



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2011년03월04일
(11) 등록번호 10-1019010
(24) 등록일자 2011년02월24일

(51) Int. Cl.
HO4N 7/26 (2006.01) HO4N 7/50 (2006.01)
HO4N 7/01 (2006.01)
(21) 출원번호 10-2008-7026885
(22) 출원일자(국제출원일자) 2007년03월13일
심사청구일자 2008년11월03일
(85) 번역문제출일자 2008년11월03일
(65) 공개번호 10-2009-0006159
(43) 공개일자 2009년01월14일
(86) 국제출원번호 PCT/US2007/063929
(87) 국제공개번호 WO 2007/114995
국제공개일자 2007년10월11일
(30) 우선권주장
60/789,048 2006년04월03일 미국(US)
(뒷면에 계속)
(56) 선행기술조사문헌
KR1020030029507 A*
KR1020060011281 A*
US5864369 A
WO200169936 A1
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
켈컴 인코포레이티드
미국 92121-1714 캘리포니아주 샌 디에고 모어하우스 드라이브 5775
(72) 발명자
텐 타오
미국 92126 캘리포니아주 샌디에고 뉴 살렘 스트리트 8680 넘버145
류 팡
미국 92122 캘리포니아주 샌디에고 코스타 베르데 불러바드 8775 넘버1207
(뒷면에 계속)
(74) 대리인
특허법인코리아나

전체 청구항 수 : 총 47 항

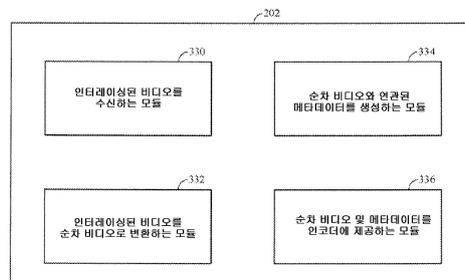
심사관 : 조우연

(54) 전처리 방법 및 장치

(57) 요약

본 발명은 일반적으로 멀티미디어 데이터 처리에 관한 것으로, 더욱 상세하게는, 데이터 압축 처리 전에 또는 데이터 압축 처리와 함께 수행되는 처리 동작에 관한 것이다. 멀티미디어 데이터를 처리하는 방법은 인터레이싱된 비디오 프레임을 수신하는 단계, 인터레이싱된 비디오 프레임에 대해 메타데이터를 획득하는 단계, 메타데이터의 적어도 일부를 이용하여 인터레이싱된 비디오 프레임을 순차 비디오로 변환하는 단계, 및 순차 비디오 및 메타데이터의 적어도 일부를 순차 비디오를 인코딩하는데 이용되는 인코더에 제공하는 단계를 포함한다. 또한, 본 발명의 방법은 인터레이싱된 비디오 프레임에 대한 공간 정보 및 양방향 모션 정보를 생성하는 단계, 및 인터레이싱된 비디오 프레임에 기초하고 공간 및 양방향 모션 정보를 이용하여 순차 비디오를 생성하는 단계를 포함한다.

대표도



(72) 발명자

스 팡

미국 92130 캘리포니아주 샌디에고 칼레 마르 데
아모니아 4460

라빈드란 비자알라크시미 알

미국 92130 캘리포니아주 샌디에고 칼레 마르 데
발레나스 4272

(30) 우선권주장

60/789,266 2006년04월04일 미국(US)

60/789,377 2006년04월04일 미국(US)

특허청구의 범위

청구항 1

멀티미디어 데이터를 처리하는 방법으로서,

인터레이싱된 비디오 프레임들을 수신하는 단계;

상기 인터레이싱된 비디오 프레임들을 순차 비디오로 변환하는 단계;

상기 순차 비디오와 연관된 메타데이터를 생성하는 단계; 및

상기 순차 비디오 및 상기 메타데이터의 적어도 일부를 상기 순차 비디오를 인코딩하는데 이용되는 인코더에 제공하는 단계를 포함하고,

상기 순차 비디오와 연관된 메타데이터를 생성하는 단계는 샷 검출을 수행하고 상기 샷 검출에 기초하여 압축 정보를 생성하는 단계를 포함하는, 멀티미디어 데이터 처리 방법.

청구항 2

제 1 항에 있어서,

상기 메타데이터를 이용하여 상기 순차 비디오를 인코딩하는 단계를 더 포함하는, 멀티미디어 데이터 처리 방법.

청구항 3

제 1 항에 있어서,

상기 인터레이싱된 비디오 프레임들을 변환하는 단계는 상기 인터레이싱된 비디오 프레임들을 디인터레이싱하는 단계를 포함하는, 멀티미디어 데이터 처리 방법.

청구항 4

제 1 항에 있어서,

상기 메타데이터는 대역폭 정보를 포함하는, 멀티미디어 데이터 처리 방법.

청구항 5

제 1 항에 있어서,

상기 메타데이터는 양방향 모션 정보를 포함하는, 멀티미디어 데이터 처리 방법.

청구항 6

제 3 항에 있어서,

상기 디인터레이싱하는 단계는,

상기 인터레이싱된 비디오 프레임들에 대한 공간 정보 및 양방향 모션 정보를 생성하는 단계; 및

상기 인터레이싱된 비디오 프레임들에 기초하고 상기 공간 정보 및 양방향 모션 정보를 이용하여 상기 순차 비디오를 생성하는 단계를 포함하는, 멀티미디어 데이터 처리 방법.

청구항 7

제 4 항에 있어서,

상기 대역폭 정보는 휘도 정보를 포함하는, 멀티미디어 데이터 처리 방법.

청구항 8

제 1 항에 있어서,

상기 메타데이터는 공간 복잡도 값을 포함하는, 멀티미디어 데이터 처리 방법.

청구항 9

제 1 항에 있어서,

상기 메타데이터는 시간 복잡도 값을 포함하는, 멀티미디어 데이터 처리 방법.

청구항 10

제 1 항에 있어서,

상기 인터레이싱된 비디오 프레임들을 변환하는 단계는 3/2 폴다운 비디오 프레임들을 역 텔레시네하는 단계를 포함하는, 멀티미디어 데이터 처리 방법.

청구항 11

제 10 항에 있어서,

상기 메타데이터는 대역폭 비율 정보를 포함하는, 멀티미디어 데이터 처리 방법.

청구항 12

제 1 항에 있어서,

상기 순차 비디오를 리사이징하는 단계를 더 포함하는, 멀티미디어 데이터 처리 방법.

청구항 13

제 12 항에 있어서,

영상 그룹 (group of picture) 정보를 결정하기 위해 상기 순차 비디오를 분할하는 단계를 더 포함하는, 멀티미디어 데이터 처리 방법.

청구항 14

삭제

청구항 15

제 1 항에 있어서,

상기 순차 비디오를 잡음 제거 (denoising) 필터로 필터링하는 단계를 더 포함하는, 멀티미디어 데이터 처리 방법.

청구항 16

제 1 항에 있어서,

상기 메타데이터는 휘도 및 크로마 정보를 포함하는, 멀티미디어 데이터 처리 방법.

청구항 17

멀티미디어 데이터를 처리하는 장치로서,

인터레이싱된 비디오 프레임들을 수신하도록 구성된 수신기;

상기 인터레이싱된 비디오 프레임들을 순차 비디오로 변환하도록 구성된 디인터레이서; 및

상기 순차 비디오와 연관된 메타데이터를 생성하고, 상기 순차 비디오 및 상기 메타데이터를 상기 순차 비디오를 인코딩하는데 이용되는 인코더에 제공하도록 구성된 분할기를 포함하고,

상기 분할기는 샷 검출을 수행하고 상기 샷 검출에 기초하여 압축 정보를 생성하도록 구성되는, 멀티미디어 데

이터 처리 장치.

청구항 18

제 17 항에 있어서,

상기 순차 비디오를 통신 모듈로부터 수신하고, 상기 제공된 메타데이터를 이용하여 상기 순차 비디오를 인코딩 하도록 구성된 인코더를 더 포함하는, 멀티미디어 데이터 처리 장치.

청구항 19

제 17 항에 있어서,

상기 디인테레이서는 공간-시간 디인테레이싱을 수행하도록 구성되는, 멀티미디어 데이터 처리 장치.

청구항 20

제 17 항에 있어서,

상기 순차 비디오를 잡음 제거하는 잡음 제거 필터를 더 포함하는, 멀티미디어 데이터 처리 장치.

청구항 21

제 17 항에 있어서,

상기 디인테레이서는 역 텔레시네 장치 (inverse teleciner) 를 포함하는, 멀티미디어 데이터 처리 장치.

청구항 22

삭제

청구항 23

제 17 항에 있어서,

상기 메타데이터는 영상 그룹 정보를 포함하는, 멀티미디어 데이터 처리 장치.

청구항 24

제 17 항에 있어서,

순차 프레임을 리사이징하도록 구성된 리샘플러 (resampler) 를 더 포함하는, 멀티미디어 데이터 처리 장치.

청구항 25

제 17 항에 있어서,

상기 메타데이터는 대역폭 정보를 포함하는, 멀티미디어 데이터 처리 장치.

청구항 26

제 17 항에 있어서,

상기 메타데이터는 양방향 모션 정보를 포함하는, 멀티미디어 데이터 처리 장치.

청구항 27

제 17 항에 있어서,

상기 디인테레이서는,

상기 인터레이싱된 비디오 프레임들에 대한 공간 정보 및 양방향 모션 정보를 생성하고,

상기 인터레이싱된 비디오 프레임들에 기초하고 상기 공간 정보 및 양방향 모션 정보를 이용하여 순차 비디오를 생성하도록 구성되는, 멀티미디어 데이터 처리 장치.

청구항 28

제 23 항에 있어서,
 상기 메타데이터는 대역폭 비율을 포함하는, 멀티미디어 데이터 처리 장치.

청구항 29

제 23 항에 있어서,
 상기 메타데이터는 휘도 정보를 포함하는, 멀티미디어 데이터 처리 장치.

청구항 30

제 17 항에 있어서,
 상기 메타데이터는 공간 복잡도 값을 포함하는, 멀티미디어 데이터 처리 장치.

청구항 31

제 17 항에 있어서,
 상기 메타데이터는 시간 복잡도 값을 포함하는, 멀티미디어 데이터 처리 장치.

청구항 32

제 17 항에 있어서,
 상기 메타데이터는 휘도 및 크로마 정보를 포함하는, 멀티미디어 데이터 처리 장치.

청구항 33

멀티미디어 데이터를 처리하는 장치로서,
 인터레이싱된 비디오를 수신하는 수단;
 상기 인터레이싱된 비디오를 순차 비디오로 변환하는 수단;
 상기 순차 비디오와 연관된 메타데이터를 생성하는 수단; 및
 상기 순차 비디오 및 상기 메타데이터의 적어도 일부를 상기 순차 비디오를 인코딩하는데 이용되는 인코더에 제공하는 수단을 포함하고,
 상기 생성 수단은 샷 검출을 수행하고 상기 샷 검출에 기초하여 압축 정보를 생성하도록 구성되는, 멀티미디어 데이터 처리 장치.

청구항 34

제 33 항에 있어서,
 상기 변환 수단은 역 텔레시네 장치를 포함하는, 멀티미디어 데이터 처리 장치.

청구항 35

제 33 항에 있어서,
 상기 변환 수단은 공간-시간 디인터레이서를 포함하는, 멀티미디어 데이터 처리 장치.

청구항 36

삭제

청구항 37

제 33 항에 있어서,

상기 생성 수단은 대역폭 정보를 생성하도록 구성되는, 멀티미디어 데이터 처리 장치.

청구항 38

제 33 항에 있어서,

순차 프레임을 리사이징하도록 리샘플링하는 수단을 더 포함하는, 멀티미디어 데이터 처리 장치.

청구항 39

제 33 항에 있어서,

상기 제공된 메타데이터를 이용하여 상기 순차 비디오를 인코딩하는 수단을 더 포함하는, 멀티미디어 데이터 처리 장치.

청구항 40

제 33 항에 있어서,

상기 순차 비디오를 잡음 제거하는 수단을 더 포함하는, 멀티미디어 데이터 처리 장치.

청구항 41

제 33 항에 있어서,

상기 메타데이터는 영상 그룹 정보를 포함하는, 멀티미디어 데이터 처리 장치.

청구항 42

제 33 항에 있어서,

상기 메타데이터는 양방향 모션 정보를 포함하는, 멀티미디어 데이터 처리 장치.

청구항 43

제 33 항에 있어서,

상기 변환 수단은,

상기 인터레이싱된 비디오 프레임들에 대한 공간 정보 및 양방향 모션 정보를 생성하고,

상기 인터레이싱된 비디오 프레임들에 기초하고 상기 공간 정보 및 양방향 모션 정보를 이용하여 순차 비디오를 생성하도록 구성되는, 멀티미디어 데이터 처리 장치.

청구항 44

제 33 항에 있어서,

상기 메타데이터는 대역폭 비율을 포함하는, 멀티미디어 데이터 처리 장치.

청구항 45

제 33 항에 있어서,

상기 메타데이터는 휘도 정보를 포함하는, 멀티미디어 데이터 처리 장치.

청구항 46

제 33 항에 있어서,

상기 메타데이터는 공간 복잡도 값을 포함하는, 멀티미디어 데이터 처리 장치.

청구항 47

제 33 항에 있어서,

상기 메타데이터는 시간 복잡도 값을 포함하는, 멀티미디어 데이터 처리 장치.

청구항 48

제 33 항에 있어서,

상기 메타데이터는 휘도 및 크로마 정보를 포함하는, 멀티미디어 데이터 처리 장치.

청구항 49

멀티미디어 데이터를 처리하는 명령들을 포함한 머신 판독가능 매체로서,

상기 멀티미디어 데이터를 처리하는 명령들은, 실행 시, 머신으로 하여금,

인터레이싱된 비디오 프레임들을 수신하게 하고,

상기 인터레이싱된 비디오 프레임들을 순차 비디오로 변환하게 하고,

상기 순차 비디오와 연관된 메타데이터를 생성하게 하고,

상기 순차 비디오 및 상기 메타데이터의 적어도 일부를 상기 순차 비디오를 인코딩하는데 이용되는 인코더에 제공하게 하고,

상기 순차 비디오와 연관된 메타데이터의 생성은 샷 검출을 수행하고 상기 샷 검출에 기초하여 압축 정보를 생성하는 것을 포함하는, 머신 판독가능 매체.

청구항 50

인터레이싱된 비디오를 수신하고,

상기 인터레이싱된 비디오를 순차 비디오로 변환하고,

상기 순차 비디오와 연관된 메타데이터를 생성하고,

상기 순차 비디오 및 상기 메타데이터의 적어도 일부를 상기 순차 비디오를 인코딩하는데 이용되는 인코더에 제공하는 구성을 포함하고,

상기 순차 비디오와 연관된 메타데이터의 생성은 샷 검출을 수행하고 상기 샷 검출에 기초하여 압축 정보를 생성하는 것을 포함하는, 프로세서.

명세서

[0001] **35 U.S.C. § 119 에 따른 우선권 주장**

[0002] 본 특허 출원은, 본원의 양수인에게 모두 양도되며 본원에서 명확히 참조로서 병합하고 있는, 2006 년 4 월 3 일 출원된 미국 가출원 제 60/789,048 호와, 2006 년 4 월 4 일 출원된 미국 가출원 제 60/789,266 호와, 2006 년 4 월 4 일 출원된 미국 가출원 제 60/789,377 호에 대한 우선권을 주장하고 있다.

[0003] **배경**

[0004] **분야**

[0005] 본 출원은 일반적으로 멀티미디어 데이터 처리에 관한 것으로, 더욱 상세하게는, 데이터 압축 처리 전에 또는 데이터 압축 처리와 함께 수행되는 처리 동작에 관한 것이다.

[0006] **배경**

[0007] **개요**

[0008] 본원에 설명되는 본 발명의 장치 및 방법의 각각은 몇몇 양태를 가지며, 그 양태들 중 어떤 단일의 양태도 그 바람직한 속성에 대해 단독으로 책임을 지는 것은 아니다. 이하, 본 발명의 범위를 제한함 없이, 본 발명의 보다 현저한 특징들을 간략하게 설명할 것이다. 본 설명을 고려한 다음에, 특히 "상세한 설명" 으로 지칭되는 부분을 읽은 다음에, 당업자라면 본 발명의 특징들이 멀티미디어 데이터 처리 장치 및 방법에 대한 개선을 어떻게 제공하는지를 이해할 것이다.

- [0009] 일 양태에서, 멀티미디어 데이터를 처리하는 방법은 인터레이싱된 비디오 프레임들을 수신하는 단계, 인터레이싱된 비디오 프레임들을 순차 비디오로 변환하는 단계, 순차 비디오와 연관된 메타데이터를 생성하는 단계, 및 순차 비디오 및 메타데이터의 적어도 일부를 순차 비디오를 인코딩하는데 이용되는 인코더에 제공하는 단계를 포함한다. 그 방법은 메타데이터를 이용하여 순차 비디오를 인코딩하는 단계를 더 포함할 수 있다. 몇몇 양태에서, 인터레이싱된 비디오 프레임은 NTSC 비디오를 포함한다. 비디오 프레임을 변환하는 단계는 인터레이싱된 비디오 프레임을 디인터레이싱하는 단계를 포함할 수 있다.
- [0010] 몇몇 양태에서, 메타데이터는 대역폭 정보, 양방향 모션 정보, 대역폭 비율, 시간 및/또는 공간 복잡도와 같은 복잡도 값, 휘도 정보를 포함할 수 있고, 공간 정보는 휘도 및/또는 크로미넌스 정보를 포함할 수 있다. 또한, 그 방법은 인터레이싱된 비디오 프레임에 대한 공간 정보 및 양방향 모션 정보를 생성하는 단계, 및 인터레이싱된 비디오 프레임에 기초하고 공간 및 양방향 모션 정보를 이용하여 순차 비디오를 생성하는 단계를 포함할 수 있다. 몇몇 양태에서, 인터레이싱된 비디오 프레임을 변환하는 단계는 3/2 풀다운 비디오 프레임을 역텔레시네하는 단계, 및/또는 순차 비디오를 리사이징하는 단계를 포함한다. 그 방법은 순차 비디오를 분할하여 영상 그룹 정보를 결정하는 단계를 더 포함할 수 있고, 분할 단계는 순차 비디오의 샷 검출을 포함할 수 있다. 또한, 몇몇 양태에서, 그 방법은 잡음 제거 필터로 순차 비디오를 필터링하는 단계를 포함한다.
- [0011] 다른 양태에서, 멀티미디어 데이터를 처리하는 장치는 인터레이싱된 비디오 프레임을 수신하도록 구성된 수신기, 인터레이싱된 비디오 프레임을 순차 비디오로 변환하도록 구성된 디인터레이서, 및 순차 비디오와 연관된 메타데이터를 생성하고, 순차 비디오 및 메타데이터를 순차 비디오를 인코딩하는데 이용되는 인코더에 제공하도록 구성된 분할기를 포함할 수 있다. 몇몇 양태에서, 그 장치는 통신 모듈로부터 순차 비디오를 수신하고, 제공된 메타데이터를 이용하여 순차 비디오를 인코딩하도록 구성된 인코더를 더 포함할 수 있다. 디인터레이서는 공간-시간 디인터레이싱 및/또는 역 텔레시네를 수행하도록 구성될 수 있다. 분할기는 샷 검출을 수행하고, 샷 검출에 기초하여 압축 정보를 생성하도록 구성될 수 있다. 몇몇 양태에서, 분할기는 대역폭 정보를 생성하도록 구성될 수 있다. 또한, 그 장치는 순차 프레임을 리사이징하도록 구성된 리샘플러를 포함할 수 있다. 메타데이터는 대역폭 정보, 양방향 모션 정보, 대역폭 비율, 휘도 정보, 콘텐츠에 관련된 공간 복잡도 값 및/또는 콘텐츠에 관련된 시간 복잡도 값을 포함할 수 있다. 몇몇 양태에서, 디인터레이서는 인터레이싱된 비디오 프레임에 기초하고 공간 및 양방향 모션 정보를 이용하여 인터레이싱된 비디오 프레임 및 순차 비디오에 대한 공간 정보 및 양방향 모션 정보를 생성하도록 구성된다.
- [0012] 다른 양태는 멀티미디어 데이터를 처리하는 장치를 포함하고, 멀티미디어 데이터 처리 장치는 인터레이싱된 비디오 프레임을 수신하는 수단, 인터레이싱된 비디오 프레임을 순차 비디오로 변환하는 수단, 순차 비디오와 연관된 메타데이터를 생성하는 수단, 및 순차 비디오 및 메타데이터 중 적어도 일부를 순차 비디오를 인코딩하는데 이용되는 인코더에 제공하는 수단을 포함한다. 몇몇 양태에서, 변환 수단은 역 텔레시네 장치 및/또는 공간-시간 디인터레이서를 포함한다. 몇몇 양태에서, 생성 수단은 샷 검출을 수행하고, 샷 검출에 기초하여 압축 정보를 생성하도록 구성된다. 몇몇 양태에서, 생성 수단은 대역폭 정보를 생성하도록 구성된다. 몇몇 양태에서, 생성 수단은 순차 프레임을 리사이징하도록 리샘플링하는 수단을 포함한다.
- [0013] 다른 양태는, 실행 시, 머신으로 하여금, 인터레이싱된 비디오 프레임을 수신하게 하고, 인터레이싱된 비디오 프레임을 순차 비디오로 변환하게 하고, 순차 비디오와 연관된 메타데이터를 생성하게 하고, 순차 비디오 및 메타데이터의 적어도 일부를 순차 비디오를 인코딩하는데 이용되는 인코더로 제공하게 하는, 멀티미디어 데이터를 처리하는 명령들을 포함한 머신 판독가능 매체를 포함한다.
- [0014] 다른 양태는, 인터레이싱된 비디오를 수신하고, 인터레이싱된 비디오를 순차 비디오로 변환하고, 순차 비디오와 연관된 메타데이터를 생성하고, 순차 비디오 및 메타데이터의 적어도 일부를 순차 비디오를 인코딩하는데 이용되는 인코더에 제공하는 구성을 포함한 프로세서를 포함한다. 인터레이싱된 비디오의 변환은 공간-시간 디인터레이싱을 수행하는 것을 포함할 수 있다. 몇몇 양태에서, 인터레이싱된 비디오의 변환은 역 텔레시네를 수행하는 것을 포함한다. 몇몇 양태에서, 메타데이터의 생성은 샷 변화의 검출에 기초하여 압축 정보를 생성하는 것을 포함한다. 몇몇 양태에서, 메타데이터의 생성은 순차 비디오의 압축 정보를 결정하는 것을 포함한다. 몇몇 양태에서, 상기 구성은 비디오를 리샘플링하여 리사이징된 순차 프레임을 생성하는 구성을 포함한다. 몇몇 양태에서, 메타데이터는 대역폭 정보, 양방향 모션 정보, 콘텐츠에 기초한 시간 또는 공간 복잡도 정보와 같은 복잡도 정보, 및/또는 압축 정보를 포함할 수 있다.

[0015] **도면의 간단한 설명**

- [0016] 도 1 은 스트리밍 멀티미디어 데이터를 전달하는 통신 시스템의 블록도이다.
- [0017] 도 2 는 전처리기를 포함하는 디지털 송신 설비의 블록도이다.
- [0018] 도 3a 는 전처리기의 예시적인 양태의 블록도이다.
- [0019] 도 3b 는 멀티미디어 데이터를 처리하는 프로세스를 도시한 흐름도이다.
- [0020] 도 3c 는 멀티미디어 데이터를 처리하는 수단을 도시한 블록도이다.
- [0021] 도 4 는 예시적인 전처리기의 동작을 도시한 블록도이다.
- [0022] 도 5 는 역 텔레시네 프로세스에 있어서 위상 결정의 도면이다.
- [0023] 도 6 은 텔레시네된 비디오를 역 텔레시네하는 프로세스를 도시한 흐름도이다.
- [0024] 도 7 은 위상 전이를 도시한 트래젝스 도면이다.
- [0025] 도 8 은 복수의 행렬을 생성하는데 이용되는 각각의 프레임을 식별하기 위한 가이드이다.
- [0026] 도 9 는 도 8 의 매트릭이 생성되는 방법을 도시한 흐름도이다.
- [0027] 도 10 은 매트릭을 처리하여 추정된 위상을 얻는 것을 도시한 흐름도이다.
- [0028] 도 11 은 결정 변수를 생성하는 시스템을 도시한 데이터 흐름도이다.
- [0029] 도 12 는 분기 정보를 평가하는데 이용되는 변수를 도시한 블록도이다.
- [0030] 도 13a, 도 13b 및 도 13c 는 하부 포락선이 계산되는 방법을 도시한 흐름도이다.
- [0031] 도 14 는 일치성 검출기의 동작을 도시한 흐름도이다.
- [0032] 도 15 는 위상 결정의 불일치를 보상하는데 이용되는 결정 변수에 대한 오프셋을 계산하는 프로세스를 도시한 흐름도이다.
- [0033] 도 16 은 폴 다운 위상이 추정된 후의 역 텔레시네 동작을 표현한다.
- [0034] 도 17 은 디인터레이서 디바이스의 블록도이다.
- [0035] 도 18 은 다른 디인터레이서 디바이스의 블록도이다.
- [0036] 도 19 는 인터레이싱된 영상의 서브샘플링 패턴의 도면이다.
- [0037] 도 20 은 W_{med} 필터링 모션 추정을 이용하여 디인터레이싱된 프레임을 생성하는 디인터레이서 디바이스의 블록도이다.
- [0038] 도 21 은 멀티미디어 데이터의 정적 영역을 결정하는 애퍼처 (aperture) 의 일 양태를 도시한다.
- [0039] 도 22 는 멀티미디어 데이터의 슬로우 모션 영역을 결정하는 애퍼처의 일 양태를 도시하는 도면이다.
- [0040] 도 23 은 모션 추정의 일 양태를 도시하는 도면이다.
- [0041] 도 24 는 모션 보상 결정에 이용되는 2 개의 모션 벡터 맵을 도시한다.
- [0042] 도 25 는 멀티미디어 데이터를 디인터레이싱하는 방법을 도시하는 흐름도이다.
- [0043] 도 26 은 공간-시간 정보를 이용하여 디인터레이싱된 프레임을 생성하는 방법을 도시한 흐름도이다.
- [0044] 도 27 은 디인터레이싱을 위해 모션 보상을 수행하는 방법을 도시한 흐름도이다.
- [0045] 도 28 은 몇몇 양태에 따라 샷 검출 및 다른 전처리 동작을 위해 구성된 프로세서를 포함한 전처리기의 블록도이다.
- [0046] 도 29 는 인코딩 복잡도 C 와 할당된 비트 B 간의 관계를 도시한다.
- [0047] 도 30 은 영상 그룹을 처리하며, 몇몇 양태에서 이용되어 비디오 프레임에서의 샷 검출에 기초하여 비디오를 인코딩할 수 있는 프로세스를 도시한 흐름도이다.
- [0048] 도 31 은 샷 검출에 대한 프로세스를 도시한 흐름도이다.

- [0049] 도 32 는 비디오의 상이한 샷 분류를 결정하는 프로세스를 도시한 흐름도이다.
- [0050] 도 33 은 샷 검출 결과에 기초하여 비디오 프레임에 프레임 압축 방식을 할당하는 프로세스를 도시한 흐름도이다.
- [0051] 도 34 는 갑작스런 장면 전환을 결정하는 프로세스를 도시한 흐름도이다.
- [0052] 도 35 는 느리게 변하는 장면을 결정하는 프로세스를 도시한 흐름도이다.
- [0053] 도 36 은 카메라 플래시라이트를 포함하는 장면을 결정하는 프로세스를 도시한 흐름도이다.
- [0054] 도 37 은 이전 프레임 MV_p 와 현재 프레임과 다음 프레임 MV_N 간의 모션 보상 벡터를 도시한다.
- [0055] 도 38 은 프레임 차이 메트릭을 결정하는데 이용되는 변수에 대한 관계를 도시한 그래프이다.
- [0056] 도 39 는 데이터 인코딩 및 레지듀얼 (residual) 계산을 도시한 블록도이다.
- [0057] 도 40 은 프레임 차이 메트릭의 결정을 도시한 블록도이다.
- [0058] 도 41 은 압축 타입을 프레임에 할당하는 절차를 도시한 흐름도이다.
- [0059] 도 42 는 1-D 다상 (poly-phase) 리샘플링의 일 예를 도시한다.
- [0060] 도 43 은 데이터 프레임의 세이프 액션 영역 및 세이프 타이틀 영역을 도시한 그래픽이다.
- [0061] 도 44 는 데이터 프레임의 세이프 액션 영역을 도시한 그래픽이다.

상세한 설명

- [0063] 다음 설명은 본 발명의 예들의 완전한 이해를 제공하기 위한 상세한 설명을 제공한다. 그러나, 당업자라면, 일 예 또는 양태의 프로세스 또는 디바이스의 모든 상세가 본원에 설명되거나 도시되어 있지는 않더라도 본 발명의 예들을 실시할 수도 있다는 것을 알 수 있다. 예를 들어, 전기 컴포넌트는, 본 발명의 예들을 지나치게 상세히 설명하여 불명확하게 되지 않도록 전기 컴포넌트의 모든 전기 접속 또는 모든 전기 소자를 도시하지는 않는 블록도로 도시될 수도 있다. 다른 경우에는, 그러한 컴포넌트, 다른 구조 및 기술을 상세히 도시하여, 본 발명의 예들을 더 설명할 수도 있다.
- [0064] 본원에는, 기존 전처리 및 인코딩 시스템의 성능을 개선하는 전처리기 및 전처리기 동작 방법에 대한 어떤 발명의 양태 및 양태들이 설명되어 있다. 그러한 전처리기는 디인터레이싱, 역 텔레시네, 필터링, 샷 타입의 식별, 메타데이터의 처리와 생성, 및 대역폭 정보의 생성을 수행하는 것을 비롯한 인코딩 준비에 있어서 메타데이터 및 비디오를 처리할 수 있다. 본원에서, "한 양태", "일 양태", "어떤 양태", 또는 "일정 양태" 에 대한 참조는, 그 양태와 관련하여 설명된 특정 피쳐, 구조, 또는 특징 중 하나 이상이 전처리기 시스템의 적어도 하나의 양태에 포함될 수 있음을 의미한다. 본 명세서의 여러 곳에서 그러한 문구의 사용은 반드시 모두가 동일한 양태를 지칭하는 것은 아니며, 기타 양태들을 서로 제외한 별개의 양태 또는 다른 양태를 지칭하는 것도 아니다. 또한, 어떤 양태에서는 나타날 수도 있지만 다른 양태에서는 나타나지 않을 수도 있는 여러 피쳐가 설명된다. 이와 유사하게, 어떤 양태를 위한 단계일 수도 있지만 다른 양태를 위한 단계는 아닐 수도 있는 여러 단계가 설명된다.
- [0065] 본원에서 이용되는 "멀티미디어 데이터" 또는 "멀티미디어" 는 비디오 데이터 (오디오 데이터를 포함할 수 있음), 오디오 데이터, 또는 비디오 데이터와 오디오 데이터 모두를 포함하는 광의의 용어이다. 본원에서 이용되는 "비디오 데이터" 또는 "비디오" 는 광의의 용어로서, 텍스트, 이미지, 및/또는 오디오 데이터를 포함한 하나 이상의 이미지 시리즈나 시퀀스 또는 하나의 이미지를 지칭하며, 달리 지정되지 않는 한, 멀티미디어 데이터를 지칭하는데 이용될 수 있고, 또는 서로 교환 가능하게 이용될 수도 있다.
- [0066] 도 1 은 스트리밍 멀티미디어를 전달하는 통신 시스템 (100) 의 블록도이다. 그러한 시스템은 도 1 에 도시된 것과 같은 다수의 단말기로 디지털 압축된 비디오를 송신하는데 적용된다. 디지털 비디오 소스는, 예를 들어, 디지털 케이블 또는 위성 피드 (satellite feed) 또는 디지털화되는 아날로그 소스일 수 있다. 비디오 소스는 송신 설비 (120) 에서 처리되고, 하나 이상의 단말기 (160) 로 네트워크 (140) 를 통한 송신을 위해 캐리어 상에서 인코딩 및 변조된다. 단말기 (160) 는 수신된 비디오를 디코딩하고, 통상 비디오의 적어도 일부를 표시한다. 네트워크 (140) 는 인코딩된 데이터를 송신하는데 적합한 임의 타입의 통신 네트워크, 유선 또는 무선을 지칭한다. 예를 들어, 네트워크 (140) 는 휴대 전화 네트워크, 유선 또는 무선 LAN (local

area network) 또는 WAN (wide area network), 또는 인터넷일 수 있다. 단말기 (160) 는 휴대 전화, PDA, 가정용이나 상업용 비디오 디스플레이 장비, 컴퓨터 (휴대용, 랩톱, 핸드헬드, PC, 및 대규모 서버 기반 컴퓨터 시스템), 및 멀티미디어 데이터를 이용할 수 있는 개인용 엔터테인먼트 디바이스를 비롯한, 데이터를 수신 및 표시할 수 있는 임의 타입의 통신 디바이스일 수 있지만, 이에 한정되는 것은 아니다.

[0067] 도 2 및 도 3 은 전처리기 (202) 의 예시적인 양태를 도시한다. 도 2 에서, 전처리기 (202) 는 디지털 송신 설비 (120) 내에 있다. 디코더 (201) 는 디지털 비디오 소스로부터 인코딩된 데이터를 디코딩하고, 메타데이터 (204) 및 비디오 (205) 를 전처리기 (202) 에 제공한다. 전처리기 (202) 는 비디오 (205) 및 메타데이터 (204) 에 대한 일정 타입의 처리를 수행하고, 처리된 메타데이터 (206; 예를 들어, 기저 계층 참조 프레임, 인헨스먼트 계층 참조 프레임, 대역폭 정보, 콘텐츠 정보) 및 비디오 (207) 를 인코더 (203) 에 제공한다. 그러한 멀티미디어 데이터의 전처리는 시각 명료도, 안티-에일리어싱, 및 데이터의 압축 효율을 개선할 수 있다. 일반적으로, 전처리기 (202) 는 디코더 (201) 에 의해 제공된 비디오 시퀀스를 수신하고, 비디오 시퀀스를 인코더에 의한 후속 처리 (예를 들어, 인코딩) 를 위해 순차 비디오 시퀀스로 변환한다. 몇몇 양태에서, 전처리기 (202) 는 역 텔레시네, 디인터레이싱, 필터링 (예를 들어, 아티팩트 제거, 디링잉, 디블록킹, 및 디노이징), 리사이징 (예를 들어, SD (standard definition) 로부터 QVGA (Quarter Video Graphics Array) 로의 공간 해상도 다운-샘플링), 및 GOP 구조 생성 (예를 들어, 복잡도 맵 생성, 장면 전환 검출, 및 페이드/플래시 검출의 계산) 을 비롯한, 다수의 동작을 위해 구성될 수 있다.

[0068] 도 3a 는 모듈 또는 컴포넌트 (본원에서 일괄하여 "모듈" 로 지칭됨) 와 함께 수신된 메타데이터 (204) 및 비디오 (205) 에 대한 전처리 동작을 수행한 다음에, 후속 처리를 위해 (예를 들어, 인코더로) 처리된 메타데이터 (206) 및 순차 비디오 (207) 를 제공하도록 구성되는 전처리기 (202) 를 도시한다. 그 모듈은 하드웨어, 소프트웨어, 펌웨어, 또는 그 조합으로 구현될 수 있다. 전처리기 (202) 는 도시되어 있는 하나 이상의 모듈을 포함하며, 모두가 아래에서 더 설명되는 역 텔레시네 (301), 디인터레이서 (302), 디노이저 (303), 에일리어스 제거기 (304), 리샘플러 (305), 디블록커/디링거 (306), 및 GOP 분할기 (307) 를 포함하는 여러 모듈을 포함할 수 있다. 또한, 전처리기 (202) 는 메모리 (308) 및 통신 모듈 (309) 을 비롯한, 비디오 및 메타데이터를 처리하는데 이용될 수도 있는 다른 적당한 모듈을 포함할 수 있다. 소프트웨어 모듈은 RAM 메모리, 플래시 메모리, ROM 메모리, EPROM 메모리, EEPROM 메모리, 레지스터, 하드 디스크, 착탈식 디스크, CD-ROM, 또는 당해 기술분야에서 공지되어 있는 임의의 다른 형태의 저장 매체에 상주할 수도 있다. 예시적인 저장 매체는 프로세서에 연결되어, 프로세서가 저장 매체로부터 정보를 판독하고, 저장 매체에 정보를 기록할 수 있도록 해준다. 다른 방법으로는, 저장 매체는 프로세서와 일체로 형성될 수도 있다. 프로세서와 저장 매체는 ASIC 에 상주할 수도 있다. ASIC 은 사용자 단말기에 상주할 수도 있다. 다른 방법으로는, 프로세서와 저장 매체는 사용자 단말기에 별도의 컴포넌트로서 상주할 수도 있다.

[0069] 도 3b 는 멀티미디어 데이터의 처리를 위한 프로세스 (300) 를 도시한 흐름도이다. 프로세스 (300) 가 시작하고, 블록 320 으로 진행하여, 인터레이싱된 비디오를 수신한다. 도 2 및 도 3 에 도시된 전처리기 (202) 는 이러한 단계를 수행할 수 있다. 몇몇 양태에서, 디코더 (예를 들어, 도 2 의 디코더 (201)) 는 인터레이싱된 데이터를 수신하고, 이 인터레이싱된 데이터를 전처리기 (202) 에 제공할 수 있다. 몇몇 양태에서, 도 3c 에 전처리기 (202) 의 일부로서 도시되어 있는 데이터 수신 모듈 (330) 은 이러한 단계를 수행할 수 있다. 그 다음에, 프로세스 (300) 는 블록 322 로 진행하여, 인터레이싱된 비디오를 순차 비디오로 변환한다. 도 2 및 도 3a 에 도시된 전처리기 (202) , 및 도 3c 의 모듈 (332) 은 이러한 단계를 수행할 수 있다. 인터레이싱된 비디오가 텔레시네되면, 블록 322 의 처리는 역 텔레시네를 수행하여 순차 비디오를 생성하는 것을 포함할 수 있다. 그 다음에, 프로세스 (300) 는 블록 324 로 진행하여, 순차 비디오와 연관된 메타데이터를 생성할 수 있다. 도 3a 의 GOP 분할기 (307) 및 도 3c 의 모듈 (334) 은 그러한 처리를 수행할 수 있다. 그 다음에, 프로세스 (300) 는 블록 326 으로 진행하여, 순차 비디오 및 메타데이터의 적어도 일부를 인코딩 (예를 들어, 압축) 을 위해 인코더에 제공한다. 도 2 및 도 3a 에 도시된 전처리기 (202), 및 도 3c 의 모듈 (336) 은 이러한 단계를 수행할 수 있다. 순차 비디오 및 연관된 메타데이터를 인코딩을 위해 다른 컴포넌트로 제공한 후에, 프로세스 (300) 가 종료할 수 있다.

[0070] 도 3c 는 멀티미디어 데이터를 처리하는 수단을 도시한 블록도이다. 본원에는 그러한 수단이 전처리기 (202) 에 포함되어 있는 것으로 도시되어 있다. 전처리기 (202) 는 모듈 (330) 과 같은 비디오를 수신하는 수단을 포함한다. 또한, 전처리기 (202) 는 모듈 (332) 과 같은 인터레이싱된 데이터를 순차 비디오로 변환하는 수단을 포함한다. 그러한 수단은, 예를 들어, 공간-시간 디인터레이서 및/또는 역 텔레시네 장치를 포함할 수 있다. 또한, 전처리기 (202) 는 모듈 (334) 과 같은 순차 비디오와 연관된 메타데이터를 생성하는

수단을 포함한다. 그러한 수단은, 상술한 것과 같이 여러 타입의 메타데이터를 생성할 수 있는 GOP 분할기 (307; 도 3a) 를 포함할 수 있다. 또한, 전처리 (202) 는 모듈 (336) 에 의해 도시된 것과 같이 순차 비디오 및 메타데이터를 인코딩을 위해 인코더에 제공하는 수단을 포함할 수 있다. 몇몇 양태에서, 그러한 수단은 도 3a 에 도시된 통신 모듈 (309) 을 포함할 수 있다. 당업자라면 알 수 있는 바와 같이, 그러한 수단은 다수의 표준적인 방법으로 구현될 수 있다.

[0071] 전처리 (202) 는 전처리 동작 중 하나 이상의 동작을 위해 획득한 메타데이터 (예를 들어, 디코더 (201) 또는 다른 소스로부터 획득함) 를 이용할 수 있다. 메타데이터는 멀티미디어 데이터의 콘텐츠 ("콘텐츠 정보") 를 설명하거나 분류하는 것과 관련된 정보를 포함할 수 있다. 특히, 메타데이터는 콘텐츠 분류를 포함할 수 있다. 몇몇 양태에서, 메타데이터는 인코딩 동작을 위해 요구된 콘텐츠 정보를 포함할 수 있다. 그러한 경우에, 전처리 (202) 는 콘텐츠 정보를 결정하고, 전처리 동작을 위해 콘텐츠 정보를 이용하고/하거나 예를 들어, 인코더 (203) 와 같은 다른 컴포넌트에 콘텐츠 정보를 제공하도록 구성될 수 있다. 몇몇 양태에서, 전처리 (202) 는 그러한 콘텐츠 정보를 이용하여, GOP 분할에 영향을 주고, 적절한 타입의 필터링을 결정하고/하거나 인코더로 통신되는 인코딩 파라미터를 결정할 수 있다.

[0072] 도 4 는 전처리에 포함될 수 있는 프로세스 블록의 예시적인 실시형태를 도시하는 것으로서, 전처리 (202) 에 의해 수행될 수 있는 처리를 도시한다. 본 예에서, 전처리 (202) 는 메타데이터 및 비디오 (204, 205) 를 수신하고, (처리된) 메타데이터 및 비디오를 포함한 출력 데이터 (206, 207) 를 인코더 (203) 에 제공한다. 통상, 전처리에 의해 수신되는 비디오에는 3 가지 타입이 있다. 첫째로, 수신된 비디오는 순차 비디오일 수 있고, 디인터레이싱은 수행되지 않아도 된다. 둘째로, 비디오 데이터는 텔레시네된 비디오일 수 있고, 인터레이싱된 비디오는 24fps 영화 시퀀스로부터 변환된다. 셋째로, 비디오는 텔레시네되지 않은 인터레이싱된 비디오일 수 있다. 전처리 (226) 는 아래에서 설명하는 바와 같이 이들 타입의 비디오를 처리할 수 있다.

[0073] 블록 401 에서, 전처리 (202) 는 수신된 비디오 (204, 205) 가 순차 비디오인지 여부를 결정한다. 몇몇 경우에, 이는 메타데이터가 그러한 정보를 포함하면 메타데이터로부터 결정될 수 있고, 또는 비디오 자체의 처리에 의해 결정될 수 있다. 예를 들어, 아래에서 설명되는 역 텔레시네 프로세스는, 수신된 비디오 (205) 가 순차 비디오인지 여부를 결정할 수 있다. 수신된 비디오가 순차 비디오인 경우에, 프로세스는 블록 407 로 진행하여, 백색 가우스 잡음과 같은 잡음을 줄이기 위해 비디오에 대해 필터링 동작을 수행한다. 비디오가 순차 비디오가 아닌 경우에, 블록 401 에서, 프로세스는 블록 404 의 위상 검출기로 진행한다.

[0074] 위상 검출기 (404) 는 텔레시네에서 생성된 비디오와 표준 방송 포맷에서 시작된 비디오를 구별한다. 비디오가 텔레시네되었다는 결정이 이루어지면 (위상 검출기 (404) 에서 나가는 "예" 결정 경로), 텔레시네된 비디오는 역 텔레시네 (406) 에서 그 원래 포맷으로 리턴된다. 중복 필드가 식별 및 제거되고, 동일 비디오로부터 유도된 필드가 완전한 이미지로 다시 만들어진다. 복원된 필름 이미지의 시퀀스가 1/24 초의 일정한 간격으로 사진 기록되므로, GOP 분할기 (412) 또는 디코더에서 수행되는 모션 추정 프로세스는 일정하지 않은 시간축을 갖는 텔레시네된 데이터보다 오히려 역 텔레시네된 이미지를 이용하여 더 정확해진다.

[0075] 일 양태에서, 위상 검출기 (404) 는 비디오 프레임의 수신 후에 어떤 결정을 행한다. 이들 결정은 (i) 현재 비디오가 텔레시네 출력으로부터의 비디오인지 및 3:2 폴 다운 위상이 도 5 에 도시된 5 개의 위상 P₀, P₁, P₂, P₃, 및 P₄ 중 하나인지 여부, 및 (ii) 비디오가 통상의 NTSC 로서 생성되었는지 여부를 포함한다. 그 결정은 위상 P₅ 로서 표시된다. 이들 결정은 도 4 에 도시된 위상 검출기 (404) 의 출력으로서 나타난다. "예" 라벨을 붙인 위상 검출기 (404) 로부터의 경로는 역 텔레시네 (406) 를 작동시켜, 동일한 사진 이미지로부터 형성된 필드를 분류하고 결합할 수 있도록 정확한 폴 다운 위상이 제공되었음을 나타낸다. 이와 유사하게, "아니오" 라벨을 붙인 위상 검출기 (404) 로부터의 경로는 디인터레이서 (405) 를 작동시켜, 외관상 NTSC 프레임을 최적 처리를 위한 필드로 분할한다. 또한, 역 텔레시네는 발명의 명칭이 "INVERSE TELECINE ALGORITHM BASED ON STATE MACHINE" 이고 본원의 양수인이 소유하고 있으며 본원에서 그 전체를 참조로서 병합하고 있는 동시 계속 중인 미국 특허 출원 [대리인 Docket No. QFDM.021A(050943)] 에 더 설명되어 있다.

[0076] 위상 검출기 (404) 는, 상이한 타입의 비디오가 어느 때라도 수신될 수도 있으므로 연속하여 비디오 프레임을 분석할 수 있다. 일 예로서, NTSC 표준에 따르는 비디오는 커머셜로서 비디오 내에 삽입될 수도 있다. 역 텔레시네 후에, 결과로서 생성된 순차 비디오는 백색 가우스 잡음을 줄이는데 이용될 수 있는 디노이저 (필터) (407) 로 송신된다.

- [0077] 통상의 NTSC 비디오가 인식되는 경우에는 (위상 검출기 (401)로부터의 "아니오" 경로), 압축을 위해 디인터레이서 (405)로 송신된다. 디인터레이서 (405)가 인터레이싱된 필드를 순차 비디오로한 다음에, 순차 비디오에 대해 디노이즈 동작이 수행될 수 있다.
- [0078] 적절한 역 텔레시네 또는 디인터레이싱 처리 후에, 블록 408에서, 순차 비디오는 에일리어스 제거 및 리샘플링 (예를 들어, 리사이징)을 위해 처리된다.
- [0079] 또한, 리샘플링 후에, 순차 비디오는 블록 410으로 진행하여, 디블록킹 및 디링잉 동작이 수행된다. 2가지 타입의 아티팩트인, "블록킹"과 "링잉"은 비디오 압축 애플리케이션에서 보통 발생한다. 압축 알고리즘이 각각의 프레임을 블록 (예를 들어, 8 × 8 블록)으로 분할하기 때문에, 블록킹 아티팩트가 발생한다. 각각의 블록은 얼마간의 작은 에러로 재구성되고, 블록의 에지에 있는 에러는 종종 이웃 블록의 에지에 있는 에러와 대비되어, 블록 경계를 볼 수 있도록 해준다. 한편, 링잉 아티팩트는 이미지 피처의 에지 주위에 있는 왜곡으로서 나타난다. 인코더가 고주파수 DCT 계수의 양자화 시에 너무 많은 정보를 버리므로, 링잉 아티팩트가 발생한다. 몇몇 예시적인 예에서, 디블록킹과 디링잉 모두는 저역통과 FIR (finite impulse response) 필터를 이용하여 이들 눈에 보이는 아티팩트를 숨길 수 있다.
- [0080] 디블록킹 및 디링잉 후에, 순차 비디오는 GOP 분할기 (412)에 의해 처리된다. GOP 포지셔닝은 샷 전환의 검출, 복잡도 맵 (예를 들어, 시간, 공간 대역폭 맵)의 생성, 및 적응 GOP 분할을 포함할 수 있다. 샷 검출은, GOP (group of picture; 영상 그룹) 내의 프레임이 장면 전환이 발생한 것을 표시하는 데이터를 나타내는 때를 결정하는 것에 관련된다. 고정된 간격으로 I-프레임을 삽입하는 대신에, 적절한 GOP 길이를 결정하고, GOP 길이에 기초하여 I-프레임을 삽입하도록, 장면 전환 검출을 비디오 인코더가 이용할 수 있다. 또한, 전처리 (202)는 멀티미디어 데이터를 인코딩하는데 이용될 수 있는 대역폭 맵을 생성하도록 구성될 수 있다. 그 대신에, 몇몇 양태에서는, 전처리기 외부에 위치한 콘텐츠 분류 모듈이 대역폭 맵을 생성한다. 적응 GOP 분할은 함께 코딩된 영상 그룹의 콤포지션 (composition)을 적응 변경할 수 있다. 도 4에 도시된 동작의 예시적인 실시형태는 아래에 설명된다.
- [0081] **역 텔레시네**
- [0082] 역 텔레시네 처리는 아래에서 설명되고, 역 텔레시네의 예시적인 실시형태는 도 4 내지 도 16을 참조하여 제공된다. 소스의 특성이 알려져 있으며 이상적으로 매칭하는 처리 형태를 선택하는데 이용되는 경우에, 비디오 압축이 최선의 결과를 제공한다. 예를 들어, 오프-디-에어 (off-the-air) 비디오는 몇몇 방법으로 발생할 수 있다. 비디오 카메라나 브로드캐스트 스튜디오 등에서 통상적인 방법으로 생성되는 브로드캐스트 비디오는 미국의 NTSC 표준을 따른다. NTSC 표준에 따라, 각각의 프레임은 2개의 필드로 이루어진다. 한 필드는 홀수 라인으로 이루어지고, 다른 필드는 짝수 라인으로 이루어진다. 이는 "인터레이싱된" 포맷으로 지칭될 수도 있다. 약 30 프레임/초로 프레임이 생성되는 동안에, 필드는 1/60 초 떨어져 있는 텔레비전 카메라 이미지의 레코드이다. 한편, 필름은 24 프레임/초로 촬영되어, 각각의 프레임은 완전한 이미지로 이루어진다. 이는 "순차" 포맷으로서 지칭될 수도 있다. NTSC 장비에서의 송신의 경우에, "순차" 비디오는 텔레시네 프로세스를 통해 "인터레이싱된" 비디오 포맷으로 변환된다. 일 양태에서, 아래에서 더 설명되는 바와 같이, 본 발명의 시스템은 비디오가 텔레시네되고 적절한 변환을 수행하여 원래 순차 프레임을 재생하는 때를 결정하는 이점을 갖는다.
- [0083] 도 5는 인터레이싱된 비디오로 변환된 순차 프레임을 텔레시네하는 효과를 도시한다. $F_1, F_2, F_3,$ 및 F_4 는 텔레시네 장치에 입력되는 순차 비디오이다. 아래의 숫자 "1" 및 "2"는 홀수 또는 짝수 필드 중 어느 하나의 표시이다. 프레임 레이트 간의 불일치를 고려하여 몇몇 필드가 반복된다는 것에 주목하자. 또한, 도 4는 풀-다운 위상 $P_0, P_1, P_2, P_3,$ 및 P_4 를 도시한다. 위상 P_0 은 동일한 제 1 필드를 갖는 2개의 NTSC 호환성 프레임 중 제 1 프레임에 의해 표시된다. 다음 4개의 프레임은 위상 $P_1, P_2, P_3,$ 및 P_4 에 대응한다. P_2 및 P_3 에 의해 표시된 프레임은 동일한 제 2 필드를 갖는다는 것에 주목하자. 필름 프레임 F_1 이 3회 스캔되므로, 2개의 동일한 연속적인 출력의 NTSC 호환성 제 1 필드가 형성된다. 필름 프레임 F_1 로부터 유도된 모든 NTSC 필드가 동일 필름 이미지로부터 선택되므로, 동시에 선택되게 된다. 필름으로부터 유도된 다른 NTSC 프레임은 1/24 초 떨어져 있는 인접한 필드를 가질 수도 있다.
- [0084] 도 4에 도시된 위상 검출기 (404)는 비디오 프레임의 수신 후에 어떤 결정을 행한다. 이들 결정은 (i) 현재 비디오가 텔레시네 출력으로부터의 비디오인지 및 3:2 풀 다운 위상이 도 5의 정의 (512)에 도시되어 있

는 5 개의 위상 P_0, P_1, P_2, P_3 , 및 P_4 중 하나인지 여부; 및 (ii) 비디오가 통상적인 NTSC 로서 생성되었는지 여부를 포함하고, 그 결정은 위상 P_5 로서 표시된다.

[0085] 이들 결정은 도 4 에 도시된 위상 검출기 (401) 의 출력으로서 나타난다. "예" 라벨을 붙인 위상 검출기 (401) 로부터의 경로는 역 텔레시네 (406) 를 작동시켜, 동일 사진 이미지로부터 형성된 필드를 분류하고 결합할 수 있도록 정확한 폴 다운 위상이 제공되는 것을 나타낸다. 이와 유사하게, "아니오" 라벨을 붙인 위상 검출기 (401) 로부터의 경로는 디인터레이서 블록 (405) 을 작동시켜, 외관상 NTSC 프레임에 최적 처리를 위한 필드로 분할한다.

[0086] 도 6 은 비디오 스트림을 역 텔레시네하는 프로세스 (600) 를 도시한 흐름도이다. 일 양태에서, 프로세스 (600) 는 도 3 의 역 텔레시네 (301) 에 의해 수행된다. 단계 651 에서 시작하여, 역 텔레시네 (301) 는 수신된 비디오에 기초하여 복수의 메트릭을 결정한다. 본 양태에서는, 동일 프레임 또는 인접 프레임으로부터 유도된 필드 간의 차이의 합인 4 개의 메트릭이 형성된다. 또한, 4 개의 메트릭은 수신된 데이터로부터 유도된 4 개의 메트릭과 6 개의 가설화된 위상 각각에 대한 이들 메트릭의 가장 가능성 있는 값 간의 거리의 유클리드 측정으로 어셈블링된다. 유클리드 합은 분기 정보로 지칭되고, 각각의 수신된 프레임에 대해 6 개의 그러한 양이 존재한다. 각각의 가설화된 위상은 뒤에 오는 위상을 가지며, 이러한 뒤에 오는 위상은 가능한 폴 다운 위상의 경우에 각각의 수신된 프레임에 따라 변한다.

[0087] 가능한 전이 경로는 도 7 에 도시되어 있고, 참조 부호 767 로 표시되어 있다. 그러한 경로가 6 개 존재한다. 결정 프로세스는 가설화된 위상의 각 경로에 대해 유클리드 거리의 합과 등가인 4 개의 측정을 유지한다. 변경된 조건에 절차가 응답하도록, 그 합에 있어서 각각의 유클리드 거리는 오래됨에 따라 줄어든다. 유클리드 거리의 합이 최소인 위상 트랙이 동작하는 위상 트랙인 것으로 생각된다. 이러한 트랙의 현재 위상은 "적용 가능한 위상" 으로 지칭된다. 이때, P_5 가 아닌 한, 선택된 위상에 기초한 역 텔레시네가 발생할 수 있다. P_5 가 선택되면, 블록 405 (도 4) 에서, 현재 프레임은 디인터레이서를 이용하여 디인터레이싱된다. 요약하면, 적용 가능한 위상은 현재 폴 다운 위상으로서 이용되거나, 유효한 NTSC 포맷을 갖는 것으로 추정된 프레임의 디인터레이싱을 명령하는 표시자로서 이용된다.

[0088] 비디오 입력으로부터 수신된 모든 프레임에 대해, 4 개의 메트릭 각각에 대한 새로운 값이 계산된다. 이들은 다음과 같이 정의된다.

$$SAD_{FS} = \sum | \text{Current Field One Value}(i,j) - \text{Previous Field One Value}(i,j) | \quad (1)$$

$$SAD_{SS} = \sum | \text{Current Field Two Value}(i,j) - \text{Previous Field Two Value}(i,j) | \quad (2)$$

$$SAD_{PO} = \sum | \text{Current Field One Value}(i,j) - \text{Previous Field Two Value}(i,j) | \quad (3)$$

$$SAD_{CO} = \sum | \text{Current Field One Value}(i,j) - \text{Current Field Two Value}(i,j) | \quad (4)$$

[0089] 용어 SAD 는 용어 "summed absolute differences" 의 약자이다. 도 8 에는 메트릭을 형성하도록 차분되는 필드가 그림으로 도시되어 있다. 아래 첨자는 필드 번호를 지칭하고, 글자는 이전 (= P) 또는 현재 (= C) 를 지칭한다. 도 8 의 괄호는 필드의 쌍 단위 (pair-wise) 차분을 지칭한다. SAD_{FS} 는 C_1 라벨을 붙인 현재 프레임의 필드 1 과 P_1 라벨을 붙인 이전 프레임의 필드 1 간의 차이 (도 8 에 제공된 정의에서 FS 라는 라벨을 붙인 괄호에 의해 걸쳐 있음) 를 지칭하고, SAD_{SS} 는 C_2 라벨을 붙인 현재 프레임의 필드 2 와 P_2 라벨을 붙인 이전 프레임의 필드 2 간의 차이 (SS 라벨을 붙인 괄호에 의해 걸쳐 있음) 를 지칭하고, SAD_{CO} 는 C_2 라벨을 붙인 현재 프레임의 필드 2 와 C_1 라벨을 붙인 현재 프레임의 필드 1 간의 차이 (CO 라벨을 붙인 괄호에 의해 걸쳐 있음) 를 지칭하며, SAD_{PO} 는 양쪽 모두가 PO 라벨을 붙인 괄호에 의해 걸쳐 있는 현재 프레임의 필드 1 과 이전 프레임의 필드 2 간의 차이를 지칭한다.

[0091] 아래에는, 각각의 SAD 를 평가하기 위한 계산 부하가 설명된다. 통상적인 NTSC 에 있어서, 약 480 개의 활성 수평 라인이 존재한다. 4:3 가로세로비 (aspect ratio) 를 가지며, 수평 방향으로 해상도가 동일하기 위해, $480 \times 4/3 = 640$ 과 등가인 수직 라인 또는 자유도가 존재해야 한다. 640×480 개 화소의 비디오 포맷은 ATSC (Advanced Television Standards Committee) 에 의해 수용된 포맷 중 하나이다. 따라서, 1/30

초의 프레임 지속기간마다, $640 \times 480 = 307,200$ 개의 새로운 화소가 생성된다. 새로운 데이터는 9.2×10^6 화소/초의 레이트로 생성되어, 이러한 시스템을 실행하는 하드웨어나 소프트웨어가 약 10 MB 이상의 레이트로 데이터를 처리한다는 것을 의미한다. 이는 그 시스템의 고속 부분 중 하나이다. 이는 하드웨어, 소프트웨어, 펌웨어, 미들웨어, 마이크로코드, 또는 그 임의의 조합에 의해 구현될 수 있다. SAD 계산기는 자립형 컴포넌트이거나, 다른 디바이스의 컴포넌트 내에 있는 하드웨어, 펌웨어, 미들웨어로서 포함되거나, 프로세서 상에서 실행되는 마이크로코드 또는 소프트웨어에 구현되거나, 또는 그 조합일 수 있다. 소프트웨어, 펌웨어, 미들웨어 또는 마이크로코드에 구현 시, 계산을 수행하는 프로그램 코드 또는 코드 세그먼트는 저장 매체와 같은 머신 판독가능 매체에 저장될 수도 있다. 코드 세그먼트는 절차, 함수, 서브프로그램, 프로그램, 루틴, 서브루틴, 모듈, 소프트웨어 패키지, 클래스, 또는 명령들, 데이터 구조들, 또는 프로그램 문장들의 임의의 조합을 표현할 수도 있다. 코드 세그먼트는 정보, 데이터, 인수 (argument), 파라미터, 또는 메모리 콘텐츠를 전달하고/하거나 수신함으로써 다른 코드 세그먼트 또는 하드웨어 회로에 연결될 수도 있다.

[0092] 도 9의 흐름도 (900)는 도 8의 관계를 명시적으로 나타낸 것으로서, 식 (1) 내지 식 (4)의 그림 표현이다. 이는 SAD_{FS} , SAD_{CO} , SAD_{SS} 및 SAD_{P0} 의 가장 최근 값이 각각 유지되는 저장 위치 (941, 942, 943 및 944)를 도시한다. 이들은 각각 4개의 절대 차이 합 계산기 (940)에 의해 각각 생성되어, 이전 제 1 필드 데이터의 휘도 값 (931), 현재 제 1 필드 데이터의 휘도 값 (932), 현재 제 2 필드 데이터의 휘도 값 (933) 및 이전 제 2 필드 데이터의 휘도 값 (934)을 처리한다. 메트릭을 정의하는 덧셈에 있어서, 용어 "값(i, j)"는 위치 i, j에 있는 휘도 값을 의미하고, 활성 화소 전부에 대해 덧셈이 이루어지지만, 활성 화소의 중요한 부분집합에 대한 덧셈이 제외되는 것은 아니다.

[0093] 도 10의 흐름도 (1000)는 텔레시네된 비디오를 검출하고, 이를 역 텔레시네하여 원래 스캔된 필름 이미지로 복원하는 프로세스를 도시하는 상세한 흐름도이다. 단계 1030에서, 도 9에 정의된 메트릭이 평가된다. 단계 1083으로 진행하여, 4개의 메트릭의 하부 포락선 값을 찾는다. SAD 메트릭의 하부 포락선은 동적으로 결정되는 양으로서, 최고 수치 플로어 (numerical floor)이며, 그 아래에서 SAD는 관통하지 않는다. 단계 1085로 진행하고, 아래의 식 (5) 내지 식 (10)에서 정의되는 분기 정보를 결정하여, 이전에 결정된 메트릭, 하부 포락선 값 및 실험적으로 결정된 상수 A를 이용할 수 있다. 연속적인 위상 값이 일치하지 않을 수도 있으므로, 단계 1087에서 이러한 외관상 불안정을 줄이기 위해 양 Δ 가 결정된다. 그 위상은, 위상 결정 시퀀스가 도 7에 도시된 당해 모델과 일치하는 경우에, 일치하는 것으로 생각된다. 그 단계 다음에, 프로세스는 단계 1089로 진행하여, 현재 Δ 값을 이용하여 결정 변수를 계산한다. 결정 변수 계산기 1089는 유도되어 1080의 블록에서 생성된 모든 정보를 이용하여 결정 변수를 평가한다. 단계 1030, 단계 1083, 단계 1085, 단계 1087 및 단계 1089는 도 6의 메트릭 결정 (651)의 확장이다. 이들 변수로부터, 위상 선택기 (1090)에 의해 적용 가능한 위상이 구해진다. 도시된 것과 같이, 결정 단계 1091는 적용 가능한 위상을 이용하여, 텔레시네된 비디오를 역 텔레시네하거나 디인터레이싱한다. 이는 도 4의 위상 검출기 (404) 동작의 보다 명시적인 표현이다. 일 양태에서, 도 10의 처리는 도 4의 위상 검출기 (404)에 의해 수행된다. 단계 1030에서 시작하여, 위상 검출기 (404)는 도 8을 참조하여 상술한 프로세스에 의해 복수의 메트릭을 결정하고, 단계 1083, 단계 1085, 단계 1087, 단계 1089, 단계 1090, 및 단계 1091을 걸쳐 진행한다.

[0094] 흐름도 (1000)는 현재 위상을 추정하는 프로세스를 도시한다. 단계 1083에서의 흐름도는 결정된 메트릭 및 하부 포락선 값을 이용하여 분기 정보를 계산하는 것을 설명한다. 분기 정보는 상술한 유클리드 거리로서 인식될 수도 있다. 분기 정보를 생성하는데 이용될 수도 있는 예시적인 식은 아래의 식 (5) 내지 식 (10)이다. 분기 정보 양은 도 12의 블록 (1209)에서 계산된다.

[0095] 처리된 비디오 데이터는 예를 들어, 프로세서에 접속된 디스크 타입 저장 매체 (예를 들어, 자기식 또는 광학식) 또는 칩 구성된 저장 매체 (예를 들어, ROM, RAM)를 포함할 수 있는 저장 매체에 저장될 수 있다. 몇몇 양태에서, 역 텔레시네 (406) 및 디인터레이서 (405)는 각각 저장 매체의 일부 또는 전부를 포함할 수 있다. 분기 정보 양은 다음 식에 의해 정의된다.

$$\text{Branch Info}(0) = (\text{SAD}_{FS} - H_S)^2 + (\text{SAD}_{SS} - H_S)^2 + (\text{SAD}_{PO} - H_P)^2 + (\text{SAD}_{CO} - L_C)^2 \quad (5)$$

$$\text{Branch Info}(1) = (\text{SAD}_{FS} - L_S)^2 + (\text{SAD}_{SS} - H_S)^2 + (\text{SAD}_{PO} - L_P)^2 + (\text{SAD}_{CO} - H_C)^2 \quad (6)$$

$$\text{Branch Info}(2) = (\text{SAD}_{FS} - H_S)^2 + (\text{SAD}_{SS} - H_S)^2 + (\text{SAD}_{PO} - L_P)^2 + (\text{SAD}_{CO} - H_C)^2 \quad (7)$$

$$\text{Branch Info}(3) = (\text{SAD}_{FS} - H_S)^2 + (\text{SAD}_{SS} - L_S)^2 + (\text{SAD}_{PO} - L_P)^2 + (\text{SAD}_{CO} - L_C)^2 \quad (8)$$

[0096]

$$\text{Branch Info}(4) = (\text{SAD}_{FS} - H_S)^2 + (\text{SAD}_{SS} - H_S)^2 + (\text{SAD}_{PO} - H_P)^2 + (\text{SAD}_{CO} - L_C)^2 \quad (9)$$

$$\text{Branch Info}(5) = (\text{SAD}_{FS} - L_S)^2 + (\text{SAD}_{SS} - L_S)^2 + (\text{SAD}_{PO} - L_P)^2 + (\text{SAD}_{CO} - L_C)^2 \quad (10)$$

[0097]

[0098]

분기 계산의 구체적인 상세는 도 12의 분기 정보 계산기 (1209)에 도시되어 있다. 계산기 (1209)에 도시된 바와 같이, 분기 정보를 전개하는 것은, SAD_{FS} 및 SAD_{SS} 의 하부 포락선 값인 L_S , SAD_{PO} 의 하부 포락선 값인 L_P , 및 SAD_{CO} 의 하부 포락선 값인 L_C 와 같은 양들을 이용한다. 하부 포락선들은 분기 정보 계산 시 거리 오프셋으로서 이용되거나, 단독으로 이용되거나, 또는 미리 결정된 상수 A 와 함께 이용되어, H_S , H_P 및 H_C 를 생성한다. 그 값들은 아래에서 설명되는 하부 포락선 추적기에 최신 값으로 유지된다. H 오프셋은 다음과 같이 정의된다.

$$H_S = L_S + A \quad (11)$$

$$H_{PO} = L_P + A \quad (12)$$

$$H_C = L_C + A \quad (13)$$

[0099]

[0100]

L_S , L_P 및 L_C 의 값들을 추적하는 프로세스는 도 13a, 도 13b 및 도 13c에 표현된다. 예를 들어, 도 13a의 상부에 도시된 L_P 에 대한 추적 알고리즘 (1300)을 고려하자. 매트릭 SAD_{PO} 는 비교기 (1305)에서 현재 L_P 값과 임계값 T_P 를 더한 것과 비교된다. 매트릭 SAD_{PO} 가 초과하는 경우에, 블록 1315에 도시된 바와 같이, 현재 L_P 값은 변하지 않는다. 매트릭 SAD_{PO} 가 초과하지 않는 경우에, 블록 1313에 도시된 바와 같이, 새로운 L_P 값은 SAD_{PO} 및 L_P 의 선형 조합으로 된다. 블록 1315에 대한 다른 양태에서, 새로운 L_P 값은 $L_P + T_P$ 이다.

[0101]

이와 유사하게, 도 13b 및 도 13c의 양들 L_S 및 L_C 가 계산된다. 동일 기능을 갖는 도 13a, 도 13b 및 도 13c의 처리 블록은 동일한 번호이지만 프라임 (' 또는 ")이 부여되어, 상이한 변수 세트에 대해 작용한다는 것을 나타낸다. 예를 들어, SAD_{PO} 및 L_C 의 선형 조합이 형성되는 경우에, 그 동작은 블록 1313'에 도시되어 있다. L_P 의 경우에서와 같이, 1315'에 대한 다른 양태는 L_C 를 $L_C + T_C$ 로 대체할 것이다.

[0102]

그러나, L_S 의 경우에, 도 13b의 알고리즘은 교대로 SAD_{PS} 와 SAD_{SS} 를 처리하여, 각각의 X 를 라벨링하는데, 그

이유는 이러한 하부 포락선이 양쪽 변수에 적용되기 때문이다. 블록 1308 내의 현재 SAD_{FS} 값이 블록 1303 에서 X 에 대한 위치로 판독된 다음에, 블록 1307 내의 현재 SAD_{SS} 값이 블록 1302 에서 X 에 대한 위치로 판독 될 때, SAD_{FS} 및 SAD_{SS} 값의 교번이 발생한다. L_p 에 대한 경우에서와 같이, 1315" 에 대한 다른 양태는 L_s 를 $L_s + T_s$ 로 대체할 것이다. 현재 하부 포락선 값 테스트 시 이용되는 임계값 및 양 A 는 실험에 의해 미리 결정된다.

[0103] 도 11 은 도 10 의 단계 1089 를 수행하는 예시적인 프로세스를 도시한 흐름도이다. 도 11 은 일반적으로 결정 변수를 갱신하는 프로세스를 도시한다. 메트릭으로부터 유도된 새로운 정보에 따라 6 개의 결정 변수 (6 개의 가능한 결정에 대응함) 가 갱신된다. 결정 변수는 다음과 같이 구해진다.

$$D_0 = \alpha D_4 + \text{Branch Info}(0) \quad (14)$$

$$D_1 = \alpha D_0 + \text{Branch Info}(1) \quad (15)$$

$$D_2 = \alpha D_1 + \text{Branch Info}(2) \quad (16)$$

$$D_3 = \alpha D_2 + \text{Branch Info}(3) \quad (17)$$

$$D_4 = \alpha D_3 + \text{Branch Info}(4) \quad (18)$$

$$D_5 = \alpha D_5 + \text{Branch Info}(5) \quad (19)$$

[0104]

[0105] 양 α 는 1 보다 작고, 자신의 과거 값들에 대한 결정 변수의 의존성을 제한하는데, α 의 이용은 그 데이터가 오래됨에 따라 각각의 유클리드 거리의 영향을 줄이는 것과 등가이다. 흐름도 (1162) 에서, 갱신될 결정 변수는 라인 (1101, 1102, 1103, 1104, 1105 및 1106) 상에서 이용 가능한 것으로서 왼쪽에 기재되어 있다. 그 다음에, 위상 전이 경로 중 하나의 위상 전이 경로 상의 각각의 결정 변수를 블록 (1100) 중 하나에서 1 보다 작은 숫자인 α 로 곱한 다음에, 오래된 결정 변수의 감소된 값을 감소된 결정 변수가 존재하였던 위상 전이 경로 상의 다음 위상에 의해 인덱싱된 분기 정보 변수의 현재 값에 더한다. 이는 블록 (1110) 에서 일어난다. 변수 D_5 는 블록 (1193) 에서 양 Δ 만큼 오프셋되고, Δ 는 블록 (1112) 에서 계산된다. 아래에서 설명하는 바와 같이, 그 양은 본 시스템에 의해 결정된 위상 시퀀스의 불일치를 줄이도록 선택된다. 블록 (1120) 에서 가장 작은 결정 변수가 구해진다.

[0106] 요약하면, 각각의 결정에 고유한 새로운 정보가 α 로 곱해진 적당한 결정 변수의 이전 값에 더해져서, 현재 결정 변수의 값을 얻는다. 새로운 메트릭을 제어할 수 있을 때에 새로운 결정을 내릴 수 있으므로, 본 기술은 모든 프레임마다 필드 1 및 2 의 수신 시 새로운 결정을 내릴 수 있다. 이들 결정 변수는 상술한 유클리드 거리의 합이다.

[0107] 적용 가능한 위상은 가장 작은 결정 변수의 아래첨자를 갖는 것이 되도록 선택된다. 결정 변수에 기초한 결정은 도 10 의 블록 1090 에서 명시적으로 행해진다. 어떤 결정이 결정 공간에서 허용된다. 블록 1091 에서 설명된 바와 같이, 이들 결정은, (i) 적용 가능한 위상이 P5 가 아니고 (비디오의 역 텔레시네), (ii) 적용 가능한 위상이 P5 이다 (비디오의 디인터레이싱) 라는 것이다.

[0108] 본질상 가변인 비디오로부터 메트릭이 도출되기 때문에, 결정의 코히런트 스트링에서 때때로 에러가 존재할 수도 있다. 이러한 기술은 도 7 과 일치하지 않는 위상 시퀀스를 검출한다. 그 동작은 도 14 에 개략적으로 도시되어 있다. 알고리즘 (1400) 은 블록 1405 에서 현재 위상 결정의 아래첨자 (=x) 를 저장하고, 블록 1406 에서 이전 위상 결정의 아래첨자 (=y) 를 저장한다. 블록 1410 에서, $x=y=5$ 가 테스트되면, 블록 1411

에서 다음과 같은 값, 즉,

[0109] $x=1, y=0$; 또는

[0110] $x=2, y=1$; 또는

[0111] $x=3, y=2$; 또는

[0112] $x=4, y=3$; 또는

[0113] $x=0, y=4$ 인지 여부가 테스트된다. 2 개의 테스트 중 어느 한쪽이라도 긍정적이면, 블록 1420 에서 일치하는 것으로 결정이 선언된다. 어느 한쪽도 긍정적이지 않으면, 도 11 의 블록 1193 에 도시되어 있는 오프셋이 도 15 에서 계산되어, P_5 와 연관된 결정 변수인 D_5 에 추가된다.

[0114] 또한, 도 15 에서 D_5 에 대한 수정은 프로세스 (1500) 의 일부로서 나타나고, 이 프로세스 (1500) 는 위상 시퀀스의 불일치에 대한 교정 액션을 제공한다. 흐름도 (1500) 의 블록 1510 에서 일치성 테스트가 실패한 것으로 가정하자. 블록 1510 에서 시작하는 "아니오" 분기를 따라 진행하여, 블록 1514 의 다음 테스트는 모든 $i < 5$ 에 대해 $D_5 > D_i$ 인지 여부, 또는 다른 방법으로는 변수 D_i ($i < 5$) 중 적어도 하나가 D_5 보다 큰지 여부에 관한 것이다. 첫 번째 경우가 유효하면, 블록 1516 에서, δ_0 를 초기값으로 갖는 파라미터 δ 가 $3\delta_0$ 로 변경된다. 두 번째 경우가 유효하면, 블록 1517 에서 δ 는 $4\delta_0$ 로 변경된다. 블록 152B 에서, Δ 의 값은 Δ_B 로 갱신되는데, 여기서,

[0115]
$$\Delta_B \equiv \max (\Delta - \delta, -40\delta_0) \tag{20}$$

[0116] 다시 블록 1510 으로 리턴하여, 결정 스트림이 일치하는 것으로 판정된 것으로 가정하자. 블록 1515 에서, 파라미터 δ 는 다음과 같이 정의되는 δ_+ 로 변경된다.

[0117]
$$\delta_+ \equiv \max (2\delta, 16\delta_0) \tag{21}$$

[0118] 블록 152A 에서, 새로운 δ 값이 Δ 에 대한 갱신 관계인 Δ_A 내에 삽입된다. 이는 다음과 같이 기재된다.

[0119]
$$\Delta_A \equiv \max (\Delta + \delta, 40\delta_0) \tag{22}$$

[0120] 그 다음에, 블록 1593 에서, 갱신된 Δ 값이 결정 변수 D_5 에 더해진다.

[0121] 도 16 은 일단 풀 다운 위상이 결정되면 역 텔레시네 프로세스가 진행되는 방법을 도시한다. 이러한 정보를 이용하여, 필드 (1605 및 1605') 는 동일 비디오 필드를 나타내는 것으로서 식별된다. 2 개의 필드는 서로 평균되고, 필드 (1606) 와 조합되어, 프레임 (1620) 을 재구성한다. 재구성된 프레임은 참조 부호가 1620' 이다. 이와 유사한 프로세스가 프레임 (1622) 을 재구성할 것이다. 프레임 (1621 및 1623) 으로부터 유도된 필드는 복제되지 않는다. 이들 프레임은 제 1 프레임과 제 2 프레임을 함께 조합하여 재구성된다.

[0122] 상술한 양태에서, 새로운 프레임이 수신될 때마다, 4 개의 새로운 메트릭 값이 구해지고, 6 겹의 가설 세트가 새로 계산된 결정 변수를 이용하여 테스트된다. 다른 처리 구조는 결정 변수를 계산하도록 적용될 수 있다.

비터비 디코더는 경로 메트릭을 형성하도록 함께 경로를 구성하는 분기의 메트릭을 더한다. 여기서 정의된 결정 변수는 유사한 규칙에 의해 형성되는데, 각각은 새로운 정보 변수의 "리키(leaky)" 합이다. (리키 합에 있어서, 결정 변수의 이전 값은 새로운 정보 데이터가 더해지기 전에 1 보다 작은 숫자로 곱해진다.) 비터비 디코더 구조는 이러한 절차의 동작을 지원하도록 수정될 수 있다.

[0123] 본 양태는 1/30 초마다 새로운 프레임이 나타나는 통상의 비디오를 처리하는 관점에서 설명되지만, 이러한 프로세스는 시간상 역방향으로 기록 및 처리되는 프레임에 적용될 수도 있다. 그 결정 공간은 동일 상태를 유지하지만, 입력 프레임 시퀀스의 시간 반전을 나타내는 중요치 않은 변화가 존재한다. 예를 들어, (여기에 도시된) 시간 반전된 모드로부터의 코히런트 텔레시네 결정의 스트림인,

[0124] $P_4 P_3 P_2 P_1 P_0$

[0125] 또한 시간상 반전될 것이다.

- [0126] 제 1 양태에 대한 이러한 변형을 이용하면, 결정 프로세스가, 성공적인 결정 수행 시, 2 회의 시도 (한 번의 시도는 시간상 순방향이며, 다른 한 번의 시도는 시간상 역방향임) 를 할 수 있다. 2 회의 시도가 독립적이지 않은 동안에는, 각각의 시도가 상이한 순서로 메트릭을 처리하도록 상이할 것이다.
- [0127] 이러한 사상은 부가물 (additional) 을 요구할 수도 있는 미래의 비디오 프레임은 저장하도록 유지된 버퍼와 관련하여 적용될 수 있다. 일 비디오 세그먼트가 순방향 처리에 있어서 허용할 수 없을 정도로 일치하지 않는 결과를 주는 것으로 판정되면, 그 절차는 버퍼로부터 미래의 프레임을 도출하고, 역방향으로 프레임을 처리함으로써 난해한 비디오 스트레치를 해결하려는 시도를 할 것이다.
- [0128] 또한, 본 출원에서 설명된 비디오의 처리는 PAL 포맷의 비디오에 적용될 수 있다.
- [0129] **디인터레이서**
- [0130] 본원에서 이용되는 "디인터레이서" 는 광의의 용어로서, 전체적으로 또는 상당 부분 인터레이싱된 멀티미디어 데이터를 처리하여 순차 멀티미디어 데이터를 형성하는 디인터레이싱 시스템, 디바이스, 또는 프로세스 (예를 들어, 프로세스를 수행하도록 구성된 소프트웨어, 펌웨어, 또는 하드웨어를 포함함) 를 설명하는데 이용될 수 있다.
- [0131] 비디오 카메라나 방송 스튜디오 등에서 통상적으로 생성되는 방송 비디오는 미국의 NTSC 표준에 따른다. 비디오를 압축하는 일반적인 방법은 비디오를 인터레이싱하는 것이다. 인터레이싱된 데이터에 있어서, 각각의 프레임은 2 개 필드 중 하나의 필드로 이루어진다. 한 필드는 프레임의 홀수 라인으로 이루어지고, 다른 필드는 프레임의 짝수 라인으로 이루어진다. 프레임이 약 30 프레임/초로 생성되는 동안에, 필드는 1/60 초 떨어져 있는 텔레비전 카메라의 이미지의 레코드이다. 인터레이싱된 비디오 신호의 각 프레임은 하나 걸러서 이미지의 수평 라인을 나타낸다. 프레임이 화면 상에 투영될 때, 비디오 신호는 짝수 라인과 홀수 라인 사이에서 교번한다. 이러한 것이 충분히 빨리, 예를 들어, 약 초당 60 프레임으로 행해질 때, 비디오 이미지는 사람 눈에 매끄럽게 보인다.
- [0132] 인터레이싱은 NTSC (미국) 및 PAL (유럽) 포맷에 기초하는 아날로그 텔레비전 방송에서 수십년 동안 이용되고 있다. 이미지의 절반만을 각각의 프레임으로 송신하면 되므로, 인터레이싱된 비디오는 전체 영상을 송신하는 경우보다 대략 절반의 대역폭을 이용하게 된다. 단말기 (160) 내부에서의 결과로서 생성된 비디오의 디스플레이 포맷은 반드시 NTSC 호환성이 있는 것은 아니며, 인터레이싱된 데이터를 쉽게 표시할 수 없다. 그 대신에, 최신 화소 기반 디스플레이 (예를 들어, LCD, DLP, LCOS, 플라즈마 등) 는 순차 스캔되어, 순차 스캔된 비디오 소스를 표시한다 (그런데, 다수의 구형 비디오 디바이스는 구형의 인터레이싱된 스캔 기술을 이용함).
 몇몇 일반적으로 이용되는 디인터레이싱 알고리즘의 예는 "Scan rate up-conversion using adaptive weighted median filtering," P. Haavisto, J. Juhola, and Y. Neuvo, *Signal Processing of HDTV II*, pp. 703-710, 1990, and "Deinterlacing of HDTV Images for Multimedia Applications," R. Simonetti, S. Carrato, G. Ramponi, and A. Polo Filisan, in *Signal Processing of HDTV IV*, pp. 765-772, 1993 에 설명되어 있다.
- [0133] 이하, 디인터레이싱의 성능을 개선하는데 단독으로 또는 조합하여 이용될 수 있으며, 디인터레이서 (405; 도 4) 에서 이용될 수 있는 시스템 및 방법에 대한 디인터레이싱 양태의 예를 설명한다. 그러한 양태는 공간-시간 필터링을 이용하여 선택된 프레임을 디인터레이싱함으로써 제 1 임시 디인터레이싱된 프레임을 결정하는 것, 양방향 모션 추정 및 모션 보상을 이용하여 선택된 프레임으로부터 제 2 임시 디인터레이싱된 프레임을 결정하는 것, 또한 제 1 및 제 2 임시 프레임을 조합하여 최종 순차 프레임을 형성하는 것을 포함할 수 있다. 공간-시간 필터링은 수평 또는 근방의 수평 에지의 블러링을 방지하는 수평 에지 검출기를 포함할 수 있는 Wmed (weighted median) 필터를 이용할 수 있다. "현재" 필드에 인접한 이전 및 후속 필드의 공간-시간 필터링은 예를 들어, 정적, 슬로우-모션, 및 패스트 모션과 같은 상이한 모션 레벨로 선택된 프레임의 일부를 분류하는 세기 모션-레벨 맵을 생성한다.
- [0134] 몇몇 양태에서, 세기 맵은 5 개의 인접한 필드 (2 개의 이전 필드, 현재 필드, 및 2 개의 다음 필드) 로부터의 화소를 포함한 필터링 애플처를 이용하여 Wmed 필터링에 의해 생성된다. Wmed 필터링은 장면 전환 및 추출하는 대상을 효과적으로 다룰 수 있도록 순방향, 역방향, 및 양방향 정적 영역 검출을 결정할 수 있다. 여러 양태에서, Wmed 필터는 필드 간 (inter-field) 필터링 모드에서 동일 패리티를 갖는 하나 이상의 필드에 걸쳐 이용될 수 있고, 임계값 기준을 미조정 (tweak) 함으로서 필드 내 (intra-field) 필터링 모드로 전환될 수 있다. 몇몇 양태에서, 모션 추정 및 보상은 루마 (화소의 세기나 밝기) 및 크로마 데이터 (화소의 컬러 정

보) 를 이용하여, 선택된 프레임의 디인터레이싱 영역을 개선하는데, 밝기 레벨은 거의 균일하지만 컬러는 상이하다. 디노이징 필터는 모션 추정 정확도를 증가시키는데 이용될 수 있다. Wmed 디인터레이싱된 임시 프레임에 디노이징 필터를 적용하여, Wmed 필터링에 의해 생성된 에일리어스 아티팩트를 제거할 수 있다. 아래에서 설명되는 디인터레이싱 방법 및 시스템은 양호한 디인터레이싱 결과를 생성하며 비교적 낮은 계산 복잡도를 가지므로, 고속으로 실행되는 디인터레이싱 구현을 허용함으로써, 그러한 구현을 디스플레이를 이용하는 휴대 전화, 컴퓨터 및 다른 타입의 전자 또는 통신 디바이스에 데이터를 제공하는데 이용되는 시스템을 비롯한 다양한 디인터레이싱 애플리케이션에 적합하게 만든다.

[0135] 본원에서, 디인터레이서 및 디인터레이싱 방법의 양태는 멀티미디어 데이터를 디인터레이싱하는데 이용되는 여러 컴포넌트, 모듈 및/또는 단계를 참조하여 설명된다.

[0136] 도 17 은 도 4 의 디인터레이서 (405) 로서 이용될 수 있는 디인터레이서 (1700) 의 일 양태를 도시한 블록도이다. 디인터레이서 (1700) 는 인터레이싱된 데이터의 적어도 일부를 공간 및 시간 ("공간-시간") 필터링하며 공간-시간 정보를 생성하는 공간 필터 (1730) 를 포함한다. 예를 들어, Wmed 는 공간 필터 (1730) 에서 이용될 수 있다. 또한, 몇몇 양태에서, 디인터레이서 (1700) 는 예를 들어, 위너 필터 (Weiner filter) 또는 웨이브렛 축소 필터와 같은 디노이징 필터 (도시생략) 를 포함한다. 또한, 디인터레이서 (1700) 는 인터레이싱된 데이터의 선택된 프레임의 모션 추정 및 보상을 제공하며 모션 정보를 생성하는 모션 추정기 (1732) 를 포함한다. 결합기 (1734) 는 공간-시간 정보와 모션 정보를 수신 및 조합하여 순차 프레임을 형성한다.

[0137] 도 18 은 디인터레이서 (1700) 의 다른 블록도이다. 디인터레이서 (1700) 내의 프로세서 (1836) 는 공간 필터 모듈 (1838), 모션 추정 모듈 (1840), 및 결합기 모듈 (1842) 을 포함한다. 외부 소스 (1848) 로부터의 인터레이싱된 멀티미디어 데이터는 디인터레이서 (1700) 내의 통신 모듈 (1844) 에 제공될 수 있다. 디인터레이서, 및 그 컴포넌트나 단계는 하드웨어, 소프트웨어, 펌웨어, 미들웨어, 마이크로코드, 또는 그 임의의 조합에 의해 구현될 수 있다. 예를 들어, 디인터레이서는 자립형 컴포넌트이거나, 다른 디바이스의 컴포넌트 내에 하드웨어, 펌웨어, 미들웨어로서 포함되거나, 프로세서 상에서 실행되는 마이크로코드나 소프트웨어에 구현되거나, 또는 그 조합으로 될 수도 있다. 소프트웨어, 펌웨어, 미들웨어 또는 마이크로코드에 구현 시, 디인터레이서 작업을 수행하는 프로그램 코드 또는 코드 세그먼트는 저장 매체와 같은 머신 판독가능 매체에 저장될 수도 있다. 코드 세그먼트는 절차, 함수, 서브프로그램, 프로그램, 루틴, 서브루틴, 모듈, 소프트웨어 패키지, 클래스, 또는 명령들, 데이터 구조들 또는 프로그램 문장들의 임의의 조합을 표현할 수도 있다. 코드 세그먼트는 정보, 데이터, 인수, 파라미터, 또는 메모리 콘텐츠를 전달하고/하거나 수신함으로써 다른 코드 세그먼트 또는 하드웨어 회로에 연결될 수도 있다.

[0138] 수신되는 인터레이싱된 데이터는, 예를 들어, 프로세서 (1836) 에 연결되는 칩 구성된 저장 매체 (예를 들어, ROM, RAM) 또는 디스크 타입 저장 매체 (예를 들어, 자기 또는 광학식) 를 포함할 수 있는 저장 매체 (1846) 내의 디인터레이서 (1700) 에 저장될 수 있다. 몇몇 양태에서, 프로세서 (1836) 는 저장 매체의 일부 또는 전부를 포함할 수 있다. 프로세서 (1836) 는 인터레이싱된 멀티미디어 데이터를 처리하여 순차 프레임을 형성하도록 구성되고, 이 순차 프레임은 다른 디바이스 또는 프로세스에 제공된다.

[0139] 텔레비전과 같은 통상의 아날로그 비디오 디바이스는 인터레이싱된 방식으로 비디오를 렌더링하는데, 즉, 그러한 디바이스는 짝수 번호를 갖는 스캔 라인 (짝수 필드) 과, 홀수 번호를 갖는 스캔 라인 (홀수 필드) 을 송신한다. 신호 샘플링 관점에서, 이는 다음과 같이 설명되는 패턴에서의 공간-시간 서브샘플링과 등가이다.

$$F(x, y, n) = \begin{cases} \Theta(x, y, n), & \text{if } y \bmod 2 = 0 \text{ 짝수 필드의 경우,} \\ \Theta(x, y, n), & \text{if } y \bmod 2 = 1 \text{ 홀수 필드의 경우,} \\ \text{소거,} & \text{otherwise,} \end{cases} \tag{23}$$

[0140] 여기서, Θ 는 원래 프레임 영상을 나타내고, F 는 인터레이싱된 필드를 나타내며, (x, y, n) 은 각각 화소의 수평, 수직, 및 시간 위치를 나타낸다.

[0142] 일반성을 잃지 않으면서, $n = 0$ 이 본 개시내용 전체에서 짝수 필드이므로, 식 (23) 은 다음과 같이 단순화될 수 있다.

$$F(x, y, n) = \begin{cases} \Theta(x, y, n), & \text{if } y \bmod 2 = n \bmod 2, \\ \text{소거}, & \text{otherwise,} \end{cases}$$

(24)

[0143]

[0144]

데시메이션 (decimation) 이 수평 방향으로 수행되므로, 서브샘플링 패턴은 다음 $n \sim y$ 좌표에서 표시될 수 있다. 도 19 에서, 원 모양과 별 모양 모두는 원래 풀-프레임 영상이 샘플 화소를 갖는 위치를 나타낸다. 인터레이싱 프로세스는 원 모양 화소에 영향을 주지 않으면서 별 모양 화소를 데시메이션한다. 수직 위치를 0 에서부터 인덱싱하므로, 짝수 필드가 상부 필드 (top field) 이고, 홀수 필드가 하위 필드 (bottom field) 인 것에 주목해야 한다.

[0145]

디인터레이서의 목적은 인터레이싱된 비디오 (필드 시퀀스) 를 인터레이싱되지 않은 순차 프레임 (프레임 시퀀스) 으로 변환하는 것이다. 즉, 짝수 필드와 홀수 필드를 보간 (interpolate) 하여 풀-프레임 영상을 "복원" 하거나 생성하는 것이다. 이는 식 (25) 에 의해 표현될 수 있다.

$$F_o(x, y, n) = \begin{cases} F(x, y, n), & y \bmod 2 = n \bmod 2, \\ F_i(x, y, n), & \text{otherwise,} \end{cases}$$

(25)

[0146]

[0147]

여기서, F_i 는 손실된 화소에 대한 디인터레이싱 결과를 나타낸다.

[0148]

도 20 은 Wmed 필터링 및 모션 추정을 이용하여 인터레이싱된 멀티미디어 데이터로부터 순차 프레임을 생성하는 디인터레이서의 일 양태의 몇몇 양상을 도시한 블록도이다. 도 20 의 상부는 현재 필드, 2 개의 이전 필드 (PP 필드 및 P 필드), 및 2 개의 후속 필드 (다음 필드 (Next Field) 및 다음 다음 필드 (Next Next Field)) 로부터의 정보를 이용하여 생성될 수 있는 모션 세기 맵 (2052) 을 도시한다. 아래에서 더 상세히 설명되는 바와 같이, 모션 세기 맵 (2052) 은 현재 프레임의 2 개 이상의 상이한 모션 레벨로 분류하거나 분할하고, 공간-시간 필터링에 의해 생성될 수 있다. 아래에서 식 (4) 내지 식 (8) 을 참조하여 설명되는 바와 같이, 몇몇 양태에서는, 모션 세기 맵 (2052) 을 생성하여, 정적 영역, 슬로우-모션 영역, 및 패스트-모션 영역을 식별한다. 예를 들어, Wmed 필터 (2054) 와 같은 공간-시간 필터는 모션 세기 맵에 기초한 기준을 이용하여 인터레이싱된 멀티미디어 데이터를 필터링하고, 공간-시간 임시 디인터레이싱된 프레임을 생성한다. 몇몇 양태에서, Wmed 필터링 프로세스는 [-1, 1] 의 수평 근방, [-3, 3] 의 수직 근방, 및 5 개의 인접한 필드의 시간 근방을 수반하고, 이는 도 20 에 도시된 5 개의 필드 (PP 필드, P 필드, 현재 필드, 다음 필드, 다음 다음 필드) 에 의해 표현되며, Z^{-1} 은 일 필드의 지연을 나타낸다. 현재 필드에 대해, 다음 필드와 P 필드는 넌-패리티 필드이고, PP 필드와 다음 다음 필드는 패리티 필드이다. 공간-시간 필터링에 이용되는 "근방" 은 필터링 동작 동안에 실제 이용되는 필드와 화소의 공간 및 시간 로케이션을 지칭하고, 예를 들어, 도 21 및 도 22 에 도시된 것과 같은 "에퍼처" 로서 나타낼 수 있다.

[0149]

또한, 디인터레이서는 디노이저 (잡음 제거 필터) (2056) 를 포함할 수 있다. 디노이저 (2056) 는 Wmed 필터 (2056) 에 의해 생성된 공간-시간 임시 디인터레이싱된 프레임을 필터링하도록 구성된다. 공간-시간 임시 디인터레이싱된 프레임을 잡음 제거함으로써, 소스 인터레이싱된 멀티미디어 데이터 시퀀스가 백색 잡음에 의해 오염되는 경우에 특히, 후속 모션 검색 프로세스를 보다 정확하게 만든다. 또한, Wmed 영상 내의 짝수 행과 홀수 행 사이의 에일리어스를 적어도 부분적으로 제거할 수 있다. 디노이저 (2056) 는 또한 아래에서 더 설명되는 웨이브렛 축소 및 웨이브렛 위너 필터를 비롯한 다양한 필터로서 구현될 수 있다.

[0150]

도 20 의 하부는 인터레이싱된 멀티미디어 데이터의 모션 정보 (예를 들어, 모션 벡터 후보, 모션 추정, 모션 보상) 를 결정하는 일 양태를 도시한다. 특히, 도 20 은, 선택된 프레임의 모션 보상된 임시 순차 프레임을 생성하는데 이용된 다음에 Wmed 임시 프레임과 결합되어 결과로서 생성된 "최종" 순차 프레임 (디인터레이싱된 현재 프레임 (2064) 으로서 도시됨) 을 형성하는 모션 추정 및 모션 보상 방식을 도시한다. 몇몇 양태에서, 인터레이싱된 멀티미디어 데이터의 모션 벡터 (MV) 후보 (또는 추정치) 는 외부 모션 추정기로부터 디인터레이서에 제공되고, 양방향 모션 추정기 및 보상기 (2068; ME/MC) 에 대한 시작 포인트를 제공하는데 이용된다. 몇몇 양태에서, MV 후보 선택기 (2072) 는 예를 들어, 디인터레이싱된 이전 프레임 (2070) 내의 블록과 같은 이전 처리된 블록의 MV 와 같이, 처리 중인 블록의 MV 후보에 대한 근방 블록에 대해 이전에 결정된 MV 를 이용

한다. 모션 보상은 이전 디인터레이싱된 프레임 (2070) 및 다음 (예를 들어, 미래) Wmed 프레임 (2058) 에 기초하여 양방향으로 행해질 수 있다. 현재 Wmed 프레임 (2060) 및 모션 보상된 (MC) 현재 프레임 (2066) 은 결합기 (2062) 에 의해 병합되거나 결합된다. 결과로서 생성되는 디인터레이싱된 현재 프레임 (2064), 즉, 순차 프레임은 ME/MC (2068) 에 다시 제공되어, 디인터레이싱된 이전 프레임 (2070) 으로서 이용되고, 또한 예를 들어, 압축 및 디스플레이 단말기로의 송신과 같은 후속 처리를 위해 디인터레이서 외부로 통신된다. 도 20 에 도시된 여러 양태는 아래에서 보다 상세히 설명된다.

[0151] 도 25 는 멀티미디어 데이터를 처리하여 인터레이싱된 프레임의 시퀀스로부터 순차 프레임의 시퀀스를 생성하는 프로세스 (2500) 를 도시한다. 일 양태에서, 순차 프레임은 도 4 에 도시된 디인터레이서 (405) 에 의해 생성된다. 블록 2502 에서, 프로세스 (2500; 프로세스 A) 는 선택된 프레임에 대한 공간-시간 정보를 생성한다. 공간-시간 정보는 멀티미디어 데이터의 모션 레벨을 분류하여 모션 세기 맵을 생성하는데 이용되는 정보를 포함할 수 있고, Wmed 임시 디인터레이싱된 프레임 및 그 프레임을 생성하는데 이용되는 정보 (예를 들어, 식 (26) 내지 식 (33) 에서 이용되는 정보) 를 포함한다. 이러한 프로세스는 도 20 의 상부에 도시된 바와 같이, Wmed 필터 (2054) 에 의해 수행될 수 있고, 그 연관된 처리는 아래에서 더 상세히 설명된다. 도 26 에 도시된 프로세스 A 에서는, 아래에서 더 설명되는 바와 같이, 블록 2602 에서, 영역들은 상이한 모션 레벨의 필드로 분류된다.

[0152] 다음에, 블록 2504 (프로세스 B) 에서, 프로세스 2500 는 선택된 프레임에 대한 모션 보상 정보를 생성한다. 일 양태에서, 도 20 의 하부에 도시되어 있는 양방향 모션 추정기/모션 보상기 (2068) 는 이러한 프로세스를 수행할 수 있다. 그 다음에, 프로세스 (2500) 는 블록 2506 으로 진행하여, 공간-시간 정보 및 모션 보상 정보에 기초하여 선택된 프레임의 필드를 디인터레이싱함으로써 선택된 프레임과 연관된 순차 프레임을 형성한다. 이는 도 20 의 하부에 도시되어 있는 결합기 (2062) 에 의해 수행될 수 있다.

[0153] **모션 세기 맵**

[0154] 각각의 프레임에 대해, 상이한 "모션" 의 영역들을 결정하도록 현재 필드에서 화소를 처리함으로써 모션 세기 맵 (2052) 을 결정할 수 있다. 3 개 카테고리의 모션 세기 맵을 결정하는 예시적인 양태는 도 21 내지 도 24 를 참조하여 아래에서 설명된다. 모션 세기 맵은 동일 패리티 필드 및 상이한 패리티 필드에서 화소를 비교하는 것에 기초하여 각각의 프레임의 영역을 정적 영역, 슬로우-모션 영역, 및 패스트 모션 영역으로서 지정한다.

[0155] **정적 영역**

[0156] 모션 맵의 정적 영역을 결정하는 것은, 일정 화소(들)의 휘도 차이가 일정 기준을 만족하는지 여부를 결정하기 위해 인접 필드의 근방에 있는 화소를 처리하는 것을 포함할 수 있다. 몇몇 양태에서, 모션 맵의 정적 영역을 결정하는 것은 5 개의 인접 필드 (현재 필드 (C), 현재 필드에 대해 시간상 이전에 있는 2 개의 필드, 및 현재 필드에 대해 시간상 다음에 있는 2 개의 필드) 의 근방에 있는 화소를 처리하여, 일정 화소(들)의 휘도 차이가 일정 임계값을 만족하는지 여부를 결정하는 것을 포함한다. 이들 5 개의 필드는 도 20 에 도시되어 있고, Z^{-1} 은 일 필드의 지연을 나타낸다. 즉, 통상, 5 개의 인접 필드는 Z^{-1} 의 시간 지연을 갖는 그러한 시퀀스에서 표시될 것이다.

[0157] 도 21 은 몇몇 양태에 따른 공간-시간 필터링에 이용될 수 있는 5 개의 필드 각각의 일정 화소를 식별하는 애퍼처를 도시한다. 애퍼처는, 왼쪽에서 오른쪽으로, 이전 이전 필드 (PP), 이전 필드 (P), 현재 필드 (C), 다음 필드 (N), 및 다음 다음 필드 (NN) 의 3×3 화소 그룹을 포함한다. 몇몇 양태에서, 현재 필드의 영역은, 식 (26) 내지 식 (28) 에 설명된 기준인, 도 21 에 도시되어 있는 화소 위치 및 그 대응 필드를 만족하는 경우에, 모션 맵에서 정적으로 간주된다.

[0158]
$$|L_P - L_N| < T_1 \tag{26}$$

[0159] 및

[0160]
$$\left(\frac{|L_{BPP} - L_B|}{2} \right) + \left(\frac{|L_{EPF} - L_E|}{2} \right) < T_1 \text{ (순방향 정적)} \tag{27}$$

[0161] 또는

$$\left| \frac{L_{BNN} - L_P}{2} \right| + \left| \frac{L_{ENN} - L_E}{2} \right| < T_1 \quad (\text{역방향 정적}), \quad (28)$$

[0162]

[0163] 여기서, T_1 은 임계값이고,

[0164] L_P 는 P 필드에 위치한 화소 P 의 휘도이고,

[0165] L_N 은 N 필드에 위치한 화소 N 의 휘도이고,

[0166] L_B 는 현재 필드에 위치한 화소 B 의 휘도이고,

[0167] L_E 는 현재 필드에 위치한 화소 E 의 휘도이고,

[0168] L_{BPP} 는 PP 필드에 위치한 화소 B_{PP} 의 휘도이고,

[0169] L_{EPP} 는 PP 필드에 위치한 화소 E_{PP} 의 휘도이고,

[0170] L_{BNN} 은 NN 필드에 위치한 화소 B_{NN} 의 휘도이며,

[0171] L_{ENN} 은 NN 필드에 위치한 화소 E_{NN} 의 휘도이다.

[0172] 임계값 T_1 은 특정 값으로 미리 결정 및 설정되고, 디인터레이싱 이외의 프로세스에 의해 결정 및 (예를 들어, 디인터레이싱되는 비디오에 대한 메타데이터로서) 제공되거나, 디인터레이싱 동안에 동적으로 결정될 수 있다.

[0173] 식 (26), 식 (27), 및 식 (28) 에서 상술한 정적 영역 기준은 적어도 2 가지 이유때문에 통상적인 디인터레이싱 기술보다 많은 필드를 이용한다. 첫 번째로, 동일 패리티 필드 간의 비교는 상이한 패리티 필드 간의 비교보다 낮은 에일리어스 및 위상-불일치를 갖는다. 그러나, 처리 중인 필드와 그 가장 인접한 동일 패리티 필드 근방 간의 최소 시간 차 (이에 따른 상관) 는 2 개의 필드로서, 그 상이한 패리티 필드 근방으로부터의 최소 시간 차보다 크다. 보다 신뢰성 있는 상이한 패리티 필드와 보다 낮은 에일리어스의 동일 패리티 필드의 조합은 정적 영역 검출의 정확도를 향상시킬 수 있다.

[0174] 또한, 도 21 에 도시된 바와 같이, 5 개의 필드는 현재 프레임 C 내의 화소 X 에 대해 과거 및 현재에서 대칭으로 분포될 수 있다. 정적 영역은 3 개의 카테고리로 세분될 수 있다: 순방향 정적 (이전 프레임에 대해 정적), 역방향 정적 (다음 프레임에 대해 정적), 또는 양방향 정적 (순방향 기준과 역방향 기준이 모두 만족되는 경우). 이러한 정적 영역의 세분화는 특히 장면 전환 및 대상 출몰 시 성능을 향상시킬 수 있다.

[0175] **슬로우-모션 영역**

[0176] 일정 화소의 휘도 값이 정적 영역으로 지정되는 기준을 만족하지 않지만 슬로우-모션 영역으로 지정되는 기준을 만족하는 경우에, 모션-맵에서 슬로우-모션 영역으로 간주될 수 있다. 아래의 식 (29) 는 슬로우-모션 영역을 결정하는데 이용될 수 있는 기준을 정의한다. 도 22 를 참조하면, 식 (29) 에서 식별된 화소 Ia, Ic, Ja, Jc, Ka, Kc, La, Lc, P 및 N 의 로케이션이 화소 X 를 중심으로 하는 애퍼처에 도시되어 있다. 그 애퍼처는 현재 필드 (C) 의 3×7 화소 근방 및 다음 필드 (N) 와 이전 필드 (P) 의 3×5 근방을 포함한다. 화소 X 는, 상기 기재된 정적 영역에 대한 기준을 만족하지 않고 애퍼처 내의 화소가 다음 식 (29) 에 도시된 기준을 만족하는 경우에, 슬로우-모션 영역의 일부로서 간주된다.

$$\left(|L_{Ia} - L_{Ic}| + |L_{Ja} - L_{Jc}| + |L_{Ka} - L_{Kc}| + |L_{La} - L_{Lc}| + |L_P - L_N| \right) / 5 < T_2 \quad (29)$$

[0178] 여기서, T_2 는 임계값이고,

[0179] L_{Ia} , L_{Ic} , L_{Ja} , L_{Jc} , L_{Ka} , L_{Kc} , L_{La} , L_{Lc} , L_P , L_N 은 각각 화소 Ia, Ic, Ja, Jc, Ka, Kc, La, Lc, P 및 N 에 대한 휘도 값이다.

[0180] 또한, 임계값 T_2 는 특정 값으로 미리 결정 및 설정되거나, 디인터레이싱 이외의 프로세스에 의해 결정 및 (예를 들어, 디인터레이싱되는 비디오에 대한 메타데이터로서) 제공되거나, 또는 디인터레이싱 동안에 동적으로 결정

될 수 있다.

[0181] 그 에지 검출 능력의 각도때문에, 일 필터는 수평인 (예를 들어, 수직으로 정렬된 것으로부터 45° 보다 큰) 에지를 블러링할 수 있다는 것에 주목해야 한다. 예를 들어, 도 22 에 도시된 에퍼처 (필터) 의 에지 검출 능력은 화소 A 및 F 또는 C 및 D 에 의해 형성된 각도에 의해 영향을 받는다. 최적으로 보간되지 않을 그러한 각도보다 수평인 임의의 에지 및 이에 따른 계단형 아티팩트가 그 에지에서 나타날 수도 있다. 몇몇 양태에서, 슬로우-모션 카테고리는 2 개의 하위 카테고리로 나뉠 수 있다. "수평 에지" 및 "기타" 는 이러한 에지 검출 효과를 설명한다. 슬로우-모션 화소는, 아래에 도시된 식 (30) 의 기준을 만족하면 수평 에지로서 분류되고, 식 (30) 의 기준을 만족하지 못하면 소위 "기타" 카테고리로서 분류될 수 있다.

[0182]
$$\{(LA + LB + LC) - (LD + LE + LF)\} < T_3 \tag{30}$$

[0183] 여기서, T_3 는 임계값이고, $LA, LB, LC, LD, LE,$ 및 LF 는 화소 A, B, C, D, E, 및 F 의 휘도 값이다.

[0184] 수평 에지 및 기타 카테고리 각각에 대해 상이한 보간 방법이 이용될 수 있다.

[0185] **패스트-모션 영역**

[0186] 정적 영역에 대한 기준 및 슬로우-모션 영역에 대한 기준을 만족하지 않으면, 화소는 패스트-모션 영역에 있는 것으로 간주될 수 있다.

[0187] 선택된 프레임에서 화소를 분류한 다음에, 프로세스 A (도 26) 는 블록 2604 로 진행하여, 모션 세기 맵에 기초하여 임시 디인터레이싱된 프레임을 생성한다. 본 양태에서, Wmed 필터 (2054; 도 20) 는 선택된 필드 및 필요한 인접 필드(들)를 필터링하여 다음과 같이 정의될 수 있는 후보 풀-프레임 이미지 F_0 를 제공한다.

$$F_0(\bar{x}, n) = \begin{cases} F(\bar{x}, n), & (y \bmod 2 = n \bmod 2) \\ \frac{1}{2}(F(\bar{x}, n-1) + F(\bar{x}, n+1)), & (\text{정적 역방향 및 순방향}) \\ F(\bar{x}, n-1) & (\text{정적 순방향이지만 순방향은 아님}) \\ F(\bar{x}, n+1) & (\text{정적 역방향이지만 역방향은 아님}) \\ \text{med}(A, B, C, D, E, F), & (\text{수평 에지를 갖지 않는 슬로우 모션}) \\ \text{med}\left(\alpha_0 \frac{A+F}{2}, \alpha_1 \frac{B+E}{2}, \alpha_2 \frac{C+D}{2}, \alpha_3 \frac{G+H}{2}\right), & (\text{수평 에지를 갖는 슬로우 모션}) \\ \frac{B+E}{2}, & (\text{패스트 모션}) \end{cases} \tag{31}$$

[0188] 여기서, a_i ($i = 0, 1, 2, 3$) 는 다음과 같이 계산되는 정수 가중치이다.

[0189]

$$\alpha_i = \begin{cases} 2 & \text{if } \beta_i = \min\{\beta_0, \beta_1, \beta_2, \beta_3\} \\ 1, & \text{otherwise} \end{cases} \tag{32}$$

[0190]

$$\beta_0 = \frac{A+F}{|A-F|}, \beta_1 = \frac{B+E}{|B-E|}, \beta_2 = \frac{C+D}{|C-D|}, \beta_3 = \frac{G+H}{|G-H|} \tag{33}$$

[0191]

[0192] 도 20 의 하부에 도시된 바와 같이, Wmed 필터링되는 임시 디인터레이싱된 프레임은 모션 추정 및 모션 보상 프로세스와 관련하여 후속 처리를 위해 제공된다.

[0193] 식 (31) 에 도시되며 상술한 바와 같이, 정적 보간은 필드 간 보간을 포함하고, 슬로우-모션 및 패스트-모션 보

간은 필드 내 보간을 포함한다. 동일 패리티 필드의 시간 (예를 들어, 필드 간) 보간이 바람직하지 않은 일정 양태에서, 시간 보간은 임계값 T_1 (식 (4) 내지 식 (6)) 을 0 ($T_1 = 0$) 으로 설정함으로써 "디스에이블" 될 수 있다. 시간 보간이 디스에이블된 상태에서 현재 필드를 처리함으로써, 모션-레벨 맵의 어떤 영역도 정적으로 분류하지 않게 되고, Wmed 필터 (2054; 도 20) 는 현재 필드 및 2 개의 인접한 난-패리티 필드에 대해 작용하는 도 22 의 애퍼처에 도시된 3 개의 필드를 이용한다.

[0194] **디노이징**

[0195] 일정 양태에서, 디노이저는 모션 보상 정보를 이용하여 더 처리되기 전에 후보 Wmed 프레임으로부터 잡음을 제거하는데 이용될 수 있다. 디노이저는 Wmed 필터에 존재하는 잡음을 제거할 수 있고, 신호의 주파수 콘텐츠에 관계없이 존재하는 신호를 유지할 수 있다. 웨이브렛 필터를 비롯한 다양한 타입의 디노이징 필터가 이용될 수 있다. 웨이브렛은 공간 영역과 스케일링 영역 모두에서 소정의 신호를 로컬화하는데 이용되는 함수 클래스이다. 웨이브렛의 기초가 되는 사상은, 웨이브렛 표현의 작은 변화가 그에 따라 원래 신호의 작은 변화를 생성하도록 상이한 스케일 또는 해상도로 신호를 분석하는 것이다.

[0196] 몇몇 양태에서, 디노이징 필터는 (4, 2) 배직교 (bi-orthogonal) 3 차 B-스플라인 웨이브렛 필터의 일 양태에 기초한다. 그러한 일 필터는 다음 순방향 및 역방향 변환에 의해 정의될 수 있다.

$$h(z) = \frac{3}{4} + \frac{1}{2}(z + z^{-1}) + \frac{1}{8}(z + z^{-2}) \quad (\text{순방향 변환}) \quad (34)$$

[0197]
[0198] 및

$$g(z) = \frac{5}{4}z^{-1} - \frac{5}{32}(1 + z^{-3}) - \frac{3}{8}(z + z^{-3}) - \frac{3}{32}(z^3 + z^{-4}) \quad (\text{역방향 변환}) \quad (35)$$

[0199]
[0200] 디노이징 필터를 적용함으로써, 잡음 환경에서 모션 보상의 정확도를 증가시킬 수 있다. 비디오 시퀀스의 잡음은 가산 백색 가우스인 것으로 가정한다. 추정된 잡음의 분산은 $\hat{\sigma}^2$ 로 표시된다. 이는 최고 주파수 부대역 계수를 0.6745 로 나눈 값의 평균 절대 편차 (median absolute deviation) 로서 추정될 수 있다. 그러한 필터의 구현은 "Ideal spatial adaptation by wavelet shrinkage," D.L. Donoho and I.M. Johnstone, *Biometrika*, vol. 8, pp. 425-455, 1994 에 더 설명되어 있으며, 이는 본원에서 그 전체를 참조로서 병합하고 있다.

[0201] 또한, 웨이브렛 축소 또는 웨이브렛 위너 필터가 디노이저로서 적용될 수 있다. 웨이브렛 축소 디노이징은 웨이브렛 변환 영역에서의 축소를 수반할 수 있고, 통상 3 개의 단계를 포함한다: 선형 순방향 웨이브렛 변환, 비선형 축소 디노이징, 및 선형 역방향 웨이브렛 변환. 위너 필터는 가산 잡음 및 블러링에 의해 열화된 이미지를 개선하는데 이용될 수 있는 MSE-최적 선형 필터이다. 그러한 필터는 당해 기술분야에서 널리 공지되어 있으며, 예를 들어, 상기 참조된 "Ideal spatial adaptation by wavelet shrinkage," 및 S.P.Ghael, A.M.Sayeed, 및 R.G.Baraniuk 에 의한 "Improvement Wavelet denoising via empirical Wiener filtering," *Proceedings of SPIE*, vol 3169, pp. 389-399, San Diego, July 1997 에 설명되어 있다.

[0202] **모션 보상**

[0203] 도 27 을 참조하면, 블록 2702 에서, 프로세스 B 는 양방향 모션 추정을 수행하고, 그 다음에, 블록 104 에서 모션 추정치를 이용하여 모션 보상을 수행하는데, 이는 도 20 에 더 도시되어 있으며 아래의 예시적인 양태에서 설명된다. Wmed 필터와 모션 보상 기반 디인터레이서 사이에는 일 필드의 "지연 (lag)" 이 존재한다. 도 23 에 도시된 바와 같이, 현재 필드 "C" 의 "잃어버린" 데이터 (화소 데이터의 원래가 아닌 (non-original) 행) 에 대한 모션 보상 정보는 이전 프레임 "P" 와 다음 프레임 "N" 모두의 정보로부터 예측된다. 현재 필드 (도 23) 에서, 실선은 원래 화소 데이터가 나가는 행을 표현하고, 점선은 Wmed-보간 화소 데이터가 나가는 행을 표현한다. 몇몇 양태에서, 모션 보상은 4 행 × 8 열 화소 근방에서 수행된다. 그러나, 이러한 화소 근방은 설명을 위한 일 예이며, 당업자라면, 다른 양태에서는 상이한 개수의 행과 상이한 개수의 열을 포함하는 화소 근방에 기초하여 모션 보상을 수행할 수도 있고, 그 선택은 예를 들어, 계산 속도, 이용 가능한 처리 전력, 또는 디인터레이싱되는 멀티미디어 데이터의 특징을 비롯한 다수의 인자에 기초할 수 있다는 것을 알 수 있다. 현재 필드가 그 행의 절반만을 가지므로, 실제 매칭되는 4 개의 행은 8 화소 × 8 화소 영역에 대응

한다.

- [0204] 도 20 을 참조하면, 양방향 ME/MC (2068) 는 SSE (sum of squared error) 를 이용하여, 디인터레이싱된 현재 프레임 (2070) 및 Wmed 다음 프레임 (2058) 에 대한 Wmed 현재 프레임 (2060) 에 대해 예측하는 블록과 예측된 블록 간의 유사도를 측정할 수 있다. 그 다음에, 모션 보상된 현재 프레임 (2066) 의 생성은 가장 유사한 매칭 블록으로부터의 화소 정보를 이용하여 원래 화소 라인 사이에 잃어버린 데이터를 채운다. 몇몇 양태에서, 양방향 ME/MC (2068) 는 디인터레이싱된 이전 프레임 (2070) 정보로부터의 화소 정보에 더 많은 가중치를 주거나 바이어싱하는데, 그 이유는, 그 프레임은 모션 보상 정보 및 Wmed 정보에 의해 생성되었지만, Wmed 다음 프레임 (2058) 은 공간-시간 필터링에 의해 디인터레이싱될 뿐이기 때문이다.
- [0205] 몇몇 양태에서는, 유사한 루마 영역을 갖지만 상이한 크로마 영역을 갖는 필드의 영역에서 매칭 성능을 개선하기 위해, (예를 들어, 하나의 4 행 × 8 열 루마 블록과 같은) 화소들 중 하나 이상의 루마 그룹 및 (예를 들어, 2 개의 2 행 × 4 열 크로마 블록 U 및 V 와 같은) 화소들 중 하나 이상의 크로마 그룹의 화소 값의 기여를 포함하는 메트릭을 이용할 수 있다. 그러한 접근법은 컬러에 민감한 영역에서의 불일치를 효과적으로 줄인다.
- [0206] 모션 벡터 (MV) 는 수직 차원으로 1/2 화소의 입도 (granularity) 및, 수평 방향으로 1/2 또는 1/4 화소의 입도를 갖는다. 프랙셔널 화소 샘플을 얻기 위해, 보간 필터를 이용할 수 있다. 예를 들어, 절반 화소 샘플을 획득하는데 이용될 수 있는 몇몇 필터는 이중 선형 (bilinear) 필터 (1, 1), H.263/AVC:의 권고 사항인 보간 필터 (1, -5, 20, 20, -5, 1), 및 6-탭 (six-tap) 헤밍 윈도우드 sinc 함수 필터 (3, -21, 147, 147, -21, 3) 를 포함한다. 1/4-화소 샘플은 이중 선형 필터를 적용함으로써 전체 화소와 절반 화소 샘플로부터 생성될 수 있다.
- [0207] 몇몇 양태에서, 모션 보상은 다양한 타입의 검색 프로세스를 이용하여, 다른 프레임 (예를 들어, 다음 프레임 또는 이전 프레임) 내의 상이한 위치에 있는 대응 데이터에 대해 현재 프레임의 일정 위치에 있는 데이터 (예를 들어, 대상을 나타내는 데이터) 를 일치시킬 수 있고, 각각의 프레임 내의 위치 차이는 대상의 모션을 나타낸다. 예를 들어, 검색 프로세스는 보다 큰 검색 영역을 커버할 수도 있는 풀 모션 검색을 이용하거나, 보다 적은 화소를 이용할 수 있는 고속 모션 검색을 이용하고/하거나 검색 패턴에서 이용되는 선택된 화소는 예를 들어, 다이아몬드 형상과 같이 특정 형상을 가질 수 있다. 고속 모션 검색의 경우에, 검색 영역은 모션 추정치, 또는 모션 후보를 중심으로 할 수 있으며, 이는 인접 프레임을 검색하기 위한 시작 포인트로서 이용될 수 있다. 몇몇 양태에서, MV 후보는 외부의 모션 추정기로부터 생성되어 디인터레이서에 제공될 수 있다. 또한, 이전 모션 보상된 인접 프레임에서 대응하는 근방으로부터의 매크로블록의 모션 벡터가 모션 추정치로부터 이용될 수 있다. 몇몇 양태에서, MV 후보는 대응하는 이전 및 다음 프레임의 매크로블록 (예를 들어, 3 매크로블록 × 3 매크로블록) 의 근방을 검색함으로써 생성될 수 있다.
- [0208] 도 24 는 2 개의 MV 맵인 MV_P 및 MV_N 의 일 예를 도시한 것으로서, 이 2 개의 MV 맵은 도 23 에 도시된 바와 같이 이전 프레임과 다음 프레임의 근방을 검색함으로써 모션 추정/보상 동안에 생성될 수 있다. MV_P 와 MV_N 모두에 있어서, 모션 정보를 결정하기 위해 처리될 블록은 "X" 로 표시되는 중심 블록이다. MV_P 와 MV_N 모두에 있어서, 처리 중인 현재 블록 X 의 모션 추정 동안에 이용될 수 있는 9 개의 MV 후보가 존재한다. 본 예에서, MV 후보 중 4 개의 MV 후보는 이전에 수행된 모션 검색으로부터 동일 필드에 존재하고, MV_P 및 MV_N 내의 보다 밝은 컬러의 블록으로 표시된다 (도 24). 보다 어두운 컬러의 블록으로 표시되는 5 개의 다른 MV 후보는 이전에 처리된 프레임의 모션 정보 (또는 맵) 로부터 복제되었다.
- [0209] 모션 추정/보상이 완료된 후에, 2 개의 보간 결과는 잃어버린 행 (도 23 에서 점선으로 표시됨) 으로부터 일어날 수도 있다: Wmed 필터에 의해 생성된 일 보간 결과 (Wmed 현재 프레임 (2060) 도 20), 및 모션 보상기의 모션 추정 처리에 의해 생성된 일 보간 결과 (MC 현재 프레임 (2066)). 통상, 결합기 (2062) 는 Wmed 현재 프레임 (2060) 과 MC 현재 프레임 (2066) 의 적어도 일부를 이용함으로써 Wmed 현재 프레임 (2060) 과 MC 현재 프레임 (2066) 을 병합하여, 현재 디인터레이싱된 프레임 (2064) 을 생성한다. 그러나, 일정 조건 하에서, 결합기 (2062) 는 현재 프레임 (2060) 또는 MC 현재 프레임 (2066) 중 하나만을 이용하여 현재 디인터레이싱된 프레임을 생성할 수도 있다. 일 예에서, 결합기 (2062) 는 Wmed 현재 프레임 (2060) 과 MC 현재 프레임 (2066) 을 병합하여, 식 (36) 에 도시된 것과 같은 디인터레이싱된 출력 신호를 생성한다.

$$F_o(\bar{x}, n) = \begin{cases} F(\bar{x}, n), & (y \bmod 2 = n \bmod 2) \\ k_2 F_{wmed}(\bar{x}, n) + (1 - k_2) F_{MC}(\bar{x} - \bar{D}, n - 1) & (otherwise) \end{cases} \quad (36)$$

[0210]

[0211] 여기서, $F(\bar{x}, n)$ 은 '가 전치 행렬인 위치 $\mathbf{x} = (x, y)^t$ 에 있는 필드 n 에서의 휘도 값을 위해 이용된다. 다음과 같이 정의된 클립 함수를 이용하면,

$$clip(0, 1, a) = 0, \text{ if } (a < 0); 1, \text{ if } (a > 1); a, \text{ otherwise} \quad (37)$$

[0213] k_1 은 다음과 같이 계산될 수 있다.

$$k_1 = clip(0, C_1, \sqrt{Diff}) \quad (38)$$

[0215] 여기서, C_1 은 강건성 파라미터 (robustness parameter) 이고, $Diff$ 는 (기존 필드로부터 선택된) 예측된 프레임 내의 이용 가능한 화소와 예측하는 프레임 화소 간의 루마이다. C_1 을 적절히 선택함으로써, 평균 제곱 에러의 상대 중요도를 조정하는 것이 가능하다. k_2 는 식 (39) 에 도시된 것과 같이 계산될 수 있다.

$$k_2 = 1 - clip\left(0, 1, (1 - k_1) \frac{|F_{wmed}(\bar{x} - \bar{y}_u, n) - F_{MC}(\bar{x} - \bar{y}_u - \bar{D}, n - 1)| + \delta}{|F_{wmed}(\bar{x}, n) - F_{MC}(\bar{x} - \bar{D}, n - 1)| + \delta}\right) \quad (39)$$

[0216]

[0217] 여기서, $\bar{x} = (x, y)$, $\bar{y}_u = (0, 1)$, \bar{D} 는 모션 벡터이고, δ 는 0 으로 나누는 것을 방지하기 위한 작은 상수이다. 또한, 필터링을 위한 클리핑 함수를 이용하여 디인터레이싱하는 것은 "De-interlacing of video data," G.D.Haan and E.B.Bellers, *IEEE Transactions on Consumer Electronics*, Vol.43, NO.3, pp.819-825, 1997 에 설명되어 있으며, 그 전체를 본원에서 참조로서 병합하고 있다.

[0218] 몇몇 양태에서, 결합기 (2062) 는 높은 PSNR 및 강건한 결과를 달성하기 위해 다음 식을 시도 및 유지하도록 구성될 수 있다.

$$|F_o(\bar{x}, n) - F_{wmed}(\bar{x}, n)| = |F_o(\bar{x} - \bar{y}_u, n) - F_{wmed}(\bar{x} - \bar{y}_u, n)| \quad (40)$$

[0219]

[0220] Wmed + MC 디인터레이싱 방식을 이용하여 필드 간 보간을 포함한 디인터레이싱 예측 방식을 필드 내 보간으로부터 분리하는 것이 가능하다. 즉, 공간-시간 Wmed 필터링은 필드 내 보간을 위해 주로 이용될 수 있지만, 필드 내 보간은 모션 보상 동안에 수행될 수 있다. 이로 인해, Wmed 결과의 피크 신호대 잡음비가 감소하지만, 모션 보상 후의 시각 품질은 더 만족스럽게 되는데, 그 이유는 정확하지 않은 필드 간 예측 모드 결정에 따라 불량 화소가 Wmed 필터링 프로세스로부터 제거될 것이기 때문이다.

[0221] 크로마 처리는 연관된 루마 처리와 양립할 수 있다. 모션 맵 생성의 관점에서, 크로마 화소의 모션 레벨은 4 개의 연관된 루마 화소의 모션 레벨을 관측함으로써 얻어진다. 그 동작은 보팅 (voting) 에 기초할 수 있다 (크로마 모션 레벨은 지배적인 루마 모션 레벨을 차용함). 그러나, 다음과 같은 통상적인 접근법을 이용할 것을 제안한다. 4 개의 루마 화소 중 어느 하나가 패스트 모션 레벨을 갖는 경우에, 크로마 모션 레벨은 패스트 모션일 것이다; 한편, 4 개의 루마 화소 중 어느 하나가 슬로우 모션 레벨을 갖는 경우에, 크로마 모션 레벨은 슬로우 모션일 것이다; 그 밖의 경우에는, 크로마 모션 레벨은 정적이다. 통상적인 접근법은 최고 PSNR 을 달성하지 못할 수도 있지만, 크로마 모션 레벨에 모호성이 존재하는 경우에는 언제나 INTER 예측을 이용하는 위험을 회피하게 된다.

[0222] 멀티미디어 데이터 시퀀스는 본원에서 설명된 Wmed 알고리즘만을 이용하거나, Wmed 알고리즘과 모션 보상된 알고리즘을 함께 이용하여 디인터레이싱되었다. 또한, 동일 멀티미디어 데이터 시퀀스는 화소 블렌딩 (또는 평균화) 알고리즘 및 "디인터레이싱 없음 (no-deinterlacing)" 케이스를 이용하여 디인터레이싱되었고, 이때, 필드는 어떤 보간이나 블렌딩 없이 단순히 결합되었다. 그 결과로서 생성된 프레임을 분석하여, PSNR 을 결

정하였고, 다음 테이블에 도시되어 있다.

[0223]

PSNR (dB) 시퀀스	디인터레이싱 없음	블렌딩	Wmed	Wmed + MC
soccer	8.955194	11.38215	19.26221	19.50528
city	11.64183	12.93981	15.03303	15.09859
crew	13.32435	15.66387	22.36501	22.58777

[0224]

Wmed 에 더해 MC 를 이용하여 디인터레이싱함으로써 한계 개선 (marginal improvement) 만이 존재하더라도, Wmed 와 MC 보간 결과를 조합함으로써 생성되는 디인터레이싱된 이미지의 시각 품질은 시각적으로 더 만족스럽게 되는데, 그 이유는, 상술한 바와 같이, Wmed 결과와 MC 결과를 조합함으로써 짝수 필드와 홀수 필드 간의 에일리어스 및 잡음을 억제하기 때문이다.

[0225]

몇몇 리샘플링 양태에서는, 영상 크기 리사이징을 위해 다상 (poly-phase) 리샘플러가 구현된다. 다운샘플링의 일 예에서, 원래 영상과 리사이징된 영상 간의 비율은 p/q 일 수 있고, 여기서, p 와 q 는 서로 소인 정수이다. 총 위상 수는 p 이다. 몇몇 양태에서, 다상 필터의 차단 주파수는 약 0.5 의 리사이징 계수에 대해 0.6 이다. 리사이징된 시퀀스의 고주파수 응답을 높이기 위해, 차단 주파수가 리사이징 비율과 정확히 일치하는 것은 아니다. 이는 필연적으로 약간의 에일리어싱을 허용한다. 그러나, 사람의 눈이 흐리고 에일리어싱이 없는 영상보다 뚜렷하지만 약간의 에일리어싱이 있는 영상을 선호한다는 것이 널리 공지되어 있다.

[0226]

도 42 는 리사이징 비율 3/4 인 경우의 위상을 도시하는 다상 리샘플링의 일 예를 도시한다. 또한, 도 42 에 도시된 차단 주파수는 3/4 이다. 원래 화소는 상기 도 42 에서 수직축으로 도시되어 있다. 또한, 그 수직축을 중심으로 sinc 함수를 도시하여 필터 파형을 표현한다. 리샘플링 비율과 정확히 동일하도록 차단 주파수를 선택하기 때문에, sinc 함수의 제로는 리사이징 후의 화소 위치와 겹치게 되며, 이는 도 42 에 교차점으로 도시되어 있다. 리사이징 후의 화소 값을 구하기 위해, 다음 식에 도시된 것과 같이 원래 화소로부터의 기여분을 합계할 수 있다.

$$v(x) = \sum_{i=-\infty}^{\infty} u(i) \times \sin c(\pi f_c (i - x)) \tag{41}$$

[0227]

여기서, f_c 는 차단 주파수이다. 상기 1-D 다상 필터는 수평 차원과 수직 차원 모두에 적용될 수 있다.

[0229]

리샘플링 (리사이징) 의 다른 양태는 오버스캔 (overscan) 을 설명하고 있다. NTSC 텔레비전 신호에 있어서, 일 이미지는 486 개의 스캔 라인을 갖고, 디지털 비디오에 있어서, 각각의 스캔 라인마다 720 개의 화소를 가질 수 있다. 그러나, 크기와 화면 포맷 간의 불일치 때문에 텔레비전에서 전체 이미지의 모두를 볼 수 있는 것은 아니다. 볼 수 없는 이미지의 일부가 오버스캔으로 지칭된다.

[0230]

가능한 한 많은 텔레비전이 볼 수 있는 영역에 유용한 정보를 브로드캐스터가 제공하는 것을 돕기 위해, SMPTE (Society of Motion Picture & Television Engineers) 는 세이프 액션 (safe action) 영역과 세이프 타이틀 (safe title) 영역으로 지칭되는 액션 프레임의 특정 크기를 규정하였다. Specifications for Safe Action and Safe Title Areas Test Pattern for Television Systems 에 대한 SMPTE RP (recommended practice) 27.3-1989 참조. SMPTE 에 의해 세이프 액션 영역은 "모든 중요한 액션이 일어나야 하는" 영역으로서 규정된다.

세이프 타이틀 영역은 "모든 유용한 정보가 다수의 가정용 텔레비전 수신기에 대한 가시성을 확보하도록 한정될 수 있는" 영역으로서 규정된다. 예를 들어, 도 43 에 도시된 바와 같이, 세이프 액션 영역 (4310) 은 화면 중심 90% 를 점유하여, 주변 모두에 5% 경계를 준다. 세이프 타이틀 영역 (4305) 는 화면의 중심 80% 를 점유하여, 10% 경계를 준다.

[0231]

도 44 를 참조하면, 세이프 타이틀 영역이 너무 작아서 그 이미지에 보다 많은 콘텐츠를 추가할 수 없기 때문에, 몇몇 브로드캐스트는 세이프 액션 영역에 텍스트를 포함시킬 것이고, 이는 백색 직사각형 윈도우 (4415) 내부에 있다. 통상, 오버스캔에서는 검은색의 경계를 볼 수 있다. 예를 들어, 도 44 에서, 검은색의 경계는 이미지의 상부 (4420) 와 하부 (4425) 에서 나타난다. 이들 검은색의 경계를 오버스캔에서 제거할 수 있는데, 그 이유는 H.264 비디오가 모션 추정 시 경계 확장을 이용하기 때문이다. 확장된 검은색의

경계는 레지듀얼 (residual) 을 증가시킬 수 있다. 통상적으로, 경계를 2% 만큼 자를 수 있고, 그 다음에, 리사이징을 행한다. 이에 따라, 리사이징을 위한 필터가 생성될 수 있다. 다상 다운샘플링 전에 오버스캔을 제거하기 위해 잘라 버림 (truncation) 이 수행된다.

[0232] **디블록킹/디링잉**

[0233] 디블록킹 처리의 일 예에서, 프레임의 경계에 있는 에지 및 디블록킹 필터 프로세스가 디스에이블되는 임의의 에지를 제외하고는, 일 프레임의 모든 4×4 블록 에지에 디블록킹 필터가 적용될 수 있다. 이러한 필터링 프로세스는, 프레임 내의 모든 매크로블록이 매크로블록 주소가 증가하는 순서로 처리되는 프레임 구성 프로세스의 완료 후에 매크로블록에 기초하여 수행될 것이다. 각각의 매크로블록에 대해, 처음에 수직 에지가 필터링되고, 좌측에서 우측으로, 그 다음에 수평 에지가 상부에서 하부로 필터링된다. 도 39 에 도시된 바와 같이, 수평 방향과 수직 방향에 대해, 루마 디블록킹 필터 프로세스는 4 개의 16-샘플 에지에 대해 수행되고, 각각의 크로마 성분에 대한 디블록킹 필터 프로세스는 2 개의 8-샘플 에지에 대해 수행된다. 이전 매크로블록에 대한 디블록킹 프로세스 동작에 의해 이미 수정되었을 수도 있는 현재 매크로블록의 위쪽 및 왼쪽에 있는 샘플 값은 현재 매크로블록에 대한 디블록킹 필터 프로세스에 대한 입력으로서 이용될 것이며, 또한 현재 매크로블록의 필터링 동안에 수정될 수도 있다. 수직 에지의 필터링 동안에 수정된 샘플 값은 동일 매크로블록에 대한 수평 에지의 필터링에 대한 입력으로서 이용될 수 있다. 디블록킹 프로세스는 루마와 크로마 성분에 대해 개별적으로 호출될 수 있다.

[0234] 디링잉 처리의 일 예에 있어서, 2-D 필터는 에지 근방의 영역을 평활화 (smooth out) 하는데 적응적으로 적용될 수 있다. 에지 화소는 블러링을 회피하기 위해 거의 필터링을 받지 않거나 전혀 필터링을 받지 않는다.

[0235] **GOP 분할기**

[0236] 이하, GOP 분할기에 포함될 수 있는 적응 GOP 분할, 샷 검출 및 대역폭 맵 생성을 비롯한 처리의 예시적인 실시 형태가 설명된다.

[0237] **대역폭 맵 생성**

[0238] 인간 시각 품질 V 는 인코딩 복잡도 C 와 할당된 비트 B (또한, 대역폭으로서 지칭됨) 모두의 함수일 수 있다. 도 29 는 이러한 관계를 도시한 그래프이다. 인코딩 복잡도 메트릭 C 는 인간의 시각 관점으로부터의 공간 및 시간 주파수를 고려한다는 것에 주목해야 한다. 인간의 눈에 더 민감한 왜곡의 경우에, 이에 대응하여 복잡도 값이 높아진다. 통상, V 는 C 에서는 단조 감소하고 있으며, B 에서는 단조 증가하고 있는 것으로 가정할 수 있다.

[0239] 일정한 시각 품질을 달성하기 위해, 대역폭 (B_i) 이 i 번째 대상 (프레임 또는 MB) 에 할당되어, 바로 아래의 2 개의 식에서 표현되는 기준을 만족하도록 인코딩된다.

$$B_i = B(C_i, V) \tag{42}$$

$$B = \sum_i B_i \tag{43}$$

[0240] 바로 위에 있는 2 개의 식에 있어서, C_i 는 i 번째 대상의 인코딩 복잡도이고, B 는 총 이용 가능한 대역폭이며, V 는 일 대상에 대해 달성된 시각 품질이다.

[0242] 인간 시각 품질을 소정의 식으로서 공식화하는 것은 어렵다. 따라서, 상기 식 세트가 정확하게 정의되는 것은 아니다. 그러나, 3-D 모델이 모든 변수에 있어서 연속인 것으로 가정하면, 대역폭 비율 $\left(\frac{B_i}{B}\right)$ 은 (C, V) 쌍의 근방 내에서 변하지 않는 것으로서 취급될 수 있다. 대역폭 비율 β_i 는 아래에 도시된 식에서 정의된다:

$$\beta_i = \frac{B_i}{B} \tag{44}$$

[0243]

[0244] 그 다음에, 비트 할당은 다음 식에서 표현된 것과 같이 정의될 수 있다.

$$\begin{aligned} \beta_i &= \beta(C_i) \\ 1 &= \sum_i \beta_i \\ \text{for} & \\ (C_i, V) &\in \delta(C_0, V_0) \end{aligned} \tag{45}$$

[0245] 여기서, δ 는 "근방" 을 나타낸다.

[0247] 인코딩 복잡도는 공간 및 시간 모두의 인간 시각 감도에 의해 영향을 받는다. 지로드 (Girod) 의 인간 시각 모델은 공간 복잡도를 정의하는데 이용될 수 있는 모델의 일 예이다. 이러한 모델은 로컬 공간 주파수와 주 위 조명을 고려한다. 그 결과로서 생성된 메트릭은 D_{csat} 로 지칭된다. 프로세스 내의 전처리 포인트에 서, 영상이 인트라 코딩될 것인지 인터 코딩될 것인지 여부는 알려지지 않고, 양쪽에 대한 대역폭 비율이 생성 된다. 상이한 비디오 대상의 β_{INTRA} 간의 비율에 따라 비트가 할당된다. 인트라 코딩된 영상의 경우에, 대역폭 비율은 다음 식에서 표현된다.

$$\beta_{INTRA} = \beta_{QINTRA} \log_{10} (1 + \alpha_{INTRA} Y^2 D_{csat}) \tag{46}$$

[0248] 상기 식에 있어서, Y 는 매크로블록의 평균 휘도 성분이고, α_{INTRA} 는 휘도 제곱 및 그 다음에 오는 D_{csat} 항에

대한 가중 계수이며, β_{QINTRA} 는 $1 = \sum_i \beta_i$ 을 보장하기 위한 정규화 계수이다. 예를 들어, $\alpha_{INTRA} = 4$ 인 값은 양호한 시각 품질을 달성한다. 콘텐츠 정보 (예를 들어, 콘텐츠 분류) 를 이용하여, 특정 비디오 콘텐츠에 대해 원하는 양호한 시각 품질 레벨에 대응하는 값으로 α_{INTRA} 를 설정할 수 있다. 일 예에서, 비 디오 콘텐츠가 "talking head" 뉴스 브로드캐스트를 포함하면, 비디오의 정보 이미지 또는 디스플레이 가능 부 분이 오디오 부분보다 낮은 중요도를 갖는 것으로 간주될 수도 있으므로, 시각 품질 레벨을 낮게 설정할 수도 있고, 그 데이터를 인코딩하는데 더 적은 비트를 할당할 수 있다. 다른 예에 있어서, 비디오 콘텐츠가 스포 츠 이벤트를 포함하면, 표시된 이미지가 시청자에게 더 중요할 수도 있으므로, 콘텐츠 정보를 이용하여, 보다 높은 시각 품질 레벨에 대응하는 값으로 α_{INTRA} 를 설정할 수도 있고, 이에 따라, 그 데이터를 인코딩하는데 더 많은 비트를 할당할 수 있다.

[0250] 이러한 관계를 이해하기 위해서는, 대역폭이 인코딩 복잡도에 따라 대수적으로 할당된다는 것에 주목해야 한다.

휘도를 제곱한 항 Y^2 은 더 큰 크기를 갖는 계수가 인코딩을 위해 더 많은 비트를 이용한다는 것을 나타낸다. 로그가 음의 값으로 되는 것을 방지하기 위해, 괄호 안의 항에 1 이 더해진다. 또한, 다른 밑을 갖는 로그를 이용할 수 있다.

[0251] 시간 복잡도는 프레임 차이 메트릭을 측정하여 결정되는 것으로서, SAD (sum of the absolute difference) 와 같은 프레임 차이 메트릭과 함께 모션 양 (예를 들어, 모션 벡터) 을 고려하여 2 개의 연속적인 프레임 간의 차 이를 측정한다.

[0252] 인터 코딩된 영상에 대한 비트 할당은 시간 복잡도뿐만 아니라 공간 복잡도도 고려할 수 있다. 이는 다음 식과 같이 표현된다.

$$\beta_{INTER} = \beta_{QINTER} \log_{10} (1 + \alpha_{INTER} \cdot SSD \cdot D_{csat} \exp(-\gamma \|MV_P + MV_N\|^2)) \tag{47}$$

[0253] 상기 식에 있어서, MV_P 와 MV_N 은 현재 MB 에 대한 순방향 및 역방향 모션 벡터이다. 인트라 코딩된 대역폭 공식 내의 Y^2 은 SSD (sum of squared difference) 로 대체됨을 알 수 있다. 상기 식에서

$\|MV_P + MV_N\|^2$ 의 역할을 이해하기 위해서는, 인간 시각 시스템의 다음 특징에 주목하자: 평활하며 예측

가능한 모션 (작은 $\|MV_P + MV_N\|^2$) 을 경험한 영역이 주의를 끌고, 인간 눈에 의해 추적될 수 있으며, 통상 정지 영역보다 임의의 많은 왜곡을 허용할 수는 없다. 그러나, 고속이거나 예측할 수 없는 모션 (큰 $\|MV_P + MV_N\|^2$) 을 경험한 영역은 추적될 수 없고, 상당한 양자화를 허용할 수 있다. 실험에는 $a_{INTRA} = 1, \gamma = 0.001$ 은 양호한 시각 품질을 달성한다는 것을 나타낸다.

[0255] **샷 검출**

[0256] 이하, 샷 검출의 예시적인 실시형태를 설명한다. 그러한 컴포넌트 및 프로세스는 GOP 분할기 (412) 에 포함될 수 있다 (도 4).

[0257] 모션 보상기 (23) 는 비디오 내의 프레임에 관한 양방향 모션 정보를 결정하도록 구성될 수 있다. 또한, 모션 보상기 (23) 는 예를 들어, SAD (sum of absolute difference) 또는 SSD (sum of squared difference) 와 같은 하나 이상의 차이 메트릭을 결정하고, 하나 이상의 프레임에 대한 휘도 정보 (예를 들어, 매크로블록 (MB) 휘도 평균 또는 차이), 휘도 히스토그램 차이, 및 프레임 차이 메트릭을 비롯한 다른 정보를 계산하도록 구성될 수 있고, 그 예들은 식 (1) 내지 식 (3) 을 참조하여 설명되어 있다. 샷 분류기는 모션 보상기에 의해 결정된 정보를 이용하여 비디오 내의 프레임을 2 개 이상의 "샷" 카테고리 분류하도록 구성될 수 있다. 인코더는 샷 분류에 기초하여 복수의 프레임을 적응 인코딩하도록 구성된다. 이하, 식 (1) 내지 식 (10) 을 참조하여 모션 보상기, 샷 분류기, 및 인코더를 설명한다.

[0258] 도 28 은 몇몇 양태에 따른 샷 검출 및 기타 전처리 동작을 위해 구성된 프로세서 (2831) 를 포함하는 전처리기 (202) 의 블록도이다. 디지털 비디오 소스는 도 4 에 도시된 것과 같이 전처리기 (202) 외부에 있는 소스에 의해 제공되어 전처리기 (202) 내의 통신 모듈 (2836) 로 통신될 수 있다. 전처리기 (202) 는 프로세서 (2831) 와 통신하는 저장 매체 (2835) 를 포함하고, 저장 매체 (2835) 와 프로세서 (2831) 모두는 통신 모듈 (2836) 과 통신한다. 프로세서 (2831) 는 모션 보상기 (2832), 샷 분류기 (2833), 및 기타 전처리 모듈 (2034) 을 포함하여, 본원에서 설명되는 것과 같이, 모션 정보를 생성하고, 비디오 데이터의 프레임 내의 샷을 분류하며 기타 전처리 테스트를 수행하도록 동작할 수 있다. 모션 보상기, 샷 분류기, 및 기타 모듈은 도 4 의 대응하는 모듈과 유사한 프로세스를 포함할 수 있고, 아래에서 설명되는 정보를 결정하도록 비디오를 처리할 수 있다. 특히, 프로세서 (2831) 는 복수의 비디오 프레임의 인접 프레임 간의 차이를 나타내는 메트릭 (이 메트릭은 양방향 모션 정보 및 휘도 정보를 포함함) 을 획득하고, 상기 메트릭에 기초하여 복수의 비디오 프레임의 샷 전환을 결정하며, 샷 전환에 기초하여 복수의 프레임을 적응 인코딩하는 구성을 가질 수 있다. 몇몇 양태에서, 그 메트릭은 프로세서 (2831; 이 프로세서는 또한 전처리기 (202) 외부에 있을 수 있음) 외부에 있는 디바이스나 프로세스에 의해 계산될 수 있고, 다른 디바이스나 메모리를 통해 직접 또는 간접적으로 프로세서 (2831) 로 통신될 수 있다. 또한, 그 메트릭은 예를 들어, 모션 보상기 (2832) 와 같은 프로세서 (2831) 에 의해 계산될 수 있다.

[0259] 전처리기 (202) 는 추가적인 처리, 인코딩 및 송신을 위해 비디오 및 메타데이터를 예를 들어, 단말기 (160; 도 1) 에 제공한다. 몇몇 양태에서, 인코딩된 데이터는 베이스 계층 및 인핸스먼트 계층을 포함할 수 있는 스케일러블 다층 인코딩된 비디오일 수 있다. 스케일러블 계층 인코딩은, 본원의 양수인이 소유하고 있고, 본원에서 그 전체를 참조로서 병합하고 있으며, 발명의 명칭이 "SCALABLE VIDEO CODING WITH TWO LAYER ENCODING AND SINGLE LAYER DECODING" 인 동시 계속중인 미국 특허 번호 [Attorney docket no. 050078] 에 더 설명되어 있다.

[0260] 몇몇 양태에서, 도 28 과 관련하여 설명된 여러 예시적인 논리 블록, 컴포넌트, 모듈 및 회로와, 본원에 개시된 다른 예 및 도면은, 범용 프로세서, DSP (digital signal processor), ASIC (application specific integrated circuit), FPGA (field programmable gate array) 나 기타 프로그램 가능 논리 디바이스, 별도의 게이트나 트랜지스터 논리, 별도의 하드웨어 컴포넌트, 또는 본원에 설명된 기능을 수행하도록 설계된 그 임의의 조합으로 구현되거나 수행될 수도 있다. 도 28 에 도시된 것과 같은 범용 프로세서는 마이크로프로세서일 수도 있지만, 다른 방법으로는, 그 프로세서는 임의의 통상적인 프로세서, 컨트롤러, 마이크로컨트롤러, 또는 상태 머신일 수도 있다. 또한, 프로세서는 컴퓨팅 디바이스의 조합, 예를 들어, DSP 와 마이크로프로세서의 조합, 복수의 마이크로프로세서, DSP 코어와 결합한 하나 이상의 마이크로프로세서, 또는 임의의 다른 그러한 구성으로서 구현될 수도 있다.

- [0261] 통상, 비디오 인코딩은 구조화된 GOP (group of picture) 에 대해 작용한다. GOP 는 보통 인트라 코딩된 프레임 (I-프레임) 에서 시작하고, 일련의 P (예측) 또는 B (양방향) 프레임으로 진행된다. 통상, I-프레임은 프레임을 표시하기 위한 모든 데이터를 저장할 수 있고, B-프레임은 이전 및 다음 프레임 내의 데이터 (예를 들어, 단지 이전 프레임으로부터 변경된 데이터를 포함하거나 다음 프레임 내의 데이터와는 상이함) 에 의존하며, P-프레임은 이전 프레임으로부터 변경된 데이터를 포함한다.
- [0262] 일반적인 이용 시, I-프레임은 인코딩된 비디오에서 P-프레임과 B-프레임에 산재된다. 크기 (예를 들어, 프레임을 인코딩하는데 이용되는 비트 수) 면에서, I-프레임은 통상 P-프레임보다 훨씬 더 크고, P-프레임은 또한 B-프레임보다 크다. 효율적인 인코딩, 송신 및 디코딩 처리를 위해, GOP 길이는 큰 I-프레임으로부터의 유효 손실을 줄이기 위해 충분히 길어야 하고, 인코더와 디코더 간의 불일치나 채널 손상을 해결하기 위해 충분히 짧아야 한다. 또한, 동일한 이유로, P 프레임 내의 매크로블록 (MB) 은 인트라코딩될 수 있다.
- [0263] 장면 전환 검출은, I-프레임을 일정한 간격으로 삽입하는 대신에, 적당한 GOP 길이를 결정하고, GOP 길이에 기초하여 I-프레임을 삽입하도록 비디오 인코더에서 이용될 수 있다. 실제 스트리밍 비디오 시스템에서, 통신 채널은 통상 비트 에러나 패킷 손실에 의해 손상된다. I 프레임 또는 I 매크로블록을 어디에 배치하는지에 따라, 디코딩된 비디오 품질과 시청 경험을 상당한 영향을 줄 수도 있다. 일 인코딩 방식은 연관된 이전 영상 또는 영상 부분으로부터의 상당한 변화를 갖는 영상 또는 영상 부분에 대해 인트라 코딩된 프레임을 이용하는 것이다. 보통, 이들 영역은 모션 추정을 이용하여 효과적·효율적으로 예측될 수 없고, 그러한 영역이 프레임 간 코딩 기술 (예를 들어, B-프레임 및 P-프레임을 이용하는 인코딩) 에서 제외되면 더 효율적으로 인코딩을 행할 수 있다. 채널 손상의 환경에서, 이들 영역은 에러 전파를 받을 가능성이 있는데, 이러한 에러 전파는 프레임 내 인코딩에 의해 감소 또는 제거될 수 있다 (또는 거의 감소 또는 제거될 수 있음).
- [0264] GOP 비디오의 일부는 2 개 이상의 카테고리 분류될 수 있는데, 각각의 영역은 특정 구현예에 따라 다를 수도 있는 상이한 프레임 내 인코딩 기준을 가질 수 있다. 일 예로서, 비디오는 3 개의 카테고리로 분류될 수 있다: 갑작스런 장면 전환, 크로스-페이딩 (cross-fading) 과 기타 느린 장면 전환, 및 카메라 플래시라이트. 갑작스런 장면 전환은 통상 카메라 조작에 의해 일어나는 이전 프레임과는 상당히 상이한 프레임을 포함한다. 이들 프레임의 콘텐츠가 이전 프레임의 콘텐츠와는 상이하기 때문에, 갑작스런 장면 전환은 I 프레임으로서 인코딩되어야 한다. 크로스-페이딩 및 기타 느린 장면 전환은 통상 카메라 샷의 컴퓨터 처리에 의해 일어나는 느린 장면 전환을 포함한다. 2 개의 상이한 장면의 점진 블렌딩은 인간 눈에는 더 만족스럽게 보일 수도 있지만, 비디오 코딩하기가 어렵다. 모션 보상은 이들 프레임의 비트레이트를 효과적으로 줄일 수 없고, 이들 프레임에 대해 더 많은 인트라 MB 가 갱신될 수 있다.
- [0265] 카메라 플래시라이트, 또는 카메라 플래시 이벤트는, 프레임의 콘텐츠가 카메라 플래시를 포함하는 경우에, 일어난다. 그러한 플래시는 지속기간이 비교적 짧고 (예를 들어, 1 프레임), 매우 밝아서, 플래시를 나타내는 프레임 내의 화소는 인접 프레임 상의 대응하는 영역에 비해 매우 높은 휘도를 나타내게 된다. 카메라 플래시라이트는 영상의 휘도를 갑작스럽고 빠르게 변화시킨다. 통상, 카메라 플래시라이트의 지속기간은 HVS (human vision system) 의 시간 마스크 지속기간보다 짧은데, 이는 통상 44 ms 가 되도록 정의된다. 인간 눈이 이들 짧은 밝기 버스트의 품질에 민감하지 않으므로, 대강 인코딩되어도 된다. 모션 보상을 이용하여 플래시라이트 프레임을 효과적으로 처리할 수 없으며 장래 프레임에 대해 열악한 예측 후보이므로, 이들 프레임의 대강 인코딩은 장래 프레임의 인코딩 효율을 줄이지 않는다. 플래시라이트로서 분류된 장면은 "인위적인" 고휘도로 인해 다른 프레임을 예측하는데 이용되지 않아야 하고, 동일한 이유로 이들 프레임을 예측하는데 다른 프레임을 효과적으로 이용할 수 없다. 일단 식별되면, 비교적 많은 양의 처리를 필요로 할 수 있기 때문에, 이들 프레임을 제거할 수 있다. 한 가지 옵션은 카메라 플래시라이트 프레임을 제거하고 그 대신에 DC 계수를 인코딩하는 것이다; 그러한 솔루션은 간단하고, 계산이 빠르며 많은 비트를 절약한다.
- [0266] 상기 프레임들 중 임의의 프레임이 검출되면, 샷 이벤트가 선언된다. 샷 검출은 인코딩 품질을 개선하는데 있어서 유용할 뿐만 아니라, 비디오 콘텐츠 검색 및 인텍싱의 식별에 있어서도 유용할 수 있다. 이하, 샷 검출 프로세스의 일 양태를 설명한다.
- [0267] 도 30 은 몇몇 양태에서 비디오 프레임 내의 샷 검출에 기초하여 비디오를 인코딩하는데 이용될 수 있는 GOP 에 대한 작용하는 프로세스 (3000) 를 도시하는 것으로서, 프로세스 (3000) 의 부분들 (또는 하위 프로세스) 는 도 30 내지 도 40 을 참조하여 설명 및 도시되어 있다. 프로세서 (2831) 는 프로세스 (3000) 를 포함하도록 구성될 수 있다. 프로세스 (3000) 가 시작한 다음에, 블록 3042 로 진행하여, 비디오 프레임에 대해 메트릭 (정보) 을 획득하는데, 이 메트릭은 인접 프레임 간의 차이를 나타내는 정보를 포함한다. 그 메트릭은 그

다음에 샷 분류에 이용될 수 있는 인접 프레임 간에 일어난 전환을 결정하기 위한 휘도 기반 정보와 양방향 모션 정보를 포함한다. 그러한 메트릭은 다른 디바이스나 프로세스로부터 획득되거나, 예를 들어, 프로세서 (2831) 에 의해 계산될 수 있다. 이하, 도 31 의 프로세스 A 를 참조하여 예시적인 메트릭 생성의 실시형태를 설명한다.

[0268] 그 다음에, 프로세스 (3000) 는 블록 3044 로 진행하여, 메트릭에 기초하여 비디오의 샷 전환을 결정한다. 비디오 프레임은 어떤 타입의 샷 (예를 들어, 갑작스런 장면 전환, 느린 장면 전환, 또는 고휘도 값 (카메라 플래시) 을 포함하는 장면) 이 프레임에 포함되어 있는지에 대한 2 개 이상의 카테고리로 분류될 수 있다. 일정 구현예의 인코딩은 다른 카테고리를 필요로 할 수도 있다. 샷 분류의 예시적인 예는 각각 도 32 의 프로세스 B 를 참조하여 설명되며, 도 34 내지 도 36 의 프로세스 D, E, 및 F 를 참조하여 더 상세히 설명된다.

[0269] 일단 프레임이 분류되면, 프로세스 (3000) 는 블록 3046 으로 진행하여, 샷 분류 결과를 이용하여 프레임을 인코딩하거나 인코딩을 위해 지정할 수 있다. 그러한 결과는 인트라 코딩된 프레임으로 프레임을 인코딩할지 예측 프레임 (예를 들어, P-프레임 또는 B-프레임) 으로 인코딩할지 여부에 영향을 줄 수 있다. 도 33 의 프로세스 C 는 샷 결과를 이용하는 인코딩 방식의 일 예를 도시한다.

[0270] 도 31 은 비디오의 메트릭을 획득하는 프로세스의 일 예를 도시한다. 도 31 은 도 30 의 블록 3042 에서 일어나는 일정 단계를 도시한다. 계속 도 31 을 참조하면, 블록 3152 에서, 프로세스 A 는 비디오의 양방향 모션 추정 및 보상 정보를 획득하거나 결정한다. 도 28 의 모션 보상기 (2832) 는 프레임에 대한 양방향 모션 추정을 수행하고, 다음 샷 분류에 이용될 수 있는 모션 보상 정보를 결정하도록 구성될 수 있다. 그 다음에, 프로세스 A 는 블록 3154 로 진행하여, 현재 또는 선택된 프레임 및 하나 이상의 인접 프레임에 대한 휘도 차이 히스토그램을 포함한 휘도 정보를 생성한다. 그 다음에, 끝으로, 프로세스 A 는 블록 3156 으로 진행하여, 프레임에 포함된 샷을 나타내는 메트릭을 계산한다. 그러한 일 메트릭은 식 (4) 및 식 (10) 의 2 개의 예에 도시되어 있는 프레임 차이 메트릭이다. 이하, 모션 정보, 휘도 정보 및 프레임 차이 메트릭을 결정하는 예시적인 예를 설명한다.

[0271] **모션 보상**

[0272] 양방향 모션 추정/보상을 수행하기 위해, 비디오 시퀀스는 프레임들 중 2 개의 프레임, 즉 가장 인접한 이웃 프레임 (과거의 일 프레임과 장래의 일 프레임) 내의 블록과 현재 프레임의 모든 8×8 블록을 일치시키는 양방향 모션 보상기로 전처리될 수 있다. 모션 보상기는 모든 블록마다 모션 벡터와 차이 메트릭을 생성한다.

도 37 은 이러한 개념을 나타내는 것으로서, 현재 프레임 C 의 화소를 이전 프레임 P 및 장래 프레임 (또는 다음) 프레임 N 과 일치시키는 일 예를 도시하는 개념을 나타낸 것으로서, 일치된 화소에 대한 모션 벡터 (이전 모션 벡터 MV_P 및 장래 모션 벡터 MV_N) 를 나타낸다. 이하, 양방향 모션 벡터 생성 및 관련 인코딩의 예시적인 양태의 간단한 설명이 계속된다.

[0273] 도 40 은 예를 들어, MPEG-4 에 있어서 모션 벡터 결정 프로세스와 예측 프레임 인코딩의 일 예를 도시한다. 도 40 에 도시된 프로세스는 도 31 의 블록 3152 에서 일어날 수 있는 예시적인 프로세스를 더 상세히 도시한다. 도 40 에서, 현재 영상 (4034) 은 5×5 매크로블록으로 이루어지고, 본 예에서 매크로블록의 수는 임의적이다. 매크로블록은 16×16 화소로 이루어진다. 화소는 8 비트 휘도 값 (Y) 과 2 개의 8 비트 색차 값 (Cr 및 Cb) 으로 정의될 수 있다.

[0274] MPEG 에 있어서, Y, Cr 및 Cb 성분은 4:2:0 포맷으로 저장될 수 있고, Cr 및 Cb 성분은 X 및 Y 방향으로 2 씩 다운샘플링된다. 따라서, 각각의 매크로블록은 256 개의 Y 성분, 64 개의 Cr 성분 및 64 개의 Cb 성분으로 이루어질 것이다. 현재 영상 (4034) 의 매크로블록 (4036) 은 현재 영상 (4034) 과는 상이한 시점에서 기준 영상 (4032) 으로부터 예측된다. 인코딩되는 현재 매크로블록 (4036) 에 대한 Y, Cr 및 Cb 값의 관점에서 가장 가까운 최적의 매칭 매크로블록 (4038) 을 찾기 위해, 기준 영상 (4032) 을 검색한다. 기준 영상 (4032) 내에 있는 최적의 매칭 매크로블록 (4038) 의 위치는 모션 벡터 (4040) 에서 인코딩된다. 기준 영상 (4032) 은, 현재 영상 (4034) 의 구성 전에 디코더가 복원할 I-프레임 또는 P-프레임일 수 있다. 최적의 매칭 매크로블록 (4038) 을 현재 매크로블록 (4036) 에서 빼면 (Y, Cr 및 Cb 성분 각각에 대한 차이를 계산), 레지듀얼 에러 (4042) 가 얻어진다. 레지듀얼 에러 (4042) 를 2D 이산 코사인 변환 (DCT) 으로 인코딩한 다음에, 양자화 (4046) 한다. 양자화 (4046) 를 수행하여, 예를 들어, 저주파수 계수에 더 많은 비트를 할당하면서 고주파수 계수에 더 적은 비트를 할당함으로써 공간 압축을 제공할 수 있다. 정보를 식별하는 기준 영상 (4034) 및 모션 벡터 (4040) 와 함께 레지듀얼 에러 (4042) 의 양자화된 계수는 현재 매크로블록 (4036) 을 나타내는 인코딩된 정보이다. 인코딩된 정보는 추후 이용을 위해 메모리에 저장되거나, 예를 들어, 에러 정

정이나 이미지 개선 (image enhancement) 을 위해 처리되거나, 또는 네트워크 (140) 를 통해 송신될 수 있다.

[0275] 인코딩된 모션 벡터 (4040) 와 함께 레지듀얼 에러 (4042) 의 인코딩된 양자화된 계수를 이용하여, 후속하는 모션 추정 및 보상을 위한 기준 프레임의 일부로서 이용되는 인코더에서 현재 매크로블록 (4036) 을 복원할 수 있다. 인코더는 이러한 P-프레임 복원을 위해 디코더의 절차를 에뮬레이션할 수 있다. 디코더의 에뮬레이션으로 인해, 인코더와 디코더 모두는 동일한 기준 영상과 함께 작업하게 된다. 이때, 추후 인터-코딩을 위해 인코더에서 행해지는지 디코더에서 행해지는지 여부에 관계없이 복원 프로세스가 제공된다. P-프레임의 복원은 기준 프레임 (또는 참조되고 있는 영상이나 프레임의 일부) 이 복원된 후에 시작할 수 있다. 인코딩된 양자화된 계수는 역양자화 (4050) 된 다음에, 2D 역 DCT, 또는 IDCT (4052) 를 수행하여, 디코딩되거나 복원된 레지듀얼 에러 (4054) 를 얻는다. 인코딩된 모션 벡터 (4040) 를 디코딩하고, 이를 이용하여 이미 복원된 기준 영상 (4032) 에서 이미 복원된 최적의 매칭 매크로블록 (4056) 을 찾는다. 그 다음에, 복원된 레지듀얼 에러 (4054) 를 복원된 최적의 매칭 매크로블록 (4056) 에 더하여, 복원된 매크로블록 (4058) 을 형성한다. 복원된 매크로블록 (4058) 은 메모리에 저장되거나, 다른 복원된 매크로블록과 함께 하나의 영상으로 또는 개별적으로 표시되거나, 또는 이미지 개선을 위해 더 처리될 수 있다.

[0276] B-프레임 (또는 양방향 예측으로 코딩된 임의의 섹션) 을 이용하는 인코딩은 현재 영상 내의 영역과 이전 영상 내의 최적의 매칭 예측 영역과 후속 영상 내의 최적의 매칭 예측 영역 간의 시간 리던던시를 활용할 수 있다. 후속하는 최적의 매칭 예측 영역과 이전에 있는 최적의 매칭 예측 영역을 결합하여, 결합된 양방향 예측 영역을 형성한다. 현재 영상 영역과 최적의 매칭 결합된 양방향 예측 영역 간의 차이가 레지듀얼 에러 (또는 예측 에러) 이다. 후속하는 기준 영상 내의 최적의 매칭 예측 영역 및 이전에 있는 기준 영상 내의 최적의 매칭 예측 영역의 위치는 2 개의 모션 벡터로 인코딩될 수 있다.

[0277] **휘도 히스토그램 차이**

[0278] 모션 보상은 모든 블록에 대해 차이 메트릭을 생성할 수 있다. 그 차이 메트릭은 SSD (sum of square difference) 또는 SAD (sum of absolute difference) 일 수 있다. 본원에서, 일반성을 잃지 않으면서, SAD 는 일 예로서 이용된다.

[0279] 모든 프레임에 대해, SAD 비율은 다음과 같이 계산된다.

$$y = \frac{\epsilon + SAD_P}{\epsilon + SAD_N} \tag{48}$$

[0280] 여기서, SAD_P 와 SAD_N 은 각각 순방향 및 역방향 차이 메트릭의 절대 차이 합이다. "0 으로 나누는" 에러를 방지하기 위해 분모는 작은 양의 수 ϵ 을 포함한다는 것에 주목하자. 또한, 분자는 분모에서 1 의 영향을 뺀것을 위해 ϵ 을 포함한다. 예를 들어, 이전 프레임, 현재 프레임, 및 다음 프레임이 동일하면, 모션 검색은 $SAD_P = SAD_N = 0$ 을 생성하게 된다. 이 경우에, 상기 계산은 0 또는 무한대 대신에 $y = 1$ 을 생성한다.

[0282] 휘도 히스토그램은 모든 프레임에 대해 계산될 수 있다. 통상, 멀티미디어 이미지는 8 비트의 휘도 깊이 (예를 들어, "빈 (bin)" 개수) 를 갖는다. 몇몇 양태에 따른 휘도 히스토그램을 계산하는데 있어서 이용되는 휘도 깊이를 16 으로 설정하여, 히스토그램을 얻을 수 있다. 다른 양태에서는, 처리되는 데이터 타입, 가용 계산 전력, 또는 기타 소정의 기준에 따라 다를 수도 있는 적절한 숫자로 휘도 깊이를 설정할 수 있다. 몇몇 양태에서, 휘도 깊이는 데이터 콘텐츠와 같은 계산되거나 수신된 메트릭에 기초하여 동적으로 설정될 수 있다.

[0283] 식 (49) 는 휘도 히스토그램 차이 (λ) 를 계산하는 일 예를 도시한다.

$$\lambda = \frac{\sum_{i=1}^{16} |N_{Pi} - N_{Ci}|}{N} \tag{49}$$

[0284] 여기서, N_{Pi} 는 이전 프레임에 대한 i 번째 빈에서의 블록 수이고, N_{Ci} 는 현재 프레임에 대한 i 번째 빈에서의

블록 수이며, N 은 일 프레임 내의 전체 블록 수이다. 이전 프레임과 현재 프레임의 휘도 히스토그램 차이가 완전히 다르면 (disjoint), $\lambda = 2$ 이다.

[0286] 도 5 의 블록 56 을 참조하여 설명되는 프레임 차이 메트릭 D 는 식 (50) 에 도시된 것과 같이 계산될 수 있다.

$$D = \frac{\gamma_C}{\gamma_P} + A\lambda(2\lambda + 1) \tag{50}$$

[0287]

$$\gamma_C = \frac{\varepsilon + SAD_P}{\varepsilon + SAD_N}, \text{ 및}$$

[0288]

여기서, A 는 애플리케이션에 의해 선택된 상수이고,

$$\gamma_P = \frac{\varepsilon + SAD_{PP}}{\varepsilon + SAD_C} \text{ 이다.}$$

[0289]

도 32 는 비디오에 대해 획득되거나 결정된 메트릭을 이용하여 샷 (또는 장면) 전환의 3 개의 카테고리를 결정하는 프로세스 B 의 일 예를 도시한다. 도 32 는 도 30 의 블록 3044 의 일 양태에서 일어나는 일정 단계를 도시한다. 도 32 를 다시 참조하면, 블록 3262 에서, 프로세스 B 는 우선 프레임이 갑작스런 장면 전환으로 지정되도록 하는 기준을 만족하는지 여부를 결정한다. 도 34 의 프로세스 D 는 이러한 결정의 일 예를 도시한다. 그 다음에, 프로세스 B 는 블록 3264 로 진행하여, 프레임이 느리게 전환하는 장면의 일부인지 여부를 결정한다. 도 35 의 프로세스 C 는 느리게 전환하는 장면을 결정하는 일 예를 도시한다. 끝으로, 블록 3366 에서, 프로세스 B 는 프레임이 카메라 플래시를 포함하는지 여부, 즉, 이전 프레임과는 상이한 큰 휘도 값을 포함하는지 여부를 결정한다. 도 36 의 프로세스 F 는 카메라 플래시를 포함하는 프레임을 결정하는 일 예를 도시한다. 이하, 이들 프로세스의 예시적인 예를 설명한다.

[0290]

갑작스런 장면 전환

[0291]

도 34 는 갑작스런 장면 전환을 결정하는 프로세스를 도시한 흐름도이다. 도 34 는 도 32 의 블록 3262 의 몇몇 양태에서 일어날 수 있는 일정 단계를 더 상세히 설명한다. 블록 3482 에서, 프레임 차이 메트릭 D 가 식 (51) 에 도시된 기준을 만족하는지 여부를 확인한다.

$$D = \frac{\gamma_C}{\gamma_P} + A\lambda(2\lambda + 1) \geq T_1 \tag{51}$$

[0292]

[0293] 여기서, A 는 애플리케이션에 의해 선택된 상수이고, T_1 은 임계값이다. 그 기준을 만족하면, 블록 3484 에서, 프로세스 D 는 그 프레임을 갑작스런 장면 전환으로서 지정하는데, 본 예에서는, 추가적인 샷 분류를 필요로 하지 않는다.

[0294]

일 예에서, 시뮬레이션은, A = 1, 및 $T_1 = 5$ 로 설정하면, 양호한 검출 성능을 달성함을 보여준다. 현재 프

레이미 갑작스런 장면 전환 프레임이면, γ_C 는 커야 하고, γ_P 는 작아야 한다. 비율 $\frac{\gamma_C}{\gamma_P}$ 를 γ_C 단독 대신에 이용하여, 그 메트릭을 콘텍스트의 액티비티 레벨로 정규화할 수 있다.

[0295]

상기 기준은 휘도 히스토그램 차이 람다 (λ) 를 비선형 방식으로 이용한다는 것에 주목해야 한다. 도 39 는 $\lambda * (2\lambda + 1)$ 이 블록 함수임을 나타낸다. λ 가 작으면 (예를 들어, 0 에 가까우면), 프리엠퍼시스 (preemphasis) 가 거의 없다. λ 가 커지면, 그 함수에 의해 더 많은 엠퍼시스가 행해진다. 1.4 보다 큰 임의의 λ 에 대해 이러한 프리엠퍼시스를 이용하여, 임계값 T_1 이 5 로 설정되면, 갑작스런 장면 전환이 검출된다.

[0296]

크로스 페이딩 및 느린 장면 전환

[0297]

또한, 도 35 는 도 32 의 블록 3264 에서 일어날 수 있는 몇몇 양태의 추가 상세를 도시한다. 도 35 를 참조하면, 블록 3592 에서, 프로세스 E 는 프레임이 느린 장면 전환을 나타내는 일련의 프레임의 일부인지 여부를

결정한다. 식 (52) 에 도시된 바와 같이, 프레임 차이 메트릭 D 가 제 1 임계값 T_1 보다 작으며 제 2 임계값 T_2 이상인 경우에, 프로세스 E 는 현재 프레임이 크로스 페이딩 또는 느린 장면 전환인 것으로 결정한다.

[0298] 일정 개수의 연속 프레임에 대해,

$$T_2 \leq D < T_1 \tag{52}$$

[0299]

여기서, T_1 은 상기 이용되는 동일 임계값이며, T_2 는 다른 임계값이다. 통상, 구현예의 차이가 가능하기 때문에, T_1 및 T_2 의 정확한 값은 정규 실험에 의해 결정된다. 기준을 만족하면, 블록 3594 에서, 프로세스 E 는 선택된 프레임의 종료를 위해 프레임을 느리게 변하는 장면 샷 분류의 일부로서 분류한다.

[0301] **카메라 플래시라이트 이벤트**

[0302] 도 36 에 도시된 프로세스 F 는 현재 프레임이 카메라 플래시라이트를 포함하는지 여부를 결정할 수 있는 프로세스의 일 예이다. 본 예시적인 양태의 카메라에 있어서, 휘도 히스토그램 통계치는 현재 프레임이 카메라 플래시라이트를 포함하는지 여부를 결정하는데 이용된다. 프로세스 F 는 블록 3602 에 도시된 것과 같이, 우선 현재 프레임의 휘도가 이전 프레임 및 다음 프레임의 휘도보다 큰지 여부를 결정함으로써, 선택된 프레임에서 카메라 플래시 이벤트를 결정한다. 현재 프레임의 휘도가 이전 프레임 및 다음 프레임의 휘도보다 크지 경우에, 그 프레임은 카메라 플래시 이벤트가 아니지만, 현재 프레임의 휘도가 큰 경우에, 그 프레임은 카메라 플래시 이벤트일 수도 있다. 블록 3604 에서, 프로세스 F 는 역방향 차이 메트릭이 임계값 T_3 보다 큰지 여부를 결정하며, 순방향 차이 메트릭이 임계값 T_4 보다 큰지 여부를 결정하고, 이들 양쪽 조건이 만족되면, 블록 3606 에서, 프로세스 F 는 현재 프레임을 카메라 플래시라이트를 갖는 것으로서 분류한다. 일 예에서, 블록 3602 에서, 식 (53) 및 식 (54) 에 도시된 것과 같이, 프로세스 F 는 현재 프레임의 평균 휘도에서 이전 프레임의 평균 휘도를 뺀 값이 임계값 T_3 이상인지 여부를 결정하고, 프로세스 F 는 현재 프레임의 평균 휘도에서 다음 프레임의 평균 휘도를 뺀 값이 임계값 T_3 이상인지 여부를 결정한다.

$$\bar{Y}_C - \bar{Y}_P \geq T_3 \tag{53}$$

$$\bar{Y}_C - \bar{Y}_N \geq T_3 \tag{54}$$

[0303]

[0304] 기준이 만족되지 않으면, 현재 프레임을 카메라 플래시라이트를 포함하는 것으로서 분류되지 않고, 프로세스 F 는 리턴한다. 기준이 만족되면, 프로세스 F 는 블록 3604 로 진행하여, 아래의 식 (55) 및 식 (56) 에 도시된 것과 같이, 역방향 차이 메트릭 SAD_P 및 순방향 차이 메트릭 SAD_N 이 일정 임계값 T_4 보다 큰지 여부를 결정한다.

$$SAD_P \geq T_4 \tag{55}$$

$$SAD_N \geq T_4 \tag{56}$$

[0305]

[0306] 여기서, \bar{Y}_C 는 현재 프레임의 평균 휘도이고, \bar{Y}_P 는 이전 프레임의 평균 휘도이고, \bar{Y}_N 은 다음 프레임의 평균 휘도이며, SAD_P 및 SAD_N 은 현재 프레임과 연관된 순방향 및 역방향 차이 메트릭이다. 기준이 만족되지 않으면, 프로세스 F 는 리턴한다.

[0307] 통상, 설명된 프로세스의 구현으로 인해 임계값을 비롯한 동작 파라미터의 차이가 발생할 때에, T_3 의 값이 정규 실험에 의해 결정된다. SAD 값은 카메라 플래시가 통상 일 프레임만을 차지하기 때문에 그 결정에 포함되고, 휘도 차이로 인해, 이러한 프레임은 순방향과 역방향 모두로부터의 모션 보상을 이용하여 양호하게 예측될 수 없다.

[0308] 몇몇 양태에서, 임계값 T_1 , T_2 , T_3 및 T_4 중 하나 이상이 미리 결정되고, 그러한 값은 인코딩 디바이스 내의 샷 분류기 내에 통합된다. 통상, 이들 임계값은 샷 검출의 특정 구현의 시험을 통해 선택된다. 몇몇 양태

에서, 임계값 T_1, T_2, T_3 및 T_4 중 하나 이상은 샷 분류기에 공급된 정보 (예를 들어, 메타데이터) 에 기초하여 또는 샷 분류기 자체에 의해 계산된 정보에 기초하여 (예를 들어, 동적으로) 처리 동안에 설정될 수 있다.

[0309] 이하, 선택된 프레임의 샷 분류에 기초하여, 비디오를 인코딩하거나, 비디오에 대한 인코딩 파라미터를 결정하는 프로세스 C 를 도시한 도 33 을 참조하자. 블록 3370 에서, 프로세스 C 는 선택된 프레임이 갑작스런 장면 전환으로서 분류되었는지 여부를 결정한다. 선택된 프레임이 갑작스런 장면 전환으로서 분류된 경우에, 블록 3371 에서, 현재 프레임은 갑작스런 장면 전환으로서 분류되고, 그 프레임을 I-프레임으로서 인코딩할 수 있으며, GOP 경계를 결정할 수 있다. 현재 프레임이 갑작스런 장면 전환으로서 분류되지 않은 경우에, 프로세스 C 는 블록 3372 로 진행하고, 현재 프레임이 느리게 변하는 장면의 일부로서 분류되는 경우에, 블록 3373 에서, 현재 프레임, 및 느리게 변하는 장면의 다른 프레임을 예측 프레임 (예를 들어, P-프레임 또는 B-프레임) 으로서 인코딩할 수 있다. 그 다음에, 프로세스 C 는 블록 3374 로 진행하여, 현재 프레임이 카메라 플래시라이트를 포함하는 플래시라이트 장면으로서 분류되었는지 여부를 확인한다. 현재 프레임이 카메라 플래시라이트를 포함하는 플래시라이트 장면으로서 분류된 경우에, 블록 3375 에서, 특별한 처리, 예를 들어, 이전 프레임의 제거, 복제, 또는 그 프레임에 대한 특정 계수의 인코딩을 위해 프레임을 식별할 수 있다. 현재 프레임이 카메라 플래시라이트를 포함하는 플래시라이트 장면으로서 분류되지 않은 경우에, 현재 프레임에 대한 어떤 분류도 이루어지지 않고, 선택된 프레임을 다른 기준에 따라 인코딩할 수 있는데, I-프레임으로서 인코딩하거나 드롭할 수 있다. 프로세스 C 는 인코더에 구현될 수 있다.

[0310] 상술한 양태에 있어서, 압축될 프레임과 그 인접한 2 개 프레임 간의 차이 양은 프레임 차이 메트릭 D 에 의해 표시된다. 일방향 휘도 변화의 상당한 양이 검출되면, 그 프레임 내의 크로스-페이드 효과를 나타낸다. 크로스-페이드가 더 현저해지면, B 프레임을 이용하여 더 많은 이득이 달성될 수도 있다. 몇몇 양태에서, 수정된 프레임 차이 메트릭은 아래 식 (57) 에 도시된 것과 같이 이용된다.

$$D_1 = \begin{cases} \left(1 - \alpha + 2\alpha \frac{|d_p - d_N|}{d_p + d_N}\right) \times D, & \text{if } Y_p - \Delta \geq Y_c \geq Y_N + \Delta \text{ or } Y_p + \Delta \leq Y_c \leq Y_N - \Delta, \\ D, & \text{otherwise,} \end{cases} \tag{57}$$

[0311] 여기서, $d_p = |Y_c - Y_p|$ 및 $d_N = |Y_c - Y_N|$ 은 각각 현재 프레임과 이전 프레임 간의 루마 차이이고, 현재 프레임과 다음 프레임 간의 루마 차이이고, Δ 는 정규 실험에서 결정될 수 있는 상수 (구현예에 따라 다를 수 있음) 를 나타내고, α 는 0 과 1 사이의 값을 갖는 가중 변수이다.

[0313] 루마 변화의 일관된 경향이 관찰되며 그 변화 세기가 충분히 크면, 수정된 프레임 차이 메트릭 D_1 은 원래 프레임 차이 메트릭 D 과는 상이할 뿐이다. D_1 은 D 이하이다. 루마의 변화가 일정하면 ($d_p = d_N$), 수정된 프레임 차이 메트릭 D_1 은 $(1 - \alpha)$ 의 최저 비율을 갖는 원래 프레임 차이 메트릭 D 보다 낮다.

[0314] 아래의 테이블 1 은 갑작스런 장면 전환 검출을 부가함으로써 성능 개선을 나타낸다. NSC (non-scene-change) 및 SC (scene-change) 경우 모두에 있어서 I-프레임의 총 수는 대략 같다. NSC 경우에 있어서, I-프레임은 전체 시퀀스 사이에서 균일하게 분포되지만, SC 경우에 있어서, I-프레임은 갑작스런 장면 전환 프레임에 할당될 뿐이다.

[0315] 통상 0.2 ~ 0.3 dB 의 개선은 PSNR 단위로 달성될 수 있음을 알 수 있다. 시뮬레이션 결과는 샷 검출기가 상술한 샷 이벤트를 결정하는데 있어서 매우 정확하다는 것을 나타낸다. 정상적인 크로스-페이드 효과를 갖는 5 개의 클립의 시뮬레이션은 $\Delta = 5.5$ 및 $\alpha = 0.4$ 에서, 0.226031 dB 의 PSNR 이득이 동일 비트레이트로 달성됨을 나타낸다.

시퀀스/메트릭	비트레이트(kbps)	평균 QP	PSNR (dB)
애니메이션 NSC	226.2403	31.1696	35.6426
애니메이션 SC	232.8023	29.8171	36.4513
음악 NSC	246.6394	32.8524	35.9337
음악 SC	250.0994	32.3209	36.1202
헤드라인 NSC	216.9493	29.8304	38.9804

헤드라인 뉴스 SC	220.2512	28.9011	39.3151
농구 NSC	256.8726	33.1429	33.5262
농구 SC	254.9242	32.4341	33.8635

[0317] **테이블 1 : 갑작스런 장면 전환 검출의 시뮬레이션 결과**

[0318] **적용 GOP 구조**

[0319] 이하, 적용 GOP 구조 동작의 예시적인 실시형태를 설명한다. 그러한 동작은 도 4의 GOP 분할기 (412)에 포함될 수 있다. 구형 비디오 압축 표준인 MPEG2는, GOP가 일정한 구조를 가질 것을 요구하지 않는다 (그러나, GOP가 일정한 구조를 가질 것이 부과될 수 있음). MPEG2 시퀀스는, 이전 영상을 참조함 없이 인코딩된 I 프레임으로 항상 시작한다. 통상, MPEG2 GOP 포맷은 I 프레임 다음에 오는 P 또는 예측 영상의 GOP에서 간격을 고정함으로써 인코더에서 미리 배열된다. P 프레임은 이전 I 또는 P 영상으로부터 부분적으로 예측된 영상을 말한다. 시작하는 I 프레임과 후속하는 P 프레임 간의 프레임은 B 프레임으로서 인코딩된다.

"B" 프레임 (B는 양방향을 나타냄)은 이전 및 다음 I 또는 P 영상을 개별적으로 또는 동시에 참조로서 이용할 수 있다. 평균적으로 I-프레임을 인코딩하는데 이용되는 비트 수는 P-프레임을 인코딩하는데 이용되는 비트 수를 초과하고, 이와 마찬가지로, 평균적으로 P-프레임을 인코딩하는데 이용되는 비트 수는 B-프레임을 인코딩하는데 이용되는 비트 수를 초과한다. 스킵된 프레임은, 이용되더라도, 그 표현을 위해 어떤 비트도 이용하지 않을 수도 있다.

[0320] 보다 최신의 압축 알고리즘에서 P-프레임 및 B-프레임을 이용하는 일 이점으로는, 프레임을 스킵함으로써 비디오 송신 크기를 줄일 수 있다는 것이다. 시간 리턴턴시가 높으면, 예를 들어, 영상 간에 변화가 거의 없으면, P, B 또는 스킵된 영상을 이용하여 비디오 스트림을 효율적으로 표현하는데, 그 이유는, 먼저 디코딩된 I 또는 P 영상을 추후에 기준으로 이용하여 다른 P 또는 B 영상을 디코딩하기 때문이다.

[0321] GOP (group of picture) 분할기는 프레임을 적용 인코딩하여 시간 리턴턴시를 최소화한다. 프레임 간의 차이가 정량화되고, 정량화된 차이에 대해 적합한 테스트가 수행된 후에 I, P, B 또는 스킵된 프레임에 의해 영상을 표현하려는 결정이 자동으로 행해진다. GOP 분할기에서의 처리는 전처리 (202)의 다른 동작에 의해 도움을 받아서, 노이즈 제거를 위한 필터링을 제공한다.

[0322] 적용 인코딩 프로세스는 "고정된" 인코딩 프로세스에서 이용 가능하지 않은 이점을 갖는다. 고정된 프로세스는 콘텐츠에서 작은 변화가 일어난 가능성을 무시하지만, 적용 절차는 훨씬 더 많은 B 프레임이 각각의 I 프레임과 P 프레임 사이에 또는 2개의 P 프레임 사이에 삽입되는 것을 허용함으로써, 프레임 시퀀스를 충분히 표현하는데 이용되는 비트 수를 줄이게 된다. 한편, 예를 들어, 고정된 인코딩 프로세스에 있어서, 비디오 콘텐츠의 변화가 상당하면, 예측된 프레임과 기준 프레임 간의 차이가 너무 크기 때문에, P 프레임의 효율은 크게 감소한다. 이들 조건 하에서, 매칭하는 대상이 모션 검색 영역 밖으로 벗어날 수도 있고, 매칭하는 대상 간의 유사도가 카메라 앵글의 변화로 인한 왜곡때문에 감소한다. 적용 인코딩 프로세스는 P 프레임이 인코딩되어야 할 때를 최적으로 결정하는데 이용될 수도 있는 이점을 갖는다.

[0323] 본원에 개시된 시스템에 있어서, 상술한 조건의 타입은 자동으로 감지된다. 본원에 설명된 적용 인코딩 프로세스는 유연하며 이들 콘텐츠의 변화에 적응하도록 만들어진다. 적용 인코딩 프로세스는 프레임 차이 메트릭을 평가하고, 이 프레임 차이 메트릭은 동일한 부가적인 거리 특성을 갖는 프레임 간의 거리의 측정으로서 간주될 수 있다. 개념적으로는, 프레임 F₁, F₂ 및 F₃이 프레임 간 거리 d₁₂ 및 d₂₃을 가지면, F₁과 F₃ 간의 거리는 적어도 d₁₂ + d₂₃인 것으로 간주된다. 프레임 할당은 거리 유사 (distance-like) 메트릭 및 다른 측정에 기초하여 이루어진다.

[0324] GOP 분할기 (412)는 수신 시 프레임에 영상 타입을 할당함으로써 동작한다. 영상 타입은 각각의 블록을 코딩하는데 이용될 수도 있는 예측 방법을 나타낸다.

[0325] I-영상은 다른 영상을 참조함 없이 코딩된다. I-영상은 독립형이기 때문에 디코딩이 시작할 수 있는 데이터 스트림 내에 액세스 포인트를 제공한다. I 인코딩 타입은, 그 이전 프레임까지의 "거리"가 장면 전환 임계값을 초과하는 경우에 프레임에 할당된다.

[0326] P-영상은 모션 보상된 예측을 위해 이전 I 또는 P 영상을 이용할 수 있다. P-영상은 인코딩을 위한 기준으로서 예측되는 블록으로부터 옮겨질 수도 있는 이전 필드 또는 프레임 내의 블록을 이용한다. 고려 중인 블

록에서 기준 블록을 뺀 다음에, 통상, 공간 중복의 제거를 위해 이산 코사인 변환을 이용하여 나머지 블록을 인코딩한다. P 프레임이 되도록 할당된 마지막 프레임과 일 프레임 간의 "거리" 가 제 2 임계값 (제 2 임계값은 통상 제 1 임계값보다 작음) 을 초과하는 경우에, 그 프레임에 P 인코딩 타입이 할당된다.

[0327] 상술한 것과 같이, B-프레임 영상은 모션 보상을 위해 이전 및 다음 P-영상 또는 I-영상을 이용할 수 있다. B 영상 내의 블록은 순방향, 역방향 또는 양방향 예측될 수 있고, 또는 다른 프레임을 참조함 없이 인트라-코딩될 수 있다. H.264 에 있어서, 기준 블록은 다수의 프레임 중에서 32 개 블록의 선형 조합일 수 있다. 프레임이 I 또는 P 타입이 되도록 할당될 수 없으면, 그 바로 이전 프레임까지의 "거리" 가 제 3 임계값 (통상, 제 2 임계값보다 작음) 보다 큰 경우에, B 타입이 되도록 할당된다. 프레임이 인코딩된 B 타입이 되도록 할당될 수 없으면, "스킵 프레임" 상태에 할당된다. 이러한 프레임은 이전 프레임의 사실상 복제본이므로 스킵될 수 있다.

[0328] 인접 프레임 간의 차이를 디스플레이 순서로 정량화하는 메트릭을 평가하는 것은 GOP 분할기 (412) 에서 일어나는 이러한 처리의 제 1 부분이다. 이러한 메트릭은 상기 참조되는 거리이고, 그 메트릭을 이용하여, 모든 프레임은 그 적절한 타입을 찾기 위해 평가된다. 이와 같이, I 프레임과 인접한 P 프레임, 또는 2 개의 연속적인 P 프레임 간의 간격은 변할 수 있다. 메트릭의 계산은 블록 기반 모션 보상기를 이용하여 비디오 프레임을 처리함으로써 시작하고, 일 블록은 비디오 압축의 기본 단위로서, 16 × 16 화소로 보통 이루어지지만, 예를 들어 8 × 8, 4 × 4 및 8 × 16 과 같은 다른 블록 크기도 가능하다. 출력에 존재하는 2 개의 인터레이싱된 필드로 이루어진 프레임의 경우에, 모션 보상은 필드 단위로 이루어지고, 기준 블록에 대한 검색은 프레임보다는 필드에서 일어난다. 현재 프레임의 제 1 필드 내의 블록의 경우에, 순방향 기준 블록은 그 다음에 오는 프레임의 필드 내에서 발견되고, 또한 역방향 기준 블록은 현재 필드 바로 전에 있는 프레임의 필드 내에서 발견된다. 현재 블록은 보상된 필드 내에 모아진다. 그 프로세스는 프레임의 제 2 필드에서 계속된다. 2 개의 보상된 필드가 결합되어, 순방향 및 역방향 보상된 프레임을 형성한다.

[0329] 역 텔레시네 (406) 에서 생성된 프레임의 경우에, 단지 복원된 필름 프레임만이 생성되므로, 기준 블록에 대한 검색은 프레임 단위로만 이루어 질 수도 있다. 또한, 2 개의 기준 블록과 2 개의 차이 (순방향 및 역방향) 를 구하여, 순방향 및 역방향 보상된 프레임을 생성한다. 요컨대, 모션 보상기는 모든 블록에 대해 모션 벡터 및 차이 메트릭을 생성한다. 순방향 차이가 평가되는지 역방향 차이가 평가되는지 여부에 따라, 이전 필드나 프레임 또는 그 바로 다음에 오는 필드나 프레임에서, 현재 고려 중인 필드나 프레임 내의 블록과 그와 가장 잘 매칭하는 블록 사이에서 메트릭의 차이가 평가된다는 것에 주목하자. 단지 휘도 값만이 이 계산에 포함된다.

[0330] 이와 같이, 모션 보상 단계는 2 가지 세트의 차이를 생성한다. 이들은 시간상 현재 프레임의 바로 앞 프레임과 바로 이전 프레임에 있는 프레임들로부터 선택된 기준 블록 내의 휘도 값들과 현재 휘도 값들의 블록 사이에 존재한다. 각각의 순방향 및 각각의 역방향 차이의 절대 값은 블록 내의 각 화소에 대해 결정되고, 그 각각은 전체 프레임에 걸쳐 개별적으로 합산된다. 프레임을 포함하는 디인터레이싱된 NTSC 필드가 처리될 때에, 양쪽 필드는 2 개의 덧셈에 포함된다. 이러한 방법으로, SAD_p 및 SAD_N, 즉, 순방향 및 역방향 차이의 합산된 절대 값이 구해진다.

[0331] 모든 프레임에 대해, SAD 비율은 다음 관계식을 이용하여 계산된다.

$$\gamma = \frac{\epsilon + SAD_p}{\epsilon + SAD_N} \tag{58}$$

[0332] 여기서, SAD_p 및 SAD_N 은 각각 순방향 및 역방향 차이의 합산된 절대 값이다. 작은 양의 수 ϵ 가 분자에 더해져서, "0 으로 나누는" 에러를 방지한다. 유사한 항 ϵ 가 분모에 더해져서, 또한, SAD_p 또는 SAD_N 이 0 에 가까워지는 경우에, γ 의 감도를 줄인다.

[0334] 다른 양태에서, 그 차이는 SSD (sum of squared difference), SAD (sum of absolute difference), 또는 SATD 일 수 있고, 화소 값의 블록은 블록 요소 내의 차이를 취하기 전에 2 차원 이산 코사인 변환을 적용함으로써 변환된다. 그 합들이 액티브 비디오의 영역에 걸쳐 평가되지만, 다른 양태에서는 보다 작은 영역이 이용될 수도 있다.

[0335] 또한, 수신된 것과 같은 (모션 보상되지 않은) 모든 프레임의 휘도 히스토그램이 계산된다. 휘도 히스토그

램은, 이용 가능한 경우에 2 차원 이산 코사인 변환을 휘도 값 블록에 적용한 결과인 16×16 계수 어레이에서, DC 계수, 즉, (0,0) 계수에 대해 작용한다. 이에 대응하여, 16×16 블록 내의 256 개의 휘도 값의 평균 값이 휘도 히스토그램에서 이용될 수도 있다. 휘도 깊이가 8 비트인 이미지에 대해, bin의 수는 16 으로 설정된다. 다음 메트릭은 히스토그램 차이를 계산한다.

$$\lambda = \frac{1}{N} \sum_{i=1}^{16} |N_{P_i} - N_{C_i}| \tag{59}$$

[0336]

상기와 같이, N_{P_i} 는 i 번째 bin에서 이전 프레임으로부터의 블록 수이고, N_{C_i} 는 i 번째 bin에 속하는 현재 프레임으로부터의 블록 수이며, N 은 프레임 내의 전체 블록 수이다.

[0337]

이들 중간 결과를 모아서, 다음과 같이 현재 프레임 차이를 형성한다.

[0338]

$$M = \frac{Y_C}{Y_P} + \lambda(2\lambda + 1), \tag{60}$$

[0339]

여기서, Y_C 는 현재 프레임에 기초한 SAD 비율이고, Y_P 는 이전 프레임에 기초한 SAD 비율이다. 장면이 평활한 모션을 가지며 그 루마 히스토그램이 거의 변하지 않으면, $M \approx 1$ 이다. 현재 프레임이 갑작스런

[0340]

장면 전환을 표시하면, Y_C 는 커질 것이고, Y_P 는 작아야 한다. Y_C 단독 대신에, 비율 $\frac{Y_C}{Y_P}$ 를 이용하므로, 문맥의 액티비티 레벨로 메트릭이 정규화된다.

[0341]

도 40 의 데이터흐름 (4100) 은 프레임 차이 메트릭을 계산하는데 이용될 수도 있는 일정 컴포넌트를 도시한다. 전처리기 (4125) 는 NTSC 소스를 갖는 비디오의 경우에 인터레이싱된 필드를 전달하고, 비디오 소스가 양방향 모션 보상기 (4133) 에 대한 역 텔레시네의 결과인 경우에 필름 이미지의 프레임을 전달한다. 양방향 모션 보상기 (4133) 는 16×16 화소의 블록으로 분할하고 각각의 블록을 이전 프레임 필드의 정해진 영역 내의 모든 16×16 블록과 비교함으로써 필드 (또는 영화의 비디오 소스의 경우에는 프레임) 에 대해 작용한다. 최적의 매치를 제공하는 블록이 선택되어, 현재 블록에서 제거된다. 그 차이의 절대 값이 취해지고, 그 결과는 현재 블록을 포함하는 256 개의 화소에 걸쳐 더해진다. 이러한 절차가 필드의 모든 현재 블록에 대해 행해진 다음에 양쪽 필드에 대해 행해지면, 양 SAD_N, 즉, 역방향 차이 메트릭이 역방향 차이 모듈 (4137) 에 의해 계산된다. 유사한 절차가 순방향 차이 메트릭 (4136) 에 의해 수행될 수도 있다. 순방향 차이 모듈 (4136) 은 시간상 현재 프레임의 바로 앞에 있는 프레임을 기준 블록 소스로서 이용하여 SAD_P, 즉, 순방향 차이 메트릭을 전개한다. 동일한 추정 프로세스는, 복원된 필름 이미지를 이용하여 수행되더라도, 역 텔레시네에서 입력 프레임이 형성될 때에 일어난다. 프레임 차이 메트릭의 계산을 완료하는데 이용될 수 있는 히스토그램은 히스토그램 차이 모듈 (4141) 에 형성될 수도 있다. 각각의 16×16 블록은 그 휘도 값의 평균에 기초하여 bin에 할당된다. 이러한 정보는 모든 256 개의 화소 휘도 값을 일 블록 내에 함께 더하고, 필요에 따라 256 으로 정규화하고, 평균 값이 배치된 bin의 카운트를 증분함으로써 형성된다. 그 계산은 각각의 프리모션 보상된 프레임에 대해 한 번 행해지고, 현재 프레임에 대한 히스토그램은 새로운 현재 프레임이 도착할 때에 이전 프레임에 대한 히스토그램이 된다. 히스토그램 차이 모듈 (4141) 에서 블록 수로 2 개의 히스토그램을 차분 및 정규화하여 식 (59) 에 정의된 λ 를 형성한다. 이들 결과는 프레임 차이 결합기 (4143) 에서 결합되고, 이 프레임 차이 결합기 (4143) 는 히스토그램 차이 모듈 (4139), 순방향 및 역방향 차이 모듈 (4136 및 4136) 에서 구한 중간 결과를 이용하여 식 (60) 에 정의된 현재 프레임 차이를 평가한다.

[0342]

흐름도 (4100) 의 시스템 및 그 컴포넌트 또는 단계들은 하드웨어, 소프트웨어, 펌웨어, 미들웨어, 마이크로코드, 또는 그 임의의 조합에 의해 구현될 수 있다. 전처리기 (4125), 양방향 모션 보상기 (4133), 순방향 및 역방향 차이 메트릭 모듈 (4136 및 4137), 히스토그램 차이 모듈 (4141), 및 프레임 차이 메트릭 결합기 (4143) 를 비롯한, 흐름도 (4100) 의 각 기능 컴포넌트는 독립형 컴포넌트로서 구현되거나, 다른 디바이스의 컴포넌트 내에 하드웨어, 펌웨어, 미들웨어로서 통합되거나, 프로세서 상에서 실행되는 마이크로코드나 소프트웨어에 구현되거나, 또는 그 조합일 수도 있다. 소프트웨어, 펌웨어, 미들웨어 또는 마이크로코드에 구현 시, 원하는 작업을 수행하는 프로그램 코드 또는 코드 세그먼트는 저장 매체와 같은 머신 판독가능 매체에 저장될 수도 있다. 코드 세그먼트는 절차, 함수, 서브프로그램, 프로그램, 루틴, 서브루틴, 모듈, 소프트웨어 패키지, 클

래스, 또는 명령들, 데이터 구조들 또는 프로그램 문장들의 임의의 조합을 나타낼 수도 있다. 코드 세그먼트는 정보, 데이터, 인수 (argument), 파라미터 또는 메모리 콘텐츠를 전달하고/하거나 수신함으로써 다른 코드 세그먼트 또는 하드웨어 회로에 연결될 수도 있다.

[0343] 수신되어 처리된 데이터는 프로세서에 접속된 예를 들어, 칩 구성된 저장 매체 (예를 들어, ROM, RAM) 또는 디스크 타입 저장 매체 (예를 들어, 자기식 또는 광학식) 를 포함할 수 있다. 몇몇 양태에서, 결합기 (4143) 는 저장 매체의 일부 또는 전부를 포함할 수 있는 저장 매체에 저장될 수 있다. 도 41 의 흐름도 (4200) 는 프레임에 압축 타입을 할당하는 프로세스를 도시한다. 일 양태에서, 식 (3) 에서 정해진 현재 프레임 차이는 프레임 할당과 관련하여 이루어진 모든 결정에 대한 기초가 된다. 결정 블록 4253 이 나타내는 바와 같이, 고려 중인 프레임이 시퀀스 내의 첫 번째 프레임이면, "예" 로 표시된 결정 경로는 블록 4255 로 진행하여, 그 프레임이 I 프레임인 것으로 선언한다. 블록 4257 에서 누산된 프레임 차이는 0 으로 설정되고, 프로세서는 (블록 4258 에서) 시작 블록 4253 으로 리턴한다. 고려 중인 프레임이 시퀀스 내의 첫 번째 프레임이 아니면, "아니오" 로 표시된 경로는 결정이 이루어진 블록 4253 으로부터 진행하여, 테스트 블록 4259 에서 현재 프레임 차이를 장면 전환 임계값에 대해 테스트한다. 현재 프레임 차이가 그 임계값보다 크면, "예" 로 표시된 결정 경로는 블록 4255 로 진행하여, 다시 I-프레임을 할당하게 된다. 현재 프레임 차이가 장면 전환 임계값보다 작으면, "아니오" 경로는 블록 4261 로 진행하여, 현재 프레임 차이를 누산된 프레임 차이에 더한다.

[0344] 흐름도를 통해 계속 진행하여, 결정 블록 4263 에서, 누산된 프레임 차이를 임계값 t 와 비교하는데, 이 임계값 t 는 일반적으로 장면 전환 임계값보다 작다. 누산된 프레임 차이가 t 보다 크면, 제어는 블록 4265 로 이동하여, 프레임이 P 프레임이 되도록 할당되고, 그 다음에, 단계 4267 에서 누산된 프레임 차이가 리셋된다. 누산된 프레임 차이가 t 보다 작으면, 제어는 블록 4263 에서 블록 4269 로 이동한다. 블록 4269 에서 현재 프레임 차이가 τ 와 비교되는데, 이 τ 는 t 보다 작다. 현재 프레임 차이가 τ 보다 작으면, 블록 4273 에서 그 프레임은 스킵되고, 현재 프레임 차이가 τ 보다 크면, 그 프레임은 β 가 되도록 할당된다.

[0345] 다른 양태에서, 다른 프레임 인코딩 복잡도 표시자 M^* 는 다음과 같이 정의된다.

[0346]
$$M^* = M \times \min(1, \alpha \max(0, SAD_p - s) \times \max(0, MV_p - m)), \quad (61)$$

[0347] 여기서, α 는 스케일러이고, SAD_p 는 순방향 모션 정보를 갖는 SAD 이고, MV_p 는 순방향 모션 보상으로 부터 모션 벡터의 화소에서 측정된 길이의 합이고, s 및 m 은 SAD_p 가 s 보다 낮거나 MV_p 가 m 보다 낮은 경우에 프레임 인코딩 복잡도 표시자를 0 으로 만드는 2 개의 임계값 숫자이다. M^* 는 도 41 의 흐름도 (4200) 에서 현재 프레임 차이 대신에 이용될 것이다. 알 수 있는 바와 같이, 순방향 모션 보상이 낮은 레벨의 움직임을 나타내는 경우에만, M^* 은 M 과 상이하다. 이 경우에, M 은 M 보다 작다.

[0348] 본원에 설명된 샷 검출 및 인코딩 양태는 플로차트, 흐름도, 구조도, 또는 블록도로서 도시되는 프로세스로서 설명될 수도 있다. 도면에 도시된 플로차트가 순차 프로세스로서 동작을 설명할 수도 있지만, 다수의 동작은 병렬로 또는 동시에 수행될 수 있다. 또한, 동작 순서가 재배열될 수도 있다. 통상, 일 프로세스는 그 동작이 완료된 때에 종료한다. 일 프로세스는 메서드, 함수, 절차, 서브루틴, 서브프로그램 등에 대응할 수도 있다. 프로세스가 함수에 대응하는 경우에, 그 종료는 호출 함수 또는 메인 함수로 그 함수가 리턴하는 것에 대응한다.

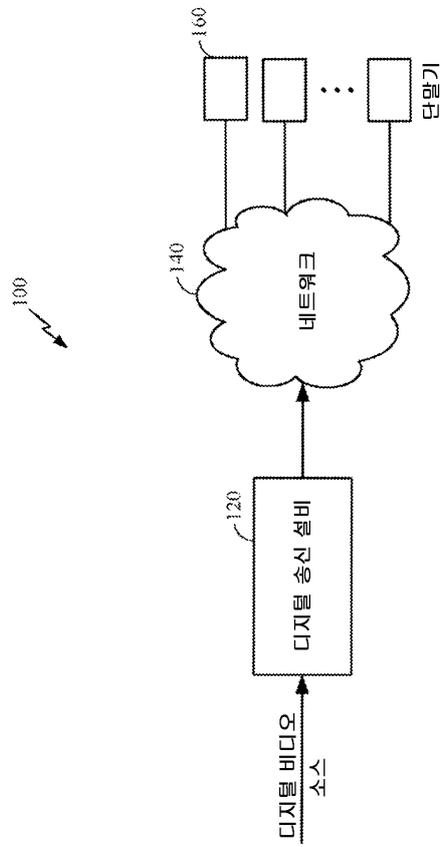
[0349] 또한, 당업자라면, 본원에 개시된 디바이스의 하나 이상의 요소가 그 디바이스의 동작에 영향을 주지 않으면서 재배열될 수도 있다는 것을 알 수 있다. 이와 유사하게, 본원에 개시된 디바이스의 하나 이상의 요소는 그 디바이스의 동작에 영향을 주지 않으면서 결합될 수도 있다. 당업자라면, 여러 상이한 테크놀러지와 기술 중 임의의 것을 이용하여 정보 및 멀티미디어 데이터를 표현할 수도 있다는 것을 알 수 있다. 또한, 당업자라면, 본원에 개시된 예와 관련하여 설명된 여러 예시적인 논리 블록, 모듈, 및 알고리즘 단계가 전자 하드웨어, 펌웨어, 컴퓨터 소프트웨어, 미들웨어, 마이크로코드, 또는 그 조합으로서 구현될 수도 있다는 것을 알 수 있다. 하드웨어와 소프트웨어의 호환성을 명확히 설명하기 위해, 여러 예시적인 컴포넌트, 블록, 모듈, 회로, 및 단계를 일반적으로 그 기능면에서 상술하였다. 그러한 기능이 하드웨어 또는 소프트웨어로 구현되는지 여부는 특정 애플리케이션 및 전체 시스템에 부과된 설계 제약에 따라 다르다. 당업자라면, 각각의 특정 애플리케이션에 대해 설명된 기능을 다양한 방법으로 구현할 수도 있지만, 그러한 구현 결정은 개시된

방법의 범위로부터 일탈하는 것으로서 해석되지 않아야 한다.

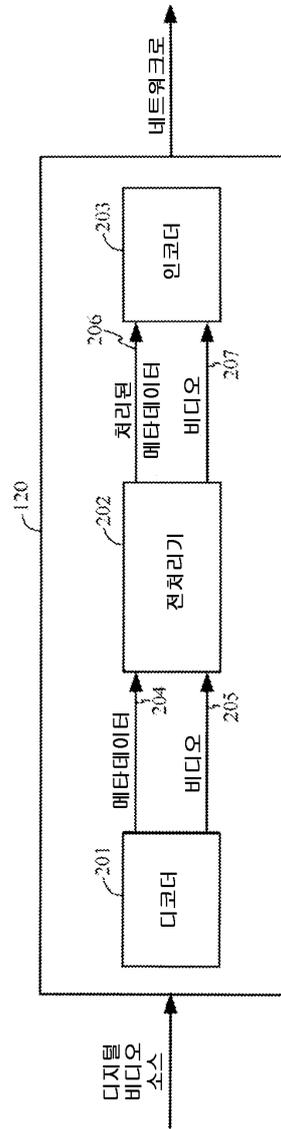
- [0350] 예를 들어, 본원에 개시된 샷 검출과 인코딩 예 및 도면과 관련하여 설명된 방법 또는 알고리즘의 단계는 하드웨어에 직접, 프로세서에 의해 실행되는 소프트웨어 모듈에, 또는 양자의 조합으로 구현될 수도 있다. 그 방법 및 알고리즘은 휴대 전화기, 컴퓨터, 랩톱 컴퓨터, PDA 및 모든 타입의 개인 및 비즈니스용 통신 디바이스로 비디오를 무선 송신하는 것을 비롯한 통신 기술에 특히 적용 가능하다. 소프트웨어 모듈은 RAM 메모리, 플래시 메모리, ROM 메모리, EPROM 메모리, EEPROM 메모리, 레지스터, 하드 디스크, 착탈식 디스크, CD-ROM, 또는 당해 기술분야에서 공지된 임의의 다른 형태의 저장 매체에 상주할 수도 있다. 예시적인 저장 매체는 프로세서에 연결되어, 프로세서가 저장 매체로부터 정보를 판독하거나 저장 매체에 정보를 기록할 수 있도록 한다. 다른 방법으로는, 저장 매체는 프로세서와 일체로 형성될 수도 있다. 프로세서 및 저장 매체는 ASIC (Application Specific Integrated Circuit) 에 상주할 수도 있다. ASIC 은 무선 모뎀에 상주할 수도 있다. 다른 방법으로는, 프로세서와 저장 매체는 무선 모뎀에 별도의 컴포넌트로서 상주할 수도 있다.
- [0351] 또한, 본원에 개시된 예와 관련하여 설명된 여러 예시적인 논리 블록, 컴포넌트, 모듈, 및 회로는 범용 프로세서, DSP (digital signal processor), ASIC (application specific integrated circuit), FPGA (field programmable gate array) 또는 다른 프로그램 가능 논리 디바이스, 별도의 게이트 또는 트랜지스터 논리, 별도의 하드웨어 컴포넌트, 또는 본원에 설명된 기능을 수행하도록 설계된 그 임의의 조합으로 구현되거나 수행될 수도 있다. 범용 프로세서는 마이크로프로세서일 수도 있지만, 다른 방법으로는, 프로세서는 임의의 통상적인 프로세서, 컨트롤러, 마이크로컨트롤러, 또는 상태 머신일 수도 있다. 또한, 프로세서는 컴퓨팅 디바이스의 조합, 예를 들어, DSP 와 마이크로프로세서의 조합, 복수의 마이크로프로세서, DSP 코어와 결합한 하나 이상의 마이크로프로세서, 또는 임의의 기타 그러한 구성으로서 구현될 수도 있다.
- [0352] 당업자라면, 개시된 예의 상술한 설명을 참조하여, 개시된 방법 및 장치를 실시하거나 이용할 수 있다. 당업자라면, 이들 예에 대한 여러 변형을 쉽게 알 수 있고, 본원에 정의된 원리는 개시된 방법 및 장치의 사상이나 범위로부터 일탈함이 없이 다른 예에 적용될 수도 있으며, 부가적인 요소가 추가될 수도 있다. 본 발명의 양태의 설명은 예시적인 것으로서, 청구항의 범위를 한정하려는 것은 아니다.

도면

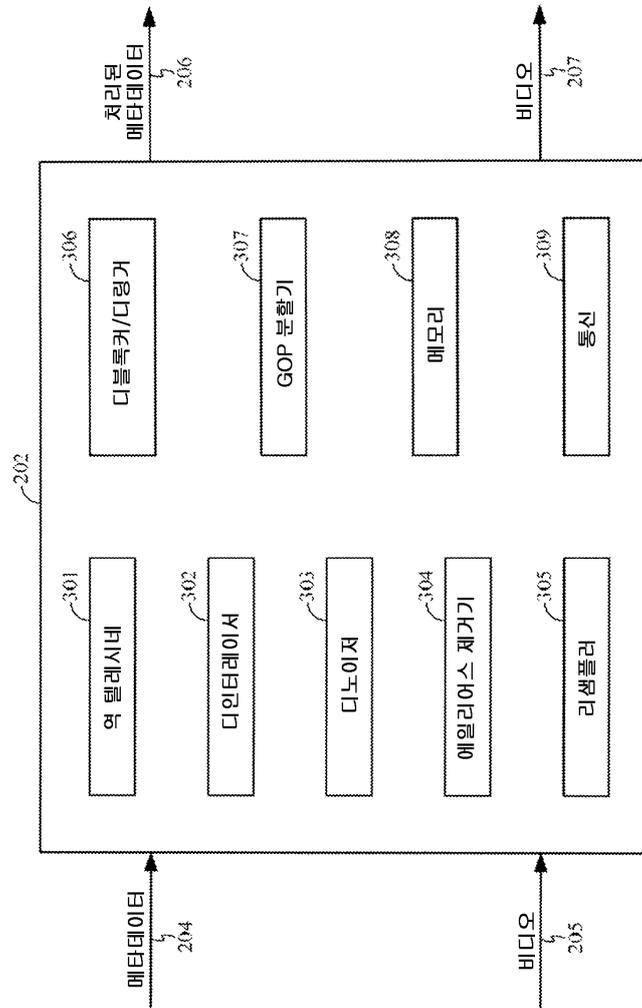
도면1



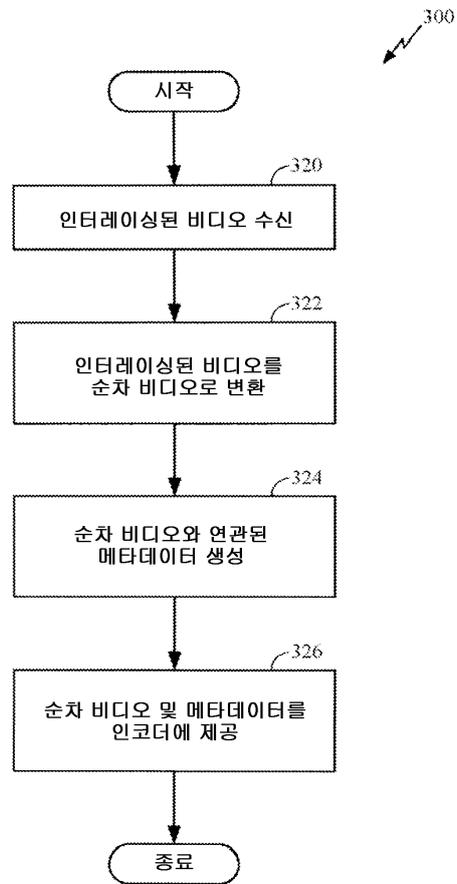
도면2



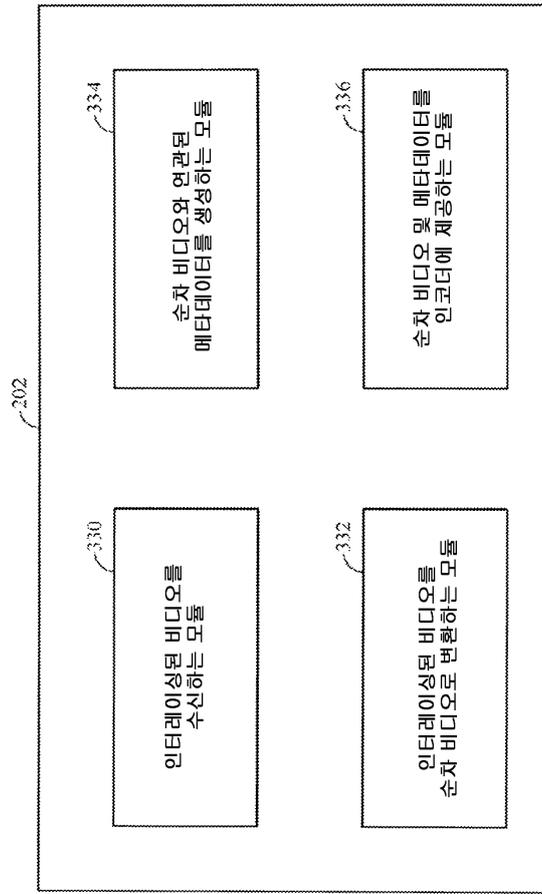
도면3a



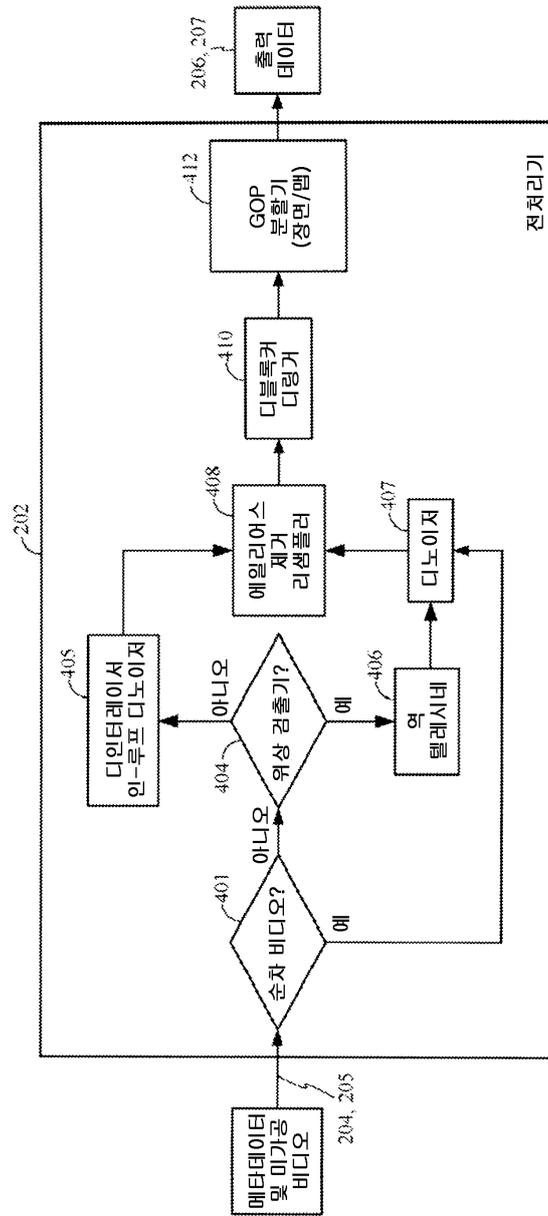
도면3b



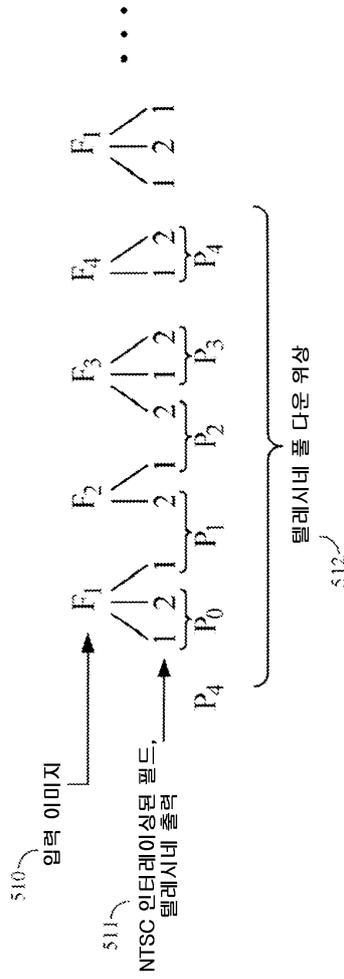
도면3c



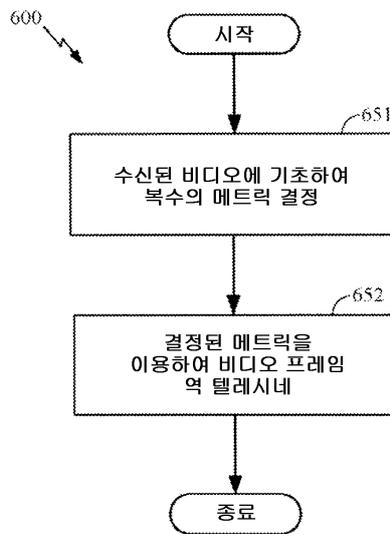
도면4



도면5

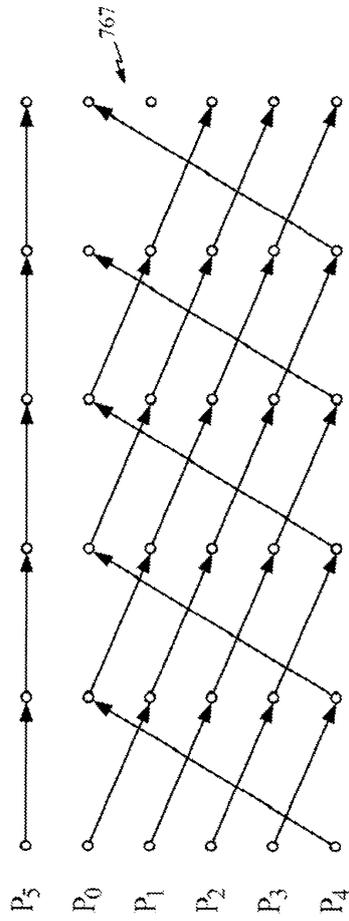


도면6

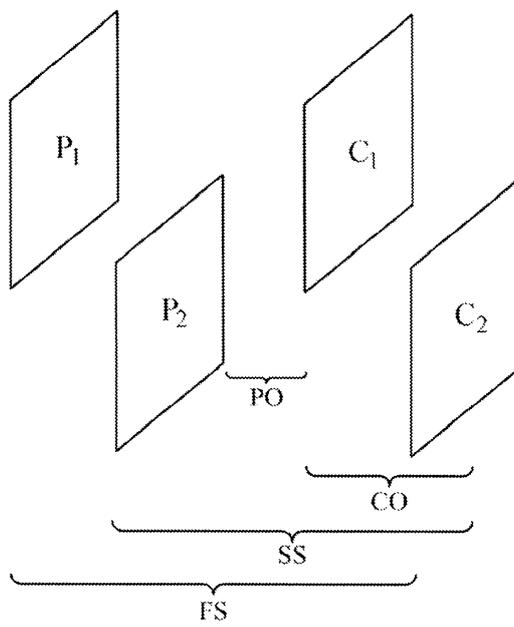


시스템 개요

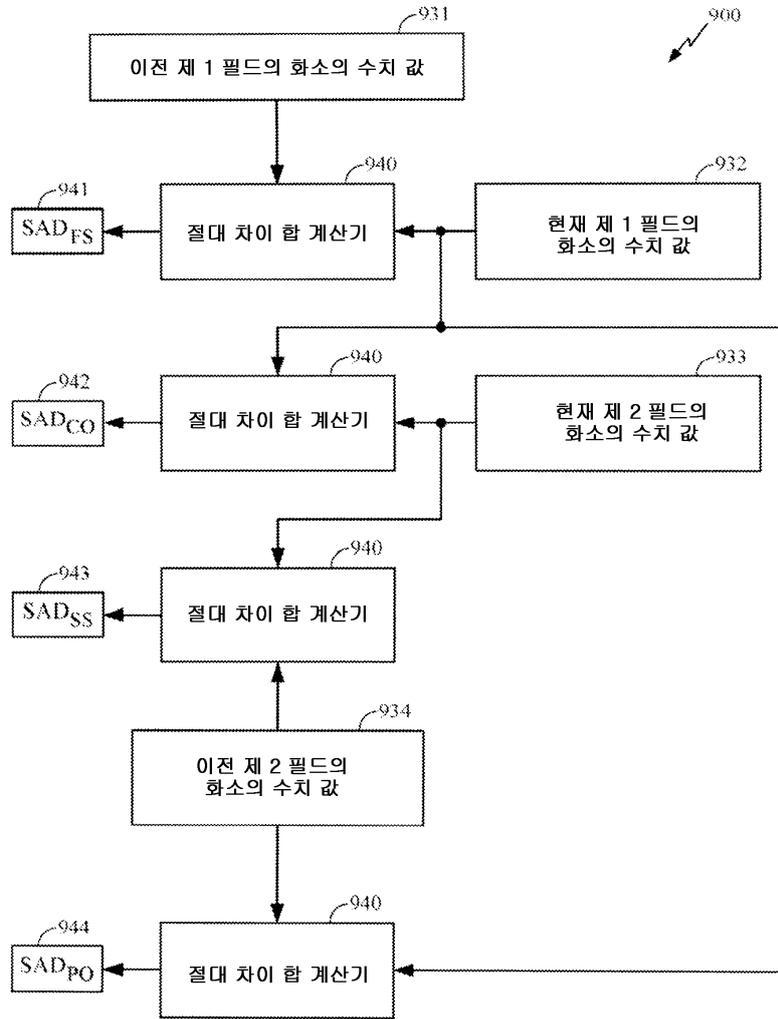
도면7



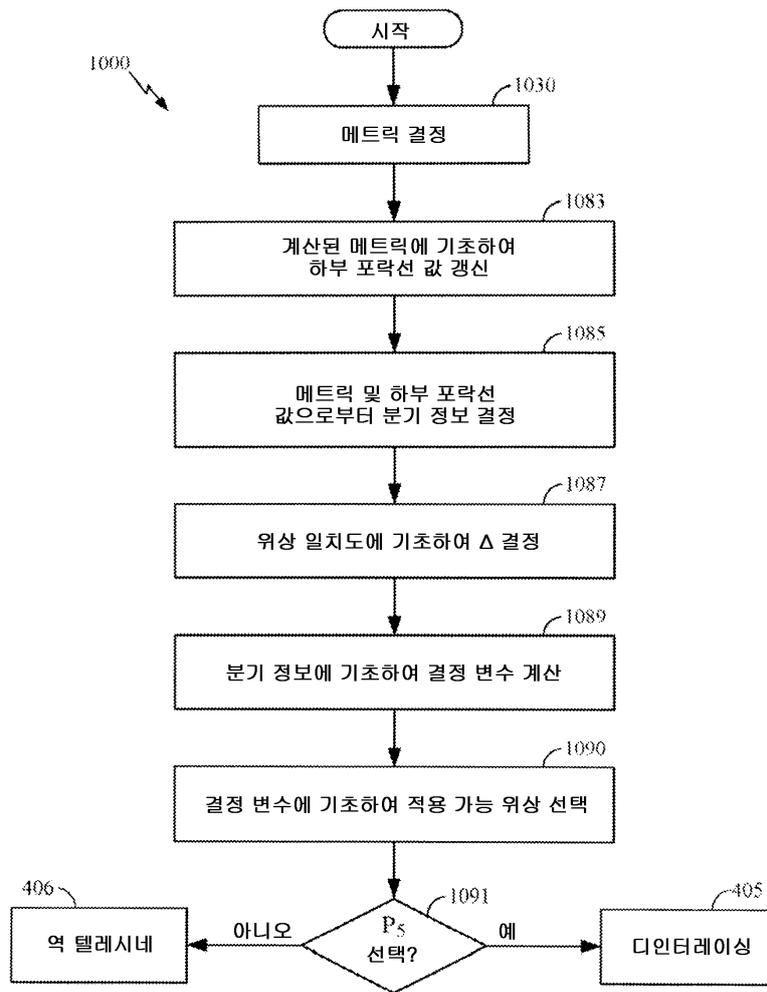
도면8



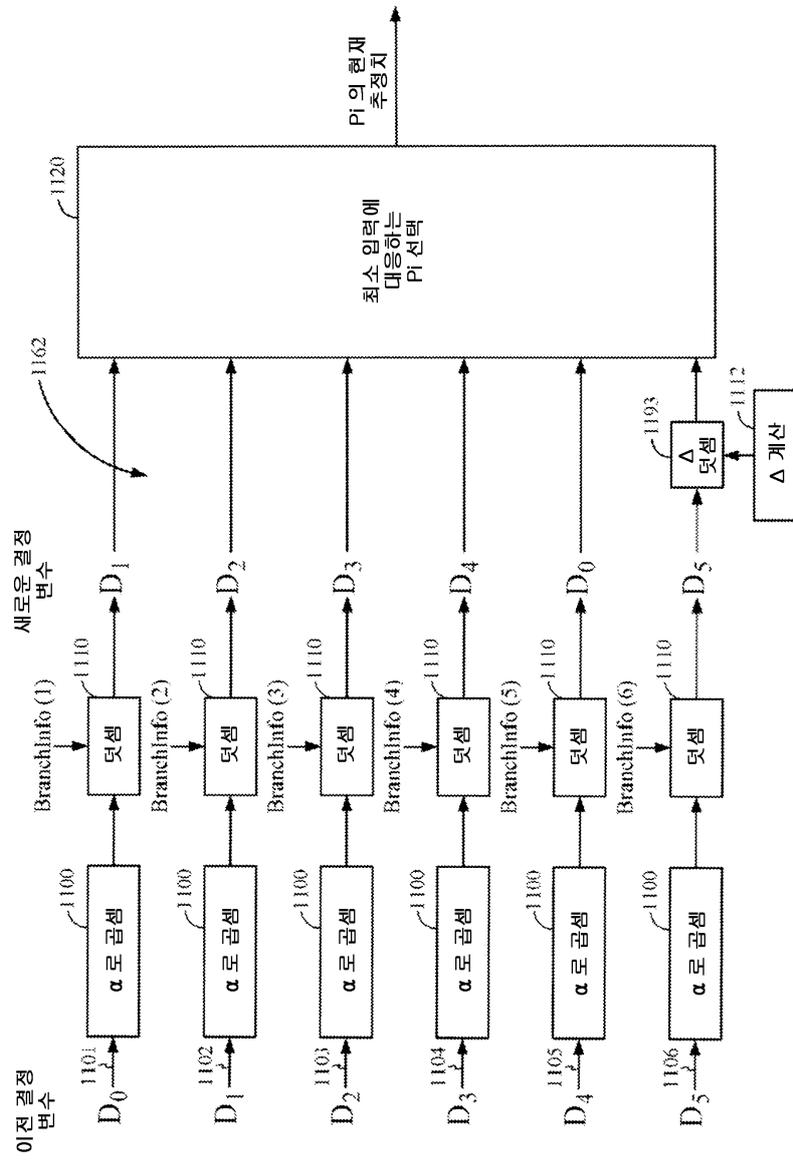
도면9



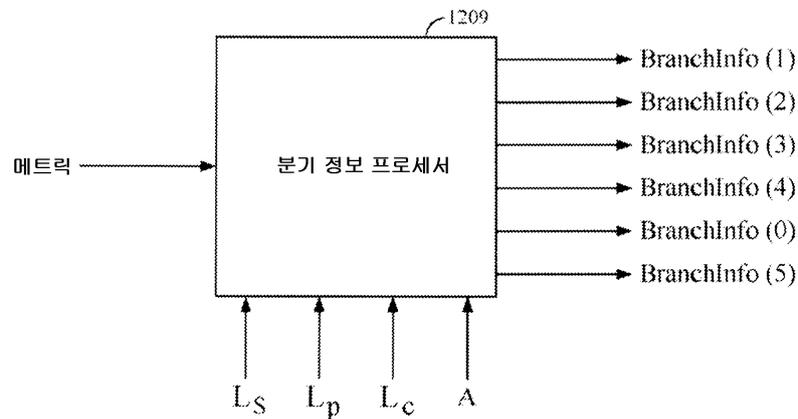
도면10



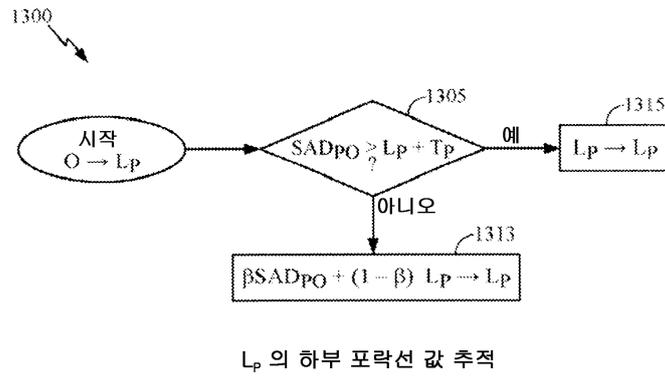
도면11



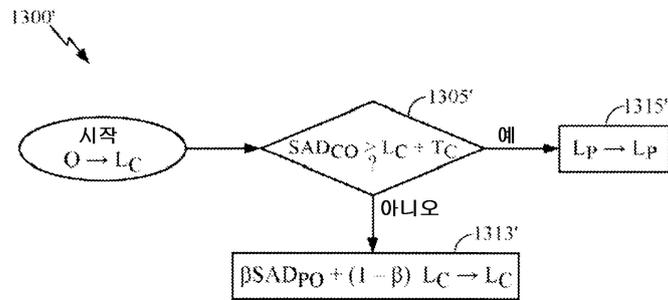
도면12



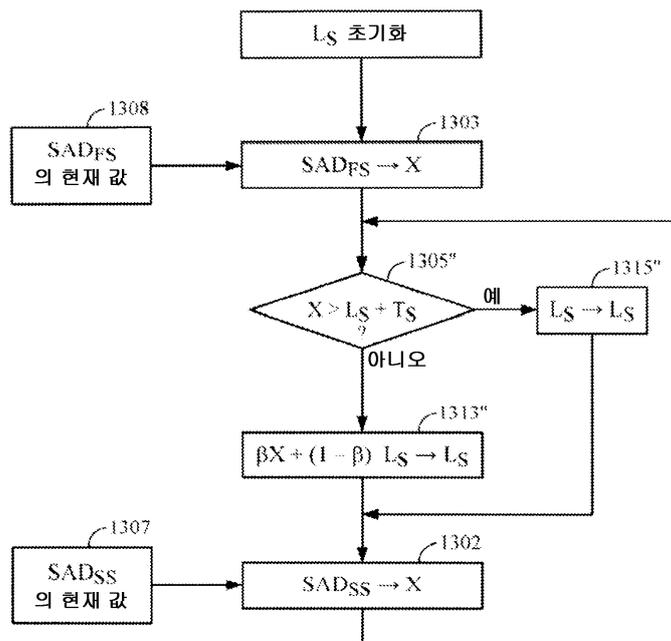
도면13a



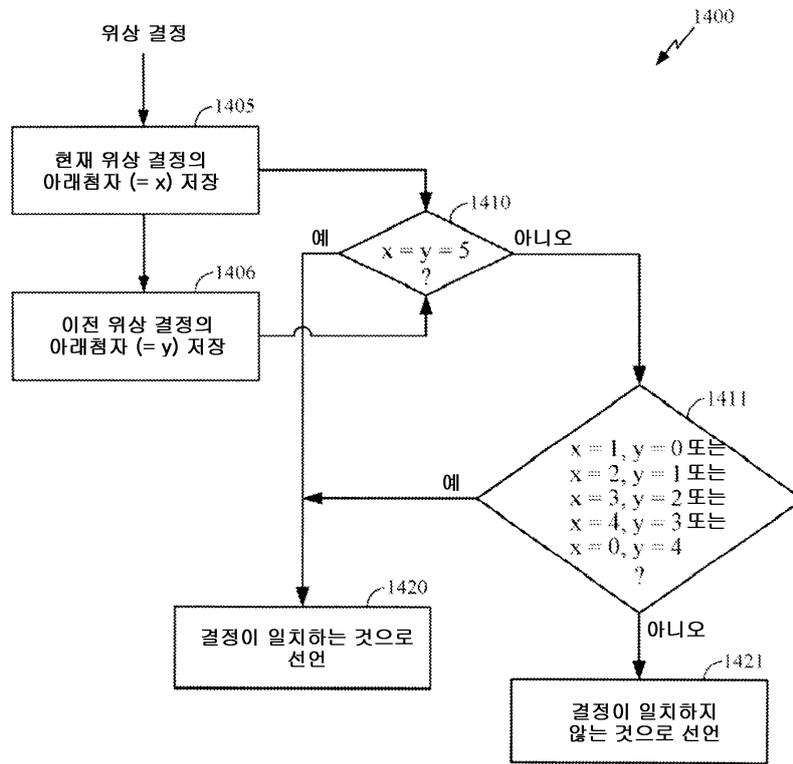
도면13b



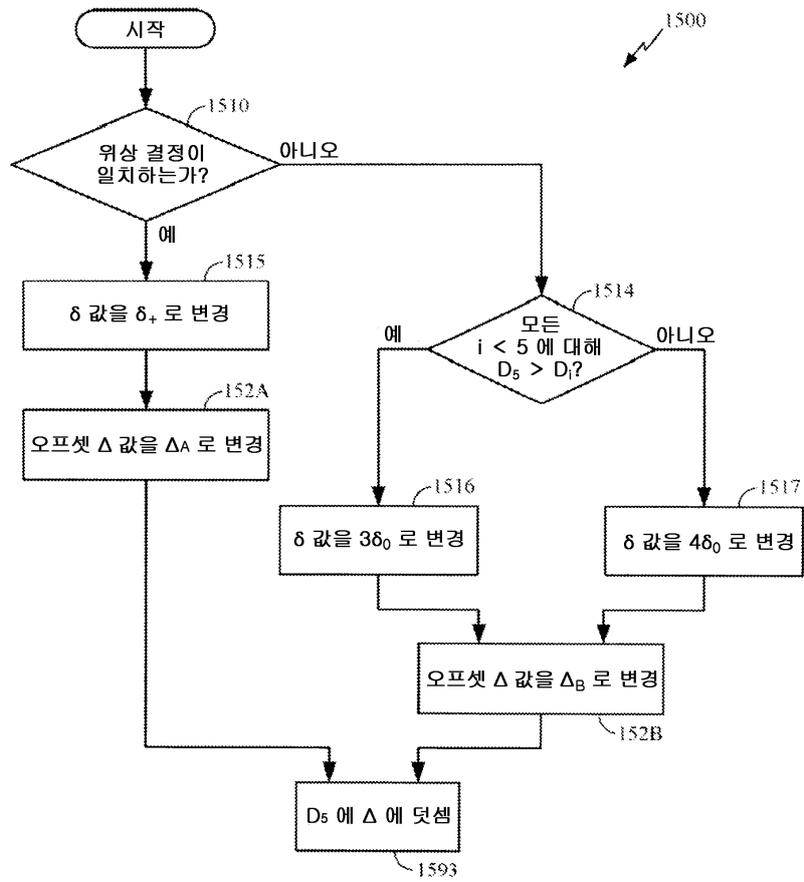
도면13c



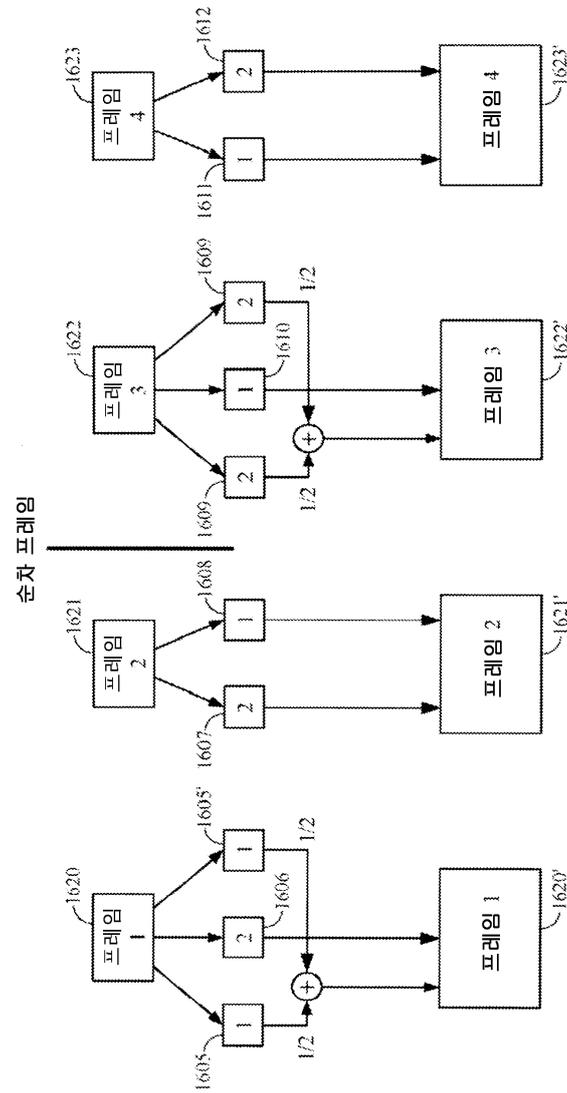
도면14



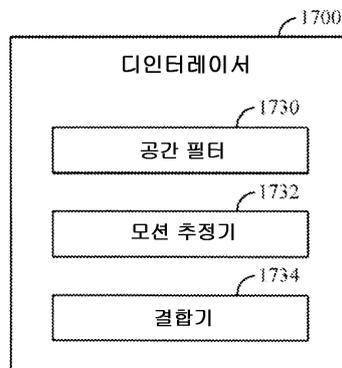
도면15



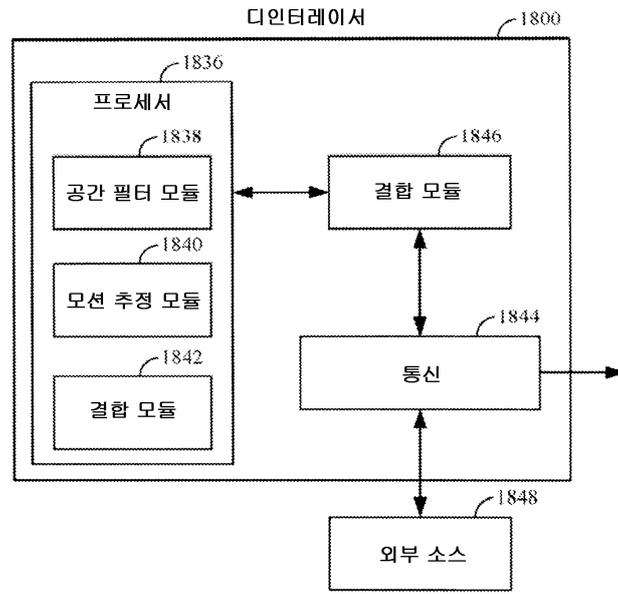
도면16



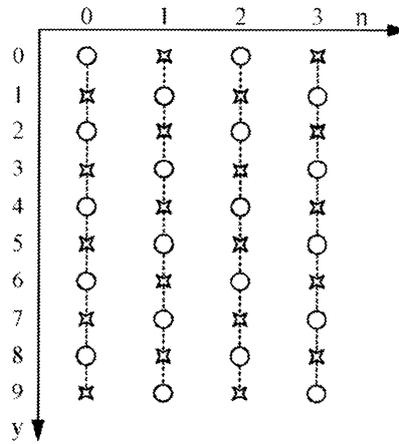
도면17



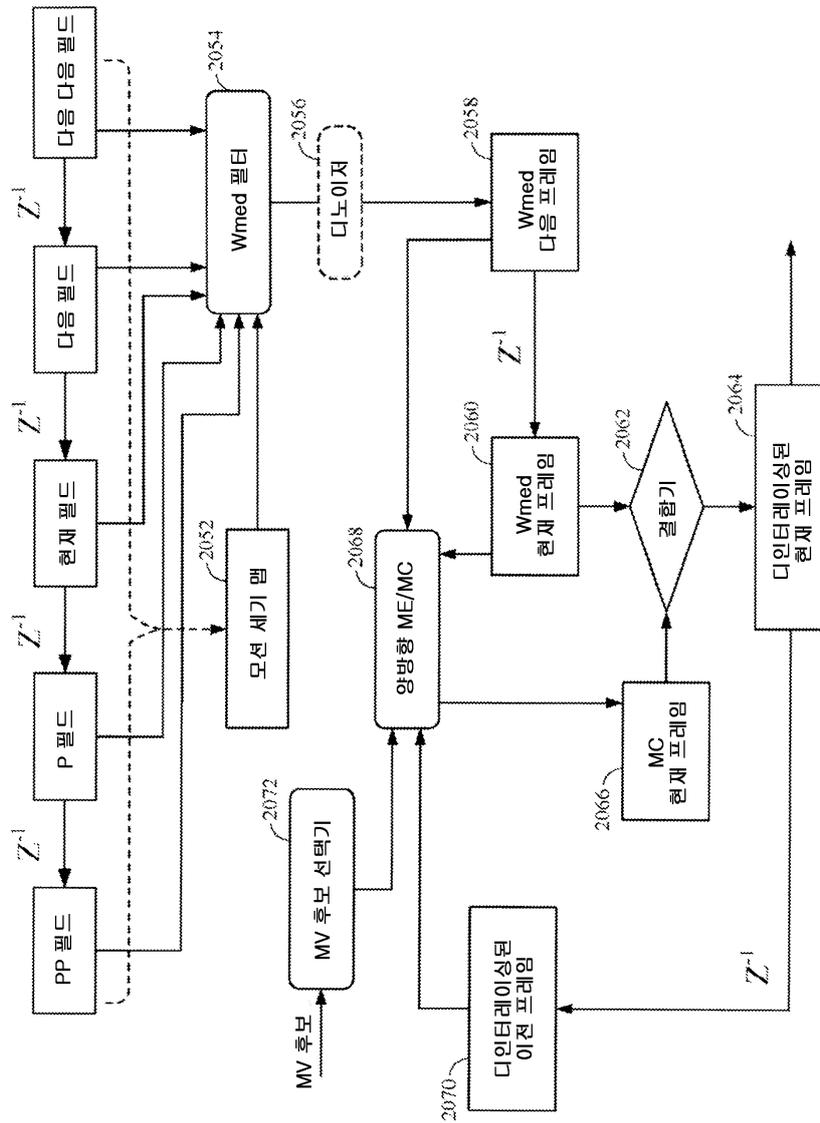
도면18



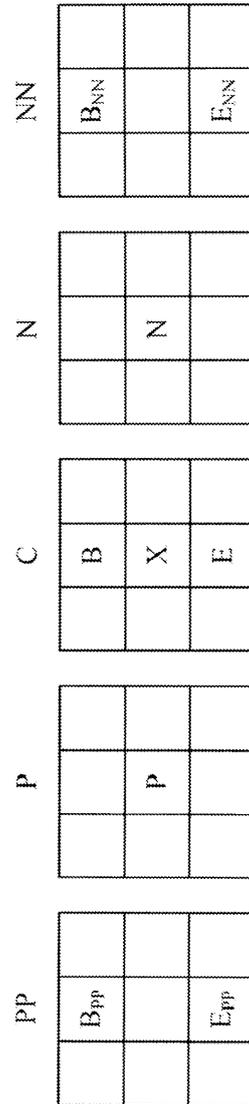
도면19



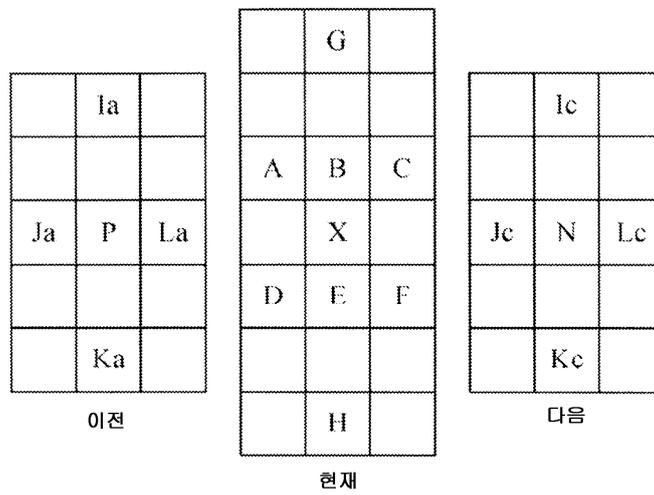
도면20



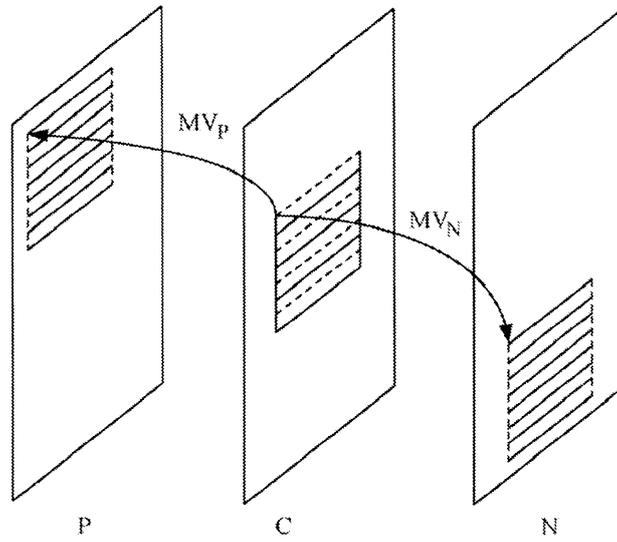
도면21



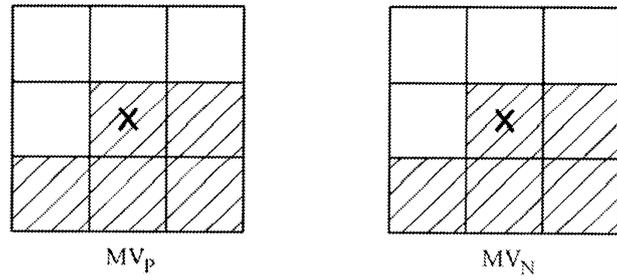
도면22



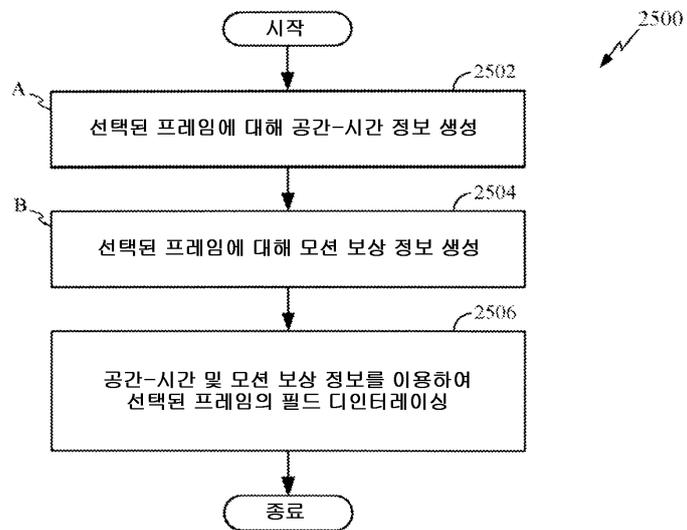
도면23



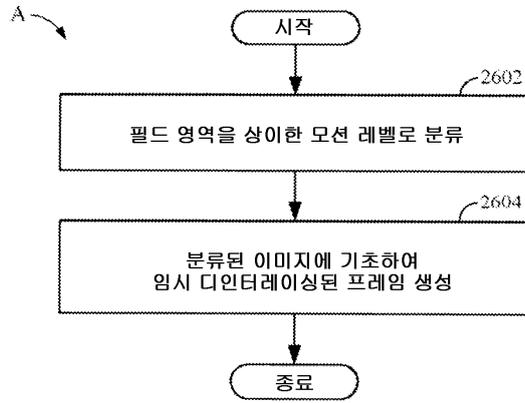
도면24



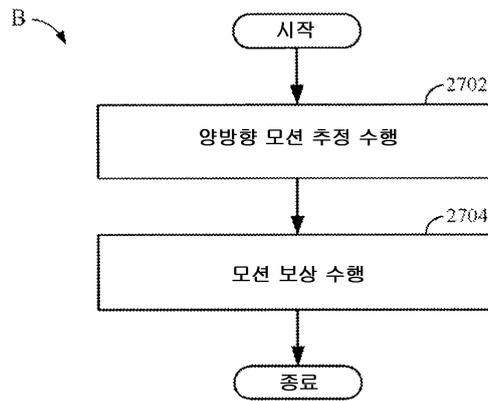
도면25



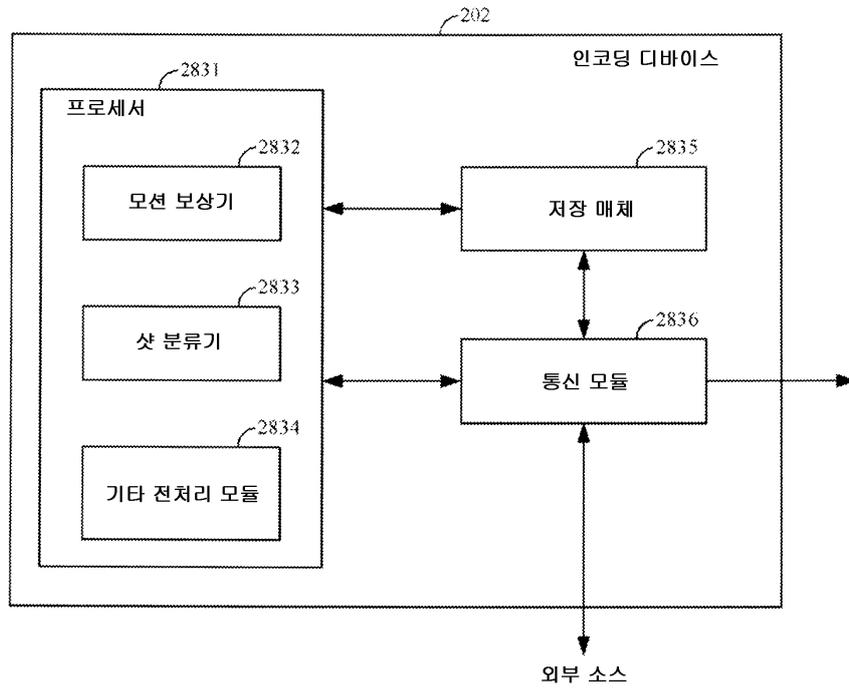
도면26



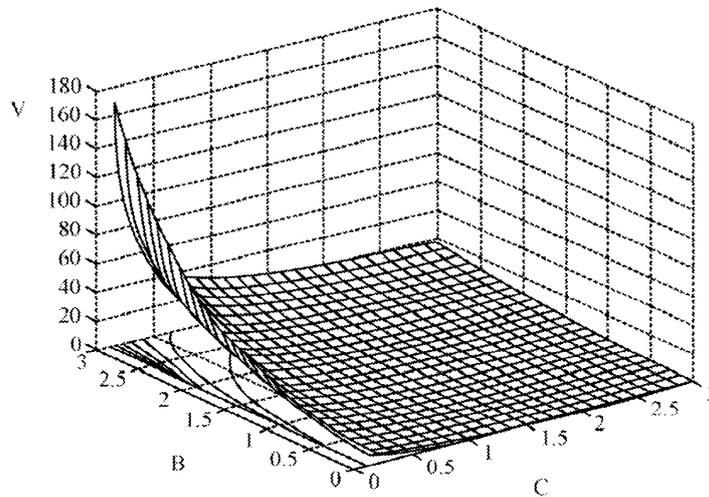
도면27



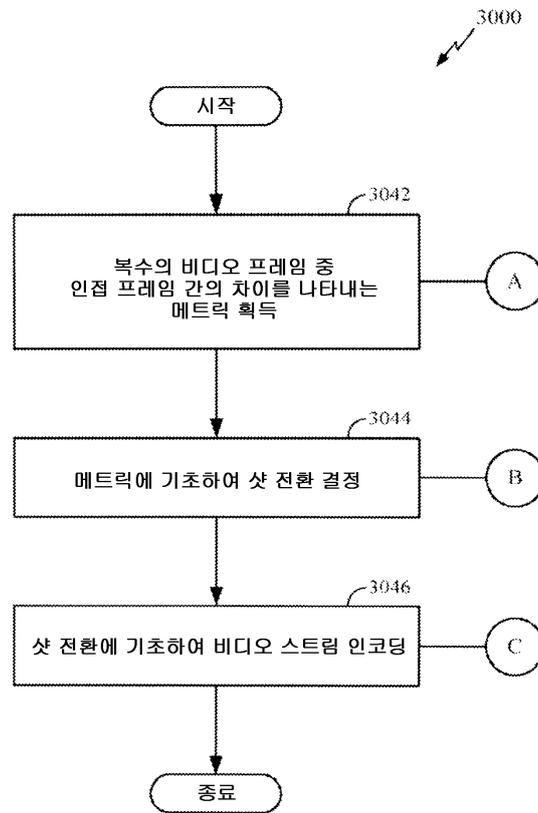
도면28



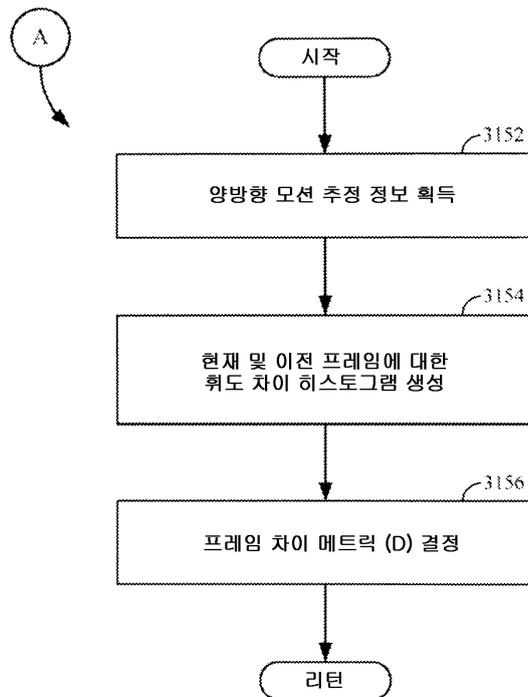
도면29



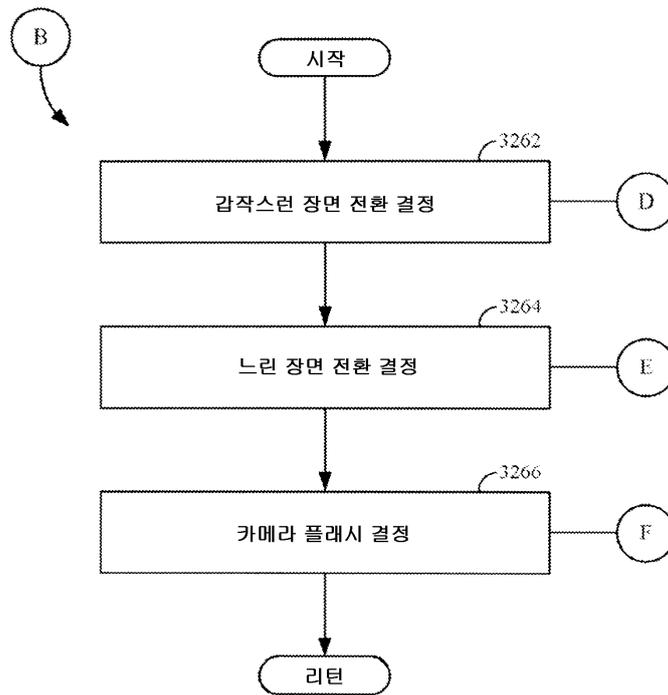
도면30



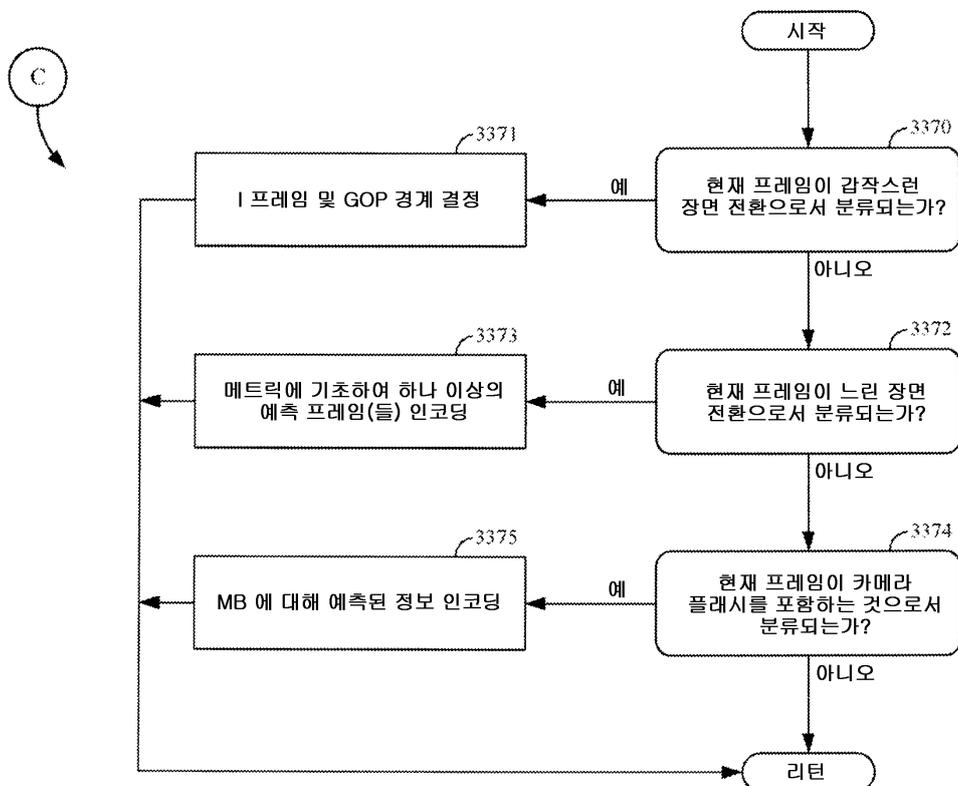
도면31



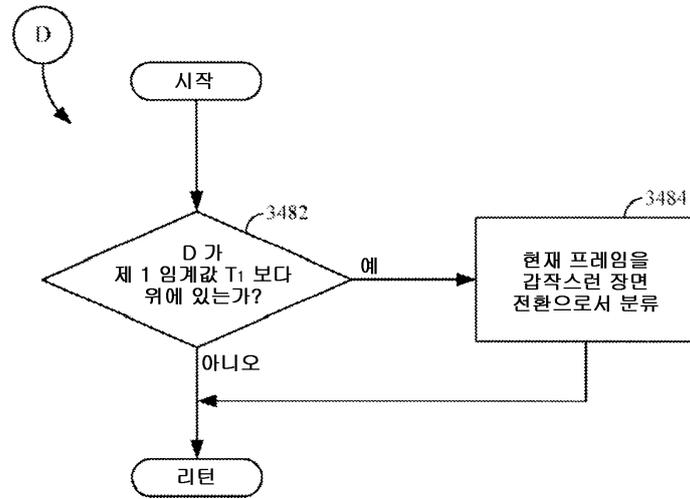
도면32



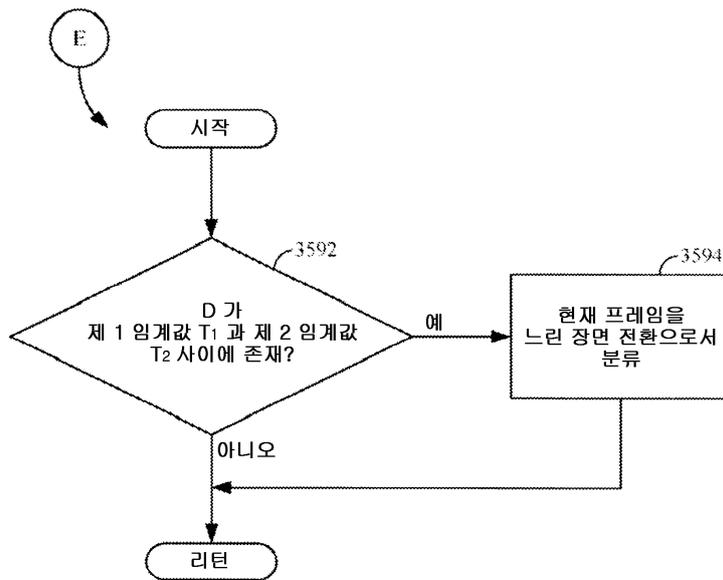
도면33



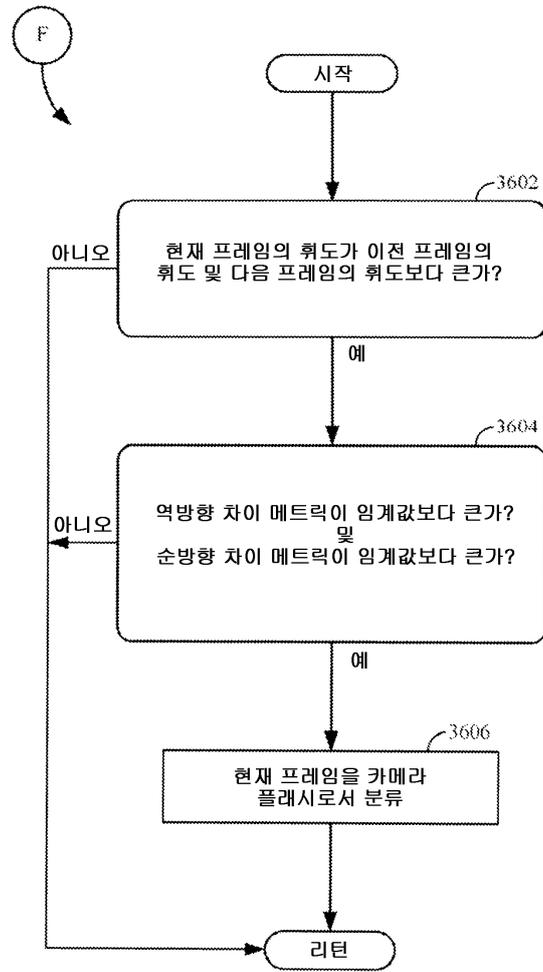
도면34



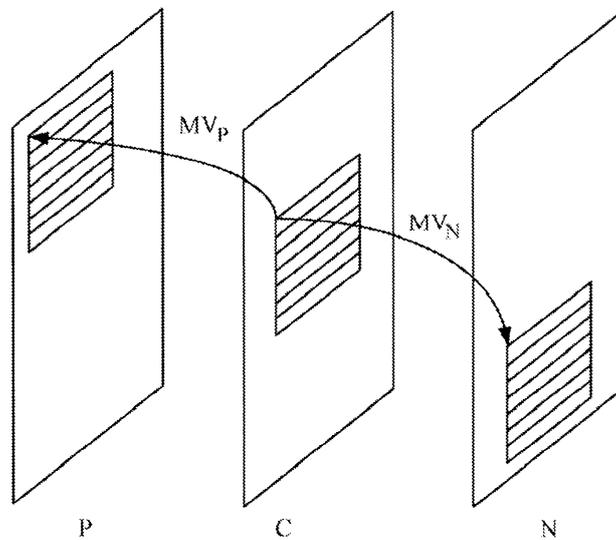
도면35



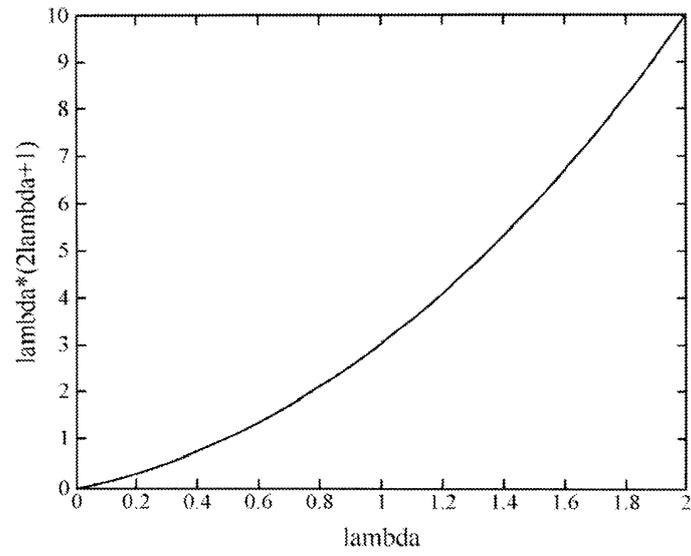
도면36



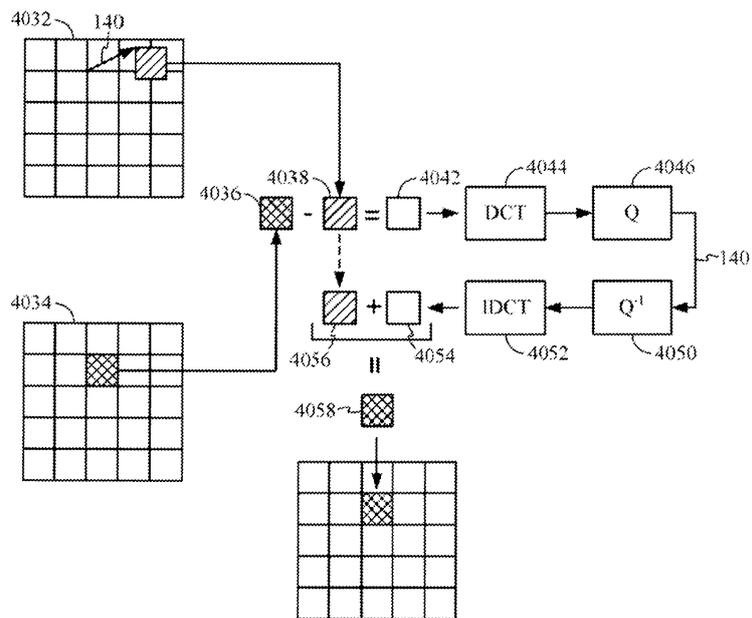
도면37



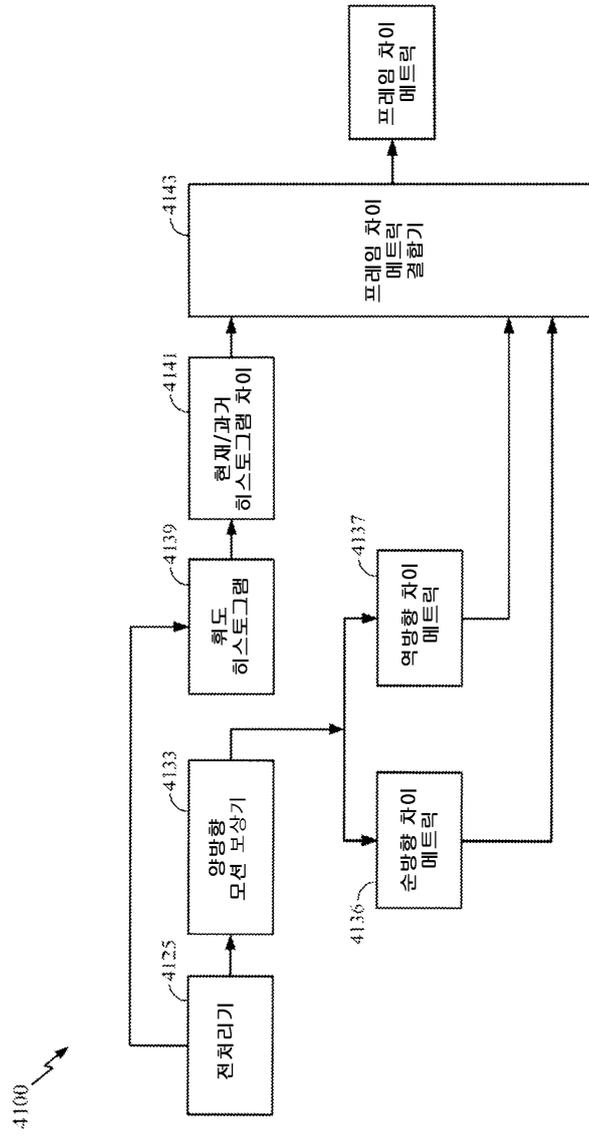
도면38



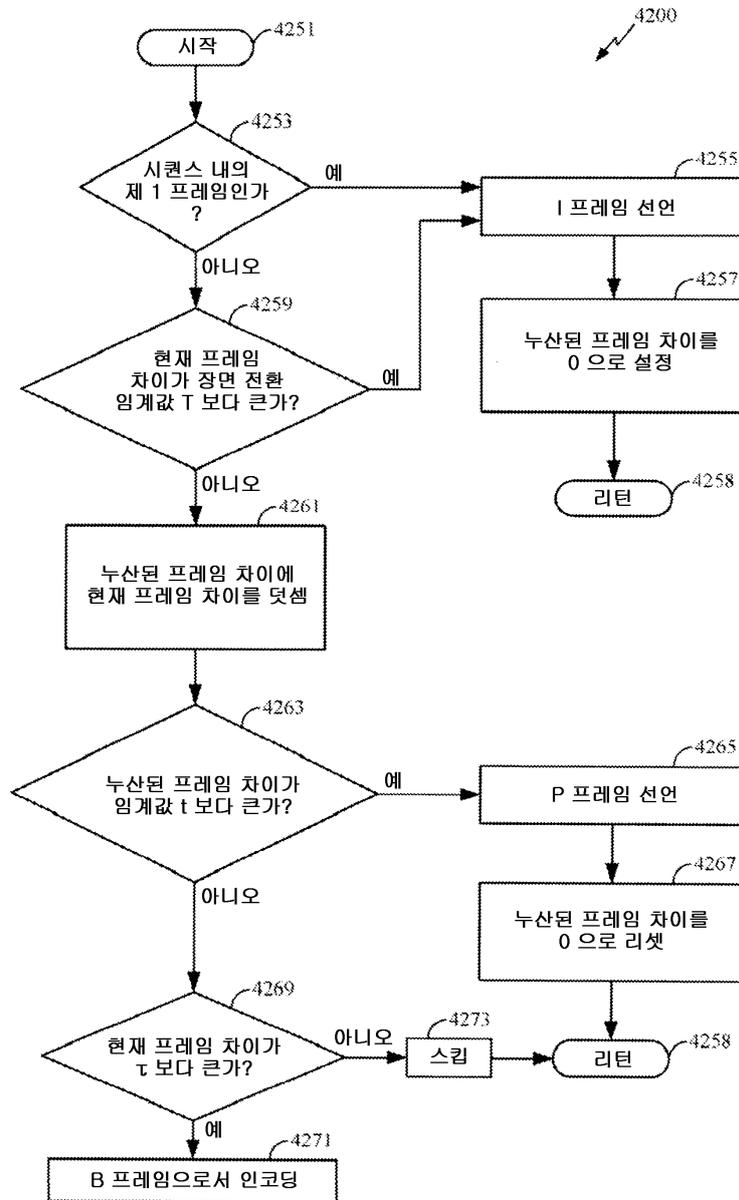
도면39



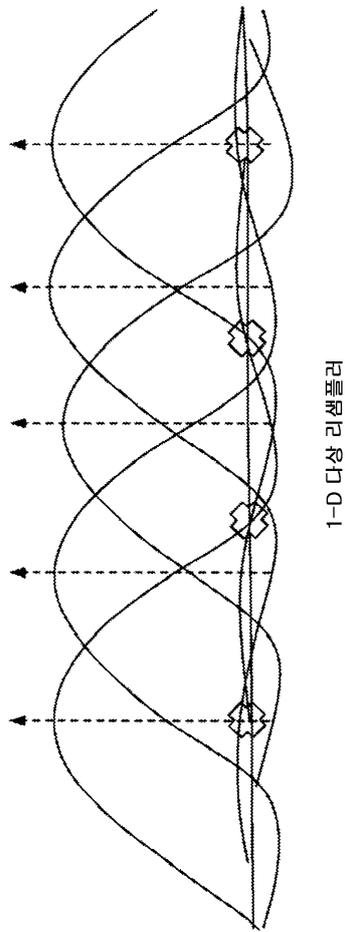
도면40



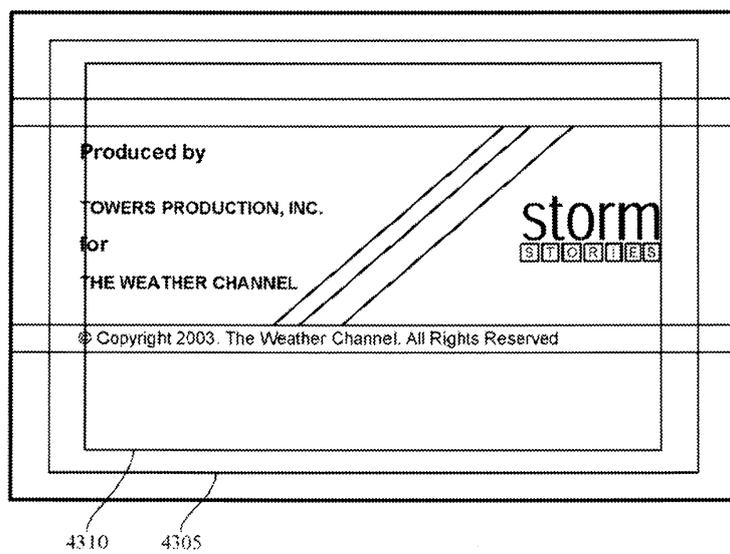
도면41



도면42



도면43



도면44

