



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2014-0034149
(43) 공개일자 2014년03월19일

(51) 국제특허분류(Int. Cl.)
H04N 19/136 (2014.01) H04N 19/142 (2014.01)
H04N 21/2662 (2011.01)
(21) 출원번호 10-2013-7022649
(22) 출원일자(국제) 2012년01월26일
심사청구일자 없음
(85) 번역문제출일자 2013년08월27일
(86) 국제출원번호 PCT/US2012/022710
(87) 국제공개번호 WO 2012/103326
국제공개일자 2012년08월02일
(30) 우선권주장
61/437,193 2011년01월28일 미국(US)
61/437,223 2011년01월28일 미국(US)

(71) 출원인
아이 이오, 엘엘씨
미국 캘리포니아 94301, 팰로앨토 스위트 넘버3,
유니버시티 애비뉴 165
(72) 발명자
게레로, 로돌포 바르가스
미국, 캘리포니아 94301, 팰러 엘토, 스위트
넘버3, 유니버시티 애비뉴 165
(74) 대리인
특허법인씨엔에스

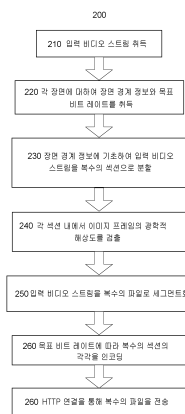
전체 청구항 수 : 총 24 항

(54) 발명의 명칭 **장면에 기초한 적응적 비트 레이트 제어**

(57) 요약

비디오 스트림을 인코딩하는 인코더가 본 명세서에 설명된다. 인코더는 입력 비디오 스트림과, 장면 전환이 발생하는 입력 비디오 스트림에서의 위치를 나타내는 장면 경계 정보와, 각 장면에 대한 목표 비트 레이트를 취득한다. 인코더는 장면 경계 정보에 기초하여 입력 비디오 스트림을 복수의 섹션으로 분할한다. 각 섹션은 시간적으로 인접한 복수의 이미지 프레임들을 포함한다. 인코더는 목표 비트 레이트에 따라 복수의 섹션의 각각을 인코딩하여, 장면에 기초하여 적응적 비트 레이트 제어를 제공한다. 비디오 품질 바가 더 낮은 비트 레이트에서 만족되면, 품질 바가 이미 만족되었기 때문에, 동일한 섹션을 더 높은 비트 레이트로 인코딩할 필요가 없다.

대표도 - 도2



특허청구의 범위

청구항 1

장면 타입을 이용하여 비디오 스트림을 인코딩하는 방법에 있어서,

입력 비디오 스트림을 취득하는 단계;

장면 전환이 발생하는 상기 입력 비디오 스트림에서의 위치를 나타내는 장면 경계 정보와 각 장면에 대한 목표 비트 레이트를 취득하는 단계;

상기 장면 경계 정보에 기초하여 상기 입력 비디오 스트림을 각각이 시간적으로 인접한 복수의 이미지 프레임에 포함하는 복수의 섹션으로 분할하는 단계; 및

상기 복수의 섹션의 각각을 상기 목표 비트 레이트에 따라 인코딩하는 단계

를 포함하는,

비디오 스트림 인코딩 방법.

청구항 2

제1항에 있어서,

각 장면에 대한 최대 컨테이너 크기를 취득하는 단계를 더 포함하는,

비디오 스트림 인코딩 방법.

청구항 3

제2항에 있어서,

상기 인코딩하는 단계는, 상기 목표 비트 레이트와 상기 최대 컨테이너 크기에 따라 상기 복수의 섹션의 각각을 인코딩하는 단계를 포함하는,

비디오 스트림 인코딩 방법.

청구항 4

제1항에 있어서,

상기 입력 비디오 스트림을 복수의 파일로 세그먼트화하는 단계를 더 포함하고, 각각의 파일은 하나 이상의 섹션을 포함하는,

비디오 스트림 인코딩 방법.

청구항 5

제1항에 있어서,

상기 입력 비디오 스트림을 데이터베이스 및 단일 비디오 파일로 세그먼트화하는 단계를 더 포함하고,

각각의 파일은 섹션을 포함하지 않거나 또는 하나 이상의 섹션을 포함하는,

비디오 스트림 인코딩 방법.

청구항 6

제1항에 있어서,
HTTP 연결을 통해 상기 복수의 파일을 전송하는 단계를 더 포함하는,
비디오 스트림 인코딩 방법.

청구항 7

제1항에 있어서,
각 섹션 내에서 상기 이미지 프레임의 최적 광학적 해상도를 검출하는 단계를 더 포함하는,
비디오 스트림 인코딩 방법.

청구항 8

제1항에 있어서,
상기 장면 타입 중 하나 이상은 상기 섹션 내에서의 상기 이미지 프레임의 광학적 해상도에 기초하여 결정되는,
비디오 스트림 인코딩 방법.

청구항 9

제1항에 있어서,
상기 섹션의 목표 비트 레이트의 하나 이상은 상기 섹션 내에서의 상기 이미지 프레임의 광학적 해상도에 기초하여 결정되는,
비디오 스트림 인코딩 방법.

청구항 10

제1항에 있어서,
상기 섹션의 비디오 이미지 크기의 하나 이상은 상기 섹션 내에서의 상기 이미지 프레임의 가장 가까운 광학적 해상도에 기초하여 결정되는,
비디오 스트림 인코딩 방법.

청구항 11

제1항에 있어서,
상기 인코딩하는 단계는, H.264/MPEG-4 AVC 표준에 기초하여 상기 복수의 섹션의 각각을 상기 목표 비트 레이트에 따라 인코딩하는 단계를 포함하는,
비디오 스트림 인코딩 방법.

청구항 12

제1항에 있어서,

주어진 장면 타입은,
 빠른 모션 장면 타입;
 정지 장면 타입;
 토크 헤드(talking head);
 문자;
 대부분의 블랙 이미지;
 짧은 장면;
 낮은 관심도 장면 타입;
 화재 장면 타입;
 물 장면 타입;
 연기 장면 타입;
 크레딧 장면 타입;
 블러(blur) 장면 타입;
 아웃 포커스 장면 타입;
 이미지 컨테이너 크기보다 낮은 해상도를 갖는 이미지 장면 타입;
 기타; 또는
 디폴트
 중 하나 이상을 포함하는,
 비디오 스트림 인코딩 방법.

청구항 13

장면 타입을 이용하여 비디오 스트림을 인코딩하는 비디오 인코딩 장치에 있어서,
 입력 비디오 스트림을 취득하고, 장면 전환이 발생하는 상기 입력 비디오 스트림에서의 위치를 나타내는 장면
 경계 정보와 각 장면에 대한 목표 비트 레이트를 취득하는 입력 모듈;
 상기 장면 경계 정보에 기초하여 상기 입력 비디오 스트림을 각각이 시간적으로 인접한 복수의 이미지 프레임을
 포함하는 복수의 섹션으로 분할하는 비디오 처리 모듈; 및
 상기 복수의 섹션의 각각을 상기 목표 비트 레이트에 따라 인코딩하는 비디오 인코딩 모듈
 을 포함하는,
 비디오 인코딩 장치.

청구항 14

제1항에 있어서,
 상기 입력 모듈은 각 장면에 대한 광학적 이미지 크기를 더 취득하는,
 비디오 인코딩 장치.

청구항 15

제14항에 있어서,
상기 비디오 인코딩 모듈은 상기 광학적 이미지 크기에 따라 상기 복수의 섹션의 각각을 더 인코딩하는,
비디오 인코딩 장치.

청구항 16

제13항에 있어서,
상기 비디오 처리 모듈은 상기 입력 비디오 스트림을 복수의 파일로 분할하고, 각각의 파일은 하나 이상의 섹션을 포함하는,
비디오 인코딩 장치.

청구항 17

제13항에 있어서,
상기 비디오 스트림은 각 세그먼트의 위치, 시작 프레임, 시간 스탬프 및 해상도를 포함하는 파일이 동반되는
단일 파일로서 인코딩되는,
비디오 인코딩 장치.

청구항 18

제13항에 있어서,
HTTP 연결을 통해 상기 복수의 파일을 전송하는 비디오 전송 모듈을 더 포함하는,
비디오 인코딩 장치.

청구항 19

제13항에 있어서,
상기 비디오 처리 모듈은, 각 섹션 내에서 상기 이미지 프레임의 광학적 해상도를 더 검출하는,
비디오 인코딩 장치.

청구항 20

제13항에 있어서,
상기 장면 타입 중 하나 이상은 상기 섹션 내에서의 상기 이미지 프레임의 광학적 해상도에 기초하여 결정되는,
비디오 인코딩 장치.

청구항 21

제13항에 있어서,
상기 섹션의 목표 비트 레이트의 하나 이상은 상기 섹션 내에서의 상기 이미지 프레임의 광학적 해상도에 기초하여 결정되는,

비디오 인코딩 장치.

청구항 22

제13항에 있어서,

상기 섹션의 비디오 품질 바(quality bar)의 하나 이상은 상기 섹션 내에서의 상기 이미지 프레임의 광학적 해상도에 기초하여 결정되는,

비디오 인코딩 장치.

청구항 23

제13항에 있어서,

상기 비디오 인코딩 모듈은, H.264/MPEG-4 AVC 표준에 기초하여 상기 복수의 섹션의 각각을 상기 목표 비트 레이트에 따라 인코딩하는,

비디오 인코딩 장치.

청구항 24

제13항에 있어서,

상기 비디오 처리 모듈에 의해 할당되는 주어진 장면 타입은,

빠른 모션 장면 타입;

정지 장면 타입;

토크 헤드(talking head);

문자;

대부분의 블랙 이미지;

짧은 장면;

낮은 관심도 장면 타입;

화재 장면 타입;

물 장면 타입;

연기 장면 타입;

스크롤 크레딧 장면 타입;

블러(blur) 장면 타입;

아웃 포커스 장면 타입;

이미지 컨테이너 크기보다 낮은 해상도를 갖는 이미지 장면 타입;

기타; 또는

디폴트

중 하나 이상을 포함하는,

비디오 인코딩 장치.

명세서

기술분야

[0001] [관련 출원에 대한 교차 참조]

[0002] 본 출원은 전문이 본 명세서에 참조로서 명시적으로 편입되는 2011년 1월 28일 출원된 미국 가특허 출원 No. 61/437,193 및 2011년 1월 28일 출원된 미국 가특허 출원 No. 61/437,223에 대한 우선권을 주장한다.

[0003] [기술분야]

[0004] 본 발명은 비디오 및 이미지 압축 기술에 관한 것으로, 더욱 상세하게는, 장면에 기초하여 적응적 비트 레이트 제어를 이용하는 비디오 및 이미지 압축 기술에 관한 것이다.

배경기술

[0005] 비디오 스트리밍이 일상적인 사용자들 사이에서 인기 및 사용이 계속 상승하고 있지만, 극복되어야 할 여러 가지 내재하는 한계가 있다. 예를 들어, 사용자들은 종종 비디오 스트림을 획득하기 위하여 제한적인 대역폭만을 갖는 인터넷을 통해 비디오를 시청하기 원한다. 예를 들어, 사용자들은 이동 전화 연결 또는 가정용 무선 연결을 통해 비디오 스트림을 획득하기 원할 수 있다. 일부 시나리오에서, 사용자들은 콘텐츠를 스포olling함으로써 충분한 대역폭의 부족을 보상한다(즉, 최종적인 시청을 위하여 로컬 스토리지에 콘텐츠를 다운로드한다). 이 방법은 여러 가지 단점으로 가득 차 있다. 먼저, 사용자는 실제의 "런타임(run-time)" 경험을 가질 수 없다 - 즉, 사용자는 프로그램을 시청할지를 결정할 때 이를 볼 수 없다. 대신에, 사용자는 프로그램을 시청하기 전에 콘텐츠가 스포olling되기 위하여 상당한 지연을 겪어야만 한다. 다른 단점은 스토리지의 유용성이다 - 제공자(provider) 또는 사용자는, 단기간이라 하더라도, 스포olling된 콘텐츠가 저장될 수 있다는 것을 보장하도록 스토리지 리소스에 대한 책임을 져야 하여, 고가의 스토리지 리소스의 불필요한 활용을 야기한다.

[0006] 비디오 스트림(통상적으로 이미지 부분과 오디오 부분을 포함한다)은, 특히 고해상도(예를 들어, HD 비디오)에서, 상당한 대역폭을 필요로 할 수 있다. 오디오는 통상적으로 훨씬 더 적은 대역폭을 필요로 하지만, 종종 고려될 필요가 여전히 있다. 하나의 스트리밍 비디오 접근 방식은 비디오 스트림을 아주 많이 압축하여 사용자가 런타임으로 또는 실질적으로 즉각적으로(즉, 실질적인 스포olling 지연을 겪지 않으면서) 콘텐츠를 시청할 수 있게 하도록 빠른 비디오 전달을 가능하게 하는 것이다. 통상적으로, 비가역 압축(lossy compression)(즉, 완전히 가역적이지 않은 압축)은 무손실 압축보다 더 많은 압축을 제공하지만, 너무 많은 비가역 압축은 바람직하지 않은 사용자 경험을 제공한다.

[0007] 디지털 비디오 신호를 전송하는데 요구되는 대역폭을 감소시키기 위하여, (비디오 데이터 압축의 목적으로) 디지털 비디오 신호의 데이터 레이트(data rate)가 실질적으로 감소될 수 있는 효율적인 디지털 비디오 인코딩을 사용하는 것이 잘 알려져 있다. 정보 처리 상호 운용(interoperability)을 보장하기 위하여, 비디오 인코딩 표준은 많은 전문적인 애플리케이션 및 소비자 애플리케이션에서 디지털 비디오의 채용을 용이하게 하는데 주요한 역할을 하여 왔다. 대부분의 영향력 있는 표준은 일반적으로 ITU-T(International Telecommunications Union) 또는 ISO/IEC(International Organization for Standardization/International Electrotechnical Committee)의 MPEG(Motion Pictures Experts Group) 15 위원회에 의해 개발되어 왔다. 권장 사항으로서 알려진 ITU-T 표준은 통상적으로 실시간 통신(예를 들어, 비디오 컨퍼런싱)에 목표를 두고 있는 반면, 대부분의 MPEG 표준은 저장(예를 들어, DVD(Digital Versatile Disc)) 및 방송(예를 들어, OVB(Digital Video Broadcast) 표준)에 대하여 최적화된다.

[0008] 현재, 대다수의 표준화된 비디오 인코딩 알고리즘은 하이브리드 비디오 인코딩에 기초한다. 하이브리드 비디오 인코딩 방법은 통상적으로 원하는 압축 이득을 획득하도록 여러 가지 상이한 무손실 압축 스킴 및 비가역 압축

스킴을 결합한다. 또한, 하이브리드 비디오 인코딩은 ITV-T 표준(H.261, H.263과 같은 H.26x 표준)과 ISO/IEC 표준(MPEG-1, MPEG-2 및 MPEG-4와 같은 MPEG-X 표준)의 근거이다. 최근의 진보된 대부분의 비디오 인코딩 표준은 현재 ITV-T 그룹 및 ISO/IEC MPEG 그룹의 합동 팀인 JVT(joint video team)에 의한 표준화 노력의 결과인 H.264/MPEG-4 AVC(advanced video coding)으로 표시되는 표준이다.

[0009] H.264 표준은 MPEG-2와 같은 확립된 표준으로부터 알려진 블록 기반 모션 보상 하이브리드 변환 코딩과 동일한 원리를 채용한다. 따라서, H.264 구문(syntax)은 화면(picture)-블록 헤더, 부분(slice)-블록 헤더 및 매크로-블록 헤더와 같은 보통의 헤더 계층과 모션-벡터, 블록-변환 계수, 양자화기 스케일(quantizer scale) 등과 같은 데이터로서 조직화된다. 그러나, H.264 표준은 비디오 데이터의 콘텐츠를 나타내는 VCL(Video Coding Layer)과 데이터의 포맷을 지정하고 헤더 정보를 제공하는 NAL(Network Adaptation Layer)을 분리한다.

[0010] 또한, H.264는 인코딩 파라미터에 대한 훨씬 증가된 선택을 허용한다. 예를 들어, 이는 16x16 매크로 블록의 더욱 정교한 분할 및 조작을 허용하여, 이에 의해 예를 들어, 모션 압축 과정이 4x4 크기와 같이 작은 매크로 블록의 세그먼트화에 수행될 수 있다. 또한, 샘플 블록의 모션 보상 예측을 위한 선택 과정이, 단지 인접한 화면 대신에, 저장된 이전에 디코딩된 다수의 화면을 포함할 수 있다. 단일 프레임 내에서의 인트라 코딩(intra coding)으로도, 동일한 프레임으로부터의 이전에 디코딩된 샘플을 이용하여 블록 예측을 형성하는 것이 가능하다. 또한, 모션 보상에 이어지는 결과에 따른 예측 오차는, 전통적인 8x8 크기 대신에, 4x4 블록 크기에 기초하여 변환되고 양자화될 수 있다. 또한, 인루프(in-loop) 디블로킹(deblocking) 필터가 이제 필수적이다.

[0011] H.264 표준은, 가능한 코딩 결정(coding decision) 및 파라미터의 개수를 확장하면서 비디오 데이터의 동일한 전체 구성을 이용한다는 점에서, H.262/MPEG-2 비디오 인코딩 구문의 확대집합(superset)으로 고려될 수 있다. 다양한 코딩 결정을 갖는 결과는 비트 레이트와 화질 사이의 양호한 트레이드 오프가 획득될 수 있다는 것이다. 그러나, H.264 표준이 블록 기반의 코딩의 전형적인 아티팩트를 상당히 감소시킬 수 있다는 것이 일반적으로 인정되지만, 다른 아티팩트를 두드러지게 할 수도 있다. H.264가 다양한 코딩 파라미터에 대한 증가된 개수의 가능한 값을 허용하고 이에 따라 인코딩 과정을 개선하기 위한 잠재성을 증가시킨다는 사실은, 또한 비디오 인코딩 파라미터의 선택에 대한 민감성을 증가시키는 결과를 초래한다.

[0012] 다른 표준과 유사하게, H.264 표준은 비디오 인코딩 파라미터를 선택하기 위한 규범적인 절차를 특정하지 않지만, 기준 구현(reference implementation)을 통해 코딩 효율, 비디오 품질 및 구현의 실현 가능성 사이의 적합한 트레이드 오프를 획득하는 것과 같이 비디오 인코딩 파라미터를 선택하는데 사용될 수 있는 다수의 기준을 기술한다. 그러나, 기술된 기준은 모든 종류의 콘텐츠 및 애플리케이션에 적합한 코딩 파라미터의 최적의 또는 적합한 선택을 항상 제공할 수 없다. 예를 들어, 기준은 비디오 신호의 특성에 최적이거나 바람직한 비디오 인코딩 파라미터의 선택을 제공하지 않거나 또는 기준은 현재의 애플리케이션에 적합하지 않은 인코딩된 신호의 특성을 획득하는데 기초할 수 있다.

[0013] CBR(constant bit rate) 인코딩 또는 VBR(variable bit rate) 인코딩을 이용하여 비디오 데이터를 인코딩하는 것이 알려져 있다. 양 경우에, 단위 시간당 비트의 수는 캐핑된다. 즉, 비트 레이트는 일부 임계값을 초과할 수 없다. 종종, 비트 레이트는 초당 비트로 표현된다. CBR 인코딩은 종종 단지 일정한 비트 레이트까지 추가로 더해지는(예를 들어, 비트 스트림을 0으로 채워) VBR 인코딩의 한 종류이다.

[0014] 인터넷과 같은 TCP/IP 네트워크는 "비트 스트림(bit stream)" 파이프가 아니지만, 임의의 시간에 전송 용량이 가변하는 최고 효율의 네트워크이다. CBR 또는 VBR 방식을 이용하여 비디오를 인코딩 및 전송하는 것은 최고 효율 네트워크에서는 이상적이지 않다. 일부 프로토콜이 인터넷을 통해 비디오를 전달하도록 설계되어 왔다. 좋은 예는 HTTP 적응적 비트 레이트 비디오 스트리밍이며, 비디오 스트림은 HTTP 연결을 통해 파일로서 전달되는 파일들로 세그먼트화된다. 이러한 파일들의 각각은 미리 결정된 재생 시간을 갖는 비디오 시퀀스이고; 그리고 비트 레이트가 가변될 수 있으며, 파일 크기가 가변될 수 있다. 따라서, 일부 파일은 다른 것보다 더 짧은

수 있다.

[0015] 따라서, 비디오 인코딩을 위한 개선된 시스템이 유익할 것이다.

[0016] 관련 기술에 대한 전문적인 예와 그와 관련된 한정은 예시적이고 비독점적인 것으로 의도된다. 관련 기술의 다른 한정은 본 명세서를 읽고 도면을 연구함에 따라 명백하게 될 것이다.

발명의 내용

[0017] 비디오 스트림을 인코딩하기 위한 인코더가 본 명세서에 설명된다. 인코더는 입력 비디오 스트림과, 장면 전환이 발생하는 입력 비디오 스트림에서의 위치를 나타내는 장면 경계 정보와, 각 장면에 대한 목표 비트 레이트를 취득한다. 인코더는 장면 경계 정보에 기초하여 입력 비디오 스트림을 복수의 섹션으로 분할한다. 각 섹션은 시간적으로 인접한 복수의 이미지 프레임들을 포함한다. 인코더는 목표 비트 레이트에 따라 복수의 장면의 각각을 인코딩하여, 장면에 기초한 적응적 비트 레이트 제어를 제공한다.

[0018] 본 발명의 내용은 아래의 발명을 실시하기 위한 구체적인 내용에서 더 설명되는 개념 중에서 선택된 것을 간단한 형태로 소개하기 위하여 제공된다. 본 발명의 내용은 청구된 대상의 주요 특징 또는 필수적인 특징을 식별하려고 의도되지 않으며, 청구된 대상의 범위를 제한하는데 사용되려고 의도되지 않는다.

도면의 간단한 설명

[0019] 본 발명의 하나 이상의 실시예들은 예로서 예시되며 유사한 도면 부호가 유사한 구성 요소를 나타내는 첨부된 도면에 의해 한정되지 않는다.

도 1은 인코더의 일례를 도시한다.

도 2는 입력 비디오 스트림을 인코딩하는 표본 방법의 단계들을 예시한다.

도 3은 본 명세서에 설명된 소정의 기술을 구현하는 인코더를 구현하는데 사용될 수 있는 처리 시스템의 블록도이다.

발명을 실시하기 위한 구체적인 내용

[0020] 본 발명의 다양한 양태가 설명될 것이다. 다음의 설명은 이러한 예들의 설명을 완전히 이해하고 가능하게 하기 위한 구체적인 상세를 제공한다. 그러나, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 이러한 상세의 많은 부분 없이 본 발명이 실시될 수 있다는 것을 이해할 것이다. 또한, 널리 알려진 일부 구조 또는 기능은 관련 설명을 불필요하게 흐리게 하는 것을 방지하기 위하여 상세히 도시되거나 설명되지 않을 수 있다. 도면이 기능적으로 분리된 컴포넌트를 도시하더라도, 이러한 도시는 단지 예시적인 목적을 위한 것이다. 이 도면에서 묘사된 컴포넌트들이 임의로 결합되거나 개별 컴포넌트로 분할될 수 있다는 것은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 명백할 것이다.

[0021] 아래에 제공된 설명에서 사용된 용어는, 본 발명의 소정의 특정 예에 대한 상세한 설명과 함께 사용되고 있더라도, 최광의의 타당한 방식으로 해석되도록 의도된다. 어떤 용어는 아래에서 강조될 수 있다; 그러나, 임의의 제한된 방식으로 해석되도록 의도되는 임의의 용어는 본 발명을 실시하기 위한 구체적인 내용 항목에서 명시적이고 구체적으로 그와 같이 정의될 것이다.

[0022] 본 명세서에서 "일 실시예", "하나의 실시예" 등에 대한 참조는, 설명되는 특정 특징, 구조 또는 특성이 본 발명의 적어도 하나의 실시예에 포함된다는 것을 의미한다. 본 명세서에서 이러한 문구의 사용은 반드시 모두 동

일한 실시예를 말하는 것은 아니다.

- [0023] 도 1은 본 발명의 일 실시예에 따른 인코더(100)의 일례를 도시한다. 인코더(100)는 입력 비디오 스트림(110)을 취득하여 입력 비디오 스트림(110)의 인스턴스를, 적어도 근사적으로, 복구하기 위해 디코더에서 디코딩될 수 있는 인코딩된 비디오 스트림(120)을 출력한다. 인코더(100)는 입력 모듈(102), 비디오 처리 모듈(104) 및 비디오 인코딩 모듈(106)을 포함한다. 인코더(100)는 하드웨어, 소프트웨어 또는 임의의 적합한 조합으로 구현될 수 있다. 인코더(100)는 비디오 전송 모듈, 파라미터 입력 모듈, 파라미터를 저장하기 위한 메모리 등과 같은 다른 컴포넌트를 포함할 수 있다. 인코더(100)는 본 명세서에 구체적 설명되지 않은 다른 비디오 처리 기능들을 수행할 수 있다.
- [0024] 입력 모듈(102)은 입력 비디오 스트림(110)을 취득한다. 입력 비디오 스트림(110)은 임의의 적합한 형태를 취할 수 있으며, 메모리와 같은 다양한 적합한 소스 중 어느 것으로부터 또는 심지어 생방송 공급으로부터 유래할 수 있다. 입력 모듈(102)은 장면 경계 정보 및 각 장면에 대한 목표 비트 레이트를 더 취득한다. 장면 경계 정보는 장면 전환이 발생하는 입력 비디오 스트림에서의 위치를 나타낸다.
- [0025] 비디오 처리 모듈(104)은 입력 비디오 스트림(110)을 분석하여, 비디오 스트림(110)을 장면 경계 정보에 기초하여 복수의 장면의 각각에 대한 복수의 섹션으로 분할한다. 각 섹션은 시간적으로 인접한 복수의 이미지 프레임을 포함한다. 일 실시예에서, 비디오 처리 모듈은 입력 비디오 스트림을 복수의 파일로 더 세그먼트화한다. 각 파일은 하나 이상의 섹션을 포함한다. 다른 실시예에서, 비디오 파일의 각 섹션의 위치, 해상도, 시간 스탬프 또는 시작 프레임 번호는 파일 또는 데이터베이스로 기록된다. 비디오 인코딩 모듈은 관련된 목표 비트 레이트 또는 비트 레이트 제한을 갖는 비디오 품질을 이용하여 각 섹션을 인코딩한다. 하나의 실시예에서, 인코더는 HTTP 연결과 같은 네트워크 연결을 통해 파일을 전송하기 위한 비디오 전송 모듈을 더 포함한다.
- [0026] 일부 실시예에서, 비디오 이미지 프레임의 광학적 해상도가 검출되어 정확한 또는 최적의 장면 비디오 치수 및 장면 분할을 결정하는데 활용된다. 광학적 해상도는 하나 이상의 비디오 이미지 프레임이 상세를 연속으로 해상(解像)할 수 있는 해상도를 말한다. 캡처 광학 장치, 기록 매체 및 원 포맷(original format)의 한계 때문에, 비디오 이미지 프레임의 광학적 해상도는 비디오 이미지 프레임의 기술적 해상도보다 훨씬 더 적을 수 있다. 비디오 처리 모듈은 각 섹션 내에서 이미지 프레임의 광학적 해상도를 검출할 수 있다. 장면 타입은 섹션 내의 이미지 프레임의 광학적 해상도에 기초하여 결정될 수 있다. 더하여, 섹션의 목표 비트 레이트가 섹션 내의 이미지 프레임의 광학적 해상도에 기초하여 결정될 수 있다. 낮은 광학적 해상도를 갖는 소정의 섹션에 대하여, 높은 비트 레이트가 섹션의 충실도(fidelity)를 유지하는데 도움을 주지 않기 때문에, 목표 비트 레이트는 더 낮을 수 있다. 또한, 전자 업스케일러(up-scaler)의 일부 경우에, 더 높은 해상도의 비디오 프레임으로 피팅하도록 낮은 해상도 이미지를 변환하는 업스케일러는 원하지 않은 아티팩트를 생성할 수 있다. 이는 특히 오래된 스케일링 기술에 해당된다. 원 해상도를 복구함으로써, 현대의 비디오 프로세서가 더욱 효율적인 방식으로 이미지를 업스케일링하여 원 이미지의 일부가 아닌 원하지 않은 아티팩트의 인코딩을 회피할 수 있게 한다.
- [0027] 비디오 인코딩 모듈은 H.264/MPEG-4 AVC 표준과 같은 임의의 인코딩 표준을 이용하여 각 섹션을 인코딩할 수 있다.
- [0028] 각 섹션은, 상이한 장면에 기초하여, 상이한 비트 레이트(즉, 500Kbps, 1 Mbps, 2 Mbps)를 수송하는 상이한 레벨의 지각 품질(perceptual quality)로 인코딩될 수 있다. 일 실시예에서, 광학 또는 비디오 품질 바(bar)가 소정의 낮은 비트 레이트, 즉 500 Kbps에서 만족되면, 인코딩 과정은 더 높은 비트 레이트에 대하여 필요하지 않을 수 있어, 더 높은 비트 레이트, 즉 1 Mbps 또는 2 Mbps에서 그 장면을 인코딩하는 필요성을 방지한다. 표 1을 참조하라. 단일 파일에서 이러한 장면들을 저장하는 경우에, 단일 파일은 더 높은 비트 레이트에서 인코딩될 필요가 있는 장면만을 저장할 것이다. 그러나 일부 경우에, 모든 장면에 대하여 높은 비트 레이트 파일(즉,

1 Mbps)에서 저장하는 것이 필요할 수 있으며(종래의 일부 오래된 적응적 비트 레이트 시스템에서), 특히 이 경우에, 저장될 섹션 또는 세그먼트는 높은 비트 레이트의 섹션 또는 세그먼트 대신에 낮은 비트 레이트, 즉 500 Kbps의 섹션 또는 세그먼트일 것이다. 따라서, 저장 공간이 절약된다(그러나, 장면을 저장하지 않는 것만큼 상당하지는 않다). 표 2를 참조하라. 단일 비디오 파일에서 다중 해상도를 지원하지 않는 시스템에 대한 것과 같은 다른 경우에, 섹션의 저장은 결정된 프레임 크기를 갖는 파일에서 발생할 수 있다. 각 해상도에서 파일의 개수를 최소화하기 위하여, 일부 시스템은 SDTV, HD720p, HD1080p와 같은 프레임 크기의 수를 제한할 것이다. 표 3을 참조하라.

표 1

장면 #	프레임 엔드 #	장면 타입	섹션 또는 인덱스	비트 레이트 (kbps)
1	29	블랙 스크린	1	단일 파일에 파일 또는 섹션이 없음
2	673	디폴트	2	1,000
3	1369	빠른 모션	3	1,000
4	1373	낮은 관심도	4	단일 파일에 파일 또는 섹션이 없음
5	1386	화재/물/연기	5	1,000
6	1411	디폴트	6	단일 파일에 파일 또는 섹션이 없음
7	1419	디폴트	7	단일 파일에 파일 또는 섹션이 없음
8	1445	빠른 모션	8	1,000
9	1455	블랙 스크린	9	단일 파일에 파일 또는 섹션이 없음
10	1469	크레딧	10	단일 파일에 파일 또는 섹션이 없음

표 2

장면 #	프레임 엔드 #	장면 타입	섹션 또는 인덱스	비트 레이트 (kbps)
1	29	블랙 스크린	1	5
2	673	디폴트	2	1,000
3	1369	빠른 모션	3	1,000
4	1373	낮은 관심도	4	600
5	1386	화재/물/연기	5	1,000
6	1411	디폴트	6	700
7	1419	디폴트	7	534
8	1445	빠른 모션	8	1,000
9	1455	블랙 스크린	9	5
10	1469	크레딧	10	120

표 3

장면 #	프레임 엔드 #	장면 타입	섹션 또는 인덱스	그룹의 이미지 크기 폭 x 높이
1	29	블랙 스크린	1	320 x 240
2	673	디폴트	2	720 x 480
3	1369	빠른 모션	3	320 x 480

4	1373	낮은 관심도	4	1280 x 720
5	1386	화재/물/연기	5	720 x 480
6	1411	디폴트	6	720 x 480
7	1419	디폴트	7	720 x 480
8	1445	빠른 모션	8	320 x 480
9	1455	블랙 스크린	9	320 x 480
10	1469	크레딧	10	720 x 480

[0032] 각 섹션은, 상이한 장면에 기초하여, 상이한 레벨의 지각 품질 및 상이한 비트 레이트로 인코딩될 수 있다. 일 실시예에서, 인코더는 입력 비디오 스트림 및 데이터베이스 또는 다른 장면 목록을 읽고, 그 다음 장면의 정보에 기초하여 비디오 스트림을 섹션들로 분할한다. 비디오에서의 장면 목록에 대한 예시적인 데이터 구조는 표 4에 도시된다. 일부 실시예에서, 데이터 구조는 컴퓨터 판독 가능한 메모리 또는 데이터베이스에 저장되어 인코더에 의해 액세스 가능할 수 있다.

표 4

[0033]

장면 #	프레임 엔드 #	장면 타입	섹션 또는 인덱스	비트 레이트 (kbps)
1	29	블랙 스크린	1	5
2	673	디폴트	2	1,000
3	1369	빠른 모션	3	1,500
4	1373	낮은 관심도	4	600
5	1386	화재/물/연기	5	1,200
6	1411	디폴트	6	700
7	1419	디폴트	7	534
8	1445	빠른 모션	8	1,300
9	1455	블랙 스크린	9	5
10	1469	크레딧	10	120

[0034] 장면의 상이한 타입은, "빠른 모션", "정지", "토크 헤드(talking head)", "문자", "대부분의 블랙 이미지", "5 프레임 이하의 짧은 장면", "블랙 스크린", "낮은 관심도", "파일", "물", "연기(smoke)", "크레딧", "블러(blur)", "아웃 포커스", "이미지 컨테이너 크기보다 낮은 해상도를 갖는 이미지" 등과 같은 장면 목록에 대하여 활용될 수 있다. 일부 경우에, 일부 장면 시퀀스는 이러한 장면에 할당된 "기타", "미지(unknown)" 또는 "디폴트" 장면 타입일 수 있다.

[0035] 도 2는 입력 비디오 스트림을 인코딩하는 방법(200)의 단계들을 도시한다. 본 방법(200)은 입력 비디오 스트림의 인스턴스를, 적어도 근사적으로, 복구하기 위해 디코더에서 디코딩될 수 있는 인코딩된 비디오 비트 스트림으로 입력 비디오 스트림을 인코딩한다. 단계 210에서, 본 방법은 인코딩된 입력 비디오 스트림을 취득한다. 단계 220에서, 본 방법은 장면 전환이 발생하는 입력 비디오 스트림에서의 위치를 나타내는 장면 경계 정보와 각 장면에 대한 목표 비트 레이트를 취득한다. 단계 230에서, 입력 비디오 스트림은 장면 경계 정보에 기초하여 복수의 섹션으로 분할되고, 각 섹션은 시간적으로 인접한 복수의 이미지 프레임에 포함한다. 그 다음, 단계 240에서, 본 방법은 각 섹션 내에서 이미지 프레임의 광학적 해상도를 검출한다. 단계 250에서, 본 발명은 입력 비디오 스트림을 복수의 파일로 세그먼트화하고, 각 파일은 하나 이상의 섹션을 포함한다. 단계 260에서, 복수의 섹션의 각각은 목표 비트 레이트에 따라 인코딩된다. 그 다음, 단계 270에서, 본 방법은 HTTP 연결을 통해 복수의 파일을 전송한다.

[0036] 입력 비디오 스트림은 일반적으로 복수의 이미지 프레임에 포함한다. 각 이미지 프레임은 일반적으로 입력 비

디오 스트림에서 별개의 "시간 위치"에 기초하여 식별될 수 있다. 실시예에서, 입력 비디오 스트림은 나누어서 또는 별개의 세그먼트로 인코더에 사용될 수 있는 스트림일 수 있다. 이러한 경우에, 인코더는 심지어 전체 입력 비디오 스트림을 취득하기 전에 롤링(rolling) 기반으로 인코딩된 비디오 비트 스트림을 스트림으로서 (예를 들어, HDTV와 같은 최종 소비자 장치로) 출력한다.

[0037] 실시예에서, 입력 비디오 스트림과 인코딩된 비디오 비트 스트림은 스트림 시퀀스로서 저장된다. 여기에서, 인코딩은 미리 수행될 수 있고, 그 다음, 인코딩된 비디오 스트림은 나중에 소비자 장치로 스트리밍될 수 있다. 여기에서, 인코딩은 소비자 장치로 스트리밍되기 전에 전체 비디오 스트림에 완전히 수행된다. 또한, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 고려될 수 있는 바와 같이, 비디오 스트림의 사전, 사후 또는 "인라인(inline)" 인코딩 혹은 그 조합의 다른 예가 본 명세서에서 소개된 기술과 함께 고려될 수 있다는 것이 이해된다.

[0038] 도 3은 인코더와 같이 임의의 전술한 기술을 구현하는데 사용될 수 있는 처리 시스템에 대한 블록도이다. 소정의 실시예에서, 도 3에 도시된 컴포넌트의 적어도 일부는 2개 이상의 물리적으로 분리되지만 연결된 컴퓨팅 플랫폼 또는 박스 사이의 분산될 수 있다는 것에 유의하여야 한다. 처리는 통상적인 서버-클래스 컴퓨터, PC, 이동 통신 장치(예를 들어, 스마트폰) 또는 임의의 다른 공지되거나 통상적인 처리/통신 장치를 나타낼 수 있다.

[0039] 도 3에 도시된 처리 시스템(301)은 하나 이상의 프로세서(310), 즉 중앙 처리 장치(CPU)와, 메모리(320)와, 이더넷 어댑터 및/또는 무선 통신 서브 시스템(예를 들어, 휴대 전화, 와이파이, 블루투스 등)과 같은 적어도 하나의 통신 장치(340)와, 하나 이상의 I/O 장치(370, 380)를 포함하며, 그 모두는 상호 연결부(390)를 통해 서로 연결된다.

[0040] 프로세서(들)(310)는 컴퓨터 시스템(301)의 동작을 제어하고, 하나 이상의 프로그래머블 범용 또는 특수 목적 마이크로프로세서, 마이크로 컨트롤러, ASIC(application specific integrated circuit), 프로그래머블 논리 장치(programmable logic device(PLD)) 또는 이러한 장치들의 조합이거나 이를 포함할 수 있다. 상호 연결부(390)는 하나 이상의 버스, 직접 연결부 및/또는 다른 종류의 물리적 연결부를 포함할 수 있으며, 본 발명이 속하는 기술 분야에서 널리 알려진 것과 같은 다양한 브릿지, 컨트롤러 및/또는 어댑터를 포함할 수 있다. 상호 연결부(390)는 하나 이상의 어댑터를 통해 PCI(Peripheral Component Interconnect) 버스, 하이퍼 트랜스포트(HyperTransport) 또는 ISA(industry standard architecture) 버스, SCSI(small computer system interface) 버스, USB(universal serial bus) 또는 IEEE(Institute of Electrical and Electronics Engineers) 표준 1394 버스(가끔 "파이어와이어(Firewire)"라 함)의 형태와 같은 하나 이상의 확장 버스에 연결될 수 있는 "시스템 버스"를 더 포함할 수 있다.

[0041] 메모리(320)는, ROM(read-only memory), RAM(random access memory), 플래시 메모리, 디스크 드라이브 등과 같은 하나 이상의 종류의 하나 이상의 메모리 장치이거나 이를 포함할 수 있다. 네트워크 어댑터(340)는 처리 시스템(301)이 통신 링크를 통해 원격 처리 시스템과 데이터를 통신할 수 있게 하기에 적합한 장치이고, 예를 들어, 종래의 전화 모뎀, 무선 모뎀, DSL(Digital Subscriber Line) 모뎀, 케이블 모뎀, 무선 트랜스ceiver, 위성 트랜스ceiver, 이더넷 어댑터 등일 수 있다. I/O 장치(370, 380)는, 예를 들어, 마우스, 트랙볼, 터치패드 등과 같은 지시 장치; 키보드; 음성 인식 인터페이스를 갖는 마이크; 오디오 스피커; 디스플레이 장치 등을 포함할 수 있다. 그러나, 적어도 일부 환경에서는 서버가 그러하듯이, 이러한 I/O 장치는 서버로서 독립적으로 동작되고 직접적인 사용자 인터페이스를 제공하지 않는 시스템에서는 불필요할 수 있다. 예시된 부품 세트에서의 다른 변동은 본 발명과 일치하는 방식으로 구현될 수 있다.

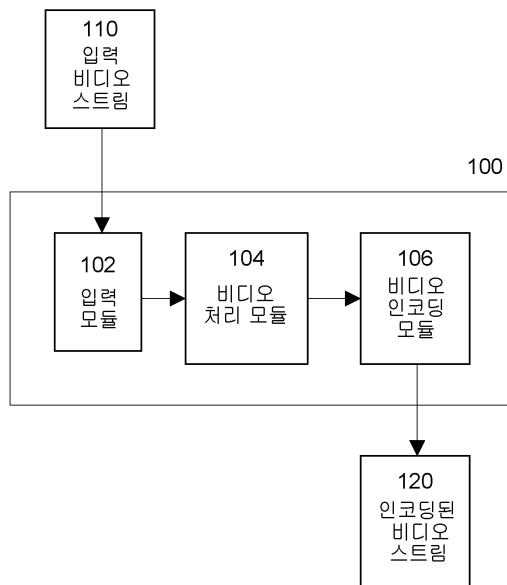
[0042] 전술한 동작을 수행하도록 프로세서(들)(310)를 프로그래밍하기 위한 소프트웨어 및/또는 펌웨어(330)는 메모리(320) 내에 저장될 수 있다. 소정의 실시예에서, 이러한 소프트웨어 또는 펌웨어는 초기에 컴퓨터 시스템(301)을 통해(예를 들어, 네트워크 어댑터(340)를 통해) 원격 시스템으로 이를 다운로드함으로써 컴퓨터 시스템

(301)에 제공할 수 있다.

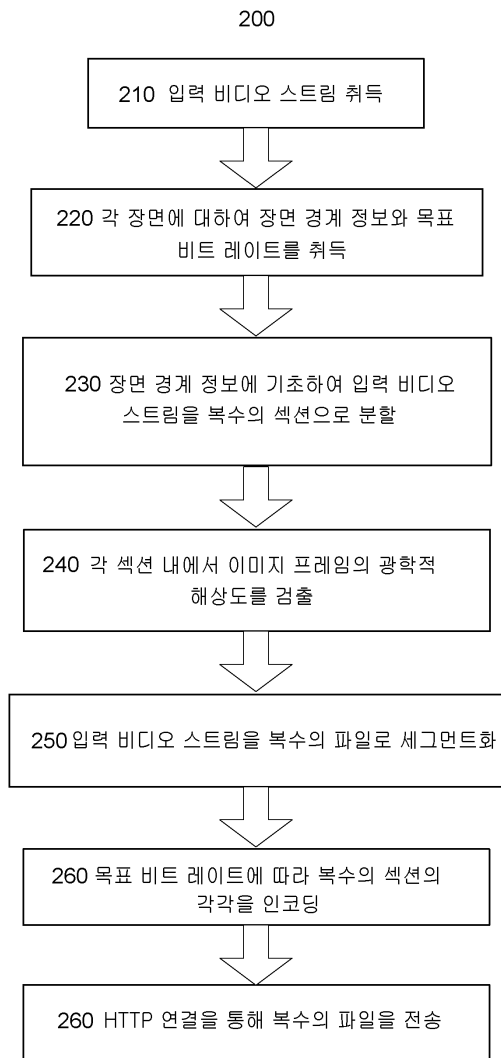
- [0043] 위에서 소개된 기술은, 예를 들어, 소프트웨어 및/또는 펌웨어로 프로그래밍된 프로그래머블 회로(예를 들어, 하나 이상의 마이크로프로세서)에 의해, 또는 특수 목적의 하드웨어 내장형(hardwired) 회로에서 전적으로, 혹은 이러한 형태의 조합으로 구현될 수 있다. 특수 목적의 하드웨어 내장형 회로는 ASIC(application-specific integrated circuit), PLD(programmable logic device), FPGA(field-programmable gate array) 등의 형태일 수 있다.
- [0044] 여기에서 소개된 기술을 구현하는데 사용하기 위한 소프트웨어 또는 펌웨어는 기계 판독 가능한 저장 매체에 저장될 수 있고, 하나 이상의 범용 또는 특수 목적 프로그래머블 마이크로프로세서에 의해 실행될 수 있다. "기계 판독 가능한 저장 매체"는, 용어가 본 명세서에서 사용되는 바와 같이, 기계(기계는, 예를 들어, 컴퓨터, 네트워크 장치, 휴대 전화, PDA(personal digital assistant), 제작 도구, 하나 이상의 프로세서를 갖는 임의의 장치 등일 수 있다)에 의해 액세스 가능한 형태로 정보를 저장할 수 있는 임의의 메커니즘을 포함한다. 예를 들어, 기계가 액세스 가능한 저장 매체는, 재기록 가능/재기록 불가능 매체(예를 들어, ROM(read-only memory); RAM(random access memory); 자기 디스크 저장 매체; 광학 저장 매체; 플래시 메모리 장치 등) 등을 포함한다.
- [0045] "논리부(logic)"라는 용어는, 본 명세서에서 사용되는 바와 같이, 예를 들어, 특정 소프트웨어 및/또는 펌웨어, 특수 목적의 하드웨어 내장형 회로 또는 그 조합으로 프로그래밍된 프로그래머블 회로를 포함할 수 있다.
- [0046] 청구된 대상의 다양한 실시예에 대한 전술한 설명이 예시 및 설명의 목적을 위해 제공되었다. 청구 대상을 개시된 특정한 형태에 한정시키거나 소진적인 것으로 의도되지 않는다. 많은 수정 및 변형이 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 자명할 것이다. 본 발명의 원리와 그의 실용적인 애플리케이션을 최선으로 설명하기 위하여 실시예들이 선택되고 설명되었으며, 이에 의해 관련 분야에서 통상의 지식을 갖는 자가 청구 대상과 다양한 실시예를 이해할 수 있게 하며, 특정 용도에 적합한 다양한 변형이 고려될 수 있게 한다.
- [0047] 여기에서 제공된 본 발명에 대한 교시 내용은 반드시 전술한 시스템일 필요가 없는 다른 시스템에 적용될 수 있다. 전술한 다양한 실시예의 요소 및 동작은 다른 실시예를 제공하기 위하여 조합될 수 있다.
- [0048] 전술한 설명은 본 발명의 소정의 실시예를 설명하고 고려되는 최선의 형태를 설명하지만, 전술한 것이 본문에서 얼마나 상세한 지에 관계없이, 본 발명은 많은 방법으로 실시될 수 있다. 실시예에 대한 상세는, 본 명세서에 개시된 발명에 의해 여전히 포함되면서, 그 상세한 구현에서 상당히 변경될 수 있다. 전술한 바와 같이, 본 발명의 소정의 특징 또는 양태를 설명할 때 사용되는 특정 용어는, 그 용어가 관련되는 발명의 임의의 특수한 특성, 특징 또는 양태에 제한되는 것으로 용어가 재정의되는 것을 의미하도록 취급되어서는 안 된다. 일반적으로, 이어지는 특허청구범위에 사용되는 용어는, 전술한 발명을 실시하기 위한 구체적인 내용 부분이 그러한 용어를 명시적으로 정의하지 않는 한, 본 발명을 명세서에 개시된 특정 실시예로 한정하도록 고려되어서는 안 된다. 따라서, 본 발명의 실제 범위는 개시된 실시예 뿐만 아니라 특허청구범위 하에서 본 발명을 실시하거나 구현하는 모든 균등한 방법을 포함한다.

도면

도면1



도면2



도면3

