



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년11월20일
(11) 등록번호 10-2604657
(24) 등록일자 2023년11월16일

(51) 국제특허분류(Int. Cl.)
G06T 9/00 (2019.01) G06N 20/10 (2019.01)
G06N 3/04 (2023.01) G06N 3/08 (2023.01)
H04N 19/14 (2014.01) H04N 19/85 (2014.01)
(52) CPC특허분류
G06T 9/002 (2013.01)
G06N 20/10 (2021.08)
(21) 출원번호 10-2022-0039403
(22) 출원일자 2022년03월30일
심사청구일자 2022년03월30일
(65) 공개번호 10-2023-0140755
(43) 공개일자 2023년10월10일
(56) 선행기술조사문헌
US20200092558 A1*
CHAOYI LIN ET AL, "A Spatial RNN Codec for End-To-End Image Compression", 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2020.08.05. 1부.*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
이상윤
서울특별시 서초구 청두곶14길 6, 202호(방배동)
강홍구
서울특별시 은평구 백련산로 38, 206동 301호(응암동, 백련산 힐스테이트2차)
(74) 대리인
민영준

전체 청구항 수 : 총 18 항

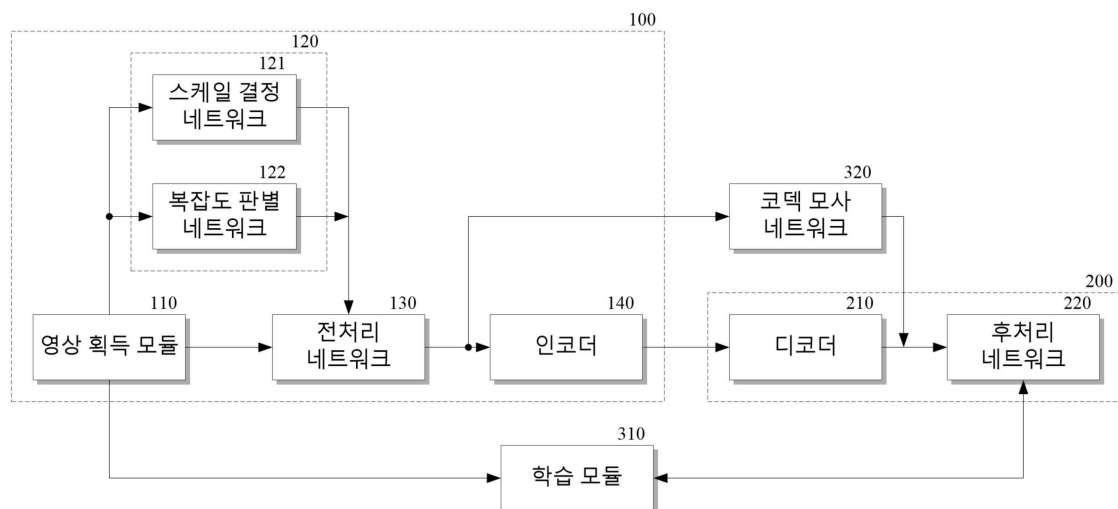
심사관 : 권영학

(54) 발명의 명칭 영상 압축 성능 개선 방법 및 장치

(57) 요약

개시된 실시예는 인가된 입력 영상에 대해 신경망 연산하여 입력 영상의 특성 따른 적어도 하나의 압축 파라미터를 획득하는 단계, 적어도 하나의 압축 파라미터에 따라 입력 영상에 대한 전처리를 수행하는 인공 신경망의 가중치를 조절하는 단계, 가중치가 조절된 인공 신경망으로 입력 영상에 대한 전처리를 수행하여 압축 영상을 획득하는 단계 및 압축 영상을 인코딩하여 인코딩 영상을 획득하는 단계를 포함하여, 입력 영상의 특성에 따라 적응적으로 입력 영상의 시간 및 공간 해상도를 압축 및 복원하므로, 압축률을 크게 향상시킬 수 있을 뿐만 아니라 향상된 압축률에도 복원 영상의 품질 저하를 최대한 억제할 수 있는 영상 압축 성능 개선 방법 및 장치를 제공한다.

대표도



(52) CPC특허분류

G06N 3/04 (2023.01)

G06N 3/084 (2023.01)

H04N 19/14 (2015.01)

H04N 19/85 (2015.01)

G06T 2207/20084 (2013.01)

명세서

청구범위

청구항 1

하나 이상의 프로세서 및 상기 하나 이상의 프로세서에 의해 실행되는 하나 이상의 프로그램들을 저장하는 메모리를 구비한 컴퓨팅 장치에 의해 수행되는 방법으로서,

인가된 입력 영상에 대해 신경망 연산하여 상기 입력 영상의 특성에 따른 적어도 하나의 압축 파라미터를 획득하는 단계;

상기 적어도 하나의 압축 파라미터에 따라 상기 입력 영상에 대한 전처리를 수행하는 인공 신경망의 가중치를 조절하는 단계;

가중치가 조절된 인공 신경망으로 상기 입력 영상에 대한 전처리를 수행하여 압축 영상을 획득하는 단계; 및

상기 압축 영상을 인코딩하여 인코딩 영상을 획득하는 단계를 포함하되,

상기 가중치를 조절하는 단계는

상기 전처리를 수행하는 인공 신경망의 다수의 연산 레이어 각각에 포함된 메인 커널과 적어도 하나의 서브 커널을 구성하는 가중치를 적어도 상기 적어도 하나의 압축 파라미터에 따른 비율로 혼합하여 각 연산 레이어의 가중치를 조절하는 영상 압축 성능 개선 방법.

청구항 2

삭제

청구항 3

제1항에 있어서, 상기 가중치를 조절하는 단계는

상기 메인 커널을 구성하는 가중치에 압축 파라미터에 따른 비율($1-\alpha$)을 가중하고, 서브 커널을 구성하는 가중치에 압축 파라미터(α)를 가중하여 합하여 가중치를 조절하는 영상 압축 성능 개선 방법.

청구항 4

제3항에 있어서, 상기 가중치를 조절하는 단계는

상기 압축 파라미터가 다수개이면, 메인 커널의 가중치 압축 파라미터에 따른 비율로 서브 커널의 가중치를 혼합하여 메인 커널의 가중치를 조절하고, 이후, 이전 압축 파라미터에 의해 조절된 메인 커널 가중치에 순차적으로 다른 압축 파라미터에 따른 비율로 서브 커널의 가중치를 혼합하여 조절하는 영상 압축 성능 개선 방법.

청구항 5

제1항에 있어서, 상기 적어도 하나의 압축 파라미터를 획득하는 단계는

둘 이상의 인공 신경망을 포함하고, 둘 이상의 인공 신경망은 각각 상기 입력 영상을 인가받아 신경망 연산하여, 상기 입력 영상의 크기를 줄일 수 있는 위한 비율을 나타내는 스케일 파라미터와 상기 입력 영상의 프레임 내 및 프레임간 복잡도를 나타내는 복잡도 파라미터를 상기 압축 파라미터로서 획득하는 영상 압축 성능 개선 방법.

청구항 6

제5항에 있어서, 상기 압축 영상을 획득하는 단계는

가중치가 조절된 인공 신경망의 다수의 연산 레이어로 상기 입력 영상에 대해 신경망 연산하여 특징맵을 출력하고, 상기 적어도 하나의 압축 파라미터에 상기 스케일 파라미터가 포함되어 있으면, 상기 스케일 파라미터에 따

라 스케일 다운 레이어가 특징맵을 다운 스케일링하여 상기 압축 영상을 획득하는 영상 압축 성능 개선 방법.

청구항 7

제1항에 있어서, 상기 인코딩 영상을 획득하는 단계는

인코딩된 영상에 상기 적어도 하나의 압축 파라미터를 포함하여 상기 인코딩 영상을 획득하는 영상 압축 성능 개선 방법.

청구항 8

제7항에 있어서, 상기 영상 압축 성능 개선 방법은

상기 인코딩 영상을 인가받아 디코딩하여 디코딩 영상을 획득하고, 상기 인코딩 영상에 포함된 상기 적어도 하나의 압축 파라미터를 추출하는 단계;

상기 적어도 하나의 압축 파라미터에 따라 상기 디코딩 영상에 대한 후처리를 수행하는 인공 신경망의 후처리 가중치를 조절하는 단계; 및

가중치가 조절된 인공 신경망으로 상기 디코딩 영상에 대해 후처리를 수행하여 복원 영상을 획득하는 단계를 더 포함하는 영상 압축 성능 개선 방법.

청구항 9

제8항에 있어서, 상기 가중치를 조절하는 단계는

상기 후처리를 수행하는 인공 신경망의 다수의 연산 레이어 각각에 포함된 메인 커널과 적어도 하나의 서브 커널을 구성하는 가중치를 적어도 상기 적어도 하나의 압축 파라미터에 따른 비율로 혼합하여 각 연산 레이어의 가중치를 조절하는 영상 압축 성능 개선 방법.

청구항 10

제9항에 있어서, 상기 복원 영상을 획득하는 단계는

가중치가 조절된 인공 신경망의 다수의 연산 레이어로 상기 입력 영상에 대해 신경망 연산하여 특징맵을 출력하고, 상기 적어도 하나의 압축 파라미터에 스케일 파라미터가 포함되어 있으면, 상기 스케일 파라미터에 따라 스케일 업 레이어가 특징맵을 업 스케일링하여 상기 복원 영상을 획득하는 영상 압축 성능 개선 방법.

청구항 11

하나 이상의 프로세서; 및 상기 하나 이상의 프로세서에 의해 실행되는 하나 이상의 프로그램들을 저장하는 메모리를 구비한 장치로서,

상기 프로세서는

인가된 입력 영상에 대해 신경망 연산하여 상기 입력 영상의 특성에 따른 적어도 하나의 압축 파라미터를 획득하고,

상기 적어도 하나의 압축 파라미터에 따라 상기 입력 영상에 대한 전처리를 수행하는 인공 신경망의 가중치를 조절하고,

가중치가 조절된 인공 신경망으로 상기 입력 영상에 대한 전처리를 수행하여 압축 영상을 획득하며,

상기 압축 영상을 인코딩하여 인코딩 영상을 획득하고,

상기 전처리를 수행하는 인공 신경망의 다수의 연산 레이어 각각에 포함된 메인 커널과 적어도 하나의 서브 커널을 구성하는 가중치를 적어도 상기 적어도 하나의 압축 파라미터에 따른 비율로 혼합하여 각 연산 레이어의 가중치를 조절하는 영상 압축 성능 개선 장치.

청구항 12

삭제

청구항 13

제11항에 있어서, 상기 프로세서는

상기 메인 커널을 구성하는 가중치에 압축 파라미터에 따른 비율($1-\alpha$)을 가중하고, 서브 커널을 구성하는 가중치에 압축 파라미터(α)를 가중하여 합하여 가중치를 조절하는 영상 압축 성능 개선 장치.

청구항 14

제13항에 있어서, 상기 프로세서는

상기 압축 파라미터가 다수개이면, 메인 커널의 가중치 압축 파라미터에 따른 비율로 서브 커널의 가중치를 혼합하여 메인 커널의 가중치를 조절하고, 이후, 이전 압축 파라미터에 의해 조절된 메인 커널 가중치에 순차적으로 다른 압축 파라미터에 따른 비율로 서브 커널의 가중치를 혼합하여 조절하는 영상 압축 성능 개선 장치.

청구항 15

제11항에 있어서, 상기 프로세서는

각각 상기 입력 영상을 인가받아 신경망 연산하는 둘 이상의 인공 신경망을 이용하여, 상기 입력 영상의 크기를 줄일 수 있는 위한 비율을 나타내는 스케일 파라미터와 상기 입력 영상의 프레임 내 및 프레임간 복잡도를 나타내는 복잡도 파라미터를 상기 압축 파라미터로서 획득하는 영상 압축 성능 개선 장치.

청구항 16

제15항에 있어서, 상기 프로세서는

가중치가 조절된 상기 다수의 연산 레이어로 상기 입력 영상에 대해 신경망 연산하여 특징맵을 출력하고, 상기 적어도 하나의 압축 파라미터에 상기 스케일 파라미터가 포함되어 있으면, 상기 스케일 파라미터에 따라 상기 특징맵을 다운 스케일링하여 상기 압축 영상을 획득하는 영상 압축 성능 개선 장치.

청구항 17

제11항에 있어서, 상기 프로세서는

인코딩된 영상에 상기 적어도 하나의 압축 파라미터를 포함하여 상기 인코딩 영상을 획득하는 영상 압축 성능 개선 장치.

청구항 18

제17항에 있어서, 상기 프로세서는

상기 인코딩 영상을 인가받아 디코딩하여 디코딩 영상을 획득하고, 상기 인코딩 영상에 포함된 상기 적어도 하나의 압축 파라미터를 추출하고,

상기 적어도 하나의 압축 파라미터에 따라 상기 디코딩 영상에 대한 후처리를 수행하는 인공 신경망의 후처리 가중치를 조절하며,

가중치가 조절된 인공 신경망으로 상기 디코딩 영상에 대해 후처리를 수행하여 복원 영상을 획득하는 영상 압축 성능 개선 장치.

청구항 19

제18항에 있어서, 상기 프로세서는

상기 후처리를 수행하는 인공 신경망의 다수의 연산 레이어 각각에 포함된 메인 커널과 적어도 하나의 서브 커널을 구성하는 가중치를 적어도 상기 적어도 하나의 압축 파라미터에 따른 비율로 혼합하여 각 연산 레이어의 가중치를 조절하는 영상 압축 성능 개선 장치.

청구항 20

제19항에 있어서, 상기 프로세서는

가중치가 조절되어 후처리를 수행하는 인공 신경망의 다수의 연산 레이어로 상기 입력 영상에 대해 신경망 연산하여 특징맵을 출력하고, 상기 적어도 하나의 압축 파라미터에 스케일 파라미터가 포함되어 있으면, 상기 스케일 파라미터에 따라 상기 특징맵을 업 스케일링하여 상기 복원 영상을 획득하는 영상 압축 성능 개선 장치.

발명의 설명

기술 분야

[0001] 개시되는 실시예들은 영상 압축 성능 개선 방법 및 장치에 관한 것으로, 입력 영상의 특성에 따라 적응적으로 시공간 해상도를 조절하여 비디오 코덱의 효율성 향상시킬 수 있는 영상 압축 성능 개선 방법 및 장치에 관한 것이다.

배경 기술

[0002] 화상 통신 및 동영상 기반의 OTT(Over-The-Top media service) 서비스가 급격히 증가함에 따라, 영상 신호를 효과적으로 전송 혹은 저장하기 위한 비디오 압축 혹은 코덱 기술에 대한 중요성이 지속적으로 제기되고 있다. 신호처리 및 확률/통계 이론을 기반으로 구현된 H.26L, HEVC (High Efficiency Video Codec) codec과 같은 기존의 레거시(legacy) 비디오 코덱은 ITU-T, MPEG 등의 국제 표준화 작업을 통해 다양한 응용 분야에서 성공적으로 상용화되고 있다.

[0003] 한편 고화질 영상을 선호하는 사용자들의 요구를 해결하기 위해 영상 디스플레이 화면 크기가 지속적으로 증가되고 있어, 영상의 시간 및 공간 상에서의 해상도 역시 급격히 증가하고 있다. 이로 인해 고화질 영상을 표현하기 위해 필요한 정보량 역시 증가하였으며, 정보량이 증가된 고화질 영상을 저장하기 위해 필요한 메모리 용량 및 스트리밍 서비스를 위한 네트워크의 용량 역시 기하 급수적으로 증가하고 있다.

[0004] 따라서 고화질 영상을 효율적으로 처리하기 위해 영상 압축 기술이 활발하게 연구되고 있다. 영상 압축 기술은 영상의 화질을 최대한 원본과 가깝게 유지하면서 적은 비트수로 영상 데이터를 표현하는 기술로써, 영상을 표현하기 위한 데이터의 양을 줄여 전송 및 저장의 효율성을 향상시키기 위한 기술이다. 코덱 또한 영상 압축 기술의 일부로 볼 수 있다.

[0005] 한편 인공 신경망을 이용한 딥러닝 기술의 발전으로 인해, 로우 레벨 컴퓨터 비전에서도 압축된 영상을 원본 영상에 가깝게 복원할 수 있는 초해상도(Super-resolution) 기법 및 프레임 보간 기법 등이 제안되었다. 딥러닝 기술을 이용하면 입력 영상의 스케일이나 프레임율이 축소될 지라도 원본 영상으로 더욱 가깝게 복원할 수 있다. 즉 기존의 핸드 크래프트(Hand-crafted) 알고리즘과 비교하여 압축 영상 복원 시, 왜곡의 발생을 최대한 줄일 수 있다. 이에 동영상 압축에서 인공 신경망을 이용하여 입력 영상의 시간 및 공간적 해상도 축소를 통한 전처리를 수행하고 축소된 영상을 기존 코덱을 이용해 영상을 부호화하여 압축률을 크게 높이고, 압축된 비트스트림을 기존 코덱의 디코더로 복호화를 수행한 후 인공 신경망으로 영상의 해상도를 복원하는 방식도 제안되었다.

[0006] 이러한 코덱과 인공 신경망을 함께 이용하는 압축 기법은 기존 비디오 코덱 모듈과 결합되어 매우 우수한 압축 성능 향상을 이루어 냈다. 다만 비디오 코덱의 전처리 및 후처리 작업을 수행하는 인공 신경망이 End-to-end 학습 방법 학습되므로, 학습 시에 가중치가 고정되어 여전히 영상의 다양한 특성을 명확히 반영하지 못한다는 한계가 있다.

[0007] 영상 압축 기술을 이용하여 인코딩 과정에서 입력 영상의 크기 및 프레임율을 줄이면, 영상의 총 정보량이 감소하여 인코딩된 비트스트림(bitstream)의 비트레이트(bitrate)가 원본 영상보다 줄어들도록 압축될 수 있다. 그러나 다양한 특성을 갖는 입력 영상을 동일한 방식으로 압축하는 경우, 입력 영상의 특성에 따라 서로 상이한 크기로 정보량 손실이 발생된다. 따라서 단순히 이중선형 다운 샘플링(bilinear downsampling) 및 프레임율 다운 샘플링을 수행했을 때, 비트레이트의 이득에 비해 복구된 영상의 왜곡(Distortion)이 월등히 커질 수도 있다. 예로서 기존에는 입력 영상 신호의 공간에 대한 해상도를 N:1 형태로 줄여서 레거시 코덱에 입력하여 부호화/복호화한 후, 초해상도 기법을 통해 복호화하여 원본 영상 해상도의 영상을 복원하였다. 또는 프레임 보간 기법을 이용하여 시간축 해상도를 복원하는 방식을 이용하였다. 그러나 입력 영상의 콘텐츠 종류나 촬영 방법 등에 따라 영상의 시간 및 공간적 특성이 변화하므로, 초해상도 기법과 프레임 보간 기법을 단순히 적용하는 경우, 복원된 영상의 화질 저하가 발생할 수 있다. 즉 입력 영상의 특성에 따라 압축 성능이 균일하지 않아 압

축에 의한 효율성이 일반적으로 향상된다고 볼 수 없다는 문제가 있다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 한국 공개 특허 제10-2018-0119753호 (2018.11.05 공개)

발명의 내용

해결하려는 과제

[0009] 개시되는 실시예들은 입력 영상을 특성에 따라 영상을 상이하게 압축 및 복원함으로써, 압축률을 향상시킬 수 있는 영상 압축 성능 개선 방법 및 장치를 제공하는데 있다.

[0010] 개시되는 실시예들은 압축률이 향상되면서도 복원 영상의 품질 저하를 최대한 억제할 수 있는 영상 압축 성능 개선 방법 및 장치를 제공하는데 있다.

과제의 해결 수단

[0011] 실시예에 따른 영상 압축 성능 개선 방법은 인가된 입력 영상에 대해 신경망 연산하여 상기 입력 영상의 특성 따른 적어도 하나의 압축 파라미터를 획득하는 단계; 상기 적어도 하나의 압축 파라미터에 따라 상기 입력 영상에 대한 전처리를 수행하는 인공 신경망의 가중치를 조절하는 단계; 가중치가 조절된 인공 신경망으로 상기 입력 영상에 대한 전처리를 수행하여 압축 영상을 획득하는 단계; 및 상기 압축 영상을 인코딩하여 인코딩 영상을 획득하는 단계를 포함한다.

[0012] 상기 가중치를 조절하는 단계는 상기 전처리를 수행하는 인공 신경망의 다수의 연산 레이어 각각에 포함된 메인 커널과 적어도 하나의 서브 커널을 구성하는 가중치를 적어도 상기 적어도 하나의 압축 파라미터에 따른 비율로 혼합하여 각 연산 레이어의 가중치를 조절할 수 있다.

[0013] 상기 가중치를 조절하는 단계는 상기 메인 커널을 구성하는 가중치에 압축 파라미터에 따른 비율($1-\alpha$)을 가중하고, 서브 커널을 구성하는 가중치에 압축 파라미터(α)를 가중하여 합하여 가중치를 조절할 수 있다.

[0014] 상기 가중치를 조절하는 단계는 상기 압축 파라미터가 다수개이면, 메인 커널의 가중치 압축 파라미터에 따른 비율로 서브 커널의 가중치를 혼합하여 메인 커널의 가중치를 조절하고, 이후, 이전 압축 파라미터에 의해 조절된 메인 커널 가중치에 순차적으로 다른 압축 파라미터에 따른 비율로 서브 커널의 가중치를 혼합하여 조절할 수 있다.

[0015] 상기 적어도 하나의 압축 파라미터를 획득하는 단계는 둘 이상의 인공 신경망을 포함하고, 둘 이상의 인공 신경망은 각각 상기 입력 영상을 인가받아 신경망 연산하여, 상기 입력 영상의 크기를 줄일 수 있는 위한 비율을 나타내는 스케일 파라미터와 상기 입력 영상의 프레임 내 및 프레임간 복잡도를 나타내는 복잡도 파라미터를 상기 압축 파라미터로서 획득할 수 있다.

[0016] 상기 압축 영상을 획득하는 단계는 가중치가 조절된 인공 신경망의 다수의 연산 레이어로 상기 입력 영상에 대해 신경망 연산하여 특징맵을 출력하고, 상기 적어도 하나의 압축 파라미터에 상기 스케일 파라미터가 포함되어 있으면, 상기 스케일 파라미터에 따라 스케일 다운 레이어가 특징맵을 다운 스케일링하여 상기 압축 영상을 획득할 수 있다.

[0017] 상기 인코딩 영상을 획득하는 단계는 인코딩된 영상에 상기 적어도 하나의 압축 파라미터를 포함하여 상기 인코딩 영상을 획득할 수 있다.

[0018] 상기 영상 압축 성능 개선 방법은 상기 인코딩 영상을 인가받아 디코딩하여 디코딩 영상을 획득하고, 상기 인코딩 영상에 포함된 상기 적어도 하나의 압축 파라미터를 추출하는 단계; 상기 적어도 하나의 압축 파라미터에 따라 상기 디코딩 영상에 대한 후처리를 수행하는 인공 신경망의 후처리 가중치를 조절하는 단계; 및 가중치가 조절된 인공 신경망으로 상기 디코딩 영상에 대해 후처리를 수행하여 복원 영상을 획득하는 단계를 더 포함할 수 있다.

[0019] 상기 가중치를 조절하는 단계는 상기 후처리를 수행하는 인공 신경망의 다수의 연산 레이어 각각에 포함된 메인 커널과 적어도 하나의 서브 커널을 구성하는 가중치를 적어도 상기 적어도 하나의 압축 파라미터에 따른 비율로 혼합하여 각 연산 레이어의 가중치를 조절할 수 있다.

[0020] 상기 복원 영상을 획득하는 단계는 가중치가 조절된 인공 신경망의 다수의 연산 레이어로 상기 입력 영상에 대해 신경망 연산하여 특징맵을 출력하고, 상기 적어도 하나의 압축 파라미터에 스케일 파라미터가 포함되어 있으면, 상기 스케일 파라미터에 따라 스케일 업 레이어가 특징맵을 업 스케일링하여 상기 복원 영상을 획득할 수 있다.

[0021] 실시예에 따른 영상 압축 성능 개선 장치는 하나 이상의 프로세서; 및 상기 하나 이상의 프로세서에 의해 실행되는 하나 이상의 프로그램들을 저장하는 메모리를 구비한 장치로서, 상기 프로세서는 인가된 입력 영상에 대해 신경망 연산하여 상기 입력 영상의 특성 따른 적어도 하나의 압축 파라미터를 획득하고, 상기 적어도 하나의 압축 파라미터에 따라 상기 입력 영상에 대한 전처리를 수행하는 인공 신경망의 가중치를 조절하고, 가중치가 조절된 인공 신경망으로 상기 입력 영상에 대한 전처리를 수행하여 압축 영상을 획득하며, 상기 압축 영상을 인코딩하여 인코딩 영상을 획득한다.

발명의 효과

[0022] 따라서, 실시예에 따른 영상 압축 성능 개선 방법 및 장치는 입력 영상의 특성을 분석하여, 입력 영상의 특성에 따라 적응적으로 입력 영상의 시간 및 공간 해상도를 압축 및 복원하므로, 압축률을 크게 향상시킬 수 있을 뿐만 아니라 향상된 압축률에도 복원 영상의 품질 저하를 최대한 억제할 수 있다.

도면의 간단한 설명

[0023] 도 1은 일 실시예에 따른 영상 압축 성능 개선 장치에서 수행되는 동작에 따라 구분된 구성을 나타낸다.

도 2는 도 1의 전처리 네트워크와 후처리 네트워크의 개략적 구성을 나타낸다.

도 3은 도 2의 전처리 네트워크와 후처리 네트워크의 다수의 연산 레이어의 상세 구성의 일 예를 나타낸다.

도 4는 도 4의 파라미터 필터의 상세 구성의 일 예를 나타낸다.

도 5는 일 실시예에 따른 영상 압축 성능 개선 방법을 나타낸다.

도 6은 일 실시예에 따른 컴퓨팅 장치를 포함하는 컴퓨팅 환경을 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0024] 이하, 도면을 참조하여 일 실시예의 구체적인 실시형태를 설명하기로 한다. 이하의 상세한 설명은 본 명세서에서 기술된 방법, 장치 및/또는 시스템에 대한 포괄적인 이해를 돕기 위해 제공된다. 그러나 이는 예시에 불과하며 본 발명은 이에 제한되지 않는다.

[0025] 일 실시예들을 설명함에 있어서, 본 발명과 관련된 공지기술에 대한 구체적인 설명이 일 실시예의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하기로 한다. 그리고, 후술되는 용어들은 본 발명에서의 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. 상세한 설명에서 사용되는 용어는 단지 일 실시예들을 기술하기 위한 것이며, 결코 제한적이지는 안 된다. 명확하게 달리 사용되지 않는 한, 단수 형태의 표현은 복수 형태의 의미를 포함한다. 본 설명에서, "포함" 또는 "구비"와 같은 표현은 어떤 특성들, 숫자들, 단계들, 동작들, 요소들, 이들의 일부 또는 조합을 가리키기 위한 것이며, 기술된 것 이외에 하나 또는 그 이상의 다른 특성, 숫자, 단계, 동작, 요소, 이들의 일부 또는 조합의 존재 또는 가능성을 배제하도록 해석되어서는 안 된다. 또한, 명세서에 기재된 "...부", "...기", "모듈", "블록" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.

[0026] 도 1은 일 실시예에 따른 영상 압축 성능 개선 장치에서 수행되는 동작에 따라 구분된 구성을 나타낸다.

[0027] 도시된 실시예에서, 각 구성들은 이하에 기술된 것 이외에 상이한 기능 및 능력을 가질 수 있고, 이하에 기술되지 것 이외에도 추가적인 구성을 포함할 수 있다. 또한, 일 실시예에서, 각 구성은 물리적으로 구분된 하나 이상의 장치를 이용하여 구현되거나, 하나 이상의 프로세서 또는 하나 이상의 프로세서 및 소프트웨어의 결합에

의해 구현될 수 있으며, 도시된 예와 달리 구체적 동작에 있어 명확히 구분되지 않을 수 있다.

- [0028] 그리고 도 1에 도시된 영상 압축 성능 개선 장치는 하드웨어, 펌웨어, 소프트웨어 또는 이들의 조합에 의해 로직회로 내에서 구현될 수 있고, 범용 또는 특정 목적 컴퓨터를 이용하여 구현될 수도 있다. 장치는 고정배선형(Hardwired) 기기, 필드 프로그램 가능한 게이트 어레이(Field Programmable Gate Array, FPGA), 주문형 반도체(Application Specific Integrated Circuit, ASIC) 등을 이용하여 구현될 수 있다. 또한, 장치는 하나 이상의 프로세서 및 컨트롤러를 포함한 시스템온칩(System on Chip, SoC)으로 구현될 수 있다.
- [0029] 뿐만 아니라 영상 압축 성능 개선 장치는 하드웨어적 요소가 마련된 컴퓨팅 장치 또는 서버에 소프트웨어, 하드웨어, 또는 이들의 조합하는 형태로 탑재될 수 있다. 컴퓨팅 장치 또는 서버는 각종 기기 또는 유무선 통신망과 통신을 수행하기 위한 통신 모듈 등의 통신장치, 프로그램을 실행하기 위한 데이터를 저장하는 메모리, 프로그램을 실행하여 연산 및 명령하기 위한 마이크로프로세서 등을 전부 또는 일부 포함한 다양한 장치를 의미할 수 있다.
- [0030] 도 1을 참조하면, 일 실시예에 따른 영상 압축 성능 개선 장치는 영상 압축 모듈(100) 및 영상 복원 모듈(200)을 포함할 수 있다. 여기서 영상 압축 모듈(100)과 영상 복원 모듈(200)은 각각 영상 압축 장치와 영상 복원 장치로 구분되어 구현될 수 있다. 이 경우 영상 압축 장치와 영상 복원 장치 각각은 통신 모듈을 더 포함하여 영상 전송 장치와 영상 수신 장치로 동작할 수도 있다.
- [0031] 영상 압축 모듈(100)은 영상을 획득하고, 획득된 영상을 저장 또는 전송하기 위해 압축하며, 영상 복원 모듈(200)은 압축되어 저장되거나 전송된 영상을 인가받아 복원하는 역할을 수행한다.
- [0032] 특히 실시예에서 영상 압축 모듈(100)은 인공 신경망을 이용하여 획득된 영상의 특성을 판별하고, 판별된 특성에 따라 적응적으로 영상의 스케일이나 압축율을 조절하여 압축을 수행하며, 영상 복원 모듈(200) 또한 영상 압축 모듈(100)이 판별한 영상 특성에 따라 적용된 영상의 스케일이나 압축률을 반영하여 복원을 수행한다. 따라서 각 영상의 특성에 따라 상이한 스케일 및 압축률로 압축 및 복원을 수행함으로써, 압축률을 향상시킬 수 있을 뿐만 아니라 복원된 영상의 품질이 저하되는 것을 최대한 억제할 수 있다. 또한 영상 압축 성능 개선 장치는 기존의 레거시 코덱을 최대한 활용할 수 있도록 함으로써, 영상 표준 영상 압축 기법과 최대한 호환 가능하도록 한다.
- [0033] 구체적으로 영상 압축 모듈(100)은 영상 획득 모듈(110), 영상 특성 추정 모듈(120), 전처리 네트워크(130) 및 인코더(140)를 포함할 수 있다.
- [0034] 영상 획득 모듈(110)은 압축 대상이 되는 입력 영상을 획득한다. 영상 획득 모듈(110)은 저장 또는 전송을 위해 데이터량이 저감되어야 하는 입력 영상을 획득하며, 이때 획득되는 입력 영상은 단일 프레임의 정지 영상이거나 다수의 프레임으로 구성된 동영상일 수도 있다. 영상 획득 모듈(110)은 입력 영상을 직접 취득하는 카메라 모듈로 구현되거나, 취득된 입력 영상이 저장된 메모리와 같은 컴퓨터 판독 가능한 저장 매체 등으로 구현될 수도 있다.
- [0035] 영상 특성 추정 모듈(120)은 영상 획득 모듈(110)에서 획득된 입력 영상에 대해 학습된 인공 신경망으로 신경망 연산을 수행하여, 입력 영상의 특성을 추정하고, 추정된 영상 특성에 따른 압축 파라미터를 획득한다. 압축 파라미터는 입력 영상에 대해 적응적으로 서로 다른 압축률로 압축할 수 있도록 하기 위해 결정되는 파라미터로서, 입력 영상의 특성을 고려하여 결정된다.
- [0036] 입력 영상이 매우 단조로운 구조의 영상이거나 단색의 평면 영상 등과 같이 단순 구조의 영상인 경우와 다수의 객체와 다양한 배경 및 색상이 포함된 복잡한 구조의 영상은 영상 특성이 매우 상이하다고 할 수 있다. 그리고 이와 같이 영상 특성이 서로 매우 상이한 영상을 동일한 압축률로 압축하고자 하는 경우, 단순 영상을 기준으로 압축을 수행하면 압축율을 크게 향상시킬 수 있지만 복원된 복잡한 영상의 품질이 크게 저하되는 문제가 발생한다. 그에 반해, 복잡한 영상을 기준으로 압축을 수행하면 영상 품질을 유지할 수 있으나, 압축률이 매우 낮아져 영상 압축의 의미가 퇴색된다.
- [0037] 이에 도 1의 실시예에서는 영상 특성 모듈(120)이 영상의 특성을 사전에 판별하여 압축 파라미터를 획득하고, 획득된 압축 파라미터에 따라 전처리 네트워크(130)가 서로 상이하게 입력 영상을 압축하도록 한다. 이때 영상 특성 모듈(120)이 영상의 특성에 따라 획득하는 압축 파라미터는 다양하게 설정될 수 있으나, 여기서는 일 예로 영상 특성 모듈(120)이 압축 파라미터로서 스케일 파라미터(β)와 복잡도 파라미터(α)를 획득하는 것으로 가정한다. 여기서 각 압축 파라미터(α , β)는 $[0, 1]$ 사이의 범위의 실수 값으로 획득될 수 있다.

- [0038] 도 1에서는 영상 특성 추정 모듈(120)이 압축 파라미터로서 스케일 파라미터(β)와 복잡도 파라미터(α)를 획득하는 것으로 가정하였으므로, 영상 특성 추정 모듈(120)은 인공 신경망으로 구현되는 스케일 결정 네트워크(121)와 복잡도 판별 네트워크(122)를 구비하는 것으로 도시하였다. 그러나, 영상 특성 추정 모듈(120)은 획득하고자 하는 영상의 특성 종류에 따라 구분되는 압축 파라미터를 각각 추출하기 위한 별도의 인공 신경망을 구비할 수 있다.
- [0039] 스케일 결정 네트워크(121)는 입력 영상을 인가받아 신경망 연산하여, 입력 영상의 해상도를 조절하기 위한 스케일 파라미터(β)를 결정한다. 입력 영상은 영상 내에 포함된 배경이나 객체 등의 구조적 형태 등에 따라 복원 가능한 수준에서 다운 스케일링 될 수 있다. 즉 입력 영상의 해상도보다 저해상도의 영상으로 스케일 다운되어도 입력 영상의 해상도로 복원된 복원 영상의 품질이 입력 영상의 품질과 크게 차이가 발생하지 않을 수 있다. 다만 상기한 바와 같이, 단일 색상이나 단순 구조 객체가 포함된 영상의 경우, 스케일 다운 비율을 크게 높이더라도 원본인 입력 영상과 큰 차이가 없는 복원 영상을 획득할 수 있는 반면, 복잡한 구조의 객체가 포함된 영상의 경우, 스케일 다운 비율을 높이면 복원 영상의 품질이 크게 저하될 수 있다. 예를 들어 3,080 X 2,106 크기를 갖는 UHD(Ultra High Definition) 해상도의 입력 영상을 동일하게 1,920 X 1,080 크기의 FHD(Full High Definition) 해상도로 압축하면, 복원된 영상은 영상의 특성에 따라 매우 상이한 영상 품질을 나타내게 된다.
- [0040] 이에 스케일 결정 네트워크(121)는 입력 영상에 대해 신경망 연산을 수행하여 입력 영상의 해상도에 따른 스케일 특징을 추정하고, 추정된 스케일 특징에 따라 입력 영상에 적합한 스케일 파라미터를 추정할 수 있다. 이때 스케일 결정 네트워크(121)는 가로 및 세로 방향에서 균일한 비율의 스케일 파라미터(β)를 획득하도록 구성될 수 있으나, 경우에 따라서는 가로 및 세로 방향에 대해 각각 스케일 파라미터(β_1 , β_2)를 획득하도록 구성될 수도 있다.
- [0041] 이는 입력 영상의 구성에 따라 가로 방향 및 세로 방향에 대해 상이한 특성이 나타날 수 있기 때문이다. 일 예로 가로수 길이나 가로등에 대한 영상인 경우, 세로 방향에 비해 가로 방향에서 더 큰 비율로 스케일 다운되더라도, 이후 용이하게 고품질의 영상을 복원할 수 있다. 반면, 세로 방향으로 더 큰 비율로 스케일 다운되면 고품질의 영상을 복원하기 어렵다. 이에 스케일 결정 네트워크(121)는 압축 성능을 향상시키면서도 이후 용이하게 영상을 복원할 수 있도록 가로 및 세로 방향에 각각에 대한 스케일 파라미터(β_1 , β_2)를 구분하여 획득할 수도 있다.
- [0042] 한편, 복잡도 판별 네트워크(122)는 입력 영상에 대해 신경망 연산을 수행하여 입력 영상의 각 프레임 내 및 인접한 프레임간 복잡도를 추정하고, 추정된 복잡도에 따라 복잡도 파라미터(α)를 획득한다. 영상의 복잡도는 영상의 해상도에 따른 스케일과도 연관이 되지만, 스케일과 별개로 영상 자체의 특성에도 큰 영향을 받는다. 예로서 단색의 벽면을 촬영한 영상과 숲이나 많은 사람이 모인 광장을 촬영한 영상의 복잡도는 스케일과 무관하게 프레임내 복잡도가 서로 상이한 것으로 볼 수 있다. 즉 프레임내 픽셀간 유사도에 따라 프레임내 복잡도가 서로 상이하게 나타날 수 있다.
- [0043] 또한 다수의 프레임으로 구성된 영상의 경우, 움직이지 않는 객체를 촬영한 영상과 빠르게 이동하는 객체를 촬영한 영상은 프레임간 복잡도가 서로 상이하다. 즉 서로 다른 프레임 사이에서 픽셀의 변화에 따라 프레임간 복잡도가 서로 상이하게 나타날 수 있다.
- [0044] 이러한 영상의 복잡도에 따른 영상 압축 및 복원은 스케일 다운 및 스케일 업에 따른 영상 압축 및 복원과 별도로 구분될 필요가 있다. 이에 본 실시예에서 복잡도 판별 네트워크(122)는 스케일 결정 네트워크(121)와 별도로 구성되어 복잡도 파라미터(α)를 획득한다.
- [0045] 전처리 네트워크(130)는 미리 학습된 인공 신경망으로 구현되어, 인가되는 입력 영상에 대해 신경망 연산을 수행하여 입력 영상을 압축한다. 이때 전처리 네트워크(130)는 영상 특성 추정 모듈(120)에서 획득된 압축 파라미터에 따라 입력 영상을 서로 다르게 압축한다. 전처리 네트워크(130)는 압축 파라미터에 따라 인공 신경망의 가중치를 가변하여 입력 영상에 대한 압축을 수행한다. 여기서는 영상 특성 추정 모듈(120)이 압축 파라미터로서 스케일 파라미터(β)와 복잡도 파라미터(α)를 획득하는 것으로 가정하였으므로, 전처리 네트워크(130)는 복잡도 파라미터(α)에 따라 입력 영상에 대해 서로 상이하게 압축을 수행하면서 스케일 파라미터(β)에 따라 입력 영상을 스케일 다운시켜 압축 영상을 획득할 수 있다.
- [0046] 인코더(140)는 압축 영상을 인가받아, 지정된 방식으로 인코딩하여 인코딩 영상을 출력한다. 여기서 인코더(140)는 기존 레거시 코덱의 인코더로 구현될 수 있다. 인코더(140) 또한 영상을 압축하기 위해 이용되지만,

인공 신경망으로 구현되는 전처리 네트워크(130)와 상이한 방식으로 영상 압축을 수행한다. 다만 인공 신경망으로 구현되는 전처리 네트워크(130)는 입력 영상의 특성에 따라 인코더(140)의 영상 압축 효율성이 최대가 되도록 입력 영상에 대해 전처리를 수행하는 것으로 볼 수도 있다. 즉 전처리 네트워크(130)는 단순히 자체적으로 수행되는 영상 압축에 의해 입력 영상이 최대 압축되도록 하는 것이 아니라, 인코더(140)가 추가적으로 인코딩하여 출력되는 인코딩 영상의 압축률이 최대가 되도록 영상에 대한 압축을 수행하도록 학습될 수 있다.

[0047] 그리고 인코더(140)는 인코딩 영상에 압축 파라미터를 추가하여 출력할 수 있다. 레거시 코덱으로 영상을 인코딩하는 경우, 다양한 인코딩 정보가 인코딩 영상에 함께 포함되며, 이러한 인코딩 정보를 위한 데이터 공간 중에는 추후 추가될 수 있는 정보를 위해 데이터가 기록되지 않은 여백 공간이 존재한다. 이에 본 실시예에서 인코더(140)는 인코딩 영상의 여백 공간에 압축 파라미터의 비트열을 사이드 정보로서 추가하여 출력할 수 있다.

[0048] 한편 영상 복원 모듈(200)은 디코더(210) 및 후처리 네트워크(220)를 포함한다. 디코더(210)는 인코더(140)에서 출력된 인코딩 영상을 인가받아 디코딩하여 디코딩 영상을 출력한다. 디코더(210) 또한 인코더(140)와 마찬가지로 레거시 코덱의 디코더로 구현되어 인코딩 영상을 디코딩할 수 있다. 그리고 디코더(210)는 인코딩 영상을 디코딩하여 디코딩 영상을 획득하면서, 인코딩 영상에 포함된 압축 파라미터를 추출하여 후처리 네트워크(220)로 전달한다.

[0049] 후처리 네트워크(220)는 인공 신경망으로 구현되어 디코더(210)로부터 디코딩 영상과 압축 파라미터를 인가받고, 인가된 압축 파라미터에 따라 디코딩 영상에 대해 서로 상이하게 신경망 연산하여 복원 영상을 획득한다. 이때 후처리 네트워크(220)는 전처리 네트워크(130)와 마찬가지로 인가된 압축 파라미터에 따라 가중치를 조절하고, 조절된 가중치에 따라 디코딩 영상에 대해 신경망 연산을 수행하여 복원 영상을 획득한다. 그리고 후처리 네트워크(220) 또한 디코더(21)에서 디코딩된 디코딩 영상을 기반으로 영상 품질을 최대 복원할 수 있도록 학습된다.

[0050] 도 2는 도 1의 전처리 네트워크와 후처리 네트워크의 개략적 구성을 나타내고, 도 3은 도 2의 전처리 네트워크와 후처리 네트워크의 다수의 연산 레이어의 상세 구성의 일 예를 나타내며, 도 4는 도 4의 파라미터 필터의 상세 구성의 일 예를 나타낸다.

[0051] 도 2에서 (a)는 전처리 네트워크(130)의 개략적 구조를 나타내고, (b)는 후처리 네트워크(220)의 개략적 구조를 나타낸다. 도 2의 (a) 및 (b)에 도시된 바와 같이, 전처리 네트워크(130)와 후처리 네트워크(220)는 각각 다수의 연산 레이어(L1 ~ Ln)를 포함하여 신경망 연산을 수행하는 인공 신경망으로 구현될 수 있다. 그리고 전처리 네트워크(130)와 후처리 네트워크(220)가 인가된 영상에 대해 스케일 다운 또는 스케일 업을 수행하는 경우, 전처리 네트워크(130)와 후처리 네트워크(220)는 다수의 연산 레이어(L1 ~ Ln) 중 마지막 레이어(Ln) 이후에 스케일 다운 레이어(Lsd) 또는 스케일 업 레이어(Lsu)를 더 포함할 수 있다.

[0052] 다수의 연산 레이어(L1 ~ Ln) 각각은 네트워크로 입력되는 영상 또는 이전 레이어에서 출력된 특징맵에 대해 학습에 의해 지정된 가중치로 연산을 수행하여 출력한다. 여기서는 전처리 네트워크(130)와 후처리 네트워크(220)가 이미지 처리에 주로 이용되는 컨볼루션 네트워크(CNN)를 기반으로 구현되는 것으로 가정하며, 이에 다수의 연산 레이어(L1 ~ Ln) 각각은 입력되는 영상 또는 특징맵에 대해 가중치로 컨볼루션 연산을 수행하여 출력할 수 있다. 다만 본 실시예에서 다수의 연산 레이어(L1 ~ Ln) 각각은 압축 파라미터에 따라 가중치를 조절하면서 연산을 수행할 수 있다. 즉 일반적인 컨볼루션 네트워크의 연산 레이어와 상이하게 압축 파라미터에 따라 가중치를 가변하여 연산을 수행하여 출력할 수 있으며, 이에 동일한 영상 또는 특징맵이 인가되더라도 압축 파라미터에 따라 서로 상이한 가중치로 연산을 수행함으로써 상이한 특징맵이 출력되도록 할 수 있다.

[0053] 도 3에서는 일 예로 다수의 연산 레이어(L1 ~ Ln) 중 i번째 연산 레이어(Li)의 상세 구성을 도시한 것으로서, 다른 연산 레이어 또한 동일한 구성을 가질 수 있다. 도 3을 참조하면, 연산 레이어(Li)는 이전 연산 레이어(Li-1)에서 출력된 특징맵(FMi-1)을 인가받는다. 다만 연산 레이어(Li)가 전처리 네트워크(130)의 첫번째 연산 레이어(L1)인 경우, 연산 레이어(L1)는 입력 영상을 인가받는다. 그리고 연산 레이어(Li)가 후처리 네트워크(220)의 첫번째 연산 레이어(L1)인 경우, 연산 레이어(L1)는 디코딩 영상을 인가받는다.

[0054] 그리고 연산 레이어(Li)는 메인 커널(MK)과 다수의 파라미터 필터(PFL1 ~ PFLm)를 포함할 수 있다. 여기서 메인 커널(MK)과 다수의 파라미터 필터(PFL1 ~ PFLm) 각각은 학습에 의해 지정된 가중치를 갖는다. 이때 메인 커널(MK)은 압축 파라미터와 무관한 가중치를 갖는다. 그러나 다수의 파라미터 필터(PFL1 ~ PFLm)는 각각 지정된 압축 파라미터에 따른 비율로 메인 커널(MK)의 가중치 또는 이전 파라미터 필터에서 조절된 가중치를 조절하기 위한 가중치를 갖는다.

- [0055] 다수의 파라미터 필터(PFL1 ~ PFLm) 각각은 메인 커널(MK)의 가중치 또는 이전 파라미터 필터에서 조절된 가중치에 자신의 가중치를 압축 파라미터에 따른 비율로 가중함으로써 연산 레이어(Li)가 다수의 압축 파라미터에 따라 가변된 가중치로 연산을 수행할 수 있도록 한다. 연산 레이어(Li)에 포함되는 파라미터 필터(PFL1 ~ PFLm)의 개수(m)는 영상 특성 추정 모듈(120)이 획득하는 압축 파라미터의 개수에 따라 결정될 수 있다. 예를 들어 상기한 바와 같이, 영상 특성 추정 모듈(120)이 복잡도 파라미터(α) 및 스케일 파라미터(β)의 2개의 압축 파라미터를 획득하는 경우, 다수의 연산 레이어(L1 ~ Ln)는 각각 2개의 파라미터 필터(PFL1, PFL2)를 포함할 수 있다. 이때 2개의 파라미터 필터(PFL1, PFL2) 중 제1 파라미터 필터(PFL1)는 복잡도 파라미터(α)에 따라 메인 커널(MK)의 가중치를 조절하고, 제2 파라미터 필터(PFL2)는 스케일 파라미터(β)에 따라 제1 파라미터 필터(PFL1)에서 조절된 메인 커널(MK)의 가중치를 추가적으로 조절할 수 있다.
- [0056] 그러나 영상 특성 추정 모듈(120)이 스케일 파라미터(β)를 가로 방향 및 세로 방향에서 구분하여 2개의 스케일 파라미터(β_1, β_2)를 획득하는 경우, 다수의 연산 레이어(L1 ~ Ln) 각각은 2개의 스케일 파라미터 또한 구분하여 3개의 파라미터 필터(PFL1 ~ PFL3)를 포함할 수 있다.
- [0057] 다수의 파라미터 필터(PFL1 ~ PFLm) 각각은 메인 커널(MK) 또는 이전 파라미터 필터의 출력에 대해 저장된 서브 가중치를 압축 파라미터에 따른 가중비로 가중하여 조절한다. 즉 다수의 파라미터 필터(PFL1 ~ PFLm) 각각이 메인 커널(MK)을 구성하는 가중치를 대응하는 압축 파라미터에 따라 조절하는 역할을 수행한다.
- [0058] 도 4에서는 일 예로 다수의 파라미터 필터(PFL1 ~ PFLm) 중 제1 파라미터 필터(PFL1)의 상세 구성을 도시하였으며, 제1 파라미터 필터(PFL1)는 압축 파라미터 중 복잡도 파라미터(α)를 인가받는 것으로 가정하였다.
- [0059] 도 4를 참조하면, 파라미터 필터(PFL1)는 필터 전이 네트워크(410), 메인 경감 모듈(420), 서브 커널(430), 서브 가중 모듈(440) 및 커널 혼합 모듈(450)을 포함할 수 있다.
- [0060] 필터 전이 네트워크(410)는 메인 커널(MK) 또는 이전 파라미터 필터의 출력을 인가받아 서브 커널(430)에 연결한다. 필터 전이 네트워크(410)는 메인 커널(MK) 또는 이전 파라미터 필터의 출력을 서브 커널(430)로 전달하기 위한 구성으로 일 예로 스킵 커넥션으로 연결되는 2개의 1 X 1 컨볼루션 필터와 2개의 1 X 1 컨볼루션 필터 사이에 위치하는 활성화 함수인 PReLU(Parametric Rectified Linear Unit)로 구현될 수 있다.
- [0061] 메인 경감 모듈(420)은 메인 커널(MK) 또는 이전 파라미터 필터의 출력을 압축 파라미터(α)에 따른 비율($1-\alpha$)로 경감시킨다. 서브 커널(430)은 미리 학습에 의해 압축 파라미터(α)의 종류에 대응하여 지정된 서브 가중치로 구성된다. 서브 가중 모듈(440)은 서브 커널(430)을 구성하는 서브 가중치에 압축 파라미터(α)를 가중한다. 그리고 커널 혼합 모듈(450)은 압축 파라미터(α)에 따른 비율($1-\alpha$)로 경감된 경감 가중치와 압축 파라미터(α)가 가중된 서브 가중치를 가산하여 압축 파라미터(α)에 따라 가중치가 조절된 커널을 획득한다.
- [0062] 그리고 가중치가 조절된 커널의 가중치로 인가되는 영상 또는 특징맵에 대해 연산을 수행하여 출력한다.
- [0063] 한편 전처리 네트워크(130)와 후처리 네트워크(220)의 스케일 변환 레이어(Lsd) 또는 스케일 업 레이어(Lsu)는 압축 파라미터(α, β) 중 스케일 파라미터(β)를 인가받는다. 그리고 인가된 스케일 파라미터(β)에 따른 해상도를 갖도록 마지막 연산 레이어(Ln)에서 출력된 특징맵을 스케일 다운 또는 스케일 업한다. 이때 스케일 변환 레이어(Lsd) 또는 스케일 업 레이어(Lsu)는 [0, 1] 범위의 값을 갖는 스케일 파라미터(β)에 대응하여 [1, 2]의 범위의 실수값의 스케일로 특징맵을 스케일 다운 또는 스케일 업할 수 있다. 그리고 스케일 파라미터(β)가 가로 방향 및 세로 방향으로 구분된 경우, 스케일 변환 레이어(Lsd)와 스케일 업 레이어(Lsu) 또한 가로 및 세로 방향으로 구분된 2개씩의 스케일 변환 레이어(Lsd)와 스케일 업 레이어(Lsu)로 구성될 수 있다.
- [0064] 이때 스케일 변환 레이어(Lsd) 또는 스케일 업 레이어(Lsu) 또한 다수의 파라미터 필터(PFL1 ~ PFLm)를 포함하여, 압축 파라미터(α, β)에 따라 가중치를 조절한 후 스케일 다운 또는 스케일 업을 수행할 수 있다.
- [0065] 영상 특성 추정 모듈(120)과 전처리 네트워크(130) 및 후처리 네트워크(220)는 인공 신경망으로 구현되므로 미리 학습되어 가중치가 설정되어야 하며, 이에 영상 압축 성능 개선 장치는 학습 시에 학습 모듈(310)을 더 구비할 수 있다. 학습 모듈(310)은 학습 영상을 영상 압축 성능 개선 장치의 입력 영상으로 입력하고, 후처리 네트워크(220)에서 출력되는 복원 영상을 인가받아 분석하여 손실을 역전파함으로써, 영상 압축 성능 개선 장치에 포함된 인공 신경망을 학습시킬 수 있다. 다만 인코더(140)와 디코더(210)로 구성되는 코덱은 미분 불가능한 양자화 연산을 가지고 있어, 손실 역전파를 위한 그라디언트의 전달이 불가능하다. 이를 해결하기 위해, 학습 시에는 코덱을 모델링하여 인코더(140)와 디코더(210)를 대체할 수 있는 코덱 모사 네트워크를 이용하여 학습을 수행할 수 있다. 레거시 코덱에 대한 코덱 모사 네트워크는 공지된 기술로서 별도로 학습이 완료된 인공 신경

망이 이용될 수 있다. 따라서 여기서는 상세하게 설명하지 않는다.

- [0066] 학습 모듈(310)은 학습 영상과 복원 영상 사이의 차이에 따른 손실을 후처리 네트워크로 역전파하여 학습을 수행할 수 있다. 이때 학습 모듈(310)은 다수의 학습 영상을 최대 신호대잡음비(Peak Signal-to-noise ratio: PSNR)와 비트레이트(bitrate)에 따라 학습 영상의 효과적인 압축 스케일을 구분하고, 프레임내 복잡도 및 프레임간 복잡도에 따라 학습 영상의 복잡도를 구분하고, 구분된 학습 영상의 압축 파라미터를 레이블링할 수 있다.
- [0067] 학습 모듈(310)은 학습 영상을 이용한 학습 시에 영상 특성 추정 모듈(120)에서 획득된 압축 파라미터와 학습 영상에 레이블된 압축 파라미터 사이의 오차를 손실로 계산하여 역전파함으로써, 영상 특성 추정 모듈(120)을 학습시킬 수 있다.
- [0068] 다만 본 실시예의 영상 압축 성능 개선 장치에서는 메인 커널과 함께 서브 커널을 포함하고 있을 뿐만 아니라 메인 커널과 서브 커널의 가중치가 압축 파라미터에 따라 서로 다른 비율로 혼합된다. 또한 압축 파라미터 또한 영상에 따라 가변된다. 따라서 다양한 압축 가능 스케일과 복잡도를 갖는 모든 학습 데이터를 이용하여 학습을 수행하는 경우, 인공 신경망의 커널 가중치가 수렴되기에 매우 오랜 시간이 소요되거나 수렴되지 않는 경우가 발생할 수 있다. 이와 같은 문제를 해소하고 학습 시간을 저감시키기 위해, 학습 모듈(310)은 우선 최소값을 갖는 압축 스케일과 복잡도를 갖는 학습 영상을 기반으로 학습을 수행한다. 이 경우, 학습 데이터의 압축 파라미터(α , β)가 모두 0의 값으로 레이블되며, 이에 따라 압축 파라미터(α , β)의 영향이 배제될 뿐만 아니라, 각 인공 신경망의 연산 레이어(L1 ~ Lm)에서도 다수의 파라미터 필터(PFL1 ~ PFLm)가 비활성화된다. 따라서 학습 모듈(310)은 각 인공 신경망의 메인 커널(MK)의 가중치를 학습시킬 수 있다. 이후, 다수의 압축 파라미터(α , β) 중 순차적으로 하나만이 최대값을 갖거나, 하나만이 최소값을 갖는 학습 영상들을 입력하여 학습을 수행함으로써, 다수의 압축 파라미터 각각에 따른 서브 커널의 가중치를 학습시킬 수 있다. 추가적으로 학습 모듈(310)은 레이블된 파라미터에 무관하게 학습 영상을 입력시킴으로써 최종적으로 커널의 가중치에 대한 세밀한 조절을 수행할 수 있다.
- [0069] 도 5는 일 실시예에 따른 영상 압축 성능 개선 방법을 나타낸다.
- [0070] 도 1 내지 도 4를 참조하여 도 5의 영상 압축 성능 개선 방법을 설명하면, 우선 저장하거나 전송할 입력 영상을 획득한다(51). 그리고 학습된 인공 신경망이 입력 영상을 인가받아 신경망 연산하여, 입력 영상의 특징을 추출하여 압축 파라미터를 획득한다(51). 이때 서로 다른 다수의 인공 신경망을 이용하여 입력 영상에 대한 서로 다른 특징을 추출함으로써 다수의 압축 파라미터를 획득할 수 있으며, 일 예로 영상의 손실을 최소화하면서 압축 가능한 해상도 특성을 나타내는 스케일 파라미터와 영상의 프레임 내 및 프레임간 복잡도를 나타내는 복잡도 파라미터를 획득할 수 있다.
- [0071] 다수의 압축 파라미터가 획득되면, 획득된 압축 파라미터를 이용하여 입력 영상에 대해 전처리를 수행하는 인공 신경망의 가중치를 조절한다(53). 이때 전처리를 수행하는 인공 신경망은 다수의 연산 레이어(L1 ~ Ln)와 적어도 하나의 스케일 다운 레이어(Lsd)를 포함하며, 다수의 연산 레이어(L1 ~ Ln) 각각은 압축 파라미터에 무관한 가중치를 갖는 하나의 메인 커널과 영상 특성별로 구분되어 획득된 다수의 압축 파라미터 각각에 따른 다수의 서브 커널을 포함할 수 있다. 그리고 다수의 연산 레이어(L1 ~ Ln) 각각은 압축 파라미터로 지정된 비율에 따라 메인 커널과 다수의 서브 커널이 혼합된 가중치를 갖는다.
- [0072] 압축 파라미터에 따라 인공 신경망의 다수의 연산 레이어(L1 ~ Ln)의 가중치가 조절되면, 전처리를 수행하는 인공 신경망은 입력 영상을 인가받아 조절된 가중치로 신경망 연산을 수행하여 특징맵을 출력하고, 적어도 하나의 스케일 다운 레이어(Lsd)는 압축 파라미터 중 스케일 파라미터에 따라 신경망 연산된 특징맵을 인가받아 스케일 다운하는 전처리를 수행하여 압축 영상을 획득한다(54).
- [0073] 압축 영상은 레거시 코덱의 인코더에서 인코딩되어 인코딩 영상으로 획득된다(55). 이때 인코더는 압축 파라미터를 인코딩 영상에 추가하여 포함시킬 수 있다.
- [0074] 인코딩 영상은 레거시 코덱의 디코더에 인가되고, 디코더는 인가된 인코딩 영상을 디코딩하여 디코딩 영상을 획득한다(56). 그리고 인코딩 영상에 포함된 압축 파라미터를 추출한다(57).
- [0075] 압축 파라미터가 추출되면, 추출된 압축 파라미터를 이용하여 디코딩 영상에 대해 후처리를 수행하는 인공 신경망의 가중치를 조절한다(58). 이때 전처리를 수행하는 인공 신경망과 유사하게 후처리를 수행하는 인공 신경망은 다수의 연산 레이어(L1 ~ Ln)와 적어도 하나의 스케일 업 레이어(Lsu)를 포함한다. 이에 다수의 연산 레이어(L1 ~ Ln) 각각은 압축 파라미터로 지정된 비율에 따라 메인 커널과 다수의 서브 커널이 혼합된 가중치를 갖

는다.

- [0076] 압축 파라미터에 따라 인공 신경망의 다수의 연산 레이어(L1 ~ Ln)의 가중치가 조절되면, 후처리를 수행하는 인공 신경망은 디코딩 영상을 인가받아 조절된 가중치로 신경망 연산을 수행하여 특징맵을 출력하고, 적어도 하나의 스케일 업 레이어(Lsu)는 압축 파라미터 중 스케일 파라미터에 따라 신경망 연산된 특징맵을 인가받아 스케일 업하는 후처리를 수행하여 복원 영상을 획득한다(59).
- [0077] 도 5에서는 각각의 과정을 순차적으로 실행하는 것으로 기재하고 있으나 이는 예시적으로 설명한 것에 불과하고, 이 분야의 기술자라면 본 발명의 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 도 5에 기재된 순서를 변경하여 실행하거나 또는 하나 이상의 과정을 병렬적으로 실행하거나 다른 과정을 추가하는 것으로 다양하게 수정 및 변형하여 적용 가능하다.
- [0078] 도 6은 일 실시예에 따른 컴퓨팅 장치를 포함하는 컴퓨팅 환경을 설명하기 위한 도면이다.
- [0079] 도시된 실시예에서, 각 구성 요소들은 이하에 기술된 것 이외에 상이한 기능 및 능력을 가질 수 있고, 이하에 기술되지 않은 것 이외에도 추가적인 구성 요소를 포함할 수 있다. 도시된 컴퓨팅 환경(60)은 컴퓨팅 장치(61)를 포함한다. 일 실시예에서, 컴퓨팅 장치(61)는 도 1에 도시된 영상 압축 성능 개선 장치에 포함된 하나 이상의 구성 요소일 수 있다.
- [0080] 컴퓨팅 장치(61)는 적어도 하나의 프로세서(62), 컴퓨터 판독 가능 저장매체(63) 및 통신 버스(65)를 포함한다. 프로세서(62)는 컴퓨팅 장치(61)로 하여금 앞서 언급된 예시적인 실시예에 따라 동작하도록 할 수 있다. 예컨대, 프로세서(62)는 컴퓨터 판독 가능 저장매체(63)에 저장된 하나 이상의 프로그램들(64)을 실행할 수 있다. 상기 하나 이상의 프로그램들(64)은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 상기 컴퓨터 실행 가능 명령어는 프로세서(62)에 의해 실행되는 경우 컴퓨팅 장치(61)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.
- [0081] 통신 버스(65)는 프로세서(62), 컴퓨터 판독 가능 저장매체(63)를 포함하여 컴퓨팅 장치(61)의 다른 다양한 구성 요소들을 상호 연결한다.
- [0082] 컴퓨팅 장치(61)는 또한 하나 이상의 입출력 장치(68)를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(66) 및 하나 이상의 통신 인터페이스(67)를 포함할 수 있다. 입출력 인터페이스(66) 및 통신 인터페이스(67)는 통신 버스(65)에 연결된다. 입출력 장치(68)는 입출력 인터페이스(66)를 통해 컴퓨팅 장치(61)의 다른 구성 요소들에 연결될 수 있다. 예시적인 입출력 장치(68)는 포인팅 장치(마우스 또는 트랙패드 등), 키보드, 터치 입력 장치(터치패드 또는 터치스크린 등), 음성 또는 소리 입력 장치, 다양한 종류의 센서 장치 및/또는 촬영 장치와 같은 입력 장치, 및/또는 디스플레이 장치, 프린터, 스피커 및/또는 네트워크 카드와 같은 출력 장치를 포함할 수 있다. 예시적인 입출력 장치(68)는 컴퓨팅 장치(61)를 구성하는 일 구성 요소로서 컴퓨팅 장치(61)의 내부에 포함될 수도 있고, 컴퓨팅 장치(61)와는 구별되는 별개의 장치로 컴퓨팅 장치(61)와 연결될 수도 있다.
- [0083] 이상에서 대표적인 실시예를 통하여 본 발명에 대하여 상세하게 설명하였으나, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

부호의 설명

- [0084]
- | | |
|------------------|------------------|
| 100: 영상 압축 모듈 | 110: 영상 획득 모듈 |
| 120: 영상 특성 추정 모듈 | 121: 스케일 결정 네트워크 |
| 122: 복잡도 판별 네트워크 | 130: 전처리 네트워크 |
| 140: 인코더 | 200: 영상 복원 모듈 |
| 210: 디코더 | 220: 후처리 네트워크 |
| 410: 필터 전이 네트워크 | 420: 메인 경감 모듈 |
| 430: 서브 커널 | 440: 서브 가중 모듈 |
| 450: 커널 혼합 모듈 | Ln1 ~ Ln: 연산 레이어 |

Lsd: 스케일 다운 레이어

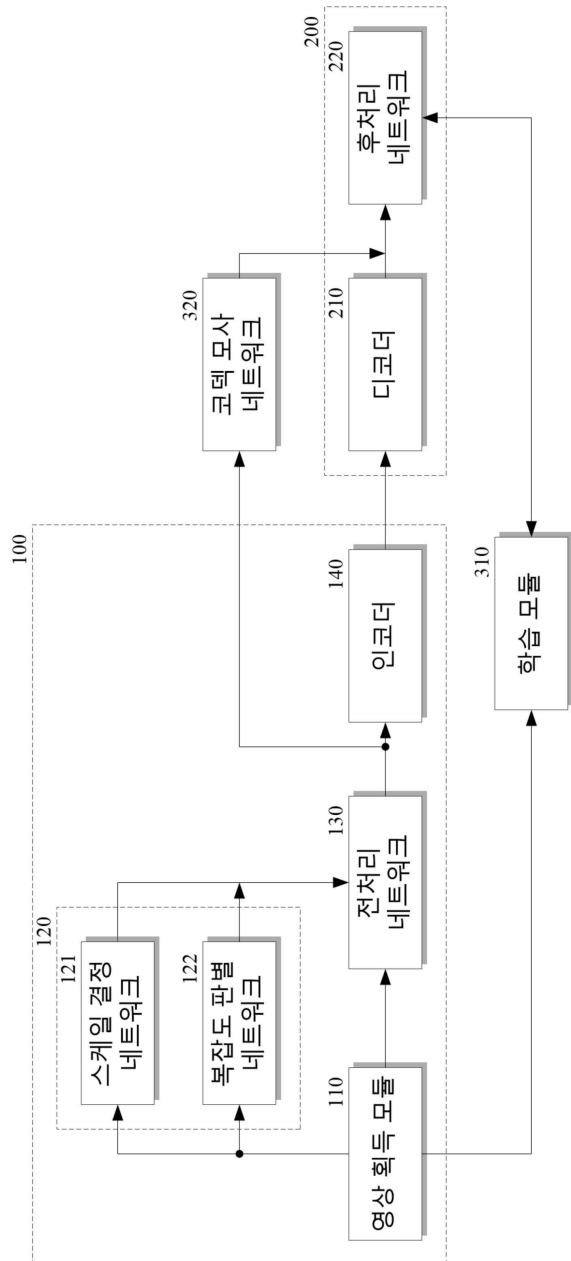
Lsu: 스케일 업 레이어

MK: 메인 커널

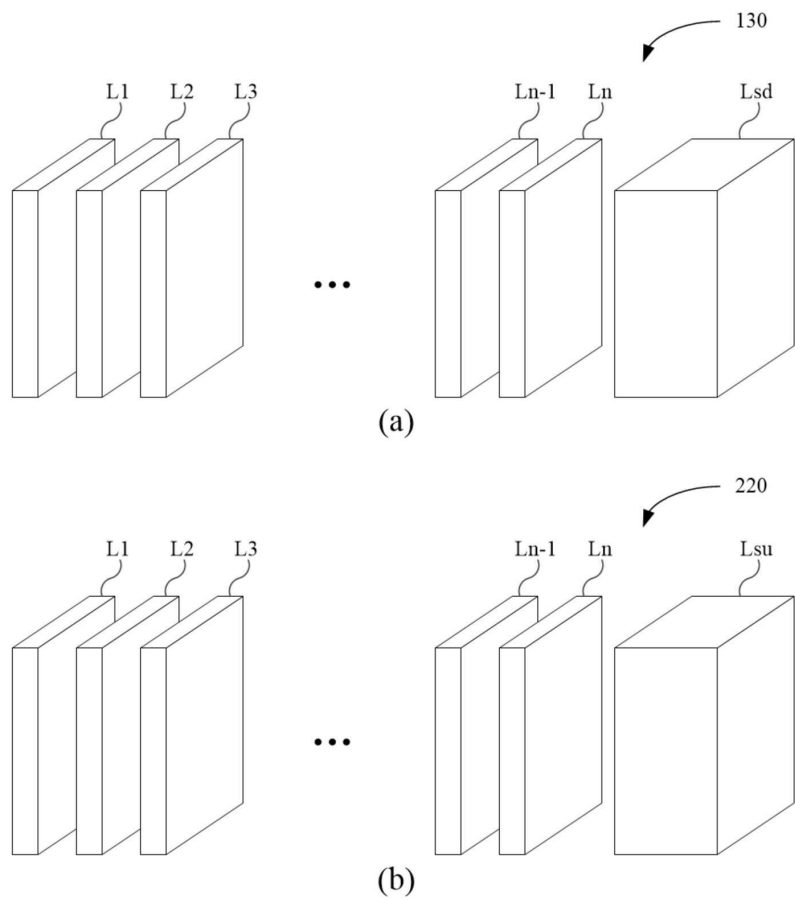
PFL1 ~ PFL_m: 파라미터 필터

도면

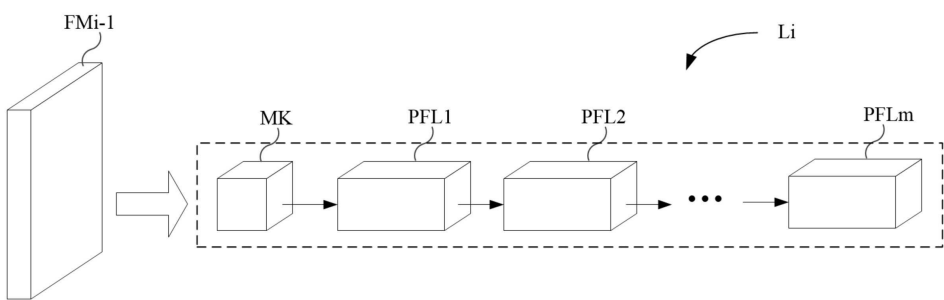
도면1



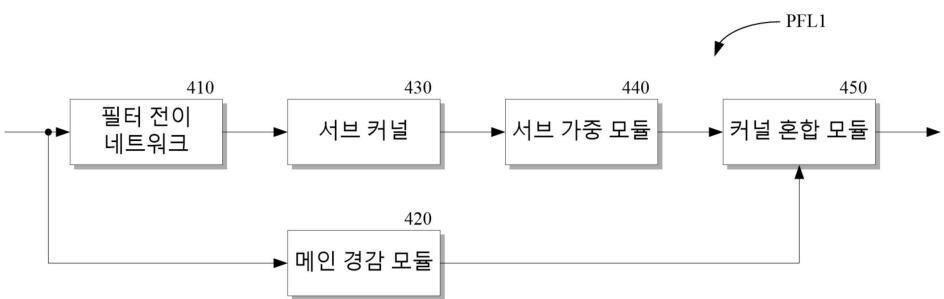
도면2



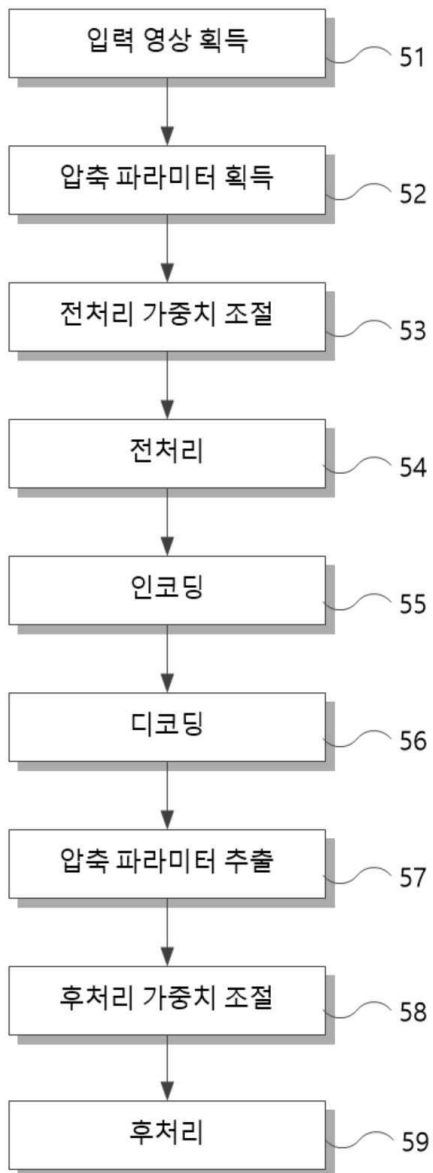
도면3



도면4



도면5



도면6

60

