



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0073666
(43) 공개일자 2022년06월03일

- (51) 국제특허분류(Int. Cl.)
H04N 19/122 (2014.01) G06N 3/04 (2006.01)
G06N 3/08 (2006.01) H04N 19/124 (2014.01)
H04N 19/132 (2014.01) H04N 19/136 (2014.01)
H04N 19/176 (2014.01) H04N 19/30 (2014.01)
H04N 19/563 (2014.01) H04N 19/60 (2014.01)
H04N 19/96 (2014.01)
- (52) CPC특허분류
H04N 19/122 (2015.01)
G06N 3/04 (2013.01)
- (21) 출원번호 10-2021-0161984
(22) 출원일자 2021년11월23일
심사청구일자 없음
- (30) 우선권주장
1020200161619 2020년11월26일 대한민국(KR)
1020210052318 2021년04월22일 대한민국(KR)
- (71) 출원인
한국전자통신연구원
대전광역시 유성구 가정로 218 (가정동)
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
- (72) 발명자
도지훈
대전광역시 유성구 가정로 218
정세윤
대전광역시 유성구 가정로 218
(뒷면에 계속)
- (74) 대리인
(유)한양특허법인

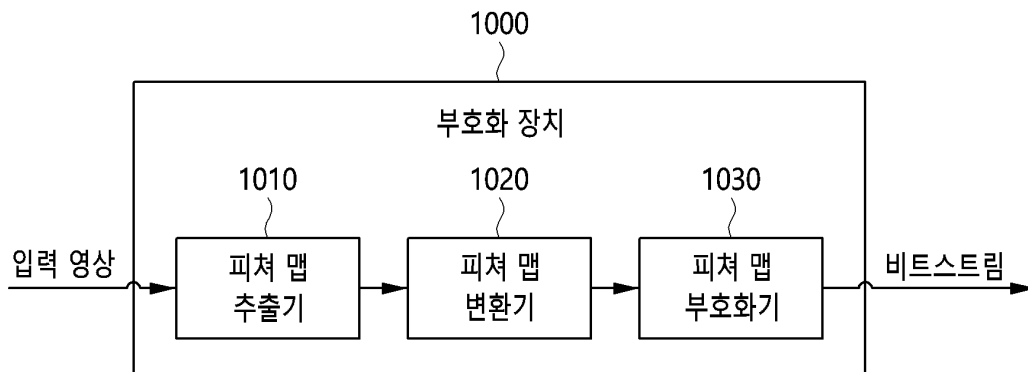
전체 청구항 수 : 총 1 항

(54) 발명의 명칭 인공 신경망의 피쳐 맵을 사용하는 영상 부호화/복호화를 위한 방법, 장치 및 기록 매체

(57) 요약

인공 신경망 피쳐 맵을 사용하는 영상 부호화/복호화를 위한 방법, 장치 및 기록 매체가 제공된다. 부호화 방법은 입력 영상으로부터 피쳐 맵을 추출하는 단계, 추출된 피쳐 맵에 대한 변환을 수행함으로써 변환된 피쳐 맵을 생성하는 단계 및 변환된 피쳐 맵에 대한 부호화를 수행하는 단계를 포함할 수 있다. 복호화 방법은 부호화된 피쳐 맵에 대한 정보에 대한 복호화를 수행함으로써 변환된 피쳐 맵을 복원하는 단계 및 복원된 변환된 피쳐 맵에 대한 역변환을 수행함으로써 복원된 피쳐 맵을 생성하는 단계를 포함할 수 있다.

대표도 - 도10



(52) CPC특허분류

G06N 3/08 (2013.01)
H04N 19/124 (2015.01)
H04N 19/132 (2015.01)
H04N 19/136 (2015.01)
H04N 19/176 (2015.01)
H04N 19/30 (2015.01)
H04N 19/563 (2015.01)
H04N 19/60 (2015.01)
H04N 19/96 (2015.01)

(72) 발명자

최진수
대전광역시 유성구 가정로 218
권형진
대전광역시 유성구 가정로 218
김동현
대전광역시 유성구 가정로 218
김연희
대전광역시 유성구 가정로 218
김중호
대전광역시 유성구 가정로 218

이주영

대전광역시 유성구 가정로 218

이태진

대전광역시 유성구 가정로 218

김휘용

대전광역시 유성구 은구비남로 34

최창균

서울특별시 강남구 언주로146길 18

반현민

경기도 수원시 영통구 삼성로320번길 62

이 발명을 지원한 국가연구개발사업

과제고유번호	1711102942
과제번호	2020-0-00011
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원(IITP)
연구사업명	한국전자통신연구원연구개발지원사업
연구과제명	[전문연구실] 기계를 위한 영상 부호화
기 여 율	1/1
과제수행기관명	한국전자통신연구원
연구기간	2020.01.01 ~ 2020.12.31

명세서

청구범위

청구항 1

입력 영상으로부터 피쳐 맵을 추출하는 단계;

상기 추출된 피쳐 맵에 대한 변환을 수행함으로써 변환된 피쳐 맵을 생성하는 단계; 및

상기 변환된 피쳐 맵에 대한 부호화를 수행하는 단계

를 포함하는, 부호화 방법.

발명의 설명

기술 분야

[0001] 본 발명은 영상 부호화/복호화를 위한 방법, 장치 및 기록 매체에 관한 것이다. 구체적으로, 본 발명은 인공 신경망의 피쳐 맵을 사용하는 영상 부호화/복호화를 위한 방법, 장치 및 기록 매체를 제공한다.

배경 기술

[0002] 머신 러닝을 사용하는 작업이 대형 서버뿐만 아니라 모바일 장치를 포함한 다양한 장치에서 광범위하게 사용되면서, 피쳐 맵(feature map)에 대한 추출 수단 및 작업 수행 수단이 하나의 장치 내에 존재하지 않고, 서로 다른 장치들 내에 각각 존재하는 경우도 점점 많아지고 있다.

[0003] 이렇게 피쳐 맵에 대한 추출 수단과 작업 수행 수단이 서로 분리되어 있는 경우에는, 추출 수단에 의한 추출된 특징 맵이 작업 수행 수단으로 전달해야만 한다. 그러나, 피쳐 맵의 데이터량은 매우 크기 때문에, 작업 수행 성능의 저하를 최소화하면서도 피쳐 맵의 데이터량을 감소시키기 위한 피쳐 맵에 대한 부호화 방법이 요구된다.

발명의 내용

해결하려는 과제

[0004] 인공 신경망을 통해 추출된 피쳐 맵에 대한 변환(conversion) 방법 및 부호화 방법을 제공함으로써, 특징 맵의 압축 비트량을 줄이면서도 작업의 성능 저하도 줄이는 장치, 방법 및 기록 매체를 제공할 수 있다.

[0005] 피쳐 맵 채널의 크기가 피쳐 맵 부호화 수단의 블록 크기와 정렬되도록 함으로써 특징 맵 부호화 성능을 향상시키는 장치, 방법 및 기록 매체를 제공할 수 있다.

과제의 해결 수단

[0006] 일 측에 있어서, 입력 영상으로부터 피쳐 맵을 추출하는 단계; 상기 추출된 피쳐 맵에 대한 변환을 수행함으로써 변환된 피쳐 맵을 생성하는 단계; 및 상기 변환된 피쳐 맵에 대한 부호화를 수행하는 단계를 포함하는 부호화 방법이 제공된다.

발명의 효과

[0007] 인공 신경망을 통해 추출된 피쳐 맵에 대한 변환 방법 및 부호화 방법을 제공함으로써, 특징 맵의 압축 비트량을 줄이면서도 작업의 성능 저하도 줄이는 장치, 방법 및 기록 매체가 제공된다.

[0008] 피쳐 맵 채널의 크기가 피쳐 맵 부호화 수단의 블록 크기와 정렬되도록 함으로써 특징 맵 부호화 성능을 향상시키는 장치, 방법 및 기록 매체가 제공된다.

도면의 간단한 설명

[0009] 도 1은 인공 신경망의 하나인 고속 지역-기반된 컨볼루션 신경망(Fast. Region-based. Convolutional Neural

Network; Fast R-CNN)을 이용하여 물체 검출 및 분류를 하는 기계 작업 결과의 예시이다.

도 2는 일 예에 따른 마스크 R-CNN의 구조를 나타낸다.

도 3은 일 예에 따른 원본 피쳐 맵을 나타낸다.

도 4는 일 예에 따른 p2 피쳐 맵을 나타낸다.

도 5는 일 예에 따른 p3 피쳐 맵을 나타낸다.

도 6은 일 예에 따른 p4 피쳐 맵을 나타낸다.

도 7은 일 예에 따른 p5 피쳐 맵을 나타낸다.

도 8은 일 예에 따른 p6 피쳐 맵을 나타낸다.

도 9는 일 예에 따른 멀티-테스크 모델 협업 인텔리전스를 나타낸다.

도 10은 일 실시예에 따른 부호화 장치의 구조도이다.

도 11은 일 실시예에 따른 복호화 장치의 구조도이다.

도 12는 일 예에 따른 피쳐 맵 변환기의 구조도이다.

도 13은 일 예에 따른 피쳐 맵 역변환기의 구조도이다.

도 14는 HEVC 및 VVC와 같은 기존 영상 부호화 장치 및 복호화 장치에 기반하여 피쳐 맵 부호화를 수행하는 예를 나타낸다.

도 15은 피쳐 맵 부호화를 위해 인공 신경망을 사용하는 예를 나타낸다.

도 16는 일 실시예에 따른 피쳐 맵 부호화 방법의 흐름도이다.

도 17은 일 실시예에 따른 피쳐 맵 복호화 방법의 흐름도이다.

도 18은 일 예에 따른 코딩 블록의 크기를 나타낸다.

도 19는 일 예에 따른 크기가 조정되기 전의 피쳐 맵의 채널을 나타낸다.

도 20은 일 예에 따른 크기가 조정된 후의 피쳐 맵의 채널을 나타낸다.

도 21은 일 예에 따른 크기의 조정의 전 및 크기의 조정의 후의 피쳐 맵의 채널의 정보의 위치들 간의 관계를 나타낸다.

도 22는 다른 일 예에 따른 크기의 조정의 전 및 크기의 조정의 후의 피쳐 맵의 채널의 정보의 위치들 간의 관계를 나타낸다.

도 23은 또 일 예에 따른 크기의 조정의 전 및 크기의 조정의 후의 피쳐 맵의 채널의 정보의 위치들 간의 관계를 나타낸다.

도 24는 일 실시예에 따른 피쳐 맵 채널들의 크기의 조정을 나타낸다.

도 25는 일 예에 따른 피쳐 맵의 구성을 나타낸다.

도 26은 일 예에 따른 공간적으로 배열된 피쳐 맵을 나타낸다.

도 27은 일 예에 따른 시간적으로 배열된 피쳐 맵을 나타낸다.

도 28은 일 예에 따른 시간적으로 배열된 피쳐 맵을 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0010] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0011] 후술하는 예시적 실시예들에 대한 상세한 설명은, 특정 실시예를 예시로서 도시하는 첨부 도면을 참조한다. 이

들 실시예는 당업자가 실시예를 실시할 수 있기에 충분하도록 상세히 설명된다. 다양한 실시예들은 서로 다르지만 상호 배타적일 필요는 없음이 이해되어야 한다. 예를 들면, 여기에 기재되어 있는 특정 형상, 구조 및 특성은 일 실시예에 관련하여 본 발명의 정신 및 범위를 벗어나지 않으면서 다른 실시예로 구현될 수 있다. 또한, 각각의 개시된 실시예 내의 개별 구성요소의 위치 또는 배치는 실시예의 정신 및 범위를 벗어나지 않으면서 변경될 수 있음이 이해되어야 한다. 따라서, 후술하는 상세한 설명은 한정적인 의미로서 취하려는 것이 아니며, 예시적 실시예들의 범위는, 적절하게 설명된다면, 그 청구항들이 주장하는 것과 균등한 모든 범위와 더불어 첨부된 청구항에 의해서만 한정된다.

- [0012] 도면에서 유사한 참조부호는 여러 측면에 걸쳐서 동일하거나 유사한 기능을 지칭한다. 도면에서의 요소들의 형상 및 크기 등은 보다 명확한 설명을 위해 과장될 수 있다.
- [0013] 본 발명에서 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들면, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. "및/또는"이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함할 수 있다.
- [0014] 어떤 구성요소(component)가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 상기의 2개의 구성요소들이 서로 간에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있으나, 상기의 2개의 구성요소들의 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소(component)가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 상기의 2개의 구성요소들의 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다.
- [0015] 실시예들에서 나타나는 구성요소들은 서로 다른 특징적인 기능들을 나타내기 위해 독립적으로 도시되는 것으로, 각 구성요소들이 분리된 하드웨어나 하나의 소프트웨어 구성단위로 이루어짐을 의미하지 않는다. 즉, 각 구성요소는 설명의 편의상 각각의 구성요소로 나열하여 포함한 것으로 각 구성요소 중 적어도 두 개의 구성요소가 합쳐져 하나의 구성요소로 이루어지거나, 하나의 구성요소가 복수 개의 구성요소로 나뉘어져 기능을 수행할 수 있고 이러한 각 구성요소의 통합된 실시예 및 분리된 실시예도 본 발명의 본질에서 벗어나지 않는 한 본 발명의 권리범위에 포함된다.
- [0016] 실시예들에서 사용된 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 실시예들에서, "포함하다" 또는 "가지다" 등의 용어는 명세서 상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다. 즉, 실시예들에서 특정 구성을 "포함"한다고 기술하는 내용은 해당 구성 이외의 구성을 배제하는 것이 아니며, 추가적인 구성 또한 본 발명의 실시 또는 본 발명의 기술적 사상의 범위에 포함될 수 있음을 의미한다.
- [0017] 실시예들에서 용어 "적어도 하나(at least one)"는 1, 2, 3 및 4와 같은 1 이상의 개수들 중 하나를 의미할 수 있다. 실시예들에서 용어 "복수(a plurality of)"는 2, 3 및 4와 같은 2 이상의 개수들 중 하나를 의미할 수 있다.
- [0018] 실시예들의 일부의 구성요소는 본 발명에서 본질적인 기능을 수행하는 필수적인 구성요소는 아니고 단지 성능을 향상시키기 위한 선택적 구성요소일 수 있다. 실시예들은, 단지 성능 향상을 위해 사용되는 구성요소가 제외된, 실시예들의 본질을 구현함에 있어 필수적인 구성요소만을 포함하여 구현될 수 있다. 단지 성능 향상을 위해 사용되는 선택적인 구성요소가 제외된 필수적인 구성요소만을 포함하는 구조도 실시예들의 권리범위에 포함된다.
- [0019] 이하에서는, 기술분야에서 통상의 지식을 가진 자가 실시예들을 용이하게 실시할 수 있도록 하기 위하여, 첨부된 도면을 참조하여 실시 형태에 대하여 구체적으로 설명한다. 실시예들을 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 명세서의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략한다. 또한, 도면 상의 동일한 구성요소에 대해서는 동일한 참조부호를 사용하고, 동일한 구성요소에 대한 중복된 설명은 생략한다.
- [0020] 이하에서, 영상은 비디오(video)를 구성하는 하나의 픽처(picture)를 의미할 수 있으며, 비디오 자체를 나타낼 수도 있다. 예를 들면, "영상의 부호화 및/또는 복호화"는 "비디오의 부호화 및/또는 복호화"를 의미할 수 있으며, "비디오를 구성하는 영상들 중 하나의 영상의 부호화 및/또는 복호화"를 의미할 수도 있다.

- [0021] 이하에서, 용어들 "비디오(video)" 및 "동영상(motion picture(s))"은 동일한 의미로 사용될 수 있으며, 서로 교체되어 사용될 수 있다.
- [0022] 이하에서, 대상 영상은 부호화의 대상인 부호화 대상 영상 및/또는 복호화의 대상인 복호화 대상 영상일 수 있다. 또한, 대상 영상은 부호화 장치로 입력된 입력 영상일 수 있고, 복호화 장치로 입력된 입력 영상일 수 있다. 또한, 대상 영상은 현재 부호화 및/또는 복호화의 대상인 현재 영상일 수 있다. 예를 들면, 용어들 "대상 영상" 및 "현재 영상"은 동일한 의미로 사용될 수 있으며, 서로 교체되어 사용될 수 있다.
- [0023] 이하에서, 용어들 "영상(image)", "픽처(picture)", "프레임(frame)" 및 "스크린(screen)"은 동일한 의미로 사용될 수 있으며, 서로 교체되어 사용될 수 있다.
- [0024] 이하에서, 대상 블록은 부호화의 대상인 부호화 대상 블록 및/또는 복호화의 대상인 복호화 대상 블록일 수 있다. 또한, 대상 블록은 현재 부호화 및/또는 복호화의 대상인 현재 블록일 수 있다. 예를 들면, 용어들 "대상 블록" 및 "현재 블록"은 동일한 의미로 사용될 수 있으며, 서로 교체되어 사용될 수 있다. 현재 블록은 부호화 시 부호화의 대상이 되는 부호화 대상 블록 및/또는 복호화 시 복호화의 대상이 되는 복호화 대상 블록을 의미할 수 있다.
- [0025] 이하에서, 용어들 "블록" 및 "유닛"은 동일한 의미로 사용될 수 있으며, 서로 교체되어 사용될 수 있다. 또는 "블록"은 특정한 유닛을 나타낼 수 있다.
- [0026] 이하에서, 용어들 "영역(region)" 및 "세그먼트(segment)"는 서로 교체되어 사용될 수 있다.
- [0027] 실시예들에서, 특정된 정보, 데이터, 플래그(flag), 인덱스(index) 및 요소(element), 속성(attribute) 등의 각각은 값을 가질 수 있다. 정보, 데이터, 플래그, 인덱스, 요소 및 속성 등의 값 "0"은 거짓(false), 논리 거짓(logical false) 또는 제1 기정의된(predefined) 값을 나타낼 수 있다. 말하자면, 값 "0", 거짓, 논리 거짓 및 제1 기정의된 값은 서로 대체되어 사용될 수 있다. 정보, 데이터, 플래그, 인덱스, 요소 및 속성 등의 값 "1"은 참(true), 논리 참(logical true) 또는 제2 기정의된(predefined) 값을 나타낼 수 있다. 말하자면, 값 "1", 참, 논리 참 및 제2 기정의된 값은 서로 대체되어 사용될 수 있다.
- [0028] 행, 열 또는 인덱스를 나타내기 위해 i 또는 j 등의 변수가 사용될 때, i의 값은 0 이상의 정수일 수 있으며, 1 이상의 정수일 수도 있다. 말하자면, 실시예들에서 행, 열 및 인덱스 등은 0에서부터 카운트될 수 있으며, 1에서부터 카운트될 수 있다.
- [0029] 실시예들에서, 용어 "하나 이상" 또는 용어 "적어도 하나"는 용어 "복수"를 의미할 수 있다. "하나 이상" 또는 "적어도 하나"는 "복수"로 교체되어 사용될 수 있다.
- [0030] 인공 신경 망과 기계 작업
- [0031] 도 1은 인공 신경망의 하나인 고속 지역-기반된 컨볼루션 신경망(Fast. Region-based. Convolutional Neural Network; Fast R-CNN)을 이용하여 물체 검출 및 분류를 하는 기계 작업 결과의 예시이다.
- [0032] 물체 분류(Object Classification), 물체 인식(Object Recognition), 물체 검출(Object Detection), 물체 영역 분할(Object Segmentation), 물체 추적(Object Tracking) 등 다양한 기계비전(Machine Vision) 작업이나 해상도 향상(Super-Resolution), 프레임 보간(Frame-Interpolation)등 다양한 영상 처리(Image Processing) 작업과 같은 기계 작업(Machine Task)에서 인공 신경망(Artificial Neural Network; ANN)을 활용하는 사례가 점점 늘어나고 있다.
- [0033] 피쳐 맵 추출 수단과 작업 수행 수단
- [0034] 기계 작업(Machine Task)을 수행하는 인공 신경망 모델은 통상 입력 데이터 혹은 입력 영상으로부터 피쳐를 추출하는 피쳐 맵 추출 수단과 추출된 피쳐에 기반하여 구체적인 기계 작업을 실제로 수행하는 작업 수행 수단으로 구성될 수 있다.
- [0035] 이때, 입력이 영상 형태를 갖는 경우 추출된 피쳐는 통상 피쳐 맵이라 불리울 수 있다.
- [0036] 실시예들에서는, 피쳐 맵이라는 표현을 사용하여 설명되지만, 맵의 형태가 아닌 다른 형태를 갖는 피쳐의 경우에도 실시예들이 동일하게 적용될 수 있다. 또한, 용어 "피쳐 맵" 또는 "피쳐 맵 정보" 등은 용어 "피쳐"로 대체될 수 있다.
- [0037] 도 2는 일 예에 따른 마스크 R-CNN의 구조를 나타낸다.

- [0038] 마스크(Mask) R-CNN은 물체 영역 분할에 사용되는 인공 신경망 모델일 수 있다.
- [0039] 이 구조에서는 피쳐 피라미드 네트워크(Feature Pyramid Network; FPN)이 피쳐 맵 추출 수단으로 사용될 수 있고, 영역 제안 네트워크(Region Proposal Network; RPN)와 관심영역 헤드(Region Of Interest Heads; ROI Heads)가 작업 수행 수단으로 사용될 수 있다.
- [0040] 피쳐 맵
- [0041] 도 3은 일 예에 따른 원본 피쳐 맵을 나타낸다.
- [0042] 도 4는 일 예에 따른 p2 피쳐 맵을 나타낸다.
- [0043] 도 5는 일 예에 따른 p3 피쳐 맵을 나타낸다.
- [0044] 도 6은 일 예에 따른 p4 피쳐 맵을 나타낸다.
- [0045] 도 7은 일 예에 따른 p5 피쳐 맵을 나타낸다.
- [0046] 도 8은 일 예에 따른 p6 피쳐 맵을 나타낸다.
- [0047] 입력이 영상인 경우, 피쳐 맵의 형태는 가로x세로의 2차원 어레이(array)로 표현될 수 있다. 하나의 계층(layer)의 피쳐 맵은 통상 여러 개의 채널로 구성될 수 있으므로, 각 계층의 피쳐 맵은 가로_크기 x 세로_크기 x 채널_개수만큼의 크기를 갖는 3차원 어레이로서 표현될 수 있다.
- [0048] 즉, 계층 k의 피쳐 맵을 F_k 라 하면, F_k 는 추출된 피쳐 값들로 구성된 3차원 어레이 $F_k[x][y][c]$ 로 표현될 수 있다. 이 때, x와 y는 각각 피쳐 값의 가로 위치와 세로 위치를 나타낼 수 있다. c는 채널 인덱스를 나타낼 수 있다.
- [0049] 도 3 내지 도 8에서는, 도 2를 참조하여 전술된 마스크 R-CNN의 FPN을 통해 추출된 피쳐 맵이 예시되었다.
- [0050] FPN에서는, 각 계층의 피쳐 맵이 256 개의 채널들로 구성될 수 있다. 도 3 내지 도 8에서는, 각 계층의 피쳐 맵들 중 첫 번째 채널에 해당하는 피쳐 맵만이 예시적으로 도시되었다.
- [0051] 여기에서, FPN에서는 계층이 더 깊어질 수록 피쳐 맵의 가로 크기 및 세로 크기가 입력 영상(즉, 원본 영상)의 가로 크기 및 세로 크기에 비해 점점 더 작아질 수 있다.
- [0052] 도 9는 일 예에 따른 멀티-테스크 모델 협업 인텔리전스를 나타낸다.
- [0053] 기계 작업이 대형 서버뿐 아니라, 모바일 장치를 포함한 다양한 장치에서 광범위하게 사용되면서 피쳐 맵 추출 수단과 작업 수행 수단이 같은 장치내에 존재하지 않고 서로 다른 장치 내에 존재하는 경우도 점점 많아지고 있다.
- [0054] 도 9에서는, 피쳐 맵 추출 수단은 모바일 장치에 존재하되, 물체 분할이나 디스패리티 맵 추정, 또는 영상 복원 등과 같은 구체적인 작업 수행 수단은 클라우드 서버에 존재하는 상황이 예시되었다.
- [0055] 이러한 상황에서는, 모바일 장치에서 추출된 피쳐 맵이 서버로 전달될 수 있고, 서버의 작업 수행 결과가 다시 모바일 장치로 전달될 수 있다.
- [0056] 이러한 예와 같이, 피쳐 맵 추출 수단과 작업 수행 수단이 분리되어 있는 경우에는 추출된 피쳐 맵을 작업 수행 수단에 전달해야만 할 수 있다. 또한, 전달 또는 저장할 피쳐 맵의 데이터량을 최소화하는 동시에 작업 수행 성능의 저하를 최소화하기 위한 피쳐 맵 부호화 방법이 요구될 수 있다.
- [0057] 또 다른 예로, 피쳐 맵 추출 수단과 작업 수행 수단이 한 장치에 존재하는 경우라 할 지라도 추출된 피쳐 맵을 저장 장치를 통해 저장할 수 있고, 이후 저장된 피쳐 맵이 작업 수행 수단에서 활용될 수 있다. 이러한 경우에도 상기와 같은 피쳐 맵 부호화 방법이 요구될 수 있다.
- [0058] 실시예들에서는, 이러한 피쳐 맵 부호화에 대한 방법 및 장치에 관한 설명이 제시될 수 있다.
- [0059] 실시예들에서, 수단은 특정 장치를 의미할 수 있다.
- [0060] 도 10은 일 실시예에 따른 부호화 장치의 구조도이다.
- [0061] 부호화 장치(1000)는 피쳐 맵 추출기(1010), 피쳐 맵 변환기(converter)(1020) 및 피쳐 맵 부호화기(1030)를

포함할 수 있다.

- [0062] 피쳐 맵 추출기(1010)는 입력 영상으로부터 피쳐 맵을 추출할 수 있다.
- [0063] 피쳐 맵 변환기(1020)는 추출된 피쳐 맵에 대한 변환(conversion)을 수행함으로써 변환된 피쳐 맵을 생성할 수 있다.
- [0064] 변환은 피쳐 맵에 대한 양자화, 패딩, 크기 조정 및 배열(arrangement) 등을 포함할 수 있다.
- [0065] 피쳐 맵 변환기(1020)에 의해 피쳐 맵은 부호화에 적합한 포맷으로 변환될 수 있다.
- [0066] 피쳐 맵 부호화기(1030)는 변환된 피쳐 맵에 대한 부호화를 수행할 수 있다.
- [0067] 피쳐 맵 부호화기(1030)는 변환된 피쳐 맵에 대한 부호화를 수행함으로써 (부호화된) 피쳐 맵에 대한 정보를 생성할 수 있다.
- [0068] 부호화는 압축을 포함할 수 있다.
- [0069] (부호화된) 피쳐 맵에 대한 정보는 비트스트림 등을 통해 복호화 장치(1100)로 전송될 수 있고, 컴퓨터 판독-가능한 기록 매체 등에 저장될 수 있다.
- [0070] 도 11은 일 실시예에 따른 복호화 장치의 구조도이다.
- [0071] 복호화 장치(1100)는 피쳐 맵 복호화기(1110) 및 피쳐 맵 역변환기(inverse-converter)(1120)를 포함할 수 있다.
- [0072] 피쳐 맵 복호화기(1110)는 비트스트림 또는 컴퓨터 판독-가능한 기록 매체 등에 저장된 (부호화된) 피쳐 맵에 대한 정보에 대한 복호화를 수행함으로써 변환된 피쳐 맵을 복원(reconstructed)할 수 있다.
- [0073] 피쳐 맵 역변환기(inverse-converter)(1120)는 (복원된) 변환된 피쳐 맵에 대한 역변환을 수행함으로써 (복원된) 피쳐 맵을 생성할 수 있다.
- [0074] 말하자면, 피쳐 맵 역변환기(1120)는 (복원된) 변환된 피쳐 맵에 대한 역변환을 수행함으로써 피쳐 맵 추출기(1010)에 의해 추출된 최초의 피쳐 맵과 유사한 형식을 갖는 (복원된) 피쳐 맵을 생성할 수 있다.
- [0075] 복원된 피쳐 맵은 작업 수행 수단에 입력될 수 있고, 다양한 기계 작업에 사용될 수 있다.
- [0076] 도 12는 일 예에 따른 피쳐 맵 변환기의 구조도이다.
- [0077] 피쳐 맵 변환기(1020)는 피쳐 맵 크기 조정기(1210), 피쳐 맵 양자화기(1220) 및 피쳐 맵 배열기(1230)를 포함할 수 있다.
- [0078] 피쳐 맵 크기 조정기(1210)는 피쳐 맵의 크기를 조정할 수 있다.
- [0079] 피쳐 맵 양자화기(1220)는 피쳐 맵에 대한 양자화를 수행하여 양자화된 피쳐 맵을 생성할 수 있다.
- [0080] 피쳐 맵 배열기(1230)는 (양자화된) 피쳐 맵에 대한 배열을 수행하여 배열된 피쳐 맵을 생성할 수 있다.
- [0081] 배열된 피쳐 맵은 전술된 변환된 피쳐 맵일 수 있다. 배열이 수행되지 않은 경우, 양자화된 피쳐 맵은 전술된 변환된 피쳐 맵일 수 있다.
- [0082] 도 13은 일 예에 따른 피쳐 맵 역변환기의 구조도이다.
- [0083] 피쳐 맵 역변환기(1120)는 피쳐 맵 재배열기(1310), 피쳐 맵 역양자화기(1320) 및 피쳐 맵 크기 재조정기(1330)를 포함할 수 있다.
- [0084] 피쳐 맵 재배열기(1310)는 (복원된) 변환된 피쳐 맵에 대한 재배열(rearrangement)을 수행하여 재배열된 피쳐 맵을 생성할 수 있다.
- [0085] 재배열된 피쳐 맵은 부호화 장치(1000)에서의 양자화된 피쳐 맵에 대응할 수 있다.
- [0086] 피쳐 맵 역양자화기(1320)는 재배열된 피쳐 맵에 대한 역양자화를 수행하여 역양자화된 피쳐 맵을 생성할 수 있다.
- [0087] 부호화 장치(1000)에서 배열이 수행되지 않은 경우 또는 피쳐 맵 재배열기(1310)에서 재배열이 수행되지 않은 경우, 피쳐 맵 역양자화기(1320)는 (복원된) 변환된 피쳐 맵에 대한 역양자화를 수행하여 역양자화된 피쳐 맵을

생성할 수 있다.

- [0088] 피쳐 맵 크기 재조정기(1330)는 역양자화된 피쳐 맵의 크기를 재조정할 수 있다.
- [0089] 복원된 피쳐 맵은 피쳐 맵 크기 재조정기(1330)에 의해 재조정된 크기를 갖는 역양자화된 피쳐 맵일 수 있다.
- [0090] 또는, 복원된 피쳐 맵은 역양자화된 피쳐 맵일 수 있고, 피쳐 맵 크기 재조정기(1330)는 복원된 피쳐 맵의 크기를 재조정할 수 있다.
- [0091] 도 14는 HEVC 및 VVC와 같은 기존 영상 부호화 장치 및 복호화 장치에 기반하여 피쳐 맵 부호화를 수행하는 예를 나타낸다.
- [0092] 도 15은 피쳐 맵 부호화를 위해 인공 신경망을 사용하는 예를 나타낸다.
- [0093] 다양한 상황들에서 피쳐 맵 부호화 시스템을 사용하기 위해서는, 해당하는 시스템이 다양한 종류들의 피쳐 맵 추출 방식들과 피쳐 맵 변환 방식들을 지원하는 것이 유리할 수 있다.
- [0094] 이러한 지원을 실시예들에서는, 피쳐 맵 추출 과정에서 사용된 파라미터나 피쳐 맵 추출 방식 등에 관한 정보(예를 들면, 피쳐 맵 정보)들도 요구되는 경우 피쳐 맵 부호화기(1030)를 통해 부호화될 수 있고, 부호화된 파라미터 및 부호화된 정보는 비트스트림이나 컴퓨터 판독-가능한 기록 매체에 포함될 수 있다. 또한, 부호화된 파라미터 및 부호화된 정보 등이 최종적으로 작업 수행 수단에 전달될 수 있다.
- [0095] 도 16는 일 실시예에 따른 피쳐 맵 부호화 방법의 흐름도이다.
- [0096] 단계(1610)에서, 피쳐 맵 추출기(1010)는 입력 영상으로부터 피쳐 맵을 추출할 수 있다.
- [0097] 단계(1620)에서, 피쳐 맵 변환기(1020)는 추출된 피쳐 맵을 부호화에 적합한 형식으로 변환할 수 있다.
- [0098] 피쳐 맵 변환기(1020)는 추출된 피쳐 맵에 대한 변환(conversion)을 수행함으로써 변환된 피쳐 맵을 생성할 수 있다.
- [0099] 피쳐 맵은 하나 이상의 채널들을 포함할 수 있다. 이하에서, 피쳐 맵 채널은 피쳐 맵의 채널을 의미할 수 있다.
- [0100] 단계(1620)은 단계들(1621, 1622 및 1623)을 포함할 수 있다.
- [0101] 단계(1621)에서, 피쳐 맵 크기 조정기(1210)는 피쳐 맵의 크기를 조정할 수 있다.
- [0102] 피쳐 맵 크기 조정기(1210)는 피쳐 맵의 각 채널의 크기가 피쳐 맵 부호화기(1030)에서 사용되는 블록의 크기에 기반한 특정 조건을 만족하도록 피쳐 맵의 각 채널의 정보에 대한 처리를 수행할 수 있다. 처리는, 업스케일링, 다운스케일링, 패딩 및 크로핑을 포함할 수 있다.
- [0103] 특정 조건은 제1 특정 조건 및 제2 특정 조건 중 하나 이상을 포함할 수 있다.
- [0104] 제1 특정 조건은, 각 채널의 가로 크기가 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기의 정수 배가 되어야 하는 것일 수 있다. 제1 특정 조건은, 각 채널의 세로 크기가 피쳐 맵 부호화기(1030)에서 사용되는 블록의 세로 크기의 정수 배가 되어야 하는 것일 수 있다. 제1 특정 조건은, 각 채널의 가로 크기 및 세로 크기가 각각 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기 및 세로 크기의 정수 배가 되어야 하는 것일 수 있다.
- [0105] 제2 특정 조건은, 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기가 각 채널의 가로 크기의 정수 배가 되어야 하는 것일 수 있다. 제2 특정 조건은, 피쳐 맵 부호화기(1030)에서 사용되는 블록의 세로 크기가 각 채널의 세로 크기의 정수 배가 되어야 하는 것일 수 있다. 제2 특정 조건은, 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기 및 세로 크기가 각 채널의 가로 크기 및 세로 크기의 정수 배가 되어야 하는 것일 수 있다.
- [0106] 단계(1622)에서, 피쳐 맵 양자화기(1220)는 피쳐 맵에 대한 양자화를 수행하여 양자화된 피쳐 맵을 생성할 수 있다.
- [0107] 단계(1623)에서, 피쳐 맵 배열기(1230)는 (양자화된) 피쳐 맵에 대한 배열을 수행하여 배열된 피쳐 맵을 생성할 수 있다.
- [0108] 배열된 피쳐 맵은 전술된 변환된 피쳐 맵일 수 있다. 배열이 수행되지 않은 경우 양자화된 피쳐 맵은 전술된 변환된 피쳐 맵일 수 있다.

- [0109] 피쳐 맵 배열기(1230)는 (양자화된) 피쳐 맵의 하나 이상의 채널들을 시간적 및/또는 공간적으로 배열할 수 있다.
- [0110] 피쳐 맵 배열기(1230)는 1) 채널들에 대한 공간적 배열, 2) 채널들에 대한 시간적 배열 및 3) 채널들에 대한 시공간적 배열 중 적어도 하나를 수행할 수 있다.
- [0111] 단계(1630)에서, 피쳐 맵 부호화기(1030)는 변환된 피쳐 맵에 대한 부호화를 수행할 수 있다.
- [0112] 피쳐 맵 부호화기(1030)는 변환된 피쳐 맵, 피쳐 맵 정보 및 피쳐 맵 변환 정보에 대한 부호화를 수행함으로써 (부호화된) 피쳐 맵에 대한 정보를 생성할 수 있다.
- [0113] 부호화는 압축을 포함할 수 있다.
- [0114] 피쳐 맵 정보는 부호화의 대상인 피쳐 맵에 관한 정보일 수 있다.
- [0115] 피쳐 맵 정보는 1) 피쳐 맵의 채널의 가로 크기, 2) 피쳐 맵의 채널의 세로 크기, 3) 피쳐 맵의 하나 이상의 채널들의 계수, 4) 피쳐 맵의 계층 정보 및 5) 피쳐 맵 추출 정보 중 적어도 하나를 포함할 수 있다.
- [0116] 피쳐 맵 변환 정보는 피쳐 맵에 대한 변환에서 사용된 파라미터에 관한 정보일 수 있다.
- [0117] 피쳐 맵 변환 정보는 1) 피쳐 맵의 하나 이상의 채널들에 대한 크기 조절에 관련된 파라미터, 2) 피쳐 맵에 대한 양자화에 관련된 파라미터 및 3) 피쳐 맵의 하나 이상의 채널들의 배열에 관련된 파라미터 중 적어도 하나를 포함할 수 있다.
- [0118] (부호화된) 피쳐 맵에 대한 정보는 비트스트림 등을 통해 복호화 장치(1100)로 전송될 수 있고, 컴퓨터 판독-가능한 기록 매체 등에 저장될 수 있다.
- [0119] 피쳐 맵 부호화기(1030)는 피쳐 맵에 대한 정보를 포함하는 비트스트림을 생성할 수 있고, 피쳐 맵에 대한 정보를 컴퓨터 판독-가능한 기록 매체 등에 저장할 수 있다.
- [0120] 도 17은 일 실시예에 따른 피쳐 맵 복호화 방법의 흐름도이다.
- [0121] 단계(1710)에서, 피쳐 맵 복호화기(1110)는 복호화 장치(1100)로 입력된 비트스트림 또는 컴퓨터 판독-가능한 기록 매체 등에 저장된 (부호화된) 피쳐 맵에 대한 정보에 대한 복호화를 수행함으로써 변환된 피쳐 맵을 복원(reconstructed)할 수 있다.
- [0122] 피쳐 맵 복호화기(1110)는 (부호화된) 피쳐 맵에 대한 정보에 대한 복호화를 수행함으로써 (복원된) 변환된 피쳐 맵, 피쳐 맵 정보 및 피쳐 맵 변환 정보를 생성할 수 있다.
- [0123] 복호화는 압축 풀기를 포함할 수 있다.
- [0124] 피쳐 맵 정보는 복호화의 대상인 피쳐 맵에 관한 정보일 수 있다.
- [0125] 피쳐 맵 정보는 1) 피쳐 맵의 채널의 가로 크기, 2) 피쳐 맵의 채널의 세로 크기, 3) 피쳐 맵의 하나 이상의 채널들의 계수, 4) 피쳐 맵의 계층 정보 및 5) 피쳐 맵 추출 정보 중 적어도 하나를 포함할 수 있다.
- [0126] 피쳐 맵 변환 정보는 피쳐 맵에 대한 변환에서 사용된 파라미터에 관한 정보일 수 있다.
- [0127] 피쳐 맵 변환 정보는 1) 피쳐 맵의 하나 이상의 채널들에 대한 크기 조절에 관련된 파라미터, 2) 피쳐 맵에 대한 양자화에 관련된 파라미터 및 3) 피쳐 맵의 하나 이상의 채널들의 배열에 관련된 파라미터 중 적어도 하나를 포함할 수 있다.
- [0128] 단계(1720)에서, 피쳐 맵 역변환기(1120)는 (복원된) 변환된 피쳐 맵에 대한 역변환을 수행함으로써 (복원된) 피쳐 맵을 생성할 수 있다.
- [0129] 말하자면, 피쳐 맵 역변환기(1120)는 (복원된) 변환된 피쳐 맵에 대한 역변환을 수행함으로써 피쳐 맵 추출기(1010)에 의해 추출된 최초의 피쳐 맵과 유사한 형식을 갖는 (복원된) 피쳐 맵을 생성할 수 있다.
- [0130] 단계(1720)는 단계들(1721, 1722 및 1723)을 포함할 수 있다.
- [0131] 단계(1721)에서, 피쳐 맵 재배열기(1310)는 (복원된) 변환된 피쳐 맵에 대한 재배열을 수행하여 재배열된 피쳐 맵을 생성할 수 있다.
- [0132] 재배열된 피쳐 맵은 부호화 장치(1000)에서의 양자화된 피쳐 맵에 대응할 수 있다.

- [0133] 피쳐 맵 재배열기(1310)는 (복원된) 변환된 피쳐 맵의 하나 이상의 채널들을 시간적 및/또는 공간적으로 배열할 수 있다.
- [0134] 피쳐 맵 재배열기(1310)는 1) 채널들에 대한 공간적 재배열, 2) 채널들에 대한 시간적 재배열 및 3) 채널들에 대한 시공간적 재배열 중 적어도 하나를 수행할 수 있다.
- [0135] 단계(1722)에서, 피쳐 맵 역양자화기(1320)는 재배열된 피쳐 맵에 대한 역양자화를 수행하여 역양자화된 피쳐 맵을 생성할 수 있다.
- [0136] 부호화 장치(1000)에서 배열이 수행되지 않은 경우 또는 피쳐 맵 재배열기(1310)에서 재배열이 수행되지 않은 경우, 피쳐 맵 역양자화기(1320)는 (복원된) 변환된 피쳐 맵에 대한 역양자화를 수행하여 역양자화된 피쳐 맵을 생성할 수 있다.
- [0137] 단계(1723)에서, 피쳐 맵 크기 재조정기(1330)는 역양자화된 피쳐 맵의 크기를 재조정할 수 있다.
- [0138] 피쳐 맵 크기 재조정기(1330)는 피쳐 맵의 각 채널의 크기가 피쳐 맵 부호화기(1030)에서 사용되는 블록의 크기에 기반한 특정 조건을 만족하도록 피쳐 맵의 각 채널의 정보에 대한 처리를 수행할 수 있다. 처리는, 업스케일링, 다운스케일링, 패딩 및 크로핑을 포함할 수 있다.
- [0139] 특정 조건은 제1 특정 조건 및 제2 특정 조건 중 하나 이상일 수 있다.
- [0140] 제1 특정 조건은, 각 채널의 가로 크기가 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기의 정수 배가 되어야 하는 것일 수 있다. 제1 특정 조건은, 각 채널의 세로 크기가 피쳐 맵 부호화기(1030)에서 사용되는 블록의 세로 크기의 정수 배가 되어야 하는 것일 수 있다. 제1 특정 조건은, 각 채널의 가로 크기 및 세로 크기가 각각 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기 및 세로 크기의 정수 배가 되어야 하는 것일 수 있다.
- [0141] 제2 특정 조건은, 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기가 각 채널의 가로 크기의 정수 배가 되어야 하는 것일 수 있다. 제2 특정 조건은, 피쳐 맵 부호화기(1030)에서 사용되는 블록의 세로 크기가 각 채널의 세로 크기의 정수 배가 되어야 하는 것일 수 있다. 제2 특정 조건은, 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기 및 세로 크기가 각각 채널의 가로 크기 및 세로 크기의 정수 배가 되어야 하는 것일 수 있다.
- [0142] 복원된 피쳐 맵은 피쳐 맵 크기 재조정기(1330)에 의해 재조정된 크기를 갖는 역양자화된 피쳐 맵일 수 있다.
- [0143] 또는, 복원된 피쳐 맵은 역양자화된 피쳐 맵일 수 있고, 피쳐 맵 크기 재조정기(1330)는 복원된 피쳐 맵의 크기를 재조정할 수 있다.
- [0144] 복원된 피쳐 맵은 작업 수행 수단에 입력될 수 있고, 다양한 기계 작업에 사용될 수 있다.
- [0145] 피쳐 맵의 채널의 크기의 조정에 사용되는 피쳐 맵 부호화기(1030)의 블록 크기 설정
- [0146] 단계(1621) 및 단계(1723)에 관련하여 아래의 설명들이 적용될 수 있다.
- [0147] 제이펙(JPEG), 에이브이씨(AVC), 에이브이원(AV1), 에이치이브이씨(HEVC) 및 브이브이씨(VVC) 등과 같은 대부분의 영상 코딩 기술에서는, 일반적으로 구현의 용이성을 위해 실제 코딩에서 사용 가능한 블록의 최대 크기가 제한될 수 있다.
- [0148] 예를 들어, AVC에서는, 코딩 블록은 가로 x 세로의 최대 크기가 16 x 16인 휘도 샘플들을 갖는 매크로블록(MacroBlock; MB)으로 제한될 수 있다.
- [0149] 또 다른 예로, HEVC 및 VVC에서는, 코딩 블록의 최대 크기는 코딩 트리 블록(Coding Tree Block; CTB)의 크기로 제한될 수 있고, 주파수 변환 블록의 최대 크기는 최대 변환 크기(Maximum Transform Size)로 제한될 수 있다.
- [0150] HEVC 및 VVC에서는, AVC와 달리, 코딩 블록의 최대 크기 및 주파수 변환 블록의 최대 크기가 실제 입력 시퀀스 특성에 따라 부호화 장치에서 결정될 수 있다.
- [0151] 이러한 최대 크기의 블록은 부호화 과정에서의 더 높은 효율의 예측 및 변환(Transform)을 위하여, 통상적으로 더 작은 블록들로 분할(Partition)될 수 있다.
- [0152] 예를 들어, HEVC의 부호화 과정에서, 코딩 블록은 코딩 블록의 실제 크기가 CTB의 크기보다 더 작게 되도록 쿼

드트리 등의 형태로 분할될 수 있다. 또한, 변환 블록도 변환 블록의 실제 크기가 최대 변환 크기보다 더 작게 되도록 쿼드트리 등의 형태로 분할될 수 있다.

[0153] 피쳐 맵의 크기는 입력 영상의 크기에 의한 영향을 받을 수 있다. 따라서, 입력 영상의 크기를 조정함에 따라 피쳐 맵의 크기도 조정될 수 있고, 이러한 조정에 따라 특피쳐맵의 크기도 코딩 블록의 크기로 조정될 수 있다.

[0154] 예를 들면, 입력 영상의 크기가 1024 x 768과 같은 특정 크기로 조절되면, 처음으로 출력되는 피쳐 맵의 크기는 블록의 크기의 배수인 256 x 192가 될 수 있다.

[0155] 실시예들에서는, 피쳐 맵의 채널의 크기 조정에는 사용되는 입력 영상 및 피쳐 맵 부호화기(1030)에서의 블록 크기가 전송된 예들과 같은 최대 블록 크기의 이하의 크기로 설정될 수 있다.

[0156] 예를 들어, 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기가 W_B 이면, W_B 는 HEVC 및 VVC에 있어서는, CTB의 크기 또는 최대 변환 크기의 이하로 설정될 수 있다.

[0157] 예를 들어, CTB의 크기가 64x64이면, W_B 는 64 이하의 정수로 설정될 수 있다. 또는, W_B 는 $W_B = 2^k$ (k 는 1 보다 더 큰 정수)가 되도록 설정될 수 있다.

[0158] 실시예에서 W_B 를 64 이하의 값인 48 또는 24 등으로 설정하는 것에 비해, $W_B = 2^k$ (k 는 1 보다 더 큰 정수)가 되도록(즉, W_B 를 64, 32, 16, 8 및 4와 같은 2의 제곱수 중 하나가 되도록) 설정하는 것이 FPN과 같이 가로 크기 및 세로 크기가 각각 1/2씩 줄어드는 방식으로 피쳐 맵이 추출되는 경우에 대응 가능하기 때문에 더 유리할 수 있다.

[0159] 또한, W_B 는 주어진 특정 조건을 충족시키는 가장 큰 값으로 설정될 수 있다. W_B 가 주어진 특정 조건을 충족시키는 가장 큰 값으로 설정됨에 따라 부호화 성능이 최대화될 수 있다.

[0160] 피쳐 맵 채널의 크기 조정 조건

[0161] 단계(1621) 및 단계(1723)의 특정 조건에 관련하여 아래의 설명들이 적용될 수 있다.

[0162] 실시예들에 의하면, 피쳐 맵의 각 채널의 크기는 앞서 설명된 피쳐 맵 부호화기(1030)에서의 블록의 크기에 기반하는 특정 조건을 충족하도록 설정되어야 할 수 있다.

[0163] 특정 조건은, 아래의 조건 1, 조건 2, 조건 3 및 조건 4 중 하나 이상일 수 있다.

[0164] [조건 1] 각 채널의 가로 크기 및/또는 세로 크기가 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기 및/또는 세로 크기의 정수 배가 되어야 함.

[0165] [조건 2] 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기 및/또는 세로 크기가 각 채널의 가로 크기 및/또는 세로 크기의 정수 배가 되어야 함.

[0166] 예를 들어, 피쳐 맵 부호화기(1030)의 블록의 가로 크기가 W_B 이고, 피쳐 맵의 채널의 가로 크기가 W_C 이면, 상기의 조건 1) 및 조건 2)는 아래와 같은 수식을 포함하는 조건 3) 및 조건 4)로서 표현될 수 있다.

[0167] [조건 3] $W_C = m \cdot W_B$ (m 은 양의 정수)

[0168] [조건 4] $W_B = n \cdot W_C$ (n 은 양의 정수)

[0169] 예를 들어, W_B 가 64인 경우, W_C 가 256이면, $m=4$ 가 되어 조건 1은 충족되지만, $n<1$ 가 되어 조건 2는 충족되지 못할 수 있다.

[0170] W_B 가 64인 또 다른 예로, 1) W_C 가 192이면, $m=3.5$ 가 되고, $n<1$ 가 되어 제1 조건 및 제2 조건의 양자가 모두 충족되지 못할 수 있다. 2) W_C 가 32이면 $m<1$ 가 되고, $n=2$ 가 되어 조건 1은 충족되지 못할 수 있고, 조건 2는 충족될 수 있다. 3) W_C 가 64이면 조건 1 및 조건 2가 모두 충족될 수 있다.

[0171] 피쳐 맵에 대한 부호화 장치(1000)에서, 블록 분할 및 계산 효율성 등의 이유로 인해, 블록의 크기(예를 들면, 가로 길이 x 세로 길이)가, 통상적으로 $2^k \times 2^l$ 과 같은 2의 제곱수가 되도록 설정이 수행될 수 있다. k 는 1 이

상의 정수일 수 있다. 1은 1 이상의 정수일 수 있다. 예를 들면, AVC, HEVC 및 VVC와 같은 동영상 처리 방식들에서는, 2의 제곱수인 크기들을 갖는 블록들이 사용될 수 있다.

[0172] 부호화 장치(1000)의 피쳐 맵 크기 조정기(1210)에서는, 피쳐 맵의 채널의 가로 크기 및/또는 세로 크기가 전술된 조건들 중 적어도 하나를 충족시키도록 업스케일링, 다운스케일링 및 패딩 등을 이용하여 채널의 크기를 조정할 수 있다.

[0173] 예를 들어, $W_b=64$ 이고, $W_c=192$ 인 경우, 크기가 조정된 채널의 가로 크기 W'_c 를 256으로 설정한 후, 피쳐 맵 크기 조정기(1210)는 업스케일링 및/또는 패딩을 이용하여 채널의 가로 크기가 W'_c 이 되도록 조정할 수 있다.

[0174] 복호화 장치(1100)의 피쳐 맵 크기 재조정기(1330)에서는, 부호화의 과정에서 크기 조정된 피쳐 맵의 채널의 크기를 채널의 원래의 크기와 같아지도록 재조정할 수 있다.

[0175] 예를 들어, 피쳐 맵 크기 조정기(1210)에서 $W_c=192$ 인 채널의 가로 크기가 $W'_c=256$ 으로 업스케일링된 경우, 피쳐 맵 크기 재조정기(1330)에서는 $W'_c=256$ 인 채널의 가로 크기가 $W_c=192$ 로 다운스케일링될 수 있다.

[0176] 예를 들어, 피쳐 맵 크기 조정기(1210)에서 $W_c=192$ 인 채널의 가로 크기가 $W'_c=256$ 이 되도록 패딩이 수행된 경우, 피쳐 맵 크기 재조정기(1330)에서는 $W'_c=256$ 인 채널의 가로 크기가 $W_c=192$ 가 되도록 패딩된 영역이 크로핑(cropping)될 수 있다.

[0177] 피쳐 맵 채널의 크기 조정

[0178] 단계(1621) 및 단계(1723)의 크기 조정에 관련하여 아래의 설명들이 적용될 수 있다.

[0179] 도 18은 일 예에 따른 코딩 블록의 크기를 나타낸다.

[0180] 도 18의 점선은 코딩 블록의 크기를 나타낸다.

[0181] 도 19는 일 예에 따른 크기가 조정되기 전의 피쳐 맵의 채널을 나타낸다.

[0182] 도 19의 음영 부분은 크기가 조정되기 전의 피쳐 맵의 채널을 나타낸다.

[0183] 도 19의 점선 부분은 크기가 조정된 후의 피쳐 맵의 채널의 크기를 나타낸다.

[0184] 도 20은 일 예에 따른 크기가 조정된 후의 피쳐 맵의 채널을 나타낸다.

[0185] 도 20에서, 피쳐 맵의 채널의 크기는 코딩 블록의 크기와 같아지도록 조정되었다.

[0186] 도 18, 도 19 및 도 20에서 도시된 예를 참조하면, 피쳐 맵 부호화기(1030)에서, 블록의 가로 크기 및 세로 크기가 각각 $W_b=64$ 및 $H_b=64$ 로 설정된 경우, 피쳐 맵의 채널의 가로 크기 및 세로 크기가 각각 $W_c=96$ 및 $H_c=48$ 일 때, 실시예들에서는 가로 크기 및 세로 크기가 각각 $W'_c=2 \cdot W_b=128$ 및 $H'_c=H_b=64$ 가 되도록 피쳐 맵의 채널의 크기가 조정될 수 있다.

[0187] 상기의 채널의 크기의 조정에서는, 채널의 크기가 더 커지도록 조정되어야 하므로, 업샘플링 및/또는 패딩 등의 방법을 이용하여 채널의 크기가 더 커질 수 있다.

[0188] 도 18, 도 19 및 도 20에서는 상기의 실시예에 의한 1) 코딩 블록의 크기, 2) 크기가 조정되기 전의 피쳐 맵의 채널의 크기 및 3) 크기가 조정된 후의 피쳐 맵의 채널의 크기가 각각 도시되었다.

[0189] 만약, 채널의 크기가 더 작아지도록 조정되는 경우에는, 다운샘플링 또는 크로핑 등의 방법을 이용하여 크기가 더 작아질 수 있다.

[0190] 업샘플링 및 다운샘플링에 의한 크기 조정 및 크기 재조정

[0191] 단계(1621) 및 단계(1723)의 크기 조정에 관련하여 아래의 설명들이 적용될 수 있다.

[0192] 전술된 실시예에서, 크기 조정 전의 채널은 96x48의 크기를 가질 수 있고, 크기 조정 후의 채널은 128x64의 크기가 되어야 할 수 있다. 따라서, 가로 및 세로의 모두가 4/3배로 업샘플링 되어야 할 수 있다.

[0193] 업샘플링은, 바이리니어(Bilinear) 방식, 바이큐빅(Bicubic) 방식, 랜스조스(Lanczos) 방식, HEVC 보간 필터, VVC 보간 필터 및 딥러닝에 의한 방식과 같이 등 다양한 방식들을 사용할 수 있으며, 가급적 단순한 방식을 사

용하는 것이 계산량의 측면에서 유리할 수 있다.

- [0194] 피쳐 맵에 대한 부호화에서 업샘플링된 채널은, 복호화에서 다시 다운샘플링됨으로써 원래의 크기와 동일한 크기로 재조정될 수 있다.
- [0195] 즉, 전술된 실시 예에서, 크기 조정 후 채널은 128x64의 크기를 가질 수 있으며, 원래의 채널의 크기는 96x48이므로, 가로 및 세로의 모두가 3/4배로 다운샘플링 되어야 할 수 있다.
- [0196] 다운샘플링은, 표본화, 저역통과 표본화, 스케일러블 비디오 코딩(Scalable Video Coding; SVC) 및 다운샘플링 필터 등 다양한 방식들을 사용할 수 있으며, 가급적 단순한 방식을 사용하는 것이 계산량 측면에서 유리할 수 있다.
- [0197] 전술된 실시예와는 반대로, 채널의 크기가 더 작아지도록 조정되는 경우, 피쳐 맵에 대한 부호화에서 다운샘플링에 의한 크기의 조정이 수행될 수 있고, 피쳐 맵에 대한 복호화에서 업샘플링을 통한 크기의 재조정이 수행될 수 있다.
- [0198] 패딩 및/또는 크로핑에 의한 크기의 조정 및 재조정
- [0199] 단계(1621) 및 단계(1723)의 크기 조정에 관련하여 아래의 설명들이 적용될 수 있다.
- [0200] 전술된 실시예에서, 크기가 조정되기 전의 채널은 96x48의 크기를 가질 수 있고, 크기가 조정된 후의 채널은 128x64의 크기가 되어야 할 수 있다. 따라서, 가로 방향으로 32 개의 샘플들의 피쳐 값들이 추가로 할당될 수 있고, 세로 방향으로 16 개의 샘플들의 피쳐 값들이 추가로 할당될 수 있다.
- [0201] 추가되는 샘플들에 대한 피쳐 값의 할당은 패딩 방법을 사용하여 도출될 수 있다. 패딩 방법은 동일 값 반복, 평균 값 채우기, 거울(mirror) 패딩 및 보간(interpolation) 등 다양한 방법들을 포함할 수 있다.
- [0202] 추가되는 샘플의 위치는 크기의 조정의 전의 채널 및 크기의 조정의 의 채널 간의 위치 관계를 어떻게 설정하느냐에 따라서 변할 수 있다.
- [0203] 도 21은 일 예에 따른 크기의 조정의 전 및 크기의 조정의 후의 피쳐 맵의 채널의 정보의 위치들 간의 관계를 나타낸다.
- [0204] 도 21에서 도시된 것과 같이, 크기의 조정의 전의 채널(도 21에서 음영으로 표시된 부분)이 크기의 조정의 후에 채널의 안쪽에 위치하도록 설정될 수 있다.
- [0205] 이러한 경우, 도 21에서 화살표로서 도시된 것과 같이, 크기의 조정의 전의 채널의 경계 부근에서의 샘플 값들(즉, 경계 부근 피쳐 값들)을 이용하여 샘플 값이 할당되지 않은 샘플들에 대한 패딩이 수행될 수 있다.
- [0206] 도 22는 다른 일 예에 따른 크기의 조정의 전 및 크기의 조정의 후의 피쳐 맵의 채널의 정보의 위치들 간의 관계를 나타낸다.
- [0207] 도 22에서 도시된 것과 같이, 크기의 조정의 전의 채널의 경계가 크기의 조정의 후의 채널의 경계와 일부 겹칠 수 있다. 크기의 조정의 전의 채널의 경계가 크기의 조정의 후의 채널의 경계와 일부 겹치도록 채널들 간의 위치 관계가 설정될 수 있다.
- [0208] 이러한 경우, 도 22에서 화살표로서 도시된 것과 같이, 경계들이 서로 겹치지 않는 방향으로만 패딩이 수행될 수 있다.
- [0209] 도 23은 또 일 예에 따른 크기의 조정의 전 및 크기의 조정의 후의 피쳐 맵의 채널의 정보의 위치들 간의 관계를 나타낸다.
- [0210] 도 23에서 도시된 것과 같이, 크기의 조정의 전의 채널의 정보는 4 개로 분할될 수 있다. 분할된 채널들의 각 채널의 경계 위치가 크기가 조정된 후의 채널의 경계 위치가 일부 겹치도록 분할된 채널들 및 크기가 조정된 후의 채널 간의 위치 관계가 설정될 수 있다.
- [0211] 이러한 경우, 도 23에서 화살표로서 도시된 것과 같이, 패딩에 있어서 크기의 조정의 전의 채널의 양 방향들의 경계들의 부분의 샘플 값들을 이용될 수 있어, 선형 보간 기법 및 큐빅 보간 기법 등과 같이 원래의 정보가 더 잘 유지되는 패딩 방법이 적용될 수 있다. 여기에서, 부분의 샘플 값들은 경계에 인접한 샘플 값들 및 경계로부터의 거리가 임계 값의 이하인 샘플 값들을 의미할 수 있다.
- [0212] 또한, 도 23에서 사선으로 표시된 영역에 대해서는, 4 개의 방향들에서의 샘플들의 정보들을 이용하여 패딩이 수

행될 수 있다.

- [0213] 피쳐 맵에 대한 부호화에서 패딩이 적용되었던 채널은 복호화에서는 원래의 크기와 동일한 크기로 재조정될 수 있다. 재조정에 있어서, 크로핑을 통해 피쳐 맵에 패딩되었던 샘플들은 버려지고, 원래의 피쳐 맵의 채널에 해당되는 샘플들만이 추출될 수 있다.
- [0214] 만약, 도 23에서와 같이, 크기의 조정의 과정에서 피쳐 맵의 채널이 분할된 경우, 크기의 재조정에서의 크로핑의 결과(말하자면, 분할된 채널들)이 다시 하나로 합쳐서 복원하는 과정이 추가로 수행되어야 할 수 있다.
- [0215] 전술된 예와는 반대로, 채널의 크기가 더 작아지도록 크기가 조정되는 경우, 피쳐 맵에 대한 부호화에서 크로핑에 의한 채널의 크기의 조정이 수행될 수 있고, 피쳐 맵에 대한 복호화에서 패딩을 통한 채널의 크기의 재조정이 수행될 수 있다.
- [0216] 업샘플링/다운샘플링 방식과 패딩/크로핑 방식을 동시에 사용하는 크기 조정 및 재조정 방법
- [0217] 단계(1621) 및 단계(1723)의 크기 조정에 관련하여 아래의 설명들이 적용될 수 있다.
- [0218] 통상 크기 조정 전과 후의 채널 크기 차이가 큰 경우에는 업/다운샘플링 방식의 크기 조정이 유리하고, 크기 차이가 작은 경우에는 패딩/크로핑 방식의 크기 조정이 유리한 경우가 많다.
- [0219] 두 가지 방식을 조합하여 성능 향상을 도모할 수 있다. 즉, 먼저 패딩/크로핑을 통해 1차적으로 크기 조정을 수행한 다음 업샘플링/다운샘플링을 수행하여 최종적인 크기 조정을 수행하거나, 반대로 업/다운샘플링을 통해 1차적으로 크기 조정을 수행한 다음 패딩/크로핑을 통해 최종적인 크기 조정을 수행할 수 있다. 이때 크기 재조정은 크기 조정시 사용되었던 순서와 반대로 수행하면 된다.
- [0220] 다계층 피쳐 맵 채널의 크기 조정
- [0221] 단계(1621) 및 단계(1723)의 크기 조정에 관련하여 아래의 설명들이 적용될 수 있다.
- [0222] 복수의 계층들을 갖는 피쳐 맵에 대한 부호화 및/또는 복호화를 수행하는 경우에는, 다수의 계층들의 각 계층별로 실시예들의 크기의 조정 및 재조정이 적용될 수 있다.
- [0223] 각 계층별로 독립적인 크기의 조정 및/또는 재조정을 수행하지 않고, 복수의 계층들에 대해 동일한 방식이 적용될 수 있다. 이러한 동일한 방식의 적용은 것은 계산 복잡도를 낮출 수 있다.
- [0224] 예를 들어, 마스크 R-CNN의 FPN과 같이, 추출된 피쳐 맵의 계층의 깊이가 하나씩 더 깊어질 수록, 피쳐 맵 채널의 가로 크기 및 세로 크기가 각각 1/2배로 줄어드는 경우에는, 가장 얇은 계층의 채널(즉, 가장 큰 채널)의 크기 W_c 및 H_c 에 기반하여 조정될 크기 W'_c 와 H'_c 가 도출될 수 있고, 이후의 계층들에 대해서는 W'_c 및 H'_c 를 각각 1/2배로 줄이는 방식에 따라 조정될 크기가 도출될 수 있다.
- [0225] 이러한 경우에는, 가장 큰 채널의 조정될 크기 W'_c 및 H'_c 가 256, 128, 64, 32 및 16 등과 같이 2의 제곱수가 되도록 설정될 수 있다. 이러한 설정에 따라 모든 계층들에 대한 크기의 조정의 결과가 피쳐 맵 부호화기(1030)에서의 블록 분할의 경계와 정렬될 가능성이 높아질 수 있으며, 따라서 이러한 설정은 부호화 성능 향상에 있어서 유리할 수 있다.
- [0226] 입력 영상의 크기의 조정에 의한 피쳐 맵 채널의 크기의 조정
- [0227] 도 24는 일 실시예에 따른 피쳐 맵 채널들의 크기의 조정을 나타낸다.
- [0228] 단계(1621) 및 단계(1723)의 크기 조정에 관련하여 아래의 설명들이 적용될 수 있다.
- [0229] 실시예들에 의하면 피쳐 맵의 각 채널의 크기는 전술된 피쳐 맵 부호화기(1030)에서 사용되는 블록의 크기에 기반하는 특정 조건을 충족하도록 설정되어야 할 수 있다.
- [0230] 특정 조건은 제1 특정 조건 및 제2 특정 조건 중 하나 이상일 수 있다.
- [0231] 제1 특정 조건은, 각 채널의 가로 크기가 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기의 정수 배가 되어야 하는 것일 수 있다. 제1 특정 조건은, 각 채널의 세로 크기가 피쳐 맵 부호화기(1030)에서 사용되는 블록의 세로 크기의 정수 배가 되어야 하는 것일 수 있다. 제1 특정 조건은, 각 채널의 가로 크기 및 세로 크기가 각각 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기 및 세로 크기의 정수 배가 되어야 하는 것일 수 있다.

- [0232] 제2 특정 조건은, 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기가 각 채널의 가로 크기의 정수 배가 되어야 하는 것일 수 있다. 제2 특정 조건은, 피쳐 맵 부호화기(1030)에서 사용되는 블록의 세로 크기가 각 채널의 세로 크기의 정수 배가 되어야 하는 것일 수 있다. 제2 특정 조건은, 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기 및 세로 크기가 각각 채널의 가로 크기 및 세로 크기의 정수 배가 되어야 하는 것일 수 있다.
- [0233] 이러한 피쳐 맵 채널의 크기의 조정은 전술된 실시예들에서와 같이 피쳐 맵에 대해 수행될 수도 있지만, 입력 영상 자체에 대해 수행될 수도 있다. 즉, 입력 영상으로부터 추출된 피쳐 맵의 각 채널의 크기가 전술된 제1 특정 조건 및/또는 제2 특정 조건을 충족시키도록 입력 영상의 크기가 조정될 수 있다.
- [0234] 추출된 피쳐 맵의 가로 크기 또는 세로 크기가 입력 영상의 가로 또는 세로 크기에 비해 $1/(2^{\text{의 제공수}})$ (예를 들면, $1/2$, $1/4$ 및 $1/8$ 등)로 줄어 드는 경우를 예로 하여, 입력 영상의 크기의 조정에 의한 피쳐 맵 채널의 크기의 조정 과정이 아래에서 보다 상세하게 설명된다.
- [0235] 전술된 피쳐 맵의 크기에 대한 제1 특정 조건 및 제2 특정 조건은 아래와 같이 입력 영상의 크기에 대한 변형된 제1 특정 조건 및 변형된 제2 특정 조건으로 변경될 수 있다. 실시예들에서는, 제1 특정 조건 및 제2 특정 조건을 대신하여 변형된 제1 특정 조건 및 변형된 제2 특정 조건이 적용될 수 있다.
- [0236] 변형된 제1 특정 조건은, 각 입력 영상의 가로 크기가 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기의 정수 배가 되어야 하는 것일 수 있다. 변형된 제1 특정 조건은, 각 입력 영상의 세로 크기가 피쳐 맵 부호화기(1030)에서 사용되는 블록의 세로 크기의 정수 배가 되어야 하는 것일 수 있다. 변형된 제1 특정 조건은, 각 입력 영상의 가로 크기 및 세로 크기가 각각 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기 및 세로 크기의 정수 배가 되어야 하는 것일 수 있다.
- [0237] 변형된 제2 특정 조건은, 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기가 각 입력 영상의 가로 크기의 정수 배가 되어야 하는 것일 수 있다. 변형된 제2 특정 조건은, 피쳐 맵 부호화기(1030)에서 사용되는 블록의 세로 크기가 각 입력 영상의 세로 크기의 정수 배가 되어야 하는 것일 수 있다. 변형된 제2 특정 조건은, 피쳐 맵 부호화기(1030)에서 사용되는 블록의 가로 크기 및 세로 크기가 각각 입력 영상의 가로 크기 및 세로 크기의 정수 배가 되어야 하는 것일 수 있다.
- [0238] 입력 영상의 가로 크기가 W 이고, 세로 크기가 H 이고, 피쳐 맵 채널의 가로 크기 및 세로 크기가 2의 제공수가 되도록 피쳐 맵 채널에 대한 크기의 조정이 수행될 때, 크기가 조정된 입력 영상의 가로 크기 W' 및 세로 크기 H' 는 아래의 수식 1 및 수식 2에 의해 획득될 수 있다.
- [0239] [수식 3]
- [0240] $W' = 2^k$
- [0241] 여기에서, 선택된 k 가 $2^{k-1} < W \leq 2^k$ 를 충족시키는 자연수이면, 피쳐 맵 채널의 가로 크기가 더 커지도록 피쳐 맵 채널의 크기에 대한 조정이 수행될 수 있다.
- [0242] 선택된 k 가 $2^k \leq W < 2^{k+1}$ 을 충족시키는 자연수이면, 피쳐 맵 채널의 가로 크기가 더 작아지도록 피쳐 맵 채널의 크기에 대한 조정이 수행될 수 있다.
- [0243] [수식 4]
- [0244] $H' = 2^l$
- [0245] 여기에서, 선택된 l 이 $2^{l-1} < H \leq 2^l$ 을 충족시키는 자연수이면, 피쳐 맵 채널의 세로 크기가 더 커지도록 피쳐 맵 채널의 크기에 대한 조정이 수행될 수 있다.
- [0246] 선택된 l 이 $2^l \leq H < 2^{l+1}$ 을 충족시키는 자연수이면, 피쳐 맵 채널의 세로 크기가 더 작아지도록 피쳐 맵 채널의 크기에 대한 조정이 수행될 수 있다.
- [0247] 예를 들면, W 가 640이고, H 가 480인 경우, W' 는 2의 제공수인 1024가 될 수 있고, H' 는 2의 제공수인 512가 될 수 있다. 따라서, 추출된 피쳐 맵의 크기도 2의 제공수가 됨에 따라, 피쳐 맵의 크기가 피쳐 맵 부호화기(1030)에서 사용되는 블록의 크기의 배수가 되는 조건이 충족될 수 있다.

- [0248] 전술된 입력 영상에 대한 크기의 조정 및 피쳐 맵 채널에 대한 크기의 조정은 서로 조합하여 사용될 수도 있다.
- [0249] 가로 크기에 대한 조정 및 세로 크기에 대한 조정은 서로 다른 대상들에 대해 각각 수행될 수 있다.
- [0250] 예를 들면, 가로 크기에 대한 조정은 입력 영상에 대해 수행될 수 있고, 세로 크기에 대한 조정은 피쳐 맵 채널에 대해 수행될 수 있다. 반대로, 가로 크기에 대한 조정은 피쳐 맵 채널에 대해 수행될 수 있고, 세로 크기에 대한 조정은 입력 영상에 대해 수행될 수 있다.
- [0251] 입력 영상에 대한 크기 조정과 피쳐 맵 채널에 대한 크기 조정의 조합
- [0252] 추출된 피쳐 맵의 가로 크기 및 세로 크기가 입력 영상의 가로 크기 및 세로 크기에 비해 각각 1/4로 줄어드는 경우, 후술될 것과 같이 가로 크기에 대한 조정은 입력 영상에 대해 수행되고, 세로 크기에 대한 조정은 피쳐 맵 채널에 대해 수행될 수 있다.
- [0253] 입력 영상의 가로 크기가 W 이고, 세로 크기가 H 이면, 가로 크기가 조정된 입력 영상의 가로 크기는 W' 일 수 있고, 가로 크기가 조정된 입력 영상의 크기는 H' 일 수 있다. (여기에서, H 및 H' 는 같을 수 있다.)
- [0254] 가로 크기가 조정된 입력 영상으로부터 추출된 피쳐의 크기는 $P' \times Q'$ 일 수 있다. 상기의 추출된 피쳐에 대해 세로 크기의 조정이 수행됨으로써 생성되는 세로 크기가 조정된 피쳐의 크기는 $P'' \times Q''$ 일 수 있다. (여기에서, P' 및 P'' 는 같을 수 있다)
- [0255] W 가 640이고, H 가 480인 경우, 전술된 입력 영상의 가로 크기에 대한 조정에 따라, W' 는 1024가 되고, H' 는 480이 될 수 있다. W' 및 H' 의 크기를 갖는 (가로 크기가 조정된) 입력 영상으로부터 추출된 피쳐의 가로 크기 P' 는 256일 수 있고, 세로 크기 Q' 는 120이 될 수 있다. $P' \times Q'$ 의 크기를 갖는 피쳐의 세로 크기에 대한 조정을 수행함으로써 생성된 (세로 크기가 조정된) 피쳐의 가로 크기 P'' 는 256일 수 있고, 세로 크기 Q'' 는 128일 수 있다.
- [0256] 복원된 피쳐 맵을 부호화의 이전의 형식으로 역변환하는 피쳐 맵 복호화기(1110)에 의한 복원된 피쳐 맵의 크기에 대한 재조정
- [0257] 복호화 장치(1100)에 의해 생성된 복원된 피쳐가 사람 또는 기계가 소비하기 위해서는 복원된 피쳐가 원래의 피쳐의 크기와 같은 크기를 갖도록 변환되어야 할 수 있다.
- [0258] 피쳐 맵 복호화기(1110)에 의해 생성된 복원된 피쳐 맵에 대한 재배열은_피쳐 맵의 하나 이상의 채널들에 대한 크기 조정에 관련된 파라미터를 사용하여 수행될 수 있다.
- [0259] 여기에서, 크기 조정에 관련된 파라미터는 전술된 크기가 조정된 피쳐의 가로 크기 P'' 및 세로 크기 Q'' 를 포함할 수 있다.
- [0260] 또는, 크기 조정에 관련된 파라미터는 1) 가로 방향으로 피쳐 맵을 구성하는 피쳐들의 개수 K_w 및 2) 세로 방향으로 피쳐 맵을 구성하는 피쳐들의 개수 K_h 를 포함할 수 있다.
- [0261] 예를 들면, 피쳐 맵이 가로 방향으로 16 개의 피쳐들을 갖고(말하자면, $K_w = 16$), 세로 방향으로 16 개의 피쳐들을 갖도록(말하자면, $K_h = 16$) 재배열된 경우, 피쳐의 가로 크기는 피쳐 맵의 가로 크기가 K_w 로 나뉘어진 결과인 P'' 일 수 있고, 피쳐의 세로 크기가 K_h 로 나뉘어진 결과인 Q'' 일 수 있다.
- [0262] 전술된 과정에 의해 결정된 재배열된 피쳐의 가로 크기 및 세로 크기는 각각 P' 및 Q' 로 변환되어야 할 수 있다. 이 때, P' 및 Q' 는 크기로 변환해야 한다. 이때, P' , Q' 의 값은 피쳐 맵 정보의 피쳐 맵 채널의 가로 크기 및 피쳐 맵 채널의 세로 크기를 통해 부호화 장치(1000)로부터 복호화 장치(1100)로 전달될 수 있다.
- [0263] 피쳐 값의 양자화 및 역양자화
- [0264] 단계(1622) 및 단계(1722)의 크기 조정에 관련하여 아래의 설명들이 적용될 수 있다.
- [0265] 피쳐의 피쳐 값의 양자화에 대해서는 균등 양자화 또는 비균등 양자화가 사용될 수 있으며, 피쳐 값은 n -비트 정수로 양자화 될 수 있다.
- [0266] 아래의 수식 3은 피쳐 값의 n -비트 균등 양자화 과정을 나타낼 수 있다.

[0267] [수식 3]

$$F_{Converted} = Round((F_{Original}) \times (2^n - 1) / Range_{max} + 2^{n-1})$$

$$IF, F_{Converted} > 2^{n-1}, \quad F_{Converted} = 2^{n-1}$$

$$F_{Converted} < 0, \quad F_{Converted} = 0$$

[0269] 아래의 수식 4는 양자화 후 복원된 피쳐 값 $F_{Original}$ 의 균등 역양자화 과정을 나타낼 수 있다.

[0270] [수식 4]

$$\hat{F} = ((\tilde{F}_{Converted} - 2^{n-1}) / (2^n - 1)) \times Range_{max}$$

[0272] 양자화에서는, 수식 3에서와 같이 정수화 과정에서 오차를 감소시키기 위해 반올림(rounding)이 적용될 수 있으며, n-비트 정수의 범위를 넘는 값들을 클리핑하는 과정이 포함될 수 있다.

[0273] 양자화 및 역양자화를 위해서는, 양자화의 결과의 비트들의 개수 n의 값 및 피쳐 값의 범위 $Range_{max}$ 의 값이 설정되어야 할 수 있다.

[0274] n의 값은 8 또는 10과 같이 통상의 영상 부호화 장치에 입력될 수 있는 값으로 설정될 수 있고, 압축 성능 향상을 위해 4 또는 6과 같은 더 작은 값으로 설정될 수도 있다.

[0275] $Range_{max}$ 의 값은 부호화될 피쳐 값들의 최대 값 및 최소 값 간의 차이일 수 있고, 기정의된 값일 수도 있다.

[0276] n의 값 및 $Range_{max}$ 의 값은 피쳐 값에 대한 양자화에 관련된 파라미터로서, 부호화 장치(1000)에서 사용된 값들이 복호화 장치(1100)에서도 동일하게 사용되어야만 할 수 있다. 따라서, n의 값 및 $Range_{max}$ 의 값은 부호화 장치(1000) 및 복호화 장치(1100)에서 동일하기 미리 정의될 수 있다. 또는, n의 값 및 $Range_{max}$ 의 값은 피쳐 맵 부호화기(1030)를 통해 비트스트림 또는 컴퓨터 판독-가능한 기록 매체 등에 포함될 수 있고, 비트스트림 또는 컴퓨터 판독-가능한 기록 매체 등을 통해 복호화 장치(1100)의 피쳐 맵 복호화기(1110)로 전달될 수 있다.

[0277] 피쳐 값에 대한 양자화는 피쳐 맵의 크기의 조정 이전에 수행될 수도 있고, 이후에 수행될 수도 있다. 계산 복잡도를 줄이는 측면에서는, 양자화가 크기의 조정보다 더 먼저 수행될 수 있다. 크기의 조정의 과정에서의 계산 오차를 줄이는 측면에서는, 양자화가 크기의 조정의 이후에 수행될 수 있다.

[0278] 피쳐 값에 대한 역양자화도 피쳐 맵의 크기의 재조정 이전에 수행될 수도 있고, 이후에 수행될 수도 있다. 계산 복잡도를 줄이는 측면에서는, 역양자화가 크기의 재조정보다 더 먼저 수행될 수 있다. 크기의 재조정의 과정에서의 계산 오차를 줄이는 측면에서는, 역양자화가 크기의 재조정의 이후에 수행될 수 있다.

[0279] 피쳐 맵의 채널들의 배열 및 재배열

[0280] 단계(1623) 및 단계(1721)의 배열 및 재배열에 관련하여 아래의 설명들이 적용될 수 있다.

[0281] 도 25는 일 예에 따른 피쳐 맵의 구성을 나타낸다.

[0282] 인공 신경망으로부터 추출된 피쳐 맵은 통상적으로 도 25에서 도시된 것과 같이 복수의 채널들로 구성될 수 있다.

[0283] 도 25에서, N_c 는 채널들의 개수를 나타낼 수 있다. W_c 는 채널의 가로 크기를 나타낼 수 있다. H_c 는 채널의 세로 크기를 나타낼 수 있다.

[0284] 복수의 개의 채널을 부호화하기 위해, 복수의 채널들은 부호화에 대한 입력 단위인 프레임(또는, 픽처)의 형태로 변환될 수 있다.

[0285] 피쳐 맵 배열기(1230)는 복수의 채널들을 프레임으로 변환함에 있어서, 공간적(Spatial) 배열 방식, 시간적

(Temporal) 배열 방식 및 시공간적(Spatio-Temporal) 배열 방식 중 하나를 사용할 수 있다.

- [0286] 이러한 방식으로 변환된 피쳐 맵의 프레임은 부호화 장치(1000)에 의해 부호화(또는, 압축)될 수 있고, 복호화 장치(1100)에 의해 복호화(또는, 복원)될 수 있다.
- [0287] 피쳐 맵 재배열기(1310)는 복원된 피쳐 맵을 나타내는 복원된 프레임을 원래의 채널의 구성의 형태로 재배열할 수 있다.
- [0288] 도 26는 일 예에 따른 공간적으로 배열된 피쳐 맵을 나타낸다.
- [0289] 도 26에서는 피쳐 맵에 대한 공간적 배열 방식이 도시되었다.
- [0290] 도 26에서 도시된 것과 같이, 공간적 배열은 피쳐 맵의 채널들을 마치 타일과 같은 형태로 가로로 m 개 및 세로로 n 개씩 배열하여 하나의 피쳐 맵 프레임을 구성하는 것을 의미할 수 있다. m 및 n 은 m 및 n 의 곱이 배열될 채널들의 총 개수 N_c 와 동일하도록 설정될 수 있다.
- [0291] 공간적으로 배열된 피쳐 맵의 프레임은 부호화 장치(1000)에서 인트라 예측(Intra Prediction)을 이용하여 부호화될 수 있다.
- [0292] 도 27는 일 예에 따른 시간적으로 배열된 피쳐 맵을 나타낸다.
- [0293] 도 27에서는 피쳐 맵에 대한 시간적 배열 방식이 도시되었다.
- [0294] 도 27에서 도시된 것과 같이, 시간적 배열은 피쳐 맵의 채널들의 각 프레임이 하나의 프레임이 되도록 시간적으로 나열되는 것을 의미할 수 있다.
- [0295] 프레임들의 개수를 설명함에 있어서, 하나의 피쳐 맵에 대한 프레임들의 개수는 상기의 피쳐 맵을 구성하는 채널들의 총 개수 N_c 와 동일하게 되도록 설정될 수 있다.
- [0296] 시간적으로 배열된 피쳐 맵의 프레임들은 부호화 장치(100)에서 인터 예측(Inter Prediction)을 이용하여 부호화될 수 있다.
- [0297] 실시예에서는, 이해를 돕기 위해, 영상 코딩 표준에서 널리 사용되는 용어인 인터 예측이 사용되었으나, 통상적인 인터 예측과의 구분을 위해서 전술된 인터 예측은 채널-간 예측(Inter-Channel Prediction)으로 명명될 수 있다.
- [0298] 도 28은 일 예에 따른 시간적으로 배열된 피쳐 맵을 나타낸다.
- [0299] 도 28에서는 피쳐 맵에 대한 시간적 배열 방식이 도시되었다.
- [0300] 도 28에서 도시된 것과 같이, 시공간적 배열은 공간적으로 배열된 프레임들을 시간적으로 배열하는 것일 수 있다.
- [0301] 시공간적 배열에 있어서, 하나의 프레임을 구성하는 채널들의 개수인 $m \times n$ 및 프레임들의 총 개수의 곱이 배열될 채널들의 총 개수 N_c 와 동일하게 되도록 설정이 수행될 수 있다.
- [0302] 피쳐 맵 정보
- [0303] 단계(1630) 및 단계(1710)에서의 피쳐 맵 정보에 관련하여 아래의 설명들이 적용될 수 있다.
- [0304] 피쳐 맵 추출 정보는 1) 피쳐 맵의 채널의 크기, 2) 채널들의 개수, 3) 계층들의 개수 및 4) 피쳐 맵이 어떤 신경망 모델에서 추출된 것인가를 식별할 수 있게 하는 정보 등을 포함할 수 있다.
- [0305] 이러한 피쳐 맵 추출 정보가 복호화 장치(1100)에서의 복호화의 이후에 수행되는 기계 작업의 수행 수단에서도 알려져야 적절한 작업 수행이 이루어질 수 있다.
- [0306] 따라서, 피쳐 맵 추출 정보는 미리 정해질 수 있으며, 피쳐 맵 부호화기(1030)에 의해 부호화된 피쳐 맵 추출 정보로 부호화될 수 있다. 부호화된 피쳐 맵 추출 정보는 피쳐 맵 복호화기(1110)에서 (복원된) 피쳐 맵 추출 정보로 복호화될 수 있다. (복원된) 피쳐 맵 추출 정보는 기계 작업의 수행 수단으로 전달될 수 있다.
- [0307] 피쳐 맵 변환 정보
- [0308] 단계(1630) 및 단계(1710)에서의 피쳐 맵 변환 정보에 관련하여 아래의 설명들이 적용될 수 있다.

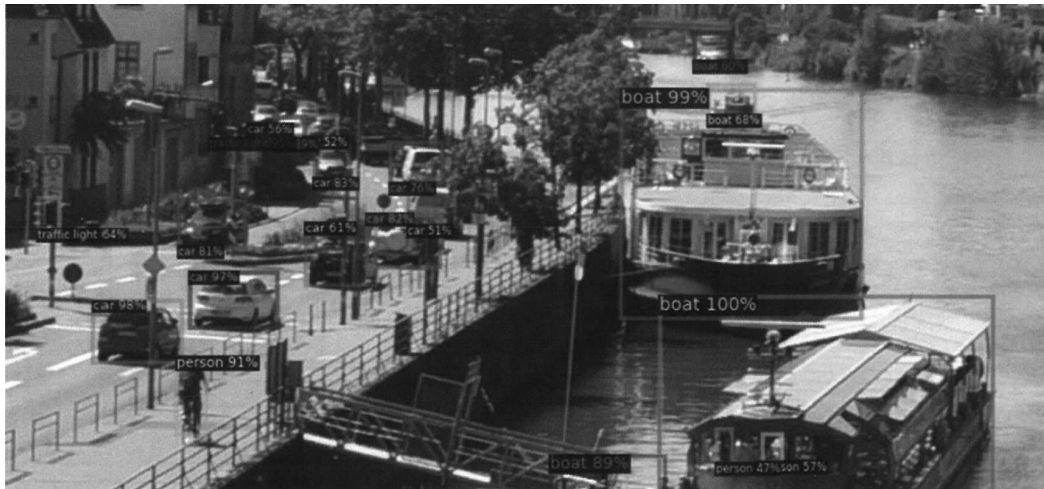
- [0309] 채널의 크기의 재조정을 통해 원래의 채널의 크기 및 모양으로의 복원을 수행하기 위해서, 피쳐 맵의 채널의 크기의 조정에서 사용되는 파라미터들이 부호화 장치(1000)로부터 피쳐 맵 크기 재조정기(1330)로 전달될 수 있다.
- [0310] 여기에서, 피쳐 맵의 채널의 크기의 조정에서 사용되는 파라미터들은 1) 업샘플링 및 다운샘플링 중 어떤 방식으로 크기의 조정이 이루어졌는가를 나타내는 정보, 2) 패딩 및 크로핑 중 어떤 방식으로 크기의 조정이 이루어졌는가를 나타내는 정보 및 3) 만약 패딩 또는 크로핑을 통해 크기의 조정이 이루어진 경우, 크기의 조정의 전의 채널 및 크기의 조정의 후의 채널 간의 위치 관계를 나타내는 정보 등을 포함할 수 있다.
- [0311] 피쳐 값에 대한 양자화에 사용되는 비트들의 개수 n 및 피쳐 값의 범위 $Range_{max}$ 도 부호화 장치(1000)로부터 피쳐 맵 역양자화기(1320)로 전달될 수 있다.
- [0312] 채널들의 재배열을 통해 원래의 채널의 구성으로의 복원을 수행하기 위해서, 피쳐 맵의 채널들에 대한 재배열에서 사용되는 파라미터들이 부호화 장치(1000)로부터 피쳐 맵 재배열기(1310)로 전달될 수 있다.
- [0313] 여기에서, 피쳐 맵의 채널들에 대한 재배열에서 사용되는 파라미터들은 1) 피쳐 맵의 채널들의 배열이 공간적 배열, 시간적 배열 및 시공간적 배열 중 어떤 배열로 이루어졌는가를 나타내는 정보 및 2) 공간적 배열 또는 시공간적 배열이 사용된 경우, 하나의 프레임을 구성하는 가로 채널들의 개수와 세로 채널들의 개수를 포함할 수 있다.
- [0314] 따라서, 피쳐 맵 변환 정보는 미리 정해질 수 있으며, 피쳐 맵 부호화기(1030)에 의해 부호화된 피쳐 맵 변환 정보로 부호화될 수 있다. 부호화된 피쳐 맵 변환 정보는 피쳐 맵 복호화기(1110)에서 (복원된) 피쳐 맵 변환 정보로 복호화될 수 있다. (복원된) 피쳐 맵 변환 정보는 기계 작업의 수행 수단으로 전달될 수 있다.
- [0315] 상기의 실시예들은 부호화 장치(1000) 및 복호화 장치(1100)에서 동일한 방법 및/또는 상응하는 방법으로 수행될 수 있다. 또한, 부호화 및/또는 복호화에 있어서 상기의 실시예들 중 하나 이상의 조합이 사용될 수 있다.
- [0316] 상기의 실시예들이 적용되는 순서는 부호화 장치(1000) 및 복호화 장치(1100)에서 서로 상이할 수 있다. 또는, 상기의 실시예들이 적용되는 순서는 부호화 장치(1000) 및 복호화 장치(1100)에서 (적어도 부분적으로) 동일할 수 있다.
- [0317] 상기의 실시예들은 루마 신호 및 크로마 신호의 각각에 대하여 수행될 수 있다. 루마 신호 및 크로마 신호에 대하여 상기의 실시예들이 동일하게 수행할 수 있다.
- [0318] 상기의 실시예들이 적용되는 블록의 형태는 정방형(square) 형태 또는 비정방형(non-square) 형태를 가질 수 있다.
- [0319] 상술된 실시예들에서, 특정된 대상에 특정된 처리를 적용함에 있어서, 특정된 조건이 요구될 수 있으며, 특정된 결정 하에 상기의 특정된 처리가 처리되는 것으로 설명된 경우, 특정된 코딩 파라미터에 기반하여 특정된 조건이 충족되는지 여부가 결정되거나, 특정된 코딩 파라미터에 기반하여 특정된 결정이 이루어지는 것으로 설명되었으면, 상기의 특정된 코딩 파라미터는 다른 코딩 파라미터로 대체될 수 있는 것으로 해석될 수 있다. 말하자면, 특정된 조건 또는 특정된 결정에 영향을 미치는 코딩 파라미터는 단지 예시적인 것으로 간주될 수 있으며, 명시된 코딩 파라미터 외에도 하나 이상의 다른 코딩 파라미터들의 결합이 상기의 명시된 코딩 파라미터의 역할을 수행하는 것으로 이해될 수 있다.
- [0320] 상술된 실시예들에서, 방법들은 일련의 단계 또는 유닛으로서 순서도를 기초로 설명되고 있으나, 본 발명은 단계들의 순서에 한정되는 것은 아니며, 어떤 단계는 상술한 바와 다른 단계와 다른 순서로 또는 동시에 발생할 수 있다. 또한, 당해 기술 분야에서 통상의 지식을 가진 자라면 순서도에 나타난 단계들이 배타적이지 않고, 다른 단계가 포함되거나, 순서도의 하나 또는 그 이상의 단계가 본 발명의 범위에 영향을 미치지 않고 삭제될 수 있음을 이해할 수 있을 것이다.
- [0321] 상술된 실시예들은 다양한 양태의 예시들을 포함한다. 다양한 양태들을 나타내기 위한 모든 가능한 조합이 기술될 수는 없지만, 해당 기술 분야의 통상의 지식을 가진 자는 명시적으로 기술된 조합 외에도 다른 조합이 가능함을 인식할 수 있을 것이다. 따라서, 본 발명은 이하의 특허청구범위 내에 속하는 모든 다른 교체, 수정 및 변경을 포함한다고 할 것이다.
- [0322] 이상 설명된 본 발명에 따른 실시예들은 다양한 컴퓨터 구성요소를 통하여 수행될 수 있는 프로그램 명령어의 형태로 구현되어 컴퓨터 판독 가능한 기록 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능한 기록 매체는 프로

그럼 명령어, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 컴퓨터 판독 가능한 기록 매체에 기록되는 프로그램 명령어는 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 분야의 당업자에게 공지되어 사용 가능한 것일 수도 있다.

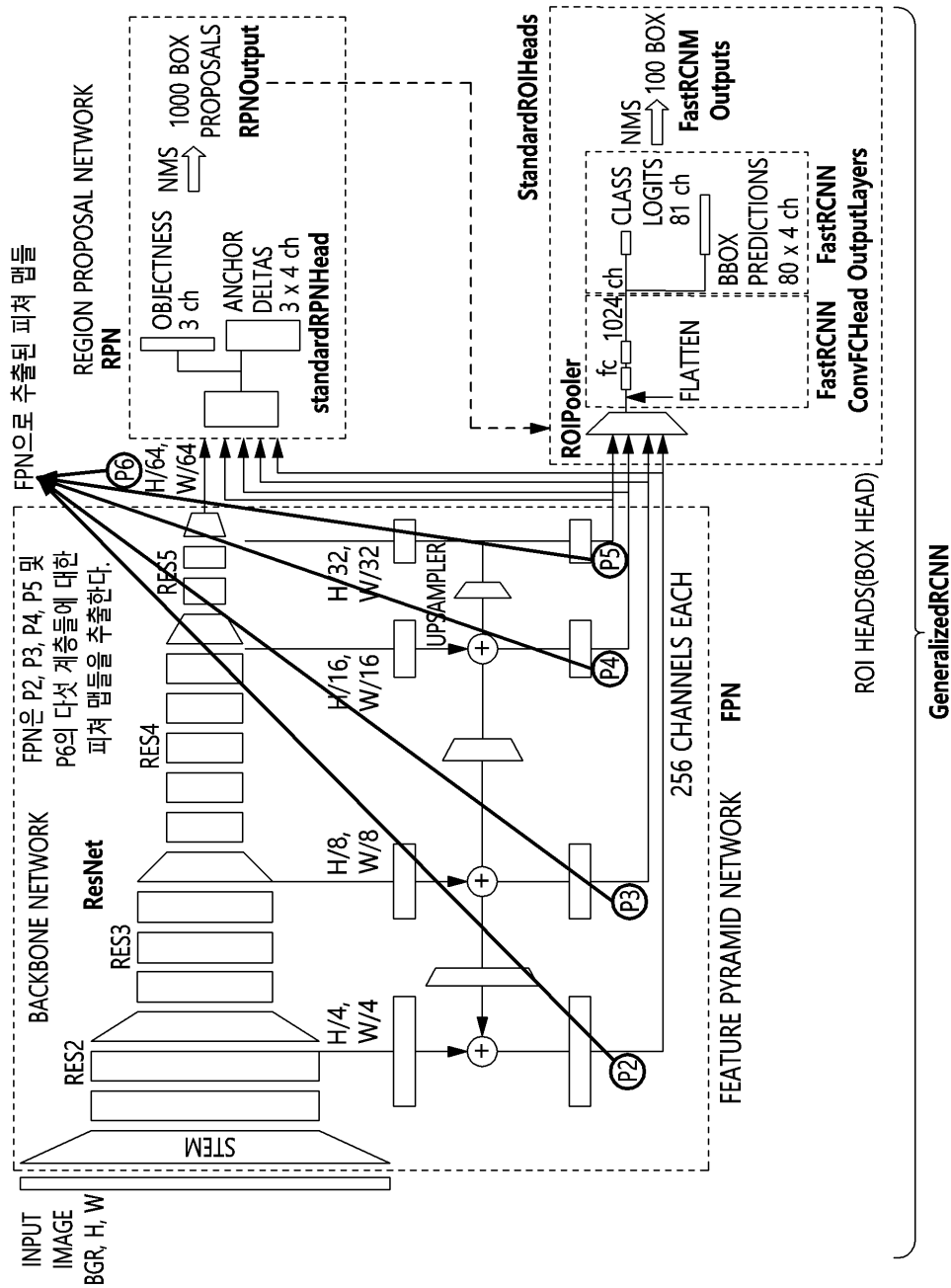
- [0323] 컴퓨터 판독 가능한 기록 매체는 본 발명에 따른 실시예들에서 사용되는 정보를 포함할 수 있다. 예를 들면, 컴퓨터 판독 가능한 기록 매체는 비트스트림을 포함할 수 있고, 비트스트림은 본 발명에 따른 실시예들에서 설명된 정보를 포함할 수 있다.
- [0324] 컴퓨터 판독 가능한 기록 매체는 비-일시적 컴퓨터 판독 가능한 매체(non-transitory computer-readable medium)를 포함할 수 있다.
- [0325] 컴퓨터 판독 가능한 기록 매체의 예에는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM, DVD와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 ROM, RAM, 플래시 메모리 등과 같은 프로그램 명령어를 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령어의 예에는, 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드도 포함된다. 상기의 하드웨어 장치는 본 발명에 따른 처리를 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0326] 이상에서 본 발명이 구체적인 구성요소 등과 같은 특정 사항들과 한정된 실시예 및 도면에 의해 설명되었으나, 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐, 본 발명이 상기의 실시예들에 한정되는 것은 아니며, 본 발명이 속하는 기술분야에서 통상적인 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형을 꾀할 수 있다.
- [0327] 따라서, 본 발명의 사상은 상기의 설명된 실시예에 국한되어 정해져서는 아니 되며, 후술하는 특허청구범위뿐만 아니라 이 특허청구범위와 균등하게 또는 등가적으로 변형된 모든 것들은 본 발명의 사상의 범주에 속한다고 할 것이다.

도면

도면1

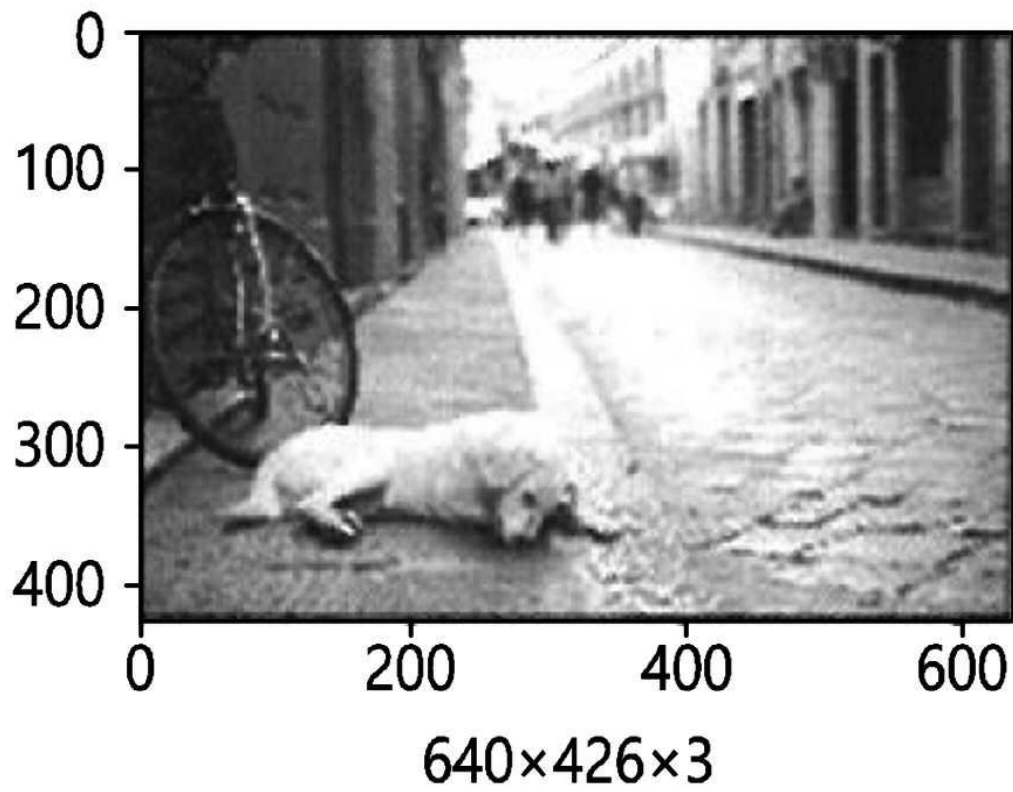


도면2

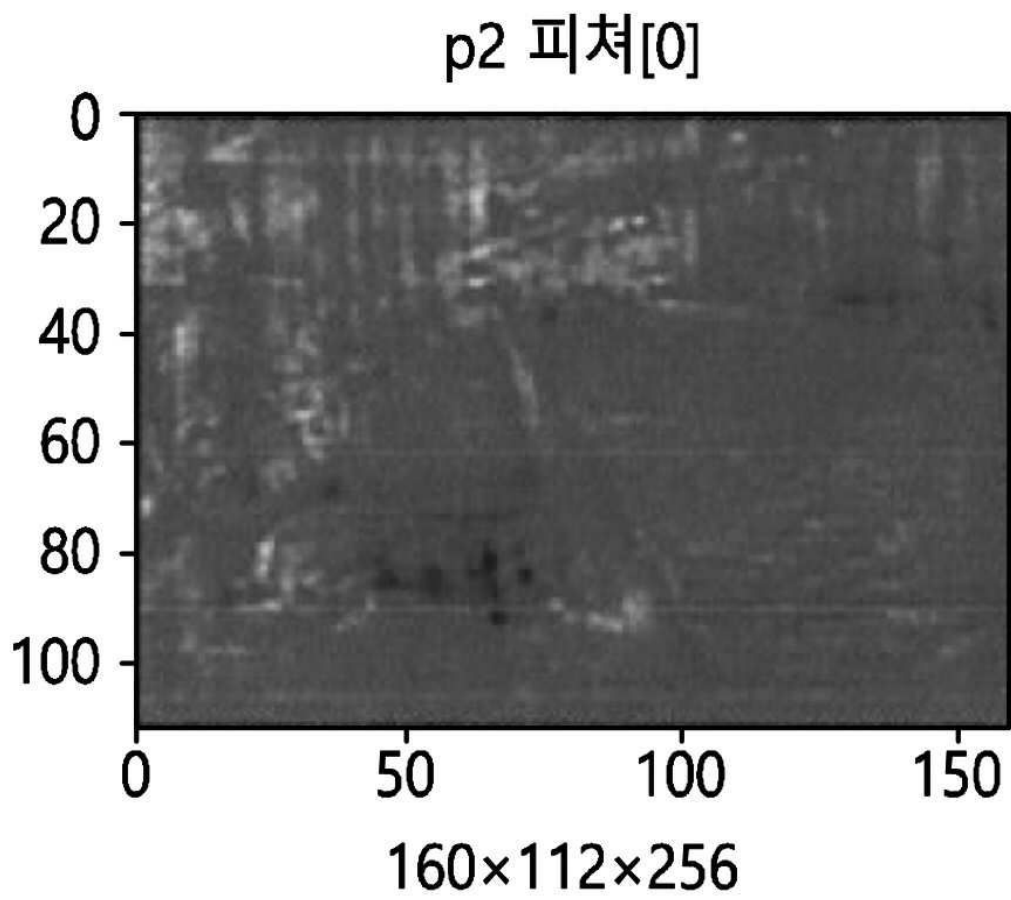


도면3

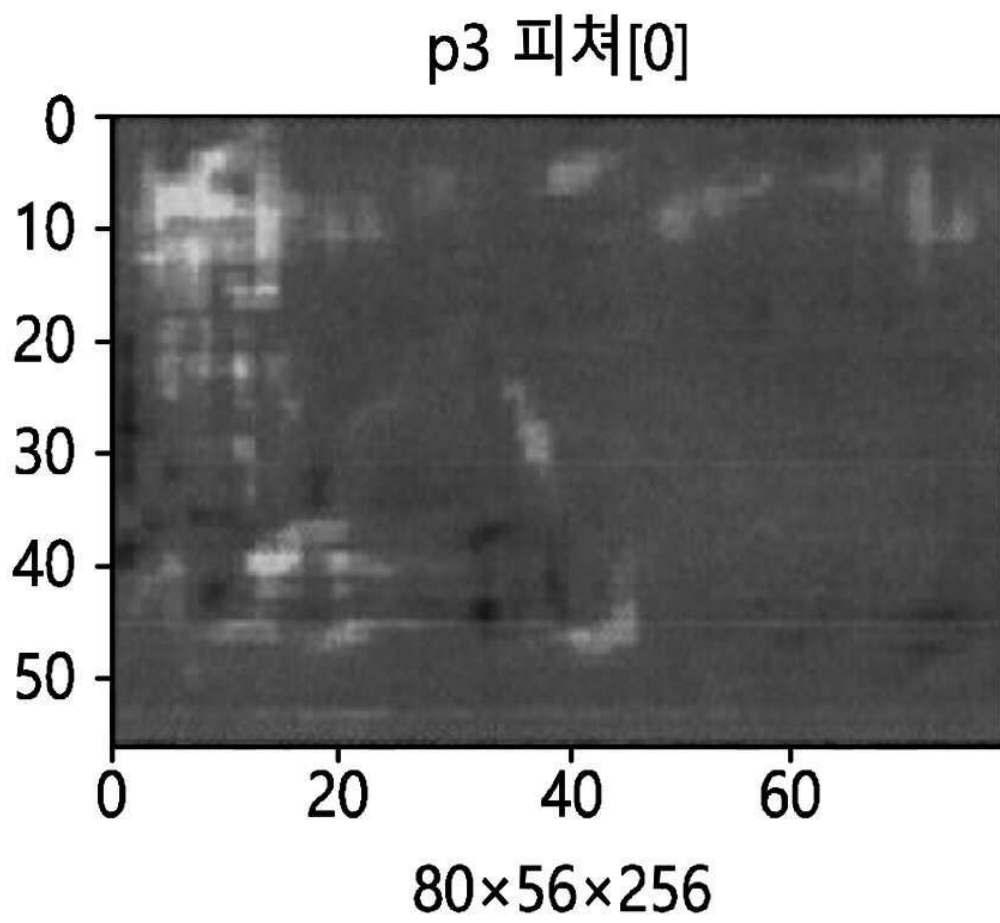
원본 피쳐 맵



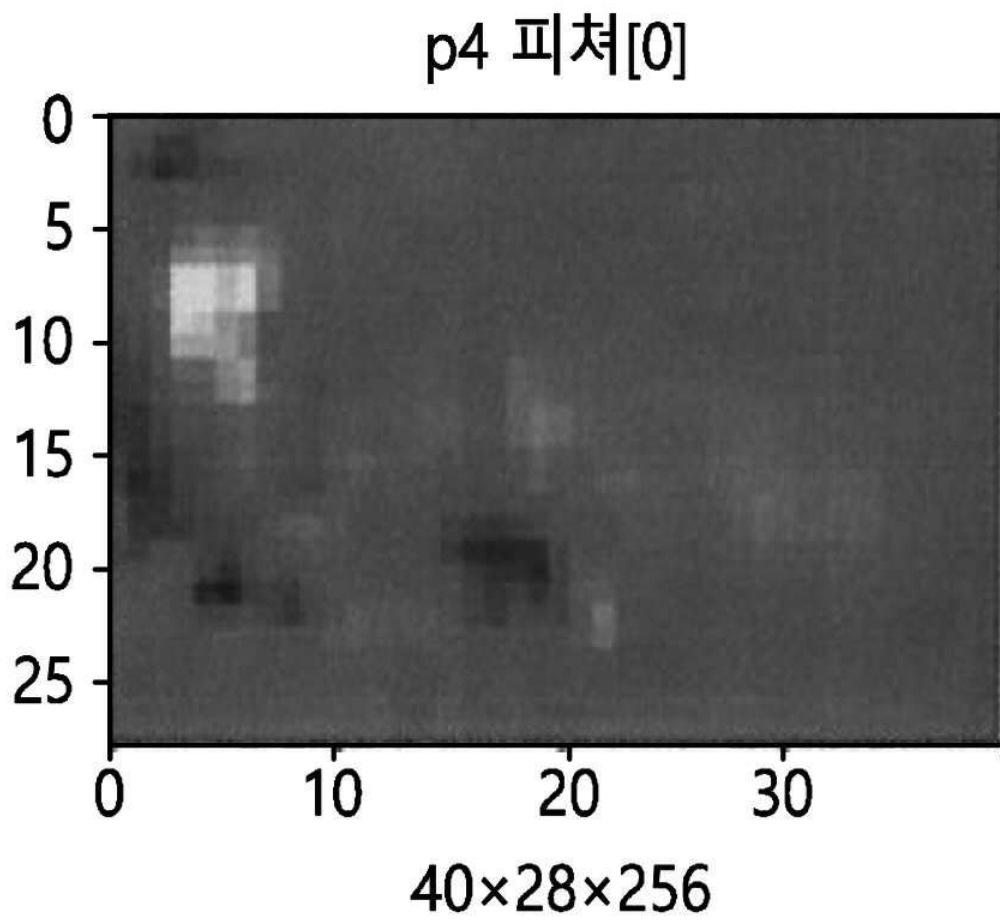
도면4



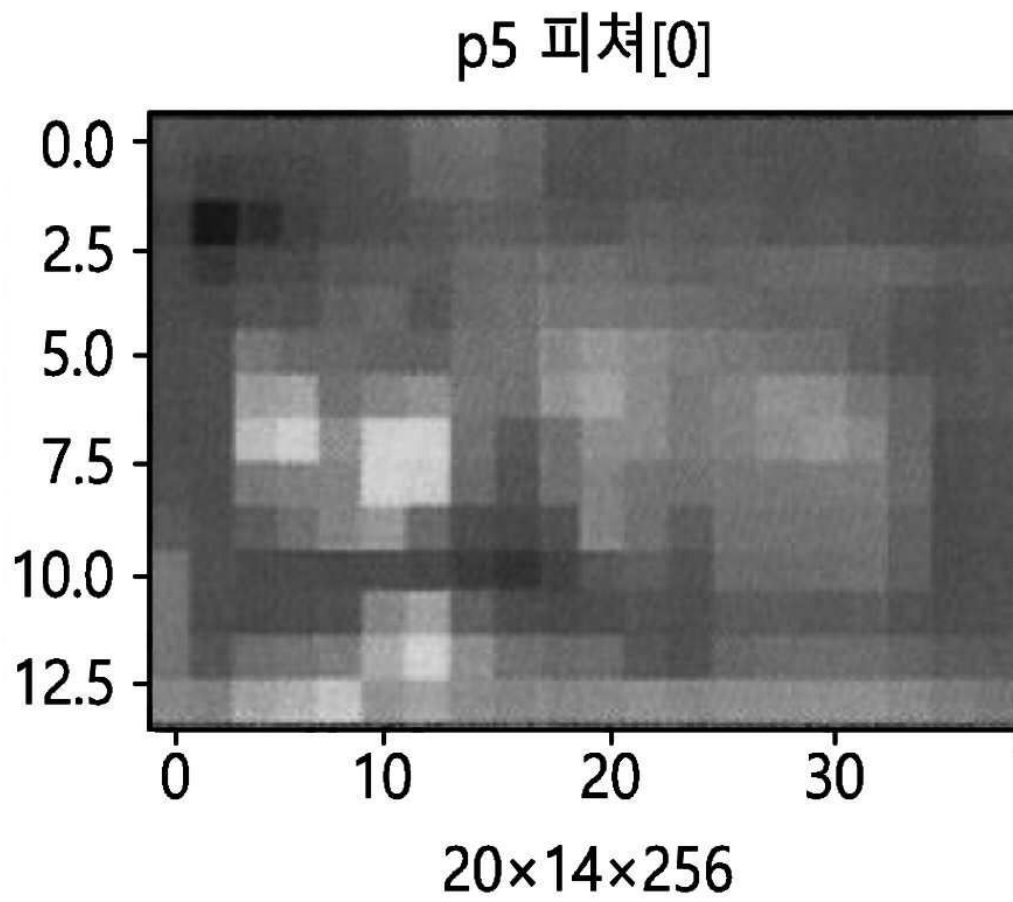
도면5



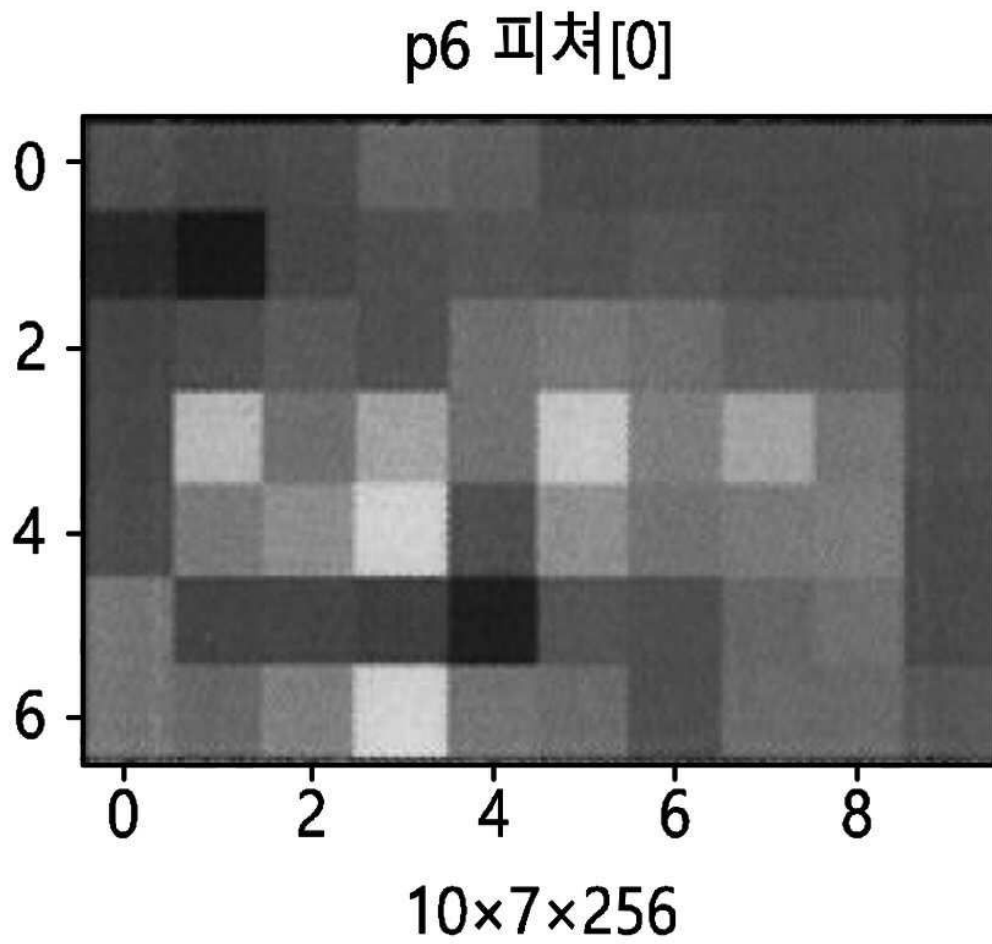
도면6



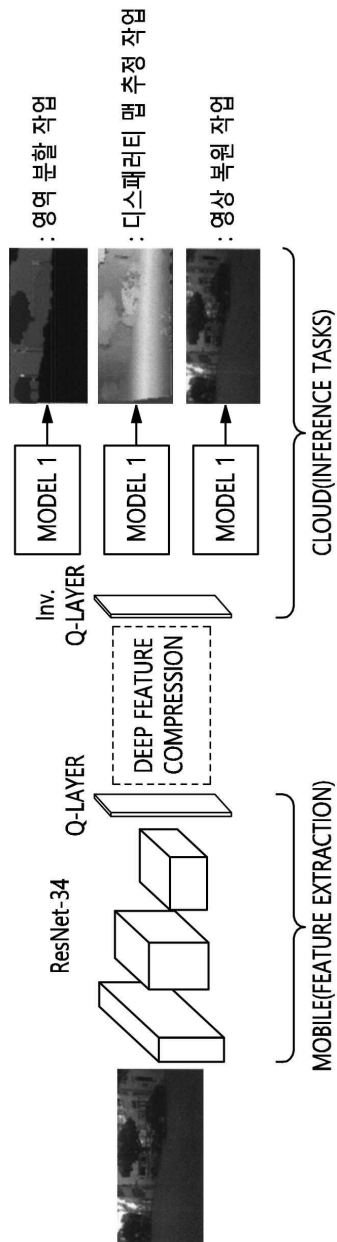
도면7



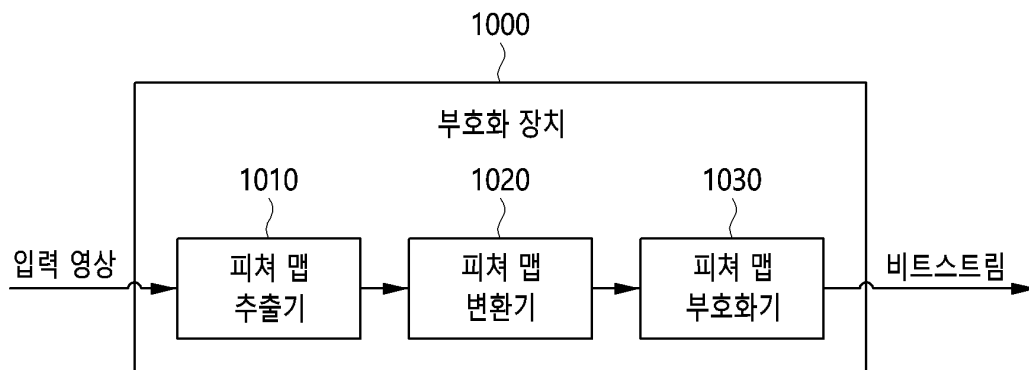
도면8



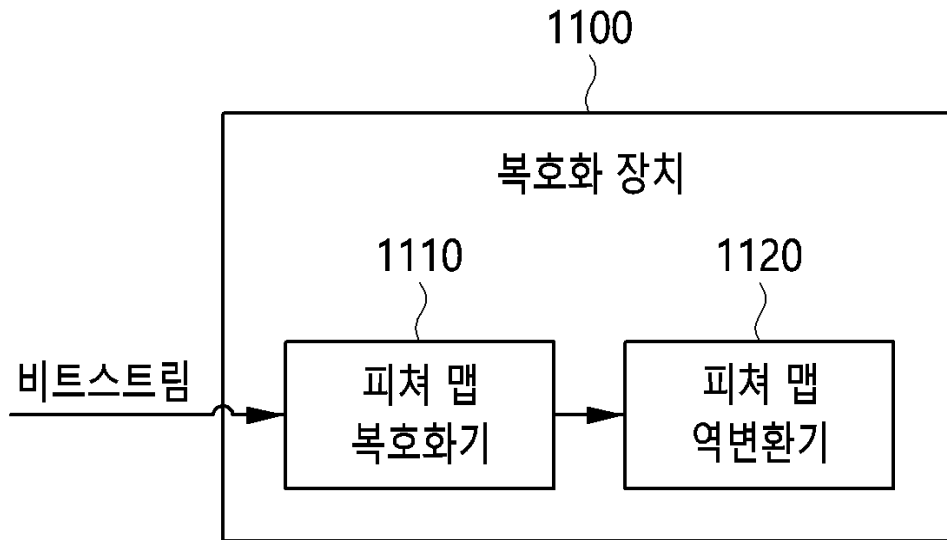
도면9



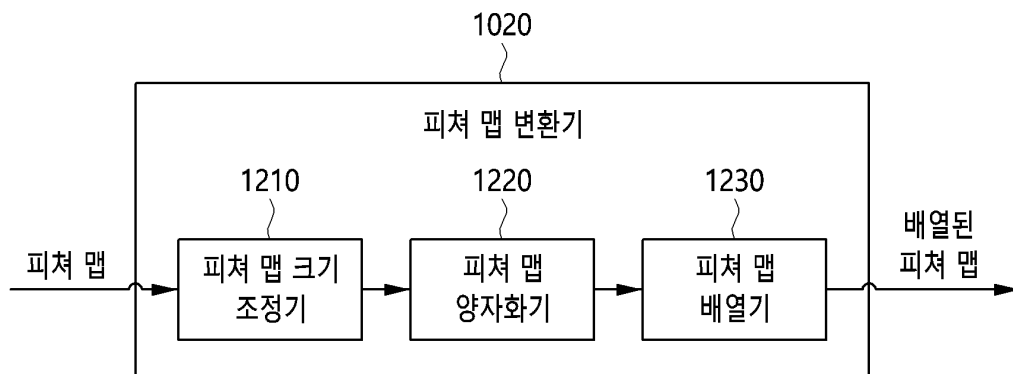
도면10



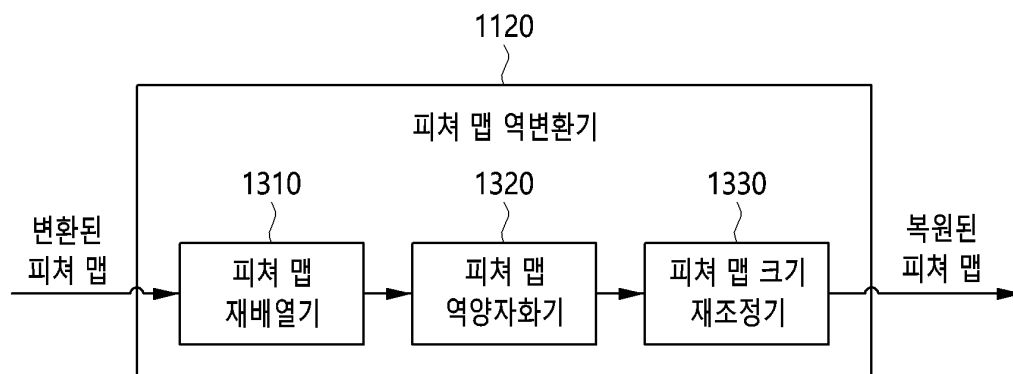
도면11



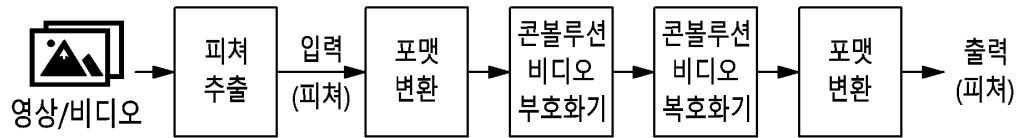
도면12



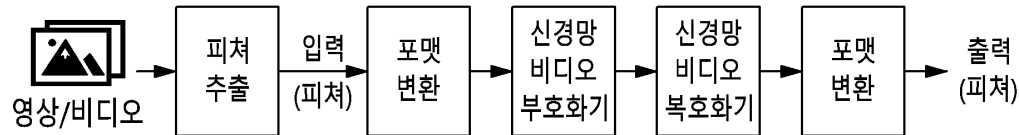
도면13



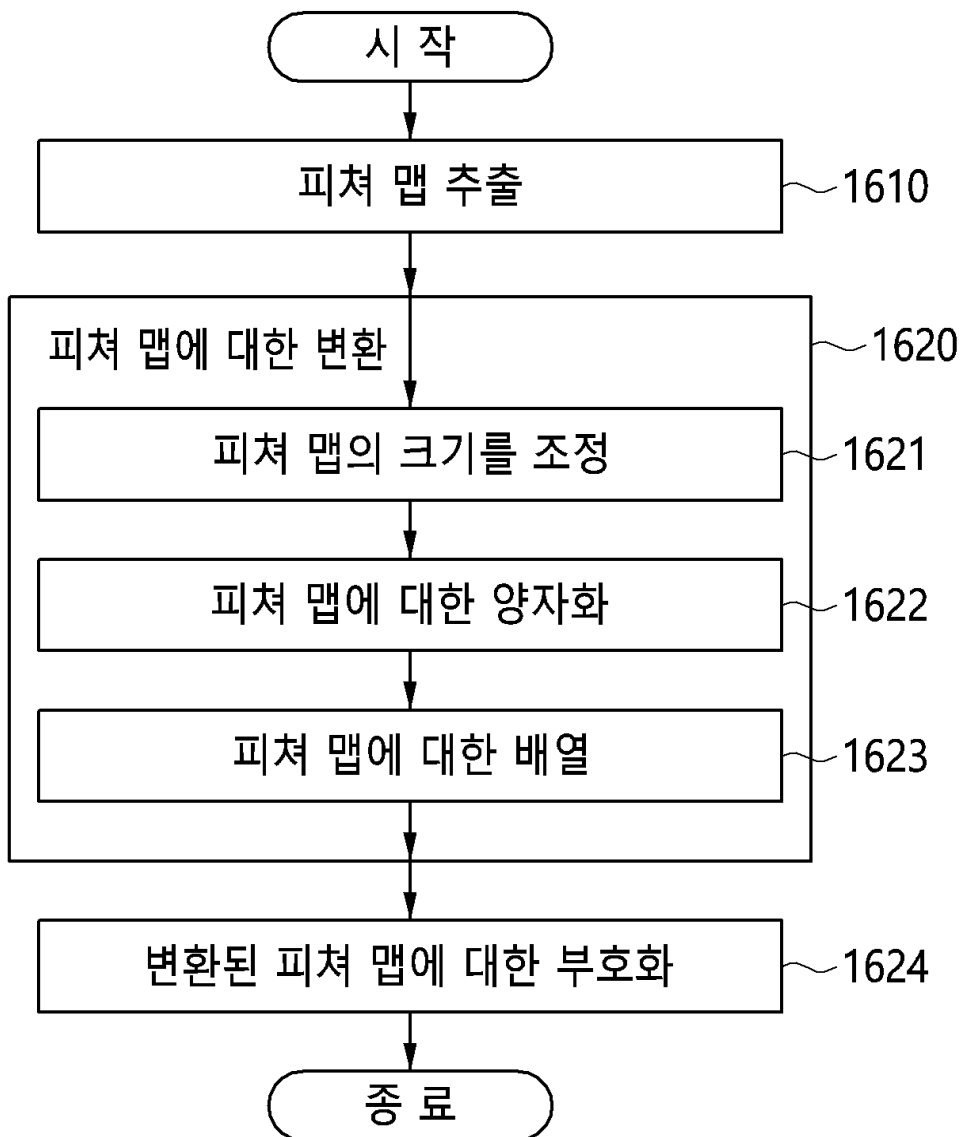
도면14



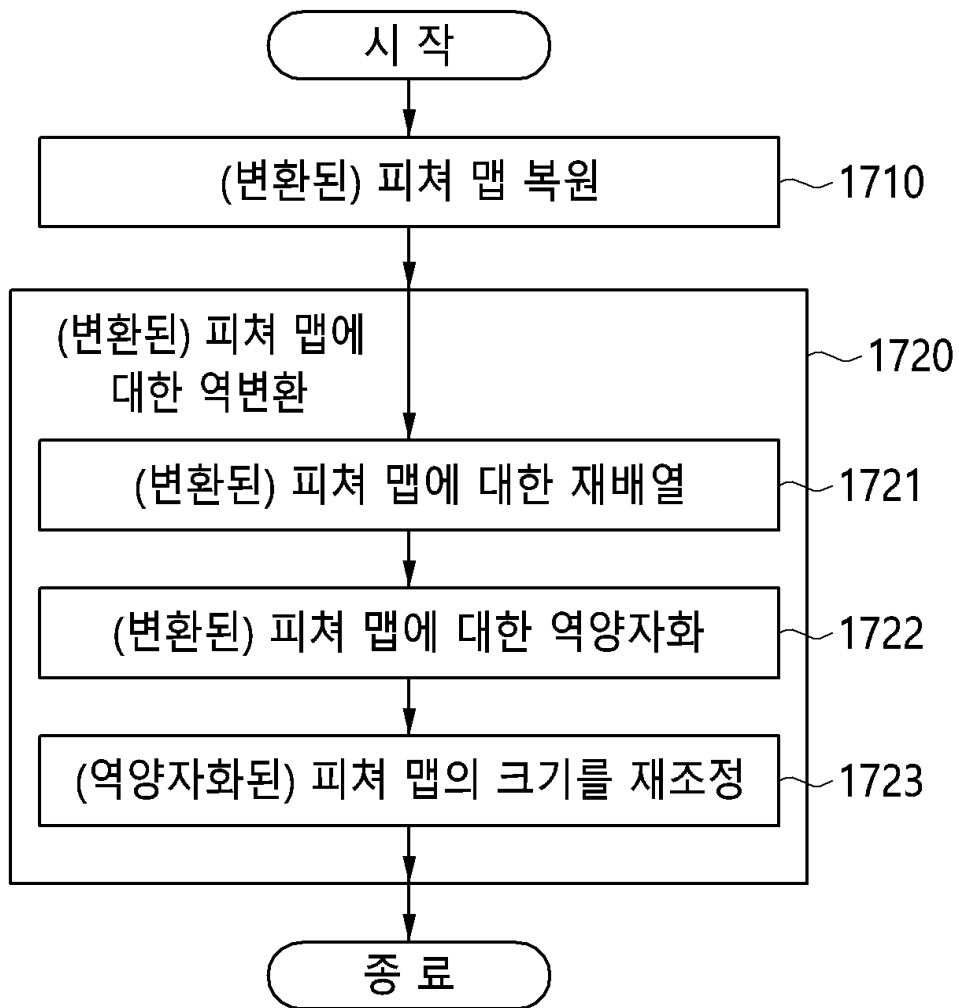
도면15



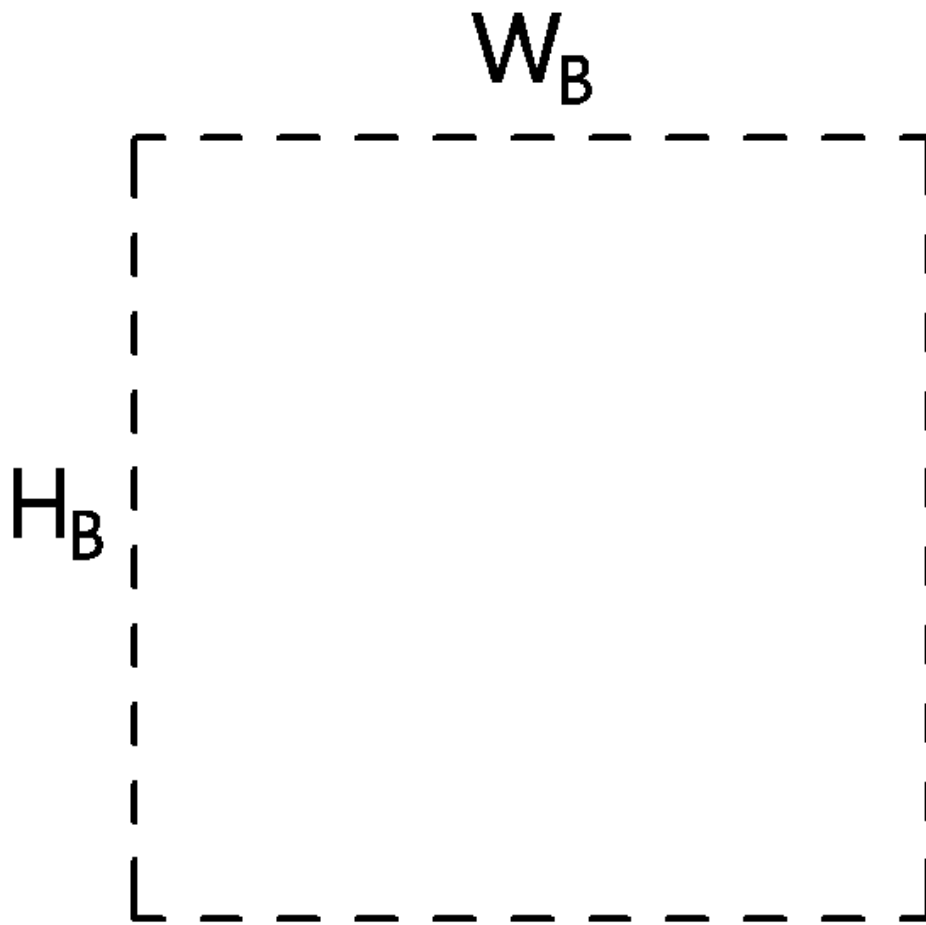
도면16



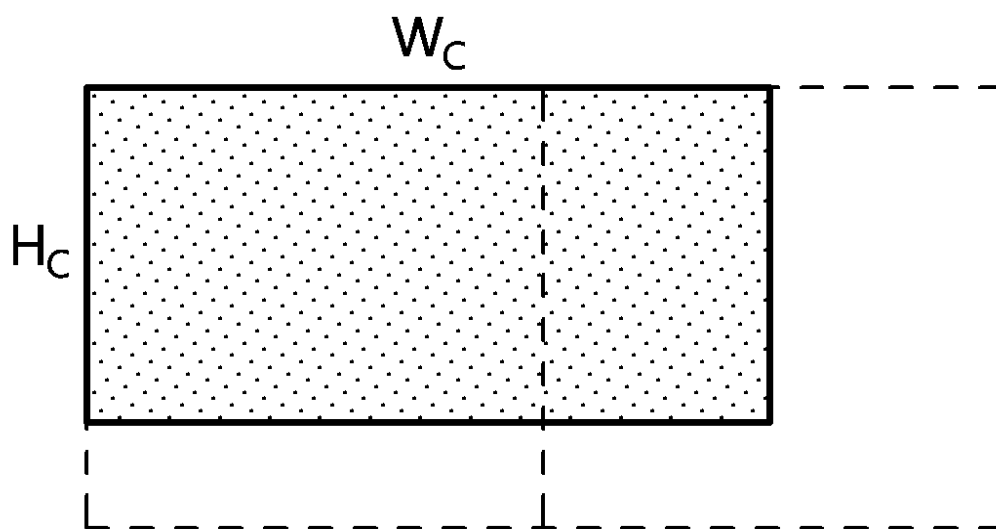
도면17



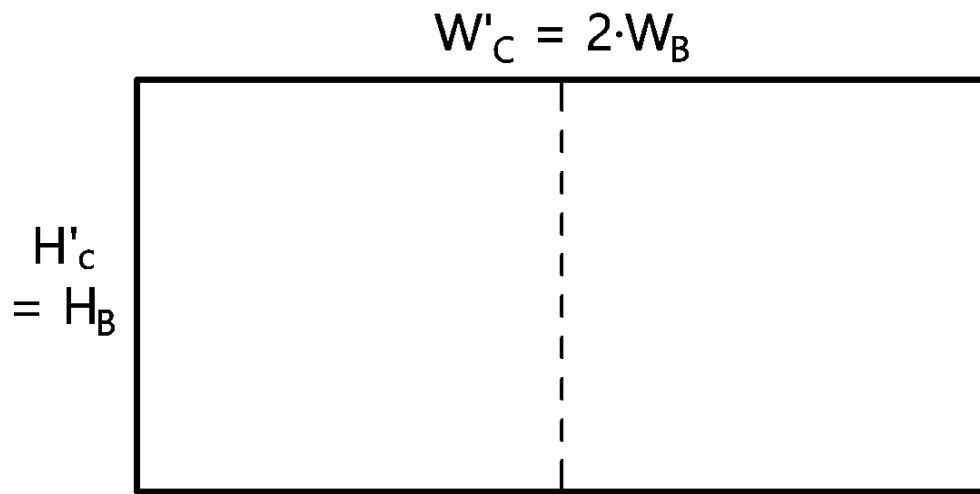
도면18



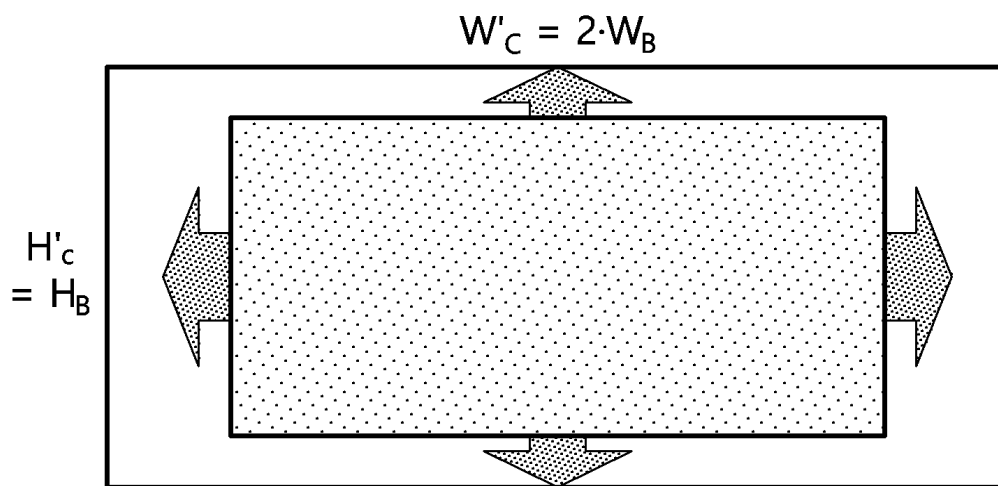
도면19



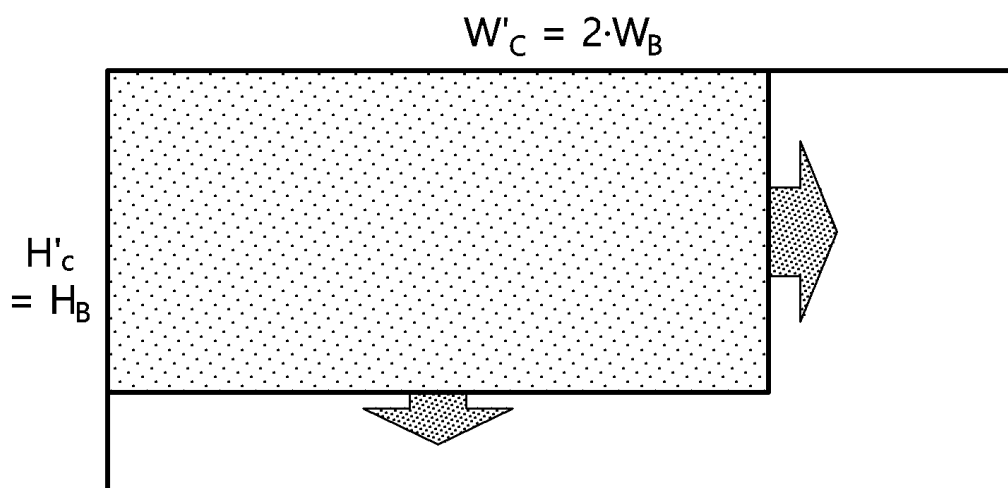
도면20



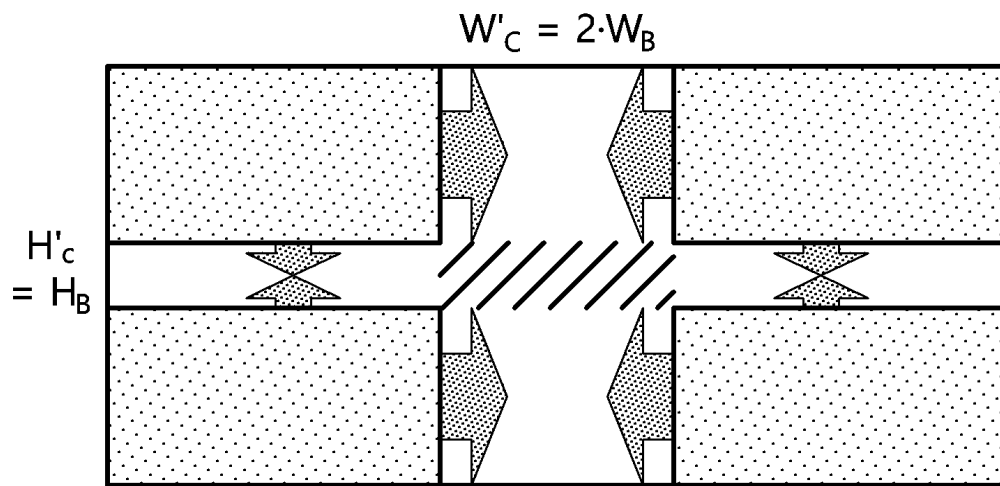
도면21



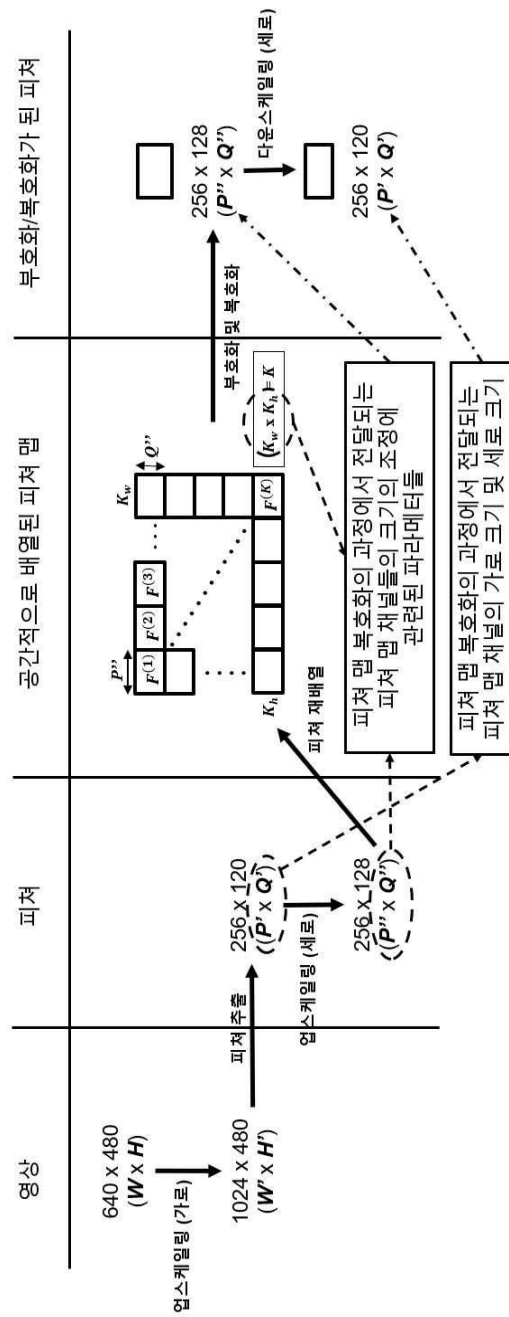
도면22



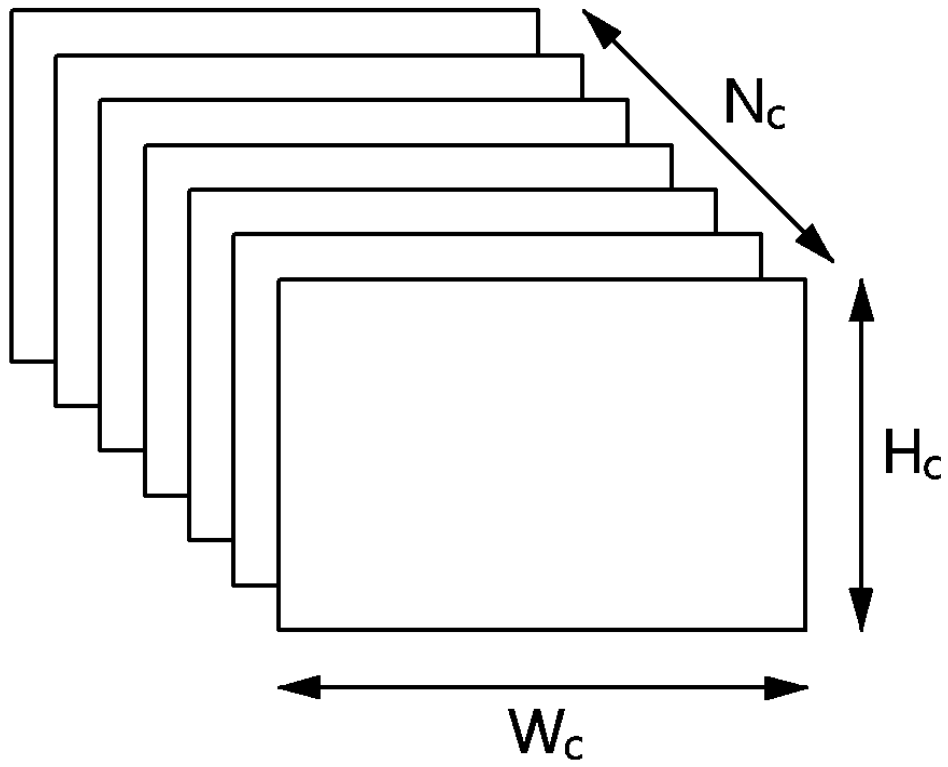
도면23



도면24

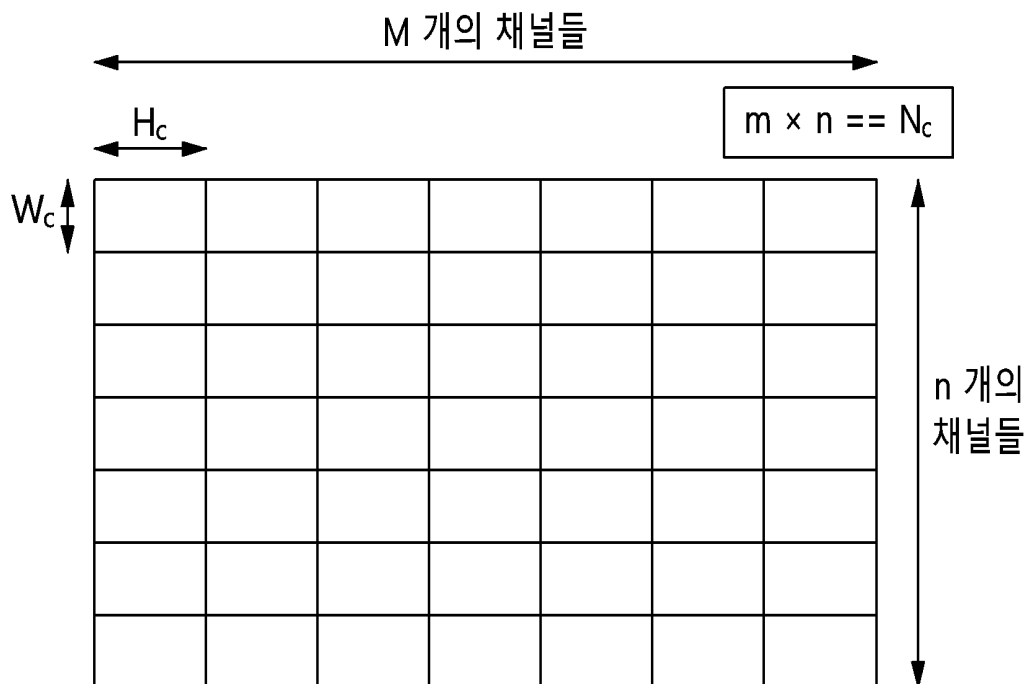


도면25

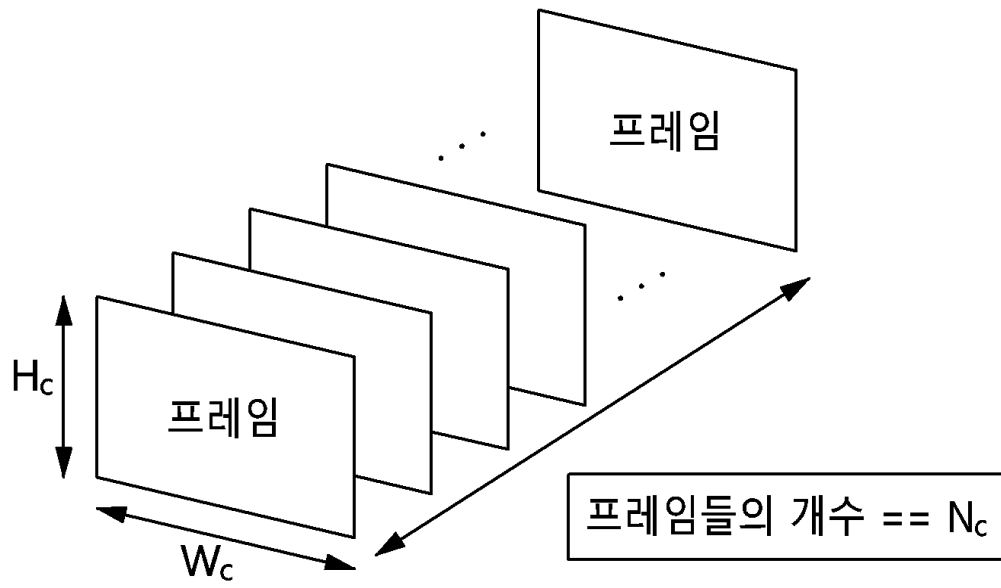


피쳐 맵의 구성

도면26



도면27



도면28

