

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2016년 12월 29일 (29.12.2016)



(10) 국제공개번호
WO 2016/208941 A1

- (51) 국제특허분류:
G06F 17/28 (2006.01) G06F 17/27 (2006.01)
- (21) 국제출원번호: PCT/KR2016/006576
- (22) 국제출원일: 2016년 6월 21일 (21.06.2016)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보:
10-2015-0088695 2015년 6월 22일 (22.06.2015) KR
- (71) 출원인: 전자부품연구원 (KOREA ELECTRONICS TECHNOLOGY INSTITUTE) [KR/KR]; 13509 경기도 성남시 분당구 새나리로, 25, Gyeonggi-do (KR).
- (72) 발명자: 문연국 (MOON, Yeon Guk); 03924 서울시 마포구 월드컵북로 54길, 11, 전자회관 10층 (상암동), Seoul (KR). 이동현 (LEE, Dong Hyun); 08070 서울시 양천구 중앙로 29길 55, 4동 401호 (신월동, 신안약수 아파트), Seoul (KR). 채승훈 (CHAE, Seung Hoon); 07569 서울시 강서구 화곡로 66길, 70 (등촌동), Seoul (KR). 윤희화 (YOON, Hee Hwa); 03401 서울시 은평구 연서로 8길 17-2, 102호 (역촌동), Seoul (KR).

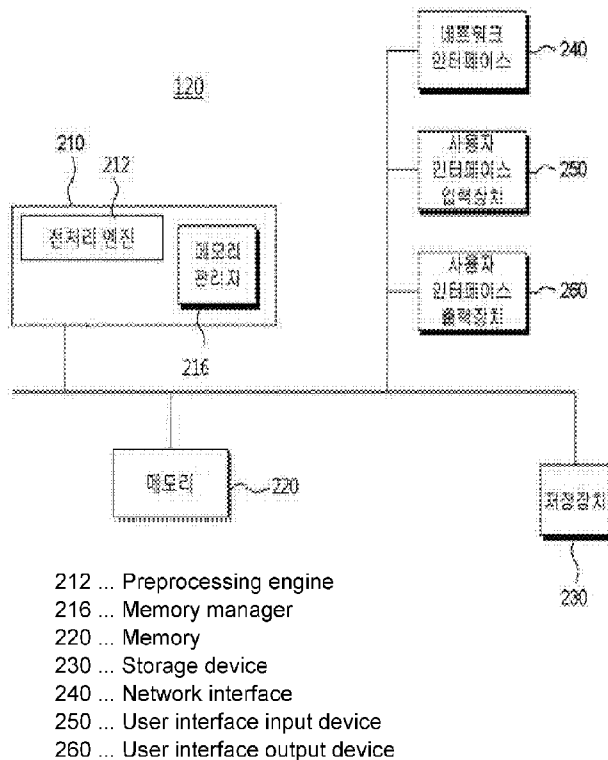
- (74) 대리인: 정부연 (JUNG, Bu Yon); 06646 서울시 서초구 반포대로 26길, 29, 원플러스빌딩, 2층, Seoul (KR).
- (81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

[다음 쪽 계속]

(54) Title: TEXT PREPROCESSING METHOD AND PREPROCESSING SYSTEM FOR PERFORMING SAME

(54) 발명의 명칭: 텍스트 전처리 방법 및 이를 수행하는 전처리 시스템

[도2]



(57) Abstract: The present invention relates to a text preprocessing method capable of increasing the accuracy of machine translations through a preprocessing process for replacing words included in a text to be translated with standard words, and a preprocessing system for performing the same. The text preprocessing system comprises: an alternative word database which stores alternative words; and a processor for operating a preprocessing engine which pre-processes an inputted text and outputs a text of the same language as the inputted text, wherein the preprocessing engine identifies, from the inputted text, a word to be replaced, and outputs a text in which the identified word to be replaced is replaced with an alternative word.

(57) 요약서: 본 발명은 번역 대상 텍스트에 포함된 용어를 표준어로 대체하는 전처리를 통해 기계 번역의 정확도를 높일 수 있는 텍스트 전처리 방법 및 이를 수행하는 전처리 시스템에 관한 것으로, 텍스트 전처리 시스템은 대체 용어를 저장하는 대체 용어 데이터베이스 및 상기 입력된 텍스트를 전처리하여 상기 입력된 텍스트와 동일한 언어의 텍스트를 출력하는 전처리 엔진을 실행하는 프로세서를 포함하되, 상기 전처리 엔진은 상기 입력된 텍스트에서 대체 대상 용어를 식별하고, 상기 식별된 대체 대상 용어를 대체 용어로 대체한 텍스트를 출력한다.

공개:

— 국제조사보고서와 함께 (조약 제 21 조(3))

명세서

발명의 명칭: 텍스트 전처리 방법 및 이를 수행하는 전처리 시스템 기술분야

- [1] 본 발명은 텍스트 번역 기술에 관한 것으로, 보다 상세하게는 번역 대상 텍스트에 포함된 용어를 표준어로 대체하는 전처리를 통해 기계 번역의 정확도를 높일 수 있는 텍스트 전처리 방법 및 이를 수행하는 전처리 시스템에 관한 것이다.

배경기술

- [2] 웹툰(Webtoon)은 웹(Web)과 카툰(Cartoon)의 합성어로 웹 코믹 플랫폼을 통칭한다. 웹툰은 모바일 디바이스 기술과 결합하여 전세계적으로 인기 있는 하나의 미디어 콘텐츠로 부상하였다. 웹툰의 세계적인 인기와 함께 다양한 언어로 웹툰의 번역이 진행되고 있다.
- [3] 현재 웹툰 번역은 사람이 직접 웹툰을 읽고 번역하는 수준에 머물러 있다. 체계적인 번역 과정 없이 개인의 능력에 의존하여 번역하는 경우가 대부분이어서 번역의 정확성 측면이나 문화적 차이를 고려한 표현의 측면에서 한계가 여실히 드러나고 있다.
- [4] 또한, 종래 기계 번역 엔진은 신조어, 한글 파괴어, 구어체, 의성어, 의태어, 사투리 등 표준어가 아닌 언어들의 번역 정확도가 확연히 떨어지고, 번역 방식, 엔진 별 DB내 구축된 용어 등에 의해 정확도가 크게 영향을 받는 문제점 있다.
- [5] 한국등록특허 제10-1099177호는 기계 번역기를 훈련하기 위한 방법 및 시스템에 관한 것으로, 다른 기계 번역기들에 의해 생성된 텍스트 입력들을 가지고 훈련되는 기계 번역기가 개시된다. 사용자 또는 다른 소스(source)에 의해 제1 언어로 된 텍스트 입력이 제공되고, 그 후 이러한 텍스트 입력이 제1 기계 번역기에 의해 번역되어 그 텍스트 입력의 제2 언어로 번역된 버전이 생성된다. 그 텍스트 입력 및 번역된 버전이 파싱되고(parsed) 훈련 아키텍처를 통과하여 전송 매핑(transfer mapping) 및 이중언어 사전(bilingual dictionary)이 개발된다. 이러한 컴포넌트들은 그 후 제2 기계 번역기에 의해 다른 텍스트 입력들을 번역할 때 사용된다.
- [6] 한국등록특허 제10-0961717호는 병렬 코퍼스를 이용한 기계번역 오류 탐지 방법 및 장치에 관한 것으로, 병렬 코퍼스를 이용하여 규칙기반의 기계번역시스템의 오류를 자동으로 탐지하고 추적함에 있어, 기계번역상 발견된 오류를 병렬 코퍼스의 목적언어 문장 즉, 정답 문장을 이용하여 오류 수정이 가능할 뿐만 아니라, 오류 유형을 분류하여 일정한 빈도수 이상의 오류들을 오류 정보로 제공함으로써, 기계번역상 오류를 탐지하는데 소요되는 시간과 노력을 획기적으로 줄일 수 있으며, 시스템 엔지니어가 탐지, 추적된 오류 정보를 통해 기계번역시스템의 성능을 쉽게 개선할 수 있도록 함으로써,

기계번역시스템의 성능 향상 효율을 극대화할 수 있다.

발명의 상세한 설명

기술적 과제

- [7] 본 발명의 일 실시예는 번역 대상 텍스트에 포함된 용어를 표준어로 대체하는 전처리를 수행하는 텍스트 전처리 방법 및 이를 수행하는 전처리 시스템을 제공하고자 한다.
- [8] 본 발명의 일 실시예는 기계 번역에 앞서 번역 대상 텍스트에 대한 전처리를 수행하여 번역 정확도를 향상시킬 수 있는 텍스트 전처리 방법 및 이를 수행하는 전처리 시스템을 제공하고자 한다.
- [9] 본 발명의 일 실시예는 신조어, 한글 파괴어, 구어체, 의성어, 의태어, 사투리 등 표준어가 아닌 언어들을 포함하는 웹툰 텍스트의 번역 정확도를 향상시킬 수 있는 텍스트 전처리 방법 및 이를 수행하는 전처리 시스템을 제공하고자 한다.

과제 해결 수단

- [10] 실시예들 중에서, 텍스트 전처리 시스템은 대체 용어를 저장하는 대체 용어 데이터베이스 및 상기 입력된 텍스트를 전처리하여 상기 입력된 텍스트와 동일한 언어의 텍스트를 출력하는 전처리 엔진을 실행하는 프로세서를 포함하되, 상기 전처리 엔진은 상기 입력된 텍스트에서 대체 대상 용어를 식별하고, 상기 식별된 대체 대상 용어를 대체 용어로 대체한 텍스트를 출력한다.
- [11] 일 실시예에서, 상기 전처리 엔진은 상기 입력된 텍스트를 형태소 단위로 분리하고 분리된 형태소의 품사를 결정하는 형태소 분석부, 상기 분리된 형태소가 대체 대상 용어에 해당하는 지 여부를 식별하는 용어 식별부 및 상기 분리된 형태소가 대체 대상 용어에 해당하는 경우, 용어의 일치 여부를 기초로 상기 대체 용어 데이터베이스에서 상기 대체 대상 용어에 대응되는 대체 용어를 검색하는 대체 용어 검색부를 포함할 수 있다.
- [12] 일 실시예에서, 상기 전처리 엔진은 상기 대체 대상 용어를 상기 검색된 대체 용어로 대체한 텍스트를 생성하는 텍스트 생성부를 더 포함할 수 있다.
- [13] 일 실시예에서, 상기 전처리 엔진은 용어의 일치 여부를 기초로 상기 대체 용어 데이터베이스에 대응되는 대체 용어가 없는 경우, 상기 입력된 텍스트의 구문을 분석하여 상기 대체 용어가 없는 용어에 대한 의미를 추정하는 구문 분석부를 더 포함하고, 상기 대체 용어 검색부는 상기 구문 분석 결과를 기초로 추정된 의미에 대응되는 대체 용어를 검색할 수 있다.
- [14] 일 실시예에서, 상기 구문 분석부는 상기 분리된 형태소를 문법에 따라 분석하여 구문 트리 구조를 생성하여 상기 용어에 대한 의미를 추정할 수 있다.
- [15] 일 실시예에서, 상기 전처리 엔진은 상기 용어와 해당 용어에 대해 추정된 의미에 대응되는 대체 용어를 연계하여 상기 대체 용어 데이터베이스에 저장하는 대체 용어 등록부를 더 포함할 수 있다.

- [16] 일 실시예에서, 상기 텍스트 전처리 시스템은 대체 대상 용어를 저장하는 대체 대상 용어 데이터베이스를 더 포함하고, 상기 전처리 엔진은 상기 대체 대상 용어 데이터베이스에 저장된 대체 대상 용어와의 일치 여부를 기초로 상기 입력된 텍스트에서 대체 대상 용어를 식별할 수 있다.
- [17] 일 실시예에서, 상기 텍스트 전처리 시스템은 기계 번역 용어를 저장하는 번역 용어 데이터베이스를 더 포함하고, 상기 전처리 엔진은 상기 번역 용어 데이터베이스에 용어가 포함되어 있는지 여부를 기초로 상기 입력된 텍스트에서 대체 대상 용어를 식별할 수 있다.
- [18] 일 실시예에서, 상기 프로세서는 입력된 텍스트를 다른 언어의 텍스트로 번역하는 기계 번역 엔진을 실행하고, 상기 기계 번역 엔진은 상기 전처리 엔진에서 출력된 텍스트를 설정된 언어로 기계 번역할 수 있다.
- [19] 실시예들 중에서, 텍스트 전처리 방법은 (a) 상기 입력된 텍스트를 형태소 단위로 분리하고 분리된 형태소의 품사를 결정하는 단계, (b) 상기 분리된 형태소가 대체 대상 용어에 해당하는 지 여부를 식별하는 단계, (c) 상기 분리된 형태소가 대체 대상 용어에 해당하는 경우, 용어의 일치 여부를 기초로 상기 대체 용어 데이터베이스에서 상기 대체 대상 용어에 대응되는 대체 용어를 검색하는 단계 및 (d) 상기 대체 대상 용어를 상기 검색된 대체 용어로 대체한 텍스트를 생성하는 단계를 포함한다.

발명의 효과

- [20] 본 발명의 일 실시예에 따른 텍스트 전처리 방법 및 이를 수행하는 전처리 시스템은 번역 대상 텍스트에 포함된 용어를 표준어로 대체하는 전처리를 수행할 수 있다.
- [21] 본 발명의 일 실시예에 따른 텍스트 전처리 방법 및 이를 수행하는 전처리 시스템은 기계 번역에 앞서 번역 대상 텍스트에 대한 전처리를 수행하여 번역 정확도를 향상시킬 수 있다.
- [22] 본 발명의 일 실시예에 따른 텍스트 전처리 방법 및 이를 수행하는 전처리 시스템은 신조어, 한글 파괴어, 구어체, 의성어, 의태어, 사투리 등 표준어가 아닌 언어들을 포함하는 웹툰 텍스트의 번역 정확도를 향상시킬 수 있다.

도면의 간단한 설명

- [23] 도 1은 본 발명의 일 실시예에 따른 텍스트 전처리 시스템을 설명하는 도면이다.
- [24] 도 2는 도 1에 있는 텍스트 전처리 서버를 나타내는 블록도이다.
- [25] 도 3은 도 2에 있는 전처리 엔진을 나타내는 블록도이다.
- [26] 도 4는 구문 분석 과정을 나타내는 블록도이다.
- [27] 도 5는 도 1에 있는 텍스트 전처리 시스템에서 수행되는 텍스트 번역 방법을 설명하는 흐름도이다.

발명의 실시를 위한 형태

- [28] 본 발명에 관한 설명은 구조적 내지 기능적 설명을 위한 실시예에 불과하므로, 본 발명의 권리범위는 본문에 설명된 실시예에 의하여 제한되는 것으로 해석되어서는 아니 된다. 즉, 실시예는 다양한 변경이 가능하고 여러 가지 형태를 가질 수 있으므로 본 발명의 권리범위는 기술적 사상을 실현할 수 있는 균등물들을 포함하는 것으로 이해되어야 한다. 또한, 본 발명에서 제시된 목적 또는 효과는 특정 실시예가 이를 전부 포함하여야 한다거나 그러한 효과만을 포함하여야 한다는 의미는 아니므로, 본 발명의 권리범위는 이에 의하여 제한되는 것으로 이해되어서는 아니 될 것이다.
- [29] 한편, 본 출원에서 서술되는 용어의 의미는 다음과 같이 이해되어야 할 것이다.
- [30] "제1", "제2" 등의 용어는 하나의 구성요소를 다른 구성요소로부터 구별하기 위한 것으로, 이들 용어들에 의해 권리범위가 한정되어서는 아니 된다. 예를 들어, 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다.
- [31] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결될 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다고 언급된 때에는 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다. 한편, 구성요소들 간의 관계를 설명하는 다른 표현들, 즉 "~사이에"와 "바로 ~사이에" 또는 "~에 이웃하는"과 "~에 직접 이웃하는" 등도 마찬가지로 해석되어야 한다.
- [32] 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한 복수의 표현을 포함하는 것으로 이해되어야 하고, "포함하다" 또는 "가지다" 등의 용어는 실시된 특징, 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것이며, 하나 또는 그 이상의 다른 특징이나 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [33] 각 단계들에 있어 식별부호(예를 들어, a, b, c 등)는 설명의 편의를 위하여 사용되는 것으로 식별부호는 각 단계들의 순서를 설명하는 것이 아니며, 각 단계들은 문맥상 명백하게 특정 순서를 기재하지 않는 이상 명기된 순서와 다르게 일어날 수 있다. 즉, 각 단계들은 명기된 순서와 동일하게 일어날 수도 있고 실질적으로 동시에 수행될 수도 있으며 반대의 순서대로 수행될 수도 있다.
- [34] 본 발명은 컴퓨터가 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현될 수 있고, 컴퓨터가 읽을 수 있는 기록 매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록 장치를 포함한다. 컴퓨터가 읽을 수 있는 기록 매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피 디스크, 광 데이터 저장 장치 등이 있으며, 또한, 캐리어 웨이브(예를 들어 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한, 컴퓨터가 읽을 수 있는 기록 매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산 방식으로

컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다.

[35]

[36] *여기서 사용되는 모든 용어들은 다르게 정의되지 않는 한, 본 발명이 속하는 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한 이상적이거나 과도하게 형식적인 의미를 지니는 것으로 해석될 수 없다.

[37]

[38] 도 1은 본 발명의 일 실시예에 따른 텍스트 전처리 시스템을 설명하는 도면이다.

[39] 도 1을 참조하면, 텍스트 전처리 시스템(100)은 사용자 단말(110), 텍스트 전처리 서버(120), 제1 데이터베이스(130) 및 제2 데이터베이스(140)를 포함하고, 이들은 네트워크를 통해 연결될 수 있다.

[40] 사용자 단말(110)은 텍스트 전처리 서버(120)와 연결된 컴퓨팅 장치에 해당하고, 예를 들어, 데스크톱, 노트북, 태블릿 PC 또는 스마트폰으로 구현될 수 있다. 일 실시예에서, 사용자 단말(110)은 데스크톱에 해당할 수 있고, 텍스트 전처리 서버(120)와 LAN을 통해 연결될 수 있다.

[41] 텍스트 전처리 서버(120)는 입력된 텍스트를 다른 언어의 텍스트로 번역하기 전에 텍스트를 전처리하여 출력할 수 있다. 일 실시예에서, 텍스트 전처리 서버(120)는 입력된 텍스트를 전처리하는 전처리 엔진을 포함하고, 전처리된 텍스트를 기계 번역하는 기계 번역 엔진을 포함하는 텍스트 번역 서버(미도시)와 연계될 수 있다. 일 실시예에서, 전처리 엔진은 입력된 텍스트에서 대체 대상 용어를 식별하고, 식별된 대체 대상 용어를 대체 용어로 대체한 텍스트를 출력할 수 있다. 기계 번역 엔진은 전처리 엔진에서 출력된 텍스트를 설정된 언어로 기계 번역하여 출력한다.

[42] 일 실시예에서, 텍스트 전처리 서버(120)는 클라우드 소싱(crowd sourcing) 번역 엔진을 포함하는 클라우드 소싱 서버(미도시)와 연계될 수 있다. 클라우드 소싱(crowd sourcing) 번역 엔진은 클라우드 소싱을 통해 구축된 데이터베이스를 이용하여 기계 번역된 텍스트를 해당 언어의 표현에 적합하도록 수정할 수 있다. 일 실시예에서, 텍스트 전처리 서버, 번역 서버 및 클라우드 소싱 서버는 하나의 시스템 내에 구비되어 구현될 수도 있고, 각각 다른 시스템으로 구현되어 처리 순서에 따라 연결될 수도 있다.

[43] 일 실시예에서, 제1 데이터베이스(130)는 대체 대상 용어를 저장하는 대체 대상 용어 데이터베이스에 해당할 수 있다. 전처리 엔진은 제1 데이터베이스(130)에 저장된 대체 대상 용어가 텍스트에 포함되어 있는 지 여부를 판단하여 텍스트에 포함된 대체 대상 용어를 식별할 수 있다.

[44] 다른 실시예에서, 제1 데이터베이스(130)는 번역 용어를 저장하는 번역 용어

데이터베이스에 해당할 수 있다. 전처리 엔진은 텍스트에 포함된 용어가 제1 데이터베이스에 포함되어 있는지 여부를 판단하여 대체 대상 용어를 식별할 수 있다. 예를 들어, 텍스트에 포함된 용어가 제1 데이터베이스(130)에 포함되어 있지 않으면, 전처리 엔진은 대체 대상 용어로 식별할 수 있다.

[45] 제2 데이터베이스(140)는 대체 용어를 저장하는 대체 용어 데이터베이스에 해당할 수 있다. 전처리 엔진은 제2 데이터베이스(140)에서 대체 대상 용어에 대응되는 대체 용어를 검색할 수 있다.

[46]

[47] 도 2는 도 1에 있는 텍스트 전처리 서버를 나타내는 블록도이다.

[48] 도 2를 참조하면, 텍스트 전처리 서버(110)는 프로세서(210), 메모리(220), 저장장치(230), 네트워크 인터페이스(240), 사용자 인터페이스 입력 장치(250) 및 사용자 인터페이스 출력 장치(260)를 포함한다.

[49] 프로세서(210)는 전처리 엔진(212) 및 메모리 관리자(216)를 실행한다. 전처리 엔진(212)은 입력된 텍스트에서 대체 대상 용어를 식별하고, 식별된 대체 대상 용어를 대체 용어로 대체한 텍스트를 출력한다.

[50] 메모리 관리자(216)는 전처리 엔진(212)에 의하여 읽혀지거나 또는 기록되는 메모리(220)에 있는 데이터를 관리한다. 메모리(220)는 휘발성 또는 비휘발성 메모리로 구현될 수 있다.

[51] 저장장치(230)는 SSD(Solid State Disk) 또는 HDD(Hard Disk Drive)와 같은 비휘발성 메모리로 구현될 수 있고, 텍스트 전처리 서버(120)에 필요한 데이터를 저장하는데 사용된다.

[52] 네트워크 인터페이스(240)는 네트워크와 연결하기 위한 장치를 포함하고, 예를 들어, LAN(Local Area Network) 통신을 위한 어댑터를 포함할 수 있다.

[53] 사용자 인터페이스 입력 장치(250)는 사용자 입력을 수신하기 위한 장치를 포함하고, 예를 들어, 마우스, 트랙볼, 터치 패드, 그래픽 태블릿, 스캐너, 터치 스크린, 키보드 또는 포인팅 장치와 같은 어댑터를 포함할 수 있다. 사용자 인터페이스 출력 장치(260)는 사용자에게 특정 정보(예를 들어, 번역된 텍스트)를 출력하기 위한 장치를 포함하고, 예를 들어, 모니터 또는 터치스크린과 같은 어댑터를 포함할 수 있다.

[54]

[55] 이하에서는 해당 시스템을 통해 웹툰의 텍스트를 전처리하는 경우를 예로 해당 시스템의 동작을 자세히 설명하기로 한다. 해당 시스템을 통해 신조어, 한글 파괴어, 구어체, 의성어, 의태어, 사투리 등 표준어가 아닌 언어들을 포함하는 웹툰 텍스트의 번역 정확도를 크게 향상시킬 수 있다.

[56]

[57] 도 3은 도 2에 있는 전처리 엔진을 나타내는 블록도이다.

[58] 도 3을 참조하면, 전처리 엔진(212)은 형태소 분석부(310), 용어 식별부(320), 대체 용어 검색부(330), 텍스트 생성부(340) 및 구문 분석부(350)를 포함한다.

- [59] 전처리 엔진(212)은 형태소 분석부(310)는 웹툰에 포함된 텍스트(이하, 웹툰 텍스트)를 입력받는다. 예를 들어, 말풍선, 해설, 의성어, 의태어 등과 같이 웹툰에 포함된 모든 텍스트가 전처리 엔진(212)에 입력될 수 있다. 일 실시예에서, 전처리 엔진(212)은 이미지에 포함된 텍스트를 인식하는 텍스트 인식 엔진으로부터 웹툰 텍스트를 입력받을 수도 있고, 사람이 읽어서 정리한 웹툰 텍스트를 입력받을 수도 있다.
- [60] 형태소 분석부(310)는 입력된 텍스트를 형태소 단위로 분리하고 분리된 형태소의 품사를 결정한다. 일 실시예에서, 형태소 분석부(310)는 분리된 각 형태소를 원형으로 복원하고 복원된 형태소를 기초로 품사를 결정할 수 있다. 형태소는 일정한 뜻을 가진 가장 작은 말의 단위로서, 최소의 유의적 단위이며, 어소라고도 한다. 예를 들어, '하늘이 맑다'는 '하늘', '이', '맑-', '-다'와 같이 4개의 형태소로 분리될 수 있으며, 해당 형태소들은 더 이상 분리되면 의미가 달라지거나 사라진다.
- [61] 예를 들어, '너들 어디 학교야!?'라는 웹툰 텍스트가 입력된 경우, 형태소 분석부(310)는 '너'(NP, 대명사), '들'(XSN, 명사 파생 접미사), '어디'(NP, 대명사), '학교'(NNG, 일반명사), '야'(JKV, 호격조사)로 형태소를 분리하고 품사를 결정할 수 있다.
- [62] 용어 식별부(320)는 형태소 분석부(310)에서 분리된 형태소가 대체 대상 용어에 해당하는 지 여부를 식별한다. 일 실시예에서, 용어 식별부(320)는 웹툰 용어 데이터베이스(대체 대상 용어 데이터베이스)에 저장된 웹툰 용어(대체 대상 용어)와 해당 형태소가 일치하는 지 여부를 기초로 텍스트에 포함된 웹툰 용어(대체 대상 용어)를 식별할 수 있다. 웹툰 용어 데이터베이스(대체 대상 용어 데이터베이스)에는 미리 구축된 웹툰 용어(대체 대상 용어)가 저장될 수 있다.
- [63] 다른 실시예에서, 용어 식별부(320)는 기계 번역 용어 데이터베이스(번역 용어 데이터베이스)에 저장된 기계 번역 용어(번역 용어)에 해당 형태소가 포함되어 있는 지 여부를 기초로 텍스트에 포함된 웹툰 용어(대체 대상 용어)를 식별할 수 있다. 예를 들어, 분리된 형태소가 기계 번역 용어에 포함되어 있지 않은 경우, 용어 식별부(320)는 해당 형태소를 웹툰 용어(대체 대상 용어)로 식별할 수 있다.
- [64] 대체 용어 검색부(330)는 형태소 분석부(310)에서 분리된 형태소가 대체 대상 용어에 해당하는 경우, 용어의 일치 여부를 기초로 대체 용어 데이터베이스(140)에서 대체 용어를 검색한다. 예를 들어, 대체 용어 검색부(330)는 대체 대상 용어와 대체 용어 데이터베이스(140)에 저장된 대체 용어를 비교하여 일치율이 기 설정된 값(threshold) 이상인 대체 용어를 검색할 수 있다. 일 실시예에서, 대체 용어는 표준어에 해당할 수 있다.
- [65] 예를 들어, 대체 용어인 '너희'에 대응되는 웹툰 용어로 '너', '니', '너' 등이 연계되어 대체 용어 데이터베이스(140)에 저장되어 있는 경우, 형태소 분석부(310)에서 분리된 '너'와 대체 용어인 '너희'에 연계된 '너'가 100% 일치하므로, 대체 용어 검색부(330)는 '너희'를 대체 용어로 검색할 수 있다.

- [66] 일 실시예에서, 일치율이 기설정된 값 이상인 대체 용어가 복수개 있는 경우, 대체 용어 검색부(330)는 대체 대상 용어의 품사 또는 대체 대상 용어로 식별된 형태소 주위의 다른 형태소와의 관계를 기초로 복수개의 대체 용어 가운데 하나의 대체 용어를 결정할 수 있다. 예를 들어, '너'와 대응되는 대체 용어가 복수개 있는 경우, 대체 용어 검색부(330)는 '너' 주위의 다른 형태소('들')의 결합관계(예를 들어, 복수의 타인을 지칭하는 결합 구조)를 고려하여 '너희'를 대체 용어로 결정할 수 있다.
- [67] 식별된 대체 대상 용어가 복수 개 있는 경우, 대체 용어 검색부(330)는 각 대체 대상 용어에 대해 대체 용어를 검색할 수 있다.
- [68] 텍스트 생성부(340)는 대체 대상 용어를 대체 용어 검색부(330)에서 검색된 대체 용어로 대체하여, 웹툰 용어를 표준어로 대체한 텍스트를 생성한다. 일 실시예에서, 전처리 엔진(212)는 전처리 되기 전의 웹툰 텍스트와 텍스트 생성부(340)에서 웹툰 용어를 표준어로 대체한 텍스트를 함께 출력할 수 있다.
- [69] 용어의 일치 여부를 기초로 대체 용어 데이터베이스(140)에 대응되는 대체 용어가 없는 경우, 구문 분석부(350)는 입력된 텍스트의 구문을 분석하여 대체 용어가 없는 용어에 대한 의미를 추정한다. 구문 분석부(350)는 형태소 분석부(310)에서 분리된 형태소를 문법에 따라 분석하여 구문 트리 구조를 생성하고, 용어에 대한 의미를 추정한다.
- [70] 도 4는 구문 분석 과정을 나타내는 블록도이다.
- [71] 구문 분석부(350)는 분리된 형태소를 병렬 해석하고, 이를 문법에 따라 분석하여 구문 트리 구조를 생성할 수 있다. 예를 들어, 도 4의 (a)와 같은 '엑쓰세대 뽀타지 모르삼! ㅋㅋㅋㅋ'이라는 웹툰 텍스트가 입력된 경우, 구문 분석부(350)는 도 4의 (b)와 같은 구문 트리 구조를 생성한다. 예를 들어, 도 4의 (a)의 웹툰 텍스트는 한글 문법(예를 들어, 문장의 형식, 주어가 문장의 가장 앞에 오도록 하는 규칙, 동사가 문장의 끝에 오도록 하는 규칙 등)에 따라 제1 목적어 '엑쓰세대', 제2 목적어 '뽀타지', 동사 '모르삼' 및 수식어 'ㅋㅋㅋㅋ'로 분류될 수 있다. 한글 문법 상 주어가 생략되어 있으므로, 구문 분석부(350)는 제3자를 지칭하는 주어 '너'를 추가하여 도 4의 (b)와 같은 구문 트리 구조를 생성할 수 있다. 도 4의 (b)의 구문 트리는 동사를 상위 노드에 배치하고, 주어, 목적어, 수식어 등을 동일한 하위 노드에 배치한 경우의 예이다.
- [72] 구문 분석부(350)는 생성된 구문 트리 구조를 기초로 용어에 대한 의미를 추정한다. 예를 들어, '모르삼'의 경우 문장의 끝에 위치하는 동사로서 의문부호(?) 앞에 위치하므로 동사가 의문형으로 표현된 것으로 추론될 수 있다. 또한, '모르삼'의 경우 '모르다'와 가장 유사하므로, 구문 분석부(350)는 '모르삼'의 경우 '모르다'의 의문형 표현으로 추론할 수 있다.
- [73] 대체 용어 검색부(330)는 구문 분석부(350)의 구문 분석 결과를 기초로 추정된 의미에 대응되는 대체 용어를 대체 용어 데이터베이스(140)에서 검색한다. 예를 들어, 대체 용어 검색부(330)는 '모르다'의 의문형 표현에 대응되는 '몰라',

'모르니' 등을 검색할 수 있다.

- [74] 텍스트 생성부(340)는 대체 대상 용어를 대체 용어 검색부(330)에서 검색된 대체 용어로 대체하여, 웹툰 용어를 표준어로 대체한 텍스트를 생성한다. 예를 들어, 도 4의 (a)의 웹툰 텍스트에 대해 대체 용어 검색부(330)가 '엑쓰'와 일치하는 대체 용어 'X', '뽀타지'와 일치하는 대체 용어 '판타지', 'ㅋㅋㅋㅋ'와 일치하는 대체 용어 '하하하하'를 검색하고, '모르삼'의 추정된 의미에 대응되는 대체 용어 '몰라'를 검색한 경우, 텍스트 생성부(340)는 검색된 대체 용어를 기초로 '너 엑스세대 판타지 몰라?! 하하하하'와 같은 텍스트를 생성할 수 있다.
- [75] 일 실시예에서, 전처리 엔진(212)은 대체 대상 용어와 추정된 의미에 대응되는 대체 용어를 연계하여 대체 용어 데이터베이스(140)에 저장하는 대체 용어 등록부(미도시)를 더 포함할 수 있다.
- [76]
- [77] *일 실시예에서, 용어의 일치 여부를 기초로 대체 용어 데이터베이스(140)에 대응되는 대체 용어가 없거나, 구문 분석 결과를 기초로 대체 용어 데이터베이스(140)에 대응되는 대체 용어가 없는 경우, 전처리 엔진(212)은 해당 용어에 대응되는 대체 용어를 운영자로부터 입력받을 수도 있다.
- [78]
- [79] 도 5는 도 1에 있는 텍스트 전처리 시스템에서 수행되는 텍스트 전처리 방법을 설명하는 흐름도이다.
- [80] 도 5를 참조하면, 텍스트 전처리 서버(120)는 입력된 텍스트를 다른 언어의 텍스트로 번역하기 전에 전처리한다. 텍스트 전처리 서버(120)는 입력된 텍스트와 동일한 언어로 전처리된 텍스트를 출력할 수 있다. 전처리 엔진(212)은 번역할 텍스트를 입력 받는다(단계 S510). 일 실시예에서, 전처리 엔진(212)은 이미지에 포함된 텍스트를 인식하는 텍스트 인식 엔진으로부터 웹툰 텍스트를 입력받을 수도 있고, 사람이 읽어서 정리한 웹툰 텍스트를 입력받을 수도 있다.
- [81] 전처리 엔진(212)은 입력된 텍스트를 형태소 단위로 분리하고 분리된 형태소의 품사를 결정한다(단계 S520). 전처리 엔진(212)은 분리된 형태소가 대체 대상 용어에 해당하는 지 여부를 식별한다(단계 S530).
- [82] 일 실시예에서, 전처리 엔진(212)은 대체 대상 용어 데이터베이스에 저장된 대체 대상 용어와의 일치 여부를 기초로 대체 대상 용어를 식별할 수 있다. 다른 실시예에서, 전처리 엔진(212)은 번역 용어 데이터베이스에 용어가 포함되어 있는지 여부를 기초로 대체 대상 용어를 식별할 수 있다.
- [83] 전처리 엔진(212)은 식별된 대체 대상 용어와 대체 용어 데이터 베이스(140)에 저장된 대체 용어의 일치 여부를 판단하여(단계 S540), 대체 대상 용어에 대응되는 대체 용어를 검색한다.
- [84] 대체 대상 용어와 대체 용어 데이터베이스(140)에 저장된 대체 용어를 비교하여 일치율이 기 설정된 값 이상인 경우, 전처리 엔진(212)은 대체 대상 용어를 대체 용어로 대체한 텍스트를 생성한다(단계 S570).

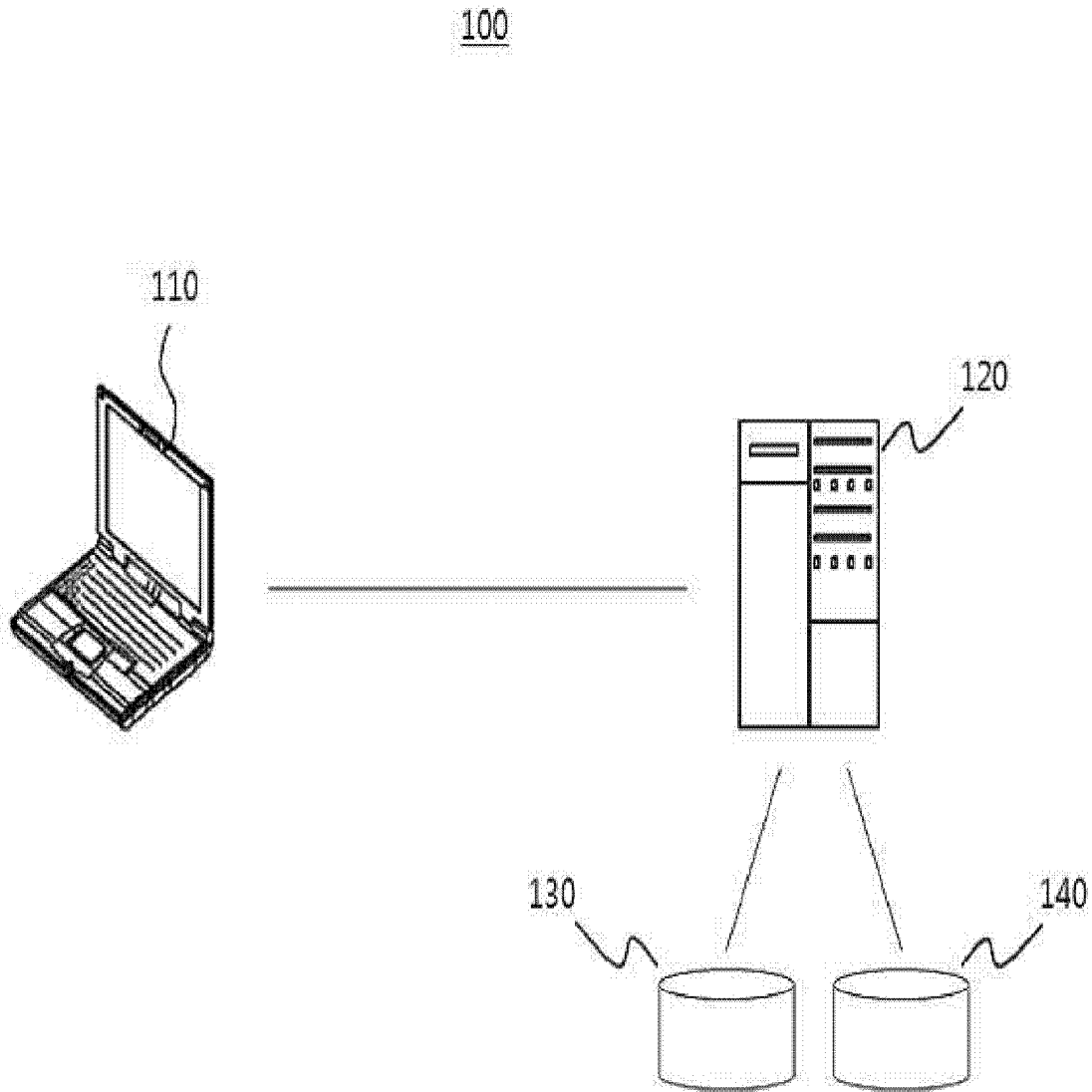
- [85] 일치율이 기 설정된 값 이상인 대체 용어가 없는 경우, 전처리 엔진(212)은 입력된 텍스트의 구문을 분석하여 대체 용어가 없는 용어에 대한 의미를 추정한다(단계 S550). 일 실시예에서, 전처리 엔진(212)은 입력된 텍스트의 구문을 분석하여 구문 트리 구조를 생성할 수 있다.
- [86] 전처리 엔진(212)은 구문 분석 결과를 기초로 추정된 의미에 대응되는 대체 용어를 검색하고(단계 S560), 대체 대상 용어를 대체 용어로 대체한 텍스트를 생성한다(단계 S570).
- [87]
- [88] 상기에서는 본 출원의 바람직한 실시예를 참조하여 설명하였지만, 해당 기술 분야의 숙련된 당업자는 하기의 특허 청구의 범위에 기재된 본 발명의 사상 및 영역으로부터 벗어나지 않는 범위 내에서 본 출원을 다양하게 수정 및 변경시킬 수 있음을 이해할 수 있을 것이다.
- [89]
- [90]

청구범위

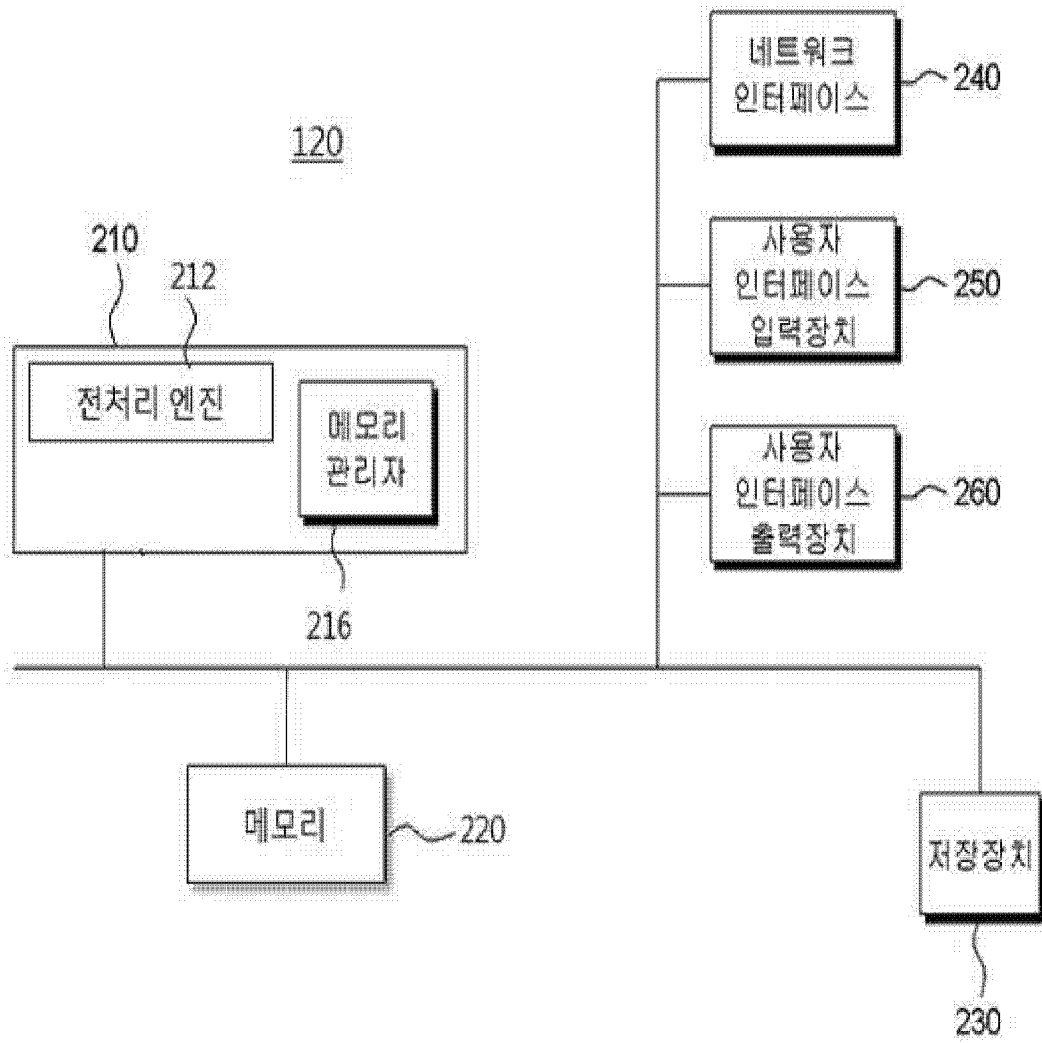
- [청구항 1] 입력된 텍스트를 다른 언어의 텍스트로 번역하기 전에 전처리하는 텍스트 전처리 시스템에 있어서,
대체 용어를 저장하는 대체 용어 데이터베이스; 및
상기 입력된 텍스트를 전처리하여 상기 입력된 텍스트와 동일한 언어의 텍스트를 출력하는 전처리 엔진을 실행하는 프로세서를 포함하되,
상기 전처리 엔진은
상기 입력된 텍스트를 형태소 단위로 분리하고 분리된 형태소의 품사를 결정하는 형태소 분석부;
상기 분리된 형태소가 대체 대상 용어에 해당하는지 여부를 식별하는 용어 식별부;
상기 분리된 형태소가 대체 대상 용어에 해당하는 경우, 용어의 일치 여부를 기초로 상기 대체 용어 데이터베이스에서 상기 대체 대상 용어에 대응되는 대체 용어를 검색하는 대체 용어 검색부; 및
용어의 일치 여부를 기초로 상기 대체 용어 데이터베이스에 대응되는 대체 용어가 없는 경우, 상기 입력된 텍스트의 구문을 분석하여 상기 대체 용어가 없는 용어에 대한 의미를 추정하는 구문 분석부를 포함하고,
상기 대체 용어 검색부는 상기 구문 분석 결과를 기초로 추정된 의미에 대응되는 대체 용어를 검색하는 텍스트 전처리 시스템.
- [청구항 2] 제1항에 있어서, 상기 전처리 엔진은
상기 대체 대상 용어를 상기 검색된 대체 용어로 대체한 텍스트를 생성하는 텍스트 생성부를 더 포함하는 것을 특징으로 하는 텍스트 전처리 시스템.
- [청구항 3] 제1항에 있어서, 상기 구문 분석부는
상기 분리된 형태소를 문법에 따라 분석하여 구문 트리 구조를 생성하여 상기 대체 용어가 없는 용어에 대한 의미를 추정하는 것을 특징으로 하는 텍스트 전처리 시스템.
- [청구항 4] 제1항에 있어서, 상기 전처리 엔진은
상기 용어와 해당 용어에 대해 추정된 의미에 대응되는 대체 용어를 연계하여 상기 대체 용어 데이터베이스에 저장하는 대체 용어 등록부를 더 포함하는 것을 특징으로 하는 텍스트 전처리 시스템.
- [청구항 5] 제1항에 있어서,
상기 대체 대상 용어를 저장하는 대체 대상 용어 데이터베이스를 더 포함하고,
상기 전처리 엔진은 상기 대체 대상 용어 데이터베이스에 저장된 대체 대상 용어와의 일치 여부를 기초로 상기 입력된 텍스트에서 대체 대상 용어를 식별하는 것을 특징으로 하는 텍스트 전처리 시스템.

- [청구항 6] 제1항에 있어서,
 기계 번역 용어를 저장하는 번역 용어 데이터베이스를 더 포함하고,
 상기 전처리 엔진은 상기 번역 용어 데이터베이스에 용어가 포함되어
 있는지 여부를 기초로 상기 입력된 텍스트에서 대체 대상 용어를
 식별하는 것을 특징으로 하는 텍스트 전처리 시스템.
- [청구항 7] 입력된 텍스트를 다른 언어의 텍스트로 번역하기 전에 전처리하고, 대체
 용어를 저장하는 대체 용어 데이터베이스와 상기 입력된 텍스트를
 전처리하여 상기 입력된 텍스트와 동일한 언어의 텍스트를 출력하는
 전처리 엔진을 실행하는 프로세서를 포함하며, 상기 프로세서에 의해서
 수행되는 텍스트 전처리 방법에 있어서,
 (a) 상기 입력된 텍스트를 형태소 단위로 분리하고 분리된 형태소의
 품사를 결정하는 단계;
 (b) 상기 분리된 형태소가 대체 대상 용어에 해당하는 지 여부를 식별하는
 단계;
 (c) 상기 분리된 형태소가 대체 대상 용어에 해당하는 경우, 용어의 일치
 여부를 기초로 상기 대체 용어 데이터베이스에서 상기 대체 대상 용어에
 대응되는 대체 용어를 검색하는 단계; 및
 (d) 용어의 일치 여부를 기초로 상기 대체 용어 데이터베이스에 대응되는
 대체 용어가 없는 경우, 상기 입력된 텍스트의 구문을 분석하여 상기 대체
 용어가 없는 용어에 대한 의미를 추정하는 단계를 포함하고,
 상기 (c) 단계는 상기 구문 분석 결과를 기초로 추정된 의미에 대응되는
 대체 용어를 검색하는 단계를 포함하는 텍스트 전처리 방법.

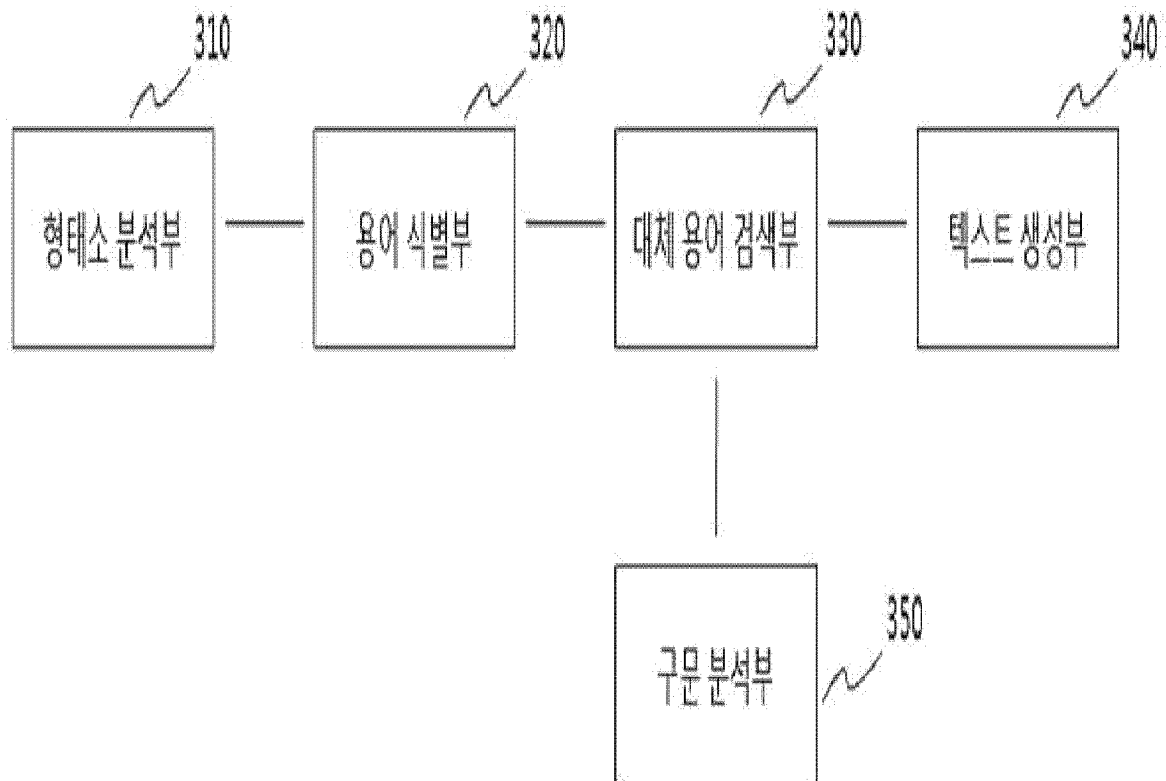
[도 1]



[도2]



[도3]

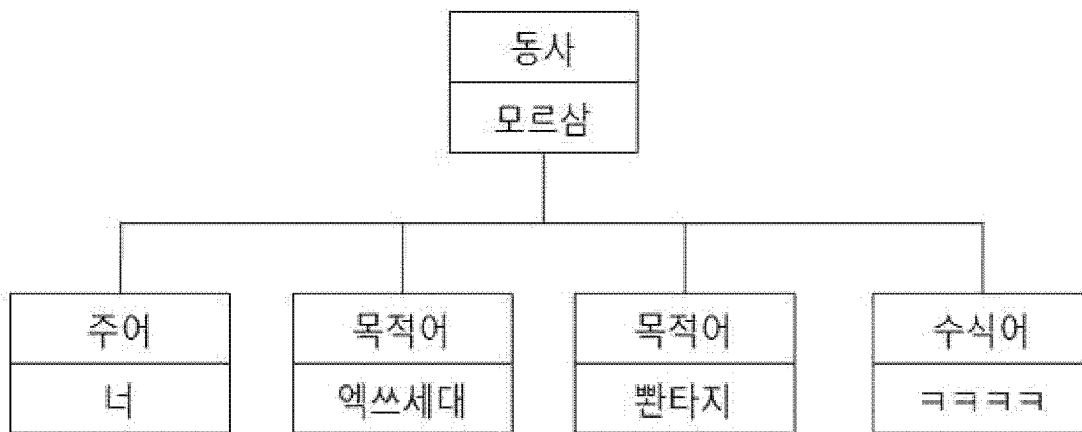
212

[도4]

• 엑쓰세대 뽀타지 모르삼?! ㅋㅋㅋㅋ

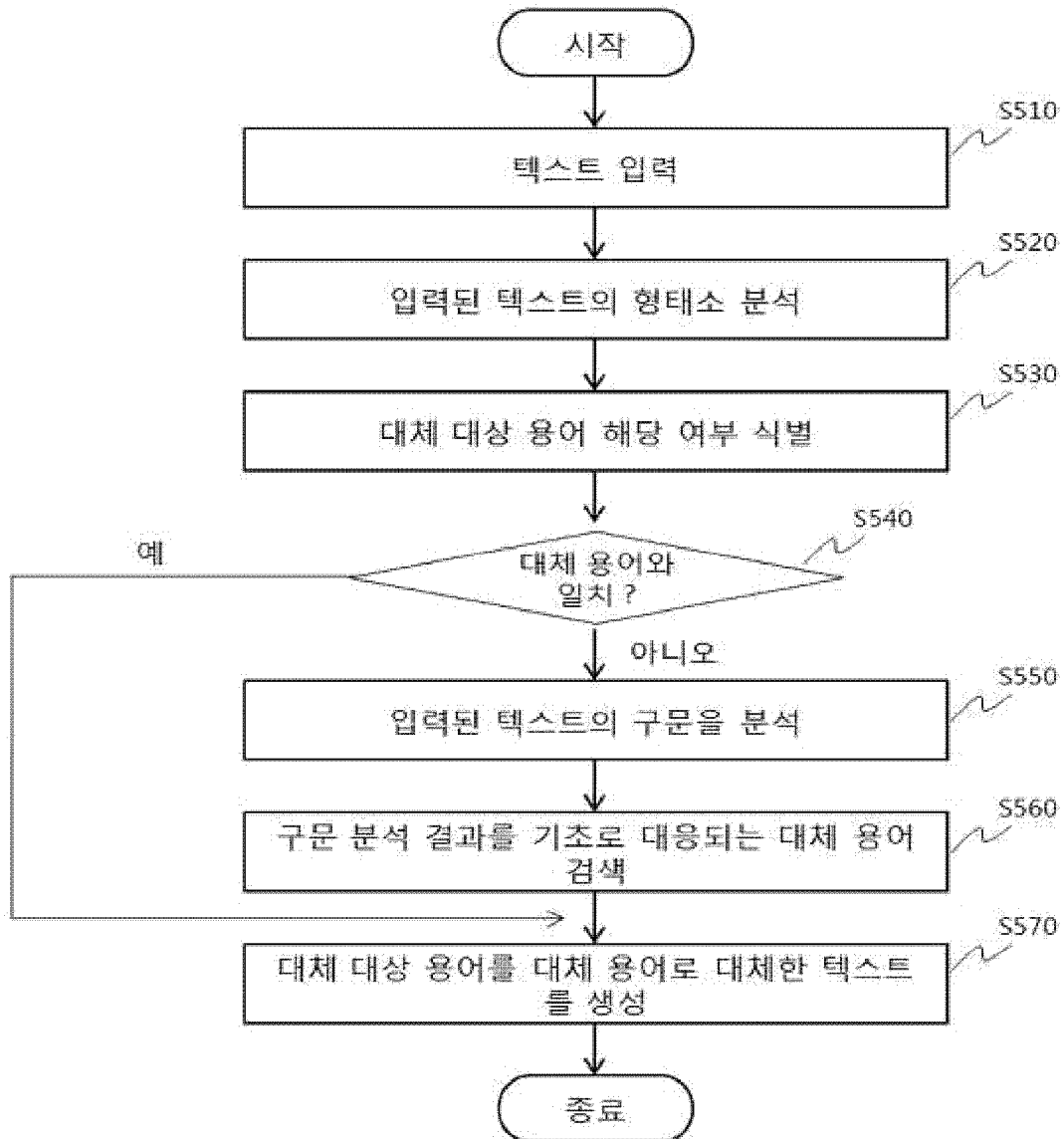
-> '엑쓰' + '세대' + '뽀타지' + '모르삼' + 'ㅋㅋㅋㅋ'

(a)



(b)

[도5]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2016/006576**A. CLASSIFICATION OF SUBJECT MATTER****G06F 17/28(2006.01)i, G06F 17/27(2006.01)i**

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F 17/28; G06F 17/27

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean Utility models and applications for Utility models: IPC as above

Japanese Utility models and applications for Utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS (KIPO internal) & Keywords: text, language, translation, alternative word, pretreatment, morpheme, part of speech, syntax

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KR 10-2012-0035077 A (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 13 April 2012 See paragraphs [0010], [0020]-[0023], [0031] and [0038]; and figure 3.	1-7
A	KR 10-0911372 B1 (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 10 August 2009 See paragraphs [0027]-[0032] and figures 1-2.	1-7
A	KR 10-0322743 B1 (SAMSUNGELECTRONICS CO., LTD.) 07 February 2002 See claims 1-4 and figures 1-4.	1-7
A	US 2010-0088085 A1 (JEON, Jae - Hun et al.) 08 April 2010 See paragraphs [0038]-[0043] and [0056]-[0057]; and figures 1-2.	1-7
A	KR 10-0837358 B1 (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 12 June 2008 See paragraphs [0041]-[0060] and figure 1.	1-7



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

20 SEPTEMBER 2016 (20.09.2016)

Date of mailing of the international search report

21 SEPTEMBER 2016 (21.09.2016)

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex-Daejeon, 189 Seonsa-ro, Daejeon 302-701,
Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2016/006576

Patent document cited in search report	Publication date	Patent family member	Publication date
KR 10-2012-0035077 A	13/04/2012	NONE	
KR 10-0911372 B1	10/08/2009	KR 10-2008-0052282 A	11/06/2008
KR 10-0322743 B1	07/02/2002	KR 10-2001-0028946 A	06/04/2001
US 2010-0088085 A1	08/04/2010	KR 10-2010-0037813 A	12/04/2010
KR 10-0837358 B1	12/06/2008	CN 101131691 A	27/02/2008
		CN 101131691 B	24/10/2012
		US 2008-0052061 A1	28/02/2008
		US 8219382 B2	10/07/2012

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G06F 17/28(2006.01)I, G06F 17/27(2006.01)I

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G06F 17/28; G06F 17/27

조사된 기술분야에 속하는 최소문헌 이외의 문헌

한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우))

eKOMPASS(특허청 내부 검색시스템) & 키워드: 텍스트, 언어, 번역, 대체 용어, 전처리, 형태소, 품사, 구문

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
A	KR 10-2012-0035077 A (한국전자통신연구원) 2012.04.13 단락 [0010], [0020]-[0023], [0031]과 [0038]; 및 도면 3 참조.	1-7
A	KR 10-0911372 B1 (한국전자통신연구원) 2009.08.10 단락 [0027]-[0032] 및 도면 1-2 참조.	1-7
A	KR 10-0322743 B1 (삼성전자 주식회사) 2002.02.07 청구항 1-4 및 도면 1-4 참조.	1-7
A	US 2010-0088085 A1 (JAE-HUN JEON 등) 2010.04.08 단락 [0038]-[0043]과 [0056]-[0057]; 및 도면 1-2 참조.	1-7
A	KR 10-0837358 B1 (한국전자통신연구원) 2008.06.12 단락 [0041]-[0060] 및 도면 1 참조.	1-7

☐ 추가 문헌이 C(계속)에 기재되어 있습니다.

☒ 대응특허에 관한 별지를 참조하십시오.

* 인용된 문헌의 특별 카테고리:

“A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌

“T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌

“E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌

“X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다.

“L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌

“Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.

“O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌

“&” 동일한 대응특허문헌에 속하는 문헌

“P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌

국제조사의 실제 완료일

2016년 09월 20일 (20.09.2016)

국제조사보고서 발송일

2016년 09월 21일 (21.09.2016)

ISA/KR의 명칭 및 우편주소



대한민국 특허청
(35208) 대전광역시 서구 청사로 189,
4동 (둔산동, 정부대전청사)

팩스 번호 +82-42-481-8578

심사관

노지명

전화번호 +82-42-481-8528



국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2012-0035077 A	2012/04/13	없음	
KR 10-0911372 B1	2009/08/10	KR 10-2008-0052282 A	2008/06/11
KR 10-0322743 B1	2002/02/07	KR 10-2001-0028946 A	2001/04/06
US 2010-0088085 A1	2010/04/08	KR 10-2010-0037813 A	2010/04/12
KR 10-0837358 B1	2008/06/12	CN 101131691 A	2008/02/27
		CN 101131691 B	2012/10/24
		US 2008-0052061 A1	2008/02/28
		US 8219382 B2	2012/07/10