



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년10월28일
(11) 등록번호 10-2722957
(24) 등록일자 2024년10월23일

- | | |
|---|--|
| <p>(51) 국제특허분류(Int. Cl.)
H04N 19/14 (2014.01) G06N 3/0464 (2023.01)
G06N 3/08 (2023.01) G06T 3/40 (2024.01)
H04N 19/119 (2014.01) H04N 19/176 (2014.01)</p> <p>(52) CPC특허분류
H04N 19/14 (2015.01)
G06N 3/0464 (2023.01)</p> <p>(21) 출원번호 10-2023-0166491
(22) 출원일자 2023년11월27일
심사청구일자 2023년11월27일</p> <p>(56) 선행기술조사문헌
KR1020020047959 A
KR1020020041812 A</p> | <p>(73) 특허권자
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)</p> <p>(72) 발명자
이채은
서울특별시 서초구 남부순환로 2183(방배동, 방배래미안타워) 102동 2102호)
정영철
충청남도 아산시 방축로 30-5(방축동, 성정빌라)206호</p> <p>(74) 대리인
특허법인아이더스</p> |
|---|--|

전체 청구항 수 : 총 2 항

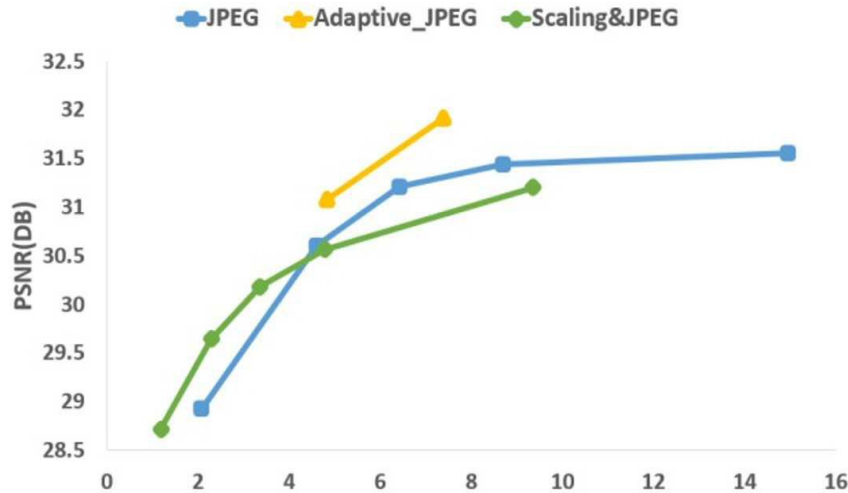
심사관 : 전용욱

(54) 발명의 명칭 프레임 보간 딥러닝 네트워크를 위한 피쳐맵 압축 방법

(57) 요약

본 발명에 의한 프레임 보간 딥러닝 네트워크를 위한 피쳐맵 압축 방법은, JPEG 기반 피쳐맵을 8×8 블록 단위로 나누고, 각각의 블록의 에지 밸류가 50 이상인 픽셀 개수를 기준으로 서로 다른 압축 수준을 부여하고, 각각의 블록의 에지 밸류가 50 이상인 픽셀 개수가 적을수록 낮은 JPEG 압축 퀄리티로 설정하는 것을 특징으로 한다.

대표도 - 도1



(52) CPC특허분류

G06N 3/08 (2023.01)
G06T 3/4046 (2024.01)
H04N 19/119 (2015.01)
H04N 19/176 (2015.01)
G06T 2207/20084 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711193592
과제번호	2021-0-02052-003
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	스마트 모빌리티를 위한 인공지능 시스템반도체 핵심 기술 개발 및 인력 양성
기 여 율	1/1
과제수행기관명	인하대학교 산학협력단
연구기간	2023.01.01 ~ 2023.12.31

공지예외적용 : 있음

명세서

청구범위

청구항 1

JPEG 기반 피쳐맵을 8×8 블록 단위로 나누고,

각각의 블록의 에지 밸류가 50 이상인 픽셀 개수를 기준으로 서로 다른 압축 수준을 부여하고,

각각의 블록의 에지 밸류가 50 이상인 픽셀 개수가 적을수록 낮은 JPEG 압축 퀄리티로 설정하는 것을 특징으로 하는 프레임 보간 딥러닝 네트워크를 위한 피쳐맵 압축 방법.

청구항 2

청구항 1에 있어서,

사용자가 어댑티브 퀄리티 1과 어댑티브 퀄리티 2 중 하나를 선택할 수 있고,

어댑티브 퀄리티 1이 선택된 상태에서는,

에지 밸류가 50 이상인 픽셀 개수가 0~4개일 때 JPEG 압축 퀄리티를 10으로,

에지 밸류가 50 이상인 픽셀 개수가 5~9개일 때 JPEG 압축 퀄리티를 20으로,

에지 밸류가 50 이상인 픽셀 개수가 10~14개일 때 JPEG 압축 퀄리티를 30으로,

에지 밸류가 50 이상인 픽셀 개수가 15~19개일 때 JPEG 압축 퀄리티를 50으로,

에지 밸류가 50 이상인 픽셀 개수가 20~24개일 때 JPEG 압축 퀄리티를 70으로,

에지 밸류가 50 이상인 픽셀 개수가 25~64개일 때 JPEG 압축 퀄리티를 90으로 설정하고,

어댑티브 퀄리티 2가 선택된 상태에서는,

에지 밸류가 50 이상인 픽셀 개수가 0~4개일 때 JPEG 압축 퀄리티를 40으로,

에지 밸류가 50 이상인 픽셀 개수가 5~9개일 때 JPEG 압축 퀄리티를 50으로,

에지 밸류가 50 이상인 픽셀 개수가 10~14개일 때 JPEG 압축 퀄리티를 60으로,

에지 밸류가 50 이상인 픽셀 개수가 15~19개일 때 JPEG 압축 퀄리티를 70으로,

에지 밸류가 50 이상인 픽셀 개수가 20~24개일 때 JPEG 압축 퀄리티를 80으로,

에지 밸류가 50 이상인 픽셀 개수가 25~64개일 때 JPEG 압축 퀄리티를 90으로 설정하는 것을 특징으로 하는 프레임 보간 딥러닝 네트워크를 위한 피쳐맵 압축 방법.

발명의 설명

기술 분야

[0001]

본 발명은 프레임 보간 딥러닝 네트워크를 위한 피쳐맵 압축 방법에 관한 것으로서, 더욱 상세하게는, 피쳐맵의 품질을 효율적으로 유지하기 위한 프레임 보간 딥러닝 네트워크를 위한 피쳐맵 압축 방법에 관한 것이다.

배경 기술

[0002]

최근 증강현실과 자율 주행 등 인공지능과 영상처리가 결합된 분야가 발전하면서, 이미지 뿐 아니라 동영상 처리를 위한 딥러닝 기술에 대한 요구 사항도 크게 증가하고 있다. 한 예로 시스템의 성능과 안정성을 높이기 위해 프레임 레이트를 높이기 위한 프레임 인터폴레이션(Frame Interpolation)이 있다. 이 기술을 이용하면, 고가의 카메라를 사용하지 않고도 더 높은 품질의 동영상을 획득할 수 있다. CNN 기반 프레임 인터폴레이션(Frame

Interpolation)에서는 두 개의 이미지를 순서대로 입력 받아 모션벡터를 추출하여 중간 프레임을 생성한다. {선행문헌 1(Liu, Ziwei, et al. "Video frame synthesis using deep voxel flow." Proceedings of the IEEE international conference on computer vision. 2017.) 참조}

[0003] 하지만, Frame Interpolation 특성상 중간 프레임을 합성하기 위해서는 두 장의 이미지에 대한 연산결과가 모두 필요하기 때문에 먼저 입력받은 프레임의 피쳐맵(Feature map)을 다음 프레임에 대한 CNN 연산이 마무리되기 전 까지 온칩 메모리에 저장해야 하는 메모리 자원 측면에서 오버헤드가 매우 크다.

[0004] 피쳐맵을 압축하기 위해 다양한 연구에서 JPEG, H.264와 같은 이미지 및 영상 압축 분야에서 사용되는 코덱을 사용하였다. 선행문헌 2(Chen, Zhuo, et al. "Data representation in hybrid coding framework for feature maps compression." 2020 IEEE International Conference on Image Processing (ICIP). IEEE, 2020.)는 프리 콰нти제이션(Pre-quantization)과 채널 칸캐터네이션(channel concatenation) 및 틸팅(tiling)등 방식을 통해 피쳐맵을 전처리하여 압축 효율을 높인다. 선행문헌3(Xie, Xiufeng, and Kyu-Han Kim. "Source compression with bounded dnn perception loss for iot edge computer vision." The 25th Annual International Conference on Mobile Computing and Networking. 2019.), 선행문헌 4(Liu, Zihao, et al. "DeepN-JPEG: A deep neural network favorable JPEG-based image compression framework." Proceedings of the 55th annual design automation conference. 2018.)는 특정 색상 또는 주파수 구성 요소가 DNN 정확도에 미치는 영향을 기반으로 표준 JPEG의 양자화 테이블을 재설계하였다. 하지만 피쳐맵의 영역별 중요성을 고려하지 않은 균일한 압축 수준을 적용한다.

[0005] 따라서, 본 논문에서는 온칩 메모리 소모를 줄이고 피쳐맵의 품질을 효율적으로 유지하기 위한 방법이 필요하다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) 한국 등록특허공보 KR 10-2305229 B1(등록일: 2021. 9. 16.)

발명의 내용

해결하려는 과제

[0007] 본 발명의 목적은, 온칩 메모리 소모를 줄이고 피쳐맵의 품질을 효율적으로 유지하기 위한 피쳐맵 압축 방법을 제공하는 것이다.

과제의 해결 수단

[0008] 본 발명에 의한 프레임 보간 딥러닝 네트워크를 위한 피쳐맵 압축 방법은, JPEG 기반 피쳐맵을 8×8 블록 단위로 나누고, 각각의 블록의 에지 밸류가 50 이상인 픽셀 개수를 기준으로 서로 다른 압축 수준을 부여하고, 각각의 블록의 에지 밸류가 50 이상인 픽셀 개수가 적을수록 낮은 JPEG 압축 퀄리티로 설정하는 것을 특징으로 한다.

[0009] 상기 피쳐맵 압축 방법에서, 사용자가 어댑티브 퀄리티 1과 어댑티브 퀄리티 2 중 하나를 선택할 수 있고, 어댑티브 퀄리티 1이 선택된 상태에서는, 에지 밸류가 50 이상인 픽셀 개수가 0~4개일 때 JPEG 압축 퀄리티를 10으로, 에지 밸류가 50 이상인 픽셀 개수가 5~9개일 때 JPEG 압축 퀄리티를 20으로, 에지 밸류가 50 이상인 픽셀 개수가 10~14개일 때 JPEG 압축 퀄리티를 30으로, 에지 밸류가 50 이상인 픽셀 개수가 15~19개일 때 JPEG 압축 퀄리티를 50으로, 에지 밸류가 50 이상인 픽셀 개수가 20~24개일 때 JPEG 압축 퀄리티를 70으로, 에지 밸류가 50 이상인 픽셀 개수가 25~64개일 때 JPEG 압축 퀄리티를 90으로 설정하고, 어댑티브 퀄리티 2가 선택된 상태에서는, 에지 밸류가 50 이상인 픽셀 개수가 0~4개일 때 JPEG 압축 퀄리티를 40으로, 에지 밸류가 50 이상인 픽셀 개수가 5~9개일 때 JPEG 압축 퀄리티를 50으로, 에지 밸류가 50 이상인 픽셀 개수가 10~14개일 때 JPEG 압축 퀄리티를 60으로, 에지 밸류가 50 이상인 픽셀 개수가 15~19개일 때 JPEG 압축 퀄리티를 70으로, 에지 밸류가 50 이상인 픽셀 개수가 20~24개일 때 JPEG 압축 퀄리티를 80으로, 에지 밸류가 50 이상인 픽셀 개수가 25~64개일 때 JPEG 압축 퀄리티를 90으로 설정하도록 할 수 있다.

발명의 효과

[0010] 본 발명에 의한 프레임 보간 딥러닝 네트워크를 위한 피쳐맵 압축 방법은, 피쳐맵의 품질을 효율적으로 유지할 수 있다.

도면의 간단한 설명

[0011] 도 1은 피쳐맵 압축실험 결과 그래프

발명을 실시하기 위한 구체적인 내용

[0012] 이하, 첨부한 도면을 참조하여 본 발명의 바람직한 실시 예를 보다 상세하게 설명한다. 그리고, 본 발명은 다수의 상이한 형태로 구현될 수 있고, 기술된 실시 예에 한정되지 않는다.

[0013] JPEG 기반 피쳐맵 적응형 압축을 위해 8x8 블록(block) 단위로 압축수준을 설정할 필요가 있다. 피쳐맵 특성상 입력 프레임의 에지(edge)에서 높은 피쳐 값을 가지는 경향을 다양한 입력 프레임과 피쳐맵의 시각화를 통해 파악하였으며, 피쳐맵의 에지(edge)를 캐니 에지 디텍션(canny edge detection)(low:20/255, high:50/255)을 통해 획득한 뒤, 8x8 block 내부의 에지 밸류(edge value)가 50이상인 픽셀 개수를 기준으로, 다음의 표 1과 같이 block단위 압축 수준을 6단계로 나눈다.

표 1

[0014]	edge value > 50 개수	64~25	24~20	19~15	14~10	9~5	4~0
	Adaptive_quality_1	90	70	50	30	20	10
	Adaptive_quality_2	90	80	70	60	50	40

[0015] 상기 표 1에서 숫자가 작을수록 낮은 JPEG 압축 퀄리티인 것을 의미하고, 이것은 높은 압축률인 것을 의미한다.

[0016] 사용자는 필요에 따라 어댑티브 퀄리티 1(Adaptive_quality_1) 또는 어댑티브 퀄리티 2(Adaptive_quality_2) 중 하나를 선택하여 압축할 수 있도록 구현하는 것이 바람직하다.

[0017] 피쳐맵의 에지(edge)를 캐니 에지 디텍션(canny edge detection)을 통해 획득하는 방법의 예는 다음 4단계를 거치는 것이다.

[0018] (1) 제1 단계 - 노이즈 제거 (Noise Reduction):

[0019] 이미지에서 노이즈를 제거하기 위해 5x5 가우시안 필터를 사용한다.

[0020] (2) 제2 단계 - 그래디언트(Gradient) 값이 높은 부분 찾기(Finding Intensity Gradient of the Image):

[0021] 가우시안 필터로 노이즈가 제거된 이미지에 Sobel 커널을 수평방향, 수직방향으로 적용하여 각 방향의 gradient를 획득한다. 이 때, 각 픽셀의 에지(edge) 그래디언트와 방향은 다음과 같이 계산된다.

$$Edge_Gradient(G) = \sqrt{\{G_x^2 + G_y^2\}}$$

[0022]

$$Angle(\theta) = \tan^{-1}(\{G_y\}/\{G_x\})$$

[0023]

[0024] (3) 제3 단계 - 최대값이 아닌 픽셀의 값을 0으로 만들기(Non-maximum Suppression):

[0025] 에지에 기여하지 아니한 픽셀을 제거하기 위해 이미지 전체를 스캔한다. 이미지를 스캔하는 동안 gradient 방향으로 스캔 구역에서 gradient의 최대값을 가진 픽셀을 찾는다.

[0026] (4) 제4 단계 - 히스테리시스 스레스홀딩(Hysteresis Thresholding)

[0027] 3단계를 거친 것들이 실제 에지인지 아닌지 판단하는 단계이다. 먼저 문턱값을 minVal, maxVal 2개 잡는다. maxVal 보다 높은 부분은 확실한 에지이고 minVal보다 낮은 부분은 에지가 아니라고 판단한다. minVal, maxVal

사이에 있는 값들은 이 픽셀들의 연결구조를 보고 에지인지 아닌지 판단한다.

[0028] 이러한 과정을 거치면 이미지에서 에지로 판단된 부분만 남게 된다.

[0029] 또한 모든 8x8 DCT 코이피션트(coefficent)에 동일한 양자화 수준을 적용하는 기존 JPEG 압축 방식과 달리, 에지(edge) 기반 블록(block) 단위 어댑티브 칼리티(adaptive quality) 양자화를 적용하기 위해, 구간별 칼리티(quality) 설정 방식을 표 1과 같이 어댑티브 칼리티(Adaptive quality) 1, 2의 2가지로 압축 수준을 나눈다. 피쳐맵 밸류(value)가 낮고, 변화가 적은 로 프리퀀시 블록(low-frequency block)에서 낮은 칼리티(quality)의 양자화를 적용하고, 반대로 높은 피쳐맵 밸류(value)와 하이 프리퀀시 블록(high-frequency block)은 높은 칼리티(quality)의 양자화를 적용하여 전체 피쳐맵의 칼리티와 압축률을 조정한다.

[0030] 추가적으로 JPEG을 통한 피쳐맵 압축률을 높이기 위한 스케일링(Scaling) 방식을 제안한다. 다음의 식에서 보인 것과 같이 Min-Max scaling을 통해 피쳐맵들의 값 범위의 상관 관계를 높여 압축률을 높인다.

$$\frac{x - \text{Min}(X)}{\text{Max}(X) - \text{Min}(X)}$$

[0031]

[0032] 이때, 피쳐맵의 채널별 평균 밸류(value)가 높은 3개의 채널을 제외한 부분만 스케일링(Scaling)을 적용하여, 보간 프레임의 칼리티(quality)를 높인다.

[0033] 상기 방법을 구현하기 위해, 선행문헌 1(Liu, Ziwei, et al. "Video frame synthesis using deep voxel flow." Proceedings of the IEEE international conference on computer vision. 2017.)의 프레임 보간 모델과 JPEG 오리지널 로스리스 커프레션(Original Lossless Compression)(ISO/IEC 10918)을 사용하였다.

[0034] 도 1은 피쳐맵 압축실험 결과 그래프이다.

[0035] 상기 실험 결과는, 피쳐맵의 크기가 32x32 이고, 각각의 블록의 크기가 8x8인 사례를 통한 실험 결과이다.

[0036] 피쳐맵을 압축하지 않은 베이스라인(baseline) 피쳐맵을 5가지 압축수준으로 일반적인 JPEG을 통해 압축한 방법(BLUE), Min(X)를 0, Max(X)를 60으로 Scaling한 방법(GREEN), 마지막으로 2가지 적응형 압축수준을 적용한 방법(YELLOW)을 통해 생성된 JPEG 비트스트림(bitstream)을 디코딩하여 생성된 피쳐맵을 프레임 보간 모델 디코더에 입력하여 생성된 보간 프레임을 원본 프레임과 PSNR을 측정한 결과를 도 1에서 확인할 수 있다.

[0037] 베이스라인(Baseline) 피쳐맵의 크기는 32.77kB이며, PSNR은 31.59dB이다. 도 1에서 JPEG압축 대비 스케일링(Scaling)&JPEG 압축 방법이 압축률은 높은 반면 PSNR은 낮은 결과를 확인할 수 있다. 특히 높은 비트레이트(bitrate)에서 JPEG압축 대비 PSNR이 낮은 결과가 측정되는데, 이는 스케일링(Scaling)을 통해 피쳐맵의 분포를 조정하여 압축률을 높인 반면, 전반적인 피쳐맵의 실제 밸류(value)가 변경되었기 때문이다. 어댑티브(Adaptive) JPEG 방식은 JPEG방식 대비 같은 압축률에서 더 높은 PSNR을 확인할 수 있다.

[0038] 어댑티브 칼리티 1(Adaptive_quality_1) 방식은 4.82kB, 31.08dB로 측정되며, 어댑티브 칼리티 2(Adaptive_quality_2)방식은 7.37kB, 31.913dB로 측정되었다.

[0039] 표1의 에지 밸류(edge value) 기준, 각 단계에서 높은 칼리티(quality)로 설정한 어댑티브 칼리티 2(Adaptive_quality_2) 방식이 압축률은 떨어지지만, 높은 PSNR이 측정되며 이를 통해 단계별 칼리티(quality)를 어댑티브(adaptive)하게 조절하는 것으로 압축률과 보간 프레임 칼리티(quality)를 조절할 수 있다.

[0040] 결과적으로, 본 출원에서 제안한 스케일링(Scaling) 방식은 높은 압축률을 보이지만 낮은 PSNR이 측정되었고, 피쳐맵 적응형 압축 방식은 일반적인 JPEG 압축대비 같은 압축률에서 높은 PSNR을 확인할 수 있었다.

도면

도면1

