



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2016년06월02일
(11) 등록번호 10-1627069
(24) 등록일자 2016년05월27일

(51) 국제특허분류(Int. Cl.)
H04N 19/59 (2014.01) H04N 19/176 (2014.01)
H04N 19/60 (2014.01)
(52) CPC특허분류
H04N 19/59 (2015.01)
H04N 19/176 (2015.01)
(21) 출원번호 10-2015-0045914
(22) 출원일자 2015년04월01일
심사청구일자 2015년04월01일
(56) 선행기술조사문헌
KR1020150032647 A*
US08767835 B2*
Je-Won Kang ET AL: "Sparse/DCT (S/DCT)
Two-Layered Representation of Prediction
Residuals for Video Coding", IEEE
TRANSACTIONS ON IMAGE PROCESSING, VOL. 22,
NO. 7, JULY 2013
KR1020140120341 A
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
이화여자대학교 산학협력단
서울특별시 서대문구 이화여대길 52 (대현동, 이
화여자대학교)
(72) 발명자
강제원
서울특별시 마포구 독막로 145, 110동 502호(창전
동, 서강쌍용에가아파트)
강민주
서울특별시 서대문구 이화여대2다길 15, 봉하숙지
하2호 (대현동)
(74) 대리인
남정길

전체 청구항 수 : 총 10 항

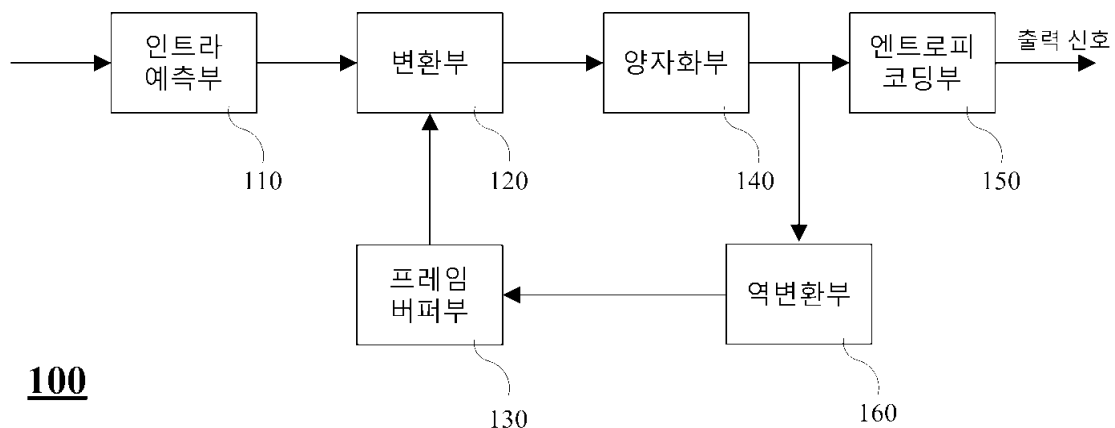
심사관 : 장석환

(54) 발명의 명칭 동영상 인코딩에서 잔차 신호에 대한 회소 변환 방법, 회소 변환을 이용한 동영상 인코딩 방법 및 회소 변환을 이용하여 부호화된 신호에 대한 디코딩 방법

(57) 요약

회소 변환을 이용한 동영상 인코딩 방법은 인코더가 $i-1$ 번째 프레임에 대한 잔차 신호에 대한 회소 변환을 수행하는 단계, 상기 인코더가 소스 영상 신호의 i 번째 프레임에 대한 인트라 예측을 수행하여 잔차 신호를 생성하는 단계, 상기 인코더가 상기 잔차 신호 중 0이 아닌 계수를 갖는 픽셀로 구성된 계수 블록 패치를 생성하는 단계 및 상기 인코더가 상기 $i-1$ 번째 프레임에 대한 회소 변환에서 사용된 변환 함수 및 상기 계수 블록 패치를 이용하여 상기 i 번째 프레임의 잔차 신호에 대한 회소 변환을 수행하는 단계를 포함한다.

대표도



(52) CPC특허분류

H04N 19/60 (2015.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 1345228037

부처명 교육부

연구관리전문기관 한국연구재단

연구사업명 기초연구사업(학술진흥)(2014)-일반연구자지원사업-기본연구

연구과제명 미래 스마트 미디어 이미지/비디오 서비스를 위한 정보 부호화 기술

기 여 율 1/1

주관기관 이화여자대학교 산학협력단

연구기간 2014.11.01 ~ 2017.04.30

명세서

청구범위

청구항 1

인코더가 소스 영상 신호의 i 번째 프레임에 대한 방향성 인트라 예측을 수행하여 잔차 신호를 생성하는 단계;

상기 인코더가 상기 잔차 신호 중 0이 아닌 계수를 갖는 픽셀로 구성된 블록 패치를 생성하는 단계; 및

상기 인코더가 상기 블록 패치를 상기 i 번째 프레임의 블록에 마스킹하고, 상기 블록에 마스킹한 값을 이용하여 상기 잔차 신호에 대한 최소 변환을 수행하는 단계를 포함하는 동영상 인코딩에서 잔차 신호에 대한 최소 변환 방법.

청구항 2

삭제

청구항 3

제1항에 있어서,

상기 최소 변환을 수행하는 단계는 아래의 수학적식을 만족하는 계수 행렬 U 와 변환 행렬 V 를 구하는 것인 동영상 인코딩에서 잔차 신호에 대한 최소 변환 방법.

$$\min_{U, V} \|X - UV^T\|_F^2 + \lambda \sum_{k=1}^p S(V_k)$$

(여기서, X 는 i 번째 프레임에 대한 입력 행렬, λ 는 상수인 계수이고, $S(V_k)$ 는 블록에 상기 계수 블록 패치를 마스킹한 값이고, k 는 블록의 식별자, F 는 matrix의 프로베니우스 놈, p 는 변환 행렬의 계수의 개수임)

청구항 4

제3항에 있어서,

상기 $S(V_k)$ 는

$$S(V_k) = \sum_{B \in \mathbf{B}} \sum_{j \in B} |(b^B(j)V_k(j))|^2)^{\frac{1}{2}} = \sum_{B \in \mathbf{B}} \|(b^B \circ V_k)\|_2$$

이고, $\mathbf{B}$ 는 상기 계수 블

록 패치이고, j 는 픽셀의 좌표값이고, $V_k(j)$ 는 j 에 대한 컬럼 벡터이고, b^B 는 $b^B(j)$ 가 B 에 속한 픽셀이면 1이고 B 에 속한 픽셀이 아니면 0인 값이고, $\circ$ 는 행렬 성분에 대한 곱셈인 동영상 인코딩에서 잔차 신호에 대한 최소 변환 방법.

청구항 5

인코더가 $i-1$ 번째 프레임에 대한 잔차 신호에 대한 최소 변환을 수행하는 단계;

상기 인코더가 소스 영상 신호의 i 번째 프레임에 대한 방향성 인트라 예측을 수행하여 잔차 신호를 생성하는 단계;

상기 인코더가 상기 잔차 신호 중 0이 아닌 계수를 갖는 픽셀로 구성된 블록 패치($\mathbf{B}$)를 생성하는 단계; 및

상기 인코더가 상기 $i-1$ 번째 프레임에 대한 최소 변환에서 사용된 변환 함수 및 상기 i 번째 프레임의 블록에 상기 블록 패치를 마스킹한 값을 이용하여 상기 i 번째 프레임의 잔차 신호에 대한 최소 변환을 수행하는 단계

를 포함하는 희소 변환을 이용한 동영상 인코딩 방법.

청구항 6

제5항에 있어서,

상기 i-1 번째 프레임에 대한 희소 변환은 아래의 수학적식을 만족하는 계수 행렬 U와 변환 행렬 V를 구하는 것인 희소 변환을 이용한 동영상 인코딩 방법.

$$\min_{\mathbf{U}, \mathbf{V}} \|\mathbf{X} - \mathbf{UV}^T\|_F^2 + \lambda \sum_{k=1}^p S(V_k)$$

(여기서, λ 는 상수인 계수이고, $S(V_k)$ 는 블록에 상기 계수 블록 패치를 마스킹한 값이고, k는 블록의 식별자, F는 matrix의 프로베니우스 놈, p는 변환 행렬의 계수의 개수임)

청구항 7

제6항에 있어서,

상기 $S(V_k)$ 는

$$S(V_k) = \sum_{B \in \mathbf{B}} \sum_{j \in B} |(b^B(j)V_k(j))|^2 \frac{1}{2} = \sum_{B \in \mathbf{B}} \|(b^B \circ V_k)\|_2$$

이고, B는 상기 계수 블록 패치이고, j는 픽셀의 좌표값이고, $V_k(j)$ 는 j에 대한 컬럼 벡터이고, b^B 는 $b^B(j)$ 가 B에 속한 픽셀이면 1이고 B에 속한 픽셀이 아니면 0인 값이고, $\circ$ 는 행렬 성분에 대한 곱셈인 희소 변환을 이용한 동영상 인코딩 방법.

청구항 8

제5항에 있어서,

상기 i 번째 프레임의 잔차 신호에 대한 희소 변환은 아래의 식으로 표현되는 상기 i 번째 프레임에 대한 계수 행렬 U_i 와 변환 행렬 V_i 를 이용하여 수행되는 희소 변환을 이용한 동영상 인코딩 방법.

$$\mathbf{U}_i = \arg \min_{\|\mathbf{U}\| \leq 1} \|\mathbf{X}_i - \mathbf{UV}_{i-1}^T\|_F^2 + \lambda \sum_{k=1}^p S(V_{k,i-1})$$

$$\mathbf{V}_i = \arg \min_{\mathbf{V}} \|\mathbf{X}_i - \mathbf{U}_i \mathbf{V}^T\|_F^2 + \lambda \sum_{k=1}^p S(V_k)$$

(여기서, X_i 는 i 번째 프레임에 대한 입력 행렬, V_{i-1} 은 i-1 번째 프레임에 대한 변환 행렬, $S(V_{k,i-1})$ 는 i-1 번째 블록에 계수 블록 패치를 마스킹한 값이고, $S(V_k)$ 는 i 번째 블록에 상기 계수 블록 패치 $\mathbf{B}$ 를 마스킹한 값, k는 블록에 대한 식별자, F는 matrix의 프로베니우스 놈, p는 변환 행렬의 계수의 개수임)

청구항 9

제8항에 있어서,

$$\text{상기 } S(V_k) \text{는 } S(V_k) = \frac{1}{2} V_k^T \text{diag}(\zeta_k) V_k + \|z\|_\beta \text{ 이고, } \zeta_{kj} = \sum_{B \in \mathbf{B}} (b_{kj}^B) z_k \text{ 이고,}$$

$z_k = \|b^B \circ V_k\|_2$ 이고, b^B 는 $b^B(j)$ 가 B에 속한 픽셀이면 1이고 B에 속한 픽셀이 아니면 0인 값이고, $\circ$ 는 행렬 성분에 대한 곱셈인 희소 변환을 이용한 동영상 인코딩 방법.

청구항 10

제5항에 있어서,

상기 인코더는 윌 왜곡 최적화 기법을 이용하여 DCT(discrete cosine transform) 보다 상기 i 번째 프레임에 대한 희소 변환의 효율이 좋은 경우에 상기 i 번째 프레임의 잔차 신호에 대한 희소 변환을 수행하는 희소 변환을 이용한 동영상 인코딩 방법.

청구항 11

디코더가 i-1 번째 프레임에 대한 희소 변환 함수를 저장하는 단계;

상기 디코더가 수신한 i 번째 프레임 신호에 대한 역 엔트로피 코딩 및 역 양자화 변환을 수행하는 단계;

상기 디코더가 역 양자화 변환 후 생성된 신호 중 0이 아닌 계수를 갖는 픽셀로 구성된 블록 패치를 생성하는 단계; 및

상기 디코더가 상기 희소 변환 함수 및 상기 블록 패치를 이용한 역 희소 변환을 수행하여 잔차 신호를 생성하는 단계를 포함하는 희소 변환을 이용한 동영상 신호에 대한 디코딩 방법.

발명의 설명

기술 분야

[0001] 이하 설명하는 기술은 학습 기반의 희소 변환을 사용하는 동영상 코딩 기법에 관한 것이다.

배경 기술

[0002] 인트라 코딩에서 방향성 예측(directional prediction) 후에 생성되는 잔차 신호는 블록 단위로 변환된다. 신호 변환에는 다양한 방식이 사용될 수 있다. HEVC에서는 기본적으로 DCT(discrete cosine transform)이 사용되고, 부가적으로 DST(discrete sine transform)가 사용되기도 한다.

[0003] 한편 희소 코딩은 신호의 일부를 줄이면서 코딩 효율을 높이는 방법으로 연구되고 있다. 동영상 코딩에서도 잔차 신호에 대한 사전 기반의 코딩 기법이 제시된 바 있다.

선행기술문헌

특허문헌

[0004] (특허문헌 0001) 미국등록특허 US 8,767,835호

발명의 내용

해결하려는 과제

[0005] 종래 사전 기반의 코딩 기법은 잔차 신호를 나타내는 계수 중 0 이 아닌 계수의 숫자를 줄이는데 초점이 맞추어졌다.

[0006] 이하 설명하는 기술은 잔차 신호에 대해 0이 아닌 계수의 숫자를 줄일 뿐만 아니라 계수의 위치를 블록 내의 특정 위치에 집중시켜 보다 효율이 좋은 동영상 코딩 방법을 제공하고자 한다.

[0007] 이하 설명하는 기술의 해결과제는 이상에서 언급된 것들에 한정되지 않으며, 언급되지 아니한 다른 해결과제들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0008] 동영상 인코딩에서 잔차 신호에 대한 회소 변환 방법은 인코더가 소스 영상 신호의 i 번째 프레임에 대한 인트라 예측을 수행하여 잔차 신호를 생성하는 단계, 인코더가 잔차 신호 중 0이 아닌 계수를 갖는 픽셀로 구성된 계수 블록 패치를 생성하는 단계 및 인코더가 계수 블록 패치를 i 번째 프레임의 블록에 적용하고, 잔차 신호에 대한 회소 변환을 수행하는 단계를 포함한다.

[0009] 회소 변환을 이용한 동영상 인코딩 방법은 인코더가 $i-1$ 번째 프레임에 대한 잔차 신호에 대한 회소 변환을 수행하는 단계, 인코더가 소스 영상 신호의 i 번째 프레임에 대한 인트라 예측을 수행하여 잔차 신호를 생성하는 단계, 인코더가 잔차 신호 중 0이 아닌 계수를 갖는 픽셀로 구성된 계수 블록 패치를 생성하는 단계 및 인코더가 $i-1$ 번째 프레임에 대한 회소 변환에서 사용된 변환 함수 및 계수 블록 패치를 이용하여 i 번째 프레임의 잔차 신호에 대한 회소 변환을 수행하는 단계를 포함한다.

[0010] 회소 변환을 이용하여 부호화된 신호에 대한 디코딩 방법은 디코더가 $i-1$ 번째 프레임에 대한 회소 변환 함수를 저장하는 단계, 디코더가 수신한 i 번째 프레임 신호에 대한 엔트로피 코딩 및 역 양자화 변환을 수행하는 단계, 디코더가 역 양자화 변환 후 생성된 잔차 신호 중 0이 아닌 계수를 갖는 픽셀로 구성된 계수 블록 패치를 생성하는 단계 및 디코더가 회소 변환 함수 및 계수 블록 패치를 이용하여 잔차 신호에 대한 역 회소 변환을 수행하는 단계를 포함한다.

발명의 효과

[0011] 이하 설명하는 기술은 방향성 인트라 예측 후의 잔차 신호를 나타내는 계수의 숫자가 적고, 0이 아닌 계수가 특정한 영역에 모여 있어서 보다 낮은 비용으로 동영상 코딩이 가능하다.

[0012] 이하 설명하는 기술의 효과는 이상에서 언급된 것들에 한정되지 않으며, 언급되지 아니한 다른 효과들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

[0013] 도 1은 인트라 예측 후 0이 아닌 계수의 패턴에 대한 그룹을 도시한 예이다.

도 2는 인트라 코딩을 위한 인코더의 구성을 도시한 예이다.

도 3은 동영상 인코딩 방법에 대한 순서도의 예이다.

도 4는 동영상 디코딩 방법에 대한 순서도의 예이다.

발명을 실시하기 위한 구체적인 내용

[0014] 이하 설명하는 기술은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 이하 설명하는 기술을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 이하 설명하는 기술의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0015] 제1, 제2, A, B 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 해당 구성요소들은 상기 용어들에 의해 한정되지는 않으며, 단지 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 이하 설명하는 기술의 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 포함한다.

[0016] 본 명세서에서 사용되는 용어에서 단수의 표현은 문맥상 명백하게 다르게 해석되지 않는 한 복수의 표현을 포함하는 것으로 이해되어야 하고, "포함한다" 등의 용어는 실시된 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 구성부 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 의

해 전달되어 수행될 수도 있음은 물론이다. 따라서, 본 명세서를 통해 설명되는 각 구성부들의 존재 여부는 기능적으로 해석되어야 할 것이며, 이러한 이유로 이하 설명하는 디코더(100)에 따른 구성부들의 구성은 이하 설명하는 기술의 목적을 달성할 수 있는 한도 내에서 대응하는 도면과는 상이해질 수 있음을 명확히 밝혀둔다.

[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.

[0019] 이하 설명하는 기술은 H.264, HEVC, HEVC/RExt(HEVC Extension) 등과 같은 동영상 코딩 기술에 적용할 수 있다. 이하 설명하는 기술은 다양한 동영상 코딩 방법 내지 표준에 적용할 수 있다. 다만 설명의 편의를 위해 이하 HEVC를 기준으로 설명하고자 한다.

[0020] 인트라 코딩에서 방향성 예측(directional prediction) 후에 생성되는 잔차 신호는 블록 단위로 변환된다. 신호 변환에는 다양한 방식이 사용될 수 있다. HEVC에서는 기본적으로 DCT(discrete cosine transform)이 사용되고, 추가적으로 DST(discrete sine transform)가 사용되기도 한다.

[0021] 이하 설명하는 기술은 사전 기반의 코딩(dictionary based video coding)에 기반한다. 사전 기반의 코딩 기법은 주로 잔차 신호에 대한 변환에 관련된 것이다.

[0022] 사전 기반의 코딩 기법은 잔차 신호를 희소 변환하여 계수의 숫자를 줄인 샘플을 사용하기도 한다. 결국 종래 기법은 잔차 신호에 대한 계수의 숫자를 줄이는 것에 초점이 있다. 그러나 종래 기법은 잔차 신호의 개수가 줄어드는 대신 학습을 위한 사건의 개수가 늘어날 수 있는 문제점이 있었다. 이하 설명하는 기술은 잔차 신호의 개수를 줄이면서 동시에 0이 아닌 잔차 신호를 일정한 영역에 집중시키는 희소 변환에 관한 것이다. 나아가 이하 설명하는 기술은 잔차 신호를 특정 영역에 집중하는 희소 변환을 활용한 동영상 코딩 방법에 관한 것이다.

[0023] 먼저 희소 변환에 대해 설명하고, 이후 동영상 코딩에 관한 내용을 설명하기로 한다.

[0024] 1. 희소 구조 변환(sparse structured transform)

[0025] 인트라 코딩에서 방향 예측 후의 잔차 신호에 대한 변환은 블록 단위로 수행된다. 입력 행렬(input matrix) X 를 구성하는 데이터 벡터 x 는 컬럼 벡터(column vector) V_k 와 계수 벡터(coefficient vector) u 의 집합의 선형적 조합에 해당한다.

[0026] V_k 는 변환 행렬 $V = \{V_1, \dots, V_p\}$ 의 원소이고, u 는 계수 행렬 U 의 원소에 해당한다. 사전 기반의 학습 과정에서 통해 V 와 U 를 동시에 학습한다. 사전 기반의 코딩 기법은 아래의 수학적 식 1과 같이 변환 후에 0이 아닌 계수가 최소가 되도록 하는 것이다. 수학적 식 1은 수학적으로 최적화 문제를 풀기위한 것이다.

수학적 식 1

$$\min_{U, V} \|X - UV^T\|_F^2 + \lambda \sum_{k=1}^p \|V_k\|_1$$

$$\text{s.t. } \|U_k\|_1 \leq 1, \forall k \in \{1, 2, \dots, p\}$$

[0028] λ 는 균일화(regularization) 파라미터이고, 수학적 식 1에서 앞에 있는 첫 번째 항과 뒤에 있는 두 번째 항에 대한 트레이드 오프(trade off) 역할을 한다. 첫 번째 항은 블록 내의 신호를 잘 나타내기 위한 것이고, 두 번째 항은 계수의 희소성과 관련된 V_k 에 대한 페널티(penalty)에 해당한다. λ 는 동영상 코딩의 목적 및 코딩 장치의 성능 등에 따라 다른 값이 사용될 수 있다. F 는 matrix의 프로베니우스 놈(Frobenius norm)으로 matrix 계수의

제곱 합에 루트를 이용하여 얻는 값이고 p 는 변환 행렬의 계수의 개수이다.

[0029] 수학식 1에서 V_k 는 계수의 개수를 줄이는 기능만을 수행한다. 이제 잔차(residue)가 일정한 영역에 모이도록 상기 수학식 1의 두 번째 항을 수정하고자 한다. 이를 위해 상기 수학식 1의 두 번째 항을 논리적 블록 패치(logical block patch) $B \in \mathbf{B}$ 로 마스킹(masking)한다. $\mathbf{B}$ 는 0이 아닌 계수의 패턴이 서로 연결된 중첩된 블록 패치의 집합에 해당한다. 이하 $\mathbf{B}$ 를 잔차 블록 패치라고 명명한다. 아래의 수학식 2는 상기 수학식 1을 B 를 이용하여 변경한 것이다.

수학식 2

$$\min_{\mathbf{U}, \mathbf{V}} \|\mathbf{X} - \mathbf{UV}^T\|_F^2 + \lambda \sum_{k=1}^p S(V_k)$$

[0031] 수학식 2의 $S(V_k)$ 는 아래의 수학식 3과 같이 정의된다.

수학식 3

$$S(V_k) = \sum_{B \in \mathbf{B}} \sum_{j \in B} |(b^B(j)V_k(j))^2|^{\frac{1}{2}} = \sum_{B \in \mathbf{B}} \|(b^B \circ V_k)\|_2$$

[0033] b^B 는 $b_{(j)}^B$ 가 픽셀 좌표 j 가 잔차 블록 패치 $\mathbf{B}$ 에 속하는 경우 1의 값을 갖고, $\mathbf{B}$ 에 속하지 않는 경우 0의 값을 갖는 벡터이다. 수학식 3에서 $\circ$ 는 행렬 성분별 곱셈(component-wise multiplication)이다. $S(V_k)$ 는 또한 블록에 있는 복수의 픽셀을 0으로 만든다(즉 계수의 희소 변환을 수행함).

[0034] 도 1은 인트라 예측 후 0이 아닌 계수의 패턴에 대한 그룹을 도시한 예이다. 도 1은 8×8 픽셀 크기의 블록을 예로 도시한다. 상기 패턴은 잔차의 분포를 고려한 것이다. 잔차는 샘플 예측의 상관 관계가 낮은 우측 하단 구석에 위치할 수 있다. 이는 예측 샘플의 상관 관계가 우측 하단으로 갈수록 낮아지기 때문이다. 따라서 블록의 우측 하단 구석 부터 범위를 넓혀가면서 잔차 블록 패치 $\mathbf{B}$ 에 속하는 패턴을 찾을 수 있을 것이다. 즉 이 과정은 해당 블록에서 블록 패치를 찾는 과정이다. 도 1은 우측 하단 구석에 있는 2×2 픽셀 영역으로부터 수직 또는 수평 방향으로 2 픽셀 씩 범위를 넓혀가면서 블록 패치를 찾을 수 있다. 블록 내의 특정 사각형은 $\mathbf{B}$ 에 속한 패턴의 교집합으로 표현될 수 있다. 예컨대, 도 1에서 우측에 도시한 영역 Z 는 블록 $A \cap B^c \cap C^c$ 로 표현된다.

[0035] 2. 동영상 인코딩

[0036] 도 2는 인트라 코딩을 위한 인코더의 구성을 도시한 예이다. 인코더(100)는 인트라 예측부(110), 인트라 예측 후의 잔차에 대한 변환부(120), 이전 프레임의 정보를 저장하는 프레임 버퍼(130), 잔차 계수에 대한 양자화부(140) 및 양자화된 신호에 대한 엔트로피 코딩을 수행하는 엔트로피 코딩부(150)를 포함한다. 또한 인코더(100)는 인코더 내부에서 인트라 예측 등을 위해 양자화된 신호를 역 변환하는 역변환부(160)도 포함한다. 인코더

(100)는 인코더에서 일반적으로 사용하는 구성도 포함하나, 도 2에서는 발명의 설명에 필요한 부분만 도시하였다.

[0037] 변환부(120)는 인트라 코딩 후의 잔차 샘플에 대해 상기 수학적 식 2에 따른 최적의 U 및 V 를 반복적으로 훈련한다.

[0038] i 번째 프레임에 대한 코딩은 $i-1$ 번째 프레임의 샘플 행렬 X_{i-1} 을 사용하여 아래의 수학적 식의 최적화 문제를 해결한다. $i-1$ 번째 프레임에 대한 V_{i-1} 는 이미 연산된 상태이다. 프레임 버퍼(130)가 이전 프레임에 대한 정보를 저장한다.

수학적 식 4

$$U_i = \arg \min_{\|U\| \leq 1} \|X_i - UV_{i-1}^T\|_F^2 + \lambda \sum_{k=1}^p S(V_{k,i-1})$$

[0039]

[0040] 수학적 식 4는 $S(V)$ 가 고정된 값인 최소 제곱 문제(least square problem)에 해당한다. 다음으로 수학적 식 4로 구한 U_i 를 이용하여 아래의 수학적 식 5의 V_i 를 연산한다.

수학적 식 5

$$V_i = \arg \min_V \|X_i - U_i V^T\|_F^2 + \lambda \sum_{k=1}^p S(V_k)$$

[0041]

[0042] 수학적 식 5에서 $S(V_k)$ 는 V 에 대해 비볼록(non convex)하다. 따라서 아래의 수학적 식 6과 같이 부가적인 z 항을 사용하여 $S(V_k)$ 를 변환할 수 있다.

수학적 식 6

$$S(V_k) = \frac{1}{2} V_k^T \text{diag}(\zeta_k) V_k + \|z\|_\beta$$

[0043]

[0044] $\zeta_{kj} = \sum_{B \in \mathbf{B}} (b_{kj}^B) z_k$ 이고, $z_k = \|b^B \circ V_k\|_2$ 이다. 수학적 식 6은 V_k 에 대한 이차 방정식 형태이고, z 는 닫혀진(closed) 항이다.

[0045] 인코딩에서 하나의 프레임에 대한 각각의 블록 패치는 인트라 코딩을 위한 방향성 예측과 함께 예측된다. 즉 도 2의 인트라 예측부(110)에서 인트라 코딩과 함께 해당 프레임에 대한 블록 패치를 예측한다.

[0046] 예컨대, $x(m)$ 이 8×8 픽셀 크기의 잔차 블록 패치이라면, $u(m)$ 은 대응되는 계수에 해당한다. 여기서 $m = 1, 2, \dots, M$ 이고, M 은 해당 프레임에 존재하는 블록의 개수이다.

[0047] 변환부(120)는 이전 프레임을 이용하여 학습하고, 인트라 예측부(110)에서 전달된 계수로부터 부터 잔차를 재구성한다. 재구성된 잔차는 프레임 버퍼(130)에 저장되고 다음 프레임에 대한 학습에 사용된다. 정리하면 최소 변환 과정에서 블록 별로 잔차 샘플에 대하여 변환을 위한 변환 함수가 생성된다. $i-1$ 번째 프레임에 대한 변환 함수가 이후 i 번째 프레임에 대한 변환 함수 학습에 사용된다.

[0048] 3. 최적의 변환 기법 선택

[0049] 인코더는 올 왜곡 최적기법(rate distortion optimization)을 사용하여 DCT 또는 전술한 학습 기반의 변환 기법을 선택할 수 있다.

[0050] 아래의 수학적 식 7을 이용하여 라그랑지안 비용(Lagrangian cost)를 최소화 하는 기법을 선택한다.

수학적 식 7

$$J = \{|x - \mathbf{V}^T u_q|^2 + \mu(R_C + R_H)\}$$

[0052] 수학적 식 7에서 첫 번째 항은 평균 제곱 에러(mean square error)로 표현한 왜곡에 관한 것이고, 두 번째 항에서 R_H 는 DCT 모드 또는 전술한 학습 기반의 변환 모드를 나타내는 모드 지시자(indicator)이다. R_C 는 추정된 비트율(bir rates)을 의미한다. 정확한 비트율을 추정하면 보다 효과적인 변환 모드 선택이 가능할 것이다. 비트율 추정에는 다양한 기법이 사용될 수 있다.

[0053] 정리하면 아래의 수학적 식 8을 이용하여 DCT 모드(T_{DCT}) 또는 제안한 학습 기반의 변환 모드(T_{PROP})를 선택할 수 있다.

수학적 식 8

$$\mathbf{T}^* = \arg \min_{\mathbf{V} \in \{T_{DCT}, T_{PROP}\}} \{J\}$$

[0055] 전술한 동영상 인코더의 동작을 정리한다. 도 3은 동영상 인코딩 방법(200)에 대한 순서도의 예이다. 동영상 인코딩 과정에서 일반적으로 수행되는 동작을 모두 도시하지 않고, 필요한 부분만을 도시하였다. 동영상 인코딩에서 수행되는 다른 과정들은 해당 분야의 통상의 지식을 가진 자가 이해하는 과정에 해당할 것이다.

[0056] 현재 인코딩하려는 프레임이 i 번째 프레임이고, 직전의 프레임은 $i-1$ 번째 프레임이다. 인코더는 $i-1$ 번째 프레임에 대한 인코딩을 완료한 상태이고, 이제 i 번째 프레임에 대한 인코딩을 수행하려는 상황이다.

[0057] 인코더는 $i-1$ 번째 프레임도 전술한 학습 기반의 최소 변환을 이용하여 인코딩하였다(210). 따라서 인코더는 $i-1$ 번째 프레임을 구성하는 각 블록에 대한 잔차 신호를 변환한 정보를 보유하고 있다.

[0058] 인코더는 i 번째 프레임에 대한 인트라 예측을 수행하여 잔차 신호를 생성한다(220). 전술한 바와 같이 여기서 인트라 예측은 방향 인트라 예측에 해당한다.

[0059] 이후 인코더는 도 1에서 설명한 바와 같이 현재 i 프레임에 대한 블록에 대해 인트라 예측 수행과 동시에 계수 블록 패치를 생성한다(230).

[0060] 인코더는 $i-1$ 번째 프레임에 대한 최소 변환에 사용된 변환 함수 및 현재 i 프레임에 대한 계수 블록 패치를 이용하여 i 번째 프레임의 잔차 신호에 대한 최소 변환을 수행한다(240). 240 과정은 전술한 수학적 식 4 및 수학적 식 5를 적용하여 수행할 수 있다.

[0061] 나아가 인코더는 올 왜곡 최적화 기법을 이용하여 DCT 또는 전술한 학습 기반의 최소 변환 중 효율이 좋은 기법을 선택적으로 사용할 수도 있다. 이와 관련해서는 수학적 식 7 내지 수학적 식 8에서 설명한 바 있다.

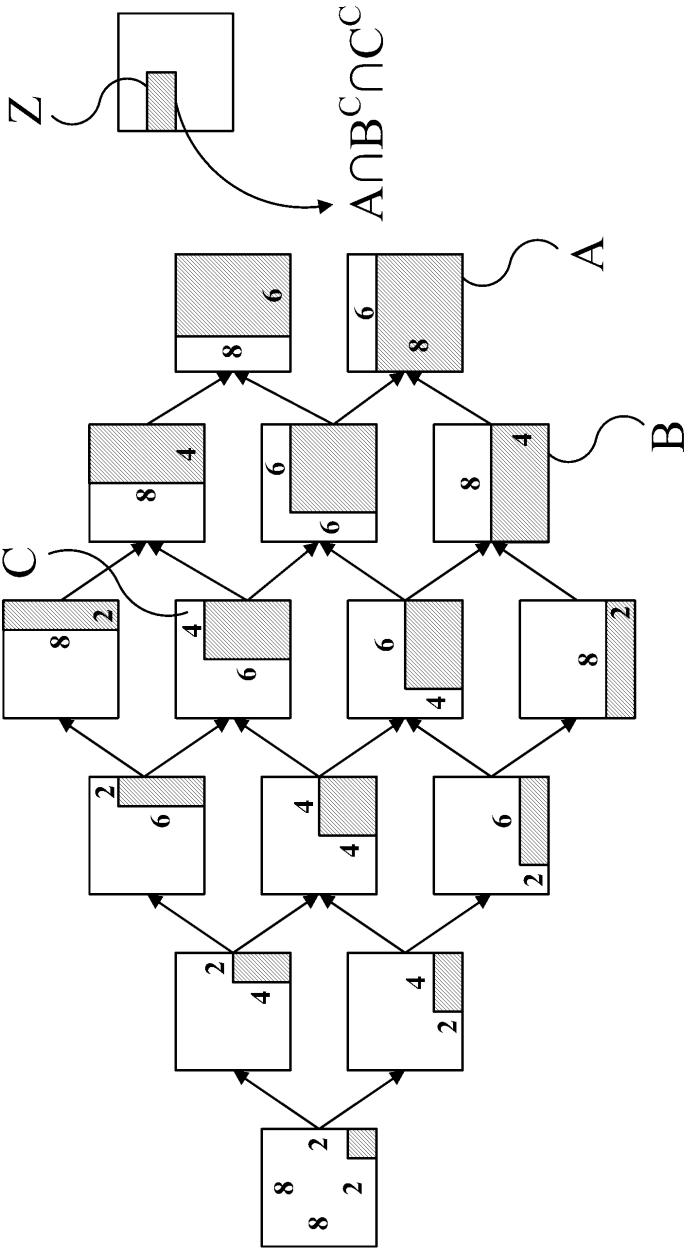
- [0062] 또한 디코더는 전술한 학습 기반의 회소 변환을 사용하여 인코딩한 신호를 디코딩할 수 있다. 도 4는 동영상 디코딩 방법(300)에 대한 순서도의 예이다. 도 4에서도 일반적으로 디코딩 과정에서 수행되는 단계는 생략하고, 필요한 부분만을 도시하였다.
- [0063] 동영상 디코딩 방법(300)은 동영상 인코딩 방법(200)의 역순으로 수행된다. 따라서 동영상 디코딩 방법(300)은 $i-1$ 번째 프레임에 대한 회소 변환 함수 또는 역 회소 변환 함수를 저장해야 한다(310).
- [0064] 디코더는 일반적인 디코딩과정과 같이 현재 i 프레임에 대한 인코딩된 신호를 역 엔트로피 코딩하고, 역 양자화를 수행한다(320).
- [0065] 디코더는 역 양자화후 생성된 잔차 신호 중 0 이 아닌 계수를 갖는 픽셀로 구성된 계수 블록 패치를 생성한다(330). 물론 계수 블록 패치는 인코더로부터 별도로 전달될 수도 있을 것이다.
- [0066] 디코더는 이제 $i-1$ 번째 프레임에 대한 회소 변환 함수 또는 역 회소 변환 함수 및 300 과정에서 생성한 계수 블록 패치를 이용하여 i 번째 프레임에 대한 역 회소 변환을 수행한다(340). 회소한 계수에 대한 역 회소 변환은 일정한 잔차 신호를 생성한다. 여기서 생성된 잔차 신호는 인코딩 과정에서 인트라 예측으로 생성된 잔차 신호에 대응된다.
- [0067] 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.

부호의 설명

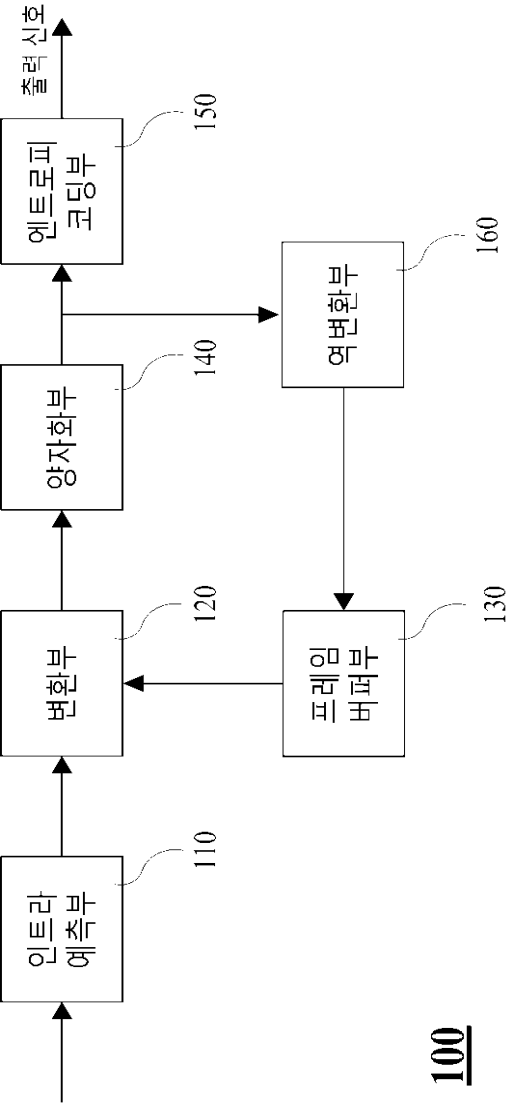
- | | | |
|--------|------------|---------------|
| [0068] | 100 : 인코더 | 110 : 인트라 예측부 |
| | 120 : 변환부 | 130 : 프레임 버퍼부 |
| | 140 : 양자화부 | 150 : 엔트로피코딩부 |
| | 160 : 역변환부 | |

도면

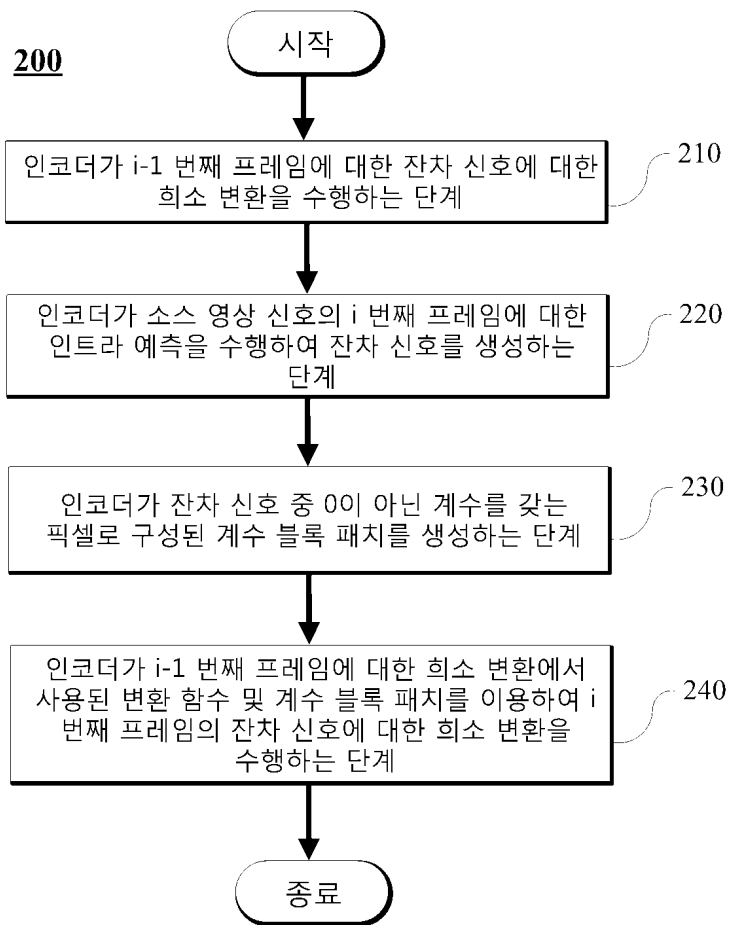
도면1



도면2



도면3



도면4

