



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2015년12월30일
(11) 등록번호 10-1581112
(24) 등록일자 2015년12월22일

(51) 국제특허분류(Int. Cl.)

G06T 7/00 (2006.01)

(21) 출원번호 10-2014-0035235

(22) 출원일자 2014년03월26일

심사청구일자 2014년03월26일

(65) 공개번호 10-2015-0111641

(43) 공개일자 2015년10월06일

(56) 선행기술조사문헌

논문1

논문2

KR1020120102738 A

논문3

(73) 특허권자

포항공과대학교 산학협력단

경상북도 포항시 남구 청암로 77 (지곡동)

(72) 발명자

김인수

경기도 광명시 목감로 372, 3층(광명동)

김대진

경상북도 포항시 남구 지곡로 155 교수아파트 7동 402호(지곡동)

(74) 대리인

특허법인이상

진체 청구항 수 : 총 19 항

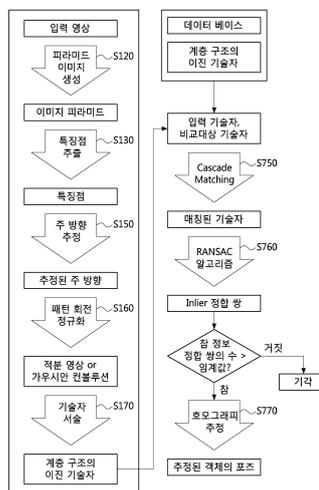
심사관 : 신재철

(54) 발명의 명칭 계층적 패턴 구조에 기반한 기술자 생성 방법 및 이를 이용한 객체 인식 방법과 장치

(57) 요약

계층적 패턴 구조 기반의 기술자 서술 방법 및 이를 이용한 이미지 내 객체의 포즈 추정 방법과 객체 인식 장치가 개시된다. 계층적 패턴 구조 기반의 기술자 서술 방법은, 입력된 분석 대상 이미지의 특징점에 대한 패치 영역에 기반하여 부모 노드를 정의하고, 부모 노드를 소정의 깊이까지 분할한 자식 노드를 정의하여 계층적 패턴 구조를 생성하는 단계와, 부모 노드와 자식 노드의 위치 좌표 및 픽셀 대표값에 기반하여 패치 영역의 주방향 벡터를 산출하는 단계와, 주방향 벡터에 기반하여 패치 영역의 회전각을 산출하고 계층적 패턴 구조를 회전각만큼 회전하는 단계를 포함하여 구성될 수 있다.

대표도 - 도11



이 발명을 지원한 국가연구개발사업

과제고유번호 ITAH0203130110010001000100100
부처명 미래창조과학부
연구관리전문기관 에스케이텔레콤(주)
연구사업명 IT명품인재양성사업
연구과제명 포스텍 미래 IT 융합연구원
기여율 1/2
주관기관 포항공과대학교 산학협력단
연구기간 2013.04.01 ~ 2014.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호 10040246
부처명 산업통상자원부
연구관리전문기관 전자부품연구원
연구사업명 산업원천기술개발사업(산업기술)
연구과제명 이동 로봇의 안정적 영상 획득을 통한 3D DEPTH 정보 획득과 실시간 객체 인식을 위한 로
봇 비전 SOC 및 모듈 개발
기여율 1/2
주관기관 포항공과대학교 산학협력단
연구기간 2013.06.01 ~ 2014.05.31

명세서

청구범위

청구항 1

이미지(image)로부터 객체를 인식하기 위해 기술자(descriptor)를 서술하는 방법에 있어서,

입력된 분석 대상 이미지의 특징점에 대한 패치(patch) 영역에 기반하여 부모 노드(parent node)를 정의하고, 상기 부모 노드를 소정의 깊이(depth)까지 분할한 자식 노드(child node)를 정의하여 계층적 패턴 구조(hierarchical structured pattern)를 생성하는 단계;

상기 부모 노드와 상기 자식 노드의 위치 좌표 및 픽셀(pixel) 대표값에 기반하여 상기 패치 영역의 주방향 벡터(master direction vector)를 산출하는 단계;

상기 주방향 벡터에 기반하여 상기 패치 영역의 회전각(rotation angle)을 산출하고 상기 계층적 패턴 구조를 상기 회전각만큼 회전하는 단계를 포함하는 계층적 패턴 구조 기반 기술자 서술 방법.

청구항 2

청구항 1에 있어서,

상기 계층적 패턴 구조를 생성하는 단계 이전에,

FAST 모서리 추출 기술, Harris Corner 검출 기술 중 적어도 하나를 이용하여 상기 특징점에 대한 패치 영역을 설정하는 단계를 더 포함하는 것을 특징으로 하는 계층적 패턴 구조 기반 기술자 서술 방법.

청구항 3

청구항 1에 있어서,

상기 회전각만큼 회전하는 단계 이후에,

상기 회전된 계층적 패턴 구조, 상기 부모 노드의 픽셀 대표값 및 상기 자식 노드의 픽셀 대표값에 기반하여 상기 기술자를 서술하는 단계를 더 포함하는 것을 특징으로 하는 계층적 패턴 구조 기반 기술자 서술 방법.

청구항 4

청구항 1에 있어서,

상기 계층적 패턴 구조는,

최상위 부모 노드를 소정의 노드수(N)만큼 분할하여 최초 자식 노드가 정의되고, 상기 최초 자식 노드는 차상위 부모 노드가 되어 상기 소정의 깊이 단계마다 자식 노드를 정의하는 재귀적(recursive) 방법으로 생성하되,

상기 소정의 노드수(N)는 자연수이고, 상기 분할은 리던던시(redundancy) 또는 익셉션(exception)이 있는 분할인 것을 특징으로 하는 계층적 패턴 구조 기반 기술자 서술 방법.

청구항 5

청구항 4에 있어서,

상기 재귀적 방법은,

상기 최상위 부모 노드의 반지름(radius)을 지름(diameter)으로 하고, 상기 최상위 부모 노드를 균등 분할하여 상기 소정의 노드수(N)의 최초 자식 노드를 생성하는 제1 분할방법을 수행하거나,

상기 소정의 노드수 N에서 1을 뺀 수만큼 상기 제1 분할방법을 수행한 이후에 상기 최상위 부모 노드와 동심원(concentric circle)인 최초 자식 노드 하나를 추가로 생성하는 제2 분할방법을 수행하거나, 또는

상기 최상위 부모 노드의 반지름보다 긴 길이의 지름으로 하고, 상기 최상위 부모 노드를 균등 분할하여 상기 소정의 노드수(N)의 최초 자식 노드를 생성하는 제3 분할방법을 수행하는 것을 특징으로 하는 계층적 패턴 구조

기본 기술자 서술 방법.

청구항 6

청구항 4에 있어서,

상기 소정의 노드수 및 상기 소정의 깊이는,

상기 특징점에 기반하여 확장 또는 축소되는 것을 특징으로 하는 계층적 패턴 구조 기반 기술자 서술 방법.

청구항 7

청구항 1에 있어서,

상기 픽셀 대표값은,

상기 부모 노드와 상기 자식 노드에 해당하는 영역의 픽셀값들의 평균이거나, 상기 영역을 가우시안 컨벌루션 (Gaussian Convolution)하여 가중치를 적용한 픽셀값인 것을 특징으로 하는 계층적 패턴 구조 기반 기술자 서술 방법.

청구항 8

청구항 1에 있어서,

상기 계층적 패턴 구조를 생성하는 단계 이전에,

상기 이미지를 실수배(real number multiple)하여 크기가 다른 복수의 이미지로 구성된 피라미드 영상 공간을 생성하고, 상기 피라미드 영상 공간 내의 이미지를 상기 분석 대상 이미지로 출력하는 단계를 더 포함하는 것을 특징으로 하는 계층적 패턴 구조 기반 기술자 서술 방법.

청구항 9

계층적 패턴 구조의 기술자를 이용하여 이미지 내 객체의 포즈를 추정하는 방법에 있어서,

상기 기술자를 소정의 개수(D)만큼의 군집으로 구분하는 단계;

상기 기술자의 i번째 군집과 미리 저장된 참조 기술자의 i번째 군집 간의 거리를 계산하는 단계;

상기 계산된 거리가 소정의 임계값보다 작다는 판단에 따라 상기 기술자의 (i+1)번째 군집과 미리 저장된 참조 기술자의 (i+1)번째 군집 간의 거리를 계산하는 단계; 및

상기 기술자의 D번째 군집과 상기 미리 저장된 참조 기술자의 D번째 군집 간의 거리가 소정의 임계값보다 작다는 판단에 따라 상기 기술자와 상기 미리 저장된 참조 기술자를 매칭하는 단계를 포함하되,

상기 i는 1을 초항(first term)으로 (D-1)까지 반복(iteration)되고,

상기 소정의 개수(D)는 상기 계층적 패턴 구조의 기술자의 대상인 특징점에 대한 패치 영역을 자식 노드로 분할한 깊이를 의미하는 것을 특징으로 하는 이미지 내 객체의 포즈 추정 방법.

청구항 10

청구항 9에 있어서,

상기 기술자의 i번째 군집과 상기 미리 저장된 참조 기술자의 i번째 군집 간의 거리가 소정의 임계값보다 크거나 같은 경우, 상기 (i+1)번째 군집 간의 거리를 계산하는 단계를 생략하고 상기 기술자와 상기 미리 저장된 참조 기술자와의 매칭을 스킵(skip)하는 것을 특징으로 하는 이미지 내 객체의 포즈 추정 방법.

청구항 11

청구항 9에 있어서,

상기 기술자를 매칭하는 단계 이후에,

상기 객체와 상기 미리 저장된 참조 기술자의 대상인 참조 객체 간의 호모그래피(homography)를 추정하는 단계

를 더 포함하는 것을 특징으로 하는 이미지 내 객체의 포즈 추정 방법.

청구항 12

청구항 11에 있어서,

상기 호모그래피를 추정하는 단계 이전에,

상기 호모그래피에 기반하여 상기 객체의 인식 여부를 판단하는 단계를 더 포함하되,

상기 인식 여부를 판단하는 방법은,

최종 매칭된 기술자들 중 RANSAC(RANdom Sample Consensus) 알고리즘을 통해 산출된 인라이어(inlier)의 개수에 기반하여 상기 객체의 인식 여부를 판단하는 제1 판단 방법,

상기 호모그래피의 정확도에 기반하여 상기 객체의 인식 여부를 판단하는 제2 판단 방법,

상기 객체를 상기 호모그래피에 기반하여 아핀 변환(affine transform)한 영역과 상기 참조 객체의 영역을 동일한 포즈로 정규화하여 템플릿 매칭한 결과 정보에 기반하여 상기 객체의 인식 여부를 판단하는 제3 판단 방법 중 어느 하나인 것을 특징으로 하는 이미지 내 객체의 포즈 추정 방법.

청구항 13

삭제

청구항 14

계층적 패턴 구조 기반의 기술자를 이용하여 이미지로부터 객체를 인식하는 장치에 있어서,

입력된 분석 대상 이미지의 특징점에 대한 패치 영역에 기반하여 부모 노드를 정의하고, 상기 부모 노드를 소정의 깊이까지 분할한 자식 노드를 정의하여 계층적 패턴 구조를 생성하는 패턴 생성부;

상기 부모 노드와 상기 자식 노드의 위치 좌표 및 픽셀 대표값에 기반하여 상기 패치 영역의 주방향 벡터를 산출하는 주방향 산출부;

상기 주방향 벡터에 기반하여 상기 패치 영역의 회전각을 산출하고 상기 계층적 패턴 구조를 상기 회전각만큼 회전하는 패턴 회전부를 포함하는 계층적 패턴 구조 기반 객체 인식 장치.

청구항 15

청구항 14에 있어서,

FAST 모서리 추출 기술, Harris Corner 검출 기술 중 적어도 하나를 이용하여 상기 특징점에 대한 패치 영역을 설정하는 영역 설정부를 더 포함하는 것을 특징으로 하는 계층적 패턴 구조 기반 객체 인식 장치.

청구항 16

청구항 14에 있어서,

상기 회전된 계층적 패턴 구조, 상기 부모 노드의 픽셀 대표값 및 상기 자식 노드의 픽셀 대표값에 기반하여 상기 기술자를 서술하는 기술자 서술부를 더 포함하는 것을 특징으로 하는 계층적 패턴 구조 기반 객체 인식 장치.

청구항 17

청구항 14에 있어서,

상기 이미지를 실수배(real number multiple)하여 크기가 다른 복수의 이미지로 구성된 피라미드 영상 공간을 생성하고, 상기 피라미드 영상 공간 내의 이미지를 상기 분석 대상 이미지로 출력하는 피라미드 영상 생성부를 더 포함하는 것을 특징으로 하는 계층적 패턴 구조 기반 객체 인식 장치.

청구항 18

청구항 14에 있어서,

상기 기술자를 상기 소정의 깊이만큼의 군집으로 구분하여 상기 기술자와 미리 저장된 참조 기술자 간의 거리를 상기 군집 단위에 기반하여 계산하고 계산된 결과에 기반하여 상기 기술자와 상기 미리 저장된 참조 기술자의 매칭 여부를 결정하는 기술자 매칭부를 더 포함하는 것을 특징으로 하는 계층적 패턴 구조 기반 객체 인식 장치.

청구항 19

청구항 18에 있어서,

상기 객체와 상기 미리 저장된 참조 기술자의 대상인 참조 객체 간의 호모그래피(homography)를 추정하는 호모그래피 추정부를 더 포함하는 것을 특징으로 하는 계층적 패턴 구조 기반 객체 인식 장치.

청구항 20

청구항 19에 있어서,

상기 호모그래피에 기반하여 상기 객체의 인식 여부를 판단하는 객체 인식부를 더 포함하되,

상기 인식 여부를 판단하는 방법은,

최종 매칭된 기술자들 중 RANSAC(RANDom Sample Consensus) 알고리즘을 통해 산출된 인라이어(inlier)의 개수에 기반하여 상기 객체의 인식 여부를 판단하는 제1 판단방법,

상기 호모그래피의 정확도에 기반하여 상기 객체의 인식 여부를 판단하는 제2 판단 방법, 또는

상기 객체를 상기 호모그래피에 기반하여 아핀 변환(affine transform)한 영역과 상기 참조 객체의 영역을 동일한 포즈로 정규화하여 템플릿 매칭한 결과 정보에 기반하여 상기 객체의 인식 여부를 판단하는 제3 판단 방법 중 어느 하나인 것을 특징으로 하는 계층적 패턴 구조 기반 객체 인식 장치.

발명의 설명

기술 분야

[0001] 본 발명은 컴퓨터 비전(vision)의 객체 인식 기술에 관한 것으로, 더욱 자세하게는 계층적 구조를 갖는 기술자(descriptor)를 고안하여 입력 영상으로부터 객체를 인식하는 방법 및 장치에 관한 것이다.

[0002] 또한 이를 위하여 입력 영상의 특징점을 분석하는 기술로서 특징점의 패치(patch) 영역을 계층적으로 분할하여 기술자를 생성하는 기술에 관한 것이다.

배경 기술

[0003] 컴퓨터 비전(computer vision)은 기계의 시각에 해당하는 부분을 연구하는 컴퓨터 과학의 최신 연구 분야에 속한다. 더 넓은 카테고리로는 인공 지능(AI, Artificial Intelligence)의 한 분야로서 컴퓨터를 사용하여 인간의 시각적인 인식 능력 일반을 재현하는 연구 분야라고 할 수 있다.

[0004] 컴퓨터 비전에서 필요한 기술로는 디지털 영상 처리, 기계 학습, 패턴 인식, 물체 인식 등이 있다. 이 중 물체 인식은 미리 학습을 한 지식정보를 바탕으로, 물체의 영상을 보고, 물체의 종류, 크기, 방향, 위치 등의 공간 정보를 실시간으로 알아내는 기술을 말한다. 이는 로봇 분야뿐만 아니라 컴퓨터 과학 분야 전체의 도전과제이다.

[0005] 이러한 컴퓨터 비전 분야에서 객체 인식(object recognition)을 위한 객체 표현 방식인 특징점 생성 기술 및 기술자 생성 기술은 물체 인식, 물체 검색, 지능형 감시 시스템, 지능형 로봇, 가상현실 및 기타 여러 응용 분야에서 기술적 수요가 증가하고 있다. 나아가, 모바일 시장의 증대에 따라 저 메모리 용량의 고인식 성능을 갖는 기술자에 대한 요구가 크게 대두되고 있다. 이와 같은 고효율의 객체를 표현하기 위한 기술자는 객체의 가려짐, 조명, 회전 등의 다양한 변화에 대해 강인한 인식 성능을 보이고 실시간 처리가 가능해야 한다.

[0006] 기존 연구에서 객체를 인식, 표현하기 위한 알고리즘으로 SIFT(Scale-Invariant Feature Transform), SURF(Seeded Up Robust Features) 등이 제안되어 활용되었다. 그렇지만 복잡한 연산으로 실시간 처리가 어려운 단점이 있었다. 최근에는 BRIEF(Binary Robust Independent Elementary Features) 이후로, ORB(Oriented FAST and Rotated BRIEF), BRISK(Binary Robust Invariant Scalable Key points), FREAK(Fast Retina Key point)

등의 이진 기술자가 제안되었으며, 기술자로서 물체 인식 성능과 속도 면에서 큰 향상을 보였다. 그러나, ORB와 BRISK의 경우 학습에 따른 특징 선택 또는 특징의 무작위 선택으로 인한 일반적인 객체의 표현에 한계를 보이며, FREAK는 빠른 속도로 처리가 가능하지만 인식 성능이 확실하게 입증되지 않았다.

발명의 내용

해결하려는 과제

- [0007] 본 발명은 영상 내 특정 물체, 얼굴, 사람 등의 의미정보를 포함하는 식별 가능한 대상에 대한 영역을 카메라 입력 영상 내 대상 객체의 위치 및 포즈를 추정하는 방법 및 장치를 제공하는 데에 그 목적이 있다.
- [0008] 이를 위하여 피라미드 영상을 생성하여 코너(corner)와 같은 특징점을 생성하고, 특징점을 중심으로 하는 패치 영역에서 주 방향을 추정하고 계층적 패턴 구조를 회전 정규화하며, 특징점을 중심으로 회전 정규화된 계층적 패턴 구조를 이용하여 제안하는 방식으로 기술된 기술자를 생성하는 방법을 제공하는 것도 본 발명의 다른 목적이다.
- [0009] 또한 본 발명의 또 다른 목적은 상술한 계층적 패턴 구조 기반의 기술자를 이용하여 사전에 저장(database)된 대상 객체 기술자와 cascade 기술자 매칭(matching)을 수행하고, 정합된 기술자들의 2차원 위치를 기반으로 객체 간 기하학적 변환 관계(homography)를 추정하는 방법을 제공하는 데 있다.

과제의 해결 수단

- [0010] 상술한 본 발명의 목적을 달성하기 위한 일 측면에 따른 계층적 패턴 구조 기반 기술자 서술 방법은, 이미지(image)로부터 객체를 인식하기 위해 기술자(descriptor)를 서술하는 방법에 있어서, 입력된 분석 대상 이미지의 특징점에 대한 패치(patch) 영역에 기반하여 부모 노드(parent node)를 정의하고, 상기 부모 노드를 소정의 깊이(depth)까지 분할한 자식 노드(child node)를 정의하여 계층적 패턴 구조(hierarchical structured pattern)를 생성하는 단계와, 부모 노드와 자식 노드의 위치 좌표 및 픽셀(pixel) 대표값에 기반하여 패치 영역의 주방향 벡터(master direction vector)를 산출하는 단계와, 주방향 벡터에 기반하여 패치 영역의 회전각(rotation angle)을 산출하고 계층적 패턴 구조를 회전각만큼 회전하는 단계를 포함하여 구성될 수 있다.
- [0011] 계층적 패턴 구조 기반 기술자 서술 방법은, 계층적 패턴 구조를 생성하는 단계 이전에, FAST 모서리 추출 기술, Harris Corner 검출 기술 중 적어도 하나를 이용하여 특징점에 대한 패치 영역을 설정하는 단계를 더 포함하여 구성될 수 있다.
- [0012] 계층적 패턴 구조 기반 기술자 서술 방법은, 회전각만큼 회전하는 단계 이후에, 회전된 계층적 패턴 구조, 부모 노드의 픽셀 대표값 및 자식 노드의 픽셀 대표값에 기반하여 기술자를 서술하는 단계를 더 포함하여 구성될 수 있다.
- [0013] 상기 계층적 패턴 구조는, 최상위 부모 노드를 소정의 노드수(N)만큼 분할하여 최초 자식 노드가 정의되고, 최초 자식 노드는 차상위 부모 노드가 되어 소정의 깊이 단계마다 자식 노드를 정의하는 재귀적(recursive) 방법으로 생성되며, 소정의 노드수(N)는 자연수이고, 분할은 리던던시(redundancy) 또는 익셉션(exception)이 있는 분할일 수 있다. 나아가 상기 재귀적 방법은, 최상위 부모 노드의 반지름(radius)을 지름(diameter)으로 하고, 최상위 부모 노드를 균등 분할하여 소정의 노드수(N)의 최초 자식 노드를 생성하는 제1 분할방법을 수행하거나, 소정의 노드수 N에서 1을 뺀 수만큼 제1 분할방법을 수행한 이후에 최상위 부모 노드와 동심원(concentric circle)인 최초 자식 노드 하나를 추가로 생성하는 제2 분할방법을 수행하거나, 또는 최상위 부모 노드의 반지름보다 긴 길이의 지름으로 하고, 최상위 부모 노드를 균등 분할하여 소정의 노드수(N)의 최초 자식 노드를 생성하는 제3 분할방법을 수행할 수 있다.
- [0014] 상기 소정의 노드수 및 상기 소정의 깊이는, 특징점에 기반하여 확장 또는 축소될 수 있다.
- [0015] 상기 픽셀 대표값은, 부모 노드와 자식 노드에 해당하는 영역의 픽셀값들의 평균이거나, 영역을 가우시안 컨벌루션(Gaussian Convolution)하여 가중치를 적용한 픽셀값일 수 있다.
- [0016] 계층적 패턴 구조 기반 기술자 서술 방법은, 계층적 패턴 구조를 생성하는 단계 이전에, 이미지를 실수배(real number multiple)하여 크기가 다른 복수의 이미지로 구성된 피라미드 영상 공간을 생성하고, 피라미드 영상 공간 내의 이미지를 분석 대상 이미지로 출력하는 단계를 더 포함하여 구성될 수 있다.
- [0017] 상술한 본 발명의 목적을 달성하기 위한 다른 측면에 따른 이미지 내 객체의 포즈 추정 방법은, 계층적 패턴 구

조의 기술자를 이용하여 이미지 내 객체의 포즈를 추정하는 방법에 있어서, 기술자를 소정의 개수(D)만큼의 군집으로 구분하는 단계와, 기술자의 i번째 군집과 미리 저장된 참조 기술자의 i번째 군집 간의 거리를 계산하는 단계와, 계산된 거리가 소정의 임계값보다 작다는 판단에 따라 기술자의 (i+1)번째 군집과 미리 저장된 참조 기술자의 (i+1)번째 군집 간의 거리를 계산하는 단계와, 기술자의 D번째 군집과 미리 저장된 참조 기술자의 D번째 군집 간의 거리가 소정의 임계값보다 작다는 판단에 따라 기술자와 미리 저장된 참조 기술자를 매칭하는 단계를 포함하되, i는 1을 초항(first term)으로 (D-1)까지 반복(iteration)될 수 있다.

[0018] 상기 이미지 내 객체의 포즈 추정 방법은, 기술자의 i번째 군집과 미리 저장된 참조 기술자의 i번째 군집 간의 거리가 소정의 임계값보다 크거나 같은 경우, (i+1)번째 군집 간의 거리를 계산하는 단계를 생략하고 기술자와 미리 저장된 참조 기술자와의 매칭을 스킵(skip)할 수 있다.

[0019] 상기 이미지 내 객체의 포즈 추정 방법은, 기술자를 매칭하는 단계 이후에, 객체와 미리 저장된 참조 기술자의 대상인 참조 객체 간의 호모그래피(homography)를 추정하는 단계를 더 포함하여 구성될 수 있다. 나아가 상기 호모그래피를 추정하는 단계 이전에, 호모그래피에 기반하여 객체의 인식 여부를 판단하는 단계를 더 포함하되, 인식 여부를 판단하는 방법은, 최종 매칭된 기술자들 중 RANSAC(RANdom Sample Consensus) 알고리즘을 통해 산출된 인라이어(inlier)의 개수에 기반하여 객체의 인식 여부를 판단하는 제1 판단 방법, 호모그래피의 정확도에 기반하여 객체의 인식 여부를 판단하는 제2 판단 방법, 객체를 호모그래피에 기반하여 아핀 변환(affine transform)한 영역과 참조 객체의 영역을 동일한 포즈로 정규화하여 템플릿 매칭한 결과 정보에 기반하여 객체의 인식 여부를 판단하는 제3 판단 방법 중 어느 하나일 수 있다.

[0020] 상기 소정의 개수(D)는, 계층적 패턴 구조의 기술자의 대상인 특징점에 대한 패치 영역을 자식 노드로 분할한 깊이와 같을 수 있다.

[0021] 상술한 본 발명의 목적을 달성하기 위한 또 다른 측면에 따른 계층적 패턴 구조 기반 객체 인식 장치는, 계층적 패턴 구조 기반의 기술자를 이용하여 이미지로부터 객체를 인식하는 장치에 있어서, 입력된 분석 대상 이미지의 특징점에 대한 패치 영역에 기반하여 부모 노드를 정의하고, 부모 노드를 소정의 깊이까지 분할한 자식 노드를 정의하여 계층적 패턴 구조를 생성하는 패턴 생성부와, 부모 노드와 자식 노드의 위치 좌표 및 픽셀 대표값에 기반하여 패치 영역의 주방향 벡터를 산출하는 주방향 산출부와, 주방향 벡터에 기반하여 패치 영역의 회전각을 산출하고 계층적 패턴 구조를 회전각만큼 회전하는 패턴 회전부를 포함하여 구성될 수 있다.

[0022] 상기 계층적 패턴 구조 기반 객체 인식 장치는, FAST 모서리 추출 기술, Harris Corner 검출 기술 중 적어도 하나를 이용하여 특징점에 대한 패치 영역을 설정하는 영역 설정부를 더 포함하여 구성될 수 있다.

[0023] 상기 계층적 패턴 구조 기반 객체 인식 장치는, 회전된 계층적 패턴 구조, 상기 부모 노드의 픽셀 대표값 및 자식 노드의 픽셀 대표값에 기반하여 기술자를 서술하는 기술자 서술부를 더 포함하여 구성될 수 있다.

[0024] 상기 계층적 패턴 구조 기반 객체 인식 장치는, 이미지를 실수배(real number multiple)하여 크기가 다른 복수의 이미지로 구성된 피라미드 영상 공간을 생성하고, 피라미드 영상 공간 내의 이미지를 분석 대상 이미지로 출력하는 피라미드 영상 생성부를 더 포함하여 구성될 수 있다.

[0025] 상기 계층적 패턴 구조 기반 객체 인식 장치는, 기술자를 소정의 깊이만큼의 군집으로 구분하여 기술자와 미리 저장된 참조 기술자 간의 거리를 군집 단위에 기반하여 계산하고 계산된 결과에 기반하여 기술자와 미리 저장된 참조 기술자의 매칭 여부를 결정하는 기술자 매칭부를 더 포함하여 구성될 수 있다. 나아가 객체와 미리 저장된 참조 기술자의 대상인 참조 객체 간의 호모그래피(homography)를 추정하는 호모그래피 추정부를 더 포함하여 구성될 수 있다. 더 나아가 호모그래피에 기반하여 객체의 인식 여부를 판단하는 객체 인식부를 더 포함하되, 인식 여부를 판단하는 방법은, 최종 매칭된 기술자들 중 RANSAC(RANdom Sample Consensus) 알고리즘을 통해 산출된 인라이어(inlier)의 개수에 기반하여 객체의 인식 여부를 판단하는 제1 판단방법, 호모그래피의 정확도에 기반하여 객체의 인식 여부를 판단하는 제2 판단 방법, 또는 객체를 호모그래피에 기반하여 아핀 변환(affine transform)한 영역과 참조 객체의 영역을 동일한 포즈로 정규화하여 템플릿 매칭한 결과 정보에 기반하여 객체의 인식 여부를 판단하는 제3 판단 방법 중 어느 하나일 수 있다.

발명의 효과

[0026] 상기와 같은 본 발명에 따른 계층적 패턴 구조에 기반한 기술자 서술 방법 및 이를 이용한 이미지 내 객체의 포즈 추정 방법과 객체 인식 장치를 사용하면, 객체의 가려짐, 조명, 회전 등의 다양한 변화에 대해 강인한 인식 성능을 보이는 기술자를 생성할 수 있다는 장점이 있다. 나아가 보다 단순해진 연산으로 실시간 처리가 가능하

며, 객체 인식 속도 면에서도 기존의 방법들과 비교하여 큰 향상이 가능하다. 이는 기존 방법들의 일반적인 객체 표현의 한계를 극복한 것이다.

[0027] 좀 더 자세하게는 ORB와 BRISK와 같은 방식에서의 학습에 따른 특징점 선택 또는 특징점의 무작위 선택으로 인한 일반적인 객체의 표현에 한계를 극복할 수 있고, FREAK와 같은 방법에서의 인식 성능이 확실하게 입증되지 않은 문제점을 해결할 수 있다.

도면의 간단한 설명

[0028] 도 1은 본 발명의 일 실시예에 따른 계층적 패턴 구조 기반 기술자 서술 방법과 그 세부 단계를 설명하기 위한 흐름도이다.

도 2는 본 발명의 일 실시예에 따른 기술자 서술을 위한 계층적 패턴 구조를 설명하기 위한 개념도이다.

도 3은 본 발명의 일 실시예에 따른 노드수 9개인 경우에 계층적 패턴 구조에서 이진 기술자를 추출하는 과정을 설명하기 위한 예시도이다.

도 4는 본 발명의 일 실시예에 따른 깊이 1을 갖는 5개의 자식 노드로 구성된 패턴 구조의 기술자 서술 결과를 설명하기 위한 예시도이다.

도 5는 본 발명의 일 실시예에 따른 피라미드 영상 공간을 생성하는 과정을 설명하기 위한 예시도이다.

도 6은 본 발명의 일 실시예에 따른 계층적 패턴 구조의 다양한 형태를 설명하기 위한 예시도이다.

도 7은 본 발명의 일 실시예에 따른 이미지 내 객체의 포즈 추정 방법과 그 세부 단계를 설명하기 위한 흐름도이다.

도 8은 본 발명의 일 실시예에 따른 계층적 패턴 구조 기반의 기술자를 이용한 cascade 기술자 매칭 과정을 설명하기 위한 개념도이다.

도 9는 본 발명의 일 실시예에 따른 코너 특징점 5개에 대한 각 계층적 패턴 구조를 설명하기 위한 예시도이다.

도 10은 본 발명의 일 실시예에 따른 기술자 매칭 과정을 통하여 객체 인식을 수행한 결과를 설명하기 위한 예시도이다.

도 11은 본 발명의 일 실시예에 따른 계층적 패턴 구조 기반 기술자 서술 방법과 이미지 내 객체의 포즈 추정 방법을 통합하여 설명하기 위한 간략한 흐름도이다.

도 12는 본 발명의 일 실시예에 따른 계층적 패턴 구조 기반 객체 인식 장치와 그 세부 구성요소를 설명하기 위한 블록도이다.

발명을 실시하기 위한 구체적인 내용

[0029] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였다.

[0030] 제1, 제2, A, B 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.

[0031] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다.

[0032] 본 출원에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함

하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0033] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥 상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0034] 먼저, 본 출원에서 사용되는 용어를 간략히 설명하면 다음과 같다.

[0035] 기술자(descriptor)란 컴퓨터에서 프로그램 단위의 성질을 정하거나 파일의 각 블록마다 레코드마다의 특성을 정의(define)하는 데 쓰이는 용어이며 기술어로도 번역될 수 있다. 이 기술자는 정보의 내용을 적절히 표시하는 짧은 단어이다. 파일 내에서 목적 레코드와 목적 프로그램 등을 발견하기 위한 키워드로서 사용되며, 어휘의 사용을 한정된 시스템 검사에 쓰여진다. 정보 검색에서는 데이터 베이스 중의 정보를 분류하거나 정보에 색인(index)을 붙이는 데 쓰여지는 이름과 종류명, 이러한 이름과 종류명을 키워드로 검색하고 목적인 것을 조사해 낼 수 있다.

[0036] 특징점이란 객체 간의 식별을 위하여 이미지 내에서 검출하는 포인트를 말한다. 어떤 화상이 입력으로 주어지고, 화상 안에 물체가 무엇인지라는 질문을 컴퓨터가 수행하기 위해선 일단 이 특징점 검출을 해야 하고, 뽑아낸 특징점들이 대체 어떤 물체일까를 매칭을 통해 최종적인 결과를 출력할 수 있게 된다.

[0037] 패치 영역이란 특징점을 중심으로 하여 특징점의 특징을 표현할 수 있는 특징점 주위의 픽셀을 포함하는 일정한 영역이다. 마치 옷을 기울 때에 대는 가죽이나 상처 부위에 일정한 면적으로 붙이는 패치와 형태가 비슷하다 하여 패치 영역이라 명명하였다.

[0038] 노드란 일반적으로 변과 함께 그래프를 구성하는 요소의 하나로서 그래프 이론적으로는 결절, 정점, 점이라고도 한다. 그래프는 점과 선으로 구성되는데, 이 점을 노드 또는 절점이라 한다. 하지만 본 발명에서는 이미지 내에서 특징점을 대표한다는 점에서 노드라고 표현되었을 뿐, 실제로는 특징점의 패치 영역을 나타내는 픽셀을 포함하는 일정한 영역을 가진 부분을 일컫는다.

[0039] 본 발명에서 깊이란 이미지 내 일정한 영역을 계층적으로 분할할 때 그 계층의 개수를 말한다. 예를 들어 깊이가 4란 이야기는 일정한 영역을 분할한 계층이 4개 존재한다는 것을 의미하며, 초기에 분할되지 않은 영역을 계층 0이라고 하면, 계층 1, 계층 2, 계층 3으로 갈수록 조밀하게 분할되는 것을 의미한다.

[0040] 리던던시는 중복이 있다는 뜻으로 본 발명의 노드 분할시 각 분할된 노드가 서로 겹쳐 있는 부분을 말한다. 즉 상호배제(ME, Mutually Exclusive)의 반대말이라고 볼 수 있다. 익셉션이란 수학에서는 차집합을 뜻하지만 본 발명에서는 패치 영역에서 각 분할된 노드를 합친 부분을 뺀 부분이 존재한다는 것을 의미한다. 즉 전체 포괄(CE, Collectively Exhaustive)의 반대말이라고 볼 수 있다.

[0041] 호모그래피란 사영 기하학(projective geometry)에서 벡터 공간의 동형(isomorphism)에 의해 유도된 사영 공간(projective spaces)의 동형을 말한다. 즉, 라인과 라인을 매핑시키는 일대일 대응을 말한다.

[0042] RANSAC 알고리즘이란 회귀분석에서 최소자승법(least squares method)의 단점인 정상적인 분포에서 완전히 틀어져 버린 노이즈 문제를 피할 수 있는 회귀 분석 방법이다. 정상적인 분포에서 완전히 벗어난 이상점들을 아웃라이어(outlier)라고 하는데 데이터에 아웃라이어가 끼어 있으면 최소자승법을 적용한 그림은 엉망이 되어 버린다. 최소자승법은 데이터들과의 잔차의 제곱의 합을 최소화하도록 모델을 찾는 반면, RANSAC은 컨센서스가 최대인, 즉 가장 많은 수의 데이터들로부터 지지를 받는 모델을 선택하는 방법이다. 위에서 가장 많은 수의 데이터들에 해당하는 것이 인라이어(inlier)이다.

[0043] 아핀 변환(Affine transformation)은 한 벡터공간을 다른 벡터공간으로 대응시키는 변환으로, 선형 변환과 평행 이동 변환의 합성으로 이루어져 있다

[0044] 이하, 본 발명에 따른 바람직한 실시예를 첨부된 도면을 참조하여 상세하게 설명한다.

- [0045] 도 1은 본 발명의 일 실시예에 따른 계층적 패턴 구조 기반 기술자 서술 방법과 그 세부 단계를 설명하기 위한 흐름도이다. 도 2는 기술자 서술을 위한 계층적 패턴 구조를 설명하기 위한 개념도이다. 도 3은 노드수 9개인 경우에 계층적 패턴 구조에서 이진 기술자를 추출하는 과정을 설명하기 위한 예시도이다. 도 4는 깊이 1을 갖는 5개의 자식 노드로 구성된 패턴 구조의 기술자 서술 결과를 설명하기 위한 예시도이다. 도 5는 본 발명의 일 실시예에 따른 피라미드 영상 공간을 생성하는 과정을 설명하기 위한 예시도이다. 도 6은 본 발명의 일 실시예에 따른 계층적 패턴 구조의 다양한 형태를 설명하기 위한 예시도이다. 도 1 내지 도 6을 참조하여 계층적 패턴 구조 기반 기술자 서술 방법을 살펴보면 다음과 같다.
- [0046] 계층적 패턴 구조 기반 기술자 서술 방법은, 이미지(image)로부터 객체를 인식하기 위해 기술자(descriptor)를 서술하는 방법에 있어서, 입력된 분석 대상 이미지의 특징점에 대한 패치(patch) 영역에 기반하여 부모 노드(parent node)를 정의하고, 부모 노드를 소정의 깊이(depth)까지 분할한 자식 노드(child node)를 정의하여 계층적 패턴 구조(hierarchical structured pattern)를 생성하는 단계(S140)와, 부모 노드와 자식 노드의 위치 좌표 및 픽셀(pixel) 대표값에 기반하여 패치 영역의 주방향 벡터(master direction vector)를 산출하는 단계(S150)와, 주방향 벡터에 기반하여 패치 영역의 회전각(rotation angle)을 산출하고 계층적 패턴 구조를 회전각만큼 회전하는 단계(S160)를 포함하여 구성될 수 있다.
- [0047] 계층적 패턴 구조 기반 기술자 서술 방법은, 계층적 패턴 구조를 생성하는 단계 이전에, FAST 모서리 추출 기술, Harris Corner 검출 기술 중 적어도 하나를 이용하여 특징점에 대한 패치 영역을 설정하는 단계(S130)를 더 포함하여 구성될 수 있다.
- [0048] 계층적 패턴 구조 기반 기술자 서술 방법은, 회전각만큼 회전하는 단계(S160) 이후에, 회전된 계층적 패턴 구조, 부모 노드의 픽셀 대표값 및 자식 노드의 픽셀 대표값에 기반하여 기술자를 서술하는 단계(S170)를 더 포함하여 구성될 수 있다.
- [0049] 계층적 패턴 구조는, 최상위 부모 노드를 소정의 노드수(N)만큼 분할하여 최초 자식 노드가 정의되고, 최초 자식 노드는 차상위 부모 노드가 되어 소정의 깊이 단계마다 자식 노드를 정의하는 재귀적(recursive) 방법으로 생성하되, 소정의 노드수(N)는 자연수이고, 분할은 리던던시(redundancy) 또는 익셉션(exception)이 있는 분할일 수 있다. 나아가 재귀적 방법은, 최상위 부모 노드의 반지름(radius)을 지름(diameter)으로 하고, 최상위 부모 노드를 균등 분할하여 소정의 노드수(N)의 최초 자식 노드를 생성하는 제1 분할방법을 수행하거나, 소정의 노드수 N에서 1을 뺀 수만큼 제1 분할방법을 수행한 이후에 최상위 부모 노드와 동심원(concentric circle)인 최초 자식 노드 하나를 추가로 생성하는 제2 분할방법을 수행하거나, 또는 최상위 부모 노드의 반지름보다 긴 길이의 지름으로 하고, 최상위 부모 노드를 균등 분할하여 소정의 노드수(N)의 최초 자식 노드를 생성하는 제3 분할방법을 수행할 수 있다. 또 나아가 소정의 노드수 및 소정의 깊이는, 특징점에 기반하여 확장 또는 축소될 수 있다.
- [0050] 픽셀 대표값은, 부모 노드와 자식 노드에 해당하는 영역의 픽셀값들의 평균이거나, 영역을 가우시안 컨벌루션(Gaussian Convolution)하여 가중치를 적용한 픽셀값일 수 있다.
- [0051] 계층적 패턴 구조 기반 기술자 서술 방법은, 계층적 패턴 구조를 생성하는 단계(S140) 이전에, 이미지를 실수배(real number multiple)하여 크기가 다른 복수의 이미지로 구성된 피라미드 영상 공간을 생성하고, 피라미드 영상 공간 내의 이미지를 분석 대상 이미지로 출력하는 단계(S120)를 더 포함하여 구성될 수 있다.
- [0052] 이진 기술자의 경우, 컴퓨터 부호의 1bit 영역으로 0 또는 1로서 표현되는 의미 정보를 갖는 이진 부호들의 집합을 의미하며, 기술자 매칭에 있어서 매우 빠른 속도로 정합이 가능하고, 임의의 두 위치에 대한 명암 값 비교와 같이 간단한 연산으로 생성 가능할 뿐만 아니라 적은 용량의 저장 메모리가 필요하다는 이점을 가질 수 있다. 본 발명에서는 새로운 계층적 패턴 구조에 기반한 기술자를 기술하는 새로운 방법을 제안하고 이를 이용하여 객체를 인식하는 방법 및 장치를 기술한다. 본 발명을 적용하기 위한 대상은 객체 내에 FAST 모서리 추출기, Harris Corner 검출기 등의 기존의 특징점 추출 방법을 사용하여 얻은 지역 특징점의 패치 영역으로 정의된 얼굴과 같은 의미 정보를 갖는 객체 영역에 대하여 적용이 가능하다.
- [0053] 특징점 영역에서 이진 기술자를 기술하기 위해 제안하는 계층적 패턴 구조의 예시는 도 2와 같다. 도 2에서처럼 패턴 구조는 부모와 자식 노드들의 재귀적 상속관계로서 구성될 수 있다. 같은 계층의 노드들은 같은 크기의 원으로 구성될 수 있으며, 이 노드들을 하나의 계층(레이어, layer)로서 정의할 수 있다. 부모에 상속되는 자식 노드의 수는 요구되는 성능에 따라 노드수를 확장 또는 축소 가능하고, 레이어의 깊이 또한 확장 축소가 가능하다. 자식 노드의 크기는 예를 들어 부모 노드의 반지름을 지름으로 하여, 360도를 노드수만큼 분할하는 위치에

서의 중심원의 형태를 취할 수 있다. 노드의 수가 증가하거나, 레이어의 깊이가 깊어질수록 좀 더 정밀한 기술자를 생성할 수 있지만, 대신 요구되는 메모리 크기와 처리 속도는 늦어진다.

[0054] 도 2는 부모 노드 1개가 총 9개의 자식 노드들을 갖는 패턴 구조로 레이어 깊이가 4인 경우 총 820개의 노드로 구성되며 부모 노드와 자식 노드가 계층적 관계로 구성되어 있음을 보여준다. 각 노드는 예를 들어 원으로 표현될 수 있는데, 자식 노드 한 개의 크기는 부모 노드의 반지름을 지름으로 하는 크기를 가질 수 있고, 노드를 대표하는 픽셀 대표값은 각 노드의 위치에서 해당 영역의 픽셀값들의 평균이나, 노드 크기에 따라 정해진 분산치를 갖는 가우시안 함수를 convolution한 화소값의 가중치된 값으로 표현될 수 있다.

[0055] 도 2에서 서술되는 기술자의 차원은 819 차원으로서 최상위 노드는 부모를 갖지 않기 때문에, 전체 노드의 개수에서 최상위 부모 노드 한 개를 제외한 전체 자식 노드의 수가 기술자의 차원이 된다. 수학식 1은 부모 노드 한 개가 갖는 자식 노드의 수에 따른 기술자의 차원의 수를 구하는 식을 나타낸다.

수학식 1

$$D = \sum_{d=1}^L c^d$$

[0056]

[0057] L : 계층의 깊이,

[0058] c : 한 개의 부모가 갖는 자식 노드의 수,

[0059] D : 기술자의 차원의 수.

[0060] 도 2에서는 9개의 자식 노드의 수를 나타낼 수 있다. 수학식 1은 레이어의 깊이가 깊어질수록 표현되는 기술자의 차원은 지수 형태로 증가함을 보여준다.

[0061] 부모 노드가 9개의 자식 노드를 갖는 패턴 구조에서 이진 기술자를 추출하는 방법은 도 3과 같다. 레이어 0의 부모 노드 한 개와 레이어 1의 자식 노드들의 픽셀 대표값을 순차적으로 비교하여 최상위 계층의 9개의 bit 부호가 결정되고, 레이어 1에서 부모 노드 한 개와 레이어 2의 자식 노드들과의 픽셀 대표값 비교를 통해 bit 부호를 기술할 수 있다. 이어서 하위 레이어도 깊이에 따라 동일하게 비트 부호를 기술할 수 있다. 도 3에서 최상위 레벨의 레이어 1의 부모 노드와 자식 노드들은 각각 채워진 보라색과 빨간색 외곽선으로 표시되었고, 레이어 2에서는 채워진 빨간색, 하늘색 외곽선, 레이어 3에서는 채워진 하늘색, 노란색 외곽선으로 각각 표현되었다. 기술자 서술은 부모 노드의 픽셀 대표값과 다음 레이어의 부모 노드에 포함되는 자식 노드들의 픽셀 대표값의 비교를 통하여 기술자를 서술할 수 있다. 아래 수학식 2는 제한한 기술자를 서술하는 과정을 나타낸 것이다.

수학식 2

$$F = \sum_{i=1}^D 2^i T(x_i)$$

$$T(x_i) = \begin{cases} 1 & \text{if } I(P(x_i)) - I(x_i) > 0 \\ 0 & \text{otherwise,} \end{cases}$$

[0062]

[0063] D : 기술자의 차원의 수,

[0064] x_i : 기술자의 i번째 bit의 부호를 결정하는 i번째 노드의 중심 위치,

- [0065] $P(x_i)$: i 번째 노드의 부모 노드의 중심 위치,
- [0066] $I(\)$: 해당 노드의 픽셀 대표값,
- [0067] $T(x_i)$: i 번째 부모 노드와 자식 노드의 대표값 비교.

[0068] 본 발명에서의 기술자 생성 방법은 각 bit의 기술 순서가 정해져 있기 때문에 객체 영역이 회전하게 되면 회전된 객체 영역에 대한 기술자가 다르게 된다. 즉, 객체 영역의 회전에 대하여 항상성이 없다. 이와 같은 회전 불변에 대한 정의는 도 4와 같이 표현될 수 있다. 도 4는 깊이 1을 갖는 5개의 자식 노드로 구성된 패턴 구조로 흰색 숫자는 각 노드를 기술하는 순서를 나타내고, 각 그림 아래의 부호의 집합은 기술된 결과를 보여준다. 도 4의 (a)는 고정된 계층적 패턴으로 동일한 객체의 회전된 영상에 대해 다른 기술 결과를 보여주고, (b)는 패턴을 90도 단위로 회전 후 기술자를 생성한 결과로 동일한 객체의 회전된 영상에 대해 같은 기술 결과를 보여준다. 도 4는 객체가 가지고 있는 회전에 불변한 특정 방향에 대하여 패턴을 회전 할 경우 회전에 대하여 항상 불변한 기술자를 생성할 수 있는 것을 의미한다. 따라서, 회전에 불변한 기술자를 추정하기 위하여 특징점 패치 영역에서 주방향을 추정하여 계층적 패턴을 회전 정규화하는 것이 필요하다. 수학적 3은 본 발명에서 계층적 패턴 구조를 이용하여 주방향을 추정하는 것을 나타낸다.

수학적 3

$$\vec{O} = \frac{1}{D} \sum_{i=1}^D (I(P(x_i)) - I(x_i)) \cdot \frac{P(x_i) - x_i}{\|P(x_i) - x_i\|}$$

- [0069]
- [0070] O : 특징점 중심으로부터 주방향에 대한 방향 벡터,
- [0071] D : 기술자의 차원으로, 즉 전체 자식 노드들의 수.

[0072] 수학적 3에서 추정된 방향 벡터로부터 아래 수학적 4에서 주방향에 대한 회전각을 추정할 수 있으며, 추정된 회전각을 이용하여 패턴을 회전시키고, 기술자를 기술함으로써 회전 불변한 기술자를 생성할 수 있다.

수학적 4

$$\theta = \tan^{-1} \frac{O_y}{O_x}$$

- [0073]
- [0074] θ : 주방향에 대한 회전각.

[0075] 본 발명에서는 크기에 불변한 기술자를 생성하기 위해서, 도 5에 나타난 것처럼 크기 공간(scale space)인 피라미드 영상을 생성하고 각 크기 레벨에 대하여 기술자를 기술할 수 있다. 피라미드 영상은 입력 영상으로부터 특정 배수로 영상의 크기를 변화시켜 영상의 크기 공간을 생성하는 방법이다. 피라미드 영상의 생성 후 각 크기 레벨의 영상에 대하여 특징점을 추정하고 각 위치에 대해 기술자를 기술함으로써 피라미드 공간 내의 객체의 크기 변화에 대한 인식을 수행할 수 있다.

- [0076] 도 6은 본 발명에서 제시한 계층적 패턴 구조의 다양한 형태의 변종이 가능함을 보여준다. 도 6의 (a)는 앞서 제시한 패턴 구조에서 중심 노드가 없는 경우의 패턴 구조를 보여준다. 이 패턴 구조의 경우 8bit로 표현이 가능하기 때문에, 컴퓨팅 환경에서 문자 타입(character, char)형의 8bit 메모리로서 표현되어 메모리가 절약되고, 매칭 과정에서 최적화 기법을 사용하기 용이하다. 또한 기술자 서술 및 매칭의 구현이나 연산이 간단하다. 그러나, 중심 노드가 없기 때문에 도 2의 9bit 계층 기술자에 비해 인식률이 낮은 단점이 있을 수 있다. 도 6의 (b)는 자식 노드의 크기가 부모 노드의 반지름의 크기보다 큰 경우의 패턴 구조로서, 패턴이 특징점 중심에 더 밀집되어 표현될 수 있다. 이것은 기술자를 기술함에 있어 패치의 중심의 정보를 더 많이 사용하는 것으로 중심에 가까운 영역일수록 기술자 생성에 영향을 더 많이 줄 수 있음을 나타낸다. 이 패턴의 경우에는 특징점 영역 내에 외곽보다 중심에 정보가 많이 밀집되어 있는 경우 도 2의 9bit 계층 기술자에 비해 인식률이 높을 수 있다.
- [0077] 도 7은 본 발명의 일 실시예에 따른 이미지 내 객체의 포즈 추정 방법과 그 세부 단계를 설명하기 위한 흐름도이다. 도 8은 계층적 패턴 구조 기반의 기술자를 이용한 cascade 기술자 매칭 과정을 설명하기 위한 개념도이다. 도 9는 코너 특징점 5개에 대한 각 계층적 패턴 구조를 설명하기 위한 예시도이다. 도 10은 기술자 매칭 과정을 통하여 객체 인식을 수행한 결과를 설명하기 위한 예시도이다. 도 7 내지 도 10을 참조하여 이미지 내 객체의 포즈 추정 방법을 살펴보면 다음과 같다.
- [0078] 이미지 내 객체의 포즈 추정 방법은, 계층적 패턴 구조의 기술자를 이용하여 이미지 내 객체의 포즈를 추정하는 방법에 있어서, 기술자를 소정의 개수(D)만큼의 군집으로 구분하는 단계(S710)와, 기술자의 i번째 군집과 미리 저장된 참조 기술자의 i번째 군집 간의 거리를 계산하는 단계(S720)와, 계산된 거리가 소정의 임계값보다 작다는 판단(S730)에 따라 기술자의 (i+1)번째 군집과 미리 저장된 참조 기술자의 (i+1)번째 군집 간의 거리를 계산하는 단계(S720)와, 기술자의 D번째 군집과 미리 저장된 참조 기술자의 D번째 군집 간의 거리가 소정의 임계값보다 작다는 판단에 따라 기술자와 미리 저장된 참조 기술자를 매칭하는 단계(S750)를 포함하되, i는 1을 초항(first term)으로 (D-1)까지 반복(iteration)될 수 있다.
- [0079] 기술자의 i번째 군집과 미리 저장된 참조 기술자의 i번째 군집 간의 거리가 소정의 임계값보다 크거나 같은 경우, (i+1)번째 군집 간의 거리를 계산하는 단계를 생략하고 기술자와 미리 저장된 참조 기술자와의 매칭을 스킵(skip)할 수 있다.
- [0080] 이미지 내 객체의 포즈 추정 방법은, 기술자를 매칭하는 단계(S750) 이후에, 객체와 미리 저장된 참조 기술자의 대상인 참조 객체 간의 호모그래피(homography)를 추정하는 단계(S770)를 더 포함하여 구성될 수 있다. 나아가 이미지 내 객체의 포즈 추정 방법은, 호모그래피를 추정하는 단계(S770) 이전에, 호모그래피에 기반하여 객체의 인식 여부를 판단하는 단계(S760)를 더 포함하되, 인식 여부를 판단하는 방법은, 최종 매칭된 기술자들 중 RANSAC(RANdom Sample Consensus) 알고리즘을 통해 산출된 인라이어(inlier)의 개수에 기반하여 객체의 인식 여부를 판단하는 제1 판단 방법과 호모그래피의 정확도에 기반하여 객체의 인식 여부를 판단하는 제2 판단 방법과 객체를 호모그래피에 기반하여 아핀 변환(affine transform)한 영역과 참조 객체의 영역을 동일한 포즈로 정규화하여 템플릿 매칭한 결과 정보에 기반하여 객체의 인식 여부를 판단하는 제3 판단 방법 중 어느 하나일 수 있다.
- [0081] 소정의 개수(D)는, 계층적 패턴 구조의 기술자의 대상인 특징점에 대한 패치 영역을 자식 노드로 분할한 깊이와 같을 수 있다.
- [0082] 입력 영상 내의 특징점들로부터 기술된 계층적 구조의 기술자들은 사전에 등록된 데이터베이스 영상에서 추출된 계층적 구조의 기술자들과의 거리 비교를 통하여 매칭 과정을 수행할 수 있다. 여기서의 매칭이란 가장 가까운 거리를 갖는 두 쌍의 기술자를 매칭되었다고 정의할 수 있다. 본 발명에서 제안하는 기술자는 이진 부호의 집합으로 해밍 거리(Hamming distance)를 이용하여 거리를 측정 할 수 있다. 이진 기술자 도메인에서의 해밍 거리는 컴퓨터 bit-wise XOR 연산을 통해 쉽게 계산 가능하다. 따라서, 기존의 소수점으로 표현되었던 기술자들에 비해 매우 빠른 속도로 기술자 매칭이 가능하다.
- [0083] 본 발명에서는 매칭을 위한 거리 측정 방법으로 해밍 거리를 이용할 뿐 아니라, 계층적 구조를 이용한 cascade 매칭 방법을 사용하여 기존 방법들에 비해 더 빠른 속도로 기술자를 매칭하는 방법을 제안한다. 도 8은 본 발명에서 제안하는 기술자를 이용한 cascade 기술자 매칭에 대한 과정을 보여준다. 여기서, cascade 매칭은 기술자

의 bit들을 각 단계로 군집을 구성하고, 앞선 단계에서 매칭이 된 경우 다음 단계로 연속적으로 매칭이 수행되는 기법을 의미한다.

[0084] 본 발명에서는 이미 사전에 정의된 계층적 구조에서 각 계층을 cascade의 각 단계로 정의하고 cascade 매칭을 수행함으로써 고속 매칭을 수행할 수 있다. 특징점을 중심으로 한 패치 영역의 전체 특성을 나타내는 처음 9개의 bit만을 이용하여 거리 측정 결과 두 기술자가 서로 다르다는 것이 판명되어지면(rejection) 더 이상의 거리 측정이 필요 없게 되어 뒤따르는 bit간 비교 연산이 생략되므로 전체 연산을 대폭 줄일 수 있음을 나타내며, rejection 되지 않는 경우 뒤따르는 bit 연산에 의한 상세한 매칭이 수행되어짐을 나타내고 있다.

[0085] 이것은 본 발명의 패턴 구조에 대해 상위 레이어는 적은 양의 정보로 패치의 대략의 특성을 나타내고, 하위 레이어는 많은 정보를 포함하는 상세한 특성을 나타내는 계층적 구조의 특성과 부합한다. 도 9는 영상에서 임의의 코너 특징점 5개를 추출하고, 해당 위치에서의 각 계층에 대한 구조를 나타낸다. 레이어 1의 경우 적은 수의 큰 노드들로 구성되어 있고, 레이어의 깊이가 깊어질수록 많은 수의 노드들로 구성되어 레이어 3의 경우 특징점을 중심으로 하는 패치 영역을 조밀하고 상세하게 표현할 수 있다.

[0086] 각 단계에 대한 매칭을 기각하기 위해서 각 단계에 대하여 임계값을 취하고, 이에 대하여 기각을 수행하는 방법을 사용할 수 있다. 최종 매칭된 기술자들은 RANSAC 알고리즘을 통하여 객체 간의 관계인 호모그래피를 추정하여 변환된 자세 및 인식 여부를 판정할 수 있다. 여기서 인식 여부는 인라이어의 개수를 판정하여 결정하거나, 호모그래피의 정확도를 판정하여 인식 여부를 판별하거나, 추정된 포즈에 대한 아핀(Affine) 변환된 영역과 사전에 저장된 객체 영상을 같은 포즈로 정규화하여 템플릿 매칭을 수행하여 매칭 정도가 일정 임계값 이상인 경우 인식으로 판정할 수 있다.

[0087] 도 10은 본 발명에서 제시한 기술자와 매칭 과정을 이용하여 객체 인식을 수행한 결과를 예시한다. 왼쪽 영상은 임의의 데이터베이스 영상이고, 우측의 영상은 데이터베이스 영상에 대한 아핀 회전을 갖는 영상을 보여준다. 도 10에서 (a)는 두 영상에 대하여 각각 약 500개 이상의 특징점을 추출 후 앞서 제시한 cascade 매칭을 수행한 결과를 보여준다. 가장 가까운 기술자의 쌍의 위치에 대하여 직선으로 표현하였다. (b)는 매칭된 500개 이상의 기술자 쌍에 대해 RANSAC 알고리즘을 적용하여 아웃라이어 제거하고, 인라이어만을 직선으로 표현하였다. (c)는 인라이어 쌍으로부터 두 영상의 호모그래피를 추정하고, 왼쪽 영상을 객체로 하여 우측 영상에 아핀된 포즈를 추정된 결과로써 녹색 사각형으로 아핀 변환된 객체의 포즈를 출력하여 보여줄 수 있다.

[0088] 도 11은 본 발명의 일 실시예에 따른 계층적 패턴 구조 기반 기술자 서술 방법과 이미지 내 객체의 포즈 추정 방법을 통합하여 설명하기 위한 간략한 흐름도이다. 도 11을 참조하여 계층적 패턴 구조 기반 기술자 서술 방법과 이를 이용한 이미지 내 객체의 포즈 추정 방법을 정리하여 간략하게 정리하여 보면 다음과 같다.

[0089] 본 발명에서 인식을 위한 객체는 영상 내 특정 물체, 얼굴, 사람 등의 의미정보를 포함하는 식별 가능한 대상에 대한 영역을 의미하고, 인식은 카메라 입력 영상 내 대상 객체의 위치 및 포즈를 추정하는 것으로 정의할 수 있다. 제안하는 기술자를 이용하여 객체를 인식하는 방법은 ① 피라미드 영상을 생성하여 코너(corner)와 같은 특징점을 생성하는 단계, ② 특징점을 중심으로 하는 패치 영역에서 주 방향을 추정하고 계층적 패턴 구조를 회전 정규화 하는 단계, ③ 특징점을 중심으로 회전 정규화된 계층적 패턴 구조를 이용하여 부호들의 집합으로 표현된 기술자를 기술하는 단계, ④ 입력 영상으로부터 추출된 이진 기술자와 사전에 저장된 대상 객체 기술자와 cascade 기술자 매칭(cascade matching)을 수행하는 단계, 그리고 ⑤ 정합된 기술자들의 2차원 위치를 기반으로 객체 간 기하학적 변환 관계(homography)를 추정하는 단계로 구성될 수 있다.

[0090] 도 11은 상기 과정에 대해 지역 특징점을 기반으로 객체를 인식하는 과정을 보여준다. 도 11의 흐름도 왼쪽 과정은 입력 영상에 대한 계층 구조의 이진 기술자 연산 과정으로, 입력 영상이 입력되면 피라미드 이미지를 생성하여 크기 공간을 생성하고, 각 크기 영상에 대하여 다수의 코너와 같은 (corneriness) 특징점을 추출하며, 추출된 특징점의 각 위치를 기반으로 본 발명에서 제시한 패턴 구조를 이용하여 주방향을 추정하고 주방향에 대하여 패턴을 회전 정규화한다. 이 과정은 보통의 경우 속도 향상을 위해 일정 각도의 패턴 회전에 대하여 사전에 룩업 테이블(lookup table)을 생성하여 회전에 대한 불필요한 연산 과정을 줄일 수 있다.

[0091] 주방향에 대하여 패턴 정규화 후 각 노드를 중심으로 하는 영역의 화소값의 총합을 적분 영상(integral image)을 이용하여 구한 후 이를 영역 크기로 나눈 화소값의 평균치나, 패턴의 각 계층의 노드 크기에 따라 정해진 분산치를 갖는 가우시안 함수를 convolution한 화소값의 가중치를 각 노드의 대표값으로 하고 이들을 앞에서 설명한 방법으로 서로 비교하여 계층 구조의 이진 기술자를 얻을 수 있다.

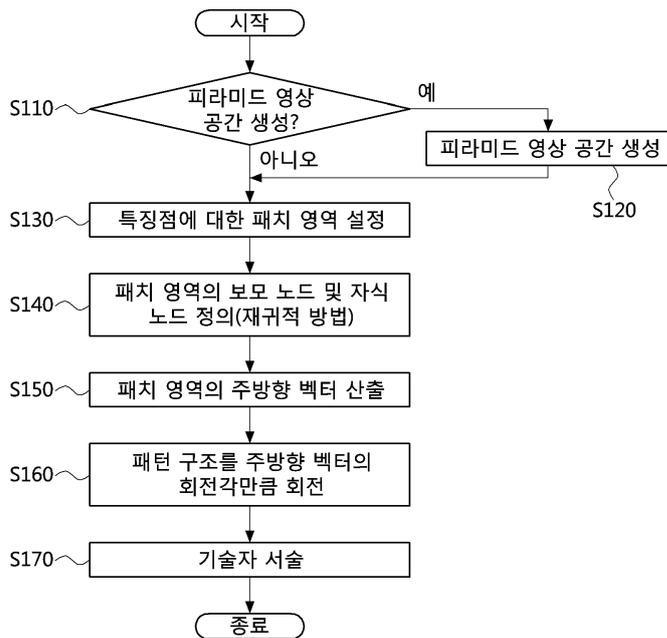
- [0092] 입력 영상을 나타내는 이진 기술자들과 저장된 데이터베이스 영상을 나타내는 이진 기술자들을 본 발명에서 제안하는 cascade 매칭 방법을 이용하여 기술자 매칭을 수행할 수 있다. 매칭된 기술자들에 대하여 RANSAC 알고리즘을 적용하여 정확하게 매칭된 인라이어 기술자 쌍을 추정하고, 인라이어 쌍이 최소 4개 이상인 경우 인라이어 쌍의 특징점 위치에 대해 호모그래피 추정함으로써 객체의 포즈를 인식할 수 있다.
- [0093] 도 12는 본 발명의 일 실시예에 따른 계층적 패턴 구조 기반 객체 인식 장치(10)와 그 세부 구성요소를 설명하기 위한 블록도이다. 도 12를 참조하여 계층적 패턴 구조 기반 객체 인식 장치(10)를 살펴보면 다음과 같다.
- [0094] 계층적 패턴 구조 기반 객체 인식 장치(10)는, 계층적 패턴 구조 기반의 기술자를 이용하여 이미지로부터 객체를 인식하는 장치에 있어서, 입력된 분석 대상 이미지의 특징점에 대한 패치 영역에 기반하여 부모 노드를 정의하고, 부모 노드를 소정의 깊이까지 분할한 자식 노드를 정의하여 계층적 패턴 구조를 생성하는 패턴 생성부(200)와, 부모 노드와 자식 노드의 위치 좌표 및 픽셀 대표값에 기반하여 패치 영역의 주방향 벡터를 산출하는 주방향 산출부(400)와, 주방향 벡터에 기반하여 패치 영역의 회전각을 산출하고 계층적 패턴 구조를 회전각만큼 회전하는 패턴 회전부(500)를 포함하여 구성될 수 있다.
- [0095] 계층적 패턴 구조 기반 객체 인식 장치(10)는, FAST 모서리 추출 기술, Harris Corner 검출 기술 중 적어도 하나를 이용하여 특징점에 대한 패치 영역을 설정하는 영역 설정부(100)를 더 포함하여 구성될 수 있다.
- [0096] 계층적 패턴 구조 기반 객체 인식 장치(10)는, 회전된 계층적 패턴 구조, 부모 노드의 픽셀 대표값 및 자식 노드의 픽셀 대표값에 기반하여 기술자를 서술하는 기술자 서술부(600)를 더 포함하여 구성될 수 있다.
- [0097] 계층적 패턴 구조 기반 객체 인식 장치(10)는, 이미지를 실수배(real number multiple)하여 크기가 다른 복수의 이미지로 구성된 피라미드 영상 공간을 생성하고, 피라미드 영상 공간 내의 이미지를 분석 대상 이미지로 출력하는 피라미드 영상 생성부(100)를 더 포함하여 구성될 수 있다.
- [0098] 계층적 패턴 구조 기반 객체 인식 장치(10)는, 기술자를 소정의 깊이만큼의 군집으로 구분하여 기술자와 미리 저장된 참조 기술자 간의 거리를 군집 단위에 기반하여 계산하고 계산된 결과에 기반하여 기술자와 미리 저장된 참조 기술자의 매칭 여부를 결정하는 기술자 매칭부(700)를 더 포함하여 구성될 수 있다. 나아가 계층적 패턴 구조 기반 객체 인식 장치(10)는, 객체와 미리 저장된 참조 기술자의 대상인 참조 객체 간의 호모그래피(homography)를 추정하는 호모그래피 추정부(800)를 더 포함하여 구성될 수 있다. 더 나아가 계층적 패턴 구조 기반 객체 인식 장치(10)는, 호모그래피에 기반하여 객체의 인식 여부를 판단하는 객체 인식부를 더 포함하되, 인식 여부를 판단하는 방법은, 최종 매칭된 기술자들 중 RANSAC(RANdom Sample Consensus) 알고리즘을 통해 산출된 인라이어(inlier)의 개수에 기반하여 객체의 인식 여부를 판단하는 제1 판단방법, 호모그래피의 정확도에 기반하여 객체의 인식 여부를 판단하는 제2 판단 방법, 또는 객체를 상기 호모그래피에 기반하여 아핀 변환(affine transform)한 영역과 참조 객체의 영역을 동일한 포즈로 정규화하여 템플릿 매칭한 결과 정보에 기반하여 객체의 인식 여부를 판단하는 제3 판단 방법 중 어느 하나일 수 있다.
- [0099] 계층적 패턴 구조 기반 객체 인식 장치(10)에 대한 설명은 앞서 설명한 계층적 패턴 구조 기반 기술자 서술 방법과 이미지 내 객체의 포즈 추정 방법에 대한 설명에서 상술한 바 있으므로 중복하여 기재하지는 않는다.
- [0100] 비록 몇몇의 측면들은 방법의 관점에서 설명되었지만, 이러한 측면들은 상응하는 장치의 설명을 나타내는 것이 명확하며, 여기서 방법의 단계는 장치에 대응한다. 특정 구현 요구들에 따르면, 발명의 실시예들은 하드웨어 또는 소프트웨어에서 구현될 수 있다. 본 발명의 실시예들은 프로그램 코드, 방법들 중 하나의 수행을 위해 동작하는 프로그램 코드를 가지는 컴퓨터 프로그램 제품으로서 수행될 수 있다.
- [0101] 이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서 본 발명에 개시된 실시 예들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시 예에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

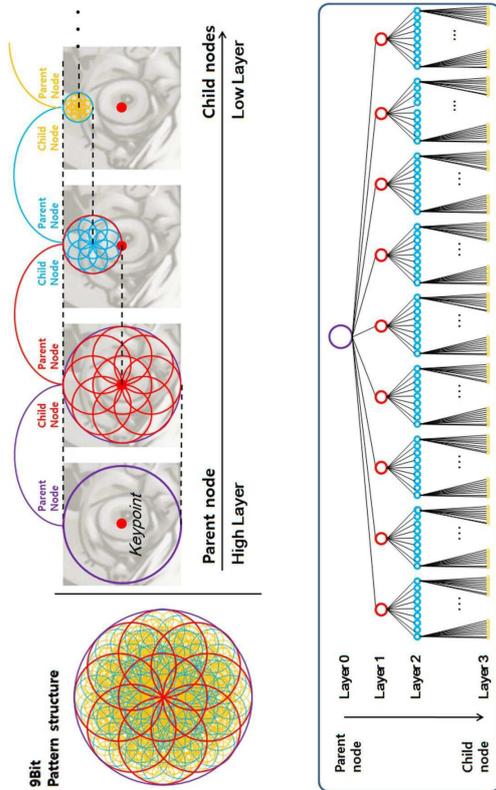
- [0102]
- | | |
|---------------------------|----------------|
| 10: 계층적 패턴 구조 기반 객체 인식 장치 | |
| 100: 피라미드 영상 생성부 | 200: 영역 설정부 |
| 300: 패턴 생성부 | 400: 주방향 산출부 |
| 500: 패턴 회전부 | 600: 기술자 서술부 |
| 700: 기술자 매칭부 | 800: 호모그래피 추정부 |

도면

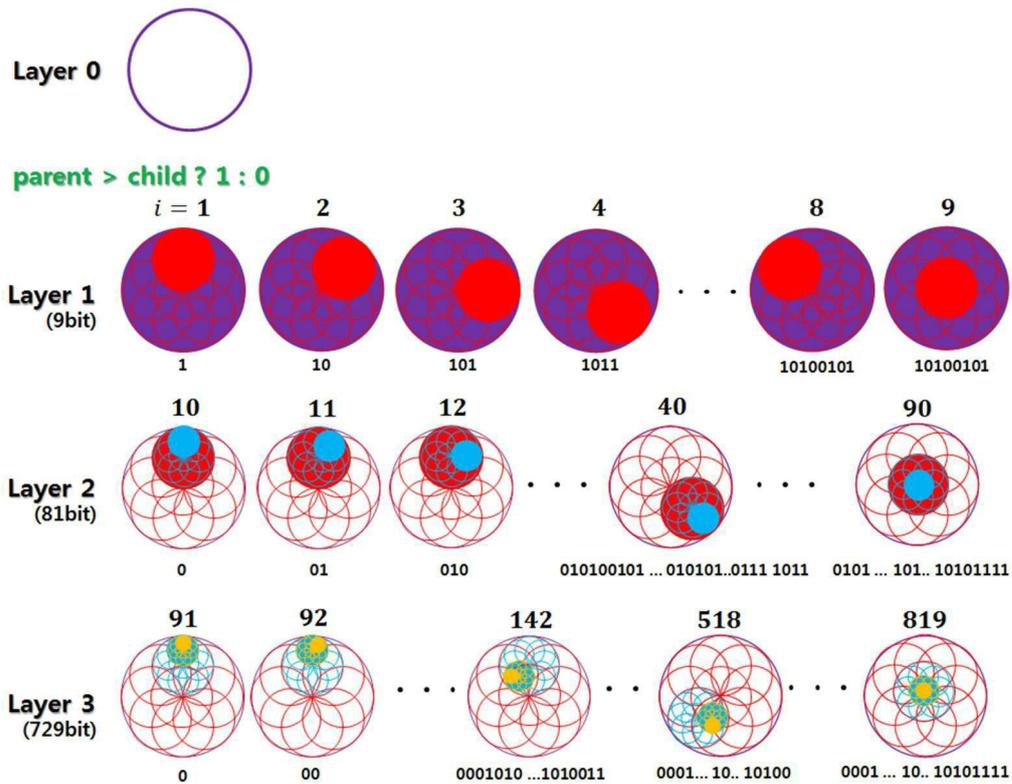
도면1



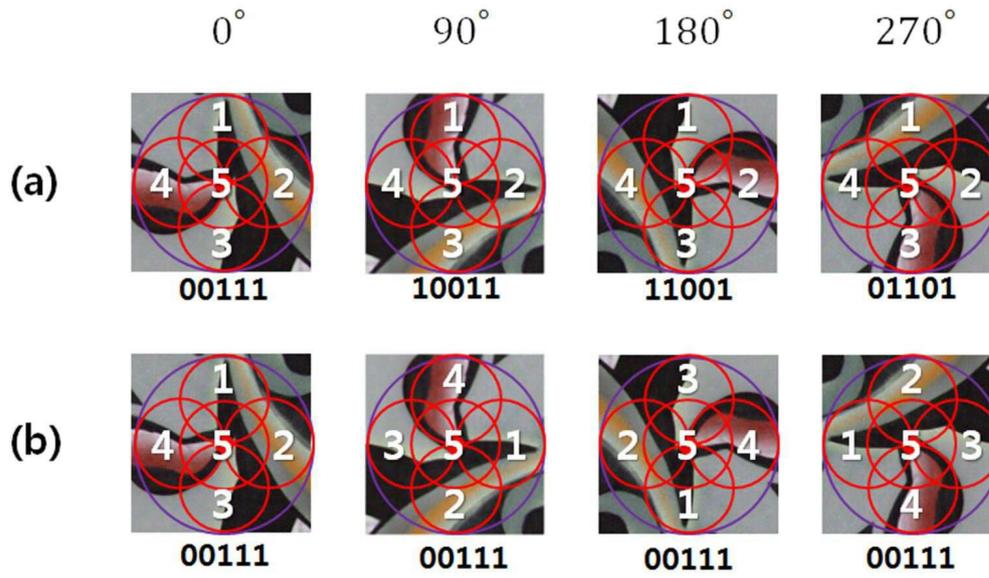
도면2



도면3

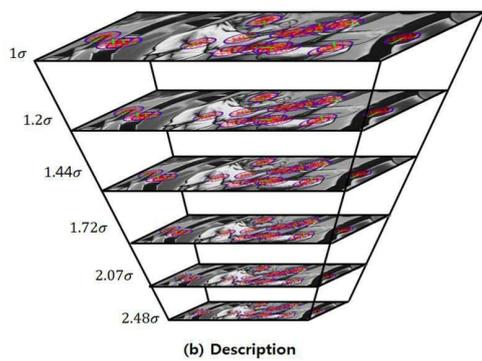
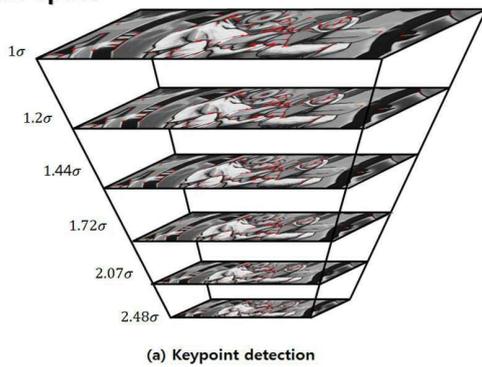


도면4

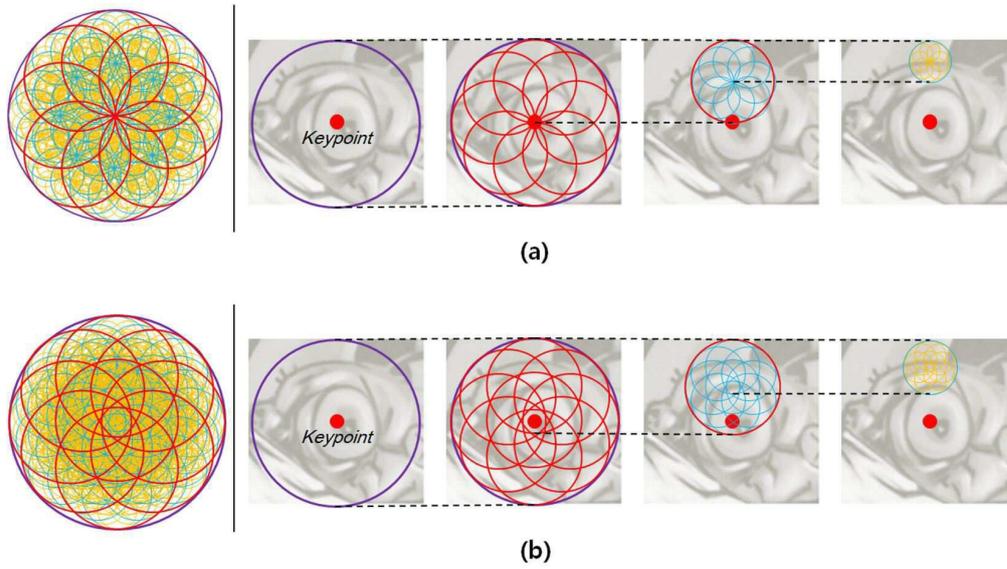


도면5

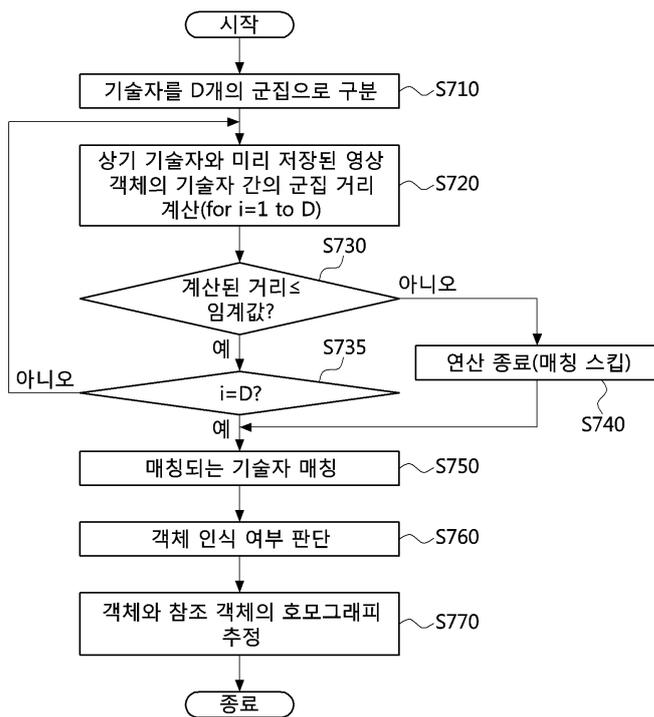
Scale space



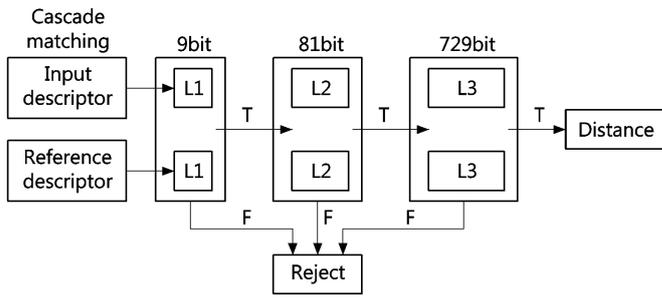
도면6



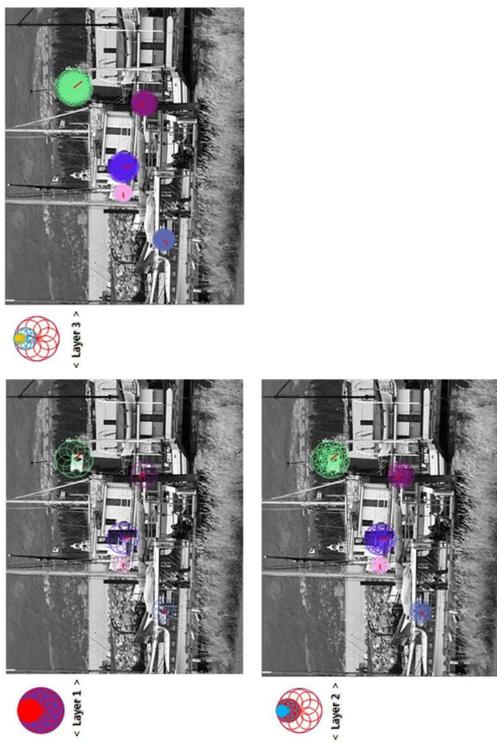
도면7



도면8



도면9



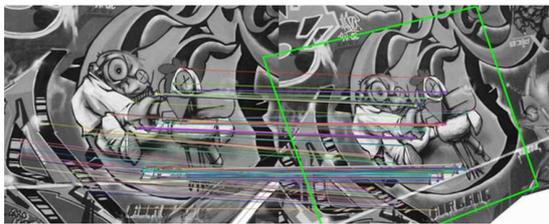
도면10



(a) Matching result

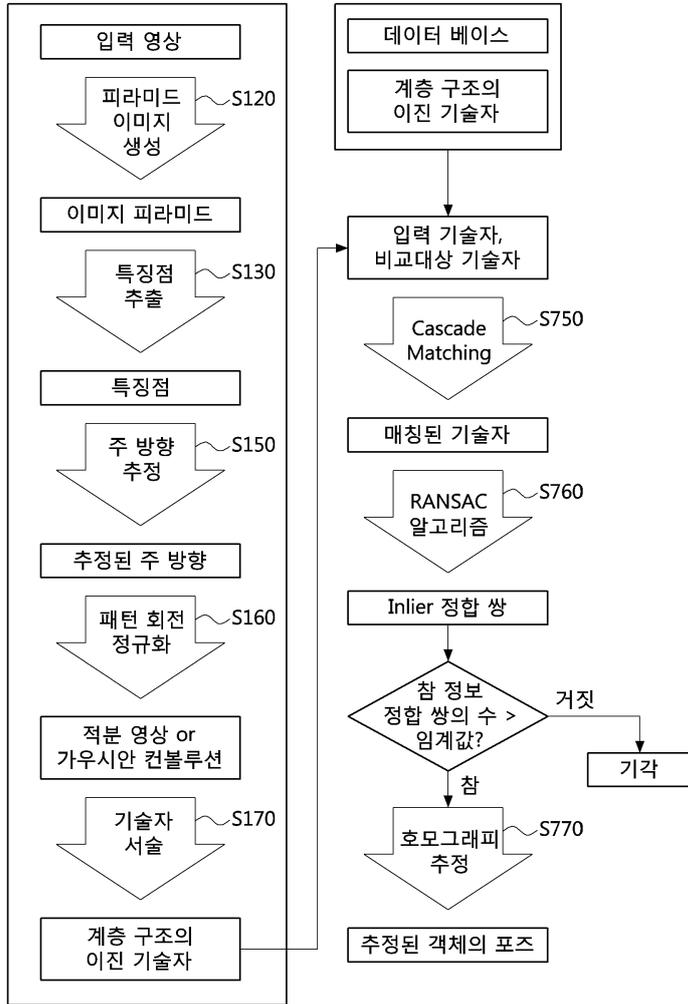


(b) Ransac result



(c) Find homography

도면11



도면12

