(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0066125
(43) 공개일자 2020년06월09일

(51) 국제특허분류(Int. Cl.)
G06K 9/00 (2006.01) *G06K 9/62* (2006.01)
(52) CPC특허분류
G06K 9/00711 (2013.01)
G06K 9/00335 (2013.01)
(21) 출원번호 10-2019-0025744
(22) 출원일자 2019년03월06일
심사청구일자 2019년03월06일
(30) 우선권주장
1020180151637 2018년11월30일 대한민국(KR)

(71) 출원인
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
(72) 발명자
이영구
경기도 수원시 영통구 청명북로 33, 431동 1704호(영통동, 청명마을삼성아파트)
김태연
경기도 수원시 영통구 봉영로1517번길 76, 631동 506호 (영통동)
(뒷면에 계속)
(74) 대리인
특허법인도담

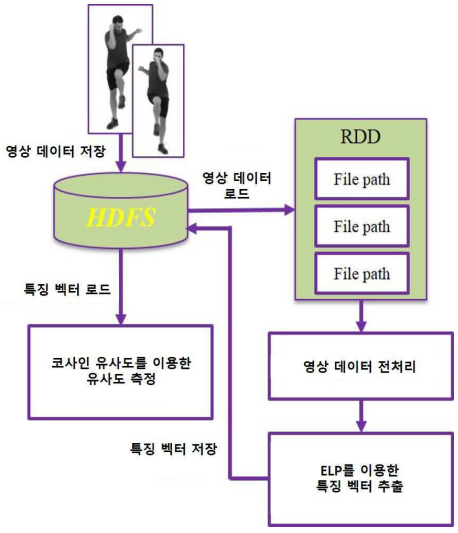
전체 청구항 수 : 총 14 항

(54) 발명의 명칭 **영상의 유사도 예측 방법 및 장치**

(57) 요 약

본 발명은 영상 데이터의 유사도를 예측하기 위한 것으로, 영상 데이터베이스로부터 제1 영상 데이터와 상기 제1 영상 데이터의 비교 대상이 되는 참고 영상 데이터를 수신하는 단계, 상기 제1 영상 데이터를 복수 개의 제2 영상 데이터로 분산 저장하는 단계, 상기 제2 영상 데이터에서 제1 프레임을 추출하는 전처리 단계, 상기 제1 프레임에서 기 설정된 크기를 갖는 에지 마스크를 검출하는 단계, 상기 에지 마스크의 영역에 대한 특징을 추출하여 제1 특징 벡터를 생성하는 단계, 상기 제1 프레임에서 생성된 하나 이상의 상기 제1 특징 벡터를 결합하여 상기 제1 프레임에 대응되는 제2 특징 벡터를 생성하는 단계, 상기 참고 영상 데이터에서 상기 제1 프레임과 대응되는 제2 프레임, 상기 제2 프레임과 연속하는 제3 및 제4 프레임을 식별하는 단계, 상기 제1 프레임과 상기 제2 내지 제4 프레임을 각각 비교하여 유사도를 측정하는 단계를 포함하는 것을 특징으로 한다.

대 표 도 - 도1



(52) CPC특허분류
G06K 9/6201 (2013.01)
(72) 발명자
엠디 아즈헐 우딘
경기도 용인시 기흥구 덕영대로 1732 전자정보대학 349호 (서천동, 경희대학교 국제캠퍼스)
이창주
경기도 수원시 영통구 매영로 432, 301호 (영통동)

이 발명을 지원한 국가연구개발사업
과제고유번호 2016-0-00406(R7120-17-1007)
부처명 과학기술정보통신부
연구관리전문기관 정보통신기술진흥센터
연구사업명 SW컴퓨팅기술개발사업
연구과제명 (기반SW-창조씨앗2단계) SIAT형 CCTV 클라우드 플랫폼 기술 개발
기 여 율 1/1
주관기관 숙명여자대학교 산학협력단
연구기간 2016.06.01 ~ 2021.12.31

명 세 서

청구범위

청구항 1

영상 데이터베이스로부터 제1 영상 데이터와 상기 제1 영상 데이터의 비교 대상이 되는 참고 영상 데이터를 수신하는 단계;

상기 제1 영상 데이터를 복수 개의 제2 영상 데이터로 분산 저장하는 단계;

상기 제2 영상 데이터에서 제1 프레임을 추출하는 전처리 단계;

상기 제1 프레임에서 기 설정된 크기를 갖는 에지 마스크를 검출하는 단계;

상기 에지 마스크의 영역에 대한 특징을 추출하여 제1 특징 벡터를 생성하는 단계;

상기 제1 프레임에서 생성된 하나 이상의 상기 제1 특징 벡터를 결합하여 상기 제1 프레임에 대응되는 제2 특징 벡터를 생성하는 단계;

상기 참고 영상 데이터에서 상기 제1 프레임과 대응되는 제2 프레임, 상기 제2 프레임과 인접한 제3 및 제4 프레임을 식별하는 단계;

상기 제1 프레임과 상기 제2 내지 제4 프레임을 각각 비교하여 유사도를 측정하고, 상기 유사도의 최대 값을 최종 유사도로 하는 단계를 포함하는 유사도 예측 방법.

청구항 2

제1항에 있어서,

상기 전처리 단계는,

상기 제2 영상 데이터에서 제1 프레임을 추출하는 단계;

상기 제1 프레임을 그레이 스케일로 변환하는 단계;

상기 제1 프레임의 크기를 변경하는 단계;

상기 제1 프레임에서 전경을 추출하는 단계를 포함하는 유사도 예측 방법.

청구항 3

제2항에 있어서,

상기 제1 프레임에서 전경을 추출하기 위하여 가우시안 정규 분포를 이용하는 것을 특징으로 하는 유사도 예측 방법.

청구항 4

제1항에 있어서,

상기 에지 마스크는,

상기 에지 마스크의 중심에 위치한 중심 픽셀;

상기 중심 픽셀에 인접한 복수 개의 이웃 픽셀을 포함하는 것을 특징으로 하는 유사도 예측 방법.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서,

상기 제1 특징 벡터를 생성하는 단계는,

상기 에지 마스크의 상기 중심 픽셀과 상기 이웃 픽셀 사이의 차이 값을 연산하는 단계;

상기 차이 값의 평균값을 연산하는 단계;

상기 차이 값과 상기 평균값을 비교하여, 상기 차이 값이 상기 평균값보다 크면 1의 값을, 그렇지 않으면 0의 값을 부여하여 제1 특징 벡터를 생성하는 단계를 포함하는 것을 특징으로 하는 유사도 예측 방법.

청구항 6

제1항에 있어서,

상기 제2 내지 제4 프레임은 서로 연속하는 것을 특징으로 하는 유사도 예측 방법.

청구항 7

제1항에 있어서,

상기 유사도를 측정하는 단계는,

상기 제1 프레임과 상기 제2 내지 제4 프레임에 코사인 유사도를 적용하는 것을 특징으로 하는 유사도 예측 방법.

청구항 8

영상 데이터베이스로부터 제1 영상 데이터와 상기 제1 영상 데이터의 비교 대상이 되는 참고 영상 데이터를 수신하는 영상 수신부;

상기 제1 영상 데이터를 복수 개의 제2 영상 데이터로 분산 저장하는 데이터 분할부;

상기 제2 영상 데이터에서 제1 프레임을 추출하는 전처리부;

상기 제1 프레임에서 기 설정된 크기를 갖는 에지 마스크를 검출하고, 상기 에지 마스크의 영역에 대한 특징을 추출하여 제1 특징 벡터를 생성하며, 상기 제1 프레임에서 생성된 하나 이상의 상기 제1 특징 벡터를 결합하여 상기 제1 프레임에 대응되는 제2 특징 벡터를 생성하는 제어부;

상기 참고 영상 데이터에서 상기 제1 프레임과 대응되는 제2 프레임, 상기 제2 프레임과 인접한 제3 및 제4 프레임을 식별하고, 상기 제1 프레임과 상기 제2 내지 제4 프레임을 각각 비교하여 유사도를 측정하며, 상기 유사도의 최대 값을 최종 유사도로 하는 유사도 측정부를 포함하는 유사도 예측 시스템.

청구항 9

제8항에 있어서,

상기 전처리부는,

상기 제2 영상 데이터에서 제1 프레임을 추출하고, 상기 제1 프레임을 그레이 스케일로 변환하며, 상기 제1 프레임의 크기를 변경하고, 상기 제1 프레임에서 전경을 추출하는 것을 특징으로 하는 유사도 예측 시스템.

청구항 10

제9항에 있어서,

상기 제1 프레임에서 전경을 추출하기 위하여 가우시안 정규 분포를 이용하는 것을 특징으로 하는 유사도 예측 시스템.

청구항 11

제8항에 있어서,

상기 에지 마스크는 상기 에지 마스크의 중심에 위치한 중심 픽셀, 상기 중심 픽셀에 인접한 복수 개의 이웃 픽셀을 포함하는 것을 특징으로 하는 유사도 예측 시스템.

청구항 12

제8항 내지 제11항 중 어느 한 항에 있어서,

상기 제어부는,

상기 에지 마스크의 상기 중심 픽셀과 상기 이웃 픽셀 사이의 차이 값을 연산하고, 상기 차이 값의 평균값을 연산하며, 상기 차이 값과 상기 평균값을 비교하여, 상기 차이 값이 상기 평균값보다 크면 1의 값을, 그렇지 않으면 0의 값을 부여하여 제1 특징 벡터를 생성하는 제1 특징 벡터 생성부를 포함하는 것을 특징으로 하는 유사도 예측 시스템.

청구항 13

제8항에 있어서,

상기 제2 내지 제4 프레임은 서로 연속하는 것을 특징으로 하는 유사도 예측 시스템.

청구항 14

제8항에 있어서,

상기 유사도 측정부는,

상기 제1 프레임과 상기 제2 내지 제4 프레임에 코사인 유사도를 적용하는 것을 특징으로 하는 유사도 예측 시스템.

발명의 설명

기 술 분 야

[0001] 본 발명은 영상의 유사도를 예측하는 방법 및 장치에 관한 것으로, 보다 자세하게는 분산 처리 시스템을 이용하여 영상에서 특징 벡터를 추출하여 유사도를 보다 신속하게 예측하는 방법 및 장치에 관한 것이다.

배 경 기 술

[0002] 최근 인터넷 기술의 발전 및 스마트 기기의 급속한 보급에 따라 방대한 양의 영상에 대한 접근성이 낮아지고 있다. 영상에서 객체를 검출하고 인식하는 기술은 영상감지 시스템에 꾸준하게 적용되어 왔으며, 컴퓨터 비전 및 기계 이해 분야의 근본적인 문제로 떠오르고 있다. 일반적으로 영상은 다양한 방식으로 구성된 표면 및 객체로 구성되어 있기 때문에, 영상에서 객체만을 분리하는 것은 지도 제작, 로봇 네비게이션, 장소 추천 등의 다양한 응용 프로그램에 있어서 주요한 이슈가 될 수 있다.

[0003] 영상에서 특징 벡터를 검출 및 인식하기 위해 다양한 특징 추출 방법이 사용되고 있으나 대부분은 이미지를 이용하여 객체를 인식하고 일부만이 영상에 적용 가능하다는 한계가 있다. 객체를 검출하는 기존의 방법 중 대표적인 것이 LBP(Local Binary Pattern)이다. LBP의 견고성, 차별성 및 적용 가능성을 향상시키기 위해 다양한 LBP 변형 방법이 연구되었으나 조명, 시점 등 픽셀의 변화에 따라 바이너리 코드가 완전히 달라질 수 있어 일관된 패턴을 생성하지 못하기 때문에 노이즈에 취약하다는 치명적인 단점이 있다.

선행기술문헌

특허문헌

[0005] (특허문헌 0001) 대한민국 등록특허 제10-1716646호, 공개일자 2014년 7월 18일

발명의 내용

해결하려는 과제

[0006] 본 발명은 전술한 문제점을 해결하기 위한 것으로서, 영상 데이터와 참조 영상 데이터의 유사도를 예측하기 위하여 영상 데이터의 특징 벡터를 추출하는 것을 일 목적으로 한다.

[0007] 또한 본 발명은 영상 데이터에서 유사도를 예측함에 있어서 분산 처리 시스템을 이용하여 연산의 속도를 향상시키는 것을 일 목적으로 한다.

[0008] 또한 본 발명은 영상 데이터에서 추출한 제1 프레임과 유사도를 연산할 기준 프레임으로 참조 영상 데이터에 포함된 복수 개의 프레임을 이용함으로써 유사도 예측의 정확성을 향상시키는 것을 일 목적으로 한다.

[0009] 또한 본 발명은 영상 데이터에서 추출한 제1 프레임의 특징 벡터를 추출함에 있어서 에지 마스크를 이용함으로써 특징 벡터 추출의 정확성을 향상시키는 것을 일 목적으로 한다.

과제의 해결 수단

[0010] 이러한 목적을 달성하기 위한 본 발명은 영상 데이터베이스로부터 제1 영상 데이터와 상기 제1 영상 데이터의 비교 대상이 되는 참고 영상 데이터를 수신하는 단계, 상기 제1 영상 데이터를 복수 개의 제2 영상 데이터로 분산 저장하는 단계, 상기 제2 영상 데이터에서 제1 프레임을 추출하는 전처리 단계, 상기 제1 프레임에서 기 설정된 크기를 갖는 에지 마스크를 검출하는 단계, 상기 에지 마스크의 영역에 대한 특징을 추출하여 제1 특징 벡터를 생성하는 단계, 상기 제1 프레임에서 생성된 하나 이상의 상기 제1 특징 벡터를 결합하여 상기 제1 프레임에 대응되는 제2 특징 벡터를 생성하는 단계, 상기 참고 영상 데이터에서 상기 제1 프레임과 대응되는 제2 프레임, 상기 제2 프레임과 인접한 제3 및 제4 프레임을 식별하는 단계, 상기 제1 프레임과 상기 제2 내지 제4 프레임을 각각 비교하여 유사도를 측정하는 단계를 포함하는 것을 일 특징으로 한다.

[0011] 또한 상기 전처리 단계는, 상기 제2 영상 데이터에서 제1 프레임을 추출하는 단계, 상기 제1 프레임을 그레이스케일로 변환하는 단계, 상기 제1 프레임의 크기를 변경하는 단계, 상기 제1 프레임에서 전경을 추출하는 단계를 포함하는 것을 일 특징으로 한다.

[0012] 나아가 상기 제1 프레임에서 전경을 추출하기 위하여 가우시안 정규 분포를 이용하는 것을 일 특징으로 한다.

[0013] 또한 상기 에지 마스크는, 상기 에지 마스크의 중심에 위치한 중심 픽셀, 상기 중심 픽셀에 인접한 복수 개의 이웃 픽셀을 포함하는 것을 일 특징으로 한다.

[0014] 나아가 상기 제1 특징 벡터를 생성하는 단계는, 상기 에지 마스크의 상기 중심 픽셀과 상기 이웃 픽셀 사이의 차이 값을 연산하는 단계, 상기 차이 값의 평균값을 연산하는 단계, 상기 차이 값과 상기 평균값을 비교하여, 상기 차이 값이 상기 평균값보다 크면 1의 값을, 그렇지 않으면 0의 값을 부여하여 제1 특징 벡터를 생성하는 단계를 포함하는 것을 일 특징으로 한다.

[0015] 또한 상기 제2 내지 제4 프레임은 서로 연속하는 것을 일 특징으로 한다.

[0016] 나아가 상기 유사도를 측정하는 단계는, 상기 제1 프레임과 상기 제2 내지 제4 프레임에 코사인 유사도를 적용

하는 것을 일 특징으로 한다.

[0017] 또한 본 발명은 영상 데이터베이스로부터 제1 영상 데이터와 상기 제1 영상 데이터의 비교 대상이 되는 참고 영상 데이터를 수신하는 영상 수신부, 상기 제1 영상 데이터를 복수 개의 제2 영상 데이터로 분산 저장하는 데이터 분할부, 상기 제2 영상 데이터에서 제1 프레임을 추출하는 전처리부, 상기 제1 프레임에서 기 설정된 크기를 갖는 에지 마스크를 검출하고, 상기 에지 마스크의 영역에 대한 특징을 추출하여 제1 특징 벡터를 생성하며, 상기 제1 프레임에서 생성된 하나 이상의 상기 제1 특징 벡터를 결합하여 상기 제1 프레임에 대응되는 제2 특징 벡터를 생성하는 제어부, 상기 참고 영상 데이터에서 상기 제1 프레임과 대응되는 제2 프레임, 상기 제2 프레임과 인접한 제3 및 제4 프레임을 식별하여 상기 제1 프레임과 상기 제2 내지 제4 프레임을 각각 비교하여 유사도를 측정하는 유사도 측정부를 포함하는 것을 일 특징으로 한다.

[0018] 나아가 상기 전처리부는, 상기 제2 영상 데이터에서 제1 프레임을 추출하고, 상기 제1 프레임을 그레이 스케일로 변환하며, 상기 제1 프레임의 크기를 변경하고, 상기 제1 프레임에서 전경을 추출하는 것을 일 특징으로 한다.

[0019] 또한 상기 제1 프레임에서 전경을 추출하기 위하여 가우시안 정규 분포를 이용하는 것을 일 특징으로 한다.

[0020] 나아가 상기 에지 마스크는 상기 에지 마스크의 중심에 위치한 중심 픽셀, 상기 중심 픽셀에 인접한 복수 개의 이웃 픽셀을 포함하는 것을 일 특징으로 한다.

[0021] 또한 상기 제어부는, 상기 에지 마스크의 상기 중심 픽셀과 상기 이웃 픽셀 사이의 차이 값을 연산하고, 상기 차이 값의 평균값을 연산하며, 상기 차이 값과 상기 평균값을 비교하여, 상기 차이 값이 상기 평균값보다 크면 1의 값을, 그렇지 않으면 0의 값을 부여하여 제1 특징 벡터를 생성하는 제1 특징 벡터 생성부를 포함하는 것을 일 특징으로 한다.

[0022] 나아가 상기 제2 내지 제4 프레임은 서로 연속하는 것을 일 특징으로 한다.

[0023] 또한 상기 유사도 측정부는, 상기 제1 프레임과 상기 제2 내지 제4 프레임에 코사인 유사도를 적용하는 것을 일 특징으로 한다.

발명의 효과

[0024] 전술한 바와 같은 본 발명에 의하면, 영상 데이터와 참조 영상 데이터의 유사도를 예측하기 위하여 영상 데이터의 특징 벡터를 추출할 수 있다.

[0025] 또한 본 발명은 영상 데이터에서 유사도를 예측함에 있어서 분산 처리 시스템을 이용하여 연산의 속도를 향상시킬 수 있다.

[0026] 또한 본 발명은 영상 데이터에서 추출한 제1 프레임과 유사도를 연산할 기준 프레임으로 참조 영상 데이터에 포함된 복수 개의 프레임을 이용함으로써 유사도 예측의 정확성을 향상시킬 수 있다.

[0027] 또한 본 발명은 영상 데이터에서 추출한 제1 프레임의 특징 벡터를 추출함에 있어서 에지 마스크를 이용함으로써 특징 벡터 추출의 정확성을 향상시킬 수 있다.

도면의 간단한 설명

[0028] 도 1은 본 발명의 일 실시 예에 의한 유사도 예측 방법의 대략적인 동작을 나타낸 도면이다.

도 2는 본 발명의 일 실시 예에 의한 유사도 예측 장치의 구성을 도시한 도면이다.

도 3은 본 발명의 일 실시 예에 의한 유사도 예측 방법을 설명하기 위한 도면이다.

도 4는 본 발명의 일 실시 예에 의한 제1 특징 벡터를 추출하는 방법을 도시한 도면이다.

도 5는 본 발명의 일 실시 예에 의한 프레임 비교 방법을 도시한 도면이다.

도 6은 본 발명의 일 실시 예에 의한 비교 프레임 수에 따른 연산 속도를 나타낸 그래프이다.

도 7은 UCF Youtube 데이터 세트의 카테고리 별 샘플 영상 데이터를 나타낸 것이다.

도 8은 유사도 예측 장치를 이용하여 UCF Youtube 데이터 세트에 대하여 측정된 유사도의 정확도를 나타낸 그래프이다.

도 9는 본 발명의 일 실시 예에 의한 유사도 예측 장치에서 사용되는 특징 추출 알고리즘과 기존 특징 추출 알고리즘을 비교한 결과를 나타낸 그래프이다.

도 10은 본 발명의 일 실시 예에 의한 Spark 클러스터의 노드의 수에 따른 소요 시간을 나타낸 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0029] 전술한 목적, 특징 및 장점은 첨부된 도면을 참조하여 상세하게 후술되며, 이에 따라 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 것이다. 본 발명을 설명함에 있어서 본 발명과 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 상세한 설명을 생략한다.

[0030] 도면에서 동일한 참조부호는 동일 또는 유사한 구성요소를 가리키는 것으로 사용되며, 명세서 및 특허청구의 범위에 기재된 모든 조합은 임의의 방식으로 조합될 수 있다. 그리고 다른 식으로 규정하지 않는 한, 단수에 대한 언급은 하나 이상을 포함할 수 있고, 단수 표현에 대한 언급은 또한 복수 표현을 포함할 수 있음이 이해되어야 한다.

[0031] 본 명세서에서 사용되는 용어는 단지 특정 예시적 실시 예들을 설명할 목적을 가지고 있으며 한정할 의도로 사용되는 것이 아니다. 본 명세서에서 사용된 바와 같은 단수적 표현들은 또한, 해당 문장에서 명확하게 달리 표시하지 않는 한, 복수의 의미를 포함하도록 의도될 수 있다. 용어 "및/또는," "그리고/또는"은 그 관련되어 나열되는 항목들의 모든 조합들 및 어느 하나를 포함한다. 용어 "포함한다", "포함하는", "포함하고 있는", "구비하는", "갖는", "가지고 있는" 등은 내포적 의미를 갖는 바, 이에 따라 이러한 용어들은 그 기재된 특징, 정수, 단계, 동작, 요소, 및/또는 컴포넌트를 특정하며, 하나 이상의 다른 특징, 정수, 단계, 동작, 요소, 컴포넌트, 및/또는 이들의 그룹의 존재 혹은 추가를 배제하지 않는다. 본 명세서에서 설명되는 방법의 단계들, 프로세스들, 동작들은, 구체적으로 그 수행 순서가 확정되는 경우가 아니라면, 이들의 수행을 논의된 혹은 예시된 그러한 특정 순서로 반드시 해야 하는 것으로 해석돼서는 안 된다. 추가적인 혹은 대안적인 단계들이 사용될 수 있음을 또한 이해해야 한다.

[0032] 또한, 각각의 구성요소는 각각 하드웨어 프로세서로 구현될 수 있고, 위 구성요소들이 통합되어 하나의 하드웨어 프로세서로 구현될 수 있으며, 또는 위 구성요소들이 서로 조합되어 복수 개의 하드웨어 프로세서로 구현될 수도 있다.

[0033] 이하, 첨부된 도면을 참조하여 본 발명에 따른 바람직한 실시 예를 상세히 설명하기로 한다.

[0034] 본 발명은 영상의 유사도를 평가하기 위하여 영상 데이터의 특징 벡터를 추출할 수 있다. 영상에서 특징 벡터를 추출하기 위한 기존의 방법 중 대표적인 것이 LBP(Local Binary Pattern)이다. LBP는 이미지의 텍스처를 분류하기 위하여 개발된 알고리즘으로, 객체 인식과 같은 다양한 영상 인식 분야에도 활용된다. LBP는 영상에 포함된 모든 픽셀에 대해 연산되는 값으로, 각 픽셀의 주변 3x3 크기를 갖는 영역에 대한 상대적인 밝기 변화를 2진수로 나타내는 인덱스 값을 의미한다. 즉, LBP는 지역적인 이진 패턴을 연산한다.

[0035] LBP의 구체적인 원리는 3x3 영역 내에서 중심에 위치하는 중심 픽셀, 중심 픽셀과 이웃하는 8개의 이웃 픽셀끼리 서로 크기를 비교하여 이웃 픽셀의 값이 중심 픽셀보다 크면 1의 값을, 그렇지 않으면 0의 값을 갖도록 하여 이진 값을 연산한다. 예를 들어 중심 픽셀의 값이 50이고 그에 대한 이웃 픽셀의 값이 {65, 90, 40, 125, 35, 15, 70, 5}이면, LBP 이진 값은 {1, 1, 0, 1, 0, 0, 1, 0}이 된다. 따라서 최종 이진 패턴은 11010010(2)가 될 수 있다.

[0036] 이와 같이 LBP는 영상에 포함된 픽셀에 대한 상대적인 값을 식별하기 때문에 밝기와 상관없이 일정한 값을 가질 수 있다. 그러나 LBP는 이웃 노드의 강도에 따라 이진 값이 완전히 달라질 수 있어 일관된 이진 패턴을 추출하지 못하기 때문에 노이즈에 취약하다는 단점이 있다.

[0037] 도 1은 본 발명의 일 실시 예에 의한 유사도 예측 방법의 대략적인 동작을 나타낸 도면이다. 도 1을 참조하면, 유사도 예측 방법은 영상 데이터를 저장하기 위하여 HDFS를 사용할 수 있다. HDFS(Hadoop Distribution File System)은 Hadoop의 데이터 저장 시스템을 의미한다. Hadoop은 대용량 데이터를 분산 처리할 수 있는 자바 기반의 오픈 소스 프레임워크이다. Hadoop은 복수 개의 서버에 데이터를 분산 저장하고, 데이터가 저장된 서버에서 동시에 데이터를 처리하는 시스템이다.

[0038] 영상 데이터에 포함된 특징을 추출하기 위하여 Spark RDD는 HDFS에 저장된 영상 데이터를 불러올 수 있다.

Spark RDD는 Spark 내의 저장소를 의미한다. Spark는 범용성 분산 플랫폼으로, 분산된 복수 개의 노드에서 연산을 수행하게 하는 범용 분산 클러스터링 플랫폼을 의미한다. Spark는 Hadoop과 유사한데, Hadoop이 Map&Reduce 작업을 디스크 기반으로 수행하기 때문에 성능의 감소가 발생하는 반면, Spark는 Map&Reduce 작업을 메모리 기반으로 수행하기 때문에 Hadoop보다 빠른 속도를 가질 수 있다.

[0039] 본 발명은 Spark RDD가 HDFS로부터 제1 영상 데이터를 불러오면, 제1 영상 데이터를 복수 개의 제2 영상 데이터로 분할하여 Spark에서 처리 가능한 노드에 분산 저장할 수 있다. 제2 영상 데이터가 저장된 각 노드에서, 제2 영상 데이터를 전처리하고, 전처리된 영상 데이터로부터 특징 벡터를 추출한 후, 추출된 특징 벡터를 HDFS에 저장할 수 있다. HDFS에 저장된 특징 벡터와 영상 데이터와 비교 대상이 되는 기 저장된 영상 데이터의 특징 벡터를 비교하여 영상 데이터의 객체를 식별할 수 있다.

[0040] 이하에서는 도 2를 참조하여, 본 발명의 일 실시 예에 의한 유사도 예측 장치를 설명한다. 도 2를 참조하면, 본 발명의 일 실시 예에 의한 유사도 예측 장치는 영상 수신부(100), 전처리부(200), 제어부(300), 데이터 분할부(400), 유사도 측정부(500), 저장부(600)를 포함할 수 있다. 제어부(300)는 에지 마스크 생성부(310), 특징 벡터 생성부(330)를 더 포함할 수 있고, 저장부(600)는 영상 데이터 저장부(미도시)와 특징 벡터 저장부(미도시)를 더 포함할 수 있다.

[0041] 영상 수신부(100)는 HDFS에서 제1 영상 데이터를 수신할 수 있다. 데이터 분할부(400)는 영상 수신부(100)로부터 수신된 제1 영상 데이터를 복수 개의 제2 영상 데이터로 분할하여 영상 데이터 저장부(600)에 분산 저장할 수 있다. 영상 데이터 저장부(600)는 Spark RDD를 따르며, 복수 개의 노드로 구성되어 있어 제2 영상 데이터를 각 노드에 분산시켜 저장함으로써 제2 영상 데이터 각각을 동시에 처리할 수 있도록 한다. 이하에서는 각 노드에서 제2 영상 데이터를 처리하기 위한 모듈을 설명한다.

[0042] 전처리부(200)는 영상 데이터 저장부(600)에 분산 저장된 제2 영상 데이터의 전처리를 수행할 수 있다. 보다 구체적으로 전처리부(200)는 제2 영상 데이터에서 하나 이상의 제1 프레임을 추출할 수 있다. 제1 프레임은 RGB 스케일이기 때문에, 전처리부(200)는 제1 프레임을 그레이 스케일로 변환한 후, 제1 프레임의 크기를 기 설정된 크기로 재설정할 수 있다. 이 때, 제1 프레임이 재설정 되는 크기는 720 * 404일 수 있으나 이 외의 다른 크기로 설정될 수 있다.

[0043] 나아가 전처리부(200)는 제1 프레임에서 전경을 추출하기 위해 배경을 제거할 수 있다. 전처리부(200)는 제1 프레임에서 전경을 추출하기 위해 가우시안 정규 분포를 이용할 수 있다. 가우시안 정규 분포는 가장 널리 사용되는 연속 확률 분포로, 일반적으로 정규 분포로 알려져 있다. 가우시안 정규 분포는 수집된 데이터에 대한 분포를 근사하는 것으로, 수학식 1을 통해 연산된다.

수학식 1

[0044]
$$P(R(x,y)) = \frac{1}{C}\sum_{i=1}^{C}\frac{1}{\sqrt{2\pi\sigma^2}}exp(\frac{-(R(x,y)-B_i(x,y))^2}{2\sigma^2})$$

[0045] 수학식 1에서, C는 배경 이미지의 수를, $R(x,y)$는 영상 데이터의 (x, y) 픽셀의 강도 값을, $B_i(x,y)$는 영상 데이터에서 추출된 배경의 (x, y) 픽셀의 강도 값을 의미할 수 있다.

[0046] 제어부(300)는 전처리가 수행된 제1 프레임에서 기 설정된 크기를 갖는 에지 마스크를 검출하고, 에지 마스크를 이용하여 제1 프레임에 대응되는 특징 벡터를 생성할 수 있다. 보다 구체적으로 제어부(300)는 에지 검출부(310)와 특징 벡터 생성부(330)를 포함할 수 있다.

[0047] 에지 검출부(310)는 제1 프레임에서 에지 마스크를 검출하기 위하여 소벨 연산자(Sobel Operator)를 이용할 수 있다. 소벨 연산자는 3 * 3 크기의 행렬에서 중심 값을 기준으로 이웃 값을 비교하여 픽셀의 변화량을 검출하는 1차 미분 알고리즘이다. 에지 검출부(310)는 소벨 연산자를 이용하여 제1 프레임에 포함된 픽셀 중 밝기의 변화가 급격한 부분을 나타내는 에지를 검출할 수 있다.

[0048] 영상 데이터에서 추출한 제1 프레임은 2차원의 속성을 갖기 때문에 제1 프레임의 가로와 세로 방향에 대해 각각 소벨 연산자를 적용할 수 있다. 에지 검출부(310)는 수학식 2를 통해 제1 프레임에서 에지 마스크를 검출할 수 있다.

수학식 2

$$\nabla f = [\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}]^T$$

$$|\nabla f| = mag(\nabla f) = [(\frac{\partial f}{\partial x})^2 + (\frac{\partial f}{\partial y})^2]^{\frac{1}{2}}$$

$$\nabla f \approx |G_x| + |G_y|$$

$$|G| = \sqrt{Gx^2 + Gy^2}$$

[0049]

[0050] 에지 검출부(310)가 생성한 에지 마스크는 에지 마스크의 중심에 위치한 중심 픽셀과 중심 픽셀에 인접한 복수개의 이웃 픽셀을 포함할 수 있다.

[0051] 특징 벡터 생성부(330)는 에지 검출부(310)에서 추출된 에지 마스크를 이용하여 제1 프레임에 대한 특징 벡터를 생성할 수 있다. 특징 벡터 생성부(330)는 제1 프레임에서 기 설정된 영역 및 크기에 대한 에지 마스크에 대한 제1 특징 벡터를 생성할 수 있다.

[0052] 특징 벡터 생성부(330)는 에지 마스크의 중심 픽셀과 이웃 픽셀의 차이 값을 연산할 수 있다. 에지 마스크가 중심 픽셀 a와 이웃 픽셀 {a1, a2, a3, a4, a5, a6, a7, a8}을 포함할 때, 특징 벡터 생성부(330)는 중심 픽셀과 이웃 픽셀의 차이 값을 {a1-a, a2-a, a3-a, a4-a, a5-a, a6-a, a7-a, a8-a}와 같이 연산할 수 있다. 이 때 a1에 대응되는 이웃 픽셀은 사용자의 설정에 따라 달라질 수 있으나 본 발명의 설명에서는 에지 마스크의 (1, 1) 위치에 대응되는 이웃 픽셀로 설명한다.

[0053] 도 4를 참조하여 유사도 예측 방법에 대한 일 실시 예를 설명하면, 에지 마스크에 포함된 중심 픽셀의 값이 32이고, 중심 픽셀과 이웃한 이웃 픽셀의 값이 {80, 85, 95, 31, 51, 23, 22, 35}이면, 특징 벡터 생성부(330)는 중심 픽셀과 이웃 픽셀의 차이 값을 {48, 53, 63, -1, 19, -9, -10, 3}과 같이 연산할 수 있다.

[0054] 특징 벡터 생성부(330)는 이웃 픽셀의 차이 값에 대한 평균 값을 연산할 수 있다. 보다 구체적으로 특징 벡터 생성부(330)는 이웃 픽셀의 차이 값의 절대 값에 대한 평균 값을 연산할 수 있다. 위 예시에서, 특징 벡터 생성부(330)는 이웃 픽셀의 차이 값의 절대 값인 {48, 53, 63, 1, 19, 9, 10, 3}에 대한 평균 값을 연산할 수 있다. 이 때 연산되는 평균 값은 25.75로, 특징 벡터 생성부는 연산된 평균 값을 내림하여 25의 값을 얻을 수 있다.

[0055] 특징 벡터 생성부(300)는 연산된 평균 값과 이웃 픽셀의 차이 값을 비교하여 제1 특징 벡터를 생성할 수 있다. 특징 벡터 생성부(330)는 이웃 픽셀의 차이 값이 평균 값보다 크면 1의 값을, 그렇지 않으면 0의 값을 부여하여 제1 특징 벡터를 생성할 수 있다.

[0056] 위 예시에서, 특징 벡터 생성부(330)는 평균 값인 25와 이웃 픽셀의 차이 값인 {48, 53, 63, -1, 19, -9, -10, 3}를 비교하여 제1 특징 벡터를 생성할 수 있다. 특징 벡터 생성부(330)는 이웃 픽셀의 차이 값과 평균 값을 비교하여 {1, 1, 1, 0, 0, 0, 0, 0}의 값을 갖는 제1 특징 벡터를 생성할 수 있다. 특징 벡터 생성부(330)가 생성한 제1 특징 벡터는 a1에 대응되는 이웃 픽셀에 따라 상이할 수 있다.

[0057] 특징 벡터 생성부(330)는 제1 프레임의 일 영역에 대하여 추출된 하나 이상의 제1 특징 벡터를 이용하여 제1 프레임의 전체 영역에 대응되는 제2 특징 벡터를 생성할 수 있다. 다시 말해서 특징 벡터 생성부(330)는 제1 프레임에 대해 추출된 하나 이상의 제1 특징 벡터를 병합하여 제1 프레임 전체에 대응되는 제2 특징 벡터를 생성할 수 있다.

[0058] 유사도 측정부(500)는 제어부(300)에서 생성된 제2 특징 벡터를 이용하여 제2 영상 데이터의 유사도를 측정할 수 있다. 유사도 측정부(500)는 제2 영상 데이터의 유사도를 측정하기 위하여 기준이 되는 참고 영상 데이터를 이용할 수 있다. 유사도 측정부(500)는 참고 영상 데이터에 포함된 하나 이상의 프레임에서 제2 영상 데이터에 포함된 하나 이상의 제1 프레임에 대응되는 제2 프레임을 식별할 수 있다. 나아가 유사도 측정부(500)는 참고 영상 데이터에서 제2 프레임의 이웃 프레임인 제3 및 제4 프레임을 식별할 수 있다. 이 때, 제2 프레임, 제3 프레임, 그리고 제4 프레임은 서로 연속한 프레임일 수 있다. 제1 프레임과 제2 내지 제4 프레임을 비교함으로써

영상 데이터와 참고 영상 데이터의 속도가 상이할 경우 제1 프레임과 제2 프레임의 유사도가 낮게 평가되는 것을 방지할 수 있다.

[0059] 참고 영상 데이터의 제2 프레임이 참고 영상 데이터의 첫 프레임일 경우, 제3 프레임은 제2 프레임의 다음 프레임이고, 제4 프레임은 제3 프레임의 다음 프레임일 수 있다. 나아가 참고 영상 데이터의 제2 프레임이 참고 영상 데이터의 중간에 위치한 프레임일 경우, 제3 프레임은 제2 프레임의 이전 프레임이고, 제4 프레임은 제2 프레임의 이후 프레임일 수 있다. 또한 참고 영상 데이터의 제2 프레임이 참고 영상 데이터의 마지막 프레임일 경우, 제3 프레임은 제2 프레임의 이전 프레임이고, 제4 프레임은 제3 프레임의 이전 프레임일 수 있다.

[0060] 도 5를 참조하면, 유사도 측정부(500)는 영상 데이터의 제1 프레임과 참고 영상 데이터의 제2 프레임 내지 제4 프레임 각각에 대한 유사도를 연산할 수 있다. 유사도 측정부(500)는 유사도를 측정하기 위해 코사인 유사도를 이용할 수 있다. 코사인 유사도(Cosine Similarity)는 두 벡터 사이의 각도에 대한 코사인 값을 이용하여 측정된 벡터 간의 유사도의 정도를 의미한다. 유사도 측정부(500)는 제1 프레임의 제2 특징 벡터와 제2 내지 제4 프레임 각각에 대한 특징 벡터의 유사도를 측정하기 위하여 수학식 3과 같은 코사인 유사도를 이용할 수 있다.

수학식 3

[0061] $$\cos\theta = \frac{F_q \cdot F_d}{\|F_q\| \cdot \|F_d\|}$$

[0062] 유사도 측정부(500)는 제1 프레임의 제2 특징 벡터와 제2 내지 제4 프레임 각각에 대한 특징 벡터의 유사도 중 최대 값을 최종 유사도로 할 수 있다.

[0063] 부가적으로 설명하면 유사도 측정부(500)는 제1 프레임의 비교 대상 프레임으로 제2 내지 제4 프레임, 즉 3개의 프레임을 식별하는데, 이는 비교 프레임 수가 3개일 때 연산 속도가 가장 최적화되어 있기 때문이다. 이는 도 6을 통해 확인할 수 있다.

[0064] 이하에서는 도 3을 이용하여 본 발명의 일 실시 예에 의한 유사도 예측 방법을 설명한다. 유사도 예측 방법에 관한 설명에 있어서 전술한 유사도 예측 시스템과 중복되는 세부 실시 예는 생략될 수 있다.

[0065] 도 3은 본 발명의 일 실시 예에 의한 유사도 예측 방법을 설명하기 위한 도면이다. 유사도 예측 방법을 수행하는 유사도 예측 장치는 서버로 구현될 수 있는 바, 이하에서는 설명의 편의를 위해 서버로 명명한다.

[0066] 도 3을 참조하면, 서버는 HDFS에서 제1 영상 데이터와 제1 영상 데이터의 비교 데이터인 참조 영상 데이터를 불러올 수 있다(S100). 서버는 불러온 제1 영상 데이터를 복수 개의 제2 영상 데이터로 분할(S200)하여 Spark RDD의 각 노드에 분산 저장할 수 있다. Spark RDD의 노드는 Mapper와 Reducer를 각각 포함하고 있다. 이하에서는 설명의 편의성을 위하여 제2 영상 데이터가 분할된 노드를 제1 노드라고 명명한다.

[0067] 서버는 제1 노드 각각에 분산 저장된 제2 영상 데이터에서 제1 프레임을 추출하고, 제1 프레임을 그레이 스케일로 변환하며, 변환된 제1 프레임의 크기를 변경하고, 제1 프레임의 전경을 추출하는 전처리를 수행할 수 있다(S300).

[0068] 서버는 전처리된 제2 영상 데이터에서 추출한 제1 프레임에서, 기 설정된 위치 및 크기를 갖는 에지 마스크를 검출(S400)할 수 있다. 서버는 에지 마스크를 검출하기 위해 소벨 연산자를 사용할 수 있다.

[0069] 서버는 검출된 에지 마스크를 이용하여 제1 특징 벡터를 생성(S500)할 수 있다. 서버는 에지 마스크에 포함된 중간 픽셀과 중간 픽셀과 인접한 이웃 픽셀의 값을 이용하여 제1 특징 벡터를 생성할 수 있다. 보다 구체적으로 서버는 중간 픽셀의 값과 이웃 픽셀의 값에 대한 차이 값을 연산하고, 연산된 차이 값의 절대 값에 대한 평균 값을 연산할 수 있다. 서버는 연산된 평균 값과 중간 픽셀의 값과 이웃 픽셀의 값에 대한 차이 값을 비교하여 제1 특징 벡터를 생성할 수 있다.

[0070] 서버는 제1 프레임에서 추출된 하나 이상의 제1 특징 벡터를 이용하여 제1 프레임에 대응되는 제2 특징 벡터를 생성할 수 있다(S600). 제1 특징 벡터는 제1 프레임에 포함된 특정 영역에 대한 벡터 값이고, 제2 특징 벡터는 제1 프레임에 대한 벡터 값이다.

[0071] 서버는 제1 프레임의 제2 특징 벡터의 유사도를 연산하기 위하여 제1 영상 데이터의 비교 대상인 참조 영상 데

이터에 포함된 하나 이상의 프레임에서 제1 프레임과 대응되는 제2 프레임을 식별할 수 있다. 나아가 서버는 참조 영상 데이터에서 제2 프레임과 인접한 제3 및 제4 프레임을 식별할 수 있다. 참고 영상 데이터의 제2 프레임이 참고 영상 데이터의 첫 프레임일 경우, 제3 프레임은 제2 프레임의 다음 프레임이고, 제4 프레임은 제3 프레임의 다음 프레임일 수 있다. 나아가 참고 영상 데이터의 제2 프레임이 참고 영상 데이터의 중간에 위치한 프레임일 경우, 제3 프레임은 제2 프레임의 이전 프레임이고, 제4 프레임은 제2 프레임의 이후 프레임일 수 있다. 또한 참고 영상 데이터의 제2 프레임이 참고 영상 데이터의 마지막 프레임일 경우, 제3 프레임은 제2 프레임의 이전 프레임이고, 제4 프레임은 제3 프레임의 이전 프레임일 수 있다.

[0072] 서버는 제1 프레임의 제2 특징 벡터와 제2 내지 제4 프레임 각각의 특징 벡터의 유사도를 연산한 후, 최대 값을 제1 프레임에 대한 유사도로 설정할 수 있다(S700).

[0073] 이하에서는 도 7 내지 도 10을 이용하여 본 발명의 일 실시 예에 의한 유사도 예측 장치의 성능을 기존 방법과 비교하기 위한 실험 및 그 결과를 설명한다.

[0074] 유사도 예측 장치의 실험에는 Hadoop, Spark 클러스터를 사용할 수 있다. Spark 클러스터는 3개의 노드를 포함할 수 있으며, 각 노드는 3.0GHz, 32GB의 성능을 갖는 4개의 코어를 가질 수 있다. 또한 실험에 사용되는 Hadoop은 2.7.1 버전, Spark는 1.6.2버전이다. 유사도 예측 장치의 실험에 있어서 사용되는 데이터는 UCF Youtube 데이터 세트로, 11개의 카테고리로 구분될 수 있다. UCF 데이터 세트는 농구 슈팅(basketball shooting), 자전거(biking), 드라이브(driving), 골프 스윙(golf swinging), 승마(horse-back riding), 축구공 저글링(soccer juggling), 그네(swinging), 테니스 스윙(tennis swinging), 트램플린 점프(trampoline jumping), 발리볼 스파이크(volleyball spiking), 개와의 산책(walking with a dog)와 같은 총 11개의 카테고리를 포함할 수 있다. UCF 데이터 세트는 각 카테고리 별 125개의 영상 데이터, 즉 총 1375개의 영상 데이터를 포함할 수 있으며 도 7은 각 카테고리 별 샘플 영상 데이터를 나타낸 것이다.

[0075] 도 8은 유사도 예측 장치를 이용하여 UCFYoutube 데이터 세트에 대하여 측정된 유사도의 정확도(mAP, mean average precision)를 나타낸 그래프이다. 도 8을 참조하면, 승마 카테고리의 정확도가 0.89로 가장 높은 수치를 나타내고, 자전거 카테고리와 테니스 스윙 카테고리가 각각 0.85, 0.84로 그 뒤를 따르고 있다. 농구 슈팅 카테고리의 정확도는 0.72로 가장 낮은 수치를 나타내고 있다.

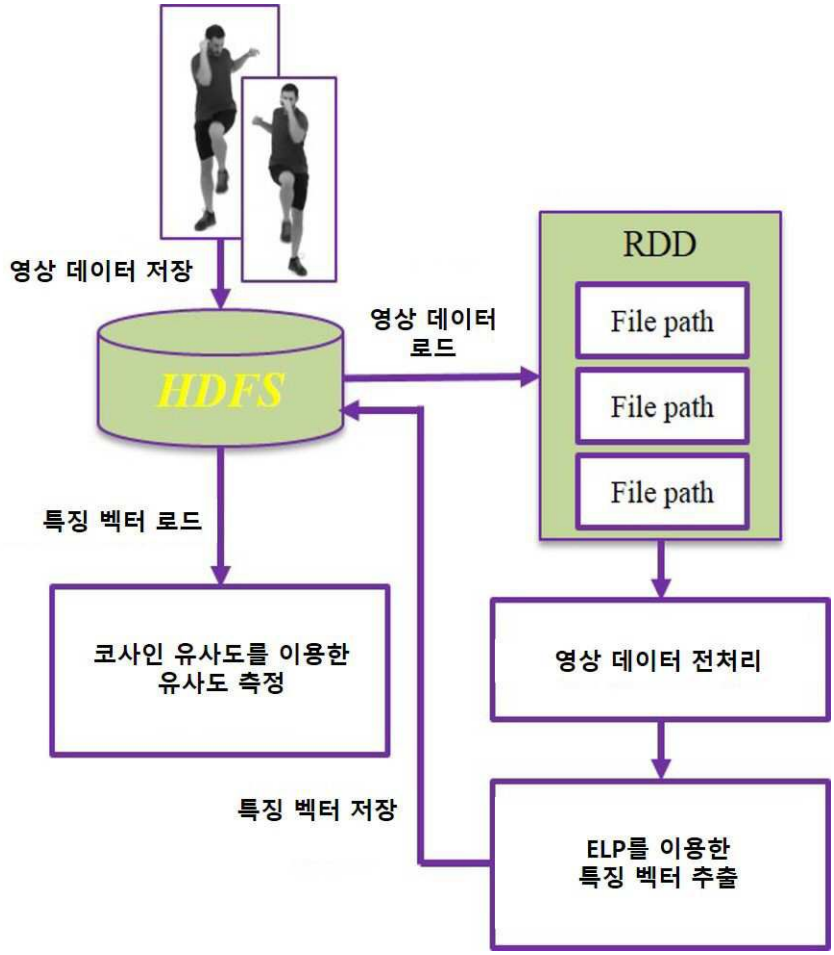
[0076] 도 9는 유사도 예측 장치에서 사용되는 특징 추출 알고리즘과 기존 특징 추출 알고리즘을 비교한 결과를 나타낸 그래프이다. 기존 알고리즘으로 LBP, MBP, LTP, MTP, Improved LBP, GLTP, DRLTP, Gradient LBP, Gradient direction and LBP를 이용할 수 있다. 도 9를 참조하면 유사도 예측 장치에서 사용되는 특징 추출 알고리즘은 0.806의 정확도를 가져 기존 알고리즘과 대비하여 정확도가 높은 것을 확인할 수 있다. 그러나 DRLTP 알고리즘의 경우 에지 정보와 질감 정보를 모두 포함하기 때문에 가장 정확한 결과 값을 가질 수 있다. 질감 정보를 포함하지 않는 알고리즘 중에서는 유사도 예측 장치에서 사용되는 특징 추출 알고리즘의 정확도가 단연 우수하다고 볼 수 있다.

[0077] 도 10은 Spark 클러스터의 노드의 수에 따른 소요 시간을 나타낸 그래프이다. 본 발명은 Spark 클러스터에서 3개의 노드를 사용하는데, 도 10을 참조하면 3개의 노드가 특징을 추출함에 있어서 가장 적은 시간을 소모하는 것을 확인할 수 있다.

[0078] 본 명세서와 도면에 개시된 본 발명의 실시 예들은 본 발명의 기술 내용을 쉽게 설명하고 본 발명의 이해를 돕기 위해 특정 예를 제시한 것뿐이며, 본 발명의 범위를 한정하고자 하는 것은 아니다. 여기에 개시된 실시 예들 이외에도 본 발명의 기술적 사상에 바탕을 둔 다른 변형 예들이 실시 가능하다는 것은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 자명한 것이다.

도면

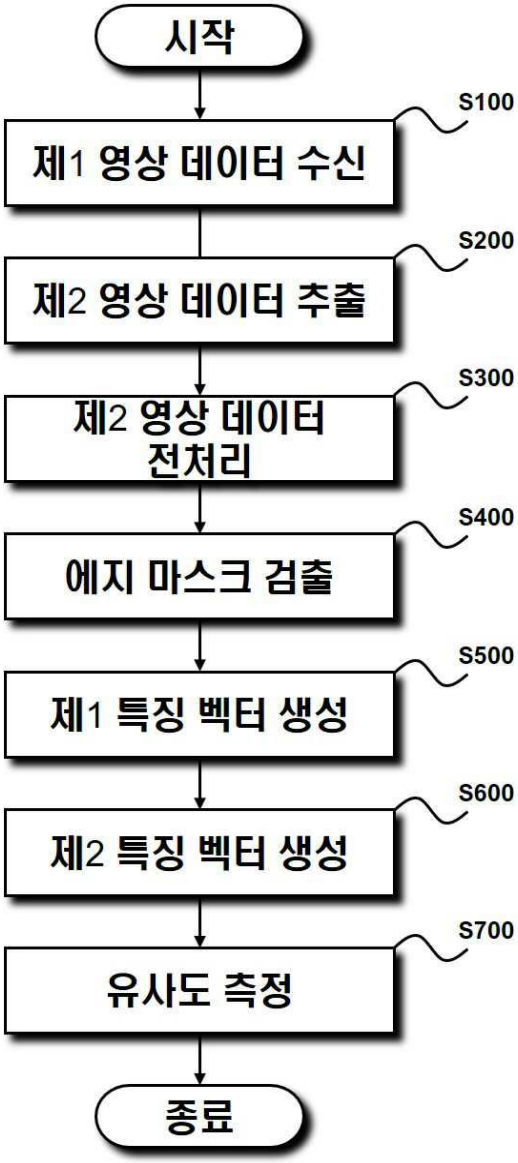
도면1


영상 데이터 저장
HDFS
영상 데이터
로드
RDD
File path
File path
File path
특징 벡터 로드
코사인 유사도를 이용한
유사도 측정
영상 데이터 전처리
특징 벡터 저장
ELP를 이용한
특징 벡터 추출

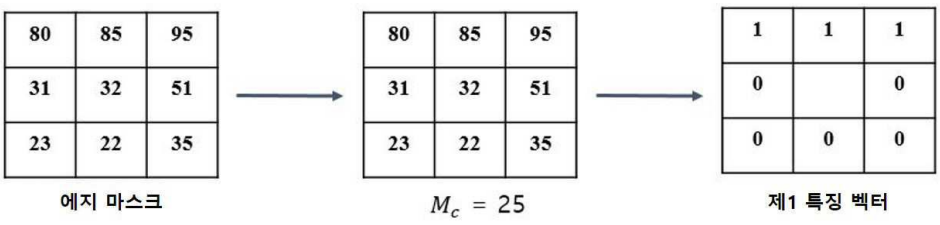
도면2


유사도 예측 장치
영상 수신부 [100]
전처리부 [200]
제어부 [300]
에지 검출부 [310]
특징 벡터 생성부 [330]
데이터 분할부 [400]
유사도 측정부 [500]
저장부 [600]

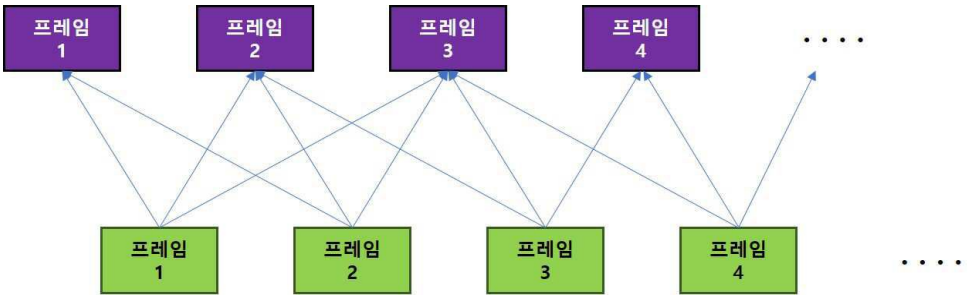
도면3


시작
S100
제1 영상 데이터 수신
S200
제2 영상 데이터 추출
S300
제2 영상 데이터 전처리
S400
에지 마스크 검출
S500
제1 특징 벡터 생성
S600
제2 특징 벡터 생성
S700
유사도 측정
종료

도면4


80 85 95
31 32 51
23 22 35
에지 마스크
80 85 95
31 32 51
23 22 35
M_c = 25
1 1 1
0 0
0 0 0
제1 특징 벡터

도면5



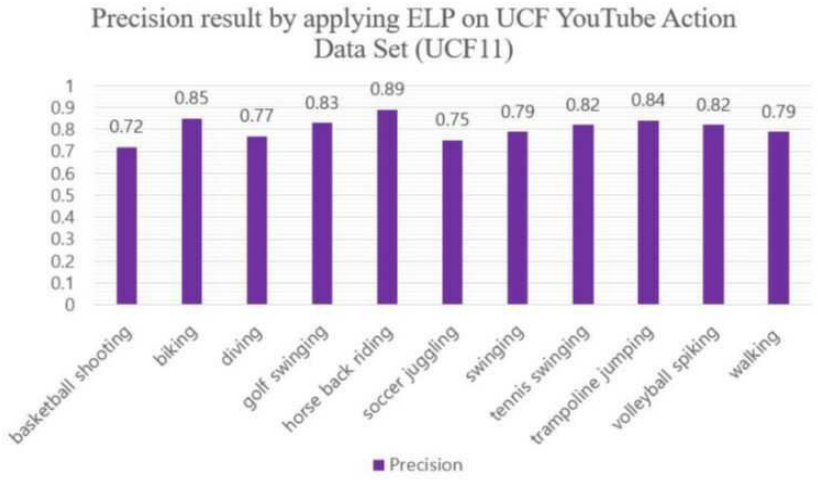
도면6

프레임 수	비교 프레임 수	연산 시간(초)
136	3	94.3782
136	4	108.8361
136	5	129.1928
272	3	170.9647
272	4	194.7050
272	5	232.8930

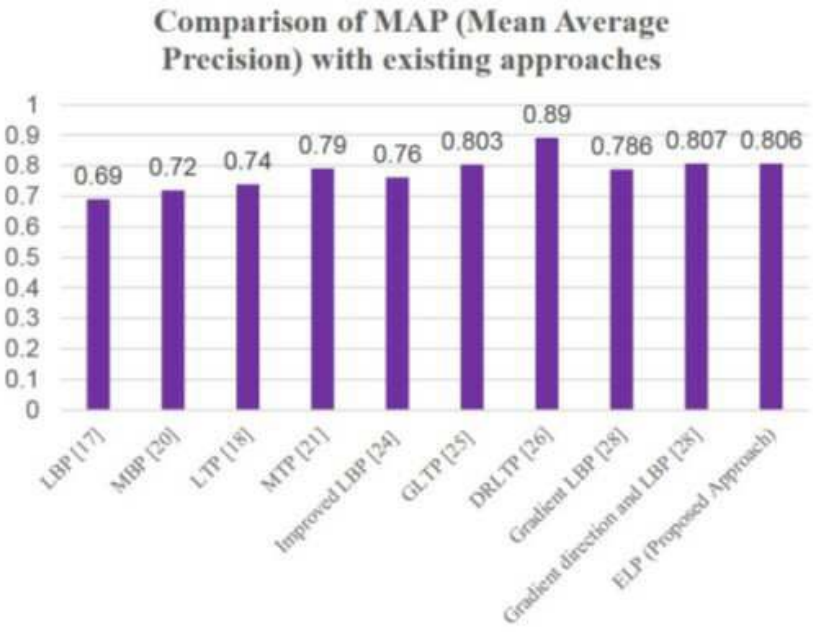
도면7



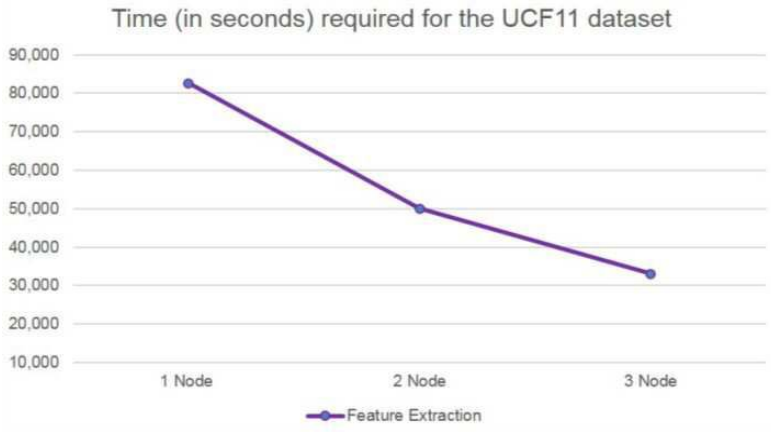
도면8


Precision result by applying ELP on UCF YouTube Action Data Set (UCF11)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.72
0.85
0.77
0.83
0.89
0.75
0.79
0.82
0.84
0.82
0.79
basketball shooting
biking
diving
golf swinging
horse back riding
soccer juggling
swinging
tennis swinging
trampoline jumping
volleyball spiking
walking
Precision

도면9


Comparison of MAP (Mean Average Precision) with existing approaches
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.69
0.72
0.74
0.79
0.76
0.803
0.89
0.786
0.807
0.806
LBP [17]
MBP [20]
LTP [18]
MTP [21]
Improved LBP [24]
GLTP [25]
DRLTP [26]
Gradient LBP [28]
Gradient direction and LBP [28]
ELP (Proposed Approach)

도면10


Time (in seconds) required for the UCF11 dataset
90,000
80,000
70,000
60,000
50,000
40,000
30,000
20,000
10,000
1 Node
2 Node
3 Node
Feature Extraction