



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2019년08월29일
 (11) 등록번호 10-2016262
 (24) 등록일자 2019년08월23일

- | | |
|--|--|
| <p>(51) 국제특허분류(Int. Cl.)
 <i>G06T 7/20</i> (2017.01)</p> <p>(52) CPC특허분류
 <i>G06T 7/20</i> (2013.01)
 <i>G06T 7/11</i> (2017.01)</p> <p>(21) 출원번호 10-2018-0062174</p> <p>(22) 출원일자 2018년05월30일
 심사청구일자 2018년05월30일</p> <p>(56) 선행기술조사문헌
 KR101366198 B1*
 Fei et al. Memorable and rich video summarization. Journal of Visual Communication and Image Representation, 2016년 12d월, pp. 207-217. 1부.*
 KR101270718 B1
 *는 심사관에 의하여 인용된 문헌</p> | <p>(73) 특허권자
 세종대학교산학협력단
 서울특별시 광진구 능동로 209 (군자동, 세종대학교)</p> <p>(72) 발명자
 백성욱
 서울특별시 광진구 아차산로 262, B동 1304호(자양동, 더샵스타시티)</p> <p>이미영
 서울특별시 강남구 언주로85길 13, 102호(역삼동, 강남아파트)
 (뒷면에 계속)</p> <p>(74) 대리인
 송인호, 윤형근, 최영중, 최관락</p> |
|--|--|

전체 청구항 수 : 총 16 항

심사관 : 김창원

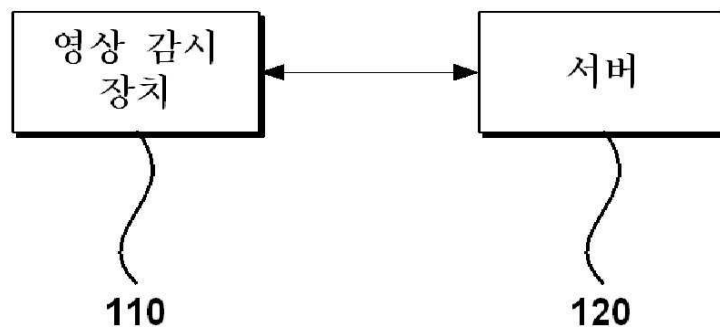
(54) 발명의 명칭 **연기를 포함하는 환경에 대한 영상 분석 방법 및 그 시스템**

(57) 요약

연기를 포함하는 환경에 대한 영상 분석 방법 및 그 시스템이 개시된다. (a) 특정 공간을 촬영한 비디오 스트림을 분석하여 연기를 포함하는지 여부에 따라 연기(smoky) 프레임들과 비연기(non-smoky) 프레임들로 분류하는 단계; (b) 상기 비연기 프레임들을 이용하여 각 물체의 움직임 이벤트 발생 여부에 따라 각 물체에 대한 샷(shot)으로 각각 분할하는 단계; (c) 상기 각 샷에 상응하는 배경 이미지를 저장하는 단계; 및 (d) 상기 각 물체를 훈련된 인공 신경망을 통해 분류하는 단계를 포함한다.

대표도 - 도1

100



(52) CPC특허분류

G06T 7/194 (2017.01)
G06T 2207/10016 (2013.01)
G06T 2207/20084 (2013.01)
G06T 2210/24 (2013.01)

(72) 발명자

칸 무하마드

서울특별시 광진구 능동로 209 세종대학교 율곡관
601B

올라 아민

서울특별시 광진구 능동로 209 세종대학교 율곡관
601B

이자즈 울 하크

서울특별시 광진구 능동로 209 세종대학교 율곡관
601B

박준렬

서울특별시 서초구 방배중앙로 112, A동 203호(방
배동, 방배파크빌2)

이 발명을 지원한 국가연구개발사업

과제고유번호	2016R1A2B4011712
부처명	과학기술정보통신부
연구관리전문기관	한국연구재단
연구사업명	개인기초연구(미래부)
연구과제명	지능형 영상 감시 시스템을 위한 다중 영상 비디오 데이터 처리 및 분석 기술
기 여 율	1/1
주관기관	세종대학교
연구기간	2017.06.01 ~ 2018.03.31

명세서

청구범위

청구항 1

- (a) 특정 공간을 촬영한 비디오 스트림을 분석하여 연기를 포함하는지 여부에 따라 연기(smoky) 프레임들과 비연기(non-smoky) 프레임들로 분류하는 단계;
- (b) 상기 비연기 프레임들을 이용하여 각 물체의 움직임 이벤트 발생 여부에 따라 각 물체에 대한 샷(shot)으로 각각 분할하는 단계;
- (c) 상기 각 샷에 상응하는 배경 이미지를 저장하는 단계;
- (d) 상기 각 물체를 훈련된 인공 신경망을 통해 분류하는 단계;
- (e) 상기 연기 프레임들을 분석하여 중요 키 프레임을 선별하는 단계; 및
- (f) 상기 중요 키 프레임과 상기 저장된 배경 이미지와 신호대 잡음비(PSNR)을 각각 도출한 후 신호대 잡음비가 가장 큰 배경 이미지를 상기 중요 키 프레임에 대한 배경 이미지로 선택하는 단계를 포함하는 영상 분석 방법.

청구항 2

제1 항에 있어서,

상기 (c) 단계는,

상기 각 샷에 포함된 프레임들을 이용하여 물체 움직임이 발생한 이후 가장 깨끗한 프레임을 배경 이미지로 선택하여 저장하는 것을 특징으로 하는 영상 분석 방법.

청구항 3

제2 항에 있어서,

상기 각 샷에 포함된 각 프레임에 대한 기억 가능성 점수 및 엔트로피를 도출한 후 상기 도출된 기억 가능성 점수 및 엔트로피가 임계치 이상인 프레임을 배경 이미지로 선택하는 것을 특징으로 하는 영상 분석 방법.

청구항 4

제3 항에 있어서,

상기 기억 가능성 점수는 CNN 모델을 기반으로 계산되며,

상기 엔트로피는 각 프레임을 HSV 색상 모델로 변환한 뒤 색상, 채도, 명도 컴포넌트 각각에 대한 히스토그램 빈의 확률을 계산하여 도출되는 것을 특징으로 하는 영상 분석 방법.

청구항 5

삭제

청구항 6

제1 항에 있어서,

상기 중요 키 프레임을 선별하는 단계는,

상기 연기 프레임들 중 연속된 두개의 프레임을 HSV 색상 모델로 변환하는 단계; 및

상기 HSV 색상 모델로 변환된 두개의 프레임간의 색조(Hue) 컴포넌트의 예지 차이가 임계치 이상인 프레임을 중요 키 프레임으로 선별하는 단계를 포함하는 것을 특징으로 하는 영상 분석 방법.

청구항 7

제1 항에 있어서,

상기 연기 프레임들을 학습된 인공 신경망에 적용하여 도출된 특징맵을 이용하여 관심 영역을 추출하는 단계; 및

상기 추출된 관심 영역을 학습된 인공 신경망에 적용하여 컨텍스트 정보를 추출하는 단계를 더 포함하는 영상 분석 방법.

청구항 8

제1 항에 따른 방법을 수행하기 위한 프로그램 코드를 기록한 컴퓨터로 판독 가능한 기록매체 제품.

청구항 9

각 지역에 설치된 영상 감시 장치에 있어서,

특정 공간을 촬영한 비디오 스트림을 분석하여 연기를 포함하는지 여부에 따라 연기(smoky) 프레임들과 비연기(non-smoky) 프레임들로 분류하는 프레임 분류부;

상기 비연기 프레임들을 이용하여 각 물체의 움직임 이벤트 발생 여부에 따라 각 물체에 대한 샷(shot)으로 각 분할한 후 상기 각 샷에 상응하는 배경 이미지를 저장하는 배경 저장부; 및

상기 각 물체를 훈련된 인공 신경망을 통해 분류하는 물체 분류부를 포함하되,

상기 연기 프레임들은 분석되어 중요 키 프레임으로 선별되되,

상기 중요 키 프레임과 상기 저장된 배경 이미지와 신호대 잡음비(PSNR)이 도출된 후 신호대 잡음비가 가장 큰 배경 이미지가 상기 중요 키 프레임에 대한 배경 이미지로 선택되는 것을 특징으로 하는 영상 감시 장치.

청구항 10

제9 항에 있어서,

상기 배경 저장부는,

상기 각 샷에 포함된 프레임들을 이용하여 물체 움직임이 발생한 이후 가장 깨끗한 프레임을 배경 이미지로 선택하여 저장하는 것을 특징으로 하는 영상 감시 장치.

청구항 11

제9 항에 있어서,

상기 배경 저장부는,

상기 각 샷에 포함된 각 프레임에 대한 기억 가능성 점수 및 엔트로피를 도출한 후 상기 도출된 기억 가능성 점수 및 엔트로피가 임계치 이상인 프레임을 배경 이미지로 선택하는 것을 특징으로 하는 영상 감시 장치.

청구항 12

영상 감시 장치와 연결된 서버에 있어서,

통신부;

적어도 하나의 명령어들을 저장하는 메모리; 및

상기 메모리에 저장된 명령어들을 수행하는 프로세서를 포함하되,

상기 프로세서에 의해 수행된 명령어들은,

상기 영상 감시 장치로부터 수신된 연기 프레임들을 분석하여 중요 키 프레임을 선별하는 단계; 및

상기 중요 키 프레임과 기저장된 배경 이미지와 신호대 잡음비(PSNR)을 각각 도출한 후 신호대 잡음비가 가장 큰 배경 이미지를 상기 중요 키 프레임에 대한 배경 이미지로 선택하는 단계를 수행하는 것을 특징으로 하는 서버.

청구항 13

제12 항에 있어서,

상기 중요 키 프레임을 선별하는 단계는,

상기 연기 프레임들 중 연속된 두개의 프레임을 HSV 색상 모델로 변환하는 단계; 및

상기 HSV 색상 모델로 변환된 두개의 프레임간의 색조(Hue) 컴포넌트의 에지 차이가 임계치 이상인 프레임을 중요 키 프레임으로 선별하는 단계를 포함하는 것을 특징으로 하는 서버.

청구항 14

제12 항에 있어서,

상기 연기 프레임들을 학습된 인공 신경망에 적용하여 도출된 특징맵을 이용하여 관심 영역을 추출하는 단계; 및

상기 배경 이미지에 상응하여 상기 영상 감시 장치에서 분류된 물체 단서를 이용하여 상기 추출된 관심 영역에 대한 문맥 정보를 추출하는 단계를 더 포함하는 서버.

청구항 15

특정 공간을 촬영한 비디오 스트림을 분석하여 연기를 포함하는지 여부에 따라 연기(smoky) 프레임들과 비연기(non-smoky) 프레임들로 분류하며, 상기 비연기 프레임들을 이용하여 각 물체의 움직임 이벤트 발생 여부에 따라 각 물체에 대한 샷(shot)으로 각각 분할한 후 상기 각 샷에 상응하는 배경 이미지를 저장하는 영상 감시 장치; 및

상기 영상 감시 장치로부터 수신된 연기 프레임들을 분석하여 중요 키 프레임을 선별하고, 상기 중요 키 프레임과 상기 저장된 배경 이미지와 신호대 잡음비(PSNR)을 각각 도출한 후 신호대 잡음비가 가장 큰 배경 이미지를 상기 중요 키 프레임에 대한 배경 이미지로 선택하여 인공 신경망을 통해 분석하는 서버를 포함하는 영상 처리 시스템.

청구항 16

제15 항에 있어서,

상기 영상 감시 장치는

상기 배경 이미지를 인공 신경망에 적용하여 각 물체를 분류하는 것을 특징으로 하는 영상 처리 시스템.

청구항 17

제15 항에 있어서,

상기 서버는,

상기 연기 프레임들을 학습된 인공 신경망에 적용하여 도출된 특징맵을 이용하여 관심 영역을 추출하며, 상기 배경 이미지에 상응하여 상기 영상 감시 장치에서 분류된 물체 단서를 이용하여 상기 추출된 관심 영역에 대한 문맥 정보를 추출하는 것을 특징으로 하는 영상 처리 시스템.

발명의 설명

기술 분야

[0001] 본 발명은 연기를 포함하는 환경에 대한 영상 분석 방법 및 그 시스템에 관한 것이다.

배경 기술

[0003] 최근에는 유선 또는 무선 네트워크의 센서를 활용하여 광원 없이 온도 변화와 유해 가스 흡수를 감지함으로써 연기 및 화재 감지에 다양한 기술이 사용되고 있으며, 배터리 및 통신 모듈에 내장된 경량 센서를 기반으로 원거리에서 보고할 수 있도록 기술이 발전하고 있다. 그러나 이러한 장치는 정기적인 유지 보수가 필요하며, 특정 영역에 설치된 센서에서 다른 정보를 추출 할 수 없는 문제점이 있다.

[0004] 연기가 자욱한 환경에서는 움직이는 물체와 건물, 지하철 또는 손을 흔들며 도움을 요청하는 사람 등의 활동을 인식하는 것이 중요하다. 피해를 입은 연기 지역의 다양한 활동과 함께 상황 별 정보를 추출하여 즉시 재난 관리 부서에 전달하여 귀중한 생명과 재산을 구할 수 있는 조기 조치를 취할 수 있으며, 최근 컴퓨터 비전 기술을 이용하여 연기를 감지하기 위한 다양한 방법들이 연구되고 있으나, 연기가 자욱한 환경에서 움직이는 물체를 식별하는 것은 매우 어려운 문제이다.

발명의 내용

해결하려는 과제

[0006] 본 발명은 연기를 포함하는 환경에 대한 영상 분석 방법 및 그 시스템을 제공하기 위한 것이다.

[0007] 또한, 본 발명은 연기가 자욱한 환경에서 구조 과정에서 도움이 될 수 있는 장면의 문맥을 분석할 수 있는 연기를 포함하는 환경에 대한 영상 분석 방법 및 그 시스템을 제공하기 위한 것이다.

과제의 해결 수단

[0009] 본 발명의 일 측면에 따르면, 연기가 자욱한 환경에서 구조 과정에서 도움이 될 수 있는 장면의 문맥을 분석할 수 있는 연기를 포함하는 환경에 대한 영상 분석 방법이 제공된다.

[0010] 본 발명의 일 실시예에 따르면, (a) 특정 공간을 촬영한 비디오 스트림을 분석하여 연기를 포함하는지 여부에 따라 연기(smoky) 프레임들과 비연기(non-smoky) 프레임들로 분류하는 단계; (b) 상기 비연기 프레임들을 이용하여 각 물체의 움직임 이벤트 발생 여부에 따라 각 물체에 대한 샷(shot)으로 각각 분할하는 단계; (c) 상기 각 샷에 상응하는 배경 이미지를 저장하는 단계; 및 (d) 상기 각 물체를 훈련된 인공 신경망을 통해 분류하는 단계를 포함하는 영상 분석 방법이 제공될 수 있다.

[0011] 상기 (c) 단계는, 상기 각 샷에 포함된 프레임들을 이용하여 물체 움직임이 발생한 이후 가장 깨끗한 프레임을 배경 이미지로 선택하여 저장할 수 있다.

[0012] 상기 각 샷에 포함된 각 프레임에 대한 기억 가능성 점수 및 엔트로피를 도출한 후 상기 도출된 기억 가능성 점수 및 엔트로피가 임계치 이상인 프레임을 배경으로 선택할 수 있다.

- [0013] 상기 기억 가능성 점수는 CNN 모델을 기반으로 계산되며, 상기 엔트로피는 각 프레임을 HSV 색상 모델로 변환한 뒤 색상, 채도, 명도 컴포넌트 각각에 대한 히스토그램 빈의 확률을 계산하여 도출될 수 있다.
- [0014] 상기 연기 프레임들을 분석하여 중요 키 프레임을 선별하는 단계; 및 상기 중요 키 프레임과 상기 저장된 배경 이미지와 신호대 잡음비(PSNR)을 각각 도출한 후 신호대 잡음비가 가장 큰 배경 이미지를 상기 중요 키 프레임에 대한 배경 이미지로 선택하는 단계를 더 포함할 수 있다.
- [0015] 상기 중요 키 프레임을 선별하는 단계는, 상기 연기 프레임들 중 연속된 두개의 프레임을 HSV 색상 모델로 변환하는 단계; 및 상기 HSV 색상 모델로 변환된 두개의 프레임간의 색조(Hue) 컴포넌트의 에지 차이가 임계치 이상인 프레임을 중요 키 프레임으로 선별하는 단계를 포함할 수 있다.
- [0016] 상기 연기 프레임들을 학습된 인공 신경망에 적용하여 도출된 특징맵을 이용하여 관심 영역을 추출하는 단계; 및 상기 추출된 관심 영역을 학습된 인공 신경망에 적용하여 컨텍스트 정보를 추출하는 단계를 더 포함할 수 있다.
- [0018] 본 발명의 다른 측면에 따르면, 연기가 자욱한 환경에서 구조 과정에서 도움이 될 수 있는 장면의 문맥을 분석할 수 있는 연기를 포함하는 환경에 대한 영상 분석 시스템이 제공된다.
- [0019] 본 발명의 일 실시예에 따르면, 각 지역에 설치된 영상 감시 장치에 있어서, 특정 공간을 촬영한 비디오 스트림을 분석하여 연기를 포함하는지 여부에 따라 연기(smoky) 프레임들과 비연기(non-smoky) 프레임들로 분류하는 프레임 분류부; 상기 비연기 프레임들을 이용하여 각 물체의 움직임 이벤트 발생 여부에 따라 각 물체에 대한 샷(shot)으로 각각 분할한 후 상기 각 샷에 상응하는 배경 이미지를 저장하는 배경 저장부; 및 상기 각 물체를 훈련된 인공 신경망을 통해 분류하는 물체 분류부를 포함하는 영상 감시 장치가 제공될 수 있다.
- [0020] 상기 배경 저장부는, 상기 각 샷에 포함된 프레임들을 이용하여 물체 움직임이 발생한 이후 가장 깨끗한 프레임을 배경 이미지로 선택하여 저장할 수 있다.
- [0021] 상기 배경 저장부는, 상기 각 샷에 포함된 각 프레임에 대한 기억 가능성 점수 및 엔트로피를 도출한 후 상기 도출된 기억 가능성 점수 및 엔트로피가 임계치 이상인 프레임을 배경으로 선택할 수 있다.
- [0023] 본 발명의 다른 실시예에 따르면, 영상 감시 장치와 연결된 서버에 있어서, 통신부; 적어도 하나의 명령어들을 저장하는 메모리; 및 상기 메모리에 저장된 명령어들을 수행하는 프로세서를 포함하되, 상기 프로세서에 의해 수행된 명령어들은, 상기 영상 감시 장치로부터 수신된 연기 프레임들을 분석하여 중요 키 프레임을 선별하는 단계; 및 상기 중요 키 프레임과 상기 저장된 배경 이미지와 신호대 잡음비(PSNR)을 각각 도출한 후 신호대 잡음비가 가장 큰 배경 이미지를 상기 중요 키 프레임에 대한 배경 이미지로 선택하는 단계를 수행하는 것을 특징으로 하는 서버가 제공될 수 있다.
- [0024] 상기 중요 키 프레임을 선별하는 단계는, 상기 연기 프레임들 중 연속된 두개의 프레임을 HSV 색상 모델로 변환하는 단계; 및 상기 HSV 색상 모델로 변환된 두개의 프레임간의 색조(Hue) 컴포넌트의 에지 차이가 임계치 이상인 프레임을 중요 키 프레임으로 선별하는 단계를 포함할 수 있다.
- [0025] 상기 연기 프레임들을 학습된 인공 신경망에 적용하여 도출된 특징맵을 이용하여 관심 영역을 추출하는 단계; 및 상기 배경 이미지에 상응하여 상기 영상 감시 장치에서 분류된 물체 단서를 이용하여 상기 추출된 관심 영역에 대한 문맥 정보를 추출하는 단계를 더 포함할 수 있다.
- [0027] 본 발명의 또 다른 실시예에 따르면, 특정 공간을 촬영한 비디오 스트림을 분석하여 연기를 포함하는지 여부에 따라 연기(smoky) 프레임들과 비연기(non-smoky) 프레임들로 분류하며, 상기 비연기 프레임들을 이용하여 각 물체의 움직임 이벤트 발생 여부에 따라 각 물체에 대한 샷(shot)으로 각각 분할한 후 상기 각 샷에 상응하는 배경 이미지를 저장하는 영상 감시 장치; 및 상기 영상 감시 장치로부터 수신된 연기 프레임들을 분석하여 중요 키 프레임을 선별하고, 상기 중요 키 프레임과 상기 저장된 배경 이미지와 신호대 잡음비(PSNR)을 각각 도출한 후 신호대 잡음비가 가장 큰 배경 이미지를 상기 중요 키 프레임에 대한 배경 이미지로 선택하여 인공 신경망을 통해 분석하는 서버를 포함하는 영상 처리 시스템이 제공될 수 있다.
- [0028] 상기 영상 감시 장치는 상기 배경 이미지를 인공 신경망에 적용하여 각 물체를 분류할 수 있다.
- [0029] 상기 서버는, 상기 연기 프레임들을 학습된 인공 신경망에 적용하여 도출된 특징맵을 이용하여 관심 영역을 추출하며, 상기 배경 이미지에 상응하여 상기 영상 감시 장치에서 분류된 물체 단서를 이용하여 상기 추출된 관심 영역에 대한 문맥 정보를 추출할 수 있다.

발명의 효과

[0031] 본 발명의 일 실시예에 따른 연기를 포함하는 환경에 대한 영상 분석 방법 및 그 시스템을 제공함으로써, 연기가 자욱한 환경에서 구조 과정에서 도움이 될 수 있는 장면의 문맥을 분석할 수 있다.

도면의 간단한 설명

- [0033] 도 1은 본 발명의 일 실시예에 따른 영상 분석 시스템을 개략적으로 도시한 도면.
- 도 2는 본 발명의 일 실시예에 따른 영상 감시 장치에서의 영상 처리 방법을 나타낸 순서도.
- 도 3은 본 발명의 일 실시예에 따른 비연기 프레임들을 이용하여 키 프레임을 선별하는 방법을 나타낸 순서도.
- 도 4는 본 발명의 일 실시예에 따른 배경 이미지를 선택하는 방법을 나타낸 순서도.
- 도 5는 본 발명의 일 실시예에 따른 서버에서의 영상 분석 방법을 나타낸 순서도.
- 도 6은 본 발명의 일 실시예에 따른 서버에서의 영상 처리 방법을 나타낸 순서도.
- 도 7은 본 발명의 일 실시예에 따른 연기 환경에 가장 적합한 배경 이미지를 선택하는 방법을 나타낸 순서도.
- 도 8은 본 발명의 일 실시예에 따른 물체 분류 방법을 나타낸 순서도.
- 도 9는 본 발명의 일 실시예에 따른 영상 감시 장치의 내부 구성을 개략적으로 도시한 블록도.
- 도 10은 본 발명의 일 실시예에 따른 서버의 내부 구성을 개략적으로 도시한 블록도.

발명을 실시하기 위한 구체적인 내용

[0034] 본 명세서에서 사용되는 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "구성된다" 또는 "포함한다" 등의 용어는 명세서상에 기재된 여러 구성 요소들, 또는 여러 단계들을 반드시 모두 포함하는 것으로 해석되지 않아야 하며, 그 중 일부 구성 요소들 또는 일부 단계들은 포함되지 않을 수도 있고, 또는 추가적인 구성 요소 또는 단계들을 더 포함할 수 있는 것으로 해석되어야 한다. 또한, 명세서에 기재된 "...부", "모듈" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어 또는 소프트웨어로 구현되거나 하드웨어와 소프트웨어의 결합으로 구현될 수 있다.

[0035] 이하, 첨부된 도면들을 참조하여 본 발명의 실시예를 상세히 설명한다.

[0037] 도 1은 본 발명의 일 실시예에 따른 영상 분석 시스템을 개략적으로 도시한 도면이다.

[0038] 도 1을 참조하면, 본 발명의 일 실시예에 따른 영상 분석 시스템(100)은 영상 감시 장치(110) 및 서버(120)를 포함하여 구성된다.

[0039] 영상 감시 장치(110)는 특정 공간(지역)을 실시간으로 촬영하여 감시하기 위한 장치이다. 예를 들어, 영상 감시 장치(110)는 CCTV일 수 있다.

[0040] 영상 감시 장치(110)는 특정 공간(지역)을 실시간으로 촬영한 후 연기가 감지되는 경우, 해당 프레임을 서버(120)로 전송할 수 있다. 반면, 특정 공간(지역)을 촬영한 영상(프레임)에서 연기가 감지되지 않은 경우, 영상 감시 장치(110)는 배경으로 이용될 키 프레임을 선별하여 데이터베이스에 저장하며, 추후 서버(120)에서 이용될 수 있도록 각 물체를 분류하여 저장할 수 있다.

[0041] 서버(120)는 영상 감시 장치(110)로부터 수신된 연기 프레임들을 이용하여 중요 키 프레임을 선별한 후 각 영상 감시 장치(110)에서 사전 처리된 배경 이미지와 각 물체 분류 정보를 이용하여 장면에 대한 문맥 정보를 추출한다.

[0042] 즉, 각 영상 감시 장치(110)는 CCTV에 포함되거나 CCTV와 연결된 로컬 장치일 수 있다. 서버(120)가 각 로컬 장치인 수백 ~ 수천개에 이르는 CCTV 영상을 실시간으로 모두 처리하는 것은 현실적으로 불가능하며, 서버 과부하를 유발할 수 있다. 따라서, 본 발명의 일 실시예에 따른 영상 분석 시스템(100)은 연기를 포함하지 않는 비연기 프레임들에 대한 사전 처리를 각 로컬 장치(즉, 영상 감시 장치(110))에서 수행되며, 연기를 포함하는 구조 활동에 직접적으로 연결되는 중요한 영상들을 서버(120)에서 선별적으로 처리하도록 할 수 있다.

[0043] 따라서, 도 2를 참조하여 각 로컬에 위치되는 영상 감시 장치에서의 영상 처리 방법에 대해 우선 설명하기로 한

다.

- [0045] 도 2는 본 발명의 일 실시예에 따른 영상 감시 장치에서의 영상 처리 방법을 나타낸 순서도이다.
- [0046] 단계 210에서 영상 감시 장치(110)는 비디오 스트림을 입력받는다. 영상 감시 장치(110)는 특정 공간(지역)을 실시간 촬영하는 장치로, 실시간으로 비디오 스트림이 지속적으로 입력될 수 있다.
- [0047] 단계 215에서 영상 감시 장치(110)는 비디오 스트림의 각 프레임을 분석하여 연기를 포함하는지 여부를 판단한다.
- [0048] 예를 들어, 영상 감시 장치(110)는 CNN 모델을 기반으로 각각의 프레임을 분석하여 연기를 포함하는지 여부를 판단할 수 있다.
- [0049] 이하, 본 명세서에서는 이해와 설명의 편의를 도모하기 위해 CNN 모델에 의해 연기를 포함하는 것으로 분류된 프레임을 연기(smoky) 프레임으로 통칭하기로 하며, 연기를 포함하지 않는 것으로 분류된 프레임을 비연기(non-smoky) 프레임으로 통칭하여 설명하기로 한다.
- [0050] 만일 연기 프레임인 경우, 단계 220에서 영상 감시 장치(110)는 연기 프레임을 서버(120)로 전송한다.
- [0051] 그러나 만일 연기 프레임이 아닌 경우(즉, 비연기 프레임인 경우), 단계 225에서 영상 감시 장치(110)는 비연기 프레임들을 이용하여 배경에 해당하는 키 프레임을 선별한다. 이에 대해서는 도 3 및 도 4를 참조하여 하기에서 보다 상세히 설명하기로 한다.
- [0052] 단계 230에서 영상 감시 장치(110)는 선별된 배경 프레임을 이용하여 움직임 이벤트가 발생한 물체를 분류한다.
- [0053] 영상 감시 장치(110)는 훈련된 CNN 모델을 이용하여 비디오 스트림을 연기를 포함하는 프레임들과 연기를 포함하지 않은 프레임을 구분한 후 비연기 프레임들을 이용하여 깨끗한 배경을 선별하고, 움직이는 각각의 물체를 분류하는 사전 작업을 수행할 수 있다.
- [0055] 도 3은 본 발명의 일 실시예에 따른 비연기 프레임들을 이용하여 키 프레임을 선별하는 방법을 나타낸 순서도이다.
- [0056] 단계 310에서 영상 감시 장치(110)는 비연기 프레임들에서 물체들의 움직임 이벤트를 기준으로 샷(shots)을 분할한다.
- [0057] 여기서, 샷(shot)은 짧은 동영상을 의미한다. 즉, 본 발명의 일 실시예에서 샷은 움직임이 있는 물체를 기준으로 추출된다. 예를 들어, 의자가 이동된 경우, 의자의 움직임이 발생한 프레임부터 의자의 움직임이 완료된 마지막 프레임까지가 의자에 대한 샷으로 분할될 수 있다.
- [0058] 이와 같은 방식으로 영상 감시 장치(110)는 비연기 프레임들을 대상으로 움직임이 있는 물체에 대한 샷(shots)을 분할(segmentation)할 수 있다.
- [0059] 단계 315에서 영상 감시 장치(110)는 분할된 샷 중 어느 하나를 선택한다.
- [0060] 단계 320에서 영상 감시 장치(110)는 선택된 샷에 포함된 프레임들 중 어느 하나를 선택한다.
- [0061] 단계 325에서 영상 감시 장치(110)는 선택된 프레임에 대한 기억 가능성 점수(memorability)와 엔트로피(entropy)를 계산한다.
- [0062] 영상 감시 장치(110)는 CNN을 기반으로 선택된 프레임에 대한 기억 가능성 점수를 계산할 수 있다. 이에 대해서는 "Understanding and Predicting Image Memorability at a Large Scale A. Khosla, A. S. Raju, A. Torralba and A. Oliva International Conference on Computer Vision (ICCV), 2015"에 상세히 설명되어 있는 바 추가적인 설명은 생략하기로 한다.
- [0063] 또한, 영상 감시 장치(110)는 선택된 프레임을 HSV 칼라 스페이스로 변환한다. 이어, 영상 감시 장치(110)는 HSV 칼라 스페이스로 변환된 선택된 프레임에 대한 색상(Hue), 채도(saturation), 명도(value) 컴포넌트를 각각 분리한 후 각각의 컴포넌트를 히스토그램으로 변환하여 양자화한다. 이때, 영상 감시 장치(110)는 8 히스토그램 빈으로 색상(Hue), 채도(saturation), 명도(value) 컴포넌트 각각을 양자화할 수 있다.
- [0064] 또한, 영상 감시 장치(110)는 히스토그램 빈의 확률(probability)를 계산한 후 이를 이용하여 해당 프레임에 대한 엔트로피를 계산할 수 있다.

- [0065] 단계 330에서 영상 감시 장치(110)는 선택된 프레임에 상응하여 계산된 기억 가능성 점수 및 엔트로피가 임계치 이상인지 여부를 판단한다.
- [0066] 만일 임계치 이상인 경우, 단계 335에서 영상 감시 장치(110)는 해당 선택된 프레임을 키 프레임으로 선별한다.
- [0067] 그러나 만일 임계치 미만인 경우, 샷 내의 다른 프레임에 대해 해당 과정을 수행하도록 할 수 있다. 즉, 이와 같은 방식으로 샷 내의 모든 프레임들을 대상으로 단계 320 내지 단계 335 과정을 반복 수행할 수 있다.
- [0068] 도 3에서는 하나의 샷에 대해 키 프레임을 선별하는 과정을 설명하였으나, 이는 추출된 모든 샷들을 대상으로 반복 수행될 수 있음은 당연하다.
- [0070] 도 4는 본 발명의 일 실시예에 따른 배경 이미지를 선택하는 방법을 나타낸 순서도이다. 도 4는 도 3과 같이 샷들을 중심으로 키 프레임이 선별된 이후 각 샷에 대한 배경을 선별하는 과정에 대해 설명하기로 한다.
- [0071] 단계 410에서 영상 감시 장치(110)는 각각의 샷에 상응하는 키 프레임들 중 어느 하나를 배경 이미지로 선택하여 저장한다.
- [0072] 예를 들어, 영상 감시 장치(110)는 샷에 상응하는 제1 키 프레임을 배경 이미지로 선택할 수 있다.
- [0073] 단계 413에서 영상 감시 장치(110)는 제2 키 프레임을 선택한다.
- [0074] 단계 415에서 영상 감시 장치(110)는 제2 키 프레임을 선택한 후 물체 변경이 임계값 이상인지 여부를 판단한다. 예를 들어, 영상 감시 장치(110)는 제2 키 프레임과 배경 이미지간의 차영상을 이용하여 물체 변경을 감지할 수 있다. 또한, 영상 감시 장치(110)는 제2 키 프레임과 배경 이미지간의 차영상의 전체 합산값이 임계값 이상인지 여부를 판단할 수 있다.
- [0075] 만일 물체 변경 정도가 임계값 이상인 경우, 단계 420에서 영상 감시 장치(110)는 제2 프레임을 배경 이미지로 갱신 저장한다.
- [0076] 그러나 만일 물체 변경 정도가 임계값 미만인 경우, 단계 425에서 영상 감시 장치(110)는 마지막 키 프레임인지 여부를 판단한다.
- [0077] 만일 마지막 키 프레임이 아니면 영상 감시 장치(110)는 다음 키 프레임을 선택한 후 단계 415로 진행한다.
- [0078] 그러나 만일 마지막 키 프레임이면 해당 샷에 대한 배경 이미지 선택 과정을 종료한다. 도 4는 각각의 샷에 대한 키 프레임들을 대상으로 수행될 수 있다. 따라서, 도 4의 과정을 통해 각각의 샷에 대한 배경 이미지를 선택함에 있어, 각 샷의 키 프레임들 중 움직임 변화가 발생한 이후 가장 깨끗한 키 프레임을 해당 샷의 배경 이미지로 선택할 수 있다.
- [0079] 영상 감시 장치(110)는 각 샷에 상응하여 선별된 배경 이미지를 데이터베이스에 저장할 수 있다.
- [0081] 도 5는 본 발명의 일 실시예에 따른 서버에서의 영상 분석 방법을 나타낸 순서도이다. 이하에서는 영상 감시 장치(110)로부터 연기 프레임들을 입력받아 분석하는 방법에 대해 설명하기로 한다.
- [0082] 단계 510에서 서버(120)는 연기 프레임들을 분석하여 중요한 키 프레임을 선별한다. 이에 대해서는 도 6을 참조하여 보다 상세히 설명하기로 한다.
- [0083] 단계 515에서 서버(120)는 선별된 중요 키 프레임들을 이용하여 배경 이미지를 선택한다.
- [0084] 예를 들어, 서버(120)는 선별된 중요 키 프레임들과 DB에 저장된 배경 이미지들간의 매칭을 통해 중요 키 프레임들과 유사한 배경 이미지를 선택할 수 있다. 이에 대해서는 도 7를 참조하여 보다 상세히 설명하기로 한다.
- [0085] 단계 520에서 서버(120)는 선택된 배경 이미지를 이용하여 관심 영역을 추출하고, 추출된 관심 영역에서의 물체를 분석한다.
- [0086] 이에 대해서는 도 8을 참조하여 보다 상세히 설명하기로 한다.
- [0088] 도 6은 본 발명의 일 실시예에 따른 서버에서의 영상 처리 방법을 나타낸 순서도이다.
- [0089] 단계 610에서 서버(120)는 연기 프레임들 중 두개의 연속 프레임(F_i, F_{i+1})을 선택하여 HSV 색상 모델로 변환한다.
- [0090] 단계 615에서 서버(120)는 HSV 색상 모델로 변환된 두 프레임의 색조 컴포넌트의 에지 차이(d 라 칭하기로 함)를

도출한다.

- [0091] 편의상 F_i 의 색조 컴포넌트를 H_i 라 가정하기로 하며, F_{i+1} 의 색조 컴포넌트를 H_{i+1} 라 가정하기로 한다. H_i 와 H_{i+1} 의 차이 도출함으로써 에지 차이가 도출될 수 있다.
- [0092] 단계 620에서 서버(120)는 도출된 두 프레임의 색조 컴포넌트의 에지 차이(d)가 임계값 이상인지 여부를 판단한다.
- [0093] 만일 임계값 미만이면, 다음 프레임을 선택한 후 단계 610으로 진행한다.
- [0094] 그러나 만일 임계값 이상이면, 단계 625에서 서버(120)는 현재 프레임을 후보 키프레임으로 선별한다.
- [0095] 단계 630에서 서버(120)는 마지막 프레임인지 여부를 판단한다.
- [0096] 만일 마지막 프레임이 아닌 경우, 다음 프레임을 선택하여 단계 610으로 진행한다.
- [0097] 그러나 만일 마지막 프레임이면, 단계 635에서 서버(120)는 후보 키프레임들을 정제하여 최종 키 프레임을 선별한다.
- [0098] 예를 들어, 서버(120)는 후보 키프레임들 중 중복되는 키 프레임을 제거함으로써 키 프레임을 정제하여 적어도 하나의 최종 키 프레임을 선별할 수 있다.
- [0100] 도 7은 본 발명의 일 실시예에 따른 연기 환경에 가장 적합한 배경 이미지를 선택하는 방법을 나타낸 순서도이다.
- [0101] 단계 710에서 서버(120)는 중요한 키 프레임들 중 어느 하나의 키 프레임을 선택한다. 이해와 설명의 편의를 도모하기 위해 선택된 키 프레임을 F_i 라 칭하기로 한다.
- [0102] 단계 715에서 서버(120)는 선택된 키 프레임과 DB에 저장된 배경 이미지간의 신호대 잡음비(PSNR: peak signal to noise ratio)을 각각 도출한다.
- [0103] 단계 720에서 서버(120)는 신호대 잡음비가 가장 낮은 배경 이미지를 해당 선택된 키 프레임에 대한 배경 이미지로 선택한다.
- [0104] 단계 725에서 서버(120)는 선택된 키 프레임이 마지막 프레임인지 여부를 판단한다.
- [0105] 만일 마지막 프레임인 경우, 도 7의 과정을 종료한다. 그러나 만일 마지막 프레임이 아닌 경우, 중요한 키 프레임들 중 다른 키 프레임에 대한 배경 이미지를 선택하기 위해 단계 710으로 진행한다.
- [0106] 이를 통해, 서버(120)는 연기를 포함하는 각각의 중요한 키 프레임에 대해 연기를 포함하지 않은 가장 유사한 깨끗한 배경 이미지를 선택할 수 있다.
- [0108] 도 8은 본 발명의 일 실시예에 따른 물체 분류 방법을 나타낸 순서도이다.
- [0109] 단계 810에서 서버(120)는 연기 프레임들을 인공 신경망(CNN)에 적용하여 특징정보를 추출한 후 분석함으로써 관심 영역을 추출한다.
- [0110] 단계 815에서 서버(120)는 배경 이미지의 물체 단서를 이용하여 추출된 관심 영역에 대한 문맥 정보를 추출한다.
- [0111] 예를 들어, 서버(120)는 추출된 관심 영역에 포함된 각각의 물체 위치 정보를 추출할 수 있다. 또한, 서버(120)는 영상 감시 장치에서 각 배경 이미지에 상응하여 훈련된 CNN을 기반으로 추출된 각 물체 단서를 이용하여 관심 영역에 포함된 각 물체 위치 정보에 상응하는 문맥 정보를 추출할 수 있다. 여기서, 문맥 정보는 물체 유형(예를 들어, 실내, 실외), 사람 수, 군중 밀도 등일 수 있다.
- [0113] 도 9는 본 발명의 일 실시예에 따른 영상 감시 장치의 내부 구성을 개략적으로 도시한 블록도이다.
- [0114] 도 9를 참조하면, 본 발명의 일 실시예에 따른 영상 감시 장치(110)는 영상 획득부(910), 통신부(915), 프레임 분류부(920), 배경 저장부(925), 물체 분류부(930), 메모리(935) 및 프로세서(940)를 포함하여 구성된다.
- [0115] 영상 획득부(910)는 특정 공간(지역)에 대한 비디오 스트림을 획득한다. 여기서, 영상 획득부(910)는 카메라일 수 있다.

- [0116] 통신부(915)는 통신망을 통해 다른 장치(예를 들어, 서버(120))와 데이터를 송수신하기 위한 수단이다.
- [0117] 프레임 분류부(920)는 비디오 스트림을 분석하여 각 프레임을 연기를 포함하는 연기 프레임과 연기를 포함하지 않는 비연기 프레임으로 분류한다.
- [0118] 프레임 분류부(920)에 의해 연기 프레임으로 분류된 프레임들은 통신부(915)를 통해 서버(120)로 전송될 수 있다. 또한, 비연기 프레임은 영상 감시 장치(110)에 의해 전처리를 위해 처리된다.
- [0119] 배경 저장부(925)는 비연기 프레임들을 이용하여 각 물체의 움직임 이벤트 발생 여부에 따라 각 물체에 대한 샷(shot)을 각각 분할하고, 각 샷에 상응하는 배경 이미지를 저장한다. 이에 대해서는 도 3 및 도 4를 이용하여 설명한 바와 동일하므로 중복되는 설명은 생략하기로 한다.
- [0120] 배경 저장부(925)에 의해 각 샷(shot)에서 각 물체에 대한 움직임 이벤트가 발생한 이후 가장 깨끗한 프레임이 배경 이미지로 선택될 수 있다. 이는 도 3 및 도 4에 설명된 바와 같다.
- [0121] 물체 분류부(930)는 각 물체를 훈련된 인공 신경망을 통해 각각 분류한다. 인공 신경망은 CNN일 수 있다. CNN을 통해 각 물체를 분류하는 방법 자체는 당업자에는 자명한 사항이므로 이에 대한 별도의 설명은 생략하기로 한다.
- [0122] 메모리(935)는 영상 감시 장치(110)를 운용하기 위해 필요한 다양한 명령어들을 저장한다.
- [0123] 프로세서(940)는 본 발명의 일 실시예에 따른 영상 감시 장치(110)의 내부 구성 요소들(예를 들어, 영상 획득부(910), 통신부(915), 프레임 분류부(920), 배경 저장부(925), 물체 분류부(930), 메모리(935) 등)를 제어하기 위한 수단이다.
- [0125] 도 10은 본 발명의 일 실시예에 따른 서버의 내부 구성을 개략적으로 도시한 블록도이다.
- [0126] 도 10을 참조하면, 본 발명의 일 실시예에 따른 서버(120)는 통신부(1010), 메모리(1015) 및 프로세서(1020)를 포함한다.
- [0127] 통신부(1010)는 통신망을 통해 다른 장치들(예를 들어, 영상 감시 장치(110))과 데이터를 송수신하기 위한 수단이다.
- [0128] 메모리(1015)는 적어도 하나의 명령어들을 저장한다.
- [0129] 프로세서(1020)는 통신부(1010) 및 메모리(1015)를 제어하며, 메모리(1015)에 저장된 명령어들을 실행할 수 있다. 프로세서(1020)에 의해 실행된 명령어들은 도 5 내지 도 7에서 설명된 각각의 단계를 수행할 수 있다. 이에 대해서는 도 5 내지 도 7에 상세히 설명되어 있는바 중복되는 설명은 생략하기로 한다.
- [0131] 본 발명의 실시 예에 따른 장치 및 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 컴퓨터 판독 가능 매체에 기록되는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 분야 통상의 기술자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media) 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.
- [0132] 상술한 하드웨어 장치는 본 발명의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0133] 이제까지 본 발명에 대하여 그 실시 예들을 중심으로 살펴보았다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 그러므로 개시된 실시 예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

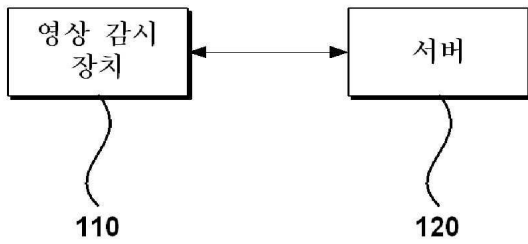
부호의 설명

- [0135] 100: 영상 분석 시스템
- 110: 영상 감시 장치
- 120: 서버

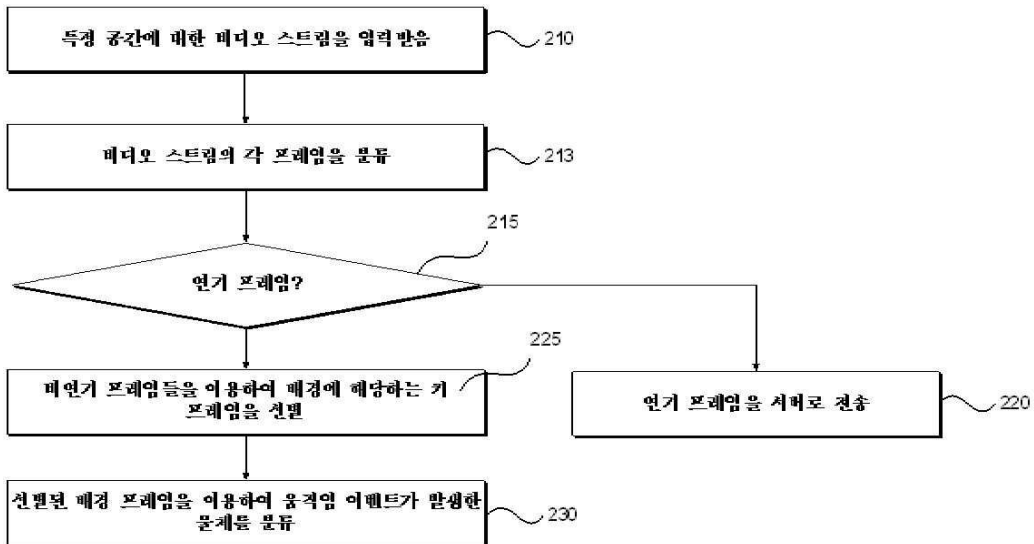
도면

도면1

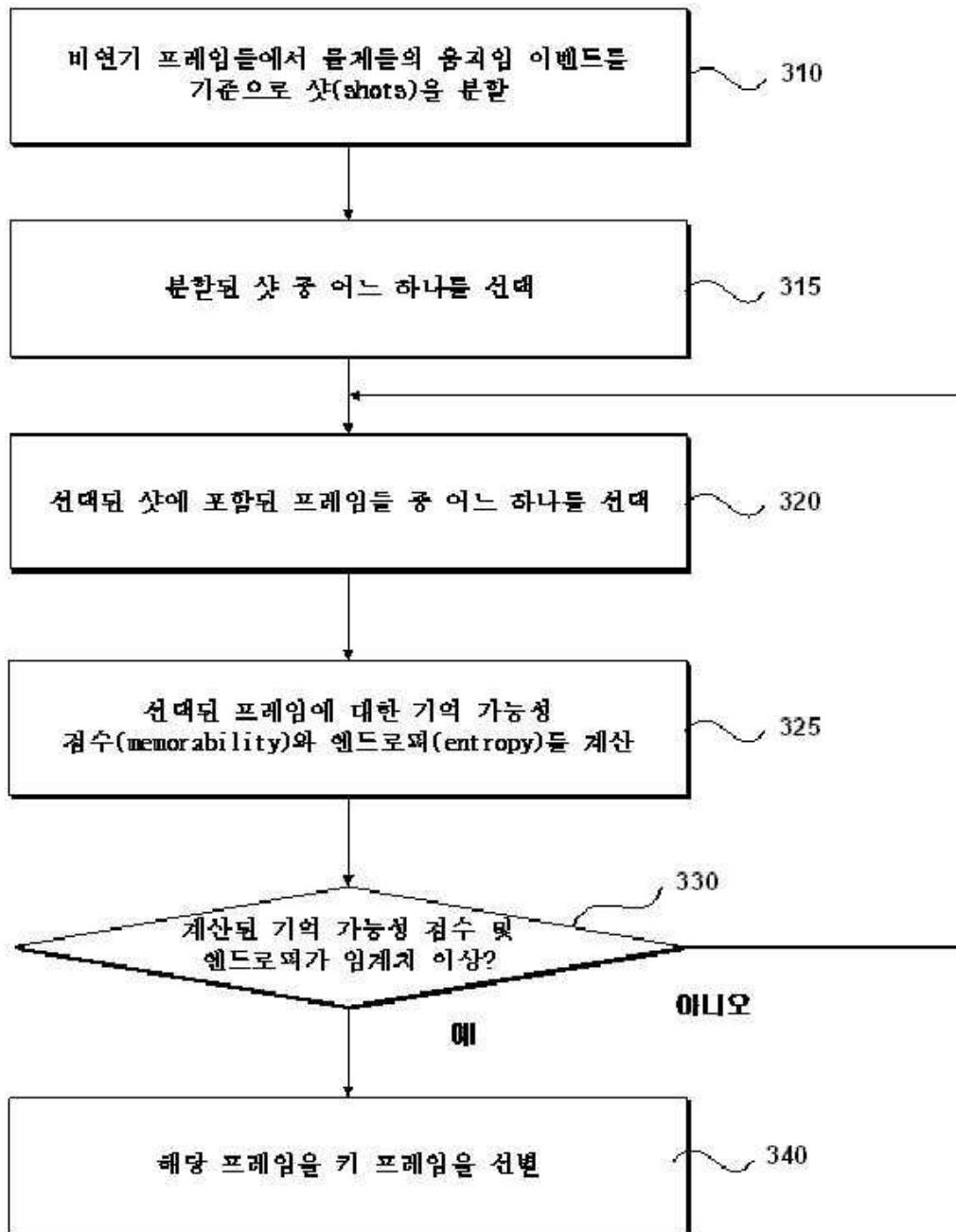
100



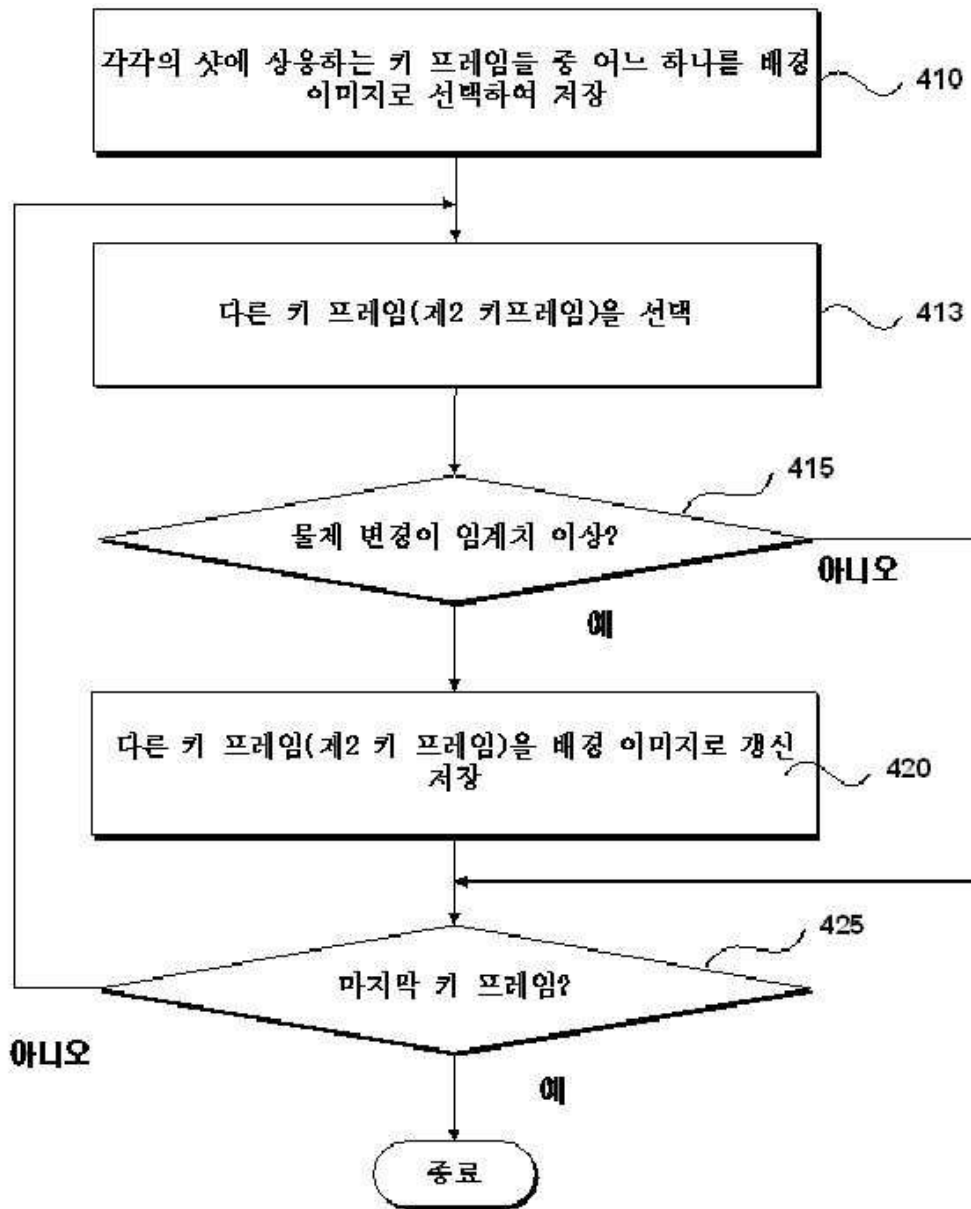
도면2



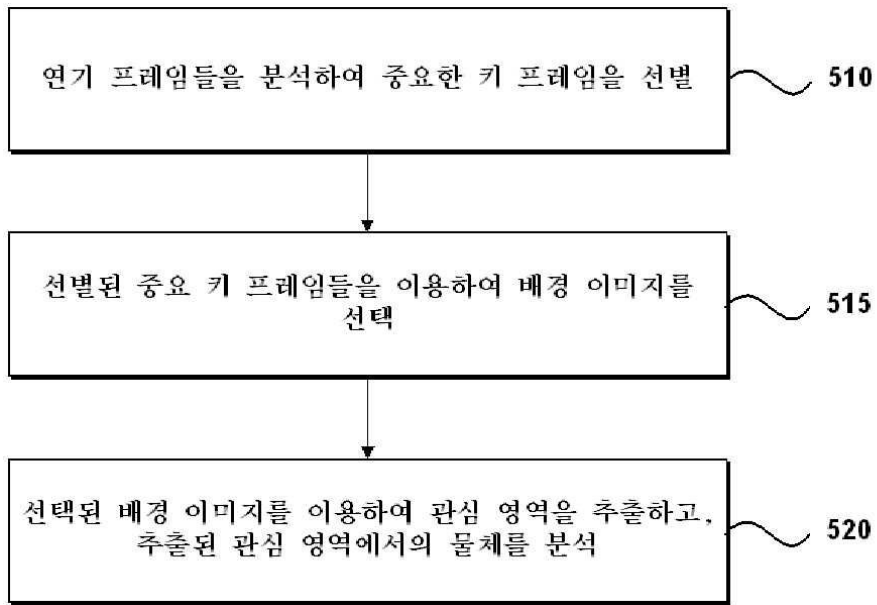
도면3



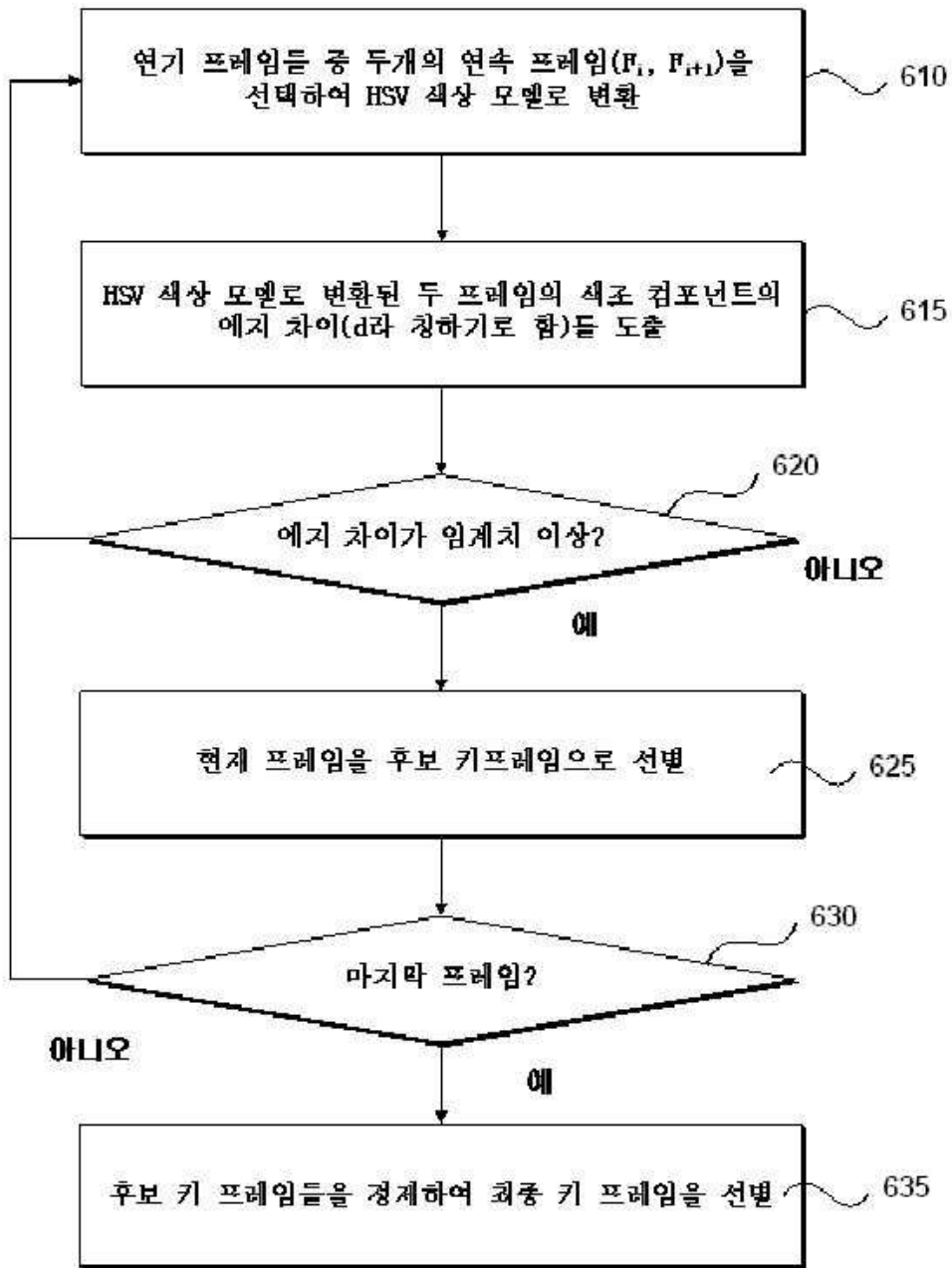
도면4



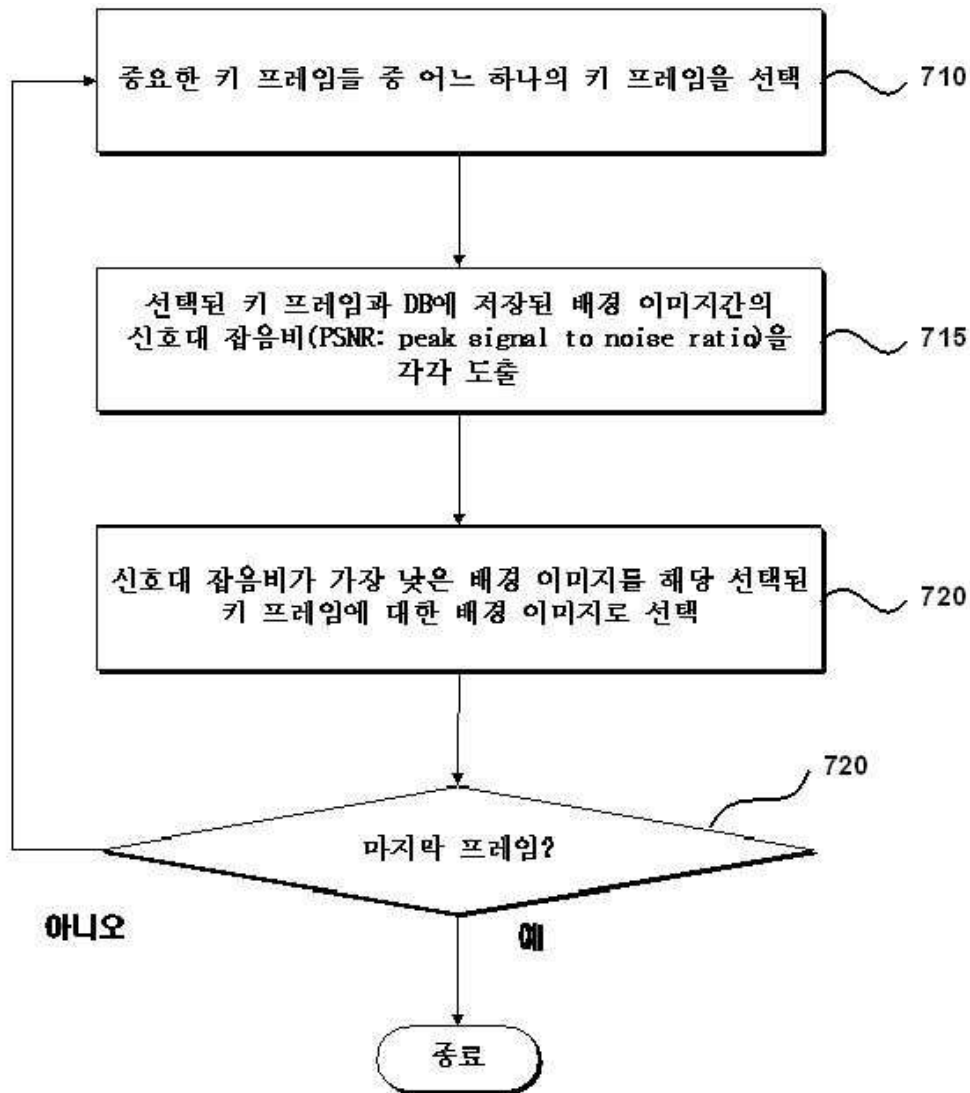
도면5



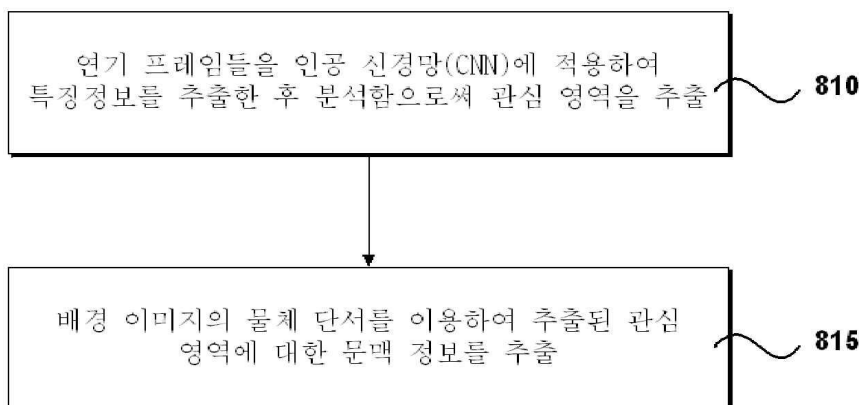
도면6



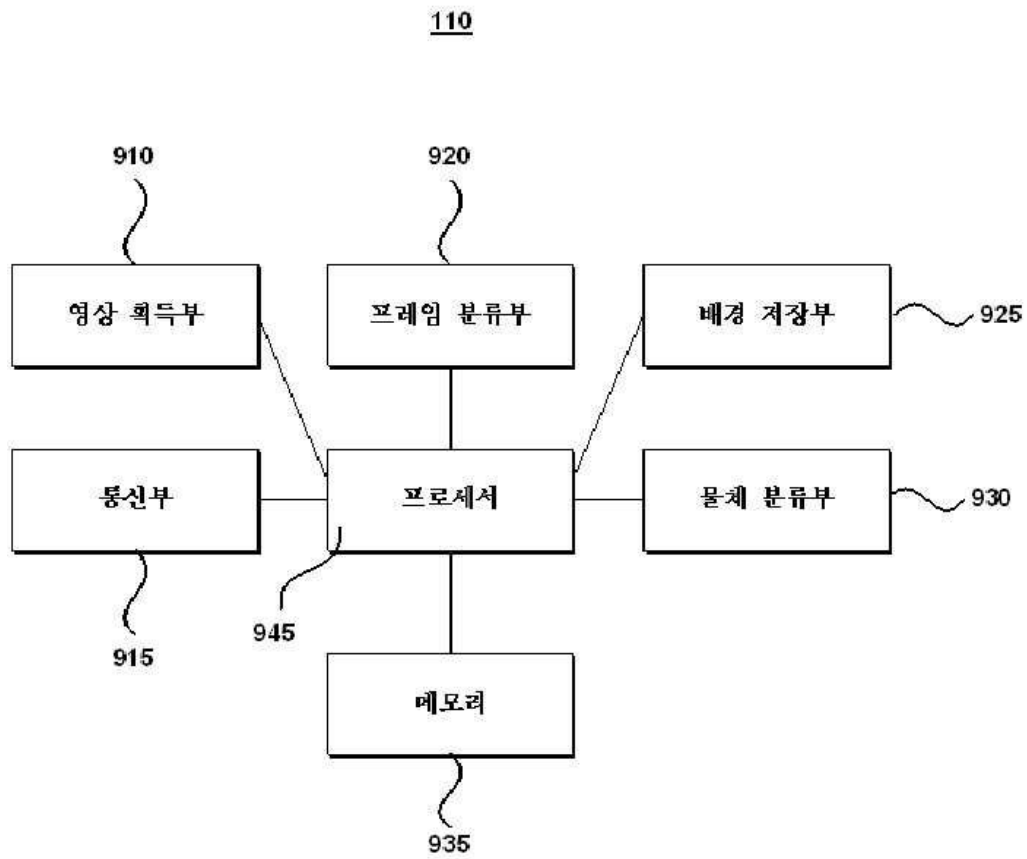
도면7



도면8



도면9



도면10

