



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0090716
(43) 공개일자 2023년06월22일

(51) 국제특허분류(Int. Cl.)
G06T 5/00 (2019.01) G06N 3/04 (2023.01)
G06N 3/08 (2023.01) G06T 5/50 (2006.01)
G06T 7/215 (2017.01)

(52) CPC특허분류
G06T 5/001 (2013.01)
G06N 3/045 (2023.01)

(21) 출원번호 10-2021-0179711
(22) 출원일자 2021년12월15일
심사청구일자 없음

(71) 출원인
삼성전자주식회사
경기도 수원시 영통구 삼성로 129 (매탄동)
포항공과대학교 산학협력단

경상북도 포항시 남구 청암로 77 (지곡동)

(72) 발명자
최재석
경기도 용인시 수지구 용구대로 2742 (죽전동, 동성1차아파트) 102동 1401호

조민수
경상북도 포항시 남구 효성로 55, 105동 2502호
(효자동, 효자풍림아이원아파트)
(뒷면에 계속)

(74) 대리인
특허법인 무한

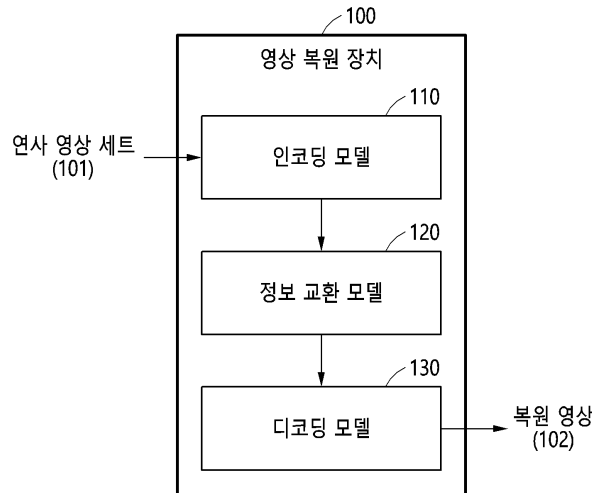
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 연사 영상 기반의 영상 복원 방법 및 장치

(57) 요약

연사 영상 기반의 영상 복원 방법 및 장치가 제공된다. 일 실시예에 따르면, 그 방법은 연사 영상 세트의 개별 영상들을 인코딩하여 개별 영상들에 대응하는 복수의 특징 표현들을 생성하고, 복수의 특징 표현들 중 레퍼런스 특징 표현을 결정하고, 레퍼런스 특징 표현, 및 복수의 특징 표현들 중 제1 특징 표현을 포함하는 제1 비교 쌍을 결정하고, 레퍼런스 특징 표현과 제1 특징 표현의 유사도 스코어 맵에 기초하여 제1 비교 쌍의 제1 모션-임베딩 특징 표현을 생성하고, 제1 모션-임베딩 특징 표현을 포함하는 복수의 모션-임베딩 특징 표현들을 융합하여 융합 결과를 생성하고, 및 융합 결과를 디코딩하여 적어도 하나의 복원 영상을 생성하는 단계들을 포함한다.

대표도 - 도1



(52) CPC특허분류

G06N 3/08 (2023.01)
G06T 5/50 (2023.01)
G06T 7/215 (2017.01)
G06T 2207/20081 (2013.01)
G06T 2207/20084 (2013.01)

(72) 발명자

김상현

경기도 안양시 동안구 관평로138번길 63, 17동
1003호 (평촌동, 초원부영아파트)

이종민

경상북도 포항시 북구 중흥로139번길 12, A동 302
호 (죽도동, 퍼스트하우스)

권기남

경기도 용인시 기흥구 기흥역로58번길 10, 205동
605호 (구갈동, 기흥역 센트럴푸르지오)

서건석

경기도 수원시 영통구 도청로 10, A동 993호 (이의
동, 광고센트럴푸르지오시티)

이형욱

경기도 수원시 영통구 대학로 98, 8101동 602호 (이
의동, 광고 호반베르디움트라엘)

명세서

청구범위

청구항 1

연사 영상 세트의 개별 영상들을 인코딩하여 상기 개별 영상들에 대응하는 복수의 특징 표현들을 생성하는 단계;

상기 복수의 특징 표현들 중 레퍼런스 특징 표현을 결정하는 단계;

상기 레퍼런스 특징 표현, 및 상기 복수의 특징 표현들 중 제1 특징 표현을 포함하는 제1 비교 쌍을 결정하는 단계;

상기 레퍼런스 특징 표현과 상기 제1 특징 표현의 유사도 스코어 맵에 기초하여 상기 제1 비교 쌍의 제1 모션-임베딩 특징 표현을 생성하는 단계;

상기 제1 모션-임베딩 특징 표현을 포함하는 복수의 모션-임베딩 특징 표현들을 융합하여 융합 결과를 생성하는 단계; 및

상기 융합 결과를 디코딩하여 적어도 하나의 복원 영상을 생성하는 단계를 포함하는 영상 복원 방법.

청구항 2

제1항에 있어서,

상기 제1 모션-임베딩 특징 표현을 결정하는 단계는

상기 레퍼런스 특징 표현의 제1 지점의 제1 특징 벡터, 및 상기 제1 특징 표현에서 상기 제1 지점의 대응 지점을 중심으로 갖는 제1 관심 블록을 결정하는 단계;

상기 제1 특징 벡터와 상기 제1 관심 블록의 채널 벡터들 간의 유사도에 기초하여 상기 제1 지점에 대응하는 제1 유사도 스코어 맵을 생성하는 단계

상기 제1 유사도 스코어 맵에 기초하여 상기 제1 지점에 대응하는 제1 모션-지각 커널을 생성하는 단계;

상기 제1 관심 블록과 상기 제1 모션-지각 커널 간의 연산에 기초하여 상기 제1 지점에 대응하는 제1 모션-임베딩 벡터를 생성하는 단계; 및

상기 제1 모션-임베딩 벡터를 포함하는 모션-임베딩 벡터들에 기초하여 상기 제1 모션-임베딩 특징 표현을 생성하는 단계

를 포함하는, 영상 복원 방법.

청구항 3

제2항에 있어서,

상기 제1 모션-임베딩 특징 표현을 결정하는 단계는

상기 레퍼런스 특징 표현의 제2 지점의 제2 특징 벡터, 및 상기 제1 특징 표현에서 상기 제2 지점의 대응 지점을 중심으로 갖는 제2 관심 블록을 결정하는 단계; 및

상기 제2 관심 블록에 기초하여 상기 제2 지점에 대응하는 제2 모션-임베딩 벡터를 생성하는 단계를

를 더 포함하고,

상기 제1 모션-임베딩 특징 표현을 생성하는 단계는

상기 제1 모션-임베딩 벡터 및 상기 제2 모션-임베딩 벡터를 포함하는 상기 모션-임베딩 벡터들에 기초하여 상

기 제1 모션-임베딩 특징 표현을 생성하는 단계를 포함하는, 영상 복원 방법.

청구항 4

제2항에 있어서,

상기 제1 모션-지각 커널을 생성하는 단계는

상기 제1 유사도 스코어 맵에 대응하는 제1 모션 정보 벡터를 결정하는 단계;

상기 제1 관심 블록과 상기 제1 유사도 스코어 맵 간의 연산에 기초하여 제1 컨텍스트 정보 벡터를 생성하는 단계; 및

상기 제1 모션 정보 벡터와 상기 제1 컨텍스트 정보 벡터를 융합하여 상기 제1 모션-지각 커널을 생성하는 단계를 포함하는, 영상 복원 방법.

청구항 5

제4항에 있어서,

상기 제1 모션 정보 벡터와 상기 제1 컨텍스트 정보 벡터를 융합하여 상기 제1 모션-지각 커널을 생성하는 단계는

제1 완전 연결 네트워크를 이용하여 상기 제1 모션 정보 벡터에 대응하는 제1 채널 벡터를 생성하는 단계;

제2 완전 연결 네트워크를 이용하여 상기 제1 컨텍스트 정보 벡터에 대응하는 제2 채널 벡터를 생성하는 단계; 및

상기 제1 채널 벡터와 상기 제2 채널 벡터를 융합하여 상기 제1 모션-지각 커널을 생성하는 단계를 포함하는, 영상 복원 방법.

청구항 6

제1항에 있어서,

상기 특징 표현들을 생성하는 단계는

적어도 하나의 컨볼루션 레이어를 포함하는 인코딩 모델을 이용하여 상기 개별 영상들을 인코딩하는 단계를 포함하고,

상기 적어도 하나의 복원 영상을 생성하는 단계는

적어도 하나의 업-컨볼루션 레이어를 포함하는 디코딩 모델을 이용하여 상기 융합 결과를 디코딩하는 단계를 포함하는,

영상 복원 방법.

청구항 7

제1항에 있어서,

상기 개별 영상들 중에 앵커 영상을 선택하는 단계를

를 더 포함하고,

상기 레퍼런스 특징 표현을 결정하는 단계는

상기 특징 표현들 중 상기 앵커 영상의 특징 표현을 상기 레퍼런스 특징 표현으로 결정하는 단계를 포함하는,

영상 복원 방법.

청구항 8

제7항에 있어서,
 상기 앵커 영상을 선택하는 단계는
 품질 기반으로 상기 개별 영상들 중에 상기 앵커 영상을 선택하는 단계를 포함하는,
 영상 복원 방법.

청구항 9

제1항에 있어서,
 상기 개별 영상들은 제1 개별 영상 및 제2 개별 영상을 포함하고,
 상기 적어도 하나의 복원 영상은 상기 제1 개별 영상의 제1 복원 영상 및 상기 제2 개별 영상의 제2 복원 영상을 포함하고,
 상기 레퍼런스 특징 표현을 결정하는 단계는
 상기 제1 개별 영상을 상기 제1 복원 영상으로 복원하는 경우, 상기 특징 표현들 중 상기 제1 개별 영상의 특징 표현을 상기 레퍼런스 특징 표현으로 결정하는 단계; 및
 상기 제2 개별 영상을 상기 제2 복원 영상으로 복원하는 경우, 상기 특징 표현들 중 상기 제2 개별 영상의 특징 표현을 상기 레퍼런스 특징 표현으로 결정하는 단계
 를 포함하는, 영상 복원 방법.

청구항 10

하드웨어와 결합되어 제1항 내지 제9항 중 어느 하나의 항의 방법을 실행시키기 위하여 컴퓨터 판독 가능한 기록매체에 저장된 컴퓨터 프로그램.

청구항 11

프로세서; 및
 상기 프로세서에서 실행가능한 명령어들을 포함하는 메모리
 를 포함하고,
 상기 명령어들이 상기 프로세서에서 실행되면, 상기 프로세서는
 연사 영상 세트의 개별 영상들을 인코딩하여 상기 개별 영상들에 대응하는 복수의 특징 표현들을 생성하고,
 상기 복수의 특징 표현들 중 레퍼런스 특징 표현을 결정하고,
 상기 레퍼런스 특징 표현, 및 상기 복수의 특징 표현들 중 제1 특징 표현을 포함하는 제1 비교 쌍을 결정하고,
 상기 레퍼런스 특징 표현과 상기 제1 특징 표현의 유사도 스코어 맵에 기초하여 상기 제1 비교 쌍의 제1 모션-임베딩 특징 표현을 생성하고,
 상기 제1 모션-임베딩 특징 표현을 포함하는 복수의 모션-임베딩 특징 표현들을 융합하여 융합 결과를 생성하고,
 상기 융합 결과를 디코딩하여 적어도 하나의 복원 영상을 생성하는,
 영상 복원 장치.

청구항 12

제11항에 있어서,
 상기 프로세서는
 상기 레퍼런스 특징 표현의 제1 지점의 제1 특징 벡터, 및 상기 제1 특징 표현에서 상기 제1 지점의 대응 지점을 중심으로 갖는 제1 관심 블록을 결정하고,

상기 제1 특징 벡터와 상기 제1 관심 블록의 채널 벡터들 간의 유사도에 기초하여 상기 제1 지점에 대응하는 제1 유사도 스코어 맵을 생성하고,

상기 제1 유사도 스코어 맵에 기초하여 상기 제1 지점에 대응하는 제1 모션-지각 커널을 생성하고,

상기 제1 관심 블록과 상기 제1 모션-지각 커널 간의 연산에 기초하여 상기 제1 지점에 대응하는 제1 모션-임베딩 벡터를 생성하고,

상기 제1 모션-임베딩 벡터를 포함하는 모션-임베딩 벡터들에 기초하여 상기 제1 모션-임베딩 특징 표현을 생성하는,

영상 복원 장치.

청구항 13

제12항에 있어서,

상기 프로세서는

상기 레퍼런스 특징 표현의 제2 지점의 제2 특징 벡터, 및 상기 제1 특징 표현에서 상기 제2 지점의 대응 지점을 중심으로 갖는 제2 관심 블록을 결정하고,

상기 제2 관심 블록에 기초하여 상기 제2 지점에 대응하는 제2 모션-임베딩 벡터를 생성하고,

상기 제1 모션-임베딩 벡터 및 상기 제2 모션-임베딩 벡터를 포함하는 상기 모션-임베딩 벡터들에 기초하여 상기 제1 모션-임베딩 특징 표현을 생성하는,

영상 복원 장치.

청구항 14

제12항에 있어서,

상기 프로세서는

상기 제1 유사도 스코어 맵에 대응하는 제1 모션 정보 벡터를 결정하고,

상기 제1 관심 블록과 상기 제1 유사도 스코어 맵 간의 연산에 기초하여 제1 컨텍스트 정보 벡터를 생성하고,

상기 제1 모션 정보 벡터와 상기 제1 컨텍스트 정보 벡터를 융합하여 상기 제1 모션-지각 커널을 생성하는,

영상 복원 장치.

청구항 15

제14항에 있어서,

상기 프로세서는

제1 완전 연결 네트워크를 이용하여 상기 제1 모션 정보 벡터에 대응하는 제1 채널 벡터를 생성하고,

제2 완전 연결 네트워크를 이용하여 상기 제1 컨텍스트 정보 벡터에 대응하는 제2 채널 벡터를 생성하고,

상기 제1 채널 벡터와 상기 제2 채널 벡터를 융합하여 상기 제1 모션-지각 커널을 생성하는,

영상 복원 장치.

청구항 16

제11항에 있어서,

상기 프로세서는,

상기 개별 영상들 중에 앵커 영상을 선택하고,

상기 특징 표현들 중 상기 앵커 영상의 특징 표현을 상기 레퍼런스 특징 표현으로 결정하는,

영상 복원 장치.

청구항 17

제11항에 있어서,

상기 개별 영상들은 제1 개별 영상 및 제2 개별 영상을 포함하고,

상기 적어도 하나의 복원 영상은 상기 제1 개별 영상의 제1 복원 영상 및 상기 제2 개별 영상의 제2 복원 영상을 포함하고,

상기 프로세서는

상기 제1 개별 영상을 상기 제1 복원 영상으로 복원하는 경우, 상기 특징 표현들 중 상기 제1 개별 영상의 특징 표현을 상기 레퍼런스 특징 표현으로 결정하고,

상기 제2 개별 영상을 상기 제2 복원 영상으로 복원하는 경우, 상기 특징 표현들 중 상기 제2 개별 영상의 특징 표현을 상기 레퍼런스 특징 표현으로 결정하는,

영상 복원 장치.

청구항 18

개별 영상들을 포함하는 연사 영상 세트를 생성하는 카메라; 및

상기 개별 영상들을 인코딩하여 상기 개별 영상들에 대응하는 복수의 특징 표현들을 생성하고,

상기 복수의 특징 표현들 중 레퍼런스 특징 표현을 결정하고,

상기 레퍼런스 특징 표현, 및 상기 복수의 특징 표현들 중 제1 특징 표현을 포함하는 제1 비교 쌍을 결정하고,

상기 레퍼런스 특징 표현과 상기 제1 특징 표현의 유사도 스코어 맵에 기초하여 상기 제1 비교 쌍의 제1 모션-임베딩 특징 표현을 생성하고,

상기 제1 모션-임베딩 특징 표현을 포함하는 복수의 모션-임베딩 특징 표현들을 융합하여 융합 결과를 생성하고,

상기 융합 결과를 디코딩하여 적어도 하나의 복원 영상을 생성하는, 프로세서

를 포함하는, 전자 장치.

청구항 19

제18항에 있어서,

상기 프로세서는

상기 레퍼런스 특징 표현의 제1 지점의 제1 특징 벡터, 및 상기 제1 특징 표현에서 상기 제1 지점의 대응 지점을 중심으로 갖는 제1 관심 블록을 결정하고,

상기 제1 특징 벡터와 상기 제1 관심 블록의 채널 벡터들 간의 유사도에 기초하여 상기 제1 지점에 대응하는 제1 유사도 스코어 맵을 생성하고,

상기 제1 유사도 스코어 맵에 기초하여 상기 제1 지점에 대응하는 제1 모션-지각 커널을 생성하고,

상기 제1 관심 블록과 상기 제1 모션-지각 커널 간의 연산에 기초하여 상기 제1 지점에 대응하는 제1 모션-임베딩 벡터를 생성하고,

상기 제1 모션-임베딩 벡터를 포함하는 모션-임베딩 벡터들에 기초하여 상기 제1 모션-임베딩 특징 표현을 생성하는,

전자 장치.

청구항 20

제19항에 있어서,
 상기 프로세서는
 상기 제1 유사도 스코어 맵에 대응하는 제1 모션 정보 벡터를 결정하고,
 상기 제1 관심 블록과 상기 제1 유사도 스코어 맵 간의 연산에 기초하여 제1 컨텍스트 정보 벡터를 생성하고,
 상기 제1 모션 정보 벡터와 상기 제1 컨텍스트 정보 벡터를 융합하여 상기 제1 모션-지각 커널을 생성하는,
 전자 장치.

발명의 설명

기술 분야

[0001] 아래 실시예들은 연사 영상 기반의 영상 복원 방법 및 장치에 관한 것이다.

배경 기술

[0002] 영상 복원은 열화 상태의 영상을 향상된 화질의 영상으로 복원하는 기술이다. 영상 복원을 위해 딥 러닝 기반의 뉴럴 네트워크가 이용될 수 있다. 뉴럴 네트워크는 딥 러닝에 기반하여 훈련된 후, 비선형적 관계에 있는 입력 데이터 및 출력 데이터를 서로 매핑함으로써 목적에 맞는 추론(inference)을 수행해낼 수 있다. 이러한 맵핑을 생성하는 훈련된 능력은 신경망의 학습 능력이라 할 수 있다. 더구나, 영상 복원과 같은 특화된 목적을 위해 훈련된 신경망은, 예를 들어 훈련하지 않은 입력 패턴에 대하여 비교적 정확한 출력을 발생시키는 일반화 능력을 가질 수 있다.

발명의 내용

해결하려는 과제

과제의 해결 수단

[0003] 일 실시예에 따르면, 영상 복원 방법은 연사 영상 세트의 개별 영상들을 인코딩하여 상기 개별 영상들에 대응하는 복수의 특징 표현들을 생성하는 단계; 상기 복수의 특징 표현들 중 레퍼런스 특징 표현을 결정하는 단계; 상기 레퍼런스 특징 표현, 및 상기 복수의 특징 표현들 중 제1 특징 표현을 포함하는 제1 비교 쌍을 결정하는 단계; 상기 레퍼런스 특징 표현과 상기 제1 특징 표현의 유사도 스코어 맵에 기초하여 상기 제1 비교 쌍의 제1 모션-임베딩 특징 표현을 생성하는 단계; 상기 제1 모션-임베딩 특징 표현을 포함하는 복수의 모션-임베딩 특징 표현들을 융합하여 융합 결과를 생성하는 단계; 및 상기 융합 결과를 디코딩하여 적어도 하나의 복원 영상을 생성하는 단계를 포함한다.

[0004] 일 실시예에 따르면, 영상 복원 장치는 프로세서; 및 상기 프로세서에서 실행가능한 명령어들을 포함하는 메모리를 포함하고, 상기 명령어들이 상기 프로세서에서 실행되면, 상기 프로세서는 연사 영상 세트의 개별 영상들을 인코딩하여 상기 개별 영상들에 대응하는 복수의 특징 표현들을 생성하고, 상기 복수의 특징 표현들 중 레퍼런스 특징 표현을 결정하고, 상기 레퍼런스 특징 표현, 및 상기 복수의 특징 표현들 중 제1 특징 표현을 포함하는 제1 비교 쌍을 결정하고, 상기 레퍼런스 특징 표현과 상기 제1 특징 표현의 유사도 스코어 맵에 기초하여 상기 제1 비교 쌍의 제1 모션-임베딩 특징 표현을 생성하고, 상기 제1 모션-임베딩 특징 표현을 포함하는 복수의 모션-임베딩 특징 표현들을 융합하여 융합 결과를 생성하고, 상기 융합 결과를 디코딩하여 적어도 하나의 복원 영상을 생성한다.

[0005] 일 실시예에 따르면, 전자 장치는 개별 영상들을 포함하는 연사 영상 세트를 생성하는 카메라; 및 상기 개별 영상들을 인코딩하여 상기 개별 영상들에 대응하는 복수의 특징 표현들을 생성하고, 상기 복수의 특징 표현들 중 레퍼런스 특징 표현을 결정하고, 상기 레퍼런스 특징 표현, 및 상기 복수의 특징 표현들 중 제1 특징 표현을 포함하는 제1 비교 쌍을 결정하고, 상기 레퍼런스 특징 표현과 상기 제1 특징 표현의 유사도 스코어 맵에 기초하여 상기 제1 비교 쌍의 제1 모션-임베딩 특징 표현을 생성하고, 상기 제1 모션-임베딩 특징 표현을 포함하는 복수의 모션-임베딩 특징 표현들을 융합하여 융합 결과를 생성하고, 상기 융합 결과를 디코딩하여 적어도 하나의

복원 영상을 생성하는, 프로세서를 포함한다.

도면의 간단한 설명

- [0006] 도 1은 일 실시예에 따른 영상 복원 장치의 동작을 개략적으로 나타낸다.
- 도 2는 일 실시예에 따른 입력 영상을 특징 표현들로 인코딩하는 동작을 나타낸다.
- 도 3은 일 실시예에 따른 특징 표현들의 모션-임베딩 특징 표현들을 생성하고, 및 모션-임베딩 특징 표현들을 융합하는 동작을 나타낸다.
- 도 4는 일 실시예에 따른 모션-임베딩 특징 표현들을 생성하는 동작을 나타낸다.
- 도 5는 일 실시예에 따른 모션 정보 벡터와 컨텍스트 정보 벡터를 융합하는 동작을 나타낸다.
- 도 6은 일 실시예에 따른 특징 표현들을 복원 영상으로 디코딩하는 동작을 나타낸다.
- 도 7은 일 실시예에 따른 데이터 흐름을 나타낸다.
- 도 8은 일 실시예에 따른 영상 복원 방법을 나타낸다.
- 도 9는 일 실시예에 따른 영상 복원 장치의 구성을 나타낸다.
- 도 10은 일 실시예에 따른 전자 장치의 구성을 나타낸다.

발명을 실시하기 위한 구체적인 내용

- [0007] 실시예들에 대한 특정한 구조적 또는 기능적 설명들은 단지 예시를 위한 목적으로 개시된 것으로서, 다양한 형태로 변경되어 구현될 수 있다. 따라서, 실제 구현되는 형태는 개시된 특정 실시예로만 한정되는 것이 아니며, 본 명세서의 범위는 실시예들로 설명한 기술적 사상에 포함되는 변경, 균등물, 또는 대체물을 포함한다.
- [0008] 제1 또는 제2 등의 용어를 다양한 구성요소들을 설명하는데 사용될 수 있지만, 이런 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 해석되어야 한다. 예를 들어, 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소는 제1 구성요소로도 명명될 수 있다.
- [0009] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다.
- [0010] 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 설명된 특징, 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함으로써 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0011] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 해당 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0012] 이하, 실시예들을 첨부된 도면들을 참조하여 상세하게 설명한다. 첨부 도면을 참조하여 설명함에 있어, 도면 부호에 관계없이 동일한 구성 요소는 동일한 참조 부호를 부여하고, 이에 대한 중복되는 설명은 생략하기로 한다.
- [0013] 도 1은 일 실시예에 따른 영상 복원 장치의 동작을 개략적으로 나타낸다. 도 1을 참조하면, 영상 복원 장치(image restoration apparatus, 100)는 연사 영상 세트(burst images set, 101)를 수신하고, 연사 영상 세트(101)에 기초하여 복원 영상(restored image, 102)을 생성하고, 복원 영상(102)을 출력할 수 있다. 연사 영상 세트(101)는 카메라(예: 카메라(1030))에 의해 생성될 수 있다. 연사 영상 세트(101)는 연속적으로 촬영된 복수의 영상들을 포함할 수 있다. 예를 들어, 연사 영상 세트(101)는 비디오 촬영 기능을 통해 생성된 비디오 영상이거나, 혹은 연사 촬영 기능을 통해 생성된 연속적인 스틸 영상일 수 있다. 연사 영상 세트(101)의 복수의 영상들의 각 영상은 개별 영상으로 지칭될 수 있다. 예를 들어, 연사 영상 세트(101)는 N개의 개별 영상들을 포함할 수 있다. 비디오의 경우 각 영상 프레임이 개별 영상에 해당할 수 있고, 연사 촬영 영상의 경우 각 스

틸 영상이 개별 영상에 해당할 수 있다.

- [0014] 카메라로 대상 객체를 촬영하여 연사 영상 세트(101)를 생성하는 것을 가정할 때, 카메라 및/또는 대상 객체의 움직임, 및/또는 주변광의 변화(예: 조도, 색상 등) 등으로 인해 연사 영상 세트(101)의 각 개별 영상은 서로 다른 다양한 특성을 가질 수 있다. 연사 영상 세트(101)가 저조도 환경과 같이 열악한 환경에서 촬영되고, 및/또는 각 개별 영상이 열화된 화질을 가지는 경우, 각 개별 영상의 다양한 특성을 적절히 조합하면 향상된 화질을 갖는 복원 영상(102)이 도출될 수 있다. 따라서, 저화질의 개별 영상들에 대한 복원 작업을 통해 고화질의 복원 영상(102)이 도출될 수 있다.
- [0015] 연사 영상 세트(101)의 생성 시 카메라 및/또는 대상 객체의 모션이 발생할 수 있고, 연사 영상 세트(101)의 개별 영상들에는 이러한 모션으로 인한 열화가 포함될 수 있다. 영상 복원 장치(100)는 이러한 모션을 고려하여 영상 복원을 수행할 수 있다. 비교적 짧은 시간에 연속적으로 생성되는 연사 영상 세트(101)의 특성 상 모션에 의한 영상 픽셀의 변화 영역은 제한적일 수 있다. 아래에서 다시 설명하겠지만, 영상 복원 장치(100)는 제한적인 크기의 유사도 스코어 맵을 이용하여 모션 분석에 필요한 연산 효율을 향상시킬 수 있다.
- [0016] 영상 복원 장치(100)는 연사 영상 세트(101)의 개별 영상들에 기초한 인코딩을 수행하여 개별 영상들에 대응하는 특징 표현들을 생성하고, 특징 표현들 간의 정보 교환을 수행하여 모션-임베딩 특징 표현들(motion-embedding feature representations)을 생성하고, 모션-임베딩 특징 표현들에 기초한 디코딩을 수행하여 복원 영상(102)을 생성할 수 있다. 영상 복원 장치(100)는 인코딩 모델(110), 정보 교환 모델(120), 및 디코딩 모델(130)을 이용하여 이러한 일련의 동작들을 수행할 수 있다.
- [0017] 인코딩 모델(110), 정보 교환 모델(120), 및 디코딩 모델(130)의 적어도 일부는 뉴럴 네트워크 모델일 수 있다. 뉴럴 네트워크는 복수의 레이어들을 포함하는 딥 뉴럴 네트워크(deep neural network, DNN)에 해당할 수 있다. 이때, 복수의 레이어들은 입력 레이어(input layer), 적어도 하나의 히든 레이어(hidden layer), 및 출력 레이어(output layer)를 포함할 수 있다.
- [0018] 딥 뉴럴 네트워크는 완전 연결 네트워크(fully connected network, FCN), 컨볼루션 뉴럴 네트워크(convolutional neural network, CNN), 및 리커런트 뉴럴 네트워크(recurrent neural network, RNN) 중 적어도 하나를 포함할 수 있다. 예를 들어, 뉴럴 네트워크 내 복수의 레이어들 중 적어도 일부는 CNN에 해당할 수 있고, 다른 일부는 FCN에 해당할 수 있다. 이 경우, CNN은 컨볼루션 레이어로 지칭될 수 있고, FCN은 완전 연결 레이어로 지칭될 수 있다.
- [0019] CNN의 경우, 각 레이어에 입력되는 데이터는 입력 특징 맵(input feature map)으로 지칭될 수 있고, 각 레이어에서 출력되는 데이터는 출력 특징 맵(output feature map)으로 지칭될 수 있다. 입력 특징 맵 및 출력 특징 맵은 특징 표현(feature representation) 혹은 액티베이션 데이터(activation data)로 지칭될 수도 있다. 컨볼루션 레이어가 입력 레이어에 해당하는 경우, 입력 레이어의 입력 특징 맵은 입력 영상일 수 있다.
- [0020] 뉴럴 네트워크는 딥 러닝에 기반하여 트레이닝된 후, 비선형적 관계에 있는 입력 데이터 및 출력 데이터를 서로 매핑함으로써 트레이닝 목적에 맞는 추론(inference)을 수행해낼 수 있다. 딥 러닝은 빅 데이터 세트로부터 영상 또는 음성 인식과 같은 문제를 해결하기 위한 기계 학습 기법이다. 딥 러닝은 준비된 트레이닝 데이터를 이용하여 뉴럴 네트워크를 트레이닝하면서 에너지가 최소화되는 지점을 찾아가는 최적화 문제 풀이 과정으로 이해될 수 있다.
- [0021] 딥 러닝의 지도식(supervised) 또는 비지도식(unsupervised) 학습을 통해 뉴럴 네트워크의 구조, 혹은 모델에 대응하는 웨이트(weight)가 구해질 수 있고, 이러한 웨이트를 통해 입력 데이터 및 출력 데이터가 서로 매핑될 수 있다. 뉴럴 네트워크의 폭과 깊이가 충분히 크면 임의의 함수를 구현할 수 있을 만큼의 용량(capacity)을 가질 수 있다. 뉴럴 네트워크가 적절한 트레이닝 과정을 통해 충분히 많은 양의 트레이닝 데이터를 학습하면 최적의 성능을 달성할 수 있다.
- [0022] 아래에서 뉴럴 네트워크가 '미리' 트레이닝된 것으로 표현될 수 있는데, 여기서 '미리'는 뉴럴 네트워크가 '시작'되기 전을 나타낼 수 있다. 뉴럴 네트워크가 '시작'되었다는 것은 뉴럴 네트워크가 추론을 위한 준비가 된 것을 의미할 수 있다. 예를 들어, 뉴럴 네트워크가 '시작'된 것은 뉴럴 네트워크가 메모리에 로드된 것, 혹은 뉴럴 네트워크가 메모리에 로드된 이후 뉴럴 네트워크에 추론을 위한 입력 데이터가 입력된 것을 포함할 수 있다.
- [0023] 영상 복원 장치(100)는 모션 정보를 이용하면서 뉴럴 네트워크 모델을 실행할 수 있다. 영상 복원 장치(100)는 개별 영상들의 특징 표현들 간의 정보 교환을 수행하여 모션-임베딩 특징 표현을 생성할 수 있다. 정보 교환은

정보 교환 모델(120)의 모션-지각 모델(motion-aware model)을 통해 수행될 수 있다. 영상 복원 장치(100)는 개별 영상들의 특징 표현들 간의 유사도 스코어 맵에 기초하여 정보 교환을 수행할 수 있다. 모션-임베딩 특징 표현은 카메라 및/또는 대상 객체의 모션에 따른 개별 영상들의 모션 정보를 포함할 수 있다. 영상 복원 장치(100)는 이러한 모션 정보를 고려하여 영상 복원을 수행하여 복원 영상(102)의 품질을 향상시킬 수 있다. 비교적 짧은 시간에 연속적으로 생성되는 연사 영상 세트(101)의 특성 상 모션에 의한 영상 픽셀의 변화 영역은 제한적일 수 있으므로, 영상 복원 장치(100)는 제한된 크기의 유사도 스코어 맵을 사용함으로써 모션 정보를 효율적으로 이용할 수 있다.

[0024] 도 2는 일 실시예에 따른 입력 영상을 특징 표현들로 인코딩하는 동작을 나타낸다. 도 2를 참조하면, 영상 복원 장치는 개별 영상(201)을 인코딩하여 특징 표현(220)을 생성할 수 있다. 개별 영상(201)의 인코딩에는 인코딩 모델이 이용될 수 있다. 개별 영상(201)이 인코딩되는 동안, 개별 영상(201)은 다양한 뉴럴 네트워크 레이어(예: 적어도 하나의 컨볼루션 레이어)를 통과할 수 있다.

[0025] 예를 들어, 인코딩 모델의 제1 레이어 및 제2 레이어는 각각 컨볼루션 레이어에 해당할 수 있다. 제1 레이어는 개별 영상(201)에 따른 컨볼루션 연산을 수행하여 제1 중간 특징 표현(211)을 생성할 수 있고, 제2 레이어는 제1 중간 특징 표현(211)에 따른 컨볼루션 연산을 수행하여 제2 중간 특징 표현(212)을 생성할 수 있다. 제3 레이어는 풀링(pooling) 레이어에 해당할 수 있다. 제3 레이어는 제2 중간 특징 표현(212)에 따른 풀링 연산을 수행하여 제3 중간 특징 표현(213)을 생성할 수 있다. 다음 레이어들은 중간 특징 표현들(213 내지 219)에 관해 컨볼루션 연산 및 풀링 연산을 교차로 수행할 수 있고, 그 결과로 특징 표현(220)이 생성될 수 있다.

[0026] 아래에서 데이터의 사이즈는 높이(height)*폭(width)*채널(channel)의 형식으로 표현될 수 있다. 예를 들어, 개별 영상(201)의 사이즈는 H*W*4일 수 있다. 중간 특징 표현들(211, 212)의 채널 사이즈는 32, 중간 특징 표현들(213, 214)의 채널 사이즈는 64, 중간 특징 표현들(215, 216)의 채널 사이즈는 128, 중간 특징 표현들(217, 218)의 채널 사이즈는 256일 수 있다. 중간 특징 표현(219) 및 특징 표현(220)의 사이즈는 H'*W'*512일 수 있다. 다만, 이러한 인코딩 모델의 레이어 배치, 및 데이터 사이즈 하나의 예시에 해당하며, 다른 다양한 레이어 배치 및 데이터 사이즈가 이용될 수 있다. 연사 영상 세트의 개별 영상들에 대해 인코딩 모델이 반복적으로 실행되어 개별 영상들에 대응하는 특징 표현들이 생성될 수 있다.

[0027] 도 3은 일 실시예에 따른 특징 표현들의 모션-임베딩 특징 표현들을 생성하고, 및 모션-임베딩 특징 표현들을 융합하는 동작을 나타낸다. 도 3을 참조하면, 영상 복원 장치는 특징 표현들(I_1 내지 I_N) 중 레퍼런스 특징 표현(I_R)을 결정하고, 레퍼런스 특징 표현(I_R), 및 특징 표현들(I_1 내지 I_N) 각각을 포함하는 비교 쌍들을 결정할 수 있다. 레퍼런스 특징 표현(I_R), 및 특징 표현들(I_1 내지 I_N)의 사이즈는 H'*W'*c일 수 있다. 연사 영상 세트의 개별 영상들의 수, 특징 표현들(I_1 내지 I_N)의 수, 비교 쌍들의 수는 모두 N개일 수 있다. 예를 들어, 제1 비교 쌍은 레퍼런스 특징 표현(I_R) 및 제1 특징 표현(I_1)을 포함할 수 있고, 제2 비교 쌍은 레퍼런스 특징 표현(I_R) 및 제2 특징 표현(I_2)을 포함할 수 있다. 제1 특징 표현(I_1)은 제1 개별 영상에 대응할 수 있고, 제2 특징 표현(I_2)은 제2 개별 영상에 대응할 수 있다.

[0028] 영상 복원 장치는 비교 쌍들에 기초하여 모션-지각 모델(310)을 실행하여 모션-임베딩 특징 표현들(311, 311N)을 생성할 수 있다. 모션-지각 모델(310)은 각 비교 쌍의 특징 표현들의 모션 정보에 기초하여 해당 특징 표현들 간의 정보 교환을 수행할 수 있다. 모션-지각 모델(310)은 비교 쌍들의 수만큼 반복적으로 실행될 수 있고, 비교 쌍들의 수에 대응하는 모션-임베딩 특징 표현들(311, 311N)이 생성될 수 있다. 모션-지각 모델(310)의 구체적인 동작은 추후 상세히 설명한다. 영상 복원 장치는 모션-임베딩 특징 표현들(311, 311N)을 융합하여 융합 결과(320)를 생성할 수 있다. 융합 결과의 사이즈는 H*W*C일 수 있다. 예를 들어, 영상 복원 장치는 엘리먼트 별 덧셈 연산(elementwise addition)을 수행하여 융합 결과(320)를 생성할 수 있다.

[0029] 도 4는 일 실시예에 따른 모션-임베딩 특징 표현들을 생성하는 동작을 나타낸다. 도 4를 참조하면, 영상 복원 장치는 레퍼런스 특징 표현(410) 및 특징 표현(420)을 포함하는 비교 쌍의 모션-임베딩 특징 표현을 생성할 수 있다. 예를 들어, 특징 표현(420)은 제1 특징 표현일 수 있고, 비교 쌍은 제1 비교 쌍일 수 있다. 도 4의 동작은 모션-지각 모델을 통해 수행될 수 있다.

[0030] 영상 복원 장치는 레퍼런스 특징 표현(410)의 지점 (i, j)의 특징 벡터(411)를 결정하고, 및 특징 표현(420)에서 지점 (i, j)을 중심으로 갖는 관심 블록(421)을 결정할 수 있다. 레퍼런스 특징 표현(410) 및 특징 표현(420)은 H'*W'*c의 사이즈를 가질 수 있고, 레퍼런스 특징 표현(410)의 지점 (i, j)과 특징 표현(420)의 지점

(i, j)는 서로 대응할 수 있다. 이들은 대응 지점으로 부를 수 있다.

- [0031] 영상 복원 장치는 특징 벡터(411)와 관심 블록(421)의 채널 벡터들 간의 유사도에 기초하여 지점 (i, j)에 대응하는 유사도 스코어 맵(430)을 생성할 수 있다. 예를 들어, 특징 벡터(411)의 사이즈는 $1 \times 1 \times c$, 관심 블록(421)의 사이즈는 $k \times k \times c$, 유사도 스코어 맵(430)의 사이즈는 $k \times k \times 1$ 일 수 있다. $k \times k$ 는 $H \times W$ 보다 상당히 작은 값에 해당할 수 있다. 연사 영상은 비교적 짧은 시간에 연속적으로 생성되므로, 모션에 의한 영상 픽셀의 변화 영역은 제한적일 수 있다. 유사도 스코어 맵(430)의 사이즈가 $k \times k$ 로 제한됨에 따라 모션 분석에 필요한 연산 효율이 향상될 수 있다. 관심 블록(421)은 $k \times k$ 개의 채널 벡터들을 포함할 수 있고, 각 채널 벡터의 사이즈는 $1 \times 1 \times c$ 일 수 있다. 유사도 스코어 맵(430)은 $k \times k$ 개의 스코어 엘리먼트들을 포함할 수 있고, 각 스코어 엘리먼트는 특징 벡터(411)와 관심 블록(421)의 채널 벡터들 간의 유사도를 나타낼 수 있다. 예를 들어, 유사도는 코사인 유사도(cosine similarity)에 해당하거나, 혹은 네트워크 기반의 상관관계(correlation)에 해당할 수 있다. 후자의 경우 이러한 상관관계를 계산하기 위한 별도의 뉴럴 네트워크가 추가될 수 있다.
- [0032] 영상 복원 장치는 유사도 스코어 맵(430)에 기초하여 지점 (i, j)에 대응하는 모션-지각 커널(462)을 생성할 수 있다. 영상 복원 장치는 관심 블록(421)과 유사도 스코어 맵(430) 간의 연산(440)에 기초하여 컨텍스트 정보 벡터(441)를 생성할 수 있다. 예를 들어, 연산(440)은 점곱(dot product)에 해당할 수 있다. 컨텍스트 정보 벡터(441)의 사이즈는 $1 \times 1 \times c$ 일 수 있다. 영상 복원 장치는 유사도 스코어 맵(430)의 벡터화(450)에 기초하여 유사도 스코어 맵(430)에 대응하는 모션 정보 벡터(451)를 결정할 수 있다. 예를 들어, 벡터화(450)는 평탄화(flatten)에 해당할 수 있다. 모션 정보 벡터(451)의 사이즈는 $1 \times 1 \times k^2$ 일 수 있다. 영상 복원 장치는 모션 정보 벡터(451)와 컨텍스트 정보 벡터(441)를 융합하여 융합 결과(461)를 생성할 수 있고, 융합 결과(461)에 기초하여 모션-지각 커널(462)을 결정할 수 있다. 융합 결과(461)의 사이즈는 $1 \times 1 \times k^2$ 일 수 있고, 모션-지각 커널(462)의 사이즈는 $k \times k \times 1$ 일 수 있다.
- [0033] 영상 복원 장치는 관심 블록(421)과 모션-지각 커널(462) 간의 연산에 기초하여 지점(i, j)에 대응하는 모션-임베딩 벡터(470)를 생성할 수 있다. 예를 들어, 해당 연산은 컨볼루션 연산일 수 있다. 모션-임베딩 벡터(470)의 사이즈는 $1 \times 1 \times c$ 일 수 있다. 영상 복원 장치는 지점(i, j)을 변경해가며 이러한 일련의 동작을 $H \times W \times c$ 만큼 반복적으로 수행하여 $H \times W \times c$ 의 모션 임베딩 벡터들을 갖는 모션 특징 표현을 생성할 수 있다. 예를 들어, 모션-임베딩 벡터(470)가 제1 지점의 제1 모션-임베딩 벡터에 해당하는 경우, 영상 복원 장치는 제2 지점의 제2 모션 임베딩 벡터를 생성할 수 있고, 제1 모션-임베딩 벡터 및 제2 모션-임베딩 벡터를 포함하는 모션-임베딩 벡터들에 기초하여 모션-임베딩 특징 표현을 생성할 수 있다. 레퍼런스 특징 표현(410) 및 특징 표현(420)이 제1 특징 쌍을 구성하는 경우, 해당 모션-임베딩 특징 표현은 제1 특징 쌍에 대응할 수 있다.
- [0034] 도 5는 일 실시예에 따른 모션 정보 벡터와 컨텍스트 정보 벡터를 융합하는 동작을 나타낸다. 도 5를 참조하면, 완전 연결 네트워크(520)를 이용하여 컨텍스트 정보 벡터(511)에 대응하는 채널 벡터(521)가 생성될 수 있고, 완전 연결 네트워크(540)를 이용하여 모션 정보 벡터(531)에 대응하는 채널 벡터(541)가 생성될 수 있다. 컨텍스트 정보 벡터(511)와 모션 정보 벡터(531)는 서로 다른 사이즈를 가질 수 있다. 완전 연결 네트워크들(520, 540)은 컨텍스트 정보 벡터(511)와 모션 정보 벡터(531)의 사이즈를 조절할 수 있다. 완전 연결 네트워크들(520, 540)에 의해 채널 벡터(521)와 채널 벡터(541)는 서로 같은 사이즈를 가질 수 있다. 예를 들어, 컨텍스트 정보 벡터(511)의 사이즈는 $1 \times 1 \times c$ 일 수 있고, 모션 정보 벡터(531)의 사이즈는 $1 \times 1 \times k^2$ 일 수 있고, 채널 벡터들(521, 541)의 사이즈는 $1 \times 1 \times c$ 일 수 있다. 완전 연결 네트워크들(520, 540)은 각각 하나 이상의 완전 연결 레이어를 포함할 수 있다.
- [0035] 채널 벡터들(521, 541)은 서로 융합되어, 완전 연결 네트워크(550)에 입력될 수 있다. 예를 들어, 채널 벡터들(521, 541)은 엘리먼트 별 덧셈 연산을 통해 융합될 수 있다. 완전 연결 네트워크(550)의 출력 벡터(551)는 도 4의 융합 결과(461)에 대응할 수 있다. 완전 연결 네트워크(550)는 채널 벡터들(521, 541)의 사이즈를 조절할 수 있다. 예를 들어, 출력 벡터(551)의 사이즈는 $1 \times 1 \times k^2$ 일 수 있다. 완전 연결 네트워크(550)는 하나 이상의 완전 연결 레이어를 포함할 수 있다. 출력 벡터(551)는 모션-지각 커널로 변환될 수 있다.
- [0036] 컨텍스트 정보 벡터(511)와 모션 정보 벡터(531)는 도 5의 동작 이외에도 다양한 방식으로 융합될 수 있다. 또한 완전 연결 네트워크(520, 540, 550)들 이외에 다른 뉴럴 네트워크가 이용될 수 있고, 나아가 도 5와 다른 뉴럴 네트워크 구조가 이용되는 것도 가능하다.
- [0037] 도 6은 일 실시예에 따른 특징 표현들을 복원 영상으로 디코딩하는 동작을 나타낸다. 도 6을 참조하면, 영상

복원 장치는 특징 표현(612)을 디코딩하여 복원 영상(627)을 생성할 수 있다. 특징 표현(612)은 모션-임베딩 특징 표현들의 융합 결과(예: 융합 결과(320))에 대응할 수 있다. 특징 표현(612)의 디코딩에는 뉴럴 네트워크 기반의 디코딩 모델이 이용될 수 있다. 특징 표현(612)이 디코딩되는 동안, 특징 표현(612)은 다양한 뉴럴 네트워크 레이어(예: 적어도 하나의 업-컨볼루션 레이어)를 통과할 수 있다.

[0038] 특징 표현들(611, 612)이 결합되어 특징 표현(613)이 결정될 수 있다. 예를 들어, 특징 표현들(611, 612)은 연쇄화(concatenation)를 통해 결합될 수 있다. 특징 표현(611)은 스킵-커넥션(skip-connection)을 통해 인코딩 모델(601)으로부터 수신될 수 있다. 예를 들어, 특징 표현(611)은 도 2의 특징 표현(218)에 대응할 수 있다. 특징 표현(614)은 특징 표현(616)으로 변환될 수 있다. 예를 들어, 특징 표현(613)에 따른 컨볼루션 연산에 기초하여 특징 표현(614)이 결정될 수 있다. 특징 표현들(611, 612)의 채널 사이즈는 256, 특징 표현(613)의 채널 사이즈는 512, 특징 표현(614)의 채널 사이즈는 256일 수 있다.

[0039] 특징 표현(614)에 따른 업-컨볼루션에 기초하여 특징 표현(614)이 결정될 수 있다. 특징 표현(615)은 스킵-커넥션을 통해 인코딩 모델(601)으로부터 수신될 수 있다. 예를 들어, 특징 표현(615)은 도 2의 특징 표현(216)에 대응할 수 있다. 특징 표현들(615, 616)이 결합되어 특징 표현(617)이 결정될 수 있다. 예를 들어, 특징 표현들(615, 616)은 연쇄화를 통해 결합될 수 있다. 특징 표현(617)은 특징 표현(618)으로 변환될 수 있다. 예를 들어, 특징 표현(617)에 따른 컨볼루션 연산에 기초하여 특징 표현(618)이 결정될 수 있다. 특징 표현들(615, 616)의 채널 사이즈는 128, 특징 표현(617)의 채널 사이즈는 256, 특징 표현(618)의 채널 사이즈는 128일 수 있다.

[0040] 특징 표현들(619 내지 626)에도 특징 표현들(615 내지 618)에 적용된 동작의 대응 동작이 적용될 수 있다. 특징 표현들(619, 620)의 채널 사이즈는 64, 특징 표현(621)의 채널 사이즈는 128, 특징 표현(622)의 채널 사이즈는 64, 특징 표현들(623, 624)의 채널 사이즈는 32, 특징 표현(625)의 채널 사이즈는 64, 특징 표현(626)의 채널 사이즈는 32일 수 있다. 특징 표현(626)에 기초하여 복원 영상(627)이 결정될 수 있다. 복원 영상(627)은 개별 영상과 동일한 사이즈를 가질 수 있다. 예를 들어, 복원 영상(627)의 사이즈는 H*W*4일 수 있다.

[0041] 도 7은 일 실시예에 따른 데이터 흐름을 나타낸다. 도 7을 참조하면, 연사 영상 세트(710)는 개별 영상들(701, 702)을 포함할 수 있다. 인코딩 모델(720)은 개별 영상들(701, 702)을 인코딩하여 개별 영상들(701, 702)에 대응하는 복수의 특징 표현들을 생성할 수 있다. 예를 들어, 인코딩 모델(720)은 제1 개별 영상(701)에 대응하는 제1 특징 표현 및 제2 개별 영상(702)에 대응하는 제2 특징 표현을 생성할 수 있다. 예를 들어, 제1 특징 표현 및 제2 특징 표현은 각각 도 2의 특징 표현(220)에 대응할 수 있다.

[0042] 정보 교환 모델(730)은 특징 표현들의 비교 쌍의 정보 교환을 수행하여 모션-임베딩 특징 표현의 융합 결과를 생성할 수 있다. 영상 복원 장치는 특징 표현들 중에 레퍼런스 특징 표현을 결정할 수 있고, 각각 레퍼런스 특징 표현과 특징 표현들 중 하나를 포함하는 비교 쌍들을 결정할 수 있다. 예를 들어, 제1 특징 표현 내지 제N 특징 표현을 포함하는 N개의 특징 표현들이 존재하고, N개의 특징 표현들 중에 레퍼런스 특징 표현이 선택되는 경우, 레퍼런스 특징 표현과 제1 특징 표현을 포함하는 제1 비교 쌍 내지 레퍼런스 특징 표현과 제N 특징 표현을 포함하는 제N 비교 쌍을 포함하는 N개의 비교 쌍들이 결정될 수 있다.

[0043] 레퍼런스 특징 표현이 결정되면, 레퍼런스 특징 표현을 중심으로 영상 복원이 수행될 수 있다. 일 실시예에 따르면, 영상 복원 장치는 개별 영상들(701, 702) 중에 하나의 앵커 영상(anchor image)을 선택하고, 특징 표현들 중 앵커 영상의 특징 표현을 레퍼런스 특징 표현으로 결정할 수 있다. 이 경우, 하나의 레퍼런스 특징에 따라 영상 복원이 수행되고, 하나의 복원 영상(750)이 생성될 수 있다. 영상 복원 장치는 품질 기반으로 개별 영상들 중에 앵커 영상을 선택할 수 있다. 보다 구체적으로, 영상 복원 장치는 개별 영상들(701, 702)의 품질을 평가할 수 있고, 개별 영상들(701, 702) 중 가장 높은 품질을 갖는 하나를 앵커 영상으로 선택할 수 있다.

[0044] 다른 일 실시예에 따르면, 개별 영상들(701, 702)이 순차적으로 앵커 영상으로 선택될 수 있다. 이 경우, 개별 영상들(701, 702)에 대응하는 N개의 특징 표현들이 순차적으로 레퍼런스 특징 표현으로 이용될 수 있고, N개의 복원 영상들이 생성될 수 있다. 예를 들어, 제1 개별 영상(701)을 제1 복원 영상으로 복원하는 경우, 특징 표현들 중 제1 개별 영상의 특징 표현이 레퍼런스 특징 표현으로 결정될 수 있고, 해당 레퍼런스 특징 표현을 이용하여 모션-임베딩 특징 표현들이 생성될 수 있다. 또한, 제2 개별 영상(702)을 제2 복원 영상으로 복원하는 경우, 특징 표현들 중 제2 개별 영상의 특징 표현이 레퍼런스 특징 표현으로 결정될 수 있고, 해당 레퍼런스 특징 표현을 이용하여 모션-임베딩 특징 표현들이 생성될 수 있다.

[0045] 디코딩 모델(740)은 모션-임베딩 특징 표현의 융합 결과를 디코딩하여 적어도 하나의 복원 영상(750)을 생성할

수 있다. 인코딩 모델(720)로부터 디코딩 모델(740)로의 화살표들은 스킵-커넥션을 나타낼 수 있고, 디코딩 모델(740)의 점선 화살표는 업-컨볼루션을 나타낼 수 있고, 디코딩 모델(740) 내의 실선 화살표는 융합을 나타낼 수 있다. 예를 들어, 융합은 연쇄화에 해당할 수 있다.

[0046] 도 8은 일 실시예에 따른 영상 복원 방법을 나타낸다. 도 8을 참조하면, 단계(810)에서 영상 복원 장치는 연사 영상 세트의 개별 영상들을 인코딩하여 상기 개별 영상들에 대응하는 특징 표현들을 생성한다. 영상 복원 장치는 적어도 하나의 컨볼루션 레이어를 포함하는 인코딩 모델을 이용하여 개별 영상들을 인코딩할 수 있다.

[0047] 단계(820)에서 영상 복원 장치는 복수의 특징 표현들 중 레퍼런스 특징 표현을 결정한다. 영상 복원 장치는 개별 영상들 중에 앵커 영상을 선택할 수 있고, 특징 표현들 중 앵커 영상의 특징 표현을 레퍼런스 특징 표현으로 결정할 수 있다. 영상 복원 장치는 품질 기반으로 개별 영상들 중에 앵커 영상을 선택할 수 있다. 개별 영상들은 제1 개별 영상 및 제2 개별 영상을 포함할 수 있고, 적어도 하나의 복원 영상은 제1 개별 영상의 제1 복원 영상 및 제2 개별 영상의 제2 복원 영상을 포함할 수 있다. 이때, 영상 복원 장치는 제1 개별 영상을 제1 복원 영상으로 복원하는 경우 특징 표현들 중 제1 개별 영상의 특징 표현을 레퍼런스 특징 표현으로 결정할 수 있고, 제2 개별 영상을 제2 복원 영상으로 복원하는 경우 특징 표현들 중 제2 개별 영상의 특징 표현을 레퍼런스 특징 표현으로 결정할 수 있다.

[0048] 단계(830)에서 영상 복원 장치는 레퍼런스 특징 표현, 및 복수의 대상 특징 표현들 중 제1 대상 특징 표현을 포함하는 제1 비교 쌍을 결정한다. 단계(840)에서 영상 복원 장치는 레퍼런스 특징 표현과 제1 대상 특징 표현의 유사도 스코어 맵에 기초하여 제1 비교 쌍의 제1 모션-임베딩 특징 표현을 생성한다.

[0049] 영상 복원 장치는 레퍼런스 특징 표현의 제1 지점의 제1 특징 벡터, 및 제1 특징 표현에서 제1 지점의 대응 지점을 중심으로 갖는 제1 관심 블록을 결정하고, 제1 특징 벡터와 제1 관심 블록의 채널 벡터들 간의 유사도에 기초하여 제1 지점에 대응하는 제1 유사도 스코어 맵을 생성하고, 제1 유사도 스코어 맵에 기초하여 제1 지점에 대응하는 제1 모션-지각 커널을 생성하고, 제1 관심 블록과 제1 모션-지각 커널 간의 연산에 기초하여 제1 지점에 대응하는 제1 모션-임베딩 벡터를 생성하고, 제1 모션-임베딩 벡터를 포함하는 모션-임베딩 벡터들에 기초하여 제1 모션-임베딩 특징 표현을 생성할 수 있다.

[0050] 영상 복원 장치는 레퍼런스 특징 표현의 제2 지점의 제2 특징 벡터, 및 제1 특징 표현에서 제2 지점의 대응 지점을 중심으로 갖는 제2 관심 블록을 결정하고, 제2 관심 블록에 기초하여 제2 지점에 대응하는 제2 모션-임베딩 벡터를 생성할 수 있다. 영상 복원 장치는 제1 모션-임베딩 벡터 및 제2 모션-임베딩 벡터를 포함하는 모션-임베딩 벡터들에 기초하여 제1 모션-임베딩 특징 표현을 생성할 수 있다.

[0051] 영상 복원 장치는 제1 유사도 스코어 맵에 대응하는 제1 모션 정보 벡터를 결정하고, 제1 관심 블록과 제1 유사도 스코어 맵 간의 연산에 기초하여 제1 컨텍스트 정보 벡터를 생성하고, 제1 모션 정보 벡터와 제1 컨텍스트 정보 벡터를 융합하여 제1 모션-지각 커널을 생성할 수 있다. 영상 복원 장치는 제1 완전 연결 네트워크를 이용하여 제1 모션 정보 벡터에 대응하는 제1 채널 벡터를 생성하고, 제2 완전 연결 네트워크를 이용하여 제1 컨텍스트 정보 벡터에 대응하는 제2 채널 벡터를 생성하고, 제1 채널 벡터와 제2 채널 벡터를 융합하여 제1 모션-지각 커널을 생성할 수 있다.

[0052] 단계(850)에서 영상 복원 장치는 제1 모션-임베딩 특징 표현을 포함하는 복수의 모션-임베딩 특징 표현들을 융합하여 융합 결과를 생성한다. 단계(860)에서 영상 복원 장치는 융합 결과를 디코딩하여 적어도 하나의 복원 영상을 생성한다. 영상 복원 장치는 적어도 하나의 업-컨볼루션 레이어를 포함하는 디코딩 모델을 이용하여 융합 결과를 디코딩할 수 있다.

[0053] 도 9는 일 실시예에 따른 영상 복원 장치의 구성을 나타낸다. 도 9를 참조하면, 영상 복원 장치(900)는 프로세서(910) 및 메모리(920)를 포함한다. 메모리(920)는 프로세서(910)에 연결되고, 프로세서(910)에 의해 실행 가능한 명령어들, 프로세서(910)가 연산할 데이터 또는 프로세서(910)에 의해 처리된 데이터를 저장할 수 있다. 메모리(920)는 비일시적인 컴퓨터 판독가능 매체, 예컨대 고속 랜덤 액세스 메모리 및/또는 비휘발성 컴퓨터 판독가능 저장 매체(예컨대, 하나 이상의 디스크 저장 장치, 플래시 메모리 장치, 또는 기타 비휘발성 솔리드 스테이트 메모리 장치)를 포함할 수 있다.

[0054] 프로세서(910)는 도 1 내지 도 8, 및 도 10의 동작을 수행하기 위한 명령어들을 실행할 수 있다. 예를 들어, 프로세서(910)는 연사 영상 세트의 개별 영상들을 인코딩하여 개별 영상들에 대응하는 복수의 특징 표현들을 생성하고, 복수의 특징 표현들 중 레퍼런스 특징 표현을 결정하고, 레퍼런스 특징 표현, 및 복수의 특징 표현들 중 제1 특징 표현을 포함하는 제1 비교 쌍을 결정하고, 레퍼런스 특징 표현과 제1 특징 표현의 유사도 스코어

맵에 기초하여 제1 비교 쌍의 제1 모션-임베딩 특징 표현을 생성하고, 제1 모션-임베딩 특징 표현을 포함하는 복수의 모션-임베딩 특징 표현들을 융합하여 융합 결과를 생성하고, 융합 결과를 디코딩하여 적어도 하나의 복원 영상을 생성할 수 있다. 그 밖에, 영상 복원 장치(900)에는 도 1 내지 도 8, 및 도 10의 설명이 적용될 수 있다.

[0055] 도 10은 일 실시예에 따른 전자 장치의 구성을 나타낸다. 도 10를 참조하면, 전자 장치(1000)는 프로세서(1010), 메모리(1020), 카메라(1030), 저장 장치(1040), 입력 장치(1050), 출력 장치(1060) 및 네트워크 인터페이스(1070)를 포함할 수 있으며, 이들은 통신 버스(1080)를 통해 서로 통신할 수 있다. 예를 들어, 전자 장치(1000)는 이동 전화, 스마트 폰, PDA, 넷북, 태블릿 컴퓨터, 랩톱 컴퓨터 등과 같은 모바일 장치, 스마트 워치, 스마트 밴드, 스마트 안경 등과 같은 웨어러블 디바이스, 데스크탑, 서버 등과 같은 컴퓨팅 장치, 텔레비전, 스마트 텔레비전, 냉장고 등과 같은 가전 제품, 도어 락 등과 같은 보안 장치, 자율주행 차량, 스마트 차량 등과 같은 차량의 적어도 일부로 구현될 수 있다. 전자 장치(1000)는 도 1의 영상 복원 장치(100) 및/또는 도 9의 영상 복원 장치(900)를 구조적 및/또는 기능적으로 포함할 수 있다.

[0056] 프로세서(1010)는 전자 장치(1000) 내에서 실행하기 위한 기능 및 명령어들을 실행한다. 예를 들어, 프로세서(1010)는 메모리(1020) 또는 저장 장치(1040)에 저장된 명령어들을 처리할 수 있다. 프로세서(1010)는 도 1 내지 도 9를 통하여 설명된 하나 이상의 동작을 수행할 수 있다. 메모리(1020)는 컴퓨터 판독가능한 저장 매체 또는 컴퓨터 판독가능한 저장 장치를 포함할 수 있다. 메모리(1020)는 프로세서(1010)에 의해 실행하기 위한 명령어들을 저장할 수 있고, 전자 장치(1000)에 의해 소프트웨어 및/또는 애플리케이션이 실행되는 동안 관련 정보를 저장할 수 있다.

[0057] 카메라(1030)는 사진 및/또는 비디오를 촬영할 수 있다. 카메라(1030)는 사진을 연속적으로 촬영하거나, 비디오를 촬영하여 연사 영상 세트를 생성할 수 있다. 연사 영상 세트가 연속적인 사진인 경우 연사 영상 세트의 각 개별 영상은 각 사진에 해당할 수 있고, 연사 영상 세트가 비디오인 경우 연사 영상 세트의 각 개별 영상은 비디오의 각 영상 프레임에 해당할 수 있다. 저장 장치(1040)는 컴퓨터 판독가능한 저장 매체 또는 컴퓨터 판독가능한 저장 장치를 포함한다. 저장 장치(1040)는 메모리(1020)보다 더 많은 양의 정보를 저장하고, 정보를 장기간 저장할 수 있다. 예를 들어, 저장 장치(1040)는 자기 하드 디스크, 광 디스크, 플래시 메모리, 플로피 디스크 또는 이 기술 분야에서 알려진 다른 형태의 비휘발성 메모리를 포함할 수 있다.

[0058] 입력 장치(1050)는 키보드 및 마우스를 통한 전통적인 입력 방식, 및 터치 입력, 음성 입력, 및 이미지 입력과 같은 새로운 입력 방식을 통해 사용자로부터 입력을 수신할 수 있다. 예를 들어, 입력 장치(1050)는 키보드, 마우스, 터치 스크린, 마이크로폰, 또는 사용자로부터 입력을 검출하고, 검출된 입력을 전자 장치(1000)에 전달할 수 있는 임의의 다른 장치를 포함할 수 있다. 출력 장치(1060)는 시각적, 청각적 또는 촉각적인 채널을 통해 사용자에게 전자 장치(1000)의 출력을 제공할 수 있다. 출력 장치(1060)는 예를 들어, 디스플레이, 터치 스크린, 스피커, 진동 발생 장치 또는 사용자에게 출력을 제공할 수 있는 임의의 다른 장치를 포함할 수 있다. 네트워크 인터페이스(1070)는 유선 또는 무선 네트워크를 통해 외부 장치와 통신할 수 있다.

[0059] 이상에서 설명된 실시예들은 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치, 방법 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0060] 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, 가상

장치(virtual equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(signal wave)에 영구적으로, 또는 일시적으로 구체화(embodiment)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0061] 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 저장할 수 있으며 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.

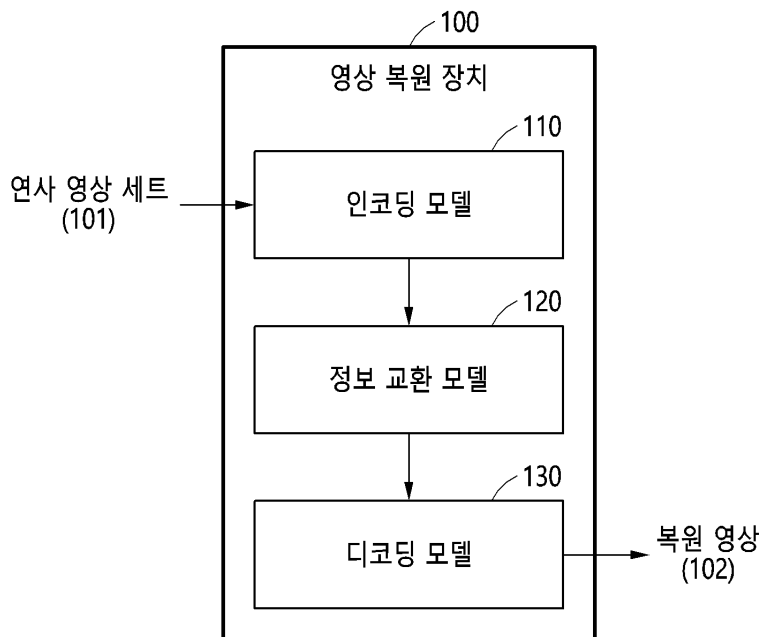
[0062] 위에서 설명한 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 또는 복수의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0063] 이상과 같이 실시예들이 비록 한정된 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 이를 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.

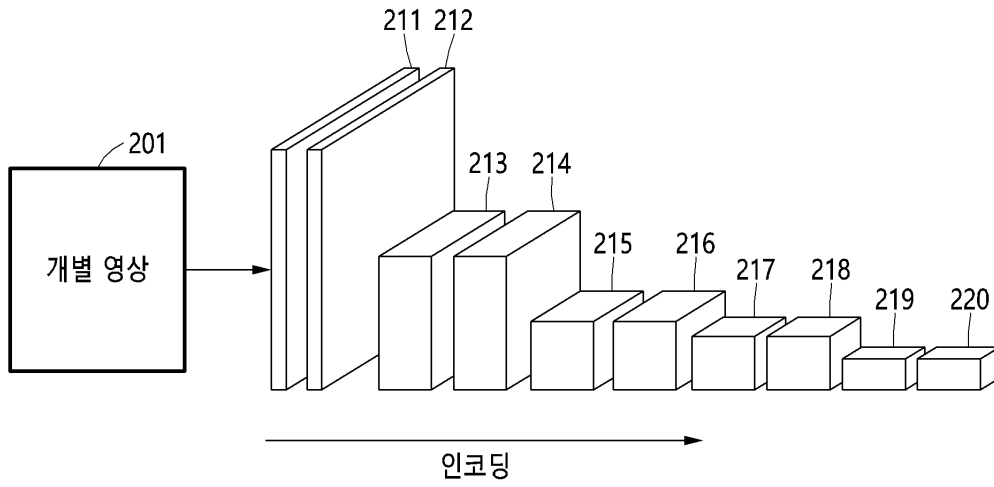
[0064] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

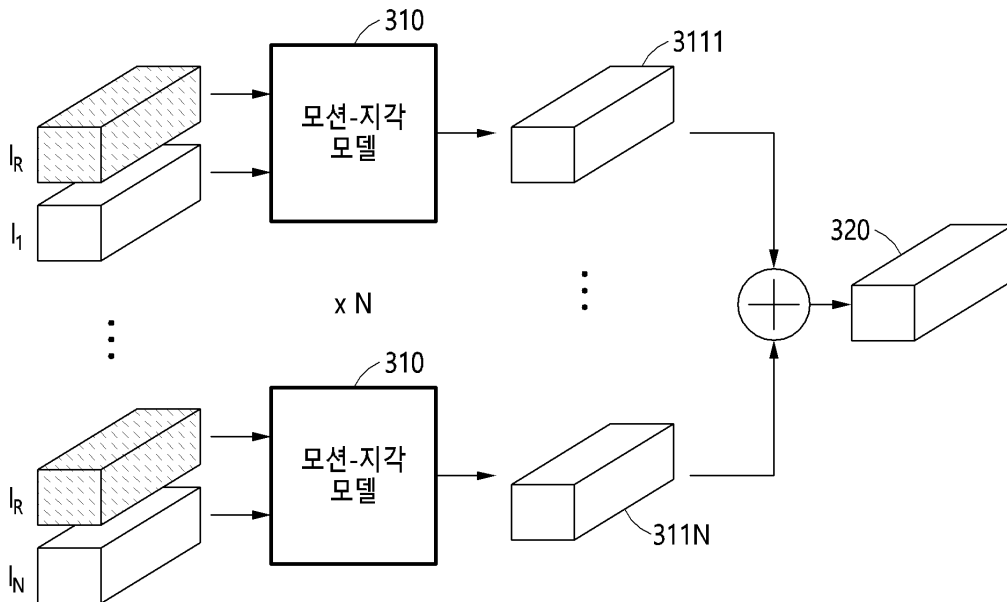
도면1



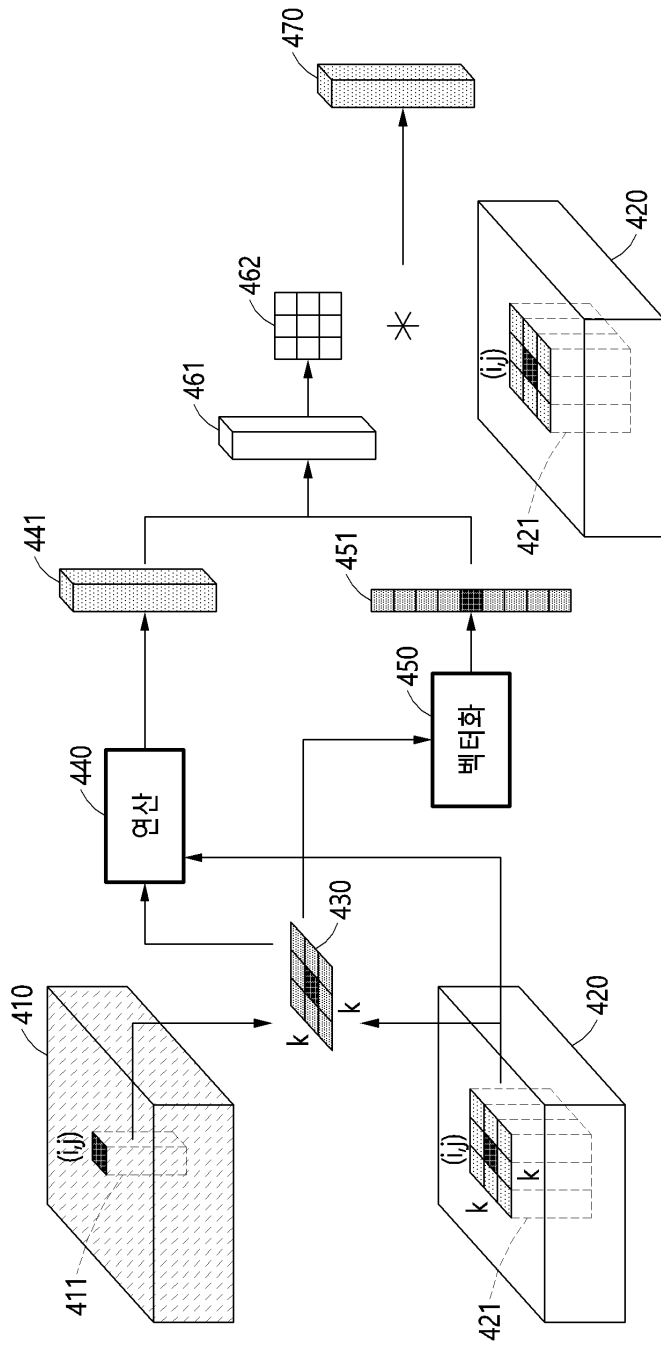
도면2



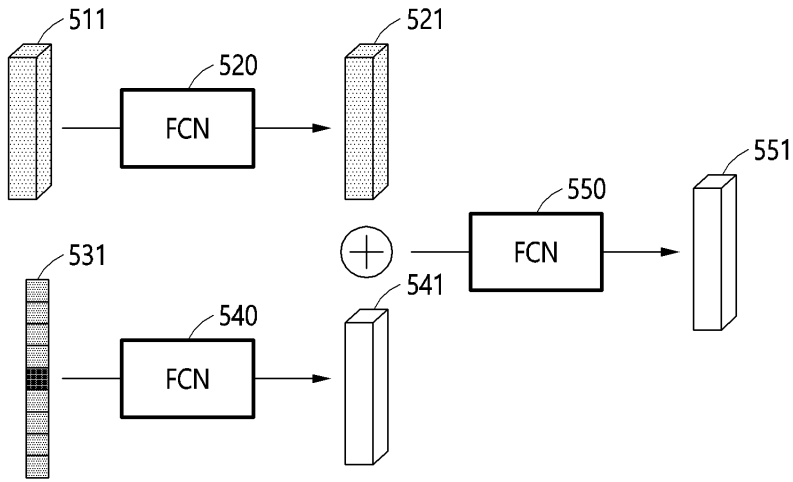
도면3



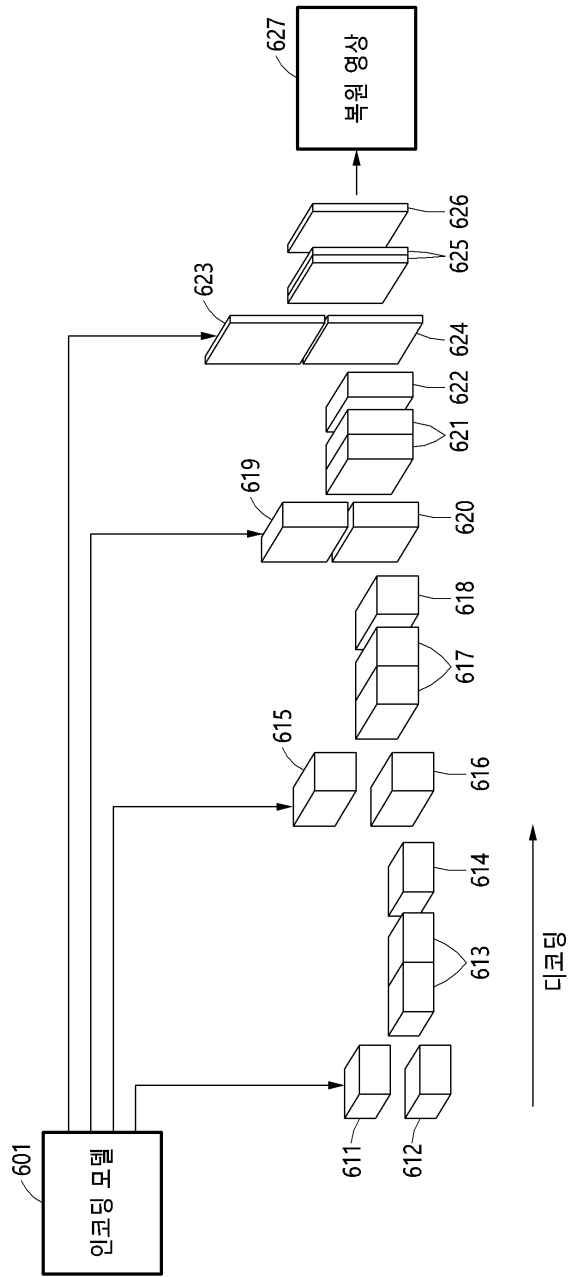
도면4



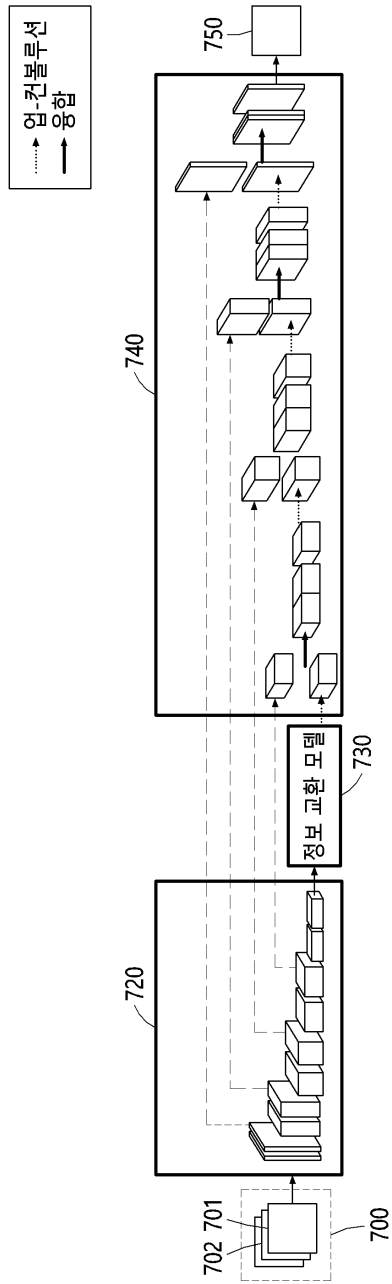
도면5



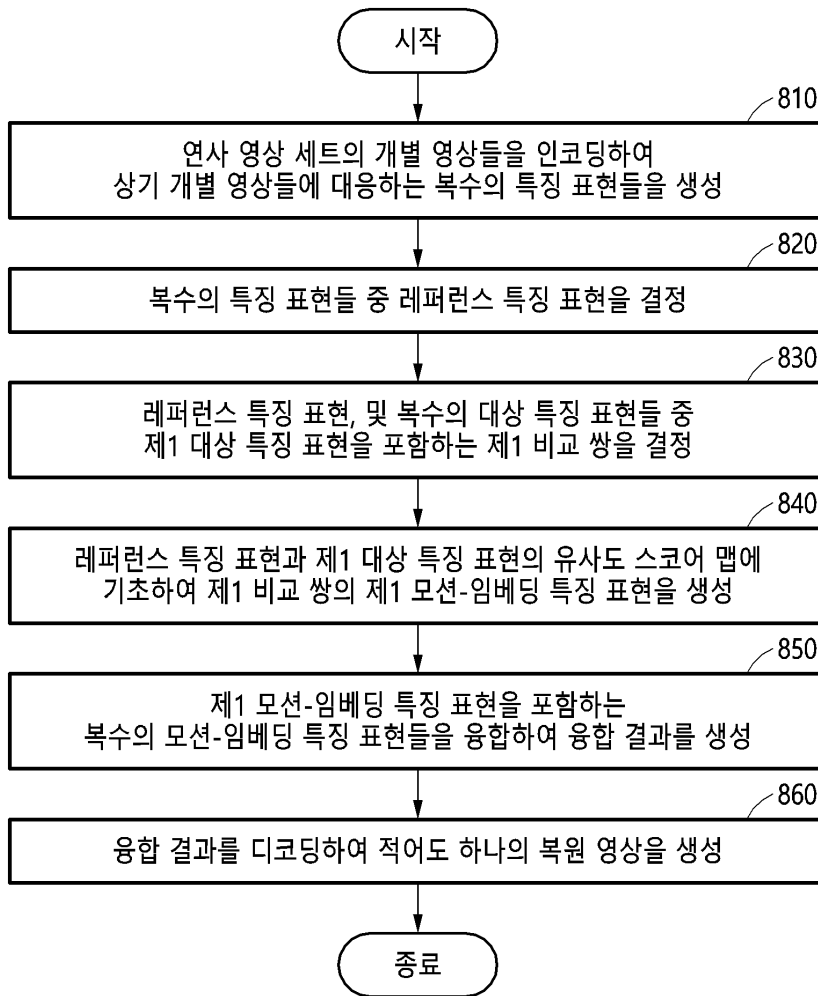
도면6



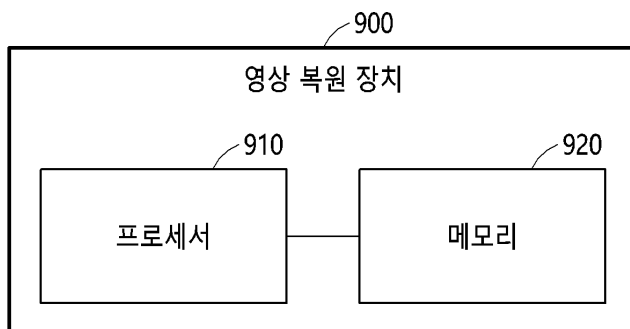
도면7



도면8



도면9



도면10

