



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년08월09일
(11) 등록번호 10-2430756
(24) 등록일자 2022년08월04일

(51) 국제특허분류(Int. Cl.)
G06T 7/11 (2017.01) G06T 7/13 (2017.01)
G06T 7/40 (2017.01) G06T 7/90 (2017.01)
(52) CPC특허분류
G06T 7/11 (2017.01)
G06T 7/13 (2017.01)
(21) 출원번호 10-2020-0019677
(22) 출원일자 2020년02월18일
심사청구일자 2020년02월18일
(65) 공개번호 10-2021-0105126
(43) 공개일자 2021년08월26일
(56) 선행기술조사문헌
Hrishikesh Bhaumik외 3인, "Detection of Gradual Transition in Videos:Approaches and Applications", ResearchGate, (2016.07.28.) 1부.*
Muhammad Sajjad외 5인, "Integrating salient colors with rotational invariant texture features for image representation in retrieval systems", SpringerLink, (2017.08.18.) 1부.*
Songhao Zhu외 1인, "Video Scene segmentation and semantic representation using a novel scheme", SpringerLink, (2008.10.08.) 1부.*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
세종대학교산학협력단
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
(72) 발명자
백성욱
서울특별시 광진구 아차산로 262, B동 1304호 (자양동, 더샵스타시티)
칸 무하마드
서울특별시 광진구 능동로 209, 세종대학교 대양AI센터 454호
(뒷면에 계속)
(74) 대리인
특허법인위더피플

전체 청구항 수 : 총 11 항

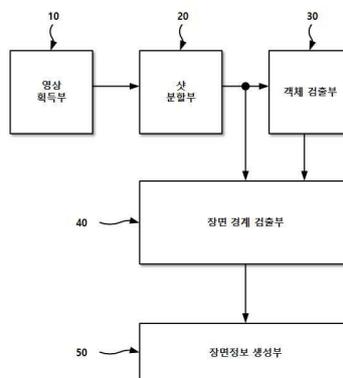
심사관 : 박금옥

(54) 발명의 명칭 객체탐지 및 집합이론을 이용한 영상의 장면 분할 장치 및 방법

(57) 요약

본 발명은 영화 등과 같은 영상의 장면을 분할하는 장면 분할 장치 및 방법에 관한 것으로, 더욱 상세하게는 영상의 프레임 열로부터 색 및 텍스처(Texture) 특징을 이용하여 샷(Shot)을 분할하고, 샷을 구성하는 각 프레임의 객체를 탐지한 후 탐지된 샷의 객체에 따른 집합이론을 적용하여 장면의 경계를 검출하여 장면을 분할하는 객체 탐지 및 집합이론을 이용한 영상의 장면 분할 장치 및 방법에 관한 것이다.

대표도 - 도1



(52) CPC특허분류

G06T 7/40 (2013.01)
G06T 7/90 (2017.01)
G06V 20/00 (2022.01)
H04N 21/845 (2013.01)

이자즈 율 하크

서울특별시 광진구 능동로 209, 세종대학교 대양A
I센터 411호

(72) 발명자

이미영

서울특별시 강남구 도산대로101길 34, 603호 (청담
동, 구산아파트)

탄비어 후세인

서울특별시 광진구 능동로 209, 세종대학교 대양A
I센터 411호

이 발명을 지원한 국가연구개발사업

과제고유번호	1345293063
부처명	교육부
과제관리(전문)기관명	한국연구재단
연구사업명	이공학개인기초연구
연구과제명	지능형 행동 기반 프로파일링을 위한 인물 성향 및 행동 연구
기 여 율	1/1
과제수행기관명	세종대학교산학협력단
연구기간	2018.06.01 ~ 2021.05.31

공지예외적용 : 있음

명세서

청구범위

청구항 1

다수의 프레임으로 구성되는 영상을 획득하여 출력하는 영상 획득부;

상기 영상 획득부를 통해 획득되는 영상에서 적어도 하나 이상의 프레임을 포함하는 샷을 검출하여 출력하는 샷 분할부;

상기 샷 분할부에서 분할된 샷에 포함되어 있는 객체를 검출하고 검출된 객체정보를 출력하는 객체 검출부;

상기 샷 및 상기 해당 샷별 객체정보를 입력받고 샷별 객체의 변화에 따른 집합이론을 적용하여 장면의 경계를 검출하고, 경계 정보를 출력하는 장면 경계 검출부; 및

상기 경계 정보에 의해 장면을 분할하고, 분할된 장면정보를 생성하여 출력하는 장면정보 생성부를 포함하고,

상기 장면 경계 검출부는,

상기 샷이 다른 샷과 동일한 객체를 포함하는지 여부에 따라 상기 장면의 경계를 검출하고,

적어도 셋 이상의 샷을 포함하도록 구성되어 순차적으로 하나의 샷 단위로 이동하는 슬라이딩 윈도우를 구비하되, 상기 슬라이딩 윈도우의 마지막 샷과 이전 샷 중 어느 하나 이상의 샷에서 검출된 객체들을 비교하여 집합이론에 따른 장면 경계를 검출하는

객체탐지 및 집합이론을 이용한 영상의 장면 분할 장치.

청구항 2

제1항에 있어서,

상기 샷 분할부는,

상기 영상 획득부로부터 입력되는 영상을 프레임으로 분할하여 출력하는 프레임 검출부;

상기 프레임 검출부에서 출력되는 프레임별로 컬러 히스토그램을 적용하여 색 특징을 추출하고, 추출된 색 특징에 대한 색 특징정보를 출력하는 색 특징정보 추출부;

상기 프레임 검출부에서 출력되는 프레임별로 로컬 바이너리 패턴(Local Binary Pattern: LBP)을 적용하여 질감 특징을 추출하고, 추출된 질감 특징에 대한 질감 특징정보를 출력하는 텍스처 특징정보 추출부; 및

연속 입력되는 프레임 중 연속되는 두 프레임에 대해 상기 색 특징정보 추출부 및 텍스처 특징정보 추출부로부터 입력되는 색 특징정보 및 질감 특징정보를 비교하여 유사성을 계산하고, 계산된 유사성이 미리 설정된 기준치를 초과하는지에 따라 샷을 분할하여 출력하는 샷 결정부를 포함하는 것을 특징으로 하는 객체탐지 및 집합이론을 이용한 영상의 장면 분할 장치.

청구항 3

제2항에 있어서,

상기 색 특징정보 추출부는,

상기 컬러 히스토그램으로, 색상, 채도, 명도로 색을 지정하는 HSV(Hue saturation value)를 적용하는 것을 특징으로 하는 객체탐지 및 집합이론을 이용한 영상의 장면 분할 장치.

청구항 4

제3항에 있어서,

상기 텍스처 특징정보 추출부는,

58개의 균일한 패턴만을 선택한 다음 5개의 2*2의 각도 구조를 적용한 290개의 질감 특징을 벡터로 구분하여 질감 특징정보를 출력하는 것을 특징으로 하는 객체탐지 및 집합이론을 이용한 영상의 장면 분할 장치.

청구항 5

제4항에 있어서,

상기 샷 결정부는,

360 크기의 특징 벡터(Feature Vector, FV)로 표현되는 단일 프레임 열 중 두 프레임 간의 유사성 점수를 하기 수학적 식 2에 의해 계산하고,

상기 360 크기의 특징 벡터는

상기 HSV 색상 모델의 색상 양자화로 획득되는 70(bin)의 색상 히스토그램과 290개의 질감 특징을 연결하여 달성되는 것을 특징으로 하는 객체탐지 및 집합이론을 이용한 영상의 장면 분할 장치.

[수학적 식 2]

$$S_{(i,j)} = \sum_{b \in bins} \min(FV_i(b), FV_j(b)), \quad \text{where } S_{(i,j)} \in [0, 1]$$

$$FV_j \in \begin{cases} \text{same shot frames,} & \text{if } S_{(i,j)} > \text{threshold} \\ \text{different shots' frames,} & \text{otherwise} \end{cases}$$

상기 수학적 식 2에서, $FV_i(a)$ 와 $FV_i(b)$ 는 상기 두 프레임에 대한 360 크기의 특징 벡터임.

청구항 6

제1항에 있어서,

상기 객체 검출부는,

분할된 각 샷에 대해 영화의 장면을 구성하는 중요한 하나 이상의 객체들에 대해 합성곱 신경망 네트워크(Convolution Neural Network: CNN) 학습을 수행한 CNN 모델에 의해 객체를 검출하는 것을 특징으로 하는 객체 탐지 및 집합이론을 이용한 영상의 장면 분할 장치.

청구항 7

삭제

청구항 8

영상 획득부가 다수의 프레임으로 구성되는 영상을 획득하여 출력하는 영상 획득 과정;

샷 분할부가 상기 영상 획득부를 통해 획득되는 영상에서 적어도 하나 이상의 프레임을 포함하는 샷을 검출하여 출력하는 샷 분할 과정;

객체 검출부가 상기 샷 분할부에서 분할된 샷에 포함된 객체를 검출하고 검출된 객체정보를 출력하는 객체 검출 과정;

장면 경계 검출부가 상기 샷 및 상기 해당 샷별 객체정보를 입력받고 샷별 객체의 변화에 따른 집합이론을 적용하여 장면의 경계를 검출하고, 경계 정보를 출력하는 장면 경계 검출 과정; 및

장면정보 생성부가 상기 경계 정보에 의해 장면을 분할하고, 분할된 장면정보를 생성하여 출력하는 장면정보 생성 과정을 포함하고,

상기 장면 경계 검출 과정은,

상기 샷이 다른 샷과 동일한 객체를 포함하는지 여부에 따라 상기 장면의 경계를 검출하고,

적어도 셋 이상의 샷을 포함하도록 구성되어 순차적으로 하나의 샷 단위로 이동하는 슬라이딩 윈도우를 구비하되, 상기 슬라이딩 윈도우의 마지막 샷과 이전 샷 중 어느 하나 이상의 샷에서 검출된 객체들을 비교하여 집합이론에 따른 장면 경계를 검출하는

객체탐지 및 집합이론을 이용한 영상의 장면 분할 방법.

청구항 9

제8항에 있어서,

상기 샷 분할 과정은,

프레임 검출부가 상기 영상 획득부로부터 입력되는 영상을 프레임으로 분할하여 출력하는 프레임 검출 단계;

색 특징정보 추출부가 상기 프레임 검출부에서 출력되는 프레임별로 컬러 히스토그램을 적용하여 색 특징을 추출하여 출력하는 색 특징정보 추출 단계;

텍스처 특징정보 추출부가 상기 프레임 검출부에서 출력되는 프레임별로 로컬 바이너리 패턴(Local Binary Pattern: LBP)을 적용하여 질감 특징을 추출하여 출력하는 텍스처 특징정보 추출 단계; 및

샷 결정부가 연속 입력되는 프레임 중 연속되는 두 프레임에 대해 상기 색 특징정보 추출부 및 텍스처 특징정보 추출부로부터 입력되는 색 특징정보 및 질감 특징정보를 비교하여 유사성을 계산하고, 계산된 유사성이 미리 설정된 기준치를 초과하는지에 따라 샷을 분할하여 출력하는 샷 결정 단계를 포함하는 것을 특징으로 하는 객체탐지 및 집합이론을 이용한 영상의 장면 분할 방법.

청구항 10

제9항에 있어서,

상기 색 특징정보 추출부는,

상기 색 특징정보 추출 단계에서 상기 컬러 히스토그램으로, 색상, 채도, 명도로 색을 지정하는 HSV(Hue saturation value)를 적용하여 색 특징정보를 추출하는 것을 특징으로 하는 객체탐지 및 집합이론을 이용한 영상의 장면 분할 방법.

청구항 11

제10항에 있어서,

상기 텍스처 특징정보 추출부는,

상기 텍스처 특징정보 추출 단계에서 58개의 균일한 패턴만을 선택한 다음 5개의 2*2의 각도 구조를 적용한 290개의 질감 특징을 벡터로 구분하여 질감 특징정보를 출력하는 것을 특징으로 하는 객체탐지 및 집합이론을 이용한 영상의 장면 분할 방법.

청구항 12

제11항에 있어서,

상기 샷 결정부는,

360 크기의 특징 벡터(Feature Vector: FV)로 표현되는 단일 프레임 열 중 두 프레임 간의 유사성 점수를 하기 수학적 식 2에 의해 계산하고,

상기 360 크기의 특징 벡터는

상기 HSV 색상 모델의 색상 양자화로 획득되는 70(bin)의 색상 히스토그램과 290개의 질감 특징을 연결하여 달성되는 것을 특징으로 하는 객체탐지 및 집합이론을 이용한 영상의 장면 분할 방법.

[수학적 식 2]

$$S_{(i,j)} = \sum_{b \in bins} \min(FV_i(b), FV_j(b)), \quad \text{where } S_{(i,j)} \in [0, 1]$$

$$FV_j \in \begin{cases} \text{same shot frames,} & \text{if } S_{(i,j)} > \text{threshold} \\ \text{different shots' frames,} & \text{otherwise} \end{cases}$$

상기 수학적 식 2에서, $FV_i(a)$ 와 $FV_i(b)$ 는 상기 두 프레임에 대한 360 크기의 특징 벡터임.

발명의 설명

기술 분야

[0001] 본 발명은 영화 등과 같은 영상의 장면을 분할하는 장면 분할 장치 및 방법에 관한 것으로, 더욱 상세하게는 영상의 프레임 열로부터 색 및 텍스처(Texture) 특징을 이용하여 샷(Shot)을 분할하고, 샷을 구성하는 각 프레임의 객체를 탐지한 후 탐지된 샷의 객체에 따른 집합이론을 적용하여 장면의 경계를 검출하여 장면을 분할하는 객체탐지 및 집합이론을 이용한 영상의 장면 분할 장치 및 방법에 관한 것이다.

배경 기술

- [0002] 일반적으로 인터넷 기술의 발전과, DVD, HDTV, 위성 TV, 셋톱박스, 디지털카메라, 스마트폰, 스마트패드 등과 같은 멀티미디어 기기 기술이 발달함에 따라 인터넷을 통한 영상의 수요가 기하급수적으로 증가하고 있다.
- [0003] 이러한 멀티미디어 기기의 발전과 영상의 급속적인 증가는 영상을 효율적으로 플레이하고, 효과적으로 관리하기 위한 연구가 진행되고 있다.
- [0004] 이러한 연구 중 하나가 영상을 적절하게 분하여 구조화하는 영상 분할 기술이다.
- [0005] 영상 분할 기술은 영상을 용이하게 저장, 색인 및 검색을 할 수 있도록 하기 위한 것으로, 영상의 중간에 광고를 삽입하기 위한 용도로도 사용되고 있다.
- [0006] 일반적으로 영상 분할 기술은 장면 단위로 분할하는 것이 일반적이며, 장면 분할을 위해 컬러 히스토그램 비교 방법, 화소 단위 비교 방법 등이 이용되고 있다.
- [0007] 화소 단위 비교 방법은 동일한 장면 내에서는 화소값의 변화가 적다는 점에 착안하여, 연속하는 한 쌍의 프레임에서 대응하는 화소값을 비교하여 얼마나 많은 변화가 발생하였는지를 측정한다. 이 방법의 구현은 간단하나 카메라 움직임에 민감하여, 움직임이 많은 영상에 적용할 경우 장면을 정확하게 구분할 수 없는 단점이 있다.
- [0008] 한편, 컬러 히스토그램 비교 방법은 동일한 장면 내의 프레임들이 상호 유사한 색상 분포를 가진다는 특성을 이용하여, 인접 프레임들의 히스토그램 차이를 임계값과 비교하여 장면을 분할한다. 이 방법은 카메라의 이동에는 덜 민감하나, 빛의 영향에 민감한 단점이 있다. 즉, 갑작스러운 조명 변화가 있는 경우 동일한 장면을 다른 장면으로 인식할 수 있고, 다른 장면임에도 불구하고 색상 분포가 유사하면 이를 검출하지 못하는 문제점이 있다.
- [0009] 이와 같이 현재는 영상의 화소값이나 색상 분포 등 RGB값을 기초 정보로 이용하기 때문에 카메라의 움직임이나 시간/공간적 연속성을 정확히 판단할 수 없어 영상의 장면을 효과적으로 분할하기 어려운 문제점이 있었다.
- [0010] 이러한 문제점을 해결하기 위해 대한민국 등록특허 제10-1050255호 [동영상 장면 분할 시스템 및 방법](이하 "

선행기술"이라 함)는 특정 포맷(MPEG)의 I-프레임 및 P-프레임의 에너지를 이용하여 장면을 분할하는 기술을 개시한다.

[0011] 그러나 상술한 선행기술은 I-프레임 및 P-프레임을 가지는 포맷의 동영상에 대해서만 장면을 검출할 수 있을 뿐, 다른 포맷의 영상에는 적용할 수 없는 문제점이 있으며, 상기 I-프레임 및 P-프레임이 정확하게 구성되었는지에 대한 신뢰성을 제공할 수 없는 문제점이 있었다.

선행기술문헌

특허문헌

[0012] (특허문헌 0001) 대한민국 등록특허 제10-1050255호(2011.07.19. 공고)

발명의 내용

해결하려는 과제

[0013] 따라서 본 발명의 목적은 영상의 프레임 열로부터 색 및 텍스처(Texture) 특징을 이용하여 샷(Shot)을 분할하고, 샷을 구성하는 각 프레임의 객체를 탐지한 후 탐지된 샷의 객체에 따른 집합이론을 적용하여 장면의 경계를 검출하여 장면을 분할하는 객체탐지 및 집합이론을 이용한 영상의 장면 분할 장치 및 방법을 제공함에 있다.

과제의 해결 수단

[0014] 상기와 같은 목적을 달성하기 위한 본 발명에 따른 객체탐지 및 집합이론을 이용한 영상의 장면 분할 장치는: 다수의 프레임으로 구성되는 영상을 획득하여 출력하는 영상 획득부; 상기 영상 획득부를 통해 획득되는 영상에서 적어도 하나 이상의 프레임을 포함하는 샷을 검출하여 출력하는 샷 분할부; 상기 샷 분할부에서 분할된 샷에 포함되어 있는 객체를 검출하고 검출된 객체정보를 출력하는 객체 검출부; 상기 샷 및 상기 해당 샷별 객체정보를 입력받고 샷별 객체의 변화에 따른 집합이론을 적용하여 장면의 경계를 검출하고, 경계 정보를 출력하는 장면 경계 검출부; 및 상기 경계 정보에 의해 장면을 분할하고, 분할된 장면정보를 생성하여 출력하는 장면정보 생성부를 포함하는 것을 특징으로 한다.

[0015] 상기 샷 분할부는, 상기 영상 획득부로부터 입력되는 영상을 프레임으로 분할하여 출력하는 프레임 검출부; 상기 프레임 검출부에서 출력되는 프레임별로 컬러 히스토그램을 적용하여 색 특징을 추출하고, 추출된 색 특징에 대한 색 특징정보를 출력하는 색 특징정보 추출부; 상기 프레임 검출부에서 출력되는 프레임별로 로컬 바이너리 패턴(Local Binary Pattern: LBP)을 적용하여 질감 특징을 추출하고, 추출된 질감 특징에 대한 질감 특징정보를 출력하는 텍스처 특징정보 추출부; 및 연속 입력되는 프레임 중 연속되는 두 프레임에 대해 상기 색 특징정보 추출부 및 텍스처 특징정보 추출부로부터 입력되는 색 특징정보 및 질감 특징정보를 비교하여 유사성을 계산하고, 계산된 유사성이 미리 설정된 기준치를 초과하는지에 따라 샷을 분할하여 출력하는 샷 결정부를 포함하는 것을 특징으로 한다.

[0016] 상기 색 특징정보 추출부는, 상기 컬러 히스토그램으로, 색상, 채도, 명도로 색을 지정하는 HSV(Hue saturation value)를 적용하는 것을 특징으로 한다.

[0017] 상기 텍스처 특징정보 추출부는, 58개의 균일한 패턴만을 선택한 다음 5개의 2*2의 각도 구조를 적용한 290개의 질감 특징을 벡터로 구분하여 질감 특징정보를 출력하는 것을 특징으로 한다.

[0018] 상기 샷 결정부는, 상기 HSV 색상 모델의 색상 양자화 및 290개의 질감 특징을 사용하여 달성되는 70빈(bin)의 색상 히스토그램을 연결함으로써 360 크기의 특징 벡터(Feature Vector: FV)로 표현되는 단일 프레임 열 중 하기 수학적 2에 의해 두 프레임 간의 유사성 점수를 계산하는 것을 특징으로 한다.

[0019] [수학적 2]

[0020]
$$S_{(i,j)} = \sum_{b \in bins} \min(FV_i(b), FV_j(b)), \quad where S_{(i,j)} \in [0, 1]$$

$$FV_j \in \begin{cases} \text{same shot frames,} & \text{if } S_{(i,j)} > \text{threshold} \\ \text{different shots' frames,} & \text{otherwise} \end{cases}$$

[0021]

[0022]

상기 객체 검출부는, 분할된 각 샷에 대해 영화의 장면을 구성하는 중요한 하나 이상의 객체들에 대해 합성곱 신경망 네트워크(Convolution Neural Network: CNN) 학습을 수행한 CNN 모델에 의해 객체를 검출하는 것을 특징으로 한다.

[0023]

상기 장면 경계 검출부는, 적어도 셋 이상의 샷을 포함하도록 구성되어 순차적으로 하나의 샷 단위로 이동하는 슬라이딩 윈도우를 구비하되, 상기 슬라이딩 윈도우의 마지막 샷과 이전 샷 중 어느 하나 이상의 샷에서 검출된 객체들을 비교하여 집합이론에 따른 장면 경계를 검출하는 것을 특징으로 한다.

[0024]

상기와 같은 목적을 달성하기 위한 본 발명에 따른 객체탐지 및 집합이론을 이용한 영상의 장면 분할 방법은: 영상 획득부가 다수의 프레임으로 구성되는 영상을 획득하여 출력하는 영상 획득 과정; 샷 분할부가 상기 영상 획득부를 통해 획득되는 영상에서 적어도 하나 이상의 프레임을 포함하는 샷을 검출하여 출력하는 샷 분할 과정; 객체 검출부가 상기 샷 분할부에서 분할된 샷에 포함된 객체를 검출하고 검출된 객체정보를 출력하는 객체 검출 과정; 장면 경계 검출부가 상기 샷 및 상기 해당 샷별 객체정보를 입력받고 샷별 객체의 변화에 따른 집합이론을 적용하여 장면의 경계를 검출하고, 경계 정보를 출력하는 장면 경계 검출 과정; 및 장면정보 생성부가 상기 경계 정보에 의해 장면을 분할하고, 분할된 장면정보를 생성하여 출력하는 장면정보 생성 과정을 포함하는 것을 특징으로 한다.

[0025]

상기 샷 분할 과정은, 프레임 검출부가 상기 영상 획득부로부터 입력되는 영상을 프레임으로 분할하여 출력하는 프레임 검출 단계; 색 특징정보 추출부가 상기 프레임 검출부에서 출력되는 프레임별로 컬러 히스토그램을 적용하여 색 특징을 추출하여 출력하는 색 특징정보 추출 단계; 상기 텍스처 특징정보 추출부가 상기 프레임 검출부에서 출력되는 프레임별로 로컬 바이너리 패턴(Local Binary Pattern: LBP)을 적용하여 질감 특징을 추출하여 출력하는 텍스처 특징정보 추출 단계; 및 샷 결정부가 연속 입력되는 프레임 중 연속되는 두 프레임에 대해 상기 색 특징정보 추출부 및 텍스처 특징정보 추출부로부터 입력되는 색 특징정보 및 질감 특징정보를 비교하여 유사성을 계산하고, 계산된 유사성이 미리 설정된 기준치를 초과하는지에 따라 샷을 분할하여 출력하는 샷 결정 단계를 포함하는 것을 특징으로 한다.

[0026]

상기 색 특징정보 추출부는, 상기 색 특징정보 추출 단계에서 상기 컬러 히스토그램으로, 색상, 채도, 명도로 색을 지정하는 HSV(Hue saturation value)를 적용하여 색 특징정보를 추출하는 것을 특징으로 한다.

[0027]

상기 텍스처 특징정보 추출부는, 상기 텍스처 특징정보 추출 단계에서 58개의 균일한 패턴만을 선택한 다음 5개의 2*2의 각도 구조를 적용한 290개의 질감 특징을 벡터로 구분하여 질감 특징정보를 출력하는 것을 특징으로 한다.

[0028]

상기 샷 결정부는, 상기 샷 결정 단계에서 상기 HSV 색상 모델의 색상 양자화 및 290개의 질감 특징을 사용하여 달성되는 70bin(bin)의 색상 히스토그램을 연결함으로써 360 크기의 특징 벡터(Feature Vector: FV)로 표현되는 단일 프레임 열 중 하기 수학식 2에 의해 두 프레임 간의 유사성 점수를 계산하는 것을 특징으로 한다.

[0029]

[수학식 2]

[0030]

$$S_{(i,j)} = \sum_{b \in \text{bins}} \min(FV_i(b), FV_j(b)), \quad \text{where } S_{(i,j)} \in [0, 1]$$

[0031]

$$FV_j \in \begin{cases} \text{same shot frames,} & \text{if } S_{(i,j)} > \text{threshold} \\ \text{different shots' frames,} & \text{otherwise} \end{cases}$$

발명의 효과

[0032]

본 발명은 프레임의 색 및 질감을 복합 적용하여 적어도 하나 이상의 프레임을 포함하는 샷(Shot)을 구분하므로 보다 정밀하게 샷을 분할할 수 있는 효과가 있다.

[0033]

또한, 본 발명은 샷 내에 포함된 객체를 탐지하고, 검출된 객체들에 집합이론을 적용하여 장면을 분할함으로써 보다 정밀하게 장면을 분할할 수 있는 효과가 있다.

도면의 간단한 설명

- [0034] 도 1은 본 발명에 따른 객체탐지 및 집합이론을 이용한 영상의 장면 분할 장치의 구성을 나타낸 도면이다
- 도 2는 본 발명의 일실시예에 따른 장면 분할 장치의 장면 분할 개념 및 방법을 나타낸 도면이다
- 도 3은 본 발명의 일실시예에 따른 장면 분할 장치의 영상 샷의 객체 검출 예를 나타낸 도면이다
- 도 4는 본 발명의 일실시예에 따른 장면 분할 장치의 슬라이딩 윈도우 및 집합이론을 이용한 장면 분할 개념을 설명하기 위한 도면이다
- 도 5는 본 발명에 따른 장면 분할 장치의 샷 분할부의 상세 구성을 나타낸 도면이다

발명을 실시하기 위한 구체적인 내용

- [0035] 이하 첨부된 도면을 참조하여 본 발명에 따른 객체탐지 및 집합이론을 이용한 영상의 장면 분할 장치의 구성 및 동작을 상세히 설명하고, 상기 장치에서의 장면 분할 방법을 설명한다.
- [0036] 본 발명에서 영상은 다수의 프레임으로 구성되고, 하나 또는 둘 이상의 상호 연관된 프레임을 포함하는 다수의 샷(Shot)으로 구성되며, 하나 이상의 샷을 포함하는 다수의 장면을 포함한다. 즉 하나의 영상은 다수의 장면을 포함하고, 하나의 장면은 적어도 하나 이상의 샷을 포함할 수 있으며, 하나의 샷은 적어도 하나 이상의 프레임을 포함할 수 있을 것이다.
- [0037] 본 발명은 프레임 단위의 색 특징정보 및 질감(Texture) 특징정보를 이용하여 샷을 구분하고, 구분된 샷 단위로 객체검출 및 집합이론을 적용하여 장면 경계를 검출하여 장면을 분할한다.
- [0038] 도 1은 본 발명에 따른 객체탐지 및 집합이론을 이용한 영상의 장면 분할 장치의 구성을 나타낸 도면이고, 도 2는 본 발명의 일실시예에 따른 장면 분할 장치의 장면 분할 개념 및 방법을 나타낸 도면이고, 도 3은 본 발명의 일실시예에 따른 장면 분할 장치의 영상 샷의 객체 검출 예를 나타낸 도면이며, 도 4는 본 발명의 일실시예에 따른 장면 분할 장치의 슬라이딩 윈도우 및 집합이론을 이용한 장면 분할 개념을 설명하기 위한 도면이다. 이하 도 1 내지 도 4를 참조하여 설명한다.
- [0039] 본 발명에 따른 장면 분할 장치는 영상 획득부(10), 샷 분할부(20), 객체 검출부(30), 장면 경계 검출부(40) 및 장면정보 생성부(50)를 포함한다.
- [0040] 영상 획득부(10)는 카메라(미도시) 등과 같은 영상 획득장치로부터 영상을 획득하거나 내부 저장장치(미도시), 웹 서버 및 클라우드 서버 등의 네트워크 저장장치 등으로부터 영상을 획득하여 출력한다.
- [0041] 샷 분할부(20)는 상기 영상 획득부(10)로부터 출력되는 영상을 입력받고, 도 2의 301단계에서 보이는 바와 같이 영상을 구성하는 프레임(1)들로부터 적어도 하나 이상의 프레임(1)을 포함하는 샷(Shot)(2)을 검출하고, 검출된 샷(2) 단위로 분할하여 출력한다. 샷 분할부(20)의 상세 구성 및 동작은 도 5를 참조하여 상세히 설명한다. 상기 영상에 대해 분할된 샷 열은 객체 검출부(30) 및 장면 경계 검출부(40)로 입력한다.
- [0042] 객체 검출부(30)는 상기 샷 분할부(20)에서 분할된 샷을 입력받고, 도 2의 302단계에서 보이는 바와 같이 상기 샷별로 포함되어 있는 객체(401)들을 검출하고, 검출된 객체(401)들에 대한 객체정보를 장면 경계 검출부(40)로 출력한다. 상기 객체는 도 3의 실시예에서 보이는 바와 같이 사람(401-1), 동물(401-2) 및 자동차(401-3) 등이 될 수 있으나, 이에 한정되지 않는다. 즉 상기 객체는 사람, 개, 고양이 등의 동물, 차, 자전거, 오토바이, 간판, 나무, 건물 등의 물체, 산, 강, 호수, 바다, 하늘 등의 배경 등이 될 수 있을 것이다.
- [0043] 상기 객체 검출부(30)는 도 2의 302단계에서 보이는 바와 같이 샷에 포함된 객체를 검출하기 위해 다양한 객체들을 포함하는 데이터 세트에 대해 학습된 다크넷(Dark net) 백엔드 프레임워크(Backend Framework) 및 YOLOv3 합성곱 신경망 네트워크(Convolution Neural Network: CNN) 모델이 적용된다. 상기 YOLOv3의 아키텍처는 53개의 합성곱 계층 시퀀스에서 3*3 및 1*1 크기의 연속 합성곱 필터를 사용하고, 각 계층 다음에는 Leaky Rectified라 불리는 비선형 활성화 기능이 적용되며, 보폭이 2인 2*2 크기의 맥스 풀링(Max-pooling) 계층이 구성된다.
- [0044] 상기 객체 검출부(30)는 하기 수학적 1과 같이 각 샷의 각 프레임에서 검출된 객체의 조합을 취하여 객체 집합을 관련 샷에 연관시킨 후, 검출된 샷의 객체에 대한 객체정보를 장면 경계 검출부(40)로 출력한다.

수학식 1

[0045]

$$S[ob_1, ob_2, \dots, ob_m] = \sum_{i=1}^n \odot [Sh(fi)]$$

[0046]

$$S_{set} = \bigcup_{j=i}^m S(ob_j)$$

[0047]

여기서, \odot 는 임의의 샷(sh)의 프레임(fi)들에서 객체를 검출하기 위한 함수를 의미하고, sh는 샷(shot)을 의미하고, fi는 샷에 포함된 프레임을 의미하며, ob는 검출된 객체이며, S는 검출된 객체 세트를 의미한다.

[0048]

장면 경계 검출부(40)는 샷 분할부(20)로부터 분할된 샷 열을 입력받고, 객체 검출부(30)로부터 검출된 객체에 대한 객체정보 입력받는다.

[0049]

장면 경계 검출부(40)는 상기 샷 열 및 샷별 객체정보를 입력받고 적어도 두 샷(2) 간의 객체정보를 집합이론에 기반하여 비교하여 객체 변화에 따른 장면 경계를 검출한다.

[0050]

상기 장면 경계 검출부(40)는 도 2의 303단계 및 도 4에서 보이는 바와 같이 적어도 둘 이상의 샷을 포함하는 슬라이딩 윈도우(511)를 구성하고, 슬라이딩 윈도우(511) 내에 포함된 마지막 샷과 이전 샷들을 비교하여 동일 장면인지 새로운 장면인지를 판단한다.

[0051]

도 4와 같이 슬라이딩 윈도우(511)가 4개의 샷을 포함하도록 구성되고 슬라이딩 윈도우(511)의 첫 번째 내지 세 번째 샷(shot2 ~ shot4)이 동일 장면인 경우를 예를 들어 설명하면, 슬라이딩 윈도우(511)의 마지막 샷인 네 번째 샷(shot5)과 첫 번째 내지 세 번째 샷(shot2~shot4) 중 어느 하나 이상과 동일한 객체를 포함하는지를 검사하고, 첫 번째 내지 세 번째 샷 중 어느 하나 이상과 동일한 객체를 포함하고 있으면 상기 네 번째 샷(shot5)을 동일한 장면의 샷으로 분류하고, 동일한 객체를 하나도 포함하고 있지 않으면 다른 장면의 샷으로 분류하여 장면 경계를 결정한다.

[0052]

그러나 도 4의 경우, 네 번째 샷(shot 5)은 슬라이딩 윈도우(511)의 이전 샷(shot2~shot4)들 중 어느 하나와 동일한 객체를 포함하므로, 상기 장면 경계 검출부(40)는 슬라이딩 윈도우(511)의 상기 마지막 샷이 이전 샷과 동일 장면인 것으로 판단하고, 슬라이딩 윈도우(511)를 다음의 샷인 shot 6을 포함하도록 이동시킨다.

[0053]

이동된 슬라이딩 윈도우(511')의 네 번째 샷인 shot6은 이전 샷(shot3~shot5)의 객체들과 동일한 객체를 포함하고 있지 않으므로 장면 경계 검출부(40)는 상기 슬라이딩 윈도우(511')의 네 번째 샷인 shot6을 다른 장면의 샷으로 결정하고 장면 경계를 shot5로 결정하고, 슬라이딩 윈도우(511')를 shot6이 첫 번째 샷이 되도록 이동시킨다.

[0054]

그런 후 상기 장면 경계 검출부(40)는 shot6을 첫 번째 샷으로 포함하는 슬라이딩 윈도우(511'')에 포함된 샷들에 대해 동일 장면인지를 판단하도록 구성하거나, 첫 4개의 샷은 무조건적으로 동일 장면으로 판단한 후 다음 샷에 대해 동일 장면의 샷인지를 판단하도록 구성할 수도 있을 것이다.

[0055]

슬라이딩 윈도우(511) 내의 샷들은 도 2의 303단계에서 보이는 바와 같이 인접한 샷 간 동일 객체를 포함하여야만 동일 장면으로 분류될 수도 있고, 도 4와 같이 슬라이딩 윈도우(511) 내의 마지막 샷이 이전 샷 중 어느 하나와 동일 객체를 포함하는 경우 동일 장면으로 분류될 수도 있을 것이다.

[0056]

후자의 경우, 슬라이딩 윈도우(511)에 포함되는 샷 수는 장면을 구성하는 최소 샷 수일 수 있으며, 상기 샷 수는 실험결과에 따라 적정하게 결정될 수 있을 것이다.

[0057]

장면정보 생성부(50)는 상기 장면 경계 검출부(40)에 의해 장면 경계가 결정되면 장면 경계에 결정에 따른 분할 장면정보를 생성하여 저장하거나 표시 수단을 통해 표시한다.

[0058]

도 5는 본 발명에 따른 장면 분할 장치의 샷 분할부의 상세 구성을 나타낸 도면이다.

[0059]

샷 분할부(20)는 프레임 검출부(21), 색 특징정보 추출부(22), 텍스처 특징정보 추출부(23) 및 샷 결정부(24)를 포함한다.

- [0060] 프레임 검출부(21)는 상기 영상 획득부(10)로부터 입력되는 영상을 프레임으로 분할하여 출력한다. 영상에 대한 분할된 프레임 열은 색 특징정보 추출부(22) 및 텍스처 특징정보 추출부(23)로 입력한다.
- [0061] 색 특징정보 추출부(22)는 입력되는 프레임 열의 프레임별로 컬러 히스토그램을 적용하여 색 특징을 추출하고 추출된 색 특징에 대한 색 특징정보를 샷 결정부(24)로 출력한다. 상기 컬러 히스토그램은 색상, 채도, 명도로 색을 지정하는 HSV(Hue, Saturation Value)가 적용되는 것이 바람직할 것이다.
- [0062] 텍스처 특징정보 추출부(23)는 상기 프레임 열의 프레임별로 로컬 바이너리 패턴(Local Binary Pattern: LBP)을 적용하여 질감 특징을 추출하고, 추출된 질감 특징에 대한 질감 특징정보를 출력한다.
- [0063] 상기 텍스처 특징정보 추출부(23)는 58개의 균일한 패턴만을 선택한 다음 5개의 2*2의 각도 구조를 적용한 290개의 질감 특징을 벡터로 구분하여 질감 특징정보를 출력한다.
- [0064] 샷 결정부(24)는 연속 입력되는 프레임 중 연속되는 두 프레임에 대해 상기 색 특징정보 추출부(22) 및 텍스처 특징정보 추출부(23)로부터 입력되는 색 특징정보 및 질감 특징정보를 비교하여 유사성을 계산하고, 계산된 유사성이 미리 설정된 기준치를 초과하는지에 따라 샷을 분할하여 출력한다.
- [0065] 구체적으로, 샷 결정부(24)는 상기 HSV 색상 모델의 색상 양자화 및 290개의 질감 특징을 사용하여 달성되는 70 빈(bin)의 색상 히스토그램을 연결함으로써 360 크기의 특징 벡터(Feature Vector: FV)로 표현되는 단일 프레임 열 중 하기 수학적 식 2에 의해 두 프레임 간의 유사성 점수를 계산하고, 계산된 유사성 점수가 미리 설정된 기준 값(Threshold)을 초과하는지의 여부에 따라 동일 샷인지 다른 샷인지를 결정한다.

수학적 식 2

[0066]
$$S_{(i,j)} = \sum_{b \in bins} \min(FV_i(b), FV_j(b)), \quad \text{where } S_{(i,j)} \in [0,1]$$

[0067]
$$FV_j \in \begin{cases} \text{same shot frames,} & \text{if } S_{(i,j)} > \text{threshold} \\ \text{different shots' frames,} & \text{otherwise} \end{cases}$$

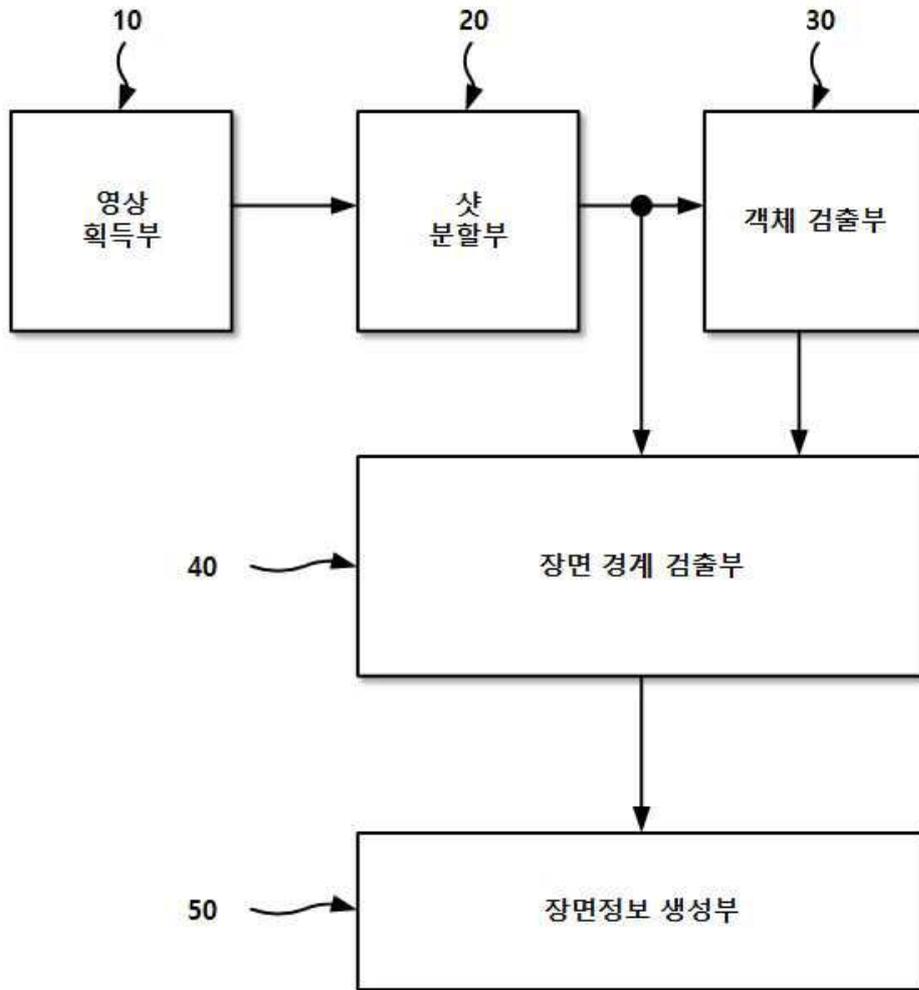
[0068] 한편, 본 발명은 전술한 전형적인 바람직한 실시예에만 한정되는 것이 아니라 본 발명의 요지를 벗어나지 않는 범위 내에서 여러 가지로 개량, 변경, 대체 또는 부가하여 실시할 수 있는 것임은 당해 기술분야에서 통상의 지식을 가진 자라면 용이하게 이해할 수 있을 것이다. 이러한 개량, 변경, 대체 또는 부가에 의한 실시가 이하의 첨부된 특허청구범위의 범주에 속하는 것이라면 그 기술사상 역시 본 발명에 속하는 것으로 보아야 한다.

부호의 설명

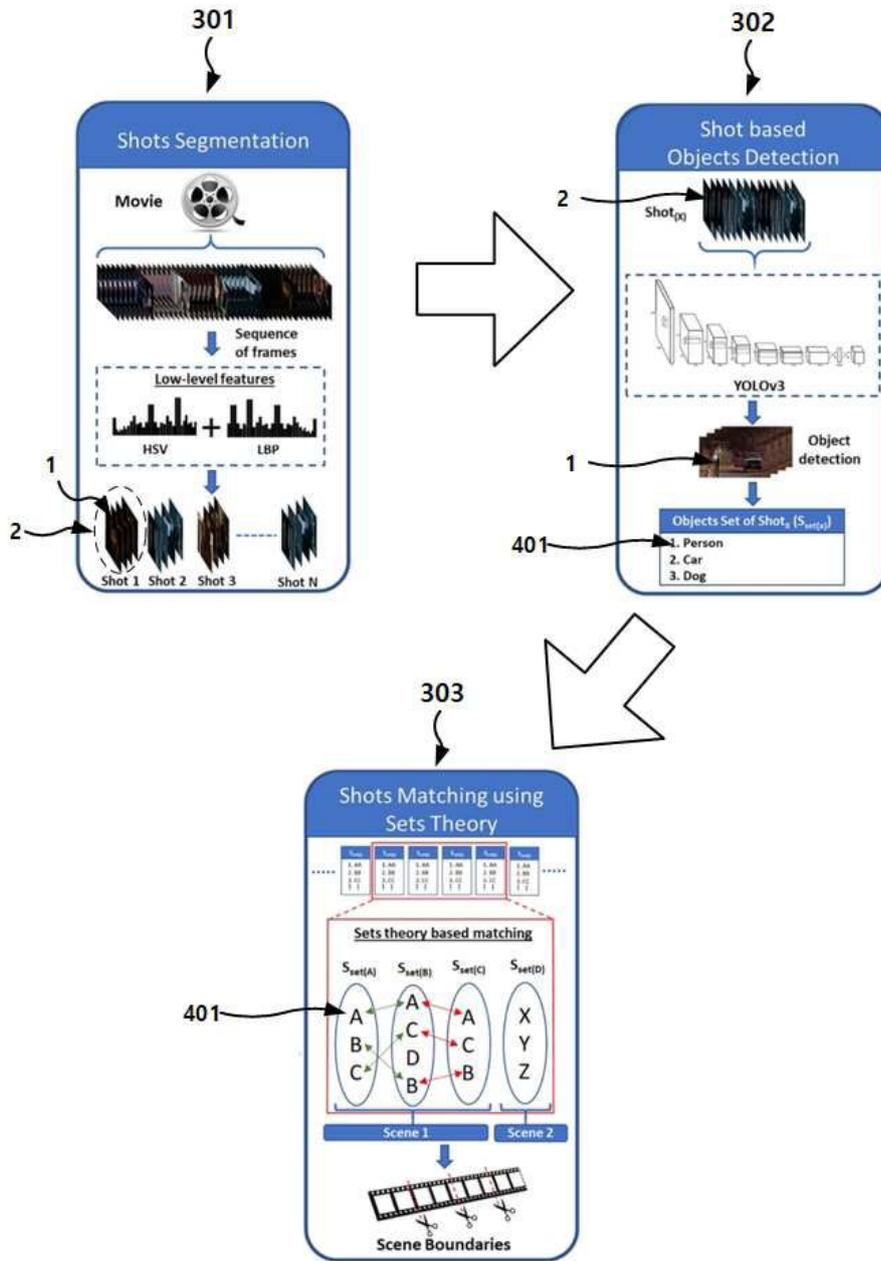
- [0069]
- | | |
|------------------|----------------|
| 10: 영상 획득부 | 20: 샷 분할부 |
| 21: 프레임 검출부 | 22: 색 특징정보 추출부 |
| 23: 텍스처 특징정보 추출부 | 24: 샷 결정부 |
| 30: 객체 검출부 | 40: 장면 경계 검출부 |
| 50: 장면정보 생성부 | |

도면

도면1



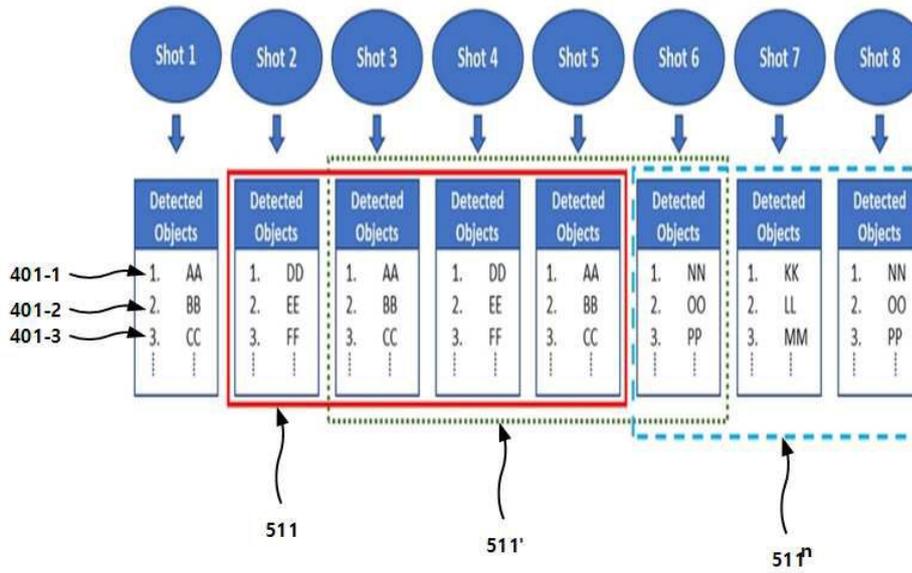
도면2



도면3



도면4



도면5

