



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년11월27일
(11) 등록번호 10-2183473
(24) 등록일자 2020년11월20일

(51) 국제특허분류(Int. Cl.)
HO4N 7/18 (2006.01) G06T 5/50 (2006.01)
HO4N 5/232 (2006.01)
(52) CPC특허분류
HO4N 7/181 (2013.01)
G06T 5/50 (2013.01)
(21) 출원번호 10-2020-0091087
(22) 출원일자 2020년07월22일
심사청구일자 2020년07월22일
(56) 선행기술조사문헌
KR1020030027575 A
KR1020120058050 A
KR1020140052357 A

(73) 특허권자
(주)대교씨엔에스
서울특별시 관악구 보라매로3길 23(봉천동, 대교빌딩 7층)
(72) 발명자
김성연
서울특별시 은평구 서오릉로 115-10, 삼익하이빌 301호
송승철
서울특별시 강서구 양천로 140, 방화현대2차아파트 204동 403호
(뒷면에 계속)
(74) 대리인
이용성, 서평강

전체 청구항 수 : 총 4 항

심사관 : 박재학

(54) 발명의 명칭 영상 모니터링 방법 및 장치

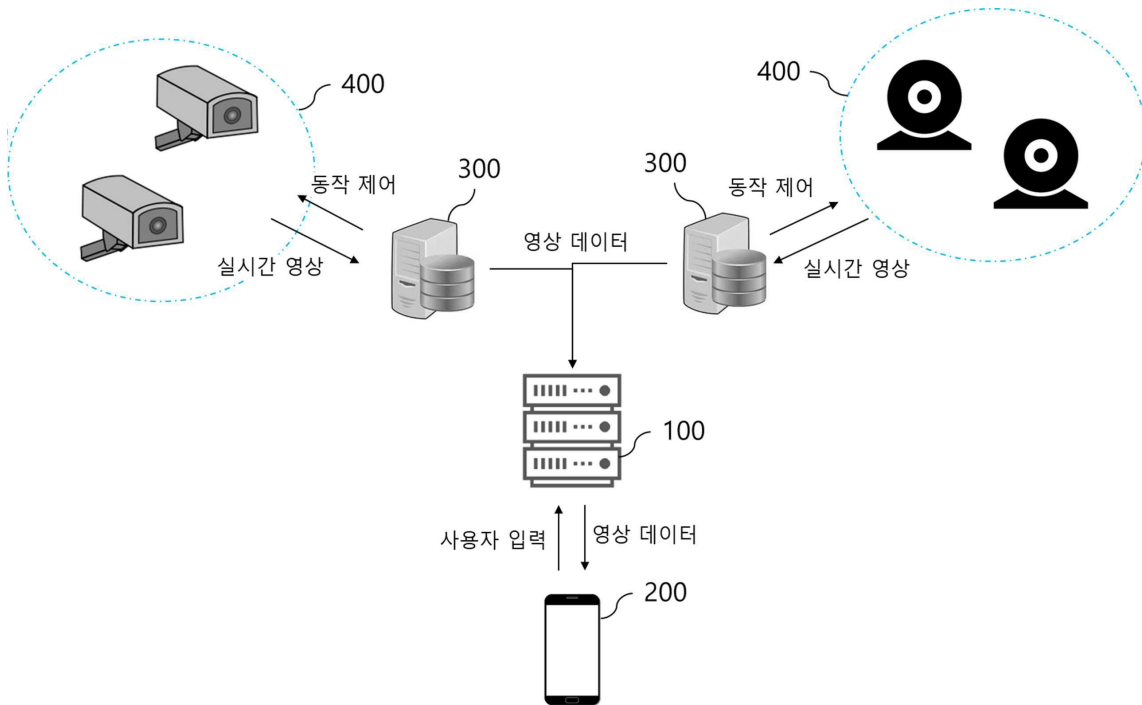
(57) 요약

본 발명에서는 영상 모니터링 장치 및 그 방법이 개시된다.

본 발명의 영상 모니터링 장치는, 복수의 CCTV들을 제어하는 로컬 서버 및 관리 단말과 무선 네트워크를 통해 통신하는 통신부; 상기 관리 단말로부터 입력되는 사용자 입력에 따라 상기 CCTV들에 대한 제어 신호를 생성하는

(뒷면에 계속)

대표도 - 도1



제어 명령 생성부; 상기 통신부를 통해 상기 CCTV들에서 촬영된 실시간 영상 또는 상기 실시간 영상을 인코딩하여 생성된 영상 데이터들을 저장하는 영상 데이터 저장부; 상기 영상 데이터들을 분석하고, 이벤트가 발생한 것으로 판단되는 영상 프레임들을 결합하여 이벤트 영상을 생성하는 이벤트 영상 생성부; 상기 CCTV들 중에서 촬영 방향 조정에 따른 상기 제어 신호를 수신하는 CCTV들을 선정하고, 선정된 CCTV들로부터 획득된 상기 영상 데이터들의 영상 프레임들을 연결하여 합성 영상을 생성하는 합성 영상 생성부; 및 상기 영상 데이터들에서 다수의 객체들을 추출하고, 추출된 객체들 중 관심 객체가 포함된 영상 프레임들을 조합하여 객체 영상을 생성하는 객체 영상 생성부를 포함한다.

(52) CPC특허분류

G06T 7/292 (2017.01)

H04N 5/23206 (2018.08)

(72) 발명자

허은지

인천광역시 부평구 안남로206번길 18 407동 101호
(산곡동, 경남4차아파트)

김경수

서울특별시 노원구 섭발로 139(공릉동, 공릉풍림아파트)

명세서

청구범위

청구항 1

복수의 CCTV들을 제어하는 로컬 서버 및 관리 단말과 무선 네트워크를 통해 통신하는 통신부;

상기 관리 단말로부터 입력되는 사용자 입력에 따라 상기 CCTV들에 대한 제어 신호를 생성하는 제어 명령 생성부;

상기 통신부를 통해 상기 CCTV들에서 촬영된 실시간 영상 또는 상기 실시간 영상을 인코딩하여 생성된 영상 데이터들을 저장하는 영상 데이터 저장부;

상기 영상 데이터들을 분석하고, 이벤트가 발생한 것으로 판단되는 영상 프레임들을 결합하여 이벤트 영상을 생성하는 이벤트 영상 생성부;

상기 CCTV들 중에서 촬영 방향 조정에 따른 상기 제어 신호를 수신하는 CCTV들을 선정하고, 선정된 CCTV들로부터 획득된 상기 영상 데이터들의 영상 프레임들을 연결하여 합성 영상을 생성하는 합성 영상 생성부;

상기 영상 데이터들에서 다수의 객체들을 추출하고, 추출된 객체들 중 관심 객체가 포함된 영상 프레임들을 조합하여 객체 영상을 생성하는 객체 영상 생성부; 및

상기 실시간 영상, 상기 이벤트 영상, 상기 합성 영상, 상기 객체 영상 중 적어도 하나의 크기와 위치를 결정하고, 결정된 크기와 위치에 따라 격자 형태의 다중 영상을 상기 관리 단말을 통해 표시하는 다중 영상 표시부를 포함하고,

상기 이벤트 영상 생성부는,

상기 CCTV들 중에서 제1 CCTV로부터 획득된 영상 데이터에서 제1 프레임의 일부 영역에 해당하는 관심 영역을 상기 관리 단말로부터 입력받고, 상기 관심 영역에서 배경 영역을 추출하고, 상기 배경 영역에서 객체 추출 알고리즘을 이용하여 움직임이 없는 적어도 하나의 고정 객체를 추출하고, 상기 CCTV들 중에서 상기 제1 CCTV를 제외한 나머지 CCTV들로부터 획득된 영상 데이터들 각각을 이용하여 배경 영상 프레임들을 생성하고, 상기 배경 영상 프레임들 중에서 상기 객체 추출 알고리즘을 이용해 상기 적어도 하나의 고정 객체와 동일성이 있는 객체를 검출하고, 상기 동일성이 있는 객체가 검출된 상기 배경 영상 프레임들과 대응하는 적어도 하나의 CCTV를 결정하고, 결정된 상기 적어도 하나의 CCTV에서 획득된 영상 데이터들을 이용하여 추가 이벤트 영상을 생성하고,

상기 객체 영상 생성부는,

상기 제1 CCTV로부터 획득된 영상 데이터를 이용하여 배경 영상 프레임을 생성하고, 생성된 배경 영상 프레임을 상기 제1 CCTV로부터 획득된 영상 데이터의 영상 프레임들 각각에 대해 차분함으로써 차분 영상 프레임들을 생성하고, 차분 영상 프레임들 각각에 포함된 화소값들의 총합이 임계값 이상인 차분 영상 프레임들을 대상으로 복수의 이동 객체들을 추출하고,

상기 복수의 이동 객체들 중에서 선택된 관심 객체를 구성하는 상기 차분 영상 프레임들의 영역들을 4개 이상의 화소값으로 구성된 복수의 객체셀들로 구획하고, 상기 객체셀들 각각에 대해 색상, 채도 및 명도의 평균값을 산출함으로써 상기 객체셀들과 대응하는 색상셀들, 채도셀들, 및 명도셀들을 결정하고, 상기 색상셀들, 상기 채도셀들, 및 상기 명도셀들을 각각 제1 특징값, 제2 특징값, 및 제3 특징값으로 설정하고, 설정된 상기 제1 특징값, 상기 제2 특징값, 및 상기 제3 특징값을 상기 객체 추출 알고리즘의 비교 기준으로 사용하여 상기 관심 객체와 동일성이 있는 객체를 상기 CCTV들로부터 획득된 영상 데이터들에서 검색하되,

상기 객체 영상 생성부는,

상기 CCTV들로부터 획득된 영상 데이터들에서 이동 객체들을 검출하고, 검출된 이동 객체들 중 상기 제1 특징값을 기반으로 상기 관심 객체와 가장 유사한 제1 후보 객체, 상기 제2 특징값을 기반으로 상기 관심 객체와 가장 유사한 제2 후보 객체, 및 상기 제3 특징값을 기반으로 상기 관심 객체와 가장 유사한 제3 후보 객체를 선정하고, 화소값 기반으로 상기 제1 후보 객체, 상기 제2 후보 객체, 및 상기 제3 후보 객체 중에서 상기 관심 객체

와 동일성이 있는 객체를 선정하는, 영상 모니터링 장치.

청구항 2

청구항 1에서,

상기 제어 신호는,

상기 CCTV들 중 적어도 하나에 대한 줌 인 제어, 조리개 값 조정, 포커스 조정, 촬영 방향 조정 중 하나를 지시하는,

영상 모니터링 장치.

청구항 3

청구항 1에서,

상기 다중 영상 표시부는,

상기 관리 단말로부터 상기 CCTV들 중 하나를 드래그 앤 드롭하는 선택 입력을 수신함에 따른 응답으로, 상기 CCTV들 중 상기 선택 입력과 대응하는 실시간 영상을 상기 다중 영상에 표시하는,

영상 모니터링 장치.

청구항 4

삭제

청구항 5

청구항 1에서,

상기 합성 영상 생성부는,

상기 선정된 CCTV들로부터 획득된 상기 영상 데이터들의 영상 프레임들 중 시간적으로 연속된 영상 프레임들의 엮어 영역을 서로 연결하여 상기 합성 영상을 생성하는,

영상 모니터링 장치.

발명의 설명

기술 분야

[0001] 본 발명은 영상 모니터링 방법에 관한 것으로, 더욱 상세하게는 다수의 카메라(예; CCTV)를 통해 수집되는 영상들을 모니터링하는 방법 및 장치에 관한 것이다.

배경 기술

[0002] 재해, 교통, 보안 등 다양한 분야에서 안전을 위협하는 여러 이벤트나 객체를 모니터링하기 위하여 CCTV(closed circuit television)가 다양한 장소에 설치되고 있다.

[0003] 설치된 다수의 CCTV들을 이용하여 촬영되는 다수의 영상 데이터들은 관리실에 설치된 중앙 관리 서버로 전송되며, 중앙 관리 서버는 이러한 다수의 영상 데이터들을 실시간으로 관리자에게 디스플레이하여 감시 활동을 수행하도록 지원한다.

[0004] 관리자가 다수의 영상 데이터들을 실시간으로 감시하고 이벤트나 객체를 확인하는 데에는 많은 주의력이 요구되며, 관리자가 자칫 중요한 이벤트나 객체를 확인하지 못하는 일이 빈번하게 발생한다.

[0005] 또한, 모니터링하는 CCTV의 개수가 많은 경우, 각각의 CCTV에서 촬영된 영상 데이터들을 일일이 교차 비교하여

원하는 장면을 확인하기 위해서는 많은 노동력이 투입되는 것이 필요하다.

[0006] 따라서, 다수의 CCTV를 통해 촬영된 영상 데이터들 중에서 관리자가 원하는 대상이나 이벤트를 확인하는 노력을 줄이고 효율적으로 모니터링할 수 있는 방안이 필요한 실정이다.

발명의 내용

해결하려는 과제

[0007] 상기와 같은 문제점을 해결하기 위한 본 발명의 목적은, 다수의 CCTV를 통해 확보된 영상 데이터들로부터 각종 이벤트나 객체를 정확하게 추출하여 원하는 목적에 따라 효율적으로 모니터링할 수 있는, 영상 모니터링 방법 및 장치를 제공하는 데 있다.

과제의 해결 수단

- [0008] 상기 목적을 달성하기 위한 본 발명의 일 측면은, 영상 모니터링 장치를 제공한다.
- [0009] 영상 모니터링 장치는, 복수의 CCTV들을 제어하는 로컬 서버 및 관리 단말과 무선 네트워크를 통해 통신하는 통신부; 상기 관리 단말로부터 입력되는 사용자 입력에 따라 상기 CCTV들에 대한 제어 신호를 생성하는 제어 명령 생성부; 상기 통신부를 통해 상기 CCTV들에서 촬영된 실시간 영상 또는 상기 실시간 영상을 인코딩하여 생성된 영상 데이터들을 저장하는 영상 데이터 저장부; 상기 영상 데이터들을 분석하고, 이벤트가 발생한 것으로 판단되는 영상 프레임들을 결합하여 이벤트 영상을 생성하는 이벤트 영상 생성부; 상기 CCTV들 중에서 촬영 방향 조정에 따른 상기 제어 신호를 수신하는 CCTV들을 선정하고, 선정된 CCTV들로부터 획득된 상기 영상 데이터들의 영상 프레임들을 연결하여 합성 영상을 생성하는 합성 영상 생성부; 상기 영상 데이터들에서 다수의 객체들을 추출하고, 추출된 객체들 중 관심 객체가 포함된 영상 프레임들을 조합하여 객체 영상을 생성하는 객체 영상 생성부; 및 상기 실시간 영상, 상기 이벤트 영상, 상기 합성 영상, 상기 객체 영상 중 적어도 하나의 크기와 위치를 결정하고, 결정된 크기와 위치에 따라 격자 형태의 다중 영상을 상기 관리 단말을 통해 표시하는 다중 영상 표시부를 포함할 수 있다.
- [0010] 상기 제어 신호는, 상기 CCTV들 중 적어도 하나에 대한 줌 인 제어, 조리개 값 조정, 포커스 조정, 촬영 방향 조정 중 하나를 지시할 수 있다.
- [0011] 상기 다중 영상 표시부는, 상기 관리 단말로부터 상기 CCTV들 중 하나를 드래그 앤 드롭하는 선택 입력을 수신함에 따른 응답으로, 상기 CCTV들 중 상기 선택 입력과 대응하는 실시간 영상을 상기 다중 영상에 표시할 수 있다.
- [0012] 상기 이벤트 영상 생성부는, 상기 관리 단말로부터 관찰하고자 하는 관심 영역을 입력받고, 상기 관심 영역의 변화가 감지된 영상 프레임들을 결합하여 상기 이벤트 영상을 생성할 수 있다.
- [0013] 상기 합성 영상 생성부는, 상기 선정된 CCTV들로부터 획득된 상기 영상 데이터들의 영상 프레임들 중 시간적으로 연속된 영상 프레임들의 엷지 영역을 서로 연결하여 상기 합성 영상을 생성할 수 있다.
- [0014] 상기 객체 영상 생성부는, 상기 CCTV들 중 제1 CCTV에서 획득된 영상 데이터의 영상 프레임들을 대상으로 적어도 하나의 이동 객체를 추출하고, 추출된 이동 객체 중 선정된 관심 객체가 검출된 영상 프레임들을 시간 순서에 따라 결합하여 상기 객체 영상을 생성할 수 있다.
- [0015] 상기 객체 영상 생성부는, 상기 제1 CCTV에서 획득된 상기 영상 데이터에서 일정한 시간 간격 내에 위치한 영상 프레임들의 화소값 평균을 산출하고, 산출된 평균을 화소값으로 갖는 배경 영상 프레임을 생성할 수 있다.
- [0016] 상기 객체 영상 생성부는, 상기 배경 영상 프레임을 상기 제1 CCTV에서 획득된 상기 영상 데이터의 영상 프레임들 각각에 대해 차분함으로써 차분 영상 프레임들을 생성하고, 상기 차분 영상 프레임들 중에서, 상기 차분 영상 프레임들 각각을 구성하는 화소값들의 총합이 임계값 이상인 차분 영상 프레임들을 선정하고, 선정된 차분 영상 프레임들에서 상기 적어도 하나의 이동 객체를 추출할 수 있다.
- [0017] 상기 이벤트 영상 생성부는, 상기 CCTV들 중 제1 CCTV에서 획득된 영상 데이터의 제1 프레임의 일부 영역에 해당하는 상기 관심 영역을 상기 관리 단말로부터 입력받고, 상기 관심 영역에서 적어도 하나의 고정 객체를 추출할 수 있다.
- [0018] 상기 이벤트 영상 생성부는, 상기 CCTV들 중 상기 제1 CCTV를 제외한 나머지 CCTV들에서 획득된 각각의 영상 데

이더들마다 하나씩 배경 영상 프레임들을 생성하고, 상기 배경 영상 프레임들을 대상으로 상기 고정 객체와 동일성이 있는 객체를 검출하고, 상기 동일성이 있는 객체가 검출된 상기 배경 영상 프레임들과 대응하는 적어도 하나의 CCTV를 결정할 수 있다.

발명의 효과

[0019] 상기와 같은 본 발명에 따른 영상 모니터링 방법 및 장치를 이용할 경우에는, 다수의 CCTV를 통해 확보된 영상 데이터들로부터 각종 이벤트나 객체를 정확하게 추출하여 원하는 목적에 따라 효율적으로 모니터링할 수 있다.

도면의 간단한 설명

- [0020] 도 1은 일 실시예에 따른 영상 모니터링 방법 및 장치에 대한 개요도이다.
- 도 2는 도 1에 따른 영상 모니터링 장치가 관리 단말에 제공하는 유저 인터페이스를 나타낸 예시도이다.
- 도 3은 도 1에 따른 영상 모니터링 장치에 대한 기능부이다.
- 도 4는 도 3에 따른 이벤트 영상 생성부의 동작을 설명하기 위한 개념도이다.
- 도 5는 도 3에 따른 객체 영상 생성부의 동작을 설명하기 위한 개념도이다.
- 도 6은 도 3에 따른 합성 영상 생성부의 동작을 설명하기 위한 개념도이다.
- 도 7은 도 1에 따른 영상 모니터링 장치의 하드웨어 구성도이다.

발명을 실시하기 위한 구체적인 내용

- [0021] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였다.
- [0022] 제1, 제2, A, B 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.
- [0023] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다.
- [0024] 본 출원에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0025] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥 상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0026] 이하, 본 발명에 따른 바람직한 실시예를 첨부된 도면을 참조하여 상세하게 설명한다.
- [0028] 도 1은 일 실시예에 따른 영상 모니터링 방법 및 장치에 대한 개요도이다. 도 2는 도 1에 따른 영상 모니터링 장치가 관리 단말에 제공하는 유저 인터페이스를 나타낸 예시도이다.
- [0029] 도 1을 참조하면, 본 발명의 일 실시예에 따른 영상 모니터링 장치(100)는, 관리 단말(200), 다수의 장소에 설

치되어 실시간으로 영상을 촬영하는 복수의 CCTV(400)들, 및 상기 CCTV(400)들의 동작을 제어하고 CCTV(400)에서 촬영된 실시간 영상을 수집하여 영상 데이터를 생성하는 로컬 서버(300)와 연동하여 영상 모니터링 방법을 수행할 수 있다.

- [0030] CCTV(400)는 설치된 장소에서 실시간으로 영상을 촬영하고, 촬영된 실시간 영상을 실시간으로 로컬 서버(300)에 전송할 수 있다. 또한, CCTV(400)는 로컬 서버(300)에서 제공되는 제어 신호에 기초하여, 팬(pan), 틸트(tilt), 줌(zoom) 등이 조정됨으로써, 다양한 각도와 방향, 위치에 따라 실시간 영상을 촬영할 수 있다. CCTV(400)는 IP 카메라, 지능형 카메라, 카메라가 내장된 드론(drone) 등을 포함할 수 있다.
- [0031] 로컬 서버(300)는, 다수의 CCTV(400)들이 설치된 장소와 동일한 건물이나 인접한 장소에 설치됨으로써, 다수의 CCTV(400)에 제어 신호를 전송하여 실시간으로 동작 제어를 수행하고, 다수의 CCTV(400)로부터 실시간 영상을 수집하고, 실시간 영상을 인코딩(encoding)하여 영상 데이터를 생성할 수 있다. 로컬 서버(300)는, NVR(network video recorder)을 포함할 수 있다.
- [0032] 예를 들어, 로컬 서버(300)는, CCTV(400)에 줌 인 제어, 조리개 값 조정, 포커스 조정, 촬영 방향 조정 등을 지시하는 제어 신호를 전송할 수 있다. 촬영 방향 조정을 지시하는 제어 신호는, CCTV(400)가 일정한 방향을 주기적으로 왕복하는 스윙(swing) 제어, 특정 경로를 따라 이동하는 트레이싱(tracing) 제어, 미리 설정된 위치로 이동하는 프리셋(preset) 제어 등을 지시하는 신호를 포함할 수 있다.
- [0033] 관리 단말(200)은, 관리자 또는 사용자가 네트워크를 통해 영상 모니터링 장치(100)에 접속하는 단말로서, 영상 모니터링 장치(100)를 통해 다수의 CCTV(400)들의 동작 제어를 수행하고, 다수의 CCTV(400)들에서 수집된 영상 데이터를 검색하여 수신받거나, 다수의 CCTV(400)들에서 촬영되고 있는 실시간 영상을 수신받을 수 있다. 이하에서, 영상 모니터링 장치(100)의 동작은, 관리 단말(200)을 통해 전송되는 사용자 입력 신호에 기초하여 수행될 수 있다.
- [0034] 관리 단말(200)의 예를 들면, 통신 가능한 데스크탑 컴퓨터(desktop computer), 랩탑 컴퓨터(laptop computer), 노트북(notebook), 스마트폰(smart phone), 태블릿 PC(tablet PC), 모바일폰(mobile phone), 스마트 워치(smart watch), 스마트 글래스(smart glass), e-book 리더기, PMP(portable multimedia player), 휴대용 게임기, 네비게이션(navigation) 장치, 디지털 카메라(digital camera), DMB(digital multimedia broadcasting) 재생기, 디지털 음성 녹음기(digital audio recorder), 디지털 음성 재생기(digital audio player), 디지털 동영상 녹화기(digital video recorder), 디지털 동영상 재생기(digital video player), PDA(Personal Digital Assistant) 등일 수 있다.
- [0035] 영상 모니터링 장치(100)는, 관리 단말(200)의 입력에 기초하여, 적어도 하나의 로컬 서버(300)와 무선 네트워크를 통해 연결되고, 로컬 서버(300)에 저장된 영상 데이터 또는 실시간 영상을 수신받으며, 로컬 서버(300)에 CCTV(400)를 제어하기 위한 제어 신호를 전송할 수 있다.
- [0036] 예를 들어, 영상 모니터링 장치(100)는, 로컬 서버 별로 구분된 다수의 CCTV들의 목록(LST)을 포함하는 사용자 인터페이스(UI)를 생성하고, 생성된 사용자 인터페이스(UI)를 영상 모니터링 장치(100)에 접속된 관리 단말(200)에 표시할 수 있다. 또한, 영상 모니터링 장치(100)는 관리 단말(200)이 목록(LST)에 표시된 다수의 CCTV(400)들 중 하나를 드래그 앤 드롭(drag and drop)함에 따라, 사용자 인터페이스(UI) 내의 지정된 카메라 영역(CAREA)에 해당 CCTV에서 촬영되는 실시간 영상을 표시할 수 있다. 이때, 영상 모니터링 장치(100)는, 카메라 영역(CAREA)에 드롭된 실시간 영상들의 개수에 상응하도록 각 실시간 영상의 크기를 조정할 수 있다. 또한, 영상 모니터링 장치(100)는, 다수의 CCTV에서 수집된 영상 데이터를 분석함으로써, 목적 객체(target object)가 포함된 객체 영상, 화재, 재난, 범죄 등과 같이 특정 목적에 따른 이벤트 장면이 포함된 이벤트 영상, 2개 이상의 영상 프레임들을 합성하여 생성된 합성 영상을 카메라 영역(CAREA)에 표시할 수 있다.
- [0038] 도 3은 도 1에 따른 영상 모니터링 장치에 대한 기능부이다.
- [0039] 도 3을 참조하면, 영상 모니터링 장치(100)는, 통신부(101), 포맷 변환부(102), 영상 데이터 저장부(103), 제어 명령 생성부(104), 이벤트 영상 생성부(105), 합성 영상 생성부(106), 객체 영상 생성부(107), 다중 영상 표시부(108)를 포함할 수 있다.
- [0040] 통신부(101)는, 관리 단말(200) 및 로컬 서버(300)와 무선 네트워크를 통해 통신하는 통신 모듈(또는 칩(chip))로서, CDMA(code division multiple access) 기술, WCDMA(wideband CDMA) 기술, TDMA(time division multiple access) 기술, FDMA(frequency division multiple access) 기술, OFDM(orthogonal frequency division multiplexing) 기술, Filtered OFDM 기술, CP(cyclic prefix)-OFDM 기술, DFT-s-OFDM(discrete

Fourier transform-spread-OFDM) 기술, OFDMA(orthogonal frequency division multiple access) 기술, SC(single carrier)-FDMA 기술, NOMA(Non-orthogonal Multiple Access) 기술, GFDM(generalized frequency division multiplexing) 기술, FBMC(filter bank multi-carrier) 기술, UFMC(universal filtered multi-carrier) 기술, SDMA(Space Division Multiple Access) 기술, 5G(fifth generation) 기술 중 적어도 하나를 지원할 수 있다.

- [0041] 포맷 변환부(102)는, 제어 명령 생성부(103)에서 생성된 제어 신호를 로컬 서버(300)에서 지원하는 제어 신호의 형식에 따라 변환하거나, 통신부(101)를 통해 수집된 영상 데이터 또는 실시간 영상 데이터를 미리 지정된 영상 포맷(H.264, MPEG4 등)으로 변환할 수 있다. 즉 포맷 변환부(102)는, 다양한 형식을 갖는 CCTV(400) 제어 신호들 사이의 호환성을 보장하고, 영상 데이터의 영상 포맷 호환성을 보장할 수 있다.
- [0042] 영상 데이터 저장부(103)는, 통신부(101)를 통해 로컬 서버(300)로부터 수신되는 실시간 영상 또는 영상 데이터를 수집하여 저장할 수 있다. 영상 데이터 저장부(103)는 별도의 독립한 데이터베이스 서버로 구현될 수도 있으며, 분산 저장소, 또는 클라우드 저장소 등으로 구현될 수도 있다.
- [0043] 제어 명령 생성부(104)는, 관리 단말(200)로부터 입력되는 사용자 입력에 따라 로컬 서버(300)와 연결된 CCTV(400)들에 대한 제어 신호를 생성할 수 있다. 예를 들어, 제어 신호는, 로컬 서버(300)와 마찬가지로 CCTV(400)에 줌 인 제어, 조리개 값 조정, 포커스 조정, 촬영 방향 조정 등을 지시하는 제어 신호를 포함할 수 있다.
- [0044] 이벤트 영상 생성부(105)는, 영상 데이터 저장부(103)에 저장된 영상 데이터들을 분석하여 특정한 이벤트가 발생한 것으로 판단되는 영상 프레임들을 결합한 이벤트 영상을 생성할 수 있다. 예를 들어, 이벤트 영상 생성부(105)는, 관리 단말(200)로부터 관찰하고자 하는 관심 영역(IA)을 입력받고, 입력받은 관심 영역(IA)의 변화가 감지된 영상 프레임들을 결합하여 이벤트 영상을 생성할 수 있다.
- [0045] 합성 영상 생성부(106)는, CCTV(400)들 중에서 제어 신호에 따른 위치 변화가 감지된 CCTV(400)들을 대상으로, 시간적으로 연속된 영상 프레임들을 결합하여 합성 영상을 생성할 수 있다. 합성 영상은, 여러 방향에서 촬영된 영상 프레임들을 결합한 영상으로, 파노라마 영상으로 지칭될 수도 있다.
- [0046] 객체 영상 생성부(107)는, 영상 데이터 저장부(103)에 저장된 영상 데이터들에서 다수의 객체들 추출하고, 추출된 객체들 중 관심 객체를 관리 단말(200)의 입력에 따라 결정하고, 결정된 관심 객체가 포함된 영상 프레임들을 조합하여 객체 영상을 생성할 수 있다.
- [0047] 다중 영상 표시부(108)는, 영상 데이터 저장부(103)에 저장된 실시간 영상이나 이벤트 영상 생성부(105)에서 획득된 이벤트 영상, 합성 영상 생성부(106)에서 획득된 합성 영상, 객체 영상 생성부(107)에서 획득된 객체 영상 사이의 크기와 위치를 결정하고, 결정된 크기와 위치에 따라 격자 형태의 다중 영상을 생성하여 카메라 영역(CAREA)에 표시할 수 있다.
- [0049] 도 4는 도 3에 따른 이벤트 영상 생성부의 동작을 설명하기 위한 개념도이다.
- [0050] 도 4를 참조하면, 이벤트 영상 생성부(105)는 제1 CCTV(CCTV1)에서 촬영된 영상 데이터의 제1 프레임(FRM1)의 일부 영역에 해당하는 관심 영역(IA)을 입력받을 수 있다. 예를 들어, 관리 단말(200)은, 특정 출입문을 출입하는 사람들의 출입 이벤트를 모니터링하기 위하여, 특정 출입문 주변 영역을 관심 영역(IA, Area of interest)으로 입력할 수 있다.
- [0051] 이벤트 영상 생성부(105)는, 관심 영역(IA)을 입력받은 제1 프레임(FRM1)을 기준으로 소정의 시간 간격 내에 위치한 영상 프레임들을 추출할 수 있다. 예를 들어, 시간 간격은 관리 단말(200)에 의해 초단위, 분단위, 시간단위, 일단위 등으로 입력받을 수 있다.
- [0052] 다음으로, 이벤트 영상 생성부(105)는, 추출된 영상 프레임들에서 시간적으로 연속된 영상 프레임들에 포함된 관심 영역(IA)과 대응하는 영역들을 서로 비교하여, 관심 영역(IA)과 대응하는 영역들의 변화가 임계값 이상인 영상 프레임들을 추출할 수 있다. 이벤트 영상 생성부(105)는, 추출된 영상 프레임들을 시간 순서로 결합하여 이벤트 영상을 생성할 수 있다.
- [0053] 예를 들어, 관심 영역(IA)과 대응하는 제1 관심 영역 및 관심 영역(IA)과 대응하는 제2 관심 영역이 속하는 2개의 영상 프레임들이 서로 시간적으로 인접하다고 할 때, 제1 관심 영역을 구성하는 화소값들과 제2 지정 영역을 구성하는 화소값들을 서로 대응하는 위치끼리 매칭하고, 매칭된 2개의 화소값끼리 차분함으로써, 차분 화소값들을 산출하고, 산출된 차분 화소값들을 모두 더한 값이 임계값 이상이면, 제1 관심 영역과 제2 관심 영역이 속하

는 영상 프레임들은 변화가 있는 것으로 판단할 수 있다.

- [0054] 한편, 본 발명의 일 실시예에서는 하나의 CCTV에서 촬영된 영상 데이터를 이용하여 이벤트 영상을 생성하는 것 뿐만 아니라, 두개의 CCTV에서 촬영된 영상 데이터를 이용하여 이벤트 영상을 생성할 수도 있다.
- [0055] 예를 들어, 관심 영역(IA)에 출입문이 포함되어 있고, 출입문을 촬영하는 CCTV가 제1 CCTV(CCTV1) 뿐만 아니라, 제2 CCTV(CCTV2)와 제3 CCTV(CCTV3)도 있다고 할 경우, 제1 CCTV(CCTV1)에서 포착되지 않은 장면이 제2 CCTV(CCTV2)와 제3 CCTV(CCTV3)에서 포착될 수도 있다.
- [0056] 따라서, 이벤트 영상 생성부(105)는, 제2 CCTV(CCTV2)와 제3 CCTV(CCTV3)에서 촬영된 영상 데이터를 포함하여 분석함으로써 이벤트 영상을 생성할 수 있다.
- [0057] 구체적으로, 이벤트 영상 생성부(105)는, 제1 CCTV(CCTV1)에서 촬영된 영상 데이터의 제1 프레임(FRM1)의 일부 영역에 해당하는 관심 영역(IA)을 입력받을 수 있다. 다음으로, 이벤트 영상 생성부(105)는, 관심 영역(IA)에서 배경 영역을 추출할 수 있다. 예를 들어, 이벤트 영상 생성부(104)는, 관심 영역(IA)을 입력받은 제1 프레임(FRM1)을 기준으로 소정의 시간 간격 내에 위치한 영상 프레임들을 추출하고, 추출된 영상 프레임들에서 시간적으로 연속된 영상 프레임들에 포함된 관심 영역(IA)과 대응하는 영역들의 화소값들 사이의 평균을 산출함으로써, 산출된 평균값을 화소값으로 갖는 배경 영역을 추출할 수 있다. 이때, 소정의 시간 간격은 느리게 이동하는 객체가 배경 영역으로 추출되는 것을 방지하기 위하여 충분히 긴 시간 간격(예를 들면 시간 단위 또는 일 단위)으로 설정될 수 있다.
- [0058] 배경 영역이 추출되면, 이벤트 영상 생성부(105)는, 배경 영역에서 객체 추출 알고리즘을 이용하여 적어도 하나의 고정 객체(FIXEDOBJ)를 추출할 수 있다. 예를 들어, 객체 추출 알고리즘은, 각종 컴퓨터 비전 기술의 딥러닝 네트워크(R-CNN, YOLO 등)를 이용한 알고리즘들이 이용될 수 있는데, 예를 들면 SSD(W. Liu, D. Anguelov, D. Erhan, C. Szegedy, and S. E. Reed. SSD: single shot multibox detector. CoRR, abs/1512.02325,2015.)가 이용될 수 있다.
- [0059] 여기서 추출되는 적어도 하나의 고정 객체(FIXEDOBJ)는 배경 영역에 해당하는 객체에 해당하므로, 움직임이 없는 객체일 수 있다.
- [0060] 이러한 고정 객체(FIXEDOBJ)가 추출되면, 제1 CCTV(CCTV1)를 제외한 나머지 CCTV(400)들에서 획득된 각각의 영상 데이터들마다 하나의 배경 영상 프레임을 생성할 수 있다. 예를 들어, 제2 CCTV(CCTV2)에서 획득된 영상 데이터에 따른 영상 프레임들 중에서 일정한 시간 간격 내에 위치한 영상 프레임들을 추출하고, 추출된 영상 프레임들에 대한 화소값의 평균을 산출함으로써 평균 화소값을 갖는 배경 영상 프레임을 생성할 수 있다.
- [0061] 다음으로, 이벤트 영상 생성부(105)는, 각 CCTV와 대응하는 배경 영상 프레임들에서 객체 추출 알고리즘을 이용하여 앞서 추출된 고정 객체(FIXEDOBJ)와 동일성이 있는 객체를 검출하고, 동일성이 있는 객체가 검출된 배경 영상 프레임들과 대응하는 적어도 하나의 CCTV를 결정할 수 있다. 예를 들어, 제2 CCTV(CCTV2)와 제3 CCTV(CCTV3)의 영상 데이터들로부터 획득된 배경 영상 프레임들에서 고정 객체(FIXEDOBJ)와 동일성이 있는 객체가 검출될 수 있다.
- [0062] 이벤트 영상 생성부(105)는, 결정된 적어도 하나의 CCTV에서 획득된 영상 데이터를 대상으로 고정 객체(FIXEDOBJ)와 동일성이 있는 객체의 위치와 대응하는 영역을 보조 관심 영역(SIA)으로 결정할 수 있다. 다음으로, 보조 관심 영역을 앞서 관리 단말(200)에서 입력받은 관심 영역인 것으로 취급하여 관심 영역으로부터 이벤트 영상을 생성한 것과 마찬가지로 추가 이벤트 영상을 생성할 수 있다.
- [0063] 예를 들어, 제2 CCTV(CCTV2)와 제3 CCTV(CCTV3)의 영상 데이터들을 대상으로 보조 관심 영역들을 결정하고, 결정된 보조 관심 영역들을 이용하여 제2 CCTV(CCTV2)에 상응하는 제2 추가 이벤트 영상과 제3 CCTV(CCTV3)에 상응하는 제3 추가 이벤트 영상을 생성할 수 있다.
- [0064] 이처럼, 일 실시예에 따른 영상 모니터링 장치(100)는, 2개 이상의 CCTV들에서 촬영된 영상 데이터들을 대상으로 고정 객체에 기반한 이벤트 영상을 생성할 수 있기 때문에, 관리자가 일일이 모든 CCTV들을 살펴보고 관심 영역을 입력할 필요가 없어 효율적으로 모니터링이 가능한 장점이 있다.
- [0066] 도 5는 도 3에 따른 객체 영상 생성부의 동작을 설명하기 위한 개념도이다.
- [0067] 도 5를 참조하면, 객체 영상 생성부(107)는, CCTV(400)들 중 제1 CCTV(CCTV1)에서 획득된 영상 데이터의 영상 프레임들을 대상으로 적어도 하나의 이동 객체(MOV OBJ)를 추출하고, 추출된 이동 객체(MOV OBJ) 중 선정된 관심

객체가 검출된 영상 프레임들을 시간순서로 결합함으로써, 객체 영상을 생성할 수 있다.

[0068] 예를 들어, 객체 영상 생성부(107)는, 제1 CCTV(CCTV1)에서 획득된 영상 데이터의 영상 프레임들을 이용하여 배경 영상 프레임을 생성할 수 있다. 앞선 도 4에서 설명한 것과 마찬가지로, 제1 CCTV(CCTV1)에서 획득된 영상 데이터의 일정한 시간 간격 내에 위치한 영상 프레임들의 화소값의 평균을 산출하고, 산출된 평균을 화소값으로 갖는 배경 영상 프레임(bgf)을 생성할 수 있다.

[0069] 다음으로, 객체 영상 생성부(107)는, 배경 영상 프레임(bgf)을 영상 프레임들 각각에 대하여 차분하여, 차분 영상 프레임들을 생성할 수 있다. 여기서 생성된 차분 영상 프레임들 각각에 포함된 화소값들의 총합이 임계값 이상인 차분 영상 프레임들은 배경 이외의 객체(즉 이동 객체)가 포함된 영상 프레임에 해당할 수 있다. 따라서, 객체 영상 생성부(107)는, 생성된 차분 영상 프레임들 각각에 포함된 화소값들의 총합이 임계값 이상인 차분 영상 프레임들을 선정하고, 선정된 차분 영상 프레임들에서 적어도 하나의 이동 객체(MOV OBJ)를 결정할 수 있다.

[0070] 객체 영상 생성부(107)는, 이동 객체(MOV OBJ)가 2개 이상 결정된 경우, 관리 단말(200)로부터 2개 이상의 이동 객체(MOV OBJ)들 중 하나를 선택하는 입력을 수신함으로써, 객체 영상을 생성할 관심 객체(이하 이동 객체와 혼용하여 지칭될 수 있음)를 결정할 수 있다. 다만, 이것은 선택적인 것으로, 반드시 이동 객체들 중 하나를 입력 받을 필요는 없으며, 모든 이동 객체들을 관심 객체들로 선정하고, 관심 객체들이 포함된 객체 영상을 생성할 수도 있다.

[0071] 객체 영상 생성부(107)는, 이동 객체(MOV OBJ)들 중 선정된 관심 객체를 구성하는 영역들을 4개 이상의 화소값으로 구성된 복수의 객체셀(ocell)들로 구획하고, 구획된 객체셀들 각각에 대하여 색상, 채도, 및 명도의 평균값을 산출함으로써, 각 객체셀(ocell)과 대응하는 색상셀(Hcell), 채도셀(Scell), 명도셀(Icell)을 결정할 수 있다.

[0072] 예를 들어, 객체셀(ocell)에 포함된 4개 이상의 화소값들과 대응하는 색상값들을 산출하고, 산출된 색상값들의 평균값을 산출함으로써, 객체셀(ocell)과 대응하는 색상셀(Hcell)을 결정할 수 있다. 이때, 화소값에 대응하는 색상값(H)은 다음의 수학식 1을 이용하여 산출될 수 있다.

수학식 1

$$H = \cos^{-1} \frac{\frac{1}{2}[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B) \cdot (G - B)}}$$

[0073]

[0074] 수학식 1을 참조하면, R, G, B는 화소값이 RGB 포맷일 때, 각각 붉은색 화소값(R), 그린 화소값(G), 블루 화소값(B)일 수 있다.

[0075] 또한, 객체셀(ocell)에 포함된 2개 이상의 화소값들과 대응하는 채도값들을 산출하고, 산출된 채도값들의 평균값을 산출함으로써, 객체셀(ocell)과 대응하는 채도셀(Hcell)을 결정할 수 있다. 이때, 화소값에 대응하는 채도값(S)은 다음의 수학식 2를 이용하여 산출될 수 있다.

수학식 2

$$S = 1 - \frac{3}{(R + G + B)} \cdot \min(R, G, B)$$

[0076]

[0077] 또한, 객체셀(ocell)에 포함된 2개 이상의 화소값들과 대응하는 명도값들을 산출하고, 산출된 명도값들의 평균값을 산출함으로써, 객체셀(ocell)과 대응하는 명도셀(Hcell)을 결정할 수 있다. 이때, 화소값에 대응하는 명도값(S)은 다음의 수학식 3을 이용하여 산출될 수 있다.

수학식 3

$$I = \frac{1}{3} \cdot (R + G + B)$$

[0078]

[0079]

[0080]

[0081]

[0082]

[0083]

[0085]

[0086]

[0087]

[0088]

[0089]

[0090]

각 객체셀(ocell)과 대응하는 색상셀(Hcell), 채도셀(Scell), 명도셀(Icell)을 결정하면, 객체 영상 생성부(107)는, 색상셀(Hcell)들을 제1 특징값, 채도셀(Scell)들을 제2 특징값, 명도셀(Icell)들을 제3 특징값으로 설정하고, 설정된 제1 내지 제3 특징값을 객체 추출 알고리즘의 비교 기준으로 사용하여 관심 객체와 유사한 객체를 제1 CCTV(CCTV)와 로컬 서버(300)를 공유하는 모든 CCTV들의 영상 데이터에 포함된 영상 프레임들에서 검색할 수 있다.

예를 들어, 모든 CCTV들의 영상 프레임들 각각에서 이동 객체(MOV OBJ)들을 검출하고, 검출된 이동 객체들 중에서 제1 특징값을 기준으로 할 때, 관심 객체와 가장 유사한 제1 후보 객체, 제2 특징값을 기준으로 할 때, 관심 객체와 가장 유사한 제2 후보 객체, 제3 특징값을 기준으로 할 때, 관심 객체와 가장 유사한 제3 후보 객체를 선정할 수 있다.

다음으로, 화소값들을 기반으로, 제1 후보 객체 내지 제3 후보 객체를 관심 객체와 비교하여 가장 동일성이 높은 동일성이 높은 후보 객체를 선정함으로써, 제1 후보 객체 내지 제3 후보 객체 중 관심 객체와 동일한 객체를 선정할 수 있다. 여기서, 화소값 기반으로 한 객체 비교는, 앞선 화소값들 사이의 차분값을 산출하는 것을 비롯하여 다양한 형태의 이미지 비교 방법이 적용될 수 있으므로, 자세한 설명은 생략한다.

관심 객체와 동일한 객체가 선정되면, 객체 영상 생성부(107)는, 관심 객체와 동일한 객체가 포함된 영상 프레임들을 시간 순서대로 연결함으로써, 객체 영상을 각 CCTV마다(또는 동일성이 있는 객체가 검출된 CCTV마다) 생성할 수 있다.

상술한 것처럼, 이동 객체(MOV OBJ)를 생성하고, 생성된 이동 객체를 4개 이상의 화소값을 갖는 객체셀(ocell)로 구획한 후, 색상셀(Hcell), 채도셀(Scell), 명도셀(Icell)을 특징값으로 비교할 경우, 비교되는 대상이 크게 감소하므로 모든 영상 프레임들을 검색할 때 발생하는 과도한 시스템 오버헤드(overhead)를 줄일 수 있고, 제1 후보 객체 내지 제3 후보 객체 중 가장 동일한 객체를 최종적으로 선정하기 때문에, 관심 객체와 유사하지만 동일하지 않은 객체가 선정될 가능성이 크게 낮아질 수 있다.

도 6은 도 3에 따른 합성 영상 생성부의 동작을 설명하기 위한 개념도이다.

도 6을 참조하면, 합성 영상 생성부(106)는, 복수의 CCTV(400)들 중에서 촬영 방향 조정에 따른 제어 신호를 수신하는 CCTV(400)들을 선정하고, 선정된 CCTV(400)들로부터 획득된 합성 타겟 영상 데이터들을 대상으로 합성 영상을 생성할 수 있다.

구체적으로, 합성 영상 생성부(106)는, 합성 타겟 영상 데이터들에서 시간적으로 인접한 영상 프레임들을 서로 비교하여, 서로 중첩되는 영역을 결정하고, 중첩되는 영역이 전체 프레임 영역의 크기보다 임계값보다 작은지 결정할 수 있다.

여기서, 중첩되는 영역이 임계값보다 큰 경우, 합성 영상 생성부(106)는, 비교된 영상 프레임들을 합성하지 않을 수 있다. 예를 들어, 임계값보다 크다면, 실질적으로 동일한 장면이 계속 촬영되는 것으로 판단할 수 있으므로, 합성 영상을 생성하는 것이 적절하지 않다.

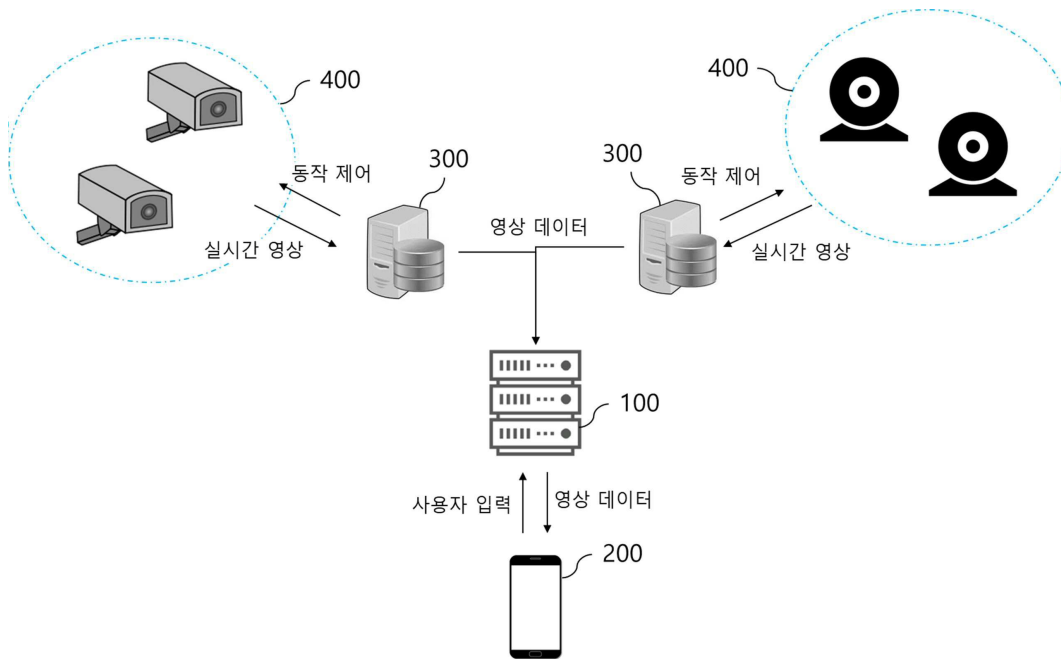
여기서, 중첩되는 영역이 임계값보다 작은 경우, 합성 영상 생성부(106)는, 중첩되는 영역이 영상 프레임의 엣지(edge) 영역을 포함하는지 판단할 수 있다. 이때, 중첩되는 영역이 영상 프레임의 엣지(edge) 영역(상하좌우 엣지들 중 하나)을 포함하는 경우, 서로 인접한 영상 프레임들 중에서 시간적으로 뒤에 있는 영상 프레임에서 중첩 영역을 제외한 나머지 영역을 시간적으로 앞선 영상 프레임(또는 기 생성된 중간 합성 영상 프레임)의 엣지 영역에 연결하여, 중간 합성 영상 프레임을 생성할 수 있다.

시간순서에 따라 중간 합성 영상 프레임에 시간적으로 뒤에 있는 영상 프레임을 연결시키는 과정은 계속 반복될 수 있다. 이때, 합성 영상 생성부(106)는, 중첩되는 영역이 영상 프레임의 엣지 영역을 포함하지 않으면, 영상 합성을 종료하고 그동안 생성된 중간 합성 영상 프레임을 합성 영상으로서 생성할 수 있다.

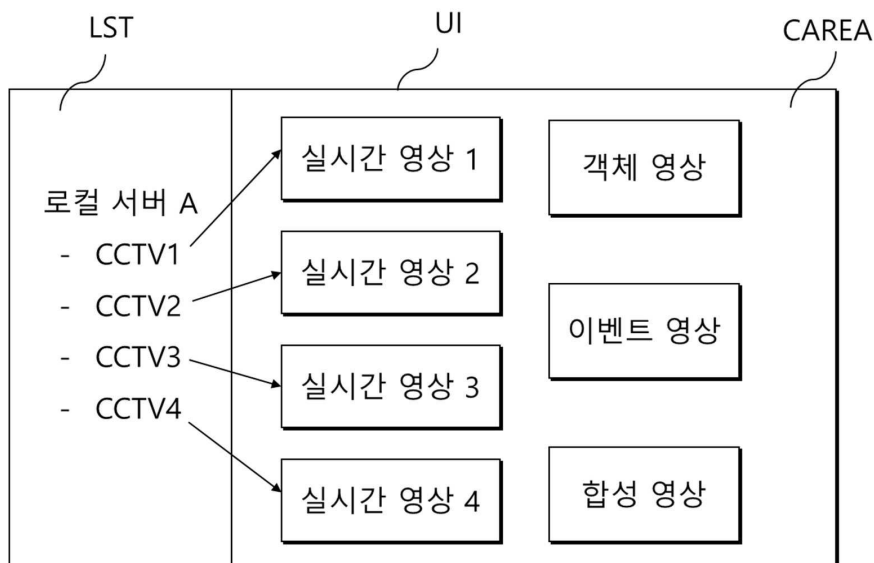
- 108: 다중 영상 표시부
- 110: 프로세서
- 120: 메모리
- 130: 송수신 장치
- 140: 입력 인터페이스 장치
- 150: 출력 인터페이스 장치
- 160: 저장 장치
- 170: 버스
- 200: 관리 단말
- 300: 로컬 서버
- 400: CCTV

도면

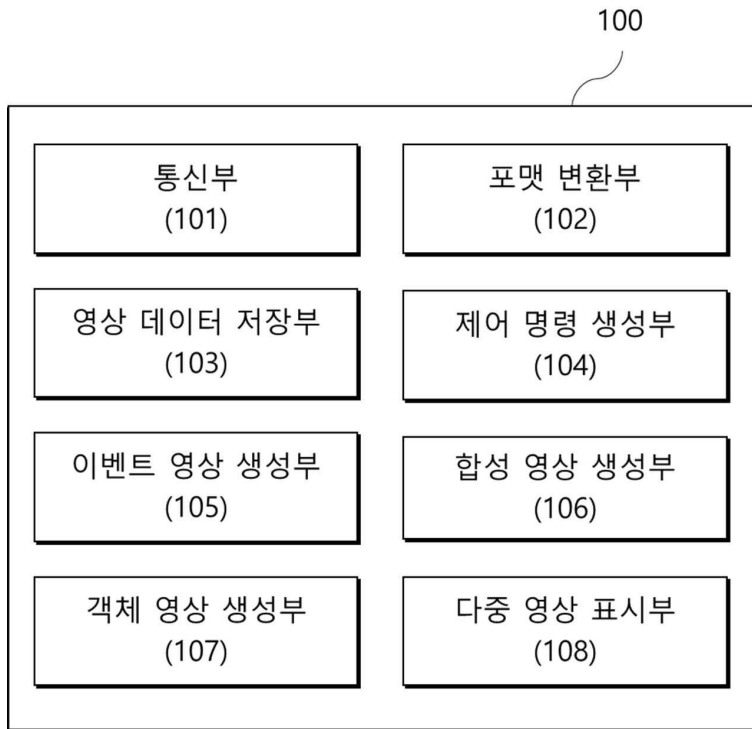
도면1



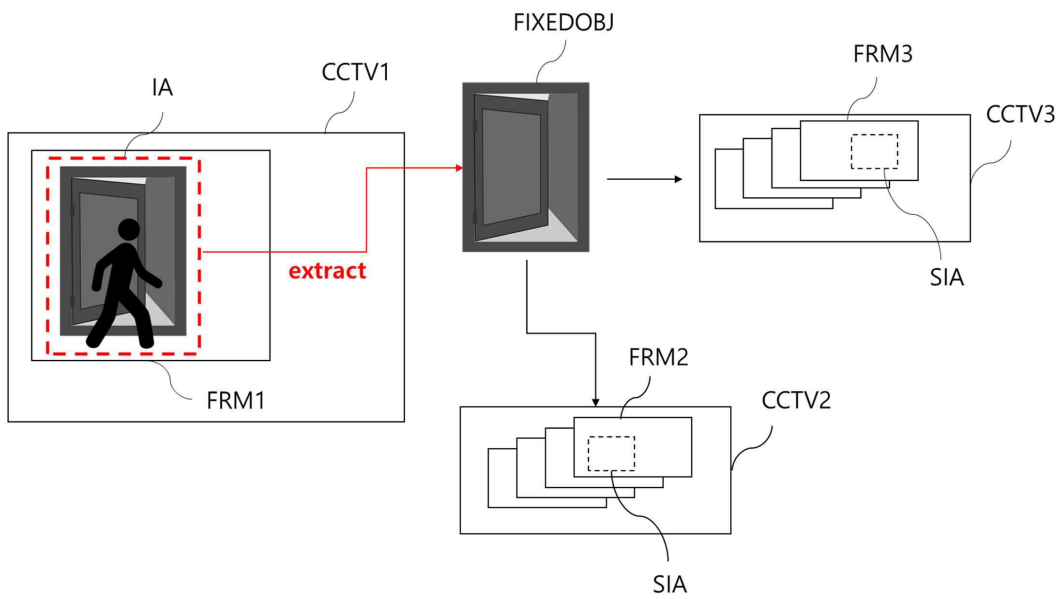
도면2



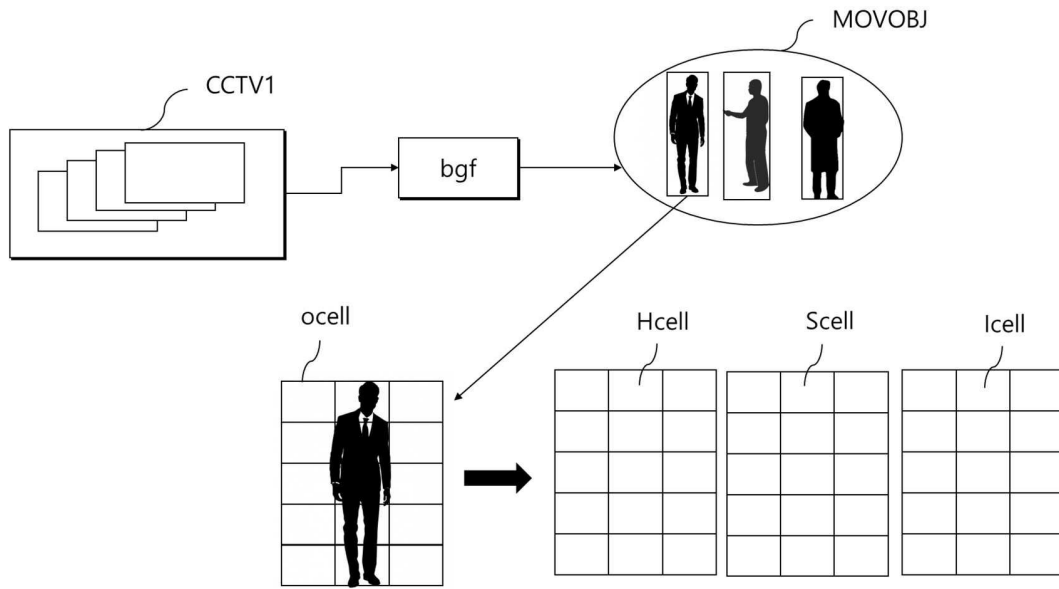
도면3



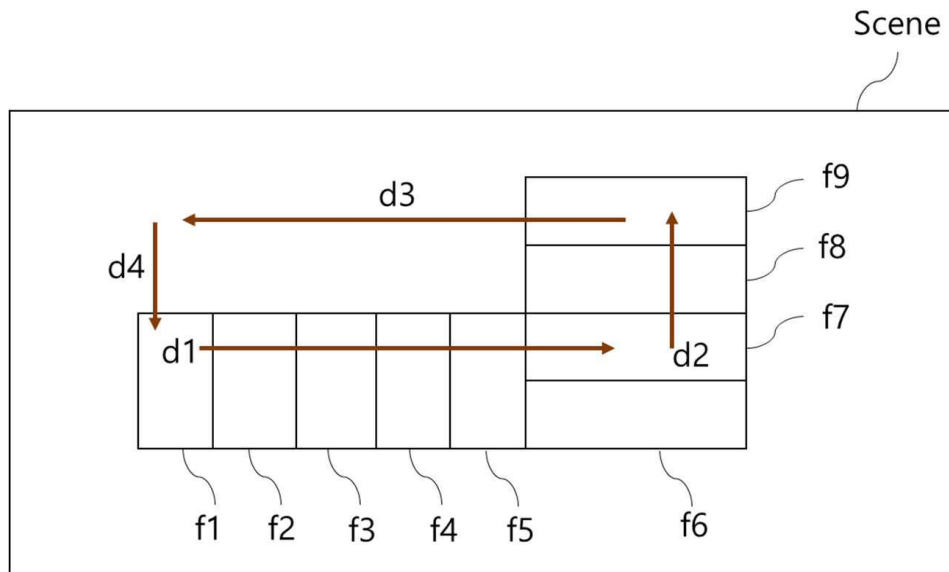
도면4



도면5



도면6



도면7

