



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년04월25일
(11) 등록번호 10-2390185
(24) 등록일자 2022년04월20일

(51) 국제특허분류(Int. Cl.)
H04N 21/854 (2011.01) G06F 40/247 (2020.01)
G06Q 50/10 (2012.01) H04N 21/845 (2011.01)
(52) CPC특허분류
H04N 21/854 (2013.01)
G06F 16/7844 (2019.01)
(21) 출원번호 10-2021-0013445
(22) 출원일자 2021년01월29일
심사청구일자 2021년01월29일
(56) 선행기술조사문헌
KR1020200081017 A
KR1020100013157 A
KR1020090022375 A

(73) 특허권자
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
(72) 발명자
장준도
서울특별시 강서구 초원로16길 41 2층
김상원
인천광역시 연수구 능허대로79번길 30 럭키송도아파트 112동 1003호
이종영
서울특별시 양천구 지양로13길 14 미림주택 나동 102호
(74) 대리인
이은철, 이수찬

전체 청구항 수 : 총 6 항

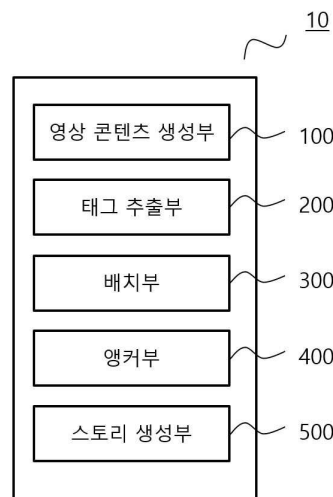
심사관 : 홍기완

(54) 발명의 명칭 **유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템 및 그 방법**

(57) 요약

본 발명은 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템에 관한 것으로써, 서사 콘텐츠의 기계적 생성을 위한 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템 및 그 방법에 관한 것이다. 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템에 의하면, 창작자들에게 적절한 스토리 진행의 예시를 제시함과 함께 작성 중인 시나리오와 유사한 내용의 상황을 화면으로 시뮬레이션해볼 수 있도록 하여 시나리오 작성 과정에서부터 반복 가능한 테스트 절차 사이클(Agile Process)을 수행하여 더욱 높은 품질의 시나리오를 빠른 시간 내에 작성할 수 있는 효과가 있다.

대표도 - 도1



(52) CPC특허분류

G06F 40/117 (2020.01)

G06F 40/247 (2020.01)

G06Q 50/10 (2015.01)

H04N 21/8455 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711103050
과제번호	2017-0-01642-004
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성(R&D)
연구과제명	인공지능을 활용한 콘텐츠 창작 기술
기 여 율	1/1
과제수행기관명	인하대학교 산학협력단
연구기간	2020.01.01 ~ 2020.12.31

명세서

청구범위

청구항 1

영상 데이터베이스 라이브러리에서 선택되는 영상 콘텐츠의 기 설정된 시간적 배열에 따라 분할 및 재배포하는 영상 콘텐츠 생성부(100);

상기 영상 콘텐츠 생성부에서 선택되는 각 영상 데이터 간에 관련된 태그를 추출하는 태그 추출부(200);

상기 태그 추출 과정에서 의미적으로 위계성을 갖는 계층형 구조에 태그를 배치하는 배치부(300);

상기 태그를 배치하는 과정에서 의미적으로 근접한 태그를 연결하는 앵커를 구성하는 앵커부(400); 및

상기 앵커부를 통해 연결되는 태그 개념의 각 영상 데이터 간에 대한 유의어 기반 유사성 가중치를 부여하고 상기 영상 콘텐츠 생성부를 통해 생성되는 콘텐츠 조합에서 유의어에 기반하여 학습된 데이터 배열순서와 유사한 영상 데이터를 선택하여 유의어 기반의 스토리를 생성하는 스토리 생성부(500);를 포함하는 것을 특징으로 하는 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템.

청구항 2

제1항에 있어서,

상기 태그 추출부는 사용자의 텍스트 형태 질의(Query)에 기반하여 검색결과를 보여주기 위해 텍스트 형태의 미리 정의된 태그 정보를 포함하는 콘텐츠 중, 동일 콘텐츠에 관련되는 태그를 추출하는 것을 특징으로 하는 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템.

청구항 3

제1항에 있어서,

상기 스토리 생성부는 콘텐츠의 원본 데이터와 함께 데이터의 내용을 언어적으로 설명하기 위한 태그 또는 설명을 포함하는 메타 데이터가 연동되어 저장되도록 하는 것을 특징으로 하는 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템.

청구항 4

제1항에 있어서,

상기 배치부는 태그 추출 과정에서 각 개념들이 의미적으로 위계성을 갖는 계층형 구조에 배치하여 포함 관계와 유의어 관계를 데이터베이스 상에 저장하는 것을 특징으로 하는 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템.

청구항 5

제1항에 있어서,

상기 앵커부는 태그를 배치하는 과정에서 의미적으로 근접한 태그를 횡적으로 연결되는 앵커를 구성하여 인접 개념을 연결하는 것을 특징으로 하는 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템.

청구항 6

(a)영상 데이터베이스 라이브러리에서 선택되는 영상 콘텐츠의 기 설정된 시간적 배열에 따라 분할 및 재배포하는 단계;

(b)상기 (a)단계에서 선택되는 각 영상 데이터 간에 관련된 태그를 추출하는 단계;

(c)상기 태그 추출 과정에서 의미적으로 위계성을 갖는 계층형 구조에 태그를 배치하는 단계;

(d)태그를 배치하는 과정에서 의미적으로 근접한 태그를 연결하는 앵커를 구성하는 단계;

(e)상기 (d)단계에서, 연결되는 태그 개념의 각 영상 데이터 간에 대한 유의어 기반 유사성 가중치를 부여하고 상기 영상 콘텐츠 생성부를 통해 생성되는 콘텐츠 조합에서 유의어에 기반하여 학습된 데이터 배열순서와 유사한 영상 데이터를 선택하여 유의어 기반의 스토리를 생성하는 단계;를 포함하는 것을 특징으로 하는 유의어 클러스터링에 기반한 서사 콘텐츠 생성방법.

발명의 설명

기술 분야

[0001] 본 발명은 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템에 관한 것으로써, 더욱 상세하게는 서사 콘텐츠의 기계적 생성을 위한 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템 및 그 방법에 관한 것이다.

배경 기술

[0002] 현재 인터넷의 발달로 사용자가 급속히 늘어나면서 웹 서비스 환경이 다양하게 변화하고 있다. 종래의 웹 서비스가 정적으로 수동적인데 반하여, 웹 서비스는 점차 동적이고 능동적으로 변화되고 있으며, 이러한 웹 서비스 변화의 흐름을 반영하기 위하여 웹 2.0이 도입되었다.

[0003] 웹 2.0이란 정보의 개방을 통해 인터넷 사용자들 간의 정보 공유와 참여를 이끌어내고, 창조된 정보의 가치를 지속적으로 증대시키기 위하여 개발된 일련의 움직임을 의미한다. 즉, 웹 2.0에서는 개방적인 웹 환경을 기반으로 네티즌들이 자유롭게 참여하고, 콘텐츠를 생산 및 재창조, 공유할 수 있다.

[0004] 웹 2.0에서 정보는 사용자에게 의하여 생산되고, 사용자가 붙인 태그에 의해 정보가 체계화된다. 사용자들은 이러한 정보를 용이하게 공유할 수 있으며, 따라서 다양한 리소스들이 상호 연관된다. 이와 같이 웹 2.0 현상은 모든 인터넷 사이트의 필수 전략이 되었으며, 웹 2.0을 성공적으로 구현하기 위하여 다양한 기법들이 소개되고 있다.

[0005] 이러한 기법들 중 하나가 태깅(tagging)이다. 태깅은 블로그와 같은 웹 문서로부터 이미지, 동영상과 같은 멀티미디어 콘텐츠에 까지 폭넓게 이용되고 있는데, 사용자가 자신이 생성한 콘텐츠에 태그를 붙임으로써 검색과 분류가 용이하게 이루어지도록 하는 것이다.

[0006] 종래, 공개특허 제2010-0013157호에서 연관 태그에 기반한 태그 클러스터링 장치에 의하면, 태그 매핑 과정에서 동일 태그의 출현 빈도를 추출하고, 연관 태그 쌍들 중 임계치 이상의 빈도수를 가지는 연관 태그들만을 추출하여 태그 클러스터를 생성하여 동의어를 나타내는 태그들을 동일한 태그로 간주한다.

[0007] 이러한 빈도 기반 텍스트 분석(Frequency Based Text Analysis)과 같은 통계적인 추측은 이론적으로 인간 활동에서 사용된 모든 언어 기록이 있다면 이 문제를 매우 잘 해결할 수 있으나 현실적으로는 수집 가능한 텍스트 정보에는 한계가 있고 지역, 연령, 관심 분야, 직종 등으로 나뉘어질 수 있는 언어 사용의 하위 그룹들의 특수한 언어 사용 습관은 이러한 일반론적 분석으로는 놓칠 수 있는 언어적 다양성을 가지고 있으며 이는 표현의 유사성을 예측하는 통계적 추측 방식에 사용되는 언어 데이터베이스도 모든 상황에서 동일한 것을 사용할 수는 없다는 문제가 있다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 공개특허 제10-2010-0013157호(2010.02.09)

발명의 내용

해결하려는 과제

[0009] 본 발명은 상술한 문제를 해결하고자 고안한 것으로, 작성 중인 시나리오와 유사한 내용의 상황을 화면으로 시뮬레이션해볼 수 있도록 하는 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템 및 그 방법을 제공함에 목적

이 있다.

과제의 해결 수단

- [0010] 본 발명의 일 측면에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템은 영상 데이터베이스 라이브러리에서 선택되는 영상 콘텐츠의 기 설정된 시간적 배열에 따라 분할 및 재배포하는 영상 콘텐츠 생성부(100); 상기 영상 콘텐츠 생성부에서 선택되는 각 영상 데이터 간에 관련된 태그를 추출하는 태그 추출부(200); 상기 태그 추출 과정에서 의미적으로 위계성을 갖는 계층형 구조에 태그를 배치하는 배치부(300); 상기 태그를 배치하는 과정에서 의미적으로 근접한 태그를 연결하는 앵커를 구성하는 앵커부(400); 및 상기 앵커부를 통해 연결되는 태그 개념의 각 영상 데이터 간에 대한 유의어 기반 유사성 가중치를 부여하고 상기 영상 콘텐츠 생성부를 통해 생성되는 콘텐츠 조합에서 유의어에 기반하여 학습된 데이터 배열순서와 유사한 영상 데이터를 선택하여 유의어 기반의 스토리를 생성하는 스토리 생성부(500);를 포함한다.
- [0011] 바람직하게 상기 태그 추출부는 사용자의 텍스트 형태 질의(Query)에 기반하여 검색결과를 보여주기 위해 텍스트 형태의 미리 정의된 태그 정보를 포함하는 콘텐츠 중, 동일 콘텐츠에 관련되는 태그를 추출한다.
- [0012] 바람직하게 상기 스토리 생성부는 콘텐츠의 원본 데이터와 함께 데이터의 내용을 언어적으로 설명하기 위한 태그 또는 설명을 포함하는 메타 데이터가 연동되어 저장되도록 한다.
- [0013] 바람직하게 배치부는 태그 추출 과정에서 각 개념들이 의미적으로 위계성을 갖는 계층형 구조에 배치하여 포함관계와 유의어 관계를 데이터베이스 상에 저장한다.
- [0014] 바람직하게 앵커부는 태그를 배치하는 과정에서 의미적으로 근접한 태그를 횡적으로 연결되는 앵커를 구성하여 인접 개념을 연결한다.
- [0015] 한편, 본 발명의 일 측면에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성방법은 (a)영상 데이터베이스 라이브러리에서 선택되는 영상 콘텐츠의 기 설정된 시간적 배열에 따라 분할 및 재배포하는 단계; (b)상기 (a)단계에서 선택되는 각 영상 데이터 간에 관련된 태그를 추출하는 단계; (c)상기 태그 추출 과정에서 의미적으로 위계성을 갖는 계층형 구조에 태그를 배치하는 단계; (d)태그를 배치하는 과정에서 의미적으로 근접한 태그를 연결하는 앵커를 구성하는 단계; (e)상기 (d)단계에서, 연결되는 태그 개념의 각 영상 데이터 간에 대한 유의어 기반 유사성 가중치를 부여하고 상기 영상 콘텐츠 생성부를 통해 생성되는 콘텐츠 조합에서 유의어에 기반하여 학습된 데이터 배열순서와 유사한 영상 데이터를 선택하여 유의어 기반의 스토리를 생성하는 단계;를 포함한다.

발명의 효과

- [0016] 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템에 의하면, 창작자들에게 적절한 스토리 진행의 예시를 제시함과 함께 작성 중인 시나리오와 유사한 내용의 상황을 화면으로 시뮬레이션해볼 수 있도록 하여 시나리오 작성 과정에서부터 반복 가능한 테스트 절차 사이클(Agile Process)을 수행하여 더욱 높은 품질의 시나리오를 빠른 시간 내에 작성할 수 있는 효과가 있다.

도면의 간단한 설명

- [0017] 도 1은 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템의 구성을 나타낸 도면이다.
- 도 2는 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템에서, 영상 콘텐츠의 시간적 배열에 따른 배치를 나타낸 도면이다.
- 도 3은 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템의 특정 영상 클립의 후속 영상 클립을 설명하기 위한 도면이다.
- 도 4는 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템의 계층형 배치 구조와 앵커 연결을 설명하기 위한 도면이다.
- 도 5는 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템의 학습된 영상 데이터와 조합될 수 있는 기계 생성결과를 나타낸 도면이다.
- 도 6은 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템의 유의어 기반의 클러스터링을 통한 시뮬레이션을 설명하기 위한 도면이다.

도 7은 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성방법을 나타낸 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0018] 본 발명의 실시예에서 제시되는 특정한 구조 내지 기능적 설명들은 단지 본 발명의 개념에 따른 실시예를 설명하기 위한 목적으로 예시된 것으로, 본 발명의 개념에 따른 실시예들은 다양한 형태로 실시될 수 있다. 또한, 본 명세서에 설명된 실시예들에 한정되는 것으로 해석되어서는 아니 되며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경물, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.
- [0019] 한편, 본 발명에서 제1 및/또는 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 한정되지는 않는다. 상기 용어들은 하나의 구성요소를 다른 구성요소들과 구별하는 목적으로만, 예컨대 본 발명의 개념에 따른 권리 범위로부터 벗어나지 않는 범위 내에서, 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소는 제1 구성요소로도 명명될 수 있다.
- [0020] 이하 첨부된 도면을 참조하여 본 발명의 실시예를 설명한다. 본 발명의 실시예를 설명함에 있어서, 관련된 공지 기능 혹은 구성에 대한 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우 그 설명을 생략하였다.
- [0021] 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템은 각 영상 데이터 간에 대한 유의어 기반 유사성 가중치를 부여하고, 기계적 콘텐츠 조합에서 유의어에 기반하여 학습된 데이터 배열순서와 유사한 데이터를 선택하여 유의어 기반의 스토리를 생성할 수 있다. 이러한 유의어 기반의 스토리 생성 엔진을 통해 기존에 만들어진 영상 클립들을 기반으로 새로운 동영상 생성함으로써, 창작자들에게 적절한 스토리 진행의 예시를 제시함과 함께 작성 중인 시나리오와 유사한 내용의 상황을 화면으로 시뮬레이션해볼 수 있도록 하여 시나리오 작성 과정에서부터 반복 가능한 테스트 절차 사이클(Agile Process)을 수행하여 더욱 높은 품질의 시나리오를 빠른 시간 내에 작성할 수 있는 효과가 있다.
- [0022] 도 1은 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템의 구성을 나타낸 도면이다. 도 1에 도시된 바와 같이, 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템(10)은 영상 콘텐츠 생성부(100), 태그 추출부(200), 배치부(300), 앵커부(400), 스토리 생성부(500)를 포함한다.
- [0023] 영상 콘텐츠 생성부(100)는 영상 데이터베이스 라이브러리에서 선택되는 영상 콘텐츠의 기 설정된 시간적 배열에 따라 분할 및 재배포하는 구성이다. 이러한 영상 콘텐츠 생성부는 도 2에 도시된 바와 같이, 서사 콘텐츠의 기계적 생성을 위해서 콘텐츠의 시간적 배열 기준에 대한 제시와 함께 특정한 재생 절차에서 가장 적합한 내용을 구성할 수 있는 맥락적 정보를 포함한다.
- [0024] 영상 콘텐츠 생성에 있어서 기존 촬영된 실사 영상을 분할 및 재배포하여 재구성하는 방식을 이용할 경우 서사적 맥락에 대한 내용의 수집은 음성 정보(음성의 크기, 높낮이, 인물의 감정상태 추정, 언어적 내용), 영상 정보(얼굴을 통한 개인 특정, 표정 정보, 장소 정보, 사물 인식) 등을 이용할 수 있다. 이러한 정보 및 기타 정보는 문자 형식으로 원본 데이터와 함께 인덱싱 되어 이후의 검색, 편집, 분류 등의 작업에 활용될 수 있다. 하지만 이러한 각종 분류 조건에 대해 최대한 자세한 언어적 메타 데이터를 구성한다고 하더라도 유사한 정보에 대해서도 다양한 방식의 언어적 메타 데이터가 구성될 수 있다는 문제가 발생한다. 아래는 영상 정보의 메타 데이터 구성에서 선택적으로, 혹은 동시에 사용될 수 있는 단어들의 예시이다.
- [0025] 사과, 과일, 다과, 디저트
- [0026] 달리다, 뛰다
- [0027] 집, 아파트, 거실
- [0028] 메타 데이터는 인간의 수작업, 혹은 기계에 의한 자동 생성에 의해 작성될 수 있으나 이 두 가지 방법 모두 영상 정보의 특정 상황을 언어적으로 기술함에 있어 다양한 표현 중 하나, 혹은 일부를 선택해야 하는 상황에 놓이게 된다. 이러한 문제는 아래와 같은 원인으로 발생한다. 다른 표현이지만 의미적으로 유사하여 선택적으로 사용될 수 있는 경우, 의미적으로 더 상위에, 혹은 하위에 위치하여 동일한 대상을 지칭하나 해석될 수 있는 표현의 범주가 달라질 수 있는 경우가 있다.
- [0029] 유의어 사전을 이용할 경우 이러한 정보에 기반하여 영상 데이터베이스와 같이 비언어적 정보로 이루어진 데이터셋에 대한 효과적인 접근이 가능해진다. 사용자의 질의에 유연하게 대처할 수 있는 검색결과, 분류 방식, 유연한 범주의 설정이 가능하며 서사 콘텐츠의 기반적 생성 역시 이러한 기술적 기반 위에서 이루어질 수 있다.

- [0030] 빈도 기반 텍스트 분석(Frequency Based Text Analysis)과 같은 통계적인 추측은 이론적으로 인간 활동에서 사용된 모든 언어 기록이 있다면 이 문제를 매우 잘 해결할 수 있으나 현실적으로는 수집 가능한 텍스트 정보에는 한계가 있고 지역, 연령, 관심 분야, 직종 등으로 나뉘어질 수 있는 언어 사용의 하위 그룹들의 특수한 언어 사용 습관은 이러한 일반론적 분석으로는 놓칠 수 있는 언어적 다양성을 가지고 있으며 이는 표현의 유사성을 예측하는 통계적 추측 방식에 사용되는 언어 데이터베이스도 모든 상황에서 동일한 것을 사용할 수는 없다는 것을 의미한다.
- [0031] 인간의 수집과 수정에 기반한 유의어 데이터베이스는 빅 데이터와 머신 러닝에 소요되는 스토리지, 컴퓨팅 자원 등에 대한 소요를 절감할 수 있으며 수집되는 콘텐츠의 형태에 따라 유연하게 수정될 수 있으며 무엇보다 연관 데이터 추천에 대한 근거를 데이터베이스 관리자가 직관적으로 확인하고 개선 사항을 추가할 수 있다는 장점을 가지고 있다. 빅 데이터에 기반한 통계 분석이 유의미하게 작동하기 위해서는 최소 수십 기가바이트 이상의 텍스트 정보를 보유하고, 지속적으로 업데이트 해야 하나 대부분의 프로젝트에서 이 정도 분량의 텍스트 정보는 실제 분석해야 하는 비디오 내의 내용 정보보다도 많은 양이 될 수 있고 수집된 텍스트 데이터가 영상 데이터베이스를 분석하는데 적합한 샘플인지의 여부도 증명은 불가하다는 점에서 유의어 데이터베이스에 기반한 확장된 검색 질의 방식은 상대적으로 중소규모의 프로젝트에서 큰 효용을 가진다고 할 수 있다.
- [0032] 도 2는 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템의 영상 콘텐츠의 시간적 배열에 따른 배치를 나타낸 도면이다. 도 2에서 영상 데이터베이스 라이브러리에서 선택된 여러 개의 영상 중에서 특정 영상은 다른 영상에 비해 앞부분에 배치될 수 있다.
- [0033] 영상 콘텐츠 생성부(100)는 영상의 조합 배치에 있어서, 특정 영상 클립에 도달했을 때에 종료되어야 하며, 이를 통해 발단부에서 시작해 결말부로 마무리되는 유한한 분량의 시간적으로 의미있게 배열된 조합된 영상을 생성할 수 있다.
- [0034] 도 3은 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템의 특정 영상 클립의 후속 영상 클립을 설명하기 위한 도면이다. 도 3에 도시된 바와 같이, 특정 영상 클립 뒤에 위치될 수 있는 유효한 후속 영상 클립은 여러 개가 있을 수 있다. 유효한 영상 클립의 목록에서 전체 스토리의 맥락을 위해 가장 적합한 클립을 선택할 수 있어야 한다. 이를 위한 과정으로 영상 콘텐츠 생성부에서 선택되는 각 영상 데이터 간에 관련된 태그를 추출하고, 이러한 태그 추출 과정에서 의미적으로 위계성을 갖는 계층형 구조에 태그를 배치하고, 태그 배치과정에서 의미적으로 근접한 태그를 연결하는 앵커를 구성하여 전체 스토리의 맥락을 위한 가장 적합한 클립을 선택할 수 있도록 한다.
- [0035] 도 4는 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템의 계층형 배치 구조와 앵커 연결을 설명하기 위한 도면이다. 각 개념들은 의미적으로 위계성을 갖는 계층형 구조에 배치될 수 있으며 이를 통해 기본적인 포함 관계와 유의어 관계를 데이터베이스 상에 저장할 수 있다. 또한 위계적 개념으로 정의하기 힘든, 다른 위계상에 존재하지만 의미적으로는 근접한 인접한 개념의 경우에는 횡적으로 연결되는 앵커를 구성하여 가능한 모든 인접 개념을 연결할 수 있고 앵커의 연결에 대해 가중치를 부여하는 것으로 보다 긴밀한 연결과 인접성이 상대적으로 적은 막연한 연결의 연결 강도를 규정할 수 있다. 초기의 앵커 구성은 통계적 방식을 이용할 수 있으나 인간이 이해할 수 있는 시각화 된 연결성을 제공할 수 있는 본 방식은 인간의 지도학습을 통해 손쉬운 지도학습 절차를 이행함으로써 데이터 학습량이 적거나, 일반적 용도가 아닌 특수한 상황 속에서의 적절한 인접 개념들을 연결해야 할 필요가 있을 때에 빠른 개념 분류 기계를 구축할 수 있게 된다.
- [0036] 또한 도 4에 도시된 바와 같이, 위계적 관계를 가진 어휘, 또는 개념들에 대한 트리 구조형 배치로 상위 범주의 개념들이 완전히 포괄할 수 없는 하위 개념에 대한 수정적 앵커를 생성하여 유기적인 연결을 제공한다. 유의어 각 개념들의 유의어 개념들에 대한 시각적인 직관적 view를 제공하는 디스플레이를 포함한다.
- [0037] 태그 추출부(200)는 영상 콘텐츠 생성부에서 선택되는 각 영상 데이터 간에 관련된 태그를 추출하는 구성이다. 이러한 태그 추출부는 사용자의 텍스트 형태 질의(Query)에 기반하여 검색결과를 보여주기 위해 텍스트 형태의 미리 정의된 태그 정보를 포함하는 콘텐츠 중, 동일 콘텐츠에 관련되는 태그를 추출한다.
- [0038] 배치부(300)는 태그 추출 과정에서 의미적으로 위계성을 갖는 계층형 구조에 태그를 배치하는 구성이다. 이러한 배치부는 태그 추출 과정에서 각 개념들이 의미적으로 위계성을 갖는 계층형 구조에 배치하여 포함 관계와 유의어 관계를 데이터베이스 상에 저장한다.
- [0039] 앵커부(400)는 태그를 배치하는 과정에서 의미적으로 근접한 태그를 연결하는 앵커를 구성한다. 앵커부는 태그를 배치하는 과정에서 의미적으로 근접한 태그를 횡적으로 연결되는 앵커를 구성하여 인접 개념을 연결한다.

- [0040] 스토리 생성부(500)는 앵커부를 통해 연결되는 태그 개념의 각 영상 데이터 간에 대한 유의어 기반 유사성 가중치를 부여하고 상기 영상 콘텐츠 생성부를 통해 생성되는 콘텐츠 조합에서 유의어에 기반하여 학습된 데이터 배열순서와 유사한 영상 데이터를 선택하여 유의어 기반의 스토리를 생성한다. 이러한 스토리 생성부는 콘텐츠의 원본 데이터와 함께 데이터의 내용을 언어적으로 설명하기 위한 태그 또는 설명을 포함하는 메타 데이터가 연동되어 저장되도록 한다.
- [0041] 도 5는 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템의 학습된 영상 데이터와 조합될 수 있는 기계 생성결과를 나타낸 도면이다. 도 5에 도시된 바와 같이, 특정 일상이라는 가중치가 부여된 유의어 기반으로 연결되는 앵커를 구성하여 거실, 휴대폰 통화를 영상데이터로 배치하고, 도로, 운전은 제2 영상데이터로 배치하며, 사무실, 회의를 제3 영상데이터로 배치하고, 침실, 수면을 제4 영상데이터로 배치한다. 이러한 배치에 따라 제1 내지 제4 영상데이터에 따라 조합될 수 있는 기계 생성결과는 제1 영상데이터의 거실, 휴대폰 통화는 안방, 메신저 채팅과 같은 유의어 기반의 클러스터링을 통해 조합될 수 있고, 제2 영상데이터의 도로, 운전은 지하도, 달리기로 조합될 수 있다. 제3 영상데이터의 사무실, 회의는 강당, 컨퍼런스, 제4 영상데이터의 침실, 수면은 거실 휴식으로 조합될 수 있다.
- [0042] 도 6은 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템의 유의어 기반의 클러스터링을 통한 시물레이션을 설명하기 위한 도면이다. 도 6에 도시된 바와 같이, 유의어 기반의 클러스터링을 통해 앵커를 구성한 시대적 배경 선택을 "조선시대"로 하여 앵커부를 통해 연결되는 태그 개념의 각 영상 데이터 간에 대한 유의어 기반 유사성 가중치를 부여할 경우 생성되는 제1 내지 제4 영상데이터에 따라 유의어 기반으로 클러스터링을 통해 조합될 수 있다. 즉, 제1 영상데이터의 거실, 휴대폰 통화는 대청, 대화로 조합되고, 제2 영상데이터의 도로, 운전은 저잣거리, 승마로 조합되며, 제3 영상데이터의 사무실, 회의는 대전, 어전회의로 조합되고, 제4 영상데이터의 침실, 수면은 사랑채, 수면으로 조합되어 영상 콘텐츠 생성부를 통해 생성되는 콘텐츠 조합에서 유의어에 기반하여 학습된 데이터 배열순서와 유사한 영상 데이터를 선택함으로써, 유의어 기반의 스토리를 생성한다. 이러한 스토리 생성부는 콘텐츠의 원본 데이터와 함께 데이터의 내용을 언어적으로 설명하기 위한 태그 또는 설명을 포함하는 메타 데이터가 연동되어 저장되도록 한다.
- [0043] 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템은 유의어 관여에 의한 데이터 베이스 분류에 기반한 맥락적 일관성을 가진 기계적인 스토리를 생성하고, 이러한 스토리의 생성을 가능하게 해주는 영상 데이터 라이브러리를 작성할 수 있다.
- [0044] 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템은 서사적 스토리의 콘텐츠 요소들을 유의어에 기반하여 분류하고 데이터베이스화하는 방식으로 효율적인 데이터베이스를 구축하기 위해서는 다음과 같은 조건이 필요하다. 오디오, 비디오, 인터랙티브 콘텐츠 등의 라이브러리를 구성함에 있어 사용자의 텍스트 형태의 질의(Query)에 기반하여 검색결과를 보여주기 위해서는 텍스트 형태로 미리 정의된 태깅 정보가 필요하다. 이러한 비 언어적 데이터의 언어적 설명을 위해서는 원본 데이터와 함께 데이터의 내용을 언어적으로 설명하기 위한 태그(Tag)나 설명(Description)와 같은 메타 데이터(Meta Data)가 연동되어 저장되어야 한다. 언어적 태그 및 설명 정보는 비 언어적 데이터의 내용을 최대한 상세히 표현할 수 있어야 하며 이를 위해 적절한 조건의 분류 기준이 필요하다.
- [0045] 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성시스템의 스토리 생성 엔진은 기존에 만들어진 영상 클립들을 기반으로 새로운 동영상을 생성함으로써 창작자들에게 적절한 스토리 진행의 예시를 제시함과 함께 작성 중인 시나리오와 유사한 내용의 상황을 화면으로 시물레이션 해볼 수 있도록 하여 시나리오 작성 과정에서부터 반복 가능한 테스트 절차 사이클(Agile Proccess)을 수행할 수 있도록 하여 더욱 높은 품질의 시나리오를 빠른 시간 내에 작성할 수 있도록 할 수 있다.
- [0046] 도 7은 본 발명의 일 실시예에 따른 유의어 클러스터링에 기반한 서사 콘텐츠 생성방법을 나타낸 흐름도이다. 도 7에 도시된 바와 같이, (701)영상 데이터베이스 라이브러리에서 선택되는 영상 콘텐츠의 기 설정된 시간적 배열에 따라 분할 및 재배치하는 단계, (702)상기 (701)단계에서 선택되는 각 영상 데이터 간에 관련된 태그를 추출하는 단계, (703)상기 태그 추출 과정에서 의미적으로 위계성을 갖는 계층형 구조에 태그를 배치하는 단계, (704)태그를 배치하는 과정에서 의미적으로 근접한 태그를 연결하는 앵커를 구성하는 단계, (705)상기 (704)단계에서, 연결되는 태그 개념의 각 영상 데이터 간에 대한 유의어 기반 유사성 가중치를 부여하고 상기 영상 콘텐츠 생성부를 통해 생성되는 콘텐츠 조합에서 유의어에 기반하여 학습된 데이터 배열순서와 유사한 영상 데이터를 선택하여 유의어 기반의 스토리를 생성하는 단계를 포함한다.
- [0047] 이상에서 설명한 본 발명은 전술한 실시예 및 첨부된 도면에 의해 한정되는 것이 아니고, 본 발명의 기술적 사

상을 벗어나지 않는 범위 내에서 여러 가지 치환, 변형 및 변경이 가능함은 당업자에게 명백할 것이다.

부호의 설명

[0048]

100 : 영상 콘텐츠 생성부

200 : 태그 추출부

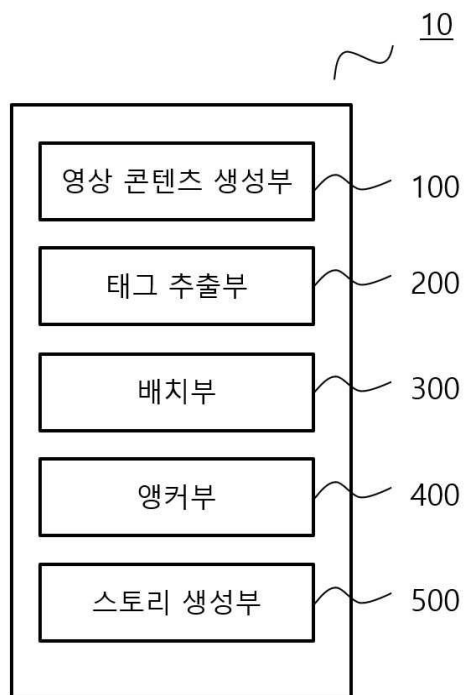
300 : 배치부

400 : 앵커부

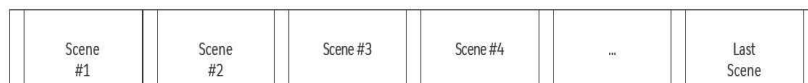
500 : 스토리 생성부

도면

도면1



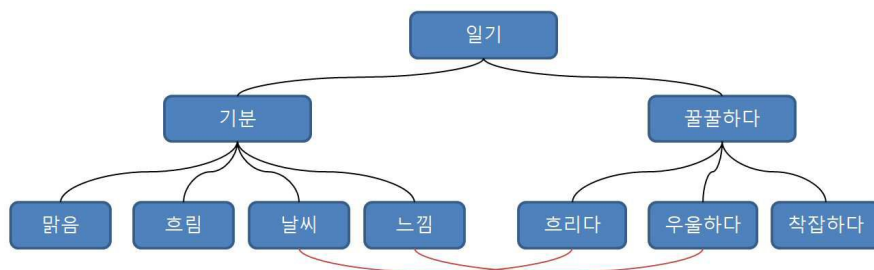
도면2



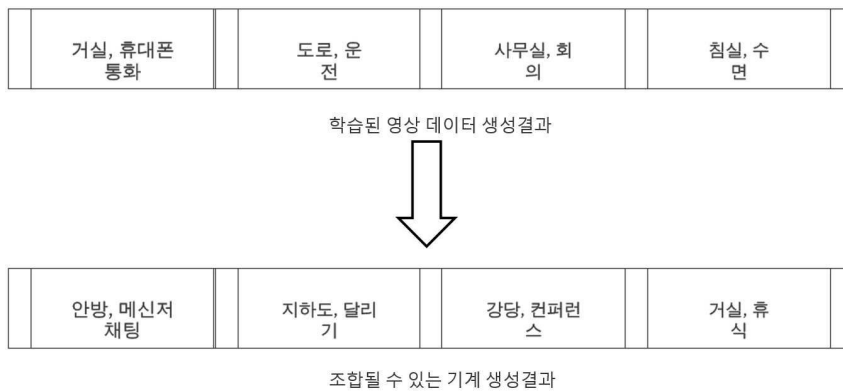
도면3

Scene #1	Scene #2	Scene #3 (후보 1)	Scene #4
		Scene #3 (후보 2)	
		Scene #3 (후보 3)	
		Scene #3 (후보 ...)	

도면4



도면5



도면6

시대적 배경 선택을 "조선시대"로 가중치를 줄 경우 생성결과

대청, 대화	저잣거리, 승마	대전, 어전회의	사랑채, 수면
--------	----------	----------	---------

도면7

