



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0071886
(43) 공개일자 2020년06월22일

(51) 국제특허분류(Int. Cl.)
G06K 9/00 (2006.01) G06K 9/48 (2006.01)
G06K 9/62 (2006.01) G06N 3/08 (2006.01)
H04N 7/18 (2006.01)

(52) CPC특허분류
G06K 9/00711 (2013.01)
G06K 9/481 (2013.01)

(21) 출원번호 10-2018-0156434
(22) 출원일자 2018년12월06일
심사청구일자 2018년12월06일

(71) 출원인
이노덱 주식회사
서울특별시 구로구 디지털로31길 61, 5층(구로동, 드림마크원데이터센터)

(72) 발명자
최승현
서울특별시 양천구 중앙로52길 16-7, 403호
이성진
경기도 광명시 오리로 801, 107동 1004호

(74) 대리인
김도형

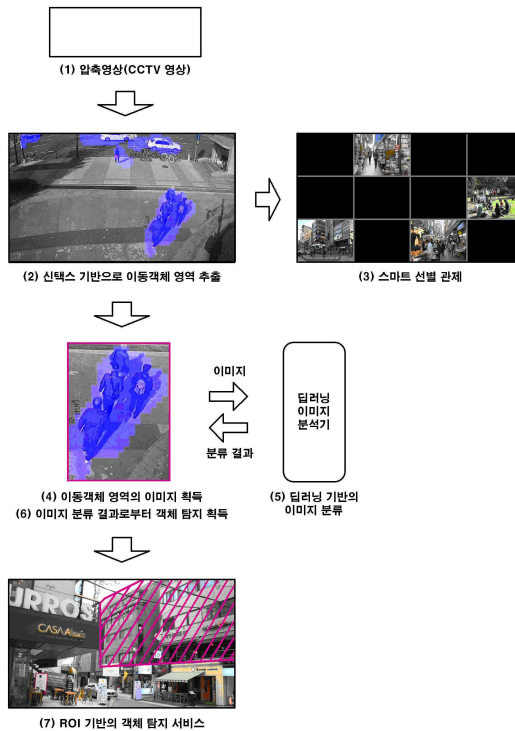
전체 청구항 수 : 총 9 항

(54) 발명의 명칭 **딥러닝 이미지 분석과 연동하는 신택스 기반의 선별 관제 제공 방법**

(57) 요약

본 발명은 일반적으로 다수의 CCTV 카메라에 대하여 스마트 영상관제를 효과적으로 수행하는 기술에 관한 것이다. 더욱 상세하게는, 본 발명은 CCTV 카메라가 생성하는 압축영상에 대해 종래기술처럼 복잡한 이미지 프로세싱을 통해 객체를 감지하는 것이 아니라 압축영상 데이터를 파싱하여 얻은 신택스 정보(예: 모션벡터, 코딩유닛)를 이용하여 객체를 감지하는 것이다.

대표도 - 도4



형) 기반으로 영상 내에서 무언가 유의미한 움직임이 존재하는 영역(즉, 이동객체 영역)을 추출하고 그에 대응하여 선별 관제를 제공하는 기술에 관한 것이다. 특히, 압축영상에서 선택스 기반으로 이동객체 영역을 추출하는 프로세스를 정교화하여 눈/비 내림, 야간 노이즈, 바람에 흔들리는 나뭇가지나 현수막, PTZ 동작에 따른 화면 이동, CCTV 지지대 진동에 따른 화면 떨림 등을 필터링함으로써 선별 관제의 효과를 높일 수 있는 기술에 관한 것이다. 특히, ROI 지정 및 객체 감지에 따른 영상관제를 수행함에 있어서 종래기술처럼 압축영상을 디코딩하여 나오는 일련의 영상 프레임에 대해 딥러닝 객체 감지를 수행하는 것이 아니라 선택스 기반으로 추출된 이동객체 영역의 이미지를 선별하여 딥러닝 이미지 분류를 수행하도록 구성함으로써 영상관제의 효율을 높일 수 있는 기술에 관한 것이다.

(52) CPC특허분류

G06K 9/6267 (2013.01)

G06N 3/08 (2013.01)

H04N 7/18 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 2017-0-00250

부처명 과학기술정보통신부

연구관리전문기관 정보통신기술진흥센터

연구사업명 정보통신방송연구개발사업

연구과제명 시 시스템 핵심 기술 개발 옛지카메라 임베디드 시스템과 영상분석 시스템의 협업 강화학습기반 지능형 국방경계 감시 시스템

기 여 율 1/1

주관기관 중앙대학교 산학협력단

연구기간 2017.04.01 ~ 2021.12.31

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 영역으로 마킹된 복수의 영상 블록들이 상호 연결된 덩어리를 상기 압축영상의 이동객체 영역으로 설정하는 제 5 단계;
상기 이동객체 영역이 설정된 CCTV 영상공급 채널을 취합하여 선별 관제를 제공하는 제 6 단계;
상기 압축영상을 디코딩하여 상기 이동객체 영역의 이미지를 획득하는 제 7 단계;
상기 이동객체 영역의 이미지에 대해 딥러닝 기반으로 이미지 분류 정보를 획득하는 제 8 단계;
상기 이동객체 영역에 대하여 프레임 이미지 내의 위치 정보와 상기 이미지 분류 정보를 결합하여 객체 감지 정보를 획득하는 제 9 단계;
를 포함하여 구성되는 딥러닝 이미지 분석과 연동하는 신택스 기반의 선별 관제 제공 방법.

청구항 2

청구항 1에 있어서,
상기 제 9 단계 이후에 수행되는,
상기 객체 감지 정보를 미리 설정된 ROI(관심영역)와 대비함으로써 ROI에 대한 객체 감지 서비스를 제공하는 단계;
를 더 포함하여 구성되는 것을 특징으로 하는 딥러닝 이미지 분석과 연동하는 신택스 기반의 선별 관제 제공 방법.

청구항 3

청구항 1에 있어서,
상기 제 7 단계는 상기 압축영상에서 이동객체 영역이 마킹된 영상 프레임을 선별적으로 디코딩하여 상기 이동객체 영역의 이미지를 획득하는 것을 특징으로 하는 딥러닝 이미지 분석과 연동하는 신택스 기반의 선별 관제 제공 방법.

청구항 4

청구항 1에 있어서,
상기 제 1 단계와 상기 제 2 단계 사이에 수행되는,
프레임 이미지의 전역에 대해 지배적인 모션벡터가 균일하게 검출되는 전역 균일 모션벡터 상황을 식별 및 필터링하는 단계;
를 더 포함하여 구성되는 것을 특징으로 하는 딥러닝 이미지 분석과 연동하는 신택스 기반의 선별 관제 제공 방법.

법.

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 제 1 단계에서 획득된 모션벡터 값을 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 값을 갖는 이웃 블록을 이동객체 영역으로 추가 마킹하는 제 c 단계;

를 더 포함하여 구성되는 것을 특징으로 하는 딥러닝 이미지 분석과 연동하는 신택스 기반의 선별 관계 제공 방법.

청구항 6

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 딥러닝 이미지 분석과 연동하는 신택스 기반의 선별 관계 제공 방법.

청구항 7

청구항 5 또는 청구항 6에 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 딥러닝 이미지 분석과 연동하는 신택스 기반의 선별 관계 제공 방법.

청구항 8

청구항 7에 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 딥러닝 이미지 분석과 연동하는 신택스 기반의 선별 관제 제공 방법.

청구항 9

하드웨어와 결합되어 청구항 1 내지 6 중 어느 하나의 항에 따른 딥러닝 이미지 분석과 연동하는 신택스 기반의 선별 관제 제공 방법을 실행시키기 위하여 매체에 저장된 컴퓨터프로그램.

발명의 설명

기술 분야

- [0001] 본 발명은 일반적으로 다수의 CCTV 카메라에 대하여 스마트 영상관제를 효과적으로 수행하는 기술에 관한 것이다.
- [0002] 더욱 상세하게는, 본 발명은 CCTV 카메라가 생성하는 압축영상에 대해 종래기술처럼 복잡한 이미지 프로세싱을 통해 객체를 감지하는 것이 아니라 압축영상 데이터를 파싱하여 얻은 신택스 정보(예: 모션벡터, 코딩유형) 기반으로 영상 내에서 무언가 유의미한 움직임이 존재하는 영역(즉, 이동객체 영역)을 추출하고 그에 대응하여 선별 관제를 제공하는 기술에 관한 것이다.
- [0003] 특히, 압축영상에서 신택스 기반으로 이동객체 영역을 추출하는 프로세스를 정교화하여 눈/비 내림, 야간 노이즈, 바람에 흔들리는 나뭇가지나 현수막, PTZ 동작에 따른 화면 이동, CCTV 지지대 진동에 따른 화면 떨림 등을 필터링함으로써 선별 관제의 효과를 높일 수 있는 기술에 관한 것이다.
- [0004] 특히, ROI 지정 및 객체 감지에 따른 영상관제를 수행함에 있어서 종래기술처럼 압축영상을 디코딩하여 나오는 일련의 영상 프레임에 대해 딥러닝 객체 감지를 수행하는 것이 아니라 신택스 기반으로 추출된 이동객체 영역의 이미지를 선별하여 딥러닝 이미지 분류를 수행하도록 구성함으로써 영상관제의 효율을 높일 수 있는 기술에 관한 것이다.

배경 기술

- [0006] 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.
- [0007] 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.
- [0008] 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.
- [0009] 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.

[0010] 이하에서는 [도 1]와 [도 2]를 참조하여 종래기술에서 CCTV 영상으로부터 객체를 감지해내는 과정을 기술한다.

[0011] [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.

[0012] [도 2]는 기존의 영상분석 솔루션에서 CCTV 영상으로부터 이동객체를 감지해내는 과정을 나타내는 순서도이다. [도 2]를 참조하면, 종래기술에서는 압축영상을 동영상 표준(예: H.264 AVC, H.265 HEVC 등)에 따라 디코딩하여 재생영상을 획득하고(S10), 그 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).

[0013] 종래기술에서 이동객체를 감지하려면 압축영상 디코딩, 다운스케일 리사이징, 영상 분석을 수행한다. 이들은 복잡도가 매우 높은 프로세스이고, 그로 인해 종래의 영상관제 시스템에서는 영상분석 서버의 채널 수용량, 즉 한 대의 영상분석 서버가 동시 처리할 수 있는 CCTV 채널 수가 상당히 제한되어 있다. 현재 고성능의 영상분석 서버가 커버할 수 있는 최대 CCTV 채널은 통상 최대 16 채널 정도이다. CCTV 관제센터에는 적게는 수백 대에서 많게는 수만 대의 CCTV 카메라로부터 촬영영상이 전달된다. 이렇게 다수의 채널로부터 전달되는 압축영상들을 동시에 처리하려면 영상분석 서버가 엄청나게 많이 필요하였고, 이는 비용 증가와 물리적 공간 확보의 어려움이라는 문제점을 유발하였다.

[0014] 한편, 최근들어 심층신경망(deep neural network)을 여러 분야에 접목시키려는 시도가 있어왔으며, 영상관제 분야에서도 CCTV 촬영영상을 딥러닝(deep learning)과 연동 처리하여 영상분석 서버의 부담을 줄이는 방식이 논의되고 있다. 그러나, 딥러닝을 이용하여 영상에서 객체를 감지한다는 것은 소프트웨어적으로 대단히 어려운 과제이다. 실험 결과에 따르면, 고성능 컴퓨터와 최신 알고리즘을 활용하더라도 연산처리 속도는 초당 6 ~ 25 프레임 정도에 불과하고 객체 감지의 정확도 또한 66% 정도로 매우 낮다. 객체 감지의 정확도가 낮은 것도 문제이지만 처리속도가 너무 떨어진다는 것은 더 큰 문제이므로 딥러닝 알고리즘이 획기적으로 개선되지 않는 한, 실제 환경에 적용하기 어려운 수준이다.

발명의 내용

해결하려는 과제

[0016] 본 발명의 목적은 일반적으로 다수의 CCTV 카메라에 대하여 스마트 영상관제를 효과적으로 수행하는 기술을 제공하는 것이다.

[0017] 특히, 본 발명의 목적은 CCTV 카메라가 생성하는 압축영상에 대해 종래기술처럼 복잡한 이미지 프로세싱을 통해 객체를 감지하는 것이 아니라 압축영상 데이터를 파싱하여 얻은 선택스 정보(예: 모션벡터, 코딩유형) 기반으로 영상 내에서 무언가 유의미한 움직임이 존재하는 영역(즉, 이동객체 영역)을 추출하고 그에 대응하여 선별 관제를 제공하는 기술을 제공하는 것이다.

[0018] 특히, 본 발명의 목적은 압축영상에서 선택스 기반으로 이동객체 영역을 추출하는 프로세스를 정교화하여 눈/비 내림, 야간 노이즈, 바람에 흔들리는 나뭇가지나 현수막, PTZ 동작에 따른 화면 이동, CCTV 지지대 진동에 따른 화면 떨림 등을 필터링함으로써 선별 관제의 효과를 높일 수 있는 기술을 제공하는 것이다.

[0019] 특히, 본 발명의 목적은 ROI 지정 및 객체 감지에 따른 영상관제를 수행함에 있어서 종래기술처럼 압축영상을 디코딩하여 나오는 일련의 영상 프레임에 대해 딥러닝 객체 감지를 수행하는 것이 아니라 선택스 기반으로 추출된 이동객체 영역의 이미지를 선별하여 딥러닝 이미지 분류를 수행하도록 구성함으로써 영상관제의 효율을 높일 수 있는 기술을 제공하는 것이다.

과제의 해결 수단

- [0021] 상기의 목적을 달성하기 위하여 본 발명에 따른 딥러닝 이미지 분석과 연동하는 신텍스 기반의 선별 관계 제공 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; 이동객체 영역으로 마킹된 복수의 영상 블록들이 상호 연결된 덩어리를 압축영상의 이동객체 영역으로 설정하는 제 5 단계; 이동객체 영역이 설정된 CCTV 영상공급 채널을 취합하여 선별 관계를 제공하는 제 6 단계; 압축영상을 디코딩하여 이동객체 영역의 이미지를 획득하는 제 7 단계; 이동객체 영역의 이미지에 대해 딥러닝 기반으로 이미지 분류 정보를 획득하는 제 8 단계; 이동객체 영역에 대하여 프레임 이미지 내의 위치 정보와 이미지 분류 정보를 결합하여 객체 감지 정보를 획득하는 제 9 단계;를 포함하여 구성된다.
- [0022] 이때, 제 7 단계는 압축영상에서 이동객체 영역이 마킹된 영상 프레임을 선별적으로 디코딩하여 이동객체 영역의 이미지를 획득하도록 구성될 수 있다.
- [0023] 이때, 본 발명에 따른 신텍스 기반의 선별 관계 제공 방법은, 객체 감지 정보를 미리 설정된 ROI(관심영역)와 대비함으로써 ROI에 대한 객체 감지 서비스를 제공하는 단계;를 더 포함하여 구성될 수 있다.
- [0024] 또한, 본 발명에 따른 신텍스 기반의 선별 관계 제공 방법은, 프레임 이미지의 전역에 대해 지배적인 모션벡터가 균일하게 검출되는 전역 균일 모션벡터 상황을 식별 및 필터링하는 단계;를 더 포함하여 구성될 수 있다.
- [0025] 또한, 본 발명에 따른 신텍스 기반의 선별 관계 제공 방법은, 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 모션벡터 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.
- [0026] 또한, 본 발명에 따른 신텍스 기반의 선별 관계 제공 방법은, 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.
- [0027] 한편, 본 발명에 따른 컴퓨터프로그램은 하드웨어와 결합되어 이상과 같은 딥러닝 이미지 분석과 연동하는 신텍스 기반의 선별 관계 제공 방법을 실행시키기 위하여 매체에 저장된 것이다.

발명의 효과

- [0029] 본 발명에 따르면 압축영상을 파싱하여 얻은 신텍스 정보에 기반하여 선별 관계를 수행하므로 복잡한 알고리즘의 이미지 분석 작업을 수행할 필요가 없어 영상분석 서버의 채널 수용량을 대폭 증가시킬 수 있는 장점이 있다. 발명자의 자체 테스트에 따르면 고성능 개인용컴퓨터를 기준으로 종래에 10여 채널에 불과하였던 처리용량을 최대 64 채널까지 확대할 수 있었다.
- [0030] 또한, 본 발명에 따르면 신텍스 기반의 선별 관계를 딥러닝 이미지 분석과 연동시킴으로써 일련의 영상 프레임에 대해 딥러닝 객체 감지를 수행하는 종래의 방식이 아니라 신텍스 기반으로 추출된 이동객체 영역의 이미지에 대해 딥러닝 이미지 분류를 수행하는 새로운 방식을 구성하였고, 이를 통해 객체 분석의 정확도 개선과 CCTV 채널 수용량의 확대라는 서로 양립하기 어려운 두가지 기술적 과제를 모두 달성할 수 있는 장점이 있다.

도면의 간단한 설명

- [0032] [도 1]은 동영상 디코딩 장치의 일반적인 구성을 나타내는 블록도.
- [도 2]는 종래기술에서 CCTV 영상으로부터 이동객체를 감지해내는 과정을 나타내는 순서도.
- [도 3]은 본 발명에 따라 딥러닝 이미지 분석과 연동하여 신택스 기반의 선별 관제를 제공하는 전체 프로세스를 나타내는 순서도.
- [도 4]는 본 발명에서 딥러닝 이미지 분석과 연동하여 신택스 기반의 선별 관제를 제공하는 개념을 나타내는 도면.
- [도 5]는 본 발명에서 선별 관제의 개념을 나타내는 도면.
- [도 6]은 여러 신경망 분석 모델 별로 정확도와 처리 속도를 비교하여 나타낸 도면.
- [도 7]은 선별 관제와 딥러닝을 연동시킨 구성의 성능 개선을 보여주는 실험 데이터를 나타낸 도면.
- [도 8]은 본 발명에서 압축영상으로부터 유효 움직임 영역을 검출하는 과정의 구현 예를 나타내는 순서도.
- [도 9]는 CCTV 영상에 대해 유효 움직임 영역 검출 과정을 적용한 결과의 일 예를 나타낸 도면.
- [도 10]은 본 발명에서 이동객체 영역에 대한 바운더리 영역을 검출하는 과정의 구현 예를 나타내는 순서도.
- [도 11]은 [도 9]의 CCTV 영상 이미지에 대해 바운더리 영역 검출 과정을 적용한 결과의 일 예를 나타낸 도면.
- [도 12]는 [도 11]의 CCTV 영상 이미지에 대해 인터플레이션을 통해 이동객체 영역을 정리한 결과의 일 예를 나타낸 도면.
- [도 13]은 이동객체 영역에 Unique ID가 할당된 일 예를 나타내는 도면.

발명을 실시하기 위한 구체적인 내용

- [0033] 이하에서는 도면을 참조하여 본 발명을 상세하게 설명한다.
- [0035] [도 3]은 본 발명에 따라 딥러닝 이미지 분석과 연동하여 신택스 기반의 선별 관제를 제공하는 전체 프로세스를 나타내는 순서도이고, [도 4]는 본 발명에서 딥러닝 이미지 분석과 연동하여 신택스 기반의 선별 관제를 제공하는 개념을 나타내는 도면이다.
- [0036] 본 발명에 따른 영상 처리 프로세스는 다수의 압축영상을 다루는 시스템, 예컨대 CCTV 영상관제 시스템 또는 CCTV 영상분석 시스템에서 영상분석 서버가 수행할 수 있다. 본 발명의 영상 처리 프로세스는 심층신경망(deep neural network)을 이용하는데, 영상분석 서버 내부에 소프트웨어 혹은 하드웨어로 딥러닝을 구현할 수도 있고, 혹은 외부의 클라우드 서버에 구현된 심층신경망을 오픈 API(Open Application Program Interface)를 통해 활용할 수도 있다.
- [0037] 먼저, [도 4]를 참조하여 본 발명에서 딥러닝 이미지 분석과 연동하여 신택스 기반의 선별 관제를 제공하는 개념을 살펴본다.
- [0038] 선별 관제는 [도 5]에 도시된 바와 같이 유의미한 이벤트가 감지되는 CCTV 영상만 모니터에 표시하는 것이다. [도 5]의 (A)는 일반 관제 방식으로서 모든 CCTV 영상이 모니터에 표시된다. 반면, [도 5]의 (B)는 선별 관제 방식으로서 유의미한 이벤트가 감지된 CCTV 영상만 모니터에 표시하며, 이를 통해 관제 요원이 효과적으로 상황을 판단할 수 있게 해준다.
- [0039] 이와 같은 선별 관제를 수행하려면 각각의 CCTV 영상에 대해 유의미한 이벤트가 존재하는지 여부를 판단해야 한다. 영상의 내용을 소프트웨어 처리를 통해 정확하게 파악하는 것은 상당히 어려운 일이기 때문에 본 발명에서는 유의미한 움직임이 존재한다면 이를 유의미한 이벤트로 추정한다. 따라서, CCTV 영상으로부터 이동객체로 보이는 부분을 식별해내야 한다.
- [0040] 본 발명에서는 압축영상의 비트스트림을 파싱하여 각 영상 블록에 대한 신택스 정보(syntax information)를 통

해 이동객체 영역을 추출한다. 이동객체 영역이 추출되지 않는 CCTV 영상은 무시하고, 이동객체 영역이 추출되는 CCTV 영상공급 채널을 취합하여 스마트 선별 관제를 제공한다.

[0041] 이때, 압축영상에서 선택스 정보를 추출하기 위한 영상 블록으로는 매크로블록(Macro Block) 및 서브블록(Sub Block) 등의 어느 하나 혹은 이들의 조합을 채택할 수 있고, 선택스 정보로는 모션벡터(Motion Vector)와 코딩 유형(Coding Type)이 바람직하다. 이러한 접근 방식은 압축영상을 파싱하여 얻는 선택스 정보에 기초한 것이기 때문에 압축영상을 전체적으로 디코딩한 후에 복잡한 이미지 분석을 수행하는 종래기술에 비해 영상 프레임의 소프트웨어 처리 속도가 대단히 빠르다. 또한, 이렇게 얻어진 이동객체 영역은 본 명세서에 첨부된 여러 이미지에서 확인되는 바와 같이 압축영상 내에 존재하는 이동객체의 경계선을 정확하게 반영하지는 못하지만 추출 처리속도가 빠르면서도 유의미한 움직임 여부 판단의 정확도가 비교적 높은 장점이 있다

[0042] 또한, 본 발명에서는 딥 러닝과 연동함으로써 CCTV 영상관제에서 좀더 고차원적인 서비스를 제공하려고 한다. 심층신경망은 영상에서 객체를 감지하거나 이미지에서 객체를 식별하는 능력이 훌륭하기 때문에 CCTV 영상관제 시스템에서 이를 활용하려는 것이다. 그런데, 종래기술에서 전술한 바와 같이 딥러닝을 CCTV 영상관제 시스템에 연동하는 것은 현재로서는 곤란하다.

[0043] [도 6]은 여러 신경망 분석 모델 별로 정확도와 처리 속도를 비교하여 나타낸 도면이다. 일반적으로 딥 러닝 알고리즘의 해결 과제는 이미지 한장이 주어졌을 때에 그 내용을 분류하는 것(예: 개 사진과 고양이 사진의 구분)과 동영상이 주어졌을 때에 그 안에서 객체를 감지하는 것(예: CCTV 영상에서 사람이 지나가는 것을 감지)이 있다. CCTV 영상관제 시스템과 연동하는 것은 후자의 과제에 해당하는데, 고성능 컴퓨터를 활용하더라도 연산처리 속도는 초당 6 ~ 25 프레임 정도에 불과하고 객체 감지의 정확도 또한 66% 정도로 매우 낮다. 처리속도와 정확도 모두에서 실제 환경에 적용이 곤란한 수준이다.

[0044] 반면, [도 6]을 참조하면, 딥 러닝을 이용하여 이미지를 분류하는 것은 처리 속도와 분류 정확도 모두 훌륭하다. 처리 속도도 4배 정도 빠르고 정확도도 97% 정도이므로 실제 환경에 적용할 수 있을 정도가 된다. 이에, 본 발명에서는 CCTV 영상관제 시스템이 딥 러닝과 연동함에 있어서 딥 러닝의 객체 감지 기능이 아니라 딥 러닝의 이미지 분류 기능을 활용하도록 처리 프로세스를 특수하게 구성하였으며 이를 통해 처리속도와 정확도 모두 크게 개선하였다.

[0045] 앞서의 과정에서 CCTV 영상에 대해 선택스 기반으로 이동객체 영역을 추출하였는데, 이어서 압축영상을 디코딩하여 이동객체 영역의 이미지를 획득한다. 이동객체 영역의 이미지를 딥러닝 이미지 분석기로 제공하여 딥 러닝 기반의 이미지 분류를 수행하며, 그 분류 결과를 얻는다. 즉, 앞서 CCTV 영상에서 무언가 유의미한 움직임을 보이는 이미지 덩어리를 발견하였다면, 딥 러닝 기반으로 그 이미지 덩어리가 무엇인지(예: 사람, 개, 자동차, 자전거 등) 식별해낸 것이다. 추가로, 그 이미지 덩어리(이동객체 영역)의 위치 정보는 해당 영상블록이 이미지 프레임에서 놓여지는 위치로부터 얻을 수 있다.

[0046] 결국, 이러한 정보를 결합하면 CCTV 영상으로부터 객체를 감지한 결과가 얻어지는 것이다. 그런데, CCTV 영상으로부터 선택스 기반으로 이동객체 영역을 추출하는 과정과 이동객체 영역에 대해 이미지를 획득하는 과정, 그리고 이동객체 영역의 이미지를 딥러닝 이미지 분석기에 입력하여 이미지 분류 결과를 얻는 과정 모두 처리 속도가 빠르고 정확도도 매우 높다. 따라서, CCTV 영상을 직접 딥 러닝 기반으로 객체 감지하는 종래의 방식에 비해 본 발명에 따른 처리 프로세스는 처리 속도와 정확도 모두에 훨씬 우수한 결과를 제공한다.

[0047] 그리고, 이상의 과정을 통하여 CCTV 영상으로부터 고속으로 객체를 탐지하였으므로, 미리 설정된 특수한 지점 혹은 관제 요원이 관제 소프트웨어 메뉴를 통해 화면 상에 지정한 특수한 지점인 ROI(관심영역)를 기준으로 객체 감지 서비스(예: 침입, 월담, 배회 등)를 제공할 수 있다.

[0049] 다음으로, [도 3]을 참조하여 본 발명에 따라 딥러닝 이미지 분석과 연동하여 선택스 기반의 선별 관제를 제공하는 과정을 살펴본다.

[0050] 단계 (S100) : 먼저, 압축영상의 모션벡터에 기초하여 압축영상으로부터 실질적으로 의미를 인정할만한 유효 움직임을 검출하며, 이처럼 유효 움직임이 검출된 영상 영역을 이동객체 영역으로 설정한다.

[0051] 이를 위해, H.264 AVC 및 H.265 HEVC 등의 동영상압축 표준에 따라서 압축영상의 코딩 유닛(coding unit)의 모션벡터와 코딩유형을 파싱한다. 이때, 코딩 유닛의 사이즈는 일반적으로 64x64 픽셀 내지 4x4 픽셀 정도이며 설계자의 선택에 따라 다양하게 설정될 수 있다.

- [0052] 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.
- [0053] 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 파악화된 영상 블록의 형태로 검출된 이동객체 영역을 서로 연결하여 유의미한 덩어리 형태를 만들어가는 결과를 얻는다.
- [0054] 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.
- [0055] 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 예상한다. 즉, 이동객체가 찍힌 영상 블록 근방에 존재하는 영상 블록에는 이동객체가 찍혀있을 가능성이 그렇지 않은 영상 블록에 비해 상대적으로 높다고 예상한다. 그러한 기술적 사상을 반영하여 (S200)에서는 이동객체 영역 주변에 존재하는 이웃 블록에 대해 상대적으로 완화된 판단기준에 따라 이동객체 영역에 해당하는지 여부를 판단한다.
- [0056] 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.
- [0057] 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 파편화되었을 때의 손실은 크기 때문이다.
- [0058] 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 11]과 [도 12]를 비교하면 명확하게 드러난다.
- [0059] 단계 (S400) : 앞서의 과정 (S100) 내지 (S300)을 통하여 압축영상에서 이동객체 영역으로 마킹된 복수의 영상 블록들이 상호 연결된 덩어리를 본 발명에서는 압축영상의 이동객체 영역으로 설정한다.
- [0060] 앞서의 과정 (S100) 내지 (S300)을 통하여 압축영상으로부터 하나이상의 이동객체 영역(region of moving object)을 획득하였다. 이렇게 획득된 이동객체 영역은 [도 12]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 개념적으로는 압축영상의 신텍스 정보에 기초하여 그 안에 하나이상의 이동객체가 포함되어 있을 것으로 추정되어

압축영상으로부터 구분된 부분이다.

- [0061] 그리고, 이동객체 영역은 개별 프레임 이미지 뿐만 아니라 일련의 영상 프레임 시퀀스에서 일종의 객체처럼 다루어질 수 있다. 영상분석 서버에서의 소프트웨어 처리를 위하여 각각의 이동객체 영역에는 고유 식별정보(Unique ID)를 할당하여 관리하는 것이 바람직하다. [도 12]와 [도 13]을 참조하면 CCTV 영상으로부터 3개의 이동객체 영역(영상블록의 덩어리)이 검출되었으며 이들에 대해 각각 001, 002, 003의 Unique ID가 할당되었다.
- [0062] 단계 (S500) : 이동객체 영역이 설정된 CCTV 영상공급 채널을 취합하여 [도 5]의 (B)에 도시된 바와 같은 선별 관제를 제공한다. 이를 통해 관제 모니터에는 유의미한 이벤트가 감지된 CCTV 영상만 표시되며, 이를 통해 관제 요원이 효과적으로 상황을 판단할 수 있게 해준다
- [0063] 단계 (S600 ~ S800) : 이하에서는, 앞서 선택스 기반으로 획득한 이동객체 영역의 추출 결과를 이용하여 CCTV 영상관제 시스템을 딥 러닝과 효과적으로 연동시키는 부분에 대해 기술한다.
- [0064] 이를 위해, 이동객체 영역의 이미지(예: 썸네일 이미지)를 압축영상을 디코딩하여 획득한다. 본 발명에서 이동객체 영역의 이미지를 얻는 이유는 딥러닝 이미지 분석기에 입력하여 분류 결과(예: 사람, 동물, 자동차, 자전거 등)를 얻기 위함이다. 딥러닝 이미지 분석기는 객체의 이미지를 전달하면 그 객체의 종류를 제공해주는 방식이므로 이러한 딥러닝 이미지 분석기와 연동하려면 이동객체 영역의 이미지가 필요한 것이다.
- [0065] 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.
- [0066] 그리고 나서, 위에서 얻은 이동객체 영역의 이미지를 딥러닝 이미지 분석기에 입력하며, 이를 통해 이미지 분류 정보를 획득한다. 즉, 이동객체 영역의 이미지에 대해 딥러닝 기반의 이미지 분류 정보를 획득한다.
- [0067] 이동객체 영역의 위치 정보는 이동객체 영역을 구성하는 해당 영상블록이 이미지 프레임에서 놓여지는 위치로부터 얻을 수 있다. 이에, 이동객체 영역의 위치 정보와 이미지 분류 정보를 결합하여 CCTV 영상에서의 객체 감지 정보를 획득할 수 있다. CCTV 영상에서 어느 이미지 프레임에서 어느 위치에 어떠한 객체가 존재하는지에 관한 정보를 얻은 것이기 때문이다.
- [0068] 즉, (S600) 내지 (S800)의 과정을 통하여 CCTV 영상에서 딥 러닝 기반으로 객체를 감지해내는 과정을 딥 러닝 기반으로 이미지를 분류하는 문제로 전환시킨 것이며, 이를 통해 CCTV 영상관제 시스템에 딥 러닝을 연동함에 있어서 높은 정확도와 빠른 속도 모두 얻을 수 있다.
- [0069] [도 7]은 선별 관제와 딥러닝을 연동시킨 구성의 성능 개선을 보여주는 실험 데이터를 나타낸 도면이다. 본 발명에 따르면 영상처리 서버를 한대 혹은 두대를 설치하는 것만으로도 엄청나게 많은 객체 분석을 수행할 수 있다는 것을 알 수 있다. 또한, 하드웨어 성능을 약간만 개선한다면 객체분석 성능을 대폭 개선할 수 있는 장점도 본 발명은 나타낸다.
- [0070] 단계 (S900) : 이상의 과정을 통해 획득한 객체 감지 정보를 이용하여 ROI 기반의 객체 감지 서비스(침입, 월담, 배회 등)를 제공한다. 일반적으로 CCTV 영상관제에 있어서 모든 공간을 동일한 수준으로 감시하는 것은 아니고, 특정 시간대(예: 자정에서 오전 6시까지)에 그러한 상황이 발생하지 않아야 정상인 지점을 선별하거나 특별히 중요한 지점을 선별하여 별도로 관리할 수 있는데, 이러한 지점을 '관심영역(Region of Interest, ROI)'라고 부른다. CCTV 영상으로부터 객체를 탐지하였으므로, 미리 설정된 특수한 지점 혹은 관제 요원이 관제 소프트웨어 메뉴를 통해 화면 상에 지정한 특수한 지점인 ROI(관심영역)를 기준으로 객체 감지 서비스(예: 침입, 월담, 배회 등)를 제공한다.
- [0071] 이때, 시간 흐름에 대응하여 CCTV 영상에서 동일한 객체에 대해 ROI 기반의 객체 감지를 수행하기 위해 동일 Unique ID가 할당된 이동객체 영역을 기준으로 프로세스를 진행할 수 있다.
- [0073] [도 8]은 본 발명에서 이동객체 영역을 추출하는 첫번째 단계로서 압축영상에서 유효 움직임(effective movement) 영역을 검출하는 과정의 구현 예를 나타내는 순서도이고, [도 9]는 CCTV 영상에 대해 유효 움직임 영역 검출 과정이 적용된 결과의 일 예를 나타내는 도면이다. [도 8]의 프로세스는 [도 3]에서 단계 (S100)에 대응한다.

- [0074] 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.
- [0075] 단계 (S120) : 압축영상의 프레임 이미지에서 프레임 이미지의 전역에 대해 지배적인 모션벡터(dominant motion vector)가 균일하게 검출되는 전역 균일 모션벡터 상황(globally uniform motion vector status)인지 여부를 식별하고 만일 그러하다면 프레임을 무시하거나 혹은 지배적인 모션벡터를 균일하게 소거함으로써 전역 균일 모션벡터 상황을 필터링 제거한다.
- [0076] 압축영상에서 유효 움직임 검출함에 있어서는 2가지를 필터링해야 하는데, 첫번째는 눈/비 내림, 야간 노이즈, 바람에 흔들리는 나뭇가지/현수막과 같은 무의미한 영상 변화이다. 이는 (S130)과 (S140)에서 모션벡터 누적값을 기준으로 이루어지는 데이터 처리를 통해 필터링된다.
- [0077] 두번째는 카메라의 PTZ 동작에 따른 화면 이동이나 CCTV 지지대의 진동에 따른 화면 떨림 등이다. 이러한 상황에서는 필연적으로 CCTV 영상이 전체적으로 크게 움직이게 되는데, 신택스 정보의 측면에서는 이미지 프레임 전방에 걸쳐서 모션벡터가 상당한 값이 나오게 된다. 따라서, 이러한 상황에서는 이후의 (S130)에서 모든 영상 블록에 대해 모션벡터 누적값이 큰 값이 얻어지게 되고, 그로 인해 실제로는 그 지점에서 별다른 일이 일어나지도 않았는데도 불구하고 프레임 이미지를 구성하는 전체 영상 블록에서 유효한 움직임이 존재하는 것처럼, 즉 무언가 큰 사건이 벌어진 것처럼 잘못 검출될 가능성이 존재한다.
- [0078] 이러한 상황에서는 프레임 이미지의 전역에 대해 지배적인 모션벡터가 균일하게 검출되는 현상이 나타난다. CCTV 영상 내에 일부 움직이는 객체가 존재할 수 있기 때문에 영상 블록 별로 약간씩은 차이가 날 수도 있지만, 것처럼 미세한 것과는 구별되는 지배적인 모션벡터가 프레임 이미지의 전역에 대해 검출된다. 이러한 전역 균일 모션벡터 상황이 검출된다면, 그 자체는 관계 요인이 주목할만한 이벤트라고 볼 수 없으므로 유효 움직임으로 검출되지 않도록 처리한다. 예를 들어, 해당 프레임 이미지를 무시하거나, 모든 영상 블록에서 얻어지는 모션벡터 값에서 지배적인 모션벡터 값(예: 이미지 프레임에서 얻은 모든 모션벡터 값의 평균값)을 균일하게 소거(substraction)할 수 있다.
- [0079] 단계 (S130) : 압축영상을 구성하는 복수의 영상 블록 별로 미리 설정된 시간(예: 500 ms) 동안의 모션벡터 누적값을 획득한다.
- [0080] 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.
- [0081] 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.
- [0082] 단계 (S140, S150) : 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.
- [0083] 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 관계 요인이 관심을 가질만한 가치가 있을 정도의 움직임을 선별하여 검출하려는 것이다. 반대로, 모션벡터가 발생하였더라도 일정 시간동안의 누적값이 제 1 임계치를 넘지 못할 정도로 작을 경우에는 영상에서의 변화가 그다지 크지않고 미미한 것으로 추정하고 검출 단계에서 무시한다.
- [0084] [도 9]는 본 발명에서 [도 8]의 과정을 통해 CCTV 영상으로부터 유효 움직임 영역을 검출한 결과를 시각적으로 나타낸 일 예이다. [도 9]에서는 제 1 임계치 이상의 모션벡터 누적값을 갖는 영상 블록이 이동객체 영역으로 마킹되어 붉은 색으로 표시되었다. [도 9]를 살펴보면 보도블럭이나 도로, 그리고 그림자가 있는 부분 등은 이동객체 영역으로 표시되지 않은 반면, 걷고있는 사람들이나 주행중인 자동차 등이 이동객체 영역으로 표시되었다.
- [0086] [도 10]은 본 발명에서 이동객체 영역을 추출하는 두번째 단계로서 이동객체 영역에 대한 바운더리 영역

(boundary area)을 검출하는 과정의 구현 예를 나타내는 순서도이고, [도 11]은 [도 9]의 CCTV 영상 이미지에 대해 바운더리 영역 검출 과정이 추가로 적용된 결과의 일 예를 나타내는 도면이다. [도 10]의 프로세스는 [도 3]에서 단계 (S200)에 대응한다.

- [0087] 앞서의 [도 9]를 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.
- [0088] 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 이웃 블록은 (S100)에 의해서는 이동객체 영역으로 마킹되지 않은 부분인데, [도 10]의 프로세스에서는 이들에 대해 좀더 살펴봄으로써 이들 이웃 블록 중에서 이동객체 영역의 바운더리에 포함될만한 것이 있는지 확인하려는 것이다.
- [0089] 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.
- [0090] 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.
- [0091] 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.
- [0092] 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.
- [0093] [도 11]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 11]을 살펴보면, 앞서 [도 9]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.
- [0095] [도 12]는 [도 11]의 CCTV 영상 이미지에 대해 인터플레이션을 통해 이동객체 영역을 정리한 결과의 일 예를 나타낸 도면이다.
- [0096] 단계 (S300)은 앞의 (S100)과 (S200)에서 검출된 이동객체 영역에 인터플레이션을 적용하여 이동객체 영역의 분할을 정리하는 과정이다. [도 11]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.
- [0097] 그에 따라, 본 발명에서는 이동객체 영역으로 마킹된 복수의 영상 블록으로 둘러싸여 하나 혹은 소수의 비마킹

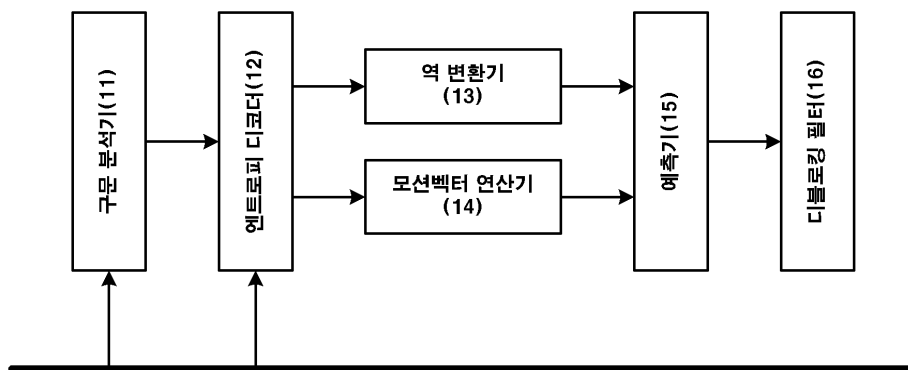
영상 블록이 존재한다면 이는 이동객체 영역으로 마킹하는데, 이를 인터폴레이션이라고 부른다. [도 11]과 대비하여 [도 12]를 살펴보면, 이동객체 영역 사이사이에 존재하던 비마킹 영상 블록이 모두 이동객체 영역이라고 마킹되었다. 이를 통해, 덩어리로 움직이는 영역은 모두 묶어서 하나의 이동객체로서 다루게 된다.

[0098] [도 9], [도 11], [도 12]를 비교하면 바운더리 영역 검출 과정과 인터폴레이션 과정을 거치면서 이동객체 영역이 실제 영상의 상황을 제대로 반영하게 되어간다는 사실을 발견할 수 있다. [도 9]에서 붉은 색으로 마킹된 덩어리로 판단한다면 영상 화면 속에 아주 작은 물체들이 다수 움직이는 것처럼 다루어질 것인데, 이는 실제와는 부합하지 않는다. 반면, [도 12]에서 파란 색으로 마킹된 덩어리로 판단한다면 어느 정도의 부피를 갖는 몇 개의 이동객체가 존재하는 것으로 다루어질 것이어서 실제 장면을 유사하게 반영한다.

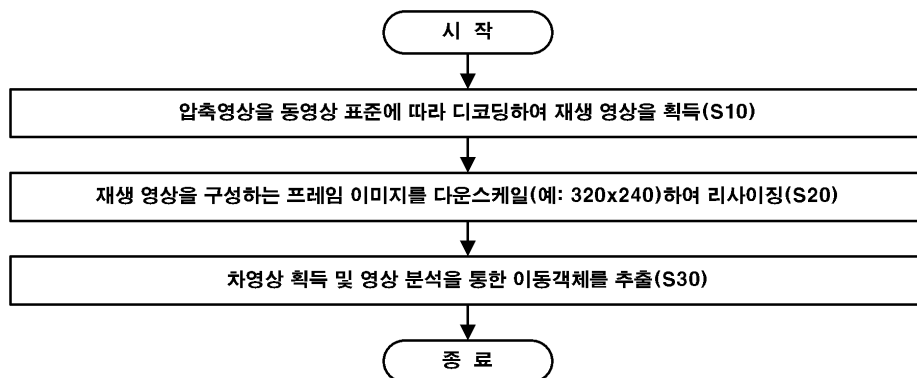
[0100] 한편, 본 발명은 컴퓨터가 읽을 수 있는 비휘발성 기록매체에 컴퓨터가 읽을 수 있는 코드의 형태로 구현되는 것이 가능하다. 이러한 비휘발성 기록매체로는 다양한 형태의 스토리지 장치가 존재하는데 예컨대 하드디스크, SSD, CD-ROM, NAS, 자기테이프, 웹디스크, 클라우드 디스크 등이 있고 네트워크로 연결된 다수의 스토리지 장치에 코드가 분산 저장되고 실행되는 형태도 구현될 수 있다. 또한, 본 발명은 하드웨어와 결합되어 특정의 절차를 실행시키기 위하여 매체에 저장된 컴퓨터프로그램의 형태로 구현될 수도 있다.

도면

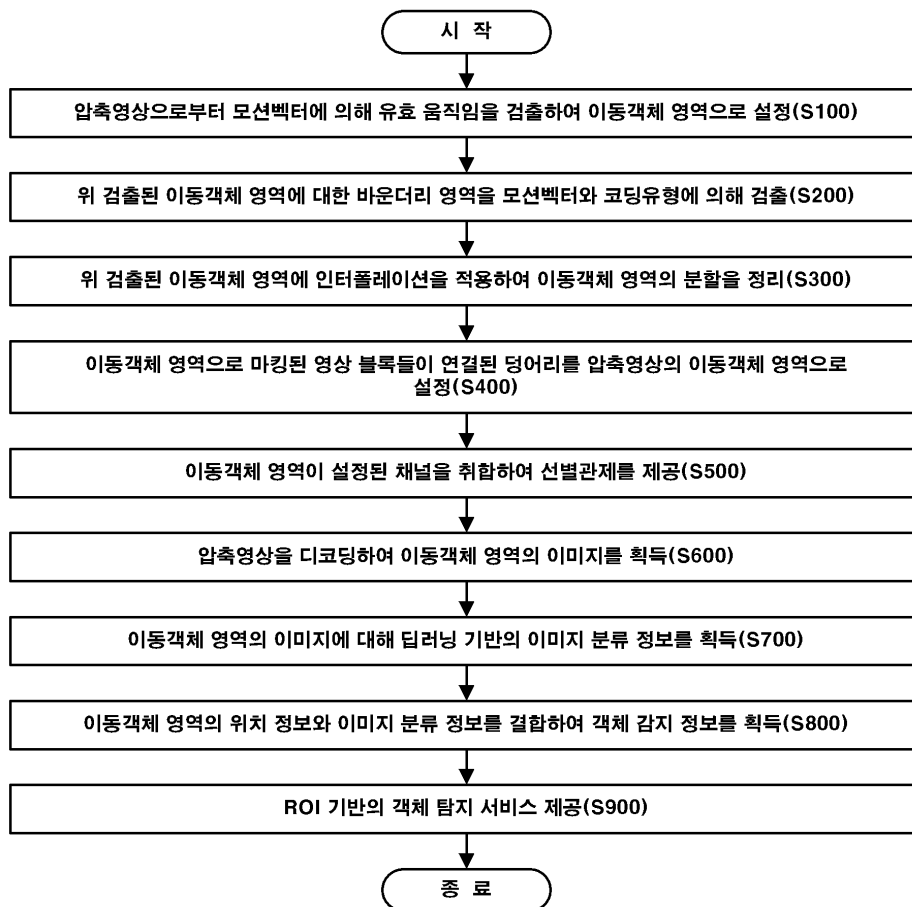
도면1



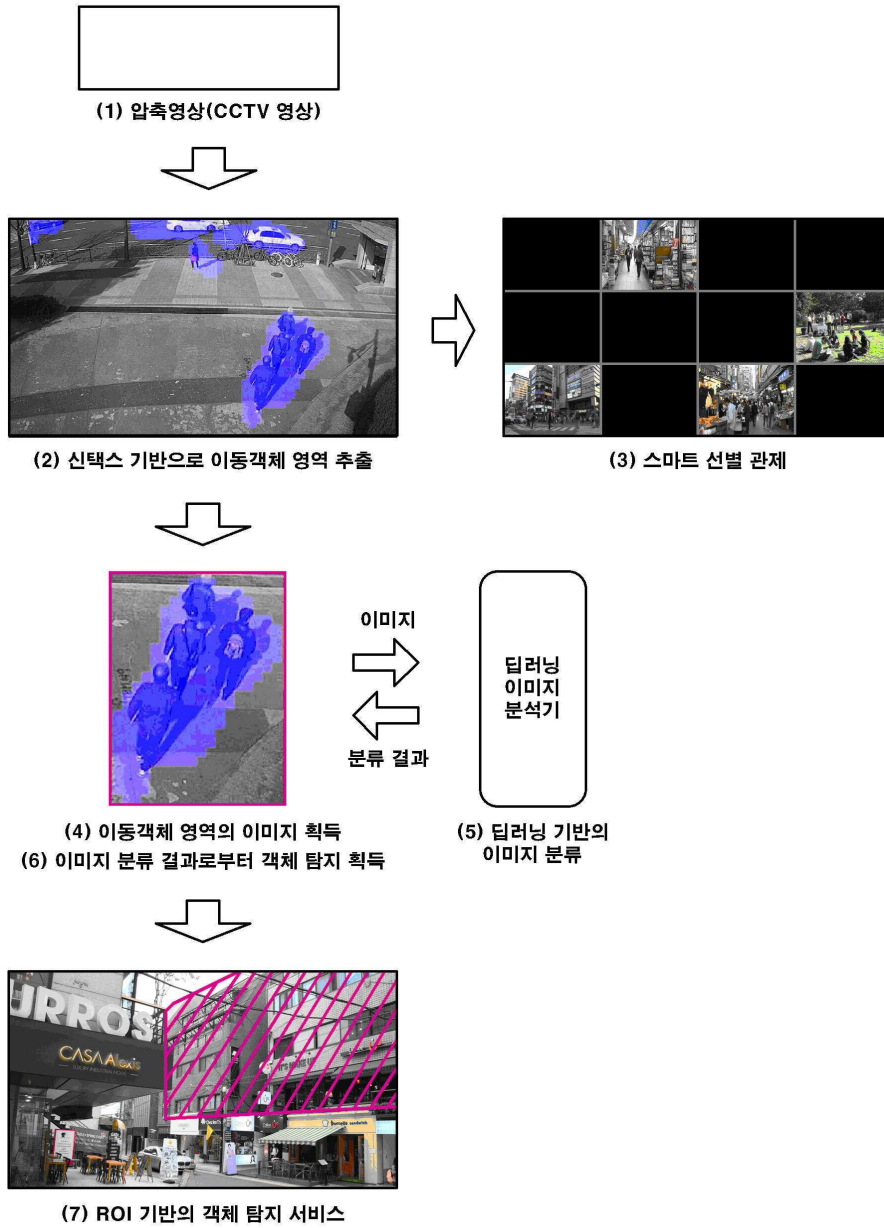
도면2



도면3



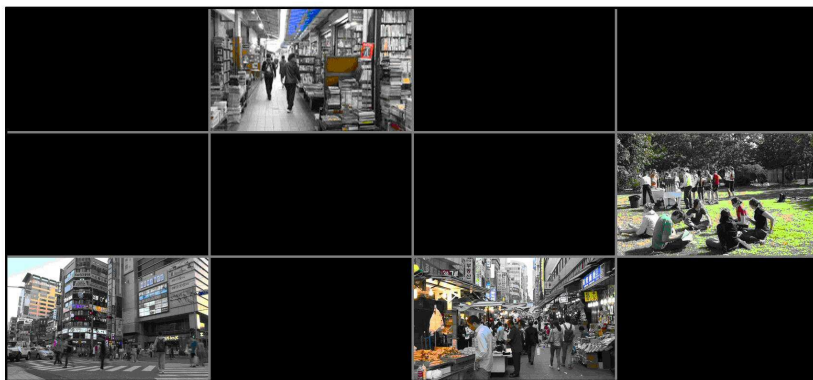
도면4



도면5



(A) 일반 관제 방식



(B) 선별 관제 방식

도면6

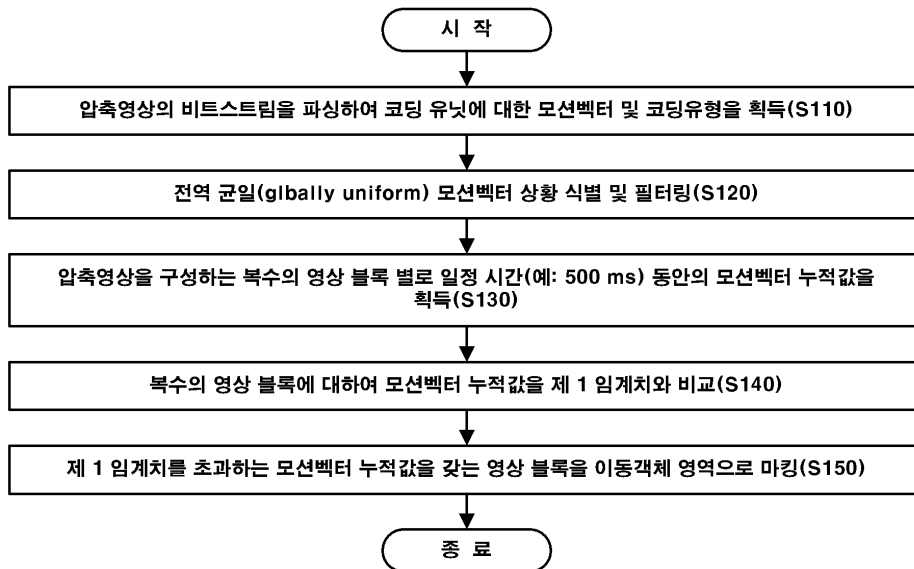
분석 모델	정확도	처리시간	FPS
머신 러닝(machine learning)	72%	가변	가변
피상적 러닝(shallow learning)	88%	5.4 ms	200
딥 러닝(이미지 분류)	97%	10~48 ms	20~100
딥 러닝(객체 감지)	66%	40~150 ms	6~25

<데이터세트: ImageNet, GPU: GTX 1080TI 기준>

도면7

구성 요소	딤러닝 기준 (이미지 분류)	증감율	테스트 결과
그래픽 카드 종류	GTX1080TI	▶ 1.5배	GV100
그래픽 카드 개수	1개	▶ 4배	4개
서버 수	1대	▶ 2배	2대
입력 이미지 사이즈	224 X 224	▶ 1~1/4배	(112~224) X (112~224)
딤러닝 모델 압축률	0%	▶ 1~2배	0~50%
객체 분석 수	100개	▶ 12~96배	1200~9600개

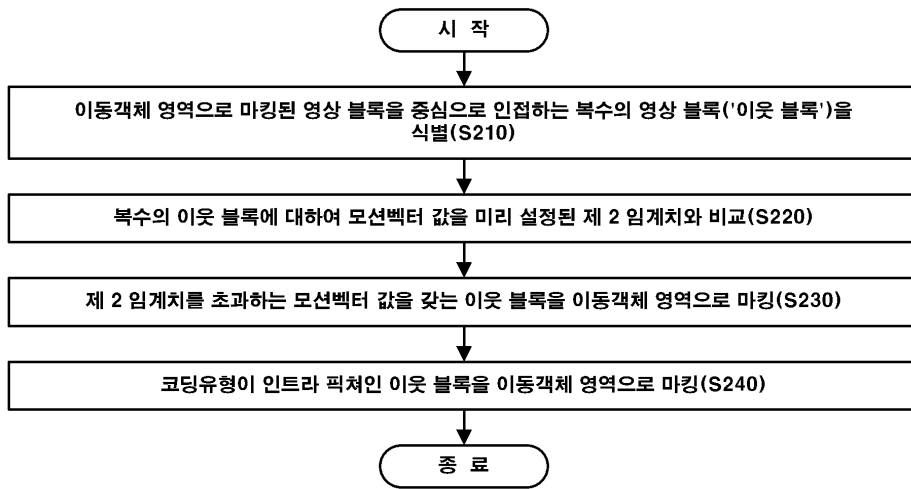
도면8



도면9



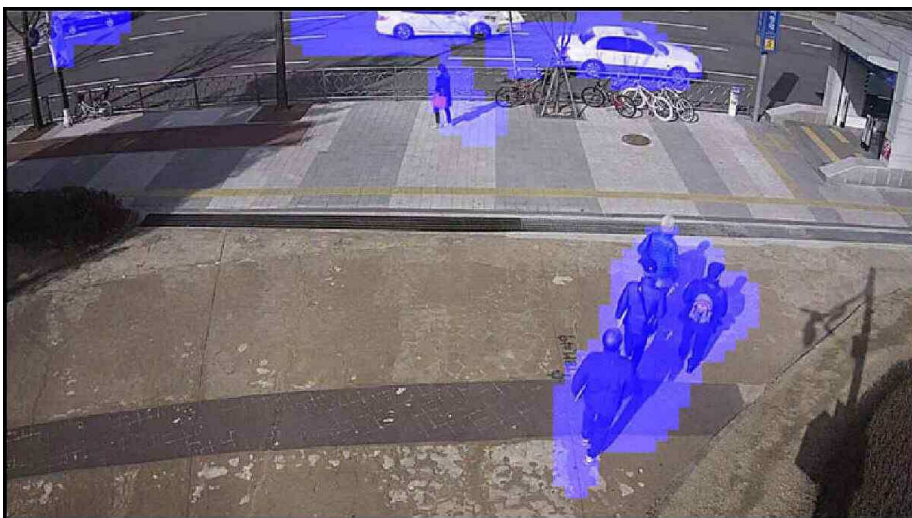
도면10



도면11



도면12



도면13

