(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0048413
(43) 공개일자 2020년05월08일

(51) 국제특허분류(Int. Cl.)
G06K 9/00 (2006.01) *G06K 9/32* (2006.01)
G06T 7/11 (2017.01)
(52) CPC특허분류
G06K 9/00744 (2013.01)
G06K 9/3233 (2013.01)
(21) 출원번호 10-2018-0130495
(22) 출원일자 2018년10월30일
심사청구일자 2018년10월30일

(71) 출원인
서강대학교산학협력단
서울특별시 마포구 백범로 35 (신수동, 서강대학교)
(72) 발명자
낭종호
서울특별시 강남구 언주로146길 18 (논현2동, 동현아파트) 1동 206호
김현곤
전라북도 전주시 완산구 오두정 1길 27-2
석혜경
부산광역시 동래구 충렬대로 314 (낙민동, 동래한양아파트) 7동 806호
(74) 대리인
이지연

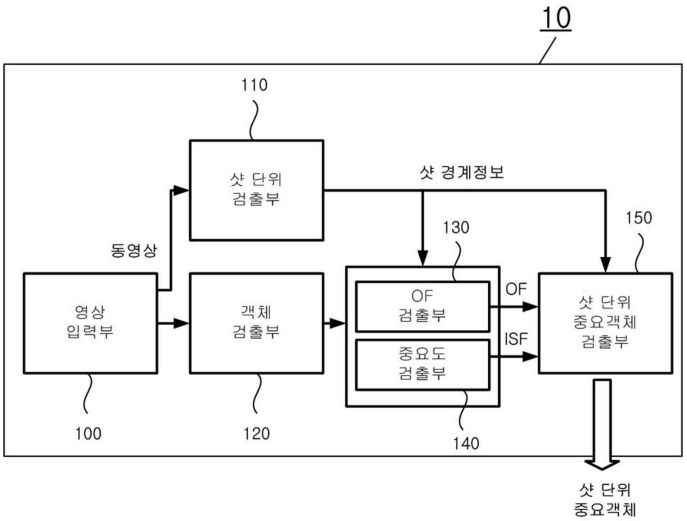
전체 청구항 수 : 총 11 항

(54) 발명의 명칭 **샷 단위 중요 객체 추출 방법 및 장치**

(57) 요 약

본 발명은 샷 단위 중요 객체 추출 방법 및 장치에 관한 것이다. 상기 샷 단위 중요 객체 추출 방법은, (a) 동영상을 샷 단위로 분할하는 단계; (b) 상기 동영상을 구성하는 각 프레임들에 대하여, 객체들을 검출하는 단계; (c) 각 샷에 대하여, 해당 샷 내에서의 객체들에 대한 객체 빈도수(Object Frequency; OF)를 획득하는 단계; (d) 각 객체에 대하여, 상기 동영상 전체에서의 중요도를 획득하는 단계; (e) 각 객체에 대하여 구한 샷 내의 객체 빈도수와 동영상 전체에서의 중요도를 이용하여, 각 샷 내의 중요 객체들을 추출하는 단계;를 구비하여 샷 단위의 중요 객체들을 추출한다.

대 표 도 - 도1



(52) CPC특허분류
G06T 7/11 (2017.01)
G06T 2207/30242 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 1711055377
부처명 과학기술정보통신부
연구관리전문기관 정보통신기술진흥센터
연구사업명 SW중심대학 지원사업
연구과제명 SW중심대학(서강대학교)
기 여 율 1/1
주관기관 서강대학교 산학협력단
연구기간 2018.01.01 ~ 2018.12.31

명 세 서

청구범위

청구항 1

영상 처리 장치로 입력된 동영상에 대하여 상기 영상 처리 장치내에 포함된 프로세서에 의해 수행되는 샷 단위 중요 객체 추출 방법에 있어서,

(a) 동영상을 샷 단위로 분할하는 단계;

(b) 상기 동영상을 구성하는 각 프레임들에 대하여, 객체들을 검출하는 단계;

(c) 각 샷에 대하여, 해당 샷 내에서의 객체들에 대한 객체 빈도수(Object Frequency; OF)를 획득하는 단계;

(d) 각 객체에 대하여, 상기 동영상 전체에서의 중요도를 획득하는 단계;

(e) 각 객체에 대하여 구한 샷 내의 객체 빈도수와 동영상 전체에서의 중요도를 이용하여, 각 샷 내의 중요 객체들을 추출하는 단계;

를 구비하여 샷 단위의 중요 객체들을 추출하는 것을 특징으로 하는 샷 단위 중요 객체 추출 방법.

청구항 2

제1항에 있어서, 상기 (c) 단계의 객체 빈도수(OF)는 특정 객체(o)가 특정 샷(s)에서 등장하는 횟수(of(o,s))인 것을 특징으로 하는 샷 단위 중요 객체 추출 방법.

청구항 3

제1항에 있어서, 상기 (d)단계의 각 객체에 대한 동영상 전체에서의 중요도는,

상기 동영상을 구성하는 모든 샷의 개수(V)와 해당 객체(o)가 검출된 샷들의 개수의 비율을 이용하여 각 객체에 대하여 계산된 샷 빈도수(shot frequency)의 역수(Inverse Shot Frequency;ISF)(isf(o,V))인 것을 특징으로 하는 샷 단위 중요 객체 추출 방법.

청구항 4

제3항에 있어서, 상기 각 객체에 대한 샷 빈도수의 역수(isf(o,V))는 아래의 수학식에 의해 구해지는 것을 특징으로 하는 샷 단위 중요 객체 추출 방법.

$$isf(o,V) = \frac{\log|V|}{|\{s \in V : o \in s\}|}$$

(여기서, $|V|$: 동영상 내의 전체 shot의 개수,
$|\{s \in V : o \in s\}|$: object o가 포함된 shot의 수임.)

청구항 5

제1항에 있어서, 상기 (e) 단계는,

각 샷별로 구한 각 객체에 대한 샷 내의 객체 빈도수(of(o,s))와 해당 객체에 대한 동영상 전체에서의 중요도를 곱해서 구한 값이 사전 설정된 문턱값 이상이면, 해당 객체를 해당 샷에서의 중요 객체로 선택하는 것을 특징으로 하는 샷 단위 중요 객체 추출 방법.

청구항 6

제3항에 있어서, 상기 (e) 단계는,

각 샷별로 구한 각 객체(o)에 대한 샷(s) 내의 객체 빈도수(of(o,s))와 해당 객체에 대한 샷 빈도수의 역수(isf(o,V))를 곱해서 구한 값($ofisf(o,s,V) = of(o,s) * isf(o,V)$)이 사전 설정된 문턱값 이상이면, 해당 객체(o)를 해당 샷(s)에서의 중요 객체로 선택하는 것을 특징으로 하는 샷 단위 중요 객체 추출 방법.

청구항 7

외부로부터 동영상을 입력받는 영상 입력부;

상기 영상 입력부로 입력된 동영상을 샷 단위로 분할하는 샷 단위 검출부;

상기 동영상을 구성하는 프레임들로부터 객체들을 검출하는 객체 검출부;

각 샷에 포함된 객체들에 대한 샷 내의 객체 빈도수(Object Frequency; OF)를 획득하는 OF 검출부;

각 객체에 대하여, 상기 동영상 전체에서의 중요도를 획득하는 중요도 검출부;

각 객체에 대하여 구한 샷 내의 객체 빈도수와 동영상 전체에서의 중요도를 이용하여, 각 샷에 대한 중요 객체를 추출하는 중요 객체 검출부;

를 구비하여, 동영상을 구성하는 각 샷의 중요 객체를 추출하는 것을 특징으로 하는 샷 단위 중요 객체 추출 장치.

청구항 8

제7항에 있어서, 상기 중요도 검출부의 각 객체에 대한 동영상 전체에서의 중요도는,

상기 동영상을 구성하는 모든 샷의 개수(V)와 해당 객체(o)가 검출된 샷들의 개수의 비율을 이용하여 각 객체에 대하여 게산된 샷 빈도수(shot frequency)의 역수(Inverse Shot Frequency;ISF)(isf(o,V))인 것을 특징으로 하는 샷 단위 중요 객체 추출 장치.

청구항 9

제8항에 있어서, 상기 각 객체에 대한 샷 빈도수의 역수(isf(o,V))는 아래의 수학식에 의해 구해지는 것을 특징으로 하는 샷 단위 중요 객체 추출 장치.

$$isf(o,V) = \frac{\log|V|}{|\{s \in V : o \in s\}|}$$

(여기서, $|V|$: 동영상 내의 전체 shot의 개수,
$|\{s \in V : o \in s\}|$: object o가 포함된 shot의 수임.)

청구항 10

제7항에 있어서, 상기 중요 객체 검출부는,

각 샷별로 구한 각 객체에 대한 샷 내의 객체 빈도수(of(o,s))와 해당 객체에 대한 동영상 전체에서의 중요도를 곱해서 구한 값이 사전 설정된 문턱값 이상이면, 해당 객체를 해당 샷에서의 중요 객체로 선택하는 것을 특징으로 하는 샷 단위 중요 객체 추출 장치.

청구항 11

제8항에 있어서, 상기 중요 객체 검출부는,

각 샷별로 구한 각 객체(o)에 대한 샷(s) 내의 객체 빈도수(of(o,s))와 해당 객체에 대한 샷 빈도수의 역수(isf(o,V))를 곱해서 구한 값($$ofisf(o,s,V) = of(o,s) * isf(o,V)$$)이 사전 설정된 문턱값 이상이면, 해당 객체(o)를 해당 샷(s)에서의 중요 객체로 선택하는 것을 특징으로 하는 샷 단위 중요 객체 추출 장치.

발명의 설명

기 술 분 야

[0001] 본 발명은 동영상의 중요 객체 추출 방법에 관한 것으로서, 더욱 구체적으로는 동영상의 각 샷에 대하여, 각 객체들의 해당 샷내에서의 객체 빈도수와 각 객체에 대한 동영상 전체에서의 중요도를 이용하여, 중요 객체들을 선택하여 제공함으로서 샷의 중요 객체에 대한 정확성을 향상시킬 수 있는 샷 단위 중요 객체 추출 방법에 관한 것이다.

배 경 기 술

[0002] 일반적으로, 하나의 미디어 동영상에 있어서 샷(shot) 단위의 메타 데이터들을 검출하여 제공하게 되며, 샷 단위의 메타 데이터들을 정확하게 제공하기 위하여 각 샷의 중요 객체들을 정확하게 추출해야 된다. 동영상 콘텐츠에서 샷(Shot)은 일반적으로 카메라가 한 번의 촬영을 시작하여 멈추기 전까지를 의미하므로, 샷은 화면전환의 경계를 나타낼 수 있으며 비디오 콘텐츠를 나누는 기본적인 단위가 되기도 한다.

[0003] 하지만, 샷의 중요 객체를 추출함에 있어서, 동영상 전체에서 전반적으로 중복되거나 자주 등장하는 객체는, 중요 객체가 아니라 샷의 위치나 상황을 판단하기 어려운 흔한 객체일 가능성이 높다. 따라서, 동영상 전체에서 전반적으로 중복되거나 자주 등장하는 객체는, 샷에 대한 상황 또는 위치등의 연관 관계를 판단하는데 악영향을 미칠수 있게 된다.

[0004] 예컨대, 드라마의 경우, 대부분의 샷에서 사람이 등장하고 드라마의 전체에서 사람이 많이 등장하므로, 드라마에서의 사람이라는 객체는 중요 객체라고 할 수 없다. 하지만, 자연 다큐멘터리의 경우, 일반적으로 사람이라는 객체는 등장하기 않기 때문에, 특정 샷에서 등장하는 사람은 중요 객체라고 할 수 있을 것이다.

[0005] 하지만, 종래의 중요 객체 추출 방법은 전술한 바와 같은 동영상의 특성을 고려하지 못함으로써, 중요 객체를 정확하게 추출해내지 못하는 문제점이 발생한다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) 한국공개특허공보 10-2015-0070629

발명의 내용

해결하려는 과제

[0007] 전술한 문제점을 해결하기 위한 본 발명의 목적은 동영상의 특성을 고려하여, 동영상에서의 샷 단위의 객체 메타 데이터를 정확하게 검출할 수 있도록 하는 샷 단위 중요 객체 추출 방법 및 장치를 제공하는 것이다.

과제의 해결 수단

[0008] 전술한 기술적 과제를 달성하기 위한 본 발명의 제1 특징은, 영상 처리 장치로 입력된 동영상에 대하여 상기 영상 처리 장치내에 포함된 프로세서에 의해 수행되는 샷 단위 중요 객체 추출 방법에 관한 것으로서, 상기 샷 단위 중요 객체 추출 방법은, (a) 동영상을 샷 단위로 분할하는 단계; (b) 상기 동영상을 구성하는 각 프레임들에 대하여, 객체들을 검출하는 단계; (c) 각 샷에 대하여, 해당 샷 내에서의 객체들에 대한 객체 빈도수(Object Frequency; OF)를 획득하는 단계; (d) 각 객체에 대하여, 상기 동영상 전체에서의 중요도를 획득하는 단계; (e)

각 객체에 대하여 구한 샷 내의 객체 빈도수와 동영상 전체에서의 중요도를 이용하여, 각 샷 내의 중요 객체들을 추출하는 단계;를 구비하여 샷 단위의 중요 객체들을 추출한다.

[0009] 전술한 제1 특징에 따른 샷 단위 중요 객체 추출 방법에 있어서, 상기 (c) 단계의 객체 빈도수(OF)는 특정 객체(o)가 특정 샷(s)에서 등장하는 횟수(of(o,s))인 것이 바람직하다.

[0010] 전술한 제1 특징에 따른 샷 단위 중요 객체 추출 방법에 있어서, 상기 (d)단계의 각 객체에 대한 동영상 전체에서의 중요도는, 상기 동영상을 구성하는 모든 샷의 개수(V)와 해당 객체(o)가 검출된 샷들의 개수의 비율을 이용하여 각 객체에 대하여 게산된 샷 빈도수(shot frequency)의 역수(Inverse Shot Frequency;ISF)(isf(o,V))인 것이 바람직하다.

[0011] 전술한 제1 특징에 따른 샷 단위 중요 객체 추출 방법에 있어서, 상기 (e) 단계는, 각 샷별로 구한 각 객체(o)에 대한 샷(s) 내의 객체 빈도수(of(o,s))와 해당 객체에 대한 샷 빈도수의 역수(isf(o,V))를 곱해서 구한 값 ($$ofisf(o,s,V) = of(o,s) * isf(o,V)$$)이 사전 설정된 문턱값 이상이면, 해당 객체(o)를 해당 샷(s)에서의 중요 객체로 선택하는 것이 바람직하다.

[0012] 본 발명의 제2 특징에 따른 샷 단위 중요 객체 추출 장치는, 외부로부터 동영상을 입력받는 영상 입력부; 상기 영상 입력부로 입력된 동영상을 샷 단위로 분할하는 샷 단위 검출부; 상기 동영상을 구성하는 프레임들로부터 객체들을 검출하는 객체 검출부; 각 샷에 포함된 객체들에 대한 샷 내의 객체 빈도수(Object Frequency; OF)를 획득하는 OF 검출부; 각 객체에 대하여, 상기 동영상 전체에서의 중요도를 획득하는 중요도 검출부; 각 객체에 대하여 구한 샷 내의 객체 빈도수와 동영상 전체에서의 중요도를 이용하여, 각 샷에 대한 중요 객체를 추출하는 중요 객체 검출부;를 구비하여, 동영상을 구성하는 각 샷의 중요 객체를 추출하여 제공한다.

[0013] 전술한 제2 특징에 따른 샷 단위 중요 객체 추출 장치에 있어서, 상기 중요도 검출부의 각 객체에 대한 동영상 전체에서의 중요도는, 상기 동영상을 구성하는 모든 샷의 개수(V)와 해당 객체(o)가 검출된 샷들의 개수의 비율을 이용하여 각 객체에 대하여 게산된 샷 빈도수(shot frequency)의 역수(Inverse Shot Frequency;ISF)(isf(o,V))인 것이 바람직하다.

[0014] 전술한 제2 특징에 따른 샷 단위 중요 객체 추출 장치에 있어서, 상기 중요 객체 검출부는, 각 샷별로 구한 각 객체(o)에 대한 샷(s) 내의 객체 빈도수(of(o,s))와 해당 객체에 대한 샷 빈도수의 역수(isf(o,V))를 곱해서 구한 값($$ofisf(o,s,V) = of(o,s) * isf(o,V)$$)이 사전 설정된 문턱값 이상이면, 해당 객체(o)를 해당 샷(s)에서의 중요 객체로 선택하는 것이 바람직하다.

발명의 효과

[0015] 본 발명에 따른 샷내 중요 객체 중요 장치 및 방법은 동영상의 특성을 고려하여, 샷 내 객체 빈도수와 객체의 동영상 내에서의 중요도를 이용하여 , 샷 단위별로 중요 객체를 선택함으로서, 중요 객체를 보다 정확하게 추출할 수 있게 된다.

[0016] 이렇게 추출된 중요 객체들을 미디어 동영상에서의 샷 단위의 객체 메타 데이터(object meta data)로 사용하는 경우, 각 샷에 포함된 특정 장면에서의 상황과 장소 검출의 정확도를 높일 수 있게 되고, 그 결과 샷과 객체의 상황 또는 장소 등과의 연관 관계를 정확하게 나타낼 수 있게 된다.

도면의 간단한 설명

[0017] 도 1은 본 발명의 바람직한 실시예에 따른 샷 단위 중요 객체 추출 장치를 전체적으로 도시한 사시도이다.

도 2는 본 발명의 바람직한 실시예에 따른 샷 단위 중요 객체 추출 장치에 있어서, 객체 검출부에 의해 하나의 프레임으로부터 객체를 검출한 것을 예시적으로 도시한 것이다.

도 3은 본 발명의 바람직한 실시예에 따른 샷 단위 중요 객체 추출 방법을 순차적으로 도시한 흐름도이다.

도 4는 본 발명의 바람직한 실시예에 따른 샷 단위 중요 객체 추출 방법을 구현한 프로그램에 대한 Pseudo code이다.

발명을 실시하기 위한 구체적인 내용

[0018] 본 발명에 따른 샷 단위 중요 객체 추출 장치 및 방법은 샷 내에서의 객체 빈도수와 각 객체에 대한 동영상 전체에서의 중요도를 모두 고려하여, 각 샷에서의 중요 객체들을 선택함으로써, 샷 단위의 중요 객체에 대한 정확도를 향상시킬 수 있게 된다.

[0019] 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시예에 따른 샷 단위 중요 객체 추출 장치의 구성 및 동작에 대하여 구체적으로 설명한다. 본 발명에 따른 샷 단위 중요 객체 추출 장치는 동영상에 대한 다양한 처리가 가능하고 CNN(Convolution Neural Network)과 같은 Deep Learning 학습이 가능한 컴퓨터 등으로 구성될 수 있으며, 상기 장치의 각 부는 프로그램과 같은 소프트웨어로 구성될 수 있다. 따라서, 본 발명에 따른 샷 단위 중요 객체 추출 방법은 상기 샷 단위 중요 객체 추출 장치를 구성하는 각 부들을 구현하는 소프트웨어들로 구성될 수 있다.

[0020] 도 1은 본 발명의 바람직한 실시예에 따른 샷 단위 중요 객체 추출 장치를 전체적으로 도시한 사시도이다. 도 1을 참조하면, 본 발명에 따른 샷 단위 중요 객체 추출 장치(10)는 영상 입력부(100), 샷 단위 검출부(110), 객체 검출부(120), OF 검출부(130), 중요도 검출부(140), 샷단위 중요 객체 검출부(150)를 구비하여, 동영상의 각 샷에 대한 중요 객체들을 추출하는 것을 특징으로 한다. 이하, 전술한 각 구성 요소들에 대하여 구체적으로 설명한다.

[0021] 상기 영상 입력부(100)는 외부로부터 동영상을 입력받는다.

[0022] 상기 샷 단위 검출부(110)는 상기 영상 입력부로 입력된 동영상에서 샷 단위를 검출한다. 샷 단위 검출 모듈은 동영상을 샷 단위로 구분하고 해당 샷의 시작 시간과 종료 시간을 출력한다. 동영상에 대한 샷 단위 검출 모듈은 당업계에 널리 알려진 다양한 방법들 중 하나로 구현될 수 있으며, 간단한 Threshold 기반의 페이드-인(fade-in), 페이드-아웃(fade-out) 감지 방법 또는 컨텐츠 인식을 통한 샷 경계 검출 방법 등 여러 가지의 감지 방법을 통해 샷 경계를 검출하여 제공한다. 본 발명에 따른 샷 단위 검출부는 동영상을 입력값으로 제공하면, 동영상을 shot 단위로 구분하고, 각 shot의 시작 시간과 종료 시간을 출력한다.

[0023] 상기 객체 검출부(120)는 상기 동영상을 구성하는 프레임들에 대하여 각 프레임에 포함된 객체들을 검출한다. 도 2는 본 발명의 바람직한 실시예에 따른 샷 단위 중요 객체 추출 장치에 있어서, 객체 검출부에 의해 하나의 프레임으로부터 객체를 검출한 것을 예시적으로 도시한 것이다. 도 2를 참조하면, 각 프레임들에 대하여 해당 프레임에 포함된 객체들을 검출한다. 동영상의 각 프레임에 있어서, 얼굴 영역이나 객체들을 검출하고, 검출된 객체들을 인식하는 다양한 기술들이 연구되고 있다. 종래의 객체 검출 라이브러리로는 Faster R-CNN(Convolution Neural Network), YOLO, Mask R-CNN 등 Deep-Learning 기반의 방법들이 다양하게 존재하고 있으며, 이들 또한 성능을 높이기 위한 많은 연구가 이루어지고 있다. 이와 같이, 객체 검출 방법은 이미 공지의 기술들이 다양하게 존재하므로, 본 발명에서는 구체적인 설명을 생략한다.

[0024] 상기 OF 검출부(130)는, 각 샷에 대하여, 해당 샷에 포함된 객체들의 객체 빈도수(Object Frequency; OF)를 획득하여 제공한다. 특정 샷(s)에서의 특정 객체(o)에 대한 객체 빈도수(of(o,s))는 수학식 1과 같이 특정 객체(o)가 특정 샷(s)에서 등장하는 횟수(f(o,s))로 구할 수 있다.

수학식 1

[0025] $$of(o,s) = f(o,s)$$

[0026] 전술한 OF는 특정한 객체가 샷내에 얼마나 자주 등장하는지를 나타내는 값이다. 따라서, 특정한 객체의 OF 의 값이 높더라도, 해당 객체가 동영상 전체에서 자주 등장하는 경우, 해당 샷의 중요 객체가 아니라 오히려 흔하게 등장하는 객체인 것을 의미하게 된다. 이와 같이 각 객체가 동영상 전체에서 자주 등장하는 횟수를 샷 빈도수(Shot Frequency; SF)라고 한다.

[0027] 상기 중요도 검출부(140)는, 각 객체에 대하여, 상기 동영상 전체에서의 중요도를 획득하여 제공한다. 각 객체에 대한 동영상 전체에서의 중요도는, 상기 동영상을 구성하는 모든 샷의 개수(V)와 해당 객체(o)가 검출된 샷들의 개수의 비율을 이용하여 각 객체에 대하여 계산된 샷 빈도수(shot frequency)의 역수(Inverse Shot

Frequency;ISF)(isf(o,V))인 것이 바람직하다. 상기 각 객체에 대한 샷 빈도수의 역수(isf(o,V))는 수학식 2에 의해 구해질 수 있다.

수학식 2

$$isf(o,V) = \frac{\log|V|}{|\{s \in V : o \in s\}|}$$

(여기서, $|V|$: 동영상 내의 전체 shot의 개수,
$|\{s \in V : o \in s\}|$: object o가 포함된 shot의 수임.)

[0028]

[0029] 여기서, V의 범위는 동영상 1개부터, 존재하는 모든 동영상까지 검색하고자 하는 범위에 따라 변할 수 있으며, of(o,s)와 isf(o,V)의 계산 방법도 V의 범위 또는 검색하고자 하는 상황에 따라 변할 수 있게 된다.

[0030] 각 객체에 대한 ISF 값은 동영상의 특징에 따라 결정된다. 예를 들어, 사람이라는 객체의 경우, 드라마나 영화와 같은 동영상에서는 자주 등장하는 객체이므로 큰 영향이 없고 따라서 샷 내의 중요 객체가 될 수 없다. 한편, 자연 다큐멘터리와 같은 동영상에서는 사람은 자주 등장하지 않는 객체이므로, 자연 다큐멘터리에서의 사람이라는 객체는 전문가의 의견이 등장하는 장면이므로 중요 상황일 가능성이 높은 것이다.

[0031] 상기 샷 단위 중요 객체 검출부(150)는, OF 검출부에 의하여 각 객체에 대하여 구한 샷 내의 객체 빈도수와 상기 중요도 검출부에 의하여 각 객체에 대하여 구한 동영상 전체에서의 중요도를 이용하여, 각 샷에 대한 중요 객체들을 추출하게 된다.

[0032] 상기 샷 단위 중요 객체 검출부는, 수학식 3과 같이 각 샷별로 구한 각 객체(o)에 대한 샷(s) 내의 객체 빈도수(of(o,s))와 해당 객체에 대한 샷 빈도수의 역수(isf(o,V))를 곱해서 구한 값이 사전 설정된 문턱값 이상이면, 해당 객체(o)를 해당 샷(s)에서의 중요 객체로 선택하게 된다.

수학식 3

[0033]
$$ofisf(o,s,V) = of(o,s) * isf(o,V)$$

[0034] 따라서, 상기 샷 단위 중요 객체 검출부는, 각 샷에 대하여, 해당 샷에 포함된 객체들 중 ofisf(o,s,V)가 사전 설정된 문턱값 이상인 객체들을 중요 객체들로 선택하여 제공하게 된다.

[0036] 이하, 전술한 샷 단위 중요 객체 추출 장치를 영상 처리 장치내의 프로세서에 의해 구현한 샷 단위 중요 객체 추출 방법에 대하여 구체적으로 설명한다.

[0037] 도 3은 본 발명의 바람직한 실시예에 따른 샷 단위 중요 객체 추출 방법을 순차적으로 도시한 흐름도이다. 도 4는 본 발명의 바람직한 실시예에 따른 샷 단위 중요 객체 추출 방법을 구현한 프로그램에 대한 Pseudo code이다.

[0038] 도 3 및 도 4를 참조하면, 본 발명에 따른 샷 단위 중요 객체 추출 방법은, 먼저 입력된 동영상을 샷 단위로 분할하며(단계 300), 동영상을 구성하는 각 프레임들에 포함된 객체들을 검출한다(단계 310).

[0039] 다음, 각 샷에 대하여, 해당 샷에 포함된 객체들의 객체 빈도수(Object Frequency; OF)를 획득한다(단계 320). 특정 샷(s)에서의 특정 객체(o)에 대한 객체 빈도수(of(o,s))는 수학식 1과 같이 특정 객체(o)가 특정 샷(s)에서 등장하는 횟수(f(o,s))로 구할 수 있다.

[0040] 다음, 각 객체에 대하여, 상기 동영상 전체에서의 중요도를 획득한다(단계 330). 각 객체에 대한 동영상 전체에서의 중요도는, 상기 동영상을 구성하는 모든 샷의 개수(V)와 해당 객체(o)가 검출된 샷들의 개수의 비율을 이용하여 각 객체에 대하여 계산된 샷 빈도수(shot frequency)의 역수(Inverse Shot Frequency;ISF)(isf(o,V))인 것이 바람직하다. 상기 각 객체에 대한 샷 빈도수의 역수(isf(o,V))는 수학식 2에 의해 구해질 수 있다.

[0041] 다음, 각 객체에 대하여 구한 샷 내의 객체 빈도수와 동영상 전체에서의 중요도를 이용하여, 각 샷 내의 중요 객체를 추출한다(단계 340). 단계 340은, 수학식 3과 같이 각 샷별로 구한 각 객체(o)에 대한 샷(s) 내의 객체 빈도수(of(o,s))와 해당 객체에 대한 샷 빈도수의 역수(isf(o,V))를 곱해서 구한 값이 사전 설정된 문턱값 이상이면, 해당 객체(o)를 해당 샷(s)에서의 중요 객체로 선택하게 된다.

[0043] 이상에서 본 발명에 대하여 그 바람직한 실시예를 중심으로 설명하였으나, 이는 단지 예시일 뿐 본 발명을 한정하는 것이 아니며, 본 발명이 속하는 분야의 통상의 지식을 가진 자라면 본 발명의 본질적인 특성을 벗어나지 않는 범위에서 이상에 예시되지 않은 여러 가지의 변형과 응용이 가능함을 알 수 있을 것이다. 그리고, 이러한 변형과 응용에 관계된 차이점들은 첨부된 청구 범위에서 규정하는 본 발명의 범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

[0044] 10 : 샷 단위 중요 객체 추출 장치

100 : 영상 입력부

110 : 샷 단위 검출부

120 : 객체 검출부

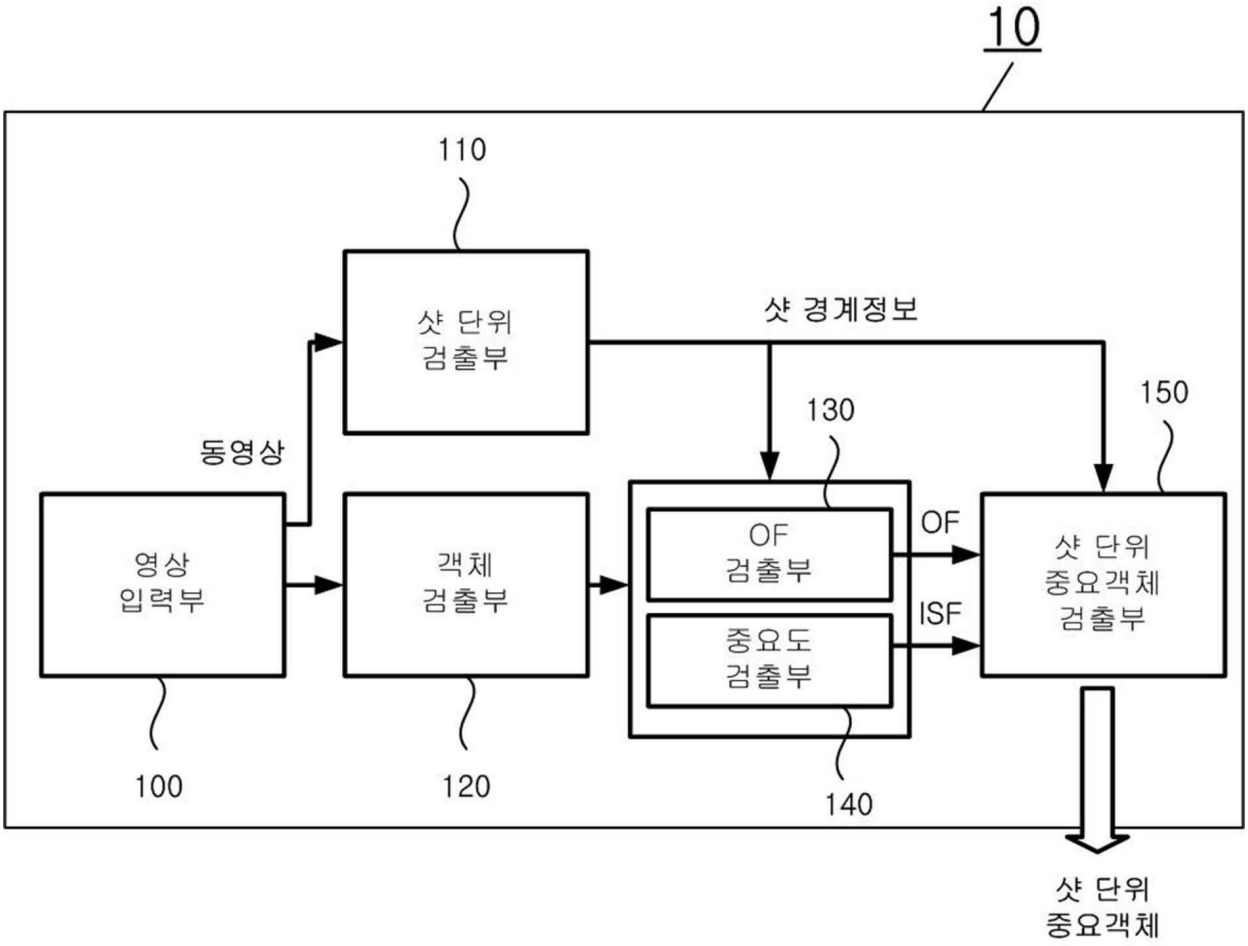
130 : OF 검출부

140 : 중요도 검출부

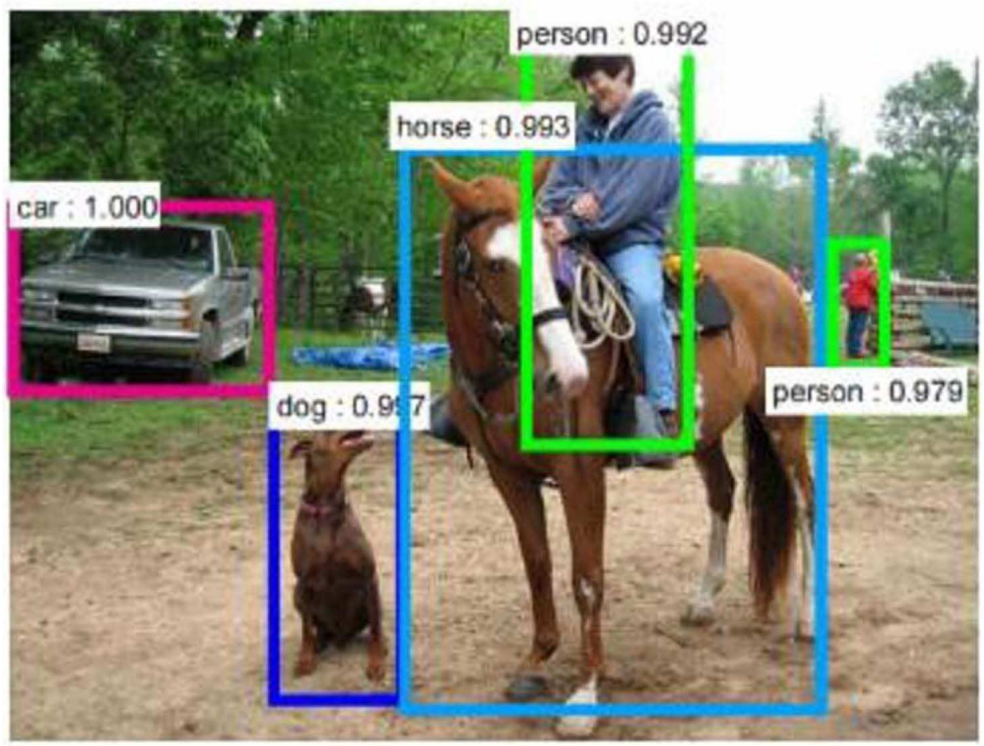
150 : 샷단위 중요 객체 검출부

도면

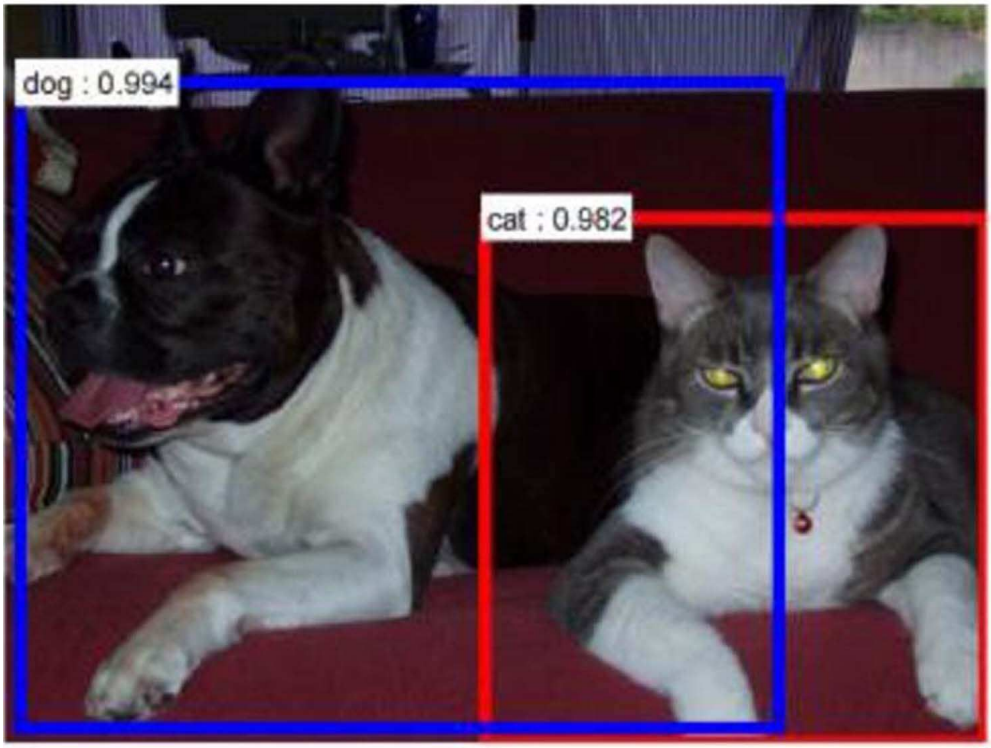
도면1


10
110
샷 단위
검출부
샷 경계정보
130
150
동영상
영상
입력부
객체
검출부
OF
검출부
중요도
검출부
OF
ISF
샷 단위
중요객체
검출부
100
120
140
샷 단위
중요객체

도면2

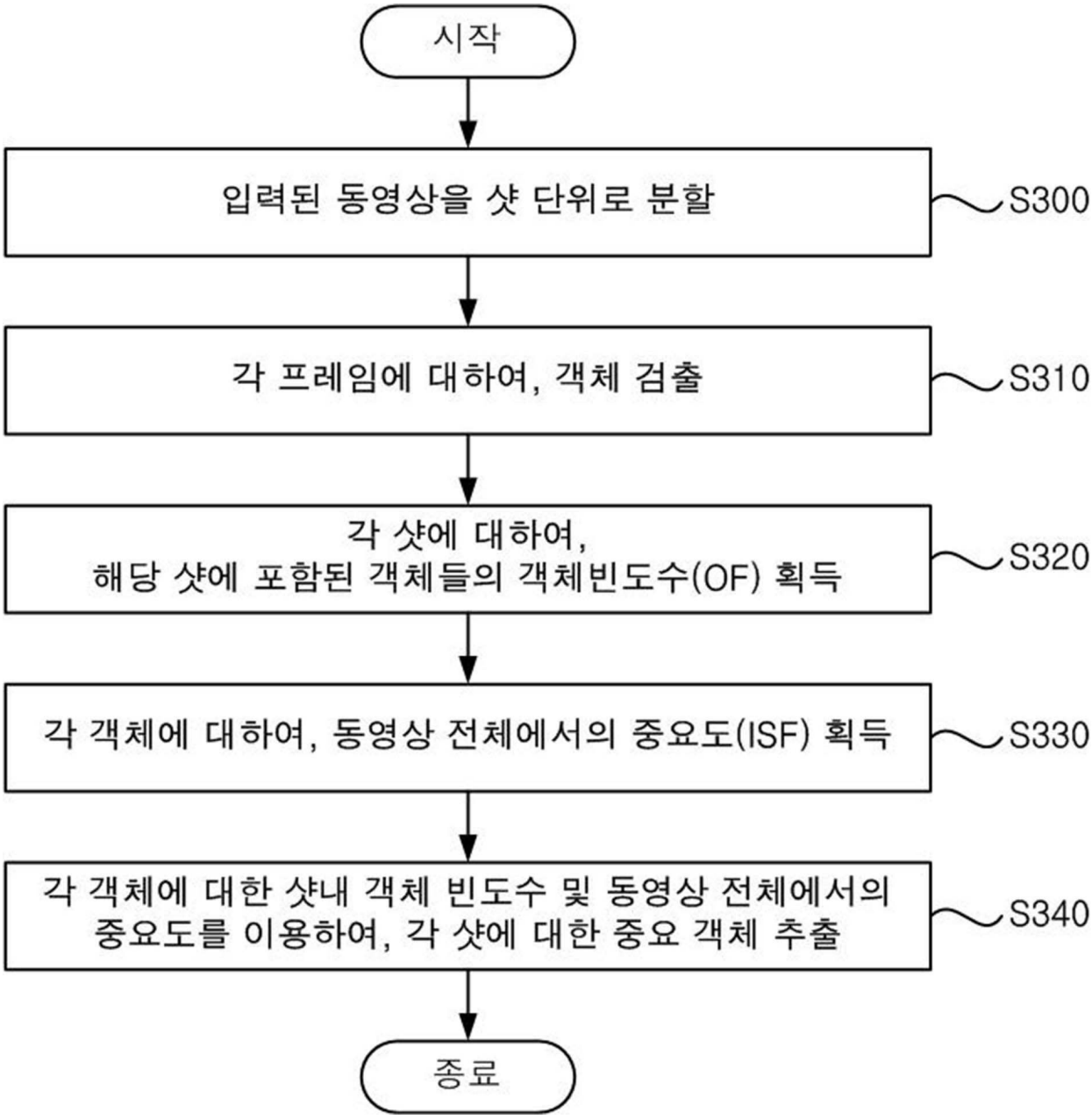

person : 0.992
horse : 0.993
car : 1.000
dog : 0.9
person : 0.979

(a)


dog : 0.994
cat : 0.982

(b)

도면3


시작
입력된 동영상을 샷 단위로 분할
S300
각 프레임에 대하여, 객체 검출
S310
각 샷에 대하여,
해당 샷에 포함된 객체들의 객체빈도수(OF) 획득
S320
각 객체에 대하여, 동영상 전체에서의 중요도(ISF) 획득
S330
각 객체에 대한 샷내 객체 빈도수 및 동영상 전체에서의
중요도를 이용하여, 각 샷에 대한 중요 객체 추출
S340
종료

도면4

1 : **Input :** video

2 : **for all** frame $\in$ video **do**

3 : Objects := Object detection(frame)

4 : shots := Shot Change Detector(video)

5 : **for all** $s \in$ shots **do**

6 : $of(o,s) :=$ number of objects in s $(o \in$ Objects$)$

7 : $isf(o,V) = \frac{\log|V|}{|\{s \in V : o \in s\}|}$ $(V := video)$

8 : $ofisf(o,s,V) := of(o,s) * isf(o,V)$

9 : **if** $ofisf(o,s,V) >$ threshold **do**

10 : $\text{Output}_{\text{shot}} := o$

11 : **End for return** Output