



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0138558  
(43) 공개일자 2018년12월31일

- |   |   |
|---|---|
| <p>(51) 국제특허분류(Int. Cl.)<br/>G06T 7/00 (2017.01) G06K 9/62 (2006.01)<br/>G06T 7/11 (2017.01)</p> <p>(52) CPC특허분류<br/>G06T 7/0002 (2013.01)<br/>G06K 9/6267 (2013.01)</p> <p>(21) 출원번호 10-2018-0120470(분할)</p> <p>(22) 출원일자 2018년10월10일<br/>심사청구일자 2018년10월10일</p> <p>(62) 원출원 특허 10-2017-0078790<br/>원출원일자 2017년06월21일<br/>심사청구일자 2017년06월21일</p> | <p>(71) 출원인<br/>에스케이텔레콤 주식회사<br/>서울특별시 중구 을지로 65 (을지로2가)</p> <p>(72) 발명자<br/>이희열<br/>서울특별시 중구 을지로 65 SKT-타워<br/>양승지<br/>서울특별시 중구 을지로 65 SKT-타워<br/>김태완<br/>서울특별시 중구 을지로 65 SKT-타워</p> <p>(74) 대리인<br/>이철희</p> |
|---|---|

전체 청구항 수 : 총 8 항

(54) 발명의 명칭 객체 검출을 위한 영상분석 서버장치 및 방법

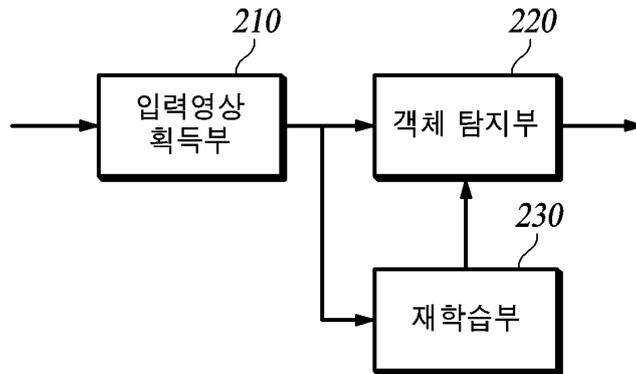
(57) 요약

객체 검출을 위한 영상분석 서버장치 및 방법을 개시한다.

본 발명의 일 실시예에 의하면, 입력영상으로부터 객체를 검출하여 인식하도록 학습된 객체 탐지부; 및 상기 입력영상 내의 움직임이 검출되지 않고 상기 객체 탐지부로부터 출력된 탐지결과에 오차가 기 설정된 임계 값을 초과하는 경우, 상기 입력영상으로부터 객체가 검출되지 않도록 상기 객체 탐지부를 재학습시키는 재학습부를 포함하는 영상분석 서버장치를 제공한다.

대표도 - 도2

120



(52) CPC특허분류

**G06T 7/11** (2017.01)

G06T 2207/20081 (2013.01)

G06T 2207/30232 (2013.01)

---

## 명세서

### 청구범위

#### 청구항 1

입력영상으로부터 객체를 검출하여 인식하도록 학습된 객체 탐지부; 및

상기 입력영상 내의 움직임이 검출되지 않고 상기 객체 탐지부로부터 출력된 탐지결과에 오차가 기 설정된 임계 값을 초과하는 경우, 상기 입력영상으로부터 객체가 검출되지 않도록 상기 객체 탐지부를 재학습시키는 재학습부

를 포함하는 영상분석 서버장치.

#### 청구항 2

제1항에 있어서,

상기 재학습부는,

상기 검출된 객체가 분류되는 하나 이상의 클래스 수와 상기 탐지결과에 신뢰도에 근거하여 상기 탐지결과에 오차를 계산하는, 영상분석 서버장치.

#### 청구항 3

제2항에 있어서,

상기 재학습부는,

목적하는 신뢰도 값인 0과 상기 탐지결과에 신뢰도 값 사이의 평균제곱오차(MSE: Mean Square Error)를 이용하여 상기 탐지결과에 오차를 계산하는, 영상분석 서버장치.

#### 청구항 4

제1항에 있어서,

상기 재학습부는,

상기 입력영상 내 객체가 존재하지 않음을 나타내는 정보 및 상기 입력영상을 학습데이터로 하여 상기 객체 탐지부를 재학습시키는, 영상분석 서버장치.

#### 청구항 5

입력영상을 획득하는 과정;

상기 입력영상으로부터 객체를 검출하여 인식하도록 학습된 객체 탐지기에 의해 탐지결과가 생성되는 과정; 및

상기 입력영상 내의 움직임이 검출되지 않고 상기 탐지결과에 오차가 기 설정된 임계 값을 초과하는 경우, 상기 입력영상으로부터 객체가 검출되지 않도록 상기 객체 탐지기를 재학습시키는 과정

을 포함하는 영상분석방법.

#### 청구항 6

제5항에 있어서,

상기 탐지결과에 오차는,

상기 검출된 객체가 분류되는 하나 이상의 클래스 수와 상기 탐지결과에 신뢰도에 근거하여 계산되는, 영상분석 방법.

#### 청구항 7

제6항에 있어서,

상기 탐지결과와 오차는,

목적하는 신뢰도 값인 0과 상기 탐지결과와 신뢰도 값 사이의 평균제곱오차(MSE: Mean Square Error)를 이용하여 계산되는, 영상분석방법.

**청구항 8**

제5항에 있어서,

상기 객체 탐지기를 재학습시키는 과정은,

상기 입력영상 내 객체가 존재하지 않음을 나타내는 정보 및 상기 입력영상을 학습데이터로 하여 상기 객체 탐지기를 재학습시키는, 영상분석방법.

**발명의 설명**

**기술 분야**

[0001] 본 발명은 객체 검출을 위한 영상분석 서버장치 및 방법에 관한 것이다.

**배경 기술**

[0002] 이 부분에 기술된 내용은 단순히 본 발명에 대한 배경 정보를 제공할 뿐 종래기술을 구성하는 것은 아니다.

[0003] CCTV(Closed Circuit Television)와 DVR(Digital Video Recorder)를 활용한 감시 시스템은 관리자가 육안으로 단순히 모니터링하고 영상을 저장하는 기술에서 더 나아가, 지능형 영상감시 시스템(Intelligent Surveillance System)으로 영역을 넓히고 있다.

[0004] 지능형 영상감시 시스템은 카메라로부터 입력되는 영상정보를 실시간으로 분석하여 객체를 탐지, 인식, 분류, 추적 등을 할 수 있는 형태의 시스템을 말한다. 지능형 영상감시 시스템은 객체가 보안과 관련된 이벤트를 발생시켰는지를 판단 및 분석하여 관리자에게 정보를 제공하거나 데이터 및 이벤트 내용을 저장한 후 사후예방 관리 및 검색의 효율성을 극대화시킬 수 있다.

[0005] CCTV 영상감시 시스템에서는 객체의 탐지, 인식, 분류, 추적 등을 위하여 컨볼루션 신경망(CNN: Convolutional Neural Network)과 같은 딥 러닝(Deep Learning) 기술을 채택하고 있다.

[0006] 컨볼루션 신경망(CNN) 기반의 기술을 이용할 경우, 각 영상 프레임마다 객체를 검출하고 검출된 객체에 대한 정보를 활용할 수 있으나, 실제로는 영상 내 객체가 존재하지 않는 경우에도 조명환경과 날씨변화 등의 요인으로 인하여 객체를 오검출하는 경우가 빈번히 발생할 수 있다.

[0007] 오검출로 인한 경보(False Alarm)는 위험 상황의 정확한 인지 및 신속한 대응을 저해한다. 특히, CCTV 영상은 유사한 배경영상이 지속되기 때문에 일단 오검출이 발생하게 되면 유사한 영상에 대해 계속하여 잘못된 경보가 발생하는 문제가 생긴다.

**발명의 내용**

**해결하려는 과제**

[0008] 본 발명의 실시예들은 CCTV와 같은 고정형 카메라 환경에서 객체의 오검출을 최소화할 수 있는 영상분석 서버장치 및 방법을 제공하고자 한다.

**과제의 해결 수단**

[0009] 본 발명의 일 실시예에 의하면, 입력영상으로부터 객체를 검출하여 인식하도록 학습된 객체 탐지부; 및 상기 입력영상 내의 움직임이 검출되지 않고 상기 객체 탐지부로부터 출력된 탐지결과와 오차가 기 설정된 임계 값을 초과하는 경우, 상기 입력영상으로부터 객체가 검출되지 않도록 상기 객체 탐지부를 재학습시키는 재학습부를 포함하는 영상분석 서버장치를 제공한다.

[0010] 또한, 본 발명의 다른 일 실시예에 의하면, 입력영상을 획득하는 과정; 상기 입력영상으로부터 객체를 검출하여

인식하도록 학습된 객체 탐지기에 의해 탐지결과가 생성되는 과정; 및 상기 입력영상 내의 움직임이 검출되지 않고 상기 탐지결과와 오차가 기 설정된 임계 값을 초과하는 경우, 상기 입력영상으로부터 객체가 검출되지 않도록 상기 객체 탐지기를 재학습시키는 과정을 포함하는 영상분석방법을 제공한다.

**발명의 효과**

[0011] 이상에서 설명한 바와 같이 본 발명의 실시예들에 의하면, 입력영상 내 움직임 정보를 이용하여 객체 검출 및 인식을 위한 신경망의 재학습 여부를 결정함으로써 영상감시시스템에서 지속적으로 오탐지가 발생하는 것을 방지하거나 최소화할 수 있는 효과가 있다.

**도면의 간단한 설명**

[0012] 도 1은 본 발명의 실시예에 따른 영상감시 시스템의 블록도이다.  
 도 2는 본 발명의 실시예에 따른 영상분석 서버장치의 블록도이다.  
 도 3은 본 발명의 실시예에 따른 영상분석 방법을 나타내는 흐름도이다.  
 도 4는 본 발명의 실시예에 따른 영상분석 방법의 일례를 나타내는 흐름도이다.

**발명을 실시하기 위한 구체적인 내용**

[0013] 이하, 본 발명의 일부 실시예들을 예시적인 도면을 통해 상세하게 설명한다. 각 도면의 구성요소들에 참조부호를 부가함에 있어서, 동일한 구성요소들에 대해서는 비록 다른 도면상에 표시되더라도 가능한 한 동일한 부호를 가지도록 하고 있음에 유의해야 한다. 또한, 본 발명을 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략한다.

[0014] 또한, 본 발명의 구성 요소를 설명하는 데 있어서, 제 1, 제 2, A, B, (a), (b) 등의 용어를 사용할 수 있다. 이러한 용어는 그 구성 요소를 다른 구성 요소와 구별하기 위한 것일 뿐, 그 용어에 의해 해당 구성 요소의 본질이나 차례 또는 순서 등이 한정되지 않는다. 명세서 전체에서, 어떤 부분이 어떤 구성요소를 '포함', '구비' 한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 '...부', '모듈' 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.

[0016] 도 1은 본 발명의 실시예에 따른 영상감시 시스템(100)의 블록도이다. 영상감시 시스템(100)은 영상촬영 장치(110), 영상분석 서버장치(120) 및 모니터링 장치(130)를 포함한다.

[0017] 영상촬영 장치(110)는 영상분석 서버장치(120)로 입력될 영상을 실시간으로 생성한다. 영상촬영 장치(110)는 CCTV용 카메라와 같은 고정형 카메라로 구현될 수 있으나, 반드시 이에 한정되는 것은 아니고 영상분석 서버장치(120)의 입력영상을 생성할 수 있는 모든 장치를 포함할 수 있다.

[0018] 영상분석 서버장치(120)는 영상촬영 장치(110)로부터 수신한 입력영상을 실시간으로 분석하여 객체를 탐지 및 인식하고, 분석결과를 기초로 사전에 정의된 하나 이상의 이벤트를 감지한다.

[0019] 또한, 영상분석 서버장치(120)는 감지된 이벤트에 따른 정보를 모니터링 장치(130)로 전송한다. 영상분석 서버장치(120)는 객체를 탐지하고 인식하도록 미리 설계된 신경망을 학습(training)하고, 학습된 신경망을 기반으로 입력영상 내 객체를 탐지 및 인식한다.

[0020] 예를 들어, 신경망으로는 컨볼루션 신경망(CNN)이 이용될 수 있다. 컨볼루션 신경망(CNN)은 입력영상에 대해 컨볼루션 연산을 수행하는 하나 이상의 컨볼루션 층(Convolution Layer) 및 컨볼루션 층의 출력을 샘플링하는 하나 이상의 풀링 층(Pooling Layer)을 포함한다.

[0021] 다만, 컨볼루션 신경망(CNN)은 예시일 뿐 반드시 이에 한정되는 것은 아니며, 순환 신경망(RNN: Recurrent Neural Network) 또는 CNN 및 RNN의 조합 등 그 밖의 다양한 신경망이 이용될 수 있다.

[0022] 모니터링 장치(130)는 영상촬영 장치(110)로부터 촬영된 영상을 수신하여 실시간으로 디스플레이할 수 있다. 또한, 모니터링 장치(130)는 영상분석 서버장치(120)로부터 수신한 정보에 따라 경보 발생 등과 같은 적절한 이벤

트를 발생시킬 수 있다.

- [0024] 도 2는 본 발명의 실시예에 따른 영상분석 서버장치(120)의 블록도이다. 영상분석 서버장치(120)는 입력영상 획득부(210), 객체 탐지부(220) 및 재학습부(230)를 포함한다. 도 2에 도시한 각 구성요소는 하드웨어 칩으로 구현될 수 있으며, 또는 소프트웨어로 구현되고 마이크로프로세서가 각 구성요소에 대응하는 소프트웨어의 기능을 실행하도록 구현될 수도 있다.
- [0025] 입력영상 획득부(210)는 영상촬영 장치(110)로부터 입력영상을 수신하여 이를 캡처한다. 입력영상 획득부(210)는 캡처한 입력영상을 객체 탐지 및 인식 신경망의 입력층(input layer)의 크기로 조정(resizing)한다. 조정된 입력영상은 객체 탐지부(220)로 전달된다. 또한, 조정된 입력영상은 객체 탐지부(220)의 재학습(update)이 필요한 경우에 재학습부(230)에도 전달된다.
- [0026] 객체 탐지부(120)는 학습된 신경망을 기반으로 입력영상으로부터 객체를 검출하여 인식한다. 객체 탐지부(120)는 신경망에 포함된 여러 개의 층(예: Convolution Layer, Max-pooling Layer 등)에서의 계산을 통해 검출 후보를 도출한다. 객체 탐지부(120)는 검출 후보에 대하여 임계값(threshold)을 적용하여 최종 탐지결과를 생성 및 출력한다.
- [0027] 출력된 탐지결과는 객체의 위치정보, 객체의 클래스정보 및 객체의 신뢰도 중 적어도 하나를 포함한다. 여기서, 객체의 위치정보는 객체가 검출된 영역(또는 Box)에 대한 정보(예: 위치좌표, 영역의 폭과 높이)를 나타내며, 객체의 클래스정보는 객체가 어떤 클래스(예: 사람, 차량 등)로 분류되는지를 나타낸다.
- [0028] 또한, 객체의 신뢰도는 객체가 최종적으로 결정된 클래스에 속할 확률을 나타낸다. 예를 들어, 객체가 존재하지 않는 경우 이상적인 신뢰도 값은 0이 된다. 하나의 입력영상 내에 복수 개의 검출영역이 존재하는 경우, 각 검출영역마다 위와 같은 탐지결과가 출력될 수 있다.
- [0029] 도 2에는 도시하지 않았지만 영상분석 서버장치(120)는 탐지결과를 기반으로 상황을 판단하기 위한 구성요소를 더 포함할 수 있다. 상황 판단을 위한 구성요소(미도시)는 탐지된 객체를 이용하여 입력영상 내에서 발생한 상황을 감지할 수 있다.
- [0030] 예를 들어, 본 구성요소는 감지하고자 하는 상황이 사람이 나타나는 상황인 경우 사람이 탐지되는지를 판단하고, 사람이 차량에서 내리는 상황을 감지하고자 하는 경우에는 사람과 차량이 동시에 탐지되는지를 판단할 수 있다. 또한, 본 구성요소는 탐지된 각각의 객체의 신뢰도 값에 따라 상황 판단을 할 수도 있다.
- [0031] 재학습부(230)는 입력영상 내의 움직임 검출 여부 및 객체 탐지부(220)로부터 출력된 탐지결과의 오차에 근거하여 객체 탐지부(220)의 재학습이 필요한지 여부를 판단한다.
- [0032] 판단 결과, 재학습이 필요한 경우 재학습부(230)는 입력영상으로부터 객체가 검출되지 않도록 객체 탐지부(220)를 재학습시킨다. 객체 탐지부(220)를 재학습시킨다는 것은 신경망에 포함된 노드들 사이의 기 학습된 파라미터(또는 가중치)들을 업데이트시킨다는 것이다. 우선, 재학습부(230)가 재학습이 필요한지 여부를 판단하는 방법에 대하여 설명한다.
- [0033] 재학습부(230)는 재학습 필요여부를 판단하기 위하여, 입력영상 내의 움직임을 계산한다. 입력영상 내 움직임을 계산하는 방법으로는, 연속된 두 개의 영상 차이를 이용하거나, 가우시안 혼합 모델(GMM: Gaussian Mixture Model), VIBE 등을 기반으로 모델링한 배경영상의 차분을 이용하는 등의 다양한 방법이 이용될 수 있다.
- [0034] 또한, 재학습부(230)는 재학습 필요여부를 판단하기 위하여, 객체 탐지부(220)로부터 출력된 탐지결과의 오차(error)를 계산한다. 탐지결과의 오차는 입력영상에 대하여 목적하는 결과(즉, target output)와 객체 탐지부(220)에서 출력된 결과(즉, estimated output) 간의 차이(loss)를 의미한다.
- [0035] 탐지결과에 포함된 정보 중 신뢰도를 예로 들어 설명하면, 본 발명의 실시예에서는 배경영상에서 객체가 탐지되지 않는 것을 목적으로 하므로, 배경영상에 있어서 검출 객체에 대한 신뢰도 값은 0인 것이 이상적이다.
- [0036] 그러나, 실제로 출력된 탐지결과에 따르면 신뢰도 값이 0보다 큰 경우도 존재하므로 탐지결과의 오차가 발생하게 된다. 재학습부(230)는 탐지결과의 신뢰도 및 검출된 객체가 분류되는 하나 이상의 클래스의 수에 근거하여 탐지결과의 오차를 계산할 수 있다.
- [0037] 입력영상이 배경영상인 경우, 실제 출력된 신뢰도 값과 원하는 신뢰도 값과의 차이가 클수록 오차 역시 커지므

로, 탐지결과의 오차는 평균제곱오차(MSE: Mean Square Error)를 이용하여 계산될 수 있다.

[0038] 구체적으로, 목적하는 신뢰도 값인 0과 상기 탐지결과의 신뢰도 값 간의 평균제곱오차를 이용하여 탐지결과의 오차가 계산될 수 있다. 본 실시예에 따른 탐지결과의 오차를 계산하기 위한 연산식은 수학적 식 1과 같다.

**수학적 식 1**

$$L = \sqrt{\frac{1}{n} \sum_i Confidence(i)^2}$$

- [0039]
- [0040] 위 식에서, L은 탐지결과의 오차, n은 객체가 분류되는 클래스의 수, Confidence(i)는 입력영상 내 하나 이상의 검출영역 중 i번째 영역에 대한 신뢰도를 나타낸다.
- [0041] 재학습부(230)는 계산된 입력영상 내의 움직임과 탐지결과의 오차를 이용하여 객체 탐지부(220)의 재학습이 필요한지 여부를 판단한다. 재학습부(230)는 우선 계산된 입력영상 내의 움직임 정보를 이용하여 객체 탐지부(220)를 재학습시킬지 여부를 판단한다.
- [0042] 본 발명의 일 목적은 입력영상에 움직임이 없는 경우에 객체가 지속적으로 오검출되는 것을 방지하고자 하는 것이므로, 움직임이 있는 것으로 계산된 경우에는 탐지결과가 잘못된 것이 아닐 확률이 높다.
- [0043] 이와 반대로, 움직임이 없는 것으로 계산된 경우에는 입력영상이 배경일 확률이 높으므로 탐지결과의 신뢰도가 낮게 나오도록(예컨대, 0에 수렴하도록) 신경망을 업데이트할 필요가 있다.
- [0044] 재학습부(230)는 입력영상 내의 움직임이 없는 것으로 판단되면, 탐지결과의 오차를 이용하여 객체 탐지부(220)의 재학습이 필요한지 여부를 판단한다. 구체적으로, 재학습부(230)는 탐지결과의 오차가 미리 설정한 임계 값 이하이면 객체 탐지부(220)의 재학습이 필요하지 않은 것으로 판단할 수 있다.
- [0045] 탐지결과의 오차가 미리 설정한 임계 값 이하라는 것은 배경영상에서 오탐지가 일어나지 않는 것을 의미하므로, 신경망을 업데이트할 필요가 없기 때문이다. 반대로, 탐지결과의 오차가 미리 설정한 임계 값 이상이면, 재학습부(230)는 객체 탐지부(220)의 재학습이 필요한 것으로 판단할 수 있다.
- [0046] 이는 입력영상 내 움직임이 없음에도 불구하고 신뢰도가 높은 객체가 탐지되는 것이므로 오탐지의 확률이 높다고 볼 수 있기 때문이다. 따라서, 이 경우에는 탐지결과의 신뢰도가 낮아지도록(예컨대, 0에 수렴하도록) 신경망을 업데이트할 필요가 있다.
- [0047] 재학습부(230)는 객체 탐지부(220)의 재학습이 필요한 것으로 결정되면, 입력영상 내 탐지객체가 존재하지 않음을 나타내는 정보를 포함하는 레이블 데이터(labeled data) 및 입력영상(즉, 오탐지된 배경영상)을 학습데이터로 하여 객체 탐지부(220)를 재학습시킨다(즉, 신경망을 업데이트한다). 레이블 데이터는 예컨대, 입력영상 내 탐지 객체가 없음을 나타내는 정보를 갖는 벡터(vector)가 될 수 있다.
- [0048] 신경망의 업데이트는 미리 수행된 신경망 학습에 사용된 방식과 같은 방식으로 수행된다. 구체적으로, 재학습부(230)는 역전파(back propagation) 기법을 이용하여 오탐지된 입력영상에 대한 탐지결과의 오차가 낮아지도록(또는 신뢰도가 낮아지도록) 객체 탐지부(220)를 재학습시킬 수 있다.
- [0049] 예를 들어, 재학습부(230)는 객체 탐지부(220) 내 신경망의 출력층(output layer)로부터 은닉층(hidden layer)을 거쳐 입력층(input layer)로 향하는 역방향으로 해당 탐지결과의 오차(또는 신뢰도)를 전파시킬 수 있다. 해당 탐지결과의 오차(또는 신뢰도)가 역방향으로 전파되는 과정에서, 오차(또는 신뢰도) 값이 감소되도록 노드들 사이의 파라미터(또는 연결 가중치)들이 업데이트될 수 있다.
- [0050] 재학습부(230)는 객체 탐지부(220)의 재학습 시, 하나의 입력영상에 대한 학습률(업데이트 비율)을 가변적으로 정할 수 있다. 학습률이 높은 경우 신경망은 빠르게 수렴하고, 반대로 학습률이 낮은 경우 신경망은 느리게 수렴한다. 신경망이 수렴하게 되면 더 이상 업데이트가 이루어지지 않는다.
- [0051] 재학습부(230)는 학습률을 조정하여 어느 정도의 빈도로 신경망을 업데이트할지 결정할 수 있다. 학습률을 높이면 신경망이 빠르게 수렴하기 때문에 적은 영상(예: 1초에 1프레임 또는 1분에 1프레임)으로 빠르게 신경망을

업데이트할 수 있다.

- [0053] 이하, 도 3 및 도 4를 참조하여 본 발명의 실시예에 따른 영상분석방법에 대하여 설명한다. 도 3은 본 발명의 실시예에 따른 영상분석방법을 나타내는 흐름도이다.
- [0054] 도 3을 참조하면, 우선 단계 S310에서 입력영상을 획득한다. 구체적으로, 영상촬영 장치로부터 전달된 입력영상을 캡처하고, 캡처한 입력영상을 객체 탐지 및 인식 신경망의 입력층(input layer)의 크기로 조정(resizing)한다.
- [0055] 입력영상이 획득되면, 입력영상으로부터 객체를 검출하여 인식하도록 학습된 객체 탐지기에 의해 탐지결과가 생성된다(S320). 객체 탐지기는 신경망에 포함된 여러 개의 층(예: Convolution Layer, Max-pooling Layer 등)에서의 계산을 통해 검출 후보를 도출한다. 객체 탐지기는 검출 후보에 대하여 임계 값(threshold)을 적용하여 최종 탐지결과를 생성 및 출력한다.
- [0056] 출력된 탐지결과는 객체의 위치정보, 객체의 클래스정보 및 객체의 신뢰도 중 적어도 하나를 포함한다. 여기서, 객체의 위치정보는 객체가 검출된 영역(또는 Box)에 대한 정보(예: 위치좌표, 영역의 폭과 높이)를 나타내며, 객체의 클래스정보는 객체가 어떤 클래스(예: 사람, 차량 등)로 분류되는지를 나타낸다.
- [0057] 객체의 신뢰도는 객체가 최종적으로 결정된 클래스에 속할 확률을 나타낸다. 예를 들어, 객체가 존재하지 않는 경우 이상적인 신뢰도 값은 0이 된다. 하나의 입력영상 내에 복수 개의 검출영역이 존재하는 경우, 각 검출영역마다 위와 같은 탐지결과가 출력될 수 있다.
- [0058] 단계 S330에서는 입력영상 내의 움직임 검출 여부 및 단계 S320의 탐지결과의 오차에 근거하여 객체 탐지기의 재학습이 필요한지 여부를 판단한다. 판단 결과, 재학습이 필요한 경우 단계 S340에서는 입력영상으로부터 객체가 검출되지 않도록 객체 탐지기를 재학습시킨다.
- [0059] 객체 탐지기를 재학습시킨다는 것은 신경망에 포함된 노드들 사이의 기 학습된 파라미터(또는 가중치)들을 업데이트시킨다는 것이다. 우선, 재학습이 필요한지 여부를 판단하는 과정에 대하여 설명한다.
- [0060] 재학습 필요여부를 판단하는 과정은, 입력영상 내의 움직임 및 탐지결과의 오차(error)를 계산하는 과정 및 계산 결과를 이용하여 재학습이 필요한지 판단하는 과정을 포함한다.
- [0061] 입력영상 내 움직임을 계산하는 방법으로는, 연속된 두 개의 영상 차이를 이용하거나, 가우시안 혼합 모델(GMM: Gaussian Mixture Model), VIBE 등을 기반으로 모델링한 배경영상의 차분을 이용하는 등의 다양한 방법이 이용될 수 있다.
- [0062] 탐지결과의 오차는 입력영상에 대하여 목적하는 결과(즉, target output)와 객체 탐지부(220)에서 출력된 결과(즉, estimated output) 간의 차이(loss)를 의미한다.
- [0063] 탐지결과에 포함된 정보 중 신뢰도를 예로 들어 설명하면, 본 발명의 실시예에서는 배경영상에서 객체가 탐지되지 않는 것을 목적으로 하므로, 배경영상에 있어서 검출 객체에 대한 신뢰도 값은 0인 것이 이상적이다. 그러나, 실제로 출력된 탐지결과에 따르면 신뢰도 값이 0보다 큰 경우도 존재하므로 탐지결과의 오차가 발생하게 된다.
- [0064] 탐지결과의 오차는 탐지결과의 신뢰도 및 검출된 객체가 분류되는 하나 이상의 클래스의 수에 근거하여 계산될 수 있다. 입력영상이 배경영상인 경우, 실제 출력된 신뢰도 값과 원하는 신뢰도 값과의 차이가 클수록 오차 역시 커지므로, 탐지결과의 오차는 평균제곱오차(MSE: Mean Square Error)를 이용하여 계산될 수 있다.
- [0065] 구체적으로, 목적하는 신뢰도 값인 0과 상기 탐지결과의 신뢰도 값 간의 평균제곱오차를 이용하여 탐지결과의 오차가 계산될 수 있다. 본 실시예에 따른 탐지결과의 오차를 계산하기 위한 연산식은 전술한 수학적 식 1과 같다.
- [0066] 입력영상 내의 움직임과 탐지결과의 오차를 계산한 이후, 계산결과를 이용하여 객체 탐지부(220)의 재학습이 필요한지 여부를 판단한다. 우선 계산된 입력영상 내의 움직임 정보를 이용하여 객체 탐지기를 재학습시킬지 여부를 판단한다.
- [0067] 본 발명의 일 목적은 입력영상에 움직임이 없는 경우에 객체가 지속적으로 오검출되는 것을 방지하고자 하는 것이므로, 움직임이 있는 것으로 계산된 경우에는 탐지결과가 잘못된 것이 아닐 확률이 높다. 그러나, 반대로 움직임이 없는 것으로 계산된 경우에는 입력영상이 배경일 확률이 높으므로 탐지결과의 신뢰도가 낮게 나오도록

(예컨대, 0에 수렴하도록) 신경망을 업데이트할 필요가 있다.

- [0068] 입력영상 내의 움직임이 없는 것으로 판단되면, 탐지결과와 오차를 이용하여 객체 탐지기의 재학습이 필요한지 여부를 판단한다. 구체적으로, 탐지결과와 오차가 미리 설정한 임계 값 이하이면 객체 탐지기의 재학습이 필요하지 않은 것으로 판단할 수 있다.
- [0069] 탐지결과와 오차가 미리 설정한 임계 값 이하라는 것은 배경영상에서 오탐지가 일어나지 않는 것을 의미하므로, 신경망을 업데이트할 필요가 없기 때문이다. 반대로, 탐지결과와 오차가 미리 설정한 임계 값 이상이면, 객체 탐지기의 재학습이 필요한 것으로 판단할 수 있다.
- [0070] 이는 입력영상 내 움직임이 없음에도 불구하고 신뢰도가 높은 객체가 탐지되는 것이므로 오탐지의 확률이 높다고 볼 수 있기 때문이다. 따라서, 이 경우에는 탐지결과와 신뢰도가 낮아지도록(예컨대, 0에 수렴하도록) 신경망을 업데이트할 필요가 있다.
- [0071] 정리하면, 입력영상 내의 움직임이 검출되지 않고 탐지결과와 오차가 임계 값 이상이면 객체 탐지기의 재학습이 필요한 것으로 판단한다.
- [0072] 단계 S330에서 객체 탐지기의 재학습이 필요한 것으로 결정되면, 입력영상 내 탐지객체가 존재하지 않음을 나타내는 정보를 포함하는 레이블 데이터(labeled data) 및 입력영상(즉, 오탐지된 배경영상)을 학습데이터로 하여 객체 탐지기를 재학습시킨다(즉, 신경망을 업데이트한다). 레이블 데이터는 예컨대, 입력영상 내 탐지 객체가 없음을 나타내는 정보를 갖는 벡터(vector)가 될 수 있다.
- [0073] 신경망의 업데이트는 미리 수행된 신경망 학습에 사용된 방식과 같은 방식으로 수행된다. 구체적으로, 역전파(back propagation) 기법을 이용하여 오탐지된 입력영상에 대한 탐지결과와 오차가 낮아지도록(또는 신뢰도가 낮아지도록) 객체 탐지기를 재학습시킬 수 있다.
- [0074] 예를 들어, 객체 탐지기 내 신경망의 출력층(output layer)로부터 은닉층(hidden layer)을 거쳐 입력층(input layer)로 향하는 역방향으로 해당 탐지결과와 오차(또는 신뢰도)를 전파시킬 수 있다. 해당 탐지결과와 오차(또는 신뢰도)가 역방향으로 전파되는 과정에서, 오차(또는 신뢰도) 값이 감소되도록 노드들 사이의 파라미터(또는 연결 가중치)들이 업데이트될 수 있다.
- [0075] 단계 S340의 재학습 시, 하나의 입력영상에 대한 학습률(업데이트 비율)을 가변적으로 정할 수 있다. 학습률이 높은 경우 신경망은 빠르게 수렴하고, 반대로 학습률이 낮은 경우 신경망은 느리게 수렴한다. 신경망이 수렴하게 되면 더 이상 업데이트가 이루어지지 않는다.
- [0076] 단계 S340에서는 학습률을 조정하여 어느 정도의 빈도로 신경망을 업데이트할지 결정할 수 있다. 학습률을 높이면 신경망이 빠르게 수렴하기 때문에 적은 영상(예: 1초에 1프레임 또는 1분에 1프레임)으로 빠르게 신경망을 업데이트할 수 있다.
- [0077] 도 4는 본 발명의 실시예에 따른 영상분석 방법의 일례를 나타내는 흐름도이다.
- [0078] 단계 S410에서 입력영상을 획득한 후, 입력영상 내 객체를 탐지 및 인식하여 탐지결과를 생성한다(S420). 그리고 생성된 탐지결과를 이용하여 상황 판단을 할 수 있다(S430). 즉, 탐지된 객체를 이용하여 입력영상 내에서 발생한 상황을 감지할 수 있다.
- [0079] 예를 들어, 감지하고자 하는 상황이 사람이 나타나는 상황인 경우 사람이 탐지되는지를 판단하고, 사람이 차량에서 내리는 상황을 감지하고자 하는 경우에는 사람과 차량이 동시에 탐지되는지를 판단할 수 있다. 또한, 탐지된 각각의 객체의 신뢰도 값에 따라 상황 판단을 할 수도 있다.
- [0080] 한편, 객체 탐지기를 재학습시킬지 여부를 판단하기 위하여 우선 입력영상 내 움직임을 계산하고(S440), 단계 S420에서 생성된 탐지결과와 오차를 추가적으로 계산한다(S450).
- [0081] 단계 S440에서 계산된 입력영상 내 움직임을 기초로 움직임의 검출 여부를 판단하고(S460), 판단결과 입력영상 내 움직임이 존재하지 않으면 단계 S450에서 계산한 탐지결과와 오차가 기 설정된 임계값보다 큰지 판단한다(S470). 판단 결과, 탐지결과와 오차가 기 설정된 임계값 이상이면 객체 탐지기를 재학습시킨다(S480).
- [0083] 도 3 및 도 4에서는 각 과정을 순차적으로 실행하는 것으로 기재하고 있으나, 반드시 이에 한정되는 것은 아니다. 다시 말해, 도 3 및 도 4에 기재된 과정을 변경하여 실행하거나 하나 이상의 과정을 병렬적으로 실행하는

것으로 적용 가능할 것이므로, 도 3 및 도 4는 시계열적인 순서로 한정되는 것은 아니다.

[0084] 도 3 및 도 4에 기재된 본 실시예에 따른 영상분석방법은 프로그램으로 구현되고 컴퓨터로 읽을 수 있는 기록매체에 기록될 수 있다. 본 실시예에 따른 영상분석방법을 구현하기 위한 프로그램이 기록되고 컴퓨터가 읽을 수 있는 기록매체는 컴퓨팅 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다.

[0085] 이상의 설명은 본 실시예의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 실시예가 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서, 본 실시예들은 본 실시예의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

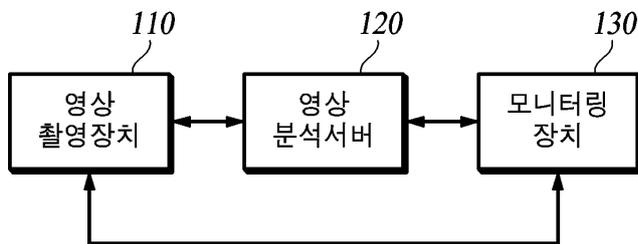
**부호의 설명**

- [0086] 100: 영상감시시스템    110: 영상촬영 장치  
 120: 영상분석서버    130: 모니터링 장치  
 210: 입력영상 획득부    220: 객체 탐지부  
 230: 재학습부

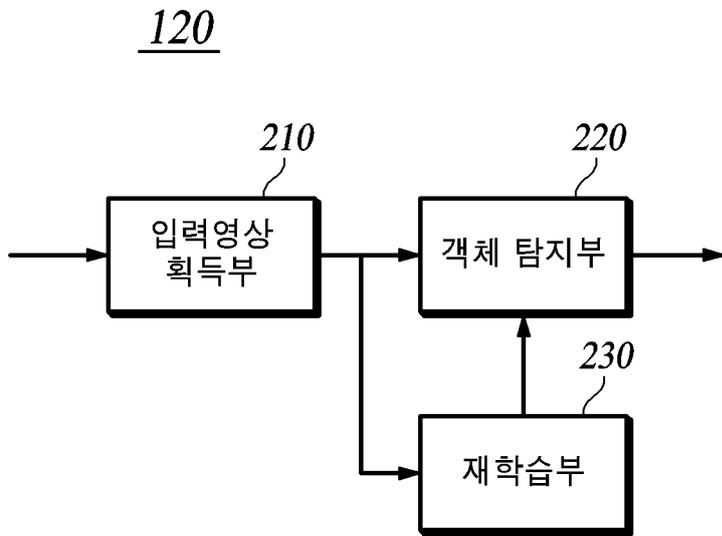
**도면**

**도면1**

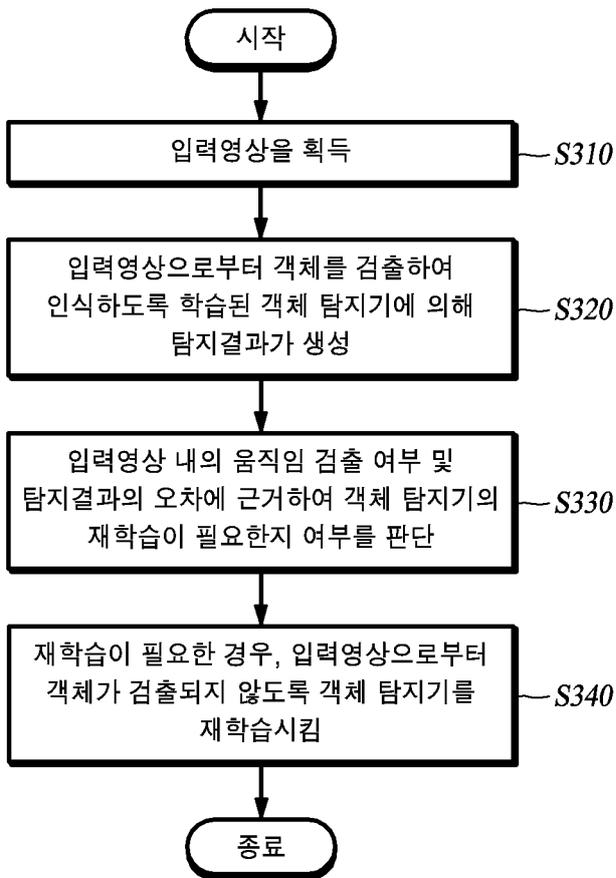
100



도면2



도면3



도면4

