■ 기술명 : (인공지능 기반 고해상도 동영상 프레임 율 고속 변환 기술)

(영문 : High-speed video frame rate conversion based on AI)

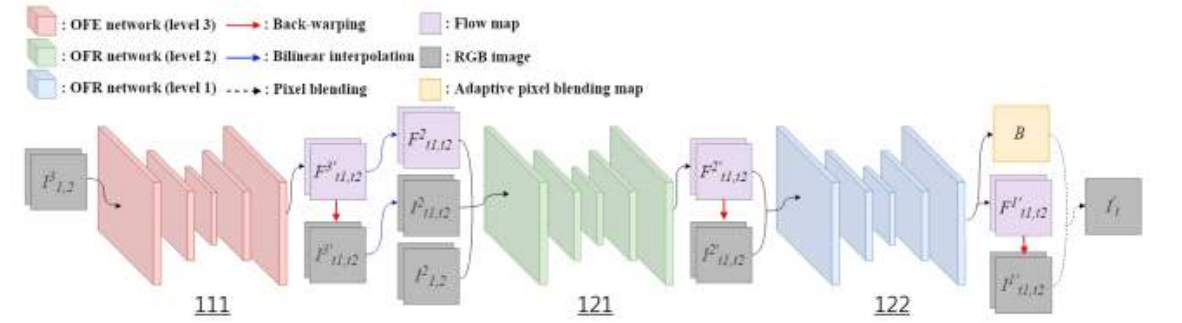
산업기술분류	정보통신 - 디지털 방송 - 디지털 방송 콘텐츠 (300203)
Key-word(국문)	인공지능, 프레임 율 변환, 고속 변환, 슬로 모션
Key-word(영문)	AI, frame rate conversion, fast conversion, slow motion

■ 기술의 개요

○ (배경) 4K 이상의 고해상도 대화면 디스플레이가 주류로 자리 잡음. 이런 디스플레이를 위한 고품질의 영상은 높은 해상도뿐만 아니라 높은 프레임 율이 요구됨. 그러나 대부분의 기존 영상은 낮은 프레임 율을 지니고 있으므로 이를 높은 프레임 율로 변환할 수 있는 기술이 요구됨. 또한 최근 스마트폰에서 240fps의 초고속 촬영을 지원하면서 인상적인 장면을 슬로 모션으로 기록하는 경우가 많아짐. 기존의 낮은 프레임 율 영상을 슬로 모션으로 보기 위해서는 높은 프레임 율로 변환할 수 있는 기술이 요구됨

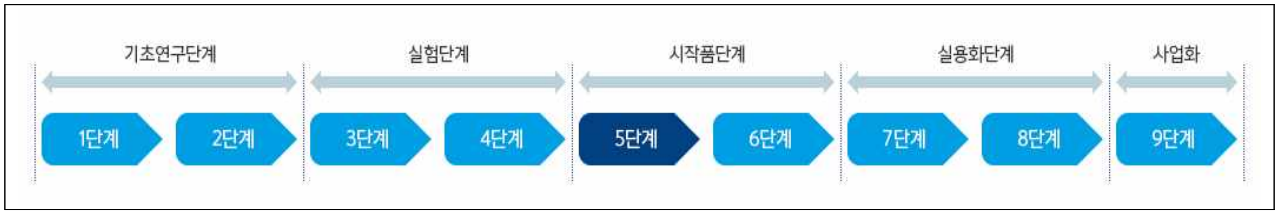
○ (개요) 최근 인공 지능 기반 프레임 율 고속 기법이 등장하여 전통적인 기법에 비해 더욱 우수한 성능을 제공함. 그러나 기존 인공 지능 기법들은 고해상도 영상에서 과도한 GPU 메모리와 연산량을 요구함. 이를 극복하기 위해 계층적 인공 지능 네트워크를 개발하여 저 복잡도 및 고성능 프레임 율 변환 기술을 개발함

<기술 개요도 >



[명칭: 계층적 구조 기반 프레임 율 변환 인공지능 네트워크]
(관련특허 : 10-2019-0137132)

■ 기술의 구현수준(TRL)



■ 구현기술 명세 및 장점(경쟁기술과의 차별성)

○ 당해단계 기술명세

- 본 기술은 고해상도 영상의 프레임 율 변환을 효과적으로 수행하기 위해 다음과 같이 3단계 네트워크로 구성되어 있음.
- (1단계: 옵티컬 플로우 산출) 첫 번째 네트워크는 원영상 해상도의 1/4로 감소된 영상을 옵티컬 플로우 네트워크의 입력으로 사용하여 옵티컬 플로우를 산출함.
- (2단계: 옵티컬 플로우 정제) 네트워크는 1/4 해상도의 옵티컬 플로우를 1/2의 해상도로 증가시켜주는 옵티컬 플로우 정제 (refinement) 네트워크임
- (3단계: 초해상화) 1/2 해상도의 중간 프레임을 원본 해상도로 증가시키는 초해상도 (Super Resolution) 수행 네트워크임.

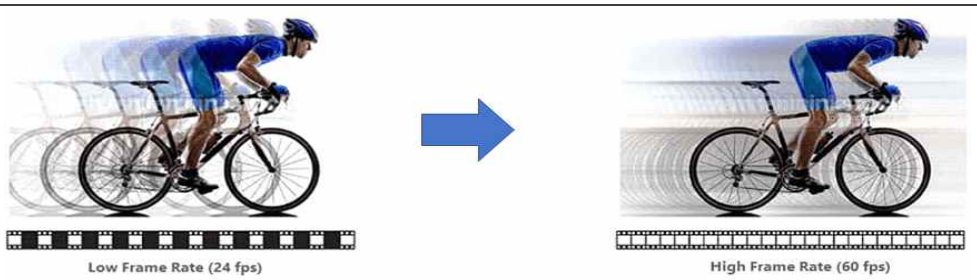
○ 추후단계 확보대상 기술명세

- 네트워크 양자화 기술을 통한 딥러닝 네트워크 추가 경량화
- 클라우드 기반 딥러닝 네트워크 시스템

○ 보유기술 장점 및 차별성(경쟁기술대비)

- 경쟁 기술 대비 더 적은 GPU 메모리를 사용하고 기존 기술 대비 최대 10배 이상의 속도 향상을 보임
- 4K와 같은 고해상도 영상에 대하여 기존 대비 더욱 높은 품질의 영상을 제공
- 조명 변화와 텍스트에 더욱 강건함.

■ 활용범위 및 응용분야



[프레임 율 업 컨버젼 분야]

[슬로 모션 비디오 생성기]

○ 프레임 율 업 컨버젼 엔진 SW
○ 슬로 모션 (slow motion) 비디오 생성기 (스마트폰, PC 등)
○ TV 및 셋톱 박스의 프레임 율 업 컨버젼 엔진

■ 지식재산권 현황

구분	발명의 명칭	출원번호 (출원일)	등록번호 (등록일)
특허	딥러닝 기반 동영상 프레임 율 변환 방법 및 장치	2018-0144595 (2018.11.21)	
특허	고해상도 동영상 프레임 율 고속 변환 방법 및 장치	2019-0137132 (2019.10.31)	
특허	데이터 변형을 통한 고해상도 동영상 프레임 율 고속 변환 방법 및 장치	2019-0167488 (2019.12.16.)	

(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년07월21일
(11) 등록번호 10-2136468
(24) 등록일자 2020년07월15일

(51) 국제특허분류(Int. Cl.)
H04N 7/01 (2006.01) *G06N 3/08* (2006.01)
(52) CPC특허분류
H04N 7/0127 (2013.01)
G06N 3/08 (2013.01)
(21) 출원번호 10-2018-0144595
(22) 출원일자 2018년11월21일
심사청구일자 2019년02월12일
(65) 공개번호 10-2020-0059627
(43) 공개일자 2020년05월29일
(56) 선행기술조사문헌
KR1020150004916 A*
US20090279111 A1*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
정진우
서울특별시 송파구 송파대로 111, 파크하비오105동 1205호
안하은
서울특별시 성북구 성북로8다길 37, 102호
김제우
경기도 성남시 분당구 수내로 181, 303동 901호
(74) 대리인
남충우

전체 청구항 수 : 총 11 항 심사관 : 박재학

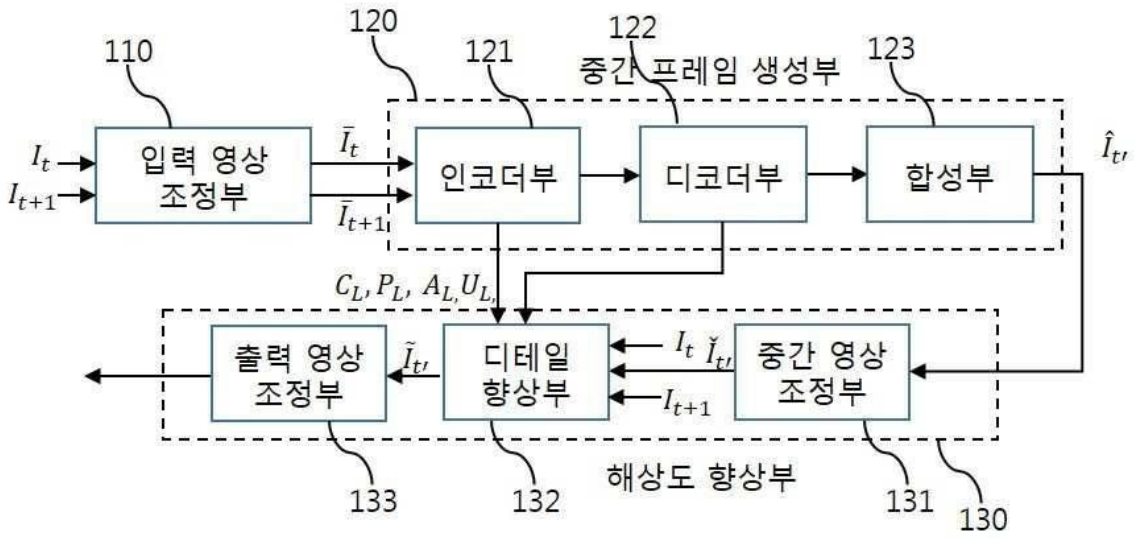
(54) 발명의 명칭 **딥러닝 기반 동영상 프레임 율 변환 방법 및 장치**

(57) 요 약

딥러닝에 기반하여 다양한 컬러 영상 포맷에 대응하는 동시에 고품질, 고속으로 프레임을 보간하는 동영상 프레임 율 변환 방법 및 장치가 제공된다. 본 발명의 실시예에 따른 동영상 프레임 율 변환 장치는 각 채널의 영상에 대한 해상도를 조정하는 조정부; 조정부에서 해상도가 조정된 t초와 t+1초의 연속된 두 영상으로 중간 영상을 생성하는 생성부; 및 생성부에서 생성된 중간 영상의 해상도를 향상시키는 향상부;를 포함한다.

이에 의해, 딥러닝 기반 동영상 프레임 율 변환에 있어, 딥러닝의 입력이 RGB 컬러 포맷 뿐만 아니라 YCbCr 등 다양한 컬러 포맷에 적용할 수 있게 되며, 프레임 보간 방법과 해상도 향상 방법을 동시에 이용하여 고품질, 고속의 프레임 율 변환이 가능해진다.

대 표 도



(52) CPC특허분류
H04N 7/0135 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 1711076193
부처명 과학기술정보통신부
연구관리전문기관 정보통신기술진흥센터
연구사업명 차세대(UHD)방송서비스활성화기술개발
연구과제명 영상콘텐츠 초고속/초고화질 변환 기술 개발
기 여 율 1/1
주관기관 (주)픽스트리
연구기간 2018.07.01 ~ 2018.12.31

명 세 서

청구범위

청구항 1

각 채널의 영상에 대한 해상도를 조정하는 조정부;

조정부에서 해상도가 조정된 t초와 t+1초의 연속된 두 영상으로 중간 영상을 생성하는 생성부; 및

생성부에서 생성된 중간 영상의 해상도를 향상시키는 향상부;를 포함하고,

중간 영상은,

t+0.5초의 영상이며,

조정부는,

각 채널의 영상에 대한 해상도를 동일하게 서브 샘플링하는 것을 특징으로 하는 동영상 프레임 율 변환 장치.

청구항 2

삭제

청구항 3

청구항 1에 있어서,

생성부는,

해상도가 조정된 영상의 특성을 계산하는 인코더부;

계산된 특성을 영상의 해상도 크기만큼 복원하는 디코더부;

조정부에서 해상도가 조정된 두 영상과 디코더의 출력을 이용하여 중간 영상을 생성하는 합성부;를 포함하는 것을 특징으로 하는 동영상 프레임 율 변환 장치.

청구항 4

청구항 1에 있어서,

향상부는,

생성부에서 생성된 중간 영상의 해상도를 원본 영상의 해상도로 향상시키는 것을 특징으로 하는 동영상 프레임 율 변환 장치.

청구항 5

청구항 4에 있어서,

향상부는,

디테일을 개선할 중간 영상의 해상도를 조정하는 제1 조정부;

원본 영상을 이용하여, 해상도가 조정된 중간 영상의 디테일을 개선하는 디테일 향상부;

제1 조정부에서 해상도가 조정되지 않은 중간 영상의 해상도를 조정하는 제2 조정부;를 포함하는 것을 특징으로 하는 동영상 프레임 율 변환 장치.

청구항 6

청구항 5에 있어서,

디테일 향상부는,

선별적으로 선택된 채널만을 입력받는 것을 특징으로 하는 동영상 프레임 율 변환 장치.

청구항 7

청구항 6에 있어서,

선택된 채널은,

밝기 신호인 것을 특징으로 하는 동영상 프레임 율 변환 장치.

청구항 8

청구항 5에 있어서,

디테일 향상부는,

생성부에서 생성되는 중간 정보를 참조하여, 중간 영상의 디테일을 개선하는 것을 특징으로 하는 동영상 프레임 율 변환 장치.

청구항 9

청구항 5에 있어서,

디테일 향상부는,

딥러닝 네트워크를 사용하는 것을 특징으로 하는 동영상 프레임 율 변환 장치.

청구항 10

각 채널의 영상에 대한 해상도를 조정하는 단계;

해상도가 조정된 t초와 t+1초의 연속된 두 영상으로 중간 영상을 생성하는 단계; 및

생성된 중간 영상의 해상도를 향상시키는 단계;를 포함하고,

중간 영상은,

t+0.5초의 영상이며,

조정 단계는,

각 채널의 영상에 대한 해상도를 동일하게 서브 샘플링하는 것을 특징으로 하는 동영상 프레임 율 변환 장치.

청구항 11

각 채널의 영상에 대한 해상도를 조정하는 조정부;

조정부에서 해상도가 조정된 t초와 t+1초의 연속된 두 영상으로 생성된 중간 영상의 해상도를 향상시키는 향상부;를 포함하고,

중간 영상은,

t+0.5초의 영상이며,

조정부는,

각 채널의 영상에 대한 해상도를 동일하게 서브 샘플링하는 것을 특징으로 하는 동영상 프레임 율 변환 장치.

청구항 12

각 채널의 영상에 대한 해상도를 조정하는 단계;

조정부에서 해상도가 조정된 t초와 t+1초의 연속된 두 영상으로 생성된 중간 영상의 해상도를 향상시키는 단계;를 포함하고,

중간 영상은,

t+0.5초의 영상이며,

조정 단계는,

각 채널의 영상에 대한 해상도를 동일하게 서브 샘플링하는 것을 특징으로 하는 동영상 프레임 율 변환 장치.

발명의 설명

기 술 분 야

[0001] 본 발명은 동영상의 프레임 율 변환 방법에 관한 것으로, 더욱 상세하게는 딥러닝 기법에 기반한 프레임 보간 기법과 해상도 증가 기법을 적용하여 고속으로 프레임 율을 증가시키는 방법 및 장치에 관한 것이다.

배 경 기 술

[0003] 1) 동영상 프레임 율 변환 기법 개요

[0004] 동영상은 연속된 정지 영상의 집합으로 구성된다. 비디오에서 정지 영상을 프레임이라고 부르며 단위 시간 당 프레임의 수를 동영상의 프레임 율 (frame rate)이라고 한다. 예를 들어 1초에 24장의 프레임으로 구성되면 프레임 율은 24 fps (frame per second)가 된다.

[0005] 프레임 율은 촬영자의 의도, 영상의 포맷, 카메라의 한계 등에 의하여 결정된다. 관찰자가 영상을 연속된 화면으로 느끼기 위해서는 어느 정도 이상의 프레임 율이 필요하고 이보다 낮을 경우 움직임이 부드럽지 않아 보인다. 이 현상은 디스플레이의 크기, 조명, 시청 거리 등에 의해 달라질 수 있다.

[0006] 이를 개선하기 위해 동영상의 프레임 율을 후처리에 의해 증가시키는 것을 동영상 프레임 율 변환이라고 한다.

[0008] 2) 종래 기술

[0009] 동영상 프레임 율을 증가시키는 가장 간단한 방법은 프레임을 반복하는 것이다. 예를 들어 30 fps 영상을 60 fps 영상으로 증가시킬 경우, 각 프레임마다 한 프레임을 반복하여 출력하는 것이다. 그러나 이 방법의 경우 동영상의 정보량은 동일하고 움직임에 대한 연속성은 변하지 않았으므로 관찰자가 느끼는 불편감은 동일하다.

[0010] 이를 해결하기 위해 연속된 프레임들을 이용하여 가상의 프레임을 생성하는 기술이 개발되었다. 즉 t 초와 t+1초 사이의 영상을 이용하여 t+0.5 초의 중간 영상을 새롭게 생성하며 이를 프레임 보간 (frame interpolation) 기술이라고 한다.

[0011] 프레임 보간은 다양한 방법이 개발되었으며 일반적으로는 다음과 같은 두 단계 과정을 거친다. 첫 번째 단계는 움직임 또는 옵티컬 플로우 (optical flow)를 획득하는 단계 이며, 두 번째 단계는 움직임 정보를 바탕으로 중간 프레임을 생성 (Synthesis)하는 단계이다.

[0012] 동영상에서 물체의 움직임이 부드럽게 보이려면 중간 영상은 물체의 움직임이 두 영상 사이의 중간에 해당되어야 한다. 따라서 물체의 움직임 정보를 가지고 있는 옵티컬 플로우를 정확하게 찾는 것이 매우 중요하다. 이에

기반한 다양한 기법들이 제안되어 왔다.

[0013] 최근 딥러닝 (Deep learning) 알고리즘이 등장하여 컴퓨터 비젼, 음성 인식 등 다양한 분야에 널리 사용되고 있으며 종래에 방법에 비해 월등한 성능을 보이고 있다. 이에 발맞추어 딥러닝을 사용한 다양한 프레임 보간 기법이 등장하였다. 이 기법들은 옵티컬 플로우와 합성에 기반한 종래의 방법보다 더욱 뛰어난 보간 결과를 보여줌에 따라 최근 지속적으로 연구되고 있다.

[0015] 3) 종래 기술 문제점

[0016] 종래의 딥러닝에 기반한 기술은 입력 영상이 RGB 형식의 컬러 포맷을 갖는다고 가정한다. 그러나 대부분의 동영상은 YCbCr 형식의 컬러 포맷으로 압축되어 저장되어 있다. 또한 색차 신호인 Cb, Cr은 서브 샘플링 (sub-sampling) 되어 있어 밝기 신호인 Y와 해상도가 다르다.

[0017] 따라서 기존 방법에 일반적인 동영상을 적용하기 위해서는 색차 신호를 업샘플링 (up-sampling)하여 밝기 신호와 영상 크기를 맞춘 후, 다시 YCbCr 컬러 형식을 RGB로 변환하는 작업을 수행하여야 한다. 이런 컬러 변환 작업들은 비효율적이고 특히 영상 크기가 커질 경우 실시간 동작을 어렵게 하는 요소로 작동한다.

[0018] 기존 방법의 두 번째 문제는 네트워크 구조로 인한 문제로 4K (3840x2160) 해상도와 같이 큰 해상도에 대하여 GPU 메모리 부족으로 연산이 불가능 하거나, 아니면 매우 느린 연산 속도를 보여준다. 이와 같은 현상은 실시간 연산을 요구하는 상용 애플리케이션에는 딥러닝을 이용한 프레임 보간 방법 적용을 어렵게 한다.

발명의 내용

해결하려는 과제

[0020] 본 발명은 상기와 같은 문제점을 해결하기 위하여 안출된 것으로서, 본 발명의 목적은, 딥러닝에 기반하여 다양한 컬러 영상 포맷에 대응하는 동시에 고품질, 고속으로 프레임을 보간하는 동영상 프레임 율 변환 방법 및 장치를 제공함에 있다.

과제의 해결 수단

[0022] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른, 동영상 프레임 율 변환 장치는 각 채널의 영상에 대한 해상도를 조정하는 조정부; 조정부에서 해상도가 조정된 t초와 t+1초의 연속된 두 영상으로 중간 영상을 생성하는 생성부; 및 생성부에서 생성된 중간 영상의 해상도를 향상시키는 향상부;를 포함한다.

[0023] 조정부는, 각 채널의 영상에 대한 해상도를 동일하게 서브 샘플링하는 것일 수 있다.

[0024] 생성부는, 해상도가 조정된 영상의 특성을 계산하는 인코더부; 계산된 특성을 영상의 해상도 크기만큼 복원하는 디코더부; 조정부에서 해상도가 조정된 두 영상과 디코더의 출력을 이용하여 중간 영상을 생성하는 합성부;를 포함할 수 있다.

[0025] 향상부는, 생성부에서 생성된 중간 영상의 해상도를 원본 영상의 해상도로 향상시키는 것일 수 있다.

[0026] 향상부는, 디테일을 개선할 중간 영상의 해상도를 조정하는 제1 조정부; 원본 영상을 이용하여, 해상도가 조정된 중간 영상의 디테일을 개선하는 디테일 향상부; 제1 조정부에서 해상도가 조정되지 않은 중간 영상의 해상도를 조정하는 제2 조정부를 포함할 수 있다.

[0027] 디테일 향상부는, 선별적으로 선택된 채널만을 입력받는 것일 수 있다.

[0028] 선택된 채널은, 밝기 신호일 수 있다.

[0029] 디테일 향상부는, 생성부에서 생성되는 중간 정보를 참조하여, 중간 영상의 디테일을 개선하는 것일 수 있다.

[0030] 디테일 향상부는, 딥러닝 네트워크를 사용하는 것일 수 있다.

[0031] 한편, 본 발명의 다른 실시예에 따른, 동영상 프레임 율 변환 장치는 각 채널의 영상에 대한 해상도를 조정하는 단계; 해상도가 조정된 t초와 t+1초의 연속된 두 영상으로 중간 영상을 생성하는 단계; 및 생성된 중간 영상의 해상도를 향상시키는 단계;를 포함한다.

[0032] 한편, 본 발명의 다른 실시예에 따른, 동영상 프레임 율 변환 장치는 각 채널의 영상에 대한 해상도를 조정하는 조정부; 조정부에서 해상도가 조정된 t초와 t+1초의 연속된 두 영상으로 생성된 중간 영상의 해상도를 향상시키

는 향상부;를 포함한다.

[0033] 한편, 본 발명의 다른 실시예에 따른, 동영상 프레임 율 변환 장치는 각 채널의 영상에 대한 해상도를 조정하는 단계; 조정부에서 해상도가 조정된 t초와 t+1초의 연속된 두 영상으로 생성된 중간 영상의 해상도를 향상시키는 단계;를 포함한다.

발명의 효과

[0035] 이상 설명한 바와 같이, 본 발명의 실시예들에 따르면, 딥러닝 기반 동영상 프레임 율 변환에 있어, 딥러닝의 입력이 RGB 컬러 포맷 뿐만 아니라 YCbCr 등 다양한 컬러 포맷에 적용할 수 있게 되며, 프레임 보간 방법과 해상도 향상 방법을 동시에 이용하여 고품질, 고속의 프레임 율 변환이 가능해진다.

도면의 간단한 설명

[0037] 도 1 : 프레임 보간 기술

도 2 : 발명의 도면

도 3 : 색차 신호 서브 샘플링, Y: 밝기 신호, Cb, Cr: 색차 신호

도 4 : 본 발명의 대표 도면

도 5 : 중간 프레임 생성부와 해상도 향상부 개념도

도 6 : 입력 영상 조정부와 디테일 향상부의 입출력 해상도간의 관계

도 7 : YCbCr 4:2:0 형식에 대한 본 발명의 예시

도 8 : YCbCr 4:2:2 형식에 대한 본 발명의 예시 1

도 9 : YCbCr 4:2:2 형식에 대한 본 발명의 예시 2

도 10 : YCbCr 4:4:4 또는 RGB 형식에 대한 본 발명의 예시 1

도 11 : YCbCr 4:4:4 또는 RGB 형식에 대한 본 발명의 예시 2

발명을 실시하기 위한 구체적인 내용

[0038] 이하에서는 도면을 참조하여 본 발명을 보다 상세하게 설명한다.

[0039] 본 발명의 실시예에 따른 동영상 프레임 율 변환 장치는, 도 2에 도시된 바와 같이, 입력 영상 조정부(110), 중간 프레임 생성부(120), 해상도 향상부(130)를 포함하여 구성된다.

[0040] 대부분의 동영상은 RGB 형식에서 YCbCr 형식으로 변환되어 사용된다. YCbCr에서 Y는 밝기 신호, Cb와 Cr은 색차 신호라고 한다. 일반적으로 대부분 동영상의 YCbCr 영상의 색차 신호는 서브 샘플링이 되어 사용된다.

[0041] YCbCr의 대표적인 서브 샘플링 기법으로는 도 2와 같이 YCbCr4:4:4, YCbCr4:2:2, YCbCr4:2:0 방법이 있다. YCbCr4:4:4는 색차신호인 Cb 와 Cr 신호를 서브샘플링 하지 않는 영상 포맷을 의미하며 Cb, Cr의 해상도와 Y의 해상도가 같다. YCbCr 4:2:2의 경우 색차 신호는 가로 방향으로 2씩 서브 샘플링 되고, YCbCr 4:2:0에서 색차 신호는 가로, 세로 2씩 서브 샘플링되었다. 이와 같이 색차 신호와 밝기 신호의 해상도는 컬러 포맷에 의해 서로 상이하다.

[0042] Y, Cb, Cr의 각각을 하나의 컬러 채널이라고 하자. 일반적인 딥러닝 네트워크의 입력은 채널 간의 해상도가 동일해야 각각의 컬러 채널을 동시에 처리할 수 있다.

[0043] 그러나 위에서 언급하였듯이 컬러 포맷에 따라 컬러 채널 간 해상도가 상이하다. 컬러 채널 간 해상도를 맞추기 위해서 본 발명의 실시예에서는 입력 신호를 조정하는 방식을 제안한다.

[0044] 도 2에서, 입력 영상 조정부(110)는 딥러닝에 적용하기 전 각 채널의 영상의 해상도를 동일하게 조절하도록 한다. 즉 채널 별로 각각 서브 샘플링하여 동일한 해상도를 갖게 하도록 한다.

[0045] 구체적으로 설명하기 위해 입력 신호가 N개의 채널로 구성되어 있고, 각각의 채널에 대한 해상도가 $H_1 \times W_1$, ... ,$H_N \times W_N$라고 하자. 여기서 H는 영상의 세로 해상도를 W는 가로 해상도를 의미한다.

[0046] 입력 영상 조정부(110)는 모든 채널의 가로 해상도와 세로 해상도를 서브 샘플링을 통해 동일한 해상도 H_M x W_M으로 변형시킨다. H_M x W_M은 가장 큰 해상도를 갖는 채널의 해상도보다는 크지 않는 것을 권장한다. 왜냐하면 딥러닝 네트워크의 입력 해상도가 커질 경우 그에 비례해 연산량이 증가하기 때문이다.

[0047] 입력 영상 조정부(110)에서는 영상의 해상도를 조절함으로써 보간 프로세서의 연산량을 줄일 수 있다. 참고로 채널 별로 별도의 네트워크로 처리할 경우 밝기와 색차신호의 움직임이 다르게 예측되어 밝기 신호와 색차 신호가 합성된 신호는 불일치 생길 수 가 있다. 이와 같이 입력 영상 조정부(110)는 채널 별로 입력된 신호의 해상도를 조절하여 네트워크에 입력으로 사용될 영상을 만든다.

[0048] 본 발명의 실시예에서 제안하는 네트워크 구조는 도 5와 같이 크게 두 가지 단계로 이루어져 있다. 첫 번째는 중간 영상을 생성하는 네트워크로 중간 프레임 생성부로 지칭한다. 이 네트워크는 기존의 중간 영상을 생성하는 네트워크가 사용될 수 있다. 두 번째 네트워크는 생성된 중간 영상을 원본 해상도와 일치하게 조절하는 네트워크로 해상도 향상부(130)라고 지칭한다.

[0049] 도 4를 참조하여 본 발명의 실시예에 따른 동영상 프레임 율 변환 장치에 대하여 보다 상세히 설명한다.

[0050] 중간 프레임 생성부(120)는 t초와 t+1초의 연속된 두 프레임의 영상($\bar{I}_t$, $\bar{I}_{t+1}$)을 입력으로 받는다. 두 프레임은 입력 영상 조정부(110)를 거쳐 두 프레임 간의 모든 채널의 영상 해상도는 일치된 상태가 된다.

[0051] 중간 프레임 생성부(120)는 인코더부(121), 디코더부(122), 합성부(123)로 구성된다.

[0052] 인코더부(121)는 컨볼루셔널 레이어 (convolutioal layer), 풀링 레이어 (Pooling layer), 활성화 레이어 (Activation layer) 등으로 구성되어 있고 이에 한정하지 않는다.

[0053] 인코더부(121)의 출력은 디코부(122)의 입력이 되고 디코더부(122)에서는 디컨볼루셔널 레이어 (deconvolutional layer) 및 활성화 레이어 (activation layer) 등으로 구성되어 있으며 이에 한정하지 않는다.

[0054] 디코더부(122)의 출력과 네트워크의 입력이 합성부(123)의 입력으로 들어가 1차 중간 (보간) 영상을 생성한다. 1차 중간 영상의 해상도는 중간 프레임 생성부(120)의 입력의 해상도와 동일하다. 디코더부(122)에서는 인코더부(121)의 출력뿐만 아니라 인코더부(121)의 중간 결과도 입력으로 받을 수 있고, 이를 이용해 성능을 개선할 수 있다.

[0055] 도 5는 인코더부(121), 디코더부(122), 합성부(123)에 대한 개념을 보여준다.

[0056] 인코더부(121)에서는 영상의 해상도가 압축되어 영상의 특성 (feature)을 계산하고, 디코더부(122)에서는 특성을 영상의 해상도 크기만큼 복원한다. 인코더부(121)의 중간 정보는 디코더부(122)에 직접적으로 전달되고 이는 도 5에서 화살표로 표현된다.

[0057] 또한 인코더부(121)와 디코더부(122)의 중간 정보도 디테일 향상부(132)에 전달되어 재활용되어 진다.

[0058] 해상도 향상부(130)는 해상도가 감소된 채널 신호에 대하여 해상도 증가를 실시한다. 중간 프레임 생성부(120)에 생성된 1차 중간 프레임의 모든 채널의 해상도는 H_M x W_M이다. 원본 영상의 해상도는 H_1 x W_1 , ... ,H_N x W_N이므로, 해상도 향상부(130)에서는 원본 영상의 해상도와 일치하게 영상의 크기를 조절한다. 1차 중간 프레임의 해상도는 원본 해상도에 비해 작은 것이 권장되므로 해상도 향상부(130)는 주로 해상도 증가 작업이 수행된다.

[0059] 해상도 향상부(130)는 중간 영상 조정부(131), 디테일 향상부(132), 출력 영상 조정부(133)로 구성되어 있고, 이 중 디테일 향상부(132)만 딥러닝 네트워크를 사용하여 많은 연산량이 필요로 한다.

[0060] 따라서 디테일 향상부(132)에 사용되는 입력은 모든 채널이 아닌 선별적으로 선택된 채널만 사용될 수 있다. 일반적으로는 밝기 신호의 대부분의 디테일이 집중되어 있으므로 디테일 향상부(132)의 입력으로는 밝기 신호만을 사용하는 것을 권장한다.

[0061] 중간 영상 조정부(131)에서는 디테일 향상부(132)에 사용될 입력의 해상도를 조절한다. 도 6은 입력 영상 조정부(110)와 디테일 향상부(132)의 입출력 해상도간의 관계를 보여준다. 원본 영상의 n 번째 채널의 해상도는 H_n x W_n이라고 하고 n 번째 채널을 디테일 향상부(132)에 적용하는 상태이다.

[0062] 디테일 향상부(132)의 입출력의 해상도는 원본 해상도와 동일해야 하고, 이를 위해 입력 영상 조정부(110)에서 1차 중간 프레임의 해상도를 조절하여 영상 크기 원본 해상도와 동일하게 한다. 조절하는 방법은 bilinear, bicubic 등의 필터를 사용할 수 있고 특정 방법에 제한되지 않는다.

[0063] 디테일 향상부(132)의 입력은 1차 중간 영상과 원본 입력 영상 I_t 와 I_{t+1} 이다. 원본 입력 영상은 비록 다른 시간의 영상들이지만 원본 해상도의 디테일들을 보존하고 있다. 따라서 디테일 향상부(132)에는 이 정보를 사용하여 1차 중간 영상의 디테일을 개선하도록 한다.

[0064] 이외에도 해상도 향상부(130)의 입력으로 중간 프레임 생성부(120)의 중간 결과들을 입력으로 받는다. 중간 프레임 생성부(120)의 중간 과정를 디테일 향상부(132)에서는 입력으로 받는다. 이 입력은 영상의 특징 (feature)을 포함하고 있으며 이 특징은 디테일 향상부(132)에 사용됨으로써 더욱 높은 성능을 제공한다.

[0065] 출력 영상 조정부(133)에서는 최종적으로 보간 프레임 각각의 채널 해상도가 원본의 채널 해상도와 동일하게 조절해 준다. 디테일 향상부(132)에서 해상도가 향상된 채널은 원본 영상의 해상도와 동일하겠지만, 다른 채널의 영상은 원본 해상도와 다를 수 있다. 다시 말하면 원본 해상도와 다른 보간 영상의 채널에 대해서만 출력 영상 조정부(133)에서 해상도 조절이 수행된다.

[0066] 본 발명의 실시예에서, 네트워크의 학습(training)은 중간 프레임 생성부(120)를 먼저 학습한 후, 디테일 향상부(132)를 학습하는 것을 권장한다. 또는 위와 같이 학습을 수행한 후, 엔드 투 엔드 (end-to-end)로 전체 네트워크를 한 번에 학습하는 것을 권장한다. 다른 방법으로는 중간 프레임 생성부(120)를 학습한 후, 중간 프레임과 디테일 향상부(132)를 동시에 학습하는 것을 권장한다.

[0068] 다음은 본 발명의 다양한 구형예들을 보여준다.

[0069] 전체 네트워크의 입력은 동영상에서 연속된 두 프레임이며, 이를 I_t 와 I_{t+1} 이라고 하자. 전체 네트워크의 출력은 $I_{t'}$ 이며 t' 은 0과 1사이의 값이다. 즉 출력 영상은 t 와 t+1 사이의 영상이 된다. 여기서 t는 시간을 의미한다. 입력 영상이 YCbCr 컬러 포맷을 갖는다고 하면 입력 영상 I_t 와 I_{t+1}는 $I_t = \{Y_t, Cb_t, Cr_t\}$ 와 $I_{t+1} = \{Y_{t+1}, Cb_{t+1}, Cr_{t+1}\}$로 표현할 수 있다. Y_t와 Y_{t+1}은 밝기 신호를, $Cb_t, Cr_t, Cb_{t+1}, Cr_{t+1}$ 은 색차 신호를 의미한다. 입력 영상 조정부(110)의 입력은 I_t 와 I_{t+1}이 되며 출력은 $\overline{I}_t = \{\overline{Y}_t, \overline{Cb}_t, \overline{Cr}_t\}$ 와 $\overline{I}_{t+1} = \{\overline{Y}_{t+1}, \overline{Cb}_{t+1}, \overline{Cr}_{t+1}\}$ 이 된다. $\overline{X}$ 은 서브 샘플링된 영상을 의미하며 $\overline{X}$사이의 해상도는 모두 같아야 한다. 다시 말하면 채널 간 해상도는 같아야 한다.

[0070] 입력 영상 조정부(110)에 대한 실시예를 YCbCr 4:2:0 컬러 형식을 갖는 동영상에 대하여 도 7과 같이 살펴보자. 밝기 신호 Y의 해상도가 H x W이면 색차 신호 Cb, Cr은 H/2 x W/2 의 해상도를 갖는다. 입력 영상 조정부(110)에서 영상 신호의 출력 해상도는 H/2 x W/2로 설정한다. 이와 같을 경우 Y의 신호만 가로, 세로 2배씩 서브 샘플링하여 H/2 x W/2로 생성한다. $\overline{Y}_t, \overline{Cb}_t, \overline{Cr}_t$과 $\overline{Y}_{t+1}, \overline{Cb}_{t+1}, \overline{Cr}_{t+1}$은 모두 H/2 x W/2의 해상도를 갖는다. 출력 영상의 해상도는 갈 컬러 채널 중 가장 작은 해상도를 갖는 채널보다 작거나 같아야 한다. 즉, YCbCr 4:2:0영상의 경우에는 출력 해상도는 H/2 x W/2 보다 같거나 작아야 한다. 다른 실시예로 YCbCr 4:2:2 영상에 대하여 살펴보자. Y의 해상도를 H x W라고 하면 Cb, Cr의 해상도는 H x W/2이다. 그럼 입력 영상 조정부(110)의 출력 해상도는 H x W/2 이거나 그 이하여야 한다. 따라서 사용자의 설정에 따라 $\overline{Y}_t, \overline{Cb}_t, \overline{Cr}_t$과 $\overline{Y}_{t+1}, \overline{Cb}_{t+1}, \overline{Cr}_{t+1}$의 해상도는 H x W/2 가 될 수 도 있고, H/2 x W/2 등도 될 수 있다.

[0071] 중간 프레임 생성부(120)의 입력은 $\overline{I}_t = \{\overline{Y}_t, \overline{Cb}_t, \overline{Cr}_t\}$ 과 $\overline{I}_{t+1} = \{\overline{Y}_{t+1}, \overline{Cb}_{t+1}, \overline{Cr}_{t+1}\}$이다. 도 7의 YCbCr 4:2:0 포맷의 경우에는 $\overline{Cb}_{t+1}, \overline{Cr}_{t+1}$ 과 Cb_{t+1}, Cr_{t+1}이 같고 t 프레임도 마찬가지다. 중간 프레임 생성부의 출력은 $\hat{I}_{t'} = \{\hat{Y}_{t'}, \hat{Cb}_{t'}, \hat{Cr}_{t'}\}$이 된다. $\hat{I}_{t'}$는 t와 t+1 프레임 사이에서 보간된 중간 프레임이며 $\overline{I}_t$ 와 $\hat{I}_{t'}$ 의 해상도는 동일하다. 중간 프레임 생성부(120)는 인코더부, 디코더부, 합성부로 구성되나 이에 한정되지 않는다. 기존의 프레임 보간 방법을 사용하여도 무관하다. 인코더부는 콘볼루셔널 레이어, 풀링 레이어, 활성 레이어로 구성되어 있다. 아래 도 처럼 입력 영상에 대하여 컨볼루션을 수행하고 풀링 레이어를 수행하고 활성 레이어를 수행한다. 풀링 레이어는 영상의 해상도를 감소시킨다. 일반적으로 가로, 세로 각각 2배식 감소시킨다. 컨볼루셔널

레이어의 출력을 C_L, 풀링 레이어의 출력을 P_L ,활성 레이어의 출력을 A_L이라고 한다. L은 레이어의 순서를 의미하며 이 각각의 레이어의 출력은 해상도 증가부의 입력이 된다. 디코더부는 인코더부에서 줄어든 해상도를 다시 증가시키면서 특징을 추출한다. 해상도 증가는 Bilinear 필터 또는 디콘볼루션 레이어 등 다양한 방법을 사용할 수 있으며 특정 방법에 제한되지 않는다. 각각의 업샘플링 레이어의 출력을 U_L 이라고 정의하며 L은 레이어의 순서를 의미한다. 각각의 업샘플링 레이어의 출력 U_L은 해상도 향상부(130)의 입력으로 사용된다. 디코더부의 최종 출력 U_F 와 $\overline{I_t}=\{\overline{Y_t},\overline{Cb_t},\overline{Cr_t}\}$, $\overline{I_{t+1}}=\{\overline{Y_{t+1}},\overline{Cb_{t+1}},\overline{Cr_{t+1}}\}$ 은 합성부의 입력이 된다. 합성은 콘볼루션 필터, 와핑 (Warping) 등 다양한 방법이 사용될 수 있으며 본 발명에서는 특정 방법에 제한되지 않는다. 합성부의 출력은 $\hat{I}_{t'}=\{\hat{Y}_{t'},\widehat{Cb}_{t'},\widehat{Cr}_{t'}\}$이 되며 이를 1차 중간 프레임으로 부른다.

[0072] 1차 중간 프레임의 해상도는 원영상에 비해서 작은 해상도를 갖는다. 따라서 해상도 향상부(130)에서는 1차 중간 프레임의 해상도를 증가시켜 원영상과 같은 해상도를 갖게 하는 동작을 수행한다. 해상도 향상부(130)의 입력은 해상도 향상부(130)의 입출력은 다양하게 정의될 수 있으며 다음에서 몇 가지 실시 예에 대하여 살펴본다.

[0073] 먼저 원본 입력 영상이 YCbCr 4:2:0 컬러 포맷이며 Y의 해상도는 H x W, Cb, Cr의 해상도는 H/2 X W/2이라고 하자. 입력 영상 조정부(110)에서 Y의 해상도를 서브샘플링하여 Cb, Cr의 해상도인 H/2 x W/2 와 일치시킨다. 따라서 중간 프레임 생성부(120)의 해상도는 Y, Cb, Cr의 경우 H/2 x W/2 가 되며, Y의 경우 원영상에 비해 해상도가 줄었지만 Cb, Cr의 경우 해상도가 변하지 않았으므로 Y 영상만 해상도 향상부(130)를 사용하여 해상도를 증가시킨다. 해상도 향상부(130)는 중간 영상 조정부(131)와 디테일 향상부(132)의 두 단계로 구성된다. 중간 영상 조정부(131)는 1차 중간 프레임을 업샘플링하여 원 영상의 크기와 동일하게 만든다.

[0074] 위의 경우 중간 영상 조정부(131)의 입력은 H/2 x W/2 의 해상도를 갖는 1차 중간 프레임 $\hat{Y}_{t'}$이고 출력은 H x W 해상도를 갖는 영상 $\check{Y}_{t'}$이다. 디테일 향상부(132)의 입력은 $\check{Y}_{t'}$와 원영상 Y_t, Y_{t+1} 이다. $\check{Y}_{t'}$의 디테일은 고해상도 디테일을 가지고 있는 Y_t, Y_{t+1}의 협력을 받아 더욱 정교하게 복원될 수 있다. 기존의 다중 프레임 (multi-frame) 해상도 향상 방법은 여러 장의 저해상도 영상을 사용하여 고해상도 영상을 사용하였다. 그러나 본 발명에서는 저해상도 영상과 고해상도 영상을 동시에 이용하여 해상도 향상을 꾀하였다. 디테일 향상부(132)는 딥러닝 네트워크로 구성되어 있으며 출력은 $\widetilde{Y}_{t'}$이다. 디테일 향상부(132)를 위한 딥러닝 네트워크는 컨벌루셔널 레이어, 풀링 레이어, 활성 레이어 등으로 구성되어 있으며, 본 논문에서는 특정 네트워크 형태로 제한하지 않는다. 디테일 향상부(132)는 중간 프레임 생성부(120)의 중간 결과물을 입력으로 받아 디테일 향상부(132)의 중간 과정에 더하거나 적층하여 (concatenation) 성능을 개선할 수 있다. 중간 생성부의 중간 결과물과 디테일 향상부(132)의 중간 결과물의 해상도는 일치해야한다. $\widetilde{Y}_{t'}$와 1차 중간 프레임의 색차 신호 $\widehat{Cb}_{t'}$, $\widehat{Cr}_{t'}$ 이 최종 결과물이 된다.

[0075] 원본 입력 영상이 YCbCr 4:2:2 컬러 포맷인 실시예에 대하여 살펴보자. Y의 해상도는 H x W, Cb, Cr의 해상도는 H x W/2 라고 하자. 도 8과 같이 입력 영상 조정부(110)에서 Y의 해상도를 서브샘플링하여 Cb, Cr의 해상도인 H x W/2 와 일치시킨다. 따라서 중간 프레임 생성부(120)의 해상도는 Y, Cb, Cr의 경우 H x W/2 가 되며, Y의 경우 원영상에 비해 해상도가 줄었지만 Cb, Cr의 경우 해상도가 변하지 않았으므로 Y 영상만 해상도 향상부(130)를 사용하여 해상도를 증가시킨다. 해상도 향상부(130)의 작동 방식은 위의 YCbCr 4:2:0 형식일 때와 동일하다. H x W의 해상도를 가지는 $\widetilde{Y}_{t'}$와 1차 H x W/2의 해상도를 가지는 중간 프레임의 색차 신호 $\widehat{Cb}_{t'}$, $\widehat{Cr}_{t'}$이 최종 결물이 된다.

[0076] YCbCr 4:2:2 컬러 포맷의 다른 실행 예시는 도 9와 같다. Y의 해상도는 H x W, Cb, Cr의 해상도는 H x W/2 라고 하자. 입력 조정부에서 Y, Cb, Cr의 해상도를 모두 H/2 x W/2로 서브 샘플링한다. 그리고 위의 YCbCr 4:2:0과 동일한 방식으로 처리해준다. 이와 같은 과정을 거칠 경우 디테일 향상부(132)의 출력은 Y신호는 H x W의 해상도를 갖으며, Cb, Cr 신호는 H/2 x W/2 의 해상도를 갖는다. Cb, Cr의 신호가 원본 신호와 다르기 때문에 출력 영상 조정부(133)에서 Cb, Cr의 해상도를 원본 신호와 동일하게 조정하여 준다. 조정 방법은 Bicubic 필터 등이 사용될 수 있으며 특정 방법에 제한 받지 않는다. 이와 같이 할 경우 위의 방법에 비해 네트워크에서 처리되는 입력 영상 크기가 작으므로 연산량과 메모리를 절약할 수 있다는 장점이 있다. 그리고 YCbCr 4:2:0에서 사

용되었던 네트워크를 공유할 수 있다는 장점이 있다.

[0077] YCbCr 4:4:4나 RGB 입력 영상이 입력인 경우를 고려하여 보자. YCbCr 4:4:4나 RGB 컬러 형식의 동영상은 모든 채널의 해상도가 동일함으로 입력 영상 조정부(110)와 해상도 향상부(130)가 사용되지 않아도 중간 프레임이 생성될 수 있다. 그러나 연산량을 감소시키기 위해서 본 발명에서는 입력 영상 조정부(110)와 해상도 향상부(130)를 사용할 수 있다. 예를 들어 입력 영상이 RGB 형식이고 각 채널의 해상도는 H x W라고 가정하자. 입력 영상 조정부(110)에서 각 채널을 H/2 x W/2 의 크기로 서브샘플링 한 후 중간 프레임 생성부(120)의 입력으로 넣는다. 그리고 H/2 x W/2의 해상도의 1차 중간 프레임은 해상도 향상부(130)에서 H x W의 해상도로 복원된다. 전체 프로세스에서 속도의 이점은 중간 프레임 생성부(120)의 연산량이 해상도 향상부(130)의 연산량보다 높다는 데에서 온다. 따라서 연산량이 높은 중간 프레임 생성부(120)는 작은 해상도로 처리하고 이를 다시 해상도 향상부(130)에서 복원하여 연산량을 줄일 수 있도록 한다. YCbCr 4:4:4 의 경우도 유사한 방식으로 적용할 수 있다. 각 채널의 해상도가 H x W이면 입력 영상 조절부에서 H/2 x W/2로 변환한다. 그리고 중간 프레임 생성부(120)에서 H/2 x W/2의 Y, Cb, Cr 결과를 생성한다. 1차 중간 프레임은 해상도 향상부(130)에 2가지 방법으로 적용될 수 있다. 첫 번째 방법은 위에서 설명한 RGB 영상처럼 도 10과 같이 Y, Cb, Cr 모든 채널에 대하여 디테일 향상부(132)의 딥러닝 네트워크를 통과시키는 것이다. 두 번째 방법은 위에서 설명한 YCbCr 4:2:2 방식처럼 도 11과 같이 Y 영역만 딥러닝 네트워크를 통과하여 고화질 영상으로 복원하고 Cb, Cr은 출력영상 조정부에서 단순 업스케일링을 통해 복원하는 것이다. 이것은 Cb, Cr은 일반적으로 디테일 영역이 적기 때문에 단순 업스케일링을 통해서도 복원이 가능하다. 이와 같은 방법은 전체 채널 영상에 대하여 딥러닝 네트워크를 사용하는 것에 비해 더욱 적은 연산량으로 복원이 가능하다. 또한 YCbCr4:2:0 네트워크와 공통으로 사용할 수 있다는 장점이 있어 범용성이 높다는 장점이 있다.

[0078] 한편, 본 실시예에 따른 장치와 방법의 기능을 수행하게 하는 컴퓨터 프로그램을 수록한 컴퓨터로 읽을 수 있는 기록매체에도 본 발명의 기술적 사상이 적용될 수 있음은 물론이다. 또한, 본 발명의 다양한 실시예에 따른 기술적 사상은 컴퓨터로 읽을 수 있는 기록매체에 기록된 컴퓨터로 읽을 수 있는 코드 형태로 구현될 수도 있다. 컴퓨터로 읽을 수 있는 기록매체는 컴퓨터에 의해 읽을 수 있고 데이터를 저장할 수 있는 어떤 데이터 저장 장치이더라도 가능하다. 예를 들어, 컴퓨터로 읽을 수 있는 기록매체는 ROM, RAM, CD-ROM, 자기 테이프, 플로피 디스크, 광디스크, 하드 디스크 드라이브, 등이 될 수 있음은 물론이다. 또한, 컴퓨터로 읽을 수 있는 기록매체에 저장된 컴퓨터로 읽을 수 있는 코드 또는 프로그램은 컴퓨터간에 연결된 네트워크를 통해 전송될 수도 있다.

[0079] 또한, 이상에서는 본 발명의 바람직한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

부호의 설명

[0081] 110 : 입력 영상 조정부

120 : 중간 프레임 생성부

121 : 인코더부

122 : 디코더부

123 : 합성부

130 : 해상도 향상부

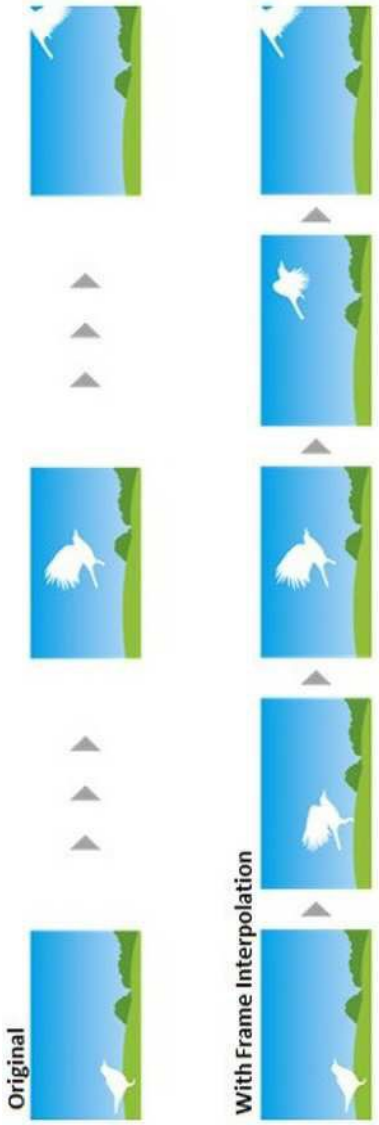
131 : 중간 영상 조정부

132 : 디테일 향상부

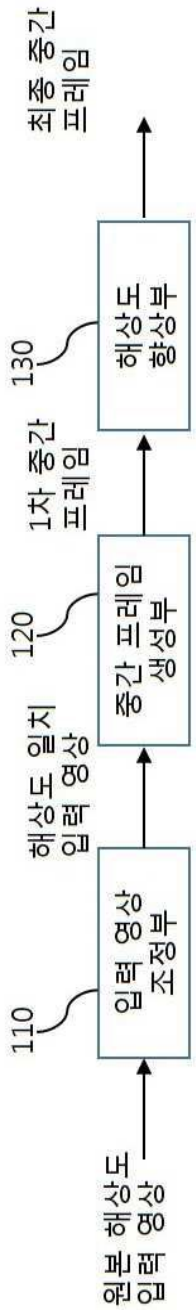
133 : 출력 영상 조정부

도면

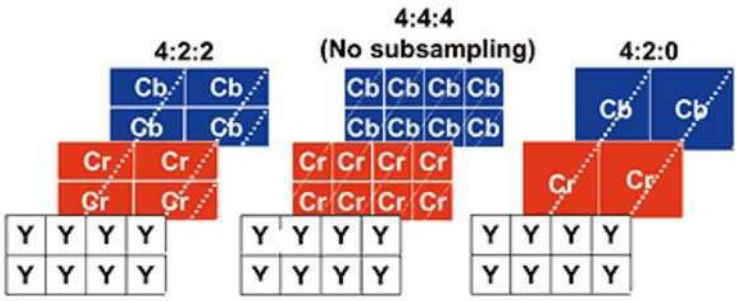
도면1



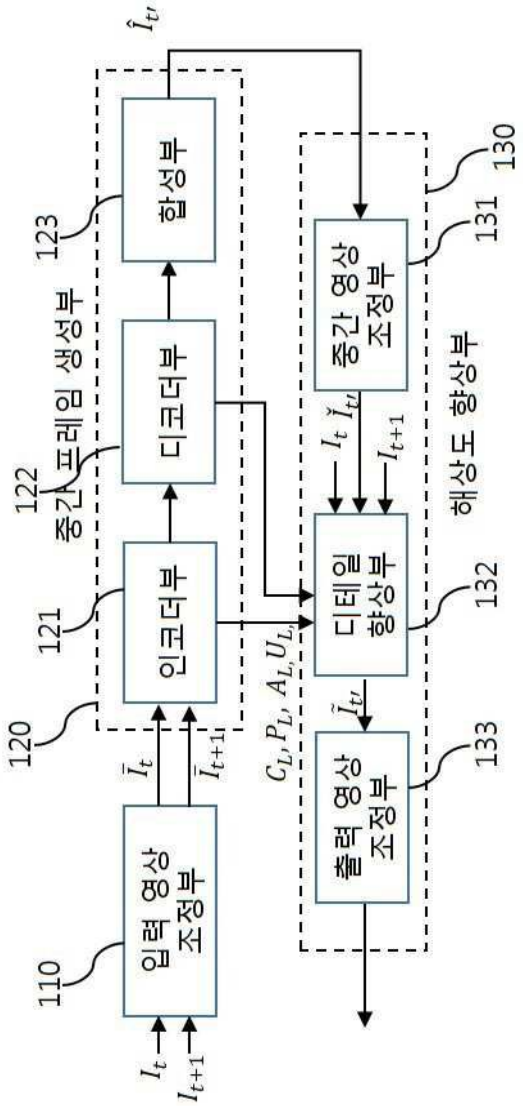
도면2


원본 해상도 입력 영상
110
입력 영상 조정부
해상도 일치 입력 영상
120
중간 프레임 생성부
1차 중간 프레임
130
해상도 향상부
최종 중간 프레임

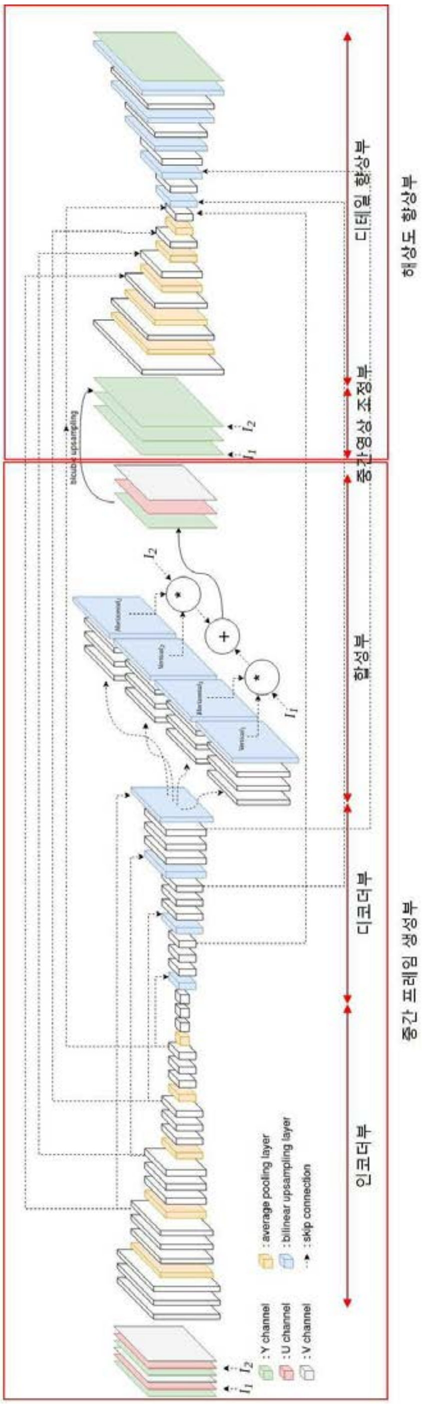
도면3


4:2:2
4:4:4
(No subsampling)
4:2:0
Cb
Cr
Y

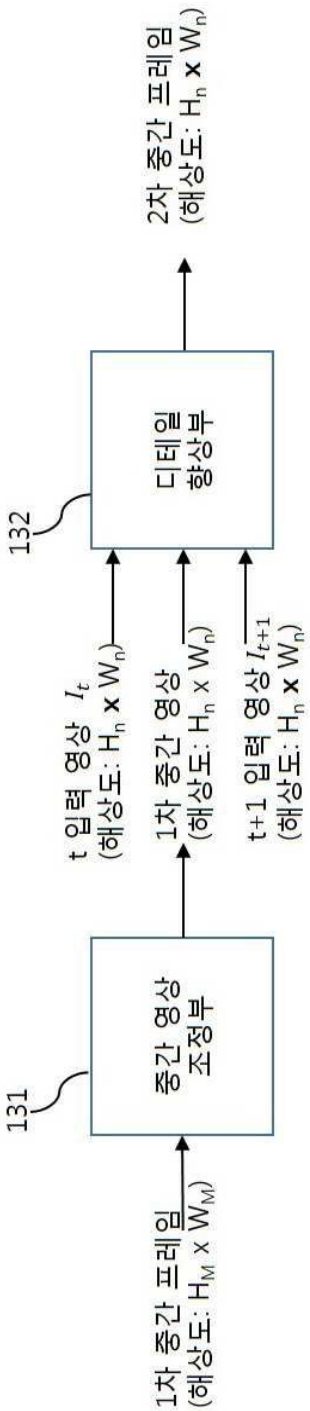
도면4


입력 영상 조정부
인코더부
디코더부
합성부
중간 프레임 생성부
중간 영상 조정부
디테일 향상부
출력 영상 조정부
해상도 향상부
110
120
121
122
123
130
131
132
133

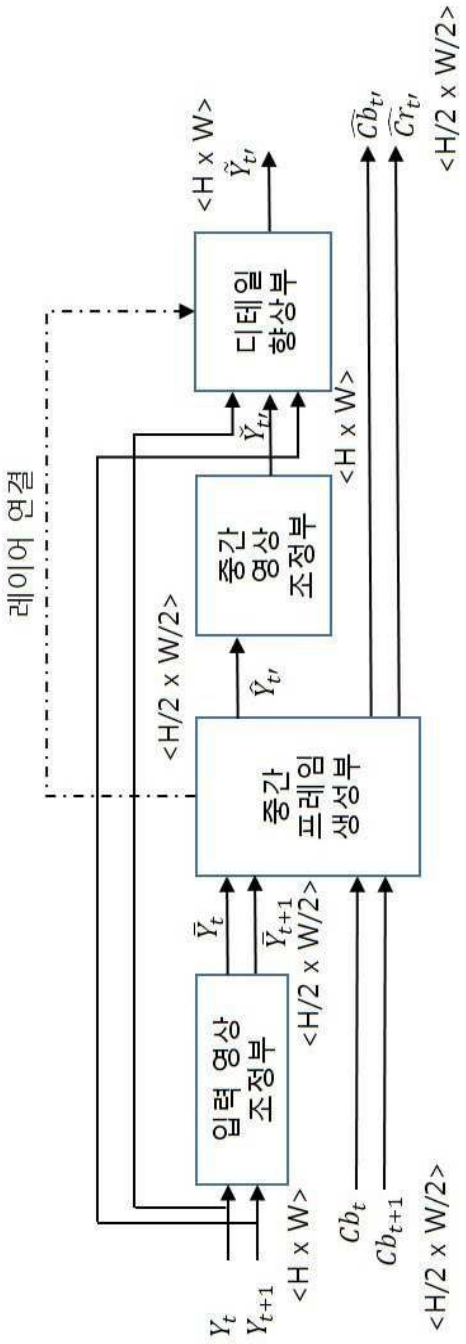
도면5


: Y channel
: U channel
: V channel
: average pooling layer
: bilinear upsampling layer
: skip connection
bicubic upsampling
인코더부
디코더부
합성부
중간영상 조정부
디테일 향상부
중간 프레임 생성부
해상도 향상부

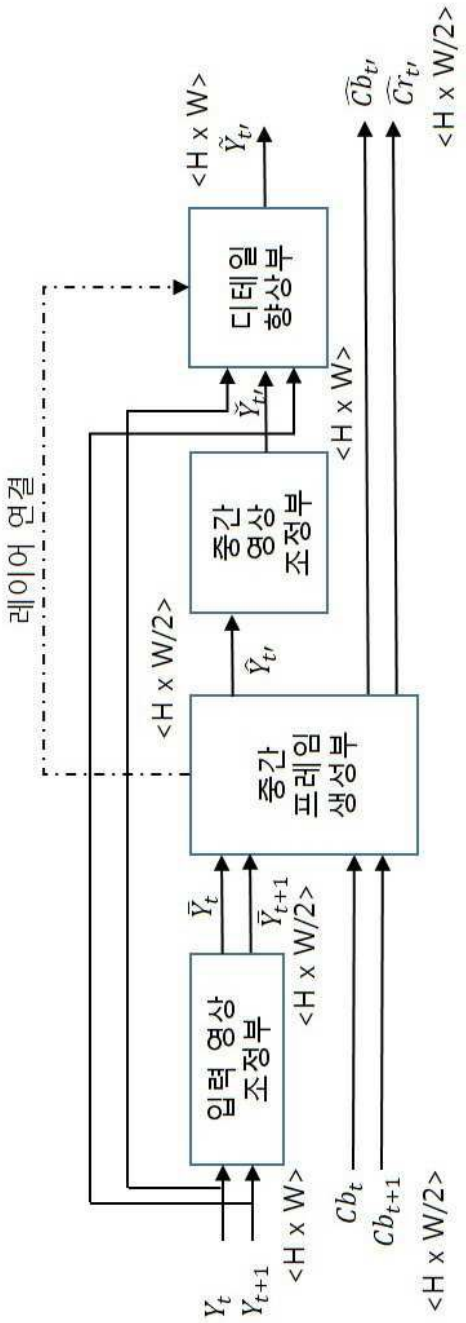
도면6


2차 중간 프레임
(해상도: H_n x W_n)
디테일
향상부
132
t 입력 영상 I_t
(해상도: H_n x W_n)
1차 중간 영상
(해상도: H_n x W_n)
t+1 입력 영상 I_{t+1}
(해상도: H_n x W_n)
중간 영상
조정부
131
1차 중간 프레임
(해상도: H_M x W_M)

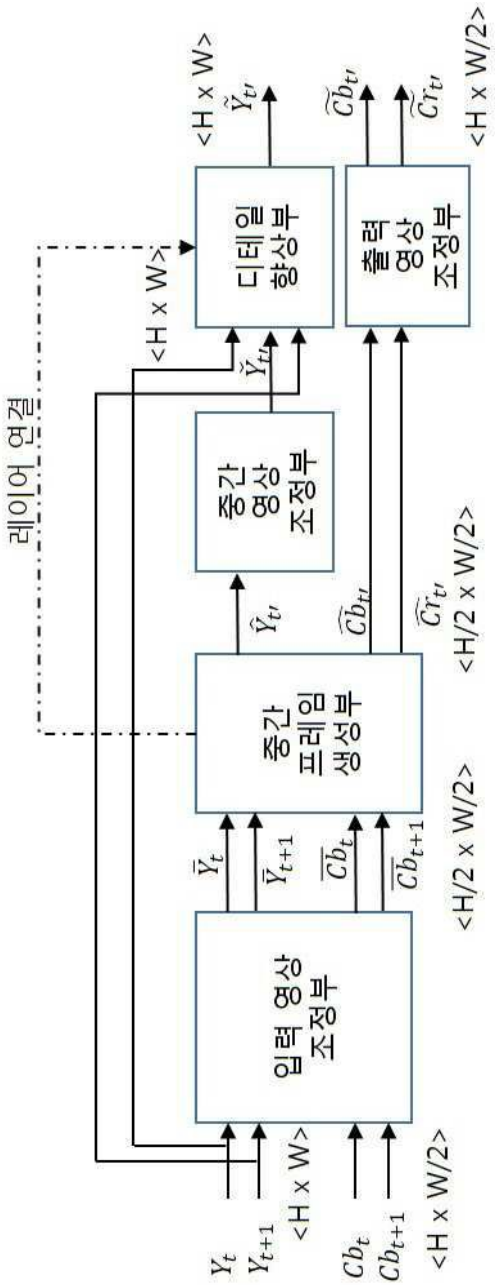
도면7


레이어 연결
입력 영상 조정부
중간 프레임 생성부
중간 영상 조정부
디테일 향상부
<H x W>
<H/2 x W/2>
Y_t
Y_{t+1}
$\bar{Y}_t$
$\bar{Y}_{t+1}$
Cb_t
Cb_{t+1}
$\hat{Y}_{t'}$
$\tilde{Y}_{t'}$
$\widetilde{Cb}_{t'}$
$\widetilde{Cr}_{t'}$

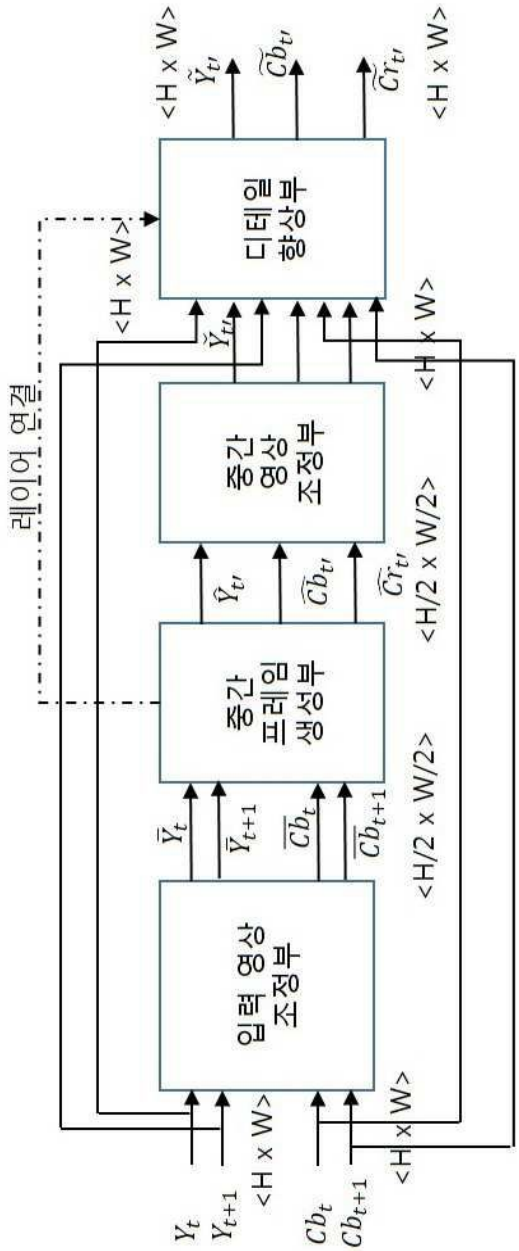
도면8


입력 영상 조정부
중간 프레임 생성부
중간 영상 조정부
디테일 향상부
레이어 연결
<H x W>
<H x W/2>
Y_t
Y_{t+1}
$\bar{Y}_t$
$\bar{Y}_{t+1}$
$\hat{Y}_t$
$\tilde{Y}_t$
$\tilde{Y}_{t'}$
Cb_t
Cb_{t+1}
$\widetilde{Cb}_{t'}$
$\widetilde{Cr}_{t'}$

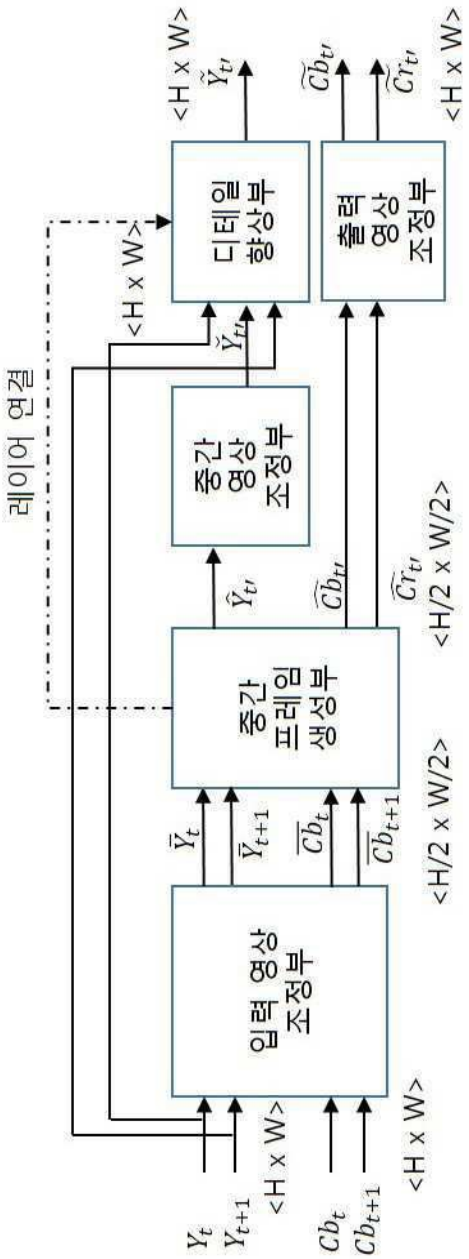
도면9


입력 영상 조정부
중간 프레임 생성부
중간 영상 조정부
디테일 향상부
출력 영상 조정부
레이어 연결
<H x W>
<H x W/2>
<H/2 x W/2>

도면10



도면11



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년04월20일
(11) 등록번호 10-2242343
(24) 등록일자 2021년04월14일

(51) 국제특허분류(Int. Cl.)
H04N 7/01 (2006.01) *H04N 21/2343* (2011.01)
H04N 21/4402 (2011.01)
(52) CPC특허분류
H04N 7/0127 (2013.01)
H04N 21/234381 (2013.01)
(21) 출원번호 10-2019-0137132
(22) 출원일자 2019년10월31일
심사청구일자 2020년02월20일
(56) 선행기술조사문헌
A Fast 4K Video Frame Interpolation Using a Hybrid Task-Based Convolutional Neural Network.*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
한국전자기술연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
정진우
서울특별시 송파구 송파대로 111, 파크하비오105동 1205호
안하은
서울특별시 성북구 성북로8다길 37, 102호
김제우
경기도 성남시 분당구 수내로 181, 303동 901호
(74) 대리인
남충우

전체 청구항 수 : 총 12 항 심사관 : 박재학

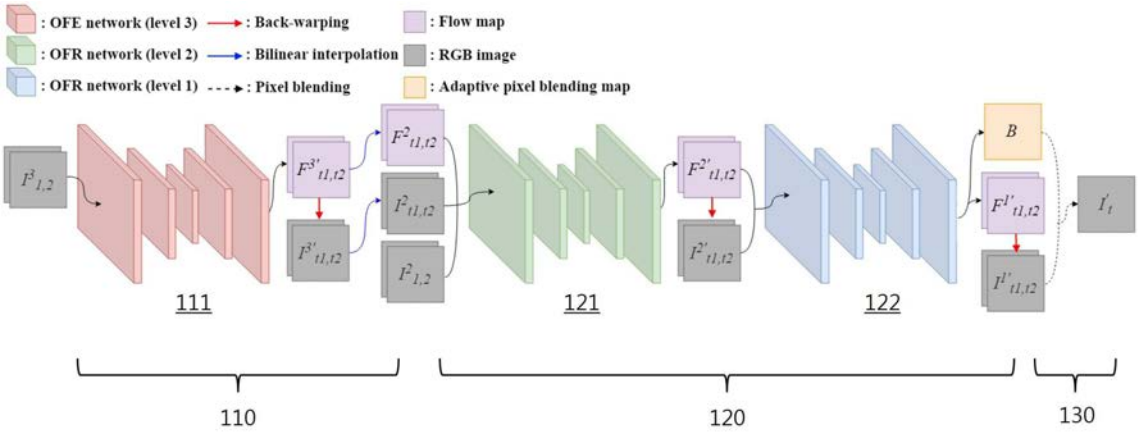
(54) 발명의 명칭 **고해상도 동영상 프레임 율 고속 변환 방법 및 장치**

(57) 요 약

고해상도 비디오 영상에 대하여 고품질, 고속으로 프레임 보간을 수행하기 위한 방안으로, 입력 고해상도 영상을 저해상도 영상으로 변환하여 고속으로 광학 흐름 지도를 생성하고 이를 원본 고해상도로 복원하여 고해상도 영상을 고속으로 보간하는 동영상 프레임 율 고속 변환 방법 및 장치가 제공된다. 본 발명의 실시예에 따른, 동영상

(뒷면에 계속)

대 표 도 - 도4



프레임 율 변환 방법은 시간적으로 연속된 고해상도의 프레임들로부터 생성한 저해상도의 프레임들들로 광학 흐름을 예측하여, 저해상도의 광학 흐름 지도들을 생성하는 제1 생성단계; 저해상도의 광학 흐름 지도들의 해상도를 단계적으로 높이면서, 고해상도의 중간 프레임들을 생성하는 제2 생성단계; 생성된 고해상도의 중간 프레임들을 블렌딩하여, 최종 보간 프레임을 생성하는 제3 생성단계:를 포함한다.

이에 의해, 입력 고해상도 영상을 저해상도 영상으로 변환하여 고속으로 광학 흐름 지도를 생성하고 이를 원본 고해상도로 복원하여 고해상도 영상을 보간함으로써, 4K와 같은 고해상도 비디오 영상에 대하여 실시간성을 요구하는 시스템 환경에서도 고품질, 고속으로 프레임 보간을 수행할 수 있게 된다.

(52) CPC특허분류
H04N 21/440281 (2013.01)
H04N 7/0137 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711081116
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	차세대(UHD)방송서비스활성화기술개발(R&D)
연구과제명	영상콘텐츠 초고속/초고화질 변환 기술 개발
기 여 율	1/1
과제수행기관명	(주)픽스트리
연구기간	2019.01.01 ~ 2019.12.31

공지예외적용 : 있음

명 세 서

청구범위

청구항 1

시간적으로 연속된 고해상도의 프레임들로부터 생성한 저해상도의 프레임들들로 광학 흐름을 예측하여, 저해상도의 광학 흐름 지도들을 생성하는 제1 생성단계;

저해상도의 광학 흐름 지도들의 해상도를 단계적으로 높이면서, 고해상도의 중간 프레임들을 생성하는 제2 생성단계;

생성된 고해상도의 중간 프레임들을 블렌딩하여, 최종 보간 프레임을 생성하는 제3 생성단계:를 포함하는 것을 특징으로 하는 동영상 프레임 율 변환 방법.

청구항 2

청구항 1에 있어서,

제1 생성단계는,

입력되는 시간적으로 연속된 프레임들의 해상도인 제1 해상도 보다 낮은 제3 해상도의 프레임들로 광학 흐름을 예측하여, 제3 해상도의 광학 흐름 지도들을 생성하는 단계;

제3 해상도의 광학 흐름 지도들을 이용하여, 제3 해상도의 중간 프레임들을 생성하는 단계;

제3 해상도의 중간 프레임들과 광학 흐름 지도들로, 제3 해상도 보다 높은 제2 해상도의 중간 프레임들과 광학 흐름 지도들을 복원하는 단계;를 포함하고,

제2 생성단계는,

제2 해상도의 중간 프레임들과 광학 흐름 지도들을 이용하여, 제2 해상도 보다 높은 제1 해상도의 중간 프레임들을 생성하는 제2 생성단계;를 포함하는 것을 특징으로 하는 동영상 프레임 율 변환 방법.

청구항 3

청구항 2에 있어서,

제2 생성단계는,

복원된 제2 해상도의 중간 프레임들과 광학 흐름 지도들로 광학 흐름을 예측하여, 제2 해상도의 광학 흐름 지도들을 생성하는 단계;

제2 해상도의 광학 흐름 지도들을 이용하여, 제2 해상도의 중간 프레임들을 생성하는 단계;

생성된 제2 해상도의 중간 프레임들과 광학 흐름 지도들로 광학 흐름을 예측하여, 제1 해상도의 광학 흐름 지도들을 생성하는 단계;

제1 해상도의 광학 흐름 지도들을 이용하여, 제1 해상도의 중간 프레임들을 생성하는 단계;를 포함하고,

제3 생성단계는,

생성된 제1 해상도의 중간 프레임들을 블렌딩하여, 최종 보간 프레임을 생성하는 것을 특징으로 하는 동영상 프레임 율 변환 방법.

청구항 4

청구항 3에 있어서,

제2 해상도의 광학 흐름 지도 생성단계는,

복원된 제2 해상도의 중간 프레임들과 광학 흐름 지도들에 제1 해상도의 프레임들을 제2 해상도로 다운 샘플링한 프레임들을 추가로 이용하여, 광학 흐름을 예측하는 것을 특징으로 하는 동영상 프레임 율 변환 방법.

청구항 5

청구항 3에 있어서,

복원 단계는,

선형 보간 방법을 이용하여, 제3 해상도의 중간 프레임들과 광학 흐름 지도들로, 제2 해상도의 중간 프레임들과 광학 흐름 지도들을 복원하는 것을 특징으로 하는 동영상 프레임 율 변환 방법.

청구항 6

청구항 3에 있어서,

제3 해상도의 광학 흐름 지도 생성단계, 제2 해상도의 광학 흐름 지도 생성단계 및 제1 해상도의 광학 흐름 지도 생성단계는,

훈련된 인공지능 모델을 이용하여, 제3 해상도의 광학 흐름 지도, 제2 해상도의 광학 흐름 지도 및 제1 해상도의 광학 흐름 지도를 생성하는 것을 특징으로 하는 동영상 프레임 율 변환 방법.

청구항 7

청구항 6에 있어서,

제3 해상도의 광학 흐름 지도 생성단계는,

제3 해상도의 프레임들과 제3 해상도의 광학 흐름 지도들을 bak-warping 연산하여 생성하고,

제2 해상도의 광학 흐름 지도 생성단계는,

제2 해상도의 프레임들과 제2 해상도의 광학 흐름 지도들을 bak-warping 연산하여 생성하며,

제1 해상도의 광학 흐름 지도 생성단계는,

제1 해상도의 프레임들과 제1 해상도의 광학 흐름 지도들을 bak-warping 연산하여 생성하는 것을 특징으로 하는 동영상 프레임 율 변환 방법.

청구항 8

청구항 6에 있어서,

제3 생성단계는,

$$B = \frac{1}{1 + e^{-k_1 k_2}}$$

위의 블렌딩 파라미터 B로 제1 해상도의 중간 프레임들을 블렌딩 연산하고,

k_1은 sigmoid 함수의 기울기를 조절하며,

K_2는 sigmoid 함수의 입력으로 주어지는 것을 특징으로 하는 동영상 프레임 율 변환 방법.

청구항 9

청구항 6에 있어서,

인공지능 모델의 훈련은,

원본 입력 영상과 보간 영상을 다시 원본 영상으로 back-warpking한 영상과의 차이를 계산하는 함수, 생성된 광학 흐름 지도들의 전과 후의 차이를 계산하는 함수, adversarial 손실 함수를 이용하는 것을 특징으로 하는 동영상 프레임 율 변환 방법.

청구항 10

시간적으로 연속된 고해상도의 프레임들로부터 생성한 저해상도의 프레임들들로 광학 흐름을 예측하여, 저해상도의 광학 흐름 지도들을 생성하는 예측부;

저해상도의 광학 흐름 지도들의 해상도를 단계적으로 높이면서, 고해상도의 중간 프레임들을 생성하는 향상부;

생성된 고해상도의 중간 프레임들을 블렌딩하여, 최종 보간 프레임을 생성하는 생성부:를 포함하는 것을 특징으로 하는 동영상 프레임 율 변환 장치.

청구항 11

저해상도의 프레임들들로 광학 흐름을 예측하여, 저해상도의 광학 흐름 지도들을 생성하는 제1 생성단계;

저해상도의 광학 흐름 지도들을 이용하여, 고해상도의 중간 프레임들을 생성하는 제2 생성단계; 및

생성된 고해상도의 중간 프레임들을 이용하여, 최종 보간 프레임을 생성하는 제3 생성단계:를 포함하는 것을 특징으로 하는 동영상 프레임 율 변환 방법.

청구항 12

시간적으로 연속된 고해상도의 프레임들로부터 생성한 저해상도의 프레임들들로 광학 흐름을 예측하여, 저해상도의 광학 흐름 지도들을 생성하는 제1 생성단계;.

저해상도의 광학 흐름 지도들의 해상도를 단계적으로 높이면서, 고해상도의 중간 프레임들을 생성하는 제2 생성단계;

생성된 고해상도의 중간 프레임들을 블렌딩하여, 최종 보간 프레임을 생성하는 제3 생성단계:를 포함하는 것을 특징으로 하는 동영상 프레임 율 변환 방법을 수행할 수 있는 프로그램이 기록된 컴퓨터로 읽을 수 있는 기록매체.

발명의 설명

기 술 분 야

[0001] 본 발명은 동영상의 프레임 율 변환 기술에 관한 것으로, 더욱 상세하게는 딥러닝 기법에 기반한 광학 흐름 추정 지도를 저해상도에서 예측하고 이를 고해상도로 복원하여 고해상도 동영상의 프레임을 고속으로 보간하는 동영상 프레임 율 고속 변환 방법 및 장치에 관한 것이다.

배 경 기 술

[0003] 1. 동영상 프레임 율 변환 기법 개요

[0004] 동영상은 연속된 정지 영상의 집합으로 구성된다. 비디오에서 정지 영상을 프레임이라고 부르며 단위 시간 당 프레임의 수를 동영상의 프레임 율 (frame rate)이라고 한다. 예를 들어 1초에 24장의 프레임으로 구성되면 프레임 율은 24 fps (frame per second)가 된다. 프레임 율은 촬영자의 의도, 영상의 포맷, 카메라의 한계 등에 의하여 결정된다.

[0005] 관찰자가 영상을 연속된 화면으로 느끼기 위해서는 어느 정도 이상의 프레임 율이 필요하고 이보다 낮을 경우 움직임이 부드럽지 않아 보인다. 이 현상은 디스플레이의 크기, 조명, 시청 거리 등에 의해 달라질 수 있다. 이를 개선하기 위해 동영상의 프레임 율을 후처리에 의해 증가시키는 것을 동영상 프레임 율 변환이라고 한다.

[0007] 2. 종래 기술

[0008] 동영상 프레임 율을 증가시키는 가장 간단한 방법은 프레임을 반복하는 것이다. 예를 들어 30 fps 영상을 60 fps 영상으로 증가시킬 경우, 각 프레임마다 한 프레임을 반복하여 출력하는 것이다.

[0009] 그러나 이 방법의 경우 동영상의 정보량은 동일하고 움직임에 대한 연속성은 변하지 않았으므로 관찰자가 느끼는 불편감은 동일하다. 이를 해결하기 위해 연속된 프레임들을 이용하여 가상의 프레임을 생성하는 기술이 개발되었다. 즉 t 초와 t+1 초 사이의 영상을 이용하여 t+0.5 초의 중간 영상을 새롭게 생성하며 이를 프레임 보간 (frame interpolation) 기술(도 1 참조)이라고 한다.

[0010] 프레임 보간은 다양한 방법이 개발되었으며 일반적으로는 다음과 같은 두 단계 과정을 거친다. 첫 번째 단계는 움직임 또는 광학 흐름 지도를 획득하는 단계 이며 두 번째 단계는 움직임 정보를 바탕으로 중간 프레임을 생성(Synthesis)하는 단계이다.

[0011] 동영상에서 물체의 움직임이 부드럽게 보이려면 중간 영상은 물체의 움직임이 두 영상 사이의 중간에 해당되어야 한다. 따라서 물체의 움직임 정보를 가지고 있는 광학 흐름 지도를 정확하게 찾는 것이 매우 중요하다. 이에 기반한 다양한 기법들이 제안되어 왔다.

[0012] 최근 딥러닝 (Deep learning) 알고리즘이 등장하여 컴퓨터 비젼, 음성 인식 등 다양한 분야에 널리 사용되고 있으며 종래에 방법에 비해 월등한 성능을 보이고 있다. 이에 발맞추어 딥러닝을 사용한 다양한 프레임 보간 기법이 등장하였다. 이 기법들은 딥러닝을 이용하여 고품질의 광학 흐름 지도를 예측하여 종래의 방법보다 더욱 뛰어난 보간 결과를 보여주고 있다.

[0014] 3. 종래 기술 문제점

[0015] 기존 방법의 문제는 네트워크 구조로 인한 문제로 4K (3840x2160) 해상도와 같이 큰 해상도에 대하여 GPU 메모리 부족으로 연산이 불가능 하거나, 매우 느린 연산 속도를 보여준다. 이와 같은 현상은 실시간 연산을 요구하는 상용 애플리케이션에 딥러닝을 이용한 프레임 보간 방법 적용을 어렵게 한다.

[0016] 또한, 고해상도 영상은 저해상도 영상들에 비하여 일반적으로 큰 움직임을 가진다. 기존 방법들은 이러한 큰 움직임에 대하여 저품질의 광학 흐름 지도를 생성하는 경향이 있으며, 이는 보간된 영상의 품질 저하를 야기하는 문제를 가진다.

발명의 내용

해결하려는 과제

[0018] 본 발명은 상기와 같은 문제점을 해결하기 위하여 안출된 것으로서, 본 발명의 목적은, 고해상도 비디오 영상에 대하여 고품질, 고속으로 프레임 보간을 수행하기 위한 방안으로, 입력 고해상도 영상을 저해상도 영상으로 변환하여 고속으로 광학 흐름 지도를 생성하고 이를 원본 고해상도로 복원하여 고해상도 영상을 고속으로 보간하는 동영상 프레임 율 고속 변환 방법 및 장치를 제공함에 있다.

과제의 해결 수단

[0020] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른, 동영상 프레임 율 변환 방법은 시간적으로 연속된 고해상도의 프레임들로부터 생성한 저해상도의 프레임들들로 광학 흐름을 예측하여, 저해상도의 광학 흐름 지도들을 생성하는 제1 생성단계; 저해상도의 광학 흐름 지도들의 해상도를 단계적으로 높이면서, 고해상도의 중간 프레임들을 생성하는 제2 생성단계; 생성된 고해상도의 중간 프레임들을 블렌딩하여, 최종 보간 프레임을 생성하

는 제3 생성단계:를 포함한다.

[0021] 제1 생성단계는, 입력되는 시간적으로 연속된 프레임들의 해상도인 제1 해상도 보다 낮은 제3 해상도의 프레임들로 광학 흐름을 예측하여, 제3 해상도의 광학 흐름 지도들을 생성하는 단계; 제3 해상도의 광학 흐름 지도들을 이용하여, 제3 해상도의 중간 프레임들을 생성하는 단계; 제3 해상도의 중간 프레임들과 광학 흐름 지도들로, 제3 해상도 보다 높은 제2 해상도의 중간 프레임들과 광학 흐름 지도들을 복원하는 단계;를 포함하고, 제2 생성단계는, 제2 해상도의 중간 프레임들과 광학 흐름 지도들을 이용하여, 제2 해상도 보다 높은 제1 해상도의 중간 프레임들을 생성하는 제2 생성단계;를 포함할 수 있다.

[0022] 제2 생성단계는, 복원된 제2 해상도의 중간 프레임들과 광학 흐름 지도들로 광학 흐름을 예측하여, 제2 해상도의 광학 흐름 지도들을 생성하는 단계; 제2 해상도의 광학 흐름 지도들을 이용하여, 제2 해상도의 중간 프레임들을 생성하는 단계; 생성된 제2 해상도의 중간 프레임들과 광학 흐름 지도들로 광학 흐름을 예측하여, 제1 해상도의 광학 흐름 지도들을 생성하는 단계; 제1 해상도의 광학 흐름 지도들을 이용하여, 제1 해상도의 중간 프레임들을 생성하는 단계;를 포함하고, 제3 생성단계는, 생성된 제1 해상도의 중간 프레임들을 블렌딩하여, 최종 보간 프레임을 생성하는 것일 수 있다.

[0023] 제2 해상도의 광학 흐름 지도 생성단계는, 복원된 제2 해상도의 중간 프레임들과 광학 흐름 지도들에 제1 해상도의 프레임들을 제2 해상도로 다운 샘플링한 프레임들을 추가로 이용하여, 광학 흐름을 예측하는 것일 수 있다.

[0024] 복원 단계는, 선형 보간 방법을 이용하여, 제3 해상도의 중간 프레임들과 광학 흐름 지도들로, 제2 해상도의 중간 프레임들과 광학 흐름 지도들을 복원하는 것일 수 있다.

[0025] 제3 해상도의 광학 흐름 지도 생성단계, 제2 해상도의 광학 흐름 지도 생성단계 및 제1 해상도의 광학 흐름 지도 생성단계는, 훈련된 인공지능 모델을 이용하여, 제3 해상도의 광학 흐름 지도, 제2 해상도의 광학 흐름 지도 및 제1 해상도의 광학 흐름 지도를 생성하는 것일 수 있다.

[0026] 제3 해상도의 광학 흐름 지도 생성단계는, 제3 해상도의 프레임들과 제3 해상도의 광학 흐름 지도들을 bak-warping 연산하여 생성하고, 제2 해상도의 광학 흐름 지도 생성단계는, 제2 해상도의 프레임들과 제2 해상도의 광학 흐름 지도들을 bak-warping 연산하여 생성하며, 제1 해상도의 광학 흐름 지도 생성단계는, 제1 해상도의 프레임들과 제1 해상도의 광학 흐름 지도들을 bak-warping 연산하여 생성하는 것일 수 있다.

[0027] 제3 생성단계는,

$$B = \frac{1}{1 + e^{-k_1 k_2}}$$

[0028]

[0029] 위의 블렌딩 파라미터 B로 제1 해상도의 중간 프레임들을 블렌딩 연산하고,

[0030] k_1은 sigmoid 함수의 기울기를 조절하며,

[0031] K_2는 sigmoid 함수의 입력으로 주어지는 것일 수 있다.

[0032] 인공지능 모델의 훈련은, 원본 입력 영상과 보간 영상을 다시 원본 영상으로 back-warpking한 영상과의 차이를 계산하는 함수, 생성된 광학 흐름 지도들의 전과 후의 차이를 계산하는 함수, adversarial 손실 함수를 이용하는 것일 수 있다.

[0033] 본 발명의 다른 측면에 따르면, 시간적으로 연속된 고해상도의 프레임들로부터 생성한 저해상도의 프레임들들로 광학 흐름을 예측하여, 저해상도의 광학 흐름 지도들을 생성하는 예측부; 저해상도의 광학 흐름 지도들의 해상도를 단계적으로 높이면서, 고해상도의 중간 프레임들을 생성하는 향상부; 생성된 고해상도의 중간 프레임들을 블렌딩하여, 최종 보간 프레임을 생성하는 생성부:를 포함하는 것을 특징으로 하는 동영상 프레임 율 변환 장치가 제공된다.

[0034] 본 발명의 또다른 측면에 따르면, 저해상도의 프레임들들로 광학 흐름을 예측하여, 저해상도의 광학 흐름 지도들을 생성하는 제1 생성단계; 저해상도의 광학 흐름 지도들을 이용하여, 고해상도의 중간 프레임들을 생성하는 제2 생성단계; 및 생성된 고해상도의 중간 프레임들을 이용하여, 최종 보간 프레임을 생성하는 제3 생성단계:를 포함하는 것을 특징으로 하는 동영상 프레임 율 변환 방법이 제공된다.

[0035] 본 발명의 또다른 측면에 따르면, 시간적으로 연속된 고해상도의 프레임들로부터 생성한 저해상도의 프레임들들로 광학 흐름을 예측하여, 저해상도의 광학 흐름 지도들을 생성하는 제1 생성단계; 저해상도의 광학 흐름 지도들의 해상도를 단계적으로 높이면서, 고해상도의 중간 프레임들을 생성하는 제2 생성단계; 생성된 고해상도의 중간 프레임들을 블렌딩하여, 최종 보간 프레임을 생성하는 제3 생성단계:를 포함하는 것을 특징으로 하는 동영상 프레임 율 변환 방법을 수행할 수 있는 프로그램이 기록된 컴퓨터로 읽을 수 있는 기록매체가 제공된다.

발명의 효과

[0037] 이상 설명한 바와 같이, 본 발명의 실시예들에 따르면, 입력 고해상도 영상을 저해상도 영상으로 변환하여 고속으로 광학 흐름 지도를 생성하고 이를 원본 고해상도로 복원하여 고해상도 영상을 보간함으로써, 4K와 같은 고해상도 비디오 영상에 대하여 실시간성을 요구하는 시스템 환경에서도 고품질, 고속으로 프레임 보간을 수행할 수 있게 된다.

도면의 간단한 설명

[0039] 도 1 : 프레임 보간 기술

도 2 : 본 발명의 실시예에 따른 동영상 프레임 율 변환 장치의 블럭도

도 3 : 피라미드 형태의 영상 표현 방법

도 4 : 광학 흐름 예측부, 광학 흐름 해상도 향상부 및 중간 프레임 생성부의 상세 구조

발명을 실시하기 위한 구체적인 내용

[0040] 이하에서는 도면을 참조하여 본 발명을 보다 상세하게 설명한다.

[0041] 도 2는 본 발명의 일 실시예에 따른 동영상 프레임 율 변환 장치의 블럭도이다. 본 발명의 실시예에 따른 동영상 프레임 율 변환 장치는, 도시된 바와 같이, 광학 흐름 예측부(110), 광학 흐름 해상도 향상부(120) 및 최종 중간 프레임 생성부(130)를 포함하여 구성된다.

[0042] 광학 흐름 예측부(110)는 입력으로 주어지는 시간적으로 연속된 프레임들의 양방향 광학 흐름을 저해상도에서 예측한다.

[0043] 기존의 방법에서는 광학 흐름 예측을 원본 해상도에서 수행하기 때문에, 고해상도 영상을 입력으로 받는 경우 많은 양의 메모리를 요구하며, 제한된 하드웨어 환경을 가지는 시스템에서는 프레임 보간을 수행 할 수 없으며, 프레임 보간을 수행한다 하더라도, 매우 느린 동작 속도를 가지는 문제를 가졌었다. 또한 고해상도 영상의 경우 영상내의 객체들의 움직임이 매우 큰 편이며, 기존의 방법들은 이러한 경우에 blur나 ghost 열화를 발생시키는 문제가 있었다.

[0044] 하지만, 본 발명의 실시예에서는 광학 흐름 예측을 저해상도에서 수행함으로써 상기 문제들을 효과적으로 해결한다.

[0045] 도 3은 본 발명의 실시예에서 사용가능한 피라미드 형태의 영상 표현 방법들을 보여준다. 도 3에서 Level 1 해상도는 원본 입력 영상의 해상도와 동일한 해상도를 가지며, level 2의 영상들은 level 1의 해상도를 2분의 1크기로 downsample한 해상도를 가진다. 마지막으로 level 3 해상도는 level 1 해상도의 4분의 1크기를 가진다.

[0046] 광학 흐름 예측부(110)는 level 3의 해상도에서 광학 흐름 지도를 생성하고, 광학 흐름 해상도 향상부(120)에서는 level 3 해상도의 광학 흐름 지도를 level 2 해상도와 level 1 해상도로 순차적으로 복원한다.

[0047] 광학 흐름 지도의 해상도 향상을 단계적으로 수행하는 것은 광학 흐름이 급작스럽게 변화하지 않도록 하여, 최종 보간된 프레임이 자연스러운 움직임을 갖도록 하기 위함이다.

[0048] 최종적으로 최종 중간 프레임 생성부(130)에서는, 복원된 고해상도 광학 흐름 지도를 이용하여 최종 중간 프레임을 생성하여 보간한다.

[0049] 도 4는 광학 흐름 예측부(110), 광학 흐름 해상도 향상부(120) 및 최종 중간 프레임 생성부(130)의 상세 구조를 도식화하여 보여준다.

[0050] 일반적으로 고품질의 광학 흐름 지도를 생성하기 위해서는 합성곱 신경망에서 충분한 크기의 receptive field가 요구된다. 하지만 receptive field가 커질수록 합성곱 신경망의 알고리즘 계산 복잡도가 증가하게 되는 문제를

가진다. 기존의 방법들에서는 원본 영상의 해상도와 동일한 크기를 가지는 광학 흐름 지도를 생성하기 때문에 4K 프레임과 같은 고해상도 영상에 대해서 느린 동작 속도를 가지는 문제를 가진다.

[0051] 이에 본 발명의 실시예에서는 원본 입력 영상의 해상도를 level 3 해상도로 줄여 광학 흐름을 예측한다. 이를 통하여 비교적 작은 크기의 receptive field를 통해서도 영상 내의 객체의 움직임을 쉽게 다룰 수 있고, 동시에 계산 연산에 필요한 메모리와 동작시간을 크게 감소시킬 수 있다.

[0052] 광학 흐름 예측부(110)는 합성곱 신경망(111)를 포함하고 있다. 합성곱 신경망(111)에서는 원본 입력 해상도의 4분의 1의 크기를 갖는 level 3 해상도의 원본 입력 영상을 이용하여 광학 흐름 지도 $F_{t1}^{3'}$와 $F_{t2}^{3'}$를 생성한다. 그리고, 광학 흐름 지도 $F_{t1}^{3'}$ 와 $F_{t2}^{3'}$ 를 이용하여, 광학 흐름 예측부(110)는 level 3 해상도의 중간 프레임 $I_{t1}^{3'}$ 와 $I_{t2}^{3'}$ 를 생성한다. $I_{t1}^{3'}$ 와 $I_{t2}^{3'}$ 는 다음의 수학식 (1)로 구할 수 있다.

[0053]
$$I_{t1}^{3'} = b_w(I_1^3, F_{t1}^{3'}), \qquad I_{t2}^{3'} = b_w(I_2^3, F_{t2}^{3'}) \tag{1}$$

[0054] 여기서 $b_w(\cdot,\cdot)$는 bak-warping 연산을 의미한다.

[0055] 다음, 광학 흐름 예측부(110)는 선형 보간 방법을 이용하여, 중간 프레임 $I_{t1}^{3'}$과 $I_{t2}^{3'}$ 그리고 광학 흐름 지도 $F_{t1}^{3'}$ 과 $F_{t2}^{3'}$로부터 원본 입력 해상도의 2분의 1크기를 갖는 level 2 해상도의 I_{t1}^{2}과 I_{t2}^{2} 그리고 level 2 해상도의 F_{t1}^{2}과 F_{t2}^{2}를 복원한다.

[0056] 복원된 영상들은 광학 흐름 해상도 향상부(120)의 입력으로 주어진다. 이 때, 2분의 1크기로 downsample 된 level 2 해상도의 원본 입력 영상을 추가적으로 사용하여 광학 흐름 지도의 해상도 향상시 품질 향상에 도움을 줄 수 있다.

[0057] 광학 흐름 해상도 향상부(120)는 2개의 합성곱 신경망을 포함하고 있다. 첫 번째 합성곱 신경망(121)에서는 $X_1 = \{I_1^2, I_2^2, F_{t1}^2, F_{t2}^2, I_{t1}^2, I_{t2}^2\}$, $X_1 \in R^{2\times\frac{H}{2}\times\frac{W}{2}}$ 를 입력으로 받아, level 2 해상도의 광학 흐름 지도 $F_{t1}^{2'}$과 $F_{t2}^{2'}$를 생성한다. 여기서 H와 W는 원본 입력 영상의 세로와 가로 크기를 의미한다.

[0058] 그리고 광학 흐름 해상도 향상부(120)는 수식 (1)과 유사하게 광학 흐름 지도 $F_{t1}^{2'}$과 $F_{t2}^{2'}$를 이용하여 level 2 해상도의 중간 프레임 $I_{t1}^{2'}$ 과 $I_{t2}^{2'}$ 를 생성한다. $I_{t1}^{2'}$ 과 $I_{t2}^{2'}$ 는 다음의 수학식 (2)로 구할 수 있다.

[0059]
$$I_{t1}^{2'} = b_w(I_1^2, F_{t1}^{2'}), \quad I_{t2}^{2'} = b_w(I_2^2, F_{t2}^{2'}) \tag{2}$$

[0060] 다음 광학 흐름 해상도 향상부(120)의 두 번째 합성곱 신경망(122)에서는 $X_2 = \{F_{t1}^{2'}, F_{t2}^{2'}, I_{t1}^{2'}, I_{t2}^{2'}\}$ 를 입력으로 받아 원본 입력 영상의 해상도인 level 1 해상도를 가지는 광학 흐름 지도 $F_{t1}^{1'}$과 $F_{t2}^{1'}$를 복원한다. 즉, $F_{t1}^{1'}$ 과 $F_{t2}^{1'}$ 는 원본 해상도인 level 1 해상도로 복원된 최종 광학 흐름 지도를 의미한다.

[0061] 그리고 광학 흐름 해상도 향상부(120)는 원본 입력 영상의 해상도인 level 1 해상도를 가지는 중간 프레임 $I_{t1}^{1'}$ 과 $I_{t2}^{1'}$ 을 다음 식 (3)을 통하여 구한다.

[0062] $$I_{t1}^{1'} = b_w(I_1^1, F_{t1}^{1'}) \quad I_{t2}^{1'} = b_w(I_2^1, F_{t2}^{1'}) \quad (3)$$

[0063] 마지막으로, 최종 중간 프레임 생성부(130)는 최종 중간 프레임 $I_{t1}^{1'}$ 과 $I_{t2}^{1'}$를 블렌딩 (blending) 하여 보간할 최종 중간 프레임을 구한다. 최종 중간 프레임 $I_t^{'}$ 을 다음 식 (4)를 통하여 구한다.

[0064] $$I_t^{'} = B \odot I_{t1}^{1'} + (1 - B) \odot I_{t2}^{1'} \quad (4)$$

[0065] 여기서 ⊙는 element-wise 곱셈 연산을 의미하며 B는 입력 영상의 blending parameter를 의미한다. 예를 들어, pixel p가 두 개의 입력 영상 I_1 과 I_2 에 모두 존재한다면 B는 0.5로 설정되며, 만약 p가 I_1에는 존재하지만 I_2 에는 존재하지 않는다면 B=1 로 설정된다.

[0066] Blending parameter를 통한 중간 프레임 생성 방법은 가려짐 영역들에 의하여 발생하는 문제를 효과적으로 해결하나, 복잡한 구조의 변화나, 너무 큰 움직임을 가지는 객체들에 대해서는 저조한 성능을 보인다. 특히 이러한 경우에는 B의 값이 0.5로 설정되는 경우가 많아, 최종적으로 생성된 중간프레임이 blur한 결과를 보이는 경우가 많다.

[0067] 본 발명의 실시예에서는 이를 해결하기 위하여,sigmoid 함수를 이용하여 적응적으로 학습을 통해 blending parameter를 생성하는 방법을 이용한다. 적응적 blending parameter B는 다음을 통하여 구한다.

[0068] $$B = \frac{1}{1 + e^{-k_1 k_2}} \quad (5)$$

[0069] 여기서, k_1과 K_2는 두 번째 광학 흐름 해상도 향상부(120)에서 출력된다. 첫 번째 parameter k_1은 sigmoid 함수의 기울기를 조절하며, 두 번째 parameter K_2는 sigmoid 함수의 입력으로 주어진다. Sigmoid 함수의 기울기 조절을 통하여 blur와 ghost 문제를 효과적으로 해결할 수 있다.

[0070] 한편, 본 발명의 실시예에서는 딥러닝 기반의 광학 흐름 예측부(110)와 광학 흐름 해상도 향상부(120)를 훈련하기 위하여 multi-scale smoothness 손실 함수, consistecny 손실 함수와 adversarial 손실 함수를 새롭게 제안한다.

[0071] Consistency 손실 함수는 원본 입력 영상과 보간 영상을 다시 원본 영상으로 back-warpking한 영상과의 차이를 계산하는 함수이다. 이를 통하여 광학 흐름 해상도 향상부(120)가 고 품질의 고해상도 광학 흐름 지도를 복원하고 보간 영상이 blur 열화되는 것을 방지 할 수 있다. 제안하는 consistency 손실 함수 l_c는 다음 식을 통하여 계산된다.

[0072] $$l_c = \| f_b(I_t, F_{1t}^{1'}) - I_1^1 \|_1 + \| f_b(I_t, F_{2t}^{1'}) - I_2^1 \|_1 \quad (6)$$

[0073] 한편, Consistency 손실 함수를 이용하여 blur 열화 문제를 방지 함에도 불구하고, 최종 보간 영상이 급작스러운 광학 흐름 변화로 인하여 over-smoothed되는 경향을 보일 수 있다.

[0074] 이를 해결하기 위하여 본 발명의 실시예에서는 mulit-scale smoothness 손실 함수를 제안한다. Multi-scale smoothness 손실함수는 광학 흐름 해상도 향상부(120)의 결과와 이전 단계의 광학 흐름 해상도 향상부(120)와 광학 흐름 예측부(110)의 결과의 차이를 계산하는 함수이다.

[0075] Multi-scale smoothness 손실함수는 regularization 역할을 수행하며, 훈련된 광학 흐름 해상도 향상부(120)가 안정적으로 고해상도의 광학 흐름 지도를 복원하도록 한다. Multi-scale smoothness 손실함수 l_s는 다음 식을 통하여 계산 된다.

$$l_s = \sum_{k=1}^{2} \| f_u(2 \odot F_{t1}^{k+1'}) - F_{t1}^{k'} \|_1$$

[0076] (7)

[0077] 여기서 f_u는 양방향 선형 보간 연산을 의미한다. 마지막으로 자연스러운 보간 영상을 생성하기 위하여 adversarial 손실 함수를 제안한다.

[0078] Adversarial 손실 함수 l_a는 다음 식을 통하여 계산된다.

$$l_a = \min_G \max_D E_{I_t \sim p_{data}(I_t)}[\log D(I_t)] + E_{(I_1, I_2) \sim \pi(I_1, I_2)}[\log(1 - D(G(I_1, I_2)))]$$

[0079] (8)

[0080] 지금까지, 고해상도 동영상 프레임 율 고속 변환 방법 및 장치에 대해 바람직한 실시예를 들어 상세히 설명하였다.

[0081] 종래 방법은 고해상도 영상을 입력시, 원본 해상도와 동일한 해상도를 가지는 광학 흐름 지도를 생성하여 프레임 보간을 수행함으로써, 많은 양의 메모리를 요구하며, 매우 느린 보간 속도를 갖는 문제점을 보였따.

[0082] 하지만, 본 발명의 실시예에서는 입력 고해상도 영상을 저해상도 영상으로 변환하여 고속으로 광학 흐름 지도를 생성하고 이를 원본 고해상도로 복원하는 형태로 동작하여, 4K 프레임과 같은 고해상도 영상을 고속으로 보간이 가능하다.

[0083] 본 발명의 실시예에 따른 고해상도 동영상 프레임 율 고속 변환 장치를 구현할 수 있는 하드웨어 구조에 대해, 이하에서 도 5를 참조하여 상세히 설명한다.

[0084] 도 5는 본 발명의 다른 실시예에 따른 고해상도 동영상 프레임 율 고속 변환 장치로 기능할 수 있는 영상 시스템의 하드웨어 구조를 나타낸 블록도이다. 본 발명의 실시예에 따른 영상 시스템은, 도 5에 도시된 바와 같이, 입력부(210), 프로세서(220), 출력부(230) 및 저장부(240)를 포함한다.

[0085] 입력부(210)는 외부 저장매체, 외부 기기, 통신망 등을 통해 영상 데이터를 입력받는 수단이고, 프로세서(220)는 입력된 영상에 대해 프레임 율 변환을 수행하기 위한 CPU들과 GPU들의 집합니다.

[0086] 동영상 프레임 율을 변환함에 있어 프로세서(220)는 전술한 실시예에서 제시한 방법을 이용한다. 저장부(240)는 프로세서(220)가 프레임 율 변환을 수행함에 있어 필요한 저장공간을 제공하는 내부 저장매체이다.

[0087] 출력부(230)는 프로세서(220)에서 프레임 율이 변환된 영상을 외부 저장매체, 외부 기기, 통신망 등으로 출력한다.

[0088] 한편, 본 실시예에 따른 장치와 방법의 기능을 수행하게 하는 컴퓨터 프로그램을 수록한 컴퓨터로 읽을 수 있는 기록매체에도 본 발명의 기술적 사상이 적용될 수 있음은 물론이다. 또한, 본 발명의 다양한 실시예에 따른 기술적 사상은 컴퓨터로 읽을 수 있는 기록매체에 기록된 컴퓨터로 읽을 수 있는 코드 형태로 구현될 수도 있다. 컴퓨터로 읽을 수 있는 기록매체는 컴퓨터에 의해 읽을 수 있고 데이터를 저장할 수 있는 어떤 데이터 저장 장치이더라도 가능하다. 예를 들어, 컴퓨터로 읽을 수 있는 기록매체는 ROM, RAM, CD-ROM, 자기 테이프, 플로피 디스크, 광디스크, 하드 디스크 드라이브, 등이 될 수 있음은 물론이다. 또한, 컴퓨터로 읽을 수 있는 기록매체에 저장된 컴퓨터로 읽을 수 있는 코드 또는 프로그램은 컴퓨터간에 연결된 네트워크를 통해 전송될 수도 있다.

[0089] 또한, 이상에서는 본 발명의 바람직한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

부호의 설명

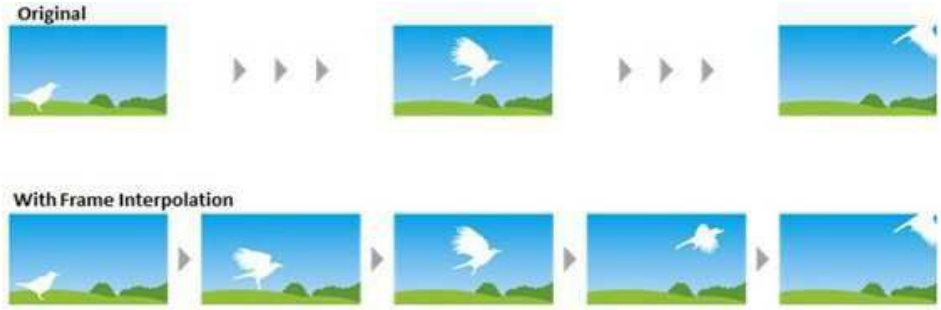
[0091] 110 : 광학 흐름 예측부

120 : 광학 흐름 해상도 향상부

130 : 최종 중간 프레임 생성부

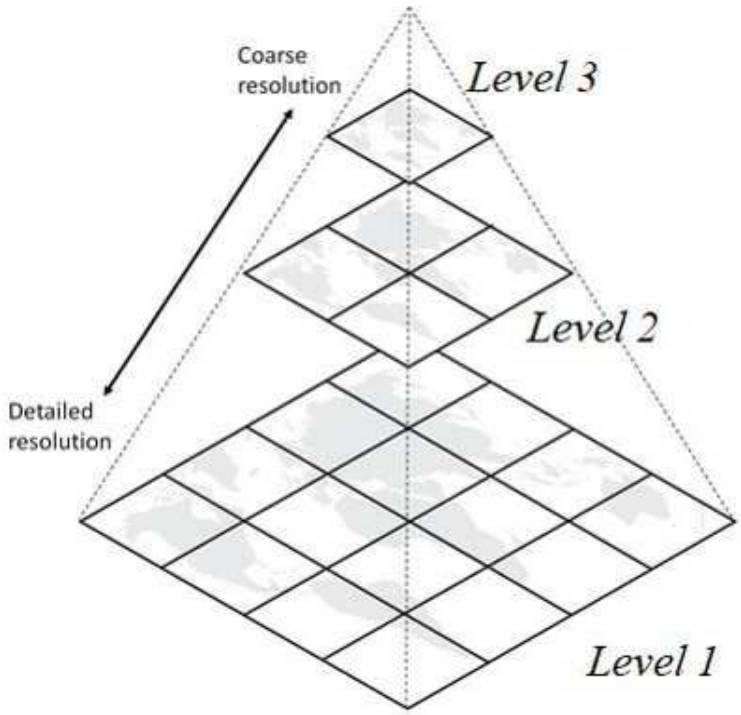
도면

도면1

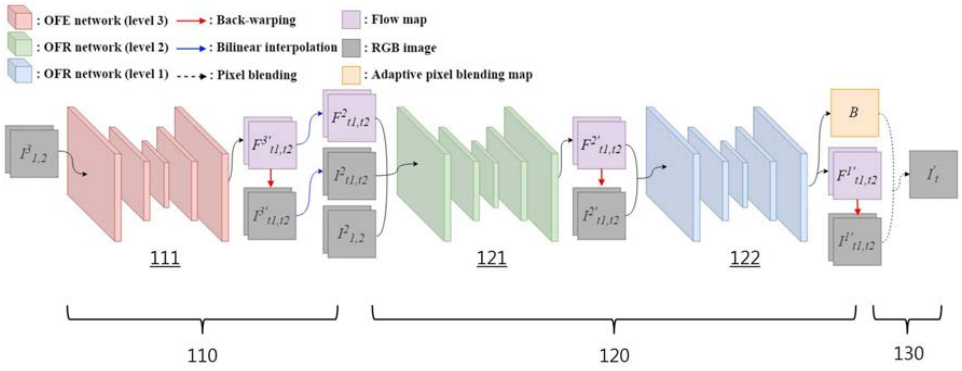


도면2

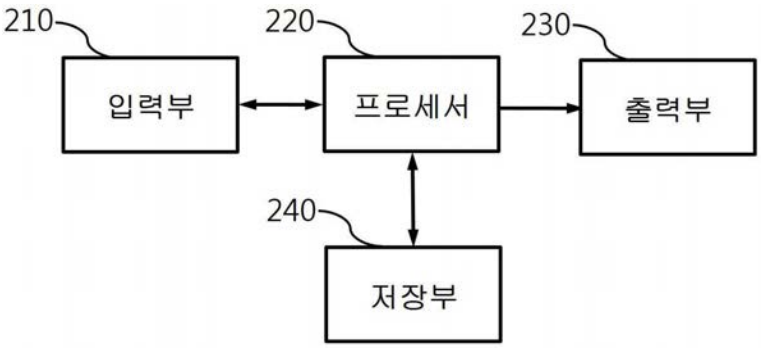
도면3



도면4


: OFE network (level 3)
: Back-warping
: Flow map
: OFR network (level 2)
: Bilinear interpolation
: RGB image
: OFR network (level 1)
: Pixel blending
: Adaptive pixel blending map
111
121
122
110
120
130

도면5


210
입력부
220
프로세서
230
출력부
240
저장부

(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년04월20일
(11) 등록번호 10-2242334
(24) 등록일자 2021년04월14일

(51) 국제특허분류(Int. Cl.)
H04N 7/01 (2006.01) *H04N 5/262* (2006.01)
(52) CPC특허분류
H04N 7/0135 (2013.01)
H04N 5/2628 (2013.01)
(21) 출원번호 10-2019-0167488
(22) 출원일자 2019년12월16일
심사청구일자 2020년02월21일
(56) 선행기술조사문헌
US20070097260 A1
US20190138889 A1
US20190139205 A1

(73) 특허권자
한국전자기술연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
정진우
서울특별시 송파구 송파대로 111, 파크하비오105동 1205호
안하은
서울특별시 성북구 성북로8다길 37, 102호
김제우
경기도 성남시 분당구 수내로 181, 303동 901호
(74) 대리인
남충우

전체 청구항 수 : 총 11 항 심사관 : 박재학

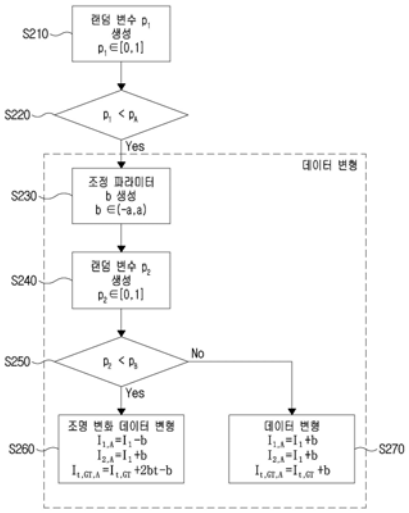
(54) 발명의 명칭 **데이터 변형을 통한 고해상도 동영상 프레임 율 고속 변환 방법 및 장치**

(57) 요 약

고해상도 비디오 영상에 대하여 고품질, 고속으로 프레임 보간을 수행하는 딥러닝에 기반 프레임 율 고속 변환 방법 및 장치가 제공된다. 본 발명의 실시예에 따른 데이터 변형 방법은 프레임 보간을 위한 중간 프레임을 생성하는 인공지능 모델의 학습에 이용될 연속하는 입력 영상들을 입력받는 단계; 입력된 입력 영상들의 영상값을 변경하여, 변형된 영상들을 생성하는 제1 생성단계; 입력된 입력 영상들의 중간 프레임의 영상값을 변경하여, 변형된 영상들의 중간 프레임을 생성하는 제2 생성단계;를 포함한다.

이에 의해, 4K 프레임과 같은 고해상도 영상을 고속으로 보간이 가능하며, Fade-in/out, 조명 변화, 줌 영상에 강인하게 학습 데이터 변형을 수행하여 이런 다양한 영상에서도 정확한 광학 흐름 지도를 생성해 낼 수 있게 하여, 고해상도 영상에 대한 고속의 강인한 프레임 보간이 가능해진다.

대 표 도 - 도12



(52) CPC특허분류
H04N 7/0127 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711081116
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	차세대(UHD)방송서비스활성화기술개발(R&D)
연구과제명	영상콘텐츠 초고속/초고화질 변환 기술 개발
기 여 율	1/1
과제수행기관명	(주)픽스트리
연구기간	2019.01.01 ~ 2019.12.31

명 세 서

청구범위

청구항 1

프레임 보간을 위한 중간 프레임을 생성하는 인공지능 모델의 학습에 이용될 연속하는 입력 영상들을 입력받는 단계;

입력된 입력 영상들의 영상값을 변경하여, 변형된 영상들을 생성하는 제1 생성단계;

입력된 입력 영상들의 중간 프레임의 영상값을 변경하여, 변형된 영상들의 중간 프레임을 생성하는 제2 생성단계;

입력 영상들과 입력 영상들의 중간 프레임의 기준 패치를 결정하는 단계;

입력 영상들 중 제1 입력 영상의 기준 패치를 축소하는 단계;

입력 영상들 중 제2 입력 영상의 기준 패치를 확대하는 단계;

축소된 제1 입력 영상의 기준 패치를 원래 크기로 조정하는 제1 조정단계;

확대된 제2 입력 영상의 기준 패치를 원래 크기로 조정하는 제2 조정단계;를 포함하는 것을 특징으로 하는 데이터 변형 방법.

청구항 2

청구항 1에 있어서,

제1 생성 단계는,

입력 영상들 중 제1 입력 영상에 대해서는 조정 파라미터를 감산하여 제1 입력 영상을 변형하고,

입력 영상들 중 제2 입력 영상에 대해서는 조정 파라미터를 가산하여 제2 입력 영상을 변경하며,

제2 생성 단계는,

중간 프레임의 시간에 따라 가변하는 조정 파라미터를 가산하여 중간 프레임을 변형하는 것을 특징으로 하는 데이터 변형 방법.

청구항 3

청구항 2에 있어서,

제1 생성 단계는,

입력 영상들 중 제1 입력 영상에 대해 조정 파라미터를 가산하여 제1 입력 영상을 변형하고,

입력 영상들 중 제2 입력 영상에 대해 조정 파라미터를 가산하여 제2 입력 영상을 변형하며,

제2 생성 단계는,

중간 프레임에 대해 조정 파라미터를 가산하여 중간 프레임을 변형하는 것을 특징으로 하는 데이터 변형 방법.

청구항 4

청구항 2 또는 청구항 3에 있어서,

조정 파라미터는,

랜덤한 값으로 생성되는 것을 특징으로 하는 데이터 변형 방법.

청구항 5

청구항 4에 있어서,

조정 파라미터는,

밝기, 감마 함수, 컨트라스트, 휴(hue), 새츄레이션(saturation) 중 어느 하나를 조정하기 위한 파라미터인 것을 특징으로 하는 데이터 변형 방법.

청구항 6

삭제

청구항 7

청구항 1에 있어서,

결정 단계는,

기준 패치의 위치와 크기를 결정하는 것을 특징으로 하는 데이터 변형 방법.

청구항 8

청구항 1에 있어서,

줌 파라미터를 랜덤하게 생성하는 단계;를 더 포함하고,

축소 단계는,

제1 입력 영상의 기준 패치를 줌 파라미터를 이용하여 축소하며,

확대 단계는,

제2 입력 영상의 기준 패치를 줌 파라미터를 이용하여 확대하는 것을 특징으로 하는 데이터 변형 방법.

청구항 9

청구항 1에 있어서,

제1 조정 단계 및 제2 조정 단계는,

확대된 제2 입력 영상의 기준 패치가 제2 입력 영상을 벗어나지 않은 것으로 판단된 경우에 수행되는 것을 특징으로 하는 데이터 변형 방법.

청구항 10

프레임 보간을 위한 중간 프레임을 생성하는 인공지능 모델의 학습에 이용될 연속하는 입력 영상들을 입력받는 입력부;

입력된 입력 영상들의 영상값을 변경하여 변형된 영상들을 생성하고, 입력된 입력 영상들의 중간 프레임의 영상값을 변경하여 변형된 영상들의 중간 프레임을 생성하는 프로세서;를 포함하고,

프로세서는,

입력 영상들과 입력 영상들의 중간 프레임의 기준 패치를 결정하고,

입력 영상들 중 제1 입력 영상의 기준 패치를 축소하며,

입력 영상들 중 제2 입력 영상의 기준 패치를 확대하고,

축소된 제1 입력 영상의 기준 패치를 원래 크기로 조정하며,

확대된 제2 입력 영상의 기준 패치를 원래 크기로 조정하는 것을 특징으로 하는 데이터 변형 시스템.

청구항 11

프레임 보간을 위한 중간 프레임을 생성하는 인공지능 모델의 학습에 이용될 연속하는 입력 영상들을 입력받는 단계;

입력된 입력 영상들의 영상값을 변경하여, 변형된 영상들을 생성하는 제1 생성단계;

입력된 입력 영상들의 중간 프레임의 영상값을 변경하여, 변형된 영상들의 중간 프레임을 생성하는 제2 생성단계;

제1 생성단계에서 생성된 변형된 영상들과 제2 생성단계에서 생성된 변형된 영상들의 중간 프레임으로 인공지능 모델을 학습시키는 단계;

입력 영상들과 입력 영상들의 중간 프레임의 기준 패치를 결정하는 단계;

입력 영상들 중 제1 입력 영상의 기준 패치를 축소하는 단계;

입력 영상들 중 제2 입력 영상의 기준 패치를 확대하는 단계;

축소된 제1 입력 영상의 기준 패치를 원래 크기로 조정하는 제1 조정단계;

확대된 제2 입력 영상의 기준 패치를 원래 크기로 조정하는 제2 조정단계;

제1 조정단계에서 조정된 제1 입력 영상의 기준 패치, 제2 조정단계에서 조정된 제2 입력 영상의 기준 패치 및 결정단계에서 결정된 중간 프레임의 기준 패치로 인공지능 모델을 학습시키는 단계;를 포함하는 것을 특징으로 하는 인공지능 모델 학습 방법.

청구항 12

프레임 보간을 위한 중간 프레임을 생성하는 인공지능 모델의 학습에 이용될 연속하는 입력 영상들의 영상값을 변경하여 변형된 영상들을 생성하고, 입력된 입력 영상들의 중간 프레임의 영상값을 변경하여 변형된 영상들의 중간 프레임을 생성하는 데이터 변형 시스템; 및

프로세서에 의해 생성된 변형된 영상들과 변형된 영상들의 중간 프레임으로 인공지능 모델을 학습시키는 고속 프레임 보간 시스템;을 포함하고,

데이터 변형 시스템은,

입력 영상들과 입력 영상들의 중간 프레임의 기준 패치를 결정하고,

입력 영상들 중 제1 입력 영상의 기준 패치를 축소하며,

입력 영상들 중 제2 입력 영상의 기준 패치를 확대하고,

축소된 제1 입력 영상의 기준 패치를 원래 크기로 조정하며,

확대된 제2 입력 영상의 기준 패치를 원래 크기로 조정하고,

고속 프레임 보간 시스템은,

조정된 제1 입력 영상의 기준 패치, 조정된 제2 입력 영상의 기준 패치 및 결정된 중간 프레임의 기준 패치로 인공지능 모델을 학습시키는 것을 특징으로 하는 인공지능 모델 학습 시스템.

발명의 설명

기 술 분 야

[0001] 본 발명은 고해상도 동영상의 프레임 율 변환 기술에 관한 것으로, 더욱 상세하게는 딥러닝 기법에 기반한 광학 흐름 추정 지도를 저해상도에서 예측하고 이를 고해상도로 복원하여 고해상도 동영상의 츠레임을 고속으로 보간하는 방법에 관한 것이다.

배 경 기 술

[0003] 1) 동영상 프레임 율 변환 기법 개요

[0004] 동영상은 연속된 정지 영상의 집합으로 구성된다. 비디오에서 정지 영상을 프레임이라고 부르며 단위 시간 당 프레임의 수를 동영상의 프레임 율(frame rate)이라고 한다. 예를 들어 1초에 24장의 프레임으로 구성되면 프레임 율은 24 fps(frame per second)가 된다. 프레임 율은 촬영자의 의도, 영상의 포맷, 카메라의 한계 등에 의하여 결정된다. 관찰자가 영상을 연속된 화면으로 느끼기 위해서는 어느 정도 이상의 프레임 율이 필요하고 이보다 낮을 경우 움직임이 부드럽지 않아 보인다. 이 현상은 디스플레이의 크기, 조명, 시청 거리 등에 의해 달라질 수 있다. 이를 개선하기 위해 동영상의 프레임 율을 후처리에 의해 증가시키는 것을 동영상 프레임 율 변환이라고 한다.

[0006] 2) 종래 기술

[0007] 동영상 프레임 율을 증가시키는 가장 간단한 방법은 프레임을 반복하는 것이다. 예를 들어 30fps 영상을 60 fps 영상으로 증가시킬 경우, 각 프레임마다 한 프레임을 반복하여 출력하는 것이다. 그러나 이 방법의 경우 동영상의 정보량은 동일하고 움직임에 대한 연속성은 변하지 않았으므로 관찰자가 느끼는 불편감은 동일하다. 이를 해결하기 위해 연속된 프레임들을 이용하여 가상의 프레임을 생성하는 기술이 개발되었다. 즉 t 초와 t+1 초 사이의 영상을 이용하여 t+0.5 초의 중간 영상을 새롭게 생성하며 이를 프레임 보간(frame interpolation) 기술이라고 한다.

[0008] 프레임 보간은 다양한 방법이 개발되었으며 일반적으로는 다음과 같은 두 단계 과정을 거친다. 첫 번째 단계는 움직임 또는 광학 흐름 지도를 획득하는 단계 이며 두 번째 단계는 움직임 정보를 바탕으로 중간 프레임을 생성(Synthesis) 하는 단계이다. 동영상에서 물체의 움직임이 부드럽게 보이려면 중간 영상은 물체의 움직임이 두 영상 사이의 중간에 해당되어야 한다. 따라서 물체의 움직임 정보를 가지고 있는 광학 흐름 지도를 정확하게 찾는 것이 매우 중요하다. 이에 기반한 다양한 기법들이 제안되어 왔다.

[0009] 최근 딥러닝(Deep learning) 알고리즘이 등장하여 컴퓨터 비젼, 음성 인식 등 다양한 분야에서 사용되고 있으며 종래에 방법에 비해 월등한 성능을 보이고 있다. 이에 발맞추어 딥러닝을 사용한 다양한 프레임 보간 기법이 등장하였다. 이 기법들은 딥러닝을 이용하여 고품질의 광학 흐름 지도를 예측하여 종래의 방법보다 더욱 뛰어난 보간 결과를 보여줌에 따라 최근 지속적으로 연구되고 있다.

[0011] 3) 종래 기술 문제점

[0012] 기존 방법의 문제는 네트워크 구조로 인한 문제로 4K(3840x2160) 해상도와 같이 큰 해상도에 대하여 GPU 메모리 부족으로 연산이 불가능 하거나, 매우 느린 연산 속도를 보여준다. 이와 같은 현상은 실시간 연산을 요구하는 상용 애플리케이션에 딥러닝을 이용한 프레임 보간 방법 적용을 어렵게한다. 또한, 고해상도 영상은 저해상도 영상들에 비하여 일반적으로 큰 움직임을 가진다. 기존 방법들은 이러한 큰 움직임에 대하여 저품질의 광학 흐름 지도를 생성하는 경향이 있으며, 이는 보간된 영상의 품질 저하를 야기하는 문제를 가진다.

[0013] 또한 기존 방법의 문제는 Fade-in/out, 조명 변화, 줌 영상을 고려하지 않아 이런 종류의 영상에 있어서는 광학 흐름 지도가 정확히 생성되지 않는 문제점이 발생한다. 따라서 이런 영상에서는 프레임 보간 성능이 급격히 저하되는 단점이 존재한다.

발명의 내용

해결하려는 과제

[0015] 본 발명은 상기와 같은 문제점을 해결하기 위하여 안출된 것으로서, 본 발명의 목적은, 고해상도 비디오 영상에 대하여 고품질, 고속으로 프레임 보간을 수행하는 딥러닝에 기반 프레임 율 고속 변환 방법 및 장치를 제공함에 있다.

과제의 해결 수단

[0017] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른, 데이터 변형 방법은 프레임 보간을 위한 중간 프레임을 생성하는 인공지능 모델의 학습에 이용될 연속하는 입력 영상들을 입력받는 단계; 입력된 입력 영상들의 영상값을 변경하여, 변형된 영상들을 생성하는 제1 생성단계; 입력된 입력 영상들의 중간 프레임의 영상값을 변경하여, 변형된 영상들의 중간 프레임을 생성하는 제2 생성단계;를 포함한다.

[0018] 제1 생성 단계는, 입력 영상들 중 제1 입력 영상에 대해서는 조정 파라미터를 감산하여 제1 입력 영상을 변형하고, 입력 영상들 중 제2 입력 영상에 대해서는 조정 파라미터를 가산하여 제2 입력 영상을 변경하며, 제2 생성 단계는, 중간 프레임의 시간에 따라 가변하는 조정 파라미터를 가산하여 중간 프레임을 변형하는 것일 수 있다.

[0019] 제1 생성 단계는, 입력 영상들 중 제1 입력 영상에 대해 조정 파라미터를 가산하여 제1 입력 영상을 변형하고, 입력 영상들 중 제2 입력 영상에 대해 조정 파라미터를 가산하여 제2 입력 영상을 변형하며, 제2 생성 단계는, 중간 프레임에 대해 조정 파라미터를 가산하여 중간 프레임을 변형하는 것일 수 있다.

[0020] 조정 파라미터는, 랜덤한 값으로 생성되는 것일 수 있다.

[0021] 조정 파라미터는, 밝기, 감마 함수, 컨트라스트, 휴(hue), 새츄레이션(saturation) 중 어느 하나를 조정하기 위한 파라미터일 수 있다.

[0022] 본 발명에 따른 데이터 변형 방법은 입력 영상들과 입력 영상들의 중간 프레임의 기준 패치를 결정하는 단계; 입력 영상들 중 제1 입력 영상의 기준 패치를 축소하는 단계; 입력 영상들 중 제2 입력 영상의 기준 패치를 확대하는 단계; 축소된 제1 입력 영상의 기준 패치를 원래 크기로 조정하는 제1 조정단계; 확대된 제2 입력 영상의 기준 패치를 원래 크기로 조정하는 제2 조정단계;를 더 포함할 수 있다.

[0023] 결정 단계는, 기준 패치의 위치와 크기를 결정하는 것일 수 있다.

[0024] 본 발명에 따른 데이터 변형 방법은 줌 파라미터를 랜덤하게 생성하는 단계;를 더 포함하고, 축소 단계는, 제1 입력 영상의 기준 패치를 줌 파라미터를 이용하여 축소하며, 확대 단계는, 제2 입력 영상의 기준 패치를 줌 파라미터를 이용하여 확대하는 것일 수 있다.

[0025] 제1 조정 단계 및 제2 조정 단계는, 확대된 제2 입력 영상의 기준 패치가 제2 입력 영상을 벗어나지 않은 것으로 판단된 경우에 수행되는 것일 수 있다.

[0026] 본 발명의 다른 측면에 따르면, 프레임 보간을 위한 중간 프레임을 생성하는 인공지능 모델의 학습에 이용될 연속하는 입력 영상들을 입력받는 입력부; 입력된 입력 영상들의 영상값을 변경하여 변형된 영상들을 생성하고, 입력된 입력 영상들의 중간 프레임의 영상값을 변경하여 변형된 영상들의 중간 프레임을 생성하는 프로세서;를 포함하는 것을 특징으로 하는 데이터 변형 시스템이 제공된다.

[0027] 본 발명의 또다른 측면에 따르면, 프레임 보간을 위한 중간 프레임을 생성하는 인공지능 모델의 학습에 이용될 연속하는 입력 영상들을 입력받는 단계; 입력된 입력 영상들의 영상값을 변경하여, 변형된 영상들을 생성하는 제1 생성단계; 입력된 입력 영상들의 중간 프레임의 영상값을 변경하여, 변형된 영상들의 중간 프레임을 생성하는 제2 생성단계; 제1 생성단계에서 생성된 변형된 영상들과 제2 생성단계에서 생성된 변형된 영상들의 중간 프레임으로 인공지능 모델을 학습시키는 단계;를 포함하는 것을 특징으로 하는 인공지능 모델 학습 방법이 제공된다.

[0028] 본 발명의 또다른 측면에 따르면, 프레임 보간을 위한 중간 프레임을 생성하는 인공지능 모델의 학습에 이용될 연속하는 입력 영상들의 영상값을 변경하여 변형된 영상들을 생성하고, 입력된 입력 영상들의 중간 프레임의 영상값을 변경하여 변형된 영상들의 중간 프레임을 생성하는 데이터 변형 시스템; 및 프로세서에 의해 생성된 변형된 영상들과 변형된 영상들의 중간 프레임으로 인공지능 모델을 학습시키는 고속 프레임 보간 시스템;을 포함하는 것을 특징으로 하는 인공지능 모델 학습 시스템이 제공된다.

발명의 효과

[0030] 이상 설명한 바와 같이, 본 발명의 실시예들에 따르면, 4K 프레임과 같은 고해상도 영상을 고속으로 보간이 가

능하며, Fade-in/out, 조명 변화, 줌 영상에 강인하게 학습 데이터 변형을 수행하여 이런 다양한 영상에서도 정확한 광학 흐름 지도를 생성해 낼 수 있게 하여, 고해상도 영상에 대한 고속의 강인한 프레임 보간이 가능해진다.

도면의 간단한 설명

[0032] 도 1은 프레임 보간 기술,

도 2는 고속 프레임 보간 시스템의 블럭도,

도 3은 피라미드 형태의 영상 표현 방법,

도 4은 저해상도 광학 흐름 예측부의 상세 블럭도,

도 5는 광학 흐름 해상도 향상부의 상세 블럭도,

도 6은 중간 프레임 시간 간격,

도 7은 중간 프레임 생성부의 상세 블럭도,

도 8은 중간 프레임 해상도 향상부의 상세 블럭도,

도 9는 학습 데이터 구성,

도 10은 데이터 변형된 학습 데이터 구성,

도 11은 밝기 조정 파라미터 결정 방법,

도 12는 학습을 위한 샘플 별 데이터 변형 과정,

도 13은 줌 데이터 변형 패치 생성 개념도,

도 14는 줌 데이터 변형 패치 생성 과정,

도 15는 줌 데이터 변형 패치 생성 순서도,

도 16은 데이터 변형 시스템의 블럭도이다.

발명을 실시하기 위한 구체적인 내용

[0033] 이하에서는 도면을 참조하여 본 발명을 보다 상세하게 설명한다.

[0034] 본 발명의 실시예에서는, 첫 번째로 딥러닝 네트워크를 이용하여 고해상도 영상에 대하 고속 프레임 보간을 수행하는 방법을 제시하고, 두 번째로 조명 변화, 페이드 인/아웃, 줌 영상에 강인한 프레임 보간 기법으로 이는 학습시 데이터를 변형(argumentation) 하는 방법을 제시한다.

[0036] 1. 고속 프레임 보간

[0037] 본 발명의 일 실시예에 따른 고속 프레임 보간 시스템은, 도 2에 도시된 바와 같이, 광학 흐름 예측부(110), 저해상도 광학 흐름 해상도 향상부(120), 중간 프레임 생성부(130), 중간 프레임 해상도 향상부(140)를 포함하여 구성된다.

[0038] 저해상도 광학 흐름 예측부(110)에서는 입력으로 주어진 시간적으로 연속된 프레임들의 양방향 광학 흐름을 저해상도에서 예측한다. 기존의 방법에서는 광학 흐름 예측을 원본 해상도에서 수행하기 때문에, 고해상도 영상을 입력으로 받는 경우 많은 양의 메모리를 요구하며, 제한된 하드웨어 환경을 가지는 시스템에서는 프레임 보간을 수행 할 수 없으며, 프레임 보간을 수행한다 하더라도, 매우 느린 동작 속도를 가지는 문제를 가진다. 또한 고해상도 영상의 경우 영상내의 객체들의 움직임이 매우 큰 편이며, 기존의 방법들은 이러한 경우에 blur나 ghost 열화를 발생시키는 문제를 갖는다. 본 발명의 실시예에서는 광학 흐름 예측을 저해상도에서 수행함으로써 상기 문제들을 효과적으로 해결한다.

[0039] 도 3은 본 발명의 실시예에서 사용하는 피라미드 형태의 영상 표현 방법을 보여준다. Level 1 해상도는 원본 입력 영상의 해상도와 동일한 해상도를 가지며, Level 2의 영상들은 Level 1의 해상도를 2분의 1크기로 downsample한 해상도를 가진다. 마지막으로 Level 3 해상도는 Level 1 해상도의 4분의 1크기를 가진다. 즉 Level 1의 원본 입력 해상도를 W(너비) x H(높이) 라고 하면, Level2의 해상도는 W/2 x H/2, Level 3의 해상도

는 W/4 x H/4 가 된다.

[0040] 저해상도 광학 흐름 예측부(110)에서는 Level 3의 해상도에서 광학 흐름 지도를 생성하고, 광학 흐름 해상도 향상부(120)에서는 Level 3 해상도의 광학 흐름 지도를 Level 2 해상도로 복원한다. 중간 프레임 생성부(130)에서는 Level 2의 해상도를 갖는 광학 흐름과 Level 2 해상도의 입력 영상을 이용하여 Level 2 해상도의 중간 프레임을 생성한다. 중간 프레임 해상도 향상부(140)에서는 Level 2 해상도의 중간 프레임을 Level 1의 원해상도로 복원한다.

[0041] 저해상도 광학 흐름 예측부(110)에서는 입력 두 프레임 사이의 광학 흐름 지도를 산출한다. 일반적으로 고품질의 광학 흐름 지로를 생성하기 위해서는 합성곱 신경망에서 충분한 크기의 receptive field가 요구된다. 하지만 receptive field가 커질수록 합성곱 신경망의 알고리즘 계산 복잡도가 증가하게 되는 문제를 가진다. 기존의 방법들에서는 원본 영상의 해상도와 동일한 크기를 가지는 광학 흐름 지도를 생성하기 때문에 4K 프레임과 같은 고해상도 영상에 대해서 느린 동작 속도를 가지는 문제를 가진다.

[0042] 본 발명의 실시예에서는 도 4에 도시된 바와 같이, 해상도 감소부(111)가 원본 입력 영상의 해상도를 Level 3 해상도로 줄인 후에, 광학 흐름 예측부(112)가 광학 흐름을 예측한다.

[0043] 해상도 감소부(111)에서는 Level 1 해상도의 원본 입력 영상을 Level 3 해상도로 줄인다. 줄이는 방법은 bilinear 또는 bicubic 필터 등을 사용할 수 있으며 특정 방법에 한정되지 않는다. 도 3에서 I_1^{L1}, I_2^{L1}의 L1은 Level 1의 해상도를, I_1^{L3}, I_2^{L3}의 L3는 Level 3의 해상도를 각각 나타낸다.

[0044] 광학 흐름 예측부(112)에서는 Level 3 저해상도로 영성으로 입력 두 프레임 사이의 광학 흐름 지도를 산출한다. 이를 통하여 비교적 작은 크기의 receptive field를 통해서도 영상 내의 객체의 움직임을 쉽게 다룰 수 있고, 동시에 계산 연산에 필요한 메모리와 동작시간을 크게 감소시킬 수 있다.

[0045] 광학 흐름 예측부(112)에서는 원본 입력 해상도의 4분의 1의 크기(즉 Level 1의 크기) 를 갖는 광학 흐름 지도 $F_{1\to2}^{L3}$와 $F_{2\to1}^{L3}$를 생성한다. 광학 흐름을 생성하는 방법은 다양한 딥러닝 네트워크 및 기존 방법들이 사용될 수 있으며 특정한 방법에 한정되지 않는다.

[0046] 광학 흐름 해상도 향상부(120)는 Level 3 광학 흐름 지도를 Level 2의 광학 흐름 지도로 해상도를 개선시키는 장치로 도 5와 같이 구성된다.

[0047] 해상도 증가부1(121)은 Level 3 해상도의 광학 흐름 지도를 Level 2의 해상도의 광학 흐름 지도로 증가시킨다. 증가 방법은 bilinear 또는 bicubic 필터 등을 사용할 수 있으며 특정 방법에 한정되지 않는다.

[0048] 광학 흐름 재조정부(123)는 입력 시간 t 에 대한 광학 흐름 지도로 재조정한다. 도 6과 같이 t는 I_1, I_2 사이에 생성할 중간 프레임의 시간을 의미한다. $F_{1\to2}^{L3}$ 의 의미는 I_1^{L3} 에서 I_2^{L3}로 진행할 때의 광학 흐름 지도를 $F_{2\to1}^{L3}$은 그 역을 나타낸다. 도 5의 $\hat{F}_{t\to1}^{L2}$는 임의의 t에서 보간된 프레임인 $\hat{I}_t^{L2}$ 에서 I_1^{L2}으로 진행되는 광학 흐름 지도로 볼 수 있고 $\hat{F}_{t\to2}^{L2}$ 도 마찬가지이다. $\hat{F}_{t\to1}^{L2}$와 $\hat{F}_{t\to2}^{L2}$는 다음과 같이 산출 될 수 있다.

$$\hat{F}_{t\to1}^{L2} = -(1-t)t\times F_{1\to2}^{L2} + t^2 \times F_{2\to1}^{L2}$$
$$\hat{F}_{t\to2}^{L2} = (1-t)^2 \times F_{1\to2}^{L2} - t(1-t)\times F_{2\to1}^{L2}$$

[0049] (1)

[0050] 해상도 감소부2(112)에서는 Level 1 해상도를 갖는 입력 원영상을 Level 2의 해상도로 감소시켜 I_1^{L2}, I_2^{L2}를 생성한다. 증가 방법은 bilinear 또는 bicubic 필터 등을 사용할 수 있으며 특정 방법에 한정되지 않는다.

[0051] 와핑부 1(124)에서는 광학 흐름 지도 $\hat{F}_{t\to1}^{L2}$와 $\hat{F}_{t\to2}^{L2}$와 Level 2의 입력 영상 I_1^{L2}, I_2^{L2}를 이용하여 임시 중간 프레임 $\hat{I}_{t\to1}^{L2}$와 $\hat{I}_{t\to2}^{L2}$를 생성한다. $\hat{I}_{t\to1}^{L2}$와 $\hat{I}_{t\to2}^{L2}$는 식 (2)와 같이 구할 수 있으며, 여기서 $b_w(\bullet,\bullet)$은 back-warping 연산을 의미한다.

[0052]
$$\hat{I}_{t\to1}^{L2} = b_w(I_1^{L2}, \hat{F}_{t\to1}^{L2})$$
$$\hat{I}_{t\to2}^{L2} = b_w(I_2^{L2}, \hat{F}_{t\to2}^{L2}) \quad (2)$$

[0053] 광학 흐름 향상부(125)에서는 와핑부1(124)의 출력 $\hat{I}_{t\to1}^{L2}$, $\hat{I}_{t\to2}^{L2}$, 광학 흐름 재조정부(123)의 출력 $\hat{F}_{t\to1}^{L2}$, $\hat{F}_{t\to2}^{L2}$, 해상도 감소부2(112)의 출력 I_1^{L2}, I_2^{L2}을 이용하여 광학 흐름을 개선한다. 즉, 위의 입력들을 $X^{L2}=\{I_1^{L2}, I_2^{L2}, \hat{F}_{t\to1}^{L2}, \hat{F}_{t\to2}^{L2}, \hat{I}_{t\to1}^{L2}, \hat{I}_{t\to2}^{L2}\}$로 묶을 수 있고 광학 흐름 향상부(125)의 입력이 된다. 광학 흐름 향상부(125)는 합성곱 신경망으로 구성되며 개선된 광학 흐름 지도 $F_{t\to1}^{L2}$, $F_{t\to2}^{L2}$를 생성한다.

[0054] 중간 프레임 생성부(130)는 광학 흐름 해상도 향상부(120)의 출력인 $F_{t\to1}^{L2}$, $F_{t\to2}^{L2}$를 이용하여 Level 2 해상도의 중간 프레임을 생성한다. 도 7은 중간 프레임 생성부(130)의 상세 구성을 나타낸다. 도시된 바와 같이, 중간 프레임 생성부(130)는 와핑부2(131)와 중간 프레임 합성부(132)로 구성된다.

[0055] 와핑부2(131)에서는 광학 흐름 지도 $F_{t\to1}^{L2}$, $F_{t\to2}^{L2}$와 Level 2의 입력 영상 I_1^{L2}, I_2^{L2}를 이용하여 임시 중간 프레임 $I_{t\to1}^{L2}$와 $I_{t\to2}^{L2}$를 생성한다. 이는 식 (3)과 같이 구할 수 있다.

[0056]
$$I_{t\to1}^{L2} = b_w(I_1^{L2}, F_{t\to1}^{L2})$$
$$I_{t\to2}^{L2} = b_w(I_2^{L2}, F_{t\to2}^{L2}) \quad (3)$$

[0057] 중간 프레임 합성부(132)에서는 생성된 임시 중간 프레임을 이용하여 Level 2 해상도의 최종 중간 프레임을 다음과 같이 생성한다.

[0058]
$$\hat{I}_t^{L2} = B \odot I_{t\to1}^{L2} + (1-B) \odot I_{t\to2}^{L2}$$

[0059] 여기서 ⊙는 element-wise 곱셈 연산을 의미하며 B는 입력 영상의 blending parameter를 의미한다. 예를 들어, pixel의 p가 두 개의 입력 영상 I_1과 I_2에 모두 존재한다면 B는 0.5로 설정되며, 만약 p가 I_1에는 존재하지만 I_2에 존재하지 않는다면 B=1 로 설정된다. B는 학습을 통해서 선택 될 수도 있고 고정된 값으로 결정될 수 있다.

[0060] 중간 프레임 해상도 향상부(140)는 Level 2 해상도의 최종 중간 프레임을 $\hat{I}_t^{L2}$를 원영상 해상도인 Level1 해상도로 복원한다. 도 8에 도시된 바와 같이, 중간 프레임 해상도 향상부(140)는 해상도 증가부2(141과 142), 와핑부3(143), 해상도 개선부(144)로 구성된다.

[0061] 해상도 증가부2(141과 142)에서는 $\hat{I}_t^{L2}$와 $F_{t\to1}^{L2}$, $F_{t\to2}^{L2}$를 Level 2 해상도에서 Level 1 해상도로 증가시켜

$\hat{I}_t^{L1}$와 $F_{t\to1}^{L1}$, $F_{t\to2}^{L1}$를 생성한다. 증가 방법은 bilinear 또는 bicubic 필터 등을 사용할 수 있으며 특정 방법에 한정되지 않는다.

[0062] 해상도 개선부(144)에서는 고해상도 영상을 입력으로 넣기 위해 원영상 정보를 입력으로 사용한다. 원영상을 바로 넣을 경우 I_1^{L1}, $\hat{I}_t^{L1}$, I_1^{L2}의 입력 간에 시간 차이가 존재하여 개선 효율이 저하된다. 이를 방지하기 위해 I_1^{L1}과 I_2^{L1}는 광학 흐름 지도를 이용하여 와핑부3(143)을 거쳐 t시간의 프레임을 각각 생성하도록 한다. 이는 식 (4)와 같이 표현할 수 있다.

[0063]
$$I_{t\to1}^{L1} = b_w(I_1^{L1}, F_{t\to1}^{L1})$$
$$I_{t\to2}^{L1} = b_w(I_2^{L1}, F_{t\to2}^{L1}) \quad (4)$$

[0064] 최종적으로 해상도 개선부(144)는 $I_{t\to1}^{L1}$, $\hat{I}_t^{L1}$, $I_{t\to2}^{L1}$를 묶어 $X^{L1} = \{ I_{t\to1}^{L1}, \hat{I}_t^{L1}, I_{t\to2}^{L1} \}$을 입력으로 하는 합성곱 신경망으로 구성된다. 해상도 개선부(144)의 출력은 Level 1의 해상도를 갖는 $\tilde{I}_t^{L1}$이 전체 네트워크의 최종 출력이 된다.

[0065] 위에서 제시한 방법으로 중간 프레임을 생성할 경우 고해상도 영상의 빠른 움직임을 고속으로 정확하게 찾을 수 있어 정확한 프레임 보간이 가능하다.

[0067] 2. 데이터 변형

[0068] 전술하였듯이 일반적인 영상은 빈번하게 플래쉬, 조명 변화, 페이드 인/아웃이 발생하게 된다. 이런 영상에 있어서는 광학 흐름 지도가 정확하게 산출되기 어렵다. 이런 원인이 발생되는 주된 이유는 학습시 데이터에 조명 변화가 발생하는 샘플이 많이 포함되지 않아서이다. 조명 변화가 있는 샘플을 많이 포함하기 위해서는 이런 데이터를 많이 수집해야 하는데 이는 많은 비용과 노력이 필요하다. 본 발명의 실시예에서는 학습 시 데이터 변형을 통해 이를 해결하도록 한다.

[0069] 일반적으로 딥러닝을 이용한 프레임 보간 학습은, 도 9와 같이 두 개의 입력 영상 I_1, I_2과 한개의 Ground Truth(GT) 영상 $I_{t,GT}$으로 구성된다. 딥러닝 네트워크에서 두 개의 입력 영상은 도 2에서의 입력 영상이 되고, 출력 영상은 $\tilde{I}_t$가 된다. 학습 시 $I_{t,GT}$와 $\tilde{I}_t$의 차이가 최소가 되게 하도록 네트워크를 갱신하게 된다.

[0070] 도 9에서 보듯이 대부분의 영상에서는 학습 데이터 간 즉, I_1, I_2, $I_{t,GT}$간에 조명 변화가 일어나지 않는다. 조명 변화가 있는 학습 데이터 샘플이 부족하므로 조명 변화가 발생할 경우 영상이 정확히 보간되지 않게 된다.

[0071] 이 문제를 해결하기 위해 도 10과 같이 데이터 변형(argumentation)을 통해 강제로 조명 변화가 발생하는 학습 데이터를 생성시킨다. 밝기 조절을 위해 밝기 및 감마(gamma) 함수 등이 사용될 수 있다. 밝기 조절의 강도는 밝기 조정 파라미터 b에 의해 결정되어 진다. 이와 같을 경우 다음과 같이 데이터의 변형을 이루도록 한다.

[0072]
$$I_{1,A} = I_1 - b$$
$$I_{2,A} = I_2 + b \quad (5)$$

[0073] GT 영상은 t의 위치에 따라 데이터 변형을 수행하도록 하며 밝기 변형은 선형적으로 변한다고 가정한다. 이는 도 11과 같은 그래프로 표현될 수 있다. X축은 GT가 위치한 시간을 Y축은 밝기 조정 파라미터 값을 나타낸다. t의 범위는 I_1과 I_2 사이로 제한되므로 0과 1을 포함하지 않는 0과 1 사이의 값이다. 예를 들어 t=0.5 이면 GT의 밝기 조정 파라미터는 0이 된다. GT의 밝기 조정 파라미터는 식 (6)과 같이 일반화 할 수 있다. 밝기 조정

파라미터의 산출은 항상 선형식일 필요는 없으며 다양한 방식으로 계산이 가능하다.

[0074]
$$I_{t,GT,A} = I_{t,GT} + 2 \times b \times t - b \quad (6)$$

[0075] 영상의 신호 범위를 0과 1사이로 정규화 했다고 가정하면, b의 범위는 -1과 1사이로 매 배치(batch)마다 랜덤하게 선택된다. 모든 학습 샘플에서 데이터 변형을 수행하지 않고 일정 확률 P 만큼 랜덤하게 데이터 변형을 수행하게 된다. 즉 데이터 변형 횟수와 강도는 모두 랜덤하게 선택되도록 한다.

[0076] 또한 데이터 변형을 입력 영상에 동일하게 하는 방법을 적용할 수 있다. 이는 특정 밝기 영상의 샘플이 부족할 때 성능이 떨어지는 점을 보완할 수 있다. 이럴 경우 데이터 변형은 식 (7)과 같이 이루어진다.

[0077]
$$I_{1,A} = I_1 + b$$
$$I_{2,A} = I_2 + b$$
$$I_{t,GT,A} = I_{t,GT} + b \quad (7)$$

[0078] 도 12는 위의 매 학습 샘플마다 이루어지는 데이터 변형 과정을 보여준다. 0과 1사이의 값을 갖는 랜덤 변수 p_1을 생성시킨다(S210). 랜덤 변수 p_1이 P_A보다 클 경우에만(S220-Yes), 데이터 변형을 수행하도록 한다(박스 점선 부분). P_A는 0과 1사이의 값으로 데이터 변형이 적용되는 빈도를 결정한다. 만약 데이터 변형을 하도록 결정되면, 조정 파라미터 b를 생성한다(S230). 조정 파라미터 b는 데이터 변형 강도를 결정하게 된다. 조정 파라미터 범위 a는 영상의 범위 데이터 변형 종류 등에 따라 다양하게 결정될 수 있다. 그 다음 조명 변화 데이터 변환을 할지를 랜덤 변수 p_2를 사용하여 결정하고(S240), 조명 변화 학습 변형을 할 경우에는(S250-Yes), 식 (5,6)을 사용하여 데이터 변형을 수행하고(S260), 그렇지 않을 경우에는(S250-No), 식 (7)을 사용하여 데이터 변형을 수행한다(S270).

[0079] 위와 같은 방법은 감마 함수, 컨트라스트, 휴(hue), 새츄레이션(saturation) 등 다양한 데이터 변형 함수에 대하여 적용이 가능하다.

[0080] 다른 실시 예로 줌(zoom) 영상을 효과적으로 대응하기 위한 데이터 변형 방법을 제시한다. 줌 영상은 시간이 진행함에 따라 물체의 크기가 변하게 된다. 이런 영상에 대하여 효율적으로 보간하기 위하여, 도 13와 같이 줌을 고려한 데이터 변형 패치를 생성하도록 한다.

[0081] 일반적으로 학습을 진행 할 시 영상의 일부분을 패치로 사용하도록 한다. 예를 들어 영상의 크기를 $W \times H$로 하고 패치의 크기를 $w_0 \times h_0$라고 하자. 입력 영상과 GT 영상의 패치의 위치는 동일하다. 도 13에서와 같이 한 개의 입력은 크기를 감소시키고 다른 한 개의 입력은 크기를 증가시키면서 줌 영상에 대응하는 패치를 생성할 수 있다.

[0082] 도 14와 도 15는 구체적인 줌 데이터 생성 패치 과정을 보여준다. 위의 데이터 변형 과정과 마찬가지로 0과 1사이의 값을 갖는 랜덤 변수 p_1을 생성시킨다(S310). 랜덤 변수 p_1가 P_A보다 클 경우에만(S320-Yes) 줌 데이터 변형을 수행하도록 한다.

[0083] 줌 데이터 변형을 수행하기로 결정되었다면 줌 파라미터 s를 랜덤하게 설정한다(S330). 줌 파라미터 s는 줌 배율의 강도를 의미하며 1이면 원래의 패치 크기와 같게 된다. 줌 파라미터가 너무 클 경우 패치 크기가 영상 크기를 벗어날 수 있는 경우가 많이 발생함으로 1과 1.2 사이의 값을 권장한다.

[0084] 다음으로는 도 14과 같이 기준 패치의 위치 $(x_{0L}, y_{0L}, x_{0R}, y_{0R})$를 결정한다(S340). 패치의 크기는 $w_0 \times h_0$이다. 기준 패치의 위치, 패치의 크기 및 줌 파라미터를 이용하여 입력 영상 I_1, I_2에서 사용될 위치를 식

(8~11)과 같이 결정한다. 식(8)과 식(10)에서 $\lfloor X \rfloor_{Even}$ 는 X 를 넘지 않는 가장 큰 짝수를 의미한다. I_1 영상에서 사용될 패치의 크기는 식 (8)과 같이 산출된다. 재계산된 패치 크기인 $w_1 \times h_1$ 를 기반으로 I_1 의 패치 영역을 식 (9)와 같이 계산한다. 이는 도 14의 I_1 영상의 빨간 박스 영역을 의미한다. I_2 영상에서 사용될 패치의 크기는 식 (10)과 같이 산출된다. 재계산된 패치 크기인 $w_2 \times h_2$ 를 기반으로 I_2 의 패치 영역을 식 (11)와 같이 계산한다. 이는 도 14의 I_2 영상의 빨간 박스 영역을 의미한다. I_1 과 I_2 의 위치는 랜덤하게 바뀔 수 있다.

[0085]
$$w_1 = \lfloor w_0 \times s \rfloor_{Even}$$
$$h_1 = \lfloor h_0 \times s \rfloor_{Even} \quad (8)$$

[0086]
$$x_{1L} = x_{0L} - \frac{w_1 - w_0}{2}$$
$$x_{1R} = x_{0R} + \frac{w_1 - w_0}{2}$$
$$y_{1L} = y_{0L} - \frac{h_1 - h_0}{2}$$
$$y_{1R} = y_{0R} + \frac{h_1 - h_0}{2} \quad (9)$$

[0087]
$$w_2 = \left\lfloor \frac{w_0}{s} \right\rfloor_{Even}$$
$$h_2 = \left\lfloor \frac{h_0}{s} \right\rfloor_{Even} \quad (10)$$

[0088]
$$x_{2L} = x_{0L} + \frac{w_0 - w_2}{2}$$
$$x_{2R} = x_{0R} - \frac{w_0 - w_2}{2}$$
$$y_{2L} = y_{0L} + \frac{h_0 - h_2}{2}$$
$$y_{2R} = y_{0R} - \frac{h_0 - h_2}{2} \quad (11)$$

[0089] 만약 식(9)에서 계산된 영역이 영상의 범위를 벗어난다면 줌 데이터 변형 과정을 수행하지 않는다. 만약 경계 조건을 만족한다면(S360-Yes), 식(9)에서 계산된 좌표를 이용하여 $w_1 \times h_1$ 의 크기를 갖는 $I_{1,patch}$를 생성한다(S370). $I_{1,patch}$ 는 $w_0 \times h_0$ 로 해상도를 조정하여 줌 데이터 변형된 패치인 $I^Z_{1,patch}$를 생성한다. 마찬가지로 식 (10)에서 계산된 좌표를 이용하여 $w_2 \times h_2$ 의 크기를 갖는 $I_{2,patch}$ 를 생성하고 $w_0 \times h_0$로 해상도를 조정하여 줌 데이터 변형된 패치인 $I^Z_{2,patch}$ 를 생성한다(S340). 최종적으로 $w_0 \times h_0$를 갖는 $I^Z_{1,patch}$, $I_{t,GT,patch}$, $I^Z_{2,patch}$ 를 이용하여 학습을 수행하도록 한다.

[0090] 지금까지 설명한 데이터 변형을 수행할 수 있는 시스템에 대해, 이하에서 16을 참조하여 상세히 설명한다. 도 16은 본 발명의 다른 실시예에 따른 데이터 변형 시스템의 블록도이다.

[0091] 본 발명의 실시예에 따른 데이터 변형 시스템은, 도 16에 도시된 바와 같이, 입력부(210), 프로세서(220), 출력부(230) 및 저장부(240)를 포함하는 컴퓨팅 시스템으로 구현 가능하다.

[0092] 입력부(210)는 외부 저장매체, 외부 기기, 통신망 등을 통해 영상들을 입력받는 수단이고, 프로세서(220)는 입력된 영상 데이터들을 변형하여 학습 영상들을 증가시킨다.

[0093] 영상 데이터를 변형함에 있어, 프로세서(220)는 전술한 실시예에서 제시한 방법을 이용한다. 저장부(240)는 프로세서(220)가 영상 데이터들을 변형함에 있어 필요한 저장공간을 제공하는 내부 저장매체이다.

[0094] 출력부(230)는 프로세서(220)에서 추가된 학습 영상들을 외부 저장매체, 외부 기기, 통신망 등으로 출력한다.

[0096] 3. 변형예

[0097] 지금까지, 딥러닝 네트워크를 이용하여 고해상도 영상에 대하 고속 프레임 보간을 수행하는 방법과 조명 변화, 페이드 인/아웃, 줌 영상에 강인한 프레임 보간 기법으로 이는 학습시 데이터를 변형하는 방법에 대해, 바람직한 실시예들을 들어 상세히 설명하였다.

[0098] 고해상도 영상을 입력시, 원본 해상도와 동일한 해상도를 가지는 광학 흐름 지도를 생성하여 프레임 보간을 수행하여 많은 양의 메모리를 요구하며, 매우 느린 보간 속도를 갖는 종래 방법과 달리, 본 발명의 실시예에서는 입력 고해상도 영상을 저해상도 영상으로 변환하여 고속으로 광학 흐름 지도를 생성하고 이를 원본 고해상도로 복원하여, 4K 프레임과 같은 고해상도 영상을 고속으로 보간이 가능하게 하였따.

[0099] 또한, 본 발명의 실시예에서는, Fade-in/out, 조명 변화, 줌 영상에 강인하게 학습 데이터 변형을 수행하여 이런 다양한 영상에서도 정확한 광학 흐름 지도를 생성해 낼 수 있게 하였다.

[0100] 한편, 본 실시예에 따른 장치와 방법의 기능을 수행하게 하는 컴퓨터 프로그램을 수록한 컴퓨터로 읽을 수 있는 기록매체에도 본 발명의 기술적 사상이 적용될 수 있음은 물론이다. 또한, 본 발명의 다양한 실시예에 따른 기술적 사상은 컴퓨터로 읽을 수 있는 기록매체에 기록된 컴퓨터로 읽을 수 있는 코드 형태로 구현될 수도 있다. 컴퓨터로 읽을 수 있는 기록매체는 컴퓨터에 의해 읽을 수 있고 데이터를 저장할 수 있는 어떤 데이터 저장 장치이더라도 가능하다. 예를 들어, 컴퓨터로 읽을 수 있는 기록매체는 ROM, RAM, CD-ROM, 자기 테이프, 플로피 디스크, 광디스크, 하드 디스크 드라이브, 등이 될 수 있음은 물론이다. 또한, 컴퓨터로 읽을 수 있는 기록매체에 저장된 컴퓨터로 읽을 수 있는 코드 또는 프로그램은 컴퓨터간에 연결된 네트워크를 통해 전송될 수도 있다.

[0101] 또한, 이상에서는 본 발명의 바람직한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

부호의 설명

[0103] 110 : 광학 흐름 예측부

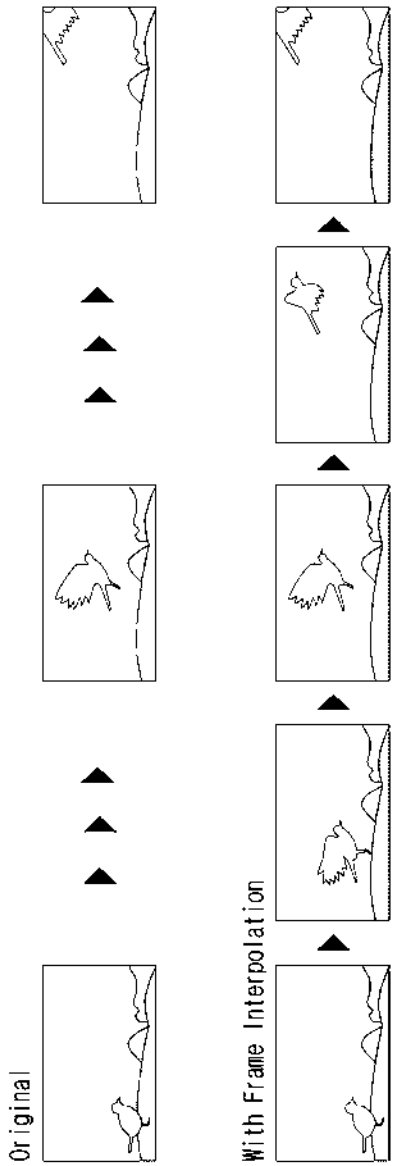
120 : 저해상도 광학 흐름 해상도 향상부

130 : 중간 프레임 생성부

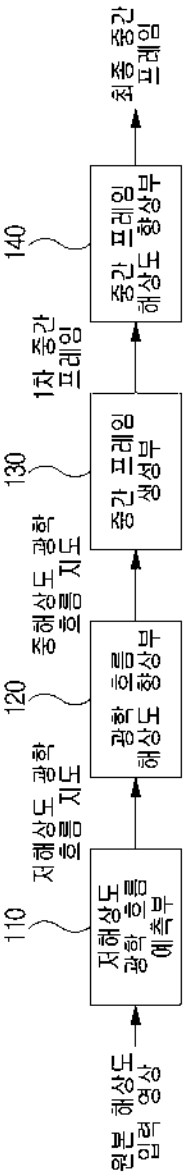
140 : 중간 프레임 해상도 향상부

도면

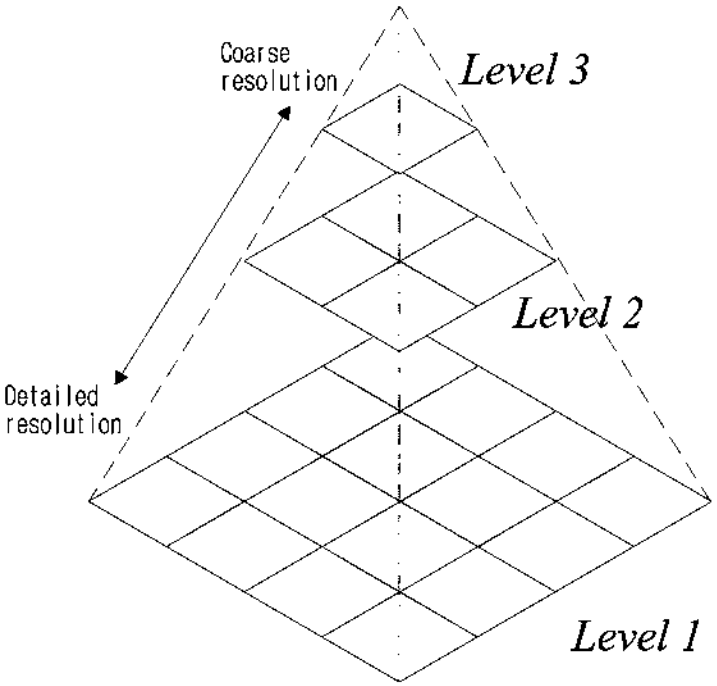
도면1



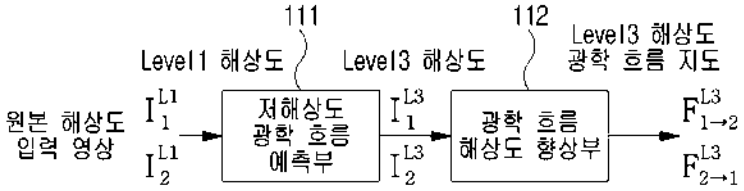
도면2


원본 해상도 입력 영상
110
저해상도 광학 흐름 지도
저해상도 광학 흐름 예측부
120
중해상도 광학 흐름 지도
광학 흐름 해상도 향상부
130
1차 중간 프레임
중간 프레임 생성부
140
중간 프레임 해상도 향상부
최종 중간 프레임

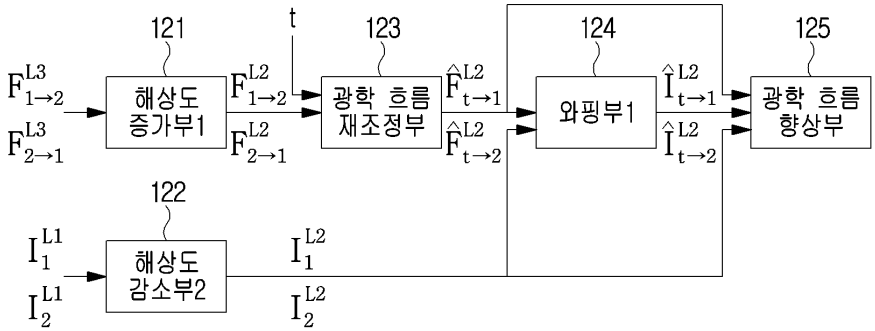
도면3


Coarse resolution
Level 3
Level 2
Detailed resolution
Level 1

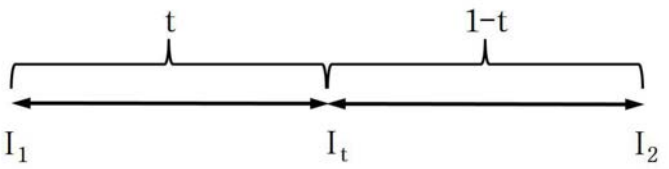
도면4


111
112
Level1 해상도
Level3 해상도
Level3 해상도 광학 흐름 지도
원본 해상도 입력 영상
I_1^{L1}
I_2^{L1}
저해상도 광학 흐름 예측부
I_1^{L3}
I_2^{L3}
광학 흐름 해상도 향상부
$F_{1\to2}^{L3}$
$F_{2\to1}^{L3}$

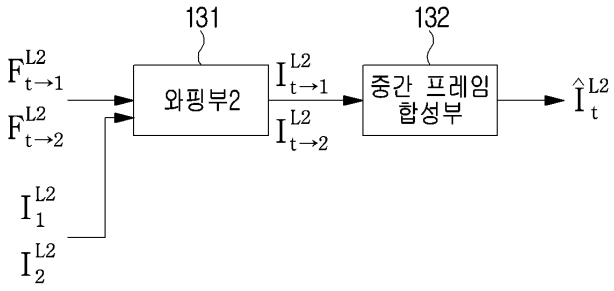
도면5


121
t
123
124
125
122
$F_{1\to2}^{L3}$
$F_{2\to1}^{L3}$
해상도 증가부1
$F_{1\to2}^{L2}$
$F_{2\to1}^{L2}$
광학 흐름 재조정부
$\hat{F}_{t\to1}^{L2}$
$\hat{F}_{t\to2}^{L2}$
와핑부1
$\hat{I}_{t\to1}^{L2}$
$\hat{I}_{t\to2}^{L2}$
광학 흐름 향상부
I_1^{L1}
I_2^{L1}
해상도 감소부2
I_1^{L2}
I_2^{L2}

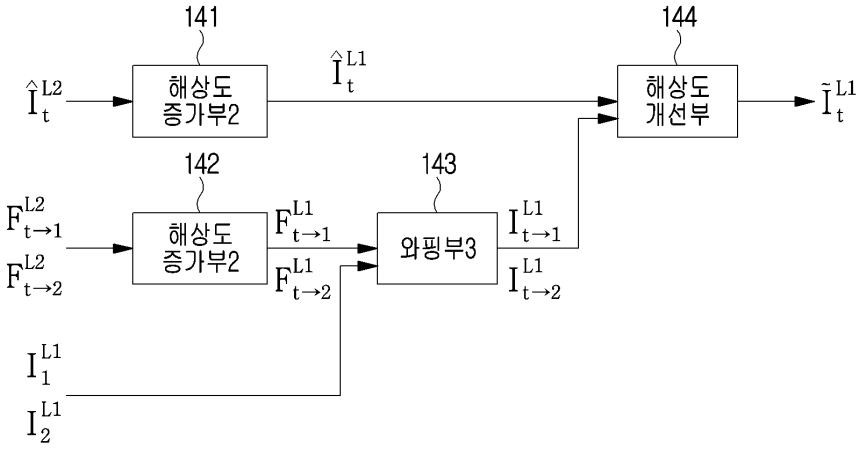
도면6



도면7



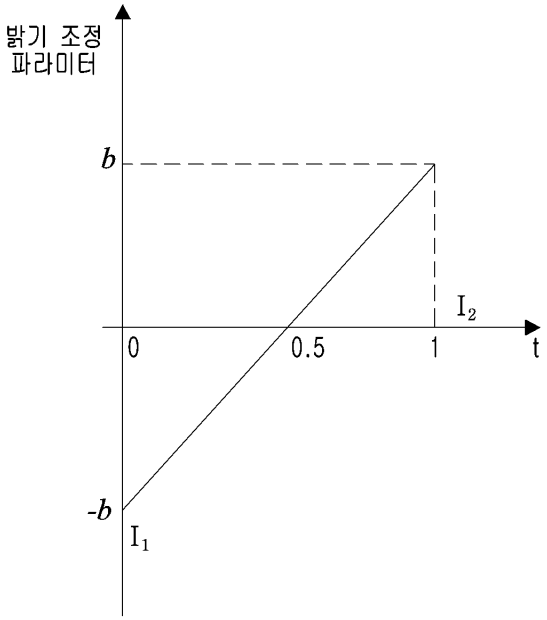
도면8



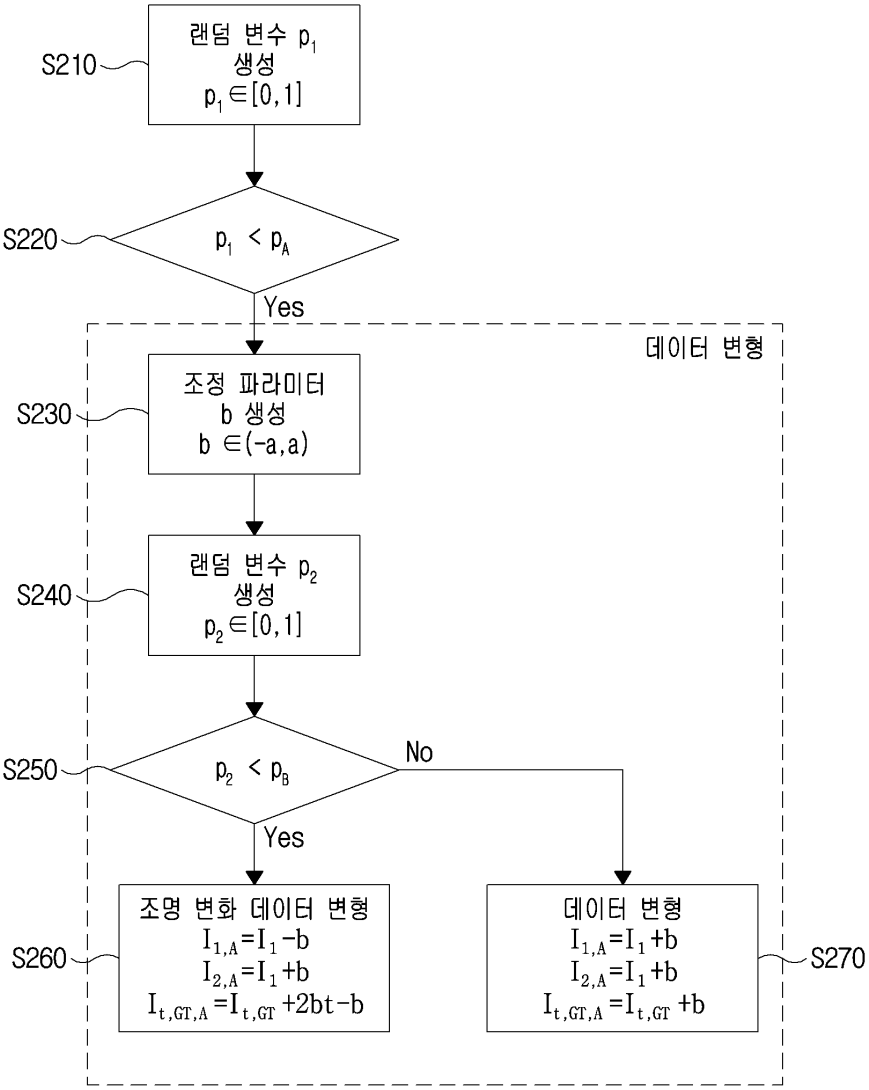
도면9

도면10

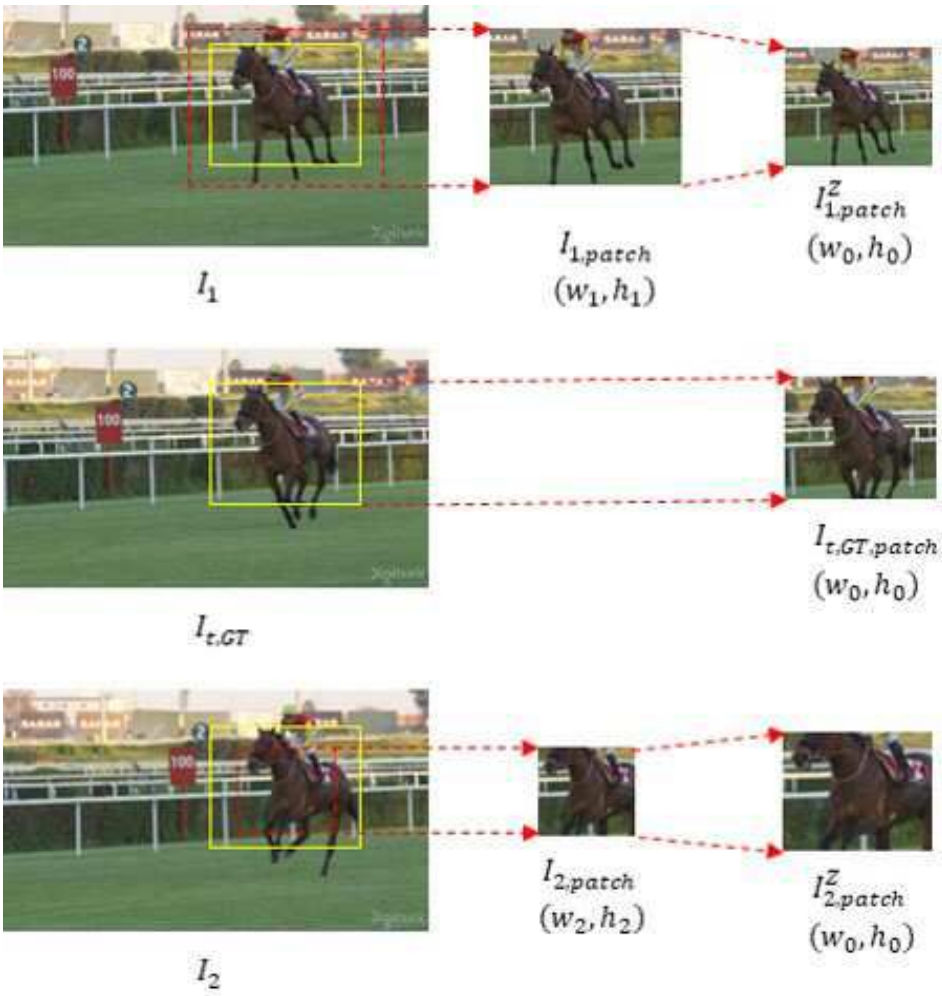
도면11



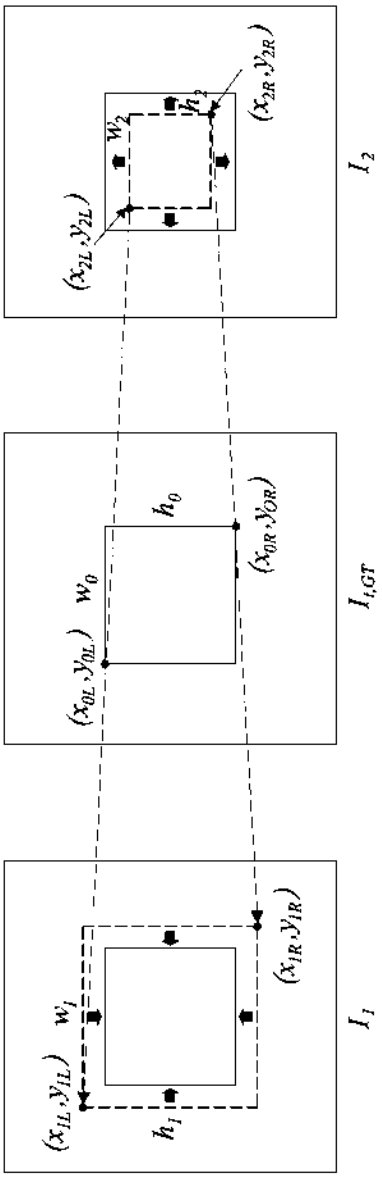
도면12


랜덤 변수 p_1 생성
$p_1 \in [0,1]$
S210
$p_1 < p_A$
S220
Yes
데이터 변형
조정 파라미터 b 생성
$b \in (-a,a)$
S230
랜덤 변수 p_2 생성
$p_2 \in [0,1]$
S240
$p_2 < p_B$
S250
No
Yes
조명 변화 데이터 변형
$I_{1,A}=I_1-b$
$I_{2,A}=I_1+b$
$I_{t,GT,A}=I_{t,GT}+2bt-b$
S260
데이터 변형
$I_{1,A}=I_1+b$
$I_{2,A}=I_1+b$
$I_{t,GT,A}=I_{t,GT}+b$
S270

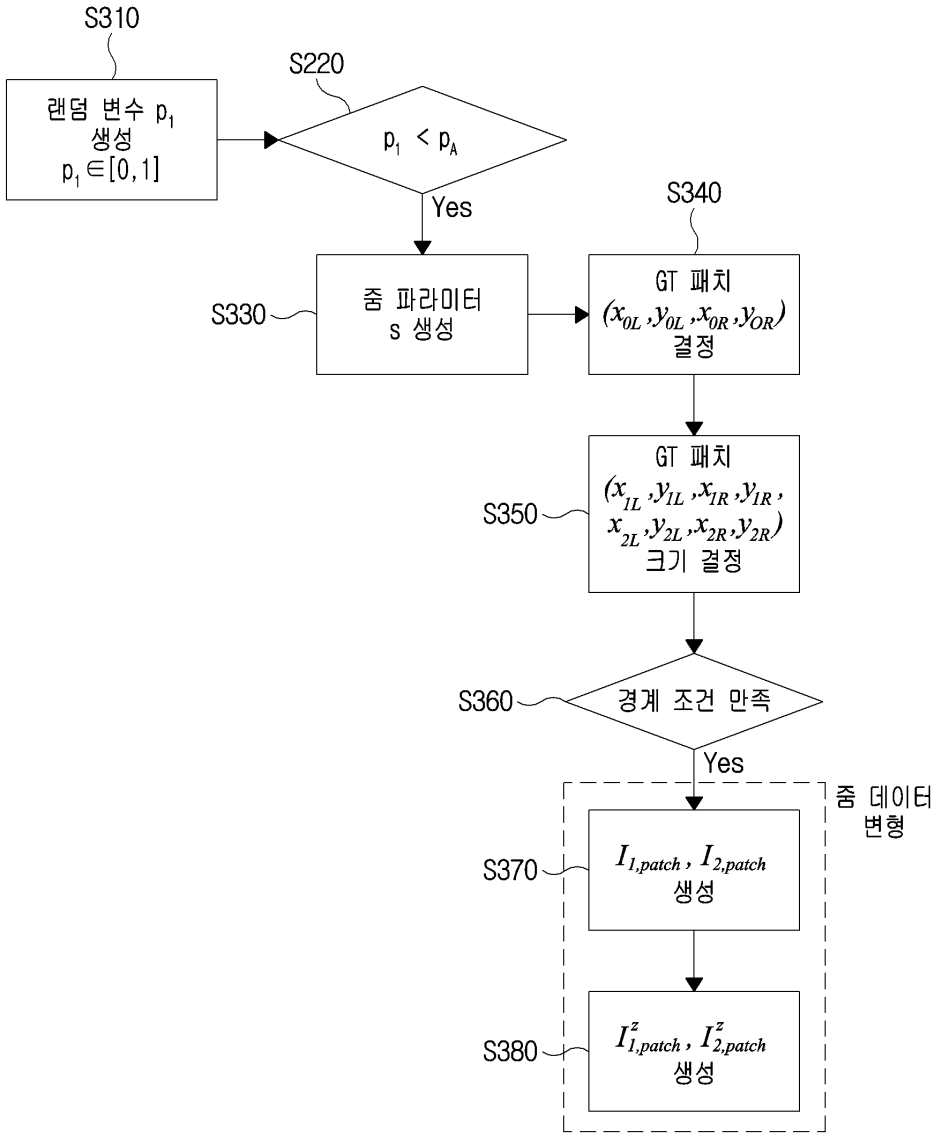
도면13


I_1
$I_{1,patch}$
(w_1, h_1)
$I^Z_{1,patch}$
(w_0, h_0)
$I_{t,GT}$
$I_{t,GT,patch}$
(w_0, h_0)
I_2
$I_{2,patch}$
(w_2, h_2)
$I^Z_{2,patch}$
(w_0, h_0)

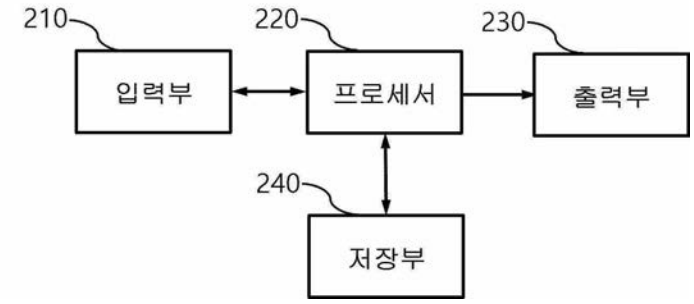
도면14


I_2
(x_{2L}, y_{2L})
w_2
h_2
(x_{2R}, y_{2R})
$I_{1,GT}$
(x_{0L}, y_{0L})
w_0
h_0
(x_{0R}, y_{0R})
I_1
(x_{1L}, y_{1L})
w_1
h_1
(x_{1R}, y_{1R})

도면15


S310
랜덤 변수 p_1 생성
$p_1 \in [0,1]$
S220
$p_1 < p_A$
Yes
S340
S330
줌 파라미터 s 생성
GT 패치 $(x_{0L}, y_{0L}, x_{0R}, y_{OR})$ 결정
GT 패치 $(x_{1L}, y_{1L}, x_{1R}, y_{1R}, x_{2L}, y_{2L}, x_{2R}, y_{2R})$ 크기 결정
S350
S360
경계 조건 만족
Yes
줌 데이터 변형
S370
$I_{1,patch}$, $I_{2,patch}$ 생성
S380
$I^z_{1,patch}$, $I^z_{2,patch}$ 생성

도면16


210
입력부
220
프로세서
230
출력부
240
저장부

■ 기술명 : 3D 영상 및 센서 기반의 보안 시스템 및 서비스 기술 (Surveillance System and Service Technology based on 3D Video and Sensor)

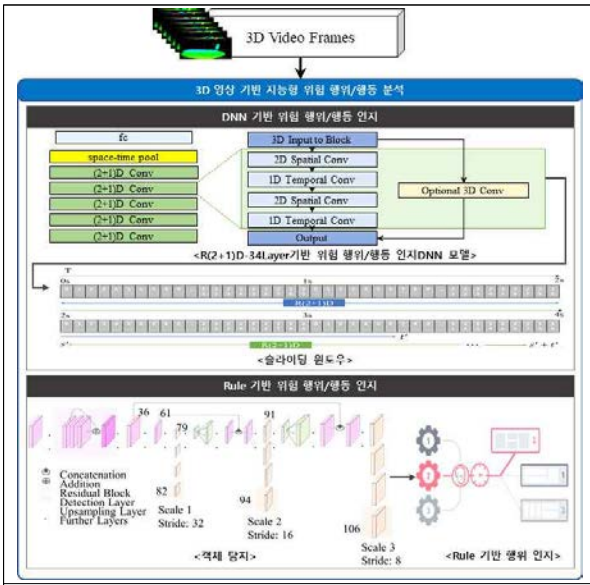
산업기술분류	정보통신- 지식정보보안 - 융합보안
Key-word(국문)	보안 시스템, 3D 영상 분석, 위치기반 서비스, 보안 이벤트 검출
Key-word(영문)	Surveillance System, 3D video Analysis, LBS, Security Event Detection

■ 기술의 개요

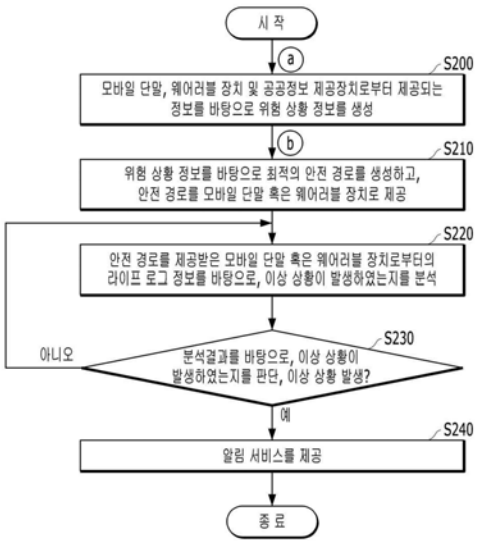
○ (배경) 기존 방식인 2D 기반 보안시스템은 보안 이벤트의 검출 정확도가 떨어지고 거짓 알람도 다수 발생하고 있어, 3D 영상 분석기반의 보안시스템에 대한 관심이 증대되는 실정임

○ (개요) 3D 영상 입력을 기반으로 객체의 움직임을 분석하고 이를 기반으로 보안 이벤트를 추출하는 지능형 보안 영상 분석 시스템과 사용자의 GPS 및 다종 센서 정보를 활용하여 안전 경로를 제공하고, 수집된 센서 정보를 분석하여 위험 상황 알림 및 경고를 발송하는 서비스 제공 기술임

< 기술 개요도 >



[3D 영상 기반 지능형 보안 영상 분석 시스템 구조]



[센서 기반 안전 경로 및 실시간 경보 제공 시스템 구조] (출처: 특허 10-2016-0175288)

■ 기술의 구현수준(TRL)

■ 기술의 장점(경쟁기술과의 차별성)

○ 지능형 보안 영상 분석 시스템 성능 향상

- 3D 영상 분석 기반 13개 보안 이벤트 검출율 : 67.5%
- 깊이(Depth) 정보 기반 다종 보안 이벤트 검출 가능 : 위장, 도구/흉기 등

○ 위치 기반 서비스와 연계된 사전 경고형 보안 서비스 제공 가능

- GPS, 온도, 적외선 센서 등과 연계한 사전 보안 경고 서비스

■ 활용범위 및 응용분야

	
[보안 감시 시스템]	[지능형 보안 영상 분석 시스템]

○ 보안 시스템 또는 보안 서비스 분야

○ 범죄 예방을 위한 실시간 경보 장치 분야

■ 지식재산권 현황

구분	발명의 명칭	출원번호 (출원일)	등록번호 (등록일)
1	인체 포즈 인지 시스템 및 방법	10-2015-0187330 (2015.12.28.)	-
2	사회 안전망 시스템 및 이의 서비스 제공방법	10-2016-0175288 (2016.12.21.)	-
3	깊이 맵 정보 기반의 인체 행위 인지 방법 및 그 장치	10-2017-0146080 (2017.11.03.)	-
4	3D 깊이 이미지 기반 객체 분리 장치 및 그 방법	10-2017-0145949 (2017.11.03.)	10-1967858 (2019.04.04.)
5	딥러닝 기반 이상 행위 인지 장치 및 방법	10-2018-0136243 (2018.11.08.)	-
6	하이브리드 기반 이상 행위 인지 시스템 및 그 방법	10-2019-0146821 (2019.11.15.)	-

(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2019년04월11일
(11) 등록번호 10-1967858
(24) 등록일자 2019년04월04일

(51) 국제특허분류(Int. Cl.)
G06T 7/11 (2017.01) *G06T 7/194* (2017.01)
G06T 7/50 (2017.01)
(52) CPC특허분류
G06T 7/11 (2017.01)
G06T 7/194 (2017.01)
(21) 출원번호 10-2017-0145949
(22) 출원일자 2017년11월03일
심사청구일자 2017년11월23일
(56) 선행기술조사문헌
US20140139633 A1*
KR1020110051416 A*
KR1020160011523 A
KR101755023 B1
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
김동칠
서울특별시 도봉구 우이천로20길 7, 103동 404호 (창동, 건영캐스빌아파트)
박성주
경기도 성남시 분당구 불정로 397, 315동 1003호 (서현동, 효자촌임광아파트)
(74) 대리인
특허법인지명

전체 청구항 수 : 총 10 항 심사관 : 진민숙

(54) 발명의 명칭 **3D 깊이 이미지 기반 객체 분리 장치 및 그 방법**

(57) 요 약

본 발명은 3D 카메라를 통해 얻은 Depth 이미지로부터 객체를 효과적으로 분리함으로써, 3D 이미지 데이터 기반 객체 검출, 이미지 이벤트 검출 등과 같은 이미지 보안 서비스에 활용할 수 있도록 한 3D 깊이 이미지 기반 객체 분리 장치 및 그 방법에 관한 것으로서, 상기 장치는, 3D 깊이 이미지를 입력하는 이미지 입력부; 및 입력된 3D 깊이 이미지의 깊이 정보를 이용하여 객체 위치 및 객체 수를 탐지한 후, 탐지된 객체들에 대한 픽셀값들중 유사 픽셀값을 가지는 픽셀간 그룹핑을 수행하고, 그룹핑된 객체 영역 이외의 배경 영역 및 천장 및 지면 영역을 순차적으로 제거하여 최종적인 객체를 추출하는 객체 분리부를 포함한다.

대 표 도

(52) CPC특허분류
G06T 7/50 (2017.01)
G06T 2207/10028 (2013.01)
G06T 2207/20084 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 1711055147
부처명 과학기술정보통신부
연구관리전문기관 정보통신기술진흥센터
연구사업명 SW컴퓨팅산업원천기술개발
연구과제명 (딥뷰-2세부) 대규모 실시간 비디오 분석에 의한 전역적 다중 관심객체 추적 및 상황 예측 기술 개발
기 여 율 1/1
주관기관 광주과학기술원
연구기간 2014.04.01 ~ 2024.02.28

명 세 서

청구범위

청구항 1

3D 깊이 이미지 기반 객체 분리 장치에 있어서,

3D 깊이 이미지를 입력하는 이미지 입력부; 및

입력된 3D 깊이 이미지의 깊이 정보를 이용하여 객체 위치 및 객체 수를 탐지한 후, 탐지된 객체들에 대한 픽셀값들중 유사 픽셀값을 가지는 픽셀간 그룹핑을 수행하고, 그룹핑된 객체 영역 이외의 배경 영역 및 천장 및 지면 영역을 순차적으로 제거하여 최종적인 객체를 추출하는 객체 분리부를 포함하되,

상기 객체 분리부는 CCA(Connected Component Analysis) 방식을 이용하여 배경 영역을 제거하는 배경 영역 제거부를 포함하고,

상기 배경 영역 제거부는 상기 3D 깊이 이미지에서 탐지된 객체에 대하여 모폴로지 침식 알고리즘을 적용한 후, 상기 모폴로지 침식 알고리즘이 적용된 상기 3D 깊이 이미지에 대하여 공간 필터링으로 미디안 필터를 적용하여 결정된 마스크 사이즈에 해당하는 이웃 픽셀값들을 크기 순서로 정렬하고 중간값을 선택한 다음, 모폴로지 팽창 알고리즘을 적용하여, 상기 객체를 제외한 배경 영역을 제거하는 것인 3D 깊이 이미지 기반 객체 분리 장치.

청구항 2

제1항에 있어서,

상기 이미지 입력부로부터 입력되는 3D 깊이 이미지의 상태를 판단하여 해당 3D 깊이 이미지에 노이즈가 포함된 경우, 해당 3D 깊이 이미지를 객체 분리부로 제공하는 이미지 상태 판단부를 더 포함하는 것인 3D 깊이 이미지 기반 객체 분리 장치.

청구항 3

제2항에 있어서,

상기 객체 분리부는,

노이즈가 포함된 3D 깊이 이미지로부터 객체의 위치 및 객체 수를 탐지하는 객체 탐지부 및

상기 배경 영역 제거부를 통해 배경 영역이 제거된 이미지로부터 천장 영역 및 지면 영역을 순차적으로 제거하여 최종적인 객체를 추출하는 천장 및 지면 영역 제거부를 포함하는 것인 3D 깊이 이미지 기반 객체 분리 장치.

청구항 4

제3항에 있어서,

상기 객체 탐지부는, 입력되는 3D 깊이 이미지의 깊이 정보를 통해 딥러닝 객체 탐지 방식을 이용하여 객체의 위치 및 객수를 탐지하여 객체의 위치 좌표값을 획득하여 저장하는 것인 3D 깊이 이미지 기반 객체 분리 장치.

청구항 5

삭제

청구항 6

제3항에 있어서,

상기 천장 및 지면 영역 제거부는, 배경영역 제거부에서 배경 영역이 제거된 깊이 이미지를 3차원 좌표로 변환한 후, 3차원 좌표들 중 임의로 3개 점을 선택하여 평면 방정식을 이용하여 평면을 결정하고,

천장 및 지면으로부터 특정 임계치까지를 평면이라고 정하고, 계산된 평면에서 임계치에 포함되는 점의 개수를 산출하며,

상기 동작을 일정 횟수 반복하여, 반복 횟수 개의 결과 중 가장 점을 많이 포함하는 평면을 지면 그리고 천장으로 결정하여 해당 영역을 제거하는 것인 3D 깊이 이미지 기반 객체 분리 장치.

청구항 7

3D 깊이 이미지 기반 객체 분리 방법에 있어서,

3D 깊이 이미지를 입력하는 단계; 및

입력된 3D 깊이 이미지의 깊이 정보를 이용하여 객체 위치 및 객체 수를 탐지한 후, 탐지된 객체들에 대한 픽셀값들중 유사 픽셀값을 가지는 픽셀간 그룹핑을 수행하고, 그룹핑된 객체 영역 이외의 배경 영역 및 천장 및 지면 영역을 순차적으로 제거하여 최종적인 객체를 추출하는 단계를 포함하되,

상기 객체를 추출하는 단계는 CCA(Connected Component Analysis) 방식을 이용하여 배경 영역(노이즈)을 제거하는 단계를 포함하고,

상기 배경 영역을 제거하는 단계는,

상기 3D 깊이 이미지에서 탐지된 객체에 대하여 모폴로지 침식 알고리즘을 적용한 후, 상기 모폴로지 침식 알고리즘이 적용된 상기 3D 깊이 이미지에 대하여 공간 필터링으로 미디안 필터를 적용하여 결정된 마스크 사이즈에 해당하는 이웃 픽셀값들을 크기 순서로 정렬하고 중간값을 선택한 다음, 모폴로지 팽창 알고리즘을 적용하여, 상기 객체를 제외한 노이즈(배경 영역)을 제거하는 것인 3D 깊이 이미지 기반 객체 분리 방법.

청구항 8

제7항에 있어서,

상기 입력되는 3D 깊이 이미지의 상태를 판단하는 단계를 더 포함하고, 상기 판단 결과, 해당 3D 깊이 이미지에 노이즈가 포함된 경우, 노이즈가 포함된 해당 3D 깊이 이미지를 이용하여 상기 객체를 추출하는 단계를 수행하는 것인 3D 깊이 이미지 기반 객체 분리 방법.

청구항 9

제8항에 있어서,

상기 객체를 추출하는 단계는,

노이즈가 포함된 3D 깊이 이미지로부터 객체의 위치 및 객체 수를 탐지하는 단계 및

상기 배경 영역이 제거된 이미지로부터 천장 영역 및 지면 영역을 순차적으로 제거하여 최종적인 객체를 추출하는 단계를 포함하는 것인 3D 깊이 이미지 기반 객체 분리 방법.

청구항 10

제9항에 있어서,

상기 객체의 위치 및 객체 수를 탐지하는 단계는, 입력되는 3D 깊이 이미지의 깊이 정보를 통해 딥러닝 객체 탐지 방식을 이용하여 객체의 위치 및 객수를 탐지하여 객체의 위치 좌표값을 획득하여 저장하는 것인 3D 깊이 이

미지 기반 객체 분리 방법.

청구항 11

삭제

청구항 12

제9항에 있어서,

상기 객체를 추출하는 단계는,

상기 배경영역 제거된 깊이 이미지를 3차원 좌표로 변환한 후, 3차원 좌표들 중 임의로 3개 점을 선택하여 평면 방정식을 이용하여 평면을 결정하는 단계;

천장 및 지면으로부터 특정 임계치까지를 평면이라고 결정하고, 상기 계산된 평면에서 임계치에 포함되는 점의 개수를 산출하는 단계; 및

상기 단계들을 일정 횟수 반복하여, 반복 횟수 개의 결과 중 가장 점을 많이 포함하는 평면을 지면 그리고 천장으로 결정하여 해당 영역을 제거하는 단계를 포함하는 것인 3D 깊이 이미지 기반 객체 분리 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 3D 깊이(3 Dimension Depth) 기반 이미지 기반 객체 분리 장치 및 그 방법에 관한 것으로서, 특히 3D 카메라를 통해 얻은 깊이 이미지로부터 객체를 효과적으로 분리함으로써, 3D 이미지 데이터 기반 객체 검출, 이미지 이벤트 검출 등과 같은 이미지 보안 서비스에 활용할 수 있도록 한 3D 깊이 이미지 기반 객체 분리 장치 및 그 방법에 관한 것이다.

배 경 기 술

[0002] 일반적으로, 카메라를 이용하여 사용자의 손 동작, 즉 제스처를 인식하는 다양한 어플리케이션이 제시되었다. 어플리케이션은 사용자의 제스처의 의미를 정확하게 인식함에 따라 인식된 의미에 대응하는 명령 또는 처리의 과정을 수행한다.

[0003] 이러한 제스처의 인식은 인간-컴퓨터 상호작용 분야에서 주로 연구되어 왔다. 스마트폰을 이용하는 원격제어, 의사소통, 게임 및 어플리케이션의 실행과 같은 제스처의 인식에 관련된 다양한 시도가 이루어지고 있다.

[0004] 그러나, 카메라를 이용해서 사용자의 제스처를 인식하는 다양한 방법이 제시되었음에도 불구하고, 경우에 따라서 이러한 제스처의 인식은 제한되거나 정확하지 않은 결과를 도출할 수 있다. 예를 들면, 사용자의 손이 사용자의 얼굴 또는 사용자의 피부색과 유사한 색을 갖는 다른 객체와 중첩될 경우, 카메라에 의해 획득된 이미지의 정보에서 사용자의 손을 얼굴 또는 다른 객체로부터 분리하는 것이 어렵고, 따라서 사용자의 실제의 제스처를 정확하게 인식할 수 없다는 문제가 발생할 수 있다. 즉, 정확한 객체를 검출하여야만 해당 객체의 제스쳐 등을 인지할 수 는 것이다.

[0005] 또한, 일반적인 2D 기반 객체 분리 기술은 주변 환경의 색상으로 인한 오류, 야간에 객체 인지의 어려움, 조명/화질의 열화 등으로 인한 오류 등 다양한 환경 변화에서 배경으로부터 객체 분리를 하는데 어려움이 있으며, 이로 인해 이미지 보안 시스템에서 2D 이미지를 기반으로 이벤트 검출을 했을 때 정확도가 낮은 문제점이 발생하게 된다.

발명의 내용

해결하려는 과제

[0006] 따라서, 본 발명은 상기한 문제점을 해결하기 위한 것으로, 본 발명의 목적은, 다양한 환경 변화가 존재하는 환경에서 3D 카메라로부터 입력되는 깊이 이미지로부터 객체를 효과적으로 분리함으로써, 3D 이미지 데이터를 기반으로 이미지 이벤트를 검출하는 이미지 보안 시스템에 적용하여 위험 이벤트 인식 성능 정확도를 향상시킬 수 있도록 한 3D 깊이 이미지 기반 객체 분리 장치 및 그 방법을 제공함에 있다.

과제의 해결 수단

[0007] 상기한 목적을 달성하기 위한 본 발명의 일 측면에 따른 3D 깊이 이미지 기반 객체 분리 장치에 따르면, 3D 깊이 이미지를 입력하는 이미지 입력부; 및 입력된 3D 깊이 이미지의 깊이 정보를 이용하여 객체 위치 및 객체 수를 탐지한 후, 탐지된 객체들에 대한 픽셀값들중 유사 픽셀값을 가지는 픽셀간 그룹핑을 수행하고, 그룹핑된 객체 영역 이외의 배경 영역 및 천장 및 지면 영역을 순차적으로 제거하여 최종적인 객체를 추출하는 객체 분리부를 포함할 수 있다.

[0008] 상기 장치는, 상기 이미지 입력부로부터 입력되는 3D 깊이 이미지의 상태를 판단하여 해당 3D 깊이 이미지에 노이즈가 포함된 경우, 해당 3D 깊이 이미지를 객체 분리부로 제공하는 이미지 상태 판단부를 더 포함할 수 있다.

[0009] 상기 객체 분리부는, 노이즈가 포함된 3D 깊이 이미지로부터 객체의 위치 및 객체 수를 탐지하는 객체 탐지부; 객체 탐지부로부터 탐지된 객체에 대한 픽셀 유사도를 이용하여 객체를 제외한 배경 영역(노이즈)을 제거하는 배경 영역 제거부; 및 상기 배경 영역 제거부를 통해 배경 영역이 제거된 이미지로부터 천장 영역 및 지면 영역을 순차적으로 제거하여 최종적인 객체를 추출하는 천장 및 지면 영역 제거부를 포함할 수 있다.

[0010] 상기 객체 탐지부는, 입력되는 3D 깊이 이미지의 깊이 정보를 통해 딥러닝 객체 탐지 방식을 이용하여 객체의 위치 및 객수를 탐지하여 객체의 위치 좌표값을 획득하여 저장할 수 있다.

[0011] 상기 배경 영역 제거부는, CCA(Connected Component Analysis) 방식을 이용하여 배경 영역(노이즈)을 제거할 수 있다.

[0012] 상기 천장 및 지면 영역 제거부는, 배경영역 제거부에서 배경 영역이 제거된 깊이 이미지를 3차원 좌표로 변환한 후, 3차원 좌표들 중 임의로 3개 점을 선택하여 평면 방정식을 이용하여 평면을 결정하고, 천장 및 지면으로부터 특정 임계치까지를 평면이라고 결정하고, 계산된 평면에서 임계치에 포함되는 점의 개수를 산출하며, 상기 동작을 일정 횟수 반복하여, 반복 횟수 개의 결과 중 가장 점을 많이 포함하는 평면을 지면 그리고 천장으로 결정하여 해당 영역을 제거할 수 있다.

[0013] 한편, 상기한 목적을 달성하기 위한 본 발명의 3D 깊이 이미지 기반 객체 분리 방법에 따르면, 3D 깊이 이미지를 입력하는 단계; 및 입력된 3D 깊이 이미지의 깊이 정보를 이용하여 객체 위치 및 객체 수를 탐지한 후, 탐지된 객체들에 대한 픽셀값들중 유사 픽셀값을 가지는 픽셀간 그룹핑을 수행하고, 그룹핑된 객체 영역 이외의 배경 영역 및 천장 및 지면 영역을 순차적으로 제거하여 최종적인 객체를 추출하는 단계를 포함할 수 있다.

[0014] 상기 방법은, 상기 입력되는 3D 깊이 이미지의 상태를 판단하는 단계를 더 포함하며, 상기 판단 결과, 해당 3D 깊이 이미지에 노이즈가 포함된 경우, 노이즈가 포함된 해당 3D 깊이 이미지를 이용하여 상기 객체를 추출하는 단계를 수행한다.

[0015] 상기 객체를 추출하는 단계는, 노이즈가 포함된 3D 깊이 이미지로부터 객체의 위치 및 객체 수를 탐지하는 단계; 상기 탐지된 객체에 대한 픽셀 유사도를 이용하여 객체를 제외한 배경 영역(노이즈)을 제거하는 단계; 및 상기 배경 영역이 제거된 이미지로부터 천장 영역 및 지면 영역을 순차적으로 제거하여 최종적인 객체를 추출하는 단계를 포함할 수 있다.

[0016] 상기 객체의 위치 및 객체 수를 탐지하는 단계는, 입력되는 3D 깊이 이미지의 깊이 정보를 통해 딥러닝 객체 탐지 방식을 이용하여 객체의 위치 및 객수를 탐지하여 객체의 위치 좌표값을 획득하여 저장할 수 있다.

[0017] 상기 배경 영역을 제거하는 단계는, CCA(Connected Component Analysis) 방식을 이용하여 배경 영역(노이즈)을 제거할 수 있다.

[0018] 상기 객체를 추출하는 단계는, 상기 배경영역 제거된 깊이 이미지를 3차원 좌표로 변환한 후, 3차원 좌표들 중 임의로 3개 점을 선택하여 평면 방정식을 이용하여 평면을 결정하는 단계; 천장 및 지면으로부터 특정 임계치까

지를 평면이라고 결정하고, 상기 계산된 평면에서 임계치에 포함되는 점의 개수를 산출하는 단계; 및 상기 단계들을 일정 횟수 반복하여, 반복 횟수 개의 결과 중 가장 점을 많이 포함하는 평면을 지면 그리고 천장으로 결정하여 해당 영역을 제거하는 단계를 포함할 수 있다.

발명의 효과

[0019] 본 발명에 따르면, 다양한 환경 변화가 존재하는 환경에서 3D 카메라로부터 입력되는 Depth 이미지로부터 객체를 효과적으로 분리함으로써, 3D 이미지 데이터를 기반으로 이미지 이벤트를 검출하는 이미지 보안 시스템에 적용하여 위험 이벤트 인식 성능 정확도를 향상시킬 수 있다.

[0020] 또한, 본 발명에 따른 다양한 환경 변화가 존재하는 환경에서 3D 카메라로부터 입력되는 깊이 이미지의 배경으로부터 객체를 효과적으로 분리함으로써, 3D 이미지 기반 인체 행위 인지 시스템의 전처리 수행 시 효과적으로 적용이 가능하며, 이미지 보안 시스템에서 관심 객체를 추출하고, 객체의 행동을 효과적으로 분석할 수 있는 효과를 가진 것이다.

도면의 간단한 설명

[0021] 도 1은 본 발명에 따른 3D 깊이 이미지 기반 객체 분리 장치에 대한 개략적인 블록 구성을 나타낸 도면.

도 2는 도 1에 도시된 객체 분리부에 대한 상세 블록 구성을 나타낸 도면.

도 3은 본 발명에 따른 3D 깊이 이미지 기반 객체 분리 예시 화면을 나타낸 도면.

도 4는 본 발명에 따른 3D 깊이 이미지 기반 객체 분리 방법에 대한 동작 플로우챠트를 나타낸 도면.

발명을 실시하기 위한 구체적인 내용

[0022] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시 예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 것이며, 단지 본 실시 예들은 본 발명의 개시가 완전하도록 하며, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 용이하게 이해할 수 있도록 제공되는 것이며, 본 발명은 청구항의 기재에 의해 정의된다. 한편, 본 명세서에서 사용된 용어는 실시 예들을 설명하기 위한 것이며 본 발명을 제한하고자 하는 것은 아니다. 본 명세서에서, 단수형은 문구에서 특별히 언급하지 않는 한 복수형도 포함한다. 명세서에서 사용되는 "포함한다(comprises)" 또는 "포함하는(comprising)"은 언급된 구성요소, 단계, 동작 및/또는 소자 이외의 하나 이상의 다른 구성요소, 단계, 동작 및/또는 소자의 존재 또는 추가를 배제하지 않는다.

[0023] 먼저, 본 발명의 구체적인 설명에 앞서, 3D 이미지 기반 행위 인지 기술 분야에서 배경으로부터 객체를 분리하는 본 발명은 객체 검출, 객체 추적, 이미지 이벤트 검출 등 이미지 보안 서비스를 정확하게 수행하기 위한 핵심 요소이다.

[0024] 이하, 본 발명에 따른 3D 깊이 이미지 기반 객체 분리 장치 및 방법에 대한 구체적인 실시예에 대하여 첨부한 도면을 참조하여 상세하게 설명하기로 한다.

[0025] 도 1은 본 발명에 따른 3D 깊이 이미지 기반 객체 분리 장치에 대한 개략적인 블록 구성을 나타낸 도면이다.

[0026] 도 1에 도시된 바와 같이, 본 발명에 따른 3D 깊이 이미지 기반 객체 분리장치는, 이미지 입력부(10), 이미지 상태 판단부(20) 및 객체 분리부(30)를 포함할 수 있다.

[0027] 이미지 입력부(10)는 획득한 3D 깊이 이미지를 이미지 상태 판단부(20)로 입력한다. 여기서, 이미지 입력부(10)는 적어도 하나 이상의 3D 카메라, 3D 이미지 센서 중 적어도 하나 일 수 있으며, 3D 깊이 이미지에는 깊이 정보를 포함할 수 있다.

[0028] 이미지 상태 판단부(20)는 이미지 입력부(10)로부터 입력되는 3D 깊이 이미지의 이미지 상태를 판단하여 이미지 상태가 좋지 않을 경우 즉, 노이즈가 많거나 배경이 복잡한 상황인 경우 이하의 객체 분리 동작을 수행한다. 만

약, 이미지 상태가 좋은 경우 별도의 객체 분리 동작을 수행하지 않고 입력된 깊이 이미지의 객체만으로 행위인지 등 다양한 동작을 수행하는 것이다.

[0029] 만약, 이미지 상태 판단부(20)에서 이미지 상태 판단 결과, 이미지 상태가 좋지 않다고 판단되는 경우, 객체 분리부(30)는 도 3 (a)와 같이 입력되는 깊이 이미지로부터 객체 위치 및 객체 수를 탐지한 후, 탐지된 객체들에 대한 픽셀값들중 유사 픽셀값을 가지는 픽셀간 그룹핑을 수행한다.

[0030] 그리고, 객체 분리부(30)는 픽셀간 그룹핑이 이루어진 후, 패당 그룹 이외의 이미지 영역을 제거한 후, 해당 객체의 천장 및 지면 영역을 제거함으로써 최종적인 원하는 객체를 분리 추출하는 것이다.

[0031] 이하, 상기 설명한 도 1에 도시된 객체 분리부(30)에 대한 상세 구성 및 동작에 대하여 도 2 및 도 3을 참조하여 살펴보기로 하자.

[0032] 도 2는 도 1에 도시된 객체 분리부에 대한 상세 블록 구성을 나타낸 도면, 도 3은 본 발명에 따른 3D 깊이 이미지 기반 객체 분리 예시 화면을 나타낸 도면이다.

[0033] 도 2에 도시된 바와 같이, 객체 분리부(30)는 객체 탐지부(31), 배경 영역 제거부(32), 천장 영역 제거부(33), 지면 영역 제거부(34) 및 객체 추출부(35)를 포함할 수 있다.

[0034] 객체 탐지부(31)는 도 3 (a)와 같이 입력되는 노이즈가 포함된 깊이 이미지로부터 깊이 정보를 이용하여 깊이 이미지내 객체들의 위치 및 객체의 수를 탐지한 후, 탐지된 객체의 좌표값을 획득한다. 여기서, 객체들의 위치 및 객체의 수를 탐지하는 방법으로, 객체 검출 속도가 빠른 딥러닝 기반 객체 탐지 방법을 사용할 수 있다. 즉, 입력되는 깊이 이미지에 대한 깊이 정보를 기반으로 다양한 깊이 정보를 기반으로 학습되어 모델링된 모델링 값에 따라 객체들의 위치 및 객체의 수를 탐지한 후, 탐지된 객체들에 대한 좌표값을 획득하는 것이다. 이때, 상기한 딥러닝 기반 객체 탐지 방법은 이미 공지된 일반적인 방법으로서, 구체적인 방법에 대한 설명은 생략하며, 이러한 방식 이외의 방법을 이용해서 객체를 탐지하여도 무방함을 이해해야 할 것이다. 여기서, 객체 탐지부(31)에서 객체 탐지 결과에 대한 화면은 도 3 (b)와 같다.

[0035] 배경 영역 제거부(32)는 상기 객체 탐지부(31)에서 탐지된 객체에 대한 좌표값을 기준으로 하여 CCA(Connected Component Analysis) 알고리즘을 수행한다. 즉, CCA 알고리즘은, 상기 객체 탐지부(31)에서 획득한 해당 객체의 좌표값에 해당되는 픽셀을 기준으로 연결된 유사 픽셀값을 분석하여 유사 픽셀값을 가지는 픽셀간 그룹핑을 수행하는 것이다.

[0036] CCA 알고리즘은, 입력되는 이미지 안의 물체를 찾아내어 분석하기 위한 목적으로 쓰레숄딩(Thresholding)에 기반한 이진화 기술들이 개발되었으며, 이진화 결과에서 노이즈를 제거하고, 결과들을 보다 정교화하기 위해서 mathematical morphology 기술들이 개발되었다. Morphology까지 거친 이미지들을 이용하면 물체를 구분하고, 물체를 인식할 수 있는데, 그 전에 픽셀들을 객체 단위로 구분하여야 한다. 하나의 객체를 구성하는 픽셀들은 이미지에서 서로 연결된 형태로 나타나게 되는데 이와 같이 서로 연결된 유사 픽셀들을 그룹핑하여 객체를 구분하는 분석 방법이 CCA 알고리즘이다. 여기서, 상기한 CCA 알고리즘에 대하여 좀 더 구체적으로 살펴보면, 다음과 같다.

[0037] CCA 알고리즘은, 공간 필터링 방식을 수행하는 미디안(Median) 필터를 이용하여 객체 이외의 배경 영역(노이즈)을 제거할 수 있다. 여기서, 미디안 필터는 마스크 사이즈를 결정하고, 결정된 마스크 사이즈에 해당하는 이웃 픽셀 값들을 크기 순서로 정렬하고, 그 중 중간 값을 선택하는 역할을 한다. 다만, 배경영역 제거부(32)에서 미디안 필터만을 사용한다면, 깊이 이미지에 포함된 객체들의 경계가 무너질 것이다.

[0038] 따라서, 배경영역 제거부(32)는 미디안 필터 이외에 모폴로지 침식(Erosion) 알고리즘 및 모폴로지 팽창(Dilation) 알고리즘을 추가적으로 이용할 수도 있다. 모폴로지 침식 알고리즘은 객체나 작은 크기의 파티클(Particle)로부터 레이어(Layer)를 벗기거나, 이미지에 존재하는 부적절한 픽셀 또는 작은 크기의 파티클을 제거하는 역할을 한다. 이와 반대로, 모폴로지 팽창 연산은 객체나 작은 크기의 파티클로부터 레이어를 확장시키거나, 모폴로지 침식 알고리즘에 의해 작아진 파티클을 원래 사이즈로 확장시키는 역할을 한다.

[0039] 배경영역 제거부(32)는 상기 획득된 깊이 이미지에 모폴로지 침식 알고리즘을 적용한 후, 공간 필터링으로서 미디안 필터를 적용하고 나서, 모폴로지 팽창(dilation) 알고리즘을 적용함으로서 효과적으로 깊이 이미지의 노이즈를 최소화 시킬 수 있다.

[0040] 결국, 배경영역 제거부(32)는 탐지된 객체를 기준으로 한 주변 배경 영역을 제거한 후, 배경 영역이 제거된 이

미지를 천장 영역 제거부(33)로 제공한다.

[0041] 천장 영역 제거부(33)는 상기 배경영역 제거부(32)를 통해 배경영역이 제거된 이미지로부터 도 3 (c)와 같이 천장 영역을 제거한다. 즉, 배경영역 제거부(32)에서 CCA 알고리즘을 수행한 후 객체 영역을 제외한 나머지 배경영역을 제거하더라도 객체의 머리 위쪽 부분은 객체 영역이라고 판단하여 제거되지 않을 확률이 높다. 따라서, 천장 영역 제거부(33)에서 평면 방정식을 통해 천장 영역을 계산하여 천장 영역을 제거하는 것이다.

[0042] 지면 영역 제거부(34)는 도 3 (c)와 같이 상기 천장 영역 제거부(33)를 통해 천장 영역이 제거된 이미지로부터 객체를 기준으로 한 지면 영역을 도 3 (d)와 같이 제거하게 된다. 즉, 상기 배경영역 제거부(32)에서 CCA 알고리즘 동작 후 객체 영역을 제외한 나머지 영역을 제거하더라도 객체의 신발 아래쪽 부분은 객체 영역이라고 판단하여 제거되지 않을 확률이 높기 때문에 평면 방정식을 통해 지면 영역을 계산하여 지면 영역을 제거하는 것이다.

[0043] 상기한 천장 영역 제거부(33) 및 지면 영역 제거부(34)에서 천장 및 지면 영역의 제거하는 방법을 좀 더 구체적으로 살펴보면, 먼저, 배경영역 제거부(31)에서 CCA 알고리즘을 통해 배경 영역이 제거된 깊이 이미지를 3차원 좌표로 변환하고, 3차원 좌표들 중 임의로 3개 점을 선택하여 평면 방정식을 이용하여 평면을 결정한다.

[0044] 그리고, 천장 및 지면으로부터 특정 임계치까지를 평면이라고 정하고, 계산된 평면에서 임계치에 포함되는 점의 개수를 구한다.

[0045] 이어, 이와 같은 동작을 일정 횟수 반복하고, 반복 횟수 개의 결과 중 가장 점을 많이 포함하는 평면을 지면 그리고 천장으로 결정하여 해당 영역을 제거하는 것이다.

[0046] 상기에서는 천장 영역을 제거한 후, 지면 영역을 제거하는 순서로 설명하였으나, 지면 영역을 먼저 제거한 후, 천장 영역을 제거하는 순서로 동작할 수도 있음을 이해해야 할 것이다.

[0047] 객체 추출부(35)는 상기와 같이 천장 영역과 지면 영역이 제거된 이미지로부터 실질적인 객체만을 추출하게 되는 것이다.

[0048] 상기와 같은 본 발명에 따른 3D 깊이 이미지 기반 객체 분리 장치의 동작과 상응하는 본 발명에 따른 3D 깊이 이미지 기반 객체 분리 방법에 대하여 첨부된 도 4를 참조하여 단계적으로 설명해 보기로 하자.

[0049] 도 4는 본 발명에 따른 3D 깊이 이미지 기반 객체 분리 방법은, 먼저, 3D 이미지 센서 또는 3D 카메라를 통해 획득된 3D 깊이 이미지가 입력되면(S401), 입력되는 3D 깊이 이미지의 이미지 상태를 판단한다. 즉, 입력되는 3D 깊이 이미지에 노이즈가 포함되어 있는지를 판단한다(S402).

[0050] 판단 결과, 이미지 상태가 좋지 않을 경우 즉, 노이즈가 많거나 배경이 복잡한 상황인 경우 이하의 객체 분리 동작을 수행한다. 만약, 이미지 상태가 좋은 경우 별도의 객체 분리 동작을 수행하지 않고 해당 이미지로부터 객체를 바로 추출하여(S407) 추출된 객체를 이용하여 행위 인지 등 다양한 동작을 수행하는 것이다.

[0051] 상기 S402 단계에서, 입력되는 3D 깊이 이미지에 노이즈가 포함되어 있는 경우, 입력되는 깊이 이미지로부터 깊이 정보를 이용하여 깊이 이미지내 객체들의 위치 및 객체의 수를 탐지한 후, 탐지된 객체의 좌표값을 획득한다(S403). 여기서, 객체들의 위치 및 객체의 수를 탐지하는 방법으로, 객체 검출 속도가 빠른 딥러닝 기반 객체 탐지 방법을 사용할 수 있다.

[0052] 상기 S403단계에서 객체 위치 및 객체수를 탐지한 후, 탐지된 객체에 대한 좌표값을 기준으로 하여 CCA 알고리즘을 수행하여 깊이 이미지의 노이즈(배경 영역)을 제거한다(S404). 즉, CCA 알고리즘은, 상기 S403단계에서 획득한 해당 객체의 좌표값에 해당되는 픽셀을 기준으로 연결된 유사 픽셀값을 분석하여 유사 픽셀값을 가지는 픽셀간 그룹핑을 수행하여 탐지된 객체를 기준으로 한 주변 배경 영역을 제거하는 것이다.

[0053] 이어, S404단계에서 CCA 알고리즘을 통해 배경영역(노이즈)이 제거되면, 배경 영역이 제거된 이미지로부터 평면 방정식을 이용하여 천장 영역을 제거한다(S405). 즉, 상기 S404단계에서 CCA 알고리즘을 통해 배경영역이 제거된 이미지로부터 도 3 (c)와 같이 천장 영역을 제거한다. 여기서, CCA 알고리즘을 수행한 후 객체 영역을 제외한 나머지 배경영역을 제거하더라도 객체의 머리 위쪽 부분은 객체 영역이라고 판단하여 제거되지 않을 확률이 높기 때문에 평면 방정식을 통해 천장 영역을 계산하여 천장 영역을 제거하는 것이다.

[0054] S405 단계에서, 천장 영역이 제거되면, 천장 영역이 제거된 이미지로부터 객체를 기준으로 한 지면 영역을 도 3

(d)와 같이 제거하게 된다(S406). 즉, 상기 S404단계에서 CCA 알고리즘 수행 후 객체 영역을 제외한 나머지 영역을 제거하더라도 객체의 신발 아래쪽 부분은 객체 영역이라고 판단하여 제거되지 않을 확률이 높기 때문에 평면 방정식을 통해 지면 영역을 계산하여 지면 영역을 제거하는 것이다.

[0055] 상기한 천장 영역 및 지면 영역의 제거하는 방법은, 먼저 CCA 알고리즘을 통해 배경 영역이 제거된 깊이 이미지를 3차원 좌표로 변환하고, 3차원 좌표들 중 임의로 3개 점을 선택하여 평면 방정식을 이용하여 평면을 결정한다.

[0056] 그리고, 천장 및 지면으로부터 특정 임계치까지를 평면이라고 정하고, 계산된 평면에서 임계치에 포함되는 점의 개수를 구한 후, 이와 같은 동작을 일정 횟수 반복하고, 반복 횟수 개의 결과 중 가장 점을 많이 포함하는 평면을 지면 그리고 천장으로 결정하여 해당 영역을 제거하는 것이다.

[0057] 상기의 실시예에서는 천장 영역을 먼저 제거한 후, 지면 영역을 제거하는 순서로 설명하였으나, 지면 영역을 먼저 제거한 후, 천장 영역을 제거할 수도 있음을 이해해야 할 것이다.

[0058] 이와 같이, S405, S406 단계를 통해 천장 영역과 지면 영역이 제거되면 실질적인 객체만이 추출되는 것이다(S407).

[0059] 본 발명에 따른 3D 깊이 이미지 기반 객체 분리 장치 및 방법을 실시예에 따라 설명하였지만, 본 발명의 범위는 특정 실시 예에 한정되는 것은 아니며, 본 발명과 관련하여 통상의 지식을 가진 자에게 자명한 범위 내에서 여러 가지의 대안, 수정 및 변경하여 실시할 수 있다.

[0060] 따라서, 본 발명에 기재된 실시 예 및 첨부된 도면들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시 예 및 첨부된 도면에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리 범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

[0061] 10 : 이미지 입력부

20 : 이미지 상태 판단부

30 : 객체 분리부

31 : 객체 탐지부

32 : 배경영역 제거부

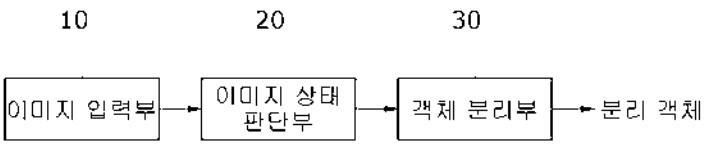
33 : 천장 영역 제거부

34 : 지면 영역 제거부

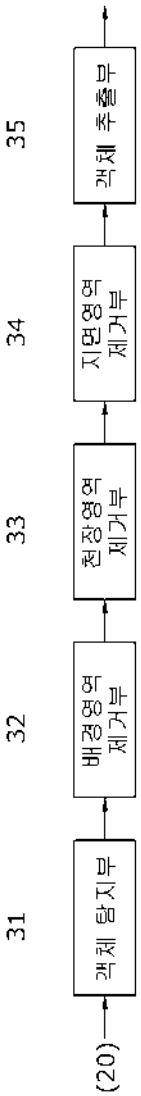
35 : 객체 추출부

도면

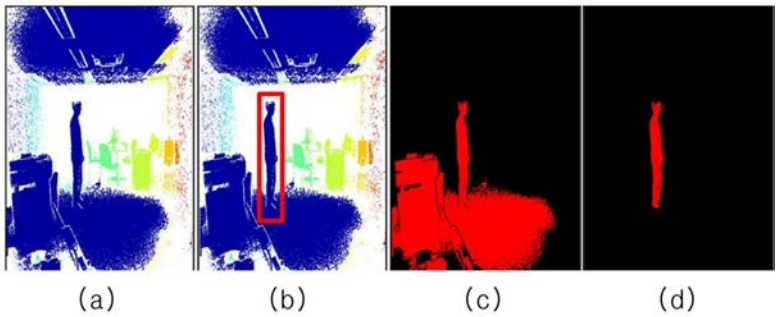
도면1



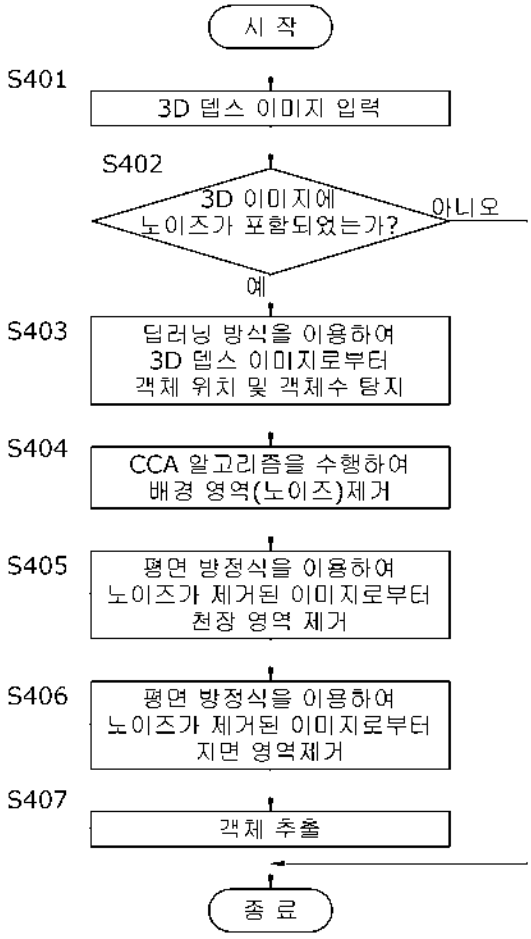
도면2



도면3



도면4


시 작
S401
3D 뎁스 이미지 입력
S402
3D 이미지에
노이즈가 포함되었는가?
아니오
예
S403
딥러닝 방식을 이용하여
3D 뎁스 이미지로부터
객체 위치 및 객체수 탐지
S404
CCA 알고리즘을 수행하여
배경 영역(노이즈)제거
S405
평면 방정식을 이용하여
노이즈가 제거된 이미지로부터
천장 영역 제거
S406
평면 방정식을 이용하여
노이즈가 제거된 이미지로부터
지면 영역제거
S407
객체 추출
종 료

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0050551
(43) 공개일자 2019년05월13일

(51) 국제특허분류(Int. Cl.)
G06K 9/00 (2006.01) *G06K 9/48* (2006.01)
G06K 9/62 (2006.01) *G06N 3/08* (2006.01)
(52) CPC특허분류
G06K 9/00335 (2013.01)
G06K 9/481 (2013.01)
(21) 출원번호 10-2017-0146080
(22) 출원일자 2017년11월03일
심사청구일자 없음

(71) 출원인
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
김동칠
서울특별시 도봉구 우이천로20길 7, 103동 404호 (창동, 건영캐스빌아파트)
박성주
경기도 성남시 분당구 불정로 397, 315동 1003호 (서현동, 효자촌임광아파트)
(74) 대리인
특허법인지명

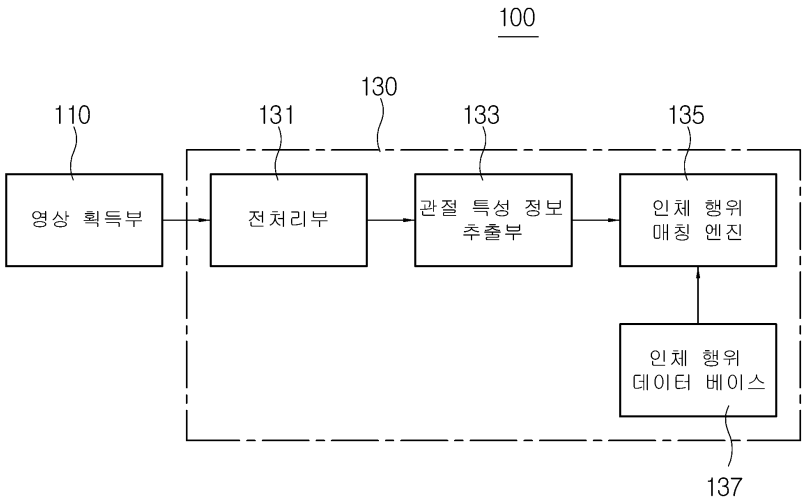
전체 청구항 수 : 총 12 항

(54) 발명의 명칭 **깊이 맵 정보 기반의 인체 행위 인지 방법 및 그 장치**

(57) 요 약

깊이 맵 정보 기반의 인체 행위 인지 방법이 개시된다. 이러한 인체 행위 인지 방법은, 영상 획득부로부터 입력된 인체 행위가 캡쳐된 깊이 맵 정보에 대해 전처리를 수행하여, 노이즈가 제거된 인체 영역을 추출하는 단계; 상기 인체 영역을 다수의 인체 부위로 분류하는 단계; 상기 다수의 인체 부위 각각의 관절 위치 좌표를 정의하는 단계; 상기 관절 위치 좌표의 변위량을 기반으로 관절 특성 정보를 추출하는 단계; 사정에 정의된 인체 행위 데이터베이스에서 상기 추출된 관절 특성 정보에 매칭되는 인체 행위 정보를 검색하고, 상기 깊이 맵 정보에 캡쳐된 인체 행위를 상기 검색된 인체 행위 정보에 정의된 인체 행위로 인지하는 단계를 포함한다.

대 표 도 - 도1



(52) CPC특허분류
G06K 9/627 (2013.01)
G06N 3/084 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 1711055147
부처명 과학기술정보통신부
연구관리전문기관 정보통신기술진흥센터
연구사업명 SW컴퓨팅산업원천기술개발
연구과제명 (딥뷰-2세부) 대규모 실시간 비디오 분석에 의한 전역적 다중 관심객체 추적 및 상황 예측 기술 개발
기 여 율 1/1
주관기관 광주과학기술원
연구기간 2014.04.01 ~ 2024.02.28

명 세 서

청구범위

청구항 1

영상 획득부로부터 입력된 인체 행위가 캡처된 깊이 맵 정보에 대해 전처리를 수행하여, 노이즈가 제거된 인체 영역을 추출하는 단계;

상기 인체 영역을 다수의 인체 부위로 분류하는 단계;

상기 다수의 인체 부위 각각의 관절 위치 좌표를 정의하는 단계;

상기 관절 위치 좌표의 변위량을 기반으로 관절 특성 정보를 추출하는 단계; 및

사전에 정의된 인체 행위 데이터베이스에서 상기 추출된 관절 특성 정보에 매칭되는 인체 행위 정보를 검색하고, 상기 깊이 맵 정보에 캡처된 인체 행위를 상기 검색된 인체 행위 정보에 정의된 인체 행위로 인지하는 단계

를 포함하는 인체 행위 인지 방법.

청구항 2

제1항에서, 상기 분류하는 단계는,

인체 영역과 인체 부위 간의 상호 관련성을 학습한 분류 모델을 이용하여 상기 인체 영역을 다수의 인체 부위로 분류하는 단계인 것인 인체 행위 인지 방법.

청구항 3

제2항에서, 상기 분류 모델은,

컨볼루션 신경망(CNN, Convolutional Neural Network) 구조의 학습기법에 따라 학습 데이터를 학습하는 것인 인체 행위 인지 방법.

청구항 4

제2항에서, 상기 분류 모델은,

인체의 외관 형상 별로 인체 부위가 분류된 학습 데이터를 학습하는 것인 인체 행위 인지 방법.

청구항 5

제1항에서, 상기 관절 위치 좌표를 정의하는 단계는,

Mean Shift 기법의 밀도 추정자(Density Estimator)를 이용하여, 상기 다수의 인체 부위 각각의 관절 위치 좌표를 정의하는 단계인 것인 인체 행위 인지 방법.

청구항 6

제1항에서, 상기 분류하는 단계는,

상기 인체 영역을 M개의 인체 부위로 분류하는 단계이고,

상기 관절 위치 좌표를 정의하는 단계는,

상기 M개로 분류된 인체 부위를 N(여기서, N은 M보다 작은 자연수)개로 다시 분류하는 단계; 및

상기 N개로 분류된 인체 부위 각각의 관절 위치 좌표를 정의하는 단계

를 포함하는 것인 인체 행위 인지 방법.

청구항 7

제1항에서, 상기 관절 특성 정보를 추출하는 단계는,

이전 프레임에서 이전 관절 위치 좌표와 현재 프레임에서 상기 이전 관절 위치 좌표에 대응하는 현재 관절 위치 좌표 간의 변위량을 벡터 형태로 나타내는 특징 벡터를 계산하는 단계; 및

상기 계산한 특징 벡터를 상기 관절 특성 정보로서 추출하는 단계

를 포함하는 것인 인체 행위 인지 방법.

청구항 8

영상 획득부로부터 입력된 인체 행위가 캡쳐된 깊이 맵 정보에 대해 전처리를 수행하여, 노이즈가 제거된 인체 영역을 추출하는 전처리부;

상기 인체 영역을 다수의 인체 부위로 분류하는 분류부;

상기 다수의 인체 부위 각각의 관절 위치 좌표를 정의하는 좌표 설정부;

상기 관절 위치 좌표의 변위량을 기반으로 관절 특성 정보를 추출하는 추출부; 및

사정에 정의된 인체 행위 데이터베이스에서 상기 추출된 관절 특성 정보에 매칭되는 인체 행위 정보를 검색하고, 상기 깊이 맵 정보에 캡쳐된 인체 행위를 상기 인체 행위 데이터베이스에서 검색된 인체 행위 정보에 정의된 인체 행위로 인지하는 인체 행위 매칭 엔진

을 포함하는 인체 행위 인지 장치.

청구항 9

제8항에서, 상기 분류부는,

인체 영역과 인체 부위 간의 상호 관련성을 학습한 분류 모델을 이용하여 상기 인체 영역을 다수의 인체 부위로 분류하는 것인 인체 행위 인지 장치.

청구항 10

제8항에서, 상기 분류 모델은,

인체 영역과 인체 부위 간의 상호 관련성을 학습하기 위해, 인체의 외관 형상 별로 인체 부위가 분류된 학습 데이터를 학습하는 것인 인체 행위 인지 장치.

청구항 11

제8항에서, 상기 좌표 설정부는,

Mean Shift 기법의 밀도 추정자(Density Estimator)를 이용하여, 상기 다수의 인체 부위 각각의 관절 위치 좌표를 정의하는 것인 인체 행위 인지 장치.

청구항 12

제8항에서, 상기 추출부는,

이전 프레임에서 이전 관절 위치 좌표와 현재 프레임에서 상기 이전 관절 위치 좌표에 대응하는 현재 관절 위치 좌표 간의 변위량을 벡터 형태로 나타내는 특징 벡터를 계산하고, 상기 계산한 특징 벡터를 상기 관절 특성 정보로서 추출하는 것인 인체 행위 인지 장치.

발명의 설명

기 술 분 야

[0001] 본 발명은 깊이 맵 정보 기반의 인체 행위 인지 방법 및 그 장치에 관한 것으로, 상세하게는, 3D 카메라를 통해 얻은 깊이 맵(Depth Map) 정보를 이용하여 인체 행위를 인지하는 방법 및 그 장치에 관한 것이다.

배 경 기 술

[0003] 최근, 영상 보안 시스템에 대한 연구개발이 활발히 진행되고 있다. 영상 보안 시스템은 영상 기반의 보안 서비스를 제공하는 시스템이다. 2차원 영상(2D) 기반의 영상 보안 시스템은 2D 영상을 이용하여 사람, 차량 등의 객체를 탐지, 분류 및 트래킹하는 영상 처리를 수행하고, 그 처리 결과로부터 객체의 행위 또는 이벤트를 인지한다. 즉, 2D 기반의 영상 보안 시스템은 2D 영상을 분석하여 객체가 특정 지점을 통과하는지, 침입했는지, 또는 배회하는 지 등을 모니터링 한다.

[0004] 그러나, 객체의 행위 또는 이벤트를 인지하기 위해, 2D 영상을 분석하는 과정에서, 분석결과의 정확도는 조명, 날씨 등과 같은 환경적인 요소에 영향을 받는다. 특히, 2D 영상은 충분한 밝기를 제공하는 않는 야간에서 객체의 행위 또는 이벤트를 인지할 수 있는 표시 품질을 제공하지 않기 때문에, 그 분석결과의 정확도가 낮다.

발명의 내용

해결하려는 과제

[0006] 본 발명에서 해결하고자 하는 과제는 전술한 바와 같이 2D 영상 분석에 대한 문제를 해결하기 위해, 3차원(3D) 카메라로부터 획득한 깊이 맵 정보를 기반으로 인체 부위를 분류하고, 분류된 인체 부위에서 관절 특성 정보를 추출하여 추출된 관절 특성 정보를 기반으로 인체 행위를 인지하는 깊이 맵 정보 기반의 인체 행위 인지 방법 및 그 장치를 제공하는 데 있다.

[0007] 본 발명에서 해결하고자 하는 과제는 이상에서 언급된 것들에 한정되지 않으며, 언급되지 아니한 다른 해결과제들은 아래의 기재로부터 당해 기술분야에 있어서의 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0009] 상술한 목적을 달성하기 위한 본 발명의 일면에 따른 인체 행위 인지 방법은, 영상 획득부로부터 입력된 인체 행위가 캡쳐된 깊이 맵 정보에 대해 전처리를 수행하여, 노이즈가 제거된 인체 영역을 추출하는 단계; 상기 인체 영역을 다수의 인체 부위로 분류하는 단계; 상기 다수의 인체 부위 각각의 관절 위치 좌표를 정의하는 단계; 및 상기 관절 위치 좌표의 변위량을 기반으로 관절 특성 정보를 추출하는 단계; 사정에 정의된 인체 행위 데이터베이스에서 상기 추출된 관절 특성 정보에 매칭되는 인체 행위 정보를 검색하고, 상기 깊이 맵 정보에 캡쳐된 인체 행위를 상기 검색된 인체 행위 정보에 정의된 인체 행위로 인지하는 단계를 포함한다.

[0010] 상술한 목적을 달성하기 위한 본 발명의 다른 일면에 따른 인체 행위 인지 장치는, 영상 획득부로부터 입력된 인체 행위가 캡쳐된 깊이 맵 정보에 대해 전처리를 수행하여, 노이즈가 제거된 인체 영역을 추출하는 전처리부;

상기 인체 영역을 다수의 인체 부위로 분류하는 분류부; 상기 다수의 인체 부위 각각의 관절 위치 좌표를 정의하는 좌표 설정부; 상기 관절 위치 좌표의 변위량을 기반으로 관절 특성 정보를 추출하는 추출부; 및 사정에 정의된 인체 행위 데이터베이스에서 상기 추출된 관절 특성 정보에 매칭되는 인체 행위 정보를 검색하고, 상기 깊이 맵 정보에 캡처된 인체 행위를 상기 인체 행위 데이터베이스에서 검색된 인체 행위 정보에 정의된 인체 행위로 인지하는 인체 행위 매칭 엔진을 포함한다.

발명의 효과

[0012] 본 발명에 따르면, 3D 카메라로부터 획득한 깊이 맵(Depth Map) 정보를 기반으로 인체 부위를 분류하고, 분류된 인체 부위에서 관절 특성 정보를 추출하여 인체 행위를 인지함으로써, 환경적인 요소에 관계없이 인체 행위를 정확하게 인지할 수 있다. 따라서, 악조건의 환경에서도 인체 행위의 정확한 인지가 가능하다.

[0013] 나아가, 본 발명의 인체 행위 인지 방법이 적용된 영상 보안 시스템은 악조건의 환경에서도 인체 행위를 정확하게 인지함으로써, 개인 신변 안전 및 범죄 예방 효과를 극대화할 수 있다.

[0014] 본 발명의 효과는 이상에서 언급된 것들에 한정되지 않으며, 언급되지 아니한 다른 효과들은 아래의 기재로부터 당해 기술분야에 있어서의 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

[0016] 도 1은 본 발명의 일 실시 예에 따른 인체 행위 인지 장치의 블록도이다.

도 2는 도 1에 도시한 관절 특성 정보 추출부의 블록도이다.

도 3은 도 2에 도시한 분류부가 인체 부위를 분류하기 위해 학습하는 학습 데이터의 일 예를 도식적으로 나타낸 도면이다.

도 4는 도 2에 도시한 좌표 설정부에 의해 설정된 관절 위치 좌표의 일 예를 도식적으로 나타낸 것이다.

도 5는 본 발명의 일 실시 예에 따른 깊이 맵 정보 기반의 인체 행위 인지 방법을 나타내는 흐름도이다.

도 6는 도 5에 도시한 단계 S520의 상세 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0017] 본 발명은 이하에서 개시되는 실시 예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 이하의 실시 예들은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 목적, 구성 및 효과를 용이하게 알려주기 위해 제공되는 것일 뿐으로서, 본 발명의 권리범위는 청구항의 기재에 의해 정의된다. 한편, 본 명세서에서 사용된 용어는 실시 예들을 설명하기 위한 것이며 본 발명을 제한하고자 하는 것은 아니다. 본 명세서에서, 단수형은 문구에서 특별히 언급하지 않는 한 복수형도 포함한다. 명세서에서 사용되는 "포함한다(comprises)" 및/또는 "포함하는(comprising)"은 언급된 구성소자, 단계, 동작 및/또는 소자가 하나 이상의 다른 구성소자, 단계, 동작 및/또는 소자의 존재 또는 추가됨을 배제하지 않는다.

[0018] 도 1은 본 발명의 일 실시 예에 따른 인체 행위 인지 장치의 블록도이다.

[0019] 도 1을 참조하면, 본 발명의 일 실시 예에 따른 객체 행위 인지 장치(100)는 객체에 대한 깊이 맵 정보를 기반으로 다양한 환경 조건에서도 객체 행위를 인지할 수 있다. 여기서, 객체는, 사람, 동물, 차량, 이동 가능한 물건 등일 수 있다.

[0020] 설명의 편의를 위해, 객체 행위 인지 장치(100)가 인체 행위를 인지하는 것으로 가정한다. 이에 따라, 이하에서는 객체 행위 인지 장치(100)는 '인체 행위 인지 장치'로 지칭하며, 객체 행동은 인체 행동으로 지칭한다.

[0021] 인체 행위를 인지할 수 있는 인체 행위 인지 장치(100)는 컴퓨터 프로세서를 구비한 전자 장치로 구현된 것일 수 있다. 전자 장치로 구현될 수 있는 인체 행위 인지 장치(100)는 다른 전자 장치 내에 내장될 수 있다. 다른 전자 장치는, 예를 들면, 모바일 폰, 데스크 탑, 서버, 영상 보안 기기, 이동 로봇 등일 수 있다. 이에 한정하지 않고, 다른 전자 장치는 인체 행위의 인지가 필요한 장치라면, 그 종류에 제한이 없다.

[0022] 인체 행위 인지 장치(100)는, 기본적으로, 영상 획득부(110) 및 영상 처리부(130)를 포함할 수 있다.

[0023] 영상 획득부(110)는 인체를 촬영한 3차원(3D) 이미지를 캡쳐하도록 구성된 이미지 센서일 수 있다. 이미지 센서는 스테레오 카메라, 깊이 카메라 등과 같이 3D 카메라로 통칭될 수 있는 모든 종류의 카메라를 포함할 수 있다.

[0024] 도시하지는 않았으나, 영상 획득부(110)는 3D 이미지로부터 깊이 맵 정보를 추출할 수 있는 수단을 포함하도록 구성될 수 있다. 깊이 맵 정보는 카메라와 객체 간의 거리를 픽셀 단위의 깊이값(깊이 정보 또는 깊이 데이터)으로 표현되는 정보로 정의할 수 있다. 깊이값은 '강도값'이라는 용어로 대체될 수 있다.

[0025] 영상 획득부(110)는 3D 이미지로부터 추출된 깊이 맵 정보(또는 깊이 이미지)를 영상 처리부(130)로 제공할 수 있다.

[0026] 영상 처리부(130)는 영상 획득부(110)로부터 제공된 깊이 맵 정보를 기반으로 인체 행위를 인식하기 위해 깊이 맵 정보를 프로세싱할 수 있다. 이러한 영상 처리부(130)는 적어도 하나의 범용의 프로세서 및/또는 그래픽 프로세서를 포함할 수 있다.

[0027] 영상 처리부(130)는 전처리부(131), 관절 특성 정보 추출부(133) 및 인체 행위 매칭 엔진(135)을 포함할 수 있다.

[0028] 전처리부(131)는 영상 획득부(130)로부터 제공된 깊이 맵 정보로부터 노이즈가 제거된 인체 정보를 추출하기 위해 전처리 과정을 수행할 수 있다. 여기서, 노이즈가 제거된 인체 정보는 인체 행위를 다수의 픽셀 좌표값으로 나타낸 정보일 수 있다.

[0029] 관절 특성 정보 추출부(133)는 전처리 과정에 의해 노이즈가 제거된 인체 정보로부터 관절 특성 정보를 추출한다.

[0030] 관절 특성 정보를 추출하기 위한, 관절 특성 정보 추출부(133)의 블록 구성도가 도 2에 도시된다.

[0031] 도 2를 참조하면, 관절 특성 정보 추출부(133)는 분류부(133-1), 좌표 설정부(133-3) 및 추출부(133-5)를 포함할 수 있다.

[0032] 분류부(133-1)는 전처리부(도 1의 131)에 의해 노이즈가 제거된 인체 정보를 M개의 인체 부위로 분류한다. 인체 부위 분류를 위해, 분류부(133-1)는 인체 부위를 분류하도록 학습된 일종의 분류 모델일 수 있다. 이러한 분류 모델은 노이즈가 제거된 인체 정보와 인체 부위 간의 상호 관련성을 학습 데이터를 이용하여 학습한 일종의 학습 모델일 수 있다. 학습 방법은, 예를 들면, 신경망 구조의 딥러닝(deep learning) 학습 기법 중에 하나인 컨볼루션 신경망(CNN, Convolutional Neural Network) 구조의 학습기법이 사용될 수 있다.

[0033] 도 3에는 분류부(133-1)가 인체 부위를 분류하기 위해 학습하는 학습 데이터의 예를 도식적으로 나타낸 것으로서, 분류부(133-1)는, 도 3에 도시된 바와 같이, 날씬한 정도 또는 뚱뚱한 정도로 구분되는 외관 형상 별로 인체 부위가 분류된 다수의 학습 데이터를 학습할 수 있다. 특별히 한정하는 것은 아니지만, 외관 형상 별로 인체 부위가 분류된 각 학습 데이터는 총 44개의 인체 부위로 분류된 학습 데이터일 수 있다.

[0034] 좌표 설정부(133-3)는 분류부(133-1)에서 분류된 인체 부위를 상기 M개보다 작은 N개의 인체부위로 다시 분류하고, Mean Shift 기법의 밀도 추정자(Density Estimator)를 이용하여, 다시 분류한 N개의 인체 부위를 N개의 관절 위치 좌표로 각각 설정한다(또는 정의한다). 즉, 좌표 설정부(133-3)는 Mean Shift 기법의 밀도 추정자(Density Estimator)를 이용하여 다시 분류한 인체 부위들 각각을 구성하는 픽셀 좌표들이 수렴하는 좌표를 관절 위치 좌표로 정의할 수 있다. 도 4에는 좌표 설정부(133-3)에서 정의한 관절 위치 좌표의 일 예를 도식적으로 나타낸 것이다.

[0035] 좌표 설정부(133-3)는 설정한 관절 위치 좌표를 추출부(133-5)로 제공할 수 있다.

[0036] 한편, 좌표 설정부(133-3)는, 예를 들면, 44개로 분류된 인체 부위를 10개의 인체 부위로 다시 분류한 경우, 10개의 인체 부위를 10개의 관절 위치 좌표로 설정할 수 있다.

[0037] 추출부(133-5)는 좌표 설정부(133-3)에서 제공하는 관절 위치 좌표의 변위량(차이값 또는 변위량)을 나타내는 특징 벡터(Feature vector)를 계산하고, 계산된 특징 벡터를 관절 특성 정보로서 추출한다. 즉, 추출부(133-5)는 이전 프레임에서 관절 위치 좌표(이하, 이전의 관절 위치 좌표)와 현재 프레임에서 상기 이전의 관절 위치 좌표에 대응하는 관절 위치 좌표(이하, 현재의 관절 위치 좌표) 간의 변위량(차이값 또는 이동량)을 관절 특성

정보로서 추출할 수 있다.

[0038] 추출부(133-5)는 추출한 관절 특성 정보를 인체 행위 매칭 엔진(135)으로 제공할 수 있다.

[0039] 다시 도 2를 참조하면, 인체 행위 매칭 엔진(135)은 사전에 정의된 인체 행위 데이터베이스(137)에서 상기 관절 특성 정보 추출부(133)로부터 제공된 관절 특성 정보에 매칭되는 인체 행위 정보를 검색하고, 관절 특성 정보에 매칭되는 인체 행위 정보가 검색되면, 검색된 인체 행위 정보에 정의된 인체 행위를 영상 획득부(110)에서 촬영한 인체의 인체 행위로 인지할 수 있다. 관절 특성 정보와 인체 행위 데이터베이스(137)에 저장된 인체 행위 정보 간의 매칭 여부를 판단하는 방법으로, SVM(Support Vector Machine) 기법이 이용될 수 있다.

[0040] 도 5는 본 발명의 일 실시 예에 따른 깊이 맵 정보 기반의 인체 행위 인지 방법을 나타내는 흐름도이다.

[0041] 도 5를 참조하면, 먼저, 단계 S510에서, 3D 카메라로부터 깊이 맵 정보(또는 깊이 이미지)가 입력되는 과정이 수행된다.

[0042] 이어, 단계 S520에서, 입력된 깊이 맵 정보(또는 깊이 이미지)로부터 노이즈가 제거된 인체 정보를 추출하기 위해, 입력된 깊이 맵 정보(또는 깊이 이미지)에 대해 전처리 과정이 수행된다. 이하, 도 6을 참조하여 전처리 과정에 대해 상세히 설명한다.

[0043] 도 6은 단계 S520의 상세하게 나타내는 흐름도이다.

[0044] 도 6을 참조하면, 먼저, 단계 S521에서, 깊이 맵 정보(또는 깊이 이미지)에서 인체를 구성하는 픽셀 좌표값을 포함하는 인체 영역을 검출하는 과정이 수행된다. 검출된 인체의 픽셀 좌표값은 메모리에 저장될 수 있다. 인체의 픽셀 좌표값을 통해 인체의 위치와 인체의 개수를 파악할 수 있다. 인체 영역을 검출하는 방법으로, 다양한 객체 검출 알고리즘이 이용될 수 있으며, 본 실시 예에서는, 객체 검출 속도가 빠른 딥러닝 기반의 객체 검출 기법이 사용될 수 있다. 이러한 딥러닝 기반의 객체 검출 기법은 본 발명의 요지를 벗어나므로 상세한 설명은 생략한다.

[0045] 이어, 단계 S523에서, 깊이 맵 정보(또는 깊이 이미지)에서 배경을 제거하기 위해, 깊이 맵 정보(또는 깊이 이미지)에 대해 연결 요소 분석(Connected Component Analysis, CCA)을 수행하여, 배경 영역을 검출한다. 구체적으로, 깊이 맵 정보(또는 깊이 이미지)에 포함된 픽셀들 중에서 유사한 픽셀값(밝기값, 강도값 또는 계조값)을 갖는 픽셀들을 그룹화하여, 배경 영역에 대응하는 픽셀 그룹을 검출한다.

[0046] 이어, 단계 S525에서, 전 단계에서 배경 영역이 검출되면, 깊이 맵 정보(또는 깊이 이미지)에서 인체 영역을 제외한 배경영역을 제거하는 과정이 수행된다.

[0047] 이어, 단계 S527에서, 배경 영역이 제거된 깊이 맵 정보(또는 깊이 이미지)에서 천장 영역을 제거하는 과정이 수행된다. 상기 단계 S523에서 수행한 연결 요소 분석(CCA)을 통해 인체 영역을 제외한 배경 영역을 제거하더라도 인체의 머리 위쪽 부분은 인체 영역으로 판단하여 제거되지 않을 확률이 높다. 이에 따라 평면 방정식을 통해 천장 영역을 계산하여 천장 영역을 제거한다.

[0048] 이어, 단계 S529에서, 천장 영역이 제거된 깊이 맵 정보(또는 깊이 이미지)에서 지면 영역을 제거하는 과정이 수행된다. 전술한 바와 유사하게, 상기 단계 S523에서 수행한 연결 요소 분석(CCA)을 통해 인체 영역을 제외한 배경 영역을 제거하더라도 인체의 신발 아래쪽 부분을 인체 영역으로 판단하여 제거되지 않을 확률이 높다. 이에 따라 평면 방정식을 통해 지면 영역을 계산하여 천장 영역을 제거한다.

[0049] 이와 같이, 천장 및 지면 영역의 제거 과정을 통해 깊이 맵 정보(또는 깊이 이미지)에서 노이즈가 제거된 인체 영역(또는 인체 정보)이 검출될 수 있다.

[0050] 한편, 단계 S527 및 S529에서, 평면 방정식을 이용하여 천장과 지면 영역을 제거하는 방법으로, RANSAC(Random sample consensus) 기법이 이용될 수 있다. 간략히 설명하면, 깊이 맵 정보(또는 깊이 이미지)에서 임의의 3개의 픽셀을 선택하고, 선택된 3개의 픽셀을 평면 방정식을 이용하여 평면을 결정하는 제1 과정이 수행된다. 이후, 지면으로부터 특정 임계치까지를 평면으로 정하고, 결정된 평면에서 임계치에 포함되는 픽셀의 개수를 구하는 제2 과정이 수행된다. 이후, 제1 및 제2 과정을 수회 반복하고, 그 반복 수행된 결과 중에서 가장 많은 픽셀을 포함하는 평면을 천정 및 바닥으로 각각 결정할 수 있다.

[0051] 다시 도 5를 참조하면, 도 6을 참조하여 설명한 전처리 과정을 통해 깊이 맵 정보(또는 깊이 이미지)로부터 천장, 지면 등과 같은 노이즈가 깨끗하게 제거된 인체 정보(또는 인체 영역)이 검출되면, 단계 S530에서, 노이즈가 제거된 인체 정보(또는 인체 영역)를 M개의 인체 부위로 분류하는 과정이 수행된다. 인체 부위를 분류하는

방법으로, 인체 정보(또는 인체 영역)과 인체 부위 간의 상호 관련성을 학습한 분류 모델이 이용될 수 있다. 이러한 분류 모델은, 예를 들면, 신경망 구조의 딥러닝(deep learning) 학습 기법 중에 하나인 컨볼루션 신경망(CNN, Convolutional Neural Network) 구조의 학습기법을 통해 학습될 수 있다.

[0052] 이어, 단계 S540에서, 전단계 S530에서 분류한 M개의 인체 부위를 N(여기서, N은 M보다 작은 자연수)개의 인체 부위로 다시 분류하고, Mean Shift 기법의 밀도 추정자(Density Estimator)를 이용하여, 다시 분류한 N개의 인체 부위를 N개의 관절 위치 좌표로 각각 설정(또는 정의)하는 과정이 수행된다.

[0053] 이어, 단계 S550에서, 이전 프레임과 현재 프레임에서의 관절 위치 좌표의 변위량(차이값 또는 변위량)을 나타내는 특징 벡터(Feature vector)를 계산하고, 계산된 특징 벡터를 관절 특성 정보로서 추출하는 과정이 수행된다.

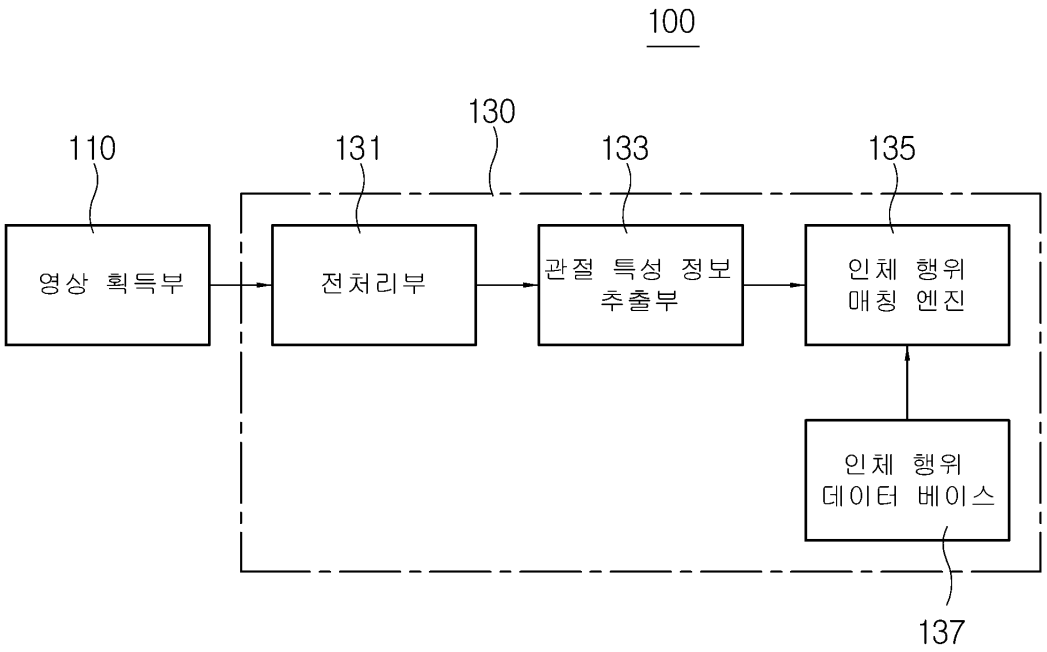
[0054] 이어, 단계 S560에서, 사전에 정의된 인체 행위 데이터베이스(137)에서 상기 관절 특성 정보 추출부(133)로부터 제공된 관절 특성 정보에 매칭되는 인체 행위 정보를 검색하고, 관절 특성 정보에 매칭되는 인체 행위 정보가 검색되면, 검색된 인체 행위 정보에 정의된 인체 행위를 영상 획득부(110)에서 촬영한 인체의 인체 행위로 인지하는 과정이 수행된다.

[0055] 이상 설명한 바와 같이, 본 발명의 장치 및 방법은 종래의 2D 영상 분석에 대한 문제점을 해결하기 위해, 3D 카메라에서 취득한 깊이 맵(Depth Map) 정보를 기반으로 인체의 부위를 분류하고, 분류된 인체 부위에서 관절 특성 정보를 추출하여 인체 행위를 인지함으로써, 다양한 환경적인 요소의 영향에 따른 인체 행위의 인지 성능 저하를 방지할 수 있고, 이러한 본 발명이 적용된 영상 보안 시스템은 다양한 환경적인 요소에 관계없이, 인체 행위에 대한 균일한 인지 성능을 유지함으로써, 개인 신변 안정 보장 및 범죄 예방을 극대화 할 수 있다.

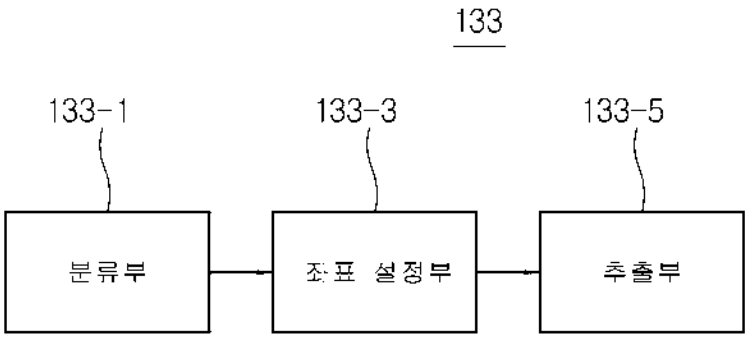
[0056] 이상에서 본 발명에 대하여 실시 예를 중심으로 설명하였으나 이는 단지 예시일 뿐 본 발명을 한정하는 것이 아니며, 본 발명이 속하는 분야의 통상의 지식을 가진 자라면 본 발명의 본질적인 특성을 벗어나지 않는 범위에서 이상에 예시되지 않은 여러 가지의 변형과 응용이 가능함을 알 수 있을 것이다. 예를 들어, 본 발명의 실시예에 구체적으로 나타난 각 구성 요소는 변형하여 실시할 수 있는 것이다. 그리고 이러한 변형과 응용에 관계된 차이점들은 첨부된 청구 범위에서 규정하는 본 발명의 범위에 포함되는 것으로 해석되어야 할 것이다.

도면

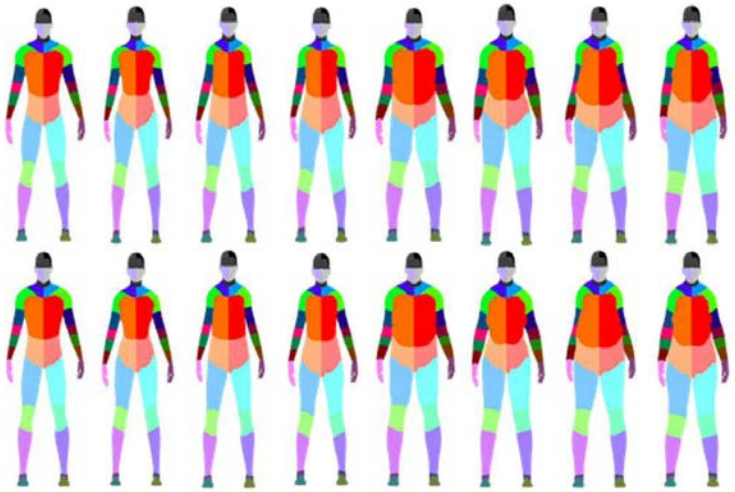
도면1



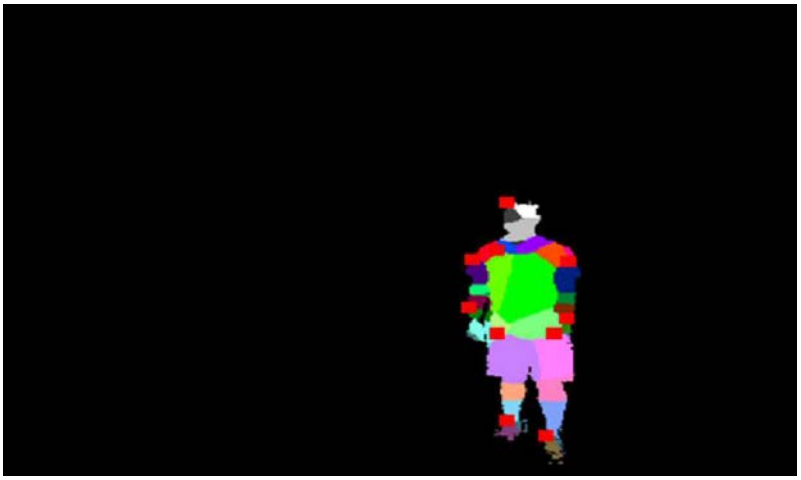
도면2



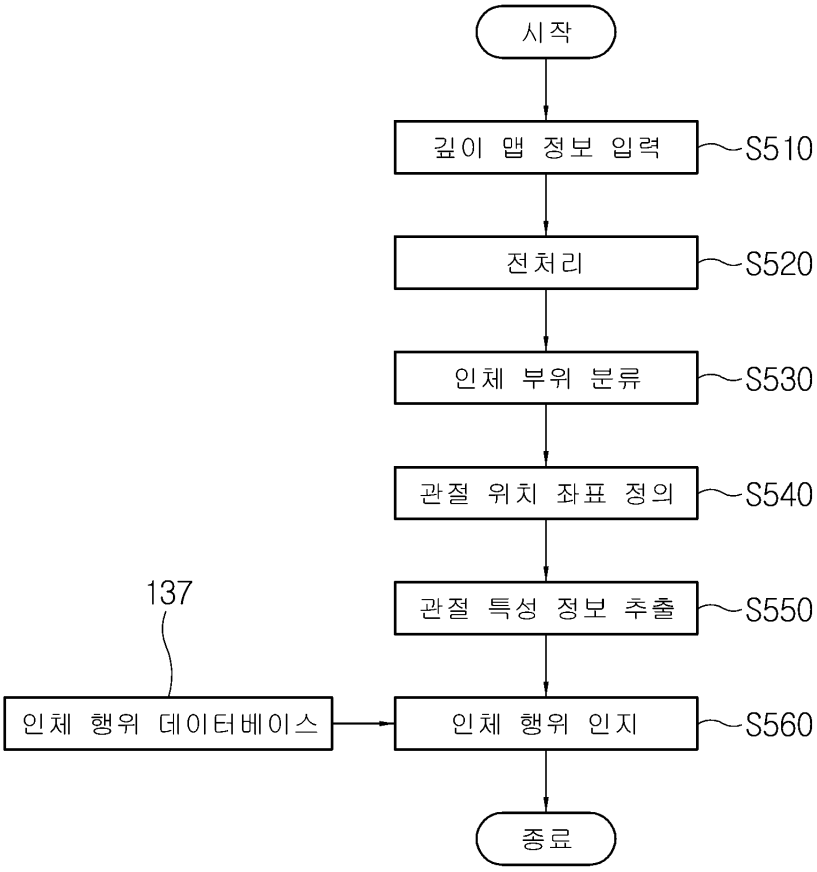
도면3



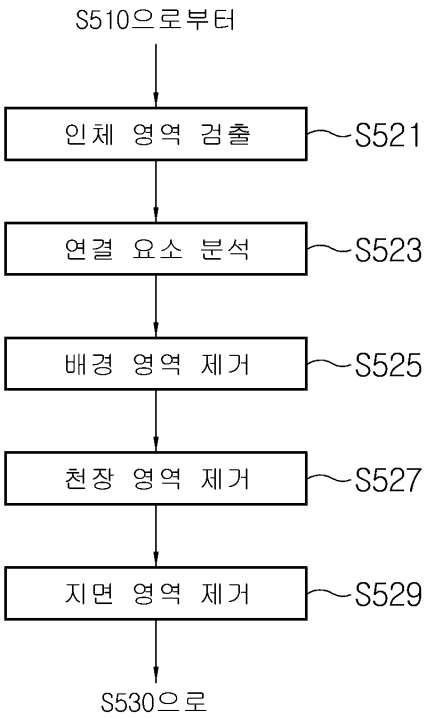
도면4



도면5


시작
깊이 맵 정보 입력
S510
전처리
S520
인체 부위 분류
S530
관절 위치 좌표 정의
S540
137
관절 특성 정보 추출
S550
인체 행위 데이터베이스
인체 행위 인지
S560
종료

도면6


S510으로부터
인체 영역 검출
S521
연결 요소 분석
S523
배경 영역 제거
S525
천장 영역 제거
S527
지면 영역 제거
S529
S530으로

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0055812
(43) 공개일자 2020년05월22일

(51) 국제특허분류(Int. Cl.)
G06K 9/00 (2006.01) *G08B 13/196* (2006.01)
(52) CPC특허분류
G06K 9/00201 (2013.01)
G06N 3/08 (2013.01)
(21) 출원번호 10-2018-0136243
(22) 출원일자 2018년11월08일
심사청구일자 없음

(71) 출원인
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
김동칠
서울특별시 도봉구 우이천로20길 7, 103동 404호 (창동, 건영캐스빌아파트)
박성주
경기도 성남시 분당구 불정로 397, 315동 1003호 (서현동, 효자촌임광아파트)
(74) 대리인
특허법인지명

전체 청구항 수 : 총 7 항

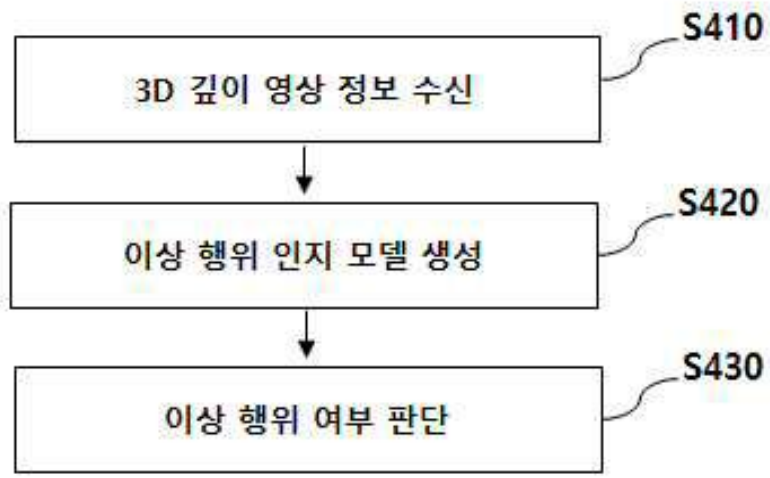
(54) 발명의 명칭 **딥러닝 기반 이상 행위 인지 장치 및 방법**

(57) 요 약

본 발명은 딥러닝을 기반으로 이상 행위를 인지하는 장치 및 그 방법에 관한 것이다.

본 발명에 따른 딥러닝 기반 이상 행위 인지 장치는 3D 영상을 수신하는 영상 수신부와, 3D 영상을 분석하여 이상 행위 여부를 판단하는 프로그램이 저장된 메모리 및 프로그램을 실행시키는 프로세서를 포함하고, 프로세서는 3D 영상에 대해 딥러닝을 이용한 시공간 영상데이터 분석을 수행하여 이상 행위 인지 모델을 생성하고, 이상 행위를 추정하는 것을 특징으로 한다.

대 표 도 - 도4



(52) CPC특허분류
G08B 13/19602 (2013.01)
G08B 13/19613 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 1711065359
부처명 과학기술정보통신부
연구관리전문기관 정보통신기술진흥센터
연구사업명 국가전략프로젝트(과기정통부)
연구과제명 (딥뷰-2세부) 대규모 실시간 비디오 분석에 의한 전역적 다중 관심객체 추적 및 상황 예측 기술 개발
기 여 율 1/1
주관기관 광주과학기술원
연구기간 2018.01.01 ~ 2018.12.31

명 세 서

청구범위

청구항 1

3D 영상을 수신하는 영상 수신부;

상기 3D 영상을 분석하여 이상 행위 여부를 판단하는 프로그램이 저장된 메모리; 및

상기 프로그램을 실행시키는 프로세서를 포함하되,

상기 프로세서는 상기 3D 영상에 대해 딥러닝을 이용한 시공간 영상데이터 분석을 수행하여 이상 행위 인지 모델을 생성하고, 이상 행위를 추정하는 것

인 딥러닝 기반 이상 행위 인지 장치.

청구항 2

제1항에 있어서,

상기 프로세서는 딥러닝 기반 영상 학습을 위한 상기 3D 영상을 수신하고, 깊이 값 차이를 계산하여 객체를 배경으로부터 분리하여 추출 및 추적하고, 인지 행위에 대한 Ground Truth 자료를 생성하고, Convolutional 3D 네트워크 구조를 기반으로 이상 행위 분석을 위한 네트워크 모델을 구현하고, 상기 이상 행위 인지 모델을 생성하는 것

인 딥러닝 기반 이상 행위 인지 장치.

청구항 3

제1항에 있어서,

상기 프로세서는 상기 이상 행위 인지 모델을 이용하여 실제 깊이 영상 데이터를 분석하고, 이상 행위가 추정되면 알림을 보안 시스템에 제공하는 것

인 딥러닝 기반 이상 행위 인지 장치.

청구항 4

(a) 3D 깊이 영상 정보를 수신하는 단계;

(b) 상기 3D 깊이 영상 정보를 이용하여 딥러닝 기반으로 시공간 분석을 수행하여 학습을 수행하고, 이상 행위 인지 모델을 생성하는 단계; 및

(c) 상기 이상 행위 인지 모델을 이용하여, 실제 깊이 영상 데이터를 분석하고 이상 행위 여부를 판단하는 단계

를 포함하는 딥러닝 기반 이상 행위 인지 방법.

청구항 5

제4항에 있어서,

상기 (b) 단계는 객체를 배경으로부터 분리하여 추출 및 추적하고, 행위 시작 및 종료 시간, 행위명에 대한 Ground Truth 자료를 생성하는 것

인 딥러닝 기반 이상 행위 인지 방법.

청구항 6

제5항에 있어서,

상기 (b) 단계는 CNN을 이용한 시공간 분석을 통해 네트워크 모델을 구현하고, 상기 이상 행위 인지 모델을 생성하는 것

인 딥러닝 기반 이상 행위 인지 방법.

청구항 7

제4항에 있어서,

상기 (c) 단계는 상기 이상 행위 인지 모델을 이용하여 실제 깊이 영상 데이터를 분석하고, 이상 행위가 추정되면 알림을 보안 시스템에 제공하는 것

인 딥러닝 기반 이상 행위 인지 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 이상 행위를 인지하는 장치 및 그 방법에 관한 것이다.

배 경 기 술

[0003] 종래 기술에 따른 2D 영상 기반의 영상 보안 시스템은 환경적인 요소에 따라 영향을 많이 받아 이상 행위 인지의 신뢰성이 크게 떨어지는 문제점이 있다.

[0004] 종래 기술에 따르면 3D 깊이(depth) 영상 기반의 이상 행위 인지 기술이 제안되고 있으나, 이는 인지의 정확도가 낮고 처리 속도가 느려 실시간으로 입력되는 영상을 처리하지 못하는 한계점이 있다.

발명의 내용

해결하려는 과제

[0006] 본 발명은 전술한 문제점을 해결하기 위하여 제안된 것으로, 3D 깊이 영상 입력을 기반으로 딥러닝을 활용한 시공간 영상 데이터 분석을 통해 이상 행위를 인지하는 장치 및 방법을 제공하는데 그 목적이 있다.

과제의 해결 수단

[0008] 본 발명에 따른 딥러닝 기반 이상 행위 인지 장치는 3D 영상을 수신하는 영상 수신부와, 3D 영상을 분석하여 이상 행위 여부를 판단하는 프로그램이 저장된 메모리 및 프로그램을 실행시키는 프로세서를 포함하고, 프로세서는 3D 영상에 대해 딥러닝을 이용한 시공간 영상데이터 분석을 수행하여 이상 행위 인지 모델을 생성하고, 이상 행위를 추정하는 것을 특징으로 한다.

[0009] 본 발명에 따른 딥러닝 기반 이상 행위 인지 방법은 3D 깊이 영상 정보를 수신하는 단계와, 3D 깊이 영상 정보를 이용하여 딥러닝 기반으로 시공간 분석을 수행하여 학습을 수행하고, 이상 행위 인지 모델을 생성하는 단계 및 이상 행위 인지 모델을 이용하여, 실제 깊이 영상 데이터를 분석하고 이상 행위 여부를 판단하는 단계를 포함하는 것을 특징으로 한다.

발명의 효과

[0011] 본 발명의 실시예에 따르면, 3D 깊이 영상을 이용하여 딥러닝 기반으로 이상 행위에 대한 실시간 인지가 가능하며, 이를 영상 보안 서비스에 적용하는 것이 가능하다.

[0012] 본 발명에 실시예에 따르면, 영상 보안 시스템에 적용되어 개인 신변 안전을 지키고 범죄 예방 효과를 극대화하는 효과가 있으며, 다양한 환경 변화에도 강건하며 실시간으로 영상 처리가 가능하여 이상 행위 추론의 신뢰성이 크게 향상되는 효과가 있다.

[0013] 본 발명의 효과는 이상에서 언급한 것들에 한정되지 않으며, 언급되지 아니한 다른 효과들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

[0015] 도 1은 본 발명의 실시예에 따른 딥러닝 기반 이상 행위 인지 장치를 나타내는 블록도이다.

도 2a 내지 도 2d는 본 발명의 실시예에 따른 깊이 영상 데이터에 대한 전처리 과정을 도시한다.

도 3은 본 발명의 실시예에 따른 딥러닝을 활용한 시공간 분석을 위한 네트워크 구조를 도시한다.

도 4는 본 발명의 실시예에 따른 딥러닝 기반 이상 행위 인지 방법을 나타내는 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0016] 본 발명의 전술한 목적 및 그 이외의 목적과 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다.

[0017] 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 이하의 실시예들은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 목적, 구성 및 효과를 용이하게 알려주기 위해 제공되는 것일 뿐으로서, 본 발명의 권리범위는 청구항의 기재에 의해 정의된다.

[0018] 한편, 본 명세서에서 사용된 용어는 실시예들을 설명하기 위한 것이며 본 발명을 제한하고자 하는 것은 아니다. 본 명세서에서, 단수형은 문구에서 특별히 언급하지 않는 한 복수형도 포함한다. 명세서에서 사용되는 "포함한다(comprises)" 및/또는 "포함하는(comprising)"은 언급된 구성소자, 단계, 동작 및/또는 소자가 하나 이상의 다른 구성소자, 단계, 동작 및/또는 소자의 존재 또는 추가됨을 배제하지 않는다.

[0020] 이하에서는, 당업자의 이해를 돕기 위하여 본 발명이 제안된 배경에 대하여 먼저 서술하고, 본 발명의 실시예에 대하여 서술하기로 한다.

[0021] 종래 기술에 따른 2D 영상 기반의 영상 보안 시스템은 사람, 차량 등의 객체를 탐지하고 분류하며, 객체를 트래킹함으로써, 특정 지점의 통과, 침입, 배회 등과 같은 단순 이상 행위를 인지한다.

[0022] 종래 기술에 따른 2D 영상 기반 영상 보안 시스템은 객체 및 이벤트를 검출하는데 조명, 날씨, 컬러 등과 같은 환경적인 요소에 의한 영향을 많이 받으며, 특히, 충분한 빛이 확보되지 않은 야간 상황에서는 객체에 대한 식별이 가능한 영상 획득 자체가 어려운 문제점이 있다.

[0023] 종래 기술에 따르면, 이러한 외부 환경 요소에 의한 문제점을 해결하기 위해, 3D Depth 기반의 이상 행위 인지 기술이 개발되고 있지만, 인지 정확도가 낮고 처리 속도가 늦어 실시간으로 입력되는 영상을 처리하기 어려운 문제점이 있다.

[0025] 본 발명은 전술한 문제점을 해결하기 위하여 제안된 것으로, 3D Depth 영상 입력 기반으로 CNN을 활용한 시공간 영상 데이터 분석을 통해 이상 행위를 분석하는 장치 및 방법을 제안하며, End-to-End 구조를 이용하여 영상 데이터를 처리하고 비교적 적은 수의 Layer 구성을 통해 실시간으로 이상 행위 인지가 가능한 장치 및 방법을 제

안한다.

[0026] 본 발명의 실시예에 따르면, 3D Depth 영상 기반으로 이상 행위를 인지하되, 다양한 환경적인 요소를 극복하고, 3D Depth 영상 및 딥러닝을 활용하여 실시간 영상 처리 및 이상 행위 인지가 가능하며, 영상 보안 시스템에 적용되어 개인 신변 안전을 보장하고 범죄를 예방하는 것이 가능하다.

[0028] 도 1은 본 발명의 실시예에 따른 딥러닝 기반 이상 행위 인지 장치를 나타내는 블록도이다.

[0029] 전술한 바와 같이, 종래 기술에 따르면, 다양한 환경 변화가 존재하는 상황에서 2D 영상을 이용하여 이상 행위를 인지하기 어렵기 때문에 효율적인 영상 보안 서비스를 수행할 수 없다.

[0030] 따라서, 본 발명의 실시예에 따르면 조명, 날씨, 컬러 등 다양한 환경 변화에 강인한 3D Depth 영상 정보에 대해 CNN(Convolutional Neural Network)을 활용한 시공간 영상 데이터 분석을 수행하여, 이상 행위를 분석한다.

[0031] 본 발명의 실시예에 따른 딥러닝 기반 이상 행위 인지 장치는 3D 영상을 수신하는 영상 수신부(100)와, 3D 영상을 분석하여 이상 행위 여부를 판단하는 프로그램이 저장된 메모리(200) 및 프로그램을 실행시키는 프로세서(300)를 포함하고, 프로세서(300)는 3D 영상에 대해 딥러닝을 이용한 시공간 영상데이터 분석을 수행하여 이상 행위 인지 모델을 생성하고, 이상 행위를 추정한다.

[0032] 프로세서(300)는 딥러닝 기반 영상 학습을 위한 3D 영상을 수신하고, 깊이 값 차이를 계산하여 객체를 배경으로부터 분리하여 추출 및 추적하고, 인지 행위에 대한 Ground Truth 자료를 생성하고, Convolutional 3D 네트워크 구조를 기반으로 이상 행위 분석을 위한 네트워크 모델을 구현하고, 이상 행위 인지 모델을 생성한다.

[0033] 프로세서(300)는 이상 행위 인지 모델을 이용하여 실제 깊이 영상 데이터를 분석하고, 이상 행위가 추정되면 알림을 보안 시스템에 제공한다.

[0034] 본 발명의 실시예에 따르면, End-to-End 구조를 이용하여 영상 데이터를 처리하고 비교적 적은 수의 Layer 구성을 통해 실시간으로 이상 행위 인지가 가능하다.

[0035] 본 발명의 실시예에 따른 프로세서(300)는 딥러닝 기반 영상 학습을 위한 Depth 영상 데이터 취득 및 전처리 단계를 수행하며, 이 때, 영상 데이터 학습을 위한 Depth 영상 데이터를 취득하고, 취득된 영상에서 배경 모델링 방법을 통해 객체를 추출한다.

[0036] 객체를 추출하는 과정은 객체 추출을 위한 Depth 영상 배경 모델링 과정, 객체와 배경 간의 Depth 값 차이 계산 과정, 객체에 해당하는 블랍(Blob)을 배경에서 분리 후 객체를 추출하는 과정, 이전 프레임과 비교하여 블랍을 중심으로 L2-norm이 최소가 되는 블랍을 동일객체로 트래킹하는 과정을 포함한다.

[0037] 도 2a 내지 2d는 본 발명의 실시예에 따른 깊이 영상 데이터에 대한 전처리 과정을 도시하는 것으로, 각각 배경 이미지 모델링, 객체/배경 깊이 값 차이 계산, 객체 추출 및 객체 트래킹 과정을 도시한다.

[0038] 본 발명의 실시예에 따른 프로세서(300)는 학습 Depth 영상 데이터 GT 생성을 수행하며, 이 때, 이상 행위 인지를 위해 행위 시작 시간, 행위 종료 시간, 행위명에 대한 GT를 생성한다.

[0039] 예컨대 인지 행위로는 배회, 펀치, 발차기, 쓰러짐, 군집, 집단 구타, 가상 경로 통과, 기물 파손, 나타남, 사라짐, 위장, 버림이 포함되고, 본 발명의 실시예에 따르면 배경에서 객체를 추출하여 학습을 수행함으로써, 배경 변화에 의한 영향을 덜 받게 되는 효과가 있다.

[0040] 본 발명의 실시예에 따른 프로세서(300)는 영상 학습 데이터에 대해 딥러닝 기반으로 시공간 분석을 수행하며, Convolutional 3D 네트워크 구조를 기반으로 이상 행위 분석을 위한 네트워크 모델을 구현하고, End-to-End 구조를 이용하여 영상 데이터를 처리하고 비교적 적은 수의 Layer 구성을 통해 실시간 이상 행위 인지가 가능하다. 도 3은 본 발명의 실시예에 따른 딥러닝을 활용한 시공간 분석을 위한 네트워크 구조를 도시한다.

[0041] 본 발명의 실시예에 따르면, 딥러닝 기반 이상 행위 인지 모델이 생성되고, 이 때 학습 영상 데이터를 딥러닝 기반 이상 행위 인지 모델에 50 epoch 정도 학습시킴으로써, 이상 행위 인지 모델을 생성한다.

[0042] 본 발명의 실시예에 따른 프로세서는 실제 Depth 영상 데이터를 입력 받아 이상 행위를 추정하고, 이벤트 알림을 수행하며, 이상 행위 인지 모델에 실제 Depth 영상 데이터를 입력 받아 이상 행위를 추정하고, 이상 행위 여부를 사용자에게 알려준다.

[0044] 도 4는 본 발명의 실시예에 따른 딥러닝 기반 이상 행위 인지 방법을 나타내는 순서도이다.

[0045] 본 발명의 실시예에 따른 딥러닝 기반 이상 행위 인지 방법은 3D 깊이 영상 정보를 수신하는 단계(S410)와, 3D 깊이 영상 정보를 이용하여 딥러닝 기반으로 시공간 분석을 수행하여 학습을 수행하고, 이상 행위 인지 모델을 생성하는 단계(S420) 및 이상 행위 인지 모델을 이용하여, 실제 깊이 영상 데이터를 분석하고 이상 행위 여부를 판단하는 단계(S430)를 포함한다.

[0046] S420 단계는 객체를 배경으로부터 분리하여 추출 및 추적하고, 행위 시작 및 종료 시간, 행위명에 대한 Ground Truth 자료를 생성한다.

[0047] S420 단계는 객체 추출을 위한 Depth 영상 배경 모델링 과정, 객체와 배경 간의 Depth 값 차이 계산 과정, 객체에 해당하는 블랍(Blob)을 배경에서 분리 후 객체를 추출하는 과정, 이전 프레임과 비교하여 블랍을 중심으로 L2-norm이 최소가 되는 블랍을 동일객체로 트래킹하는 과정을 포함한다.

[0048] S420 단계는 Ground Truth 자료 생성을 수행함에 있어서, 행위 시작 시간, 행위 종료 시간, 행위명에 대한 GT를 생성한다.

[0049] S420 단계는 영상 학습 데이터에 대해 딥러닝 기반으로 시공간 분석을 수행하며, Convolutional 3D 네트워크 구조를 기반으로 이상 행위 분석을 위한 네트워크 모델을 구현한다.

[0050] S430 단계는 실제 깊이 영상 데이터를 입력 받아 이상 행위를 추정하고, 이벤트 알림을 수행하며, 이상 행위 인지 모델에 실제 깊이 영상 데이터를 입력 받아 이상 행위를 추정하고, 이상 행위 여부를 사용자에게 알려준다.

[0052] 한편, 본 발명의 실시예에 따른 딥러닝 기반 이상 행위 인지 방법은 컴퓨터 시스템에서 구현되거나, 또는 기록매체에 기록될 수 있다. 컴퓨터 시스템은 적어도 하나 이상의 프로세서와, 메모리와, 사용자 입력 장치와, 데이터 통신 버스와, 사용자 출력 장치와, 저장소를 포함할 수 있다. 전술한 각각의 구성 요소는 데이터 통신 버스를 통해 데이터 통신을 한다.

[0053] 컴퓨터 시스템은 네트워크에 커플링된 네트워크 인터페이스를 더 포함할 수 있다. 프로세서는 중앙처리 장치(central processing unit (CPU))이거나, 혹은 메모리 및/또는 저장소에 저장된 명령어를 처리하는 반도체 장치일 수 있다.

[0054] 메모리 및 저장소는 다양한 형태의 휘발성 혹은 비휘발성 저장매체를 포함할 수 있다. 예컨대, 메모리는 ROM 및 RAM을 포함할 수 있다.

[0055] 따라서, 본 발명의 실시예에 따른 딥러닝 기반 이상 행위 인지 방법은 컴퓨터에서 실행 가능한 방법으로 구현될 수 있다. 본 발명의 실시예에 따른 딥러닝 기반 이상 행위 인지 방법이 컴퓨터 장치에서 수행될 때, 컴퓨터로 판독 가능한 명령어들이 본 발명에 따른 인지 방법을 수행할 수 있다.

[0056] 한편, 상술한 본 발명에 따른 딥러닝 기반 이상 행위 인지 방법은 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현되는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록 매체로는 컴퓨터 시스템에 의하여 해독될 수 있는 데이터가 저장된 모든 종류의 기록 매체를 포함한다. 예를 들어, ROM(Read Only Memory), RAM(Random Access Memory), 자기 테이프, 자기 디스크, 플래시 메모리, 광 데이터 저장장치 등이 있을 수 있다. 또한, 컴퓨터로 판독 가능한 기록매체는 컴퓨터 통신망으로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 읽을 수 있는 코드로서 저장되고 실행될 수 있다.

[0058] 이제까지 본 발명의 실시예들을 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 실시예들은 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.

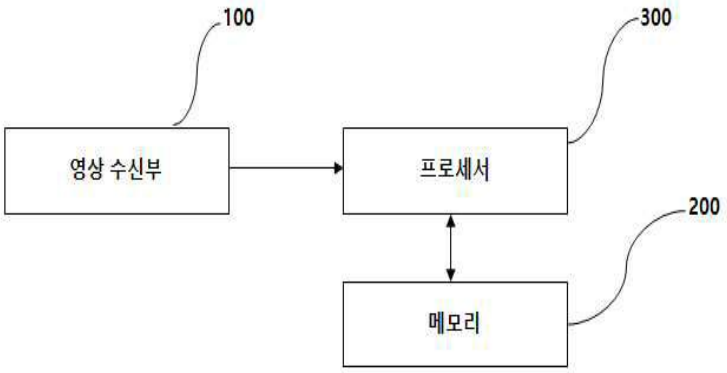
부호의 설명

[0060] 100: 영상 수신부 200: 메모리

300: 프로세서

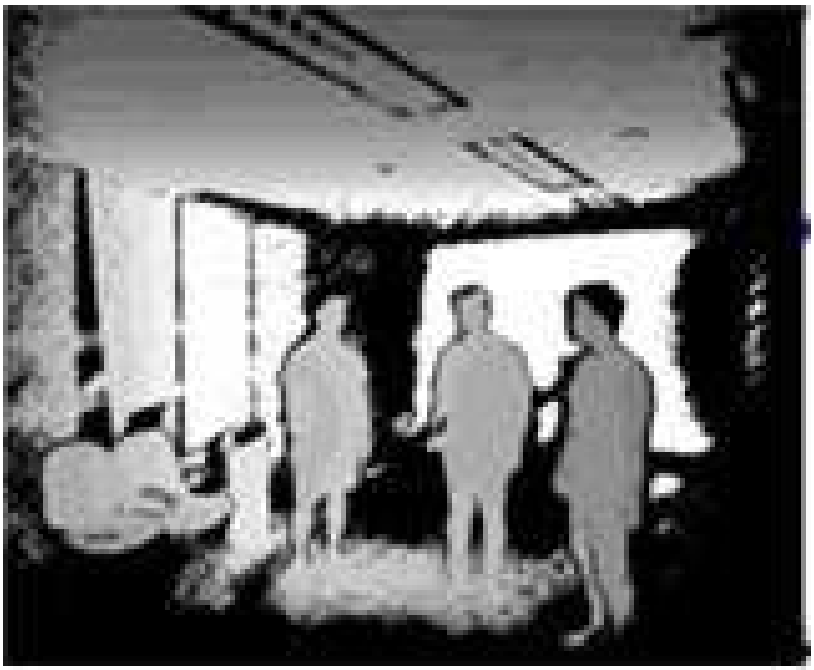
도면

도면1

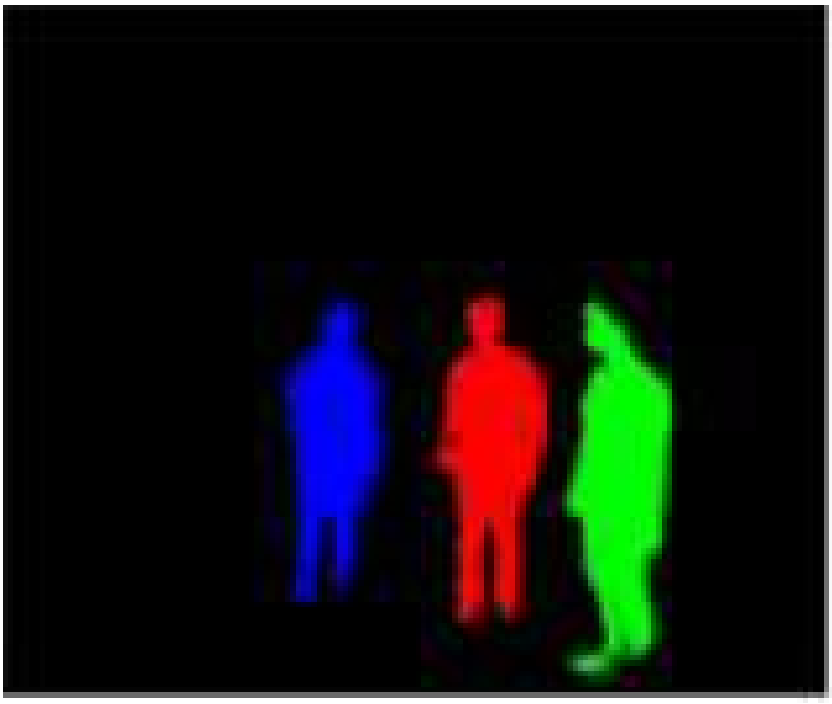


도면2a

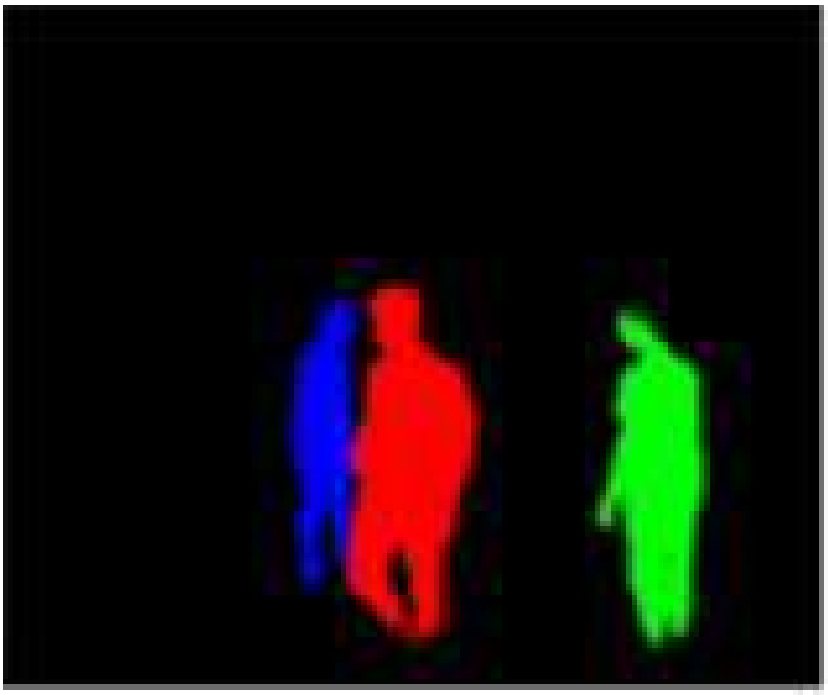
도면2b



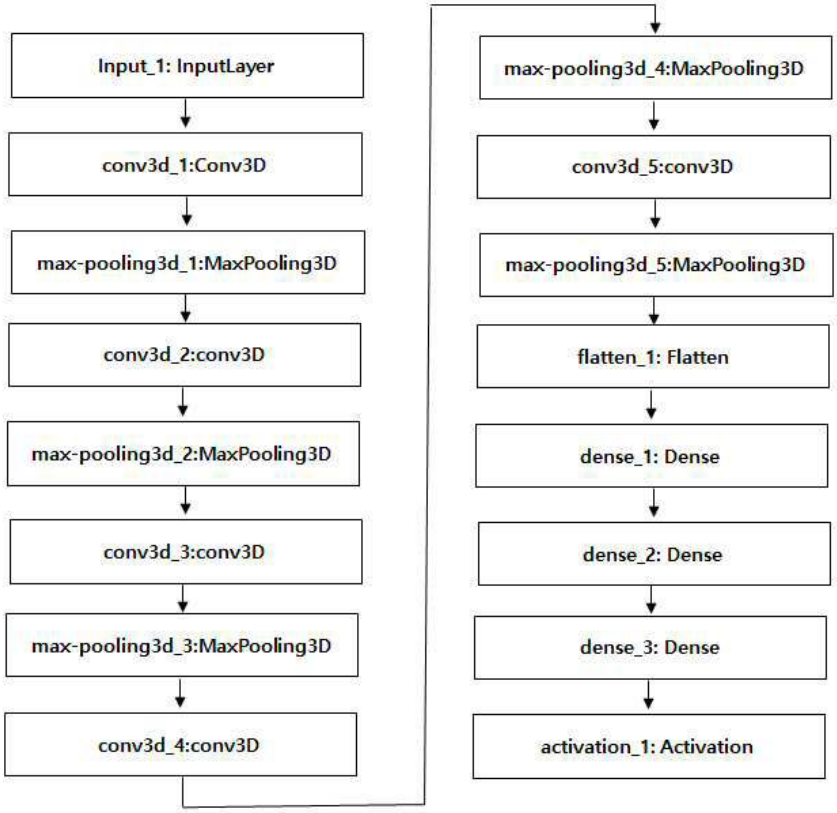
도면2c



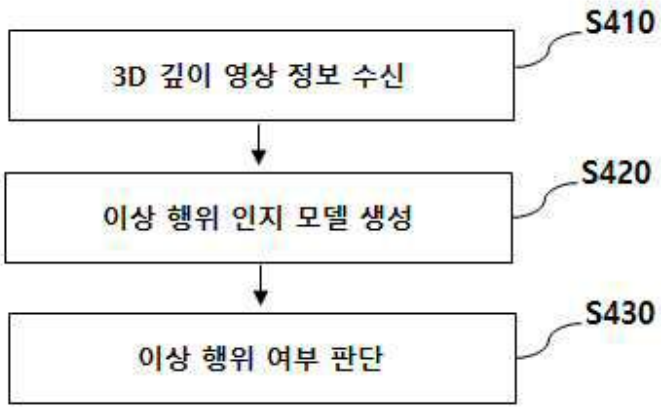
도면2d



도면3


Input_1: InputLayer
conv3d_1:Conv3D
max-pooling3d_1:MaxPooling3D
conv3d_2:conv3D
max-pooling3d_2:MaxPooling3D
conv3d_3:conv3D
max-pooling3d_3:MaxPooling3D
conv3d_4:conv3D
max-pooling3d_4:MaxPooling3D
conv3d_5:conv3D
max-pooling3d_5:MaxPooling3D
flatten_1: Flatten
dense_1: Dense
dense_2: Dense
dense_3: Dense
activation_1: Activation

도면4



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0059411
(43) 공개일자 2021년05월25일

(51) 국제특허분류(Int. Cl.)
G06K 9/00 (2006.01) *G06N 3/08* (2006.01)
H04N 7/18 (2006.01)
(52) CPC특허분류
G06K 9/00335 (2013.01)
G06N 3/08 (2013.01)
(21) 출원번호 10-2019-0146821
(22) 출원일자 2019년11월15일
심사청구일자 없음

(71) 출원인
한국전자기술연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
김동칠
서울특별시 도봉구 우이천로20길 7, 103동 404호 (창동, 건영캐스빌아파트)
박성주
경기도 성남시 분당구 불정로 397, 315동 1003호 (서현동, 효자촌임광아파트)
(74) 대리인
특허법인지명

전체 청구항 수 : 총 8 항

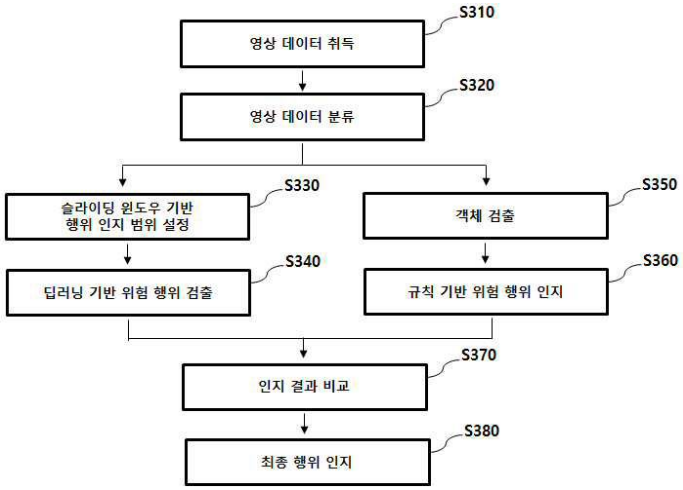
(54) 발명의 명칭 **하이브리드 기반 이상 행위 인지 시스템 및 그 방법**

(57) 요 약

본 발명은 하이브리드 기반 이상 행위 인지 시스템 및 그 방법에 관한 것이다.

본 발명에 따른 하이브리드 기반 이상 행위 인지 시스템은 영상 데이터를 수신하는 입력부와, 영상 데이터를 이용하여 이상 행위를 인지하는 프로그램이 저장된 메모리 및 프로그램을 실행시키는 프로세서를 포함하고, 프로세서는 딥러닝 및 규칙 기반의 인지 방식을 혼합하여 행위 인지 결과를 비교하여, 이상 행위를 인지하는 것을 특징으로 한다.

대 표 도 - 도3



(52) CPC특허분류
G08B 13/19602 (2013.01)
H04N 7/18 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 2014-3-00077
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 정보통신방송기술개발
연구과제명 대규모 실시간 비디오 분석에 의한 전역적 다중 관심객체 추적 및 상황 예측 기술 개발
기 여 율 1/1
과제수행기관명 광주과학기술원
연구기간 2014.01.01 ~ 2024.02.28

명 세 서

청구범위

청구항 1

영상 데이터를 수신하는 입력부;

상기 영상 데이터를 이용하여 이상 행위를 인지하는 프로그램이 저장된 메모리; 및

상기 프로그램을 실행시키는 프로세서를 포함하되,

상기 프로세서는 딥러닝 및 규칙 기반의 인지 방식을 혼합하여 행위 인지 결과를 비교하여, 상기 이상 행위를 인지하는 것

인 하이브리드 기반 이상 행위 인지 시스템.

청구항 2

제1항에 있어서,

상기 프로세서는 상기 영상 데이터를 분류하고, 슬라이딩 윈도우를 이용하여 딥러닝 기반으로 이상 행위를 인지할 범위를 설정하며, 3D 컨볼루션을 통해 프레임 내에 객체 존재 여부 및 움직임 존재 여부를 검출하는 것

인 하이브리드 기반 이상 행위 인지 시스템.

청구항 3

제1항에 있어서,

상기 프로세서는 탐지된 객체의 위치 정보를 이용하여 규칙 기반으로 이상 행위를 인지하는 것

인 하이브리드 기반 이상 행위 인지 시스템.

청구항 4

제1항에 있어서,

상기 프로세서는 딥러닝 기반의 인지 방식을 이용한 인지 결과와, 규칙 기반의 인지 방식을 이용한 인지 결과를 비교하여, 최종 행위를 인지하고 출력하는 것

인 하이브리드 기반 이상 행위 인지 시스템.

청구항 5

(a) 영상 데이터를 수신하는 단계;

(b) 상기 영상 데이터를 이용하여 딥러닝 기반으로 위험 행위를 인지하는 단계;

(c) 상기 영상 데이터를 이용하여 규칙 기반으로 위험 행위를 인지하는 단계; 및

(d) 상기 (b) 단계 및 (c) 단계에서의 인지 결과를 비교하여, 최종 행위를 출력하는 단계

를 포함하는 하이브리드 기반 이상 행위 인지 방법.

청구항 6

제5항에 있어서,

상기 (b) 단계는 슬라이딩 윈도우를 이용하여 이상 행위 인지 범위를 설정하고, 해당 프레임 내에 객체 및 움직임 존재 여부를 검출하는 것

인 하이브리드 기반 이상 행위 인지 방법.

청구항 7

제5항에 있어서,

상기 (c) 단계는 객체의 위치 정보를 이용하여 이상 행위를 인지하는 것

인 하이브리드 기반 이상 행위 인지 방법.

청구항 8

제5항에 있어서,

상기 (d) 단계는 상기 (b) 단계를 통해 획득된 행위별 인지 확률 및 상기 (c) 단계를 통해 획득된 행위 인지 결과를 비교하는 것

인 하이브리드 기반 이상 행위 인지 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 하이브리드 기반 이상 행위 인지 시스템 및 그 방법에 관한 것이다.

배 경 기 술

[0003] 딥러닝 기술은 영상 분석에 활용되어, 객체 검출, 객체 추적, 이상 상황 탐지 등 다양한 보안 영역에서 좋은 성능을 보인다.

[0004] 이에 따라, 영상 보안 영역에서 객체 검출 등의 영상 분석 기술은 대부분 딥러닝 기반으로 개발되어 운영되고 있다.

[0005] 그런데, 딥러닝만을 활용한 기술은 다양한 환경에서 다양한 형태의 이상 행위를 인지해야 하는 경우, 오히려 이상 행위 인지의 성능 저하가 발생하는 문제점이 있다.

발명의 내용

해결하려는 과제

[0007] 본 발명은 전술한 문제점을 해결하기 위하여 제안된 것으로, CCTV 카메라로부터 입력 받은 영상에서 이상 행위를 인지함에 있어서, 다양한 환경적인 요소를 극복하고, 기존 영상 보안 시스템에서 사용되는 영상 분석 기술보다 이상 행위 인지의 정확도를 향상시켜 개인 신변 안전을 보장하고 범죄를 예방하는 것이 가능한 이상 행위 인지 시스템 및 방법을 제공하는데 그 목적이 있다.

과제의 해결 수단

[0009] 본 발명에 따른 하이브리드 기반 이상 행위 인지 시스템은 영상 데이터를 수신하는 입력부와, 영상 데이터를 이용하여 이상 행위를 인지하는 프로그램이 저장된 메모리 및 프로그램을 실행시키는 프로세서를 포함하고, 프로세서는 딥러닝 및 규칙 기반의 인지 방식을 혼합하여 행위 인지 결과를 비교하여, 이상 행위를 인지하는 것을 특징으로 한다.

[0010] 본 발명에 따른 하이브리드 기반 이상 행위 인지 방법은 영상 데이터를 수신하는 단계와, 영상 데이터를 이용하여 딥러닝 기반으로 위험 행위를 인지하는 단계와, 영상 데이터를 이용하여 규칙 기반으로 위험 행위를 인지하는 단계 및 인지 결과를 비교하여 최종 행위를 출력하는 단계를 포함하는 것을 특징으로 한다.

발명의 효과

[0012] 본 발명에 따르면, 이상 행위를 인지함으로써 위험 상황을 예측하는 영상 보안 서비스 영역에 활용되어, 개인 신변 안전 및 범죄 예방 효과를 극대화하는 효과가 있다.

[0013] 본 발명에 따르면, 딥러닝 방법와 규칙 방법을 혼합함으로써, 다양한 환경 변화에서도 신뢰성 높은 이상 행위 추론이 가능한 효과가 있다.

[0014] 본 발명의 효과는 이상에서 언급한 것들에 한정되지 않으며, 언급되지 아니한 다른 효과들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

[0016] 도 1은 본 발명의 실시예에 따른 이상 행위 인지 시스템을 도시한다.

도 2는 본 발명의 실시예에 따른 이상 행위 검출을 도시한다.

도 3은 본 발명의 실시예에 따른 이상 행위 인지 방법을 나타내는 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0017] 본 발명의 전술한 목적 및 그 이외의 목적과 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다.

[0018] 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 이하의 실시예들은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 목적, 구성 및 효과를 용이하게 알려주기 위해 제공되는 것일 뿐으로서, 본 발명의 권리범위는 청구항의 기재에 의해 정의된다.

[0019] 한편, 본 명세서에서 사용된 용어는 실시예들을 설명하기 위한 것이며 본 발명을 제한하고자 하는 것은 아니다. 본 명세서에서, 단수형은 문구에서 특별히 언급하지 않는 한 복수형도 포함한다. 명세서에서 사용되는 "포함한다(comprises)" 및/또는 "포함하는(comprising)"은 언급된 구성소자, 단계, 동작 및/또는 소자가 하나 이상의 다른 구성소자, 단계, 동작 및/또는 소자의 존재 또는 추가됨을 배제하지 않는다.

[0021] 종래 기술에 따른 영상 보안 시스템은 CCTV 영상 분석 시, 딥러닝 기술만을 활용하여 객체 검출, 객체 추적, 위험 상황 인지 등과 같은 영상 분석을 하는 반면, 본 발명의 실시예에 따르면, 다양한 환경 및 사람의 상태 변화에 적응적으로 대응하기 위해 규칙 기반 기술을 딥러닝 방법과 혼합하여 이상 행위를 인지한다.

[0022] 도 1은 본 발명의 실시예에 따른 이상 행위 인지 시스템을 도시한다.

[0023] 본 발명의 실시예에 따른 이상 행위 인지 시스템(100)은 영상 데이터를 수신하는 입력부(110)와, 영상 데이터를 이용하여 이상 행위를 인지하는 프로그램이 저장된 메모리(120) 및 프로그램을 실행시키는 프로세서(130)를 포함하고, 프로세서(130)는 딥러닝 및 규칙 기반의 인지 방식을 혼합하여 행위 인지 결과를 비교하여, 이상 행위를 인지한다.

[0024] 프로세서(130)는 영상 데이터를 분류하고, 슬라이딩 윈도우를 이용하여 딥러닝 기반으로 이상 행위를 인지할 범

위를 설정하며, 3D 컨볼루션을 통해 프레임 내에 객체 존재 여부 및 움직임 존재 여부를 검출한다.

[0025] 프로세서(130)는 탐지된 객체의 위치 정보를 이용하여 규칙 기반으로 이상 행위를 인지한다.

[0026] 프로세서(130)는 딥러닝 기반의 인지 방식을 이용한 인지 결과와, 규칙 기반의 인지 방식을 이용한 인지 결과를 비교하여, 최종 행위를 인지하고 출력한다.

[0028] 도 3은 본 발명의 실시예에 따른 이상 행위 인지 방법을 나타내는 순서도이다.

[0029] 본 발명의 실시예에 따른 이상 행위 인지 방법은 (a) 영상 데이터를 수신하는 단계와, (b) 영상 데이터를 이용하여 딥러닝 기반으로 위험 행위를 인지하는 단계와, (c) 영상 데이터를 이용하여 규칙 기반으로 위험 행위를 인지하는 단계 및 (d) 인지 결과를 비교하여 최종 행위를 출력하는 단계를 포함한다.

[0030] (b) 단계는 슬라이딩 윈도우를 이용하여 이상 행위 인지 범위를 설정하고, 해당 프레임 내에 객체 및 움직임 존재 여부를 검출한다.

[0031] (c) 단계는 객체의 위치 정보를 이용하여 이상 행위를 인지한다.

[0032] (d) 단계는 (b) 단계를 통해 획득된 행위 별 인지 확률 및 (c) 단계를 통해 획득된 행위 인지 결과를 비교한다.

[0033] S310 단계는 CCTV로부터 영상 데이터를 취득한다.

[0034] S320 단계는 딥러닝 모델 및 규칙 기반 기술에 활용하기 위해, 영상 데이터를 분류한다.

[0035] S330 단계는 슬라이딩 윈도우 기술을 이용하여 딥러닝 기반으로 이상 행위를 인지할 범위를 설정한다.

[0036] 이 때, CCTV로 입력되는 영상의 기설정 범위(예: 3.2s)에 해당하는 시계열 데이터를 이용하여 행위 분석을 수행한다.

[0037] S340 단계는 딥러닝 기반으로 위험 행위 검출 모델을 통해 해당 영상 내에 이상 행위가 있는지 검출한다.

[0038] 이 때, 3D 컨볼루션을 통해 해당 프레임 내에 사람이 존재하고, 움직임이 있는지 여부를 검출한다.

[0039] S350 단계는 영상 데이터가 분류되면, 해당 영상 프레임 내에 객체가 존재하는지 검출한다.

[0040] 이 때, 객체 인지 딥러닝 모델을 이용하여 해당 영상 프레임 내 객체가 존재하는지 검출한다.

[0041] 객체가 검출되면, S360 단계는 규칙기반으로 이상 행위를 인지한다.

[0042] 이 때, S350 단계에서 탐지된 객체의 위치 정보 (x, y)를 이용하여 사람의 이상 행위를 인지한다.

[0043] 예컨대, 사람이 특정 범위를 서성거리면 '배회' 이벤트로 탐지한다.

[0044] S370 단계는 S340 단계에서의 딥러닝 기반 위험 행위 검출 결과와, S360 단계에서의 규칙 기반 위험 행위 인지 결과를 비교한다.

[0045] 본 발명의 실시예에 따르면, 딥러닝 방법과 규칙 기반 방법을 혼합하여 행위 인지 결과를 비교하며, 딥러닝 기반 위험 행위 검출 과정을 통해 획득된 행위별 인지 확률과, 규칙기반으로 나온 행위 인지 결과를 비교하여, 최종 행위를 인지한다.

[0046] S380 단계는 S370 단계의 인지 결과 비교에 따라, 이상 행위를 최종적으로 검출하며, 도 2에 도시한 바와 같이 최종적으로 결정된 행위명을 표출한다(Action Event, "펀치").

[0048] 한편, 본 발명의 실시예에 따른 이상 행위 인지 방법은 컴퓨터 시스템에서 구현되거나, 또는 기록매체에 기록될 수 있다. 컴퓨터 시스템은 적어도 하나 이상의 프로세서와, 메모리와, 사용자 입력 장치와, 데이터 통신 버스와, 사용자 출력 장치와, 저장소를 포함할 수 있다. 전술한 각각의 구성 요소는 데이터 통신 버스를 통해 데이터 통신을 한다.

[0049] 컴퓨터 시스템은 네트워크에 커플링된 네트워크 인터페이스를 더 포함할 수 있다. 프로세서는 중앙처리 장치(central processing unit (CPU))이거나, 혹은 메모리 및/또는 저장소에 저장된 명령어를 처리하는 반도체 장치일 수 있다.

[0050] 메모리 및 저장소는 다양한 형태의 휘발성 혹은 비휘발성 저장매체를 포함할 수 있다. 예컨대, 메모리는 ROM 및 RAM을 포함할 수 있다.

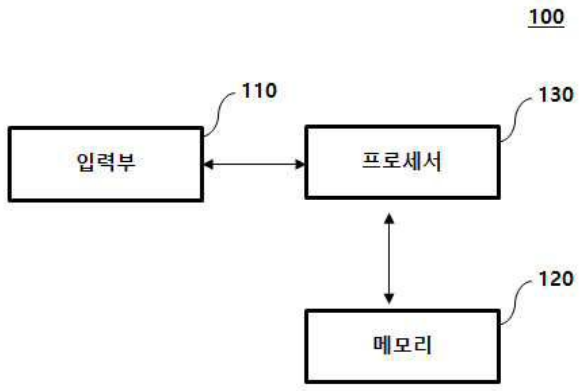
[0051] 따라서, 본 발명의 실시예에 따른 이상 행위 인지 방법은 컴퓨터에서 실행 가능한 방법으로 구현될 수 있다. 본 발명의 실시예에 따른 이상 행위 인지 방법이 컴퓨터 장치에서 수행될 때, 컴퓨터로 판독 가능한 명령어들이 본 발명에 따른 인지 방법을 수행할 수 있다.

[0052] 한편, 상술한 본 발명에 따른 이상 행위 인지 방법은 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현되는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록 매체로는 컴퓨터 시스템에 의하여 해독될 수 있는 데이터가 저장된 모든 종류의 기록 매체를 포함한다. 예를 들어, ROM(Read Only Memory), RAM(Random Access Memory), 자기 테이프, 자기 디스크, 플래시 메모리, 광 데이터 저장장치 등이 있을 수 있다. 또한, 컴퓨터로 판독 가능한 기록매체는 컴퓨터 통신망으로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 읽을 수 있는 코드로서 저장되고 실행될 수 있다.

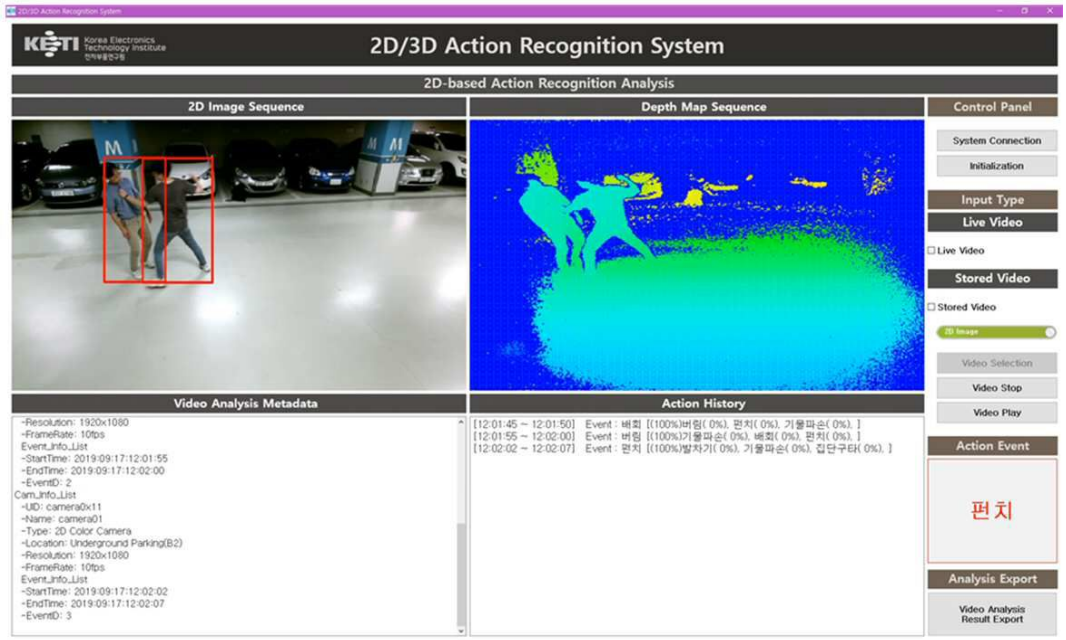
[0054] 이제까지 본 발명의 실시예들을 중심으로 살펴보았다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 그러므로 개시된 실시예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

도면

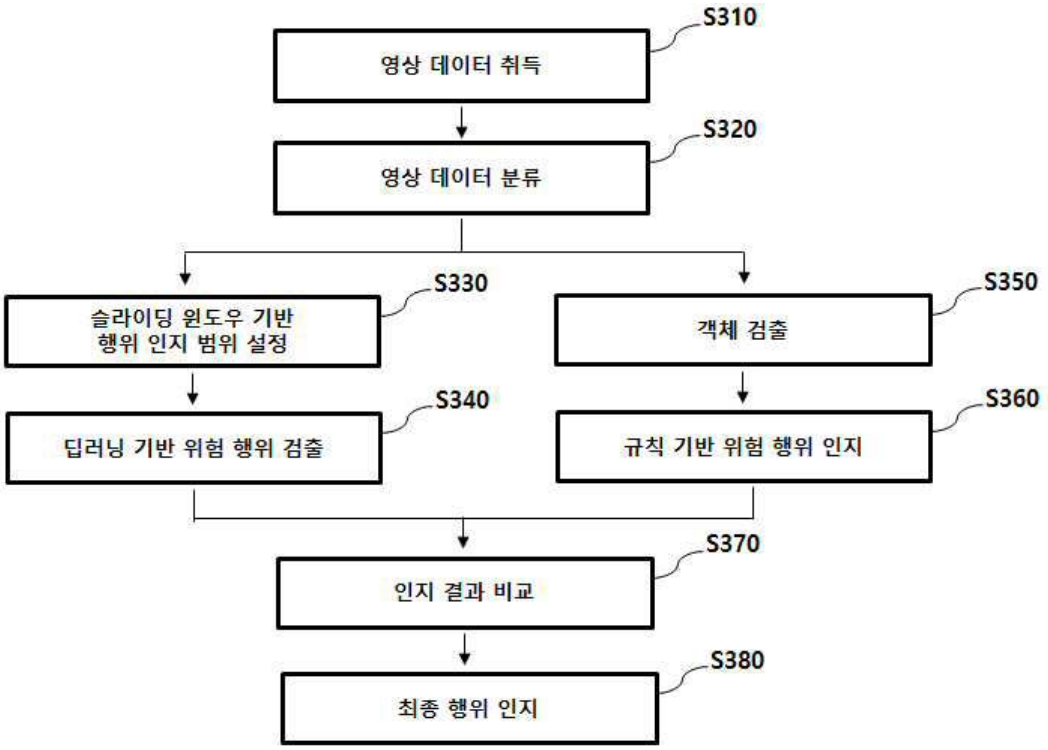
도면1



도면2


2D/3D Action Recognition System
2D-based Action Recognition Analysis
2D Image Sequence
Depth Map Sequence
Control Panel
System Connection
Initialization
Input Type
Live Video
Live Video
Stored Video
Stored Video
Video Selection
Video Stop
Video Play
Action Event
펀치
Analysis Export
Video Analysis Result Export
Video Analysis Metadata
Action History

도면3


S310
영상 데이터 취득
S320
영상 데이터 분류
S330
슬라이딩 윈도우 기반 행위 인지 범위 설정
S340
딥러닝 기반 위험 행위 검출
S350
객체 검출
S360
규칙 기반 위험 행위 인지
S370
인지 결과 비교
S380
최종 행위 인지

■ 기술명 : 제한된 네트워크 환경에서의 음향센서 기기 및 음향분석 시스템

(Acoustic Sensor Device and Acoustic Analysis System in Restricted Network Environments)

산업기술분류	전기전자/가정용 기기 및 전자응용 기기/음성정보기술 응용기기(200602)
Key-word(국문)	음향센서, 음향분석, 제한 네트워크
Key-word(영문)	acoustic sensor device, restricted network

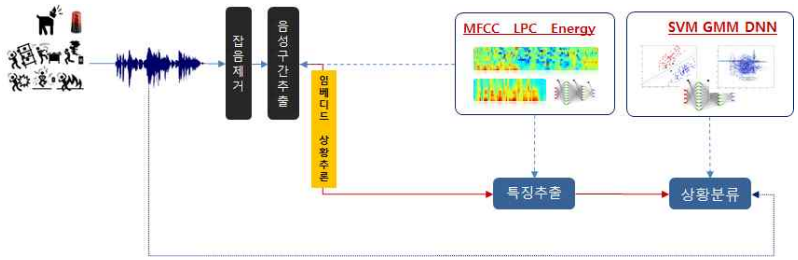
■ 기술의 개요

○ (배경) 최근 가정, 사무실, 작업장 등에 AI 스피커 및 각종 IoT 기기 등이 많이 보급되어 있으며, 이들 공간에서 발생하는 소리를 감지해 각종 상황을 인지하고 판단하는 기술에 대한 수요가 증가하고 있음

○ (개요) 본 기술은 실내에서 발생하는 소리를 분석해서 상황을 인지하고 알림을 발생시키는 음향센서 기기로, 음향입력장치(마이크로폰)와 입력신호 분석 프로세서로 구성됨

<기술 개요도 >

①입력신호 ▸ ②전처리(잡음 제거) ▸ ③특징 추출/분류 ▸ ④특별상황 판단



■ 기술의 구현수준(TRL)

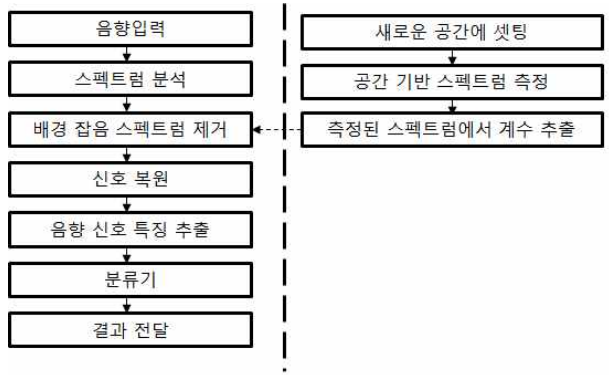
■ 기술의 장점(경쟁기술과의 차별성)

○ 경쟁 기술

- **(일정한 성능 구현 어려움)** 클라우드 또는 WIFI 등의 이용이 제한된 구역, 즉 제한된 네트워크 환경에서 음향기기 성능은 각 공간마다 다른 음향 상황으로 일정한 성능 구현이 어려움
- **(시스템 재설계 및 비용증가)** 새로운 환경에서 적응하기 위해서는 음향센서 기기 분류기를 재학습하거나 알고리즘 재설계가 필요하며, 이에 따른 추가 비용이 발생함

○ 본 기술의 장점

- **(분류기 재학습 불필요)** 특정장소에 음향센서 기기를 셋팅할 때, 배경 잡음을 분석하고, 이를 기반으로 음향 특징 추출 프로세스를 구현하여 분류기를 학습시키기 때문에, 분류기의 적응 또는 재학습 필요성이 없음
- **(음향센서 성능저하 방지)** 음향센서가 독립적으로 구동하기 때문에 특정장소에서 성능이 급격히 저하되는 것을 방지할 수 있음
- **(간편한 프로세스)** 특정 장소의 잡음 정보에 대해서 실시간 분석은 불가능하지만, 복잡한 프로세싱과 네트워크 사용이 불필요 함



음향 분석 프로세스: 음향 센서 기기 셋팅시 작업

■ 활용범위 및 응용분야

<table>
<tr><td> 
</td><td> 
</td></tr>
<tr><td>[안전/보안용 IoT 기기]</td><td>[독립 음향센서 IoT 기기]</td></tr>
<tr><td>  </td><td>  </td></tr>
<tr><td>[인공지능 스피커]</td><td>[자동화 설비 건전성 예측진단]</td></tr>
</table>

○ 가정/학교/어린이집 등의 IoT 센서 기반의 안전/보안 시스템에 적용
○ 독립적인 형태의 음향센서로서 홈 IoT 기기에 적용
○ 인공지능 스피커 등 포터블 IoT 기기 등에 적용
○ 공장/창고 등 자동화 설비 건전성 예측 진단 등에 적용

■ 지식재산권 현황

구분	발명의 명칭	출원번호 (출원일)	등록번호 (등록일)
특허	제한된 네트워크 환경에서의 음향 센서 기기 및 음향 분석 시스템	2018-0146810 (2018.11.23)	-

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0061258
(43) 공개일자 2020년06월02일

(51) 국제특허분류(Int. Cl.)
G10L 25/18 (2013.01) *G10L 19/02* (2006.01)
G10L 21/0272 (2013.01)
(52) CPC특허분류
G10L 25/18 (2013.01)
G10L 19/02 (2013.01)
(21) 출원번호 10-2018-0146810
(22) 출원일자 2018년11월23일
심사청구일자 없음

(71) 출원인
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
장달원
경기도 김포시 김포한강11로 179,504동 1002호
이종설
경기도 파주시 미래로 422, 112동 1201호 한빛마을1단지
(74) 대리인
박종한

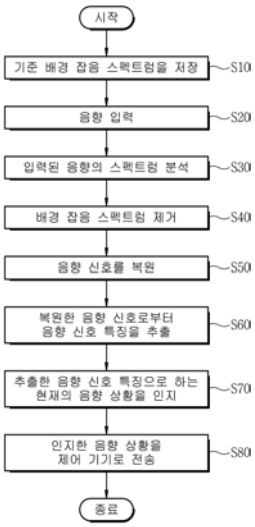
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 **제한된 네트워크 환경에서의 음향 센서 기기 및 음향 분석 시스템**

(57) 요 약

본 발명은 제한된 네트워크 환경에서의 음향 센서 기기 및 음향 분석 시스템에 관한 것으로, 실내의 제한된 네트워크 환경에서 입력된 음향을 분석하여 지금 현재의 음향 상황에 대한 인지를 수행하기 위한 것이다. 본 발명에 따른 음향 센서 기기는 기 저장된 기준 배경 잡음 스펙트럼을 기반으로 실내 환경에서 입력된 음향으로부터 배경 잡음 스펙트럼을 제거한 후 필요한 음향 신호 특징을 추출하여 현재의 음향 상황에 대한 인지를 수행하기 때문에, 기존의 네트워크 환경에서 적응 알고리즘을 사용하지 않더라도, 기준 배경 잡음 스펙트럼을 이용하여 다양한 실내 환경에서 적응적으로 지금 현재의 음향 상황에 대한 인지를 수행할 수 있다.

대 표 도 - 도3



(52) CPC특허분류
G10L 21/0208 (2013.01)
G10L 21/0272 (2013.01)
G10L 25/15 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 20000111
부처명 산업부
연구관리전문기관 한국산업기술평가관리원
연구사업명 글로벌전문기술개발사업
연구과제명 (R)영상/음향 상황인지 학습을 위한 상황수집센서와 스마트미러 및 센서 일체형 스마트비서 개발
기 여 율 1/1
주관기관 주식회사 엘컴텍
연구기간 2018.04.01 ~ 2020.12.31

명 세 서

청구범위

청구항 1

복수의 실내 환경에 대한 기준 배경 잡음 스펙트럼를 저장하는 저장부;

특정 실내 환경에서 음향을 입력받는 음향 입력부;

상기 입력된 음향의 스펙트럼을 분석하는 스펙트럼 분석부;

상기 기준 배경 잡음 스펙트럼을 기반으로 분석한 음향의 스펙트럼에 포함된 배경 잡음 스펙트럼을 제거하는 배경 잡음 제거부;

상기 배경 잡음 스펙트럼에 제거된 음향으로부터 음향 신호 특징을 추출하는 음향 신호 추출부; 및

추출한 음향 신호 특징을 수신하여 상기 특정 실내 환경에서의 음향 상황을 인지하는 분류기;

를 포함하는 제한된 네트워크 환경에서의 음향 센서 기기.

청구항 2

제1항에 있어서,

상기 기준 배경 잡음 스펙트럼은 복수의 실내 환경에서 수집된 배경 잡음 스펙트럼의 평균적인 배경 잡음 스펙트럼인 것을 특징으로 하는 제한된 네트워크 환경에서의 음향 센서 기기.

청구항 3

제1항에 있어서,

상기 스펙트럼 분석부는 입력된 음향을 N 포인트로 FFT 수행하여 N 포인트 스펙트럼을 생성하고, 생성한 N 포인트 스펙트럼을 주파수 밴드별로 에너지를 산출하여 상기 입력된 음향의 스펙트럼을 생성하는 것을 특징으로 하는 제한된 네트워크 환경에서의 음향 센서 기기.

청구항 4

제3항에 있어서,

상기 배경 잡음 제거부는 상기 기준 배경 잡음 스펙트럼을 주파수 밴드별로 에너지를 산출하여 상기 입력된 음향의 스펙트럼의 주파수 밴드별 에너지에서 차감하여 상기 입력된 음향의 배경 잡음 스펙트럼을 제거하는 것을 특징으로 하는 제한된 네트워크 환경에서의 음향 센서 기기.

청구항 5

제1항에 있어서,

제어 기기와 통신을 수행하고, 상기 분류기의 제어에 따라 인지한 음향 상황을 상기 제어 기기로 전송하는 통신부;

를 더 포함하는 것을 특징으로 하는 제한된 네트워크 환경에서의 음향 센서 기기.

청구항 6

특정 실내 환경에 설치되어 음향을 입력받아 상기 특정 실내 환경에서의 음향 상황을 인지하여 제어 기기로 전송하는 음향 센서 기기; 및

상기 음향 센서 기기로부터 음향 상황을 수신하고, 수신한 음향 상황에 따른 기능을 수행하는 제어 기기;를 포함하고,

상기 음향 센서 기기는,

복수의 실내 환경에 대한 기준 배경 잡음 스펙트럼를 저장하는 저장부;

상기 특정 실내 환경에서 음향을 입력받는 음향 입력부;

상기 입력된 음향의 스펙트럼을 분석하는 스펙트럼 분석부;

상기 기준 배경 잡음 스펙트럼을 기반으로 분석한 음향의 스펙트럼에 포함된 배경 잡음 스펙트럼을 제거하는 배경 잡음 제거부;

상기 배경 잡음 스펙트럼에 제거된 음향으로부터 음향 신호 특징을 추출하는 음향 신호 추출부; 및

추출한 음향 신호 특징을 수신하여 상기 특정 실내 환경에서의 음향 상황을 인지하는 분류기;

를 포함하는 제한된 네트워크 환경에서의 음향 분석 시스템.

청구항 7

제6항에 있어서,

상기 기준 배경 잡음 스펙트럼은 복수의 실내 환경에서 수집된 배경 잡음 스펙트럼의 평균적인 배경 잡음 스펙트럼인 것을 특징으로 하는 제한된 네트워크 환경에서의 음향 분석 시스템.

청구항 8

제6항에 있어서,

상기 스펙트럼 분석부는 입력된 음향을 N 포인트로 FFT 수행하여 N 포인트 스펙트럼을 생성하고, 생성한 N 포인트 스펙트럼을 주파수 밴드별로 에너지를 산출하여 상기 입력된 음향의 스펙트럼을 생성하고,

상기 배경 잡음 제거부는 상기 기준 배경 잡음 스펙트럼을 주파수 밴드별로 에너지를 산출하여 상기 입력된 음향의 스펙트럼의 주파수 밴드별 에너지에서 차감하여 상기 입력된 음향의 배경 잡음 스펙트럼을 제거하는 것을 특징으로 하는 제한된 네트워크 환경에서의 음향 분석 시스템.

발명의 설명

기 술 분 야

[0001] 본 발명은 음향을 인식해서 활용하는 기술에 관한 것으로, 더욱 상세하게는 실내의 제한된 네트워크 환경에서 입력된 음향을 분석하여 지금 현재의 음향 상황에 대한 인지를 수행하는 제한된 네트워크 환경에서의 음향 센서 기기 및 음향 분석 시스템에 관한 것이다.

배 경 기 술

[0002] 최근 정보통신기술(ICT)을 기반으로 다양한 사물을 인터넷으로 연결하고 사람과 사물, 사물과 사물 간 정보를 교환하며 상호 소통하는 이른바 사물 인터넷(IoT, Internet of Things)에 대한 수요가 다양한 분야에 걸쳐 높아지고 있다.

[0003] 생활 밀접도와, 서비스 활용도 측면 등을 고려한다면 위 사물 인터넷 관련 기술 중 음향 센서 기기에서 소리를 감지한 후, 특정한 조건에 부합할 때 사용자에게 알려주는 음향 분석 기술을 그 대표적인 예로 들 수 있겠다.

[0004] 음향 센서 기기는 가정이나 사무실 등 실내 공간에 존재하기 위해서 개발되는 것이며, 제어 기기의 제어를 받고, 자체적으로 음향 입력을 분석해서 제어 기기로 분석 결과를 전달하는 역할을 수행한다.

[0005] 음향 센서 기기에 있어서 음향 환경의 변화는 성능을 크게 떨어뜨리는 요인이 될 수 있다. 이와 같이 음향 신호를 입력 받아서 분석하는 음향 분석 시스템은 환경의 변화에 따라서 성능이 저하되는 같은 문제를 가지고 있다.

[0006] 이러한 문제를 극복하기 위한 방법은 기존의 시스템이 새로운 음향 환경에서의 음향 신호의 입력에 따라서 점점 적응해 가는 방법이 있다. 이를 위해서, 분류기(classifier)의 적응 방법이 많이 연구되고 있다. 즉 실내의 음향 환경은 각 방의 특성에 따라 다르기 때문에, 적절한 성능을 내기 위해서는 음향 환경에 적응하는 과정을 거쳐야 하고, 이러한 음향 환경에 따른 적응은 온라인으로 연결된 시스템에서 가능한 것이다.

[0007] 하지만 기존의 음향 센서 기기는 온라인으로 적응 알고리즘을 돌릴 정도의 네트워크 대역을 갖고 있지 못하고

대용량의 데이터를 주고받을 수 없기 때문에, 온라인으로 분류기를 적응하는 방법이 아닌 다른 방법으로 음향 환경에 적응이 필요한 실정이다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 한국등록특허 제10-1634356호 (2016.06.22. 등록)

발명의 내용

해결하려는 과제

[0009] 따라서 본 발명의 목적은 실내의 제한된 네트워크 환경에서 입력된 음향 신호를 분석하여 지금 현재의 음향 상황에 대한 인지를 수행하는 제한된 네트워크 환경에서의 음향 센서 기기 및 음향 분석 시스템을 제공하는 데 있다.

과제의 해결 수단

[0010] 상기 목적을 달성하기 위하여, 본 발명은 복수의 실내 환경에 대한 기준 배경 잡음 스펙트럼를 저장하는 저장부; 특정 실내 환경에서 음향을 입력받는 음향 입력부; 상기 입력된 음향의 스펙트럼을 분석하는 스펙트럼 분석부; 상기 기준 배경 잡음 스펙트럼을 기반으로 분석한 음향의 스펙트럼에 포함된 배경 잡음 스펙트럼을 제거하는 배경 잡음 제거부; 상기 배경 잡음 스펙트럼에 제거된 음향으로부터 음향 신호 특징을 추출하는 음향 신호 추출부; 및 추출한 음향 신호 특징을 수신하여 상기 특정 실내 환경에서의 음향 상황을 인지하는 분류기;를 포함하는 제한된 네트워크 환경에서의 음향 센서 기기를 제공한다.

[0011] 상기 기준 배경 잡음 스펙트럼은 복수의 실내 환경에서 수집된 배경 잡음 스펙트럼의 평균적인 배경 잡음 스펙트럼이다.

[0012] 상기 스펙트럼 분석부는 입력된 음향을 N 포인트로 FFT 수행하여 N 포인트 스펙트럼을 생성하고, 생성한 N 포인트 스펙트럼을 주파수 밴드별로 에너지를 산출하여 상기 입력된 음향의 스펙트럼을 생성한다.

[0013] 상기 배경 잡음 제거부는 상기 기준 배경 잡음 스펙트럼을 주파수 밴드별로 에너지를 산출하여 상기 입력된 음향의 스펙트럼의 주파수 밴드별 에너지에서 차감하여 상기 입력된 음향의 배경 잡음 스펙트럼을 제거한다.

[0014] 본 발명에 따른 음향 센서 기기는 제어 기기와 통신을 수행하고, 상기 분류기의 제어에 따라 인지한 음향 상황을 상기 제어 기기로 전송하는 통신부;를 더 포함한다.

[0015] 그리고 본 발명은 특정 실내 환경에 설치되어 음향을 입력받아 상기 특정 실내 환경에서의 음향 상황을 인지하여 제어 기기로 전송하는 음향 센서 기기; 및 상기 음향 센서 기기로부터 음향 상황을 수신하고, 수신한 음향 상황에 따른 기능을 수행하는 상기 제어 기기;를 포함하는 제한된 네트워크 환경에서의 음향 분석 시스템을 제공한다.

발명의 효과

[0016] 본 발명에 따른 음향 센서 기기는 기 저장된 기준 배경 잡음 스펙트럼을 기반으로 실내 환경에서 입력된 음향으로부터 배경 잡음 스펙트럼을 제거한 후 필요한 음향 신호 특징을 추출하여 현재의 음향 상황에 대한 인지를 수행하기 때문에, 실내의 제한된 네트워크 환경에서 입력된 음향 신호를 분석하여 지금 현재의 음향 상황에 대한 인지를 수행할 수 있다.

[0017] 이와 같이 본 발명에 따른 음향 센서 기기는, 네트워크 환경에서 적응 알고리즘을 사용하지 않더라도, 기준 배경 잡음 스펙트럼을 이용하여 다양한 실내 환경에서 적응적으로 지금 현재의 음향 상황에 대한 인지를 수행할 수 있다.

도면의 간단한 설명

[0018] 도 1은 본 발명의 실시예에 따른 제한된 네트워크 환경에서의 음향 분석 시스템을 보여주는 블록도이다.

도 2는 도 1의 음향 센서 기기의 세부 구성을 보여주는 블록도이다.

도 3은 본 발명의 실시예에 따른 제한된 네트워크 환경에서의 음향 분석 방법에 따른 흐름도이다.

도 4는 도 3의 기준 배경 잡음 스펙트럼을 저장하는 단계에 대한 상세 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0019] 하기의 설명에서는 본 발명의 실시예를 이해하는데 필요한 부분만이 설명되며, 그 이외 부분의 설명은 본 발명의 요지를 흩트리지 않는 범위에서 생략될 것이라는 것을 유의하여야 한다.

[0020] 이하에서 설명되는 본 명세서 및 청구범위에 사용된 용어나 단어는 통상적이거나 사전적인 의미로 한정해서 해석되어서는 아니 되며, 발명자는 그 자신의 발명을 가장 최선의 방법으로 설명하기 위해 용어의 개념으로 적절하게 정의할 수 있다는 원칙에 입각하여 본 발명의 기술적 사상에 부합하는 의미와 개념으로 해석되어야만 한다. 따라서 본 명세서에 기재된 실시예와 도면에 도시된 구성은 본 발명의 바람직한 실시예에 불과할 뿐이고, 본 발명의 기술적 사상을 모두 대변하는 것은 아니므로, 본 출원시점에 있어서 이들을 대체할 수 있는 다양한 균등물과 변형예들이 있을 수 있음을 이해하여야 한다.

[0021] 이하, 첨부된 도면을 참조하여 본 발명의 실시예를 보다 상세하게 설명하고자 한다.

[0022] 도 1은 본 발명의 실시예에 따른 제한된 네트워크 환경에서의 음향 분석 시스템을 보여주는 블록도이다.

[0023] 도 1을 참조하면, 본 실시예에 따른 음향 분석 시스템(100)은 실내의 제한된 네트워크 환경에서 실내에서 발생되는 음향을 입력받아 적응적으로 음향을 분석하여 지금 현재의 음향 상황에 대한 인지를 수행하는 시스템이다. 이러한 본 실시예에 다른 음향 분석 시스템(100)은 음향 센서 기기(10)와 제어 기기(20)를 포함한다. 음향 센서 기기(10)는 특정 실내 환경에 설치되어 음향을 입력받아 특정 실내 환경에서의 음향 상황을 인지하여 제어 기기(20)로 전송한다. 그리고 제어 기지는 음향 센서 기기(10)로부터 음향 상황을 수신하고, 수신한 음향 상황에 따른 기능을 수행한다.

[0024] 여기서 음향 센서 기기(10)는 다양한 실내 환경에 설치되어 해당 실내 환경에서 발생되는 음향을 수집 및 분석하여 음향 상황을 인지한다. 음향 센서 기기(10)는 인지한 음향 상황을 제어 기기(20)로 전송한다.

[0025] 본 실시예에 따른 음향 센서 기기(10)는 복수의 실내 환경에 대한 기준 배경 잡음 스펙트럼을 저장하고, 기준 배경 잡음 스펙트럼을 기반으로 입력되는 음향에 대한 분석 및 인지를 수행한다.

[0026] 즉 음향 센서 기기(10)는 기 저장된 기준 배경 잡음 스펙트럼을 기반으로 실내 환경에서 입력된 음향으로부터 배경 잡음 스펙트럼을 제거한 후 필요한 음향 신호 특징을 추출하여 현재의 음향 상황에 대한 인지를 수행하기 때문에, 실내의 제한된 네트워크 환경에서 입력된 음향 신호를 분석하여 지금 현재의 음향 상황에 대한 인지를 수행할 수 있다.

[0027] 이와 같이 음향 센서 기기(10)는, 네트워크 환경에서 적응 알고리즘을 사용하지 않더라도, 기준 배경 잡음 스펙트럼을 이용하여 다양한 실내 환경에서 적응적으로 지금 현재의 음향 상황에 대한 인지를 수행할 수 있다.

[0028] 음향 센서 기기(10)는 제어 기기(20)와 기준 배경 잡음 스펙트럼 또는 음향 상황을 인지한 신호를 송수신하는 정도의 통신을 수행하며, 대용량의 데이터를 송수신하는 네트워크 환경은 아니다.

[0029] 그리고 제어 기기(20)는 음향 센서 기기(10)로부터 수신한 음향 상황에 따른 기능을 수행한다. 예컨대 수신한 음향 상황이 "TV 끄기"인 경우, 제어 기기(20)는 연결된 TV를 오프시킨다. 수신한 음향 상황이 "특정 음악 송출"인 경우, 제어 기기(20)는 해당하는 특정 음악을 찾아서 TV, 오디오, 자체 스피커, 무선 스피커와 같은 음향 출력 기기를 통해서 출력할 수 있다.

[0030] 이와 같은 본 실시예에 따른 음향 센서 기기(10)에 대해서 도 1 및 도 2를 참조하여 설명하면 다음과 같다. 여기서 도 2는 도 1의 음향 센서 기기(10)의 세부 구성을 보여주는 블록도이다.

[0031] 본 실시예에 따른 음향 센서 기기(10)는 저장부(11), 음향 입력부(12), 스펙트럼 분석부(13), 배경 잡음 제거부(14), 음향 신호 추출부(14) 및 분류기(16)를 포함한다. 저장부(11)는 복수의 실내 환경에 대한 기준 배경 잡음 스펙트럼를 저장한다. 음향 입력부(12)는 특정 실내 환경에서 음향을 입력받는다. 스펙트럼 분석부(13)는 입력된 음향의 스펙트럼을 분석한다. 배경 잡음 제거부(14)는 기준 배경 잡음 스펙트럼을 기반으로 분석한 음향의 스펙트럼에 포함된 배경 잡음 스펙트럼을 제거한다. 음향 신호 추출부(14)는 배경 잡음 스펙트럼에 제거된 음향으로부터 음향 신호 특징을 추출한다. 그리고 분류기(16)는 추출한 음향 신호 특징을 수신하여 특정 실내 환경

에서의 음향 상황을 인지한다. 그 외 본 실시예에 따른 음향 센서 기기(10)는 통신부(17)를 더 포함할 수 있다.

[0032] 통신부(17)는 제어 기기(20)와 통신을 수행하고, 분류기(16)의 제어에 따라 인지한 음향 상황을 제어 기기(20)로 전송한다. 통신부(17)는 제어 기기(20)와 유무선 통신 방식으로 통신을 수행한다. 무선 통신 방식으로는 블루투스, 지그비, 와이파이 등과 같은 근거리 통신 방식이 사용될 수 있다.

[0033] 저장부(11)는 음향 센서 기기(10)의 동작 제어시 필요한 프로그램과, 그 프로그램의 수행 중에 발생되는 정보를 저장한다. 저장부(11)는 복수의 실내 환경에 대한 기준 배경 잡음 스펙트럼를 저장한다. 저장부(11)는 기준 배경 잡음 스펙트럼을 디폴트 형태로 저장할 수 있고, 제어 기기(20)로부터 수신하여 저장할 수 있다.

[0034] 여기서 기준 배경 잡음 스펙트럼은 복수의 실내 환경에서 수집된 배경 잡음 스펙트럼의 평균적인 배경 잡음 스펙트럼일 수 있다. 기준 배경 잡은 스펙트럼은 제어 기기(20)가 설정하거나 별도의 기기에서 설정할 수 있다. 예컨대 제어 기기(20)는 복수의 실내 환경에 대한 배경 잡음을 압력받는다. 제어 기기(20)는 입력된 배경 잡음의 스펙트럼을 분석한다. 제어 기기(20)는 분석한 배경 잡음 스펙트럼의 평균치를 산출한다. 제어 기기(20)는 산출한 평균치를 기준 배경 잡음 스펙트럼으로 설정한다. 그리고 제어 기기(20)는 설정한 기준 배경 잡음 스펙트럼을 저장하고, 저장한 기준 배경 잡음 스펙트럼을 음향 센서 기기(10)로 제공한다.

[0035] 음향 입력부(12)는 실내에서 발생되는 음향을 입력받는다. 예컨대 음향 입력부(12)는 마이크로폰일 수 있다.

[0036] 스펙트럼 분석부(13)는 입력된 음향을 N 포인트로 FFT 수행하여 N 포인트 스펙트럼을 생성한다. 스펙트럼 분석부(13)는 생성한 N 포인트 스펙트럼을 주파수 밴드별로 에너지를 산출하여 입력된 음향의 스펙트럼을 생성할 수 있다.

[0037] 배경 잡음 제거부(14)는 기준 배경 잡음 스펙트럼을 기반으로 생성한 음향의 스펙트럼으로부터 배경 잡음을 제거한다. 즉 배경 잡음 제거부(14)는 기준 배경 잡음 스펙트럼을 주파수 밴드별로 에너지를 산출하여 입력된 음향의 스펙트럼의 주파수 밴드별 에너지에서 차감하여 입력된 음향의 배경 잡음 스펙트럼을 제거한다.

[0038] 음향 신호 추출부(14)는 배경 잡음 스펙트럼이 제거된 음향 스펙트럼을 음향 신호로 복원한다. 음향 신호 추출부(14)는 복원한 음향 신호로부터 음향 신호 특징을 추출한다.

[0039] 그리고 분류기(16)는 음향 신호 추출부(14)로부터 추출한 음향 신호 특징을 수신하고, 수신한 음향 신호 특징으로부터 특정 실내 환경에서의 음향 상황을 인지한다. 분류기(16)는 인지한 음향 상황을 통신부(17)를 통해서 제어 기기(20)로 전송한다.

[0040] 한편 기존에 음향 센서 기기를 개발하는 데 있어서 문제점은 해당 음향 센서 기기가 설치되는 특정 실내 환경에 맞게 구체적으로 정해진 음향 특징과 분류기를 개발하게 된다. 그런데 실제 음향 센서 기기는 다양한 실내 환경에 설치되어 사용되기 때문에, 실내 환경에 따라서 음향 상황 인지도가 크게 달라질 수 있다. 즉 실내 환경도 장소에 따라서 배경 잡음의 크기나 형태가 전혀 다를 수 있다.

[0041] 이런 문제점을 해소하기 위해서, 본 실시예에 따른 음향 센서 기기(10)는 애초에 음향 특징을 추출할 때 배경 잡음에 대한 성분을 제거하고 음향 신호 특징을 추출한다. 그리고 추출한 음향 신호 특징을 기반으로 분류기(16)를 학습할 수 있도록, 음향 센서 기기(10)에 기준 배경 잡음 스펙트럼을 제공한다.

[0042] 기존의 연구들이 분류기를 재학습하거나, 분류기에 입력을 추가적으로 계속 주어서 새롭게 갱신하는 방법을 활용해서 환경의 변화에 대해서 적응하는 방법을 많이 연구해왔다. 하지만 실제 음향 센서 기기에서 이런 방법을 사용하는 경우, 음향 센서 기기가 분류기를 학습시킬 수 있는 외부프로세서와 네트워크 연결이 되어서 외부에서 분류기를 학습하고 이를 다시 다운로드 받는 과정을 거쳐야 한다. 이런 문제점은 하드웨어적 비용을 증가시킨다.

[0043] 따라서 본 실시예에서는 음향 센서 기기(10)를 셋팅할 때, 배경 잡음을 분석하고 평균적인 배경 잡음을 기준 배경 잡음 스펙트럼으로 설정하고, 기준 배경 잡음 스펙트럼을 기반으로 배경 잡음을 제거한 후 음향 신호 특징을 추출할 수 있도록 한다.

[0044] 이로 인해 본 실시예에 따른 음향 센서 기기(10)는, 네트워크 환경에서 적응 알고리즘을 사용하지 않더라도, 기준 배경 잡음 스펙트럼을 이용하여 다양한 실내 환경에서 적응적으로 지금 현재의 음향 상황에 대한 인지를 수행할 수 있다.

[0045] 이와 같은 본 실시예에 따른 음향 분석 방법을 도 1 내지 도 4를 참조하여 설명하면 다음과 같다. 여기서 도 3은 본 발명의 실시예에 따른 제한된 네트워크 환경에서의 음향 분석 방법에 따른 흐름도이다. 그리고 도 4는 도

3의 기준 배경 잡음 스펙트럼을 저장하는 단계에 대한 상세 흐름도이다.

[0046] 먼저 S10단계에서 음향 센서 기기(10)는 기준 배경 잡음 스펙트럼을 저장한다. 기준 배경 잡음 스펙트럼은 디폴트 형태로 음향 센서 기기(10)에 저장되어 있을 수도 있고, 제어 기기(20)로부터 수신하여 저장할 수 있다.

[0047] S10단계에 따른 제어 기기(20)로부터 기준 배경 잡음 스펙트럼을 수신하여 저장하는 과정을 도 4를 참조하여 설명하면 다음과 같다.

[0048] S11단계에서 제어 기기(20)는 복수의 실내 환경에 대한 배경 잡음을 입력받는다. 다음으로 S13단계에서 제어 기기(20)는 입력된 배경 잡음의 스펙트럼을 분석한다. 다음으로 S15단계에서 제어 기기(20)는 분석한 배경 잡음 스펙트럼의 평균치를 산출한다. 이어서 S17단계에서 제어 기기(20)는 산출한 평균치를 기준 배경 잡음 스펙트럼으로 설정한다. 그리고 S19단계에서 제어 기기(20)는 설정한 기준 배경 잡음 스펙트럼을 저장하고, 저장한 기준 배경 잡음 스펙트럼을 음향 센서 기기(10)로 제공한다.

[0049] 다음으로 S20단계에서 음향 센서 기기(10)는 음향 센서 기기(10)가 설치된 특정 실내 환경에서 음향을 입력받는다.

[0050] 다음으로 S30단계에서 음향 센서 기기(10)는 입력된 음향의 스펙트럼을 분석한다. 즉 음향 센서 기기(10)는 입력된 음향을 N 포인트로 FFT 수행하여 N 포인트 스펙트럼을 생성한다. 그리고 음향 센서 기기(10)는 생성한 N 포인트 스펙트럼을 주파수 밴드별로 에너지를 산출하여 입력된 음향의 스펙트럼을 생성한다.

[0051] 다음으로 S40단계에서 음향 센서 기기(10)는 기준 배경 잡음 스펙트럼을 기반으로 분석한 음향의 스펙트럼에 포함된 배경 잡음 스펙트럼을 제거한다. 즉 음향 센서 기기(10)는 기준 배경 잡음 스펙트럼을 주파수 밴드별로 에너지를 산출한다. 음향 센서 기기(10)는 기준 배경 잡음 스펙트럼의 주파수 밴드별 에너지에서 입력된 음향의 스펙트럼의 주파수 밴드별 에너지를 차감하여 입력된 음향의 배경 잡음 스펙트럼을 제거한다.

[0052] 다음으로 S50단계에서 음향 센서 기기(10)는 배경 잡음 스펙트럼이 제거된 음향 스펙트럼을 음향 신호로 복원한다.

[0053] 다음으로 S60단계에서 음향 센서 기기(10)는 복원한 음향 신호로부터 음향 신호 특징을 추출한다.

[0054] 이어서 S70단계에서 음향 센서 기기(10)는 추출한 음향 신호 특징으로부터 특정 실내 환경에서의 음향 상황을 인지한다.

[0055] 그리고 S80단계에서 음향 센서 기기(10)는 인지한 음향 상황을 제어 기기(20)로 전송한다.

[0056] 이후 제어 기기(20)는 수신한 음향 상향에 따른 기능을 수행한다.

[0057] 한편, 본 명세서와 도면에 개시된 실시예들은 이해를 돕기 위해 특정 예를 제시한 것에 지나지 않으며, 본 발명의 범위를 한정하고자 하는 것은 아니다. 여기에 개시된 실시예들 이외에도 본 발명의 기술적 사상에 바탕을 둔 다른 변형예들이 실시 가능하다는 것은, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게는 자명한 것이다.

부호의 설명

[0058] 10 : 음향 센서 기기

11 : 저장부

12 : 음향 입력부

13 : 스펙트럼 분석부

14 : 배경 잡음 제거부

15 : 음향 신호 추출부

16 : 분류기

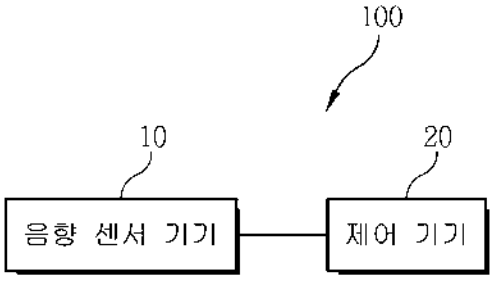
17 : 통신부

20 : 제어 기기

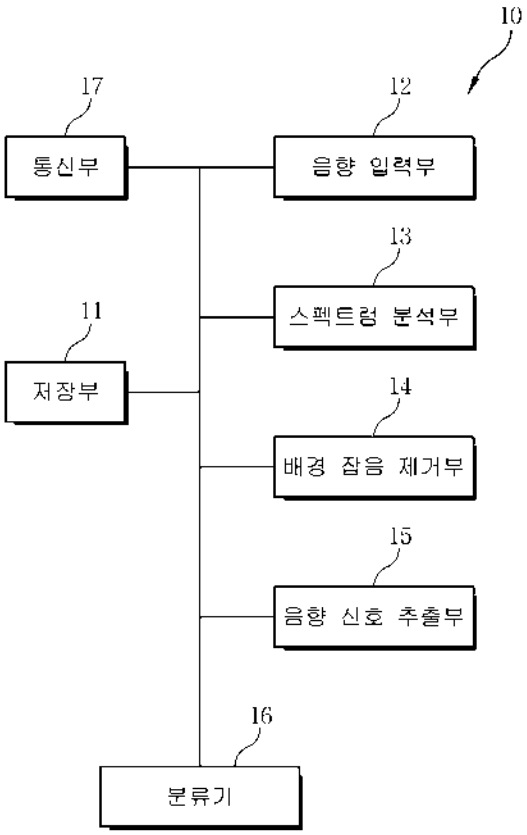
100 : 음향 분석 시스템

도면

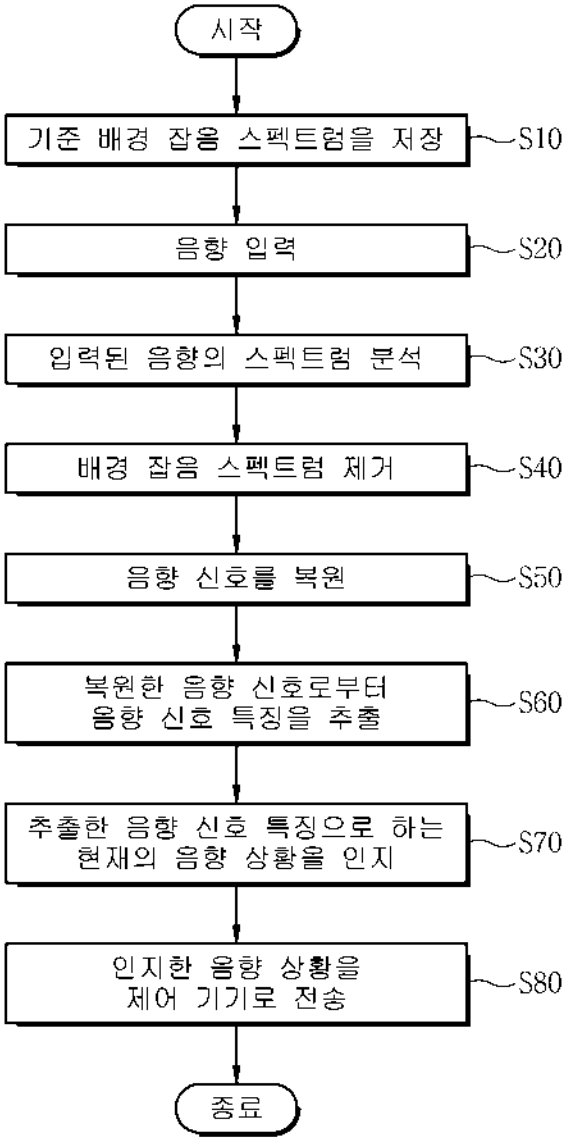
도면1



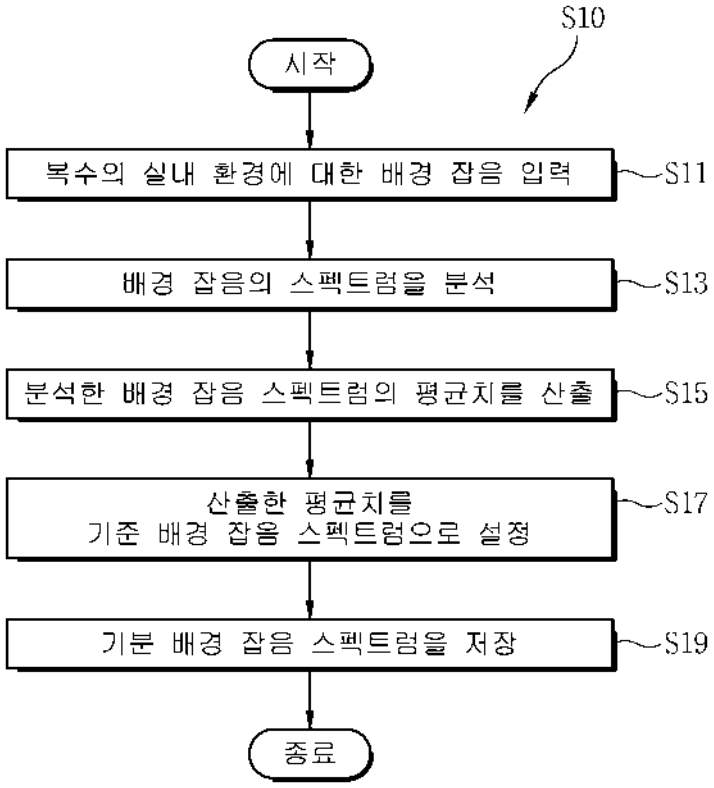
도면2



도면3


시작
기준 배경 잡음 스펙트럼을 저장
S10
음향 입력
S20
입력된 음향의 스펙트럼 분석
S30
배경 잡음 스펙트럼 제거
S40
음향 신호를 복원
S50
복원한 음향 신호로부터
음향 신호 특징을 추출
S60
추출한 음향 신호 특징으로 하는
현재의 음향 상황을 인지
S70
인지한 음향 상황을
제어 기기로 전송
S80
종료

도면4


S10
시작
복수의 실내 환경에 대한 배경 잡음 입력
S11
배경 잡음의 스펙트럼을 분석
S13
분석한 배경 잡음 스펙트럼의 평균치를 산출
S15
산출한 평균치를
기준 배경 잡음 스펙트럼으로 설정
S17
기분 배경 잡음 스펙트럼을 저장
S19
종료

■ 기술명 : 분산 HEVC 트랜스코더 기술
(Technology for HEVC Distributed Transcoder System SW)

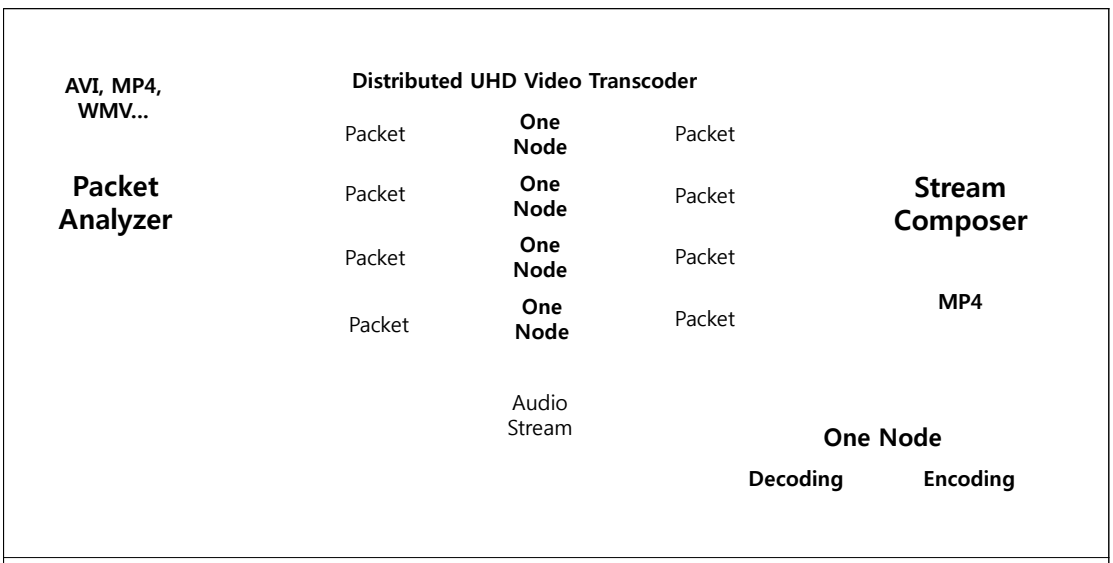
산업기술분류	정보통신/이동통신/이동통신 단말기(300103) 방송/방송미디어장비 · 단말/단말기기
Key-word(국문)	고효율 코덱, 소프트웨어 인코더, 트랜스코더, 분산 병렬 처리
Key-word(영문)	HEVC, SW encoder, transcoder, distributed parallel processing

■ 기술의 개요

○ (배경) 최근 UHD 방송의 UHD 영상과 인터넷 스트리밍 분야의 FHD(Full HD) 영상의 초고압축에 대한 요구가 높아짐에 따라 막대한 연산량이 소요되는 UHD 영상의 트랜스코딩에 분산 병렬 처리를 통한 고속화 기술 필요

○ (개요) 본 기술은 ISO/IEC 23008-2 HEVC 비디오 코덱 표준을 만족하면서 다양한 동영상(MP4, MKV, TS, AVI, WMV 등)을 입력으로 받아 다양한 해상도의 HEVC/MP4 미디어로 실시간 분산 트랜스코딩할 수 있는 시스템 SW 기술

< 기술 개요도 >



[클라우드 기반 분산 HEVC 트랜스코더 시스템 SW]

■ 기술의 구현수준(TRL)

■ 기술의 장점(경쟁기술과의 차별성)

○ ISO/IEC 23008-2 HEVC 비디오 코덱 표준 만족
- Main / Main 10 / Main Still Picture Profile 만족
- Main 10/12 4:2:2 Profile 만족

○ 4K@60p UHD 콘텐츠 실시간 분산 트랜스코딩 지원
- 하나의 동영상 입력에 대해 최대 4개의 해상도 동시 출력 기능 지원
- 분산 처리 구조이므로, 컴퓨팅 노드를 추가함으로써 성능 확장이 용이

■ 활용범위 및 응용분야

[N 스크린 - 인터넷 동영상 서비스]

[HEVC 비디오 콘텐츠]

■ 지식재산권 현황

구분	발명의 명칭	출원번호 (출원일)	등록번호 (등록일)
특허	HEVC 부호화기에서의 적응적인 CU depth 범위 예측 방법	2014-0086735 (2014.07.10)	10-1616461 (2016.04.22)
특허	HEVC 인코더를 위한 고속 모드 결정 방법	2015-0057710 (2015.04.24)	

(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2016년04월29일
(11) 등록번호 10-1616461
(24) 등록일자 2016년04월22일

(51) 국제특허분류(Int. Cl.)
H04N 19/70 (2014.01) *H04N 19/50* (2014.01)
(21) 출원번호 10-2014-0086735
(22) 출원일자 2014년07월10일
심사청구일자 2014년07월10일
(65) 공개번호 10-2016-0007929
(43) 공개일자 2016년01월21일
(56) 선행기술조사문헌
KR1020140008984 A*
KR1020120082606 A
KR1020140076652 A
KR1020140076646 A
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
김용환
경기 안양시 동안구 귀인로 294, 306동 502호 (평촌동, 꿈마을건영아파트)
김동혁
경기 광주시 초월읍 현산로 130-72, 102동 202호 (위드팰리스)
박지영
경기 성남시 분당구 정자일로 248, 613동 2802호 (정자동, 파크뷰)
(74) 대리인
남충우

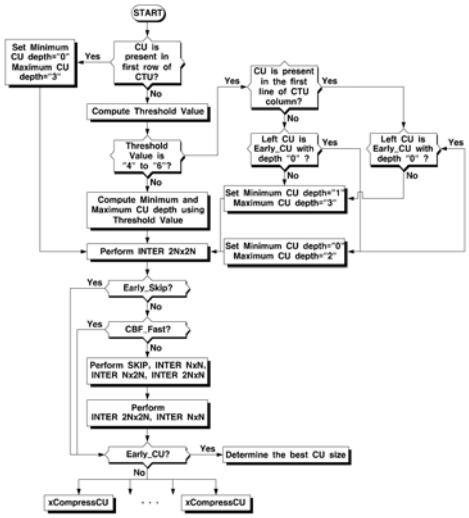
전체 청구항 수 : 총 4 항 심사관 : 조우연

(54) 발명의 명칭 **HEVC 부호화기에서의 적응적인 CU depth 범위 예측 방법**

(57) 요 약

HEVC 부호화기의 속도 향상을 위한 적응적인 CU depth 범위 예측 방법이 제공된다. 본 발명의 실시예에 따른 CU depth 범위 예측 방법은, 부호화할 현재 CU의 주변 CU들 중 일부 z-오더 인덱스에 대한 depth들을 획득하여, 현재 CU의 depth 범위를 예측한다. 이때, 획득하는 것은 한 라인 또는 한 z-오더 인덱스에 대한 depth들이다. 이에 의해, HEVC 영상 부호화기에서 연산량이 절감되어, 속도 향상을 기대할 수 있다.

대 표 도 - 도6



이 발명을 지원한 국가연구개발사업
과제고유번호 1711009201
부처명 미래창조과학부
연구관리전문기관 서울산업통상진흥원
연구사업명 선울전략산업 지원사업
연구과제명 클라우드 기반 UHD 방송콘텐츠 스트리밍 서비스 기술 개발
기 여 율 1/1
주관기관 전자부품연구원
연구기간 2013.12.01 ~ 2016.09.30

명 세 서

청구범위

청구항 1

부호화할 현재 CU(coding unit)의 주변 CU들 중 일부 z-오더 인덱스(z-order index)에 대한 depth들을 획득하는 단계; 및

상기 획득단계에서 획득한 depth들을 이용하여, 상기 현재 CU의 depth 범위를 예측하는 단계;를 포함하고,

상기 예측단계는,

상기 획득단계에서 획득한 depth들을 합산하여 임계값을 산출하는 단계; 및

상기 임계값을 기초로, 상기 현재 CU의 depth 범위를 예측하는 단계;를 포함하고,

상기 획득단계는,

주변 CU들 중, Left CU 및 Above CU에 대해서는 상기 현재 CU에 인접한 한 라인에 포함된 z-오더 인덱스에 대한 depth들 중 최대값을 획득하고, Above-Left CU 및 Above-Right CU에 대해서는 상기 현재 CU에 가장 인접한 한 포인트의 z-오더 인덱스에 대한 depth를 획득하며,

상기 예측단계는,

상기 임계값이 특정 범위에 속하는 경우, 상기 현재 CU의 depth 범위를,

Left CU 또는 Above CU가, depth가 0이고 Early CU 조건을 만족하면, 제1 depth 범위로 예측하고,

만족하지 않으면, 제2 depth 범위로 예측하는 것을 특징으로 하는 CU depth 범위 예측 방법.

청구항 2

삭제

청구항 3

삭제

청구항 4

삭제

청구항 5

제 1항에 있어서,

상기 현재 CU의 depth 범위는,

상기 임계값에 비례하는 것을 특징으로 하는 CU depth 범위 예측 방법.

청구항 6

삭제

청구항 7

제 1항에 있어서,

프레임에서의 첫 번째 라인의 CTB들에 포함된 CU들의 depth 범위는 0~3으로 예측하는 단계;를 더 포함하는 것을

특징으로 하는 CU depth 범위 예측 방법.

청구항 8

부호화할 현재 CU(coding unit)의 주변 CU들 중 일부 z-오더 인덱스(z-order index)에 대한 depth들을 획득하는 단계; 및

상기 획득단계에서 획득한 depth들을 이용하여, 상기 현재 CU의 depth 범위를 예측하는 단계;를 포함하고,

상기 예측단계는,

상기 획득단계에서 획득한 depth들을 합산하여 임계값을 산출하는 단계; 및

상기 임계값을 기초로, 상기 현재 CU의 depth 범위를 예측하는 단계;를 포함하고,

상기 획득단계는,

주변 CU들 중, Left CU 및 Above CU에 대해서는 상기 현재 CU에 인접한 한 라인에 포함된 z-오더 인덱스에 대한 depth들 중 최대값을 획득하고, Above-Left CU 및 Above-Right CU에 대해서는 상기 현재 CU에 가장 인접한 한 포인트의 z-오더 인덱스에 대한 depth를 획득하며,

상기 예측단계는,

상기 임계값이 특정 범위에 속하는 경우, 상기 현재 CU의 depth 범위를,

Left CU 또는 Above CU가, depth가 0이고 Early CU 조건을 만족하면, 제1 depth 범위로 예측하고,

만족하지 않으면, 제2 depth 범위로 예측하는 것을 특징으로 하는 CU depth 범위 예측 방법을 수행할 수 있는 프로그램이 기록된 컴퓨터로 읽을 수 있는 기록매체.

발명의 설명

기 술 분 야

[0001] 본 발명은 영상의 효율적인 부호화 방법에 관한 것으로, 더욱 상세하게는 현재 CU의 주변 CU 정보를 이용하여 HEVC 표준 영상의 효율적인 고속 부호화를 가능하게 하는 방법 및 시스템에 관한 것이다.

배 경 기 술

[0002] 1) HEVC(High Efficiency Video Coding)

[0003] HEVC는 H.264/AVC 이후에 개발된 새로운 비디오 코딩 표준이고, H.264/AVC에 비해 대략 50% 정도의 압축률 향상을 제공한다. HEVC의 코딩 구조는 H.264/AVC와 거의 유사하지만 다른 점이 두 가지 존재한다.

[0004] 첫째는, 매우 유연한 쿼드트리(quadtree) 파티션 구조를 이용하여 비디오 데이터가 압축된다는 점이다. 쿼드트리 파티션 구조에 의해 HEVC는 다양한 해상도의 영상을 최적의 파티션으로 압축할 수 있다.

[0005] HEVC의 압축 구조는 CU (coding unit), PU(prediction unit), TU(transform unit)라는 세 가지로 구성된다.

[0006] CU는 H.264/AVC의 매크로블록(macroblock)과 유사한 것으로, 압축을 위한 기본 단위를 나타낸다. HEVC에서 CU의 크기는 H.264/AVC의 고정된 매크로블록 크기(16x16)와는 다르게 다양한 크기를 가질 수 있다. 현재 HEVC에서 CU 크기는 8x8에서 64x64까지 가질 수 있으며, 가장 큰 CU(largest CU : LCU)와, 가장 작은 CU (smallest CU : SCU)를 SPS(sequence parameter set)에 기술하게 되어있다. 또한, 하나의 CU 내에서 PU는 1~2개로 구성되는 반면, TU는 하나의 CU내에서 또 다른 쿼드트리 구조로 다양한 크기를 가질 수 있다.

[0007] H.264/AVC와 비교하여 두 번째로 다른 점은, 코딩 효율을 위해 HEVC의 표준에 새로운 코딩 툴들이 포함된다는 것이다. 새로운 인트라 예측 방향이 추가되었으며, large transform(16x16, 32x32), AMP(asymmetric motion partitions), SAO(sample adaptive offset) 등이 추가 되었다.

[0008] 도 1은 HEVC 부호화기의 블록도를 나타내며, 특히, CU의 경우 모든 시퀀스에 대해 고정된 범위에 대해 인코딩을 수행한다. CU의 depth 값이 작을수록 변화가 적고 동일한 영역이 많이 존재하는 영상에 적합하고, 반대로 클수록 변화가 많이 존재하는 영상이나 사물이 많은 영상에 적합하다. 그래서 고정된 범위를 갖는 CU의 depth 값에 대해 모든 부호화 과정을 진행하는 것은 부호화 시간을 증가시키며 비효율적이다. 따라서 영상의 특성에 따라 불필요한 CU의 분할 과정과 이에 발생하는 계산 과정을 줄여 부호화의 속도를 증가시킬 필요성이 있다.

[0009] 결과적으로, HEVC는 H.264/AVC에 비해 부호화 과정이 복잡해 졌으며, 따라서 빠른 부호화를 위한 구현 방법이 필요하게 된 것이다.

[0010] 2) HEVC 기반 부호화 블록 결정 방법

[0011] HEVC는 쿼드트리(quard-tree) 구조를 사용하여 계층에 따라 재귀적으로 CU를 더 작은 유닛으로 분할한다. 최적의 CU 크기를 결정하기 위하여 각 CU에서 각각의 PU 크기 및 TU 크기에 대해 가장 작은 비용을 갖는 블록 크기를 조사하게 되며, 이때 CU 크기 마다 PU는 다양한 모양의 예측블록으로 분할하여 비용조사를 하고, TU는 재귀적인 트리 구조로 분할을 하여 각각의 크기에 대해 비용조사를 한다.

[0012] 도 2는 임의 크기의 CU로부터 최적의 PU 크기 및 TU 크기 결정 방법을 나타낸다. 최적의 비용을 갖는 CU를 찾기 위해 순차적으로 가장 큰 크기인 LCU(large coding unit)부터 가장 작은 크기인 SCU(small coding unit)까지 재귀적인 쿼드트리 구조에 따라 비용을 계산하게 된다. 각 CU 크기 별로 매번 PU의 여러 모드에 대한 예측 비용을 전부 계산하며, 각 PU에 대한 비용조사를 할 때, TU의 비용조사 또한 수행한다. TU는 4개의 하위블록으로 재귀적으로 분할될 수 있다.

[0013] HEVC의 부호화 과정에서 CU에 따라 PU, TU의 조합 수가 매우 크며 최적의 CU, PU, TU를 결정하기 위해 많은 계산 과정이 필요하다는 것을 알 수 있다. 이를 해결하기 위해 도 2에 도시된 아래의 ①, ②, ③ 과 같은 고속화 방법들이 연구되었다.

[0014] ① Early_SKIP

[0015] Inter 2Nx2N 모드를 조사한 다음 Coded Block Flag(CBF) 값이 0 이고 Motion vector difference(MVD)가 (0,0) 값을 갖는 경우 다른 형태의 PU들을 조사하지 않고, 바로 CU를 분할하여 다음 depth의 CU 조사과정을 이어간다. 여기서, CBF는 0이 아닌 잔차 신호의 존재 유무를 알려주는 flag이며, MVD는 참조 영상과 현재 영상과의 움직임 정보에 대한 차이 값이다.

[0016] ② CBF_Fast

[0017] CBF의 값에 따른 고속 PU 결정 방법(CBF Fast Mode Setting; CFM)으로, CBF의 값에 따라 현재 비용조사 중인 PU에서 다른 형태의 PU로 더 이상 비용조사를 하지 않고 조기종료 하는 방법이다. 즉 CBF가 0인 경우, 남은 PU들을 검색하지 않는다.

[0018] ③ Early_CU

[0019] CU의 조사과정을 조기 종료하는 방법으로, CU의 PU에 대한 조사 과정 중에 최적 모드가 SKIP 모드로 결정되면 하위 블록으로의 재귀 CU 분할을 하지 않는다.

[0020] 이러한 HEVC의 고속화 방법들이 연구되었지만 아직 실시간 환경에서 사용하기에는 부호화 시간이 많이 소비된다. 따라서, 기존 고속화 방법들과 더불어 실시간 환경을 위한 새로운 HEVC 구조 설계와 고속화 방법의 개발이 필수적이다.

발명의 내용

해결하려는 과제

[0021] 본 발명은 상기와 같은 문제점을 해결하기 위하여 안출된 것으로서, 본 발명의 목적은, H.264/AVC에 비해 부호화 과정이 복잡하게 된 HEVC 부호화기에서 Temporally co-located CU 정보를 이용하지 않으면서도 최적의 CU depth의 범위를 예측하는 방법을 제공함에 있다.

과제의 해결 수단

[0022] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른, CU depth 범위 예측 방법은, 부호화할 현재 CU(coding unit)의 주변 CU들 중 일부 z-오더 인덱스(z-order index : 4×4)에 대한 depth들을 획득하는 단계; 및 상기 획득단계에서 획득한 depth들을 이용하여, 상기 현재 CU의 depth 범위를 예측하는 단계;를 포함한다.

[0023] 그리고, 상기 획득단계는, 주변 CU들 중, 일부에 대해서는 한 라인에 포함된 z-오더 인덱스에 대한 depth들 중 최대값을 획득하고, 나머지에 대해서는 한 z-오더 인덱스에 대한 depth를 획득할 수 있다.

[0024] 또한, 상기 일부에 해당하는 주변 CU들은, Left CU 및 Above CU를 포함하고, 상기 나머지에 해당하는 주변 CU들은, Above-Left CU 및 Above-Right CU를 포함할 수 있다.

[0025] 그리고, 상기 예측단계는, 상기 획득단계에서 획득한 depth들을 합산하여 임계값을 산출하는 단계; 및 상기 임계값을 기초로, 상기 현재 CU의 depth 범위를 예측하는 단계;를 포함할 수 있다.

[0026] 또한, 상기 현재 CU의 depth 범위는, 상기 임계값에 비례할 수 있다.

[0027] 그리고, 상기 임계값이 특정 범위에 속하는 경우, 상기 현재 CU의 depth 범위는, Left 또는 Above CU의 depth가 0이고, Early CU 조건을 만족하면, 제1 depth 범위로 예측하고, 만족하지 않으면, 제2 depth 범위로 예측할 수 있다.

[0028] 또한, 본 발명의 일 실시예에 따른, CU depth 범위 예측 방법은, 프레임에서의 첫 번째 라인의 CTB들에 포함된 CU들의 depth 범위는 0~3으로 예측하는 단계;를 더 포함할 수 있다.

[0029] 한편, 본 발명의 다른 실시예에 따른, 컴퓨터로 읽을 수 있는 기록매체는, 부호화할 현재 CU(coding unit)의 주변 CU들 중 일부 z-오더 인덱스에 대한 depth들을 획득하는 단계; 및 상기 획득단계에서 획득한 depth들을 이용하여, 상기 현재 CU의 depth 범위를 예측하는 단계;를 포함하는 것을 특징으로 하는 CU depth 범위 예측 방법을 수행할 수 있는 프로그램이 기록된다.

발명의 효과

[0030] 이상 설명한 바와 같이, 본 발명의 실시예들에 따르면, HEVC 부호화기에서 Temporally co-located CU 정보를 이용하지 않으면서도 최적의 CU depth의 범위를 예측할 수 있게 된다. 이에 따라, HEVC 영상 부호화기에서 연산량이 절감되어, 속도 향상을 기대할 수 있다.

[0031] 또한, 본 발명의 실시예들에 따르면, HEVC 부호화기의 속도 향상을 위한 CU의 depth 범위 예측시 주변 CU 정보만을 이용하여 부호화 속도를 향상시킬 수 있다. 이에 의해, 메모리 사용량과 데이터를 불러오는 시간을 줄일 수 있다.

도면의 간단한 설명

[0032] 도 1은 종래의 HEVC 부호화기를 도시한 블록도

도 2는 재귀적 CU, PU, TU 결정 과정의 설명에 제공되는 흐름도,

도 3은 부호화 될 CU의 depth 예측에 참고 되는 CU들을 도시한 도면,

도 4는 임계값(Threshold Value)에 따른 현재 CU의 depth 범위를 나타낸 표,

도 5는 임계값이 4~6인 경우 Left CU와 Above CU의 정보를 활용하는 방법을 나타낸 도면,

도 6은 본 발명의 실시예에 따른, HEVC 부호화기의 속도 향상을 위한 CU의 depth 범위 예측 방법의 설명에 제공되는 순서도, 그리고,

도 7은 프레임 단위에서의 CU depth 범위 예측 처리 방법을 도시한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0033] 이하에서는 도면을 참조하여 본 발명을 보다 상세하게 설명한다.

[0034] 본 발명의 실시예에 따른 HEVC 부호화기의 속도 향상을 위한 CU의 depth 범위 예측 방법은 본적으로 CU 레벨에서부터 시작한다.

[0035] 도 3에는 현재 부호화되려는 CU의 depth를 예측하기 위한 방법으로, 주변(spatially neighboring) CU들의 depth 정보를 이용하는 방법이 도시되어 있다.

[0036] 먼저, 부호화될 CU의 depth 범위를 정확하게 예측하기 위해 현재 CU의 주변 CU들(Left CU, Above-Left CU, Above CU, Above-Right CU)의 depth 정보를 획득한다.

[0037] 구체적으로, 도 3에 도시된 바와 같이, Left CU와 Above CU의 경우, 현재 CU와 높은 상관도를 가지고 있기 때문에 가장 근접한 한 라인에 포함된 z-오더 인덱스(z-order index : 4×4)들에 대한 depth 정보를 획득한다. 하지만, Above-Left CU와 Above-Right CU는 상대적으로 낮은 상관도를 가지고 있기 때문에 가장 근접한 한 z-오더 인덱스에 대해서만 depth 정보만을 획득한다.

[0038] 그리고, 획득된 정보들로부터 아래의 수학식 1에 따라 Depth들($Depth_0$, $Depth_1$, $Depth_2$ 및 $Depth_3$)을 산출한다.

[0039] [수학식 1]

[0040] $Depth_0$ = MAX(Left CU depth)

[0041] $Depth_1$ = MAX(Above CU depth)

[0042] $Depth_2$ = Above-Left CU depth

[0043] $Depth_3$ = Above-Right CU depth

[0044] $Depth_0$은 Left CU에서 현재 CU에 인접한 한 라인에 포함된 z-오더 인덱스들에 대한 depth들 중 최대값이고, $Depth_1$은 Above CU에서 현재 CU에 인접한 한 라인에 포함된 z-오더 인덱스들에 대한 depth들 중 최대값이다.

[0045] 한편, $Depth_2$는 Above-Left CU에서 현재 CU에 가장 인접한 한 z-오더 인덱스의 depth이고, $Depth_3$은 Above-Right CU에서 현재 CU에 가장 인접한 한 z-오더 인덱스의 depth이다.

[0046] 이후, 수학식 1을 통해 산출한 Depth들($Depth_0$, $Depth_1$, $Depth_2$ 및 $Depth_3$)로부터 임계값(Threshold Value)을 산출하는데, 임계값은 아래의 수학식 2에 나타난 바와 같이 Depth들($Depth_0$, $Depth_1$, $Depth_2$ 및 $Depth_3$)의 합을 의미한다.

[0047] [수학식 2]

[0048] $$Threshold\ Value = \sum_{i=0}^{3} Depth_i$$

[0049] 영상의 특성에 맞게 CU의 depth가 결정된다. 따라서, 현재 CU를 중심으로 주변 CU들의 depth 값을 합한 임계값이 크다면 변화가 많은 영역임을 나타낸다. 반대로 이 임계값이 작다면 변화가 적은 영역임을 나타낸다.

[0050] 이에 따라, 임계값을 이용하여, 현재 CU의 depth 범위를 결정할 수 있다. 도 4는 임계값에 따른 현재 CU(부호화될 CU)의 depth 범위를 나타낸 표이다.

[0051] 도 4에 도시된 표를 참조하면, 주변 CU들의 depth 값이 작아 임계값이 작을수록 현재 CU의 최적 depth는 0, 1, 2가 선택될 확률이 높다. 반면, 주변 CU들의 depth 값이 커서 임계값이 클수록 현재 CU의 최적 depth는 1, 2, 3이 선택될 확률이 높음을 확인할 수 있다.

[0052] 한편, 도 4에 도시된 바와 같이, 임계값이 4~6 값인 경우, CU의 depth가 0~3까지 다양한 범위로 혼용됨을 확인할 수 있다. 구체적으로, CU의 depth 범위는 0~2 또는 1~3가 된다.

[0053] 어떠한 범위로 예측할지 결정하기 위해, CU의 위치에 따라 Left CU 혹은 Above CU의 정보를 이용한다. 구체적으로, 1) Left 또는 Above CU의 depth가 0이고, Early CU 조건을 만족한다면, 현재 CU의 depth 범위를 0~2로 예측하고, 2) 이를 만족하지 않는다면, 현재 CU의 depth 범위를 1~3으로 예측한다.

[0054] 도 5에 임계값이 4~6인 경우 Left CU와 Above CU의 정보를 활용하는 방법을 나타내었다.

[0055] 도 5에서, ①번 영역의 CTU들은 왼쪽 첫 번째 CTU 라인으로서 Left CU가 존재하지 않는 영역이다. 따라서, 해당 영역에 대해서 depth를 결정할 경우에는 Above CU의 정보를 활용한다. 한편, ②번 영역은 Left CU와 Above CU가 둘 다 존재하는 영역이지만, 부호화기의 속도 향상을 위해 Left CU의 정보만을 이용하여 depth를 결정한다.

[0056] 도 6은 본 발명의 실시예에 따른, HEVC 부호화기의 속도 향상을 위한 CU의 depth 범위 예측 방법의 설명에 제공되는 순서도이다.

[0057] 본 발명의 실시예에 따른 CU depth 범위 예측 방법에서는, 고속 부호화 기법인 Early SKIP, CBF Fast, Early CU이 활용된다.

[0058] 구체적으로, 도 3에 도시된 바와 같이 현재 CU(부호화될 CU)의 주변 CU들에 대한 depth 정보와 수학식 1을 이용하여 Depth들($Depth_0$, $Depth_1$, $Depth_2$ 및 $Depth_3$)을 산출하고, 산출된 Depth들과 수학식 2를 이용하여 임계값을 산출한다.

[0059] 다음, 산출된 임계값에 대해 도 4에 도시된 표를 참조하여, 현재 CU(부호화될 CU)의 depth 범위를 예측한다.

[0060] 도 4와 도 6을 통해 할 수 있는 바와 같이, 1) 임계값이 "0"인 경우, CU의 depth 범위는 "0,1"로 예측하고, 2) 임계값이 "1~3"인 경우, CU의 depth 범위는 "0~2"으로 예측하며, 3) 임계값이 "7~12"인 경우, CU의 depth 범위는 "1~3"으로 예측한다.

[0061] 한편, 임계값이 "4~6"인 경우, 1) Left 또는 Above CU의 depth가 0이고, Early CU 조건을 만족한다면, 현재 CU의 depth 범위를 "0~2"로 예측하고, 2) 이를 만족하지 않는다면, 현재 CU의 depth 범위를 "1~3"으로 예측한다.

[0062] 이후, 예측된 depth 범위에서 최소 depth 값에 대해 Inter 2Nx2N 과정을 수행하고, Early SKIP, CBF Fast 조건을 비교한다. 만약, Early SKIP, CBF Fast 조건이 참이라면, Early CU 조건으로 수행하고, 거짓이라면 SKIP, Inter NxN, Inter Nx2N, Inter 2NxN을 수행한다.

[0063] 한편, AMP는 고속 부호화 환경에서 속도를 저하시키기 때문에 생략한다. 이후, Intra 2Nx2N, Intra NxN을 수행하고 Early CU 조건을 비교한다. 만약, Early CU 조건이 참이라면 해당하는 CU depth를 최적의 depth로 결정하고, 거짓이라면 depth 값을 예측된 범위에서 1만큼 증가시켜 다시 Inter 2Nx2N 과정부터 진행하게 된다. 따라서 현재 부호화될 CU의 예측된 depth 범위에 대해서만 RDO(Rate Distortion Optimization)를 수행하여 최적의 CU depth를 찾는다.

[0064] 도 7에는 프레임 단위에서의 CU depth 범위 예측 처리 방법을 도시하였다. 본 발명의 실시예에 따른 CU depth 범위 예측 처리 방법에서는, 주변 (spatially neighboring) CU의 정보만을 이용하여 현재 CU의 depth 범위를 예측하기 때문에 주변 CU들의 정확한 정보가 필요하다.

[0065] 따라서, 주변 CU들의 정확한 정보를 획득하기 위해서는 도 7에 도시된 바와 같이 프레임에서의 첫 번째 라인의 CTB들은 본 발명의 실시예에서 제시한 방법을 적용하지 않는다. 즉, 첫 번째 라인의 CTB들에 포함되는 CU들의 depth는 0~3으로 예측한다.

[0066] 첫 번째 라인의 CTB들은 정확한 예측을 위해 사용되어질 주변 CU들(Above, Above-Left, Above-Right CU들)의 정보들이 부족하기 때문이다.

[0067] 본 발명의 실시예에 따른 HEVC 부호화기의 속도 향상을 위한 CU의 depth 범위 예측 방법은 주변 CU 정보만을 이용하여 부호화 속도를 향상시킬 수 있다. 이에 의해, 메모리 사용량과 데이터를 불러오는 시간을 줄일 수 있다.

[0068] 또한, 이상에서는 본 발명의 바람직한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야

에서 통상의 지식을 가진자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

부호의 설명

[0069] $Depth_0$: MAX(Left CU depth)

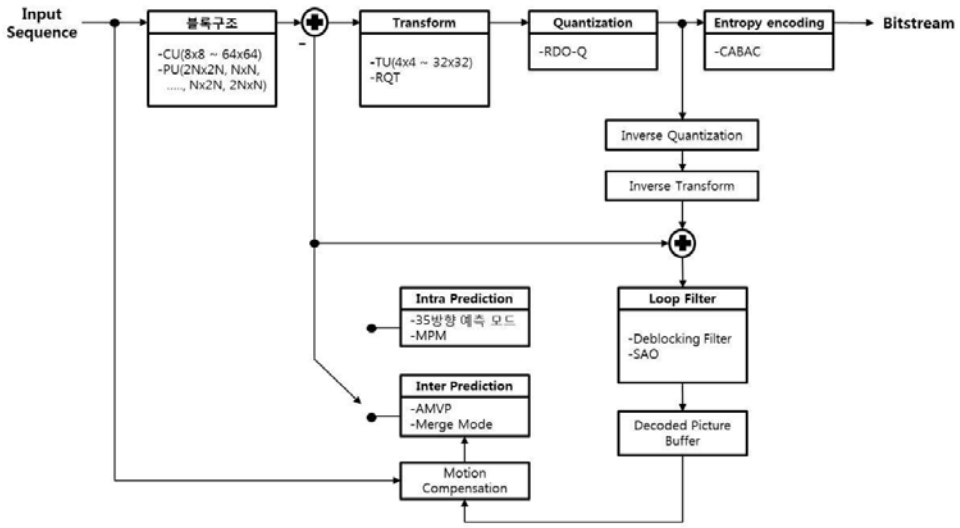
$Depth_1$: MAX(Above CU depth)

$Depth_2$: Above-Left CU depth

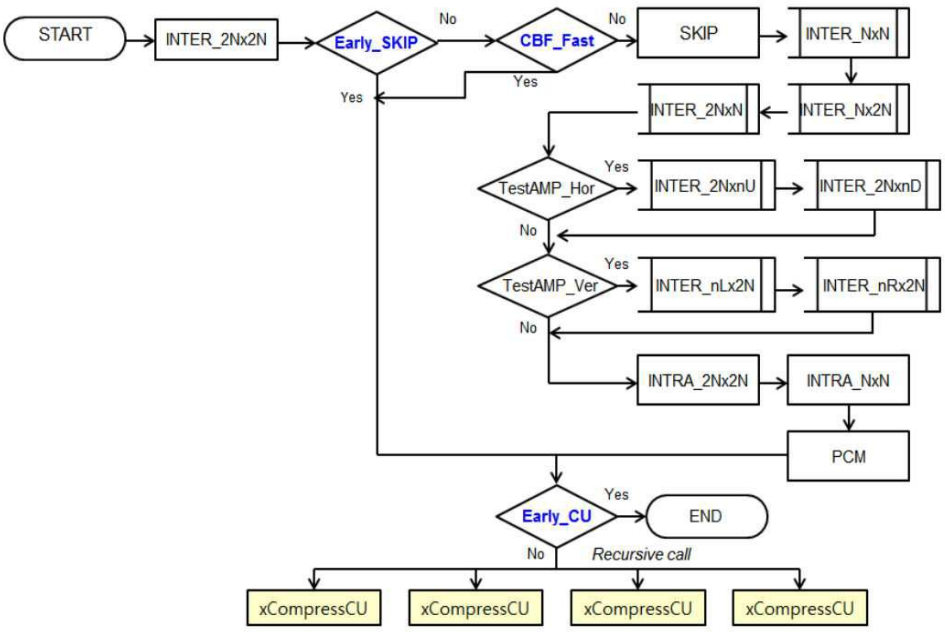
$Depth_3$: Above-Right CU depth

도면

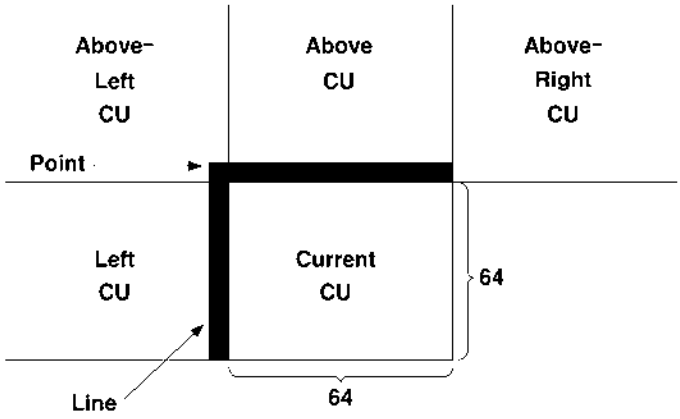
도면1



도면2



도면3

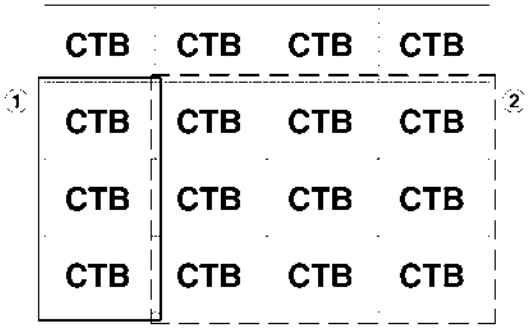


Spatially neighboring CU information

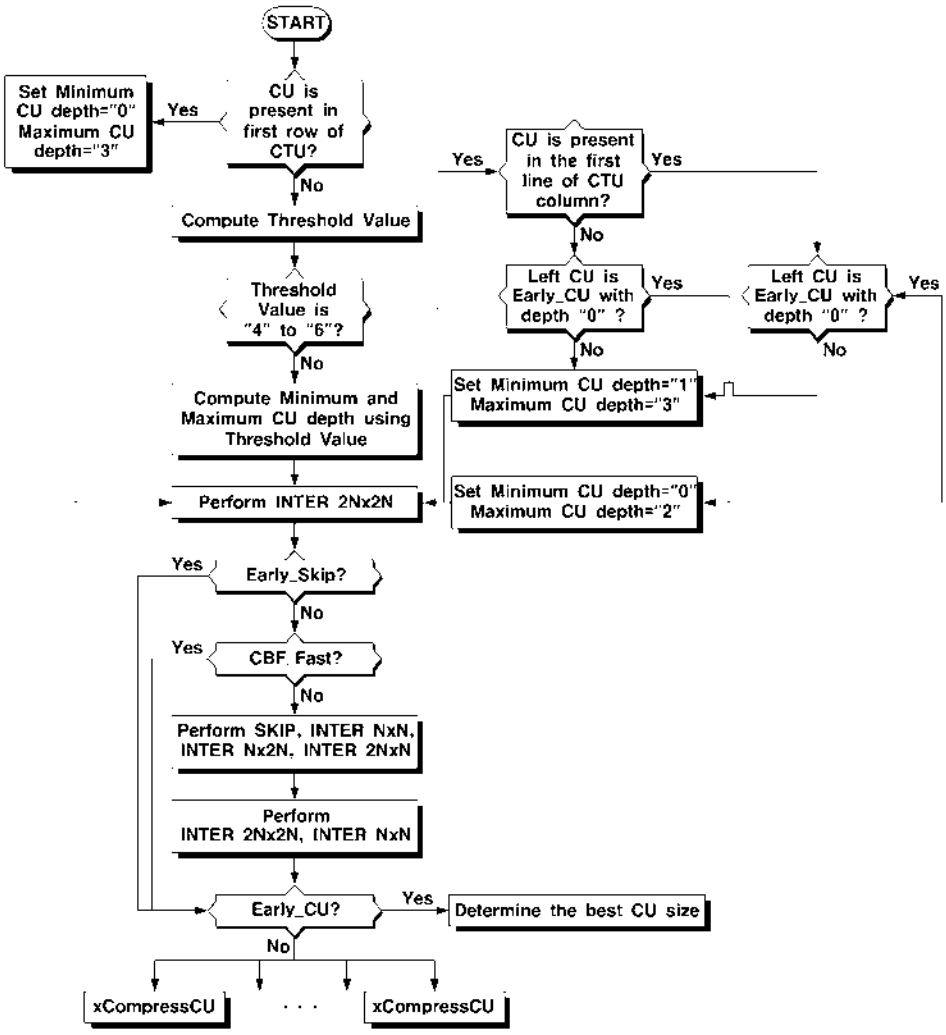
도면4

Threshold Value 범위	부호화 될 CU의 depth 범위
Threshold Value = 0	0,1
1 ≤ Threshold Value ≤ 3	0,1,2
4 ≤ Threshold Value ≤ 6	0,1,2 1,2,3
7 ≤ Threshold Value ≤ 12	1,2,3

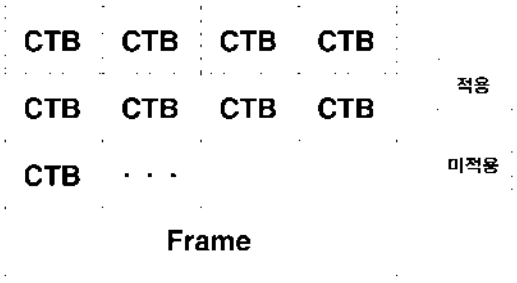
도면5



도면6


START
CU is present in first row of CTU?
Yes
Set Minimum CU depth="0" Maximum CU depth="3"
No
Compute Threshold Value
Threshold Value is "4" to "6"?
No
Compute Minimum and Maximum CU depth using Threshold Value
Perform INTER 2Nx2N
Yes
Early_Skip?
No
Yes
CBF Fast?
No
Perform SKIP, INTER NxN, INTER Nx2N, INTER 2NxN
Perform INTER 2Nx2N, INTER NxN
Early_CU?
Yes
Determine the best CU size
No
xCompressCU
. . .
xCompressCU
Yes
CU is present in the first line of CTU column?
Yes
No
Left CU is Early_CU with depth "0" ?
Yes
No
Set Minimum CU depth="1" Maximum CU depth="3"
Left CU is Early_CU with depth "0" ?
Yes
No
Set Minimum CU depth="0" Maximum CU depth="2"

도면7


CTB CTB CTB CTB
CTB CTB CTB CTB
CTB . . .
적용
미적용
Frame

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2016-0127231
(43) 공개일자 2016년11월03일

(51) 국제특허분류(Int. Cl.)
H04N 19/103 (2014.01) *H04N 19/119* (2014.01)
H04N 19/122 (2014.01)
(52) CPC특허분류
H04N 19/103 (2015.01)
H04N 19/119 (2015.01)
(21) 출원번호 10-2015-0057710
(22) 출원일자 2015년04월24일
심사청구일자 없음

(71) 출원인
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
김용환
경기도 안양시 동안구 귀인로 294 306동 502호 (평촌동,꿈마을건영아파트)
김동혁
경기도 광주시 초월읍 현산로 130-72 102동 202호 (도평리,위드팰리스)
박영수
경기도 용인시 기흥구 동백죽전대로 455-10, 104동 1301호 (동백동, 해든마을동문굿모닝힐아파트)
(74) 대리인
남충우

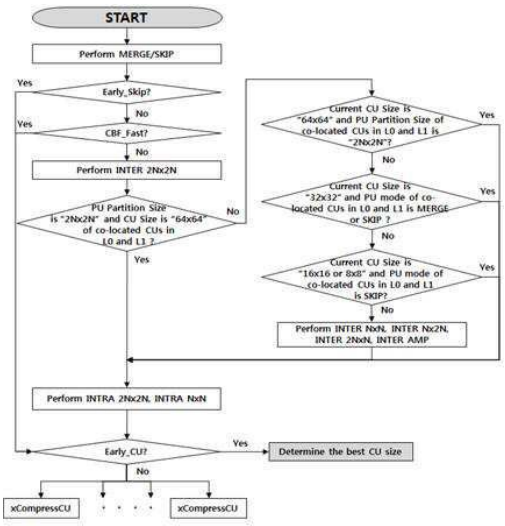
전체 청구항 수 : 총 6 항

(54) 발명의 명칭 **HEVC 인코더를 위한 고속 모드 결정 방법**

(57) 요 약

HEVC 인코더를 위한 고속 모드 결정 방법이 제공된다. 본 발명의 실시예에 따른 CU의 PU 모드 결정 방법은, 현재 CU의 크기 및 현재 CU에 대한 참조 픽처들에서 Co-located CU들의 PU 분할 크기들과 PU 모드들을 기초로, 남아 있는 PU 모드의 탐색을 스킵한다. 이에 의해, HEVC 부호화기에서 Temporally Co-located CU의 두 가지 best PU 정보를 이용하여 현재 CU의 INTER PU 모드 탐색을 조기에 종료할 수 있어, 부호화 속도는 증가시키고, 필요한 연산량은 감소시킬 수 있게 된다.

대 표 도 - 도16



(52) CPC특허분류
H04N 19/122 (2015.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 1711009201
부처명 미래창조과학부
연구관리전문기관 정보통신기술진흥센터
연구사업명 첨단융복합콘텐츠기술개발(미래부)
연구과제명 클라우드 기반 UHD 방송콘텐츠 스트리밍 서비스 기술 개발
기 여 율 1/1
주관기관 전자부품연구원
연구기간 2013.12.01 ~ 2014.09.30

명 세 서

청구범위

청구항 1

현재 CU의 크기 및 상기 현재 CU에 대한 참조 픽처들에서 Co-located CU들의 PU 분할 크기들과 PU 모드들을 파악하는 단계; 및

파악된 PU 분할 크기들과 PU 모드들을 기초로, 남아 있는 PU 모드의 탐색을 스킵하는 단계;를 포함하는 것을 특징으로 하는 CU의 PU 모드 결정 방법.

청구항 2

청구항 1에 있어서,

상기 스킵 단계는,

INTER Nx2N, INTER 2NxN, INTER AMP(Asymmetric Motion Partitions)의 탐색을 스킵하는 것을 특징으로 하는 CU의 PU 모드 결정 방법.

청구항 3

청구항 1에 있어서,

상기 스킵 단계는,

상기 Co-located CU들의 CU 크기가 64x64이고, 상기 Co-located CU들의 PU 분할 크기가 PART_2Nx2N 이면, 상기 탐색을 스킵하는 것을 특징으로 하는 CU의 PU 모드 결정 방법.

청구항 4

청구항 3에 있어서,

상기 스킵 단계는,

상기 현재 CU의 크기가 64x64이고, 상기 Co-located CU들의 PU 분할 크기가 2Nx2N이면, 상기 탐색을 스킵하는 것을 특징으로 하는 CU의 PU 모드 결정 방법.

청구항 5

청구항 4에 있어서,

상기 스킵 단계는,

상기 현재 CU의 크기가 32x32이고, 상기 Co-located CU들의 PU 모드가 MERGE 또는 SKIP이면, 상기 탐색을 스킵하고,

상기 현재 CU의 크기가 16x16 또는 8x8이고, 상기 Co-located CU들의 PU 모드가 SKIP이면, 상기 탐색을 스킵하는 것을 특징으로 하는 CU의 PU 모드 결정 방법.

청구항 6

현재 CU의 크기 및 상기 현재 CU에 대한 참조 픽처들에서 Co-located CU들의 PU 분할 크기들과 PU 모드들을 파악하는 단계; 및

파악된 PU 분할 크기들과 PU 모드들을 기초로, 남아 있는 PU 모드의 탐색을 스킵하는 단계;를 포함하는 것을 특징으로 하는 CU의 PU 모드 결정 방법을 수행할 수 있는 프로그램이 기록된 컴퓨터로 읽을 수 있는 기록매체.

발명의 설명

기 술 분 야

[0001] 본 발명은 영상의 효율적인 부호화 방법에 관한 것으로, 더욱 상세하게는 CU 부호화 과정에서 시간적 및 공간적으로 상관도가 높으면서도 기부호화된 CU들의 최적의 PU 모드 정보들을 이용하여 현재 CU의 PU 모드를 고속으로 결정하는 방법 및 장치에 관한 것이다.

배 경 기 술

[0002] 1) HEVC(High Efficiency Video Coding)

[0003] HEVC는 H.264/AVC 이후에 개발된 새로운 비디오 코딩 표준이고, H.264/AVC에 비해 대략 50% 정도의 압축율 향상을 제공한다. HEVC의 코딩 구조는 H.264/AVC와 거의 유사하지만 다른 점이 두 가지 존재한다.

[0004] 첫째는, 매우 유연한 쿼드트리(quadtree) 파티션 구조를 이용하여 비디오 데이터가 압축된다는 점이다. 쿼드트리 파티션 구조에 의해 HEVC는 다양한 해상도의 영상을 최적의 파티션으로 압축할 수 있다. HEVC의 압축 구조는 CU(coding unit), PU(prediction unit), TU(transform unit) 세 가지로 구성된다.

[0005] CU는 H.264/AVC의 매크로블록(macroblock)과 유사한 것으로, 압축을 위한 기본 단위를 나타낸다. HEVC에서 CU의 크기는 H.264/AVC의 고정된 매크로블록 크기(16x16)와는 다르게 다양한 크기를 가질 수 있다. 현재 HEVC에서 CU 크기는 8x8에서 64x64까지 가질 수 있으며, 가장 큰 CU(largest CU : LCU)와, 가장 작은 CU(smallest CU : SCU)를 sequence parameter set(SPS)에 기술하게 되어있다. 또한 하나의 CU내에서 PU는 1~2개로 구성되는 반면, TU는 하나의 CU내에서 또 다른 쿼드트리 구조로 다양한 크기를 가질 수 있다.

[0006] H.264/AVC와 비교하여 두 번째로 다른 점은, 코딩 효율을 위해 HEVC의 표준에 새로운 코딩 툴들이 포함된다는 것이다. 새로운 인트라 예측 방향이 추가되었으며, large transform(16x16, 32x32), asymmetric motion partitions(AMP), sample adaptive offset(SAO) 등이 추가되었다.

[0007] 도 1은 HEVC 부호화기의 블록도를 나타내며, 특히, PU의 경우 Early_SKIP, CBF_Fast의 고속 부호화 방법을 적용하지 않는다면 부호화되는 CU depth에 대해 모든 PU 모드를 탐색한다. 그래서, 모든 PU 모드에 대해서 탐색을 진행하는 것은 부호화 시간을 증가시키며 비효율적이다. 따라서, 영상의 특성에 따라 불필요한 PU 모드의 탐색과 이에 발생하는 계산 과정을 줄여 부호화의 속도를 증가시킬 필요성이 있다.

[0008] 결과적으로, HEVC는 H.264/AVC에 비해 부호화 과정이 복잡해 졌으며, 따라서 빠른 부호화를 위한 구현 방법이 필요하게 되었다.

[0009] 2) HEVC 기반 부호화 블록 결정 방법

[0010] HEVC는 쿼드트리(quard-tree) 구조를 사용하여 계층에 따라 재귀적으로 CU를 더 작은 유닛으로 분할한다. 최적의 CU 크기를 결정하기 위하여 각 CU에서 각각의 PU 크기 및 TU 크기에 대해 가장 작은 비용을 갖는 블록 크기를 조사하게 되며 이때 CU 크기 마다 PU는 다양한 모양의 예측블록으로 분할하여 비용을 조사하고, TU는 재귀적인 트리 구조로 분할을 하여 각각의 크기에 대해 비용을 조사한다. 여기에서 비용이란 화질과 압축율을 동시에 고려하여 QP에 따라 계량화한 값이다. 예를 들어, 상대적으로 비용이 적다는 의미는 동일 화질에서 고압축이 가능하거나, 동일 압축율에서 고화질이 가능하다는 의미가 된다.

[0011] 도 2는 임의 크기의 CU로부터 최적의 PU 크기 및 TU 크기 결정 방법을 나타낸다. 최적의 비용을 갖는 CU를 찾기 위해 순차적으로 가장 큰 크기인 LCU(large coding unit)부터 가장 작은 크기인 SCU(small coding unit)까지 재귀적인 쿼드트리 구조에 따라 비용을 계산하게 된다. 각 CU 크기 별로 매번 PU의 여러 모드에 대한 예측

비용을 전부 계산하며, 각 PU에 대한 비용조사를 할 때, TU의 비용조사 또한 수행한다. TU는 4개의 하위블록으로 재귀적으로 분할될 수 있다.

[0012] HEVC의 부호화 과정에서 CU에 따라 PU, TU의 조합 수가 매우 많으며 최적의 CU, PU, TU를 결정하기 위해 많은 계산 과정이 필요하다는 것을 알 수 있다. 이를 해결하기 위해 도 2에서 아래와 같은 고속화 방법들이 개발되었다.

[0013] ① Early_SKIP

[0014] 하나의 CU에서 첫 번째로 INTER 2Nx2N 모드를 조사한 다음 Coded Block Flag(CBF) 값이 0이고 Motion vector difference(MVD)가 (0,0)값을 갖는 경우 다른 형태의 PU 모드를 조사하지 않고, 바로 CU를 분할하여 다음 깊이의 CU 조사과정을 이어간다. 여기서 CBF는 0이 아닌 잔차 신호의 존재 유무를 알려주는 flag이며, MVD는 주변 블록 또는 이전 프레임의 동일 위치 블록과의 움직임 정보에 대한 차이 값이다.

[0015] ② CBF_Fast

[0016] CBF의 값에 따른 고속 PU 결정 방법(CBF Fast Mode Setting; CFM)으로, INTER_2Nx2N PU 의 CBF가 0인 경우, 남아 있는 다른 형태의 PU 모드들을 검색하지 않는다.

[0017] ③ Early_CU

[0018] CU의 조사과정을 조기 종료하는 방법으로, CU의 PU에 대한 조사 과정 중에 최적 모드가 SKIP 모드로 결정되면 하위 블록으로의 재귀 CU 분할을 하지 않는다.

[0019] 이러한 HEVC의 고속화 방법들이 등장하였지만 아직 실시간 환경에서 사용하기에는 부호화 시간이 많이 소비된다. 따라서 기존 고속화 방법들과 더불어 실시간 환경을 위한 새로운 HEVC 구조 설계와 고속화 방법의 개발이 필수적이다.

[0020] HEVC 부호화기에서 PU 모드를 조기 선택하는 종래 기술은 아래와 같다.

[0021] 3.1) 종래 기술 I

[0022] CU depth level을 고려하여 Upper depth의 upperBestMode와 UpperIntraCost를 이용한 fast mode decision scheme이다. 이 방법은 현재 CU depth의 best PU 모드 탐색을 위해 상위 CU depth의 best PU 모드 정보(upperBestMode)와 상위 CU depth의 INTRA 모드의 RDCost 정보(UpperIntraCost)를 사용한다.

[0023] 도 3은 PU 모드 탐색을 조기에 종료하기 위해 현재 CU depth에서 사용되는 상위 CU depth의 두 가지 정보(BestMode, IntraCost)를 나타낸다.

[0024] 도 4는 fast PU mode decision의 전체 과정을 나타낸다. 이 방법은 총 3가지 조건을 가지며 해당되는 조건을 만족한다면 현재 CU depth의 PU 모드 search를 종료하게 된다.

[0025] ① 조건 1을 만족 ((현재 CU의 depth>0) && (상위 CU depth의 best PU 모드가 SKIP) && (현재 CU의 best PU 모드가 SKIP/MERGE))

[0026] - skip 되는 PU 모드 : INTER NxN, INTER SMP, INTER AMP, INTRA 2Nx2N, INTRA NxN를 수행하지 않는다.

[0027] ② 조건 2, 3을 동시에 만족 ((현재 CU의 depth>0) && (상위 CU depth의 best PU 모드가 INTER) && (현재 CU의 best PU 모드가 INTER AMP) && (현재 CU depth의 best PU 분할 크기가 2Nx2N) && (현재 CU depth의 best RDCost<상위 CU depth의 INTRA RDCost/5))

[0028] - skip되는 PU 모드 : INTRA 2Nx2N, INTRA NxN를 수행하지 않는다.

[0029] 이 방법은 상위 CU depth의 best PU 정보를 이용해야하기 때문에 현재 CU depth가 0일 경우, 이 방법을 수행하

지 않는다.

[0030] 이러한 CU depth의 상관관계를 이용한 PU 모드 검색 skip 구조는 크게 2가지 단점을 갖고 있다.

[0031] (1) 상위 CU depth의 INTRA RDCost를 임의의 값으로 나눈 임계값은 영상에 따라 정확도가 떨어져 고속 부호화 방법으로 적합하지 않다.

[0032] (2) 상위 CU depth의 INTRA RDCost 값을 메모리에 저장하고 있어야 하기 때문에 메모리 사용량의 증가와 부호화 지연을 초래하여 고속화 방법으로 적합하지 않다.

[0033] 3.2) 종래 기술 II

[0034] 현재 CU depth의 best PU 모드 결정을 위해 현재 CU를 중심으로 하여 공간적/시간적으로 존재하는 CU들, 그리고 상위 CU depth들 간의 상관관계를 이용한다.

[0035] 도 5는 현재 CU가 가지는 다양한 상관관계를 나타낸다. 여기서 현재 CU를 중심으로 CU1, CU2, CU3 는 공간적으로 주변에 존재하는 CU들을 나타내며, CU4, CU5는 시간적으로 존재하는 CU들을 나타낸다. 그리고 CU6, CU7 는 현재 CU의 상위 depth를 가지는 CU들이다. 이 중에서 현재 CU의 best PU 모드와 높은 확률로 같은 best PU 모드를 갖는 CU는 공간적으로 주변에 존재하는 CU1, CU2, CU3 이다.

[0036] 도 6은 Random Access configuration으로 부호화 할 때, 프레임간의 시간적 계층 구조와 부호화 순서를 나타낸다. Level1은 프레임 거리로 8이며 Level2, Level3, Level4는 각각 4, 2, 1의 프레임 거리를 갖는다. 여기서 프레임 거리가 작을수록 현재 CU와 시간적으로 존재하는 CU의 best PU 모드가 같아질 확률은 높아진다.

[0037] 이 방법은 먼저 도 5를 참조하여 예측 집합으로 아래의 식 (1)을 정의한다.

[0038] $$\Omega = CU_1, CU_2, CU_3, CU_4, CU_5, CU_6, CU_7 \quad (1)$$

[0039] 그 후, 영상 영역의 특징에 따라 다양한 weight factor를 정의한다. weight factor는 homogeneous 영역의 경우 큰 크기의 PU 블록을, 복잡하거나 rich texture의 경우에는 작은 크기의 PU 블록을 사용해야 한다는 특성을 이용하여 도 7의 표와 같이 정의한다.(INTER NXN의 경우, HM7에서 수행하지 않기 때문에 제외함)

[0040] 도 7의 표에서 γ는 영상 영역의 특징에 따른 PU 모드의 weight factor를 나타낸다. 다음으로 block motion complexity(BMC)를 식 (2)와 같이 정의한다.

[0041] $$BMC = \frac{\sum_{i=1}^{N} W(i,l) \cdot k_i \cdot \gamma_i}{\sum_{i=1}^{N} W(i,l) \cdot k_i} \quad (2)$$

[0042] 여기서 N은 CU들의 개수인 7이며, w(i,l)는 각 모드의 weight factor 함수를, 그리고 γ_i는 CU_i의 모드를 위한 weight factor를 나타낸다. k_i는 CU_i의 존재 유무를 확인해주는 파라미터로서 존재한다면 1을, 존재하지 않는다면 0의 값을 갖는다. w(i,l)는 식 (3),(4)와 같이 정의한다.

[0043] $$W(i,l) = w_i + f(i,l) \quad (3)$$

[0044] $$f(i,l) = \partial\,(l \cdot T_{level}) \cdot T_t \cdot Tk_i \quad (4)$$

[0045] 식 (3)은 adaptive weighting factor 함수이다. 여기서 w_i는 adjacent CU의 weight factor로서 $\sum_{i=1}^{N} w_i = 1$ 이다. w_i의 default value는 도 8의 표, 프레임 거리 차에 따른 weight factor는 도 9의 표와 같다.

[0046] adaptive weighting factor 함수의 l 파라미터는 도 6의 시간적으로 계층적인 Level의 값으로서 GOP가 8일 때, 1,2,3,4 중에 한 가지 값을 가진다. 식 (4)는 공간적, 시간적 또는 depth correlation의 가중치로서 T_{level}는 0.02이며, Tk_i는 control value로서 CU_4, CU_5는 1, CU_6는 -2, 나머지 CU는 0의 값을 가진다. T_t 또한 control

value로서 0.5의 값을 가진다.

[0047] 이렇게 정의된 BMC 값을 이용하여 식 (5)와 같이 임의의 임계값을 결정한다.

[0048]
$$\begin{cases} BMC < Th_1 \text{ --------- } 1 \\ Th_1 \le BMC < Th_2 \text{ --- } 2 \\ Th_2 \le BMC \text{ ------- } 3 \end{cases} \quad (5)$$

[0049] 식 (5)에서 T_1의 값은 1, T_2는 3이다. BMC 값이 조건 1을 만족한다면 slow homogeneous motion으로서 SKIP, INTER 2Nx2N 모드만 수행하고, 조건 2를 만족한다면 medium motion으로서 SKIP, INTER 2Nx2N, INTER SMP 모드를 수행하고, 조건 3을 만족한다면 모든 PU 모드를 수행한다.

[0050] 이러한 임계값을 가지는 PU decision 구조는 크게 2가지 단점을 갖고 있다.

[0051] (1) 영상 영역의 특성에 따라 변경되어야할 파라미터 값이 많고 임계값의 정확도가 떨어지기 때문에 성능 저하를 초래한다.

[0052] (2) 임계값을 결정하기 위해 값이 정의 내려진 파라미터가 존재하고 고려해야할 정보(공간적, 시간적, 상위 CU depth 정보들)가 많기 때문에 복잡도가 증가하여 고속 부호화 성능이 떨어진다.

[0053] 3.3) 종래 기술 III

[0054] 참조 픽처 리스트 L0와 L1의 Temporally Co-located CU와 그 주변에 있는 CU의 PU 모드 정보를 이용하여 현재 CU의 INTER PU 모드 탐색을 조기에 종료하는 구조이다. 이 방법에서는 현재 CU와 Temporally Co-located CU의 크기를 이용하여 총 세 가지 조건을 정의하고 각각의 발생 확률의 결과값을 바탕으로 조기 종료 조건을 제시하였다.

[0055] ①번 조건 : Lco < Lcu 일 경우(현재 CU의 크기가 Co-located CU의 크기 보다 클 경우)

[0056] 하단에서 정의한 식 (6)의 두 가지 조건인 Con1과 Con2을 이용하여 6가지 표준 영상에 대해 특정한 발생 확률을 계산하였다.

[0057] Con1 = {Lco < Lcu}, Con2 = {Lco < Lcu} ∩ {Pcu = PART_2Nx2N} (6)

[0058] 여기서 Lco는 Temporally Co-located CU의 CU 크기를, Lcu는 현재 CU 크기를 나타내며 Pcu는 현재 CU의 best PU 모드가 PART_2Nx2N일 경우를 나타낸다.

[0059] 각각의 발생 확률에 대한 결과는 도 10과 같으며 여기서 Pi,c는 전체 CU의 개수에서 Con1을 만족할 확률을, Pi,e는 전체 CU의 개수에서 Con2를 만족할 확률을, Pi는 Con2를 만족하면서 best PU 모드의 분할 크기가 PART_2Nx2N일 확률을 나타낸다. 결론적으로 Pi의 확률이 높을수록 Con1의 조건에 대한 정확도가 높다는 것을 나타낸다. 이 방법에서는 Con1의 정확도가 높다는 것을 들어 ①번 조건을 만족할 때는 PART_2Nx2N만을 수행하도록 하였다.

[0060] 따라서, 도 10에 제시된 표를 바탕으로, Con1을 만족한다면 INTER 모드의 분할 크기가 PART_2Nx2N일 경우만 탐색하고 나머지 INTER PU 모드는 스킵한다.

[0061] ②번 조건 : Lco == Lcu 일 경우(현재 CU의 크기가 Temporally Co-located CU의 크기와 같을 경우)

[0062] 하단에서 정의한 식 (7)의 두 가지 조건인 Suri와 C1을 이용하여 6가지 시퀀스에 대해 특정한 발생 확률을 계산하였다.

[0063]
$$Sur_i = \bigcap_{j=0}^{3}((L_j = L_{co}) \cap (P_j = PART2N \times 2N)) \cup (L_j > L_{co}), \quad C1 = (L_{cu} < L_{co}) \cap Sur_i \quad (7)$$

[0064] 여기서 Lj는 Temporally Co-located CU의 주변 CU들의 CU 크기를, Pj는 Temporally Co-located CU를 둘러싼 주변 CU들의 분할 크기를 나타낸다.

[0065] 도 11의 표에서 Ri,s는 전체 Co-located CU의 개수에서 Suri를 만족할 확률을, Ri,d는 Suri를 만족하는 CU의 개수에서 C1을 만족할 확률을, Ri,e는 C1의 전체 개수에서 현재 CU가 C1을 만족하면서 best PU 모드가

PART_2Nx2N일 확률을 나타낸다. Ri는 C1을 만족하면서 현재 CU가 Co-located CU보다 작고 PU 분할 크기가 PART_2Nx2N이 아닐 확률을 나타낸다. 결론적으로 현재 CU의 크기가 64x64일 경우에 Ri,e의 확률이 높다는 것을 확인할 수 있고 이를 이용하여, 도 12와 같이 조건을 정리하였다.

[0066] ③번 조건 : Lco > Lcu 일 경우(현재 CU의 크기가 Co-located CU의 크기 보다 작을 경우)

[0067] ③번 조건을 만족한다면 현재 CU의 INTER PU 모드를 예측하기 어렵다. 따라서 ③번 조건의 경우, Temporally Co-located CU의 best PU 모드에 따라 현재 CU의 best PU 모드는 PART_2Nx2N, PART_2NxN, PART_Nx2N 중에서 한 가지를 선택하여 수행하도록 하였다. 또한 ③번 조건에서는 임의의 수식의 발생 확률을 계산해 INTER PU 모드 탐색의 조기 종료 조건을 설계한 것이 아니라 Temporally Co-located CU의 best PU 모드를 바탕으로 현재 CU의 PU 모드를 결정하였으며 이를 도 13과 같이 정의하였다.

[0068] 이러한 INTER PU 모드 탐색을 조기 종료하는 구조는 크게 2가지 단점을 갖고 있다.

[0069] (1) INTER PU 모드 탐색을 조기 종료하는 조건들을 결정하기 위해서 사용된 시퀀스들이 한정되어 있기 때문에 조건식에 대한 정확도가 떨어진다. 따라서 고속 부호화에 적합하지 않다.

[0070] (2) INTER PU 모드 탐색을 조기 종료하는 조건에서 현재 CU 크기가 Temporally Co-located CU 크기 보다 작을 경우, INTER AMP를 고려하지 않기 때문에 부호화 성능 저하가 발생하여 UHD급 영상과 같이 고품질/고용량 해상도를 갖는 영상에는 적합하지 않는다.

[0071] HEVC는 H.264/AVC에 비해 부호화 과정이 복잡해 졌으며, 따라서 고속 부호화를 위한 새로운 방법이 광범위하게 연구&개발 되어질 것으로 예상된다. SW 기반 부호화기의 고속화 기법 중에 특정한 PU 모드 비용 조사를 스킵하는 구조 또한 유망한 방법 중의 하나이다.

[0072] 효과적인 PU 모드 비용 조사 스킵을 위해서는 현재 CU의 best PU 모드 탐색의 조기 종료 조건에 이용되는 정보가 중요하므로, HEVC 부호화기에서 현재 CU의 INTER PU 모드 탐색을 조기에 종료할 수 있도록 하기 위한 방안의 모색이 요청된다.

발명의 내용

해결하려는 과제

[0073] 본 발명은 상기와 같은 문제점을 해결하기 위하여 안출된 것으로서, 본 발명의 목적은, HEVC 부호화기에서 Temporally Co-located CU의 두 가지 best PU 정보를 이용하여 현재 CU의 INTER PU 모드 탐색을 조기에 종료하는 방법을 제공함에 있다.

과제의 해결 수단

[0074] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른, CU의 PU 모드 결정 방법은, 현재 CU의 크기 및 상기 현재 CU에 대한 참조 픽처들에서 Co-located CU들의 PU 분할 크기들과 PU 모드들을 파악하는 단계; 및 파악된 PU 분할 크기들과 PU 모드들을 기초로, 남아 있는 PU 모드의 탐색을 스킵하는 단계;를 포함한다.

[0075] 그리고, 상기 스킵 단계는, INTER Nx2N, INTER 2NxN, INTER AMP(Asymmetric Motion Partitions)의 탐색을 스킵할 수 있다.

[0076] 또한, 상기 스킵 단계는, 상기 Co-located CU들의 CU 크기가 64x64이고, 상기 Co-located CU들의 PU 분할 크기가 PART_2Nx2N 이면, 상기 탐색을 스킵할 수 있다.

[0077] 그리고, 상기 스킵 단계는, 상기 현재 CU의 크기가 64x64이고, 상기 Co-located CU들의 PU 분할 크기가 2Nx2N이면, 상기 탐색을 스킵할 수 있다.

[0078] 또한, 상기 스킵 단계는, 상기 현재 CU의 크기가 32x32이고, 상기 Co-located CU들의 PU 모드가 MERGE 또는 SKIP이면, 상기 탐색을 스킵하고, 상기 현재 CU의 크기가 16x16 또는 8x8이고, 상기 Co-located CU들의 PU 모

드가 SKIP이면, 상기 탐색을 스킵할 수 있다.

[0079] 한편, 본 발명의 다른 실시예에 따른, 컴퓨터로 읽을 수 있는 기록매체에는, 현재 CU의 크기 및 상기 현재 CU에 대한 참조 픽처들에서 Co-located CU들의 PU 분할 크기들과 PU 모드들을 파악하는 단계; 및 파악된 PU 분할 크기들과 PU 모드들을 기초로, 남아 있는 PU 모드의 탐색을 스킵하는 단계;를 포함하는 것을 특징으로 하는 CU의 PU 모드 결정 방법을 수행할 수 있는 프로그램이 기록된다.

발명의 효과

[0080] 이상 설명한 바와 같이, 본 발명의 실시예들에 따르면, HEVC 부호화기에서 Temporally Co-located CU의 두 가지 best PU 정보를 이용하여 현재 CU의 INTER PU 모드 탐색을 조기에 종료할 수 있어, 부호화 속도는 증가시키고, 필요한 연산량은 감소시킬 수 있게 된다.

도면의 간단한 설명

[0081] 도 1은 HEVC 부호화기 블록도,

도 2는 재귀적 CU, PU, TU 결정 과정을 나타낸 도면,

도 3은 current CU depth에서 upper CU depth의 부호화 정보를 사용하는 방법을 나타낸 도면,

도 4는 고속 PU 모드 결정 과정을 나타낸 도면,

도 5는 현재 CU가 가지는 다양한 상관관계를 나타낸 도면,

도 6은 Random Access configuration에서의 프레임 계층 구조 및 부호화 순서를 나타낸 도면,

도 7은 PU 모드 별 weight factor의 정의를 나타낸 표,

도 8은 CUs의 weight factor(default setting)를 나타낸 표,

도 9는 CUs의 weight factor(프레임 거리가 1과 2일 경우)를 나타낸 표,

도 10은 종래기술 III의 ①번 조건에 대한 발생 확률을 나타낸 표,

도 11은 종래기술 III의 ②번 조건에 대한 발생 확률을 나타낸 표,

도 12는 종래기술 III의 ②번 조건에 대한 INTER PU 모드 스킵 방법,

도 13은 종래기술 III의 ③번 조건에 대한 INTER PU 모드 스킵 방법,

도 14는 본 발명의 일 실시예에 따라 현재 CU의 PU 모드 탐색을 조기에 종료하기 위해 참고되는 Temporally Co-located CU 정보를 나타낸 도면,

도 15는 본 발명의 일 실시예에 따른 INTER PU 모드 탐색의 조기 종료 방법을 나타낸 도면, 그리고,

도 16은 본 발명의 일 실시예에 따른 INTER PU 모드 탐색의 조기 종료 방법의 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0082] 이하에서는 도면을 참조하여 본 발명을 보다 상세하게 설명한다.

[0083] 본 발명의 실시예에서 제안하는 HEVC 부호화기의 속도 향상을 위한 PU 모드 예측을 조기에 종료하는 방법은 기본적으로 CU 레벨에서부터 시작한다.

[0084] 도 14는 현재 부호화되려는 CU depth의 best PU 모드 탐색을 조기에 종료하기 위해 Temporally Co-located CU들의 best PU 모드 정보를 이용하는 방법을 나타낸 도면이다.

[0085] Temporally Co-located CU들은 현재 부호화되는 CU와 상호 상관도가 높으며 best PU 모드가 동일할 확률이 높다. 따라서, 현재 부호화되는 CU의 PU 모드 탐색을 조기에 종료하기 위해서는 현재 부호화되는 CU의 Co-located CU들(Picture List인 L0, L1의 CUs)의 best PU 모드 정보들을 알아야 한다.

[0086] 이 때, 이용되는 PU 모드 정보들은 best PU 분할 크기, best PU 모드 정보이며 Picture는 참조 픽처 리스트인 L0, L1에서 현재 픽처와 가장 가까운 Picture를 각각 이용한다.

[0087] 도 15는 INTER PU 모드 탐색의 조기 종료 방법을 위한 주요 알고리듬을 나타낸다.

[0088] 도 15에 나타난 바와 같이, INTER PU 모드 탐색의 조기 종료 방법은 MERGE/SKIP 그리고 INTER 2Nx2N을 수행한 후, 제안한 조기 종료 조건을 비교하게 된다.

[0089] 여기서 MERGE/SKIP은, ① Early_SKIP, ② CBF_Fast, ③ Early_CU의 적용을 위해, 그리고 INTER 2Nx2N는 QP 값이 증가함에 따라 best PU 모드의 분할 크기가 PART_2Nx2N이 선택될 확률이 증가하기 때문에 수행해야 한다.

[0090] 그 다음, 현재 부호화되는 CU의 크기에 따라 다른 조기 종료 조건을 가지게 되며, 조건 생성에 필요한 Temporally Co-located CU의 best PU 정보를 획득하기 위해 참조 픽처 리스트인 L0와 L1에서 현재 Picture와 가장 가까운 Picture들을 이용하게 된다.

[0091] 따라서 L0와 L1이 존재할 경우에만 INTER PU 모드 탐색의 조기 종료 방법을 적용하며 해당 프레임이 I 프레임일 경우에는 INTRA PU 모드(INTRA 2Nx2N, INTRA NxN)만 수행되기 때문에 제안한 방법을 적용하지 않는다.

[0092] 본 발명의 일 실시예에서는, 현재 CU의 크기가 64x64일 경우, L0 그리고 L1의 Co-located CU들의 best PU 분할 크기가 2Nx2N을 만족한다면 남아있는 INTER Nx2N, INTER 2NxN, INTER AMP(Asymmetric Motion Partitions)의 탐색을 스킵하게 되고 만족하지 않는다면 모두 수행하게 된다.

[0093] 현재 CU의 크기가 32x32일 경우에는, L0 그리고 L1의 Co-located CU들의 best PU 모드가 MERGE 또는 SKIP을 만족한다면 INTER Nx2N, INTER 2NxN, INTER AMP의 탐색을 스킵하고 만족하지 않는다면 모두 수행하게 된다.

[0094] 나머지 CU의 크기(16x16 또는 8x8)에 대해서는 L0 그리고 L1의 Co-located CU들의 best PU 모드가 SKIP일 경우에만 INTER Nx2N, INTER 2NxN, INTER AMP의 탐색을 스킵하고 만족하지 않는다면 모두 수행하게 된다.

[0095] INTER PU 모드 탐색의 조기 종료 방법은 현재 CU의 크기가 64x64일 경우에는 Temporally Co-located CU의 best PU 분할 크기를, 현재 CU가 나머지 다른 크기의 경우에는 best PU 모드 정보를 이용하도록 구현하였다.

[0096] 도 16에는 HEVC 부호화기의 속도 향상을 위한 현재 부호화되는 CU의 PU 모드 탐색의 조기 종료 방법에 대한 전체 순서도를 도시하였다. 도 16의 방법은 도 15에서 제시한 L0, L1의 두 가지 best PU 정보(best PU 분할 크기, best PU 모드)를 이용한다는 것을 확인할 수 있다.

[0097] 본 발명의 실시예에서는, 고속 부호화 방법인 Early_SKIP, CBF_Fast, Early_CU가 적용된다. 현재 부호화되는 CU의 best PU 모드 탐색을 조기에 종료하기 위해 MERGE/SKIP을 수행하고 Early SKIP, CBF Fast 조건을 비교하게 된다.

[0098] 만약, Early_SKIP, CBF_Fast 조건이 참이라면 해당 CU의 best PU 모드는 MERGE/SKIP(분할 크기는 2Nx2N)이 되고 다음 깊이의 CU 분할을 위해 Early_CU 조건을 비교하게 된다. 만약 조건이 거짓이라면 INTER 2Nx2N 모드를 수행한다.

[0099] INTER 2Nx2N을 수행한 다음, L0와 L1의 Temporally Co-located CU의 크기가 64x64이고 best PU 분할 크기가 2Nx2N의 조건을 만족한다면 남아있는 INTER SMP(Symmetric Motion partitions) 그리고 INTER AMP를 수행하지 않으며, 만족하지 않는다면 도 15에서 제시한 조기 종료 조건을 비교하게 된다.

[0100] 마지막으로, INTRA 2Nx2N 또는 INTRA NxN의 PU 모드 탐색이 완료된다면 탐색된 모든 PU 모드들에 대해서만 RDO(Rate Distortion Optimization)를 수행하여 현재 부호화되는 CU에 대한 best PU 모드를 결정한다.

[0101] 본 발명의 실시예에 따른 HEVC 부호화기의 속도 향상을 위한 INTER PU 모드 탐색 조기 종료 방법은 Temporally Co-located CU의 best PU 정보들을 이용하여 INTER PU 모드 탐색 과정을 조기에 종료시킴으로써 부호화 속도를 향상시킬 수 있다.

[0102] 이는 종래 기술들이 현재 CU를 중심으로 Temporally Co-located CU의 best PU 모드 정보와 Spatially neighboring CU의 best PU 모드 정보, CU depth간 상관관계를 이용한 임계값을 활용한 반면, Spatially neighboring CU의 best PU 모드 정보와 CU 깊이간의 RDCost 비교를 배제하여 메모리 사용량과 데이터를 불러오는 시간을 줄일 수 있다.

[0103] 또한, 이상에서는 본 발명의 바람직한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시

예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

부호의 설명

[0104] current CU : 현재 CU

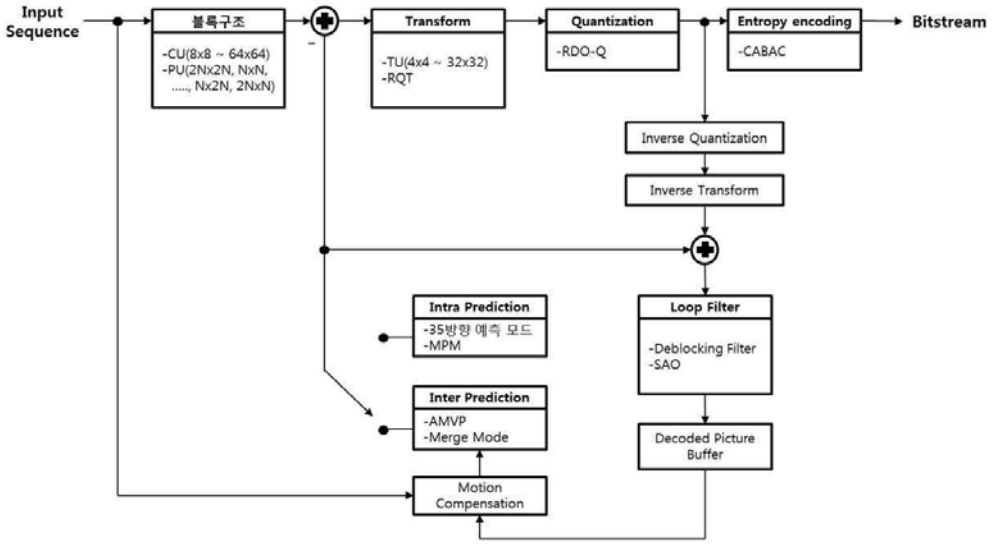
Co-located CU : 다른 픽처에서 현재 CU와 동일한 위치의 CU

best PU Partition size : 최적 PU 분할 크기

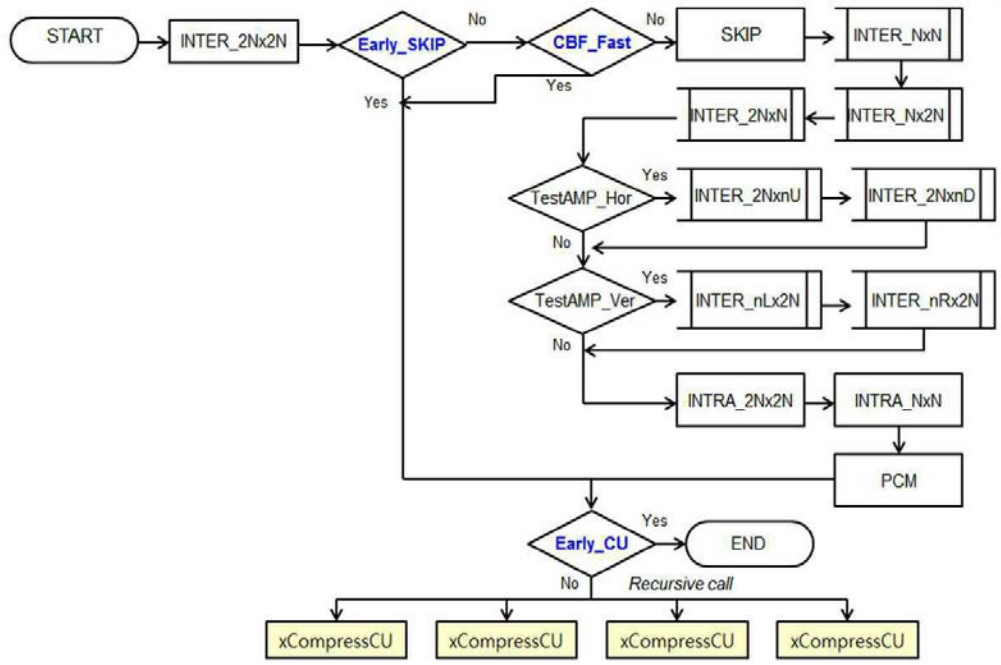
best PU mode : 최적 PU 모드

도면

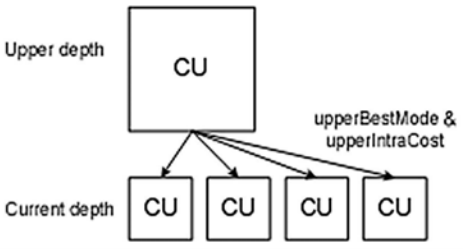
도면1



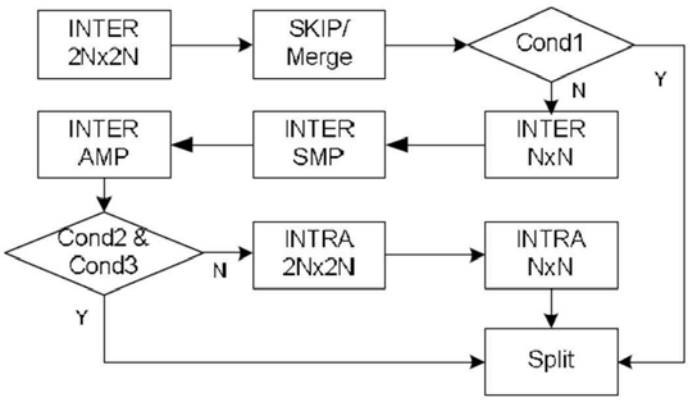
도면2



도면3

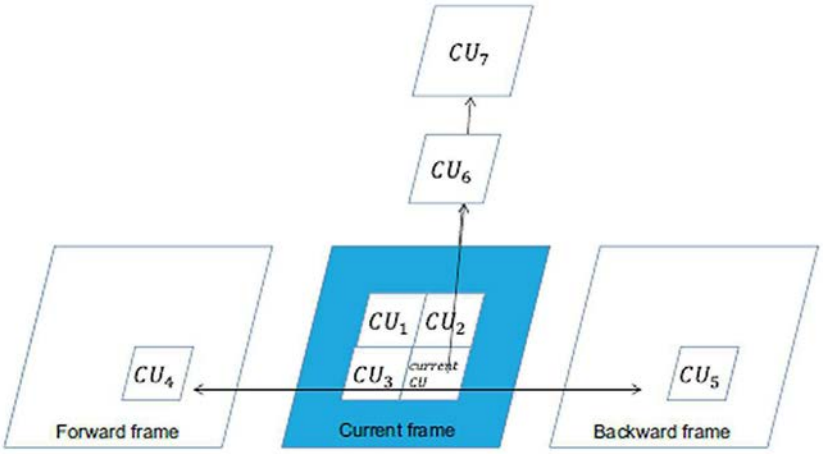


도면4

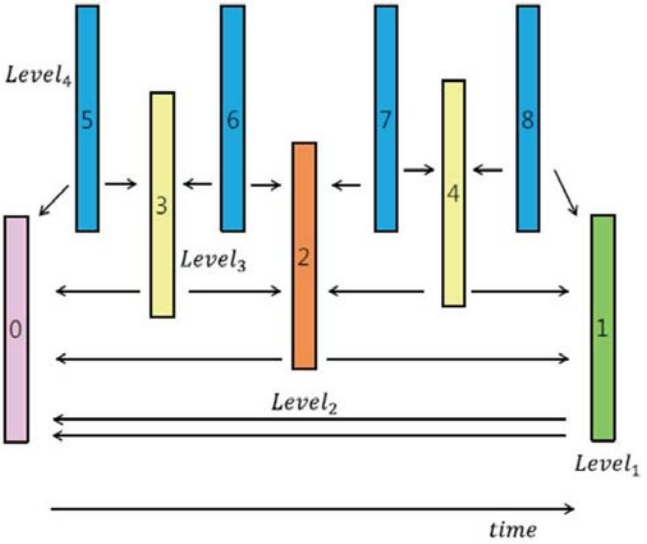


Cond1 : uiDepth>0 && upperBestMode==SKIP && BestCost is changed?
Cond2 : uiDepth>0 && upperBestMode==Inter && BestCost is changed?
Cond3 : eParantPartSize==2Nx2N && BestCost< upperIntraCost/5

도면5



도면6



도면7

Class	Mode	Motion activity	Mode Weight factor (γ_i)
1	SKIP	Homogeneous region with motionless	0
2	Inter 2N×2N	Slow motion	1
3	Inter 2N×N, N×2N	Moderate motion	2
4	Inter 2N×nU, 2N×nD nL×2N, nR×2N	Fast or complex motion	3
5	Intra 2N×2N, N×N	Highly-textured region with scene changes or in fast motion	4

도면8

Index(i) in classified CUs	1	2	3	4	5	6	7
w_i	0.2	0.2	0.2	0.05	0.05	0.2	0.1

도면9

Index(i) in classified CUs	1	2	3	4	5	6	7
w_i	0.2	0.2	0.2	0.1	0.1	0.1	0.1

도면10

sequences	QP	64×64			32×32			16×16		
		$P_{0,e}$ (%)	$P_{0,e}$ (%)	P_0 (%)	$P_{1,e}$ (%)	$P_{1,e}$ (%)	P_1 (%)	$P_{2,e}$ (%)	$P_{2,e}$ (%)	P_2 (%)
Four People 1,280x720_60	22	6.56	5.36	81.7	4.51	3.28	72.8	3.80	2.58	67.9
	27	7.64	6.43	84.2	5.20	3.88	74.6	4.78	3.34	69.8
	32	8.18	6.95	85.0	6.10	4.67	76.5	5.53	3.91	70.7
	37	9.32	8.06	86.5	7.93	6.10	76.9	6.06	4.38	72.3
Johnny 1,280x720_60	22	11.39	9.10	79.9	8.13	6.02	74.1	5.65	3.88	68.6
	27	12.30	10.28	83.6	8.64	6.58	76.1	6.55	4.64	70.8
	32	12.82	10.78	84.1	9.75	7.55	77.4	7.45	5.33	71.6
	37	13.30	11.37	85.5	10.42	8.18	78.5	8.40	6.13	73.0
Kristen And Sara 1280x720_60	22	14.77	12.50	84.6	13.20	10.38	78.6	7.36	5.42	73.7
	27	15.65	13.93	89.5	13.46	10.89	80.9	7.90	5.87	74.3
	32	15.83	14.44	91.2	14.37	11.74	81.7	8.60	6.57	76.4
	37	16.73	15.53	92.8	14.88	12.42	83.5	9.45	7.38	78.1
Cactus 1920 × 1,080_50	22	9.71	8.32	85.7	8.67	6.83	78.8	4.53	3.33	73.5
	27	10.10	8.87	87.8	9.84	7.83	79.6	5.32	4.01	75.4
	32	11.08	9.75	88.0	9.96	8.10	81.3	5.70	4.38	76.8
	37	12.62	11.43	90.4	10.21	8.42	82.5	6.81	5.33	78.3
Kimono1 1920 × 1,080_24	22	21.02	19.74	93.9	14.93	12.48	83.6	12.72	10.02	78.8
	27	21.86	20.77	95.0	15.38	13.13	85.4	13.65	11.04	80.9
	32	22.00	21.21	96.4	16.64	14.34	86.2	13.74	11.14	81.1
	37	23.48	22.75	96.9	17.88	15.86	88.7	14.81	11.32	83.2
ParkScene 1920 × 1080_24	22	18.51	17.20	92.9	13.53	11.70	86.5	13.14	10.51	80.0
	27	19.13	18.42	96.3	14.43	12.73	88.2	13.83	11.23	81.2
	32	21.85	21.26	97.3	15.59	13.97	89.6	14.63	12.06	82.5
	37	22.50	21.98	97.7	16.28	14.70	90.3	15.56	13.12	84.3

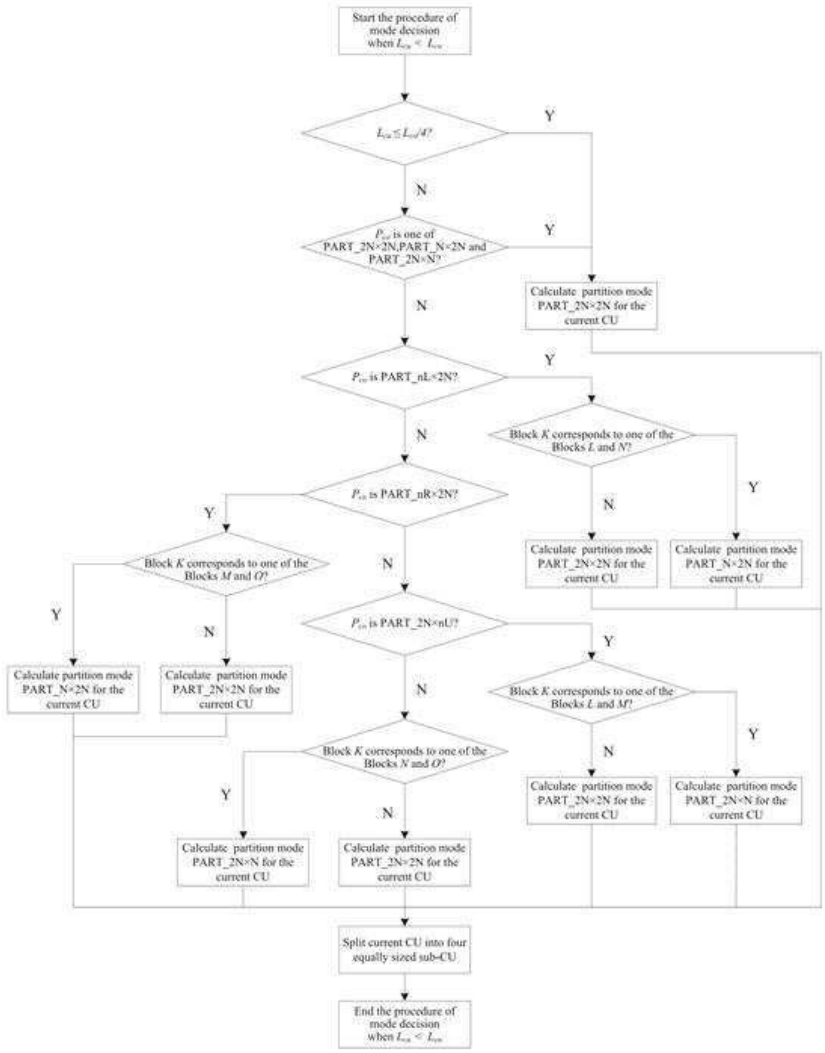
도면11

sequences	QP	64×64				32×32				16×16			
		$R_{0,a}$	$R_{0,d}$	$R_{0,e}$	R_0	$R_{1,a}$	$R_{1,d}$	$R_{1,e}$	R_1	$R_{2,a}$	$R_{2,d}$	$R_{2,e}$	R_2
Basketball-Drill 832×480	22	29.91	12.48	89.90	1.26	30.85	21.78	67.27	7.13	29.54	30.30	51.19	14.79
	27	31.17	11.96	90.38	1.15	32.09	20.74	67.47	6.80	30.92	29.80	53.66	13.80
	32	32.82	11.50	91.48	0.98	33.43	18.95	68.13	6.04	32.57	28.73	59.94	11.94
	37	34.08	10.18	92.34	0.78	35.14	17.07	66.26	5.76	34.18	27.89	60.31	11.07
Partyscreen 832×480	22	29.35	11.89	91.34	1.03	28.32	21.57	67.92	6.92	29.35	30.32	51.98	14.56
	27	30.71	10.61	91.41	0.91	29.83	19.85	69.60	6.03	30.82	28.23	52.45	13.42
	32	31.46	10.04	91.83	0.82	30.64	18.69	70.68	5.48	32.78	27.50	56.18	12.05
	37	33.20	9.44	92.91	0.67	32.89	16.96	70.22	5.05	34.07	26.81	59.31	10.91
Fourpeople 1,280×720	22	33.42	9.87	92.30	0.76	39.29	18.69	68.65	5.86	50.73	27.57	55.80	12.46
	27	34.90	8.01	93.10	0.55	40.80	16.46	68.09	5.26	52.26	26.55	57.10	11.39
	32	36.63	7.50	93.47	0.49	42.28	14.38	65.44	4.97	53.76	24.93	59.65	10.06
	37	37.45	7.02	94.30	0.40	43.77	13.58	66.47	4.54	55.19	23.60	61.82	9.01
Johnny 1,280×720	22	34.09	9.15	91.26	0.80	40.60	16.92	69.74	5.12	51.80	27.70	52.78	13.08
	27	35.86	8.59	92.72	0.63	42.05	15.17	69.92	4.70	52.91	25.49	54.88	11.50
	32	37.84	8.01	93.51	0.52	43.52	15.12	70.97	4.39	55.78	24.88	58.08	10.43
	37	39.67	7.75	93.94	0.47	45.26	14.71	72.60	4.03	57.27	22.29	59.40	9.05
Kimono1 1920×1080	22	47.18	6.00	94.33	0.34	51.61	13.45	70.33	3.99	66.66	25.23	62.5	9.46
	27	49.27	5.56	95.86	0.23	53.98	12.78	71.32	3.67	68.85	23.60	66.40	7.93
	32	50.54	5.18	96.53	0.18	55.54	11.84	71.45	3.38	70.23	21.18	69.93	6.37
	37	52.78	4.65	97.63	0.11	57.62	11.37	72.30	3.15	71.92	20.51	70.36	6.08
BQTerrace 1920×1080	22	46.19	6.80	94.26	0.39	49.45	13.88	71.33	3.98	65.25	23.79	58.65	9.86
	27	47.06	6.21	96.04	0.25	51.78	13.46	72.79	3.66	67.34	22.19	61.29	8.59
	32	48.86	5.79	96.55	0.20	52.94	13.00	73.54	3.44	68.85	20.55	63.36	7.53
	37	50.28	5.46	96.89	0.17	54.34	12.46	74.24	3.21	70.95	19.32	65.22	6.72

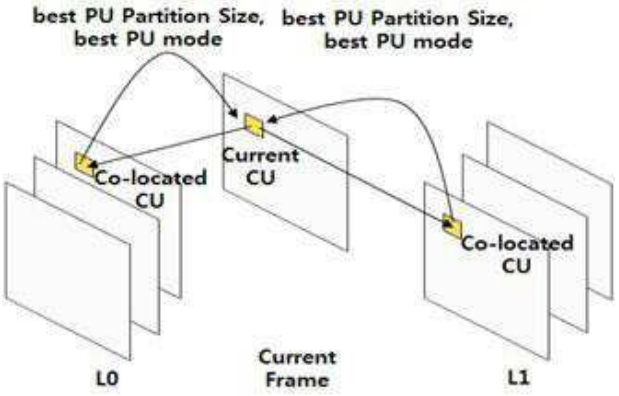
도면12

If Size of CU == Size of CoCU
Then Bestmode of CU= Bestmode of CoCU
If Size of CU == 64×64
Then If (Sizes of surCUs ==64×64) and (Bestmode of surCUs == '2 N×2 N')
Then Bestmodes of SubCUs (32×32, 16×16, 8×8)= '2 N×2 N'
Else split CU
EndIf
Else split CU
EndIf
EndIf

도면13


Start the procedure of mode decision when $L_{cu} < L_{cu}$
$L_{cu} \le L_{cu}/4$?
Y
N
P_{cu} is one of PART_2N×2N,PART_N×2N and PART_2N×N?
Y
Calculate partition mode PART_2N×2N for the current CU
N
P_{cu} is PART_nL×2N?
Y
Block K corresponds to one of the Blocks L and N?
N
Y
N
P_{cu} is PART_nR×2N?
Y
Block K corresponds to one of the Blocks M and O?
N
Calculate partition mode PART_2N×2N for the current CU
Calculate partition mode PART_N×2N for the current CU
Y
N
P_{cu} is PART_2N×nU?
Y
Calculate partition mode PART_N×2N for the current CU
Calculate partition mode PART_2N×2N for the current CU
N
Block K corresponds to one of the Blocks L and M?
Y
Block K corresponds to one of the Blocks N and O?
N
Calculate partition mode PART_2N×2N for the current CU
Calculate partition mode PART_2N×N for the current CU
Y
N
Calculate partition mode PART_2N×N for the current CU
Calculate partition mode PART_2N×2N for the current CU
Split current CU into four equally sized sub-CU
End the procedure of mode decision when $L_{cu} < L_{cu}$

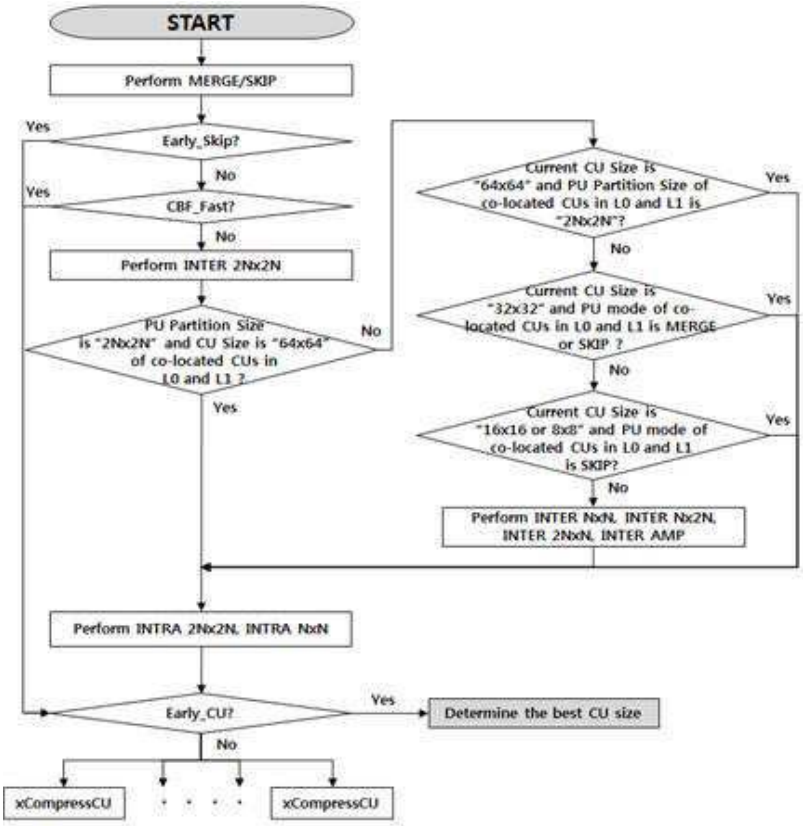
도면14


best PU Partition Size, best PU mode
best PU Partition Size, best PU mode
Co-located CU
Current CU
Co-located CU
L0
Current Frame
L1

도면15

```
Perform MERGE/SKIP and INTER 2N×2N
If (Size of CU == 64×64)
   Then If (Co-located CU's best PU Partition Size of L0 and L1 is '2N×2N')
         Then Skip the Remaining Symmetric and Asymmetric partition size
      Else Perform the every partition size
   EndIf
Else If (Size of CU == 32×32)
   Then If ((Co-located CU's best PU mode of L0 and L1 is MERGE or SKIP))
         Then Skip the Remaining Symmetric and Asymmetric PU partition size
      Else Perform the every partition size
   EndIf
Else
   Then If (Co-located CU's best PU mode of L0 and L1 is SKIP)
         Then Skip the Remaining Symmetric and Asymmetric PU partition size
      Else Perform the every partition size
   EndIf
EndIf
```

도면16



■ 기술명 : 고성능 영상 스무딩 기술
(High performance Image Smoothing Technology)

산업기술분류	디지털 콘텐츠 / 실감형 영상 콘텐츠 / 3D・UHD
Key-word(국문)	고성능 영상 스무딩 기술
Key-word(영문)	High performance Image Smoothing

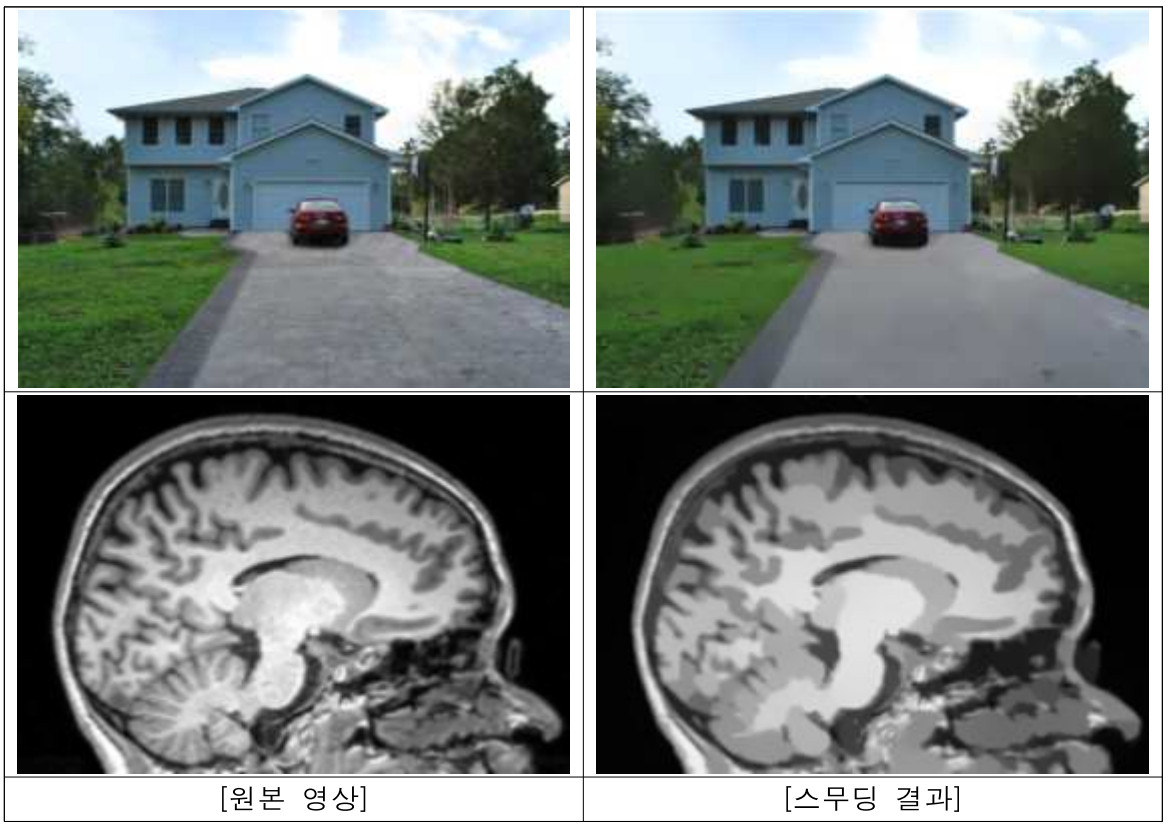
■ 기술의 개요

○ (배경) 영상에 포함된 중요도가 낮은 정보는 영상을 이용한 다양한 응용에 역효과를 발생시킬 수 있으므로 영상의 주요 특성 부분은 부각시키고 중요도가 낮은 부분은 제거하는 이미지 스무딩 기술 필요

○ (개요) RGB영상의 gradient, Laplacian, diagonal 미분 및 분리 최적화를 통해 영상의 주요성분을 살리고 미미한 성분을 제거하는 스무딩 영상 획득 기술

- 영상을 영역별로 분리 처리하여 부자연스러운 데이터의 보간, 노이즈 제거 가능

< 기술 개요도 >



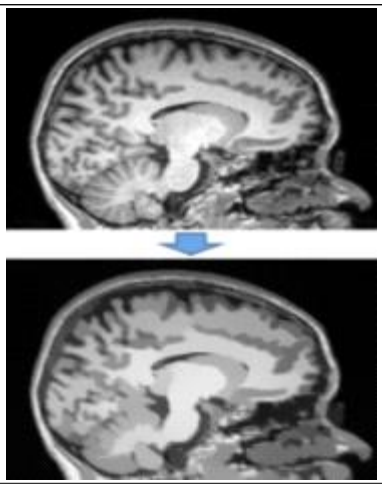
[원본 영상] [스무딩 결과]

■ 기술의 구현수준(TRL)

■ 기술의 장점(경쟁기술과의 차별성)

○ 기존 기술 대비 최대신호 대 잡음비(PSNR)가 약 1.5dB 높음
○ 영상의 주요 특성 추출 및 노이즈 제거 기능을 통해 영상 전처리 모듈로 사용 가능
○ 영상 편집 SW의 효과 모듈로 사용 가능
○ 3D depth를 생성하거나 후처리 하는 엔진으로 사용 가능
○ 의료 영상 분석에 적용 가능

■ 활용범위 및 응용분야

	
[콘텐츠 저작 툴]	[영상 전처리 엔진(분리/인식), 의료영상 분석]

■ 지식재산권 현황

구분	발명의 명칭	출원번호 (출원일)	등록번호 (등록일)
특허	뎁스맵 등심선 생성 방법 및 시스템	2014-0153108 (2014.11.05)	10-1709974 (2017.02.20)
특허	영상 스무딩 방법 및 장치	2014-0165076 (2014.11.25)	-

(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2017년02월27일
(11) 등록번호 10-1709974
(24) 등록일자 2017년02월20일

(51) 국제특허분류(Int. Cl.)
H04N 13/00 (2016.01) *H04N 13/02* (2006.01)
(21) 출원번호 10-2014-0153108
(22) 출원일자 2014년11월05일
심사청구일자 2015년03월10일
(65) 공개번호 10-2016-0054157
(43) 공개일자 2016년05월16일
(56) 선행기술조사문헌
KR1020110135786 A*
WO2013081304 A1*
KR1020090080556 A*
KR101089344 B1
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
조충상
경기도 성남시 분당구 장미로 55 116동 805호 (야탑동,장미마을코오롱아파트)
고민수
경기도 양주시 광적면 화합로 75-5 (효촌리)
(뒷면에 계속)
(74) 대리인
남충우

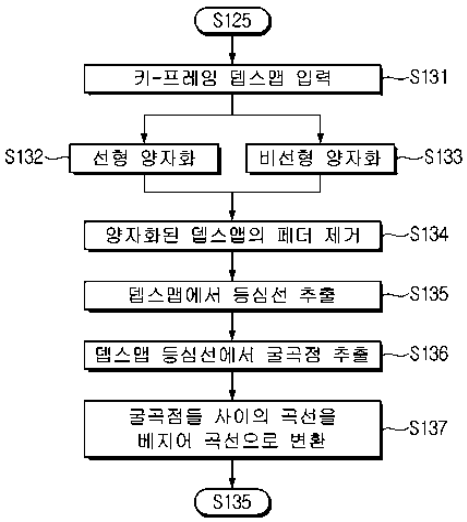
전체 청구항 수 : 총 3 항 심사관 : 정성훈

(54) 발명의 명칭 **뎁스맵 등심선 생성 방법 및 시스템**

(57) 요 약

뎁스맵 등심선 생성 방법 및 시스템이 제공된다. 본 발명의 실시예에 따른 뎁스맵 등심선 생성 방법은, 뎁스맵의 뎁스값들을 다수의 양자화 단계들로 양자화 하고, 양자화된 뎁스맵에서 등심선을 생성한다. 이에 의해, 2D 영상 콘텐츠를 3D 영상 콘텐츠로 변환하는 과정에서, 뎁스맵 등심선을 자동으로 생성할 수 있게 되어, 수작업을 통해 그래픽툴에서 뎁스맵 경계의 곡선을 그리는 작업이 필요 없게 된다. 이에 의해, 필요한 인력의 감소는 물론, 작업 속도를 크게 향상시킬 수 있게 된다.

대 표 도 - 도3



(72) 발명자

신화선

경기도 용인시 기흥구 보정로 26 101동 1601호 (보정동,상록데시앙아파트)

강주형

대전광역시 동구 송촌남로19번길 15 204호 (용전동)

이 발명을 지원한 국가연구개발사업

과제고유번호 10048070

부처명 산업통상자원부

연구관리전문기관 한국산업기술평가관리원

연구사업명 (산업부)기술료지원사업

연구과제명 양자화 기반 2D 영상콘텐츠의 3D변환 기술과 이를 활용한 입체영상 저작도구 개발 및 시험적용

기 여 율 1/1

주관기관 전자부품연구원

연구기간 2013.12.01 ~ 2014.11.30

명 세 서

청구범위

청구항 1

뎁스맵의 뎁스값들을 다수의 양자화 단계들로 양자화 하는 단계;

양자화된 뎁스맵에서 영역 크기가 임계 크기 보다 작은 영역을 페더 영역으로 설정하는 단계;

상기 페더 영역을 뎁스값 차이가 가장 작은 주변 영역의 뎁스값으로 변환하는 단계;

상기 변환단계에 의해 변환된 뎁스맵에서 등심선을 생성하는 단계;

상기 등심선의 굴곡점들을 추출하는 단계; 및

상기 굴곡점들 사이의 곡선을 다른 타입의 곡선으로 변환하는 단계;를 포함하는 것을 특징으로 하는 뎁스맵 등심선 생성 방법.

청구항 2

청구항 1에 있어서,

상기 양자화 단계는,

뎁스맵의 뎁스값들을 선형 양자화 또는 비선형 양자화하는 것을 특징으로 하는 뎁스맵 등심선 생성 방법.

청구항 3

삭제

청구항 4

삭제

청구항 5

삭제

청구항 6

뎁스맵의 뎁스값들을 다수의 양자화 단계들로 양자화 하는 단계; 및

양자화된 뎁스맵에서 영역 크기가 임계 크기 보다 작은 영역을 페더 영역으로 설정하는 단계;

상기 페더 영역을 뎁스값 차이가 가장 작은 주변 영역의 뎁스값으로 변환하는 단계;

상기 변환단계에 의해 변환된 뎁스맵에서 등심선을 생성하는 단계;

상기 등심선의 굴곡점들을 추출하는 단계; 및

상기 굴곡점들 사이의 곡선을 다른 타입의 곡선으로 변환하는 단계;를 포함하는 것을 특징으로 하는 뎁스맵 등심선 생성 방법을 수행할 수 있는 프로그램이 기록된 컴퓨터로 읽을 수 있는 기록매체.

발명의 설명

기 술 분 야

[0001] 본 발명은 3D 영상 처리에 관한 것으로, 더욱 상세하게는 2D 영상 콘텐츠를 3D 영상 콘텐츠로 변환하는데 있어 필요한 뎁스맵을 생성하고 활용하는 방법에 관한 것이다.

배 경 기 술

[0002] 수요에 비해 공급이 매우 부족하여, 3D 영상 콘텐츠의 부재가 심각한 현 상황에서, 3D 영상 변환 기술은 3D 영상 콘텐츠 확보 측면에서 그 중요성이 증대하고 있다.

[0003] 3D 영상 변환 기술은 아날로그나 디지털로 촬영된 기존의 2D 영상을 콘텐츠를 3D 영상 콘텐츠로 변환시키는 기술이다. 변환 기술의 기법들은 다양하지만 어느 방법이든 입력된 2D 영상에 포함된 단안 입체 정보를 해석하여 단안 영상의 뎁스를 추정한 뎁스맵을 생성하여 시차에 맞는 좌우 영상을 출력하는 것이 기본 원리이다.

[0004] 고화질을 유지하기 위해, 3D 영상 변환 기술은 매 프레임마다 모두 수작업을 거치고 있으며, 이 때문에 다수의 인력과 오랜 작업 시간이 소요되는 바, 궁극적으로는 3D 영상 콘텐츠의 생산성 문제가 시장 형성의 큰 걸림돌로 작용하고 있다.

[0005] 또한, 3D 영상 변환 작업은 매우 노동 집약적이라는 특징이 있다. 정밀한 변환 작업 시, 작업 과정에서 영상의 매 프레임마다 모두 수작업을 거쳐야하기 때문에 다수의 인력과 오랜 작업 시간이 필요하다. 완성도 높은 3D 영상을 원할수록 이러한 작업의 강도는 더욱 높아진다.

[0006] 2D 영상의 3D 영상 변환은 먼저 영상 정보를 분석해 사물과 배경, 전경과 후경 등을 분리하고, 각각의 사물과 배경에 대해 입체 값을 부여해 입체 영상으로 만들어내는 과정으로 이루어진다.

[0007] 자동 방식에서는 사물과 배경의 분리, 뎁스값 부여가 자동으로 이루어지는 데 반해 수작업에서는 기술자의 노하우에 의거해 수동적인 작업을 거친다. 이러한 변환에는 오토 컨버팅(auto converting)과 수작업 혼합(semi-auto), 완전 수작업의 3가지 작업 방식이 있다.

[0008] 오토 컨버팅은 변환 장치나 소프트웨어가 설치되어있으면 자동으로 변환해주기 때문에 비용이 저렴하지만 입체 품질이 매우 낮다. 수작업으로 이루어지는 수동방식은 매우 우수한 입체감을 얻을 수 있지만, 많은 인력과 시간, 그리고 비용이 많이 든다는 단점이 있다.

[0009] 현재 진행되고 있는 기존 영화의 3D 영상으로의 변환은 대부분이 프레임 별로 작업하는 완전 수작업 방식으로 이루어지고 있다. 오토 컨버팅과 수작업 방식의 중간인 반자동 또는 혼합방식의 변환 방식의 시스템으로 작업하는 변환 업체들이 많으며, 작업량의 비율이 자동과 수동 중 어떤 방식에 더 비중을 두느냐에 따라달리 변환 작업을 하고 있는 시점이다.

[0010] 앞으로도 자동으로 3D 영상으로 변환해주는 시스템은 발전할 것이나, 오토컨버팅 방식은 3D 영상 콘텐츠의 부재를 때우기 위한 일시적인 방편일 뿐 점점 사라질 것으로 예상되고 있다. 완성도 높은 변환 작업에는 수동방식의 장점을 살리고 단점을 극복할 수 있는 다양한 효율적인 기술의 개발이 필요하다.

발명의 내용

해결하려는 과제

[0011] 본 발명은 상기와 같은 문제점을 해결하기 위하여 안출된 것으로서, 본 발명의 목적은, 2D 영상 콘텐츠를 3D 영상 콘텐츠로 변환하는 과정에서, 연속적인 뎁스값을 갖는 뎁스맵으로부터 등심선을 생성하는 방법 및 장치를 제공함에 있다.

과제의 해결 수단

[0012] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른, 뎁스맵 등심선 생성 방법은, 뎁스맵의 뎁스값들을 다수의 양자화 단계들로 양자화 하는 단계; 및 양자화된 뎁스맵에서 등심선을 생성하는 단계;를 포함한다.

[0013] 그리고, 상기 양자화 단계는, 뎁스맵의 뎁스값들을 선형 양자화 또는 비선형 양자화할 수 있다.

[0014] 또한, 본 발명의 일 실시예에 따른 뎁스맵 등심선 생성 방법은, 양자화된 뎁스맵의 경계 부분에 존재하는 페더를 제거하는 단계;를 더 포함할 수 있다.

[0015] 그리고, 본 발명의 일 실시예에 따른 뎁스맵 등심선 생성 방법은, 양자화된 뎁스맵에서 영역 크기가 임계 크기보다 작은 영역을 페더 영역으로 설정하는 단계; 및 상기 페더 영역을 뎁스값 차이가 가장 작은 주변 영역의 뎁스값으로 변환하는 단계;를 더 포함할 수 있다.

[0016] 또한, 본 발명의 일 실시예에 따른 뎁스맵 등심선 생성 방법은, 상기 등심선의 굴곡점들을 추출하는 단계; 및 상기 굴곡점들 사이의 곡선을 다른 타입의 곡선으로 변환하는 단계;를 더 포함할 수 있다.

[0017] 한편, 본 발명의 다른 실시예에 따른, 컴퓨터로 읽을 수 있는 기록매체에는, 뎁스맵의 뎁스값들을 다수의 양자화 단계들로 양자화 하는 단계; 및 양자화된 뎁스맵에서 등심선을 생성하는 단계;를 포함하는 것을 특징으로 하는 뎁스맵 등심선 생성 방법을 수행할 수 있는 프로그램이 기록된다.

발명의 효과

[0018] 이상 설명한 바와 같이, 본 발명의 실시예들에 따르면, 2D 영상 콘텐츠를 3D 영상 콘텐츠로 변환하는 과정에서, 뎁스맵 등심선을 자동으로 생성할 수 있게 되어, 수작업을 통해 그래픽툴에서 뎁스맵 경계의 곡선을 그리는 작업이 필요 없게 된다. 이에 의해, 필요한 인력의 감소는 물론, 작업 속도를 크게 향상시킬 수 있게 된다.

[0019] 또한, 본 발명의 실시예들에 따르면, 뎁스맵 등심선의 모든 구간을 베지어 곡선으로 변환하기 때문에 미세 보정이 용이하다는 장점이 있다.

도면의 간단한 설명

[0020] 도 1은 본 발명이 적용가능한 3D 콘텐츠 제작 방법의 설명에 제공되는 흐름도,

도 2는, 도 1의 부연 설명에 제공되는 도면,

도 3은 본 발명의 일 실시예에 따른, 뎁스맵 등심선 생성 방법의 설명에 제공되는 흐름도,

도 4는 뎁스맵 선형 양자화 결과를 예시한 도면,

도 5는 뎁스맵 비선형 양자화 결과를 예시한 도면,

도 6은 양자화된 뎁스맵의 페더를 나타낸 도면,

도 7은 양자화된 뎁스맵에 레이블링 기법을 이용하여 페더 영역을 추출한 결과를 예시한 도면,

도 8은 페더 영역을 제거하는 기법의 설명에 제공되는 도면,

도 9는 페더 영역이 제거된 양자화된 뎁스맵을 예시한 도면,

도 10은 양자화된 뎁스맵에서 양자화 단계별로 생성한 마스크를 도시한 도면,

도 11은 각 양자화 단계별 등심선을 나타낸 도면,

도 12는 양자화된 뎁스맵 등심선을 나타낸 도면,

도 13은 등심선 내 굴곡 정도를 계산하는 방법을 나타낸 도면,

도 14는 굴곡점 후보들에서 최종 굴곡점을 추출한 결과를 나타낸 도면,

도 15는 베지어 곡선을 이용하여 등심선을 재생성한 뎁스맵을 나타낸 도면, 그리고,

도 16은 본 발명의 다른 실시예에 따른 3D 콘텐츠 제작 시스템의 블럭도이다.

발명을 실시하기 위한 구체적인 내용

[0021] 이하에서는 도면을 참조하여 본 발명을 보다 상세하게 설명한다.

[0022] **1. 3D 콘텐츠 제작**

[0023] 도 1은 본 발명이 적용가능한 3D 콘텐츠 제작 방법의 설명에 제공되는 흐름도이다. 도시된 3D 콘텐츠 제작 방법은, 2D 콘텐츠로부터 3D 콘텐츠를 생성하는 과정이다.

[0024] 도 1에 도시된 바와 같이, 먼저 2D 영상 콘텐츠를 입력받아(S105), 샷(Shot) 단위로 분할한다(S110). S110단계에 의해 2D 영상 콘텐츠는 다수의 샷들로 분할된다. 동일 샷에 포함된 프레임들은 동일/유사한 배경을 갖는다.

[0025] S110단계에서의 샷 분할은 전문 아티스트에 의한 수작업으로 수행될 수도 있고, 유사한 프레임들을 샷 단위로 구분하는 프로그램을 이용하여 자동으로 수행될 수도 있다.

[0026] 이후, 하나의 샷을 선정하고(S115), 선정된 샷에서 키-프레임을 선정한다(S120). S115단계에서의 샷 선정은, 2D 영상 콘텐츠에서 샷의 시간 순서에 따라 순차적으로 자동 선정(즉, 첫 번째 샷을 먼저 선정하고, 이후 다음 샷을 선정)할 수 있음은 물론, 전문 아티스트의 판단에 의해 선정 순서를 다른 순서로 정할 수도 있다. 이하에서는, 샷 선정이 샷의 시간 순서 따라 순차적으로 이루어지는 것을 상정하겠다.

[0027] 아울러, S120단계에서의 키-프레임 선정도 샷을 구성하는 프레임들 중 시간적으로 가장 앞선 프레임(즉, 첫 번째 프레임)을 키-프레임으로 자동 선정할 수 있음은 물론, 전문 아티스트의 판단에 의해 그 밖의 다른 프레임을 키-프레임으로 선정할 수도 있다. 이하에서는, 키-프레임이 샷의 첫 번째 프레임인 것을 상정하겠다.

[0028] 다음, S120단계에서 선정된 키-프레임에 대한 뎁스맵(Depth Map)을 생성한다(S125). S125단계에서의 뎁스맵 생성 역시 프로그램을 이용한 자동 생성은 물론, 전문 아티스트의 수작업에 의한 생성도 가능하다.

[0029] 이후, S125단계에서 생성된 뎁스맵의 등심선을 생성한다(S130). S130단계에서, 뎁스맵 등심선은 프로그램을 이용하여 뎁스맵을 뎁스에 따라, 다수의 양자화 단계들로 양자화 처리하여 생성한다.

[0030] 다음, S130단계에서 생성된 등심선을 다음-프레임에 매치시키다(S135). 움직임 객체가 있거나 화각이 이동하여, 다음-프레임이 이전-프레임인 키-프레임과 다른 부분이 있는 경우, 키-프레임의 등심선은 다음-프레임에 정확히 매치되지 않는다.

[0031] 이에 따라, 다음-프레임에 완전히 매치되도록, 등심선을 부분적으로 이동 처리(매치 무브)한다(S140). S140단계에서의 등심선 이동 처리 역시 프로그램을 이용하여 자동으로 수행된다. S140단계에 의해, 다음-프레임에 완전하게 매치된 등심선이 생성된다.

[0032] 이후, S140단계에서 매치 무브된 등심선을 기반으로, 다음-프레임의 뎁스맵을 생성한다(S145). S145단계에서 생성되는 다음-프레임의 뎁스맵은 양자화된 상태의 뎁스맵이다.

[0033] 이에, 뎁스맵을 보간하여, 다음-프레임의 뎁스맵을 완성한다(S150). S150단계에서의 뎁스맵 보간 처리 역시 프로그램을 이용하여 자동으로 수행된다.

[0034] S135단계 내지 S150단계는, 샷을 구성하는 모든 프레임들에 대해 완료될 때까지 반복된다(S155). 즉, 샷의 두 번째 프레임에 대한 뎁스맵이 완성되면, 두 번째 프레임의 등심선을 세 번째 프레임에 매치시키고(S135), 등심선 매치 무브를 수행한 후에(S140), 매치 무브된 등심선을 기반으로, 세 번째 프레임의 뎁스맵을 생성하고(S145), 보간을 통해 뎁스맵을 완성하게 되며(S150), 세 번째 프레임 이후의 프레임들에 대해서도 동일한 작업이 수행된다.

[0035] 샷을 구성하는 모든 프레임들에 대해 뎁스맵 생성이 완료되면(S155-Y), 두 번째 샷에 대해 S115단계 내지 S155단계를 반복하며, 두 번째 샷을 구성하는 모든 프레임들에 대해 뎁스맵 생성이 완료되면, 이후의 샷들에 대해서도 동일한 작업이 수행된다.

[0036] 2D 영상 콘텐츠를 구성하는 모든 샷들에 대해 위 절차가 완료되면(S160-Y), 지금까지 생성된 뎁스맵들을 이용하여, 2D 영상 콘텐츠를 구성하는 프레임들을 3D 변환처리한다(S165). 그리고, S165단계에서 변환처리 결과를 3D 영상 콘텐츠로 출력한다(S170).

[0037] 도 2는, 도 1의 부연 설명에 제공되는 도면이다. 도 2에는,

[0038] 1) S115단계에서 선정된 샷을 구성하는 프레임들 중 첫 번째 프레임을 키-프레임으로 선정하고(S120),

[0039] 2) 선정된 키-프레임에 대한 뎁스맵을 생성한 후에(S125),

[0040] 3) 뎁스맵의 등심선을 추출하고(S130),

[0041] 4) 추출한 등심선을 다음-프레임에 매치시켜, 매치 무브한 후(S135, S140),

[0042] 5) 매치 무브된 등심선을 기반으로, 다음-프레임의 뎁스맵을 생성하고(S145),

[0043] 6) 생성된 뎁스맵을 보간하여, 다음-프레임의 뎁스맵을 완성(S150)하는 과정이 도식적으로 나타나 있다.

[0044] **2. 뎁스맵 등심선 생성**

[0045] 도 1과 도 2에 도시된 S130단계의 뎁스맵 등심선 생성 과정에 대해, 도 3을 참조하여 상세히 설명한다. 도 3은 본 발명의 일 실시예에 따른, 뎁스맵 등심선 생성 방법의 설명에 제공되는 흐름도이다.

[0046] 도 3에 도시된 바와 같이, 먼저, 키-프레임의 뎁스맵을 입력받고(S131), 입력된 뎁스맵에 대해 선형 양자화(S132) 또는 비선형 양자화(S133)를 수행한다. S131단계에서 입력되는 뎁스맵에서 뎁스들값은 연속적인 반면, S132단계 또는 S133단계에 의해 뎁스맵의 뎁스값들은 다수의 양자화 단계들로 양자화 된다.

[0047] S132단계 또는 S133단계에 적용할 양자화 단계의 개수는 필요와 사양에 따라 지정할 수 있다.

[0048] 이후, 양자화된 뎁스맵의 경계 부분들에 존재하는 페더를 제거하고(S134), 페더가 제거된 양자화된 뎁스맵에서 등심선을 추출한다(S135). S135단계에서 추출하는 등심선은 뎁스맵에서 동일한 뎁스값을 갖는 화소들을 연결한 선이다.

[0049] 다음, 뎁스맵의 등심선에서 굴곡점들을 추출하고(S136), 굴곡점들 사이의 곡선을 베지어 곡선으로 변환하여(S137), 뎁스맵 등심선을 완성한다. S136단계에서 추출되는 굴곡점들은 등심선에서 방향 변화가 큰 지점들을 말한다.

[0050] S137단계에서의 변환은, 2개의 굴곡점들과 그 사이의 1/4, 3/4 지점의 중간점을 계산하여, 총 4개의 포인트를 이용하여 3차 베지어 커브의 제어 포인트들을 계산하고, 계산된 제어 포인트들을 이용하여 베지어 곡선으로 변환하는 과정에 의한다.

[0051] 이하에서, 도 3을 구성하는 단계들에 대해, 보다 구체적으로 설명한다.

[0052] **3. 키-프레임 뎁스맵 선형 양자화**

[0053] 키-프레임 뎁스맵 선형 양자화는, 사용자가 설정한 양자화 단계의 개수에 따라 뎁스맵의 뎁스값 범위를 동일한 크기로 구분하여 뎁스맵을 양자화 하는 기법으로 아래의 식 (1)로 나타낼 수 있다.

[0054]
$$Q(x) = \Delta \cdot \left(\left\lfloor \frac{x - x_{Min}}{\Delta} \right\rfloor + \frac{1}{2} \right) + x_{Min}$$

$$\Delta = \frac{x_{Max} - x_{Min}}{step} \qquad (1)$$

[0055] 여기서, x는 입력된 키-프레임 뎁스맵의 뎁스값을 나타내며 x_{Max}, x_{Min}은 각 각 뎁스맵에서의 최고 뎁스값, 최소 뎁스값을 나타낸다. 그리고, step은 양자화 단계의 개수를 나타내며 Δ는 양자화 단계의 크기를 나타낸다. 마지막으로, Q(x)는 양자화된 뎁스값을 나타낸다.

[0056] 위 식 (1)로 표현된 바와 같이, 양자화 단계의 크기는, 뎁스맵에서의 최고 뎁스값과 최소 뎁스값을 파악하고, 이들의 차를 사용자가 설정한 양자화 단계의 개수로 나누어 산출함을 알 수 있다.

[0057] 위 식 (1)을 이용하면, 입력된 키-프레임 뎁스맵이 나타낼수 있는 뎁스값의 범위를 동일한 크기의 양자화 단계로 구분한 양자화 뎁스맵을 얻을 수 있다. 도 4에는 선형 양자화 기법을 통해 얻은 양자화 뎁스맵과 그 히스토그램을 나타내었다. 도 4에서, (a)는 입력된 키-프레임 뎁스맵이고, (b)는 양자화 단계의 크기를 10으로 설정하여 선형 양자화한 뎁스맵이며, (c)는 양자화 단계의 크기를 20으로 설정하여 선형 양자화한 뎁스맵이다.

[0058] **4. 키-프레임 뎁스맵 비선형 양자화**

[0059] 키-프레임 뎁스맵 비선형 양자화는, 사용자가 설정한 양자화 단계의 개수에 따라 각 양자화 단계의 크기를 키-프레임 뎁스맵에 최적화된 크기로 계산하여 키-프레임 뎁스맵을 양자화하는 기법이다.

[0060] 최적화된 양자화 단계의 크기를 계산하기 위해 K-means 군집화 기법을 사용할 수 있다. 아래의 (2)는 K-Means 군집화 기법을 이용한 키-프레임 뎁스맵 비선형 양자화의 식을 나타낸다.

[0061] $$V = \sum_{i=1}^{k} \sum_{j \in S_i} \left| x_j - \mu_i \right|^2 \quad (2)$$

[0062] 여기서, k는 양자화 단계의 개수를 나타내며, μ_i는 i번째 양자화 단계의 평균값을 나타낸다. 또한, S_i는 i번째 양자화 단계에 속하는 화소들의 집합을 나타내며, x_j는 집합에 속하는 화소의 뎁스값을 나타낸다. 그리고, V는 전체 분산을 나타낸다.

[0063] 키-프레임 뎁스맵 비선형 양자화는, 사용자가 설정한 양자화 단계의 개수에 따라 양자화 단계의 크기를 동일하게 나누고, 각 단계의 중심값을 양자화 단계의 초기 평균값들로 정한다. 이후, 뎁스맵의 각 화소들을 초기 평균값들과 비교하여 가장 가까운 값을 찾고 그 양자화 단계의 집합으로 구분한다. 그리고, 전체 뎁스맵이 구분되면 각 양자화 단계에 속한 집합의 평균값을 다시 계산하여 양자화 단계의 평균값으로 갱신한다. 이 과정을 현재의 평균값들과 갱신된 평균값들이 변하지 않거나 전체 분산값이 더 이상 작아지지 않을 때까지 반복한다. 모든 과정이 완료되면 양자화 단계의 집합에 속한 화소들을 그 집합의 평균값으로 대체하여 비선형 양자화를 수행한다.

[0064] 도 5에는 비선형 양자화 기법을 통해 얻은 양자화 뎁스맵과 그 히스토그램을 나타내었다. 도 5에서, (a)는 입력된 키-프레임 뎁스맵이고, (b)는 양자화 단계의 크기를 10으로 설정하여 비선형 양자화한 뎁스맵이며, (c)는 양자화 단계의 크기를 20으로 설정하여 비선형 양자화한 뎁스맵이다.

[0065] **5. 양자화된 뎁스맵의 페더 제거**

[0066] 원본 뎁스맵의 경계에는 부드러운 변화를 위한 페더 영역이 존재한다. 선형 또는 비선형 양자화 기법을 통해 생성된 양자화 뎁스맵의 경계 부분에는 페더 영역이 작은 영역들로 나타난다.

[0067] 도 6에는 양자화된 뎁스맵의 페더를 나타내었다. 도 6의 (a)에는 원본 뎁스맵을, (b)에는 양자화된 뎁스맵을 각각 나타내었다.

[0068] 양자화된 뎁스맵에서 페더 영역을 추출하기 위해, 레이블링 기법을 이용한다. 레이블링 기법은 영상에서 주변 화소와 연속으로 같은 값을 갖는 영역을 하나의 영역으로 분리하는 기법이다. 레이블링 기법을 이용하여 양자화된 뎁스맵을 각각의 영역으로 분리하면, 페더들은 주변 뎁스값들과 다른 뎁스값을 갖기 때문에 하나의 작은 영역으로 분리된다.

[0069] 페더 영역은 페더 영역이 아닌 다른 영역과는 다르게 매우 작은 영역이기 때문에 영역에 속한 화소 수가 매우 작은 특징이 있다. 따라서, 레이블링 기법으로 분리된 영역에 속한 화소 수가 일정 임계치보다 작을 경우 페더 영역으로 추출한다.

[0070] 도 7에는 양자화된 뎁스맵에 레이블링 기법을 이용하여 페더 영역을 추출한 예를 나타내었다. 도 7의 (a)에는 레이블링으로 페더 영역을 추출한 결과를 나타내었고, 도 7의 (b)에는 페더 영역을 확대하여 나타내었다.

[0071] 도 8은 페더 영역을 제거하는 기법의 설명에 제공되는 도면이다. 도 8에서, A 영역은 페더 영역을 나타내고, B 영역과 C 영역은 비-페더 영역(페더가 아닌 영역)을 나타낸다.

[0072] 페더 영역을 제거하기 위해, 페더 영역인 A 영역의 주변 화소의 뎁스값을 탐색하고, 탐색된 뎁스값과 페더 영역의 뎁스값의 차이값을 구한다. 그리고, 가장 작은 차이값을 보이는 주변 화소의 뎁스값으로 페더 영역의 뎁스값을 대체하여 페더 영역을 제거 한다.

[0073] 이 기법에 의해 페더 영역이 제거된 양자화된 뎁스맵을 도 9에 예시하였다. 도 9의 (a)에는 페더 영역이 제거

되지 않은 양자화된 뎁스맵을 나타내었고, 도 9의 (b)에는 페더 영역이 제거된 양자화된 뎁스맵을 나타내었다.

[0074] **6. 양자화 뎁스맵 기반 등심선 추출**

[0075] 양자화 뎁스맵 기반 등심선 추출은, 페더가 제거된 양자화 뎁스맵에서 같은 뎁스값을 갖는 영역을 잇는 폐곡선을 생성하는 과정이다.

[0076] 이를 위해, 도 10에 도시된 바와 같이, 먼저 양자화된 뎁스맵에서 양자화 단계별로 마스크를 생성한다. 다음, 도 11에 도시된 바와 같이, 각 양자화 단계별로 생성된 마스크에서 외곽선 추출을 수행한다. 이에 의해, 각 양자화 단계별 등심선을 얻을 수 있다.

[0077] 이후, 각 양자화 단계별로 등심선을 통합하면, 도 12에 도시된 바와 같이, 양자화된 뎁스맵 등심선을 얻을 수 있게 된다.

[0078] **7. 등심선 내 굴곡점 추출**

[0079] 굴곡점은 등심선에서 방향의 변화가 큰 지점을 의미한다. 추출된 등심선에서 각각의 곡선을 구분하기 위해 굴곡점을 추출한다. 도 13에 등심선 내 굴곡 정도를 계산하는 방법을 나타내었고, 아래의 식 (3)은 등심선 내의 기준 화소의 굴곡정도를 계산하는 식이다.

$$D = \sum_{i=0}^{7} \left| FV_i - PV_i \right|$$

$$FV_i = \sum_{d=1}^{S} \{W_H - C(d-1)\} \times B_{i,d}, \quad PV_i = \sum_{d=1}^{S} \{W_H - C(d-1)\} \times B_{i,d}$$

[0080] $$C = (W_H - W_L)/S \qquad (3)$$

[0081] 여기서, D는 기준 화소의 굴곡 정도를 나타낸다. 그리고, FV는 기준 화소 이전의 화소들의 방향 성분 분포를 나타내고, PV는 기준 화소 이후들의 화소의 방향 성분 분포를 나타낸다. W_H와 W_L은 각각 기준 화소의 방향과 가장 비슷한 방향의 가중치와 기준 화소의 방향과 가장 다른 방향의 가중치를 나타내며, S는 기준 화소 이전, 이후의 방향 분포를 계산할 화소수를 나타내고, d는 기준 화소와의 거리를 나타낸다. 마지막으로 $B_{i,d}$는 d 거리에 있는 화소가 i 방향이면 1의 값, 아니면 0의 값을 갖는다. 등심선의 한 화소의 굴곡 정도를 구하기 위해 앞, 뒤의 화소들의 방향 변화 분포를 계산하며 앞, 뒤의 분포의 차이값을 구해서 기준 화소의 방향 변화의 정도를 계산할 수 있다. 등심선 내의 모든 화소에 대해 굴곡 정도 D를 계산하고 이값이 일정 임계치를 넘을 경우 그 화소를 굴곡점 후보로 정한다.

[0082] 굴곡점 후보들은 등심선에 변화가 시작되는 부분부터 연속적으로 나타난다. 따라서, 연속적인 굴곡점 후보들에서 최대 굴곡 정도를 갖는 굴곡점 후보를 최종 굴곡점으로 추출한다. 도 14에는 굴곡점 후보들에서 최종 굴곡점을 추출한 결과를 나타내었다. 구체적으로, 도 14의 (a)에는 굴곡점 후보들을 녹색으로 나타내었고, 도 14의 (b)에는 최종 굴곡점을 녹색으로 나타내었다.

[0083] **8. 굴곡점 사이 곡선의 베지어 곡선 변환**

[0084] 각 양자화 단계의 등심선은 추출된 굴곡점을 양끝점으로 하여 각각의 곡선들로 나눌 수 있는데, 이 곡선들을 각각 3차 베지어 곡선으로 변환한다. 본 발명의 실시예에서는 양끝의 굴곡점과 그 사이의 중간점을 계산하고, 이 점들을 이용하여 베지어 곡선 식을 추정한다. 이를 통해, 추청된 베지어 곡선의 제어 포인트들을 얻는다. 아래의 식 (4)는 3차 베지어 곡선의 식을 나타낸다.

[0085] $$B(t) = P_0(1-t)^3 + 3P_1t(1-t)^2 + 3P_2t^2(1-t) + P_3t^3 \qquad (4)$$

[0086] 여기서, P는 베지어 곡선의 제어 포인트를 나타내며, t는 0~1사이의 곡선의 변화구간을 나타낸다. 마지막으로, B(t)는 t구간일 때 베지어 곡선의 포인트를 나타낸다. 식 (4)를 식 (5)와 같이 t값을 나타내는 T행렬, 제어 포

인트를 나타내는 P행렬, t의 계수들을 나타내는 M행렬들의 곱의 형태로 나타낼 수 있다. 행렬의 곱의 형태로 변환하면 식 (6)과 같이 나타낼 수 있다.

[0087]
$$T=\begin{bmatrix} t^3 & t^2 & t & 1 \end{bmatrix} \qquad M=\begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad P=\begin{bmatrix} P_0 \\ P_1 \\ P_2 \\ P_3 \end{bmatrix} \tag{5}$$

[0088]
$$B(t)=TMP \tag{6}$$

[0089] 아래의 식 (7)과 같이 입력으로 받는 양끝의 굴곡점과 그 사이의 중간점들의 행렬인 K는 다시 x성분과 y성분으로 각각 나누어 생각할 수 있다. 본 발명의 실시예에서는 중간점으로 1/4, 3/4 지점의 2개의 값을 계산하여 총 4개의 점을 입력으로 사용한다.

[0090]
$$K=\begin{bmatrix} K_0 \\ K_1 \\ \cdots \\ K_n \end{bmatrix} \rightarrow K_i=\{x_i, y_i\} \rightarrow x=\begin{bmatrix} x_0 \\ x_1 \\ \cdots \\ x_n \end{bmatrix}, \quad y=\begin{bmatrix} y_0 \\ y_1 \\ \cdots \\ y_n \end{bmatrix} \tag{7}$$

[0091] 아래의 식 (8)과 같이 입력된 점들의 값들과 곡선의 처음과 끝인 굴곡점을 이용하여 비율을 계산하면 그 입력점의 대략적인 t값을 구할 수가 있다. 모든 입력 점들에 대해 t값을 각각 구해 이를 행렬로 나타내면 아래의 식 (9)와 같다.

[0092]
$$t_i=\frac{|d_i-d_{i-1}|}{\sum_{j=2}^{n}|d_i-d_{i-1}|}$$

$$d_i=\sum_{j=2}^{n}|K_j-K_{j-1}| \tag{8}$$

[0093]
$$\hat{T}=\begin{bmatrix} t_1^3 & t_1^2 & t_1 & 1 \\ t_2^3 & t_2^3 & t_2 & 1 \\ \cdots & & & \\ t_n^3 & t_n^3 & t_n & 1 \end{bmatrix} \tag{9}$$

[0094] 최적의 곡선을 만들기 위해 실제 등심선의 점들과 베지어 곡선을 통해 복원한 점들과의 화소 위치 차이를 구해서 생성한 베지어 곡선의 오차를 계산한다. 식 (10)은 y성분에 대해 오차를 구하는 식을 나타내며 이를 행렬의 계산으로 정리한 식 또한 나타낸다.

[0095]
$$E(P_y)=\sum_{i=1}^{n}(y_i-B(t_i))^2 \rightarrow E(P_y)=(y-\hat{T}MP_y)^T(y-\hat{T}MP_y) \tag{10}$$

[0096] 식 (10)에서 오차값이 최소가 될 때의 Py행렬의 값이 최적의 베지어 곡선의 제어 포인트의 y성분이 된다. 식 (11)은 최소오차 기반의 Py의 값을 구하는 과정을 나타낸다.

[0097]
$$\frac{\partial E}{\partial P}=0=-2\hat{T}^T(y-\hat{T}MP_y) \rightarrow P_y=M^{-1}(\hat{T}^T\hat{T})^{-1}\hat{T}^TY \tag{11}$$

[0098] 계산을 통해 최적의 베지어 곡선의 y성분 포인트를 얻었다면 입력을 x성분으로 바꾸어 같은 과정을 반복하면 최적의 x성분 또한 얻을 수 있다. 도 15는 변환된 베지어 곡선의 포인트들을 이용하여 등심선을 재생성한 뎁스맵을 나타내었다.

[0099] **9. 3D 콘텐츠 제작 시스템**

[0100] 도 16은 본 발명의 다른 실시예에 따른 3D 콘텐츠 제작 시스템의 블럭도이다. 본 발명의 실시예에 따른 3D 콘텐츠 제작 시스템(200)은, 도 16에 도시된 바와 같이, 2D 영상 입력부(210), 뎁스맵 생성부(220), 3D 변환부(230) 및 3D 영상 출력부(240)를 포함한다.

[0101] 뎁스맵 생성부(220)는 키-프레임 뎁스맵을 이용하여, 2D 영상 입력부(210)를 통해 입력되는 2D 영상 프레임들의 뎁스맵을 생성한다. 키-프레임 뎁스맵은, 2D 영상에서 샷 마다 하나씩 생성되는데, 프로그램에 의한 자동 생성은 물론, 전문 아티스트에 의한 수동 생성 모두가 가능하다.

[0102] 뎁스맵 생성부(220)는 키-프레임 뎁스맵의 등심선을 추출하고, 추출한 등심선을 다음-프레임에 매치시켜 매치 무브한 후, 매치 무브된 등심선을 기반으로 다음-프레임의 뎁스맵을 생성하고 보간하여 다음-프레임의 뎁스맵을 완성하는 절차에 의해, 2D 영상 프레임들에 대한 뎁스맵들을 생성한다.

[0103] 3D 변환부(230)는 뎁스맵 생성부(220)에서 생성된 뎁스맵들을 이용하여, 2D 영상 프레임들을 3D 변환처리한다. 그리고, 3D 영상 출력부(240)는 3D 변환부(230)에 의한 변환처리 결과를 3D 영상으로 출력한다.

[0104] 또한, 이상에서는 본 발명의 바람직한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

부호의 설명

[0105] 200 : 3D 콘텐츠 제작 시스템

210 : 2D 영상 입력부

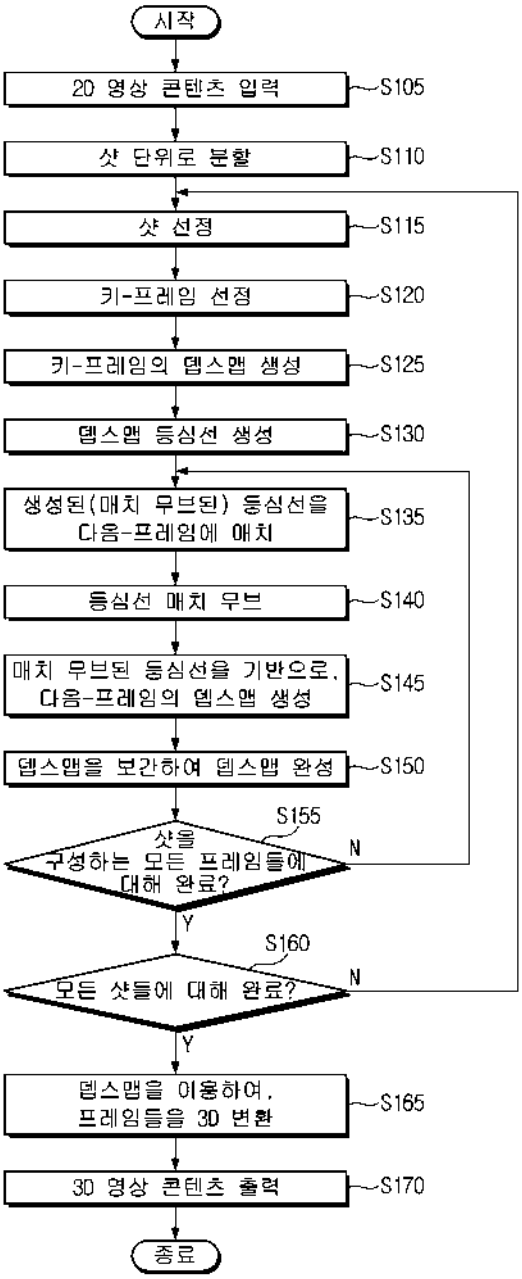
220 : 뎁스맵 생성부

230 : 3D 변환부

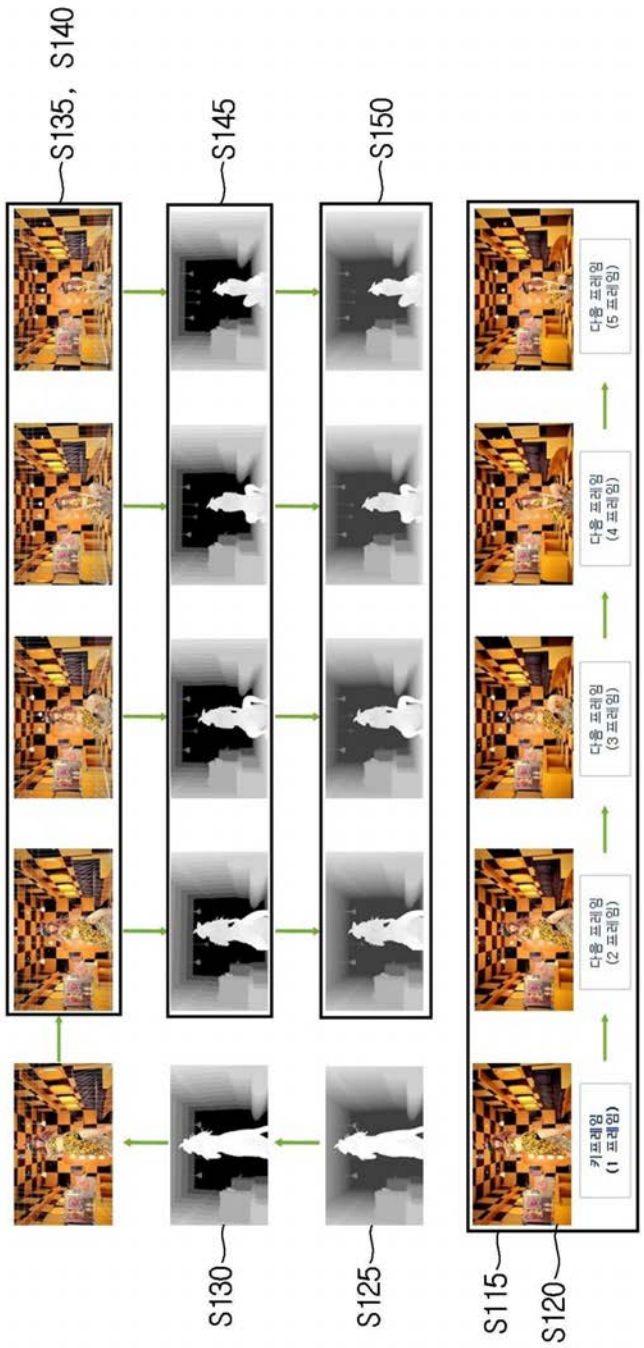
240 : 3D 영상 출력부

도면

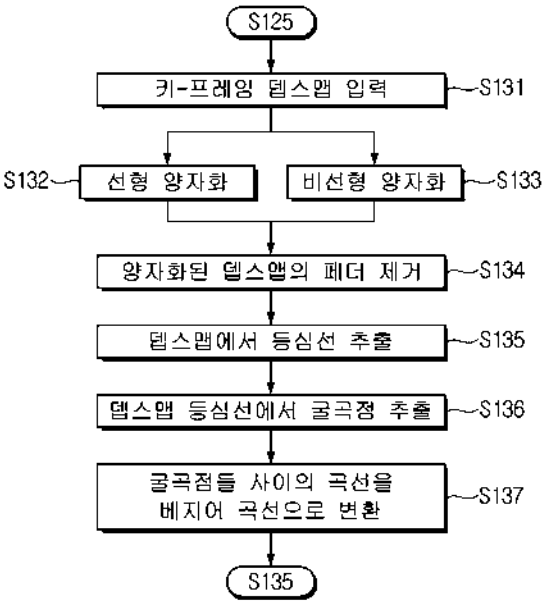
도면1


시작
2D 영상 콘텐츠 입력
S105
샷 단위로 분할
S110
샷 선정
S115
키-프레임 선정
S120
키-프레임의 뎁스맵 생성
S125
뎁스맵 등심선 생성
S130
생성된(매치 무브된) 등심선을 다음-프레임에 매치
S135
등심선 매치 무브
S140
매치 무브된 등심선을 기반으로, 다음-프레임의 뎁스맵 생성
S145
뎁스맵을 보간하여 뎁스맵 완성
S150
S155
샷을 구성하는 모든 프레임들에 대해 완료?
N
Y
S160
모든 샷들에 대해 완료?
N
Y
뎁스맵을 이용하여, 프레임들을 3D 변환
S165
3D 영상 콘텐츠 출력
S170
종료

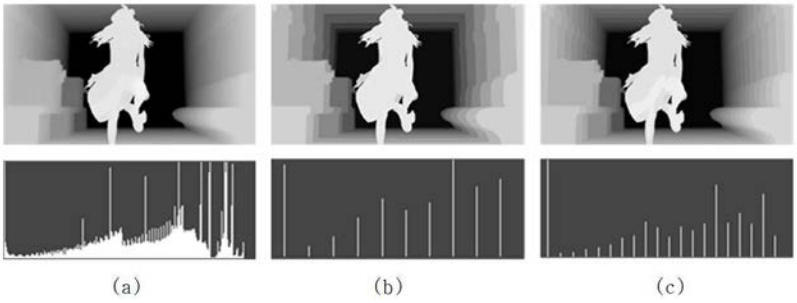
도면2


S135, S140
S145
S150
S130
S125
S115
S120
키프레임 (1 프레임)
다음 프레임 (2 프레임)
다음 프레임 (3 프레임)
다음 프레임 (4 프레임)
다음 프레임 (5 프레임)

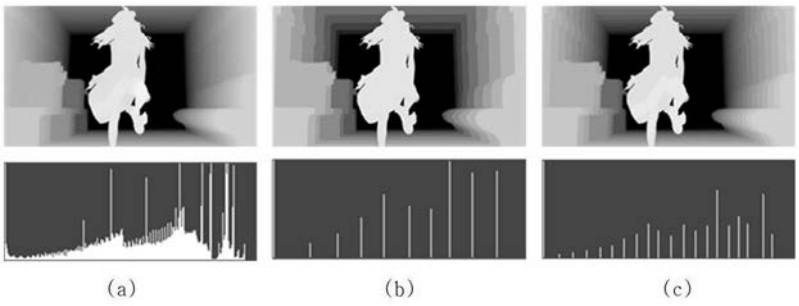
도면3


S125
키-프레임 뎁스맵 입력
S131
S132
선형 양자화
비선형 양자화
S133
양자화된 뎁스맵의 페더 제거
S134
뎁스맵에서 등심선 추출
S135
뎁스맵 등심선에서 굴곡점 추출
S136
굴곡점들 사이의 곡선을
베지어 곡선으로 변환
S137
S135

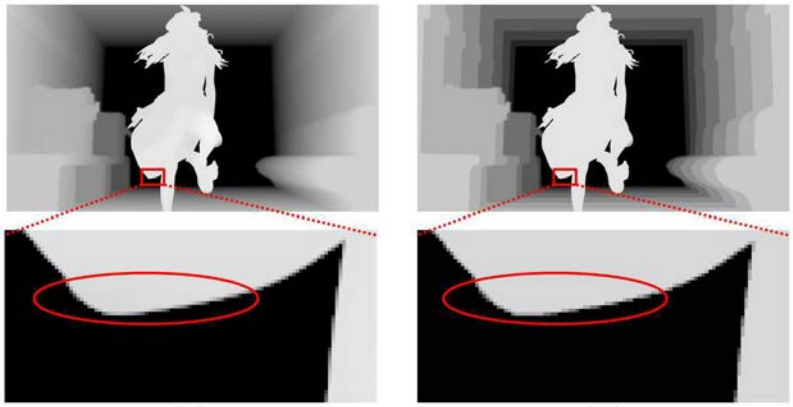
도면4


(a)
(b)
(c)

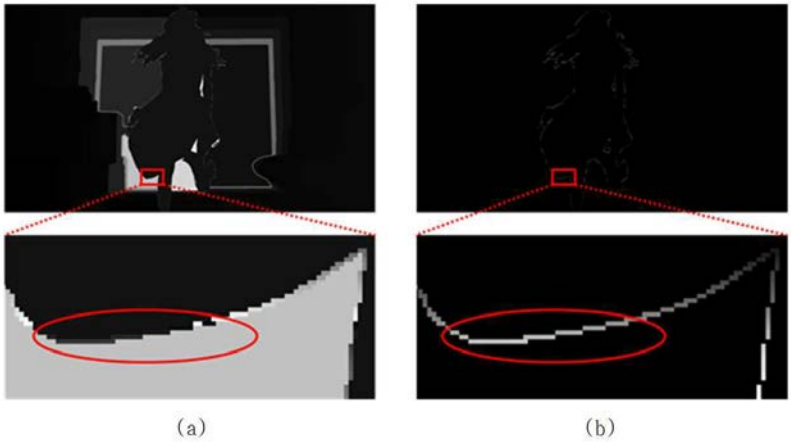
도면5


(a)
(b)
(c)

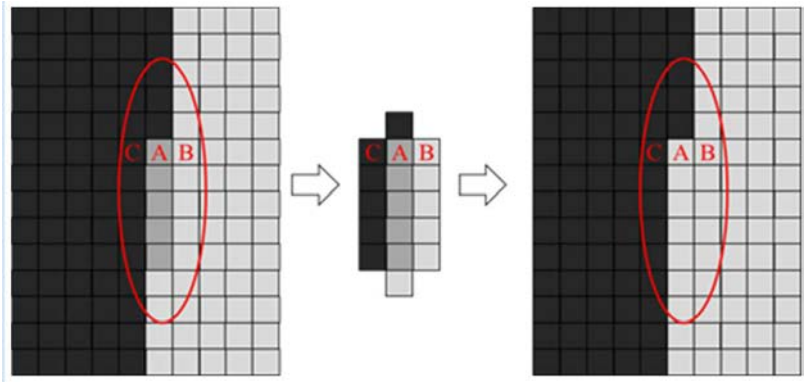
도면6



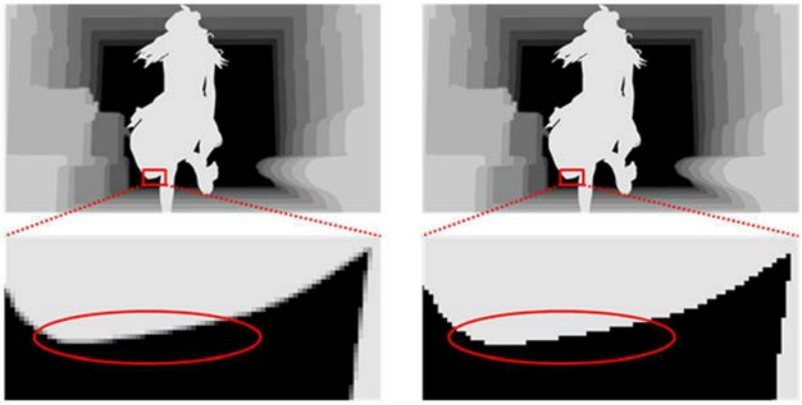
도면7



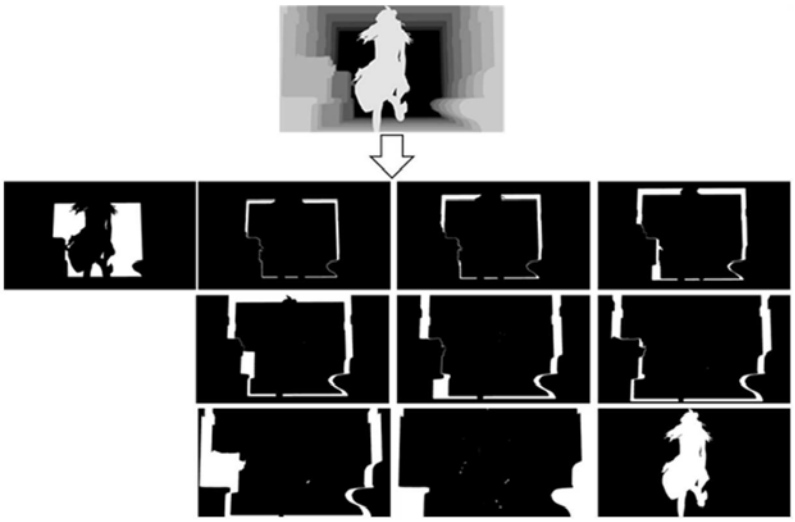
도면8



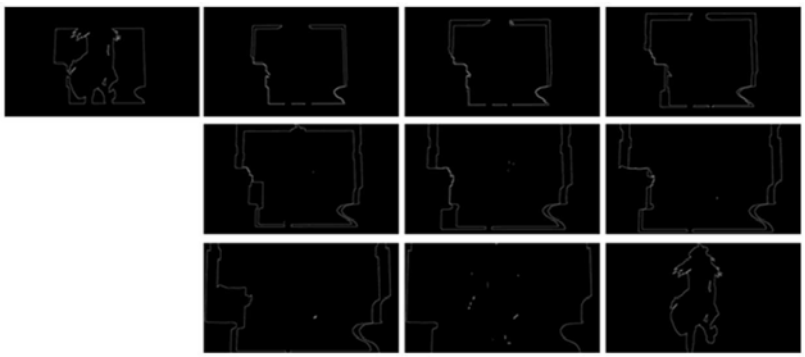
도면9



도면10

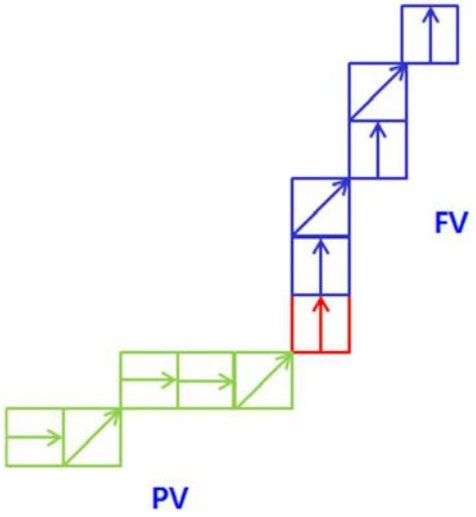


도면11

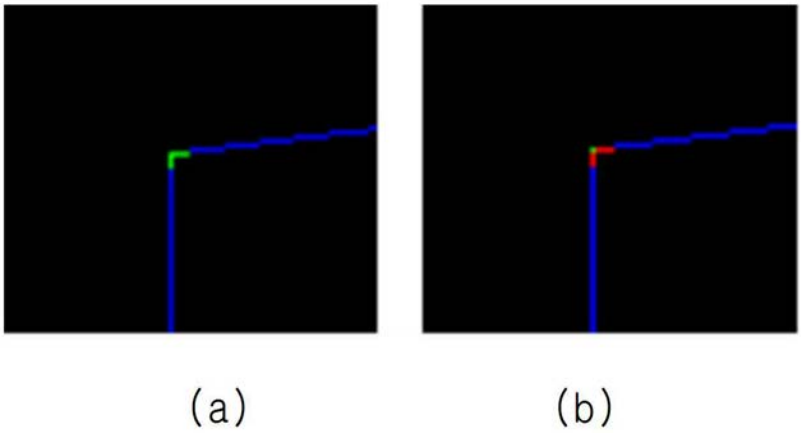


도면12

도면13

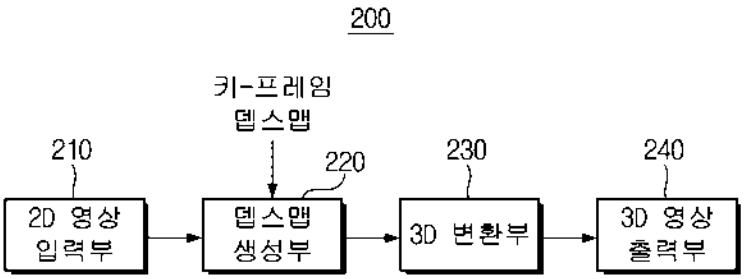


도면14



도면15

도면16



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2016-0062771
(43) 공개일자 2016년06월03일

(51) 국제특허분류(Int. Cl.)
G06T 5/00 (2006.01)
(21) 출원번호 10-2014-0165076
(22) 출원일자 2014년11월25일
심사청구일자 없음

(71) 출원인
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
조충상
경기 성남시 분당구 장미로 55, 116동 805호 (야탑동, 장미마을아파트)
신화선
경기 용인시 기흥구 보정로 26, 101동 1601호 (보정동, 신촌마을상록데시앙)
강주형
대전 서구 계룡로 356, 102호 (갈마동, 낙원빌라)
(74) 대리인
남충우

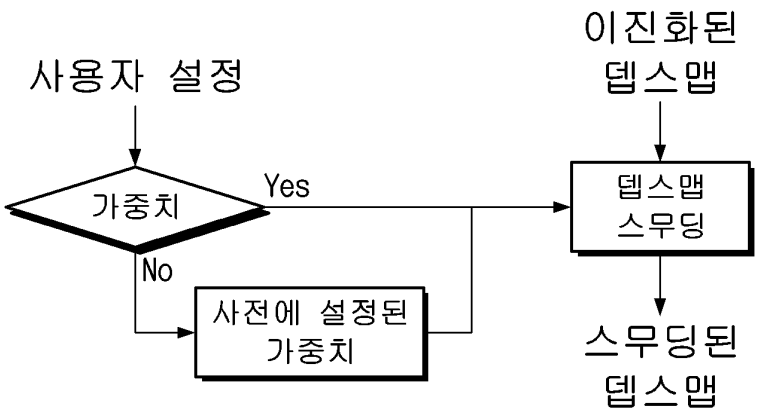
전체 청구항 수 : 총 6 항

(54) 발명의 명칭 **영상 스무딩 방법 및 장치**

(57) 요 약

영상 스무딩 방법 및 장치가 제공된다. 본 발명의 실시예들에 따른 영상 스무딩 방법은, 영상 스무딩에 이용되는 코스트 함수에 다양한 항들을 적응적/선택적으로 추가하여 영상 스무딩을 수행한다. 이에 의해, 에지 부분에서의 스무딩 처리가 우수해져, 궁극적으로는 영상 전체의 품질을 향상시킬 수 있게 된다.

대 표 도 - 도7



이 발명을 지원한 국가연구개발사업
과제고유번호 10048070
부처명 산업통상자원부
연구관리전문기관 한국산업기술평가관리원
연구사업명 (산업부)기술료지원사업
연구과제명 양자화 기반 2D 영상콘텐츠의 3D변환 기술과 이를 활용한 입체영상 저작도구 개발 및 시험 적용
기 여 율 1/1
주관기관 전자부품연구원
연구기간 2013.12.01 ~ 2014.11.30

명 세 서

청구범위

청구항 1

양자화된 영상을 입력받는 단계;

입력된 영상을 스무딩하는 단계; 및

스무딩된 영상을 출력하는 단계;를 포함하고,

상기 스무딩 단계는,

스무딩 영상의 2차 편미분이 반영된 항을 포함하는 코스트 함수를 이용하여, 상기 입력된 영상을 스무딩하는 것을 특징으로 하는 영상 스무딩 방법.

청구항 2

청구항 1에 있어서,

상기 코스트 함수는,

스무딩 영상의 x에 관한 2차 편미분 및 스무딩 영상의 y에 관한 2차 편미분이 반영된 항을 포함하는 것을 특징으로 하는 영상 스무딩 방법.

청구항 3

청구항 2에 있어서,

상기 코스트 함수는,

스무딩 영상의 xy에 관한 편미분이 반영된 항을 더 포함하는 것을 특징으로 하는 영상 스무딩 방법.

청구항 4

청구항 3에 있어서,

상기 코스트 함수는,

스무딩 영상의 x에 관한 1차 편미분 및 스무딩 영상의 y에 관한 1차 편미분이 반영된 항을 더 포함하는 것을 특징으로 하는 영상 스무딩 방법.

청구항 5

청구항 1에 있어서,

항들의 가중치들은,

사용자의 입력에 의해 설정되는 것을 특징으로 하는 영상 스무딩 방법.

청구항 6

양자화된 영상을 입력받는 단계;

입력된 영상을 스무딩하는 단계; 및

스무딩된 영상을 출력하는 단계;를 포함하고,

상기 스무딩 단계는,

스무딩 영상의 2차 편미분이 반영된 항을 포함하는 코스트 함수를 이용하여, 상기 입력된 영상을 스무딩하는 것을 특징으로 하는 영상 스무딩 방법을 수행할 수 있는 프로그램이 기록된 컴퓨터로 읽을 수 있는 기록매체.

발명의 설명

기 술 분 야

[0001] 본 발명은 영상 처리에 관한 것으로, 더욱 상세하게는 이진화된 영상을 스무딩하여 연속적인 영상을 생성하는 방법에 관한 것이다.

배 경 기 술

[0002] 가장 일반적인 영상 스무딩 알고리즘은 아래의 수학식 1에 제시된 코스트 함수를 이용하는 방법이다.

수학식 1

[0003] $$F=\min_{S,h,v}\left\{\sum_{p}\left(S_p-I_p\right)^2+\lambda C(h,v)+\beta\left[\left(\partial_x S_p-h_p\right)^2+\left(\partial_y S_p-v_p\right)^2\right]\right\}$$

[0004] 기존의 영상 스무딩 방법에서 가장 문제가 되는 것은, 영상이 급격히 변하는 에지 부분에서 스무딩 처리가 만족스럽지 못하다는 점이다.

[0005] 도 1에 나타난 원본 영상을 양자화한 영상에 대해, 위 수학식 1에 따라 스무딩한 결과를 도 2에 나타내었다. 도 1과 도 2에서는 기존의 영상 스무딩에서 문제가 되는 에지 부분을 중점적으로 나타내었다.

[0006] 도 2를 통해 알 수 있는 바와 같이, 기존 방식에 따르면, 양자화된 영상을 스무딩하는 경우, 에지 부분에서 스무딩이 제대로 이루어지지 않았음을 확인할 수 있는데, 이는 영상 전반의 품질을 떨어뜨리는 요인으로 작용하게 된다.

발명의 내용

해결하려는 과제

[0007] 본 발명은 상기와 같은 문제점을 해결하기 위하여 안출된 것으로서, 본 발명의 목적은, 에지 부분에서의 스무딩 처리가 우수한 영상 스무딩 방법 및 장치를 제공함에 있다.

과제의 해결 수단

[0008] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른, 영상 스무딩 방법은, 양자화된 영상을 입력받는 단계; 입력된 영상을 스무딩하는 단계; 및 스무딩된 영상을 출력하는 단계;를 포함하고, 상기 스무딩 단계는, 스무딩 영상의 2차 편미분이 반영된 항을 포함하는 코스트 함수를 이용하여, 상기 입력된 영상을 스무딩한다.

[0009] 그리고, 상기 코스트 함수는, 스무딩 영상의 x에 관한 2차 편미분 및 스무딩 영상의 y에 관한 2차 편미분이 반영된 항을 포함할 수 있다.

[0010] 또한, 상기 코스트 함수는, 스무딩 영상의 xy에 관한 편미분이 반영된 항을 더 포함할 수 있다.

[0011] 그리고, 상기 코스트 함수는, 스무딩 영상의 x에 관한 1차 편미분 및 스무딩 영상의 y에 관한 1차 편미분이 반영된 항을 더 포함할 수 있다.

[0012] 또한, 항들의 가중치들은, 사용자의 입력에 의해 설정될 수 있다.

[0013] 한편, 본 발명의 다른 실시예에 따른, 컴퓨터로 읽을 수 있는 기록매체에는, 양자화된 영상을 입력받는 단계; 입력된 영상을 스무딩하는 단계; 및 스무딩된 영상을 출력하는 단계;를 포함하고, 상기 스무딩 단계는, 스무딩 영상의 2차 편미분이 반영된 항을 포함하는 코스트 함수를 이용하여, 상기 입력된 영상을 스무딩하는 것을 특징으로 하는 영상 스무딩 방법을 수행할 수 있는 프로그램이 기록된다.

발명의 효과

[0014] 이상 설명한 바와 같이, 본 발명의 실시예들에 따르면, 영상 스무딩에 이용되는 코스트 함수에 다양한 항들을 적응적/선택적으로 추가하여 영상 스무딩을 수행할 수 있게 되는 바, 에지 부분에서의 스무딩 처리가 우수해져, 궁극적으로는 영상 전체의 품질을 향상시킬 수 있게 된다.

도면의 간단한 설명

[0015] 도 1은 원본 영상을 나타낸 도면,

도 2는, 도 1의 원본 영상을을 양자화한 영상에 대해, 기존의 방법으로 스무딩한 결과를 나타낸 도면,

도 3은 본 발명이 적용가능한 3D 콘텐츠 제작 방법의 설명에 제공되는 흐름도,

도 4는, 도 3의 부연 설명에 제공되는 도면,

도 5는, 도 1의 원본 영상을을 양자화한 영상에 대해, 제안된 방법으로 스무딩한 결과를 나타낸 도면,

도 6은 38개 파일에 대한 PSNR dB 측정 결과를 나타낸 도면,

도 7은 코스트 함수의 가중치 설정 방법의 설명에 제공되는 도면, 그리고,

도 8은 본 발명의 또 다른 실시예에 따른 3D 콘텐츠 제작 시스템의 블럭도이다.

발명을 실시하기 위한 구체적인 내용

[0016] 이하에서는 도면을 참조하여 본 발명을 보다 상세하게 설명한다.

[0017] **1. 3D 콘텐츠 제작**

[0018] 도 3은 본 발명이 적용가능한 3D 콘텐츠 제작 방법의 설명에 제공되는 흐름도이다. 도시된 3D 콘텐츠 제작 방법은, 2D 콘텐츠로부터 3D 콘텐츠를 생성하는 과정이다.

[0019] 도 3에 도시된 바와 같이, 먼저 2D 영상 콘텐츠를 입력받아(S105), 샷(Shot) 단위로 분할한다(S110). S110단계에 의해 2D 영상 콘텐츠는 다수의 샷들로 분할된다. 동일 샷에 포함된 프레임들은 동일/유사한 배경을 갖는다.

[0020] S110단계에서의 샷 분할은 전문 아티스트에 의한 수작업으로 수행될 수도 있고, 유사한 프레임들을 샷 단위로 구분하는 프로그램을 이용하여 자동으로 수행될 수도 있다.

[0021] 이후, 하나의 샷을 선정하고(S115), 선정된 샷에서 키-프레임을 선정한다(S120). S115단계에서의 샷 선정은, 2D 영상 콘텐츠에서 샷의 시간 순서에 따라 순차적으로 자동 선정(즉, 첫 번째 샷을 먼저 선정하고, 이후 다음 샷을 선정)할 수 있음은 물론, 전문 아티스트의 판단에 의해 선정 순서를 다른 순서로 정할 수도 있다. 이하에서는, 샷 선정이 샷의 시간 순서 따라 순차적으로 이루어지는 것을 상정하겠다.

[0022] 아울러, S120단계에서의 키-프레임 선정도 샷을 구성하는 프레임들 중 시간적으로 가장 앞선 프레임(즉, 첫 번째 프레임)을 키-프레임으로 자동 선정할 수 있음은 물론, 전문 아티스트의 판단에 의해 그 밖의 다른 프레임을

키-프레임으로 선정할 수도 있다. 이하에서는, 키-프레임이 샷의 첫 번째 프레임인 것을 상정하겠다.

[0023] 다음, S120단계에서 선정된 키-프레임에 대한 뎁스맵(Depth Map)을 생성한다(S125). S125단계에서의 뎁스맵 생성 역시 프로그램을 이용한 자동 생성은 물론, 전문 아티스트의 수작업에 의한 생성도 가능하다.

[0024] 이후, S125단계에서 생성된 뎁스맵의 등심선을 생성한다(S130). S130단계에서, 뎁스맵 등심선은 프로그램을 이용하여 뎁스맵을 뎁스에 따라, 다수의 양자화 단계들로 양자화 처리하여 생성한다.

[0025] 다음, S130단계에서 생성된 등심선을 다음-프레임에 매치시키다(S135). 움직임 객체가 있거나 화각이 이동하여, 다음-프레임이 이전-프레임인 키-프레임과 다른 부분이 있는 경우, 키-프레임의 등심선은 다음-프레임에 정확히 매치되지 않는다.

[0026] 이에 따라, 다음-프레임에 완전히 매치되도록, 등심선을 부분적으로 이동 처리(매치 무브)한다(S140). S140단계에서의 등심선 이동 처리 역시 프로그램을 이용하여 자동으로 수행된다. S140단계에 의해, 다음-프레임에 완전하게 매치된 등심선이 생성된다.

[0027] 이후, S140단계에서 매치 무브된 등심선을 기반으로, 다음-프레임의 뎁스맵을 생성한다(S145). S145단계에서 생성되는 다음-프레임의 뎁스맵은 양자화된 상태의 뎁스맵이다.

[0028] 이에, 뎁스맵을 보간하여, 다음-프레임의 뎁스맵을 완성한다(S150). S150단계에서의 뎁스맵 보간 처리 역시 프로그램을 이용하여 자동으로 수행된다.

[0029] S135단계 내지 S150단계는, 샷을 구성하는 모든 프레임들에 대해 완료될 때까지 반복된다(S155). 즉, 샷의 두 번째 프레임에 대한 뎁스맵이 완성되면, 두 번째 프레임의 등심선을 세 번째 프레임에 매치시키고(S135), 등심선 매치 무브를 수행한 후에(S140), 매치 무브된 등심선을 기반으로, 세 번째 프레임의 뎁스맵을 생성하고(S145), 보간을 통해 뎁스맵을 완성하게 되며(S150), 세 번째 프레임 이후의 프레임들에 대해서도 동일한 작업이 수행된다.

[0030] 샷을 구성하는 모든 프레임들에 대해 뎁스맵 생성이 완료되면(S155-Y), 두 번째 샷에 대해 S115단계 내지 S155단계를 반복하며, 두 번째 샷을 구성하는 모든 프레임들에 대해 뎁스맵 생성이 완료되면, 이후의 샷들에 대해서도 동일한 작업이 수행된다.

[0031] 2D 영상 콘텐츠를 구성하는 모든 샷들에 대해 위 절차가 완료되면(S160-Y), 지금까지 생성된 뎁스맵들을 이용하여, 2D 영상 콘텐츠를 구성하는 프레임들을 3D 변환처리한다(S165). 그리고, S165단계에서 변환처리 결과를 3D 영상 콘텐츠로 출력한다(S170).

[0032] 도 4는, 도 3의 부연 설명에 제공되는 도면이다. 도 4에는,

[0033] 1) S115단계에서 선정된 샷을 구성하는 프레임들 중 첫 번째 프레임을 키-프레임으로 선정하고(S120),

[0034] 2) 선정된 키-프레임에 대한 뎁스맵을 생성한 후에(S125),

[0035] 3) 뎁스맵의 등심선을 추출하고(S130),

[0036] 4) 추출한 등심선을 다음-프레임에 매치시켜, 매치 무브한 후(S135, S140),

[0037] 5) 매치 무브된 등심선을 기반으로, 다음-프레임의 뎁스맵을 생성하고(S145),

[0038] 6) 생성된 뎁스맵을 보간하여, 다음-프레임의 뎁스맵을 완성(S150)하는 과정이 도식적으로 나타나 있다.

[0039] **2. 뎁스맵 스무딩**

[0040] 등심선으로 표현된 양자화된 뎁스맵을 보간하여, 연속적인 뎁스맵을 획득하기 위해서는, 뎁스맵을 스무딩 처리하여야 한다. 뎁스맵 스무딩 방법에 대해, 이하에서 상세히 설명한다.

[0041] 본 발명의 실시예에 따른 뎁스맵 스무딩을 위한 코스트 함수는 아래의 수학식 2와 같다.

수학식 2

[0042]
$$F=\min_{S,h,v}\left\{\sum_{p}\left(S_p-I_p\right)^2+\lambda C(h,v)+\beta\left[\left(\partial_x S_p-h_{x,p}\right)^2+\left(\partial_y S_p-v_{y,p}\right)^2\right]+\gamma\left[\left(\partial_{xx} S_p-h_{xx,p}\right)^2+\left(\partial_{yy} S_p-v_{yy,p}\right)^2\right]+\alpha\left(|\nabla S|-|\nabla I|\right)\right\}$$
$$|\nabla S|=\left[\left(\partial_x S\right)^T\partial_x S+\left(\partial_x S\right)^T\partial_x S\right]$$
$$\alpha\left(|\nabla S|-|\nabla I|\right)=\alpha\left\{\left[\left(\partial_x S\right)^T\partial_x S+\left(\partial_x S\right)^T\partial_x S\right]-\left[\left(\partial_x I\right)^T\partial_x I+\left(\partial_x I\right)^T\partial_x I\right]\right\}$$

[0043] 위 수학식 2를 통해 알 수 있는 바와 같이, 본 발명의 일 실시예에 따른 뎁스맵 스무딩을 위한 코스트 함수는, 스무딩 영상(S_p)의 x에 관한 1차 편미분과 스무딩 영상(S_p)의 y에 관한 1차 편미분이 반영된 항[가중치가 β인 항]을 포함한다는 점에서, 수학식 1에 표시된 기존의 코스트 함수와 동일하다.

[0044] 하지만, 본 발명의 일 실시예에 따른 뎁스맵 스무딩을 위한 코스트 함수는, 가중치가 α인 항과 가중치가 γ인 항을 더 포함한다는 점에서, 수학식 1에 표시된 기존의 코스트 함수와 차이가 있다.

[0045] 가중치가 α인 항과 가중치가 γ인 항은 스무딩 영상(S_p)의 2차 편미분이 반영되어 있다는 점에서, 스무딩 영상(S_p)의 1차 편미분이 반영된 가중치가 β인 항과 차이가 있다.

[0046] 구체적으로, 가중치가 α인 항과 가중치가 γ인 항은 스무딩 영상(S_p)의 x에 관한 2차 편미분 및 스무딩 영상(S_p)의 y에 관한 2차 편미분이 반영되어 있다.

[0047] 한편, 가중치가 α인 항은 아래의 수학식 3과 같이 변경이 가능하다. 뿐만 아니라, 가중치가 γ인 항 없이 가중치가 α인 항만 코스트 함수에 포함되도록 구현가능하다.

[0048] 반대로, 가중치가 α인 항 없이 가중치가 γ인 항만이 코스트 함수에 포함되도록 구현가능하는 것도 가능함은 물론이다.

수학식 3

[0049]
$$F=\min_{S,h,v}\left\{\sum_{p}\left(S_p-I_p\right)^2+\lambda C(h,v)+\beta\left[\left(\partial_x S_p-h_p\right)^2+\left(\partial_y S_p-v_p\right)^2\right]+\alpha\left(|S_*|-|\nabla I_*|\right)\right\}$$
$$S_*\in\left\{\partial_x S,\partial_y S,\partial_{xx} S,\partial_{yy} S,\partial_{xy} S\right\}$$
$$\frac{\partial\alpha\left(|\nabla S|-|\nabla I|\right)}{\partial S}=2\alpha\left\{\left(\partial_x\right)^T\partial_x S+\left(\partial_y\right)^T\partial_y S+\left(\partial_{xx}\right)^T\partial_{xx} S+\left(\partial_{yy}\right)^T\partial_{yy} S+\left(\partial_{xy}\right)^T\partial_{xy} S\right\}$$

[0050] 위 수학식 3을 통해 알 수 있는 바와 같이, 본 발명의 다른 실시예에 따른 뎁스맵 스무딩을 위한 코스트 함수는, 스무딩 영상(S_p)의 xy에 관한 편미분이 반영된 항[가중치가 α인 항]을 포함한다.

[0051] 수학식 3에 제시된 코스트 함수에 따라 산출되는 스무딩 영상(S_p)은 아래의 수학식 4와 같다.

수학식 4

[0052]
$$S_p=\frac{I_p+h_p\partial_x+v_p\partial_y+h_{xx,p}\partial_{xx}+v_{yy,p}\partial_{yy}}{\left(I+\beta\partial_x^T\partial_x+\beta\partial_y^T\partial_y+\gamma\partial_{xx}^T\partial_{xx}+\gamma\partial_{yy}^T\partial_{yy}+\alpha\left[\left(\partial_x\right)^T\partial_x+\left(\partial_y\right)^T\partial_y+\left(\partial_{xx}\right)^T\partial_{xx}+\left(\partial_{yy}\right)^T\partial_{yy}+\left(\partial_{xy}\right)^T\partial_{xy}\right]\right)}$$

[0053] **3. 성능 비교**

[0054] 기존 방식과의 성능 비교를 위해, 도 1에 나타난 원본 영상(뎁스맵)을 양자화한 영상(뎁스맵)에 대해, 본 발명의 실시예에 따라 스무딩한 결과를 도 5에 나타내었다. 도 2와 도 5를 비교하면, 기존의 스무딩 방법에 비해

에지 부분에서의 스무딩 결과가 보다 우수하게 나타났음을 확인할 수 있다.

[0055] 38개 파일에 대해 기존 방식과 본 발명의 실시예에 따른 스무딩 방법에 대한 PSNR dB 측정 결과를 도 6에 나타내었다. 도 6을 통해서도, 본 발명의 실시예에 따른 스무딩 방법의 우수성을 확인할 수 있다.

[0056] **4. 영상 스무딩 방법 및 장치**

[0057] 지금까지, 뎁스맵 스무딩을 위한 코스트 함수와 이에 따른 스무딩 영상(S_p)에 대해 상세히 설명하였다.

[0058] 코스트 함수에 나타나는 각 항들의 가중치들(α, β, γ)은 사용자에 의해 설정 가능하며, 뎁스맵의 사양/특성에 따라 자동으로 설정가능하다. 따라서, 도 7에 도시된 바와 같이, 사용자가 가중치들을 설정한 경우 그에 따라 뎁스맵 스무딩이 수행되는 반면, 사용자가 가중치들을 설정하지 않으면 사전 설정된 가중치들에 따라 뎁스맵 스무딩이 이루어진다.

[0059] 도 8은 본 발명의 또 다른 실시예에 따른 3D 콘텐츠 제작 시스템의 블럭도이다. 본 발명의 실시예에 따른 3D 콘텐츠 제작 시스템(200)은, 도 8에 도시된 바와 같이, 2D 영상 입력부(210), 뎁스맵 생성부(220), 3D 변환부(230) 및 3D 영상 출력부(240)를 포함한다.

[0060] 뎁스맵 생성부(220)는 키-프레임 뎁스맵을 이용하여, 2D 영상 입력부(210)를 통해 입력되는 2D 영상 프레임들의 뎁스맵을 생성한다. 키-프레임 뎁스맵은, 2D 영상에서 샷 마다 하나씩 생성되는데, 프로그램에 의한 자동 생성은 물론, 전문 아티스트에 의한 수동 생성 모두가 가능하다.

[0061] 뎁스맵 생성부(220)는 키-프레임 뎁스맵의 등심선을 추출하고, 추출한 등심선을 다음-프레임에 매치시켜 매치 무브한 후, 매치 무브된 등심선을 기반으로 다음-프레임의 뎁스맵을 생성하고 보간하여 다음-프레임의 뎁스맵을 완성하는 절차에 의해, 2D 영상 프레임들에 대한 뎁스맵들을 생성한다. 이 과정에서, 뎁스맵 생성부(220)는 위에서 제시한 코스트 함수를 이용하여 뎁스맵 스무딩을 수행한다.

[0062] 3D 변환부(230)는 뎁스맵 생성부(220)에서 생성된 뎁스맵들을 이용하여, 2D 영상 프레임들을 3D 변환처리한다. 그리고, 3D 영상 출력부(240)는 3D 변환부(230)에 의한 변환처리 결과를 3D 영상으로 출력한다.

[0063] **5. 변형예**

[0064] 지금까지, 뎁스맵 스무딩 방법과 이를 적용한 컨텐츠 제작 방법에 대해 바람직한 실시예들을 들어 상세히 설명하였다.

[0065] 위 실시예에서 언급한 뎁스맵은 영상의 일종이다. 뎁스맵(뎁스 영상) 이외의 다른 영상에 대해서도 본 발명의 기술적 사상이 적용될 수 있음은 물론이며, 다른 영상에는 의료 영상도 포함될 수 있음은 물론이다.

[0066] 또한, 이상에서는 본 발명의 바람직한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

부호의 설명

[0067] 200 : 3D 콘텐츠 제작 시스템

210 : 2D 영상 입력부

220 : 뎁스맵 생성부

230 : 3D 변환부

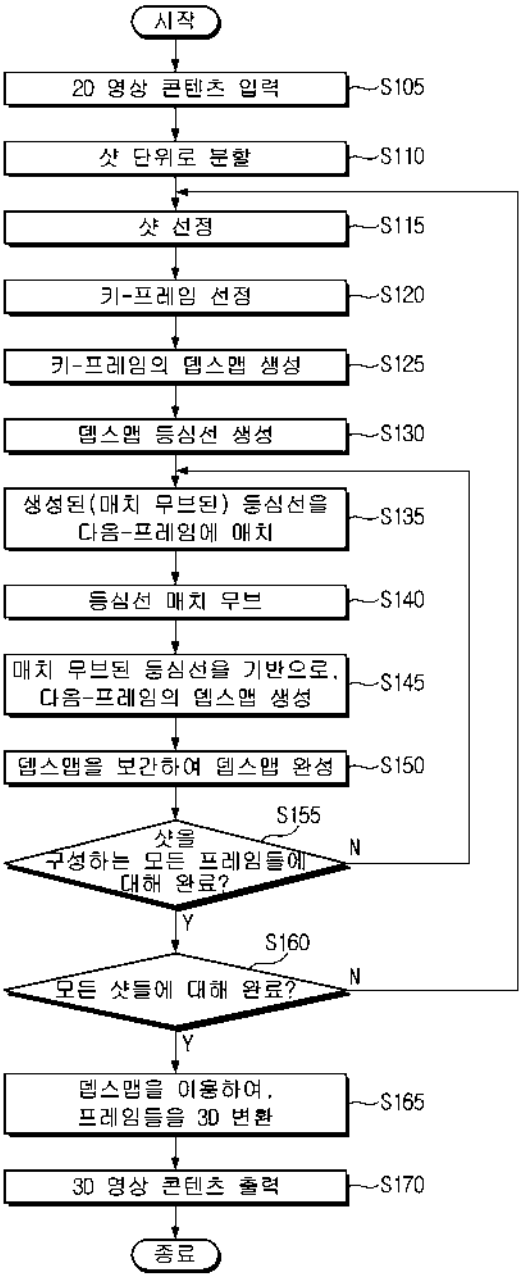
240 : 3D 영상 출력부

도면

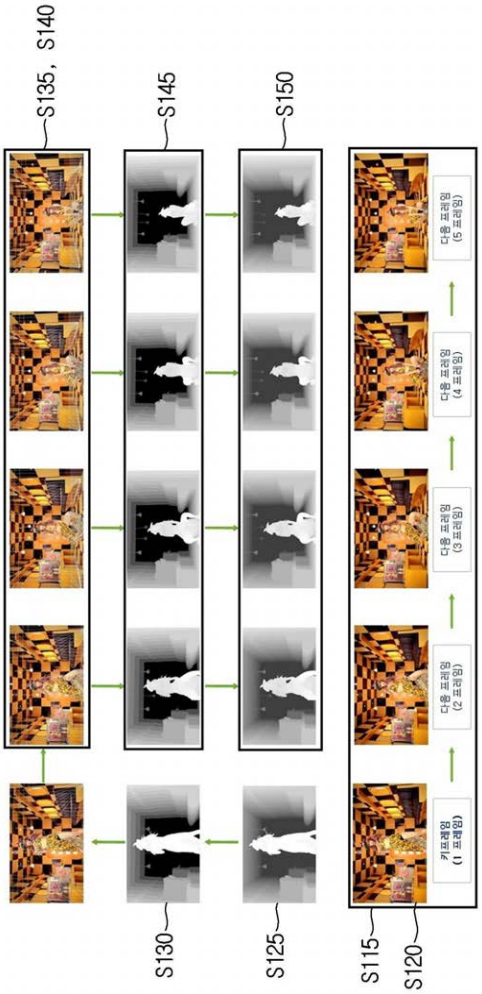
도면1

도면2

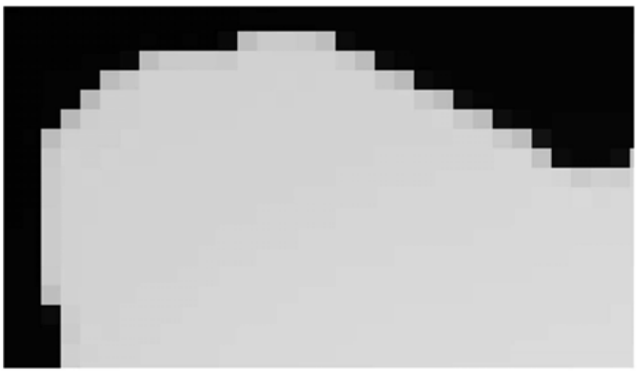
도면3


시작
2D 영상 콘텐츠 입력
S105
샷 단위로 분할
S110
샷 선정
S115
키-프레임 선정
S120
키-프레임의 뎁스맵 생성
S125
뎁스맵 등심선 생성
S130
생성된(매치 무브된) 등심선을 다음-프레임에 매치
S135
등심선 매치 무브
S140
매치 무브된 등심선을 기반으로, 다음-프레임의 뎁스맵 생성
S145
뎁스맵을 보간하여 뎁스맵 완성
S150
S155
샷을 구성하는 모든 프레임들에 대해 완료?
N
Y
S160
모든 샷들에 대해 완료?
N
Y
뎁스맵을 이용하여, 프레임들을 3D 변환
S165
3D 영상 콘텐츠 출력
S170
종료

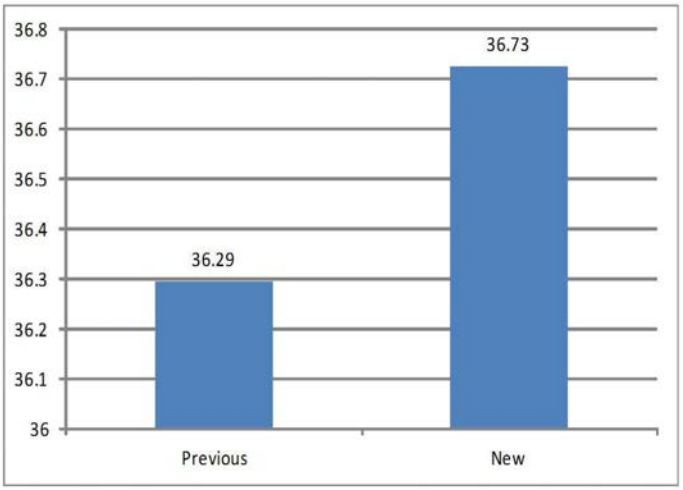
도면4


S135, S140
S145
S150
S130
S125
S115
S120
키프레임
(1 프레임)
다음 프레임
(2 프레임)
다음 프레임
(3 프레임)
다음 프레임
(4 프레임)
다음 프레임
(5 프레임)

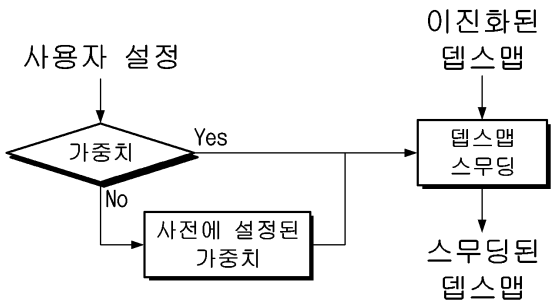
도면5



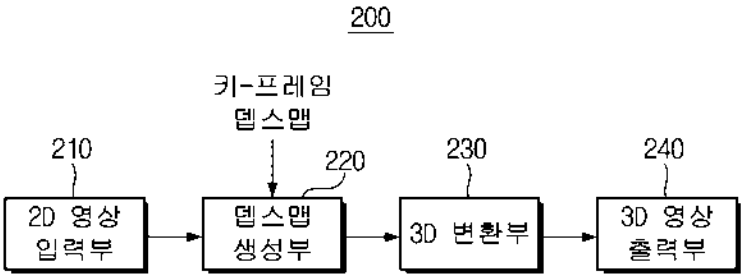
도면6


36.8
36.7
36.6
36.5
36.4
36.3
36.2
36.1
36
36.73
36.29
Previous
New

도면7


사용자 설정
이진화된
뎁스맵
가중치
Yes
No
뎁스맵
스무딩
사전에 설정된
가중치
스무딩된
뎁스맵

도면8


200
키-프레임
뎁스맵
210
220
230
240
2D 영상
입력부
뎁스맵
생성부
3D 변환부
3D 영상
출력부

■ 기술명 : 지능형 카메라 용 영상 신호처리 기술
(Image Signal Processing Technology for Intelligent Camera)

산업기술분류	전기・전자 / 영상・음향기기 / 카메라 및 캠코더(200806)
Key-word(국문)	영상 흔들림 보정, 노이즈 제거, 컬러 보정, 롤링셔터 보정, 비디오 컬러 변환
Key-word(영문)	Digital Image Stabilizer (DIS), Noise Reduction (NR), Color Correction, Rolling Shutter Correction, Video Color Transfer

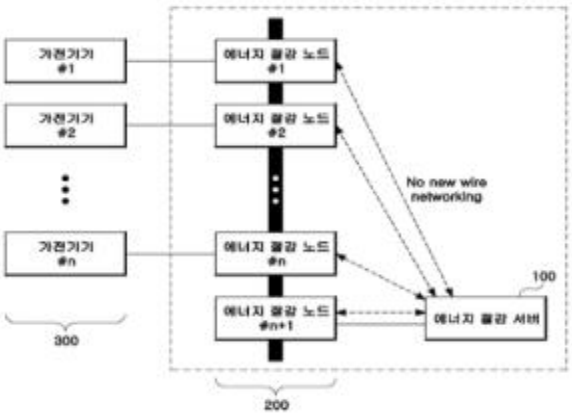
■ 기술의 개요

○ (배경) 많은 전자기기에 카메라가 탑재되고 있지만 아직 사용자의 니즈를 충족하기에는 부족하며, 흔들림 보정, 노이즈 제거, 컬러 보정, 롤링셔터 보정 등의 기능이 포함되어 사용자 만족도를 높이는 카메라 기술 필요

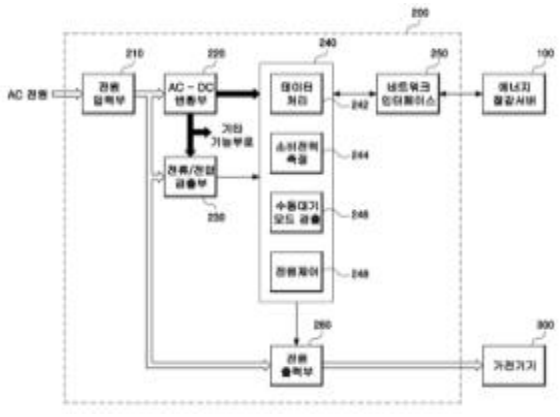
○ (개요) 사용자가 카메라를 사용할 때 손떨림으로 인한 영상 흔들림을 보정하며, 카메라의 빠른 움직임에 의해 발생할 수 있는 롤링셔터 왜곡도 동시에 제거 기술

- 저조도에서 효과적으로 노이즈 제거 및 실제 색에 가장 가깝거나 사용자가 원하는 색으로 컬러 보정 가능

< 기술 개요도 >



[에너지 절감 시스템 구조]



[에너지 절감 시스템 클라이언트 구조]

■ 기술의 구현수준(TRL)

■ 기술의 장점(경쟁기술과의 차별성)

○ 실시간 흔들림 보정 가능(30fps)
- Semi-global 방식, Key-frame 방식 적용
- 롤링셔터 왜곡을 동시에 보정 가능

○ 저전력 노이즈 제거 가능
- 기존 기술 대비 5%이상 최대신호대잡음비(PSNR) 향상, 30%이상 전력 절감

○ 비선형적 컬러 보정 가능
- 카메라의 비선형적 특성을 고려한 효과적인 컬러 보정
- 카메라 종류/조명 변화 및 컬러 리터치 등의 다양한 컬러 변환 표현

■ 활용범위 및 응용분야

<table>
<tr><td> 
</td><td> 
</td></tr>
<tr><td>[지능형 카메라 시스템 - 화질 개선]</td><td>[모바일 카메라 시스템 - 롤링셔터 왜곡 보정]</td></tr>
<tr><td> 
</td><td>  </td></tr>
<tr><td>[자동차 카메라 시스템 - 흔들림 보정]</td><td>[로봇 어플리케이션 - 객체 인식 성능 향상]</td></tr>
</table>

■ 지식재산권 현황

구분	발명의 명칭	출원번호 (출원일)	등록번호 (등록일)
특허	동영상의 일괄 컬러 변환 방법 및 그 기록매체	2013-0098569 (2013.08.20)	10-1492060 (2015.02.04)
특허	대응 영상의 컬러 보정 방법 및 그 기록매체	2011-0052031 (2011.05.31)	10-1227936 (2013.01.24)
특허	이미지 센서 노이즈 모델 기반 필터링 방법 및 이를 적용한 영상 시스템	2014-0069494 (2014.06.09)	10-1585963 (2016.01.11)

(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2013년01월30일
(11) 등록번호 10-1227936
(24) 등록일자 2013년01월24일

(51) 국제특허분류(Int. Cl.)
H04N 9/07 (2006.01) *H04N 9/67* (2006.01)
(21) 출원번호 10-2011-0052031
(22) 출원일자 2011년05월31일
심사청구일자 2011년05월31일
(65) 공개번호 10-2012-0133417
(43) 공개일자 2012년12월11일
(56) 선행기술조사문헌
JP2005292132 A
KR1020100076351 A
JP2011048507 A
KR100668073 B1

(73) 특허권자
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
황영배
서울특별시 서초구 방배선행길 1, 우성 아파트 106동 607호 (방배동)
김제우
경기도 성남시 분당구 내정로 55, 우성아파트 320동 603호 (정자동, 상록마을)
최병호
경기도 용인시 수지구 대지로 27, 103동 1306호 (죽전동, 대주한신아파트)
(74) 대리인
노철호, 남충우

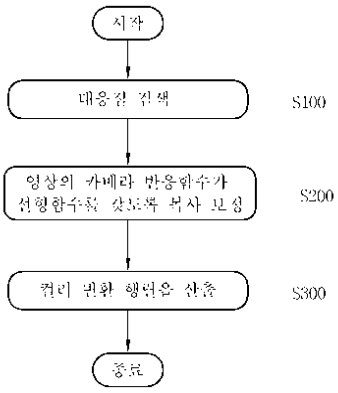
전체 청구항 수 : 총 8 항 심사관 : 신재철

(54) 발명의 명칭 **대응 영상의 컬러 보정 방법 및 그 기록매체**

(57) 요 약

대응 영상의 컬러 보정 방법 및 그 기록매체가 개시된다. 본 발명의 일 실시예에 따른 대응 영상의 컬러 보정 방법은 대응 영상의 대응점을 검색하는 제1단계; 상기 대응 영상의 카메라 반응함수가 선형함수를 갖도록 복사 보정(radiometric calibration)하는 제2단계; 및 컬러 변환 행렬을 산출하여 컬러를 보정 하는 제3단계를 포함하는 것을 특징으로 한다.

대 표 도 - 도1



이 발명을 지원한 국가연구개발사업

과제고유번호	10037388
부처명	지식경제부
연구사업명	산업원천기술개발사업
연구과제명	양안식 3DTV 방송용 카메라 개발
주관기관	전자부품연구원
연구기간	2010.06.01 ~ 2011.06.30

특허청구의 범위

청구항 1

동일한 물체나 장면을 포함하는 2 이상의 영상인 대응 영상에서 동일한 물리적 위치에 있으면서 동일한 특징(feature)을 갖는 대응점을 검색하는 제1단계(S100);

상기 대응 영상의 카메라 반응함수가 선형함수를 갖도록 복사 보정(radiometric calibration)하는 제2단계(S200); 및

컬러 변환 행렬을 산출하여 컬러를 보정 하는 제3단계(S300)를 포함하는 것을 특징으로 하는 대응 영상의 컬러 보정 방법.

청구항 2

제1항에 있어서,

상기 제1단계(S100)는,

코너점을 기반으로 하는 특징량 검출 방법을 이용하여 특징량을 검출하는 단계(S110);

상기 특징량을 매칭하여 제1매칭점(N)을 도출하는 단계(S120);

상기 제1매칭점(N)에서 잘못된 매칭(outlier)을 제거하여 제2매칭점(M)을 도출하는 단계(S130); 및

상기 특징량 검출의 적합성을 판단하는 단계(S140)를 포함하는 것을 특징으로 하는 대응 영상의 컬러 보정 방법.

청구항 3

제2항에 있어서,

상기 특징량 검출의 적합성을 판단하는 단계(S140)는,

상기 제2매칭점(M)의 수를 상기 제1매칭점(N)의 수로 나눈 값이 0.7 미만인 경우에,

불변특징량을 기반으로 하는 특징량 검출 방법을 이용하여 특징량을 재검출하는 단계(S142); 및

상기 특징량을 매칭하고, 잘못된 매칭을 제거하여 제2매칭점(M2)을 재도출하는 단계(S144)를 포함하는 것을 특징으로 하는 대응 영상의 컬러 보정 방법.

청구항 4

제3항에 있어서,

상기 제1단계(S100)는 대응점 결정 적합성을 판단하는 단계(S150)를 더 포함하고,

상기 대응점 결정 적합성을 판단하는 단계(S150)는,

제2매칭점(M 또는 M2)의 R,G,B 컬러 분포값이 전체 영상의 R,G,B 컬러분포 값의 20% 이상 또는 미만인지 여부를 판단하는 것을 특징으로 하는 대응 영상의 컬러 보정 방법.

청구항 5

제4항에 있어서,

상기 제2매칭점(M 또는 M2)의 R,G,B 컬러 분포값이 전체 영상의 R,G,B 컬러분포 값의 20% 이상인 경우에는,

상기 제2매칭점(M 또는 M2)을 대응점으로 결정하는 단계(S160)를 포함하고,

상기 제2매칭점(M 또는 M2)의 R,G,B 컬러 분포값이 전체 영상의 R,G,B 컬러분포 값의 20% 미만인 경우에는,

밀집 스테레오 매칭을 이용하여 매칭점을 재도출하는 단계(S170); 및

재도출된 상기 매칭점을 대응점으로 결정하는 단계(S172)를 더 포함하는 것을 특징으로 하는 대응 영상의 컬러 보정 방법.

청구항 6

제1항에 있어서,

상기 제2단계(S200)는,

카메라의 반응 함수를 산출하는 단계(S210);

상기 반응 함수의 역함수를 산출하는 단계(S220); 및

산출된 상기 역함수를 영상에 적용하여 선형인 카메라 반응 함수를 갖는 영상으로 변환하는 단계(S230)를 포함하는 것을 특징으로 하는 대응 영상의 컬러 보정 방법.

청구항 7

제1항에 있어서,

상기 제3단계(S300)는,

하기 수학식 1에 따른 최소 제곱(least square) 방법으로 컬러 변환 행렬을 산출하는 방법, 하기 수학식 2에 따른 3 X 3 RGB 컬러 변환 행렬을 산출하는 방법, 또는 하기 수학식 3에 따른 비선형 다항식을 갖는 컬러 변환 행렬을 산출하는 방법 중에서 어느 하나인 것을 특징으로 하는 대응 영상의 컬러 보정 방법.

[수학식 1]

$$\sum_{x=1}^{NS}(\overrightarrow{Ic_s}-(a_c\overrightarrow{Tc_s}+b_c))^2,\ c\in R,G,B$$

(여기에서, Ic_s는 목표로 하는 영상의 RGB 컬러, Tc_s는 변환하려고 하는 입력 영상의 RGB 컬러, a_c 및 b_c는 두 컬러를 일치시키기 위한 계수임.)

[수학식 2]

$$\sum_{s=1}^{NS}(\overrightarrow{I_s}-T_{RGB}\cdot\overrightarrow{T_s})^2$$

(여기에서, I_s는 입력 영상 내의 S번째 대응점의 R, G, B 벡터, T_{RGB}는 컬러 변환 행렬의 각 성분, T_s는 목표 영상 내의 S번째 대응점의 R, G, B 벡터임.)

[수학식 3]

$$\sum_{k=1}^{D}\left(t_{rc_k}Ir_s^k + t_{gc_k}Ig_s^k + t_{bc_k}Ib_s^k\right) + t_{c0} \simeq Tc_s$$

(여기에서, D는 비선형적 변환에서의 차수, $t_{rck}, t_{gck}, t_{bck}$는 각 차수의 각 컬러에 대항하는 계수, t_{c0}는 0차에서의 계수, Ir^k_s, Ig^k_s, Ib^k_s는 입력영상의 각 차수의 R,G,B 값, Tc_s는 목표 영상의 한 채널(R,G,B중 하나)의 값임.)

청구항 8

제1항 내지 제7항 중 어느 한 항에 기재된 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

명 세 서

기 술 분 야

[0001] 본 발명은 대응 영상의 컬러를 보정하는 방법 및 상기 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록 매체에 관한 것이다.

배 경 기 술

[0002] 일반적으로, 3차원 영상 정보 획득에 사용되는 방법 중 하나는 여러 방향의 시야각(view)으로부터 획득한 2차원의 영상정보를 조합하는 것이다. 이 때, 동일한 물체나 장면을 포함하는 다수의 영상에서 같은 물리적 위치에 있으면서 동일한 특징을 갖는 대응점들을 찾는 일은 매우 중요하다.

[0003] 특히, 다수의 카메라를 사용하여 획득한 영상의 경우에는 획득한 영상의 컬러 값들이 동일하지 않아 발생하는 컬러 차를 보정하는 기술이 요구되고 있다. 이러한 컬러 차를 보정하기 위한 방법으로 종래에는 맥베스 컬러차트(Macbeth color checker chart)를 이용하는 방법이 있다.

[0004] 한국공개특허 제10-2010-0035497호에서는 컬러 보정을 하기 위해서 맥베스 컬러차트를 각 카메라에서 보이도록 영상을 획득한 후에 각 컬러 패치의 색을 픽셀 평균을 내서 구한 후에, 상기 패치의 색을 동일하게 만드는 단일 변환 행렬을 구함으로써 컬러 보정을 하는 방법을 개시하고 있다. 그러나, 상술한 방법에서는 다음과 같은 문제점이 있다.

[0005] 첫째, 맥베스 컬러차트를 이용한 컬러 보정 방법은 카메라의 하드웨어 적인 특성만을 고려하므로 실제 촬영 시 발생할 수 있는 다른 조건들(예를 들면, 조명, 날씨 변화)을 반영할 수 없어 컬러 보정을 수시로 해주어야 한다. 둘째, 카메라의 피라미터가 자동적으로 변화하는 경우에서는 적용될 수 없다. 셋째, 맥베스 컬러차트 영상은 균일하므로 자동적인 대응점 매칭이 어렵다. 넷째, 동일 모델의 카메라라 하더라도 카메라의 피라미터마다 고유의 비선형 카메라 반응 함수(Camera response function)이 존재하므로, 변환 행렬만을 이용해서는 컬러 보정이 정확히 이루어지지 않는다.

[0006] 따라서, 종래 맥베스 컬러차트를 이용한 컬러 보정 방법들의 상기 문제점들을 해결하기 위한 방안이 모색되고 있다.

발명의 내용

해결하려는 과제

[0007] 본 발명의 실시예들은 맥베스 컬러차트를 이용하지 않아 자동적인 대응점 매칭이 가능하고, 다양한 조건 하에서도 적용 가능한 대응 영상의 컬러 보정 방법 및 그 기록매체를 제공하고자 한다.

과제의 해결 수단

[0008] 본 발명의 일 측면에 따르면, 대응 영상의 대응점을 검색하는 제1단계; 상기 대응 영상의 카메라 반응 함수가 선형함수를 갖도록 복사 보정(radiometric calibration)하는 제2단계; 및 컬러 변환 행렬을 산출하여 컬

러를 보정 하는 제3단계를 포함하는 것을 특징으로 하는 대응 영상의 컬러 보정 방법이 제공될 수 있다.

[0009] 또한, 상기 제1단계는, 코너점을 기반으로 하는 특징량 검출 방법을 이용하여 특징량을 검출하는 단계; 상기 특징량을 매칭하여 제1매칭점(N)을 도출하는 단계; 상기 제1매칭점(N)에서 잘못된 매칭(outlier)을 제거하여 제2매칭점(M)을 도출하는 단계; 및 상기 특징량 검출의 적합성을 판단하는 단계를 포함하는 것을 특징으로 할 수 있다.

[0010] 한편, 상기 특징량 검출의 적합성을 판단하는 단계는, 상기 제2매칭점(M)의 수를 상기 제1매칭점(N)의 수로 나눈 값이 0.7 미만인 경우에는, 불변특징량을 기반으로 하는 특징량 검출 방법을 이용하여 특징량을 재검출하는 단계; 및 상기 특징량을 매칭하고, 잘못된 매칭을 제거하여 제2매칭점(M2)을 재도출하는 단계를 포함하는 것을 특징으로 할 수 있다.

[0011] 또한, 상기 제1단계는 대응점 결정 적합성을 판단하는 단계를 더 포함하고, 상기 대응점 결정 적합성을 판단하는 단계는, 제2매칭점(M 또는 M2)의 R,G,B 컬러 분포값이 전체 영상의 R,G,B 컬러분포 값의 20% 이상 또는 미만인지 여부를 판단하는 것을 특징으로 할 수 있다.

[0012] 이 때, 상기 제2매칭점(M 또는 M2)의 R,G,B 컬러 분포값이 전체 영상의 R,G,B 컬러분포 값의 20% 이상인 경우에는, 상기 제2매칭점(M 또는 M2)을 대응점으로 결정하는 단계를 포함하고,

[0013] 상기 제2매칭점(M 또는 M2)의 R,G,B 컬러 분포값이 전체 영상의 R,G,B 컬러분포 값의 20% 미만인 경우에는, 밀집 스테레오 매칭을 이용하여 매칭점을 재도출하는 단계; 및 재도출된 상기 매칭점을 대응점으로 결정하는 단계를 더 포함하는 것을 특징으로 할 수 있다.

[0014] 한편, 상기 제2단계는, 카메라의 반응 함수를 산출하는 단계; 상기 반응 함수의 역함수를 산출하는 단계; 및 산출된 상기 역함수를 영상에 적용하여 선형인 카메라 반응 함수를 갖는 영상으로 변환하는 단계를 포함하는 것을 특징으로 할 수 있다.

[0015] 또한, 상기 제3단계는, 하기 수학식 1에 따른 최소 제곱(least square) 방법으로 컬러 변환 행렬을 산출하는 방법, 하기 수학식 2에 따른 3 X 3 RGB 컬러 변환 행렬을 산출하는 방법, 또는 하기 수학식 3에 따른 비선형 다항식을 갖는 컬러 변환 행렬을 산출하는 방법 중에서 어느 하나인 것을 특징으로 할 수 있다.

[0016] [수학식 1]

[0017]
$$\sum_{x=1}^{NS}(\overrightarrow{Ic_s}-(a_c\overrightarrow{Tc_s}+b_c))^2,\ c\in R,G,B$$

[0018] (여기에서, Ic_s는 목표로 하는 영상의 RGB 컬러, Tc_s는 변환하려고 하는 입력 영상의 RGB 컬러, a_c 및 b_c는 두 컬러를 일치시키기 위한 계수임.)

[0019] [수학식 2]

[0020]
$$\sum_{s=1}^{NS}(\overrightarrow{I_s}-T_{RGB}\cdot\overrightarrow{T_s})^2$$

[0021] (여기에서, I_s는 입력 영상 내의 S번째 대응점의 R, G, B 벡터, T_{RGB}는 컬러 변환 행렬의 각 성분, T_s는 목표 영상 내의 S번째 대응점의 R, G, B 벡터임.)

[0022] [수학식 3]

[0023]
$$\sum_{k=1}^{D}\left(t_{rc_k}Ir_s^k+t_{gc_k}Ig_s^k+t_{bc_k}Ib_s^k\right)+t_{c0}\simeq Tc_s$$

[0024] (여기에서, D는 비선형적 변환에서의 차수, t_{rck}, t_{gck}, t_{bck}는 각 차수의 각 컬러에 대항하는 계수, t_{c0}는 0차에서의 계수, Ir^k_s, Ig^k_s, Ib^k_s는 입력영상의 각 차수의 R,G,B 값, Tc_s는 목표 영상의 한 채널(R,G,B중 하나)의 값임.)

[0025] 본 발명의 다른 측면에 따르면, 상기 기재된 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체가 제공될 수 있다.

발명의 효과

[0026] 본 발명의 실시예들은 맥베스 컬러차트를 사용하지 않음으로써, 자동적인 대응점 매칭이 가능하고 카메라 파라미터가 자동적으로 변하거나 조명 또는 날씨 변화와 같은 다양한 조건하에서도 컬러 보정을 가능하게 할 수 있다.

[0027] 또한, 대응 영상의 카메라 반응함수가 선형함수를 갖도록 복사 보정함으로써, 비선형 카메라 반응 함수에 의한 오차를 제거할 수 있다.

[0028] 또한, 컬러 변환 행렬을 3×3 RGB 컬러 변환 행렬뿐만이 아니라, 최소 제곱 방법 또는 비선형적 다항식을 이용하여 산출 가능하다.

도면의 간단한 설명

[0029] 도 1은 본 발명의 일 실시예에 따른 대응 영상의 컬러 보정 방법의 흐름도이다.

도 2는 도 1의 대응점 검색 단계를 세부적으로 나타내는 흐름도이다.

도 3은 코너점을 기반으로 하는 특징량 검출 방법을 이용하여 특징량을 검출한 결과를 나타내는 도면이다.

도 4는 도3에서 검출된 코너점을 탬플릿 매칭시킨 결과를 나타내는 도면이다.

도 5는 기술자 기반의 특징량 매칭 결과를 나타내는 도면이다.

도 6은 불변특징량을 기반으로 하는 특징량 검출 방법을 이용하여 특징량을 검출한 결과를 나타내는 도면이다.

도 7은 도 1의 복사 보정 단계를 세부적으로 나타내는 흐름도이다.

도 8은 다수 카메라의 반응 함수를 나타내는 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0030] 이하, 첨부된 도면을 참조하여 본 발명의 실시예들에 대하여 상세히 설명하도록 한다.

[0031] 도 1은 본 발명의 일 실시예에 따른 대응 영상의 컬러 보정 방법(S)의 흐름도이다. 도 1을 참조하면, 대응 영상의 컬러 보정 방법(S)은 대응 영상의 대응점을 검색하는 제1단계(S100), 상기 대응 영상의 카메라 반응함수가 선형함수를 갖도록 복사 보정(radiometric calibration)하는 제2단계(S200) 및 컬러 변환 행렬을 산출하여 컬러를 보정 하는 제3단계(S300)를 포함한다. 이하에서는 각 단계에 대하여 상세히 설명하도록 한다.

[0032] 도 2는 도 1의 대응점 검색 단계(S100)을 세부적으로 나타내는 흐름도이다.

[0033] 여기에서 대응점이란 동일한 물체나 장면을 포함하는 다수의 영상에서 같은 물리적 위치에 있으면서 동일한 특징을 갖는 점들을 의미한다. 대응점을 검색하는 방법은 다양하게 존재할 수 있으나, 본 발명의 일 실시예에서는 대응 영상의 특징량을 추출하여 매칭시키는 방법을 주된 방법으로 하고, 모든 픽셀에 대하여 밀집 스테레오 매칭시키는 방법을 부수적 방법으로 하였다. 이에 대해서는 후술하도록 한다.

[0034] 대응점 검색 단계(S100)는 코너점을 이용하여 특징량을 검출하는 단계(S110)와, 상기 특징량을 매칭하여 제1매칭점(N)을 도출하는 단계(S120)와, 상기 제1매칭점(N)에서 잘못된 매칭(아웃라이어, outlier)을 제거하여 제2매칭점(M)을 도출하는 단계 및 상기 특징량 검출의 적합성을 판단하는 단계(S140)를 포함할 수 있다.

[0035] 여기에서 특징량은 영상에서 특징(feature)이 되는 점(point)들을 의미하며, 대체로 영상에서의 코너점 또는 경계선에서 검출될 수 있다. 특징량을 검출하는 방법은 다양하게 존재할 수 있으나, 본 발명의 일 실시예에서는 코너점을 기반으로 하는 특징량 검출 방법을 주된 방법으로 하고, 불변 특징량을 기반으로 하는 특징량 검출 방법을 부수적 방법으로 하였다.

[0036] 본 발명의 일 실시예에서는 대응 영상의 대응점을 검색하기 위하여 코너점을 기반으로 하는 특징량 검출 방법을 우선적으로 수행할 수 있다.

[0037] 상기 코너점을 기반으로 하는 특징량 검출 방법은 소위 해리스 코너 검출기라 불리는 것으로, 코너점이 양방향으로 곡률이 높다는 점에 착안하여 고유값과 코너응답함수를 이용하여 회전에 불변하는 특징을 찾는 검출 방법을 의미한다.

[0038] 우선, 코너점을 이용하여 특징량을 검출한다. 코너점(Harris corner)은 하기 [식1]과 같이 계산될 수 있다.

[0039] [식1]

[0040] $$A=\sum_{u}\sum_{v}w(u,v)\begin{pmatrix} I_x^2 & I_xI_y \\ I_xI_y & I_y^2 \end{pmatrix}=\begin{bmatrix} <I_x^2> & <I_xI_y> \\ <I_xI_y> & <I_y^2> \end{bmatrix}$$

[0041] 여기에서, A는 소위 Harris matrix라고 불리며, 코너점을 구하기 위해 계산되는 매트릭스를 의미한다. w(u,v)는 가우시안 웨이트를 위한 가중치로, u 및 v는 A라는 매트릭스를 구하기 위해서 더해지는 마스크의 위치를 의미한다. I는 영상에서의 한 포인트로, I_x는 x 방향으로의 미분값, I_y는 y 방향으로의 미분값을 의미한다. 우측 매트릭스의 성분에서 <>표시는 가우시안 웨이트와 마스크 내의 합계가 이루어진 형태를 의미한다.

[0042] 상기와 같이 구해진 코너점의 정도는 하기 [식2]와 같이 계산될 수 있다.

[0043] [식2]

[0044] $$M_c=\lambda_1\lambda_2-\kappa(\lambda_1+\lambda_2)^2=\det(A)-\kappa trace^2(A)$$

[0045]

[0046] 여기에서, M_c는 코너 정도를 나타내는 것으로, 클수록 코너의 정도가 크다는 것을 의미한다. 한편, λ_1, λ_2는 매트릭스 A의 고유값(eigenvalue)이고, k는 상수, det는 행렬값, trace는 외각합을 의미한다. 한편, 상기 [식1] 및 [식2]는 공지된 것으로, 구체적인 설명은 생략하도록 한다.

[0047] 상기 코너점을 기반으로 하는 특징량 검출 방법은 계산량이 작아 특징량 검출 속도가 빠르고, 코너점에서 특징점이 바로 추출될 확률이 높아 신뢰성이 높다는 장점이 있다. 도 3에는 코너점을 기반으로 하는 특징량 검출 방법을 이용하여 특징량을 검출한 결과를 나타내었다(이상 S110).

[0048] 다음으로, 검출된 특징량을 매칭하여 제1매칭점(N)을 도출한다. 특징량을 매칭하는 방법은 한정되지 않는다. 예를 들면, 특징량 매칭 방법으로 템플릿 매칭(template matching) 또는 기술자 매칭(descriptor matching)을 이용할 수 있다.

[0049] 상기 템플릿 매칭은 특징점을 중심으로 일정 크기의 제1블록을 만들고, 상기 블록을 매칭해야 할 대상 영상의 검색 영역의 각 점에 대해서 동일한 크기의 제2블록을 만든 다음, 상기 제1블록과 상기 제2블록을 비교하여 가장 유사한 블록을 갖는 점을 매칭점으로 도출하는 방법을 의미한다. 도4에는 도3에서 검출된 코너점을 템플릿 매칭시킨 결과를 나타내었다.

[0050] 한편, 상기 기술자 매칭은 특징점 근처를 기술자로 기술한 다음, 대응 영상 간의 가장 비슷한 기술자를 갖는 경우에 두 특징점을 매칭점으로 도출하는 방법을 의미한다. 상기 기술자 매칭을 수행하는 알고리즘으로는 예를 들면, SIFT(Scale Invariant Feature Transform)가 있으며, 상기 SIFT는 영상에서의 도함수(gradient)를 구하여 세부영역을 구성하고, 각 세부영역에서 픽셀들의 기울기 방향 히스토그램을 구분하여 전체 영역을 구성하는 방법을 사용한다. 도 5에는 기술자 기반의 특징량 매칭 결과를 나타내었다.

[0051] 상기 탬플릿 매칭 또는 상기 기술자 매칭은 대응 영상의 상태에 따라 적절히 선택되어 수행될 수 있다. 예를 들어, 대응 영상이 상대적으로 근접 거리에 있어 검색 영역을 제한할 수 있는 경우에는 템플릿 매칭을 수행 가능하고, 대응 영상의 스케일이 변하거나 회전 등의 변형이 일어난 경우에는 기술자 매칭을 수행할 수

있다. 한편, 상기 제1매칭점(N)은 상기 매칭 방법들에 의해 매칭된 후 얻어진 매칭점들을 의미한다(이상 S120).

[0052] 다음으로, 상기 제1매칭점(N)에서 잘못된 매칭(아웃라이어, outlier)을 제거하여 제2매칭점(M)을 도출한다. 상기 제2매칭점(M)은 잘못된 매칭을 제거하여 최종적으로 얻어진 매칭점들을 의미한다.

[0053] 잘못된 매칭을 제거하는 방법은 예를 들면, 호모그래피(homography)와 같은 기하학적 모델 및 RANSAC(Random Sample Consensus)와 같은 추정 방법론을 이용할 수 있다. 상기 호모그래프 및 RANSAC을 이용하여 잘못된 매칭을 제거하는 방법은 공지된 것인 바, 상세한 설명은 생략하기로 한다(이상 S130).

[0054] 다음으로, 특징량 검출의 적합성을 판단한다. 여기에서 특징량 검출의 적합성을 판단한다는 의미는, 상기 코너점을 기반으로 하는 특징량 검출 방법의 적정성을 검토함을 의미한다.

[0055] 이는 상기 코너점을 기반으로 하는 특징량 검출 방법은 특징량 검출 속도가 빠르고 신뢰성이 높다는 장점이 있으나, 영상의 스케일 변화 또는 회전 변화에 대응할 수 없기 때문에 특정 경우에는 적용될 수 없으므로 적정성을 검토하여, 부적합한 경우 다른 방식의 특징량 검출 방법을 적용시키기 위함이다.

[0056] 구체적으로, 제2매칭점(M)의 수를 제1매칭점(N)의 수로 나눈 값(M/N)이 0.7 이상인 경우에는 코너점을 기반으로 하는 특징량 검출 방법이 적합함을 의미한다. 반대로, 제2매칭점(M)의 수를 제1매칭점(N)의 수로 나눈 값(M/N)이 0.7 미만인 경우에는, 코너점을 기반으로 하는 특징량 검출 방법이 부적합함을 의미하므로, 불변특징량을 기반으로 하는 특징량 검출 방법을 선택하여 상기 과정들을 재수행한다.

[0057] 여기에서 상기 불변특징량을 기반으로 하는 특징량 검출 방법은 영상내에서 대상 영역의 불변한 특징 또는 부분적으로 불변한 특징을 이용하여 특징량을 검출하는 방법으로 SIFT(Scale Invariant Feature Transform), SURF(Speeded Up Robust Features) 등과 같은 알고리즘이 이용될 수 있다. 상기 불변특징량을 기반으로 하는 특징량 검출 방법은 공지된 것으로, 구체적인 설명은 생략하도록 한다.

[0058] 즉, 불변특징량을 이용하여 특징량을 재검출하고(S142), 재검출된 상기 특징량을 매칭한 후, 잘못된 매칭을 제거하여 제2매칭점(M2)을 재도출한다(S144). 여기에서 특징량을 매칭하는 방법(탬플릿 매칭 또는 기술자 매칭) 및 잘못된 매칭을 제거하는 방법(호모그래피 및 RANSAC)은 전술한 것과 동일 또는 유사할 수 있으므로 중복 설명은 제외하도록 한다.

[0059] 한편, 상기 제2매칭점(M2)은 불변특징량을 기반으로 하는 특징량 검출 방법에 따라 도출되는 값으로, 코너점을 기반으로 하는 특징량 검출 방법에 따라 도출되는 제2매칭점(M)과는 구분되는 값을 가진다.

[0060] 도 6에는 불변특징량을 기반으로 하는 특징량 검출 방법에 따라 특징량을 검출한 결과를 나타내었다. 도 6을 참조하면, 원의 크기는 해당 점에서의 스케일을 의미하며, 원의 중심에서 원의 임의의 한 점까지 그려진 선은 해당 점에서의 방위(orientation)을 의미한다.

[0061] 이와 같이, 본 발명의 일 실시예에서는 코너점을 기반으로 하는 특징량 검출 방법을 우선적으로 적용시키되, 조건을 만족하지 않는 경우에는 불변특징량을 기반으로 하는 특징량 검출 방법을 적용하도록 할 수 있다(이상 S140).

[0062] 다음으로, 대응점 결정의 적합성을 판단한다. 여기에서 대응점 결정의 적합성을 판단한다는 의미는, 특징량을 추출하여 대응점을 결정하는 방법의 적정성을 검토함을 의미한다.

[0063] 이는 특징량을 추출하여 대응점을 결정하는 방법(코너점 기반 또는 불변특징량 기반)은 대응점이 적어 매칭 시간이 상대적으로 적게 걸리며 대응점의 정확성을 매우 높게 확보할 수 있다는 장점이 있으나, 대응점이 주로 코너점 또는 경계선에서 추출되기 때문에 대응점에서 얻어진 컬러 페어가 다양한 색을 가질 수 없기 때문에, 적정성을 검토하여 부적합한 경우 다른 방식의 대응점 결정 방법을 적용시키기 위함이다.

[0064] 구체적으로, 대응점 결정 적합성을 판단하는 단계는 제2매칭점(M 또는 M2)의 R, G, B 컬러 분포값이 전체 영상의 R, G, B 컬러분포값의 20% 이상 또는 미만인지 여부를 판단한다. 이 때, 상기 제2매칭점의 R, G, B 컬러 분포값이 전체 영상의 R, G, B 컬러 분포 값의 20% 이상인 경우에는, 특징량을 추출하여 대응점을 결정하는 방식이 적합함을 의미하는 바, 상기 제2매칭점을 대응점으로 결정한다(S160).

[0065] 반면에, 상기 제2매칭점의 R, G, B 컬러 분포값이 전체 영상의 R, G, B 컬러 분포 값의 20% 미만인 경우에는, 특징량을 추출하여 대응점을 결정하는 방식이 부적합함을 의미하는 바, 밀집 스테레오 매칭을 이용하여 대응점을 결정하는 방법을 선택하도록 한다.

[0066] 상기 밀집 스테레오 매칭을 이용하여 대응점을 결정하는 방법은 모든 픽셀에 대하여 스테레오 매칭을 수행하여 대응점을 검색하는 방법을 의미하며, 밀집 스테레오 매칭을 이용하여 매칭점을 재도출하고(S170), 재도출된 상기 매칭점을 대응점으로 결정하게 된다(S172).

[0067] 상기 밀집 스테레오 매칭은 상기 특징량을 추출하여 대응점을 검색하는 방법에 비하여 매칭 시간이 오래 걸리고 신뢰도가 다소 떨어질 수 있으나, 컬러 페어가 영상 장면의 모든 컬러를 포함한다는 장점이 있다. 여기에서 매칭 방법은 전술한 것과 동일 또는 유사할 수 있으므로 중복 설명은 제외하도록 한다.

[0068] 이와 같이, 본 발명의 일 실시예에서는 특징량을 추출하여 대응점을 검색하는 방법을 우선적으로 적용시키되, 조건을 만족하지 않는 경우에는 밀집 스테레오 매칭을 이용하여 대응점 검색을 수행하도록 할 수 있다(이상 S160 및 S170).

[0069]

[0070] 도 7은 도 1의 복사 보정 단계(S200)를 세부적으로 나타내는 흐름도이고, 도 8은 다수 카메라의 반응 함수를 나타내는 그래프이다.

[0071] 도 7을 참조하면, 복사 보정 단계(S200)는 카메라의 반응 함수를 산출하는 단계(S210)와, 상기 반응 함수의 역함수를 산출하는 단계(S220) 및 산출된 상기 역함수를 영상에 적용하여 선형인 카메라 반응 함수를 갖는 영상으로 변환하는 단계(S230)를 포함한다. 한편, 보정을 위한 대응 컬러는 앞서 설명한 대응점 검색 단계(S100)에서 얻어질 수 있다.

[0072] 여기에서, 복사 보정(radiometric calibration)이란 영상의 카메라 반응 함수를 구하여, 이의 역함수를 영상에 가함으로써 카메라 반응 함수가 선형함수가 되도록 보정하는 것을 의미한다.

[0073] 또한, 상기 카메라 반응 함수(camera response function)란 카메라의 영상을 통해 얻은 관찰(measurement)과 장면의 복사휘도(scene radiance)간의 관계를 의미한다.

[0074] 일반적으로, 디지털 카메라는 렌즈를 통해서 들어온 빛의 양을 CCD 및 CMOS와 같은 이미지 센서를 통해 전기 신호로 변환하고 이를 다시 디지털 신호로 전환한다. 이 때, 감시 또는 머신 비전에 사용되는 카메라는 빛의 크기에 선형 비례하는 디지털 값을 그대로 출력하기 때문에 선형의 카메라 반응 함수를 가지지만, 대부분의 사용자들이 사용하는 디지털 카메라 또는 방송용 카메라의 경우에는 카메라의 종류, 주변 환경등에 따라 선형의 카메라 반응 함수를 가지지 않는다. 그러므로, 대응 영상의 카메라 반응 함수 차이에 의한 컬러 차이를 최소화 시킬 필요가 있다.

[0075] 한편, 이와 관련하여 도 8은 여러 종류의 필름 카메라 또는 비디오 카메라의 반응 함수가 다양하게 존재하는 것을 보여주고 있다. 상기 도 8에서 X축은 카메라의 영상을 통해 얻은 관찰에 해당하고, Y축은 장면의 복사휘도에 해당한다.

[0076] 상기 카메라 반응 함수는 하기 [식3]과 같이 표현될 수 있다.

[0077] [식 3]

[0078] $$M = g(I)$$

[0079] 여기에서 g함수가 카메라 반응 함수이고, I는 장면의 복사휘도, M은 장면(카메라 영상)에 해당한다.

[0080] 따라서, 카메라 반응 함수인 g함수의 역함수(f함수)를 상기 장면(M)에 가하면, 장면의 복사휘도를 나타내는 영상으로 변환될 수 있다. 이는 하기 [식 4]와 같이 표현 될 수 있다.

[0081] [식 4]

[0082] $$I = f(M), where \ \ f = g^{-1}$$

[0083] 이와 같이, 본 발명의 일 실시예에서는 비선형인 카메라 반응 함수를 갖는 영상을 선형의 카메라 반응 함수를 갖는 영상으로 변환할 수 있다(이상 S200).

[0084] 상기와 같은 복사 보정 단계(S200)를 거친 후에는, 컬러 변환 행렬을 산출하여 대응 영상 간의 컬러 차이를 줄이게 된다. 상기 컬러 변환 행렬을 산출하는 방법은 다양할 수 있다.

[0085] 예를 들면, 상기 컬러 변환 행렬을 산출하는 방법으로 하기 [식5] 내지 [식7] 중 어느 하나에 따른 방법을 사용할 수 있다.

[0086] [식 5]

[0087] $$\sum_{x=1}^{NS}(\overrightarrow{Ic_s}-(a_c\overrightarrow{Tc_s}+b_c))^2,\ c\in R,G,B$$

[0088] 상기 [식 5]는 최소 제곱(least square) 방법에 따른 컬러 변환 행렬을 산출하는 식에 해당한다. 여기에서, Ic_s는 목표로 하는 영상의 RGB 컬러, Tc_s는 변환하려고 하는 입력 영상의 RGB 컬러, a_c 및 b_c는 두 컬러를 일치시키기 위한 계수를 의미한다.

[0089] [식 6]

[0090] $$\sum_{s=1}^{NS}(\overrightarrow{I_s}-T_{RGB}\bullet\overrightarrow{T_s})^2$$

[0091] 상기 [식 6]은 3×3 RGB 컬러 변환 행렬을 산출하는 식에 해당한다. 여기에서, I_s는 입력 영상 내의 S번째 대응점의 R, G, B 벡터, T_{RGB}는 컬러 변환 행렬의 각 성분, T_s는 목표 영상 내의 S번째 대응점의 R, G, B 벡터를 의미한다.

[0092] [식 7]

[0093] $$\sum_{k=1}^{D}\left(t_{rc_k}Ir_s^k+t_{gc_k}Ig_s^k+t_{bc_k}Ib_s^k\right)+t_{c0}\simeq Tc_s$$

[0094] 상기 [식 7]은 비선형 다항식을 갖는 컬러 변환 행렬을 산출하는 식에 해당한다. 여기에서, D는 비선형적 변환에서의 차수, t_{rck}, t_{gck}, t_{bck}는 각 차수의 각 컬러에 대항하는 계수, t_{c0}는 0차에서의 계수, Ir^k_s, Ig^k_s, Ib^k_s는 입력영상의 각 차수의 R,G,B 값, Tc_s는 목표 영상의 한 채널(R,G,B중 하나)의 값을 의미한다. 한편, 상기 [식 5] 내지 [식 7]은 공지된 것으로, 구체적인 설명은 생략하도록 한다.

[0095] 이와 같이, 컬러 변환 행렬을 산출하여 대상 영상의 모든 픽셀에 대하여 컬러 변환 행렬을 적용시키면 대응 영상의 컬러를 동일하게 할 수 있다(이상 S300).

[0096] 상술한 본 발명에 따른 대응 영상의 컬러 보정 방법은 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현할 수 있다. 컴퓨터가 읽을 수 있는 기록메체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피디스크, 광 데이터 저장장치 등이 있으며, 또한 캐리어 웨이브의 형태로 구현되는 것도 포함한다. 또한, 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다. 그리고, 상기 대응 영상의 컬러 보정 방법을 구현하기 위한 기능적인(function) 프로그램, 코드 및 코드 세그먼트들은 본 발명이 속하는 기술분야의 프로그래머들에 의해 용이하게 추론될 수 있다.

[0097] 상술한 것과 같이, 본 발명의 실시예들은 맥베스 컬러차트를 사용하지 않음으로써, 자동적인 대응점 매

칭이 가능하고 카메라 파라미터가 자동적으로 변하거나 조명 또는 날씨 변화와 같은 다양한 조건하에서도 컬러 보정을 가능하게 할 수 있다. 또한, 대응 영상의 카메라 반응함수가 선형함수를 갖도록 복사 보정함으로써, 비선형 카메라 반응 함수에 의한 오차를 제거할 수 있다. 또한, 컬러 변환 행렬을 3X3 RGB 컬러 변환 행렬뿐만이 아니라, 최소 제곱 방법 또는 비선형적 다항식을 이용하여 산출 가능하다.

[0098] 이상, 본 발명의 실시예들에 대하여 설명하였으나, 해당 기술 분야에서 통상의 지식을 가진 자라면 특허청구범위에 기재된 본 발명의 사상으로부터 벗어나지 않는 범위 내에서, 구성 요소의 부가, 변경, 삭제 또는 추가 등에 의해 본 발명을 다양하게 수정 및 변경시킬 수 있을 것이며, 이 또한 본 발명의 권리범위 내에 포함된다고 할 것이다.

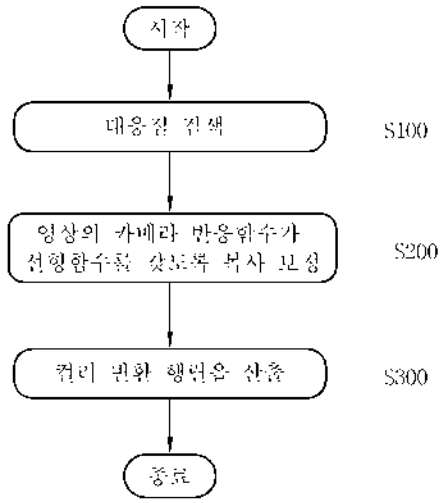
부호의 설명

[0099] N: 제1매칭점

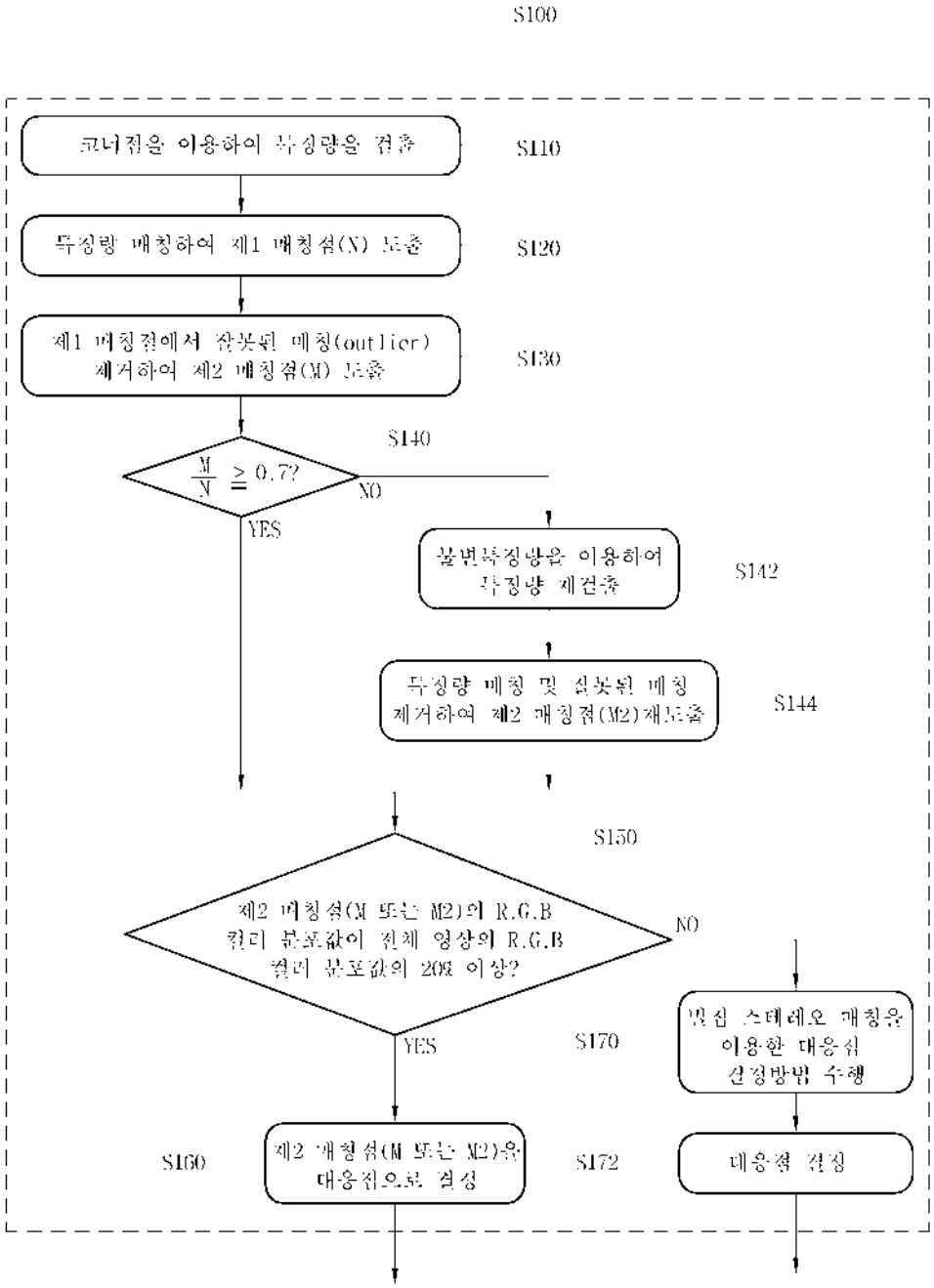
M, M2: 제2매칭점

도면

도면1



도면2

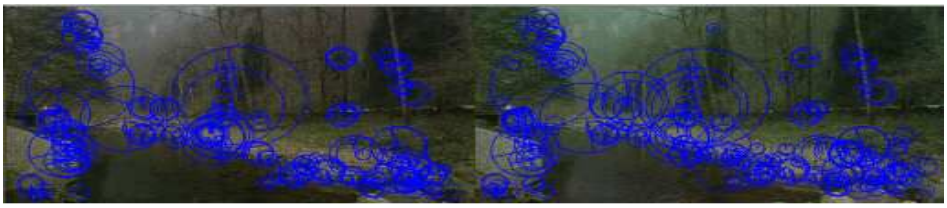

S100
코너점을 이용하여 특징량을 검출
S110
특징량 매칭하여 제1 매칭점(N) 도출
S120
제1 매칭점에서 잘못된 매칭(outlier) 제거하여 제2 매칭점(M) 도출
S130
S140
M/N ≥ 0.7?
NO
YES
불변특징량을 이용하여 특징량 재검출
S142
특징량 매칭 및 잘못된 매칭 제거하여 제2 매칭점(M2)재도출
S144
S150
제2 매칭점(M 또는 M2)의 R,G,B 컬러 분포값이 전체 영상의 R,G,B 컬러 분포값의 20% 이상?
NO
YES
S170
멀티 스테레오 매칭을 이용한 대응점 결정방법 수행
S160
제2 매칭점(M 또는 M2)을 대응점으로 설정
S172
대응점 결정

도면3

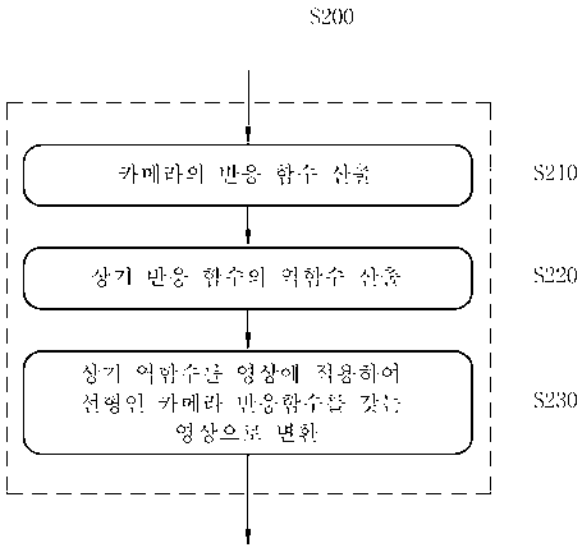
도면4

도면5

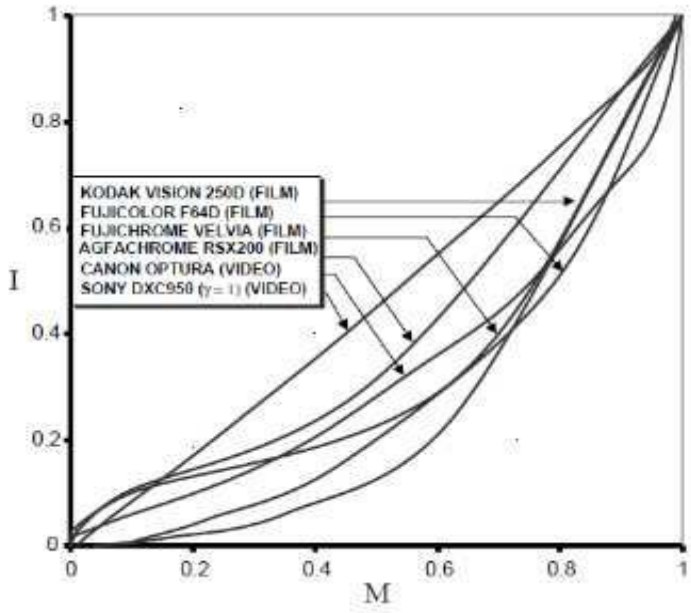
도면6



도면7



도면8


KODAK VISION 250D (FILM)
FUJICOLOR F64D (FILM)
FUJICHROME VELVIA (FILM)
AGFACHROME RSX200 (FILM)
CANON OPTURA (VIDEO)
SONY DXC950 (γ= 1) (VIDEO)
I
M
0
0.2
0.4
0.6
0.8
1

(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2015년02월12일
(11) 등록번호 10-1492060
(24) 등록일자 2015년02월04일

(51) 국제특허분류(Int. Cl.)
H04N 9/64 (2006.01) *H04N 9/79* (2006.01)
(21) 출원번호 10-2013-0098569
(22) 출원일자 2013년08월20일
심사청구일자 2013년08월20일
(56) 선행기술조사문헌
KR1020110016505 A*
JP2008300981 A*
KR1020120133417 A*
KR1020110071854 A
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
황영배
서울 서초구 방배선행길 1, 106동 607호 (방배동, 방배우성아파트)
김선주
인천 연수구 컨벤시아대로42번길 95, 1005동 2003호 (송도동, 더샵 엑스포)
김제우
경기 성남시 분당구 수내로 181, 309동 905호 (분당동, 샛별마을우방아파트)
(74) 대리인
남충우, 노철호

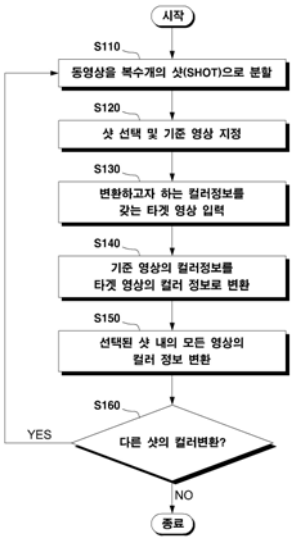
전체 청구항 수 : 총 7 항 심사관 : 이상래

(54) 발명의 명칭 **동영상의 일괄 컬러 변환 방법 및 그 기록매체**

(57) 요 약

동영상의 일괄 컬러 변환 방법 및 그 기록매체가 개시된다. 본 발명의 일 실시예에 따른 동영상의 일괄 컬러 변환 방법은 동영상을 복수개의 샷(SHOT)으로 분할하는 1단계; 샷들 중에서 어느 하나의 샷을 선택하고, 샷을 이루는 영상들 중에서 적어도 하나의 기준 영상을 지정하는 2단계; 기준 영상과 동일한 영상을 포함하는 것으로, 기준 영상과 다른 컬러 정보를 가지는 타겟 영상을 입력하는 3단계; 기준 영상의 컬러 정보를 타겟 영상의 컬러 정보로 변환하는 4단계; 및 변환된 기준 영상의 컬러 정보로 선택된 샷을 이루는 전체 영상들의 컬러 정보를 모두 변환하는 5단계를 포함한다.

대 표 도 - 도1



이 발명을 지원한 국가연구개발사업

과제고유번호 10043450

부처명 지식경제부

연구관리전문기관 산업기술평가관리원(KEIT)

연구사업명 산업융합원천기술개발사업

연구과제명 8K UHD 및 4K S3D(stereoscopic 3D) 콘텐츠의 획득/저장/Ingest 및 전송용 비디오 서버 기술 개발

기 여 율 1/1

주관기관 전자부품연구원

연구기간 2012.09.01 ~ 2015.08.31

특허청구의 범위

청구항 1

동영상의 일괄 컬러 변환 방법에 있어서,

동영상을 복수개의 샷(SHOT)으로 분할하는 1단계;

상기 샷들 중에서 어느 하나의 샷을 선택하고, 상기 샷을 이루는 영상들 중에서 적어도 하나의 기준 영상을 지정하는 2단계;

상기 기준 영상과 동일한 영상을 포함하는 것으로, 상기 기준 영상과 다른 컬러 정보를 가지는 타겟 영상을 입력하는 3단계;

상기 기준 영상의 컬러 정보를 상기 타겟 영상의 컬러 정보로 변환하는 4단계; 및

상기 변환된 기준 영상과 타겟 영상 간의 변환 관계로 상기 선택된 샷을 이루는 전체 영상들의 컬러 정보를 모두 변환하는 5단계를 포함하고,

상기 2단계에서의 기준 영상은 상기 선택된 샷의 첫 번째 영상이며,

상기 동영상의 일괄 컬러 변환 방법은,

상기 첫 번째 영상과 상기 선택된 샷의 다른 영상들의 픽셀 값의 차이를 계산하여 상기 차이가 임계값 이상인 영상들 중에서 상기 차이가 가장 큰 영상을 새로운 기준 영상으로 추가하는 2-1 단계를 더 포함하는 동영상의 일괄 컬러 변환 방법.

청구항 2

청구항 1에 있어서,

상기 5단계 이후에,

상기 2단계부터 5단계까지 반복 수행하는 6단계를 더 포함하는 동영상의 일괄 컬러 변환 방법.

청구항 3

청구항 1에 있어서,

상기 1단계에서의 동영상 분할은 두 개의 인접하는 영상의 픽셀 값의 차이를 계산하여 상기 차이가 임계값 미만이면 동일 샷으로 판단하고, 상기 임계값 이상이면 다른 샷으로 판단하여 분할하는 방법을 이용하는 동영상의 일괄 컬러 변환 방법.

청구항 4

삭제

청구항 5

삭제

청구항 6

동영상의 일괄 컬러 변환 방법에 있어서,

동영상을 복수개의 샷(SHOT)으로 분할하는 1단계;

상기 샷들 중에서 어느 하나의 샷을 선택하고, 상기 샷을 이루는 영상들 중에서 적어도 하나의 기준 영상을 지정하는 2단계;

상기 기준 영상과 동일한 영상을 포함하는 것으로, 상기 기준 영상과 다른 컬러 정보를 가지는 타겟 영상을 입력하는 3단계;

상기 기준 영상의 컬러 정보를 상기 타겟 영상의 컬러 정보로 변환하는 4단계; 및

상기 변환된 기준 영상과 타겟 영상 간의 변환 관계로 상기 선택된 샷을 이루는 전체 영상들의 컬러 정보를 모두 변환하는 5단계를 포함하고,

상기 2단계에서의 기준 영상은 상기 선택된 샷의 첫 번째 영상이며,

상기 동영상의 일괄 컬러 변환 방법은,

상기 5단계 이후에,

상기 변환된 전체 영상들의 컬러 정보를 각각 상기 타겟 영상의 컬러 정보와 비교하는 5-1단계;

상기 변환된 전체 영상들 중 상기 타겟 영상과의 컬러 정보 차이가 임계값 이상인 영상이 존재하는 경우에, 상기 컬러 정보 차이가 가장 큰 영상을 기준 영상에 추가하는 5-2단계;

상기 4단계 및 5단계를 재수행하는 5-3단계를 더 포함하는 동영상의 일괄 컬러 변환 방법.

청구항 7

청구항 1에 있어서,

상기 타겟 영상은 상기 기준 영상과 동일한 카메라를 사용하여 찍은 영상이되 다른 영상 세팅을 갖는 영상이거나, 상기 기준 영상과 다른 카메라를 사용하여 찍은 영상이거나, 상기 기준 영상과 동일한 장면을 포함하는 것으로 인터넷을 통해 다운로드한 영상이거나, 상기 기준 영상을 영상 편집 소프트웨어를 사용하여 영상 처리한 영상인 동영상의 일괄 컬러 변환 방법.

청구항 8

청구항 1에 있어서,

상기 4단계는,

상기 기준 영상 및 타겟 영상에서 동일한 물리적 위치에 있으면서 동일한 특징(feature)을 갖는 대응점을 검색하는 4-1단계;

상기 기준 영상 및 타겟 영상의 카메라 반응함수가 선형함수를 갖도록 복사 보정(radiometric calibration)하는 4-2단계; 및

컬러 변환 행렬을 산출하여 컬러 정보를 변환하는 4-3단계를 포함하는 동영상의 일괄 컬러 변환 방법.

청구항 9

청구항 1 내지 청구항 3 및 청구항 6 내지 청구항 8 중 어느 한 항에 기재된 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

명 세 서

기 술 분 야

[0001] 본 발명은 컬러 변환 방법에 관한 것으로, 보다 상세하게는 동영상의 컬러를 일괄적으로 변환하기 위한 동영상의 일괄 컬러 변환 방법과 상기 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록 매체에 관한 것이다.

배 경 기 술

[0002] 인터넷 기술 및 스트림 전송 기술의 발달로 인해 다양한 멀티미디어 콘텐츠들이 활발하게 제작 및 유통되고 있다. 특히 동영상 콘텐츠들의 경우에는 스마트 디바이스 및 소셜 네트워크의 확산으로 큰 폭의 성장세를 보이고 있는 추세이다.

[0003] 이러한 동영상들은 다양한 신(SCENE) 및 컷(CUT)으로 구분된 부분 영상들을 조합 및 편집하여 이루어지는 것이 대부분이다. 이 때, 장비나 기후 등의 동영상 촬영 조건에 따라 하나의 동영상을 이루는 다수의 부분 영상들은 서로 컬러 정보가 상이한 경우가 매우 많은 바, 경우에 따라서는 전체 동영상에 있어 균일한 색감을 유지해야

하는 경우가 발생한다. 그리고 이 경우에 이용될 수 있는 기술이 동영상의 컬러 변환 기술이다.

[0004] 동영상은 수 많은 연속 프레임들이 모여 이루어지는 바, 동영상 전체의 컬러 변환을 위해서는 상기 프레임들을 모두 RAW로 저장한 후에 적절한 영상 처리(예컨대 화이트 밸런스, 감마 코렉션 등)를 거치게 된다. 그러나 RAW는 데이터 용량이 매우 크기 때문에 RAW 형태로 동영상이 존재하기는 사실상 어려우며, 전체 동영상의 컬러를 변환하기 위해서는 수많은 프레임들을 일일이 컬러 변환하여야 하므로 시간이 매우 오래 걸린다는 단점이 있다.

[0005] 이에 동영상의 컬러를 변환하기 위한 편집 툴(TOOL)들이 존재하고 있으나, 이러한 툴들은 컬러의 절대 값을 변환시키는 바, 애초부터 영상 간 존재하는 촬영 조건 차이에 따른 불균질한 컬러 차이를 완전히 보정하기 어려우므로 한계가 있다.

발명의 내용

해결하려는 과제

[0006] 본 발명의 실시예들은 동영상의 전체 영상 또는 부분 영상에 대하여 일괄적으로 컬러를 변환시킬 수 있는 동영상의 일괄 컬러 변환 방법 및 그 기록매체를 제공하고자 한다.

과제의 해결 수단

[0007] 본 발명의 일 측면에 따르면, 동영상을 복수개의 샷(SHOT)으로 분할하는 1단계; 상기 샷들 중에서 어느 하나의 샷을 선택하고, 상기 샷을 이루는 영상들 중에서 적어도 하나의 기준 영상을 지정하는 2단계; 상기 기준 영상과 동일한 영상을 포함하는 것으로, 상기 기준 영상과 다른 컬러 정보를 가지는 타겟 영상을 입력하는 3단계; 상기 기준 영상의 컬러 정보를 상기 타겟 영상의 컬러 정보로 변환하는 4단계; 및 상기 변환된 기준 영상과 타겟 영상 간의 변환 관계로 상기 선택된 샷을 이루는 전체 영상들의 컬러 정보를 모두 변환하는 5단계를 포함하는 동영상의 일괄 컬러 변환 방법이 제공될 수 있다.

[0008] 또한, 상기 5단계 이후에, 상기 2단계부터 5단계까지 반복 수행하는 6단계를 더 포함할 수 있다.

[0009] 또한, 상기 1단계에서의 동영상 분할은 두 개의 인접하는 영상의 픽셀 값의 차이를 계산하여 상기 차이가 임계값 미만이면 동일 샷으로 판단하고, 상기 임계값 이상이면 다른 샷으로 판단하여 분할하는 방법을 이용할 수 있다.

[0010] 또한, 상기 2단계에서의 기준 영상은 상기 선택된 샷의 첫 번째 영상일 수 있다.

[0011] 또한, 상기 첫 번째 영상과 상기 선택된 샷의 다른 영상들의 픽셀 값의 차이를 계산하여 상기 차이가 임계값 이상인 영상들 중에서 상기 차이가 가장 높은 영상을 새로운 기준 영상으로 추가하는 2-1 단계를 더 포함할 수 있다.

[0012] 또한, 상기 5단계 이후에, 상기 변환된 전체 영상들의 컬러 정보를 각각 상기 타겟 영상의 컬러 정보와 비교하는 5-1단계; 상기 변환된 전체 영상들 중 상기 타겟 영상과의 컬러 정보 차이가 임계값 이상인 영상이 존재하는 경우에, 상기 컬러 정보 차이가 가장 큰 영상을 기준 영상에 추가하는 5-2단계; 상기 4단계 및 5단계를 재수행하는 5-3단계를 더 포함할 수 있다.

[0013] 또한, 상기 타겟 영상은 상기 기준 영상과 동일한 카메라를 사용하여 찍은 영상이되 다른 영상 세팅을 갖는 영상이거나, 상기 기준 영상과 다른 카메라를 사용하여 찍은 영상이거나, 상기 기준 영상과 동일한 장면을 포함하는 것으로 인터넷을 통해 다운로드한 영상이거나, 상기 기준 영상을 영상 편집 소프트웨어를 사용하여 영상 처리한 영상일 수 있다.

[0014] 또한, 상기 4단계는, 상기 기준 영상 및 타겟 영상에서 동일한 물리적 위

[0015] 치에 있으면서 동일한 특징(feature)을 갖는 대응점을 검색하는 4-1단계; 상기 기준 영상 및 타겟 영상의 카메라 반응함수가 선형함수를 갖도록 복사 보정(radiometric calibration)하는 4-2단계; 및 컬러 변환 행렬을 산출하여 컬러 정보를 변환하는 4-3단계를 포함할 수 있다.

[0016] 본 발명의 다른 측면에 따르면, 본 발명의 일 측면에 따른 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체가 제공될 수 있다.

발명의 효과

[0017] 본 발명의 실시예들은 동영상을 이루는 샷(SHOT)에 대하여 최소 하나의 영상의 컬러를 변환할 경우, 상기 샷을 이루는 다른 영상에 대해서도 일괄적으로 컬러를 변환할 수 있으므로 매우 효율적으로 동영상의 컬러를 일괄 변환할 수 있다.

[0018] 또한, 변환하고자 하는 컬러 정보를 갖는 타겟 영상으로 다른 카메라 셋팅, 다른 카메라 또는 영상 편집 소프트웨어로 영상처리된 영상을 이용 가능하므로 동영상의 각 영상의 RAW를 별도로 저장할 필요가 없어 간편하다.

도면의 간단한 설명

[0019] 도 1은 본 발명의 일 실시예에 따른 동영상의 일괄 컬러 변환 방법을 나타내는 순서도이다.

도 2는 동영상에서 샷(SHOT)의 개념을 나타내는 영상 이미지이다.

도 3은 도 1에서 S120 단계를 세부적으로 나타내는 순서도이다.

도 4는 기준 영상과 타겟 영상의 일 예시를 나타내는 영상 이미지이다.

도 5는 도 1에서 S140 단계를 세부적으로 나타내는 순서도이다.

도 6은 도 1에서 S150 이후에 추가되는 단계를 세부적으로 나타내는 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0020] 이하, 첨부된 도면을 참조하여 본 발명의 실시예들에 대하여 구체적으로 설명하도록 한다.

[0021] 도 1은 본 발명의 일 실시예에 따른 동영상의 일괄 컬러 변환 방법(이하, 일괄 컬러 변환 방법)을 나타내는 순서도이다. 이하에서는 도 1에 도시된 각 단계별로 본 발명을 설명하도록 한다.

[0022] (1) 1단계(S110)

[0023] 1단계는 동영상을 복수개의 샷(SHOT)으로 분할하는 단계이다.

[0024] 본 명세서에 기재된 동영상의 계층적 구조에 대하여 간략히 설명하면, 동영상은 수 없이 많은 연속 프레임(FRAME)이 모여 이루어진다. 즉, 프레임은 동영상을 이루는 기본 단위에 해당한다. 본 명세서에서 프레임은 '영상'과 동일한 의미로 기재되었음을 밝혀둔다.

[0025] 복수개의 연속적인 프레임은 하나의 샷(SHOT)을 구성할 수 있다. 샷은 전체적으로 유사한 영상을 가진 프레임들의 집합으로 정의될 수 있다. 즉, 동영상은 복수개의 샷으로 구성된다. 일반적으로 동영상 내에서 장면 전환이 이루어진다 함은 샷(SHOT)이 달라졌다는 의미로도 파악될 수 있다.

[0026] 복수개의 샷은 의미적인 관계에 따라 하나의 신(SCENE)을 구성할 수 있다. 이러한 신(SCENE)은 동영상 편집자(내지 연출자)의 의도에 따라 달라질 수 있다. 즉, 편집자가 복수개의 샷(SHOT)을 어떻게 그룹화 하는지에 따라 신(SCENE)들은 다양하게 달라질 수 있다. 그러므로 하나의 신(SCENE) 내에서는 전혀 다른 영상들이 존재할 수도 있다.

[0027] 따라서 본 발명의 1단계에서 동영상을 분할한다는 의미는 신(SCENE)의 분할을 의미하는 것이 아니라, 전체적으로 유사한 영상을 가진 프레임의 집합인 샷(SHOT)의 분할을 의미함을 밝혀둔다.

[0028] 관련하여 도 2에서는 동영상에서의 샷(SHOT)의 개념을 나타낸다. 도 2는 하나의 뉴스(NEWS) 동영상을 15개의 샷(SHOT)으로 분할한 모습을 나타내고 있다. 설명하였듯이 각 샷들은 전체적으로 유사한 영상을 가진 프레임들의 집합이다. 도 2에서 각 샷의 위에 배치된 이미지는 각 샷에서 첫 번째 프레임이고, 아래에 배치된 이미지는 각 샷에서 마지막 프레임이다.

[0029] 동영상을 복수개의 샷으로 분할하는 방법에 대하여 설명한다.

[0030] 동영상을 복수개의 샷으로 분할하는 방법은 다양하게 존재할 수 있으며, 예컨대 두 개의 인접 영상(프레임)의 픽셀 값의 차이를 계산하여 상기 차이가 임계값 미만이면 동일 샷으로 판단하고, 상기 차이가 임계값 이상이면 다른 샷으로 판단하여 분할하는 방법을 이용할 수 있다(샷 경계 검출법, shot boundary detection). 여기에서 상기 임계값은 기 설정될 수 있다.

[0031] 또는, 영상 간의 전체 컬러 히스토그램(color histogram)을 비교하고, 상기 컬러 히스토그램의 차이가 임계값 미만이면 동일 샷으로 판단하고, 상기 차이가 임계값 이상이면 다른 샷으로 판단하여 분할하는 방법을 이용할

수 있다.

[0032] 또는, 동영상에서 장면의 변화가 많지 않거나 동영상 길이가 상대적으로 짧은 경우에는 사용자가 수동으로 샷을 분할하는 방법을 사용하는 것도 가능하다.

[0033] 또는, 동영상에 오디오 정보가 포함되어 있는 경우에는 상기 오디오 정보를 이용하는 것도 가능하다. 예컨대 사람의 목소리는 개인마다 고유한 피치 정보가 존재하는 바, 피치 변화를 분석하여 화자가 바뀌는 것을 검출함으로써 샷을 분할하는 방법을 사용할 수 있다.

[0034] 이렇듯 동영상을 복수개의 샷으로 분할하는 방법은 공지된 다양한 방법을 사용할 수 있으며, 상기 언급된 방법으로만 한정되는 것은 아니다(이상 S110).

[0035] (2) 2단계(S120)

[0036] 2단계는 분할된 샷(SHOT)들 중에서 어느 하나의 샷을 선택하고, 상기 샷을 이루는 영상(프레임)들 중에서 적어도 하나의 기준 영상을 지정하는 단계이다.

[0037] 상술하였듯이 하나의 샷은 복수개의 영상으로 구성된다. 상기 영상들은 유사한 픽셀 값 및 컬러 히스토그램값을 갖는다.

[0038] 기준 영상은 샷을 이루는 복수개의 영상들 중에서 적어도 하나 이상일 수 있다. 예컨대 기준 영상은 선택된 샷의 첫 번째 영상(프레임)일 수 있다(DEFAULT값). 본 명세서에서는 선택된 샷의 첫 번째 영상을 제1 기준 영상으로 지칭하기로 한다.

[0039] 기준 영상은 복수개가 지정될 수도 있다. 동일한 샷(SHOT) 내에서도 상대적으로 다른 영상들과 픽셀 값 및 컬러 히스토그램값이 차이가 큰 영상이 속할 수 있으며, 이 경우에는 컬러 변환 결과가 저하될 수 있기 때문이다. 본 명세서에서는 제1 기준 영상 이외에 기준 영상으로 추가되는 영상을 제2 기준 영상으로 지칭하기로 한다.

[0040] 도 3에서는 도 1의 S120 단계를 세부적으로 나타내는 순서도를 도시하고 있다. 도 3에서와 같은 단계를 거쳐 제2 기준 영상이 지정될 수 있다.

[0041] 구체적으로, 제1 기준 영상이 지정되면(S121) 제1 기준 영상과 선택된 샷의 다른 영상들의 컬러 정보 차이를 계산한다. 동일 샷 내에 있는 영상들은 비교적 유사한 컬러 정보를 가지고 있으나, 영상들끼리는 컬러 정보 차이가 존재할 수 있으며 경우에 따라서는 상기 차이가 클 수도 있다(S122). 한편, 본 명세서에서 컬러 정보란 영상으로부터 얻어지는 RGB(Red-Green-Blue)와 같은 색 모델뿐만 아니라 영상의 밝기, 채도, 색상 등과 같은 컬러 특성을 의미한다.

[0042] 이 때, 기 설정된 임계값(예컨대 임계값은 10%)과 비교하였을 때에(S123), 상기 컬러 정보 차이(제1 기준 영상과 선택된 샷의 다른 영상들 사이)가 임계값 미만인 경우에는 영상들 간의 컬러 정보가 비교적 유사한 수준에 있음을 의미하는 바, 제2 기준 영상을 추가 지정할 필요가 없다. 따라서 다음 단계인 S130으로 넘어갈 수 있다. 그러나, 상기 컬러 정보 차이가 임계값 이상인 경우에는 컬러 정보 차이가 가장 큰 영상을 제2 기준 영상으로 추가함으로써(S124) 후술할 컬러 변환 결과에서의 신뢰성을 향상시킬 수 있다.

[0043] 한편, 제2 기준 영상은 컬러 변환이 이루어진 후에도 추가적으로 지정될 수 있으며, 이에 대해서는 후술하기로 한다(이상 S120).

[0044] (3) 3단계(S130)

[0045] 3단계는 타겟 영상을 입력하는 단계이다. 본 명세서에서 타겟 영상은 기준 영상과 동일한 영상을 포함하는 것으로, 기준 영상과는 다른 컬러 정보를 가지는 영상을 의미한다. 즉, 타겟 영상은 기준 영상의 컬러 정보의 변환 대상이 되는 컬러 정보를 가진 영상이다.

[0046] 타겟 영상은 기준 영상과 동일한 영상을 포함하면 되고 특정 종류의 것으로 한정되지 않는다.

[0047] 예컨대, 타겟 영상은 기준 영상과 동일한 카메라를 사용하여 찍은 영상이되 다른 영상 세팅(여기에서 영상 세팅은 화이트 밸런스, 감마 코렉션, 픽처 스타일, 밝기 조정 등의 영상 세팅을 의미함)을 갖는 영상이거나, 기준 영상과 다른 카메라를 사용하여 찍은 영상일 수 있다(카메라가 다르면 스펙트럼 감도, 사용되는 감마, 스타일 등이 달라짐). 또한 타겟 영상은 기준 영상과 동일한 장면을 포함할 경우에 인터넷을 통해 다운로드한 영상일 수도 있다. 또한 타겟 영상은 기준 영상을 영상 편집 소프트웨어(포토샵, 페인트샵과 같은 편집툴)를 이용하여 영상처리한 영상일 수도 있다.

[0048] 관련하여 도 4는 기준 영상과 타겟 영상의 일 예시를 나타내는 영상 이미지이다. 좌측 이미지는 DSLR을 이용해서 찍은 영상이고, 우측 이미지는 스마트폰을 이용해서 찍은 영상이다. 촬영 장비가 다르므로 양 자의 색감은 상이하다. 이 때, 좌측 이미지가 기준 영상이고 우측 이미지가 타겟 영상이라 가정할 때에, 좌측 이미지는 본 발명의 컬러 변환 방법에 의해 우측 이미지와 같은 색감을 가지도록 변환될 수 있다. 즉, DSLR을 이용해서 찍은 영상을 마치 스마트폰으로 찍은 영상처럼 컨버젼할 수 있다. 물론 그 역의 경우도 가능하다(이상 S130).

[0049] (4) 4단계(S140)

[0050] 4단계는 S130에서와 같이 타겟 영상이 입력되면, 기준 영상의 컬러 정보를 타겟 영상의 컬러 정보로 변환하는 단계이다(컬러 보정). 컬러 정보 변환 방법은 본 기술분야에서 공지된 다양한 컬러 보정 방법을 이용할 수 있다.

[0051] 관련하여 본 발명의 발명자들이 출원하여 등록 받은 한국등록특허 제10-1227936호에서는 대응 영상의 컬러 보정 방법이 기재되어 있으며, 본 발명에서 S140 단계는 상기 특허에 기재된 대응 영상의 컬러 보정 방법이 적용될 수 있다. 즉, 본 발명의 S140 단계는 상기 특허의 내용을 포함할 수 있으며, 이하에서 간략히 설명하도록 한다.

[0052] 도 5는 도 1에서 S140 단계를 세부적으로 나타내는 순서도이다.

[0053] 도 5를 참조하면, 상기 4단계는 상기 기준 영상 및 타겟 영상에서 동일한 물리적 위치에 있으면서 동일한 특징(feature)을 갖는 대응점을 검색하는 4-1 단계와, 상기 기준 영상 및 타겟 영상의 카메라 반응함수가 선형함수를 갖도록 복사 보정(radiometric calibration, 영상의 카메라 반응 함수를 구하고, 이의 역함수를 영상에 가함으로써 카메라 반응 함수가 선형함수가 되도록 보정하는 것을 의미하며 카메라 반응 함수는 카메라의 영상을 통해 얻은 관찰(measurement)과 장면의 복사휘도간의 관계를 의미함)하는 4-2 단계와, 컬러 변환 행렬을 산출하여 컬러 정보를 변환하는 4-3 단계를 포함할 수 있다. 상기 각 단계에 대한 세부적인 설명은 한국등록특허 제10-1227936호에 구체적으로 기재되어 있는 바, 여기에서는 생략하도록 한다.

[0054] 이로써 기준 영상의 컬러 정보는 타겟 영상의 컬러 정보로 변환된다(이상 S140).

[0055] (5) 5단계(S150)

[0056] 5단계는 상기 변환된 기준 영상과 타겟 영상 간의 변환 관계로 상기 선택된 샷을 이루는 전체 영상들의 컬러 정보를 모두 변환하는 단계이다. 즉, 상기 4단계에서 기준 영상의 컬러 정보가 변환되면 상기 기준 영상과 타겟 영상 간의 변환 관계가 도출되고, 이를 토대로 동일한 샷 내에 있는 모든 영상들의 컬러 정보가 일괄적으로 변환될 수 있다.

[0057] 상술한 것과 같이 동일한 샷 내에 있는 모든 영상들은 기준 영상과 컬러 정보가 비교적 유사한 영상들이다. 따라서 동일한 샷 내의 모든 영상들을 변환된 기준 영상의 컬러 정보로 변환하는 것은 상대적으로 속도가 빠를 뿐만 아니라 신뢰성이 높아질 수 있다.

[0058] 한편, 동일한 샷 내의 모든 영상들에 대하여 일괄적으로 컬러 정보를 변환한 후에는 변환 결과에 따라 기준 영상을 추가하여(제2 기준 영상) 컬러 변환을 재수행하는 것이 가능하다.

[0059] 앞서 언급한 것처럼 동일 샷 내에 있는 영상들은 비교적 유사한 컬러 정보를 가지고 있으나, 영상들끼리는 컬러 정보 차이가 존재할 수 있으며 경우에 따라서는 상기 차이가 클 수도 있다. 따라서 기준 영상은 처음부터 복수개가 지정될 수 있으나, 동일한 샷 내의 모든 영상들에 대하여 컬러 정보를 변환한 후에 그 변환 결과에 따라 기준 영상을 추가로 지정하는 것이 가능하다.

[0060] 관련하여 도 6은 도 1에서 S150 이후에 추가되는 단계를 세부적으로 나타내는 순서도이다.

[0061] 도 6을 참조하면, S150 단계에서 변환된 전체 영상들(동일 샷 내에 있는 영상들임)의 컬러 정보를 각각 상기 타겟 영상의 컬러 정보와 비교한다(S151). 이 때, 컬러 정보가 변환된 전체 영상들 중에서 상기 타겟 영상과의 컬러 정보 차이가 임계값 미만인 경우에는 전체 영상들에 대한 컬러 변환이 적절하게 수행된 것으로 볼 수 있다. 따라서 다음 단계로 넘어갈 수 있다(S160).

[0062] 그러나 컬러 정보가 변환된 전체 영상들 중에서 상기 타겟 영상과의 컬러 정보 차이가 임계값 이상인 영상이 존재하는 경우에는(S152), 전체 영상들에 대한 컬러 변환이 적절하게 수행되지 않은 것으로 볼 수 있다. 따라서 기준 영상을 추가 지정할 필요가 있다. 그러므로 이와 같은 경우에는 컬러 정보 차이가 임계값 이상인 영상 중에서도 그 차이가 가장 큰 영상(가장 컬러 변환 결과가 나쁘게 나온 영상에 해당함)을 제2 기준 영상으로 추가한 다음(S153), 다시 S140 단계부터 재수행함으로써 컬러 변환의 신뢰성을 향상시킬 수 있다. 제2 기준 영상이

추가되어도 타겟 영상은 변하지 않으므로 가장 컬러 변환 결과가 나쁘게 나온 영상이 보다 타겟 영상에 가깝도록 변환되는 결과를 낳는다(이상 S150).

[0063] (6) 6단계(S160)

[0064] 상술한 5단계까지의 과정이 종료되면 동영상을 이루는 복수개의 샷(SHOT) 중에서 선택된 샷(SHOT)에 대한 컬러 변환이 이루어진 상태이다. 따라서 다른 샷에 대하여도 컬러 변환을 수행하고자 할 때에는 새로운 샷을 선택하고 상술한 S120 단계부터 S150 단계까지 반복 수행하게 된다. 상기와 같은 과정을 반복 수행하면 동영상 전체 영상에 대하여 원하는 타겟 영상과 동일 또는 유사한 컬러 정보를 갖도록 영상들이 일괄적으로 변환될 수 있다.

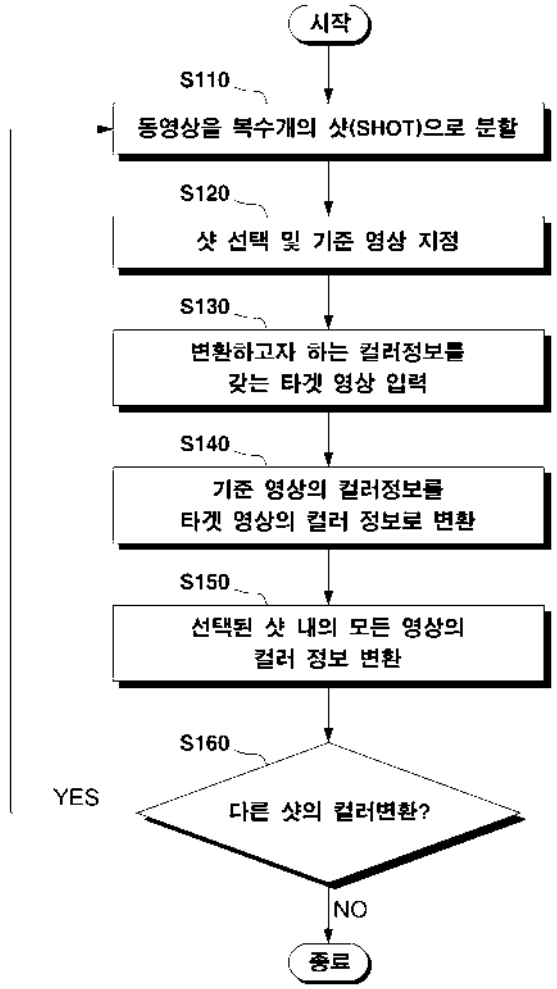
[0065] 상술한 바와 같이, 본 발명의 실시예들은 동영상을 이루는 샷(SHOT)에 대하여 최소 하나의 영상의 컬러를 변환할 경우, 상기 샷을 이루는 다른 영상에 대해서도 일괄적으로 컬러를 변환할 수 있으므로 매우 효율적으로 동영상의 컬러를 일괄 변환할 수 있다. 또한, 변환하고자 하는 컬러 정보를 갖는 타겟 영상으로 다른 카메라 셋팅, 다른 카메라 또는 영상 편집 소프트웨어로 영상처리된 영상을 이용 가능하므로 동영상의 각 영상의 RAW를 별도로 저장할 필요가 없어 간편하다는 장점을 갖는다.

[0066] 이상에서 상술한 본 발명에 따른 동영상의 일괄 컬러 변환 방법은 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현할 수 있다. 컴퓨터가 읽을 수 있는 기록메체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피디스크, 광 데이터 저장장치 등이 있으며, 또한 캐리어 웨이브의 형태로 구현되는 것도 포함한다. 또한, 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다. 그리고, 상기 대응 영상의 컬러 보정 방법을 구현하기 위한 기능적인(function) 프로그램, 코드 및 코드 세그먼트들은 본 발명이 속하는 기술분야의 프로그래머들에 의해 용이하게 추론될 수 있다.

[0067] 이상, 본 발명의 실시예들에 대하여 설명하였으나 해당 기술 분야에서 통상의 지식을 가진 자라면 특허청구범위에 기재된 본 발명의 사상으로부터 벗어나지 않는 범위 내에서, 구성 요소의 부가, 변경, 삭제 또는 추가 등에 의해 본 발명을 다양하게 수정 및 변경시킬 수 있을 것이며, 이 또한 본 발명의 권리범위 내에 포함된다고 할 것이다.

도면

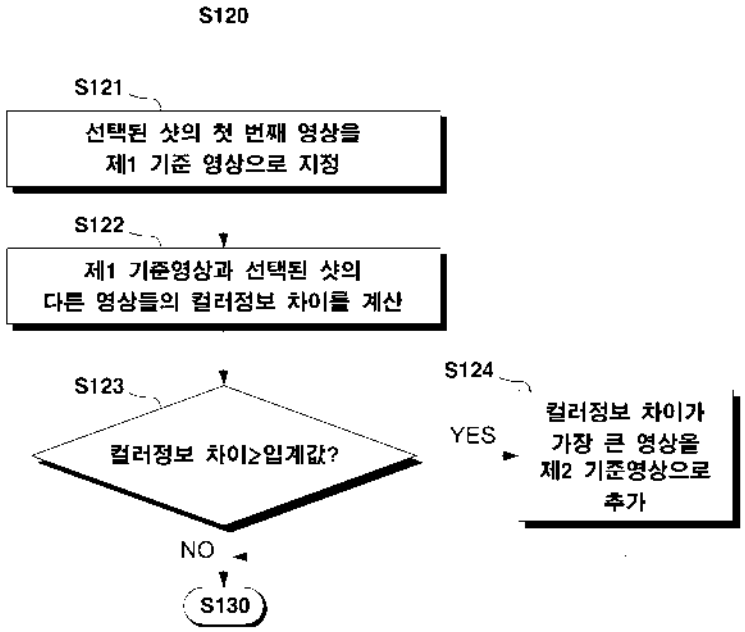
도면1


시작
S110
동영상을 복수개의 샷(SHOT)으로 분할
S120
샷 선택 및 기준 영상 지정
S130
변환하고자 하는 컬러정보를
갖는 타겟 영상 입력
S140
기준 영상의 컬러정보를
타겟 영상의 컬러 정보로 변환
S150
선택된 샷 내의 모든 영상의
컬러 정보 변환
S160
YES
다른 샷의 컬러변환?
NO
종료

도면2

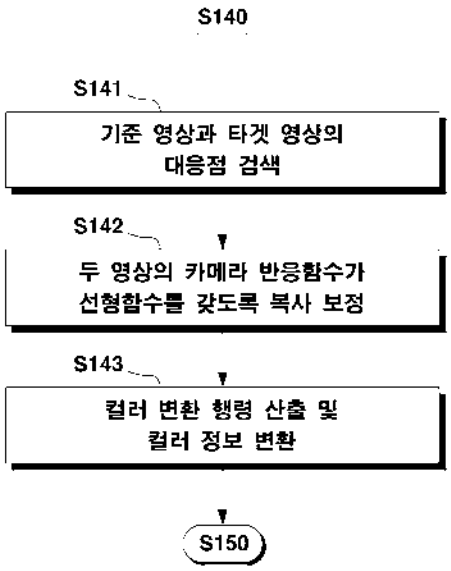

Shot 1
Shot 2
Shot 3
Shot 4
Shot 5
Shot 6
Shot 7
Shot 8
First Frame
Last Frame
Shot 9
Shot 10
Shot 11
Shot 12
Shot 13
Shot 14
Shot 15
First Frame
Last Frame

도면3

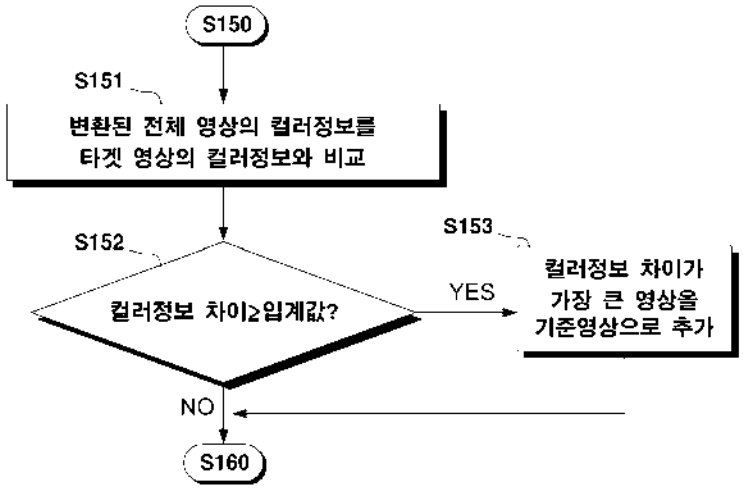

S120
S121
선택된 샷의 첫 번째 영상을
제1 기준 영상으로 지정
S122
제1 기준영상과 선택된 샷의
다른 영상들의 컬러정보 차이를 계산
S123
컬러정보 차이≥입계값?
YES
S124
컬러정보 차이가
가장 큰 영상을
제2 기준영상으로
추가
NO
S130

도면4

도면5


S140
S141
기준 영상과 타겟 영상의
대응점 검색
S142
두 영상의 카메라 반응함수가
선형함수를 갖도록 복사 보정
S143
컬러 변환 행렬 산출 및
컬러 정보 변환
S150

도면6


S150
S151
변환된 전체 영상의 컬러정보를
타겟 영상의 컬러정보와 비교
S152
컬러정보 차이≥임계값?
YES
S153
컬러정보 차이가
가장 큰 영상을
기준영상으로 추가
NO
S160

■ 기술명 : 객체 인식/추적 IP 기술
(Technology for Object Recognition/Tracking SW/HW IP)

산업기술분류	기계·소재/로봇・자동화 기계/로봇 비젼 및 생산자동화 기술(100503)
Key-word(국문)	로봇 비전, 객체 인식, 객체 추적, SoC, 모바일 로봇, 로봇 어플리케이션 SDK
Key-word(영문)	Robot Vision, Object Recognition, Object Tracking, SoC, Mobile robot, Robot application SDK

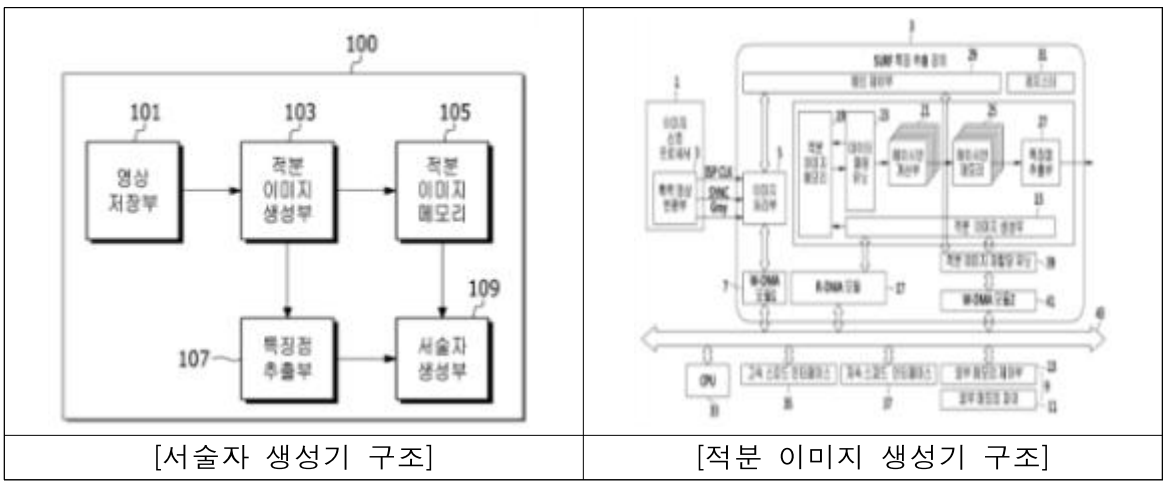
■ 기술의 개요

○ (배경) 저사양의 임베디드 보드에 구현할 수 있으며, 다양한 지능형 영상인식 분야에 적용될 수 있는 새로운 영상기반 객체 인식, 객체 추적 기술 필요

○ (개요) 영상 기반 지능적 정보 추출 및 동작을 위한 객체 인식/추적 기술을 SW로 구현하고, 높은 복잡도를 가진 알고리즘을 전용 HW로 구현

- 고사양 객체 인식이 특징인 SURF(Speeded Up Robust Feature)를 전용 HW로 구현
- 저 복잡도, 실시간 처리가 가능한 새로운 형태의 SW/SoC 구조 및 구현 기술

< 기술 개요도 >



[서술자 생성기 구조] [적분 이미지 생성기 구조]

■ 기술의 구현수준(TRL)

■ 기술의 장점(경쟁기술과의 차별성)

○ 고정확 및 고속 처리 가능

* 기존 기술 대비 약 10% 정확도 향상, 외부 DDR Interface 접목(AXI 구조 설계)

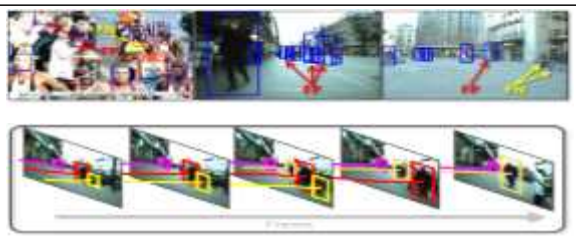
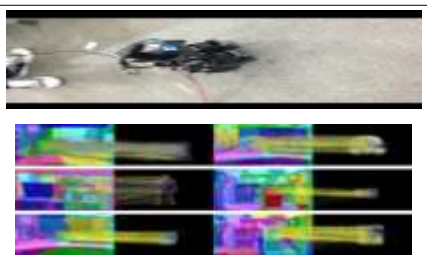
* 기존 기술 대비 약 17% 속도 향상, 다양한 기능을 단일 플랫폼에서 실시간 처리 가능 (30fps@VGA)

○ 요구 성능에 따른 가변 설계 가능

- SW/SoC 간의 Configurable한 구조를 이용하여 Target 플랫폼에 가장 적합한 구조 설계 및 적용 가능

○ 빠른 상용화/소량 상품화의 경우 임베디드 및 FPGA로의 적용, 대량 생산의 경우 SoC로의 변경이 가능한 구조

■ 활용범위 및 응용분야

<table>
<tr><td>  </td><td>  </td></tr>
<tr><td>[객체인식 FPGA 동작 및 SoC 구현]</td><td>[다중 객체 추적을 통한 이상행위 감지]</td></tr>
<tr><td>  </td><td>  </td></tr>
<tr><td>[객체인식 기능이 구현된 블랙박스]</td><td>[객체 인식을 통한 로봇 어플리케이션]</td></tr>
</table>

■ 지식재산권 현황

구분	발명의 명칭	출원번호 (출원일)	등록번호 (등록일)
특허	SURF 하드웨어 장치 및 적분 이미지 생성 방법	2013-0151188 (2013.12.06)	10-1531037 (2015.06.17)
특허	SURF 하드웨어 장치 및 적분 이미지 메모리 관리 방법	2013-0150958 (2013.12.05)	10-1531038 (2015.06.17)

구분	발명의 명칭	출원번호 (출원일)	등록번호 (등록일)
특허	하드웨어 장치 및 적분 이미지 생성 방법	2013-0045330 (2013.04.24)	10-1418524 (2014.07.04)
PCT	ＳＵＲＦ 하드웨어 장치 및 적분 이미지 생성 방법	KR2013/011268 (2013.12.06)	
PCT	ＳＵＲＦ 하드웨어 장치 및 적분 이미지 메모리 관리 방법	KR2013/011270 (2013.12.06)	
특허	깊이 정보를 이용한 물체 인식 방법 및 장치	2012-0051871 (2012.05.16)	10-1916460 (2018.11.01)
특허	영상 특성 기반 분할 기법을 이용한 물체 검출 방법 및 이를 적용한 영상 처리 장치	2013-0045556 (2013.04.24)	10-1517805 (2015.04.29)
PCT	저사양 영상 기기에서의 실시간 영상 인식 방법	KR2014/012832 (2014.12.24)	
특허	컬러 영상과 뎁스 영상을 이용한 움직임 영역 추출 방법 및 시스템	2014-0183245 (2014.12.18)	
특허	객체 검출과 추적을 결합한 객체 추적 방법 및 장치	2014-0187797 (2014.12.24)	

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2016-0078545
(43) 공개일자 2016년07월05일

(51) 국제특허분류(Int. Cl.)
G06T 7/20 (2006.01)
(21) 출원번호 10-2014-0187797
(22) 출원일자 2014년12월24일
심사청구일자 없음

(71) 출원인
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
김정호
경기도 성남시 분당구 황새울로258번길 35 (수내동) 트레벨오피스텔 907호
최병호
경기 용인시 수지구 대지로 27, 103동 1306호 (죽전동, 한신아파트)
(74) 대리인
남충우

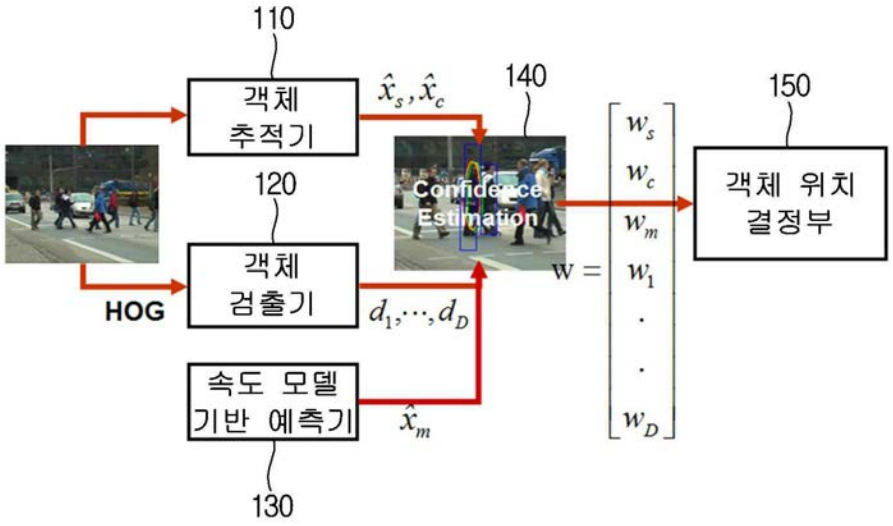
전체 청구항 수 : 총 6 항

(54) 발명의 명칭 **객체 검출과 추적을 결합한 객체 추적 방법 및 장치**

(57) 요 약

객체 검출과 추적을 결합한 객체 추적 방법 및 장치가 제공된다. 본 발명의 실시예에 따른 객체 추적 장치는, 추적기의 추적 결과에 대한 추적 신뢰도 및 검출기의 검출 결과에 대한 검출 신뢰도를 계산하고, 추적 결과와 추적 신뢰도 및 검출 결과와 검출 신뢰도를 기반으로 객체 위치를 결정한다. 이에 의해, 복잡한 환경에서도 정확한 객체 추적이 가능해짐은 물론, 가려짐이 발생하는 경우에도 강인한 객체 추적을 할 수 있게 된다.

대 표 도 - 도2



이 발명을 지원한 국가연구개발사업

과제고유번호	GK14D0200
부처명	미래창조과부/교육부
연구관리전문기관	재단법인 기가코리아사업단
연구사업명	(미래부)Giga KOREA사업
연구과제명	실시간 인터렉션을 제공하는 초다시점 단말기술개발
기 여 율	1/1
주관기관	한국과학기술연구원
연구기간	2013.09.01 ~ 2018.04.30

명 세 서

청구범위

청구항 1

영상에서 객체를 추적하는 추적기;

상기 영상에서 상기 객체를 검출하는 검출기;

상기 추적기의 추적 결과에 대한 추적 신뢰도 및 상기 검출기의 검출 결과에 대한 검출 신뢰도를 계산하는 계산부; 및

상기 추적 결과와 상기 추적 신뢰도 및 상기 검출 결과와 상기 검출 신뢰도를 기반으로 객체 위치를 결정하는 결정부;를 포함하는 것을 특징으로 하는 객체 추적 장치.

청구항 2

청구항 1에 있어서,

상기 객체의 움직임을 예측하여 객체 위치를 예측하는 예측기;를 더 포함하고,

상기 계산부는,

상기 예측 결과에 대한 예측 신뢰도를 더 계산하고,

상기 결정부는,

상기 예측 결과와 상기 예측 신뢰도를 더 참조하여, 상기 객체 위치를 결정하는 것을 특징으로 하는 객체 추적 장치.

청구항 3

청구항 2에 있어서,

상기 결정부는,

상기 추적 결과에 의한 객체 위치에 상기 추적 신뢰도를 가중하고, 상기 검출 결과에 의한 객체 위치에 상기 검출 신뢰도를 가중하며, 상기 예측 결과에 의한 객체 위치에 상기 예측 신뢰도를 가중하여, 신뢰도들이 가중된 위치들로부터 상기 객체 위치를 결정하는 것을 특징으로 하는 객체 추적 장치.

청구항 4

청구항 3에 있어서,

상기 결정부는,

대각선 성분들은 상기 추적 결과, 상기 검출 결과 및 상기 예측 결과와 추적 대상 객체의 유사도를 나타내는 성분들이고, 비대각선 성분들은 결과들 간의 거리를 나타내는 성분들인 행렬을 이용하여, 상기 신뢰도들을 계산하는 것을 특징으로 하는 객체 추적 장치.

청구항 5

청구항 4에 있어서,

상기 결정부는,

상기 행렬을 고유치 분해하여 가장 큰 고유치를 갖는 고유치 벡터를 찾아, 상기 신뢰도들을 계산하는 것을 특징으로 하는 객체 추적 장치.

청구항 6

영상에서 객체를 추적하는 단계;

상기 영상에서 상기 객체를 검출하는 단계;

추적 결과에 대한 추적 신뢰도 및 검출 결과에 대한 검출 신뢰도를 계산하는 단계; 및

상기 추적 결과와 상기 추적 신뢰도 및 상기 검출 결과와 상기 검출 신뢰도를 기반으로 객체 위치를 결정하는 단계;를 포함하는 것을 특징으로 하는 객체 추적 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 객체 추적에 관한 것으로, 더욱 상세하게는 관심 대상이 되는 객체를 추적하는 방법 및 장치에 관한 것이다.

배 경 기 술

[0002] 기존의 객체 추적에서는, 먼저 영상으로부터 객체를 검출하고, 검출된 결과와 추적하고자 하는 객체의 유사성을 측정하여, 추적 대상 객체를 결정한다. 하지만, 도 1에 나타난 바와 같이, 객체 검출의 경우 오검출률이 높고, 또는 찾고자하는 객체가 검출되지 않을 수 있기 때문에 추적의 성공을 보장할 수 없다.

[0003] 또한, 객체 추적기만을 이용하는 경우 추적 객체가 가려지게 되면 추적이 제대로 이루어지지 않으며, 추적 객체가 다시 나타난 경우 다시 객체를 추적하는 것이 어렵다.

[0004] 이에, 복잡한 환경이나 가려짐 발생에도 강인한 특성을 보이는 객체 추적을 위한 방안의 모색이 요청된다.

발명의 내용

해결하려는 과제

[0005] 본 발명은 상기와 같은 문제점을 해결하기 위하여 안출된 것으로서, 본 발명의 목적은, 객체 검출 결과와 추적 결과 및 결과들에 대한 신뢰도를 기반으로 객체 위치를 결정하는 객체 추적 방법 및 장치를 제공함에 있다.

과제의 해결 수단

[0006] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른, 객체 추적 장치는, 영상에서 객체를 추적하는 추적기; 상기 영상에서 상기 객체를 검출하는 검출기; 상기 추적기의 추적 결과에 대한 추적 신뢰도 및 상기 검출기의 검출 결과에 대한 검출 신뢰도를 계산하는 계산부; 및 상기 추적 결과와 상기 추적 신뢰도 및 상기 검출 결과와 상기 검출 신뢰도를 기반으로 객체 위치를 결정하는 결정부;를 포함한다.

[0007] 그리고, 본 발명의 일 실시예에 따른 객체 추적 장치는, 상기 객체의 움직임을 예측하여 객체 위치를 예측하는 예측기;를 더 포함하고, 상기 계산부는, 상기 예측 결과에 대한 예측 신뢰도를 더 계산하고, 상기 결정부는, 상기 예측 결과와 상기 예측 신뢰도를 더 참조하여, 상기 객체 위치를 결정할 수 있다.

[0008] 또한, 상기 결정부는, 상기 추적 결과에 의한 객체 위치에 상기 추적 신뢰도를 가중하고, 상기 검출 결과에 의한 객체 위치에 상기 검출 신뢰도를 가중하며, 상기 예측 결과에 의한 객체 위치에 상기 예측 신뢰도를 가중하

여, 신뢰도들이 가중된 위치들로부터 상기 객체 위치를 결정할 수 있다.

[0009] 그리고, 상기 결정부는, 대각선 성분들은 상기 추적 결과, 상기 검출 결과 및 상기 예측 결과와 추적 대상 객체의 유사도를 나타내는 성분들이고, 비대각선 성분들은 결과들 간의 거리를 나타내는 성분들인 행렬을 이용하여, 상기 신뢰도들을 계산

[0010] 또한, 상기 결정부는, 상기 행렬을 고유치 분해하여 가장 큰 고유치를 갖는 고유치 벡터를 찾아, 상기 신뢰도들을 계산할 수 있다.

[0011] 한편, 본 발명의 다른 실시예에 따른, 객체 추적 방법은, 영상에서 객체를 추적하는 단계; 상기 영상에서 상기 객체를 검출하는 단계; 추적 결과에 대한 추적 신뢰도 및 검출 결과에 대한 검출 신뢰도를 계산하는 단계; 및 상기 추적 결과와 상기 추적 신뢰도 및 상기 검출 결과와 상기 검출 신뢰도를 기반으로 객체 위치를 결정하는 단계;를 포함한다.

발명의 효과

[0012] 이상 설명한 바와 같이, 본 발명의 실시예들에 따르면, 객체 검출 결과와 추적 결과 및 결과들에 대한 신뢰도를 기반으로 객체 위치를 결정하여, 복잡한 환경에서도 정확한 객체 추적이 가능해진다. 나아가, 가려짐이 발생하는 경우에도 강인한 객체 추적을 할 수 있게 된다.

도면의 간단한 설명

[0013] 도 1은 종래의 방법에 의한 객체 추적 결과를 예시한 사진,

도 2는 본 발명의 일 실시예에 따른 객체 추적 장치의 블럭도,

도 3은 객체 추적/검출 결과들을 예시한 사진,

도 4는 결과들에 대한 신뢰도들을 예시한 그래프, 그리고,

도 5 내지 도 7은, 도 2에 도시된 객체 추적 장치에 의한 객체 추적 결과를 예시한 사진들이다.

발명을 실시하기 위한 구체적인 내용

[0014] 이하에서는 도면을 참조하여 본 발명을 보다 상세하게 설명한다.

[0015] 도 2는 본 발명의 일 실시예에 따른 객체 추적 장치의 블럭도이다. 본 발명의 실시예에 따른 객체 추적 장치는, 객체 검출 결과와 추적 결과를 이용하여 추적하고자 하는 객체의 위치를 결정한다.

[0016] 이를 위해, 본 발명의 실시예에 따른 객체 추적 장치는, 다양한 검출기와 추적기를 이용하여 획득한 추적 결과들에 대해 신뢰도를 계산하는데, 신뢰도는 추적 대상 객체의 외형 모델과의 유사도 및 추적 결과들 간의 공간적 유사도를 고려하여 계산한다. 객체 추적의 성공률을 높이기 위함이다.

[0017] 이와 같은 기능을 수행하는 본 발명의 실시예에 따른 객체 추적 장치는, 객체 추적기(Object Tracker)(110), 객체 검출기(Object Detector)(120), 속도 모델 기반 예측기(Velocity Model based Predictor)(130), 신뢰도 계산부(Confidence Estimator)(140) 및 객체 위치 결정부(150)를 포함한다.

[0018] 객체 추적기(110)는 영상에서 추적 대상 객체를 추적한다. 객체 추적기(110)는, 컬러 히스토그램을 이용한 입자 필터(particle filter) 기반의 객체 추적 기법, 외곽정보 특징량을 이용한 입자 필터 기반의 객체 추적 등을 사용가능하며, 그 밖의 다른 객체 추적 기법을 더 사용하는 것을 배제하지 않는다.

[0019] 객체 검출기(120)는 영상에서 추적 대상 객체를 검출한다. 객체 검출기(120)로, HoG 기법, SVM 기법 등을 사용하며, 추적 대상 객체와 비슷한 외형을 갖는 객체를 영상으로부터 찾는다.

[0020] 속도 모델 기반 예측기(130)는 이전의 객체 추적 결과들을 이용하여 객체의 움직임을 예측함으로써, 추적 대상 객체의 위치를 예측한다.

[0021] 객체 추적기(110), 객체 검출기(120) 및 속도 모델 기반 예측기(130)에 의해 추적 대상 객체에 대한 위치 정보

들과 크기 정보들이 획득된다.

[0022] 신뢰도 계산부(140)는 객체 추적기(110)의 추적 결과에 대한 추적 신뢰도, 객체 검출기(120)의 검출 결과에 대한 검출 신뢰도, 속도 모델 기반 예측기(130)의 예측 결과에 대한 예측 신뢰도를 계산한다.

[0023] 이를 위해, 신뢰도 계산부(140)는 먼저 아래의 수학식 1에 제시된 행렬 S를 생성한다.

[0024] [수학식 1]

[0025] $$S = \begin{bmatrix} S_{1,1} & S_{1,2} & S_{1,3} & \cdots & S_{1,N} \\ S_{2,1} & S_{2,2} & S_{2,3} & & S_{2,N} \\ S_{3,1} & S_{3,2} & S_{3,3} & & S_{3,N} \\ \\ S_{N,1} & S_{N,2} & S_{N,3} & & S_{N,N} \end{bmatrix}$$

[0026] $$S_{i,j} = \begin{cases} \text{Photometric similarity} \\ p(z_t^s \mid \theta_i) p(z_t^c \mid \theta_i) c(\theta_i) & \text{if } i = j \\ \text{Spatial consistency} \\ \lambda \exp\left(-\left|\theta_i - \theta_j\right| / \sigma\right) & \text{else if } \min(i,j) \le 3 \\ 0 & \text{else} \end{cases}$$

[0027] 행렬 S의 대각선 성분들은, 추적 대상 객체의 외곽 정보 모델, 컬러 모델 및 학습된 분류기를 기반으로 계산한 유사도를 나타낸다. 또한, 행렬 S의 비대각선 성분들은 추적 결과, 검출 결과 및 예측 결과 간의 거리 차이를 나타낸다.

[0028] 이에, 행렬 S의 대각선 성분들은 추적 대상 객체에 대한 외형의 유사도를, 행렬 S의 비대각선 성분들은 추적/검출/예측 결과들 간의 공간적 유사도를 각각 나타낸다. 이에, 추적 대상 객체와 추적/검출/예측 결과가 유사할 경우 대각선 성분들은 큰 값을 가지게 된다. 또한, 다른 결과와 공간적으로 유사할 경우 비대각선 성분의 값이 커지게 된다.

[0029] 이에 따라, 각 추적/검출/예측 결과들에 대한 신뢰도 ω는 아래의 수학식 2를 이용하여 $\omega^T S \omega$를 최대로 하는 벡터를 계산하여, 각 결과들에 대한 신뢰도를 결정할 수 있다.

[0030] [수학식 2]

[0031] $$w^* = \arg max\,(w^T S w)$$

[0032] 신뢰도 ω 계산은, 먼저 행렬 S를 고유치 분해(eigenvalue decomposition)하고, 이로부터 가장 큰 고유치(eigenvalue)를 갖는 고유치 벡터(eigenvector)를 찾는 과정에 의한다.

[0033] 이에 의해, 객체 추적기(110)의 추적 결과에 대한 추적 신뢰도, 객체 검출기(120)의 검출 결과에 대한 검출 신뢰도, 속도 모델 기반 예측기(130)의 예측 결과에 대한 예측 신뢰도가 계산된다.

[0034] 객체 위치 결정부(150)는 추적 결과와 추적 신뢰도, 검출 결과와 검출 신뢰도, 예측 결과와 예측 신뢰도를 기반으로 객체 위치를 결정한다. 구체적으로, 객체 위치 결정부(150)는, 추적 결과에 의한 객체 위치에 추적 신뢰도를 가중하고, 검출 결과에 의한 객체 위치에 검출 신뢰도를 가중하며, 예측 결과에 의한 객체 위치에 예측 신뢰도를 가중하여, 신뢰도들이 가중된 위치들로부터 최종 객체 위치를 결정한다.

[0035] 즉, 최종 객체 위치 $\hat{x}$는 아래의 수학식 3으로 나타낼 수 있다.

[0036] [수학식 3]

$$\hat{x} = \sum_{i=1}^{N} w_i x_i$$

[0037]

[0038] 여기서, x_i는 추적/검출/예측 결과에 의한 객체 위치를 나타내며, w_i는 추적/검출/예측 신뢰도를 나타낸다.

[0039] 도 2에 도시된 객체 추적 장치를 이용한 객체 추적/검출/예측 결과들을 도 3에 예시하였고, 이 결과들에 대한 신뢰도들을 도 4에 예시하였다.

[0040] 한편, 도 2에 도시된 객체 추적 장치를 이용한 객체 추적 결과를 도 5 내지 도 7에 제시하였다. 도 5 내지 도 7에 제시된 바와 같이, 도 2에 도시된 객체 추적 장치는 가려짐이 발생하거나 복잡한 환경에서도 강인하게 대상 객체를 정확하게 추적함을 확인할 수 있다.

[0041] 또한, 이상에서는 본 발명의 바람직한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

부호의 설명

[0042] 110 : 추적기(Object Tracker)

120 : 객체 검출기(Object Detector)

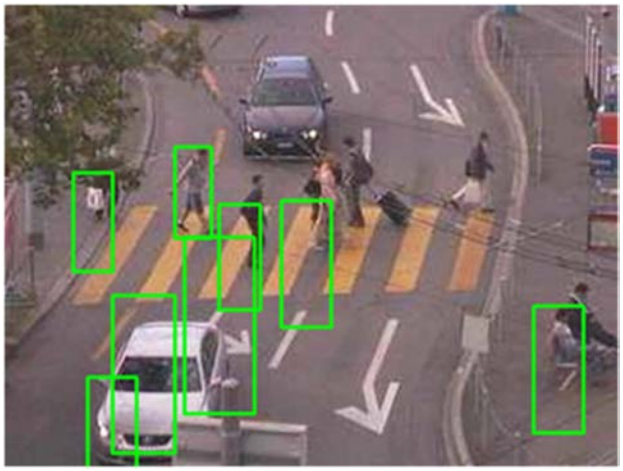
130 : 속도 모델 기반 예측기(Velocity Model based Predictor)

140 : 신뢰도 계산부(Confidence Estimator)

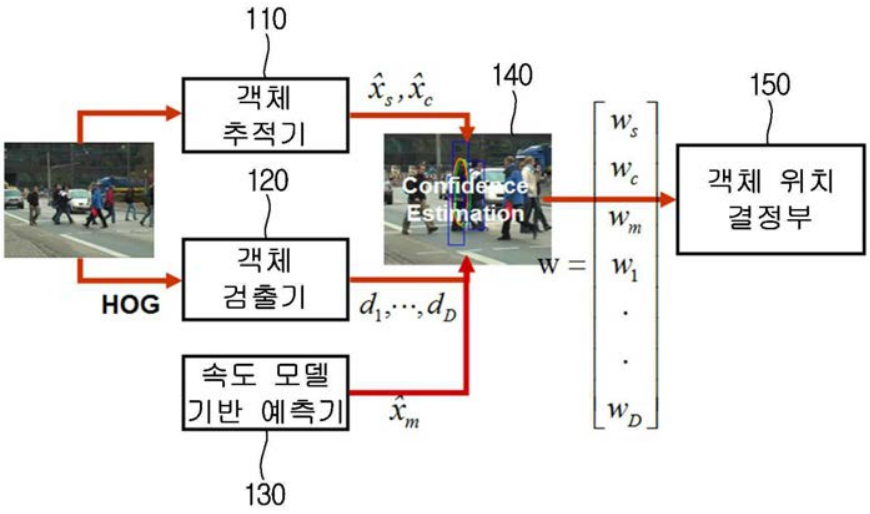
150 : 객체 위치 결정부

도면

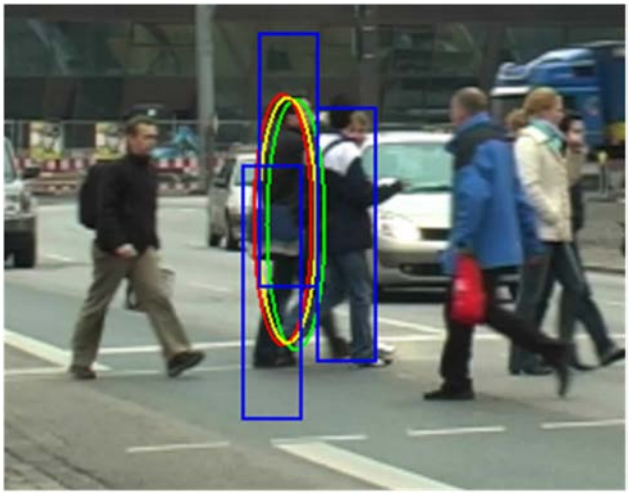
도면1



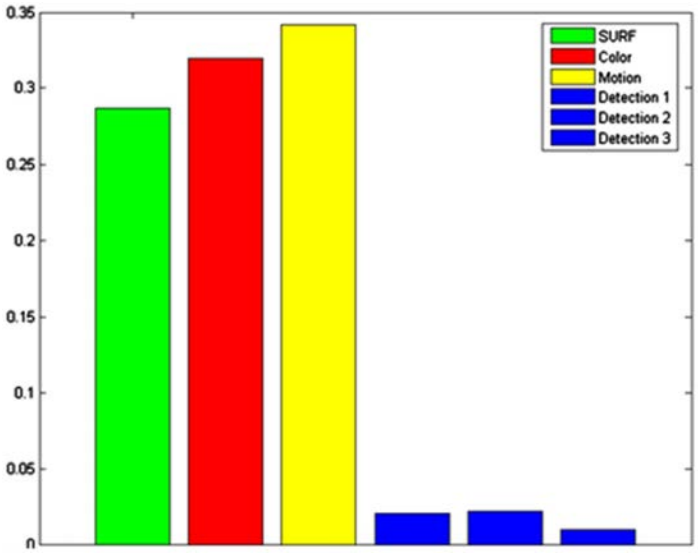
도면2


110
객체
추적기
$\hat{x}_s, \hat{x}_c$
140
150
120
객체
검출기
HOG
$d_1, \cdots, d_D$
Confidence
Estimation
W =
w_s
w_c
w_m
w_1
w_D
객체 위치
결정부
속도 모델
기반 예측기
$\hat{x}_m$
130

도면3



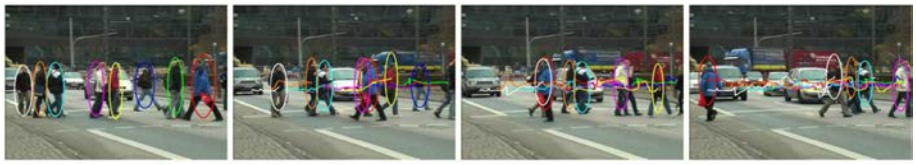
도면4



도면5

도면6

도면7



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년11월08일
(11) 등록번호 10-1916460
(24) 등록일자 2018년11월01일

(51) 국제특허분류(Int. Cl.)
G06T 7/00 (2017.01) *G06K 9/46* (2006.01)
(21) 출원번호 10-2012-0051871
(22) 출원일자 2012년05월16일
심사청구일자 2017년03월08일
(65) 공개번호 10-2013-0128097
(43) 공개일자 2013년11월26일
(56) 선행기술조사문헌
JP2012083855 A*
KR1020050082252 A*
KR1020110047814 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
황영배
서울 서초구 방배선행길 1, 106동 607호 (방배동, 방배우성아파트)
배주한
서울 광진구 동일로26길 52, (화양동)
(뒷면에 계속)
(74) 대리인
남충우, 노철호

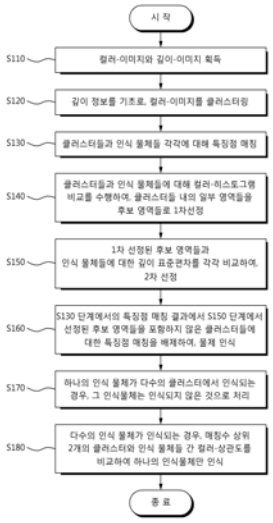
전체 청구항 수 : 총 6 항 심사관 : 박상철

(54) 발명의 명칭 **깊이 정보를 이용한 물체 인식 방법 및 장치**

(57) 요 약

깊이 정보를 이용한 물체 인식 방법 및 장치가 제공된다. 본 물체 인식 방법은, 컬러-이미지와 깊이-이미지를 획득하는 단계, 깊이-이미지에 나타난 깊이 정보를 기초로, 컬러-이미지를 다수의 클러스터들로 클러스터링하는 단계 및 클러스터링된 클러스터들을 기반으로 물체 인식을 수행하는 단계를 포함한다. 이에 의해, 특징점 기반의 물체 인식에서 물체의 특징점이 별로 없는 경우나 복잡 배경을 가진 물체의 경우에도, 보다 신뢰성 있게 물체 인식율을 높일 수 있게 된다.

대 표 도 - 도1



(72) 발명자

최병호

경기 용인시 수지구 대지로 27, 103동 1306호 (죽전동, 한신아파트)

김제우

경기 성남시 분당구 황새울로 54, 320동 603호 (정자동, 상록마을우성아파트)

이 발명을 지원한 국가연구개발사업

과제고유번호 10040246

부처명 지식경제부

연구관리전문기관 한국산업기술평가관리원

연구사업명 산업원천기술개발사업

연구과제명 이동 로봇의 안정적 영상 획득을 통한 3D Depth 정보 획득과 실시간 객체 인식을 위한 로봇비전 soC 및 모듈 개발

기 여 율 1/1

주관기관 전자부품연구원

연구기간 2011.06.01 ~ 2015.05.31

명 세 서

청구범위

청구항 1

컬러-이미지와 깊이-이미지를 획득하는 단계;

깊이-이미지에 나타난 깊이 정보를 기초로, 컬러-이미지를 다수의 클러스터들로 클러스터링하는 단계; 및

클러스터링된 클러스터들을 기반으로, 물체 인식을 수행하는 단계;를 포함하고,

물체 인식 수행 단계는,

클러스터들 각각과 인식 물체들 각각에 대해 특징점 매칭을 수행하는 단계;

클러스터들 내의 일부 영역들을 후보 영역들로 선정하는 단계; 및

선정된 후보 영역들을 포함하는 클러스터들 밖에서 이루어진 특징점 매칭을 배제하고, 물체 인식을 수행하는 단계;를 포함하며,

후보 영역 선정 단계는,

클러스터들 각각에 대한 컬러-히스토그램들과 인식 물체들 각각에 대한 컬러-히스토그램들을 비교하여, 클러스터들 내의 일부 영역들을 후보 영역들로 선정하는 제1 선정 단계; 및

제1 선정 단계에서 선정된 후보 영역을 구성하는 픽셀들의 깊이 값들에 대한 표준편차와 인식 물체의 이미지를 구성하는 픽셀들의 깊이 값들에 대한 표준편차의 차가 임계치 미만인 후보 영역과 인식 물체를 선정하는 제2 선정단계;를 포함하는 것을 특징으로 하는 물체 인식 방법.

청구항 2

삭제

청구항 3

제 1항에 있어서,

물체 인식 수행 단계는,

클러스터들 내의 일부 영역들을 후보 영역들로 선정하는 단계;

후보 영역들 각각과 인식 물체들 각각에 대해 특징점 매칭을 수행하는 단계; 및

특징점 매칭 결과를 이용하여, 물체 인식을 수행하는 단계;를 더 포함하는 것을 특징으로 하는 물체 인식 방법.

청구항 4

삭제

청구항 5

제 1항에 있어서,

물체 인식 수행 단계는,

클러스터들 내의 일부 영역들을 후보 영역들로 선정하는 제3 선정 단계;

제3 선정단계에서 선정된 후보 영역들 각각과 인식 물체들 각각에 대해 특징점 매칭을 수행하는 단계;

제3 선정단계에서 선정된 후보 영역들 중 일부를 다시 선정하는 제4 선정 단계; 및

제4 선정 단계에서 선정된 후보 영역들 밖에서 이루어진 특징점 매칭을 배제하고, 물체 인식을 수행하는 단계;를 더 포함하는 것을 특징으로 하는 물체 인식 방법.

청구항 6

제 1항에 있어서,

다수의 클러스터에서 인식된 인식 물체는 인식되지 않은 것으로 처리하는 단계;를 더 포함하는 것을 특징으로 하는 물체 인식 방법.

청구항 7

제 1항에 있어서,

컬러-이미지에서 다수의 인식 물체가 인식되는 경우, 클러스터와 인식 물체들 간의 컬러-상관도를 비교하여, 하나의 인식 물체를 선정하여 인식된 것으로 처리하는 단계;를 더 포함하는 것을 특징으로 하는 물체 인식 방법.

청구항 8

컬러-이미지와 깊이-이미지를 획득하는 카메라;

인식 물체들의 이미지가 DB화 되어 있는 저장부; 및

깊이-이미지에 나타난 깊이 정보를 기초로 컬러-이미지를 다수의 클러스터들로 클러스터링하고, 클러스터링된 클러스터들을 기반으로 물체 인식을 수행하는 프로세서;를 포함하고,

프로세서는,

클러스터들 각각과 인식 물체들 각각에 대해 특징점 매칭을 수행하며, 클러스터들 내의 일부 영역들을 후보 영역들로 선정하고, 선정된 후보 영역들을 포함하는 클러스터들 밖에서 이루어진 특징점 매칭을 배제하고, 물체 인식을 수행하며,

클러스터들 각각에 대한 컬러-히스토그램들과 인식 물체들 각각에 대한 컬러-히스토그램들을 비교하여 클러스터들 내의 일부 영역들을 후보 영역들로 선정하고, 선정된 후보 영역을 구성하는 픽셀들의 깊이 값들에 대한 표준편차와 인식 물체의 이미지를 구성하는 픽셀들의 깊이 값들에 대한 표준편차의 차가 임계치 미만인 후보 영역과 인식 물체를 선정하는 것을 특징으로 하는 물체 인식 장치.

발명의 설명

기 술 분 야

[0001] 본 발명은 물체 인식 방법 및 장치에 관한 것으로, 더욱 상세하게는 깊이 정보를 이용한 물체 인식 방법 및 장치에 관한 것이다.

배 경 기 술

[0002] 촬영된 이미지에서 찾고자 하는 물체가 존재하는지 여부를 인식하기 위한 방법으로 특징점 매칭 기법이 주로 이용되고 있다.

[0003] 이 기법은 이미지 상에서 찾고자 하는 물체의 존재여부는 물론 존재 위치까지 비교적 정확하게 알아낼 수 있지만, 경우에 따라 정확도가 떨어지는 문제를 드러내고 있다.

[0004] 구체적으로, 배경에 의해 잘못된 물체 인식이 발생할 수 있는데, 이는 배경이 복잡할수록 더욱 그러하다. 뿐만 아니라, 특징 기술자가 식별력이 적은 경우나, 인식 물체가 균일한 경우에도 잘못된 특징점 대응이 나타나, 부정확한 인식을 유발한다.

[0005] 한편, 물체 인식을 위한 특징점 매칭은 이미지 전체에 대해 수행되기 때문에 처리량의 증가로 인한 인식 속도 저하를 유발하는데, 이미지의 해상도가 높아지는 추세를 고려한다면 속도 저하는 더욱 가중될 것이라는데 문제가 있다.

발명의 내용

해결하려는 과제

[0006] 본 발명은 상기와 같은 문제점을 해결하기 위하여 안출된 것으로서, 본 발명의 목적은, 특징점 기반의 물체 인식에서 물체의 특징점이 별로 없는 경우나 복잡 배경을 가진 물체의 경우에도, 보다 신뢰성 있게 물체 인식율을 높일 수 있는 물체 인식 방법 및 장치를 제공함에 있다.

과제의 해결 수단

[0007] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른, 물체 인식 방법은, 컬러-이미지와 깊이-이미지를 획득하는 단계; 깊이-이미지에 나타난 깊이 정보를 기초로, 컬러-이미지를 다수의 클러스터들로 클러스터링하는 단계; 및 클러스터링된 클러스터들을 기반으로, 물체 인식을 수행하는 단계;를 포함한다.

[0008] 그리고, 물체 인식 수행 단계는, 클러스터들 각각과 인식 물체들 각각에 대해 특징점 매칭을 수행하는 단계; 클러스터들 내의 일부 영역들을 후보 영역들로 선정하는 단계; 및 선정된 후보 영역들을 포함하는 클러스터들 밖에서 이루어진 특징점 매칭을 배제하고, 물체 인식을 수행하는 단계;를 포함할 수 있다.

[0009] 또한, 물체 인식 수행 단계는, 클러스터들 내의 일부 영역들을 후보 영역들로 선정하는 단계; 후보 영역들 각각과 인식 물체들 각각에 대해 특징점 매칭을 수행하는 단계; 및 특징점 매칭 결과를 이용하여, 물체 인식을 수행하는 단계;를 포함할 수 있다.

[0010] 그리고, 후보 영역 선정 단계는, 클러스터들 각각에 대한 컬러-히스토그램들과 인식 물체들 각각에 대한 컬러-히스토그램들을 비교하여, 클러스터들 내의 일부 영역들을 후보 영역들로 선정하는 제1 선정 단계; 및 제1 선정 단계에서 선정된 후보 영역과 인식 물체의 깊이 통계치들을 비교하고, 통계치 차가 임계치 미만인 후보 영역과 인식 물체를 선정하는 제2 선정단계;를 포함할 수 있다.

[0011] 또한, 물체 인식 수행 단계는, 클러스터들 내의 일부 영역들을 후보 영역들로 선정하는 제1 선정 단계; 제1 선정단계에서 선정된 후보 영역들 각각과 인식 물체들 각각에 대해 특징점 매칭을 수행하는 단계; 제1 선정단계에서 선정된 후보 영역들 중 일부를 다시 선정하는 제2 선정 단계; 및 제2 선정 단계에서 선정된 후보 영역들 밖에서 이루어진 특징점 매칭을 배제하고, 물체 인식을 수행하는 단계;를 포함할 수 있다.

[0012] 그리고, 다수의 클러스터에서 인식된 인식 물체는 인식되지 않은 것으로 처리하는 단계;를 더 포함할 수 있다.

[0013] 또한, 컬러-이미지에서 다수의 인식 물체가 인식되는 경우, 클러스터와 인식 물체들 간의 컬러-상관도를 비교하여, 하나의 인식 물체를 선정하여 인식된 것으로 처리하는 단계;를 더 포함할 수 있다.

[0014] 한편, 본 발명의 다른 실시예에 따른, 물체 인식 장치는, 컬러-이미지와 깊이-이미지를 획득하는 카메라; 인식 물체들의 이미지가 DB화 되어 있는 저장부; 및 깊이-이미지에 나타난 깊이 정보를 기초로 컬러-이미지를 다수의 클러스터들로 클러스터링하고, 클러스터링된 클러스터들을 기반으로 물체 인식을 수행하는 프로세서;를 포함한다.

발명의 효과

[0015] 이상 설명한 바와 같이, 본 발명에 따르면, 특징점 기반의 물체 인식에서 물체의 특징점이 별로 없는 경우나 복

잡 배경을 가진 물체의 경우에도, 보다 신뢰성 있게 물체 인식율을 높일 수 있게 된다.

[0016] 구체적으로, 깊이 정보를 이용하여 이미지를 클러스터링하고, 컬러-히스토그램 비교를 통해 클러스터 내의 후보 영역을 선정하여 영상 인식에 이용하기 때문에 물체 인식율이 높아짐은 물론, 특징점 매칭 및 그 후속 처리에 소요되는 연산량을 크게 줄일 수 있게 된다.

[0017] 뿐만 아니라, 후보 영역과 인식 물체 간의 깊이 통계치 비교를 통해, 후보 영역을 검증할 수 있게 되어, 물체 인식율을 더욱 높일 수 있게 된다.

[0018] 아울러, 깊이 정보를 이용하더라도 인식이 어려울 수 있는 비슷한 물체에 대해서는 컬러 상관도를 보조 지표로 활용하여 유사도를 더 측정하므로 인식율을 더욱 증가시킬 수 있게 된다.

도면의 간단한 설명

[0019] 도 1은 본 발명의 일 실시예에 따른 깊이 정보를 이용한 물체 인식 방법의 설명에 제공되는 흐름도,

도 2는, 컬러-이미지, 깊이-이미지 및 Depth Proximity Clustering 결과를 나타낸 도면,

도 3은, 클러스터들 중에서 선정한 3개의 후보 영역을 컬러-이미지에 나타낸 도면,

도 4는 인식 물체들의 깊이 표준 편차를 예시한 도면,

도 5와 도 6은, 하나의 인식 물체가 다수의 클러스터에서 인식되는 경우를 예시한 도면,

도 7과 도 8은, 컬러-이미지에서 다수의 인식 물체가 인식되는 경우를 에 예시한 도면,

도 9는 본 발명의 따른 실시예에 따른 깊이 정보를 이용한 물체 인식 방법의 설명에 제공되는 흐름도,

도 10은 본 발명의 또 다른 실시예에 따른 깊이 정보를 이용한 물체 인식 방법의 설명에 제공되는 흐름도, 그리고,

도 11은, 도 1, 도 9 및 도 10에 도시된 물체 인식 방법을 수행할 수 있는 물체 인식 장치를 도시한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0020] 이하에서는 도면을 참조하여 본 발명을 보다 상세하게 설명한다.

[0021] **1. 깊이 정보를 이용한 물체 인식 #1**

[0022] 본 발명의 일 실시예에 따른 깊이 정보를 이용한 물체 인식 방법에서는 컬러-이미지를 클러스터링하고 클러스터 내에 인식하고자 하는 물체가 존재할 가능성이 높은 후보로 선정된 영역을 검증함에 있어 깊이 정보를 이용한다.

[0023] 도 1은 본 발명의 일 실시예에 따른 깊이 정보를 이용한 물체 인식 방법의 설명에 제공되는 흐름도이다.

[0024] 도 1에 도시된 바와 같이, 먼저 컬러-이미지와 깊이-이미지를 획득한다(S110). 컬러-이미지와 깊이-이미지는 RGB-카메라와 D-카메라를 이용하여 각각 생성가능하다.

[0025] 이후, 깊이-이미지에 나타난 깊이 정보를 기초로, 컬러-이미지를 다수의 클러스터들로 클러스터링한다(S120). 엄밀하게, S120단계에서의 클러스터링에서는 깊이 정보 외에 위치 정보가 더 반영되는데, 반영율은 깊이 정보가 더 크다. 여기서, 위치 정보는 깊이(d)에 수직한 평면상의 위치(x,y)에 대한 정보를 말한다.

[0026] S120단계에서 수행되는 클러스터링은 깊이가 비슷하고 거리가 가까운 픽셀들을 하나의 클러스터에 클러스터링하는 것이다.

[0027] S120단계에 의해, 컬러-이미지 상에 존재하는 '인식하고자 하는 물체'(이하, '인식 물체'로 약칭)는 어느 한 클러스터에 소속될 가능성이 매우 높다. 이는, 하나의 인식 물체를 구성하는 픽셀들의 깊이는 비슷하고 거리가 가깝기 때문이다.

[0028] 인식 물체는 컬러-이미지에서 찾고자 하는 물체로, DB에 이미지가 저장되어 있다.

[0029] 다음, 클러스터들 각각과 인식 물체들 각각에 대해 특징점 매칭을 수행한다(S130). 클러스터링된 클러스터가 k개이고, DB에 구축되어 있는 인식 물체의 개수가 n개인 경우, S130단계는,

[0030] 11) 클러스터-1과 인식 물체-1 간의 특징점 매칭을 수행하고,

[0031] 12) 클러스터-1과 인식 물체-2 간의 특징점 매칭을 수행하고,

[0032] ...

[0033] 1n) 클러스터-1과 인식 물체-n 간의 특징점 매칭을 수행하고,

[0034] 21) 클러스터-2와 인식 물체-1 간의 특징점 매칭을 수행하고,

[0035] 22) 클러스터-2와 인식 물체-2 간의 특징점 매칭을 수행하고,

[0036] ...

[0037] 2n) 클러스터-2와 인식 물체-n 간의 특징점 매칭을 수행하고,

[0038] ...

[0039] k1) 클러스터-k와 인식 물체-1 간의 특징점 매칭을 수행하고,

[0040] k2) 클러스터-k와 인식 물체-2 간의 특징점 매칭을 수행하고,

[0041] ...

[0042] kn) 클러스터-k와 인식 물체-n 간의 특징점 매칭을 수행하는 절차로 수행된다.

[0043] 이후, 클러스터들 내의 일부 영역들을 후보 영역들로 1차 선정한다(S140). S140단계는, S120단계에서 클러스터링된 클러스터들 각각에 대해 수행되며, 후보 영역 선정은 클러스터와 인식 물체의 컬러-히스토그램 비교를 통해 수행된다.

[0044] 구체적으로 S140단계는,

[0045] 11) 클러스터-1에서 인식 물체-1의 컬러-히스토그램과 유사한 컬러-히스토그램을 갖는 영역이 존재하는 경우, 그 영역을 후보 영역으로 선정하고,

[0046] 12) 클러스터-1에서 인식 물체-2의 컬러-히스토그램과 유사한 컬러-히스토그램을 갖는 영역이 존재하는 경우, 그 영역을 후보 영역으로 선정하고,

[0047] ...

[0048] 1n) 클러스터-1에서 인식 물체-n의 컬러-히스토그램과 유사한 컬러-히스토그램을 갖는 영역이 존재하는 경우, 그 영역을 후보 영역으로 선정하고,

[0049] 21) 클러스터-2에서 인식 물체-1의 컬러-히스토그램과 유사한 컬러-히스토그램을 갖는 영역이 존재하는 경우, 그 영역을 후보 영역으로 선정하고,

[0050] 22) 클러스터-2에서 인식 물체-2의 컬러-히스토그램과 유사한 컬러-히스토그램을 갖는 영역이 존재하는 경우, 그 영역을 후보 영역으로 선정하고,

[0051] ...

[0052] 2n) 클러스터-2에서 인식 물체-n의 컬러-히스토그램과 유사한 컬러-히스토그램을 갖는 영역이 존재하는 경우, 그 영역을 후보 영역으로 선정하고,

[0053] ...

[0054] k1) 클러스터-k에서 인식 물체-1의 컬러-히스토그램과 유사한 컬러-히스토그램을 갖는 영역이 존재하는 경우, 그 영역을 후보 영역으로 선정하고,

[0055] k2) 클러스터-k에서 인식 물체-2의 컬러-히스토그램과 유사한 컬러-히스토그램을 갖는 영역이 존재하는 경우, 그 영역을 후보 영역으로 선정하고,

[0056] ...

[0057] kn) 클러스터-k에서 인식 물체-n의 컬러-히스토그램과 유사한 컬러-히스토그램을 갖는 영역이 존재하는 경우, 그 영역을 후보 영역으로 선정한다.

[0058] 다음, S140단계에서 1차 선정된 후보 영역과 인식 물체 간의 깊이 표준 편차를 각각 비교하여, 후보 영역들을 2차로 선정한다(S150).

[0059] 여기서, 깊이 표준 편차는 픽셀들에 대한 깊이 값들의 표준 편차로, 후보 영역의 깊이 표준 편차는 후보 영역을 구성하는 픽셀들의 깊이 값들에 대한 표준편차이고, 인식 물체의 깊이 표준 편차는 인식 물체의 이미지를 구성하는 픽셀들의 깊이 값들에 대한 표준편차이다.

[0060] 인식 물체의 깊이 표준 편차는 미리 계산하여 DB에 저장해 놓는 것이 바람직하다. 한편, 표준 편차를 이외의 다른 통계치, 예를 들면, 분산, 평균, 중간값 등으로 대체하는 것도 가능하다.

[0061] S140단계에서,

[0062] 1) 클러스터-1의 일부 영역이 '인식 물체-1'과 컬러-히스토그램이 유사하여 '후보 영역-1'로 선정되었고,

[0063] 2) 클러스터-3의 일부 영역이 '인식 물체-1'과 컬러-히스토그램이 유사하여 '후보 영역-2'로 선정되었고,

[0064] 3) 클러스터-2의 일부 영역이 '인식 물체-2'와 컬러-히스토그램이 유사하여 '후보 영역-3'으로 선정되었으며,

[0065] 4) 클러스터-4의 일부 영역이 '인식 물체-2'와 컬러-히스토그램이 유사하여 '후보 영역-4'으로 선정되었으며,

[0066] 5) 클러스터-3의 일부 영역이 '인식 물체-3'과 컬러-히스토그램이 유사하여 '후보 영역-5'로 선정되었고,

[0067] 6) 클러스터-5의 일부 영역이 '인식 물체-4'와 컬러-히스토그램이 유사하여 '후보 영역-6'으로 선정었으며,

[0068] 7) 클러스터-6의 일부 영역이 '인식 물체-4'와 컬러-히스토그램이 유사하여 '후보 영역-7'로 선정된 경우를 가정하면,

[0069] S150에서는,

[0070] 1) 후보 영역-1의 깊이 표준 편차와 인식 물체-1의 깊이 표준 편차의 차가 임계치 미만이면, '후보 영역-1과 인식 물체-1'을 선정하고,

[0071] 2) 후보 영역-2의 깊이 표준 편차와 인식 물체-1의 깊이 표준 편차의 차가 임계치 미만이면, '후보 영역-2와 인식 물체-1'을 선정하며,

[0072] 3) 후보 영역-3의 깊이 표준 편차와 인식 물체-2의 깊이 표준 편차의 차가 임계치 미만이면, '후보 영역-3과 인식 물체-2'를 선정하며,

[0073] 4) 후보 영역-4의 깊이 표준 편차와 인식 물체-2의 깊이 표준 편차의 차가 임계치 미만이면, '후보 영역-4와 인식 물체-2'를 선정하고,

[0074] 5) 후보 영역-5의 깊이 표준 편차와 인식 물체-3의 깊이 표준 편차의 차가 임계치 미만이면, '후보 영역-5와 인식 물체-3'를 선정하며,

[0075] 6) 후보 영역-6의 깊이 표준 편차와 인식 물체-4의 깊이 표준 편차의 차가 임계치 미만이면, '후보 영역-6과 인식 물체-4'를 선정하고,

[0076] 7) 후보 영역-7의 깊이 표준 편차와 인식 물체-4의 깊이 표준 편차의 차가 임계치 미만이면, '후보 영역-7과 인식 물체-4'를 선정하게 된다.

[0077] 이후, S130단계에서의 특징점 매칭 결과에서 S150단계에서 선정된 후보 영역들을 포함하지 않는 클러스터들에 대한 특징점 매칭을 배제하고, 물체 인식을 수행한다(S160). 즉, S130단계에서의 특징점 매칭 결과 중 S150단계에서 선정된 후보 영역들을 포함하는 클러스터들에 대한 특징점 매칭만을 남기고, 물체 인식을 수행한다.

[0078] S150단계에서, '후보 영역-1과 인식 물체-1', '후보 영역-2와 인식 물체-1', '후보 영역-3과 인식 물체-2', '후보 영역-5와 인식 물체-3' 및 '후보 영역-6과 인식 물체-4'가 선정된 경우,

[0079] 후보 영역-1은 클러스터-1에 포함되고, 후보 영역-3은 클러스터-2에 포함되며, 후보 영역-2와 후보 영역-5는 클러스터-3에 포함되고, 후보 영역-6은 클러스터-5에 포함되므로,

[0080] S160단계에서는, 클러스터-1, 클러스터-2, 클러스터-3 및 클러스터-5에 대한 특징점 매칭만을 남기고, 나머지 클러스터들에 대한 특징점 매칭을 배제하고, 물체 인식을 수행한다

[0081] 한편, 하나의 인식 물체가 다수의 클러스터에서 인식되는 경우, 그 인식 물체는 인식되지 않은 것으로 처리할 수 있다(S170). 구체적으로, 특징점 매칭수가 첫 번째로 많은 클러스터와 두 번째로 많은 클러스터 간에 특징점 매칭수의 비율이 특정 범위 내인 인식 물체는, 인식되지 않은 것으로 처리하는 것이다.

[0082] 예를 들어, 인식 물체-1이 클러스터-1과의 특징점 매칭수(①)가 "100"이고, 클러스터-3과의 특징점 매칭수(②)가 "90"인 경우와 같이, 두 번째 매칭수(②)가 첫 번째 매칭수(①)의 80% 이상이서 인식 물체-1이 클러스터-1과 클러스터-3 모두에서 인식된 것으로 취급할 수 있는 경우, 인식 물체-1은 인식되지 않은 것으로 처리한다.

[0083] 동일한 인식 물체가 다수의 클러스터에 인식되는 것은 실제 그런 것보다, 영상 인식 과정에서의 오류에 기인했을 확률이 더 높기 때문이며, 잘못된 인식을 미연에 방지하기 위함이다.

[0084] 또한, 컬러-이미지의 적어도 하나의 클러스터에서 다수의 인식 물체가 인식되는 경우, 특징점 매칭수 상위 2개의 클러스터와 인식 물체들 간의 컬러-상관도를 비교하여 하나의 인식 물체만을 선정하고, 선정된 인식 물체가 인식된 것으로 처리한다(S180).

[0085] 예를 들어, 인식 물체-2와 클러스터-2의 특징점 매칭수(③)가 "90"이고, 인식 물체-3과 클러스터-3의 특징점 매칭수(④)가 "60"이며, 인식 물체-4와 클러스터-5의 특징점 매칭수(⑤)가 "70"인 경우와 같이, 두 번째 매칭수(⑤)가 첫 번째 매칭수(③)의 70% 이상이서 컬러-이미지에서 인식 물체-2와 인식 물체-4가 모두 인식된 것으로 취급할 수 있는 경우이다.

[0086] 이 경우에는, '인식 물체-2와 클러스터-2의 컬러-상관도'와 '인식 물체-4와 클러스터-5의 컬러-상관도'를 비교하고, 컬러-상관도가 큰 인식 물체가 컬러-이미지에서 인식된 것으로 처리하는 것이다.

[0087] 만약, 인식 물체-2와 클러스터-2의 특징점 매칭수(③)가 "90"이고, 인식 물체-3과 클러스터-3의 특징점 매칭수(④)가 "40"이며, 인식 물체-4와 클러스터-5의 특징점 매칭수(⑤)가 "30"인 경우와 같이, 두 번째 매칭수(④)가 첫 번째 매칭수(③)의 70% 미만인 경우에는, 컬러-상관도 비교 없이, 매칭수가 가장 많은 인식 물체-2가 컬러-이미지의 클러스터-2에서 인식된 것으로 처리한다.

[0088] 지금까지, 깊이 정보를 이용한 물체 인식 방법의 전체 과정에 대해 상세히 설명하였다. 이하에서, 물체 인식 방법을 구성하는 각 단계들에 대해 상세히 설명한다.

[0089] **2. Depth Proximity Clustering(도 1의 S120)**

[0090] Depth Proximity Clustering은 S120단계에서의 클러스터링을 명명한 것으로, 복잡한 배경으로 인해 발생하는 특징점 매칭의 부정확도를 줄이기 위해, 컬러-이미지를 물체와 배경으로 분리하여 클러스터링한다.

[0091] 클러스터링은 K-means 클러스터링을 응용하여, 깊이 정보와 위치 정보를 참조로 깊이가 비슷하고 거리가 가까운 픽셀들이 하나의 클러스터에 클러스터링할 수 있는데, 구체적으로 아래의 수학식 1에 따라 수행가능하다.

수학식 1

[0092] $$\arg_S \min \sum_{i}^{k} \sum_{x_j \in S_j} \| X_j - \mu_i \|$$

[0093] 여기서, k는 클러스터링의 개수, μ_i는 i번째 클러스터의 S_i의 기하학적 중심, X_j는 $[\alpha_1 x, \alpha_2 y, \alpha_3 d]$로 표현되는 벡터이다. x와 y는 픽셀의 x좌표와 y좌표이고, d는 깊이 값이다. α_1, α_2 및 α_3은 가중치로, 픽셀의 위치(x,y) 보다 깊이에 의한 영향을 더 반영하기 위해, α_1=0.05, α_2 = 0.05 및 α_3 = 0.9로 설정함이 바람직하지만, 필요에 따라 변경가능함은 물론이다.

[0094] 도 2의 좌측 상부에는 특정 공간에 대한 컬러-이미지를 나타내었고, 좌측 하부에는 그 특정 공간에 대한 깊이-이미지를 나타내었으며, 우측에는 Depth Proximity Clustering 결과를 나타내었다.

[0095] 3. Color Histogram based Object Detection(도 1의 S140)

[0096] 특징점이 많이 존재하지 않는 균일한 물체를 복잡한 배경에서 물체 인식을 할 경우 잘못된 특징점 매칭이 많이 나타난다. 따라서, 본 실시예에서는 color histogram back-projection을 이용한 물체 탐지 알고리즘을 사용하여, 클러스터 내에서도 인식 물체가 존재할 가능성이 높은 일부 영역을 후보 영역으로 선정하도록 하였다.

[0097] 도 3에는 도 2의 클러스터링 결과에 나타난 클러스터들에 대해 Color Histogram based Object Detection 하여 선정한 3개의 후보 영역을 컬러-이미지에 나타내었다.

[0098] 4. Equality Test of Standard Deviation(도 1의 S150)

[0099] Equality Test of Standard Deviation는 Color Histogram based Object Detection을 통해 선정된 후보 영역들과 인식 물체들에 대해 깊이 표준 편차를 비교하여, 인식 물체가 존재할 가능성이 높은 후보 영역들을 검증 또는 재선정하는 절차로 이해될 수 있다.

[0100] 도 4에는 인식 물체들인 책, 인형, 쿠션에 대한 깊이 표준 편차를 나타내었다. 도 4에 도시된 바와 같이 하나의 물체에 대한 깊이 표준 편차는 물체의 모양과 상관없이 대부분 가우시안 분포(Gaussian distribution)를 따른다.

[0101] 따라서, 인식 물체와 후보 영역의 랜덤 샘플이 정규 분포를 따른다고 가정할 경우 표준편차의 비율은 F 분포를 따른다. 그리고, 신뢰도는 α이고, 자유도는 n_x-1, n_y-1이며, 자유도는 랜덤 샘플의 개수로 설정할 경우 다음과 같은 신뢰 구간을 갖는데, 신뢰 구간을 벗어나는 매칭은 제거되고 나머지가 신뢰할 수 있는 매칭으로 간주된다.

[0102] 신뢰 구간 : $(F_{1-\alpha/2},n_x-1,n_y-1 \quad F_{\alpha/2},n_x-1,n_y-1)$

[0103] 5. Depth Layered Feature Matching with Geometric Constraints(도 1의 S160)

[0104] S130단계에서의 특징점 매칭으로, 스케일과 회전에 불변한 특징점 매칭 기법을 사용함이 바람직한데, 다른 종류의 특징점 매칭 기법을 사용하는 것을 배제하는 것은 아니다.

[0105] 한편, 렌즈에 의해 발생하는 radial distortion을 고려하지 않을 경우, 일반적으로 두 이미지 간의 관계는 하나의 행렬로 표현할 수 있다. 이미지에서 대응점들이 건물 벽과 같은 평면상에 여러 개가 존재하는 경우 하나의 평면상에 존재하는 것으로 근사할 수 있으며 이 관계를 Homography라고 한다.

[0106] Depth Proximity Clustering을 통해 생성된 클러스터와 인식 물체 간의 homography는, 인식 물체와 배경의 특징점 매칭이 배제되기 때문에, 컬러-이미지와 인식 물체 간의 homography 보다 정확하다.

[0107] 이에 더 나아가, Color Histogram based Object Detection과 Equality Test of Standard Deviation(도 1의 S140 및 S150)을 통해 선정한 후보 영역들을 포함하는 클러스터들과 인식 물체 간의 homography는, 인식 물체와 배경의 특징점 매칭이 더욱 더 배제되기 때문에, 클러스터와 인식 물체 간의 homography 보다 정확하다.

[0108] 6. Final decision with consistency check(도 1의 S170 및 S180)

[0109] Final decision with consistency check는, 하나의 인식 물체가 다수의 클러스터에서 인식되는 경우와 컬러-이미지에서 다수의 인식 물체가 인식되는 경우, 최종적인 물체 인식을 위한 절차이다.

[0110] 하나의 인식 물체가 다수의 클러스터에서 인식되는 경우를 도 5와 도 6에 예시하였다. 도 5와 도 6에 도시된 바에 따르면, 인식 물체인 "쿠션"이 각기 다른 클러스터들에서 많은 특징점 매칭수를 보이고 있음을 확인할 수 있다. 전술했던 바와 같이, 양자 간에 특징점 매칭수의 비율이 특정 범위 내이면, 컬러-이미지에서 "쿠션"은 인식되지 않은 것으로 처리한다.

[0111] 컬러-이미지에서 다수의 인식 물체가 인식되는 경우를 도 7과 도 8에 예시하였다. 도 7과 도 8에는, 컬러-이미지에서 특징점 매칭수가 상위 2개인 "책"과 "인형"에 대한 특징점 매칭이 나타나 있다.

[0112] 양자 간에 특징점 매칭수의 비율이 특정 범위 밖이면, 예를 들어 "책"과 클러스터 간 특징점 매칭수가 "인형"과 클러스터 간 특징점 매칭수 보다 현저히 많은 경우, "책"이 인식된 것으로 취급한다.

[0113] 반면, 양자 간에 특징점 매칭수의 비율이 특정 범위 내이면, 클러스터와 "책" 간의 컬러-상관도와 클러스터와 "인형" 간의 컬러-상관도를 비교하여, 컬러-상관도가 큰 인식 물체가 클러스터 내에서 인식된 것으로 처리한다.

[0114] **7. 깊이 정보를 이용한 물체 인식 #2**

[0115] 이하에서는, 본 발명의 다른 실시예에 따른 깊이 정보를 이용한 물체 인식 방법의 설명에 대해 도 9를 참조하여 설명하되, 전술한 설명과 동일한 부분에 대해서는 그 설명을 간략히 하겠다.

[0116] 도 9에 도시된 바와 같이, 먼저 컬러-이미지와 깊이-이미지를 획득한다(S210). 이후, 깊이-이미지에 나타난 깊이 정보를 기초로, 컬러-이미지를 다수의 클러스터들로 클러스터링한다(S220).

[0117] 이후, 클러스터들 내의 일부 영역들을 후보 영역들로 1차 선정한다(S230). S230단계는, S220단계에서 클러스터링된 클러스터들 각각에 대해 수행되며, 후보 영역 선정은 클러스터와 인식 물체의 컬러-히스토그램 비교를 통해 수행된다.

[0118] 다음, 후보 영역들 각각과 인식 물체들 각각에 대해 특징점 매칭을 수행한다(S240). 도 9의 S240단계에서의 특징점 매칭은 "후보 영역"과 인식 물체 간의 특징점 매칭이라는 점에서, 도 1의 S130단계에서의 "클러스터"와 인식 물체 간의 특징점 매칭과 차이가 있다.

[0119] 이후, S230단계에서 1차 선정된 후보 영역과 인식 물체 간의 깊이 표준 편차가 유사한 후보 영역들을 2차로 선정한다(S250).

[0120] 다음, S240단계에서의 특징점 매칭 결과에서 S250단계에서 선정된 "후보 영역들" 밖에서 이루어진 특징점 매칭을 배제하고, 물체 인식을 수행한다(S260). 즉, S240단계에서의 특징점 매칭 결과 중 S250단계에서 선정된 후보 영역들에 포함된 특징점 매칭만을 남기고, 물체 인식을 수행한다.

[0121] S260단계에서 배제되는 특징점 매칭들은 2차 선정된 후보 영역들 밖에서 이루어진 특징점 매칭이라는 점에서, 2차 선정된 후보 영역들을 포함하는 클러스터들 밖에서 이루어진 특징점 매칭을 배제하는 도 1의 S160단계와 차이가 있다.

[0122] 한편, 하나의 인식 물체가 다수의 "후보 영역"에서 인식되는 경우, 그 인식 물체는 인식되지 않은 것으로 처리할 수 있다(S270). 구체적으로, 특징점 매칭수가 첫 번째로 많은 후보 영역과 두 번째로 많은 후보 영역 간에 특징점 매칭수의 비율이 특정 범위 내인 인식 물체는, 인식되지 않은 것으로 처리하는 것이다.

[0123] 또한, 컬러-이미지의 적어도 하나의 후보 영역에서 다수의 인식 물체가 인식되는 경우, 특징점 매칭수가 상위 2개인 후보 영역과 인식 물체들 간의 컬러-상관도를 비교하여 하나의 인식 물체만을 선정하고, 선정된 인식 물체가 인식된 것으로 처리한다(S280).

[0124] 구체적으로, S280단계도, 특징점 매칭수가 첫 번째로 많은 인식 물체와 두 번째로 많은 인식 물체 간에 특징점 매칭수의 비율이 특정 범위 내인 후보 영역에 대해서만 수행한다. 그렇지 않으면, 특징점 매칭수가 첫 번째로 많은 인식 물체가 인식된 것으로 처리한다.

[0125] S270단계와 S280단계에서 처리 기준은 후보 영역이라는 점에서, 처리 기준이 클러스터인 도 1의 S170단계 및 S180단계와 차이가 있다.

[0126] **8. 깊이 정보를 이용한 물체 인식 #3**

[0127] 이하에서는, 본 발명의 또 다른 실시예에 따른 깊이 정보를 이용한 물체 인식 방법의 설명에 대해 도 10을 참조하여 설명한다.

[0128] 도 10에 도시된 물체 인식 방법은, 후보 영역 1차 선정 및 2차 선정을 수행한 후에(S330, S340), 2차 선정된 후보 영역들 각각과 인식 물체들 각각에 대해 특징점 매칭을 수행하여(S350), 물체를 인식한다(S360)는 점에서, 도 9에 도시된 물체 인식 방법과 차이가 있을 뿐, 나머지 사항은 동일하므로 이에 대한 상세한 설명은

생략한다.

[0129] **9. 물체 인식 장치**

[0130] 도 11은 도 1, 도 9 및 도 10에 도시된 물체 인식 방법을 수행할 수 있는 물체 인식 장치를 도시한 도면이다. 도 11에 도시된 바와 같이, 본 발명이 적용가능한 물체 인식 장치는, RGB-카메라(210), D-카메라(420), 프로세서(430), 저장부(440) 및 디스플레이(450)를 구비한다.

[0131] RGB-카메라(410)는 컬러-이미지를 생성하고, D-카메라(420)는 깊이-이미지를 생성한다. RGB-카메라(410)와 D-카메라(420)는 도 11에 도시된 바와 같이 별도로 구성할 수도 있지만, 하나의 카메라로 구성할 수도 있다.

[0132] 저장부(440)에는 인식 물체들의 이미지들이 DB화 되어 있다. 한편, 저장부(440)에는 인식 물체들에 대한 깊이 표준 편차들도 함께 DB화 되어 있다.

[0133] 프로세서(430)는 RGB-카메라(410)에서 생성된 컬러-이미지, D-카메라(420)에서 생성된 깊이-이미지 및 저장부(440)에는 저장된 인식 물체들을 이용하여, 도 1, 도 9 및 도 10에 도시된 물체 인식 방법을 수행한다.

[0134] 디스플레이(450)에는 RGB-카메라(410)에 의해 생성된 컬러-이미지가 표시되는데, D-카메라(420)에 의해 생성된 깊이-이미지를 이용하여 입체-이미지로 표시될 수도 있다. 또한, 디스플레이(450)에는 프로세서(430)에 의해 수행된 물체 인식 결과가 디스플레이된다.

[0135] 또한, 이상에서는 본 발명의 바람직한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

부호의 설명

[0136] 410 : RGB-카메라

420 : D-카메라

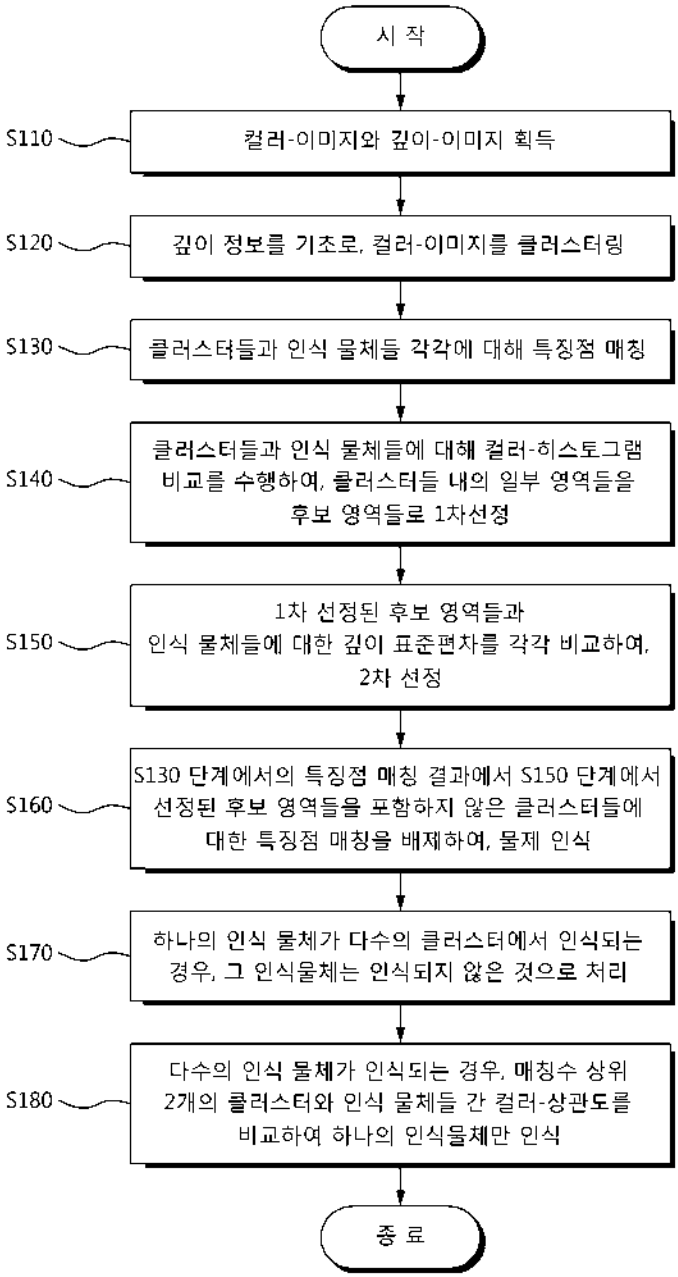
430 : 프로세서

440 : 저장부

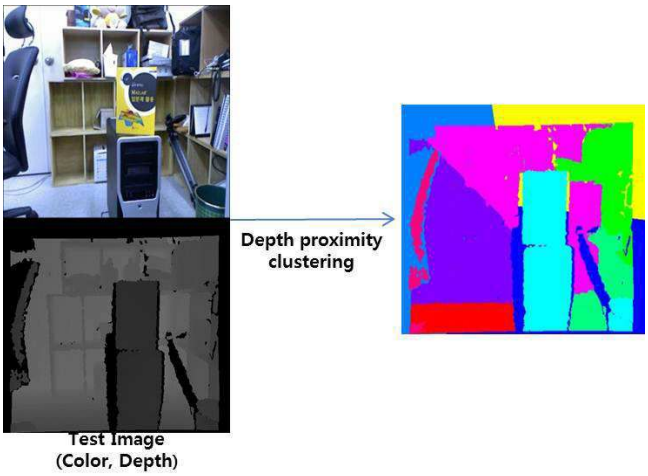
450 : 디스플레이

도면

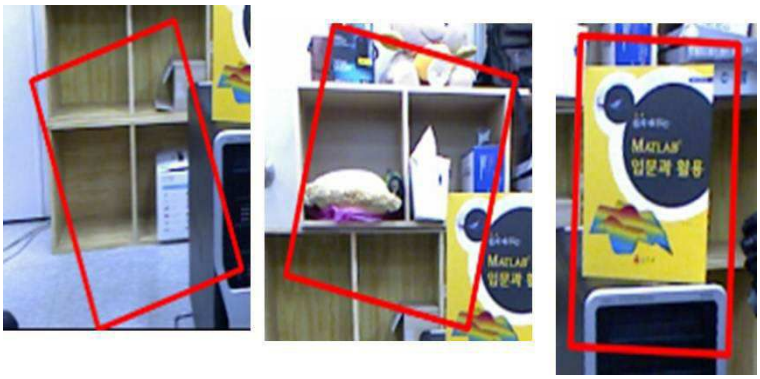
도면1


시 작
S110
컬러-이미지와 깊이-이미지 획득
S120
깊이 정보를 기초로, 컬러-이미지를 클러스터링
S130
클러스터들과 인식 물체들 각각에 대해 특징점 매칭
S140
클러스터들과 인식 물체들에 대해 컬러-히스토그램 비교를 수행하여, 클러스터들 내의 일부 영역들을 후보 영역들로 1차선정
S150
1차 선정된 후보 영역들과 인식 물체들에 대한 깊이 표준편차를 각각 비교하여, 2차 선정
S160
S130 단계에서의 특징점 매칭 결과에서 S150 단계에서 선정된 후보 영역들을 포함하지 않은 클러스터들에 대한 특징점 매칭을 배제하여, 물제 인식
S170
하나의 인식 물체가 다수의 클러스터에서 인식되는 경우, 그 인식물체는 인식되지 않은 것으로 처리
S180
다수의 인식 물체가 인식되는 경우, 매칭수 상위 2개의 클러스터와 인식 물체들 간 컬러-상관도를 비교하여 하나의 인식물체만 인식
종 료

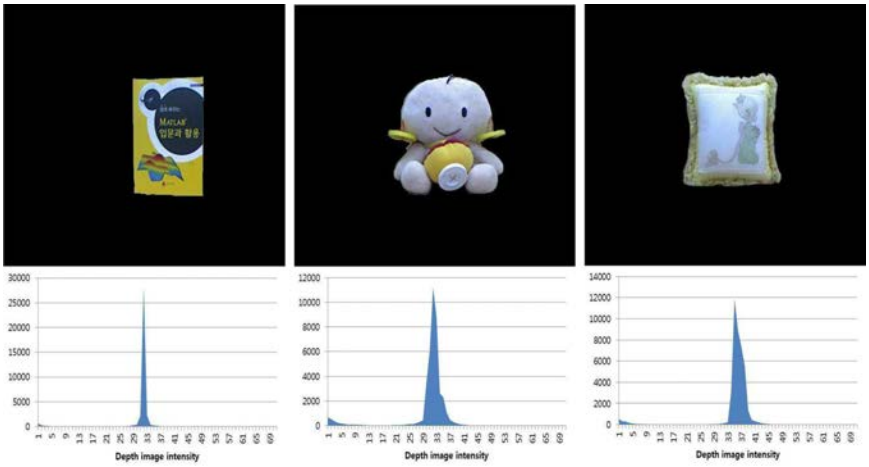
도면2


Depth proximity clustering
Test Image
(Color, Depth)

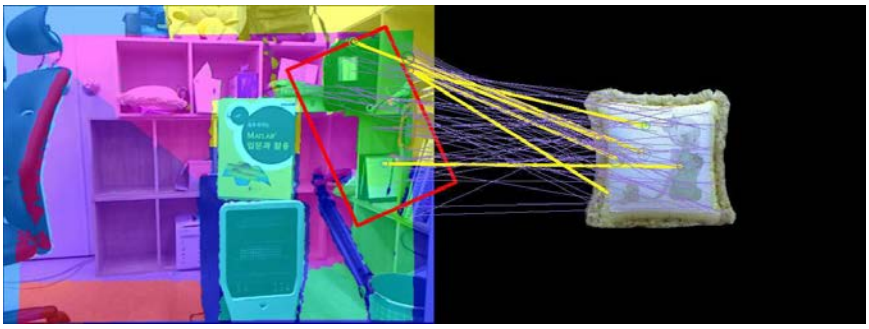
도면3



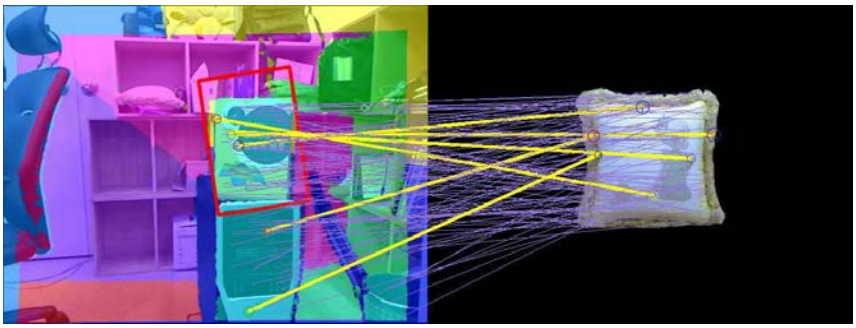
도면4


Depth image intensity
Depth image intensity
Depth image intensity

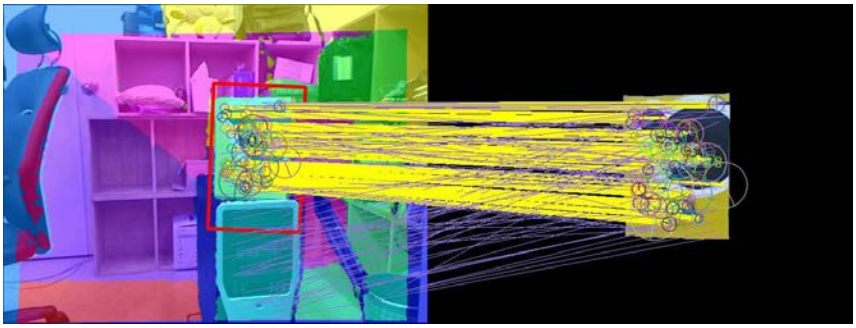
도면5



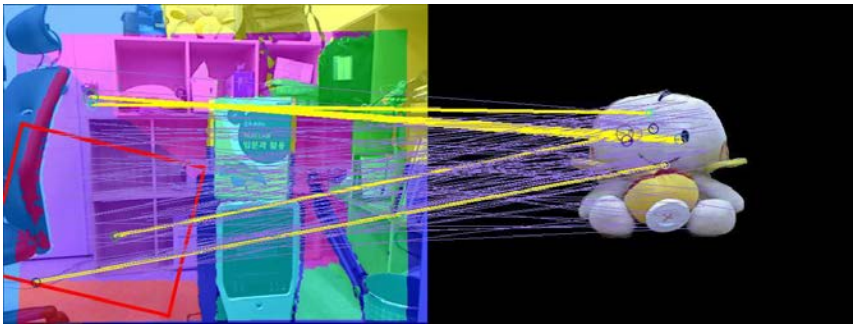
도면6



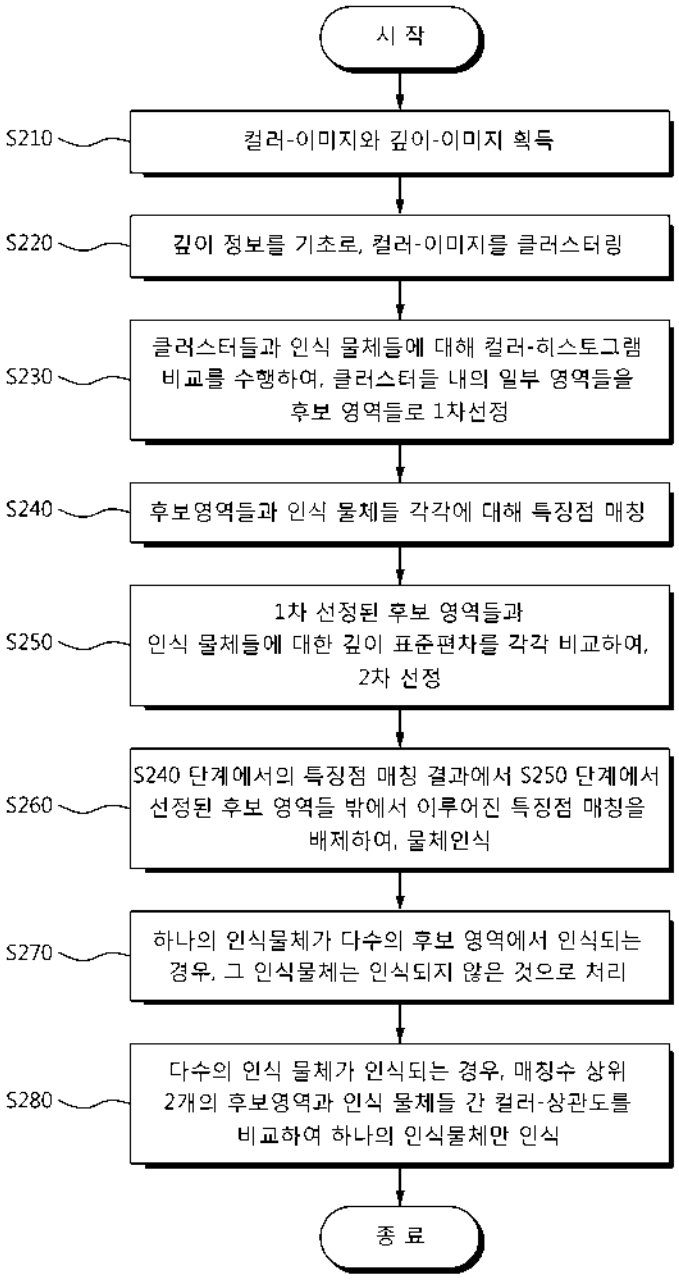
도면7



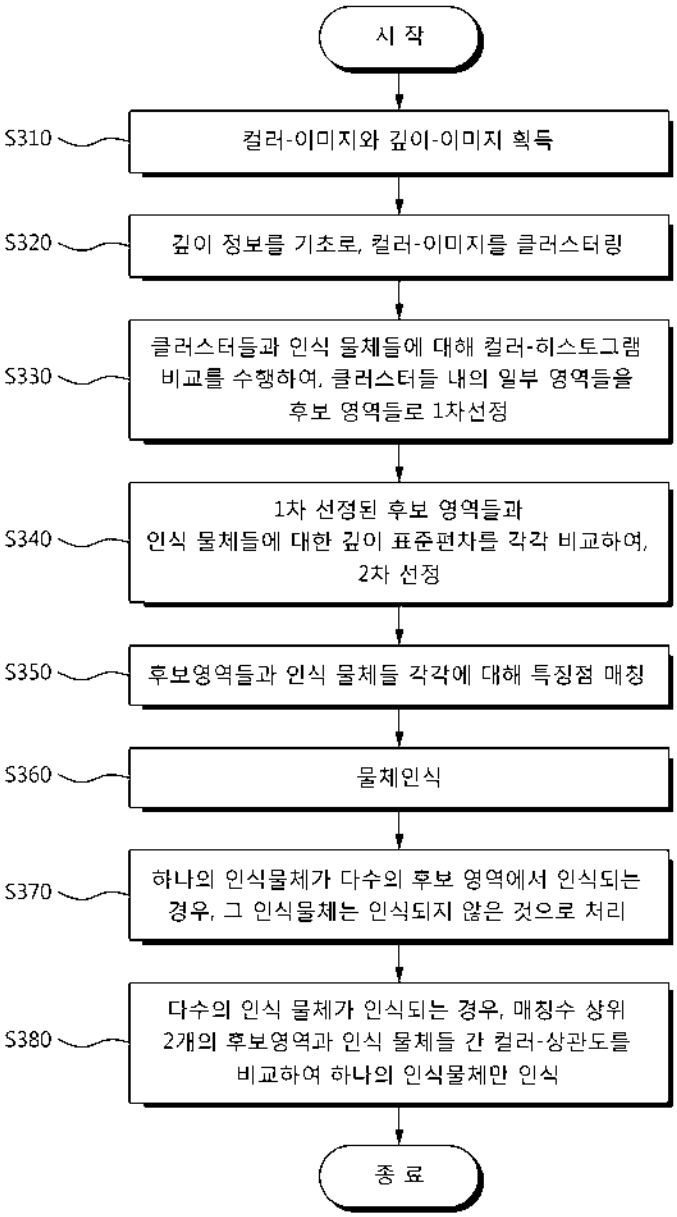
도면8



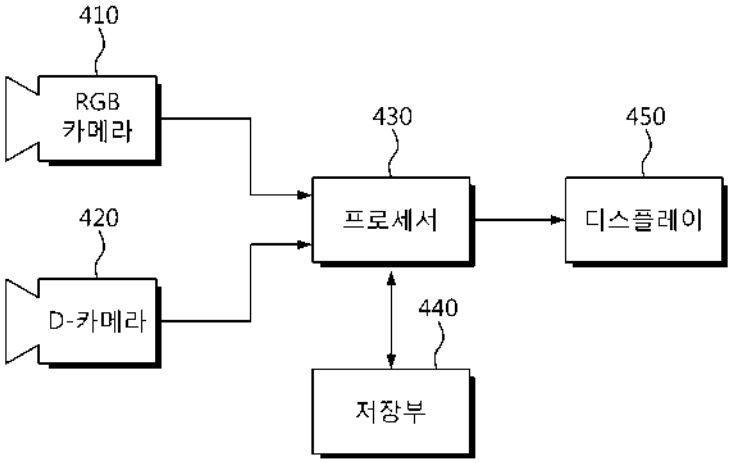
도면9


시 작
S210
컬러-이미지와 깊이-이미지 획득
S220
깊이 정보를 기초로, 컬러-이미지를 클러스터링
S230
클러스터들과 인식 물체들에 대해 컬러-히스토그램 비교를 수행하여, 클러스터들 내의 일부 영역들을 후보 영역들로 1차선정
S240
후보영역들과 인식 물체들 각각에 대해 특징점 매칭
S250
1차 선정된 후보 영역들과 인식 물체들에 대한 깊이 표준편차를 각각 비교하여, 2차 선정
S260
S240 단계에서의 특징점 매칭 결과에서 S250 단계에서 선정된 후보 영역들 밖에서 이루어진 특징점 매칭을 배제하여, 물체인식
S270
하나의 인식물체가 다수의 후보 영역에서 인식되는 경우, 그 인식물체는 인식되지 않은 것으로 처리
S280
다수의 인식 물체가 인식되는 경우, 매칭수 상위 2개의 후보영역과 인식 물체들 간 컬러-상관도를 비교하여 하나의 인식물체만 인식
종 료

도면10


시 작
S310
컬러-이미지와 깊이-이미지 획득
S320
깊이 정보를 기초로, 컬러-이미지를 클러스터링
S330
클러스터들과 인식 물체들에 대해 컬러-히스토그램 비교를 수행하여, 클러스터들 내의 일부 영역들을 후보 영역들로 1차선정
S340
1차 선정된 후보 영역들과 인식 물체들에 대한 깊이 표준편차를 각각 비교하여, 2차 선정
S350
후보영역들과 인식 물체들 각각에 대해 특징점 매칭
S360
물체인식
S370
하나의 인식물체가 다수의 후보 영역에서 인식되는 경우, 그 인식물체는 인식되지 않은 것으로 처리
S380
다수의 인식 물체가 인식되는 경우, 매칭수 상위 2개의 후보영역과 인식 물체들 간 컬러-상관도를 비교하여 하나의 인식물체만 인식
종 료

도면11



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2015년05월06일
(11) 등록번호 10-1517805
(24) 등록일자 2015년04월29일

(51) 국제특허분류(Int. Cl.)
G06T 7/00 (2006.01)
(21) 출원번호 10-2013-0045556
(22) 출원일자 2013년04월24일
심사청구일자 2013년04월24일
(65) 공개번호 10-2014-0127044
(43) 공개일자 2014년11월03일
(56) 선행기술조사문헌
김종배, "안전 운전 지원을 위한 도로 영상에서 시각 주의 영역 검출", 한국전자공학회 논문지 제49권 SC편 제1호, 2012.01. pp.94-102.*
JP2002288658 A
JP2003189235 A
JP2009037625 A
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
김정호
경기 성남시 분당구 매화로48번길 17-1, 205호 (야탑동)
최병호
경기 용인시 수지구 대지로 27, 103동 1306호 (죽전동, 한신아파트)
(뒷면에 계속)
(74) 대리인
남충우, 노철호

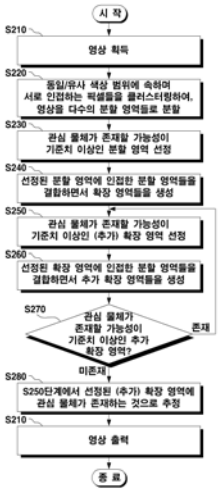
전체 청구항 수 : 총 8 항 심사관 : 진민숙

(54) 발명의 명칭 **영상 특성 기반 분할 기법을 이용한 물체 검출 방법 및 이를 적용한 영상 처리 장치**

(57) 요 약

영상 특성 기반 분할 기법을 이용한 물체 검출 방법 및 이를 적용한 영상 처리 장치가 제공된다. 본 발명의 실시예에 따른 물체 검출 방법은, 영상의 특성을 기초로 영상을 다수의 분할 영역들로 분할하여, 분할 영역들을 기반으로 물체를 검출한다. 이에 의해, 물체 검출 성공률을 높일 수 있고, 윈도우로 영상 전체를 스캔하는 방식이 아니기 때문에 물체 검출이 매우 빠른 속도로 이루어지게 된다.

대 표 도 - 도2



(72) 발명자

황영배

서울 서초구 방배선행길 1, 106동 607호 (방배동, 방배우성아파트)

배주한

서울 광진구 동일로26길 52, (화양동)

이 발명을 지원한 국가연구개발사업

과제고유번호 10040246

부처명 지식경제부

연구관리전문기관 한국산업기술평가관리원

연구사업명 산업원천기술개발사업

연구과제명 이동 로봇의 안정적 영상 획득을 통한 3D Depth 정보 획득과 신시간 객체 인식을 위한 로봇 비전 SoC 및 모듈 개발

기 여 율 1/1

주관기관 전자부품연구원

연구기간 2011.06.01 ~ 2015.05.31

명 세 서

청구범위

청구항 1

영상의 특성을 기초로, 영상을 다수의 분할 영역들로 분할하는 단계; 및

상기 분할단계에서 분할된 분할 영역들을 기반으로, 물체를 검출하는 단계;를 포함하고,

상기 분할 영역들에 기반한 물체 검출단계는,

상기 물체가 존재할 가능성이 기준치 이상인 분할 영역을 선정하는 단계;

상기 선정단계에서 선정된 분할 영역에 인접한 분할 영역들 중 적어도 하나를 결합하면서 적어도 하나의 확장 영역들을 생성하는 단계; 및

상기 확장 영역들을 기반으로, 상기 물체를 검출하는 단계;를 포함하는 것을 특징으로 하는 물체 검출 방법.

청구항 2

제 1항에 있어서,

상기 분할 영역들은,

영상값이 유사 범위 내에 속하며, 서로 인접하는 픽셀들의 집합인 것을 특징으로 하는 물체 검출 방법.

청구항 3

제 2항에 있어서,

상기 영상값은,

색상 및 휘도 중 적어도 하나인 것을 특징으로 하는 물체 검출 방법.

청구항 4

삭제

청구항 5

제 1항에 있어서,

상기 확장 영역들에 기반한 물체 검출단계는,

상기 확장 영역들에 대해서만, 상기 물체를 검출하는 것을 특징으로 하는 물체 검출 방법.

청구항 6

제 1항에 있어서,

상기 확장 영역들에 기반한 물체 검출단계는,

상기 물체가 존재할 가능성이 기준치 이상인 확장 영역을 선정하는 단계;

상기 선정단계에서 선정된 확장 영역에 인접한 분할 영역들 중 적어도 하나를 결합하면서 적어도 하나의 추가

확장 영역들을 생성하는 단계; 및

상기 추가 확장 영역들을 기반으로, 상기 물체를 검출하는 단계;를 포함하는 것을 특징으로 하는 물체 검출 방법.

청구항 7

제 6항에 있어서,

상기 추가 확장 영역들에 기반한 물체 검출단계는,

상기 물체가 존재할 가능성이 기준치 이상인 추가 확장 영역이 없으면, 상기 선정된 확장 영역에 상기 물체가 존재하는 것으로 추정하는 단계;를 포함하는 것을 특징으로 하는 물체 검출 방법.

청구항 8

제 7항에 있어서,

상기 확장 영역들에 기반한 물체 검출단계는,

상기 물체가 존재할 가능성이 기준치 이상인 추가 확장 영역이 있으면, 상기 추가 확장 영역에 인접한 분할 영역들 중 적어도 하나를 결합하면서 적어도 하나의 추가 확장 영역들을 생성하는 단계; 및

상기 생성단계에서 생성된 추가 확장 영역들을 기반으로, 상기 물체를 검출하는 단계;를 더 포함하는 것을 특징으로 하는 물체 검출 방법.

청구항 9

영상을 획득하는 획득부; 및

상기 획득부에서 획득된 영상의 특성을 기초로 영상을 다수의 분할 영역들로 분할하고, 분할된 분할 영역들을 기반으로 물체를 검출하는 프로세서;를 포함하고,

상기 프로세서는,

상기 물체가 존재할 가능성이 기준치 이상인 분할 영역을 선정하고, 선정된 분할 영역에 인접한 분할 영역들 중 적어도 하나를 결합하면서 적어도 하나의 확장 영역들을 생성하며, 상기 확장 영역들을 기반으로 상기 물체를 검출하는 것을 특징으로 하는 영상 처리 장치.

청구항 10

삭제

발명의 설명

기 술 분 야

[0001] 본 발명은 물체 검출 방법에 관한 것으로, 더욱 상세하게는 영상으로부터 차량이나 사람과 같은 관심 물체를 검출하기 위한 물체 검출 방법 및 이를 적용한 영상 처리 장치에 관한 것이다.

배 경 기 술

[0002] 물체 검출을 위해서는, 정방형의 윈도우(Window)를 영상 전체에서 스캔(Scan) 하면서, HoG(Histogram of Oriented Gradients)와 에지 등과 같은 특징량(Feature)를 추출하고, 이를 미리 학습한 SVM(Support Vector

Machine) 등의 분류기(Classifier)를 통하여 관심 물체의 유무를 판단하게 된다.

[0003] 이때 발생하는 문제점은 영상에 존재하는 물체들은 다양한 조건으로부터 외형의 변형이 크다는 것이다. 예를 들면, 관심 물체의 자세 변화, 카메라 시점 변화, 부분적인 가려짐(Occlusion) 등으로 인하여 영상에서 물체의 크기와 모양은 다양하게 변화하는 것이다.

[0004] 이에 대처하기 위해, 다양한 크기 및 모양의 윈도우를 영상 전체에 스캔하고 있는데, 이는 많은 처리 시간을 필요로 한다는 문제가 있다.

[0005] 뿐만 아니라, 사람 또는 차량과 같이 물체의 종류가 다를 경우 다른 창의 크기와 모양이 요구되기 때문에 검출하고자 하는 물체의 종류에 따라서 동일한 작업을 다시 수행해야 하는 문제점도 있다.

발명의 내용

해결하려는 과제

[0006] 본 발명은 상기와 같은 문제점을 해결하기 위하여 안출된 것으로서, 본 발명의 목적은, 영상 특성 기반 분할 기법을 이용하여 고속의 물체 검출을 가능하게 하는 물체 검출 방법 및 이를 적용한 영상 처리 장치를 제공함에 있다.

과제의 해결 수단

[0007] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른, 물체 검출 방법은, 영상의 특성을 기초로, 영상을 다수의 분할 영역들로 분할하는 단계; 및 상기 분할단계에서 분할된 분할 영역들을 기반으로, 물체를 검출하는 단계;를 포함한다.

[0008] 그리고, 상기 분할 영역들은, 영상값이 유사 범위 내에 속하며, 서로 인접하는 픽셀들의 집합일 수 있다.

[0009] 또한, 상기 영상값은, 색상 및 휘도 중 적어도 하나일 수 있다.

[0010] 그리고, 상기 분할 영역들에 기반한 물체 검출단계는, 상기 물체가 존재할 가능성이 기준치 이상인 분할 영역을 선정하는 단계; 상기 선정단계에서 선정된 분할 영역에 인접한 분할 영역들 중 적어도 하나를 결합하면서 적어도 하나의 확장 영역들을 생성하는 단계; 및 상기 확장 영역들을 기반으로, 상기 물체를 검출하는 단계;를 포함할 수 있다.

[0011] 또한, 상기 확장 영역들에 기반한 물체 검출단계는, 상기 확장 영역들에 대해서만, 상기 물체를 검출할 수 있다.

[0012] 그리고, 상기 확장 영역들에 기반한 물체 검출단계는, 상기 물체가 존재할 가능성이 기준치 이상인 확장 영역을 선정하는 단계; 상기 선정단계에서 선정된 확장 영역에 인접한 분할 영역들 중 적어도 하나를 결합하면서 적어도 하나의 추가 확장 영역들을 생성하는 단계; 및 상기 추가 확장 영역들을 기반으로, 상기 물체를 검출하는 단계;를 포함할 수 있다.

[0013] 또한, 상기 추가 확장 영역들에 기반한 물체 검출단계는, 상기 물체가 존재할 가능성이 기준치 이상인 추가 확장 영역이 없으면, 상기 확장 영역에 상기 물체가 존재하는 것으로 추정하는 단계;를 포함할 수 있다.

[0014] 그리고, 상기 확장 영역들에 기반한 물체 검출단계는, 상기 물체가 존재할 가능성이 기준치 이상인 추가 확장 영역이 있으면, 상기 추가 확장 영역에 인접한 분할 영역들 중 적어도 하나를 결합하면서 적어도 하나의 추가 확장 영역들을 생성하는 단계; 및 상기 생성단계에서 생성된 추가 확장 영역들을 기반으로, 상기 물체를 검출하는 단계;를 더 포함할 수 있다.

[0015] 한편, 본 발명의 다른 실시예에 따른, 영상 처리 장치는, 영상을 획득하는 획득부; 및 상기 획득부에서 획득된 영상의 특성을 기초로 영상을 다수의 분할 영역들로 분할하고, 분할된 분할 영역들을 기반으로 물체를 검출하는 프로세서;를 포함한다.

[0016] 그리고, 상기 분할 영역들은, 영상값이 유사 범위 내에 속하며, 서로 인접하는 픽셀들의 집합일 수 있다.

발명의 효과

[0017] 이상 설명한 바와 같이, 본 발명의 실시예들에 따르면, 영상 특성 기반 분할 기법을 기초로 물체 검출 성공률을 높일 수 있게 된다. 뿐만 아니라, 다양한 윈도우로 영상 전체를 스캔하는 방식이 아니기 때문에, 물체 검출이 매우 빠른 속도로 이루어지게 된다.

도면의 간단한 설명

[0018] 도 1은 본 발명이 적용가능한 영상 처리 장치의 블럭도,

도 2는 본 발명의 바람직한 실시예에 따른 물체 검출 방법의 설명에 제공되는 흐름도,

도 3은, 도 2에 도시된 물체 검출 방법의 부연 설명에 제공되는 도면,

도 4a 내지 도 4c는, 도 2에 도시된 물체 검출 방법의 부연 설명에 제공되는 도면,

도 5는 본 실시예에 따른 영상 특성 기반 분할 기법을 이용한 물체 검출 결과를 예시한 도면, 그리고,

도 6은 본 실시예에 따른 방법과 기존의 윈도우 스캔방식의 검출 성능을 비교한 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0019] 이하에서는 도면을 참조하여 본 발명을 보다 상세하게 설명한다.

[0020] 도 1은 본 발명이 적용가능한 영상 처리 장치의 블럭도이다. 본 발명이 적용가능한 영상 처리 장치(100)는, 도 1에 도시된 바와 같이, 영상 획득부(110), 영상 프로세서(120) 및 영상 출력부(130)를 포함한다.

[0021] 영상 획득부(110)는 영상을 획득하고, 획득된 영상을 영상 프로세서(120)에 제공한다. 여기서, 영상 획득부(110)에 의한 영상 획득 방법은 그 종류를 불문한다. 즉, 영상 획득부(110)는 영상 촬영, 저장매체, 네트워크 통신 등으로부터 영상을 획득할 수 있다.

[0022] 영상 프로세서(120)는 영상 획득부(110)로부터 제공되는 영상에서 관심 물체를 검출한다. 관심 물체를 검출하기 위해, 영상 프로세서(120)는 영상의 색상에 기반하여 영상을 분할하고, 분할된 영역에서 추출한 특징량을 분류기에 대입하여 물체 검출을 수행한다.

[0023] 영상 출력부(130)는 영상 획득부(110)에서 획득된 영상에 영상 프로세서(120)의 물체 검출 결과를 나타내어 표시한다.

[0024] 영상 프로세서(120)에 의한 물체 검출 과정에 대해, 도 2를 참조하여 상세히 설명한다. 도 2는 본 발명의 바람직한 실시예에 따른 물체 검출 방법의 설명에 제공되는 흐름도이다.

[0025] 도 2에 도시된 바와 같이, 영상 획득부(110)로부터 영상이 획득되면(S210), 영상 프로세서(120)는 동일/유사 색상 범위에 속하며 서로 인접하는 픽셀들을 클러스터링하여, 영상을 다수의 분할 영역들로 분할한다(S220). 도 3의 중앙에 나타난 이미지는, 아래에 나타난 이미지를 도 2의 S220단계에 따라 다수의 분할 영역들로 분할한 결과를 예시한 것이다.

[0026] 이후, 영상 프로세서(120)는 관심 물체가 존재할 가능성이 기준치 이상인 분할 영역을 선정한다(S230). S230단계에서의 선정은, 분할 영역에서 특징량을 추출하고, 추출된 특징량을 분류기에 대입하여 산출되는 가능성을 참고하여 수행될 수 있다.

[0027] 부연 설명을 위해, 도 4a에는 S220단계를 통해 영상을 17개의 분할 영역들로 분할한 결과를 나타내었고, 도 4b에는 S230단계를 통해 분할 영역-6이 선정된 결과를 나타내었다.

[0028] 도 4b에서는 분할 영역이 하나만 선정된 것을 상정하였으나, 이는 설명의 편의를 위한 것으로, 분할 영역은 다수가 선정될 수 있으며, 실제로 다수의 분할 영역들이 선정됨이 일반적일 것이다.

[0029] 다시, 도 2를 참조하여, S230단계 이후부터 설명한다.

[0030] 영상 프로세서(120)는 S230단계에서 선정된 분할 영역에 인접한 분할 영역들 중 적어도 하나를 결합하면서 적어도 하나의 확장 영역들을 생성한다(S240). 도 4b에 도시된 바와 같이 분할 영역-6이 선정되었다면, 분할 영역-6에 인접한 분할영역들은 분할 영역-1,2,3,5,7,8,9이다.

[0031] 따라서, S240단계에서는,

[0032] 1) 분할 영역-6에 분할 영역-1,2,3,5,7,8,9 중 1개를 결합한 확장 영역들,

[0033] 2) 분할 영역-6에 분할 영역-1,2,3,5,7,8,9 중 2개를 결합한 확장 영역들,

[0034] 3) 분할 영역-6에 분할 영역-1,2,3,5,7,8,9 중 3개를 결합한 확장 영역들,

[0035] 4) 분할 영역-6에 분할 영역-1,2,3,5,7,8,9 중 4개를 결합한 확장 영역들,

[0036] 5) 분할 영역-6에 분할 영역-1,2,3,5,7,8,9 중 5개를 결합한 확장 영역들,

[0037] 6) 분할 영역-6에 분할 영역-1,2,3,5,7,8,9 중 6개를 결합한 확장 영역들,

[0038] 7) 분할 영역-6에 분할 영역-1,2,3,5,7,8,9 모두를 결합한 확장 영역을 생성하게 된다.

[0039] 다음, 영상 프로세서(120)는 S240단계에서 생성된 확장 영역들 중 관심 물체가 존재할 가능성이 기준치 이상인 확장 영역을 선정한다(S250). S250단계에서의 선정은, 확장 영역에서 특징량을 추출하고, 추출된 특징량을 분류기에 대입하여 산출되는 가능성을 참고하여 수행될 수 있다.

[0040] 부연 설명을 위해, 도 4c에는 S250단계를 통해 분할 영역-5,6,8,9가 결합된 확장 영역이 선정된 결과를 나타내었다. 도 4c에서는 확장 영역이 하나만 선정된 것을 상정하였으나, 이는 설명의 편의를 위한 것으로, 확장 영역은 다수가 선정될 수 있으며, 실제로 다수의 확장 영역들이 선정됨이 일반적일 것이다.

[0041] 다시, 도 2를 참조하여, S250단계 이후부터 설명한다.

[0042] 영상 프로세서(120)는 S250단계에서 선정된 확장 영역에 인접한 분할 영역들 중 적어도 하나를 결합하면서 적어도 하나의 추가 확장 영역들을 생성한다(S260). 도 4c에 도시된 바와 같이 분할 영역-5,6,8,9가 결합된 확장 영역이 선정되었다면, 확장 영역에 인접한 분할영역들은 분할 영역-1,2,3,4,7,10,11,12,13이다.

[0043] 따라서, S260단계에서는,

[0044] 1) 확장 영역에 분할 영역-1,2,3,4,7,10,11,12,13 중 1개를 결합한 추가 확장 영역들,

[0045] 2) 확장 영역에 분할 영역-1,2,3,4,7,10,11,12,13 중 2개를 결합한 추가 확장 영역들,

[0046] ...

[0047] 9) 확장 영역에 분할 영역-1,2,3,4,7,10,11,12,13 모두를 결합한 추가 확장 영역을 생성하게 된다.

[0048] 다음, 영상 프로세서(120)는 S260단계에서 생성된 추가 확장 영역들 중 관심 물체가 존재할 가능성이 기준치 이상인 추가 확장 영역을 선정한다(S270). S270단계에서의 선정은, 추가 확장 영역에서 특징량을 추출하고, 추출된 특징량을 분류기에 대입하여 산출되는 가능성을 참고하여 수행될 수 있다.

[0049] 만약, S270단계에서 관심 물체가 존재할 가능성이 기준치 이상인 추가 확장 영역이 없는 경우(S270-미존재). S250단계에서 선정된 확장 영역에 관심 물체가 존재하는 것으로 추정한다(S280).

[0050] 그리고, 영상 출력부(130)는 S210단계에서 획득된 영상에 S280단계에서의 물체 검출 결과를 나타내어 표시한다(S290). S290단계에서의 영상 출력 결과는 도 3의 위에 나타난 이미지와 같이 나타난다.

[0051] 한편, S270단계에서 관심 물체가 존재할 가능성이 기준치 이상인 추가 확장 영역이 존재하는 경우에는(S270-존재), S250단계부터 반복 수행한다.

[0052] 지금까지, 영상 특성 기반 분할 기법을 이용한 물체 검출 방법 및 이를 적용한 영상 처리 장치에 대해 바람직한 실시예들을 들어 상세히 설명하였다.

[0053] 위 실시예에서 영상 분할은 영상을 구성하는 픽셀들의 색상에 기반하는 것을 상정하였으나, 이는 바람직한 일 예에 해당한다. 색상 이외에 휘도나 기타 다른 영상값으로 대체될 수 있음은 물론 2 이상의 영상값을 조합하는 경우도 가능하다.

[0054] 도 5에는 본 실시예에 따른 영상 특성 기반 분할 기법을 이용한 물체 검출 결과를 예시한 도면이고, 도 6은 본

실시예에 따른 방법(HGS+SVM, HGS+GPC)과 기존의 윈도우 스캔방식(SW+SVM, SW+GPC)의 검출 성능을 비교한 그래프이다.

[0055] 도 6을 통해, 본 실시예에 따른 물체 검출 방법은 기존의 방법 보다 우수한 성능을 나타냈음을 확인할 수 있다.

[0056] 또한, 이상에서는 본 발명의 바람직한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

부호의 설명

[0057] 100 : 영상 처리 장치

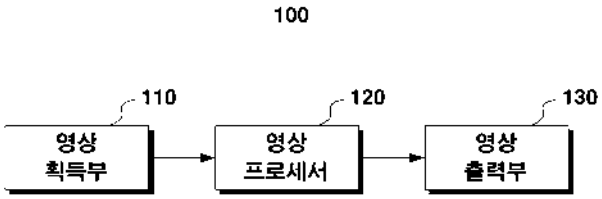
110 : 영상 획득부

120 : 영상 프로세서

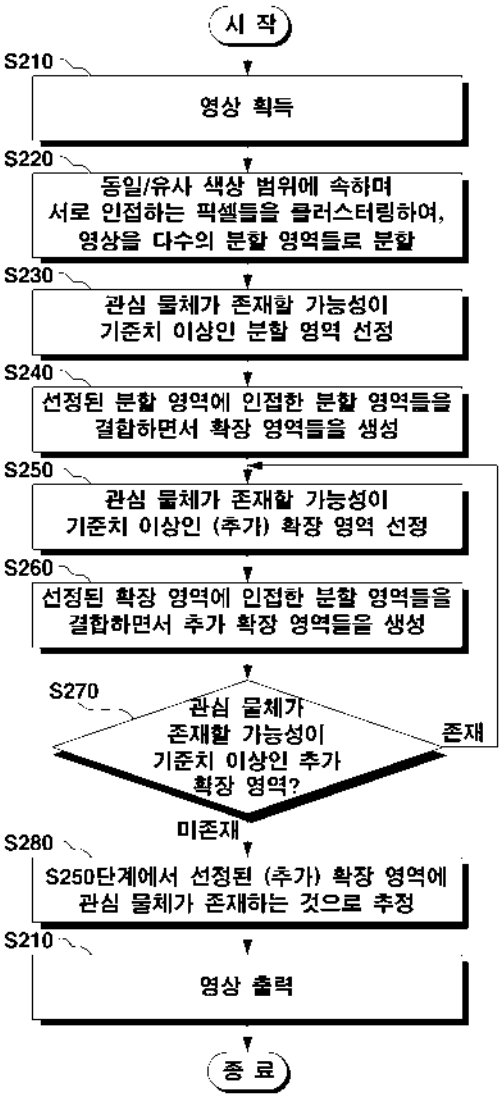
130 : 영상 출력부

도면

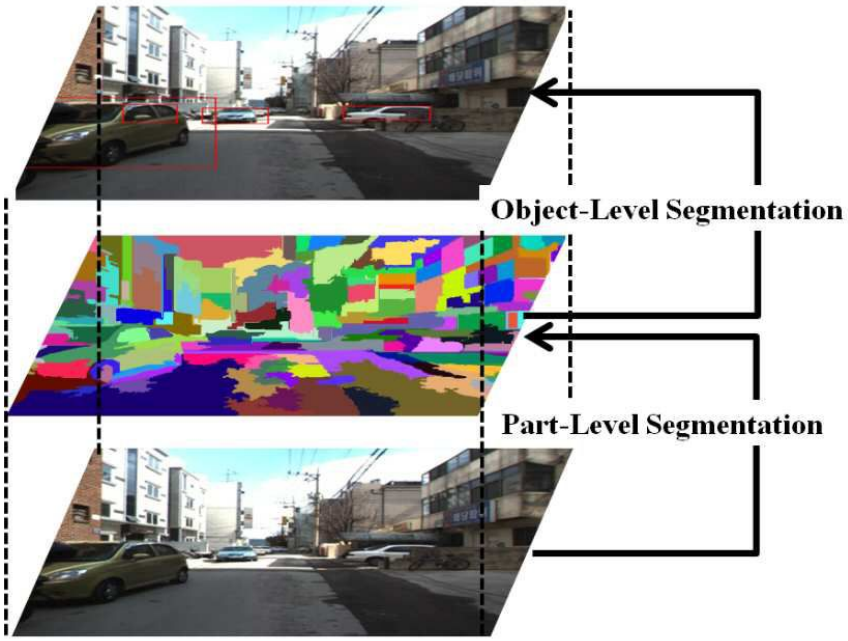
도면1



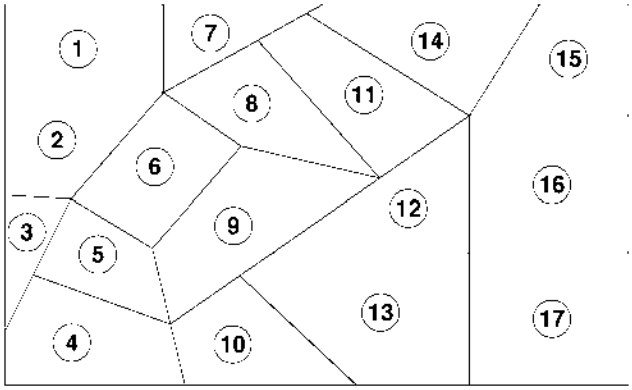
도면2


시 작
S210
영상 획득
S220
동일/유사 색상 범위에 속하며
서로 인접하는 픽셀들을 클러스터링하여,
영상을 다수의 분할 영역들로 분할
S230
관심 물체가 존재할 가능성이
기준치 이상인 분할 영역 선정
S240
선정된 분할 영역에 인접한 분할 영역들을
결합하면서 확장 영역들을 생성
S250
관심 물체가 존재할 가능성이
기준치 이상인 (추가) 확장 영역 선정
S260
선정된 확장 영역에 인접한 분할 영역들을
결합하면서 추가 확장 영역들을 생성
S270
관심 물체가
존재할 가능성이
기준치 이상인 추가
확장 영역?
존재
미존재
S280
S250단계에서 선정된 (추가) 확장 영역에
관심 물체가 존재하는 것으로 추정
S210
영상 출력
종 료

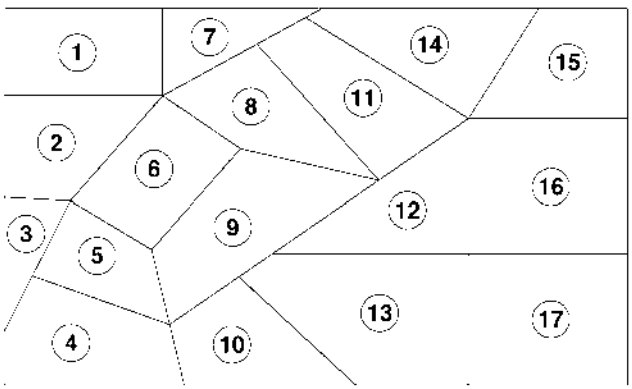
도면3



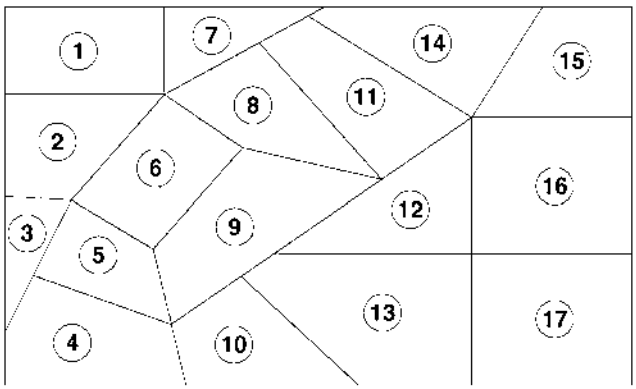
도면4a



도면4b

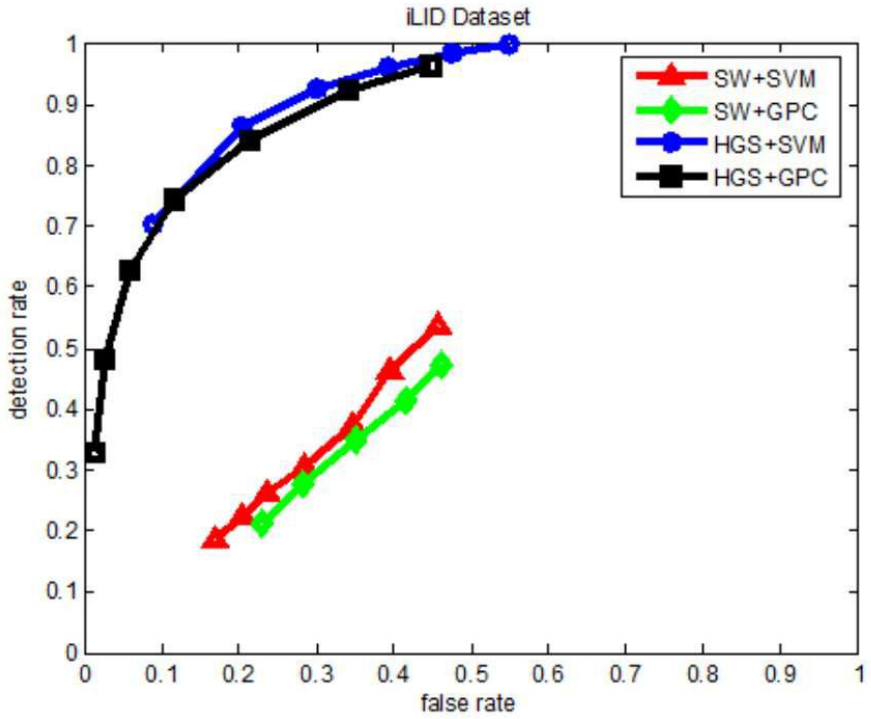


도면4c



도면5

도면6


iLID Dataset
SW+SVM
SW+GPC
HGS+SVM
HGS+GPC
detection rate
false rate
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2016-0074819
(43) 공개일자 2016년06월29일

(51) 국제특허분류(Int. Cl.)
G06T 7/20 (2006.01)
(21) 출원번호 10-2014-0183245
(22) 출원일자 2014년12월18일
심사청구일자 없음

(71) 출원인
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
김정호
경기 성남시 분당구 매화로48번길 17-1, 205호 (야탑동)
최병호
경기 용인시 수지구 대지로 27, 103동 1306호 (죽전동, 한신아파트)
(뒷면에 계속)
(74) 대리인
남충우

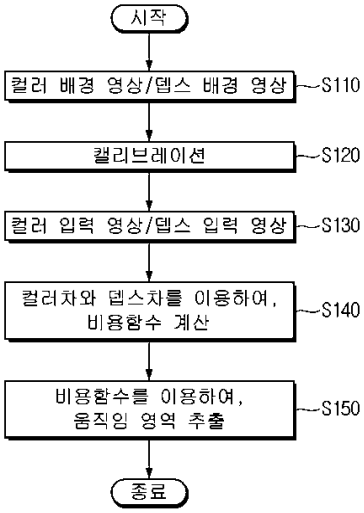
전체 청구항 수 : 총 6 항

(54) 발명의 명칭 **컬러 영상과 뎁스 영상를 이용한 움직임 영역 추출 방법 및 시스템**

(57) 요 약

컬러 영상과 뎁스 영상를 이용한 움직임 영역 추출 방법 및 시스템이 제공된다. 본 발명의 실시예에 따른, 움직임 영역 추출 방법은, '컬러 입력 영상과 컬러 배경 영상의 컬러차'와 '뎁스 입력 영상과 뎁스 배경 영상의 뎁스차'를 기초로, 움직임 영역을 추출하한다. 이에 의해, 조명 변화 및 센서 잡음에 강인한 움직임 영역 추출이 가능하여, 움직임 추정 결과가 기존 방법 대비 우수하다.

대 표 도 - 도2



(72) 발명자

윤주홍

경기도 화성시 동탄반석로 277, 111동 1103호

배주한

서울 광진구 동일로26길 52 (화양동)

이 발명을 지원한 국가연구개발사업

과제고유번호 GK14D0200

부처명 미래창조과학부/교육부

연구관리전문기관 재단법인 기가코리아사업단

연구사업명 (미래부)Giga KOREA사업

연구과제명 실시간 인터렉션을 제공하는 초다시점 단말기술개발

기 여 율 1/1

주관기관 한국과학기술연구원

연구기간 2013.09.01 ~ 2018.04.30

명 세 서

청구범위

청구항 1

컬러 영상을 획득하는 단계;

뎁스 영상을 획득하는 단계; 및

'컬러 입력 영상과 컬러 배경 영상의 컬러차'와 '뎁스 입력 영상과 뎁스 배경 영상의 뎁스차'를 기초로, 움직임 영역을 추출하는 단계;를 포함하는 것을 특징으로 하는 움직임 영역 추출 방법.

청구항 2

청구항 1에 있어서,

상기 추출단계는,

상기 컬러차에 제1 가중치를 부여하고, 상기 뎁스차에 제2 가중치를 부여하는 것을 특징으로 하는 움직임 영역 추출 방법.

청구항 3

청구항 1에 있어서,

상기 추출단계는,

주변 픽셀과의 유사도를 더 고려하여 상기 움직임 영역을 추출하는 것을 특징으로 하는 움직임 영역 추출 방법.

청구항 4

청구항 3에 있어서,

상기 추출단계는,

상기 컬러차, 상기 뎁스차 및 상기 유사도에 기초한 비용 함수를 최소화하는 픽셀을 움직임 영역으로 추출하는 것을 특징으로 하는 움직임 영역 추출 방법.

청구항 5

청구항 4에 있어서,

상기 비용 함수는 아래의 수학식들로 정의되는 것을 특징으로 하는 움직임 영역 추출 방법.

$$E(L)=\sum_{p} D_p(L_p)+\sum_{(p,q)\in N} V_{p,q}(L_p,L_q)$$

$$D_p(L_p)=\begin{cases}\dfrac{1-\exp(-\lambda p_t)}{1+\exp(-\lambda p_t)} & L_p=0\\ \dfrac{2\exp(-\lambda p_t)}{1+\exp(-\lambda p_t)} & L_p=1\end{cases}$$

$$p_t = \alpha\left(\|r_t - r_d\| + \|g_t - g_d\| + \|b_t - b_d\|\right) + \beta\left(\|d_t - d_d\|\right)$$

$$V_{p,q}\left(L_p, L_q\right) = \begin{cases} c & L_p = L_q \\ 0 & L_p \neq L_q \end{cases}$$

청구항 6

컬러 영상을 생성하는 컬러-카메라;

뎁스 영상을 생성하는 뎁스-카메라; 및

'컬러 입력 영상과 컬러 배경 영상의 컬러차'와 '뎁스 입력 영상과 뎁스 배경 영상의 뎁스차'를 기초로, 움직임 영역을 추출하는 영상 프로세서;를 포함하는 것을 특징으로 하는 영상 시스템.

발명의 설명

기 술 분 야

[0001] 본 발명은 영상 처리에 관한 것으로, 더욱 상세하게는 영상에서 움직임 영역을 추출하는 방법 및 시스템에 관한 것이다.

배 경 기 술

[0002] 영상에서 움직이는 물체를 검출하기 위해, 도 1에 도시된 바와 같이, 배경 영상을 미리 확보하고, 입력 영상이 획득되었을 때 두 영상의 컬러 값 차이를 계산하여 움직이는 물체에 의한 변화의 유무를 판단하는 기법이 있다.

[0003] 하지만, 컬러 영상의 경우 조명 변화 또는 영상의 잡음으로 인하여 움직이는 물체를 판단하는데 어려움이 있다.

[0004] 한편, 뎁스 카메라는 조명 변화에는 비교적 강인하지만 뎁스 측정 결과에 대한 오차가 크기 때문에 단독으로 움직이는 영역을 검출하기 위한 방안으로 사용이 어렵다는 문제가 있다.

[0005] 이에, 조명 변화와 영상 잡음은 물론 뎁스 측정 오차에 대해서도 강인한 움직임 영역 추출 방법의 모색이 요청된다.

발명의 내용

해결하려는 과제

[0006] 본 발명은 상기와 같은 문제점을 해결하기 위하여 안출된 것으로서, 본 발명의 목적은, 컬러 영상과 뎁스 영상을 융합하여 조명 변화 및 센서 잡음에 강인한 움직임 영역 추출 방법 및 시스템을 제공함에 있다.

과제의 해결 수단

[0007] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른, 움직임 영역 추출 방법은, 컬러 영상을 획득하는 단계; 뎁스 영상을 획득하는 단계; 및 '컬러 입력 영상과 컬러 배경 영상의 컬러차'와 '뎁스 입력 영상과 뎁스 배경 영상의 뎁스차'를 기초로, 움직임 영역을 추출하는 단계;를 포함한다.

[0008] 그리고, 상기 추출단계는, 상기 컬러차에 제1 가중치를 부여하고, 상기 뎁스차에 제2 가중치를 부여할 수 있다.

[0009] 또한, 상기 추출단계는, 주변 픽셀과의 유사도를 더 고려하여 상기 움직임 영역을 추출할 수 있다.

[0010] 그리고, 상기 추출단계는, 상기 컬러차, 상기 뎁스차 및 상기 유사도에 기초한 비용 함수를 최소화하는 픽셀을 움직임 영역으로 추출할 수 있다.

[0011] 또한, 상기 비용 함수는 아래의 수학식들로 정의될 수 있다.

[0012] $$E(L)=\sum_{p} D_p(L_p)+\sum_{(p,q)\in N} V_{p,q}(L_p,L_q)$$

[0013] $$D_p(L_p)=\begin{cases}\dfrac{1-\exp(-\lambda p_t)}{1+\exp(-\lambda p_t)} & L_p=0\\ \dfrac{2\exp(-\lambda p_t)}{1+\exp(-\lambda p_t)} & L_p=1\end{cases}$$

[0014] $$p_t=\alpha\left(\|r_t-r_d\|+\|g_t-g_d\|+\|b_t-b_d\|\right)+\beta\left(\|d_t-d_d\|\right)$$

[0015] $$V_{p,q}(L_p,L_q)=\begin{cases}c & L_p=L_q\\ 0 & L_p\neq L_q\end{cases}$$

[0016] 한편, 본 발명의 다른 실시예에 따른, 영상 시스템은, 컬러 영상을 생성하는 컬러-카메라; 뎁스 영상을 생성하는 뎁스-카메라; 및 '컬러 입력 영상과 컬러 배경 영상의 컬러차'와 '뎁스 입력 영상과 뎁스 배경 영상의 뎁스차'를 기초로, 움직임 영역을 추출하는 영상 프로세서;를 포함한다.

발명의 효과

[0017] 이상 설명한 바와 같이, 본 발명의 실시예들에 따르면, 컬러 영상과 뎁스 영상을 융합하여 조명 변화 및 센서 잡음에 강인한 움직임 영역 추출이 가능하여, 움직임 추정 결과가 기존 방법 대비 우수하다. 이에, 영상에서 사람이나 차량 등의 움직이는 물체를 정확하게 추출하는 것이 가능해진다.

도면의 간단한 설명

[0018] 도 1은 종래의 움직임 영역 추출방법의 설명에 제공되는 도면,

도 2는 본 발명의 일 실시예에 따른 움직임 영역 추출 방법의 설명에 제공되는 흐름도,

도 3은 컬러 배경 영상과 뎁스 배경 영상을 예시한 사진,

도 4는 컬러 입력 영상과 뎁스 입력 영상을 예시한 사진,

도 5는 컬러 영상만을 이용한 움직임 영역 추출 결과,

도 6 및 도 7은 컬러 영상과 뎁스 영상을 모두 이용한 움직임 영역 추출 결과, 그리고,

도 8은 본 발명의 다른 실시예에 따른 영상 시스템의 블럭도이다.

발명을 실시하기 위한 구체적인 내용

[0019] 이하에서는 도면을 참조하여 본 발명을 보다 상세하게 설명한다.

[0020] 도 2는 본 발명의 일 실시예에 따른 움직임 영역 추출 방법의 설명에 제공되는 흐름도이다. 본 발명의 실시예에 따른 움직임 영역 추출 방법은, 배경 영상과 입력 영상을 비교하여 움직임 영역을 추출하는데, 이때 영상으로 컬러 영상과 뎁스 영상을 모두 이용한다.

[0021] 이를 위해, 도 2에 도시된 바와 같이, 먼저, 움직임 물체가 없는 배경에 대한 컬러 영상과 뎁스 영상을 획득한다(S110). 표기의 편의를 위해, 배경에 대한 컬러 영상은 컬러 배경 영상으로, 배경에 대한 뎁스 영상은 뎁스 배경 영상으로, 각각 표기한다.

[0022] 도 3의 좌측에는 컬러 배경 영상을, 도 3의 우측에는 뎁스 배경 영상을, 각각 예시하였다.

[0023] 다음, 컬러 배경 영상과 뎁스 배경 영상을 이용하여, 컬러 영상을 생성하는 영상-카메라와 뎁스 영상을 생성하는 뎁스-카메라 간의 기하학적 위치 관계를 계산하기 위한 캘리브레이션(Calibration)을 수행한다(S120).

[0024] 이후, 영상-카메라에서 생성된 컬러 영상과 뎁스-카메라에서 생성된 뎁스 영상을 입력받는다(S130). 표기의 편의를 위해, 영상-카메라로부터 생성/입력된 컬러 영상은 컬러 입력 영상으로, 뎁스-카메라로부터 생성/입력된 뎁스 영상은 뎁스 입력 영상으로, 각각 표기한다.

[0025] 도 4의 좌측에는 움직임 물체가 있는 컬러 입력 영상을, 도 4의 우측에는 뎁스 입력 영상을, 각각 예시하였다.

[0026] 그리고, 컬러차와 뎁스차를 이용하여 비용 함수를 계산한다(S140). 여기서, 컬러차는 컬러 입력 영상과 컬러 배경 영상의 차를 말하고, 뎁스차는 뎁스 입력 영상과 뎁스 배경 영상의 차를 말한다.

[0027] 비용 함수(에너지 함수) E(L)은 아래의 수학식 1과 같다.

수학식 1

[0028] $$E(L)=\sum_{p}D_p(L_p)+\sum_{(p,q)\in N}V_{p,q}(L_p,L_q)$$

[0029] 위 수학식 1에서 Dp(Lp)는, 컬러차와 뎁스차를 나타내는 척도로 아래의 수학식 2와 같다.

수학식 2

[0030] $$D_p(L_p)=\begin{cases}\dfrac{1-\exp(-\lambda p_t)}{1+\exp(-\lambda p_t)} & L_p=0\\ \dfrac{2\exp(-\lambda p_t)}{1+\exp(-\lambda p_t)} & L_p=1\end{cases}$$

[0031] 위 수학식 2에서 p_t는 컬러차와 뎁스차의 가중치 합으로, 아래의 수학식 3과 같다. 여기서, 컬러차에 대한 가중치 α와 뎁스차에 대한 가중치 β는 0.5로 동일하게 설정할 수 있지만, 그와 다르게 설정하는 것도 가능함은 물론이다.

수학식 3

[0032] $$p_t=\alpha\left(\|r_t-r_d\|+\|g_t-g_d\|+\|b_t-b_d\|\right)+\beta\left(\|d_t-d_d\|\right)$$

[0033] 위 수학식 2와 수학식 3에 나타난 바와 같이, p_t가 크면 L=0 이기 위한 Dp 값이 커지게 되고, L=1이기 위한 Dp 값은 작아지므로 비용 함수 E(L)을 최소화하기 위해서 L값이 1로 결정된다. 여기서, L=1은 움직임 영역, L=0은 움직이지 않는 배경임을 의미한다.

[0034] 한편, 위 수학식 1에서 Vp,q(Lp,Lq)는, 픽셀 p와 주변 픽셀 q의 유사도를 나타내는 척도로 아래의 수학식 4와 같다.

수학식 4

[0035] $$V_{p,q}(L_p,L_q)=\begin{cases}c & L_p=L_q\\ 0 & L_p\neq L_q\end{cases}$$

[0036] 위 수학식 4에 따르면, 두 개의 이웃하는 픽셀의 레이블(L) 값이 다른 값인 경우 Vp,q는 c가 되고, 두 개의 이웃하는 픽셀의 레이블(L) 값이 같은 경우 0을 가지게 되기 때문에, 이웃하는 두 픽셀은 같은 레이블 값을 가지도록 한다.

[0037] 이와 같이, 비용 함수에 주변 화소들에 대한 정보를 반영한 것은, 움직임 추정 결과의 잡음을 제거하기 위함이다.

[0038] 이후, S140단계에서 계산된 비용함수를 최소화하는 픽셀들을 움직임 영역으로 추출한다(S150).

[0039] 도 5에는 컬러 영상만을 이용한 움직임 영역 추출 결과를 나타내었고, 도 6에는 컬러 영상과 뎁스 영상을 모두 이용한 움직임 영역 추출 결과를 나타내었다. 도 5와 도 6을 비교하면, 컬러 영상만을 이용한 경우 보다 뎁스 영상을 함께 이용한 경우가, 움직임 영역이 보다 정확하게 추출되었음을 확인할 수 있다.

[0040] 도 7에는 컬러 영상과 뎁스 영상을 모두 이용하는 경우에, 움직임 물체를 변형시키면서 움직임 영역을 추출한 결과를 나타내었다. 도 8에 나타난 바와 같이, 어떠한 상황에서도 움직임 영역 추출 성능이 우수하였음을 확인할 수 있다.

[0041] 도 8은 본 발명의 다른 실시예에 따른 영상 시스템의 블럭도이다. 도 8에 도시된 바와 같이, 본 발명의 실시예에 따른 영상 시스템은, 영상-카메라(210), 뎁스-카메라(220), 영상 프로세서(230) 및 출력부(240)를 포함한다.

[0042] 영상-카메라(210)는 컬러 영상을 생성하는데, 생성되는 컬러 영상에는 전술한 컬러 배경 영상과 컬러 입력 영상이 포함된다.

[0043] 뎁스-카메라(220)는 뎁스 영상을 생성하는데, 생성되는 뎁스 영상에는 전술한 뎁스 배경 영상과 뎁스 입력 영상이 포함된다.

[0044] 영상 프로세서(230)는 영상-카메라(210)에서 생성된 컬러 영상과 뎁스-카메라(220)에서 생성된 뎁스 영상을 이용하여, 도 2에 도시된 알고리즘에 따라 움직임 영역을 추출한다.

[0045] 출력부(240)는 영상 프로세서(230)에 의한 움직임 영역 추출 결과를 시각적으로 또는 정보로 출력한다.

[0046] 또한, 이상에서는 본 발명의 바람직한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

부호의 설명

[0047] 210 : 영상-카메라

220 : 뎁스-카메라

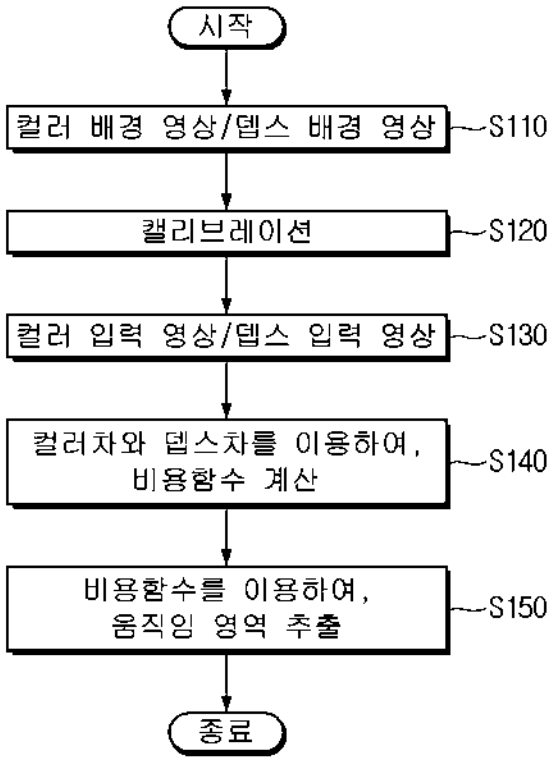
230 : 영상 프로세서

240 : 출력부

도면

도면1

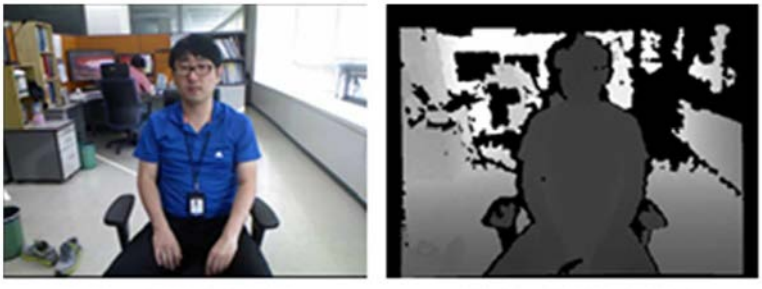
도면2


시작
컬러 배경 영상/뎁스 배경 영상
S110
캘리브레이션
S120
컬러 입력 영상/뎁스 입력 영상
S130
컬러차와 뎁스차를 이용하여,
비용함수 계산
S140
비용함수를 이용하여,
움직임 영역 추출
S150
종료

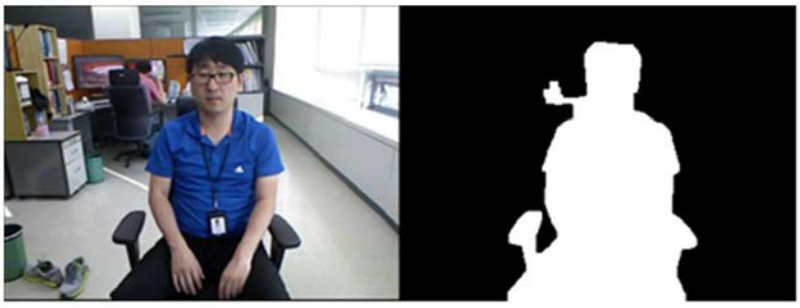
도면3


컬러 배경 영상
뎁스 배경 영상

도면4

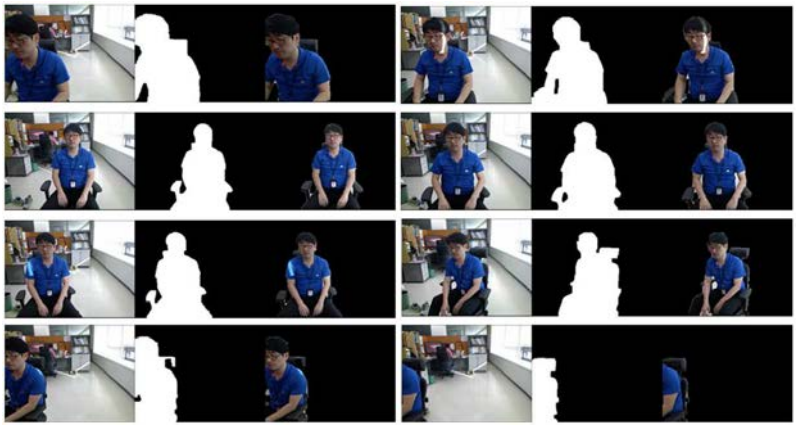

컬러 입력 영상
뎁스 입력 영상

도면5

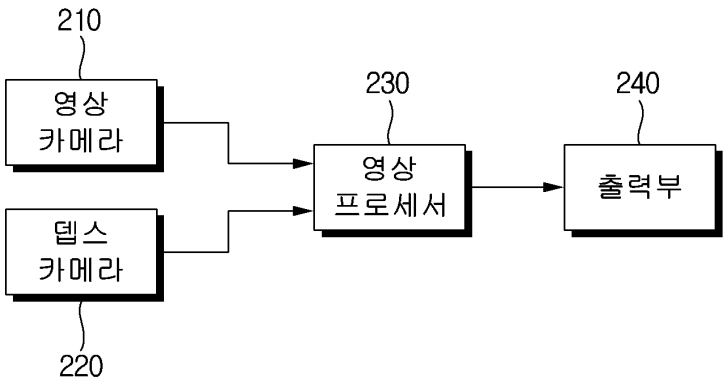


도면6

도면7



도면8



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2014년08월06일
(11) 등록번호 10-1418524
(24) 등록일자 2014년07월04일

(51) 국제특허분류(Int. Cl.)
G06T 1/20 (2006.01) *G06T 3/00* (2006.01)
G06T 5/00 (2006.01)
(21) 출원번호 10-2013-0045330
(22) 출원일자 2013년04월24일
심사청구일자 2013년04월24일
(56) 선행기술조사문헌
류재경 외 2, "객체인식 시스템을 위한 적분이미지의 메모리 축소 방법", 한국통신학회 종합학술발표회 (2010.02.)
류재경 외 2, "SURF 알고리즘 기반 특징점 추출기의 FPGA 설계", 멀티미디어학회논문지 제14권 제3호 (2011.03)
박종일, "병렬 구조를 이용한 SURF 알고리즘 특징점 추출기의 SoC설계", 서경대학교 대학원, 석사학위논문 (2013.02.)
KR1020110002043 A
전체 청구항 수 : 총 8 항

(73) 특허권자
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
최병호
경기 용인시 수지구 대지로 27, 103동 1306호 (죽전동, 한신아파트)
김제우
경기 성남시 분당구 수내로 181, 309동 905호 (분당동, 샛별마을우방아파트)
(뒷면에 계속)
(74) 대리인
유미특허법인

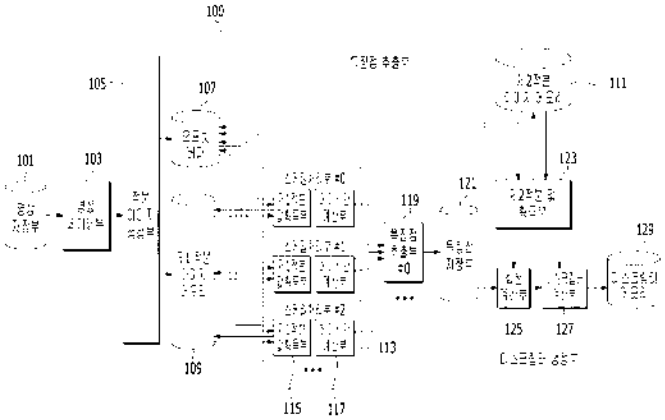
심사관 : 김평수

(54) 발명의 명칭 **하드웨어 장치 및 적분 이미지 생성 방법**

(57) 요 약

하드웨어 장치 및 적분 이미지 생성 방법이 개시된다. 여기서, 영상의 특징점 및 서술자를 연산하는 하드웨어 장치는 입력받은 흑백 영상 전체를 페이딩시켜 최대 비트 수를 기 정의된 비트로 감소시키는 영상 페이딩부, 상기 영상 페이딩부로부터 출력되는 흑백 영상의 화소값을 순차적으로 계산하여 이전 라인까지 누적된 화소값을 오프셋으로 버퍼링하며, 다음 라인의 처음 화소부터 누적된 화소값을 차분값으로 저장하는 적분 이미지 생성부, 그리고 상기 오프셋 및 상기 차분값을 이용하여 상기 특징점을 추출하는데 필요한 제1 적분 이미지의 화소값을 산출하는 제1 적분값 획득부를 포함한다.

대 표 도



(72) 발명자

이상설

인천 남구 경인북길 59-21, 201호 (용현동, 영동하이츠)

황영배

서울 서초구 방배선행길 1, 106동 607호 (방배동, 방배우성아파트)

장성준

서울 서초구 바우뫼로 91, 103동 1401호 (양재동, 우성아파트)

김정호

경기 성남시 분당구 매화로48번길 17-1, 205호 (야탑동)

이 발명을 지원한 국가연구개발사업

과제고유번호	10040246
부처명	지식경제부
연구사업명	산업원천기술개발사업
연구과제명	이동 로봇의 안정적 영상 획득을 통한 3D Depth 정보 획득과 실시간 객체 인식을 위한 로봇비전 SoC 및 모듈 개발
기 여 율	1/1
주관기관	전자부품연구원
연구기간	2011.06.01~2015.05.31

특허청구의 범위

청구항 1

영상의 특징점 및 서술자를 연산하는 하드웨어 장치로서,

입력받은 흑백 영상 전체를 페이딩시켜 최대 비트 수를 기 정의된 비트로 감소시키는 영상 페이딩부,

상기 영상 페이딩부로부터 출력되는 흑백 영상의 화소값을 순차적으로 계산하여 이전 라인까지 누적된 화소값을 오프셋으로 버퍼링하며, 다음 라인의 처음 화소부터 누적된 화소값을 차분값으로 저장하는 적분 이미지 생성부, 그리고

상기 오프셋 및 상기 차분값을 이용하여 상기 특징점을 추출하는데 필요한 제1 적분 이미지의 화소값을 산출하는 제1 적분값 획득부

를 포함하는 하드웨어 장치.

청구항 2

제1항에 있어서,

상기 영상 페이딩부로부터 출력되는 흑백 영상의 이전 라인의 마지막 행 또는 마지막 열의 누적된 화소값인 오프셋이 저장되는 오프셋 버퍼, 그리고

상기 이전 라인의 다음 라인의 처음 화소부터 계산되어 상기 다음 라인의 마지막 행 또는 마지막 열의 누적된 화소값인 차분값이 저장되는 제1 적분 이미지 메모리를 더 포함하고,

상기 제1 적분값 획득부는,

상기 오프셋 버퍼로 접근하여 상기 제1 적분 이미지에 해당하는 이전 라인의 오프셋을 획득하고, 상기 제1 적분 이미지 메모리로 접근하여 상기 이전 라인의 다음 라인의 차분값을 획득하며, 획득한 오프셋 및 획득한 차분값을 합산하여 상기 제1 적분 이미지의 화소값을 계산하는 하드웨어 장치.

청구항 3

제2항에 있어서,

상기 오프셋 버퍼는,

하나 이상의 라인을 포함하는 그룹 단위로 상기 그룹 내 마지막 행 또는 마지막 열의 누적된 화소의 적분값인 오프셋이 상기 그룹 별로 인덱싱되어 저장되는 하드웨어 장치.

청구항 4

제2항에 있어서,

상기 제1 적분값 획득부는 스케일 단위로 복수개이고,

상기 오프셋 버퍼 및 상기 제1 적분 이미지 메모리를 복수개의 상기 제1 적분값 획득부가 서로 공유하는 하드웨어 장치.

청구항 5

제3항에 있어서,

상기 이전 라인의 다음 라인의 처음 화소부터 계산되어 상기 다음 라인의 마지막 행 또는 마지막 열의 누적된 화소값인 차분값이 저장되는 제2 적분 이미지 메모리 그리고

상기 오프셋 버퍼로 접근하여 상기 서술자를 생성하는데 필요한 제2 적분 이미지에 해당하는 오프셋을 획득하고, 상기 제2 적분 이미지 메모리로 접근하여 상기 오프셋 다음 라인의 차분값을 획득하며, 획득한 오프셋 및 획득한 차분값을 합산하여 상기 제2 적분 이미지의 화소값을 계산하는 하드웨어 장치.

청구항 6

하드웨어 장치가 입력받은 영상의 특징점 및 서술자를 연산하기 위한 적분 이미지를 생성하는 방법으로서,

입력받은 흑백 영상 전체에 대해 최대 비트 수를 기 정의된 비트로 감소시키는 페이딩을 수행하는 단계,

상기 페이딩된 흑백 영상의 이전 라인의 마지막 행 또는 마지막 열의 누적된 화소값인 오프셋을 계산하는 단계,

상기 이전 라인의 다음 라인의 처음 화소부터 계산되어 누적된 상기 다음 라인의 마지막 행 또는 마지막 열의 화소값인 차분값을 계산하는 단계, 그리고

상기 오프셋 및 상기 차분값을 이용하여 상기 적분 이미지의 화소값을 산출하는 단계

를 포함하는 적분 이미지 생성 방법.

청구항 7

제6항에 있어서,

상기 차분값을 계산하는 단계는,

상기 특징점을 추출하는데 필요한 제1 적분 이미지의 제1 차분값을 계산하여 저장하는 단계, 그리고

상기 서술자를 추출하는데 필요한 제2 적분 이미지의 제2 차분값을 계산하여 저장하는 단계

를 포함하는 적분 이미지 생성 방법.

청구항 8

제7항에 있어서,

상기 산출하는 단계는,

상기 오프셋 및 상기 제1 차분값을 합산하여 상기 제1 적분 이미지의 화소값을 산출하는 단계, 그리고

상기 오프셋 및 상기 제2 차분값을 합산하여 상기 제2 적분 이미지의 화소값을 산출하는 단계

를 포함하는 적분 이미지 생성 방법.

명 세 서

기 술 분 야

[0001] 본 발명은 하드웨어 장치 및 적분 이미지 생성 방법에 관한 것으로, 특히, 객체○장면 인식 및 하드웨어○시스템-온-칩 기술 분야에 관한 것이다.

배 경 기 술

[0002] SURF(Speeded Up Robust Feature) 알고리즘은 사물의 특징점을 추출하는 대표적인 알고리즘 중의 하나이다.

[0003] SURF 알고리즘은 입력 흑백 영상을 토대로 내부적으로 적분 이미지를 재생성하여 객체○장면 인식을 수행한다. 그런데 입력 해상도가 커짐에 따라 최대 누적값 또한 커진다. 결국 이를 표현하기 위해 필요로 하는 비트 수가 커져 전체 메모리 사용량이 늘어나는 문제가 있다.

[0004] 640×480 해상도의 경우, 적분 이미지의 필요한 최대 비트 수는 27 비트가 된다. 27비트의 307200(640×480)개 값을 저장하는 것은 하드웨어 및 시스템-온-칩 구현에 있어 현실적으로 불가능하다. 알고리즘의 구현을 위해서 이러한 적분 이미지 메모리 사용량 감소 방법이 절실히 요구되고 있다.

발명의 내용

해결하려는 과제

[0005] 따라서, 본 발명이 이루고자 하는 기술적 과제는 적분 이미지 메모리의 크기를 감소시키는 하드웨어 장치 및 적분 이미지 생성 방법을 제공하는 것이다.

과제의 해결 수단

[0006] 본 발명의 하나의 특징에 따르면, 하드웨어 장치는 영상의 특징점 및 서술자를 연산하는 하드웨어 장치로서, 입력받은 흑백 영상 전체를 페이딩시켜 최대 비트 수를 기 정의된 비트로 감소시키는 영상 페이딩부, 상기 영상 페이딩부로부터 출력되는 흑백 영상의 화소값을 순차적으로 계산하여 이전 라인까지 누적된 화소값을 오프셋으로 버퍼링하며, 다음 라인의 처음 화소부터 누적된 화소값을 차분값으로 저장하는 적분 이미지 생성부, 그리고 상기 오프셋 및 상기 차분값을 이용하여 상기 특징점을 추출하는데 필요한 제1 적분 이미지의 화소값을 산출하는 제1 적분값 획득부를 포함한다.

[0007] 상기 하드웨어 장치는,

[0008] 상기 영상 페이딩부로부터 출력되는 흑백 영상의 이전 라인의 마지막 행 또는 마지막 열의 누적된 화소값인 오프셋이 저장되는 오프셋 버퍼, 그리고 상기 이전 라인의 다음 라인의 처음 화소부터 계산되어 상기 다음 라인의 마지막 행 또는 마지막 열의 누적된 화소값인 차분값이 저장되는 제1 적분 이미지 메모리를 더 포함하고,

[0009] 상기 제1 적분값 획득부는,

[0010] 상기 오프셋 버퍼로 접근하여 상기 제1 적분 이미지에 해당하는 이전 라인의 오프셋을 획득하고, 상기 제1 적분 이미지 메모리로 접근하여 상기 이전 라인의 다음 라인의 차분값을 획득하며, 획득한 오프셋 및 획득한 차분값을 합산하여 상기 제1 적분 이미지의 화소값을 계산할 수 있다.

[0011] 상기 오프셋 버퍼는,

[0012] 하나 이상의 라인을 포함하는 그룹 단위로 상기 그룹 내 마지막 행 또는 마지막 열의 누적된 화소의 적분값인 오프셋이 상기 그룹 별로 인덱싱되어 저장될 수 있다.

[0013] 상기 제1 적분값 획득부는 스케일 단위로 복수개이고,

[0014] 상기 오프셋 버퍼 및 상기 제1 적분 이미지 메모리를 복수개의 상기 제1 적분값 획득부가 서로 공유할 수 있다.

[0015] 상기 하드웨어 장치는,

[0016] 상기 이전 라인의 다음 라인의 처음 화소부터 계산되어 상기 다음 라인의 마지막 행 또는 마지막 열의 누적된 화소값인 차분값이 저장되는 제2 적분 이미지 메모리 그리고 상기 오프셋 버퍼로 접근하여 상기 서술자를 생성하는데 필요한 제2 적분 이미지에 해당하는 오프셋을 획득하고, 상기 제2 적분 이미지 메모리로 접근하여 상기 오프셋 다음 라인의 차분값을 획득하며, 획득한 오프셋 및 획득한 차분값을 합산하여 상기 제2 적분 이미지의 화소값을 계산할 수 있다.

[0017] 본 발명의 다른 특징에 따르면, 적분 이미지 생성 방법은 하드웨어 장치가 입력받은 영상의 특징점 및 서술자를 연산하기 위한 적분 이미지를 생성하는 방법으로서, 입력받은 흑백 영상 전체에 대해 최대 비트 수를 기 정의된 비트로 감소시키는 페이딩을 수행하는 단계, 상기 페이딩된 흑백 영상의 이전 라인의 마지막 행 또는 마지막 열의 누적된 화소값인 오프셋을 계산하는 단계, 상기 이전 라인의 다음 라인의 처음 화소부터 계산되어 누적된 상기 다음 라인의 마지막 행 또는 마지막 열의 화소값인 차분값을 계산하는 단계, 그리고 상기 오프셋 및 상기 차분값을 이용하여 상기 적분 이미지의 화소값을 산출하는 단계를 포함한다.

[0018] 상기 차분값을 계산하는 단계는,

[0019] 상기 특징점을 추출하는데 필요한 제1 적분 이미지의 제1 차분값을 계산하여 저장하는 단계, 그리고 상기 서술자를 추출하는데 필요한 제2 적분 이미지의 제2 차분값을 계산하여 저장하는 단계를 포함할 수 있다.

[0020] 상기 산출하는 단계는,

[0021] 상기 오프셋 및 상기 제1 차분값을 합산하여 상기 제1 적분 이미지의 화소값을 산출하는 단계, 그리고 상기 오프셋 및 상기 제2 차분값을 합산하여 상기 제2 적분 이미지의 화소값을 산출하는 단계를 포함할 수 있다.

발명의 효과

[0022] 본 발명의 실시예에 따르면, 페이딩 기법을 통해 입력 영상의 비트 수를 감소시킨다. 또한, 입력 영상의 화소 적분값을 한 라인의 마지막 열에서의 누적된 값을 별도로 저장하여 오프셋으로 사용하고, 새로운 라인의 적분 이미지를 생성 시에는 바로 이 오프셋을 가지고 있기 때문에 처음 화소부터 다시 누적하여 차분값을 형성하며,

오프셋 및 차분값을 합산하여 적분 이미지를 생성함으로써, 전체 화소의 적분값으로 모두 저장하지 않아도 되어 적분 이미지 메모리의 크기를 감소시킬 수 있다.

도면의 간단한 설명

[0023] 도 1은 본 발명의 실시예에 따른 SURF 하드웨어 장치의 구성도이다.

도 2는 본 발명의 실시예에 따른 오프셋 버퍼의 구성도이다.

도 3은 본 발명의 실시예에 따른 적분 이미지 메모리의 구성도이다.

도 4는 본 발명의 실시예에 따른 스케일 처리부 및 특징점 추출부의 구현 예시도이다.

도 5는 본 발명의 실시예에 따른 적분 이미지 생성 및 획득 방법을 나타낸 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0024] 아래에서는 첨부한 도면을 참고로 하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 발명은 여러가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

[0025] 명세서 전체에서, 어떤 부분이 어떤 구성 요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성 요소를 제외하는 것이 아니라 다른 구성 요소를 더 포함할 수 있는 것을 의미한다.

[0026] 또한, 명세서에 기재된 "…부", "…모듈" 의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.

[0027] 이하, 도면을 참조로 하여 본 발명의 실시예에 따른 하드웨어 장치 및 적분 이미지 생성 방법에 대하여 상세히 설명한다.

[0028] 도 1은 본 발명의 실시예에 따른 SURF(Speeded Up Robust Feature) 하드웨어 장치의 구성도이고, 도 2는 본 발명의 실시예에 따른 오프셋 버퍼의 구성도이고, 도 3은 본 발명의 실시예에 따른 적분 이미지 메모리의 구성도이며, 도 4는 본 발명의 실시예에 따른 스케일 처리부 및 특징점 추출부의 구현 예시도이다.

[0029] 먼저, 도 1을 참조하면, SURF 하드웨어 장치(100)는 특징점(Feature point) 및 서술자(Descriptor)를 생성한다. 이때, 특징점을 추출하는 스케일 공간(Scale space)을 표현하기 위해 이미지 피라미드(Image pyramids)를 생성하고 생성된 이미지 피라미드에서 특징점을 추출한다. 이때, 사용되는 적분 이미지(Integral image)는 SURF 알고리즘의 핵심이다.

[0030] SURF 하드웨어 장치(100)는 영상 저장부(101), 영상 페이딩부(103), 적분 이미지 생성부(105), 오프셋 버퍼(107), 제1 적분 이미지 메모리(109), 제2 적분 이미지 메모리(111), 스케일 처리부(113), 제1 적분값 획득부(115), 헤이시안 계산부(117), 특징점 추출부(119), 특징점 저장부(121), 제2 적분값 획득부(123), 회전 계산부(125), 서술자 계산부(127) 및 서술자 저장부(129)를 포함한다.

[0031] 이때, 특징점 저장부(121)를 경계로 좌우 측 2개의 큰 블록으로 생성되는데,좌측은 특징점 추출부, 우측은 서술자 생성부로 구분된다.

[0032] 영상 저장부(101)는 입력받은 흑백 영상을 저장한다.

[0033] 영상 페이딩부(103)는 영상 저장부(101)에 저장된 입력받은 흑백 영상 전체를 페이딩(Fading)시켜 최대 비트 수를 기 정의된 비트로 감소시킨다.

[0034] 흑백 영상의 경우 아날로그 입력을 0부터 255까지로 양자화하여 디지털 값을 사용한다. 여기서, 페이딩 기법은 256개의 양자화 레벨 수를 감소시키는 것이다. 양자화 레벨 수를 줄였을 때, 영상의 명암이나 색은 변할 수 있지만 영상 자체는 크게 변하지 않는다. 객체 인식의 경우는 영상의 명암이나 색에 기반하는 것이 아니기 때문에 인식률에 있어 양자화 레벨 수에 의한 영향은 거의 없다.

[0035] 양자화 레벨 수를 256개에서 192개로 감소시켰을 경우 표현 비트는 8비트에서 6비트로 감소하고, 640×480 해상도의 영상에서 적분 이미지 메모리의 사용량은 약 8% 감소된다.

[0036] 적분 이미지 생성부(105)는 영상 페이딩부(103)로부터 전달받은 페이딩된 흑백 영상을 구성하는 화소들 각각의 화소값을 라인 방향으로 순차적으로 계산하여 누적한다.

[0037] 여기서, 라인은 영상의 행(Row) 단위로 배치된 복수개의 화소로 이루어지거나 또는 영상의 열(Column) 단위로 배치된 복수개의 화소로 이루어질 수 있다.

[0038] 적분 이미지 생성부(105)는 하나 이상의 라인을 포함하는 그룹 단위로 화소값을 순차적으로 계산하여 누적한다. 이때, 그룹의 마지막 행 또는 마지막 열의 누적된 화소값을 해당 그룹의 오프셋(off-set)으로 산출하여 오프셋 버퍼(107)에 저장한다.

[0039] 또한, 적분 이미지 생성부(105)는 오프셋에 기반하여 차분값(Defferential value)을 산출하여 제1 적분 이미지 메모리(109) 및 제2 적분 이미지 메모리(121)에 각각 저장한다.

[0040] 이때, 적분 이미지 생성부(105)는 차분값을 계산시 이전 라인의 누적된 화소값과 관계없이 이전 라인의 다음 라인의 처음 행 또는 처음 열의 화소값부터 새롭게 계산하여 누적한다.

[0041] 오프셋 버퍼(107)는 적분 이미지 생성부(105)가 산출한 오프셋을 그룹 별로 인덱싱하여 저장한다. 그리고 오프셋 버퍼(107)는 레지스터(register)로 구현되어 특징점 추출부와 서술자 생성부의 동시 접근에도 충돌이 발생하지 않는다.

[0042] 이때, 오프셋 버퍼(107)는 도 3과 같이 구현될 수 있다.

[0043] 도 3의 (a)를 참조하면, 적분 이미지 생성부(105)는 N개의 라인(Row or Column #0 ~ Row or Column #N)을 포함하는 그룹 단위로 버퍼링된 오프셋(P1)은 도 3의 (b)와 같이 그룹 별로 오프셋 버퍼(107)에 저장된다.

[0044] 즉, 제1 그룹(RG #0)은 N개의 라인(Row or Column #0 ~ Row or Column #N)을 포함하고, 제1 그룹(RG #0)의 오프셋은 N번째 라인(Row or Column #N)의 마지막 행 또는 마지막 열의 누적된 화소값이 저장된다. 마찬가지로 제2 그룹(RG #1)의 오프셋은 2N번째 라인(Row or Column #2N)의 마지막 행 또는 마지막 열의 누적된 화소값이 저장된다.

[0045] 이때, 640×480, 8비트 흑백 영상의 경우, 2-라인의 그룹 단위로 오프셋을 산출하면, 적분 이미지의 화소값의 표현 비트는 종래 27비트에서 19비트로 감소하게 된다. 이로 인해 적분 이미지 메모리의 사용량을 약 30% 감소시킬 수 있다.

[0046] 다시, 도 1을 참조하면, 제1 적분 이미지 메모리(109)는 특징점 추출에 필요한 적분 이미지의 차분값을 저장한다.

[0047] 제2 적분 이미지 메모리(111)는 서술자 추출에 필요한 적분 이미지의 차분값을 저장한다.

[0048] 적분 이미지 메모리를 이처럼 제1 적분 이미지 메모리(109) 및 제2 적분 이미지 메모리(111)로 구분하여 독립적으로 구성하는 이유는 제1 적분값 획득부(115) 및 제2 적분값 획득부(123)가 순차적으로 동작하지 않고 동시에 동작할 수 있기 때문이다. 만약 적분 이미지 메모리를 하나로 구현하면, 동시 동작시 적분 이미지 메모리 접근 충돌이 발생할 수 있다.

[0049] 또한, 제1 적분값 획득부(115) 및 제2 적분값 획득부(123)의 중간에 FIFO로 인터페이싱하기 때문에 제1 적분값 획득부(115) 및 제2 적분값 획득부(123)는 서로 독립적으로 동작할 수 있다.

[0050] 제1 적분 이미지 메모리(109)는 특징점 추출을 위한 전용 메모리이다. 그리고 복수의 스케일 처리부(113)가 서로 공유하는 메모리이다.

[0051] 제1 적분 이미지 메모리(109)는 페이딩된 입력 흑백 영상 내에서 특정 사각형 영역에서 적분 이미지의 화소값에 대해 필터링을 하게 되는데 특정 사각형 영역의 크기는 크지 않다. 여기서, 특정 사각형 영역은 최대 51×51 크기 혹은 그 이상의 사각형 박스 일 수 있다.

[0052] 제2 적분 이미지 메모리(111)는 서술자 추출을 위한 전용 메모리이다. 제2 적분 이미지 메모리(111)는 페이딩된 입력 흑백 영상 프레임의 적분 이미지의 화소값을 모두 저장한다. 이때, 특징점을 중심으로 상당히 큰 영역의 적분 이미지를 저장한다.

[0053] 여기서, 제1 적분 이미지 메모리(109) 및 제2 적분 이미지 메모리(111)는 도 4와 같이 구현될 수 있다.

[0054] 도 4를 참조하면, 제1 적분 이미지 메모리(109) 및 제2 적분 이미지 메모리(111)는 영상의 전체 라인(Row or

Column #0 ~ Row or Column #N) 별로 차분값이 인덱싱되어 저장된다. 즉, 차분값은 라인 별로 해당하는 라인의 처음 행 또는 처음 열(0,0)의 화소값부터 순차적으로 계산되어 마지막 행 또는 마지막 열의 누적된 화소값이다.

[0055] 다시, 도 1을 참조하면, 스케일 처리부(113)는 특징점 추출시에 복수개의 스케일에 대해 병렬 연산을 한다. 이때, 스케일은 전술한 제1 적분 이미지 메모리(109)의 특정 사각형 영역에 대한 필터링 작업이이고, 예컨대 6개의 스케일을 포함할 수 있다. 따라서, 스케일 처리부(113)는 스케일의 개수만큼 복수개로 구현될 수 있다. 그리고 파이프라인으로 구현되어 병렬로 6개 혹은 그 이상으로 구성되어 연산을 수행할 수 있다.

[0056] 이때, 각각의 스케일 처리부(113)는 특정 사각형의 크기를 달리하여 필터링을 한다. 예를 들어, 각각의 스케일 처리부(113)는 9×9, 15×15, 21×21, 27×27, 39×39, 51×51 크기의 필터 크기로 각각 구현될 수 있다.

[0057] 이러한 스케일 처리부(113) 각각은 제1 적분값 획득부(115) 및 헤이시안(Hessian) 계산부(117)를 포함한다.

[0058] 여기서, 제1 적분값 획득부(115)는 오프셋 버퍼(107)에 접근하여 획득한 오프셋 및 제1 적분 이미지 메모리(109)에 접근하여 획득한 차분값을 합산하여 헤이시안 계산부(117)에서 필요로 하는 적분 이미지의 화소값을 계산한다.

[0059] 즉, 제1 적분값 획득부(115)는 헤이시안 계산부(117)에서 필요로 하는 적분 이미지를 구성하는 복수의 라인 중에서 해당하는 그룹에 인덱싱된 오프셋을 오프셋 버퍼(107)로부터 획득한다. 그리고 그룹의 다음 라인의 차분값을 제1 적분 이미지 메모리(109)로부터 획득한다. 이처럼 획득한 오프셋 및 차분값을 합산하여 적분 이미지 화소값을 계산하여 헤이시안 계산부(117)로 출력한다.

[0060] 도 3 및 도 4를 다시 참조하면, 제1 적분값 획득부(115)는 제1 그룹(RG #0)의 오프셋 및 제1 그룹(RG #0)의 다음 라인(Row or Column N+1)의 차분값을 계산하여 적분 이미지 화소값을 연산할 수 있다.

[0061] 헤이시안 계산부(117)는 제1 적분값 획득부(115)로부터 전달받은 적분 이미지 화소값을 이용하여 박스 필터(Box filter) 연산을 수행하여 헤이시안 행렬식을 계산한다.

[0062] 특징점 추출부(119)는 헤이시안 계산부(117)가 계산한 헤이시안 행렬식을 이용하여 이미지 피라미드를 생성하여 특징점을 추출한다.

[0063] 이때, 특징점 추출부(119)는 스케일 처리부(113)의 개수에 대응하여 복수개 형성되며, 도 5와 같이 구현될 수 있다.

[0064] 도 5를 참조하면, 스케일 처리부(113)가 총 6개라고 하면, 특징점 추출부(119)는 총 4개로 구현된다. 즉, 6개의 스케일 처리부(113)의 인덱스는 각각 S0, S1, S2, S3, S4, S5이고, 4개의 특징점 추출부(119)의 인덱스는 각각 F0, F1, F2, F3이다.

[0065] 여기서, 스케일 처리부(113)가 6개면, 제1 적분값 획득부(115) 및 헤이시안 계산부(117) 역시 각각 6개로 구현됨은 자명하다.

[0066] 이때, F0은 S0, S1, S2와 연결되고, F1은 S0, S1, S3과 연결도며, F2는 S1, S3, S4와 연결되고, F3은 S2, S4, S5와 연결된다. 이처럼, 하나의 특징점 추출부(119)는 3개의 스케일 처리부(113)가 출력하는 특징점을 토대로 최종 특징점을 추출한다.

[0067] 특징점 저장부(121)는 특징점 추출부(119)가 추출하는 특징점을 저장하며, FIFO(First In First Out)로 구현된다.

[0068] 제2 적분값 획득부(123)는 서술자 생성에 필요한 적분 이미지를 구성하는 복수의 라인 중에서 해당하는 그룹에 인덱싱된 오프셋을 오프셋 버퍼(107)로부터 획득한다. 그리고 그룹의 다음 라인의 차분값을 제2 적분 이미지 메모리(111)로부터 획득한다. 이처럼 획득한 오프셋 및 차분값을 합산하여 적분 이미지 화소값을 계산하여 회전 계산부(125) 및 서술자 계산부(127)에게 각각 출력한다.

[0069] 회전 계산부(125)는 제2 적분값 획득부(123)가 획득한 적분 이미지 화소값을 토대로 회전값을 계산한다. 특징점 저장부(121)에서 추출된 특징점의 좌표 및 스케일에 기반하여 특징점을 중심으로 특정 영역의 적분값을 가지고 특징점의 주 방향을 계산한다. 이를 통해 현재 특징점이 어느 정도 회전되어 있는지 알 수 있게 되고 이는 서술자를 계산할 때 이용하는 적분값의 영역을 결정하는 기준으로 사용된다.

[0070] 서술자 계산부(127)는 제2 적분값 획득부(123)가 획득한 적분 이미지 화소값을 토대로 서술자를 계산한다. 그리고 특징점 저장부(121)에 저장된 특징점에 서술자를 부여한다. 서술자는 특징점을 중심으로 스케일에 따른 특정

영역의 적분값을 사용하며 Harr-wavelet 필터링 후, dx, dy, |dx|, |dy|, 총 4개로 계산 및 표현된다. 또한, 서술자를 계산하기 위해 사용하는 적분값의 영역은 앞서 계산된 회전(각)에 의해 조정된다.

[0071] 서술자 저장부(129)는 서술자 계산부(127)가 계산한 서술자들이 저장된다.

[0072] 도 5는 본 발명의 실시예에 따른 적분 이미지 생성 및 획득 방법을 나타낸 순서도이다. 즉, 전술한 SURF 하드웨어 장치(100)의 적분 이미지 생성 및 획득 동작을 나타낸 순서도이다.

[0073] 도 5를 참조하면, 영상 페이딩부(103)가 입력받은 흑백 영상 전체에 대해 페이딩 기법을 적용하여 최대 비트 수를 기 정의된 비트로 감소시킨다(S101).

[0074] 적분 이미지 생성부(105)는 페이딩된 흑백 영상의 하나 이상의 라인을 포함하는 그룹 단위로 화소값을 순차적으로 계산하여 그룹의 마지막 행 또는 마지막 열의 누적된 화소값을 해당 그룹의 오프셋(off-set)으로 산출(S103)하여 오프셋 버퍼(107)에 저장한다(S105).

[0075] 또한, 적분 이미지 생성부(105)는 라인 별로 라인의 마지막 행 또는 마지막 열의 누적된 화소값을 해당 라인의 차분값(Defferential value)으로 산출(S107)하여 제1 적분 이미지 메모리(109) 및 제2 적분 이미지 메모리(121)에 각각 저장한다(S109).

[0076] 다음, 제1 적분값 획득부(115)는 헤이시안 계산부(117)에서 필요로 하는 적분 이미지를 구성하는 복수의 라인 중에서 해당하는 그룹에 인덱싱된 오프셋을 오프셋 버퍼(107)로부터 획득한다(S111). 그리고 그룹의 다음 라인의 차분값을 제1 적분 이미지 메모리(109)로부터 획득한다(S113).

[0077] 다음, 제1 적분값 획득부(115)는 S111 단계에서 획득한 오프셋 및 S113 단계에서 획득한 차분값을 합산하여 적분 이미지 화소값을 계산(S115)하여 헤이시안 계산부(117)로 출력한다.

[0078] 또한, 제2 적분값 획득부(123)는 서술자 생성에 필요한 적분 이미지를 구성하는 복수의 라인 중에서 해당하는 그룹에 인덱싱된 오프셋을 오프셋 버퍼(107)로부터 획득한다(S117). 그리고 그룹의 다음 라인의 차분값을 제2 적분 이미지 메모리(111)로부터 획득한다(S119).

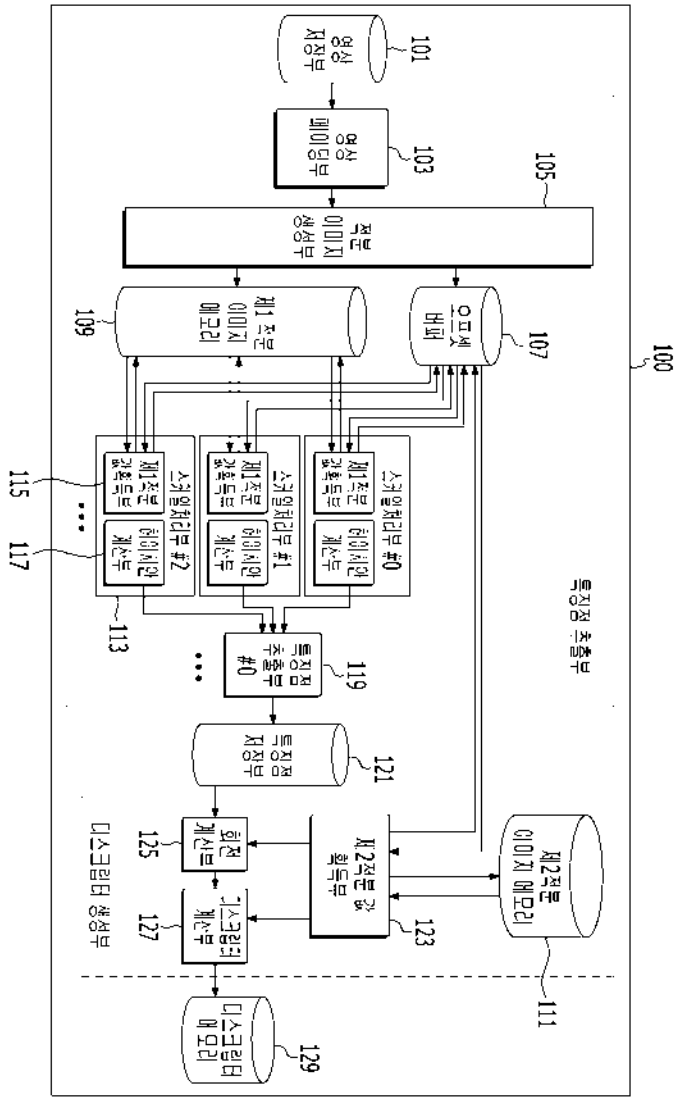
[0079] 다음, 제2 적분값 획득부(123)는 S119 단계에서 획득한 오프셋 및 S121 단계에서 획득한 차분값을 합산하여 적분 이미지 화소값을 계산(S121)하여 회전 계산부(125) 및 서술자 계산부(127)에게 각각 출력한다.

[0080] 이상에서 설명한 본 발명의 실시예는 장치 및 방법을 통해서만 구현이 되는 것은 아니며, 본 발명의 실시예의 구성에 대응하는 기능을 실현하는 프로그램 또는 그 프로그램이 기록된 기록 매체를 통해 구현될 수도 있다.

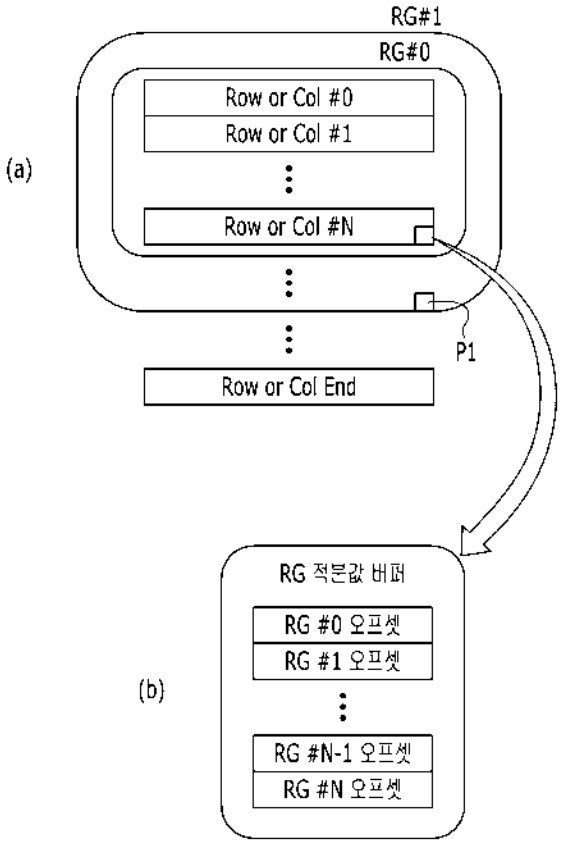
[0081] 이상에서 본 발명의 실시예에 대하여 상세하게 설명하였지만 본 발명의 권리범위는 이에 한정되는 것은 아니고 다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한 당업자의 여러 변형 및 개량 형태 또한 본 발명의 권리범위에 속하는 것이다.

도면

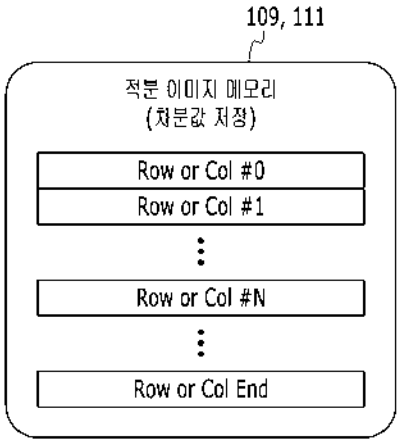
도면1


영상 저장부
101
영상 메인당부
103
105
제1 조문 이미지 에모리
109
오므셋 버퍼
107
100
스케일처리부 #2
스케일처리부 #1
스케일처리부 #0
제1조문 값획득부
히어세연 계산부
115
117
113
특징점 추출부 #0
119
특징점 저장부
121
회전 계산부
125
기스컬러 계산부
127
제2조문 값 획득부
123
제2조문 이미지 에모리
111
디스크립터 생성부
디스크립터 에모리
129
특징점 추출부

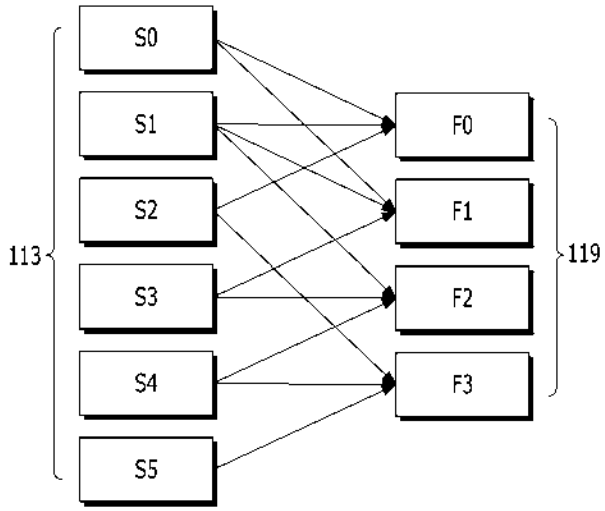
도면2


RG#1
RG#0
Row or Col #0
Row or Col #1
Row or Col #N
Row or Col End
P1
(a)
(b)
RG 적분값 버퍼
RG #0 오프셋
RG #1 오프셋
RG #N-1 오프셋
RG #N 오프셋

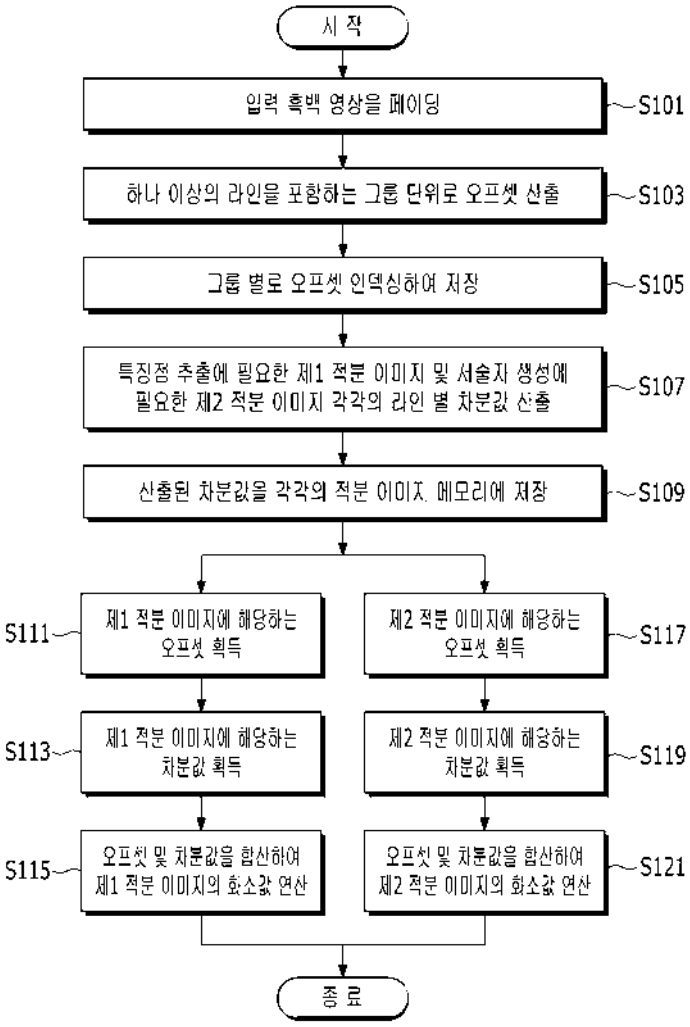
도면3


109, 111
적분 이미지 메모리
(차분값 저장)
Row or Col #0
Row or Col #1
Row or Col #N
Row or Col End

도면4



도면5



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 7

【변경전】

"제1 적분 이지미"

【변경후】

"제1 적분 이미지"