



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0067682  
(43) 공개일자 2020년06월12일

(51) 국제특허분류(Int. Cl.)  
G06K 9/62 (2006.01) G06N 3/08 (2006.01)  
H04N 19/20 (2014.01) H04N 19/70 (2014.01)  
(52) CPC특허분류  
G06K 9/627 (2013.01)  
G06N 3/08 (2013.01)  
(21) 출원번호 10-2018-0154787  
(22) 출원일자 2018년12월04일  
심사청구일자 2018년12월04일

(71) 출원인  
이노덱 주식회사  
서울특별시 구로구 디지털로31길 61, 5층(구로동, 드림마크원데이터센터)  
(72) 발명자  
김태우  
서울특별시 성북구 정릉로21길 21, 101동 902호  
이성진  
경기도 광명시 오리로 801, 107동 1004호  
(74) 대리인  
김도형

전체 청구항 수 : 총 8 항

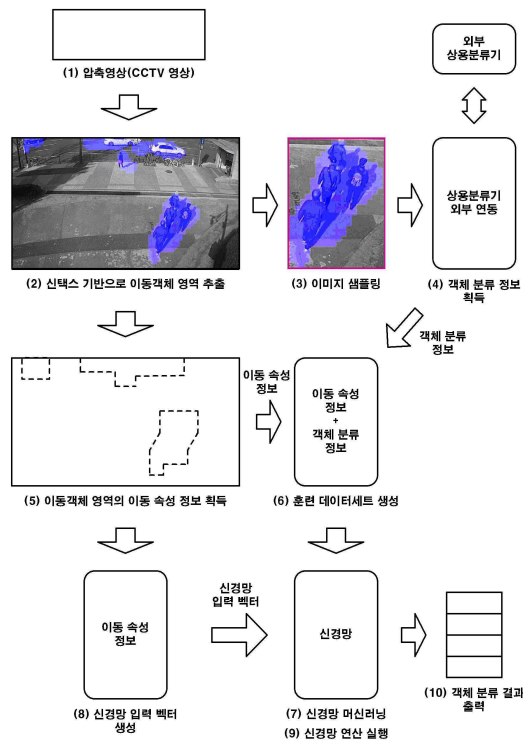
(54) 발명의 명칭 상용분류기 외부 연동 학습형 신경망을 이용한 압축영상에 대한 선택스 기반의 객체 분류 방법

(57) 요약

본 발명은 일반적으로 H.264 AVC 및 H.265 HEVC 등의 압축영상으로부터 객체 분류를 효과적으로 수행하는 기술에 관한 것이다. 더욱 상세하게는, 본 발명은 예컨대 CCTV 카메라가 생성하는 압축영상에 대해 종래기술처럼 복잡한 이미지 프로세싱을 통해 객체를 인식 및 분류하는 것이 아니라 압축영상 데이터를 파싱하여 얻어지는 선택스 정

(뒷면에 계속)

대표도 - 도4



보(예: 모션벡터, 코딩유형)를 활용하여 영상 내에서 무언가 유의미한 움직임이 존재하는 영역, 즉 이동객체 영역을 추출하고 그 이동객체 영역의 이동 속성 정보를 신경망에 입력하여 객체 분류 결과를 빠르게 획득하는 기술에 관한 것이다. 특히, 압축영상에서 추출되는 다수의 이동객체 영역에 대한 이미지를 외부의 상용분류기(예: Azure 클라우드 서비스)로 전달하여 객체 분류 정보를 제공받는 방식으로 다수의 이동객체 영역에 대해 이동 속성 정보와 객체 분류 정보의 조합을 획득한 후에 이들을 활용하여 신경망을 학습함으로써 다양한 촬영환경에서 얻어지는 대량의 데이터를 사람의 개입 없이 소프트웨어 처리를 통해 고속으로 신경망에 학습시키는 것이 가능하게 되어 객체 분류 결과의 신뢰성을 높일 수 있는 기술에 관한 것이다.

(52) CPC특허분류

*H04N 19/20* (2015.01)

*H04N 19/70* (2015.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 2014-3-00123

부처명 과학기술정보통신부

연구관리전문기관 정보통신기술진흥센터

연구사업명 정보통신 방송연구개발사업

연구과제명 (1세부) 실시간 대규모 영상 데이터 이해 예측을 위한 고성능 비주얼 디스커버리 플랫폼

개발

기여율 1/1

주관기관 한국전자통신연구원

연구기간 2014.04.01 ~ 2024.02.29

## 명세서

### 청구범위

#### 청구항 1

압축영상의 비트스트림을 파싱하여 코딩 유닛에 대한 모션벡터 및 코딩유형을 획득하는 제 1 단계;  
 압축영상을 구성하는 복수의 영상 블록 별로 미리 설정된 시간동안의 모션벡터 누적값을 획득하는 제 2 단계;  
 상기 복수의 영상 블록에 대하여 상기 모션벡터 누적값을 미리 설정된 제 1 임계치와 비교하는 제 3 단계;  
 상기 제 1 임계치를 초과하는 모션벡터 누적값을 갖는 영상 블록을 이동객체 영역으로 마킹하는 제 4 단계;  
 상기 이동객체 영역에 대한 이동 속성 정보를 획득하는 제 5 단계;  
 상기 압축영상을 디코딩하여 상기 이동객체 영역의 이미지를 획득하는 제 6 단계;  
 상기 이동객체 영역의 이미지를 외부 상용분류기와 연동하여 객체 분류 정보를 획득하는 제 7 단계;  
 신경망 학습 모드에서, 상기 이동객체 영역에 대하여 상기 이동 속성 정보와 상기 객체 분류 정보의 조합을 훈련 데이터셋으로 설정하여 신경망에 대한 머신러닝을 수행하는 제 8 단계;  
 객체 식별 모드에서, 객체 분류 대상으로 설정된 이동객체 영역의 이동 속성 정보를 상기 신경망에 입력하여 신경망 연산 결과로부터 객체 분류 결과를 획득하는 제 9 단계;  
 를 포함하여 구성되는 상용분류기 외부 연동 학습형 신경망을 이용한 압축영상에 대한 신텍스 기반의 객체 분류 방법.

#### 청구항 2

청구항 1에 있어서,  
 상기 제 5 단계는 상기 압축영상을 구성하는 일련의 프레임 시퀀스로부터 획득되는 복수의 이동객체 영역을 분석하여 상기 이동객체 영역에 대한 이동 속성 정보를 획득하고,  
 상기 이동 속성 정보는 이동궤적 정보, 박스크기 정보, 순간속도 정보를 포함하여 구성되는 것을 특징으로 하는 상용분류기 외부 연동 학습형 신경망을 이용한 압축영상에 대한 신텍스 기반의 객체 분류 방법.

#### 청구항 3

청구항 1에 있어서,  
 상기 제 6 단계는 상기 압축영상에서 이동객체 영역이 마킹된 영상 프레임을 선별적으로 디코딩하여 상기 이동객체 영역의 이미지를 획득하는 것을 특징으로 하는 상용분류기 외부 연동 학습형 신경망을 이용한 압축영상에 대한 신텍스 기반의 객체 분류 방법.

#### 청구항 4

청구항 1에 있어서,  
 상기 제 4 단계와 상기 제 5 단계 사이에 수행되는,  
 상기 이동객체 영역을 중심으로 그 인접하는 복수의 영상 블록(이하, '이웃 블록'이라 함)을 식별하는 제 a 단계;  
 상기 복수의 이웃 블록에 대하여 상기 제 1 단계에서 획득된 모션벡터 값을 미리 설정된 제 2 임계치와 비교하

는 제 b 단계;

상기 복수의 이웃 블록 중에서 상기 제 b 단계의 비교 결과 상기 제 2 임계치를 초과하는 모션벡터 값을 갖는 이웃 블록을 이동객체 영역으로 추가 마킹하는 제 c 단계;

를 더 포함하여 구성되는 것을 특징으로 하는 상용분류기 외부 연동 학습형 신경망을 이용한 압축영상에 대한 신택스 기반의 객체 분류 방법.

#### 청구항 5

청구항 1에 있어서,

상기 제 4 단계와 상기 제 5 단계 사이에 수행되는,

상기 이동객체 영역을 중심으로 그 인접하는 복수의 영상 블록(이하, '이웃 블록'이라 함)을 식별하는 제 a 단계;

상기 복수의 이웃 블록에 대하여 상기 모션벡터 누적값을 상기 제 1 임계치보다 작은 값으로 미리 설정된 제 2 임계치와 비교하는 제 b 단계;

상기 복수의 이웃 블록 중에서 상기 제 b 단계의 비교 결과 상기 제 2 임계치를 초과하는 모션벡터 누적값을 갖는 이웃 블록을 이동객체 영역으로 추가 마킹하는 제 c 단계;

를 더 포함하여 구성되는 것을 특징으로 하는 상용분류기 외부 연동 학습형 신경망을 이용한 압축영상에 대한 신택스 기반의 객체 분류 방법.

#### 청구항 6

청구항 4 또는 청구항 5에 있어서,

상기 제 c 단계 이후에 수행되는,

상기 복수의 이웃 블록 중에서 코딩유형이 인트라 픽처인 이웃 블록을 이동객체 영역으로 추가 마킹하는 제 d 단계;

를 더 포함하여 구성되는 것을 특징으로 하는 상용분류기 외부 연동 학습형 신경망을 이용한 압축영상에 대한 신택스 기반의 객체 분류 방법.

#### 청구항 7

청구항 6에 있어서,

상기 제 d 단계 이후에 수행되는,

상기 복수의 이동객체 영역에 대하여 인터플레이션을 수행하여 이동객체 영역으로 둘러싸인 미리 설정된 갯수 이하의 비마킹 영상 블록을 이동객체 영역으로 추가 마킹하는 제 e 단계;

를 더 포함하여 구성되는 것을 특징으로 하는 상용분류기 외부 연동 학습형 신경망을 이용한 압축영상에 대한 신택스 기반의 객체 분류 방법.

#### 청구항 8

하드웨어와 결합되어 청구항 1 내지 5 중 어느 하나의 항에 따른 상용분류기 외부 연동 학습형 신경망을 이용한 압축영상에 대한 신택스 기반의 객체 분류 방법을 실행시키기 위하여 매체에 저장된 컴퓨터프로그램.

### 발명의 설명

**기술 분야**

- [0001] 본 발명은 일반적으로 H.264 AVC 및 H.265 HEVC 등의 압축영상으로부터 객체 분류를 효과적으로 수행하는 기술에 관한 것이다.
- [0002] 더욱 상세하게는, 본 발명은 예컨대 CCTV 카메라가 생성하는 압축영상에 대해 종래기술처럼 복잡한 이미지 프로세싱을 통해 객체를 인식 및 분류하는 것이 아니라 압축영상 데이터를 파싱하여 얻어지는 선택스 정보(예: 모션 벡터, 코딩유형)를 활용하여 영상 내에서 무언가 유의미한 움직임이 존재하는 영역, 즉 이동객체 영역을 추출하고 그 이동객체 영역의 이동 속성 정보를 신경망에 입력하여 객체 분류 결과를 빠르게 획득하는 기술에 관한 것이다.
- [0003] 특히, 압축영상에서 추출되는 다수의 이동객체 영역에 대한 이미지를 외부의 상용분류기(예: Azure 클라우드 서비스)로 전달하여 객체 분류 정보를 제공받는 방식으로 다수의 이동객체 영역에 대해 이동 속성 정보와 객체 분류 정보의 조합을 획득한 후에 이들을 활용하여 신경망을 학습함으로써 다양한 촬영환경에서 얻어지는 대량의 데이터를 사람의 개입 없이 소프트웨어 처리를 통해 고속으로 신경망에 학습시키는 것이 가능하게 되어 객체 분류 결과의 신뢰성을 높일 수 있는 기술에 관한 것이다.

**배경 기술**

- [0005] 최근에는 범죄예방이나 사후증거 확보 등을 위해 CCTV를 이용하는 영상관제 시스템을 구축하는 것이 일반적이다. 지역별로 다수의 CCTV 카메라를 설치해둔 상태에서 이들 CCTV 카메라가 생성하는 영상을 모니터에 표시하고 만일을 위해 스토리지 장치에 저장해둔다. 범죄나 사고가 발생하는 장면을 관제 요원이 발견하게 되면 그 즉시 적절하게 대처하는 한편, 필요에 따라서는 사후증거 확보를 위해 스토리지에 저장되어 있는 영상을 검색하는 것이다.
- [0006] 그런데, CCTV 카메라의 설치 현황에 비해 관제 요원의 수는 매우 부족한 것이 현실이다. 이처럼 제한된 인원으로 영상 감시를 효과적으로 수행하려면 CCTV 영상을 모니터 화면에 단순 표시하는 것만으로는 충분하지 않다. 각각의 CCTV 영상에 존재하는 객체의 움직임을 감지하여 실시간으로 해당 영역에 무언가 추가 표시함으로써 효과적으로 발견되도록 처리하는 것이 바람직하다. 이러한 경우에 관제 요원은 CCTV 영상 전체를 균일한 관심도를 가지고 지켜보는 것이 아니라 객체 움직임이 있는 부분을 중심으로 CCTV 영상을 감시하면 된다.
- [0007] 특히, 영상관제 시스템이 CCTV 영상에서 객체의 움직임을 자동 감지할 뿐만 아니라 당해 객체를 분류(예: 사람, 자동차, 동물, 깃발)할 수 있다면 영상관제의 효율을 더욱 제고할 수 있다. 특정한 종류의 객체만을 선별적으로 감시할 수 있을 뿐만 아니라, 스토리지에 저장되어 있는 대규모의 CCTV 영상 데이터로부터 사후 증거를 빠르게 검출해낼 수 있게 된다.
- [0008] 최근에 설치되는 CCTV 카메라는 고해상도(예: Full HD) 및 고프레임(예: 초당 24프레임)의 제품이 채택되고 있기 때문에 네트워크 대역폭과 스토리지 공간의 부담을 고려하여 H.264 AVC 및 H.265 HEVC 등과 같은 고압축율의 복잡한 영상압축 기술이 채택되고 있다. CCTV 카메라 장치는 촬영 영상을 영상압축 기술에 따라 인코딩하여 생성한 압축영상을 제공하고, CCTV 영상을 활용하는 측에서는 해당 기술규격에 따라 역으로 압축영상에 대한 디코딩을 수행한다. 따라서, 영상압축 기술이 적용된 CCTV 영상에서 객체를 인식하고 그 객체가 어떠한 종류인지, 예를 들어 사람인지, 자동차인지, 동물인지, 깃발인지 분류하려면 종래에는 압축영상을 디코딩하여 재생영상, 즉 압축이 풀려있는 원래 영상을 얻은 후에 내용 분석을 위해 이미지를 처리하는 과정이 필요하였다.
- [0009] 이하에서는 [도 1]와 [도 2]를 참조하여 종래기술에서 CCTV 영상으로부터 객체를 인식하고 분류해내는 과정을 기술한다.
- [0010] [도 1]은 H.264 AVC 기술규격에 따른 동영상 디코딩 장치의 일반적인 구성을 나타내는 블록도이다. [도 1]을 참조하면, H.264 AVC에 따른 동영상 디코딩 장치는 구문분석기(11), 엔트로피 디코더(12), 역 변환기(13), 모션벡터 연산기(14), 예측기(15), 디블로킹 필터(16)를 포함하여 구성된다. 이들 하드웨어 모듈이 압축영상의 데이터를 순차적으로 처리함으로써 압축영상에서 압축을 풀고 원래의 영상 데이터를 복원해낸다. 이때, 구문분석기(11)는 압축영상의 코딩 유닛에 대해 모션벡터 및 코딩유형을 파싱해낸다. 이러한 코딩 유닛(coding unit)은 일반적으로는 매크로블록이나 서브 블록과 같은 영상 블록이다.
- [0011] [도 2]는 기존의 영상분석 솔루션에서 CCTV 영상으로부터 객체 분류를 수행하는 과정을 나타내는 순서도이다.

[도 2]를 참조하면, 종래기술에서는 압축영상을 H.264 AVC 및 H.265 HEVC 등에 따라 디코딩하고(S10), 재생영상의 프레임 이미지들을 작은 이미지, 예컨대 320x240 정도로 다운스케일 리사이징을 한다(S20). 이때, 다운스케일 리사이징을 하는 이유는 이후 영상분석 과정에서의 프로세싱 부담을 그나마 줄이기 위한 것이다. 그리고 나서, 리사이징된 프레임 이미지들에 대해 차영상(differentials)을 구한 후에 영상 분석을 통해 이동객체를 추출해낸다(S30). 그리고 나서, 일련의 프레임 이미지에 대한 영상 내용 분석을 통해 이동객체를 인식하여 객체 분류를 달성한다(S40).

[0012] 종래기술에서 이동객체를 추출 및 분류하려면 압축영상 디코딩, 다운스케일 리사이징, 영상 분석을 수행한다. 이들은 복잡도가 매우 높은 프로세스이고, 그로 인해 종래의 영상관제 시스템에서는 한 대의 영상분석 서버가 동시 처리할 수 있는 용량이 상당히 제한되어 있다. 현재 고성능의 영상분석 서버가 커버할 수 있는 최대 CCTV 채널은 통상 최대 16 채널이다. 다수의 CCTV 카메라가 설치되므로 영상관제 시스템에는 다수의 영상분석 서버가 필요하였고, 이는 비용 증가와 물리적 공간 확보의 어려움이라는 문제점을 유발하였다.

### 선행기술문헌

#### 비특허문헌

[0014] (비특허문헌 0001) 김태우, 김형현, 김평강 "이동객체의 메타데이터 필터링을 이용한 관심객체 추출 시스템 설계", 한국컴퓨터 종합학술대회 vol. 43 no. 12. pp. 1351-1355. 2016

(비특허문헌 0002) 김태우, 김형현, 김평강, 이유나 "이동체적 분석을 통해 쓰레기 무단 투기를 확률적으로 판단하는 지능형 시스템 설계", 한국컴퓨터종합학술대회 논문집 pp.805-807. 2017

(비특허문헌 0003) T. Kim, H. Kim, P. Kim, "Object Classifications in CCTV with Trajectory and Size Based on Neural Network" CUTE 2017

### 발명의 내용

#### 해결하려는 과제

[0015] 본 발명의 목적은 일반적으로 H.264 AVC 및 H.265 HEVC 등의 압축영상으로부터 객체 분류를 효과적으로 수행하는 기술을 제공하는 것이다.

[0016] 특히, 본 발명의 목적은 예컨대 CCTV 카메라가 생성하는 압축영상에 대해 종래기술처럼 복잡한 이미지 프로세싱을 통해 객체를 인식 및 분류하는 것이 아니라 압축영상 데이터를 파싱하여 얻어지는 선택스 정보(예: 모션벡터, 코딩유형)를 활용하여 영상 내에서 무언가 유의미한 움직임이 존재하는 영역, 즉 이동객체 영역을 추출하고 그 이동객체 영역의 이동 속성 정보를 신경망에 입력하여 객체 분류 결과를 빠르게 획득하는 기술을 제공하는 것이다.

[0017] 특히, 본 발명의 목적은 압축영상에서 추출되는 다수의 이동객체 영역에 대한 이미지를 외부의 상용분류기(예: Azure 클라우드 서비스)로 전달하여 객체 분류 정보를 제공받는 방식으로 다수의 이동객체 영역에 대해 이동 속성 정보와 객체 분류 정보의 조합을 획득한 후에 이들을 활용하여 신경망을 학습함으로써 다양한 촬영환경에서 얻어지는 대량의 데이터를 사람의 개입 없이 소프트웨어 처리를 통해 고속으로 신경망에 학습시키는 것이 가능하게 되어 객체 분류 결과의 신뢰성을 높일 수 있는 기술을 제공하는 것이다.

#### 과제의 해결 수단

[0019] 상기의 목적을 달성하기 위하여 본 발명에 따른 상용분류기 외부 연동 학습형 신경망을 이용한 압축영상에 대한 선택스 기반의 객체 분류 방법은, 압축영상의 비트스트림을 파싱하여 코딩 유닛에 대한 모션벡터 및 코딩유형을 획득하는 제 1 단계; 압축영상을 구성하는 복수의 영상 블록 별로 미리 설정된 시간동안의 모션벡터 누적값을 획득하는 제 2 단계; 복수의 영상 블록에 대하여 모션벡터 누적값을 미리 설정된 제 1 임계치와 비교하는 제 3 단계; 제 1 임계치를 초과하는 모션벡터 누적값을 갖는 영상 블록을 이동객체 영역으로 마킹하는 제 4 단계; 이

동객체 영역에 대한 이동 속성 정보를 획득하는 제 5 단계; 압축영상을 디코딩하여 이동객체 영역의 이미지를 획득하는 제 6 단계; 이동객체 영역의 이미지를 외부 상용분류기와 연동하여 객체 분류 정보를 획득하는 제 7 단계; 신경망 학습 모드에서, 이동객체 영역에 대하여 이동 속성 정보와 객체 분류 정보의 조합을 훈련 데이터 세트로 설정하여 신경망에 대한 머신러닝을 수행하는 제 8 단계; 객체 식별 모드에서, 객체 분류 대상으로 설정된 이동객체 영역의 이동 속성 정보를 신경망에 입력하여 신경망 연산 결과로부터 객체 분류 결과를 획득하는 제 9 단계;를 포함하여 구성된다.

[0020] 이때, 제 5 단계는 압축영상을 구성하는 일련의 프레임 시퀀스로부터 획득되는 복수의 이동객체 영역을 분석하여 이동객체 영역에 대한 이동 속성 정보를 획득하도록 구성될 수 있다. 또한, 이동 속성 정보는 이동객체 정보, 박스크기 정보, 순간속도 정보를 포함하여 구성될 수 있다. 또한, 제 6 단계는 압축영상에서 이동객체 영역이 마킹된 영상 프레임을 선별적으로 디코딩하여 이동객체 영역의 이미지를 획득하도록 구성될 수 있다.

[0021] 또한, 본 발명은 제 4 단계와 제 5 단계 사이에, 이동객체 영역을 중심으로 그 인접하는 복수의 영상 블록(이하, '이웃 블록'이라 함)을 식별하는 제 a 단계; 복수의 이웃 블록에 대하여 제 1 단계에서 획득된 모션 벡터 값을 미리 설정된 제 2 임계치와 비교하는 제 b 단계; 복수의 이웃 블록 중에서 제 b 단계의 비교 결과 제 2 임계치를 초과하는 모션벡터 값을 갖는 이웃 블록을 이동객체 영역으로 추가 마킹하는 제 c 단계; 복수의 이웃 블록 중에서 코딩유형이 인트라 픽처인 이웃 블록을 이동객체 영역으로 추가 마킹하는 제 d 단계; 복수의 이동객체 영역에 대하여 인터플레이션을 수행하여 이동객체 영역으로 둘러싸인 미리 설정된 갯수 이하의 비마킹 영상 블록을 이동객체 영역으로 추가 마킹하는 제 e 단계;를 더 포함하여 구성될 수 있다.

[0022] 또한, 본 발명은 제 4 단계와 제 5 단계 사이에, 이동객체 영역을 중심으로 그 인접하는 복수의 영상 블록(이하, '이웃 블록'이라 함)을 식별하는 제 a 단계; 복수의 이웃 블록에 대하여 모션벡터 누적값을 제 1 임계치보다 작은 값으로 미리 설정된 제 2 임계치와 비교하는 제 b 단계; 복수의 이웃 블록 중에서 제 b 단계의 비교 결과 제 2 임계치를 초과하는 모션벡터 누적값을 갖는 이웃 블록을 이동객체 영역으로 추가 마킹하는 제 c 단계; 복수의 이웃 블록 중에서 코딩유형이 인트라 픽처인 이웃 블록을 이동객체 영역으로 추가 마킹하는 제 d 단계; 복수의 이동객체 영역에 대하여 인터플레이션을 수행하여 이동객체 영역으로 둘러싸인 미리 설정된 갯수 이하의 비마킹 영상 블록을 이동객체 영역으로 추가 마킹하는 제 e 단계;를 더 포함하여 구성될 수 있다.

[0023] 한편, 본 발명에 따른 컴퓨터프로그램은 하드웨어와 결합되어 이상과 같은 상용분류기 외부 연동 학습형 신경망을 이용한 압축영상에 대한 신텍스 기반의 객체 분류 방법을 실행시키기 위하여 매체에 저장된 것이다.

**발명의 효과**

[0025] 본 발명에 따르면 압축영상을 파싱하여 얻은 신텍스 정보에 기반하여 객체 분류를 수행하므로 이미지 분석을 수행할 필요가 없어 영상분석 서버의 객체 분류 성능을 대폭 개선할 수 있는 장점이 있다.

[0026] 또한, 본 발명에 따르면 다양한 촬영환경에서 얻어지는 대량의 데이터를 사람의 개입 없이 소프트웨어 처리를 통해 고속으로 신경망에 학습시키는 것이 가능하여 객체 분류 결과의 신뢰성을 높일 수 있는 장점이 있다.

**도면의 간단한 설명**

- [0028] [도 1]은 동영상 디코딩 장치의 일반적인 구성을 나타내는 블록도.
- [도 2]는 종래기술에서 CCTV 영상으로부터 객체 분류를 수행하는 과정을 나타내는 순서도.
- [도 3]는 본 발명에 따라 상용분류기 외부 연동 학습형 신경망을 이용하여 압축영상에 대한 신텍스 기반의 객체 분류를 수행하는 전체 프로세스를 나타내는 순서도.
- [도 4]은 본 발명에서 상용분류기 외부 연동 학습형 신경망을 이용하여 압축영상에 대한 신텍스 기반의 객체 분류를 수행하는 개념을 나타내는 도면.
- [도 5]는 본 발명에서 이동객체 영역에 대한 이동 속성 정보의 일 예를 나타내는 도면.
- [도 6]은 본 발명에서 신경망이 출력하는 객체 분류 결과의 일 예를 나타내는 도면.
- [도 7]은 본 발명에서 압축영상으로부터 유효 움직임 영역을 검출하는 과정의 구현 예를 나타내는 순서도.

[도 8]은 CCTV 영상에 대해 유효 움직임 영역 검출 과정을 적용한 결과의 일 예를 나타낸 도면.

[도 9]는 본 발명에서 이동객체 영역에 대한 바운더리 영역을 검출하는 과정의 구현 예를 나타내는 순서도.

[도 10]은 [도 8]의 CCTV 영상 이미지에 대해 바운더리 영역 검출 과정을 적용한 결과의 일 예를 나타낸 도면.

[도 11]은 [도 10]의 CCTV 영상 이미지에 대해 인터플레이션을 통해 이동객체 영역을 정리한 결과의 일 예를 나타낸 도면.

[도 12]은 이동객체 영역에 Unique ID가 할당된 일 예를 나타내는 도면.

**발명을 실시하기 위한 구체적인 내용**

- [0029] 이하에서는 도면을 참조하여 본 발명을 상세하게 설명한다.
  
- [0031] [도 3]는 본 발명에 따라 압축영상으로부터 상용분류기 외부 연동 학습형 신경망을 이용하여 압축영상에 대한 신택스 기반의 객체 분류를 수행하는 전체 프로세스를 나타내는 순서도이고, [도 4]은 본 발명에서 상용분류기 외부 연동 학습형 신경망을 이용하여 압축영상에 대한 신택스 기반의 객체 분류를 수행하는 개념을 나타내는 도면이다.
  
- [0032] 본 발명에 따른 객체 분류 프로세스는 다수의 압축영상을 다루는 시스템, 예컨대 CCTV 영상관제 시스템 또는 CCTV 영상분석 시스템에서 영상분석 서버가 수행할 수 있다. 본 발명의 객체 분류 프로세스는 신경망(neural network)을 이용하는데, 영상분석 서버 내부에 소프트웨어 혹은 하드웨어로 신경망을 구현할 수도 있고, 혹은 외부의 클라우드 서버에 구현된 신경망을 오픈 API(Open Application Program Interface)를 통해 활용할 수도 있다.
  
- [0033] 먼저, [도 4]을 참조하여 본 발명에서 상용분류기 외부 연동 학습형 신경망을 이용하여 압축영상에 대해 신택스 기반으로 객체 분류를 수행하는 개념을 살펴본다.
  
- [0034] 객체분류를 수행하려면 먼저 압축영상으로부터 이동객체로 보이는 부분을 식별해내야 한다. 본 발명에서는 압축 영상의 비트스트림을 파싱하여 각 영상 블록에 대한 신택스 정보(syntax information)를 통해 이동객체 영역을 빠르게 추출한다. 영상 블록으로는 매크로블록(Macro Block) 및 서브블록(Sub Block) 등의 어느 하나 혹은 이들의 조합을 채택할 수 있고, 신택스 정보로는 모션벡터(Motion Vector)와 코딩유형(Coding Type)이 바람직하다. 이렇게 얻어진 이동객체 영역은 본 명세서에 첨부된 여러 이미지에서 확인되는 바와 같이 압축영상 내에 존재하는 이동객체의 경계선을 정확하게 반영하지는 못하지만 추출 처리속도가 빠르면서도 신뢰도가 비교적 높은 장점이 있다.
  
- [0035] 본 발명에서는 이동객체 영역의 이동 속성 정보를 신경망(neural network)에 입력하여 그 신경망 연산 결과에 의해 객체 분류 결과를 얻는다. 이동 속성 정보로는 이동궤적 정보(trajjectory), 박스크기 정보(size of bounding box), 순간속도 정보(instantaneous speed)의 조합이 가능하다. 이들 정보는 객체의 크기와 행동 특성에 직접적으로 관련된 것이어서 이들 정보와 객체 종류와의 연관성이 학습되어 있다면 이들 정보로부터 객체 분류하는 것이 가능하다.
  
- [0036] 이러한 접근 방식은 압축영상을 파싱하여 얻는 신택스 정보에 기초한 것이기 때문에 압축영상을 전체적으로 디코딩한 후에 복잡한 이미지 분석을 수행하는 종래기술에 비해 객체 분류 처리 속도가 대단히 빠르다.
  
- [0037] 이러한 방식의 객체 분류가 가능하려면 이동 속성 정보와 객체 종류 간의 연관성이 신경망에 미리 학습되어 있어야 한다. 그러나, 신경망을 충분한 정도로 학습시키는 것은 현실에서는 매우 곤란하다.
  
- [0038] 훈련 데이터(training data)가 많을수록 신경망의 객체 분류 결과의 신뢰도는 올라갈 것이므로 실제 필드에 적용할 수준의 신뢰도를 확보하려면 대규모의 훈련 데이터세트가 필요하다. 또한, 지역별로 이동객체의 종류와 이동 속성은 다르게 나타날 수 있으며, CCTV 카메라의 설치 환경도 각양각색일 것이어서 동일한 종류의 객체라 하더라도 CCTV 카메라마다 조금씩 상이한 이동 속성 정보를 나타낼 것이다. 따라서, 모든 지역과 모든 사용환경에서 신뢰도있게 사용하기 위해서는 다양한 환경에서 얻어진 훈련 데이터가 필요하다.
  
- [0039] 본 발명에서는 신택스 기반으로 이동객체 영역을 추출하는 기법에 외부의 사용분류기(예: Azure 클라우드 서비스)와의 연동을 통해 객체 분류 정보를 획득하는 것을 결합시키는 방식을 제안한다.

- [0040] 앞서 압축영상에 대해 선택스 기반으로 이동객체 영역을 추출하였으며, 이어서 압축영상을 디코딩하여 이동객체 영역의 이미지를 획득한다. 이때, 압축영상을 전체적으로 디코딩하는 것 보다는 이동객체 영역이 마킹된 영상 프레임을 선별적으로 디코딩하는 것이 속도 측면에서 바람직하다. 또한, 영상 프레임 내에서도 이동객체 영역에 해당하는 영상 데이터만 디코딩한다면 속도 측면에서 더욱 바람직하다. 이렇게 얻어진 이동객체 영역의 이미지를 외부의 사용분류기로 전달하여 이로부터 객체 분류 정보를 획득하는 것이다.
- [0041] 즉, 압축영상으로부터 선택스 기반으로 이동객체 영역을 추출하면, 일측으로는 해당 이동객체 영역에 대해 이동 속성 정보를 획득하는 한편, 타측으로는 이동객체 영역의 이미지를 외부 상용분류기에 전달하여 그 이동객체 영역에 대한 객체 분류 정보를 획득하는 것이다. 이를 통해, 해당 이동객체 영역과 관련하여 이동 속성 정보와 객체 분류 정보의 조합을 얻을 수 있으며, 이를 신경망에 대한 훈련 데이터로 삼는 것이다.
- [0042] 이러한 과정은 사람의 개입 없이 소프트웨어 처리로 수행될 수 있으므로 고속 작업이 가능하다. 여러 지역에 설치된 다수의 CCTV 카메라로부터 지속적으로 얻어지는 압축영상으로부터 다수의 이동객체 영역을 추출할 수 있는데, 이들 이동객체 영역에 대하여 선택스 정보를 처리하여 이동 속성 정보를 구성하고 외부의 상용분류기로부터 객체 분류 정보를 얻음으로써 대량의 훈련 데이터를 대단히 짧은 시간에 효과적으로 확보할 수 있다. 본 발명에서는 이렇게 확보된 훈련 데이터셋을 이용하여 신경망을 학습(머신러닝)시킨다. 신경망 이론(예: 딥러닝 이론)에 따르면, 이러한 머신러닝 과정을 통하여 신경망은 이동 속성 정보를 카테고리화(categorization)하는 성질을 갖추게 된다.
- [0043] 이처럼 신경망에 대한 학습을 수행한 이후에 해당 신경망을 이용하여 객체 분류를 수행하게 된다. 압축영상에서 선택스 기반으로 이동객체 영역을 추출하고, 일련의 프레임 시퀀스로부터 획득되는 복수의 이동객체 영역을 분석하여 그 이동객체 영역에 대한 이동 속성 정보를 획득한 후에, 이를 신경망에 입력하여 신경망 연산을 수행하면 된다. 신경망 연산 결과로부터 해당 이동객체 영역의 객체 분류 결과를 얻을 수 있다. 이러한 일련의 과정은 이미지 분석과는 달리 연산의 복잡도가 낮기 때문에 객체 분류 결과를 신속하게 얻을 수 있다.
- [0044] 다음으로, [도 3]를 참조하여 본 발명에 따라 압축영상으로부터 상용분류기 외부 연동 학습형 신경망을 이용하여 압축영상에 대한 선택스 기반의 객체 분류를 수행하는 과정을 살펴본다.
- [0045] 단계 (S100) : 먼저, 압축영상의 모션벡터에 기초하여 압축영상으로부터 실질적으로 의미를 인정할만한 유효 움직임을 검출하며, 이처럼 유효 움직임이 검출된 영상 영역을 이동객체 영역으로 설정한다.
- [0046] 이를 위해, H.264 AVC 및 H.265 HEVC 등의 동영상압축 표준에 따라서 압축영상의 코딩 유닛(coding unit)의 모션벡터와 코딩유형을 파악한다. 이때, 코딩 유닛의 사이즈는 일반적으로 64x64 픽셀 내지 4x4 픽셀 정도이며 설계자의 선택에 따라 다양하게 설정될 수 있다.
- [0047] 각 영상 블록에 대해 미리 설정된 일정 시간(예: 500 msec) 동안 모션벡터를 누적시키고, 그에 따른 모션벡터 누적값이 미리 설정된 제 1 임계치(예: 20)을 초과하는지 검사한다. 만일 그러한 영상 블록이 발견되면 해당 영상 블록에서 유효 움직임이 발견된 것으로 보고 이동객체 영역으로 마킹한다. 그에 따라, 모션벡터가 발생하였더라도 일정 시간동안의 누적값이 제 1 임계치를 넘지 못하는 경우에는 영상 변화가 미미한 것으로 추정하고 무시한다.
- [0048] 단계 (S200) : 앞의 (S100)에서 검출된 이동객체 영역에 대하여 그 주변 영역을 모션벡터와 코딩유형에 기초하여 검사함으로써 이들 이동객체 영역의 바운더리가 대략적으로 어디까지인지 확장해나간다. 이러한 과정을 통해서 앞서 (S100)에서 파편화된 영상 블록의 형태로 검출된 이동객체 영역을 서로 연결하여 유의미한 덩어리 형태를 만들어가는 결과를 얻는다.
- [0049] 앞의 (S100)에서는 엄격한 판단기준에 따라 영상 블록들을 선별함으로써 압축영상 내에서 이동객체임이 확실해 보이는 영상 블록을 검출하여 이동객체 영역으로 마킹하였다. 이번의 (S200)에서는 이렇게 (S100)에서 이동객체 영역으로 마킹되었던 영상 블록 주변에 위치하는 다른 영상 블록들을 검사한다. 이들을 본 명세서에서는 편의상 '이웃 블록'이라고 부른다. 이들 이웃 블록에 대해서는 앞서 (S100)에 적용하였던 판단기준에 비해 상대적으로 완화된 판단기준에 따라 이동객체 영역에 해당하는지 여부를 판단한다.
- [0050] 압축영상에서 매크로블록이나 서브블록 등은 매우 작은 사이즈이다. 따라서 CCTV 촬영영상과 같이 사람, 자동차, 자전거, 동물 등을 촬영한 영상이라면 그 속성상 이동객체가 하나의 영상 블록에만 나타나기는 곤란하고 여러 영상 블록에 걸쳐서 나타날 것으로 예상된다. 즉, 이동객체가 찍힌 영상 블록 근방에 존재하는 영상 블록에는 이동객체가 찍혀있을 가능성이 그렇지 않은 영상 블록에 비해 상대적으로 높다고 예상된다. 그러한 기술

적 사상을 반영하여 (S200)에서는 이동객체 영역 주변에 존재하는 이웃 블록에 대해 상대적으로 완화된 판단기준에 따라 이동객체 영역에 해당하는지 여부를 판단한다.

[0051] 바람직하게는 각각의 이웃 블록을 검사하여, 현재 프레임에서 검출된 모션벡터 값이 미리 설정된 제 2 임계치 (예: 0) 이상이거나 코딩유형이 인트라 픽처(Intra Picture)일 경우에는 해당 영상 블록도 이동객체 영역으로 마킹한다. 다른 실시예로는, 이웃 블록에 대해 앞서 (S100)에서 산출하였던 모션벡터 누적값이 제 2 임계치(예: 5) 이상이거나 코딩유형이 인트라 픽처일 경우에는 해당 영상 블록도 이동객체 영역으로 마킹할 수 있다. 이때, 제 2 임계치는 제 1 임계치에 비해 작은 값으로 설정되는 것이 논리적으로 타당하다.

[0052] 개념적으로는, 유효 움직임이 발견되어 이동객체 영역의 근방에서 어느 정도의 움직임이 있는 영상 블록이라면 이는 앞의 이동객체 영역과 한 덩어리일 가능성이 높기 때문에 이동객체 영역이라고 마킹하는 것이다. 또한, 인트라 픽처의 경우에는 모션벡터가 존재하지 않기 때문에 모션벡터에 기초하여 이동객체 영역인지 여부를 판단하는 것이 불가능하다. 이에, 이동객체 영역으로 이미 검출된 영상 블록에 인접하여 위치하는 인트라 픽처라면 기 추출된 이동객체 영역과 함께 한 덩어리를 이루는 것으로 추정한다. 이동객체 영역이 아닌 영상 블록 하나가 이동객체 영역에 포함되었을 때의 손실은 별로 크지 않은 반면, 이동객체 영역이 과편화되었을 때의 손실은 크기 때문이다.

[0053] 단계 (S300) : 앞의 (S100)과 (S200)에서 검출된 이동객체 영역에 인터폴레이션(interpolation)을 적용하여 이동객체 영역의 분할(fragmentation)을 정리한다. 앞의 과정에서는 영상 블록 단위로 이동객체 영역 여부를 판단하였기 때문에 실제로는 하나의 이동객체(예: 사람)임에도 불구하고 중간중간에 이동객체 영역으로 마킹되지 않은 영상 블록이 존재하여 여러 개의 이동객체 영역으로 분할되는 현상이 발생할 수 있다. 그에 따라, 이동객체 영역으로 마킹된 복수의 영상 블록으로 둘러싸여 하나 혹은 소수의 비마킹 영상 블록이 존재한다면 이들은 이동객체 영역으로 추가로 마킹한다. 이를 통해, 여러 개로 분할되어 있는 이동객체 영역을 하나로 뭉쳐지도록 만들 수 있는데, 이와 같은 인터폴레이션의 영향은 [도 10]과 [도 11]을 비교하면 명확하게 드러난다.

[0054] 앞서의 과정 (S100) 내지 (S300)을 통하여 압축영상으로부터 하나이상의 이동객체 영역(region of moving object)을 획득하였다. 이렇게 획득된 이동객체 영역은 [도 11]에 파란 색으로 표시된 것으로서 일련의 과정에서 이동객체 영역에 속한다고 마킹해둔 다수의 영상블록들이 서로 연결되어 뭉쳐진 덩어리이다. 각각의 단계 (S100) 내지 (S300)에서는 영상블록 단위로 이동객체 영역에 속하는지 여부를 판단하여 마킹하였으나, 최종적으로는 이들이 뭉쳐져서 이룬 영상블록의 덩어리가 이동객체 영역으로 다루어진다. 이러한 이동객체 영역은 압축 영상의 신택스 정보에 기초하여 그 안에 하나이상의 이동객체가 포함되어 있을 것으로 추정된 부분이다. 소프트웨어 처리를 위하여 각각의 이동객체 영역에는 고유 식별정보(Unique ID)를 할당하여 관리하는 것이 바람직하다. [도 12]은 이동객체 영역에 Unique ID가 할당된 일 예를 나타낸다.

[0055] 단계 (S400) : 다음으로, 이상의 과정을 통해 압축영상으로부터 추출된 이동객체 영역에 대한 이동 속성 정보를 획득한다. 이동 속성 정보는 영상 내에서 객체의 존재 및 이동 형태와 연관성을 갖는 정보를 나타내는 것으로서, 개별 프레임 이미지로부터 획득 가능한 정보(예: 모션벡터의 2차원 패턴)로 구성될 수도 있고, 일련의 프레임 시퀀스로부터 획득 가능한 정보(예: 이동궤적)로 구성될 수도 있으며, 이들의 조합으로 구성될 수도 있다. 본 명세서에서는 이동궤적을 기준으로 본 발명을 기술한다. 이동궤적은 압축영상을 구성하는 일련의 프레임 시퀀스로부터 획득되는 복수의 이동객체 영역을 분석하여 얻을 수 있다.

[0056] [도 5]는 본 발명에서 이동객체 영역에 대한 이동 속성 정보가 이동궤적 정보, 박스크기 정보, 순간속도 정보로 구성된 예를 나타내는 도면이다. [도 5]를 참조하면, 영상 내에 네 개의 이동객체 영역이 추출되었으며, 이들 네 개의 이동객체 영역에 대해서 각각 이동 속성 정보가 도출되었다. 이러한 이동 속성 정보는 신경망의 학습 과정에 활용될 수도 있고, 신경망에 입력 벡터로 제공되어 객체 분류 과정에 활용될 수도 있다.

[0057] 이때, 이동궤적 정보나 순간속도 정보는 프레임 시퀀스에서 추출된 여러 이동객체 영역을 분석해야 얻을 수 있는 정보이다. 이를 위해서는 [도 12]에 도시된 바와 같이 이동객체 영역을 마치 하나의 객체인 것처럼 간주하고 압축영상에서 일련의 프레임 이미지를 넘어가면서 그 객체의 이동 궤적을 추적하고 순간적인 이동속도를 측정해야 한다.

[0058] 이 과정에서 이동객체 영역이라고 마킹되어진 서로 연결된 영상 블록의 덩어리가 일련의 영상 프레임 앞뒤 간에 동일한지 아닌지 판단할 수 있어야 한다. 본 발명에서는 원본 영상 이미지를 해석하지 않고 영상 블록이 이동객체 영역인지 여부만 체크하였기 때문에 앞뒤의 영상 프레임에서 이동객체 영역의 덩어리가 실제로 동일한지 아닌지 확인할 수 없다. 즉, 영상에 포함된 이미지 내용을 파악하지 않기 때문에 예컨대 동일 지점에서 앞뒤 프레

임 간에 고양이가 개로 치환되었을 때에 그러한 변화를 식별하지 못한다. 하지만, 프레임 간의 시간간격이 매우 짧다는 점과 CCTV 카메라의 관찰 대상은 통상의 속도로 움직인다는 점을 감안하면 이러한 일이 벌어질 가능성은 매우 낮다.

- [0059] 이에, 본 발명에서는 앞뒤 프레임에서 이동객체 영역의 덩어리 간에 중첩되는 영상 블록의 비율 혹은 갯수가 일정 임계치 이상인 것들을 동일한 이동객체 영역이라고 추정한다. 이러한 접근방식에 의하면 원본 영상의 내용을 모르더라도 특정의 이동객체 영역이 움직이고 있는 것인지 아니면 새로운 이동객체 영역이 신규로 나타난 것인지 아니면 기존의 이동객체 영역이 사라진 것인지 판단할 수 있다. 이러한 판단은 정확도는 종래기술에 비해 낮지만 데이터 처리 속도를 획기적으로 높일 수 있어 실제 적용에서 유리하다.
- [0060] 단계 (S500, S600) : 다음으로, 이동객체 영역의 이미지(예: 썸네일 이미지)를 압축영상을 디코딩하여 획득한다. 본 발명에서 이동객체 영역의 이미지를 얻는 이유는 외부의 상용분류기와 연동하기 위함이다. 상용분류기는 객체의 이미지를 전달하면 그 객체의 종류를 제공해주는 방식이므로 이러한 상용분류기와 연동하려면 이동객체 영역의 이미지가 필요한 것이다.
- [0061] 이때, 압축영상을 전체적으로 디코딩하는 것 보다는 이동객체 영역이 마킹된 영상 프레임을 선별적으로 디코딩하는 것이 속도 측면에서 바람직하다. 또한, 영상 프레임 내에서도 이동객체 영역에 해당하는 영상 데이터만 디코딩한다면 속도 측면에서 더욱 바람직하다. 즉, 압축영상을 디코딩하여 프레임 이미지를 구성한 후에 이동객체 영역 부분을 오려낼 수도 있고, 이동객체 영역에 해당하는 압축영상 데이터만 선별적으로 디코딩할 수도 있다.
- [0062] 그리고 나서, 위에서 얻은 이동객체 영역의 이미지를 외부 상용분류기와 연동하여 객체 분류 정보를 획득한다. 일 실시예로서, 이동객체 영역에 대한 썸네일 이미지를 Azure 클라우드 서비스에게 질의(query)함으로써 해당 이동객체 영역의 범주(예: 사람, 자동차 등)에 대한 해답을 제공받는다.
- [0063] 단계 (S700) : 신경망에 이동객체 영역의 이동 속성 정보를 입력하여 객체 분류 결과를 얻기 위해서는 그 신경망이 그에 대응하는 성질을 갖도록 학습(머신러닝)이 미리 이루어져 있어야만 한다. 따라서, 본 발명에서 (S700)과 같은 신경망 학습 모드의 프로세스가 수행되어야 한다.
- [0064] 이상의 과정을 통하여, 압축영상으로부터 획득되는 다수의 이동객체 영역의 각각에 대하여 이동 속성 정보를 획득하였을 뿐만 아니라, 썸네일 이미지를 상용분류기에 전달함으로써 객체 분류 정보도 획득하였다. 즉, 다수의 이동객체 영역의 각각에 대하여 이동 속성 정보와 객체 분류 정보의 조합이 획득된 것인데, (S700)에서는 이들 조합을 신경망을 위한 훈련 데이터셋으로 설정하여 신경망에 대한 머신러닝을 수행한다. 신경망 이론에 따르면, 이러한 머신러닝을 통하여 해당 신경망은 이동 속성 정보를 카테고리화하는 성질을 갖추게 된다.
- [0065] CCTV 영상관제 시스템의 예를 들면, CCTV 카메라는 다양한 각도, 높이, 줌 값을 가지고 있으므로 각각의 카메라에서 촬영된 영상은 조금씩 상이한 특성을 나타내게 된다. 실제 현장에서 사용하기 위해서는 이러한 카메라를 모두 수용해야 하는데, 개별 카메라마다 일일이 사람이 확인하여 신경망을 학습시킨다면 엄청난 시간과 비용이 발생하게 되어 비현실적이다. 반면, 본 발명에서는 다양한 촬영환경에 대응할 수 있는 대규모의 훈련 데이터셋을 사람의 개입 없이 소프트웨어 처리를 통해서 확보하고 신경망 학습에 적용할 수 있다.
- [0066] 단계 (S800) : 이처럼 신경망에 대한 학습을 수행한 이후에 해당 신경망을 이용하여 객체 분류를 수행하게 된다. 객체 분류 대상으로 설정된 이동객체 영역의 이동 속성 정보를 신경망에 입력하면, 그 신경망의 연산 결과로부터 이동객체 영역에 대한 객체 분류 결과를 획득할 수 있다. 이 과정은 연산의 복잡도가 낮기 때문에 객체 분류 결과를 신속하게 얻을 수 있다. 실제 시스템 운용 상황에서는 압축영상으로부터 실시간으로 획득되는 이동객체 영역이 객체 분류 대상이며, 그에 따라 지속적으로 객체 분류 결과가 도출된다.
- [0067] [도 6]은 본 발명에서 신경망이 출력하는 객체 분류 결과의 일 예를 나타내는 도면이다. 객체 분류 결과의 신뢰도를 높이기 위해서는 신경망의 연산 결과 값을 여러가지가 가능하도록 구성하고(예: 8가지), 이 중에서 소수의 값에 대해 유의미한 객체 분류(예: 사람, 자동차)를 할당하는 한편, 나머지 값에 대해서는 무의미한 객체 분류(예: 클러터)를 할당하는 것이 바람직하다.
- [0068] 한편, 이동객체 영역, 이동객체 영역의 이미지, 객체 분류 결과의 조합은 영상관제 장치로 제공되는데, 영상관제 장치는 객체 분류 결과를 다양한 형태로 활용할 수 있다. 예를 들어, CCTV 모니터 화면에서 객체 분류별로 상이한 색상을 지정하여 디스플레이할 수 있다. 또한, 범죄해결 단서를 확보할 목적으로 영상 검색을 수행할 때에 스토리지에 저장되어 있는 대규모의 CCTV 촬영 영상으로부터 검색 대상을 특정 분류로 한정함으로써 검색 대상 모집단의 범위를 축소시킬 수 있으며, 이를 통해 자료 검색 속도를 높일 수 있다.

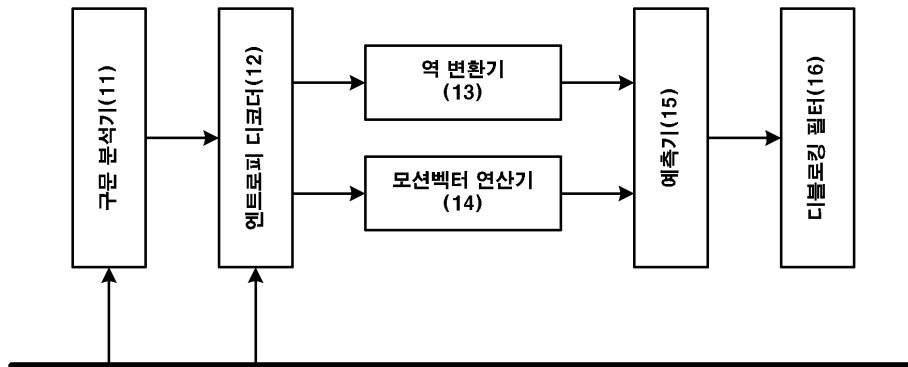
- [0070] [도 7]은 본 발명에서 이동객체 영역을 추출하는 첫번째 단계로서 압축영상에서 유효 움직임(effective movement) 영역을 검출하는 과정의 구현 예를 나타내는 순서도이고, [도 8]은 CCTV 영상에 대해 유효 움직임 영역 검출 과정이 적용된 결과의 일 예를 나타내는 도면이다. [도 7]의 프로세스는 [도 3]에서 단계 (S100)에 대응한다.
- [0071] 단계 (S110) : 먼저, 압축영상의 코딩 유닛을 파싱하여 모션벡터 및 코딩유형을 획득한다. [도 1]을 참조하면, 동영상 디코딩 장치는 압축영상의 스트림에 대해 H.264 AVC 및 H.265 HEVC 등과 같은 동영상압축 표준에 따라 구문분석(헤더 파싱) 및 모션벡터 연산을 수행한다. 이러한 과정을 통하여 압축영상의 코딩 유닛에 대하여 모션벡터와 코딩유형을 파싱해낸다.
- [0072] 단계 (S120) : 압축영상을 구성하는 복수의 영상 블록 별로 미리 설정된 시간(예: 500 ms) 동안의 모션벡터 누적값을 획득한다.
- [0073] 이 단계는 압축영상으로부터 실질적으로 의미를 인정할만한 유효 움직임, 예컨대 주행중인 자동차, 달려가는 사람, 서로 싸우는 군중들이 있다면 이를 검출하려는 의도를 가지고 제시되었다. 흔들리는 나뭇잎, 잠시 나타나는 고스트, 빛의 반사에 의해 약간씩 변하는 그림자 등은 비록 움직임은 있지만 실질적으로는 무의미한 객체이므로 검출되지 않도록 한다.
- [0074] 이를 위해, 미리 설정된 일정 시간(예: 500 msec) 동안 하나이상의 영상 블록 단위로 모션벡터를 누적시켜 모션벡터 누적값을 획득한다. 이때, 영상 블록은 매크로블록과 서브블록을 포함하는 개념으로 사용된 것이다.
- [0075] 단계 (S130, S140) : 복수의 영상 블록에 대하여 모션벡터 누적값을 미리 설정된 제 1 임계치(예: 20)와 비교하며, 제 1 임계치를 초과하는 모션벡터 누적값을 갖는 영상 블록을 이동객체 영역으로 마킹한다.
- [0076] 만일 이처럼 일정 이상의 모션벡터 누적값을 갖는 영상 블록이 발견되면 해당 영상 블록에서 무언가 유의미한 움직임, 즉 유효 움직임이 발견된 것으로 보고 이동객체 영역으로 마킹한다. 예컨대 영상관제 시스템에서 사람이 뛰어가는 정도로 관제 요원이 관심을 가질만한 가치가 있을 정도의 움직임을 선별하여 검출하려는 것이다. 반대로, 모션벡터가 발생하였다더라도 일정 시간동안의 누적값이 제 1 임계치를 넘지 못할 정도로 작을 경우에는 영상에서의 변화가 그다지 크지않고 미미한 것으로 추정하고 검출 단계에서 무시한다.
- [0077] [도 8]은 본 발명에서 [도 7]의 과정을 통해 CCTV 영상으로부터 유효 움직임 영역을 검출한 결과를 시각적으로 나타낸 일 예이다. [도 8]에서는 제 1 임계치 이상의 모션벡터 누적값을 갖는 영상 블록이 이동객체 영역으로 마킹되어 붉은 색으로 표시되었다. [도 8]을 살펴보면 보도블럭이나 도로, 그리고 그림자가 있는 부분 등은 이동객체 영역으로 표시되지 않은 반면, 걷고있는 사람들이나 주행중인 자동차 등이 이동객체 영역으로 표시되었다.
- [0079] [도 9]는 본 발명에서 이동객체 영역을 추출하는 두번째 단계로서 이동객체 영역에 대한 바운더리 영역(boundary area)을 검출하는 과정의 구현 예를 나타내는 순서도이고, [도 10]은 [도 8]의 CCTV 영상 이미지에 대해 바운더리 영역 검출 과정이 추가로 적용된 결과의 일 예를 나타내는 도면이다. [도 9]의 프로세스는 [도 3]에서 단계 (S200)에 대응한다.
- [0080] 앞서의 [도 8]을 살펴보면 이동객체에 해당되는 영상블록이 제대로 마킹되지 않았으며 일부에 대해서만 마킹이 이루어진 것을 발견할 수 있다. 걷고있는 사람이나 주행중인 자동차를 살펴보면 객체의 전부가 마킹되지 않고 그 일부의 영상블록만 마킹되었다. 또한, 하나의 이동객체에 대해 복수의 이동객체 영역이 형성된 것도 많이 발견된다. 예를 들어, 자동차를 살펴보면 복수 개의 이동객체 영역이 형성되어 있다. 이는 (S100)에서 채택한 이동객체 영역의 판단 기준이 일반 영역을 필터링 아웃하는 데에는 매우 유용하지만 상당히 엄격한 것이었음을 의미한다. 따라서, (S100)에서 마킹된 이동객체 영역을 중심으로 그 주변의 영상블록들을 검토하고 일정 기준을 만족한다면 이동객체 영역을 추가로 마킹해줌으로써 결과적으로는 이동객체 영역의 바운더리를 검출해나가는 과정이 필요하다.
- [0081] 단계 (S210) : 먼저, 앞의 (S100)에 의해 이동객체 영역으로 마킹된 영상 블록을 중심으로 하여 인접하는 복수의 영상 블록을 식별한다. 이들은 본 명세서에서는 '이웃 블록(neighboring blocks)'이라고 부른다. 이들 이웃블록은 (S100)에 의해서는 이동객체 영역으로 마킹되지 않은 부분인데, [도 9]의 프로세스에서는 이들에 대해 좀더 살펴봄으로써 이들 이웃 블록 중에서 이동객체 영역의 바운더리에 포함될만한 것이 있는지 확인하려는 것

이다.

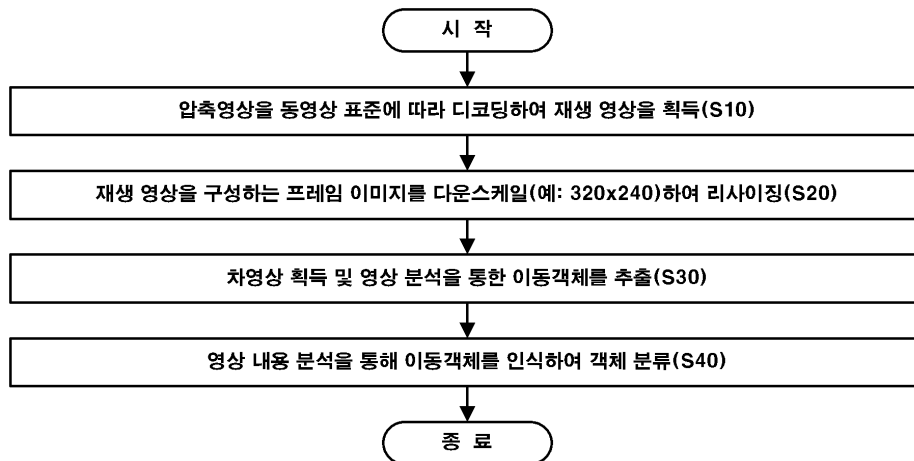
- [0082] 단계 (S220, S230) : 복수의 이웃 블록에 대하여 모션벡터 값을 미리 설정된 제 2 임계치와 비교하고, 제 2 임계치를 초과하는 모션벡터 값을 갖는 이웃 블록을 이동객체 영역으로 마킹한다. 실질적으로 의미를 부여할만한 유효 움직임이 인정된 이동객체 영역에 인접하여 위치하고 그 자신에 대해서도 어느 정도의 움직임이 발견되고 있다면 그 영상 블록은 촬영 영상(예: CCTV 영상)의 특성상 앞의 이동객체 영역과 한 덩어리일 가능성이 높다. 따라서, 이러한 이웃 블록도 이동객체 영역이라고 마킹한다.
- [0083] 이를 구현하는 제 1 실시예로서, 각각의 이웃 블록을 검사하여, 현재 프레임에서 검출된 모션벡터 값이 미리 설정된 제 2 임계치(예: 0) 이상인 경우에 해당 영상 블록도 이동객체 영역으로 마킹한다.
- [0084] 한편, 제 2 실시예로서, 이웃 블록에 대해 앞서 (S100)에서 산출하였던 모션벡터 누적값이 미리 설정된 제 2 임계치(예: 5) 이상인 경우에는 해당 영상 블록도 이동객체 영역으로 마킹할 수 있다. 이때, 제 2 임계치는 제 1 임계치에 비해 작은 값으로 설정되는 것이 타당하다.
- [0085] 단계 (S240) : 또한, 복수의 이웃 블록 중에서 코딩유형이 인트라 픽처인 것을 이동객체 영역으로 마킹한다. 인트라 픽처의 경우에는 모션벡터가 존재하지 않기 때문에 해당 이웃 블록이 이동객체 영역에 해당되는지 여부를 모션벡터에 기초하여 판단하는 것이 불가능하다. 이동객체 영역으로 이미 검출된 영상 블록에 인접하여 위치하는 인트라 픽처라면 기 추출된 이동객체 영역과 함께 한 덩어리를 이루는 것으로 추정하는 것이 바람직하다. 이동객체 영역이 아닌 영상 블록 하나가 이동객체 영역에 포함되었을 때의 손실은 별로 크지 않은 반면, 이동객체 영역이 과편화되었을 때의 손실은 크기 때문이다.
- [0086] [도 10]은 본 발명에서 CCTV 영상에 바운더리 영역 검출 과정까지 적용된 결과를 시각적으로 나타낸 도면인데, 이상의 과정을 통해 이동객체 영역으로 마킹된 다수의 영상 블록을 파란 색으로 표시하였다. [도 10]을 살펴보면, 앞서 [도 8]에서 붉은 색으로 표시되었던 이동객체 영역의 근방으로 파란 색의 이동객체 영역은 좀더 확장되었으며 이를 통해 CCTV로 촬영된 실제 영상과 비교할 때 이동객체를 전부 커버할 정도가 되었다는 사실을 발견할 수 있다.
- [0088] [도 11]은 [도 10]의 CCTV 영상 이미지에 대해 인터플레이션을 통해 이동객체 영역을 정리한 결과의 일 예를 나타낸 도면이다.
- [0089] 단계 (S300)은 앞의 (S100)과 (S200)에서 검출된 이동객체 영역에 인터플레이션을 적용하여 이동객체 영역의 분할을 정리하는 과정이다. [도 10]을 살펴보면 파란 색으로 표시된 이동객체 영역 사이사이에 비마킹 영상 블록이 발견된다. 이렇게 중간중간에 비마킹 영상 블록이 존재하게 되면 이들이 다수의 개별적인 이동객체인 것처럼 간주될 수 있다. 이렇게 이동객체 영역이 과편화되면 단계 (S400)의 결과가 부정확해질 수 있고, 이동객체 영역의 갯수가 많아져서 단계 (S400)의 프로세스가 복잡해지는 문제도 있다.
- [0090] 그에 따라, 본 발명에서는 이동객체 영역으로 마킹된 복수의 영상 블록으로 둘러싸여 하나 혹은 소수의 비마킹 영상 블록이 존재한다면 이는 이동객체 영역으로 마킹하는데, 이를 인터플레이션이라고 부른다. [도 10]과 대비하여 [도 11]을 살펴보면, 이동객체 영역 사이사이에 존재하던 비마킹 영상 블록이 모두 이동객체 영역이라고 마킹되었다. 이를 통해, 덩어리로 움직이는 영역은 모두 묶어서 하나의 이동객체로서 다루게 된다.
- [0091] [도 8], [도 10], [도 11]을 비교하면 바운더리 영역 검출 과정과 인터플레이션 과정을 거치면서 이동객체 영역이 실제 영상의 상황을 제대로 반영하게 되어간다는 사실을 발견할 수 있다. [도 8]에서 붉은 색으로 마킹된 덩어리로 판단한다면 영상 화면 속에 아주 작은 물체들이 다수 움직이는 것처럼 다루어질 것인데, 이는 실제와는 부합하지 않는다. 반면, [도 11]에서 파란 색으로 마킹된 덩어리로 판단한다면 어느 정도의 부피를 갖는 몇 개의 이동객체가 존재하는 것으로 다루어질 것이어서 실제 장면을 유사하게 반영한다.
- [0093] 한편, 본 발명은 컴퓨터가 읽을 수 있는 비휘발성 기록매체에 컴퓨터가 읽을 수 있는 코드의 형태로 구현되는 것이 가능하다. 이러한 비휘발성 기록매체로는 다양한 형태의 스토리지 장치가 존재하는데 예컨대 하드디스크, SSD, CD-ROM, NAS, 자기테이프, 웹디스크, 클라우드 디스크 등이 있고 네트워크로 연결된 다수의 스토리지 장치에 코드가 분산 저장되고 실행되는 형태로 구현될 수 있다. 또한, 본 발명은 하드웨어와 결합되어 특정의 절차를 실행시키기 위하여 매체에 저장된 컴퓨터프로그램의 형태로 구현될 수도 있다.

도면

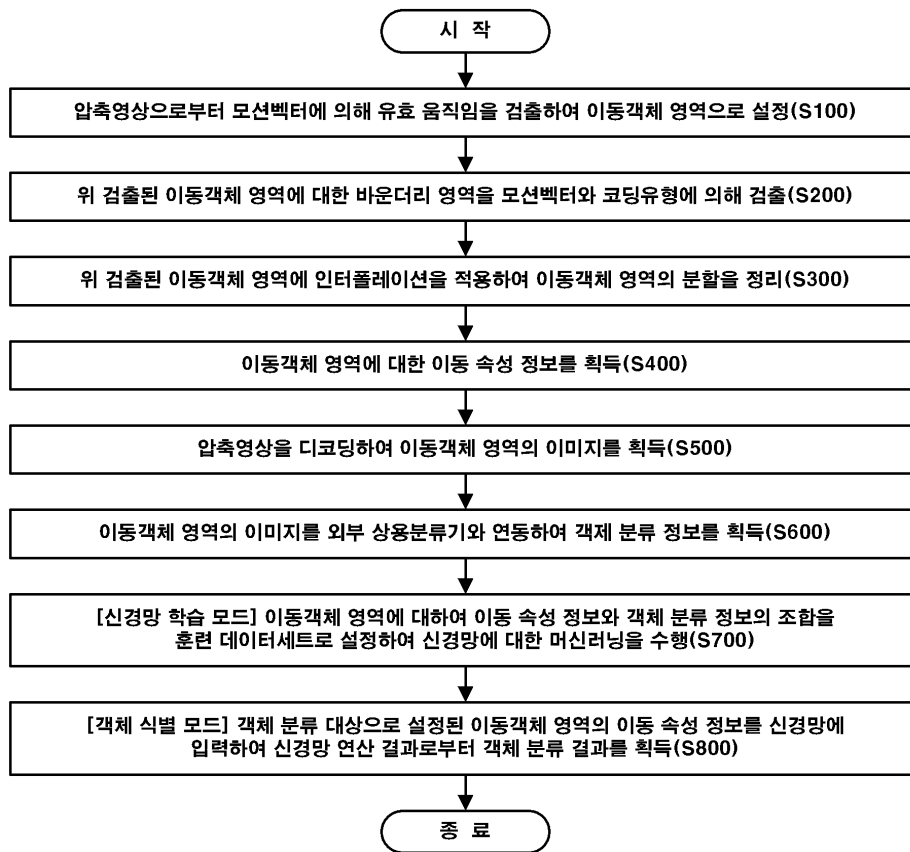
도면1



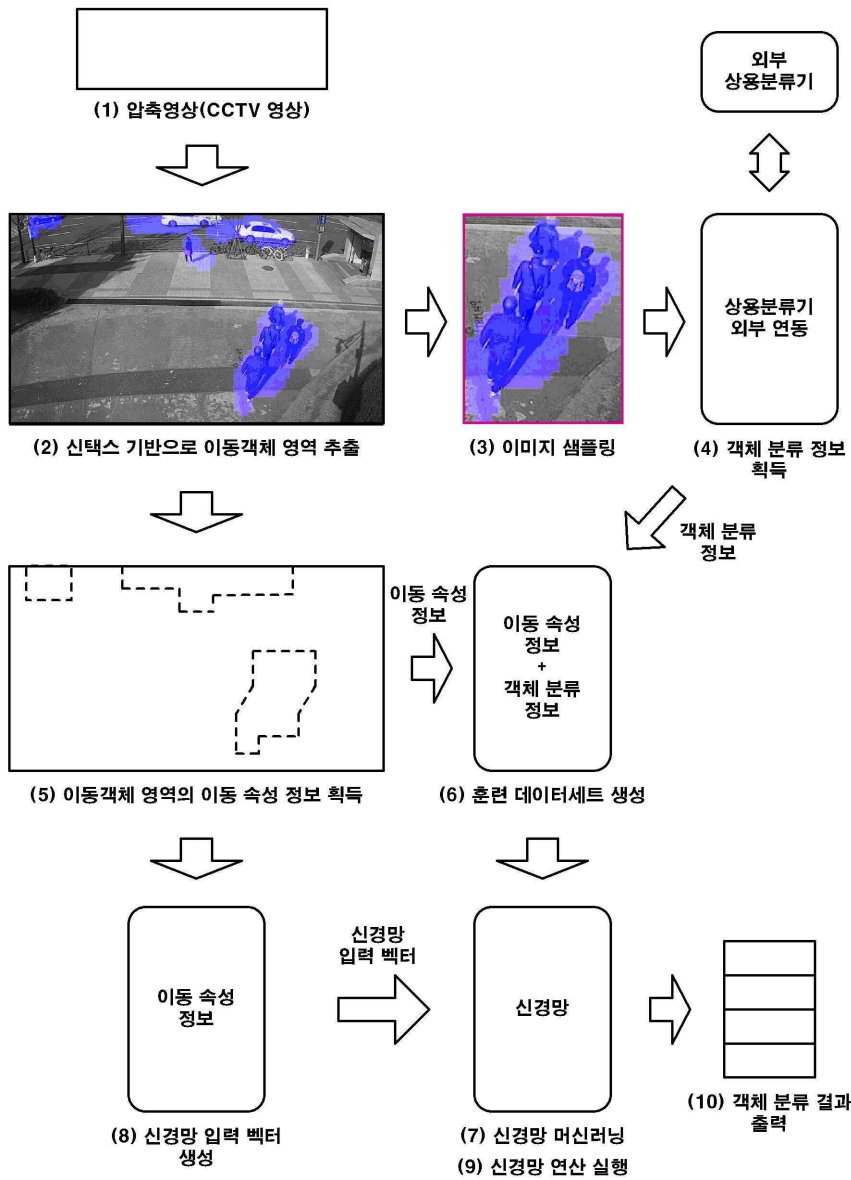
도면2



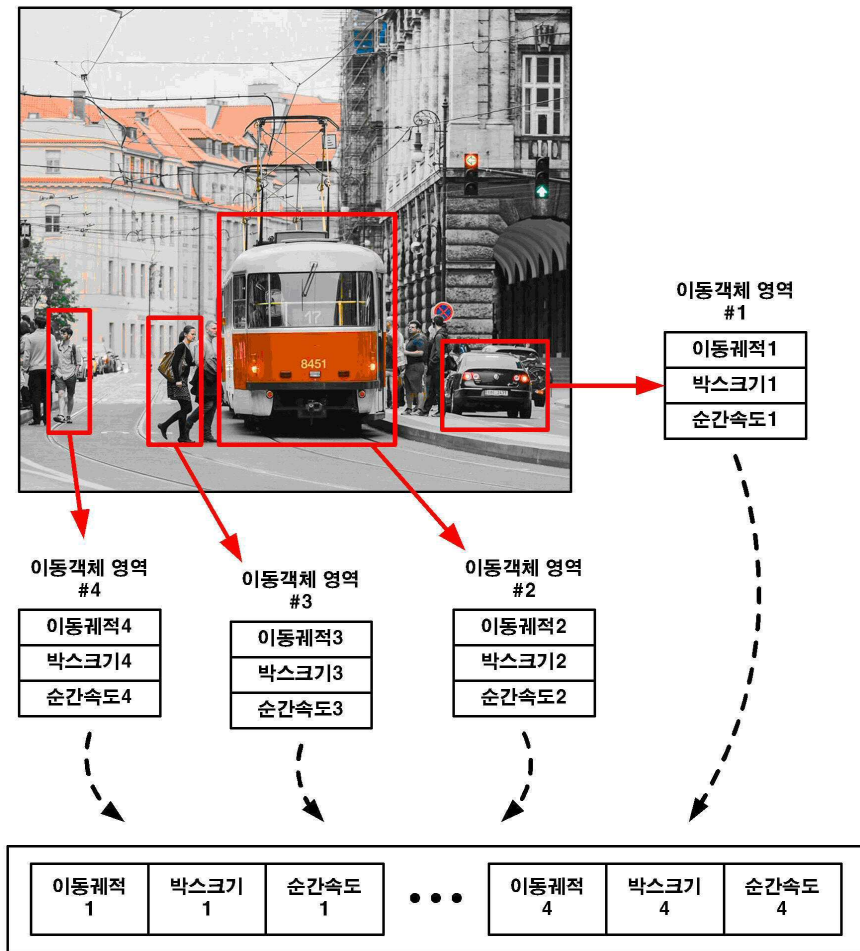
도면3



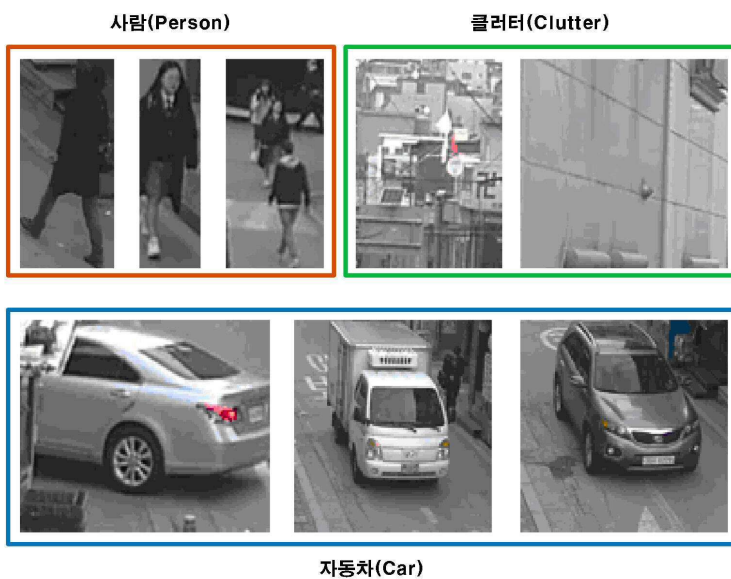
도면4



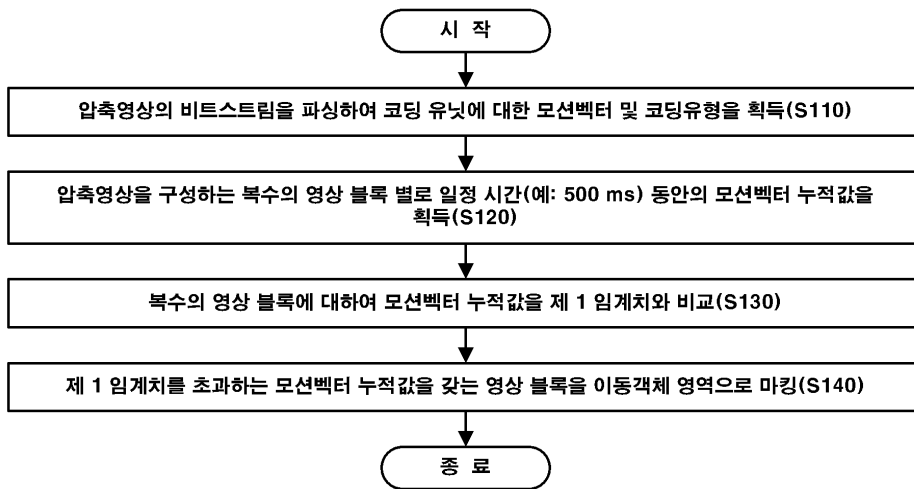
도면5



도면6



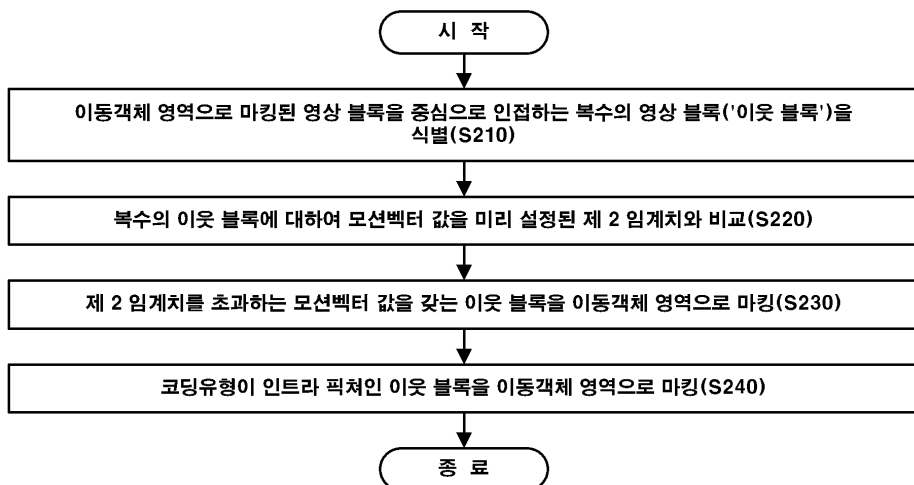
도면7



도면8



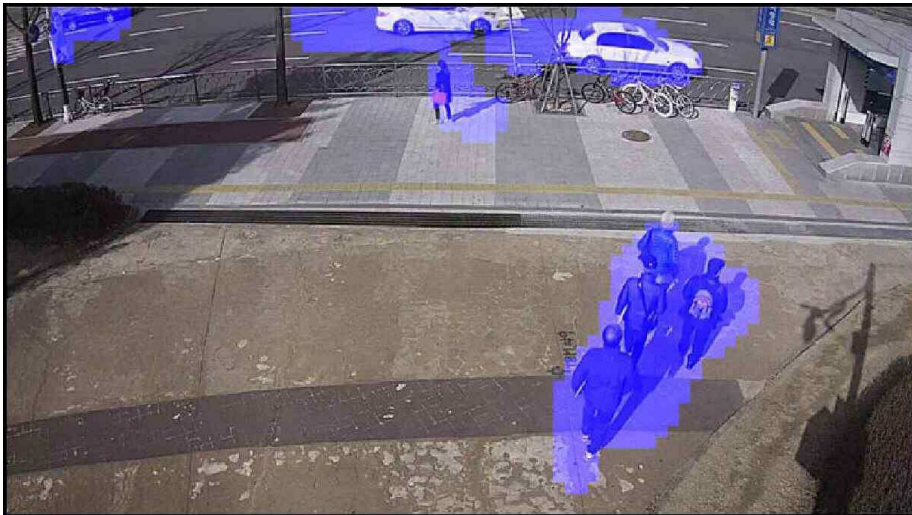
도면9



도면10



도면11



도면12

