

(19)대한민국특허청(KR)  
(12) 공개특허공보(A)

(51) Int. Cl.<sup>7</sup>  
G06F 17/30

(11) 공개번호 10-2005-0033075  
(43) 공개일자 2005년04월08일

(21) 출원번호 10-2005-7003343  
(22) 출원일자 2005년02월25일  
    번역문 제출일자 2005년02월25일  
(86) 국제출원번호 PCT/IB2003/003401  
    국제출원출원일자 2003년07월31일

(87) 국제공개번호 WO 2004/019224  
    국제공개일자 2004년03월04일

(30) 우선권주장 02078516.8 2002년08월26일 유럽특허청(EPO)(EP)

(71) 출원인 코닌클리케 필립스 일렉트로닉스 엔.브이.  
네델란드왕국, 아인드호펜, 그로네보르스베그 1  
(72) 발명자 스니즈더 프레디  
네델란드, 아아 아인드호펜 5656, 홀스트란 6  
파우루센 이고르 더블유. 에프.  
네델란드, 아아 아인드호펜 5656, 홀스트란 6

(74) 대리인 정상구  
신현문  
이범래

심사청구 : 없음

(54) 비디오 이미지들의 시퀀스에서 콘텐츠 속성을 검출하는 유닛 및 방법

명세서

기술분야

본 발명은 저-레벨 특징들에 기초하여 데이터 스트림에서 콘텐츠 속성을 검출하는 방법에 관한 것이다.

또한, 본 발명은 저-레벨 특징들에 기초하여 데이터 스트림에서 콘텐츠 속성을 검출하는 유닛에 관한 것이다.

또한, 본 발명은 그와 같은 유닛을 포함하는 이미지 처리 장치에 관한 것이다.

또한, 본 발명은 그와 같은 유닛을 포함하는 오디오 처리 장치에 관한 것이다.

배경기술

사람들이 생활하는 거실들로부터 액세스되고 소비될 수 있는 비디오 정보의 양은 계속 증가되고 있다. 이런 경향은 미래의 텔레비전 수신기들 및 개인용 컴퓨터들에 의해 제공되는 기술 및 기능성의 통합에 의해 더 가속화될 수 있다. 흥미로운 비디오 정보를 얻기 위해, 사용자들이 관련된 비디오 정보를 추출하는데 도움을 주고, 대용량의 이용가능한 비디오 정보를 통해 효과적으로 네비게이트하기 위한 툴들이 필요하다. 기존의 콘텐츠-기반 비디오 색인 및 검색 방법들은 상기 응용들에서 요청되는 툴들을 제공하지 않는다. 그러한 방법들의 대부분은 다음의 세가지 카테고리들로 분류될 수 있다: 1)비디오의 신택스 구조화; 2)비디오 분류; 및 3)의미적 추출(extraction of semantics).

제 1 카테고리의 작업은 쇼트 바운더리(shot boundary) 검출 및 키 프레임 추출, 쇼트 클러스터링, 콘텐츠 생성의 테이블, 비디오 요약 및 비디오 스키밍(video skimming)에 집중되어 있다. 대개, 이러한 방법들은 계산을 하는데 단순하고, 그 수행이 비교적 견고하다. 하지만, 그 결과들은 반드시 의미적으로 중요하거나 적절하지는 않다. 소비자-지향 응용들에 대해, 의미적으로 적절치 않은 결과들은 사용자를 혼란시키고, 검색 또는 브라우징 작업을 방해하도록 만들 수 있다.

제 2 카테고리 즉, 비디오 분류의 작업은 비디오 시퀀스들을 뉴스, 스포츠, 액션 영화들, 근접-촬영들, 대중 등과 같은 카테고리들로 분류하려는 것이다. 이러한 방법들은 사용자들이 조악한 레벨에서 비디오 시퀀스들의 브라우징을 용이하게 할

수 있는 분류 결과들을 제공한다. 더 정교한 레벨에서의 비디오-콘텐츠 분석은 아마도, 사용자들이 찾고 있는 것을 발견하는데 있어 더 효율적으로 돕는데 필요하다. 사실, 소비자들은 객체들, 행동들 및 이벤트들을 기술하는 키워드들과 같은, 더 정확한 의미의 라벨들로 검색 항목들을 종종 나타낸다.

제 3 카테고리 즉, 의미적 추출은 특정 도메인들에 대개 특정되어 있다. 예를 들어, 풋볼 경기들, 축구 경기들, 농구 경기들, 야구 경기들 및 감시 하의 사이트들에서 이벤트를 검출하는 방법이 검출하도록 제안되었다. 이러한 방법들의 이점들은, 검출된 이벤트들이 의미적으로 의미있고 보통 사용자들에 중요하다는 것이다. 하지만 결점은, 이들 대부분의 방법들이 방송 프로그램들에서의 편집 패턴들과 같은 특정 아티팩트들에 상당히 의존한다는 것이며, 이는 다른 이벤트들의 검출로 확장하려는 것을 어렵게 한다.

서두부에서 기술된 종류의 방법의 실시에는 논문 "A Semantic Event-Detection Approach and Its Application to Detecting Hunts in Wildlife Video", by Niels Haering, Richard J. Qian, and M. Ibrahim Sezan, in IEEE Transactions on Circuits and systems for Video Technology, Vol. 10, No. 6, September 2000로부터 공지된다. 상기 논문에서, 의미적 이벤트 검출에 대한 신축적 해결 방향으로의 계산 방법 및 여러 알고리즘 구성요소들이 제안된다. 자동화된 이벤트-검출 알고리즘은 비디오 콘텐츠에서 의미적으로 중요한 이벤트들의 검출을 촉진하고, 빠른 브라우징에 대한 의미적으로 중요한 하이이라이트들을 발생시키도록 돕는다. 그것은 다른 도메인들내의 다른 이벤트들을 검출하는데 적용되는 신축적 계산 접근이다. 3개의 레벨 비디오 이벤트-검출 알고리즘이 제안된다. 제 1 레벨은 색, 텍스처 및 움직임 특징들같은 비디오 이미지들로부터 저-레벨 특징들을 추출한다.

**발명의 상세한 설명**

본 발명의 목적은 비교적 견고한 서두부에서 기술된 종류의 방법을 제공하는 것이다.

본 발명의 이러한 목적은:

- 저-레벨 특징들의 시퀀스로부터 행동 특징을 결정하는 단계;
- 상기 결정된 행동 특징이, 행동 특징 공간 내의 행동 특징들의 미리 결정된 클러스터들의 세트로부터 어떤 클러스터에 속하는 지를 결정하는 단계;
- 상기 결정된 행동 특징 및 상기 결정된 클러스터에 기초하여 콘텐츠 속성 존재(content property presence)의 신뢰 레벨을 결정하는 단계; 및
- 상기 콘텐츠 속성 존재의 상기 결정된 신뢰 레벨에 기초하여 상기 콘텐츠 속성을 검출하는 단계를 포함하는 방법에 의해 달성된다.

콘텐츠 속성을 검출하는 저-레벨 특징들을 적용하는데 있어 문제는 저-레벨 특징들의 편차가 비교적 높다는 것이다. 저-레벨 특징들의 시퀀스로부터 행동 특징들을 추출하고 결정된 클러스터 및 행동 특징에 기초하여 신뢰 레벨을 결정함으로써, 분산은 의미있는 정보의 손실없이 감소한다. 상기 방법의 이점은, 다른 시간 스케일에서 다른 콘텐츠 속성들 예를 들어, 또한 장르들을 제외한 장면 변화들 같은 이벤트들을 검출하는 포괄적 접근이다.

데이터 스트림은 일련의 비디오 이미지들 또는 오디오 데이터에 대응할 수 있다. 저-레벨 특징들은 콘텐츠에 관한 별로 가공되지 않은 정보를 제공하고, 시간내의 저 정보 밀도를 갖는다. 저 레벨 특징들은 데이터 스트림의 샘플들에서 예를 들어, 이미지들의 경우에서의 픽셀 값들에서 단순 동작들에 기초한다. 동작들은 덧셈, 뺄셈 및 곱셈들을 포함할 수 있다. 예를 들어, 저 레벨 특징들은 평균 프레임 루미넌스, 프레임에서의 루미넌스 분산, 평균 MAD(절대 평균 편차)같은 특징들이다. 예를 들어, 높은 MAD 값들은 콘텐츠에서의 많은 움직임 또는 액션을 지시할 수 있고, 높은 분산은 콘텐츠의 타입에 관한 어떤 것을 말할 수 있다. 예를 들어, 광고들 및 카툰들은 높은 루미넌스 값들을 가진다. 대안으로, 저-레벨 특징들은 움직임 평가 처리로부터 유도된 파라미터들 예를 들어, 움직임 벡터들의 크기 또는 디코딩 처리로부터 유도된 파라미터들 예를 들어, DCT 계수들에 대응한다.

행동 특징들은 저-레벨 특징들의 행동에 관련된다. 그것은 예를 들어, 시간 함수로서 저-레벨 특징의 값들은 행동 특징에 의해 포함된다. 행동 특징의 값은 저-레벨 특징의 다중 값들의 결합에 의해 계산된다.

본 발명에 따른 실시예에서, 결정된 행동 특징은 시퀀스에서의 저-레벨 특징들 중 제 1 특징의 값들의 제 1 평균을 포함한다. 그것은 평균 값이 시퀀스의 시간 윈도우에서의 저-레벨 특징들 중 제 1 특징에 대해 계산된다는 것을 의미한다. 평균 값을 계산하는 것은 상대적으로 쉽다. 다른 이점은 평균 값을 계산하는 것이 분산을 감소시키기에 좋은 수단이라는 것이다. 저-레벨 특징으로부터의 행위 특징들을 추출하는 대안의 접근들은 다음의 것이다:

- 윈도우에서 저-레벨 특징의 표준 편차를 계산;
- 윈도우에서 저-레벨 특징의 푸리에 변환의 N 가장 중요한 파워 스펙트럼 값들을 취득;
- 윈도우에서 N 가장 중요한 주요 구성요소들을 취득. Christopher M. Bishop, "Neural Networks for Pattern Recognition", Oxford university press, 1995를 참조. T. Kohonen, "Self-Organizing Maps", Springer, 2001, ISBN 3-540-67921-9를 또한 참조.
- 윈도우에서 장면 변화들 또는 블랙 프레임들과 같은 저-레벨 이벤트들의 주파수 및/또는 밀도를 적용.

바람직하게 결정된 행동 특징은 시퀀스에서의 저-레벨 특징들 중 제 2 특징의 값들의 제 2 평균을 포함한다. 상기 경우에서, 행동 특징은 개별의 저-레벨 특징들에 각각 관련된, 다중 요소들 포함하는 벡터이다. 대안으로, 행동 특징은 하나의 저-레벨 특징에 각각 관련된 다중 요소들 예를 들어, 루미넌스의 평균 및 표준 편차를 포함한다. 하나의 저-레벨 특징을 주목하거나 다중 저-레벨 특징들을 개별적으로 주목하는 것은 장르 타입 또는 발생하는 이벤트의 타입에 관한 충분한 정보를 가장 바람직하게는 제공하지 못하지만, 다중 저-레벨 특징들의 결합 행동을 함께 주목하는 것은 훨씬 더 많은 정보를 제공하고 훨씬 더 많은 식별 능력을 제공한다.

본 발명에 따른 방법의 실시예에서, 콘텐츠 속성 존재의 신뢰 레벨은 행동 특징들의 결정된 클러스터의 모델에 기초하여 결정된다. 바람직하게, 상기 모델은 단순하고 견고하기에 선형 모델이다. 설계 단계 동안 행동 특징들의 다수의 예들은 테스트 데이터에 대해 결정된다. 예를 들어, 이 테스트 데이터는 주석이 달린 비디오 이미지들의 시간들이 될 수 있다. 주석은 이러한 비디오 이미지들의 각각에 대해, 상기 이미지들이 콘텐츠 속성을 갖는지 또는 아닌지 예를 들어, 상기 이미지들이 특정 장르인지 아닌지 알려지고 지시되는 것을 의미한다. 테스트 데이터의 행동 특징들의 분포의 구분에 의해 복수의 미리 결정된 클러스터들은 확립되었다. 이러한 미리 결정된 클러스터들의 각각에 대한 모델 및 클러스터 중심이 계산되었다. 검출 단계 동안 즉, 본 발명에 따른 방법을 적용할 때, 적합한 클러스터는 특정 행동 특징으로 결정된다. 이용된 클러스터 방법에 따라, 이것은 상기 특정 행동과 상기 여러 클러스터 중심들간의 유클리드 거리들을 계산함으로써 행해질 수 있다. 최소 유클리드 거리들은 특정 행동 특징이 속하는 미리 결정된 클러스터를 이끈다. 특정 행위 특징들을 위한 상기 적합하게 미리 결정된 클러스터의 모델의 평가로써, 대응하는 신뢰 레벨은 결정된다. 상기 신뢰 레벨은 모델 설계 단계 동안 이용된 주석 데이터를 갖는 특정 행동 특징에 대한 미리 결정된 클러스터의 모델의 적합에 관련된다. 또는 다른 말로, 특정 행동 특징은 상기 콘텐츠 속성에 실제로 대응하는 확률의 측정이다.

대안으로, 콘텐츠 속성 존재의 신뢰 레벨은 신경 네트워크로 결정된다.

본 발명에 따른 방법의 실시예에서, 상기 콘텐츠 속성을 검출하는 단계는 상기 콘텐츠 존재의 상기 신뢰 레벨을 미리 결정된 임계치와 비교함으로써 행해진다. 예를 들어, 상기 콘텐츠 속성의 신뢰 레벨이 미리 결정된 임계치보다 더 높다면 그때, 상기 데이터 스트림은 상기 콘텐츠 속성을 포함한다. 임계치를 이용한 이점은 상대적으로 쉽다는 것이다.

본 발명에 따른 방법의 실시예는 상기 콘텐츠 속성 존재의 상기 신뢰 레벨을 다른 행동 특징에 대응하는 다른 신뢰 레벨으로 비교하여 외층 필터링을 더 포함한다. 상기 신뢰 레벨이, 콘텐츠 속성이 데이터 스트림에 의해 실제로 포함되는 정확한 지시인지를 결정하도록 다중 행동 특징들은 선택적으로 적용된다. 바람직하게, 특정 행동 특징 주변의 시간 윈도우에서 다중 행동 특징들에 대응하는 상기 신뢰 레벨들은 외층 필터링으로 이용된다. 본 발명에 따른 이러한 실시예의 이점은 상대적으로 견고하고 단순하다는 것이다.

본 발명에 따른 방법의 실시예는 상기 비디오 이미지들 중 어느 것이 상기 콘텐츠 속성을 가지는 상기 일련의 비디오 이미지들의 일부에 대응하는 지를 결정하는 단계를 더 포함한다. 저-레벨 특징들의 시퀀스로부터 행동 특징들을 추출함으로써 예를 들어, 평균 넘으로써, 시간 이동은 상기 콘텐츠 속성 및 상기 콘텐츠 속성을 갖는 일련의 비디오 이미지들의 일부의 실제적인 시작의 검출에 도입된다. 예를 들어, 일련의 비디오 이미지들은 카툰의 일부 및 카툰에 속하지 않는 다른 일부를 포함하는 것으로 검출된다. 카툰으로부터 비-카툰으로 실제의 변환은 일련의 비디오 이미지에서의 카툰의 검출을 이끄는 행동 특징의 예에 기초하고, 시간 관련 파라미터들 예를 들어, 저-레벨 특징들로부터 행동 특징들을 검출하도록 이용된 윈도우의 크기에 기초하여 결정된다.

본 발명에 따른 방법의 실시예에서, 전자 프로그램 가이드(EPG)로부터의 데이터가 상기 콘텐츠 속성의 검출에 적용된다. 전자 프로그램 가이드로부터의 데이터와 같은 상위 레벨 데이터는 상기 콘텐츠 속성을 검출하는 방법의 견고함을 증가시키기 위해 매우 적합하다. 그것은 검출 문제에 문맥(context)을 제공한다. 검출기가 풋볼 시합들을 검출하도록 만드는 것은 이 검출기가 EPG에 의해 지시된 스포츠 프로그램들의 비디오 스트림들에 한정될 때 더 쉬어진다.

본 발명에 따른 방법의 실시예는:

- 상기 결정된 행동 특징이, 상기 행동 특징 공간(300) 내의 행동 특징들의 미리 결정된 클러스터들의 상기 세트로부터 어떤 또 다른 클러스터에 속하는지를 결정하는 단계;
- 상기 결정된 행동 특징 및 상기 또다른 결정된 클러스터에 기초하여 다른 콘텐츠 속성 존재의 또 다른 신뢰 레벨을 결정하는 단계; 및
- 상기 다른 콘텐츠 속성 존재의 상기 또다른 결정된 신뢰 레벨에 기초하여 다른 콘텐츠 속성을 검출하는 단계를 더 포함한다.

본 발명의 따른 이 실시예의 이점은, 다른 콘텐츠의 속성이 상대적으로 거의 추가 노력없이 검출될 수 있다는 것이다. 가장 비용이 많이 드는 계산들 예를 들어, 저-레벨 특징들의 계산 및 행동 특징들의 계산은 분배된다. 상대적으로 단순한 처리 단계들만이 다른 콘텐츠 속성의 부가의 검출에 대해 특징된다. 이런 실시예로 예를 들어, 비디오 이미지들의 시퀀스가 카툰에 대응하는지와 비디오 이미지들의 시퀀스가 야생 영화에 대응하는지를 검출하는 것이 가능하다.

본 발명의 또다른 목적은 상대적으로 견고한 검출을 수행하도록 설계된 서두부에서 기술된 종류의 유닛을 제공하는 것이다.

본 발명의 이런 목적은 상기 유닛이:

- 상기 저-레벨 특징들의 시퀀스로부터 행동 특징을 결정하는 제 1 결정 수단;

- 상기 결정된 행동 특징이, 행동 특징 공간 내의 행동 특징들의 미리 결정된 클러스터들의 세트로부터 어떤 클러스터에 속하는지를 결정하는 제 2 결정 수단;
- 상기 결정된 행동 특징 및 상기 결정된 클러스터에 기초하여 콘텐츠 속성 존재의 신뢰 레벨을 결정하는 제 3 결정 수단; 및
- 상기 콘텐츠 속성 존재의 상기 결정된 신뢰 레벨에 기초하여 상기 콘텐츠 속성을 검출하는 검출 수단을 포함함으로써 달성된다.

서두부에서 기술된 이미지 처리 장치에서의 본 발명에 따른 유닛의 실시예를 적용하는 것이 유리하다. 이미지 처리 장치는 부가의 구성요소들 예를 들어, 이미지들을 디스플레이하는 디스플레이 장치, 이미지들의 기억을 위한 기억 장치 또는 비디오 압축 즉, 예를 들어 MPEG 표준 또는 H26L에 따른 엔코딩 또는 디코딩을 위한 비디오 압축 장치를 포함할 수 있다. 이미지 처리 장치는 다음의 응용들 중 하나를 지원할 수 있다:

- 장르 또는 이벤트 정보에 기초한 기록된 데이터의 검색;
- 장르 및 이벤트 정보에 기초한 데이터의 자동 기록;
- 재생 동안, 동일 장르를 갖는 기억된 데이터 스트림들간의 호핑(hopping);
- 재생 동안, 이벤트에서 동일 타입의 이벤트로의 호핑으로 예를 들어, 풋볼 득점에서 풋볼 득점으로 호핑;
- 어떤 장르가 다른 채널 상에서 방송된다면 사용자에게 알리기. 예를 들어, 사용자가 한 채널에서 시청하고 있을 때 풋볼 시합이 다른 채널 상에서 시작하는 것이 사용자에게 알려진다.
- 특정 이벤트이 발생했다면 사용자에게 알리기. 예를 들어, 사용자가 한 채널을 시청하고 있지만 풋볼 득점이 다른 채널 상에서 일어난 것을 사용자에게 알려진다.
- 비디오 카메라로 모니터되는 방에서 발생된 무엇인가를 경비원에게 알리기.

방법의 변경 및 그에 따른 변화들은 기술된 유닛의 변경들 및 그에 따른 변화들에 대응할 수 있다.

본 발명에 따른 방법, 유닛 및 이미지 처리 장치의 여러 측면들은 이하에서 기술된 구현들 및 실시예들과 첨부된 도면들에 의해 분명해지고 명백해 질 것이다.

### 도면의 간단한 설명

- 도 1A는 저-레벨 특징들의 예들 및 상기 저-레벨 특징들로부터 추출된 행동 특징들을 도시한 도면;
- 도 1B는 도 1A로부터의 행동 특징 벡터들을 대해 가장 적합한 클러스터들의 예들을 도시한 도면;
- 도 1C는 도 1A의 행동 특징 벡터들 및 도 1B에서의 가장 적합한 클러스터들에 기초하여 결정되는 신뢰 레벨을 도시한 도면;
- 도 1D는 도 1C의 신뢰 레벨들의 임계치 및 외층 이동(outlier removal)한 후 최종 출력을 도시한 도면;
- 도 2는 데이터 스트림에서 콘텐츠 속성을 검출하는 유닛을 개략적으로 도시한 도면.
- 도 3은 행동 특성 벡터들의 다수의 클러스터들을 포함하는 행동 특징 공간을 개략적으로 도시한 도면.
- 도 4는 저-레벨 특징들에 기초하여 처리를 분석하는 콘텐츠의 불력도를 개략적으로 도시한 도면.
- 도 5는 본 발명에 따른 이미지 처리 장치의 요소들을 개략적으로 도시한 도면.

### 실시예

동일 참조 번호들은 도면들을 통해 유사한 부분들을 나타내는데 이용된다.

예시에 의해 본 발명에 따른 방법은 이하에서 설명될 것이다. 상기 예시는 카툰 검출에 관련된다. 도 1A 내지 1D에서, 예시에 포함된 일부 곡선들이 도시된다. 카툰 검출에서 이용된 저-레벨 특징들은 MPEG2 엔코더로부터 추출된다. 엔코딩에 대해 이용된 GOP(Group Of Pictures; 픽처들의 그룹) 길이는 12 였다. 일부 특징들은 단지 이용가능한 모든 I-프레임이고, 다른 것은 이용가능한 모든 프레임이다. 이용된 저-레벨 AV 특징들의 개관에 대해 표 1을 주목하라. 이 예에서, 어떤 오디오 특징들도 오디오 특징들을 제외하곤 이용되지 않았다.

표 1: 저-레벨 특징들의 예시들

도 1A은 이러한 저-레벨 특징들로부터 추출된 저-레벨 특징들의 예시들 및 행동 특징들을 도시한다. 도 1A는 모든 프레임(104) 및 데이터 스트림의 예시적인 부분의 모든 I-프레임에 대한 총 프레임 루미넌스(102)에 대한 MAD를 도시한다. 데이터 스트림은 6분분량의 비디오 이미지들에 대응하고, 비-카툰에서 카툰 재료로의 변환을 포함한다. 변환의 위치는 수직 선(101)으로 마크된다. 행동 특징들 때문에, 시간 윈도우에 걸쳐 저 레벨 특징들의 평균(106, 108) 및 표준 편차(110, 112)가 계산된다. 평균 및 표준 편차가 계산되기 전에, 저-레벨 특징들은 정규화된다. 계산된 평균 값들 및 표준 편차 값들은 행동 특징 벡터를 형성하기 위해 벡터에서 적층된다. 윈도우의 모든 GOP는 쉬프트되고, 새로운 행동 특징 벡터는 계산된다. 이용된 윈도우 길이는 250 GOP's이고, 대략 2분 분량이다. GOP 내의 통계들에 기초하여 프레임을 평균내는 것은 더욱 견고한 특징들을 제공한다. 예를 들어, MAD는 매우 큰 동적 범위를 가진다: 쇼트 컷(shot cut)이 발생할 때, 값은 크기의 순서들이 될 수 있고, 콘텐츠에서 많은 움직임이 없을 때, 더 높아질 것이다.

설계 단계에서, 행동 특징 벡터 공간은 자기-조직화 맵(Self-Organizing Map)을 이용하는 클러스터들로 분할된다. T. kohonen의 "Self-Organizing Maps", Springer, 2001, ISBN 3-540-67921-9를 보자. 자기-조직화 맵은 행동 특징 공간을 클러스터할 수 있어, 행동 특징 공간의 행동 특징 벡터 분산의 뛰어난 표현을 형성한다. SOM의 클러스터들은 맵에서 공간적으로 조직되고, 우리의 경우에서, 맵은 클러스터들을 포함하는 유닛들의 3x3으로 구성된다. 이 예에서, 공간 구성 특징은 이용되지 않지만, 맵 상에서 위치가 정보를 제공하기 때문에, 검출 품질을 더 개선시킬 수 있다. 다시 말해서, 9개의 미리 결정된 클러스터들이 존재한다. 설계 단계 동안, SOM에서의 모든 클러스터에 대해, 국부 선형 분류 모델이 또한 제작되었다.

각 행동 특징 벡터에 대한 검출 위상에서, 적합한 클러스터가 결정된다. 그것은 SOM이 행동 특징 벡터를 이용하여 평가된다는 것을 의미한다. 상기 평가는 행동 특징 벡터에 가장 적합한 클러스터를 지시하는 클러스터 인덱스 내에서 발생한다. 도 1B는 예시적인 데이터 스트림의 행동 특징 벡터들에 가장 적합한 클러스터 색인들을 도시한다.

검출 위상에서, 선택된 클러스터에 속하는 모델은 행동 특징 벡터를 이용하여 평가된다. 각 평가는 신뢰-레벨 즉, "카툰 신뢰(cartoon-ness confidence)"가 된다. 도 1C는 각 GOP에 대한 "카툰 신뢰"를 도시한다 즉, 도 1C는 도 1A의 행동 특징 벡터들 및 도 1B의 클러스터 인덱스들에 기초하여 결정된 신뢰 레벨을 도시한다. 도시된 신뢰 레벨은 반드시 엄격한 개연적 판단에서의 신뢰는 아닌데, 값들이 범위 0과 1사이의 범위에 존재하지 않기 때문이다.

모든 GOP를 요약하기 위해, 새로운 행동 특징 벡터는 계산되고, 클러스터 인덱스는 이러한 행동 특징 벡터들에 가장 적합하도록 발견된다. 따라서, 모든 GOP의 단지 하나의 국부 선형 모델은 계산된 행동 특징 벡터 상에 평가된다.

임계치에 의해, 콘텐츠 속성은 검출된다. 즉, 신뢰 레벨을 미리 결정된 임계치와 비교함으로써 데이터 스트림은 카툰에 속하는 이미지들을 포함하는 것이 검출된다. 미리 결정된 임계치는 설계 단계 동안 결정되었다. 도 1C의 하위 부분은 임계치의 출력(118)을 도시한다. "카툰 신뢰"가 미리 결정된 임계치 이상이라면, 출력(118)은 1이고, "카툰 신뢰"가 미리 결정된 임계치 미만이라면, 출력은 0이다.

임계치의 출력(118)에서, 일부 외층(120 내지 126)들이 존재한다. 그것은 출력(118)에서 스파이크들이 존재한다는 것을 의미한다. 필터링함으로써, 이러한 외층(120 -126)들은 제거된다. 이러한 필터링은 다음처럼 작동한다. 시간 윈도우 내에서 임계치에 의해 결정된 분류의 확율은 양(예를 들어, 1)으로 결정된다. 상기 확율이 제 2 미리 결정된 임계치보다 더 높다면, 카툰이 나타난다고 결정되고, 그렇지 않다면, 어떤 카툰도 나타나지 않는다. 외층 이동 윈도우 길이 및 제 2 미리 결정된 임계치는 위상 설계동안 계산되어진다.

카툰이 데이터 스트림에 의해 나타난 비디오 시퀀스에서 나타나도록 결정된 후, 그것은 카툰의 시작 및 끝을 결정하도록 요청될 수 있다. 여러 시간 윈도우들의 길이들을 고려하여 예를 들어, 행동 특징들 및 외층 이동을 추출하여, 가장 나쁜-경우의 시작 및 끝은 계산될 수 있다. 가장 나쁜 경우의 시작(103) 및 끝은 완전한 카툰이 시작(103) 및 끝 내에 있을 때 매우 높은 확실성이 존재한다는 것이다. 이것은 흥미로운데, 본 발명에 따른 이미지 처리 장치의 사용자는 카툰이 이미 시작된 후에 검출된 가툰의 재생을 시작하거나 카툰이 끝나기 전에 재생을 정지함으로써 성가시게 된다. 예시적인 데이터 스트림의 계산된 가장 나쁜 경우의 시작(103)은 도 1D에서 도시된다.

도 2는 저-레벨 특징들에 기초하여 데이터 스트림의 콘텐츠 속성을 검출하는 유닛(200)을 개략적으로 도시한다. 유닛(200)은 :

- 입력 접속기(212)에 제공된 저-레벨 특징들(102, 104)의 시퀀스들로부터 행동 특징들(106 내지 112)를 추출하는 추출 유닛(202). 저-레벨 특징들은 비디오 또는 오디오 데이터에 기초하여 계산될 수 있다. 행동 특징들은 스칼라들 또는 벡터들이 될 수 있다.
- 행동 특징들이 속하는 행동 특징 공간(300) 내의 행동 특징들(318 내지 328)의 미리 결정된 클러스터들(302 내지 316) 중 하나를 결정하는 제 1 결정 유닛(204). 또한, 도 1B 및 도 3을 참조.
- 행동 특징들(318 내지 328)의 선택된 클러스터들(302 내지 316)의 기초하여 개별적인 행동 특징들의 신뢰 레벨들을 결정하는 제 2 결정 유닛(206). 또한 도 1C 및 도 3을 참조.
- 행동 특징들의 신뢰 레벨들에 기초하여 콘텐츠 속성을 검출하는 분류 유닛(208). 선택적으로, 이 분류 유닛(208)은 도 1D에 관련되어 기술된 외층 이동 필터들을 포함한다.

- 콘텐츠 속성을 갖는 시퀀스의 부분의 시작을 계산하는 시작 및 종료 계산 유닛(210). 이 시작 계산 유닛(210)은 도 1에 관련되어 기술된다. 이 시작 계산 유닛은 선택적이다. 콘텐츠 속성을 검출하는 추출 유닛(202), 제 1 결정 유닛(204), 제 2 결정 유닛(204), 분류 유닛(208) 및 유닛(200)의 시작 및 종료 계산 유닛은 하나의 처리기를 이용하여 구현될 수 있다. 일반적으로, 이러한 기능들은 소프트웨어 제품의 제어 아래 수행된다. 실행 동안 일반적으로 소프트웨어 프로그램 제품은 RAM 같은 메모리에 로딩되고, 그것으로부터 실행된다. 프로그램은 ROM, 하드 디스크 또는 자기 및/또는 광 기억 장치와 같은 후순위 메모리로부터 로딩되거나 인터넷과 같은 네트워크를 통해 로딩될 수 있다. 선택적으로 응용 특정 집적 회로는 계산된 기능을 제공한다.

상기 방법은 하드웨어 검출 유닛들에 대한 설계 템플릿을 제공하고, 모든 유닛에서 구성요소들은 같지만, 설계 파라미터들은 다르다.

도 3은 행동 특징 벡터들(318 내지 328)의 클러스터들(302 내지 316)의 번호를 포함하는 행동 특징 공간(300)을 개략적으로 도시한다. 도 3에 도시된 행동 특징 공간(300)은 다중-차원 공간이다. 행동 특징 공간(300)의 축들의 각각은 행동 특징 벡터들(318 내지 328)의 각각 요소들에 대응한다. 행동 특징 공간(300) 내의 각 클러스터(302 내지 316)는 콘텐츠의 모드로서 해석될 수 있다. 예를 들어, 콘텐츠 속성이 "비디오 이미지들의 시퀀스에서의 카툰"에 대응하는 경우에서, 제 1 클러스터(302)는 빠른 이동 캐릭터를 갖는 카툰의 제 1 모드에 대응할 수 있다. 원칙적으로, 클러스터는 특정 콘텐츠 속성에 독립한다; 한 클러스터는 변화하는 루미넌스를 갖는 빠른 이동 재료를 가리킬 수 있다. 그 때, 로컬 모델에 의해 나타내진 관계는, 낮은 루미넌스를 갖는 특징 벡터들은 카툰을 갖지 않지만, 높은 루미넌스를 갖는 벡터들은 카툰들을 갖는 것을 언급한다. 다른 클러스터들에서, 다른 관계는 (상기 클러스터에 속하는 국부 모델에 의해 기술되어) 존재할 수 있다. 제 2 클러스터(316)는 느린 이동 캐릭터들을 갖는 카툰의 제 2 모드에 대응할 수 있고, 제 3 클러스터(306)는 야간의 카툰 장면 에 대응할 수 있다.

클러스터들(302 내지 316)의 각각에 대해, 모델은 설계 단계 동안 결정된다. 그것은 적어도 스케어 방법으로 한 세트의 방

정식들의 해를 구하여 결정되는 선형 모델이 될 수 있다. N 요소들을 갖는 행동 특징 벡터  $x$  의 한 예에 대해, 선형 모델  $M_i$ 에 대한 방정식은 수학식 1에서 주어진다:

설계 단계 동안, 파라미터들  $\alpha_k$  ( $1 \leq k \leq N$ )의 N 값들 및 파라미터  $\beta_i$ 의 N 값들은 결정되어야 한다. 설계 단계 동안, 테스트 데이터의 특정 행동 특징 벡터가 콘텐츠 속성을 갖지 않는 데이터 예를 들어, 비디오 이미지의 일부에 대응하는 테스트 데이터의 일부에 대응하면 y의 값은 0이고, 테스트 데이터의 특정 행동 특징 벡터가 콘텐츠 속성을 갖는 데이터의 일부에 대응하면 y의 값은 1이다.

검출 단계에서, y의 값은 목표 데이터의 특정 행동 특징 벡터에 대한 신뢰 레벨에 대응한다. y의 이 후자 값은 파라미터들  $\alpha_k$  ( $1 \leq k \leq N$ ) 및 파라미터  $\beta_i$ 의 공지된 값들을 갖는 타겟 데이터의 특정 행동 특징 벡터에 대한 평가하는 수학식 1에 의해 쉽게 발견된다.

도 4는 데이터 스트림의 계산된 저-레벨 특징들에 기초하여 처리를 분석하는 콘텐츠의 블러드도를 개략적으로 도시한다. 저-레벨 특징들은 행동 특징들의 추출(402)에 대한 입력들이다. 이러한 행동 특징들은 예를 들어, 비디오 시퀀스를 나타내는 데이터 스트림이 카툰(404) 또는 광고(406) 또는 스포츠 게임(408)을 포함하는지를 검출하기 위해 다중 결정 처리들(404 내지 408)로 이용된다. 관련된 데이터 스트림들의 EPG 정보로부터 유도된 데이터 스트림 또는 정적 데이터에 대응하는 EPG로부터의 정보는 데이터 스트림을 분석하기 위해 선택적으로 적용된다.

선택적으로, 제 1 결정 처리들(408)로부터의 중간 결과들(414)은 제 2 결정 처리(406)에 제공되고, 제 2 결정 처리(306)으로부터의 결과들(412)은 제 3 결정 처리(404)에 제공된다. 이러한 결정 처리들(404 내지 408)은 다른 시간 스케일들 즉, 예를 들어 장면 변화들 및 광고 분리들을 갖는 단기로부터 예를 들어, 하이 라이트들, 비디오 클립들, 유사 콘텐츠들을 갖는 중기 및 장르 인식 및 사용자 선호 인식을 갖는 장기까지에 대응할 수 있다. 결정 처리들(404 내지 408)의 최종 결과들은 선택적으로 결합된다(410). 예를 들어 원칙적으로, 408로부터의 정보는 또한 직접 404로 갈 수 있다.

도 5는 본 발명에 따른 이미지 처리 장치(500)의 요소들을 개략적으로 도시하고, 본 발명은 다음을 포함한다:

- 일부 처리가 수행된 후 디스플레이되는 이미지들을 나타내는 데이터 스트림을 수신하는 수신 유닛(502). 상기 신호는 안테나 또는 케이블을 통해 수신된 방송 신호가 될 수 있지만, 또한 VCR(비디오 카세트 녹화기) 또는 DVD(디지털 비디오 디스크)와 같은 기억 장치로부터의 신호가 될 수 있다. 신호는 입력 접속기(510)에 제공된다.

- 도 1A 내지 1D에 관련하여 기술된 저-레벨 특징들에 기초하여 데이터 스트림 내의 콘텐츠 속성을 검출하는 유닛(504);

- 콘텐츠 속성을 기초하여 콘텐츠 속성을 검출하는 유닛(504)에 의해 제어되는 콘텐츠 이미지 처리 유닛(506). 상기 이미지 처리 유닛(506)은 예를 들어, 상기 유닛(504)이, 잡음 감소의 양이 증가되는 카툰에 대응하는 데이터 스트림을 검출하는 경우에서, 잡음 감소를 수행하도록 배열될 수 있다.

- 처리된 이미지들을 디스플레이하는 디스플레이 장치(508). 상기 디스플레이 장치는 선택적이다.

상기-언급된 실시예들은 본 발명을 제한하려는 것이 아니라 설명을 위한 것이고, 당업자는 첨부된 청구항들의 범위를 벗어나지 않고 대안의 실시예들을 설계할 수 있을 것이다. 상기 청구항들에서, 괄호들 간의 위치된 어떤 참조 번호들은 청구항을 제한하는 것으로 해석되지 않을 것이다. 단어 '포함하는'은 청구항들에 열거되지 않은 요소들 또는 단계들의 존재를

배제하지 않을 것이다. 요소들에 선행하는 단어 "a" 또는 "an"은 복수의 그와 같은 요소들의 존재를 배제하지 않을 것이다. 본 발명은 여러 구별된 요소들을 포함하는 하드웨어에 의하고, 적합하게 프로그램된 컴퓨터에 의해 구현될 수 있다. 여러 수단을 열거하는 장치 청구항들에서, 이러한 여러 수단은 하드웨어의 하나 및 동일 항목에 의해 구현될 수 있다.

(57) 청구의 범위

**청구항 1.**

저-레벨 특징들에 기초하여 데이터 스트림에서 콘텐츠 속성을 검출하는 방법에 있어서,

- 상기 저-레벨 특징들의 시퀀스로부터 행동 특징을 결정하는 단계;
- 상기 결정된 행동 특징이, 행동 특징 공간 내의 행동 특징들의 미리 결정된 클러스터들의 세트로부터 어떤 클러스터에 속하는지를 결정하는 단계;
- 상기 결정된 행동 특징 및 상기 결정된 클러스터에 기초하여 콘텐츠 속성 존재(content property presence)의 신뢰 레벨을 결정하는 단계; 및
- 상기 콘텐츠 속성 존재의 상기 결정된 신뢰 레벨에 기초하여 상기 콘텐츠 속성을 검출하는 단계를 포함하는, 콘텐츠 속성 검출 방법.

**청구항 2.**

제 1 항에 있어서,

상기 데이터 스트림은 일련의 비디오 이미지들에 대응하는, 콘텐츠 속성 검출 방법.

**청구항 3.**

제 1 항에 있어서,

상기 결정된 행동 특징은 상기 시퀀스에서의 상기 저-레벨 특징들 중 제 1 특징의 값들의 제 1 평균을 포함하는, 콘텐츠 속성 검출 방법.

**청구항 4.**

제 3 항에 있어서,

상기 결정된 행동 특징은 상기 시퀀스에서의 상기 저-레벨 특징들 중 제 2 특징의 값들의 제 2 평균을 포함하는, 콘텐츠 속성 검출 방법.

**청구항 5.**

제 1 항에 있어서,

상기 콘텐츠 속성 존재의 상기 신뢰 레벨은 상기 행동 특징들의 결정된 클러스터의 모델에 기초하여 결정되는, 콘텐츠 속성 검출 방법.

**청구항 6.**

제 5 항에 있어서,

상기 행동 특징들의 결정된 클러스터의 상기 모델은 선형 모델인, 콘텐츠 속성 검출 방법.

**청구항 7.**

제 1 항에 있어서,

상기 콘텐츠 속성 존재의 상기 신뢰 레벨은 신경 회로망(neural network)으로 결정되는, 콘텐츠 속성 검출 방법.

### 청구항 8.

제 1 항에 있어서,

상기 콘텐츠 속성을 검출하는 단계는 상기 콘텐츠 존재의 상기 신뢰 레벨을 미리 결정된 임계치와 비교함으로써 행해지는, 콘텐츠 속성 검출 방법.

### 청구항 9.

제 1 항에 있어서,

상기 콘텐츠 속성 존재의 상기 신뢰 레벨을 다른 행동 특징에 대응하는 다른 신뢰 레벨과 비교하여 외층 필터링(outlier filtering)하는 단계를 더 포함하는, 콘텐츠 속성 검출 방법.

### 청구항 10.

제 2 항에 있어서,

상기 비디오 이미지들 중 어느 것이 상기 콘텐츠 속성을 가지는 상기 일련의 비디오 이미지들의 일부에 대응하는 지를 결정하는 단계를 더 포함하는, 콘텐츠 속성 검출 방법.

### 청구항 11.

제 1 항에 있어서,

EPG로부터의 데이터가 상기 콘텐츠 속성의 검출에 적용되는, 콘텐츠 속성 검출 방법.

### 청구항 12.

제 1 항에 있어서,

- 상기 결정된 행동 특징이, 상기 행동 특징 공간(300) 내의 행동 특징들의 미리 결정된 클러스터들의 상기 세트로부터 어떤 또 다른 클러스터에 속하는 지를 결정하는 단계;
- 상기 결정된 행동 특징 및 상기 또다른 결정된 클러스터에 기초하여 다른 콘텐츠 속성 존재의 또다른 신뢰 레벨을 결정하는 단계; 및
- 상기 다른 콘텐츠 속성 존재의 상기 또다른 결정된 신뢰 레벨에 기초하여 다른 콘텐츠 속성을 검출하는 단계를 더 포함하는, 콘텐츠 속성 검출 방법.

### 청구항 13.

저-레벨 특징들에 기초하여 데이터 스트림에서 콘텐츠 속성을 검출하는 유닛에서 있어서,

- 상기 저-레벨 특징들의 시퀀스로부터 행동 특징을 결정하는 제 1 결정 수단;
- 상기 결정된 행동 특징이, 행동 특징 공간 내의 행동 특징들의 미리 결정된 클러스터들의 세트로부터 어떤 클러스터에 속하는 지를 결정하는 제 2 결정 수단;

- 상기 결정된 행동 특징 및 상기 결정된 클러스터에 기초하여 콘텐츠 속성 존재의 신뢰 레벨을 결정하는 제 3 결정 수단; 및
- 상기 콘텐츠 속성 존재의 상기 결정된 신뢰 레벨에 기초하여 상기 콘텐츠 속성을 검출하는 검출 수단을 포함하는, 콘텐츠 속성 검출 유닛.

#### 청구항 14.

이미지 처리 장치에 있어서,

- 비디오 이미지들의 시퀀스를 나타내는 데이터 스트림을 수신하는 수신 수단;
- 저-레벨 특징들에 기초하여 상기 비디오 이미지들의 시퀀스에서 콘텐츠 속성을 검출하는, 제 13 항에서 청구된 검출 유닛; 및
- 상기 콘텐츠 속성에 기초하여 콘텐츠 속성을 검출하는 상기 유닛에 의해 제어되는 이미지 처리 유닛을 포함하는, 이미지 처리 장치.

#### 청구항 15.

제 13 항에 있어서,

상기 이미지 처리 유닛은 기억 장치를 포함하는, 이미지 처리 장치.

#### 청구항 16.

제 13 항에 있어서,

상기 이미지 처리 유닛은 비디오 이미지 압축 장치를 포함하는, 이미지 처리 장치.

#### 청구항 17.

오디오 처리 장치에 있어서,

오디오를 나타내는 데이터 스트림을 수신하는 수신 수단;

저-레벨 특징들에 기초하여 상기 오디오에서의 콘텐츠속성을 검출하는, 제 13 항에서 청구된 유닛; 및

상기 콘텐츠 속성에 기초하여, 콘텐츠 속성을 검출하는 상기 유닛에 의해 제어된 오디오 처리 유닛을 포함하는, 오디오 처리 장치.

#### 요약

저-레벨 특징들에 기초하여 데이터 스트림에서의 콘텐츠 속성의 검출 방법이 제안된다. 상기 방법은 저-레벨 특징들의 시퀀스로부터 행동 특징(예를 들어, 320)을 결정하는 단계(202); 상기 결정된 행동 특징(320)이 속하는 행동 특징 공간(300) 내의 행동 특징들(318 내지 328)의 미리 결정된 클러스터들(304) 중 하나를 결정하는 단계; 행동 특징들의 상기 결정된 클러스터(304) 및 상기 결정된 행동 특징에 기초하여 콘텐츠 속성 존재(content property presence)의 신뢰 레벨을 결정하는 단계; 및 상기 콘텐츠 속성 존재의 상기 신뢰 레벨에 기초하여 상기 콘텐츠 속성을 검출하는 단계를 포함한다.

#### 대표도

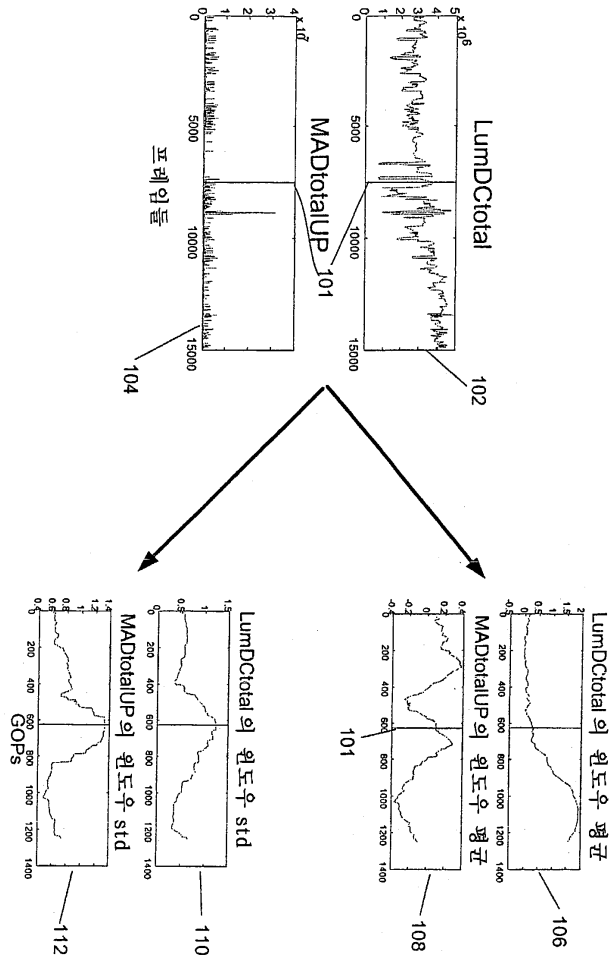
도 3

#### 색인어

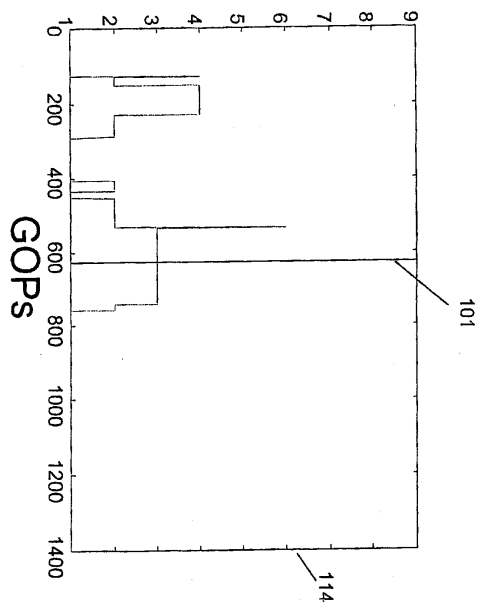
저-레벨 특징, 콘텐츠 속성, 클러스터, 신뢰 레벨, 검출

도면

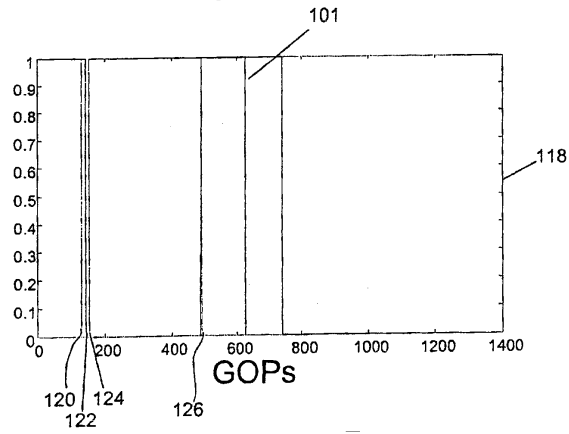
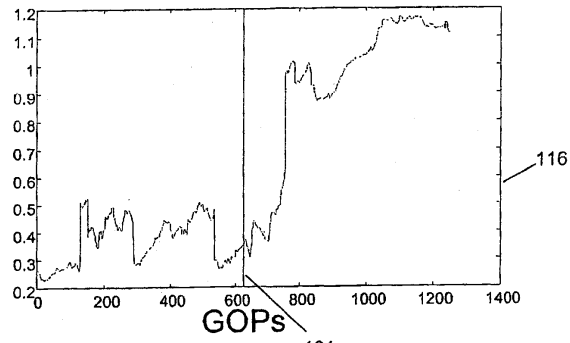
도면1a



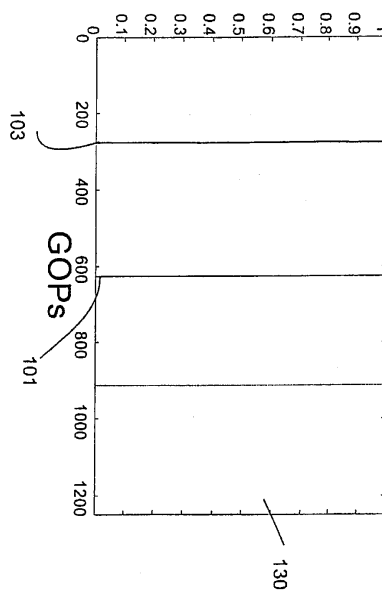
도면1b



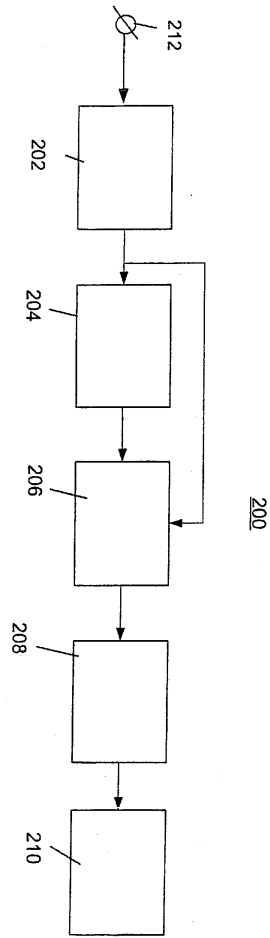
도면1c



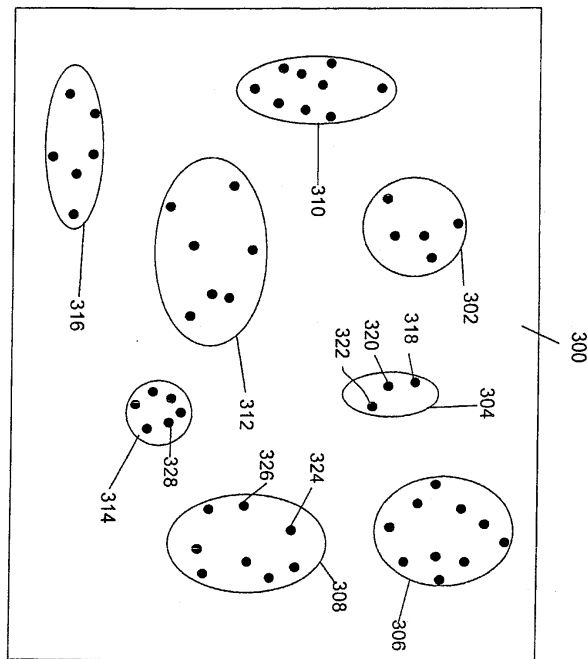
도면1d



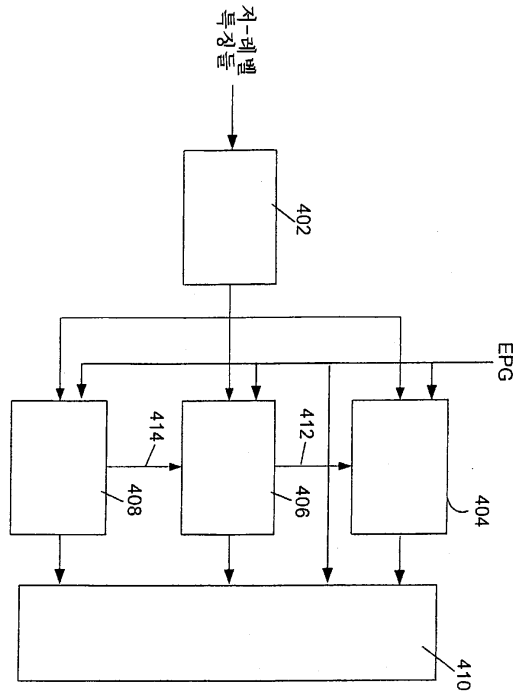
도면2



도면3



도면4



도면5

