포함하는 데이터베이스부와, 타겟 이미지 및 상기 복수 개의 지역 이미지 각각으로부터 특징을 추출하도록 기 학습된 특징 추출부와, 상기 특징 추출부가 추출한 상기 복수 개의 지역 이미지 각각에 대한 특징과 상기 특징 추출부가 추출한 상기 타겟 이미지에 대한 특징 간의 유사도를 비교하는 유사도 비교부와, 상

기 유사도 비교부가 비교한 결과에 기초해서 상기 복수 개의 지역 이미지 중 어느 하나를 선별하고, 상기 선별된 지역 이미지에 포함된 위치 정보를 이용해서 상기 타겟 이미지에 대한 위치 정보를 추출하는 위치 검색결과 출력부를 포함한다.

[0009] 이때, 상기 타겟 이미지 및 상기 복수 개의 지역 이미지 각각은, 항공기에서 촬영된 영상이고, 상기 특징 추출부의 학습에 이용된 입력 데이터는, 소정의 지역을 제1 시점에 촬영한 기준 이미지, 상기 지역을 상기 제1 시점과 상이한 제2 시점에 촬영한 긍정 이미지 및 상기 지역과 상이한 지역을 촬영한 부정 이미지를 포함하고, 정답 데이터는, 상기 기준 이미지에 대한 기준 특징 벡터, 상기 긍정 이미지에 대한 긍정 벡터 및 상기 부정 이미지에 대한 부정 특징 벡터를 포함하며, 상기 특징 추출부는, 상기 기준 특징 벡터와 상기 긍정 특징 벡터 간의 차이는 최소화시키고 상기 기준 특징 벡터와 상기 부정 특징 벡터 간의 차이는 최대화시키는 손실 함수를 이용해서 학습된 것이다.

[0010] 또한, 상기 특징 추출부가 추출한 특징은, 상기 특징 추출부에 포함된 장단기 메모리(long-short term memory, LSTM)를 통과함으로써, 가중치가 부여된다.

[0011] 한편, 본 발명의 일 실시 예에 따른 위치 정보 추출 방법은 위치 정보를 포함하는 지역 이미지를 복수 개 마련하는 단계와, 기 학습된 머신 러닝 모델이 상기 복수 개의 지역 이미지 각각으로부터 추출한 특징을 확보하는 단계와, 소정의 타겟 이미지를 입력받으면, 상기 기 학습된 머신 러닝 모델이 상기 타겟 이미지로부터 추출한 특징을 확보하는 단계와, 상기 복수 개의 지역 이미지 각각에 대한 특징과 상기 타겟 이미지에 대한 특징 간의 유사도를 비교하는 단계와, 상기 유사도에 대한 비교 결과에 기초해서 상기 복수 개의 지역 이미지 중 어느 하나를 선별하는 단계와, 상기 선별된 지역 이미지에 포함된 위치 정보를 이용해서 상기 타겟 이미지에 대한 위치 정보를 추출하는 단계를 포함한다.

## *발명의 효과*

[0013] 본 발명에 따르면, 전통적인 컴퓨터 비전 방법이 아닌 딥러닝 네트워크를 사용함으로써, 크기가 크고 많은 객체가 존재하는 항공이미지에 대한 분석을 좀 더 효과적으로 수행할 수 있다.

[0014] 또한, 전술한 본 발명의 학습된 딥러닝 네트워크에 의하면, 대응 영상 간에 서로 다른 시간대에 촬영된 차이를 최소화할 수 있다. 네트워크를 유사도 학습을 통해 학습시킴으로써, 영상의 밝기나 그림자 혹은 사람이나 자동차 같은 변화가 많은 요소의 특징 기여도는 줄어들게 되고, 건물이나 도로의 형태 등 잘 변하지 않는 요소들의 특징 기여도는 증가하게 된다. 이러한 특징들을 통해 영상 위치검색을 시행한다면, 촬영된 영상의 위치를 정확하게 찾을 수 있다.

## *도면의 간단한 설명*

[0015] 도 1은 본 발명의 실시예에 따른 위치 정보 추출 장치를 개략적으로 나타낸 구성도이고,

도 2는 도 1의 특징 추출부를 개략적으로 나타낸 구성도이며,

도 3은 유사도 기반의 학습 네트워크에서 튜플 구조 이미지의 예시를 나타낸 도면이고,

도 4는 도 1의 특징 추출부에서 삼중 네트워크를 나타낸 예시도이며,

도 5는 도 2의 유사도 기반의 학습 네트워크에서 학습의 결과 예시를 나타낸 도면이고,

도 6은 일반적인 순환 신경망 구조(왼쪽)와 양방향 순환 신경망 구조(오른쪽)를 각각 나타낸 도면이며,

도 7은 입력 이미지를 전분류 학습과 유사도 및 주의 기법 학습의 특징 추출 예시를 나타낸 도면이고,

도 8은 입력 영상에 대한 검색 결과의 예시를 나타낸 도면이며,

도 9는 본 발명의 실시예에 따른 위치 정보 추출 방법을 개략적으로 나타낸 흐름도이다.

## *발명을 실시하기 위한 구체적인 내용*

[0016] 먼저, 본 명세서 및 청구범위에서 사용되는 용어는 본 발명의 다양한 실시 예들에서의 기능을 고려하여 일반적인 용어들을 선택하였다. 하지만, 이러한 용어들은 당 분야에 종사하는 기술자의 의도나 법률적 또는 기술적 해석 및 새로운 기술의 출현 등에 따라 달라질 수 있다. 또한, 일부 용어는 출원인이 임의로 선정한 용어일 수 있다. 이러한 용어에 대해서는 본 명세서에서 정의된 의미로 해석될 수 있으며, 구체적인 용어 정의가 없으면 본

명세서의 전반적인 내용 및 당해 기술 분야의 통상적인 기술 상식을 토대로 해석될 수도 있다.

[0017] 또한, 본 명세서에 첨부된 각 도면에 기재된 동일한 참조 번호 또는 부호는 실질적으로 동일한 기능을 수행하는 부품 또는 구성요소를 나타낸다. 설명 및 이해의 편의를 위해서 서로 다른 실시 예들에서도 동일한 참조번호 또는 부호를 사용하여 설명하도록 한다. 즉, 복수의 도면에서 동일한 참조 번호를 가지는 구성 요소를 모두 도시하고 있다고 하더라도, 복수의 도면들이 하나의 실시 예를 의미하는 것은 아니다.

[0018] 또한, 본 명세서 및 청구범위에서는 구성요소들 간의 구별을 위하여 '제1', '제2' 등과 같이 서수를 포함하는 용어가 사용될 수 있다. 이러한 서수는 동일 또는 유사한 구성 요소들을 서로 구별하기 위하여 사용하는 것이며, 이러한 서수 사용으로 인하여 용어의 의미가 한정 해석되어서는 안될 것이다. 일 예로, 이러한 서수와 결합된 구성 요소는 그 숫자에 의해 사용 순서나 배치 순서 등이 제한 해석되어서는 안된다. 필요에 따라서는, 각 서수들은 서로 교체되어 사용될 수도 있다.

[0019] 본 명세서에서 단수의 표현은 문맥상 명백하게 다름을 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, '포함하다' 또는 '구성하다' 등의 용어는 명세서 상에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성 요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0020] 또한, 본 발명의 실시 예에서, 어떤 부분이 다른 부분과 연결되어 있다고 할 때, 이는 직접적인 연결뿐 아니라, 다른 매체를 통한 간접적인 연결의 경우도 포함한다. 또한 어떤 부분이 어떤 구성 요소를 포함한다는 의미는, 특별히 반대되는 기재가 없는 한 다른 구성 요소를 제외하는 것이 아니라 다른 구성 요소를 더 포함할 수 있다는 것을 의미한다.

[0021] 이하, 첨부된 도면을 참조하여 본 발명을 더욱 구체적으로 설명하기로 한다.

[0022] 도 1은 본 발명의 실시예에 따른 위치 정보 추출 장치를 개략적으로 나타낸 구성도이고, 도 2는 도 1의 특징 추출부를 개략적으로 나타낸 구성도이다.

[0023] 본 발명의 실시예에 따른 위치 정보 추출 장치(100)는 도 1에 도시된 바와 같이, 위치 정보를 포함하는 지역 이미지를 복수개 포함하는 데이터 베이스부(110)와, 타겟 이미지와 상기 데이터 베이스부(110)에 포함된 복수개의 지역 이미지 각각으로부터 특징을 추출하도록 기 학습된 특징 추출부(120)와, 상기 특징 추출부(120)가 추출한 상기 복수 개의 지역 이미지 각각에 대한 특징과 상기 특징 추출부(120)가 추출한 상기 타겟 이미지에 대한 특징 간의 유사도를 비교하는 유사도 비교부(130)와, 상기 유사도 비교부(130)가 비교한 결과에 기초하여 상기 복수 개의 지역 이미지 중 어느 하나를 선별하고 상기 선별된 지역 이미지에 포함한 위치 정보를 이용해서 상기 타겟 이미지에 대한 위치 정보를 추출하는 위치 검색결과 출력부(140)를 포함한다.

[0024] 여기서, 상기 타겟 이미지 및 상기 복수 개의 지역 이미지 각각은, 항공기에서 촬영된 영상이다.

[0025] 상기 특징 추출부(120)의 학습에 이용된 입력 데이터는, 소정의 지역을 제1 시점에 촬영한 기준 이미지, 상기 지역을 상기 제1 시점과 상이한 제2 시점에 촬영한 긍정 이미지 및 상기 지역과 상이한 지역을 촬영한 부정 이미지를 포함하고, 정답 데이터는, 상기 기준 이미지에 대한 기준 특징 벡터, 상기 긍정 이미지에 대한 긍정 벡터 및 상기 부정 이미지에 대한 부정 특징 벡터를 포함하며, 상기 특징 추출부(120)는, 상기 기준 특징 벡터와 상기 긍정 특징 벡터 간의 차이는 최소화시키고 상기 기준 특징 벡터와 상기 부정 특징 벡터 간의 차이는 최대화시키는 손실 함수를 이용해서 학습된 것이다.

[0026] 상기 특징 추출부(120)가 추출한 특징은, 상기 특징 추출부(120)에 포함된 장단기 메모리(long-short term memory, LSTM)를 통과함으로써, 가중치가 부여된다.

[0027] 상기 특징 추출부(120)는 도 2에서와 같이, 타겟 이미지와 지역 이미지 간에 서로 다른 시간대에 촬영된 차이를 줄이기 위한 유사도 기반의 학습 네트워크(121)와 주의 기법 기반의 학습 모듈(122)로 구성되어 있다.

[0028] 상기 특징 추출부(120)는 딥 러닝(Deep Learning)으로 통칭되는 인공신경망 기반 인식/검출 시스템 등에서 구현될 수 있는데, 딥 러닝은 인공신경망을 이용하여 데이터를 군집화하거나 분류하는데 사용하는 기술로서, 수많은 데이터 속에서 패턴을 발견하여 인간이 사물을 구분하듯 컴퓨터가 데이터를 나누는 것을 말한다.

[0029] 이러한 딥 러닝 기술 중 CNN(Convolutional Neural Network) 기술은 인간의 시각을 처리하는 시신경 구조를 모사하여 만든 인공신경망 모델 중에서 가장 대표적인 기술로서, 합성곱 신경망이라고도 한다. 이미지를 이해하고

이로부터 고수준의 추상화된 정보를 추출하거나 새로운 질감을 가진 그림을 그리는 등과 같은 다양한 영상 처리 또는 컴퓨터 비전 분야 등에서 많이 연구되고 있으며, 비디오 영상 중에서 수천 가지의 서로 다른 물체를 정확하게 인식하고 이를 기반으로 새로운 물체를 인식하는 학습을 추가할 수도 있다.

[0030] 도 3은 유사도 기반의 학습 네트워크에서 튜플 구조 이미지의 예시를 나타낸 도면으로서, 항공 영상의 위치 검색을 위한 네트워크의 학습을 위해서는 도 3과 같이 튜플 형태의 데이터 집합이 필요하다. 1개의 튜플은 기준 영상, 긍정 영상, 부정 영상으로 구성된 3개의 영상으로 구성한다.

[0031] 먼저, 긍정 영상은 기준 영상과 정확하게 대응하는 장소를 담고 있지만, 다른 시간에 촬영된 영상으로 정의하며, 부정 이미지로는 기준 이미지와 다른 장소를 담고 있는 영상으로 정의한다.

[0032] 즉, 도 3에서와 같이 기준 영상은 A 장소로서 t1 의 시점에 대한 영상이고, 긍정 영상은 A 장소로서 t2의 시점에 대한 영상이며, 부정 영상은 A 장소와는 전혀 다른 B 장소의 영상이다.

[0033] 도 4는 도 1의 특징 추출부에서 삼중 네트워크를 나타낸 예시도이다.

[0034] 여기서, 도 3에서 정해진 3개의 기준, 긍정, 부정 영상은 도 4의 네트워크에 투입되어 특징을 추출하게 된다. 이때 네트워크의 구조는 3개의 흐름을 가지고 있다. 그러나 3개의 네트워크는 모두 매개변수를 공유하는 구조를 가지고 있기 때문에 1개의 네트워크를 사용하는 효과를 얻게 된다.

[0035] 이렇게 구성된 네트워크를 통과시켜 기준 영상 특징(FA), 긍정 영상 특징(Fp), 부정 영상 특징(FN)을 얻게 된다. 이후, 학습을 위해 긍정거리(Pd)와 부정거리(Nd)를 다음과 같이 정의한다.

*수학식 1*

[0036] $$P_d = ||F_A - F_P||_2$$

[0037] $$N_d = ||F_A - F_N||_2$$

[0038] 여기서, 상기 긍정거리(Pd)와 부정거리(Nd)를 정의하는 거리 개념은 기본적으로 두 영상을 얼마나 비슷하게 인식하느냐와 밀접한 관련이 있다. 특징 간의 거리가 가까운 경우, 두 가지 영상을 같은 특징으로 요약하여 같게 인식한다는 의미이다.

[0039] 반대로 특징 간의 거리가 먼 경우 두 가지 영상을 다른 특징으로 요약하고 있으며, 다르게 인식하고 있다고 말할 수 있다. 긍정거리(Pd)의 경우, 같은 지역 영상 간의 차이이므로 가까울수록 바람직하다. 반대로 부정거리(Nd)의 경우, 다른 지역의 영상 간의 차이이므로, 멀어질수록 바람직하다.

[0040] 도 5는 도 2의 유사도 기반의 학습 네트워크에서 학습의 결과 예시를 나타낸 도면이다.

[0041] 도 5에서와 같이, 기준 영상 특징(FA), 긍정 영상 특징(Fp), 부정 영상 특징(FN)은 유사도 학습을 통해 왼쪽 그림에서 오른쪽 그림으로 변화한다.

[0042] 이러한 개념을 이용하여, 다음과 같이 손실함수(L)를 정의한다.

*수학식 2*

[0043] $$L = P_d + \max(0, P_d - N_d + m)$$

[0044] 이후, 손실함수(L)를 0으로 만드는 방향으로 학습이 이루어지게 된다. 긍정거리(Pd)와 부정거리(Nd)가 0보다 크기 때문에 손실함수(L)을 0으로 만들기 위해서는 긍정거리(Pd)를 0으로, 긍정거리와 부정거리의 차(Pd -Nd + m)를 0으로 만드는 학습이 동시에 진행된다.

[0045] 먼저 긍정거리(Pd)가 0이 되는 학습은 기준 이미지의 특징과 긍정 이미지의 특징 간의 유클리드 거리 차이를 0으로 만드는 것을 목적으로 한다. 그 결과 특정 장소에 대해, 2개의 다른 시점에 찍힌 지역의 사진이 공통적인 내용 혹은 잘 변하지 않는 영상 내용(도로, 건물 등)을 특징으로 요약하게 학습된다.

[0046] 반대로 잘 변하는 영상 내용 (사람, 차)등은 특징에 대한 기여도가 떨어지게 된다. 또한, 긍정거리와 부정거리의 차(Pd -Nd + m)를 0으로 만드는 학습은 기준 영상에 대해, 긍정거리가 부정거리보다 m만큼 가까워지는 방향으로 학습되게 된다.

[0047] 그 결과 같은 지역 특징 간의 차가 다른 지역 간 특징 간의 차이보다 m만큼 작아지게 학습함으로써 같은 지역의 영상을 더 유사하게 판단한다.

[0048] 상기 주의 기법 기반 학습 모듈(122)은 항공 영상에는 한 장에 많은 객체가 존재하기 때문에 특징을 추출할 때, 어떤 특징을 더 중요하게 생각하는지에 따라서 검색 결과가 완전히 달라진다. 따라서 본 발명에서는 추출되는 특징들의 상대 중요도를 파악하여 중요한 특징에 가중치를 주는 주의(Attention) 기법 기반의 학습 모듈을 통해 추출한다.

[0049] 도 6은 일반적인 순환 신경망 구조(왼쪽)와 양방향 순환 신경망 구조(오른쪽)를 각각 나타낸 도면이다.

[0050] 도 6에서와 같이, 추출된 특징의 상대적인 중요도를 파악하는 방법으로 양방향 순환 신경망(Bidirectional Recurrent Neural Network)을 사용하며, Recurrent Neural Network의 구조 중 LSTM(Long short-Term Memory)을 사용한다.

*수학식 3*

[0051] $$F_t=\sigma(W_f[h_{t-1},x_t]+b_f)$$

[0052] $$i_t=\sigma(W_i[h_{t-1},x_t]+b_i)$$ : input gate

[0053] $$C'=\tanh(W_c[h_{t-1},x_t]+b_C)$$

[0054] $$C_t=f_t\odot C_{t-1}+i_t\odot C_t$$

[0055] $$o_t=\sigma(W_o[h_{t-1},x_t]+b_o)$$ : output gate

[0056] $$h_t=o_t\odot\tanh(C_t)$$

[0057] LSTM은 어떤 정보를 잃어버릴지(forget gate), 어떤 정보를 기억할지(input gate), 어떤 정보를 내보낼지(output gate)를 정하는 각각의 게이트(gate)가 존재한다.

[0058] 따라서 이러한 구조는 입력으로 주어지는 특징들의 상대적인 중요도를 파악하는데 적합한 구조라고 할 수 있다. 이전의 정보인 Cell state($C_t$-1)을 얼마나 반영할 것인지는 forget gate($F_t$)을 통해 조절하고, 현재 입력값($x_t$)과 이전의 출력값($h_{t-1}$)을 얼마나 반영할 것인지는 input gate(it)을 사용하여 조절한다. 여기서 Θ는 요소별 곱셈을 뜻하는 Hadamard product 연산자이다.

[0059] 각각의 gate에서는 가중치(W)을 곱하고, 편향치(b)를 더한 후 출력을 0에서 1 사이 값으로 반환하는 활성화 함수인 시그모이드 함수( σ ) 또는, -1에서 1 사이의 값으로 반환하는 쌍곡선 탄젠트 함수(tanh)를 사용하여 결과를 출력한다.

*수학식 4*

[0060] $$F''=F\times\sigma(LSTM_{forward}(F)+LSTM_{backward}(F))$$

[0061] 도 6에 도시된 바와 같이, 일반적인 순환신경망의 구조에서는 이전의 출력값($h_{t-1}$)을 한 방향으로만 전달하기 때문에 결국 모든 출력 값을 고려한 활성화 값을 구할 수 있는 위치는 마지막만 존재하게 된다.

[0062] 하지만 양방향 순환신경망의 경우 이전의 출력값($h_{t-1}$)을 단방향이 아닌 양방향으로 전달하기 때문에, 결국 모든

위치에서 전체를 고려한 상대적인 활성화 값을 계산할 수 있게 된다.

[0063] 따라서 상기 언급한 LSTM에 추출된 특징 F를 입력하고 양방향의 출력값을 각각 더하면 상대적 활성화 값을 구할 수 있고, 이 상대적 활성화 값을 중요도로 표현하기 위해 시그모이드 함수를 사용하여 0에서 1 사이 값으로 변환시킨 후 추출된 특징 F와 곱 합성을 해줌으로써 주어졌던 특징의 각 채널의 중요도를 반영한 주의(attention) 특징 F*을 얻게 된다.

[0064] 상기 유사도 비교부(130)는 훈련된 네트워크를 통하여, 사전에 데이터베이스부(110) 내의 영상들에 대한 특징들을 추출하여 저장한다. 촬영 영상이 투입되면, 영상을 훈련된 네트워크를 통하여 특징을 추출하여 유클리드 거리가 가장 가까운 영상을 찾아낸다.

[0065] 특징 p = (p1, p2,…, pn)와 특징 q = (q1, q2,…, qn)가 있을 때, 두 특징 사이의 거리는 다음과 같다.

*수학식 5*

[0066] $$||p-q||=\sqrt{(p-q)\cdot(p-q)}=\sqrt{||p||^2+||q||^2-2p\cdot q}$$

[0067] 도 7은 입력 이미지를 전분류 학습과 유사도 및 주의 기법 학습의 특징 추출 예시를 나타낸 도면이다.

[0068] 도 7에서와 같이, 입력 이미지를 전달받아 분류 학습을 통해 학습하여 특징을 추출하고 계속해서 유사도 및 주의 기법 학습을 통해 영상 특징을 추출한다.

[0069] 도 8은 입력 영상에 대한 검색 결과의 예시를 나타낸 도면이다.

[0070] 도 9는 본 발명의 실시예에 따른 위치 정보 추출 방법을 개략적으로 나타낸 흐름도이다.

[0071] 본 발명의 실시예에 따른 위치 정보 추출 방법은 도 9에 도시된 바와 같이, 위치 정보를 포함하는 지역 이미지를 복수개 마련한다(S110).

[0072] 여기서, 지역 이미지는 입력 이미지 촬영 이전 시점의 영상들로 이루어져 있으며, 각각의 이미지들은 어느 위치를 촬영한 것인지에 대한 위치 정보(GPS 정보, 장소 이름 등)를 담고 있다.

[0073] 이어서, 기 학습된 머신 러닝 모델이 상기 복수 개의 지역 이미지 각각으로부터 추출한 특징을 확보한다(S120).

[0074] 이어서, 소정의 타겟 이미지를 입력받으면, 상기 기 학습된 머신 러닝 모델이 상기 타겟 이미지로부터 추출한 특징을 확보한다(S130).

[0075] 이어서, 상기 복수 개의 지역 이미지 각각에 대한 특징과 상기 타겟 이미지에 대한 특징 간의 유사도를 비교한다(S140).

[0076] 이어서, 상기 유사도에 대한 비교 결과에 기초해서 상기 복수 개의 지역 이미지 중 어느 하나를 선별한다(S150).

[0077] 그리고 상기 선별된 지역 이미지에 포함된 위치 정보를 이용해서 상기 타겟 이미지에 대한 위치 정보를 추출한다(S160).

[0078] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다.

[0079] 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다.

[0080] 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0081] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0082] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0083] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

[0084] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

## *부호의 설명*

[0085] 110: 데이터 베이스부

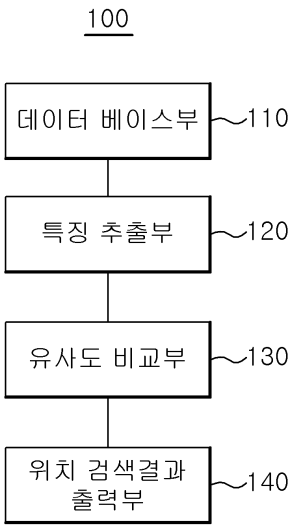
120: 특징 추출부

130: 유사도 비교부

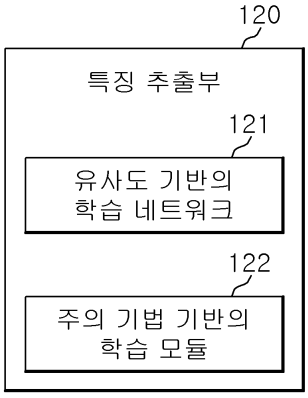
140: 위치 검색결과 출력부

## 도면

### 도면1


100
데이터 베이스부
110
특징 추출부
120
유사도 비교부
130
위치 검색결과 출력부
140

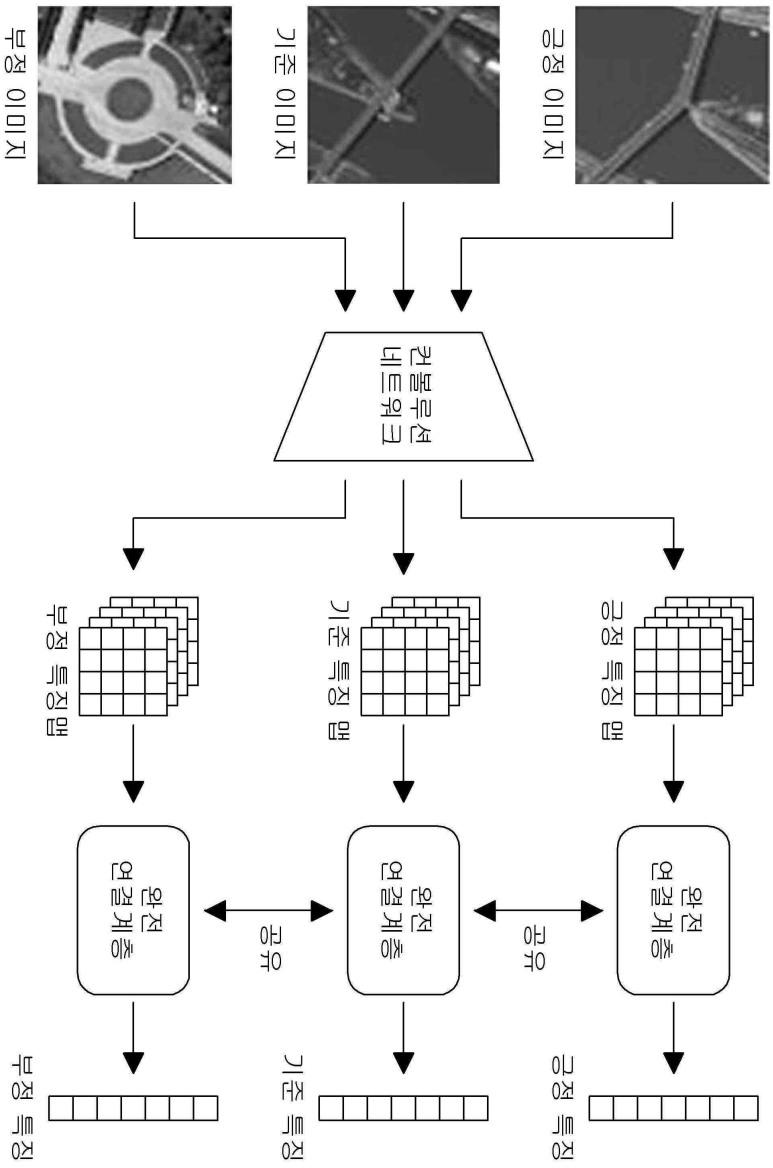
### 도면2


120
특징 추출부
121
유사도 기반의 학습 네트워크
122
주의 기법 기반의 학습 모듈

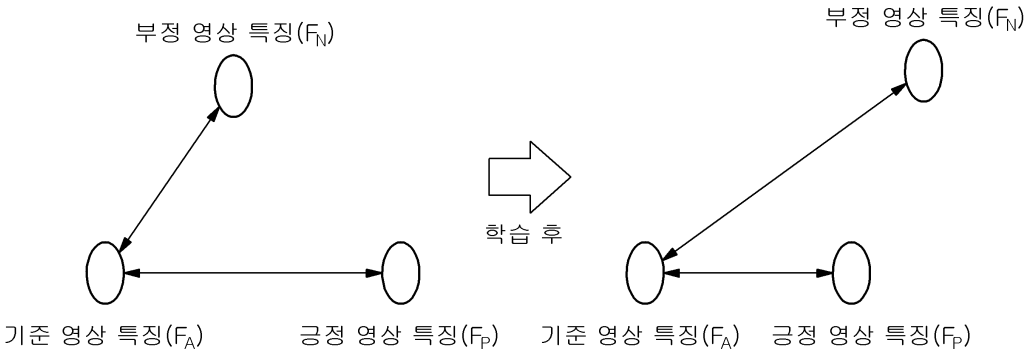
도면3


기준 이미지
A 장소, $t_1$시점
긍정 이미지
A 장소, $t_2$시점
부정 이미지
B 장소

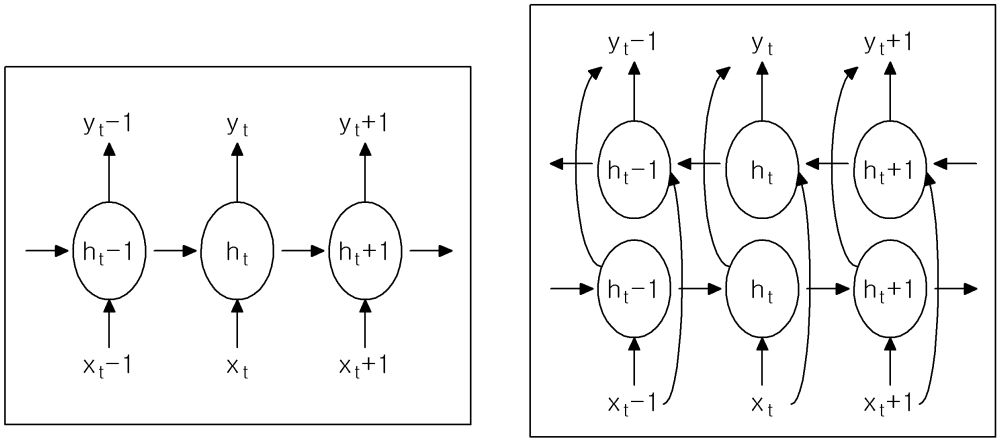
도면4


부정 이미지
기준 이미지
긍정 이미지
컨볼루션 네트워크
부정 특징맵
기준 특징 맵
긍정 특징 맵
완전 연결계층
공유
완전 연결계층
공유
완전 연결계층
부정 특징
기준 특징
긍정 특징

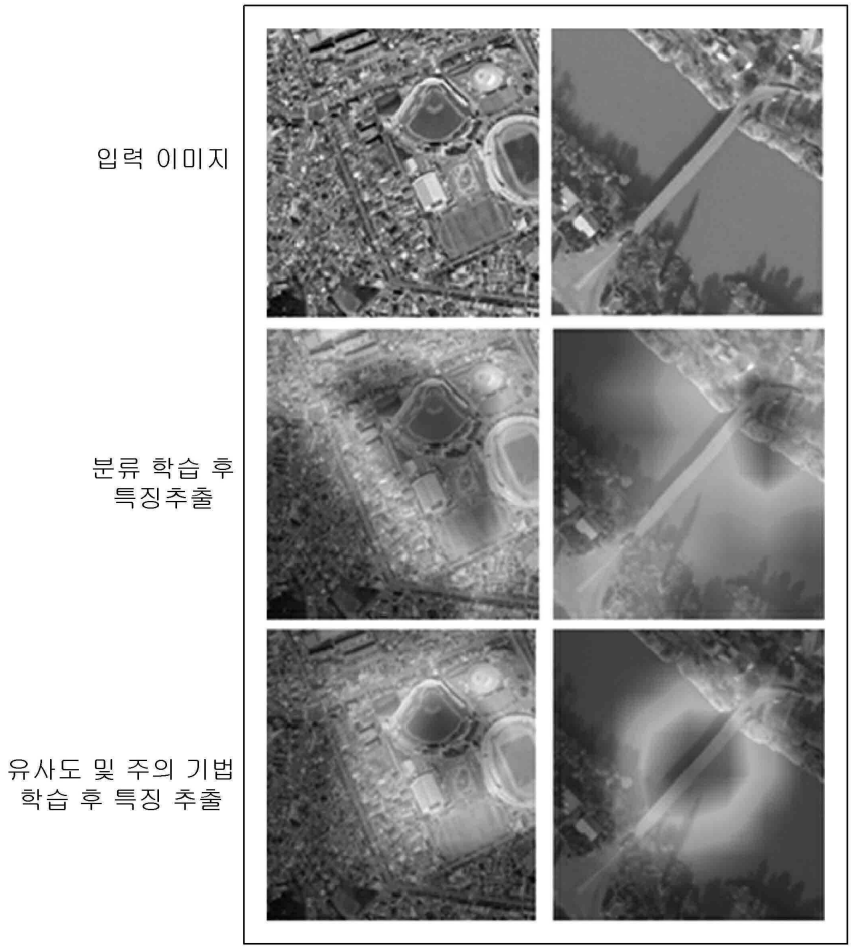
도면5


부정 영상 특징(F_N)
학습 후
부정 영상 특징(F_N)
기준 영상 특징(F_A)
긍정 영상 특징(F_P)
기준 영상 특징(F_A)
긍정 영상 특징(F_P)

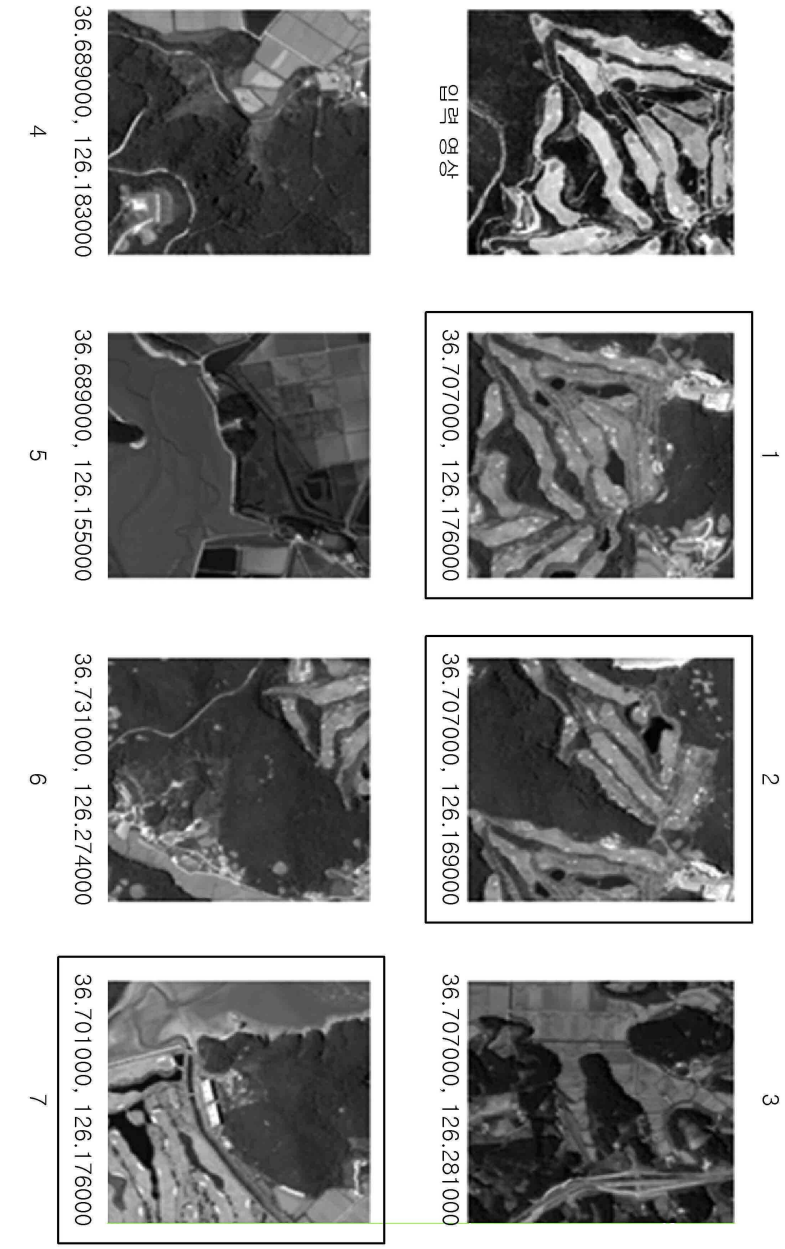
도면6


y_t−1
y_t
y_t+1
h_t−1
h_t
h_t+1
x_t−1
x_t
x_t+1

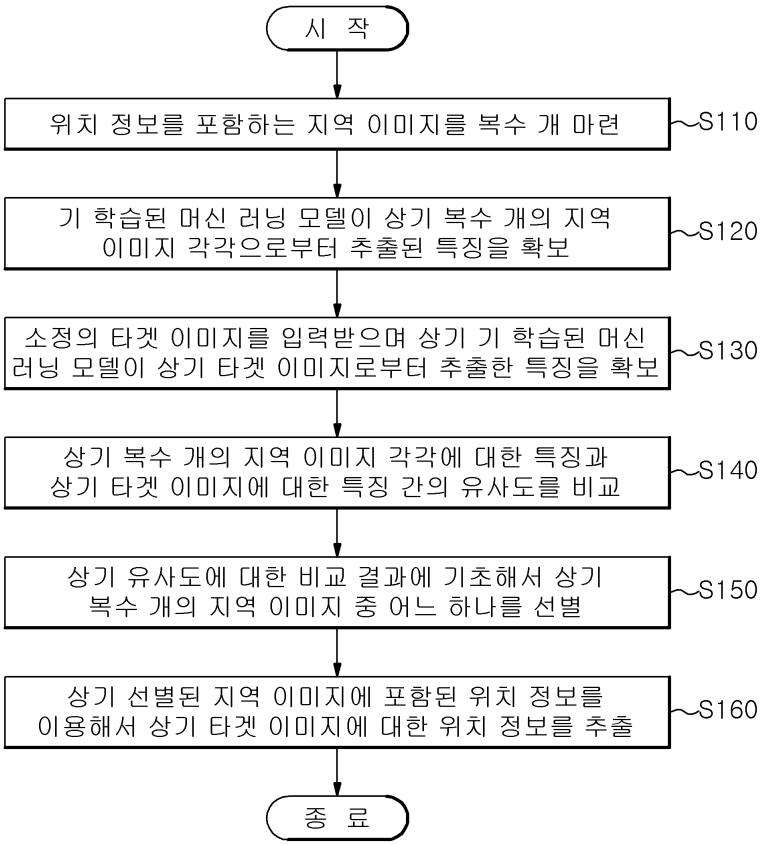
도면7



도면8


입력 영상
1
36.707000, 126.176000
2
36.707000, 126.169000
3
36.707000, 126.281000
4
36.689000, 126.183000
5
36.689000, 126.155000
6
36.731000, 126.274000
7
36.701000, 126.176000

도면9


시 작
위치 정보를 포함하는 지역 이미지를 복수 개 마련
S110
기 학습된 머신 러닝 모델이 상기 복수 개의 지역 이미지 각각으로부터 추출된 특징을 확보
S120
소정의 타겟 이미지를 입력받으며 상기 기 학습된 머신 러닝 모델이 상기 타겟 이미지로부터 추출한 특징을 확보
S130
상기 복수 개의 지역 이미지 각각에 대한 특징과 상기 타겟 이미지에 대한 특징 간의 유사도를 비교
S140
상기 유사도에 대한 비교 결과에 기초해서 상기 복수 개의 지역 이미지 중 어느 하나를 선별
S150
상기 선별된 지역 이미지에 포함된 위치 정보를 이용해서 상기 타겟 이미지에 대한 위치 정보를 추출
S160
종 료