



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2007년11월16일  
(11) 등록번호 10-0776529  
(24) 등록일자 2007년11월08일

(51) Int. Cl.

G11B 20/10 (2006.01)

(21) 출원번호 10-2001-7014472

(22) 출원일자 2001년11월13일

심사청구일자 2006년03월13일

번역문제출일자 2001년11월13일

(65) 공개번호 10-2002-0006632

공개일자 2002년01월23일

(86) 국제출원번호 PCT/JP2001/001982

국제출원일자 2001년03월13일

(87) 국제공개번호 WO 2001/69936

국제공개일자 2001년09월20일

(30) 우선권주장

JP-P-2000-00068720 2000년03월13일 일본(JP)

60/204,729 2000년05월16일 미국(US)

(56) 선행기술조사문헌

WO 9852356 A

전체 청구항 수 : 총 15 항

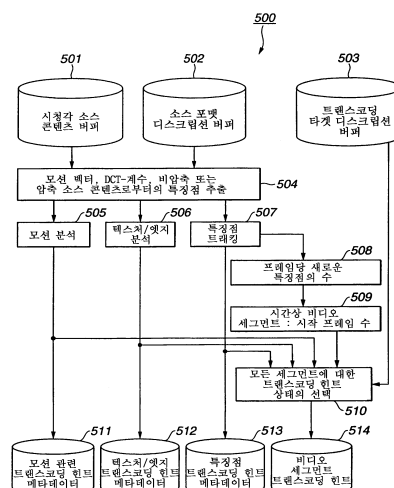
심사관 : 이강하

(54) 간결한 트랜스코딩 힌트 메타데이터를 생성하는 방법 및 장치

(57) 요약

멀티미디어 디스크립션(description)의 간결한 표현 및 다른 (예를 들면, MPEG) 압축 콘텐츠 표현들 사이에서의 트랜스코딩(transcoding)을 위한 트랜스코딩 힌트 메타데이터를 추출하고, 프레임 비율, 비트 레이트, 세션(session) 크기, 양자화 파라미터들과 같은 (예를 들면, MPEG 압축된) 비트스트림 파라미터들, 및 화상 코딩 타입 구조(예를 들면, 화상들의 그룹 또는 "GOP" (Group Of Pictures))를 조작하고, A/V 콘텐츠를 분류하고, 또한 멀티미디어 정보를 검색하기 위한 오디오/비디오 (또는 시청각, "A/V") 신호 처리 장치 및 방법이 제공된다.

대표도 - 도5



(81) 지정국

국내특허 : 오스트레일리아, 캐나다, 중국, 일본,  
대한민국, 미국

EP 유럽특허 : 오스트리아, 벨기에, 스위스, 독일,  
스페인, 프랑스, 영국, 그리스, 아일랜드, 이탈리  
아, 룩셈부르크, 모나코, 네덜란드, 포르투갈, 스  
웨덴, 핀란드, 사이프러스, 터어키

---

**특허청구의 범위**

청구항 1

삭제

청구항 2

삭제

청구항 3

삭제

청구항 4

삭제

청구항 5

삭제

청구항 6

삭제

청구항 7

삭제

청구항 8

삭제

청구항 9

삭제

청구항 10

삭제

청구항 11

삭제

청구항 12

삭제

청구항 13

삭제

청구항 14

삭제

청구항 15

삭제

청구항 16

삭제

청구항 17

삭제

청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

**청구항 49**

삭제

**청구항 50**

삭제

**청구항 51**

삭제

**청구항 52**

삭제

**청구항 53**

삭제

**청구항 54**

삭제

**청구항 55**

삭제

**청구항 56**

삭제

**청구항 57**

삭제

**청구항 58**

삭제

**청구항 59**

콘텐츠 처리 방법으로서:

단말(terminal)에 대한 단말 정보를 제 1 메모리에 저장하는 단계;

상기 콘텐츠 및 상기 콘텐츠를 처리하기 위한 콘텐츠 정보를 제 2 메모리에 저장하는 단계;

상기 콘텐츠 정보 및 상기 단말 정보에 기초하여 트랜스코딩 힌트들(transcoding hints)을 추출하는 단계; 및

상기 트랜스코딩 힌트들에 기초하여 상기 콘텐츠를 변환하는 단계를 포함하는, 콘텐츠 처리 방법.

**청구항 60**

제 59 항에 있어서, 상기 트랜스코딩 힌트 추출 단계는:

제 1 GOP 구조를 갖는 압축 이미지 데이터의 제 1 비트스트림을 수신하는 단계;

상기 제 1 비트스트림으로부터 제 1 모션 정보를 획득하는 단계;

제 1 분할(segmentation)의 텍스처/엣지 정보(texture/edge information)를 획득하는 단계;

상기 제 1 비트스트림으로부터 특징점들(feature points) 및 상기 특징점들과 연관된 모션 정보를 획득하는 단

계; 및

상기 제 1 비트스트림으로부터 관심있는 정보의 영역을 획득하는 단계를 포함하는, 콘텐츠 처리 방법.

#### 청구항 61

제 60 항에 있어서, 상기 트랜스코딩 힌트 추출 단계는 상기 제 1 모션 정보를 트랜스코딩 힌트들로서 저장하는 단계를 더 포함하는, 콘텐츠 처리 방법.

#### 청구항 62

제 60 항에 있어서, 상기 트랜스코딩 힌트 추출 단계는 모션 관련 트랜스코딩 힌트들을 파라메트릭 모션 모델(parametric motion model)의 파라미터들로서 표현하는 단계를 더 포함하는, 콘텐츠 처리 방법.

#### 청구항 63

제 62 항에 있어서, 상기 트랜스코딩 힌트 추출 단계는, 후속하는 직사각형의 비디오 프레임들 내의 글로벌 모션(global motion)을 기술(describe)하도록 상기 파라메트릭 모션 모델을 채용하는 단계를 더 포함하는, 콘텐츠 처리 방법.

#### 청구항 64

제 62 항에 있어서, 상기 트랜스코딩 힌트 추출 단계는, 임의 형태의 규정된 영역 내의 모션을 기술하도록 상기 파라메트릭 모션 모델을 채용하는 단계를 더 포함하는, 콘텐츠 처리 방법.

#### 청구항 65

제 64 항에 있어서, 상기 파라메트릭 모션 모델은, MPEG-4 내에서 이용되는 상기 임의 형태의 규정된 영역 내의 상기 모션을 기술하도록 채용되는, 콘텐츠 처리 방법.

#### 청구항 66

제 60 항에 있어서, 상기 트랜스코딩 힌트 추출 단계는 모션 관련 트랜스코딩 힌트들을, 상기 압축 이미지 데이터의 제 1 비트스트림에 포함된 모션 벡터들의 어레이로서 표현하는 단계를 더 포함하는, 콘텐츠 처리 방법.

#### 청구항 67

제 60 항에 있어서, 상기 트랜스코딩 힌트 추출 단계는 모션 관련 트랜스코딩 힌트들을, 상기 압축 이미지 데이터의 제 1 비트스트림에 포함된 모션 벡터들로부터 유도된 모션 벡터들의 어레이로서 표현하는 단계를 더 포함하는, 콘텐츠 처리 방법.

#### 청구항 68

제 60 항에 있어서, 상기 트랜스코딩 힌트 추출 단계는 모션 관련 트랜스코딩 힌트들을, 후속하는 프레임들 내에서 트래킹되는 모션 벡터들과 연관된 특징점들의 리스트로서 표현하는 단계를 더 포함하는, 콘텐츠 처리 방법.

#### 청구항 69

제 60 항에 있어서, 상기 트랜스코딩 힌트 추출 단계는 모션 관련 트랜스코딩 힌트들을, 후속하는 프레임들 내의 임의 형태의 영역들 내에서 트래킹되는 모션 벡터들과 연관된 특징점들의 리스트로서 표현하는 단계를 더 포함하는, 콘텐츠 처리 방법.

#### 청구항 70

제 60 항에 있어서, 상기 트랜스코딩 힌트 추출 단계는 텍스처 관련 트랜스코딩 힌트들을, DCT-계수들의 리스트 및 DCT-계수들의 리스트로부터 유도된 측정(평균치, 최소치, 최대치, 분산 및 표준 편차 중 하나) 중 하나로서 표현하는 단계를 더 포함하는, 콘텐츠 처리 방법.

#### 청구항 71

제 60 항에 있어서, 상기 트랜스코딩 힌트 추출 단계는 엣지 관련 트랜스코딩 힌트들을, DCT-계수들의 리스트 및 DCT-계수들의 리스트로부터 유도된 측정(평균치, 최소치, 최대치, 분산 및 표준 편차 중 하나) 중 하나로서 표현하는 단계를 더 포함하는, 콘텐츠 처리 방법.

## 청구항 72

제 60 항에 있어서, 상기 트랜스코딩 힌트 추출 단계는 특징점들 및 상기 특징점들과 연관된 모션 벡터들의 리스트를 표현하는 단계를 더 포함하는, 콘텐츠 처리 방법.

## 청구항 73

제 60 항에 있어서, 상기 트랜스코딩 힌트 추출 단계는, 하나의 프레임에서 다음 프레임으로 손실되고 새로운 다수의 특징점들을 이용함으로써 부호화-복잡성 관련 트랜스코딩 힌트들(encoding-complexity-related transcoding hints)을 후속하는 프레임들 내에서 트래킹되는 특징점들의 수명 리스트로부터 유도된 복잡성 메트릭(complexity metric)으로서 표현하는 단계를 더 포함하는, 콘텐츠 처리 방법.

## 명세서

### 기술 분야

<1> 본 발명은, 다른 (예를 들면, MPEG) 압축 콘텐츠 표현들 간의 트랜스코딩(transcoding)을 위해 멀티미디어 디스크립션(description) 및 트랜스코딩 힌트 메타데이터(transcoding hints metadata)의 간결한 표현을 추출하고, 프레임 비율, 비트 레이트, 세션(session) 크기, 양자화 파라미터와 같은 (예를 들어, MPEG 압축된) 비트스트림 파라미터들, 및 화상의 그룹 또는 "GOP (Group Of Pictures)"와 같은 화상 코딩 타입 구조를 조작하고, A/V 콘텐츠를 분류하고, 또한 멀티미디어 정보를 검색하는 오디오/비디오 (또는 시청각, "A/V") 신호 처리 방법 및 A/V 신호 처리 장치에 관한 것이다.

### 배경 기술

- <2> 광, 무선, 및 유선 네트워크를 통해 A/V 콘텐츠가 증가적으로 전송되고 있다. 이들 네트워크는 다른 네트워크 대역폭 제한치에 의해 특징지어지기 때문에, 주관적인 시각적 신호 품질을 변화시키도록 유발하는 다른 비트 레이트로 A/V 콘텐츠를 표현할 필요가 있다. A/V 콘텐츠의 압축 표현에 대한 추가 요구조건은 화면 크기, 계산 능력, 및 A/V 단말의 메모리 제한치에 의해 부여된다.
- <3> 따라서, 예를 들어 MPEG (Moving Pictures Experts Group)에 의해 규정된 압축 포맷으로 저장된 A/V 콘텐츠는 예컨대, 다른 비트 레이트, 프레임 비율, 화면 크기로 변환되어야 하며, 다른 A/V 단말들의 부호화 복잡도 및 메모리 제한치 변화에 따라 변환되어야 한다.
- <4> 다른 네트워크 대역폭 및 다른 A/V 단말에 대해 동일한 A/V 콘텐츠의 다중 압축 표현들을 저장할 필요를 없애기 위해, 압축 MPEG 포맷으로 저장된 A/V 콘텐츠는 다른 MPEG 포맷으로 트랜스코딩(transcoding) 될 수 있다.
- <5> 비디오 트랜스코딩에 대해서는 다음을 참고한다:
- <6> WO09838800A1: O.H. Werner, N.D. Wells, M.J. Knee: 개선된 양자화의 디지털 압축 부호화 (Digital Compression Encoding with improved quantization), 1999, 적응적 양자화 구조를 제안;
- <7> US5870146: Zhu; Qin-Fan: 디지털 비디오 트랜스코딩을 위한 디바이스 및 방법 (Device and method for digital video transcoding), 1999;
- <8> WO09929113A1: Nilsson, Michael, Erling; Ghanbari, Mohammed: 트랜스코딩 (Transcoding), 1999;
- <9> US5805224: Keesman; Gerrit J, Van Otterloo; Petrus J.: 비디오 신호를 트랜스코딩하는 방법 및 디바이스 (Method and Device for Transcoding Video Signal), 1998;
- <10> WO09943162AL Golin, Stuart, Jay: 비디오 시퀀스를 트랜스코딩하는 모션 벡터 외삽법 (Motion vector extrapolation for transcoding video sequences), 1999;
- <11> US5838664: Polomski; Mark D.: 디지털 트랜스코딩을 갖춘 화상 회의 시스템 (Video teleconferencing system with digital transcoding), 1998;



- <12> WO09957673A2: Balliol, Nicolas: 데이터 스트림의 트랜스코딩 (Transcoding of a data stream), 1999;
- <13> US5808570: Bakhmutsky; Michael: 호프만-트랜스코딩 및 고성능 가변 길이 디코더를 이용하는 2-워드 비트스트림 분할과 이들을 쌍-정합시키는 디바이스 및 방법 (Device and Method for pair-matching Huffman-Transcoding and high performance variable length decoder with two-word bitstream segmentation which utilizes the same), 1998;
- <14> WO09905870A2: Lemaguet, Yann: 비디오 시퀀스들 간의 스위칭 방법 및 대응하는 디바이스(Method of Switching between Video Sequences and corresponding Device), 1999; 및
- <15> WO09923560A1: LUDWIG, Lester; BROWN, William; YUL, Inn, J.; VUONG, Anh, T.; VANDERLIPPE, Richard; BURNETT, Gerald; LAUWERS, Chris; LUI, Richard; APPLEBAUM, Daniel: 스케일링 가능한 네트워크화 멀티미디어 시스템 및 응용 (Scalable networked multimedia system and application), 1999.
- <16> 그러나, 비디오 트랜스코딩에 대한 이들 특허 중 어떠한 것도, A/V 트랜스코딩을 용이하게 하는데 트랜스코딩 힌트 메타데이터 정보를 이용하는 것을 설명하거나 제안하지 않는다.
- <17> "SMPTE" (The Society of Motion Picture and Television)는 MPEG-2 비디오 기록 데이터 세트 상의 텔레비전에 대한 표준 (327M-2000)을 제안하였고, 이는 소스 포맷의 모든 매크로블록 (macroblock) 마다 256 비트를 이용하여 메타데이터를 재부호화하도록 제공된다. 그러나, 트랜스코딩 힌트 메타데이터의 이러한 추출 및 표현은 몇 가지 단점을 갖는다. 예를 들어, 제안된 표준에 따라, 트랜스코딩 힌트 메타데이터 (GOP 구조, 양자화 셋팅, 모션 벡터 등과 같은)는 A/V 소스 콘텐츠의 모든 단일 프레임 및 매크로블록마다 추출된다. 이 방법은 상세하고 콘텐츠 적응 트랜스코딩 힌트를 제공하는 이점을 제공하여, 주관적인 A/V 이중성(duality)을 널리 보존하면서 트랜스코딩을 용이하게 한다. 그러나, 트랜스코딩 힌트 데이터는 매우 크다. 제안된 표준의 하나의 특정 실시에서, MPEG 비디오의 매크로블록 당 256 비트의 트랜스코딩 힌트 데이터가 저장된다. 이 많은 양의 트랜스코딩 힌트 메타데이터는 로컬 (예를 들면, 가정) A/V 콘텐츠 서버에 방송을 제공하는 것이 쉽지 않다. 결과적으로, 트랜스코딩 힌트 메타데이터에 대해 제안된 표준은 방송 스튜디오 응용에 제한된다.
- <18> 트랜스코딩 힌트 메타데이터 추출 및 표현을 위한 또 다른 기술은 특정한 비트 레이트를 갖는 압축 A/V 소스 콘텐츠를 또 다른 압축 포맷 및 비트 레이트로 트랜스코딩하기 위해 일반 트랜스코딩 힌트 메타데이터를 수집하는 단계를 포함한다. 그러나, 이 기술은 트랜스코딩된 콘텐츠의 특징적인 특성을 고려하지 않는다는 단점이 있다. 예를 들어, 소스 콘텐츠에서, A/V 특성은 제한된 양의 모션 및 적은 상세들 (예를 들면, 뉴스 앵커 화면)을 갖는 A/V 세그먼트를 고속 모션 및 수많은 상세들 (예를 들면, 스포츠 이벤트 화면)을 갖는 또 다른 A/V 세그먼트로 변화시킬 수 있다. 이 기술에 따라, 비디오 세그먼트들 양자 모두의 다른 특성을 적절하게 표현하지 않은 잘못된 트랜스코딩 힌트 메타데이터가 선택되므로, A/V 신호 품질이 열악하고 잘못된 비트 레이트 할당을 제공하게 된다.

### 발명의 상세한 설명

- <19> 상기를 고려하여, 본 발명의 목적은 간결한 A/V-콘텐츠 적응 멀티미디어 디스크립션 및 트랜스코딩 힌트 메타데이터 표현(transcoding hints metadata representation)을 추출하는 방법 및 장치를 제공하는 것이다.
- <20> 본 발명의 또 다른 목적은, 트랜스코딩 방법에 대한 요구조건 중 하나로 상당한 지연이 없는 실시간 실행 및 계산의 복잡도를 갖지 않는 트랜스코딩 방법 및 장치를 제공하는 것이다. 트랜스코딩 방법에 대한 제 2 요구조건은 주관적인 A/V 신호 품질을 가능한 한 많이 보전하는 것이다. 다양한 압축 타겟 포맷에 대한 이들 요구조건 모두를 달성하는 트랜스코딩 방법을 용이하게 하기 위해, 트랜스코딩 힌트 메타데이터는 미리 생성되어 압축 A/V 콘텐츠와 함께 또는 개별적으로 저장될 수 있다. 본 발명의 또 다른 목적은, 저장 크기를 감소시키고 멀티미디어 디스크립션 및 트랜스코딩 힌트 메타데이터의 배급 (예를 들면, 로컬 A/V 콘텐츠 서버에 대한 방송)을 용이하게 하도록 매우 간결한 표현을 제공하는 것이다.
- <21> 따라서, 본 발명의 목적은: 1) 트랜스코딩 처리를 통해 A/V 신호 품질을 보전하고, 2) 최소 지연으로 실시간 응용을 가능하게 하기 위해 계산의 복잡도를 제한하는 트랜스코딩 시스템을 제공하는 것이다. 본 발명의 한 실시예에 따라, 트랜스코딩 힌트를 포함하는 추가 데이터 (메타데이터)는 압축 A/V 콘텐츠에 연관될 수 있다.
- <22> 본 발명의 다른 목적 및 이점은 명세서 및 도면으로부터 일부 밝혀지고 명확해진다. 본 발명은 자동적인 트랜스코딩 힌트 메타데이터 추출 및 간결한 표현을 제공하는 장치 및 방법에 관한 것이다.

- <23> 본 발명은 지원하는 트랜스코딩 메타데이터를 이용하여 압축 A/V 콘텐츠를 하나의 압축 포맷으로부터 또 다른 포맷의 A/V 콘텐츠로 트랜스코딩하는 분야에 관한 것이다. 용어 트랜스코딩은 압축 포맷을 변화시키는 것에 제한되지 않고 (예를 들면, MPEG-2 포맷에서 MPEG-4 포맷으로의 변환), 프레임-비율 변환, 비트 레이트-변환, 세션-크기 변환, 화면-크기 변환, 화상 코딩 타입 변환 등을 포함한다.
- <24> 본 발명은 또한 비디오에서 다른 화면 활동의 등급으로 상술된 트랜스코딩 힌트 상태를 이용하여 자동적으로 비디오를 분류하는데 적용될 수 있다.
- <25> 따라서, 본 발명은 다음의 상세한 설명에서 예시화된 바와 같이 여러 단계 및 다른 것에 대해 하나 이상의 이와 관련된 단계, 또한 이러한 단계를 실시하도록 구성된 구조 특성, 소자의 조합 및 부품 배열을 실현하는 장치를 구비하고, 본 발명의 범위는 청구항에서 표시된다.

## 실시예

- <42> 본 발명의 더 완전한 이해를 위해, 다음의 설명 및 첨부한 도면을 참조한다.
- <43> 도 1은 본 발명의 한 실시예에 따라 가정용 네트워크 환경에서 트랜스코딩하기 위한 시스템(100)의 일반적인 개요도를 도시한다. 도 1에 도시된 바와 같이, A/V 콘텐츠 서버(102)는 A/V 콘텐츠 저장기(103), A/V 트랜스코딩 유닛(106), 트랜스코딩 힌트 메타데이터 추출 유닛(104), 및 A/V 트랜스코딩 힌트 메타데이터 저장 버퍼(105)를 포함한다. A/V 콘텐츠 저장기(103)는, 비트 레이트를 변화시키고 주관적인 신호 품질을 변화시키는 동시에 다양한 소스로부터 압축 A/V 자료들을 저장한다. 예를 들어, AV 콘텐츠 저장기(103)는 휴대용 DV(Digital Video) 비디오 카메라(111)로부터의 가정용 비디오, MPEG-4 인터넷 카메라(112)로부터의 매우 낮은 비트 레이트 (10 kbit/s)을 갖는 MPEG-4 압축 비디오, 및 일부 경우에서 이미 트랜스코딩 힌트 메타데이터와 연관되어 있는 방송 서비스(101)로부터 약 5 Mbit/s의 메인 레벨 ("MP@ML") 압축 방송 비디오에서의 MPEG-2 메인 프로파일을 포함할 수 있다. A/V 콘텐츠 서버(102)는 또한 상당히 더 높은 비트 레이트의 고선명도 압축 MPEG 비디오를 포함할 수 있다.
- <44> 도 1에 도시된 바와 같이, A/V 콘텐츠 서버(102)는, 유선 기반 또는 무선 가정용 네트워크일 수 있는 네트워크(113)에 연결된다. 제한되지는 않지만, 무선 MPEG-4 A/V PDA (Personal Digital Assistant)(107), 고선명도 텔레비전 오락용 고선명도 A/V 단말(108), A/V 게임 콘솔 (console)(109), 및 ITU-T (International Telecommunications Union Technical Standards Group) 기반의 비디오폰(110)을 포함하는 네트워크(113)에는 다른 특성을 갖는 여러 A/V 단말이 또한 부착될 수 있다. A/V 단말(107, 108, 109, 110)은 다른 비트 레이트의 전송 특성으로 (케이블이나 무선 링크로 인하여) 가정용 네트워크(113)에 부착될 수 있다.
- <45> 더욱이, 무선 비디오 PDA(107)는 예를 들어 계산 능력, 저장 메모리, 화면 크기, 비디오 프레임 비율, 및 네트워크 비트 레이트에 대해 제한될 수 있다. 그러므로, A/V 트랜스코딩 유닛(106)은 무선 전송을 위해, 예를 들어, European 25 fps (frames per second)의 5 Mbit/s MPEG-2 방송 비디오 및 A/V 콘텐츠 서버(102)에 포함된 720 x 480 pel을 MPEG-4 500 kbit/s 15 fps 비디오로 트랜스코딩하고 무선 MPEG-4 비디오 PDA(107)에 의한 352 x 240 pel 디스플레이 상에 디스플레이한다. A/V 트랜스코딩 유닛(106)은 A/V 콘텐츠의 압축 소스 비트 레이트를 특정한 각 타겟 A/V 단말 (107, 108, 109, 및 110)의 특성으로 실시간 트랜스코딩하기 위해 버퍼(105)로부터의 트랜스코딩 힌트 메타데이터를 이용한다. 트랜스코딩 힌트 메타데이터는 트랜스코딩 힌트 메타데이터 추출 유닛(104)에서 생성되거나, 방송 서비스(101)에 의해 제공될 수 있다.
- <46> 도 1에 도시된 바와 같이, 소스 포맷의 압축 비트스트림(116)은 (이후 "제1 비트스트림") A/V 콘텐츠 버퍼(103)로부터 A/V 트랜스코딩 유닛(106)으로 전달된다. 타겟 포맷의 비트스트림은 (이후 "제 2 비트스트림") 트랜스코딩 유닛(106)에서 트랜스코딩한 이후에 가정용 네트워크(113)로 전달된다. 가정용 네트워크(113)로부터, 예를 들어, 압축 DV 포맷의 콘텐츠는 링크(114)를 통해 A/V 콘텐츠 저장기(103)에 저장된다.
- <47> 도 2는 본 발명의 한 실시예에 따른 트랜스코딩 힌트 추출, 트랜스코딩 힌트 저장, 및 트랜스코딩 처리를 설명한다. 도 2에 도시된 바와 같이, 버퍼(201)는 소스 포맷의 A/V 콘텐츠를 포함한다. 버퍼(202)는 비트 레이트, 압축 방법, GOP 구조, 화면 크기, 인터레이스 (interlace) 또는 순차 (progressive) 포맷 등과 같은 소스 포맷의 디스크립션을 포함한다. 버퍼(203)는 비트 레이트, 압축 방법, GOP 구조, 화면 크기, 인터레이스 또는 순차 포맷 등과 같은 타겟 포맷의 디스크립션을 포함한다. 트랜스코딩 힌트 추출 유닛(207)은 A/V 버퍼(201)로부터 압축 소스 포맷의 A/V 콘텐츠, 버퍼(202)로부터 소스 포맷 디스크립션, 및 버퍼(203)로부터 트랜스코딩 타겟 포맷 디스크립션을 판독한다. 트랜스코딩 힌트 추출 유닛(207)에 의해 트랜스코딩 힌트가 계산된 이후에, 트랜스코딩 힌트는 트랜스코딩 힌트 메타데이터 버퍼(206)에 저장된다. A/V 트랜스코딩 유닛(205)은 A/V 콘텐츠 버퍼

(201)로부터 소스 포맷의 제1 비트스트림(204)을 판독하고, 버퍼(206)에 저장된 트랜스코딩 힌트 메타데이터를 통해 소스 포맷을 타겟 포맷으로 변환한다. A/V 트랜스코딩 유닛(205)은 새로운 압축 타겟 포맷의 제 2 비트스트림(208)을 A/V 타겟 버퍼 저장기(209)에 출력하여 저장한다.

<48> 도 3 및 도 4는 본 발명의 한 실시예에 따른 트랜스코딩 힌트 메타데이터 조직의 원리를 설명한다. MPEG-기반의 비디오 압축은 예측 가능한 방법을 이용하여, 연속적인 프레임 사이의 변화가 부호화된다. 한 프레임에서 다음 프레임 사이에 많은 수의 변화를 포함하는 비디오 콘텐츠는 프레임 사이의 변화가 작은 비디오 콘텐츠 보다 다른 재부호화 파라미터 셋팅을 요구한다 (비트 레이트를 제한하면서 주관적인 신호 품질을 유지하기 위해). 그러므로, 재부호화 파라미터를 미리 결정하는 것이 중요하다. 트랜스코딩 힌트 메타데이터 선택은 주로 예측 불가능한 시각 콘텐츠의 양과 특성에 의존한다. 새로운 시각 콘텐츠는 이전 프레임으로부터 예측될 수 없고, DCT-계수를 이용하여 비트 레이트 강조 부호화될 수 있다. 이와 같이, 본 발명의 방법은 프레임 당 새로운 콘텐츠의 양을 결정하기 위해 이전 프레임으로부터 현재 프레임까지 트래킹(tracking)되지 않은 새로운 특징점들의 수를 이용한다.

<49> 도 3은 비디오의 프레임수 (수평축, 시간축)에 따라 프레임 당 새로운 특징점의 수에 대한 그래프를 도시한다. 섹션(301)은 후속하는 프레임 사이에 매우 작은 양의 새로운 콘텐츠만이 나타나는 경우의 비디오 세그먼트 일부이므로, 각 트랜스코딩 힌트 메타데이터 (예를 들면, 큰 GOP 크기, 낮은 프레임 비율, 낮은 비트 레이트, ...)가 선택될 수 있다. 섹션(302)은 프레임 당 약간 더 높은 새로운 특징점의 수를 포함하고, 이는 트랜스코딩 힌트 메타데이터를 설명하는 상태가 선택되어 이 상황에 최적의 트랜스코딩 파라미터를 제공함을 의미한다 (예를 들면, 약간 더 작은 GOP 크기, 더 높은 비트 레이트). 섹션(303)은 프레임 당 높은 새로운 특징점의 수를 갖는 트랜스코딩 메타데이터 힌트 상태(transcoding metadata hints state)를 설명하므로, 화면 당 높은 새로운 콘텐츠량을 갖는다. 이와 같이, 더 작은 M 값 (I/P-프레임 거리) 및 더 높은 비트 레이트가 선택된다.

<50> 도 4는 3개의 이산 트랜스코딩 힌트 메타데이터 상태로 구성된 트랜스코딩 힌트 메타데이터 상태도에 대한 기본 조직예를 도시한다. 모든 이산 트랜스코딩 상태는 GOP 구조, 양자화기 파라미터(quantizer parameter), 비트 레이트, 화면 크기 등에 대한 메타데이터를 포함할 수 있다. 이들 트랜스코딩 힌트 파라미터는 고정값을 갖거나 또 다른 파라미터의 함수일 수 있다. 예를 들어, GOP 길이는 프레임 당 새로운 특징점의 수에 대한 이산 함수가 될 수 있고, 양자화기 파라미터는 DCT 계수로부터 유도된 텍스처 액티비티(texture activity) 및 엣지(edge)의 함수가 될 수 있다. 본 예에서 3개의 트랜스코딩 힌트 메타데이터 상태 각각은 3개의 다른 부호화 상황을 수용하도록 선택될 수 있다. 도 4에 도시된 바와 같이, 상태 "3"(403)은 프레임 당 높은 모션량과 낮은 새로운 콘텐츠량에 대해 선택되고, 이러한 콘텐츠에 대한 트랜스코딩 힌트 메타데이터의 최적 상태를 나타낸다. 상태 "2"(402)는 낮은 모션량과 높은 엣지 액티비티를 갖는 높은 콘텐츠량에 대해 선택되어, 높은 비트수가 소모될 것을 요구할 수 있다. 상태 "1"(401)은 예를 들어, 낮은 화면 액티비티를 갖는 A/V 콘텐츠에 대해 트랜스코딩 처리를 수용하도록 선택된다. 또한, 다른 크로스페이딩(crossfading) 효과, 갑작스런 화면 변화, 또는 두 화면 사이의 흑색 화상과 같이, 비디오 편집 효과에 대해 제공되는 다른 특수한 트랜스코딩 힌트 메타데이터 상태가 있다. 비디오 편집 효과의 위치는 수동적으로, 반자동적으로, 또는 완전히 자동적으로 검출될 수 있다.

<51> 도 5는 본 발명의 한 실시예에 따라 압축 및 비압축 소스 콘텐츠로부터 트랜스코딩 힌트 메타데이터를 추출하는 것을 설명한다. 도 5에 도시된 바와 같이, 시스템(500)은 A/V 소스 콘텐츠 버퍼(501), 소스 포맷 디스크립션 버퍼(502), 및 타겟 포맷 디스크립션 버퍼(503)를 포함한다.

<52> 메모리(504)는 압축 또는 비압축 영역으로부터 특징점 추출, DCT-계수, 및 모션 벡터를 저장하도록 포함된다. 압축 영역에서, P- 및 B-매크로블록으로부터의 모션 벡터는 비트스트림으로부터 직접 추출될 수 있다. 그러나, 인트라-매크로블록 (Intra-macroblock)에 대해서는 모션 벡터가 없다. 그러므로, B- 및 P- 매크로블록에 대해 구해진 모션 벡터는 I-매크로블록에 대해 보간될 수 있다. (Roy Wang, Thomas Huang: "MPEG 영역에서의 고속 카메라 모션 분석 (Fast Camera motion Analysis in MPEG domain)", IEEE International Conference on Image Processing, ICIP 99, Kobe, Japan, 1999년 10월을 참고) 인트라-매크로블록의 블록에 대한 DCT 계수는 비트스트림으로부터 직접 추출될 수 있다. P- 및 B-매크로블록에 대해, 제한된 수의 DCT-계수들 (DC 및 2AC 계수들)은 Shih-Fu Chang, David G. Messerschmid: "MC-DCT 압축 비디오의 조작 및 합성 (Manipulation and Composition of MC-DCT compressed video)", IEEE Journal on Selected Areas in Communications, vol. 8, 1996에 의해 설명된 방법에 의해 구해질 수 있다. 압축 영역 특징점 추출 및 모션 추정에 대한 모범적인 방법은, 여기서 참조로 통합된 Peter Kuhn: "압축 영역 특징점 등록 및 모션 추정을 위한 방법 및 장치 (Method and Apparatus for compressed domain feature point registration and motion estimation)", 1999년 12월, PCT 특허에 의해 개시된다. 일부 경우, AV 소스 콘텐츠는 MPEG-1, MPEG-2, MPEG-4, ITU-T H.261, 및 ITU-T H.263

에 의해 이용되는 DCT 및 모션 보상 원리에 기초하지 않는 비압축 포맷 또는 압축 포맷에서만 이용가능하다. DV 포맷에 대해서는 DCT-계수만이 이용가능한 경우가 될 수 있다. 이들 경우, 모션 벡터는 예를 들어, Peter Kuhn: "MPEG-4 모션 추정을 위한 알고리즘, 복잡도 분석, 및 VLSI 구조 (Algorithms, Complexity Analysis and VLSI Architectures for MPEG-4 Motion Estimation)", Kluwer Academic Publishers, 1999의 모션 추정 방법에 의해 구해질 수 있다. DCT-계수는 K.R. Rao, P. Yip: "이산 코사인 변환 - 알고리즘, 이점, 응용 (Discrete Cosine Transform - Algorithms, Advantages, Applications)", Academic Press 1990의 블록-기반의 DCT-변환을 실행함으로써 구해질 수 있다. 펠 영역(pel domain ; 비압축 영역)에서의 특징점은 예를 들어 Bruce D. Lucas, Takeo Kanade: "스테레오 비전에 대한 응용을 갖는 반복적 등록 기술 (An iterative registration technique with an application to stereo vision)", International Joint Conference on Artificial Intelligence, pp. 674-679, 1981에 의해 구해질 수 있다.

<53> 모션 분석 부분(505)은 메모리(504)내의 모션 벡터 표현으로부터 파라메트릭 모션 모델의 파라미터를 추출한다. 파라메트릭 모션 모델은 6 및 8개 파라미터를 갖고, 파라메트릭 모션 추정은 M. Tekalp: "디지털 비디오 처리 (Digital Video Processing)", Prentice Hall, 1995에서 설명된 방법에 의해 구해질 수 있다. 모션 표현을 이용하는 목적은 지연 및 속도를 이유로 하는 트랜스코더에서의 모션 추정을 제거하는 것이다. 그러므로, 소스 비트스트림으로부터의 모션에 대한 입력 표현은 출력 표현 (타겟 비트스트림)을 유도하는데 이용될 수 있다. 예를 들어, 화면-크기 재조정, 인터레이스-순차 변환(interlaced-progressive conversion) 등이 모션 표현에 많이 의존한다. 모션 표현에 대한 파라미터는 또한 GOP 구조에서의 코드화 결정에 이용될 수 있다. 텍스처/엠티지 분석 부분(506)은 예를 들어, K.R. Rao, P. Yip: "이산 코사인 변환 - 알고리즘, 이점, 응용 (Discrete Cosine Transform - Algorithms, Advantages, Applications)", Academic Press 1990 또는 K.W. Chum, K.W. Lim, H.D. Cho, J.B. Ra: "비디오 부호화에 대한 적응적 인지 양자화 알고리즘 (An adaptive perceptual quantization algorithm for video encoding)", IEEE Transactions on Consumer Electronics, Vol. 39, No. 3, 1993년 8월에서 설명된 바와 같이, 비트스트림으로부터 추출된 DCT-계수에 기초할 수 있다.

<54> 압축 영역에 대한 특징점 트래킹 부분(507)은 여기서 참고로 포함되는 Peter Kuhn: "압축 영역 특징점 등록 및 모션 추정을 위한 방법 및 장치 (Method and Apparatus for compressed domain feature point registration and motion estimation)", PCT 특허, 1999년 12월에서 설명된 기술을 이용한다. 프로세서(510)는 프레임 당 새로운 특징점의 수를 계산한다. 프로세서(509)는 임시 비디오 세그먼트를 계산하고, 프로세서(510)는 모든 세그먼트에 대한 트랜스코딩 힌트 상태를 계산한다. 본 발명의 한 실시예에 따른 이들 계산 방법들은 도 6, 도 7, 및 도 8을 참고로 이후 상세히 설명된다.

<55> 메모리(511)는 모션-관련 트랜스코딩 힌트 메타데이터를 포함한다. 메모리(512)는 텍스처/엠티지 관련 트랜스코딩 힌트 메타데이터를 포함하고, 메모리(513)는 특징점 트랜스코딩 힌트 메타데이터를 포함하며, 이들은 모두 도 15를 참고로 이후 상세히 설명된다. 메모리(514)는 비디오 세그먼트 트랜스코딩 힌트 선택 메타데이터를 포함하고, 이는 도 16을 참고로 설명된다. 이제는 트랜스코딩 힌트 메타데이터의 자동 추출, 간결한 표현, 및 용도가 설명된다.

<56> 도 6은 본 발명의 한 실시예에 따른 비디오 분할 및 트랜스코딩 힌트 상태 선택 처리를 설명한다. 단계(601)에서는 일부 변수가 초기화된다. 변수 "frame"은 소스 비트스트림의 현재 프레임수이고, "nframes"은 새로운 비디오 세그먼트 (또는 GOP, group of pictures)내에서의 프레임수이다. 다른 변수들은 이 루틴내에서만 이용된다. 단계(602)에서는 GOP내의 프레임수가 증가된다. 단계(603)에서는 새로운 세그먼트/GOP가 그 프레임내에서 시작되는지 여부를 결정하고, 이에 대한 상세한 내용은 도 7을 참고로 상세히 논의된다. 그런 경우 ("예"), 제어는 단계(604)로 전해지고, 그렇지 않은 경우에는 단계(615)로 전해진다. 단계(604)에서는 변수 "last\_gop\_start"가 "new\_gop\_start"의 값으로 초기화된다. 단계 (608, 609)에서는 변수 "frame"이 1 보다 큰 경우 변수 "last\_gop\_stop"이 "frame-1"로 설정된다. 그렇지 않은 경우에는 단계(610)에서 "last\_gop\_stop"이 1로 설정된다. 도 8에서 상세히 설명되는 다음 단계(611)는 모션 파라미터(605), 텍스처/엠티지 파라미터(606), 및 특징점 데이터(607)에 기초하여 트랜스코딩 힌트 상태를 결정한다. 단계(612)에서는 트랜스코딩 힌트 메타데이터가 트랜스코딩 힌트 메타데이터 버퍼로 출력된다. 본 발명의 한 실시예에 따라, 트랜스코딩 힌트 메타데이터는 "nframes" (GOP내에서의 프레임수), 모든 파라미터를 갖는 트랜스코딩 힌트 상태, 및 새로운 GOP의 시작 프레임수 ("new\_gop\_start")를 구비한다. 그 이후에, 변수 "nframes"는 0으로 설정되고, 현재 프레임수 "frame"에는 변수 "new\_gop\_start"가 주어진다. 이어서, 단계(615)에서는 소스 비트스트림의 모든 프레임이 처리되었나를 결정하도록 테스트된다. 그렇지 않은 경우 ("아니오")에는 제어가 단계(614)로 전해져 프레임수가 증가되고, 처리는 단계(602)로부터 시작되어 반복된다. 그렇지 않은 경우에는 처리가 종료된다.



- <57> 도 7은 본 발명의 한 실시예에 따라 새로운 비디오 세그먼트 또는 GOP의 시작 프레임 및 끝 프레임을 결정하는 방법을 설명한다. 단계(701)에서는 도 6으로부터의 변수 "nframes"이 M (I/P 프레임 거리)의 정수 배수인가 여부를 결정한다. 그런 경우, "아니오"가 선택되고, 단계(702)에서는 현재 프레임 번호가 제1 프레임인가 여부를 결정한다. 그런 경우 ("아니오"), 제어는 단계(703)로 진행되어, "nframes"이 GOP내에서 프레임의 최소수 "gop\_min" 보다 큰가 여부를 결정한다. 단계(702)에서 결과가 "예"인 경우, 새로운 GOP는 단계(705)에서 시작된다. 단계(703)에서 결과가 "예"인 경우, 새로운 GOP는 단계(705)에서 시작된다. 단계(703)에서 결과가 "아니오"인 경우, 제어는 단계(704)로 전해지고, "nframes"이 GOP내에서 프레임의 최대수 "gop\_max" 보다 큰가 여부를 결정한다. 단계(704)에서 결과가 "예"인 경우, GOP는 단계(706)에서 폐쇄되고, 그렇지 않은 경우에는 처리가 종료된다.
- <58> 도 8은 본 발명의 한 실시예에 따라 특정한 GOP 또는 A/V 세그먼트가 프레임 당 새로운 특징점의 수만을 고려하여 트랜스코딩 힌트 상태를 선택하는 처리를 설명한다. 설명되는 기본 개념에 기초하여, 상술된 파라메트릭 모션 추정으로부터의 모션 파라미터 뿐만 아니라 DCT-계수로부터 얻어진 텍스처/엡지 파라미터를 이용하여 유사한 결정 구조가 실시될 수 있다. 설명된 등급 또는 알고리즘은 또한 모션, 엡지 액티비티, 프레임 당 새로운 콘텐츠 등에 대해 A/V 자료들을 분류하여 더 높은 레벨의 A/V 분류에 이르게 이용될 수 있음을 주목한다. 이러한 경우, 트랜스코딩 힌트 상태는 다른 자료들의 특정한 등급을 나타낸다. 이제 도 8을 참고로, 단계(801)에서는 변수 "frame\_no", "last\_gop\_start", "sum", 및 "new\_seg"가 초기화된다. 변수 "frame\_no"에는 "last\_gop\_start" 파라미터의 콘텐츠가 주어지고, 변수 "sum" 및 "new\_seg"는 0으로 초기화된다. 이어서, 단계(802)에서는 변수 "sum"의 콘텐츠가 현재 프레임의 새로운 특징점의 수 ("frame\_no") 만큼 증가된다. 단계(803)에서는 변수 "frame\_no"가 변수 "last\_gop\_stop" 보다 작은가 여부를 결정한다. 그런 경우 ("예"), 단계(802)가 반복되고, 그렇지 않은 경우에는 제어가 단계(804)로 전해진다. 단계(804)에서는 변수 "sum"의 값이 소정의 파라미터 "summax"의 1/8 보다 작은지 여부를 결정한다. 파라미터 "summax"는 프레임 "last\_gop\_start" 및 "last\_gop\_stop" 사이의 프레임 수 만큼 곱하여져 프레임에서 프레임으로 트래킹될 수 있는 특징점의 최대수를 나타내는 상수이다. 이는 본 발명의 한 실시예에 따라 값 200을 갖는다. 단계(804)에서의 결과가 "예"이면, 파라미터가 도 8의 도표 1에 도시된 바와 같은 단계(806)에서 트랜스코딩 힌트 상태(1)가 선택된다. 그렇지 않은 경우, 단계(805)에서는 변수 "sum"의 값이 미리 결정된 파라미터 "summax"의 1/4 보다 작은가 여부를 결정한다. 그런 경우 ("예"), 트랜스코딩 힌트 상태(2)는 도표 1에 도시된 바와 같이 단계(807)에서 선택된다. 그렇지 않은 경우 ("아니오")에는 트랜스코딩 힌트 상태(3) (도표 1에 도시된 바와 같이)가 단계(808)에서 선택되고, 처리는 종료된다. 단계 (804, 805)에서의 결정 임계값은 트랜스코딩 힌트 상태의 규정 및 수에 의존함을 주목하여야 한다.
- <59> 트랜스코딩 힌트 메타데이터 디스크립션
- <60> 메타데이터 디스크립션에서는 의사 C-코드 (pseudo C-code) 스타일이 이용될 수 있다. 긴급 MPEG-7 메타데이터 표준에서 규정된 바와 같이, 디스크립션에 대한 약자 D 및 디스크립션 구조(Description Schemes)에 대한 약자 DS가 이용된다.
- <61> 도 9는 본 발명의 한 실시예에 따라 일반 A/V DS(901)내에서 트랜스코딩 힌트 메타데이터의 구조적인 조직을 도시한다. 도 9에 도시된 바와 같이, 세그먼트 DS(904) 및 미디어 정보 DS(902)는 일반 A/V DS(901)로부터 유도된다. 세그먼트 분해(906)는 세그먼트 DS(904)로부터 유도되고, 비디오 세그먼트 DS(907) 및 이동 영역 DS(907)은 세그먼트 분해(906)로부터 유도된다. 도 14를 참고로 상세히 설명될 세그먼트-기반의 트랜스코딩 힌트 DS(909)는 비디오 세그먼트 DS(907)로부터 유도된다. 비디오 세그먼트 DS(907)는 도 16을 참고로 상세히 설명될 하나 또는 수개의 트랜스코딩 힌트 상태 DS(911)를 액세스한다. 이동 영역 DS(908)로부터, 도 14를 참고로 상세히 설명될 세그먼트-기반의 트랜스코딩 힌트 DS(910)는 이동 영역에 대해 유도되고, 도 16을 참고로 상세히 설명될 하나 또는 수개의 트랜스코딩 힌트 상태 DS(912)를 액세스한다. 매체 정보 DS(902)에 대해, 매체 프로파일 DS(903)가 유도된다. 매체 프로파일 DS(903)로부터, 도 10을 참고로 설명될 일반 트랜스코딩 힌트 DS(905)가 유도된다.
- <62> 도 10은 트랜스코딩 힌트 DS(1001)의 구조적인 조직을 도시하고, 이는 도 11을 참고로 설명될 소스 포맷 규정 DS(1002)의 한 예와, 도 12를 참고로 설명될 타겟 포맷 규정 DS(1003)에 대한 하나 또는 수개의 예로 구성된다. 부가하여, 트랜스코딩 힌트 DS(1001)는 도 13을 참고로 설명될 일반 트랜스코딩 힌트 DS(1004)의 한 선택적인 예와, 도 15를 참고로 설명될 한가지 선택적인 트랜스코딩 부호화 복잡도 DS(1005)로 구성된다.
- <63> 도 11은 본 발명의 한 실시예에 따라 전체적인 A/V 콘텐츠 또는 특정한 A/V 세그먼트에 연관된 소스 포맷 규정

트랜스코딩 힌트 메타데이터 (예를 들면, 도 10의 소스 포맷 규정 DS(1002))를 도시한다. 도 11에 도시된 바와 같이, 관련된 디스크립터 및 설명 구조는 다음을 포함할 수 있다:

- <64> \* bitrate는 타입 <int>이고, 소스 A/V 데이터 스트림의 초당 비트 레이트를 설명한다.
- <65> \* size\_of\_pictures는 타입 <2\*int>이고, x 및 y 방향으로 소스 A/V 포맷의 화상 크기를 설명한다.
- <66> \* number\_of\_frames\_per\_second는 타입 <int>이고, 소스 콘텐츠의 초당 프레임의 수를 나타낸다.
- <67> \* pel\_aspect\_ratio는 타입 <float>이고, 펄 중횡비를 설명한다.
- <68> \* pel\_colour\_depth는 타입 <int>이고, 칼라 깊이를 설명한다.
- <69> \* usage\_of\_progressive\_interlaced\_format은 크기 <1 비트>이고, 소스 포맷이 순차 또는 인터레이스 포맷인가 여부를 설명한다.
- <70> \* usage\_of\_frame\_field\_pictures는 크기 <1 비트>이고, 프레임 또는 필드 화상이 이용되는지 여부를 설명한다.
- <71> \* compression method는 타입 <int>이고, 소스 포맷에 이용되는 압축 방법을 규정하고, MPEG-1, MPEG-2, MPEG-4, DV, H.263, H.261 등을 포함하는 리스트로부터 선택될 수 있다. 모든 압축 방법에 대해, 여기서는 파라미터가 더 규정될 수 있다.
- <72> \* GOP\_structure는 I, P, B 상태의 런-길이 (run length) 부호화 데이터 필드이다. 예를 들어, MPEG-2 비디오에 I-프레임만이 있는 경우, 압축 영역에서 DV 포맷으로의 직접적인 변환이 가능하다.
- <73> 도 12는 본 발명의 한 실시예에 따라 전체적인 A/V 콘텐츠 또는 특정한 A/V 세그먼트에 연관된 타겟 포맷 규정 트랜스코딩 힌트 메타데이터를 도시한다. 도 12에 도시된 바와 같이, 관련된 디스크립터 및 디스크립션 구조는 다음을 포함할 수 있다:
- <74> \* bitrate는 타입 <int>이고, 타겟 A/V 데이터 스트림의 초당 비트 레이트를 설명한다.
- <75> \* size\_of\_pictures는 타입 <2\*int>이고, x 및 y 방향으로 타겟 A/V 포맷의 화상 크기를 설명한다.
- <76> \* number\_of\_frames\_per\_second는 타입 <int>이고, 타겟 콘텐츠의 초당 프레임의 수를 나타낸다.
- <77> \* pel\_aspect\_ratio는 타입 <float>이고, 펄 중횡비를 설명한다.
- <78> \* pel\_colour\_depth는 타입 <int>이고, 칼라 깊이를 설명한다.
- <79> \* usage\_of\_progressive\_interlaced\_format은 크기 <1 비트>이고, 타겟 포맷이 순차 또는 인터레이싱될 필요가 있는지 여부를 설명한다.
- <80> \* usage\_of\_frame\_field\_pictures는 크기 <1 비트>이고, 프레임 또는 필드 화상이 이용되는지 여부를 설명한다.
- <81> \* compression method는 타입 <int>이고, 타겟 포맷에 이용되는 압축 방법을 규정하고, MPEG-1, MPEG-2, MPEG-4, DV, H.263, H.261 등을 포함하는 리스트로부터 선택될 수 있다. 모든 압축 방법에 대해, 여기서는 파라미터가 더 규정될 수 있다.
- <82> \* GOP\_structure는 I, P, B 상태의 선택적인 런-길이 부호화 데이터 필드이다. 이 선택적인 파라미터로, 고정된 GOP 구조가 강요될 수 있다. 고정된 GOP 구조는 예를 들어, 특정한 위치에 있는 I-프레임이 비디오 편집을 용이하게 시키는데 유용하다.
- <83> 도 13은 본 발명의 한 실시예에 따라 전체적인 A/V 콘텐츠 또는 특정한 A/V 세그먼트에 연관된 일반 트랜스코딩 힌트 메타데이터 (예를 들면, 도 11의 일반 트랜스코딩 힌트 DS(1004))를 설명한다. 도 13에 도시된 바와 같이, 관련된 디스크립터 및 디스크립션 구조는 다음을 포함할 수 있다:
- <84> \* use\_region\_of\_interest\_DS는 <1 비트>의 길이를 갖고, 관심있는 디스크립션 구조의 영역이 트랜스코딩 힌트로 이용가능한가 여부를 나타낸다.
- <85> \* region\_of\_interest\_DS가 이용되는 경우, motion\_trajectory\_D와 함께 shape\_D (예를 들어 다음 중 하나가 될 수 있는: boundary\_box\_D, MB\_shape\_D, 또는 임의의 다른 shape\_D)는 관심 영역을 공간적 및 시간적으로 설명하는데 이용될 수 있다. MB\_shape\_D는 객체 형상 디스크립션에 매크로블록 (16x16) 크기의 블록을 이용할 수

있다. Motion\_trajectory\_DS는 이미 시간 표시를 포함하므로, region\_of\_interest\_DS의 시작 프레임 및 종료 프레임이 규정될 수 있다. region\_of\_interest\_DS는 각 shape\_DS의 크기 및 각 motion\_trajectory\_DS의 크기를 가질 수 있다. 트랜스코딩 응용에서, region\_of\_interest\_DS는 예를 들어, 배경 보다 관심있는 영역내의 블록에 더 많은 비트를 쓰는데 (또는 양자화기를 각각 수정하는데) 이용될 수 있다. MPEG-4에 대한 또 다른 트랜스코딩 응용은 분리된 MPEG-4 객체에 의해 관심있는 영역을 설명하고 배경과 같이 다른 MPEG-4 보다 관심있는 영역에 더 높은 비트 레이트 및 더 높은 프레임 비율을 쓰는 것이다. region\_of\_interest\_DS의 추출은 자동적으로 또는 수동적으로 실행될 수 있다.

<86> \* use\_editing\_effects\_transcoding\_hints\_DS는 <1 비트>의 길이를 갖고, 정보가 편집 효과에 기초하여 트랜스코딩 힌트에 이용가능한가를 나타낸다.

<87> \* camera\_flash는 카메라 플래쉬가 일어나는 경우 모든 엔트리가 프레임수를 설명하는 엔트리 (entry)의 리스트이다. 그러므로, 디스크립터의 길이는 <int>로 곱하여진 카메라 플래쉬 이벤트의 수이다. 트랜스코딩 응용에서는 Peter Kuhn: "MPEG-4 모션 추정에 대한 알고리즘, 복잡도 분석, 및 VLSI 설계 (Algorithms, Complexity Analysis and VLSI Architectures for MPEG-4 motion estimation)", Kluwer Academic Publishers, 1999에서 설명된 휘도(luminance) 차이에 기초하여 대부분의 비디오 (재) 인코더/트랜스코더가 모션 추정 방법을 이용하므로, camera\_flash 디스크립터는 매우 유용하다. 휘도-기반의 모션 추규정 경우, 2개의 연속적인 프레임 (플래쉬를 가진 것, 플래쉬를 가지지 않은 것)의 2개 매크로블록 사이의 평균 절대 에러는 예측하기에 너무 높고, 카메라 플래쉬를 갖는 프레임이 높은 비트 레이트 비용을 갖는 인트라-프레임으로 부호화되어야 한다. 그러므로, 트랜스코딩 힌트 디스크립터 구조 ("DS")내에서 카메라 플래쉬를 나타내는 것은 예를 들어, 적절한 비트 레이트로 앵커 프레임으로부터 카메라 플래쉬를 갖는 프레임을 예측하는데 휘도 정정 모션 추정 방법 또는 다른 수단을 이용하도록 허용한다.

<88> \* cross\_fading은 모든 엔트리가 크로스페이딩의 시작 프레임 및 종료 프레임을 설명하는 엔트리의 리스트이다. 그러므로, 이 디스크립터의 길이는 크로스페이딩 이벤트의 수의 2배 <int>이다. 트랜스코딩 힌트 메타데이터에서 크로스페이딩 이벤트를 나타내는 것은 크로스페이딩 동안 비트 레이트/양자화기를 제어하는데 매우 유용하다. 크로스페이딩 동안, 예측은 일반적으로 제한된 용도이므로, 예측 에러 코딩에 대해 비트 레이트를 증가시킨다. 크로스페이딩 동안 화면이 통상적으로 흐려지므로, 비트 레이트 증가는 양자화기 스케일, 비트 레이트, 또는 비율 제어 파라미터를 각각 조정함으로써 제한될 수 있다.

<89> \*black\_pictures는 모든 엔트리가 흑색 화상의 시퀀스의 시작 프레임 및 종료 프레임을 설명하는 엔트리의 리스트이다. 화면 사이에는 특히, 가정용 비디오에서 흑색 화상이 일어난다. 실험적으로, 일련의 흑색 화상은 예측이 단지 제한된 용도이기 때문에 모션-보상된 DCT 코드에서 비트 레이트를 증가시키는 것으로 나타난다. 그러므로, 이 트랜스코딩 힌트 디스크립터는 양자화기 스케일, 비트 레이트, 또는 비율 제어 파라미터를 각각 조정함으로써 흑색 화상 동안 비트 레이트를 제한시키는데 이용될 수 있다.

<90> \* fade\_in은 cross\_fading과 유사하고, 페이드인의 시작 프레임 및 종료 프레임을 결정하는 다수의 엔트리로 설명된다. 크로스페이딩과 비교하여, 페이드인은 흑색 화상으로부터 시작하므로, 양자화기 스케일, 비트 레이트, 또는 비율 제어 파라미터를 각각 조정함으로써, 일종의 눈에 대한 마스킹 (masking) 효과가 페이드인 동안 비트 레이트를 제한하는데 이용될 수 있다.

<91> \* fade\_out은 장면 이후에, 일련의 흑색 화상이 설명되는 것을 제외하고 fade\_in과 유사하다.

<92> \* abrupt\_change는 페이딩이 나타나지 않고 어디서 갑작스런 장면 또는 샷 (shot)이 변하는지를 나타내는 타입 <int>의 단일 프레임수의 리스트에 의해 설명된다. 이들 이벤트는 예를 들어, 도 3의 매우 높고 예리한 피크에 의해 표시된다. 이들 피크는 새로운 카메라 샷 또는 장면의 시작을 나타낸다. abrupt\_change 편집 효과는 페이딩 효과와 대조적이다. 두 비디오 세그먼트 사이에 갑작스런 변화가 일어날 때, 사람의 시각적 인지는 새로운 A/V 세그먼트의 상세한 내용을 인식하고 적응하는데 수 millisecond를 필요로 한다. 사람 눈의 이러한 지연 효과는 예를 들어, 장면 또는 샷의 갑작스런 변화 이후에 비트 레이트를 감소시키거나 비디오 세그먼트의 제1 프레임에 대한 양자화기 스케일을 수정하도록 비디오 트랜스코딩에 유리하게 이용될 수 있다.

<93> \* use\_motion\_transcoding\_hints\_DS는 <1 비트>의 길이를 갖고, 모션-관련 트랜스코딩 힌트 메타데이터의 이용을 나타낸다.

<94> \* number of regions는 후속하는 모션-관련 트랜스코딩 힌트 메타데이터가 유효한 영역의 수를 나타낸다.

<95> \* for\_every\_region은 <1 비트> 길이의 필드로, 영역이 직사각형 또는 임의의 형상인가 여부를 나타낸다. 영역

이 임의의 형상인 경우, 영역 디스크립터 (예를 들어, 형상 디스크립터 및 모션 궤도 디스크립터로 구성된)가 이용된다. 직사각형 영역인 경우, 직사각형 영역의 크기가 이용된다. 이 영역내의 모션 필드는 파라메트릭 모션 모델에 의해 설명되고, 이는 프레임의 시퀀스 또는 모든 프레임에 대한 수개의 파라미터에 의해 결정된다. 트랜스코딩에서, 소스 비디오의 실제 모션에 대한 모션 표현은 (재)부호화 부분의 계산적으로 복잡한 모션 추경 검색 영역을 제한하고, 또한 빠르고 효과적인 인터레이스/디인터레이스 (프레임/필드) 변환을 위해 비디오 내의 모션량에 의존하여 GOP (Group of Pictures) 구조를 결정하는데 이용될 수 있다. 모션 표현은 또한 비디오의 크기 변환을 위해 유리하게 이용될 수 있다.

- <96> 도 14는 본 발명의 한 실시예에 따라 일정한 특성을 도시하는 A/V 세그먼트에 대한 (재)부호화/트랜스코딩 셋팅을 결정하는데 이용될 수 있는 세그먼트-기반의 트랜스코딩 힌트 메타데이터 (예를 들면, 도 9의 세그먼트-기반의 트랜스코딩 힌트 DS (909, 910)를 도시한다. 도 14에 도시된 바와 같이, 관련된 디스크립터 및 디스크립션 구조는 다음을 포함할 수 있다:
- <97> \* start\_frame은 타입 <int>이고, A/V 세그먼트의 트랜스코딩 힌트 메타데이터의 시작부에서 프레임수를 설명한다.
- <98> \* nframes는 타입 <int>이고, A/V 세그먼트의 길이를 설명한다.
- <99> \* I\_frame\_location은 A/V 세그먼트내에서 I-프레임의 위치를 설명하는데 몇가지 가능성을 제공한다.
- <100> \* select\_one\_out\_of\_following은 크기 <2 비트>이고, 다음의 4가지 I-프레임 위치 설명 방법 중 하나를 선택한다.
- <101> \* first frame은 크기 <1 비트>이고, 디폴트 (default) I-프레임 위치이다. 이 방법은 제1 프레임만이 A/V 세그먼트의 인트라 프레임으로 또 다른 예측을 위해 앵커로 이용되고 A/V 세그먼트내의 다른 모든 프레임이 P- 또는 B-프레임인 경우 A/V 세그먼트를 설명한다.
- <102> \* List of frames는 A/V 세그먼트내에서 인트라-프레임의 프레임수의 리스트를 제공한다. 이 방법은 A/V 세그먼트내에서 인트라-프레임의 위치를 임의로 설명하도록 허용한다. 이 리스트 내의 k 프레임에 대해, 이 디스크립터의 크기는 <k\*int>이다.
- <103> \* first\_frame\_and\_every\_k\_frames는 타입 <int>이고, 여기서 세그먼트내의 제1 프레임은 인트라이고 k는 A/V 세그먼트내에서 I-프레임의 간격을 설명한다.
- <104> \* no\_I\_frame은 크기 <1 비트>이고, A/V 세그먼트내에서 I-프레임이 이용되지 않는 경우를 설명하고, 이는 A/V 세그먼트의 부호화가 이전 세그먼트에서의 앵커 (인트라-프레임)에 기초할 때 유용하다.
- <105> \* quantizer\_scale은 타입 <int>이고, A/V 세그먼트에 대한 초기 양자화기 스케일값을 설명한다.
- <106> \* target\_bitrate는 타입 <int>이고, A/V 세그먼트에서 초당 타겟 비트 레이트를 설명한다.
- <107> \* target\_min\_bitrate는 크기 <int>이고, A/V 세그먼트에서 초당 최소 타겟 비트 레이트를 설명한다 (선택적).
- <108> \* target\_max\_bitrate는 크기 <int>이고, A/V 세그먼트에서 초당 최대 타겟 비트 레이트를 설명한다 (선택적).
- <109> \* use\_transcoding\_states는 크기 <1 비트>이고, 트랜스코딩 힌트 상태가 A/V 세그먼트에 이용되는지 여부를 설명한다.
- <110> \* transcoding\_state\_nr은 타입 <int>이고, 세그먼트에 대한 트랜스코딩 힌트 메타데이터 상태를 제공한다. 트랜스코딩 힌트 메타데이터 상태는 트랜스코딩 힌트 상태의 테이블에서 엔트리에 대한 포인터이다. 트랜스코딩 힌트 상태의 테이블은 수개의 엔트리를 가질 수 있고, 여기서 새로운 엔트리는 트랜스코딩 힌트 파라미터에 의해 부가 또는 삭제될 수 있다. 단일 트랜스코딩 힌트 상태의 트랜스코딩 힌트 메타데이터는 도 16을 참고로 설명된다.
- <111> \* add\_new\_transcoding\_state는 크기 <1 비트>이고, 연관된 정보를 갖는 새로운 트랜스코딩 상태가 트랜스코딩 힌트 테이블에 추가되어야 하는지 여부를 설명한다. add\_new\_transcoding\_state 신호가 "예"이면, 새로운 트랜스코딩 힌트 상태의 파라미터 리스트가 주어진다. 파라미터 리스트의 크기는 한 트랜스코딩 힌트 상태의 파라미터의 수와 트랜스코딩 힌트 상태의 수에 의해 결정된다.
- <112> \* remove\_transcoding\_state는 트랜스코딩 상태가 제거될 수 있는지 여부를 나타내는 크기 <1 비트>의 플래그이



다. 트랜스코딩 상태가 제거될 수 있는 경우, 제거되는 트랜스코딩 상태의 상태수 (타입 $\text{int}$ )가 주어진다.

- <113> \* use\_encoding\_complexity\_description은 크기 <1 비트>이고, 도 15에서 규정된 바와 같이 보다 상세한 부호화 복잡도 설명이 이용되어야 하는지 여부를 알린다.
- <114> 도 15는 본 발명의 한 실시예에 따라 전체적인 A/V 콘텐츠 또는 특정한 A/V 세그먼트에 연관될 수 있는 코드화 복잡도 트랜스코딩 힌트 메타데이터를 도시한다. 부호화 복잡도 메타데이터는 비율 제어에 이용될 수 있고, 양자화 및 비트 레이트 셋팅을 결정한다.
- <115> \* use\_feature\_points는 크기 <1 비트>이고, 특징점 기반의 복잡도 평가 데이터의 이용을 나타낸다.
- <116> \*select\_feature\_point\_method는 크기 <2 비트>이고, 특징점 방법을 선택한다.
- <117> \* 프레임 당 number\_of\_new\_feature\_points는 도 3에 도시된 바와 같이 프레임 당 새로운 특징점의 수에 대한 리스트를 설명하고, 이는 크기 <nframes\*int>이다. 이 메트릭(metric)은 프레임 당 새로운 콘텐츠의 양을 나타낸다.
- <118> \* feature\_point\_metrics는 한 세그먼트내에서 프레임 당 새로운 특징점에 기초하여 메트릭의 리스트를 설명한다. 메트릭들은 다음 수단으로 <int> 값의 순서 리스트로 나타내진다: 프레임 당 새로운 특징점의 수에 대한 평균치, 최대치, 최소치, 분산, 표준 편차.
- <119> \* use\_equation\_description은 프레임 당 부호화 복잡도의 수식-기반의 디스크립션에 대한<int> 포인터이다.
- <120> \* use\_motion\_description은 크기 <1 비트>이고, 모션-기반의 복잡도 디스크립션의 이용을 나타낸다.
- <121> \* select\_motion\_method는 크기 <4 비트>이고, 모션 설명 방법을 선택한다.
- <122> \* param\_k\_motion은 크기 <nframes\*k\*int>이고, 전반적인 파라메트릭 모션 모델의 모든 단일 프레임에 대해 k 파라미터를 설명한다.
- <123> \* motion\_metrics는 모션 벡터의 크기에 기초하여 전체적인 세그먼트에 대한 메트릭들의 리스트를 설명한다. 메트릭들은 다음 수단으로 <int> 값의 순서 리스트로 나타내진다: 매크로블록 모션 벡터의 평균치, 최대치, 최소치, 분산, 표준 편차.
- <124> \* block\_motion\_field는 m\*m 블록 크기의 모션 필드의 모든 벡터를 설명하고, 크기 <nframes\*int\*size\_x\*size\_y/(m\*m)>이다.
- <125> \* use\_texture\_edge\_metrics는 텍스처 또는 엣지 메트릭들이 이용될 때 설정되는 플래그이고, 크기 <1 비트>이다.
- <126> \* select\_texture\_edge\_metrics는 크기 <4 비트>이고, 다음으로부터 어떤 텍스처 메트릭이 이용되는지를 결정한다.
- <127> \* DCT\_block\_energy는 한 블록의 모든 DCT-계수들의 합이고 프레임 내의 모든 블록에 대해 규정된다. 그 크기는 <size\_y\*size\_x\*nframes\*int/64>이다.
- <128> \* DCT\_block\_activity는 DC 계수 없이 한 블록의 모든 DCT-계수의 합으로 규정된다. 이는 프레임내에서 모든 블록에 대해 규정되고, 크기 <size\_y\*size\_x \*nframes\*int/64>이다.
- <129> \* DCT\_energy\_metric은 각 블록의 각 DCT 에너지에 기초하는 전체적인 세그먼트에 대한 메트릭들의 리스트를 설명한다. 메트릭들은 다음 수단으로 <int> 값의 순서 리스트로 나타내진다: 모든 각기 DCT 에너지 메트릭들의 평균치, 최대치, 최소치, 분산, 표준 편차. 디스크립터의 크기는 <6\*int>이다. 이 디스크립터의 다른 실시는 비디오 세그먼트의 모든 단일 프레임에 대해 DCT 에너지 메트릭을 설명하는 것이다.
- <130> \* DCT\_activity\_metric은 각 블록의 개별 DCT 모션에 기초하여 전체적인 세그먼트에 대한 메트릭들의 리스트를 설명한다. 메트릭들은 다음 수단으로 <int> 값의 순서 리스트로 나타내진다: 모든 각기 DCT 액티비티 메트릭들의 평균치, 최대치, 최소치, 분산, 표준 편차. 디스크립터의 크기는 <6\*int>이다. 이 디스크립터의 다른 실시는 비디오 세그먼트의 모든 단일 프레임에 대해 DCT 액티비티 메트릭을 설명하는 것이다.
- <131> 도 16은 본 발명의 한 실시예에 따라 전체적인 시청각 콘텐츠 또는 특정한 A/V 세그먼트에 연관된 트랜스코딩 힌트 상태 메타데이터를 도시한다. 관련된 디스크립터 및 디스크립션 구조는 다음을 포함할 수 있다:

- <132> \* M은 타입 <int>이고, I-프레임/P-프레임 거리를 설명한다.
- <133> \* bitrate\_fraction\_for\_I는 타입 <float>이고, I 프레임에 이용가능한 A/V 세그먼트에 대해 규정된 비트 레이트 일부를 설명한다.
- <134> \* bitrate\_fraction\_for\_P는 타입 <float>이고, P 프레임에 이용될 수 있는 A/V 세그먼트에 대해 규정된 비트 레이트 일부를 설명한다. B-프레임에 대한 비트 레이트 일부는 100%의 퍼센트에서 나머지 부분이다.
- <135> \* quantizer\_scale\_ratio\_I\_P는 타입 <float>이고, I- 및 P-프레임 사이에서 양자화기 스케일의 관계 (이 세그먼트에 대해 규정된 바와 같은)를 나타낸다.
- <136> \* quantizer\_scale\_ratio\_I\_B는 타입 <float>이고, I- 및 P-프레임 사이에서 양자화기 스케일의 관계 (이 세그먼트에 대해 규정된 바와 같은)를 나타낸다. 비트 레이트 디스크립터 (bitrate\_fraction\_for\_I < bitrate\_fraction\_for\_P), quantizer\_scale\_ratio 디스크립터 (quantizer\_scale\_ratio\_I\_P, quantizer\_scale\_ratio\_I\_B), 또는 다음의 비율-제어 파라미터가 필수적임을 주목한다.
- <137> \* X\_I, X\_P, X\_B는 frame\_vbv\_complexities로 각각 타입 <int>이고, 프레임 기반의 압축 타겟 포맷의 경우에 규정된다 (예를 들면, 도 12). 이들 및 다음의 VBV (Virtual Buffer Verifier) 복잡도 조정은 선택적이고, 소스 콘텐츠 특성 및 타겟 포맷 규정에 따라 비율 제어 구조를 수정하는데 이용될 수 있다.
- <138> \* X\_I top, X\_P top, X\_B top은 상단 필드에 대한 field\_vbv\_complexities로 각각 타입 <int>이고, 필드 기반의 압축 타겟 포맷의 경우에 규정된다 (예를 들면, 도 12).
- <139> \* X\_I bot, X\_P bot, X\_B bot는 바닥 필드에 대한 field\_vbv\_complexities로 각각 타입 <int>이고, 필드 기반의 압축 타겟 포맷의 경우에 규정된다 (예를 들면, 도 12).
- <140> 이와 같이, 선행하는 설명으로부터 명백해지는 것 중에서, 상기에 설명된 목적들이 효과적으로 이루어짐을 볼 수 있고, 본 발명의 의도 및 범위에서 벗어나지 않고 설명된 구조에서 상기 방법을 실행할 때 특정한 변화가 일어날 수 있기 때문에, 상기 설명에 포함되고 첨부된 도면에 도시된 모든 내용은 제한적인 의미가 아니라 설명적인 의미로 해석되도록 의도된다.
- <141> 또한, 다음의 청구항은 여기서 설명된 본 발명의 일반적인 특성 및 특정한 특성과 언어적인 문제로 그에 포함되는 것이라 말하여질 수 있는 본 발명의 범위에 대한 모든 언급을 모두 포함하도록 의도되는 것으로 이해하여야 한다.

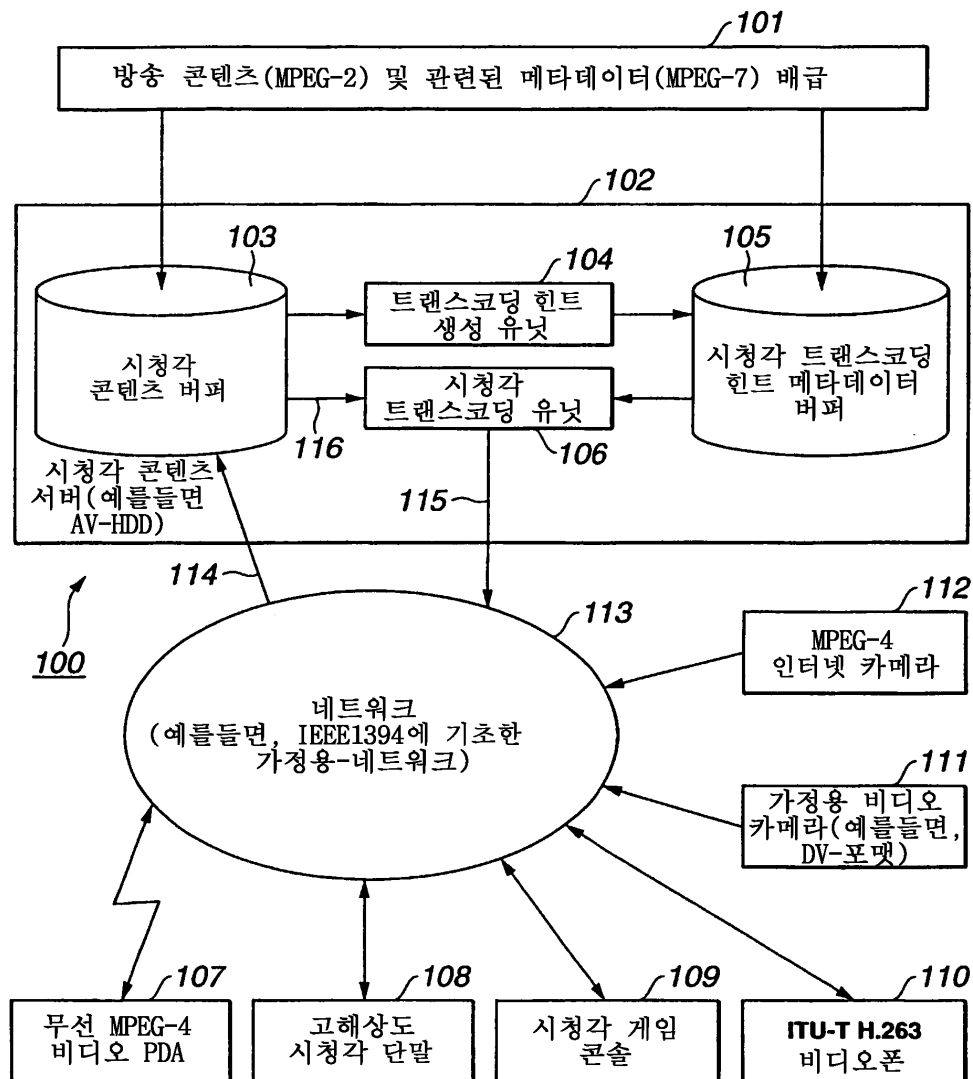
### 도면의 간단한 설명

- <26> 도 1은 본 발명의 한 실시예에 따른 다양한 A/V 단말들을 갖춘 가정용 네트워크에서의 트랜스코딩 시스템(transcoding system)을 도시하는 시스템 개요도.
- <27> 도 2는 본 발명의 한 실시예에 따른 트랜스코딩 힌트 추출(Group of Pictures, "GOP")을 예시하는 도면.
- <28> 도 3은 본 발명의 한 실시예에 따라 프레임 당 새로운 특징점들(feature points)의 수에 따라 트랜스코딩 상태를 선택하는 예를 예시하는 도면.
- <29> 도 4는 본 발명의 한 실시예에 따라 3개의 상태를 갖는 트랜스코딩 힌트 상태도의 예를 도시하는 도면.
- <30> 도 5는 본 발명의 한 실시예에 따라 압축 및 비압축 소스 콘텐츠로부터의 트랜스코딩 힌트 메타데이터의 추출을 예시하는 도면.
- <31> 도 6은 본 발명의 한 실시예에 따라 비디오 분할 및 트랜스코딩 힌트 상태 선택 처리를 도시하는 도면.
- <32> 도 7은 본 발명의 한 실시예에 따라 새로운 비디오 세그먼트 (또는 새로운 GOP)의 경계를 결정하는 방법을 도시하는 도면.
- <33> 도 8은 본 발명의 한 실시예에 따라 트랜스코딩 힌트 상태를 선택하는 방법에 대한 알고리즘을 도시하는 도면.
- <34> 도 9는 본 발명의 한 실시예에 따른 트랜스코딩 힌트 메타데이터의 구조적인 조직의 개요도.
- <35> 도 10은 본 발명의 한 실시예에 따른 일반 트랜스코딩 힌트 메타데이터 디스크립션의 구조적인 조직을 도시하는 도면.

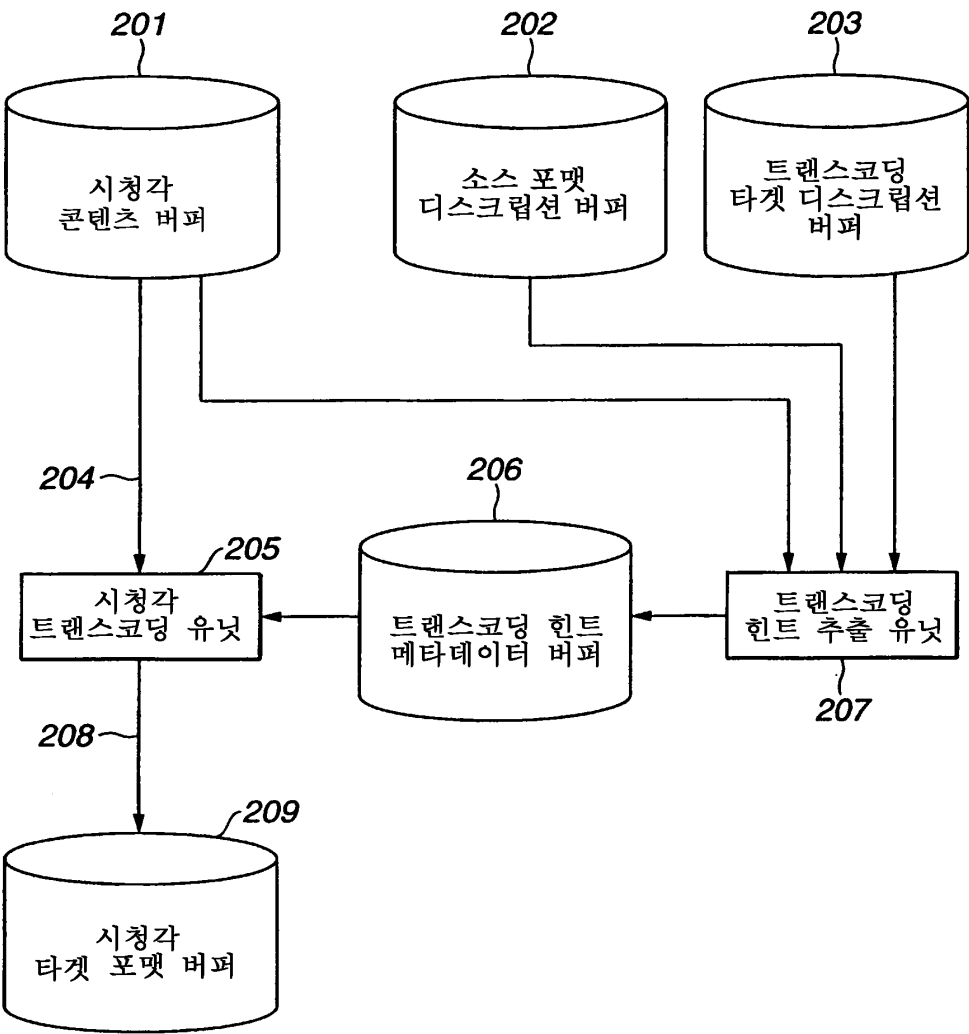
- <36> 도 11은 본 발명의 한 실시예에 따른 소스 포맷 규정에 대한 트랜스코딩 힌트 메타데이터를 도시하는 도면.
- <37> 도 12는 본 발명의 한 실시예에 따른 타겟 포맷 규정에 대한 트랜스코딩 힌트 메타데이터를 도시하는 도면.
- <38> 도 13은 본 발명의 한 실시예에 따른 일반 트랜스코딩 힌트 메타데이터 표현을 도시하는 도면.
- <39> 도 14는 본 발명의 한 실시예에 따른 세그먼트-기반의 트랜스코딩 힌트 메타데이터를 도시하는 도면.
- <40> 도 15는 본 발명의 한 실시예에 따른 부호화 복잡도 트랜스코딩 힌트 메타데이터를 도시하는 도면.
- <41> 도 16은 본 발명의 한 실시예에 따른 트랜스코딩 힌트 상태 메타데이터를 도시하는 도면.

## 도면

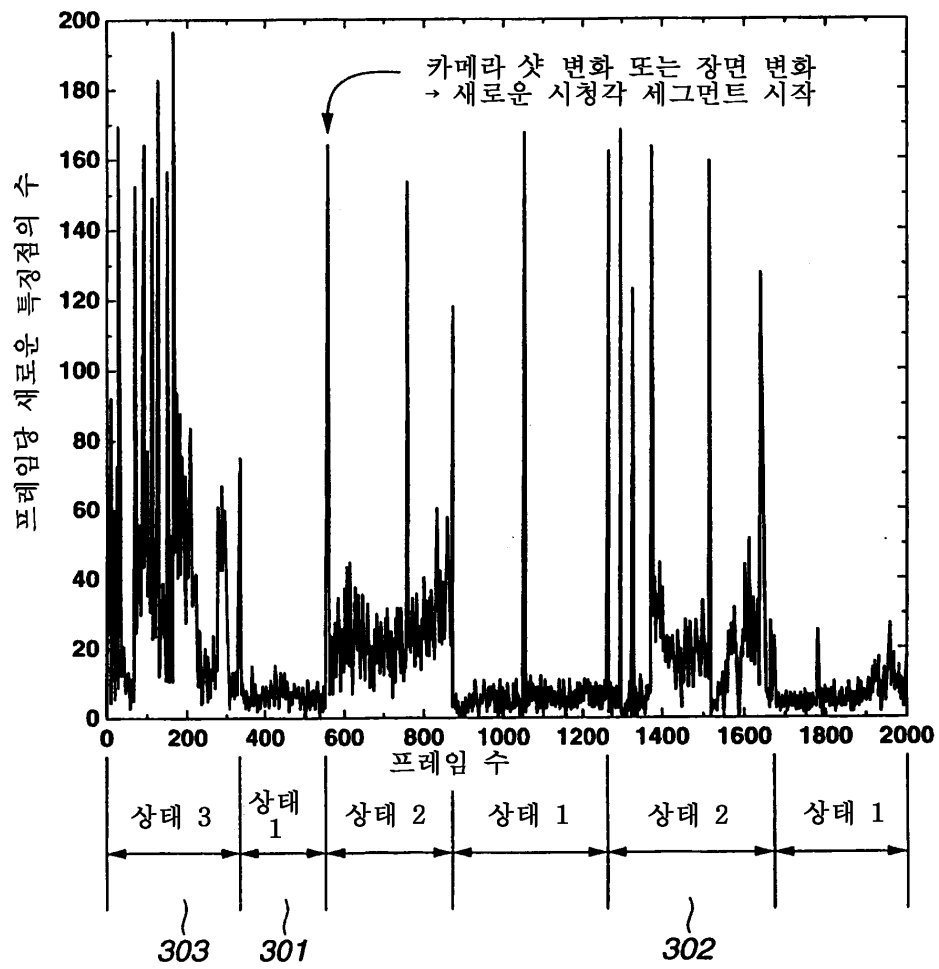
도면1



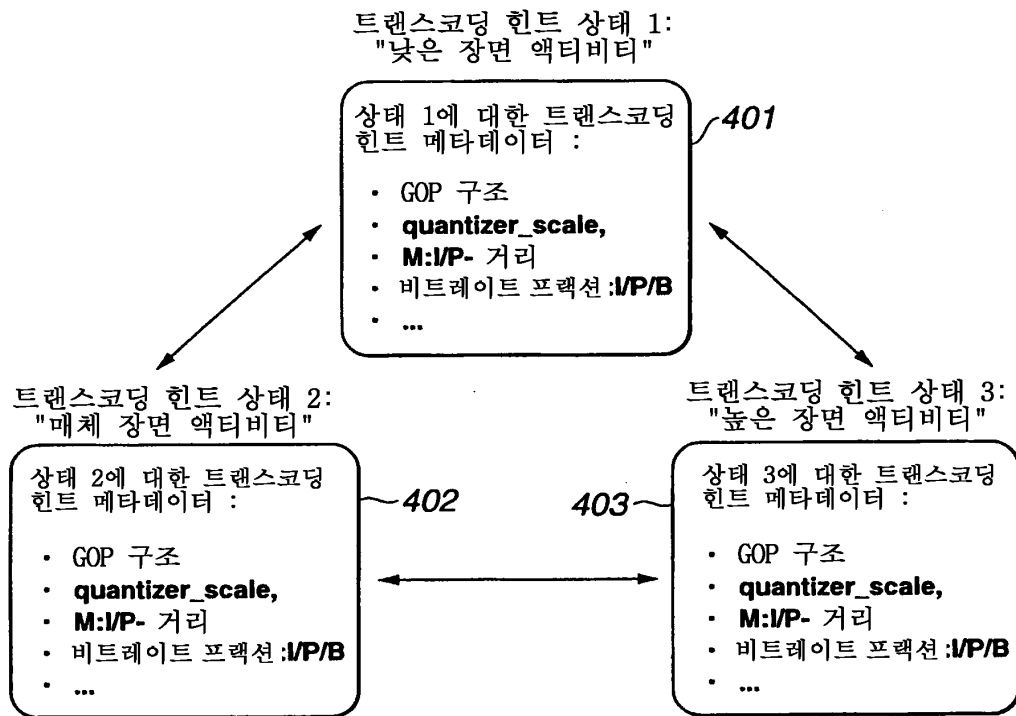
도면2



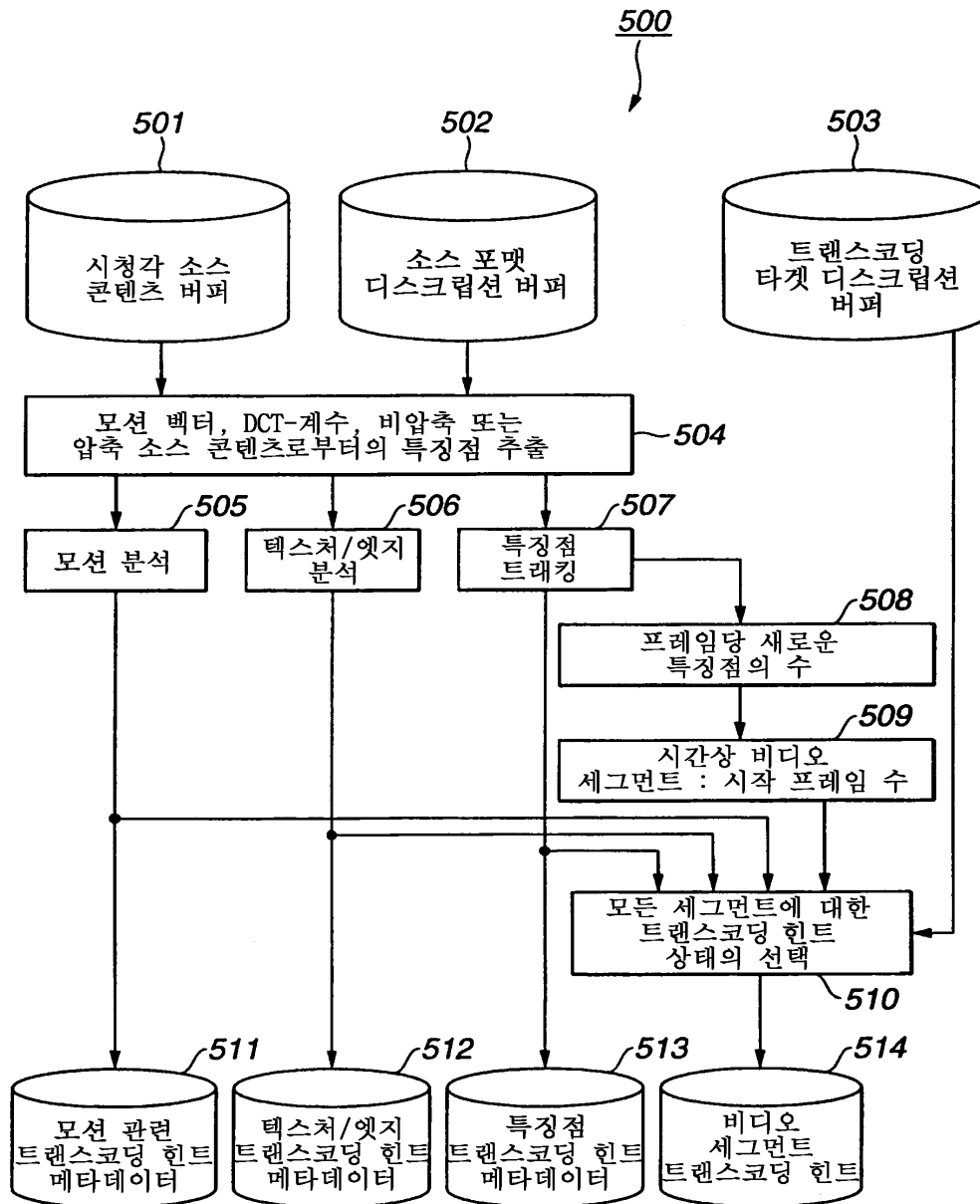
도면3



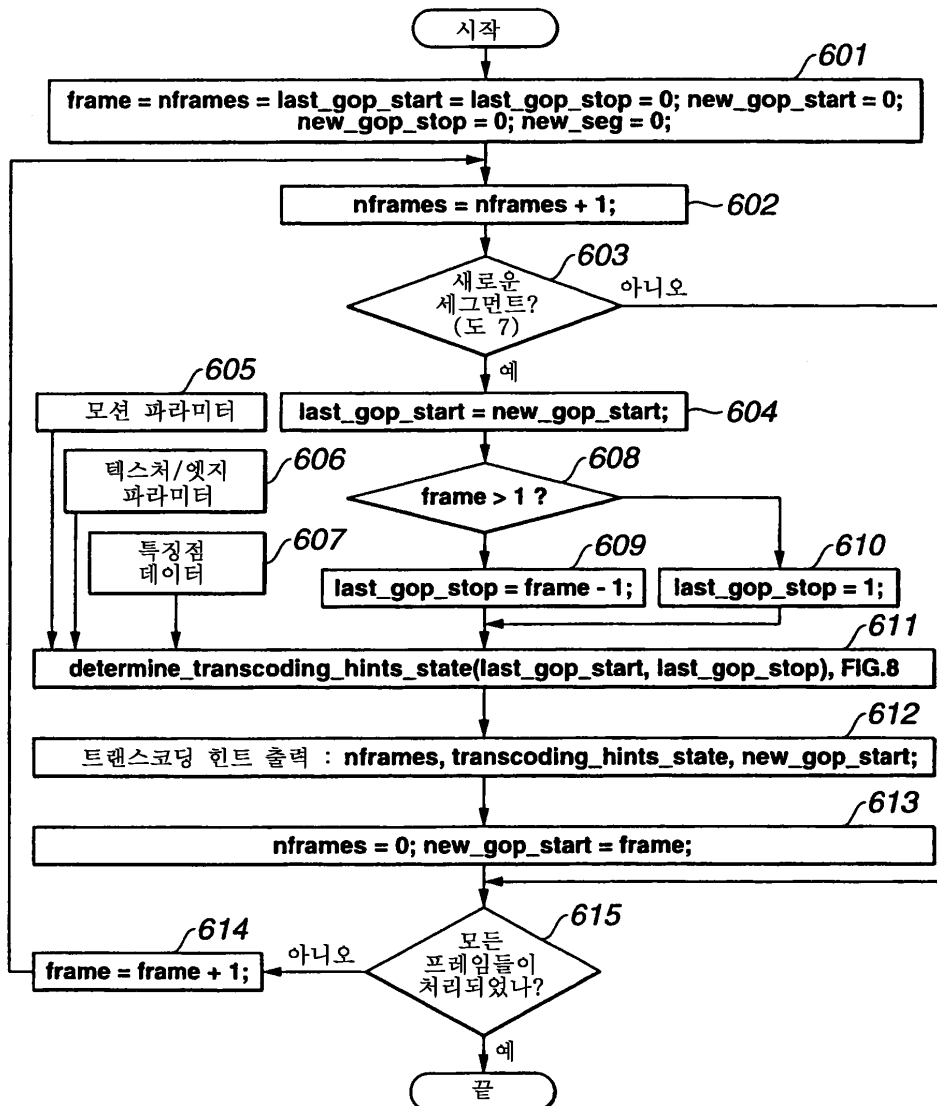
도면4



도면5

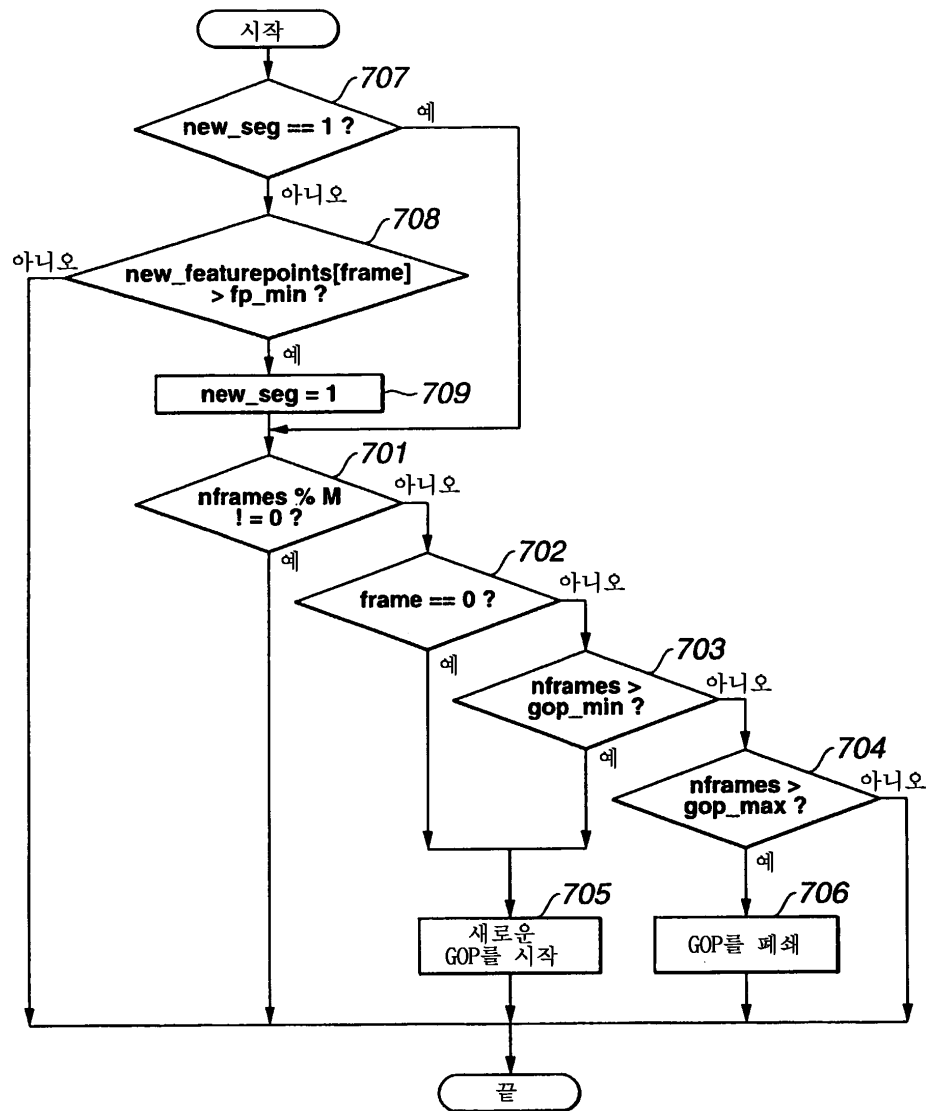


도면6





도면7



도면8

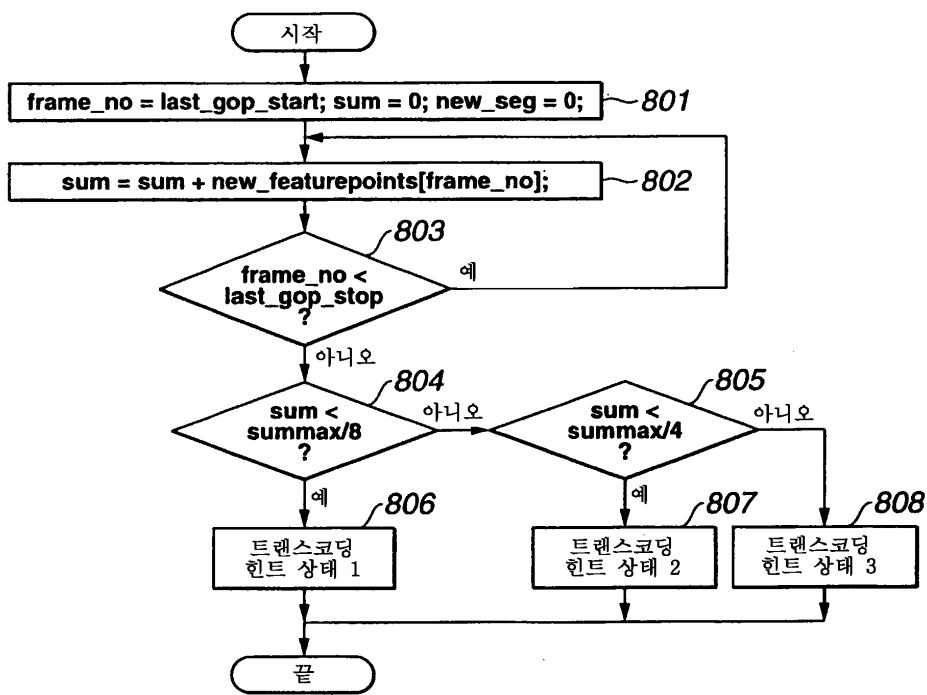
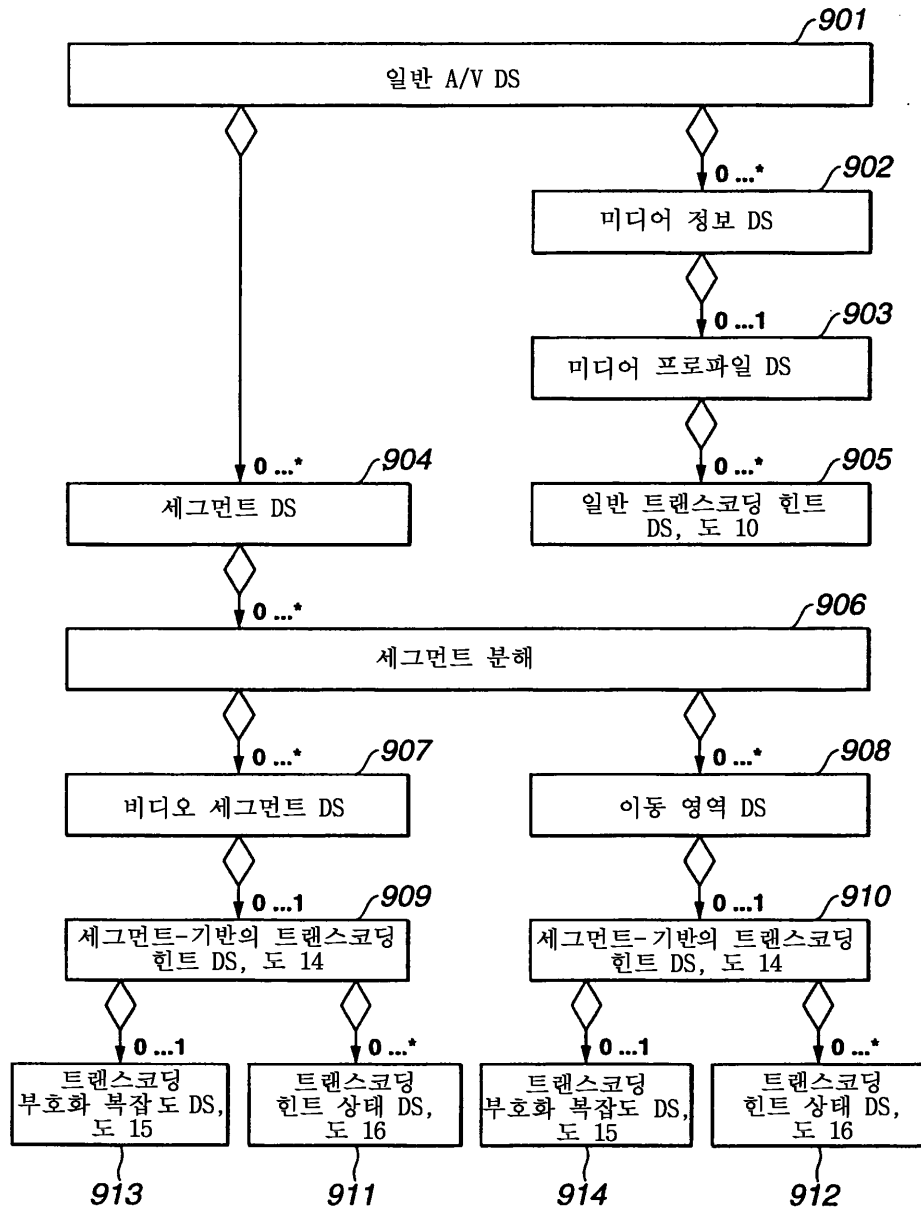


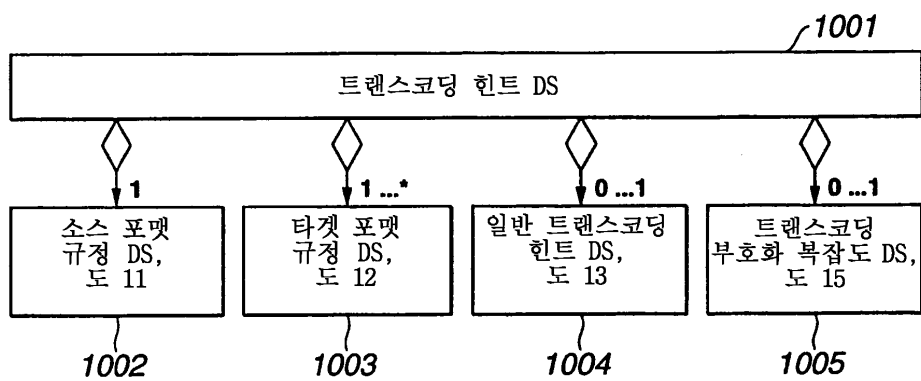
도표 1 : 트랜스코딩 힌트 상태 도표(실례)

상태	M	bitrate_fraction_for_I	bitrate_fraction_for_p
1	5	0.8	0.15
2	4	0.85	0.1
3	3	0.9	0.05

도면9



도면10



도면11

1: bitrate <int>  
 2: size\_of\_pictures <2\*int>  
 3: number\_of\_frames\_per\_second <int>  
 4: pel\_aspect\_ratio <float>  
 5: pel\_colour\_depth <int>  
 6: usage\_of\_progressive\_interlaced\_format <1 bit>  
 7: usage\_of\_frame\_field\_pictures <1bit>  
 8: compression method <int>  
 9:       one out of list {MPEG-1, MPEG-2, MPEG-4, DV, H.263, H.261, .... }  
 10:       { further parameters for compression method }  
 11:       GOP\_structure (Runlength coding)

도면12

1: bitrate <int>  
 2: size\_of\_pictures <2\*int>  
 3: number\_of\_frames\_per\_second <int>  
 4: pel\_aspect\_ratio <float>  
 5: pel\_colour\_depth <int>  
 6: usage\_of\_progressive\_interlaced\_format <1 bit>  
 7: usage\_of\_frame\_field\_pictures <1bit>  
 8: compression\_method <int>  
 9:       one\_out\_of\_list {MPEG-1, MPEG-2, MPEG-4, DV, H.263, H.261, .... }  
 10:       { further parameters for compression method }  
 11:       GOP\_structure (Runlength coding)

도면13

```

1: use_region_of_interest_DS: <1bit>
2:     region_of_interest_DS:
3:         shape_D: select one of {boundary_box_D, MB_shape_D, shape_D}
4:         motion_trajectory_D
5:
6: use_editing_effects_transcoding_hints_DS: <1bit>
7:     camera_flash {frame1, frame2, .... framek} <k*int>
8:     cross_fading {(start_frame, end_frame), ...} <k*(int, int)>
9:     black_pictures {(start_frame, end_frame), ...} <k*(int, int)>
10:    fade_in {(start_frame, end_frame), ...} <k*(int, int)>
11:    fade_out {(start_frame, end_frame), ...} <k*(int, int)>
12:    abrupt_change {frame1, frame2, .... framek} <k*int>
13:
14: use_motion_transcoding_hints_DS: <1 bit>
15:     number_of_regions: <int>
16:     for_every_region:
17:         is_region_rectangular_shaped (y/n) : <1bit>
18:         if_arbitrarily_shaped: use region D for this region
19:         describe_parametric_object_motion_for_this_region

```

도면14

```

1: start_frame <int>
2: nframes <int>
3: I_frame_location:
4:     select_one_out_of_the_following: <2 bit>
5:         first_frame (default)
6:         list_of_frames {frame1, frame2, ...., framek} <k*int>
7:         first_frame_and_every_k_frames <int>
8:         no_I_frame
9: quantizer_scale <int>
10: target_bitrate <int>
11: target_min_bitrate <int>
12: target_max_bitrate <int>
13: use_transcoding_states (y/n) <1 bit>
14: transcoding_state_nr <int>
15: add_new_transcoding_state (y/n) <1bit>
16:     if yes: {list of parameters}
17: remove_transcoding_state (y/n) <1bit>
18:     if yes: state_nr <int>
19: use_encoding_complexity_description (y/n) <1 bit>
20:     if yes: encoding_complexity_description_scheme

```

도면15

1:	use_feature_points (y/n) <1bit>
2:	select_feature_point_method <2 bits>
3:	number_of_new_feature_points <nframes * int>
4:	feature_point_metrics {mean, max, min, var, stddev} <5* int>
5:	use_equation_description (y/n) <1bit>
6:	use_motion_description (y/n) <1bit>
7:	select_motion_method <4 bits>
8:	param_k_motion <nframes * k * int>
9:	motion_metrics {min, max, sum, var, stddev} <5*int>
10:	block_motion_field < nframes*int*size_x*size_y / (m*m) >
11:	use_texture_edge_metrics (y/n) <1bit>
12:	select_texture_edge_metrics <4 bits>
13:	DCT_block_energy <size_y*size_x*nframes*int/64>
14:	DCT_block_activity <size_y*size_x*nframes*int/64>
15:	DCT_energy_metric {mean, min, max, sum, var, stddev} <6*int>
16:	DCT_activity_metric {mean, min, max, sum, var, stddev} <6*int>

도면16

1:	M: I/P distance <int>
2:	bitrate_fraction_for_I <float>
3:	bitrate_fraction_for_P <float> /* bitrate_fraction of B is rest to 100 %)
4:	quantizer_scale_ratio_I_P <float>
5:	quantizer_scale_ratio_I_B <float>
6:	if_frame: /* see target format transcoding hints */
7:	X_I, X_P, X_B <3*int> /* frame_vbv_complexities */
8:	if_field:
9:	X_I_top, X_P_top, X_B_top <3*int> /* field_top_vbv_complexities */
10:	X_I_bot, X_P_bot, X_B_bot <3*int> /* field_bottom_vbv_complexities */