(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년03월22일
(11) 등록번호 10-2650764
(24) 등록일자 2024년03월20일

(51) 국제특허분류(Int. Cl.)
G06T 3/40 (2024.01) *G06T 11/60* (2006.01)
G06V 10/46 (2022.01) *G06V 10/74* (2022.01)
G06V 30/19 (2022.01)
(52) CPC특허분류
G06T 3/4053 (2024.01)
G06T 11/60 (2013.01)
(21) 출원번호 10-2022-0173576
(22) 출원일자 2022년12월13일
심사청구일자 2022년12월13일
(56) 선행기술조사문헌
KR1020220059288 A
US20190378242 A1

(73) 특허권자
한국기술교육대학교 산학협력단
충청남도 천안시 동남구 병천면 충절로 1600 (한국기술교육대학교내)
(72) 발명자
김덕수
세종특별자치시 남세종로 302, 610동 506호(소담동, 새샘마을6단지)
최대준
충청남도 천안시 동남구 병천면 충절로 1600
(74) 대리인
특허법인오암

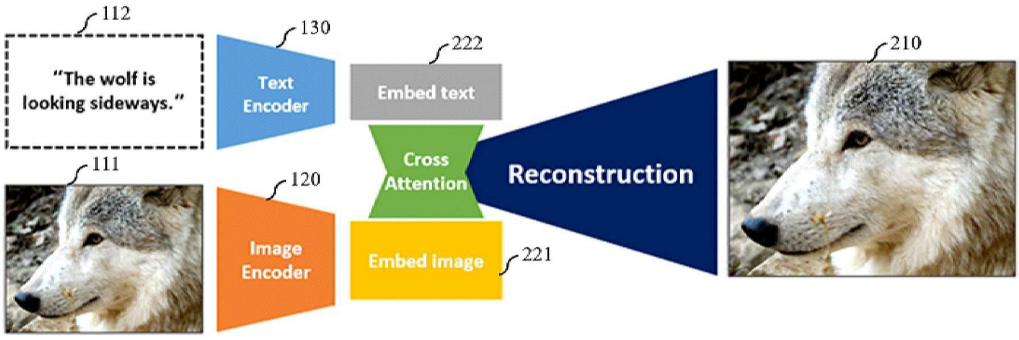
전체 청구항 수 : 총 3 항 심사관 : 이정은

(54) 발명의 명칭 **텍스트를 활용한 이미지 초해상화 방법 및 시스템**

(57) 요 약

일 실시 예에 따른 이미지의 초해상화 방법은, 제1 해상도를 갖는 상기 이미지 및 상기 이미지와 관련된 텍스트를 수신하는 동작, 상기 이미지 및 상기 텍스트를 각각 임베딩 함으로써, 임베딩 이미지 및 임베딩 텍스트를 생성하는 동작, 상기 임베딩 이미지에 포함된 복수의 이미지 벡터로부터 상기 이미지와 관련된 복수의 특징들을 추출하는 동작, 상기 임베딩 텍스트와 상기 임베딩 이미지 간의 유사도를 판단하는 동작 및 상기 유사도 및 상기 복수의 특징들 중 적어도 일부에 기반하여, 상기 제1 해상도보다 높은 제2 해상도를 갖는 고해상도 이미지를 생성하는 동작을 포함할 수 있다.

대 표 도 - 도3



(52) CPC특허분류
G06V 10/469 (2023.08)
G06V 10/761 (2023.08)
G06V 30/19093 (2023.08)

이 발명을 지원한 국가연구개발사업

과제고유번호	1345356194
과제번호	LINC3.0-2022-31
부처명	교육부
과제관리(전문)기관명	한국연구재단
연구사업명	산학연협력고도화지원
연구과제명	3단계 산학연 협력 선도 대학 육성사업(LINC3.0)(0.5)
기 여 율	1/1
과제수행기관명	한국기술교육대학교
연구기간	2022.03.01 ~ 2023.02.28

명 세 서

청구범위

청구항 1

이미지의 초해상화 방법에 있어서,

제1 해상도를 갖는 상기 이미지 및 상기 이미지와 관련된 텍스트를 수신하는 동작;

상기 이미지 및 상기 텍스트를 각각 임베딩 함으로써, 임베딩 이미지 및 임베딩 텍스트를 생성하는 동작;

상기 임베딩 이미지에 포함된 복수의 이미지 벡터로부터 상기 이미지와 관련된 복수의 특징들을 추출하는 동작;

상기 임베딩 텍스트와 상기 임베딩 이미지 간의 유사도를 판단하는 동작; 및

상기 유사도 및 상기 복수의 특징들에 기반하여, 상기 제1 해상도보다 높은 제2 해상도를 갖는 고해상도 이미지를 생성하는 동작을 포함하며,

상기 임베딩 이미지 및 상기 임베딩 텍스트를 생성하는 동작은,

상기 이미지 및 상기 텍스트를 각각 임의의 공간에 투영함으로써, 벡터 값을 갖는 상기 임베딩 이미지 및 상기 임베딩 텍스트를 생성하는 방식으로 수행되고,

상기 이미지 벡터로부터 상기 복수의 특징들을 추출하는 동작은,

상기 임베딩 이미지에 포함된 상기 복수의 이미지 벡터들 간의 유사도를 측정하는 동작, 및 상기 측정된 유사도에 기반하여, 상기 복수의 특징들을 추출하는 동작을 포함하며,

상기 임베딩 텍스트와 상기 임베딩 이미지 간의 유사도를 판단하는 동작은,

상기 임베딩 이미지에 대한 상기 복수의 특징들과 상기 임베딩 텍스트 간의 유사도를 판단하는 방식으로 수행되고,

상기 고해상도 이미지를 생성하는 동작은:

상기 유사도에 기반하여, 상기 복수의 특징들 중 상기 임베딩 텍스트와 임계 값 이상의 유사도를 갖는 특징들을 추출하는 동작, 및 상기 추출된 특징들에 기반하여, 상기 고해상도 이미지를 생성하는 동작을 포함하는 것을 특징으로 하는, 초해상화 방법.

청구항 2

삭제

청구항 3

삭제

청구항 4

삭제

청구항 5

이미지의 초해상화 시스템에 있어서,

제1 해상도를 갖는 상기 이미지를 수신하고, 수신한 상기 이미지를 임베딩 함으로써, 임베딩 이미지를 생성하는 이미지 인코더;

상기 이미지와 관련된 텍스트를 수신하고, 수신한 상기 텍스트를 임베딩 함으로써, 임베딩 텍스트를 생성하는 텍스트 인코더; 및

상기 이미지 인코더 및 상기 텍스트 인코더와 연결되는 프로세서를 포함하고,

상기 이미지 인코더는,

상기 이미지를 임의의 공간에 투영함으로써, 벡터 값을 갖는 상기 임베딩 이미지를 생성하는 방식으로 수행하고, 상기 임베딩 이미지에 포함된 복수의 이미지 벡터들 간의 유사도를 측정하며,

상기 텍스트 인코더는,

상기 텍스트를 임의의 공간에 투영함으로써, 벡터 값을 갖는 상기 임베딩 텍스트를 생성하는 방식으로 수행하고,

상기 프로세서는:

상기 임베딩 이미지에 포함된 상기 복수의 이미지 벡터로부터 상기 이미지와 관련된 복수의 특징들을 추출하는 동작;

상기 임베딩 텍스트와 상기 임베딩 이미지 간의 유사도를 판단하는 동작; 및

상기 유사도 및 상기 복수의 특징들에 기반하여, 상기 제1 해상도보다 높은 제2 해상도를 갖는 고해상도 이미지를 생성하는 동작을 포함하며,

상기 이미지와 관련된 복수의 특징들을 추출하는 동작에서 상기 이미지 인코더를 통해 측정된 상기 유사도에 기반하여, 상기 복수의 특징들을 추출하고,

상기 임베딩 텍스트와 상기 임베딩 이미지 간의 유사도를 판단하는 동작에서 상기 임베딩 이미지에 대한 상기 복수의 특징들과 상기 임베딩 텍스트 간의 유사도를 판단하는 방식으로 수행하며,

상기 고해상도 이미지를 생성하는 동작에서 상기 유사도에 기반하여, 상기 복수의 특징들 중 상기 임베딩 텍스트와 임계 값 이상의 유사도를 갖는 특징들을 추출하는 동작, 및 상기 추출된 특징들에 기반하여, 상기 고해상도 이미지를 생성하는, 초해상화 시스템.

청구항 6

삭제

청구항 7

삭제

청구항 8

삭제

청구항 9

이미지의 초해상화 방법이 실행될 수 있는 컴퓨터 프로그램이 저장된 기록매체에 있어서, 상기 컴퓨터 프로그램의 실행에 의한 동작은,

제1 해상도를 갖는 상기 이미지 및 상기 이미지와 관련된 텍스트를 수신하는 동작;

상기 이미지 및 상기 텍스트를 각각 임베딩 함으로써, 임베딩 이미지 및 임베딩 텍스트를 생성하는 동작;

상기 임베딩 이미지에 포함된 복수의 이미지 벡터로부터 상기 이미지와 관련된 복수의 특징들을 추출하는 동작;

상기 임베딩 텍스트와 상기 임베딩 이미지 간의 유사도를 판단하는 동작; 및

상기 유사도 및 상기 복수의 특징들에 기반하여, 상기 제1 해상도보다 높은 제2 해상도를 갖는 고해상도 이미지를 생성하는 동작을 포함하며,

상기 임베딩 이미지 및 상기 임베딩 텍스트를 생성하는 동작은,

상기 이미지 및 상기 텍스트를 각각 임의의 공간에 투영함으로써, 벡터 값을 갖는 상기 임베딩 이미지 및 상기 임베딩 텍스트를 생성하는 방식으로 수행되고,

상기 이미지 벡터로부터 상기 복수의 특징들을 추출하는 동작은,

상기 임베딩 이미지에 포함된 상기 복수의 이미지 벡터들 간의 유사도를 측정하는 동작, 및 상기 측정된 유사도에 기반하여, 상기 복수의 특징들을 추출하는 동작을 포함하며,

상기 임베딩 텍스트와 상기 임베딩 이미지 간의 유사도를 판단하는 동작은,

상기 임베딩 이미지에 대한 상기 복수의 특징들과 상기 임베딩 텍스트 간의 유사도를 판단하는 방식으로 수행되고,

상기 고해상도 이미지를 생성하는 동작은:

상기 유사도에 기반하여, 상기 복수의 특징들 중 상기 임베딩 텍스트와 임계 값 이상의 유사도를 갖는 특징들을 추출하는 동작, 및 상기 추출된 특징들에 기반하여, 상기 고해상도 이미지를 생성하는 동작을 포함하는 것을 특징으로 하는, 컴퓨터 프로그램이 저장된 기록 매체.

발명의 설명

기 술 분 야

[0001] 본 문서에서 개시되는 다양한 실시 예들은, 텍스트를 활용한 이미지 초해상화 방법 및 시스템에 관한 것이다. 보다 상세하게는, 이미지 및 이미지와 관련된 텍스트에 대한 임베딩 및 유사도 판단을 통해, 입력된 이미지보다 높은 해상도를 갖는 고해상도 이미지를 생성하는 방법 및 시스템에 관한 것이다.

[0002] 본 발명은 "뉴스미디어 품질개선을 위한 딥러닝 기반 초해상화 기술 연구"의 연구결과로 고안된 것이다.

배 경 기 술

[0003] 저해상도 이미지를 고해상도 이미지로 변환함에 있어서, 종래 기법은 보간법을 이용하여 중간 이미지를 생성하고, 선형 매핑을 사용하여 중간 이미지를 고해상도 이미지로 변환하였다. 이러한 종래 기법에서, 보간된 중간 이미지의 패치 수는 저해상도 이미지의 패치 수보다 많아지므로, 고해상도 이미지의 생성에 소요되는 시간이 길었다.

[0004] 한편, 최근 합성곱 신경망(convolutional neural network, CNN) 기반의 초해상화(super resolution) 기법에 관한 연구가 다양하게 발전하면서, 종래의 로컬 패치(local patch) 기반 또는 사전 학습 기반의 초해상화와 대비하여 높은 수준의 초해상화가 가능해지고 있다.

[0005] 합성곱 신경망 기반의 초해상화 기법으로는, 픽셀들의 MSE(Mean Square Error)를 최소화하는 방향으로 인공신경망(artificial neural network)을 학습시키는 PSNR-Oriented 기법, 인지 손실(perceptual loss)을 더 고려하여 인공신경망을 학습시키는 SRGAN(Super Resolution Generative Adversarial Network) 기법 등이 있다. 인지 손실이란 입력 영상과 정답 영상을 다른 딥러닝(deep learning) 기반의 영상 분류 신경망에 입력하여 얻은 피쳐맵(feature map) 간의 차를 최소화하도록 필터 파라미터를 학습시킨다.

[0006] PSNR-Oriented 기법은, 초해상화 성능의 메트릭(metric) 중 하나인 PSNR(Peak-Signal-to-Noise Ration) 값은 향상시키나, 출력 영상과 정답 영상 간의 픽셀별 오차가 줄어들도록 인공신경망을 학습시키므로 출력 영상이 과도하게 편평화(smoothing)되는 문제가 있다. 한편, SRGAN 기법은 인지 손실이 고려되도록 인공신경망을 학습시켜 초해상화 성능을 인지적으로 향상시키나, 출력 영상에 구조적 왜곡이나 잡음을 발생시키는 문제가 있다.

발명의 내용

해결하려는 과제

[0008] 이러한 기존 초해상화 방법들은 이미지만을 이용하여 해상도와 품질을 높여왔다. 그러나, 입력된 이미지에 해당하는 정보들을 활용하여 초해상화 하는 것은 이미지에 매우 많은 도메인이 존재함으로 인하여 각 도메인에 대응하여 품질이 높은 이미지를 생성하는데 한계가 존재한다.

[0009] 본 개시에서 이루고자 하는 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급하지 않은 또다른 기술적 과제들은 아래의 기재로부터 본 개시가 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하

게 이해될 수 있을 것이다.

과제의 해결 수단

[0011] 일 실시 예에 따른 이미지의 초해상화 방법은, 제1 해상도를 갖는 상기 이미지 및 상기 이미지와 관련된 텍스트를 수신하는 동작, 상기 이미지 및 상기 텍스트를 각각 임베딩 함으로써, 임베딩 이미지 및 임베딩 텍스트를 생성하는 동작, 상기 임베딩 이미지에 포함된 복수의 이미지 벡터로부터 상기 이미지와 관련된 복수의 특징들을 추출하는 동작, 상기 임베딩 텍스트와 상기 임베딩 이미지 간의 유사도를 판단하는 동작 및 상기 유사도 및 상기 복수의 특징들 중 적어도 일부에 기반하여, 상기 제1 해상도보다 높은 제2 해상도를 갖는 고해상도 이미지를 생성하는 동작을 포함할 수 있다.

[0012] 일 실시 예에 따르면, 상기 고해상도 이미지를 생성하는 동작은, 상기 유사도에 기반하여, 상기 복수의 특징들 중 상기 임베딩 텍스트와 임계 값 이상의 유사도를 갖는 특징들을 추출하는 동작, 및 상기 추출된 특징들에 기반하여, 상기 고해상도 이미지를 생성하는 동작을 포함할 수 있다.

[0013] 일 실시 예에 따르면, 상기 이미지 벡터로부터 상기 복수의 특징들을 추출하는 동작은, 상기 임베딩 이미지에 포함된 상기 복수의 이미지 벡터들 간의 유사도를 측정하는 동작, 및 상기 측정된 유사도에 기반하여, 상기 복수의 특징들을 추출하는 동작을 포함할 수 있다.

[0014] 일 실시 예에 따르면, 상기 임베딩 이미지 및 상기 임베딩 텍스트를 생성하는 동작은, 상기 이미지 및 상기 텍스트를 각각 임의의 공간에 투영함으로써, 벡터 값을 갖는 상기 임베딩 이미지 및 상기 임베딩 텍스트를 생성하는 것을 특징으로 할 수 있다.

[0015] 또는, 일 실시 예에 따른 이미지의 초해상화 시스템은, 제1 해상도를 갖는 상기 이미지를 수신하고, 수신한 상기 이미지를 임베딩 함으로써, 임베딩 이미지를 생성하는 이미지 인코더, 상기 이미지와 관련된 텍스트를 수신하고, 수신한 상기 텍스트를 임베딩 함으로써, 임베딩 텍스트를 생성하는 텍스트 인코더 및 상기 이미지 인코더 및 상기 텍스트 인코더와 연결되는 프로세서를 포함하고, 상기 프로세서는, 상기 임베딩 이미지에 포함된 복수의 이미지 벡터로부터 상기 이미지와 관련된 복수의 특징들을 추출하고, 상기 임베딩 이미지 및 상기 임베딩 텍스트 간의 유사도를 판단하고, 상기 유사도 및 상기 복수의 특징들 중 적어도 일부에 기반하여, 상기 제1 해상도보다 높은 제2 해상도를 갖는 고해상도 이미지를 생성할 수 있다.

[0016] 일 실시 예에 따르면, 상기 프로세서는, 상기 유사도에 기반하여, 상기 복수의 특징들 중 상기 임베딩 텍스트와 임계 값 이상의 유사도를 갖는 특징들을 추출하고, 상기 추출된 특징들에 기반하여, 상기 고해상도 이미지를 생성할 수 있다.

[0017] 나아가, 일 실시 예에 따르면, 상기 이미지 인코더는, 상기 임베딩 이미지에 포함된 상기 복수의 이미지 벡터들 간의 유사도를 측정하고, 상기 프로세서는, 상기 측정된 유사도에 기반하여, 상기 복수의 특징들을 추출할 수 있다.

[0018] 일 실시 예에 따르면, 상기 이미지 인코더는, 상기 이미지를 임의의 제1 공간에 투영함으로써, 벡터 값을 갖는 상기 임베딩 이미지를 생성하고, 상기 텍스트 인코더는, 상기 텍스트를 임의의 제2 공간에 투영함으로써, 벡터 값을 갖는 상기 임베딩 텍스트를 생성할 수 있다.

[0019] 또는, 일 실시 예에 따른 이미지의 초해상화 방법이 실행될 수 있는 컴퓨터 프로그램이 저장된 기록매체에 있어서, 상기 컴퓨터 프로그램의 실행에 의한 동작은, 제1 해상도를 갖는 상기 이미지 및 상기 이미지와 관련된 텍스트를 수신하는 동작, 상기 이미지 및 상기 텍스트를 각각 임베딩 함으로써, 임베딩 이미지 및 임베딩 텍스트를 생성하는 동작, 상기 임베딩 이미지에 포함된 복수의 이미지 벡터로부터 상기 이미지와 관련된 복수의 특징들을 추출하는 동작, 상기 임베딩 텍스트와 상기 임베딩 이미지 간의 유사도를 판단하는 동작 및 상기 유사도 및 상기 복수의 특징들 중 적어도 일부에 기반하여, 상기 제1 해상도보다 높은 제2 해상도를 갖는 고해상도 이미지를 생성하는 동작을 포함할 수 있다.

발명의 효과

[0021] 본 개시의 다양한 실시 예에 따르면, 이미지 정보 이외에도 텍스트의 정보를 추가적으로 제공하여 초해상화를 수행함으로써, 초해상화 출력의 품질을 향상시킬 수 있다.

[0022] 또한, 본 개시의 다양한 실시 예에 따르면, 이미지에 포함된 도메인 중 텍스트와 관련된 도메인을 추출하여 초

해상화에 활용함으로써, 초해상화의 품질을 향상시키고, 초해상화에 소요되는 시간 및 비용을 절감할 수 있다.

[0023] 본 발명에서 얻을 수 있는 효과는 이상에서 언급한 효과들로 제한되지 않으며, 언급하지 않은 또 다른 효과들은 아래의 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

[0025] 본 발명의 특정한 바람직한 실시예들의 상기에서 설명한 바와 같은 또한 다른 측면들과, 특징들 및 이득들은 첨부 도면들과 함께 처리되는 하기의 설명으로부터 보다 명백하게 될 것이다.

도 1은 일 실시 예에 따른 초해상화 시스템의 블록도다.

도 2는 일 실시 예에 따른 초해상화 방법을 나타내는 흐름도다.

도 3은 일 실시 예에 따라 텍스트를 활용한 초행상화 시스템을 나타내는 개념도다.

도 4는 일 실시 예에 따른 이미지로부터 특징을 추출하는 구성을 도시하는 개념도다.

도면의 설명과 관련하여, 동일 또는 유사한 구성요소에 대해서는 동일 또는 유사한 참조 부호가 사용될 수 있다.

발명을 실시하기 위한 구체적인 내용

[0026] 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 수 있으며, 단지 본 실시 예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 도면부호는 동일 구성 요소를 지칭한다.

[0027] 본 발명의 실시 예들을 설명함에 있어서 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략할 것이다. 그리고 후술되는 용어들을 본 발명의 실시 예에서의 기능을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다. 그러므로 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.

[0028] 이때, 처리 흐름도 도면들의 각 블록과 흐름도 도면들의 조합들은 컴퓨터 프로그램 인스트럭션들에 의해 수행될 수 있음을 이해할 수 있을 것이다. 이들 컴퓨터 프로그램 인스트럭션들은 범용 컴퓨터, 특수용 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비의 프로세서에 탑재될 수 있으므로, 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비의 프로세서를 통해 수행되는 그 인스트럭션들이 흐름도 블록(들)에서 설명된 기능들을 수행하는 수단을 생성하게 된다. 이들 컴퓨터 프로그램 인스트럭션들은 특정 방식으로 기능을 구현하기 위해 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비를 지향할 수 있는 컴퓨터 이용 가능 또는 컴퓨터 판독 가능 메모리에 저장되는 것도 가능하므로, 그 컴퓨터 이용가능 또는 컴퓨터 판독 가능 메모리에 저장된 인스트럭션들은 흐름도 블록(들)에서 설명된 기능을 수행하는 인스트럭션 수단을 내포하는 제조 품목을 생산하는 것도 가능하다. 컴퓨터 프로그램 인스트럭션들은 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비 상에 탑재되는 것도 가능하므로, 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비 상에서 일련의 동작 단계들이 수행되어 컴퓨터로 실행되는 프로세스를 생성해서 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비를 수행하는 인스트럭션들은 흐름도 블록(들)에서 설명된 기능들을 실행하기 위한 단계들을 제공하는 것도 가능하다.

[0029] 또한, 각 블록은 특정된 논리적 기능(들)을 실행하기 위한 하나 이상의 실행 가능한 인스트럭션들을 포함하는 모듈, 세그먼트 또는 코드의 일부를 나타낼 수 있다. 또, 몇 가지 대체 실행 예들에서는 블록들에서 언급된 기능들이 순서를 벗어나서 발생하는 것도 가능함을 주목해야 한다. 예컨대, 잇달아 도시되어 있는 두 개의 블록들은 사실 실질적으로 동시에 수행되는 것도 가능하고 또는 그 블록들이 때때로 해당하는 기능에 따라 역순으로 수행되는 것도 가능하다.

[0030] 이 때, 본 실시 예에서 사용되는 '~부'라는 용어는 소프트웨어 또는 FPGA(field-Programmable Gate Array) 또는 ASIC(Application Specific Integrated Circuit)과 같은 하드웨어 구성요소를 의미하며, '~부'는 어떤 역할들을 수행한다. 그렇지만 '~부'는 소프트웨어 또는 하드웨어에 한정되는 의미는 아니다. '~부'는 어드레싱 할

수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 따라서, 일 예로서 '~부'는 소프트웨어 구성요소들, 객체지향 소프트웨어 구성요소들, 클래스 구성요소들 및 태스크 구성요소들과 같은 구성요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로코드, 회로, 데이터, 데이터베이스, 데이터 구조들, 테이블들, 어레이들, 및 변수들을 포함한다. 구성요소들과 '~부'들 안에서 제공되는 기능은 더 작은 수의 구성요소들 및 '~부'들로 결합되거나 추가적인 구성요소들과 '~부'들로 더 분리될 수 있다. 뿐만 아니라, 구성요소들 및 '~부'들은 디바이스 또는 보안 멀티미디어카드 내의 하나 또는 그 이상의 CPU들을 재생시키도록 구현될 수도 있다.

[0031] 본 발명의 실시예들을 구체적으로 설명함에 있어서, 특정 시스템의 예를 주된 대상으로 할 것이지만, 본 명세서에서 청구하고자 하는 주요한 요지는 유사한 기술적 배경을 가지는 여타의 통신 시스템 및 서비스에도 본 명세서에 개시된 범위를 크게 벗어나지 아니하는 범위에서 적용 가능하며, 이는 당해 기술분야에서 숙련된 기술적 지식을 가진 자의 판단으로 가능할 것이다.

[0032] 이하에서, 첨부된 도면을 참조하여 본 발명의 일 실시 예에 따른 초해상화 시스템(100) 및 초해상화 방법에 대하여 설명하도록 한다.

[0033] 도 1은 일 실시 예에 따른 초해상화 시스템의 블록도다. 도 2는 일 실시 예에 따른 초해상화 방법을 나타내는 흐름도다. 도 3은 일 실시 예에 따라 텍스트를 활용한 초행상화 시스템을 나타내는 개념도다. 도 4는 일 실시 예에 따른 이미지로부터 특징을 추출하는 구성을 도시하는 개념도다.

[0034] 도 1을 참조하면, 일 실시 예에 따른 초해상화 시스템(100)은 이미지(111)를 입력 받아 임베딩 이미지(221)를 생성하는 이미지 인코더(120), 이미지(111)와 관련된 텍스트(112)를 입력 받아 임베딩 텍스트(222)를 생성하는 텍스트 인코더(130), 및 프로세서(110)를 포함할 수 있다.

[0035] 본 발명에 따른 초해상화 시스템(100)의 이미지 인코더(120)는 이미지(111)의 특징을 추출하는 방법을 학습하도록 구성된 기계 학습 모델로 참조될 수 있다.

[0036] 또한, 본 발명에 따른 초해상화 시스템(100)의 텍스트 인코더(130)는 텍스트(112)를 벡터 표현으로 변환하는 방법을 학습하도록 구성된 기계 학습 모델로 참조될 수 있다.

[0037] 즉, 본 발명에 따른 이미지 인코더(120) 및 텍스트 인코더(130)는 각각 이미지(111)와 텍스트(112)를 임베딩하여 임의의 공간에 투영함으로써, 벡터 값을 갖도록 변환하는 방법을 학습하도록 구성된 기계 학습 모델로 참조될 수 있다. 또한, 본 발명에 따른 이미지 인코더(120) 및 텍스트 인코더(130)는 Transformer 모델로 지칭될 수 있다.

[0038] 또한, 상술한 특정 기계 학습 모델(이하, 모델)은 이러한 딥러닝을 통해 이미지(111)의 형상과 공간 정보를 유지하여 연산을 수행하고 여러 개의 필터링을 활용하여 이미지의 특징을 추출하도록 학습된 모델로 이해될 수 있다. 예를 들어, 본 발명에 따른 모델은 DNN(deep neural network), RNN(recurrent neural network), BRDNN(bidirectional recurrent deep neural network), 합성곱 신경망(convolutional neural network, CNN)와 같은 기계 학습 모델로 이해될 수 있다. 나아가, 본 발명에 따른 모델은 인공 지능 모델(artificial intelligence model) 또는 딥러닝 모델(deep learning model)로 지칭될 수 있다.

[0039] 이 때, 프로세서(110)는 초해상화 시스템(100)의 전반적인 동작을 제어할 수 있다. 일 실시 예에 따른 프로세서(110)는 이미지 인코더(120) 및 텍스트 인코더(130)를 이용하여, 이미지(111) 및 텍스트(112)의 정보로부터 보다 높은 해상도를 갖는 고해상도 이미지(210)를 생성할 수 있다. 따라서, 이하에서는, 입력된 이미지(111)보다 높은 해상도를 갖는 고해상도 이미지(210)를 생성하기 위한 전반적인 동작이 프로세서(110)의 제어에 의해 수행되는 것으로 설명한다.

[0040] 한편, 초해상화 시스템(100)은 입력부(미도시)를 더 포함할 수 있다. 일 실시 예에 따른 입력부는 관리자로부터 입력되는 이미지(111) 또는 이미지(111)와 관련된 텍스트(112)에 대한 입력을 위한 것으로서, 입력부는 관리자와 초해상화 시스템(100) 사이의 매개체가 될 수 있다. 보다 구체적으로, 입력부는 관리자로부터 이미지(111) 및 텍스트(112)에 대한 전자 파일 형태의 데이터를 수신하는 입력 수단을 의미할 수 있다. 이를 위해, 초해상화 시스템(100)의 입력부는 전자 파일 형태의 데이터를 수신할 수 있도록 구성될 수 있다.

[0041] 도 2 및 도 3을 함께 참조하면, 일 실시 예에 따른 프로세서(110)는 동작 201에서 제1 해상도를 갖는 이미지(111) 및 이미지(111)와 관련된 텍스트(112)를 수신할 수 있다.

[0042] 이 때, 텍스트(112)는 이미지(111)의 특징 중 적어도 일부를 설명하는 문장 또는 단어로 이해될 수 있으나, 이에 한정되는 것은 아니며, 이미지(111)를 설명하는 다양한 형태의 텍스트 정보로 참조될 수 있다.

[0043] 예를 들어, 프로세서(110)는 제1 해상도를 갖고 “늑대”의 형상을 포함하는 이미지(111) 및 “늑대는 측면을 바라보고 있다”와 같은 텍스트(112)를 수신할 수 있다.

[0044] 일 실시 예에 따르면, 프로세서(110)는 입력부를 통해 이미지(111)와 텍스트(112)를 동시에 수신할 수 있으나, 이에 한정되는 것은 아니며, 이미지(111)와 텍스트(112)를 순차적으로 수신할 수 있다.

[0045] 다른 실시 예(미도시)에 따르면, 이미지 인코더(120)는 제1 해상도를 갖는 이미지(111)를 수신하고, 텍스트 인코더(130)는 이미지(111)와 관련된 텍스트(112)를 수신할 수 있다.

[0046] 나아가, 일 실시 예에 따른 프로세서(110)는 동작 203에서, 이미지(111) 및 텍스트(112)를 각각 임베딩 함으로써, 임베딩 이미지(221) 및 임베딩 텍스트(222)를 생성할 수 있다.

[0047] 보다 구체적으로, 일 실시 예에 따른 프로세서(110)는 동작 203에서 이미지 인코더(120)를 제어함으로써, 입력된 이미지(111)를 특정 공간에 투영시켜 임베딩 이미지(221)를 생성할 수 있다.

[0048] 또한, 일 실시 예에 따른 프로세서(110)는 동작 203에서 텍스트 인코더(130)를 제어함으로써, 입력된 텍스트(112)를 특정 공간에 투영시켜 임베딩 텍스트(222)를 생성할 수 있다.

[0049] 이 때, 임베딩 이미지(221) 및 임베딩 텍스트(222)는 각각의 정보를 가진 채로 임의의 공간에 투영되어 존재할 수 있다.

[0050] 이를 통해, 프로세서(110)는 입력된 이미지(111) 및 텍스트(112)를 학습된 기계 학습 모델이 처리 가능하도록 임베딩 형태로 변환할 수 있다.

[0051] 나아가, 일 실시 예에 따른 도 4를 참조하면, 프로세서(110)는 동작 205에서, 임베딩 이미지(221)로부터 이미지(111)와 관련된 복수의 특징들을 추출할 수 있다.

[0052] 보다 구체적으로, 프로세서(110)는 동작 205에서, 이미지 인코더(120)를 제어함으로써, 임베딩 이미지(221)에 포함된 복수의 이미지 벡터들(또는, 임베딩 벡터들) 간의 유사도를 측정할 수 있다. 나아가, 프로세서(110)는 측정한 유사도에 기반하여 이미지(111)와 관련된 복수의 특징들을 추출할 수 있다.

[0053] 다른 실시 예(미도시)에 따르면, 프로세서(110)는 이미지 인코더(120)를 제어함으로써, 이미지(111)를 복수의 패치 단위로 분할하고, 복수의 패치 간의 유사도를 측정함으로써, 이미지(111)와 관련된 복수의 특징들을 추출할 수 있다.

[0054] 상술한 구성을 통해, 본 발명에 따른 초해상화 시스템(100)은 입력된 이미지(111)로부터 이미지(111)를 고해상도로 재구성(reconstruction)하기 위해 요구되는 특징들을 추출할 수 있다.

[0055] 나아가, 일 실시 예에 따른 프로세서(110)는 동작 207에서, 임베딩 이미지(221)와 임베딩 텍스트(222) 간의 유사도를 판단할 수 있다.

[0056] 보다 구체적으로, 프로세서(110)는 텍스트(112)와 이미지(111) 간의 유사도를 판단하도록 학습된 기계 학습 모델을 이용하여, 임베딩 텍스트(112)와 임베딩 이미지(111)의 유사도를 판단할 수 있다.

[0057] 이 때, 상기 기계 학습 모델은 텍스트(112)와 이미지(111) 간의 유사도를 판단하기 위한 “cross attention” 기법을 수행하도록 학습된 기계 학습 모델로 이해될 수 있다. 또한, 상술한 모델은 프로세서(110) 내부에 포함될 수 있으나, 이에 한정되는 것은 아니며, 별도의 구성으로 구비되고 프로세서(110)와 연결될 수 있다.

[0058] 일 실시 예에 따르면, 프로세서(110)는 이미지(111)의 복수의 특징들과 임베딩 텍스트(222) 간의 유사도를 판단할 수 있다.

[0059] 나아가, 프로세서(110)는 이미지(111)의 복수의 특징들 중 임베딩 텍스트(222)와 기 설정된 임계 값 이상의 유사도를 갖는 특징들을 추출할 수 있다.

[0060] 일 실시 예에 따르면, 프로세서(110)는 동작 209에서, 동작 207을 통해 판단된 유사도 및 복수의 특징들 중 적어도 일부에 기반하여, 제1 해상도 보다 높은 제2 해상도를 갖는 고해상도 이미지(210)를 생성할 수 있다.

[0061] 보다 구체적으로, 프로세서(110)는 동작 209에서, 임베딩 이미지(221)로부터 판단된 이미지(111)의 복수의 특징들 중 임베딩 텍스트(222)와 임계 값 이상의 유사도를 갖는 특징들에 기반하여, 이미지(111)보다 높은 해상도를

갖는 고해상도 이미지(210)를 생성할 수 있다.

[0062] 예를 들어, 프로세서(110)는 이미지(111)와 관련된 복수의 특징들 중 입력된 텍스트(112)(“늑대는 측면을 바라보고 있다”)와 유사한 특징들에 기반하여, 이미지(111)보다 높은 해상도를 갖는 고해상도 이미지(210)를 생성할 수 있다.

[0063] 즉, 본 발명에 따른 초해상화 시스템(100)(또는, 프로세서(110))은 이미지(111) 및 이미지(111)와 관련된 텍스트(112)에 기반하여, 입력된 이미지(111)보다 높은 해상도를 가진 고해상도 이미지(210)를 재구성(reconstruction) 할 수 있다.

[0064] 상술한 구성을 통해, 본 발명에 따른 초해상화 시스템(100)은 이미지(111) 및 이미지(111)와 관련된 텍스트(112)를 함께 고려함으로써, 초해상화에 따른 출력 값(예: 고해상도 이미지(210))의 품질을 향상시킬 수 있다.

[0065] 또한, 상술한 구성을 통해, 본 발명에 따른 초해상화 시스템(100)은 이미지(111)에 포함된 도메인 중 텍스트(112)와 관련된 도메인을 추출하여 초해상화에 활용함으로써, 초해상화의 품질을 향상시키고, 초해상화에 소요되는 시간 및 비용을 절감할 수 있다.

[0066] 이와 같이, 본 발명의 다양한 실시예들은 특정 관점에서 컴퓨터 리드 가능 기록 매체(computer readable recording medium)에서 컴퓨터 리드 가능 코드(computer readable code)로서 구현될 수 있다. 컴퓨터 리드 가능 기록 매체는 컴퓨터 시스템에 의해 리드될 수 있는 데이터를 저장할 수 있는 임의의 데이터 저장 디바이스이다. 컴퓨터 리드 가능 기록 매체의 예들은 읽기 전용 메모리(read only memory: ROM)와, 랜덤-접속 메모리(random access memory: RAM)와, 컴팩트 디스크- 리드 온니 메모리(compact disk-read only memory: CD-ROM)들과, 마그네틱 테이프(magnetic tape)들과, 플로피 디스크(floppy disk)들과, 광 데이터 저장 디바이스들, 및 캐리어 웨이브(carrier wave)들(인터넷을 통한 데이터 송신 등)을 포함할 수 있다. 컴퓨터 리드 가능 기록 매체는 또한 네트워크 연결된 컴퓨터 시스템들을 통해 분산될 수 있고, 따라서 컴퓨터 리드 가능 코드는 분산 방식으로 저장 및 실행된다. 또한, 본 발명의 다양한 실시예들을 성취하기 위한 기능적 프로그램들, 코드, 및 코드 세그먼트(segment)들은 본 발명이 적용되는 분야에서 숙련된 프로그래머들에 의해 쉽게 해석될 수 있다.

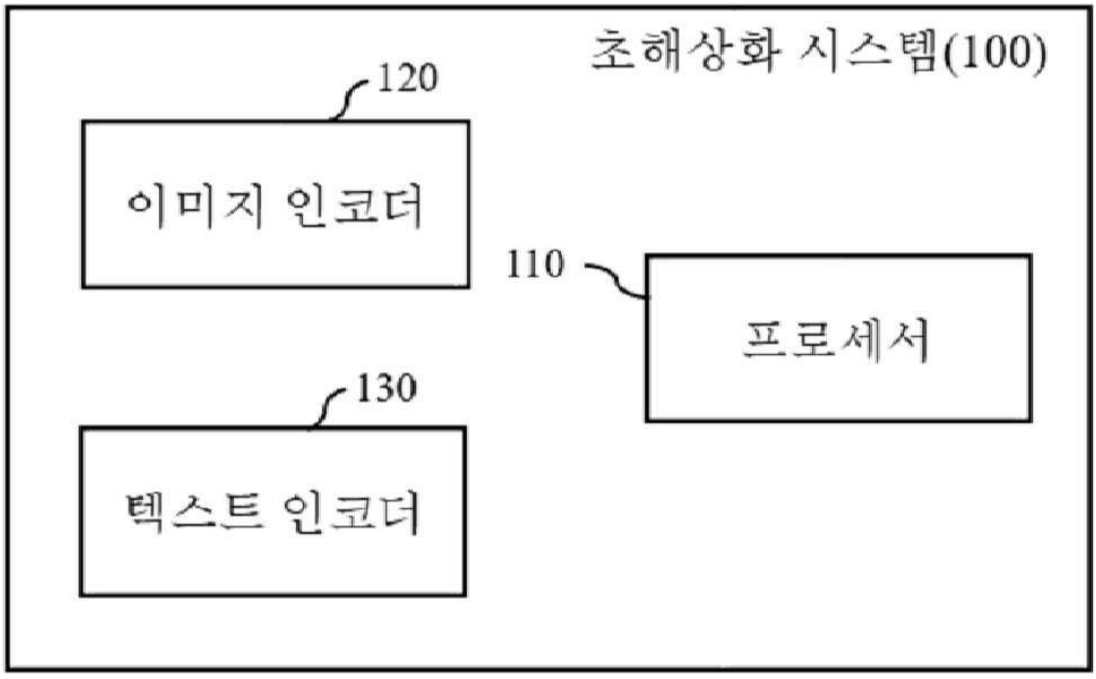
[0067] 또한 본 발명의 다양한 실시예들에 따른 장치 및 방법은 하드웨어, 소프트웨어 또는 하드웨어 및 소프트웨어의 조합의 형태로 실현 가능하다는 것을 알 수 있을 것이다. 이러한 소프트웨어는 예를 들어, 삭제 가능 또는 재기록 가능 여부와 상관없이, ROM 등의 저장 장치와 같은 휘발성 또는 비휘발성 저장 장치, 또는 예를 들어, RAM, 메모리 칩, 장치 또는 집적 회로와 같은 메모리, 또는 예를 들어 콤팩트 디스크(compact disk: CD), DVD, 자기 디스크 또는 자기 테이프 등과 같은 광학 또는 자기적으로 기록 가능함과 동시에 기계(예를 들어, 컴퓨터)로 읽을 수 있는 저장 매체에 저장될 수 있다. 본 발명의 다양한 실시예들에 따른 방법은 제어 모듈 및 메모리를 포함하는 컴퓨터 또는 휴대 단말에 의해 구현될 수 있고, 이러한 메모리는 본 발명의 실시예들을 구현하는 명령들을 포함하는 프로그램 또는 프로그램들을 저장하기에 적합한 기계로 읽을 수 있는 저장 매체의 한 예임을 알 수 있을 것이다.

[0068] 따라서, 본 발명은 본 명세서의 청구항에 기재된 장치 또는 방법을 구현하기 위한 코드를 포함하는 프로그램 및 이러한 프로그램을 저장하는 기계(컴퓨터 등)로 읽을 수 있는 저장 매체를 포함한다. 또한, 이러한 프로그램은 유선 또는 무선 연결을 통해 전달되는 통신 신호와 같은 임의의 매체를 통해 전자적으로 이송될 수 있고, 본 발명은 이와 균등한 것을 적절하게 포함한다.

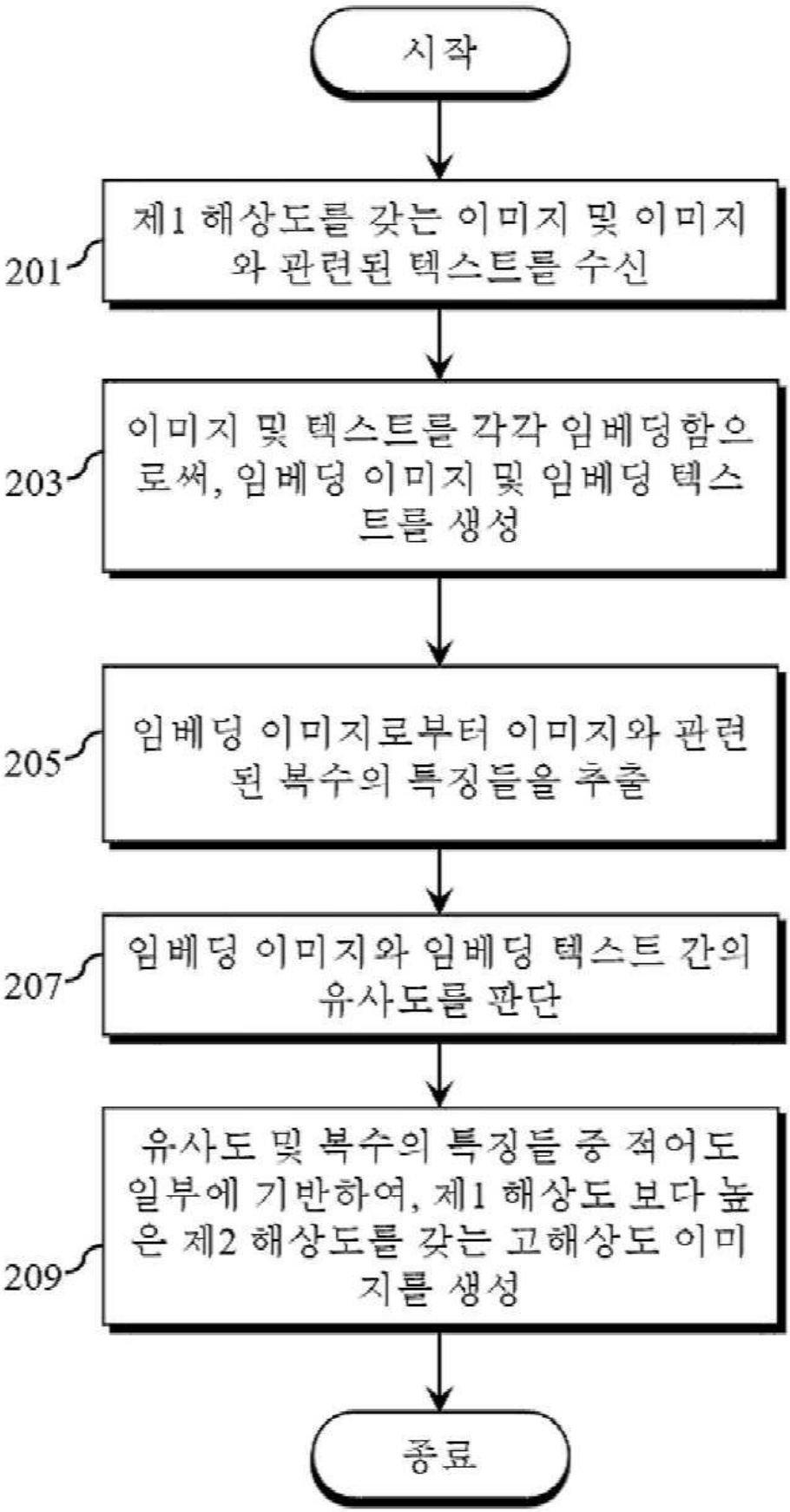
[0069] 본 명세서와 도면에 개시된 본 발명의 실시 예들은 본 발명의 기술 내용을 쉽게 설명하고, 본 발명의 이해를 돕기 위해 특정 예를 제시한 것일 뿐이며, 본 발명의 범위를 한정하고자 하는 것은 아니다. 또한 앞서 설명된 본 발명에 따른 실시예들은 예시적인 것에 불과하며, 당해 분야에서 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 범위의 실시예가 가능하다는 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호 범위는 다음의 특허청구범위에 의해서 정해져야 할 것이다.

도면

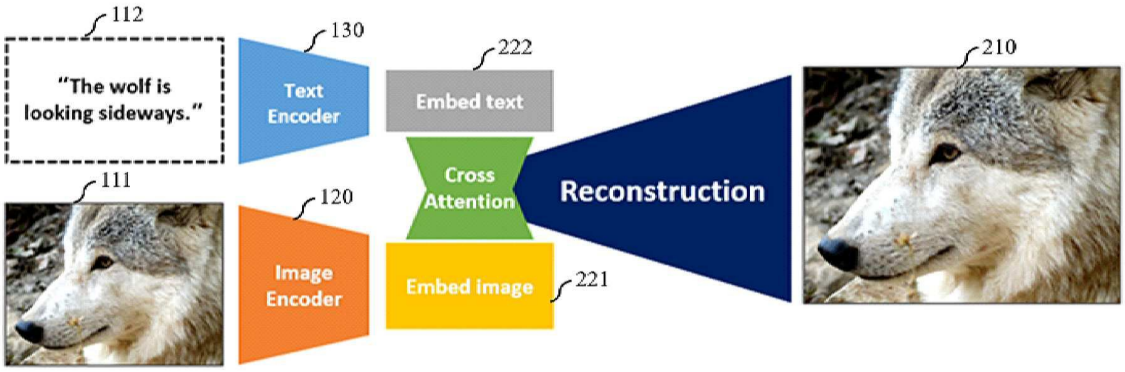
도면1



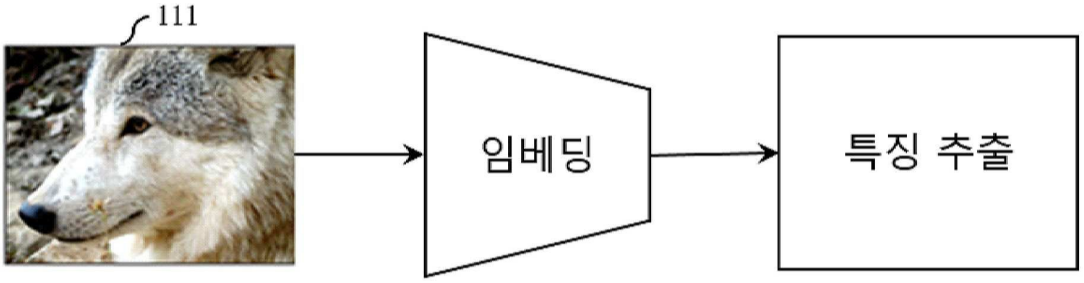
도면2


시작
201
제1 해상도를 갖는 이미지 및 이미지와 관련된 텍스트를 수신
203
이미지 및 텍스트를 각각 임베딩함으로써, 임베딩 이미지 및 임베딩 텍스트를 생성
205
임베딩 이미지로부터 이미지와 관련된 복수의 특징들을 추출
207
임베딩 이미지와 임베딩 텍스트 간의 유사도를 판단
209
유사도 및 복수의 특징들 중 적어도 일부에 기반하여, 제1 해상도 보다 높은 제2 해상도를 갖는 고해상도 이미지를 생성
종료

도면3


112
"The wolf is looking sideways."
130
Text Encoder
222
Embed text
210
111
120
Image Encoder
Cross Attention
Reconstruction
Embed image
221

도면4


111
임베딩
특징 추출