



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2007-0107722
(43) 공개일자 2007년11월07일

(51) Int. Cl.

H04N 7/24 (2006.01) *H04N 11/02* (2006.01)
G06T 9/00 (2006.01) *H04N 11/02* (2006.01)
G06T 9/00 (2006.01)

(21) 출원번호 10-2007-7019634
(22) 출원일자 2007년08월28일
심사청구일자 없음
번역문제출일자 2007년08월28일
(86) 국제출원번호 PCT/US2006/001907
국제출원일자 2006년01월20일
(87) 국제공개번호 WO 2006/083567
국제공개일자 2006년08월10일

(30) 우선권주장
60/648,094 2005년01월28일 미국(US)
60/653,810 2005년02월17일 미국(US)

(71) 출원인

유클리드 디스커버리스, 엘엘씨

미국 매사추세츠 콩코드 슈트 212 모누먼트 스웨어 30 (우:01742)

(72) 발명자

파세, 찰스, 폴

미국 05763 베몬트 노쓰 치텐덴 스미쓰 팜 로드 70

웨이스, 존

미국 02476 매사추세츠 알링턴 그랜드 뷰 로드 49

(74) 대리인

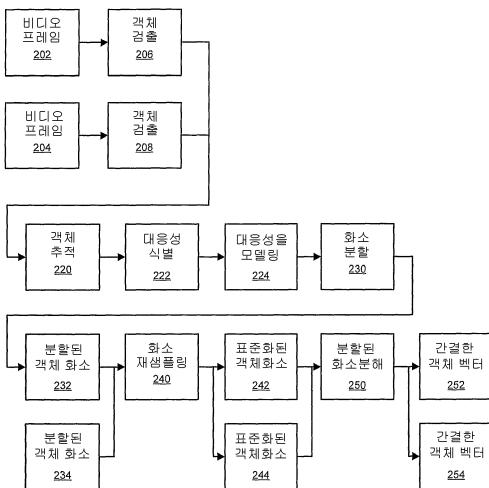
남상선

전체 청구항 수 : 총 10 항

(54) 비디오 데이터를 프로세싱하는 장치 및 방법

(57) 요 약

비디오 데이터를 프로세싱하는 장치 및 방법이 기재된다. 본 발명은 데이터와 이 데이터의 특수한 매개변수화를 위한 맞춤 모델 사이의 일치를 평가하기 위해 사용될 수 있는 비디오 데이터의 표현을 제공한다. 이에 의해, 상이한 매개변수화 기법의 비교와, 특정 데이터의 연속된 비디오 프로세싱에 대한 최적의 것의 선택을 가능하게 한다. 상기 표현은 보다 큰 프로세스의 일부로써 또는 비디오 데이터를 프로세싱하는 피드백 메커니즘으로써 중간 형태로 이용될 수 있다. 중간 형태로 이용될 때, 본 발명은 비디오 데이터의 저장, 강화, 정제, 특정 추출, 압축, 코딩, 및 전송을 위한 프로세스에서 사용될 수 있다. 본 발명은 강고하고 효율적인 방식으로 돌출 정보를 추출하도록 기능하는 한편, 비디오 데이터 소스와 일반적으로 연관된 문제를 해결한다.

대표도 - 도2

특허청구의 범위

청구항 1

다수의 비디오 프레임으로부터 비디오 신호 데이터의 인코딩된 형태를 생성하기 위한 컴퓨터 장치로서,

둘 이상의 프레임들 사이에서 상기 객체의 대응하는 엘리먼트를 식별하는 수단;

모델링된 대응성(correspondence)을 생성하도록 이러한 대응성을 모델링하는 수단;

상기 객체와 연관된 상기 비디오 프레임들내에 있는 화소 데이터를 재샘플링(resampling)하는 수단 - 상기 재샘플링 수단은 상기 모델링된 대응성을 이용함 -; 및

상기 재샘플링된 화소 데이터의 공간적 위치를 복원하는 수단 - 상기 복원 프로세스는 상기 모델링된 대응성을 이용함 -; 을 포함하며,

상기 객체는 하나 이상의 객체이며,

상기 재샘플링된 데이터는 상기 데이터의 중간 형태인,

다수의 비디오 프레임으로부터 비디오 신호 데이터의 인코딩된 형태를 생성하기 위한 컴퓨터 장치.

청구항 2

제1항에 있어서,

비디오 프레임 시퀀스에 있는 객체를 검출하는 수단; 및

비디오 프레임 시퀀스의 둘 이상의 프레임들을 통해 상기 객체를 추적하는 수단 -상기 객체 검출 수단 및 추적 수단은 비올라/존스(Viola/Jones) 페이스 검출 알고리즘을 포함함- 을 더 포함하며,

이에 의해, 상기 객체가 추적되는,

다수의 비디오 프레임으로부터 비디오 신호 데이터의 인코딩된 형태를 생성하기 위한 컴퓨터 장치.

청구항 3

제1항에 있어서,

상기 객체와 연관된 상기 화소 데이터를 상기 비디오 프레임 시퀀스내의 다른 화소 데이터로부터 분할(segmenting)하는 수단 - 상기 분할 수단은 일시적인 통합을 포함함 -; 및

원래의 비디오 프레임을 생성하도록 상기 복원된 화소를 연관된 분할 데이터와 함께 조립하는 수단을 더 포함하며,

이에 의해 상기 객체가 비디오 프레임으로부터 분할되는,

다수의 비디오 프레임으로부터 비디오 신호 데이터의 인코딩된 형태를 생성하기 위한 컴퓨터 장치.

청구항 4

제1항에 있어서,

상기 컴퓨터 장치는 대응성 측정을 글로벌 모션의 모델로 통합하는 수단을 더 포함하며,

상기 대응성 모델링 수단은 2차원의 아핀(affine) 모션 모델의 솔루션에 대한 강고한 샘플링 콘센서스(consensus)를 포함하며,

상기 대응성 모델링 수단은 상기 시퀀스에 있는 둘 이상의 비디오 프레임들 사이에서 블록 기반의 모션 추정으로부터 생성된 유한 차분에 기초한 샘플링 모집단을 포함하며,

이에 의해 상기 대응성 모델이 팩토링(factoring)되어 글로벌 모델로 되는,

다수의 비디오 프레임으로부터 비디오 신호 데이터의 인코딩된 형태를 생성하기 위한 컴퓨터 장치.

청구항 5

제1항에 있어서,

상기 정규화된 객체 화소 데이터를 인코딩된 표현으로 분해하는 수단; 및

인코딩된 표현으로부터 상기 정규화된 객체 화소 데이터를 재구성하는 수단을 더 포함하며,

상기 분해 수단은 주성분 분석(Principal Component Analysis)을 포함하며,

상기 재구성 수단은 주성분 분석을 포함하며,

이에 의해 상기 중간 데이터가 추가로 인코딩되는,

다수의 비디오 프레임으로부터 비디오 신호 데이터의 인코딩된 형태를 생성하기 위한 컴퓨터 장치.

청구항 6

제5항에 있어서,

상기 객체는, 다른 객체들이 제거될 때 상기 프레임의 잔여물, 비객체(non-object)이며,

이에 의해 상기 프레임의 비객체 화소들이 객체 화소로 모델링되는,

다수의 비디오 프레임으로부터 비디오 신호 데이터의 인코딩된 형태를 생성하기 위한 컴퓨터 장치.

청구항 7

제5항에 있어서,

표준 비디오 데이터로써의 상기 재샘플링된 화소를 종래의 비디오 압축 프로세스에 공급하는 수단; 및

모델 대응성 데이터를 대응하는 인코딩된 비디오 데이터와 함께 저장 및 전송하는 수단을 더 포함하며,

이에 의해, 상기 분할되고 재샘플링된 화소가 종래의 비디오 압축/압축해제 방법과 결합되며,

이에 의해, 상기 압축/압축해제 방법은 상기 종래의 비디오 압축 방법이 압축 효율을 증가시킬 수 있게 하는,

다수의 비디오 프레임으로부터 비디오 신호 데이터의 인코딩된 형태를 생성하기 위한 컴퓨터 장치.

청구항 8

제1항에 있어서,

상기 객체에 대응하는 화소 상부에 놓인 2차원 메시를 정의하는 수단; 및

로컬 모션의 모델로의 대응성을 측정하는 수단을 더 포함하며,

상기 대응성 측정은 상기 시퀀스에 있는 둘 이상의 비디오 프레임들 사이에서 블록 기반의 모션 추정으로부터 생성된 유한 차분에 기초한 정점 변위를 포함하며,

이에 의해, 상기 대응성 모델이 팩토링되어 로컬 변형 모델로 되는,

다수의 비디오 프레임으로부터 비디오 신호 데이터의 인코딩된 형태를 생성하기 위한 컴퓨터 장치.

청구항 9

제8항에 있어서,

상기 객체에 대응하는 현저한 이미지 특징들을 식별하는 수단을 더 포함하며,

상기 식별 수단은 이미지 경사도 해리스 응답(image gradient Harris response)의 분석이며,

이에 의해, 상기 정점들은 이산 이미지 특징(feature)에 대응하는,

다수의 비디오 프레임으로부터 비디오 신호 데이터의 인코딩된 형태를 생성하기 위한 컴퓨터 장치.

청구항 10

이산 선형 부분공간(subspace)에 존재하는 데이터 벡터를 분리(separating)하기 위한 컴퓨터 장치로서,
 데이터 벡터의 세트에 부분공간 분할(segmentation)을 수행하는 수단; 및
 암시적인 벡터 공간에 탄젠트 벡터 분석의 적용을 통해 부분공간 분할 기준을 강제하는 수단을 포함하며,
 상기 부분공간 분할 방법은 GPCA이며,
 상기 암시적인 벡터 공간은 베로네세 맵(Veronese Map)이며,
 상기 탄젠트 공간 제약은 르장드르(Legendre) 변환인,
 이산 선형 부분공간에 존재하는 데이터 벡터를 분리하기 위한 컴퓨터 장치.

명세서

기술 분야

- <1> 본원은 2005년 2월 17일 출원되고 발명의 명칭이 "System and Method For Video Compression Employing Principal Component Analysis"인 미국 출원 No. 60/653,810 및 2005년 1월 28일 출원되고 발명의 명칭이 "System and Method For Video Compression Employing Principal Component Analysis"인 미국 출원 No. 60/648,094에 우선권을 주장한다. 본원은 2005년 11월 16일 출원된 미국 출원 No. 11/280,625의 일부 계속 출원이며, 이는 2005년 9월 20일 출원된 미국 출원 No. 11/230,686의 일부 계속 출원이며, 이는 2005년 7월 28일 출원된 미국 출원 No. 11/191,562의 일부 계속 출원이다. 전술한 출원들 각각은 그 전체 내용이 본원에 참고문헌으로 포함된다.
- <2> 본 발명은 일반적으로 디지털 신호 프로세싱의 분야에 관한 것이며, 보다 구체적으로는 신호 또는 이미지 데이터의 효율적인 표현 및 프로세싱을 위한 컴퓨터 장치 및 컴퓨터 구현 방법에 관한 것이며, 가장 구체적으로 비디오 데이터의 효율적인 표현 및 프로세싱을 위한 컴퓨터 장치 및 컴퓨터 구현 방법에 관한 것이다.

배경기술

- <3> 본원 발명이 속하는 종래 기술의 일반적인 시스템 설명은 도 1에서와 같이 표현될 수 있다. 여기서, 블록 다이어그램은 통상의 종래 기술의 비디오 프로세싱 시스템을 도시한다. 이러한 시스템은 통상적으로 이하의 단계: 입력 단계(102), 프로세싱 단계(104), 출력 단계(106), 및 하나 이상의 데이터 저장 메커니즘(108)을 포함한다.
- <4> 입력 단계(102)는 카메라 센서, 카메라 센서 어레이, 영역 탐색 센서, 또는 저장 메커니즘으로부터 데이터를 검색하는 수단을 포함할 수 있다. 입력 단계는 인위적 및/또는 자연적으로 발생하는 현상의 시간 상관 시퀀스를 나타내는 비디오 데이터를 제공한다. 데이터의 돌출(salient) 컴포넌트는 잡음 또는 다른 원치 않는 신호에 의해 차폐 또는 오염될 수 있다.
- <5> 데이터 스트림, 어레이 또는 패킷 형태의 비디오 데이터는 예정된 전달 프로토콜에 따라 직접 또는 중간 저장 요소(108)를 통해 프로세싱 단계(104)로 제공될 수 있다. 프로세싱 단계(104)는 원하는 세트의 비디오 데이터 프로세싱 연산을 실행하기 위해 전용 아날로그 또는 디지털 장치, 또는 중앙처리장치(CPU), 디지털 신호 처리기(DSP), 또는 필드 프로그램 가능한 게이트 어레이(FPGA)와 같은 프로그램 가능한 장치의 형태를 가질 수도 있다. 프로세싱 단계(104)는 통상적으로 하나 이상의 CODEC(코더/디코더)을 포함한다.
- <6> 출력 단계(106)는 신호, 디스플레이, 또는 사용자 또는 외부 장치에 영향을 줄 수 있는 다른 응답을 생성한다. 통상적으로, 출력 장치는 지시자 신호, 디스플레이, 하드카피, 저장소내의 프로세싱된 데이터의 표현을 생성하기 위해 사용되거나, 원격 사이트로 데이터의 전송을 개시하기 위해 사용된다. 이는 이후의 프로세싱 연산에서 사용하기 위한 중간 신호 또는 제어 파라미터를 제공하기 위해 사용될 수도 있다.
- <7> 저장소는 이러한 시스템에서 선택적인 엘리먼트로 제공된다. 사용될 때, 저장소 엘리먼트(108)는 판독 전용 저장 매체와 같이 비휘발성이거나, 동적 랜덤 액세스 메모리(RAM)와 같이 휘발성일 수도 있다. 단일 비디오 프로세싱 시스템이, 입력, 프로세싱 및 출력 단계와 다양한 관계를 갖는 여러 형태의 저장소 엘리먼트를 포함하는 것이 이상한 것은 아니다. 이러한 저장 엘리먼트의 예는 입력 버퍼, 출력 버퍼, 및 프로세싱 캐시를 포함한다.
- <8> 도1의 비디오 프로세싱 시스템의 주목적은 입력 데이터를 프로세싱하여 특정 애플리케이션에서 의미 있는 출력을 생성하는 것이다. 이러한 목표를 달성하기 위해, 잡음 감소 또는 제거, 특성 추출, 객체 분할

(segmentation) 및/또는 정규화, 데이터 카테고리화, 이벤트 검출, 편집, 데이터 선택, 데이터 재-코딩, 및 트랜스코딩을 포함하는 다양한 프로세싱 연산이 사용될 수도 있다.

- <9> 불완전하게 제한된 데이터를 생성하는 많은 데이터 소스는 인간에게, 특히 음향 및 시각 이미지에 중요하다. 대부분의 경우, 이러한 소스 신호들의 본질적인 특징은 효율적인 데이터 프로세싱이라는 목적으로 악영향을 미친다. 소스 데이터의 고유한 변화가능성은 공학적 가설을 유도하는데 사용되는 순수한 경험 및 발견적 방법으로부터 발생하는 예리를 도입하지 않고 신뢰가능하고 효율적인 방법으로 데이터를 프로세싱하는데 장애가 된다. 입력 데이터가 자연히 또는 고의로 좁게 정의된 특징 세트(가령, 심볼 값들 또는 좁은 대역폭의 제한된 세트)로 제한된 경우, 애플리케이션들에 대해 이러한 변화가능성은 감소된다. 이러한 제한들 모두에 의해, 낮은 상업적 가치를 갖는 프로세싱 기술이 아주 빈번히 생긴다.
- <10> 단일 프로세싱 시스템의 설계는 입력으로 사용된 소스 신호의 예상된 특성 및 시스템의 의도된 사용에 의해 영향을 받는다. 대부분의 경우, 요구되는 성능 효율은 중요한 설계 인자일 것이다. 순차적으로, 성능 효율은, 이용가능한 데이터 저장소와 대비한 프로세싱될 데이터량과, 이용가능한 계산력과 대비한 어플리케이션의 계산 복잡도에 의해 영향을 받는다.
- <11> 통상의 비디오 프로세싱 방법은, 느린 데이터 통신 속도, 큰 저장소의 필요성, 및 교란시키는 의식적 인공물의 형태로 나타나는 다수의 비효율을 겪는다. 이들은, 사람들이 비디오 데이터를 이용하고 조작하기 위해 원하는 다양한 방식으로 인해, 그리고 사람들이 시각적 정보의 소정 형태에 대해 갖는 선천적 민감도로 인해, 심각한 문제가 될 수 있다.
- <12> "최적의" 비디오 프로세싱 시스템은 원하는 세트의 프로세싱 연산을 실행하는데 있어서 효율적이고, 신뢰가능하고 강건하다. 이러한 연산은 저장, 송신, 디스플레이, 압축, 편집, 암호화, 강화(enhancement), 카테고리화, 특성 검출, 및 데이터의 인식을 포함한다. 부차적인 연산은 다른 정보 소스와 이러한 프로세싱된 데이터의 통합을 포함한다. 마찬가지로 중요한 것은, 비디오 프로세싱 시스템의 경우, 출력이 의식적 인공물의 도입을 방지함으로써 인간 시각과 호환가능해야 한다는 것이다.
- <13> 비디오 프로세싱 시스템은 그 속도, 효율, 및 품질이 입력 데이터의 임의의 특정 특성의 세부사항에 의존하지 않으면, "강고한(robust)" 것으로 표현될 수 있다. 강고성은 또한 입력의 일부가 오류인 때 연산을 실행할 수 있는 능력과 관련된다. 많은 비디오 프로세싱 시스템은 일반적 부류의 애플리케이션에 대해 허용할 만큼 충분히 강고하지 않다 - 시스템에 개발에 사용되었던 좁게 한정된 동일 데이터에 대한 적용만을 제공한다.
- <14> 돌출한 정보(salient information)는, 센싱된 현상의 신호 특성과 매칭되지 않는 입력 엘리먼트의 샘플링 레이트로 인해, 연속한 값의 데이터 소스를 이산화할 때 손실될 수 있다. 또한, 신호의 강도가 센서의 한계를 초과하여 포화될 때 손실이 존재한다. 유사하게, 입력 데이터의 정밀도가 감소된 경우 정보는 손실되는데, 이는 입력 데이터의 전체 범위의 값이 이산 값의 세트에 의해 표현된 경우 모든 양자화 프로세스에서 발생하며, 이로 인해 데이터 표현의 정확성이 감소된다.
- <15> 총 변화가능성(ensemble variability)은 데이터 또는 정보 소스의 부류내의 모든 예측불가성을 말한다. 시각 정보의 데이터 표현은 시각 정보가 통상적으로 제한되지 않기 때문에 매우 큰 총 변화가능성 등급을 갖는다. 시각 데이터는 센서 어레이에 입사하는 빛에 의해 형성될 수 있는 모든 공간 어레이 시퀀스 또는 공간-시간 시퀀스를 표현할 수 있다.
- <16> 시각 현상의 모델링에서, 비디오 프로세서는 통상적으로 데이터가 표현되거나 변역되는 방식에 대해 소정 세트의 제한 및/또는 구조를 부과한다. 그 결과, 이러한 방법은, 출력의 품질; 출력에 대해 부여되는 신뢰; 및 데이터에 대해 신뢰성 있게 실행될 수 있는 이후의 처리 과제의 유형에; 영향을 주는 체계적인 오류를 도입할 수 있다.
- <17> 양자화 방법은 비디오 프레임 내의 데이터의 정확도를 감소시키지만 한편으로는 데이터의 통계적 변화를 유지하려고 한다. 통상적으로, 비디오 데이터가 분석되면 데이터 값들의 분포는 확률 분포에 모인다. 데이터를 공간 주파수의 혼합으로 특징지우기 위해 데이터를 위상 공간으로 투영함으로써, 거부감이 덜 한 방법으로 정확도 감소를 방지시키는 방법도 있다. 심하게 사용될 경우, 이러한 양자화 방법은 지각적으로 받아들이기 어려운 색상을 생성하고 비디오 프레임의 원래 평탄한 영역에서 급격한 픽셀레이션(pixilation)을 일으킬 수 있다.
- <18> 차등(differential) 코딩이 또한 통상적으로 데이터의 로컬 공간 유사성을 이용하기 위해 사용된다. 프레임의 일부의 데이터는 그 프레임내의 유사 데이터 부근에, 그리고 또한 후속 프레임내의 유사 위치에 클러스터링되는 경향이 있다. 데이터를 공간적으로 이와 인접한 데이터에 관하여 표현하는 것은 이후 양자화와 결합될 수 있으

며, 그 최종 결과는, 주어진 정밀도에 대해 그 차이를 나타내는 것이 데이터의 절대값을 이용하는 것보다 더욱 정확하다는 것이다. 이러한 가정은 흑백 비디오 또는 낮은 컬러 비디오와 같이, 원본 비디오 데이터의 스펙트럼 해상도가 제한된 때 잘 들어맞는다. 비디오의 스펙트럼 해상도가 증가함에 따라, 유사성의 가정은 현저히 어긋난다. 이러한 어긋남은 비디오 데이터의 정밀도를 선택적으로 보존할 수 없음에 기인한다.

<19> 나머지 코딩은, 표현의 에러가 원본 데이터의 정밀도를 원하는 레벨의 정확성까지 복원하기 위해 추가로 차등 엔코딩된다는 점에서, 차등 엔코딩과 유사하다.

<20> 이러한 방법의 변형들은, 비디오 데이터를, 공간 위상 및 스케일에 있어서의 데이터 상관관계를 노출시키는 대안의 표현으로 변환시키려고 한다. 일단 비디오 데이터가 이러한 방식으로 변환되었다면, 그 후 양자화 및 차등 코딩 방법이 변환된 데이터에 적용될 수 있어서, 돌출 이미지 특성의 보존을 증가시킨다. 널리 보급된 이러한 변환 비디오 압축 기술 중 두 개는 이산 코사인 변환(DCT) 및 이산 웨이블릿 변환(DWT)이다. DCT 변환에서의 에러는 비디오 데이터 값의 넓은 변화에서 나타나며, 따라서, DCT는 통상적으로 이러한 잘못된 상관을 국부화시키기 위해 비디오 데이터의 블록에 대해 일반적으로 사용된다. 이러한 국부화로부터의 인공물은 종종 블록의 경계를 따라 나타난다. DWT의 경우, 기본 기능과 소정의 텍스쳐 사이에 부정합이 존재할 때 더욱 복잡한 인공물이 발생하며, 이는 흐려짐(blurring) 효과를 일으킨다. DCT 및 DWT의 부정적 영향에 대항하기 위해, 표현의 정밀도가 정밀한 대역폭을 희생하고 낮은 왜곡으로 증가된다.

발명의 상세한 설명

<21> 본 발명은 현존하는 최신의 방법에 대해 계산적 이점 및 분석적 이점 모두를 제공하는 컴퓨터로 구현되는 비디오 프로세싱 방법이다. 주된 본 발명의 방법은 선형 분해 방법, 공간 분할 방법, 및 공간 정규화 방법의 통합이다. 공간적으로 제약된 비디오 데이터는 선형 분해 방법의 강고성 및 적용가능성을 크게 증가시킨다. 부가적으로, 공간적 정규화에 대응하는 데이터의 공간적 분할 공간 정규화 자체로부터 유도된 이점을 증가시키기 위하여 추가로 기능할 수 있다.

<22> 특히, 본 발명은 신호 데이터를 하나 이상의 유용한 표현으로 효율적으로 프로세싱할 수 있는 수단을 제공한다. 본 발명은 많은 공통적으로 발생하는 데이터 세트를 프로세싱하는 데 효율적이며 비디오 및 이미지 데이터를 프로세싱하는 데 특히 효율적이다. 본 발명의 방법은 데이터를 분석하며 프로세싱 및 인코딩을 용이하게 하기 위하여 상기 데이터의 하나 이상의 간결한 표현을 제공한다. 각각의 새롭고, 보다 간결한 데이터 표현은 계산 프로세싱과, 전송 대역폭과, 많은 어플리케이션(비디오 데이터의 인코딩, 압축, 전송, 분석, 저장 및 디스플레이를 포함하나 이에 제한되지 않음)에 대한 저장 요건에 있어서의 감소를 가능하게 한다. 본 발명은 비디오 데이터의 돌출 요소들의 식별 및 추출을 위한 방법을 포함하며, 데이터의 프로세싱 및 표현에 있어서의 우선순위 결정을 가능하게 한다. 노이즈 및 다른 원치하는 신호의 부분은 낮은 우선도로 식별되므로, 추가적인 프로세싱은 비디오 신호의 보다 높은 우선도 부분들을 분석하고 표현하는 데 집중될 수 있다. 결과로써, 비디오 신호는 이전에 가능했던 것 보다 훨씬 간결하게 표현된다. 그리고 정밀도에 있어서의 손실은 개념적으로 중요치 않은 비디오 신호의 일부분에 집중된다.

실시예

<36> 비디오 신호 데이터에서, 비디오의 프레임들은, 투사되어 이미지화 된 3차원 장면을 2차원 이미지화 표면상에 일반적으로 도시하는 이미지의 시퀀스로 조립된다. 각각의 프레임 또는 이미지는 샘플링된 신호에 대한 이미지 센서 응답을 나타내는 화소(pixel)로 구성된다. 종종, 샘플링된 신호는, 반사, 굴절 또는 방사되고 2차원 센서 어레이에 의해 샘플링된 전자기 에너지(가령 전자기, 음향 등)에 대응한다. 연속한 순차적 샘플링에 의해, 프레임 당 두 개의 공간 차원 및 비디오 시퀀스에서 프레임의 순서에 대응하는 시간 차원을 갖는 시공간적 데이터 스트림을 발생된다.

<37> 도2에 도시된 바와 같이, 본 발명은 신호 데이터를 분석하고, 돌출 컴포넌트를 식별한다. 신호가 비디오 데이터로 구성될 때, 시공간적 스트림은 얼굴과 같은 종종 특정한 객체인 돌출 컴포넌트를 나타낸다. 식별 프로세스는 돌출 컴포넌트의 존재 및 중요성을 정량화하고, 이를 정량화된 돌출 컴포넌트 중 하나 이상의 가장 중요한 것을 선택한다. 이것은 현재 설명된 프로세싱 이후에 또는 이와 동시에 다른 덜 돌출된 컴포넌트를 식별 및 프로세싱하는 것을 제한하는 것은 아니다. 전술한 돌출 컴포넌트가 이어 추가로 분석되어 가변 및 불변의 서브 컴포넌트를 식별한다. 불변 서브 컴포넌트의 식별은 컴포넌트의 소정 특성을 모델링하는 프로세스이며, 상기 모델링에 의해, 컴포넌트가 원하는 정확도 레벨까지 동기화되게 하는 모델의 파라미터화를 나타낸다.

<38> 일 실시예에서, 전방(foreground) 객체가 검출 및 추적된다. 객체의 화소는 비디오의 각각의 프레임으로부터 식별 및 분할된다. 블록 기반 모션 추정이 다수의 프레임내의 분할된 객체에 적용된다. 이어 이러한 모션 추정은 더 높은 계층의 모션 모델로 통합된다. 모션 모델은 공통 공간 구성에 대한 객체의 인스턴스(instance)를 뒤틀기(warp)하는데 사용된다. 소정의 데이터의 경우, 이러한 구성에서, 객체의 더 많은 특성이 정렬된다. 이러한 정규화는 다수의 프레임에 대한 객체의 화소의 값의 선형 분해가 컴팩트하게 표현될 수 있게 해준다. 객체의 외관에 속하는 돌출 정보는 이러한 컴팩트한 표현에 포함된다.

<39> 본 발명의 바람직한 실시예는 전방 비디오 객체의 선형 분해를 상술한다. 객체는 공간적으로 정규화됨으로써, 컴팩트한 선형 외관 모델을 생성한다. 또 다른 바람직한 실시예는 공간 정규화에 앞서 비디오 프레임의 후방으로부터 전방 객체를 추가로 분할한다.

<40> 본 발명의 바람직한 실시예는 소량의 모션을 실행하면서 카메라에 이야기하는 사람의 비디오에 본 발명을 적용한다.

<41> 본 발명의 바람직한 실시예는 공간 변환을 통해 잘 표현될 수 있는 비디오내의 소정의 객체에 본 발명을 적용한다.

<42> 본 발명의 바람직한 실시예는 비디오의 둘 이상의 프레임들 간의 한정된 차이점을 결정하기 위해 블록 기반 모션 추정을 이용한다. 더욱 효율적인 선형 분해를 제공하기 위해 더 높은 등급의 모션 모델이 상기 한정된 차이로부터 인수화(factoring)된다.

검출 및 추적

<44> 일단 신호의 구성 돌출 컴포넌트가 결정되면, 이러한 컴포넌트들은 제한될 수도 있으며, 모든 다른 신호 컴포넌트는 감소되거나 제거될 수도 있다. 돌출 컴포넌트를 검출하는 프로세스는 도 2에 도시되어 있으며, 비디오 프레임(202)은 하나 이상의 객체 검출(206) 프로세스에 의해 처리되어, 하나 이상의 객체는 식별되고, 이후 추적된다. 제한된 컴포넌트는 비디오 데이터의 중간 형태를 나타낸다. 이어 이러한 중간 데이터는 현재의 비디오 프로세싱 방법에 통상적으로 이용가능하지 않은 기술을 이용하여 엔코딩될 수 있다. 중간 데이터가 여러 형태로 존재하기 때문에, 표준 비디오 엔코딩 기술이 이러한 여러 중간 형태를 엔코딩하도록 사용될 수도 있다. 각각의 예에 대해, 본 발명은 가장 효율적인 엔코딩 기술을 결정하고 이용한다.

<45> 바람직한 일 실시예에서, 돌출 분석 프로세스는 돌출 신호 모드를 검출 및 분류한다. 이러한 프로세스의 일 실시예는 응답 신호를 생성하기 위해 구체적으로 지정된 공간 필터들의 조합을 이용하는데, 이 응답 신호의 강도는 비디오 프레임 내에 있는 객체의 검출된 돌출성과 관련이 있다. 분류기는 상이한 공간 스케일에서 그리고 상이한 비디오 프레임의 위치에서 적용된다. 분류기로부터의 응답의 강도는 돌출 신호 모드의 존재의 가능성을 나타낸다. 현저한 돌출 객체가 중심에 있을 경우, 프로세스는 이와 부합하여 이 객체를 강한 응답으로 분류한다. 돌출 신호 모드의 검출은 비디오 시퀀스에서의 돌출 정보에 대한 이후의 프로세싱 및 분석을 가능하게 함으로써 본 발명을 차별화한다.

<46> 하나 이상의 비디오 프레임에서 돌출 신호 모드의 검출 위치가 주어지면, 본 발명은 돌출 신호 모드의 불변 특성을 분석한다. 부가적으로, 본 발명은 불변 특성에 대해 "덜 돌출된" 신호 모드인 나머지 신호를 분석한다. 불변 특성의 식별은 중복되는 정보를 감소시키고 신호 모드를 분할(즉, 분리)하기 위한 기반을 제공한다.

특성 포인트 추적

<48> 본 발명의 일 실시예에서, 하나 이상의 프레임의 공간 위치는 공간 강도 필드 경사도(gradients) 분석을 통해 결정된다. 이러한 특징은 "라인들"의 몇 개의 교차부에 대응하며, 상기 라인들은 느슨하게는 "코너"라고 기재될 수 있다. 이러한 실시예는, 강한 코너들인 동시에 서로와 공간적으로 분리되어 있는 코너들의 세트를 선택하는데, 본원에서는 이 코너들의 세트를 특성 포인트(feature point)라고 한다. 또한, 광학 플로우의 계층적 다해상도 추정을 이용하면 시간에 따른 특성 포인트의 변환 변위의 결정이 가능하다.

<49> 도 2에서, 객체 검출 프로세스(208)로부터 검출 인스턴스를 도출하는 객체 추적(220) 프로세스가 도시되어 있고, 추가로 다수의 비디오 프레임들(202 및 204)에 대하여 하나 이상의 검출된 객체들의 특성들의 대응성을 식별하는 프로세스(222)가 도시되어 있다.

<50> 특성 추적의 비제한적 실시예가 사용될 수 있어서, 특성은 블록 기반 모션 추정과 같은 더욱 일정한 경사도 분석 방법을 자격 검증하는데 이용될 수 있다.

<51> 다른 실시예는 특성 추적에 기초한 모션 추정의 예측을 예상한다.

객체 기반 검출 및 추적

<53> 본원 발명의 비제한적 실시예에서, 강고한 객체 분류기가 비디오의 프레임에서 얼굴들을 추적하기 위해 사용된다. 이러한 분류기는 얼굴들에 대해 트레이닝된 지향된 에지에 대한 직렬 응답에 기초한다. 이러한 분류기에서, 에지는 기본 하르(Haar) 특성 및 이를 특성을 45도 만큼 회전한 것의 세트로서 정의된다. 직렬 분류기는 아다부스트(AdaBoost) 알고리즘의 변형이다. 게다가, 응답 계산은 합산된 영역 테이블의 사용을 통해 최적화될 수 있다.

로컬 등록

<55> 등록은 두 개 이상의 비디오 프레임에서 식별된 객체의 엘리먼트들 사이의 대응의 할당을 포함한다. 이러한 대응은 비디오 데이터에서 시간적으로 구별되는 포인트에서 비디오 데이터 사이의 공간적 관계를 모델링하는데 기초가 된다.

<56> 다양한 제한되지 않는 등록 수단이, 주지의 알고리즘 및 이를 알고리즘의 파생 알고리즘에 관하여 실행하기 위한 특정 실시예 및 이들의 관련된 감소를 설명하기 위해 본 발명에 대해 기술된다.

<57> 시공간적 시퀀스에서 명백한 광학 플로우를 모델링하는 한 가지 수단은 비디오 데이터의 두 개 이상의 프레임으로부터의 유한 차분 필드의 생성을 통해 달성될 수 있다. 대응성이 공간 및 강도 의미에서 특정의 불변성 제약을 따르면 광학 플로우 필드는 개략적으로 추정될 수 있다.

<58> 도 3에 도시된 것처럼, 프레임(302 또는 304)은 데시메이션 프로세스(306) 또는 일부 다른 서브 샘플링 프로세스(가령, 저역필터)를 통해 공간적으로 서브 샘플링된다. 이러한 공간적으로 감소된 이미지(310 및 312)는 또한 추가로 서브 샘플링될 수 있다.

다이아몬드 검색

<60> 비디오 프레임의 블록으로의 비중첩 분할이 주어지면, 각각의 블록에 대한 정합(match)을 위하여 비디오의 이전 프레임을 탐색한다. 전체 검색 블록 기반(FSBB) 블록 모션 추정은, 현재 프레임에 있는 블록과 비교될 때 가장 적은 에러를 갖는 비디오의 이전 프레임의 위치를 탐색한다. FSBB를 실행하는 것은 계산적으로 매우 고가이며, 종종 로컬화된 모션의 가정에 기초한 다른 모션 추정 방식에 비해 양호한 정합을 도출하지 않는다. 다이아몬드 검색 블록 기반(DSBB) 경사도 하강 모션 추정은 블록에 대한 최상의 정합을 향한 에러 경사도를 반복적으로 트래버스법으로 측정(traverse)하기 위해 다양한 크기의 다이아몬드형 검색 패턴을 이용하는 FSBB에 대한 일반적인 대안이다.

<61> 본 발명의 일 실시예에서, DSBB는 자신의 값이 더 높은 등급의 모션 모델로 이후에 분해되는 유한 차분을 생성하기 위해 비디오의 하나 이상의 프레임들 사이에서 이미지 경사도 필드의 분석에서 사용된다.

<62> 기술 분야의 당업자는 블록 기반 모션 추정이 규칙적인 메시의 정점들의 분석의 등가로서 보여질 수 있다는 것을 알 것이다.

메시(mesh) 기반 모션 추정

<64> 메시 기반 예측은 비디오 프레임의 불연속(discrete) 영역들을 묘사하기 위하여 에지들에 의해 연결된 정점들의 기하학적 메시를 이용한 후, 이후에 메시 정점들의 위치에 의해 제어되는 변형 모델을 통해 후속 프레임들에 있는 이들 영역의 변형과 이동을 예측한다. 정점들이 이동될 때, 정점들에 의해 정의된 영역내의 화소들은 현재의 프레임을 예측하기 위하여 또한 이동된다. 최초 화소 값의 상대적인 이동 및 결과적인 근사는 화소 위치를 그 화소의 주변에 있는 정점들의 위치와 연관시키는 몇몇 보간법을 통해 수행된다. 순수한 변환과 비교하여 스케일링 및 회전의 부가적인 모델링은, 이러한 동작이 비디오 신호내에 존재할 때, 프레임의 화소들의 보다 정밀한 예측을 생성할 수 있다.

<65> 일반적으로 메시 모델은 규칙적인 것 또는 적응성이 있는 것으로 정의될 수 있다. 규칙적인 메시 모델은 기초가 되는 신호 특성을 고려하지 않고 설계되는 반면, 적응성이 있는 방법은 정점들과 에지들을 기초가 되는 비디오 신호의 특성들과 관련하여 공간적으로 배치하고자 시도한다.

<66> 비디오 내의 이미지화된 객체가 메시내의 에지들과 잘 대응하는 공간적 불연속성을 갖는다고 가정하면, 규칙적인 메시 표현은, 모션 또는 균등하게는 모션 내의 고유한 변형이 예측되거나 모델링될 수 있는 수단을

제공한다.

- <67> 적응성이 있는 메시는 규칙적인 메시보다 기본이 되는 비디오 신호의 특성에 대해 실질적으로 더 많은 고려를 하여 형성된다. 부가적으로, 이러한 메시의 적응성이 있는 본성에 의해 시간에 따른 메시의 다양한 정체가 가능하다.
- <68> 본 발명은 메시, 및 균등하게는 화소 등록을 수행하기 위하여, 동질성 기준을 이용하여 정점 검색 정렬을 조정 한다. 이종 강도 경사도와 공간적으로 연관되어 있는 정점들은 보다 동질성이 있는 경사도를 갖는 것 이전에 추정되는 모션이다.
- <69> 바람직한 실시예에서, 메시의 정점 모션 추정은 동일하거나 거의 동일한 동질성을 갖는 정점들에 대한 모션 추정의 공간적 플러드-충진(flood-filling)을 통해 부가적으로 우선시된다.
- <70> 바람직한 실시예에서, 최초의 메시 공간 구조 및 최종 메시 구조는 표준 그래픽 충진 루틴을 이용하여 매핑 이미지를 면 식별자들로 충진함으로써 면 레벨에서 서로에 대해 매핑된다. 각각의 삼각형과 연관된 아핀(affine) 변환은 변환 테이블에서 빠르게 찾을 수 있고 하나의 메시내의 면과 연관된 화소 위치들은 다른 메시내의 위치로 빠르게 변환될 수 있다.
- <71> 바람직한 실시예에서, 각각의 모션 추정 정합과 연관된 나머지 오차를 추정하기 위하여, 예비적인 모션 추정이 정점들에 대해 이루어진다. 이러한 예비적 추정은 정점들의 모션 추정 순서를 우선시하기 위하여 부가적으로 사용된다. 이러한 나머지 오차 분석의 이점은 더 적은 왜곡과 연관된 모션 추정은 결과적으로 보다 개연성이 있는 메시 토폴로지(topology)를 유지하게 될 것이라는 점이다.
- <72> 바람직한 실시예에서, 메시 정점 모션 추정은 몇몇 제한된 범위로 스케일 다운되며, 메시가 보다 전역적으로 (globally) 최적이고 토폴로지 면에서 정확한 솔루션에 접근할 수 있도록 다수의 모션 추정이 여러 번의 반복을 통해 이루어진다.
- <73> 바람직한 실시예에서, 각각의 정점에 중심이 있는 사각형의 타일 근접성을 이용하는 블록 기반 모션 추정이 사용되어 보간된 다각형 근접성을 고려하여 정점 변위를 결정한다. 공간 보간을 회피하는 것과 오차 경사도 감소를 위한 화소의 뒤틀림에 부가하여, 이 기법도 또한 모션 추정의 병렬 계산을 가능하게 한다.
- <74> **위상 기반 모션 추정**
- <75> 종래 기술에서, 블록 기반 모션 추정은 통상적으로 하나 이상의 공간 매치에서 초래하는 공간 탐색으로서 통상적으로 실행되었다. 도3에 도시된 바와 같이, 위상 기반 정규화 크로스 상관(PNCC)은 현재 프레임 및 이전 프레임으로부터의 블록을 "위상 공간"으로 변환하고 이를 두 블록들의 크로스 상관을 탐색한다. 크로스 상관은 자신의 위치가 두 블록들 사이의 에지의 "위상 시프트"에 대응하는 값들의 필드로서 표현된다. 이러한 위치들은 임계를 통해 분리되고, 그로 인해 공간 좌표로 역변환된다. 공간 좌표는 개별 에지 변위이며, 모션 벡터에 대응한다.
- <76> PNCC의 장점은 비디오 스트림에서 개인/노출 조정의 공차를 가능하게 하는 콘트라스트 마스킹을 포함한다. 또한, PNCC는 공간적으로 기반된 모션 추정으로부터 많은 반복을 취할 수도 있는 단일 단계로부터 경과를 허용한다. 게다가, 모션 추정은 서브 픽셀 정확도이다.
- <77> 본 발명의 일 실시예는 자신의 값이 더 높은 등급의 모션 모델로 이후에 분해되는 유한 차분을 생성하기 위해 비디오의 하나 이상의 프레임들 사이의 이미지 경사도 필드의 분석에서 PNCC를 이용한다.
- <78> **글로벌 등록**
- <79> 일 실시예에서, 본 발명은 유한 차분 추정의 필드로부터 하나 이상의 선형 모델을 분해한다. 이러한 샘플링이 발생하는 필드는 유한 차분의 일반적 모집단으로서 언급된다. 설명된 방법은 RANSAC 알고리즘의 추정과 유사한 강고한 추정을 이용한다.
- <80> 도 4에 도시된 것처럼, 글로벌 모션 모델링의 경우에, 유한 차분은 일반적인 모집단 풀(404)로 수집되는 변형 모션 추정(402)이며, 상기 풀(402)은 이러한 모션 추정의 랜덤 샘플링(410)에 의해 반복적으로 프로세싱되며 선형 모델은 이를 샘플들로부터 학습된다(420). 그 후 결과는 모집단(404)을 조정하기 위해 사용되어 랜덤 프로세스를 통해 찾을 때, 모델에 대한 이상점(outlier)의 배제를 통해 선형 모델을 더 잘 분류한다.
- <81> 선형 모델 추정 알고리즘의 일 실시예에서, 모션 모델 추정기는 선형의 최소 제곱 해(least square solution)를

기반으로 한다. 이러한 의존성에 의해, 추정기는 이상점 데이터에 의해 떨어져 나간다. RANSAC에 기초하면, 개시된 방법은 데이터의 부분집합의 반복적인 추정을 통해 이상점의 효과에 역행하는 간접한 방법이며, 데이터의 중요한 부분집합을 설명할 모션 모델을 탐색한다. 각각의 탐색에 의해 생성된 모델은 이 모델이 표현하는 데이터의 백분율에 대해 테스팅된다. 만일 충분한 수의 반복이 있다면, 데이터의 가장 큰 부분집합과 맞는 모델이 발견될 것이다.

<82> 도 4에서 이해되고 설명된 바와 같이, 본 발명은 유한 차분(샘플)의 초기 샘플링 및 선형 모델의 최소 제곱 추정을 포함하는 알고리즘의 변형의 형태인 RANSAC 알고리즘을 초과하는 혁신을 개시한다. 종합 예라는 해결된 선형 모델을 이용하여 일반 모집단에서 모든 샘플에 대해 평가된다. 랭크는 자신의 나머지가 예정된 임계치에 일치하는 샘플의 수에 기초하여 선형 모델에 할당되며, 이러한 랭크는 "후보 컨센서스"로 고려된다.

<83> 초기 샘플링, 해법, 및 랭킹은 반복적으로 결과 표준이 충족될 때까지 실행된다. 일단 표준이 충족되면, 최고의 랭크를 갖는 선형 모델은 모집단의 최종 컨센서스로 고려된다.

<84> 옵션 정밀화 단계는 후보 모델에 대한 최상의 적합 등급에서 샘플들의 서브세트를 반복적으로 분석하는 단계, 및 하나 이상의 샘플 부가가 전체 서브세트에 대한 나머지 예러 임계치를 초과할 때까지 서브세트의 크기를 증가시키는 단계를 포함한다.

<85> 도 4에 도시된 것처럼, 글로벌 모델 추정 프로세스(450)는 컨센서스 랭크 수용성 테스트가 만족(452)될 때까지 반복된다. 랭크가 얻어지지 않았을 때, 유한 차분(404)의 모집단은 선형 모델을 보이기 위한 노력으로 발견된 모델에 대하여 분류된다. 최고의(가장 높은 랭크) 모션 모델이 프로세스(460)에서 해집합에 부가된다. 그 후, 모델이 프로세스(470)에서 재추정된다. 완결시에, 모집단(404)은 다시 정렬된다.

<86> 본 발명의 설명된 제한적이지 않은 실시예는 특정 선형 모델에 대응하는 다른 파라미터 벡터 공간에서 서브 공간 메니폴드를 결정하기 위해 유한 차분 벡터의 필드로서 전술된 벡터 공간을 샘플링하는 일반적 방법으로서 추가로 일반화된다.

<87> 광역 등록 프로세스의 추가적인 결과는, 광역 등록 프로세스와 로컬 블록 프로세스 사이의 차가 로컬 등록 나머지를 초래한다는 것이다. 이러한 나머지는 로컬 모델을 근사화하는데 있어서 광역 모델의 예러이다.

정규화

<89> 정규화는 표준, 또는 공통의 공간 구성을 향한 공간 강도 필드의 재샘플링을 의미한다. 이러한 상대적인 공간 구성이 상기 구성들 사이에서 반전가능한 공간 변화일 때, 재샘플링 및 화소의 동반하는 보간은 토폴로지 한계 까지 반전가능하다. 본 발명의 정규화 방법은 도 5에 개시된다.

<90> 두 개 이상의 공간 강도 필드가 정규화될 때, 증가된 계산 효율은 중간 정규화 계산을 유지함으로써 달성될 수도 있다.

<91> 등록을 위해, 또는 등가적으로 정규화를 위해 이미지를 샘플링하는데 사용된 공간 변환 모델은 광역 및 로컬 모델을 포함한다. 광역 모델은 변환으로부터 투영으로 등급을 증가시킨다. 로컬 모델은 블록 또는 더욱 복잡하게 구분적으로 선형 메시에 의해 기본적으로 결정된 이웃한 화소에 대한 보간을 적용하는 유한 차분이다.

<92> 정규화된 강도 필드에 대한 원본 강도 필드의 보간은 강도 필드의 서브세트에 기초한 PCA 출현 모델의 선형성을 증가시킨다.

<93> 도 2에 도시된 것처럼, 객체 화소(232 & 234)는 객체 화소(242 & 244)의 정규화된 버전을 생성하도록 다시 샘플링(240)될 수 있다.

메시 기반 정규화

<95> 본 발명의 추가적인 실시예는 특징점을 바둑판 모양으로 하여 삼각형 기반의 메시로 만들고, 메시의 정점들이 추적되고, 각각의 삼각형의 정점들의 상대적인 위치는 이러한 세 개의 정점들과 일치하는 평면에 대해 수직인 3 차원 표면을 추정하기 위해 사용된다. 수직한 표면이 카메라의 튀어나온 축과 일치할 때, 이미지화된 삼각형에 대응하는 객체의 최소 변형된 렌더링을 제공할 수 있다. 수직인 직교 표면을 선호하는 경향이 있는 정규화된 이미지를 생성하는 것은 이후의 외관 기반의 PCA 모델의 선형성을 증가시키는 중간 데이터 형태를 보존하는 화소를 생성할 수 있다.

<96> 또 다른 실시예는 글로벌 모션 모델을 암시적으로 모델링하기 위하여 통상적인 블록 기반 모션 추정을

이용한다. 하나의 비 제한적인 실시예에서, 본 방법은 통상적인 블록 기반 모션 추정/예측에 의해 기술된 모션 벡터들로부터 글로벌 유사(affine) 모션 모델을 팩토링한다.

<97> 도 9는 글로벌 및 로컬 정규화를 결합하는 방법을 도시한다.

점진적 기하구조 정규화

<99> 공간적 불연속의 분류가 사용되어 불연속이 메시 에지들과 일치할 때 불연속을 암시적으로 모델링하기 위하여 테셀레이션 된(tessellation) 메시를 정렬한다.

<100> 동종의 영역 경계는 다각형 케적에 의해 근사화된다. 케적은 각각의 다각형 정점의 돌출 우선도(saliency priority)를 결정하기 위하여 연속적으로 더 낮은 정밀도로 연속적으로 근사화된다. 정점 우선도는 공유된 정점들에 대한 정점 우선도를 보존하기 위하여 영역들을 가로질러 전파된다.

<101> 본 발명의 일 실시예에서, 다각형 분해 방법은 필드의 동종 분류와 연관된 경계들의 우선화를 가능하게 한다. 화소들은 스펙트럼 유사성과 같은 동종성 기준에 따라 분류된 후, 분류 라벨이 영역들로 공간적으로 연결된다. 추가적인 바람직한 비제한적인 실시예에서, 4- 또는 8- 연결성(connectedness) 기준이 공간 연결성을 결정하기 위해 적용된다.

<102> 바람직한 실시예에서, 이들 공간 영역들의 경계는 그 후 다각형으로 이산화(discretization)된다. 모든 동종의 영역들에 대한 모든 다각형의 공간적 덧씌움은 그 후 예비적인 메시로 테셀레이션되고 결합된다. 이 메시의 정점들은 여러 기준을 이용하여 분해되어, 원래의 메시의 많은 지각력있는 돌출성을 유지하는 보다 간단한 메시 표현을 생성한다.

<103> 바람직한 실시예에서, 본 명세서의 다른 부분에 개시되는 이미지 등록 방법은 강한 이미지 경사도(gradients)를 이용하여 이러한 높은 우선도의 정점들을 향해 편향된다. 결과적인 변형 모델은 이미지화된 객체의 기하구조와 연관된 공간적 불연속을 유지하는 경향이 있다.

<104> 바람직한 실시예에서, 능동 케적(active contour)이 영역 경계를 정제하기 위해 사용된다. 각각의 다각형에 대한 능동 케적은 한 번의 반복을 전파하도록 허용된다. 각각의 능동 케적 정점의 "변형" 또는 모션은 평균화한 동작으로 결합되어 모두 멤버쉽을 갖고 있는 암시된 메시의 강제적인 전파를 가능하게 한다.

<105> 바람직한 실시예에서, 상이한 영역의 케적의 부분이기도 한 인접한 정점들에 대해 갖고 있는 인접 정점들의 수의 카운트에 정점들이 할당된다. 이러한 다른 정점들은 반대인 것으로 정의된다. 1의 카운트를 갖는 정점의 경우에, 방대 정점을 갖지 않으며, 따라서 보존될 필요가 있다. 2개의 인접한 반대 정점 모두는 1의 카운트를 갖고(이러한 2개의 정점들은 상이한 다각형 내에 있고 서로와 인접함을 의미), 그 후 하나의 정점은 다른 것에 대해 분석된다. 1의 정점이 2의 값을 갖는 이웃하는 다각형 정점과 대향할 때, 1의 카운트를 갖는 정점은 2의 카운트를 갖는 정점으로 분해되고, 그 정점 카운트는 1이 된다. 그러므로 만약 또 다른 이웃하는 대향 정점이 존재한다면, 이 정점은 다시 분해될 수 있다. 이 경우에 대하여, 원래의 정점 카운트를 저장하는 것이 중요하므로, 정점이 분해될 때, 우리는 원래의 정점카운트에 기초하여 분해의 방향을 편향시킬 수 있다. 이와 같이 정점 a는 정점 b로 분해된 후, 정점 b는 정점 c로 분해되지 않을 것이며, 대신 정점 c는 정점 b로 분해되어야 하는데 이는 b가 이미 한 번의 분해에서 사용되었기 때문이다.

<106> 바람직한 실시예에서, T 접합점들이 특이하게 프로세싱된다. 이들 접합점은 인접한 다각형내에 점(point)을 가지지 않는 다각형내의 점들이다. 이 경우, 각각의 다각형 정점은 먼저 이미지 포인트 맵상에 플로팅되고, 이 맵은 정점의 공간상의 위치와 정점의 다각형 식별자를 식별한다. 그 후, 각각의 다각형 경계선이 트래버스법으로 측정되고, 또 다른 다각형으로부터의 인접한 정점들이 존재하는지를 확인하기 위해 테스트된다. 만약 다른 영역으로부터의 이웃하는 정점들이 존재하면, 이들은 각각 현재의 다각형으로부터의 이웃하는 정점을 갖고 있는지를 확인하기 위해 테스트된다. 그렇지 않으면, 현재의 점이 현재 다각형의 정점으로써 부가된다. 이러한 가외의 테스트는 다른 다각형내에 있는 고립된 정점들이 T 접합점을 생성하기 위해 사용됨을 보증한다. 그렇지 않으면, 이는 새로운 정점들을 단지 부가할 것이다 -이 영역은 이미 매칭하는 정점을 가지고 있었다-. 따라서 이웃하는 정점이 이러한 현재의 영역에 의해 대향되지 않기만 하면 대향하는 정점이 부가된다. 추가의 실시예에서, T 접합을 검출하는 능률이 마스크 이미지를 사용함으로써 증가된다. 다각형 정점들은 순차적으로 방문되고, 마스크는 정점의 화소들이 다각형 정점에 속하는 것으로 식별되도록 업데이트된다. 그 후, 다각형 경계 화소가 트래버스법으로 측정되고 이들이 다각형 정점과 일치한다면, 현재의 다각형내의 정점으로써 기록된다.

<107> 바람직한 실시예에서, 스펙트럼 영역이 하나 이상의 중첩하는 동종 이미지 경사 영역에 의해 다시 맵핑되었고,

또 다른 동종 스펙트럼 영역도 중첩할 때, 이미 다시 맵핑된 모든 영역들에는 현재 다시 맵핑되고 있는 그 영역들과 동일한 라벨을 부여받는다. 따라서 본질적으로, 만약 스펙트럼 영역이 두 개의 동종 영역들에 의해 중첩된다면, 이를 두 개의 동종 영역들에 의해 중첩되는 모든 스펙트럼 영역들은 동일한 라벨을 얻을 것이고, 따라서 하나의 스펙트럼 영역은 두 개의 동종 영역들을 대신하여 하나의 동종 영역에 의해 진정하게 커버될 것 같다.

<108> 본 발명의 일 실시예에서, 인접 합병(merge) 영역을 찾기 위해서는 영역 리스트보다는 영역 맵을 프로세싱하는 것이 유리하다. 추가적인 실시예에서, 스펙트럼 분할 분류기는 비동종 영역을 이용하여 분류기를 트레이닝하기 위해 수정될 수 있다. 이에 의해 프로세싱은 스펙트럼 영역의 에지에 집중할 수 있다. 또한, 에지들을 이용하는 것(가령, 캐니(canny) 에지 검출기)에 기초하여 상이한 분할을 부가하는 것과, 그 후 이를 다각형의 최초 집합을 식별하기 위하여 능동 궤적으로 피딩(feeding)하는 것은 동종 영역의 더 큰 판별을 가능하게 한다.

로컬 정규화

<110> 본 발명은 공간시간 스트림내의 화소들이 '로컬' 방식으로 등록될 수 있는 수단을 제공한다.

<111> 한 가지 이러한 로컬화된 방법은, 이미지화된 사상(事象) 또는 구체적으로 이미지화된 객체의 로컬 변형과 관련하여 표면적인 이미지 밝기 불변성 모호성을 해결할 때 이미지화된 사상의 로컬화된 간섭성(coherency)이 설명되도록 화소들을 분석하는 수단을 제공하기 위하여 기하학적 메시의 공간 어플리케이션을 사용한다.

<112> 이러한 메시가 사용되어 로컬 정규화의 수단으로써 이미지 평면에서의 표면 변형의 구분적으로(piece-wise) 선형인 모델을 제공한다. 이미지화된 사상은 비디오내의 모션과 비교하여 비디오 스트림의 시간 분석이 높을 때 이러한 모델과 종종 대응한다. 모델 가정에 대한 예외는 다음과 같은 다양한 기법, 즉 위상적 제약, 주변 정점 제한, 및 화소와 이미지 경사 영역의 동종성의 분석을 포함하는 다양한 기법을 통하여 다루어진다.

<113> 일 실시예에서, 특징점은 삼각형 엘리먼트(이의 정점이 특징점에 대응함)로 구성된 메시를 생성하기 위해 사용된다. 대응하는 특징점들은, 로컬 변형 모델을 생성하기 위하여, 삼각형 및 이에 대응하여 화소들의 보간된 "왜곡(warping)"을 암시하는 다른 프레임이다.

<114> 도 7은 이러한 객체 메시의 생성을 도시하고 있다. 도 8은 프레임을 로컬방식으로 정규화하기 위한 이러한 객체 메시의 이용을 도시하고 있다.

<115> 한 가지 바람직한 실시예에서, 맵의 각 화소가 얻어진 삼각형을 식별하는 삼각형 맵이 생성된다. 또한, 각각의 삼각형에 대응하는 아핀 변환은 최적화 단계로써 미리 계산된다. 그리고 추가적으로, 로컬 변형 모델을 생성할 때, 샘플링 할 소스 화소의 좌표를 결정하기 위하여 공간 좌표를 이용하여 앵커 이미지(anchor image)가 미리 트래버스법으로 측정된다. 이 샘플링된 화소는 현재의 화소 위치를 대체할 것이다.

<116> 다른 실시예에서, 로컬 변형은 글로벌 변형 후에 수행된다. 이전에 개시된 명세서에서, 글로벌 정규화는, 글로벌 등록 방법이 비디오의 둘 이상의 프레임에서 화소를 공간적으로 정규화하기 위해 사용되는 프로세스로써 기재되었다. 결과적인 글로벌하게 정규화된 비디오 프레임은 로컬 방식으로 추가로 정규화될 수 있다. 이러한 두 방법의 결합은 로컬 정규화를 해에 글로벌하게 도착된 정제로 제약한다. 이는 로컬 방법이 해결할 것이 요구되는 모호성을 현저히 감소시킨다.

<117> 또 다른 제한되지 않는 실시예에서, 특징점들, 또는 "정규 메시"의 경우에 있어서의 정점들은, 이를 점의 주변에서 이미지 경사의 분석을 통해 적합화(qualification)된다. 이 이미지 경사는 직접적으로나, 해리스 응답(Harris response)과 같은 몇 가지 간접 계산을 통해 계산될 수 있다. 또한, 이러한 점들은 이미지 경사의 하강과 연관된 모션 추정 에러 또는 공간적 제약에 의해 필터링될 수 있다. 적합화된 점들은 많은 테셀레이션(tessellation) 기법 중 하나에 의해 메시에 대한 기초로 사용될 수 있고, 결과적으로 그 요소들이 삼각형인 메시를 생성한다. 각각의 삼각형에 대하여 아핀 모델이 점들과 이들의 잔여 모션 벡터에 기초하여 생성된다.

<118> 바람직한 실시예에서, 삼각형 아핀 파라미터의 리스트가 유지된다. 이 리스트는 반복되고 현재/이전의 점 리스트가 (정점 조사 맵을 이용하여) 구성된다. 현재/이전의 점 리스트는, 상기 삼각형에 대한 아핀 파라미터를 계산하는 변환을 추정하기 위해 사용되는 루틴으로 통과된다. 아핀 파라미터 또는 모델은, 그 후 삼각형 아핀 파라미터 리스트에 저장된다.

<119> 추가적인 실시예에서, 본 방법은 삼각형 식별자 이미지 맵을 트래버스법으로 측정하며, 맵에 있는 각각의 화소는, 화소가 멤버쉽을 갖는 메시 내의 삼각형에 대한 식별자를 포함한다. 그리고 삼각형에 속하는 각각의 화소에 대하여, 이 화소에 대한 대응하는 글로벌 변형 및 로컬 변형 좌표가 계산된다. 이러한 좌표는, 다시, 대응

하는 화소를 샘플링하고 그 값을 대응하는 "정규화" 위치에 적용하기 위하여 사용된다.

<120> 추가적인 실시예에서, 공간적 제약이 이미지 경사의 검색으로부터 생성된 이미지 강도(intensity) 대응 세기(strength) 및 밀도(density)에 기초하여 점들에 적용된다. 점들은, 모션 추정이 이미지 강도 나머지의 몇몇 놈(norm)에 기초하여 행해진 후, 정렬된다. 그 후 점들은 공간 밀도 제약에 기초하여 필터링된다.

<121> 추가적인 실시예에서, 스펙트럼 공간 분할이 사용되며, 작은 동종 스펙트럼 영역이 공간적 인접성, 이들의 밀도의 유사성, 및/또는 색상에 기초하여 이웃하는 영역들과 병합된다. 그 후, 동종 병합은 동종 텍스처(이미지 경사)의 영역과의 중첩에 기초하여 스펙트럼 영역을 서로 결합하기 위해 사용된다. 그 후 추가적인 실시예는 작은 영역이 보다 큰 영역에 의해 둘러싸여 있는 중앙-포위 점(center-surround point)을 메시의 정점을 지원하기 위한 적합화된 관점으로 사용한다. 추가의 제한되지 않는 실시예에서, 중앙 포위 점은 크기가 3×3 또는 5×5 또는 7×7 화소인 하나의 화소내에 바운딩(bounding) 박스가 존재하는 영역으로 정의되며, 이 바운딩 박스에 대한 공간 이미지 경사는 코너 모양이다. 이 영역의 중심은 코너로 분류될 수 있고, 이 위치를 유리한 정점 위치로 추가로 적합화한다.

<122> 추가적인 실시예에서, 수평 및 수직 화소 유한 차분 이미지가 사용되어 각각의 메시 에지의 세기를 분류한다. 만약 에지가 공간 위치와 일치하는 많은 유한 차분을 가진다면, 에지와 따라서 이 에지의 정점은 이미지화된 사상의 로컬 변형에 매우 중요한 것으로 고려된다. 만약 에지의 유한 차분의 합의 평균들 사이에 큰 미분계수 차가 존재하면, 아마 영역 에지는 양자화 단계가 아닌 텍스처 변경 에지에 대응한다.

<123> 추가적인 실시예에서, 공간 밀도 모델 종결 조건이 사용되어 메시 정점의 프로세싱을 최적화한다. 검출 사각형의 최초(outset)의 공간 영역의 대부분을 커버하는 충분한 수의 점들이 검사되었을 때, 프로세싱은 종결될 수 있다. 종결은 점수를 생성한다. 프로세싱에 입력되는 정점 및 특징점은 이 점수에 의해 정렬된다. 점들이 기존 점들에 공간적으로 너무 가깝다거나, 점이 이미지 경사에 있는 에지에 대응하지 않는다면, 이 점은 무시된다. 그렇지 않다면, 점의 주변에 있는 이미지 경사가 하강되고, 경사의 나머지가 제한을 초과하면, 이 점도 또한 무시된다.

규칙적인 메시 정규화(Regular Mesh Normalization)

<124> 본 발명은 전술한 로컬 정규화 방법을 규칙적인 메시를 이용하여 확장한다. 이 메시는 기초가 되는 화소와 관련 없이 구성되지만, 검출된 객체에 대응하여 위치되며 크기를 갖는다.

<125> 검출된 객체 영역이 주어지면, 공간 프레임 위치 및 표면의 크기를 나타내는 스케일은 표면 영역의 개시부에 대하여 정규 메시를 생성한다. 바람직한 실시예에서, 사각형 메시의 윤곽을 그리기 위하여 중첩하지 않는 세트의 타일을 사용한 후 삼각형 메시 요소를 갖는 규칙적인 메시를 생성하도록 타일들의 대각선 분할(partitioning)을 사용한다. 추가적인 바람직한 실시예에서, 타일들은 종래의 비디오 압축 알고리즘(예, MPEG-4 AVC)에서 사용되는 것들과 비례한다.

<126> 바람직한 실시예에서, 전술한 메시와 연관된 정점들은 트레이닝을 위해 사용되는 비디오의 특정 프레임들에서 이 정점들을 둘러싸고 있는 화소 영역의 분석을 통해 우선순위가 부여된다. 이러한 영역에 대한 경사의 분석은 로컬 이미지 경사에 의존하는 각각의 정점과 연관된 프로세싱(가령 블록 기반의 모션 추정)에 관하여 신뢰를 제공한다.

<127> 다수의 프레임에서의 정점 위치의 대응성은 이미지 경사의 간단한 하강을 통해 발견된다. 바람직한 실시예에서, 이는 블록 기반 모션 추정을 통해 얻어진다. 본 실시예에서, 높은 신뢰도의 정점은 높은 신뢰도의 대응성을 가능하게 한다. 낮은 신뢰도의 정점 대응성은 높은 신뢰도의 정점 대응성으로부터의 추론을 통해 모호한 이미지 경사를 해결함으로써 암시적으로 이루어진다.

<128> 한 가지 바람직한 실시예에서, 규칙적인 메시는 개시부 추적 사각형에 대하여 생성된다. 타일들은 16×16 으로 생성되고, 대각선 방향으로 절단되어, 삼각형 메시를 형성한다. 이러한 삼각형들의 정점들은 모션 추정된다. 모션 추정은 각각의 점들이 가지는 텍스처의 유형에 의존한다. 텍스처는 세 개의 클래스들, 코너, 에지, 및 균일면(homogeneous)으로 나누어지는데, 이는 또한 정점들의 프로세싱의 순서를 정의하기도 한다. 코너 정점은 이웃하는 정점 추정을 이용, 다시 말해, (이용가능하다면) 이웃하는 점들의 모션 추정이 예측 모션 벡터용으로 이용되고, 모션 추정이 각각의 것에 적용된다. 가장 낮은 안 좋은(mad) 오차를 제공하는 모션벡터가 이 정점 모션 벡터로써 사용된다. 코너에 대해 사용되는 검색 전략은 모든 것(넓고, 작고, 최초의 것)이 사용된다. 에지들에 대하여, 다시 가장 가까운 주변 모션 벡터들이 예측 모션 벡터로 사용되고, 최소량의 오차를 갖는 것이 사용된다. 에지에 대한 검색 전략은 작고 최초의 것이다. 균일면에 대해서는, 이웃하는 정점이 검색되고 가장

낮은 오차를 갖는 모션 추정이 사용된다.

<130> 한 가지 바람직한 실시예에서, 각각의 삼각형 정점에 대한 이미지 경사가 계산되고, 클래스 및 크기(magnitude)에 기초하여 정렬된다. 따라서 코너들이 예지 이전이며, 예지들은 균일면 이전이다. 코너에 대해서는, 강한 코너는 약한 코너 이전이며, 예지에 대해서는 강한 예지가 약한 예지 이전이다.

<131> 한 가지 바람직한 실시예에서, 각각의 삼각형에 대한 로컬 변형은 상기 삼각형과 연관된 모션 추정에 기초한다. 각각의 삼각형은 이에 대해 추정된 아핀(affine)을 갖는다. 삼각형이 위상적으로(topologically) 반전되지 않거나, 변질되지 않는다면, 삼각형의 부분인 화소들은 얻어진 추정 아핀에 기초하여 현재 이미지를 샘플링하기 위해 사용된다.

<132> 대안의 실시예에서, 화소 영역들은 트레이닝 기간 동안, 낮은 공간 변형에 기초하여 적합화될 수 있다. 적은 변형은 로컬 변형 모델링의 관점에서 보다 높은 우선도를 나타낸다. 예시 기반의 모션 추정 및 정규화는 시종 일관 발생하는 화소들의 가장 큰 수를 포함하는 경향이 있는 메시 면(facet)을 선호하도록 치우쳐질 수 있다. 본 기법의 추가적인 실시예에서, 예지 검출 응답이 면들을 추가로 적합화하기 위해 사용된다.

분할(Segmentation)

<134> 추가로 설명된 분할 프로세스를 통해 식별된 공간 불연속은 이를 각각의 경계의 지리적 파라미터화를 통해 효율적으로 엔코딩되며, 이는 공간 불연속 모델로 불린다. 이러한 공간 불연속 모델은 엔코딩의 서브세트에 대응하는 더욱 더 간결한 한계 설명을 가능하게 하는 진보된 방법으로 엔코딩될 수도 있다. 진보한 엔코딩은 공간 불연속의 돌출 특성의 대부분을 유지하면서 공간 기하를 우선으로 하는 강고한 수단을 제공한다.

<135> 도 2에 도시된 것처럼, 일단 객체의 특정의 대응이 시간에 대해 추적(220)되고 모델링(224)되었다면, 이러한 모션/변형 모델에 대한 충실한 지지가 객체에 대응하는 화소를 분할(230)하기 위해 사용될 수 있다. 이 프로세스는 비디오(202&204)내의 검출된 다수의 객체(206&208)에 대해 반복될 수 있다. 이러한 프로세싱의 결과는 분할된 객체 화소(232)이다.

<136> 본 발명에 의해 사용되는 불변 특성 분석의 일 형태는 공간 불연속의 식별에 초점을 맞춘다. 이러한 불연속은 예지, 음영, 폐색(occlusion), 라인, 코너, 또는 비디오의 하나 이상의 이미지화된 프레임의 화소들 사이에서 갑작스럽고 식별 가능한 분리를 유발하는 소정의 다른 가시적 특성을 드러낸다. 추가적으로, 유사하게 채색된 및/또는 텍스처화된 객체 사이의 미묘한 공간 불연속은 비디오 프레임에서 객체의 화소가 객체 그 자체와 관련한 코히어런트한 모션을 진행하지만, 서로에 대해 상이한 모션인 경우 드러난다. 본 발명은 돌출 신호 모드와 관련한 공간 불연속을 강고하게 식별하기 위해 스펙트럼, 텍스처, 및 모션 분할의 조합을 이용한다.

시간적 분할

<138> 변환 모션 벡터의 시간적 통합, 또는 등가적으로 더 높은 등급의 모델로의 유한 차분 측정은 종래 기술에 설명된 모션 분할의 형태이다.

<139> 본 발명의 일 실시예에서, 비디오에서 객체 모션의 유한 차분을 나타내는 모션 벡터의 밀집한 필드가 생성된다. 이러한 파생물은 타일의 일정한 분할을 통해, 또는 공간 분할과 같은 소정의 초기 절차에 의해 서로 공간적으로 그룹화된다. 각각의 그룹의 "파생물(derivatives)"은 선형 최소 제곱 추정을 이용하여 더 높은 등급의 모션 모델로 통합된다. 이어 최종 모션 모델은 k-수단 클러스터링(k-means clustering) 기술을 이용하여 모션 모델 공간에서 벡터로서 클러스터링된다. 파생물은 어떤 클러스터가 이들에게 최상으로 적합한지에 기초하여 분류된다. 이어 클러스터 라벨은 공간 분할의 전개로서 공간적으로 클러스터링된다. 프로세스는 공간 분할이 안정될 때까지 연속된다.

<140> 본 발명의 추가의 실시예에서, 주어진 구경의 모션 벡터는 구경에 대응하는 화소 위치의 세트로 보간된다. 이러한 보간에 의해 한정된 블록이 객체 경계에 대응하는 화소에 미치면, 최종 분류는 블록의 소정의 변칙 대각 분할이다.

<141> 종래 기술에서, 파생물을 통합하기 위해 사용된 최소 제곱 추정기는 이상점에 매우 민감하다. 민감도는 반복이 넓게 발산하는 포인트에 대한 모션 모델링 클러스터링 방법을 바이어스하는 모션 모델을 생성할 수 있다.

<142> 본 발명에서 모션 분할 방법은 비디오의 두 개 이상의 프레임에 대한 명확한 화소 모션의 분석을 통해 공간 불연속을 식별한다. 명확한 모션은 비디오의 프레임에 대한 일치에 대해 분석되고 파라미터의 모션 모델로 통합된다. 이러한 일정한 모션과 관련된 공간 불연속이 식별된다. 모션 분할은 시간적 변화가 모션에 의해 유발되

기 때문에, 시간적 분할로서 언급될 수 있다. 그러나 시간적 변화는 로컬 변형, 조명 변화 등과 같은 소정의 다른 현상에 의해 유발될 수도 있다.

<143> 전술한 방법을 통해, 정규화 방법에 대응하는 돌출 신호 모드는 몇몇 배경 감산 방법 중 하나를 통해 주위의 신호 모드(배경 또는 비객체)로부터 식별 및 분리될 수 있다. 종종, 이러한 방법은 배경을 각각의 시간 인스턴스에서 최소의 변화량을 나타내는 화소로서 통계적으로 모델링한다. 변화는 화소 값 차로서 특징지워질 수 있다.

<144> 분할 경계 기반 글로벌 변형 모델링은, 객체 둘레의 경계를 생성한 후, 이 경계를 객체의 검출된 중심쪽으로 경계 정점들이 이종 이미지 경사와 일치하는 위치를 얻을 때까지 봉괴시킴으로써 얻어질 수 있다. 모션 추정들은 이러한 새로운 정점 위치들에 대하여 수집되고, 강고한 아핀 추정이 글로벌 변형 모델을 발견하기 위해 사용된다.

<145> 분할 메시 정점 이미지 경사 하강에 기반한 유한 차분은 글로벌 변형 모델로 통합된다.

객체 분할

<147> 도 13에 있는 블록도는 객체 분할의 일 바람직한 실시예를 도시한다. 프로세스는 정규화된 이미지의 총합(1302)으로부터 시작되며, 상기 총합은 그 후 총합 사이에서 쌍 방향으로(pair-wise) 차분화된다. 이러한 차분은 그 후 엘리먼트 방향으로(element-wise) 누적 베피에 누적(1306)된다. 누적 베피는 보다 현저한 오차 영역을 식별하기 위하여 임계점설정(threshold)(1310)된다. 임계점설정된 엘리먼트 마스크는 그 후 누적된 오차 영역(1310)의 공간적 지원을 결정하기 위하여 형태학적으로 분석된다(1312). 형태학적 분석(1312)의 결과적인 추출(1314)은 그 후 객체와 일치하는 누적된 오차 영역상의 이후의 프로세싱에 집중하기 위하여 검출된 객체 위치와 비교된다(1320). 고립된 공간 영역(1320)의 경계는 그 후 정점 외부가 생성(1324)되는 다각형을 이용하여 근사화된다(1322). 외부의 궤적은 그 후 정점의 위치를 능동 궤적 분석(1332)을 위해 더욱 양호하게 초기화하기 위하여 조정된다(1330). 일단 능동 궤적 분석(1332)이 누적된 오차 공간에 있는 저 에너지 해로 수렴되었다면, 궤적은 최종 궤적(1334)으로 사용되고 궤적 내에 속박된 화소들은 객체화소일 것으로 간주되며, 궤적의 외부에 있는 화소들은 비객체 화소들인 것으로 간주된다.

<148> 바람직한 실시예에서, 모션 분할은 돌출 이미지 모드의 검출된 위치 및 스케일이 주어진다면 어려울 수 있다. 거리 변환이 사용되어 검출된 위치로부터의 모든 화소의 거리를 결정한다. 최대 거리와 연관된 화소 값이 유지되면, 배경의 합리적인 모델이 해결될 수 있다. 다시 말해서, 주변 신호는 신호 차분 메트릭을 이용하여 일시적으로 재샘플링된다.

<149> 추가적인 실시예는 각각의 화소에 거리를 할당하기 위하여 현재의 검출 위치에 대해 거리 변환을 사용하는 것을 포함한다. 화소에 대한 거리가 소정의 최대 화소 거리 테이블에 있는 거리보다 크다면, 화소 값이 기록된다. 적절한 트레이닝 기간 후에, 이 화소에 대한 최대 거리가 크다면, 화소는 배경 화소일 가능성은 갖는 것으로 추정된다.

<150> 주변 신호의 모델이 주어지면, 각각의 순간에서의 완전한 돌출 신호 모드가 차분화된다. 각각의 이러한 차분은 공간적으로 정규화된 신호 차분(절대 차분)으로 재샘플링될 수 있다. 이러한 차분들은 그후 서로에 대해 정렬되고 누적된다. 이러한 차분들이 돌출 신호 모드에 대해 공간적으로 정규화되었으므로, 차분의 피크는 돌출 신호 모드와 연관된 화소 위치에 대부분 대응할 것이다.

<151> 본 발명의 일 실시예에서, 트레이닝 기간이 정의되는데, 여기서 객체 검출 위치가 결정되고 이를 위치의 중심은, 비객체 화소가 될 가장 큰 가능성을 가질 배경 화소를 생성하기 위하여 프레임 차분을 허용하는 이러한 위치로부터 이격된 검출 위치를 이용하여 최적 프레임 수를 결정하기 위하여 사용된다.

<152> 본 발명의 일 실시예에서, 능동 궤적 모델링은, 누적된 오차 "이미지"에서 궤적 정점 위치를 결정함으로써 비객체 배경으로부터 전경(foreground) 객체를 분할하기 위하여 사용된다. 바람직한 실시예에서, 능동 궤적 예지들은 검출된 객체의 스케일과 동일한 크기가 되도록 세분되어 더 큰 자유도를 생성한다. 바람직한 실시예에서, 최종 궤적 위치는 규칙적으로 이격된 궤적을 생성하도록 가장 가까운 규칙적 메시 정점에 스냅핑(snapping)된다.

<153> 객체 분할의 한 가지 제한되지 않는 실시예에서, 일시적으로 쌍방향의 이미지에 대한 오차 이미지 필터 응답을 생성하기 위해 배향된 핵(kernel)이 사용된다. 총체의 모션 방향에 직교하여 배향되는 필터에 대한 응답은, 배경에 대한 모션이 배경의 폐색 및 누출로부터 발생할 때 오차 표면을 증가시키는 경향이 있다.

<154> 정규화된 이미지의 총합에 대한 정규화된 이미지 프레임 강도 벡터는 잔여 벡터를 생성하는 하나 이상의 기준 프레임으로부터 차분화된다. 이러한 잔여 벡터는 누적된 잔여 벡터를 형성하기 위하여 엘리먼트 방향으로 누적된다. 이러한 누적된 잔여 벡터는 그 후 객체 및 비객체 화소의 공간적 분할을 위한 공간적 객체 경계를 정의하기 위하여 공간적으로 조사(probe)된다.

<155> 일 바람직한 실시예에서, 누적된 잔여 벡터의 최초의 통계적 분석이 수행되어 누적된 잔여 벡터의 임계점을 설정하기 위하여 사용될 수 있는 통계적 임계값이 얻어진다. 수축(erosion) 및 이어지는 팽창(dilation) 형태학상의 연산을 통해, 예비적인 객체 영역 마스크가 생성된다. 그 후 영역의 궤적 다각형 점들이 분석되어 이러한 점들의 볼록 껍질(convex hull)을 드러낸다. 볼록 껍질은 그 후 능동 궤적 분석 방법을 위한 최초의 궤적으로 이용된다. 능동 궤적은 이것이 객체의 누적된 잔여물의 공간적 경계에 수렴할 때까지 전파된다. 추가의 바람직한 실시예에서, 예비적인 궤적의 에지들은 모든 에지 길이들에 대하여 최소의 에지 길이가 얻어질 때까지 중간점 정점들을 부가함으로써 더 세분화된다. 이러한 추가적인 실시예는 객체의 외곽선을 보다 정확히 맞추기 위하여 능동 궤적 모델의 자유도를 증가시키도록 되어 있다.

<156> 바람직한 실시예에서, 개선된 궤적이 사용되어 궤적에 의해 암시되는 다각형을 덮어씌우고 정규화된 이미지내의 다각형을 덮어씌움으로써 객체의 화소를 나타내는 화소 마스크를 생성한다.

비객체의 분해(resolution)

<158> 도 12에 도시된 블록도는 비객체 분할, 또는 등가적으로, 배경 분해의 일 바람직한 실시예를 개시한다. 배경 버퍼의 초기화(1206) 및 최초의 최대 거리 값(1204)을 이용하여, "안정도"를 검출된 객체 위치(1202)로부터의 가장 긴 거리와 연관시킴으로써 프로세스는 가장 안정된 비객체 화소를 결정하도록 동작한다. 새롭게 검출된 객체 위치(1202)가 주어지면, 프로세스는 각각의 화소 위치를 체크한다(1210). 각각의 화소 위치(1210)에 대하여, 검출된 객체 위치(1202)로부터의 거리가 거리 변환을 이용하여 계산된다. 만약 이 화소에 대한 거리가 최대 거리 버퍼(1204)내의 미리 저장된 위치보다 크다면(1216), 이전의 값은 현재의 값(1218)으로 대체되며 화소 값은 화소 버퍼에 기록된다(1220).

<159> 분해된 배경 이미지가 주어지면, 이미지와 현재의 프레임 사이의 오차는 공간적으로 정규화되고 시간적으로 누적될 수 있다. 이러한 분해된 배경 이미지는 "배경 분해" 섹션에서 기재된다.

<160> 결과적인 누적된 오차는 그 후 최초의 궤적을 제공하도록 임계점 설정된다. 궤적은 그 후 궤적 변형에 대항하여 균형 오차 잔여물로 공간적으로 전파된다.

경사도 분할

<162> 텍스처 분할 방법, 또는 등가적으로 강도 경사도 분할은 비디오의 하나 이상의 프레임에서 화소의 로컬 경사도를 분석한다. 경사도 응답은 비디오 프레임내의 화소 위치에 국부적인 공간 불연속을 특징지우는 통계 측정이다. 그 후 몇몇 공간 클러스터링 기술 중 하나가 공간 영역으로 경사도 응답을 결합하는데 사용된다. 이러한 영역의 경계는 하나 이상의 비디오 프레임의 공간적 불연속을 식별하는데 유용하다.

<163> 본 발명의 일 실시예에서, 컴퓨터 그래픽 텍스처 생성으로부터의 합산된 영역 테이블 개념은 강도 필드의 경사도의 계산을 촉진할 목적으로 사용된다. 네 개의 추가 연산과 결합된 4개의 검색(lookup)을 통해 원본 필드의 소정의 직사각형의 합을 용이하게 하는 누진적으로 합산된 값들의 필드가 생성된다.

<164> 추가의 실시예는 이미지에 대해 생성된 해리스 응답(Harris response)을 이용하며, 각각의 화소에 이웃한 화소는 균일면, 에지 또는 코너로 분류된다. 응답 값은 이러한 정보로부터 생성되며 프레임내의 각각의 엘리먼트에 대한 에지 상태 또는 코너 상태의 정도를 표시한다.

다중 스케일 경사도 분석

<166> 본 발명의 일 실시예는 수 개의 공간 스케일을 통해 이미지 경사도 값들을 생성함으로써 이미지 경사도 지원을 추가로 강제한다. 이 방법은 이미지 경사도를 적합화하는 데 도움이 될 수 있어서, 상이한 스케일에서의 공간 불연속이 서로를 지원하는 데 사용된다 - "에지"가 수 개의 상이한 공간 스케일에서 식별되는 한 에지는 "돌출"되어 있어야 한다. 보다 적합화된 이미지 경사도는 보다 돌출된 특징에 대응하는 경향이 있다.

<167> 바람직한 실시예에서, 텍스처 응답 필드가 먼저 생성되고, 이 필드의 값이 그후 k-수단 비닝(binning)/분할(partitioning)에 기초하여 수 개의 저장소(bin)로 적합화된다. 원본 이미지 경사도 값들은 그 후 각각의 저장소를 한 번의 반복이 분기점 분할을 적용할 수 있는 값들의 간격으로 이용하여 점진적으로 프로세싱된다. 이러

한 접근법의 이점은 균일성이 강한 공간적 치우침에 대해 상대적인 의미로 정의된다는 것이다.

<168>

스펙트럼 분할

<169>

스펙트럼 분할 방법은 비디오 화소에서 흑백, 그레이 스케일, 또는 컬러 화소의 통계적 확률 분포를 분석한다. 스펙트럼 분류기는 이러한 화소의 확률 분포에 대한 클러스터링 연산을 실행함으로써 구성된다. 이어 분류기는 하나 이상의 화소를 확률 클래스에 속하는 것으로 분류하기 위해 사용된다. 최종 확률 클래스 및 그 화소에는 클래스 라벨이 주어진다. 이어 이러한 클래스 라벨은 뚜렷한 경계를 갖는 화소의 영역에 공간적으로 관련된다. 이러한 경계는 하나 이상의 비디오 프레임의 공간 불연속을 식별한다.

<170>

본 발명은 비디오의 프레임내의 화소를 분할하기 위해 스펙트럼 분류에 기초한 공간 분할을 이용할 수 있다. 더욱이, 영역들 사이의 대응은 앞선 세그먼트에서 영역을 갖는 스펙트럼 영역의 중첩에 기초하여 결정될 수 있다.

<171>

비디오 프레임이 비디오 프레임내의 객체에 대응하는 더욱 큰 영역으로 공간적으로 연결된 연속한 컬러 영역으로 개략적으로 구성된 경우, 채색된(또는 스펙트럼의) 영역의 식별 및 추적은 비디오 시퀀스내의 객체의 이후의 분할을 용이하게 할 수 있다.

<172>

배경 분할

<173>

기재된 발명은 검출된 객체와 비디오의 각각의 프레임내의 각각의 개별 화소 사이의 공간 거리 측정의 시간적 최대값에 기초하는 비디오 프레임 배경 모델링을 위한 방법을 포함한다. 객체의 검출된 위치가 주어진다면, 거리 변환에 적용되어, 프레임 내의 각각의 화소에 대한 스칼라 거리 값을 생성한다. 각각의 화소에 대하여, 모든 비디오 프레임에 대한 최대 거리의 맵이 유지된다. 최대값이 처음으로 할당될 때, 또는 이후에 새롭고 상이한 값으로 업데이트 될 때, 이 비디오 프레임에 대한 대응하는 화소는 "분해된 배경" 프레임내에 유지된다.

<174>

외관 모델링

<175>

비디오 프로세싱의 통상의 목적은 종종 비디오 프레임의 시퀀스의 외관을 모델링 및 유지하는 것이다. 본 발명은 프로세싱의 사용을 통해 강고하고 넓게 적용 가능한 방식으로 적용되는 제한된 외관 모델링 기술을 가능하게 하는 것이 목적이다. 전술한 등록, 분할, 및 정규화는 이러한 목적을 위해 설명된다.

<176>

본 발명은 외관 변화 모델링의 수단을 개시한다. 선형 모델의 경우, 외관 변화 모델링의 주된 기반은 선형 상관을 활용하는 컴팩트한 기반을 나타내는 특징 벡터의 분석이다. 공간 강도 필드 화소를 나타내는 특징 벡터는 외관 변화 모델로 어셈블링될 수 있다.

<177>

대안의 실시예에서, 외관 변화 모델은 화소의 분할된 서브세트로부터 계산된다. 더욱이, 특징 벡터는 공간적으로 비중첩 특징 벡터로 분할될 수 있다. 이러한 공간 분해는 공간 타일링으로 달성될 수도 있다. 계산 효율은 더욱 광역의 PCA 방법의 차원 감소를 희생하지 않고 이러한 시간적 양상을 프로세싱을 통해 달성할 수도 있다.

<178>

외관 변화 모델을 생성할 때, 공간 강도 필드 정규화는 공간 변화의 PCA 모델링을 감소시키기 위해 사용될 수 있다.

<179>

PCA

<180>

출현 변화 모델을 생성하는 바람직한 수단은, 패턴 벡터로서의 비디오의 프레임을 트레이닝 매트릭스, 또는 양상블, 및 트레이닝 매트릭스에 대한 중요 컴포넌트 분석(PCA)의 애플리케이션으로 조립시키는 것이다. 이러한 확장이 생략되면, 최종 PCA 변환 매트릭스는 비디오의 이후의 프레임을 분석 및 합성하기 위해 사용된다. 생략의 레벨에 기초하여, 화소의 원본 외관의 품질의 변하는 레벨이 얻어질 수 있다.

<181>

패턴 벡터의 구성 및 분해의 특정한 수단은 기술 분야의 당업자에게 잘 알려져 있다.

<182>

주위 신호로부터 돌출 신호 모드의 공간 분할 및 이러한 모드의 공간 정규화가 주어지면, 화소 그 자체, 또는 등가적으로 최종 정규화된 신호의 외관은 화소 외관의 표현에 대한 비트 레이트와 근사 에러 사이의 직접적인 교환을 허용하는 낮은 랭크 파라미터화를 이용하여 선형으로 상관된 컴포넌트로 팩토링될 수 있다.

<183>

도 2에 도시된 것처럼, 데이터의 크기에 있어서 간결한 버전(252&254)을 생성하기 위하여, 정규화된 객체 화소(242&244)는 벡터 공간으로 투사되고 선형 대응이 PCA와 같은 분해 프로세스(250)를 이용하여 모델링될 수 있다.

<184>

연속 PCA

<185>

PCA는 PCA 변환을 이용하여 패턴을 PCA 계수로 엔코딩한다. 패턴이 PCA 변환에 의해 더욱 잘 표현되면, 패턴을 엔코딩하기 위해 계수가 덜 필요하다. 트레이닝 패턴의 획득과 엔코딩 될 패턴의 획득 사이에서 시간이 경과함에 따라 패턴 벡터가 열화될 수 있음을 인식하면, 변환을 업데이트하는 것은 열화의 반대 작용에 도움을 준다. 새로운 변환을 생성하기 위한 대안으로서, 현재 패턴의 연속 업데이트가 소정의 경우 더욱 계산 효율적이다.

<186>

많은 최신 비디오 압축 알고리즘은 하나 이상의 다른 프레임으로부터 비디오의 프레임을 예측한다. 예측 모델은 각각의 예상 프레임을, 다른 프레임의 대응하는 패치로 매칭된 비중첩 타일 및 오프셋 모션 벡터에 의해 파라미터화된 관련된 변환 위치로 분할하는 것에 공통적으로 기초한다. 프레임 인덱스와 선택적으로 결합된 이러한 공간 변위는 타일의 "모션 예측된" 버전을 제공한다. 만일 예측의 에러가 소정의 임계치 미만이면, 타일의 화소는 나머지 인코딩에 적합하며; 압축 효율에 있어서 대응하는 이득이 존재한다. 그렇지 않으면, 타일의 화소는 직접 엔코딩된다. 이러한 타입의 타일 기반 - 대안으로는 블록 기반이라 칭함 - 모션 예측 방법은 화소를 포함하는 타일을 해석함으로써 비디오를 모델링한다. 비디오내의 이미지화된 사상(事象)이 이러한 타입의 모델링을 지지하면, 대응하는 엔코딩 효율이 증가한다. 이러한 모델링 제한은, 블록 기반의 예측에 고유한 해석적 가정을 따르기 위하여 시간적 분해의 소정 레벨, 또는 초당 프레임의 수가 모션을 겪는 이미지화된 객체에 대해 제공된다고 가정한다. 이러한 해석 모델에 대한 다른 요구는 소정의 시간적 분해에 대한 공간적 변위가 제한되어야 한다는 것인데; 즉, 예측이 유도되는 대상이 되는 프레임과 예측되고 있는 프레임 사이의 시간차는 절대 시간의 상대적으로 짧은 양이어야 한다. 이러한 시간적 분해 및 모션 제한은 비디오 스트림에 존재하는 소정의 과잉의 비디오 신호 컴포넌트의 식별 및 모델링을 용이하게 한다.

<187>

나머지 기반 분해

<188>

MPEG 비디오 압축에서, 현재 프레임은 모션 벡터를 이용하여 이전의 프레임을 보상하는 모션에 의해 구성되며, 이어 보상 블록에 대한 나머지 업데이트의 애플리케이션이 뒤따르며, 끝으로 충분한 부합을 갖지 않는 소정의 블록이 새로운 블록으로서 엔코딩된다.

<189>

나머지 블록에 대응하는 화소는 모션 벡터를 통해 이전의 프레임의 화소로 맵핑된다. 결과는 나머지 값들의 연속한 애플리케이션을 통해 합성될 수 있는 비디오를 통한 화소의 시간적 경로이다. 이러한 화소는 PCA를 이용하여 최상으로 표현될 수 있는 화소로서 식별된다.

<190>

폐색 기반 분해(Acclusion-based Decomposition)

<191>

본 발명의 추가의 확장은 블록에 제공된 모션 벡터가 화소를 이동시킴으로써 이전의 프레임이 폐색(커버링)되게 할 것인지를 결정한다. 각각의 폐색 이벤트의 경우, 폐색 화소를 새로운 층으로 분할한다. 또한 히스토리가 없는 드러난 화소가 존재한다. 드러난 화소는 현재 프레임에서 이들을 적합하게 하고, 히스토리컬 적합성이 상기 층에 대해 행해질 소정의 층으로 배치된다.

<192>

화소의 시간적 연속은 분할 및 상이한 층으로의 화소의 결합을 통해 지원된다. 일단 안정한 층 모델이 형성되면, 각각의 층의 화소는 코히어런트한 모션 모델에 대한 멤버십에 기초하여 그룹화될 수 있다.

<193>

서브 대역 시간 양자화

<194>

본 발명의 택일적 실시에는 각각의 프레임을 서브 대역 이미지로 분해하기 위해 이산 코사인 변환(DCT) 또는 이산 웨이블릿 변환(DWT)을 이용한다. 이어 중요 컴포넌트 분석(PCA)이 각각의 이러한 "서브 대역" 비디오에 적용된다. 개념은 비디오 프레임의 서브대역 분해가 원본 비디오 프레임과 비교하여 서브 대역 중 하나에서 공간 변화를 감소시킨다는 것이다.

<195>

이동 객체(사람)의 비디오의 경우, 공간 변화는 PCA에 의해 모델링된 변화를 지배하는 경향이 있다. 서브 대역 분해는 소정의 하나의 분해 비디오에서 공간 변화를 감소시킨다.

<196>

DCT의 경우, 소정의 하나의 서브 대역에 대한 분해 계수는 공간적으로 서브 대역 비디오로 배열된다. 예를 들어, DCT 계수는 각각의 블록으로부터 취해지며, 원본 비디오의 우표 버전으로 보이는 서브 대역 비디오로 정렬된다. 이는 모든 다른 서브 대역에 대해 반복되며, 최종 서브 대역 비디오는 PCA를 이용하여 각각 프로세싱된다.

<197>

DWT의 경우, 서브 대역은 DCT에 대해 설명된 방식으로 이미 정렬된다.

<198> 비제한 실시예에서, PCA 계수의 생략이 변화된다.

웨이블릿

<200> 데이터가 이산 웨이블릿 변환(DWT)을 이용하여 분해되면, 다수의 대역 통과 데이터 세트는 보다 낮은 공간 해상도로 귀결된다. 단일의 스칼라 값만이 생길 때까지, 변환 프로세스는 반복적으로 유도된 데이터에 적용될 수 있다. 분해된 구조의 스칼라 엘리먼트들은 계층적 부모/자식 방식으로 통상적으로 관련된다. 최종 데이터는 다해상도 계층 구조 및 유한 차분을 포함한다.

<201> DWT가 공간 강도 필드에 적용될 경우, 많은 자연적으로 발생한 이미지 사상은 낮은 공간 주파수로 인해 제1 또는 제2 저대역 통과 유도 데이터 구조에 의해 작은 지각있는 손실로 표현된다. 계층적 구조의 생략은 높은 주파수 공간 데이터가 제공되지 않거나 잡음으로 고려되지 않으면, 컴팩트한 표현을 제공한다.

<202> PCA가 작은 수의 계수로 정확한 개조를 달성하기 위해 사용될 수 있는 반면, 변환 그 자체는 매우 클 수 있다. 이러한 "초기" 변환의 크기를 줄이기 위해, 웨이블릿 분해의 내장된 제로 트리(EZT) 구조가 변환 매트릭스의 점진적으로 더욱 정교한 버전을 구축하는데 사용될 수 있다.

부분공간 분류

<204> 기술분야에서 숙달된 자에 의해 잘 이해될 수 있는 것처럼, 이산 샘플링 사상 데이터 및 유도 데이터는 대수 벡터 공간에 대응하는 데이터 벡터들의 세트로 표현될 수 있다. 이러한 데이터 벡터들은, 제한되지 않는 방식으로, 분할된 객체의 정규화된 외관에 있는 화소와, 모션 파라미터와, 2차원 또는 3차원상의 특징 또는 정점의 구조적 위치를 포함한다. 이러한 벡터들 각각은 벡터 공간에 존재하며, 공간의 기하구조의 분석이 샘플링된 벡터 또는 파라미터 벡터의 간결한 표현을 생성하기 위해 사용될 수 있다. 유용한 기하구조 상태는 컴팩트한 부분공간을 형성하는 파라미터 벡터에 의해 유형화된다. 하나 이상의 부분공간이 혼합되어, 표면적으로 보다 복잡한 단일 부분공간을 생성할 때, 구성요소의 부분공간들은 식별하기가 어려워질 수 있다. 원본 벡터들의 소정의 상호작용(가령 내적)을 통해 생성되는 고차원 벡터 공간에서 데이터를 검사함으로써 이러한 부분공간의 분리를 가능하게 하는 여러 분할 방법이 존재한다.

<205> 벡터 공간을 분할하는 한 가지 방법은 다항식을 표현하는 베로네세(Veronese) 벡터 공간으로 벡터를 투사하는 것을 포함한다. 이 방법은 일반화된 PCA 또는 GPCA로 기술분야에서 잘 알려져 있다. 이러한 투사를 통해, 다항식에 대한 표준이 발견되고, 그룹화되며, 이러한 표준과 연관된 원본 벡터는 함께 그룹화될 수 있다. 이러한 기술의 이용에는 시간에 대해 추적되는 2차원 공간 점 대응을 3차원 구조 모델 및 이 3차원 모델의 모션으로 팩토링하는 것이다.

<206> GPCA 기법은 정의된 대로 적용될 때 불완전하며, 데이터 벡터가 적은 노이즈로 생성될 때만 결과가 생성된다. 선행 기술은 GPCA 알고리즘을 안내하기 위하여 감독 사용자 개입을 취한다. 이 제약은 본 기술의 잠재성을 크게 제한한다.

<207> 본 발명은 노이즈 및 공통 차원의 존재시에 다수의 부분공간의 식별 및 분할을 강고하게 처리하기 위하여 GPCA 방법의 개념적 기초를 연장한다. 이러한 혁신은 종래 기술의 상태에 대해 기술의 감독되지 않는 개선을 제공한다.

<208> 종래 기술에서 GPCA는 이러한 법선 벡터의 탄젠트 공간에 관련 없이 베로네세 맵의 다항식의 법선 벡터에 대해 연산한다. 본 발명의 방법은 베로네세 맵에서 통상적으로 발견되는 법선 벡터의 공간에 직교하는 탄젠트 공간을 발견하기 위하여 GPCA를 연장한다. 이 "탄젠트 공간", 또는 베로네세 맵의 부분공간은, 그 후 베로네세 맵을 팩토링하기 위해 사용된다.

<209> 탄젠트 공간은 위치 및 탄젠트 평면 좌표축 사이의 르장드르(Legendre) 변환의 적용과 평면 웨이브 확장을 통해 식별되며, 이는 기하구조 객체의 표현에 있어서의 이중성, 구체적으로는 베로네세 맵의 다항식에 대한 법선의 탄젠트를 드러낸다. 이산 르장드르 변환은 법선 벡터에 대응하는 유도체의 제약된 형태를 정의하기 위하여 볼록 분석을 통해 적용된다. 이러한 접근법은 노이즈의 존재시에 법선 벡터의 계산에 의해 데이터 벡터를 분할하기 위해 사용된다. 이러한 볼록성 분석은 보다 강고한 알고리즘을 제공하기 위하여 GPCA와 일체화된다.

<210> 본 발명은 GPCA를 적용할 때 반복적인 팩토링화 접근법을 이용한다. 특히, 종래기술에서 발견되는 유도체 기반 구현은 본원에 기재된 아주 동일한 GPCA를 통해 분류된 데이터 벡터들의 총합을 정제하기 위해 확장된다. 반복적으로 적용되면, 이 기법은 베로네세 맵에서 후보 법선 벡터를 강고하게 검색하는 데 이용할 수 있으며, 그 후 확장된 GPCA 기법을 이용하여 이를 벡터들을 적합화한다. 팩토링화 단계에 대하여, 벡터의 정제된 집합과

연관된 원본 데이터는 원본 데이터 집합으로부터 제거된다. 나머지 데이터 집합은 혁신된 GPCA 기법으로 유사하게 분석될 수 있다. 이러한 혁신은 감독되지 않는 방법으로 GPCA 알고리즘을 이용하는 데 중요하다. 도 11은 데이터 벡터들의 반복적인 정체를 도시한다.

<211> GPCA 기법으로의 본 발명의 연장은 베로네세 다항식 벡터 공간에 다수의 근들이 존재하는 경우 더욱 유리하다. 또한, 선행기술 기법은 베로네세 맵에서의 법선들이 벡터 공간 축과 평행할 때 변질되는 경우를 겪지만, 본원 방법은 변질되지 않는다.

<212> 도 10은 기본적인 다항식 피팅 및 차분화의 방법을 도시한다.

하이브리드 공간 정규화 압축

<214> 본 발명은 비디오 스트림을 둘 이상의 "정규화된" 스트림으로 분할하는 것의 부가를 통해 블록 기반의 모션 예측된 코딩 방식의 효율성을 확장한다. 이러한 스트림들은 그 후 개별적으로 인코딩되어, 종래의 코덱의 해석 모션 가정이 유효하도록 한다. 정규화된 스트림을 디코딩할 때, 스트림들은 이들의 적절한 위치로 비정규화되고 원본 비디오 시퀀스를 생성하도록 서로 합성된다.

<215> 일 실시예에서, 하나 이상의 객체들이 비디오 스트림에서 검출되며 각각의 개별 객체들과 연관된 화소들이 비객체 화소를 남겨둔 채 이후에 분할된다. 다음으로, 글로벌 공간 모션 모델이 객체 및 비객체에 대해 생성된다. 글로벌 모델은 객체 및 비객체 화소를 공간적으로 정규화하기 위해 사용된다. 이러한 정규화는 비디오 스트림으로부터 비해석(non-translational) 모션을 효율적으로 제거했으며, 폐색 상호반응이 최소화된 비디오의 집합을 제공했다. 이들은 모두 본 발명의 방법의 유용한 특징이다.

<216> 공간적으로 정규화된 화소를 갖는 객체 및 비객체의 새로운 비디오는 종래의 블록 기반 압축 알고리즘에 입력으로써 제공된다. 비디오를 디코딩할 때, 글로벌 모션 모델 파라미터들이 사용되어 디코딩된 프레임을 비정규화하며, 객체 화소들은 서로 혼합되고 원본 비디오 스트림의 근사화를 생성하기 위하여 비객체 화소상에 혼합된다.

<217> 도 6에 도시된 것처럼, 하나 이상의 객체들(630&650)에 대한 이전에 검출된 객체 인스턴스(206&208)는 종래의 비디오 압축 방법(632)의 개별 인스턴스로 각각 프로세싱된다. 부가적으로, 객체의 분할(230)으로부터 생겨난 비객체(602)도 종래의 비디오 압축(632)을 이용하여 압축된다. 각각의 이러한 개별 압축 코딩(632)의 결과는 각각의 비디오 스트림에 개별적으로 대응하는 각각에 대한 개별적인 종래의 인코딩된 스트림(634)이다. 몇몇 점(point)에서는, 가능하게는 전송 후에는, 이러한 중간 인코딩된 스트림(234)은 정규화된 비객체의 합성(610) 및 다수의 객체들(638&658)로 압축해제될 수 있다(636). 이러한 합성된 화소들은 이들의 비정규화된 버전(622, 642&662)으로 비정규화(640)될 수 있어서, 합성 프로세스(670)가 객체 및 비객체 화소들을 결합하여 전체 프레임의 합성(672)을 형성할 수 있도록 서로에 대해 화소들을 공간적으로 정확히 위치시킨다.

하이브리드 코덱의 통합

<219> 본원발명에서 기재된 것처럼 종래의 블록 기반 압축 알고리즘과 정규화-분할 방식을 결합함에 있어서, 결과로 발생된 수 개의 본 발명의 방법이 존재한다. 주로, 특수화된 데이터 구조 및 요구되는 통신 프로토콜이 존재한다.

<220> 주된 데이터 구조들은 글로벌 공간 변형 파라미터 및 객체 분할 구체화 마스크를 포함한다. 주된 통신 프로토콜은 글로벌 공간 변형 파라미터와 객체 분할 구체화 마스크의 전송을 포함하는 층들이다.

도면의 간단한 설명

<23> 도 1은 종래 기술의 비디오 프로세싱 시스템을 도시하는 블록도이다.

<24> 도 2는 비디오를 프로세싱하는 주요한 모듈을 보여주는 본 발명의 개관을 제공하는 블록도이다.

<25> 도 3은 본 발명의 모션 추정 방법을 도시하는 블록도이다.

<26> 도 4는 본 발명의 광역 등록 방법을 도시하는 블록도이다.

<27> 도 5는 본 발명의 정규화 방법을 도시하는 블록도이다.

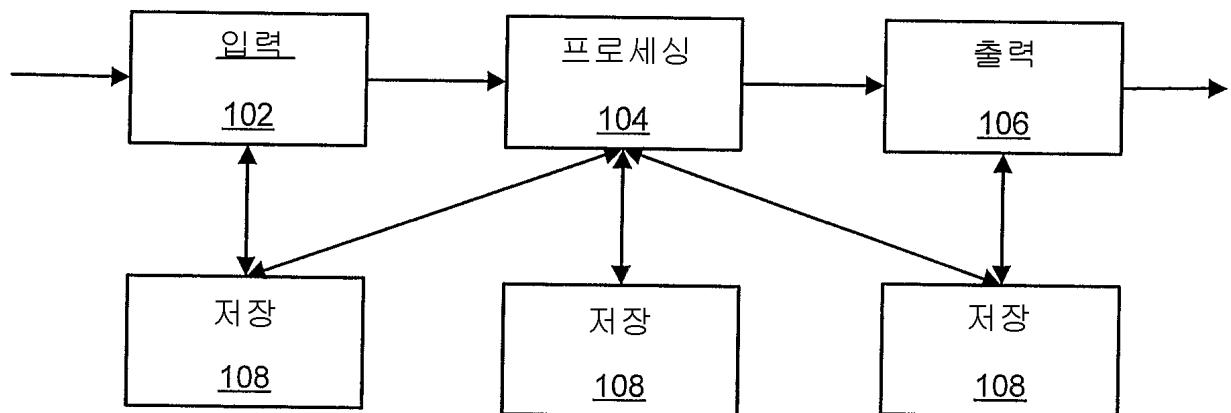
<28> 도 6은 하이브리드 공간 정규화 압축 방법을 도시하는 블록도이다.

<29> 도 7은 국부 정규화에서 사용되는 본 발명의 메쉬 생성 방법을 도시하는 블록도이다.

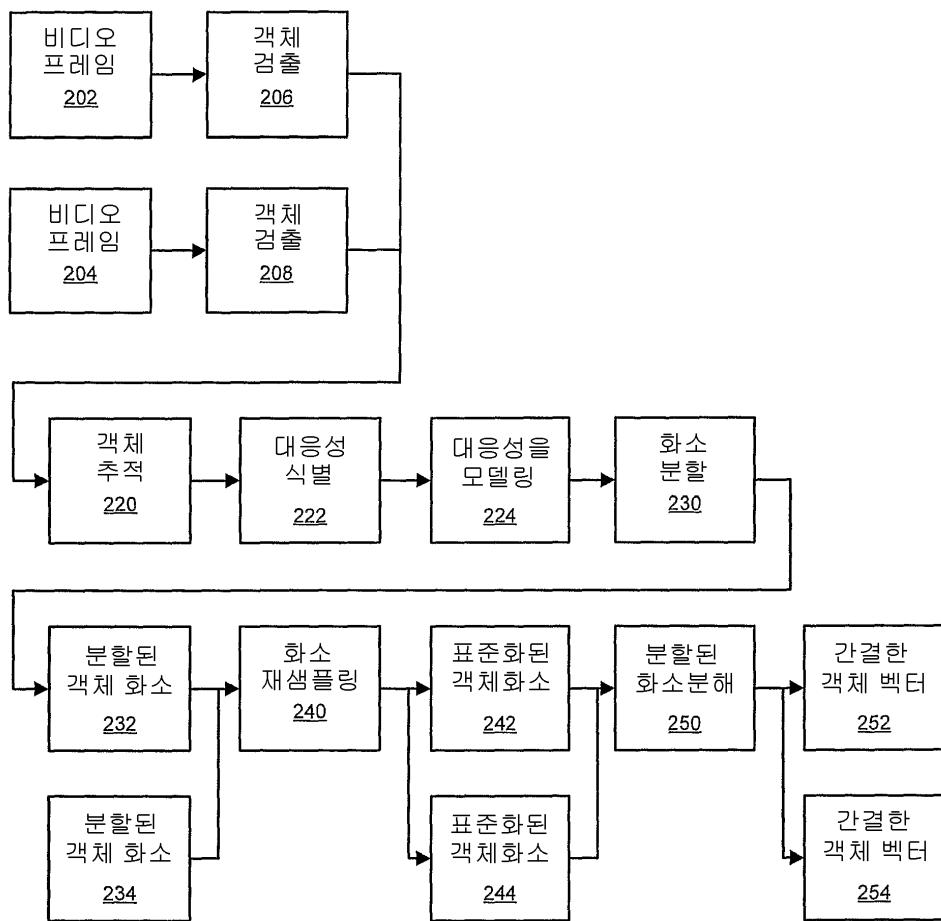
- <30> 도 8은 국부 정규화에서 사용되는 본 발명의 예상 기반의 정규화 방법을 도시하는 블록도이다.
- <31> 도 9는 본 발명의 결합된 전역 및 국부 정규화 방법을 도시하는 블록도이다.
- <32> 도 10은 본 발명의 GPCA-기초 다항식 맞춤(fitting) 및 차등화 방법을 도시하는 블록도이다.
- <33> 도 11은 본 발명의 반복적 GPCA 정제 방법을 도시하는 블록도이다.
- <34> 도 12는 배경 분해방법을 도시하는 블록도이다.
- <35> 도 13은 본 발명의 객체 분할 방법을 도시하는 블록도이다.

도면

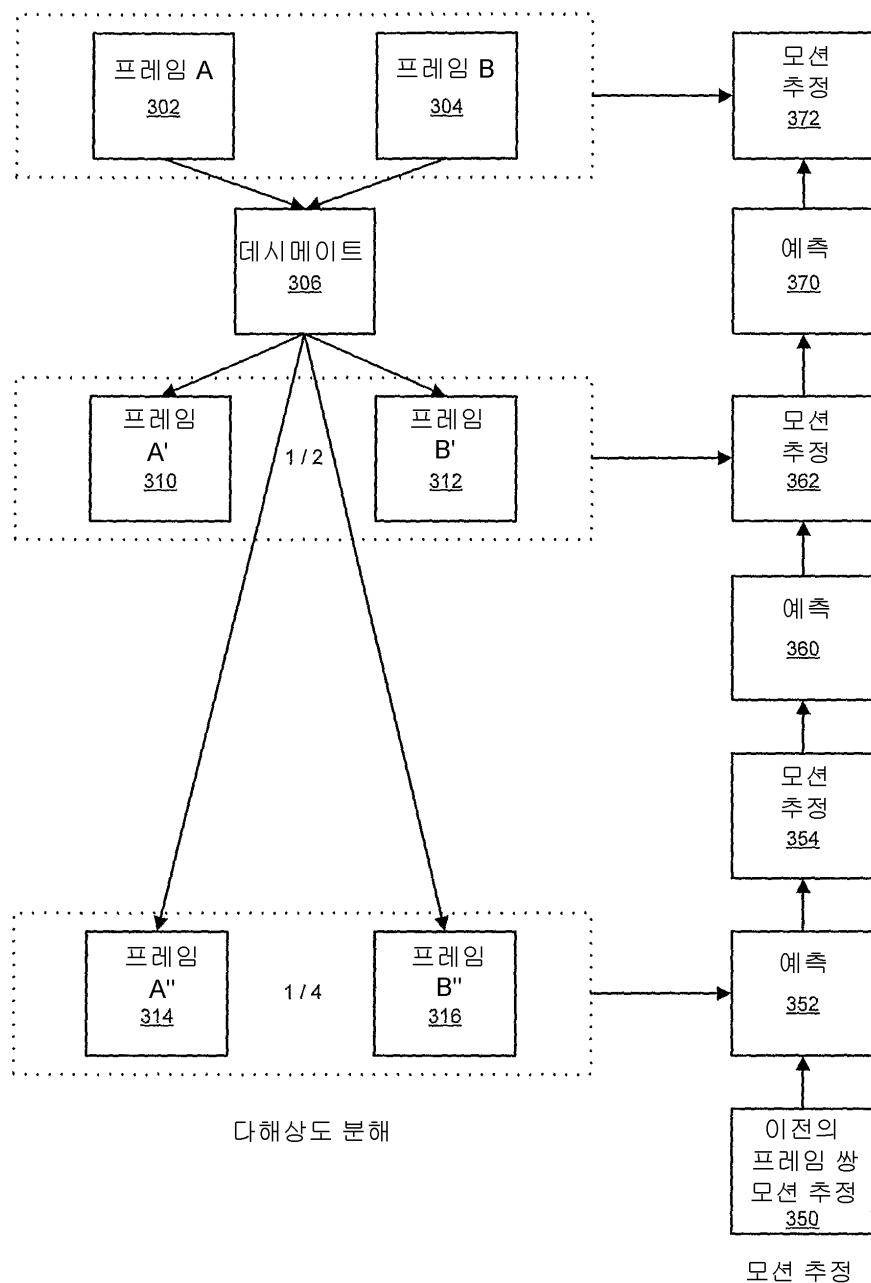
도면1



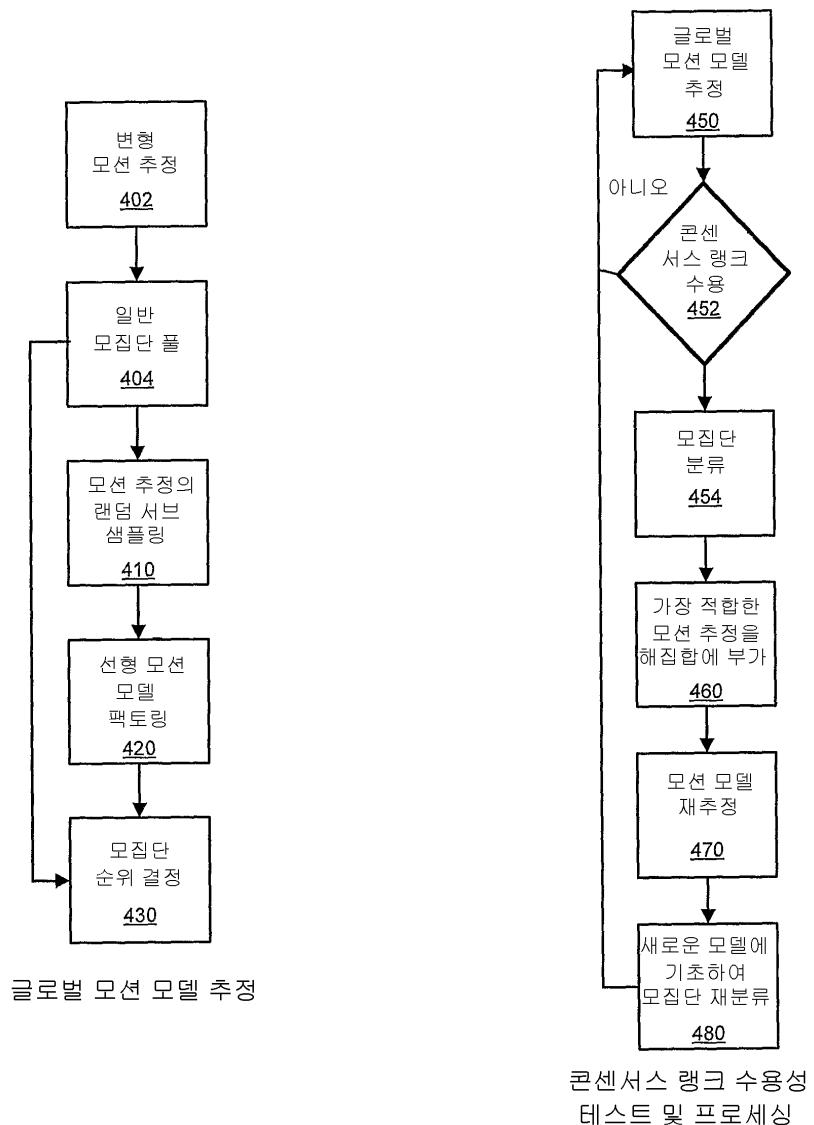
도면2



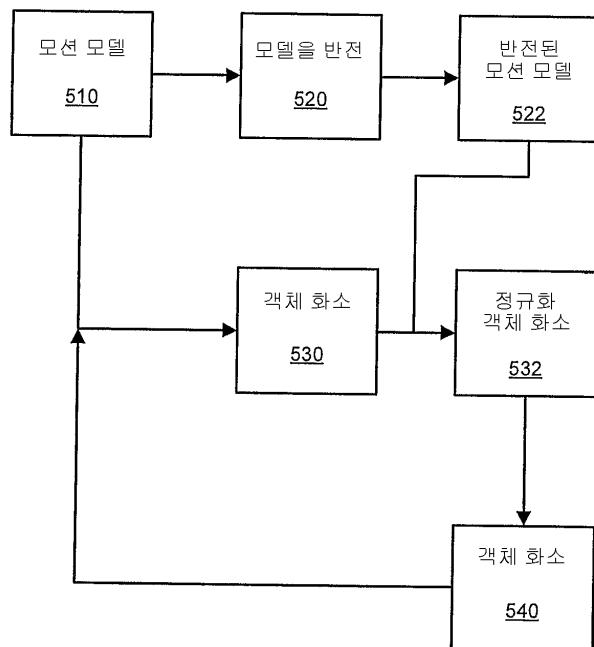
도면3



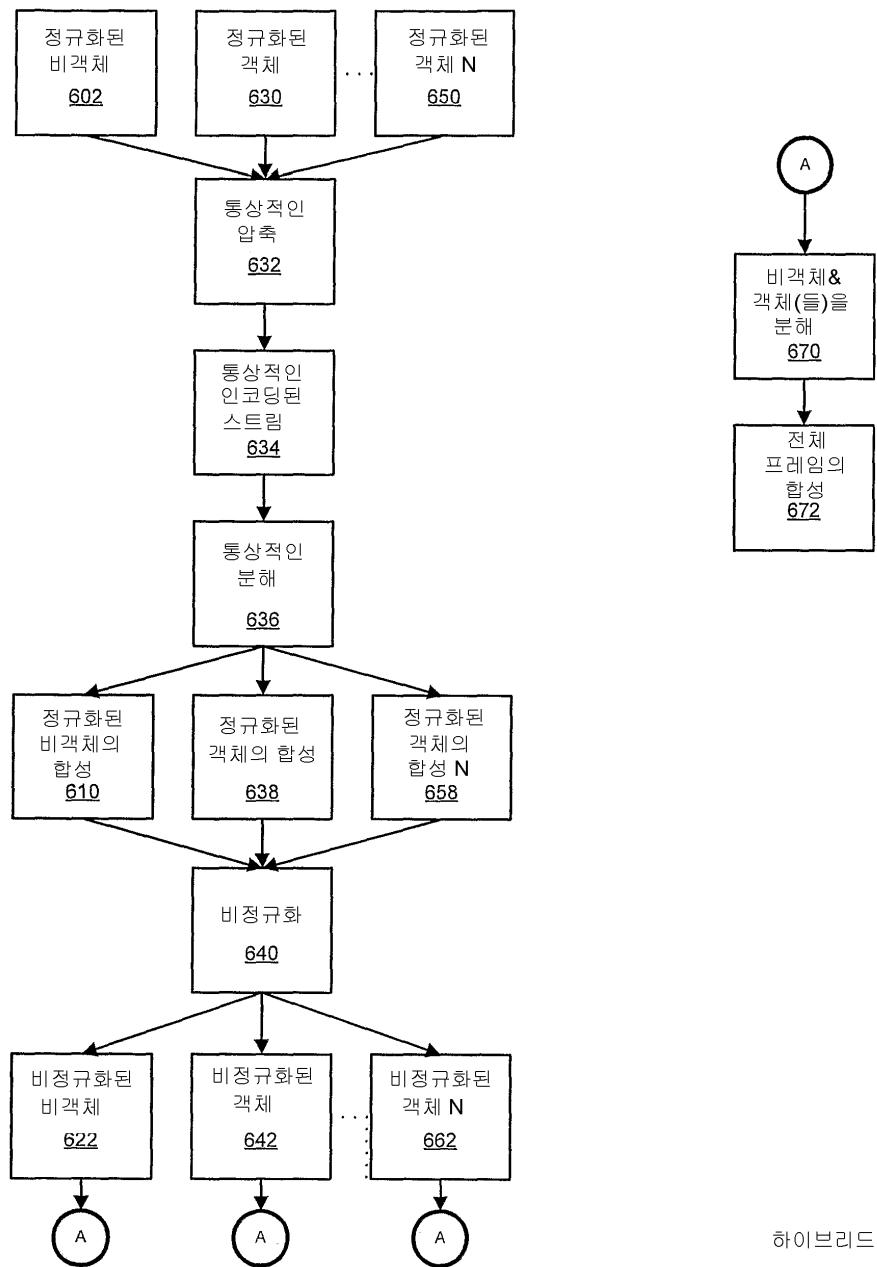
도면4



도면5

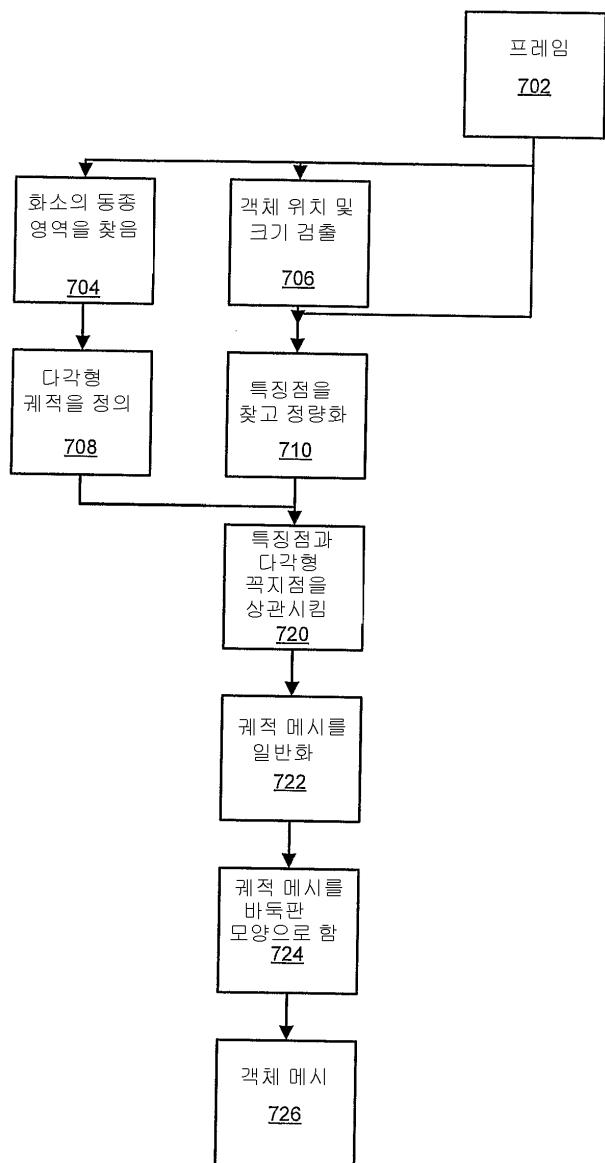


도면6

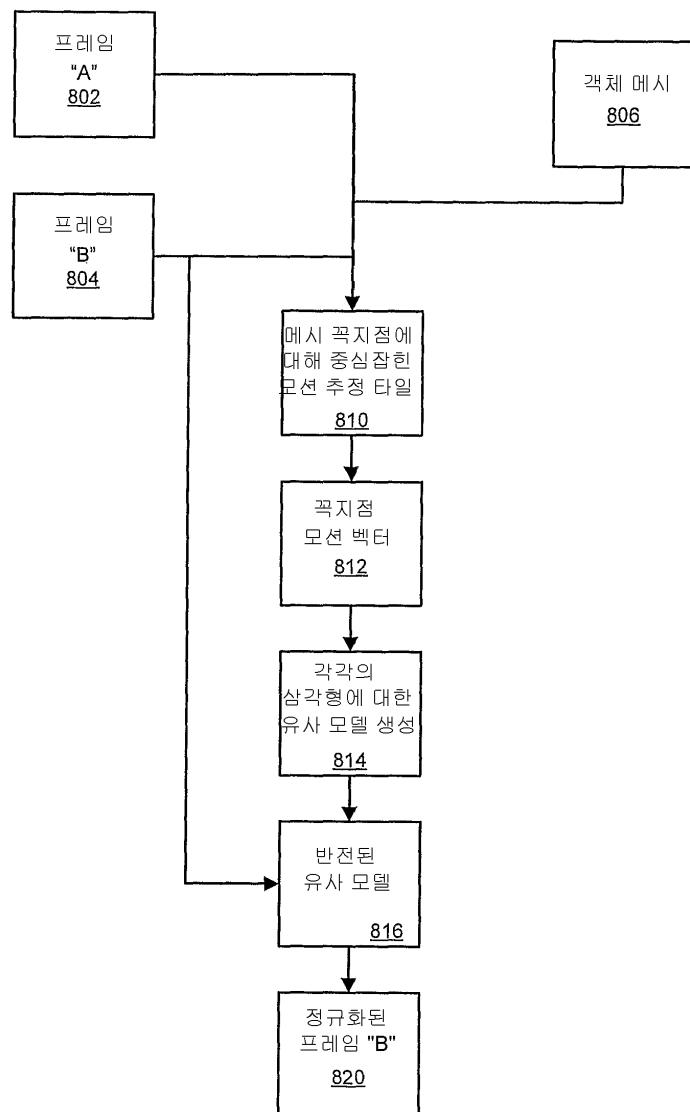


하이브리드

도면7

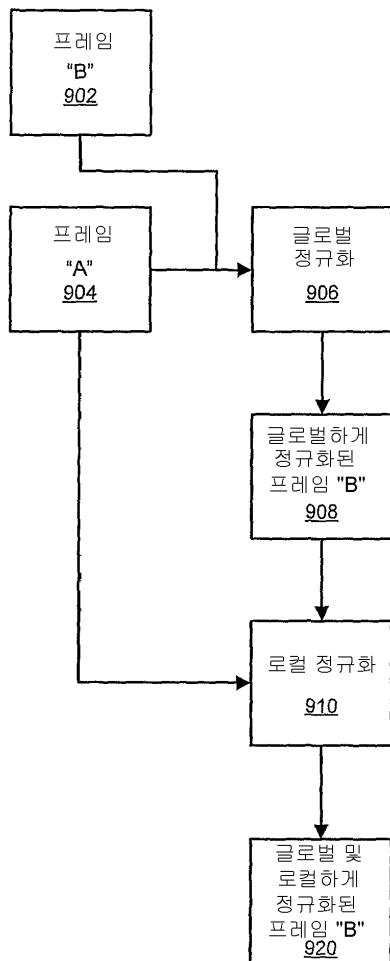
Local Normalization -
Mesh Generation

도면8



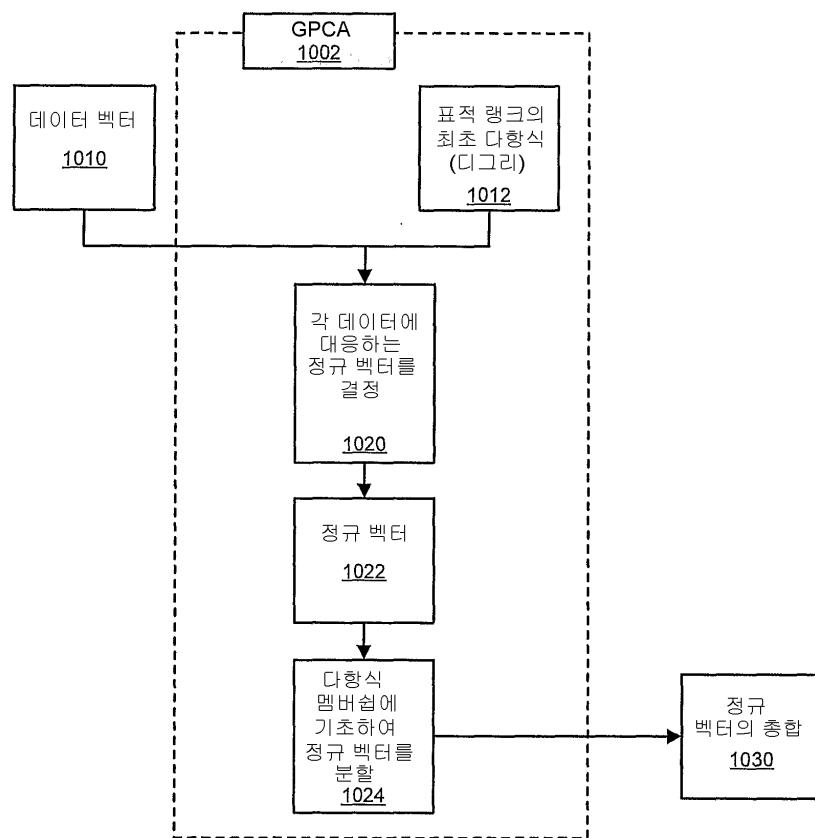
Local Normalization -
Mesh Based Normalization

도면9



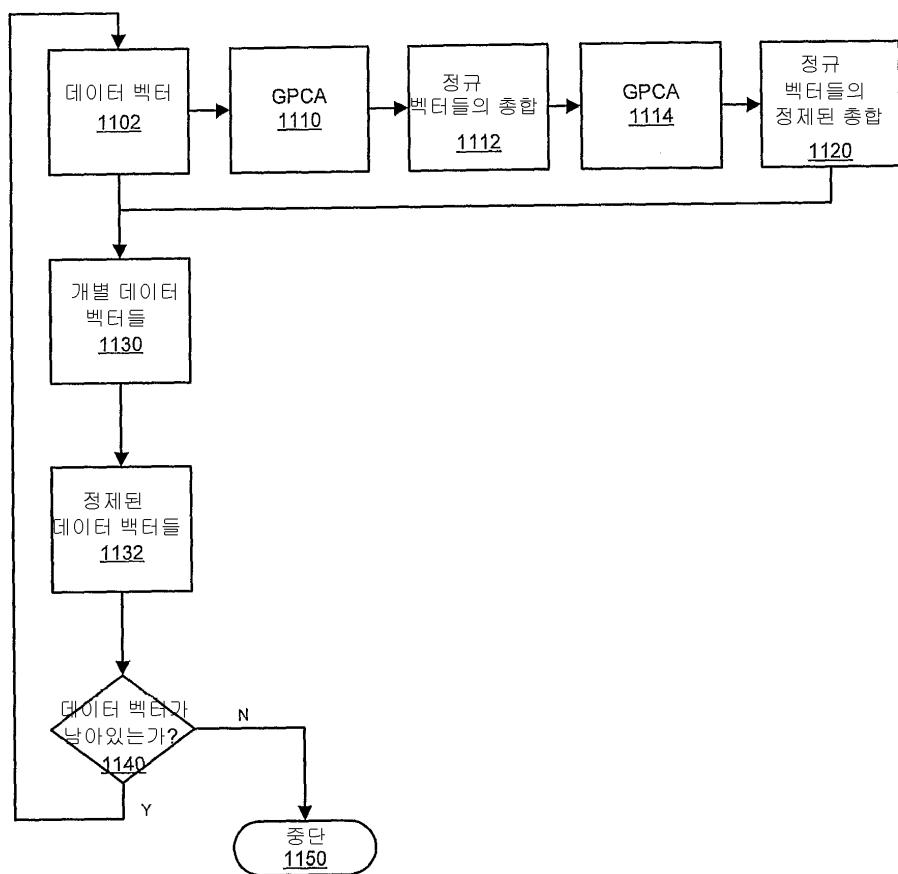
결합된 글로벌 & 로컬 정규화

도면10



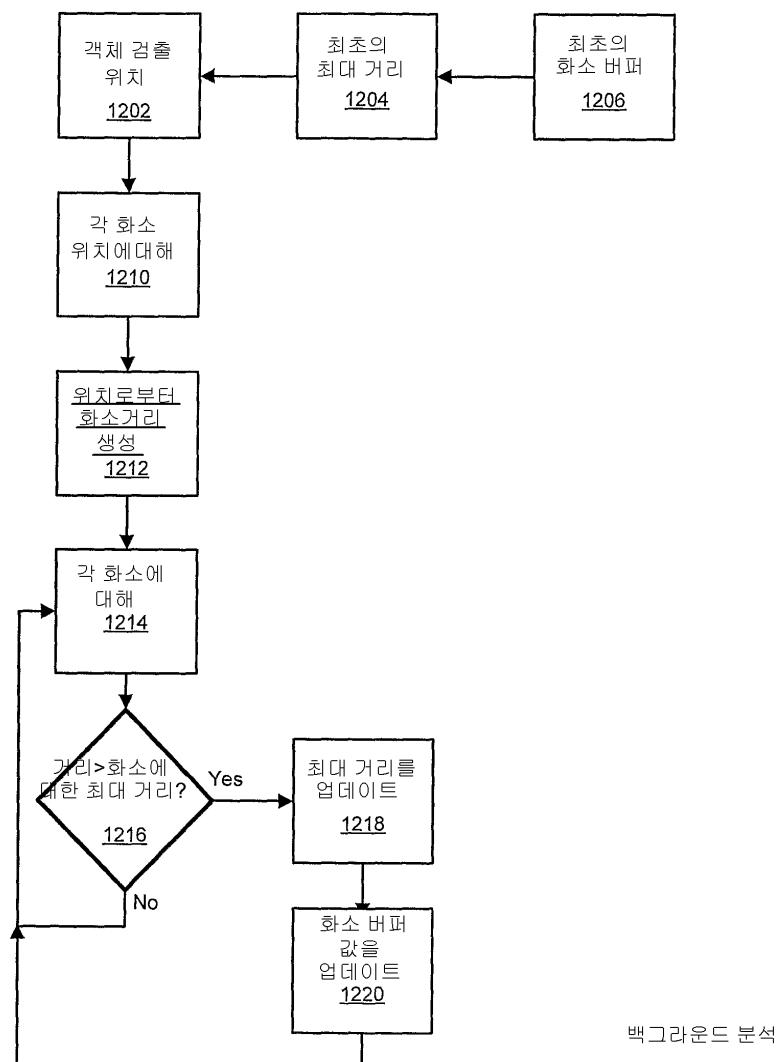
GPCA - 기본적인 다항식 피팅 & 미분

도면11



GPCA - 순환적인 GPCA 정제

도면12



백그라운드 분석

도면13

