



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2010년05월27일
(11) 등록번호 10-0959573
(24) 등록일자 2010년05월17일

(51) Int. Cl.
H04N 7/24 (2006.01) H04N 7/50 (2006.01)
(21) 출원번호 10-2006-7008830(분할)
(22) 출원일자(국제출원일자) 2003년01월22일
심사청구일자 2008년01월21일
(85) 번역문제출일자 2006년05월04일
(65) 공개번호 10-2006-0069527
(43) 공개일자 2006년06월21일
(62) 원출원 특허 10-2004-7011467
원출원일자(국제출원일자) 2003년01월22일
심사청구일자 2005년01월21일
(86) 국제출원번호 PCT/FI2003/000053
(87) 국제공개번호 WO 2003/063505
국제공개일자 2003년07월31일
(30) 우선권주장
10/306,942 2002년11월29일 미국(US)
20020127 2002년01월23일 핀란드(FI)
(56) 선행기술조사문헌
W0200147283 A1
Multimedia Signal Processing, 1998 IEEE
Workshop, Dec. 1998, S. Wenger, "Temporal
scalability using P-pictures for low-latency
applications", pp. 559-564.

(73) 특허권자
노키아 코퍼레이션
핀란드핀-02150 에스푸 카일알라텐티에 4
(72) 발명자
한누크셀라 미스카
핀란드 핀-33710 탐페레 쿡카니이틴카투 4 비
(74) 대리인
리앤목특허법인

전체 청구항 수 : 총 29 항

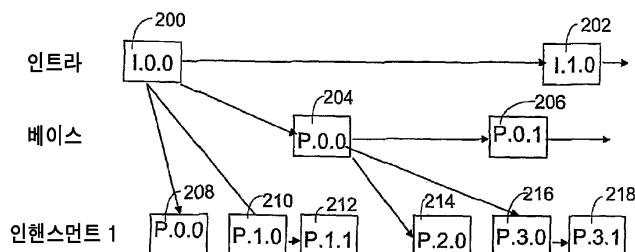
심사관 : 조재신

(54) 비디오 코딩시 이미지 프레임들의 그룹핑

(57) 요약

본 발명은 압축 비디오 시퀀스 디코딩 방법에 관한 것으로, 비디오 시퀀스의 이미지 프레임들은 디코딩과 관련된 버퍼 메모리로 입력된다. 비디오 시퀀스는 이미지 프레임들의 넘버링시의 적어도 한 번의 불연속과 관련된 표시를 포함하고, 그 표시는 비디오 시퀀스로부터 복호화된다. 그런 다음, 표시에 반응하여, 버퍼 메모리가 이미지 프레임의 넘버링시의 불연속에 대응되는 이미지 프레임들을 제공하도록 설정되고, 버퍼 메모리에 의해 제공되는 이미지 프레임들이 디코딩 프로세스시 사용된다. 표시는 비디오 시퀀스 내 이미지 프레임들의 넘버링시 적어도 한 번의 불연속이 고의적인 것임을 알려주는 것으로, 버퍼 메모리에 의해 제공되는 이미지 프레임들이 디코더에 존재하지 않는 이미지 프레임들 대신 사용된다.

대표도 - 도2



특허청구의 범위

청구항 1

디코딩과 관련된 버퍼 메모리로 이미지 프레임들이 입력되는, 압축된 비디오 시퀀스를 디코딩하는 방법에 있어서,

비디오 시퀀스로부터 이미지 프레임들의 넘버링 (번호매김; numbering)에 있어 적어도 한 번의 불연속과 관련한 표시를 복호화하는 단계;

상기 표시에 반응하여, 버퍼 메모리가 상기 이미지 프레임의 넘버링시의 불연속에 대응되는 이미지 프레임들을 제공하도록 설정하는 단계; 및

디코딩 프로세스시 상기 버퍼 메모리 내 이미지 프레임들을 사용하는 단계를 포함함을 특징으로 하는 디코딩 방법.

청구항 2

제1항에 있어서, 상기 표시는, 상기 비디오 시퀀스 내 이미지 프레임 넘버링시의 적어도 한 불연속이 고의적인 것임을 알리는 것임을 특징으로 하는 디코딩 방법.

청구항 3

제1항 또는 제2항에 있어서,

이미지 프레임들의 넘버링시의 불연속으로 인해 디코더에 존재하지 않는 이미지 프레임들 대신, 상기 버퍼 메모리 내 이미지 프레임들을 이용하는 단계를 더 포함함을 특징으로 하는 디코딩 방법.

청구항 4

제1항 또는 제2항에 있어서,

상기 버퍼 메모리 내 이미지 프레임들은, 이미지 프레임들에 대한 슬라이딩 윈도우 버퍼링 (sliding window buffering)에 사용됨을 특징으로 하는 디코딩 방법.

청구항 5

제1항 또는 제2항에 있어서,

상기 버퍼 메모리가 이미지 프레임들의 넘버링시의 불연속에 대응되는 이미지 프레임들을 제공하도록 설정된 다음, 해당 비디오 시퀀스에 구비된 이미지 프레임들의 상기 버퍼 메모리의 입력이 올바른 이미지 프레임 번호부터 재개하는 단계를 더 포함함을 특징으로 하는 디코딩 방법.

청구항 6

제1항 또는 제2항에 있어서,

상기 버퍼 메모리로, 상기 이미지 프레임들의 넘버링시의 불연속에 대응되는 보충(filler) 프레임들을 입력하는 단계를 더 포함함을 특징으로 하는 디코딩 방법.

청구항 7

제6항에 있어서, 상기 보충 프레임들은 실제 비디오 시퀀스에 속하지 않는다는 것을 나타내는 식별자에 의해 표시됨을 특징으로 하는 디코딩 방법.

청구항 8

제1항 또는 제2항에 있어서, 상기 버퍼 메모리가 상기 이미지 프레임의 넘버링시의 불연속에 대응되는 이미지 프레임들을 제공하도록 설정하는 단계는 메모리 표시에 따라 수행되고, 그에 따라 어떤 데이터도 상기 버퍼 메모리로 입력되지 않음을 특징으로 하는 디코딩 방법.

청구항 9

제3항에 있어서,

디코딩 프로세스시 이미지 프레임들의 넘버링시의 불연속에 대응하는 한 이미지 프레임에 대한 참조를 기반으로, 상기 압축된 비디오 시퀀스 내 에러를 검출하는 단계를 더 포함함을 특징으로 하는 디코딩 방법.

청구항 10

디코딩과 관련하여 이미지 프레임들이 입력되는 버퍼 메모리에 연결되어, 압축된 비디오 시퀀스를 복호화하는 비디오 디코더에 있어서,

비디오 시퀀스로부터 이미지 프레임들의 넘버링 (번호매김; numbering)에 있어 적어도 한 번의 불연속과 관련한 표시를 복호화하고;

상기 표시에 반응하여, 버퍼 메모리가 상기 이미지 프레임의 넘버링시의 불연속에 대응되는 이미지 프레임들을 제공하도록 설정하고;

디코딩 프로세스시 상기 버퍼 메모리 내 이미지 프레임들을 사용하도록 구성됨을 특징으로 하는 비디오 디코더.

청구항 11

제10항에 있어서, 상기 표시는, 상기 비디오 시퀀스 내 이미지 프레임 넘버링시의 적어도 한 불연속이 고의적인 것임을 알리는 것임을 특징으로 하는 비디오 디코더.

청구항 12

제10항 또는 제11항에 있어서,

이미지 프레임들의 넘버링시의 불연속으로 인해 디코더에 존재하지 않는 이미지 프레임들 대신, 상기 버퍼 메모리 내 이미지 프레임들을 이용하도록 더 구성됨을 특징으로 하는 비디오 디코더.

청구항 13

제10항 또는 제11항에 있어서,

상기 버퍼 메모리 내 이미지 프레임들은, 이미지 프레임들에 대한 슬라이딩 윈도우 버퍼링 (sliding window buffering)에 사용됨을 특징으로 하는 비디오 디코더.

청구항 14

제10항 또는 제11항에 있어서,

상기 버퍼 메모리가 이미지 프레임들의 넘버링시의 불연속에 대응되는 이미지 프레임들을 제공하도록 설정된 다 음, 해당 비디오 시퀀스에 구비된 이미지 프레임들의 상기 버퍼 메모리로의 입력을 올바른 이미지 프레임 번호부터 재개하도록 더 구성됨을 특징으로 하는 비디오 디코더.

청구항 15

제10항 또는 제11항에 있어서,

상기 버퍼 메모리로, 상기 이미지 프레임들의 넘버링시의 불연속에 대응되는 보충(filler) 프레임들을 입력하도록 더 구성됨을 특징으로 하는 비디오 디코더.

청구항 16

제15항에 있어서, 상기 보충 프레임들은 실제 비디오 시퀀스에 속하지 않는다는 것을 나타내는 식별자에 의해 표시됨을 특징으로 하는 비디오 디코더.

청구항 17

제10항 또는 제11항에 있어서, 상기 메모리 표시에 따라 상기 버퍼 메모리가 상기 이미지 프레임의 넘버링시의 불연속에 대응되는 이미지 프레임들을 제공하도록 설정하도록 더 구성되고, 그에 따라 어떤 데이터도 상기 버퍼

메모리로 입력되지 않음을 특징으로 하는 비디오 디코더.

청구항 18

비디오 시퀀스의 이미지 프레임들이 디코딩과 관련되는 버퍼 메모리에 입력될 때, 압축된 비디오 시퀀스를 디코딩하기 위해, 데이터 처리 장치를 통해 실행되는 컴퓨터 프로그램을 포함하는 컴퓨터 판독가능 저장매체에 있어서,

상기 컴퓨터 프로그램은,

비디오 시퀀스로부터 이미지 프레임들의 넘버링 (번호매김; numbering)에 있어 적어도 한 번의 불연속과 관련한 표시를 복호화하기 위한 컴퓨터 프로그램 코드;

상기 표시에 반응하여, 상기 이미지 프레임의 넘버링시의 불연속에 대응되는 이미지 프레임들을 제공하도록 버퍼 메모리를 설정하기 위한 컴퓨터 프로그램 코드; 및

디코딩 프로세스시 상기 버퍼 메모리 내 이미지 프레임들을 포함하기 위한 컴퓨터 프로그램 코드를 포함함을 특징으로 하는 컴퓨터 판독가능 저장매체.

청구항 19

제18항에 있어서, 상기 표시는, 상기 비디오 시퀀스 내 이미지 프레임 넘버링시의 적어도 한 불연속이 고의적인 것임을 알리는 것임을 특징으로 하는 컴퓨터 판독가능 저장매체.

청구항 20

제18 또는 제19항에 있어서,

이미지 프레임들의 넘버링시의 불연속으로 인해 디코더에 존재하지 않는 이미지 프레임들 대신, 상기 버퍼 메모리 내 이미지 프레임들을 포함하도록 하는 컴퓨터 프로그램 코드를 더 포함함을 특징으로 하는 컴퓨터 판독가능 저장매체.

청구항 21

제18 또는 제19항에 있어서,

상기 버퍼 메모리 내 이미지 프레임들을, 이미지 프레임들에 대한 슬라이딩 윈도우 버퍼링 (sliding window buffering)에 사용하도록 하는 컴퓨터 프로그램 코드를 더 포함함을 특징으로 하는 컴퓨터 판독가능 저장매체.

청구항 22

제18 또는 제19항에 있어서,

상기 버퍼 메모리가 이미지 프레임들의 넘버링시의 불연속에 대응되는 이미지 프레임들을 제공하도록 설정된 다음, 해당 비디오 시퀀스에 구비된 이미지 프레임들의 상기 버퍼 메모리로의 입력을 올바른 이미지 프레임 번호부터 재개하기 위한 컴퓨터 프로그램 코드를 더 포함함을 특징으로 하는 컴퓨터 판독가능 저장매체.

청구항 23

제18 또는 제19항에 있어서,

상기 버퍼 메모리로, 상기 이미지 프레임들의 넘버링시의 불연속에 대응되는 보충(filler) 프레임들을 입력하게 하는 컴퓨터 프로그램 코드를 더 포함함을 특징으로 하는 컴퓨터 판독가능 저장매체.

청구항 24

제23항에 있어서,

상기 보충 프레임들이 실제 비디오 시퀀스에 속하지 않는다는 것을 나타내는 식별자로서 상기 보충 프레임들을 나타내도록 하는 컴퓨터 프로그램 코드를 더 포함함을 특징으로 하는 컴퓨터 판독가능 저장매체.

청구항 25

제18 또는 제19항에 있어서,

상기 메모리 표시에 따라 상기 이미지 프레임의 넘버링시의 불연속에 대응되는 이미지 프레임들을 제공하도록 상기 버퍼 메모리를 설정하여, 그에 따라 어떤 데이터도 상기 버퍼 메모리로 입력되지 않도록 하는 컴퓨터 프로그램 코드를 더 포함함을 특징으로 하는 컴퓨터 판독가능 저장매체.

청구항 26

비디오 시퀀스를 부호화하는 비디오 인코더에 있어서,

이미지 프레임들의 넘버링시 불연속을 일으키도록 적어도 하나의 이미지 프레임이 비디오 시퀀스로부터 배제되도록 비디오 시퀀스를 부호화하고;

상기 비디오 시퀀스 안에 이미지 프레임들의 넘버링 불연속과 관련한 표시를 부호화하도록 구성됨을 특징으로 하는 비디오 인코더.

청구항 27

제26항에 있어서, 상기 표시는, 상기 비디오 시퀀스 내 이미지 프레임 넘버링시의 불연속이 고의적인 것임을 알리는 것임을 특징으로 하는 비디오 인코더.

청구항 28

부호화된 비디오 시퀀스를 수신하고, 비디오 시퀀스의 비트 레이트를 조정하는 장치에 있어서,

상기 비디오 시퀀스로부터 적어도 한 개의 이미지 프레임을 배제하여, 이미지 프레임들의 넘버링시 불연속을 일으키도록 됨을 특징으로 하는 장치.

청구항 29

제28항에 있어서,

상기 표시는, 상기 비디오 시퀀스 내 이미지 프레임 넘버링시의 불연속이 고의적인 것임을 알리는 것임을 특징으로 하는 장치.

청구항 30

삭제

청구항 31

삭제

명세서

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

[0009] 본 발명은 멀티미디어 파일들의 그룹핑에 관한 것으로, 보다 상세하게는 특히 스트리밍과 관련한 비디오 파일들의 그룹핑에 관한 것이다.

[0010] '스트리밍'이라는 용어는 데이터의 보통 오디오 및 비디오 파일들과 같은 멀티미디어 데이터인 데이터의 동시적 송신 및 재생을 일컫는 것으로, 수신자는 전송될 모든 데이터가 수신되기 전에 이미 데이터 재생을 시작한다. 멀티미디어 데이터 스트리밍 시스템은 스트리밍 서버와, 일반적으로 통신 네트워크를 통해 스트리밍 서버로의 데이터 접속을 설정하기 위해 수신자가 이용하는 단말 장치들을 포함한다. 스트리밍 서버로부터 수신자들은 저장되었거나 실시간인 멀티미디어 데이터를 검색하며, 그리고 나서 그 멀티미디어 데이터의 재생이, 단말에 포함된 스트리밍 어플리케이션에 의해, 가장 바람직하게는 데이터 전송 실시간으로 시작될 수 있다.

- [0011] 스트리밍 서버의 관점에서 보면, 스트리밍은 단말로의 일반 스트리밍(normal streaming)이나 연속 다운로드(progressive downloading)으로서 이뤄질 수 있다. 일반 스트리밍시 멀티미디어 데이터 및/또는 데이터 콘텐츠의 전송은, 전송 비트율이 단말 장치의 재생 레이트와 실질적으로 일치한다는 것을 보장하거나, 전송에 사용된 통신 네트워크가 데이터 전송의 애로사항을 야기할 때 그 전송 비트율이 통신 네트워크에 사용 가능한 대역폭과 실질적으로 일치함을 보장함으로써 제어된다. 연속 다운로드시 멀티미디어 및/또는 데이터 콘텐츠의 전송은 반드시 조금이라도 방해 받을 필요가 없지만, 멀티미디어 파일들은 보통 전송 프로토콜 플로우 제어에 의해 수신자에게 전송된다. 그리고 나서 단말들은 서버로부터 전송된 데이터의 동일 사본을 수신, 저장 및 재생하며, 그러면 그 사본은 통신 네트워크를 통해 다시 스트리밍을 시작할 필요 없이 이 단말에서 나중에 다시 재생될 수 있다. 단말에 저장된 멀티미디어 파일들은 그러나 보통 매우 크기 때문에, 그 파일들의 단말로의 전송은 시간을 소모하게 되며, 거대한 저장 메모리 용량을 필요로 한다. 이것이 흔히 일반 스트림이 선호되는 이유가 된다.
- [0012] 멀티미디어 파일 내 비디오 파일들은 수많은 정지 영상 프레임들을 구비하며, 이들은 연달아(보통 초당 15에서 30 프레임씩) 빠르게 디스플레이 되어져 동영상의 효과를 만든다. 이미지 프레임들은 보통 실질적으로 변함 없는 상태를 유지하는 이미지 정보에 의해 결정되는 다수의 정적 배경 오브젝트들과, 어느 정도 변화하는 이미지 정보에 의해 결정되는 약간의 동적 오브젝트들을 구비한다. 연속적으로 디스플레이되는 이미지 프레임에 의해 구성되는 정보는 일반적으로 매우 유사하다. 즉, 일련의 이미지 프레임들은 상당한 량의 여분(redundancy)을 포함한다. 비디오 파일들에 나타나는 여분은 공간적, 시간적, 및 스펙트럼적 여분들로 나뉜다. 공간적 여분은 인접한 이미지 픽셀들의 상호 상관과 관련되고, 시간적 여분은 잇따른 프레임들 내 특정 이미지 오브젝트들에서 발생하는 변화와 관련되며, 스펙트럼적 여분은 한 이미지 프레임 내 상이한 컬러 성분들의 상관과 관련된다.
- [0013] 비디오 파일들 내의 데이터 량을 줄이기 위해, 이미지 데이터는 이미지 프레임들의 여분 정보량을 줄임으로써 보다 작은 형태로 압축될 수 있다. 또, 부호화 중에, 현재 사용되는 대부분의 비디오 인코더들은 비디오 정보에서 덜 중요한 이미지 프레임 섹션의 화질을 저하시킨다. 또, 많은 비디오 코딩 방법들이 이미지 데이터로부터 코딩된 비트 스트림의 여분이 VLC(가변장 코딩)라고 알려진 압축 패러미터들의 효율적이고도 손실없는 코딩에 의해 감축되는 것을 가능하게 한다.
- [0014] 또, 많은 비디오 코딩 방법들은 일련의 이미지 프레임들의 상술한 시간적 여분을 활용한다. 이 경우, 모션-보상된 시간 예측이라고 알려진 방식이 사용된다. 즉, 한 비디오 시퀀스 내 이미지 프레임들 중 일부(보통 대부분)의 콘텐츠들이, 잇따른 이미지 프레임들의 특정 오브젝트들 또는 영역들의 변화 추적에 의해 그 시퀀스 내 다른 프레임들로부터 예측되어진다. 비디오 시퀀스는, 이미지 정보가 모션-보상된 시간적 예측을 사용해 결정되어지지 않은 약간의 압축 이미지 프레임들을 항상 구비한다. 이러한 프레임들을 INTRA(인트라)-프레임들, 또는 I-프레임들이라 부른다. 이와 마찬가지로, 이전 이미지 프레임들로부터 예측된 모션-보상 비디오 시퀀스 이미지 프레임들을 INTER(인터)-프레임들, 또는 P-프레임들(Predicted)이라 부른다. P-프레임들의 이미지 정보는 하나의 I-프레임과 가능하면 하나 이상의 이전에 코딩된 P-프레임들을 이용해 정해진다. 한 프레임을 잃으면, 그 프레임에 종속된 프레임들이 더 이상 올바르게 디코딩될 수 없다.
- [0015] 한 I-프레임이 보통 화상 그룹(Group of Pictures, GOP)이라고 규정된 비디오 시퀀스를 일으키며, 그것의 P-프레임들은 오직 상기 I-프레임과 해당 GOP의 이전 P-프레임들에 기초해 정해질 수 있다. 다음 I-프레임은 새로운 화상 그룹(GOP)을 시작하며, 그에 따라 구비된 이미지 정보는 이전 GOP의 프레임들에 기초해 정해질 수 없다. 즉, 화상 그룹들은 시간적으로 오버랩(중복)되지 않으며, 각 화상 그룹은 따로 따로 디코딩 될 수 있다는 것이다. 또, 많은 비디오 압축 방식들이 양방향 예측 B-프레임들(Bi-directional)을 이용하며, 이 프레임들은 화상 그룹 GOP 안의 두 앵커(anchor) 프레임들 (I- 및 P-프레임들, 또는 두개의 P-프레임들) 사이에 설정되고, B-프레임의 이미지 정보는 이전의 앵커 프레임과 그 B-프레임에 이어지는 것 모두로부터 예측된다. 따라서 B-프레임들은 P-프레임들 보다 높은 화질의 이미지 정보를 제공하나, 일반적으로 앵커 프레임들로 사용되지는 않기 때문에, 비디오 시퀀스로부터 B-프레임들을 제거하는 것이 일련의 이미지들의 화질을 저하시키지 않는다. 그러나, 어떤 것도 B-프레임들이 앵커 프레임들로도 역시 사용되는 것을 막지는 않으며, 다만 그런 경우에는 B-프레임들에 종속된 프레임들의 화질을 저하시키지 않으면서 B-프레임들이 비디오 시퀀스로부터 제거될 수는 없다.
- [0016] 각 비디오 프레임은 직사각형 이미지 영역으로 된 모든 픽셀들의 컬러 성분들(Y, U, V 와 같은)을 포함하는 매크로블록들로서 알려진 것으로 분할될 수 있다. 보다 구체적으로 말해, 하나의 매크로블록은 컬러 성분 당 적어도 하나의 블록으로 이뤄지며, 블록들 각각은 관련 이미지 영역 내 하나의 컬러 레벨의 컬러 값들(Y, U, 또는 V 등의)을 포함한다. 블록들의 공간 분해능은 매크로블록들의 분해능과는 다를 수 있는데, 예를 들면, U- 및 V

성분들은 Y-성분 분해능의 오직 절반만을 이용해 디스플레이될 수 있다. 매크로블록들은 가령, 일반적으로 이미지의 스캐닝(주사) 순서로 선택되는 매크로블록들의 그룹들인 슬라이스들로 더 그룹핑될 수 있다. 시간적 예측은 비디오 코딩 방법들에서 보통 이미지 프레임에 특정하게 되는 대신 블록이나 매크로블록에 특정하게 수행된다.

[0017] 비디오 파일들의 융통성있는 스트리밍을 허용하도록, 많은 비디오 코딩 시스템들이 스케일 가능(scalable) 코딩을 이용하며, 이 코딩에서 한 시퀀스의 일부 요소들이나 요소 그룹들이 그 비디오 시퀀스의 다른 부분들에 대한 재구성에 영향을 미치지 않으면서 제거될 수 있다. 스케일 가능성은 보통 이미지 프레임들을 다수의 계층적 레이어들로 그룹핑함으로써 구현된다. 기본 계층 스케일 가능성의 이미지 프레임들로 코딩된 이미지 프레임들은 수신단에서의 비디오 정보 디코딩을 위해 강제되는 것들만을 구비한다. 따라서 각 화상 그룹 GOP의 기본 계층은 하나의 I-프레임과 필요한 개수의 P-프레임들을 포함한다. 하나 이상의 인헨스먼트(enhancement) 계층들이 기본 계층 아래에서 결정될 수 있으며, 이 인헨스먼트 계층들 각각은 상위의 계층과 비교해 비디오 코딩의 품질을 향상시킨다. 따라서 인헨스먼트 계층은 하나 이상의 상위 계층 이미지들로부터의 모션-보상에 기반해 예측된 P- 또는 B-프레임들을 포함한다. 그 프레임들은 보통 수리적 시리즈에 따라 번호가 매겨진다.

[0018] 스트리밍시, 전송 비트율은 사용될 대역폭이나 수신자의 최대 디코딩 또는 비트율 값에 기반하여 제어될 수 있어야 한다. 비트율은 스트리밍 서버나, 인터넷 라우터 또는 이동 통신 네트워크의 기지국과 같은 통신 네트워크의 어떤 구성요소에서 제어될 수 있다. 비트율을 제어하기 위한 스트리밍 서버의 가장 간단한 수단이 전송시 고도의 정보 콘텐츠를 가진 B-프레임들을 누락시키는 것이다. 또, 스트리밍 서버가 비디오 스트림에서 전송될 스케일 가능 계층들의 개수를 결정할 수도 있으며, 그에 따라 스케일 가능 계층들의 개수는 새로운 화상 그룹 GOP가 시작될 때 항상 변화될 것이다. 서로 다른 비디오 시퀀스 코딩 방법들을 사용하는 것 역시 가능하다. 그에 따라, B-프레임들은 인헨스먼트 계층들의 다른 P-프레임들과 마찬가지로 통신 네트워크 구성 요소의 비트 스트림으로부터 제거될 수 있다.

[0019] 상술한 구성은 여러 결함들을 수반한다. ITU-T(국제 전기통신 연합, 전기통신 표준화 섹터) 표준 H.263과 같은 많은 코딩 방법들은 참조 화상 선택이라 불리는 절차에 익숙하다. 참고 화상 선택시 적어도 P-이미지의 일부가 시간 도메인의 바로 전 P-프레임이 아닌 적어도 다른 한 이미지로부터 예측되었다. 선택된 참조 이미지는 코딩된 비트 스트림이나 비트 스트림 헤더 필드로, 이미지에 특정하거나, 이미지-세그먼트에 특정하거나 (슬라이스나 매크로블록들의 그룹 등), 매크로블록에 특정하거나, 또는 블록에 특정하게 시그널링된다. 참조 화상 선택은, 코딩될 이미지에 시간적으로 잇따르는 이미지들로부터 예측 역시 이뤄질 수 있도록 일반화될 수 있다. 또, 참조 화상 선택은 B-프레임들을 포함해 시간적으로 예측된 프레임 종류들을 모두 커버하도록 일반화될 수 있다. 참조 이미지로서 화상 그룹 GOP를 시작하는 I-이미지에 선행하는 적어도 하나의 이미지를 선택하는 것 역시 가능하므로, 참조 화상 선택을 이용하는 화상 그룹이 반드시 독자적으로 디코딩될 수 있는 것은 아니다. 또, 스트리밍 서버나 네트워크 구성요소에서 스케일 가능성이나 코딩 방식의 조정이 어려워지는데, 그것은 상이한 이미지 그룹들 사이의 어떤 종속성이라도 검출될 수 있도록 하기 위해, 비디오 시퀀스는 긴 시간 주기 동안 디코딩, 분석 및 버퍼링되어야 하기 때문이다.

[0020] 또 다른 문제가 다른 비디오 시퀀스 중간에 한 비디오 시퀀스를 삽입하는 것과 관련되며, 이것은 보통 이미지 번호 매기기에서의 불연속성을 야기하였다. 비디오 시퀀스 이미지들의 번호 매김은 일반적으로 이미지 프레임들의 손실을 검출하기 위해 이용된다. 그러나, 만일 상업 광고와 같은 별개의 비디오 시퀀스가 한 비디오 시퀀스 안에 삽입될 때, 그 별개의 비디오 시퀀스는 보통 별개의 이미지 번호매김 것으로 주어지는데, 이것은 오리지널 비디오 시퀀스의 오름차순(ascending) 이미지 번호 매김과 일치하지 않는다. 따라서 수신 단말은 그 불일치한 이미지 번호매김을 잃어버린 이미지 프레임들의 신호라고 해석하고 잃어버렸다고 추정된 이미지 프레임들을 재구성하거나 그것을 재전송하라고 요구하는데 불필요한 동작을 시작한다.

[0021] 인코더와 디코더에서 기준 화상들을 버퍼링 및 인덱싱할 때 이와 비슷한 문제에 맞닥뜨릴 수 있다. 인덱싱의 일부로서, 인코더는 어떤 남아있는 이미지 프레임들의 시간적 예측 프로세스를 위한 기준 이미지 프레임들로서 쓸데없는 복수의 이미지 프레임들을 고의적으로 제거할 수 있다. 디코더에 그러한 제거 이유를 알리기 위한 프로세스 역시 존재하지 않는다: 가령, 복수 개의 이미지 프레임들이 제거되면, 수신 단말은 이러한 고의적 제거 내용을 프로토콜 에러라고 불필요한 해석을 할지도 모른다.

발명이 이루고자 하는 기술적 과제

[0022] 본 발명의 목적은, 디코더로 하여금 인코더가 고의적으로 제거한 이미지 프레임들을 고려할 수 있게 하는 개선

된 방법 및 그 방법을 구현하는 장치를 제공하는 데 있다.

발명의 구성 및 작용

- [0023] 이제, 본 발명의 다양한 양태들에는 독립 청구 범위들에서 기술한 것을 특징으로 하는 방법, 비디오 인코더, 비디오 디코더, 비디오 신호, 비디오 시퀀스의 비트 레이트를 조정하기 위한 장치이자 바람직하게는 스트리밍 시스템 구성 요소 및 컴퓨터 프로그램들이 포함된다.
- [0024] 본 발명의 바람직한 실시예들은 종속 청구범위들에 개시되어 있다.
- [0025] 본 발명은 압축된 비디오 시퀀스를 디코딩한다는 개념에 기반하는 것으로, 이때 이미지 프레임들이 디코딩과 관련된 버퍼 메모리로 입력된다. 비디오 시퀀스는 이미지 프레임들의 넘버링(번호 매김)시 적어도 한 번의 불연속과 관련된 표시를 포함하며, 그 표시는 비디오 시퀀스로부터 복호화된다. 그러면, 그 표시에 응하여, 버퍼 메모리가 이미지 프레임들의 넘버링시의 불연속에 대응하는 이미지 프레임들을 제공하도록 설정되고, 그 버퍼 메모리 내 이미지 프레임들이 디코딩 프로세스시 사용된다. 상기 표시는, 비디오 시퀀스 내 임지 프레임들의 넘버링시 적어도 한 번의 불연속이 고의적인 것이라는 정보로서 제공되고, 버퍼 메모리에서 발생된 이미지 프레임들이 디코더에 존재하지 않는 이미지 프레임들 대신에 사용됨이 바람직하다. 누락된 이미지 프레임들이 이미지 프레임 넘버링시 불연속을 일으킨다.
- [0026] 일실시예에 따르면, 버퍼 메모리가 이미지 프레임 넘버링시의 불연속에 해당하는 이미지 프레임들을 제공하도록 설정된 후, 해당 비디오 시퀀스에 구비된 이미지 프레임들의 버퍼 메모리로의 입력은 올바른 이미지 프레임 번호부터 재개된다.
- [0027] 일실시예에 따르면, 이미지 프레임 넘버링시의 불연속에 해당하는 보충(filler) 프레임들이 버퍼 메모리로 입력된다. 상기 보충 프레임들은 이들이 실제 비디오 시퀀스에 속하는 것이 아님을 나타내는 식별자에 의해 표시된다.
- [0028] 일실시예에 따르면, 버퍼 메모리가 이미지 프레임 넘버링시의 불연속에 해당하는 이미지 프레임들을 제공하도록 설정하는 단계는, 메모리 표시에 의해 수행되고, 그에 따라 아무 데이터도 버퍼 메모리로 입력되지 못한다.
- [0029] 본 발명에 의한 절차의 이점은 이미지 프레임 넘버링시의 간극(gap)이 고의적이라는 정보를 디코더에 제공함으로써, 디코더가 불필요한 여러 정정 동작을 시작하지 않게 된다는 데 있다. 또 다른 이점은 인코더 및 디코더의 버퍼 메모리들이 동조(synchronism)가 유지될 수 있어, 기준 화상 선택 처리에 대한 거침없는 동작이 가능하게 된다는 데 있다.
- [0030] 이하에서, 범용 멀티미디어 데이터 스트리밍 시스템이 개시되며, 그 시스템의 기본 원리들은 임의의 전기통신 시스템과 관련되어 적용될 수 있다. 본 발명은 여기서 스트리밍 시스템이라고 특정하게 참조하여 설명되며 그 시스템에서 멀티미디어 데이터가 가장 바람직하게는 IP 메트릭과 같은 패킷-교환 데이터 프로토콜을 이용하는 전기통신 네트워크를 통해 전송되지만, 본 발명은 고정 전화 네트워크 PSTN/SDN(공공 교환 전화 네트워크/통합 서비스 디지털 네트워크)과 같은 회로 교환 네트워크들이나 이동 통신 네트워크 PLMN(공공 지상 이동 네트워크)에서도 똑같이 잘 구현될 수 있다. 또한 본 발명은 일반 스트리밍 및 연속 다운로드 둘 모두의 형태로 멀티미디어 파일들의 스트리밍에서 적용될 수 있고, 가령 비디오 통화를 구현하기 위해 적용될 수 있다.
- [0031] 여기서 본 발명이 스트리밍 시스템을 특정하게 참조하여 설명되고 있고 또한 그 시스템 안에 적용됨이 바람직하지만, 본 발명은 스트리밍 시스템에만 한정되지 않으며, 디코딩 되어야 할 비디오 파일이 어떻게 다운로드되고 어디에서 다운로드 되는지와 관계없이 임의의 비디오 재생 시스템 안에서 적용될 수 있다. 따라서 본 발명은 가령 비디오 재생에 사용 가능한 가변 프로세싱 용량과 관련하여, DVD 디스크나 어떤 다른 컴퓨터 메모리 캐리어로부터 다운로드될 비디오 파일의 재생에 적용될 수 있다. 특히, 본 발명은 대역폭 제한을 필요로 하는 전기통신 시스템에 통상적으로 사용되는 낮은 비트율의 여러가지 비디오 코딩에 적용될 수 있다. 한 예가 ITU-T 표준 H.263과 (아마도 나중에 H.264가 된) H.26L에 정의된 시스템이다. 이들과 관련해, 본 발명은 가령 이동국들에 적용될 수 있고, 이 경우, 이동국이 비디오 재생이 아닌 다른 어플리케이션을 실행하기 위해 사용될 때에도, 비디오 재생은 변화하는 전송 용량이나 채널 품질과 현재 사용중인 프로세서 파워 모두에 대해 맞춰지도록 행하여 질 수 있다.
- [0032] 명료성을 위해, 본 발명은 아래에서 이미지 프레임 코딩과, 이미지 프레임 레벨의 시간적 예측에 관한 서술을 제공함으로써 설명될 것임을 더 유의해야 한다. 그러나, 실제로 코딩 및 시간적 예측은 통상적으로 상술한 바와 같이 블록 또는 매크로블록 레벨에 대해 일어난다.

- [0033] 도 1을 참조할 때, 일반적 멀티미디어 스트리밍 시스템이 도시되며, 이것은 본 발명의 절차를 적용하기 위한 바람직한 시스템이다.
- [0034] 멀티미디어 데이터 스트리밍 시스템은 보통, 비디오 카메라 및 마이크로폰, 또는 메모리 캐리어에 저장된 비디오 이미지나 컴퓨터 그래픽 파일들과 같은 하나 이상의 멀티미디어 소스들(100)을 포함한다. 서로 다른 멀티미디어 소스들(100)로부터 얻은 비가공 데이터는 편집기라고도 불릴 수 있는 인코더(102)의 멀티미디어 파일 안에 결합된다. 하나 이상의 멀티미디어 소스들(100)로부터 도달한 비가공 데이터는 먼저 인코더(102) 안에 포함된 포획수단(104)을 이용해 포획되며, 그 포획수단(104)은 일반적으로 서로 다른 인터페이스 카드, 드라이버 소프트웨어, 또는 카드의 기능을 제어하는 어플리케이션 소프트웨어로서 구현될 수 있다. 예를 들어, 비디오 데이터는 비디오 포획 카드와 그 관련 소프트웨어를 이용해 획득될 수 있다. 비디오 포획 카드가 관련될 때, 포획수단(104)의 출력은, 보통 예를 들어 YUV 4:2:0 포맷이나 모션-JPEG 이미지 포맷의 비압축 비디오 프레임들과 같은, 압축되지 않거나 약간만 압축된 데이터 플로우일 수 있다.
- [0035] 편집기(106)는 서로 다른 미디어 플로우들을 함께 링크하여 비디오 및 오디오 플로우들이 원하는 대로 동시에 재생될 수 있도록 한다. 편집기(106)는 또한 가령 프레임 레이트를 절반으로 하거나, 공간적 해상도(resolution)를 줄임으로써 비디오 플로우와 같은 각각의 미디어 플로우를 편집한다. 동기되었다라도 독자적인 미디어 플로우들은 압축기(108)에서 압축되며, 여기서 각 미디어 플로우는 그 미디어 플로우에 적합한 압축기를 이용해 독자적으로 압축된다. 예를 들어, YUV 4:2:0 포맷의 비디오 프레임들은 ITU-T 권고 H.263 또는 H.26L에 의한 저비트율 비디오 코딩을 이용해 압축될 것이다. 개별적으로 동기된 압축 미디어 플로우들은 보통 멀티플렉서(110)에서 인터리빙되며, 인코더(102)로부터 얻어지는 출력은 복수개의 미디어 플로우들로 된 데이터를 구비하여 멀티미디어 파일이라 부를 수 있는 단일의 규칙적 비트 플로우가 된다. 멀티미디어 파일의 생성이 반드시 복수개의 미디어 플로우들을 하나의 파일로 멀티플렉싱(다중화)할 것을 요구하지는 않는다는 것에 유의해야 한다. 그러나, 스트리밍 서버가, 미디어 플로우들을 전송하기 전에 그 미디어 플로우들을 인터리빙할 것이다.
- [0036] 멀티미디어 파일들은 스트리밍 서버(112)로 전송되고, 그러면 스트리밍 서버는 실시간 스트리밍이나 연속 다운로드 형식으로 스트리밍을 수행할 수 있다. 연속 다운로드시, 멀티미디어 파일들은 먼저 서버(112)의 메모리에 저장되며, 수요가 발생하면 전송을 위해 그 메모리에서 상기 멀티미디어 파일들이 검색된다. 실시간 스트리밍시 편집기(102)는 스트리밍 서버(112)로 멀티미디어 파일들의 연속되는 미디어 플로우를 전송하고, 서버(112)는 그 플로우를 클라이언트(114)에게 직접 보낸다. 추가 옵션으로서, 실시간 스트리밍은 또한 멀티미디어 파일들이 서버(112)로부터 액세스 가능한 저장부 안에 저장되도록 수행될 수 있고, 그 저장부로부터 실시간 스트리밍이 행해질 수 있고 수요가 일어나면 멀티미디어 파일들의 연속적인 미디어 플로우가 개시된다. 상기의 경우, 편집기(102)가 반드시 스트리밍을 꼭 제어하지는 않는다. 스트리밍 서버(112)가 사용 가능한 대역폭이나 클라이언트(114)의 최대 디코딩 및 재생율에 관해 멀티미디어 데이터의 트래픽 정형화를 수행하며, 스트리밍 서버는 가령 전송시 B-프레임들을 누락시키거나 스케일 가능 계층들의 개수를 조정함으로써 미디어 플로우의 비트율을 조정할 수 있다. 또, 스트리밍 서버(112)는 다중화된 미디어 플로우의 헤더 필드들을 변형해 크기를 줄이고 멀티미디어 데이터를 사용중인 전기통신 네트워크의 전송에 적합한 데이터 패킷들로 감싼다. 클라이언트(114)는 일반적으로, 적절한 제어 프로토콜을 이용해 적어도 어느 정도까지는 서버(112)의 동작을 조정할 수 있다. 클라이언트(114)는 적어도, 원하는 멀티미디어 파일이 클라이언트로 전송되게 선택될 수 있도록 하는 방식으로 서버(112)를 제어할 수 있고, 그에 더하여 클라이언트는 통상적으로 멀티미디어 파일의 전송을 중단 및 인터럽트할 수 있다.
- [0037] 클라이언트(114)가 멀티미디어 파일을 수신하고 있을 때, 그 파일은 먼저 디멀티플렉서(116)로 보내지고, 여기서 그 멀티미디어 파일을 구성하는 미디어 플로우들이 분리된다. 그러면 그 분리된, 압축 미디어 플로우들은 압축해제기(118)로 보내져서 각각의 분리된 미디어 플로우가 그 각각에 적합한 압축해제기에 의해 압축해제된다. 압축해제되고 재구성된 미디어 플로우들은 재생부(120)로 보내져서 그들의 동기 데이터에 의해 올바른 페이스(pace)로 렌더링되고 표시수단(124)으로 보내진다. 실제 표시수단(124)은 가령 컴퓨터나 이동국 디스플레이, 및 스피커 수단을 구비할 수 있다. 클라이언트(114)는 또한 통상적으로 최종 사용자가 보통은 사용자 인터페이스를 통해 제어하는, 상술한 제어 프로토콜을 통한 서버의 동작과 최종 사용자가 내린 명령에 기초한 재생부(120)의 동작 모두를 제어하는 제어부(122)를 포함한다.
- [0038] 스트리밍 서버(112)로부터 클라이언트(114)로의 멀티미디어 파일들의 전송은 전기통신 네트워크를 통해 발생함을 유의해야 하며, 그 전송 경로는 일반적으로 복수개의 전기통신 네트워크 구성요소들을 포함한다. 따라서 사용가능한 대역폭이나 클라이언트(114)의 최대 디코딩 및 재생율에 관한 멀티미디어 데이터의 트래픽 정형화를, 적어도 일부가 스트리밍 서버와 관련해 위에서 설명한 것과 동일한 방식으로 수행할 수 있는 적어도 몇몇의 네트워크

구성요소들이 존재할 수 있다.

- [0039] 스케일 가능 코딩이 도 2에 도시된 예 및 본 발명의 바람직한 실시예를 참조해 이하에서 설명될 것이다. 도 2는 인트라(INTRA) 프레임, 또는 I-프레임인 제1프레임(200)으로, 그 이미지 정보가 모션-보상 시간적 예측을 이용하지 않고 결정되는, 독립적으로 결정된 비디오 프레임들 중 하나를 구비하는 압축된 비디오 시퀀스의 일부를 도시한다. I-프레임(200)은 인트라(INTRA) 계층이라 부를 수 있는 제1스케일 가능 계층 위에 놓여진다. 각 스케일 가능 계층에는 계층 번호와 같은 고유한 식별자가 할당된다. 그에 따라 인트라 계층은 가령 번호 0이나, 가령 어떤 하나의 문자와 같은 다른 문자/숫자 식별자, 또는 문자와 숫자의 조합 형태의 식별자가 주어질 수 있다.
- [0040] 이와 유사하게, 하나 이상의 비디오 프레임들의 그룹들로 이뤄진 서브-시퀀스들이 각각의 스케일 가능 계층마다 정해지고, 한 그룹안의 이미지들 중 적어도 하나(보통은 최초이거나 마지막의 것)는 적어도 보통 다른 상위의, 또는 동일 스케일 가능 계층의 비디오 프레임으로부터 시간적으로 예측되고, 나머지 비디오 프레임들은 통상적으로, 오직 동일한 서브-시퀀스의 비디오 프레임들로부터, 혹은 아마도 상기 제2서브-시퀀스의 하나 이상의 비디오 프레임들로부터도 시간적으로 예측되어진다. 한 서브-시퀀스는 상기 제2서브-시퀀스를 제외한 다른 서브-시퀀스들과 무관하게 독자적으로 디코딩될 수 있다. 각각의 스케일 가능 계층의 서브-시퀀스들은 그 스케일 가능 계층의 제1서브-시퀀스가 번호 0으로 주어져 시작되는 것과 같이 연속적 번호 매기기를 이용한 고유 식별자가 할당된다. I-프레임(200)은 독자적으로 정해지고 또한 다른 이미지 프레임들과 무관하게 수신시 독자적으로 디코딩될 수 있으므로 역시 개별적 서브-시퀀스 방식으로 형성된다.
- [0041] 본 발명의 실질적 양상은 서브-시퀀스가 종속되는 서브-시퀀스들의 입장에서 각 서브-시퀀스를 결정하는 것이다. 즉, 한 서브-시퀀스는 그 서브-시퀀스의 이미지 프레임들을 예측하기 위해 바로 사용되었던 모든 서브-시퀀스들에 대한 정보를 포함한다. 이 정보는 비디오 시퀀스 비트 스트림으로 시그널링되며 실질적 이미지 정보와는 별개의 것이 바람직하므로, 나머지 이미지 데이터의 디코딩에 영향을 주지 않으면서 독자적으로 디코딩되고 제거될 수 있는 비디오 시퀀스 부분을 결정하는 것이 용이하다는 이유 때문에 비디오 시퀀스의 이미지 데이터가 바람직하게 조정될 수 있다.
- [0042] 다음으로, 각 서브-시퀀스 내에서, 그 서브-시퀀스의 비디오 프레임들에는 가령, 그 서브-시퀀스의 첫번째 비디오 프레임에 번호 0이 주어지는 것으로 시작하는 연속적 번호 매기기를 이용한 이미지 번호들이 주어진다. I-프레임(200) 역시 독립적인 서브-시퀀스를 형성하므로, 그 이미지 번호는 0이 된다. 도 2에서, I-프레임(200)은 유형 (1)에, 프레임의 서브 시퀀스 식별자 및 이미지 번호 (0.0)인 것을 보인다.
- [0043] 도 2는 또한 인트라 계층의 다음 I-프레임(202)을 도시하고 있으며, 그 프레임 역시 모션-보상된 시간적 예측을 이용하지 않고 결정되었던, 독자적으로 정해진 비디오 프레임이다. I-프레임들의 시간적 전송 주파수는 비디오 코딩, 이미지 정보 콘텐츠 및 사용될 대역폭과 관련된 많은 요인들에 좌우되며, 어플리케이션이나 어플리케이션 환경에 따라 I-프레임들이 가령 0.5에서 10초의 간격의 비디오 시퀀스로 전송된다. I-프레임(202)은 독립적으로 디코딩될 수 있기 때문에 역시 독자적인 서브-시퀀스를 형성한다. 이것이 인트라 계층의 두번째 서브-시퀀스이기 때문에, I-프레임(202)의 서브-시퀀스 식별자에 대한 연속 번호는 1이다. 또, I-프레임(202) 역시 독자적 서브-시퀀스를 형성하기 때문에, 즉, 그것이 그 서브-시퀀스 내 유일한 비디오 프레임이므로, 그 이미지 번호는 0이다. 따라서 I-프레임(202)은 식별자 (I.1.0)으로 지정될 수 있다. 마찬가지로, 인트라 계층의 다음 I-프레임의 식별자는 (I.2.0)이 되고, 계속 이런 방식으로 식별자가 정해진다. 결국, 이미지 정보가 모션-보상된 시간적 예측을 이용해 정해진 것이 아닌, 독자적으로 정해진 I-프레임들만이 제1스케일 가능 계층, 즉 인트라 계층 안에 코딩된다. 서브-시퀀스들이 서로 서로 구별될 수 있으면, 이들은 다른 종류의 번호 매김 방식이나 다른 식별자들을 이용해 정해질 수도 있다.
- [0044] 가령 계층 번호 1을 갖는 다음 스케일 가능 계층을 베이스(Base) 계층이라 칭할 때, 그 계층은 코딩되고, 일반적으로 이전 이미지 프레임들로부터만(즉, 이 경우에는 상위의 인트라 계층의 I-프레임들로부터) 예측되는 모션-보상 인터(INTER) 또는 P-프레임들을 구비한다. 도 2에 도시된 베이스 계층의 제1P-프레임(204)의 이미지 정보는 인트라 계층의 I-프레임(200)을 이용해 정해진다. P-프레임(204)은 베이스 계층의 첫번째 서브-시퀀스를 시작하며, 그에 따라 P-프레임(204)의 서브-시퀀스 식별자는 0이 된다. 또, P-프레임(204)이 베이스 계층의 제1서브-시퀀스의 제1이미지 프레임이므로, 그 P-프레임(204)의 이미지 번호는 0이다. 따라서 P-프레임(204)은 (P.0.0)로 식별될 수 있다.
- [0045] 시간적으로 이어지는 베이스 계층의 P-프레임(206)은 이전 P-프레임(204)으로부터 예측된다. 따라서 P-프레임들(204 및 206)은 동일한 서브-시퀀스에 속하고, 그에 따라 P-프레임(206) 역시 서브-시퀀스 식별자 0를 받는다. P-프레임(206)이 서브-시퀀스 0의 두번째 이미지 프레임이므로, P-프레임(206)의 이미지 번호는 1이다.

로, P-프레임(206)은 (P.0.1)로 식별될 수 있다.

- [0046] 베이스 계층에 이어 계층 번호 2를 갖는 스케일 가능 계층을 인핸스먼트(enhancement) 계층 1이라 부른다. 이 계층은 코딩되고, 여기서는 인트라 계층의 I-프레임들이나 베이스 계층의 P-프레임들이나 이전 이미지 프레임들로부터만 예측되는 모션-보상 P-프레임들을 구비한다. 도 2는 인핸스먼트 계층 1의 제1이미지 프레임(208)과 제2 이미지 프레임(210)을 도시하며, 이들은 모두 인트라 계층의 제1이미지 프레임(200)으로부터만 예측된다. P-프레임(208)은 인핸스먼트 계층 1의 첫번째 서브-시퀀스를 시작하며, 따라서 P-프레임의 그 서브-시퀀스 식별자는 0이 된다. 또, P-프레임(208)이 상기 서브-시퀀스의 최초의 이미지 프레임이므로 P-프레임(208)은 이미지 번호 0을 받는다. 따라서 P-프레임(208)은 (P.0.0)으로 식별될 수 있다.
- [0047] 제2이미지 프레임(210) 역시 인트라 계층의 제1이미지 프레임(200)으로부터만 예측되므로, P-프레임(210)은 인핸스먼트 계층 1의 제2서브-시퀀스를 시작하고 그 P-프레임(210)의 서브-시퀀스 식별자는 그래서 1이 된다. P-프레임(210)이 그 서브-시퀀스의 첫번째 이미지 프레임이므로, P-프레임(210)의 이미지 번호는 0이다. P-프레임은 따라서 (P.1.0)으로 식별될 수 있다. 시간적으로 이어지는 인핸스먼트 계층 1의 P-프레임(212)은 이전의 P-프레임(210)으로부터 예측된다. P-프레임들(210 및 212)은 동일한 서브-시퀀스에 속하므로, 그 P-프레임(212) 역시 서브-시퀀스 식별자 1을 받는다. P-프레임(212)은 서브-시퀀스 1의 두번째 이미지 프레임이므로, P-프레임은 이미지 번호 1을 받아 (P.1.1)로 식별될 수 있다.
- [0048] 인핸스먼트 계층 1의 시간적으로 네 번째 이미지 프레임(214)은 베이스 계층의 제1이미지 프레임(204)로부터 예측된다. 따라서 P-프레임(214)은 제3서브-시퀀스를 시작하므로 P-프레임(214)은 서브-시퀀스 식별자 2를 받는다. 또, P-프레임(214)이 그 서브-시퀀스 안에서 최초의 유일한 이미지 프레임이므로 P-프레임(214)의 이미지 번호는 0이다. 따라서, P-프레임(208)은 (P.2.0)으로 식별될 수 있다.
- [0049] 인핸스먼트 계층 1의 시간적으로 다섯번째인 이미지 프레임(216) 역시 베이스 계층의 제1이미지 프레임(204)으로부터만 예측되므로, 그 P-프레임(216)은 인핸스먼트 계층 1의 네 번째 서브-시퀀스를 시작하여, P-프레임(216)의 서브-시퀀스 식별자가 3이 된다. 또, P-프레임(216)이 해당 서브-시퀀스에서 첫번째이므로, P-프레임(216)의 이미지 번호는 0이 된다. 따라서 P-프레임(216)은 (P.3.0)으로 식별될 수 있다. 시간적으로 이어지는 인핸스먼트 계층 1의 P-프레임(218)은 이전 P-프레임(216)으로부터 예측된다. P-프레임들(216 및 218)이 동일한 서브-시퀀스에 속하므로, P-프레임(218)의 서브-시퀀스 식별자 역시 3이 된다. P-프레임(218)이 서브-시퀀스 3의 두번째 이미지 프레임이므로, P-프레임(218)의 이미지 번호는 1이고 따라서 P-프레임(218)의 식별자는 (P.3.1)이 된다.
- [0050] 설명의 단순성 및 명료성을 위해, 위에서 개시된 것은 단지 I- 및 P-프레임들만을 연관짓는다. 그러나, 이 분야의 당업자라면 본 발명의 스케일 가능한 비디오 코딩이 상술한 B-프레임들 및 적어도 SI-프레임들, SP-프레임들 및 MH-프레임들과 같은 다른 기존의 이미지 프레임 유형들을 이용해서도 구현될 수 있음을 알 수 있을 것이다. SI-프레임들은 I-프레임들에 해당하지만, SP-프레임과 함께 동일 이미지 재구성을 가능하게 한다. SP-프레임은 SI-프레임이나 다른 SP-프레임과 함께 동일 이미지가 재구성될 수 있게 하는 특정한 코딩을 경험하는 P-프레임이다. SP-프레임들은 일반적으로, 한 액세스 포인트 또는 스캐닝 포인트가 요망되거나 비디오 스트림의 코딩 패러미터들의 변화가 가능해야 하는 포인트들로의 비디오 시퀀스 안에 놓여진다. 프레임들은 에러 정정 및 에러 내성 증대를 위해 사용될 수도 있다. SP-프레임들은, 이들이 SP- 또는 SI-유형의 다른 비디오 프레임으로 대체될 수 있도록 규정된다는 점을 제외하면 이전 프레임들로부터 예측되는 일반적인 P-프레임들과 동일하며, 새 프레임에 대한 디코딩 결과는 비디오 스트림 안에 있었던 오리지널 SP-프레임의 디코딩 결과와 동일하게 된다. 즉, 비디오 스트림안에 있었던 것을 대체하는데 사용된 새 SP-프레임은 다른 시퀀스나 비디오 스트림으로부터 예측되며 또한 그 재구성 프레임은 동일한 콘텐츠를 포함한다. SP-프레임들은 가령 본 출원인의 이전 출원 PCT/FI02/00004에서 설명되고 있다.
- [0051] B-프레임들과 유사하게, 모션-보상 예측에 기반한 MH(Multi Hypothesis) 프레임들의 매크로블록들은 두 개의 다른 프레임들로부터 예측되나, 반드시 한 MH-프레임 다음에 위치되는 것은 아니다. 보다 자세히 말하면, 예측된 매크로블록들은 두개의 다른 프레임들의 두 매크로블록들의 평균치로서 산출된다. 두 프레임 대신, MH-프레임 매크로블록들이 다른 한 프레임으로부터 자연스럽게 예측될 수도 있다. 참조 이미지들은 매크로블록에 따라 변화할 있다, 즉, 하나의 동일 이미지 안의 모든 매크로블록들이 꼭 동일한 프레임들을 이용해 예측되는 것은 아니다.
- [0052] 따라서 서브-시퀀스는 비디오 시퀀스의 어떤 특정 시간 주기를 커버한다. 동일 계층이나 다른 계층들의 서브-시퀀스들은 일부 혹은 전부가 중복될 수 있다. 만일 시간적으로 중복된 이미지 프레임들이 동일 계층 위에 존

재하면, 그 프레임들은 동일 이미지 콘텐츠의 대안적 표현들로 해석되어 어떤 모드의 이미지 표현이라도 사용될 수 있게 된다. 한편, 만일 서로 다른 계층들에서 시간적으로 중복된 이미지 프레임들이 존재하면, 그 이미지 프레임들은 동일 이미지 콘텐츠에 대한 상이한 표현들을 생성하며, 그에 따라 표현들마다 화질에서 차이가 나게 된다. 말하자면, 이미지 품질은 하위 계층에서 보다 우수하다.

[0053] 도 2를 참조한 상기 내용은 본 발명의 실시예에 따른 스케일 가능 코딩 구성, 계층 구조 및 이미지 프레임들의 번호매기기에 대해 예시한 것이다. 이 실시예에서, 인트라 계층만이 I-프레임들을 구비하며, 베이스 계층은 인트라 계층으로부터 수신된 정보를 이용해서만 디코딩될 수 있다. 이와 마찬가지로, 인헨스먼트 계층 1의 디코딩은 일반적으로 베이스 계층 및 인트라 계층 모두로부터의 정보를 필요로 한다.

[0054] 스케일 가능 계층의 수는 위에서와 같이 3 개로 한정되지 않으며, 임의의 개수의 인헨스먼트 계층들이라도 충분한 스케일 가능성을 제공하는데 필요하다고 생각되면 사용될 수 있다. 결과적으로, 인헨스먼트 계층 2의 계층 수는 4이고, 인헨스먼트 계층 3의 계층 수는 5가 되는 식이다. 상술한 예의 몇몇 이미지 프레임들에는 동일한 식별자가 주어지므로 (가령, 이미지 프레임들(204, 208)은 모두 그 식별자가 (P.0.0)), 식별자에 계층 번호를 포함함으로써 각 이미지 프레임은 고유하게 식별될 수 있고, 동시에 각 이미지 프레임의 다른 이미지 프레임들에 대한 종속성이 바람직하게 정해진다. 따라서 각 이미지 프레임은 고유하게 식별되며, 이미지 프레임(204)의 식별자는, 가령 (P.1.0.0), 또는 단순히 (1.0.0)이 되고, 마찬가지로 이미지(208)의 식별자는 (P.2.0.0), 또는 (2.0.0)이 된다.

[0055] 본 발명의 바람직한 실시예에 따르면, 참조 이미지 프레임의 개수는, 예를 들어 0에서 255 사이의 정수와 같이 특정한 소정 문자-숫자 시리즈들에 따라 정해진다. 패러미터 값이 관련 시리즈의 최대값 N(가령, 255)에 도달할 때, 패러미터 값의 결정은 처음부터, 즉 시리즈의 최소값(가령, 0)에서부터 시작된다. 따라서 이미지 프레임은 동일한 이미지 번호가 다시 사용되는 지점까지 특정 서브-시퀀스 안에서 고유하게 식별된다. 서브-시퀀스 식별자는 또한 특정한, 소정 수리적 시리즈에 따라 정해질 수도 있다. 서브-시퀀스 식별자의 값이 시리즈의 최대값 N에 도달할 때, 식별자를 정하는 것은 그 시리즈의 처음부터 다시 시작된다. 그러나, 서브-시퀀스에는 (동일 계층 안에서) 사용 중인 식별자가 할당될 수 없다. 사용 중인 시리즈는 또한 수리적 방식 외에 다른 방법으로 정해질 수도 있다. 다른 한 대안으로서, 할당된 식별자가 다시 사용되지 않는다는 것을 감안하여 무작위의 서브-시퀀스 식별자들을 할당하는 것이다.

[0056] 사용자가 비디오 시퀀스의 중간에 비디오 파일 검색을 시작하고자 할 때 이미지 프레임들에 대한 번호 매기기의 문제가 발생한다. 가령, 사용자가 국부적으로 저장된 비디오 파일을 앞으로 검색하고자 하거나 특정 지점의 스트리밍 파일을 검색하고자 할 때; 사용자가 한 무작위 지점으로부터 스트리밍 파일의 재생을 시작할 때; 또는 재생될 비디오 파일이 그 재생을 방해하는 에러를 포함하고 있음이 검출되거나 그 에러에 뒤 이은 지점에서 재생을 다시 재개할 필요가 있을 때, 그러한 경우들이 일어난다. 비디오 파일이 이전의 검색 후 한 무작위 지점에서 재개될 때, 보통 이미지 번호 매기기의 불연속성이 발생한다. 일반적으로 디코더는 이것을 이미지 프레임들의 비의도적 손실로 해석하여 잃어버렸다고 추정한 이미지 프레임들을 재구성하려고 불필요하게 시도한다.

[0057] 본 발명의 바람직한 실시예에 따르면, 이러한 것은 디코더에서, 비디오 파일의 무작위 지점에서 활성화되는, 독자적으로 디코딩 가능한 화상 그룹 GOP에 한 시작 이미지를 규정함으로써 회피될 수 있으며, 그 시작 이미지의 번호는 0으로 정해진다. 따라서 그 독자적으로 디코딩 가능한 이미지 그룹은 인트라 계층의 서브-시퀀스일 수 있고, 가령, 이 경우 I-프레임이 그 시작 이미지로서 사용되며, 만일 베이스 계층에서 일어난 스케일링이 이용될 때 독자적으로 디코딩 가능한 이미지 그룹은 베이스 계층의 서브-시퀀스이고, 이 경우 통상적으로 I-프레임인, 그 서브-시퀀스의 제1이미지 프레임이 시작 이미지로서 보통 사용된다. 결과적으로, 무작위 지점에서 활성화될 때, 디코더는, 독자적으로 디코딩 가능한 서브-시퀀스의 제1이미지 프레임, 바람직하게는 I-프레임을 0이라고 정한다. 디코딩될 서브-시퀀스 역시 (가령 상술한 문자-숫자 시리즈가 처음부터 시작될 때) 식별자가 0인 다른 이미지 프레임들을 구비할 수 있으므로, 서브-시퀀스의 시작, 즉 제1이미지 프레임은 그 이미지 프레임의 일부분 중 헤더 필드에 부가되는 개별 플래그 등에 의해 디코더로 지시될 수 있다. 이것은 디코더로 하여금 이미지 번호들을 바르게 해석하고 비디오 시퀀스 이미지 프레임들로부터 서브-시퀀스를 개시하는 바른 이미지 프레임을 알 수 있도록 한다.

[0058] 상술한 번호 매김 시스템은 본 발명의 고유한 이미지 프레임 식별을 어떻게 수행하고 그에 따라 동시에 어떻게 이미지 프레임들 사이의 상호의존성을 나타내도록 하는지에 대한 오직 하나의 예를 제공한다. 그러나, ITU-T 표준 H.263 및 H.26L에 따른 비디오 코딩 방법들과 같은, 본 발명의 방법이 적용될 수 있는 비디오 코딩 방법들은 가변장 코드들을 사용하는 코드 테이블들을 사용한다. 가변장 코드들이 코딩 계층 번호들을 위해 사용될

때, 가령, 하위 코드워드 인덱스, 즉 작은 계층 번호는 더 짧은 코드워드를 나타낸다. 실제로 본 발명의 스케일 가능 코딩은 베이스 계층이 인트라 계층 보다 훨씬 더 많은 이미지 프레임들로 이뤄질 수 있도록 대부분의 경우들에서 사용될 것이다. 이것은 인트라 계층상에서 보다 베이스 계층 상에서, 하위 인덱스, 즉 보다 작은 계층 번호의 사용을 정당화하는데, 그 이유는 코딩된 비디오 데이터의 양이 그에 따라 바람직하게 감소되기 때문이다. 결국, 인트라 계층은 계층 번호 1이 할당되고 베이스 계층은 계층 번호 0이 주어진다. 이와 다른 선택사항으로서, 인트라 계층 번호 보다 베이스 계층 번호를 코딩하는데 더 적은 비트들을 사용해 코드가 만들어질 수 있고, 이 경우 실제 계층 번호치는 생성된 코드 길이 쪽과는 무관하다.

[0059] 또, 본 발명의 바람직한 제2실시예들에 따르면, 스케일 가능 계층들의 수는 낮게 유지되어야 할 때, 특별히 제1스케일 가능 계층은 인트라 계층과 베이스 계층 모두를 구비하도록 코딩될 수 있다. 코딩 계층성의 관점에서 볼 때, 이것을 계획하는 가장 간단한 방법이 인트라 계층을 통제로 배제하고, 이미지 정보가 모션-보상된 시간적 예측을 이용해 결정된 것이 아니며 독자적으로 규정된 I-프레임들과, 이미지 프레임들이 동일 계층의 I-프레임들로부터 예측된 모션-보상된 P-프레임들인, 이전 프레임들로부터 예상된 이미지 프레임들 모두로 이뤄진 코딩 프레임들이 베이스 계층에 제공되도록 하는 것이다. 따라서 계층 번호 0는 베이스 계층에 대해 사용될 수 있고, 인헨스먼트 계층들이 비디오 시퀀스 안에 코딩될 때, 인헨스먼트 계층 1은 계층 번호 1이 할당된다. 이것이 도 3a 및 3b를 참조하여 이하에서 예시될 것이다.

[0060] 도 3a는 스케일 불가능한 비디오 시퀀스 구조를 보이며, 이 구조에서 모든 이미지 프레임들이 동일한 스케일 가능 계층, 즉 베이스 계층 상에 위치된다. 비디오 시퀀스는 I-프레임 (I.0.0)으로 제1서브-시퀀스를 시작하는 제1이미지 프레임(300)을 구비한다. 이미지 프레임(300)은 그 서브-시퀀스의 제2이미지 프레임(302), 즉 P-프레임 (P.0.1)을 예측하는데 사용되고, 프레임(302)은 그 서브-시퀀스의 제3이미지 프레임(304), 즉 P-프레임 (P.0.2)을 예측되는데 사용되며, 프레임(304)은 다음 이미지 프레임(306), 즉 P-프레임 (P.0.3)을 예측하는데 사용된다. 그리고 나서 비디오 시퀀스에는 그 안에서 코딩된 I-프레임 (I.1.0), 즉 I-프레임(308)이 제공되며, 그것은 비디오 시퀀스의 제2서브-시퀀스를 시작하게 한다. 이런 종류의 스케일 불가능 코딩은 가령 사용중인 어플리케이션이 스케일 가능 코딩이 사용되는 것을 허용하지 않거나 그 사용의 필요가 없을 때 사용될 수 있다. 회로 교환 비디오 폰 어플리케이션에서, 예를 들어, 채널 대역폭은 일정하게 유지되고 비디오 시퀀스는 실시간으로 코딩되므로, 보통 스케일 가능 코딩을 할 필요가 없게 된다.

[0061] 도 3b는 필요할 때 인트라- 및 베이스 계층 결합에 스케일 가능성이 어떻게 추가될 수 있는지에 대한 예를 도시한다. 여기에서도 역시, 비디오 시퀀스 베이스 계층은 I-프레임 (I.0.0)이고 베이스 계층의 제1서브-시퀀스를 시작하는 제1이미지 프레임(310)을 포함한다. 이미지 프레임(310)은 서브-시퀀스의 제2이미지 프레임(312), 즉 P-프레임 (P.0.1)을 예측하는데 사용되고, 제2이미지 프레임(312)은 서브-시퀀스의 제3이미지 프레임(314), 즉, P-프레임 (P.0.2)를 예측하는데 사용된다. 그러나, 인헨스먼트 계층 1 역시 이 비디오 시퀀스 안에 코딩되고 제1서브-시퀀스를 포함한다. 그 제1서브-시퀀스의 오직 하나인 유일한 이미지 프레임(316)이 P-프레임 (P.0.0)이며, 이것은 베이스 계층의 제1이미지 프레임(310)으로부터 예측된다. 인헨스먼트 계층의 제2서브-시퀀스의 제1이미지 프레임(318)이 베이스 계층의 제2이미지 프레임(312)로부터 예측되고, 그에 따라 P-프레임은 (P.1.0)이 된다. 인헨스먼트 계층의 다음 이미지 프레임(320)은 같은 계층의 이전 이미지 프레임(318)으로부터 다시 예측되고, 동일한 서브-시퀀스에 속하므로, 식별자가 (P.1.1)이 된다.

[0062] 본 발명의 이러한 실시예에서, 한 베이스 계층 서브-시퀀스가 다른 베이스 계층 서브-시퀀스에 종속되더라도, 베이스 계층의 서브-시퀀스들은 독립적으로 디코딩될 수 있다. 베이스 계층 서브-시퀀스들의 디코딩은 베이스 계층 및/또는 인헨스먼트 계층 1의 제2서브-시퀀스로부터의 정보를 필요로 하며, 인헨스먼트 계층 2의 서브-시퀀스들은 인헨스먼트 계층 1 및/또는 인헨스먼트 계층 2의 서브-시퀀스로부터의 정보를 필요로 하는 식이 된다. 한 실시예에 따르면, I-프레임들은 베이스 계층에만 유일하게 한정되는 것이 아니며, 인헨스먼트 계층들 역시 I-프레임들을 구비할 수 있다.

[0063] 상기 실시예들 이면의 기본 사상이, 한 서브-시퀀스가 자신이 종속된 모든 서브-시퀀스들, 즉 문제의 서브-시퀀스의 이미지 프레임들 중 적어도 하나를 예측하는데 사용되었던 모든 서브-시퀀스들에 대한 정보를 구비한다는 것이다. 그러나, 일실시예에 따르면, 한 서브-시퀀스가 그 서브-시퀀스에 종속된 모든 서브-시퀀스들에 대한, 즉 그 서브-시퀀스의 적어도 한 이미지 프레임을 이용해 예측되었던 적어도 한 이미지 프레임이 있는 모든 서브-시퀀스들에 대한 정보를 포함하는 것 역시 가능하다. 후자의 경우 종속성이 보통 시간상 앞쪽으로 결정되므로, 다음에 설명될 방식에 따른 코딩시 이미지 프레임 버퍼들이 바람직하게 활용될 수 있다.

[0064] 상기 모든 실시예들에서 이미지 프레임의 번호 매기기는 서브-시퀀스에 특정한 것이다. 즉 새로운 서브-시퀀스

는 항상 처음부터 번호매김을 시작한다. 따라서 개별 이미지 프레임의 식별은 계층 번호, 서브-시퀀스 식별자 및 정해질 이미지 프레임 번호를 필요로 한다. 본 발명의 바람직한 실시예에 따르면, 이미지 프레임들은 코딩 순서 상의 연속 참조 이미지 프레임들이 하나씩 증가되는 번호들로서 표시되는 연속 번호매김 방식을 이용해 독립적으로 번호 매김될 수 있다. 계층 번호들과 서브-시퀀스 식별자들과 관련하여, 상술한 번호매김 방식이 사용될 수 있다. 이것은 계층 번호와 서브-시퀀스 식별자를 사용하지 않고 필요할 때마다 각 이미지 프레임이 고유하게 식별될 수 있도록 한다.

[0065] 이것은 베이스 계층이 시간상으로 첫번째 I-프레임(400) (I.0.0)을 포함하는 도 4a에 도시된 예로서 예시된다. 이 프레임은 인헨스먼트 계층 1의 제1이미지 프레임(402), 즉 (P.0.1)을 예측하는데 사용되며, 제1이미지 프레임(402)은 동일한 서브-시퀀스 (서브-시퀀스 식별자 0)에 속한 제2이미지 프레임(404), 즉 (P.0.2)를 예측하는데 사용되고, 제2이미지 프레임(404)은 동일한 서브-시퀀스의 제3이미지 프레임(406), 즉 (P.0.3)을 예측하는데 사용되고, 제3이미지 프레임(406)은 제4이미지 프레임(408) (P.0.4)을 예측하는데 사용되고, 마지막으로, 제4프레임은 제5이미지 프레임(410) (P.0.5)을 예측하는데 사용된다. 시간상 다음 비디오 시퀀스 이미지 프레임(412)은 베이스 계층 위에 위치하고, 여기서 그것은 I-프레임(400)과 동일한 서브-시퀀스 안에 있지만, 시간적으로는 단지 일곱번째로 코딩된 이미지 프레임일 뿐이므로, 그 식별자는 (P.0.6)이 된다. 그 일곱번째 프레임은 인헨스먼트 계층 1의 제2서브-시퀀스의 제1이미지 프레임(414), 즉 (P.1.7)을 예측하는데 사용되고, 제1이미지 프레임(414)은 동일한 서브-시퀀스(서브-시퀀스 식별자 1)에 속하는 제2이미지 프레임(416), 즉 (P.1.8)을 예측하는데 사용되고, 제2이미지 프레임(416)은 제3이미지 프레임(418) (P.1.9)을 예측하는데 사용되고, 제3이미지 프레임은 제4이미지 프레임(420) (P.1.10)을 예측하는데 사용되며, 마지막으로 제4이미지 프레임은 제5이미지 프레임(422) (P.1.11)을 예측하는데 사용된다. 다시, 시간적으로 그 다음 비디오 시퀀스 이미지 프레임(424)이 베이스 계층상에 위치하고, 그것은 I-프레임(400) 및 P-프레임(412)과 동일한 서브-시퀀스 안에 있지만, 시간적으로는 13번째 코딩된 이미지 프레임에 불과하므로 그 식별자는 (P.0.12)가 된다. 예시의 명료함을 위해, 상기 실시예에 대한 설명에서는 계층 식별자들을 포함하지 않았지만, 스케일 가능성을 구현하기 위해, 보통 계층 식별자가 이미지 프레임 식별자들의 일부로서 비디오 시퀀스와 합쳐져 시그널링되어야 한다는 것은 명백한 사실이다.

[0066] 도 4b 및 4c는 도 4a에 도시된 비디오 시퀀스의 이미지 프레임들을 그루핑하기 위한 다른 선택적 실시예들을 도시한다. 도 4b의 이미지 프레임들은 서브-시퀀스에 따라 번호가 매겨진다. 즉, 새 서브-시퀀스는 항상 처음부터(0 부터) 번호 매기기를 시작한다. 도 4c는 이와 달리, 베이스 계층의 P-프레임들이 SP-프레임 쌍들로 대체되어 이미지 정보의 동일한 재구성을 가능하게 한다는 것만을 빼면 도 4a에 사용된 것에 상응하는 이미지 프레임 번호 매김을 이용한다.

[0067] 상술한 바와 같이, 본 발명의 방식은 B-프레임들을 이용해 구현될 수도 있다. 이것의 한 예가 도 5a, 5b, 및 5c에 도시된다. 도 5a는 시간 도메인 상의 한 비디오 시퀀스를 보인다. 그 시퀀스는 P-프레임들인 P1, P4, 및 P7과 그들 사이에 놓인 B-프레임들을 포함하고, 시간적 예측과 관련해 B-프레임들의 종속성이 화살표로서 도시되고 있다. 도 5b는 비디오 시퀀스 이미지 프레임들의 바람직한 그루핑을 보인 것으로, 이때 도 5a에서 보여지는 상호 종속성들이 표시되고 있다. 도 5b는 서브-시퀀스에 특정한 이미지 프레임 번호 매기기를 도시하고 있으며, 여기서 새 서브-시퀀스는 항상 0부터 이미지 프레임들의 번호 매김을 시작한다. 도 5c는 시간적 예측 순으로 연속되는 이미지 프레임 번호 매기기를 도시하고 있고, 여기서 이어지는 참조 프레임은 항상 이전 인코딩된 참조 프레임으로서 다음 이미지 번호를 받는다. 이미지 프레임 (B1.8) (및 (B2.10))은 어떤 다른 프레임에도 참조 예측 프레임으로서 제공되지 않으므로, 이미지 프레임 번호 매기는데 영향을 미치지 않는다.

[0068] 상기 예들은 비디오 시퀀스 코딩의 스케일 가능성이 본 발명의 방법을 이용해 어떻게 조정될 수 있는지에 대한 서로 다른 선택적 대안들을 도시한다. 그 비디오 시퀀스를 재생하는 단말 장치의 관점으로 볼 때, 더 많은 스케일 가능 계층들이 사용가능하거나, 더 많은 스케일 가능 계층들을 디코딩할 수 있을수록, 더 좋은 화질을 얻게 된다. 즉, 이미지 정보량과 그 정보를 전송하는데 사용되는 비트율의 증가가 시간 또는 공간적 해상도나 이미지 데이터의 공간적 품질을 향상시킨다. 마찬가지로, 더 많은 수의 스케일 가능 계층들은 디코딩을 수행하는 단말 장치의 프로세싱 용량에 대해 훨씬 더 높은 요구들을 역시 마련한다.

[0069] 또, 상기 예들은 서브-시퀀스들을 이용함으로써 얻어지는 이점을 예시한다. 이미지 프레임 식별자들을 가지고, 서브 시퀀스 안의 각 이미지 프레임의 다른 이미지 프레임들에 대한 종속성들이 분명한 방식으로 표시된다. 따라서 서브-시퀀스는, 그 비디오 시퀀스의 일련의 이미지 프레임들에 대한 디코딩에 영향을 주지 않으면서 필요할 때 비디오 시퀀스로부터 누락될 수 있는 독립적인 통일체를 형성한다. 그 경우, 동일하고/하거나 그에 종속되는 하위의 스케일 가능 계층들 상의 해당 서브-시퀀스 및 다른 서브-시퀀스들의 이미지 프레임들만이 디코

당되지 않는다.

- [0070] 비디오 시퀀스와 함께 전송되는 이미지 프레임 식별자 데이터는 비디오 시퀀스 헤더 필드들이나 비디오 시퀀스 전송에 상요될 전송 프로토콜의 헤더 필드들에 포함됨이 바람직하다. 즉, 예측된 이미지 프레임들의 식별자 데이터는 코딩된 비디오 시퀀스의 이미지 데이터 안에 포함되지 않고, 항상 헤더 필드들 안에 들어감으로써, 이미지 프레임들의 종속성들은 실제 비디오 시퀀스의 이미지들을 디코딩하지 않고도 검출될 수 있다. 이미지 프레임들의 식별자 데이터는, 비디오 시퀀스가 전송을 위해 코딩되고 있을 때 가령 스트리밍 서버의 버퍼 메모리에 저장될 수 있다. 또, 서브-시퀀스들은 각 스케일 가능 계층 위에서 독립적으로 디코딩될 수 있는데, 이는 한 서브-시퀀스의 이미지 프레임은 동일한 스케일 가능 계층의 다른 서브-시퀀스들에 종속되지 않기 때문이다.
- [0071] 본 발명의 일실시예에 따르면, 서브-시퀀스에 구비된 이미지 프레임들은 동일한 스케일 가능 계층의 다른 서브-시퀀스들에도 종속될 수 있다. 그러면 이 종속성은 트래픽 정형화(shaping)를 수행하는 스트리밍 서버 등으로 시그널링되어야 하는데, 이는 동일 계층 상에 위치한 독립적 서브-시퀀스들은 전송될 비디오 시퀀스로부터 별도로 제거될 수 없기 때문이다. 시그널링을 수행하는 바람직한 방법이, 가령 해당 서브-시퀀스가 의존하는 계층-서브-시퀀스 쌍들을 리스트화함으로써, 전송될 이미지 프레임 식별자들 안에 그 종속성을 포함시키도록 하는 것이다. 이것은 또한 동일한 스케일 가능 계층의 다른 서브-시퀀스로부터의 종속성을 나타내는 바람직한 방법을 제공한다.
- [0072] 상기 예들은 이미지 프레임들이 이전 이미지 프레임들로부터 시간적으로 예측되는 상황을 예시한다. 그러나 어떤 코딩 방식들에서는, 참조 화상 선택이, 시간적으로 이어지는 이미지 프레임들로부터의 이미지 프레임들의 이미지 정보에 대해 예측하는 것 역시 포함하도록 더 확장되었다. 참조 화상 선택은 시간적으로 상이한 스케일 가능 이미지 프레임 구조들을 생성하기 위해 가장 다각화된 수단을 제공하고 비디오 시퀀스의 예러 감도(sensitivity)가 줄어들 수 있도록 한다. 참조 화상 선택에 기반한 코딩 기술들 중 하나가 인트라 프레임 연기(postponement)이다. 인트라 프레임은 비디오 시퀀스 안에서 시간적으로 "바른" 위치에 자리하지 않고 그 위치가 시간적으로 뒤에 온다. 인트라 프레임의 "바른" 위치와 실제 위치 사이에 있는 비디오 시퀀스 이미지 프레임들은 해당 인트라 프레임으로부터 시간적으로 역행하는 방향으로 예측된다. 이것은 당연히, 디스플레이될 모든 이미지 프레임들이 재현 순서에 맞게 코딩 및 배치될 수 있도록, 코딩되지 않은 이미지 프레임들이 충분히 긴 시간 주기 동안 버퍼링될 것을 요구한다. 본 발명에 따른, 인트라 프레임 전송 및 그 관련 서브-시퀀스들의 결정은 도 6을 참조해 이하에서 설명될 것이다.
- [0073] 도 6a는 어떤 비디오 시퀀스 부분을 도시하며, 여기서 인트라 프레임은 단일 I-프레임(600)을 포함하고, 이것은 도 6에 도시된 위치로 시간상 이동되었으나, 그 비디오 시퀀스에서의 "올바른" 위치는 첫번째 이미지 프레임이었을 것이다. "올바른" 위치와 실제 위치(600) 사이의 비디오 시퀀스 이미지 프레임들은 시간적으로 역행하여 예측된다. 이는 인헨스먼트 계층 1 안에 코딩된 서브-시퀀스에 의해 예시되며, 시간적으로 역행한 첫번째 예측 이미지 프레임(602)인 P-프레임 (P.0.0)을 포함한다. 이 프레임은 이전 이미지 프레임(604), 즉 P-프레임 (P.0.1)의 시간적 예측을 위해 사용되고, 이미지 프레임(604)은 이미지 프레임(606), 즉 P-프레임 (P.0.2)를 예측하는데 사용되며, 마지막으로 프레임(606)은 이미지 프레임(608), 즉, 이 비디오 시퀀스에서 I-프레임(600)의 "올바른" 위치였을 위치에 있는 P-프레임 (P.0.3)을 예측하는데 사용된다. 이와 마찬가지로, 베이스 계층상의 I-프레임(600)이 네 개의 P-프레임들(610, 612, 614, 616) (P.0.0), (P.0.1), (P.0.2), 및 (P.0.3)을 포함하는 서브-시퀀스의 시간적으로 전진하는 방향의 예측을 위해 사용된다.
- [0074] 이 예에서, 역행 예측 이미지 프레임들이 전진 예측 이미지 계층들 보다 하위 계층에 위치한다는 사실은, 예시할 목적으로 역행 예측 이미지 프레임들이 이 코딩 예에서는 전진 예측 이미지 프레임들 보다 주관적으로 덜 중요하다고 간주되고 있음을 나타낸다. 당연히, 상기 서브-시퀀스들은 둘 다 동일한 계층에 위치될 수 있으며, 그 경우 그들은 같다고 간주되고, 역행 예측 서브-시퀀스가 상위 계층에 있는 경우 그 시퀀스가 주관적으로 더 중요하다고 간주될 수 있다.
- [0075] 도 6b 및 6c는 도 6a에 따른 비디오 시퀀스를 코딩하는 어떤 다른 선택사항을 도시한 것이다. 도 6b에서 전진 및 역행 예측 서브-시퀀스들은 베이스 계층 위에 위치하며, I-프레임은 인트라 계층에만 위치한다. 이 계층의 전진 예측 서브-시퀀스는 제2서브-시퀀스이고 그 서브-시퀀스 식별자는 1이다. 도 6c에서, I-프레임과 그에 기반한 전진 예측 서브-시퀀스는 베이스 계층에 위치하고, 한편 역행 예측 서브-시퀀스는 인헨스먼트 계층 1에 위치된다.
- [0076] 또, 본 발명의 바람직한 실시예에 따르면, 상술한 스케일 가능성은 비디오 시퀀스로서의 장면 전환이라고 알려진 코딩을 위해 활용될 수 있다. 뉴스, 뮤직 비디오 및 영화 보여주기과 같은 비디오물은 흔히 각각의 이미지 구

성 장면들 사이에서 빠른 장면 바꾸기(cut)들을 포함한다. 때때로 이 장면 바꾸기들은 급작스럽지만, 장면 전환이라고 알려지는 방식은 흔하게 사용되는 것으로 한 장면에서 다른 장면으로의 전환은 디밍(dimming), 와이핑(wiping), 모자이크 디졸브(dissolving) 또는 이전 장면의 이미지 프레임들을 스크롤링하고 마찬가지로 나중 장면의 이미지 프레임들을 제공함으로써 발생된다. 코딩 효율성의 관점에서 볼 때, 장면 전환의 비디오 코딩은 흔히 가장 문제가 많은데, 이는 장면 전환 도중에 나타나는 이미지 프레임들이 장면을 제거 및 개시하는 이미지 프레임들에 대한 정보를 포함하고 있기 때문이다.

[0077] 통상적인 장면 전환인 페이딩(fading)은 제1장면의 이미지 프레임들의 강도나 휘도를 0으로 점진적으로 줄이면서 제2장면의 이미지 프레임들의 강도는 점진적으로 최대값까지 증가시킴으로써 수행된다. 이 장면 전환을 크로스-페이디드(cross-faded, 교차 페이딩) 장면 전환이라고 한다.

[0078] 일반적으로 말할 때, 컴퓨터로 만든 이미지는 계층들이나 이미지 오브젝트들로 이뤄진다고 생각될 수 있다. 각 오브젝트는 적어도 세 가지 정보 유형들을 참조해 규정될 수 있다. 그 세 정보 유형은, 이미지 오브젝트의 구조, 그 모양과 투과도, 및 이미지의 배경 및 다른 이미지 오브젝트들과 관련된 계층 순서(깊이)를 말한다. 모양과 투과도는 흔히 알파 평면이라고 알려진 것을 이용해 정해지는 것으로, 불투명도를 측정하고 그 값은 보통 각 이미지 오브젝트마다 따로 따로 정해지며, 보통 비투과적이라고 판단되는 배경은 배제한다.

[0079] 그래서, 배경과 같은 비투과 이미지 오브젝트의 알파 평면 값은 1.0으로 정해질 수 있는 반면, 완전 투과 이미지 오브젝트의 알파 평면 값은 0.0이 된다. 그 사이 값들은, 배경 및 해당 이미지 오브젝트에 비해 더 큰 깊이를 가진 적어도 부분적으로 중복되는 이미지 오브젝트들에 비한, 화면 내 특정 이미지 오브젝트의 가시도의 세기를 규정한다.

[0080] 이미지 오브젝트들의 모양, 투과도 및 깊이 위치(depth position)에 따른 계층들에서의 이미지 오브젝트들의 중첩을 장면 합성이라 부른다. 실제로, 그 방식은 가중된 평균치들의 사용에 기반한다. 먼저, 배경에 가장 가까운, 즉 그 깊이 위치에 따라 가장 깊은 이미지 오브젝트가 배경 위에 놓여지고 그 둘로 된 결합 이미지가 생성된다. 결합 이미지의 픽셀 값들은 배경 이미지 및 해당 이미지 오브젝트의 알파 평면 값들로 가중된 평균으로서 형성된다. 결합 이미지의 알파 평면 값은 1.0으로 정해지고, 그런 다음 다음 이미지 오브젝트에 대한 배경 이미지로서 사용된다. 이 프로세스는 모든 이미지 오브젝트들이 그 이미지에 부가될 때까지 계속된다.

[0081] 이하에서, 본 발명의 바람직한 실시예에 따른 방식이 설명될 것이며, 여기서 비디오 시퀀스 스케일 가능 계층들은 상술한 이미지 프레임들의 이미지 오브젝트들 및 그들의 정보 유형과 결합되어, 장면 전환에 대해 역시 양호한 압축 효율을 가진 스케일 가능 비디오 코딩을 제공한다.

[0082] 본 발명의 이러한 실시예가, 하나는 교차 페이딩 장면 전환을 다른 하나는 급격한 장면 전환 등을 이용하여, 예를 들어가며 단순화된 방식으로 이하에서 설명될 것이다. 장면 전환 중에 디스플레이될 이미지 프레임들은 보통 두 중첩 이미지 프레임들로 이뤄지며, 그 이미지 프레임들 중 제1이미지 프레임은 제1이미지 장면을 제2이미지 프레임은 제2이미지 장면을 포함한다. 이미지 프레임들 중 하나는 배경 이미지로서 사용되고 전경(foreground) 이미지로 불리는 다른 하나는 그 배경 이미지 위에 위치된다. 배경 이미지의 불투명도, 즉 비투과도 값은 일정하다. 즉, 픽셀-고유의 알파 평면값들은 조정되지 않는다.

[0083] 본 발명의 이러한 실시예에서, 배경과 전경 이미지들은 모두 스케일 가능 계층에 따라 규정될 수 있다. 이것도 도 7에 예시되고 있는데, 도 7은 본 발명의 장면 전환 도중에 두 개의 서로 다른 장면들의 이미지 프레임들이 어떻게 스케일 가능 계층들 위에 놓여질 수 있는가에 대한 예를 보인 것이다. 도 7은 베이스 계층 상에 위치한 제1 (제거) 장면의 제1이미지 프레임(700)을 보인다. 이미지 프레임(700)은 모션 보상된 시간적 예측을 이용해 결정되지 않았던 이미지 정보를 포함하는 I-프레임일 수 있고, 아니면 이전 이미지 프레임들로부터 예측된 모션 보상 이미지 프레임인 P-프레임일 수 있다. 제2 (시작) 장면의 코딩은 시간적으로 다음 이미지 프레임 도중에 시작하며, 본 발명에 따라 그 장면의 이미지 프레임들은 베이스 계층 상에 역시 위치된다. 그러면 제2 (제거) 장면의 이미지 프레임들(702, 704)은 인헨스먼트 계층 1에 위치한다. 이들 이미지 프레임들은 보통 P-프레임들이다.

[0084] 이 실시예에서, 제2 (시작) 장면의 이미지 프레임들은 적어도 장면 전환 도중에는 베이스 계층 상에 위치된다. 그 장면의 제1이미지 프레임(706)은 보통 I-프레임이며, 그 장면의 연속되는 이미지 프레임들을 시간적으로 예측하는데 사용된다. 결국, 제2장면의 연속 이미지 프레임들은, 도 7에 도시된 프레임들(708 및 710)과 같이 시간적으로 예측된 프레임들인, 통상적 P-프레임들이다.

[0085] 본 발명의 바람직한 실시예에 따르면, 이러한 이미지 프레임들의 스케일 가능 계층들 상의 위치시키기는, 베이

스 계층에 있는 이미지 계층을 항상 최대 불투명도(100%), 또는 비투과도의 배경 이미지로 결정함으로써 교차 페이딩 장면 전환을 구현하는데 사용될 수 있다. 장면 전환 도중, 인헨스먼트 계층들 위에 놓여진 이미지 프레임들은 배경 이미지 위로 놓여져서 그들의 불투명도는 가령 적절한 필터들에 의해 그 프레임들이 비투과적인 것에서 투과적인 것으로 차츰 바뀌도록 조정된다.

[0086] 도 7의 비디오 시퀀스에서, 제1베이스 계층 이미지 프레임(700) 도중에는 그 보다 하위의 스케일 가능 계층들 위에 어떤 이미지 프레임들도 존재하지 않는다. 이 때의 시간 동안, 제1이미지 프레임(700) 만이 비디오 시퀀스 안에 코딩된다.

[0087] 베이스 계층의 다음 이미지 프레임(706)은 새(제2의) 장면을 개시시키며, 그 동안 이미지 프레임(706)에는 그것을 배경 이미지로 놓는 깊이 위치가 제공되고, 그 불투명도 값은 최대로 설정된다. 베이스 계층의 이미지 프레임(706)과 시간적으로 동일하게, 인헨스먼트 계층 1 위에 제거(제1) 장면의 이미지 프레임(702)이 존재한다. 교차 페이딩 장면 전환이 일어날 수 있도록 하기 위해, 프레임(702)의 투과도가 증가되어야 한다. 도 7의 예는 이미지 프레임(702)의 불투명도가 67%로 정해지고, 또 이미지 프레임(702)에 그것을 전경 이미지로 결정하는 깊이 위치가 주어진다고 진제한다. 이 순간 동안, 이미지 프레임들(706 및 702)을 결합한 이미지가 비디오 시퀀스 안에 코딩되고, 이때 이미지(706)는 배경 상의 보다 약한 이미지로서 보여지고 이미지(702)는 그 불투명도 값이 실질적으로 높기(67%) 때문에 전면에서의 보다 강한 이미지로서 보여진다.

[0088] 시간적으로 이어지는 다음 이미지 프레임 도중에는, 베이스 계층 상의 제2장면의 제2이미지 프레임(708)이 존재하고, 이에 따라 프레임(708)에는 자신을 배경 이미지로 정하는 깊이 위치가 주어지고 그 불투명도 값은 최대로 정해진다. 인헨스먼트 계층 1은 시간적으로 동시에 제거되는 (제1) 장면의 마지막 이미지 프레임(704)을 더 포함하며, 그 프레임의 불투명도 값은 33%로 정해진다. 또한, 이미지 프레임(704)에는 그 자신을 전경 이미지로도 정하는 깊이 위치 역시 주어진다. 결국, 이 순간 동안, 이미지 프레임들(708 및 704)이 결합된 이미지가 비디오 시퀀스 안에 코딩되고, 이때 이미지(708)는 배경 위의 보다 강한 이미지로서 디스플레이되고 이미지(704)는 그 불투명도 값이 더 이상 33% 이상이 아니기 때문에 전경 위의 보다 약한 이미지로서 디스플레이된다.

[0089] 시간상 다음에 이어지는 이미지 프레임 동안, 베이스 계층은 제2장면의 제3이미지 프레임(710)을 포함한다. 제1장면이 제거되었기 때문에, 이미지 프레임(710) 만이 비디오 시퀀스 안에 코딩되며, 제2장면의 디스플레이는 프레임(710)부터 지속된다.

[0090] 예로 든 상기 개시 내용은 코딩 효율성의 면에서 볼 때 유리한 방식으로 교차 페이딩 장면 전환을 구현하기 위해, 본 발명에 따른 이미지 프레임들의 스케일 가능 계층들 상으로의 위치시키기(positioning)를 설명하였다. 그러나, 비디오 시퀀스가 전송되거나 디코딩되고 있을 때, 비디오 시퀀스의 비트율이 데이터 전송을 위한 대역폭의 최대값 및/또는 사용 가능한 단말장치 디코딩율에 따라 조정되어야만 하는 상황이 발생한다. 이러한 유형의 비트율 제어는, 장면 전환이 종래의 비디오 코딩 방법을 이용해 구현될 때 문제를 일으킨다.

[0091] 본 발명의 바람직한 실시예에는 이제 하나 이상의 스케일 가능 계층들, 또는 그 안에 포함된 독자적으로 디코딩 가능한 서브-시퀀스들이 비디오 시퀀스로부터 제거될 수 있도록 함으로써, 비디오 시퀀스의 비트율이 떨어질 수 있게 하고, 또 동시에 비디오 시퀀스가 이미지 주파수를 감소시키지 않으면서 디코딩 될 수 있게 한다. 도 7에 따른 이미지 프레임 위치시키기에 있어, 이러한 것은 인헨스먼트 계층 1을 비디오 시퀀스로부터 제거함으로써 구현될 수 있다. 따라서 비디오 시퀀스는 베이스 계층의 이미지 프레임들(700, 706, 708, 및 710)을 디스플레이 하는 데만 사용된다. 즉, 제1(제거) 장면으로부터 제2(개시) 장면으로의 직접 전환이 급격한 장면 전환의 형태, 즉 제1장면의 이미지 프레임(700)에서 제2장면을 시작하게 하는 I-이미지 프레임(706)으로 바로 전환하는 형태로 일어난다. 따라서 이러한 전환은 교차 페이딩 장면 전환이 아니라 급격한 장면 전환이 된다. 그럼에도 불구하고, 이 장면 전환은 비디오 시퀀스 이미지의 품질에 영향을 주지 않으면서 바람직한 방식으로 수행될 수 있으며, 보는 사람은 보통 어떤 불안하거나 잘못된 방식으로든 교차 페이딩 장면 전환 대신 수행된 급격한 장면 전환을 경험하지 않는다. 이와 반대로, 종래의 구성은 스케일 가능 계층들이 제거되는 것을 허용하지 않기 때문에, 보는 사람은 화면이 단속적이고도 불안하다는 것을 알게 될 것이다.

[0092] 따라서 본 발명은 비디오 시퀀스의 서로 다른 서브-시퀀스들에 대한 정보, 즉 그들의 평균 비트율, 전체 비디오 시퀀스와 관련된 위치, 듀레이션 및 계층들에 관한 상호의존성의 정보를 포함하는 스트리밍 서버에서, 멀티미디어 데이터 트래픽 정형화를 수행하는 바람직한 수단을 제공한다. 스트리밍 서버는 또 데이터 전송에 사용가능한 대역폭 및/또는 단말 장치의 디코딩율의 최대값을 결정한다. 이 정보에 기초해, 스트리밍 서버는 얼마나 많은 스케일 가능 계층들과 어떤 서브-시퀀스들이 비디오 시퀀스로서 전송되는지를 결정한다. 그러면, 스케일 가능 계층들의 개수에 대한 대강의 조정을 먼저 수행하고 그 다음에 보다 정밀한 서브-시퀀스-고유의 조정이 쉽게

이뤄질 수 있게 됨으로써 필요할 때마다 비트율 제어가 이뤄질 수 있다. 가장 간단한 것으로서, 비트율 제어 수단이 특정 서브-시퀀스가 비디오 시퀀스에 추가될 것인지 제거될 것인지에 대한 서브-시퀀스-고유의 결정을 내릴 수 있다. 제거할 경우, 전체 서브-시퀀스들을 비디오 시퀀스에서 제거함이 권장할 만한데, 이는 개별 이미지들의 제거가 동일한 서브-시퀀스의 다른 이미지들에 에러를 일으킬 수 있기 때문이다. 같은 이유로, 하위 인핸스먼트 계층의 모든 서브-시퀀스들도 그들이 제거된 상위 계층의 서브-시퀀스에 의존하고 있었다면 제거되어야 한다. 하나의 동일한 스케일 가능 계층 상에서 독립적인 서브-시퀀스들이 존재하면, 이전의 서브 시퀀스에 의존하는 서브-시퀀스들은 그 이전의 서브-시퀀스가 제거될 때 제거되어야 한다.

[0093] 이미지 프레임 식별자 데이터가 전송될 비디오 시퀀스에 추가될 때, 트래픽 정형화 또한, 가령 인터넷 라우터, 다른 게이트웨이들, 또는 기지국이나 이동 통신 네트워크의 기지국 제어기 등의, 비디오 시퀀스 전송에 이용될 전기통신 네트워크 구성요소에서 수행될 수 있다. 서브-시퀀스 정보를 유지하고 처리할 수 있도록 된 네트워크 구성요소들을 위해, 트래픽 정형화는 여분의 메모리와 프로세싱 용량을 가져야만 한다. 이런 이유 때문에 네트워크 내에서 수행될 트래픽 정형화는 DiffServ, 즉 어떤 IP-기반 네트워크들에 의해 지원되는 절차인 차별화된 서비스들과 같은 간단한 처리 방식을 이용해 실행될 가능성이 크다. DiffServ 방식에 있어서, 각 IP 데이터 패킷에는 우선순위가 할당되어져, 보다 높은 우선순위의 데이터 패킷들이 낮은 우선순위의 패킷들 보다 빠르고 안전하게 수신자에게 전달된다. 이것은 스케일 가능 계층에 고유한 것만이 아닌 서브-시퀀스에 고유한 우선순위를 정하여 매우 진보한 우선순위화를 가능하게 함으로써 본 발명의 스케일 가능성에 바람직하게 적용된다.

[0094] 전송될 비디오 시퀀스에 이미지 프레임 식별자 데이터를 추가시키는 많은 선택 가능한 방법들이 존재한다. 또, 어떤 식별자 데이터를 비디오 시퀀스 안에 포함시키지 않는 것 역시 가능하며, 이 경우 트래픽 정형화는 스트리밍 서버에서만 수행된다. 식별자 데이터는 비디오 시퀀스의 헤더 필드들 또는 RTP (실시간 프로토콜)과 같은, 사용할 전송 프로토콜의 헤더 필드들 안에 포함될 수 있다. 바람직한 실시예에 따르면, 식별자 데이터는 추가 보완적 개선 정보(SEI, Supplemental Enhancement Information) 메커니즘을 이용해 전송될 수 있다. SEI는 비디오 데이터 콘텐츠와 동기되어 전송되는 데이터 운반 메커니즘을 제공하여, 비디오 시퀀스의 디코딩 및 디스플레이를 돕는다. 특히 계층 및 서브-시퀀스 정보를 전송하는데 사용될 때의 SEI 메커니즘이 ITU-T 표준 문서 ITU-T 권고안 H.264(ISO/IEC 14496 -10:2002), 부록 D에 상세히 개시되어 있다. 별개의 전송 프로토콜 또는 메커니즘이 식별자 데이터 전송에 사용되는 이런 경우들에서, 트래픽 정형화는 전송 경로에 있는 네트워크 구성요소들 중 하나에서 역시 수행될 수 있다. 또, 수신 단말 장치가 디코딩을 제어할 수 있다.

[0095] 인코더 또는 디코더가 참조 화면 선택을 지원할 때, 비디오 시퀀스 코딩은 서로 다른 이미지 프레임들 사이의 관계가 하나 이상의 다른 이미지 프레임들로부터 시간적으로 예측될 수 있도록, 디코딩된 이미지 프레임들이 그 코딩 이전에 버퍼링될 것을 필요로 한다. 이미지 프레임 버퍼링은 적어도 두 가지 서로 다른 방식인, 슬라이딩 윈도우(sliding windowing)이나 적응적 버퍼 메모리 제어로서 이뤄질 수 있다. 슬라이딩 윈도우 방식, 마지막으로 코딩된 M 이미지 프레임들이 버퍼로 사용된다. 버퍼 메모리에 있는 프레임들은 디코딩 및 재구성 형식으로 되어 있어, 코딩시 참조 이미지들로서 사용될 수 있도록 한다. 코딩이 진행될 때, 이미지 프레임 버퍼링은 FIFO 원리(Fisrt-In-Fisrt-Out, 먼저 들어간 것이 먼저 나오는 원리)에 따라 동작한다. 통상적 B-이미지들과 같이 참조 이미지로 사용되지 않는 이미지들은 버퍼에 저장될 필요가 없다. 이와 다른 선택 사항으로서, 버퍼링은 적응적 버퍼 메모리 제어로 구현될 수도 있으며, 이 경우 이미지 버퍼링은 FIFO 원리에 국한되지 않으며 필요로 되지 않는 이미지 프레임들이 프로세스 도중에 버퍼로부터 제거될 수 있고, 혹은 마찬가지로 어떤 이미지 프레임들이 나중의 이미지 프레임들에 대한 참조 이미지들로서 필요로 될 때에는 보다 긴 시간 주기 동안 버퍼에 저장될 수 있다. 알려진 참조 화상 선택은 버퍼 메모리에 있는 이미지 프레임들을 특정 순서로 색인화함으로써 구현되고, 그 이미지 색인들은 모션 보상 등과 관련하여 한 이미지를 참조하는데 사용된다. 이러한 색인화 방식은, 가령, 모션-보상 참조 이미지들이 시그널링되어야 할 때 특정 이미지를 참조하는데 있어, 일반적으로 이미지 번호를 사용하는 것과 비교해 더 우수한 압축 효율을 제공한다.

[0096] 상기 참조 이미지 색인화 방법은 에러 전송에 민감한데, 그 이유는 전송자의 인코더의 버퍼와 수신자의 디코더의 버퍼가 상호적으로 상응하는 재구성 이미지들을 동일한 순서로 포함하고 있어서 인코더 및 디코더가 모두 동일한 색인화 순서를 형성함을 보장할 수 있어야 하기 때문이다. 이미지 프레임들이 인코더 및 디코더의 버퍼들에서 서로 다른 순서로 색인화되면, 디코더에서 틀린 참조 이미지가 사용될 수 있다. 이를 방지하기 위해, 디코더는 인코더가 고의적으로 비디오 시퀀스로부터 제거했던 이미지 프레임들과 서브-시퀀스들을 감안하도록 제어될 수 있어야 한다는 것이 중요하게 된다. 이 경우 이미지 프레임 번호 매기기는, 디코더가 통상적으로 에러들이라고 해석하여 잃어버렸다고 해석한 이미지 프레임들을 재구성하려고 시도하는, 공백들을 포함할 수 있다. 이러한 이유 때문에, 인코더가 디코더에게, 전송된 이미지 프레임들의 이미지 번호 매기기에 있어서의 불연속성

이 의도적이라는 것을 알릴 수 있는 것이 중요하다.

- [0097] 이에 반응하여, 그리고 슬라이딩 윈도우가 이미지 프레임을 버퍼링하는데 사용되었다고 할 때, 디코더는 버퍼 메모리를, 그 내용이 완전히 무작위적이고 잃어버린 이미지 수에 해당하는, 다수의 이미지 프레임들이라고 간주한다. 이 무작위적인 이미지 프레임들은, 그 프레임들이 실제 비디오 시퀀스에 속하지 않고 단지 버퍼 메모리 관리용으로 입력된 버퍼 칸 채우기 프레임들임을 나타내도록 "무효(Invalid)"라는 식별자가 지정된다. 버퍼 채우기 프레임은 당연히 메모리 지시자들만을 이용해 구현될 수 있다. 즉 어떤 데이터도 그 버퍼 메모리에 입력되지 않음이 바람직하며, 다만 메모리 관리는 포괄적 "무효" 프레임에 대한 참고 사항만을 저장하는데 이용된다. 실제 비디오 시퀀스의 이미지 프레임들의 입력은 누락된 이미지 번호들에 의해 지시된 채우기 프레임들의 개수가 버퍼에 입력된 다음에, 바른 이미지 프레임 번호로부터 재개되며, 이러한 것이 인코더 및 디코더의 버퍼 메모리들이 바람직하게는 동기 상태로 유지될 수 있도록 한다. 디코딩 도중에 이미지 번호에 대한 참조가 검출되고 나서 그것이 버퍼에 위치한 채우기 프레임을 나타내는 것임이 밝혀질 때, 실제 참조 이미지를 재구성하기 위해 디코더에서 가령 인코더에 해당 참조 이미지를 재전송하라는 요청을 함으로써 여러 정정 조치가 시작된다.
- [0098] 또, 본 발명의 방법은 별개의 버퍼 메모리들이 서로 다른 스케일 가능 계층들에 대해 사용될 수 있도록, 혹은 이와 마찬가지로 서브-시퀀스에 고유하게 사용될 수 있도록 한다. 따라서 각각의 스케일 가능 계층들은 개념적으로 분리되고 슬라이딩 윈도우 원리에 따라 동작하는 개별적인 버퍼 메모리를 포함한다. 이와 유사하게, 각각의 서브-시퀀스 역시, 개념적으로 분리되어 역시 슬라이딩 윈도우 원리에 따라 동작하는 버퍼 메모리가 제공된다. 이는 한 서브-시퀀스가 중단될 때 그 버퍼 메모리가 항상 비어 있게 됨을 의미한다. 각각의 버퍼 메모리들은 일반적인 슬라이딩 윈도우 버퍼링이 부적절하여 대신 실제로 적응적 버퍼 메모리 관리가 필요로 되는 어떤 상황에서 시그널링에 대한 수요를 줄일 수 있는 바람직한 방식으로 이용될 수 있다.
- [0099] H.26L 표준은 화상 순서 카운트를 출력 순서 상의 화상 위치로 규정한다. H.26L 표준에 명시된 디코딩 프로세스는 화상 순서 카운트를 이용해 B 슬라이스의 참조 화상들에 대한 디폴트 색인 순서를 결정하고, 모션 벡터 예측시의 벡터 스케일링과 B 슬라이스의 함축 모드 가중 예측(implicit mode weighted prediction)에 대한 프레임들 및 필드들 사이의 화상 순서 차를 나타내고, 디코딩 순서로 연속된 슬라이스들이 서로 다른 화상들에 속할 때를 결정한다. 화상 순서 카운트는 각 화상별로 코딩 및 전송된다.
- [0100] 본 발명의 일실시예에서, 디코더는 화상 순서 카운트를 이용하여 화상들이 시간적으로 중복됨을, 즉 같은 화상 순서 카운트를 가진 화상들이 시간적으로 중복된다고 추론한다. 디코더가 최상위 수신 계층의 화상만을 출력함이 바람직하다. 계층 정보가 없을 때, 디코더는 디코딩 순서에서 시간적으로 가장 최근에 중복된 화상이 최상위 수신 계층상에 존재한다고 추론한다.
- [0101] 위에서는 스케일 가능한 압축 비디오 시퀀스를 발생할 목적으로 비디오 프레임들을 코딩하는 방법이 설명되었다. 실제 방법은 도 1의 압축기(108)와 같이, 기존의 어떤 비디오 인코더라도 될 수 있는 비디오 인코더 내에서 수행된다. 예를 들어, ITU-T 권고안 H.263 또는 H.26L에 따른 비디오 인코더가 사용될 수 있으며, 그 비디오 인코더는 본 발명에 의해, 서브-시퀀스의 적어도 일부가 I-프레임들을 코딩함으로써 형성되는 제1서브-시퀀스를 비디오 시퀀스 안에 생성하고; 서브-시퀀스의 적어도 일부가 적어도 P- 또는 B-프레임들을 코딩함으로써 형성되고 제1서브-시퀀스의 적어도 한 비디오 프레임으로부터 예측되는 적어도 제2서브-시퀀스를 비디오 시퀀스 안에 형성하고; 적어도 제2서브-시퀀스의 비디오 프레임들의 식별 데이터를 비디오 시퀀스 안에 정하도록 구성된다.
- [0102] 본 발명의 방법에 따르면, 특정한 스케일 가능 계층의 각 서브-시퀀스는 독립적으로 디코딩될 수 있음이 바람직한데, 이때 당연히 상위 스케일 가능 계층들 및 같은 스케일 가능 계층의 아마도 다른 서브-시퀀스들로부터의 의존성이 감안되어야 한다. 상술한 것과 같이 스케일 가능하게 압축된 비디오 시퀀스는, 적어도 일부가 적어도 I-프레임들을 코딩함으로써 형성된 것인 비디오 시퀀스의 제1서브-시퀀스를 디코딩하고, 적어도 일부가 적어도 P- 또는 B-프레임들을 코딩함으로써 형성된 것이고 구비된 적어도 한 비디오 프레임이 제1서브-시퀀스의 적어도 한 비디오 프레임으로부터 예측되었던 비디오 시퀀스의 적어도 제2서브-시퀀스를 디코딩하며, 비디오 시퀀스의 적어도 제2서브-시퀀스에 구비된 비디오 프레임들의 식별자 및 종속성 데이터를 결정하고, 그 서브-시퀀스들의 종속성들에 기반해 비디오 시퀀스의 적어도 일부를 재구성함으로써 디코딩된다.
- [0103] 실제 디코딩은 도 1의 압축해제기(118)와 같은 임의의 알려진 비디오 디코더인 비디오 디코더에서 일어난다. 예를 들어, ITU-T 권고안 H.263 또는 H.26L에 따른 저비트율 비디오 디코더가 사용되며, 이러한 디코더는 본 발명에서, 일부가 I-프레임들을 코딩하여 형성된 것인 비디오 시퀀스의 제1서브-시퀀스를 디코딩하고; 적어도 일

부가 적어도 P- 또는 B-프레임들을 코딩함으로써 형성된 것이고 포함된 적어도 한 비디오 프레임이 제1서브-시퀀스의 적어도 한 비디오 프레임으로부터 예측된 것인 비디오 시퀀스의 적어도 제2서브-시퀀스를 디코딩하도록 구성된다. 비디오 디코더는, 비디오 시퀀스의 적어도 제2서브-시퀀스에 구비된 비디오 프레임들의 식별자 및 종속성 데이터를 결정하도록 되고, 그 서브-시퀀스들의 종속성들에 기반해 비디오 시퀀스의 적어도 일부를 재구성하도록 구성된다.

[0104] 본 발명의 스트리밍 시스템의 동작의 실질적 양상은, 인코더와 디코더가 인코더는 스트리밍 서버와 동작가능하게 연결되고 디코더는 수신 단말 장치와 동작가능하게 연결되도록 위치된다는 것이다. 그러나, 스트리밍 시스템의 서로 다른 구성요소들, 특히 단말 장치들이 멀티미디어 파일들의 두 방향 전송, 즉 송신 및 수신을 가능하게 하는 기능들을 포함할 수도 있다. 그러면 인코더와 디코더는 인코더 및 디코더 성능을 모두 합친 비디오 코덱이라 알려진 것의 형태로 구현될 수 있다.

[0105] 본 발명에 따르면, 상술한 스트리밍 시스템의 기능적 구성 요소들과, 스트리밍 서버, 비디오 인코더, 비디오 디코더 및 단말과 같은 구성요소들은 소프트웨어, 하드웨어 솔루션, 또는 그 둘의 결합에 의해 구현됨이 바람직하다. 본 발명의 코딩 및 디코딩 방법은 본 발명의 프로세스 단계들을 실행하기 위한 컴퓨터 판독가능 명령들을 구비한 컴퓨터 소프트웨어로서의 구현에 특히 적합하다. 인코더 및 디코더를 구현하는 바람직한 방법은 이들을, 개인용 컴퓨터(PC) 또는 이동국과 같은 컴퓨터 유사 장치에 의해 실행될 수 있는 프로그램 코드로서 저장 수단에 저장하여, 해당 장치에 코딩/디코딩 기능을 제공하는 것이다.

[0106] 다른 선택사항이 본 발명을 비디오 신호로서 구현하는 것으로, 상기 비디오 신호는 스케일 가능하게 압축된 비디오 시퀀스를 구비하고, 비디오 시퀀스는 적어도 제1 및 제2프레임 포맷에 따라 코딩된 비디오 프레임들을 포함하며, 이때 제1프레임 포맷에 따른 비디오 프레임들은 다른 비디오 프레임들에 독립적이고 제2프레임 포맷의 비디오 프레임들은 다른 비디오 프레임들 중 적어도 하나로부터 예측된다. 본 발명에 따르면, 해당 비디오 신호는, 적어도 일부가 적어도 제1프레임 포맷의 비디오 프레임들을 코딩함으로써 형성된 것인 적어도 제1서브-시퀀스; 적어도 일부가 적어도 제2프레임 포맷의 비디오 프레임들을 코딩함으로써 형성된 것이고, 포함하는 적어도 한 비디오 프레임이 제1서브-시퀀스의 적어도 한 비디오 프레임으로부터 예측되는 제2서브-시퀀스; 및 제2서브-시퀀스에 속하는 비디오 프레임들을 결정하는 적어도 한 데이터 필드를 포함한다.

[0107] 이 분야의 당업자에게는 기술이 발전함에 따라 본 발명의 기본적인 사상이 여러 방식으로 구현될 수 있음이 자명할 것이다. 따라서 본 발명과 그 실시예들은 상술한 예들에만 한정되지 않고 청구항들의 범주 안에서 변형될 수 있다.

발명의 효과

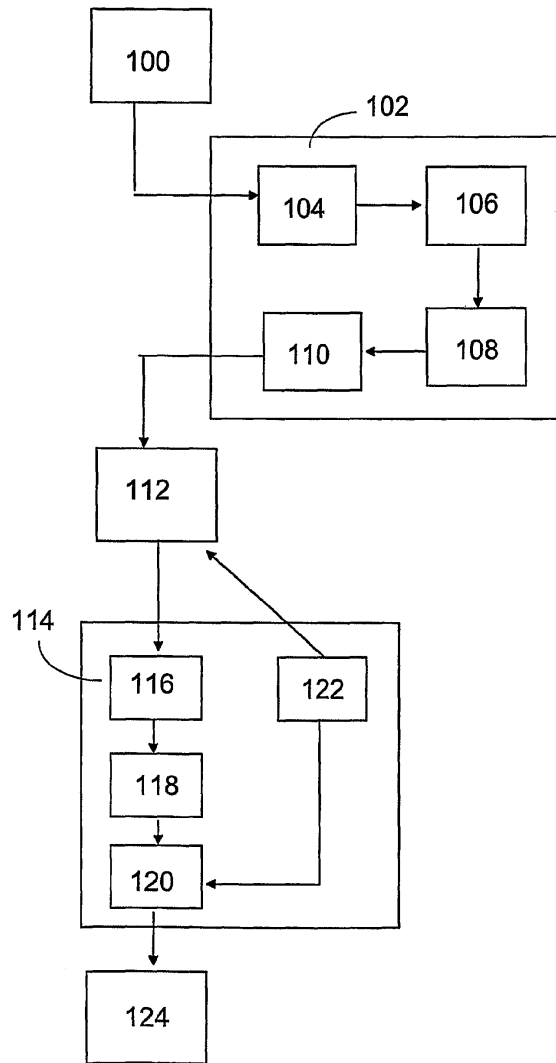
[0108] 본 발명에 의해, 이미지 프레임 넘버링시의 간극(gap)이 고의적이라는 정보를 디코더에 제공함으로써, 디코더가 불필요한 에러 정정 동작을 시작하지 않게 되고, 인코더와 디코더의 버퍼 메모리들의 동조(synchronism)가 유지될 수 있어, 기준 화상 선택 처리에 대한 거침없는 동작이 가능하게 된다.

도면의 간단한 설명

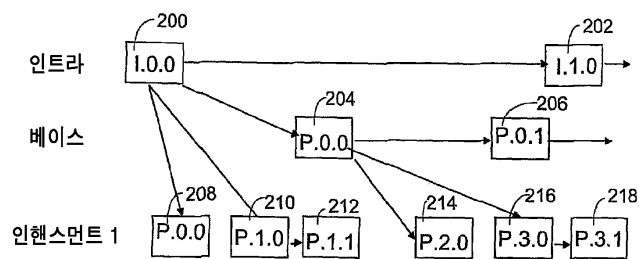
- [0001] 이하에서, 본 발명은 첨부된 도면 및 바람직한 실시예들과 관련지어 설명될 것이다.
- [0002] 도 1은 본 발명의 스케일 가능 코딩 계층이 적용될 수 있는 일반 멀티미디어 데이터 스트리밍 시스템을 도시한 것이다.
- [0003] 도 2는 본 발명의 바람직한 실시예의 스케일 가능 코딩 계층을 도시한 것이다.
- [0004] 도 3a 및 3b는 스케일 가능성의 조정을 위한 본 발명의 실시예들을 도시한 것이다.
- [0005] 도 4a, 4b 및 4c는 이미지 번호매김을 조정하기 위한 본 발명의 실시예들을 도시한 것이다.
- [0006] 도 5a, 5b 및 5c는 한 스케일 가능 코딩 계층에서 B-프레임들을 이용하기 위한 본 발명의 실시예들을 도시한 것이다.
- [0007] 도 6a, 6b 및 6c는 참조 화상 선택과 관련하여 본 발명의 바람직한 스케일 가능 코딩 계층들을 도시한 것이다.
- [0008] 도 7은 코딩 장면 천이를 위한 본 발명의 바람직한 실시예에 따른 구성을 도시한 것이다.

도면

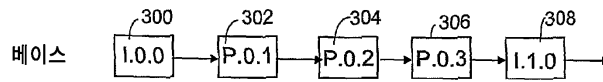
도면1



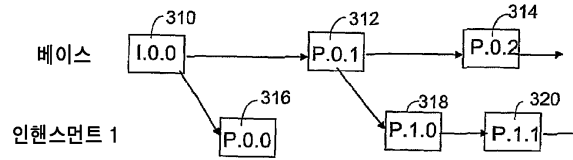
도면2



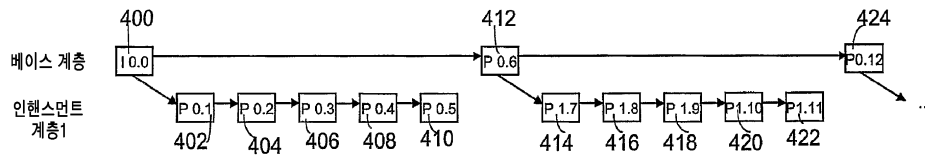
도면3a



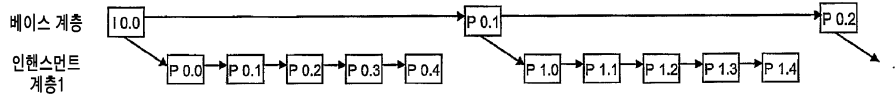
도면3b



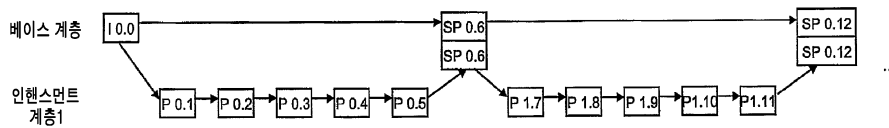
도면4a



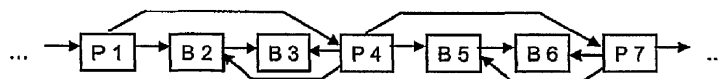
도면4b



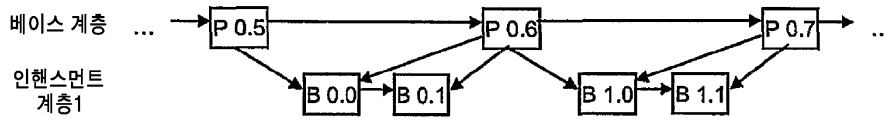
도면4c



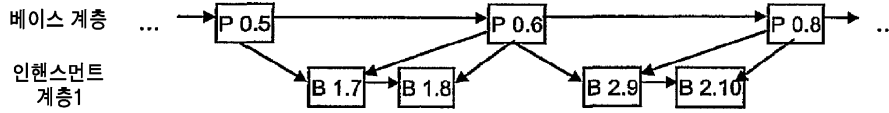
도면5a



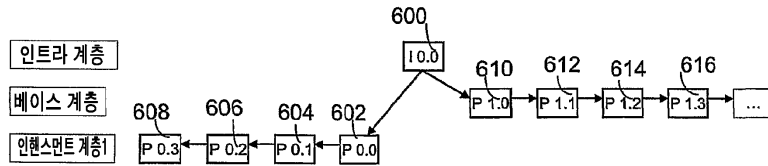
도면5b



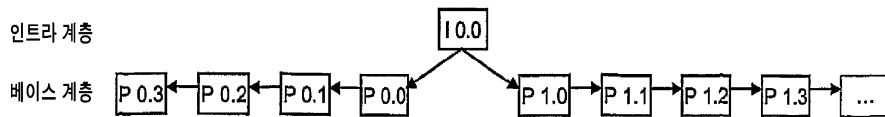
도면5c



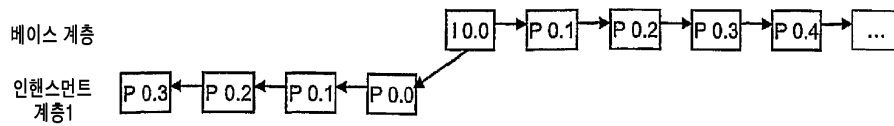
도면6a



도면6b



도면6c



도면7

