



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2011년02월16일
(11) 등록번호 10-1015365
(24) 등록일자 2011년02월09일

(51) Int. Cl.
H04N 5/91 (2006.01)
(21) 출원번호 10-2004-0021965
(22) 출원일자 2004년03월31일
심사청구일자 2009년03월04일
(65) 공개번호 10-2004-0088351
(43) 공개일자 2004년10월16일
(30) 우선권주장
10/405,971 2003년04월01일 미국(US)
(56) 선행기술조사문헌
논문1: 한국데이터베이스학회
논문2: 정보과학회논문지
논문3: 대한전자공학회
JP2002027411 A

(73) 특허권자
마이크로소프트 코퍼레이션
미국 워싱턴주 (우편번호 : 98052) 레드몬드 원
마이크로소프트 웨이
(72) 발명자
장, 동
중국100080
베이징지춘로드하이디안디스트리트넘버49베이징시
그마센터3에프
왕, 위진
중국100080베이징지춘로드넘버49
베이징시그마센터5에프
장, 홍-지앙
중국100015베이징차왕디스트리트징선로드
(74) 대리인
주성민, 이중희, 백만기

전체 청구항 수 : 총 39 항

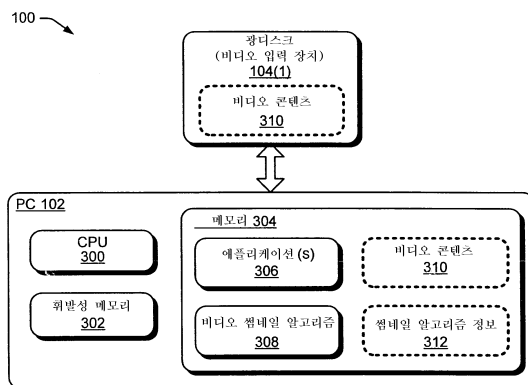
심사관 : 구대성

(54) 시각적으로 대표되는 비디오 썸네일을 생성하는 컴퓨터 구현 방법, 컴퓨팅 장치, 컴퓨터, 비디오 장치, 비디오 기록 장치 및 컴퓨터 판독가능 기록매체

(57) 요약

알고리즘은 비디오 썸네일로 사용하기 위해 비디오 시퀀스로부터 현저한 비디오 프레임을 식별한다. 비디오 썸네일의 식별은 프레임 양호도에 기초한다. 알고리즘은 프레임의 컬러 히스토그램을 계산하고, 그 컬러 히스토그램의 엔트로피와 표준편차를 산출한다. 프레임 양호도는 엔트로피와 표준편차의 가중 조합이다. 비디오 시퀀스에 관하여 가장 높은 값의 프레임 양호도를 가진 비디오 프레임이 그 비디오 시퀀스의 비디오 썸네일로 결정된다.

대표도 - 도3



특허청구의 범위

청구항 1

프로세서에 의해 실행가능한 컴퓨터-프로그램 명령어들을 포함하는 유형의(tangible) 컴퓨터 판독가능 기록매체로서, 상기 컴퓨터-프로그램 명령어들은 상기 프로세서에 의해 실행되는 경우에,

비디오 시퀀스(video sequence)를 수신하는 단계;

히스토그램 엔트로피(histogram entropy) 및 히스토그램 표준편차(histogram standard deviation)로부터 상기 비디오 시퀀스 내 프레임에 대한 프레임 양호도(frame goodness measure)를 계산하는 단계; 및

사용자가 비디오 데이터를 브라우징하도록 상기 프레임 양호도에 기초하여 시각적으로 대표되는 썸네일(thumbnail)을 나타내는 단계

를 포함하는 동작들을 수행하는, 컴퓨터 판독가능 기록매체.

청구항 2

제 1 항에 있어서,

상기 프레임 양호도에서 최대 양호도를 결정하도록 구성된 프로세서 실행가능 명령어들을 더 포함하는 컴퓨터 판독가능 기록매체.

청구항 3

제 2 항에 있어서,

상기 컴퓨터-프로그램 명령어들은,

상기 최대 양호도에 대응하는 프레임으로부터 비디오 썸네일을 생성하는 명령어들을 더 포함하는 컴퓨터 판독가능 기록매체.

청구항 4

제 1 항에 있어서,

각각의 프레임 양호도는 히스토그램 엔트로피 및 표준편차 엔트로피의 가중 조합(weighted combination)인 컴퓨터 판독가능 기록매체.

청구항 5

제 1 항에 있어서,

상기 컴퓨터-프로그램 명령어들은,

각각의 히스토그램 엔트로피를 계산하는 명령어들을 더 포함하는 컴퓨터 판독가능 기록매체.

청구항 6

제 5 항에 있어서,

상기 각각의 히스토그램 엔트로피를 계산하는 상기 컴퓨터-프로그램 명령어들은 컬러 히스토그램을 계산하는 명령어들을 포함하는 컴퓨터 판독가능 기록매체.

청구항 7

제 1 항에 있어서,

상기 컴퓨터-프로그램 명령어들은,

상기 각각의 히스토그램 표준편차를 계산하는 명령어들을 더 포함하는 컴퓨터 판독가능 기록매체.

청구항 8

제 7 항에 있어서,

상기 각각의 히스토그램 표준편차를 계산하는 상기 컴퓨터-프로그램 명령어들은 컬러 히스토그램을 계산하는 명령어들을 더 포함하는 컴퓨터 판독가능 기록매체.

청구항 9

제 1 항에 있어서,

상기 프레임 양호도를 계산하는 상기 컴퓨터-프로그램 명령어들은,

프레임을 복수의 블록들로 세분하는 명령어들;

각각의 블록에 대한 블록 양호도를 계산하는 명령어들;

최소 블록 양호도를 결정하는 명령어들; 및

상기 프레임에 대한 프레임 양호도로서 상기 최소 블록 양호도를 선택하는 명령어들

을 포함하는 컴퓨터 판독가능 기록매체.

청구항 10

제 1 항에 있어서,

상기 프레임 양호도를 계산하는 상기 컴퓨터-프로그램 명령어들은,

컬러 히스토그램을 계산하는 명령어들;

상기 컬러 히스토그램의 엔트로피를 계산하는 명령어들;

상기 컬러 히스토그램의 표준 편차를 계산하는 명령어들; 및

상기 엔트로피 및 상기 표준 편차의 가중 조합으로서 상기 양호도를 계산하는 명령어들

을 포함하는 컴퓨터 판독가능 기록매체.

청구항 11

프로세서에 의해 실행가능한 컴퓨터-프로그램 명령어들을 포함하는 유형의 컴퓨터 판독가능 기록매체로서, 상기 컴퓨터-프로그램 명령어들은 상기 프로세서에 의해 실행되는 경우에,

비디오 시퀀스를 시간적으로 세그먼트화(temporally segmenting)하는 단계;

히스토그램 엔트로피 및 히스토그램 표준편차를 이용하여 특정 시간 세그먼트 내의 프레임 양호도를 계산하는 단계;

상기 프레임 양호도에서, 상기 특정 시간 세그먼트에 대한 로컬 최대 양호도(local maximum goodness measure)를 결정하는 단계; 및

사용자가 비디오 데이터를 브라우징하도록 상기 로컬 최대 양호도에 기초하여 시각적으로 대표되는 썸네일을 나타내는 단계

를 포함하는 동작들을 수행하는, 컴퓨터 판독가능 기록매체.

청구항 12

제 11 항에 있어서,

상기 컴퓨터-프로그램 명령어들은,

상기 로컬 최대 양호도와 연관된 프레임으로부터 비디오 썸네일을 생성하는 명령어들을 더 포함하는, 컴퓨터 판독가능 기록매체.

청구항 13

제 11 항에 있어서,

상기 특정 시간 세그먼트 내의 매 n번째 프레임- 상기 n은 1 이상임 -에 대하여 프레임 양호도가 계산되는 컴퓨터 판독가능 기록매체.

청구항 14

제 11 항에 있어서,

각 양호도는 히스토그램 엔트로피 및 표준편차 엔트로피의 가중 조합인 컴퓨터 판독가능 기록매체.

청구항 15

제 11 항에 있어서,

상기 특정 시간 세그먼트는,
 상기 비디오 시퀀스의 시작 시간 세그먼트와,
 상기 비디오 시퀀스의 중간 시간 세그먼트와,
 상기 비디오 시퀀스의 종결 시간 세그먼트
 중에서 선택되는 컴퓨터 판독가능 기록매체.

청구항 16

제 11 항에 있어서,

상기 프레임 양호도를 계산하는 상기 컴퓨터-프로그램 명령어들은,
 프레임을 복수의 블록으로 세분하는 명령어들;
 각각의 블록에 대한 블록 양호도를 계산하는 명령어들;
 최소 블록 양호도를 결정하는 명령어들; 및
 상기 프레임에 대한 프레임 양호도로서 상기 최소 블록 양호도를 선택하는 명령어들
 을 포함하는 컴퓨터 판독가능 기록매체.

청구항 17

프로세서에 의해 실행가능한 컴퓨터-프로그램 명령어들을 포함하는 유형의 컴퓨터 판독가능 기록매체로서, 상기 컴퓨터-프로그램 명령어들은 상기 프로세서에 의해 실행되는 경우에,
 비디오 시퀀스를 복수의 비디오 샷들로 세그먼트화하는 단계;
 첫 번째 비디오 샷 내의 비디오 프레임의 컬러 히스토그램을 계산하는 단계;
 상기 컬러 히스토그램의 엔트로피를 계산하는 단계;
 상기 컬러 히스토그램의 표준편차를 계산하는 단계;
 상기 엔트로피 및 상기 표준편차의 가중 조합으로서 프레임 양호도를 계산하는 단계; 및
 사용자가 비디오 데이터를 브라우징하도록 상기 프레임 양호도에 기초하여 시각적으로 대표되는 썸네일을 나타 내는 단계
 를 포함하는 동작들을 수행하는, 컴퓨터 판독가능 기록매체.

청구항 18

제 17 항에 있어서,

상기 컴퓨터-프로그램 명령어들은,

상기 첫 번째 비디오 샷 내의 부가적 비디오 프레임들에 대한 부가적 프레임 양호도를 계산하는 명령어들; 및 프레임 양호도들을 비교함으로써 상기 첫 번째 비디오 샷에 대한 로컬 최대 프레임 양호도를 결정하는 명령어들을 더 포함하는, 컴퓨터 판독가능 기록매체.

청구항 19

제 18 항에 있어서,

상기 컴퓨터-프로그램 명령어들은,

상기 비디오 시퀀스 내의 부가적 비디오 샷들에 대한 부가적 로컬 최대 프레임 양호도를 계산하는 명령어들;

로컬 최대 프레임 양호도들을 비교함으로써 상기 비디오 시퀀스에 대한 썸네일 비디오 프레임을 결정하는 명령어들을

을 더 포함하는, 컴퓨터 판독가능 기록매체.

청구항 20

제 19 항에 있어서,

상기 결정하는 상기 컴퓨터-프로그램 명령어들은,

가장 높은 값을 갖는 로컬 최대 프레임 양호도를 식별하는 명령어들; 및

상기 썸네일 비디오 프레임으로서 상기 가장 높은 값을 갖는 로컬 최대 프레임 양호도에 대응하는 비디오 프레임을 선택하는 명령어들을

을 더 포함하는, 컴퓨터 판독가능 기록매체.

청구항 21

프로세서에 의해 실행가능한 컴퓨터-프로그램 명령어들을 포함하는 유형의 컴퓨터 판독가능 기록매체로서, 상기 컴퓨터-프로그램 명령어들은 상기 프로세서에 의해 실행되는 경우에,

비디오 시퀀스를 복수의 샷들로 세그먼트화하는 단계;

상기 복수의 샷들의 각각으로부터 프레임을 선택하는 단계;

각각의 선택된 프레임에 대해, 히스토그램 엔트로피 및 히스토그램 표준편차의 함수로서 프레임 양호도를 계산하는 단계; 및

상기 프레임 양호도의 함수로서 선택된 프레임들로부터 썸네일 프레임을 결정하는 단계

를 포함하는 동작들을 수행하고, 상기 썸네일 프레임은, 사용자가 비디오 데이터를 브라우징하는 것을 돕도록 사용자에게 프리젠테이션하기 위한 것인, 컴퓨터 판독가능 기록매체.

청구항 22

제 21 항에 있어서,

상기 결정하는 상기 컴퓨터-프로그램 명령어들은,

상기 프레임 양호도에서 최대 프레임 양호도를 식별하는 명령어들; 및

상기 최대 프레임 양호도와 연관된 프레임으로서 상기 썸네일 프레임을 결정하는 명령어들을

을 더 포함하는 컴퓨터 판독가능 기록매체.

청구항 23

제 21 항에 있어서,

상기 양호도를 계산하는 상기 컴퓨터-프로그램 명령어들은,

컬러 히스토그램을 계산하는 명령어들;
 상기 컬러 히스토그램의 엔트로피를 계산하는 명령어들;
 상기 컬러 히스토그램의 표준편차를 계산하는 명령어들; 및
 상기 엔트로피 및 상기 표준편차의 가중 조합으로서 상기 양호도를 계산하는 명령어들
 을 더 포함하는, 컴퓨터 판독가능 기록매체.

청구항 24

제 21 항에 있어서,
 상기 세그먼트화하는 상기 컴퓨터-프로그램 명령어들은,
 샷 경계 검출(shot boundary detection)을 수행하는 명령어들을 더 포함하는, 컴퓨터 판독가능 기록매체.

청구항 25

비디오 시퀀스 내의 비디오 프레임의 컬러 히스토그램을 계산하는 단계;
 상기 컬러 히스토그램의 엔트로피를 계산하는 단계;
 상기 컬러 히스토그램의 표준편차를 계산하는 단계;
 상기 엔트로피 및 상기 표준편차의 가중 조합으로서 프레임 양호도를 계산하는 단계; 및
 사용자가 비디오 데이터를 브라우징하도록, 상기 프레임 양호도에 기초하여 시각적으로 대표되는 비디오 썸네일
 을 디스플레이하는 단계
 를 포함하는, 컴퓨터 구현 방법(computer-implemented method).

청구항 26

제 25 항에 있어서,
 상기 비디오 시퀀스 내의 부가적 비디오 프레임들에 대한 부가적 프레임 양호도를 계산하는 단계; 및
 프레임 양호도들을 비교함으로써 상기 비디오 시퀀스에 대한 썸네일 비디오 프레임을 결정하는 단계
 를 더 포함하는, 컴퓨터 구현 방법.

청구항 27

제 26 항에 있어서,
 상기 결정하는 단계는,
 최대 프레임 양호도를 식별하는 단계; 및
 상기 썸네일 비디오 프레임으로서 상기 최대 프레임 양호도에 대응하는 비디오 프레임을 선택하는 단계
 를 더 포함하는, 컴퓨터 구현 방법.

청구항 28

비디오 시퀀스를 시간적으로 세그먼트화하는 단계;
 히스토그램 엔트로피 및 히스토그램 표준편차를 이용하여 특정 시간 세그먼트 내의 프레임 양호도를 계산하는
 단계;
 상기 프레임 양호도에서, 상기 특정 시간 세그먼트에 대한 로컬 최대 양호도를 결정하는 단계; 및
 사용자가 비디오 데이터를 브라우징하도록 상기 로컬 최대 양호도에 기초하여 시각적으로 대표되는 비디오 썸네
 일을 디스플레이하는 단계

를 포함하는, 컴퓨터 구현 방법.

청구항 29

제 28 항에 있어서,

상기 로컬 최대 양호도와 연관된 프레임으로부터 비디오 썸네일을 생성하는 단계를 더 포함하는, 컴퓨터 구현 방법.

청구항 30

비디오 시퀀스를 복수의 샷들로 세그먼트화하는 단계;

상기 복수의 샷들의 각각으로부터 프레임을 선택하는 단계;

각각의 선택된 프레임에 대해, 히스토그램 엔트로피 및 히스토그램 표준편차의 함수로서 프레임 양호도를 계산하는 단계;

상기 프레임 양호도의 함수로서 선택된 프레임들로부터 썸네일 프레임을 결정하는 단계; 및

사용자가 비디오 데이터를 브라우징하도록 상기 썸네일 프레임을 디스플레이하는 단계

를 포함하는, 컴퓨터 구현 방법.

청구항 31

제 30 항에 있어서,

상기 결정하는 단계는,

상기 프레임 양호도에서 최대 프레임 양호도를 식별하는 단계; 및

상기 최대 프레임 양호도와 연관된 프레임으로서 상기 썸네일 프레임을 결정하는 단계

를 더 포함하는, 컴퓨터 구현 방법.

청구항 32

컴퓨팅 장치로서,

프로세서; 및

상기 프로세서에 연결된 메모리

를 포함하고,

상기 메모리는 비디오 콘텐츠, 및 상기 프로세서에 의해 실행가능한 컴퓨터-프로그램 명령어들을 포함하고,

상기 컴퓨터-프로그램 명령어들은 상기 프로세서에 의해 실행되는 경우에, 비디오 썸네일 알고리즘을 구현하도록 구성되고,

상기 비디오 썸네일 알고리즘은, 상기 비디오 콘텐츠 내의 복수의 비디오 프레임들의 각각에 대한 프레임 양호도를 계산함으로써 시각적으로 대표하는 비디오 썸네일 프레임을 결정하도록 구성되고, 각각의 프레임 양호도는 프레임 히스토그램 엔트로피 및 프레임 히스토그램 표준편차의 가중 조합을 포함하는 컴퓨팅 장치.

청구항 33

비디오 장치로서,

프로세서; 및

상기 프로세서에 연결된 메모리

를 포함하고,

상기 메모리는 상기 프로세서에 의해 실행가능한 컴퓨터-프로그램 명령어들을 포함하고, 상기 컴퓨터-프로그램

명령어들은,

비디오 콘텐츠를 기록하는 단계; 및

비디오 썸네일을 결정하는 단계

를 포함하는 동작들을 수행하고,

상기 결정하는 단계는, 상기 비디오 콘텐츠 내의 복수의 비디오 프레임들의 각각에 대한 프레임 양호도를 계산하는 단계를 포함하고,

각각의 프레임 양호도는 프레임 히스토그램 엔트로피 및 프레임 히스토그램 표준편차의 가중 조합에 기초하는 비디오 장치.

청구항 34

제 33 항에 있어서,

비디오 썸네일 알고리즘을 구현하도록 구성된 하드웨어 회로 블록을 더 포함하는 비디오 장치.

청구항 35

비디오 시퀀스 내의 비디오 프레임의 컬러 히스토그램을 계산하는 수단;

상기 컬러 히스토그램의 엔트로피를 계산하는 수단;

상기 컬러 히스토그램의 표준편차를 계산하는 수단;

상기 엔트로피 및 상기 표준편차의 가중 조합으로서 프레임 양호도를 계산하는 수단; 및

사용자가 비디오 데이터를 브라우징하도록 상기 프레임 양호도에 기초하여 시각적으로 대표되는 비디오 썸네일을 디스플레이하는 수단

을 포함하는 비디오 기록 장치.

청구항 36

비디오 시퀀스를 수신하는 수단;

히스토그램 엔트로피 및 히스토그램 표준편차로부터 상기 비디오 시퀀스 내의 프레임에 대한 프레임 양호도를 계산하는 수단; 및

사용자가 비디오 데이터를 브라우징하도록 상기 프레임 양호도에 기초하여 시각적으로 대표되는 비디오 썸네일을 디스플레이하는 수단

을 포함하는 컴퓨터.

청구항 37

비디오 시퀀스 내의 비디오 프레임의 컬러 히스토그램을 계산하는 수단;

상기 컬러 히스토그램의 엔트로피를 계산하는 수단;

상기 컬러 히스토그램의 표준편차를 계산하는 수단;

상기 엔트로피 및 상기 표준편차의 가중 조합으로서 프레임 양호도를 계산하는 수단; 및

사용자가 비디오 데이터를 브라우징하도록 상기 프레임 양호도에 기초하여 시각적으로 대표되는 비디오 썸네일을 디스플레이하는 수단

을 포함하는 컴퓨터.

청구항 38

비디오 시퀀스를 시간적으로 세그먼트화하는 수단;

히스토그램 엔트로피 및 히스토그램 표준편차를 이용하여 특정 시간 세그먼트 내의 프레임 양호도를 계산하는

수단;

상기 프레임 양호도에서, 상기 특정 시간 세그먼트에 대한 로컬 최대 양호도를 결정하는 수단; 및

사용자가 비디오 데이터를 브라우징하도록 상기 최대 양호도에 기초하여 시각적으로 대표되는 비디오 썸네일을 디스플레이하는 수단

을 포함하는 컴퓨터.

청구항 39

비디오 시퀀스를 복수의 샷들로 세그먼트화하는 수단;

상기 복수의 샷들의 각각에서 프레임을 선택하는 수단;

각각의 선택된 프레임에 대하여 히스토그램 엔트로피 및 히스토그램 표준편차의 함수로서 프레임 양호도를 계산하는 수단;

상기 프레임 양호도의 함수로서 선택된 프레임으로부터 썸네일 프레임을 결정하는 수단; 및

사용자가 비디오 데이터를 브라우징하도록 상기 썸네일 프레임을 디스플레이하는 수단

을 포함하는 컴퓨터.

청구항 40

삭제

명세서

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

- [0016] 본 발명은 비디오 표현기술에 관한 것이며, 보다 구체적으로 비디오 시퀀스의 두드러진(salient) 콘텐츠를 표현하는 썸네일(thumbnail)의 생성에 관한 것이다.
- [0017] 멀티미디어 컴퓨팅 기술과 네트워크형 통신의 급격한 발전은 디지털 라이브러리에 저장된 디지털 멀티미디어(예컨대, 비디오, 오디오, 정지영상)의 양을 크게 증가시켰다. 그러나, 그와 같은 멀티미디어 데이터, 특히 비디오 데이터에 액세스하는 방법들은 그러한 멀티미디어 양의 증가에 보조를 맞추지 못하고 있다. 텍스트 기반형 문서에 대한 통상적 검색 시스템은 무관한 문서를 걸러내고, 관심의 대상이 되는 문서를 보다 더 조사하도록 돕는 빠른 문서 정보 개관을 위하여 문서 대용물(document surrogate)(예컨대, 키워드, 요약) 브라우징을 허용한다.
- [0018] 그러나, 비디오 고유의 특성 때문에, 통상적 문서 대용물 및 텍스트 지향형 브라우징 메카니즘은 비디오 데이터 액세스에는 유용하지 못하다. 비디오 데이터는 공간적 및 시간적 표현과 엄청난 볼륨으로 인하여 단순히 언어를 사용해서는 적절히 묘사할 수 있는 수준을 넘어서는 비디오 및 오디오 정보를 전달한다. 따라서, 비디오 '요약'(즉, 비디오 시퀀스에서 검출한 표현 정지영상들)의 사용은 콘텐츠 기반형 브라우징과 비디오 데이터에의 액세스가 가능하도록 하는 방법으로서 매우 흥미로운 것이다.
- [0019] 비디오 콘텐츠 브라우징/액세스를 위한 기존의 방법들은 비디오 요약이나 개요로서 사용하기 위하여 비디오 시퀀스로부터 샷 경계(shot boundaries)를 검출하고 핵심 프레임(key frames)을 추출하는 과정을 포함한다. 비디오 샷은 한 대의 카메라로부터 기록된 비디오 프레임의 연속 시퀀스이다. 비디오 샷들은 비디오 시퀀스의 구성 블록(building block)을 형성한다. 샷 경계 검출의 목적은 비디오 시퀀스를 핵심 프레임이 추출될 수 있는 복수 개의 비디오 샷들로 세그먼트화하는 것이다. 핵심 프레임은 각 샷의 두드러진 콘텐츠에 대한 썸네일 표현을 제공하는 비디오 프레임이다. 핵심 프레임의 사용은 비디오 색인에 필요한 데이터의 양을 줄이고, 비디오 콘텐츠를 조직화하고 브라우징하는 방법을 제공한다.
- [0020] 핵심 프레임 추출은 상당한 노력이 필요한 중요한 문제가 되어왔다. 핵심 프레임 추출에 자주 사용되는 한 가

지 용이한 기법은 각 비디오 샷의 첫 번째 프레임을 그 샷의 핵심 프레임으로 선택하는 것이다. 이러한 기법은 계산상 저비용이지만, 비디오 샷의 두드러진 시각적 콘텐츠를 효과적으로 캡처(capture)하는 데에는 일반적으로 실패한다. 핵심 프레임 추출의 또 다른 기법은 컬러 특성 및 비디오 프레임간의 움직임 등과 같은 여러 시각적 기준의 사용과 분석을 포함하고 있다. 그러한 기법들은 두드러진 시각적 콘텐츠의 캡처를 향상시킬 수는 있을 것이나, 계산상 고비용이기 쉽다. 따라서, 핵심 프레임 추출 기법들이 개선되어왔음에도 불구하고, 그러한 기법들은 상당한 계산상 비용이 든다는 단점이나, 비디오 데이터로부터 두드러진 시각적 콘텐츠를 효과적으로 캡처하지 못한다는 단점 등을 계속 지녀왔다.

[0021] 또한, 그와 같은 핵심 프레임 추출을 위한 종래의 기법들은 전체 비디오 시퀀스를 표현하고자 하는 핵심 프레임을 결정하는 것이 아니다. 오히려, 그와 같은 기법들은 비디오 시퀀스 내의 특정 비디오 샷만을 표현하는 핵심 프레임을 결정할 뿐이다.

발명이 이루고자 하는 기술적 과제

[0022] 따라서, 다양한 비디오 데이터의 콘텐츠 기반형 브라우징을 용이하게 하는 방식으로 비디오 시퀀스의 두드러진 콘텐츠를 정확하게 묘사하는 전체 비디오 시퀀스 표현 방법이 필요하게 된다.

[0023] 본 명세서에서는 비디오 시퀀스의 두드러진 콘텐츠를 시각적으로 표현하는 비디오 프레임을 식별함으로써 비디오 시퀀스에 대한 비디오 썸네일을 생성하는 시스템 및 방법에 관하여 개시한다. 비디오 썸네일 알고리즘은 비디오 프레임에 대한 히스토그램 엔트로피와 히스토그램 표준편차의 가중 조합의 함수로서 프레임 양호도(frame goodness measure)를 계산한다. 비디오 시퀀스 내의 여러 프레임의 프레임 양호도 비교에 따라 특정 프레임이 비디오 썸네일로서 선택된다.

[0024] 일실시예에서, 프레임 양호도는 비디오 시퀀스 내의 각 비디오 프레임에 대하여 계산된다. 프레임 양호도 계산은 프레임에 대한 히스토그램 계산을 포함한다. 그런 다음 그 히스토그램에 관하여 엔트로피와 표준편차가 계산된다. 그런 다음 그 엔트로피와 표준편차의 가중 조합으로서 프레임 양호도가 계산된다. 각 비디오 프레임에 대한 프레임 양호도가 서로 비교되어 어떤 프레임이 가장 두드러진 것인지를 결정한다. 가장 높은 값의 프레임 양호도를 가진 프레임이 비디오 썸네일로서 선택된다.

[0025] 또 다른 실시예에서는, 비디오 시퀀스가 시간적으로 세그먼트화되고, 비디오 시퀀스의 특정 세그먼트(예컨대, 비디오 시퀀스의 시작 세그먼트, 중간 세그먼트, 또는 종결 세그먼트) 내의 프레임들에 대해서 프레임 양호도가 계산된다. 그 특정 비디오 세그먼트에 관하여 로컬 최대 프레임 양호도가 결정된다. 그런 다음, 그 로컬 최대 프레임 양호도와 연관된 비디오 프레임이 그 비디오 시퀀스에 대한 비디오 썸네일로서 선택된다.

[0026] 또 다른 실시예에서는, 비디오 시퀀스가 비디오 샷들로 세그먼트화되고, 각 샷 내의 프레임들에 관하여 프레임 양호도가 계산된다. 로컬 최대 프레임 양호도가 각 비디오 샷마다 계산된다. 비디오 시퀀스에 대한 비디오 썸네일은 가장 높은 값의 로컬 최대 프레임 양호도와 연관된 비디오 프레임으로서 선택된다.

[0027] 도면 전체에 걸쳐 유사한 구성 요소 및 특징을 참조하기 위하여 동일한 참조 번호를 사용하였다.

발명의 구성 및 작용

[0028] 비디오 파일 또는 시퀀스 내의 비디오 프레임에서 두드러진 콘텐츠 함축 비디오 썸네일을 식별하는 시스템 및 방법에 관하여 설명한다. 법정 요건을 만족시키고자 상기 주제를 한정적으로 설명한다. 그러나, 이러한 설명 자체가 개시된 주제의 범위를 제한하려는 것은 아니다. 오히려, 이러한 설명은 청구범위 발명이, 현재 또는 장래의 기술과 더불어, 본 명세서에 개시된 것과 상이한 구성 요소 또는 비슷한 구성 요소의 조합을 포함하는 또 다른 방식으로도 구현될 수 있음을 고려하여 기술되었다.

[0029] 본 명세서에서 설명되는 시스템 및 방법은 비디오 시퀀스에서 비디오 썸네일로 사용하기 위한 두드러진 비디오 프레임을 식별하는 썸네일 식별 알고리즘을 구현한다. 비디오 썸네일의 식별은 프레임 양호도에 근거한다. 알고리즘은 프레임의 컬러 히스토그램을 계산하고, 그런 다음 그 컬러 히스토그램의 엔트로피와 표준편차를 계산한다. 프레임 양호도는 엔트로피와 표준편차의 가중 조합이다. 비디오 시퀀스에서 가장 높은 값의 양호도를 갖는 비디오 프레임이 비디오 시퀀스의 비디오 썸네일로 결정된다. 개시된 시스템 및 방법의 이점은 비디오 파일의 콘텐츠를 정확하게 표현하는 비디오 썸네일의 생성을 포함한다.

[0030] 예시적 환경

[0031] 도 1 및 도 2는 비디오 파일에 대한 비디오 썸네일을 식별하기에 적합한 예시적 환경을 도시하고 있다. 도 1의

예시적 환경(100)은 컴퓨터(102)와 하나 이상의 비디오 입력 소스(104)를 포함한다.

- [0032] 비디오 입력 소스(104)는 예컨대, 휴대용 저장 매체(104(1))(예를 들어, 자기디스크, 미디어 카드, 광디스크), 비디오 기록 장치(104(2)), 또는 인터넷, 기업용 네트워크, 또는 가정용 네트워크 등의 네트워크(104(3)) 등을 포함하는 비디오 콘텐츠를 컴퓨터(102)로 전송할 수 있는 임의의 유형의 장치 또는 통신 네트워크일 수 있다.
- [0033] 비디오 기록 장치(104(2))는 예컨대, VCR, TV, 및 컴퓨터(102)와 같은 개인용 컴퓨터 등을 통해서 추후 재생하기 위한 라이브-모션(live-motion) 비디오 및 오디오를 기록할 수 있는 다양한 디지털 기록 장치 중 하나일 수 있다. 비디오 기록 장치(104(2))는 일반적으로 i.LINK(IEEE 1394)나 파이어와이어(FireWire) 디지털 인터페이스를 사용해서 컴퓨터(102)에 직접 연결될 수 있고, 따라서 비디오 콘텐츠가 컴퓨터(102)상에서 직접 편집될 수 있다.
- [0034] 컴퓨터(102)는, 다양한 소스(104)로부터 비디오 콘텐츠를 수신하고 예컨대, 미디어 플레이어 등을 통한 재생을 위하여 그 비디오 콘텐츠를 조작할 수 있는 여러 컴퓨터 장치들로 구현될 수 있다. 이와 달리, 컴퓨터(102)는 전형적으로 전자메일, 일정조정, 작업 편성, 워드 프로세싱, 웹 브라우징 등의 통상적 컴퓨터 기능들을 수행할 수도 있다. 본 실시예에서, 컴퓨터(102)는 Microsoft[®]사의 Windows[®] 운영 체제와 같은 개방형 플랫폼 운영 체제를 실행한다. 컴퓨터(102)는 예를 들어, 데스크탑 컴퓨터, 서버 컴퓨터, 랩탑 컴퓨터, 또는 또 다른 형태의 개인용 컴퓨터(PC)로서 구현될 수도 있다. 컴퓨터(102) 구현에 대한 예시적 일예가 이하 도 13을 참조하여 보다 상세히 설명된다.
- [0035] 도 3의 실시예를 참조하여 이하에서 보다 상세히 설명되는 바와 같이, 컴퓨터(102)는, 일반적으로 비디오 콘텐츠 소스(104)로부터 비디오 시퀀스나 파일(즉, 비디오 콘텐츠)을 수신하고 비디오 시퀀스에서 그 시각적 콘텐츠가 실제적으로 전체 비디오 시퀀스를 나타내는 비디오 썸네일로 사용하기에 적합한 비디오 프레임을 식별하도록 구성된다.
- [0036] 도 2의 예시적 환경(200)은 독립형 비디오 기록 장치(202)를 포함하고 있다. 비디오 기록 장치(202)는 실제적으로 도 1의 비디오 기록 장치(104(2))와 비슷하게 구성될 수 있다. 따라서, 비디오 기록 장치(202)는 일반적으로 i.LINK(IEEE 1394)나 파이어와이어(FireWire) 디지털 인터페이스를 사용해서 컴퓨터에 직접 연결될 수 있다. 그러나, 도 2의 독립형 비디오 기록 장치(202)는, 그와 같은 장치가 일반적으로 캡처된 비디오 시퀀스에서 그 시각적 콘텐츠가 실제적으로 전체 비디오 시퀀스를 나타내는 비디오 썸네일로 사용하기에 적합한 비디오 프레임을 식별하도록 (컴퓨터(102)와 마찬가지로) 구성될 수 있음을 나타내고자 의도된 것이다.
- [0037] 예시적 실시예
- [0038] 도 3은 도 1의 환경(100)에서 구현될 수 있는 컴퓨터(102) 및 비디오 입력 소스(104)를 나타내는 블록도이다. 컴퓨터(102)는 데스크탑이나 랩탑 PC와 같은 PC(개인용 컴퓨터)로서 구현된다. 비디오 입력 소스(104)는 광디스크(104(1))로 구현된다.
- [0039] PC(102)는 프로세서(300), 휘발성 메모리(302)(즉, RAM), 그리고 비휘발성 메모리(304)(예컨대, ROM, 하드디스크, 플로피디스크, CD-ROM 등)를 포함한다. 비휘발성 메모리(304)는 일반적으로 컴퓨터/프로세서 판독가능 명령, 데이터 구조, 프로그램 모듈 및 PC(102)를 위한 기타 데이터를 저장한다. PC(102)는 일반적으로 메모리(304)에 저장되고 프로세서(300)에서 실행되는 여러 애플리케이션 프로그램(306)을 실행한다. 그러한 애플리케이션 프로그램(306)으로는 예컨대, 워드 프로세서, 스프레드시트, 브라우저, 멀티미디어 플레이어, 일러스트레이터(illustrators), CAD 툴(computer-aided design tools) 등을 구현하는 소프트웨어 프로그램을 포함할 수 있다. PC(102)의 예시적 일예가 도 3을 참조하여 이하 보다 상세히 설명될 것이다.
- [0040] 애플리케이션 프로그램(306)에 더하여, PC(102)는 비디오 썸네일 알고리즘(308)을 구현한다. 도 3에서는 프로세서(300) 상에서 실행 가능한 별도의 독립형 소프트웨어 애플리케이션으로 표시되었지만, 비디오 썸네일 알고리즘(308)은 더 상위레벨 애플리케이션(306)의 기능으로서 구현될 수도 있다. 그러한 애플리케이션(306)으로는 일반적으로 브라우저, 비디오 소프트웨어 편집 제품, VCR 제어 시스템, MPEG 리퍼 어플리케이션 등이 포함되지만 이로써 제한되는 것은 아니다.
- [0041] 일실시예에서, 비디오 썸네일 알고리즘(308)은, 비디오 입력 장치(104)(예를 들어, 광디스크(104(1)))로부터 수신한 비디오 콘텐츠(310), 예컨대 도 4에 도시된 비디오 시퀀스(400)를 분석하고, 그 비디오 시퀀스(400)에서 실질적으로 전체 비디오 시퀀스(400)의 콘텐츠를 표현하는 두드러진 비디오 썸네일으로써 하나의 비디오 프레임(402)을 식별(즉, 추출)하도록 구성된다. 비디오 썸네일을 식별하기 위한 계산을 수행하는 프로세스에 있어서,

비디오 썸네일 알고리즘(308)은 썸네일 알고리즘 정보(312)를 생성하여 이용한다. 비디오 콘텐츠(310)와 썸네일 알고리즘 정보(312)가 비휘발성 메모리(304)에 존재하는 것으로 표현되었지만, 이들은 그 전체 또는 일부분으로서 휘발성 메모리(302)에 존재할 수도 있음을 알아야 한다.

[0042] 도 4는 비디오 시퀀스(400)를 나타내는 상위레벨 표현이다. 비디오 시퀀스(400)는 전체 길이의 비디오 콘텐츠(310)를 나타내는 임의의 수의 프레임들(402(N))(1부터 N까지)을 나타낸다. 비디오 썸네일 알고리즘(308)(도 3)에 의해 전체 비디오 시퀀스(400)의 두드러진 콘텐츠를 나타내는 비디오 썸네일로 선택된 비디오 프레임을 표시하는 특정 비디오 프레임(402(N-R))이 비디오 프레임(400) 아래에 도시되어 있다.

[0043] 도 3의 실시예를 다시 참조하여, 비디오 썸네일 알고리즘(308)은 비디오 시퀀스(400)의 각 프레임에 대한 프레임 양호도를 계산함으로써 비디오 썸네일을 결정한다. 프레임 양호도는 주어진 비디오 프레임이 시각적으로 전체 비디오 시퀀스(400)의 두드러진 콘텐츠를 표현하는 정도를 나타낸다. 따라서, 프레임 양호도는 전체 비디오 시퀀스(400)를 표현하는 비디오 썸네일로서 가장 대표적이고 시각적으로 만족스러운 프레임(402)을 추출하는 방법이 된다.

[0044] 비디오 프레임(402)에 대한 프레임 양호도를 계산하기 위해, 비디오 썸네일 알고리즘(308)은 먼저 비디오 프레임의 컬러 히스토그램에 대한 두 가지 계산(즉, 엔트로피와 표준편차)을 수행한다. 컬러 히스토그램은 영상을 이루는 픽셀들의 통계적 표시를 제공하는 중요한 주지의 영상 분석 툴이다. 그 통계적 표시는 영상에 어떤 컬러의 픽셀들이 존재하는지 그리고 각 컬러마다 얼마나 많은 픽셀이 존재하는지 나타낸다. 컬러 히스토그램은 수평축의 각 엔트리는 픽셀이 가질 수 있는 가능한 컬러들 중의 하나인 막대그래프로 표현될 수 있다.

[0045] 정해진 순서는 아니지만, 비디오 프레임의 컬러 히스토그램에 대해 비디오 썸네일 알고리즘(308)이 수행하는 첫 번째 계산은 엔트로피 계산이다. 엔트로피는 임의변수의 불확실성 정도(measure of uncertainty)이다. X 를 이산 임의변수라 하고, x 를 X 의 가능한 결과 세트라 한다. 임의변수 X 의 확률 질량 함수는 $p(x)$ 이다. 따라서, X 의 엔트로피는 다음의 수학적식으로 정의된다.

수학식 1

$$p(x_j) = \frac{h(x_j)}{\sum_i (x_j)}$$

[0046]

수학식 2

$$Entropy = -\sum_i h(x_j) \log \frac{h(x_j)}{\sum_i x_i}$$

[0047]

[0048] 컬러 데이터 점(픽셀)의 확률이 균일 분포이면, 결과는 불확실하고 엔트로피는 최대값을 갖는다. 보다 구체적으로, 컬러 영상(즉, 컬러 비디오 프레임(402))에 최대 수의 컬러 값이 존재하고 그 컬러들이 균일하게 분포되어 있다면, 그 영상의 엔트로피가 최대화된다. 그 영상은 가장 다채로운 영상으로 간주된다. 또한, 보통 그러한 영상은 컬러를 덜 가진 다른 영상들보다 더 나은 시각적 효과를 갖는다.

[0049] 그러나, 영상의 컬러 데이터 점(픽셀)이 매우 불균등한 확률 질량 함수를 가질 때, 결과는 작은 결과 세트 안에 속하게 되고, 따라서 그 결과는 더 확실하고(즉, 불확실성이 낮고) 엔트로피는 낮게 된다. 따라서, 영상(즉, 비디오 프레임(402))에 최소 수의 컬러(예를 들어, 한 가지 컬러)가 존재하면, 영상의 엔트로피가 최소화된다. 오직 하나의 컬러를 갖는 컬러 영상에 있어서, 영상 히스토그램의 엔트로피는 영이다. 보통 그러한 영상은 더 많은 컬러를 지니는 다른 영상들보다 더 나쁜 시각적 효과를 갖는다. 따라서, 비디오 프레임(402)의 컬러 히스토그램 엔트로피는 프레임 양호도의 첫 부분으로 유용하다.

[0050] 또한, 정해진 순서는 아니지만, 비디오 프레임의 컬러 히스토그램에 대해 비디오 썸네일 알고리즘(308)이 수행하는 두 번째 계산은 표준편차이다. 컬러 히스토그램의 표준편차는 영상 히스토그램 내에서 컬러 퍼짐의 측정값(a measure of the spread of color)이며, 다음의 수학적식으로 정의된다.

수학식 3

$$SD = \sqrt{\frac{\sum_{i=1}^N (h(x_i) - M)^2}{N}}$$

[0051]

수학식 4

$$M = \frac{\sum_{i=1}^N h(x_i)}{N}$$

[0052]

[0053]

영상 히스토그램의 표준편차는 영상의 컬러 범위를 표시한다. 영상의 컬러 수치 범위가 넓을수록, 영상의 콘트라스트(contrast)가 더 커진다. 더 큰 콘트라스트를 갖는 영상은 더 적은 콘트라스트를 갖는 영상보다 인간의 시계(visual system)에 시각적으로 더 흥미를 유발하는 경향이 있다. 따라서, 비디오 프레임(402)의 컬러 히스토그램 표준편차는 프레임 양호도의 두 번째 부분으로 유용하다.

[0054]

일반적으로, 보다 다채로운 비디오 프레임(즉, 컬러 히스토그램 엔트로피가 높은)과 높은 콘트라스트의 비디오 프레임(즉, 컬러 히스토그램 표준편차가 높은)은 시각적으로 바람직한 프레임이다. 따라서, 프레임 양호도는 다음과 같이 비디오 프레임의 컬러 히스토그램의 엔트로피와 표준편차의 조합으로 정의될 수 있다.

수학식 5

$$G = \omega_1 Entropy + \omega_2 SD$$

[0055]

수학식 6

$$\omega_1 + \omega_2 = 1$$

[0056]

[0057]

또한, 'Entropy'와 'SD' 모두는 프레임 양호도를 계산하기 위한 수학식 5를 사용하기 전에 수학식 6을 이용하여 [0,1] 범위로 정규화 됨을 알아야 한다. 따라서, 프레임 양호도는 엔트로피와 표준편차의 가중 조합이 되므로 필요에 따라 어느 한 쪽이 강조될 수 있다.

[0058]

G값이 클수록, 비디오 프레임의 시각적 품질은 더 좋아진다. 그러므로, 도 4를 다시 참조하면, 비디오 썸네일 알고리즘(308)은 비디오 시퀀스(400)의 각 프레임에 대한 프레임 양호도를 계산하는 것에 더하여, 프레임 양호도를 서로 비교하고 어떤 것이 가장 큰 값인지 결정한다. 그런 다음 비디오 썸네일 알고리즘(308)은 산출된 가장 큰 프레임 양호도를 가지는 프레임(예를 들어, 402(N-R))을 비디오 시퀀스(400)의 콘텐츠를 시각적으로 두드러지게 표현하는 비디오 썸네일로 선택한다.

[0059]

전술한 프레임 양호도는 비디오 시퀀스에 대한 표현 비디오 썸네일로서의 비디오 프레임을 결정하기 위한 여러 방식에서 사용될 수 있다. 도 4에 관하여 전술한 것처럼, 예를 들어 비디오 썸네일 알고리즘(308)에 관한 일실 시예는, 비디오 시퀀스(400)안의 각 프레임에 대한 프레임 양호도를 계산하고 그 다음 모든 프레임 양호도를 비교하여 가장 높게 산출된 프레임 양호도를 가진 프레임의 위치를 파악함으로써 비디오 썸네일을 결정한다. 비디오 시퀀스에 대해 프레임 양호도를 적용하는 또 다른 예들이 도 5 내지 도 7과 관련하여 이하 설명된다.

[0060]

또 다른 실시예에서, 비디오 썸네일 알고리즘(308)은 개별적 비디오 프레임에 프레임 양호도를 적용하기에 앞서 비디오 시퀀스(500(도 5))를 세그먼트화한다. 도 5는 비디오 시퀀스(500)의 특정 세그먼트 내의 비디오 프레임들에 프레임 양호도를 적용하기 이전에 비디오 시퀀스를 시간적으로 세그먼트화하는 비디오 썸네일 알고리즘(308)의 일실시예를 나타내고 있다. 비디오 시퀀스(500)는 비디오 프레임의 시작 세그먼트(502), 비디오 프레임의 중간 세그먼트(504), 및 비디오 프레임의 종결 세그먼트(506)로 시간적으로 세그먼트화된다. 각 세그먼트는 대개 그 세그먼트의 시간 길이에 대응하는 많은 수의 프레임을 포함한다. 예를 들면, 초당 30프레임을 가지는 비디오 시퀀스의 10초짜리 세그먼트는 그 세그먼트 내에 300개의 프레임을 가지게 될 것이다. 도 5에 도시된 시간 세그먼트화는 예로써 도시된 것 일뿐, 이로써 제한을 두고자 한 것은 아님에 유의하여야 한다. 따라서, 비디오 시퀀스에 대한 어떠한 세그먼트화 방법도 가능하며, 세그먼트 내의 각 비디오 프레임에 대한 프레임 양호도의 적용을 용이하게 하는데 유용할 수 있다.

[0061]

프레임 양호도가 임의의 시간 세그먼트(예컨대, 시작 세그먼트(502), 중간 세그먼트(504), 종결 세그먼트(506))

내의 비디오 프레임에 적용될 수 있지만, 도 5는 비디오 썸네일 알고리즘(308)이 비디오 프레임의 시작 세트(502)로부터의 비디오 프레임에 프레임 양호도를 적용하는 경우를 나타낸다. 그런 다음, 비디오 썸네일 알고리즘(308)은 그 시작 비디오 프레임(502)의 프레임 양호도를 다른 것과 비교하고, 시작 세그먼트(502)에 대한 로컬 최대 프레임 양호도(local maximum frame goodness measure)를 결정한다. 비디오 프레임(508)은 시작 프레임 세그먼트(502) 내에서 가장 높은 프레임 양호도(즉, 로컬 최대값)를 가지는 것으로 표시되어 있다. 그러므로, 비디오 프레임(508)은 비디오 시퀀스(500)를 가장 잘 나타내는 비디오 썸네일로 선택(즉, 추출)된다. 대부분의 경우, 그 시간 세그먼트 내의 비디오 프레임에 대한 프레임 양호도를 산출하고 비교함으로써 특정 시간 세그먼트(예컨대, 참조번호(502, 504, 506))로 비디오 썸네일의 탐색을 제한하는 것이 전체 비디오 시퀀스(500)의 모든 프레임에 대한 프레임 양호도를 산출하고 비교하여 나오게 될 비디오 썸네일로부터 결과적 비디오 썸네일(508)의 시각적 품질을 크게 축소시키지는 않는다.

[0062] 또 다른 실시예에서, 비디오 썸네일 알고리즘(308)은 각 비디오 프레임에 프레임 양호도를 적용하기에 앞서 비디오 시퀀스(600(도 6))를 비디오 샷들로 세그먼트화한다. 비디오 샷은 비디오 시퀀스에서의 연속 시퀀스 또는 프레임 세그먼트이며, 대개 비디오의 구성 블록을 규정한다. 예로써, 도 6은, 각각 비디오 시퀀스(600)로부터의 많은 프레임을 각각 포함하는 표현 비디오 샷들(602, 604, 606, 608)을 도시한다. 비디오 샷(602, 604, 606, 608)은 비디오 시퀀스(600)를 구성한다. 비디오 샷은 임의의 특정 비디오 콘텐츠를 나타낼 수 있다. 대부분의 비디오 샷은 비디오 콘텐츠에서의 임계 편차(threshold deviation)에 의해 구별될 수 있는 정해진 경계를 가지고 있다. 비디오 썸네일 알고리즘(308)으로 구현될 수 있는 다양한 효과의 샷 경계 검출 기법이 있다. 그러나, 그러한 검출 기법은 당업자에게 이미 주지되어 있으며, 따라서 더 자세히 설명하지는 않을 것이다.

[0063] 도 6의 실시예에서, 비디오 썸네일 알고리즘(308)은 일반적으로 각 비디오 샷(예컨대, 참조번호(602, 604, 606, 608))의 로컬 최대 프레임 양호도를 결정하고 여러 비디오 샷의 로컬 최대 프레임 양호도를 비교해서 전체 비디오 시퀀스(600)에 대한 글로벌 최대 양호도(global maximum goodness measure)를 결정한다. 글로벌 최대 양호도는 비디오 시퀀스(600)에서 어떤 비디오 프레임이 전체 비디오 시퀀스(600)를 가장 시각적으로 대표하는지 나타내고, 따라서 비디오 썸네일 알고리즘(308)에 의해 전체 시퀀스(600)에 대한 비디오 썸네일로서 어떤 비디오 프레임이 선택(즉, 추출)되는지 나타낸다.

[0064] 도 7은 비디오 썸네일 알고리즘(308)이 비디오 프레임(700)을 블록들(702)로 분할하고 비디오 프레임(700) 내의 각 블록(702)에 대한 양호도를 계산하는 또 다른 실시예에 관하여 도시하고 있다. 비디오 프레임(700)에 대한 모든 양호도는 최소값의 블록 양호도가 되도록 결정된다. 따라서, 블록(702(1))이 프레임(700)에서 가장 높은 블록 양호도를 가지고 블록(702(2))이 가장 낮은 블록 양호도를 갖는다면, 프레임(700)에 대한 양호도는 가장 낮은 값의 블록(702(2))에서 계산된 양호도가 되도록 결정된다. 이러한 기법은 각 비디오 프레임(700)에 대한 양호도가 비디오 프레임(700) 내의 임의의 블록(702)에 있어서 적어도 그 최소 양호도 만큼은 될 것임을 보증한다. 이 기법은 전술한 각 실시예에도 적용 가능하다. 즉, 전술한 실시예에서 프레임 양호도가 결정될 때마다, 비디오 프레임 내의 복수의 블록에서 최소 블록 양호도의 선택에 기초하여 결정될 수 있다는 것이다.

[0065] 블록 양호도의 계산은 일반적으로 전술한 프레임 양호도의 계산과 동일하다. 전술한 바와 같이, 주어진 프레임의 프레임 양호도는, 먼저 프레임의 컬러 히스토그램 엔트로피와 프레임의 컬러 히스토그램 표준편차를 계산하고 엔트로피와 표준편차를 조합함으로써 산출된다. 도 7은, 비디오 썸네일 알고리즘(308)이 어떻게 주어진 프레임(700)을 복수의 블록(702)으로 분할하고 프레임의 각 블록(702)에 대한 엔트로피, 표준편차, 및 양호도 계산을 수행하는지를 도시한다. 그런 다음 비디오 썸네일 알고리즘(308)은 프레임(700)에 대한 전체 프레임 양호도로 될 최소값을 갖는 블록 양호도를 선택한다.

[0066] 도 8은 도 2의 환경(200)에서 구현될 수 있는 독립형 비디오 기록장치(202)를 도시한 블록도이다. 비디오 장치(202)는 실제적으로 도 1의 비디오 기록장치(104(2))와 동일한 방식으로 구성될 수 있다. 도 8의 블록도 표현은 도 2의 비디오 기록장치(202)가 그 비디오 썸네일의 시각적 콘텐츠가 컴퓨터(102)와 관련하여 본 명세서에서 전술한 것과 비슷한 방식으로 실제적으로 전체 비디오 시퀀스를 나타내도록(컴퓨터(102)처럼) 일반적으로 비디오 시퀀스로부터 비디오 썸네일을 식별하도록 구성될 수 있음을 나타내고자 한 것이다. 따라서, 비디오 장치(202)는 본 명세서의 시스템 및 방법이 구현될 수 있는 또 다른 예시적 환경일 뿐이다.

[0067] 도 8을 참조하면, 비디오 장치(202)는 일반적으로 프로세서(800), 휘발성 메모리(802)(즉, RAM), 및 비휘발성 메모리(804)(예컨대, ROM)를 포함한다. 비디오 장치(202)는 또한 일반적으로 기록된 비디오 콘텐츠의 전송이 가능하도록 휴대용 저장매체 능력(도시되지 않음) 및/또는 i.LINK(IEEE 1394)나 파이어와이어(FireWire) 디지털

인터페이스(도시되지 않음) 등의 디지털 인터페이스를 포함한다. 운영 체제 소프트웨어(806)가 메모리(804)에 저장되고, 비디오 캡처 소자(808)와 같은 다양한 비디오 기능 및 다양한 소자들을 제어하기 위해 프로세서(800)에서 실행될 수 있다. 비디오 캡처 소자(808)는 초점 렌즈, 전하 결합 소자(CCD), 및 아날로그/디지털 변환기 등의 전형적 비디오 소자들을 포함한다. 운영 체제 소프트웨어(806)는 일반적으로 사용자 제어 인터페이스(810)를 통해서 입력되는 사용자 입력에 응답하여 시각 정보를 수신하고, 정보를 전자 신호로 변환하고, 그것을 비디오 콘텐츠(310)로 저장하는 등의 기능을 수행한다. 뷰파인더(812)는 일반적으로 사용자가 비디오 콘텐츠(310) 및 인터페이스(810)를 통해 제어되는 기타 정보를 사용자가 볼 수 있도록 해준다. 비디오 기록장치(202)의 일반적 동작 및 기능은 이미 주지되어 있다.

[0068] 비디오 장치(202)의 비디오 썸네일 알고리즘(308)은 컴퓨터(102)와 관련하여 앞서 논의한 바와 비슷한 방식으로 기능하도록 구성된다. 따라서, 도 4 내지 도 7과 관련한 앞서의 설명은 비디오 장치(202)에 동일하게 적용 가능하다. 그러므로, 비디오 콘텐츠(310)가 비디오 장치(202)의 전형적 동작을 통하여 기록되면, 비디오 썸네일 알고리즘(308)은 기록된 비디오 시퀀스(즉, 비디오 파일)를 시각적으로 표현하는 비디오 썸네일을 결정하도록 동작한다. 계산 결과와 기타 알고리즘 정보는 썸네일 알고리즘 정보(312)로서 임시로 저장된다. 비디오 썸네일 알고리즘(308)에 의해 생성된 비디오 썸네일은 뷰파인더(812)를 통해 뷰잉(view)/브라우저될 수 있고, 사용자 제어 인터페이스(810)를 통하여 조작될 수 있다.

[0069] 예시적 방법

[0070] 이제, 비디오 시퀀스를 나타내는 두드러진 콘텐츠 함축 비디오 썸네일을 식별하기 위한 예시적 방법들이 도 9, 도 10, 도 11 및 도 12의 흐름도를 주로 참조하여 설명될 것이다. 이러한 방법들은 일반적으로 도 3 내지 도 8과 관련하여 앞서 논의된 예시적 실시예에 적용된다. 개시된 방법의 구성 요소들은 임의의 적절한 수단에 의해 수행될 수도 있는바, 예컨대 프로세서 판독가능 매체에 정의된 프로세서 판독가능 명령들의 실행에 의하여 수행될 수도 있다.

[0071] 본 명세서에서 사용되는 '프로세서 판독가능 매체'란 프로세서가 사용하거나 실행하는 명령을 포함하거나, 저장하거나, 통신하거나, 전파하거나 또는 전송할 수 있는 임의의 수단일 수 있다. 프로세서 판독가능 매체는, 제한 없이, 전기적, 자기적, 광학적, 전자기적, 적외선 또는 반도체 시스템, 장비, 장치, 또는 전파매체일 수 있다. 프로세서 판독가능 매체의 보다 구체적인 예로는 특히 하나 이상의 와이어로 이루어진 전기적 접속부, 휴대용 컴퓨터 디스켓, RAM, ROM, EPROM 또는 플래시메모리, 광섬유, 그리고 휴대용 CDROM 등이 포함된다.

[0072] 도 9는 전체 비디오 시퀀스의 두드러진 콘텐츠를 시각적으로 나타내는 비디오 썸네일을 식별하기 위한 예시적 방법(900)에 관하여 도시하고 있다. 블록(902)에서, 비디오 시퀀스의 한 비디오 프레임에 대한 컬러 히스토그램이 계산된다. 비디오 시퀀스는 다양한 비디오 소스(예컨대, 휴대용 저장 매체, 네트워크 접속부, 비디오 카메라)로부터 컴퓨터가 수신한 비디오 콘텐츠이거나 비디오 기록장치가 기록한 비디오 콘텐츠일 수 있다. 따라서, 방법(900)에 기재된 계산은 컴퓨터, 비디오 기록장치, 또는 임의의 유사 장치에서 수행될 수 있다. 블록(904)에서, 컬러 히스토그램의 엔트로피가 앞서 논의한 수학식 1 및 2에 따라 계산된다. 블록(906)에서, 컬러 히스토그램의 표준편차는 전술한 수학식 3 및 4에 의해 계산된다. 블록(908)에서, 프레임 양호도가 계산된다. 프레임에 대한 프레임 양호도는 전술한 수학식 5 및 6에 따라 엔트로피와 표준편차의 가중 조합으로서 계산된다.

[0073] 블록(910)에서, 비디오 시퀀스의 부가적 비디오 프레임에 대한 부가적 프레임 양호도가 동일한 방식으로 계산된다. 블록(912)에서, 최대 프레임 양호도가 계산된 모든 프레임 양호도 중에서 식별된다. 블록(914)에서, 최대 프레임 양호도에 대응하는 비디오 프레임이 전체 비디오 시퀀스의 두드러진 콘텐츠를 시각적으로 나타내는 비디오 썸네일로서 선택된다.

[0074] 도 10은 비디오 시퀀스 내의 비디오 프레임에 대한 양호도를 계산하는 또 다른 방법(1000)에 관하여 도시하고 있다. 블록(1002)에서, 비디오 프레임은 복수의 블록들로 세분된다. 블록의 수는 임의의 적절한 값일 수 있고, 예컨대 하나의 비디오 프레임을 구성하는 16×16 블록 그룹일 수 있다. 블록(1004)에서, 블록 그룹 내의 각 블록에 대한 양호도가 계산된다. 블록 양호도는, 그 계산이 전체 비디오 프레임이 아니라 비디오 프레임 일부에 대해서만 적용된다는 점을 제외하고는, 프레임 양호도에 관하여 앞서 설명한 것과 같은 방식(즉, 같은 수학적식을 사용함)으로 계산된다. 양호도가 각 블록마다 계산되면, 블록(1006)에 도시된 바와 같이 최소 블록 양호도가 결정된다. 블록(1008)에서, 최소 블록 양호도가 비디오 프레임의 전체 양호도로 선택된다.

[0075] 도 11은 전체 비디오 시퀀스의 두드러진 콘텐츠를 시각적으로 나타내는 비디오 썸네일을 식별하는 또 다른 예시

적 방법(1100)에 관하여 도시하고 있다. 블록(1102)에서, 비디오 시퀀스가 시간적으로 세그먼트화된다. 비디오 시퀀스는 예컨대, 시퀀스를 시작 세그먼트, 중간 세그먼트, 및 종결 세그먼트로 세그먼트화하는 등 어떠한 기초에 의하여도 시간적으로 세그먼트화될 수 있다. 블록(1104)에서, 비디오 시퀀스의 특정 세그먼트 내의 프레임에 대해 프레임 양호도가 계산된다. 양호도는 전술한 바와 같이 프레임 히스토그램 엔트로피와 히스토그램 표준편차를 기초로 하여 계산된다. 블록(1106)에서, 특정 세그먼트에 대한 로컬 최대 양호도가 결정된다. 블록(1108)에서, 비디오 썸네일은 로컬 최대 양호도와 연관된 프레임으로서 선택된다.

[0076] 도 12는 전체 비디오 시퀀스의 두드러진 콘텐츠를 시각적으로 나타내는 비디오 썸네일을 식별하는 또 다른 예시적 방법(1200)에 관하여 도시하고 있다. 블록(1202)에서, 비디오 시퀀스는 비디오 샷들로 세그먼트화된다. 블록(1204)에서, 첫 번째 비디오 샷 내의 비디오 프레임에 대한 컬러 히스토그램이 계산된다. 블록(1206)에서, 컬러 히스토그램의 엔트로피가 전술한 수학적 식 1 및 2에 따라 계산된다. 블록(1208)에서, 컬러 히스토그램의 표준편차가 전술한 수학적 식 3 및 4에 따라 계산된다. 블록(1210)에서, 프레임 양호도가 산출된다. 프레임의 프레임 양호도는 전술한 수학적 식 5 및 6에 따라 엔트로피와 표준편차의 가중 조합으로서 계산된다.

[0077] 방법(1200)의 블록(1212)에서, 첫 번째 비디오 샷 내의 부가적 비디오 프레임에 대한 부가적 프레임 양호도가 계산된다. 그런 다음, 블록(1214)에서 첫 번째 비디오 샷에 대한 로컬 최대 양호도가 결정된다. 블록(1216)에서, 첫 번째 비디오 샷에 대한 것과 동일한 방식으로 부가적 비디오 샷에 대한 부가적 로컬 최대 프레임 양호도가 계산된다. 블록(1218)에서, 가장 높은 값의 로컬 최대 프레임 양호도를 가진 비디오 프레임이 전체 비디오 시퀀스의 두드러진 콘텐츠를 시각적으로 나타내는 비디오 썸네일로 선택된다.

[0078] 하나 이상의 방법들이 흐름도 및 흐름도의 각 블록과 연관된 텍스트에 의해 개시되었지만, 각 블록은 반드시 그것들이 표시된 순서대로 수행되어야 할 필요는 없으며, 이와 다른 순서로 수행되어도 마찬가지로의 장점을 얻을 수 있다는 것을 알아야 한다. 또한, 그 방법들은 상호 배타적이지 않고 단독적으로 또는 서로 조합하여서도 수행될 수 있다.

[0079] 예시적 컴퓨터

[0080] 도 13은 컴퓨터(102) 구현에 이용될 수 있는 적절한 컴퓨팅 환경(1300)의 일례를 도시하고 있다. 이하 설명에 의하여, 컴퓨터(102)는, 비디오 썸네일 알고리즘(308)을 갖춘 경우, 도 3과 관련해서 앞서 설명된 제 1 예시적 실시예에 따라 여기 개시된 원리를 구현하는 임의의 범용 또는 특수 목적 컴퓨팅 플랫폼 클래스를 표현하고자 의도된 것임이 명확해질 것이다. 비디오 썸네일 알고리즘(308)은 도 3의 문맥에서 소프트웨어 애플리케이션으로 도시되었지만, 이와 달리 컴퓨터(102)는 비디오 썸네일 알고리즘(308)의 하드웨어 구현을 지원할 수도 있음을 알아야 한다. 이와 관련해서, 비디오 썸네일 알고리즘(308)에 관한 설명이 없다면, 컴퓨터(102)에 관한 다음의 설명은 단지 예시적인 것이며, 더 큰 또는 더 적은 기능을 갖춘 컴퓨터로 대체될 수도 있을 것이다.

[0081] 예컨대, 컴퓨터(102)는 컴퓨터, 개인용 컴퓨터, 서버 컴퓨터, 핸드헬드, 랩탑 장치, 휴대용 통신장치, 멀티프로세서 시스템, 마이크로프로세서 시스템, 마이크로프로세서 기반형 시스템, 소비자 전자장치, VCR, 비디오 미디어 플레이어, 주문형 비디오 장치, 및 기타 관련 장치 등과 같이 비디오를 보여줄 수 있는 임의의 비디오 표현 장치(video observation device)로 구현될 수 있다.

[0082] 컴퓨팅 환경(1300)은 컴퓨터(1302) 형태의 범용 컴퓨팅 시스템을 포함한다. 컴퓨터(1302) 소자는 하나 이상의 프로세서나 프로세싱 유닛(1304), 시스템 메모리(1306), 및 프로세서(1304)를 포함한 여러 시스템 소자들을 시스템 메모리(1306)에 연결하는 시스템 버스(1308)를 포함할 수 있으며, 이로써 제한되는 것은 아니다.

[0083] 시스템 버스(1308)는 메모리 버스나 메모리 제어기, 주변 버스, 가속형 그래픽 포트, 및 다양한 버스 아키텍처 중 임의의 것을 이용한 로컬 버스 또는 프로세서를 포함하는 여러 버스 구조중 하나 이상을 나타낸다. 예를 들면, 그러한 아키텍처에는 ISA(Industry Standard Architecture) 버스, MCA(Micro Channel Architecture) 버스, EISA(Enhanced ISA) 버스, VESA(Video Electronics Standards Association) 로컬 버스, 및 메자닌(Mezzanine) 버스로도 알려진 PCI(Peripheral Component Interconnects) 버스 등이 포함될 수 있다.

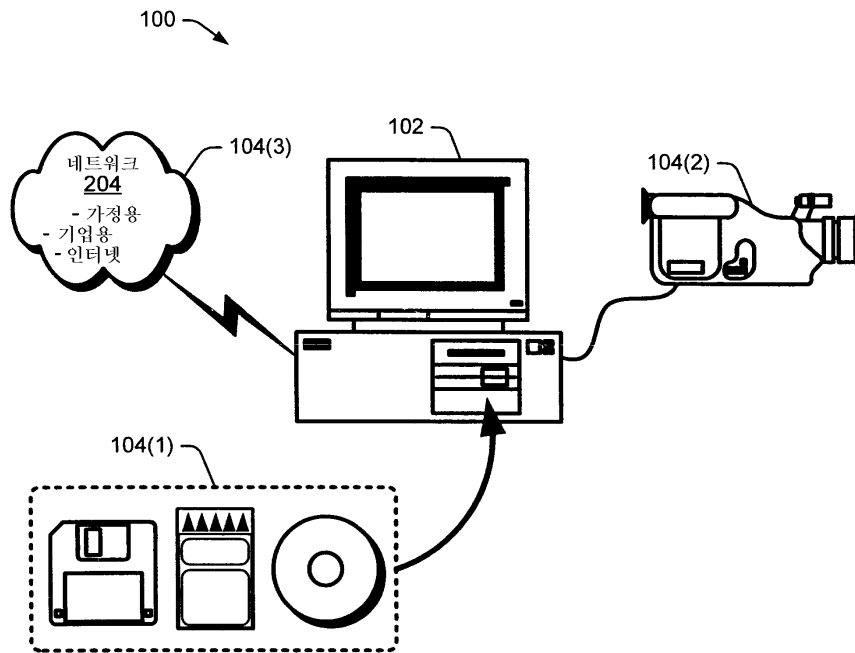
[0084] 컴퓨터(1302)는 일반적으로 다양한 컴퓨터 판독가능 매체를 포함하고 있다. 그와 같은 매체는 컴퓨터(1302)에 의하여 액세스 가능한 임의의 이용가능 매체일 수 있고, 휘발성 및 비휘발성 매체, 분리형 및 비분리형 매체 모두를 포함할 수 있다. 시스템 메모리(1306)는 RAM(1310)과 같은 휘발성 메모리 형태 및/또는 ROM(1312)과 같은 비휘발성 메모리 형태의 컴퓨터 판독가능 매체를 포함한다. 예컨대, 시동 동안 컴퓨터(1302) 내부의 구성 요소들 간에 정보 전송을 돕는 기본적 루틴을 포함하는 BIOS(basic input/output system)(1314)는 ROM(1312)에 저장되어 있다. 일반적으로, RAM(1310)은 즉시 액세스 가능한, 및/또는 프로세싱 유닛(1304)에 의해서 즉시 작동

되는 데이터 및/또는 프로그램 모듈을 저장하고 있다.

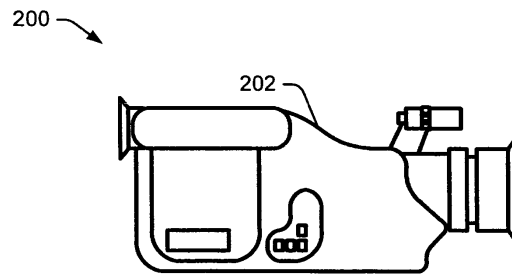
- [0085] 컴퓨터(1302)는 또한 이와 다른 이동가능/이동불가능, 휘발성/비휘발성 컴퓨터 저장 매체를 포함할 수도 있다. 예로써, 도 13은 비분리형 비휘발성 자기 매체(도시되지 않음)로부터 판독하고 그에 기록하는 하드디스크 드라이브(1316), 분리형 비휘발성 자기 디스크(1320)(예컨대, 플로피 디스크)로부터 판독하고 그에 기록하는 자기디스크 드라이브(1318), CD-ROM, DVD-ROM, 기타 광학 매체 등과 같은 분리형 비휘발성 광디스크(1324)로부터 판독하고 및/또는 그에 기록하는 광디스크 드라이브(1322)를 도시하고 있다. 하드디스크 드라이브(1316), 자기디스크 드라이브(1318), 및 광디스크 드라이브(1322)는 하나 이상의 데이터 매체 인터페이스(1326)에 의해 각각 시스템 버스(1308)로 연결된다. 이와 달리, 하드디스크 드라이브(1316), 자기디스크 드라이브(1318), 및 광디스크 드라이브(1322)는 SCSI 인터페이스(도시되지 않음)에 의하여 시스템 버스(1308)로 연결될 수 있다.
- [0086] 디스크드라이브와 그와 연관된 컴퓨터 판독가능 매체는 컴퓨터(1302)를 위한 컴퓨터 판독가능 명령, 데이터 구조, 프로그램 모듈, 및 기타 데이터의 비휘발성 저장소를 제공한다. 본 실시예는 하드디스크(1316), 분리형 자기디스크(1320), 및 분리형 광디스크(1324)를 도시하고 있지만, 자기 카세트 또는 기타 자기 저장장치, 플래시 메모리 카드, CD-ROM, DVD 또는 기타 광학 저장장치, RAM, ROM, EEPROM 등과 같이 컴퓨터에 의하여 액세스 가능한 데이터를 저장할 수 있는 또 다른 유형의 컴퓨터 판독가능 매체가 예시적 컴퓨팅 시스템 및 환경을 구현하는데 이용될 수 있음을 알아야 한다.
- [0087] 예컨대, 운영 체제(1326), 하나 이상의 애플리케이션 프로그램(1328), 기타 프로그램 모듈(1330), 및 프로그램 데이터(1332)를 포함하는 임의의 수의 프로그램 모듈이 하드디스크(1316), 자기디스크(1320), 광디스크(1324), ROM(1312), 및/또는 RAM(1310)에 저장될 수 있다. 그와 같은 운영 체제(1326), 하나 이상의 애플리케이션 프로그램(1328), 기타 프로그램 모듈(1330), 및 프로그램 데이터(1332) 각각(또는 이들의 일부의 조합)은 사용자 네트워크 액세스 정보를 위한 캐시 방식(caching scheme)의 일실시예를 포함할 수 있다.
- [0088] 컴퓨터(1302)는 통신 매체로 식별되는 다양한 컴퓨터/프로세서 판독가능 매체를 포함할 수 있다. 통신 매체는 일반적으로 반송파 또는 기타 전송 메커니즘 등의 변조 데이터 신호에 컴퓨터 판독가능 명령, 데이터 구조, 프로그램 모듈, 또는 기타 데이터를 구현하고 임의의 정보 전달 매체를 포함한다. '변조 데이터 신호'라는 용어는 신호 내의 정보를 인코딩하도록 설정되거나 변환된 특성을 하나 또는 그 이상을 갖는 신호를 의미한다. 통신 매체는 예를 들면, 유선 네트워크 또는 직접 유선 접속 등과 같은 유선 매체와, 음향, RF, 적외선, 기타 무선 매체 등의 무선 매체를 포함하되, 이로써 제한되지는 않는다. 이들로 구성된 임의의 조합 역시 컴퓨터 판독가능 매체의 범위에 포함된다.
- [0089] 사용자는 키보드(1334) 및 포인팅 장치(1336)(예컨대, '마우스') 등의 입력 장치를 통하여 컴퓨터 시스템(1302)으로 명령과 정보를 입력할 수 있다. 또 다른 입력 장치(1338)(구체적으로 도시되지 않음)에는 마이크로폰, 조이스틱, 게임패드, 위성 안테나, 직렬포트, 스캐너 등이 포함될 수 있다. 이들 및 기타 입력 장치들이 시스템 버스(1308)에 연결된 입출력 인터페이스(1340)를 통해 프로세싱 유닛(1304)에 연결되어 있지만, 병렬포트, 게임포트, 또는 범용직렬버스(USB) 등과 같은 또 다른 인터페이스 및 버스 구조에 의해 연결될 수도 있다.
- [0090] 모니터(1342) 또는 또 다른 형태의 디스플레이 장치는 또한 비디오 어댑터(1344)와 같은 인터페이스를 통해 시스템 버스(1308)에 연결될 수 있다. 모니터(1342)에 더하여, 또 다른 출력 주변 장치로는 입출력 인터페이스(1340)를 통해서 컴퓨터(1302)에 연결될 수 있는 스피커(도시되지 않음) 및 프린터(1346) 등의 소자가 포함될 수 있다.
- [0091] 컴퓨터(1302)는 원격 컴퓨팅 장치(1348)와 같은 하나 이상의 원격 컴퓨터로의 논리적 연결을 사용하는 네트워크 환경에서 작동할 수 있다. 예를 들면, 원격 컴퓨팅 장치(1348)는 개인용 컴퓨터, 휴대용 컴퓨터, 서버, 라우터, 네트워크 컴퓨터, 피어 장치, 또는 기타 공통 네트워크 노드 등이 될 수 있다. 원격 컴퓨팅 장치(1348)는 컴퓨터 시스템(1302)과 관련하여 본 명세서에서 설명된 구성 요소 및 특성 전체 또는 많은 부분을 포함할 수 있는 휴대용 컴퓨터로서 도시되었다.
- [0092] 컴퓨터(1302)와 원격 컴퓨팅 장치(1348)간의 논리적 연결은 LAN(1350) 및 일반 WAN(1352)으로 도시되었다. 그와 같은 네트워크 환경은 사무실, 기업 광역의 컴퓨터 네트워크, 인트라넷, 및 인터넷에서 일반적이다. LAN 네트워킹 환경에서 구현될 경우, 컴퓨터(1302)는 네트워크 인터페이스나 어댑터(1354)를 통해 로컬 네트워크(1350)로 연결된다. WAN 네트워킹 환경에서 구현될 경우, 컴퓨터(1302)는 일반적으로 광역 네트워크(1352)를 통하여 통신을 형성하는 모뎀(1356) 기타 수단을 포함한다. 컴퓨터(1302)에 대해 내장형이거나 외장형일 수 있

도면

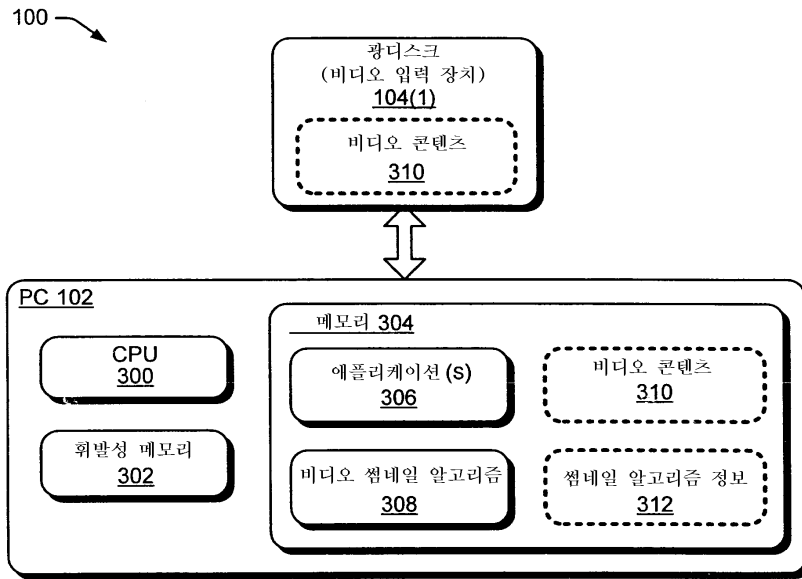
도면1



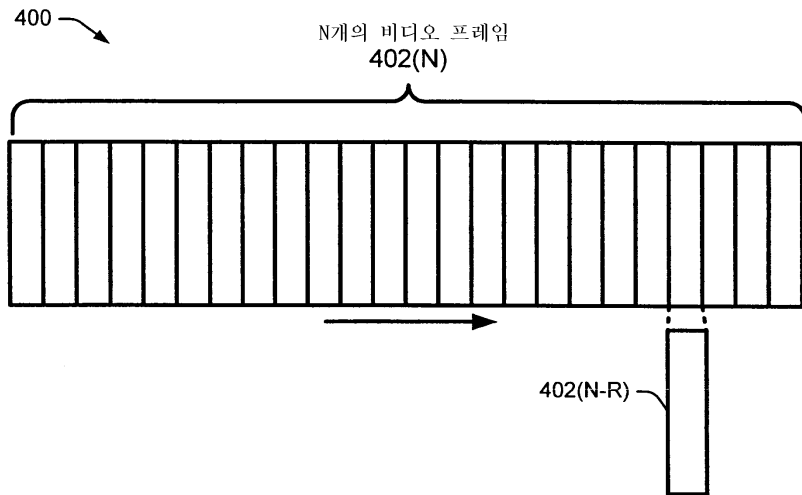
도면2



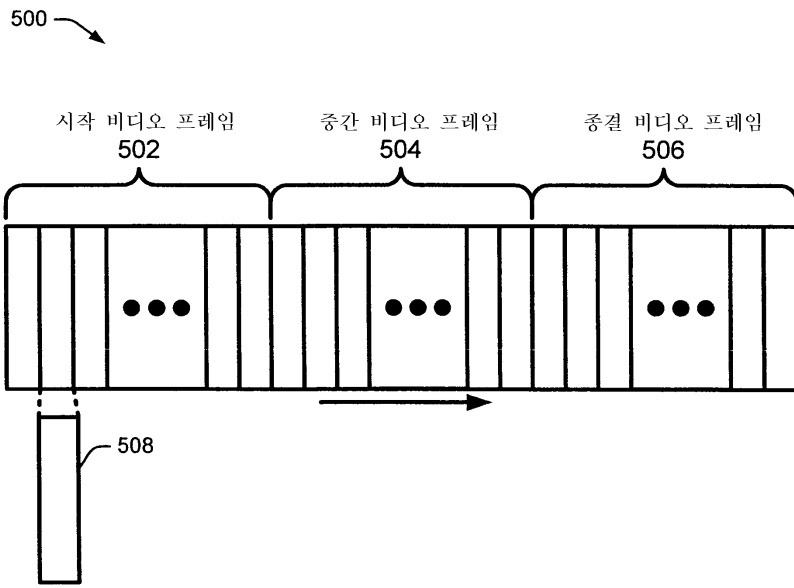
도면3



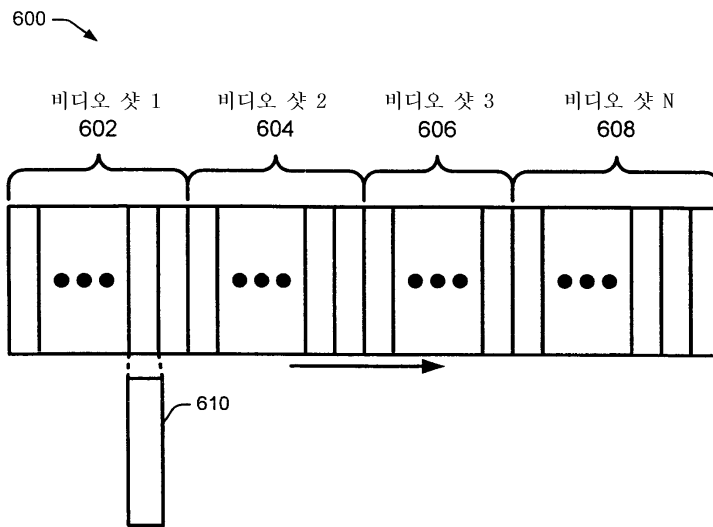
도면4



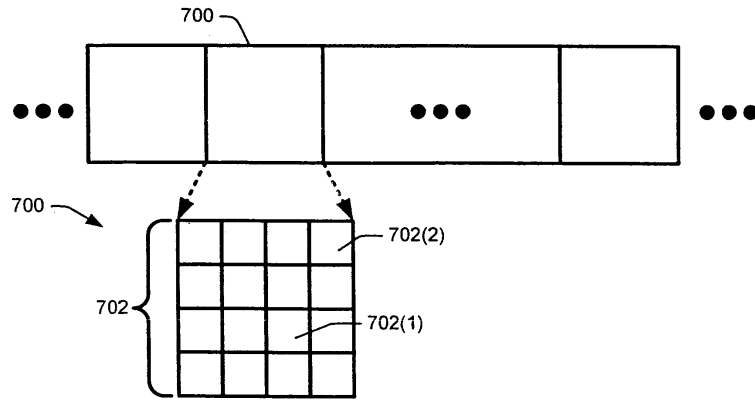
도면5



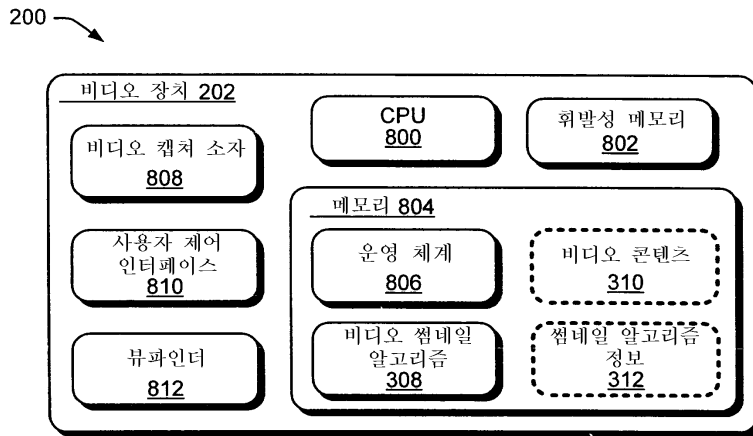
도면6



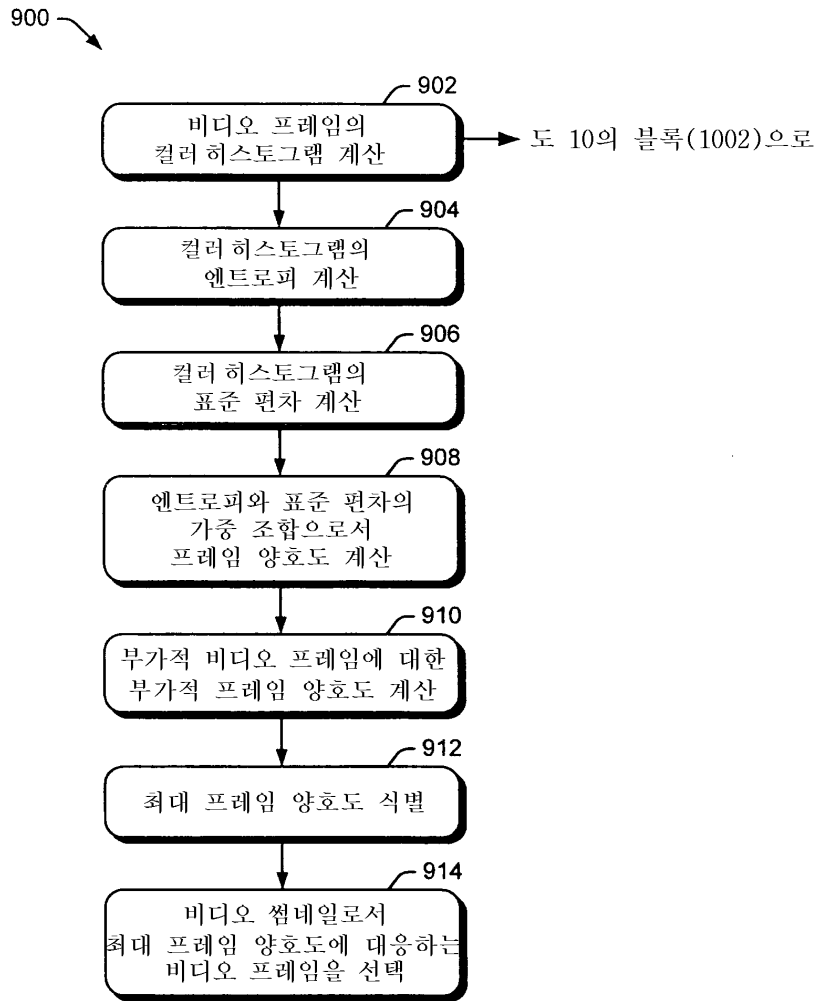
도면7



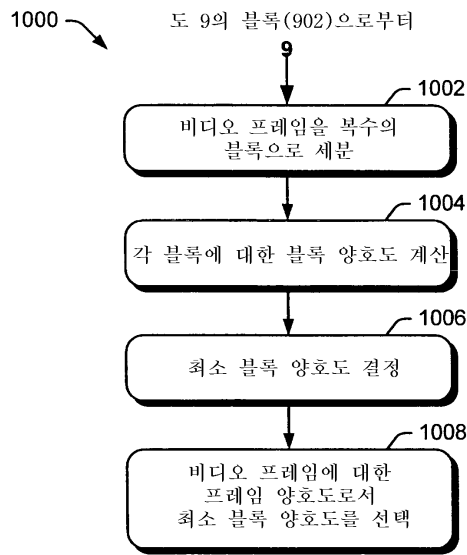
도면8



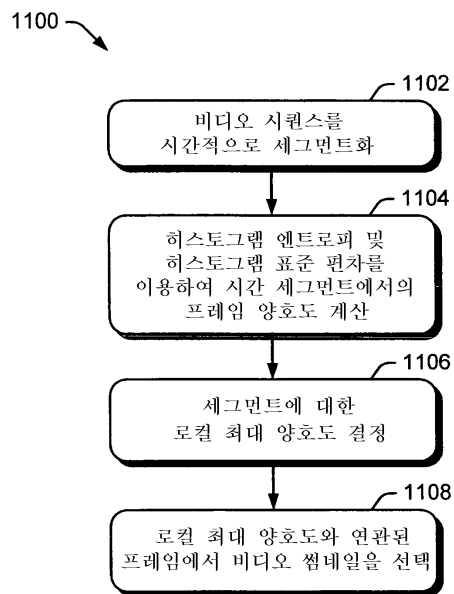
도면9



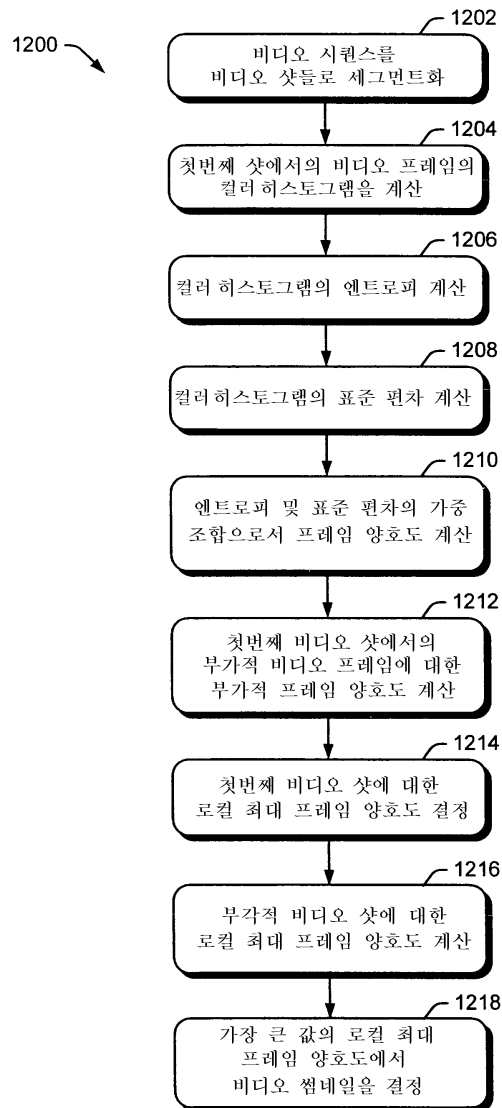
도면10



도면11



도면12



도면13

