



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2016년03월11일  
(11) 등록번호 10-1602898  
(24) 등록일자 2016년03월07일

(51) 국제특허분류(Int. Cl.)  
G06F 3/14 (2006.01) G06F 17/00 (2006.01)  
(21) 출원번호 10-2014-0154168  
(22) 출원일자 2014년11월07일  
심사청구일자 2014년11월07일  
(56) 선행기술조사문헌  
KR1020140101906 A\*  
JP2010135925 A\*  
KR1020140003305 A\*  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
아주대학교산학협력단  
경기도 수원시 영통구 월드컵로 206 (원천동)  
(72) 발명자  
이경원  
서울특별시 성북구 길음로 118, 407동 704호 (길음동, 대림e편한세상아파트)  
김기남  
서울특별시 성동구 상원길 63, 107동 1301호 (성수동1가, 쌍용아파트)  
하효지  
서울시 영등포구 국제금융로7길 1, A동 707호(여의도동 수정아파트)  
(74) 대리인  
특허법인충정

전체 청구항 수 : 총 21 항

심사관 : 신현상

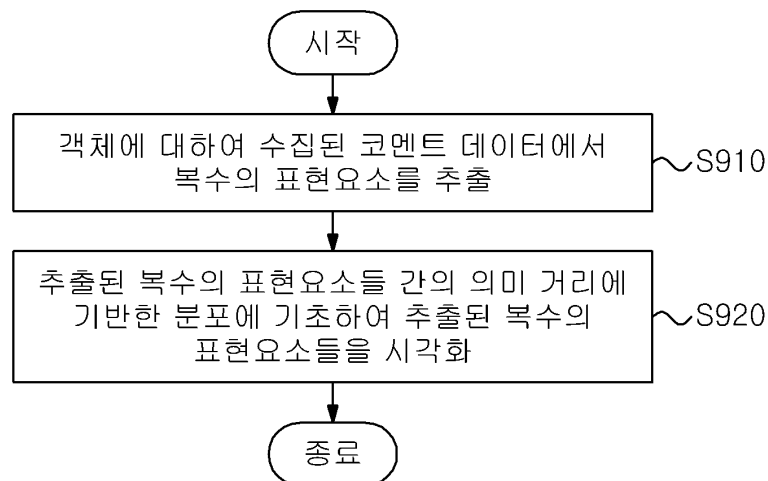
(54) 발명의 명칭 **객체의 코멘트 데이터를 이용한 데이터 시각화 방법 및 시스템**

(57) 요약

본 발명은 객체의 코멘트 데이터에서 수집된 표현요소를 이용한 시각화 방법 및 시스템에 관한 기술로, 예를 들어, 객체가 하나의 콘텐츠라면, 콘텐츠를 소비하는 소비자가 객체에 대하여 감정 또는 의견을 표현한 코멘트에 나타나는 표현요소를 시각화하는 기술에 관한 것이다.

본 발명은 객체에 대하여 사용자 코멘트에서 나타나는 사용자의 감정 또는 의견이 표현된 표현요소를 시각적으로 구현하는 것으로 기존의 객체 정보가 제공하는 제작회사, 가격 등 객관적인 정보 뿐만 아니라 사용자가 객체를 이용함으로써 느끼는 감정, 또는 의견을 표현한 표현요소를 분석하여 객체를 새로이 이용하려는 사용자에게 객체 선택의 기준으로 삼을 수 있는 정보를 제공할 수 있다.

대표도 - 도9



이 발명을 지원한 국가연구개발사업

과제고유번호 S-2013-A0403-00010

부처명 교육부

연구관리전문기관 한국연구재단

연구사업명 인문사회기초연구

연구과제명 상황별 감정어휘 분포맵을 이용한 영화추천 시스템의 시각화(2년차)

기여율 1/1

주관기관 아주대학교

연구기간 2012.05.01 ~ 2014.04.30

---

## 명세서

### 청구범위

#### 청구항 1

사용자가 선택한 객체에 대하여 수집된 코멘트 데이터에서 복수의 표현요소들을 추출하는 단계; 및

상기 추출된 복수의 표현요소들 간의 의미 거리에 기반한 분포에 기초하여 상기 추출된 복수의 표현요소들을 시각화하는 단계;

를 포함하고,

상기 시각화된 복수의 표현요소들 각각의 평면 또는 공간 상의 거리는 상기 복수의 표현요소들 각각 간의 의미 거리에 기초하여 결정되는 데이터 시각화 방법.

#### 청구항 2

제1항에 있어서,

상기 추출된 표현요소들의 상기 코멘트 데이터 내에서 추출된 빈도수를 측정하는 단계;

를 더 포함하고,

상기 표현요소들을 시각화하는 단계는

상기 추출된 표현요소들을 상기 측정된 표현요소들의 빈도수에 따라 시각화 하는 것을 특징으로 하는 데이터 시각화 방법.

#### 청구항 3

제1항에 있어서,

상기 표현요소들을 추출하는 단계 이후에,

상기 추출된 표현요소들과 기존에 추출된 표현요소들을 비교하는 단계; 및

상기 추출된 표현요소들 중 신규한 표현요소가 추가되었는지 여부를 확인하는 단계

를 더 포함하는 데이터 시각화 방법.

#### 청구항 4

제3항에 있어서,

상기 표현요소들을 시각화하는 단계는

상기 추출된 표현요소들 중 신규한 표현요소가 추가된 경우, 상기 기존에 추출된 표현요소들 중 상기 신규한 표현요소와 의미 거리가 일정 기준 이내인 하나 이상의 인접 표현요소를 결정하는 단계;

상기 결정된 하나 이상의 인접 표현요소들로부터의 의미 거리에 기반하여 상기 신규한 표현요소의 의미상 위치를 결정하는 단계;

를 포함하고,

상기 신규한 표현요소의 의미상 위치를 결정하는 단계는

상기 하나 이상의 인접 표현요소들 중 제1 인접 표현요소와 상기 신규한 표현요소의 의미 거리가 상기 하나 이상의 인접 표현요소들 중 제2 인접 표현요소와 상기 신규한 표현요소의 의미 거리보다 가까운 경우에는 상기 신규한 표현요소의 시각화되는 위치가 상기 제2 인접 표현요소의 시각화된 위치보다 상기 제1 인접 표현요소의 시각화된 위치에 더 가깝도록 결정되는 데이터 시각화 방법.

**청구항 5**

제1항에 있어서,  
상기 표현요소들을 추출하는 단계 이후에,  
상기 추출된 표현요소의 상기 사용자가 선택한 객체와 관련된 유효성을 판단하는 단계; 및  
상기 사용자가 선택한 객체에 대하여 상기 추출된 표현요소가 유효하지 않은 경우, 상기 유효하지 않은 표현요소를 제거하는 단계;  
를 더 포함하는 데이터 시각화 방법.

**청구항 6**

제5항에 있어서,  
상기 표현요소들을 추출하는 단계 이후에,  
상기 추출된 표현요소들의 상기 코멘트 데이터 내에서 추출된 빈도수를 측정하는 단계;  
를 더 포함하고,  
상기 추출된 표현요소의 유효성을 판단하는 단계는  
상기 추출된 표현요소의 측정된 빈도수를 반영하여 상기 추출된 표현요소의 유효성을 판단하는 것을 특징으로 하는 데이터 시각화 방법.

**청구항 7**

제1항에 있어서,  
상기 표현요소들을 추출하는 단계 이후에,  
상기 추출된 표현요소가 상기 사용자가 선택한 객체 이외의 다른 객체에서 추출되는 빈도수를 식별하는 단계; 및  
상기 추출된 표현요소가 상기 사용자가 선택한 객체 이외의 다른 객체 중 일정 수 이상의 객체에서 일정 빈도수 이상으로 추출되었는지 여부를 판정하는 단계;  
를 더 포함하는 데이터 시각화 방법.

**청구항 8**

제1항에 있어서,  
상기 표현요소들을 추출하는 단계 이후에,  
상기 표현요소가 추출된 빈도수를 측정하는 단계; 및  
상기 측정된 빈도수에 따라 상기 측정된 빈도수에 가중치를 부여하여, 상기 측정된 빈도수를 조정하는 단계;  
를 더 포함하고,  
상기 표현요소들을 시각화하는 단계는  
상기 조정된 빈도수를 반영하여 상기 표현요소들을 시각화하는 것을 특징으로 하는 데이터 시각화 방법.

**청구항 9**

제1항에 있어서,  
상기 표현요소들을 추출하는 단계 이후에,  
상기 코멘트 데이터에서 상기 표현요소가 추출된 빈도수를 측정하는 단계;  
상기 사용자가 선택한 객체에서 상기 표현요소가 출현하는 빈도수와 상기 측정된 빈도수를 비교하는 단계; 및

상기 사용자가 선택한 객체에서 상기 표현요소가 출현하는 빈도수와 상기 측정된 빈도수의 비교 결과에 따라, 상기 표현요소가 추출된 빈도수에 가중치를 부여하여 상기 측정된 빈도수를 조정하는 단계를 더 포함하는 데이터 시각화 방법.

**청구항 10**

제2항에 있어서,

상기 표현요소들을 추출하는 단계는

표준화된 표현요소가 미리 저장된 데이터베이스 내에, 상기 추출된 표현요소가 저장되어 있는지 여부를 탐색하는 단계; 및

상기 추출된 표현요소가 상기 데이터베이스 내에 저장되어 있지 않으면, 상기 추출된 표현요소와 가장 의미 거리가 가까운 상기 데이터베이스 상의 표준화된 표현요소를 상기 추출된 표현요소의 대표 표현요소로 식별하는 단계;

를 포함하고,

상기 빈도수를 측정하는 단계는

상기 추출된 표현요소가 상기 코멘트 데이터 내에서 추출된 빈도수를 상기 식별된 대표 표현요소가 상기 코멘트 데이터 내에서 추출된 빈도수에 합산하고,

상기 표현요소들을 시각화하는 단계는

상기 대표 표현요소를 상기 합산된 빈도수를 반영하여 시각화하는 것을 특징으로 하는 데이터 시각화 방법.

**청구항 11**

제2항에 있어서,

상기 표현요소들을 시각화하는 단계는,

상기 표현요소들을 포함하며 상기 표현요소들 간의 의미 거리를 반영하여 시각화되는 다차원척도 분석 지도 (Multi-Dimensional Scaling map: MDS map)를 배경으로 하여 상기 표현요소들을 시각화하는 것을 특징으로 하는 데이터 시각화 방법.

**청구항 12**

제1항 내지 제11항 중 어느 한 항의 방법을 실행하기 위한 프로그램이 기록되어 있는 것을 특징으로 하는 컴퓨터에서 판독 가능한 기록 매체.

**청구항 13**

사용자가 선택한 객체에 대한 코멘트 데이터를 저장하는 스토리지 장치;

상기 저장된 코멘트 데이터에서 복수의 표현요소들을 추출하는 표현요소 추출부; 및

상기 추출된 복수의 표현요소들 간의 의미 거리에 기반한 분포에 기초하여 상기 추출된 복수의 표현요소들을 시각화하는 시각화부;

를 포함하고,

상기 시각화된 복수의 표현요소들 각각의 평면 또는 공간 상의 거리는 상기 복수의 표현요소들 각각 간의 의미 거리에 기초하여 결정되는 데이터 시각화 시스템.

**청구항 14**

제13항에 있어서,

상기 추출된 표현요소들의 상기 코멘트 데이터 내에서 추출된 빈도수를 측정하는 빈도수 측정부;

를 더 포함하고,

상기 시각화부는

상기 추출된 표현요소들을 상기 측정된 표현요소들의 빈도수에 따라 시각화 하는 것을 특징으로 하는 데이터 시각화 시스템.

**청구항 15**

제13항에 있어서,

상기 추출된 표현요소들과 기존에 추출된 표현요소들을 비교하는 표현요소 비교부; 및

상기 추출된 표현요소들 중 신규한 표현요소가 추가되었는지 여부를 확인하는 신규 표현요소 확인부;

를 더 포함하는 데이터 시각화 시스템.

**청구항 16**

제13항에 있어서,

상기 추출된 표현요소의 상기 사용자가 선택한 객체와 관련된 유효성을 판단하고, 상기 사용자가 선택한 객체에 대하여 상기 추출된 표현요소가 유효하지 않은 경우, 상기 유효하지 않은 표현요소를 제거하는 유효성 판단부;

를 더 포함하는 데이터 시각화 시스템.

**청구항 17**

제16항에 있어서,

상기 추출된 표현요소들의 상기 코멘트 데이터 내에서 추출된 빈도수를 측정하는 빈도수 측정부;

를 더 포함하고,

상기 유효성 판단부는

상기 추출된 표현요소의 측정된 빈도수를 반영하여 상기 추출된 표현요소의 유효성을 판단하는 것을 특징으로 하는 데이터 시각화 시스템.

**청구항 18**

제13항에 있어서,

상기 추출된 표현요소가 상기 사용자가 선택한 객체 이외의 다른 객체에서 추출되는 빈도수를 식별하여, 상기 추출된 표현요소가 상기 사용자가 선택한 객체 이외의 다른 객체 중 일정 수 이상의 객체에서 일정 빈도수 이상으로 추출되었는지 여부를 판정하는 빈도수 측정부;

를 더 포함하는 데이터 시각화 시스템.

**청구항 19**

제13항에 있어서,

상기 표현요소가 추출된 빈도수를 측정하고, 상기 측정된 빈도수에 따라 상기 표현요소의 빈도수에 가중치를 부여하여, 상기 표현요소의 빈도수를 조정하는 빈도수 조정부;

를 더 포함하고,

상기 시각화부는

상기 조정된 빈도수를 반영하여 상기 표현요소들을 시각화하는 것을 특징으로 하는 데이터 시각화 시스템.

**청구항 20**

제13항에 있어서,

상기 저장된 코멘트 데이터에서 상기 표현요소가 추출된 빈도수를 측정하고, 상기 사용자가 선택한 객체에서 상

기 표현요소가 출현하는 빈도수와 상기 측정된 빈도수를 비교하는 빈도수 측정부; 및  
 상기 사용자가 선택한 객체에서 상기 표현요소가 출현하는 빈도수와 상기 측정된 빈도수의 비교 결과에 따라,  
 상기 표현요소가 추출된 빈도수에 가중치를 부여하여 상기 표현요소가 추출된 빈도수를 조정하는 빈도수 조정부  
 를 더 포함하는 데이터 시각화 시스템.

**청구항 21**

제14항에 있어서,  
 상기 표현요소 추출부는  
 표준화된 표현요소가 미리 저장된 데이터베이스 내에, 상기 추출된 표현요소가 저장되어 있는지 여부를 탐색하  
 는 표현요소 탐색부; 및  
 상기 추출된 표현요소가 상기 데이터베이스 내에 저장되어 있지 않으면, 상기 추출된 표현요소와 가장 의미 거  
 리가 가까운 상기 데이터베이스 상의 표준화된 표현요소를 상기 추출된 표현요소의 대표 표현요소로 식별하는  
 대표 표현요소 식별부;  
 를 포함하고,  
 상기 빈도수 측정부는  
 상기 추출된 표현요소가 상기 코멘트 데이터 내에서 추출된 빈도수를 상기 식별된 대표 표현요소가 상기 코멘트  
 데이터 내에서 추출된 빈도수에 합산하고,  
 상기 시각화부는  
 상기 대표 표현요소를 상기 합산된 빈도수를 반영하여 시각화하는 것을 특징으로 하는 데이터 시각화 시스템.

**발명의 설명**

**기술 분야**

[0001] 본 발명은 객체의 코멘트 데이터를 이용한 데이터 시각화 방법 및 시스템에 관한 기술로, 보다 상세하게는 사용  
 자 또는 소비자의 감정 또는 의견이 표현된 요소를 시각화하는 기술에 관한 것이다.

[0002] 본 발명은 교육부 및 한국연구재단의 인문사회기초연구사업의 일환으로 수행한 연구로부터 도출된 것이다[과제  
 관리번호: S-2013-A0403-00010, 과제명: 상황별 감정어휘 분포맵을 이용한 영화추천 시스템의 시각화].

**배경 기술**

[0003] 일반적으로, 영화, 음악, 문학 작품, 등 콘텐츠를 소비하거나, 상품, 또는 서비스를 이용한 사용자들은 콘텐츠,  
 상품, 또는 서비스(이하 "객체"라 함)를 이용한 감정 또는 의견을 코멘트(리뷰)형식으로 표현하게 되고, 아직  
 객체를 이용하지 못한 사용자들이나 객체에 대한 정보를 얻고자 하는 사용자들은 객체를 미리 이용해본 사용자  
 들이 남긴 코멘트(리뷰)를 참고하여 정보를 얻게 된다.

[0004] 사용자는 객체에 대한 정보를 얻고자 하는데, 객체에 대한 코멘트 데이터는 텍스트에 기반하여 구성되어 있으므  
 로, 사용자가 코멘트 데이터를 참고하여 객체에 대한 정보를 얻는 데 걸리는 시간이 상당히 길게 소요되는 문제  
 점이 있었다. 특히, 객체에 대한 코멘트 데이터의 양이 방대하거나, 다수의 사용자가 코멘트를 남겼거나, 오랜  
 시간 동안 코멘트 데이터가 누적된 경우에는 사용자가 코멘트 데이터의 내용을 읽는 것만으로도 상당한 노력이  
 필요하였다.

[0005] 따라서 이러한 문제점에 대응하고자 코멘트 데이터 상의 어휘를 기반으로 하여 코멘트 또는 객체를 검색하고,  
 사용자로 하여금 코멘트 및 객체에 대한 탐색 시간을 단축하도록 할 수 있는 기술에 대한 연구가 진행되었다.

[0006] 이러한 콘텐츠에 대한 코멘트 정보를 이용하여 콘텐츠를 검색하는 방법의 일 예가 한국등록특허 제10-0917784호  
 "콘텐츠에 대한 코멘트를 기반으로 한 집단 감성 정보 검색방법 및 시스템"에 기술되어 있다.

[0007] 상기 선행기술은 인터넷 상의 각종 콘텐츠에 달린 코멘트를 수집하여 검색용 데이터베이스(이하 DB라고 한다)를  
 작성하고 이 검색용DB를 이용하여 감성적인 질의에 대해 객관적이고 신뢰할 수 있는 순위 결과를 보여 주는 검색

색 방법 및 시스템을 제공하는데 그 목적이 있다. 특히 감성적인 단어가 포함된 질의에 대하여 감성적인 단어가 코멘트 상에 나타나는 빈도를 반영하여 객체의 추천 우선 순위를 조정하는 기술이다.

[0008] 그러나 위의 선행기술은 객체에 대한 코멘트로부터 감성 단어를 검색해 내는 기술에 대해서는 언급하고 있지만, 객체에 대하여 사용자에게 기대되는 전반적인 감정 또는 의견이 효과적으로 도시되는 것은 아니고, 다수의 코멘트가 존재하는 객체가 우선적으로 추천되는 등의 한계가 있다.

[0009] 이는 위의 선행기술이 텍스트 기반의 감정단어 검색을 채택하였기 때문에 생기는 한계로서, 하나의 객체(콘텐츠, 상품 또는 서비스)에 대하여 기대되는 전반적인 감정 또는 의견을 효과적으로 보여줄 수 있는 기술의 개발이 요구된다.

### 선행기술문헌

#### 특허문헌

[0010] (특허문헌 0001) 한국등록특허 제10-0917784호 (등록일 2009.09.10)

### 발명의 내용

#### 해결하려는 과제

[0011] 본 발명은 상기와 같은 종래 기술의 문제점을 해결하고자 도출된 것으로서, 객체(콘텐츠, 상품 또는 서비스)에 대하여 기존의 사용자 코멘트에서 나타나는 사용자의 감정 또는 의견이 표현된 요소들을 시각화하는 것으로 기존의 객체 정보가 제공하는 제작회사, 가격 등 객관적인 정보뿐만 아니라 사용자가 객체를 이용하고 표현하는 감정 또는 의견을 분석하여 객체를 새로이 이용하려는 사용자에게 객체 선택의 기준으로 삼을 수 있는 정보를 제공하는 것을 목적으로 한다.

[0012] 본 발명은 하나의 객체에 대하여 감정 또는 의견이 표현된 복수의 요소들을 의미 거리(semantic distance) 기반으로 시각화함으로써, 객체에 대하여 표현된 감정 또는 의견의 전체적인 분포를 직관적으로 시각화할 수 있는 방법 및 시스템을 제공하는 것을 목적으로 한다.

[0013] 본 발명은 하나의 객체에 대하여 대표적으로 표현된 감정 또는 의견을 시각화할 수도 있지만, 복수의 표현 요소(expression element)들을 상대적인 의미 거리 기반으로 시각화함으로써, 객체에 대하여 표현된 복수의 표현 요소들 간의 상대적인 거리 및 분포를 직관적으로 인식할 수 있는 수단을 제공하는 것을 목적으로 한다.

[0014] 또한 본 발명은 텍스트에 한정하지 않고, 이모티콘이나 아이콘 등 감정 또는 의견을 표현할 수 있는 다양한 비언어적 요소를 모두 반영하여 의미 거리 기반의 분포를 시각화할 수 있는 수단을 제공하는 것을 목적으로 한다. 또한, 텍스트라는 제약에서 벗어나 자유로울 수 있기 때문에, 다양한 외국어로 표현된 의견 또는 감정까지도 망라하여 하나의 프레임 안에서 시각화할 수 있는 수단을 제공할 수도 있다.

[0015] 또한 본 발명은 다양한 경로를 통하여 얻어질 수 있는, 예를 들어, 웹사이트에서 수집할 수 있는 코멘트 데이터에서 각 감정 또는 의견이 표현된 표현요소의 빈도를 계산하여 그 결과를 쉽게 이해할 수 있도록 시각적인 그래프로 제공하는 것을 목적으로 한다.

#### 과제의 해결 수단

[0016] 상기와 같은 목적을 달성하기 위하여, 본 발명의 일 실시예에 따른 표현요소(expression element)를 시각화 하는 방법은 사용자가 선택한 객체에 대하여 수집된 코멘트 데이터에서 복수의 표현요소들을 추출하는 단계 및 상기 추출된 복수의 표현요소들 간의 의미 거리에 기반한 분포에 기초하여 상기 추출된 복수의 표현요소들을 시각화하는 단계를 포함한다.

[0017] 이때, 상기 추출된 표현요소들의 상기 코멘트 데이터 내에서 추출된 빈도수를 측정하는 단계를 더 포함하고, 상기 표현요소들을 시각화하는 단계는 상기 추출된 표현요소들을 상기 측정된 표현요소들의 빈도수에 따라 시각화하는 것을 특징으로 한다.

[0018] 또한, 상기 표현요소들을 추출하는 단계 이후에, 상기 추출된 표현요소들과 기존에 추출된 표현요소들을 비교하는 단계 및 상기 추출된 표현요소들 중 신규한 표현요소가 추가되었는지 여부를 확인하는 단계를 더 포함할 수



있다.

- [0019] 상기 표현요소들을 시각화하는 단계는 상기 추출된 표현요소들 중 신규한 표현요소가 추가된 경우, 상기 기존에 추출된 표현요소들 중 상기 신규한 표현요소와 의미 거리가 일정 기준 이내인 하나 이상의 인접 표현요소를 결정하는 단계 및 상기 결정된 하나 이상의 인접 표현요소들로부터의 의미 거리에 기반하여 상기 신규한 표현요소의 의미상 위치를 결정하는 단계를 포함한다.
- [0020] 또한, 상기 표현요소들을 추출하는 단계 이후에, 상기 추출된 표현요소의 유효성을 판단하는 단계 및 상기 사용자가 선택한 객체에 대하여 상기 추출된 표현요소가 유효하지 않은 경우, 상기 유효하지 않은 표현요소를 제거하는 단계를 포함할 수 있으며, 상기 표현요소들을 추출하는 단계 이후에, 상기 추출된 표현요소들의 상기 코멘트 데이터 내에서 추출된 빈도수를 측정하는 단계를 더 포함하고, 상기 추출된 표현요소의 유효성을 판단하는 단계는 상기 추출된 표현요소의 측정된 빈도수를 반영하여 상기 추출된 표현요소의 유효성을 판단할 수 있다.
- [0021] 상기 표현요소들을 추출하는 단계 이후에, 상기 추출된 표현요소가 상기 사용자가 선택한 객체 이외의 다른 객체에서 추출되는 빈도수를 식별하는 단계, 상기 추출된 표현요소가 상기 사용자가 선택한 객체 이외의 다른 객체 중 일정 수 이상의 객체에서 일정 빈도수 이상으로 추출되었는지 여부를 판정하는 단계를 더 포함할 수 있으며, 상기 표현요소들을 추출하는 단계 이후에, 상기 표현요소가 추출된 빈도수를 측정하는 단계 및 상기 측정 추출된 빈도수에 따라 상기 측정된 표현 요소의 빈도수에 가중치를 부여하여, 상기 측정된 빈도수를 조정하는 단계를 더 포함하고, 상기 표현요소들을 시각화하는 단계는 상기 조정된 빈도수를 반영하여 상기 표현요소들을 시각화하는 것을 특징으로 할 수 있다.
- [0022] 또한, 상기 표현요소들을 추출하는 단계 이후에, 상기 코멘트 데이터에서 상기 표현요소가 추출된 빈도수를 측정하는 단계, 상기 사용자가 선택한 객체에서 상기 표현요소가 출현하는 빈도수와 상기 측정된 빈도수를 비교하는 단계 및 상기 사용자가 선택한 객체에서 상기 표현요소가 출현하는 빈도수와 상기 측정된 빈도수의 비교 결과에 따라, 상기 표현요소가 추출된 빈도수에 가중치를 부여하여 상기 측정된 빈도수를 조정하는 단계를 더 포함할 수 있다.
- [0023] 상기 표현요소들을 추출하는 단계는 표준화된 표현요소가 미리 저장된 데이터베이스 내에, 상기 추출된 표현요소가 저장되어 있는지 여부를 탐색하는 단계 및 상기 추출된 표현요소가 상기 데이터베이스 내에 저장되어 있지 않으면, 상기 추출된 표현요소와 가장 의미 거리가 가까운 상기 데이터베이스 상의 표준화된 표현요소를 상기 추출된 표현요소의 대표 표현요소로 식별하는 단계를 포함하고, 상기 빈도수를 측정하는 단계는 상기 추출된 표현요소가 상기 코멘트 데이터 내에서 추출된 빈도수를 상기 식별된 대표 표현요소가 상기 코멘트 데이터 내에서 추출된 빈도수에 합산하고, 상기 표현요소들을 시각화하는 단계는 상기 대표 표현요소를 상기 합산된 빈도수를 반영하여 시각화하는 것을 특징으로 할 수 있다.
- [0024] 또한, 상기 표현요소들을 시각화하는 단계는, 상기 표현요소들을 포함하는 다차원척도 분석 지도(Multi-Dimensional Scaling map: MDS map)를 배경으로 하여 상기 표현요소들을 시각화하는 것을 특징으로 한다.
- [0025] 본 발명의 일 실시예에 따른 표현요소를 시각화 하는 시스템은 사용자가 선택한 객체에 대한 코멘트 데이터를 저장하는 스토리지 장치 상기 저장된 코멘트 데이터에서 복수의 표현요소들을 추출하는 표현요소 추출부 및 상기 추출된 복수의 표현요소들 간의 의미 거리에 기반한 분포에 기초하여 상기 추출된 복수의 표현요소들을 시각화하는 시각화부를 포함한다.

**발명의 효과**

- [0026] 본 발명에 따르면 객체를 이미 사용해본 사람들이 객체를 사용하고 느낀 표현요소(expression element)들을 시각화 그래프를 통해 확인함으로써 해당 객체를 사용하기 전 사용자들이 느끼는 표현요소들을 직관적으로 분석이 가능하여, 객체를 사용하는 사람의 입장에서는 해당 객체에 대해 사람들이 어떠한 감정을 가지는지, 또는 객체를 선택하는 사용자가 원하는 객체를 손쉽게 선택할 수 있는 효과가 있다.
- [0027] 또한, 본 발명을 통하여 생성된 시각화 그래프를 스크립트 프로그램으로 웹 사이트상으로 제공됨으로써 많은 사용자들을 대상으로 동시에 제공될 수 있다.
- [0028] 또한, 본 발명은 별도의 프로그램 설치 없이 브라우저상의 웹 페이지를 통해 제공할 수 있으므로 코멘트 데이터가 갱신될 때마다 개발자가 새로운 데이터 관리나 배포의 절차 없이 실시간으로 분석 결과를 사용자는 제공 받을 수 있다.

- [0029] 또한, 본 발명은 정부 또는 공공기관이 어떠한 정책이나 계획을 발표하여 사람들이 이에 대하여 인터넷을 통해 의사를 표현하는 경우, 정책에 대한 여론의 반응을 직관적으로 확인할 수도 있다.
- [0030] 또한, 인터넷에서 발생한 기업에서 발생하는 각종 사고 또는 외부에서 발생하는 기업에 대한 여론을 수집하고 이를 분석한 여론 반응의 변화를 실시간으로 파악할 수도 있으며, 기업은 이러한 정보를 이용하여 사내 위기관리 프로토콜을 통해 대응할 수도 있다.
- [0031] 또한, 하나의 객체에 대하여 감정 또는 의견이 표현된 복수의 요소들을 의미 거리(semantic distance) 기반으로 시각화함으로써, 객체에 대하여 표현된 감정 또는 의견의 전체적인 분포를 직관적으로 사용자에게 제공할 수 있다.
- [0032] 또한, 하나의 객체에 대하여 대표적으로 표현된 감정 또는 의견을 시각화할 수도 있지만, 복수의 표현 요소들을 상대적인 의미 거리 기반으로 시각화함으로써, 객체에 대하여 표현된 복수의 표현 요소들 간의 상대적인 거리 및 분포를 직관적으로 인식할 수 있다.
- [0033] 또한, 텍스트에 한정하지 않고, 이모티콘이나 아이콘 등 감정 또는 의견을 표현할 수 있는 다양한 비언어적 요소를 모두 반영하여 의미 거리 기반의 분포를 시각화할 수 있으며, 텍스트라는 제약에서 벗어나 자유로울 수 있기 때문에, 다양한 외국어로 표현된 의견 또는 감정까지도 망라하여 하나의 프레임 안에서 시각화할 수도 있다.

**도면의 간단한 설명**

- [0034] 도 1은 본 발명의 일 실시예에 따른 감정어휘 분포맵 제작을 위해 선정된 감정어휘를 나타낸 도면이다.
- 도 2는 도 1에 도시된 각 감정어휘의 TF-IDF 스코어의 최대치를 나타낸 도면이다.
- 도 3은 도 1에 도시된 각 감정어휘들 중 최종 선정된 36개의 감정어휘를 나타낸 도면이다.
- 도 4는 본 발명의 일 실시예에 따른 감정어휘 분포맵을 나타낸 도면이다.
- 도 5 내지 도 8은 본 발명의 일 실시예에 따른 객체의 코멘트 데이터에서 추출한 표현요소를 히트맵(Heat-map)형태로 나타낸 도면이다.
- 도 9는 본 발명의 일 실시예에 따른 표현요소를 시각화 하는 방법에 대한 순서도를 나타낸 도면이다.
- 도 10은 본 발명의 일 실시예에 따른 표현요소들의 빈도수에 따라 측정된 표현요소를 시각화 하는 방법에 대한 순서도를 나타낸 도면이다.
- 도 11은 본 발명의 일 실시예에 따른 신규 어휘가 추가되었는지를 확인하는 순서도를 나타낸 도면이다.
- 도 12는 본 발명의 일 실시예에 따른 신규한 표현요소가 추가되는 경우에 따른 순서도를 나타낸 도면이다.
- 도 13은 본 발명의 일 실시예에 따른 표현요소의 유효성을 판단하는 과정을 나타낸 도면이다.
- 도 14는 본 발명의 일 실시예에 따른 표현요소의 유효성을 표현요소의 빈도수를 기준을 판단하는 과정을 나타낸 도면이다.
- 도 15는 본 발명의 일 실시예에 따른 표현요소의 유효성을 판단하는 과정을 구체화하여 나타낸 순서도이다.
- 도 16은 본 발명의 일 실시예에 따른 특정 표현요소가 집중되어 있는 경우 표현요소의 영향력을 조절하는 방법에 대한 순서도를 나타낸 도면이다.
- 도 17은 본 발명의 일 실시예에 따른 특정 표현요소가 특정 객체에서 실제로 나타나는 빈도수가 낮은 경우 가중치를 부여하는 방법에 대한 순서도를 나타낸 도면이다.
- 도 18은 본 발명의 일 실시예에 따른 표현요소를 미리 저장된 표준형의 표현요소로 매핑하고 빈도수를 측정하는 방법에 대한 순서도를 나타낸 도면이다.
- 도 19는 본 발명의 일 실시예에 따른 표현요소를 시각화 하는 시스템을 나타낸 도면이다.
- 도 20은 본 발명의 일 실시예에 따른 신규 표현요소를 확인하여 표현요소를 시각화 하는 시스템을 나타낸 도면이다.
- 도 21은 본 발명의 일 실시예에 따른 표현요소의 빈도수를 측정 및 조정하여 표현요소를 시각화 하는 시스템을

나타낸 도면이다.

도 22는 본 발명의 일 실시예에 따른 표현요소 추출부를 상세히 나타낸 도면이다.

도 23 내지 도 27는 본 발명의 일 실시예에 따라 각기 다른 시각화 방법을 나타낸 도면이다.

도 28는 본 발명의 일 실시예에 따른 히트맵(Heat-map) 시각화 방법을 3차원으로 응용하여 나타낸 도면이다.

도 29는 본 발명의 일 실시예에 따른 표현요소 시각화 방법을 등고선으로 나타낸 도면이다.

도 30은 도 29에 도시한 등고선 맵을 3차원으로 나타낸 도면이다.

도 31 내지 도 33은 본 발명의 일 실시예에 따른 의미지도를 기반으로 한 활용 방법을 나타낸 도면이다.

**발명을 실시하기 위한 구체적인 내용**

[0035] 상기 목적 외에 본 발명의 다른 목적 및 특징들은 첨부 도면을 참조한 실시 예에 대한 설명을 통하여 명백히 드러나게 될 것이다.

[0036] 본 발명의 바람직한 실시예를 첨부된 도면들을 참조하여 상세히 설명한다. 본 발명을 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략한다.

[0037] 그러나, 본 발명이 실시예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.

[0038] 도 1은 본 발명의 일 실시예에 따른 감정어휘 분포맵 제작을 위해 선정된 감정어휘를 나타낸 도면이다.

[0039] 본 발명은 객체의 코멘트 데이터에서 수집된 표현요소(expression element)를 이용한 시각화 방법 및 시스템에 관한 것으로서, 객체는 사용자가 선택한 영화, 상품, 소셜, 게임, 여행 등 사람의 감정이 포함되는 객체를 의미하며, 사람의 감정이 포함되는 객체에 대한 코멘트 또는 리뷰로부터 나타나는 감정을 시각화 할 수 있다.

[0040] 본 발명의 일 실시예로서, 객체는 영화로 한정하여 영화에 대한 코멘트 데이터를 이용한 시각화 방법 및 시스템으로도 설명할 수 있다.

[0041] 영화의 코멘트 데이터는 사용자에게 따라서 구축된 웹 서비스 통해 수집된 데이터를 이용할 수 있으며, 또는 대형 포털 및 동호회 게시판에 축적되는 코멘트 데이터를 프로그램을 이용하여 개별적으로 수집할 수도 있다.

[0042] 본 발명의 일 실시예로서, 영화에 대한 코멘트 데이터로부터 사용자의 감정을 포함하고 있는 감정어휘 수집을 자동화하기 위하여 데이터를 수집할 수 있는 웹 크롤러를 이용할 수 있으며, 크롤러는 대형 포털(네이버, 다음 등) 영화 홈페이지에서 특정 영화의 댓글과 코멘트들을 정제되지 않은 데이터 형태로 수집하고, 수집된 데이터를 연구에 사용 가능한 데이터로 가공할 수 있으며, 정제된 데이터를 분석하여 감정어휘를 추출할 수 있다. 이에 따라 크롤러를 통하여 수집되는 감정어휘는 영화를 보는 상황과 연결시켜 추후에 사용자의 이용 동기에 맞는 영화를 추천할 수도 있다.

[0043] 영화에 나타난 감정어휘의 빈도를 시각화하기 위해서는 2차원 평면상에 각 감정어의 위치를 지정해야 한다. 이를 위해 감정어간의 상관관계를 이용하여 2차원 상의 위치 좌표를 도출할 수 있다. 감정어휘의 분포맵을 제작하기 위하여 한덕웅, 강혜자(2000)의 한국어 정서 용어들의 적절성과 경험 빈도에 대한 연구를 참고하여 834개의 정서용어 중에서 영화를 봤을 때 느낄 수 있는 감정어휘만을 분류하였다. 이때, 아주대학교의 국어국문학과박사 전문가 1명과 본 발명의 발명자 2명이 함께 서로 의견취합이 가능한 감정어휘만을 골라 최종 100개의 감정어휘를 선별하였다.

[0044] 또한, 전문가 분석을 통한 감정어휘 선별작업 이외에도 영화를 시청하였을 때, 사용자들이 가장 많이 느끼는 감정어휘를 선별하기 위해 선정된 100개의 감정어휘를 토대로 최종감정어휘 선정을 위한 서베이(survey)를 실시하였다. 서베이는 아주대학교의 미디어학과 학생 30명을 대상으로 영화를 봤을 때 느낄 수 있는 감정에 대한 간단한 개념 설명을 거친 뒤에, 전문가 분석을 통해 얻어진 100개의 감정어휘에 대해서 영화를 보는 상황일 때 해당 감정어휘를 느낄 수 있는 정도가 어떻게 되는지를 조사 하였다. 실제 설문에서는 ‘여러분이 지금까지 보신 여러 장르의 영화 스토리를 생각하신 뒤 해당 영화를 봤을 때 다음에 제시된 감정어휘들을 느끼는 정도가 어떠한지 답하시기 바랍니다.’ 와 같이 시작하였으며, 각 감정어휘에 대해서 리커트 7점 척도(Likert-type scale)로 응답하게 하여 1점은 ‘전혀 관련 없다.’ 를 의미하고 7점은 ‘매우 관련 있다.’ 를 의미하도록 질문하였다.

- [0045] 본 연구에서는 사용자의 이용 동기를 이용해 영화를 추천한다는 취지에 부합하고자 영화를 봤을 때 가장 잘 느낄 수 있는 감정어휘를 수집하기 위해 전문가 분석 및 사용자 서베이를 실시하였으며, 사용자들이 설문한 리커트 7점 척도 정보를 토대로 관련성이 높은 감정어휘를 선별하기 위해 평균분석을 통해서 평균이 상대적으로 낮은 감정어휘(4.00 ‘보통이다.’를 뜻하는 수치 이하) 32개를 추가적으로 제거하여 영화 추천에 적합한 68개의 감정어휘를 선별하였다.
- [0046] 도 1은 이렇게 선별된 영화 추천에 적합한 68개의 감정어휘를 나타낸 도면이다.
- [0047] 도 2는 도 1에 도시된 각 감정어휘의 TF-IDF 스코어의 최대치를 나타낸 도면이다.
- [0048] 도 2는 도 1에 설명된 68개의 감정어휘에 실제 영화데이터를 비교하여 영향력이 미미한 감정어휘를 추가로 제거하기 위하여, 영화의 코멘트 또는 리뷰에 나타나는 각 감정어휘의 TF-IDF 스코어를 도출하고, 각 감정어휘에 나타날 수 있는 TF-IDF 스코어의 최대치를 나타낸 도면이다.
- [0049] 이때, TF(단어 빈도수, term frequency)는 특정한 단어가 문서 내에 얼마나 자주 등장하는지 나타내는 값을 의미하며, DF(Document Frequency)는 특정 단어가 나타난 문서의 수를 의미하며, 이 값의 역수를 IDF(inverse document frequency)라고 한다.
- [0050] 도 3은 도 1에 도시된 각 감정어휘들 중 최종 선정된 36개의 감정어휘를 나타낸 도면이다.
- [0051] 도 2에 도시된 도면은 TF-IDF 스코어가 도출된 각 감정어휘들 중에서 '경악하다'의 경우 모든 영화에서 TF-IDF 스코어의 비율이 0.8% 이하로 나타났으며, 반면에 '달콤하다'의 경우에는 적어도 한 개의 영화에서는 TF-IDF 스코어의 비율이 42%에 달하는 것을 의미한다.
- [0052] 이때, 도 3은 TF-IDF 스코어의 비율이 10% 미만인 감정어휘를 제거하고 최종적으로 선택된 36개의 감정어휘를 나타낸 도면이다.
- [0053] 도 4는 본 발명의 일 실시예에 따른 감정어휘 분포맵을 나타낸 도면이다.
- [0054] 도 3에 도시한 최종 군집화된 36개의 감정어휘를 2차원 평면에 각 감정어휘 간의 의미 거리를 도출하기 위하여 36개의 감정어휘를 바탕으로 유사하거나 상이한 감정어휘 간의 거리를 측정하여 상관관계를 분석한 다음 다차원 척도 분석(Multi-Dimensional Scaling: MDS)을 이용할 수 있다.
- [0055] 이때, 다차원척도 분석이란 개체들 간의 상대적인 거리를 계산하여 사람이 인지 할 수 있는 평면상에 상대적인 거리로 나타내는 통계와 연관된 기술로 정보 시각화에서는 데이터 내의 유사성 및 비유사성을 측정하기 위한 배경기술이다.
- [0056] 다차원척도법의 장점은 상대적인 거리만을 알고 있는 개체들의 의미 지도를 작성할 수 있으며 물리적인 거리뿐만 아니라 심리적인 거리에 근거하여서도 의미 지도를 작성할 수 있다.
- [0057] 본 발명의 일 실시예에 따른 다차원척도 분석을 위해 경기도 및 서울 소재 대학교 20대 남학생 11명, 여학생 9명으로 총 20명을 실험 대상으로 하여 36개의 감정어휘에 대해 의미상 거리 서베이를 실시하였으며, 서베이는 가로축 세로축 36개의 감정어휘를 배치한 설문지를 만들고(68x68), 감정어휘간의 거리가 가장 가깝다고 느껴지면 3점, 가장 멀다고 느껴지면 -3점을 주는 방식의 리커트 척도를 이용하여 체크하는 형식으로 구성하였다. 20명이 기록한 데이터를 바탕으로 다양한 네트워크 분석기법이 활용 가능한 UCINET 프로그램을 사용 하였고, 이에 따라 영화 36개 감정 어휘 간의 의미상의 거리에 선정된 68개의 감정어휘를 기반한 Metric MDS를 도 4에 나타내었다.
- [0058] 그 결과, X축의 양(+)의 방향으로는 대표어 “Happy”, “Surprise”와 관련된 감정어휘가 분포되었으며, X축의 음(-)의 방향으로는 대표어 “Anger”, “Disgust”와 관련된 감정어휘가 분포되었다. 그리고 Y축의 양(+)의 방향으로는 대표어 “Fear”, “Surprise”와 관련된 감정어휘가 분포되었으며, Y축의 음(-)의 방향으로는 대표어 “Sad”, “Boring”과 관련된 감정어휘가 분포되었다.
- [0059] 이에 따라, 감정어휘의 성격 상 X축의 양(+)의 방향은 긍정적인 감정어휘들이 분포되었고, X축의 음(-)의 방향으로서는 부정적인 감정어휘들이 분포됨을 알 수 있다.
- [0060] 또한, Y축의 양(+)의 방향은 동적인(감정을 느낄 때 비교적 큰 제스처를 취할 수 있는) 감정어휘들이 분포되었고, Y축의 음(-)의 방향으로서는 정적인(감정을 느낄 때 비교적 작은 제스처를 취할 수 있는) 감정어휘들이 분포됨을 알 수 있다.

- [0061] 그리고 'Happy', 'Sad', 'Anger', 'Fear', 'Disgust', 'Boring' 대표어와 관련된 어휘들은 각각의 단어가 뚜렷하게 군집이 되는 것을 볼 수 있는데, 대표어 'Surprise' 에 대해서는 'Happy' 대표어 군집과 'Fear' 대표어 군집에 나뉘어서 분포한다는 것을 알 수 있다. 이는 사용자가 영화를 봤을 때 'بخ찬 기쁨으로 인해 놀라운 감정이 생기는 경우' 와 '갑작스럽게 등장하는 공포로 인해서 놀라운 감정이 생기는 경우' 가 지배적이기 때문인 것으로 해석할 수 있다.
- [0062] 도 5 내지 도 8은 본 발명의 일 실시예에 따른 객체의 코멘트 데이터에서 추출한 표현요소를 히트맵(Heat-map)형태로 나타낸 도면이다.
- [0063] 도 1 내지 도 4에서 설명한 영화에 대한 코멘트 데이터로부터 추출된 감정어휘를 시각화하기 위해서는 MDS Map을 구성하고 있는 감정어휘의 빈도수가 필요하다. 상위 과정을 통해 선별된 코멘트 데이터와 감정어휘사전을 비교하여 각 영화에서의 감정어휘 빈도수를 측정한다.
- [0064] 또한, 영화의 성격과 관계없이 자주 등장하는 특정 어휘의 가중치를 낮추기 위해 TF-IDF 스코어를 계산하여 수치를 조정한다. 최종적으로 선별된 각 감정어휘의 TF-IDF 스코어를 이용하여 시각화 할 수 있다.
- [0065] 최종 시각화 그래프는 감정어휘의 MDS Map을 배경으로 하고, 사각형의 작은 셀로 구성된 히트맵(Heat-map)으로 나타낼 수 있다. 이때, 모든 셀은 0의 수치로 초기화 되어 있으며, 해당 셀에 위치한 감정어휘의 TF-IDF스코어에 따라 셀의 수치가 증가한다. 셀이 가진 수치가 높아질수록 다른 색으로 변함으로써 해당 감정어휘 TF-IDF스코어의 높고 낮음을 확인 할 수 있다. 또한 수치가 올라간 셀은 주위 셀의 수치에 영향을 미침으로써 그래프의 모습은 지형도의 모습을 띄게 된다.
- [0066] 도 5는 영화 '설국열차'에 대한 관람객들의 코멘트 데이터에 나타나는 감정어휘의 분포를 시각화한 그래프이다. 도 5에 도시된 것처럼 영화 '설국열차'에서 관객들은 재미있고 대단하다는 반응을 보이고 있으며, 안타깝고 지루하다는 감정 또한 높은 빈도를 보이고 있다. 실제로 '설국열차'에 대한 영화 코멘트를 살펴보면 영화에 대해 실망한 관객들의 리뷰가 많은 것을 볼 수 있다.
- [0067] 도 6은 영화 '극락도 살인사건'을 히트맵(Heat-map) 형태로 시각화한 도면이며, 공포영화인 '극락도 살인사건'에서 나타난 관람객들의 감정 중 가장 높은 감정어휘가 '놀라다'이며, 그 밖에 공포와 관련된 감정어휘의 빈도가 높게 나타난 것을 확인할 수 있다.
- [0068] 도 7은 영화 '돈 크라이 마미'를 히트맵(Heat-map) 형태로 시각화한 도면이며, 실제 범죄사건을 모티브로 제작된 '돈 크라이 마미'에 대한 관람객의 감정어휘의 분포는 '분노하다'와 '격분하다'에 많은 빈도수를 나타내는 것을 확인할 수 있다.
- [0069] 도 8은 영화 '위낭소리'를 히트맵 형태로 시각화한 도면이며, '위낭소리'의 경우 관객들의 감정은 '슬프다'와 '감동적이다'에 높은 빈도를 보이는 것으로 나타났다.
- [0070] 위와 같은 예시를 통하여 영화를 시청한 후 작성된 코멘트에서 수집된 코멘트 데이터가 영화의 장르 및 스토리 특성에 부합하여 감정어휘 패턴이 나타남을 알 수 있다.
- [0071] 본 발명의 일 실시예로서, 영화에 대한 코멘트 데이터를 이용하여 감정어휘를 추출하고 이를 시각화 하는 방법에 대하여 기술하였으나, 영화 뿐만 아니라 인간관계와 사회관계에서 사람들이 갖거나 드러내는 사고, 의도, 평가, 의견, 논증, 반박 같은 인지활동과 감정, 정서, 욕망, 태도 등의 감성적 반응을 대상으로 할 수도 있다.
- [0072] 또한, 본 발명이 적용될 수 있는 대상은 감성적인 부분에선 인간의 감정, 정서, 욕망, 태도 등을 포함하며, 인지 부분에선 사고, 의도, 평가, 의견, 논증, 반박 등을 포함한다. 또한, 관계 부분에선 문화콘텐츠, 인간관계(소통, 갈등), 사회관계(다문화 Homo hundred 등), 테크놀로지와의 관계(문화지체 등)을 포함한다.
- [0073] 도 9는 본 발명의 일 실시예에 따른 표현요소를 시각화 하는 방법에 대한 순서도를 나타낸 도면이다.
- [0074] 표현요소를 시각화 하는 방법은 사용자가 선택한 객체에 대하여 수집된 코멘트 데이터에서 복수의 표현요소들을 추출하고(S910), 추출된 복수의 표현요소들 간의 의미 거리에 기반한 분포에 기초하여 추출된 복수의 표현요소들을 시각화(S920)한다.
- [0075] 이때, 객체에 대한 코멘트 데이터는 영화에 대한 리뷰, 상품에 대한 상품평, 소설 리뷰, 게임 리뷰, 여행 리뷰, 서비스에 대한 평가 등 사람들의 감정이 포함되는 모든 코멘트 데이터를 의미한다.
- [0076] 또한, 표현요소는 코멘트 데이터에서 추출되는 사람들의 감정을 나타내는 단어, 문단, 이모티콘 등을 포함한다.

- [0077] 또한, 복수의 표현요소들간의 의미 거리에 기반한 분포에 기초하여 시각화 하는 방법은 도 1 내지 도 4에서 설명된 다차원척도 분석 지도(Multi-Dimensional Scaling map: MDS map)을 기반으로 히트맵(Heat-map) 형태 또는 등고선 등의 모양으로 나타낼 수 있다.
- [0078] 도 10은 본 발명의 일 실시예에 따른 표현요소들의 빈도수에 따라 측정된 표현요소를 시각화 하는 방법에 대한 순서도를 나타낸 도면이다.
- [0079] 표현요소를 시각화 하는 방법은 사용자가 선택한 객체에 대하여 수집된 코멘트 데이터에서 복수의 표현요소들을 추출하고(S910), 추출된 표현요소들의 코멘트 데이터 내에서 추출된 빈도수를 측정하여(S930), 추출된 복수의 표현요소들 간의 의미 거리에 기반한 분포에 기초하여 추출된 복수의 표현요소들을 시각화(S920)할 수 있다. 이때, 추출된 표현요소들을 측정된 표현요소들의 빈도수에 따라 표현요소들을 포함하는 다차원척도 분석 지도(Multi-Dimensional Scaling map: MDS map)를 배경으로 하여 히트맵(Heat-map), 등고선 등의 모양으로 시각화 할 수 있다.
- [0080] 또한, 표현요소들을 추출하여, 표현요소들의 빈도수를 측정할 때 표준화된 표현요소가 표준형이 아닌 경우, 사전 상에 저장된 표준형의 표현요소로 매핑하고, 그 매핑된 사전 상의 표준형의 표현요소 기준으로 각 객체의 코멘트 데이터에서의 빈도수를 측정할 수 있다.
- [0081] 도 11은 본 발명의 일 실시예에 따른 신규 어휘가 추가되었는지를 확인하는 순서도를 나타낸 도면이다.
- [0082] 도 9에 도시된 표현요소를 시각화 하는 방법은 사용자가 선택한 객체에 대하여 수집된 코멘트 데이터에서 복수의 표현요소들을 추출하고(S910) 추출된 표현요소들과 기존에 추출된 표현요소들을 비교한다(S1110). 이후, 추출된 표현요소들 중 신규한 표현요소가 추가되었는지 여부를 확인하여(S1120), 추출된 복수의 표현요소들 간의 의미 거리에 기반한 분포에 기초하여 추출된 복수의 표현요소들을 시각화(S920)할 수 있다. 이 때 신규한 표현요소의 의미를 찾는 과정은 문맥 기반 분석(context-based analysis) 등의 기법을 통하여 실행될 수 있다.
- [0083] 도 12는 본 발명의 일 실시예에 따른 신규한 표현요소가 추가되는 경우에 따른 순서도를 나타낸 도면이다.
- [0084] 도 11에서 신규한 표현요소가 추가되었는지 확인(S921)하고, 신규한 표현요소가 추가되었을 경우, 기존에 추출된 표현요소들 중 신규한 표현요소와 의미 거리가 일정 기준 이내인 하나 이상의 인접 표현요소를 결정한다(S922).
- [0085] 이때 일정 기준이라 함은, 기존에 추출한 표현요소들 중 신규한 표현요소와 가장 가까운 N개를 기준으로 삼을 수도 있고, 또는 기존에 추출한 표현요소들 중 신규한 표현요소와 의미 거리가 r 이내인 경우를 기준으로 삼을 수도 있다.
- [0086] 이후, 결정된 하나 이상의 인접 표현요소들로부터의 의미 거리에 기반하여 신규한 표현요소의 의미상 위치를 결정하여(S923), 위치가 결정된 신규한 표현요소를 시각화 한다(S924).
- [0087] 이때, 신규한 표현요소와 인접 표현요소의 의미가 유사할수록 신규한 표현요소와 인접 표현요소 간의 의미 거리가 가까워지도록 가중치를 부여하여 위치를 결정할 수도 있다. 즉, 제1 인접 표현요소가 제2 인접 표현요소보다 신규한 표현요소의 의미와 유사한 경우에는, 신규한 표현요소와 제1 인접 표현요소 간의 거리가 신규한 표현요소와 제2 인접 표현요소 간의 거리보다 더 짧게 되도록 신규한 표현요소의 위치가 결정될 수 있다. 이 때, 표현요소 간의 의미 유사성은 문맥 기반 분석을 통하여 얻어질 수도 있고, 또는 다수 인원에게 대한 설문 조사 등의 다양한 방법을 이용하여 얻어질 수도 있다.
- [0088] 도 13은 본 발명의 일 실시예에 따른 표현요소의 유효성을 판단하는 과정을 나타낸 도면이다.
- [0089] 도 9에 도시된 표현요소를 시각화 하는 방법에서 사용자가 선택한 객체에 대하여 수집된 코멘트 데이터에서 복수의 표현요소들을 추출하고(S910), 추출된 표현요소의 유효성을 판단한다(S1310). 이때, 사용자가 선택한 객체에 대하여 추출된 표현요소가 유효하지 않은 경우, 유효하지 않은 표현요소를 제거한다(S1320).
- [0090] 이후, 추출된 복수의 표현요소들 간의 의미 거리에 기반한 분포에 기초하여 추출된 복수의 표현요소들 중 유효하지 않은 표현요소가 제거된 복수의 표현요소들을 시각화(S920)할 수 있다. 특정 표현요소의 유효하지 않은 것으로 판단하는 기준으로는 특정 표현요소의 의미가 다른 표현요소들과 현저히 다르거나, 특정 표현요소의 빈도가 기준값 미만으로 현저하게 적게 나타나거나, 또는 특정 표현요소가 특정 콘텐츠만이 아닌 다수의 콘텐츠에 변별력 없이 일정한 비율로 나타나는 경우(이 경우에는 진정한 리뷰라기보다는 기계적으로 반복되는 홍보, 또는 공지 사항 등일 수가 있음) 등을 들 수 있다.

- [0091] 도 14는 본 발명의 일 실시예에 따른 표현요소의 유효성을 표현요소의 빈도수를 기준을 판단하는 과정을 나타낸 도면이다.
- [0092] 도 13에서 설명된 표현요소를 시각화 하는 방법에서 사용자가 선택한 객체에 대하여 수집된 코멘트 데이터에서 복수의 표현요소들을 추출하고(S910), 추출된 표현요소들의 코멘트 데이터 내에서 추출된 빈도수를 측정한다(S1410).
- [0093] 이후, 표현요소의 빈도수를 이용하여 추출된 표현요소의 유효성을 판단한다(S1310). 이때, 사용자가 선택한 객체에 대하여 추출된 표현요소가 유효하지 않은 경우, 유효하지 않은 표현요소를 제거한다(S1320).
- [0094] 도 15는 본 발명의 일 실시예에 따른 표현요소의 유효성을 판단하는 과정을 구체화하여 나타낸 순서도이다.
- [0095] 도 9에 도시된 표현요소를 시각화 하는 방법에서 사용자가 선택한 객체에 대하여 수집된 코멘트 데이터에서 복수의 표현요소들을 추출하고(S910), 추출된 표현요소가 사용자가 선택한 객체 외의 다른 객체에서 추출되는 빈도수를 식별한다(S1510). 이후, 추출된 표현요소가 사용자가 선택한 객체 이외의 다른 객체 중 일정 수 이상의 객체에서 일정 빈도수 이상으로 추출되었는지 여부를 판정하고(S1520), 일정 수 이상의 객체에서 일정 빈도수 이상으로 추출된 표현요소에 대하여 가중치를 조정한다(S1530). 이후, 추출된 표현요소들 간의 의미 거리에 기반한 분포에 기초하여 추출된 복수의 표현요소들을 시각화(S920)한다.
- [0096] 이에 따라 표현요소가 모든 객체(콘텐츠)에 대하여 변별력 없이 동등하게 나타나는 경우, 유효하지 않는 것으로 간주할 수 있다.
- [0097] 도 16은 본 발명의 일 실시예에 따른 특정 표현요소가 집중되어 있는 경우 표현요소의 영향력을 조절하는 방법에 대한 순서도를 나타낸 도면이다.
- [0098] 도 9에 도시된 표현요소를 시각화 하는 방법에서 사용자가 선택한 객체에 대하여 수집된 코멘트 데이터에서 복수의 표현요소들을 추출하고(S910), 표현요소가 추출된 빈도수를 측정한다(S1610). 이후, 측정된 빈도수에 따라 측정된 빈도수에 가중치를 부여하여, 측정된 빈도수를 조정한다(S1620). 이때, 조정된 빈도수를 반영하여 표현요소들을 시각화할 수 있다(S920).
- [0099] 이에 따라, 특정 표현요소가 특정 콘텐츠에 과다 집중되어 나타나는 경우에 가중치를 조정하여 특정 표현요소의 영향력을 조절할 수 있다. 즉, 특정 표현요소가 과도하게 집중되어 나타나는 경우, 그로 인하여 다른 표현요소들의 영향력이 지나치게 과소평가될 우려가 있기 때문에 특정 표현요소의 영향력을 조절하는 경우이다.
- [0100] 도 17은 본 발명의 일 실시예에 따른 특정 표현요소가 특정 객체에서 실제로 나타나는 빈도수가 낮은 경우 가중치를 부여하는 방법에 대한 순서도를 나타낸 도면이다.
- [0101] 도 9에 도시된 표현요소를 시각화 하는 방법에서 사용자가 선택한 객체에 대하여 수집된 코멘트 데이터에서 복수의 표현요소들을 추출하고(S910), 코멘트 데이터에서 표현요소가 추출된 빈도수를 측정한다(S1710). 이후, 사용자가 선택한 객체에서 표현요소가 출현하는 빈도수와 측정된 빈도수를 비교하여(S1720), 사용자가 선택한 객체에서 표현요소가 출현하는 빈도수와 측정된 빈도수의 비교 결과에 따라, 표현요소가 추출된 빈도수에 가중치를 부여하여 측정된 빈도수를 조정한다(S1730).
- [0102] 이에 따라, 특정 표현요소가 특정 객체(콘텐츠/영화)에서 실제로 나타나는 빈도수와 코멘트 데이터에서 나타나는 빈도수를 비교하여 코멘트 데이터에서 나타나는 빈도수가 낮은 경우에는 낮은 가중치를 부여할 수 있다.
- [0103] 도 18은 본 발명의 일 실시예에 따른 표현요소를 미리 저장된 표준형의 표현요소로 매핑하고 빈도수를 측정하는 방법에 대한 순서도를 나타낸 도면이다.
- [0104] 도 10에 도시된 표현요소를 시각화 하는 방법에서 사용자가 선택한 객체에 대하여 수집된 코멘트 데이터에서 복수의 표현요소들을 추출(S910)할 때, 표준화된 표현요소가 미리 저장된 데이터베이스 내에 추출된 표현요소가 저장되어 있는지 여부를 탐색하고(S911), 추출된 표현요소가 데이터베이스 내에 저장되어 있지 않으면, 추출된 표현요소와 가장 의미 거리가 가까운 상기 데이터베이스 상의 표준화된 표현요소를 추출된 표현요소의 대표 표현요소로 식별한다(S912).
- [0105] 이후, 추출된 표현요소들의 코멘트 데이터 내에서 추출된 빈도수를 측정할 때(S930), 추출된 표현요소가 코멘트 데이터 내에서 추출된 빈도수를 식별된 대표 표현요소가 코멘트 데이터 내에서 추출된 빈도수에 합산하고, 추출된 표현요소들 간의 의미 거리에 기반한 분포에 기초하여 추출된 복수의 표현요소들을 시각화할 때(S920), 대표

표현요소를 합산된 빈도수를 반영하여 시각화 한다.

- [0106] 이에 따라, 표현요소가 표준형이 아닌 경우, 감정어휘사전 상에 저장된(미리 저장된 데이터베이스 내에) 표준형의 표현요소로 매핑하고, 매핑된 감정어휘사전 상의 표준형의 표현요소 기준으로 각 객체의 코멘트 데이터에서의 빈도수를 측정할 수 있다.
- [0107] 도 19는 본 발명의 일 실시예에 따른 표현요소를 시각화 하는 시스템을 나타낸 도면이다.
- [0108] 표현요소를 시각화 하는 시스템(1900)은, 예를 들어 컴퓨팅 시스템일 수 있으며, 스토리지 장치(1910) 및 프로세서(1920)를 포함한다. 이때, 프로세서(1920)는 표현요소 추출부(1930), 빈도수 측정부(1940), 유효성 판단부(1950), 시각화부(1960)를 포함한다.
- [0109] 스토리지 장치는(1910)는 사용자가 선택한 객체에 대한 코멘트 데이터를 저장하고, 표현요소 추출부(1930)는 스토리지 장치(1910)에 저장된 코멘트 데이터에서 복수의 표현요소들을 추출하여 시각화부(1960)에서 추출된 복수의 표현요소들 간의 의미 거리에 기반한 분포에 기초하여 추출된 복수의 표현요소들을 시각화 한다.
- [0110] 또한, 빈도수 측정부(1940)에서 추출된 표현요소들의 코멘트 데이터 내에서 추출된 빈도수를 측정하는 경우 시각화부(1960)는 추출된 표현요소들을 측정된 표현요소들의 빈도수에 따라 시각화 할 수도 있으며, 이때, 유효성 판단부(1950)가 추출된 표현요소의 유효성을 판단하고, 사용자가 선택한 객체에 대하여 추출된 표현요소가 유효하지 않은 경우, 유효하지 않은 표현요소를 제거할 수도 있다.
- [0111] 또한, 빈도수 측정부(1940)는 표현요소 추출부(1930)에서 추출된 표현요소가 사용자가 선택한 객체 이외의 다른 객체에서 추출되는 빈도수를 식별하여, 추출된 표현요소가 사용자가 선택한 객체 이외의 다른 객체 중 일정 수 이상의 객체에서 일정 빈도수 이상으로 추출되었는지 여부를 판정할 수 있다.
- [0112] 이때, 객체에 대한 코멘트 데이터는 영화에 대한 리뷰, 상품에 대한 상품평, 소셜 리뷰, 게임 리뷰, 여행 리뷰, 서비스에 대한 평가 등 사람들의 감정이 포함되는 모든 코멘트 데이터를 의미한다.
- [0113] 또한, 표현요소는 코멘트 데이터에서 추출되는 사람들의 감정을 나타내는 단어, 문단, 이모티콘 등을 포함하고, 복수의 표현요소들간의 의미 거리에 기반한 분포에 기초하여 도 1 내지 도 4에서 설명된 다차원척도 분석 지도(Multi-Dimensional Scaling map: MDS map)을 기반으로 히트맵(Heat-map) 형태 또는 등고선 등의 모양으로 시각화할 수 있다.
- [0114] 도 20은 본 발명의 일 실시예에 따른 신규 표현요소를 확인하여 표현요소를 시각화 하는 시스템을 나타낸 도면이다.
- [0115] 신규 표현요소를 확인하여 표현요소를 시각화 하는 시스템(1900)은 스토리지 장치(1910) 및 프로세서(1920)를 포함한다. 이때, 프로세서(1920)는 표현요소 추출부(1930), 표현요소 비교부(1970), 신규 표현요소 확인부(1980), 시각화부(1960)를 포함한다.
- [0116] 스토리지 장치는(1910)는 사용자가 선택한 객체에 대한 코멘트 데이터를 저장하고, 표현요소 추출부(1930)는 스토리지 장치(1910)에 저장된 코멘트 데이터에서 복수의 표현요소들을 추출하여 시각화부(1960)에서 추출된 복수의 표현요소들 간의 의미 거리에 기반한 분포에 기초하여 추출된 복수의 표현요소들을 시각화 한다.
- [0117] 이때, 표현요소 비교부(1970)는 표현요소 추출부(1930)에서 추출된 표현요소들과 기존에 추출된 표현요소들을 비교하고, 신규 표현요소 확인부(1980)는 추출된 표현요소들 중 신규한 표현요소가 추가되었는지 여부를 확인한다.
- [0118] 도 21은 본 발명의 일 실시예에 따른 표현요소의 빈도수를 측정 및 조정하여 표현요소를 시각화 하는 시스템을 나타낸 도면이다.
- [0119] 표현요소를 시각화 하는 시스템(1900)은 스토리지 장치(1910) 및 프로세서(1920)를 포함한다. 이때, 프로세서(1920)는 표현요소 추출부(1930), 표현요소 비교부(1970), 신규 표현요소 확인부(1980), 시각화부(1960)를 포함한다.
- [0120] 스토리지 장치는(1910)는 사용자가 선택한 객체에 대한 코멘트 데이터를 저장하고, 표현요소 추출부(1930)는 스토리지 장치(1910)에 저장된 코멘트 데이터에서 복수의 표현요소들을 추출하여 시각화부(1960)에서 추출된 복수의 표현요소들 간의 의미 거리에 기반한 분포에 기초하여 추출된 복수의 표현요소들을 시각화 한다.
- [0121] 빈도수 측정부(1940)는 표현요소 추출부(1930)에서 추출된 표현요소가 사용자가 선택한 객체 이외의 다른 객체



에서 추출되는 빈도수를 식별하여, 추출된 표현요소가 사용자가 선택한 객체 이외의 다른 객체 중 일정 수 이상의 객체에서 일정 빈도수 이상으로 추출되었는지 여부를 판정할 수 있다.

[0122] 이때, 빈도수 조정부(1990)는 측정된 표현요소의 빈도수에 따라 표현요소의 빈도수에 따라 표현요소의 빈도수에 가중치를 부여하여, 표현요소의 빈도수를 조정할 수 있다.

[0123] 이때, 시각화부(1960)는 빈도수 조정부(1990)에서 조정된 빈도수를 반영하여 표현요소들을 시각화할 수 있다.

[0124] 이에 따라, 특정 표현요소가 특정 콘텐츠에 과하게 집중되어 나타나는 경우에도 가중치를 낮추어 그 표현요소의 영향력을 조절할 수 있다.

[0125] 또한, 빈도수 측정부(1940)는 저장된 코멘트 데이터에서 표현요소가 추출된 빈도수를 측정하고, 사용자가 선택한 객체에서 표현요소가 출현하는 빈도수와 식별된 빈도수를 비교할 수 있으며, 이때, 빈도수 조정부(1990)는 사용자가 선택한 객체에서 표현요소가 출현하는 빈도수와 식별된 빈도수의 비교 결과에 따라, 표현요소가 추출된 빈도수에 가중치를 부여하여 표현요소가 추출된 빈도수를 조정할 수 있다.

[0126] 이에 따라, 특정 표현요소가 특정 객체(콘텐츠/영화)에서 실제로 나타나는 빈도수와 코멘트 데이터에서 나타나는 빈도수를 비교하여 코멘트 데이터에서 나타나는 빈도수가 낮은 경우에는 낮은 가중치를 부여할 수 있다.

[0127] 도 22는 본 발명의 일 실시예에 따른 표현요소 추출부를 상세히 나타낸 도면이다.

[0128] 도 19 내지 도 21에 기재된 표현요소 추출부(1930)는 표현요소 탐색부(1931)와 표현요소 식별부(1932)를 포함한다.

[0129] 표현요소 탐색부(1931)는 표준화된 표현요소가 미리 저장된 데이터베이스 내에, 추출된 표현요소가 저장되어 있는지 여부를 탐색하고, 표현요소 식별부(1932)는 추출된 표현요소가 데이터베이스 내에 저장되어 있지 않으면, 추출된 표현요소와 가장 의미 거리가 가까운 데이터베이스 상의 표준화된 표현요소를 추출된 표현요소의 대표 표현요소로 식별한다.

[0130] 이때, 빈도수 측정부(1940)는 추출된 표현요소가 코멘트 데이터 내에서 추출된 빈도수를 식별된 대표 표현요소가 코멘트 데이터 내에서 추출된 빈도수에 합산하고, 시각화부(1960)는 대표 표현요소를 합산된 빈도수에 반영하여 시각화한다.

[0131] 이에 따라, 표현요소가 표준형이 아닌 경우, 감정어휘사전 상에 저장된(미리 저장된 데이터베이스 내에) 표준형의 표현요소로 매핑하고, 그 매핑된 감정어휘사전 상의 표준형의 표현요소 기준으로 각 객체의 코멘트 데이터에서의 빈도수를 측정할 수 있다.

[0132] 도 23 내지 도 27는 본 발명의 일 실시예에 따라 각기 다른 시각화 방법을 나타낸 도면이다.

[0133] 도 23은 본 발명의 일 실시예에 따라 히트맵 형태 외에 시각화 그래프로서, 본 발명을 산점도(Scatter plot) 형태로 나타낸 그래프이며, 이 경우에는 표현어휘의 빈도수에 따라 빈도수가 많을수록 색상이 붉게 표현될 수 있다. 도 24는 Small Multiples 형태로 나타낸 그래프이다.

[0134] 도 25는 본 발명을 등고선(Contour Lines) 형태로 나타낸 도면이며, 이 경우에는 표현어휘의 빈도수에 따라 빈도수가 많을수록 상대적으로 높은 값을 가지게 되어 높이가 높게 표현될 수 있다. 도 26는 코로플레스 맵(Choropleth Maps)으로 나타낸 도면이다. 이 경우에는 본 발명이 반드시 직사각형의 정형화된 형태에서 벗어나 지도의 일부 등 자연지형 또는 자연물의 형상에서도 구현될 수 있다.

[0135] 도 27은 통계지도(Cartograms)로 본 발명을 나타낸 도면이며, 이 경우에는 사용자가 선택한 코멘트 데이터 또는 의견에 따라서 시·도·군에 따라 각각의 표현어휘가 나오는 경우 높은 빈도의 표현어휘를 사용자에게 제공하여 사용자가 선택한 높은 빈도의 표현어휘가 지도상에 표시되도록 표현할 수도 있다.

[0136] 도 28는 본 발명의 일 실시예에 따른 히트맵(Heat-map) 시각화 방법을 3차원으로 응용하여 나타낸 도면이다.

[0137] 본 발명에서 기재된 히트맵(Heat-map) 형태의 시각화는 2차원 평면상에 나타내고 있지만 같은 성질을 유지하면서 3차원과 같은 입체적인 형태로도 변형이 가능하다.

[0138] 표현요소의 빈도수에 따라 차원의 변형, 각도, 픽셀의 크기, 색상을 조절할 수 있으며, 도 28은 본 발명에서 기재된 히트맵(Heat-map) 형태를 3차원 형태로 나타낸 도면이다.

[0139] 도 29는 본 발명의 일 실시예에 따른 표현요소 시각화 방법을 등고선으로 나타낸 도면이다.

- [0140] 도 29는 표현어휘의 빈도수에 따라 2차원 등고선으로 나타낸 도면으로, 표현어휘의 빈도수에 따라 등고선의 색상 및 크기를 조절할 수 있다.
- [0141] 도 30은 도 29에 도시한 등고선 맵을 3차원으로 나타낸 도면이다.
- [0142] 도 30은 도 29에 도시된 2차원 등고선 맵을 표현어휘의 빈도수에 따라 3차원 등고선으로 나타낸 도면으로, 표현어휘의 빈도수에 따라 등고선의 색상, 높낮이 및 크기를 조절할 수 있다.
- [0143] 도 31 내지 도 33은 본 발명의 일 실시예에 따른 의미지도를 기반으로 한 활용 방법을 나타낸 도면이다.
- [0144] 도 31은 본 발명에서 사용되는 다차원척도 분석 지도(Multi-Dimensional Scaling map: MDS map)를 포지셔닝을 활용하는 일 실시예를 나타내며, MDS맵 포지셔닝은 2개의 축을 사용하여 4가지 속성을 나타내는 기존의 포지셔닝에 비하여 MDS맵 상에 나타난 다양한 속성을 기준으로 다차원의 포지셔닝이 가능하다.
- [0145] 자동차 기업을 예로 들면, 2사분면에 위치한 Audi의 경우 BMW와 같은 사분면에 위치하여 있지만 좀더 미래 지향적인(Future-oriented)의 이미지에 가깝게 위치해 있는 것을 알 수 있다. 마찬가지로 4사분면에 위치한 SM의 경우 KIA에 비해 편안한(Relaxed) 이미지에 좀더 가까운 것을 알 수 있다.
- [0146] 이러한 MDS Map을 활용한 포지셔닝은 기업 이미지뿐만 아니라 MDS상에 나타나는 특징에 따라 도 32 및 도33과 같이 상품, 인물 및 캐릭터의 이미지 포지셔닝에도 활용이 가능하다.
- [0147] 상술한 본 발명의 실시예들은 하나의 객체(콘텐츠)에 대한 코멘트(리뷰) 데이터로부터 추출된 표현요소(어휘, 이모티콘, 감정, 평가, 의견을 포함하는 요소)를 중심으로 기술되었다. 그러나 본 발명의 사상은 하나의 객체에 대한 코멘트 데이터의 표현요소를 하나의 감정지도 안에서 직관적으로 표현하는 경우에 한정되는 것은 아니다.
- [0148] 즉, 본 발명의 또 다른 실시예에 따르면 사용자에게 의한 편집 메뉴 또는 둘 이상의 객체에 대한 비교 분석 기능을 제공하는 메뉴가 제공될 수 있다. 이 때 사용자는 제1 객체와 제2 객체를 선택하여 제1 객체에 대한 리뷰 내의 표현요소들과 제2 객체에 대한 리뷰 내의 표현요소들을 비교할 수 있다. 이 때 비교 메뉴로는 제1 객체에 대한 리뷰 내의 표현요소들과 제2 객체에 대한 리뷰 내의 표현요소들 간의 집합 연산(합집합, 교집합, 차집합) 등을 수행하여 양 집합을 비교할 수 있으며, 교집합, 합집합 또는 차집합에 대하여 다시 시각화를 실행할 수 있는 re-draw 메뉴도 제공될 수 있다.
- [0149] 또한 동일한 객체에 대해서도 하나의 시각화 데이터만이 존재하는 것이 아니고, 시계열적인 버전 관리에 따라 둘 이상의 시간 버전(또는 시간 레이어)에 따른 시각화 버전이 존재할 수도 있으며, 시간에 따른 노드의 위치와 속성의 변화를 추적할 수도 있다. 이 때 시간에 따른 각 노드(표현요소)의 속성은 면적, 색상 등으로 나타내어 질 수 있으며, 빈도, 집중도 등을 반영할 수 있다. 이를 테면 히트맵(Heat-Map)이 그 하나의 예가 될 수 있음은 앞에서 설명한 바와 같다.
- [0150] 본 발명의 일 실시 예에 따른 표현요소 시각화 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 본 발명의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0151] 이상과 같이 본 발명에서는 구체적인 구성 요소 등과 같은 특정 사항들과 한정된 실시예 및 도면에 의해 설명되었으나 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐, 본 발명은 상기의 실시예에 한정되는 것은 아니며, 본 발명이 속하는 분야에서 통상적인 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다.
- [0152] 따라서, 본 발명의 사상은 설명된 실시예에 국한되어 정해져서는 아니 되며, 후술하는 특허청구범위뿐 아니라 이 특허청구범위와 균등하거나 등가적 변형이 있는 모든 것들은 본 발명 사상의 범주에 속한다고 할 것이다.

부호의 설명

1910: 스토리지 장치

1920: 프로세서

도면

도면1

68개의 감정어휘					
가련하다	가슴아프다	감격스럽다	감동적이다	거북하다	겁나다
격분하다	경악하다	경이롭다	경쾌하다	공포스럽다	광장하다
구역질나다	그립다	기겁하다	기쁘다	끔찍하다	나른하다
놀라다	늘어지다	달콤하다	담담하다	대단하다	더럽다
두렵다	등골이 서늘하다	무료하다	무섭다	무시무시하다	분노하다
분하다	불결하다	불쌍하다	불쾌하다	불편하다	서글프다
서럽다	섬뜩하다	소름끼치다	속상하다	슬프다	신경질나다
신나다	쉴쉴하다	안타깝다	암울하다	애석하다	애절하다
역겹다	오싹하다	우울하다	유쾌하다	으스스하다	잔인하다
재미있다	즐겁다	증오스럽다	지루하다	징그럽다	차분하다
통쾌하다	평온하다	행복하다	혐오스럽다	환상적이다	활기 있다
황홀하다	흥겹다				

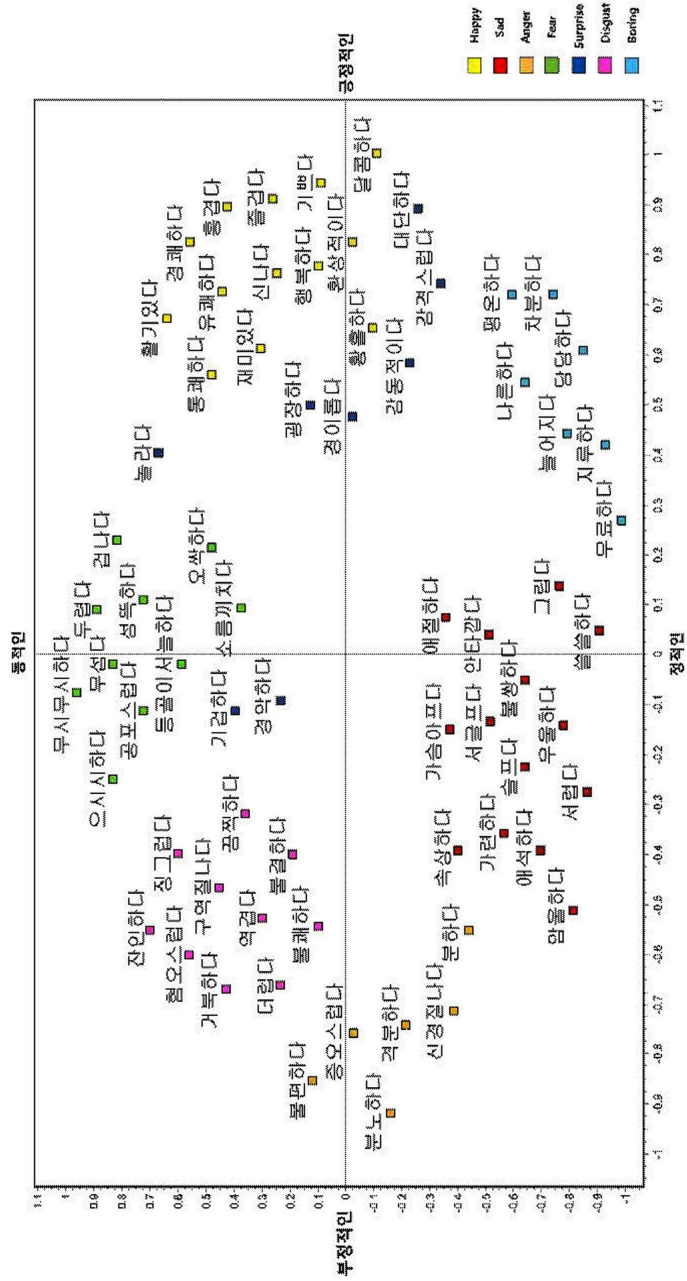
도면2

경쾌하다	가련하다	경악하다	서럽다	담담하다	으시시하다	끔찍하다	섬뜩하다
0.002525	0.005773	0.008493	0.016623	0.016993	0.021914	0.023694	0.025571
기겁하다	늘어지다	두렵다	속상하다	증오스럽다	겁나다	불쌍하다	기쁘다
0.027224	0.027751	0.03335	0.034018	0.034087	0.035402	0.042074	0.051313
징그럽다	애절하다	우울하다	분하다	차분하다	불쾌하다	그립다	거북하다
0.052592	0.053775	0.055982	0.068981	0.079125	0.079163	0.083	0.0838
더럽다	신경질나다	혐오스럽다	암울하다	흥겹다	무료하다	불편하다	역겹다
0.085487	0.08675	0.090305	0.094547	0.09533	0.096808	0.097098	0.097621
구역질나다	감격스럽다	서글프다	경이롭다	광장하다	무시무시하다	쉴쉴하다	애석하다
0.104709	0.104717	0.112113	0.113636	0.121043	0.121759	0.123827	0.125668
분노하다	안타깝다	활기있다	가슴아프다	평온하다	격분하다	불결하다	나른하다
0.130046	0.130744	0.132016	0.140839	0.14449	0.148861	0.152184	0.168187
통쾌하다	오싹하다	신나다	유쾌하다	행복하다	재미있다	대단하다	잔인하다
0.170107	0.175868	0.186403	0.187977	0.194248	0.211447	0.234945	0.247608
지루하다	등골이 서늘하다	무섭다	환상적이다	공포스럽다	즐겁다	감동적이다	황홀하다
0.249294	0.256714	0.268716	0.275393	0.27846	0.326984	0.341209	0.352601
놀라다	소름끼치다	달콤하다	슬프다				
0.394018	0.407559	0.422425	0.433089				

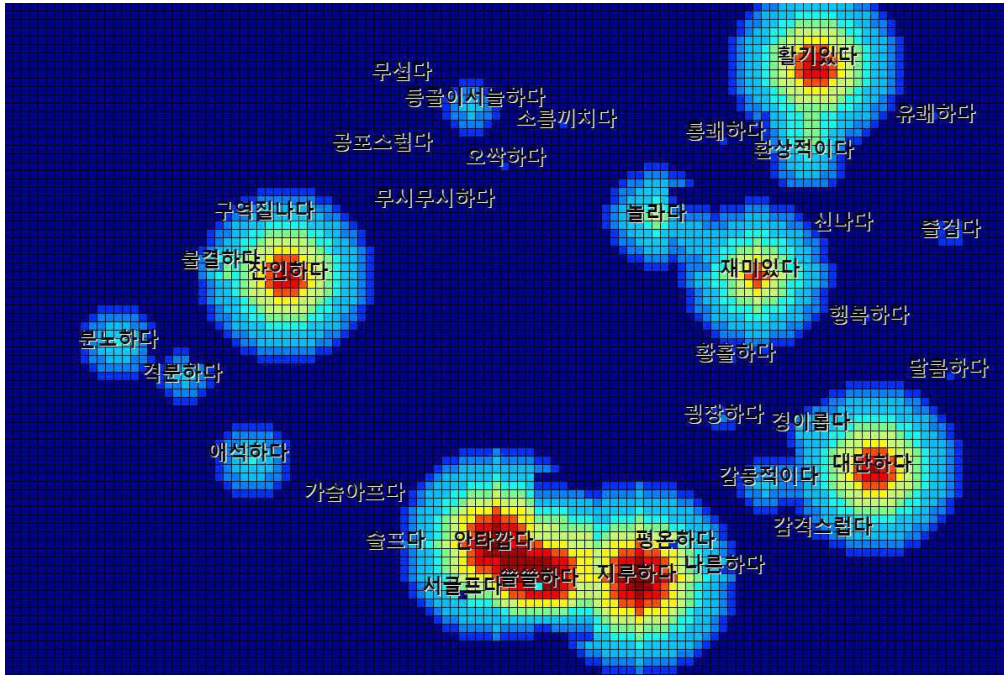
도면3

구역질나다	감격스럽다	서글프다	경이롭다	굉장하다	무시무시하다
0.104709195	0.104716879	0.112112533	0.113635779	0.121043081	0.121758791
쉴쉴하다	애석하다	분노하다	안타깝다	활기있다	가슴아프다
0.12382667	0.12566764	0.130046071	0.130743697	0.132015568	0.140838894
평온하다	격분하다	불결하다	나른하다	통쾌하다	오싹하다
0.144490291	0.148861125	0.152184309	0.168186738	0.170106636	0.175867531
신나다	유쾌하다	행복하다	재미있다	대단하다	잔인하다
0.186403236	0.18797709	0.194247664	0.211447121	0.234945365	0.247608174
지루하다	등골이서늘하다	무섭다	환상적이다	공포스럽다	즐겁다
0.249293951	0.256713888	0.26871569	0.275393151	0.278460331	0.326983704
감동적이다	황홀하다	놀라다	소름끼치다	달콤하다	슬프다
0.341209046	0.352600559	0.39401762	0.407559275	0.422425395	0.433089246

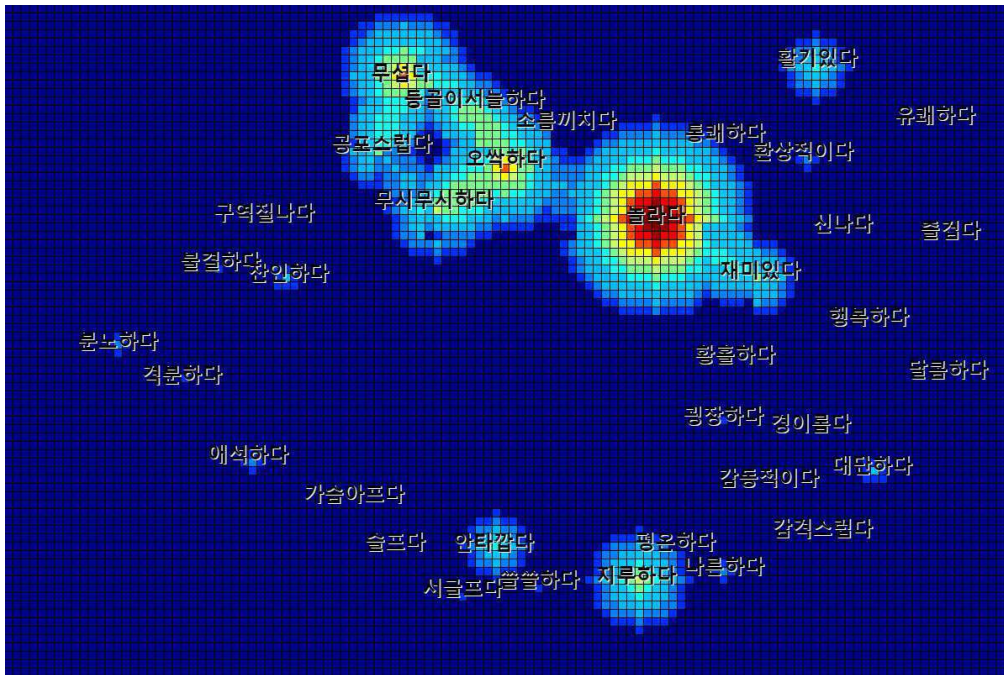
도면4



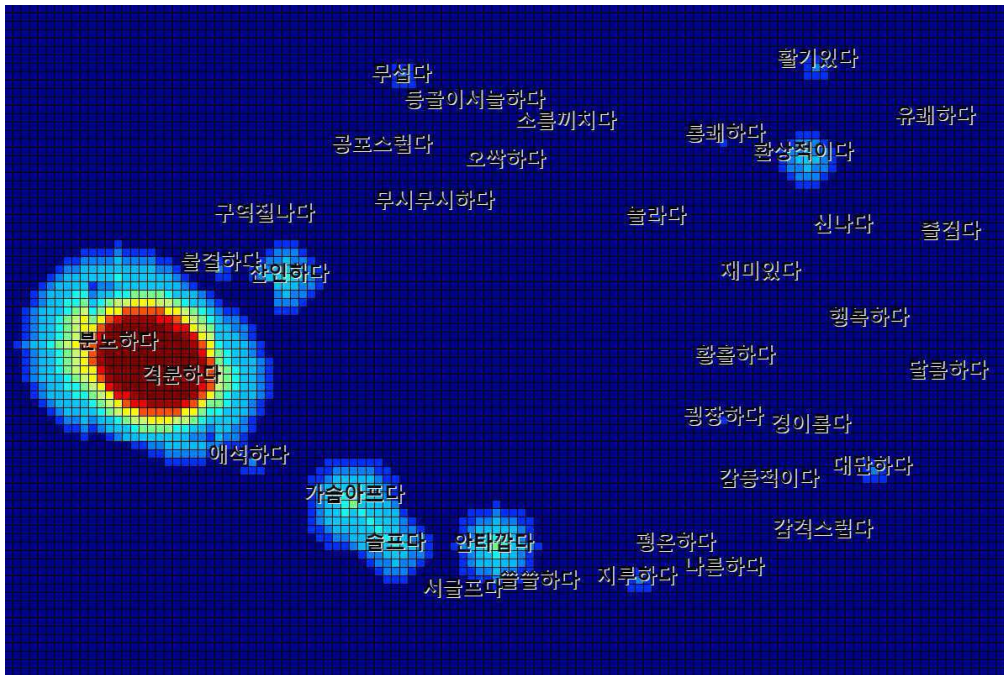
도면5



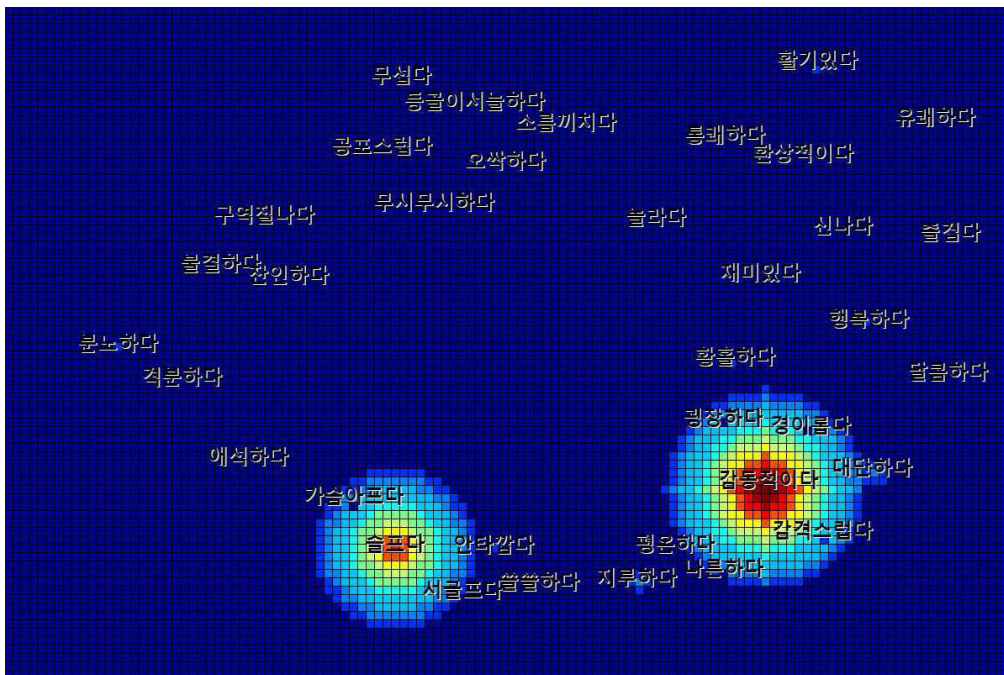
도면6



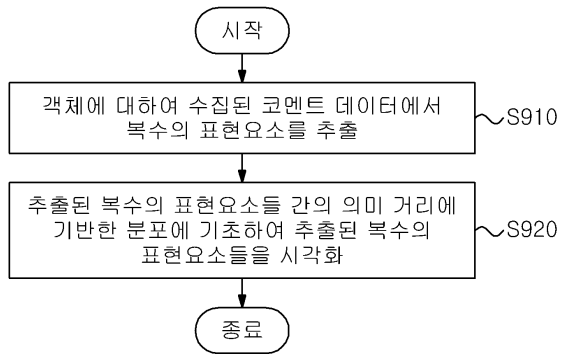
도면7



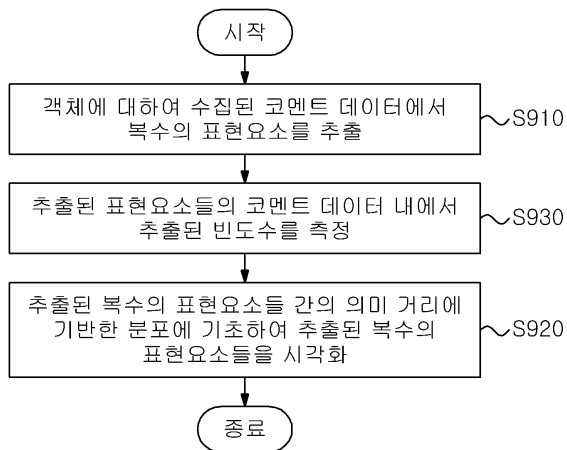
도면8



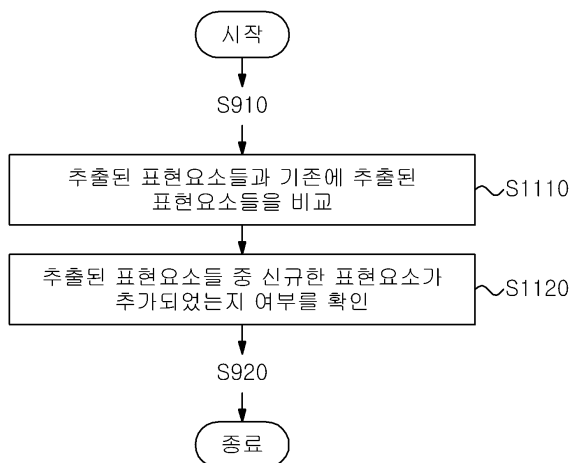
도면9



도면10

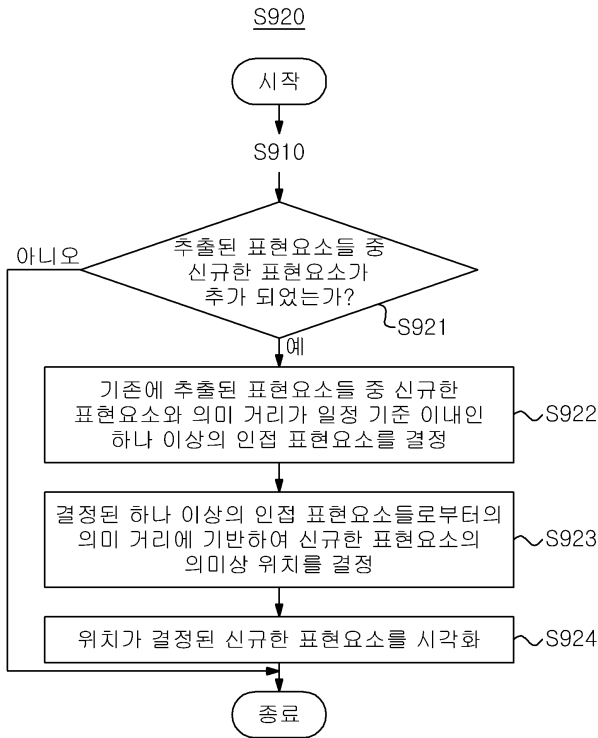


도면11

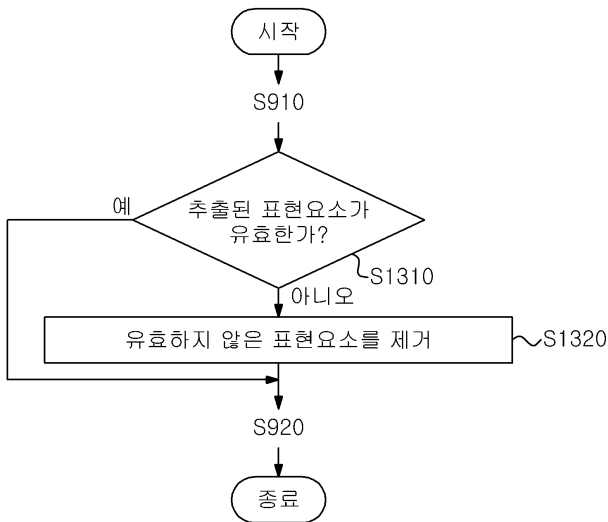




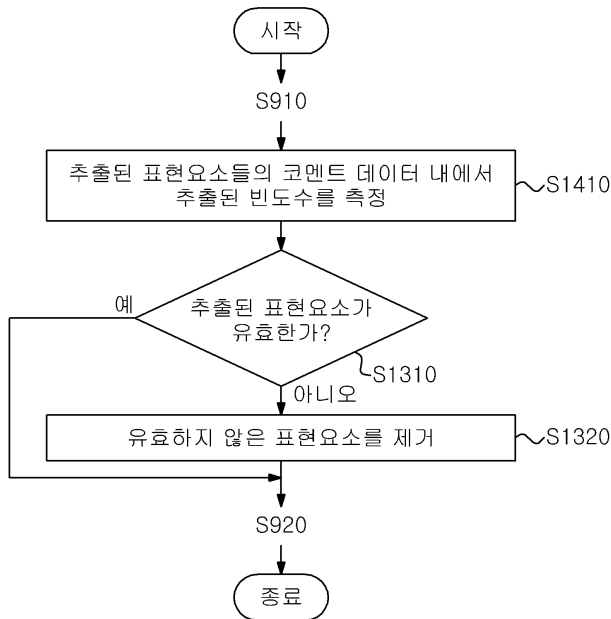
도면12



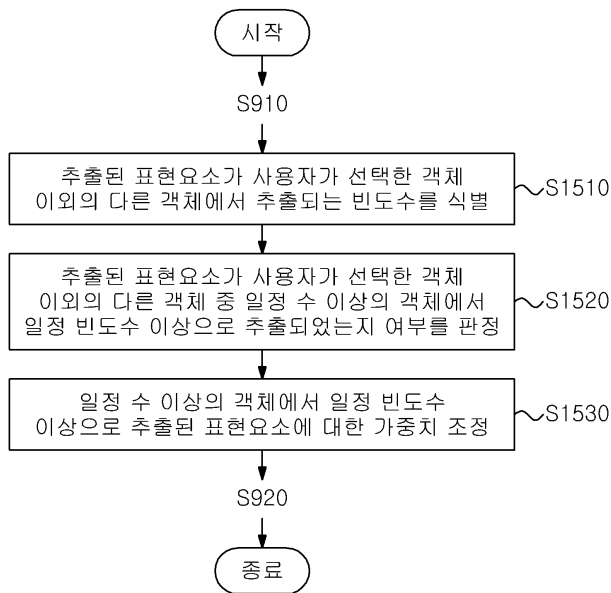
도면13



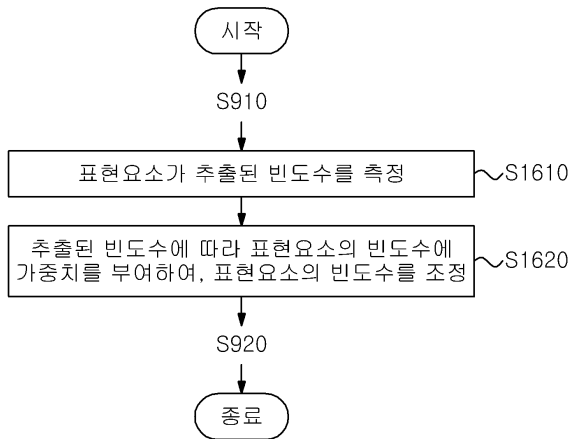
도면14



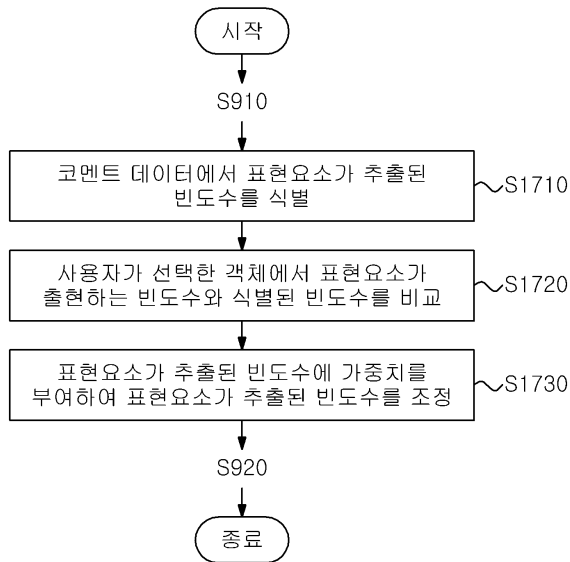
도면15



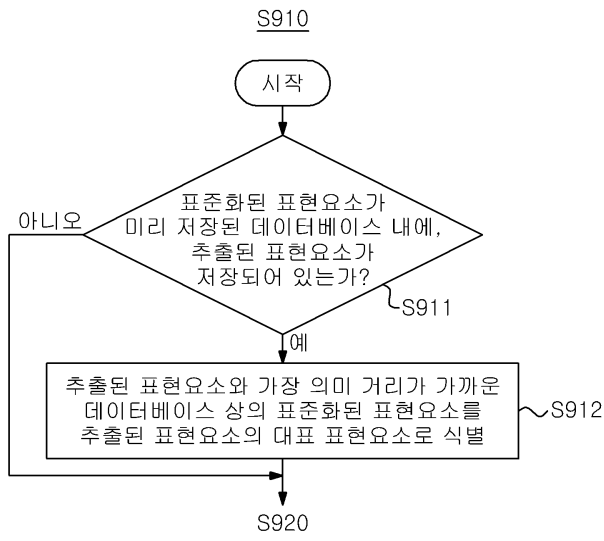
도면16



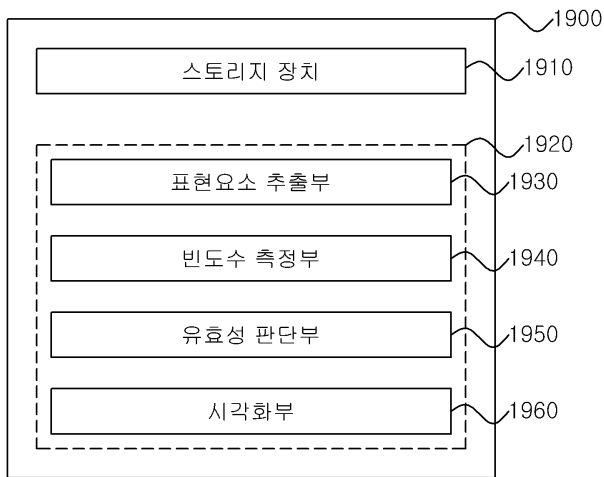
도면17



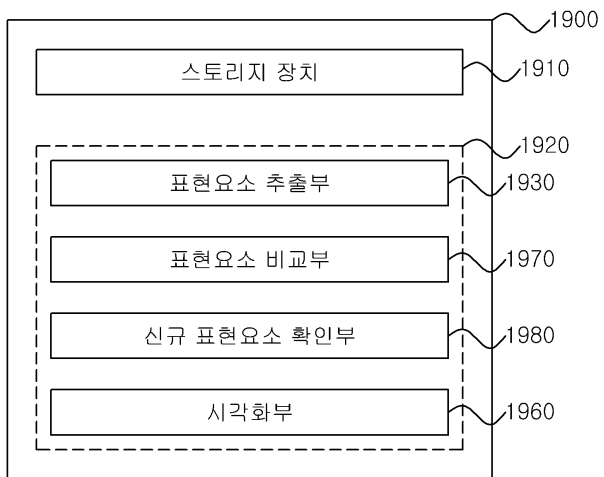
도면18



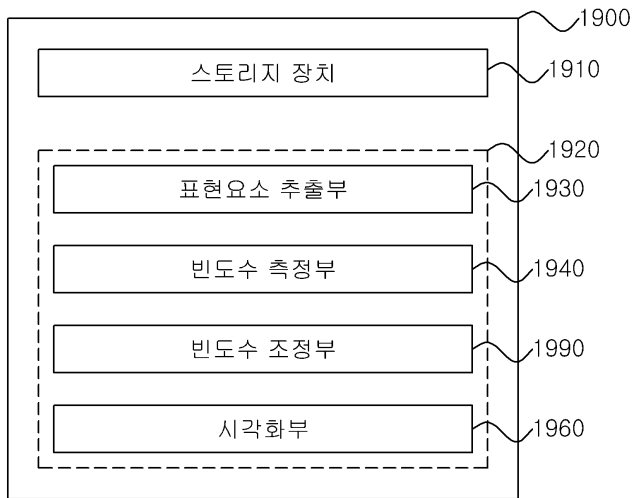
도면19



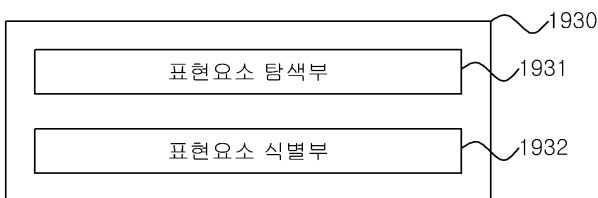
도면20



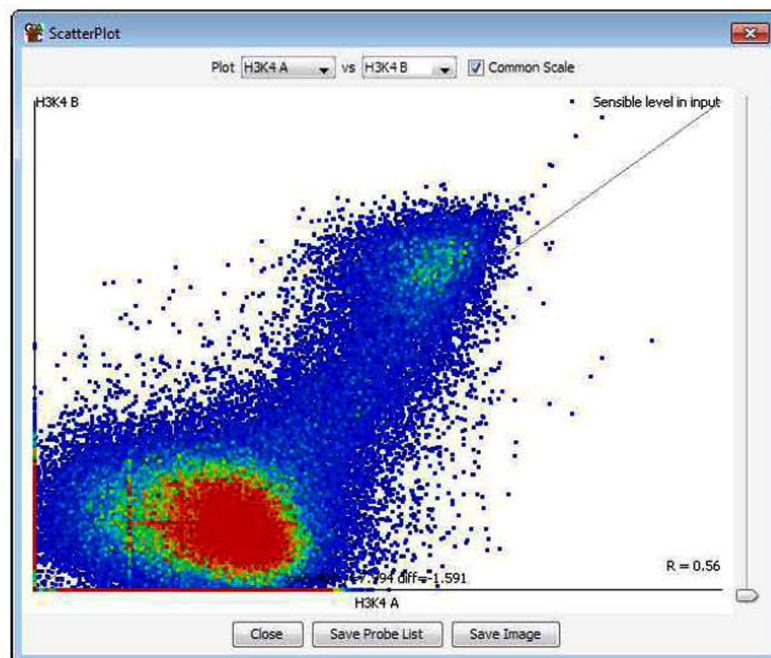
도면21



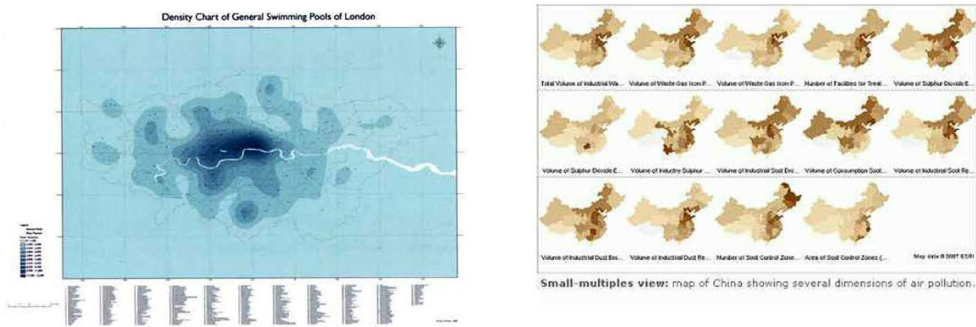
도면22



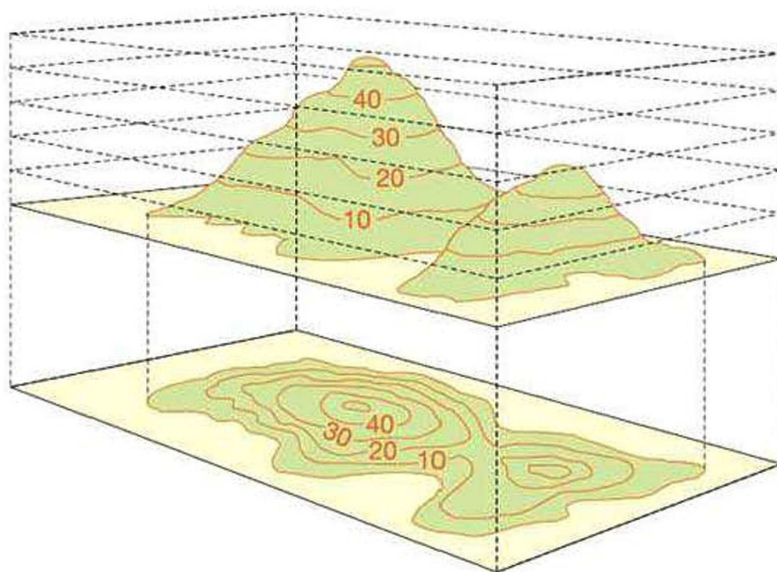
도면23



도면24



도면25



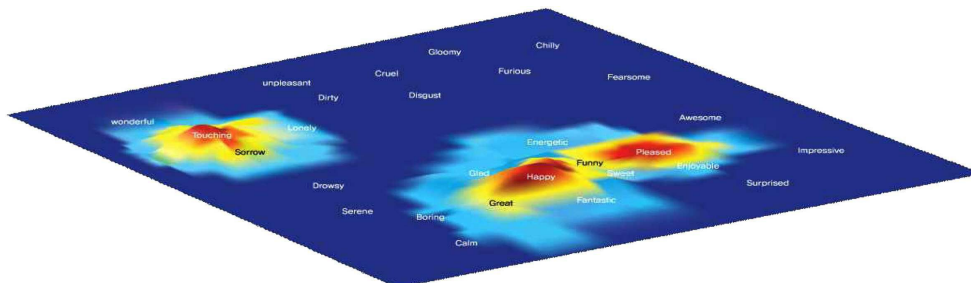
도면26



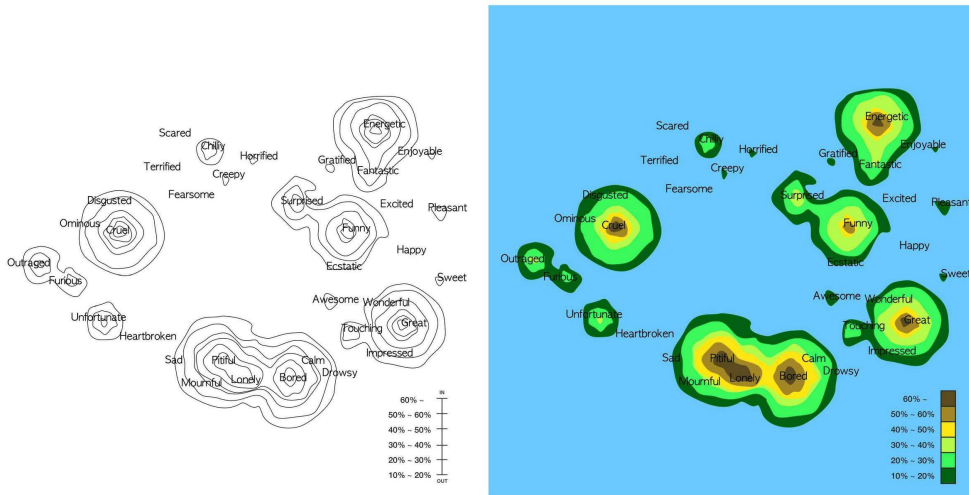
도면27



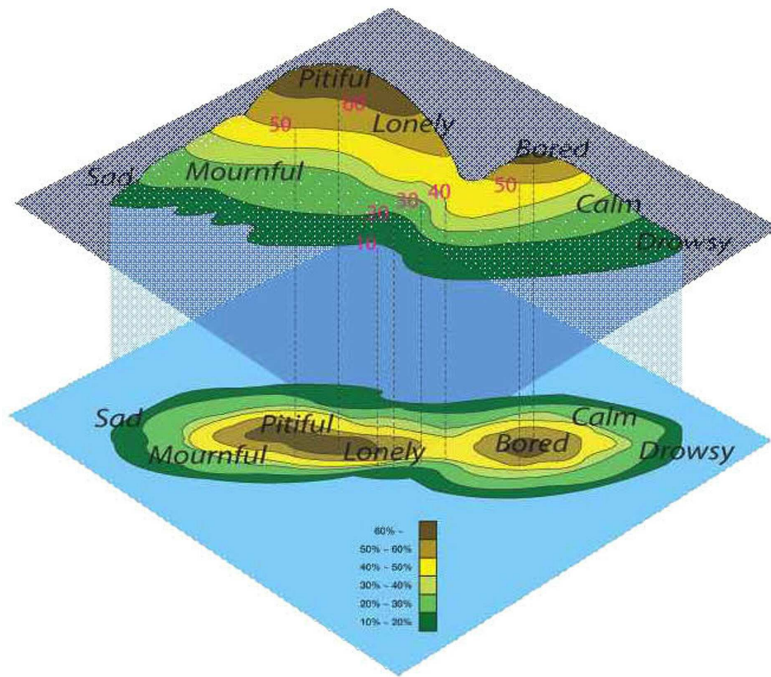
도면28



도면29

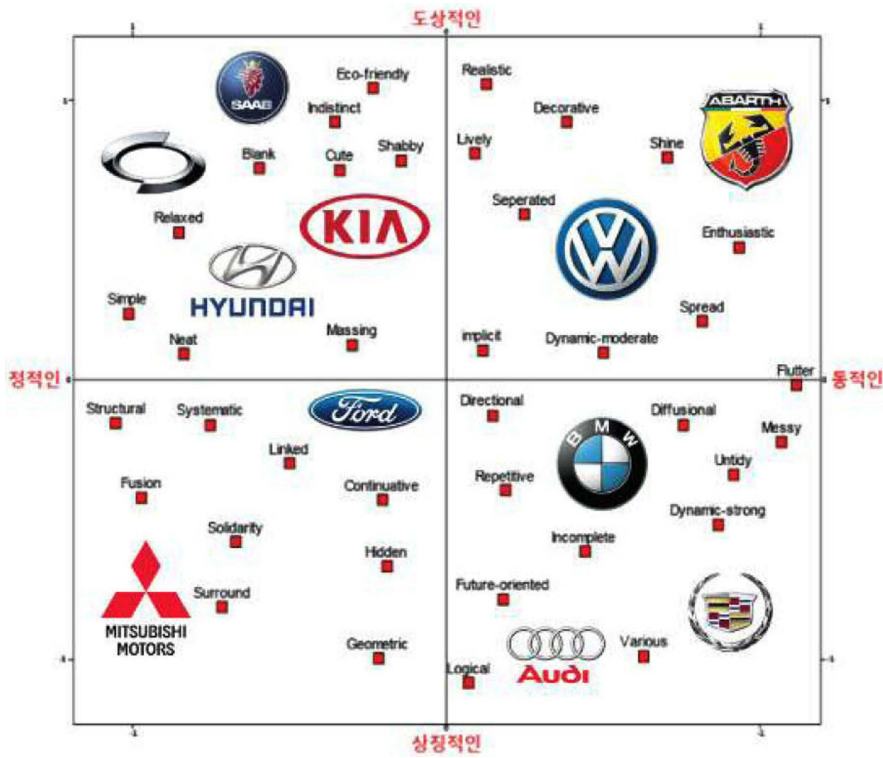


도면30

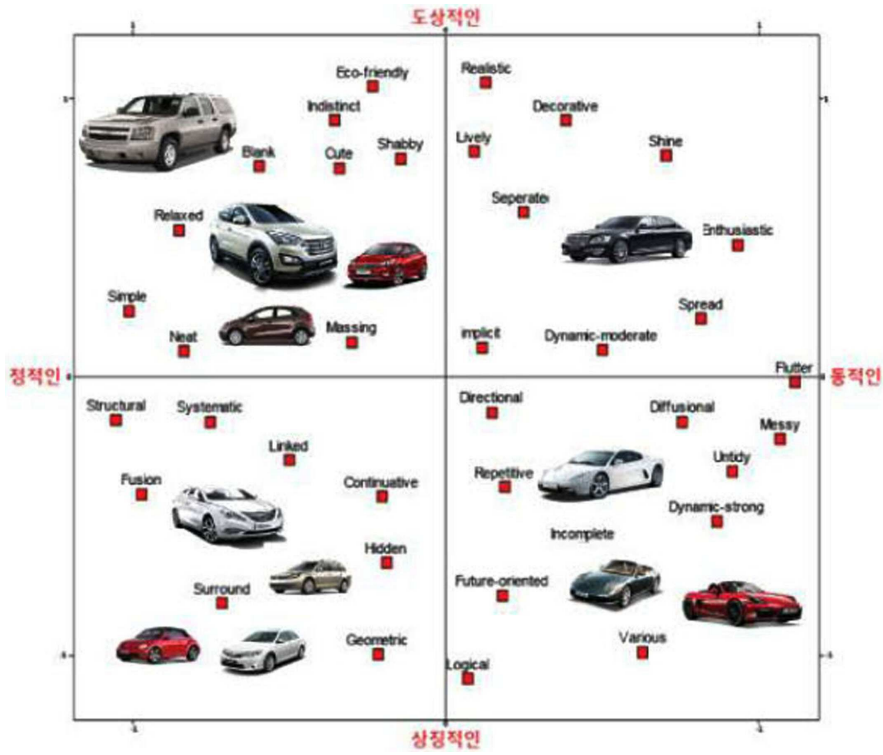




도면31



도면32



도면33

