



**(19) 대한민국특허청(KR)**  
**(12) 등록특허공보(B1)**

(45) 공고일자 2017년02월13일  
 (11) 등록번호 10-1700836  
 (24) 등록일자 2017년01월23일

(51) 국제특허분류(Int. Cl.)  
 G06Q 50/10 (2012.01) G06F 17/30 (2006.01)  
 G06Q 30/02 (2012.01) G06Q 50/30 (2012.01)  
 (52) CPC특허분류  
 G06Q 50/10 (2015.01)  
 G06F 17/30424 (2013.01)  
 (21) 출원번호 10-2015-0183069  
 (22) 출원일자 2015년12월21일  
 심사청구일자 2015년12월21일  
 (56) 선행기술조사문헌  
 KR1020120014277 A\*  
 KR1020120101233 A\*  
 KR1020140101906 A  
 KR101491628 B1  
 \*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
 한국과학기술정보연구원  
 대전광역시 유성구 대학로 245 (어은동)  
 (72) 발명자  
 이태석  
 서울특별시 동대문구 홍릉로25길 10 (청량리동)  
 김완중  
 서울특별시 성동구 행당로 82, 112동 1302호 (행당동, 행당 한진타운)  
 김혜선  
 서울특별시 송파구 양재대로 1218, 103동 1204호 (방이동, 올림픽선수기자촌아파트)  
 (74) 대리인  
 특허법인(유)화우

전체 청구항 수 : 총 7 항

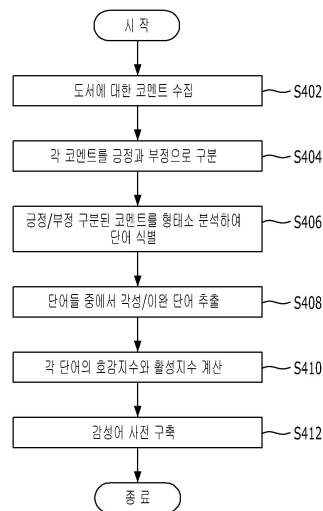
심사관 : 박상현

(54) 발명의 명칭 **도서의 코멘트를 이용한 감성어 사전 구축 방법 및 시스템**

**(57) 요약**

본 발명은 도서의 코멘트를 이용한 감성어 사전 구축 방법 및 시스템에 관한 것으로서, (a) 도서에 대한 코멘트를 수집하는 단계, (b) 상기 수집된 코멘트의 점수를 이용하여 각 코멘트를 긍정과 부정으로 분류하는 단계, (c) 상기 분류된 코멘트들을 형태소 분석하여 단어를 식별하는 단계, (d) 상기 식별된 단어들에서 각성 단어 또는 이완 단어를 추출하는 단계, (e) 상기 식별된 단어들에 대한 호감지수와 상기 각성 단어 또는 이완 단어에 대한 활성지수를 계산하는 단계, (f) 각 단어들에 대한 호감지수와 활성지수를 저장하는 단계를 포함한다.

**대표도** - 도4



(52) CPC특허분류

*G06Q 30/0255* (2013.01)

*G06Q 50/30* (2015.01)

공지예외적용 : 있음

---

**명세서**

**청구범위**

**청구항 1**

감성 사전 구축 장치가 도서의 코멘트를 이용하여 감성어 사전을 구축하는 방법에 있어서,

- (a) 도서에 대한 코멘트를 수집하는 단계;
- (b) 상기 수집된 코멘트의 점수를 이용하여 각 코멘트를 긍정과 부정으로 분류하는 단계;
- (c) 상기 분류된 코멘트들을 형태소 분석하여 단어를 식별하는 단계;
- (d) 구비된 각성 및 이완 사전을 이용하여 상기 식별된 단어들에서 각성 단어 또는 이완 단어를 추출하는 단계;
- (e) 상기 식별된 단어들에 대한 호감지수와 상기 각성 단어 또는 이완 단어에 대한 활성지수를 계산하는 단계;
- (f) 각 단어들에 대한 호감지수와 활성지수를 통계 분석하여 감성어 사전을 구축하는 단계; 및
- (g) 상기 계산된 호감지수와 활성지수의 합으로 상기 도서의 감성 극성을 판단하는 단계;를 포함하고,

상기 활성지수는 하기 수학적식을 이용하여 구하는 것을 특징으로 하는 도서의 코멘트를 이용한 감성어 사전 구축 방법

$$\text{활성지수}(t) = \frac{\text{Count}_{\text{각성}}(t) - \text{Count}_{\text{이완}}(t)}{\max[\text{Count}_{\text{각성}}(t), \text{Count}_{\text{이완}}(t)]}$$

여기서, t는 단어,  $\text{Count}_{\text{각성}}(t)$ 는 단어 t를 포함하는 각성 코멘트의 수,  $\text{Count}_{\text{이완}}(t)$ 는 단어 t를 포함하는 이완 코멘트의 수입.

**청구항 2**

삭제

**청구항 3**

제1항에 있어서,

상기 호감지수는 하기 수학적식을 이용하여 구하는 것을 특징으로 하는 도서의 코멘트를 이용한 감성어 사전 구축 방법

$$\text{호감지수}(t) = \frac{\text{Count}_{\text{긍정}}(t) - \text{Count}_{\text{부정}}(t)}{\max[\text{Count}_{\text{긍정}}(t), \text{Count}_{\text{부정}}(t)]}$$

여기서, t는 단어,  $\text{Count}_{\text{긍정}}(t)$ 는 단어 t를 포함하는 긍정 코멘트의 수,  $\text{Count}_{\text{부정}}(t)$ 는 단어 t를 포함하는 부정 코멘트의 수입.

**청구항 4**

삭제

**청구항 5**

제1항에 있어서,

상기 (f)단계는,

각 단어들에 대한 호감지수를 통계 분석하여 호감도 사전을 구축하고, 상기 각성 단어 또는 이완 단어에 대한 활성지수를 통계 분석하여 활성도 사전을 구축하는 단계; 및

상기 호감도 사전과 상기 활성도 사전을 포함하는 감성어 사전을 구축하는 단계를 포함하는 도서의 코멘트를 이용한 감성어 사전 구축 방법.

**청구항 6**

제1항, 제3항, 제5항 중 어느 한 항에 있어서,

상기 도서의 코멘트를 이용한 감성어 사전 구축 방법을 컴퓨터에서 실행시키기 위한 프로그램이 기록된 컴퓨터에서 판독 가능한 기록 매체.

**청구항 7**

도서에 대한 코멘트를 수집하는 수집부;

상기 수집된 코멘트의 점수를 이용하여 각 코멘트를 긍정과 부정으로 분류하고, 상기 분류된 코멘트들을 형태소 분석하여 단어를 식별하며, 구비된 각성 및 이완 사전을 이용하여 상기 식별된 단어들에서 각성 단어 또는 이완 단어를 추출하는 식별부;

상기 식별된 단어들에 대한 호감지수와 상기 각성 단어 또는 이완 단어에 대한 활성지수를 계산하는 계산부; 및

상기 계산된 호감지수와 활성지수의 합으로 상기 도서의 감성 극성을 판단하는 감성 인지부;를 포함하고,

상기 활성지수는 하기 수학식을 이용하여 구하는 것을 특징으로 하는 감성 사전 구축 장치

[수학식]

$$\text{활성지수}(t) = \frac{\text{Count}_{\text{각성}}(t) - \text{Count}_{\text{이완}}(t)}{\max[\text{Count}_{\text{각성}}(t), \text{Count}_{\text{이완}}(t)]}$$

여기서, t는 단어,  $\text{Count}_{\text{각성}}(t)$ 는 단어 t를 포함하는 각성 코멘트의 수,  $\text{Count}_{\text{이완}}(t)$ 는 단어 t를 포함하는 이완 코멘트의 수임.

**청구항 8**

제7항에 있어서,

상기 호감지수는 하기 수학식을 이용하여 구하는 것을 특징으로 하는 감성 사전 구축 장치

$$\text{호감지수}(t) = \frac{\text{Count}_{\text{긍정}}(t) - \text{Count}_{\text{부정}}(t)}{\max[\text{Count}_{\text{긍정}}(t), \text{Count}_{\text{부정}}(t)]}$$

여기서, t는 단어,  $\text{Count}_{\text{긍정}}(t)$ 는 단어 t를 포함하는 긍정 코멘트의 수,  $\text{Count}_{\text{부정}}(t)$ 는 단어 t를 포함하는 부정 코멘트의 수임.

**청구항 9**

삭제

**청구항 10**

제7항에 있어서,

상기 계산부는, 각 단어들에 대한 호감지수를 통계 분석하여 호감도 사전을 구축하고, 상기 각성 단어 또는 이완 단어에 대한 활성지수를 통계 분석하여 활성도 사전을 구축한 후, 상기 호감도 사전과 상기 활성도 사전을 포함하는 감성어 사전을 구축하는 것을 특징으로 하는 감성 사전 구축 장치.

**발명의 설명**

**기술 분야**

[0001] 본 발명은 도서의 코멘트를 이용한 감성어 사전 구축 방법 및 시스템에 관한 것으로서, 더욱 상세하게는 도서에 대한 감성을 긍정과 부정의 호감도와 각성과 이완의 활성도를 조합한 8개로 분류하여 감성어 사전을 구축하는 도서의 코멘트를 이용한 감성어 사전 구축 방법 및 시스템에 관한 것이다.

**배경 기술**

[0002] 감성은 감각 및 정보자극을 받은 사람이 표출하는 것이다. 감성의 표현은 개인적인 생활경험을 통해 각기 다르게 표현되며, 동일한 감각 및 정보자극을 주더라도 서로 다른 감성이 발생할 수 있다.

[0003] 감성을 다룰 때, 고려할 사항은 객관적인 감성과 개인적 상태 및 변수에 따른 주관적 감성이 함께 섞이는 것이다. 개인적인 요인으로는 연령, 성별, 교육, 건강, 심리상태 등이 있다. 그러므로 감성은 개인마다 조금씩 다르게 표현될 수 있음을 충분히 감안해야 한다. 공통된 요소를 중심으로 감성을 처리하고 이해하기 위해 사회적 요인과 문화적인 요인을 이용한다. 사회적 요인과 문화적 요인을 통해 공통된 개인 감성을 식별할 수 있기 때문이다.

[0004] 감성을 분석 활용하는 분야가 최근에 매우 다양하게 늘어나고 있고, 새로운 가치로 조명되고 있다. 하지만, 아직도 정확도나 활용 가치에 대한 기준은 명확하지 않다.

[0005] 최근, 영화에 대한 감성과 여러 종류의 온라인 콘텐츠에 대한 감성정보를 추출하여 추천에 활용하는 사례가 있지만, 감성정보는 그 영역에 따라 조금씩 다른 감성을 보여주기 때문에 영화나 온라인 콘텐츠의 감성 사전을 그대로 다른 영역에서 활용할 수 없다. 결국 개별 영역에 맞는 감성 사전을 구축해야 한다. 영화의 경우에는 온라인에서 수집할 수 있는 영화평을 이용하여, 감성 사전을 구축할 수 있다.

[0006] 그러나, 종래에는 긍정 문장과 부정 문장을 인식하고, 긍정/부정에 대한 점수가 부여된 감성 사전을 구축함으로써, 감성 사전의 활용에서도 긍정과 부정에 대한 분석만 가능하다는 제한이 있다.

[0007] 또한, 감성 사전을 단순히 긍정과 부정으로만 구축하는 것은 추천 서비스에서 활용도가 떨어지는 문제가 있다.

**선행기술문헌**

**특허문헌**

[0008] (특허문헌 0001) 선행기술1: 한국등록특허 제1,355,997호(2014.01.28 공고)

**발명의 내용**

**해결하려는 과제**

[0009] 본 발명의 목적은 도서에 대한 서평을 이용하여 감성어 사전을 구축할 수 있는 도서의 코멘트를 이용한 감성어 사전 구축 방법 및 시스템을 제공함에 있다.

[0010] 본 발명의 다른 목적은 도서에 대한 감성을 긍정과 부정의 호감도와 각성과 이완의 활성도를 조합한 8개로 분류하여 감성어 사전을 구축할 수 있는 도서의 코멘트를 이용한 감성어 사전 구축 방법 및 시스템을 제공하는 것이다

다.

[0011] 한편, 본 발명이 이루고자 하는 기술적 과제는 이상에서 언급한 기술적 과제로 제한되지 않으며, 이하에서 설명할 내용으로부터 통상의 기술자에게 자명한 범위 내에서 다양한 기술적 과제가 포함될 수 있다.

**과제의 해결 수단**

[0012] 본 발명의 일 측면에 따르면, (a) 도서에 대한 코멘트를 수집하는 단계, (b) 상기 수집된 코멘트의 점수를 이용하여 각 코멘트를 긍정과 부정으로 분류하는 단계, (c) 상기 분류된 코멘트들을 형태소 분석하여 단어를 식별하는 단계, (d) 상기 식별된 단어들에서 각성 단어 또는 이완 단어를 추출하는 단계, (e) 상기 식별된 단어들에 대한 호감지수와 상기 각성 단어 또는 이완 단어에 대한 활성지수를 계산하는 단계, (f) 각 단어들에 대한 호감지수와 활성지수를 통계 분석하여 감성어 사전을 구축하는 단계를 포함하는 도서의 코멘트를 이용한 감성어 사전 구축 방법이 제공된다.

[0013] 상기 (d)단계는, 기 구비된 각성/이완 사전을 이용하여 각성 단어 또는 이완 단어를 추출할 수 있다.

[0014] 상기 호감지수는 하기 수학식을 이용하여 구할 수 있다.

[0015] 
$$\text{호감지수}(t) = \frac{\text{Count}_{\text{긍정}}(t) - \text{Count}_{\text{부정}}(t)}{\max[\text{Count}_{\text{긍정}}(t), \text{Count}_{\text{부정}}(t)]}$$

[0016] 여기서, t는 단어,  $\text{Count}_{\text{긍정}}(t)$ 는 단어 t를 포함하는 긍정 코멘트의 수,  $\text{Count}_{\text{부정}}(t)$ 는 단어 t를 포함하는 부정 코멘트의 수일 수 있다.

[0017] 상기 활성지수는 하기 수학식을 이용하여 구할 수 있다.

[0018] 
$$\text{활성지수}(t) = \frac{\text{Count}_{\text{각성}}(t) - \text{Count}_{\text{이완}}(t)}{\max[\text{Count}_{\text{각성}}(t), \text{Count}_{\text{이완}}(t)]}$$

[0019] 여기서, t는 단어,  $\text{Count}_{\text{각성}}(t)$ 는 단어 t를 포함하는 각성 코멘트의 수,  $\text{Count}_{\text{이완}}(t)$ 는 단어 t를 포함하는 이완 코멘트의 수일 수 있다.

[0020] 상기 (f)단계는, 각 단어들에 대한 호감지수를 통계 분석하여 호감도 사전을 구축하고, 상기 각성 단어 또는 이완 단어에 대한 활성지수를 통계 분석하여 활성도 사전을 구축하는 단계, 상기 호감도 사전과 상기 활성도 사전을 포함하는 감성어 사전을 구축하는 단계를 포함할 수 있다.

[0021] 본 발명의 다른 측면에 따르면, 도서에 대한 코멘트를 수집하는 수집부, 상기 수집된 코멘트의 점수를 이용하여 각 코멘트를 긍정과 부정으로 분류하고, 상기 분류된 코멘트들을 형태소 분석하여 단어를 식별하며, 상기 식별된 단어들에서 각성 단어 또는 이완 단어를 추출하는 식별부, 상기 식별된 단어들에 대한 호감지수와 상기 각성 단어 또는 이완 단어에 대한 활성지수를 계산하는 계산부를 포함하는 감성어 사전 구축 장치가 제공된다.

[0022] 상기 계산부는, 각 단어들에 대한 호감지수를 통계 분석하여 호감도 사전을 구축하고, 상기 각성 단어 또는 이완 단어에 대한 활성지수를 통계 분석하여 활성도 사전을 구축한 후, 상기 호감도 사전과 상기 활성도 사전을 포함하는 감성어 사전을 구축할 수 있다.

**발명의 효과**

[0023] 본 발명에 따르면, 도서에 대한 서평을 수집하고, 수집된 서평을 이용하여 감성 사전을 구축할 수 있다.

[0024] 또한, 도서에 대한 감성을 긍정과 부정의 호감도와 각성과 이완의 활성도를 조합한 8개로 분류하여 감성어 사전을 구축함으로써, 비슷한 성향의 도서와 사람에 대해 짝을 지어 추천하는 서비스에 활용할 수 있다.

[0025] 또한, 기존의 긍정부정 감성어사전 기반에서 각성/이완이라는 활성도 사전을 구축함으로써, 보다 정교한 감성을 쉽게 처리할 수 있다.

[0026] 한편, 본 발명의 효과는 이상에서 언급한 효과들로 제한되지 않으며, 이하에서 설명할 내용으로부터 통상의 기술자에게 자명한 범위 내에서 다양한 효과들이 포함될 수 있다.

**도면의 간단한 설명**

- [0027] 도 1은 본 발명의 실시예에 따른 도서의 코멘트를 이용한 감성어 사전 구축 시스템을 나타낸 도면이다.
- 도 2는 본 발명의 실시예에 따른 감성 사전 구축 장치의 구성을 개략적으로 나타낸 블록도이다.
- 도 3은 본 발명의 실시예에 따른 감성 인지 모델을 설명하기 위한 도면이다.
- 도 4는 본 발명의 실시예에 따른 도서 코멘트를 이용한 감성 사전 구축 방법을 나타낸 흐름도이다.

**발명을 실시하기 위한 구체적인 내용**

- [0028] 이하, 첨부된 도면들을 참조하여 본 발명에 따른 '도서의 코멘트를 이용한 감성어 사전 구축 방법 및 시스템'을 상세하게 설명한다. 설명하는 실시 예들은 본 발명의 기술 사상을 당업자가 용이하게 이해할 수 있도록 제공되는 것으로 이에 의해 본 발명이 한정되지 않는다. 또한, 첨부된 도면에 표현된 사항들은 본 발명의 실시 예들을 쉽게 설명하기 위해 도식화된 도면으로 실제로 구현되는 형태와 상이할 수 있다.
- [0029] 한편, 이하에서 표현되는 각 구성부는 본 발명을 구현하기 위한 예일 뿐이다. 따라서, 본 발명의 다른 구현에서는 본 발명의 사상 및 범위를 벗어나지 않는 범위에서 다른 구성부가 사용될 수 있다.
- [0030] 또한, 각 구성부는 순전히 하드웨어 또는 소프트웨어의 구성만으로 구현될 수도 있지만, 동일 기능을 수행하는 다양한 하드웨어 및 소프트웨어 구성들의 조합으로 구현될 수도 있다. 또한, 하나의 하드웨어 또는 소프트웨어에 의해 둘 이상의 구성부들이 함께 구현될 수도 있다.
- [0031] 또한, 어떤 구성요소들을 '포함'한다는 표현은, '개방형'의 표현으로서 해당 구성요소들이 존재하는 것을 단순히 지칭할 뿐이며, 추가적인 구성요소들을 배제하는 것으로 이해되어서는 안 된다.
- [0032] 도 1은 본 발명의 실시예에 따른 도서의 코멘트를 이용한 감성어 사전 구축 시스템을 나타낸 도면이다.
- [0033] 도 1을 참조하면, 도서의 코멘트를 이용한 감성어 사전 구축 시스템은 도서에 대한 정보가 저장된 도서정보 제공 서버(100), 통신망을 통해 도서정보 제공 서버(100)로부터 도서 정보를 제공받아 감성어 사전을 구축하는 감성사전 구축 장치(200)를 포함한다.
- [0034] 도서정보 제공 서버(100)에는 각 도서에 대한 코멘트들이 등록되어 있다. 여기서, 코멘트는 서평, 트윗(tweet), 댓글 등을 포함할 수 있고, 코멘트에는 해당 도서에 대한 점수가 포함되어 있다. 도서정보 제공 서버(100)는 예컨대, 알라딘 사이트 등을 포함할 수 있다.
- [0035] 감성사전 구축 장치(200)는 도서에 대한 감성을 긍정과 부정의 호감도와 각성과 이완의 활성도를 조합한 8개로 분류한 감성어 사전을 구축한다. 즉, 감성사전 구축 장치(200)는 도서정보 제공 서버(100)로부터 각 도서에 대한 코멘트들을 수집하고, 그 수집된 코멘트에 포함된 단어들에 대한 호감지수와 활성지수를 계산하여 감성어 사전을 구축한다.
- [0036] 이러한 감성 사전 구축 장치(200)에 대한 상세한 설명은 도 2를 참조하기로 한다.
- [0037] 도 2는 본 발명의 실시예에 따른 감성 사전 구축 장치의 구성을 개략적으로 나타낸 블록도, 도 3은 본 발명의 실시예에 따른 감성 인지 모델을 설명하기 위한 도면이다.
- [0038] 도 2를 참조하면, 감성 사전 구축 장치(200)는 데이터베이스(230), 수집부(210), 식별부(220), 계산부(240), 제어부(260)를 포함한다.
- [0039] 데이터베이스(230)는 각성/이완 단어 사전과 감성어 사전이 저장된다.
- [0040] 각성/이완 단어 사전(232)에는 각성 단어와 이완 단어가 저장되어 있다. 각성 단어와 이완 단어는 감정의 강도를 표현하는 단어로, 강한 감정분출과 약한 감정분출의 정도로 표시된다. 이러한 각성 단어와 이완 단어는 예컨대, 매우, 아주, 조금 등과 같이 부사어로 표현되는 경우가 많다.
- [0041] 감성어 사전(234)은 각 단어에 대한 호감지수가 저장된 호감도 사전과 활성지수가 저장된 활성도 사전이 저장된다. 여기서, 단어는 감성어를 말하고, 예컨대, 신나는, 즐거운, 평온한, 화나는, 힘나는, 슬픈, 우울한, 나른한, 평온한 등을 포함할 수 있다.
- [0042] 수집부(210)는 도서에 대한 코멘트를 수집한다. 여기서, 코멘트는 서평, 트윗(tweet), 댓글 등을 포함할 수 있

고, 코멘트에는 해당 도서에 대한 점수가 포함되어 있다. 점수는 2, 4, 6, 8, 10 등의 숫자가 단계로 나누어진 형태일 수 있다.

[0043] 식별부(220)는 수집부(210)에서 수집된 코멘트의 점수를 이용하여 각 코멘트를 긍정과 부정으로 분류하고, 분류된 코멘트들을 형태소 분석하여 단어를 식별하며, 식별된 단어들에서 각성 단어 또는 이완 단어를 추출한다.

[0044] 즉, 식별부(220)는 코멘트에 포함된 점수가 기 설정된 일정 점수 이상이면, 그 코멘트를 긍정으로 분류하고, 일정 점수 이상이 아니면, 부정으로 분류한다. 그런 후, 식별부(220)는 분류된 코멘트들을 형태소 분석하여 단어를 식별하고, 각성/이완 단어 사전을 이용하여 식별된 단어들중에서 각성 단어와 이완 단어를 추출한다.

[0045] 계산부(240)는 식별부(220)에서 식별된 단어들에 대한 호감지수와 각성 단어와 이완 단어에 대한 활성지수를 계산한다.

[0046] 계산부(240)는 수학적 1을 이용하여 단어들에 대한 호감지수를 계산한다. 여기서, 호감지수는 호감도가 높은 코멘트와 호감도가 낮은 코멘트에서 상대적으로 출현한 단어의 빈도로 표현하고, -1 ~ 1의 값을 가진다.

**수학적 1**

[0047] 
$$\text{호감지수}(t) = \frac{\text{Count}_{\text{긍정}}(t) - \text{Count}_{\text{부정}}(t)}{\max[\text{Count}_{\text{긍정}}(t), \text{Count}_{\text{부정}}(t)]}$$

[0048] 여기서, t는 단어,  $\text{Count}_{\text{긍정}}(t)$ 는 단어 t를 포함하는 긍정 코멘트의 수,  $\text{Count}_{\text{부정}}(t)$ 는 단어 t를 포함하는 부정 코멘트의 수를 나타낸다.

[0049] 계산부(240)는 수학적 1을 이용하여 호감지수를 계산하고, 호감도 사전을 구축한다. 이때, 계산부(240)는 통계 분석 툴인 R로 회귀분석을 활용하여 호감도 사전을 구축할 수 있다.

[0050] 계산부(240)는 수학적 2를 이용하여 각성단어 또는 이완단어에 대한 활성지수를 계산한다. 여기서, 활성 지수는 코멘트의 각성 이완의 정도에 따라 상대적으로 출현한 단어의 빈도로 표현하고, -1 ~ 1의 값을 가진다.

**수학적 2**

[0051] 
$$\text{활성지수}(t) = \frac{\text{Count}_{\text{각성}}(t) - \text{Count}_{\text{이완}}(t)}{\max[\text{Count}_{\text{각성}}(t), \text{Count}_{\text{이완}}(t)]}$$

[0052] 여기서, t는 단어,  $\text{Count}_{\text{각성}}(t)$ 는 단어 t를 포함하는 각성 코멘트의 수,  $\text{Count}_{\text{이완}}(t)$ 는 단어 t를 포함하는 이완 코멘트의 수를 나타낸다.

[0053] 활성 지수는 호감지수와 달리 각성/이완 단어 사전(232)에 있는 수식(부사)어 단어만을 대상으로 한다. 즉, 활성지수는 감정의 강도를 표현한다고 할 수 있다. 네티즌의 리뷰중에 평점이 2점과 10점은 감정의 강도가 크다고 할 수 있으며 6점은 강도가 작다고 볼 수 있다. 또한 강도는 “매우, 아주, 조금”과 같이 부사어로 표현하는 경우가 많으므로 부사어만으로 활성지수를 계산하고, 활성도 사전을 구축한다. 이때, 계산부(240)는 통계분석 툴인 R로 회귀분석을 활용하여 활성도 사전을 구축할 수 있다.

[0054] 계산부(240)는 구축된 호감도 사전과 활성도 사전을 감성어 사전(234)에 저장한다. 감성어 사전(234)은 각 단어들에 대한 호감지수와 활성지수가 저장된다.

[0055] 한편, 수집부(210), 식별부(220), 계산부(240) 각각은 컴퓨팅 장치상에서 프로그램을 실행하기 위해 필요한 프로세서 등에 의해 각각 구현될 수 있다. 이처럼 수집부(210), 식별부(220), 계산부(240)는 물리적으로 독립된 각각의 구성에 의해 구현될 수도 있고, 하나의 프로세서 내에서 기능적으로 구분되는 형태로 구현될 수도 있다.

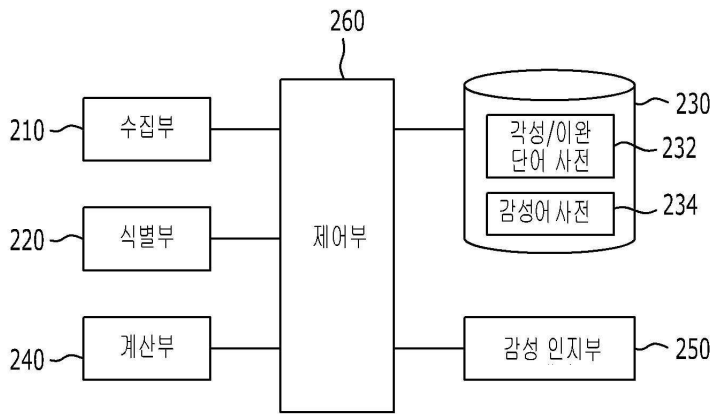
[0056] 제어부(260)는 데이터베이스(230), 수집부(210), 식별부(220), 계산부(240)를 포함하는 감성사전 구축 장치(200)의 다양한 구성부들의 동작을 제어하는 구성이다.



- [0057] 이러한 제어부(260)는 적어도 하나의 연산 장치를 포함할 수 있는데, 여기서 상기 연산 장치는 범용적인 중앙연산장치(CPU), 특정 목적에 적합하게 구현된 프로그래머블 디바이스 소자(CPLD, FPGA), 주문형 반도체 연산장치(ASIC) 또는 마이크로 컨트롤러 칩일 수 있다.
- [0058] 상기와 같이 구성된 감성 사전 구축 장치(200)는 각 도서의 호감지수와 활성지수를 이용하여 각 도서의 감성 극성을 판단하는 감성 인지부(250)를 더 포함할 수 있다. 이때, 감성 인지부(250)는 긍정-부정, 각성-이완의 두 축으로 구분하여 8가지 감성을 표현하는 감성 인지 모델을 이용하여 감성 극성을 판단할 수 있다.
- [0059] 감성 인지 모델에 대해 도 3을 참조하면, 감성 인지 모델은 긍정-부정, 각성-이완의 두 축으로 구분하여 8가지 감성을 표현한다. 따라서, 감성 인지부(250)는 각 축에 계산부(240)에서 계산된 호감지수와 활성지수의 위치를 인지하고, 그 인지된 위치의 값을 이용하여 감성 극성을 판단할 수 있다.
- [0060] 감성 인지 모델을 이용하여 도서의 감성을 해석하고, 독자와 도서를 추천할때 보다 세분화된 감성을 이용할 수 있도록 할 수 있다.
- [0061] 호감 지수는 호감도가 높은 코멘트와 호감도가 낮은 코멘트에서 상대적으로 출현한 단어의 빈도로 표현하며, -1 ~ 1의 값을 가지고, 활성 지수는 코멘트의 각성 이완의 정도에 따라 상대적으로 출현한 단어의 빈도로 표현하며, -1 ~ 1의 값을 가진다. 감성 지수의 중립상태에 대한 임계값은 -0.3 ~ 0.3 으로 설정하여, 감성 인지 모델을 생성한다.
- [0062] 감성 인지부(250)는 도서평에 포함된 단어의 감성지수(호감지수와 활성지수)의 합으로 어떤 글의 감성 극성을 판단할 수 있다. 양수이면, 긍정(신나는, 즐거운, 평온한) 느낌과 각성(화나는, 힘나는, 신나는) 느낌으로 해석한다. 음수이면, 부정(화나는, 슬픈, 우울한) 느낌과 이완(우울한, 나른한, 평온한) 느낌으로 해석한다. 두 지수를 조합하여 8가지 감성 분류에 따른 독자의 성향이나 도서의 감성을 예측하거나 비교하고 추천할 수 있다.
- [0063] 도 4는 본 발명의 실시예에 따른 도서 코멘트를 이용한 감성 사전 구축 방법을 나타낸 흐름도이다.
- [0064] 도 4를 참조하면, 감성 사전 구축 장치는 도서에 대한 코멘트를 수집하고(S402), 수집된 코멘트에 포함된 점수를 이용하여 각 코멘트를 긍정과 부정으로 분류한다(S404).
- [0065] 그런 후, 감성 사전 구축 장치는 분류된 코멘트들을 형태소 분석하여 단어를 식별하고(S406), 식별된 단어들에 서 각성 단어 또는 이완 단어를 추출한다(S408).
- [0066] 그런 후, 감성 사전 구축 장치는 식별된 단어들에 대한 호감지수와 각성 단어 또는 이완 단어에 대한 활성지수를 계산한다(S410).
- [0067] 그런 후, 감성 사전 구축 장치는 각 단어들에 대한 호감지수와 활성지수를 저장하여 감성어 사전을 구축한다(S412). 즉, 감성 사전 구축 장치는 통계분석 툴인 R로 회귀분석을 활용하여 호감도 사전과 활성도 사전을 구축한다. 이때, 구축된 호감도 사전과 활성도 사전이 감성어 사전일 수 있다.
- [0068] 이러한 도서의 코멘트를 이용한 감성어 사전 구축 방법은 프로그램으로 작성 가능하며, 프로그램을 구성하는 코드들 및 코드 세그먼트들은 당해 분야의 프로그래머에 의하여 용이하게 추론될 수 있다. 또한, 도서의 코멘트를 이용한 감성어 사전 구축 방법에 관한 프로그램은 전자장치가 읽을 수 있는 정보저장매체(Readable Media)에 저장되고, 전자장치에 의하여 읽혀지고 실행될 수 있다.
- [0069] 이와 같이, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자는 본 발명이 그 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 실시 형태로 실시될 수 있다는 것을 인지할 수 있을 것이다. 따라서 이상에서 기술한 실시 예들은 예시적인 것일 뿐이며, 그 범위를 제한해놓은 한정적인 것이 아닌 것으로 이해해야만 한다. 또한, 도면에 도시된 순서도들은 본 발명을 실시함에 있어서 가장 바람직한 결과를 달성하기 위해 예시적으로 도시된 순차적인 순서에 불과하며, 다른 추가적인 단계들이 제공되거나, 일부 단계가 삭제될 수 있음은 물론이다.
- [0070] 본 명세서에서 기술한 기술적 특징과 이를 실행하는 구현물은 디지털 전자 회로로 구현되거나, 본 명세서에서 기술하는 구조 및 그 구조적인 등가물 등을 포함하는 컴퓨터 소프트웨어, 펌웨어 또는 하드웨어로 구현되거나, 이들 중 하나 이상의 조합으로 구현 가능하다. 또한 본 명세서에서 기술한 기술적 특징을 실행하는 구현물은 컴퓨터 프로그램 제품, 다시 말해 처리 시스템의 동작을 제어하기 위하여 또는 이것에 의한 실행을 위하여 유형의 프로그램 저장매체 상에 인코딩된 컴퓨터 프로그램 명령어에 관한 모듈로서 구현될 수도 있다.
- [0071] 컴퓨터로 판독 가능한 기록 매체는 기계로 판독 가능한 저장 장치, 기계로 판독 가능한 저장 기판, 메모리 장치



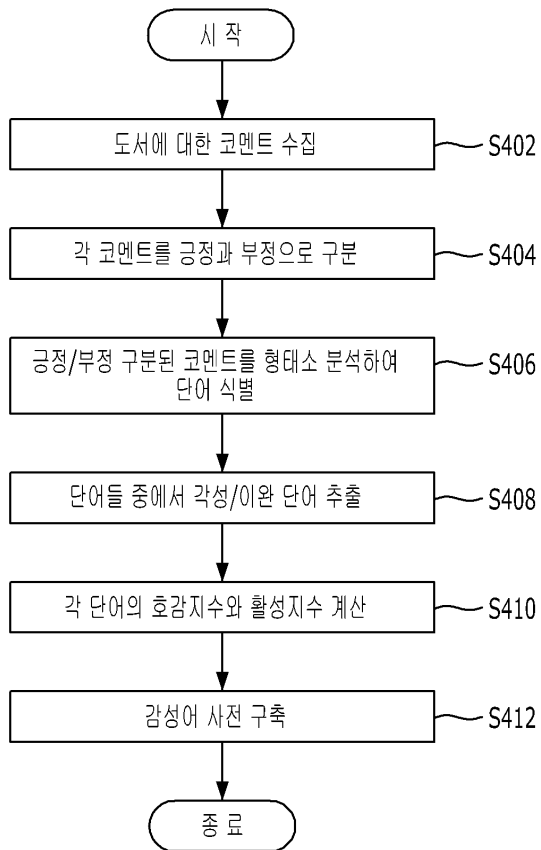
도면2



도면3



도면4



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 제 6 항

【변경전】

제1항 내지 제5항 중 어느 한 항에 있어서

【변경후】

제1항, 제3항, 제5항 중 어느 한 항에 있어서