



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2014년04월17일
(11) 등록번호 10-1386513
(24) 등록일자 2014년04월11일

(51) 국제특허분류(Int. Cl.)

G06T 7/00 (2006.01)

(21) 출원번호 10-2012-0129155

(22) 출원일자 2012년11월14일

심사청구일자 2012년11월14일

(56) 선행기술조사문헌

KR101178333 B1

홍준혁 외2, "HOOF와 Random Forest 이용한 휴먼 행동 인식", 한국컴퓨터종합학술대회 논문집 Vol.39, No.1(B), pp.450-452, 2012.06.

조수선, "의미적 연관태그와 이미지 내용정보를 이용한 웹 이미지 분류", 한국인터넷정보학회 제11권 제3호, pp.15-24, 2010.06.

(73) 특허권자

계명대학교 산학협력단

대구광역시 달서구 달구벌대로 1095 (신당동)

(72) 발명자

고병철

대구광역시 수성구 동대구로 345 범어역우방유셀 3306호

홍준혁

경상북도 경산시 동지로19길 6

남재열

대구광역시 수성구 지범로40길 11 서한화성아파트 103동 1401호

(74) 대리인

김건우

전체 청구항 수 : 총 17 항

심사관 : 진민숙

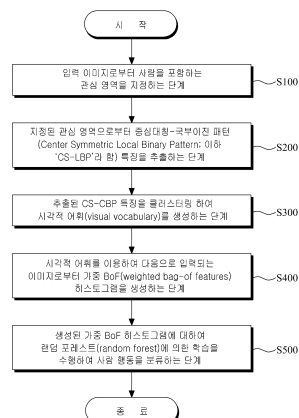
(54) 발명의 명칭 공간적 피라미드 기반의 가중 비오에프 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법

(57) 요약

본 발명은 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법에 관한 것으로서, 보다 구체적으로는 (1) 입력 이미지로부터 사람을 포함하는 관심 영역을 지정하는 단계, (2) 지정된 관심 영역으로부터 중심대칭-국부이진 패턴(Center Symmetric Local Binary Pattern; 이하 'CS-LBP' 라 함) 특징을 추출하는 단계, (3) 추출된 CS-LBP 특징을 클러스터링 하여 시각적 어휘(visual vocabulary)를 생성하는 단계, (4) 시각적 어휘를 이용하여 다음으로 입력되는 이미지로부터 가중 BoF(weighted bag-of features) 히스토그램을 생성하는 단계 및 (5) 생성된 가중 BoF 히스토그램에 대하여 랜덤 포레스트(random forest)에 의한 학습을 수행하여 사람 행동을 분류하는 단계를 포함하는 것을 그 구성상의 특징으로 한다.

본 발명에서 제안하고 있는 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법에 따르면, 입력 이미지로부터 사람을 포함하는 관심 영역을 지정하고, 지정된 관심 영역으로부터 중심대칭-국부이진 패턴(Center Symmetric Local Binary Pattern; 이하 'CS-LBP' 라 함) 특징을 추출하여, 추출된 CS-LBP 특징을 클러스터링 하여 시각적 어휘(visual vocabulary)를 생성한 후, 시각적 어휘를 이용하여 다음으로 입력되는 이미지로부터 가중 BoF(weighted bag-of features) 히스토그램을 생성하고, 생성된 가중 BoF 히스토그램에 대하여 랜덤 포레스트(random forest)에 의한 학습을 수행하여 사람 행동을 분류함으로써, 실시간으로 사람 행동을 인식함과 동시에 신속하고 정확하게 사람 행동을 인식하는 것이 가능하다.

대 표 도 - 도1



이 발명을 지원한 국가연구개발사업

과제고유번호 2012-05-A-03-031-12-010100

부처명 교육과학기술부

연구사업명 지역혁신 인력양성 사업

연구과제명 시각장애인을 위한 비전기반 장면해석 기술 개발

기 여 율 1/1

주관기관 계명대학교

연구기간 2012.05.01 ~ 2013.03.31

특허청구의 범위

청구항 1

- (1) 입력 이미지로부터 사람을 포함하는 관심 영역을 지정하는 단계;
- (2) 지정된 상기 관심 영역으로부터 중심대칭-국부이진 패턴(Center Symmetric Local Binary Pattern; 이하 ‘CS-LBP’ 라 함) 특징을 추출하는 단계;
- (3) 추출된 상기 CS-LBP 특징을 클러스터링하여 시각적 어휘(visual vocabulary)를 생성하는 단계;
- (4) 상기 시각적 어휘를 이용하여 다음으로 입력되는 이미지로부터 가중 BoF(weighted bag-of features) 히스토그램을 생성하는 단계; 및
- (5) 상기 생성된 가중 BoF 히스토그램에 대하여 랜덤 포레스트(random forest)에 의한 학습을 수행하여 사람 행동을 분류하는 단계를 포함하는 것을 특징으로 하는, 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법.

청구항 2

제1항에 있어서, 상기 단계 (2)는,

- (2-1) 상기 관심 영역을 균등 격자(regular grid)로 분할하는 단계; 및
- (2-2) 분할된 상기 균등 격자의 각 패치로부터 CS-LBP 특징을 추출하는 단계를 포함하는 것을 특징으로 하는, 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법.

청구항 3

제2항에 있어서, 상기 단계 (2-1)에서,

상기 균등 격자는 8 픽셀의 간격을 갖는 것을 특징으로 하는, 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법.

청구항 4

제2항에 있어서, 상기 단계 (2-2)에서,

상기 패치는 16×16 크기를 갖는 것을 특징으로 하는, 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법.

청구항 5

제2항에 있어서, 상기 단계 (2-2)에서,

상기 각 패치로부터 16개의 CS-LBP 특징이 추출되는 것을 특징으로 하는, 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법.

청구항 6

제1항에 있어서, 상기 단계 (3)에서는,

k-평균 클러스터링 기법을 시각적 어휘를 생성하는 것을 특징으로 하는, 공간적 피라미드 기반의 가중 BoF 히스

토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법.

청구항 7

제6항에 있어서,

상기 시각적 어휘의 크기는 200($k=200$)인 것을 특징으로 하는, 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법.

청구항 8

제1항에 있어서, 상기 단계 (4)는,

(4-1) 상기 다음으로 입력되는 이미지에서 지정된 관심 영역에 대해 복수 개의 세분화된 레벨을 갖는 공간적 피라미드(spatial pyramid)를 생성하는 단계;

(4-2) 각각의 상기 레벨로부터 국부 BoF 히스토그램을 생성하는 단계;

(4-3) 상기 국부 BoF 히스토그램의 특징을 상기 시각적 어휘를 구성하는 시각적 단어(visual word)에 할당하고, 상기 시각적 단어의 특징에 대해 가중치를 산출하는 단계;

(4-4) 산출된 상기 가중치를 각각의 상기 국부 BoF 히스토그램에 누적시키는 단계; 및

(4-5) 누적된 각각의 상기 국부 BoF 히스토그램을 하나의 특징 벡터로 연결하는 단계를 포함하는 것을 특징으로 하는, 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법.

청구항 9

제8항에 있어서, 상기 단계 (4-1)에서는,

세 개의 세분화된 레벨(레벨 0, 레벨 1 및 레벨 2)을 갖는 공간적 피라미드를 생성하는 것을 특징으로 하는, 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법.

청구항 10

제9항에 있어서, 상기 레벨 0은,

1(1×1)개의 블록을 갖는 것을 특징으로 하는, 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법.

청구항 11

제9항에 있어서, 상기 레벨 1은,

4(2×2)개의 블록을 갖는 것을 특징으로 하는, 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법.

청구항 12

제9항에 있어서, 상기 레벨 2는,

16(4×4)개의 블록을 갖는 것을 특징으로 하는, 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법.

청구항 13

제8항에 있어서, 상기 단계 (4-3)은,

(4-3-1) 상기 시각적 단어(visual word) 중 각각의 상기 국부 BoF 히스토그램의 특징에 가장 가까운 복수 개의 시각적 단어를 검출하는 단계;

(4-3-2) 각각의 상기 특징과 상기 시각적 단어 사이의 유클리디안 거리(Euclidean Distance)를 계산하는 단계; 및

(4-3-3) 상기 계산된 거리에 대한 가중 함수(weighting function)를 이용하여 각각의 상기 복수 개의 시각적 단어에 대해 가중치를 산출하는 단계를 포함하는 것을 특징으로 하는, 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법.

청구항 14

제13항에 있어서, 상기 단계 (4-3)은,

상기 이미지에서 상기 시각적 단어에 대한 가중치를 0으로 세팅하는 전처리 단계를 더 포함하는 것을 특징으로 하는, 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법.

청구항 15

제1항에 있어서, 상기 단계 (5)는,

(5-1) 결정 트리의 최대 트리 깊이를 선택하는 단계; 및

(5-2) 상기 결정 트리를 성장시키는 단계를 포함하는 것을 특징으로 하는, 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법.

청구항 16

제15항에 있어서, 상기 단계 (5)는,

(5-3) 모든 트리에서의 각 분포 확률의 평균을 산출하는 단계를 더 포함하는 것을 특징으로 하는, 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법.

청구항 17

제15항에 있어서, 상기 단계 (5-2)는,

(5-2-1) 훈련 세트로부터 새로운 부트스트랩 샘플을 선택하고, 상기 부트스트랩 샘플을 이용하여 가지치기되지 않은(unpruned) 트리를 성장시키는 단계;

(5-2-2) 상기 트리의 각 내부 노드가 랜덤하게 변수를 선택하고, 선택된 상기 변수만을 이용하여 최상(best)의 스플릿 함수(split function)를 결정하는 단계; 및

(5-2-3) 상기 트리를 최대 트리 깊이로 성장시키는 단계를 포함하는 것을 특징으로 하는, 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법.

명세서

기술분야

[0001] 본 발명은 사람 행동 인식 방법에 관한 것으로서, 보다 구체적으로는 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법에 관한 것이다.

배경 기술

[0002] 멀티미디어 데이터양의 급격한 증가와 함께, 멀티미디어 데이터의 자동 분석이 새로운 연구 분야로 발전하고 있다. 특히, 사람 행동 인식은, 비디오 감시, 인덱싱(indexing), 검색 및 사람-컴퓨터 상호작용과 같은 광범위한 적용 때문에, 컴퓨터 비전에서 중요한 문제 중 하나이다. 비디오에서 사람 행동을 인식하는 대부분의 기존 방법은 모션 정보를 사용하는데, 그 이유는 모션 정보가 행동 분류를 위한 식별되는 단서를 제공하기 때문이다. 한편, 모션 단서가 다양한 행동을 구별하기 위한 정보를 제공하지만, 많은 행동은 하나의 이미지로부터 확인될 수 있다. 서 있거나 독서와 같은 몇몇 타입의 행동은, 사실상 정적 상태인바, 비디오를 이용할 수 있다고 하더라도 정적 단서 기반의 인식 방법을 필요로 할 수 있다.

[0003] 반면, 스틸 이미지에서의 행동 인식은 그동안 자세히 연구되지 않았다. 하지만, 스틸 이미지에서의 사람 행동 인식은, 이미지 분류, 주석(annotation), 인덱싱, 및 검색에서 많은 잠재적인 응용을 가지고 있으며, 사람의 정적 포즈 및 외관(appearance)은 행동을 추론하는데 중요한 단서이다. 예컨대, 걷는 것과 뛰는 것은 특정한 포즈와 관련되는 반면, 자전거나 말을 타는 것은 포즈 및 외관과 관련되며, 이들은 모두 스틸 이미지에서 변별적 특징이 된다. 따라서 최근에는 스틸 이미지에서 행동 인식을 하기 위한 다양한 연구가 진행되고 있다.

[0004] Ikizler 등(N. Ikizler, R.G. Cinbis, S. Pehlivan, and P. Duygulu. Recognizing actions from still images. International Conference of Pattern Recognition, 1-4, 2008)은 사람의 모습을 분석하고, 사각형의 원형 히스토그램을 사용한 공간적 및 방향성 비닝(spatial and orientational binning)을 통해 이를 표현하는 것을 제안하였다. 그리고 나서, 콤팩트하고 차별되는 특징을 얻기 위해 LDA(linear discriminant analysis)를 사용하고, 행동 분류를 위해 이진화 SVM(support vector machines)을 사용하는 것을 개시하였다.

[0005] Thureau 및 Hlavac(C. Thureau and V. Hlavac. Pose Primitive based human action recognition in videos or still images. IEEE International Conference on Pattern Recognition, 1-8, 2008)는 포즈 프리미티브(pose primitive)를 기반으로 사람 행동을 인식하는 것을 제안하였다. 즉, 포즈 표현을 위해, 기존 HOG 기반의 포즈 기술자를 연장하는 비음수 행렬분해를 이용하며, 행동 클래스는 포즈 프리미티브의 히스토그램에 의해 표현되고, 행동 인식은 단순 히스토그램 비교에 기초하는 것을 개시하였다.

[0006] 하지만 상기 개시된 연구들은, 신체 부위를 식별하고 신체 구조의 사전 지식을 이용하여 전체 포즈를 구축함으로써 행동을 인식하기 위한 단서로서, 포즈 표현(pose representation)을 사용하는 방법을 제안하고 있지만, 이러한 방법들은 전형적인 행동 인식을 위해서는 긍정적인 결과를 얻을 수 있는 반면, 큰 폐색 및 카메라 각도의 큰 변화로 인하여 이미지에서 부위를 감지하는 것이 어려운 경우에는 긍정적인 결과를 얻지 못하는 문제점이 있다.

[0007] 또한, Maji 등(S. Maji, L. Bourdev, and J. Malik. Action recognition from a distributed representation of pose and appearance. IEEE International Conference on Computer Vision and Pattern Recognition, 3177-3184, 2011)은 포즈 및 멀티-스케일 방식 측면의 고유 모호성(inherent ambiguity)을 캡처하기 위해 poselet 활성 벡터를 사용하는 것을 제안하였다.

[0008] Wang 등(J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong. Locality-constrained linear coding for image classification, IEEE International Conference on Computer Vision and Pattern Recognition, 3360-3367, 2010)은 종래의 SPM에서 VQ 코딩 대신에 LLC(Locality-constrained Linear Coding,)라 일컫는 효과적인 코딩 방식을 제안하였다. 즉, LLC는, 각 기술자를 국부 좌표계로 나타내기 위해 지역 제약(locality

constraint)을 이용하고, 표현된 좌표는 최종 표현을 생성하기 위한 최대 풀링(pooling)에 의해 통합된다.

[0009] Delaitre 등(V. Delaitre, I. Laptev, and J. Sivic, Recognizing human action in still images: a study of bag-of-features and partial-based representations. British Machine Vision Conference, 1-11, 2010)은 부분 기반(part-based) 잠재 SVM 접근과 결합한 BoF를 사용하여 스틸 이미지에 대한 행동 인식을 제안하였다. 게다가, 그들은 (i) 통계적 및 부분 기반 표현을 결합함으로써 행동 인식 성능을 향상시키고, (ii) 배경과 사람-중심의 기술(description)을 통합하는 것을 제안하였다.

[0010] 하지만, 상기 개시된 방법들은 계산 속도가 느릴 뿐만 아니라 차원 수가 커서 실시간을 사람 행동을 인식할 수 없는 문제점이 있다.

발명의 내용

해결하려는 과제

[0011] 본 발명은 기존에 제안된 방법들의 상기와 같은 문제점들을 해결하기 위해 제안된 것으로서, 입력 이미지로부터 사람을 포함하는 관심 영역을 지정하고, 지정된 관심 영역으로부터 중심대칭-국부이진 패턴(Center Symmetric Local Binary Pattern; 이하 ‘CS-LBP’ 라 함) 특징을 추출하여, 추출된 CS-LBP 특징을 클러스터링 하여 시각적 어휘(visual vocabulary)를 생성한 후, 시각적 어휘를 이용하여 다음으로 입력되는 이미지로부터 가중 BoF(weighted bag-of features) 히스토그램을 생성하고, 생성된 가중 BoF 히스토그램에 대하여 랜덤 포레스트(random forest)에 의한 학습을 수행하여 사람 행동을 분류함으로써, 실시간으로 사람 행동을 인식함과 동시에 신속하고 정확하게 사람 행동을 인식하는 것이 가능한, 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법을 제공하는 것을 그 목적으로 한다.

과제의 해결 수단

[0012] 상기한 목적을 달성하기 위한 본 발명의 특징에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법은,

[0013] (1) 입력 이미지로부터 사람을 포함하는 관심 영역을 지정하는 단계;

[0014] (2) 지정된 상기 관심 영역으로부터 중심대칭-국부이진 패턴(Center Symmetric Local Binary Pattern; 이하 ‘CS-LBP’ 라 함) 특징을 추출하는 단계;

[0015] (3) 추출된 상기 CS-LBP 특징을 클러스터링하여 시각적 어휘(visual vocabulary)를 생성하는 단계;

[0016] (4) 상기 시각적 어휘를 이용하여 다음으로 입력되는 이미지로부터 가중 BoF(weighted bag-of features) 히스토그램을 생성하는 단계; 및

[0017] (5) 상기 생성된 가중 BoF 히스토그램에 대하여 랜덤 포레스트(random forest)에 의한 학습을 수행하여 사람 행동을 분류하는 단계를 포함하는 것을 그 구성상의 특징으로 한다.

[0018] 바람직하게는, 상기 단계 (2)는,

[0019] (2-1) 상기 관심 영역을 균등 격자(regular grid)로 분할하는 단계; 및

[0020] (2-2) 분할된 상기 균등 격자의 각 패치로부터 CS-LBP 특징을 추출하는 단계를 포함할 수 있다.

[0021] 더욱 바람직하게는, 상기 단계 (2-1)에서,

[0022] 상기 균등 격자는 8 픽셀의 간격을 가질 수 있다.

[0023] 더욱 바람직하게는, 상기 단계 (2-2)에서,

- [0024] 상기 패치는 16×16 크기를 가질 수 있다.
- [0025] 더욱 바람직하게는, 상기 단계 (2-2)에서,
- [0026] 상기 각 패치로부터 16개의 CS-CBP 특징이 추출될 수 있다.
- [0027] 바람직하게는, 상기 단계 (3)에서는,
- [0028] k-평균 클러스터링 기법을 시각적 어휘를 생성할 수 있다.
- [0029] 더욱 바람직하게는,
- [0030] 상기 시각적 어휘의 크기는 $200(k=200)$ 일 수 있다.
- [0031] 바람직하게는, 상기 단계 (4)는,
- [0032] (4-1) 상기 다음으로 입력되는 이미지에서 지정된 관심 영역에 대해 복수 개의 세분화된 레벨을 갖는 공간적 피라미드(spatial pyramid)를 생성하는 단계;
- [0033] (4-2) 각각의 상기 레벨로부터 국부 BoF 히스토그램을 생성하는 단계;
- [0034] (4-3) 상기 국부 BoF 히스토그램의 특징을 상기 시각적 어휘를 구성하는 시각적 단어(visual word)에 할당하고, 상기 시각적 단어의 특징에 대해 가중치를 산출하는 단계;
- [0035] (4-4) 산출된 상기 가중치를 각각의 상기 국부 BoF 히스토그램에 누적시키는 단계; 및
- [0036] (4-5) 누적된 각각의 상기 국부 BoF 히스토그램을 하나의 특징 벡터로 연결하는 단계를 포함할 수 있다.
- [0037] 더욱 바람직하게는, 상기 단계 (4-1)에서는,
- [0038] 세 개의 세분화된 레벨(레벨 0, 레벨 1 및 레벨 2)을 갖는 공간적 피라미드를 생성할 수 있다.
- [0039] 더더욱 바람직하게는, 상기 레벨 0은,
- [0040] $1(1 \times 1)$ 개의 블록을 가질 수 있다.
- [0041] 더더욱 바람직하게는, 상기 레벨 1은,
- [0042] $4(2 \times 2)$ 개의 블록을 가질 수 있다.
- [0043] 더더욱 바람직하게는, 상기 레벨 2는,
- [0044] $16(4 \times 4)$ 개의 블록을 가질 수 있다.
- [0045] 더욱 바람직하게는, 상기 단계 (4-3)은,
- [0046] (4-3-1) 상기 시각적 단어(visual word) 중 각각의 상기 국부 BoF 히스토그램의 특징에 가장 가까운 복수 개의 시각적 단어를 검출하는 단계;
- [0047] (4-3-2) 각각의 상기 특징과 상기 시각적 단어 사이의 유클리디안 거리(Euclidean Distance)를 계산하는 단계; 및

- [0048] (4-3-3) 상기 계산된 거리에 대한 가중 함수(weighting function)를 이용하여 각각의 상기 복수 개의 시각적 단어에 대해 가중치를 산출하는 단계를 포함할 수 있다.
- [0049] 더더욱 바람직하게는, 상기 단계 (4-3)은,
- [0050] 상기 이미지에서 상기 시각적 단어에 대한 가중치를 0으로 세팅하는 전처리 단계를 더 포함할 수 있다.
- [0051] 바람직하게는, 상기 단계 (5)는,
- [0052] (5-1) 결정 트리의 최대 트리 깊이를 선택하는 단계; 및
- [0053] (5-2) 상기 결정 트리를 성장시키는 단계를 포함할 수 있다.
- [0054] 더욱 바람직하게는, 상기 단계 (5)는,
- [0055] (5-3) 모든 트리에서의 각 분포 확률의 평균을 산출하는 단계를 더 포함할 수 있다.
- [0056] 더욱 바람직하게는, 상기 단계 (5-2)는,
- [0057] (5-2-1) 훈련 세트로부터 새로운 부트스트랩 샘플을 선택하고, 상기 부트스트랩 샘플을 이용하여 가지치기되지 않은(unpruned) 트리를 성장시키는 단계;
- [0058] (5-2-2) 상기 트리의 각 내부 노드가 랜덤하게 변수를 선택하고, 선택된 상기 변수만을 이용하여 최상(best)의 스플릿 함수(split function)를 결정하는 단계; 및
- [0059] (5-2-3) 상기 트리를 최대 트리 깊이로 성장시키는 단계를 포함할 수 있다.

발명의 효과

- [0060] 본 발명에서 제안하고 있는 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법에 따르면, 입력 이미지로부터 사람을 포함하는 관심 영역을 지정하고, 지정된 관심 영역으로부터 중심대칭-국부이진 패턴(Center Symmetric Local Binary Pattern; 이하 ‘CS-LBP’ 라 함) 특징을 추출하여, 추출된 CS-LBP 특징을 클러스터링 하여 시각적 어휘(visual vocabulary)를 생성한 후, 시각적 어휘를 이용하여 다음으로 입력되는 이미지로부터 가중 BoF(weighted bag-of features) 히스토그램을 생성하고, 생성된 가중 BoF 히스토그램에 대하여 랜덤 포레스트(random forest)에 의한 학습을 수행하여 사람 행동을 분류함으로써, 실시간으로 사람 행동을 인식함과 동시에 신속하고 정확하게 사람 행동을 인식하는 것이 가능하다.

도면의 간단한 설명

- [0061] 도 1은 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 순서도.
- 도 2는 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S200에 대한 세부 순서도.
- 도 3은 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S400에 대한 세부 순서도.
- 도 4는 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법에서, 공간적 피라미드를 생성한 상태를 나타낸 도면.
- 도 5는 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S430에 대한 세부 순서도.
- 도 6은 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S500에 대한 세부 순서도.

도 7은 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S500에 의해 처리된 상태를 나타낸 도면.

도 8은 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S520에 대한 세부 순서도.

도 9는 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 성능평가 실험에서, 가장 높은 정확률을 가진 10개의 행동 클래스의 confusion table을 나타낸 도면.

도 10은 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 성능평가 실험에서, 평균 정확률 측정 결과를 나타낸 도면.

도 11은 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법에 의해 높은 정확률로 분류된 이미지의 예를 나타낸 도면.

도 12는 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법에 의해 거짓 분류된 이미지의 예를 나타낸 도면.

발명을 실시하기 위한 구체적인 내용

[0062] 이하, 첨부된 도면을 참조하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 본 발명을 용이하게 실시할 수 있도록 바람직한 실시 예를 상세히 설명한다. 다만, 본 발명의 바람직한 실시 예를 상세하게 설명함에 있어, 관련된 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략한다. 또한, 유사한 기능 및 작용을 하는 부분에 대해서는 도면 전체에 걸쳐 동일한 부호를 사용한다.

[0063] 덧붙여, 명세서 전체에서, 어떤 부분이 다른 부분과 ‘연결’되어 있다고 할 때, 이는 ‘직접적으로 연결’되어 있는 경우뿐만 아니라, 그 중간에 다른 소자를 사이에 두고 ‘간접적으로 연결’되어 있는 경우도 포함한다. 또한, 어떤 구성요소를 ‘포함’한다는 것은, 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있다는 것을 의미한다.

[0064] 도 1은 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 순서도이다. 도 1에 나타낸 바와 같이, 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법은, 입력 이미지로부터 사람을 포함하는 관심 영역을 지정하는 단계(S100), 지정된 관심 영역으로부터 중심대칭-국부이진 패턴(Center Symmetric Local Binary Pattern; 이하 ‘CS-LBP’라 함) 특징을 추출하는 단계(S200), 추출된 CS-LBP 특징을 클러스터링하여 시각적 어휘(visual vocabulary)를 생성하는 단계(S300), 시각적 어휘를 이용하여 다음으로 입력되는 이미지로부터 가중 BoF(weighted bag-of features) 히스토그램을 생성하는 단계(S400), 및 생성된 가중 BoF 히스토그램에 대하여 랜덤 포레스트(random forest)에 의한 학습을 수행하여 사람 행동을 분류하는 단계(S500)를 포함하여 구성될 수 있다.

[0065] 단계 S100에서는, 입력 이미지로부터 관심 영역을 지정하며, 보다 구체적으로는 입력 이미지에서 사람을 포함하는 영역을 크로핑(cropping)하여, 크로핑 된 영역을 관심 영역으로 지정한다.

[0066] 단계 S200에서는, 단계 S100에 의해 지정된 관심 영역으로부터 중심대칭-국부이진 패턴(Center Symmetric Local Binary Pattern; 이하 ‘CS-LBP’라 함) 특징을 추출하며, 도 2는 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S200에 대한 세부 순서도이다. 도 2에 나타낸 바와 같이, 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S200은, 관심 영역을 균등 격자(regular grid)로 분할하는 단계(S210) 및 분할된 균등 격자의 각 패치로부터 CS-LBP 특징을 추출하는 단계(S220)를 포함할 수 있다.

- [0067] 단계 S210에서는, 단계 S100에 의해 지정된 관심 영역을 분할하며, 보다 구체적으로는, 8 픽셀의 간격을 갖는 밀집된 균등 격자(dense regular grid)로 분할함으로써, 이를 관심 포인트 대신에 사용한다.
- [0068] 단계 S220에서는, 단계 S210에 의해 분할된 균등 격자의 각 패치로부터 CS-LBP 특징을 추출하는데, 여기서 패치란 균등 격자를 구성하는 각각의 셀(cell)을 지칭하며, 각 패치는 16×16 크기를 갖는 것이 바람직하다.
- [0069] 일반적으로, 이미지에서 객체의 검출과 분류를 위해 많이 사용되는 기술자(descriptor)로 국부이진 패턴(Local Binary Pattern; 이하 'LBP' 라 함)이나 SIFT(scale-invariant feature transform) 등을 사용한다. 하지만, 이들은 우수한 성능에 비해 특징의 차원 수가 커서 실시간 객체 추적이 불가능한 문제점이 있다. 따라서 본 발명에 따른 사람 행동 인식 방법에서는, LBP와 비슷한 성능을 가지는 반면, 더 적은 특징 차원을 갖는 질감 특징으로서 CS-LBP를 사용하는 것이 바람직하다.
- [0070] CS-LBP는 중심 픽셀의 이웃하는 픽셀들의 관계를 나타내는 기술자로서, 이는 LBP와 동일한 성능을 나타내면서 차원 수를 현저히 줄일 수 있다. 즉, CS-LBP는 중심 픽셀과 이와 중심 대칭인 픽셀 쌍을 비교하는 반면, LBP는 중심과 각 픽셀을 비교한다는 점에서 서로 상이하며, 이를 통해 CS-LBP는 조명 변화에 강하고, 단일 톤(monotonic)의 그레이 레벨의 변화에 대한 견고성(robustness)과 같은 특징을 유지하면서 계산을 단순화하는 장점이 있다. 보다 구체적으로, LBP는 256개의 상이한 이진화 패턴을 생성하는 반면, CS-LBP는 각 패치로부터 오직 16개의 CS-LBP 특징만을 생성한다. 한편, CS-LBP 추출 후, 패치의 전체 기울기가 크기(magnitude)에서 너무 완만하지(weak) 않는 한, L2 정규화 단계가 각 패치 안에서 수행되는 것이 바람직하다.
- [0071] 단계 S300에서는, 단계 S200에 의해 추출된 CS-LBP 특징을 클러스터링하여 시각적 어휘(visual vocabulary)를 생성하며, 보다 구체적으로는 k-평균 클러스터링(k-means clustering)을 사용하여 시각적 어휘(visual vocabularies)인 코드북(codebook)을 생성하며, 이때 시각적 어휘의 크기는 $k=200$ 으로 결정하는 것이 바람직하다. 각 클러스터는 시각적 어휘 내의 시각적 단어(코드워드, codeword)로 처리되고, 이미지의 국부 특징을 시각적 어휘와 매핑(mapping)함으로써, 각 시각적 단어의 존재 또는 총수에 따라 이미지를 특징 히스토그램으로 기술할 수 있다.
- [0072] 단계 S400에서는, 단계 S300에 의해 생성된 시각적 어휘를 이용하여 다음으로 입력되는 이미지로부터 가중 BoF(weighted bag-of features) 히스토그램을 생성하며, 단계 S400에 대하여 도 3 내지 도 5를 참조하여 상세히 설명하도록 한다.
- [0073] 도 3은 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S400에 대한 세부 순서도이다. 도 3에 나타난 바와 같이, 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S400은, 다음으로 입력되는 이미지에서 지정된 관심 영역에 대해 복수 개의 세분화된 레벨을 갖는 공간적 피라미드(spatial pyramid)를 생성하는 단계(S410), 각각의 레벨로부터 국부 BoF 히스토그램을 생성하는 단계(S420), 국부 BoF 히스토그램의 특징을 시각적 어휘를 구성하는 시각적 단어(visual word)에 할당하고, 시각적 단어의 특징에 대해 가중치를 산출하는 단계(S430), 산출된 가중치를 각각의 국부 BoF 히스토그램에 누적시키는 단계(S440) 및 누적된 각각의 국부 BoF 히스토그램을 하나의 특징 벡터로 연결하는 단계(S450)를 포함할 수 있다.
- [0074] 단계 S410에서는, 다음으로 입력되는 이미지에서 지정된 관심 영역에 대해 복수 개의 세분화된 레벨을 갖는 공간적 피라미드(spatial pyramid)를 생성하며, 도 4는 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법에서, 공간적 피라미드를 생성한 상태를 나타낸

도면이다. 도 4에 나타난 바와 같이, 세 개의 세분화된 레벨(레벨 0, 레벨 1 및 레벨 2)을 갖는 공간적 피라미드를 생성하는 것이 바람직하며, 레벨 0, 레벨 1 및 레벨 2는 각각 $1(1 \times 1)$, $4(2 \times 2)$ 및 $16(4 \times 4)$ 블록을 가질 수 있다.

[0075] 단계 S420에서는, 각각의 레벨로부터 국부 BoF 히스토그램을 생성한다. 즉, 각 레벨의 해상도(resolution)에 대해 국부 BoF 히스토그램이 추정되며, 예컨대, 도 4에서 k 가 200이면, 추정된 BoF 히스토그램의 특징 차원은 레벨 0에서 $200(1 \times 1 \times 200)$ 이고, 마찬가지로, 레벨 1에서는 800이고, 레벨 2에서는 3200이며, 결과적으로 BoF 히스토그램의 총 차원은 $4200(3200+800+200)$ 이 된다.

[0076] 단계 S430에서는, 국부 BoF 히스토그램의 특징을 시각적 어휘를 구성하는 시각적 단어(visual word)에 할당하고, 시각적 단어의 특징에 대해 가중치를 산출한다.

[0077] 일반적으로, K -차원 BoF 히스토그램은 이진 가중치(binary weighting)에 의해 패치로부터 추정되는데, 이진 가중치는 시각적 단어의 존재 또는 부존재를 각각 1 및 0으로 나타낸다. 모든 가중치 방식은, 각각의 키포인트가 가장 유사한 시각적 단어로 매핑된다는 점에서, 어휘에서 가장 가까운 이웃의 검색을 수행한다.

[0078] 한편, 시각적 단어의 중요성을 강조하기 위해, 이미지 내에서 각각의 키포인트의 가장 가까운 시각적 단어만을 검색하는 방법 대신, N 개의 가장 가까운 시각적 단어를 선택하고, 거리의 합에 따라 상이한 가중치를 할당하는 소프트-가중치 방식(soft-weighting approach)을 채택하는 것이 바람직하다. 하지만, 시각적 어휘의 크기가 커지면 두 개의 유사한 포인트가 상이한 클러스터 내에 위치할 수 있기 때문에, 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S430에서는, 각각의 거리에 따라 상이한 가중치를 N 개의 가장 가까운 시각적 단어에 할당하며, 이에 대해서는 도 5를 참조하여 상세히 설명하도록 한다.

[0079] 도 5는 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S430에 대한 세부 순서도이다. 도 5에 나타난 바와 같이, 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S430은, 시각적 단어(visual word) 중 각각의 국부 BoF 히스토그램의 특징에 가장 가까운 복수 개의 시각적 단어를 검출하는 단계(S4310), 각각의 특징과 시각적 단어 사이의 유클리디안 거리(Euclidean Distance)를 계산하는 단계(S432) 및 계산된 거리에 대한 가중 함수(weighting function)를 이용하여 각각의 복수 개의 시각적 단어에 대해 가중치를 산출하는 단계(S433)를 포함할 수 있다.

[0080] 단계 S431에서는, 시각적 단어(visual word) 중 각각의 국부 BoF 히스토그램의 특징에 가장 가까운 복수 개의 시각적 단어를 검출한다. 한편, 단계 S431 이전에, 이미지에서 시각적 단어에 대한 가중치를 0으로 세팅하는 전처리 단계를 수행하는 것이 바람직하다. 즉, K -차원 BoF 히스토그램 $T=\{t_1, \dots, t_k\}$ 의 모든 요소를 0으로 세팅한다. 각 요소 t_k 는 이미지 I 에서 시각적 단어 k 의 가중치를 나타낸다.

[0081] 단계 S432에서는, 각각의 특징과 시각적 단어 사이의 유클리디안 거리(Euclidean Distance)를 계산하며, 보다 구체적으로는 하기의 수학적 식 1에 따라 특징과 시각적 단어 사이의 유클리디안 거리(D)를 계산한다.

수학적 식 1

[0082]
$$D_{i=1 \dots N} = \|P - V_i\|$$

[0083] 여기서, P는 패치, i는 이미지이고, i=0이다.

[0084] 단계 S433에서는, 단계 S432에 의해 계산된 거리에 대한 가중 함수(weighting function)를 이용하여 각각의 복수 개의 시각적 단어에 대해 가중치를 산출하며, 보다 구체적으로는, 하기의 수학식 2 및 수학식 3에 따라 시각적 단어 각각에 대해 가중치를 산출한다.

수학식 2

$$[0085] \quad t_k = t_k + W(D_i)$$

수학식 3

$$[0086] \quad W(D_i) = \frac{1}{\exp(T \cdot D_i)}, T > 0$$

[0087] 수학식 2에 따라 패치 P의 기여도(contribution)를 D_i 에 상응하는 t_k 의 요소에 추가하며, 수학식 3에서, T는 가중치 함수 W()에서 D_i 의 영향을 최대화(최소화)하기 위해 선택될 수 있다. T가 1이면, D_i 의 변화는 기하급수적으로 W()에 반영된다. 지수 가중은 국부 특징 검색(relevance)에서의 변화에 더욱 민감하고, 큰 성능 향상을 야기할 수 있으며, 바람직하게는 T를 0.5로 세팅한다.

[0088] 단계 S440에서는, 단계 S430에 의해 산출된 가중치를 각각의 국부 BoF 히스토그램에 누적시키며, 예컨대, 도 4에서, 각 블록 내의 국부 BoF 히스토그램은 레벨 0, 레벨 1 및 레벨 2 각각에 대하여 0.25, 0.25 및 0.5의 가중치가 산출되어 누적될 수 있다.

[0089] 단계 S450에서는, 단계 S440에 의해 누적된 각각의 국부 BoF 히스토그램을 하나의 특징 벡터로 연결한다.

[0090] 단계 S500에서는, 단계 S400에 의해 생성된 가중 BoF 히스토그램에 대하여 랜덤 포레스트(random forest)에 의한 학습을 수행하여 사람 행동을 분류한다.

[0091] 랜덤 포레스트는 몇 가지 타입의 랜덤화를 사용하여 자란 각 트리를 가진 결정 트리 앙상블 분류기로서, 단일 결정 트리보다는 이진 결정 트리의 집합으로 하향식(top-down)을 기본으로 하며, 이러한 결정 트리를 기반으로 빠른 훈련 속도를 가질 뿐만 아니라 많은 양의 데이터를 처리하기 위한 용량도 가지고 있다.

[0092] 이하에서는, 단계 S500에 대하여 도 6 및 도 7을 참조하여 상세히 설명하도록 한다.

[0093] 도 6은 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S500에 대한 세부 순서도이고, 도 7은 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S500에 의해 처리된 상태를 나타낸 도면이다. 도 6에 나타난 바와 같이, 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S500은, 결정 트리의 최대 트리 깊이를 선택하는 단계(S510) 및 결정 트리를 성장시키는 단계(S520)를 포함할 수 있고, 모든 트리에서의 각 분포 확률의 평균을 산출하는 단계(S530)를 더 포함할 수 있다.

[0094] 단계 S510에서는, 결정 트리의 최대 트리 깊이를 선택하여 입력한다. 즉, 트리가 연장된 최대 깊이(D_{target}), 모든 행동 샘플을 포함한 훈련 세트의 개수(S_n)를 입력하며, 초기에는 $i=0$, $j=0$, $k=0$, $F_i=1.0$ 으로 설정되고, n 부트스트랩 샘플과 이들의 BoF 히스토그램 기술자를 훈련 세트 S_n 에 할당하며 하부집합으로 분할된다.

[0095] 단계 S520에서는, 결정 트리를 성장시키며, 도 8은 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S520에 대한 세부 순서도이다. 도 8에 나타난 바와 같이, 본 발명의 일실시예에 따른 공간적 피라미드 기반의 가중 BoF 히스토그램 및 랜덤 포레스트를 이용한 사람 행동 인식 방법의 단계 S520은, 훈련 세트로부터 새로운 부트스트랩 샘플을 선택하고, 부트스트랩 샘플을 이용하여 가지치기되지 않은(unpruned) 트리를 성장시키는 단계(S521), 트리의 각 내부 노드가 랜덤하게 변수를 선택하고, 선택된 변수만을 이용하여 최상(best)의 스플릿 함수를 결정하는 단계(S522) 및 트리를 최대 트리 깊이로 성장시키는 단계(S523)를 포함할 수 있다.

[0096] 단계 S521에서는, 훈련 세트(S_n)로부터 새로운 n 부트스트랩 샘플(n bootstrap samples)을 선택하고, n 부트스트랩 샘플을 이용하여 가지치기되지 않은(unpruned) 트리를 성장시킨다.

[0097] 단계 S522에서는, 트리의 각 내부 노드가 랜덤하게 변수(variables)를 선택하고, 선택된 변수만을 이용하여 최상의 스플릿 함수를 결정한다. 즉, 다른 p -th 변수를 사용하여, 스플릿 함수 $f(v_p)$ 는 하기의 수학적 식 4에 따라, 훈련데이터를 왼쪽(I_l) 및 오른쪽(I_r) 서브세트로 반복적으로 분할한다.

수학적 식 4

$$I_l = \{p \in I_m \mid f(v_p) < t\},$$

$$I_r = I_n \setminus I_l.$$

[0098]

[0099] 여기서, 기준 값(threshold) t 는 영역 $t \in (\min_p f(v_p), \max_p f(v_p))$ 에서 스플릿 함수 $f(v_p)$ 에 의해 랜덤하게 선택된다.

[0100] 단계 S523에서는, 트리를 최대 트리 깊이로 성장시킨다. 각 결정 트리가 훈련되면, 도 7에 나타난 바와 같이, 트리의 양상불이 만들어진다.

[0101] 단계 S530에서는, 모든 트리에서의 각 분포 확률의 평균을 산출하며, 보다 구체적으로는 하기의 수학적 식 5에 따라 모든 트리 $L=(l_1, l_2, \dots, l_T)$ 에서의 각 분포 확률의 평균을 구하며, 이를 통해 최종 클래스를 선택한다.

수학적 식 5

$$P(c_i | L) = \frac{1}{T} \sum_{t=1}^T P(c_i | l_t)$$

[0102]

[0103] 여기서, T 는 트리의 개수이고, $p(c_i/L)$ 이 최댓값을 가지면, 입력 이미지의 최종 클래스로 C_i 를 선택한다. 예컨대, 도 7에서, 테스트 이미지는 두 번째 클래스 내(자전거를 타기)에 배치되는데, 이는 최대 사후 확률을 가지

기 때문이다.

- [0104] 본 발명은 이하의 실험 예에 의하여 더욱 상세히 설명되나, 본 발명이 이하의 실험 예에 의해 어떤 식으로든 제한되는 것은 아니다.
- [0105] [실험 예 1]
- [0106] 사람 행동 인식 성능 평가
- [0107] 본 발명에 따른 사람 행동 인식 방법의 성능을 평가하기 위해 하기와 같이 실험을 진행하였다.
- [0108] 뛰기, 손뼉치기, 풍선 불기, 나무 자르기 및 우산 잡기와 같은 40개의 다양한 일상적인 사람 행동을 포함한 Stanford 40 Action Dataset을 사용하여 실험을 수행하였다. 데이터베이스는 각 클래스에 대해 180~300 이미지를 제공하며, 총 9532 이미지를 제공한다. 각 클래스 내 이미지는 사람 포즈, 외관 및 배경 클러스터에 큰 변화가 있다.
- [0109] 본 발명에 따른 사람 행동 인식 방법의 성능은 평균 정확률(average precision; AP)을 측정함으로써 평가하였고, 이를 LLC(Locality-constrained Linear Coding) 방법을 통해 측정된 평균 정확률과 비교하였고, 그 결과를 도 9 및 도 10에 나타내었다.
- [0110] 도 9는 가장 높은 정확률을 가진 10개의 행동 클래스의 confusion table을 나타낸다. 도 9에 나타난 바와 같이, 최고의 행동 인식은 달리기 행동에서 일어나며, 이는 65%의 정확률을 가진다. 반대로, “프리스비(frisby) 던지기”는 42%의 정확률을 가지는데, 이것은 테스트 이미지의 배경 클러스터 및 사람의 다양한 포즈 때문인 것으로 사료된다.
- [0111] 또한, 도 10은 본 발명에 따른 사람 행동 인식 방법의 평균 정확률을 나타내며, 도 10에 나타난 바와 같이, 본 발명에 따른 사람 행동 인식 방법은 35%의 평균 정확률의 LLC와 유사한 분류 성능을 보였다. 보다 구체적으로, 도 11에 나타난 바와 같이, 달리기, 나무 자르기 및 자전거 타기와 같은 몇몇 상위 클래스는 높은 분류 성능을 얻었는데, 이는 사람의 독특한 포즈 및 배경과 구별할 수 있는 특징 때문인 것으로 사료된다. 반대로, 도 12에 나타난 바와 같이, 전화하기, 사진 찍기와 같은 몇몇 하위 클래스 낮은 분류 성능을 얻었는데, 이는 복잡한 배경, 카메라 뷰포인트 또는 불특정하고 유사한 사람 포즈 때문인 것으로 사료된다.
- [0112] [실험 예 2]
- [0113] 사람 행동 인식 처리 속도 평가
- [0114] 본 발명에 따른 사람 행동 인식 방법의 처리 속도를 평가하기 위해 하기와 같이 실험을 진행하였다.
- [0115] 즉, 본 발명에 따른 사람 행동 인식 방법과 LLC 방법을 테스트 데이터에 적용하고, 평균 처리 시간을 초로 측정함으로써 평가하였다. 그 결과, LLC 방법 및 본 발명에 따른 사람 행동 인식 방법에 대한 평균 처리 속도는, 동일한 시스템 환경을 사용할 때, 각각 이미지당 0.3s 및 0.1s이었다.
- [0116] 따라서 본 발명에 따른 사람 행동 인식 방법의 랜덤 포레스트 분류기가 LLC와 유사한 성능을 가지면서 높은 분류 속도로 많은 양의 데이터를 처리하는 용량을 가진다는 것을 확인하였다.

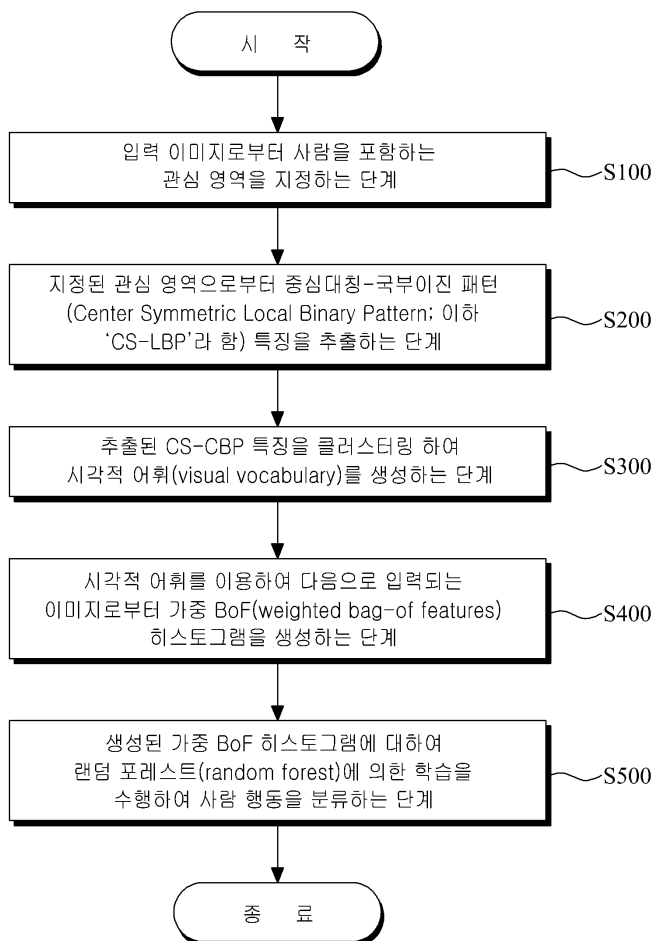
[0117] 이상 설명한 본 발명은 본 발명이 속한 기술 분야에서 통상의 지식을 가진 자에 의하여 다양한 변형이나 응용이 가능하며, 본 발명에 따른 기술적 사상의 범위는 아래의 특허청구범위에 의하여 정해져야 할 것이다.

부호의 설명

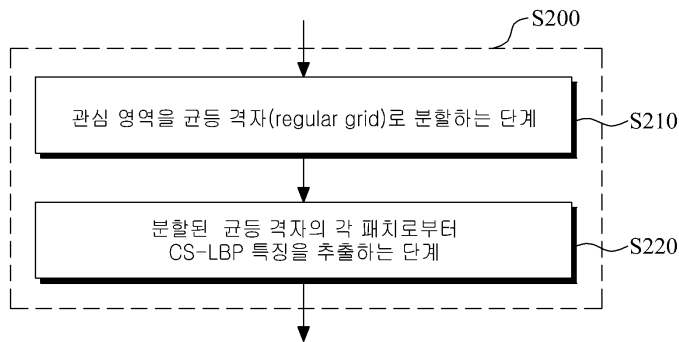
- [0118] S100: 입력 이미지로부터 사람을 포함하는 관심 영역을 지정하는 단계
- S200: 지정된 관심 영역으로부터 중심대칭-국부이진 패턴(Center Symmetric Local Binary Pattern; 이하 'CS-LBP' 라 함) 특징을 추출하는 단계
- S300: 추출된 CS-LBP 특징을 클러스터링 하여 시각적 어휘(visual vocabulary)를 생성하는 단계
- S400: 시각적 어휘를 이용하여 다음으로 입력되는 이미지로부터 가중 BoF(weighted bag-of features) 히스토그램을 생성하는 단계
- S500: 생성된 가중 BoF 히스토그램에 대하여 랜덤 포레스트(random forest)에 의한 학습을 수행하여 사람 행동을 분류하는 단계

도면

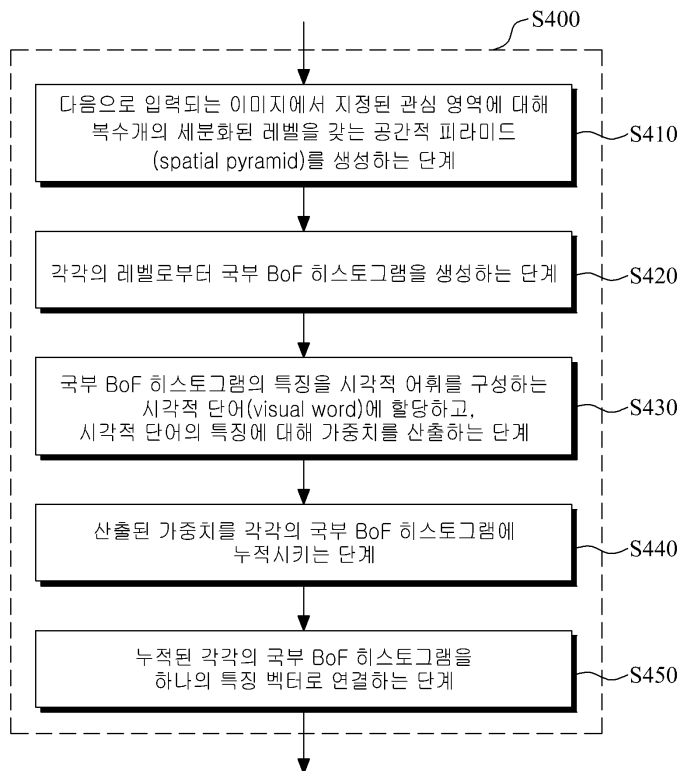
도면1



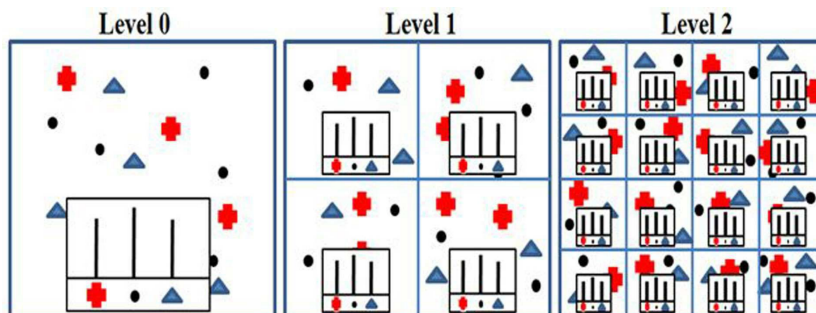
도면2



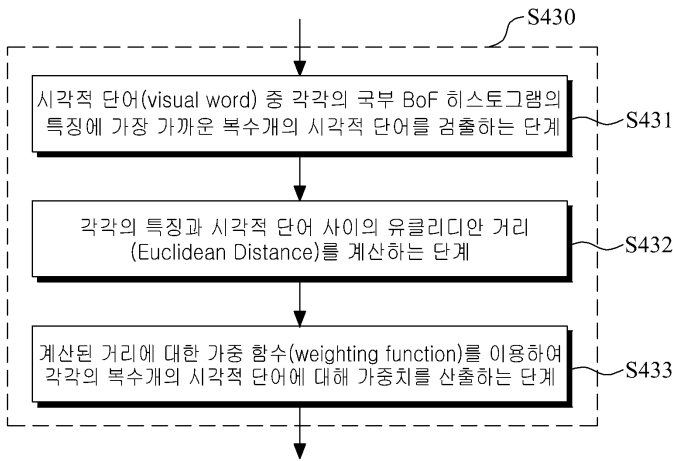
도면3



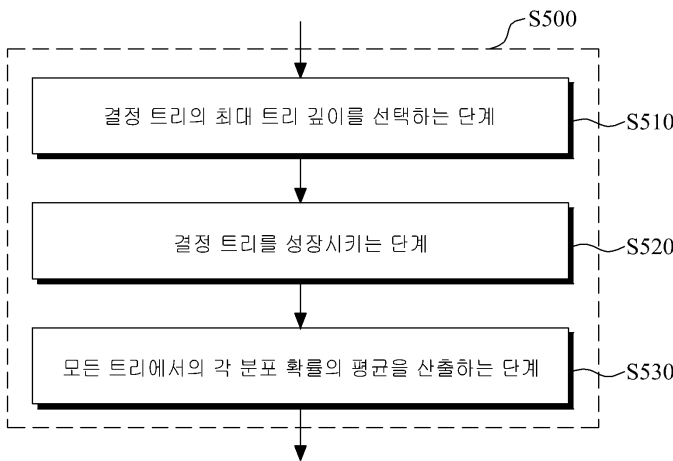
도면4



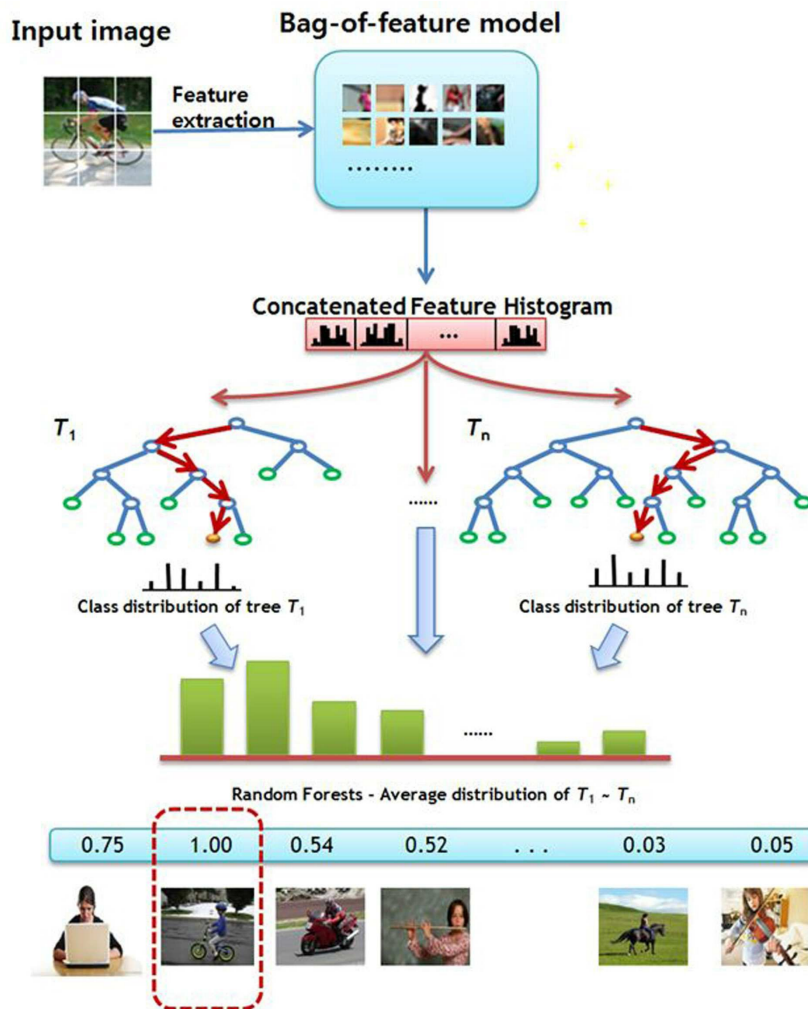
도면5



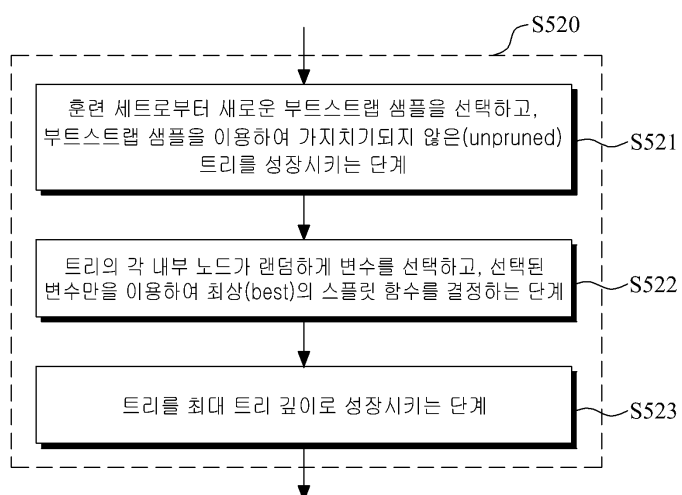
도면6



도면7



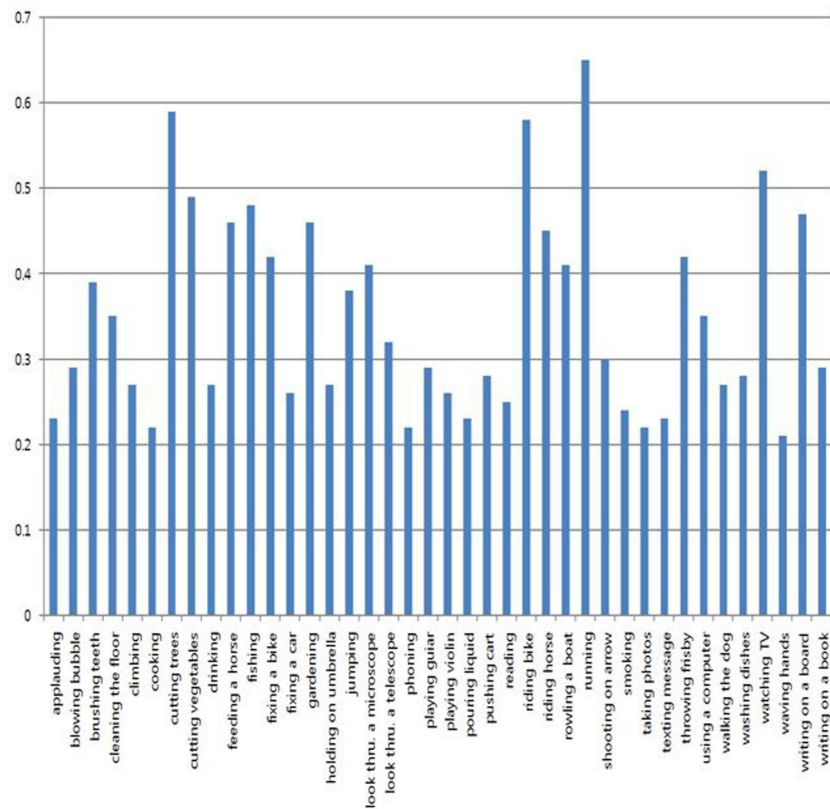
도면8



도면9

Writing a board	0.47	0.04	0.02	0.04	0.01	0.03	0.04	0.02	0.02	0.02
Watching TV	0.02	0.52	0.01	0.03	0.02	0.04	0.02	0.05	0.06	0.01
Feeding a horse	0.01	0.01	0.46	0.04	0.01	0.03	0.01	0.02	0.03	0.04
Throwing frisby	0.01	0.03	0.01	0.42	0.04	0.02	0.01	0.02	0.04	0.01
running	0.01	0.02	0.03	0.01	0.65	0.03	0.02	0.04	0.01	0.01
Riding bike	0.03	0.01	0.02	0.01	0.03	0.58	0.04	0.03	0.01	0.02
Cutting trees	0.01	0.01	0.03	0.07	0.02	0.01	0.59	0.01	0.02	0.03
Cutting vegetables	0.02	0.03	0.01	0.02	0.03	0.01	0.03	0.49	0.02	0.01
fishing	0.03	0.01	0.03	0.02	0.01	0.02	0.02	0.03	0.48	0.03
gardening	0.04	0.02	0.02	0.03	0.02	0.03	0.01	0.02	0.04	0.46
Writing a board										
Watching TV										
Feeding a horse										
Throwing frisby										
running										
Riding bike										
Cutting trees										
Cutting vegetables										
fishing										
gardening										

도면10



도면11



도면12

