

## Bibliographic Data

Int.Cl.	G06Q 50/00   G06F 40/109   G10L 15/26   G06Q 50/30   G06N 3/02   G06F 40/58
Published Date	20220530
Registration No.	1024028840000
Registration Date	20220524
Application No.	1020210095076
Application Date	20210720
Requested Date of Examination	20210720
Agent.	TEHERAN PATENT FIRM
Inventor	PARK,HyounJun   PARK,HyoJun
Applicant	AlmondMedia Co., Ltd.
Rightholder	AlmondMedia Co., Ltd.

### 발명의 명칭

자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 시스템

### Title of Invention

SYSTEM FOR PROVIDING AUTO TRANSLATE BASED GLOBAL ONLINE COMMUNITY SERVICE

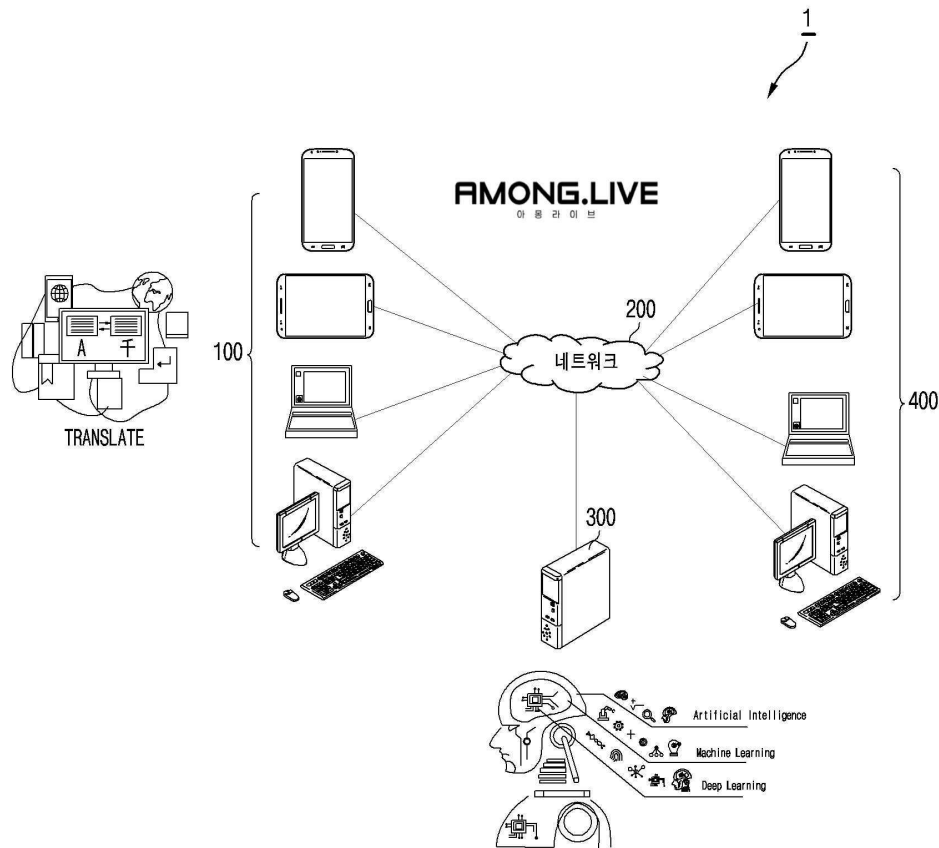
### 요약

자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 시스템이 제공되며, 언어설정을 감지하여 적어도 하나의 종류의 멀티미디어 콘텐츠 내 포함된 다국어 텍스트를 감지된 언어로 자동번역하여 출력하는 사용자 단말 및 적어도 하나의 주제를 가지는 게시판 내 적어도 하나의 종류의 멀티미디어 콘텐츠에 대응하는 게시글이 등록되면 게시글을 업로드하는 업로드부, 사용자 단말에서 게시글을 감지된 언어로 출력하도록 자동번역을 수행하는 자동번역부, 게시글에 댓글이 포함된 경우 댓글을 감지된 언어로 자동번역하여 출력하는 댓글번역부를 포함하는 커뮤니티 서비스 제공 서버를 포함한다.

### Abstract

The machine translation base global on-line community service providing system is provided and the community service providing server which includes the upload part, the automatic translation unit, and the reply translation if the posted writing corresponding to the multimedia contents of the kind of the user terminal which senses the language institution and which automatically translates the multi-lingual text included with the multimedia contents of at least one kind to the sensed language and outputted and inside of the bulletin board at least one having at least one theme is registered is included. The automatic translation unit performs the machine translation in the user terminal in order to output the posted writing to the sensed language. The reply translation automatically translates the reply to the sensed language in case the reply is included in the posted writing and it outputs.

### 대표도면 (Representative drawing)



## 청구의 범위

### 청구 1항:

언어설정을 감지하여 적어도 하나의 종류의 멀티미디어 콘텐츠 내 포함된 다국어 텍스트를 감지된 언어로 자동번역하여 출력하는 사용자 단말; 및 커뮤니티 서비스 제공 서버를 포함하되,

상기 커뮤니티 서비스 제공 서버는,

적어도 하나의 주제를 가지는 게시판 내 상기 적어도 하나의 종류의 멀티미디어 콘텐츠에 대응하는 게시글이 등록되면 상기 게시글을 업로드하는 업로드부;

상기 사용자 단말에서 상기 게시글을 상기 감지된 언어로 출력하도록 자동번역을 수행하는 자동번역부;

상기 게시글에 댓글이 포함된 경우 상기 댓글을 상기 감지된 언어로 자동번역하여 출력하는 댓글번역부;

상기 사용자 단말에서 출력하는 멀티미디어 콘텐츠가 오디오 콘텐츠인 경우, 상기 오디오 콘텐츠 내 음성발화한 발화자의 음성을 포함하여 확보된 다수의 동영상 내 크리에이터의 음성을 신경망 모델로 학습시키고, 이후 음성발화를 텍스트로 변환하는 STT(Speech to Text)를 실시하고 자동번역을 실시한 이후에, 신경망 모델로 학습된 상기 음성발화한 발화자의 음성 사용자 단말에 설정된 언어로 자동번역된 텍스트를 입력하여 오디오 콘텐츠를 출력시킴으로써 상기 자동번역된 텍스트에 대하여 상기 음성발화한 발화자에 의한 실시간 더빙이 이루어지도록 하는 자동더빙부;

## Scope of Claims

### Claim 1:

The language institution is sensed and the multi-lingual text included with the multimedia contents my of at least one kind automatically is translated to the sensed language and the user terminal outputted, and the community service providing server are included ; the multimedia contents outputted in the reply translation :

user terminal which it automatically translates the reply to the above-mentioned sensed language in case the reply is included in the automatic translation unit :

posted writing which performs the machine translation in the upload part :

user terminal which \*\*\*s the posted writing if the posted writing corresponding to the multimedia contents of the kind of inside of the bulletin board at least one in which

community service providing server has the theme of

at least one is registered in order to output the posted writing to the above-mentioned sensed language, and it outputs is the audio content ; and the voice of the multiple inside of the moving picture the creators secured including the voice of the speaker done with the audio content shade tolerance firing is cultivated to the neural network model. Then, the STT (Speech to Text) c

상기 적어도 하나의 종류의 멀티미디어 콘텐츠 내 신경망 기반 기계번역 모델을 이용하여 자동번역이 된 기계번역 모델 기반 자동번역 결과물에 대하여 서로 다른 주체에 귀속된 자동번역된 결과물을 제공받는 복수의 사용자 단말로부터 평점을 수집하고, 수집된 평점이 기 설정된 점수 이하인 경우, 상기 기 설정된 점수 이하의 평점을 갖는 기계번역 모델 기반 자동번역 결과물을 적어도 하나의 관리자 단말로 제공하여 검수를 요청하는 클라우드소싱부;

상기 사용자 단말에서 원문 텍스트의 문장을 기계번역 모델 기반으로 자동 번역한 자동번역된 텍스트의 문장이 출력되도록 할 때, 사용자 단말의 사용자가 자동번역된 텍스트의 문장 중 오역된 단어에 대응하는 영역을 지정하여 일정 시간 탭한 경우, 상기 일정 시간 탭된 오역된 단어에 대하여 수정요청이 이루어진 것으로 인식하여 상기 원문 텍스트의 문장 상에서 상기 오역된 단어에 대응하는 오역 대응 원문 단어의 영역을 표시되도록 하는 부분오역검수부;

신경망 기반 기계번역 모델 중 자기집중(Self-Attention) 기반 Transformer 모델을 이용하여 자동번역을 실시하는 인공지능부;

상기 자동번역을 실시하기 이전에 다국어 텍스트를 언어별로 분리한 다음 감성분석을 진행하고, 극성치를 종합하는 다중언어 감성분석 프로세스를 진행하는 감성분석부;

상기 사용자 단말에서 출력하는 멀티미디어 콘텐츠가 웹툰인 경우, 웹툰 내 텍스트를 OCR로 추출할 때 추출된 영역의 좌표 및 크기를 저장하고, 텍스트를 자동번역한 후 번역된 글자수를 카운트하여 상기 영역의 크기 내에 포함되도록 글씨 크기, 폭 및 간격을 조절하여 상기 추출된 영역으로 오버레이하는 웹툰관리부; 및

상기 사용자 단말에서 상기 자동번역된 텍스트의 수정을 한 경우, 수정 후 버전과 수정 전 버전을 동시에 게재하여 번역평가를 로그로 누적하는 수정평가부를 포함하고,

상기 자동번역된 텍스트의 문장 내 상기 오역된 단어는, 상기 수정요청에 응답하여 상기 사용자 단말에서 상기 오역 대응 원문 단어에 대응하는 정역 단어로 즉시 바꾸거나 관리자에 의한 검수 과정을 통해 정역 단어로 바뀌어 반영되고,

상기 자동더빙부에 의하면, 상기 실시간 더빙의 수행 후 상기 출력되는 오디오 콘텐츠 내 발화자의 발화 목소리에 대응하는 언어는 상기 실시간 더빙의 수행 전 상기 오디오 콘텐츠 내 발화자의 발화 목소리에 대응하는 언어와는 다른 언어이고,

상기 실시간 더빙의 수행 후 상기 출력되는 오디오 콘텐츠에서 발화자가 발화하는 상기 다른 언어의 발화 목소리와 톤과 음정은 상기 실시간 더빙의 수행 전 상기 오디오 콘텐츠 내 발화자와 동일 발화자가 발화하는 발화 목소리와 톤과 음정을 갖는 것이며,

상기 적어도 하나의 종류의 멀티미디어 콘텐츠는, 적어도 하나의 종류의 소셜미디어, HTML 웹게임, 유튜브 라이브 및 웹툰을 포함하는 것인, 자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 시스템.

onverting the voice firing into the text is performed and the machine translation is performed ; and the grade is collected from multiple user terminals which receive the outcome which becomes about the machine translation model-based automatic translation result water in which the machine translation is using the multimedia contents my neural network base machine translation model of the kind of the automatic dubbing part :

at least one in which the real-time dubbing by the speaker which does about the text which becomes with the machine translation with the voice firing is comprised of machine translation returned to the different subject by plating the text consisting of the language set up in the user terminal with machine translation on the voice of the speaker which is done by the neural network model with the voice firing learned and printing out the audio content The collected grade is the predetermined score or less ; the sentence of the text which becomes with machine translation automatically translates the sentence of the original text to the machine translation model-based is outputted in the crowd sourcing part :

user terminal providing the machine translation model-based automatic translation result water having the grade less than the above-mentioned predetermined score to at least one manager terminal and requests the inspection ; the user of the user terminal designates the domain corresponding with the mistranslated word among the sentence of the text which becomes with machine translation and it taps the preset time ; and it recognizes that the modify request about the mistranslated word which the preset time is tapped is made and after the multi-lingual text is separated according to language before performing the artificial intelligence part :

machine translation performing the machine translation on the sentence of the original text among the part mistranslation inspection unit :

neural network base machine translation model which is the domain of the mistranslation action text word corresponding to the above-mentioned mistranslated word indicated using the self-convergence (Self-Attention) base Transformer model the sensitivity analysis is progressed. The multimedia contents outputted in the sensitivity analysis part :

user terminal progressing the multi-lingual sensitivity analysis process synthesizing the polarity value is the webtoon ; the coordinate and size of the extracted domain are stored when extracting inside of the webtoon the text in OCR ; font size, the case of modifying of the text, and the version after the correction and the correction evaluation portion are included in order to count the number of characters which is the text translated after doing the machine translation and be included within the size of the domain ; immediately the word mistranslated inside of the sentence the above of the text which becomes with

machine translation changes into the forward reverse word corresponding with the mistranslation action text word in response to the modify request in the user terminal or it is changed to the forward reverse word through the inspection process by the administrator and it i

s reflected ; and the language corresponding to the firing voice of the above-mentioned inside of the audio content the outputted speaker according to

automatic dubbing part after the performance of the real-time dubbing is the language different from the language corresponding to the firing voice of the performance advanced audio content my speaker of the real-time dubbing, and the case of modifying of the text consists of the extracted domain in the webtoon administration unit :

of overlaying and user terminal with the machine translation it adjusts the width and gap; and the version after the correction and the correction evaluation portion simultaneously reports the patient to call version and accumulates the translation evaluation under the log.

The machine translation base global on-line community service providing system in which the firing voice of the above-mentioned dissimilar language in which the speaker starts from the above-mentioned outputted audio content after the performance of the real-time dubbing and tone and musical interval includes the multimedia contents, of the kind of

at least one is the social media of at least one kind it has as the firing voice, HTML web game, and the U-tube live and webtoon, and as to the multimedia contents, of the kind of

at least one is the social media of at least one kind it has as the, the same speaker as the performance advanced audio content my speaker of the real-time dubbing starts tone and musical interval.

**청구 2항:**

제 1 항에 있어서,

상기 적어도 하나의 종류의 멀티미디어 콘텐츠는,

이미지, 동영상, 음성 및 텍스트 중 어느 하나 또는 적어도 하나의 조합으로 생성된 콘텐츠인 것을 특징으로 하는 자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 시스템.

**Claim 2:**

As for claim 1, the machine translation base global on-line community service providing system which is the contents generated to any one of multimedia contents, of the kind of

at least one is

image, the moving picture, and voice and text or at least one combination.

**청구 3항:**

삭제

**Claim 3:**

Deletion.

**청구 3항:**

삭제

**Claim 3:**

Deletion.

**청구 4항:**

삭제

**Claim 4:**

Deletion.

**청구 4항:**

삭제

**Claim 4:**

Deletion.

**청구 5항:**

삭제

**Claim 5:**

Deletion.

청구 5항: 삭제	Claim 5: Deletion.
청구 6항: 삭제	Claim 6: Deletion.
청구 6항: 삭제	Claim 6: Deletion.
청구 7항: 삭제	Claim 7: Deletion.
청구 7항: 삭제	Claim 7: Deletion.
청구 8항: 삭제	Claim 8: Deletion.
청구 8항: 삭제	Claim 8: Deletion.
청구 9항: 삭제	Claim 9: Deletion.
청구 9항: 삭제	Claim 9: Deletion.
청구 10항: 삭제	Claim 10: Deletion.
청구 10항: 삭제	Claim 10: Deletion.

## 기술분야

본 발명은 자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 시스템에 관한 것으로, 이미지, 동영상, 음성 및 텍스트를 실시간으로 자동번역하여 제공하는 글로벌 커뮤니티 플랫폼을 제공한다.

## Technical Field

The invention provides the image as the machine translation base global on-line community service providing system, the moving picture, and voice, and the global community platform which automatically translates the text to on a real time basis and provided.

## 배경기술

전 세계적으로 문화가 서로 융합됨에 따라 비즈니스 측면에서의 고객도 다언어권으로 바뀌어 가고 있으며, 각 언어권의 사용자들이 동일한 콘텐츠를 보고 서로 간의 경험이나 느낌 등을 한 사이트에 공유할 수 있다. 한국도 한류의 영향으로 동북아 및 동남아에 한류 마케팅이 가능하게 되었는데, 동일한 콘텐츠에 여러 가지 언어로 쓰인 리뷰나 댓글이 등장하는 현상이 점점 많아지고 있다. 대표적으로 동영상 공유 사이트인 유튜브나 사진 공유를 위한 인스타그램, 호텔예약 사이트 Expedia 등에서 사람들은 언어에 제한 받지 않고 자신의 의견이나 관점을 댓글로 자유롭게 표현하고 있다. 그 이유는 Facebook이나 Twitter와 같이 지인들 사이의 정보 공유나 의사소통을 돕는 유형이 상호 이해 가능한 언어로 작성되는 것과는 달리, 이 사이트들은 공유된 콘텐츠에 대해 각자의 생각을 담기 때문에 다양한 언어로 구성되기 때문이다.

## Background Art

The client at the business side is changed to the multi language Keown as the culture fuses to universal and the contents in which the users of each language are identical can be examined and the experience between them or the feeling etc. can be shared in one site. The current-limiting marketing the Korea was possible the current-limiting in the Northeast Asia and Southeast Asia. The review used as various languages in the same contents or the phenomenon that the reply makes an appearance more and more increases. Remarkably, in the instar gram for the U-tube which is the moving picture share site or the photograph share, the hotel reservation site Expedia etc, the people does not receive the limit in language and its own opinion or the viewpoint is

freely expressed in the reply. Contrary to that the type in which reason for the helps with the information sharing between the acquaintance or the communication like the Facebook or the Twitter is prepared to the language which is understandable reciprocity because its own thinking is put in about the contents in which these sites are shared it is due to consist of the multiplex language.

이때, 실시간 스트리밍, 웹툰 및 소셜미디어 내의 다국어를 실시간으로 번역하는 플랫폼이 연구 및 개발되었는데, 선행기술인 한국공개특허 제2020-0090355호(2020년07월29일 공개), 한국공개특허 제2017-0008357호(2017년01월24일 공개) 및 한국등록특허 제10-1743686호(2017년06월20일 공고)에는, 멀티미디어 콘텐츠의 실시간 스트리밍에서 실시간 번역을 지원하도록 카메라를 통해 입력되는 영상으로부터 입술모양을 추출하고 음성발화로부터 텍스트를 추출한 후 기 설정된 언어로 번역 후 입력영상에 번역된 텍스트를 자막으로 추가하는 구성과, 웹툰에서 사용빈도가 기준값 이상인 구어체 문장 및 어구를 번역데이터베이스에 저장하고, 자동번역가능한 번역단위로 분할하며, 번역단위별 형태소 분석을 통하여 언어를 감지하고 번역단위에 대해 기 번역된 데이터가 존재하면 유사도에 기반하여 번역을 수행하며, 번역문장에 대한 번역 불일치 횟수에 기반하여 품질을 점검하는 구성과, 인터넷 채팅, 이메일, 텍스트 기반 통신, 온라인 포스팅 및 소셜미디어를 포함하는 텍스트를 번역할 때 정확하지 않거나 오류가 있는 번역에 대한 정정을 크라우드 소싱으로 제출하도록 하고, 번역의 정확도를 평가하기 위한 보상 시스템을 구축하는 구성이 각각 개시되어 있다.

Then, when translating the text including real time streaming, the platform, the composition department, and the composition department, internet chat, E-mail, text base communication, online posting and the social media which stores the sentence of spoken language and words and phrases, the correction about the translation which is not exact or in which the translation has the error is submitted the crowd sourcing The platform translates the multi-lingual within the webtoon and social media on a real time basis is the research and KR2020-0090355 A (2020 year July 29 disclosure) which developed is the prior art. The composition department extracts the lip shape from the image inputted through the camera in KR2017-0008357 A (2017 year January 24 disclosure) and KR10-1743686 B (2017 year June 20 notification) in order to support the real time translation in the real time streaming of the multimedia contents and , adds the translated text to the text from the voice firing after doing the extraction to the predetermined language after the translation in the input image to the subtitle. As to the composition department, internet chat, E-mail, text base communication, online, the usage frequency is the reference value or greater in the translation database in the webtoon and it divides by unit of the translation which is possible with machine translation and it performs the translation based on the similarity if it senses language through the translate unit morphological analysis and already translated data exist about the translate unit and checks out quality based on the translation inconsistent count about the translated sentence. The configuration of building the compensation system for evaluating the accuracy of the translation is disclosed.

다만, 번역을 하게 될 경우, 언어와 관련된 배경이나 문화, 용어사용의 차이 등으로 본래 문서에 있는 모든 단어나 문법을 정확히 표현할 수 없고, 번역 알고리즘에서 사용하는 사전에 의해 번역하는 것 자체에 착오가 존재하는데, 번역 과정에서 단어의 뜻이 달라지는 경우가 있어 감성의 미묘한 왜곡이 일어나는 단점을 가진다. 그러므로 언어의 번역과정에서 일어나는 극성치의 변경 현상을 최소화하는 방법이 필요하다. 이에, 다중언어로 수집되는 텍스트를 바로 번역하지 않고 언어별로 텍스트를 분리한 다음 감성분석을 진행하여 각각의 극성치를 종합한 후 번역을 진행함으로써 번역 알고리즘이나 기계번역에서 발생하는 오류를 제로화할 수 있는 플랫폼의 연구 및 개발이 요구된다.

But when it translates all words or the grammar which is inherently in the document to background associated with language or the culture, the difference of the use of term etc. cannot be accurately expressed and it translates the mistake exists with the prior used in the interpretation algorithm in itself. It has the case where the meaning of word is changed from the translation process it has the disadvantage that the distortion of being subtle of nature occurs. Therefore, it need the method for minimizing the change phenomenon of the polarity value occurring in the translation process of language. Thus, the research and development of the platform which immediately can not translate the collected text into the multi-lingual and zero the error generated in the interpretation algorithm or the machine translation after it separates the text according to language by progressing the sensitivity analysis and the translation each polarity value after doing the synthesis are requested.

## 발명의 내용

해결하고자 하는 과제

## Summary of Invention

Problem to be solved

본 발명의 일 실시예는, 글로벌 온라인 커뮤니티 내의 동영상, 이미지, 음성 및 텍스트를 사용자 설정언어로 실시간 번역하여 제공하고, 웹툰과 같은 이미지 내 텍스트는 이미지로부터 텍스트를 추출하는 기술을 적용하여 실시간 자동번역한 후 다시 웹툰 내 이미지 내에 삽입 및 오버레이하도록 하며, STT (Speech to Text)로 음성을 텍스트로 변환한 후 자동번역한 텍스트를 멀티미디어 콘텐츠에 제공할 때 발화자의 음성을 학습한 후 자동번역한 텍스트를 TTS(Text to Speech)로 제공함으로써 화자의 목소리로 번역된 언어를 발화하는 것과 같은 느낌을 받을 수 있도록 하고, 다중언어로 수집되는 텍스트를 바로 번역하지 않고 언어별로 텍스트를 분리한 다음 감성분석을 진행하여 각각의 극성치를 종합한 후 번역을 진행함으로써 번역 알고리즘이나 기계번역에서 발생하는 오류를 제로화할 수 있는, 자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 방법을 제공할 수 있다. 다만, 본 실시예가 이루고자 하는 기술적 과제는 상기된 바와 같은 기술적 과제에 한정되지 않으며, 또 다른 기술적 과제들이 존재할 수 있다.

A preferred embodiment of the present invention translates the moving picture within the global on-line community, the image, and voice and text into the user setting language on a real-time basis and it provides and it applies the technology in which in image text like the webtoon extracts the text from the image and it again inserts after doing the real time automatic translation within inside of the webtoon the image and it overlays and the technology receives the feeling the language translated into the voice of the speaker by providing the text which automatically translates the voice of the speaker after doing learning when providing voice the text which automatically translates after doing the conversion to the STT (Speech to Text) to the text to the multimedia contents to the TTS (Text to Speech) and it can provide the machine translation base global on-line community service providing method for zeroing the error generated in the interpretation algorithm or the machine translation by it progresses the sensitivity analysis after it immediately does not translate the collected text into the multi-lingual and it separates the text according to language and progressing the translation each polarity value after doing the synthesis . But it is not restricted to the technical problem in which the technical problem in which this embodiment is comprised is described vividly and another technical problems can exist.

### 과제해결 수단

상술한 기술적 과제를 달성하기 위한 기술적 수단으로서, 본 발명의 일 실시예는, 언어설정을 감지하여 적어도 하나의 종류의 멀티미디어 콘텐츠 내 포함된 다국어 텍스트를 감지된 언어로 자동번역하여 출력하는 사용자 단말 및 적어도 하나의 주제를 가지는 게시판 내 적어도 하나의 종류의 멀티미디어 콘텐츠에 대응하는 게시글이 등록되면 게시글을 업로드하는 업로드부, 사용자 단말에서 게시글을 감지된 언어로 출력하도록 자동번역을 수행하는 자동번역부, 게시글에 댓글이 포함된 경우 댓글을 감지된 언어로 자동번역하여 출력하는 댓글번역부를 포함하는 커뮤니티 서비스 제공 서버를 포함한다.

### Means to solve the problem

A preferred embodiment of the present invention it is the technical means for achieving the above-described technical problem includes the user terminal which senses the language institution and which automatically translates the multi-lingual text included with the multimedia contents my of at least one kind to the sensed language and outputted and the community service providing server which includes the upload part \*\*\*ing the posted writing, the automatic translation unit, and the reply translation if the posted writing corresponding to the multimedia contents of the kind of the inside of the bulletin board at least one having at least one theme is registered. The automatic translation unit performs the machine translation in the user terminal in order to output the posted writing to the sensed language. The reply translation automatically translates the reply to the sensed language in case the reply is included in the posted writing and it outputs.

### 발명의 효과

전술한 본 발명의 과제 해결 수단 중 어느 하나에 의하면, 글로벌 온라인 커뮤니티 내의 동영상, 이미지, 음성 및 텍스트를 사용자 설정언어로 실시간 번역하여 제공하고, 웹툰과 같은 이미지 내 텍스트는 이미지로부터 텍스트를 추출하는 기술을 적용하여 실시간 자동번역한 후 다시 웹툰 내 이미지 내에 삽입 및 오버레이하도록 하며, STT(Speech to Text)로 음성을 텍스트로 변환한 후 자동번역한 텍스트를 멀티미디어 콘텐츠에 제공할 때 발화자의 음성을 학습한 후 자동번역한 텍스트를 TTS(Text to Speech)로 제공함으로써 화자의 목소리로 번역된 언어를 발화하는 것과 같은 느낌을 받을 수 있도록 하고, 다중언어로 수집되는 텍스트를 바로 번역하지 않고 언어별로 텍스트를 분리한 다음 감성분석을 진행하여 각각의 극성치를 종합한 후 번역을 진행함으로써 번역 알고리즘이나 기계번역에서 발생하는 오류를 제로화할 수 있다.

### Effects of the Invention

The moving picture within the above-described global on-line community, the image, and voice and text are translated into the user setting language on a real-time basis and it provides and the technology in which in image text like the webtoon extracts the text from the image is applied and it again inserts after doing the real time automatic translation within inside of the webtoon the image and it overlays and the language which is voice translated into the STT (Speech to Text) into the voice of the speaker by providing the text which automatically translates the voice of the speaker after doing learning when providing the text which automatically translates after doing the conversion to the text to the multimedia contents to the TTS (Text

to Speech) feeling is received and the collected text can not be immediately translated into the multi-lingual and after the text is separated according to language by the sensitivity analysis being progressed and progressing the translation each polarity value after doing the synthesis the error generated in the interpretation algorithm or the machine translation can be zeroed.

## 발명을 실시하기 위한 구체적인 내용

아래에서는 첨부한 도면을 참조하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 본 발명의 실시예를 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

명세서 전체에서, 어떤 부분이 다른 부분과 #34#연결#34#되어 있다고 할 때, 이는 #34#직접적으로 연결#34#되어 있는 경우뿐 아니라, 그 중간에 다른 소자를 사이에 두고 #34#전기적으로 연결#34#되어 있는 경우도 포함한다. 또한 어떤 부분이 어떤 구성요소를 #34#포함#34#한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미하며, 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.

명세서 전체에서 사용되는 정도의 용어 #34#약#34#, #34#실질적으로#34# 등은 언급된 의미에 고유한 제조 및 물질 허용오차가 제시될 때 그 수치에서 또는 그 수치에 근접한 의미로 사용되고, 본 발명의 이해를 돕기 위해 정확하거나 절대적인 수치가 언급된 개시 내용을 비양심적인 침해자가 부당하게 이용하는 것을 방지하기 위해 사용된다. 본 발명의 명세서 전체에서 사용되는 정도의 용어 #34#~(하는) 단계#34# 또는 #34#~의 단계#34#는 #34#~를 위한 단계#34#를 의미하지 않는다.

본 명세서에 있어서 #39#부(部)#39#란, 하드웨어에 의해 실현되는 유닛(unit), 소프트웨어에 의해 실현되는 유닛, 양방을 이용하여 실현되는 유닛을 포함한다. 또한, 1 개의 유닛이 2 개 이상의 하드웨어를 이용하여 실현되어도 되고, 2 개 이상의 유닛이 1 개의 하드웨어에 의해 실현되어도 된다. 한편, #39#~부#39#는 소프트웨어 또는 하드웨어에 한정되는 의미는 아니며, #39#~부#39#는 어드레싱 할 수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 따라서, 일 예로서 #39#~부#39#는 소프트웨어 구성요소들, 객체 지향 소프트웨어 구성요소들, 클래스 구성요소들 및 태스크 구성요소들과 같은 구성요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로코드, 회로, 데이터, 데이터베이스, 데이터 구조들, 테이블들, 어레이들 및 변수들을 포함한다. 구성요소들과

## Description of Embodiments

In the lower part, the embodiment of the invention is particularly illustrated in the technical field in which the invention belongs with reference to the attached drawing so that a person skilled in the art easily performs. But it is not restricted to the embodiment which here it illustrates while the invention can be implemented as the form which the various disagrees. And in drawing, the part which had no concern with the description it specifically illustrated the invention omitted and the reference numeral which was similar was adhered through the specification whole about the part which was similar.

In the specification whole, when it becomes with the part and the " connection " , any kind of part is different even in case of the device different in the center not only the case, where this consists of the " direct with the connection " being placed in the interval and consisting of the " electrical with the connection " it includes. Moreover, when any kind of part any kind of element the " inclusion " , it means the dissimilar element without the material in which this is specially opposed not being excluded but maying further include the dissimilar element and the presence of one or more the dissimilar characteristic or the number, the step, the operation, the element, the components or the things assembling these or the part does not exclude the possibility in advance and it has to be understood.

When manufacture and the material allowable error inherent in the term " about " of the extent of being used in the specification whole, and the meaning that the etc. is substantially mentioned with " are presented, or it is used for the meaning near to the value in the value and so that it prevent the dishonest pirate unreasonably from using the disclosure that the absolute value which is exact or is mentioned, it is used. The step " of the term " ~ (it does) step " of the extent or the "~ of being used in the specification whole of the invention does not mean the step " for the "~.

In this specification, the ' part ' includes the realized unit, the unit realized by software, and the unit realized using both with hardware. Moreover, one unit is realized using at least 2 hardwares and the unit more than 2 is realized by one hardware. In the meantime, it is not meaning that the '~ part ' is restricted to software or hardware and it is configured to be in the storage medium which the '~ part ' can address and one or more processors is configured to be brought back. Therefore, as an example, the '~ part ' comprises software components, object-oriented software components, the elements like the class components and task components, processes, functions, the segments of the attributes, procedures, subroutines, the program code the drivers, firmware, the micro code, the circuit, data, the database, dat



#39#~부#39#들 안에서 제공되는 기능은 더 작은 수의 구성요소들 및 #39#~부#39#들로 결합되거나 추가적인 구성요소들과 #39#~부#39#들로 더 분리될 수 있다. 뿐만 아니라, 구성요소들 및 #39#~부#39#들은 디바이스 또는 보안 멀티미디어카드 내의 하나 또는 그 이상의 CPU들을 재생성하도록 구현될 수도 있다.

본 명세서에 있어서 단말, 장치 또는 디바이스가 수행하는 것으로 기술된 동작이나 기능 중 일부는 해당 단말, 장치 또는 디바이스와 연결된 서버에서 대신 수행될 수도 있다. 이와 마찬가지로, 서버가 수행하는 것으로 기술된 동작이나 기능 중 일부도 해당 서버와 연결된 단말, 장치 또는 디바이스에서 수행될 수도 있다.

본 명세서에서 있어서, 단말과 매핑(Mapping) 또는 매칭(Matching)으로 기술된 동작이나 기능 중 일부는, 단말의 식별 정보(Identifying Data)인 단말기의 고유번호나 개인의 식별 정보를 매핑 또는 매칭한다는 의미로 해석될 수 있다.

이하 첨부된 도면을 참고하여 본 발명을 상세히 설명하기로 한다.

도 1은 본 발명의 일 실시예에 따른 자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 시스템을 설명하기 위한 도면이다. 도 1을 참조하면, 자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 시스템(1)은, 적어도 하나의 사용자 단말(100), 커뮤니티 서비스 제공 서버(300), 적어도 하나의 업로더 단말(400)을 포함할 수 있다. 다만, 이러한 도 1의 자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 시스템(1)은, 본 발명의 일 실시예에 불과하므로, 도 1을 통하여 본 발명이 한정 해석되는 것은 아니다.

이때, 도 1의 각 구성요소들은 일반적으로 네트워크(Network, 200)를 통해 연결된다. 예를 들어, 도 1에 도시된 바와 같이, 적어도 하나의 사용자 단말(100)은 네트워크(200)를 통하여 커뮤니티 서비스 제공 서버(300)와 연결될 수 있다. 그리고, 커뮤니티 서비스 제공 서버(300)는, 네트워크(200)를 통하여 적어도 하나의 사용자 단말(100), 적어도 하나의 업로더 단말(400)과 연결될 수 있다. 또한, 적어도 하나의 업로더 단말(400)은, 네트워크(200)를 통하여 커뮤니티 서비스 제공 서버(300)와 연결될 수 있다.

여기서, 네트워크는, 복수의 단말 및 서버들과 같은 각각의 노드 상호 간에 정보 교환이 가능한 연결 구조를 의미하는 것으로, 이러한 네트워크의 일 예에는 근거리 통신망(LAN: Local Area Network), 광역 통신망(WAN: Wide Area Network), 인터넷(WWW: World Wide Web), 유무선 데이터 통신망, 전화망, 유무선 텔레비전 통신망 등을 포함한다. 무선 데이터 통신망의 일례에는 3G, 4G, 5G, 3GPP(3rd Generation Partnership Project), 5GPP(5th Generation Partnership Project), LTE(Long Term Evolution), WIMAX(World Interoperability for Microwave Access), 와이파이(Wi-Fi), 인터넷(Internet), LAN(Local Area Network), Wireless LAN(Wireless Local Area Network), WAN(Wide Area Network), PAN(Personal Area Network), RF(Radio Frequency

a constructs, tables, arrays, and variables. The function of being provided with elements in '~ part 's is more combined to the elements and '~ part 's of the small number or the function can be separated from additional elements from '~ part 's. Besides, it can be implemented so that the element and '~ part 's bring back one or more CPUs within the device or the security multimedia card.

In this specification, it can be instead of performed in the operation which the terminal, and the apparatus or the device perform and which is described or the server in which the part is connected to the corresponding terminal, and the apparatus or the device among the function. It can be performed in the operation which the server performs like this and which is described or the terminal in which the part is connected among the function to the pertinent server, and the apparatus or the device.

It has in this specification and it can be construed as the operation described to the terminal and mapping or the matching or the meaning of the part mapping the generic number or the individual identifying information of the terminal which is the identifying information (Identifying Data) of the terminal or matching among the function.

The invention decides to be particularly illustrated the drawing attached below.

Figure 1 is a drawing illustrating the machine translation base global on-line community service providing system according to a preferred embodiment of the present invention. Referring to Figure 1, it may include at least one user terminal (100), the community service providing server (300), and at least one up loader terminal (400). But the machine translation base global on-line community service providing system (1) of such drawing 1 is nothing but a preferred embodiment of the present invention. Therefore the invention is not analyzed limitedly through fig. 1.

Then, generally each element of fig. 1 s are connected through networks (Network, 200). For example, as shown in FIG. 1, at least one user terminal (100) can be connected through the network (200) to the community service providing server (300). And the community service providing server (300) can be connected through the network (200) to at least one user terminal (100), and at least one up loader terminal (400). Moreover, at least one up loader terminal (400) can be connected through the network (200) to the community service providing server (300).

Here, the network includes the local area network (LAN: Local Area Network), the wide area network (WAN: Wide Area Network), the internet (WWW: World Wide Web), the internet service provider, the telephone network, wire and wireless television communication network etc in an example of such network the connection structure in which the information exchange is possible between each node condition call is meant, such as the multiple terminals and servers. In an example of the wireless data communication net, 3G, 4G, 5G, 3 generation partnership project (3rd Generation Partnership Project), 5 GPP (5th Generation Partnership Project), the LTE (Long Term Evolution), the WIMAX (World Interopera

y), 블루투스(Bluetooth) 네트워크, NFC(Near-Field Communication) 네트워크, 위성 방송 네트워크, 아날로그 방송 네트워크, DMB(Digital Multimedia Broadcasting) 네트워크 등이 포함되나 이에 한정되지는 않는다.

하기에서, 적어도 하나의 라는 용어는 단수 및 복수를 포함하는 용어로 정의되고, 적어도 하나의 라는 용어가 존재하지 않더라도 각 구성요소가 단수 또는 복수로 존재할 수 있고, 단수 또는 복수를 의미할 수 있음은 자명하다 할 것이다. 또한, 각 구성요소가 단수 또는 복수로 구비되는 것은, 실시예에 따라 변경가능하다 할 것이다.

적어도 하나의 사용자 단말(100)은, 자동번역 기반 글로벌 온라인 커뮤니티 서비스 관련 웹 페이지, 앱 페이지, 프로그램 또는 애플리케이션을 이용하여 글로벌 온라인 커뮤니티에 접속하여 멀티미디어 콘텐츠를 출력하는 사용자의 단말일 수 있다. 이때, 사용자 단말(100)은 단말 내 언어설정값을 제공하거나 직접 커뮤니티 페이지에서 설정하고, 멀티미디어 콘텐츠를 설정된 언어로 자동번역된 콘텐츠로 출력하는 단말일 수 있다.

여기서, 적어도 하나의 사용자 단말(100)은, 네트워크를 통하여 원격지의 서버나 단말에 접속할 수 있는 컴퓨터로 구현될 수 있다. 여기서, 컴퓨터는 예를 들어, 네비게이션, 웹 브라우저(WEB Browser)가 탑재된 노트북, 데스크톱(Desktop), 랩톱(Laptop) 등을 포함할 수 있다. 이때, 적어도 하나의 사용자 단말(100)은, 네트워크를 통해 원격지의 서버나 단말에 접속할 수 있는 단말로 구현될 수 있다. 적어도 하나의 사용자 단말(100)은, 예를 들어, 휴대성과 이동성이 보장되는 무선 통신 장치로서, 네비게이션, PCS(Personal Communication System), GSM(Global System for Mobile communications), PDC(Personal Digital Cellular), PHS(Personal HandypHONE System), PDA(Personal Digital Assistant), IMT(International Mobile Telecommunication)-2000, CDMA(Code Division Multiple Access)-2000, W-CDMA(W-Code Division Multiple Access), Wibro(Wireless Broadband Internet) 단말, 스마트폰(Smartphone), 스마트 패드(Smartpad), 태블릿 PC(Tablet PC) 등과 같은 모든 종류의 핸드헬드(Handheld) 기반의 무선 통신 장치를 포함할 수 있다.

커뮤니티 서비스 제공 서버(300)는, 자동번역 기반 글로벌 온라인 커뮤니티 서비스 웹 페이지, 앱 페이지, 프로그램 또는 애플리케이션을 제공하는 서버일 수 있다. 그리고, 커뮤니티 서비스 제공 서버(300)는, 적어도 하나의 게시판을 마련하고 적어도 하나의 게시판 내 적어도 하나의 업로더 단말(400)에서 멀티미디어 콘텐츠를 업로드하는 경우, 멀티미디어 콘텐츠의 언어를 멀티미디어 콘텐츠와 매핑하여 저장하는 서버일 수 있다. 또한, 커뮤니티 서비스 제공 서버(300)는, 사용자 단말(100)에서 멀티미디어 콘텐츠를 선택하여 출력하는 경우 사용자가 설정한 언어로 자동번역하여 출력되도록 하는 서버일 수 있다.

bility for Microwave Access), the Wi-Fi, the internet, the LAN (Local Area Network), the Wireless LAN (Wireless Local Area Network), the WAN (Wide Area Network), the PAN (Personal Area Network), the RF (Radio Frequency), the Bluetooth network, the NFC (Near-Field Communication) network, the satellite broadcasting network, the analog broadcasting network, the DMB (Digital Multimedia Broadcasting) network etc is included but it is not thus restricted.

Below, the term is defined as the term including the singular number and plurality and each element can exist as the suspension of water supply or the plurality although the term is nonexistent and the suspension of water supply or the plurality can be meant it is obvious it will do. Moreover, it is changeable it will do according to the embodiment that each element is equipped to the suspension of water supply or the plurality.

At least one user terminal (100) can be the terminal of the user connecting to the global on-line community using the machine translation base global online community service linked web page, the app page, and the program or the application and outputs the multimedia contents. Then, inside of the terminal the language institution value is provided or the user terminal (100) sets up in the direct community page and it can be the terminal outputting the multimedia contents to the contents consisting of the language set up with machine translation.

Here, at least one user terminal (100) can be materialized in computer it can connect to the server or the terminal of the remote location through the network. Here, the computer, for example, can comprise the notebook, in which navigation, and the web browser are mounted the desktop, the laptop etc. Then, at least one user terminal (100) can be implemented as the terminal connecting to the server or the terminal of the remote location through the network. At least one user terminal (100), for example, can comprise the wireless communication apparatus of the hand held (Handheld) base of all kinds like the navigation, the PCS (Personal Communication System), the GSM (Global System for Mobile communications), the PDC (Personal Digital Cellular), the PHS (Personal HandypHONE System), the PDA (Personal Digital Assistant), the IMT (International Mobile Telecommunication) -2000, the CDMA (Code Division Multiple Access) -2000, the wide band-code division multiple access (W-Code Division Multiple Access), the Wibro (Wireless Broadband Internet) terminal, the smartphone, the smart pad (Smartpad), the tablet PC etc it is the wireless communication apparatus in which the portability and mobility are guaranteed.

The community service providing server (300) can be the server providing the machine translation base global online community service web page, the app page, and the program or the application. And the community service providing server (300) can be the server mapping the case of preparing for at least one bulletin board and storing the multimedia contents in the uploader terminal (400) of at least one inside of the bulletin board at least one, and the language of the multimedia contents with the multimedia contents and stores. Moreover, in the user terminal (100), the community service providing server (300) can be the server choosing the multimedia contents and automatically translates to the language

ge which the user sets up in case the outputs and is o utputted.

여기서, 커뮤니티 서비스 제공 서버(300)는, 네트워크를 통하 여 원격지의 서버나 단말에 접속할 수 있는 컴퓨터로 구현될 수 있다. 여기서, 컴퓨터는 예를 들어, 네비게이션, 웹 브라우 저(WEB Browser)가 탑재된 노트북, 데스크톱(Desktop), 랩 톱(Laptop) 등을 포함할 수 있다.

Here, the community service providing server (300) can be materialized in computer it can connect to the serv er or the terminal of the remote location through the n etwork. Here, the computer, for example, can comprise the notebook, in which navigation, and the web browse r are mounted the desktop, the laptop etc.

적어도 하나의 업로더 단말(400)은, 자동번역 기반 글로벌 온 라인 커뮤니티 서비스 관련 웹 페이지, 앱 페이지, 프로그램 또는 애플리케이션을 이용하는 업로더의 단말일 수 있다. 이 때, 업로더 단말(400)도 사용자 단말(100)과 같이 이용하는 이용자 입장일 수도 있고 그 역도 성립하지만, 본 발명의 일 실시예에서는 설명의 편의를 위하여 역할을 나누어 설명하 기로 한다.

At least one up loader terminal (400) can be the termin al of the up loader taking advantage of the machine tra nslation base global online community service linked we b page, the app page, and the program or the applicati on. Then, it can be the user position which the up load er terminal (400) uses like the user terminal (100) and t he reverse is established. However for the sake of con venience in illustration, the role is divided and it decide s to illustrate.

여기서, 적어도 하나의 업로더 단말(400)은, 네트워크를 통하 여 원격지의 서버나 단말에 접속할 수 있는 컴퓨터로 구현될 수 있다. 여기서, 컴퓨터는 예를 들어, 네비게이션, 웹 브라우 저(WEB Browser)가 탑재된 노트북, 데스크톱(Desktop), 랩 톱(Laptop) 등을 포함할 수 있다. 이때, 적어도 하나의 업로 더 단말(400)은, 네트워크를 통해 원격지의 서버나 단말에 접 속할 수 있는 단말로 구현될 수 있다. 적어도 하나의 업로더 단말(400)은, 예를 들어, 휴대성과 이동성이 보장되는 무선 통신 장치로서, 네비게이션, PCS(Personal Communication System), GSM(Global System for Mobile communicatio ns), PDC(Personal Digital Cellular), PHS(Personal Hand yphone System), PDA(Personal Digital Assistant), IMT (International Mobile Telecommunication)-2000, CDMA (Code Division Multiple Access)-2000, W-CDMA(W-Co de Division Multiple Access), Wibro(Wireless Broadban d Internet) 단말, 스마트폰(Smartphone), 스마트 패드(S martpad), 태블릿 PC(Tablet PC) 등과 같은 모든 종류의 핸드 헬드(Handheld) 기반의 무선 통신 장치를 포함할 수 있다.

Here, at least one up loader terminal (400) can be mat erialized in computer it can connect to the server or th e terminal of the remote location through the network. Here, the computer, for example, can comprise the not ebook, in which navigation, and the web browser are m ounted the desktop, the laptop etc. Then, at least one up loader terminal (400) can be implemented as the ter minal connecting to the server or the terminal of the re mote location through the network. At least one up loa der terminal (400), for example, can comprise the wirel ess communication apparatus of the hand held (Handhe ld) base of all kinds like the navigation, the PCS (Perso nal Communication System), the GSM (Global System f or Mobile communications), the PDC (Personal Digital C ellular), the PHS (Personal Handyphone System), the P DA (Personal Digital Assistant), the IMT (International Mobile Telecommunication) -2000, the CDMA (Code Divi sion Multiple Access) -2000, the wide band-code divisi on multiple access (W-Code Division Multiple Access), t he Wibro (Wireless Broadband Internet) terminal, the s martphone, the smart pad (Smartpad), the tablet PC et c it is the wireless communication apparatus in which t he portability and mobility are guaranteed.

도 2는 도 1의 시스템에 포함된 커뮤니티 서비스 제공 서버를 설명하기 위한 블록 구성도이고, 도 3 및 도 4는 본 발명의 일 실시예에 따른 자동번역 기반 글로벌 온라인 커뮤니티 서비스 가 구현된 일 실시예를 설명하기 위한 도면이다.

Figure 2 is a block diagram for illustrating the communi ty service providing server included in the system of fig. 1. Figures 3 and 4 are the drawing illustrating the embodi ment in which the machine translation base global onli ne community service according to a preferred embodi ment of the present invention is implemented.

도 2를 참조하면, 커뮤니티 서비스 제공 서버(300)는, 업로드 부(310), 자동번역부(320), 댓글번역부(330), 인공지능부(340), 감성분석부(350), 웹툰관리부(360), 자동더빙부(370), 크라우드소싱부(380), 부분오역검수부(390) 및 수정평가부 (391)를 포함할 수 있다.

Referring to Figure 2, the community service providing server (300) comprises the upload part (310), the auto matic translation unit (320), the reply translation (33 0), the artificial intelligence part (340), the sensitivity analysis part (350), the webtoon administration unit (3 60), the automatic dubbing part (370), the crowd sour cing part (380), the part mistranslation inspection unit (390) and the correction evaluation portion (391).

본 발명의 일 실시예에 따른 커뮤니티 서비스 제공 서버(300) 나 연동되어 동작하는 다른 서버(미도시)가 적어도 하나의 사 용자 단말(100) 및 적어도 하나의 업로더 단말(400)로 자동 번역 기반 글로벌 온라인 커뮤니티 서비스 애플리케이션, 프 로그램, 앱 페이지, 웹 페이지 등을 전송하는 경우, 적어도 하 하나의 사용자 단말(100) 및 적어도 하나의 업로더 단말(400) 은, 자동번역 기반 글로벌 온라인 커뮤니티 서비스 애플리케 이션, 프로그램, 앱 페이지, 웹 페이지 등을 설치하거나 열 수

It sets up the machine translation base global online co mmunity service application, program, app page, the w eb page etc or the community service providing server (300) according to a preferred embodiment of the pres ent invention or the dissimilar server (not illustrated) w hich connected operates at least one user terminal (10 0) and at least one up loader terminal (400) can open t he machine translation base global online community se

있다. 또한, 웹 브라우저에서 실행되는 스크립트를 이용하여 서비스 프로그램이 적어도 하나의 사용자 단말(100) 및 적어도 하나의 업로더 단말(400)에서 구동될 수도 있다. 여기서, 웹 브라우저는 웹(WWW: World Wide Web) 서비스를 이용할 수 있게 하는 프로그램으로 HTML(Hyper Text Mark-up Language)로 서술된 하이퍼 텍스트를 받아서 보여주는 프로그램을 의미하며, 예를 들어 넷스케이프(Netscape), 익스플로러(Explorer), 크롬(Chrome) 등을 포함한다. 또한, 애플리케이션은 단말 상의 응용 프로그램(Application)을 의미하며, 예를 들어, 모바일 단말(스마트폰)에서 실행되는 앱(App)을 포함한다.

도 2를 설명하기 이전에 기계번역에 대한 기본개념을 이하에서 서술한다. 이하에 정리된 개념들은 도 2에서 중복하여 설명하지 않는다.

기계번역(Machine Translation)은 규칙 기반, 통계 기반, 신경망 기반의 방법으로 이어져 연구되어 왔다.

003c#규칙기반 기계번역003e#

규칙기반 기계번역(Rule-based Machine Translation, RBMT)은 언어를 단어와 문법 규칙의 집합으로 간주하고 언어 규칙에 따라 번역하고자 한 연구이다. 언어학적 관점에서 보면, 기계번역은 입력문장을 형태소적, 구문적, 의미적으로 분석(Analysis)하고, 분석 결과의 표현 형태를 출력문장 생성을 위한 표현형태로 변환(Transfer)한 후, 출력문장의 표현형태로부터 최종문장을 생성(Generation)하는 3단계를 거친다. RBMT에서 한국어와 일본어처럼 언어적 유사성이 큰 언어쌍은 형태소 분석 후 변환을 수행하는 직접 번역방식으로 개발되었다. 중간언어(Interlingua) 방식은 다국어번역에 적합한 방법론으로, 문장의 의미를 언어 독립적으로 나타낼 수 있는 표현 형태를 정의한 후 분석과 생성만으로 번역을 수행하고자 한 방식이다. 대부분의 언어쌍은 구조분석에 기반한 간접 변환 방식으로 개발되었는데, 이를 위해서는 언어의 형태소 및 구문 현상을 문법적으로 나타낼 수 있어야 하므로, GPSG(Generalized Phrase Structure Grammar), HPSG(Head-driven Phrase Structure Grammar), LFG(Lexical Functional Grammar, LFG), DG(Dependency Grammar)등의 다양한 문법에 기반한 방법들이 연구되었다. 다만 모든 언어 현상을 규칙화하는 데 한계를 느끼고, 말뭉치에서 추출한 확률 정보에 기반하여 번역하고자 하는 시도가 생겨났다.

003c#통계기반 기계번역003e#

통계기반 기계번역(Statistical Machine Translation, SMT)는 원시언어 문장  $f^1 = f_1, \dots, f_j$ 과 그에 대응되는 목적언어

service application, program, app page, the web page etc with at least one user terminal (100) and at least one up loader terminal (400) in case the transmission. Moreover, the service program can be driven using the script performed in the web browser in at least one user terminal (100) and at least one up loader terminal (400). Here, the hypertext which is describe according to the HTML (Hyper Text Mark-up Language) is received as the program in which the web browser uses the web (WWW: World Wide Web) service and the Netscape the program showing is meant, explorer, and the chrome etc are included. Moreover, the application means the application program on the terminal and the app performed in for example, the mobile terminal (smartphone) is included.

Hereinafter the basic concepts about the machine translation is described before illustrating fig. 2. Concepts put in order in less than overlap in 2 and it does not illustrate.

The machine translation was connected to the rule base, the statistics based, and the method of the neural network base and the method had been being studied.

003c# rule base machine translation 003e#.

It is the research that tries to regard as the gathering of the grammatical rule and word and that the rule based machine translation (Rule-based Machine Translation, RBMT) tries to translate language according to the language rule. If it looks around the linguistics viewpoint, the machine translation goes the morpheme, the syntactic, and 3 step which semantically semantically lies the final sentence to the representation for the output sentence production after doing the conversion (Transfer) from the representation of the output sentence the production (Generation) through the input sentence. In RBMT, the language pair in which the grading for Glasgow coma scale similarity was large was developed as the direct translation mode performing the conversion like the Korean and Japanese after the morphological analysis. The intermediate language (Interlingua) mode is the mode which tries to perform the translation to the analysis and production after it defines the representation which can show the meaning of the sentence as the language independence to the methodology which is suitable for the multilingual translation. It was developed with the indirect conversion method furnace. Most of language pairs were based on the structural analysis it was for this and the morpheme and syntax phenomenon of language could be grammatically shown. Therefore based methods were studied in the various grammar of the GPSG (Generalized Phrase Structure Grammar), the HPSG (Head-driven Phrase Structure Grammar), the LFG (Lexical Functional Grammar, LFG), the DG (Dependency Grammar) etc. But all language phenomena were regulated but the limit was realized and the attempt to translate based on the probability information extracted from the language corpus appeared.

003c# statistics based machine translation 003e#.

When the object language sentence  $e^1 = e_1$  in which the statistics based machine translation (Statistical Mac

문장  $e^1 = e_1, \dots, e_l$ 이 주어졌을 때, 노이즈 채널 모델(Noisy Channel Model)에 따라 확률  $P(e^1 \# 124 \# 1 \# 124 \# | \# 124 \# f \# 124 \# ^j 1)$ 을 최대화시키는 문장 (tilde)e를 최종 번역 문장으로 결정한다.

$P(e^1 \# 124 \# 1 \# 124 \# | \# 124 \# f \# 124 \# ^j 1)$ 는 베이저안 규칙에 의해  $P(f^j \# 124 \# 1 \# 124 \# | \# 124 \# e \# 124 \# ^l 1)$ 과  $P(e^1 1)$ 로 나뉠 수 있다. 전자를 번역모델(Translation Model), 후자를 언어모델(Language Model)이라 칭하는데 SMT를 구성하는 핵심 요소들이다. 단어 번역시에 주변 단어들을 문맥 정보로 반영하기 위해, 단어가 아닌 구(Phrase)를 번역의 기본 단위로 하여  $P(f^j \# 124 \# 1 \# 124 \# | \# 124 \# e \# 124 \# ^l 1)$ 를 모델링한 SMT를 구단위 SMT(Phrase-based SMT, PBSMT)라고 한다. 여기서의 구는 RBMT에서의 언어학적 정의가 아니라 표층적으로 단순 연속된 단어열을 뜻한다. 따라서 PBSMT는 주어진 원시언어 문장을 연속된 구들의 조각으로 나눈 후, 각각을 구단위 번역패턴을 이용하여 번역하고 이를 최종 결합하는 과정으로 이루어진다.

구단위 번역 패턴은 원시언어 문장과 목적언어 문장이 1:1로 매핑된 양국어 코퍼스로부터 추출되며, 원시 언어 구에 대응하는 목적언어 구가 확률을 기반으로 계산된 점수와 함께 저장된다. 일반적으로 구단위 번역 패턴은 GIZA++과 같은 단어 정렬 툴킷을 이용하여 코퍼스의 원시언어 문장과 목적언어 문장의 단어들을 정렬한 후, 정렬 빈도정보에 기반하여 추출된다. 언어모델은 번역된 문장이 얼마나 자연스러운가를 판단하는 역할을 하며 목적언어 코퍼스에 기반하여 구축된다.

PBSMT는 가장 대중적인 SMT 모델이지만 연속 단어열을 번역의 단위로 사용하므로 한국어와 영어처럼 어순이 크게 다른 언어간의 번역에 어려움이 있었다. 이를 해결하기 위해 하위 구(Sub-Phrase)를 가질 수 있는 구(Hierarchical Phrase)를 번역의 기본 단위로 하는 HSMT(Hierarchical Phrase-Based Statistical Machine Translation)이 연구되었고 SOTA(State-Of-The-Art) PBSMT의 성능을 뛰어넘었다.

#### 003c#신경망기반 기계번역003e#

초기의 신경망기반 기계번역(Neural Machine Translation, NMT) 연구들은 SMT의 기본 구조를 그대로 유지하면서 언어 모델 등의 하위모듈을 신경망으로 대체하는 형태로 진행되었는데, 이러한 접근법은 SMT의 근본적 문제인 하위모듈 기반 튜닝의 한계를 극복할 수 없었다. 이때, NMT를 하나의 신경망으로 모델링하여 학습함으로써 SOTA SMT의 성능을 뛰어넘는 NMT 시대를 열게 되었다. 이에, 본 발명의 일 실시예에 따른 자동번역은 NMT, 즉 신경망 모델을 기반으로 하되 신경망 모델 중에서도 자기집중(Self-Attention) 기반 모델인 Transformer를 이용하도록 한다. 물론, 상술한 모델 이외에도 다양한 신경망 모델이 이용될 수 있음은 자명하다 할 것이다.

hine Translation, SMT) was corresponded to that to the primitive language sentence  $f^j 1 = f_1$ , and the  $f_j$ , and the  $e_l$  were given, the sentence (tilde) e maximizing the probability  $P(e^{(sup)} | (/sup) | 1 | | f | (/sup) | 1 | | | f | (/sup) | 1)$  is determined with the final translated sentence according to the noisy channel model.

$P(e^{(sup)} | (/sup) | 1 | | f | (/sup) | 1 | | | f | (/sup) | 1)$  can be classified into  $P(f^{(sup)} | (/sup) | 1 | | e | (/sup) | (/sup) | 1)$  and  $P(e^{(sup)} | (/sup) | 1)$  with the Bayesian rule. It is core elements which call the electron the translation model, and the latter because of being the language model but configure the SMT. In order that around words are reflected in the word translation to the contextual information SMT deciding on the globe (Phrase) which is not word to the base unit of the translation and models  $P(f^{(sup)} | (/sup) | 1 | | e | (/sup) | (/sup) | 1)$  is done as the unit of word SMT (Phrase-based SMT, PBSMT). It is not linguistics definition at RBMT but the globe of the women rare book means the word string continued to the superficial layer with basic. Therefore, PBSMT comprises the process after it divides the given primitive language sentence into the piece of the consecutive floor heater of translating each using the unit of word translation pattern and combining this with final.

The unit of word translation pattern is extracted from the primitive language sentence and two countries corpus in which the object language sentence is mapped to 1:1 and it is stored with the score in which the object language outlet corresponding to the source language outlet is calculated based on the probability. Generally, the unit of word translation pattern is the primitive language sentence of the corpus and words of the object language sentence extracted using the word arrangement toolkit like the GIZA++ after doing the queue based on the queue order of frequency. While to determining whether the sentence in which the language model is translated is natural or not it is constructed based on the object language corpus.

Since PBSMT was the SMT model which was most popular but the consecutive word heat was used by unit of the translational the difficulty was in the translation between the language in which the word order was different like the Korean and English. To solve this, the HSMT (Hierarchical Phrase-Based Statistical Machine Translation) deciding on the globe (Hierarchical Phrase) having the low rank globe (Sub-Phrase) to the base unit of the translation was studied and the performance of the SOTA (State-Of-The-Art) PBSMT was jumped over.

#### 003c# neural network base machine translation 003e#.

Archaic neural network base machine translation (Neural Machine Translation, NMT) researches were progressed as the form which replaced the low module including the language model etc. with the neural network while it like that maintained the basic structure of SMT. The limit of the low module base tuning in which such approach was the basic problem of SMT could not be overcome. Then, it was opening the NMT era which jumped over the performance of the SOTA SMT by modeling NMT by one neural network and learning. Thus, the machine translation according to a preferred embodiment of the present invention does based on the NMT, in other words, the neural network model and the Transformer

which is the self-convergence (Self-Attention) base model is used among the neural network model. Of course, the neural network model which is various besides the above-described model can be used it is obvious it will do.

도 2를 참조하면, 업로드부(310)는, 적어도 하나의 주제를 가지는 게시판 내 적어도 하나의 종류의 멀티미디어 콘텐츠에 대응하는 게시글이 등록되면 게시글을 업로드할 수 있다. 적어도 하나의 종류의 멀티미디어 콘텐츠는, 이미지, 동영상, 음성 및 텍스트 중 어느 하나 또는 적어도 하나의 조합으로 생성된 콘텐츠일 수 있다. 적어도 하나의 종류의 멀티미디어 콘텐츠는, 적어도 하나의 종류의 소셜미디어, HTML 웹게임, 유튜브 라이브 및 웹툰을 포함할 수 있다.

Referring to Figure 2, as to the upload part (310), if the posted writing corresponding to the multimedia contents of the kind of the inside of the bulletin board at least one having at least one theme is registered, it can be the posted writing. It can be the contents generated to any one of multimedia contents, is the image, the moving picture, and voice and text or at least one combination of at least one kind. The multimedia contents of at least one kind comprises the social media of at least one kind, the HTML web game, and the U-tube live and the webtoon.

자동번역부(320)는, 사용자 단말(100)에서 게시글을 감지된 언어로 출력하도록 자동번역을 수행할 수 있다. 사용자 단말(100)은, 언어설정을 감지하여 적어도 하나의 종류의 멀티미디어 콘텐츠 내 포함된 다국어 텍스트를 감지된 언어로 자동번역하여 출력할 수 있다.

In the user terminal (100), the automatic translation unit (320) can perform the machine translation in order to output the posted writing to the sensed language. The user terminal (100) senses the language institution and it automatically translates the multi-lingual text included with the multimedia contents of at least one kind to the sensed language and it can output.

댓글번역부(330)는, 게시글에 댓글이 포함된 경우 댓글을 감지된 언어로 자동번역하여 출력할 수 있다.

In case the reply is included in the posted writing it automatically translates the reply to the sensed language and it can output.

인공지능부(340)는, 신경망 기반 기계번역 모델 중 자기집중(Self-Attention) 기반 Transformer 모델을 이용하여 자동번역을 실시할 수 있다. 자기집중 기반 NMT인 Transformer는 일반적인 RNN, CNN을 적용하지 않고, 주의집중 기법에 전적으로 의존한다. 이를 위해 적용되던 인코더-디코더 주의집중과 함께 인코더 및 디코더에 자기집중 기법(Self-Attention)을 적용하는데, 인코더-디코더 주의집중과 자기집중은 이하의 수학적 식 2와 같이 정의된다.

The artificial intelligence part (340) can perform the machine translation among the neural network base machine translation model using the self-convergence (Self-Attention) base Transformer model. The Transformer which is the base NMT among the magnetic house does not apply the general RNN, and CNN and it entirely depends on the cathexis technique. For this, the technique (Self-Attention) is applied to the encoder and decoder with the applied encoder - decoder cathexis among the magnetic house. The encoder - decoder cathexis and self-convergence are defined like the equation 2 of less than.

여기서 Q는 질의(Query), K와 V는 쌍으로 존재하는 키(Key)와 값(Value)이며,  $d_k$ 는 K의 차원을 나타낸다. 인코더-디코더 주의집중은 입력단어와 출력단어 간의 연관관계를 고려하며, 인코더 자기집중은 입력 단어들끼리, 디코더 자기집중은 출력단어들끼리의 연관 관계를 고려한다.

Here, the key (Key) and the value, the  $d_k$  in which the query, and K and V Q exists as pair show the dimension of K. The encoder - decoder cathexis considers the associative relationship between the output word and the input word and the encoder self-convergence the decoder self-convergence considers the associative relationship of the output words with input words.

Transformer는 또한 멀티헤드 주의집중(Multi-Head Attention)의 개념을 추가로 도입하고, 이를 통해 질의 Q, 키 K, 값 V를 대상으로 h개의 서로 다른 선형투영(Linear Projection)을 수행함으로써, h개의 다른 부분공간 상에서의 주의집중 점수를 계산할 수 있게 하였다. 멀티 헤드 주의집중은 수학적 식 3과 같다.

The Transformer additionally introduced moreover, the concept of the plural head cathexis (Multi-Head Attention) and through this, the different linear projection of the h was performed for Q of quality, the key K, and the value V. In that way the cathexis score on the dissimilar subspace of the h was calculated. The plural head cathexis is the same as that of the equation 3.

일반적인 절대위치 기반 위치 인코딩은 강건성 등에서 문제가 있기 때문에, Transformer에서는 사인함수와 코사인함수를 이용한 정현파 위치 인코딩(Sinusoidal Positional Encoding)을 수행할 수 있다. 품질과 속도가 함께 높아질 수 있고, 멀티헤드 자기집중(Multi-Head Self-Attention) 기법의 뛰어난 문장분석 능력으로 인해, Transformer의 인코더와 디코더는 각각 BERT(Bidirectional Encoder Representations from Transformers)와 GPT(Generative Pre-trained Transf

In the general absolute position base position encoding is the robustness etc, it has the problem. Therefore the sine function and the sinusoidal wave position encoding (Sinusoidal Positional Encoding) using the cosine function can be performed in the Transformer. Quality and speed can become together higher and the encoder and decoder of the Transformer can be used as the basic structure of the GPT (Generative Pre-trained Transform

ormer)의 기본구조로도 이용될 수 있다.

감성분석부(350)는, 자동번역을 실시하기 이전에 다국어 텍스트를 언어별로 분리한 다음 감성분석을 진행하고, 극성치를 종합하는 다중언어 감성분석 프로세스를 진행할 수 있다. 이를 위하여, 입력받은 언어 별로 문장을 분리하는(Segmentation) 과정을 거쳐 일본어, 중국어, 한국어 및 영어 단어 즉, 단어를 단위로 나누어 저장한다. 예를 들어, 한국어 형태소 분석기 Rhino 1.0을 사용할 수 있고, 중국어 Segmentation Tool로는 Jieba를 사용할 수 있으며, 중국어 형태소 분석기로는 Stanford POS Tagger 중 Chinese Stanford Tagger를 사용할 수 있다.

다음으로, 4 개국 단어를 형태소 분석기에 돌려 입력 받은 4 개국 문장에 포함된 명사, 동사, 형용사, 부사의 개수를 센다. 또한 단어 단위로 저장한 4 개국어의 4 개국 감성사전을 통해 4 개국어의 감성점수를 구한다. 그 외에 댓글이나 코멘트에 많이 등장하는 이모티콘도 감성을 가지고 있기 때문에 다중언어 감성분석에 추가할 수 있다. 이모티콘 감성사전은 Emoji Sentiment Ranking 1.0을 사용할 수 있다. 마지막으로 4 개국어 및 이모티콘의 감성점수 그리고 명사, 동사, 형용사, 부사, 어근 및 감성 이모티콘의 개수를 통해 전체 다중언어의 감성점수를 구하게 된다.

003c#다중언어 감성분석 프로세스003e#

다중언어 감성분석을 위한 기법은 크게 3 단계로 구성될 수 있다. 우선 감성분석을 하고자 하는 문장을 입력받으면 그 문장을 단어 단위로 쪼갠 후 형태소 분석기를 통해 각 단어의 형태소를 판단하고, 형태소 중 명사, 동사, 형용사, 부사의 개수를 센다. 특히 한글의 경우 단어구성이 중국어를 포함한 4개국어와 다르기 때문에 한글은 명사, 동사, 형용사, 부사 외에 어근도 추가하여 센다. 또한 이모티콘의 감성값이 긍정 혹은 부정이면 이모티콘의 개수를 센다. 이때, 한국어와 중국어의 경우를 예로 들면, 한국어는  $\sum W_k$ , 중국어는  $\sum W_c$ , 이모티콘은  $\sum E_c$ 라고 표현한다.

두 번째 단계에서는 감성사전을 통한 감성분석을 진행한다. 만약 분리된 단어가 감성사전에 속해 있는 단어이면, 감성사전 중의 감성점수로 그 단어의 감성값을 반영한다. 이를 수학적 4 및 수학적 5로 표현할 수 있다.  $S_j$ 는 특정 단어의 감성점수,  $W_j$ 는 다중 언어로 구성된 구절을 Segmentation과정을 거쳐서 얻은 특정단어이며  $SDW$ 는 감성사전에 포함되어 있는 감성단어이다. 특히 수학적 5는 특정단어가 감성사전 속 단어에 포함되어 있지 않는 조건을 만족하는 경우로 이럴 때에는 특정단어의 감성점수  $S_j$ 값은 0이 된다.

마지막으로 언어별로 얻은 총 감성점수를 명사, 동사, 형용사,

er) and BERT (Bidirectional Encoder Representations from Transformers) due to the excel sentence analysis capability of the plural head self-convergence (Multi-Head Self-Attention) technique.

After the sensitivity analysis part (350) separates the multi-lingual text according to language before performing the machine translation it progresses the sensitivity analysis and it can progress the multi-lingual sensitivity analysis process synthesizing the polarity value. For this, japanese, Chinese, and the Korean and English word, in other words, word are divided after the process of especially separating the sentence with language (Segmentation) it receives into the unit and it stores. For example, the Korean morpheme analyzer Rhino 1.0 can be used and the Jieba can be used as the Chinese Segmentation Tool and the Chinese Stanford Tagger can be used as the Chinese lexical analyzer among the Stanford POS Tagger.

Next, the number of noun, verb, adjective, float included in 4 country sentence which turns 4 country word in the lexical analyzer and received are counted. Moreover, 4 foundation of the state the sensibility score of 4 foundation of the state is found for the sake through 4 country sensibility prior it stores by unit of word. Besides, because the emoticon very much shown up in the reply or the comment has the sensibility it can add to the multi-lingual sensitivity analysis. The emoticon sensibility prior can use the Emoji Sentiment Ranking 1.0. Finally, the sensibility score of the whole multi-lingual is saved through the number of sensibility score of 4 foundation of the state and emoticon and noun, verb, adjective, float, primitive and sensibility emoticons.

003c# multi-lingual sensitivity analysis process 003e#.

A lot the technique for the multi-lingual sensitivity analysis can consist of 3 step. First, the number of noun, verb, adjective, float among the morpheme the sentence to the sensitivity analysis is receiveds are counted. Especially, in case of the Hangeul, because it is different from 4 country the word structure includes Chinese the primitive adds besides the noun, verb, adjective, float and the Hangeul is strong. Moreover, the number of emoticon the sensibility value of the emoticon is the affirmation or the negations are counted. Then, for example, if the case of Chinese and Korean is given , the Korean  $\sum W_k$  , and Chinese  $\sum W_c$  , and the emoticon express a  $s \sum E_c$ .

In the second step, the sensitivity analysis through the sensibility prior is progressed. If the if is the word in which the separated word belongs to the sensibility prior , the sensibility value of the word is reflected to the sensibility score among the sensibility prior. This can be expressed in the Equation 4 and equation 5. The  $S_j$  is the sensibility score of the specific word, and the  $W_j$  are the specific word which obtains the paragraph consisting of the multi-lingual after the Segmentation process the emotional word in which  $SDW$  is included in the sensibility prior. Especially, when the specific word is satisfied the condition that is not included with the sensibility prior inside word the equation 5 is like this, the sensibility score  $S_j$  value of the specific word 0 becomes.

Finally, the result which comes out after it divides the

부사, 어근(한국어의 경우에만 해당) 및 긍정 및 부정 감성값이 존재하는 이모티콘 개수(이모티콘의 경우에만 해당)의 합으로 나눈 후 나온 결과와 각 언어에서 할당된 가중치를 반영하여 최종 얻고자 하는 다중언어 감성점수를 구하게 된다. 이것을 수학적 식 6을 통해 계산할 수 있다.

$W_c$ 는 중국어가 전체 다중언어에 차지하는 가중치,  $W_k$ 는 한국어가 차지하는 가중치이며  $W_e$ 는 이모티콘이 차지하는 가중치이다.  $S_c$ 는 감성분석을 하고자 하는 모든 중국어의 감성점수  $S_j$ 의 합이고,  $S_k$ 는 중국어와 마찬가지로 모든 한국어 감성점수  $S_j$ 의 합이며,  $S_e$ 는 모든 이모티콘의 감성점수  $S_j$ 의 합이다.  $S$ 는 다중언어의 감성점수이다.

웹툰관리부(360)는, 사용자 단말(100)에서 출력하는 멀티미디어 콘텐츠가 웹툰인 경우, 웹툰 내 텍스트를 OCR로 추출할 때 추출된 영역의 좌표 및 크기를 저장하고, 텍스트를 자동번역한 후 번역된 글자 수를 카운트하여 영역의 크기 내에 포함 되도록 글씨 크기, 폭 및 간격을 조절하여 추출된 영역으로 오버레이할 수 있다.

자동더빙부(370)는, 사용자 단말(100)에서 출력하는 멀티미디어 콘텐츠가 오디오 콘텐츠인 경우, 음성발화를 텍스트로 변환하는 STT(Speech to Text)를 실시한 후 자동번역을 실시한 후, 음성발화한 발화자의 음성에 자동번역된 텍스트를 입혀 출력시킬 수 있다. 인공지능 스피커나 챗봇을 구동할 때, STT-NLP-TTS의 과정을 거치게 되는데 이미 녹음된 성우의 목소리로 텍스트를 읽어내려가는 방법을 적용하게 된다. 이를 그대로 동영상에 적용하는 경우, 예를 들어, 일본 유명 크리에이터나 연예인의 목소리로 목소리를 학습하고, STT-NMT-TTS의 과정을 거치게 되면 실제 그 연예인이 한국어로 NMT로 자동번역된 문장을 읽어내려가는 듯한 효과를 주기 때문에 실시간 더빙을 한 것과 같은 느낌을 주면서도 그 연예인이나 크리에이터의 목소리로 듣게 되므로 거부감도 없을 수 있다.

클라우드소싱부(380)는, 적어도 하나의 종류의 멀티미디어 콘텐츠 내 자동번역이 된 결과물에 대하여 복수의 사용자 단말(100)로부터 평가를 수집한 후, 평점이 기 설정된 점수 이하인 경우 적어도 하나의 관리자 단말로 자동번역이 된 결과물을 추출하여 검수를 요청할 수 있다. 이때, 사용자 단말(100)은 서로 다른 주체에 귀속된 복수의 사용자 단말(100)일 수 있다.

부분오역검수부(390)는, 사용자 단말(100)에서 자동번역된 텍스트가 출력될 때, 사용자 단말(100)에서 설정한 언어에 대응하는 텍스트를 선택하는 경우, 원본 텍스트의 언어로 표시

total sensibility score which obtains from language by the sum total of the emoticon number (the corresponding in case of the emoticon) in which the noun, verb, adjective, float, the primitive (the corresponding in case of the Korean), the affirmation and negative emotion value exist and the weighted value allocated in each language are reflected and the multi-lingual sensibility score to get with final is saved. This can be calculated through the equation 6.

The  $W_c$  is the weighted value which Chinese occupies in the whole multi-lingual, and the  $W_k$  are the weighted value which the Korean occupies the weighted value which as to  $W_e$ , the emoticon occupies. It is the sum total of the sensibility score  $S_j$  of all Chinesees in which the  $S_c$  the sensitivity analysis and the  $S_k$  is the sum total of all Korean sensibility score  $S_j$ s like Chinese and the  $S_e$  is the sum total of the sensibility score  $S_j$  of all emoticons.  $S$  is the sensibility score of the multi-lingual.

In case the multimedia contents which the webtoon administration unit (360) outputs in the user terminal (100) is the webtoon it stores the coordinate and size of the extracted domain when extracting inside of the webtoon the text in OCR and it can overlay to the domain which adjusts the font size, and the width and gap so that it counts the number of characters which is the text translated after doing the machine translation and the is extracted domain is included within the size of the domain

In case the multimedia contents which the automatic dubbing part (370) outputs in the user terminal (100) is the audio content after it performs the STT (Speech to Text) converting the voice firing into the text it performs the machine translation it plates the text which becomes with machine translation on the voice of the speaker done with the voice firing and it can print out. When the artificial intelligence speaker or the chatbot is operated, the method for passing through the process of the STT-NLP-TTS but reading the text from the voice of the already recorded actor specializing in dubbing films is applied. Voice is learned to the case, of like that applying this to the moving picture for example, the voice of the Japanese famous creator or the entertainer and if the process of the STT-NMT-TTS is passed through, since the real-time dubbing is heard in the voice of the entertainer or the creator even when giving the feeling like one the resistance therefore can do the effect that the entertainer actually reads the sentence consisting of NMT with machine translation from the Korean not have with cycle.

In case the grade is the predetermined score or less it extracts the outcome in which the machine translation is to at least one manager terminal evaluation about the outcome in which the crowd sourcing part (380) becomes the multimedia contents my machine translation of at least one kind from multiple user terminals (100) after doing the collection and it can request the inspection. Then, may be the multiple user terminals (100) in which the user terminal (100) is returned to the different subject.

In the user terminal (100), when the text which becomes with machine translation is outputted, the text is indicated as the case of choosing the text corresponding



되도록 하고, 사용자 단말(100)에서 자동번역된 텍스트의 수정요청을 영역을 지정한 후 전송하도록 한다. 도 4i와 같이 원본 텍스트의 언어는 영어로 #34#Look at how tough he#39#s trying to be! Instead of being scary he looks so cute! This was taken by Will Burrard-Lucas. You can check more of his wildlife photos on his Instagram page. #34#이고, 자동번역된 한국어는 #34#그가 얼마나 힘든지 보십시오! 무섭다기 보다는 너무 귀엽게 생겼어요! 이것은 Will Burrard-Lucas가 찍은 것입니다. 그의 야생 동물 사진을 더 많이 확인할 수 있습니다. 인스타그램 페이지#34#라고 번역되었다. 대부분의 기계번역은 어느 정도는 문맥이 이상하지만 뜻은 통하는 경우가 존재하는데, 문장이 전부 오역되는 경우는 없고 어느 한 단어가 어절이 이상하여 문장 자체가 이상해보이는 경우가 대부분이다.

이때, #34#tough#34#는 #34#힘든#34#이라는 뜻도 있지만 이는 거의 사물이나 일, 또는 대상화된 객체에 쓰일 뿐 인간이나 동물과 같이 생명이 있는 객체에게는 #34#터프한#34#, #34#용감한#34# 등으로 번역되어야 하지만, 자동번역이기 때문에 문맥을 간과하고 #34#힘든#34#으로 번역되었고 이 부분 때문에 #34#그가 얼마나 힘든지 보십시오#34#와 같이 귀여운 아기 사자사진에 맞지 않는 문장으로 번역되었고, 결과적으로 이상한 문장이 출력되었다. 여기서, 사용자 단말(100)에서 #34#힘든지#34#에 대응하는 원문 텍스트가 무엇인지 영역을 설정하고 길게 탭하는 경우, 원문 텍스트가 #34#tough#34#로 도시되게 되고, 원문 텍스트를 보고 사용자 단말(100)에서 #34#터프한#34#으로 바꾸거나 바꾸도록 요청하는 경우 자동번역에서 #34#힘든#34#이 #34#터프한#34#으로 바로 반영되거나 또는 관리자의 검수를 받는 과정 후에 반영될 수 있다. 또는 두 가지 버전을 모두 제시한 뒤 한국어를 사용하는 사용자에게 고르도록 할 수도 있는데 이는 수정평가부(391)에서 이루어진다.

수정평가부(391)는, 사용자 단말(100)에서 자동번역된 텍스트의 수정을 한 경우, 수정 후 버전과 수정 전 버전을 동시에 게재하여 번역평가를 로그로 누적할 수 있다.

이하, 상술한 도 2의 커뮤니티 서비스 제공 서버의 구성에 따른 동작 과정들 도 3 및 도 4를 예로 들어 상세히 설명하기로 한다. 다만, 실시예는 본 발명의 다양한 실시예 중 어느 하나일 뿐, 이에 한정되지 않음은 자명하다 할 것이다.

도 3을 참조하면, (a) 커뮤니티 서비스 제공 서버(300)는 적어도 하나의 업로더 단말(4000)로부터 적어도 하나의 종류의 언어로 업로드되는 멀티미디어 콘텐츠를 업로드받고 이를 각 게시판에 업로드한다. 그리고, 커뮤니티 서비스 제공 서버(300)는 사용자 단말(100)에서 언어를 설정하거나 단말 내 설정

to the language which the part mistranslation inspection unit (390) sets up in the user terminal (100), and the language of the original text and it transmits the domain in the modify request of the text which becomes in the user terminal (100) with machine translation after doing the designation. The language of the original text is the "Look at how tough he's trying to be! Instead of being scary he looks so cute! This was taken by Will Burrard-Lucas. You can check more of his wildlife photos on his Instagram page." like the drawing 4i to English and it was so adorably generated rather than it looked at whether as to the Korean, which became with machine translation the " that was hard it was terrible this is. The Will Burrard-Lucas films. Its wildlife photograph can be more very much confirmed. It was translated as the Instagram page ". The context as to most of machine translations, about a is odd but the case where the meaning is put through exists. The case where the sentence is all mistranslated does not have and the separate word of one word is odd and the case where the sentence itself seems to be strange is the majority.

Then, the meaning called this had the "tough" with 34 but it was hard with " this was nearly used in the object, the task, or the object which became an object but it had to be translated into the etc. with 34. However it was brave with " it was tough like human or animal to the object in which human had life with " because it was the machine translation the context was overlooked and it was translated with 34 and it was hard with " it looked whether the " that was therefore hard with teeth it was translated into the sentence which did not fit for the baby lion photograph which was cute with 34 and the consequently strange sentence was outputted. Here, in the user terminal (100), the case it fixes the domain whether it is hard with " or not whether the original text corresponding with 34 is the what and or not of lengthways tapping, and the original text are illustrated to the "tough" and it is immediately reflected with 34 or it changes with 34 or it reads the original text and it is tough in the user terminal (100) with " it is tough in case the request with 34 it is hard in the machine translation with " in order to change, it can be reflected after the process of taking the inspection of the administrator. Or it is regular to the user who uses the Korean after the user presents two versions but this is comprised of the correction evaluation portion (391).

In the user terminal (100), the correction evaluation portion (391) simultaneously reports the version and patient to call version after the case of modifying of the text which becomes with machine translation, and the correction and it can accumulate the translation evaluation under the log.

Hereinafter, for example, the operation process according to the configuration of the above-described community service providing server of fig. 2 figures 3 and 4 decides to be particularly illustrated. But the embodiment is any one among the various embodiment of the invention. It is not thus restricted it is obvious it will do.

Referring to Figure 3, the multimedia contents in which (a) community service providing server (300) is uploaded from at least one uploader terminal (4000) to the language of at least one kind is uploaded and this is recorded in each bulletin board. And in the community service

된 언어를 제공하면 이에 맞게 자동으로 다국어 설정 언어로 자동번역한 후 제공할 수 있다. 그리고, (c) 커뮤니티 서비스 제공 서버(300)는 음성 또는 동영상과 같이 화자가 음성발화를 하는 콘텐츠가 존재하는 경우, 화자의 음성 데이터에 자동번역된 텍스트를 입혀 실시간 더빙이 되는 것과 같은 효과를 제공할 수도 있다. 물론, 초기에는 충분한 학습 데이터셋이 없어서 그 사람의 목소리나 톤, 음정과 같이 정확한 특색을 따라하기는 어렵겠지만, 다수의 동영상을 확보할 수 있는 크리에이터를 시작으로 충분한 학습 데이터셋을 확보한 후 보이스 톤 등을 학습시키면 실시간 더빙이 되는 효과를 얻을 수 있고 흥미요소도 줄 수 있다. (d) 그리고, 웹툰의 경우에도 이미지 형태로 제공되는 경우가 많은데 이를 OCR로 텍스트가 있었던 영역의 크기, 위치좌표 등을 모두 저장해둔 후, 자동번역된 텍스트의 글자 수를 카운트한 후 크기에 맞게 들어갈 수 있도록 글자크기나 간격 등을 조정하여 덧씌우는 작업을 수행할 수 있다.

본 발명의 일 실시예는 각 지역별 또는 국가별 커뮤니티는 존재하지만 실제 텍스트 기반으로 전 세계가 이야기를 나눌 수 있는 커뮤니티는 없다는 점에서 출발한다. 이에 각국의 언어를 사용자의 언어로 변환하여 각국 어느 나라에서도 하나의 플랫폼 내에서 소통을 할 수 있도록 하는 것을 목적으로 한다. 이때 사용자 맞춤 자동번역을 제공하고, 게시판에 게시글을 업로드하고 사용자 맞춤 콘텐츠를 이용할 수 있도록 한다. 또 HTML5 웹 게임이나, 유튜브 라이브 또, 4 개국어 자동 번역 웹툰을 제공함으로써 다른 플랫폼과의 차별화를 두고, 웹툰의 경우도 4a와 같이 번역 및 자동삽입될 수 있으며, 도 4b와 같이 SEO 최적화 개발과 각 게시판 별 전문가를 섭외함으로써 사용자를 플랫폼으로 유도하는 홍보효과를 누릴 수 있다. 이와 같은 본 발명의 플랫폼(아몽라이브, 가칭)은 도 4d와 같이 구현될 수 있으나 이에 한정되는 것은 아니고 실시예에 따라 언제든지 변경 및 변형될 수 있음은 자명하다 할 것이다.

이와 같은 도 2 내지 도 4의 자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 방법에 대해서 설명되지 아니한 사항은 앞서 도 1을 통해 자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 방법에 대하여 설명된 내용과 동일하거나 설명된 내용으로부터 용이하게 유추 가능하므로 이하 설명을 생략하도록 한다.

도 5는 본 발명의 일 실시예에 따른 도 1의 자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 시스템에 포함된 각 구성들 상호 간에 데이터가 송수신되는 과정을 나타낸 도면이다. 이하, 도 5를 통해 각 구성들 상호간에 데이터가 송수신되는 과정의 일 예를 설명할 것이나, 이와 같은 실시예로 본원이 한정 해석되는 것은 아니며, 앞서 설명한 다양한 실시예들에 따

providing server (300) is the user terminal (100), if language is set up or the language established with inside of the terminal is provided, accordingly the multi-lingual can be automatically provided to the setting language after doing the machine translation. And the text which becomes with machine translation is plated on audio data of the speaker and the contents in which the speaker (c) community service providing server (300) does the voice firing like voice or the moving picture can provide the effect that the real-time seems to be dubbed concerning in case the presence. Of course, in the initial, it be difficult to there will be no enough learning data set and will repeat the exact feature like voice or the tone of the human, and the musical interval. However if tone etc. are cultivated after enough learning data set is secured beginning with the creator securing multiple moving pictures, the effect that the real-time is dubbed can be obtained and the interesting element can give. (d) And in case of the webtoon, after there is much case of being provided to the type of image but this the size of the domain, in which the domain, had the text the constellation etc. is altogether stored to OCR the operation which controls the character size or the gap etc. so that the puts on food operation enters under size after it counts the number of characters of the text which becomes with machine translation and can be performed.

Each local or each country community a preferred embodiment of the present invention exists but the world wide actually originates with the text based in that the community having the story does not have. Thus, the language of every country is converted into the language of the user and every country makes in every country a country in one platform to the purpose having mutual understanding. Then, the customization machine translation is provided and the posted writing is embedded in the bulletin board and the customization contents is used. And and the differencing with the dissimilar platform is put with the HTML5 web game, or the U-tube live by providing with 4 foundation of the state machine translation webtoon and it can be automatically inserted in case of the webtoon like the drawing 4a with the translation and the effect of public relation which induces the user like the drawing 4b to the platform by casting the SEO optimization development and each bulletin board much expert can be enjoyed. Platform (\*\*\* live, and the tentative name) of the invention can be implemented like 4d but it can not be thus restricted and it be any time transformed into according to the embodiment with the change the change is obvious it will do.

Since the easily is assumable from the content that is identical with the content explained through 1 about the machine translation base global on-line community service providing method or is explained the matter which is not explained about the machine translation base global on-line community service providing method of this kind of the figures 2 through 4 the before hereinafter omits the description.

Figure 5 is a drawing showing the process where data are transmitted and received between each configuration included in the machine translation base global on-line community service providing system of figure 1 according to a preferred embodiment of the present invention. Hereinafter, an example of the process wh

라도 5에 도시된 데이터가 송수신되는 과정이 변경될 수 있는 것은 기술분야에 속하는 당업자에게 자명하다.

도 5를 참조하면, 커뮤니티 서비스 제공 서버는, 적어도 하나의 주제를 가지는 게시판 내 적어도 하나의 종류의 멀티미디어 콘텐츠에 대응하는 게시글이 등록되면 게시글을 업로드하고(S5100), 사용자 단말에서 게시글을 감지된 언어로 출력하도록 자동번역을 수행한다(S5200).

커뮤니티 서비스 제공 서버는, 게시글에 댓글이 포함된 경우 댓글을 감지된 언어로 자동번역하여 출력한다(S5300).

상술한 단계들(S5100~S5300)간의 순서는 예시일 뿐, 이에 한정되지 않는다. 즉, 상술한 단계들(S5100~S5300)간의 순서는 상호 변동될 수 있으며, 이 중 일부 단계들은 동시에 실행되거나 삭제될 수도 있다.

이와 같은 도 5의 자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 방법에 대해서 설명되지 아니한 사항은 앞서 도 1 내지 도 4를 통해 자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 방법에 대하여 설명된 내용과 동일하거나 설명된 내용으로부터 용이하게 유추 가능하므로 이하 설명을 생략하도록 한다.

도 5를 통해 설명된 일 실시예에 따른 자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 방법은, 컴퓨터에 의해 실행되는 애플리케이션이나 프로그램 모듈과 같은 컴퓨터에 의해 실행가능한 명령어를 포함하는 기록 매체의 형태로도 구현될 수 있다. 컴퓨터 판독 가능 매체는 컴퓨터에 의해 액세스될 수 있는 임의의 가용 매체일 수 있고, 휘발성 및 비휘발성 매체, 분리형 및 비분리형 매체를 모두 포함한다. 또한, 컴퓨터 판독가능 매체는 컴퓨터 저장 매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함한다.

전술한 본 발명의 일 실시예에 따른 자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 방법은, 단말기에 기본적으로 설치된 애플리케이션(이는 단말기에 기본적으로 탑재된 플랫폼이나 운영체제 등에 포함된 프로그램을 포함할 수 있음)에 의해 실행될 수 있고, 사용자가 애플리케이션 스토어 서버, 애플리케이션 또는 해당 서비스와 관련된 웹 서버 등의 애플리케이션 제공 서버를 통해 마스터 단말기에 직접 설치한 애플리케이션(즉, 프로그램)에 의해 실행될 수도 있다. 이러한 의미에서, 전술한 본 발명의 일 실시예에 따른 자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 방법은 단말기에 기본적으로

ere data are each other transmitted and received through fig. 5 with each configuration will be illustrated. But the present application is not analyzed limitedly to this kind of the embodiment and the process where the data is illustrated in Figure 5 is transmitted and received according to various embodiments illustrated can be changed it is obvious to the person skilled in the art belonging to the technical field.

Referring to Figure 5, as to the community service providing server, if the posted writing corresponding to the multimedia contents of the kind of the inside of the bulletin board at least one having at least one theme is registered, it performs the machine translation in the user terminal in order to output the posted writing to the sensed language (S5200).

As to the community service providing server, in case the reply is included in the posted writing it automatically translates the reply to the sensed language and it outputs (S5300).

The above-described order between the step (S5100~S5300) is the example. It is not thus restricted. That is, the above-described order between the step (S5100~S5300) can be mutually changed and part steps are simultaneously performed among this or it can be deleted.

Since the easily is assumable from the content that is identical with 1 through the content explained through fig. 4 about the machine translation base global on-line community service providing method or is explained the matter which is not explained about the machine translation base global on-line community service providing method of this kind of fig. 5 the before hereinafter omits the description.

The machine translation base global on-line community service providing method according to the embodiment explained through fig. 5 can be implemented with the computer such as the application performed with the computer and program module in the form of the recording medium including the executable instruction. It can be the arbitrary available media in which the computer readable medium can be accessed with the computer and the volatile non-volatile medium, and the decoupling-type and non-removable media the altogether are included. Moreover, the computer readable medium altogether may include the computer storage medium. The volatile non-volatility in which the computer storage medium is implemented as the computer readable instruction, the data construct, and the arbitrary method for the storage of the information like the program module or other data or technology, and the decoupling-type and non-removable media the altogether are included.

The machine translation base global on-line community service providing method according to an above-described preferred embodiment of the present invention can be performed by the application (the program included in the platform or the operating system etc. this is mounted in the terminal can be included) elementarily installed at the terminal and the user can be performed by the direct application (in other words, the program) set up through the application providing server including the application store server, the web server associated

설치되거나 사용자에 의해 직접 설치된 애플리케이션(즉, 프로그램)으로 구현되고 단말기에 등의 컴퓨터로 읽을 수 있는 기록매체에 기록될 수 있다.

with the application or the target service etc. in the master terminal. In such meaning, the above-described machine translation base global on-line community service providing method according to a preferred embodiment of the present invention is elementarily installed at the terminal or the terminal can be implemented by the user as the direct installed application (in other words, the program) and it can be recorded in the terminal in the computer-readable medium of back.

전술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 예를 들어, 단일형으로 설명되어 있는 각 구성 요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.

The above-described description of the invention is the thing for the example and it will be able to understand that it is easy to the dissimilar detailed form after a person skilled in the art of the technical field in which the invention belongs does not change the technical mapping of the invention or the essential characteristic deformation is possible. Therefore, in the above, with illustrative and it is not restrictive described embodiments have to understand from every respect. For example, each element explained with the single format is dispersed and it can be performed and it can be performed to the form in which elements which are explained similarly that it is dispersed are combined.

본 발명의 범위는 상기 상세한 설명보다는 후술하는 특허청구 범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

The scope of the present invention is shown than the detailed explanation with the patent claim which will be described later and all changes drawn from the meaning of the patent claim and range and equality concept or the transformed form are included in the scope of the present invention and it has to be interpreted.

## **도면에 대한 간단한 설명**

도 1은 본 발명의 일 실시예에 따른 자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 시스템을 설명하기 위한 도면이다.

도 2는 도 1의 시스템에 포함된 커뮤니티 서비스 제공 서버를 설명하기 위한 블록 구성도이다.

도 3 및 도 4는 본 발명의 일 실시예에 따른 자동번역 기반 글로벌 온라인 커뮤니티 서비스가 구현된 일 실시예를 설명하기 위한 도면이다.

도 5는 본 발명의 일 실시예에 따른 자동번역 기반 글로벌 온라인 커뮤니티 서비스 제공 방법을 설명하기 위한 동작 흐름도이다.

## **Brief explanation of the drawing**

Figure 1 is a drawing illustrating the machine translation base global on-line community service providing system according to a preferred embodiment of the present invention.

Figure 2 is a block diagram for illustrating the community service providing server included in the system of fig. 1.

Figures 3 and 4 are the drawing illustrating the embodiment in which the machine translation base global online community service according to a preferred embodiment of the present invention is implemented.

Figure 5 is an operational flow chart for illustrating the machine translation base global on-line community service providing method according to a preferred embodiment of the present invention.

## **면책안내**

본 문서는 특허 및 과학기술문헌 전용의 첨단 자동번역 시스템을 이용해 생성되었습니다. 따라서 부분적으로 오역의 가능성이 있으며, 본 문서를 자격을 갖춘 전문 번역가에 의한 번역물을 대신하는 것으로 이용되어서는 안 됩니다. 시스템 및 네트워크의 특성때문에 발생한 오역과 부분 누락, 데이터의 불일치 등에 대하여 본원은 법적인 책임을 지지 않습니다. 본 문서는 당사의 사전 동의 없이 권한이 없는 일반 대중을 위해 DB 및 시스템에 저장되어 재생, 복사, 배포될 수 없음을 알려드립니다.

(The document produced by using the high-tech machine translation system for the patent and science & technology literature. Therefore, the document can include the mistranslation, and it should not be used as a translation by a professional translator. We hold no legal liability for inconsistency of mistranslation, partial omission, and data generated by feature of system and network. We would like to inform you that the document cannot

t be regenerated, copied, and distributed by being stored in DB and system for unauthorized general public without our consent.)