(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년11월08일
(11) 등록번호 10-1916874
(24) 등록일자 2018년11월02일

(51) 국제특허분류(Int. Cl.)
H04N 21/81 (2011.01) *G06F 17/27* (2006.01)
G06F 17/30 (2006.01) *H04N 21/8549* (2011.01)
(52) CPC특허분류
H04N 21/8133 (2013.01)
G06F 17/2755 (2013.01)
(21) 출원번호 10-2017-0135538
(22) 출원일자 2017년10월19일
심사청구일자 2017년10월19일
(56) 선행기술조사문헌
JP2004159192 A*
JP5896900 B2*
KR1020170004154 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
충남대학교산학협력단
대전광역시 유성구 대학로 99 (궁동, 충남대학교)
(72) 발명자
이영석
대전광역시 유성구 죽동로 321, 102동 103호(죽동, 금성백조예미지)
남현규
대구광역시 달서구 장기로 242 우방드림시티아파트 108동 303호
(74) 대리인
특허법인 신우

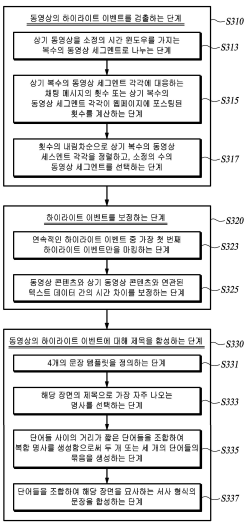
전체 청구항 수 : 총 5 항 심사관 : 홍기완

(54) 발명의 명칭 **자동으로 동영상 하이라이트 영상의 제목을 생성하는 방법, 장치 및 컴퓨터 판독가능 기록 매체**

(57) 요 약

통신망을 통해 사용자 단말과 통신하는 서버가 제공된다. 본 발명에 따른 서버는, 통신망을 통해 사용자 단말과 통신하는, 동영상 컨텐츠의 제목을 자동으로 생성하는 서버로서, 상기 통신망을 통해 사용자 단말과 통신하도록 구성되는 통신부; 상기 사용자 단말로부터 입력되는 텍스트 정보를 수신하여 상기 통신부를 통하여 저장하도록 구성된 저장부; 및 상기 저장된 텍스트 정보를 분석하여 동영상의 제목을 생성하도록 구성되는 제어부를 포함하고, 상기 제어부는 상기 저장된 텍스트 정보를 4개의 템플릿 중 하나로 재구성하도록 구성되며 - 상기 4개의 템플릿은 두 개 또는 세 개의 단어의 묶음(bag), 구절, 또는 주어, 동사, (목적어)가 포함되는 완벽한 문장임-; 상기 제어부는 상기 텍스트 정보의 형태소 분석 후에, 각각의 토큰에 단어 클래스로 태깅하고, 가장 자주 나오는 단어를 선택하여 상기 동영상의 제목을 생성하도록 구성되는 동영상 컨텐츠의 제목을 자동 생성 서버이다.

대 표 도 - 도3



(52) CPC특허분류
G06F 17/3082 (2013.01)
H04N 21/8549 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 NRF-2016R1D1A1A09916326
부처명 소프트웨어 연구소
연구관리전문기관 한국연구재단
연구사업명 일반연구자지원
연구과제명 인터넷 커뮤니티 빅데이터 활용 분석 서비스
기 여 율 1/2
주관기관 충남대학교
연구기간 2016.11.01 ~ 2017.10.31
이 발명을 지원한 국가연구개발사업
과제고유번호 IITP-2017-2016-0-00304
부처명 소프트웨어연구소
연구관리전문기관 ITRC
연구사업명 ITRC
연구과제명 금융 공격침해 이상징후 탐지 및 예측 기술 연구
기 여 율 1/2
주관기관 충남대학교
연구기간 2017.01.01 ~ 2017.12.31

명 세 서

청구범위

청구항 1

통신망을 통해 사용자 단말과 통신하는, 동영상 컨텐츠의 제목을 자동으로 생성하는 서버로서,

상기 통신망을 통해 상기 사용자 단말로부터, 상기 동영상 컨텐츠와 연관되는 적어도 하나의 채팅 메시지를 수신하도록 구성되는 통신부;

상기 사용자 단말로부터 수신한 채팅 메시지를 저장하도록 구성된 저장부; 및

상기 저장된 채팅 메시지의 텍스트 정보를 분석하여 동영상의 제목을 생성하도록 구성되는 제어부를 포함하고,

상기 제어부는

상기 동영상 컨텐츠를 소정의 시간 윈도우를 가지는 복수의 동영상 세그멘트로 나누고,

상기 저장된 채팅 메시지를 분석하여, 상기 복수의 동영상 세그멘트 각각에 대응하는 채팅 메시지의 횟수를 결정하고,

상기 복수의 동영상 세그멘트 중 채팅 메시지의 횟수가 높은 순서로 소정의 수의 동영상 세그멘트들을 선택하여, 하이라이트 이벤트들을 검출하고,

상기 하이라이트 이벤트들 중 인접한 시간 윈도우에 연속하는 하이라이트 이벤트들 중 가장 첫번째 하이라이트 이벤트만을 마킹하여, 상기 연속하는 하이라이트 이벤트들을 하나의 하이라이트 이벤트로 보정하고,

상기 저장된 채팅 메시지 중 상기 하이라이트 이벤트와 연관된 채팅 메시지를 형태소 분석하여, 상기 하이라이트 이벤트와 연관된 채팅 메시지에서 가장 자주 나오는 명사에 기초하여 상기 하이라이트 이벤트의 제목을 생성하도록 구성되는,

동영상 컨텐츠의 제목을 자동으로 생성하는 서버.

청구항 2

통신망을 통해 사용자 단말과 통신하며 서버에서 수행되는 동영상 컨텐츠의 제목을 자동으로 생성하는 방법으로서,

상기 사용자 단말로부터, 상기 동영상 컨텐츠와 연관되는 적어도 하나의 채팅 메시지를 수신하는 단계;

상기 동영상 컨텐츠를 소정의 시간 윈도우를 가지는 복수의 동영상 세그멘트로 나누는 단계;

상기 사용자 단말로부터 수신한 적어도 하나의 채팅 메시지를 분석하여, 상기 복수의 동영상 세그멘트 각각에 대응하는 채팅 메시지의 횟수를 결정하는 단계;

상기 복수의 동영상 세그멘트 중 채팅 메시지의 횟수가 높은 순서로 소정의 수의 동영상 세그멘트들을 선택하여, 하이라이트 이벤트들을 검출하는 단계;

상기 하이라이트 이벤트들 중 인접한 시간 윈도우에 연속하는 하이라이트 이벤트들 중 가장 첫번째 하이라이트 이벤트만을 마킹하여, 상기 연속하는 하이라이트 이벤트들을 하나의 하이라이트 이벤트로 보정하는 단계; 및

상기 사용자 단말로부터 전송되는 적어도 하나의 채팅 메시지 중 상기 하이라이트 이벤트와 연관된 채팅 메시지를 형태소 분석하여, 상기 하이라이트 이벤트와 연관된 채팅 메시지에서 가장 자주 나오는 명사에 기초하여 상기 하이라이트 이벤트의 제목을 생성하는 단계

를 포함하는, 동영상 컨텐츠의 제목을 자동으로 생성하는 방법.

청구항 3

삭제

청구항 4

삭제

청구항 5

제2항에 있어서,

상기 동영상 컨텐츠의 재생 시간과 상기 동영상 컨텐츠와 연관된 채팅 메시지 간의 시간 차이를 보정하는 단계를 포함하는 동영상 컨텐츠의 제목을 자동으로 생성하는 방법.

청구항 6

제2항에 있어서,

상기 하이라이트 이벤트에 대해 제목을 생성하는 단계는

4개의 문장 템플릿을 정의하는 단계 - 상기 4개의 문장 템플릿은 두 개 또는 세 개의 단어 묶음(bag)이거나, 구절, 또는 주어, 동사, (목적어)가 포함되는 완벽한 문장임 -;

상기 하이라이트 이벤트와 연관된 채팅 메시지를 형태소 분석하여 각각의 토큰에 단어 클래스로 태깅하고, 상기 태깅된 토큰에 워드2벡 모델(Word2Vec model)을 적용하여 단어들의 묶음을 생성하는 단계; 및

상기 생성된 묶음의 단어들을 조합하여 제목을 생성하는 단계;

를 포함하는 동영상 컨텐츠의 제목을 자동으로 생성하는 방법.

청구항 7

하나 이상의 명령어가 수록된 컴퓨터 판독가능 기록 매체로서, 상기 하나 이상의 명령어는 컴퓨터에 의해 실행될 경우 상기 컴퓨터로 하여금 제2항 및 제5항 내지 제6항 중 어느 한 항에 따른 방법을 수행하도록 하는, 컴퓨터 판독가능 기록 매체.

발명의 설명

기 술 분 야

[0001] 본 발명은, 동영상 비디오 클립을 생성하는 방법에 관한 것이며, 보다 구체적으로는 사용자의 관여 없이 자동으로 제목을 포함하는 비디오 클립을 생성하는 방법 및 장치 등에 관한 것이다.

배 경 기 술

[0002] 비디오 컨텐츠의 양이 폭발적으로 증가하고 다양한 매체에서 비디오 컨텐츠를 제공함에 따라 소비자는 선호하는 컨텐츠를 선택하여 시청하는 경향이 점점 더 강해지고 있다. 또한, 짧은 시간에 원하는 영상을 선택하여 시청하기 위해서 전체 동영상 컨텐츠 중에서도 원하는 장면을 강조 표시하는 짧은 비디오 클립을 선호하는 경향 또한 증가하고 있다.

[0003] 통상적으로 짧은 비디오 클립을 작성하기 위해서는 편집자가 전체 동영상 중 하이라이트 부분을 직접 선택하여 작성하는 것이 일반적이다. 한편, 대한민국 특허출원번호 제10-2017-0029933호에는, 댓글 또는 해당 동영상에 대한 누적정보(반복 재생 횟수, 캡쳐 횟수, 공유 횟수 등)를 통해서 자동으로 하이라이트 영상을 편집하는 방법이 개시되어 있다. 그러나, 이러한 시스템은, 영상 자체를 편집하는 방법에 관한 것으로, 영상에 적절한 제목을 부여하는 방법에 관한 것은 아니다. 또한, 이러한 시스템은 동영상 검색을 위한 검색어로서 사용자가 스스로 검색 정보를 별도 입력해야 하고, 그렇지 않은 경우 검색이 어려워 수많은 동영상 중 식별되어 발견되기 어렵다는 문제를 내포하고 있다.

선행기술문헌

특허문헌

[0004] (특허문헌 0001) 대한민국 특허출원번호 제10-2017-0029933호

발명의 내용

해결하려는 과제

[0005] 따라서, 본 개시는 사용자의 관여 없이 자동으로 비디오 클립을 생성하는 방법 및 장치를 제공할 수 있다.

[0006] 본 개시는 자동으로 생성된 비디오 클립에 제목을 부여하는 방법 및 장치를 제공할 수 있다.

과제의 해결 수단

[0007] 본 발명의 일 특징에 의하면, 통신망을 통해 사용자 단말과 통신하는, 동영상 컨텐츠의 제목을 자동으로 생성하는 서버가 제공된다. 상기 서버는 상기 통신망을 통해 사용자 단말과 통신하도록 구성되는 통신부; 상기 사용자 단말로부터 입력되는 텍스트 정보를 수신하여 상기 통신부를 통하여 저장하도록 구성된 저장부; 및 상기 저장된 텍스트 정보를 분석하여 동영상의 제목을 생성하도록 구성되는 제어부를 포함하고, 상기 제어부는 상기 저장된 텍스트 정보를 4개의 템플릿 중 하나로 재구성하도록 구성되며 - 상기 4개의 템플릿은 두 개 또는 세 개의 단어의 묶음(bag), 구절, 또는 주어, 동사, (목적어)가 포함되는 완벽한 문장임-; 상기 제어부는 상기 텍스트 정보의 형태소 분석 후에, 각각의 토큰에 단어 클래스로 태깅하고, 가장 자주 나오는 단어를 선택하여 상기 동영상의 제목을 생성하도록 구성된다.

[0008] 본 발명의 다른 특징에 의하면, 통신망을 통해 사용자 단말과 통신하며 서버에서 수행되는 동영상 컨텐츠의 제목을 자동으로 생성하는 방법이 제공된다. 상기 방법은 상기 동영상 컨텐츠와 연관되는, 상기 사용자 단말로부터 전송되는 텍스트를 수신하는 단계; 상기 동영상 컨텐츠의 하이라이트 이벤트를 검출하는 단계; 상기 검출된 하이라이트 이벤트를 보정하는 단계; 및 상기 동영상 컨텐츠의 상기 하이라이트 이벤트에 대해 제목을 합성하는 단계를 포함하는 동영상 컨텐츠의 제목을 자동으로 생성하는 방법이 제공된다.

[0009] 일 실시예에 있어서, 상기 동영상 컨텐츠의 하이라이트 이벤트를 검출하는 단계는 상기 동영상 컨텐츠를 소정의 시간 윈도우를 가지는 복수의 동영상 세그멘트로 나누는 단계; 상기 복수의 동영상 세그멘트 각각에 대응하는 채팅 메시지의 횟수 또는 상기 복수의 동영상 세그멘트 각각이 웹페이지에 포스팅된 횟수를 계산하는 단계; 및 상기 횟수의 내림차순으로 상기 복수의 동영상 세그멘트 각각을 정렬하고, 소정의 수의 동영상 세그멘트를 선택하는 단계를 포함할 수 있다.

[0010] 일 실시예에 있어서, 상기 검출된 하이라이트 이벤트를 보정하는 단계는 인접한 시간 윈도우에 연속적인 하이라이트 이벤트가 속하고 이는 동일한 장면(Scene)에 속하는 경우를 제거하기 위해 가장 첫 번째 하이라이트 이벤트만을 마킹함으로써 중복된 하이라이트 이벤트를 방지하는 단계를 포함할 수 있다.

[0011] 일 실시예에 있어서, 상기 검출된 하이라이트 이벤트를 보정하는 단계는상기 동영상 콘텐츠의 재생 시간과 상기 동영상 콘텐츠와 연관된 텍스트 데이터 간의 시간 차이를 보정하는 단계를 포함할 수 있다.

[0012] 일 실시예에 있어서, 상기 동영상의 하이라이트 이벤트에 대해 제목을 합성하는 단계는 4개의 문장 템플릿을 정의하는 단계 - 상기 4개의 문장 템플릿은 두 개 또는 세 개의 단어 묶음(bag)이거나, 구절, 또는 주어, 동사, (목적어)가 포함되는 완벽한 문장임 -; 상기 텍스트를 형태소 분석을 하여 각각의 토큰에 단어 클래스로 태깅하는 단계; 상기 태깅된 토큰에 워드2벡 모델(Word2Vec model)을 적용하여 단어들의 묶음을 생성하는 단계; 및 상기 계산된 거리가 짧은 단어들의 두 개 또는 세 개의 단어들을 조합하여 제목을 생성하는 단계를 포함할 수 있다.

[0013] 본 발명의 또 다른 특징에 의하면, 하나 이상의 명령어가 수록된 컴퓨터 판독가능 기록 매체로서, 하나 이상의 명령어는 컴퓨터에 의해 실행될 경우 컴퓨터로 하여금 전술한 방법들 중 어느 하나의 방법을 수행하도록 하는, 컴퓨터 판독가능 기록 매체가 제공된다.

발명의 효과

[0014] 본 개시에 의하면, 사용자의 입력에 의존하지 않고, 자동으로 동영상 하이라이트 비디오 클립을 생성할 수 있는 방법 및 장치 등이 제공될 수 있다.

[0015] 본 개시에 의하면, 생성된 비디오 클립에 사용자의 개입 없이 자동으로 적절한 제목을 부여할 수 있다.

도면의 간단한 설명

[0016] 도 1은, 본 발명의 일 실시예에 따른 네트워크 환경의 일 예를 도시한 도면이다.

도 2는, 도 1의 서버의 예시적 구성을 개략적으로 도시한 도면이다.

도 3은, 본 발명의 일 실시예에 따른, 서버 상의 게시판에 제목과 함께 동영상 하이라이트를 업로드 하기 위해 서버 상에서 수행되는 동작을 보여주는 개략적 동작 흐름도이다.

도 4은 도 3에서 설명한 동작에 따라 대중 상호 웹사이트(mass interaction website: 유투브, 트위치, 네이버 스포츠 도는 디시인사이트)에서 제목과 함께 동영상 하이라이트를 생성하는 예시적인 방법을 도시하는 도면이다.

도 5는 각 장르에서 제목과 함께 동영상 하이라이트 이벤트를 제공하는 일 실시예를 도시한 도면이다.

도 6은 각각의 비디오 카테고리에 대해서 4 가지의 제목 타입의 평균 점수를 나타내는 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0017] 이하, 첨부 도면을 참조하여 본 발명의 실시예에 관하여 상세히 설명한다. 이하에서는, 본 발명의 요지를 불필요하게 흐릴 우려가 있다고 판단되는 경우, 이미 공지된 기능 및 구성에 관한 구체적인 설명을 생략한다. 또한, 이하에서 설명하는 내용은 어디까지나 본 발명의 일 실시예에 관한 것일 뿐 본 발명이 이로써 제한되는 것은 아님을 알아야 한다.

[0018] 본 명세서에서 사용되는 용어는 단지 특정한 실시예를 설명하기 위해 사용되는 것으로 본 발명을 한정하려는 의도에서 사용된 것이 아니다. 예를 들면, 단수로 표현된 구성요소는 문맥상 명백하게 단수만을 의미하지 않는다면 복수의 구성요소를 포함하는 개념으로 이해되어야 한다. 또한, 본 발명의 명세서에서, '포함하다' 또는 '가지다' 등의 용어는 명세서 상에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것일 뿐이고, 이러한 용어의 사용에 의해 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 배제하려는 것은 아니다.

[0019] 본 명세서에 기재된 실시예에 있어서 '블록' 또는 '부'는 적어도 하나의 기능이나 동작을 수행하는 기능적 부분을 의미하며, 하드웨어 또는 소프트웨어로 구현되거나 하드웨어와 소프트웨어의 결합으로 구현될 수 있다. 또한, 복수의 '블록' 또는 '부'는, 특정한 하드웨어로 구현될 필요가 있는 '블록' 또는 '부'를 제외하고는, 적어도 하나의 소프트웨어 모듈로 일체화되어 적어도 하나의 프로세서로 구현될 수 있다.

[0020] 덧붙여, 달리 정의되지 않는 한 기술적 또는 과학적인 용어를 포함하여, 본 명세서에서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의된 용어들은, 관련 기술의 문맥상 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 발명의 명세서에서 명백하게 달리 정의하지 않는 한 과도하게 제한 또는 확장하여 해석되지 않는다는 점을 알아야 한다.

[0021] 이하, 첨부된 도면을 참조하여, 본 발명의 실시예에 대해 구체적으로 설명하기로 한다.

[0022] 도 1은, 본 발명의 일 실시예에 따른, 네트워크 환경(100)의 구성을 개략적으로 도시한 도면이다. 도시된 바에 의하면, 네트워크 환경(100)은, 하나 이상의 사용자 단말(102a-102n), 통신망(104), 서버(106)를 포함한다. 이러한 도 1은 발명의 설명을 위한 일례로 사용자 단말의 수나 서버의 수가 도 1과 같이 한정되는 것은 아니다.

[0023] 본 발명의 일 실시예에 의하면, 복수의 사용자 단말(102a-102n)은, 사용자가 이용하는 각종 고정 또는 휴대용 통신 단말일 수 있다. 복수의 사용자 단말(102a-102n) 각각은, 예컨대 데스크탑, 랩탑, 스마트 폰, 태블릿 PC, 기타 다양한 통신 단말일 수 있다. 본 발명의 일 실시예에 의하면, 사용자 단말(102a-102n)은, 통신망(104)을 통해 서버(106)와 통신하고 필요한 정보를 송수신할 수 있다. 본 발명의 일 실시예에 의하면, 복수의 사용자 단말(102a-102n) 각각은, 통신망(104)을 통해 서버(106)에 접속하여 서버(106)에 의해서 운영되는 게시판의 각 동영상을 검색하거나 해당 게시판에 새로운 동영상을 업로드할 수 있다.

[0024] 본 발명의 일 실시예에 의하면, 통신망(104)은, 임의의 유선 또는 무선 통신망일 수 있으며, 예컨대 WiFi망, LAN망, WAN망, 인터넷망 등을 포함할 수 있으며, 본 발명이 이로써 제한되는 것은 아니다.

[0025] 본 발명의 일 실시예에 의하면, 서버(106)는, 소정의 목적에 따라, 통신망(104)을 통해 복수의 사용자 단말(102a-102n)과 통신하며, 명령, 코드, 파일, 컨텐츠, 서비스 등을 제공하는 컴퓨터 장치 또는 복수의 컴퓨터 장치들로 구현될 수 있다.

[0026] 일례로, 서버는 네트워크(140)를 통해 접속한 사용자 단말(102a-102n)로회원들 어플리케이션의 설치를 위한 파일을 제공할 수 있다. 이 경우 사용자 단말(102a-102n)은 서버(106)로부터 제공된 파일을 이용하여 어플리케이션을 설치할 수 있다. 또한, 사용자 단말(102a-102n)이 포함하는 운영체제(Operating System: OS) 및 적어도 하나의 프로그램의 제어에 따라 서버(106)에 접속하여 서버(106)가 제공하는 서비스나 컨텐츠를 제공받을 수 있다. 예를 들어, 사용자 단말(102a-102n)이 어플리케이션의 제어에 따라 네트워크(140)를 통해 서비스 요청 메시지를 서버(106)로 전송하며, 서버(106)는 서비스 요청 메시지에 대응하는 코드를 사용자 단말(102a-102n)로 전송할 수 있고, 사용자 단말(102a-102n)은 어플리케이션의 제어에 따라 코드에 따른 화면을 구성하여 표시함으로써 사용자에게 컨텐츠를 제공할 수 있다.

[0027] 도 2는 본 발명의 일 실시예에 따라 도 1의 서버의 예시적 구성을 개략적으로 도시한 도면이다. 도시된 바에 의하면, 서버(102)는 제어부(204), 저장부(206), 및 통신부(208)를 포함할 수 있다.

[0028] 본 발명의 일 실시예에 의하면, 제어부(204)는 저장부(206) 및 통신부(208)에 연결되어 이들을 제어하기 위해 다양한 컴퓨터 프로그램을 실행하고 그 동작을 제어하도록 구성될 수 있다. 제어부(204)는 한 구성 요소로부터 데이터를 수신하여 다른 구성 요소에 의해 사용되기에 적합한 포맷으로 데이터 포맷팅을 하도록 구성될 수 있다. 제어부(204)는 저장부(206)와 통신 가능한 방식으로 연결될 수 있다. 다양한 실시예에서, 제어부(204)는, 예컨대 통신부(208)과 같은 다른 구성 요소와 일체로 통합 구성될 수 있다.

[0029] 제어부(204)는 통신부(208)를 통해 텍스트 정보를 수신하고, 수신된 텍스트 정보를 분석하여 동영상의 제목을 생성할 수 있다. 본 발명의 일 실시예에 의하면, 제어부(204)는, 저장부(206)에 저장된 동영상 컨텐츠 및 연관된 텍스트 정보를 이용하여 동영상 컨텐츠의 하이라이트 이벤트를 검출하고, 동영상 비디오 클립에 적절한 제목을 생성할 수 있다.

[0030] 본 발명의 일 실시예에 의하면, 제어부(204)는, 텍스트를 4개의 문장 템플릿 중 하나로 구분할 수 있다. 일 실시예에서, 동영상의 제목은 두 개 또는 세 개의 단어의 묶음(bag), 구절, 또는 주어, 동사, (목적어)가 포함되는 완벽한 문장일 수 있다. 제어부(204)는, 채팅 메시지 또는 댓글의 형태소 분석 후에, 각각의 토큰에 단어 클래스로 태깅을 할 수 있다.

[0031] 일 실시예에서, 제어부(204)는 해당 장면의 제목으로 가장 자주 나오는 단어를 선택할 수 있다. 왜냐하면 가장 많이 나오는 단어는 해당 장면의 상황을 효과적으로 설명할 수 있기 때문이다. 일 실시예에서, 제어부(204)는 주인공의 이름을 해당 장면의 제목으로 선택할 수 있다. 일 실시예에서, 제어부(204)는 비속어 등의 부적절한 단어를 사전에 제거할 수 있다.

[0032] 일 실시예에서, 제어부(204)는 단어들 사이의 거리가 짧은 단어들을 조합하여 복합 명사를 생성함으로써 두 개 또는 세 개의 단어들의 묶음을 생성할 수 있다. 일 실시예에서, 제어부(204)는 워드2벡 모델(Word2Vec model)을 사용하여 단어들의 묶음을 생성할 수 있다.

[0033] 일 실시예에서, 제어부(204)는, 두 개 또는 세 개의 단어가 주어졌을 때, 단어들을 조합하여 해당 장면을 묘사하는 서사 형식의 문장을 합성할 수 있다. 장면 제목 문장을 합성하기 위해, 제어부(204)는 적절한 명사를 선택하고 문장에 들어갈 수 있는 객체와 동사를 고려할 수 있다. 문장의 주어로 주로 언급된 명사와 다른 부분을 추출하고, 명사와 문장의 다른 부분의 유사도를 계산할 수 있다. 제어부(104)는 가장 짧은 단어 벡터와 토큰의 쌍을 선택하여 문장을 완성할 수 있다.

[0034] 저장부(206)는, 도 1의 통신망(104) 및 통신부(208)를 통하여, 단말(102)로부터 전송되어 온 동영상 정보 등을 저장 및 갱신할 수 있다. 본 발명의 일 실시예에 의하면, 저장부(206)는 또한, 도 1의 통신망(102) 및 통신부(208)를 통하여, 단말(102a-102c) 각각으로부터 실시간으로 전송되어 온 동영상 관련 채팅 메시지, 댓글, 비디오 클립 등의 정보를 저장 및 갱신할 수 있다.

[0035] 저장부(206)는 또한, 도 1의 통신망(102) 및 통신부(208)를 통하여, 외부 컨텐츠 제공 서버(도시되지 않음)로부터 전송되어 온 동영상 컨텐츠를 저장 및 갱신할 수 있다.

[0036] 저장부(206)는 플래시 메모리 타입(flash memory type), 하드 디스크 타입(hard disk type), 멀티미디어 카드(MultiMedia Card: MMC), 카드 타입의 메모리(예를 들어, SD(Secure Digital) 카드 또는 XD(eXtream Digital)

카드 등), RAM(Random Access Memory), SRAM(Static Random Access Memory), ROM(Read-Only Memory), EEPROM(Electrically Erasable Programmable Read-Only Memory), PROM(Programmable Read-Only Memory), 자기 메모리, 자기 디스크 및 광 디스크 중 어느 하나의 저장 매체로 구현될 수 있으나, 당업자라면 저장부(206)의 구현 형태가 이에 한정되는 것이 아님을 알 수 있을 것이다.

[0037] 본 발명의 일 실시예에 의하면, 통신부(208)는 서버(106)가 통신망(104)을 통해 사용자 단말(102)와 통신할 수 있게 한다. 본 발명의 일 실시예에 의하면, 통신부(208)는 제어부(104)에서 처리된 정보를 정해진 프로토콜에 따라 통신망(104)을 통하여 사용자 단말(102)로 전송할 수 있다. 본 발명의 일 실시예에 의하면, 통신부(208)는, 도 1의 통신망(104)의 프로토콜에 따라, 송수신되는 정보를 적절히 인코딩 또는 디코딩할 수 있다. 본 발명의 일 실시예에 의하면, 통신부(208)는, 사용자 단말(102)로부터 각종 정보를 수신하고 적절한 처리를 행할 수 있다.

[0038] 이상으로 설명한 실시예는, 하드웨어적 측면에서 응용 주문형 집적 회로(Application Specific Integrated Circuits: ASICs), 디지털 신호 처리기(Digital Signal Processors: DSPs), 디지털 신호 처리 소자(Digital Signal Processing Devices: DSPDs), 프로그램 가능 논리 소자(Programmable Logic Devices: PLDs), 현장 프로그램 가능 게이트 어레이(Field-Programmable Gate Arrays: FPGAs), 프로세서(processors), 제어기(controllers), 마이크로 컨트롤러(micro-controllers) 및 마이크로 프로세서 (microprocessors) 중 적어도 하나를 이용하여 구현될 수 있다.

[0039] 절차나 단계 또는 기능을 포함하는 실시예들은 적어도 하나의 기능 또는 동작을 수행하게 하는, 하드웨어 플랫폼(platform) 상에서 실행 가능한 펌웨어 (firmware)/소프트웨어 모듈로 구현될 수 있다. 이 경우, 펌웨어/소프트웨어 모듈은 적절한 프로그램 언어(program language)로 쓰여진 소프트웨어 애플리케이션 (application)에 의해 구현될 수 있다.

[0040] 도 3은, 본 발명의 일 실시예에 따른, 서버 상의 게시판에 제목과 함께 동영상 하이라이트를 업로드 하기 위해 서버 상에서 수행되는 동작을 보여주는 개략적 동작 흐름도이다.

[0041] 동영상의 하이라이트 이벤트를 검출하는 단계로부터 시작한다(S310). 일 실시예에서, 채팅 메시지와 댓글을 분석하여 하이라이트 장면을 찾을 수 있다. 예컨대, 상기 동영상을 소정의 시간 윈도우를 가지는 복수의 동영상 세그멘트로 나누고(S313), 상기 복수의 동영상 세그멘트 각각에 대응하는 채팅 메시지의 횟수 또는 상기 복수의 동영상 세그멘트 각각이 웹페이지에 포스팅된 횟수를 계산한다(S315). 이 후, 횟수의 내림차순으로 상기 복수의 동영상 세그멘트 각각을 정렬하고, 소정의 수의 동영상 세그멘트를 선택한다(S317).

[0042] 이후, 하이라이트 이벤트를 보정할 수 있다(S320). 일 실시예에서, 하이라이트 이벤트를 보정하는 것은 실제로 인접한 시간 윈도우에 연속적인 하이라이트 이벤트가 속하고 이는 동일한 장면에 속하는 경우가 많으므로 가장 첫 번째 하이라이트 이벤트만을 마킹함으로써 중복된 하이라이트 이벤트를 방지할 수 있다(S321).

[0043] 일 실시예에서, 동영상 콘텐츠와 상기 동영상 콘텐츠와 연관된 텍스트 데이터 간의 시간 차이를 보정할 수 있다(S323). 예컨대, 사용자는 텔레비전, PC 또는 스마트 폰에서 동영상 콘텐츠를 시청하면서 관련된 텍스트 데이터를 생성할 수 있다. 따라서, 비디오 컨텐츠가 시작되는 시간과 그에 관련된 텍스트 데이터가 생성되는 시간은 일치하지 않을 수 있다. 일반적으로 사용자가 동영상 컨텐츠를 시청하는 시간은 미디어의 유형과 통신 채널에 따라 달라질 수 있다. 예를 들어, 지상파 디지털 텔레비전은 동영상을 MPEG2-TS로 인코딩하여 라디오 웨이브를 통해 전송되고, 스마트폰 상의 동영상 컨텐츠를 무선 랜, CDMA, 3G 또는 LTE 를 통해 전송되는데, 무선 채널을 이용하는 스마트폰의 경우, 단대단 딜레이는 라디오 웨이브 또는 유선 환경의 경우 보다 크게 되는데, 왜냐하면 동영상 컨텐츠가 추가적으로 MPEG4 또는 H.264로 인코딩되어야 하고, 무선 링크의 지연(latency)도 있기 때문이다. 또한, 동영상 콘텐츠와 그에 관련된 텍스트 데이터 간의 시간 차이를 크게 하는 것 중, 웹사이트 상의 텍스트 메시지의 타임 스탬프(time stamp)로 인한 지연도 있다. 이는 사용자가 채팅 또는 댓글 메시지를 다 작성한 후 인터넷으로 전송할 때 타임 스탬프가 생성되기 때문이다. 따라서, 동영상 장면과 그에 대응하는 텍스트 메시지 간의 시간 차이를 고려해야 보다 정확한 하이라이트 이벤트를 선택할 수 있다. 이러한 측면에서, 본 발명은 TV 프로그램 또는 비디오 스트리밍 사이트의 시간표를 참조로 하고, 웹 사이트의 채팅 메시지의 '시작' 또는 '출발' 과 같은 키워드를 기초로 비디오 컨텐츠의 시작 시간을 보정하였다.

[0044] 이 후에, 동영상의 하이라이트 이벤트에 대해 제목을 합성하는 단계(S330)이다.

[0045] 일 실시예에서, 동영상 하이라이트 이벤트를 작성하기 위해서 4개의 문장 템플릿을 정의할 수 있다(S331). 일 실시예에서, 동영상의 제목은 두 개 또는 세 개의 단어 묶음(bag)이거나, 구절, 또는 주어, 동사, (목적어)가

포함되는 완벽한 문장일 수 있다. 일 실시예에서, 채팅 메시지 또는 댓글을 형태소 분석을 한 후, 각각의 토큰에 단어 클래스로 태깅을 할 수 있다.

[0046] 일 실시예에서, 명사에 주목하여 해당 장면의 제목으로 가장 자주 나오는 명사를 선택할 수 있는데 가장 많이 나오는 명사는 해당 장면의 상황을 효과적으로 설명할 수 있기 때문이다(S333). 일 실시예에서, 주인공의 이름을 해당 장면의 제목으로 선택할 수 있다. 일 실시예에서, 비속어 등의 부적절한 단어를 사전에 제거할 수 있다.

[0047] 다음으로, 단어들 사이의 거리가 짧은 단어들을 조합하여 복합 명사를 생성함으로써 두 개 또는 세 개의 단어들의 묶음을 생성할 수 있다(S335). 일 실시예에서, 워드2벡 모델(Word2Vec model)을 사용하여 단어들의 묶음을 생성할 수 있다.

[0048] 다음으로, 두 개 또는 세 개의 단어가 주어졌을 때, 단어들을 조합하여 해당 장면을 묘사하는 서사 형식의 문장을 합성할 수 있다(S337). 장면 제목 문장을 합성하기 위해, 문장에 포함될 수 있는 적절한 동사, 목적어 및 명사를 선택할 수 있다. 문장의 주어로서 주로 언급된 명사를 고려할 수 있고, 명사와 그 외 부분을 추출한 후, 명사와 그 외 부분의 유사도를 계산할 수 있다. 가장 짧은 단어 벡터와 토큰의 쌍을 선택하여 문장을 완성할 수 있다.

[0049] 도 4은 도 3에서 설명한 동작에 따라 대중 상호 웹사이트(mass interaction website: 유투브, 트위치, 네이버 스포츠 도는 디시인사이트)에서 제목과 함께 동영상 하이라이트를 생성하는 예시적인 방법을 도시하는 도면이다.

[0050] 도 4에 도시된 바와 같이 대중 상호 웹사이트에서 데이터를 수집하고, 하이라이트 이벤트를 검출한 후 하이라이트 제목을 합성하고, 제목과 함께 하이라이트 영상을 제공할 수 있다.

[0051] 도 5는 각 장르에서 제목과 함께 동영상 하이라이트 이벤트를 제공하는 일 실시예를 도시한 도면이다.

[0052] 도 5에 따르면, 생성된 제목은 일반적으로 각 장면의 등장인물의 이름을 포함하고 단어는 일반적으로 각 장면의 주연 배우와 연관되어 있음을 알 수 있다. 또한, 합성된 명사는 등장인물의 이름과 연관된 단어 사이의 높은 유사도 메트릭을 가지고 해당 장면의 배우들의 행동을 기술하고 있음을 알 수 있다. 완성된 문장 또한 등장인물의 이름과 행동을 표현하고 있다.

[0053] 드라마의 경우에, 배우의 외모와 관련되어 묘사하는 단어 및 해당 장면의 연기에 관한 단어를 관찰할 수 있었으며, 오락물의 경우, '재밌는', '흥미로운', '놀라운' 등의 감정을 나타내는 단어가 빈출하였다.

[0054] 스포츠의 경우에는 점수 또는 선수의 파울 플레이에 관한 많은 용어를 발견할 수 있었고, 또는 플레이를 평가하는 단어를 많이 발견할 수 있었다.

[0055] 온라인 게임의 경우, 캐릭터, 약어 및 비어에 관한 많은 다양한 단어와 게임에 특정된 용어들이 있었다.

[0056] 정확도를 측정하기 위해서, 20개의 동영상 콘텐츠에 대해 80개의 제목을 선택하였고 수작업으로 3개의 등급으로 나누어 제목의 정확도를 평가하였다. 12명의 지원자로부터 설문조사를 통해 20개의 동영상 콘텐츠의 80개의 제목을 평가하였다. 설문조사 결과, 정확도는 67.5%였고, 80개의 제목 중에 54개가 동일한 등급을 받았다.

[0057] 높은 등급을 받은 제목은 흥미로운 표현 또는 신조어가 포함되거나 적절한 표현으로 된 제목임을 알 수 있었다. 반면, 낮은 등급을 받은 제목은 어색한 단어의 조합 또는 동영상 컨텐츠와 제목과의 불일치인 경우였다.

[0058] 도 6은 각각의 비디오 카테고리에 대해서 4 가지의 제목 타입의 평균 점수를 나타내는 그래프이다.

[0059] 전체적으로, 조합 단어로 된 제목이 장르 전체에 걸쳐 높은 점수를 받았는데, 서로의 거리가 짧은 단어가 각 장면에 보다 의미있는 제목을 형성하는 경향이 있기 때문이다. 두 개 또는 세 개의 단어의 묶음으로 된 제목은 일반적으로 드라마나 예능에 높은 점수를 받았다. 반면, 문장으로 된 제목은 낮은 점수를 받았다. 사용자들은 문장에 어색한 단어의 조합이 있거나 타이틀과 동영상 컨텐츠 간에 불일치가 있는 경우 낮은 점수를 주는 경향이 있음을 알 수 있다.

[0060] 반면, 세 개의 단어로 된 제목은 두 개의 단어로 된 제목 보다 낮은 점수를 받았는데, 빈출 상위의 세 개의 단어는 종종 동영상 장면과 직접 관련되지 않은 하나의 단어를 포함하기 때문으로 분석된다. 평균적으로, 사람들은 단어의 조합으로 된 제목을 선호하는 것으로 보인다.

[0061] 본원에 개시된 실시예들에 있어서, 도시된 구성 요소들의 배치는 발명이 구현되는 환경 또는 요구 사항에 따라

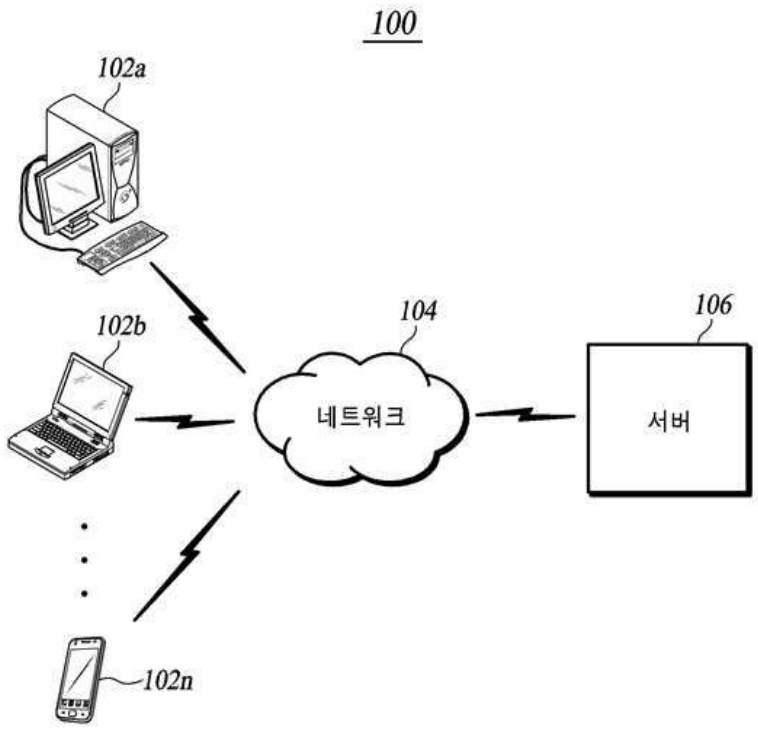
달라질 수 있다. 예컨대, 일부 구성 요소가 생략되거나 몇몇 구성 요소들이 통합되어 하나로 실시될 수 있다. 또한 일부 구성 요소들의 배치 순서 및 연결이 변경될 수 있다.

[0062] 본 발명 및 그 다양한 기능적 구성 요소들은 특정 실시예들로 설명되었으나, 본 발명은 하드웨어, 소프트웨어, 펌웨어, 미들웨어 또는 이들의 조합으로 구현될 수 있으며, 시스템, 서브시스템, 구성 요소들 또는 이들의 서브 구성 요소들로 활용될 수 있음을 이해하여야 한다. 소프트웨어로 구현되는 경우, 본 발명의 요소들은 필요한 작업들을 수행하기 위한 명령어들/코드 세그먼트들이 될 수 있다. 프로그램 또는 코드 세그먼트들은 프로세서 판독가능 매체와 같은 머신 판독가능 매체, 컴퓨터 프로그램 제품 내에 저장될 수 있다. 머신 판독가능 매체 또는 프로세서 판독가능 매체는 머신(예컨대, 프로세서, 컴퓨터 등)에 의해 판독되고 실행 가능한 형태로 정보를 저장 또는 전송할 수 있는 임의의 매체를 포함할 수 있다.

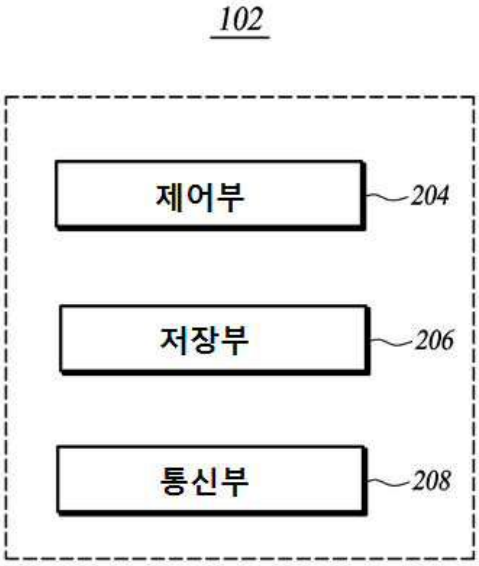
[0063] 이상에서는 본 발명의 다양한 실시예들에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예들에 한정되지 아니하며, 상술한 실시예들은 첨부하는 특허청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 다양하게 변형 실시될 수 있음은 물론이고, 이러한 변형 실시예들이 본 발명의 기술적 사상이나 범위와 별개로 이해되어져서는 아니 될 것이다. 따라서, 본 발명의 기술적 범위는 오직 첨부된 특허청구범위에 의해서만 정해져야 할 것이다.

도면

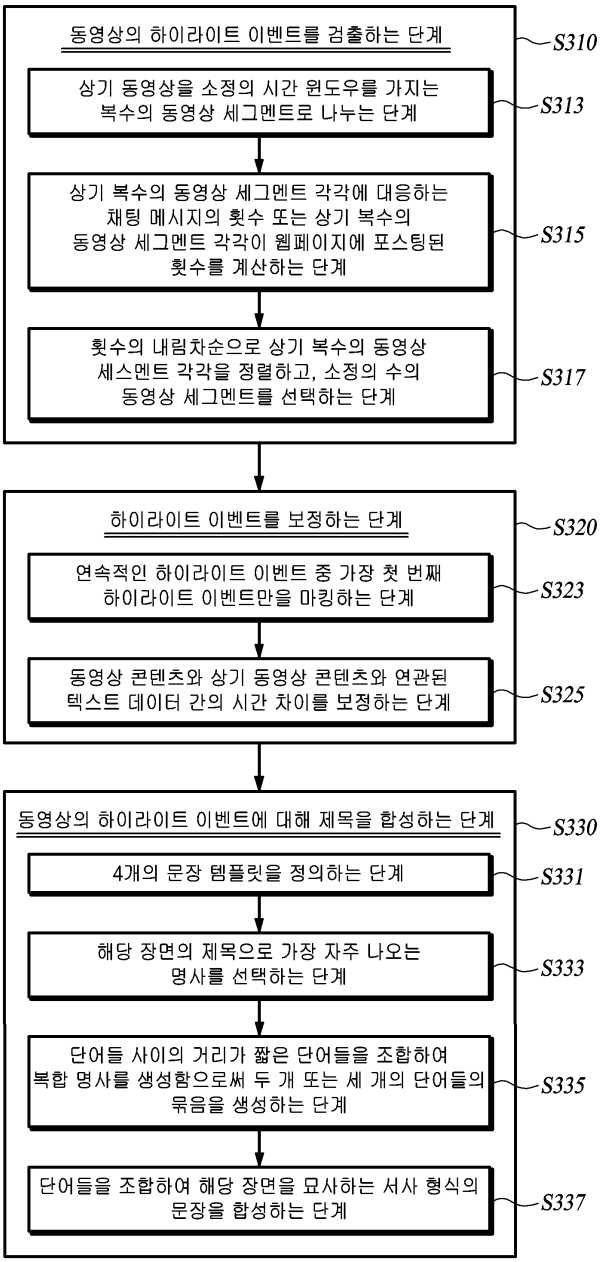
도면1



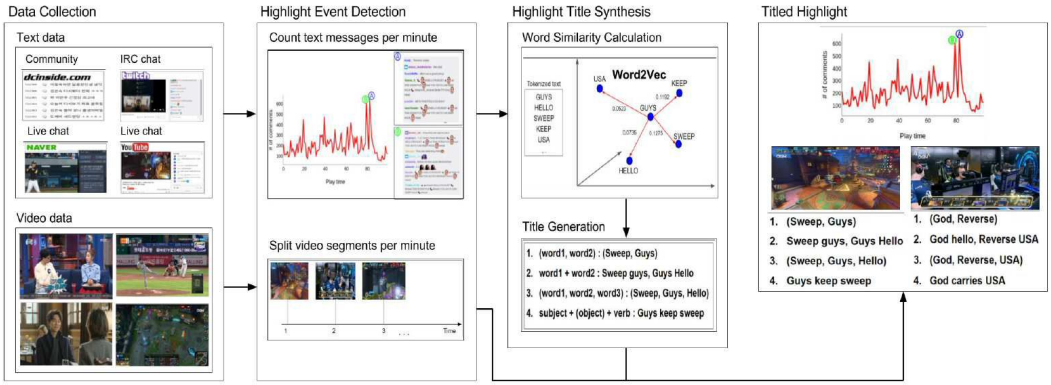
도면2



도면3


동영상의 하이라이트 이벤트를 검출하는 단계
S310
상기 동영상을 소정의 시간 윈도우를 가지는 복수의 동영상 세그멘트로 나누는 단계
S313
상기 복수의 동영상 세그멘트 각각에 대응하는 채팅 메시지의 횟수 또는 상기 복수의 동영상 세그멘트 각각이 웹페이지에 포스팅된 횟수를 계산하는 단계
S315
횟수의 내림차순으로 상기 복수의 동영상 세스멘트 각각을 정렬하고, 소정의 수의 동영상 세그멘트를 선택하는 단계
S317
하이라이트 이벤트를 보정하는 단계
S320
연속적인 하이라이트 이벤트 중 가장 첫 번째 하이라이트 이벤트만을 마킹하는 단계
S323
동영상 콘텐츠와 상기 동영상 콘텐츠와 연관된 텍스트 데이터 간의 시간 차이를 보정하는 단계
S325
동영상의 하이라이트 이벤트에 대해 제목을 합성하는 단계
S330
4개의 문장 템플릿을 정의하는 단계
S331
해당 장면의 제목으로 가장 자주 나오는 명사를 선택하는 단계
S333
단어들 사이의 거리가 짧은 단어들을 조합하여 복합 명사를 생성함으로써 두 개 또는 세 개의 단어들의 묶음을 생성하는 단계
S335
단어들을 조합하여 해당 장면을 묘사하는 서사 형식의 문장을 합성하는 단계
S337

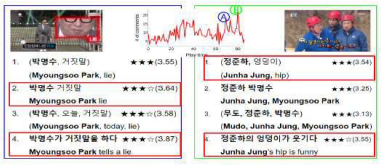
도면4



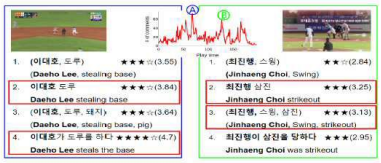
도면5



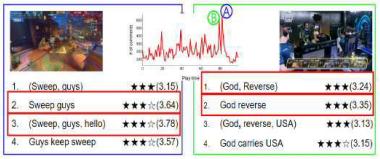
(a) Drama (Goblin, text from Dcinside, 2017-01-20)



(b) Entertainment (MBC Infinite-Challenge, text from Dcinside, 2017-04-29)



(c) Sports (KBO, text from Naver Sports, 2017-04-20)



(d) Online game (Overwatch APEX, text from Youtube, Twitch, 2017-05-06)

도면6

