



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0136176
(43) 공개일자 2022년10월07일

(51) 국제특허분류(Int. Cl.)
G06T 9/00 (2019.01) G06M 3/04 (2006.01)
G06N 3/08 (2006.01)
(52) CPC특허분류
G06T 9/002 (2013.01)
G06M 3/04 (2013.01)
(21) 출원번호 10-2022-0037097
(22) 출원일자 2022년03월25일
심사청구일자 없음
(30) 우선권주장
1020210041562 2021년03월31일 대한민국(KR)

(71) 출원인
현대자동차주식회사
서울특별시 서초구 현릉로 12 (양재동)
기아 주식회사
서울특별시 서초구 현릉로 12 (양재동)
이화여자대학교 산학협력단
서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)
(72) 발명자
강제원
서울특별시 마포구 서강로 53 서강쌍용예가아파트 104-904호
유채화
서울특별시 은평구 통일로 925, 409호
박승욱
경기도 용인시 수지구 태봉로 17, 403동 302호
(74) 대리인
이철희

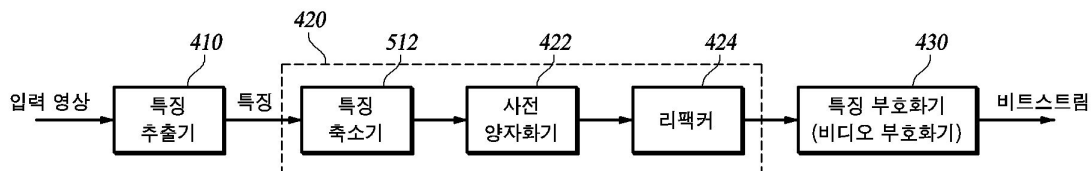
전체 청구항 수 : 총 19 항

(54) 발명의 명칭 특징맵 축소를 이용하는 머신 비전 데이터 코딩 장치 및 방법

(57) 요약

특징맵 축소를 이용하는 머신 비전 데이터 코딩 장치 및 방법에 관한 개시로서, 본 실시예는, 머신 태스크 특화된 딥러닝 모델에 의해 추출된 특징맵의 크기를 감소시키기 위해, 특징맵의 공간과 채널 측면의 불필요성을 감소시키는 희소화(sparsification) 방법, 및 텐서 분해(tensor decomposition)에 기반하는 특징맵 분해 방법을 이용하는 VCM(Video Coding for Machines) 코딩 장치 및 방법을 제공한다.

대표도



(52) CPC특허분류

G06N 3/082 (2013.01)

G06N 3/088 (2013.01)

명세서

청구범위

청구항 1

머신 비전 부호화 장치가 수행하는, 특징맵을 부호화하는 부호화 방법에 있어서,

딥러닝(deep learning) 모델을 이용하여 입력 영상으로부터 상기 특징맵을 추출하는 단계, 여기서, 상기 특징맵은 상기 딥러닝 모델의 중간 계층으로부터 생성됨;

상기 특징맵의 크기를 축소하여 축소(reduced) 특징맵을 생성하는 단계;

상기 축소 특징맵의 데이터 타입을 변환하고, 상기 축소 특징맵을 재배열하여 변형(converted) 특징맵을 생성하는 단계; 및

비디오 부호화기를 이용하여 상기 변형 특징맵을 부호화함으로써 비트스트림을 생성하는 단계

를 포함하는 것을 특징으로 하는, 부호화 방법.

청구항 2

제1항에 있어서,

상기 특징맵은,

동일한 높이와 너비를 갖는 2D 특징맵을 채널의 개수만큼 포함하는 것을 특징으로 하는, 부호화 방법.

청구항 3

제1항에 있어서,

상기 축소 특징맵을 생성하는 단계는,

상기 특징맵의 공간 또는 채널 측면에서의 특징맵 희소화(feature map sparsification)를 기반으로 상기 특징맵을 축소하는 것을 특징으로 하는, 부호화 방법.

청구항 4

제3항에 있어서,

상기 축소 특징맵을 생성하는 단계는,

상기 특징맵 희소화 및 텐서 분해를 결합하여 상기 특징맵을 축소하는 것을 특징으로 하는, 부호화 방법.

청구항 5

제3항에 있어서,

상기 축소 특징맵을 생성하는 단계는,

상기 특징맵을 구성하는 2D 특징맵 내에서, 영역의 특징 값이 기설정된 임계치보다 작은 경우, 상기 영역의 특징 값을 0으로 설정하는 것을 특징으로 하는, 부호화 방법.

청구항 6

제3항에 있어서,

상기 축소 특징맵을 생성하는 단계는,

기설정된 채널 간격을 갖는 두 개의 2D 특징맵 간의 거리를 산정하는 단계;

상기 거리가 기설정된 임계치보다 작은 경우에 해당하는 2D 특징맵의 짝들(pairs)을 모두 선택하는 단계; 및

상기 선택된 짝들 각각에 대해, 하나의 2D 특징맵의 모든 값을 0으로 희소화하거나 상기 하나의 2D 특징맵을 삭제하는 단계

를 포함하는 것을 특징으로 하는, 부호화 방법.

청구항 7

제6항에 있어서,

상기 비트스트림을 생성하는 단계는,

상기 하나의 2D 특징맵이 희소화된 경우, 희소화된 2D 특징맵, 및 상기 기설정된 채널 간격을 부호화하고, 상기 하나의 2D 특징맵이 삭제된 경우, 삭제된 2D 특징맵의 인덱스, 및 상기 기설정된 채널 간격을 부호화하는 것을 특징으로 하는, 부호화 방법.

청구항 8

제1항에 있어서,

상기 축소 특징맵을 생성하는 단계는,

텐서 분해(tensor decomposition)를 기반으로 상기 특징맵을 축소하는 것을 특징으로 하는, 부호화 방법.

청구항 9

제8항에 있어서,

상기 축소 특징맵을 생성하는 단계는,

터커 분해(Tucker decomposition)를 이용하여, 상기 특징맵을 하나의 커널 텐서(kernel tensor), 및 3 개의 인자 행렬들(factor matrices)로 분해하는 것을 특징으로 하는, 부호화 방법.

청구항 10

제8항에 있어서,

상기 축소 특징맵을 생성하는 단계는,

CP(Canonical Polyadic) 분해를 이용하여, 상기 특징맵을 P(여기서, P는 자연수) 개의 랭크(rank) 1 텐서들로 분해하는 것을 특징으로 하는, 부호화 방법.

청구항 11

딥러닝(deep learning) 모델을 이용하여 입력 영상으로부터 특징맵을 추출하는 특징 추출기, 여기서, 상기 특징맵은 상기 딥러닝 모델의 중간 계층으로부터 생성됨;

상기 특징맵의 크기를 축소하여 축소(reduced) 특징맵을 생성하는 특징 축소기(feature reducer);

상기 축소 특징맵의 데이터 타입을 변환하기 위해 상기 축소 특징맵을 양자화하는 사전 양자화기(pre-quantizer);

상기 양자화된 특징맵을 비디오 시퀀스의 형태로 재배열하여 변형(converted) 특징맵을 생성하는 리패커(repacker); 및

비디오 부호화기를 이용하여 상기 변형 특징맵을 부호화함으로써 비트스트림을 생성하는 특징 부호화기

를 포함하는 것을 특징으로 하는, 머신 비전 부호화 장치.

청구항 12

제11항에 있어서,

상기 특징 축소기는,

상기 특징맵의 공간 또는 채널 측면에서의 특징맵 희소화(feature map sparsification)를 기반으로 상기 특징맵

을 축소하는 것을 특징으로 하는, 머신 비전 부호화 장치.

청구항 13

제11항에 있어서,

상기 특징 축소는,

텐서 분해(tensor decomposition)를 기반으로 상기 특징맵을 축소하는 것을 특징으로 하는, 머신 비전 부호화 장치.

청구항 14

머신 비전 복호화 장치가 수행하는 복호화 방법에 있어서,

비디오 복호화기를 이용하여 비트스트림으로부터 변형 특징맵을 복호화하는 단계;

상기 변형 특징맵을 재배열하고, 상기 재배열된 변형 특징맵의 데이터 타입을 변환하여 축소(reduced) 특징맵을 복원하는 단계; 및

상기 축소 특징맵의 크기를 확장하여 복원 특징맵을 생성하는 단계

를 포함하되,

상기 복원 특징맵은 머신 비전 부호화 장치 내 딥러닝 모델의 중간 계층으로부터 생성된 특징맵에 대응하는 것을 특징으로 하는, 복호화 방법.

청구항 15

제14항에 있어서,

상기 복원 특징맵은,

동일한 높이와 너비를 갖는 2D 특징맵을 채널의 개수만큼 포함하는 것을 특징으로 하는, 복호화 방법.

청구항 16

제14항에 있어서,

상기 축소 특징맵이 채널 측면에서의 특징맵 희소화(feature map sparsification)를 기반으로 축소된 경우, 기 설정된 채널 간격 및 희소화된 2D 특징맵을 복호화하거나 상기 기설정된 채널 간격 및 삭제된 2D 특징맵의 인덱스를 복호화하는 단계를 더 포함하는 것을 특징으로 하는, 복호화 방법.

청구항 17

제16항에 있어서,

상기 복원 특징맵을 생성하는 단계는,

상기 희소화된 2D 특징맵에 대해 상기 기설정된 채널 간격의 앞 또는 뒤의 복원 2D 특징맵을 상기 희소화된 2D 특징맵의 위치에 복사하거나, 상기 삭제된 2D 특징맵에 대해 상기 인덱스를 참조하여 상기 기설정된 채널 간격의 앞 또는 뒤의 복원 2D 특징맵을 상기 삭제된 2D 특징맵의 위치에 복사함으로써, 상기 복원 특징맵을 생성하는 것을 특징으로 하는, 복호화 방법.

청구항 18

제14항에 있어서,

상기 복원 특징맵을 생성하는 단계는,

상기 축소 특징맵이 터커 분해(Tucker decomposition)를 이용하여 축소된 경우, 상기 축소 특징맵을 구성하는 커널 텐서(kernel tensor) 및 인자 행렬들(factor matrices)을 이용하여 상기 복원 특징맵을 생성하는 것을 특징으로 하는, 복호화 방법.

청구항 19

제14항에 있어서,

상기 복원 특징맵을 생성하는 단계는,

상기 축소 특징맵이 CP(Canonical Polyadic) 분해를 이용하여 축소된 경우, 상기 축소 특징맵을 구성하는 P(여기서, P는 자연수) 개의 랭크(rank) 1 텐서들을 이용하여 상기 복원 특징맵을 생성하는 것을 특징으로 하는, 복호화 방법.

발명의 설명

기술 분야

[0001] 본 개시는 특징맵 축소를 이용하는 머신 비전 데이터 코딩 장치 및 방법에 관한 것이다.

배경 기술

[0002] 이하에 기술되는 내용은 단순히 본 실시예와 관련되는 배경 정보만을 제공할 뿐 종래기술을 구성하는 것이 아니다.

[0003] 딥러닝(deep learning) 기술의 및 컴퓨팅 파워의 발전과 함께 머신 비전 어플리케이션이 폭발적으로 증가함에 따라 비디오 트래픽(video traffic)의 대부분이 기계에 의해 이용될 것이라는 인식이 확산되고 있다. 머신 투 머신(machine-to-machine) 어플리케이션은 향후 인터넷 비디오 트래픽의 가장 큰 사용량을 차지할 것으로 예상된다. 따라서, 기계가 사용하는 비디오 데이터의 정보를 최적화하는 방안은 비디오 처리 기술의 혁신 및 신규 솔루션의 상용화 측면에서 핵심 요소가 될 수 있다.

[0004] 기존의 비디오 코딩 방식은 특정 비트 전송률 제약 조건 하에서 최고의 영상 또는 영상 품질을 목표로 하므로, 인간의 시각에 최적화되어 있다. 반면, 머신 비전을 위한 코딩의 경우, 복원 영상/비디오가 반드시 높은 시각적 성능을 가질 필요는 없다. 지연 시간(latency)과 규모(scale)에 대해 엄격한 제한사항이 있는 커넥티드 차량, IoT(Internet of Things) 장치, 초대형 비디오 감시 네트워크, 스마트 시티, 품질 검사 등의 분야가 출현함에 따라 머신 비전을 위한 새로운 패러다임이 등장하여 머신 비전을 목표로 하는 신규 이미지/비디오 코딩 방법의 필요성이 대두되었다.

[0005] 이에 따라, 표준화 기구인 MPEG(Moving Picture Expert Group)에서는 머신 비전을 위한 표준화의 필요성이 논의되고, 머신 비전 데이터를 위한 압축 코딩, 및 휴먼-머신 하이브리드 시야를 위한 압축 코딩을 제공하는 차세대 비디오 코딩 장치로서 VCM(Video Coding for Machines)이 제안되었다.

[0006] VCM 코딩 장치의 구조에는 다양한 변형이 있을 수 있으나, 기본적인 VCM 코딩 장치의 구조는 도 1에 예시된 바와 같다. 센서의 출력인 비디오가 입력되었을 때, VCM 부호화기는 머신 비전을 위한 정보로서 특징을 추출(feature extraction)하고, 이를 필요에 맞게 변형(feature conversion)한 후, 부호화(feature encoding)한다. 또한, VCM 부호화기는 입력 영상 또는 비디오를 부호화할 때 부호화된 특징을 참조할 수 있다. 최종적으로, VCM 부호화기는 머신 비전을 위한 특징과 입력 영상(또는, 잔차 영상)을 부호화하여 비트스트림을 생성한다. VCM 부호화기는 특징 및 비디오를 부호화하여 생성되는 각각의 비트스트림을 다중화(multiplexing)하여 함께 전송한다.

[0007] VCM 복호화기는 전송된 비트스트림으로부터 특징 비트스트림과 비디오 비트스트림으로 역다중화(demultiplexing)한 후, 특징과 비디오를 각각 복호화한다. 이때, VCM 복호화기는, 비디오의 복호화 시, 복원된 특징을 참조할 수 있다. 복원된 특징은 역변형(inverse conversion)된 후, 머신 비전과 인간의 시각에 동시에 사용될 수 있다.

[0008] 또한, 도 1의 예시에서, VCM 부호화기에서의 특징 추출, VCM 복호화기에서의 머신 비전용 태스크 수행 등에 딥러닝 모델을 적용하기 위해, 신경망 연결부(interface for Neural Network)가 이용될 수 있다.

[0009] 특히, 머신 비전을 위한 정보의 대표적인 유형은, 머신 태스크 특화된 딥러닝 모델에 의해 추출되는 특징맵(feature map)이다. 영상/비디오를 전송하는 대신 VCM 부호화기가 특징맵을 전송하면, VCM 복호화기를 포함하는 클라우드 서버에서의 연산 부하가 크게 감소될 수 있고, 개인 정보 보호에도 이점이 있다. 하지만, 특징맵은 일반적인 영상/비디오에 비해 다수의 채널이 적층된(stacked) 형태의 배열이므로, 근본적으로 그 크기가 크다. 또

한, 딥러닝 모델의 특정 계층의 출력 특징맵의 크기가 입력 영상/비디오보다 작을 수 있지만, 딥러닝 모델이 변경될 때마다 대응되는 계층이 달라진다는 문제도 있다.

[0010] 특징맵의 크기가 입력 영상/비디오보다 과도하게 큰 경우, 전술한 연산 부하 감소, 개인 정보 보호 등의 이점이 퇴색될 수 있다. 따라서, 추출된 특징맵의 특성을 반영하여 그 크기를 감소시키는 추가적인 방법이 고안될 필요가 있다.

발명의 내용

해결하려는 과제

[0011] 본 개시는, 머신 태스크 특화된 딥러닝 모델에 의해 추출된 특징맵의 크기를 감소시키기 위해, 특징맵의 공간과 채널 측면의 불필요성을 감소시키는 희소화(sparsification) 방법, 및 텐서 분해(tensor decomposition)에 기반하는 특징맵 분해 방법을 이용하는 VCM(Video Coding for Machines) 코딩 장치 및 방법을 제공하는 데 목적이 있다.

과제의 해결 수단

[0012] 본 개시의 실시예에 따르면, 머신 비전 부호화 장치가 수행하는, 특징맵을 부호화하는 부호화 방법에 있어서, 딥러닝(deep learning) 모델을 이용하여 입력 영상으로부터 상기 특징맵을 추출하는 단계, 여기서, 상기 특징맵은 상기 딥러닝 모델의 중간 계층으로부터 생성됨; 상기 특징맵의 크기를 축소하여 축소(reduced) 특징맵을 생성하는 단계; 상기 축소 특징맵의 데이터 타입을 변환하고, 상기 축소 특징맵을 재배열하여 변형(converted) 특징맵을 생성하는 단계; 및 비디오 부호화기를 이용하여 상기 변형 특징맵을 부호화함으로써 비트스트림을 생성하는 단계를 포함하는 것을 특징으로 하는, 부호화 방법을 제공한다.

[0013] 본 개시의 다른 실시예에 따르면, 딥러닝(deep learning) 모델을 이용하여 입력 영상으로부터 특징맵을 추출하는 특징 추출기, 여기서, 상기 특징맵은 상기 딥러닝 모델의 중간 계층으로부터 생성됨; 상기 특징맵의 크기를 축소하여 축소(reduced) 특징맵을 생성하는 특징 축소기(feature reducer); 상기 축소 특징맵의 데이터 타입을 변환하기 위해 상기 축소 특징맵을 양자화하는 사전 양자화기(pre-quantizer); 상기 양자화된 특징맵을 비디오 시퀀스의 형태로 재배열하여 변형(converted) 특징맵을 생성하는 리패커(repacker); 및 비디오 부호화기를 이용하여 상기 변형 특징맵을 부호화함으로써 비트스트림을 생성하는 특징 부호화기를 포함하는 것을 특징으로 하는, 머신 비전 부호화 장치를 제공한다.

[0014] 본 개시의 다른 실시예에 따르면, 머신 비전 복호화 장치가 수행하는 복호화 방법에 있어서, 비디오 복호화기를 이용하여 비트스트림으로부터 변형 특징맵을 복호화하는 단계; 상기 변형 특징맵을 재배열하고, 상기 재배열된 변형 특징맵의 데이터 타입을 변환하여 축소(reduced) 특징맵을 복원하는 단계; 및 상기 축소 특징맵의 크기를 확장하여 복원 특징맵을 생성하는 단계를 포함하되, 상기 복원 특징맵은 머신 비전 부호화 장치 내 딥러닝 모델의 중간 계층으로부터 생성된 특징맵에 대응하는 것을 특징으로 하는, 복호화 방법을 제공한다.

발명의 효과

[0015] 이상에서 설명한 바와 같이 본 실시예에 따르면, 머신 태스크 특화된 딥러닝 모델에 의해 추출된 특징맵에 대해, 특징맵의 공간과 채널 측면의 불필요성을 감소시키는 희소화 방법, 및 텐서 분해에 기반하는 특징맵 분해 방법을 이용하는 VCM 코딩 장치 및 방법을 제공함으로써, 특징맵의 크기를 감소시키는 것이 가능해지는 효과가 있다.

[0016] 또한 본 실시예에 따르면, 머신 태스크 특화된 딥러닝 모델에 의해 추출된 특징맵에 대해, 특징 변형의 한 단계로서 특징맵의 공간과 채널 측면의 불필요성을 감소시키는 희소화 방법, 및 텐서 분해에 기반하는 특징맵 분해 방법을 이용하는 VCM 코딩 장치 및 방법을 제공함으로써, 특징 변형 전후에 존재하는 특징 추출 단계 및 특징 부호화 단계와는 독립적으로 특징맵 축소를 사용하는 것이 가능해지는 효과가 있다.

도면의 간단한 설명

[0017] 도 1은 VCM 코딩 장치를 개념적으로 나타내는 블록도이다.

도 2는 본 개시에 일 실시예에 따른 딥러닝 모델의 계층별 특징맵을 나타내는 예시도이다.

도 3은 프런트 단과 클라우드 서버의 동작을 개념적으로 나타내는 예시도이다.

도 4는 본 개시의 일 실시예에 따른, 중간 특징 압축을 수행하는 VCM 코딩 장치를 개념적으로 나타내는 블록도이다.

도 5는 본 개시의 일 실시예에 따른, 중간 특징맵 축소를 이용하는 VCM 부호화 장치를 개념적으로 나타내는 블록도이다.

도 6은 본 개시의 일 실시예에 따른, 중간 특징맵 복원을 이용하는 VCM 복호화 장치를 개념적으로 나타내는 블록도이다.

도 7은 본 개시의 일 실시예에 따른, 특징맵 회소화 기반의 특징맵 축소 방법을 나타내는 예시도이다.

도 8은 본 개시의 다른 실시예에 따른, 특징맵 회소화 기반의 특징맵 축소 방법을 나타내는 예시도이다.

도 9는 본 개시의 일 실시예에 따른, 터커 분해를 이용하는 특징맵 축소 방법을 나타내는 예시도이다.

도 10은 본 개시의 일 실시예에 따른, CP 분해를 이용하는 특징맵 축소 방법을 나타내는 예시도이다.

도 11는 본 개시의 일 실시예에 따른, 중간 특징맵 축소를 이용하는 VCM 부호화 방법을 나타내는 순서도이다.

도 12는 본 개시의 일 실시예에 따른, 중간 특징맵 복원을 이용하는 VCM 복호화 방법을 나타내는 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0018] 이하, 본 개시의 일부 실시예들을 예시적인 도면을 참조하여 상세하게 설명한다. 각 도면의 구성요소들에 참조 부호를 부가함에 있어서, 동일한 구성요소들에 대해서는 비록 다른 도면상에 표시되더라도 가능한 한 동일한 부호를 가지도록 하고 있음에 유의해야 한다. 또한, 본 실시예들을 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 실시예들의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략한다.

[0019] 본 실시예는 특징맵 축소를 이용하는 머신 비전 데이터 코딩 장치 및 방법에 관한 내용을 개시한다. 보다 자세하게는, 머신 태스크 특화된 딥러닝 모델에 의해 추출된 특징맵의 크기를 감소시키기 위해, 특징맵의 공간과 채널 측면의 불필요성을 감소시키는 회소화(sparsification) 방법, 및 텐서 분해(tensor decomposition)에 기반하는 특징맵 분해 방법을 이용하는 VCM(Video Coding for Machines) 코딩 장치 및 방법을 제공한다.

[0020] 여기서, VCM 코딩 장치는 VCM 부호화 장치와 VCM 복호화 장치를 포함한다.

[0021] 한편, 인간의 시각에 최적화되도록 비디오 신호를 부호화 및 복호화하는 기존의 코덱을 비디오 부호화기 및 비디오 복호화기로 나타낸다.

[0022] I. 중간 특징맵 코딩(intermediate deep feature coding)

[0023] 계층적 구조를 갖는 딥러닝 모델의 각 계층(layer)에서 출력되는 다차원의 배열은 통상적으로 특징(feature), 특징맵(feature map), 딥 특징(deep feature), 또는 딥러닝 특징맵(deep learning feature map)으로 명칭된다. 도 2의 예시와 같이, 각 계층별 특징맵은 공간적 상관성이 있는 2차원의 배열이 적층된(stacked) 형태를 가진다. 이때, 2차원 배열의 높이 H와 너비 W는 해당 특징맵의 높이와 너비로 명칭되고, 스택된 2차원 배열의 개수는 해당 특징맵의 채널의 개수 C로 명칭된다. 영상처리에 적합한 것으로 알려진 CNN(Convolutional Neural Network) 구조에서는, 일반적으로 계층이 깊어질수록 풀링(pooling)에 의해 특징맵의 높이와 너비의 크기가 작아지고 채널의 크기는 증가한다. 또한, 직관적으로, 단채널의 2차원 특징맵은 하나의 프레임으로 간주될 수 있고, 적층된 특징맵들은 연속된 하나의 비디오 시퀀스로 간주될 수 있다.

[0024] 네트워크 인프라(network infra)의 발달로 최근 클라우드 기반 애플리케이션들(cloud-based applications)이 등장하고 있다. 특히, 프런트 단(front-end) 장치는 사용자 또는 물리적 세계로부터 정보를 획득한 후, 획득된 정보를 클라우드 서버(즉, 데이터 센터)로 전송하고, 클라우드 서버는 획득된 정보를 이용하여 추가적인 프로세스와 분석을 수행한다. 보안 감시 카메라, 웨어러블 기기(wearable device) 등과 같은 실제 환경에 배치된 프런트 단 장치는, 도 3에 예시된 바와 같이, 시각적 분석을 위해 클라우드 서버에 전송되는 방대한 데이터를 획득할 수 있다. 예컨대, i) 자율 주행 차량에서의 보행자 감지, 사람/차량 식별 및 번호판 인식, ii) 휴대용 장치(예를 들어, 스마트폰, 스마트 안경 등)의 애플리케이션에서의 얼굴 인식, 랜드마크 탐색 및 물체 감지 iii) 감시 카메라에서의 감시 영상 획득 및 사람/차량 인식 등에 있어서, 많은 딥러닝 기반의 비전 모델들이 이러한 클라우드 기반 패러다임에 적용될 수 있다.

- [0025] 한편, 이러한 클라우드 기반 패러다임에서, 프런트 단과 클라우드 서버 간의 비디오의 압축과 전송은 전통적인 '압축 후 분석(compress-then-analyse)' 방식에 의존할 수 있다. 프런트 단 장치는 시각적 데이터를 캡처하고 압축한 후, 코드화된 비트스트림을 클라우드 서버로 전송한다. 클라우드 서버는 비트스트림을 디코딩한 후 형상 추출 및 시각적 분석을 수행한다. 하지만, 비디오 감시 및 IoT와 같은 시나리오에서는, 방대한 양의 프런트 단 장치가 수백 만 개의 비트스트림들을 동시에 생성할 수 있다. 따라서, 프런트 단의 측면에서, 비디오 신호 레벨에서의 시각적 압축은 전송 부담이 크고 일반적으로 감당할 수 없는 규모일 수 있다. 또한, 클라우드 서버의 측면에서, 특징 추출을 위해 동시에 실행되는 수많은 딥러닝 모델들의 연산 부하가 상당한 어려움으로 작용될 수 있다.
- [0026] 이에 반해, 프런트 단에서 추출한 특징을 압축한 후 클라우드 서버로 전송하는 '분석 후 압축 (analyze-then-compress)' 방식에서는, 시각적 신호 대신 추출된 특징이 태스크 분석에 사용되므로, 전송한 '압축 후 분석'의 실현 가능한 대안이 될 수 있다. 딥러닝 기반 특징들에 있어서, 딥러닝 모델의 최상위 계층(즉, 가장 깊은 계층)의 특징들이 하위 계층(즉, 얕은 계층)보다 간결(compact)하고 분석에 바로 사용될 수 있기 때문에, 보통 최상위 계층(즉, 가장 깊은 계층)의 특징들이 클라우드 단에 전송될 수 있다. 이러한 시나리오에서는, 특징 간 비교와 같은 경량 작업들이 클라우드 서버에서 수행되고, 특징 추출과 같은 작업들은 프런트 단에 분산된다. 또한, 이러한 특징 전송은 개인 정보 보호에도 유리하다. 개인 정보가 용이하게 노출될 수 있는 시각적 신호의 전송과 비교하여, 특징 전송은 가시적 정보 공개를 회피할 수 있다.
- [0027] 한편, 딥러닝 모델이 특정 태스크에 대해 설계 및 트레이닝되고, 최상위 계층의 특징들은 매우 추상적이고 태스크 특화되기 때문에, 최상위 계층의 특징들을 압축하는 것은 일반화되기 어려울 수 있다. 즉, 표준화에 따라 압축된 특징들이 다른 시나리오에 적용되기 어려울 수 있는데, 이러한 문제는 딥러닝 모델의 특징 압축에 대한 표준을 만드는 데 있어서 장애물이 될 수 있다.
- [0028] 전송한 바와 같은 사안들을 고려하여, 특징의 가용성(availability)을 유지하면서도, 클라우드 서버에서의 연산 부하를 줄이기 위한 방안으로, 중간 특징 압축(intermediate deep feature compression) 방식이 존재한다. 이 방식은 전송한 바와 같은 극단적인 '분석 후 압축'과 '압축 후 분석' 간의 절충으로 간주되며, 연산 부하, 통신 비용 및 일반화 능력 사이의 균형을 잘 제공할 수 있다.
- [0029] 딥러닝 모델은 일반적으로 계층 구조를 가지므로, 딥러닝 모델은 단일 특징 추출기가 아닌 적층된 특징 추출기들의 조합으로 간주될 수 있다. 상위 계층의 특징은 큰 수용 필드(receptive field)를 바탕으로 추상적이고 태스크 특화적인 반면, 하위 계층의 특징은 작은 수용 필드를 가지고 2D 특징맵에 많은 위치 정보를 부호화하므로 광범위한 분석에 적합할 수 있다. 따라서, 최종 계층 대신 중간 계층의 특징을 압축 및 전송하는 경우, 프런트 단에서 수행되는 태스크에 따라, 클라우드 서버는 적절한 특징을 유연하게 요청할 수 있다.
- [0030] 중간 특징 코딩은 VCM 표준의 파이프라인에서 매우 중요한데, 도 1에 예시된 VCM 코딩 장치의 구조에서 특징 변형, 특징 부호화, 및 특징 복호화 순서로 적용될 수 있다. 이하, 도 4의 도시를 이용하여, 중간 특징 코딩을 기술한다. 한편, 도 4의 도시에서는, 입력 영상의 비디오 부호화 및 비디오 복호화를 수행하는 구성요소들은, 본 실시예와 연관성이 크지 않으므로 그 도시를 생략한다.
- [0031] 이하의 설명에서, 특징맵은 딥러닝 모델의 중간 계층의 출력인 중간 특징맵을 나타낸다.
- [0032] 도 4는 본 개시의 일 실시예에 따른, 중간 특징 압축을 수행하는 VCM 코딩 장치를 개념적으로 나타내는 블록도이다.
- [0033] VCM 부호화 장치는 특징 추출기(feature extractor, 410), 특징 변형기(feature converter, 420) 및 특징 부호화기(430)의 전부 또는 일부를 포함한다. 또한, VCM 복호화 장치는 특징 복호화기(450) 및 특징 역변형기(feature inverse converter, 460)의 전부 또는 일부를 포함한다.
- [0034] 우선, 특징 부호화기(430)와 특징 복호화기(450)는 기존 비디오 코딩 장치를 이용하여 구현될 수 있다.
- [0035] VCM 부호화 장치 내 특징 추출기(410)는 딥러닝 모델을 기반으로, 입력 영상으로부터 머신 비전을 위한 태스크의 특징맵을 추출한다. 추가적으로, 특징 추출기(410) 내 딥러닝 모델은 머신 비전 태스크를 수행하여, 예컨대, 머신 비전 태스크의 분석 결과를 생성할 수 있다.
- [0036] 특징 추출기(410) 내 딥러닝 모델로는, 머신 비전 태스크의 분석에 적합한 특징맵을 생성할 수 있는 어느 신경망 모델이든 이용될 수 있다. 예컨대, 딥러닝 모델은 다수의 인코딩 계층들(layers)로 구현된 CNN 기반의 모델일 수 있다. 이때, 하나의 인코딩 계층은 컨볼루션 레이어(convolutional layer)와 풀링 레이어(pooling

layer)를 포함한다.

- [0037] 특징 변형기(420)는, 특징맵에 대해 데이터 타입 변환, 특징맵 채널 재정렬 등을 위해 사전 양자화기(pre-quantizer, 422) 및 리팩커(repacker, 424)를 포함한다.
- [0038] 딥러닝 모델에 의해 추출된 특징맵은 기존의 비디오 코딩 장치의 입력과 호환되지 않는 숫자 유형을 가질 수 있다. 따라서, 특징 부호화기(430)와의 입력 호환성을 맞추기 위해, 사전 양자화기(422)는 사전 양자화(pre-quantization)를 먼저 수행한다. 예를 들어, 일반적인 딥러닝 모델의 특징맵들은 float32 타입인 반면, 기존의 비디오 코딩 장치는 8 또는 그 이상의 비트의 정수형 입력을 이용한다. 또한, 중간 특징맵의 분포 분석에 따라 다른 사전 양자화 방법이 사용될 수 있다.
- [0039] 리팩커(424)는 사전 양자화된 N (여기서, N 은 자연수) 개의 특징맵 $R^{N \times H \times W \times C}$ 을 비디오 코딩 장치의 입력 형식에 맞도록 비디오 시퀀스와 같은 포맷인 $R^{H' \times W' \times NC}$ 로 리팩킹한다. 여기서, H 와 W 는 2D 특징맵의 높이와 너비, C 는 채널의 개수이다. $H' \times W'$ 은, 비디오 코딩 장치의 입력 프레임 크기에 적합하도록 원본 특징맵의 크기 $H \times W$ 가 패딩에 의해 확장된 것이다. R 은 실수를 나타낸다. 또한, 입력 영상 중 하나의 프레임으로부터 딥러닝 모델이 추출하는 특징맵이 $H \times W \times C$ 이므로, $N \times H \times W \times C$ 는 N 개의 프레임들로부터 추출된 특징맵들에 해당된다.
- [0040] 이하의 설명에서, 특징맵은 $F \in R^{H \times W \times C}$ 로 나타내고, 특징맵을 구성하는 2D 특징맵을 $f \in R^{H \times W}$ 로 나타낸다.
- [0041] 다른 실시예로서, 리팩커(424)는 사전 양자화된 N (여기서, N 은 자연수) 개의 특징맵 $R^{N \times H \times W \times C}$ 을 하나의 프레임으로 리팩킹할 수 있다. 예컨대, 전술한 바와 같이 특징 부호화기(430)가 기존의 비디오 코딩 장치인 경우, 인터 예측 모드 대신 인트라 예측 모드만을 사용하기 위해, 채널별 특징맵들이 하나의 수퍼 프레임으로 리팩킹될 수 있다. 이러한 리팩킹된 수퍼 프레임에 대해, 비디오 코딩 장치의 인트라 예측만을 사용하여 더 좋은 압축 효율이 달성될 수 있다. 이하, 이러한 리팩킹 과정을 프레임 패킹으로 명칭한다.
- [0042] 특징 부호화기(430)는 변형된 특징맵을 부호화하여 해당되는 비트스트림을 생성한다.
- [0043] VCM 복호화 장치 내 특징 복호화기(450)는 비트스트림으로부터 변형된 특징맵을 복호화한다.
- [0044] 특징 역변형기(460)는, VCM 부호화 장치에 의한 특징맵 변형의 역과정인 특징맵 역변형을 비트스트림에 적용하기 위해, 디리팩커(derepacker, 462) 및 사후 역양자화기(post-dequantizer, 464)를 포함한다.
- [0045] 디리팩커(462)는 비디오 시퀀스와 같은 포맷으로 리팩킹된 특징맵을 원본 특징맵 크기로 변형한다.
- [0046] 사후 역양자화기(464)는 정수값을 갖는 특징맵을 부동소수점 타입으로 역양자화한다. 이러한 복원된 특징맵은, 전술한 바와 같이 머신 비전 태스크 분석을 위해 이용될 수 있다.
- [0047] 시각적 신호 대신 중간 특징맵을 전송할 때 압축된 중간 특징맵이 비디오/이미지 스트림보다 더 클 수 있다는 우려가 존재한다. 하지만, 딥러닝 모델 상위 계층의 특징들은 일반적으로 하위 계층의 특징보다 더 작은 압축 크기를 가지며, 특히, 최상위 계층으로 근접할수록 해당되는 특징들이 원본 시각 정보보다 좋은 압축률을 갖는 것이 알려져 있다. 즉, 어떤 종류의 특징은 영상/비디오 비트스트림보다 적은 대역폭을 사용할 수 있다. 따라서, 이런 경우, 중간 특징맵을 전송하는 것이 전송 대역폭 효율화 측면에서 의미가 있다.
- [0048] 한편, 중간 특징맵이 영상/비디오 비트 스트림보다 작은 크기로 압축되지 못한다고 하더라도, 특징맵을 전송하는 것은 여전히 의미가 있다. 많은 적용 사례에서 연산 자원이 대역폭보다 더 많이 소비되므로, 중간 특징맵 전송을 이용하여 클라우드 서버의 연산 부하를 프린트 단 장비로 분산시키는 것이 더욱 중요할 수 있다.
- [0049] **II. 중간 특징맵 축소(intermediate deep feature reduction)**
- [0050] 이하, 도 5 및 도 6의 도시를 이용하여, 특징 변형 단계에서 중간 특징맵을 축소하는 방법, 및 특징 역변형 단계에서 중간 특징맵을 복원하는 방법을 기술한다.
- [0051] 도 5는 본 개시의 일 실시예에 따른, 중간 특징맵 축소를 이용하는 VCM 부호화 장치를 개념적으로 나타내는 블록도이다.
- [0052] 도 5의 도시에 신규로 추가된 특징 축소기(feature reducer, 512)를 제외한 나머지 구성요소들은 도 4에 도시된 VCM 부호화 장치의 구성요소들과 동일하다. 전술한 바와 같이, 특징 변형 단계는 특징 부호화를 위해 기존 비디오 코딩 장치를 이용하는 것을 전제하므로, 특징맵을 사전 양자화하기 전에, 특징 축소기(512)는 여러 가지 특징맵 축소 방법을 추가적으로 수행하여, 원본 특징맵의 크기를 감축시킴으로써, 특징맵의 부호화 효율을 향상시

킬 수 있다.

- [0053] 한편, 특징맵 축소를 위해 특징 축소기(512)는, 특징맵 희소화(feature map sparsification) 기반의 특징맵 축소, 및 텐서 분해(tensor decomposition) 기반의 특징맵 축소의 두 가지 방법을 사용할 수 있다. 이러한 두 가지 방법을 개별적으로 사용되거나 동시에 사용함으로써, 특징 축소기(512)는 특징맵 축소를 극대화할 수 있다. 또한, 두 가지 방법이 함께 사용되는 경우, 사용 순서에는 제약이 없다. 본 실시예에 따른 특징맵 축소를 위한 방법은, 전술한 바와 같은 특징맵의 프레임 패킹과 같은 별개의 전처리 방법과도 조합되어 함께 사용될 수 있다.
- [0054] 도 6은 본 개시의 일 실시예에 따른, 중간 특징맵 복원을 이용하는 VCM 복호화 장치를 개념적으로 나타내는 블록도이다.
- [0055] 도 6의 도시에 신규로 추가된 특징 복원기(feature reconstructor, 612)를 제외한 나머지 구성요소들은 도 4에 도시된 VCM 복호화 장치의 구성요소들과 동일하다. 전술한 바와 같이, 특징 역변형 단계는 특징 복호화를 위해 기존 비디오 코딩 장치를 이용하는 것을 전제하므로, 특징 축소기(512)에 대응하는 특징 복원기(612)는 사후 역양자화된 특징맵, 즉 축소된 특징맵으로부터 원본 특징맵을 복원한다. 구체적인 특징맵 복원 방법은 VCM 부호화 장치에서 사용한 특징맵 축소 방법에 의존할 수 있다.
- [0056] 이하, 특징 축소기(512) 및 특징 복원기(612)가 수행하는 특징맵 축소화 방법들을 기술한다. 도 5 및 도 6에 도시에서는, 특징 부호화/복호화를 위해 기존 비디오 코딩 장치를 이용하는 것을 전제하나, 이하에 기술되는 특징맵 축소화 방법들은 반드시 이에 한정되지 않는다. 즉, 본 실시예에 따른 특징맵 축소화 방법들은, 특징 추출을 위한 딥러닝 모델, 또는 코딩 방법과 관계없이 추출된 특징맵의 크기를 축소시키기 위해 활용될 수 있다.
- [0057] 먼저, 특징맵 축소화 방법들 중 특징맵 희소화 기반의 특징맵 축소는 특징맵의 공간 또는 채널 측면에 존재하는 희소성(sparsity)을 활용한다.
- [0058] 딥러닝 모델이 중간 특징맵을 생성 시, 입력 영상/프레임의 공간적 특성 또는 트레이닝되는 필터의 특성에 따라 일부 영역 및 일부 채널이 크게 활성화(activation)될 수 있다. 이에 반해, 작게 활성화된 영역은 다음 계층에 입력되더라도 연속적으로 작은 활성화를 유도하고, 최종적으로 딥러닝 모델의 마지막 계층에서의 판단에 작은 영향을 준다. 따라서, 이렇게 작은 활성화를 출력하는 영역의 값을 0으로 희소화시키더라도, 딥러닝 모델의 판단 성능이 크게 저하되지 않은 채로 특징맵의 크기 자체는 축소될 수 있다. 즉, 일정 기준 값보다 작게 활성화된 특징맵 내 영역을 희소화시키는 특징맵 희소화를 기반으로, 특징맵의 부호화 효율이 향상될 수 있다.
- [0059] 이하, 도 7의 예시를 이용하여, 공간 측면에 존재하는 희소성에 기초하여 특징 축소기(512)가 특징맵을 축소하는 방법을 기술한다.
- [0060] 도 7은 본 개시의 일 실시예에 따른, 특징맵 희소화 기반의 특징맵 축소 방법을 나타내는 예시도이다.
- [0061] 특징 축소기(512)는 중간 특징맵 $F \in \mathbb{R}^{H \times W \times C}$ 을 구성하는 단일 2D 특징맵 $f \in \mathbb{R}^{H \times W}$ 내에서 활성화 정도에 따라 작은 활성화를 출력하는 영역을 희소화한다. 예컨대, 특징맵의 특정 채널의 2D 특징맵 f 의 영역(예컨대, 픽셀)을 특정 값에 따라 숫자로 표현했을 때, 특징 축소기(512)는 특정 임계치보다 작은 영역의 특정 값을 완전히 0으로 설정한다. 도 7의 예시에서, 임계치를 1.5로 설정할 경우, 1로 표시된 영역들이 0으로 설정될 수 있다. 이렇게 적절한 임계치를 설정하여, 이 임계치보다 작은 영역의 특정 값을 0으로 설정하더라도, 딥러닝 모델의 성능은 큰 영향을 받지 않는다. 특징 축소기(512)는, 특징맵의 모든 채널에 해당하는 C 개의 2D 맵들 $\{f_1, f_2, \dots, f_C\}$, 또는 일부 2D 맵들 $\{f_1, f_2, \dots, f_{C'}\}$ ($C' < C$)에 대해 이러한 공간 측면의 특징맵 희소화를 수행할 수 있다.
- [0062] 이하, 도 8의 예시를 이용하여, 채널 측면에 존재하는 희소성에 기초하여 특징 축소기(512)가 특징맵을 축소하는 방법을 기술한다.
- [0063] 도 8은 본 개시의 다른 실시예에 따른, 특징맵 희소화 기반의 특징맵 축소 방법을 나타내는 예시도이다.
- [0064] 특징 축소기(512)는, 일정 채널 간격 S 만큼 인접한 두 채널의 2D 특징맵이 비슷한 경우, 즉, 두 특징맵 간 거리가 특정 임계치보다 작은 경우, 두 2D 특징맵 간 과잉(redundancy)을 감소시키기 위해, 하나의 2D 특징맵을 0으로 설정하거나 온전히 채널 하나를 삭제함으로써 원본 특징맵의 채널 수를 감소시켜 전체 특징맵의 크기를 감소시킬 수 있다.

- [0065] 원본 특징맵에 대해, 특징 축소기(512)는 일정 채널 간격 S 를 갖는 두 개의 2D 특징맵 간의 거리 $\|f_{i+S} - f_i\|$ ($1 \leq i \leq C-S$)를 계산한다. 이때, 거리는 L1, L2 등의 놈(norm)이 사용될 수 있다. 예컨대, 도 8의 예시에서, 채널 간격 S 는 1로 설정되어 있다. 특징 축소기(512)는 특정 임계치 d 보다 작은 거리를 갖는 두 2D 특징맵을 매우 유사한 특징맵들로 간주한다. 특징 축소기(512)는, 유사한 2D 특징맵을 갖는 채널의 짝들(pairs)을 모두 선택한 후, 각 짝별로 두 2D 특징맵 중 뒷 채널(또는 앞 채널)의 2D 특징맵의 모든 값을 0으로 최소화할 수 있다. 또는, 특징 축소기(512)는, 도 8에 예시된 바와 같이, 각 짝별로 하나의 채널을 완전히 삭제할 수 있다. 특징 축소기(512)는 유사하지 않은 채널들의 2D 특징맵만을 사용하여 전체 채널 수가 축소된 중간 특징맵 F' 을 생성할 수 있다.
- [0066] VCM 부호화 장치는 최소화된 채널들이 존재하는 경우, 채널 간격 S 값, 및 각 최소화된 채널의 2D 특징맵을 부호화한 후, VCM 복호화 장치로 전송한다. 각 최소화된 채널의 2D 특징맵의 모든 값이 0이므로, VCM 부호화 장치는 각 최소화된 채널의 2D 특징맵을 용이하게 부호화할 수 있으므로, 부호화 효율을 크게 향상시킬 수 있다. 또는, 삭제된 채널들이 존재하는 경우, VCM 부호화 장치는 채널 간격 S 값, 및 각 삭제된 채널의 인덱스를 부호화한 후, VCM 복호화 장치로 전송할 수 있다. 삭제된 채널에 대해 $H \times W$ 크기의 2D 특징맵을 부호화하는 대신 VCM 부호화 장치는 인덱스에 해당하는 정수 값만을 부호화함으로써, 부호화 효율을 크게 향상시킬 수 있다.
- [0067] 한편, 최소화되거나 삭제된 채널에 대해, VCM 복호화 장치 내 특징 복원기(612)는 복원된 축소 특징맵, 복원된 채널 스트라이드 S 값, 또는 삭제된 채널 인덱스 정보를 참조할 수 있다. 모든 값이 0으로 복원되는 최소화 채널에 대해, VCM 복호화 장치는 최소화된 채널 위치에 S 간격 앞(또는 뒤) 채널의 2D 특징맵을 복사함으로써, 축소 전 원본 특징맵을 복원한다. 또는, 삭제된 채널에 대해, VCM 복호화 장치는 삭제된 채널 인덱스 정보를 참조하여 삭제된 채널 위치에 S 간격 앞(또는 뒤) 채널의 2D 특징맵을 복사함으로써, 축소 전 원본 특징맵을 복원한다.
- [0068] 이렇게 복원된 특징맵은 원본 특징맵과 같은 크기를 갖는 배열이며, 이후, VCM 복호화 장치는 복원된 특징맵을 머신 비전 태스크의 분석에 이용할 수 있다.
- [0069] 다른 실시예로서, VCM 부호화 장치는 각 채널의 가중치를 표현하는 $1 \times C$ 의 벡터에 포함된 요소값 c_i 를 해당되는 특징맵 f_i 에 곱한 후, 특징맵을 압축하여 전송할 수 있다. VCM 복호화 장치는, 특징맵을 복호 시, 수신된 특징맵을 c_i 로 나누어서 복원 특징맵을 생성할 수 있다. 이때, c_i 는 딥러닝 기반의 주의 모델(attention model), 또는 딥러닝 기반의 SE(Squeeze and Excitation) 모델의 SE 블록으로부터 획득될 수 있다.
- [0070] 이하, 텐서 분해를 이용하는 특징맵 축소화 방법의 예시를 설명한다.
- [0071] 전술한 바와 같은 다차원 배열 형태의 특징맵은, 텐서 분해를 이용하여 저차원의 간단한 텐서들 여러 개로 분해되어 표현될 수 있다. 이때, 분해된 텐서들의 크기 총합은 원본 특징맵의 크기보다 감소될 수 있다.
- [0072] 한편, 특징맵 축소는 특정 텐서 분해 방법에 한정되지 않으며, 아래에서 설명한 예시들 외에 행렬 분해(matrix decomposition, 예컨대 SVD(Singular Value Decomposition), CUR 분해(factorization), QR 분해(factorization) 등) 및 HOSVD(Higher Order SVD) 등과 같은 다양한 텐서 분해 방법들이 적용될 수 있다. 텐서 분해 방법의 적용 후, VCM 부호화 장치는 원본 특징맵 대신 분해된 저차원의 텐서들을 특징 부호화에 활용할 수 있다.
- [0073] 텐서 분해의 일 예로서, 도 9의 예시를 이용하여, 터커 분해(Tucker decomposition)에 기초하여 특징 축소기(512)가 특징맵을 축소하는 방법을 기술한다.
- [0074] 도 9는 본 개시의 일 실시예에 따른, 터커 분해를 이용하는 특징맵 축소 방법을 나타내는 예시도이다.
- [0075] 특징 축소기(512)는, 도 9에 예시된 바와 같이, 터커 분해를 이용하여, 단일 영상에 대한 중간 특징맵 $F \in R^{n_1 \times n_2 \times n_3}$ 를 하나의 커널 텐서(kernel tensor) $K \in R^{r_1 \times r_2 \times r_3}$, 및 3 개의 인자 행렬들(factor matrices) $F^{(i)} \in R^{r_i \times n_i}$ ($i \in \{1, 2, 3\}$)로 분해할 수 있다. 이때, r_i 는 i -rank로 표현되고, 특징맵을 $F^* \in R^{n_i \times n_1 \dots n_{i-1} n_{i+1} \dots n_3}$ 형태의 행렬로 생성하는 경우, 이 행렬의 랭크를 나타낸다. 하나의 커널 텐서, 및 3 개의 인자 행렬들을 이용하여, 특징맵 F 는 수학적 1에 나타낸 바와 같이, 텐서들의 연산으로 근사될 수 있다.

수학식 1

$$F \approx K \times_1 F^{(1)} \times_2 F^{(2)} \times_3 F^{(3)}$$

여기서, 연산자 $\times_1$, $\times_2$ 및 $\times_3$ 는 수학식 2와 같이 표현될 수 있다.

수학식 2

$$(K \times_1 F^{(1)})(n_1, r_2, r_3) = \sum_{i_1=1}^{r_1} K(i_1, r_2, r_3) F^{(1)}(i_1, n_1)$$

$$(K \times_2 F^{(2)})(r_1, n_2, r_3) = \sum_{i_2=1}^{r_2} K(r_1, i_2, r_3) F^{(2)}(i_2, n_2)$$

$$(K \times_3 F^{(3)})(r_1, r_2, n_3) = \sum_{i_3=1}^{r_3} K(r_1, r_2, i_3) F^{(3)}(i_3, n_3)$$

터커 분해에 따라, 원본 특징맵의 데이터 크기 $n_1 n_2 n_3$ 는 ' $r_1 r_2 r_3 + r_1 n_1 + r_2 n_2 + r_3 n_3$ '로 감축된다.

VCM 부호화 장치는, 비디오 부호화기를 이용하여 분해된 텐서들 K , $F^{(1)}$, $F^{(2)}$ 및 $F^{(3)}$ 각각을 부호화한 후, 생성된 비트스트림들을 다중화(multiplexing)하여 VCM 복호화 장치로 전송할 수 있다.

한편, VCM 복호화 장치는, 비디오 복호화기를 이용하여 분해된 텐서들의 비트스트림을 복호화한 후, 도 4에 예시된 바와 같은 디리팩터(462) 및 사후 역양자화기(464)에 복호화된 텐서들을 적용한다. VCM 복호화 장치는 역양자화된 분해 텐서들을 이용하여, 수학식 1에 따라 원본 특징맵을 복원할 수 있다.

텐서 분해의 다른 실시예로서, 도 10의 예시를 이용하여, CP(Canonical Polyadic) 분해에 기초하여 특징 축소기(512)가 특징맵을 축소하는 방법을 기술한다.

도 10은 본 개시의 일 실시예에 따른, CP 분해를 이용하는 특징맵 축소 방법을 나타내는 예시도이다.

특징 축소기(512)는, 도 10에 예시된 바와 같이, CP 분해를 이용하여, 중간 특징맵 $F \in R^{n_1 \times n_2 \times n_3}$ 을 랭크 1 텐서들의 선형 조합으로 분해할 수 있다. 즉, 특징맵 F 는 수학식 3에 나타낸 바와 같이 P (여기서, P 는 자연수) 개의 랭크 1 텐서들에 기초하여 근사될 수 있다.

수학식 3

$$F \approx \sum_{p=1}^P a^p \circ b^p \circ c^p$$

여기서, 연산자 $\circ$ 는 외적을 나타낸다.

VCM 부호화 장치는, 수학식 4에 나타낸 바와 같이, 크기가 같은 랭크 1 텐서들끼리 결합(concatenation)하여 텐서들 A , B 및 C 를 생성한 후, 이들을 부호화한다.

수학식 4

$$A = [a_1, a_2, \dots, a_p] \in R^{n_1 \times P}$$

$$B = [b_1, b_2, \dots, b_p] \in R^{n_2 \times P}$$

$$C = [c_1, c_2, \dots, c_p] \in R^{n_3 \times P}$$

[0088]

[0089]

CP 분해에 따라, 원본 특징맵의 데이터 크기 $n_1 n_2 n_3$ 는 $(n_1 + n_2 + n_3)P$ 로 감축된다.

[0090]

다른 실시예로서, VCM 부호화 장치는 분해된 P 개의 랭크 1 텐서를 부호화할 수 있다.

[0091]

VCM 부호화 장치는, 분해된 텐서들을 부호화한 비트스트림을 VCM 복호화 장치로 전송한다.

[0092]

한편, VCM 복호화 장치는, 비디오 복호화기를 이용하여 분해된 텐서들의 비트스트림을 복호화한 후, 도 4에 예시된 바와 같은 디리팩커(462) 및 사후 역양자화기(464)에 복호화된 텐서들을 적용한다. VCM 복호화 장치는 역양자화된 분해 텐서들을 이용하여, 수학식 3에 따라 원본 특징맵을 복원할 수 있다.

[0093]

다른 실시예로서, VCM 부호화 장치는 $P' (< P)$ 개의 랭크 1 텐서들만을 결합한 텐서들을 부호화하여, VCM 복호화 장치로 전송할 수 있다. VCM 복호화 장치는 복호화된 P' 개의 랭크 1 텐서들을 이용하여, 원본 특징맵을 복원할 수 있다.

[0094]

이하, 도 11 및 도 12의 도시를 이용하여, 중간 특징맵 축소를 이용하는 VCM 부호화 방법, 및 중간 특징맵 복원을 이용하는 VCM 복호화 방법을 기술한다.

[0095]

도 11는 본 개시의 일 실시예에 따른, 중간 특징맵 축소를 이용하는 VCM 부호화 방법을 나타내는 순서도이다.

[0096]

VCM 부호화 장치는 딥러닝 모델을 이용하여 입력 영상으로부터 특징맵을 추출한다(S1100). 여기서, 특징맵은 딥러닝 모델의 중간 계층으로부터 생성된 중간 특징맵을 나타낸다. 특징맵은 동일한 높이와 너비를 갖는 2D 특징맵을 채널의 개수만큼 포함한다.

[0097]

VCM 부호화 장치는 특징맵의 크기를 축소하여 축소(reduced) 특징맵을 생성한다(S1102).

[0098]

VCM 부호화 장치는 특징맵의 공간 또는 채널 측면에서의 특징맵 회소화를 기반으로 특징맵의 크기를 축소한다. 또는, VCM 부호화 장치는 텐서 분해를 기반으로 특징맵을 축소할 수 있다. 이러한 두 가지 방법을 개별적으로 사용되거나 동시에 사용함으로써, VCM 부호화 장치는 특징맵 축소를 극대화할 수 있다. 또한, 두 가지 방법이 함께 사용되는 경우, 사용 순서에는 제약이 없다.

[0099]

특징맵의 공간 측면에서의 회소성을 이용하기 위해, VCM 부호화 장치는 특징맵을 구성하는 2D 특징맵 내에서, 영역의 특징 값을 기설정된 임계치와 비교한다. 특징 값이 기설정된 임계치보다 작은 경우, 해당되는 영역의 특징 값은 0으로 설정될 수 있다.

[0100]

다른 실시예로서, 특징맵의 채널 측면에서의 회소성을 이용하기 위해, VCM 부호화 장치는 기설정된 채널 간격을 갖는 두 개의 2D 특징맵 간의 거리를 산정한다. 산정된 거리가 기설정된 임계치보다 작은 경우에 해당하는 2D 특징맵의 짝들(pairs)을 모두 선택한 후, VCM 부호화 장치는 선택된 짝들 각각에 대해, 하나의 2D 특징맵의 모든 값을 0으로 회소화하거나 하나의 2D 특징맵을 삭제한다.

[0101]

한편, 텐서 분해 방법 중 터커 분해를 이용하여, VCM 부호화 장치는 특징맵을 하나의 커널 텐서, 및 3 개의 인자 행렬들로 분해할 수 있다. 이때, 수학식 1에 나타난 바와 같이, 특징맵은 하나의 커널 텐서, 및 3 개의 인자 행렬들에 의해 근사될 수 있다.

[0102]

다른 실시예로서, 텐서 분해 방법 중 CP 분해를 이용하여, VCM 부호화 장치는 특징맵을 P 개의 랭크 1 텐서들로 분해할 수 있다. 이때, 수학식 3에 나타난 바와 같이, 특징맵은 P 개의 랭크 1 텐서들에 기초하여 근사될 수 있다.

[0103]

VCM 부호화 장치는, 축소 특징맵의 데이터 타입을 변환하기 위해 축소 특징맵을 양자화한다(S1104).

[0104]

VCM 부호화 장치는 양자화된 특징맵을 비디오 시퀀스의 형태로 재배열하여 변형(converted) 특징맵을 생성한다

(S1106).

- [0105] VCM 부호화 장치는 비디오 부호화기를 이용하여 변형 특징맵을 부호화함으로써 비트스트림을 생성한다(S1108).
- [0106] 한편, 특징맵의 채널 측면에서의 희소성을 이용하는 경우, VCM 부호화 장치는 다음과 같이 비트스트림을 생성한다. 희소화된 특징맵에 대해 VCM 부호화 장치는 2D 특징맵, 및 기설정된 채널 간격을 부호화한 후, VCM 복호화 장치로 전송할 수 있다. 또는, 삭제된 특징맵에 대해 VCM 부호화 장치는 삭제된 2D 특징맵의 인덱스, 및 기설정된 채널 간격을 부호화한 후, VCM 복호화 장치로 전송할 수 있다.
- [0107] 터커 분해를 이용하는 경우, VCM 부호화 장치는 커널 텐서 및 인자 행렬들 각각을 부호화하여 대응하는 비트스트림들을 생성한 후, VCM 복호화 장치로 전송할 수 있다.
- [0108] CP 분해를 이용하는 경우, VCM 부호화 장치는 P 개의 랭크 1 텐서들 각각을 부호화하여 대응하는 비트스트림들을 생성한 후, VCM 복호화 장치로 전송할 수 있다.
- [0109] CP 분해를 이용하는 경우의 다른 실시예로서, VCM 부호화 장치는 크기가 같은 랭크 1 텐서들끼리 결합한 텐서들을 생성할 수 있다. VCM 부호화 장치는 이들을 부호화하여 대응하는 비트스트림들을 생성한 후, VCM 복호화 장치로 전송할 수 있다.
- [0110] 도 12는 본 개시의 일 실시예에 따른, 중간 특징맵 복원을 이용하는 VCM 복호화 방법을 나타내는 블록도이다.
- [0111] 여기서, 중간 특징맵은 머신 비전 부호화 장치 내 딥러닝 모델의 중간 계층으로부터 생성된 특징맵을 나타낸다. 특징맵은 동일한 높이와 너비를 갖는 2D 특징맵을 채널의 개수만큼 포함한다.
- [0112] VCM 복호화 장치는 비디오 복호화기를 이용하여 비트스트림으로부터 변형 특징맵을 복호화한다(S1200).
- [0113] 한편, 축소 특징맵이 채널 측면에서의 특징맵 희소화를 기반으로 축소된 경우, VCM 복호화 장치는 다음과 같이 비트스트림을 복호화한다. 희소화된 2D 특징맵에 대해 VCM 복호화 장치는 기설정된 채널 간격 및 희소화된 2D 특징맵을 복호화한다. 또는, 삭제된 2D 특징맵에 대해 VCM 복호화 장치는 기설정된 채널 간격 및 삭제된 2D 특징맵의 인덱스를 복호화한다.
- [0114] VCM 복호화 장치는 비디오 시퀀스 형태의 변형 특징맵을 재배열한다(S1202).
- [0115] VCM 복호화 장치는, 재배열된 변형 특징맵의 데이터 타입을 변환하기 위해, 재배열된 변형 특징맵을 역양자화하여 축소 특징맵을 복원한다(S1204).
- [0116] VCM 복호화 장치는 축소 특징맵의 크기를 확장하여 복원 특징맵을 생성한다(S1206). 여기서, 복원 특징맵은 전술한 중간 특징맵에 대응한다.
- [0117] 축소 특징맵이 채널 측면에서의 특징맵 희소화를 기반으로 축소된 경우, VCM 복호화 장치는 다음과 같이 복원 특징맵을 생성한다. 희소화된 2D 특징맵에 대해, VCM 복호화 장치는 기설정된 채널 간격의 앞 또는 뒤의 복원 2D 특징맵을 희소화된 2D 특징맵의 위치에 복사함으로써, 복원 특징맵을 생성할 수 있다. 또는, 삭제된 2D 특징맵에 대해, VCM 복호화 장치는 복호화된 인덱스를 참조하여 기설정된 채널 간격의 앞 또는 뒤의 복원 2D 특징맵을 삭제된 2D 특징맵의 위치에 복사함으로써, 복원 특징맵을 생성할 수 있다.
- [0118] 다른 실시예로서, 축소 특징맵이 터커 분해를 이용하여 축소된 경우, 수학적 식 1에 나타난 바와 같이, VCM 복호화 장치는 축소 특징맵을 구성하는 커널 텐서 및 인자 행렬들을 이용하여 복원 특징맵을 근사할 수 있다.
- [0119] 또다른 실시예로서, 축소 특징맵이 CP 분해를 이용하여 축소된 경우, 수학적 식 3에 나타난 바와 같이, VCM 복호화 장치는 축소 특징맵을 구성하는 P 개의 랭크 1 텐서들을 이용하여 복원 특징맵을 근사할 수 있다.
- [0120] 본 명세서의 흐름도/타이밍도에서는 각 과정들을 순차적으로 실행하는 것으로 기재하고 있으나, 이는 본 개시의 일 실시예의 기술 사상을 예시적으로 설명한 것에 불과한 것이다. 다시 말해, 본 개시의 일 실시예가 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 개시의 일 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 흐름도/타이밍도에 기재된 순서를 변경하여 실행하거나 각 과정들 중 하나 이상의 과정을 병렬적으로 실행하는 것으로 다양하게 수정 및 변형하여 적용 가능할 것이므로, 흐름도/타이밍도는 시계열적인 순서로 한정되는 것은 아니다.
- [0121] 이상의 설명에서 예시적인 실시예들은 많은 다른 방식으로 구현될 수 있다는 것을 이해해야 한다. 하나 이상의 예시들에서 설명된 기능들 혹은 방법들은 하드웨어, 소프트웨어, 펌웨어 또는 이들의 임의의 조합으로 구현될 수 있다. 본 명세서에서 설명된 기능적 컴포넌트들은 그들의 구현 독립성을 특히 더 강조하기 위해 "...부

(unit)" 로 라벨링되었음을 이해해야 한다.

[0122] 한편, 본 실시예에서 설명된 다양한 기능들 혹은 방법들은 하나 이상의 프로세서에 의해 관독되고 실행될 수 있는 비일시적 기록매체에 저장된 명령어들로 구현될 수도 있다. 비일시적 기록매체는, 예를 들어, 컴퓨터 시스템에 의하여 관독가능한 형태로 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 예를 들어, 비일시적 기록매체는 EPROM(erasable programmable read only memory), 플래시 드라이브, 광학 드라이브, 자기 하드 드라이브, 솔리드 스테이트 드라이브(SSD)와 같은 저장매체를 포함한다.

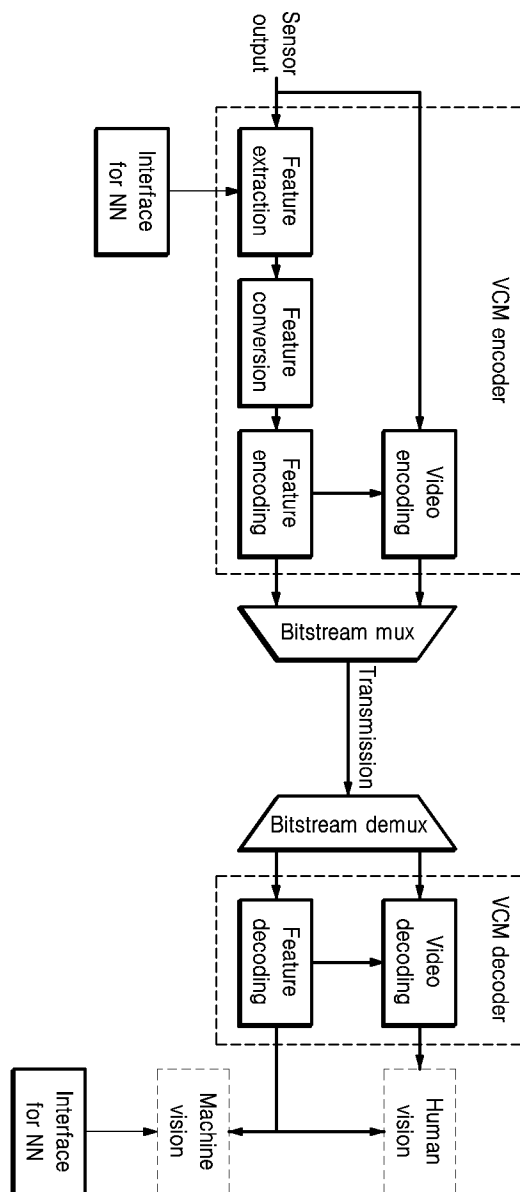
[0123] 이상의 설명은 본 실시예의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 실시예가 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서, 본 실시예들은 본 실시예의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

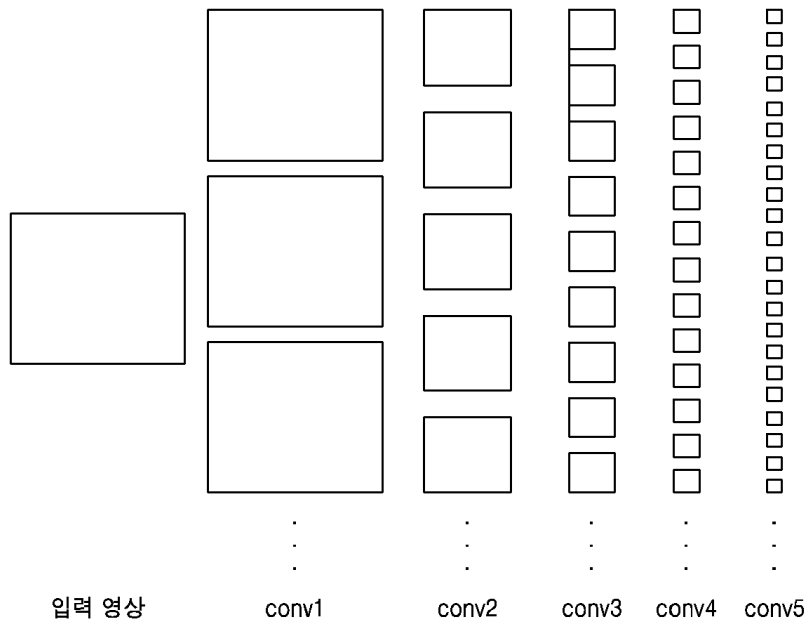
[0124] 410: 특징 추출기
420: 특징 변형기
430: 특징 부호화기
450: 특징 복호화기
460: 특징 역변형기
512: 특징 축소기
612: 특징 복원기

도면

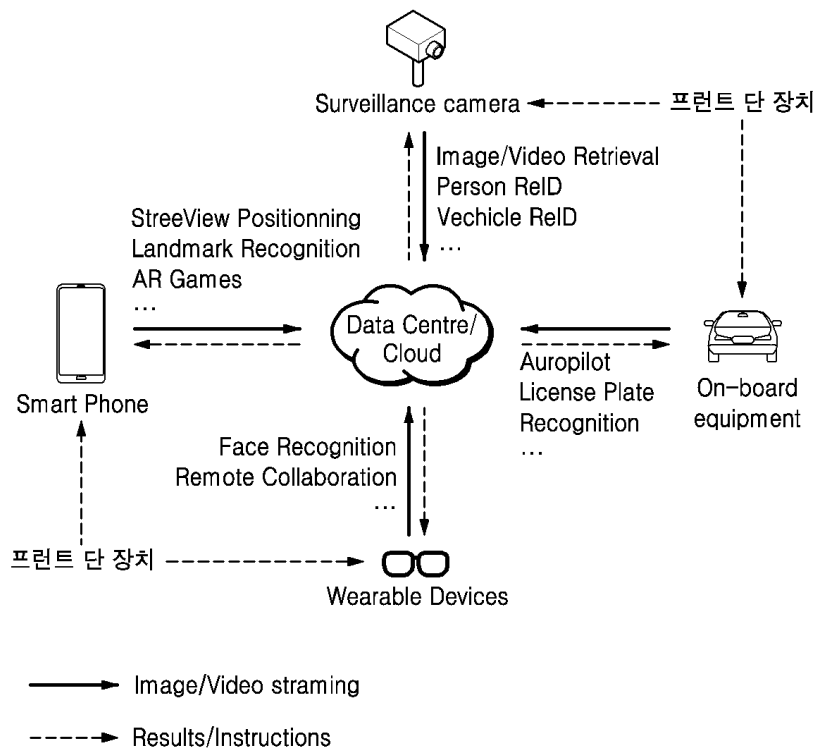
도면1



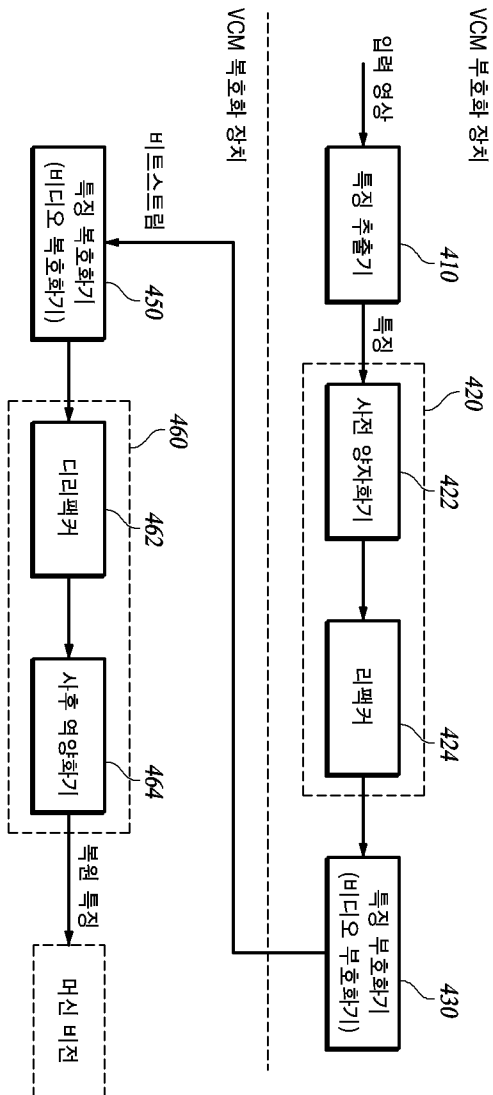
도면2



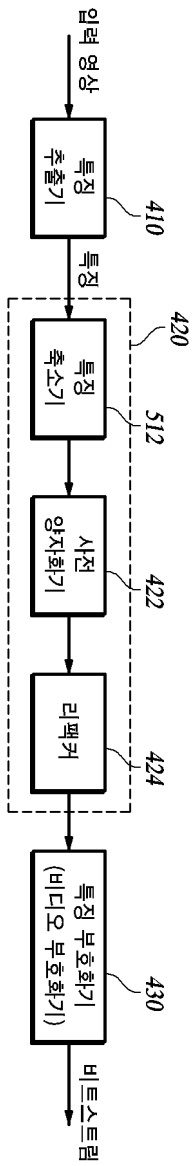
도면3



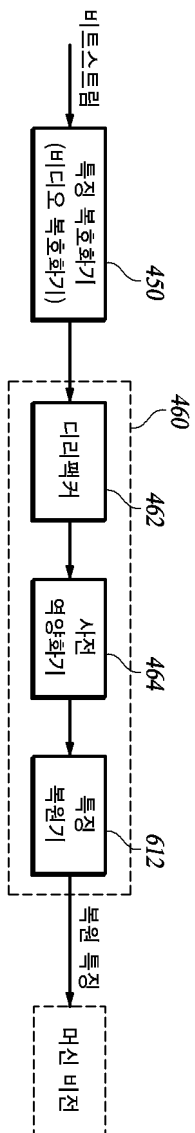
도면4



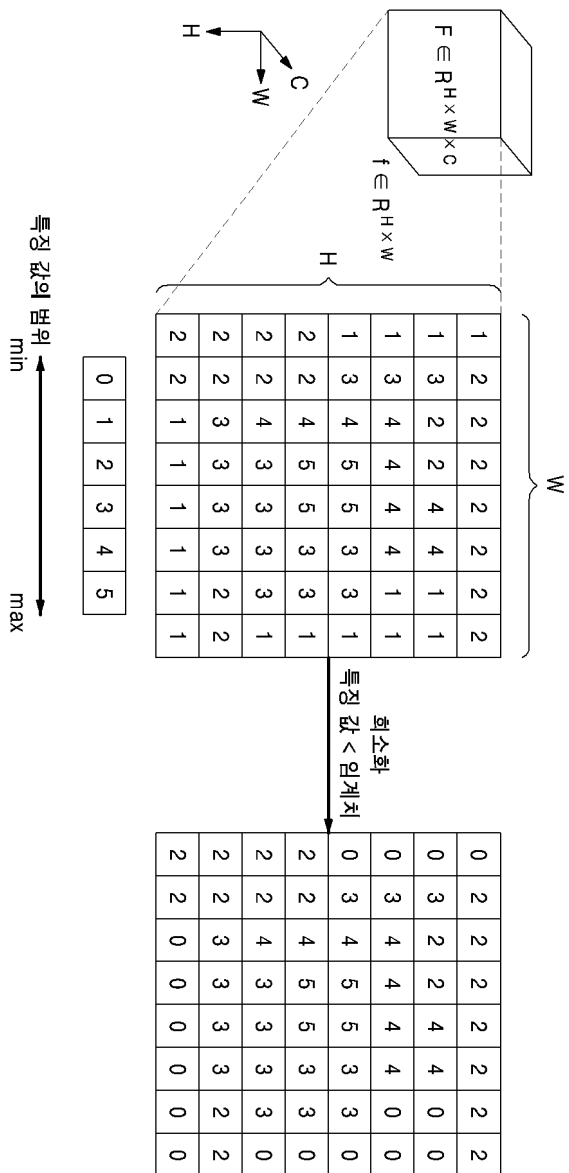
도면5



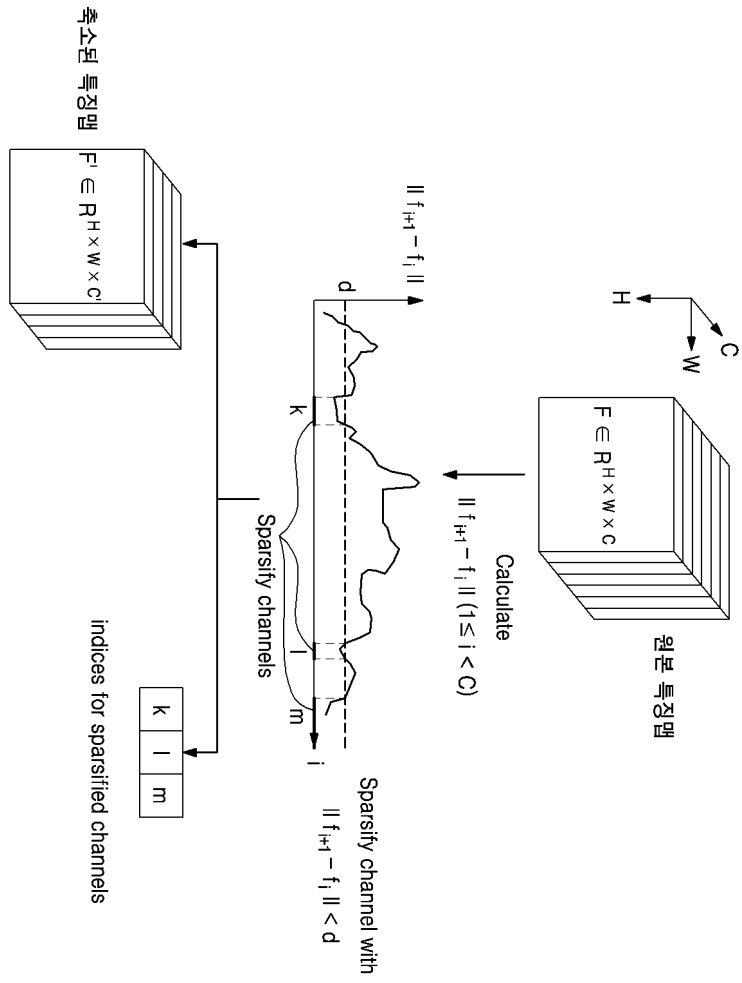
도면6



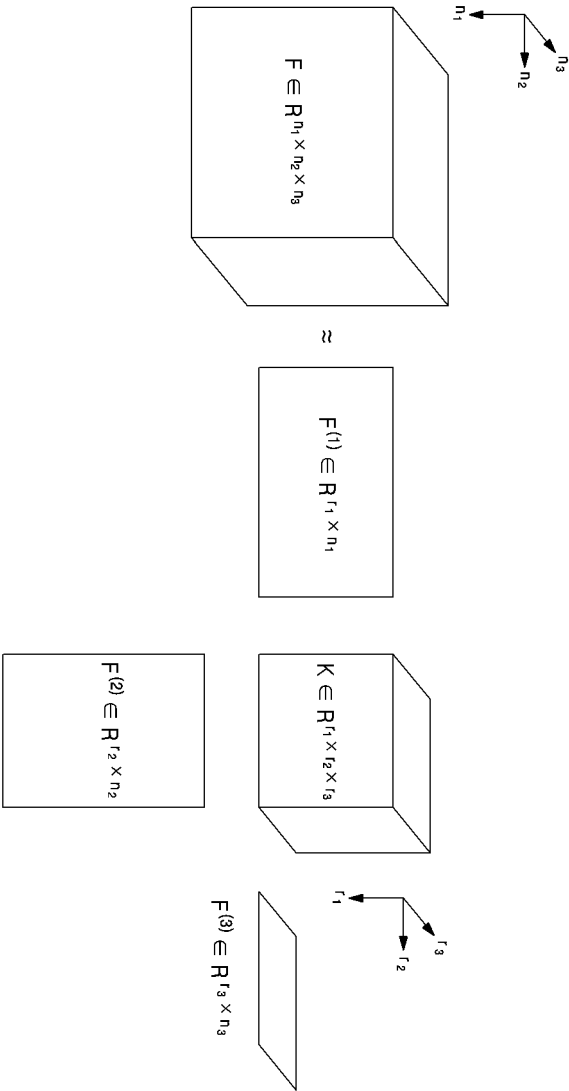
도면7



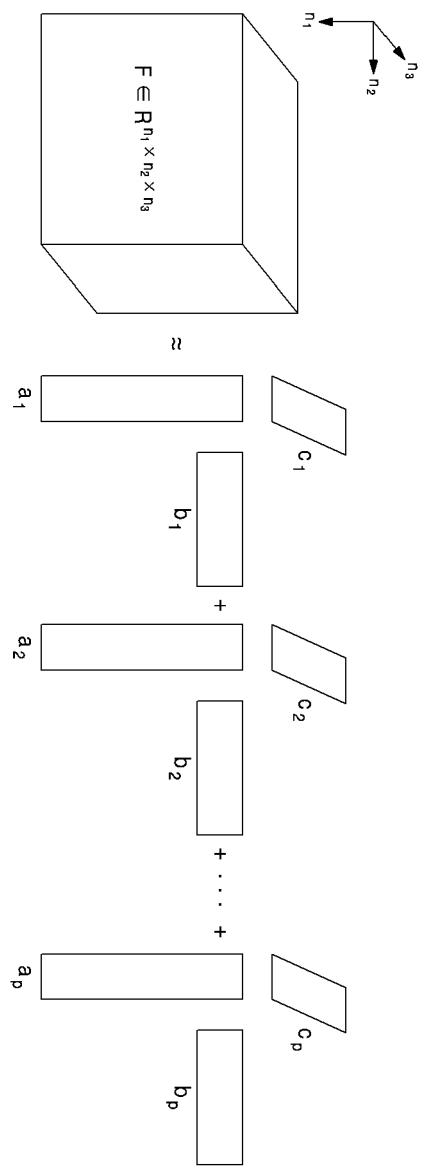
도면8



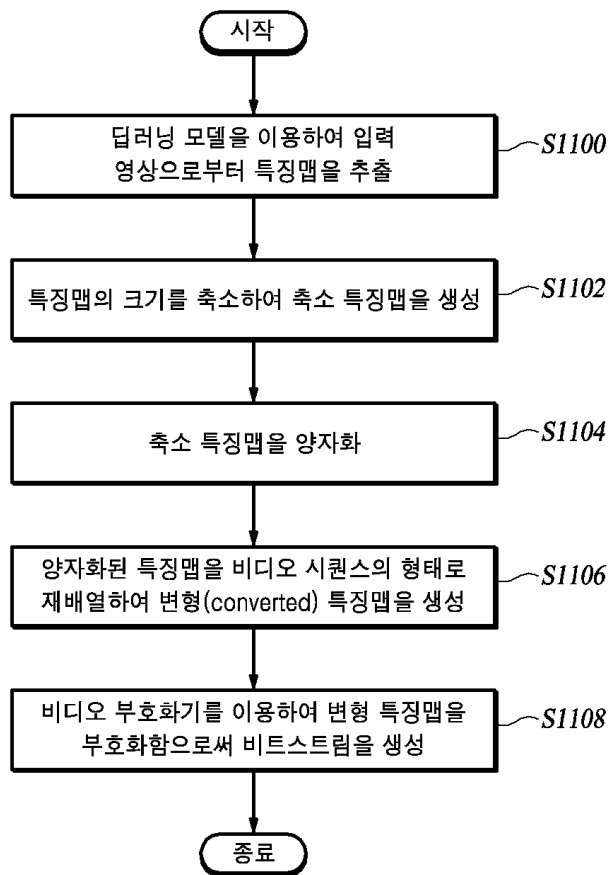
도면9



도면10



도면11



도면12

