



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2013년12월20일
(11) 등록번호 10-1344193
(24) 등록일자 2013년12월16일

(51) 국제특허분류(Int. Cl.)
G06T 9/00 (2006.01) H04N 7/30 (2006.01)
H04N 7/32 (2006.01)
(21) 출원번호 10-2008-7016381
(22) 출원일자(국제) 2007년01월08일
심사청구일자 2012년01월06일
(85) 번역문제출일자 2008년07월04일
(65) 공개번호 10-2008-0085156
(43) 공개일자 2008년09월23일
(86) 국제출원번호 PCT/US2007/000195
(87) 국제공개번호 WO 2007/081752
국제공개일자 2007년07월19일
(30) 우선권주장
60/756,846 2006년01월06일 미국(US)
(뒷면에 계속)
(56) 선행기술조사문헌
JP11316708 A
JP2005070938 A
KR1020020031193 A
전체 청구항 수 : 총 22 항

(73) 특허권자
마이크로소프트 코포레이션
미국 워싱턴주 (우편번호 : 98052) 레드몬드 원
마이크로소프트 웨이
(72) 발명자
설리반, 게리, 제이.
미국 98052-6399 워싱턴주 레드몬드 원 마이크로
소프트 웨이
(74) 대리인
제일특허법인

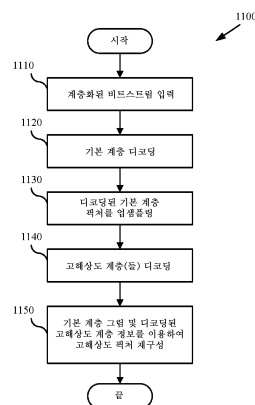
심사관 : 이주미

(54) 발명의 명칭 다중 해상도 비디오 코딩 및 디코딩을 위한 리샘플링 및픽처 크기조정 동작

(57) 요약

공간적으로 크기조정가능한 비디오 코딩 및 디코딩과 같은 적용에서 픽처 크기조정을 위한 고정밀도 위치 계산을 위한 기술 및 도구가 설명된다. 일 양태에서, 비디오 픽처의 리샘플링은 리샘플링 크기조정 인수에 따라 수행된다. 리샘플링은 리샘플링된 어레이에서 위치 i, j 에서 샘플값의 계산을 포함한다. 상기 계산은 유도된 수평 또는 수직 서브샘플 위치 x 또는 y 를 업샘플링 크기조정 인수의역 (대략적이거나 정확한)에 $2n$ 값을 곱함으로써 부분적으로 한 값에 근사시키는 것을 포함하는 방식으로 계산하는 것을 포함한다. 상기 근사시키는 것은 반올림 또는 인접한 정수에 근사시키는 올림이나 내림 함수와 같은 다른 종류의 근사일 수 있다. 상기 샘플값은 필터를 이용하여 보간된다.

대표도 - 도11



(30) 우선권주장

60/786,573 2006년03월27일 미국(US)

60/829,515 2006년10월13일 미국(US)

특허청구의 범위

청구항 1

비디오 인코더 또는 디코더를 구현하는 컴퓨팅 장치를 사용하여 다해상도 비디오 코딩 또는 디코딩을 하기 위한 리샘플링 방법으로서,

상기 컴퓨팅 장치는 프로세싱 유닛과 메모리를 포함하고,

상기 방법은 상기 비디오 인코더 또는 디코더를 구현하는 컴퓨팅 장치로 수평 리샘플링 크기조정 인수(horizontal resampling scale factor)에 따라 이미지 데이터의 리샘플링을 수행하는 단계 - 상기 리샘플링은 리샘플링되는 어레이 내 수평 위치 i 에서의 샘플값의 계산을 포함함 - 를 포함하며,

상기 계산은 수평 서브샘플 위치 x 를 식 $x=(i*C+D)>>S$ 와 결과가 수학적으로 등가인 방식으로 계산하는 단계 - 여기서 C 는 상기 수평 리샘플링 크기조정 인수의 역(inverse)을 곱한 $2^{S_{FF}}$ 에 등가인 값의 근사값 구하기(approximating)를 행함으로써 유도되고, F , C , D 및 S 는 정수값이되, F 는 상기 수평 서브샘플 위치 x 의 분수 구성요소(fractional component) 내 비트의 수에 기초하고, D 는 오프셋이며, S 는 동적 범위 및 정밀도를 설정함 - 를 포함하는

리샘플링 방법.

청구항 2

비디오 인코더 또는 디코더를 구현하는 컴퓨팅 장치를 사용하여 다중 해상도 비디오 코딩 또는 디코딩을 하기 위한 리샘플링 방법으로서,

상기 컴퓨팅 장치는 프로세싱 유닛과 메모리를 포함하고,

상기 방법은 상기 비디오 인코더 또는 디코더를 구현하는 컴퓨팅 장치로 수직 리샘플링 크기조정 인수에 따라 이미지 데이터의 리샘플링을 수행하는 단계 - 상기 리샘플링은 리샘플링되는 어레이 내 수직 위치 j 에서의 샘플값의 계산을 포함함 - 를 포함하며,

상기 계산은 수직 서브샘플 위치 y 를 식 $y=(j*C+D)>>S$ 와 결과가 수학적으로 등가인 방식으로 계산하는 단계 - 여기서 C 는 상기 수직 리샘플링 크기조정 인수의 역을 곱한 $2^{S_{VF}}$ 에 등가인 값의 근사값 구하기를 행함으로써 유도되고, F , C , D 및 S 는 정수값이되, F 는 상기 수직 서브샘플 위치 y 의 분수 구성요소 내 비트의 수에 기초하고, D 는 오프셋이며, S 는 동적 범위 및 정밀도를 설정함 - 를 포함하는

리샘플링 방법.

청구항 3

비디오 인코더 또는 디코더를 구현하는 컴퓨팅 장치를 사용하여 다중 해상도 비디오 코딩 또는 디코딩을 하기 위한 업샘플링(upsampling) 방법으로서,

상기 컴퓨팅 장치는 프로세싱 유닛과 메모리를 포함하고,

상기 방법은 상기 비디오 인코더 또는 디코더를 구현하는 컴퓨팅 장치로 수평 업샘플링 크기조정 인수 및 수직 업샘플링 크기조정 인수에 따라 비디오 픽처(picture)의 업샘플링을 수행하는 단계 - 상기 업샘플링은 업샘플링되는 어레이 내 수평 위치 i 및 수직 위치 j 에서의 보간되는(interpolated) 샘플값의 계산을 포함함 - 를 포함하며,

상기 계산은,

수평 서브샘플 위치 x 를 식 $x=(i*C+D)>>S$ 와 결과가 수학적으로 등가인 방식으로 계산하는 단계 - 여기서 C 는 상기 수평 업샘플링 크기조정 인수의 역을 곱한 $2^{S_{HF}}$ 에 등가인 값의 근사값 구하기를 행함으로써 유도되고, F , C , D 및 S 는 정수값이되, F 는 상기 수평 서브샘플 위치 x 의 분수 구성요소 내 비트의 수에 기초하고, D 는 오프셋이며, S 는 동적 범위 및 정밀도를 설정함 - ;

수직 서브샘플 위치 y 를 식 $y=(j*C+D)>>S$ 와 결과가 수학적으로 등가인 방식으로 계산하는 단계 - 여기서 C 는 상기 수직 업샘플링 크기조정 인수의 역을 곱한 2^{SF} 에 등가인 값의 근사값 구하기를 행함으로써 유도되고, F , C , D 및 S 는 정수값이되, F 는 상기 수직 서브샘플 위치 y 의 분수 구성요소 내 비트의 수에 기초하고, D 는 오프셋이며, S 는 동적 범위 및 정밀도를 설정함 - ; 및

상기 수평 서브샘플 위치 x 및 상기 수직 서브샘플 위치 y 에 의해 정의되는위치 x , y 에서 샘플값을 보간하는 단계를 포함하는

리샘플링 방법.

청구항 4

제 3 항에 있어서,

상기 계산은,

상기 수평 서브샘플 위치 x 의 F 최하위 비트에 기초하여 수평 리샘플링 필터를 선택하는 단계; 및

상기 수평 서브샘플 위치 x 의 나머지 상위 비트에 기초하여 필터링될 저해상도 샘플을 선택하는 단계를 더 포함하고,

상기 위치 x , y 에서 샘플값을 보간하는 단계는, 상기 선택된 저해상도 샘플에 기초하여 그리고 상기 선택된 수평 리샘플링 필터를 이용하여 상기 샘플값을 보간하는 단계를 포함하는

리샘플링 방법.

청구항 5

제 4 항에 있어서,

상기 수평 서브샘플 위치 x 의 상기 F 최하위 비트의 적어도 하나의 값에 대해 적용되는 수평 리샘플링 필터는 3 이상의 논-제로(non-zero) 필터 탭 값을 갖는 유한 임펄스 응답 필터(finite impulse response filter)인

리샘플링 방법.

청구항 6

제 4 항에 있어서,

상기 수평 서브샘플 위치 x 의 상기 F 최하위 비트에 대해 0이 아닌 모든 값에 대해 적용되는 수평 리샘플링 필터는 4개의 논-제로 필터 탭 값을 갖는 유한 임펄스 응답 필터인

리샘플링 방법.

청구항 7

제 3 항에 있어서,

상기 계산은,

상기 수직 서브샘플 위치 y 의 F 최하위 비트에 기초하여 수직 리샘플링 필터를 선택하는 단계; 및

상기 수직 서브샘플 위치 y 의 나머지 상위 비트에 기초하여 필터링될 저해상도 샘플을 선택하는 단계를 더 포함하고,

상기 위치 x , y 에서 샘플값을 보간하는 단계는, 상기 선택된 저해상도 샘플에 기초하여 그리고 상기 선택된 수직 리샘플링 필터를 이용하여 상기 샘플값을 보간하는 단계를 포함하는

리샘플링 방법.

청구항 8

제 7 항에 있어서,

상기 수직 서브샘플 위치 y 의 상기 F 최하위 비트의 적어도 하나의 값에 대해 적용되는 수직 리샘플링 필터는 3개 이상의 논-제로 필터 탭 값을 갖는 유한 임펄스 응답 필터인

리샘플링 방법.

청구항 9

제 7 항에 있어서,

상기 수직 서브샘플 위치 y 의 상기 F 최하위 비트에 대해 0이 아닌 모든 값에 대해 적용되는 수직 리샘플링 필터는 4개의 논-제로 필터 탭 값을 갖는 유한 임펄스 응답 필터인

리샘플링 방법.

청구항 10

제 3 항에 있어서,

상기 업샘플링은 하나 이상의 Mitchell-Netravalli 리샘플링 필터를 이용하여 수행되는

리샘플링 방법.

청구항 11

제 3 항에 있어서,

상기 업샘플링은 하나 이상의 Catmull-Rom 리샘플링 필터를 이용하여 수행되는

리샘플링 방법.

청구항 12

제 3 항에 있어서,

F , C , D 또는 S 의 수직 또는 수평 값 중 적어도 하나는 상기 샘플값이 크로마 샘플값인지 루마 샘플값인지에 따라 달라지는

리샘플링 방법.

청구항 13

제 3 항에 있어서,

상기 식 $x=(i*C+D)>>S$ 와 결과가 수학적으로 등가인 방식은 식 $x=((i*C+D)>>S)+E$ 의 구현 - 여기서, E 는 제2 오프셋임 - 을 포함하는

리샘플링 방법.

청구항 14

제 3 항에 있어서,

상기 업샘플링은 대역폭 제어 파라미터에 의해 제어되는 필터 탭값을 갖는 하나 이상의 리샘플링 필터를 이용하여 수행되는

리샘플링 방법.

청구항 15

제 3 항에 있어서,

상기 업샘플링은 계층화된 공간적으로 크기조정가능한(layered spatially-scalable) 비디오 디코딩 프로세스에서 수행되는

리샘플링 방법.

청구항 16

제 3 항에 있어서,

상기 업샘플링은 계층화된 공간적으로 크기조정가능한 비디오 인코딩 프로세스에서 수행되는 리샘플링 방법.

청구항 17

제 3 항에 있어서,

상기 업샘플링은 기준 픽처 리샘플링에 대해 수행되는 리샘플링 방법.

청구항 18

제 3 항에 있어서,

상기 F의 값은 4 이고 상기 S의 값은 12인 리샘플링 방법.

청구항 19

제 3 항에 있어서,

상기 근사값 구하기는 반올림하는 것을 포함하는 리샘플링 방법.

청구항 20

제 3 항에 있어서,

상기 역(inverse)은 근사시킨(approximate) 역인 리샘플링 방법.

청구항 21

제 3 항에 있어서,

상기 F, C, D 및 S 정수값 중 적어도 하나는 상기 수직 서브샘플 위치 y의 계산에 대한 경우가 상기 수평 서브샘플 위치 x의 계산에 대한 경우와 상이한 리샘플링 방법.

청구항 22

비디오 인코더 또는 디코더를 구현하는 컴퓨팅 장치를 사용하여 다중 해상도 비디오 코딩 또는 디코딩을 하기 위한 리샘플링 방법으로서,

상기 컴퓨팅 장치는 프로세싱 유닛과 메모리를 포함하고,

상기 방법은 상기 비디오 인코더 또는 디코더를 구현하는 컴퓨팅 장치로, 수평 업샘플링 크기조정 인수 및 수직 업샘플링 크기조정 인수에 따라 비디오 픽처의 업샘플링을 수행하는 단계 - 상기 업샘플링은 업샘플링되는 어레이 내 수평 위치 i 및 수직 위치 j에서의 보간되는 샘플값의 계산을 포함함 - 를 포함하며,

상기 계산은,

수평 서브샘플 위치 x를 식 $x = ((2^J * i + Q) * C + D) \gg S$ 와 결과가 수학적으로 등가인 방식으로 계산하는 단계 - 여기서 C는 상기 수평 업샘플링 크기조정 인수의 역을 곱한 2^{S+P} 에 등가인 값의 근사값 구하기를 행함으로써 유도되고, F, C, D, S, J 및 Q는 정수이되, F는 상기 수평 서브샘플 위치 x의 분수 구성요소 내 비트의 수에

기초하고, D 및 Q는 오프셋들이며, J는 상기 리샘플링이 루마에 대한 것인지 또는 크로마에 대한 것인지에 기초하고, S는 동적 범위 및 정밀도를 설정함 - ;

수직 서브샘플 위치 y를 식 $y=((2^J * j + Q) * C + D) \gg S$ 와 결과가 수학적으로 등가인 방식으로 계산하는 단계 - 여기서 C는 상기 수직 업샘플링 크기조정 인수의 역을 곱한 2^{SF} 에 등가인 값의 근사값을 구하기를 행함으로써 유도되고, F, C, D, S, J 및 Q는 정수이되, F는 상기 수직 서브샘플 위치 y의 분수 구성요소 내 비트의 수에 기초하고, D 및 Q는 오프셋들이며, J는 상기 리샘플링이 루마에 대한 것인지 또는 크로마에 대한 것인지에 기초하고, S는 동적 범위 및 정밀도를 설정함 - ; 및

상기 수평 서브샘플 위치 x 및 상기 수직 서브샘플 위치 y에 의해 정의되는 위치 x, y 에서 샘플값을 보간하는 단계를 포함하는

리샘플링 방법.

명세서

기술분야

[0001] 디지털 비디오를 인코딩/디코딩하는 기술 및 도구가 설명된다.

배경 기술

[0002] DVD, 인터넷을 통한 음악 전송, 및 디지털 카메라가 널리 대중화됨에 따라, 디지털 미디어가 흔하게 되었다. 엔지니어들은 다양한 기술을 이용하여 디지털 오디오, 비디오 및 이미지를 품질을 유지하면서도 효율적으로 처리한다. 이러한 기술을 이해하려면, 오디오, 비디오 및 이미지 정보가 컴퓨터에서 어떻게 표현되고 처리되는지 이해하는 것이 도움이 된다.

[0003] I. 컴퓨터에서 미디어 정보의 표현

[0004] 컴퓨터는 미디어 정보를 그 정보를 표현하는 일련의 숫자로서 처리한다. 예를 들면, 단일 숫자는 픽처(picture)의 각 기본 소영역에 대한 적, 녹, 청과 같은 컬러 구성요소의 밝기의 세기 또는 세기를 표현할 수 있어서, 이 픽처의 디지털 표현은 이러한 숫자들의 하나 이상의 어레이로 구성된다. 이러한 숫자 각각은 샘플이라 할 수도 있다. 컬러 이미지의 경우, 각각의 기본 영역의 컬러를 표현하기 위해 둘 이상의 샘플을 사용하는 것이 통상적이고, 전형적으로 3개의 샘플이 사용된다. 하나의 기본 영역에 대한 이들 샘플의 집합을 픽셀이라 할 수 있는데, "픽셀"은 "픽처 요소"의 개념을 가리키는 줄임말이다. 예컨대, 하나의 픽셀은 기본 영역을 표현하는 데 필요한 적색광, 녹색광, 청색광의 세기를 나타내는 3개의 샘플로 구성될 수 있다. 이러한 픽셀 타입을 RGB 픽셀이라 한다. 샘플 깊이, 해상도, 및 프레임 속도 (비디오의 경우)를 포함하여 몇 개의 요인이 미디어 정보의 품질에 영향을 미친다.

[0005] 샘플 깊이는 샘플을 표현하는 데 사용될 수 있는 숫자 범위를 나타내는 통상적으로 비트로 측정되는 속성이다. 샘플에 대해 더 많은 값들이 가능할 때, 숫자가 세기에서의 더 미묘한 변화 및/또는 더 큰 범위의 값들을 표현(capture)할 수 있기 때문에 품질이 더 높을 수 있다. 해상도는 일반적으로 일정 지속시간 (오디오의 경우) 또는 공간 (이미지 또는 개별 비디오 픽처의 경우)에 대한 샘플의 수를 말한다. 공간 해상도가 더 높은 이미지는 다른 이미지보다 더 생생해 보이고 더 식별가능한 유용한 세부사항을 포함하는 경향이 있다. 프레임 속도는 비디오에 대한 시간적 해상도에 대한 통상적인 용어이다. 프레임 속도가 더 높은 비디오는 다른 비디오보다 자연 물체의 평탄한 움직임(motion)을 더 잘 모방하는 경향이 있고, 마찬가지로 시간적 차원에 있어서 더 많은 세부사항을 포함한다고 간주될 수 있다. 이 모든 요인에 대해서, 표 1 에 도시된 바와 같이, 고품질에 대한 절충(tradeoff)은 샘플 깊이, 해상도, 및 프레임 속도를 표현하는 데 필요한 비트 전송률(bit rate)의 면에서 정보를 저장하고 전달하는 비용이다.

[0006]

픽셀당 비트수 (샘플 깊이 × 픽셀당 샘플수)	해상도 (픽셀수, 폭×높이)	프레임 속도 (초당 프레임수)	비트 전송률 (초당 백만 비트)
8(0-225 값, 단색)	160×120	7.5	1.2
24(0-255 값, RGB)	320×240	15	27.6
24(0-255 값, RGB)	640×480	30	221.2
24(0-255 값, RGB)	1280×720	60	1327.1

- [0007] 표 1: 원시 비디오의 서로 다른 품질 수준에 대한 비트 전송률
- [0008] 고품질 비디오 (HDTV 등)를 저장 및 전송하기 위해 필요한 높은 비트 전송률에도 불구하고, 회사들 및 소비자들은 고품질 콘텐츠를 생성, 배포, 및 재생하기 위해 점점 더 컴퓨터에 의존한다. 이러한 이유로, 엔지니어들은 디지털 미디어의 비트 전송률을 줄이기 위해 압축 (소스 코딩 또는 소스 인코딩이라고도 함)을 이용한다. 압축은 정보를 더 낮은 비트 전송률 형태로 변환시킴으로써 정보 저장 및 전송 비용을 줄인다. 압축은 손실이 없을 수 있는데, 이 경우, 비디오의 품질은 손상되지 않지만 비트 전송률의 감소는 비디오의 복잡도에 의해 제한된다. 그렇지 않으면, 압축은 손실이 있을 수 있는데, 이 경우, 비디오의 품질은 손상되지만 비트 전송률의 감소는 더욱 현저하다. 압축풀기 (디코딩이라고도 함)는 압축된 형태로부터 원래 정보의 버전을 재구성한다. "코덱"은 인코더/디코더 시스템이다.
- [0009] 일반적으로, 비디오 압축 기술은 "인트라(intra)" 압축 및 "인터(inter)" 압축 또는 예측 압축을 포함한다. 비디오 픽처의 경우, 인트라 압축 기술은 개별 픽처들을 압축한다. 인터 압축 기술은 이전 및/또는 이후 픽처를 참조하여 픽처를 압축한다.
- [0010] II. 다중 해상도 비디오 및 공간 확장성
- [0011] 표준 비디오 인코더는 대상 비트 전송률이 임의의 문턱값 아래로 떨어질 때 성능이 현저히 떨어진다. 양자화 및 다른 손실 처리단계에서 왜곡이 생긴다. 낮은 비트 전송률에서, 고주파수 정보는 심하게 왜곡되거나 완전히 손실될 수 있다. 그 결과, 심각한 아티팩트(artifacts)가 발생하여 재구성된 비디오의 품질의 현저한 저하를 초래할 수 있다. 전송 및 처리 기술이 발전함에 따라 이용가능한 비트 전송률이 증가하더라도, 제한된 비트 전송률로 높은 시각적 품질을 유지하는 것이 여전히 비디오 코덱 설계의 기본 목표이다. 기존의 코덱은 제한된 비트 전송률로 시각적 품질을 개선하기 위해 여러가지 방법을 이용한다.
- [0012] 다중 해상도 코딩은 서로 다른 공간 해상도로 비디오의 인코딩을 허용한다. 감소된 해상도 비디오는 정보 손실의 대가로 현저히 더 낮은 비트 전송률로 인코딩될 수 있다. 예를 들면, 종래의 비디오 인코더는 전 해상도 (full-resolution) 비디오를 다운샘플링하여 (다운샘플링 필터를 이용), 이를 수직 및/또는 수평 방향으로 감소된 해상도로 인코딩할 수 있다. 각 방향으로 해상도를 절반으로 줄이는 것은 인코딩된 픽처 크기의 치수를 절반으로 감소시킨다. 인코더는 감소된 해상도 코딩을 디코더로 신호로 보낸다(signal). 디코더는 감소된 해상도 인코딩을 나타내는 정보를 수신하고 이 수신된 정보로부터 디스플레이 전에 이 픽처 크기를 증가시키기 위해 어떻게 이 감소된 해상도 비디오의 업샘플링 방법을 (업샘플링 필터를 이용)확인한다. 그러나, 인코더가 비디오 픽처를 다운샘플링하고 인코딩할 때 손실되었던 정보는 업샘플링된 픽처에서도 여전히 보이지 않는다.
- [0013] 공간적으로 크기조정가능한(spatially scalable) 비디오는 다층 계층 접근법을 이용하여, 인코더가 하나 이상의 인핸스먼트(enhancement) 계층에서 원본 비디오로부터의 더 높은 해상도 정보를 보유하면서 기본 계층(base layer)에서 공간 해상도 (및 비트 전송률)를 감소시킬 수 있게 한다. 예컨대, 기본 계층 인트라-픽처는 감소된 해상도로 코딩될 수 있는 반면에, 수반하는 인핸스먼트 계층 인트라-픽처는 더 높은 해상도로 코딩될 수 있다. 마찬가지로, 기본 계층 예측 픽처는 인핸스먼트 계층 예측 픽처를 수반할 수 있다. 디코더는 더 낮은 해상도 재구성 픽처를 얻기 위해 더 낮은 해상도로 기본 계층 픽처만을 디코딩하거나, 더 높은 해상도 재구성 픽처를 얻기 위해 기본 계층 및 인핸스먼트 계층을 디코딩하도록 선택할 수 있다 (비트 전송률 제약 및/또는 다른 기준에 기초함). 기본 계층이 디스플레이된 픽처보다 더 낮은 해상도로 인코딩될 때 (다운샘플링이라고도 함), 인코딩된 픽처 크기는 디스플레이된 픽처보다 실제로 더 작다. 디코더는 재구성된 픽처를 크기조정(resizing)하기 위해 계산을 수행하고 업샘플링 필터를 사용하여 상기 재구성된 픽처의 적절한 위치에서 보간된(interpolated) 샘플 값들을 생성한다. 그러나, 공간적으로 크기조정가능한(scalable) 비디오를 사용하는 종래의 코덱은 유연성없는 업샘플링 필터와 부정확하거나 값비싼 (계산 시간 또는 비트 전송률 면에서) 픽처 크기조정 기술때문에 손해를 입었다.
- [0014] 디지털 비디오에 대한 비디오 압축 및 압축풀기의 중요성을 감안하면, 비디오 압축 및 압축풀기 분야가 상당히 개발되어 있음은 놀라운 일이 아니다. 그러나, 종래의 비디오 압축 및 압축풀기 기술의 이점이 무엇이던간에, 이들은 다음에 설명하는 기술 및 도구의 장점이 없다.

발명의 상세한 설명

- [0015] 본 발명의 요약은 개념의 발체를 간단한 형태로 소개하기 위해 제공되고, 아래 발명의 실시를 위한 형태에서 더

구체적으로 설명된다. 본 발명의 요약은 본 발명의 주요 특징이나 본질적인 특징을 확인하거나, 본 발명의 범위를 한정할 목적으로 이용되는 것이 아니다.

[0016] 요약하면, 발명의 실시를 위한 형태는 다중 해상도 및 계층화된 공간적으로 크기조정가능한 비디오 코딩 및 디코딩을 위한 다양한 기술 및 도구에 대한 것이다.

[0017] 예를 들면, 발명의 실시를 위한 형태는 공간적으로 크기조정가능한 비디오 코딩 및 디코딩과 같은 응용에서 픽처 크기조정(resizing)을 위한 고정밀 위치계산을 위한 다양한 기술 및 도구에 대한 것이다. 공간적으로 크기조정가능한 비디오 코딩 및 디코딩과 같은 응용에서 픽처 크기조정을 위한 고정밀 위치계산을 위한 다양한 기술 및 도구가 설명된다. 일 양태에서, 비디오 픽처의 리샘플링은 리샘플링 크기조정 인수에 따라 수행된다. 상기 리샘플링은 리샘플링된 어레이의 위치 i, j 에서 샘플값의 계산을 포함한다. 상기 계산은 상기 업샘플링 크기조정 인수의 역(inverse) (근사하거나(approximate) 또는 정확한)에 2^n 값을 곱함으로써 (또는 2^n 값을 업샘플링 크기조정 인수 또는 업샘플링 크기조정 인수의 근사치로 나눔으로써) 어떤 값을 부분적으로 근사시키는 단계(approximating)를 포함하는 방식으로 도출된 수평 또는 수직 부분-샘플 위치 x 또는 y 를 계산하는 것을 담당한다. 지수 n 은 분수 구성요소(fractional component)에서 비트의 수를 나타내는 정수 F 를 포함하는 2개의 정수의 합일 수 있다. 상기 근사시키는 단계는 반올림(rounding), 또는 바로 가까운의 정수에 근사시키는 올림(ceiling) 또는 내림(floor) 함수와 같은 다른 종류의 근사단계일 수 있다. 샘플 값은 필터를 사용하여 보간된다.

[0018] 상기 설명된 기술의 대안은, 일 구현에서 샘플 위치 계산 프로세스나 그 복잡성을 크게 변경하지 않고서 계산에서 대략적으로 하나의 별도의 비트의 정밀도를 제공하는 변경된 샘플 위치 계산을 제공한다. 또다른 대안은 샘플 위치 계산이 4:2:2 및 4:4:4 샘플링 구조로 동작하는 방법에 관한 것이다. 이러한 샘플링 구조에 대한 이 대안 기술들은 크로마(chroma) 및 루마(luma) 샘플링 그리드의 해상도가 특정 디멘전에서 동일할 때마다 루마 및 크로마 샘플 위치 계산을 함께 고정시킨다.

[0019] 추가적인 특징 및 장점은 첨부도면을 참조하여 다양한 실시예들에 대한 다음의 상세한 설명으로부터 명확해질 것이다.

실시예

[0034] 이하, 다중 해상도 및 계층화된 공간적으로 크기조정가능한 비디오 코딩 및 디코딩 기술 및 도구에 대한 실시예들에 대해 설명하기로 한다.

[0035] 여기 설명된 다양한 기술 및 도구는 독립적으로 사용될 수 있다. 이 중 일부는 결합하여 사용될 수 있다 (예컨대, 결합된 인코딩 및/또는 디코딩 프로세스의 서로 다른 단계에서).

[0036] 다양한 기술들이 처리 액트(acts)의 흐름도를 참조하여 아래에 설명된다. 이 흐름도에 도시된 다양한 처리 액트는 더 적은 액트로 통합되거나 더 많은 액트로 분리될 수 있다. 간단하게 하기 위해, 다른 곳에서 설명된 액트에 대한 특정 흐름도에 도시된 액트의 관계는 흔히 도시되지 않는다. 많은 경우에, 흐름도에서의 액트는 재정의될 수 있다.

[0037] 상세한 설명의 많은 부분은 비디오 정보의 표현, 코딩 및 디코딩에 대해 논의한다. 비디오 정보를 표현, 코딩 및 디코딩하기 위해 여기 설명된 기술 및 도구는 오디오 정보, 정지(still) 이미지 정보 또는 다른 미디어 정보에 적용될 수 있다.

[0038] I. 컴퓨팅 환경 (Computing Environment)

[0039] 도 1은 상기 실시예들이 구현될 수 있는 적절한 컴퓨팅 환경 (100)의 일반화된 예를 도시한다. 이 컴퓨팅 환경 (100)은 사용 또는 기능의 범위에 대한 어떤 제한을 시사하려는 것이 아니며, 기술 및 도구는 다양한 범용 또는 특수목적의 컴퓨팅 환경에서 구현될 수 있다.

[0040] 도 1을 참조하면, 컴퓨팅 환경 (100)은 적어도 하나의 처리부 (110) 및 메모리 (120)를 포함한다. 도 1에서, 이 가장 기본적인 구성 (130)은 점선 내에 포함된다. 처리부 (110)는 컴퓨터 실행가능한 명령어를 실행하고 실행 또는 가상 프로세서일 수 있다. 다중처리 시스템에서, 다중 처리부는 처리 능력을 증가시키기 위해 컴퓨터 실행가능한 명령어를 실행한다. 메모리 (120)는 휘발성 메모리 (예컨대, 레지스터, 캐시, RAM), 비휘발성 메모리 (예컨대, ROM, EEPROM, 플래시 메모리), 또는 이 둘의 조합일 수 있다. 메모리 (120)는 상기 기술 및 도구 중 하나 이상을 이용하여 비디오 인코더 또는 디코더를 구현하는 소프트웨어 (180)를 저장한다.

- [0041] 컴퓨팅 환경은 추가적인 특징을 가질 수 있다. 예컨대, 컴퓨팅 환경 (100)은 저장 장치 (140), 하나 이상의 입력 장치 (150), 하나 이상의 출력장치 (160), 및 하나 이상의 통신 연결부 (170)를 포함한다. 버스, 컨트롤러, 또는 네트워크와 같은 상호접속 메커니즘 (도시되지 않음)은 컴퓨팅 환경 (100)의 구성요소들을 상호 연결시킨다. 일반적으로, 운영체제 소프트웨어 (도시되지 않음)는 컴퓨팅 환경 (100)에서 실행되는 다른 소프트웨어에 운영 환경을 제공하고, 컴퓨팅 환경 (100)의 구성요소들의 활동을 조정한다.
- [0042] 저장 장치 (140)는 이동식 또는 이동불가식일 수 있고, 자기 디스크, 자기 테이프 또는 카세트, CD-ROM, DVD, 플래시 메모리, 또는 정보를 저장하기 위해 사용될 수 있고 컴퓨팅 환경 (100) 내에서 액세스될 수 있는 기타 다른 매체를 포함한다. 저장 장치 (140)는 비디오 인코더 또는 디코더를 구현하는 소프트웨어 (180)를 위한 명령어를 저장한다.
- [0043] 입력장치 (150)는 키보드, 마우스, 펜, 터치 스크린, 또는 트랙볼과 같은 터치 입력장치, 음성입력장치, 스캐닝 장치, 또는 컴퓨팅 환경 (100)에 입력을 제공하는 또다른 장치일 수 있다. 오디오 또는 비디오 인코딩의 경우, 입력장치 (150)는 사운드 카드, 비디오 카드, TV 튜너 카드, 또는 오디오 또는 비디오 입력을 아날로그 또는 디지털 형태로 수신하는 유사한 장치, 또는 오디오 또는 비디오 샘플을 컴퓨팅 환경 (100)으로 읽어들이는 CD-ROM, CD-RW 또는 DVD일 수 있다. 출력장치 (160)는 디스플레이, 프린터, 스피커, CD- 또는 DVD-라이터 (writer), 또는 컴퓨팅 환경 (100)으로부터 출력을 제공하는 또다른 장치일 수 있다.
- [0044] 통신 연결부 (170)는 통신 매체를 통해 다른 컴퓨팅 엔티티(entity)로의 통신을 가능하게 한다. 통신 매체는 컴퓨터 실행가능한 명령어, 오디오 또는 비디오 입력 또는 출력, 또는 다른 데이터와 같은 정보를 변조된 데이터 신호로 전송한다. 변조된 데이터 신호는 신호내의 정보를 인코딩하는 방식으로 설정 또는 변경된 하나 이상의 특성을 갖는 신호이다. 예를 들면, 통신 매체는 전기적, 광학적, RF, 적외선, 음향적 또는 다른 캐리어로 구현되는 유선 또는 무선 기술들을 포함하지만 이에 한정되는 것은 아니다.
- [0045] 상기 기술 및 도구는 컴퓨터 판독가능한 매체의 일반적인 관점으로 설명될 수 있다. 컴퓨터 판독가능한 매체는 컴퓨팅 환경 내에서 액세스될 수 있는 임의의 이용가능한 매체이다. 예를 들면, 컴퓨팅 환경 (100)에서, 컴퓨터 판독가능한 매체는 메모리 (120), 저장 장치 (140), 통신 매체, 및 이들의 조합을 포함하지만 이에 한정되는 것은 아니다.
- [0046] 상기 기술 및 도구는, 하나 이상의 대상 실제 프로세서 또는 가상 프로세서 상에서 컴퓨팅 환경에서 실행되고 있는 프로그램 모듈에 포함된 것과 같은 컴퓨터 실행가능한 명령어의 일반적인 관점으로 설명될 수 있다. 일반적으로, 프로그램 모듈은 특정 태스크를 수행하거나 특정 추상 데이터 타입을 구현하는 루틴, 프로그램, 라이브러리, 오브젝트, 클래스, 컴포넌트, 데이터 구조 등을 포함한다. 이 프로그램 모듈의 기능은 다양한 실시예에서 희망하는 대로 프로그램 모듈들 사이에서 결합 또는 분리될 수 있다. 프로그램 모듈에 대한 컴퓨터 실행가능한 명령어는 로컬 또는 분산형 컴퓨팅 환경 내에서 실행될 수 있다.
- [0047] 설명을 위해, 상세한 설명은 컴퓨팅 환경에서의 컴퓨터 동작을 설명하기 위해 "인코드", "디코드" 및 "선택"과 같은 용어를 사용한다. 이 용어들은 컴퓨터에 의해 수행되는 연산들에 대한 고수준 추상개념들이고, 인간에 의해 수행되는 행위와 혼동해서는 안된다. 이 용어들에 대응하는 실제 컴퓨터 연산은 구현에 따라 가변적일 수 있다.
- [0048] II. 일예의 비디오 인코더 및 디코더
- [0049] 도 2는 본 발명의 바람직한 실시예들이 구현될 수 있는 일예의 비디오 인코더 (200)의 블록도이다. 도 3은 본 발명의 바람직한 실시예들이 구현될 수 있는 일반화된 비디오 디코더 (300)의 블록도이다.
- [0050] 인코더 (200) 및 디코더 (300) 내의 모듈 사이에 도시된 관계는 인코더 및 디코더에서의 정보의 일반적인 흐름을 도시하고, 단순화를 위해 다른 관계는 도시하지 않았다. 특히, 도 2 및 도 3은 비디오 시퀀스, 픽처, 슬라이스, 매크로블록, 블록 등에 사용되는 인코더 설정, 모드, 테이블 등을 나타내는 부차적인 정보는 통상 도시하지 않았다. 이러한 부차적인 정보는 일반적으로 부차적인 정보의 엔트로피 인코딩 후에 출력 비트스트림에서 보내진다. 출력 비트스트림의 형식은 구현에 따라 가변적일 수 있다.
- [0051] 인코더 (200) 및 디코더 (300)는 비디오 픽처를 처리하는데, 이는 비디오 프레임, 비디오 필드 또는 프레임과 필드의 조합일 수 있다. 픽처 및 매크로블록 수준에서의 비트스트림 구문(syntax) 및 의미(semantics)는 프레임 또는 필드가 사용되는지 여부에 의존할 수 있다. 또한 매크로블록 조직 및 전체 타이밍에 대한 변화가 있을 수 있다. 인코더 (200) 및 디코더 (300)는 블록기반이고 프레임에 대해 4:2:0 매크로블록 형식을 이용하는데,

각 매크로블록은 4개의 8×8 밝기 블록 (때로 하나의 16×16 매크로블록으로 취급됨) 및 2개의 8×8 색차(chrominance) 블록을 포함한다. 필드의 경우, 동일 또는 다른 매크로블록 조직 및 형식이 사용될 수 있다. 8×8 블록은 예컨대 주파수 변환 및 엔트로피 인코딩 단계와 같은 서로 다른 단계에서 더 분할될 수 있다. 일예의 비디오 프레임 조직에 대해 아래에서 더 상세히 설명하기로 한다. 다른 경우에, 인코더 (200) 및 디코더 (300)는 개체(object) 기반이고, 서로 다른 매크로블록 또는 블록 형식을 이용하거나, 8×8 블록 및 16×16 매크로블록이 아닌 다른 크기나 구성의 샘플 집합에 대한 연산을 수행한다.

[0052] 원하는 압축의 실행 및 타입에 따라, 인코더 또는 디코더의 모듈이 추가, 생략 또는 다중의 모듈로 분할, 다른 모듈과 결합, 및/또는 유사한 모듈로 대체될 수 있다. 다른 실시예에서, 서로 다른 모듈 및/또는 다른 모듈 구성을 갖는 인코더 또는 디코더가 하나 이상의 상기 기술을 수행한다.

[0053] A. 비디오 프레임 조직

[0054] 일부 구현에서, 인코더 (200) 및 디코더 (300)는 다음과 같이 조직된 비디오 프레임을 처리한다. 한 프레임은 비디오 신호의 공간 정보의 라인들을 포함한다. 순차(progressive) 비디오 스캐닝의 경우, 이 라인들은 동시에 샘플링된 장면(scene) 내용의 스냅샷을 나타내고 프레임의 최상단부터 최하단까지의 전체 장면을 커버하는 샘플을 포함한다. 순차 비디오 프레임은 도 4에 도시된 매크로블록 (400)과 같은 매크로블록으로 분할된다. 매크로블록 (400)은, 종래의 $4:2:0$ 매크로블록 형식 다음에, 4개의 8×8 밝기 블록 (Y_1 부터 Y_4) 및 상기 4개의 밝기 블록과 함께 위치하지만 수평 및 수직으로 절반의 해상도를 갖는 2개의 8×8 색차 블록을 포함한다. 8×8 블록은 예컨대 주파수 변환 (예컨대, 8×4 , 4×8 , 또는 4×4 DCT) 및 엔트로피 인코딩 단계와 같은 서로 다른 단계에서 더 분할될 수 있다. 순차 I-프레임은 인트라 코딩된(intra-coded) 순차 비디오 프레임인데, 여기서 "인트라"라는 용어는 이전에 코딩된 다른 픽처의 내용으로부터의 예측을 포함하지 않는 코딩 방법을 말한다. 순차 P-프레임은 현재의 픽처와는 시간적으로 다른 때에 하나 이상의 다른 픽처로부터의 예측을 이용하여 코딩된 순차 비디오 프레임이고 (때때로 일부 상황에서는 순방향(forward) 예측이라고도 함), 순차 B-프레임은 일부 영역에서 다중의 예측 값들의 평균 (가능하면 가중치)을 포함하는 인터 프레임 예측을 이용하여 코딩된 순차 비디오 프레임이다 (양방향 예측이라고도 함). 순차 P- 및 B-프레임은 인트라 코딩된 매크로블록은 물론 다양한 종류의 인터 프레임 예측 매크로블록을 포함할 수 있다.

[0055] 비월 비디오 프레임 스캐닝은 한 장면의 교번하는 일련의 두 종류의 스캔으로 구성되는데, 하나는 상위 필드로서 프레임의 짝수 라인 (0번 라인, 2번 라인, 4번 라인 등)을 포함하고, 다른 하나는 하위 필드로서 프레임의 홀수 라인 (1번 라인, 3번 라인, 5번 라인 등)을 포함한다. 이 2개의 필드는 일반적으로 2개의 서로 다른 스냅샷 순간을 표현한다. 도 5A는 비월 비디오 프레임 (500)의 일부를 도시하는데, 비월 비디오 프레임 (500)의 좌측 상단부에서 상위 필드 및 하위 필드의 교번하는 라인들을 포함한다.

[0056] 도 5B는 프레임 (530)으로서 인코딩/디코딩을 위해 조직된 도 5A의 비월 비디오 프레임 (500)을 도시한다. 비월 비디오 프레임 (500)은 도 4에 도시된 바와 같이 $4:2:0$ 형식을 이용하는 매크로블록 또는 매크로블록 (531 및 532)과 같은 다른 영역으로 분할된다. 밝기 평면에서, 각 매크로블록 (531 및 532)은 하위 필드로부터의 8개 라인과 교번하는 상위 필드로부터의 8개 라인, 즉 총 16 라인을 포함하고, 각 라인은 16개 샘플 길이이다. (이 픽처의 매크로블록 또는 다른 이러한 영역으로의 실제 조직 및 매크로블록 (531 및 532) 내의 밝기 블록 및 색차 블록의 배치는 도시되지 않았고, 실제로 서로 다른 인코딩 결정에 대해 그리고 서로 다른 비디오 코딩 설계마다 가변적일 수 있다.) 주어진 매크로블록 내에서, 상위 필드 정보 및 하위 필드 정보는 다양한 단계들(phases) 중 임의의 단계에서 함께 또는 별도로 코딩될 수 있다.

[0057] 비월 I-프레임은 2개의 필드를 포함하는 인트라 코딩된 비월 비디오 프레임이고, 여기서 각 매크로블록은 하나 또는 두개의 필드에 대한 정보를 포함한다. 비월 P-프레임은 인터 프레임 예측을 이용하여 코딩된 2개의 필드를 포함하는 비월 비디오 프레임이고, 여기서 각 매크로블록은 비월 B-프레임처럼 하나 또는 두개의 필드에 대한 정보를 포함한다. 비월 P- 및 B-프레임은 인트라 코딩된 매크로블록뿐만 아니라 다양한 종류의 인터 프레임 예측 매크로블록을 포함할 수 있다.

[0058] 도 5C는 필드 (560)로서 인코딩/디코딩을 위해 조직된 도 5A의 비월 비디오 프레임 (500)을 도시한다. 비월 비디오 프레임 (500)의 2개의 필드 각각은 매크로블록으로 분할된다. 상위 필드는 매크로블록 (561)과 같은 매크로블록으로 분할되고, 하위 필드는 매크로블록 (562)과 같은 매크로블록으로 분할된다. (매크로블록은 도 4에 도시된 바와 같이 $4:2:0$ 형식을 이용하고, 이 픽처의 매크로블록 또는 다른 이러한 영역으로의 조직 및 이 매크로블록 내의 밝기 블록 및 색차 블록의 배치는 도시되지 않았고 가변적일 수 있다.) 밝기 평면에서, 매크로블록 (561)은 상위 필드로부터의 16개 라인을 포함하고 매크로블록 (562)은 하위 필드로부터의 16개 라인을 포함

하는데, 각 라인은 16개의 샘플 길이이다.

- [0059] 비월 I-필드는 비월 비디오 프레임의 단일의 별개로 표현되는 필드이다. 비월 P-필드는 비월 B-필드처럼 인터-픽처 예측을 이용하여 코딩된 비월 비디오 프레임의 단일의 별개로 표현된 필드이다. 비월 P- 및 B-필드는 인트라 코딩된 매크로블록뿐만 아니라 서로 다른 종류의 인터-픽처 예측 매크로블록을 포함할 수 있다.
- [0060] 필드로서 인코딩/디코딩을 위해 조직된 비월 비디오 프레임은 서로 다른 필드 타입의 다양한 조합을 포함할 수 있다. 예를 들면, 이러한 프레임은 최상위 및 하위 필드에서 동일한 필드 타입 (I-필드, P-필드, 또는 B-필드)을 가지거나 또는 각 필드에서 서로 다른 필드 타입을 가질 수 있다.
- [0061] 픽처라는 용어는 일반적으로 원본의 소스 코딩 또는 재구성된 이미지 데이터의 프레임 또는 필드를 말한다. 순차-스캔 비디오의 경우, 픽처는 일반적으로 순차 비디오 프레임이다. 비월 비디오의 경우, 픽처는 상황에 따라 비월 비디오 프레임, 프레임의 상위 필드, 또는 프레임의 하위 필드를 말할 수 있다.
- [0062] 도 5D는 비디오 프레임의 각 필드에 대한 루마 샘플 위치에 대한 4:2:0 크로마 샘플 위치의 6가지 예시 공간 정렬을 도시한다.
- [0063] 다른 경우에는, 인코더 (200) 및 디코더 (300)는 개체 기반이고, 서로 다른 매크로블록 형식 (예컨대, 4:2:2 또는 4:4:4) 또는 블록 형식을 이용하고, 또는 8×8 블록 및 16×16 매크로블록이 아닌 다른 크기 또는 구성의 샘플 집합에 대한 연산을 수행한다.
- [0064] B. 비디오 인코더
- [0065] 도 2는 일예의 비디오 인코더 시스템 (200)의 블록도이다. 인코더 시스템 (200)은 현재 픽처 (205) (예컨대, 순차 비디오 프레임, 비월 비디오 프레임 또는 비월 비디오 프레임의 필드)을 포함하는 일련의 비디오 픽처를 수신하고, 압축 비디오 정보 (295)를 출력으로서 생성한다. 비디오 인코더의 특정 실시예는 일반적으로 상기 인코더 (200)의 변형 또는 보충 버전을 사용한다.
- [0066] 인코더 시스템 (200)은 인트라 코딩 (인트라) 픽처 (I-픽처) 및 인터-픽처 예측 (인터) 픽처 (P- 또는 B-픽처)에 대한 인코딩 프로세스를 사용한다. 설명을 위해, 도 2는 인코더 시스템 (200)을 통한 I-픽처용 경로 및 인터-픽처 예측 픽처용 경로를 도시한다. 인코더 시스템 (200)의 구성요소 중 다수는 I-픽처 및 인터-픽처 예측 픽처를 압축하기 위해 사용된다. 이들 구성요소에 의해 수행되는 해당 동작은 압축되는 정보의 종류에 따라 가변적일 수 있다.
- [0067] 인터-픽처 예측 픽처는 하나 이상의 다른 픽처 (일반적으로 기준 픽처라고 함)으로부터의 예측 (또는 차)의 면에서 표현된다. 예측 잔차(prediction residual)는 예측된 것과 원래 픽처 간의 차이이다. 이와 반대로, I-픽처는 다른 픽처에 상관없이 압축된다. I-픽처는 I-픽처 자신의 다른 부분으로부터의 데이터를 이용하여 I-픽처의 일부분을 예측하기 위하여 공간 예측 또는 주파수-도메인 예측 (즉, 인트라-픽처 예측)을 이용할 수 있다. 그러나, 간단하게 하기 위해, 이러한 I-픽처는 이 설명 부분에서 "예측" 픽처라 하지 않으므로, "예측 픽처"라는 용어는 인터-픽처 예측 픽처 (예컨대, P- 또는 B-픽처)로 이해될 수 있다.
- [0068] 현재 픽처 (205)가 예측 픽처라면, 움직임 추정기 (210)는 하나 이상의 기준 픽처, 예컨대 픽처 저장소 (220)에 버퍼링된 재구성된 이전의 픽처 (225)에 대해 현재 픽처 (205)의 매크로블록 또는 다른 샘플 집합의 움직임을 추정한다. 움직임 추정기 (210)는 하나 이상의 시간적으로 이전의 기준 픽처 및 하나 이상의 시간적으로 다음 기준 픽처 (예컨대, 양방향 예측 픽처의 경우)에 대해 움직임을 추정할 수 있다. 따라서, 인코더 시스템 (200)은 다중 기준 픽처를 위한 별도의 저장소 (220 및 222)를 이용할 수 있다.
- [0069] 움직임 추정기 (210)는 전(full) 샘플, 1/2 샘플, 1/4 샘플, 또는 다른 증분으로 움직임을 추정할 수 있고, 픽처별로 또는 다른 기준으로 움직임 추정의 해상도를 전환할 수 있다. 움직임 추정기 (210) (및 보상기 (230))는 또한 프레임 당 (per-frame) 또는 다른 기준으로 기준 픽처 샘플 보간의 타입들 사이 (예컨대, 큐빅 컨볼루션 보간 (cubic convolution interpolation)과 이중선형 보간 (bilinear interpolation)사이)에서 전환할 수 있다. 이 움직임 추정의 해상도는 수평적 및 수직적으로 동일하거나 또는 서로 다를 수 있다. 움직임 추정기 (210)는 부차적인 정보로서 미분 움직임 벡터 정보와 같은 움직임 정보 (215)를 출력한다. 인코더 (200)는 예컨대 움직임 벡터에 대한 하나 이상의 예측자(predictor)를 계산하고, 움직임 벡터와 예측자 사이의 차를 계산하고, 이 차를 엔트로피 코딩함으로써 움직임 정보 (215)를 인코딩한다. 움직임 벡터를 재구성하기 위해, 움직임 보상기 (230)는 예측자를 움직임 벡터 차 정보와 결합한다.
- [0070] 움직임 보상기 (230)는 재구성된 움직임 벡터를 재구성된 픽처 (225)에 적용하여 움직임 보상된 예측 (235)을

형성한다. 그러나, 예측은 거의 완전하지 않고, 움직임 보상 예측 (235)과 원래의 현재 픽처 (205) 사이의 차는 예측 잔차 (245)이다. 이 픽처의 추후 재구성 동안, 움직임 보상 예측 (235) 보다 원래의 현재 픽처 (205)에 더 가까운 재구성 픽처를 얻기 위해 예측 잔차 (245)의 근사치가 움직임 보상 예측 (205)에 가산된다. 그러나, 손실 압축에서, 일부 정보는 원래의 현재 픽처 (205)로부터 여전히 손실된다. 다른 경우에, 움직임 추정기 및 움직임 보상은 다른 형태의 움직임 추정/보상을 적용한다.

[0071] 주파수 변환기 (260)는 공간 도메인 비디오 정보를 주파수 도메인 (즉, 스펙트럼) 데이터로 변환한다. 블록 기반 비디오 코딩의 경우, 주파수 변환기 (260)는 일반적으로 DCT (이산 코사인 변환; discrete cosine transform), DCT의 변형, 또는 다른 블록 변환을 샘플 데이터 또는 예측 잔차 데이터의 블록에 적용하여, 주파수 도메인 변환 계수의 블록을 생성한다. 다른 경우에, 주파수 변환기 (260)는 푸리에 변환과 같은 다른 종류의 주파수 변환을 적용하거나 웨이블릿(wavelet) 또는 서브밴드(sub-band) 분석을 이용한다. 주파수 변환기 (260)는 8×8 , 8×4 , 4×8 , 4×4 , 또는 다른 크기의 주파수 변환을 적용할 수도 있다.

[0072] 양자화기 (270)는 주파수 도메인 변환 계수의 블록을 양자화한다. 양자화기는 픽처 단위, 매크로블록 단위, 또는 다른 기준으로 가변하는 양자화 스텝 크기에 따라 스칼라 양자화를 변환 계수에 적용하는데, 여기서 양자화 스텝 크기는 디코더 역(inverse) 양자화기 프로세스에서 이산(discrete) 표현가능한 재구성 포인트들 사이의 균일하게 이격된 간격을 지배하는 제어 파라미터이며, 디코더의 역 양자화기 프로세스는 인코더의 역양자화기 프로세스 (276)에서 중복될 수 있다. 다른 경우에, 양자화기는 다른 종류의 양자화를 주파수 도메인 변환 계수, 예컨대 불균일 재구성 포인트를 갖는 스칼라 양자화기, 벡터 양자화기, 또는 비적응(non-adaptive) 양자화에 적용하거나, 또는 주파수 변환을 이용하지 않는 인코더 시스템에서 공간 도메인 데이터를 직접 양자화한다. 적응 양자화 외에도, 인코더 (200)는 프레임 누락(dropping), 적응 필터링, 또는 다른 레이트(rate) 제어 기술을 이용할 수 있다.

[0073] 재구성된 현재 픽처가 추후의 움직임 추정/보상을 위해 필요할 때, 역양자화기 (276)는 양자화된 주파수 도메인 변환 계수에 대해 역 양자화를 수행한다. 역 주파수 변환기 (266)는 주파수 변환기 (260)의 동작의 역을 수행하여, 재구성된 예측 잔차 근사(approximation) (예측 픽처의 경우) 또는 재구성된 I-픽처 근사를 생성한다. 현재 픽처 (205)가 I-픽처인 경우, 재구성된 I-픽처 근사는 재구성된 현재 픽처 근사 (도시되지 않음)로서 취해진다. 현재 픽처 (205)가 예측 픽처라면, 재구성된 예측 잔차 근사는 움직임 보상 예측 (235)에 추가되어 재구성된 현재 픽처 근사를 형성한다. 하나 이상의 픽처 저장소 (220 및 222)는 추후 픽처의 움직임 보상 예측에서 기준 픽처로서 사용하기 위해 재구성된 현재 픽처 근사를 버퍼링한다. 인코더는 디블록킹(de-blocking) 필터 또는 다른 픽처 정련(refining) 프로세스를 재구성된 프레임에 적용하여 불연속성을 적응적으로 매끄럽게 하고 픽처 근사를 하나 이상의 픽처 저장소 (220, 222)로 저장하기 전에 픽처로부터 다른 아티팩트를 제거할 수 있다.

[0074] 엔트로피 코더 (280)는 양자화기 (270)의 출력뿐 아니라 임의의 부차적인 정보 (예컨대, 움직임 정보 (215), 양자화 스텝 크기)를 압축한다. 일반적인 엔트로피 코딩 기술은 산술 코딩, 차분 코딩, 허프만 (Huffman) 코딩, 런 렱스(run length) 코딩, 램펠-지프 (Lempel-Ziv) 코딩, 사전(dictionary) 코딩, 및 이들의 조합을 포함한다. 엔트로피 코더 (280)는 일반적으로 서로 다른 종류의 정보 (예컨대, 저주파수 계수, 고주파수 계수, 영주파수 계수, 서로 다른 종류의 부차적인 정보)에 대한 서로 다른 코딩 기술을 이용하고, 특정 코딩 기술 내에서 다중의 코드 테이블 중에서 선택할 수 있다.

[0075] 엔트로피 코더 (280)는 압축된 비디오 정보 (295)를 멀티플렉서 ("MUX") (290)에 제공한다. MUX (290)는 버퍼를 포함할 수 있고, 버퍼 충만 수준 표시자(indicator)는 레이트 제어를 위한 비트 전송률 적응 모듈로 피드백될 수 있다. MUX (290) 전 또는 후에, 압축된 비디오 정보 (295)가 네트워크를 통한 전송을 위해 채널 코딩될 수 있다. 채널 코딩은 오류 검출 및 정정 데이터를 압축된 비디오 정보 (295)에 적용할 수 있다.

[0076] C. 비디오 디코더

[0077] 도 3은 일예의 비디오 디코더 시스템 (300)의 블록도이다. 디코더 시스템 (300)은 압축된 비디오 픽처 시퀀스에 대한 정보 (395)를 수신하고 재구성된 픽처 (305) (예컨대, 순차 비디오 프레임, 비월 비디오 프레임 또는 비월 비디오 프레임의 필드)를 포함하는 출력을 생성한다. 비디오 디코더의 특정 실시예는 일반적으로 일반화된 디코더 (300)의 변형 또는 보충 버전을 이용한다.

[0078] 디코더 시스템 (300)은 예측된 픽처 및 I-픽처를 압축풀기한다. 설명을 위해, 도 3은 디코더 시스템 (300)을 통한 I-픽처용 경로와 예측 픽처용 경로를 도시한다. 디코더 시스템 (300)의 구성요소 중 다수는 I-픽처 및 예

측 픽처 모두를 압축풀기 위해 사용된다. 이들 구성요소에 의해 수행되는 해당 동작은 압축풀기되고 있는 정보의 종류에 따라 가변적일 수 있다.

[0079] DEMUX (390)는 압축된 비디오 시퀀스에 대한 정보 (395)를 수신하고 수신된 정보를 엔트로피 디코더 (380)에 이 용가능하게 한다. DEMUX (390)는 지터 버퍼 및 다른 버퍼를 포함할 수 있다. DEMUX (390) 이전 또는 이내에서, 압축된 비디오 정보는 오차 검출 및 정정을 위해 채널 디코딩 및 처리될 수 있다.

[0080] 엔트로피 디코더 (380)는 엔트로피 코딩된 양자화 데이터 뿐만 아니라 엔트로피 코딩된 부차적인 정보 (예컨대, 움직임 정보 (315), 양자화 스텝 크기)를 엔트로피 디코딩하는데, 일반적으로 인코더에서 수행된 엔트로피 인코딩의 역(inverse)을 적용한다. 엔트로피 디코딩 기술은 산술 디코딩, 차분 디코딩, 허프만 디코딩, 런 렱스 디코딩, 램펠-지프 디코딩, 사전 디코딩, 및 이들의 조합을 포함한다. 엔트로피 디코더 (380)는 일반적으로 서로 다른 종류의 정보 (예컨대, 저주파수 계수, 고주파수 계수, 영주파수 계수, 서로 다른 종류의 부차적인 정보)에 대한 서로 다른 디코딩 기술을 이용하고, 특정 디코딩 기술 내에서 다중 코드 테이블 중에서 선택할 수 있다.

[0081] 디코더 (300)는 예컨대 움직임 벡터에 대한 하나 이상의 예측자를 계산하고, 움직임 벡터 차를 엔트로피 디코딩 하고 (엔트로피 디코더 (380)에서), 디코딩된 움직임 벡터 차를 재구성된 움직임 벡터에 대한 예측자와 결합함으로써 움직임 정보 (315)를 디코딩한다.

[0082] 움직임 보상기 (330)는 재구성되고 있는 픽처 (305)의 예측 (335)을 형성하기 위해 움직임 정보 (315)를 하나 이상의 기준 픽처 (325)에 적용한다. 예를 들면, 움직임 보상기 (330)는 하나 이상의 매크로블록 움직임 벡터를 사용하여 샘플 블록을 찾거나 기준 픽처 (325) 내의 샘플들 사이의 미소(fractional) 위치를 보간한다. 하나 이상의 픽처 저장소 (예컨대, 픽처 저장소 (320 및 322))는 기준 픽처로서 사용하기 위해 이전의 재구성된 픽처를 저장한다. 일반적으로, B-픽처는 둘 이상의 기준 픽처 (예컨대, 적어도 하나의 시간적으로 이전의 기준 픽처 및 적어도 하나의 시간적으로 미래 기준 픽처)를 갖는다. 따라서, 디코더 시스템 (300)은 다중 기준 픽처를 위한 별도의 픽처 저장소 (320 및 322)를 가질 수 있다. 움직임 보상기 (330)는 전(full) 샘플, 1/2 샘플, 1/4 샘플, 또는 다른 증분으로 움직임을 보상할 수 있고, 픽처별로 또는 다른 기준으로 움직임 보상의 해상도를 전환할 수 있다. 움직임 보상기 (330)는 또한 프레임 당(per-frame) 또는 다른 기준으로 기준 픽처 샘플 보간의 타입들 사이에서 (예컨대, 큐빅 컨볼루션 보간과 이중선형 보간 사이) 전환할 수 있다. 움직임 보상의 해상도는 수평 및 수직으로 동일 또는 상이할 수 있다. 다른 경우에, 움직임 보상기는 다른 타입의 움직임 보상을 적용한다. 움직임 보상기에 의한 예측은 거의 완전하지 않아서, 디코더 (300)는 예측 잔차를 재구성한다.

[0083] 역 양자화기 (370)는 엔트로피 디코딩된 데이터를 역 양자화한다. 일반적으로, 역 양자화기는 균일한 스칼라 역양자화를 픽처별, 매크로블록 기준, 또는 다른 기준으로 가변하는 재구성 스텝 크기를 갖는 엔트로피 디코딩된 데이터에 적용한다. 다른 경우에, 역 양자화기는 다른 타입의 역 양자화를 데이터, 예컨대, 비균일, 벡터 또는 비적응(non-adaptive) 역양자화에 적용하거나, 역 주파수 변환을 이용하지 않는 디코더 시스템에서 공간도메인 데이터를 직접적으로 역 양자화한다.

[0084] 역 주파수 변환기 (360)는 역 양자화 주파수 도메인 변환 계수를 공간 도메인 비디오 정보로 변환한다. 블록 기반 비디오 픽처의 경우, 역 주파수 변환기 (360)는 역 DCT ("IDCT"), IDCT의 변형, 또는 다른 역 블록 변환을 주파수 변환 계수의 블록에 적용하여, I-픽처 또는 예측 픽처에 대해 각각 샘플 데이터 또는 인터-픽처 예측 잔차 데이터를 생성한다. 다른 경우에, 역 주파수 변환기 (360)는 역 푸리에 변환과 같은 또다른 종류의 역 주파수 변환을 적용하거나 웨이블릿 또는 서브밴드 합성을 이용한다. 역 주파수 변환기 (360)는 8×8, 8×4, 4×8, 4×4 또는 다른 크기 역 주파수 변환을 적용할 수 있다.

[0085] 예측 픽처의 경우, 디코더 (300)는 재구성된 예측 잔차 (345)를 움직임 보상된 예측 (335)과 결합하여 재구성된 픽처 (305)를 형성한다. 디코더가 추후 움직임 보상을 위해 재구성된 픽처 (305)를 필요로 할 때, 하나 이상의 픽처 저장소 (예컨대, 픽처 저장소 (320))는 다음 픽처를 예측할 때 사용하기 위해 재구성된 픽처 (305)를 버퍼링한다. 일부 실시예에서, 디코더 (300)는 디블로킹 필터 또는 다른 픽처 정련(refining) 프로세스를 재구성된 픽처에 적용하여 불연속성을 적응적으로 매끄럽게 하고 재구성된 픽처 (305)를 하나 이상의 픽처 저장소 (예컨대, 픽처 저장소 (320))에 저장하기 전이나 또는 디코딩된 비디오 플레이-아웃(play-out) 동안 디코딩된 픽처를 디스플레이하기 전에 픽처로부터 다른 아티팩트(artifacts)를 제거한다.

[0086] III. 다중 해상도 인코딩 및 디코딩의 개요

[0087] 비디오는 서로 다른 해상도로 인코딩 (및 디코딩)될 수 있다. 이 설명의 목적을 위해, 다중 해상도 인코딩 및 디코딩은 프레임 기반 코딩 및 디코딩 (예컨대, 기준 픽처 리샘플링) 또는 계층화된 (공간 크기조정가능 이라고

도 함) 코딩 및 디코딩으로서 설명될 수 있다. 다중 해상도 인코딩 및 디코딩은 또한 비월 비디오 및 필드 기반 인코딩 및 디코딩, 그리고 해상도 특정 기준 또는 다른 기준으로 프레임 기반 및 필드 기반 인코딩 및 디코딩 사이의 전환을 포함할 수 있다. 그러나, 순차 비디오의 프레임 코딩은 개념 설명의 단순화 목적으로 이 개요에서 논의된다.

[0088] A. 프레임 기반 다중 해상도 인코딩 및 디코딩

[0089] 프레임 기반 다중 해상도 코딩에서, 인코더는 서로 다른 해상도로 입력 픽처를 인코딩한다. 인코더는 픽처별로 또는 다른 기준으로 픽처에 대한 공간 해상도를 선택한다. 예를 들면, 기준 픽처 리샘플링에서, 기준 픽처는 인코딩되고 있는 픽처와는 다른 해상도로 인코딩되면 리샘플링될 수 있다. 리샘플링이란 용어는 샘플링된 신호의 픽처 영역 또는 다른 섹션을 표현하기 위해 사용되는 샘플의 수를 증가(업샘플링) 또는 감소(다운샘플링)하는 것을 설명하기 위해 사용된다. 단위 영역당 또는 신호 섹션당 샘플의 수를 샘플링의 해상도라고 한다.

[0090] 공간 해상도는 예컨대 이용가능한 비트 전송률의 감소/증가, 양자화 스텝 크기의 감소/증가, 입력 비디오 콘텐츠에서 움직임량의 감소/증가, 비디오 콘텐츠의 다른 속성(예컨대, 강한 에지(strong edges)의 존재, 텍스트, 또는 더 낮은 해상도로 심하게 왜곡될 수 있는 다른 콘텐츠), 또는 다른 기준에 기초하여 선택될 수 있다. 공간 해상도는 수직, 수평, 또는 수직 및 수평 디멘전에서 가변적일 수 있다. 수평 해상도는 수직 해상도와 동일하거나 다를 수 있다. 디코더는 상보적 기술을 이용하여 인코딩된 프레임을 디코딩한다.

[0091] 인코더가 현재의 픽처 또는 현재의 픽처 내의 영역에 대한 공간 해상도를 선택한 경우, 인코더는 원래의 픽처를 코딩하기 전에 바람직한 해상도로 이를 리샘플링한다. 인코더는 공간 해상도의 선택을 디코더에 신호로 보낼 수 있다.

[0092] 도 6은 픽처의 프레임 기반 다중 해상도 인코딩에 대한 기술(600)을 도시한다. 도 2의 인코더(200)와 같은 인코더는 픽처에 대한 해상도를 설정한다(610). 예를 들면, 인코더는 상기 기준 또는 다른 기준을 고려한다. 인코더는 그 해상도로 픽처를 인코딩한다(620). 인코딩될 모든 픽처의 인코딩이 끝나면(630), 인코더는 종료한다. 그렇지 않으면, 인코더는 다음 픽처에 대한 해상도를 설정(610)하고 계속 인코딩한다. 다른 경우에, 픽처의 서로 다른 부분에 대해 달리 해상도를 설정하거나 일군의 또는 일련의 픽처에 대한 해상도를 선택하는 것처럼, 인코더는 픽처 레벨과는 다른 레벨로 해상도를 설정한다.

[0093] 인코더는 인트라 픽처 뿐만 아니라 예측 픽처도 인코딩할 수 있다. 도 8은 인트라 픽처 및 인터-픽처 예측 픽처의 프레임 기반 다중 해상도 인코딩에 대한 기술(800)을 도시한다. 먼저, 인코더는 단계(810)에서 인코딩될 현재 픽처가 인트라 픽처인지 예측 픽처인지를 체크한다. 현재 픽처가 인트라 픽처이면, 인코더는 단계(820)에서 현재 픽처에 대한 해상도를 설정한다. 픽처가 예측 픽처이면, 인코더는 현재 픽처에 대한 해상도를 설정하기 전에 단계(830)에서 기준 픽처에 대한 해상도를 설정한다. 현재 픽처에 대한 해상도를 설정한 후에, 인코더는 그 해상도로 현재 픽처를 인코딩한다(840). 픽처에 대한 해상도를 설정하는 것은(현재의 소스 픽처든지 저장된 기준 픽처든지) 선택된 해상도에 매칭하기 위해 픽처를 리샘플링하고 선택된 해상도를 디코더에 나타내기 위해 신호를 인코딩하는 단계를 포함할 수 있다. 인코딩될 모든 픽처의 인코딩이 완료되면(850), 인코더는 종료한다. 그렇지 않으면, 인코더는 추가 픽처를 계속 인코딩한다. 다른 경우에, 인코더가 다른 방식으로 예측 픽처를 처리한다.

[0094] 디코더는 인코딩된 픽처를 디코딩하고, 필요하면, 디스플레이 전에 픽처를 리샘플링한다. 인코딩된 픽처의 해상도처럼, 디코딩된 픽처의 해상도는 다수의 다른 방식으로 조정될 수 있다. 예를 들면, 디코딩된 픽처의 해상도는 출력 디스플레이 장치나 출력 디스플레이 장치의 영역(예컨대, "PIP (picture-in-picture)" 또는 PC 데스크탑 윈도우 디스플레이)의 해상도에 맞춰지도록 조정될 수 있다.

[0095] 도 7은 픽처의 프레임 기반 다중 해상도 디코딩에 대한 기술(700)을 도시한다. 도 3의 디코더(300)와 같은 디코더는 픽처에 대한 해상도를 설정한다(710). 예를 들면, 디코더는 인코더로부터 해상도 정보를 얻는다. 디코더는 그 해상도로 픽처를 디코딩한다(720). 디코딩될 모든 픽처의 디코딩이 완료되면(730), 디코더는 종료한다. 그렇지 않으면, 디코더는 다음 픽처에 대한 해상도를 설정하고(710) 계속 디코딩한다. 다른 경우에, 디코더는 픽처 레벨과는 다른 레벨로 해상도를 설정한다.

[0096] 디코더는 인트라 픽처 뿐만 아니라 예측 픽처를 디코딩할 수 있다. 도 9는 인트라 픽처 및 예측 픽처의 프레임 기반 다중 해상도 디코딩에 대한 기술(900)을 도시한다.

[0097] 먼저, 디코더는 디코딩될 현재 프레임이 인트라 픽처인지 예측 픽처인지 체크한다(910). 현재 픽처가 인트라 픽처이면, 디코더는 현재 픽처에 대한 해상도를 설정한다(920). 픽처가 예측 픽처이면, 디코더는 현재 픽처에

대한 해상도를 설정 (920)하기 전에, 기준 픽처에 대한 해상도를 설정한다 (930). 기준 픽처의 해상도를 설정하는 것은 선택된 해상도에 매칭하기 위해 저장된 기준 픽처를 리샘플링하는 것을 포함한다. 현재 픽처에 대한 해상도를 설정 (920)한 후, 디코더는 그 해상도로 현재 픽처를 디코딩한다 (940). 디코딩될 모든 픽처의 디코딩이 완료되면 (950), 디코더는 종료한다. 그렇지 않으면, 디코더는 계속 디코딩한다.

[0098] 디코더는 일반적으로 인코더에서 사용되는 동일한 해상도로 픽처를 디코딩한다. 다른 경우에, 디코더는 예컨대 디코더에 대해 이용가능한 해상도가 인코더에서 사용된 것과 정확히 동일하지 않은 경우처럼 서로 다른 해상도로 픽처를 디코딩한다.

[0099] B. 계층화된 다중 해상도 인코딩 및 디코딩

[0100] 계층화된 다중 해상도 코딩에서, 인코더는 계층에서 비디오를 인코딩하는데, 각 계층은 서로 다른 해상도로 비디오를 디코딩하기 위한 정보를 갖는다. 이런 식으로, 인코더는 둘 이상의 해상도로 비디오에서 적어도 몇개의 개별 픽처를 인코딩한다. 디코더는 계층들의 서로 다른 결합을 처리함으로써 하나 이상의 해상도로 비디오를 디코딩할 수 있다. 예를 들어, 제1 계층 (기본 계층이라고도 함)은 더 낮은 해상도로 비디오를 디코딩하기 위한 정보를 포함하고, 하나 이상의 다른 계층 (인핸스먼트(enhancement) 계층이라고도 함)은 더 높은 해상도로 비디오를 디코딩하기 위한 정보를 포함한다.

[0101] 기본 계층은 그 자체가 독립적으로 디코딩가능한 비트스트림으로 지정될 수 있다. 따라서, 이러한 설계에서, 기본 계층만을 디코딩하는 디코더는 기본 계층의 더 낮은 해상도로 유효한 디코딩된 비트스트림을 생성할 것이다. 인핸스먼트 계층을 이용한 더 높은 해상도 픽처의 적절한 디코딩은 일부 또는 모든 인코딩된 기본 계층 데이터 및 가능하면 하나 이상의 인핸스먼트 계층을 디코딩하는 것을 필요로 할 수 있다. 기본 계층 및 하나 이상의 다른 더 높은 해상도 계층을 디코딩하는 디코더는 기본 계층만을 디코딩하는 디코더 보다 더 높은 해상도 콘텐츠를 생성할 수 있을 것이다. 두개, 세개 또는 그 이상의 계층은 두개, 세개 또는 그 이상의 서로 다른 해상도를 허용하도록 사용될 수 있다. 다른 경우에, 더 높은 해상도 계층은 그자체가 독립적으로 디코딩가능한 비트스트림일 수 있다. (이러한 설계는 종종 사이멀캐스트(simulcast) 다중 해상도 인코딩 접근법이라고 한다.)

[0102] 도 10은 서로 다른 해상도로 디코딩하는 것을 허용하기 위해 비트스트림 계층을 인코딩하는 기술 (1000)을 도시한다. 도 2의 인코더 (200)와 같은 인코더는 입력으로서 전 해상도(full-resolution) 비디오 정보를 취한다 (1010). 인코더는 전 해상도 비디오 정보를 다운샘플링하고 (1020), 다운샘플링된 정보를 이용하여 기본 계층을 인코딩한다 (1030). 인코더는 기본 계층 및 고해상도(higher-resolution) 비디오 정보를 이용하여 하나 이상의 고해상도 계층을 인코딩한다 (1040). 고해상도 계층은 전 해상도로 디코딩을 허용하는 계층이거나 중간 해상도로 디코딩을 허용하는 계층일 수 있다. 그후 인코더는 둘 이상의 인코딩된 계층을 포함하는 계층화된 비트스트림을 출력한다. 다른 경우에, 고해상도 계층의 인코딩은 (1040) 기본 계층 정보를 사용할 수 없어서 사이멀캐스트 다중 해상도 인코딩 접근법을 위해 고해상도 계층 데이터의 독립적인 디코딩을 사용할 수 있다.

[0103] 인코더는 도 10에 도시된 기본 개요를 따르는 여러 방법으로 다중 해상도 계층 인코딩을 달성할 수 있다. 더 많은 정보를 위해, 미국특허 제6,510,177호, 또는 MPEG-2 표준 또는 다른 비디오 표준을 참조하기 바란다.

[0104] 도 11은 서로 다른 해상도로 비디오 디코딩을 허용하기 위해 비트스트림 계층을 디코딩하는 기술 (1100)을 도시한다. 도 3의 디코더 (300)와 같은 디코더는 입력으로서 계층화된 비트스트림을 취한다 (1110). 계층들은 저 해상도 계층 (기본 계층) 및 고해상도 정보를 포함하는 하나 이상의 계층을 포함한다. 고해상도 계층들은 독립적으로 인코딩된 픽처를 포함할 필요가 없고, 일반적으로 고해상도 계층들은 고해상도 및 저해상도 버전의 픽처들 간의 차를 설명하는 잔차(residual) 정보를 포함한다. 디코더는 기본 계층을 디코딩하고 (1120), 고해상도 디코딩이 필요하면, 디코더는 디코딩된 기본 계층 픽처를 원하는 해상도로 업샘플링한다 (1130). 디코더는 하나 이상의 고해상도 계층을 디코딩하고 (1140), 디코딩된 고해상도 정보를 업샘플링되고 디코딩된 기본 계층 픽처와 결합하여 고해상도 픽처를 형성한다 (1150). 원하는 해상도 수준에 따라, 고해상도 픽처는 전 해상도 픽처 또는 중간 해상도 픽처일 수 있다. 더 많은 정보를 위해, 미국특허 제6,510,177호, 또는 MPEG-2 표준 또는 다른 비디오 표준을 참조하기 바란다.

[0105] 디코더는 일반적으로 인코더에서 사용된 해상도 중 하나로 픽처를 디코딩한다. 다른 경우에, 디코더에 대해 이용가능한 해상도는 인코더에서 사용된 것과 정확히 동일하지는 않다.

[0106] IV. 크기조정가능한 비디오 코딩 및 디코딩을 위한 리샘플링 필터

[0107] 이 절은 크기조정가능한 비디오 코딩 및 디코딩에 대한 기술 및 도구에 대해 설명한다. 일부 설명된 기술 및 도구는 계층화된 (또는 공간상 크기조정가능한) 상황에서 설명되었지만, 일부 설명된 기술 및 도구는 프레임 기

반 (또는 기준 픽처 샘플링) 상황 또는 리샘플링 필터를 포함하는 일부 다른 상황에서 사용될 수도 있다. 또한, 일부 설명된 기술 및 도구는 리샘플링 픽처의 상황에서 설명되었지만, 일부 설명된 기술 및 도구는 고해상도 신호의 예측으로 생기는 리샘플링 잔차 또는 차 신호에 대해 사용될 수 있다.

- [0108] 크기조정이 가능한 비디오 코딩 (SVC:scalable video coding)은 디지털 비디오 코딩의 한 종류로, 이것은 더 큰 비트스트림의 부분집합이 일부 응용에 대해 수용가능한 품질을 갖는 디코딩된 픽처를 생성하기 위해 디코딩될 수 있게 한다 (비록 이러한 픽처 품질은 전체 고 비트 전송률 (higher-bit-rate) 비트스트림을 디코딩함으로써 생성되는 품질보다 더 낮을 수 있다). 한가지 잘 알려진 SVC 종류는 공간 확장성 (spatial scalability) 또는 해상도 확장성 (resolution scalability)이라고 한다. 공간 SVC 설계에서, 인코딩 프로세스 (또는 인코딩 프로세스의 범위의 정확한 정의에 따라, 인코딩 프로세스 전에 수행될 사전처리 기능)는 일반적으로 비디오를 더 낮은 해상도로 다운샘플링, 저해상도 (lower-resolution) 디코딩 프로세스를 가능하게 하기 위해 상기 저해상도 비디오를 인코딩, 및 고해상도 (higher-resolution) 비디오의 픽처에서의 샘플의 값의 예측으로서 사용하기 위해 이 저해상도 디코딩된 픽처의 업샘플링을 포함한다. 고해상도 비디오에 대한 디코딩 프로세스는 저해상도 비디오 (또는 그 중 일부)를 디코딩하고, 이 업샘플링된 비디오를 고해상도 비디오의 픽처에서의 샘플의 값의 예측으로서 사용하는 것을 포함한다. 이러한 설계는 리샘플링 필터의 사용을 필요로 한다. 특히, 코덱 설계는 디코더 및 인코더에서 업샘플링 필터의 사용을 포함하고 인코더 또는 인코딩 프리-프로세서 (pre-processor)에서 다운샘플링 필터의 사용을 포함한다. 이러한 설계에서 사용되는 업샘플링 필터에 특히 초점을 맞춰 논의하기로 한다. 일반적으로, 업샘플링 프로세스는 드리프트(drift)라고 하는 현상을 피하기 위해 인코더 및 디코더에서 동일하도록 설계되는데, 드리프트는 인코딩 및 디코딩 동안에 동일 신호에 대한 다른 예측의 사용에 의해 야기된 오차의 누적을 말한다.
- [0109] 일부 공간 SVC 설계의 한가지 단점은 디코딩 프로세스에서 저품질 필터 (예컨대, 투-탭 이중선형 필터 (two-tap bilinear filters))의 사용이다. 고품질 필터를 사용하면 비디오 품질에 유리할 것이다.
- [0110] 공간 SVC는 필터의 리샘플링 비율에서의 고도의 유연성을 가능하게 하는 리샘플링 필터를 포함할 수 있다. 그러나, 이것은 개발될 이러한 필터의 각각 서로 다른 "단계(phase)"에 대한 다수의 특정 필터 설계 및 인코더 및 디코더의 구현에서 저장될 이 필터들의 "탭(tap)" 값을 필요로 할 수 있다.
- [0111] 또한, 인코더가 공간 SVC에 대해 사용된 리샘플링 필터의 흐릿함(blurriness) 정도를 제어할 수 있게 하는 것이 비디오 품질에 유리할 수 있다. 따라서, 리샘플링의 각 "단계(phase)"가 업샘플링 또는 다운샘플링을 위해 설계되기 위하여, 이 프로세스에서 도입될 희망하는 흐릿한 정도에 따라, 선택할 몇개의 서로 다른 필터를 갖는 것이 유리할 수 있다. 업샘플링 동안 수행될 흐릿한 정도의 선택은 디코딩 프로세스에서 사용하기 위해 전달되는 정보로서 인코더로부터 디코더로 전송될 수 있다. 이 별도의 유연성은, 인코더 또는 디코더에서 저장될 필요가 있는 필요한 탭 값의 수를 크게 증가시키므로 설계를 더 복잡하게 한다.
- [0112] 다양한 단계(phases) 및 다양한 흐릿한 정도를 갖는 다양한 리샘플링 필터를 특정하기 위해 통일된 설계가 사용될 수 있다. 한가지 가능한 해법은 미첼-네트라발리 (Mitchell-Netravali) 필터 설계 방법을 사용하는 것이다. 이들 문제에 대한 미첼-네트라발리 필터 설계 방법의 직접적인 적용은 인코더 또는 디코더에서 계산될 양에 대한 가능한 값의 과도한 동적 범위의 형태로 여분의 계산 리소스를 필요로 할 수 있다. 예를 들면, 이러한 한가지 설계는 범용 CPU 및 DSP에서 통상 사용되는 16비트 또는 32비트 처리 요소보다는 45비트 산술 처리의 사용을 필요로 할 수 있다. 이 문제를 해결하기 위해, 몇가지 설계 방법을 제공한다.
- [0113] 일반적인 SVC 설계는 공간 확장성을 위한 표준(normative) 업샘플링 필터를 필요로 한다. 임의의 리샘플링 비율 (확장된 공간 확장성(extended spatial scalability)으로 알려진 특징)을 지원하기 위해, 리샘플링 비율에 대해 상당한 유연성을 포함하는 업샘플링 필터 설계가 설명된다. 또다른 핵심 양태는 루마 및 크로마의 상대적 정렬이다. 다양한 정렬 구조 (예컨대, 4:2:0 크로마에 대한 H.261/MPEG-1 vs. MPEG-2 정렬 및 H.264/MPEG-4 AVC)가 단일 계층 접근에서 발견되므로, 인코더가 디코더에게 필터링을 적절히 적용하는 법을 나타내는 간단한 방법을 이용하여, 설명되는 기술 및 도구는 유연한 다양한 정렬을 지원한다.
- [0114] 설명되는 기술 및 도구는 고품질 업샘플링 및 양호한 안티-앨리어싱(anti-aliasing)이 가능한 업샘플링 필터를 포함한다. 특히, 설명되는 기술 및 도구는 공간 확장성을 위해 이전의 이중선형 필터 설계에 의해 제공된 것을 넘어서는 품질을 갖는다. 설명되는 기술 및 도구는 양호한 신호처리 빈도 행위를 제공할 뿐만 아니라 시간적으로 만족을 주는 고품질 업샘플링 필터를 갖는다. 설명되는 기술 및 도구는 특정하기 쉽고 탭 값을 유지하기 위해 많은 메모리 저장 테이블을 필요로 하지 않는 필터 설계를 포함하고, 필터링 동작 자체는 계산적으로 연산하기 간단하다. 예를 들면, 설명되는 기술 및 도구는 지나치게 길지 않고 과도한 수학적 정밀도나 지나치게 복잡

한 수학적 기능을 필요로 하지 않는 필터를 갖는다.

[0115] 이 절은 하나 이상의 다음과 같은 특징을 갖는 설계를 설명한다:

- [0116] - 루마/크로마 페이즈(phase) 정렬의 유연성
- [0117] - 리샘플링 비율의 유연성
- [0118] - 주파수 특성의 유연성
- [0119] - 시각적 고품질
- [0120] - 너무 적거나 너무 많지 않는 필터 탭 (예컨대, 4와 6 사이)
- [0121] - 특정하기 간단함
- [0122] - 연산하기 간단함 (예컨대, 실제 워드 길이의 산술을 이용)

[0123] A. 미첼-네트라발리 (Mitchell-Netravali) 업샘플링 필터

[0124] 설명되는 기술 및 도구는 분리가 가능한 필터링 접근법을 이용한다. 따라서, 다음 논의는 주로 1차원 신호의 처리에 관한 것인데, 그 이유는 2차원 케이스는 1차원 케이스의 간단한 분리가 가능한 적용이기 때문이다. 이것은 먼저 개념적으로 연속인(conceptually-continuous) 임펄스 응답 $h(x)$ 에 기초한 2-파라미터 집합의 필터 (a two-parameter set of filters)를 제안하는데, 이는 다음과 같이 주어진다:

수학식 1

$$h(x) = \frac{1}{6} * \begin{cases} (12-9b-6c)|x|^3 - (18-12b-6c)|x|^2 + (6-2b)|x| & |x| < 1 \\ -(b+6c)|x|^3 + (6b+30c)|x|^2 - (12b+48c)|x| + (8b+24c) & 1 \leq |x| < 2 \\ 0 & \text{그외의 경우} \end{cases}$$

[0126] 여기서, b 및 c 는 2개의 파라미터이다. 상대적 페이즈 오프셋 위치 $0 \leq x < 1$ 의 경우, 이 커널(kernel)은 다음 행렬식에 의해 주어진 탭 값을 갖는 4-탭 FIR (무한 임펄스 응답; finite impulse response) 필터를 생성한다.

수학식 2

$$\frac{1}{6} * \begin{bmatrix} 1 & x & x^2 & x^3 \end{bmatrix} * \begin{bmatrix} b & 6-2b & b & 0 \\ -(3b+6c) & 0 & 3b+6c & 0 \\ 3b+12c & -18+12b+6c & 18-15b-12c & -6c \\ -(b+6c) & 12-9b-6c & -(12-9b-6c) & b+6c \end{bmatrix}$$

[0128] 실제로, x 에 대한 FIR 필터 커널은 단순히 역순으로 $1-x$ 에 대한 FIR 필터 커널이므로 0 부터 1/2 까지의 x 범위만을 고려해도 충분하다.

[0129] 이 설계는 흥미롭고 유용한 많은 속성을 갖는다. 그중 몇가지는 다음과 같다:

[0130] - 필터 탭 값을 계산하는 데 삼각함수(trigonometric functions), 초월함수(transcendental functions) 또는 무리수(irrational-number) 처리가 불필요하다. 사실, 이러한 필터에 대한 탭 값은 몇개의 간단한 연산만으로도 직접 계산될 수 있다. 이들은 사용될 페이즈 및 파라미터의 다양한 가능한 값들을 위해 저장될 필요가 없어서, 필요할 때 간단히 계산될 수 있다. (그래서, 이러한 필터의 사용을 표준화하기 위해, 몇개의 공식만이 필요하다. 숫자들의 큰 테이블이나 코사인 또는 베셀함수(Bessel functions)와 같은 함수들에 근접(approximate)시키려는 표준화된 시도가 불필요하다)

[0131] - 최종 필터는 4 탭을 갖는다. 이것은 매우 실용적인 숫자이다.

[0132] - 필터는 메인 로브(main lobe)의 각 면에서 오직 하나의 사이드로브(sidelobe)를 갖는다. 따라서 과도한 링잉 현상(ringing artifacts)을 발생시키지 않을 것이다.

[0133] - 필터는 평탄한 임펄스 응답을 갖는다. 그 값 및 1차 도함수는 모두 연속적이다.

[0134] - 필터는 단위이득 DC 응답을 갖는데, 이는 업샘플링되고 있는 정보에서 전반적인 밝기 증폭 또는 감쇠가 없다는 것을 의미한다.

[0135] - 이 필터 군(filter)의 멤버들은 "Lanczos-2" 설계 및 "Catmull-Rom" 설계와 같은 공지의 우수한 필터들의 비

교적 상당히 유사한 것들을 포함한다.

- [0136] 또한, 설명되는 기술 및 도구는 시각적으로 만족스러운 필터들의 선택을 위해 2개의 파라미터들 사이의 특정 관계를 포함한다. 이 관계는 다음과 같이 표현될 수 있다.

수학식 3

$$c = \frac{1}{2}(1-b)$$

[0137]

- [0138] 이것은 단일 대역폭 제어 파라미터 b 에 대한 자유도 (degrees of freedom)를 감소시킨다. 이 파라미터는 필터에 의해 도입된 별도의 흐릿한 정도를 제어한다. 값 $b=0$ 과 관련된 이 군의 멤버는 널리 알려진 우수한 Catmull-Rom 업샘플링 필터 (Keys "큐빅 컨볼루션" 보간 필터로도 알려짐)이다.

- [0139] Catmull-Rom 업샘플링 필터는 Mitchell-Netravali 필터 군의 모든 멤버들에 대해 발견되는 기본 장점들 이외에 많은 자체의 우수한 특성을 갖는다.

- [0140] - 이것은 "보간" 필터이다 - 즉, 페이즈 값 $x=0$ 및 $x=1$ 의 경우, 이 필터는 1에 해당하는 단일 논-제로 (non-zero) 탭을 갖는다. 즉, 업샘플링된 신호는 각 업샘플링된 곡선부(upsampled curve segment)의 에지(edges)에서 입력 샘플의 값을 통해 정확히 통과할 것이다.

- [0141] - 입력 샘플의 집합이 포물선 (또는 직선, 또는 일정값)을 형성하면, 출력점(output points)은 정확히 포물선 (또는 직선 또는 일정값) 상에 놓일 것이다.

- [0142] 실제로, 어떤 방법에서는, Catmull-Rom 업샘플러는 이러한 이유로 이 길이의 최선의 업샘플링 필터로 간주될 수 있다 - 비록 약간의 별도의 흐릿함(blurring) 도입 (b 증가)이 때로 시각적으로 더 만족스러울 수 있다. 또한 약간의 별도의 흐릿함을 도입하는 것은 낮은 비트 전송률 압축 아티팩트를 흐릿하게 하는 것(blur out)을 도울 수 있고 따라서 진정한 업샘플링된 픽처의 비에너 (Wiener) 필터 (노이즈 필터링을 위해 사용되는 공지의 필터) 추정기(estimator)에 가깝게 동작한다.

- [0143] 수학식 3의 수학식 2로의 단순한 대체로 다음과 같은 탭 값이 얻어진다:

수학식 4

$$\frac{1}{6} * \begin{bmatrix} 1 & x & x^2 & x^3 \end{bmatrix} * \begin{bmatrix} b & 6-2b & b & 0 \\ -3 & 0 & 3 & 0 \\ 6-3b & -15+9b & 12-9b & -3+3b \\ -(3-2b) & 9-6b & -(9-6b) & 3-2b \end{bmatrix}$$

[0144]

- [0145] 9명의 전문 시청자 및 500개 이상의 샘플을 이용한 주관적 테스트에 기초하여 다음과 같은 사항이 보고되었다:

- [0146] - 사용가능한 범위는 $0 \leq b \leq 5/3$ 로 보고된다;
- [0147] - $0 \leq b \leq 1/2$ 는 시각적으로 "만족스러운" 것으로 분류되고, $b=1/3$ 은 시각적으로 기분좋은 것으로 보고된다;
- [0148] - $b > 1/2$ 는 "흐릿한" 것으로서 분류되고, $b=3/2$ 은 과도하게 흐릿한 것으로서 보고된다.

- [0149] B. 대역폭 제어 파라미터의 정수화

- [0150] 수학식 4에서 6으로 나누는 것은 바람직하지 않을 수 있다. 대신에, 무한한 정밀도가 디코더 설계의 일부로서 비현실적임에 따라, 대역폭 제어 파라미터 및 필터 탭 값을 정수화하는 것이 바람직할 수 있다. 다음과 같이 정의된 새로운 정수값 변수 a 를 이용한 대체를 고려한다:

수학식 5

$$a = (b/6) * 2^s$$

[0151]

[0152] 여기서 S는 정수 시프트 인수이고 a는 정수화된 대역폭 제어 파라미터로서 작용하는 무부호(unsigned) 정수이다. 파라미터 a는 비트스트림에서 비디오 시퀀스 레벨로 인코더에 의해 구문(syntax) 요소로서 인코딩될 수 있다. 예컨대, 파라미터 a는 가변길이 또는 고정길이 코드로 명시적으로 인코딩되거나, 다른 정보와 함께 코딩되거나, 명시적으로 신호보내질 수 있다. 다른 경우에, 파라미터 a는 비트스트림에서 다른 수준으로 신호된다.

[0153] 정수화는 다음과 같은 정수화된 탭값으로 된다.

수학식 6

$$[1 \ x \ x^2 \ x^3]^* \begin{bmatrix} a & 2^S - 2a & a & 0 \\ -2^{S-1} & 0 & 2^{S-1} & 0 \\ 2^S - 3a & -5 \cdot 2^{S-1} + 9a & 2^{S+1} - 9a & -2^{S-1} + 3a \\ -(2^{S-1} - 2a) & 3 \cdot 2^{S-1} - 6a & -(3 \cdot 2^{S-1} - 6a) & 2^{S-1} - 2a \end{bmatrix}$$

[0155] 이 결과는 이진 산술 처리에서 S 위치에 의해 크기감소될 필요가 있을 것이다.

[0156] 만약 a가 0 부터 M 까지의 범위를 가지면, b는 0 부터 $6 \cdot M / 2^S$ 까지의 범위를 갖는다. M에 대한 몇가지 가능한 유용한 선택은 다음을 포함한다:

[0157] - $M = 2^{(S-2)} - 1$, 결과적으로 b의 범위는 0 부터 $3/2 - 6/2^S$ 까지이다.

[0158] - $M = \text{Ceil}(2^S/6)$, 이는 $2^S/6$ 보다 크거나 같은 가장 작은 정수로 되어, b의 범위가 0 부터 1 보다 약간 더 큰 값 까지 된다.

[0159] - $M = 2^{(S-3)} - 1$, 결과적으로 b의 대략적인 범위는 0 부터 $3/4 - 6/2^S$ 까지이다.

[0160] 상기 M에 대한 선택은 가장 유용한 케이스를 커버하기에 충분히 광범위하고, 여기서 첫번째 선택 ($M = 2^{(S-2)} - 1$)은 3가지 선택 중 가장 광범위하다. S에 대한 유용한 범위는 6과 8 사이이다.

[0161] 예컨대, S=7 및 $M = 2^{(S-2)} - 1$, 즉, M=31 을 고려한다. 다른 경우에, M 및 S 의 다른 값들도 사용될 수 있다.

[0162] C. 분수-샘플 포지셔닝(positioning)의 정수화

[0163] 다음에 x 값의 세분성 (granularity)을 고려한다. 실제로, x 도 근사화해야 한다. 예를 들어, 정수 i 를 다음과 같이 정의할 수 있다:

수학식 7

$$x = i \div 2^F$$

[0165] 여기서, F는 지원된 분수-샘플 (fractional-sample) 위치 정밀도를 나타낸다. 충분히 정확한 리샘플링 연산의 일례로서, $F \geq 4$ (1/16 이상의 샘플 포지셔닝 정밀도)를 고려한다. 이것은 다음과 같은 정수화된 필터 탭 값으로 된다:

수학식 8

$$[1 \ i \ i^2 \ i^3]^* \begin{bmatrix} a \cdot 2^{3F} & (2^S - 2a) \cdot 2^{3F} & a \cdot 2^{3F} & 0 \\ -2^{2F+S-1} & 0 & 2^{2F+S-1} & 0 \\ (2^S - 3a) \cdot 2^F & (-5 \cdot 2^{S-1} + 9a) \cdot 2^F & (2^{S+1} - 9a) \cdot 2^F & (-2^{S-1} + 3a) \cdot 2^F \\ -(2^{S-1} - 2a) & 3 \cdot 2^{S-1} - 6a & -(3 \cdot 2^{S-1} - 6a) & 2^{S-1} - 2a \end{bmatrix}$$

[0167] 예를 들어, F=4 를 고려한다. 이 결과는 3F+S 위치에 의해 크기감소될 필요가 있을 것이다.

[0168] 상기 행렬의 모든 항은 인수 2를 공통적으로 포함한다 (S는 1 보다 크다고 가정). 따라서, 다음과 같이 탭 값을 공식화할 수 있다:

수학식 9

$$[1 \ i \ i^2 \ i^3]^* \begin{bmatrix} a*2^{3F-1} & (2^{S-1}-a)*2^{3F} & a*2^{3F-1} & 0 \\ -2^{2F+S-2} & 0 & 2^{2F+S-2} & 0 \\ (2^S-3a)*2^{F-1} & (-5*2^{S-1}+9a)*2^{F-1} & (2^{S+1}-9a)*2^{F-1} & (-2^{S-1}+3a)*2^{F-1} \\ -(2^{S-2}-a) & 3*(2^{S-2}-a) & -3*(2^{S-2}-a) & 2^{S-2}-a \end{bmatrix}$$

여기서 각 탭 값을 2로 나누었다. 이 결과는 단지 $3F+S-1$ 위치에 의해 크기감소될 필요가 있을 것이다.

크기감소의 경우, 함수 $\text{RoundingRightShift}(p, R)$ 을 입력값 p 에 대해 계산된 R 비트 (반올림(rounding)과 함께)의 우측 시프트의 출력으로서 정의하고 다음과 같이 계산된다:

수학식 10

$$\text{RoundingRightShift}(p, R) = \begin{cases} (p + 2^{R-1}) \gg R & \text{for } R = 2, 3, 4, \dots \\ (p) \gg R & \text{for } R = 0 \text{ or } 1 \end{cases}$$

여기서, " $\gg$ " 표기는 2의 보수 이진법을 이용하여 이진법 우측 시프트 연산자를 나타낸다. 다른 경우에, 우측 시프트를 반올림하는 것은 달리 수행된다.

우측 시프트를 반올림하기 위한 몇가지 예 적용이 아래에 제공된다.

D. 동적 범위 고려

픽처를 N 비트의 샘플 비트 깊이로 필터링하고 반올림하기 전에 2차원적으로 그렇게 하면, 이 결과를 $2*(3F+S-1)$ 위치로 다운-시프팅하고 그 출력을 N 비트 범위로 클리핑(clipping)하기 전에 누산기에서 $2*(3F+S-1)+N+1$ 비트의 동적 범위를 필요로 할 것이다. 예를 들어, $F=4$, $S=7$ 및 $N=8$ 인 경우, 필터링된 결과를 계산하기 위해 45 비트 누산기를 사용할 필요가 있을 수 있다.

다음 하위 절(subsections)에서 이 문제를 경감하는 몇가지 접근법을 설명하기로 한다. 이 접근법은 개별적으로 또는 서로 결합하여 이용될 수 있다. 상기 동적 범위 경감 접근법의 변형이 상기 설명에 기초하여 가능함은 당연하다.

1. 제 1 예 동적 범위 경감 접근법

수평 필터링이 먼저 수행되고 다음에 수직 필터링이 수행되는 예를 고려한다. 2차원 처리 파이프라인의 임의의 점에 대해 W 비트의 최대 워드 길이를 고려한다. 제 1 동적 범위 경감 접근법에서, 필터링을 달성하기 위해 이 프로세스의 제 1 (수평) 단계의 출력에서 R_H 비트의 반올림 우측 시프트를, 이 프로세스의 제 2 (수직) 단계의 출력에서 R_V 비트의 반올림 우측 시프트를 이용한다.

다음을 계산한다:

수학식 11

$$2 * (3F + S - 1) + N + 1 - R_H = W$$

따라서,

수학식 12

$$R_H = 2 * (3F + S - 1) + N + 1 - W$$

제 2 (수직) 단계의 우측 시프트는 다음으로부터 계산될 수 있다:

수학식 13

$$R_H + R_V = 2 * (3F + S - 1)$$

[0186] 따라서,

수학식 14

$$R_V = 2 * (3F + S - 1) - R_H$$

[0187]

[0188] 예컨대, F=4, S=7, N=8 및 W=32인 경우, R_H=13 및 R_V=23을 얻는다. 따라서, 동적 범위의 45 비트 대신에, 반올림 우측 시프트를 이용하여 동적 범위는 32비트로 감소된다. 다른 수의 비트의 우측 시프트가 W의 다른 값에 대해 이용될 수 있다.

[0189] 2. 제 2 예 동적 범위 경감 접근법

[0190] 제 2 동적범위 경감 접근법은 페이지 포지셔닝의 정밀도를 감소시키거나 (즉, F 감소), 필터 대역폭 조정 파라미터의 세분성을 감소시키거나 (즉, S 감소), 제 1 단계의 출력의 정밀도를 감소시키거나 (즉, R_H 증가) 보다는 탭 값의 정밀도를 감소시키는 것을 포함한다.

[0191] 수학식 9에 의해 얻어진 4개의 정수 탭값을 [t₋₁, t₀, t₁, t₂]라고 표기한다. 4개의 필터 탭값의 합은 2^{3F+S-1} 일 것이다. 즉,

수학식 15

$$t_{-1} + t_0 + t_1 + t_2 = 2^{3F+S-1}$$

[0192]

[0193] 이것은 상기 제 2 예의 동적범위 경감 접근법의 중요한 특성이다. 왜냐하면 4개의 모든 입력 샘플이 동일 값을 가질 때마다, 출력은 이 동일한 값을 가질 것이기 때문이다.

[0194] 수학식 10에서 발견된 반올림 우측 시프팅의 이 예의 정의를 이용하고, 탭값에 대한 우측 시프트 양 R_i 가 주어지면, 다음과 같이 정의된다:

$$u_{-1} = \text{RoundingRightShift}(t_{-1}, R_i);$$

$$u_1 = \text{RoundingRightShift}(t_1, R_i);$$

$$u_2 = \text{RoundingRightShift}(t_2, R_i);$$

$$u_0 = 2^{3F+S-1} - u_{-1} - u_1 - u_2.$$

[0195]

[0196] 다음에 탭값 [t₋₁, t₀, t₁, t₂] 대신에 탭값 [u₋₁, u₀, u₁, u₂] 으로 필터링을 수행한다. R_i의 값에서 1의 각 증가는 산술 누산기에서 필요한 동적 범위의 하나 적은 비트와, 추후 처리 단계에서 수행될 우측 시프팅의 하나 적은 비트를 나타낸다.

[0197] 3. 제 3 예 동적 범위 경감 접근법

[0198] 이전의 설계는, 프로세스의 제 1 단계후 우측 시프팅의 양을 페이지 포지셔닝 변수 i의 값의 함수로 한다는 점에서, 개념상 유사하지만 제 1 예 동적 범위 경감 접근법과는 다른 트릭(trick)을 사용한다.

[0199] 수학식 9에 도시된 필터 탭값은 i 값이 2^K의 정수배인 때 K 영값(zero-valued) LSB를 포함할 것임을 알 수 있다. 따라서, 필터링 프로세스의 제 2 단계가 2^K의 정수배인 페이지 포지셔닝 변수 i를 사용하면, 제 2 단계의 탭값을 K 비트만큼 우측 시프트하고 제 1 단계에 대한 우측 시프팅의 양을 K 비트만큼 감소시킬 수 있다.

[0200] 이것은 일반화된 리샘플링 비율(factor)을 연산할 때를 추적하기가 다소 어려워질 수 있다. 그러나, 간단한 리샘플링 비율 2:1 또는 다른 간단한 비율을 수행할 때, 필터링 프로세스의 제2 단계에 사용하는 모든 페이지는 동일 곱 2^K를 포함하여 이 접근법이 이 특별한 경우들에서 적용될 수 있게 한다는 점을 인식하기는 용이하다.

[0201] V. 위치 계산 기술 및 도구

[0202] 공간 SVC에 대한 포지셔닝 정보를 계산하는 기술 및 도구에 대해 설명하기로 한다.

- [0203] 일부 기술 및 도구는 워드 길이 B에 초점을 맞추는 방법 및 이 워드 길이의 제약 내에서 계산의 정밀도를 최적화하는 방법에 대한 것이다. 이 정밀도를 선택하고 몇가지 필요한 워드 길이를 필요로 하는 대신에, 이 새로운 방법을 적용하면 실제 구현에서 더 높은 정밀도를 가져오게 되고 이 기술의 효과적인 적용의 범위를 넓히게 될 것인데, 그 이유는 이 제약 내에서 정확도를 최대화하기 위해 모든 이용가능한 워드 길이를 이용하기 때문이다.
- [0204] 일부 기술 및 도구는 a) 좌표계의 원점을 오프셋하고 b) 정밀도와 워드 길이/동적 범위 사이의 더 나은 절충(trade-off)을 위해 부호화(signed) 정수보다는 무부호(unsigned) 정수를 이용한다. 계산에서의 사소한 증가는 원점 오프셋 향을 각 계산된 위치에 가산하기 위해 필요하다.
- [0205] 일부 기술 및 도구는 생성될 샘플의 스트링의 서로 다른 섹션들의 계산을 서로 다른 처리 단계로 분할하는 것에 관한 것으로, 여기서 좌표계의 원점은 각 단계의 시작에서 변경된다. 계산 요건에서의 또다른 사소한 증가를 이용하여 정밀도와 워드 길이/동적 범위 사이의 더 나은 절충을 제공한다 (임의의 별도의 계산이 각 단계의 시작에서 수행되므로). 이 기술이 그 논리적 극한까지 취해진다면, 곱셈연산에 대한 필요성이 제거될 수 있고 정밀도와 워드 길이/동적 범위 사이의 절충이 더 개선될 수 있다. 그러나, 어떤 별도의 연산은 모든 샘플에 대해 수행될 필요가 있을 것이다 ("각 단계"에 필요한 별도의 계산은 모든 단계가 하나의 샘플만을 포함할 때 모든 샘플에 대해 필요해지기 때문에).
- [0206] 일반적인 주제로서, 계산된 결과들의 정밀도, 처리 요소의 워드 길이/동적 범위, 및 이 처리에 관련된 수학적 연산의 수 및 종류 (예컨대, 시프트, 덧셈 및 곱셈 연산) 사이의 바람직한 절충을 달성하기 위해 이 처리의 위치 계산부에 대한 설계에 대해 설명하기로 한다.
- [0207] 예를 들어, 설명된 기술 및 도구는 B 비트 (예컨대, 32 비트) 산술을 이용하여 유연한 정밀도 계산을 허용한다. 이것은 공간 SVC 인코더/디코더가 계산을 위해 서로 다른 산술 (예컨대, 16비트 또는 64비트 산술)로 변환할 필요없이 서로 다른 이미지 크기를 유연하게 수용하게 할 수 있다. 유연한 정밀도 B 비트 (예컨대, 32비트) 산술을 이용하여, 인코더/디코더는 유연한 비트 수를 분수 구성요소에 상용할 수 있다. 이것은 정수 구성요소를 표현하는 필요한 비트의 수가 감소 함에 따라 (예컨대, 더 작은 프레임 크기의 경우), 계산에 대한 증가된 정밀도를 허용한다. 정수 구성요소를 표현하기 위해 필요한 비트의 수가 증가함에 따라 (예컨대, 더 큰 프레임 크기의 경우), 인코더/디코더는 정수 구성요소에 대해 더 많은 비트를 사용하고 분수 구성요소에 대해 더 적은 비트를 사용할 수 있어서, 정밀도는 감소하지만 B 비트 산술을 유지한다. 이런 방법으로, 서로 다른 정밀도 및 서로 다른 프레임 크기 사이의 변경이 크게 단순해진다.
- [0208] 이 절은 일예 구현에 대한 구체적인 세부사항을 포함한다. 그러나, 여기 설명된 구체사항은 여기 설명된 원리에 따라 다른 구현에서 변경될 수 있음은 물론이다.
- [0209] A. 도입 및 위치 계산 원리
- [0210] 정확도의 큰 손실 없이 계산 요건이 훨씬 더 적은 위치 및 페이지 정보를 계산하는 기술에 대해 설명하기로 한다. 예를 들면, 설명되는 기술은 예컨대, 명목상의 동적 범위 요건을 현저히 (수십 비트) 감소시킴으로써 계산 요건을 크게 줄일 수 있다. 기본 및 인헨스먼트 계층에서 사용될 수 있는 다양한 가능한 크로마 위치를 고려하면, 루마 샘플에 상관하여 리샘플링된 크로마 샘플의 적당한 포지셔닝을 제공하는 해법을 찾는 것이 바람직하다. 이에 따라, 설명된 기술들은 루마와 크로마 위치 사이의 서로 다른 관계를 갖는 비디오 형식에 대한 위치를 계산하기 위해 조정이 이루어지게 할 수 있다.
- [0211] 확장된 공간 확장성을 위해 설계된 이전의 업샘플링 방법은 저해상도 계층을 업샘플링할 때 위치 및 페이지 정보를 계산하는 다소 번거로운 방법을 이용한다. 이것은 분모(denominator)의 업시프팅된 대략적인 역(inverse)을 크기조정하는데, 이렇게 하면 분자(numerator)가 증가할 때 (즉, 업샘플링 프로세스가 좌측에서 우측으로, 또는 위에서 아래로 진행할 때) 역 개산(approximation)에서 반올림 오차의 증대가 초래된다. 비교하면, 여기 설명되는 기술들은 우수한 정확도를 갖고 계산을 단순화한다. 특히, 위치 계산에서 동적 범위 및 우측 시프팅 양을 수십 비트 감소시키는 기술들에 대해 설명한다.
- [0212] 예를 들어, SVC 공간 업샘플링에서 사용하기 위한 정수 위치 및 페이지 포지셔닝 변수 i ($i=0..2^F-1$)를 얻기 위한 포지셔닝 정보를 계산하는 기술이 설명된다.
- [0213] 설명된 기술들은 리샘플링 프로세스를 순방향 기준 픽처 리샘플링 보다는 공간 크기조정가능한 비디오 코딩의 응용에 적용한다. 이 공간 크기조정가능한 비디오 코딩의 응용에서, 임의의 단순화가 적용될 수 있다. 일반적인 와핑(warping) 프로세스 보다는, 픽처 크기조정 연산만을 필요로 한다. 이것은 각 디멘전 마다 분리가능한

설계될 수 있다.

[0214] B. 위치 계산 설계

[0215] 새로운 (업샘플링된) 어레이에서 개념적으로 L 부터 $R > L$ 까지의 실수값 범위에 있는 샘플들의 스트링의 생성으로서 각 디멘전 (x 또는 y) 에서 문제 진술문을 고려한다. 이 실수 범위는 기준 저해상도 어레이에서 L' 부터 $R' > L'$ 까지의 범위에 대응하는 것이다.

[0216] $L \leq T \leq R$ 인 새로운 어레이에서의 위치 T 의 경우, 새로운 어레이에서의 위치에 대응하는 기준 어레이에서의 위치를 계산할 필요가 있다. 이것은 위치 $T' = L' + (T - L) * (R' - L') \div (R - L)$ 일 것이다.

[0217] L 부터 R 까지의 범위의 크기조정을 고려하는 대신에, 정수 $M > 0$ 을 정의하고 L 부터 $L + 2^M$ 까지의 범위를 동일 크기조정 비율 $(R' - L') \div (R - L)$ 로 크기조정하는 것을 고려한다. 기준 샘플 좌표에서의 대응하는 범위는 L' 부터 R'' 까지이고, 여기서 $R'' = L' + 2^M * (R' - L') \div (R - L)$ 이다. 만약 M 이 충분히 크면, 즉 $M \geq \text{Ceil}(\text{Log}_2(R - L))$ 이면, $R'' \geq R'$ 이다. (이 조건이 수학식의 적당한 기능에 꼭 필요한 것은 아니더라도 지금은 이 조건은 아래의 개념을 설명하기 위해 유효하다고 가정한다.)

[0218] 포지셔닝 계산을 위해 위치 L' 와 R'' 사이의 선형 보간을 이용할 수 있다. 위치 L 은 위치 L' 로 매핑되고, 위치 $T \geq L$ 은 위치 $((2^M - (T - L)) * L' + (T - L) * R'') \div 2^M$ 로 매핑된다. 이것은 연산의 분모를 2의 거듭제곱으로 변환하여, 이진 우측 시프트에 의해 대체되게 함으로써 나눗셈 연산의 계산상의 복잡성을 감소시킨다.

[0219] 적절한 수정이 계산을 정수화하기 위해 이루어질 수 있다. L' 및 R'' 의 값을 $1 \div 2^G$ 의 정수값으로 반올림하여 (여기서 G 는 정수), L' 는 $k \div 2^G$ 에 의해 근사적으로 계산되고 R'' 는 $r \div 2^G$ 에 의해 근사적으로 계산된다 (k 와 r 은 정수). 이 조정을 이용하여, 위치 T 가 위치 $((2^M - (T - L)) * k + (T - L) * r) \div 2^{(M+G)}$ 로 매핑되게 한다.

[0220] T 와 L 의 관련 값들은 $1 \div 2^J$ 의 정수값이어서 (J 는 정수), $T - L = j \div 2^J$ 이다. 이 조정을 이용하여, 위치 T 가 위치 $((2^{(M+J)} - j) * k + j * r) \div 2^{(M+G+J)}$ 로 매핑되게 한다.

[0221] 상기 IV 절을 다시 참조하면, 리샘플링 필터의 분수 페이즈는 $1 \div 2^F$ 단위로 정수이다. 그래서 이 유닛에서 계산된 위치는

수학식 16

[0222] $\text{Round}(((2^{(M+J)} - j) * k + j * r) \div 2^{(M+G+J-F)})$, 또는

[0223] $t' = ((2^{(M+J)} - j) * k + j * r + 2^{(M+G+J-F-1)}) \gg (M+G+J-F)$,

[0224] 또는 더 간단히

수학식 17

[0225] $t' = (j * C + D) \gg S$

[0226] 여기서

수학식 18

[0227] $S = M + G + J - F$

수학식 19

[0228] $C = r - k$

수학식 20

[0229] $D = (k \ll (M+J)) + (1 \ll (S-1))$

- [0230] $1 \div 2^F$ 의 가장 가까운 곱으로의 계산된 위치의 반올림에 앞서 여기 설명된 방법에서 생성된 유일한 오차 (L과 R 및 L'와 R'의 표현에서 오류가 없다고 가정)(두 설계에 존재하는 오차)는 $1 \div 2^G$ 의 가장 가까운 곱으로의 위치 R'의 반올림으로부터의 반올림 오류이다. 이 양은 G+M이 비교적 크면 매우 작다. 실제로, 이 오차의 원천은 약 $(T-L) \div 2^{(G+M+1)}$ 의 크기에 밀접하게 인접되어 있고, 이 결과의 계산을 위한 워드 길이 요건은 적당하고, 모듈로(modulo) 산술은 이 결과의 정수부분이 워드 길이를 최소화하도록 분리되게 하거나 계산이 다른 유사한 방법으로 분해될 수 있게 한다.
- [0231] 예컨대, F는 4 이상일 수 있다. (일부 적용에서는, F=3 또는 F=2 만 충분할 수 있다.) J의 예시 값들은 루마 위치 계산의 경우 J=1이고 크로마 샘플 위치의 경우 J=2 를 포함한다. 이들 J의 예시 값들에 대한 이론적 설명은 아래에 있다.
- [0232] 1. 부호화 B 비트 산술을 이용한 제 1 예 단순화된 위치 계산 기술
- [0233] $R' > 0$ 및 $L' > -R'$ 이면, $1 \div 2^F$ 의 단위의 정수로서 업샘플링될 픽처에서 계산될 모든 위치 t'는 -2^Z 와 2^Z-1 사이에 놓일 것이고, 여기서 $Z = \text{Ceil}(\text{Log}_2(R')) + F$ 이다. $(j * C + D)$ 계산의 워드 길이가 B 비트이고, 부호화 2의 보수 산술의 사용을 가정하면, $B-1 \geq Z+S$ 임을 필요로 할 수 있다. 이 제한이 엄격하면, 즉 $B-1 = Z+M+G+J-F$ 이면, 높은 정확도가 달성된다.
- [0234] 합리적으로 작은 픽처 크기의 경우 (예컨대, 현재 H.264/MPEG-4 AVC 표준에서 4.2 수준까지의 수준에 있어서), B=32는 워드 길이로 사용될 수 있다. B의 다른 값들도 사용될 수 있다. 매우 큰 픽처의 경우, 더 큰 B가 사용될 수 있다. 계산은 16비트 또는 다른 프로세서에서 사용을 위해 더 작은 워드 길이 서브 계산으로 용이하게 분해될 수 있다.
- [0235] 나머지 2개의 자유도는 M과 G이다. 그들의 관계는, L'를 $k \div 2^G$ 로 표현할때 반올림 오차에 대한 필요성을 피하기 위해 충분히 G가 크기만 하면, 유연하다. 따라서, SVC에 대한 다음 절에서 논의되는 쟁점에 기초하여, G=2를 선택하면,
- [0236] $M = B + F - (G + J + Z + 1)$
- [0237] 즉,
- [0238] $M = 32 + 4 - (2 + 1 + Z + 1)$
- [0239] 즉,
- [0240] $M = 32 - Z$ 이다.
- [0241] 예를 들어, B=32 및 L'=0을 갖는 1000 루마 샘플의 폭을 갖는 픽처의 루마 어레이를 업샘플링하고 싶으면, 이 제1예 위치 계산 기술을 이용하여 F=4, G=2, J=1, M=18, S=17, 및 Z=14를 사용할 수 있다.
- [0242] T가 R에 매우 근사 (또는 동일)하고 R'가 2의 정수 거듭제곱에 매우 근사 (또는 동일)할 때, 특히 $(T-L) * (R'-L') \div 2^F$ 가 클 때 (예컨대, 1/2 보다 클 때), 가설적으로 상한이 1만큼 어겨지는 것이 가능할 수 있다. 이러한 케이스를 다루기 위한 조정은 간단하지만, 여기서는 더이상 이에 관해 논의하지 않는다.
- [0243] 2. 무부호 B 비트 산술을 이용한 제2예 위치 계산 기술
- [0244] 저해상도 픽처에서 계산될 모든 위치가 0 보다 크거나 같으면 (이것은 좌표계의 원점에 적절한 오프셋을 가산함으로써 참이 될 수 있는 것이다), 부호화 2의 보수 산술 보다는 무부호 정수 산술을 이용하여 $t' = (j * C + D)$ 를 계산하는 것이 더 나은 선택일 수 있다. 이것은 계산에서 오버플로(overflow)없이 하나 더 많은 비트의 동적 범위를 허용하여 (즉, B-1 비트보다는 B 비트의 동적 범위 크기를 이용할 수 있다), M (또는 G)과 S를 각각 1씩 증가하고 계산된 결과의 정확도를 더 증가시킨다. 따라서, 좌표계의 원점을 조정하기 위해 오프셋 E를 포함한 후에, 계산의 형태는 단지 $t' = (j * C + D) \gg S$ 이기 보다는 $t' = ((j * C + D') \gg S) + E$ 일 것이다.
- [0245] 원점 오프셋 E가 필요하지 않을 때를 식별함으로써 무부호 산술을 포함하는 이 더 정확한 방법에 관한 세부사항을 다음과 같이 더 제공한다.
- [0246] - 상기한 바와 같이 B, F, G, J 및 Z에 대한 값들을 선택한다.

- [0247] - $M=B+F-(G+J+Z)$ 를 설정한다.
- [0248] - 수학식 18, 19 및 20 에서 각각 규정된 바와 같이 S, C 및 D를 계산한다. 여기서 D는 부호화 수로 계산된다.
- [0249] - D가 0 보다 크거나 같으면, 원점 오프셋은 필요하지 않고 (즉, E를 사용안함) 계산은 무부호 산술을 이용하여 간단히 $t'=(j*C+D)>>S$ 로서 수행될 수 있고 그 결과는 상기 V.B.1 절에서 설명된 제1예 위치 계산 기술보다 더 높은 정확도를 가질 것이다.
- [0250] 무부호 정수를 이용한 계산을 가능하게 함으로써 정확도를 증가시키는 것 이외에, 원점을 오프셋하는 것은, 때로 Z 값의 감소를 가능하게 함으로서 향상된 정확도를 제공하기 위해 사용될 수 있다. 원점 오프셋이 없다면, Z는 R'의 함수이다. 그러나, 원점 오프셋이 있으면, Z를 R'-L'의 함수로 할 수 있고, 이것이 Z의 더 작은 값으로 된다면 이것은 계산을 더욱 정확하게 할 것이다.
- [0251] 원점을 오프셋하는 한가지 방법을 보여줌으로써 무부호 산술을 포함하는 더 정확한 방법에 대한 세부사항을 더 제공하여, D' 및 E를 다음과 같이 유도한다.
- [0252] - 상기한 바와 같이, B, F, G 및 J에 대한 값을 선택한다.
- [0253] - $Z=\text{Ceil}(\text{Log2}(R'-L'))+F$ 를 설정한다.
- [0254] - $M=B+F-(G+J+Z)$ 를 설정한다.
- [0255] - 상기 수학식 18, 19 및 20에 각각 규정된 바와 같이 S, C 및 D를 계산한다. 여기서, D는 부호화 수로서 계산된다.
- [0256] - $E=D>>S$ 를 설정한다.
- [0257] - $D'=D-(E<<S)$ 를 설정한다.
- [0258] - 위치 계산은 $t'=((j*C+D')>>S)+E$ 로서 수행될 수 있다.
- [0259] D' 와 E (및 M, S 및 Z)가 이런 식으로 계산되면, 수학식 $t'=((j*C+D')>>S)+E$ 의 수학적 결과는, (j*C+D)의 값이 때로 0 부터 2^B-1 까지의 값의 범위의 외부에 있을 수 있는 반면 (j*C+D')의 값은 그렇지 않을 것이라는 점을 제외하고는, 실제로 이론적으로 수학식 $t'=(j*C+D)>>S$ 의 결과와 동일할 것이다.
- [0260] 예를 들어, B=32 및 L'=0 인 1000 루마 샘플의 폭을 갖는 픽처의 루마 어레이를 업샘플링하고 싶으면, 이 제2예 위치 계산 기술을 이용하여 F=4, G=2, J=1, M=19, S=18 및 Z=14 를 이용할 수 있다. 동등하게 잘 작용할 또다른 가능성은, j*C+D'의 모든 값들이 음수가 안되도록 원점을 오프셋하고 이에 따라 무부호 산술을 이용하여 0 부터 2^B-1 까지의 B비트 계산 범위의 사용을 허용하기 보다는, 부호화 산술을 이용하여 $-2^{(B-1)}$ 부터 $2^{(B-1)}-1$ 까지의 B비트 계산 범위의 사용을 허용하기 위해 또다른 $2^{(B-1)}$ 만큼 원점을 우측으로 더 오프셋하는 것일 것이다.
- [0261] 이전의 절에서 제1예 위치 계산 기술에서와 같이, T가 R에 매우 근사 (또는 동일)하고 R'-L'가 2의 정수 거듭제곱에 매우 근사 (또는 동일)할 때 "코너 케이스 (corner case)" 조정이 필요할 수 있다.
- [0262] 3. 위치 계산을 위한 다중 단계 기술의 예
- [0263] 동일한 수학식, 예컨대, 생성될 샘플의 범위를 커버하는 j의 모든 값 (즉, L과 R 사이의 T의 모든 값)에 대해 동일한 변수 값 C, D', S 및 E를 가지고 $t'=((j*C+D')>>S)+E$ 를 이용하여 계산을 수행할 수 있도록 설계가 되어 있는 방법을 논의하였다. 이제 어떻게 이 가정이 완화되어 더 높은 정확도 및/또는 감소된 계산 동적 범위 요건을 가능하게 하는지에 대해 논의한다.
- [0264] 일반적으로, 리샘플링 프로세스는 동일하게 이격된 위치에서 연속적인 샘플의 스트링을 생성하기 위해 좌측에서 우측 (또는 위에서 아래로) 진행한다. 상기 V.B.2 절에서 설명된 제2예 위치 기술에서, 오프셋 파라미터 E를 이용한 원점의 변경이, 위치 계산의 (j*C+D') 부분을 계산하기 위해 사용된 레지스터의 B 비트 동적 범위를 활용하기 위해 어떻게 이용될 수 있는지를 보여주었다.
- [0265] 이전의 절에서 D의 S 최하위 비트만이 D'에서 유지되었고 나머지는 E로 이동되었음을 상기하기 바란다. 따라서, (j*C+D')의 계산을 위한 남은 주요 쟁점은 j*C의 크기이다.
- [0266] T와 L은 $1 \div 2^J$ 의 정수 곱임을 상기하라. 일반적으로, 예컨대 연속적으로 생성되는 샘플들 사이의 2^J 의 간격을

가지고 고해상도 픽처에서의 정수값 증분으로 샘플의 스트링을 생성하기 위해 업샘플링 프로세스를 수행한다. 따라서, p 와 N 의 어떤 값에 대해 $i=0$ 부터 $N-1$ 까지에 대해 위치 $T_i=(p+i*2^J)\div 2^J$ 에 대응하는 위치 t'_i 를 계산하고자 한다.

[0267] 이 프로세스는 p 및 N 의 일부 값에 대해 도 12의 의사코드(pseudo-code) (1200)에 도시된 바와 같이 의사코드로 요약될 수 있다. i 가 N 쪽으로 증가함에 따라, q 의 값이 증가하고, q 의 최대값은 B 비트의 이용가능한 동적범위 내에 유지되어야 한다. q 에 대해 계산된 최대값은 $(p+(N-1)*2^J)*C+D'$ 이다.

[0268] 이제, 이런 식으로 하나의 루프에서 모든 샘플을 생성하는 대신에, 프로세스를 다중 단계, 예컨대 2단계로 분할하는 것을 고려한다. 예를 들어, 2단계 프로세스에서, 제1 단계는 제1 $N_0 < N$ 샘플을 생성하고, 제2 단계는 나머지 $N-N_0$ 샘플을 생성한다. 또한, p 는 루프에 대해 상수이므로, 그 임팩트를 제1 단계 전에 D' 및 E 로 이동시킬 수 있다. 이렇게 하면 도 13에서 의사코드 (1300)에 도시된 2단계 프로세스가 된다.

[0269] 의사코드 (1300)에서 각 단계의 시작에, 원점이 리셋되어서 이 단계에 대한 q 의 제1값의 S 최하위 비트 이외는 모두 E 로 이동되었다 (즉, 제1 단계에서 E_0 로, 제2 단계에서 E_1 으로). 따라서, 2단계의 각각의 동작 동안, q 는 더 작은 동적범위를 필요로 한다. 이 프로세스를 이런 식으로 단계들로 나눈 후, q 의 최대값은 $N_0*C'+D_0$ 또는 $((N-N_0-1)*C'+D_1$ 중 큰 것일 것이다. 그러나, D_0 와 D_1 은 각각 S 비트의 무부호 동적 범위 보다 크지 않으므로, 이것은 일반적으로 이전에 설명된 단일 단계 설계에서 보다 더 작은 최대값일 것이다. 이 단계에서 생성된 샘플의 수 (즉, 제1 단계에서 N_0 , 제2 단계에서 $N-N_0$)는 관련된 계산에 대한 동적 범위에 영향을 미칠 수 있다. 예를 들어, 각 단계에서 샘플의 더 작은 수를 이용하면 관련된 계산에 대해 더 작은 동적 범위로 될 것이다.

[0270] 각 단계는 더 많은 단계로 더 분할될 수 있고, 따라서 N 개의 전체 샘플의 생성은 이러한 더 작은 단계의 수로 더 분해될 수 있다. 예를 들어, 이 프로세스는 동일한 크기의 단계로 분리될 수 있어서 예컨대 8 또는 16개의 연속적인 샘플의 블록들이 각 단계에서 생성된다. 이 기술은 q 계산을 위한 동적범위 B 의 필요한 비트 수를 줄이거나, 동적 범위를 동일하게 유지하면서 계산의 정밀도를 증가시키기 위해 (S 및 $G+M$ 을 증가) 사용될 수 있고, 아니면 이 두 장점의 혼합하여 사용될 수 있다.

[0271] 위치 계산 프로세스를 단계들로 분해하는 이 기술은, 예컨대 샘플이 오디오 신호의 경우 아날로그-디지털 컨버터로부터 도착할 때 샘플링 레이트 변환을 수행할 때처럼, 입력 샘플의 매우 긴 스트링을 따라 (개념적으로 이 스트링은 무한히 길 수 있다) 연속적인 리샘플링 프로세스를 수행하기 위해 사용될 수도 있다. 명확히, 이 프로세스를 유한크기의 단계로 분해하고 원점을 단계 마다 점증적으로 재설정하지 않는다면, 무한히 긴 샘플 스트링은 이전의 절에서 설명된 기술들에 의해 처리될 수 없는데, 이는 처리 워드 길이에서 무한한 동적 범위를 필요로 할 것이기 때문이다. 그러나, 이 기술을 유효하게 무한한 스트링 길이에 적용할 때의 어려움은, $1\div 2^6$ 의 곱을 나타내는 정수 단위에서의 가정적 벤치마크(benchmark) 위치 L' 및 R' 의 표현에 의해 반올림 오차가 발생하지 않을 때 유효하게 무한한 길이로의 적용이 유용할 것이므로, 이러한 기술에서 대한 실질적인 제한은 아니다.

[0272] 다중 단계 위치 계산 기술이 적용될 수 있는 환경에서, 전체 레이트 변환 프로세스를 통해 위치 계산의 연산에 있어서 반올림 오차의 "드리프팅" 누적이 없는 무한 길이의 샘플 스트링을 따라 계산이 수행되는 방법을 제공한다.

[0273] 4. 위치 계산의 증분 동작의 예

[0274] 상기 다중 단계 분해 개념에 대한 흥미로운 특별 케이스는 각 단계에서 생성될 샘플의 수가 단계당 하나의 샘플로 감소된 때이다. 도 14에서 의사코드 (1400)는 $i=0$ 부터 $N-1$ 까지에 대해 N 개의 위치 t'_i 를 생성하는 프로세스를 나타낸다.

[0275] 이 프로세스는 업샘플링 프로세스로서 설명되므로 (동일한 원리가 다운샘플링 프로세스에도 적용될 수 있지만), i 의 각 증분마다, 고해상도 픽처에서 1의 간격이 있고 따라서 저해상도 픽처에서 1 미만 또는 1의 증분이 있음을 알고 있다. 저해상도 픽처에서 공간 위치에서 1의 증분은 C' 에 대해 $2^{(S+F)}$ 의 값에 해당한다. 또한, $D' < 2^S$ 임을 알고 있다. 따라서, $q=C'+D'$ 는 0부터 $2^{(S+F)}+2^S$ 미만까지의 범위를 갖고, 따라서 q 는 무부호 정수 산술을 이용하여 $B=S+F+1$ 비트 미만의 동적 범위 요건으로 계산될 수 있다. 일 구현에서, 이 동적 범위 요건은 픽처

크기에 대해 불변이다 (즉, R' 또는 $R'-L'$ 값에 종속하지 않는다).

- [0276] 크기조정가능한 비디오 코딩 및 많은 다른 이러한 적용의 경우, 1에 매우 근사한 업샘플링 비율을 지원할 실제적인 필요성이 없을 수 있다. 이러한 적용에서, C' 는 실제로 $S+F$ 비트 이상을 필요로 하지 않는다고 가정할 수 있다.
- [0277] 예를 들어, $B=32$ 및 $L'=0$ 을 갖는 1000 루마 샘플의 폭을 갖는 픽처의 루마 어레이를 업샘플링하고 싶으면, 이 방법을 이용하여 $F=4$, $G=2$, $J=1$, $M=29$, $S=28$, 및 $Z=14$ 를 이용할 수 있다. 이 결과는 B 의 더 작은 값이 더 합리적인 선택처럼 보일 정도로 매우 정확할 것이다.
- [0278] 다른 경우에, $B=16$ 및 $L'=0$ 을 갖는 1000 루마 샘플의 폭을 갖는 픽처의 루마 어레이를 업샘플링하고 싶으면, 이 방법을 이용하여 $F=4$, $G=2$, $J=1$, $M=13$, $S=12$, 및 $Z=14$ 를 이용할 수 있다.
- [0279] 수행될 업샘플링 연산의 상황에 대해 더 많이 알면, 더 많은 최적화 기회가 있을 수 있다. 예를 들어, 업샘플링 비율이 2보다 훨씬 더 크면, 동적 범위 요건은 또다른 비트만큼 감소될 것이고, 4, 16 등등보다 더 큰 업샘플링 비율에 대해 계속 감소될 것이다.
- [0280] 이 절에서 상기 증분 위치 계산 기술의 예를 참조하여 설명된 변경 중 어느 것도 (상기 다중 단계 위치 계산 기술의 예에 관련된) C , D 및 S 의 주어진 값에 대해 위치 t'_i 의 실제 계산된 값에 전혀 영향을 주지 않는다. 계산을 지원하기 위해 필요한 동적 범위만이 변경된다.
- [0281] 이런 형태의 분해에 대한 의사코드 (1400)에서의 내부 루프는 임의의 곱셈 연산을 전혀 필요로 하지 않는다. 이 사실은 일부 컴퓨팅 프로세서에서 감소된 계산 시간을 제공하는 데에 도움이 될 수 있다.
- [0282] 5. 추가 설명
- [0283] $2:1$, $3:2$ 등과 같은 통상의 리샘플링 비율의 경우 $-1 \div 2^G$ 단위의 정수로서 위치 L' 및 R' 에 근사 (approximating)시키기 위해 어떤 반올림도 필요하지 않는 케이스 -, 이들 방법을 사용할 때 전혀 반올림 오차가 없다 ($1 \div 2^F$ 단위의 정수로 최종 결과를 반올림할 때 반올림 오차가 유도될 수 있는 경우 제외; 여기서 상기 오차는 위치 계산 방법에 상관없이 존재할 오차임).
- [0284] C. 루마 및 크로마 위치 및 관계
- [0285] 루마 샘플링 그리드 인덱스 좌표에 대해 완전히 새로운 (업샘플링된) 픽처 및 기준 픽처 어레이의 정확한 정렬을 가정하면, 현재 픽처 좌표에서의 위치 L 및 R 은 $L=-1/2$ 및 $R=W-1/2$ 이고, 여기서 W 는 관련된 리샘플링 디멘전에 따라 수직 또는 수평으로의 이미지에서의 샘플의 수이다. 마찬가지로, 이미지 공간 좌표계의 원점을 그리드 인덱스(grid index) 0의 위치의 좌측 (또는 위)로의 하프-샘플(half-sample)로 설정하고 이미지 공간 좌표로부터 그리드 인덱스 값으로 변환할 때 $1/2$ 를 추가할 수 있어서, 공간 좌표계에서 계산을 수행할 때 음수를 다룰 필요가 없다.
- [0286] 기준 (저해상도) 픽처에서의 위치 L' 및 R' 는 동일한 방식으로 샘플링 그리드 좌표를 참조하고, 이 경우 W 는 새로운 픽처에서 보다는 기준 픽처에서의 샘플의 수이다.
- [0287] 크로마 샘플링 그리드의 경우 (새로운 픽처든지 기준 픽처든지 간에), 상황은 다소 덜 간단하다. 루마에 대해 크로마 샘플의 지정된 정렬을 구성하기 위해, 루마 샘플에 의해 표현된 직사각형과 동일하도록 크로마 샘플에 의해 표현된 이미지 직사각형을 고려한다. 이는 다음과 같은 케이스를 생성한다:
- [0288] - 수평으로, $4:2:0$ 크로마 샘플링 타입 0, 2 및 4의 경우 (도 5D 참조), 현재 픽처 좌표는 $L=-1/4$ 및 $R=W-1/4$ 로 정의된다.
- [0289] - 수평으로, $4:2:0$ 크로마 샘플링 타입 3, 1 및 5의 경우 (도 5D 참조), 현재 픽처 좌표는 $L=-1/2$ 및 $R=W-1/2$ 로 정의된다.
- [0290] - 수직으로, $4:2:0$ 크로마 샘플링 타입 2 및 3의 경우 (도 5D 참조), 현재 픽처 좌표는 $L=-1/4$ 및 $R=W-1/4$ 로 정의된다.
- [0291] - 수직으로, $4:2:0$ 크로마 샘플링 타입 0 및 1의 경우 (도 5D 참조), 현재 픽처 좌표는 $L=-1/2$ 및 $R=W-1/2$ 로 정의된다.
- [0292] - 수직으로, $4:2:0$ 크로마 샘플링 타입 4 및 5의 경우 (도 5D 참조), 현재 픽처 좌표는 $L=-3/4$ 및 $R=W-3/4$ 로

정의된다.

- [0293] - 수평으로, 4:2:2 크로마 샘플링의 경우, 업계 실무에서 통상적으로 사용되는 4:2:2 샘플링에 대한 현재 픽처 좌표는 $L=-1/4$ 및 $R=W-1/4$ 로 정의된다.
- [0294] - 수직으로, 4:2:2 크로마 샘플링의 경우, 업계 실무에서 통상적으로 사용되는 4:2:2 샘플링에 대한 현재 픽처 좌표는 $L=-1/2$ 및 $R=W-1/2$ 로 정의된다.
- [0295] - 수평 및 수직으로, 4:4:4 크로마 샘플링의 경우, 현재 픽처 좌표는 $L=-1/2$ 및 $R=W-1/2$ 로 정의된다.
- [0296] 오프셋은 좌표계의 원점을 충분히 위치 L의 좌측에 놓고 음수로 작업할 필요를 피하기 위해 사용될 수 있다.
- [0297] 정수 좌표 및 분수 페이지 오프셋 나머지는, 분수 오프셋 L을 보상하기 위하여 업샘플링된 어레이에서 생성된 샘플들의 정수 좌표 위치를 조정하고 그 후 V.B 절의 끝에서 설명된 변환을 적용함으로써 계산된다. 개념적으로, 이 결과를 F 비트만큼 우측으로 시프팅하면 기준 픽처로의 정수 좌표 포인터로 되고, 좌측 시프트된 정수 좌표 (F 비트 만큼 시프트된)를 빼는 것은 페이지 오프셋 나머지를 제공한다.
- [0298] D. 업샘플링에 대한 위치 계산에 대한 별도의 정밀도
- [0299] 이 절은 상기 V.C.4 절의 위치 계산 방법을 H.264 SVC Extension에 사용될 수 있는 업샘플링 프로세스와 같은 특정 업샘플링 프로세스에 매핑하는 방법을 설명한다. 이 위치 계산은 순차 및 비월 프레임 형식 뿐만 아니라 다양한 크로마 형식에서 루마 및 크로마 채널에 대한 정밀도를 최대화하기 위해 매우 유연한 방법으로 적용된다. 이 절에서 설명된 기술은 구현에 따라서 그리고 서로 다른 업샘플링 프로세스마다 가변적일 수 있다.
- [0300] 상기 위치 계산에서 (상기 V.A-C 절에서), 크기재조정(rescaling) 파라미터 (변수 C, 이하 수학식에서 ΔX (또는 ΔY)로 표기)는 2^J (루마의 경우 $J=1$, 크로마의 경우 $J=2$)인 크기조정 비율로 크기증가되어 각 샘플 위치를 좌측에서 우측으로 또는 위에서 아래로 생성하기 위해 추가된 증분을 형성한다. 크기조정은 크기증가된 (up-scaled) 증분이 16비트로 맞춰지도록 선택되었다.
- [0301] 1. 크기조정 위치 계산에 대한 최대 정밀도
- [0302] 위치 계산 방법을 적용하는 직접적인 방법은, 각 샘플 위치를 좌에서 우로 또는 위에서 아래로 생성하기 위해 추가된 증분을 형성하기 위해 크기조정 비율 2^J (루마의 경우 $J=1$, 크로마의 경우 $J=2$)로 크기재조정 파라미터를 크기증가시키는 것이다. 이 크기조정 파라미터는 크기증가된 증분이 16비트와 같은 특정 워드 길이로 맞춰질 것을 보장하도록 선택된다. 더욱 유연한 설계는 위치 정밀도를 최대화하기 위해 다음 절에서 설명된다.
- [0303] a. 루마 채널
- [0304] "직접적인" 루마 위치 계산법은 $F=4$ 및 $S=12$ (수평방향을 따라)에 대한 다음 예 수학식을 이용하여 요약될 수 있다:
- [0305] $\Delta X = \text{Floor}(((\text{BasePicWidth} \ll 15) + (\text{ScaledBaseWidth} \gg 1)) \div \text{ScaledBaseWidth})$
- [0306] $x_f = ((2 * (x_P - \text{ScaledBaseLeftOffset}) + 1) * \Delta X - 30720) \gg 12$
- [0307] 여기서, BasicPicWidth는 기본 계층 또는 저해상도 픽처의 수평 해상도이고; ScaledBaseWidth는 고해상도 픽처 영역 또는 윈도우의 수평 해상도이고; ΔX 는 중간 크기조정 파라미터로서, 이 경우에는 업샘플링 비율의 역의 32768 배의 반올림된 개산치(approximation)이고; x_P 는 고해상도 픽처에서 샘플 위치를 나타내고; ScaledBaseLeftOffset은 고해상도 픽처에서 픽처 윈도우의 상대적 위치를 나타내고; Floor()는 인수(argument) 보다 작거나 동일한 최대 정수를 나타낸다. 정수값 30720은, 상기 V.C 절의 시작에서 논의된 바와 같이, 우측 시프트 전에 반올림 오프셋으로서 2^{S-1} 을 가산하고 루마 샘플링 그리드 기준 위치의 하프-샘플 오프셋에 대해 $2^S * 2^F / 2$ 를 감산하여 생긴다.
- [0308] x_P 의 각 증분은 수학식 내에서 $2 * \Delta X$ 의 증분으로 된다는 것은 주목할 만하다. 양 $2 * \Delta X$ 의 LSB는 항상 0이고, 1비트의 계산 정밀도는 실질적으로 낭비되고 있다. 대략적으로 하나의 여분 비트의 정밀도는 복잡도의 상당한 증가없이 이들 수학식을 다음과 같이 변경함으로써 얻어질 수 있다:
- [0309] $\Delta X = \text{Floor}(((\text{BasePicWidth} \ll 16) + (\text{ScaledBaseWidth} \gg 1)) \div \text{ScaledBaseWidth})$

- [0310] $xf = ((xP - \text{ScaledBaseLeftOffset}) * \text{deltaX} + (\text{deltaX} \gg 1) - 30720) \gg 12$
- [0311] 또는 (약간) 더 정확한 형태는 다음과 같다:
- [0312] $\text{deltaXa} = \text{Floor}(((\text{BasePicWidth} \ll 16) + (\text{ScaledBaseWidth} \gg 1)) \div \text{ScaledBaseWidth})$
- [0313] $\text{deltaXb} = \text{Floor}(((\text{BasePicWidth} \ll 15) + (\text{ScaledBaseWidth} \gg 1)) \div \text{ScaledBaseWidth})$
- [0314] $xf = ((xP - \text{ScaledBaseLeftOffset}) * \text{deltaXa} + \text{deltaXb} - 30720) \gg 12$
- [0315] 이들 두 형태 중 후자는 더 높은 정확도 및 무시가능한 복잡도 영향 때문에 제안된다 (비록 정밀도 차도 매우 작은 것 같아도).
- [0316] 나눗셈 계산이 수행되기 어려운 처리 구조에서, 이들 수학적 중 하나의 결과를 갖는 것은 다른 것의 계산을 단순화시킬 수 있다. deltaXa 의 값은 항상 $2 * \text{deltaXa}$ 더하기 또는 빼기 1의 범위에 있을 것이다. 따라서, 다음의 단순화된 규칙은 deltaXa 의 계산을 위해 나눗셈 연산을 수행할 필요를 제거하기 위해 유도될 수 있다.
- [0317] $\text{deltaXa} = (\text{deltaXb} \ll 1)$
- [0318] $\text{remainderDiff} = (\text{BasePicWidth} \ll 16) + (\text{ScaledBaseWidth} \gg 1) - \text{deltaXa}$
- [0319] $\text{if}(\text{remainderDiff} < 0)$
- [0320] $\text{deltaXa}--$
- [0321] $\text{else if}(\text{remainderDiff} \geq \text{ScaledBaseWidth})$
- [0322] $\text{deltaXa}++$
- [0323] b. 크로마 채널
- [0324] 4:2:0 샘플링에 대해 (루마의 경우 상술한 바와 같이 $J=1$ 보다는 크로마의 경우 $J=2$ 를 이용) 크로마 위치의 표현을 가능하게 하기 위해 설계의 이 부분에서 2의 인수 곱셈기 (factor-of-two multiplier) 대신에 4의 인수 곱셈기 (factor-of-four multiplier)가 크로마 채널을 위해 사용될 수 있다. 따라서, "직접적인(direct)" 수학적 식은 다음과 같다:
- [0325] $\text{deltaXC} = \text{Floor}(((\text{BasePicWidthC} \ll 14) + (\text{ScaledBaseWidthC} \gg 1)) \div$
- [0326] $\text{ScaledBaseWidthC})$
- [0327] $xfC = (((4 * (xC - \text{ScaledBaseLeftOffsetC}) +$
- [0328] $(2 + \text{scaledBaseChromaPhaseX})) * \text{deltaXC})$
- [0329] $+ 2048) \gg 12) - 4 * (2 + \text{baseChromaPhaseX})$
- [0330] 여기서, baseChromaPhaseX 및 $\text{scaledBaseChromaPhaseX}$ 는 저해상도 및 고해상도 픽처에 대한 크로마 샘플링 그리드 위치 오프셋을 각각 나타낸다. 이들 파라미터의 값은 인코더로부터 디코더로 보내진 정보로서 명시적으로 전달될 수 있거나, 적용에 따라 결정되는 특정 값을 가질 수 있다. 모든 다른 변수는 크로마 채널로의 적용을 나타내기 위해 추가적인 "C" 접미사를 갖는 루마 채널에 대해 정의된 것과 유사하다.
- [0331] xC 의 각 증분은 수학적 내에서 $4 * \text{deltaXC}$ 의 증분으로 된다. 따라서, 대략적으로 2개의 여분의 비트의 정밀도가 복잡도의 실질적인 증가없이 이들 수학적 식을 다음과 같이 변경함으로써 얻어질 수 있다:
- [0332] $\text{deltaXC} = \text{Floor}(((\text{BasePicWidthC} \ll 16) + (\text{ScaledBaseWidthC} \gg 1)) \div$
- [0333] $\text{ScaledBaseWidthC})$
- [0334] $xfC = (((xC - \text{ScaledBaseLeftOffsetC}) * \text{deltaXC}$
- [0335] $+ (2 + \text{scaledBaseChromaPhaseX})) * ((\text{deltaXC} + K) \gg 2)$
- [0336] $+ 2048) \gg 12) - 4 * (2 + \text{baseChromaPhaseX})$
- [0337] 여기서, $K=0, 1$ 또는 2 이다. $K=0$ 을 사용하면 별도의 연산을 피할 수 있을 것이다. $K=1$ 또는 $K=2$ 를 사용하면 약간 더 높은 정확도를 가질 것이다.

- [0338] 대응하는 약간 더 정확한 형태는 다음과 같이 될 것이다:
- [0339] $\text{deltaXCa} = \text{Floor}(((\text{BasePicWidthC} \ll 16) + (\text{ScaledBaseWidthC} > 1)) \div$
- [0340] $\text{ScaledBaseWidthC})$
- [0341] $\text{deltaXCb} = \text{Floor}(((\text{BasePicWidthC} \ll 14) + (\text{ScaledBaseWidthC} > 1)) \div$
- [0342] $\text{ScaledBaseWidthC})$
- [0343] $\text{xC} = (((\text{xC} - \text{ScaledBaseLeftOffsetC}) * \text{deltaXCa} +$
- [0344] $(2 + \text{scaledBaseChromaPhaseX})) * \text{deltaXCb})$
- [0345] $+ 2048) >> 12) - 4 * (2 + \text{baseChromaPhaseX})$
- [0346] 루마 경우와 같이, 복잡도 차가 무시할만한 것 같으므로 (비록 정밀도 차도 매우 작은 것 같지만) 후자의 변형이 더 바람직하다.
- [0347] c. 비월 필드 좌표
- [0348] 픽처의 좌표계에 대한 기준은 일반적으로 루마 프레임 좌표에서 하프-샘플 위치에 기초하여, 결과적으로 상술한 바와 같이 루마 좌표 기준 위치의 경우 크기조정 인수 2로 된다. 루마 프레임 좌표에서 하프-샘플 시프트는 4:2:0 크로마 프레임 좌표에서 쿼터-샘플(quarter-sample) 시프트에 상응하는데, 이는 상술한 바와 같이 크로마 좌표에 대한 크기조정에서 인수 2 보다는 인수 4를 현재 사용한다.
- [0349] 수평적으로 프레임을 나타내는 코딩된 픽처와 비월 비디오의 단일 필드를 나타내는 것들에 대한 연산에 있어서 실질적인 차이가 없다. 그러나, 코딩된 픽처가 단일 필드를 나타낼 때, 루마 프레임 수직 좌표에서의 하프-샘플 위치 시프트는 루마 필드 수직 좌표에서의 쿼터-샘플 위치 시프트에 대응한다. 따라서, 크기조정 인수 2가 아닌 4가 수직 루마 좌표 위치의 계산에서 적용되어야 한다.
- [0350] 마찬가지로, 코딩된 픽처가 단일 필드를 나타낼 때, 루마 프레임 수직 좌표에서의 하프-샘플 위치 시프트는 크로마 필드 수직 좌표에서의 1/8 샘플 위치 시프트에 대응한다. 따라서, 크기조정 인수 4가 아니라 8이 수직 크로마 좌표 위치의 계산에서 적용되어야 한다.
- [0351] 코딩된 필드 픽처에서 수직 좌표 위치의 계산을 위한 이들 크기조정 인수는 코딩된 프레임 픽처에서 증분 계산을 위해 상술된 바와 같은 방식으로 deltaY 수직 증분 계산에 포함될 수 있다. 이 경우, 적용되는 증가된 크기조정 인수때문에, 정밀도 향상은 대략적으로 루마 위치에 대한 2비트의 추가된 정밀도 및 크로마에 대한 3비트의 추가된 정밀도가 된다 (수직으로).
- [0352] 2. 4:2:2 및 4:4:4 크로마 제한 및 정제(refinement)
- [0353] V.D.1.b 절의 위치 계산 방법은 루마에서 보다 크로마에 대해 서로 다른 곱셈 인수의 사용을 필요로 한다. 이는 4:2:0 비디오의 경우 이해되고 또한 수평적으로 4:2:2 비디오의 경우 합리적이지만, 수직적으로 4:2:2 비디오의 경우 또는 수평적 또는 수직적으로 4:4:4 비디오의 경우에는 불필요한데, 그 이유는 이 경우 루마 및 크로마 해상도가 동일하고 따라서 추측컨대 루마 및 크로마 샘플이 동일한 장소를 공유하기 때문이다.
- [0354] 그 결과, V.D.1.b 절의 방법은 루마 및 크로마 해상도가 일부 디멘전에서 동일하고 페이즈 시프트가 의도되지 않은 때에도 루마 및 크로마 위치를 결정하기 위한 별도의 계산을 필요로 할 수 있는데, 이것은 반올림이 두 케이스에 약간 달리 수행될 것이기 때문이다. 이것은 바람직하지 않은 현상이므로, 이 절에서는 4:2:2 및 4:4:4 샘플링 구조에서 사용하기 위한 다른 크로마 취급이 제안된다.
- [0355] a. 4:2:2 수직 및 4:4:4 수평 및 수직 위치
- [0356] 4:2:2 비디오의 수직 디멘전 및 4:4:4 비디오의 수직 및 수평 디멘전의 경우, 크로마 페이즈의 사용자 지정 제어를 위한 명백한 필요성이 없다. 따라서, 크로마 해상도가 일부 디멘전에서 루마 해상도와 동일할 때마다, 크로마 위치의 계산을 위한 수학적식은, 크로마 샘플링 형식이 특정 디멘전에서 루마 및 크로마에 대해 동일한 해상도를 가질 때마다 루마 및 크로마에 대해 정확히 동일한 위치를 계산하게 되도록 수정되어야 한다. 한가지 선택은 크로마 위치 변수를 루마 위치 변수와 동일하게 설정하는 것이고, 또다른 선택은 크로마 위치 수학적식들이 동일한 결과를 갖도록 설정하는 것이다.

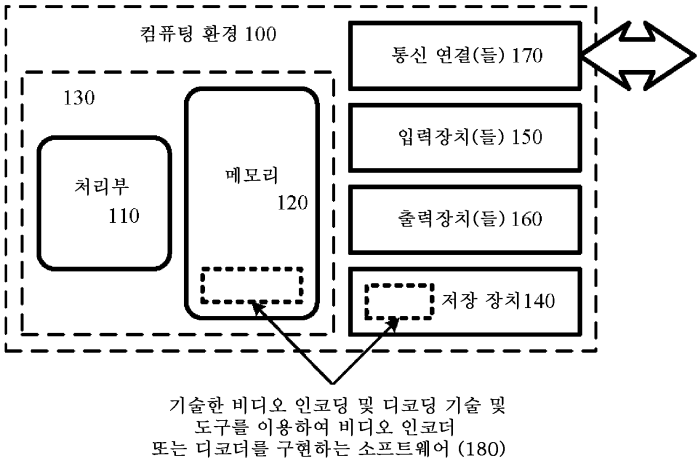
- [0357] b. 4:2:2 수평 위치
- [0358] 4:2:2 비디오에 대해 수평적으로 크로마 페이즈 조정을 허용하는 것에 기능적 문제는 없지만, 예컨대 V.D.1.b 절의 수학적식에서 scaledBaseChromaPhaseX 또는 BaseChromaPhaseX 에 대해 -1 값에 대응하는 것과 같은, 4:2:2 에 대해 사용하고 있는 오직 한 종류의 수평 서브샘플링 구조가 있다면, 컬러 샘플링 형식이 4:2:2 일 때마다 이들 값들이 사용되게 하는 것을 고려하는 것이 바람직할 수 있다.
- [0359] VI. 확장 및 대안
- [0360] 여기 설명된 기술 및 도구는 예컨대 ITU-T 국제표준권고 H.263의 부록 P에 기재되어 있는 바와 같이 기준 픽처 리샘플링을 이용한 다중 해상도 비디오 코딩에도 적용될 수 있다.
- [0361] 여기 설명된 기술 및 도구는 또한 픽처 샘플 어레이의 업샘플링 뿐만 아니라 잔차 데이터 신호 또는 다른 신호의 업샘플링에도 적용될 수 있다. 예를 들면, 여기 설명된 기술 및 도구는 예컨대 ITU-T 국제표준권고 H.263의 부록 Q에 기재되어 있는 바와 같이 감소된 해상도 업데이트 코딩을 위한 잔차 데이터 신호의 업샘플링에도 적용될 수 있다. 또다른 예로서, 여기 설명된 기술 및 도구는 공간 크기조정가능한 비디오 코딩을 위한 설계에서 저해상도 잔차 신호로부터 고해상도 잔차 신호의 예측을 위한 잔차 데이터 신호의 업샘플링에도 적용될 수 있다. 또다른 예로서, 여기 설명된 기술 및 도구는 공간 크기조정가능한 비디오 코딩을 위한 설계에서 움직임 벡터 필드의 업샘플링에도 적용될 수 있다. 또다른 예로서, 여기 설명된 기술 및 도구는 그래픽 이미지, 정지 사진 픽처, 오디오 샘플 신호 등의 업샘플링에도 적용될 수 있다.
- [0362] 본 발명의 원리를 다양한 실시예를 참조하여 설명하고 예시하였지만, 상기 실시예들은 이러한 원리로부터 벗어남없이 정렬 및 세부사항에 있어서 수정될 수 있음은 인식될 것이다. 달리 나타내지 않았으면, 여기 설명된 프로그램, 프로세스 또는 방법들은 임의의 특정 유형의 컴퓨팅 환경에 관련되거나 한정되지 않음은 물론이다. 다양한 유형의 범용 또는 특수목적의 컴퓨팅 환경들이 사용될 수 있고 또는 여기 설명된 교시 내용에 따른 동작을 수행할 수 있다. 소프트웨어로 나타낸 상기 실시예들의 요소들은 하드웨어로 구현될 수 있고 그 반대의 경우도 가능하다.
- [0363] 본 발명의 원리가 적용될 수 있는 많은 가능한 실시예들의 관점에서, 다음의 청구범위 및 이의 균등물의 범위 및 사상 내에 포함될 수 있는 모든 이러한 실시예들을 나의 발명으로서 청구한다.
- 도면의 간단한 설명**
- [0020] 도 1은 본 발명의 바람직한 실시예들이 구현될 수 있는 적절한 컴퓨팅 환경의 블록도이다.
- [0021] 도 2는 본 발명의 바람직한 실시예들이 구현될 수 있는 일반화된 비디오 인코더 시스템의 블록도이다.
- [0022] 도 3은 본 발명의 바람직한 실시예들이 구현될 수 있는 일반화된 비디오 디코더 시스템의 블록도이다.
- [0023] 도 4는 본 발명의 바람직한 실시예들에서 사용되는 매크로블록 형식의 도면이다.
- [0024] 도 5A는 비월(interlaced) 비디오 프레임의 일부를 나타내는 도면으로, 상위 필드 및 하위 필드의 교번하는 라인을 도시하고, 도 5B는 프레임으로서 인코딩/디코딩을 위해 구성된 비월 비디오 프레임을 나타내는 도면이고, 도 5C는 필드로서 인코딩/디코딩을 위해 구성된 비월 비디오 프레임을 나타내는 도면이다.
- [0025] 도 5D는 비디오 프레임의 각 필드에 대해서 루마 샘플 위치에 대한 4:2:0 크로마 샘플 위치의 6가지 예의 공간 정렬을 도시한다.
- [0026] 도 6은 비디오의 다중 해상도 인코딩에 대한 일반화된 기술을 도시하는 흐름도이다.
- [0027] 도 7은 비디오의 다중 해상도 디코딩에 대한 일반화된 기술을 도시하는 흐름도이다.
- [0028] 도 8은 인트라-픽처 및 인터-픽처 예측 픽처들의 다중 해상도 인코딩 기술을 나타내는 흐름도이다.
- [0029] 도 9는 인트라-픽처 및 인터-픽처 예측 픽처들의 다중 해상도 디코딩 기술을 나타내는 흐름도이다.
- [0030] 도 10은 서로 다른 해상도로 비디오 디코딩을 가능하게 하기 위해 공간적으로 크기조정가능한 비트스트림 계층을 인코딩하는 기술을 도시하는 흐름도이다.
- [0031] 도 11은 서로 다른 해상도로 비디오 디코딩을 가능하게 하기 위해 공간적으로 크기조정가능한 비트스트림 계층을 디코딩하는 기술을 도시하는 흐름도이다.

[0032] 도 12 및 도 13은 일예의 다중 단계 위치 계산 기술에 대한 의사 코드(psuedo-code)를 도시하는 코드 도면이다.

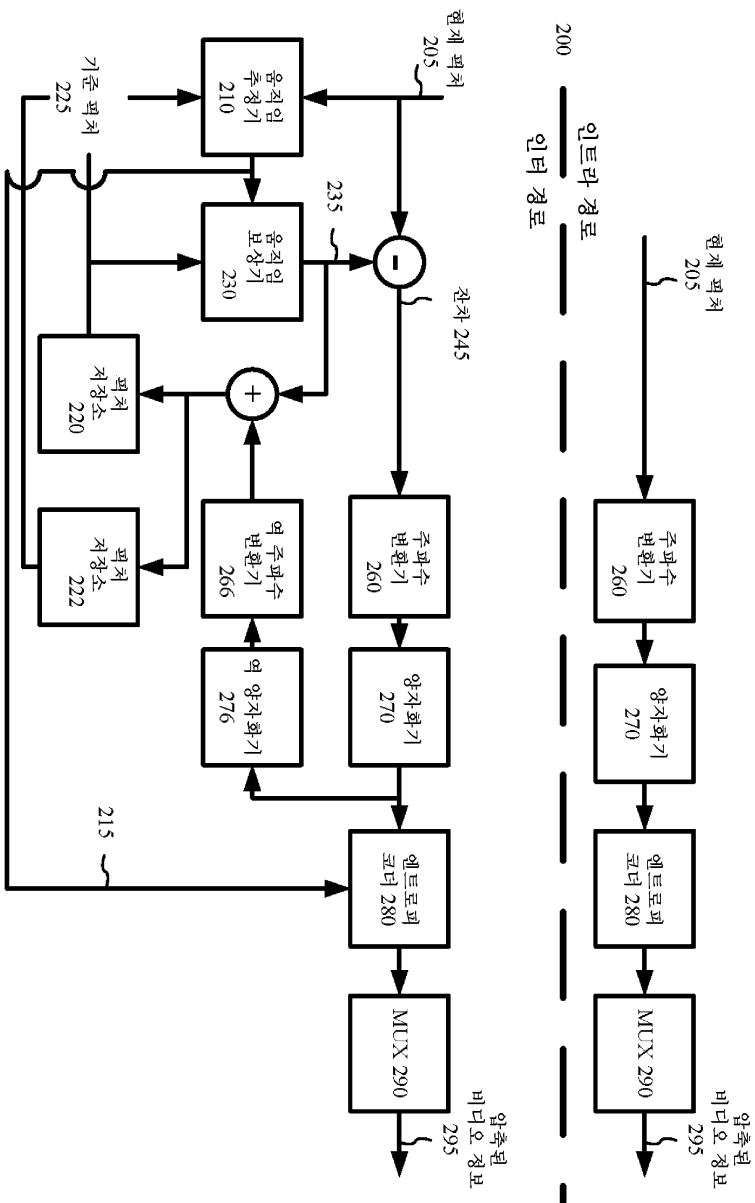
[0033] 도 14는 일예의 증분 위치 계산 기술에 대한 의사 코드를 도시하는 코드 도면이다.

도면

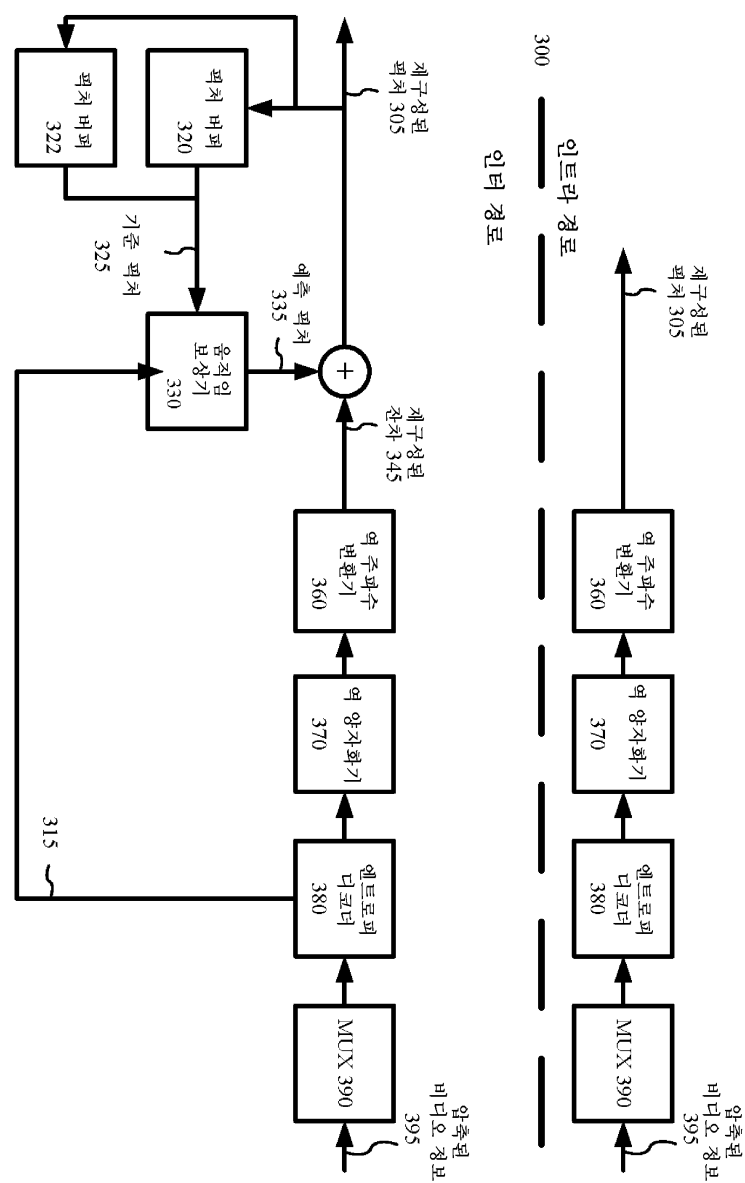
도면1



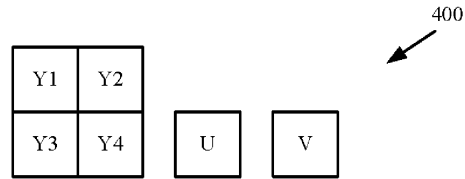
도면2



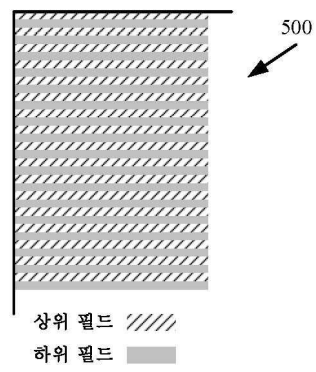
도면3



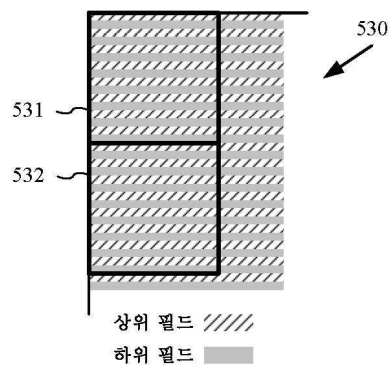
도면4



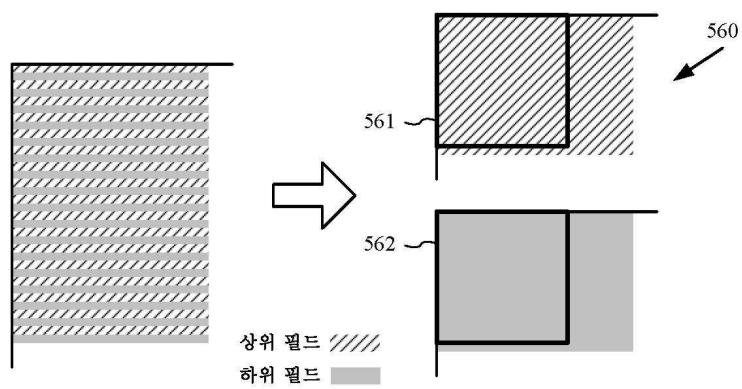
도면5A



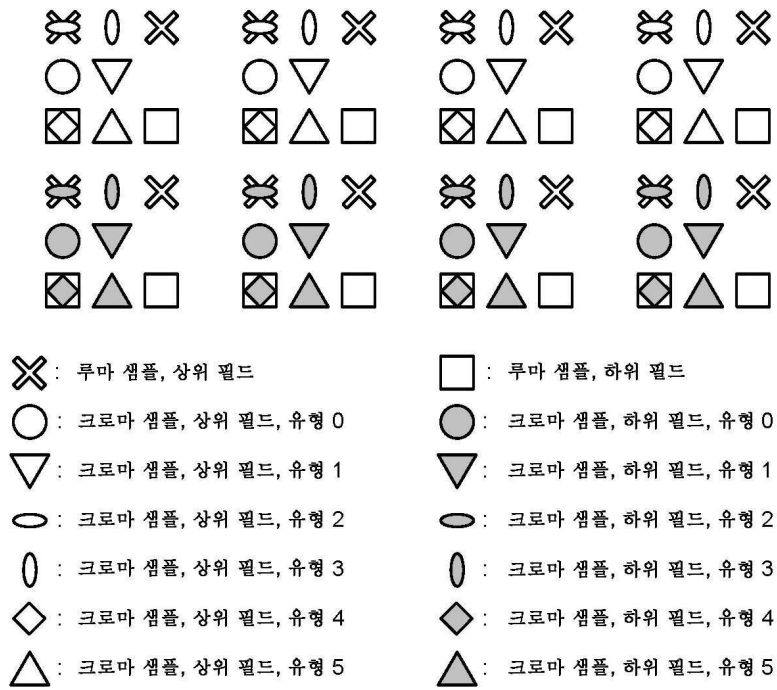
도면5B



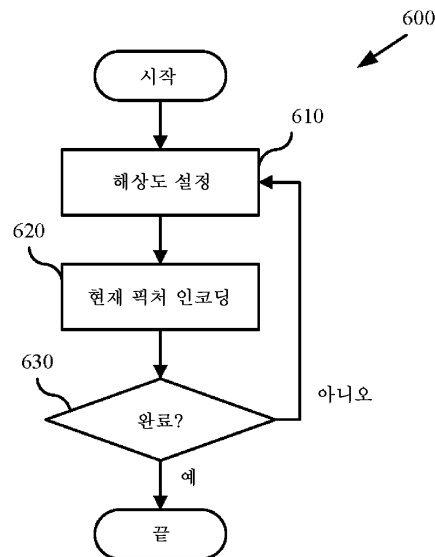
도면5C



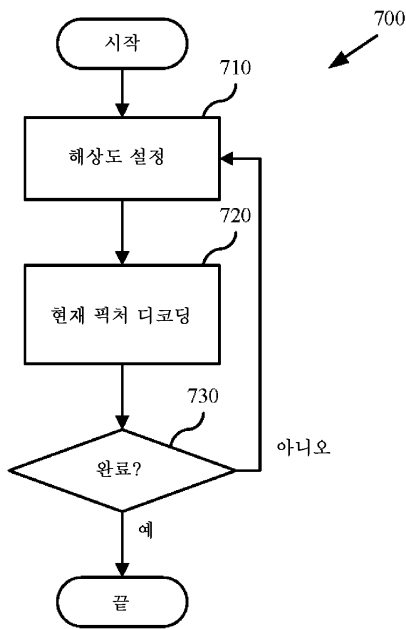
도면5D



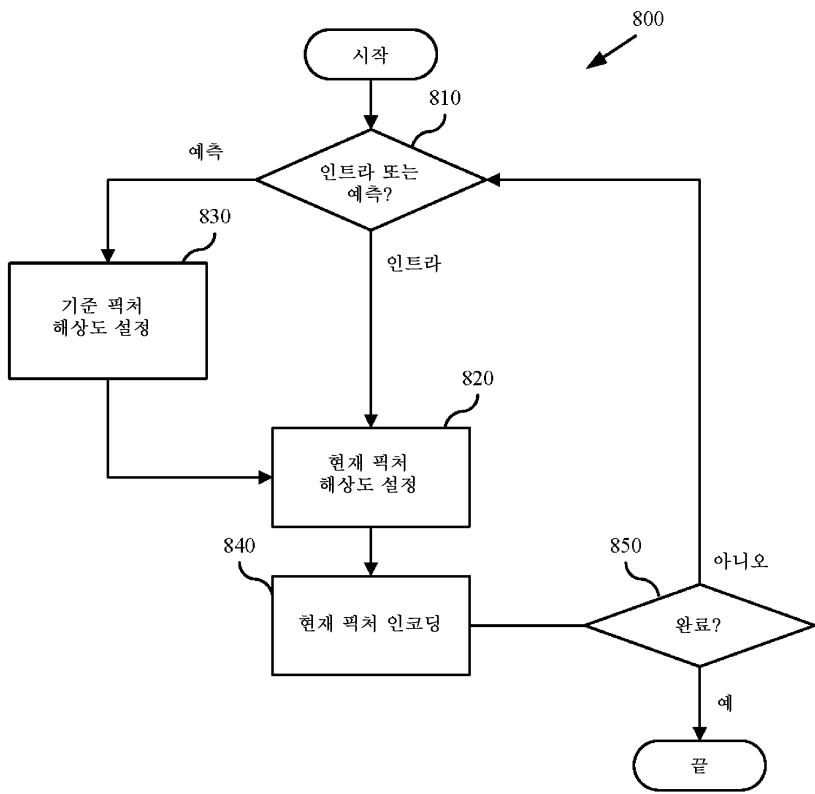
도면6



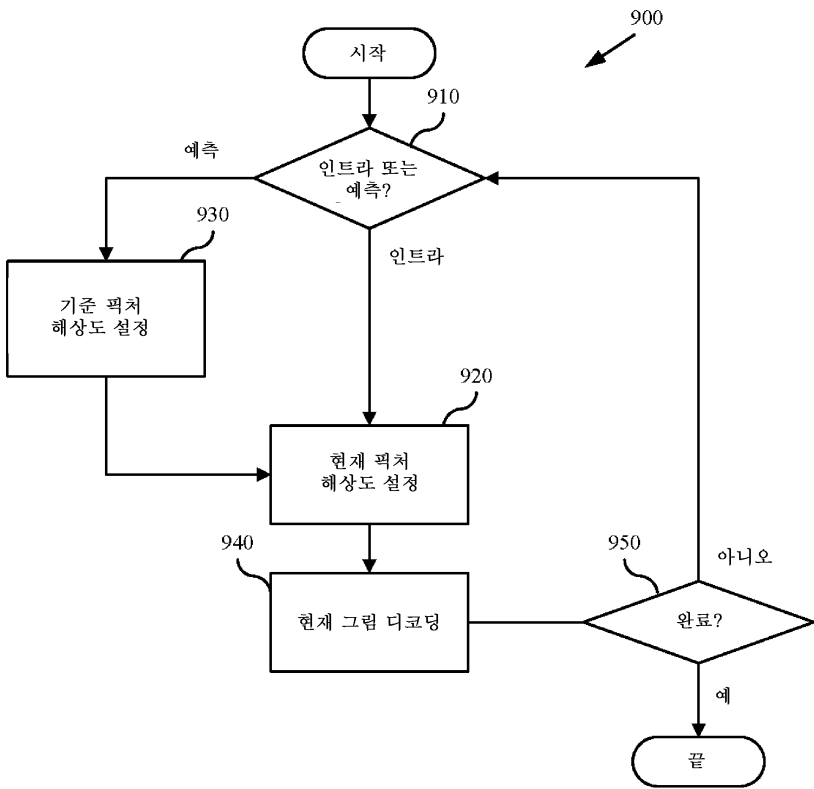
도면7



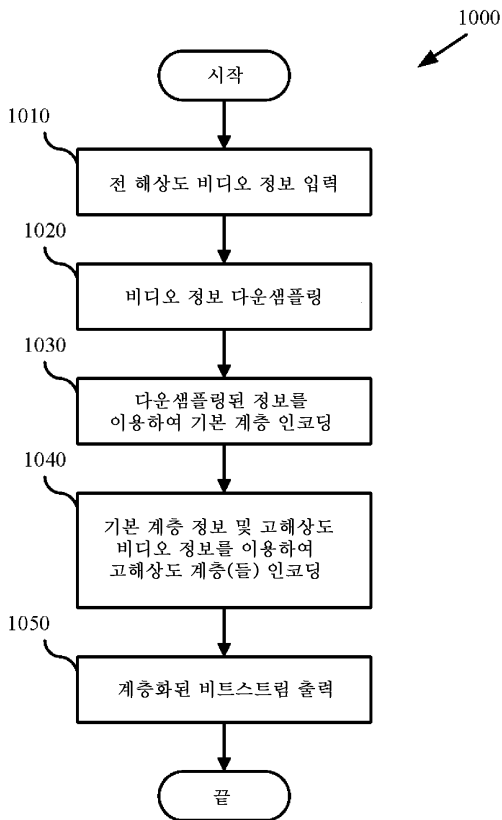
도면8



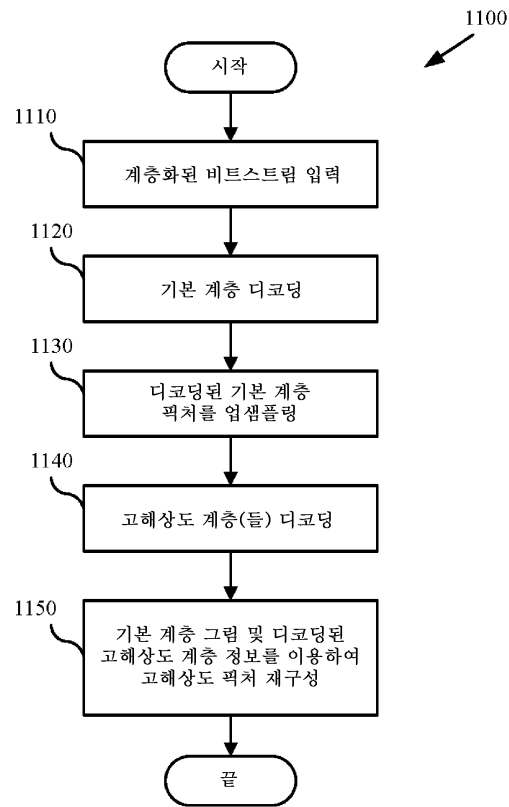
도면9



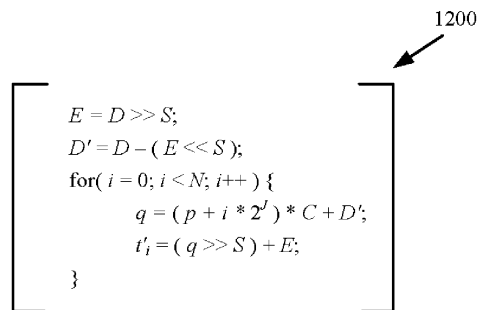
도면10



도면11



도면12



도면13

1300
↙

```

C' = C << J;
q = p * C + D;
/* First Stage */
w = q >> S;
E0 = w;
D0 = q - ( w << S );
t0 = E0
for( i = 1; i < N0; i++ ) {
    q = i * C' + D0;
    t'i = ( q >> S ) + E0;
}
/* Second Stage */
q = N0 * C' + D0;
w = q >> S;
E1 = E0 + w;
D1 = q - ( w << S );
tN0 = E1;
for( i = 1; i < N - N0; i++ ) {
    q = i * C' + D1;
    tN0+i = ( q >> S ) + E1;
}
    
```

도면14

1400
↙

```

C' = C << J;
E = 0;

q = p * C + D;
w = q >> S;
D' = q - ( w << S );
E += w;
t0 = E;

for( i = 1; i < N; i++ ) {
    q = C' + D'
    w = q >> S;
    D' = q - ( w << S );
    E += w;
    t'i = E;
}
    
```