



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2020년03월02일  
(11) 등록번호 10-2083166  
(24) 등록일자 2020년02월25일

(51) 국제특허분류(Int. Cl.)  
G06T 3/40 (2006.01) G06T 5/00 (2019.01)  
(52) CPC특허분류  
G06T 3/4053 (2013.01)  
G06T 5/007 (2013.01)  
(21) 출원번호 10-2019-0046702  
(22) 출원일자 2019년04월22일  
심사청구일자 2019년04월22일  
(56) 선행기술조사문헌  
KR1020180062422 A\*  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
한국과학기술원  
대전광역시 유성구 대학로 291(구성동)  
(72) 발명자  
김문철  
대전광역시 서구 둔산로 155, 108동 805호 (둔산동, 크로바아파트)  
김수예  
대전광역시 유성구 대학로 291 (구성동, 한국과학기술원)  
(74) 대리인  
특허법인 무한

전체 청구항 수 : 총 20 항

심사관 : 박상철

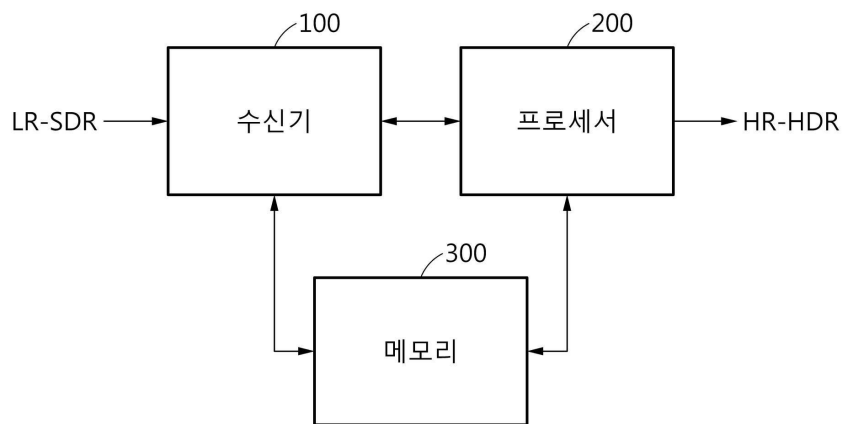
(54) 발명의 명칭 이미지 처리 방법 및 장치

(57) 요약

이미지 처리 방법 및 장치가 개시된다. 일 실시예에 따른 이미지 처리 방법은, 이미지를 수신하는 단계와, 상기 이미지를 필터링하여 제1 레이어 및 제2 레이어를 생성하는 단계와, 상기 제1 레이어에 기초하여 제1 변조(modulation)를 수행함으로써 제1 특징맵을 추출하는 단계와, 상기 제1 레이어 및 제2 레이어에 기초하여 제2 변조를 수행함으로써 제2 특징맵을 추출하는 단계와, 상기 제1 특징맵 및 상기 제2 특징맵에 기초하여 컨벌루션 연산을 포함하는 잔차 연산을 수행함으로써 출력 이미지를 생성하는 단계를 포함한다.

대표도 - 도1

10



(52) CPC특허분류

G06T 2207/20016 (2013.01)

G06T 2207/20024 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 1711082864

부처명 과학기술정보통신부

연구관리전문기관 정보통신기술진흥센터

연구사업명 방송통신산업기술개발(R&D)

연구과제명 스마트 방송 미디어를 위한 지능형 고실감 영상처리 연구

기 여 율 1/1

주관기관 한국과학기술원

연구기간 2019.01.01 ~ 2019.12.31

---

## 명세서

### 청구범위

#### 청구항 1

이미지를 수신하는 단계;

상기 이미지를 상이한 하나 이상의 신호 성분 레이어로 분해하는 단계;

상기 하나 이상의 신호 성분 레이어에 포함된 제1 레이어에 기초하여 제1 변조(modulation)을 수행함으로써 제1 특징맵을 추출하는 단계;

상기 하나 이상의 신호 성분 레이어에 포함된 제2 레이어 및 상기 제1 레이어에 기초하여 제2 변조를 수행함으로써 제2 특징맵을 추출하는 단계; 및

상기 제1 특징맵 및 상기 제2 특징맵에 기초하여 컨벌루션 연산을 포함하는 잔차 연산을 수행함으로써 초해상화(super-resolution)와 역톤매핑(inverse tone mapping)이 동시에 상기 이미지에 수행된 출력 이미지를 생성하는 단계를

포함하고,

상기 제1 레이어와 상기 제2 레이어는 상기 이미지의 이미지 성분에 대해 상이하게 분해된 이미지 처리 방법.

#### 청구항 2

제1항에 있어서,

상기 분해하는 단계는,

상기 이미지를 필터링하는 단계;

필터링된 이미지와 상기 이미지를 연결(concatenate)하여 상기 제1 레이어를 생성하는 단계;

상기 이미지에 대해 상기 필터링된 이미지에 기초한 연산을 수행하는 단계; 및

연산이 수행된 이미지와 상기 이미지를 연결하여 상기 제2 레이어를 생성하는 단계

를 포함하는 이미지 처리 방법.

#### 청구항 3

제1항에 있어서,

상기 제1 특징맵을 추출하는 단계는,

상기 제1 레이어에 기초하여 비선형화 및 컨벌루션 연산을 포함하는 잔차 연산을 수행함으로써 잔차 특징맵을 생성하는 단계; 및

상기 잔차 특징맵 및 상기 제1 레이어에 기초하여 상기 제1 변조를 수행함으로써 상기 제1 특징맵을 추출하는 단계

를 포함하는 이미지 처리 방법.

#### 청구항 4

제3항에 있어서,

상기 제2 특징맵을 추출하는 단계는,

상기 제2 레이어 및 상기 잔차 특징맵에 기초하여 상기 제2 변조를 수행함으로써 잔차 스킵 변조 특징맵을 생성하는 단계; 및

상기 잔차 스킵 변조 특징맵 및 상기 제1 변조의 출력에 기초하여 비선형화 및 컨벌루션 연산을 포함하는 잔차 스킵 연산을 수행함으로써 상기 제2 특징맵을 추출하는 단계

를 포함하는 이미지 처리 방법.

#### 청구항 5

제3항에 있어서,

상기 잔차 특징맵을 생성하는 단계는,

상기 제1 레이어에 적어도 한 번의 비선형화 및 적어도 한 번의 컨벌루션 연산을 수행하는 단계; 및

상기 적어도 한 번의 비선형화의 입력과 상기 적어도 한 번의 컨벌루션 연산의 출력을 더하여 상기 잔차 특징맵을 생성하는 단계

를 포함하는 이미지 처리 방법.

#### 청구항 6

제3항에 있어서,

상기 잔차 특징맵 및 상기 제1 레이어에 기초하여 상기 제1 변조를 수행함으로써 상기 제1 특징맵을 추출하는 단계는,

상기 잔차 특징맵에 적어도 한 번의 비선형화 및 적어도 한 번의 컨벌루션 연산을 수행하는 단계;

상기 제1 레이어에 적어도 한 번의 컨벌루션 연산 및 적어도 한 번의 비선형화를 수행하여 제1 공유 변조 특징맵을 생성하는 단계;

적어도 한 번의 비선형화와 적어도 한 번의 컨벌루션 연산이 수행된 잔차 특징맵 및 상기 제1 공유 변조 특징에 기초한 원소별 곱셈(element-wise multiplication)을 수행하는 단계; 및

상기 원소별 곱셈의 출력과 상기 잔차 특징맵을 더하여 상기 제1 특징맵을 추출하는 단계

를 포함하는 이미지 처리 방법.

#### 청구항 7

제4항에 있어서,

상기 잔차 스킵 변조 특징맵을 생성하는 단계는,

상기 제2 레이어에 기초한 특징맵과 상기 잔차 특징맵을 연결(concatenate)하고, 비선형화 및 디멘전 감소(dimension reduction)를 수행하여 제1 스킵 특징맵을 생성하는 단계;

상기 제1 스킵 특징맵에 적어도 한 번의 컨벌루션 연산 및 비선형화를 수행하는 단계;

상기 제2 레이어에 적어도 한 번의 컨벌루션 연산 및 적어도 한 번의 비선형화를 수행하여 제2 공유 변조 특징맵을 생성하는 단계;

상기 적어도 한 번의 컨벌루션 연산과 비선형화가 수행된 제1 스킵 특징맵 및 상기 제2 공유 변조 특징에 기초한 원소별 곱셈을 수행하는 단계; 및

상기 원소별 곱셈의 출력과 상기 제2 레이어에 기초한 특징맵을 더하여 상기 잔차 스킵 변조 특징맵을 생성하는

단계

를 포함하는 이미지 처리 방법.

### 청구항 8

제4항에 있어서,

상기 제2 특징맵을 추출하는 단계는,

상기 잔차 스킵 변조 특징맵과 상기 제1 변조의 출력을 연결하고 비선형화 및 디멘전 감소를 수행하여 제2 스킵 특징맵을 생성하는 단계;

상기 제2 스킵 특징맵에 적어도 한 번의 컨벌루션 연산 및 비선형화를 수행하는 단계; 및

상기 적어도 한 번의 컨벌루션 연산 및 비선형화가 수행된 제2 스킵 특징맵과 상기 잔차 스킵 변조 특징맵을 더하여 상기 제2 특징맵을 추출하는 단계

를 포함하는 이미지 처리 방법.

### 청구항 9

제1항에 있어서,

상기 출력 이미지를 생성하는 단계는,

상기 제1 특징맵과 상기 제2 특징맵을 연결하여 비선형화 및 디멘전 감소를 수행하는 단계;

상기 디멘전 감소의 출력에 기초하여 비선형화 연산 및 컨벌루션 연산을 포함하는 상기 잔차 연산을 적어도 한번 수행함으로써 복수의 채널을 생성하는 단계;

상기 복수의 채널을 결합하는 단계; 및

상기 이미지를 업샘플링하여 결합된 복수의 채널과 더함으로써 상기 출력 이미지를 생성하는 단계

를 포함하는 이미지 처리 방법.

### 청구항 10

제9항에 있어서,

상기 잔차 연산은,

적어도 한 번의 비선형화;

적어도 한 번의 컨벌루션 연산; 및

상기 적어도 한 번의 비선형화의 입력과 상기 적어도 한 번의 컨벌루션 연산의 출력의 합

을 포함하는 이미지 처리 방법.

### 청구항 11

이미지를 수신하는 수신기; 및

상기 이미지를 상이한 하나 이상의 신호 성분 레이어로 분해하고, 상기 하나 이상의 신호 성분 레이어에 포함된 제1 레이어에 기초하여 제1 변조(modulation)를 수행함으로써 제1 특징맵을 추출하고, 상기 하나 이상의 신호 성분 레이어에 포함된 제2 레이어 및 상기 제1 레이어에 기초하여 제2 변조를 수행함으로써 제2 특징맵을 추출하고, 상기 제1 특징맵 및 상기 제2 특징맵에 기초하여 컨벌루션 연산을 포함하는 잔차 연산을 수행함으로써 초

해상화(super-resolution)와 역톤매핑(inverse tone mapping)이 동시에 상기 이미지에 수행된 출력 이미지를 생성하는 프로세서

를 포함하고,

상기 제1 레이어와 상기 제2 레이어는 상기 이미지의 이미지 성분에 대해 상이하게 분해된 이미지 처리 장치.

**청구항 12**

제11항에 있어서,

상기 프로세서는,

이미지를 필터링하고, 필터링된 이미지와 상기 이미지를 연결(concatenate)하여 상기 제1 레이어를 생성하고, 상기 이미지에 대해 상기 필터링된 이미지에 기초한 연산을 수행하고, 연산이 수행된 이미지와 상기 이미지를 연결하여 상기 제2 레이어를 생성하는

이미지 처리 장치.

**청구항 13**

제11항에 있어서,

상기 프로세서는,

상기 제1 레이어에 기초하여 비선형화 및 컨벌루션 연산을 포함하는 잔차 연산을 수행함으로써 잔차 특징맵을 생성하고, 상기 잔차 특징맵 및 상기 제1 레이어에 기초하여 상기 제1 변조를 수행함으로써 상기 제1 특징맵을 추출하는

이미지 처리 장치.

**청구항 14**

제13항에 있어서,

상기 프로세서는,

상기 제2 레이어 및 상기 잔차 특징맵에 기초하여 상기 제2 변조를 수행함으로써 잔차 스킵 변조 특징맵을 생성하고, 상기 잔차 스킵 변조 특징맵 및 상기 제1 변조의 출력에 기초하여 비선형화 및 컨벌루션 연산을 포함하는 잔차 스킵 연산을 수행함으로써 상기 제2 특징맵을 추출하는

이미지 처리 장치.

**청구항 15**

제13항에 있어서,

상기 프로세서는,

상기 제1 레이어에 적어도 한 번의 비선형화 및 적어도 한 번의 컨벌루션 연산을 수행하고, 상기 적어도 한 번의 비선형화의 입력과 상기 적어도 한 번의 컨벌루션 연산의 출력을 더하여 상기 잔차 특징맵을 생성하는

이미지 처리 장치.

**청구항 16**

제13항에 있어서,

상기 프로세서는,

상기 잔차 특징맵에 적어도 한 번의 비선형화 및 적어도 한 번의 컨벌루션 연산을 수행하고, 상기 제1 레이어에 적어도 한 번의 컨벌루션 연산 및 적어도 한 번의 비선형화를 수행하여 제1 공유 변조 특징을 생성하고, 적어도 한 번의 비선형화와 적어도 한 번의 컨벌루션 연산이 수행된 잔차 특징맵 및 상기 제1 공유 변조 특징에 기초한 원소별 곱셈(element-wise multiplication)을 수행하고, 상기 원소별 곱셈의 출력과 상기 잔차 특징맵을 더하여 상기 제1 특징맵을 추출하는

이미지 처리 장치.

#### 청구항 17

제14항에 있어서,

상기 프로세서는,

상기 제2 레이어에 기초한 특징맵과 상기 잔차 특징맵을 연결(concatenate)하고, 비선형화 및 디멘전 감소(dimension reduction)를 수행하여 제1 스킵 특징맵을 생성하고, 상기 제1 스킵 특징맵에 적어도 한 번의 컨벌루션 연산 및 비선형화를 수행하고, 상기 제2 레이어에 적어도 한 번의 컨벌루션 연산 및 적어도 한 번의 비선형화를 수행하여 제2 공유 변조 특징을 생성하고, 상기 적어도 한 번의 컨벌루션 연산과 비선형화가 수행된 제1 스킵 특징맵 및 상기 제2 공유 변조 특징에 기초한 원소별 곱셈을 수행하고, 상기 원소별 곱셈의 출력과 상기 제2 레이어에 기초한 특징맵을 더하여 상기 잔차 스킵 변조 특징맵을 생성하는

이미지 처리 장치.

#### 청구항 18

제14항에 있어서,

상기 프로세서는,

상기 잔차 스킵 변조 특징맵과 상기 제1 변조의 출력을 연결하고 비선형화 및 디멘전 감소를 수행하여 제2 스킵 특징맵을 생성하고, 상기 제2 스킵 특징맵에 적어도 한 번의 컨벌루션 연산 및 비선형화를 수행하고, 상기 적어도 한 번의 컨벌루션 연산 및 비선형화가 수행된 제2 스킵 특징맵과 상기 잔차 스킵 변조 특징맵을 더하여 상기 제2 특징맵을 추출하는

이미지 처리 장치.

#### 청구항 19

제11항에 있어서,

상기 프로세서는,

상기 제1 특징맵과 상기 제2 특징맵을 연결하여 비선형화 및 디멘전 감소를 수행하고, 상기 디멘전 감소의 출력에 기초하여 비선형화 연산 및 컨벌루션 연산을 포함하는 상기 잔차 연산을 적어도 한 번 수행함으로써 복수의 채널을 생성하고, 상기 복수의 채널을 결합하고, 상기 이미지를 업샘플링하여 결합된 복수의 채널과 더함으로써 상기 출력 이미지를 생성하는

이미지 처리 장치.

#### 청구항 20

제19항에 있어서,  
 상기 잔차 연산은,  
 적어도 한 번의 비선형화;  
 적어도 한 번의 컨벌루션 연산; 및  
 상기 적어도 한 번의 비선형화의 입력과 상기 적어도 한 번의 컨벌루션 연산의 출력의 합  
 을 포함하는 이미지 처리 장치.

**발명의 설명**

**기술 분야**

[0001] 아래 실시예들은 인공 신경망을 이용하여 초해상화(super-resolution)와 역톤매핑(inverse tone mapping)이 동시에 수행된 이미지를 생성하는 방법 및 장치에 관한 것이다.

**배경 기술**

[0002] 현대의 소비자 디스플레이(예를 들어, TV)는 HDR(High Dynamic Range) 및 최대 8K-UHD(Ultra High Definition)의 HR(High Resolution)을 렌더링하는 진보된 기능을 가진다. 결과적으로, UHD HDR TV/방송 및 스트리밍은 고품질 프리미엄 서비스로 부상했다.

[0003] 그러나, 원본 UHD HDR 비디오 콘텐츠가 없기 때문에 기존의 LR(Low Resolution) SDR(Standard Dynamic Range) 비디오를 UHD HDR 버전으로 변환하려면 적절한 변환 기술이 필요하다.

[0004] 일반적인 HDR TV는 최대 4K-UHD(3840×2160)의 공간 해상도(spatial resolution)를 지원하고, 최근, 8K-UHD(7680×4320) TV가 출시되었지만, 4K 및 8K-UHD 콘텐츠가 부족하다. SR(Super-Resolution)과 ITM(Inverse Tone-Mapping)은 고품질 비디오 콘텐츠를 생성하는데 있어 중요하다.

**발명의 내용**

**해결하려는 과제**

[0005] 실시예들은 인공 신경망을 이용하여 이미지를 처리하는 기술을 제공할 수 있다.

**과제의 해결 수단**

[0006] 일 실시예에 따른 이미지 처리 방법은, 이미지를 수신하는 단계와, 상기 이미지를 하나 이상의 신호 성분 레이어로 분해하는 단계와, 상기 하나 이상의 신호 성분 레이어에 포함된 제1 레이어에 기초하여 제1 변조(modulation)를 수행함으로써 제1 특징맵을 추출하는 단계와, 상기 하나 이상의 신호 성분 레이어에 포함된 제2 레이어 및 상기 제1 레이어에 기초하여 제2 변조를 수행함으로써 제2 특징맵을 추출하는 단계와, 상기 제1 특징맵 및 상기 제2 특징맵에 기초하여 컨벌루션 연산을 포함하는 잔차 연산을 수행함으로써 초해상화와 역톤매핑이 동시에 수행된 출력 이미지를 생성하는 단계를 포함한다.

[0007] 상기 분해하는 단계는, 상기 이미지를 필터링하는 단계와, 필터링된 이미지와 상기 이미지를 연결(concatenate)하여 상기 제1 레이어를 생성하는 단계와, 상기 이미지에 대해 상기 필터링된 이미지에 기초한 연산을 수행하는 단계와, 연산이 수행된 이미지와 상기 이미지를 연결하여 상기 제2 레이어를 생성하는 단계를 포함할 수 있다.

[0008] 상기 제1 특징맵을 추출하는 단계는, 상기 제1 레이어에 기초하여 비선형화 및 컨벌루션 연산을 포함하는 잔차 연산을 수행함으로써 잔차 특징맵을 생성하는 단계와, 상기 잔차 특징맵 및 상기 제1 레이어에 기초하여 상기 제1 변조를 수행함으로써 상기 제1 특징맵을 추출하는 단계를 포함할 수 있다.

[0009] 상기 제2 특징맵을 추출하는 단계는, 상기 제2 레이어 및 상기 잔차 특징맵에 기초하여 상기 제2 변조를 수행함으로써 잔차 스킵 변조 특징맵을 생성하는 단계와, 상기 잔차 스킵 변조 특징맵 및 상기 제1 변조의 출력에 기초하여 비선형화 및 컨벌루션 연산을 포함하는 잔차 스킵 연산을 수행함으로써 상기 제2 특징맵을 추출하는 단계를 포함할 수 있다.



- [0010] 상기 잔차 특징맵을 생성하는 단계는, 상기 제1 레이어에 적어도 한 번의 비선형화 및 적어도 한 번의 컨벌루션 연산을 수행하는 단계와, 상기 적어도 한 번의 비선형화의 입력과 상기 적어도 한 번의 컨벌루션 연산의 출력을 더하여 상기 잔차 특징맵을 생성하는 단계를 포함할 수 있다.
- [0011] 상기 잔차 특징맵 및 상기 제1 레이어에 기초하여 상기 제1 변조를 수행함으로써 상기 제1 특징맵을 추출하는 단계는, 상기 잔차 특징맵에 적어도 한 번의 비선형화 및 적어도 한 번의 컨벌루션 연산을 수행하는 단계와, 상기 제1 레이어에 적어도 한 번의 컨벌루션 연산 및 적어도 한 번의 비선형화를 수행하여 제1 공유 변조 특징을 생성하는 단계와, 적어도 한 번의 비선형화 및 적어도 한 번의 컨벌루션 연산이 수행된 잔차 특징맵 및 상기 제1 공유 변조 특징에 기초한 원소별 곱셈(element-wise multiplication)을 수행하는 단계와, 상기 원소별 곱셈의 출력과 상기 잔차 특징맵을 더하여 상기 제1 특징맵을 추출하는 단계를 포함할 수 있다.
- [0012] 상기 잔차 스킵 변조 특징맵을 생성하는 단계는, 상기 제2 레이어에 기초한 특징맵과 상기 잔차 특징맵을 연결(concatenate)하고, 비선형화 및 디멘전 감소(dimension reduction)를 수행하여 제1 스킵 특징맵을 생성하는 단계와, 상기 제1 스킵 특징맵에 적어도 한 번의 컨벌루션 연산 및 비선형화를 수행하는 단계와, 상기 제2 레이어에 적어도 한 번의 컨벌루션 연산 및 적어도 한 번의 비선형화를 수행하여 제2 공유 변조 특징을 생성하는 단계와, 상기 적어도 한 번의 컨벌루션 연산과 비선형화가 수행된 제1 스킵 특징맵 및 상기 제2 공유 변조 특징에 기초한 원소별 곱셈을 수행하는 단계와, 상기 원소별 곱셈의 출력과 상기 제2 레이어에 기초한 특징맵을 더하여 상기 잔차 스킵 변조 특징맵을 생성하는 단계를 포함할 수 있다.
- [0013] 상기 제2 특징맵을 추출하는 단계는, 상기 잔차 스킵 변조 특징맵과 상기 제1 변조의 출력을 연결하고 비선형화 및 디멘전 감소를 수행하여 제2 스킵 특징맵을 생성하는 단계와, 상기 제2 스킵 특징맵에 적어도 한 번의 컨벌루션 연산 및 비선형화를 수행하는 단계와, 상기 적어도 한 번의 컨벌루션 연산 및 비선형화가 수행된 제2 스킵 특징맵과 상기 잔차 스킵 변조 특징맵을 더하여 상기 제2 특징맵을 추출하는 단계를 포함할 수 있다.
- [0014] 상기 출력 이미지를 생성하는 단계는, 상기 제1 특징맵과 상기 제2 특징맵을 연결하여 비선형화 및 디멘전 감소를 수행하는 단계와, 상기 디멘전 감소의 출력에 기초하여 비선형화 연산 및 컨벌루션 연산을 포함하는 상기 잔차 연산을 적어도 한 번 수행함으로써 복수의 채널을 생성하는 단계와, 상기 복수의 채널을 결합하는 단계와, 상기 이미지를 업샘플링하여 결합된 복수의 채널과 더함으로써 상기 출력 이미지를 생성하는 단계를 포함할 수 있다.
- [0015] 상기 잔차 연산은, 적어도 한 번의 비선형화와, 적어도 한 번의 컨벌루션 연산과, 상기 적어도 한 번의 비선형화의 입력과 상기 적어도 한 번의 컨벌루션 연산의 출력의 합을 포함할 수 있다.
- [0016] 일 실시예에 따른 이미지 처리 장치는, 이미지를 수신하는 수신기와, 상기 이미지를 하나 이상의 신호 성분 레이어로 분해하고, 상기 하나 이상의 신호 성분 레이어에 포함된 상기 제1 레이어에 기초하여 제1 변조(modulation)를 수행함으로써 제1 특징맵을 추출하고, 상기 하나 이상의 신호 성분 레이어에 포함된 제2 레이어 및 상기 제1 레이어에 기초하여 제2 변조를 수행함으로써 제2 특징맵을 추출하고, 상기 제1 특징맵 및 상기 제2 특징맵에 기초하여 컨벌루션 연산을 포함하는 잔차 연산을 수행함으로써 초해상화와 역톤매핑이 동시에 수행된 출력 이미지를 생성하는 프로세서를 포함한다.
- [0017] 상기 프로세서는, 이미지를 필터링하고, 필터링된 이미지와 상기 이미지를 연결(concatenate)하여 상기 제1 레이어를 생성하고, 상기 이미지에 대해 상기 필터링된 이미지에 기초한 연산을 수행하고, 연산이 수행된 이미지와 상기 이미지를 연결하여 상기 제2 레이어를 생성할 수 있다.
- [0018] 상기 프로세서는, 상기 제1 레이어에 기초하여 비선형화 및 컨벌루션 연산을 포함하는 잔차 연산을 수행함으로써 잔차 특징맵을 생성하고, 상기 잔차 특징맵 및 상기 제1 레이어에 기초하여 상기 제1 변조를 수행함으로써 상기 제1 특징맵을 추출할 수 있다.
- [0019] 상기 프로세서는, 상기 제2 레이어 및 상기 잔차 특징맵에 기초하여 상기 제2 변조를 수행함으로써 잔차 스킵 변조 특징맵을 생성하고, 상기 잔차 스킵 변조 특징맵 및 상기 제1 변조의 출력에 기초하여 비선형화 및 컨벌루션 연산을 포함하는 잔차 스킵 연산을 수행함으로써 상기 제2 특징맵을 추출할 수 있다.
- [0020] 상기 프로세서는, 상기 제1 레이어에 적어도 한 번의 비선형화 및 적어도 한 번의 컨벌루션 연산을 수행하고, 상기 적어도 한 번의 비선형화의 입력과 상기 적어도 한 번의 컨벌루션 연산의 출력을 더하여 상기 잔차 특징맵을 생성할 수 있다.
- [0021] 상기 프로세서는, 상기 잔차 특징맵에 적어도 한 번의 비선형화 및 적어도 한 번의 컨벌루션 연산을 수행하고,

상기 제1 레이어에 적어도 한 번의 컨벌루션 연산 및 적어도 한 번의 비선형화를 수행하여 제1 공유 변조 특징을 생성하고, 적어도 한 번의 비선형화와 적어도 한 번의 컨벌루션 연산이 수행된 잔차 특징맵 및 상기 제1 공유 변조 특징에 기초한 원소별 곱셈(element-wise multiplication)을 수행하고, 상기 원소별 곱셈의 출력과 상기 잔차 특징맵을 더하여 상기 제1 특징맵을 추출할 수 있다.

[0022] 상기 프로세서는, 상기 제2 레이어에 기초한 특징맵과 상기 잔차 특징맵을 연결(concatenate)하고, 비선형화 및 디멘전 감소(dimension reduction)를 수행하여 제1 스킵 특징맵을 생성하고, 상기 제1 스킵 특징맵에 적어도 한 번의 컨벌루션 연산 및 비선형화를 수행하고, 상기 제2 레이어에 적어도 한 번의 컨벌루션 연산 및 적어도 한 번의 비선형화를 수행하여 제2 공유 변조 특징을 생성하고, 상기 적어도 한 번의 컨벌루션 연산과 비선형화가 수행된 제1 스킵 특징맵 및 상기 제2 공유 변조 특징에 기초한 원소별 곱셈을 수행하고, 상기 원소별 곱셈의 출력과 상기 제2 레이어에 기초한 특징맵을 더하여 상기 잔차 스킵 변조 특징맵을 생성할 수 있다.

[0023] 상기 프로세서는, 상기 잔차 스킵 변조 특징맵과 상기 제1 변조의 출력을 연결하고 비선형화 및 디멘전 감소를 수행하여 제2 스킵 특징맵을 생성하고, 상기 제2 스킵 특징맵에 적어도 한 번의 컨벌루션 연산 및 비선형화를 수행하고, 상기 적어도 한 번의 컨벌루션 연산 및 비선형화가 수행된 제2 스킵 특징맵과 상기 잔차 스킵 변조 특징맵을 더하여 상기 제2 특징맵을 추출할 수 있다.

[0024] 상기 프로세서는, 상기 제1 특징맵과 상기 제2 특징맵을 연결하여 비선형화 및 디멘전 감소를 수행하고, 상기 디멘전 감소의 출력에 기초하여 비선형화 연산 및 컨벌루션 연산을 포함하는 상기 잔차 연산을 적어도 한 번 수행함으로써 복수의 채널을 생성하고, 상기 복수의 채널을 결합하고, 상기 이미지를 업샘플링하여 결합된 복수의 채널과 더함으로써 상기 출력 이미지를 생성할 수 있다.

[0025] 상기 잔차 연산은, 적어도 한 번의 비선형화와, 적어도 한 번의 컨벌루션 연산과, 상기 적어도 한 번의 비선형화의 입력과 상기 적어도 한 번의 컨벌루션 연산의 출력의 합을 포함할 수 있다.

**도면의 간단한 설명**

[0026] 도 1은 일 실시예에 따른 이미지 처리 장치의 개략적인 블록도를 나타낸다.

도 2는 도 1에 도시된 이미지 처리 장치가 사용하는 네트워크 아키텍처의 예를 나타낸다.

도 3은 도 2에 도시된 네트워크 아키텍처를 단순화한 네트워크 아키텍처를 나타낸다.

도 4는 도 1에 도시된 이미지 처리 장치의 동작의 순서도를 나타낸다.

도 5a는 수신기가 수신한 이미지의 예를 나타낸다.

도 5b는 제1 레이어의 예를 나타낸다.

도 5c는 제2 레이어의 예를 나타낸다.

도 5d는 중간 특징맵의 예를 나타낸다.

도 5e는 도 5b의 제1 레이어에 컨벌루션 연산을 수행한 결과를 나타낸다.

도 5f는 도 5c의 제2 레이어에 컨벌루션 연산을 수행한 결과를 나타낸다.

도 5g는 제1 특징맵의 예를 나타낸다.

도 5h는 제2 특징맵의 예를 나타낸다.

도 6a는 변조 전 이미지의 예를 나타낸다.

도 6b는 제1 변조된 이미지의 예를 나타낸다.

도 6c는 제2 변조된 이미지의 예를 나타낸다.

도 7a는 종래 방식으로 처리된 이미지의 예를 나타낸다.

도 7b는 도 1에 도시된 이미지 처리 장치에 의해 처리된 이미지의 예를 나타낸다.

도 7c는 그라운드 트루스 이미지의 예를 나타낸다.

도 8a는 도 1에 도시된 이미지 처리 장치의 입력 이미지의 예를 나타낸다.

도 8b는 도 8a의 이미지를 도 1에 도시된 이미지 처리 장치가 처리한 이미지를 나타낸다.

**발명을 실시하기 위한 구체적인 내용**

- [0027] 이하에서, 첨부된 도면을 참조하여 실시예들을 상세하게 설명한다. 그러나, 실시예들에는 다양한 변경이 가해질 수 있어서 특허출원의 권리 범위가 이러한 실시예들에 의해 제한되거나 한정되는 것은 아니다. 실시예들에 대한 모든 변경, 균등물 내지 대체물이 권리 범위에 포함되는 것으로 이해되어야 한다.
- [0028] 실시예에서 사용한 용어는 단지 설명을 목적으로 사용된 것으로, 한정하려는 의도로 해석되어서는 안된다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 명세서 상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0029] 제1 또는 제2등의 용어를 다양한 구성요소들을 설명하는데 사용될 수 있지만, 구성요소들은 용어들에 의해서 한정되어서는 안 된다. 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만, 예를 들어 실시예의 개념에 따른 권리 범위로부터 이탈되지 않은 채, 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소는 제1 구성요소로도 명명될 수 있다.
- [0030] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 실시예가 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥 상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0031] 또한, 첨부 도면을 참조하여 설명함에 있어, 도면 부호에 관계없이 동일한 구성 요소는 동일한 참조부호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다. 실시예를 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 실시예의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.
- [0033] 도 1은 일 실시예에 따른 이미지 처리 장치의 개략적인 블록도를 나타낸다.
- [0034] 이미지 처리 장치(10)는 이미지를 수신하여 처리할 수 있다. 이미지는 빛의 굴절이나 반사 등에 의하여 이루어진 물체의 상을 포함하는 것으로, 선이나 색채를 이용하여 사물의 형상을 나타낸 것을 의미할 수 있다. 예를 들어, 이미지는 컴퓨터가 처리할 수 있는 형태로 된 정보로 이루어질 수 있다.
- [0035] 이미지 처리 장치(10)는 LR-LDR(Low Resolution-Low Dynamic Range) 이미지(이하, 수신한 이미지 또는 이미지라 함)를 처리하여 HR-HDR(High Resolution-High Dynamic Range) 이미지를 생성할 수 있다.
- [0036] 이미지 처리 장치(10)는 인공 신경망(Artificial Neural Network)을 이용하여 이미지를 처리할 수 있다. 예를 들어, 이미지 처리 장치(10)는 CNN(Convolutional Neural Network)를 이용하여 LR-LDR 이미지를 HR-HDR 이미지로 변환할 수 있다. 이하에서, LDR은 SDR과 동일하게 사용될 수 있다.
- [0037] 즉, 이미지 처리 장치(10)는 초해상화(super resolution)와 역톤매핑(inverse tone mapping)을 동시에 수행할 수 있다.
- [0038] 이미지 처리 장치(10)는 LR SDR(Low Resolution Standard Dynamic Range) 비디오에서 HR HDR(High Resolution High Dynamic Range) 버전으로의 매핑(mapping)을 학습하는 딥 SR-ITM(Deep Super Resolution-Inverse Tone Mappin)이라고 불리는 공동으로 학습가능한 CNN 기반 엔드투엔드(end-to-end) 네트워크를 이용하여 이미지를 처리할 수 있다.
- [0039] 자연 영상/씬(natural scene)은 씬의 물리적 밝기(physical brightness)를 나타내는 라디언스 맵(radiance map)에 의해 표현될 수 있다.
- [0040] 카메라 장치는 일반적으로  $cd/m^2$ (제곱 미터당 칸델라)로 휘도를 캡처할 수 있다. 디스플레이 매체에서 씬을 재현하기 위해서는, 감마 커브(gamma curve)와 같은 OETF(Opto-Electrical Transfer Function)을 통해 선형 도메인(linear domain)의 휘도가 픽셀 도메인으로 변환되어야 할 수 있다.
- [0041] ITM(Inverse Tone-Mapping)에서는 일반적으로, LDR 이미지가 일반적으로 디스플레이 참조(display-referred)인

반면, HDR 이미지는 광학적 지각(휘도 예측)에 있어 씬 참조(scene-referred) 또는 디스플레이 참조일 수 있다.

- [0042] 기존의 ITM 방식은 픽셀 도메인에서 LDR 이미지로부터 HDR 이미지에서의 휘도 값을 예측하는 것을 목적으로 하기 때문에, 광학 디스플레이 참조(optical display-referred) ITM(oDR-ITM)으로 지칭될 수 있다.
- [0043] 소비자 HDR TV의 배포(distribution)는 LDR 이미지가 픽셀 도메인에서 디스플레이 형식의 HDR 이미지로 직접 업 컨버팅되는(up-converted) 전기 적 관점(electrical sense)의 디스플레이 참조(display-referred) ITM(eDR-ITM)에 대한 필요성을 높일 수 있다. 이는 곧바로 디스플레이 형식의 HDR 이미지를 예측 가능토록하여, 효율적으로 HDR 비디오 콘텐츠의 부족을 해결할 수 있다.
- [0044] 일반적인 HDR 디스플레이 형식은 PQ(Perceptual Quantizer) 또는 HLG(Hybrid Log Gamma) OETF와 10 bits/pixel 의 비트 깊이(bit-depth)를 갖는 BT.2020 컬러 컨테이너(container)를 포함할 수 있다.
- [0045] eDR-ITM은 소비자 HDR 디스플레이(예를 들어, HDR TV)와 관련된 ITM 어플리케이션에 대해 oDR-ITM 또는 씬 참조 ITM 보다 적합하며 효율적일 수 있다.
- [0046] 이미지 처리 장치(10)는 IoT 장치, Machine-type 통신 장치 또는 휴대용 전자 장치 등으로 구현될 수 있다.
- [0047] 휴대용 전자 장치는 랩탑(laptop) 컴퓨터, 이동 전화기, 스마트 폰(smart phone), 태블릿(tablet) PC, 모바일 인터넷 디바이스(mobile internet device(MID)), PDA(personal digital assistant), EDA(enterprise digital assistant), 디지털 스틸 카메라(digital still camera), 디지털 비디오 카메라(digital video camera), PMP(portable multimedia player), PND(personal navigation device 또는 portable navigation device), 휴대용 게임 콘솔(handheld game console), e-북(e-book), 스마트 디바이스(smart device)로 구현될 수 있다. 예를 들어, 스마트 디바이스는 스마트 워치(smart watch) 또는 스마트 밴드(smart band)로 구현될 수 있다.
- [0048] 이미지 처리 장치(10)는 수신기(100), 프로세서(200) 및 메모리(300)를 포함한다. 수신기(100)는 이미지를 수신할 수 있다. 예를 들어, 수신기(100)가 수신하는 이미지는 LR-SDR(Low Resolution-Standard Dynamic Range) 이미지를 포함할 수 있다. 수신기(100)는 이미지를 프로세서(200)로 출력할 수 있다.
- [0049] 프로세서(200)는 수신기(100)가 수신한 데이터 및 메모리(300)에 저장된 데이터를 처리할 수 있다. 프로세서(200)는 메모리(300)에 저장된 데이터를 처리할 수 있다. 프로세서(200)는 메모리(300)에 저장된 컴퓨터로 읽을 수 있는 코드(예를 들어, 소프트웨어) 및 프로세서(200)에 의해 유발된 인스트럭션(instruction)들을 실행할 수 있다.
- [0050] "프로세서(200)"는 목적하는 동작들(desired operations)을 실행시키기 위한 물리적인 구조를 갖는 회로를 가지는 하드웨어로 구현된 데이터 처리 장치일 수 있다. 예를 들어, 목적하는 동작들은 프로그램에 포함된 코드(code) 또는 인스트럭션들(instructions)을 포함할 수 있다.
- [0051] 예를 들어, 하드웨어로 구현된 데이터 처리 장치는 마이크로프로세서(microprocessor), 중앙 처리 장치(central processing unit), 프로세서 코어(processor core), 멀티-코어 프로세서(multi-core processor), 멀티프로세서(multiprocessor), ASIC(Application-Specific Integrated Circuit), FPGA(Field Programmable Gate Array)를 포함할 수 있다.
- [0052] 프로세서(200)는 수신한 이미지를 하나 이상의 신호 성분 레이어로 분해할 수 있다. 하나 이상의 신호 성분 레이어는 제1 레이어 및 제2 레이어를 포함할 수 있다. 예를 들어, 제1 레이어 및 제2 레이어는 상이한 이미지일 수 있다.
- [0053] 프로세서(200)는 수신한 이미지를 필터링하여 제1 레이어 및 제2 레이어를 생성할 수 있다. 프로세서(200)는 이미지를 필터링할 수 있다. 예를 들어, 필터링은 유도 필터링(guided filtering)을 포함할 수 있다.
- [0054] 프로세서(200)는 필터링된 이미지와 수신한 이미지를 연결(concatenate)하여 제1 레이어를 생성할 수 있다. 프로세서(200)는 이미지에 대해 필터링된 이미지에 기초한 연산을 수행하고, 연산이 수행된 이미지와 수신한 이미지를 연결하여 제2 레이어를 생성할 수 있다. 예를 들어, 필터링된 이미지에 기초한 연산은 원소별 나눗셈(element-wise division)을 포함할 수 있다.
- [0055] 프로세서(200)는 제1 레이어에 기초하여 제1 변조(modulation)를 수행함으로써 제1 특징맵을 추출할 수 있다. 프로세서(200)는 제1 레이어에 기초하여 비선형화 및 컨벌루션 연산을 포함하는 잔차 연산을 수행함으로써 잔차 특징맵을 생성할 수 있다. 예를 들어, 비선형화는 ReLU(Rectified Linear Unit)를 포함할 수 있다. 또한, 제1 변조는 원소별 곱셈을 포함할 수 있다.



- [0056] 구체적으로, 프로세서(200)는 제1 레이어에 적어도 한 번의 비선형화 및 적어도 한 번의 컨벌루션 연산을 수행하고, 적어도 한 번의 비선형화의 입력과 적어도 한 번의 컨벌루션 연산의 출력을 더하여 잔차 특징맵을 생성할 수 있다.
- [0057] 프로세서(200)는 잔차 특징맵 및 제1 레이어에 기초하여 제1 변조를 수행함으로써 제1 특징맵을 추출할 수 있다. 구체적으로, 프로세서(200)는 잔차 특징맵에 적어도 한 번의 비선형화 및 적어도 한 번의 컨벌루션 연산을 수행할 수 있다. 프로세서(200)는 제1 레이어에 적어도 한 번의 컨벌루션 연산 및 적어도 한 번의 비선형화를 수행하여 제1 공유 변조 특징을 생성할 수 있다.
- [0058] 프로세서(200)는 적어도 한 번의 비선형화와 적어도 한 번의 컨벌루션 연산이 수행된 잔차 특징맵 및 제1 공유 변조 특징에 기초한 원소별 곱셈(element-wise multiplication)을 수행할 수 있다.
- [0059] 프로세서(200)는 제1 공유 변조 특징에 대하여 적어도 한 번의 컨벌루션 연산 및 비선형화를 수행한 후에 적어도 한 번의 비선형화와 적어도 한 번의 컨벌루션 연산이 수행된 잔차 특징맵과 원소별 곱셈을 수행할 수 있다. 프로세서(200)는 원소별 곱셈의 출력과 잔차 특징맵을 더하여 제1 특징맵을 추출할 수 있다.
- [0060] 프로세서(200)는 제1 레이어 및 제2 레이어에 기초하여 제2 변조를 수행함으로써 제2 특징맵을 추출할 수 있다. 프로세서(200)는 제2 레이어 및 잔차 특징맵에 기초하여 제2 변조를 수행함으로써 잔차 스킵 변조 특징맵을 생성할 수 있다. 예를 들어, 제2 변조는 원소별 곱셈 연산을 포함할 수 있다.
- [0061] 구체적으로, 프로세서(200)는 제2 레이어에 기초한 특징맵과 잔차 특징맵을 연결(concatenate)하고, 비선형화 및 디멘전 감소(dimension reduction)를 수행하여 제1 스킵 특징맵을 생성할 수 있다.
- [0062] 프로세서(200)는 생성한 제1 스킵 특징맵에 적어도 한 번의 컨벌루션 연산 및 비선형화를 수행할 수 있다. 프로세서(200)는 제2 레이어에 적어도 한 번의 컨벌루션 연산 및 적어도 한 번의 비선형화를 수행하여 제2 공유 변조 특징을 생성할 수 있다.
- [0063] 프로세서(200)는 적어도 한 번의 컨벌루션 연산과 비선형화가 수행된 제1 스킵 특징맵 및 제2 공유 변조 특징에 기초한 원소별 곱셈을 수행할 수 있다. 프로세서(200)는 원소별 곱셈의 출력과 제2 레이어에 기초한 특징맵을 더하여 잔차 스킵 변조 특징맵을 생성할 수 있다.
- [0064] 프로세서(200)는 잔차 스킵 변조 특징맵 및 제1 변조의 출력에 기초하여 비선형화 및 컨벌루션 연산을 포함하는 잔차 스킵 연산을 수행함으로써 제2 특징맵을 추출할 수 있다.
- [0065] 구체적으로, 프로세서(200)는 잔차 스킵 변조 특징맵과 제1 변조의 출력을 연결하고, 비선형화 및 디멘전 감소를 수행하여 제2 스킵 특징맵을 생성할 수 있다.
- [0066] 프로세서(200)는 제2 스킵 특징맵에 적어도 한 번의 컨벌루션 연산 및 비선형화를 수행하고, 적어도 한 번의 컨벌루션 연산 및 비선형화가 수행된 제2 스킵 특징맵과 잔차 스킵 변조 특징맵을 더하여 제2 특징맵을 추출할 수 있다.
- [0067] 프로세서(200)는 제1 특징맵 및 제2 특징맵에 기초하여 컨벌루션 연산을 포함하는 잔차 연산을 수행함으로써 초해상화와 역톤매핑이 동시에 수행된 출력 이미지를 생성할 수 있다.
- [0068] 구체적으로, 프로세서(200)는 제1 특징맵과 제2 특징맵을 연결하여 비선형화 및 디멘전 감소를 수행할 수 있다. 프로세서(200)는 디멘전 감소의 출력에 기초하여 비선형화 연산 및 컨벌루션 연산을 포함하는 잔차 연산을 적어도 한 번 수행함으로써 복수의 채널을 생성할 수 있다. 예를 들어, 프로세서(200)는 10 번의 잔차 연산을 수행할 수 있다.
- [0069] 잔차 연산은 적어도 한 번의 비선형화, 적어도 한 번의 컨벌루션 연산 및 적어도 한 번의 비선형화의 입력과 적어도 한 번의 컨벌루션 연산의 출력의 합을 포함할 수 있다.
- [0070] 프로세서(200)는 생성한 복수의 채널을 결합하고, 수신한 이미지를 업샘플링하여 결합된 복수의 채널과 더함으로써 출력 이미지를 생성할 수 있다.
- [0071] 메모리(300)는 프로세서(200)에 의해 실행가능한 인스트럭션들(또는 프로그램)을 저장할 수 있다. 예를 들어, 인스트럭션들은 프로세서(200)의 동작 및/또는 프로세서(200)의 각 구성의 동작을 실행하기 위한 인스트럭션들을 포함할 수 있다.
- [0072] 메모리(300)는 휘발성 메모리 장치 또는 비휘발성 메모리 장치로 구현될 수 있다.

- [0073] 휘발성 메모리 장치는 DRAM(dynamic random access memory), SRAM(static random access memory), T-RAM(thyristor RAM), Z-RAM(zero capacitor RAM), 또는 TTRAM(Twin Transistor RAM)으로 구현될 수 있다.
- [0074] 불휘발성 메모리 장치는 EEPROM(Electrically Erasable Programmable Read-Only Memory), 플래시(flash) 메모리, MRAM(Magnetic RAM), 스핀전달토크 MRAM(Spin-Transfer Torque(STT)-MRAM), Conductive Bridging RAM(CBRAM), FeRAM(Ferroelectric RAM), PRAM(Phase change RAM), 저항 메모리(Resistive RAM(RRAM)), 나노 튜브 RRAM(Nanotube RRAM), 폴리머 RAM(Polymer RAM(PoRAM)), 나노 부유 게이트 메모리(Nano Floating Gate Memory(NFGM)), 홀로그래픽 메모리(holographic memory), 분자 전자 메모리 소자(Molecular Electronic Memory Device), 또는 절연 저항 변화 메모리(Insulator Resistance Change Memory)로 구현될 수 있다.
- [0076] 도 2는 도 1에 도시된 이미지 처리 장치가 사용하는 네트워크 아키텍처의 예를 나타낸다.
- [0077] 도 2를 참조하면, 프로세서(200)는 단일 LR SDR 이미지로부터 HR HDR 이미지를 HDR 픽셀 도메인(디스플레이 형식)에서 직접적으로 생성할 수 있다. 프로세서(200)는 복원된 국부 디테일 및 대비(local detail and contrast)를 이용하여 신호 진폭을 확장하고, 동시에 추가적인 공간 정보를 복합적으로 예측할 수 있다.
- [0078] 프로세서(200)는 유도 필터(210)를 통한 분해(decomposition)과 국부 대비 변조(local contrast modulation)을 이용하는 CNN(Convolutional Neural Network) 기반 아키텍처를 이용하여 이미지를 처리할 수 있다.
- [0079] 이미지(또는 입력 이미지)는 전반적 밝기(intensity)를 포함하는 베이스 레이어(base layer)와 세부적인 디테일을 포함하는 디테일 레이어(detail layer)로 분해될 수 있다.
- [0080] 프로세서(200)는 심층 SR-ITM 네트워크 구조를 이용하여 이미지의 고유한 저주파 및 고주파 성분에 집중하여 이미지를 처리할 수 있다. 이미지 처리 장치(10)는 특징(feature)을 추출(extraction)하는 동안 강화된 SR-ITM을 위해서 분리된 베이스 레이어 및 디테일 레이어 경로에서 국부 대비 변조를 수행할 수 있다. 이를 통해 네트워크 전체 신호 진폭을 확장하면서 국부 디테일 및 대비를 복원(restoring)할 수 있다.
- [0081] 프로세서(200)는 4K 비디오 데이터에 대한 학습을 수행하고, TV용 HDR 10 표준을 따르는 실제 UHD HDR 환경에서 동작할 수 있다.
- [0082] 프로세서(200)는 단일 LR-SDR 이미지에 조인트 SR 및 eDR-ITM 문제를 수행할 수 있다. 프로세서(200)는 복원된 디테일과 향상된 대비를 가진 고품질 SR-ITM을 위한 잔차(residual), 스킵(skip), 변조(modulation) 블록의 다양한 조합을 기반으로 한 심층 네트워크를 이용하여 이미지를 처리할 수 있다. 이미지 처리 장치(10)는 심층 SR-ITM 네트워크에 입력되는 이미지의 고유한 저주파 및 고주파 성분에 집중할 수 있도록 이미지를 분해할 수 있다.
- [0083] 프로세서(200)는 딥 SR-ITM을 통해 SR 및 ITM을 동시에 공동으로(jointly) 학습하고 수행할 수 있다. 프로세서(200)는 잔차(residual) 블록, 스킵(skip) 블록 및 변조(modulation) 블록의 상이한 조합을 이용하여 증가된 특징 생성 다양성(feature generation diversity)을 갖는 매우 깊은 네트워크(very deep network)를 제공할 수 있다.
- [0084] 프로세서(200)는 입력 신호를 별도의 레이어들로 분해함으로써, 저주파(low frequency) 성분 및 고주파(high frequency) 성분에 따라 구별되는 잠재적인(latent) 특성을 잡아내어, 조인트 SR-ITM 태스크에 대한 공간 해상도 및 신호 진폭을 증가시키면서 디테일(detail) 및 대비(contrast)를 향상시킬 수 있다.
- [0085] 프로세서(200)는 심층 SR-ITM을 수행할 수 있다. 프로세서(200)는 조인트 SR-ITM을 위한 신호 분해 및 변조에 기초한 심층 잔차 네트워크를 통해 단일 LR-SDR 이미지로부터 BT.2020 및 PQ-OETF의 디스플레이 형식의 HR-HDR 이미지를 생성할 수 있다.
- [0086] 프로세서(200)는 특징 추출을 향상시킬 뿐만 아니라, 대비의 국부 변조를 향상시킬 수 있다.
- [0087] 프로세서(200)는 수신한 이미지(예를 들어, LR-SDR 이미지)를 분해할 수 있다. 프로세서(200)는 이미지를 경계 보존 저역 통과 필터(edge-preserving low pass filter)를 통해 베이스 레이어 및 디테일 레이어로 분해할 수 있다. 예를 들어, 경계 보존 저역 통과 필터는 유도 필터(guided filter(210))를 포함할 수 있다.
- [0088] 프로세서(200)는 유도 필터링을 통해 이미지로부터 베이스 레이어를 획득할 수 있고, 이미지(또는 입력 이미지)를 베이스 레이어로 나누어 줌으로써 디테일 레이어를 획득할 수 있다. 디테일 레이어의 획득은 수학적 1과 같이 나타낼 수 있다.

**수학식 1**

$$I_d = I \circledast I_b$$

여기서,  $\circledast$  는 원소별 나눗셈을 의미할 수 있다.

프로세서(200)는 디테일 레이어  $I_d$ 와 입력 이미지  $I$ 를 연결하여 제2 레이어를 생성할 수 있다. 프로세서(200)는 베이스 레이어  $I_b$ 와 입력 이미지  $I$ 를 연결(concatenate)하여 제1 레이어를 생성할 수 있다. 마찬가지로, 프로세서(200)는 디테일 레이어  $I_d$ 와 입력 이미지  $I$ 를 연결하여 제2 레이어를 생성할 수 있다. 예를 들어, 프로세서(200)는 베이스 레이어 및 디테일 레이어와 입력 이미지를 채널 방향으로 연결할 수 있다.

도 2의 예시에서,  $I_b^{in}$ 는 제1 레이어를 의미할 수 있고,  $I_d^{in}$ 는 제2 레이어를 의미할 수 있다. 제1 레이어 및 제2 레이어는 수학식 2와 같이 나타낼 수 있다.

**수학식 2**

$$I_b^{in} = [I \ I_b], \text{ and } I_d^{in} = [I \ I_d].$$

프로세서(200)는 제1 레이어와 제2 레이어를 두 개의 구별된 특징 추출 경로를 통해 처리하여 특징맵을 추출할 수 있다. 네트워크 용량(capacity) 확장의 관점에서, 프로세서(200)는 전체 깊이를 증가시키는 대신 다양한 특징을 추출하기 위해 컨벌루션 레이어들을 할당할 수 있다(두 개의 브랜치(branch)가 별도로 처리되므로).

프로세서(200)는 입력 이미지를 저주파 정보가 지배적인 경계 보존된 흐린 이미지(edge-preserved blurred image)를 포함하는 베이스 레이어와 고주파 정보(예를 들어, 경계 또는 텍스처(texture))가 지배적인 나머지를 포함하는 디테일 레이어로 분해함으로써, 입력 이미지의 상이한 특성에 집중할 수 있다.

원본 입력 이미지는 유용한 정보를 포함하고 있기 때문에, 프로세서(200)는 분해된 베이스 레이어 및 디테일 레이어와 원본 이미지를 연결하여 처리할 수 있다. 또한, 프로세서(200)는 추가적인 분해의 형태로 잔차 학습을 이용하여 이미지를 처리할 수 있다.

프로세서(200)는 도 2의 예시와 같은 심층 SR-ITM 네트워크 구조를 이용하여 이미지를 처리할 수 있다. 네트워크는 잔차(residual), 스킵(skip) 및 변조(modulation)의 조합으로 이루어진 4 개의 블록(block)을 포함할 수 있다.

4 개의 블록은 잔차 블록(Res Block(230)), 잔차 변조 블록(ResModBlock(250)), 잔차 스킵 블록(ResSkipBlock(290)) 및 잔차 스킵 변조 블록(ResSkipModBlock(270))을 포함할 수 있다.

$i$  번째 잔차 블록(230)  $RB_i$ 의 출력은 수학식 3과 같이 나타낼 수 있다.

**수학식 3**

$$RB_i(x) = (Conv \circ RL \circ Conv \circ RL)(x) + x = C_{RB}(x) + x.$$

여기서,  $x$ 는  $i$  번째 블록의 입력을 의미하고, Conv는 컨벌루션 레이어를 의미하고, RL은  $ReLU(ReLU(\cdot) = \max(0, \cdot))$ 을 의미할 수 있다.

잔차 변조 블록(250)에는 잔차 블록(230)에 추가적인 곱셈 변조(multiplicative modulation) 성분이 추가될 수 있다. 잔차 변조 블록(250)은 베이스 레이어에 기초한 제1 공유 변조 특징(Shared Modulation Feature, SMF<sub>b</sub>)을 이용할 수 있다. 공유 변조 특징은 제1 레이어에 컨벌루션 연산 및 비활성화를 수행하여 획득될 수 있다.

[0109] 예를 들어, 프로세서(200)는 제1 레이어( $I_b^{in}$ )에 3 번의 컨벌루션 레이어 및 3 번의 ReLU 활성화를 번갈아 수행하여 제1 공유 변조 특징을 생성할 수 있다. 제1 공유 변조 특징은 수학적 식 4와 같이 생성될 수 있다.

**수학적 식 4**

[0111]  $SMF_b = (RL \circ Conv \circ RL \circ Conv \circ RL \circ Conv) (I_b^{in})$ .

[0113] 잔차 변조 블록(250)의 출력  $RMB_i$ 는 수학적 식 5와 같이 나타낼 수 있다.

**수학적 식 5**

[0115]  $RMB_i(x) = C_{RB}(x) \odot \{(Conv \circ RL \circ Conv) (SMF_b)\} + x$

[0117] 여기서,  $\odot$ 는 변조를 위한 원소별 곱셈(element-wise) 곱셈을 의미할 수 있다.

[0118] 베이스 레이어 경로의 마지막 특징 추출 레이어의 출력인 또는, 제1 특징맵 ( $FE_b$ )은 수학적 식 6과 같이 나타낼 수 있다.

**수학적 식 6**

[0120]  $FE_b = (RMB_m \circ RB_m^b \circ \dots \circ RMB_1 \circ RB_1^b \circ Conv) (I_b^{in})$

[0122] 프로세서(200)는 복수의 잔차 블록(230) 및 잔차 변조 블록(250)을 적용함으로써 제1 특징맵을 생성할 수 있다. 수학적 식 6에서, 프로세서(200)는 베이스 레이어 경로에서 m 개의 잔차 블록(230) 및 m 개의 잔차 변조 블록(250)을 통해 제1 특징맵을 생성할 수 있다.

[0123] 디테일 레이어를 처리하는 경로에서 스킵 성분이 추가될 수 있다. 스킵 성분은 네트워크의 유연성(flexibility)를 증가시킬 수 있다. 세 번째 유형의 블록인 잔차 스킵 블록(290)은 베이스 레이어를 처리하는 경로에서 잔차 변조 블록(250)의 특징과 연결될 수 있다. i 번째 잔차 스킵 블록(290)인  $RSB_i$ 의 출력은 수학적 식 7과 같이 나타낼 수 있다.

**수학적 식 7**

[0125]  $RSB_i(x) = (Conv \circ RL \circ Conv \circ DR \circ RL) ([x \ RMB_i]) + x$ .

[0127] 여기서, DR은  $1 \times 1$  컨벌루션을 이용하는 디멘전 감소(dimension reduction) 레이어를 의미할 수 있다.  $[x \ y]$ 는 수학적 식 2에서 나타난 바와 같이, 채널 방향으로 x 및 y의 연결을 의미할 수 있다.

[0128] DR 레이어는 확장된 입력으로부터 어떤 정보를 얼마나 통과시킬지 제어하는 선택 모듈(selection module)과 같이 동작할 수 있다.

[0129] 마지막으로, 잔차 스킵 변조 블록(270)은 디테일 레이어로부터 획득된 제2 공유 특징인  $SMF_d$ 와 베이스 레이어 경로의 잔차 블록(230)의 특징을 이용할 수 있다. 제1 공유 특징  $SMF_b$ 와 유사하게 제2 공유 특징  $SMF_d$ 는 수학적 식 8과 같이 나타낼 수 있다.



수학식 8

[0131]  $SMF_d = (RL \circ Conv \circ RL \circ Conv \circ RL \circ Conv) (I_d^m)$

[0133] 제2 공유 특징을 이용한 i 번째 잔차 스킵 변조 블록(270)의 출력  $RSMB_i$ 는 수학식 9와 같이 나타낼 수 있다.

수학식 9

[0135]  $RSMB_i(x) = \{(Conv \circ RL \circ Conv \circ DR \circ RL) ([x \ RB_i^b])\} \odot \{(Conv \circ RL \circ Conv) (SMF_d)\} + x.$

[0137] 디테일 레이어 경로의 마지막 출력인 제2 특징맵  $FE_d$ 는 수학식 10과 같이 나타낼 수 있다.

수학식 10

[0139]  $FE_d = (RSMB_m \circ RSB_{m-1} \circ \dots \circ RSB_1 \circ RSMB_1 \circ RB_1^d \circ Conv) (I_d^m).$

[0141] 프로세서(200)는 제2 레이어  $I_d^m$ 를 한 번의 컨벌루션 레이어 및 단일 잔차 블록(230)을 통과시키고, 잔차 스킵 변조 블록(270)과 잔차 스킵 블록(290)이 번갈아 적용시킬 수 있다. 예를 들어, 프로세서(200)는 m 개의 잔차 스킵 변조 블록(270), m-1 개의 잔차 스킵 블록(290) 및 1 개의 잔차 블록( $RB^d$ , 디테일 레이어 경로의 잔차 블록(230))을 이용하여 제2 특징맵  $FE_d$ 를 생성할 수 있다.

[0142] 컨벌루션 레이어의 컨벌루션 연산은 모든 픽셀 위치들에 대하여 동일한 컨벌루션 필터를 사용하기 때문에 컨벌루션 연산은 공간적으로 불변(spatially invariant)하다.

[0143] 이러한 특성에 따라, 컨벌루션 필터는 공간적으로 변하지 않기 (동일하기) 때문에 네트워크의 용량을 제한할 수 있다. 프로세서(200)는 원소별 곱셈에 의해 공간적으로 변동하는 변조를 적용함으로써 단순한 CNN에 의한 모델링에 비해서 더욱 복잡한 매핑이 가능하도록 한다.

[0144] 이러한 변조는 국부 대비 및 디테일을 복원하는 것이 중요한 ITM 태스크에서 매우 유용할 수 있다. 또한, 프로세서(200)는 변조를 위해 공유 및 적응(adaptive) 특징을 이용함으로써 복잡성을 감소시키면서 깊이 특정 변조(depth specific modulation)을 수행할 수 있다.

[0145] 심층 SR-ITM 네트워크는 통합 파트를 포함할 수 있다. 통합 파트를 이용하여, 프로세서(200)는 베이스 레이어 및 디테일 레이어 경로로부터 생성한 제1 특징맵과 제2 특징맵을 통합하여 처리함으로써 HR HDR 출력 이미지를 생성할 수 있다.

[0146] 프로세서(200)는 통합 파트에서 잔차 블록(230)을 사용할 수 있다. 통합 파트의 잔차 블록(230)의 입력  $x_{int}$ 은 수학식 11과 같이 나타낼 수 있다.

수학식 11

[0148]  $x_{int} = (Conv \circ DR \circ RL) ([FE_b \ FE_d])$

[0150] n 번째 잔차 블록(230)의 출력  $y_{int}$ 는 수학식 12와 같이 나타낼 수 있다.

수학식 12

[0152] 
$$y_{int} = (RB_n^i \circ \dots \circ RB_1^i) (x_{int}).$$

[0154] 프로세서(200)는 복수의 잔차 블록(230)을 통한 잔차 연산을 수행하여 복수의 채널을 수행할 수 있다. 복수의 채널을 결합하여 출력 이미지를 생성할 수 있다. 최종 출력 이미지인 HR HDR 예측  $\hat{J}$ 은 수학식 13과 같이 나타낼 수 있다.

수학식 13

[0156] 
$$\hat{J} = (Conv \circ PS \circ RL \circ Conv \circ RL \circ Conv \circ RL) (y_{int}) + Bic(I)$$

[0158] 여기서, PS는 복수의 채널을 결합하는 픽셀 셔플(pixel shuffle)을 의미할 수 있고, Bic(I)는 입력 이미지 I를 바이큐빅 보간(interpolation)한 결과를 의미할 수 있다.

[0159] 상술한 바와 같이, 프로세서(200)는 픽셀 셔플을 이용한 결합을 이용하여 HR-SDR 이미지를 생성할 수 있다. 예를 들어, 4 개의 채널이 제1, 제2, 제3 및 제4 채널인 경우, 프로세서(200)는 제1 채널의 (1, 1)위치의 화소를 출력될 이미지의 (1, 1) 위치의 화소로 배치하고, 제2 채널의 (1, 1)화소를 출력될 이미지의 (1, 2) 위치의 화소로 배치할 수 있다.

[0160] 마찬가지로, 프로세서(200)는 제3 채널의 (1, 1) 화소를 출력될 이미지의 (2, 1) 위치의 화소로 배치하고, 제4 채널의 (1, 1) 화소를 출력될 이미지의 (2, 2) 위치의 화소로 배치할 수 있다.

[0161] 모든 채널의 (1, 1) 화소가 배치한 후에, 프로세서(200)는 제1 채널의 (1, 2) 화소를 복원될 채널의 (1, 3) 위치의 화소로 배치할 수 있다. 프로세서(200)는 모든 채널에 대하여 상술한 동작을 수행할 수 있다.

[0162] 프로세서(200)는 학습을 쉽게 수행하고, 예측 정확도를 향상시키기 위해서 4 가지 유형의 잔차 블록(230) 내부에 국부 잔차 학습을 수행할 뿐만 아니라 전역 잔차 학습(global residual learning)을 적용시킬 수 있다.

[0164] 도 3은 도 2에 도시된 네트워크 아키텍처를 단순화한 네트워크 아키텍처를 나타낸다.

[0165] 도 3을 참조하면, 프로세서(200)는 심층 SR-ITM을 간소화한 도 3의 예시와 같은 토이 네트워크(toy network)를 이용하여 이미지를 처리할 수 있다. 토이 네트워크를 통해, 프로세서(200)는 입력의 분해, 애블레이션(ablation) 연구 및 상이한 유형의 변조의 효과를 분석할 수 있다.

[0166] 토이 네트워크는 실험의 효율적인 관리를 위해 단순하지만, 원본 심층 SR-ITM의 실험결과를 반영하는 것이 가능하다. 토이 네트워크는 입력에 대한 분해와 지역 잔차 연결, 베이스 레이어에서 디테일 레이어 경로로의 스킵 연결 및 각 경로에서의 변조를 포함할 수 있다.

[0168] 도 4는 도 1에 도시된 이미지 처리 장치의 동작의 순서도를 나타낸다.

[0169] 도 4를 참조하면, 수신기(100)는 이미지를 수신할 수 있다(510). 이미지는 LR-SDR 이미지를 포함할 수 있다.

[0170] 프로세서(200)는 수신한 이미지를 필터링하여 제1 레이어 및 제2 레이어를 생성할 수 있다(530). 프로세서(200)는 이미지를 필터링할 수 있다. 예를 들어, 필터링은 유도 필터링을 포함할 수 있다.

[0171] 프로세서(200)는 필터링된 이미지와 수신한 이미지를 연결하여 제1 레이어를 생성할 수 있다. 프로세서(200)는 이미지에 대해 필터링된 이미지에 기초한 연산을 수행하고, 연산이 수행된 이미지와 수신한 이미지를 연결하여 제2 레이어를 생성할 수 있다. 예를 들어, 필터링된 이미지에 기초한 연산은 원소별 나눗셈(element-wise division)을 포함할 수 있다.

[0172] 프로세서(200)는 제1 레이어에 기초하여 제1 변조(modulation)를 수행함으로써 제1 특징맵을 추출할 수 있다(550). 프로세서(200)는 제1 레이어에 기초하여 비선형화 및 컨벌루션 연산을 포함하는 잔차 연산을 수행함으로써 잔차 특징맵을 생성할 수 있다. 예를 들어, 비선형화는 ReLU(Rectified Linear Unit)를 포함할 수 있다.

또한, 제1 변조는 원소별 곱셈을 포함할 수 있다.

- [0173] 구체적으로, 프로세서(200)는 제1 레이어에 적어도 한 번의 비선형화 및 적어도 한 번의 컨벌루션 연산을 수행하고, 적어도 한 번의 비선형화의 입력과 적어도 한 번의 컨벌루션 연산의 출력을 더하여 잔차 특징맵을 생성할 수 있다.
- [0174] 프로세서(200)는 잔차 특징맵 및 제1 레이어에 기초하여 제1 변조를 수행함으로써 제1 특징맵을 추출할 수 있다. 구체적으로, 프로세서(200)는 잔차 특징맵에 적어도 한 번의 비선형화 및 적어도 한 번의 컨벌루션 연산을 수행할 수 있다. 프로세서(200)는 제1 레이어에 적어도 한 번의 컨벌루션 연산 및 적어도 한 번의 비선형화를 수행하여 제1 공유 변조 특징을 생성할 수 있다.
- [0175] 프로세서(200)는 적어도 한 번의 비선형화와 적어도 한 번의 컨벌루션 연산이 수행된 잔차 특징맵 및 제1 공유 변조 특징에 기초한 원소별 곱셈(element-wise multiplication)을 수행할 수 있다.
- [0176] 프로세서(200)는 제1 공유 변조 특징에 대하여 적어도 한 번의 컨벌루션 연산 및 비선형화를 수행한 후에 적어도 한 번의 비선형화와 적어도 한 번의 컨벌루션 연산이 수행된 잔차 특징맵과 원소별 곱셈을 수행할 수 있다. 프로세서(200)는 원소별 곱셈의 출력과 잔차 특징맵을 더하여 제1 특징맵을 추출할 수 있다.
- [0177] 프로세서(200)는 제1 레이어 및 제2 레이어에 기초하여 제2 변조를 수행함으로써 제2 특징맵을 추출할 수 있다(570). 프로세서(200)는 제2 레이어 및 잔차 특징맵에 기초하여 제2 변조를 수행함으로써 잔차 스킵 변조 특징맵을 생성할 수 있다. 예를 들어, 제2 변조는 원소별 곱셈 연산을 포함할 수 있다.
- [0178] 구체적으로, 프로세서(200)는 제2 레이어에 기초한 특징맵과 잔차 특징맵을 연결(concatenate)하고, 비선형화 및 디멘전 감소(dimension reduction)를 수행하여 제1 스킵 특징맵을 생성할 수 있다.
- [0179] 프로세서(200)는 생성한 제1 스킵 특징맵에 적어도 한 번의 컨벌루션 연산 및 비선형화를 수행할 수 있다. 프로세서(200)는 제2 레이어에 적어도 한 번의 컨벌루션 연산 및 적어도 한 번의 비선형화를 수행하여 제2 공유 변조 특징을 생성할 수 있다.
- [0180] 프로세서(200)는 적어도 한 번의 컨벌루션 연산과 비선형화가 수행된 제1 스킵 특징맵 및 제2 공유 변조 특징에 기초한 원소별 곱셈을 수행할 수 있다. 프로세서(200)는 원소별 곱셈의 출력과 제2 레이어에 기초한 특징맵을 더하여 잔차 스킵 변조 특징맵을 생성할 수 있다.
- [0181] 프로세서(200)는 잔차 스킵 변조 특징맵 및 제1 변조의 출력에 기초하여 비선형화 및 컨벌루션 연산을 포함하는 잔차 스킵 연산을 수행함으로써 제2 특징맵을 추출할 수 있다.
- [0182] 구체적으로, 프로세서(200)는 잔차 스킵 변조 특징맵과 제1 변조의 출력을 연결하고, 비선형화 및 디멘전 감소를 수행하여 제2 스킵 특징맵을 생성할 수 있다.
- [0183] 프로세서(200)는 제2 스킵 특징맵에 적어도 한 번의 컨벌루션 연산 및 비선형화를 수행하고, 적어도 한 번의 컨벌루션 연산 및 비선형화가 수행된 제2 스킵 특징맵과 잔차 스킵 변조 특징맵을 더하여 제2 특징맵을 추출할 수 있다.
- [0184] 프로세서(200)는 제1 특징맵 및 제2 특징맵에 기초하여 컨벌루션 연산을 포함하는 잔차 연산을 수행함으로써 출력 이미지를 생성할 수 있다(590).
- [0185] 구체적으로, 프로세서(200)는 제1 특징맵과 제2 특징맵을 연결하여 비선형화 및 디멘전 감소를 수행할 수 있다. 프로세서(200)는 디멘전 감소의 출력에 기초하여 비선형화 연산 및 컨벌루션 연산을 포함하는 잔차 연산을 적어도 한 번 수행함으로써 복수의 채널을 생성할 수 있다. 예를 들어, 프로세서(200)는 10 번의 잔차 연산을 수행할 수 있다.
- [0186] 잔차 연산은 적어도 한 번의 비선형화, 적어도 한 번의 컨벌루션 연산 및 적어도 한 번의 비선형화의 입력과 적어도 한 번의 컨벌루션 연산의 출력의 합을 포함할 수 있다.
- [0187] 프로세서(200)는 생성한 복수의 채널을 결합하고, 수신한 이미지를 업샘플링하여 결합된 복수의 채널과 더함으로써 출력 이미지를 생성할 수 있다.
- [0189] 도 5a는 수신기가 수신한 이미지의 예를 나타내고, 도 5b는 제1 레이어의 예를 나타내고, 도 5c는 제2 레이어의 예를 나타낸다.
- [0190] 도 5d는 중간 특징맵의 예를 나타내고, 도 5e는 도 5b의 제1 레이어에 컨벌루션 연산을 수행한 결과를

나타내고, 도 5f는 도 5c의 제2 레이어에 컨벌루션 연산을 수행한 결과를 나타낸다.

- [0191] 도 5g는 제1 특징맵의 예를 나타내고, 도 5h는 제2 특징맵의 예를 나타낸다.
- [0192] 도 5a 내지 5g를 참조하면, 이미지 처리 장치(10)의 성능이 평가될 수 있다. 256 채널을 갖는 픽셀 셔플 이전의 레이어를 제외한 모든 컨벌루션 필터들은 64 개의 출력 채널을 갖는 3×3 크기를 가지고, 출력 레이어는 3 개의 채널을 가질 수 있다.
- [0193] 이미지 처리 장치(10)가 사용하는 네트워크는 컬러 컨테이너가 변환되면서 3 개의 YUV 채널 모두에 대하여 학습되고 테스트될 수 있다. 네트워크 구조에서, m은 3이고, n은 10일 수 있다. m은 수학적 6, 10에서 사용된 것과 동일한 의미를 가질 수 있고, n은 수학적 12에서 사용된 것과 동일한 의미를 가질 수 있다.
- [0194] 실험은 SR에 대하여 SF(Scale Factor) 4를 이용했고, 2 개의 픽셀 셔플러(shuffler)가 구현될 수 있고, 둘 사이에는 컨벌루션 레이어가 존재할 수 있다. sf가 언급되지 않은 모든 결과는 SF=2에 대한 것일 수 있다.
- [0195] YouTube에서 총 59818 프레임의 4K-UHD HDR 비디오(BT.2020 with PQ-OETF) 10 개가 수집될 수 있다. 이 중 7 개의 비디오가 학습(44K 프레임)에 사용되고, 3 개는 테스트에 사용될 수 있다.
- [0196] 각 프레임에서 160×160 크기의 20 내지 40 개의 패치들이 크롭될 수 있다. 프레임간 높은 일관성(high coherence)를 피하기 위해 패치들은 10 내지 80의 프레임 스트라이드(stride) 범위로 추출될 수 있다.
- [0197] 테스트를 위해 테스트 비디오에서 28 가지 상이한 썬들이 선택될 수 있다. 대응 HDR 이미지로부터 SDR 이미지를 얻기 위한 톤매핑(tone-mapping)을 위해서 HDR Toolbox에서의 구현으로부터 선택된 19가지 상이한 TMO(Tone Mapping Operator)들이 이용될 수 있다. 이 경우, HDR 비디오가 선형화되고 컬러 컨테이너가 톤 매핑 전에 변환(BT. 2020에서 BT. 709로) 된 후, 감마 인코딩(gamma-encoded)될 수 있다.
- [0198] 하지만, 이러한 방법은 부자연스러운 색을 보이고, 어두운 장면에서 불안정한 결과를 초래할 수 있다. LR SDR 데이터가 자연스러워 보이고 가능한 실제 SDR 비디오에 가깝게 보여야 하기 때문에, YouTube의 자동 변환 프로세스를 통해 SDR 비디오 쌍들이 수집되고, 바이큐빅 필터링으로 다운 스케일링될 수 있다.
- [0199] 학습에 있어서, L2 손실(loss)이 사용될 수 있고, Adam 옵티마이저(optimizer) 및 Xavier 초기화 방법(initialization method)이 사용될 수 있다. L2 손실은 HR HDR 예측이 픽셀 도메인에 있는 것과 같이 작동할 수 있다.
- [0200] 심층 SR-ITM은 가중치(weight)에 대해  $5 \times 10^{-7}$ , 바이어스(bias)에 대해  $5 \times 10^{-8}$ 의 학습률(learning rate)로 변조 과정 없이 사전 학습(pre-trained)될 수 있다. 사전 학습 후에, 네트워크는 가중치에 대해  $10^{-7}$ , 바이어스에 대해  $10^{-8}$ 의 학습률에서 변조 과정과 함께 전체적으로 학습될 수 있다.
- [0201] 미니 배치(mini batch) 크기는 16일 수 있고, 전체 학습 및 테스트 과정은 MatConvNet 패키지를 이용하여 구현될 수 있다.
- [0202] 토이 네트워크를 위해서 독립적으로 수집된 SDR-HDR 쌍이 사용될 수 있다. 상술한 학습 방식과 마찬가지로, 토이 네트워크도 변조 없이 사전 학습된 후에, 전체적으로 학습될 수 있다. 다른 파라미터들은 심층 SR-ITM 네트워크와 동일하게 설정된 상태에서 학습률은  $10^{-6}$ 으로 설정될 수 있다.
- [0203] 모든 HDR 결과의 시각화는 MPC-HC 플레이어로 madVR 렌더러(renderer)를 통해 획득될 수 있다.
- [0204] 실험결과를 획득하기 위해서 토이 네트워크가 입력 이미지(이미지 레이어), 베이스 레이어 및 디테일 레이어의 상이한 조합을 입력으로 1 내지 3의 경로들을 갖도록 변경될 수 있다. 모든 경우에, 변조는 수행되지 않을 수 있다. 실험 조건에 따른 성능(PSNR(Peak Signal to Noise Ratio), SSIM(Structural Similarity))은 표 1과 같이 나타낼 수 있다.

표 1

	(a)	(b)	(c)	(d)	(e)
경로(pass)	1	1	2	2	3
이미지	√	√	x	√	√
베이스/디테일	x	√	√	√	√
스택(stack)	x	√	x	√	x

PSNR(dB)	38.11	38.25	38.11	<b>38.46</b>	38.21
SSIM	0.9905	0.9907	0.9911	<b>0.9916</b>	0.9914

- [0207] 표 1의 (a) 열에서, 이미지만이 단일 경로 네트워크에 입력될 수 있다. (b) 열에서, 단일 경로 네트워크에 입력되기 전에 모든 3 개의 레이어들이 스택될 수 있다. (c) 열은 이미지 레이어를 스택하지 않은 채로 도 3과 같이 구현된 경우를 의미할 수 있다.
- [0208] (d) 열은 도 3과 동일한 형태로 구현된 경우를 의미할 수 있고, 네트워크에 입력되기 전에 베이스 및 디테일 레이어는 이미지 레이어와 스택될 수 있다.
- [0209] (e) 열은 도 4와 유사하지만 경로가 3개이고, 이미지, 베이스 및 디테일 레이어 각각에 대응하는 3 개의 경로가 존재하는 경우를 의미할 수 있다.
- [0210] 필터 파라미터의 수는 정합될 수 있고, 모든 네트워크에 대하여 잔차 학습이 수행될 수 있다. 표 1의 결과를 분석하면 다음과 같을 수 있다.
- [0211] 단일 경로 구조에 대해서, (b) 열과 같이 입력을 분해하고, 모든 3 개 레이어들을 스택하면 (a) 열에 비해서 PSNR(Peak Signal to Noise Ratio)이 0.14 dB 향상될 수 있다. 2 개 경로 구조의 경우, 분해 및 스택((d) 열 구조)은 0.35 dB의 PSNR 이득을 가져올 수 있다.
- [0212] 베이스 레이어 및 디테일 레이어에 대하여 별도의 경로를 설계하는 경우((d) 열) (b) 열에 비해서 0.21dB의 PSNR 이득을 가져올 수 있다. (b), (d) 및 (e)를 비교하면 이미지 레이어가 스택되는 것이 중요하다는 것을 확인할 수 있다.
- [0213] 베이스 레이어 및 디테일 레이어를 별도의 경로로 설계하는 것이 가장 중요하고, (d)와 같이 각 경로에 이미지 레이어를 가이드선(guidance)로 스택하면 (a) 열과 같이 분해가 없는 단일 경로 구조에 비해 0.35 dB의 성능 향상을 가져올 수 있다.
- [0214] 전체적인 컬러 밝기(intensity)를 갖는 베이스 레이어와, 신호 진폭에서 섬세한 변화(delicate variation)을 갖는 디테일 레이어의 분해된 이미지 신호에 대하여 특징을 추출하는 것은 향상된 대비에 대해서 전체적인 진폭을 확장하면서 세부 디테일을 복원하는데 도움이 될 수 있다.
- [0215] 도 5a 내지 5c는 각각 입력으로 사용된 이미지, 베이스 및 디테일의 3 개 레이어를 나타내고, 도 5d 내지 5h는 중간 특징 맵들을 나타낼 수 있고, 도 5e, 5f는 메인 네트워크에서 베이스 레이어 및 디테일 레이어 경로의 첫 번째 컨벌루션 레이어를 통과한 후의 이미지를 나타낼 수 있다.
- [0216] 디테일 레이어 경로에서 생성된 특징맵이 경계 및 텍스처를 포함하고, 베이스 레이어 경로에서 생성된 특징맵들은 전체적인 밝기 및 전역 대비를 포함하는 것을 확인할 수 있다. 도 5d는 상술한 수학적 12의  $y_{int}$ 를 나타내는데 그레이 스케일 이미지와 비슷한 것을 확인할 수 있다.
- [0217] 이미지 처리 장치(10)의 성능을 평가하기 위해, 유도 필터 분해, 스킵 연결 및 변조와 같은 네트워크의 상이한 성분들에 대한 애플리케이션 연구가 수행될 수 있다. 토이 네트워크로 수행된 애플리케이션 연구의 결과는 표 2와 같이 나타낼 수 있다.

**표 2**

[0219]		(a)	(b)	(c)	(d)	(e)
	유도 필터	x	√	√	√	√
	스킵	x	x	√	x	√
	변조	x	x	x	√	√
	PSNR(dB)	38.11	38.46	38.47	<b>38.52</b>	38.50
	SSIM	0.9905	0.9916	0.9918	<b>0.9920</b>	<b>0.9920</b>

- [0220] 유도 필터 분해 및 변조는 각각 0.35 dB 및 0.06 dB PSNR 이득을 가져와서 분해 및 변조가 없는 단순한 단일 경로 네트워크((a) 열)에 비해서 총 0.41 dB의 이득을 가져올 수 있다.
- [0221] 베이스 레이어(base layer)에서 디테일 레이어로의 스킵 연결의 효과는 미미할 수 있다. 이는 전체 밝기에 대



한 충분한 정보를 이미 포함하고 있는 디테일 레이어의 스택된(stacked) 이미지 때문일 수 있다. 그러나, 스킵 연결(skip connection)은 네트워크에서 정보 흐름을 도울 수 있다.

[0223] 도 6a는 변조 전 이미지의 예를 나타내고, 도 6b는 제1 변조된 이미지의 예를 나타내고, 도 6c는 제2 변조된 이미지의 예를 나타낸다.

[0224] 도 6a 내지 6c를 참조하면, 표 3과 같은 공유 변조 특징을 추출하기 위해서 이미지, 베이스 및 디테일 레이어를 이용하는 상이한 조합이 사용될 수 있다.

**표 3**

[0226] SMF <sub>b</sub> 의 입력	이미지	베이스	[이미지 베이스]
SMF <sub>d</sub> 의 입력	이미지	디테일	[이미지 디테일]
PSNR (dB)	38.43	38.44	<b>38.52</b>
SSIM	0.9918	0.9918	<b>0.9920</b>

[0227] 표 3에서 [x y]는 x와 y의 연결(concatenation)을 의미할 수 있다.

[0228] 표 3에서 확인할 수 있듯이, 분해된 레이어들을 스택하는 것은 최대 0.09 dB의 PSNR 이득을 가져올 수 있고, 이는 네트워크 입력에 대하여 덜 효과적일 수 있다(표 1에서 0.35 dB PSNR 이득).

[0229] 이것은 전체 HR HDR 이미지를 복원하기 위한 메인 경로(main pass)에 대해서는 통합된 이미지가 필요하고, 베이스 레이어에서의 전체적인 밝기 및 디테일 레이어에서의 고주파 성분과 같은 구체적인 특징들이 변조에 충분하다는 것을 의미할 수 있다.

[0230] 변조맵(modulation map)은 각 변조 블록에서 메인 브랜치 특징맵들에 곱해진 변조 특징들을 나타낼 수 있다. 도 6a 내지 도 6c의 예시는 각각, 입력 이미지, 베이스 레이어 경로의 변조맵(예를 들어, 제1 변조된 특징맵) 및 디테일 레이어 경로의 변조맵(예를 들어, 제2 변조된 특징맵)를 나타낼 수 있다.

[0231] 도 6b 및 도 6c의 변조맵은 각각 베이스 레이어 경로의 RMB1 및 디테일 레이어 경로의 RSMB 1을 시각화한 것일 수 있다.

[0232] 도 5b 내지 5h의 중간 특징맵들과 같이 변조는 베이스 레이어에 대해서는 대체로 세기에 대하여 수행되고, 디테일 레이어에 대해서는 경계 및 디테일에 대하여 수행됨을 확인할 수 있다.

[0234] 도 7a는 종래 방식으로 처리된 이미지의 예를 나타내고, 도 7b는 도 1에 도시된 이미지 처리 장치에 의해 처리된 이미지의 예를 나타내고, 도 7c는 그라운드 트루스 이미지의 예를 나타낸다.

[0235] 도 7a 내지 도 7c를 참조하면, SR 및 ITM을 순차적으로 수행하기 위해서 하나의 SR 방식에 다양한 ITM 방식이 캐스케이드(cascade)로 연결될 수 있다. 이미지들은 각각의 테스트 프레임에 대하여 최대 밝기가 1000 cd/m<sup>2</sup>이 되도록 스케일링될 수 있다. 이미지들은 리사이징되는 대신에 32의 가장 가까운 배수가 되도록 크롭되어서 유효한 픽셀들만 그라운드 트루스(ground truth) 이미지와 비교될 수 있다.

[0236] 비교를 위해 HDR Toolbox를 이용할 수 있고, 모든 방법에 대하여 최대 밝기는 1000 cd/m<sup>2</sup>, HDR TV를 위한 HDR 10 표준(HDR 10 standard)을 따를 수 있다.

[0237] SDR 입력은 ITM 프로세스 전에 선형화될 수 있고, 모든 ITM 방식에 대하여 컬러는 RGB 709에서 RGB 2020으로 변환될 수 있다. 또한, ITM 후에 PQ-OETF가 적용될 수 있고, 포스트 ITM 파이프라인이 적용될 수 있다.

[0238] 성능의 평가는 5 개의 항목에 대해서 수행될 수 있다. 5 개의 항목은 PSNR, 다중 노출 PSNR(multi-exposure PSNR, mPSNR), SSIM, 멀티 스케일 SSIM(multi-scale SSIM, MS-SSIM) HDR-VDP-2.2.1(High Dynamic Range Visual-Difference Predictor-2.2.1)을 포함할 수 있다.

[0239] mPSNR에 대해서, 파이프라인에서 8 비트 양자화는 10비트 양자화로 변경될 수 있고, PSNR은 -3 부터 3까지 노출 값에 대한 평균일 수 있다.

[0240] HDR-VDP에 대해서, 선형화된 Y 채널이 degree 당 pixel(pixel per degree)가 30으로 설정된 휘도(luminance) 조건에서 비교될 수 있다.

[0241] 도 7a와 7b에서 확인할 수 있듯이, 이미지 처리 장치(10)는 종래 방식의 캐스캐이드(직렬연결) 구조에 비하여 대비 및 디테일을 강화한다는 것을 확인할 수 있다.

[0242] 상술한 5 개의 항목은 SF가 2, 4인 조건에 대해서 비교될 수 있다. 비교 결과는 표 4와 같이 나타낼 수 있다. ±는 표준 편차를 의미할 수 있다.

표 4

방식	스케일	PSNR (dB)	mPSNR (dB)	SSIM	MS-SSIM	HDR-VDP
종래 방식의 캐스캐이드	×2	23.59±0.95	25.48±2.07	0.6504±0.1545	0.9737±0.0155	57.24±3.62
이미지 처리 장치	×2	<b>35.48±4.74</b>	<b>37.69±5.76</b>	<b>0.9745±0.0133</b>	<b>0.9837±0.0096</b>	<b>61.19±5.84</b>
종래 방식의 캐스캐이드	×4	23.46±0.99	25.34±2.09	0.6325±0.1621	0.9670±0.0150	56.70±3.72
이미지 처리 장치	×4	<b>33.61±4.32</b>	<b>35.73±5.24</b>	<b>0.9561±0.0259</b>	<b>0.9748±0.0109</b>	<b>56.07±6.83</b>

[0245] 이미지 처리 장치(10)이 종래 방식에 비해서 우수한 성능을 보이는 것을 확인할 수 있다.

[0246] 이미지 처리 장치(10)는 NVIDIA TITAN Xp GPU 상에서 SF=2, 4인 경우에 대해 4K 프레임을 생성하는데 5.85 초, 5.05 초가 소요되었다. 이미지 처리 장치의 전체 필터 파라미터의 수는 SF=2인 경우 2.5M이고, SF=4 인 경우 2.64M 일 수 있다. 종래 기술의 필터 파라미터는 43M 이상일 수 있다.

[0248] 도 8a는 도 1에 도시된 이미지 처리 장치의 입력 이미지의 예를 나타내고, 도 8b는 도 8a의 이미지를 도 1에 도시된 이미지 처리 장치가 처리한 이미지를 나타낸다.

[0249] 도 8a 및 도 8b를 참조하면, 이미지 처리 장치(10)는 딥 SR-ITM을 이용하여 지역 대비(local contrast)를 높여 우수한 주관 품질(subjective quality)을 나타내고, 종래의 딥러닝 기반 SR 및 ITM 기반 방식에 비해서 5 개의 상이한 이미지 품질 측정 항목(quality metric)들에서 우수한 성능을 가질 수 있다.

[0250] 이미지 처리 장치(10)는 톤매핑 되지 않은 4K SDR인 다양한 비디오를 처리하여 HR HDR 비디오를 생성할 수 있다. 도 8a는 그라운드 트루스 이미지가 존재하지 않는 비디오 프레임을 나타낼 수 있고, 이미지 처리 장치(10)는 도 8a의 이미지를 처리하여 도 8b와 같은 HR-HDR 이미지를 생성할 수 있다.

[0251] 이미지 처리 장치(10)가 복원을 통해 대비 및 디테일을 강화한다는 것을 확인할 수 있다.

[0252] 이미지 처리 장치(10)는 입력 SDR 이미지를 저주파 및 고주파 정보로 나뉘는 베이스 레이어 및 디테일 레이어로 분해하여 처리하는 공동 SR 및 eDR ITM 프레임 워크를 제공할 수 있다.

[0253] 입력 분해를 통해서, 이미지 처리 장치(10)는 공간 업스케일링(spatial up-scaling)을 위한 디테일 레이어에 의해 고주파 디테일을 정확하게 예측할 수 있고, ITM 태스크를 위한 베이스 레이어에 의해 HDR 밝기 컨텍스트로 전체 세기(intensity)를 확장할 수 있다.

[0254] 이미지 처리 장치(10)는 변조를 통해 공간적으로 변하는 연산을 도입함으로써 이미지 신호 진폭에서 국부 대비를 향상시킬 수 있다. 이미지 처리 장치(10)는 픽셀 도메인에서 HR HDR 이미지를 직접적으로 생성하여, UHD HDR 소비자 디스플레이를 위해 고품질의 시각 콘텐츠를 생성할 수 있다.

[0256] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

참가지이다.

[0257] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상장치(virtual equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(signal wave)에 영구적으로, 또는 일시적으로 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

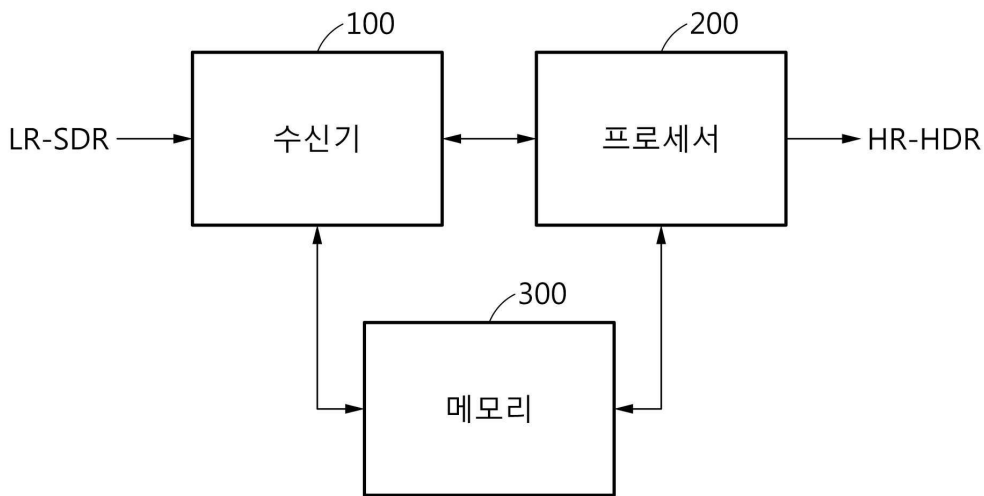
[0258] 이상과 같이 실시예들이 비록 한정된 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기를 기초로 다양한 기술적 수정 및 변형을 적용할 수 있다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

[0259] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 청구범위의 범위에 속한다.

**도면**

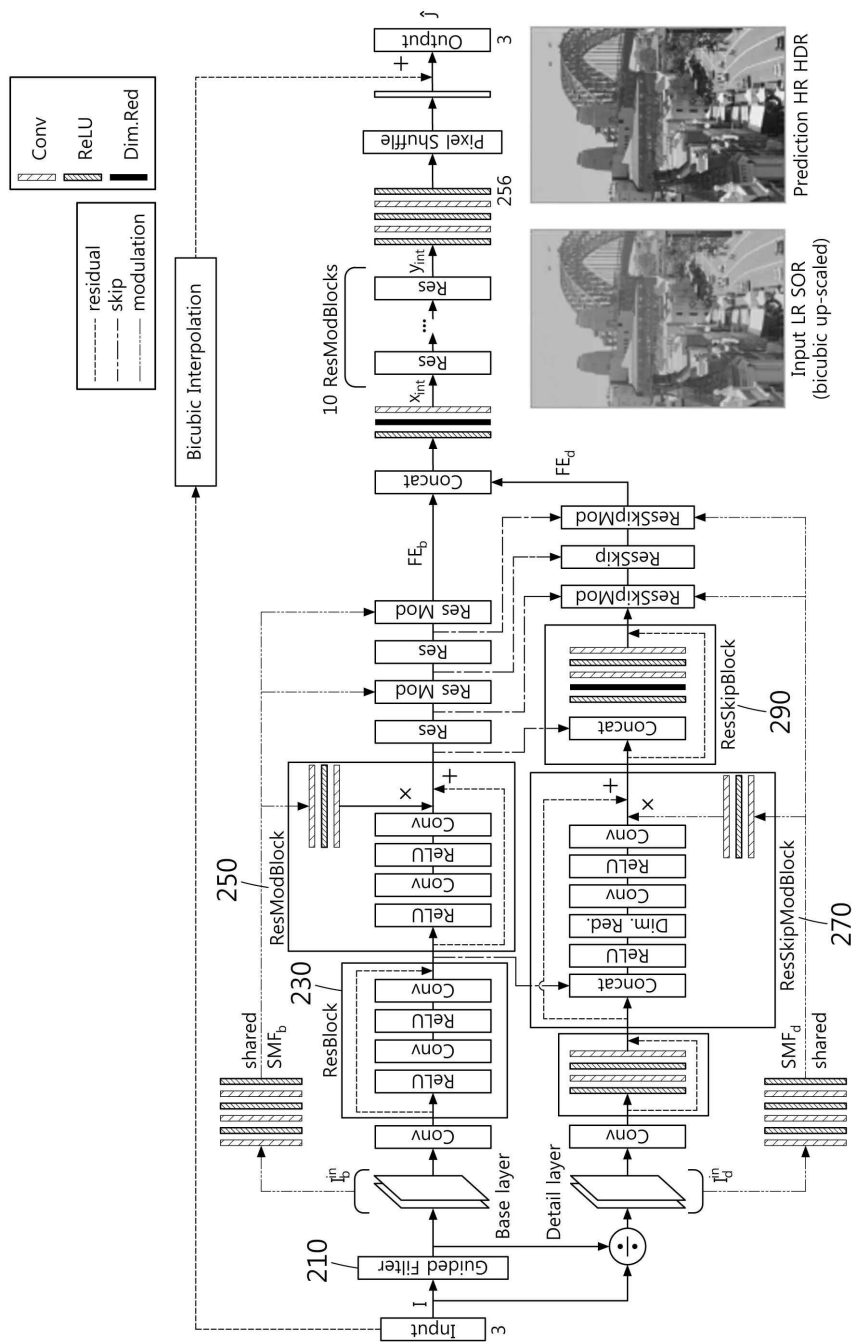
**도면1**

10

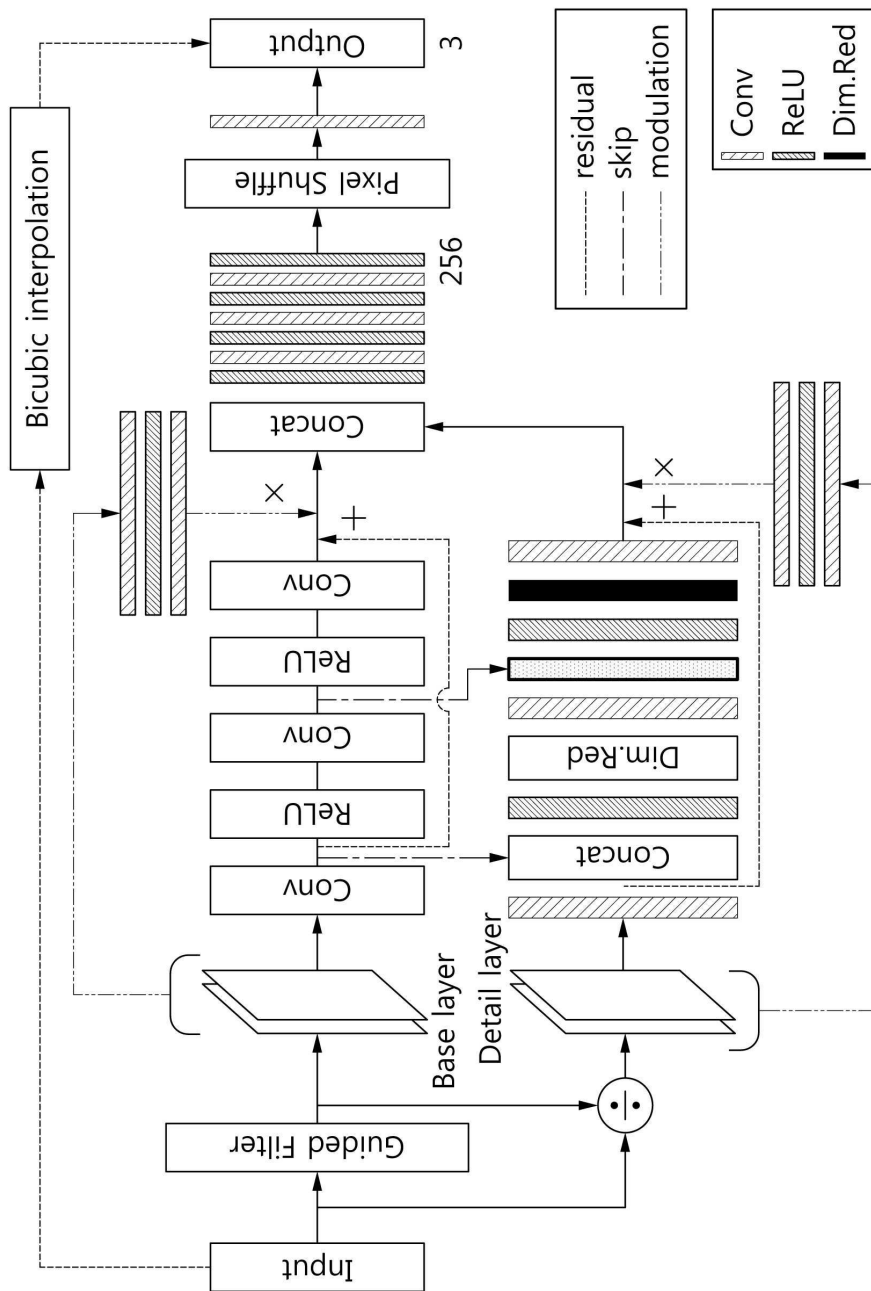




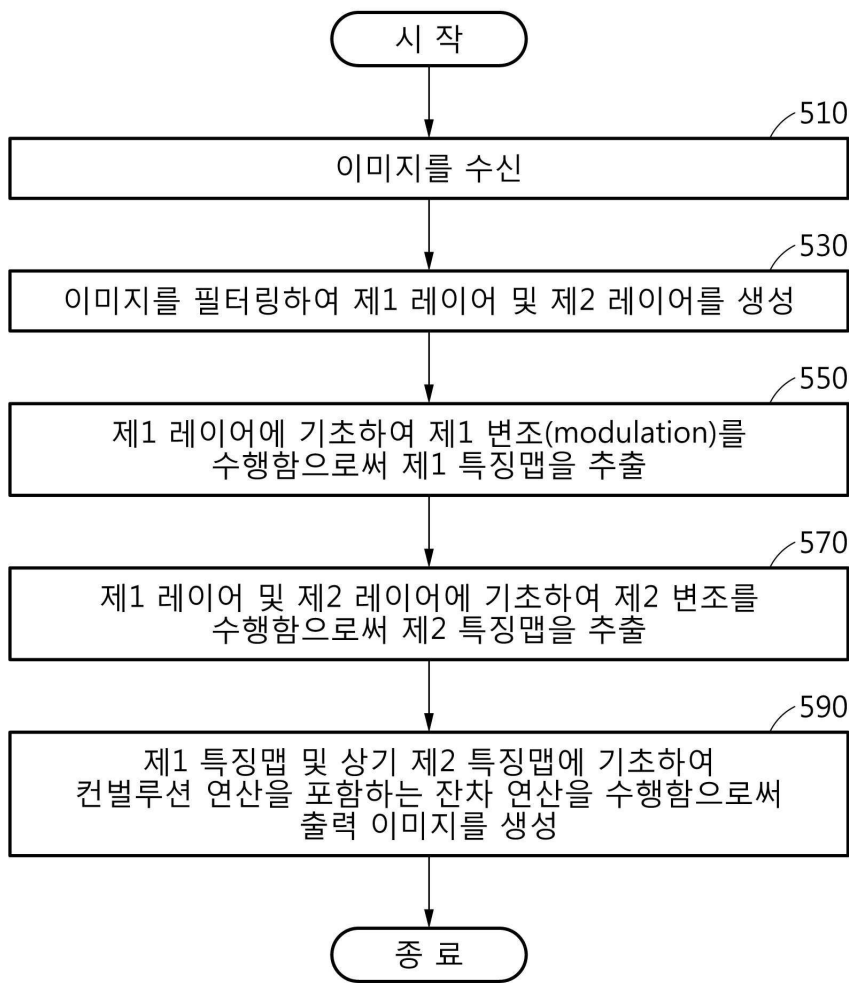
도면2



도면3



도면4



도면5a



도면5b



도면5c



도면5d



도면5e





도면5f



도면5g



도면5h



도면6a



도면6b



도면6c



도면7a



도면7b



도면7c



도면8a



도면8b

