



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2025년03월13일
(11) 등록번호 10-2780973
(24) 등록일자 2025년03월10일

(51) 국제특허분류(Int. Cl.)
G06T 9/00 (2019.01) G06F 17/15 (2006.01)
G06N 3/0464 (2023.01) H04N 19/124 (2014.01)
H04N 19/13 (2014.01) H04N 19/147 (2014.01)

(52) CPC특허분류
G06T 9/002 (2013.01)
G06F 17/153 (2013.01)

(21) 출원번호 10-2023-0058660
(22) 출원일자 2023년05월04일
심사청구일자 2023년05월04일
(65) 공개번호 10-2024-0161901
(43) 공개일자 2024년11월13일
(56) 선행기술조사문헌

David Minnen 외 2명, "Joint Autoregressive and Hierarchical Priors for Learned Image Compression" in Proc. Int. Conf. Neural Information Processing Systems, 페이지 1-10, 2018. 1부.*

ROBERT A. COHEN 외 2명, "Lightweight Compression of Intermediated Neural Network Features for Collaborative Intelligence" IEEE Open Journal of Circuits and Systems, 페이지 350-362, 2021.05.13. 1부.*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자
이종석
인천광역시 연수구 송도과학로 85(송도동)
전근우

인천광역시 연수구 경원대로119번길 21, 116동 904호 (동춘동, 연수2차풍림아파트)

유승은
서울특별시 서대문구 신촌로9길 37-5, 서윤하우스 408호 (창천동)

(74) 대리인
특허법인(유한)아이시스

전체 청구항 수 : 총 13 항

심사관 : 권영학

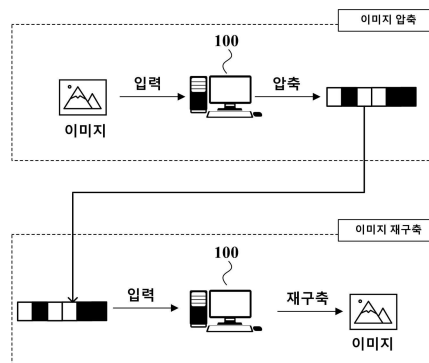
(54) 발명의 명칭 이미지 압축 방법 및 장치

(57) 요약

이미지 압축 방법은 영상 처리 장치가 이미지를 입력 받는 단계; 상기 영상 처리 장치가 입력 받은 이미지를 압축 모델에 입력하여 이미지를 압축하는 단계; 및 상기 영상 처리 장치가 상기 압축된 이미지를 상기 압축 모델에 입력하여 이미지를 복원하는 단계; 포함한다.

(뒷면에 계속)

대표도 - 도1



상기 압축 모델은 컨볼루션(Convolution) 연산을 수행하는 모델을 포함한다.

상기 압축 모델은 상기 입력 받은 이미지를 Leaky-Clip 모듈 및 라운딩 함수(Rounding Function)를 이용해 양자화(Quantization) 한 뒤, 상기 양자화 된 입력 받은 이미지와 미리 학습된 양자화된 가중치(Weight)를 이용해서 컨볼루션 연산을 수행한다.

상기 Leaky-Clip 모듈은 상기 압축 모델의 추론(Inference) 과정에서는 Clip함수의 기능을 수행하고, 상기 압축 모델의 학습(Training) 과정에서는 수정된 Clip 함수의 기능을 수행할 수 있다.

상기 Clip 함수는 내부영역(Inside)의 기울기는 1, 외부영역(Outside)의 기울기는 0인 함수일 수 있다.

상기 수정된 Clip함수는 내부영역의 기울기는 1, 외부영역의 기울기는 0초과 1이하인 함수일 수 있다.

(52) CPC특허분류

- G06N 3/0464 (2023.01)
- H04N 19/124 (2015.01)
- H04N 19/13 (2015.01)
- H04N 19/147 (2015.01)
- G06T 2207/20084 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711155041
과제번호	2021R1A2C2011474
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)
연구과제명	딥러닝 기반 차세대 이미지 압축 기술
기 여 율	1/1
과제수행기관명	연세대학교
연구기간	2022.03.01 ~ 2023.02.28

공지예외적용 : 있음

명세서

청구범위

청구항 1

영상 처리 장치가 이미지를 입력 받는 단계;

상기 영상 처리 장치가 상기 입력 받은 이미지를 압축 모델에 입력하여 이미지를 압축하는 단계; 및

상기 영상 처리 장치가 상기 압축된 이미지를 상기 압축 모델에 입력하여 이미지를 복원하는 단계; 포함하되,

상기 압축 모델은 컨볼루션(Convolution) 연산을 수행하는 모델을 포함하고,

상기 압축 모델은 학습(Training) 과정에 있어서 학습 이미지 및 학습 대상 가중치 각각을 스케일 파라미터로 나눈 값을 Leaky-Clip 모듈 및 라운딩 함수(Rounding Function)을 이용해 양자화(Quantization)한 뒤, 상기 양자화된 학습 이미지 및 상기 양자화된 학습 대상 가중치에 스케일 파라미터를 곱한 값을 기반으로 컨볼루션 연산을 수행함으로써 학습되고,

상기 압축 모델은 추론(Inference) 과정에 있어서, 상기 입력 받은 이미지를 입력 정보 스케일 파라미터로 나눈 값을 상기 Leaky-Clip 모듈 및 상기 라운딩 함수를 이용하여 양자화 한 뒤, 상기 추론과정에서 양자화한 결과와 미리 학습된 양자화된 가중치를 이용하여 컨볼루션을 수행한 뒤, 상기 컨볼루션 결과에 상기 입력 정보 스케일 파라미터 및 가중치 스케일 파라미터를 곱하는 연산을 수행하고,

상기 Leaky-Clip 모듈은 상기 압축 모델의 순전파(Foward) 과정에서는 Clip함수의 기능을 수행하고, 상기 압축 모델의 역전파(Backward) 과정에서는 수정된 Clip 함수의 기능을 수행하고,

상기 Clip 함수는 내부영역(Inside)의 기울기는 1, 외부영역(Outside)의 기울기는 0인 함수이고,

상기 수정된 Clip함수는 내부영역의 기울기는 1, 외부영역의 기울기는 0초과 1이하인 함수인, 이미지 압축 방법.

청구항 2

제1항에 있어서

상기 압축 모델은 메인 인코더, 메인 디코더, 하이퍼 인코더, 하이퍼 디코더, 산술부호화(Arithmetic Encoder, AE) 및 산술복호화(Arithmetic Decoder, DE)를 포함하는, 이미지 압축 방법.

청구항 3

삭제

청구항 4

삭제

청구항 5

제1항에 있어서

상기 압축 모델은 비트스트림 길이를 최소화하고 재구성된 이미지 왜곡을 최소화하는 항(term) 및 양자화 에러를 최소화하는 항을 포함하는 손실함수를 이용해 학습된, 이미지 압축 방법.

청구항 6

제5항에 있어서,

상기 손실함수는 아래의 수학적식(L)인, 이미지 압축 방법.

(수학식)

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_{RD} + \mathcal{L}_{SQE} \\ \mathcal{L}_{RD} &= R + \lambda \cdot D \\ &= \mathbf{E}_{\mathbf{x} \sim p_{\mathbf{x}}} [-\log_2 p_{\hat{\mathbf{y}}|\hat{\mathbf{z}}}(\hat{\mathbf{y}}|\hat{\mathbf{z}}) - \log_2 p_{\hat{\mathbf{z}}}(\hat{\mathbf{z}})] + \lambda \cdot \mathbf{E}_{\mathbf{x} \sim p_{\mathbf{x}}} [d(\mathbf{x}, \hat{\mathbf{x}})] \\ \mathcal{L}_{SQE} &= \beta \left\{ \sum (w - \hat{w})^2 + \sum (a - \hat{a})^2 \right\} \end{aligned}$$

상기 수학식에서 \mathcal{L}_{RD} 는 압축된 비트스트림의 길이를 최소화하고 재구축된 왜곡(Distortion)을 최소화하는 항이다. 상기 수학식에서 λ 는 R 과 D 사이의 트레이드-오프(trade-off)를 조절한다. 상기 수학식에서 $p_{\mathbf{x}}$ 는 원본 이미지(Original image)의 분포(distribution)을 의미한다. 상기 수학식에서 R 는 $\hat{\mathbf{y}}$ 의 추정 엔트로피와 $\hat{\mathbf{z}}$ 의 추정 엔트로피의 합을 의미한다. 상기 수학식에서 D 는 원본 이미지와 재구축된 이미지 사이의 평균 제곱 에러(Mean squared error, MSE)를 의미한다. 상기 수학식에서 \mathcal{L}_{SQE} 는 거듭제곱 양자화 에러(Squared Quantization Error Loss)에 대한 항이다. 상기 수학식에서 w 는 학습 대상 가중치이다. 상기 수학식에서 \hat{w} 는 학습 대상 가중치를 양자화한 후 역양자화 하였을 때의 값이다. 상기 수학식에서 a 는 입력 정보(또는 학습 이미지)이다. 상기 수학식에서 \hat{a} 는 학습 이미지를 양자화 한 후 역양자화 하였을 때의 값이다. 상기 수학식에서 β 는 학습 대상과 대상을 양자화한 값의 차가 발산하는 것을 방지하기 위해 설정해주는 아주 작은 값이다.

청구항 7

제1항에 있어서,

상기 압축 모델은 상기 입력 받은 이미지를 8비트(Bit)로 표현가능한 정수로 양자화 하는, 이미지 압축 방법.

청구항 8

제1항에 있어서,

상기 수정된 Clip함수의 외부영역의 기울기는 0.01인, 이미지 압축 방법.

청구항 9

이미지를 입력 받는 입력장치;

상기 입력 받은 이미지를 압축 모델에 입력하여 이미지를 압축하고, 상기 압축된 이미지를 상기 압축 모델에 입력하여 이미지를 복원하는 연산장치; 및

상기 압축 모델을 저장하는 저장장치;를 포함하되,

상기 압축 모델은 컨볼루션(Convolution) 연산을 수행하는 모델을 포함하고,

상기 압축 모델은 학습(Training) 과정에 있어서 학습 이미지 및 학습 대상 가중치 각각을 스케일 파라미터로 나눈 값을 Leaky-Clip 모듈 및 라운딩 함수(Rounding Function)을 이용해 양자화(Quantization)한 뒤, 상기 양자화된 학습 이미지 및 상기 양자화된 학습 대상 가중치에 스케일 파라미터를 곱한 값을 기반으로 컨볼루션 연산을 수행함으로써 학습되고,

상기 압축 모델은 추론(Inference) 과정에 있어서, 상기 입력 받은 이미지를 입력 정보 스케일 파라미터로 나눈 값을 상기 Leaky-Clip 모듈 및 상기 라운딩 함수를 이용하여 양자화 한 뒤, 상기 추론과정에서 양자화된 결과와 미리 학습된 양자화된 가중치를 이용하여 컨볼루션을 수행한 뒤, 상기 컨볼루션 결과에 상기 입력 정보 스케일 파라미터 및 가중치 스케일 파라미터를 곱하는 연산을 수행하고,

상기 Leaky-Clip 모듈은 상기 압축 모델의 순전파(Foward) 과정에서는 Clip함수의 기능을 수행하고, 상기 압축

모델의 역전파(Backward) 과정에서는 수정된 Clip 함수의 기능을 수행하고,

상기 Clip 함수는 내부영역(Inside)의 기울기는 1, 외부영역(Outside)의 기울기는 0인 함수이고,

상기 수정된 Clip함수는 내부영역의 기울기는 1, 외부영역의 기울기는 0초과 1이하인 함수인, 이미지 압축 장치.

청구항 10

제9항에 있어서

상기 압축 모델은 메인 인코더, 메인 디코더, 하이퍼 인코더, 하이퍼 디코더, 산술부호화(Arithmetic Encoder, AE) 및 산술복호화(Arithmetic Decoder, DE)를 포함하는, 이미지 압축 장치.

청구항 11

삭제

청구항 12

삭제

청구항 13

제9항에 있어서

상기 압축 모델은 비트스트림 길이를 최소화하고 재구축된 이미지 왜곡을 최소화하는 항(term) 및 양자화 에러를 최소화하는 항을 포함하는 손실함수를 이용해 학습된, 이미지 압축 장치.

청구항 14

제13항에 있어서,

상기 손실함수는 아래의 수학식(L)인, 이미지 압축 장치.

(수학식)

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_{RD} + \mathcal{L}_{SQE} \\ \mathcal{L}_{RD} &= R + \lambda \cdot D \\ &= \mathbf{E}_{\mathbf{x} \sim p_{\mathbf{x}}} [-\log_2 p_{\hat{\mathbf{y}}|\hat{\mathbf{z}}}(\hat{\mathbf{y}}|\hat{\mathbf{z}}) - \log_2 p_{\hat{\mathbf{z}}}(\hat{\mathbf{z}})] + \lambda \cdot \mathbf{E}_{\mathbf{x} \sim p_{\mathbf{x}}} [d(\mathbf{x}, \hat{\mathbf{x}})] \\ \mathcal{L}_{SQE} &= \beta \left\{ \sum (w - \hat{w})^2 + \sum (a - \hat{a})^2 \right\} \end{aligned}$$

상기 수학식에서 \mathcal{L}_{RD} 는 압축된 비트스트림의 길이를 최소화하고 재구축된 왜곡(Distortion)을 최소화하는 항이다. 상기 수학식에서 λ 는 R 과 D 사이의 트레이드-오프(trade-off)를 조절한다. 상기 수학식에서 $p_{\mathbf{x}}$ 는 원본 이미지(Original image)의 분포(distribution)을 의미한다. 상기 수학식에서 R 는 $\hat{\mathbf{y}}$ 의 추정 엔트로피와 $\hat{\mathbf{z}}$ 의 추정 엔트로피의 합을 의미한다. 상기 수학식에서 D 는 원본 이미지와 재구축된 이미지 사이의 평균 제곱 에러(Mean squared error, MSE)를 의미한다. 상기 수학식에서 \mathcal{L}_{SQE} 는 거듭제곱 양자화 에러(Squared Quantization Error Loss)에 대한 항이다. 상기 수학식에서 w 는 학습 대상 가중치이다. 상기 수학식에서 \hat{w} 는 학습 대상 가중치를 양자화한 후 역양자화 하였을 때의 값이다. 상기 수학식에서 a 는 입력 정보(또는 학습 이미지)이다. 상기 수학식에서 \hat{a} 는 학습 이미지를 양자화 한 후 역양자화 하였을 때의 값이다. 상기 수학식에서 β 는 학습 대상과 대상 을 양자화한 값의 차가 발산하는 것을 방지하기 위해 설정해주는 아주 작은 값이다.

청구항 15

제9항에 있어서,

상기 압축 모델은 상기 입력 받은 이미지를 8비트(Bit)로 표현가능한 정수로 양자화 하는, 이미지 압축 장치.

청구항 16

제9항에 있어서,

상기 수정된 Clip함수의 외부영역의 기울기는 0.01인, 이미지 압축 장치.

청구항 17

제1항, 제2항, 제5항 내지 제8항 중 어느 하나의 항에 기재된 이미지 압축 방법을 실행하기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

발명의 설명

기술 분야

[0001] 이하 설명하는 기술은 이미지 압축 방법 및 장치에 대한 것이다.

배경 기술

[0002] 일반적으로 이미지는 많은 양의 정보를 가지고 있다. 이에 이미지를 처리하거나 전송할때는 많은 자원이 소모된다. 이를 해결하기 위해 이미지 압축(Image Compression) 기술이 개발되었다. 이미지 압축은 이미지를 더 적은 데이터로 표현하는 기술이다. 이미지 압축 기술은 이미지를 더 적은 데이터로 표현하여 전송하거나 공유한 뒤, 이를 다시 복원하는 방식으로 많이 활용된다.

[0003] 학습 기반 이미지 압축(Learned Image Compression, LIC)는 이미지를 학습 기반 모델을 통해 이미지를 압축하는 방법이다. 최근에는 딥러닝 등의 인공지능망을 이용한 기술이 발전되기 시작하면서, 인공지능망 기반의 이미지를 압축할 수 있는 방법들이 개발되고 있다.

선행기술문헌

특허문헌

[0004] (특허문헌 0001) 한국 등록특허공보 10-2384255

발명의 내용

해결하려는 과제

[0005] 학습 기반 이미지 압축을 이용해 이미지를 압축하는 방법은 많이 활용되고 있다. 하지만 학습 기반 이미지 압축을 이용해 이미지를 압축하는 방법은 일반적으로 많은 양의 컴퓨터 자원 등이 필요하다. 예를 학습 기반 이미지 압축 방법은 종래 이미지 압축 방법인 JPEG(Joint Photographic Experts Group)과 비교해 볼 때, 메모리사용량이나 계산 복잡도 면에서 효율이 좋지 아니한다.

[0006] 이하 설명하는 기술은 위와 같은 문제를 해결하기 위한 기술이다. 이하 설명하는 기술은 입력 정보를 양자화 하거나 양자화된 가중치를 이용하는 압축 모델을 이용해 많은 양의 자원이 없이도 이미지를 압축할 수 있는 방법을 제공한다.

과제의 해결 수단

[0007] 과제를 해결하기 위한 이미지 압축 방법은 영상 처리 장치가 이미지를 입력 받는 단계; 상기 영상 처리 장치가 입력 받은 이미지를 압축 모델에 입력하여 이미지를 압축하는 단계; 및 상기 영상 처리 장치가 상기 압축된 이미지를 상기 압축 모델에 입력하여 이미지를 복원하는 단계; 포함한다.

- [0008] 상기 압축 모델은 컨볼루션(Convolution) 연산을 수행하는 모델을 포함한다.
- [0009] 상기 압축 모델은 상기 입력 받은 이미지를 Leaky-Clip 모듈 및 라운딩 함수(Rounding Function)를 이용해 양자화(Quantization) 한 뒤, 상기 양자화 된 입력 받은 이미지와 미리 학습된 양자화된 가중치(Weight)를 이용해서 컨볼루션 연산을 수행한다.
- [0010] 상기 Leaky-Clip 모듈은 상기 압축 모델의 추론(Inference) 과정에서는 Clip함수의 기능을 수행하고, 상기 압축 모델의 학습(Training) 과정에서는 수정된 Clip 함수의 기능을 수행할 수 있다.
- [0011] 상기 Clip 함수는 내부영역(Inside)의 기울기는 1, 외부영역(Outside)의 기울기는 0인 함수일 수 있다.
- [0012] 상기 수정된 Clip함수는 내부영역의 기울기는 1, 외부영역의 기울기는 0초과 1이하인 함수일 수 있다.

발명의 효과

- [0013] 이하 설명하는 기술을 이용하면 많은 양의 자원을 사용하지 않고서, 이미지를 압축할 수 있다. 이하 설명하는 기술을 이용하면 Clip함수 외부영역에 있는 정보도 잘 이용함으로써 압축 모델을 잘 학습시킬 수 있다. 이하 설명하는 기술을 이용하면, 양자화 오차를 최소화하면서 이미지 압축을 할 수 있다.

도면의 간단한 설명

- [0014] 도1은 영상 처리 장치(100)가 이미지를 압축하는 전체적인 과정을 도시한 것이다.
- 도2는 압축 모델의 일 실시예를 보여준다.
- 도3은 QConv의 작동방식을 보여준다.
- 도4는 Leaky-Clip 모듈이 이용하는 Clip함수 및 수정된 Clip함수를 보여준다.
- 도5는 종래 손실함수를 이용하는 경우에 발생할 수 있는 문제점을 보여준다.
- 도6은 영상 처리 장치(200)의 구성의 예이다.

발명을 실시하기 위한 구체적인 내용

- [0015] 이하 설명하는 기술은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있다. 명세서의 도면에 이하 설명하는 기술의 특정 실시 형태가 기재될 수 있다. 그러나, 이는 이하 설명하는 기술의 설명을 위한 것이며 이하 설명하는 기술을 특정한 실시 형태에 대해 한정하려는 것이 아니다. 따라서 이하 설명하는 기술의 사상 및 기술 범위에 포함되는 모든 변경 물, 균등 물 내지 대체 물이 이하 설명하는 기술에 포함하는 것으로 이해되어야 한다.
- [0016] 이하 사용되는 용어에서 단수의 표현은 문맥상 명백하게 다르게 해석되지 않는 한 복수의 표현을 포함하는 것으로 이해되어야 하고, "포함한다" 등의 용어는 기재된 특징, 개수, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함을 의미하는 것이지, 하나 또는 그 이상의 다른 특징들이나 개수, 단계 동작 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 배제하지 않는 것으로 이해되어야 한다.
- [0017] 도면에 대한 상세한 설명을 하기에 앞서, 본 명세서에서의 구성부들에 대한 구분은 각 구성부가 담당하는 주기능 별로 구분한 것에 불과함을 명확히 하고자 한다. 즉, 이하에서 설명할 2개 이상의 구성부가 하나의 구성부로 합쳐지거나 또는 하나의 구성부가 보다 세분화된 기능별로 2개 이상으로 분화되어 구비될 수도 있다. 그리고 이하에서 설명할 구성 부 각각은 자신이 담당하는 주기능 이외에도 다른 구성부가 담당하는 기능 중 일부 또는 전부의 기능을 추가적으로 수행할 수도 있으며, 구성부 각각이 담당하는 주기능 중 일부 기능이 다른 구성부에 의해 전담되어 수행될 수도 있음은 물론이다.
- [0018] 또, 방법 또는 동작 방법을 수행함에 있어서, 상기 방법을 이루는 각 과정들은 문맥상 명백하게 특정 순서를 기재하지 않은 이상 명기된 순서와 다르게 일어날 수 있다. 즉, 각 과정들은 명기된 순서와 동일하게 일어날 수도 있고 실질적으로 동시에 수행될 수도 있으며 반대의 순서대로 수행될 수도 있다.
- [0020] 이하 영상 처리 장치가 이미지를 압축하는 전체적인 과정을 살펴본다.
- [0021] 도1은 영상 처리 장치(100)가 이미지를 압축하는 전체적인 과정을 도시한 것이다.
- [0022] 이미지 압축 과정은 영상 처리 장치가 이미지를 입력 받는 단계; 상기 영상 처리 장치가 입력 받은 이미지를 압

축 모델에 입력하여 이미지를 압축하는 단계; 및 상기 영상 처리 장치가 상기 압축된 이미지를 상기 압축 모델에 입력하여 이미지를 복원하는 단계; 포함한다.

- [0023] 압축 모델은 컨볼루션(Convolution) 연산을 수행하는 모델을 포함할 수 있다.
- [0024] 압축 모델은 상기 입력 받은 이미지를 Leaky-Clip 모듈 및 라운딩 함수(Rounding Function)를 이용해 양자화(Quantization) 한 뒤, 상기 양자화 된 입력 받은 이미지와 미리 학습된 양자화된 가중치(Weight)를 이용해서 컨볼루션 연산을 수행할 수 있다.
- [0025] Leaky-Clip 모듈은 상기 압축 모델의 추론(Inference) 과정에서는 Clip함수의 기능을 수행하고, 상기 압축 모델의 학습(Training) 과정에서는 수정된 Clip 함수의 기능을 수행할 수 있다.
- [0026] Clip 함수는 내부영역(Inside)의 기울기는 1, 외부영역(Outside)의 기울기는 0인 함수일 수 있다.
- [0027] 수정된 Clip함수는 내부영역의 기울기는 1, 외부영역의 기울기는 0초과 1이하인 함수일 수 있다.
- [0029] 이하 압축 모델에 대해 구체적으로 설명한다.
- [0030] 본 발명의 압축 모델은 기본적으로 “Joint autoregressive and hierarchical priors for learned image compression(D. Minnen, 2018)” 논문에서 제시한 압축 모델 기반의 모델일 수 있다
- [0031] 도2는 압축 모델의 일 실시예를 보여준다.
- [0032] 압축 모델은 메인 인코더(g_a), 하이퍼 인코더(h_a), 메인 디코더(g_s), 하이퍼 디코더(h_s), 산술부호화(Arithmetic Encoder, AE) 및 산술복호화(Arithmetic Decoder, DE)를 포함할 수 있다.
- [0033] 메인 인코더는 입력 받은 이미지(input image)로부터 특징(y)을 추출한다. 하이퍼 인코더는 메인 인코더의 출력값을 받아 특징(z)을 더 추출한다. 메인 디코더는 엔트로피 코딩의 출력값을 입력 받아 이미지를 재 구축(\hat{x})한다. 하이퍼 디코더는 엔트로피 파라미터(Entropy parameter, μ_y, σ_y)을 예측한다. 산술부호화 및 산술복호화는 엔트로피 코딩을 하는데 이용된다.
- [0034] 수학식1는 메인 인코더(g_a)와 메인 디코더(g_s)에 출력되는 값을 표현한 것이다.

수학식 1

$$y = g_a(x), \hat{y} = [y], \hat{x} = g_s(\hat{y})$$

- [0035]
- [0036] 수학식1에서 x 는 입력 이미지(input image)를 의미한다. 수학식1에서 y 는 입력 이미지 x 의 레이턴트(latent)를 의미한다. 수학식1에서 \hat{y} 는 양자화된 레이턴트(Quantized latent) y 를 의미한다. 수학식1에서 $[\]$ 는 라운딩 함수(Rounding function)를 의미한다. 라운딩 함수는 입력 받은 값을 반올림하는 함수일 수 있다. 수학식1에서 \hat{x} 는 \hat{y} 로 부터 재구축된 이미지(Reconstructed image)를 의미한다.
- [0037] 수학식2는 하이퍼 인코더(h_a)와 하이퍼 디코더(h_s)에서 출력되는 값을 표현할 것이다.

수학식 2

$$z = h_a(y), \hat{z} = [z], (\mu_y, \sigma_y) = h_s(\hat{z})$$

[0038]

- [0039] 수학식2에서 y 는 입력 이미지 x 의 레이턴트(latent)를 의미한다. 수학식2에서 z 는 하이퍼 인코더가 출력한 레이턴트(Latent)를 의미한다. 수학식2에서 \hat{z} 는 양자화된 레이턴트 z 를 의미한다. 수학식2에서 $[\]$ 는 라운딩 함수를 의미한다. 수학식2에서 μ_y, σ_y 는 가우시안 분포를 따르는 엔트로피 파라미터(평균과 스케일)을 의미한다. 엔트로피 파라미터는 엔트로피 코딩을 하는데 이용될 수 있다. 즉 \hat{y} 의 엔트로피를 추측하는데 이용될 수 있다.
- [0040] 압축 모델의 메인 인코더, 메인 디코더, 하이퍼 인코더 및 하이퍼 디코더는 활성화 함수(Activation Function)로 ReLu 함수를 이용할 수 있다.
- [0042] 이하 QConv의 작동 방식을 살펴본다.
- [0043] 압축 모델의 메인 인코더, 메인 디코더, 하이퍼 인코더 및 하이퍼 디코더는 복수개의 양자화 컨볼루션 레이어(Quantized Convolution Layer, 이하 QConv)를 포함할 수 있다.
- [0044] QConv는 복수개의 필터를 이용해 입력 정보로부터 특징을 추출할 수 있다. 일 실시예로 QConv는 5 x 5의 크기를 가지는 128개의 필터(커널)를 이용해 입력 정보로부터 특징을 추출할 수 있다.
- [0045] QConv는 입력 받은 값을 양자화 한 뒤 컨볼루션 연산을 수행할 수 있다. QConv가 입력 받은 값을 양자화 한 뒤 계산함으로써 연산에 필요한 자원을 감소할 수 있다.
- [0046] QConv는 학습(Training) 과정과 학습 후 추론(Inference) 과정에서의 작동 방식이 다르다. 학습 과정에서 QConv의 가중치 w 는 업데이트 된다. 추론 과정에서 QConv는 정수값으로 이루어진 가중치 \bar{w} 와 양자화된 \bar{a} 를 이용해서 컨볼루션 연산을 수행한다.
- [0047] 도3은 QConv의 작동방식을 보여준다. 도3의 (a)는 학습 과정에서의 작동 방식을 보여준다. 도3의 (b)는 추론 과정에서의 작동 방식을 보여준다.
- [0048] QConv의 학습과정에서의 작동 방식을 살펴본다(도3의 (a)).
- [0049] QConv는 학습 이미지 및 학습 대상 가중치를 Leaky-Clip 모듈 및 라운딩 함수를 이용해 양자화 한 뒤, 양자화된 학습 이미지 및 양자화된 학습 대상 가중치를 이용해 컨볼루션 연산을 수행함으로써 학습될 수 있다.
- [0050] 다시 말하면 다음과 같다. 학습과정에서 QConv는 32비트의 실수값(float 32)을 가지는 가중치(w)와 입력 정보(a)를 입력 받는다. 입력 정보(a)는 입력 이미지 또는 이전 층의 출력 값일 수 있다. 입력 정보(a)는 활성화 값(Activations)이라고 불릴 수도 있다. QConv는 입력 받은 값을 스케일 파라미터(s_w, s_a)로 나눈 뒤, Leaky-Clip 모듈에 입력한다. QConv는 Leaky-Clip 모듈의 출력값에 라운딩 함수를 처리하여 8비트의 정수값(Int 8)을 가지는 가중치(\bar{w})와 정보(\bar{a})를 계산한다. QConv는 정수값을 가지는 가중치와 정수값 각각에 스케일 파라미터를 곱한 뒤, 곱한 값(\hat{w}, \hat{a})을 기반으로 컨볼루션 연산을 수행한다.
- [0051] QConv의 추론과정에서의 작동 방식을 살펴본다(도3의 (b)).
- [0052] QConv는 입력 받은 이미지를 입력 정보 스케일 파라미터로 나눈 값을 Leaky-Clip 모듈 및 라운딩 함수를 이용해 양자화 한 뒤 컨볼루션 연산을 수행하고, 컨볼루션 연산결과에 입력 정보 스케일 파라미터 및 가중치 스케일 파라미터를 곱할 수 있다.
- [0053] 다시 말하면 다음과 같다. 추론과정에서QConv는 32비트의 실수값(Float 32)을 가지는 입력 정보(a)를 입력 받는다. Conv는 입력 정보를 스케일 파라미터(s_a)로 나눈 뒤, Leaky-clip 모듈에 입력한다. QConv는 Leaky-clip 모듈의 출력값에 라운딩 함수를 처리하여 8비트 정수값을 가지는 정보(\bar{a})를 계산한다. QConv는 정수값을 가지는 정보와 학습된 가중치(\bar{w})를 기반으로 컨볼루션 연산을 수행한다. 컨볼루션 연산이 끝난 값(\bar{a})에 스케일 파라미터(s_w, s_a)를 곱한다.

[0054] 수학식3는 가중치(w)의 양자화 과정을 표현한 식이다. 다만 후술하듯이 수학식3는 Forward과정에서 이용되는 식이고, Backward 과정에서는 수정된 Clip함수가 이용된다.

수학식 3

[0055]
$$\bar{w} = \left[clip \left(\frac{w}{s_w}, L_w, U_w \right) \right]$$

[0056] 수학식3에서 w 는 가중치(weight)를 의미한다. 수학식3에서 \bar{w} 는 w 를 양자화 한 것을 의미한다. 수학식3에서 s_w 는 가중치의 양자화 파라미터를 의미한다. 수학식3에서 L_w 및 U_w 는 Leaky-Clip 모듈에서의 최소값 및 최대값을 의미한다. 가중치 양자화 파라미터(s_w)는 역 양자화된(Dequantized)된 가중치 \hat{w} 의 스텝 사이즈(step size)를 결정하는데 이용된다. 즉 \hat{w} 는 $\hat{w} = \bar{w} * s_w$ 이다.

[0057] 수학식3의 가중치 양자화 과정과 유사하게 입력 정보(a)의 양자화 과정은 수학식4와 같이 표현될 수 있다. 다만 후술하듯이 수학식4는 Forward과정에서 이용되는 식이고, Backward 과정에서는 수정된 Clip함수가 이용된다.

수학식 4

[0058]
$$\bar{a} = \left[clip \left(\frac{a}{s_a}, L_a, U_a \right) \right]$$

[0059] 수학식4에서 a 는 입력 받는 정보를 의미한다. 수학식4에서 s_a 는 정보의 양자화 파라미터를 의미한다. 정보의 양자화 파라미터(s_a)는 가중치의 양자화 파라미터와 유사하다. 즉 정보의 양자화 파라미터는 역 양자화된 정보 \hat{a} 의 스텝 사이즈를 결정하는데 이용된다. 즉 \hat{a} 는 $\hat{a} = \bar{a} * s_a$ 이다.

[0060] 일 실시예로 가중치와 입력 정보를 8비트로 양자화 하는 경우를 본다. \bar{w} 는 -127 내지 127 사이의 정수 값을 가질 수 있다. \bar{a} 는 0 내지 255의 사이의 정수 값을 가질 수 있다. \hat{w} 는 -127 x s_w 내지 127 x s_w 사이의 정수 값을 가질 수 있다. \hat{a} 는 0 내지 255 x s_a 사이의 정수 값을 가질 수 있다. 가중치의 클리핑 범위(Clipping range)는 $L_w * s_w$ 내지 $U_w * s_w$ 로 결정된다. 입력값의 클리핑 범위(Clipping range)는 $L_a * s_a$ 내지 $U_a * s_a$ 로 결정된다.

[0061] 압축 모델의 메인 디코더 및 하이퍼 디코더는 양자화 디 컨볼루션 레이어(Quantized Deconvolution layer, 이하 QDeconv)를 포함할 수 있다. QDeconv는 QConv한 방식으로 디 컨볼루션 연산(또는 Up sampling)을 수행할 수 있다.

[0063] 이하 Leaky-clip 모듈에 대해 설명한다.

[0064] 도4는 Leaky-Clip 모듈이 이용하는 Clip함수 및 수정된 Clip함수를 보여준다.

[0065] Leaky-clip 모듈은 Forward과정에서는 Clip 함수를 이용해 클리핑을 진행한다(도4의 (a)). Leaky-Clip모듈은 Backward과정에서는 수정된 Clip함수를 이용해 클리핑을 진행한다(도4의 (b)).

[0066] 수정된 Clip 함수는 일반적인 Clip함수와는 다르게 외부(Outside)부분에 기울기가 0이 아니다(Slope = γ) .

[0067] Leaky-Clip 모듈이 Backward과정에서 수정된 Clip함수를 이용해 클리핑을 하는 이유는 단순한 Clip 함수만 이용하면 학습에 영향을 많이 주는 외부(outside)에 있는 값이 학습 영향을 줄 수 없기 때문이다. 즉 단순한 Clip함수는 외부영역에서 기울기 값이 0이 되어 있기 때문에, 외부에 있는 값이 학습에 영향을 주지 못하는 문제가 있었다.

[0068] 표1은 전술한 문제점이 발생할 수 있는 한 경우를 보여준다.

표 1

a_i	$\sum \mathcal{G}_i$	# nonzero \mathcal{G}_i	# positive \mathcal{G}_i	# negative \mathcal{G}_i
Inside	1.53×10^{-5}	6.14×10^5	3.10×10^5	3.04×10^5
Outside	-2.00×10^{-5}	11	5	6

[0069]

[0070] 표1은 Clip함수의 내부(inside) 및 외부(outside)에 있는 입력 정보(a_i)의 개수(i)를 보여준다. 입력 정보의 개수는 \mathcal{G}_i 가 0이 아닌 것(non-zero), 양의 값인 것(Positive) 및 음의 값인 것(Negative)으로 분류되었다.

[0071] 표1의 \mathcal{G}_i 는 손실함수의 항 중 L_{RD} 를 정보의 스케일 파라미터로 미분한 것을 의미한다(수학식5). L_{RD} 는 후술하는 바와 같이 손실함수 중에서 압축된 비트스트림의 길이를 최소화하고 재구축된 왜곡(Distortion)을 최소화하는 항이다

수학식 5

$$\mathcal{G}_i = \frac{\partial a_i}{\partial s_a} \cdot \frac{\partial \mathcal{L}_{RD}}{\partial a_i}$$

[0072]

[0073] 외부(Outside)에 있는 입력 정보의 개수는 총 11개이다. 내부에 있는 입력 정보(Inside)의 개수는 6.14×10^5 이다. 즉 외부에 있는 입력 정보의 개수가 내부에 있는 입력정보의 개수보다 적다. 하지만 내부와 외부에 있는 입력값에 따른 기울기의 변화량의 총합($\sum_i \mathcal{G}_i$)은 유사하다. 즉 외부의 11개의 입력 정보에 따른 손실함수의 기울기가 내부의 60만개의 입력 정보에 따른 손실함수의 기울기를 이길만큼 영향력이 크다. 하지만 단순히 clip 함수만을 이용할 경우 외부 기울기는 제거되기 때문에 학습에 이용되지 못한다. 이는 기울기가 0이기 때문이다.

[0074] 따라서 Leaky-Clip 모듈은 Forward 과정에서는 Clip 모듈을 그대로 이용하나, Backward 과정에서는 외부에 기울기를 내부의 기울기로 일부 조정하여 남겨준다. 예를 들어 외부의 기울기를 1/100으로 조정하여 남겨둘 수 있다. 이렇게 하여 Clip 바깥쪽에 중요한 기울기를 학습에 이용할 수 있다. 외부의 기울기(γ)는 상황에 따라 0과 1사이 값을 가질 수 있다.

[0076] 이하 본 발명의 학습과정에서 사용되는 손실함수에 대해 살펴본다.

[0077] 일반적으로 압축 모델은 압축된 비트스트림의 길이를 최소화하고 재구축된 이미지의 왜곡을 최소화하는 방향으로 학습된다.

[0078] 따라서 수학식6와 같은 손실함수를 이용할 수 있다.

수학식 6

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_{RD} \\ \mathcal{L}_{RD} &= R + \lambda \cdot D \\ &= \mathbf{E}_{x \sim p_x} [-\log_2 p_{\hat{y}|\hat{z}}(\hat{y}|\hat{z}) - \log_2 p_{\hat{z}}(\hat{z})] + \lambda \cdot \mathbf{E}_{x \sim p_x} [d(x, \hat{x})] \end{aligned}$$

[0079]

[0080]

수학식6에서 L_{RD} 는 압축된 비트스트림의 길이를 최소화하고 재구축된 왜곡(Distortion)을 최소화하는 항(Terms)이다. 수학식6에서 λ 는 R 과 D 사이의 트레이드-오프(trade-off)를 조절한다. 수학식6에서 p_x 는 원본 이미지(Original image)의 분포(distribution)을 의미한다. 수학식6에서 R 는 \hat{y} 의 추정 엔트로피와 \hat{z} 의 추정 엔트로피의 합을 의미한다. 수학식6에서 D 는 원본 이미지와 재구축된 이미지 사이의 평균 제곱 에러(Mean squared error, MSE)를 의미한다.

[0081]

더 나아가 압축 모델은 양자화 에러를 감소시키는 방향으로도 학습되어야 한다. 즉 양자화 되기 전과 후의 오차가 작아야 한다. 하지만 수학식6와 같은 손실함수만을 이용하면 양자화 에러를 줄이는 방향으로 학습되기 어려웠다.

[0082]

도5는 종래 손실함수를 이용하는 경우에 발생할 수 있는 문제점을 보여준다. 도5는 정보(a), 역 양자화된 정보 \hat{a} 및 스케일 파라미터(s_a) 사이의 관계를 보여준다.

[0083]

도5의 (a)는 정보의 스케일 파라미터 값이 1인 경우이다. 도5의 (b)는 정보의 스케일 파라미터 값이 0.67(=2/3)인 경우이다.

[0084]

스케일 파라미터(s_a)가 커지면 클리핑 범위 바깥 쪽에서(a_1 및 a_4) 양자화 에러가 감소한다. 반면에 스케일 파라미터(s_a)가 커지면 클리핑 범위 안쪽에서 (a_2 및 a_3)에서 양자화 에러가 커진다. (도5의 (a))

[0085]

반대로 스케일 파라미터(s_a)가 작아지면 클리핑 범위 바깥 쪽에서(a_1 및 a_4) 양자화 에러가 커진다. 또한 스케일 파라미터(s_a)가 작아지면 클리핑 범위 안쪽에 (a_2 및 a_3)에서 양자화 에러가 감소한다. (도5의 (b))

[0086]

따라서 학습과정에서 양자화 에러를 최소화하기 위해서 최적화된 스케일 파라미터(s_a)를 찾아야 한다. 하지만 종래 손실함수만을 이용하는 경우 클리핑 범위 안쪽(Inside)에서 스케일 파라미터를 증가시키려고 하는 경우가 많았다.

[0087]

예를 들어 클리핑 범위 안쪽(inside)에 있는 a_i 에 따른 \mathcal{G}_i 의 부호는 정보의 스케일 파라미터를 감소시키기 위해서 주로 양(Positive)의 값을 가질 것으로 기대된다. 왜냐하면 스케일 파라미터가 작아질수록 클리핑 범위 안쪽에 있는 값들에 대한 양자화 에러가 감소되기 때문이다.

[0088]

하지만 표1에서 봤듯이 클리핑 범위 안쪽에 있는 정보의 약 절반 중에서는 음의 기울기를 형성한다. 이에 전술한 수학식6와 같은 손실함수만을 이용하는 경우, 압축 모델이 양자화 에러가 감소하는 방향으로 학습이 잘 되지 않을 수 있다.

[0089]

따라서 수학식6에 손실함수에 양자화 에러에 대한 항을 추가하여 수정할 필요가 있다.

[0090]

수학식7은 양자화 에러(L_{SQE})에 대한 항을 추가한 손실함수(L)이다.

수학식 7

$$\mathcal{L} = \mathcal{L}_{RD} + \mathcal{L}_{SQE}$$

$$\mathcal{L}_{RD} = R + \lambda \cdot D$$

$$= \mathbf{E}_{\mathbf{x} \sim p_{\mathbf{x}}} [-\log_2 p_{\hat{\mathbf{y}}|\hat{\mathbf{z}}}(\hat{\mathbf{y}}|\hat{\mathbf{z}}) - \log_2 p_{\hat{\mathbf{z}}}(\hat{\mathbf{z}})] + \lambda \cdot \mathbf{E}_{\mathbf{x} \sim p_{\mathbf{x}}} [d(\mathbf{x}, \hat{\mathbf{x}})]$$

$$\mathcal{L}_{SQE} = \beta \left\{ \sum (w - \hat{w})^2 + \sum (a - \hat{a})^2 \right\}$$

[0091]

[0092]

수학식7에서 \mathcal{L}_{SQE} 는 거듭제곱 양자화 에러(Squared Quantization Error Loss)에 대한 항이다. 수학식7에서 w 는 학습 대상 가중치이다. 수학식7에서 \hat{w} 는 학습 대상 가중치를 양자화한 후 역양자화 하였을 때의 값이다. 수학식7에서 a 는 입력 정보(또는 학습 이미지)이다. 수학식7에서 \hat{a} 는 학습 이미지를 양자화 한 후 역양자화 하였을 때의 값이다. 수학식 7에서 β 는 학습 대상과 대상을 양자화한 값의 차가 발산하는 것을 방지하기 위해 설정해주는 아주 작은 값이다.

[0094]

이하 연구자가 압축 모델을 구축하고 성능을 실험한 실험결과에 대해 살펴본다.

[0095]

실험을 위해 학습과정에서는 Flickr2W 데이터 셋을 이용하였다. 모든 입력 이미지는 256 x 256패치(Patch)가 되도록 랜덤하게 크로핑 되었다. 학습 후 평가 과정에서는 Kodak 데이터 셋을 이용하였다. 8의 배치 사이즈로 Adam optimizer 방식을 통해 베이스라인(Baseline) 모델을 백만회 반복 학습시켰다. 학습률(Learning rate)은 10^{-4} 로 설정한 하고, 마지막 10만회 학습과정에서는 10^{-5} 로 설정하였다. 양자화 모델은 사전 학습(Pre-trained)된 모델로부터 시작하여 학습되었다. 학습을 위해 CompressAI PyTorch library가 이용되었다.

[0096]

성능 평가를 위한 대조군으로 LSQ (learned step size quantization)방식 및 PAMS(parameterized max Scale)방식이 이용되었다. 성능 평가를 위한 지표로는 Bjontegaard delta bitrate (BD-Rate)가 이용되었다. BD-Rate는 동일한 퀄리티에서 비트레이트(Bitrate)의 증가율을 의미한다. 또한 성능 평가를 위한 지표로 Bjontegaard delta PSNR (BD-SNR)가 이용되었다. BD-SNR은 동일한 비트레이트(Bitrate)에서의 퀄리티 게인(Quality gain)을 의미한다.

[0097]

표2는 대조군과 비교한 실험결과 중 하나이다.

표 2

Quant.	Bit-width	BD-Rate ↓	BD-SNR (dB) ↑	Model size
Full-prec.	32	0%	0	26.43 MB
PAMS		+59.04%	-1.40	6.61 MB
LSQ	8	+1.98%	-0.07	6.62 MB
Proposed		-0.54%	+0.02	6.62 MB
PAMS		N.A.	-5.36	3.30 MB
LSQ	4	+68.17%	-1.68	3.31 MB
Proposed		+56.31%	-1.40	3.31 MB

[0098]

[0099]

PAMS 방식 및 LSQ 방식을 이용한 경우보다 전술한 압축 모델(Proposed)을 이용한 경우에 가장 낮은 BD-Rate 및 가장 높은 BD-SNR을 가지는 것을 확인할 수 있다. 특히 8비트로 양자화 하는 경우에 성능이 좋은 것을 확인할 수 있다.

[0100]

또한 전술한 압축 모델(Proposed)의 크기는 PAMS 방식 및 LSQ방식을 이용한 방식과 비교할 때 사이즈의 크기가

많이 차이 나지 않음을 확인할 수 있다.

[0102] 표3는 Ablation Study에 대한 실험 결과이다.

표 3

Bit-width	SQE	Leaky-Clip	BD-Rate ↓
8-bit	✓	-	-0.08%
	✓	✓	-1.01%
4-bit	✓	-	-6.90%
	✓	✓	-7.64%

[0103]

[0104] 표3는 SQE 손실과 Leaky-Clip 모듈을 적용할 때의 BD-Rate가 최소인 것을 보여준다. 즉 본 발명에 따르면 8비트로 양자화 한 경우에도 성능이 많이 감소되지 않는 것을 확인할 수 있다.

[0106] 이하 영상 처리 장치의 구성에 대하여 설명한다.

[0107] 도6은 영상 처리 장치의 구성의 예이다.

[0108] 영상 처리 장치(200)는 도1에서 설명한 영상 처리 장치(100)에 해당할 수 있다. 영상 처리 장치는 전술한 이미지 압축 방법을 수행할 수 있다.

[0109] 영상 처리 장치(200)는 물리적으로 다양한 형태로 구현될 수 있다. 예를 들어 영상 처리 장치(200)는 PC, 노트북, 스마트기기, 서버 또는 데이터처리 전용 칩셋 등의 형태를 가질 수 있다.

[0110] 영상 처리 장치(200)는 입력장치(210), 저장장치(220), 연산장치(230), 출력장치(240), 인터페이스 장치(250) 및 통신장치(260)를 포함할 수 있다.

[0111] 입력장치(210)는 일정한 명령 또는 데이터를 입력 받는 인터페이스 장치(키보드, 마우스, 터치스크린 등)를 포함할 수도 있다. 입력장치(210)는 별도의 저장장치(USB, CD, 하드디스크 등)를 통하여 정보를 입력 받는 구성을 포함할 수도 있다. 입력장치(210)는 입력 받는 데이터를 별도의 측정장치를 통하여 입력 받거나, 별도의 DB를 통하여 입력 받을 수도 있다. 입력장치(210)는 유선 또는 무선 통신을 통해 데이터를 입력 받을 수 있다.

[0112] 입력장치(210)는 이미지 압축 방법을 실행하는데 필요한 정보 및 모델을 입력 받을 수 있다. 입력장치(210)는 이미지를 입력 받을 수 있다. 입력장치(210)는 압축 모델을 입력 받을 수 있다.

[0113] 저장장치(220)는 입력장치(210)를 통해 입력 받은 정보 및 모델을 저장할 수 있다. 저장장치(220)는 연산장치(230)가 연산하는 과정에서 생성되는 정보를 저장할 수 있다. 즉 저장장치(220)는 메모리를 포함할 수 있다. 저장장치(220)는 연산장치(230)가 계산한 결과를 저장할 수 있다.

[0114] 저장장치(210)는 이미지 압축 방법을 실행하는데 필요한 정보 및 모델을 저장할 수 있다. 저장장치(210)는 이미지를 저장할 수 있다. 저장장치(210)는 압축 모델을 저장할 수 있다.

[0115] 연산장치(230)는 데이터를 처리하고, 일정한 연산을 처리하는 프로세서, AP, 프로그램이 임베디드 된 칩과 같은 장치일 수 있다. 연산장치(230)는 영상 처리 장치(200)를 제어하는 제어신호를 생성할 수 있다.

[0116] 연산장치(230)는 이미지 압축 방법을 실행하는데 필요한 연산을 수행할 수 있다. 연산장치(230)는 입력 받은 이미지를 압축 모델에 입력하여 이미지를 압축할 수 있다. 연산장치(230)는 압축된 이미지를 압축 모델에 입력하여 이미지를 복원할 수 있다.

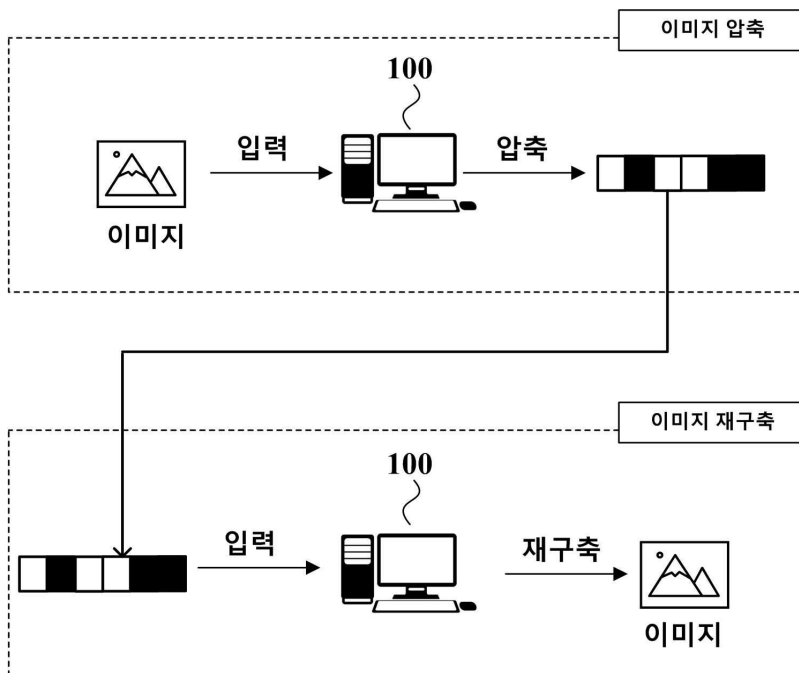
[0117] 출력장치(240)는 일정한 정보를 출력하는 장치가 될 수도 있다. 출력장치(240)는 데이터 과정에 필요한 인터페이스, 입력된 데이터, 분석결과 등을 출력할 수도 있다. 출력장치(240)는 디스플레이, 문서를 출력하는 장치, 등과 같이 물리적으로 다양한 형태로 구현될 수도 있다.

[0118] 인터페이스 장치(250)는 외부로부터 일정한 명령 및 데이터를 입력 받는 장치일 수 있다. 인터페이스 장치(250)는 물리적으로 연결된 입력 장치 또는 외부 저장장치로부터 이미지 및 모델을 입력 받을 수 있다.

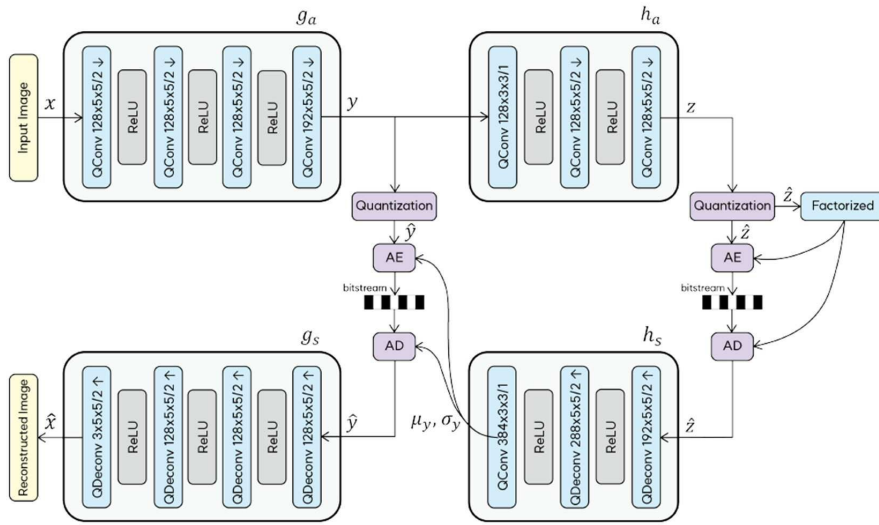
- [0119] 인터페이스 장치(250)는 영상 처리 장치(200)를 제어하기 위한 제어신호를 입력 받을 수 있다. 인터페이스 장치(250)는 영상 처리 장치(200)가 분석한 결과를 출력할 수 있다.
- [0120] 통신장치(260)는 유선 또는 무선 네트워크를 통해 일정한 정보를 수신하고 전송하는 구성을 의미할 수 있다. 통신장치(260)는 영상 처리 장치(200)를 제어하는데 필요한 제어 신호를 수신할 수 있다. 통신장치(260)는 영상 처리 장치(200)가 분석한 결과를 전송할 수 있다. 통신장치(260)는 압축 모델이 압축한 이미지를 다른 장치에 전송할 수 있다. 또는 통신장치(260)는 다른 장치가 압축한 이미지를 전송받을 수 있다.
- [0122] 전송한 이미지 압축 방법은 컴퓨터에서 실행될 수 있는 실행가능한 알고리즘을 포함하는 프로그램(또는 어플리케이션)으로 구현될 수 있다.
- [0123] 상기 프로그램은 일시적 또는 비일시적 판독 가능 매체(non-transitory computer readable medium)에 저장되어 제공될 수 있다.
- [0124] 비일시적 판독 가능 매체는 레지스터, 캐쉬, 메모리 등과 같이 짧은 순간 동안 데이터를 저장하는 매체가 아니라 반영구적으로 데이터를 저장하며, 기기에 의해 판독(reading)이 가능한 매체를 의미한다. 구체적으로는, 상술한 다양한 어플리케이션 또는 프로그램들은 CD, DVD, 하드 디스크, 블루레이 디스크, USB, 메모리카드, ROM (read-only memory), PROM (programmable read only memory), EPROM(Erasable PROM, EPROM) 또는 EEPROM(Electrically EPROM) 또는 플래시 메모리 등과 같은 비일시적 판독 가능 매체에 저장되어 제공될 수 있다.
- [0125] 일시적 판독 가능 매체는 스태틱 램(Static RAM, SRAM), 다이내믹 램(Dynamic RAM, DRAM), 싱크로너스 디램(Synchronous DRAM, SDRAM), 2배속 SDRAM(Double Data Rate SDRAM, DDR SDRAM), 증강형 SDRAM(Enhanced SDRAM, ESDRAM), 동기화 DRAM(Synclink DRAM, SLDRAM) 및 직접 램버스 램(Direct Rambus RAM, DRRAM) 과 같은 다양한 RAM을 의미한다.
- [0126] 본 실시예 및 본 명세서에 첨부된 도면은 전술한 기술에 포함되는 기술적 사상의 일부를 명확하게 나타내고 있는 것에 불과하며, 전술한 기술의 명세서 및 도면에 포함된 기술적 사상의 범위 내에서 당업자가 용이하게 유추할 수 있는 변형 예와 구체적인 실시예는 모두 전술한 기술의 권리범위에 포함되는 것이 자명하다고 할 것이다.

도면

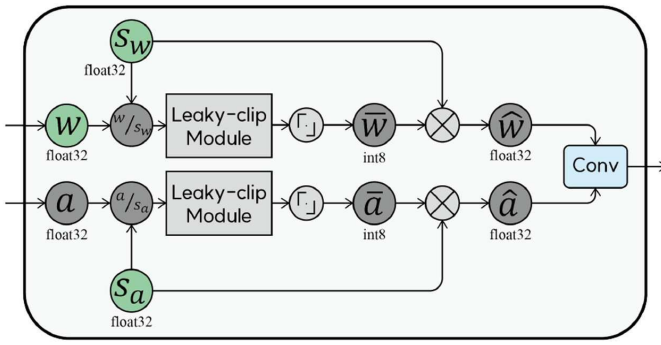
도면1



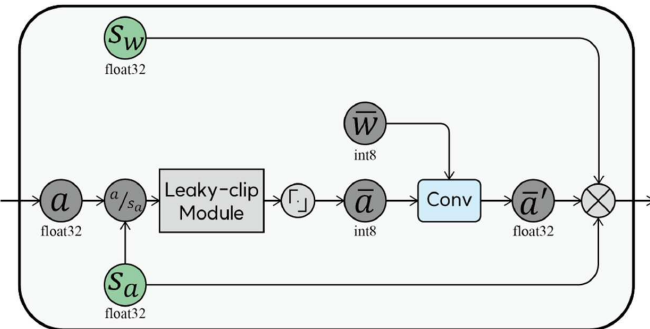
도면2



도면3

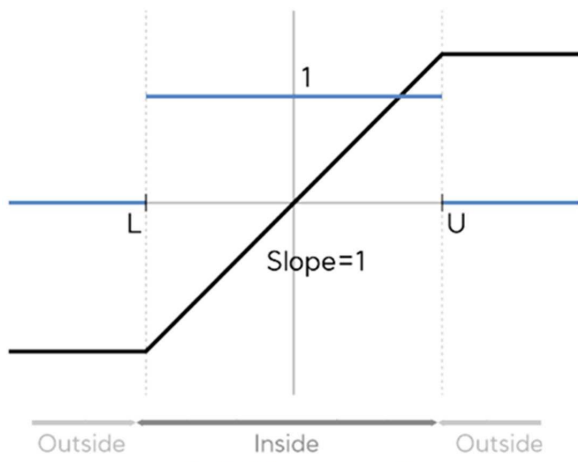


(a)

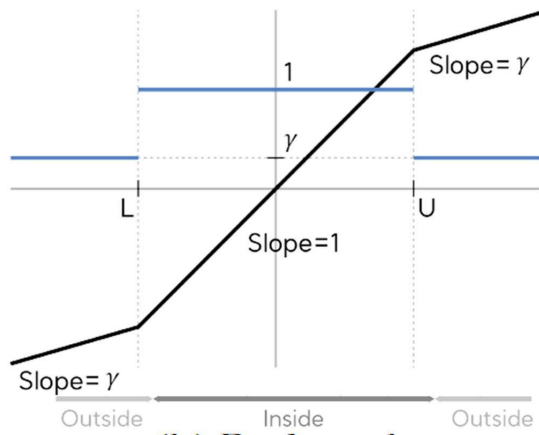


(b)

도면4

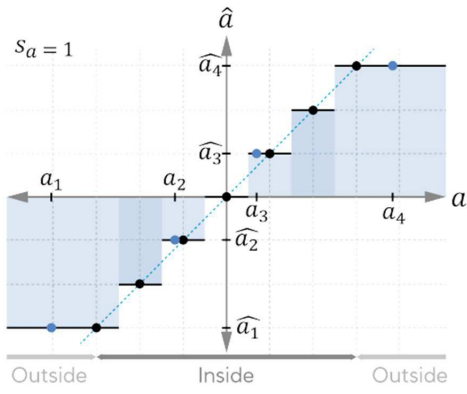


(a) Forward

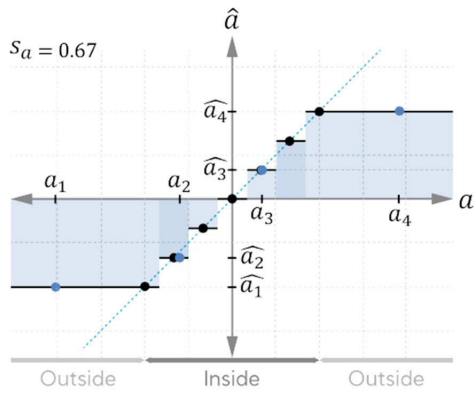


(b) Backward

도면5



(a) $s_a = 1$



(b) $s_a = 0.67$

도면6

