



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2021년04월27일

(11) 등록번호 10-2245682

(24) 등록일자 2021년04월22일

(51) 국제특허분류(Int. Cl.)

H04N 19/85 (2014.01) G06N 3/04 (2006.01)

G06N 3/08 (2006.01) H04N 19/44 (2014.01)

(52) CPC특허분류

H04N 19/85 (2015.01)

G06N 3/04 (2013.01)

(21) 출원번호 10-2019-0143346

(22) 출원일자 2019년11월11일

심사청구일자 2019년11월11일

(56) 선행기술조사문헌

KR1020180100976 A\*

(뒷면에 계속)

(73) 특허권자

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

이상윤

서울특별시 서초구 청계산로7길 43, 504동 1302호(신원동, 서초포레스타5단지)

김태오

서울특별시 중랑구 동일로 752, 109동 304호(중화동, 중화한신아파트)

(뒷면에 계속)

(74) 대리인

민영준

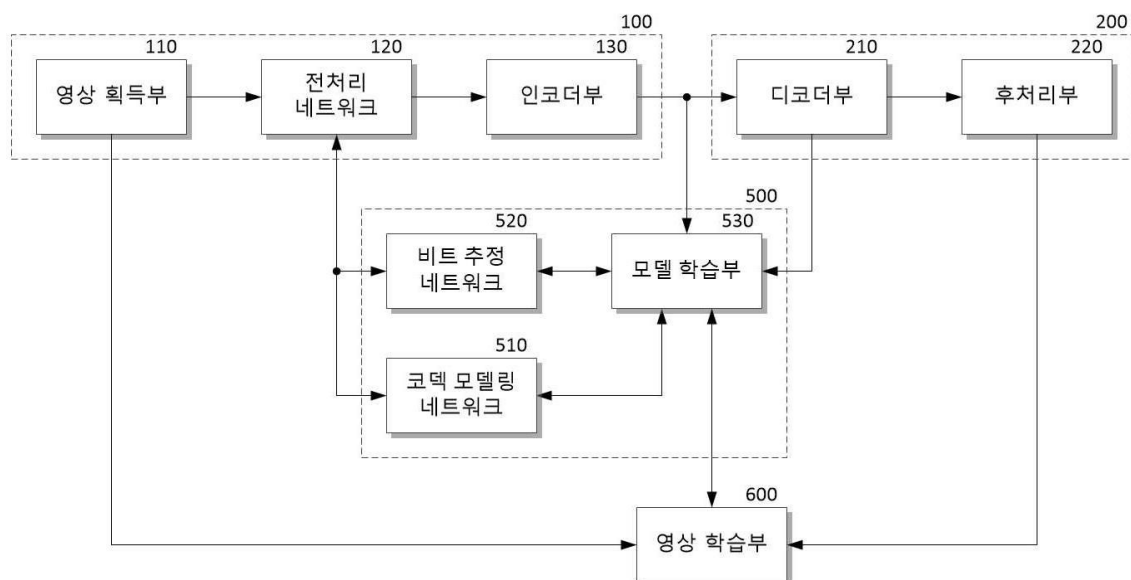
전체 청구항 수 : 총 12 항

심사관 : 조우연

(54) 발명의 명칭 영상 압축 장치, 이의 학습 장치 및 방법

**(57) 요약**

본 발명은 로우 데이터 영상을 획득하는 영상 획득부, 로우 데이터 영상을 인가받아 미리 학습된 패턴 추정 방식에 따라 전처리하는 전처리 네트워크 및 전처리된 영상을 인가받아 기지정된 표준 압축 기법에 따라 압축하여 압축 영상을 출력하는 인코더부를 포함하고, 전처리 네트워크는 학습 시에 추가되며, 인공 신경망으로 구현되고 인코더부와 압축 영상을 인가받아 디코딩하여 복호 영상을 획득하는 디코더부를 포함하는 표준 코덱부를 모의하도록 미리 학습된 코덱 모델링부에서 출력되는 모의 복호 영상을 기반으로 획득된 복원 영상과 로우 데이터 영상을 비교하여 획득되는 복원 오차가 코덱 모델링부를 통해 역전파되어 미리 학습된 영상 압축 장치, 이의 학습 장치 및 방법을 제공할 수 있다.

**대표도**

(52) CPC특허분류

*G06N 3/084* (2013.01)

*H04N 19/44* (2015.01)

(72) 발명자

**손한빈**

서울특별시 서대문구 신촌로11길 49, 101호(창천동)

**이형민**

서울특별시 강북구 오현로 56, 101동 404호(미아동, 꿈의숲 롯데캐슬)

(56) 선행기술조사문헌

W02019009449 A1\*

조승현 등, “딥 러닝 기반의 이미지와 비디오 압축 기술 분석”, 방송공학회논문지 제23권 제3호, p383-394, (2018.05.31.)\*

KR1020170059040 A

KR1020190043930 A

KR102022648 B1

W02017222140 A1

KR100286108 B1

\*는 심사관에 의하여 인용된 문헌

---

## 명세서

### 청구범위

#### 청구항 1

로우 데이터 영상을 획득하는 영상 획득부;

상기 로우 데이터 영상을 인가받아 미리 학습된 패턴 추정 방식에 따라 전처리하는 전처리 네트워크; 및

전처리된 영상을 인가받아 기지정된 표준 압축 기법에 따라 압축하여 압축 영상을 출력하는 인코더부를 포함하고,

상기 전처리 네트워크는

학습 시에 추가되며, 인공 신경망으로 구현되고 상기 인코더부와 상기 압축 영상을 인가받아 디코딩하여 복호 영상을 획득하는 디코더부를 포함하는 표준 코덱부를 모의하도록 미리 학습된 코덱 모델링부에서 출력되는 모의 복호 영상을 기반으로 획득된 복원 영상과 상기 로우 데이터 영상을 비교하여 획득되는 복원 오차가 상기 코덱 모델링부를 통해 역전파되어 미리 학습되어 있으며,

상기 코덱 모델링부는

상기 인코더부와 상기 디코더부를 포함하는 상기 표준 코덱부를 모의하도록 학습되어 압축 영상을 인가받아 복호 영상에 대응하는 모의 복호 영상을 출력하는 코덱 모델링 네트워크;

상기 전처리 네트워크에서 전처리된 영상을 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 전처리된 영상으로부터 상기 인코더부에서 출력되는 압축 영상의 비트수를 추정하는 비트 추정 네트워크; 및

상기 코덱 모델링 네트워크와 비트 추정 네트워크를 학습시키기 위한 모델 학습 시에 상기 디코더부에서 획득되는 복호 영상과 상기 모의 복호 영상을 비교하여 모델링 오차를 획득하여 상기 코덱 모델링 네트워크로 역전파하고, 상기 비트 추정 네트워크에서 추정된 비트수와 상기 인코더부에서 출력되는 압축 영상의 비트수를 비교하여 비트 오차를 획득하여 비트 추정 네트워크로 역전파하는 모델 학습부를 포함하는 영상 압축 장치.

#### 청구항 2

삭제

#### 청구항 3

제1 항에 있어서, 상기 영상 압축 장치는

상기 전처리 네트워크를 학습시키기 위한 전처리 학습 시에 추가되어 상기 로우 데이터 영상과 상기 복호 영상 또는 상기 모의 복호 영상을 기지정된 방식으로 후처리하여 획득되는 복원 영상을 비교하여 복원 오차를 획득하고, 획득된 복원 오차를 상기 코덱 모델링 네트워크와 상기 비트 추정 네트워크를 통해 상기 전처리 네트워크로 역전파하는 영상 학습부를 더 포함하는 영상 압축 장치.

#### 청구항 4

제3 항에 있어서, 상기 모델 학습부는

전처리 학습시에 상기 모델 학습에 의해 설정된 상기 코덱 모델링 네트워크와 상기 비트 추정 네트워크의 가중치를 고정하는 영상 압축 장치.

#### 청구항 5

제4 항에 있어서, 상기 영상 학습부는

상기 복호 영상 또는 상기 모의 복호 영상으로부터 상기 복원 영상을 획득하도록 패턴 추정 방식이 미리 학습된 인공 신경망으로 구현된 후처리 네트워크에 의해 후처리되는 경우, 상기 복원 오차를 상기 후처리 네트워크로

우선 역전파하고, 상기 코덱 모델링 네트워크 및 상기 비트 추정 네트워크를 거쳐 상기 전처리 네트워크로 역전파되도록 하는 영상 압축 장치.

#### 청구항 6

로우 데이터 영상을 인가받아 미리 학습된 패턴 추정 방식에 따라 전처리하는 전처리 네트워크와 전처리된 영상을 인가받아 기지정된 표준 압축 기법에 따라 압축하여 압축 영상을 출력하는 인코더부를 포함하는 영상 압축 장치의 학습 장치로서,

인공 신경망으로 구현되고 상기 인코더부와 상기 압축 영상을 인가받아 디코딩하여 복호 영상을 획득하는 디코더부를 포함하는 표준 코덱부를 모의하도록 미리 학습되어 상기 전처리된 영상으로부터 모의 복호 영상을 획득하는 코덱 모델링부; 및

상기 코덱 모델링부에서 출력되는 상기 모의 복호 영상을 기반으로 획득된 복원 영상과 상기 로우 데이터 영상을 비교하여 복원 오차를 획득하고, 획득된 복원 오차를 상기 코덱 모델링부를 통해 역전파하여 상기 전처리 네트워크를 학습시키는 영상 학습부를 포함하되,

상기 코덱 모델링부는

상기 인코더부와 상기 디코더부를 포함하는 상기 표준 코덱부를 모의하도록 학습되어 압축 영상을 인가받아 복호 영상에 대응하는 모의 복호 영상을 출력하는 코덱 모델링 네트워크;

상기 전처리 네트워크에서 전처리된 영상을 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 전처리된 영상으로부터 상기 인코더부에서 출력되는 압축 영상의 비트수를 추정하는 비트 추정 네트워크; 및

상기 코덱 모델링 네트워크와 비트 추정 네트워크를 학습시키기 위한 모델 학습 시에 상기 디코더부에서 획득되는 복호 영상과 상기 모의 복호 영상을 비교하여 모델링 오차를 획득하여 상기 코덱 모델링 네트워크로 역전파하고, 상기 비트 추정 네트워크에서 추정된 비트수와 상기 인코더부에서 출력되는 압축 영상의 비트수를 비교하여 비트 오차를 획득하여 비트 추정 네트워크로 역전파하는 모델 학습부를 포함하는 영상 압축 장치의 학습 장치.

#### 청구항 7

삭제

#### 청구항 8

제6 항에 있어서, 상기 영상 학습부는

상기 전처리 네트워크를 학습시키기 위한 전처리 학습 시에 상기 로우 데이터 영상과 상기 복호 영상 또는 상기 모의 복호 영상을 기지정된 방식으로 후처리하여 획득되는 복원 영상을 비교하여 복원 오차를 획득하고, 획득된 복원 오차를 상기 코덱 모델링 네트워크와 상기 비트 추정 네트워크를 통해 상기 전처리 네트워크로 역전파하는 영상 학습부를 더 포함하는 영상 압축 장치의 학습 장치.

#### 청구항 9

제8 항에 있어서, 상기 모델 학습부는

전처리 학습시에 상기 모델 학습에 의해 설정된 상기 코덱 모델링 네트워크와 상기 비트 추정 네트워크의 가중치를 고정하는 영상 압축 장치의 학습 장치.

#### 청구항 10

제9 항에 있어서, 상기 영상 학습부는

상기 복호 영상 또는 상기 모의 복호 영상으로부터 상기 복원 영상을 획득하도록 패턴 추정 방식이 미리 학습된 인공 신경망으로 구현된 후처리 네트워크에 의해 후처리되는 경우, 상기 복원 오차를 상기 후처리 네트워크로 우선 역전파하고, 상기 코덱 모델링 네트워크 및 상기 비트 추정 네트워크를 거쳐 상기 전처리 네트워크로 역전파하는 영상 압축 장치의 학습 장치.

## 청구항 11

로우 데이터 영상을 인가받아 미리 학습된 패턴 추정 방식에 따라 전처리하는 전처리 네트워크와 전처리된 영상을 인가받아 기지정된 표준 압축 기법에 따라 압축하여 압축 영상을 출력하는 인코더부를 포함하는 영상 압축 장치의 학습 방법에 있어서,

상기 인코더부와 상기 압축 영상을 인가받아 디코딩하여 복호 영상을 획득하는 디코더부를 포함하는 표준 코덱 부를 모의하도록 미리 학습된 인공 신경망을 이용하여 상기 전처리된 영상으로부터 모의 복호 영상을 획득하는 단계; 및

상기 모의 복호 영상을 기반으로 획득된 복원 영상과 상기 로우 데이터 영상을 비교하여 복원 오차를 획득하고, 획득된 복원 오차를 상기 인공 신경망을 통해 역전파하여 상기 전처리 네트워크를 학습시키는 단계를 포함하되,

상기 모의 복호 영상을 획득하는 단계 이전, 모델 학습 단계를 더 포함하고,

상기 모델 학습 단계는 상기 디코더부에서 획득되는 복호 영상과 상기 모의 복호 영상을 비교하여 모델링 오차를 획득하고, 추정된 비트수와 상기 인코더부에서 출력되는 압축 영상의 비트수를 비교하여 비트 오차를 획득하여 각각 대응하는 인공 신경망으로 역전파하는 영상 압축 장치의 학습 방법.

## 청구항 12

제11 항에 있어서, 상기 모의 복호 영상을 획득하는 단계는

패턴 추정 방식이 미리 학습된 상기 인공 신경망을 이용하여 압축 영상을 인가받아 복호 영상에 대응하는 모의 복호 영상을 획득하는 단계; 및

상기 모의 복호 영상을 획득하는 인공 신경망과 별도의 인공 신경망을 이용하여 미리 학습된 패턴 추정 방식에 따라 상기 전처리된 영상으로부터 상기 인코더부에서 출력되는 압축 영상의 비트수를 추정하는 단계를 포함하는 영상 압축 장치의 학습 방법.

## 청구항 13

삭제

## 청구항 14

제11 항에 있어서, 상기 전처리 네트워크를 학습시키는 단계는

상기 전처리 네트워크를 학습시키기 위한 전처리 학습 시에 상기 로우 데이터 영상과 상기 복호 영상 또는 상기 모의 복호 영상을 기지정된 방식으로 후처리하여 획득되는 복원 영상을 비교하여 복원 오차를 획득하는 단계; 및

획득된 복원 오차를 상기 인공 신경망들을 통해 상기 전처리 네트워크로 역전파하는 단계를 포함하는 영상 압축 장치의 학습 방법.

## 청구항 15

제14 항에 있어서, 상기 전처리 네트워크를 학습시키는 단계는

상기 복호 영상 또는 상기 모의 복호 영상으로부터 상기 복원 영상을 획득하도록 패턴 추정 방식이 미리 학습된 인공 신경망으로 구현된 후처리 네트워크에 의해 후처리되는 경우, 상기 복원 오차를 상기 후처리 네트워크로 우선 역전파하고, 상기 인공 신경망들을 거쳐 상기 전처리 네트워크로 역전파하는 영상 압축 장치의 학습 방법.

## 발명의 설명

## 기술 분야

본 발명은 영상 압축 장치, 이의 학습 장치 및 방법에 관한 것으로, 영상 압축 모델링을 이용하여 표준 호환 가능한 영상 압축 장치, 이의 학습 장치 및 방법에 관한 것이다.

[0001]

## 배경 기술

- [0002] 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.
- [0003] 영상 압축 기술은 크게 블록 기반 코딩 기법과 예측 기반 코딩 기법으로 구분된다. 블록 기반 코딩 기법은 일반적으로 영상을 일정 픽셀을 포함하는 블록 단위로 분할한 뒤, 주파수 변환(Fourier Transform), 양자화(Quantization) 및 엔트로피 코딩(Entropy Coding) 과정을 통해 인코딩하여 영상을 압축하게 된다. 구체적으로는 영상을 주파수 변환하여 저주파 성분과 고주파 성분으로 분리하고, 사람의 시각이 저주파 성분에 많이 영향을 받는다는 점을 고려하여 고주파 성분을 양자화를 통해서 나누어 준다. 즉 양자화를 통해 영상 데이터의 정보를 최소화한다. 그리고 엔트로피 코딩을 통해서 자주 발생하는 값은 적은 비트 기호를, 자주 발생하지 않는 값은 많은 비트 기호를 할당하여 최종적으로 이진 신호로 변환하게 된다.
- [0004] 블록 기반 코딩 기법으로 압축된 영상의 디코딩은 인코딩의 역순으로 진행하며, 이 과정에서 각 주파수 성분별 양자화 테이블(Quantization Table)이 디코딩 과정에서 필요로 하게 된다.
- [0005] 예측 기반 코딩 기법에서 변환-양자화-엔트로피 코딩의 과정은 블록 기반 코딩 기법과 동일하지만 변환 전, 예측(Prediction)이라는 과정이 추가된다. 예측 기반 코딩은 화면 내 많은 부분이 중복적으로 발생하는 점에서 착안하며, 특히 동영상의 경우에는 이전 프레임과의 상관 관계가 높다는 점을 이용한다. 예측 과정에서는 지금까지 압축된 영상 정보를 이용해서 현재 블록을 예측한다. 예측 후 그 차이값만 변환 과정으로 보내어, 원본 데이터를 그대로 변환하는 것 보다 적은 양의 정보만을 부호화한다.
- [0006] 예측 기반 코딩 기법의 디코딩 과정에서는 인코딩 과정과 동일한 예측 알고리즘을 사용하여 이진 신호와 합산하여 원본 영상을 추정하게 된다.
- [0007] 이러한 영상 압축의 성능은 원본 영상의 화질을 나타내는 최대 신호대 잡음비(PSNR: Peak Signal to Noise Ratio)을 최대한 유지하면서, 더 적은 이진 비트, 즉 높은 비트율(Bitrate)로 영상을 표현하여야 한다. 이 과정은 양자화 정도(Quantization-level)에 따라서 조절이 가능하며, 양자화를 더 강하게 하면 더 많은 정보가 손실되어 영상의 화질이 저하되고, 양자화를 약하게 하면 많은 정보가 유지되어 영상의 화질이 향상되는 트레이드 오프(Trade-off) 관계를 가지고 있으며, 영상 압축 성능은 최대 신호대 잡음비-비트율(PSNR-Bitrate)의 곡선에서 더 높은 지점에 위치하는지를 기준으로 판단하게 된다.
- [0008] 이러한 영상 압축 기술은 영상을 많은 미디어 매체들에서 호환성 등의 문제로 기본적으로 표준화되어 있다. 대표적으로 블록 기반 코딩 기법에서는 JPEG(Joint Photographic Experts Group), JPEG2000 등의 정지 영상 압축 표준이 규정되어 있으며, 예측 기반 코딩 기법에서는 H.264 및 HEVC등과 같이 동영상 압축 표준이 규정되어 주로 이용되고 있다.
- [0009] 그러나 최근 정보 통신 기술의 발전으로 인해 표준화된 영상 압축 기법만으로는 사용자의 요구를 만족시키기 어렵다. 이에 표준화된 영상 압축 기법을 기반으로 압축률을 향상하는 많은 기술이 제안되어 왔다. 표준 영상 압축 기법 기반 압축률 향상 기술로는 대표적으로 예측 성능 향상, 변환, 양자화 및 화질 향상 등이 있다. 예측 성능 향상은 새로운 예측 모드 추가, 머신러닝 학습 기반의 예측 등을 사용하여 예측 정확도를 향상시켜, 그 차이만을 변환하는 예측 기반 코딩 방법에서 차이를 줄이는 것을 이용하여 비트 수를 감소시키는 방법이다. 변환에서는 이산 코사인 변환(Discrete Cosine Transform)을 주로 사용하지만, 더 나은 기저(basis)를 갖는 주파수 도메인 변환(Frequency Domain Transform)을 이용하거나, 기저를 머신 러닝으로 학습하는 방법을 사용하여 더 많은 정보를 저주파로 압축하는 방법이 있다. 양자화는 더 많은 정보가 있는 곳은 정교하게, 더 적은 정보가 있을 것으로 예측되는 구간은 크게 양자화를 적용하여 비트 수를 감소시키는 방법이 연구되었다.
- [0010] 그러나 이러한 방법들 기본적으로 압축 표준 알고리즘 내부를 변경할 것을 요구하기 때문에 기존에 규정된 표준 압축 기법과 호환이 어렵다는 한계가 있다.
- [0011] 이에 가능한 표준 호환성을 유지하면서 화질을 향상시키기 위해, 압축 후 디코딩 된 중간 결과 또는 최종 결과의 화질을 원본에 가깝게 복원하는 방식으로 픽셀 적응적인 노이즈 감소 필터를 사용하거나 컨볼루션 신경망(Convolutional Neural Network: 이하 CNN)과 같은 인공 신경망을 이용하는 압축 후 화질 향상 기법이 제안된 바 있으나, 압축 후 화질 향상 기법은 원본 영상이 아닌 디코딩된 영상을 기반으로 화질을 향상시키므로, 압축 및 디코딩 과정에서 소실된 원본 영상의 정보로 인해 화질 향상 성능에 제약이 있다는 단점이 존재한다.

## 선행기술문헌

### 특허문헌

[0012] (특허문헌 0001) 한국 공개 특허 제10-2018-0119753호 (2018.11.05 공개)

## 발명의 내용

### 해결하려는 과제

- [0013] 본 발명의 목적은 표준 영상 압축 기법과 호환 가능하면서 복원 영상의 화질을 향상시킬 수 있는 영상 압축 장치, 이의 학습 장치 및 방법을 제공하는데 있다.
- [0014] 본 발명의 다른 목적은 표준 영상 압축 기법을 모의한 코텍 모델링 네트워크를 이용하여 표준 영상 압축 기법과 호환가능하면서 복원 영상의 화질을 향상시키기 위해 구비되는 전처리 네트워크를 학습시킬 수 있는 영상 압축 장치, 이의 학습 장치 및 방법을 제공하는데 있다.
- [0015] 본 발명의 또 다른 목적은 표준 영상 압축 기법을 모의한 코텍 모델링 네트워크와 함께 압축된 영상의 비트 스트림 구조를 모의하기 위한 비트 추정 네트워크를 더 포함하여 표준 영상 압축 기법에서의 압축률을 개선하여 복원 영상의 화질을 향상시킬 수 있는 영상 압축 장치, 이의 학습 장치 및 방법을 제공하는데 있다.

### 과제의 해결 수단

- [0016] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 영상 압축 장치는 로우 데이터 영상을 획득하는 영상 획득부; 상기 로우 데이터 영상을 인가받아 미리 학습된 패턴 추정 방식에 따라 전처리하는 전처리 네트워크; 및 전처리된 영상을 인가받아 기지정된 표준 압축 기법에 따라 압축하여 압축 영상을 출력하는 인코더부를 포함한다.
- [0017] 상기 전처리 네트워크는 학습 시에 추가되며, 인공 신경망으로 구현되고 상기 인코더부와 상기 압축 영상을 인가받아 디코딩하여 복호 영상을 획득하는 디코더부를 포함하는 표준 코텍부를 모의하도록 미리 학습된 코텍 모델링부에서 출력되는 모의 복호 영상을 기반으로 획득된 복원 영상과 상기 로우 데이터 영상을 비교하여 획득되는 복원 오차가 상기 코텍 모델링부를 통해 역전파되어 미리 학습될 수 있다.
- [0018] 상기 코텍 모델링부는 상기 인코더부와 상기 디코더부를 포함하는 상기 표준 코텍부를 모의하도록 학습되어 압축 영상을 인가받아 복호 영상에 대응하는 모의 복호 영상을 출력하는 코텍 모델링 네트워크; 상기 전처리 네트워크에서 전처리된 영상을 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 전처리된 영상으로부터 상기 인코더부에서 출력되는 압축 영상의 비트수를 추정하는 비트 추정 네트워크; 및 상기 코텍 모델링 네트워크와 비트 추정 네트워크를 학습시키기 위한 모델 학습 시에 상기 디코더부에서 획득되는 복호 영상과 상기 모의 복호 영상을 비교하여 모델링 오차를 획득하여 상기 코텍 모델링 네트워크로 역전파하고, 상기 비트 추정 네트워크에서 추정된 비트수와 상기 인코더부에서 출력되는 압축 영상의 비트수를 비교하여 비트 오차를 획득하여 비트 추정 네트워크로 역전파하는 모델 학습부를 포함할 수 있다.
- [0019] 상기 영상 압축 장치는 상기 전처리 네트워크를 학습시키기 위한 전처리 학습 시에 추가되어 상기 로우 데이터 영상과 상기 복호 영상 또는 상기 모의 복호 영상을 기지정된 방식으로 후처리하여 획득되는 복원 영상을 비교하여 복원 오차를 획득하고, 획득된 복원 오차를 상기 코텍 모델링 네트워크와 상기 비트 추정 네트워크를 통해 상기 전처리 네트워크로 역전파하는 영상 학습부를 더 포함할 수 있다.
- [0020] 상기 모델 학습부는 전처리 학습시에 상기 모델 학습에 의해 설정된 상기 코텍 모델링 네트워크와 상기 비트 추정 네트워크의 가중치를 고정할 수 있다.
- [0021] 상기 영상 학습부는 상기 복호 영상 또는 상기 모의 복호 영상으로부터 상기 복원 영상을 획득하도록 패턴 추정 방식이 미리 학습된 인공 신경망으로 구현된 후처리 네트워크에 의해 후처리되는 경우, 상기 복원 오차를 상기 후처리 네트워크로 우선 역전파하고, 상기 코텍 모델링 네트워크 및 상기 비트 추정 네트워크를 거쳐 상기 전처리 네트워크로 역전파시킬 수 있다.
- [0022] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 영상 압축 장치의 학습 장치는 로우 데이터 영상을



인가받아 미리 학습된 패턴 추정 방식에 따라 전처리하는 전처리 네트워크와 전처리된 영상을 인가받아 기지정된 표준 압축 기법에 따라 압축하여 압축 영상을 출력하는 인코더부를 포함하는 영상 압축 장치의 학습 장치로서, 인공 신경망으로 구현되고 상기 인코더부와 상기 압축 영상을 인가받아 디코딩하여 복호 영상을 획득하는 디코더부를 포함하는 표준 코덱부를 모의하도록 미리 학습되어 상기 전처리된 영상으로부터 모의 복호 영상을 획득하는 코덱 모델링부; 및 상기 코덱 모델링부에서 출력되는 상기 모의 복호 영상을 기반으로 획득된 복원 영상과 상기 로우 데이터 영상을 비교하여 복원 오차를 획득하고, 획득된 복원 오차를 상기 코덱 모델링부를 통해 역전파하여 상기 전처리 네트워크를 학습시키는 영상 학습부를 포함한다.

[0023] 상기 목적을 달성하기 위한 본 발명의 또 다른 실시예에 따른 영상 압축 장치의 학습 방법은 로우 데이터 영상을 인가받아 미리 학습된 패턴 추정 방식에 따라 전처리하는 전처리 네트워크와 전처리된 영상을 인가받아 기지정된 표준 압축 기법에 따라 압축하여 압축 영상을 출력하는 인코더부를 포함하는 영상 압축 장치의 학습 방법에 있어서, 상기 인코더부와 상기 압축 영상을 인가받아 디코딩하여 복호 영상을 획득하는 디코더부를 포함하는 표준 코덱부를 모의하도록 미리 학습된 인공 신경망을 이용하여 상기 전처리된 영상으로부터 모의 복호 영상을 획득하는 단계; 및 상기 모의 복호 영상을 기반으로 획득된 복원 영상과 상기 로우 데이터 영상을 비교하여 복원 오차를 획득하고, 획득된 복원 오차를 상기 인공 신경망을 통해 역전파하여 상기 전처리 네트워크를 학습시키는 단계를 포함한다.

### 발명의 효과

[0024] 따라서, 본 발명의 실시예에 따른 영상 압축 장치, 이의 학습 장치 및 방법은 전처리 네트워크를 제공하여 표준 압축 기법으로 영상을 압축 및 복호할 때, 압축률을 크게 높여 표준 영상 압축 기법과 호환 가능하면서 복원 영상의 화질을 향상시킬 수 있다. 또한 표준 영상 압축 기법을 모의한 코덱 모델링 네트워크와 함께 압축된 영상의 비트 스트림 구조를 모의하기 위한 비트 추정 네트워크를 더 포함하여 오차가 전처리 네트워크로 역전파될 수 있도록 하여 전처리 네트워크를 용이하게 학습시킬 수 있다.

### 도면의 간단한 설명

[0025] 도 1은 본 발명의 일 실시예에 따른 영상 압축 시스템의 구조를 나타낸다.  
 도 2는 도 1의 코덱 모델링 네트워크의 개략적 구조를 나타낸다.  
 도 3은 도 1의 비트 추정 네트워크의 개략적 구조를 나타낸다.  
 도 4는 본 발명의 다른 실시예에 따른 영상 압축 시스템의 구조를 나타낸다.  
 도 5는 본 실시예에 따른 영상 압축 시스템의 성능을 시뮬레이션한 결과를 나타낸다.  
 도 6은 본 발명의 일 실시예에 따른 영상 압축 방법 및 이의 학습 방법을 나타낸다.

### 발명을 실시하기 위한 구체적인 내용

[0026] 본 발명과 본 발명의 동작상의 이점 및 본 발명의 실시예에 의하여 달성되는 목적을 충분히 이해하기 위해서는 본 발명의 바람직한 실시예를 예시하는 첨부 도면 및 첨부 도면에 기재된 내용을 참조하여야만 한다.

[0027] 이하, 첨부한 도면을 참조하여 본 발명의 바람직한 실시예를 설명함으로써, 본 발명을 상세히 설명한다. 그러나, 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 설명하는 실시예에 한정되는 것이 아니다. 그리고, 본 발명을 명확하게 설명하기 위하여 설명과 관계없는 부분은 생략되며, 도면의 동일한 참조부호는 동일한 부재임을 나타낸다.

[0028] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라, 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 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.

[0029] 도 1은 본 발명의 일 실시예에 따른 영상 압축 시스템의 구조를 나타내고, 도 2는 도 1의 코덱 모델링 네트워크의 개략적 구조를 나타내며, 도 3은 도 1의 비트 추정 네트워크의 개략적 구조를 나타낸다.

[0030] 도 1을 참조하면, 본 실시예에 따른 영상 압축 시스템은 압축되어야 할 영상을 획득하고, 미리 학습된 패턴 추정 방식에 따라 획득된 영상에 대해 전처리하고, 지정된 표준에 따라 전처리된 영상을 압축하여 전달하는 영상



압축부(100)와 전송된 압축 영상에 적용된 표준 압축 기법에 대응하는 방식으로 디코딩하여 압축 영상을 디코딩하여 영상을 복원하는 영상 복원부(200)를 포함할 수 있다.

- [0031] 영상 압축부(100)는 압축되어야 할 영상을 획득하는 영상 획득부(110), 인공 신경망으로 구현되어 미리 학습된 패턴 추정 방식에 따라 획득된 영상에 대해 전처리하는 전처리 네트워크(120) 및 전처리된 영상을 미리 지정된 표준 압축 기법에 따라 인코딩하여 압축하는 인코더부(130)를 포함할 수 있다.
- [0032] 우선 영상 획득부(110)는 압축되어야 할 로우(raw) 데이터 영상을 획득한다. 여기서 로우 데이터 영상은 정지 영상일 수도 있으나 동영상일 수도 있다. 영상 획득부(110)는 로우 데이터 영상을 직접 취득하는 카메라 장치로 구현될 수도 있으며, 카메라에서 획득된 영상을 저장하는 저장 장치 등으로 구현될 수 있다.
- [0033] 전처리 네트워크(120)는 본 실시예에 따라 영상 압축부(100)에 추가되는 구성으로 영상 복원부(200)의 후처리부(220)에서 후처리된 영상이 영상 획득부(110)에서 획득된 로우 데이터 영상과 최대한 동일한 영상으로 획득될 수 있도록 로우 데이터 영상에 전처리를 수행한다. 전처리 네트워크(120)는 CNN과 같은 인공 신경망으로 구성될 수 있으며, 미리 학습된 패턴 추정 방식에 따라 로우 데이터 영상의 패턴을 추정하여 전처리를 수행할 수 있다.
- [0034] 인코더부(130)는 전처리된 영상을 인가받아 미리 지정된 표준 압축 기법에 따라 인코딩하여 압축한다. 여기서 인코더부(130)는 블록 기반 코딩 기법 또는 예측 기반 코딩 기법을 이용하여 전처리된 영상을 압축할 수 있으며, 일례로 획득된 로우 데이터 영상이 정지 영상이면, JPEG, JPEG2000 등의 압축 표준 기법을 적용할 수 있으며, 동영상이면 H.264 및 HEVC 등의 압축 표준 기법을 적용할 수 있다. 그러나 본 발명은 이에 한정되지 않으며, 인코더부(130)는 이외에도 다양한 압축 표준 기법에 따라 영상을 압축할 수 있다.
- [0035] 인코더부(130)에 의해 압축된 영상은 비트 스트림 형태로 영상 복원부(200)로 전송될 수 있다. 상기한 바와 같이, 영상 압축 기술은 영상의 화질을 최대한 원본과 가깝게 유지하면서 적은 비트수로 영상 데이터를 표현하여 전송 및 저장의 효율성을 향상시키기 위한 기술이므로, 인코더부(130)에서 압축되어 전송되는 압축 영상은 로우 데이터 영상에 비해 적은 비트수를 가지게 되며, 인코더부(130)의 압축 영상의 비트수는 압축률에 따라 가변될 수 있다.
- [0036] 한편, 영상 복원부(200)는 압축 영상을 인가받고 기기정된 방식으로 디코딩하는 디코더부(210) 및 디코딩된 영상에 기지정된 방식으로 후처리하여 영상의 화질을 개선하는 후처리부(220)를 포함할 수 있다.
- [0037] 디코더부(210)는 압축된 압축 영상을 인가받고, 압축 영상에 적용된 표준 압축 기법에 대응하는 디코딩 방식으로 압축 영상을 디코딩하여 복호 영상을 획득한다. 즉 디코더부(210)는 영상 압축부(100)의 인코더부(130)에서 압축 영상에 적용한 표준 압축 기법에 대응하는 디코딩 기법으로 압축 영상을 디코딩하여 복호 영상을 획득한다. 여기서 디코더부(210)는 인코더부(130)에서 적용된 표준 압축 기법을 미리 알고 있을 수 있으며, 전송된 압축 영상의 파일 포맷 등에 기반하여 적용될 파일 기법을 판별할 수도 있다. 즉 파일의 확장자나 파일 헤더 및 메타 데이터 등을 이용하여 압축 기법을 판별하고 이에 대응하는 디코딩 기법을 적용할 수 있다.
- [0038] 후처리부(220)는 디코더부(210)에서 디코딩된 복호 영상을 인가받고, 영상 화질 향상을 위해 지정된 다양한 후처리 기법 중 적어도 하나를 적용하여 복호 영상을 가능한 원본 데이터 영상에 가깝도록 복원한다. 후처리 기법은 다양하게 공지되어 있으므로 여기서는 상세한 설명은 생략한다. 그리고 경우에 따라서 후처리부(220)는 생략될 수도 있다.
- [0039] 상기한 바와 같이, 본 실시예에 따른 영상 압축 시스템에서는 영상 압축부(100)에 전처리 네트워크(120)가 구비되어 영상 복원부(200)에서 획득되는 복원 영상의 화질이 더욱 개선되도록 한다. 특히 전처리 네트워크(120)는 인코더부(130)의 전단에 배치되어 인코더부(130)에서 영상이 압축되기 이전에 로우 데이터 영상에 미리 전처리를 수행하여 복원 영상의 화질이 개선되도록 한다.
- [0040] 이는 표준 영상 압축 기법이 적용되는 인코더부(130)와 디코더부(210)가 수정되지 않고서도 복원 영상의 화질이 개선되도록 하므로, 표준 압축 기법의 호환성을 유지하면서도 화질을 개선할 수 있도록 한다.
- [0041] 다만 상기한 바와 같이, 전처리 네트워크(120)는 인공 신경망으로 구현될 수 있으며, 인공 신경망으로 구현되는 전처리 네트워크(120)가 화질 개선의 성능을 나타내기 위해서는 미리 학습이 수행되어야 한다.
- [0042] 그리고 전처리 네트워크(120)를 학습시키기 위해서는 전처리 네트워크(120)에서 전처리되고 압축 및 디코딩된 영상, 또는 후처리부(220)에 의해 후처리까지 수행된 복원 영상과 영상 획득부(110)가 획득한 로우 데이터 영상

과의 차이를 나타내는 오차를 분석하고, 분석된 오차를 전처리 네트워크(120)로 역전파해야 한다.

- [0043] 그러나 인공 신경망에서 오차의 역전파를 통한 학습은 미분 가능함을 전제로 수행되는 잘 알려진 사실이다. 따라서 오차가 전처리 네트워크(120)로 역전파되기 위해서는 영상을 압축 및 디코딩하는 코덱으로서 기능하는 인코더부(130) 및 디코더부(210)가 미분 가능한 연산을 수행해야 한다.
- [0044] 그러나 표준 압축 기법에 따라 지정된 연산을 수행하는 인코더부(130) 및 디코더부(210)는 일반적으로 미분 가능한 연산을 수행하지 않으며, 이에 전처리 네트워크(120)의 학습이 용이하지 않다는 문제가 있다.
- [0045] 이에 본 발명에서는 전처리 네트워크(120)의 학습을 위한 학습부를 더 포함할 수 있으며, 학습부는 코덱 모델링부(500)와 영상 학습부(600)로 구성될 수 있다.
- [0046] 도 1에서 표준 압축 기법에 따라 영상을 압축하고 디코딩하는 인코더부(130)와 디코더부(210)는 표준 코덱부의 구성으로 볼 수 있다. 그리고 코덱 모델링부(500)는 인코더부(130)와 디코더부(210)를 포함하는 표준 코덱부를 모델링하기 위한 구성으로, 오차 역전파가 불가능한 인코더부(130)와 디코더부(210)를 대신하여 오차 역전파가 가능하도록 인공 신경망으로 구성되고 학습이 수행될 수 있다.
- [0047] 상기한 바와 같이 전처리 네트워크(120)는 복원 영상의 화질을 개선하기 위해 구비되므로, 전처리 네트워크(120)의 학습 시에는 인코더부(130)가 전처리된 영상을 압축하는 방식과 디코더부(210)가 압축된 영상을 디코딩하는 방식을 알지 못해도 무관하다. 즉 압축 및 디코딩된 결과에 따른 오차가 전처리 네트워크(120)로 역전파 가능하면 된다. 이에 본 실시예에서는 표준 코덱부를 모의한 코덱 모델링부(500)를 제공함으로써, 전처리 네트워크(120)를 학습시킬 수 있도록 한다.
- [0048] 코덱 모델링부(500)는 코덱 모델링 네트워크(510), 비트 추정 네트워크(520) 및 모델 학습부(530)를 포함할 수 있다.
- [0049] 코덱 모델링 네트워크(510)는 표준 코덱부를 모의하기 위해 구비된다. 코덱 모델링 네트워크(510)는 전처리 네트워크(120)에 의해 전처리된 영상을 인가받고, 디코더부(210)에서 출력되는 복호 영상과 가급적 동일한 영상을 출력하도록 미리 학습되는 인공 신경망으로 구현될 수 있다.
- [0050] 코덱 모델링 네트워크(510)는 전처리 네트워크(120)가 최적의 압축률을 낼 수 있는 효율적 표현(Compact Representation)을 획득할 수 있도록, 표준 코덱부를 모의하면서, 미분 가능한 연산으로 오차에 따른 피드백 기울기(Feedback Gradient)를 전달하는 것이다.
- [0051] 따라서 코덱 모델링 네트워크(510)는 표준 코덱부의 기능을 최대한 동일하게 모의하도록 구성되어야 하며, 학습 시에만 이용되므로 충분한 복잡도를 가져도 무방하다. 코덱 모델링 네트워크(510)는 일례로 CNN 또는 RNN(Recurrent Neural Network)으로 구현될 수 있다.
- [0052] 도 2는 코덱 모델링 네트워크(510)가 블록 기반 코딩 기법에 따른 표준 코덱을 모의하는 것으로 가정하여 CNN으로 구현된 경우의 일례를 도시하였다. 도 2에 도시된 바와 같이, CNN으로 구현된 코덱 모델링 네트워크(510)는 인코더부 모의부(511)와 디코더부 모의부(512)를 포함할 수 있다. 인코더부 모의부(511)는 전처리된 영상을 인가받아 기지정된 크기의 블록 단위로 재배열하고, 재배열된 데이터를 각각 컨볼루션 레이어(CL)와 누설 ReLU(Leaky Rectified Linear Unit) 레이어(ReLU)를 포함하는 다수의 인코딩 필터를 통해 반복 필터링하여 출력한다. 디코더부 모의부(512)는 각각 컨볼루션 레이어(CL)와 누설 ReLU(Leaky Rectified Linear Unit) 레이어(ReLU) 및 인코더부 모의부(511) 다수의 인코딩 필터 중 대응하는 컨볼루션 필터의 컨볼루션 레이어가 결합된 다수의 디코딩 필터를 통해 인코더부 모의부(511)의 출력을 반복 필터링하여 출력하도록 구성될 수 있다.
- [0053] 만일 코덱 모델링 네트워크(510)가 예측 기반 코딩 기법에 따른 표준 코덱을 모의하는 경우, RNN 구조를 기반으로 구성되어 모델링된 블록이 다음 블록 모델링에 이용되도록 할 수 있다.
- [0054] 코덱 모델링 네트워크(510)는 전처리 네트워크(120)로부터 전처리된 영상을 인가받아 디코더부(210)에서 출력되는 복호 영상에 대응하는 모의 복호 영상을 출력하도록 학습되어야 하며, 모델 학습부(530)에 의해 학습이 수행될 수 있다.
- [0055] 모델 학습부(530)는 코덱 모델링 네트워크(510)에서 출력되는 모의 복호 영상과 디코더부(210)에서 출력되는 복호 영상을 인가받아, 모의 복호 영상과 복호 영상 사이의 모델링 오차를 계산하고 계산된 모델링 오차를 코덱 모델링 네트워크(510)로 역전파하여 코덱 모델링 네트워크(510)를 학습시킬 수 있다. 이때, 모델 학습부(530)는 모의 복호 영상과 복호 영상 사이의 픽셀간 평균 오차(Mean Squared Error) 또는 평균 절대값 오차(Mean

Absolute Error)를 손실 함수(Loss Function)로 적용하여 모델링 오차를 계산하고 계산된 모델링 오차를 코덱 모델링 네트워크(510)로 역전파할 수 있다.

- [0056] 한편 비트 추정 네트워크(520)는 전처리 네트워크(120)에서 전처리된 영상으로부터 인코더부(130)에서 압축된 압축 영상의 비트수를 추정하여 코덱 모델링 네트워크(510)가 더욱 정교하게 표준 코덱부를 모의할 수 있도록 하기 위한 구성이다.
- [0057] 만일 코덱 모델링 네트워크(510)를 단순히 표준 코덱부와 가능한 동일한 결과를 도출하도록 학습시키는 경우, 이후 전처리 네트워크(120)는 압축 영상의 비트수를 고려하지 않는 형태로 학습이 수행된다. 그러나 본 실시예에서 전처리 네트워크(120)는 인코더부(130)의 압축 효율성을 최적화하여 화질을 향상시키기 위한 전처리를 수행하기 위해 추가되므로, 전처리 네트워크(120)는 인코더부(130)의 압축율을 고려하여 학습이 수행되어야 한다.
- [0058] 이에 본 실시예에서는 전처리 네트워크(120)가 인코더부(130)의 압축율을 고려하여 학습될 수 있도록, 전처리된 영상을 인가받아 압축 영상의 비트수를 추정하는 비트 추정 네트워크(520)를 더 포함한다. 비트 추정 네트워크(520) 또한 도 3에 도시된 바와 같이, CNN으로 구현될 수 있다. 도 3을 참조하면 비트 추정 네트워크(520)는 전처리된 영상을 인가받아 기지정된 크기의 블록 단위로 재배열하고, 재배열된 데이터를 각각 컨볼루션 레이어(CL)와 누설 ReLU(Leaky Rectified Linear Unit) 레이어(ReLU) 및 배치 정규화(Batch Normalization) 레이어(BAT)를 포함하고, 컨볼루션 레이어(CL)의 차수가 증가되는 다수의 압축 필터를 통해 반복 필터링하는 비트 압축부(521) 및 각각 컨볼루션 레이어(CL)와 누설 ReLU(Leaky Rectified Linear Unit) 레이어(ReLU) 및 배치 정규화(Batch Normalization) 레이어(BAT)를 포함하되 컨볼루션 레이어(CL)의 차수가 감소되는 다수의 추출 필터를 통해 반복 필터링하여 비트수를 추정하는 비트 추출부(522)를 포함할 수 있다.
- [0059] 그리고 비트 추정 네트워크(520) 또한 모델 학습부(530)에 의해 학습이 수행될 수 있다. 모델 학습부(530)는 비트 추정 네트워크(520)가 전처리된 영상으로부터 추정된 추정 비트수와 인코더부(130)가 실제 전처리된 영상을 압축하여 획득한 압축 영상의 비트수의 차이를 계산하고, 계산된 비트수 차이를 비트 오차로 획득하여 비트 추정 네트워크(520)로 역전파함으로써, 비트 추정 네트워크(520)를 학습시킨다. 즉 비트 추정 네트워크(520)가 전처리된 영상으로부터 정확하게 압축 영상의 비트수를 추정할 수 있도록 학습시킨다.
- [0060] 코덱 모델링부(500)의 코덱 모델링 네트워크(510)와 비트 추정 네트워크(520)는 표준 코덱부를 모의하기 위한 모델 학습이 수행되는 동안 모델 학습부(530)에서 역전파되는 모델링 오차와 비트 오차에 따라 컨볼루션 레이어의 가중치를 가변하여 학습이 수행된다. 모델 학습은 모델링 오차와 비트 오차가 기지정된 기준 오차 이하가 될 때까지 또는 기지정된 반복 횟수만큼 반복되어 학습이 수행될 수 있다.
- [0061] 코덱 모델링부(500)에 대한 모델 학습이 완료되면, 코덱 모델링 네트워크(510)와 비트 추정 네트워크(520)의 가중치는 고정되고 이후 가변되지 않는다.
- [0062] 한편, 영상 학습부(600)는 학습된 코덱 모델링부(500)를 이용하여 전처리 네트워크(120)를 학습시키기 위해 구비된다. 전처리 네트워크(120)는 모델 학습이 완료된 이후, 전처리 학습 단계에서 학습이 수행될 수 있다.
- [0063] 영상 학습부(600)는 영상 획득부(110)로부터 로우 데이터 영상을 인가받는 한편, 후처리부(220)에서 후처리된 복원 영상을 인가받아 로우 데이터 영상과 복원 영상 사이의 차이를 복원 오차로 계산한다. 여기서 복원 오차 또한 모의 복호 영상과 복호 영상 로우 데이터 영상과 복원 영상 사이의 픽셀간 평균 오차 또는 평균 절대값 오차를 손실 함수로 적용하여 획득될 수 있다. 그리고 획득된 복원 오차를 역전파하여 전처리 네트워크(120)를 학습시킨다.
- [0064] 다만 상기한 바와 같이, 영상 학습부(600)에서 획득된 복원 오차는 피드백 기율기로 역전파되어야 하며, 디코더부(210) 및 인코더부(130)를 통해 역전파될 수 없다. 이에 영상 학습부(600)는 획득된 복원 오차를 표준 코덱부를 모의한 코덱 모델링부(500)를 통해 전처리 네트워크(120)로 역전파함으로써, 전처리 네트워크(120)를 학습시킬 수 있다.
- [0065] 이때, 표준 코덱부를 모의한 코덱 모델링부(500)의 코덱 모델링 네트워크(510)와 비트 추정 네트워크(520)는 모델 학습이 수행되어 가중치가 설정되면, 이후 전처리 학습 시에는 설정된 가중치가 변경되지 않는다. 이는 코덱 모델링부(500)가 표준 코덱부를 모의한 특성을 항상 그대로 유지해야만 전처리 네트워크(120)를 정확하게 학습시킬 수 있기 때문이다.
- [0066] 도 1에서는 영상 학습부(600)가 코덱 모델링부(500)의 모델 학습부(530)를 통해 복원 오차를 역전파하는 것으로 도시하였으나, 영상 학습부(600)는 직접 코덱 모델링 네트워크(510)와 비트 추정 네트워크(520)로 복원 오차를

역전파하여 전처리 네트워크(120)를 학습시킬 수 있다. 이 경우 코덱 모델링부(500)의 모델 학습부(530)는 모델 학습이 완료되면 제거되어도 무방하다.

- [0067] 또한 여기서는 설명의 편의를 위하여 영상 학습부(600)와 모델 학습부(530)를 별도의 구성으로 도시하였으나, 영상 학습부(600)와 모델 학습부(530)는 통합되어도 무방하다.
- [0068] 여기서 코덱 모델링부(500)와 영상 학습부(600)를 포함하는 학습부는 전처리 네트워크(120)를 학습시키기 위해 추가된 구성으로 전처리 네트워크(120)의 학습이 완료되면 제거될 수 있다.
- [0069] 도 4는 본 발명의 다른 실시예에 따른 영상 압축 시스템의 구조를 나타낸다.
- [0070] 도 1의 영상 압축 시스템과 비교하면, 도 4의 영상 압축 시스템은 영상 복원부(200)가 지정된 연산을 수행하는 후처리부(220) 대신 인공 신경망으로 구현되어 학습된 패턴에 따라 디코더부(210)에서 복호된 복호 영상으로부터 화질이 개선된 복원 영상을 획득하도록 구성되는 후처리 네트워크(230)를 포함한다.
- [0071] 영상 복원부(200)가 후처리 네트워크(230)를 포함하는 경우, 전처리 네트워크(120) 뿐만 아니라 후처리 네트워크(230) 또한 학습이 수행될 필요가 있으며, 전처리 네트워크(120)와 후처리 네트워크(230)가 함께 학습이 수행되는 경우, 화질이 크게 개선된 복원 영상을 획득할 수 있다. 즉 압축률을 표준 코덱부가 압축율을 크게 높여 적은 비트수로도 로우 데이터 영상에 매우 가까운 복원 영상을 획득할 수 있다.
- [0072] 후처리 네트워크는 다양한 인공 신경망으로 구현될 수 있으며, 일례로 기존의 영상 압축 시스템에서 후처리를 위해 이용된 SRCNN(Super-Resolution CNN) 또는 ARCNN(Artifact Removal CNN) 등이 이용될 수 있다.
- [0073] 도 1의 영상 압축 시스템에서는 후처리부(220)가 학습될 필요가 없었으므로, 영상 학습부(600)가 후처리부(220)에서 획득된 복원 영상을 인가받아 곧바로 코덱 모델링부(500)로 역전파하였다.
- [0074] 그러나 도 4와 같이 후처리 네트워크(230)를 전처리 네트워크(120)와 함께 학습시켜야 하는 경우, 오차의 역전파 경로에 후처리 네트워크(230)가 추가로 포함되어야 한다. 이에 도 4에서는 영상 학습부(600)가 영상 획득부(110)로부터 로우 데이터 영상을 인가받는 한편, 후처리 네트워크(230)에서 후처리된 복원 영상을 인가받아 로우 데이터 영상과 복원 영상 사이의 차이를 복원 오차로 계산하고, 계산된 복원 오차를 후처리 네트워크(230)로 역전파하여 후처리 네트워크(230)가 학습되도록 한다. 그리고 모델 학습부(530)는 후처리 네트워크(230)에서 역전파되는 오차를 인가받아 코덱 모델링 네트워크(510)와 비트 추정 네트워크(520)를 통해 전처리 네트워크(120)를 역전파한다. 즉 전처리 네트워크(120)와 후처리 네트워크(230)가 함께 학습이 수행되도록 한다. 이때 코덱 모델링 네트워크(510)와 비트 추정 네트워크(520)는 상기한 바와 같이, 모델 학습 시에 학습된 상태를 계속 유지한다.
- [0075] 결과적으로 도 4의 영상 압축 시스템은 전처리 네트워크(120) 뿐만 아니라 후처리 네트워크(230) 또한 인공 신경망으로 구성하고, 서로 연계하여 동시 학습시킴으로써, 표준 코덱을 이용하여 압축 및 디코딩을 수행할 때 압축률을 크게 높일 수 있다. 즉 복원 영상의 화질을 개선할 수 있다.
- [0076] 또한 표준 코덱부를 구성하는 인코더부(130)와 디코더부(210)에서 수행되는 동작의 변경없이 전처리 및 후처리를 이용하여 화질을 개선하므로, 표준 압축 기법과의 호환성을 유지하면서 화질을 개선할 수 있다.
- [0077] 여기서도 코덱 모델링부(500)와 영상 학습부(600)를 포함하는 학습부는 전처리 네트워크(120)와 후처리 네트워크(230)를 학습시키기 위해 추가된 구성으로 전처리 네트워크(120)의 학습이 완료되면 제거될 수 있다.
- [0078] 도 5는 본 실시예에 따른 영상 압축 시스템의 성능을 시뮬레이션한 결과를 나타낸다.
- [0079] 도 5에서 (a)는 영상 획득부(110)가 획득한 로우 데이터 영상으로 나비의 날개 이미지를 나타내고, (b)는 전처리 네트워크(120)에 의해 전처리된 영상을 나타내며, (c)는 JPEG 표준 압축 기법에 따라 압축 및 복호한 영상을 나타낸다. 그리고 (d)는 후처리 네트워크(230)에 의해 후처리된 영상을 나타낸다.
- [0080] 한편, (e)는 전처리되지 않고 로우 데이터 영상을 압축 및 복호한 영상을 나타내고, (f)는 ARCNN에 의해 후처리된 영상을 나타낸다.
- [0081] 도 5에서 (d)와 (f)를 비교하면, 유사한 화질(PSNR)을 갖는 복원 영상을 획득하지만, (c) 및 (e)를 비교하면, 픽셀당 비트수(Bit Per Pixel: BPP)에서 (c)는 0.2548인데 비해 (e)는 0.3613으로 전처리를 수행하는 경우의 압축률이 월등하게 높다는 것을 알 수 있다.
- [0082] 도 6은 본 발명의 일 실시예에 따른 영상 압축 방법 및 이의 학습 방법을 나타낸다.



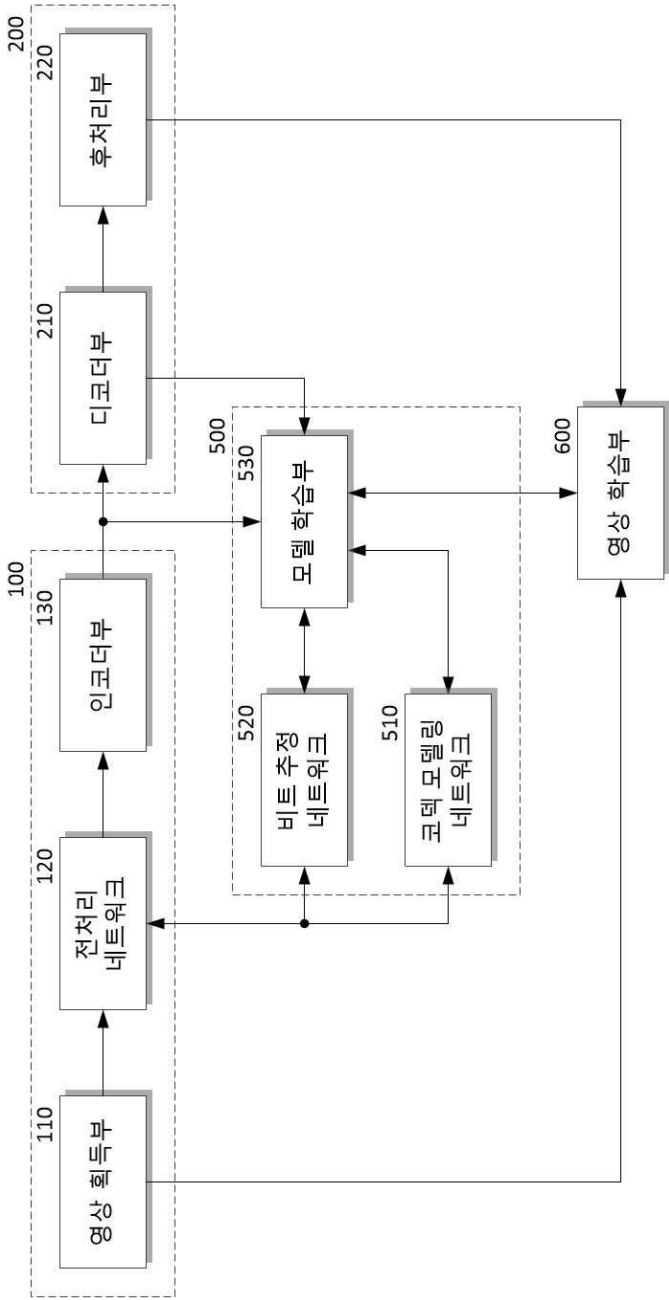
- [0083] 도 1 내지 도 4를 참조하여, 도 6의 영상 압축 방법 및 이의 학습 방법을 설명하면, 우선 압축하고자 하는 로우 데이터 영상을 획득한다(S11). 그리고 획득된 로우 데이터 영상을 학습된 패턴 추정 방식에 따라 전처리한다(S12). 전처리된 영상이 획득되면, 미리 지정된 표준 압축 기법에 따라 전처리된 영상을 압축하여 압축 영상을 획득하여 전송한다(S13). 이후 압축 영상을 인가받아 디코딩하여 복호 영상을 획득한다(S14).
- [0084] 그리고 모델 학습 단계인지 판별한다(S15). 만일 모델 학습 단계인 것으로 판별되면, 표준 압축 기법에 따라 전처리된 영상을 압축 및 디코딩한 결과 출력하도록 모의한 코덱 모델링 네트워크(510)를 이용하여 모의 복호 영상을 획득한다(S16). 또한 전처리된 영상으로부터 표준 압축 기법에 따라 압축된 압축 영상의 비트수를 추정하도록 구성된 비트 추정 네트워크(520)를 이용하여 압축 영상의 비트수를 추정한다(S17).
- [0085] 이후 획득된 모의 복호 영상과 복호 영상을 비교하여 모델링 오차를 계산하고, 압축 영상의 비트수와 추정 비트수를 비교하여 비트 오차를 계산한다(S18). 그리고 계산된 모델링 오차와 비트 오차를 코덱 모델링 네트워크(510)와 비트 추정 네트워크(520) 각각으로 역전파하여 코덱 모델링 네트워크(510)와 비트 추정 네트워크(520)를 학습시킨다(S19).
- [0086] 코덱 모델링 네트워크(510)와 비트 추정 네트워크(520)는 모델링 오차와 비트 오차가 기지정된 기준 모델링 오차 및 기준 비트 오차 이하가 되거나 기지정된 횟수에 도달할 때까지 반복 학습될 수 있다. 그리고 코덱 모델링 네트워크(510)와 비트 추정 네트워크(520)에 대한 모델 학습이 완료되면 코덱 모델링 네트워크(510)와 비트 추정 네트워크(520)의 가중치는 고정된다.
- [0087] 한편, 모델 학습이 아닌 것으로 판별되면, 획득된 복호 영상에 대해 기지정된 후처리를 수행하여 복원 영상을 획득한다(S20). 경우에 따라 후처리는 미리 학습된 패턴 추정 방식에 따라 수행될 수 있다. 복원 영상이 획득되면, 전처리 학습 단계인지 판별한다(S21). 전처리 학습 단계인 것으로 획득된 복원 영상과 로우 데이터 영상을 비교하여 복원 오차를 계산한다(S22). 그리고 계산된 복원 오차를 코덱 모델링 네트워크(510)와 비트 추정 네트워크(520)를 통해 역전파하여 전처리 네트워크(120)를 학습시킨다(S23). 이때, 만일 후처리를 수행하는 후처리부(220)가 인공 신경망으로 구현된 후처리 네트워크(230)이면, 계산된 복원 오차는 후처리 네트워크(230)로 우선 역전파되고, 후처리 네트워크(230)에서 코덱 모델링 네트워크(510)와 비트 추정 네트워크(520)를 통해 전처리 네트워크(120)로 역전파됨으로써, 전처리 네트워크(120)와 후처리 네트워크(230)를 동시에 학습시킨다.
- [0088] 그러나 전처리 학습 단계가 아닌 것으로 판별되면, 종료된다.
- [0089] 본 발명에 따른 방법은 컴퓨터에서 실행시키기 위한 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다. 여기서 컴퓨터 판독가능 매체는 컴퓨터에 의해 액세스될 수 있는 임의의 가용 매체일 수 있고, 또한 컴퓨터 저장 매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함하며, ROM(판독 전용 메모리), RAM(랜덤 액세스 메모리), CD(컴팩트 디스크)-ROM, DVD(디지털 비디오 디스크)-ROM, 자기 테이프, 플로피 디스크, 광데이터 저장장치 등을 포함할 수 있다.
- [0090] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다.
- [0091] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

## 부호의 설명

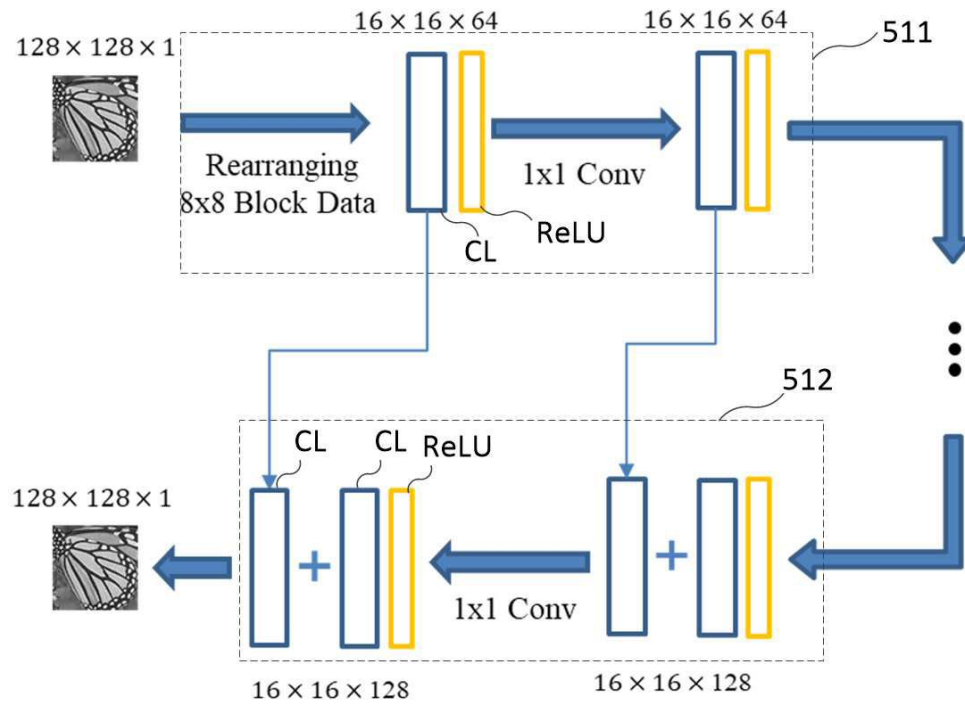
- [0092]
- |                 |                  |
|-----------------|------------------|
| 100: 영상 압축부     | 110: 영상 획득부      |
| 120: 전처리 네트워크   | 130: 인코더부        |
| 200: 영상 복원부     | 210: 디코더부        |
| 220: 후처리부       | 230: 후처리 네트워크    |
| 500: 코덱 모델링부    | 510: 코덱 모델링 네트워크 |
| 520: 비트 추정 네트워크 | 530: 모델 학습부      |
| 600: 영상 학습부     |                  |

도면

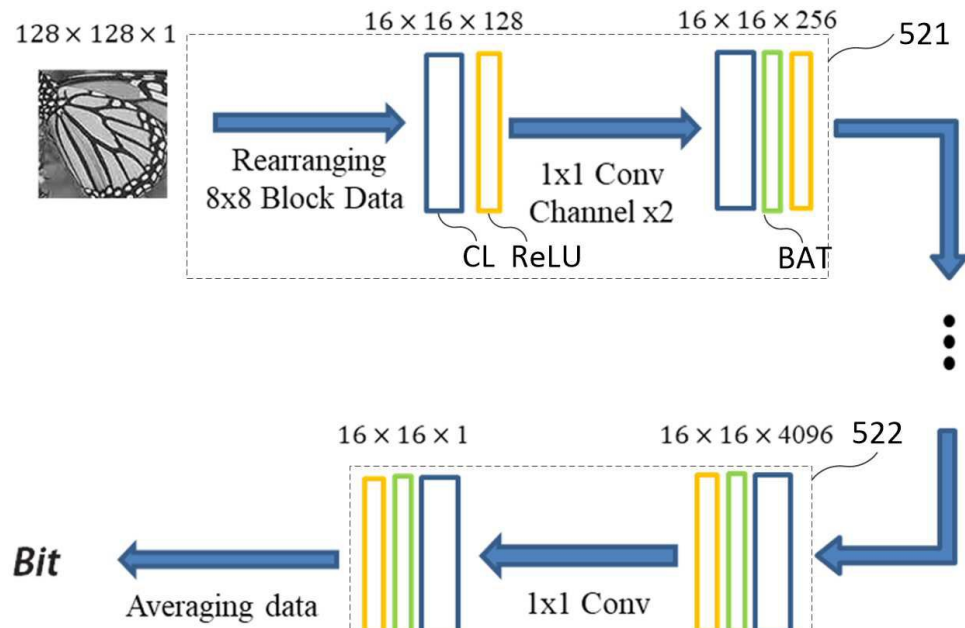
도면1



도면2

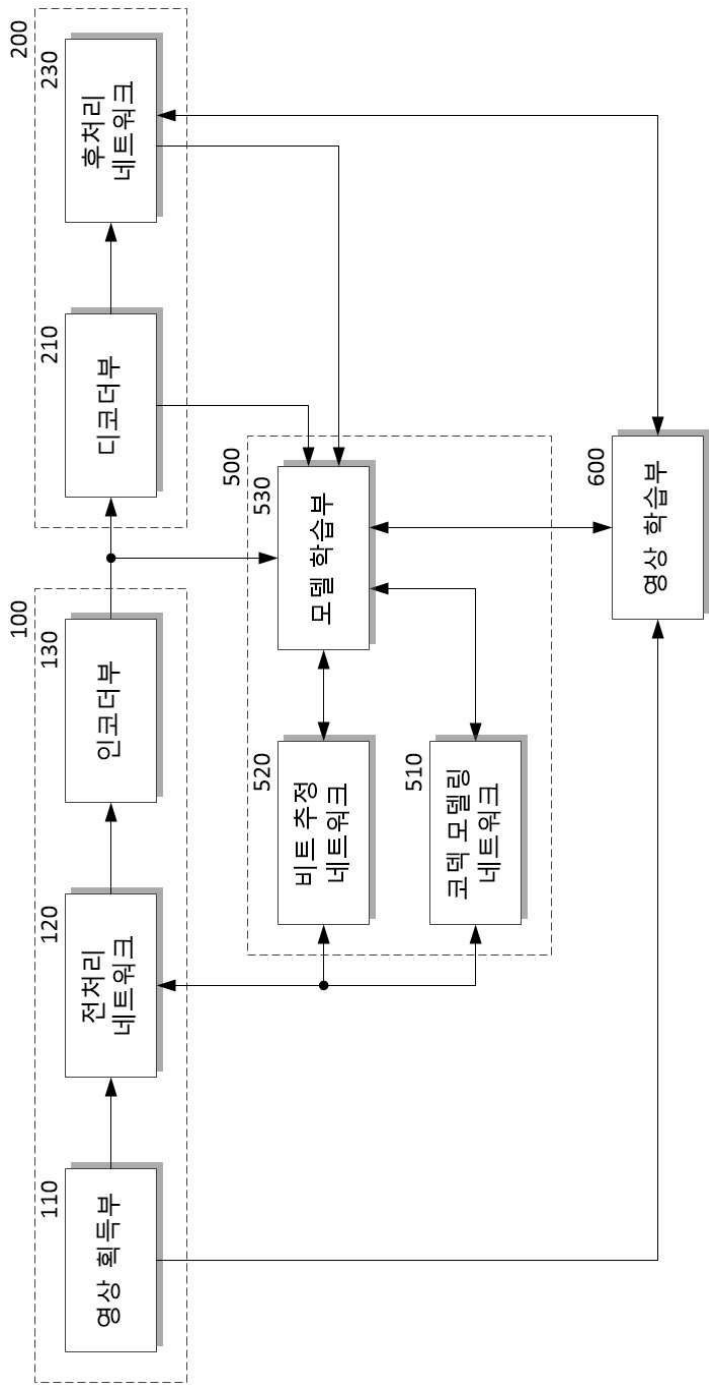


도면3

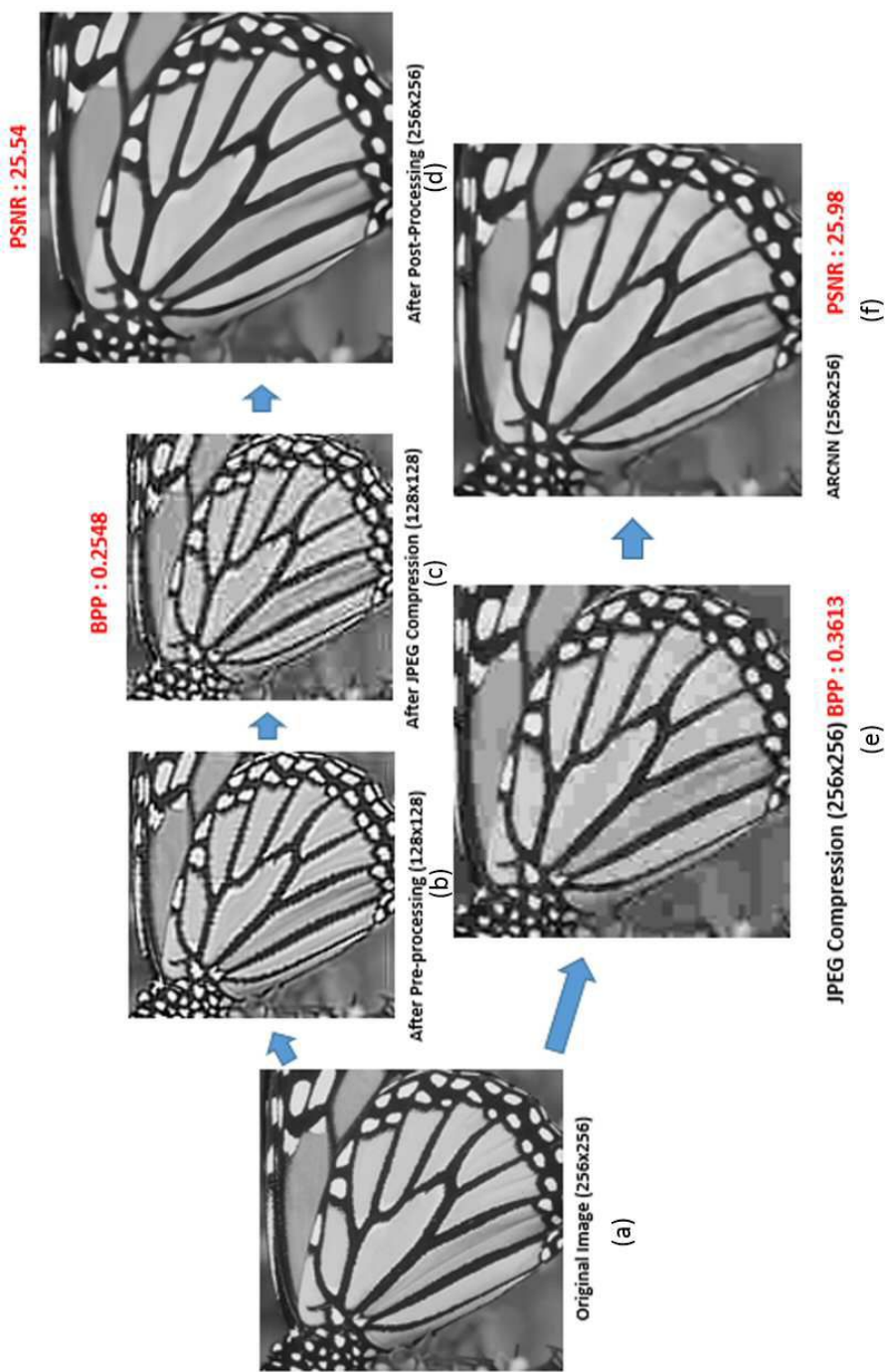




도면4



도면5



도면6

