



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년04월25일
(11) 등록번호 10-2661537
(24) 등록일자 2024년04월23일

(51) 국제특허분류(Int. Cl.)
G06Q 50/20 (2012.01) G06F 40/20 (2020.01)
G09B 19/06 (2006.01) G10L 15/04 (2006.01)
(52) CPC특허분류
G06Q 50/20 (2013.01)
G06F 40/20 (2022.01)
(21) 출원번호 10-2023-0152636
(22) 출원일자 2023년11월07일
심사청구일자 2023년11월07일
(56) 선행기술조사문헌
KR1020210059995 A*
KR1020210109914 A*
JP2008139654 A*
KR1020180105501 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
전훈철
서울특별시 서초구 서초중앙로 188, B동 1009호
(서초동, 아크로비스타)
(72) 발명자
전훈철
서울특별시 서초구 서초중앙로 188, B동 1009호
(서초동, 아크로비스타)
(74) 대리인
특허법인해안

전체 청구항 수 : 총 2 항

심사관 : 이종경

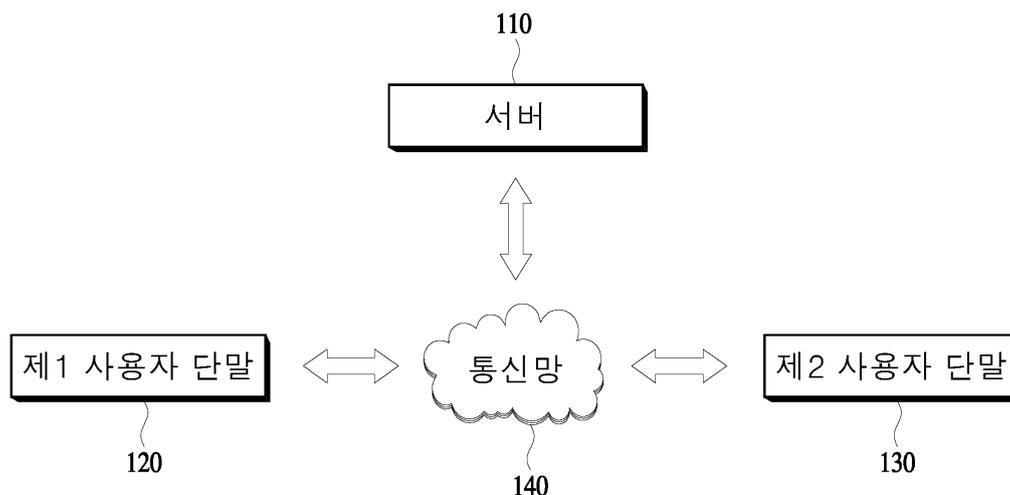
(54) 발명의 명칭 가상 공간에서 언어 학습을 제공하기 위한 장치, 시스템, 및 방법

(57) 요약

가상 공간에서 언어 학습을 제공하기 위한 장치, 시스템, 및 방법이 개시된다. 본 개시의 일 실시예에 따른 시스템은, 어플리케이션에 의해 제공되는 가상 공간에 접속하여 제1 음성 정보를 전송하는 제1 사용자 단말, 가상 공간에 접속하여 제1 음성 정보 및 제1 음성 정보에 대응되는 제1 스크립트 정보에 대해 제2 음성 정보를 전송하는 제2 사용자 단말, 어플리케이션을 제공하고, 제1 음성 정보에 따른 제1 스크립트 정보와 제2 음성 정보에 따른 제2 스크립트 정보를 기초로 제2 사용자 단말을 이용하는 학습자의 언어 학습에 대한 숙련도를 평가하는 서버를 포함한다.

대표도 - 도1

100



(52) CPC특허분류

G09B 19/06 (2013.01)

G10L 15/04 (2013.01)

명세서

청구범위

청구항 1

어플리케이션에 의해 제공되는 가상 공간에 접속하여 제1 음성 정보를 전송하는 제1 사용자 단말;

상기 가상 공간에 접속하여 상기 제1 음성 정보에 대응되는 제1 스크립트 정보에 대해 제2 음성 정보를 전송하는 제2 사용자 단말; 및

상기 어플리케이션을 상기 제1 사용자 단말 및 상기 제2 사용자 단말에 제공하는 서버를 포함하고,

상기 서버는,

상기 제1 음성 정보를 기초로 상기 제1 스크립트 정보를 생성하고, 상기 제2 음성 정보를 기초로 제2 스크립트 정보를 생성하고,

상기 제1 스크립트 정보를 기초로 상기 제1 사용자 단말을 이용해 소리 내어 말을 하는 교육자의 언어 행위의 총량을 나타내는 제1 발화량 정보를 생성하고, 상기 제2 스크립트 정보를 기초로 상기 제2 사용자 단말을 이용해 소리 내어 말을 하는 학습자의 언어 행위의 총량을 나타내는 제2 발화량 정보를 생성하고,

상기 제2 발화량 정보의 제2 발화량이 상기 제1 발화량 정보의 제1 발화량보다 크거나 같은 경우에, 상기 제2 스크립트 정보를 기초로 상기 학습자의 언어 학습에 대한 숙련도를 평가하고,

상기 제2 스크립트 정보에서 서로 다른 종류의 단어 종류의 개수 및 상기 단어 종류의 빈도수를 측정하고,

상기 단어 종류의 개수 및 상기 단어 종류의 빈도수를 기초로 상기 학습자의 언어 학습에 대한 숙련도를 평가하고,

상기 제2 스크립트 정보에서 기준에 발생한 언어 표현과 다른 언어 표현의 발생량, 상기 단어의 개수 및 상기 단어의 빈도수가 높을수록, 상기 학습자의 언어 학습에 대한 숙련도를 높게 평가하고,

상기 제2 발화량이 상기 제1 발화량보다 작은 경우에, 상기 제1 스크립트 정보에 대응되는 문장에 포함되는 단어들 각각에 매칭되는 오브젝트들을 포함하는 월드를 상기 가상 공간 내에서 구현하는 것을 특징으로 하는, 시스템.

청구항 2

삭제

청구항 3

삭제

청구항 4

삭제

청구항 5

제1 항에 있어서,

상기 서버는,

상기 학습자의 취학 연령 및 상기 숙련도에 따른 복수의 월드들을 설정하는 것을 특징으로 하는, 시스템.

발명의 설명

기술 분야

[0001] 본 개시는 전자 장치 및 그의 동작 방법에 관한 것이다. 보다 상세하게는, 본 개시는 가상 공간에서 언어 학습을 제공하기 위한 장치, 시스템, 및 방법에 관한 것이다.

배경 기술

[0002] 최근 외국어 교육에 대한 관심이 증대되고 있다. 대부분의 언어에서 약 1000개의 단어들만 회화자가 일상 회화에서 말하고자 하는 것의 약 85%를 차지한다. 또한, 일상 회화의 약 98%의 경우, 3000개의 단어로 의사소통이 이루어진다. 즉, 특정 언어를 학습하고자 하는 학습자가 해당 언어에 대한 약 3000개의 단어를 알고 있다면, 해당 학습자는 해당 언어로 의사소통을 할 수 있다.

[0003] 예시적인 언어를 학습하는 과정은, 교육자가 해당 언어로 말하고, 학습자는 교육자의 말하기를 듣고 따라한다. 이때, 교육자의 말하기가 입력으로 가정하고, 학습자의 말하기가 출력으로 가정하면, 입력 대비 출력이 더 많을수록 언어 학습 효과가 있다. 그러나, 대부분의 국가에서 외국어 특히 영어를 학습하는 과정에서, 입력 대비 출력이 상대적으로 적어 비효율적이다.

[0004] 한편, 코로나 등을 이유로 비대면 학습이 증가하고 있는데, 대면 학습에 비해 교육 성취도가 상대적으로 낮다는 문제점이 있을 수 있다. 이에, 말하기 위주의 교육이 가능한 가상 공간에서 비대면으로 언어를 효율적으로 학습하며, 이에 따라 학습자의 언어 학습에 대한 숙련도를 평가하기 위한 시스템이 필요하다.

선행기술문헌

특허문헌

[0005] (특허문헌 0001) 대한민국 공개특허공보 제10-2010-0043393호, 2010.04.29 공개

발명의 내용

해결하려는 과제

[0006] 본 개시에 개시된 실시예는, 어린아이의 모국어 취득과정과 유사한 방식으로 언어를 교육하기 위해서, 가상 공간에서 언어 학습을 제공하는 시스템을 구축하는 데 그 목적이 있다.

[0007] 본 개시가 해결하고자 하는 과제들은 이상에서 언급된 과제로 제한되지 않으며, 언급되지 않은 또 다른 과제들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0008] 상술한 기술적 과제를 달성하기 위한 본 개시의 일 측면에 따른 시스템은, 어플리케이션에 의해 제공되는 가상 공간에 접속하여 제1 음성 정보를 전송하는 제1 사용자 단말, 가상 공간에 접속하여 제1 음성 정보 및 제1 음성 정보에 대응되는 제1 스크립트 정보에 대해 제2 음성 정보를 전송하는 제2 사용자 단말, 및 어플리케이션을 제1 사용자 단말 및 제2 사용자 단말에 제공하는 서버를 포함한다. 서버는, 제1 음성 정보를 기초로 제1 스크립트 정보를 생성하고, 제2 음성 정보를 기초로 제2 스크립트 정보를 생성하고, 제1 스크립트 정보를 기초로 제1 사용자 단말을 이용해 소리 내어 말을 하는 교육자의 언어 행위의 총량을 나타내는 제1 발화량 정보를 생성하고, 제2 스크립트 정보를 기초로 제2 사용자 단말을 이용해 소리 내어 말을 하는 학습자의 언어 행위의 총량을 나타내는 제2 발화량 정보를 생성하고, 제2 발화량 정보의 제2 발화량이 제1 발화량 정보의 제1 발화량보다 크거나 같은 경우에, 제2 스크립트 정보를 기초로 학습자의 언어 학습에 대한 숙련도를 평가한다.

발명의 효과

[0009] 본 개시의 기술적 과제 해결 수단에 의하면, 가상 공간 상에서 예상치 못한 이벤트에 대한 대응능력과, 미리 정해진 챗봇보다는 의외성에 발현되는 대응능력을 언어적으로 소통하는 것으로 발휘함으로써, 학습자의 언어 능력을 향상시키는 효과가 있다.

[0010] 본 개시의 효과들은 이상에서 언급된 효과로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

- [0011] 도 1은 본 개시에 따른 시스템을 나타낸 도면이다.
- 도 2는 도 1의 서버의 구성을 나타낸 도면이다.
- 도 3은 본 개시에 따른 방법을 나타낸 순서도이다.
- 도 4는 본 개시의 예시적인 실시예들을 실행하기 위한 웹화면을 예시적으로 나타낸 도면이다.
- 도 5 및 도 6은 본 개시의 예시적인 실시예를 실행한 화면을 예시적으로 나타낸 도면들이다.

발명을 실시하기 위한 구체적인 내용

- [0012] 본 개시 전체에 걸쳐 동일 참조 부호는 동일 구성요소를 지칭한다. 본 개시가 실시예들의 모든 요소들을 설명하는 것은 아니며, 본 개시가 속하는 기술분야에서 일반적인 내용 또는 실시예들 간에 중복되는 내용은 생략한다. 명세서에서 사용되는 '부, 모듈, 부재, 블록'이라는 용어는 소프트웨어 또는 하드웨어로 구현될 수 있으며, 실시예들에 따라 복수의 '부, 모듈, 부재, 블록'이 하나의 구성요소로 구현되거나, 하나의 '부, 모듈, 부재, 블록'이 복수의 구성요소들을 포함하는 것도 가능하다.
- [0013] 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 직접적으로 연결되어 있는 경우뿐 아니라, 간접적으로 연결되어 있는 경우를 포함하고, 간접적인 연결은 무선 통신망을 통해 연결되는 것을 포함한다.
- [0014] 또한 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.
- [0015] 명세서 전체에서, 어떤 부재가 다른 부재 "상에" 위치하고 있다고 할 때, 이는 어떤 부재가 다른 부재에 접해 있는 경우뿐 아니라 두 부재 사이에 또 다른 부재가 존재하는 경우도 포함한다.
- [0016] 제1, 제2 등의 용어는 하나의 구성요소를 다른 구성요소로부터 구별하기 위해 사용되는 것으로, 구성요소가 전술된 용어들에 의해 제한되는 것은 아니다.
- [0017] 단수의 표현은 문맥상 명백하게 예외가 있지 않는 한, 복수의 표현을 포함한다.
- [0018] 각 단계들에 있어 식별부호는 설명의 편의를 위하여 사용되는 것으로 식별부호는 각 단계들의 순서를 설명하는 것이 아니며, 각 단계들은 문맥상 명백하게 특정 순서를 기재하지 않는 이상 명기된 순서와 다르게 실시될 수 있다.
- [0019] 이하 첨부된 도면들을 참고하여 본 개시의 작용 원리 및 실시예들에 대해 설명한다.
- [0020] 본 명세서에서 '본 개시에 따른 시스템'은 연산처리를 수행하여 사용자에게 결과를 제공할 수 있는 다양한 장치들이 모두 포함된다. 예를 들어, '본 개시에 따른 시스템'은, 컴퓨터, 서버 장치 및 휴대용 단말기를 모두 포함하거나, 또는 어느 하나의 형태가 될 수 있다.
- [0021] 여기에서, 컴퓨터는 예를 들어, 웹 브라우저(WEB Browser)가 탑재된 노트북, 데스크톱(desktop), 랩톱(laptop), 태블릿 PC, 슬레이트 PC 등을 포함할 수 있다.
- [0022] 휴대용 단말기는 예를 들어, 휴대성과 이동성이 보장되는 무선 통신 장치로서, 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, 스마트 폰(Smart Phone) 등과 같은 모든 종류의 핸드헬드(Handheld) 기반의 무선 통신 장치와 시계, 반지, 팔찌, 발찌, 목걸이, 안경, 콘택트 렌즈, 또는 머리 착용형 장치(head-mounted-device(HMD) 등과 같은 웨어러블 장치를 포함할 수 있다.
- [0023] 서버 장치는 외부 장치와 통신을 수행하여 정보를 처리하는 서버로써, 애플리케이션 서버, 컴퓨팅 서버, 데이터베이스 서버, 파일 서버, 게임 서버, 메일 서버, 프록시 서버 및 웹 서버 등을 포함할 수 있다.
- [0024] 본 개시에 따른 인공지능과 관련된 기능은 프로세서와 메모리를 통해 동작된다. 프로세서는 하나 또는 복수의 프로세서로 구성될 수 있다. 이때, 하나 또는 복수의 프로세서는 CPU, AP, DSP(Digital Signal Processor) 등

과 같은 범용 프로세서, GPU, VPU(Vision Processing Unit)와 같은 그래픽 전용 프로세서 또는 NPU와 같은 인공지능 전용 프로세서일 수 있다. 하나 또는 복수의 프로세서는, 메모리에 저장된 기 정의된 동작 규칙 또는 인공지능 모델에 따라, 입력 데이터를 처리하도록 제어한다. 또는, 하나 또는 복수의 프로세서가 인공지능 전용 프로세서인 경우, 인공지능 전용 프로세서는, 특정 인공지능 모델의 처리에 특화된 하드웨어 구조로 설계될 수 있다.

[0025] 기 정의된 동작 규칙 또는 인공지능 모델은 학습을 통해 만들어진 것을 특징으로 한다. 여기서, 학습을 통해 만들어진다는 것은, 기본 인공지능 모델이 학습 알고리즘에 의하여 다수의 학습 데이터들을 이용하여 학습됨으로써, 원하는 특성(또는, 목적)을 수행하도록 설정된 기 정의된 동작 규칙 또는 인공지능 모델이 만들어짐을 의미한다. 이러한 학습은 본 개시에 따른 인공지능이 수행되는 기기 자체에서 이루어질 수도 있고, 별도의 서버 및/또는 시스템을 통해 이루어 질 수도 있다. 학습 알고리즘의 예로는, 지도형 학습(supervised learning), 비지도형 학습(unsupervised learning), 준지도형 학습(semi-supervised learning) 또는 강화 학습(reinforcement learning)이 있으나, 전술한 예에 한정되지 않는다.

[0026] 인공지능 모델은 하나의 인공지능 모델일 수 있고, 복수의 인공지능 모델로 구현될 수도 있다. 인공지능 모델은 뉴럴 네트워크(또는 인공 신경망)로 구성될 수 있으며, 기계학습과 인지과학에서 생물학의 신경을 모방한 통계학적 학습 알고리즘을 포함할 수 있다. 뉴럴 네트워크는 시냅스의 결합으로 네트워크를 형성한 인공 뉴런(노드)이 학습을 통해 시냅스의 결합 세기를 변화시켜, 문제 해결 능력을 가지는 모델 전반을 의미할 수 있다. 뉴럴 네트워크의 뉴런은 가중치 또는 바이어스의 조합을 포함할 수 있다. 뉴럴 네트워크는 하나 이상의 뉴런 또는 노드로 구성된 하나 이상의 레이어(layer)를 포함할 수 있다. 예시적으로, 장치(100)는 input layer, hidden layer, output layer를 포함할 수 있다. 장치(100)를 구성하는 뉴럴 네트워크는 뉴런의 가중치를 학습을 통해 변화시킴으로써 임의의 입력(input)으로부터 예측하고자 하는 결과(output)를 추론할 수 있다.

[0027] 프로세서는 뉴럴 네트워크를 생성하거나, 뉴럴 네트워크를 훈련(train, 또는 학습(learn))하거나, 수신되는 입력 데이터를 기초로 연산을 수행하고, 수행 결과를 기초로 정보 신호(information signal)를 생성하거나, 뉴럴 네트워크를 재훈련(retrain)할 수 있다. 뉴럴 네트워크의 모델들은 GoogleNet, AlexNet, VGG Network 등과 같은 CNN(Convolution Neural Network), R-CNN(Region with Convolution Neural Network), RPN(Region Proposal Network), RNN(Recurrent Neural Network), S-DNN(Stacking-based deep Neural Network), S-SDNN(State-Space Dynamic Neural Network), Deconvolution Network, DBN(Deep Belief Network), RBM(Restrcted Boltzman Machine), Fully Convolutional Network, LSTM(Long Short-Term Memory) Network, Classification Network 등 다양한 종류의 모델들을 포함할 수 있으나 이에 제한되지는 않는다. 프로세서는 뉴럴 네트워크의 모델들에 따른 연산을 수행하기 위한 하나 이상의 프로세서를 포함할 수 있다. 예를 들어 뉴럴 네트워크는 심층 뉴럴 네트워크(Deep Neural Network)를 포함할 수 있다.

[0028] 뉴럴 네트워크는 CNN(Convolutional Neural Network), RNN(Recurrent Neural Network), 퍼셉트론(perceptron), 다층 퍼셉트론(multilayer perceptron), FF(Feed Forward), RBF(Radial Basis Network), DFF(Deep Feed Forward), LSTM(Long Short Term Memory), GRU(Gated Recurrent Unit), AE(Auto Encoder), VAE(Variational Auto Encoder), DAE(Denoising Auto Encoder), SAE(Sparse Auto Encoder), MC(Markov Chain), HN(Hopfield Network), BM(Boltzmann Machine), RBM(Restricted Boltzmann Machine), DBN(Depp Belief Network), DCN(Deep Convolutional Network), DN(Deconvolutional Network), DCIGN(Deep Convolutional Inverse Graphics Network), GAN(Generative Adversarial Network), LSM(Liquid State Machine), ELM(Extreme Learning Machine), ESN(Echo State Network), DRN(Deep Residual Network), DNC(Differentiable Neural Computer), NTM(Neural Turning Machine), CN(Capsule Network), KN(Kohonen Network) 및 AN(Attention Network)를 포함할 수 있으나 이에 한정되는 것이 아닌 임의의 뉴럴 네트워크를 포함할 수 있음은 통상의 기술자가 이해할 것이다.

[0029] 본 개시의 예시적인 실시예에 따르면, 프로세서는 GoogleNet, AlexNet, VGG Network 등과 같은 CNN(Convolution Neural Network), R-CNN(Region with Convolution Neural Network), RPN(Region Proposal Network), RNN(Recurrent Neural Network), S-DNN(Stacking-based deep Neural Network), S-SDNN(State-Space Dynamic Neural Network), Deconvolution Network, DBN(Deep Belief Network), RBM(Restrcted Boltzman Machine), Fully Convolutional Network, LSTM(Long Short-Term Memory) Network, Classification Network, Generative Modeling, eXplainable AI, Continual AI, Representation Learning, AI for Material Design, 자연어 처리를 위한 BERT, SP-BERT, MRC/QA, Text Analysis, Dialog System, GPT-3, GPT-4, 비전 처리를 위한 Visual Analytics, Visual Understanding, Video Synthesis, ResNet 데이터 지능을 위한 Anomaly Detection,

Prediction, Time-Series Forecasting, Optimization, Recommendation, Data Creation 등 다양한 인공지능 구조 및 알고리즘을 이용할 수 있으며, 이에 제한되지 않는다.

[0030]

도 1은 본 개시에 따른 시스템을 나타낸 도면이다.

[0031]

도 1을 참조하면, 영어를 포함하여 대부분의 언어에서 약 1000개의 단어가 일상 회화에서 회화자가 말하고자 하는 것의 약 85%를 차지할 수 있다. 또한, 일상 회화의 약 98%의 경우, 3000개의 단어로 의사소통이 이루어진다. 즉, 특정 언어를 학습하고자 하는 학습자가 해당 언어에 대한 약 3000개의 단어를 알고 있다면, 해당 학습자는 해당 언어로 의사소통을 할 수 있다. 시스템(100)은 High Frequency Language의 현상에 근거하여 가상 공간에 노출된 학습자의 단어에 대한 빈도수와 종류(예를 들어 서로 다른 단어 종류의 수)를 측정하여 학습자의 언어 숙련도를 평가할 수 있다. 예를 들어, 원어민, 고급 학습자, 중급 학습자, 및 초급 학습자 순으로 그 단어 종류의 수 및/또는 단어 종류의 빈도수가 다르게 나타날 수 있다. 시스템(100)은 PLEPS(Play Language Evaluation Program of Speaking)로서 제공될 수 있다. 시스템(100)은, 기존의 말하기에 대한 평가, 즉 주로 사람에 의한 주관적 평가와는 다른 방법을 채택하고 있다. 이러한 평가가 가능한 이유는 본 개시의 시스템(100)에 의해 제공되는 수업 환경이 목표어가 정해져 있지 않은 모국어 환경이고, 실생활과 흡사한 환경(예, 공간, 맥락, 상호작용대상, 비선형적 상황 등)으로 조성되기 때문이다. 학습자가 언어 발달의 가장 중요한 요소 중 하나인 단어를 암기하지 않고, 시스템(100)은 학습자의 경험이나 행동을 통해 학습자가 직접적으로 체득하게 할 수 있다. 시스템(100)과 협업을 진행하고 있는 Glee English에서 33명의 학생들에게 “Do you know What Furnace is?” 라고 물었을 때, 33명의 아이들은 모두 수업 초반에는 그 뜻을 모른다고 했다. 하지만, 11시간 동안의 수업이 진행된 후, Glee English에서 같은 질문을 하자 33명의 아이들은 "무기를 만들 때 쓰는 도구요", "고기를 구워 먹었을 때도 썼어요"라고 대답을 했다. 만약 "난로가 무엇인지 답해보세요"라고 질문이 한국인에게 주어진다면, 한국인 학생들은 아마 "난로"의 쓰임새나 사용처 등을 이야기할 것이다. 반대로 일반 영어수업을 들은 아이들에게 “Furnace가 뭐예요?”라는 질문이 주어진다면, "난로요“, "용광로요"라는 암기식을 답을 하는 것 제2 외국어의 기본적 단어 학습법과는 다르게 모국어 단어의 체득방식의 교육이 시스템(100)에 의해 자연스럽게 이루어 질 수 있다. 모국어 습득 이론에 따르면, 생후 11개월이 된 아이가 주변에서 들리는 낱말과 말소리를 듣고 흉내내기 시작하며, 아이가 생후 18개월이 된 무렵에 사물에 대한 명칭과 행동에 대한 정의 등을 2개 내지 3개의 낱말로 이루어진 문장으로 말하기 시작할 수 있다. 시스템(100)에 의해 제공되는 초급 수업에서는, 걸음마와 같은 조작성, 많은 시각적 및 청각적 소재로 하여금 말소리와 낱말, 행동들과 언어를 훈련하고 만 60시간 후에는 점진적으로 "전보문식(telegraphic)" 발화가 시작될 수 있도록 구성되어 있다. 예를 들면, 시스템(100)에서 제공되는 가상 공간에 포함되는 월드는, 초급 수업에서 걸을 수 있는 경로(path)를 오브젝트로 하는 등 비교적 단순하게 구현될 수 있다. 또한, 학습자의 숙련도가 증가함에 따라, 도구의 사용과 행동 범위의 자유로움 등으로 더 많은 문장을 구사할 수 있도록 월드가 자연스럽게 의문문과 부정문, 문장의 연결, 단어의 확장등의 결과를 제공하기 위한 다양하고 복잡한 오브젝트들로 구성될 수 있다. 시스템(100)은, 학습자가 전문적 단어의 확장과 복문 구성을 자연스럽게 구사할 수 있도록, 고급 수업 과정에서 원어민 기준으로 취학 연령의 수학이나 과학 등의 수업 내용을 오브젝트로 하는 워드로 구성할 수 있다.

[0032]

시스템(100)은 서버(110), 제1 사용자 단말(120), 제2 사용자 단말(130), 및 통신망(140)을 포함할 수 있다.

[0033]

서버(110)는, 컴퓨팅 장치로 지칭될 수 있다. 서버(110)는, 통신망(140)을 통해 어플리케이션을 제1 사용자 단말(120) 및 제2 사용자 단말(130)에 제공할 수 있다. 어플리케이션은, 어플리케이션 프로그램 또는 프로그램일 수 있고, PLEPS(Play Language Evaluation Program of Speaking)로 지칭될 수 있으며, 가상 공간을 제공할 수 있다. 어플리케이션에 의해 제공되는 가상 공간은, 하나 이상의 워드를 포함할 수 있다. 워드는 다양한 환경을 나타내기 위한 복수의 오브젝트들을 포함할 수 있다. 예를 들면, 워드에 포함된 오브젝트들은, 사람, 동물, 식물, 사물, 바다, 땅, 공기, 바람, 날씨 등의 환경뿐만 아니라 상상 속에서 존재하는 것들(예, 용, 좀비 등) 등을 모두 포함할 수 있다.

[0034]

서버(110)는, 제1 음성 정보를 기초로 제1 스크립트 정보를 생성하고, 제2 음성 정보를 기초로 제2 스크립트 정보를 생성하고, 제1 스크립트 정보를 기초로 제1 사용자 단말을 이용해 소리 내어 말을 하는 교육자의 언어 행위의 총량을 나타내는 제1 발화량 정보를 생성하고, 제2 스크립트 정보를 기초로 제2 사용자 단말을 이용해 소리 내어 말을 하는 학습자의 언어 행위의 총량을 나타내는 제2 발화량 정보를 생성하고, 제2 발화량 정보의 제2 발화량이 제1 발화량 정보의 제1 발화량보다 크거나 같은 경우에, 제2 스크립트 정보를 기초로 학습자의 언어 학습에 대한 숙련도를 평가할 수 있다.

[0035]

일부 실시예들에서, 서버(110)는, 거대하고 복잡한 워드에 약 3개 내지 4개의 표식에서 적절한 표현에 대한 학

습이 이루어지도록, 다양한 월드를 생성할 수 있다. 해당 월드의 표식에 나타내는 문장은, 실제 그 지점에서 경험, 행동, 감정, 감각과 연결되는 표현으로, 해당 표식이 있는 월드의 오브젝트는 문장에 포함되는 단어에 매칭될 수 있다. 따라서, 학습자는 자연스럽게 언어를 체득할 수 있다. 예를 들면, "헨젤과 그레텔이 빵 부스러기로 가야할 길을 잃지 않았던 것처럼, 당신은 path를 따라 가서 길을 잃지 않을 수 있습니다" 라는 문장이 주어진 월드에서, 해당 월드의 제1 오브젝트는, 경로(path)이고, 제2 오브젝트들은 경로 상에 놓인 빵 부스러기들이며, 제3 오브젝트들은 경로에 위치하는 헨젤 캐릭터와 그레텔 캐릭터일 수 있다.

[0036] 일부 실시예들에서, 가상 공간 내의 월드에 놓인 원어민, 고급 학습자, 중급 학습자, 및 초급 학습자 순으로 그 단어 종류의 수 및/또는 단어 종류의 빈도수가 다르게 나타날 수 있다. 향후에 이러한 현상이 가상 공간 내에 존재하는 월드(구체적으로 예를 들면, 월드에 포함된 오브젝트들)을 구성 및 생성하는데 지표(index)가 될 수 있다. 따라서, 서버(110)는, 가상 공간 내의 월드에 놓인 원어민, 고급 학습자, 중급 학습자, 및 초급 학습자에 따라 다양한 오브젝트들로 구성되는 월드를 구현하면, 학습자는 해당 월드 내에서 오브젝트들을 이용하여 언어를 학습할 수 있다. 이에 따라, 서버(110)는 해당 학습자의 언어 학습에 대한 숙련도를 측정할 수 있다. 또한, 단어 종류의 수 및/또는 단어 종류의 빈도수가 외에 가상 공간 내에서 이루어지는 수업 시간 동안에서 학습자의 점유율(Possession), 및 학습자가 사용하는 표현의 증감 등을 측정해 실질적 모국어화 정도를 측정할 수 있다. 서버(110)는, 말하기 위주의 교육이 가능한 가상 공간에서 비대면으로 언어를 효율적으로 학습하여, 이에 따라 학습자의 언어 학습에 대한 숙련도를 평가할 수 있다.

[0037] 서버(110)는, 모국어 체득과정에 대한 데이터를 수집할 수 있다. 대부분의 모국어 체득과정에 대한 연구는 한정된 샘플(sample)에 대해 횡단 또는 종단 연구를 통해 이루어질 수 있다. 하지만 모든 표본에 대한 환경이 동일하기 어려울 수 있다. 예를 들어, 놀이공원에서 새끼 사자가 어미 사자랑 노는 모습을 봤을 때 어떤 언어를 발화하는지, 여러 마리의 실험쥐를 각기 다른 미로에 넣어서 결과를 얻는 것과 같이, 또는, 실제 동물원에서의 학습자는 새끼 사자와 어미사자와 노는 모습을 볼 수도 있고 못 볼 수도 있지만, 가상 환경에서는 늘 새끼 사자와 어미사자가 놀고 있는 환경을 만들어 줌으로써 수 많은 학습자들이 같은(통제된) 환경에 노출되고 그 때, 서버(110)에 의해 제공되는 다양하고 많은 수업에 대한 데이터는 외국어를 모국어로서 체득하는 과정의 여러 표본에 대한 종횡단 연구의 자료가 될 수 있다. 이러한 데이터는 어느 공간에서 어떤 말들이 공통적으로 발화되는 지 등에 활용될 수 있다. 또한, 이러한 데이터는 AI의 Deep Learning을 위한 데이터로 활용될 수 있고, 현재 원어민이 하고 있는 수업의 일부분을 NPC(Non player Character)가 아닌 스스로 생각하는 AI NPC로 바꿀 수 있는 데이터로 생성될 수 있으며, 향후 서버(110)에 의해 제공되는 서비스의 고도화와 인공지능 원본 데이터(Raw Data)로서 제공될 수 있다. 수없이 많은 영어(제2외국어) 교육은 "모국어 습득방식"을 표방하고 있고, 기존의 EFL 방식과는 다르게 원어민이 그 언어를 가르치며 모국어처럼 습득된다고 말한다. 중요한 것은 교수와 학습자의 교육이 이뤄지는 곳은 대부분 "교실"이라는 점이고, 우리가 우리말을 익히는 데는 교실이 아닌 수많은 "환경"에 노출되며 언어를 체득해 간다는데 있어서 근본적인 차이가 보인다. 그렇다고 그러한 환경에 놓인다고 해서 무조건 외국어가 모국어처럼 체득되는 건 아니다. 모국어 체득의 가장 중요한 교실 밖 환경은 공간, 상황, 맥락 특히 부모님처럼 학습자의 감정이나 need를 언어가 정확히 통하지 않는다 하더라도 같이 행동, 공감하며 그런 한 경험을 언어로 연결지을 수 있어야 한다. 즉 정교화된 모국어 체득환경이 주어져야 제2외국어가 모국어화 될 수 있다

[0038] 제1 사용자 단말(120)은, 어플리케이션에 의해 제공되는 가상 공간에 접속하여 제1 음성 정보를 서버(110) 및 제2 사용자 단말(130)에 전송할 수 있다. 어플리케이션은 서버(110)에 의해 제공될 수 있다. 제1 사용자 단말(120)을 사용하는 사용자는 예를 들어 언어 학습을 지도하고 학습자에게 교육시키는 교육자 또는 학습을 지도하는 지도자일 수 있다. 교육자는 제1 사용자 단말(120)을 통해 교육하고자 하는 언어로 대화를 수행할 수 있으며, 제1 사용자 단말(120)은 교육자의 음성을 포함하는 제1 음성 정보를 전송할 수 있다. 제1 사용자 단말(120)은 제1 음성 정보에 대응되는 제1 스크립트 정보를 시각적으로 표시할 수 있다. 제1 스크립트 정보는 서버(110)에 의해 생성되며, 제1 스크립트 정보는 제1 음성 정보에 대응되는 문자, 단어, 및 문장으로 구성되는 스크립트를 포함할 수 있다. 제1 사용자 단말(120)은 제2 음성 정보를 청각적으로 출력할 수 있고, 제2 음성 정보에 대응되는 제2 스크립트 정보를 시각적으로 표시할 수 있다. 제2 음성 정보는 제2 사용자 단말(130)로부터 수신된 정보이고, 제2 스크립트 정보는 제2 음성 정보에 대한 스크립트 정보일 수 있다. 제1 사용자 단말(120)은 교육자에 대응되는 가상의 교육자 캐릭터 및/또는 학습자에 대응되는 가상의 학습자 캐릭터를 어플리케이션의 가상 공간에 표시할 수 있다. 교육자가 제1 사용자 단말(120)을 통해 교육자 캐릭터를 조종할 수 있으며, 교육자 캐릭터는 교육자의 입력에 의해 가상 공간 내에서 자유롭게 움직일 수 있다. 교육자 캐릭터는 교육자가 말한 언어 표현, 즉 문장의 의미에 매칭되는 행동을 수행할 수 있다. 이를 위해, 제1 사용자 단말(120)은 마이크, 입력 키 등을 포함하는 사용자 입력부, 스피커 등의 음성 출력부, 화면 등의 디스플레이부, 및 프로세서 등을 포

함할 수 있다.

- [0039] 제2 사용자 단말(130)은, 어플리케이션의 가상 공간에 접속하여 제1 음성 정보를 수신할 수 있다. 또한, 제2 사용자 단말(130)은, 가상 공간에서 제1 음성 정보를 청각적으로 출력할 수 있다. 또한, 제2 사용자 단말(130)은, 제1 스크립트 정보를 시각적으로 표시할 수 있다. 제2 사용자 단말(130)은, 제1 음성 정보에 대응되는 제1 스크립트 정보에 대해 제2 음성 정보를 전송할 수 있다. 제2 사용자 단말(130)을 사용하는 사용자는 예를 들어 언어를 학습하고자 하는 학습자, 학생 등일 수 있다. 예를 들어, 학습자가 제2 사용자 단말(130)을 통해 표시된 제1 스크립트 정보를 보고 따라서 말을 하면, 제2 사용자 단말(130)은 학습자가 말한 음성을 포함하는 제2 음성 정보를 서버(110) 및 제1 사용자 단말(110)에 전송할 수 있다. 제2 사용자 단말(130)은 제2 음성 정보에 대응되는 제2 스크립트 정보를 시각적으로 표시할 수 있다. 제2 스크립트 정보는 서버(110)에 의해 생성될 수 있다. 제2 사용자 단말(130)은 가상 공간에서 교육자 캐릭터 및/또는 학습자에 대응되는 가상의 학습자 캐릭터를 어플리케이션의 가상 공간에 표시할 수 있다. 학습자가 제2 사용자 단말(130)을 통해 학습자 캐릭터를 조종할 수 있으며, 학습자 캐릭터는 학습자의 입력에 의해 가상 공간 내에서 자유롭게 움직일 수 있다. 학습자 캐릭터는 교육자 캐릭터의 행동, 제1 음성 정보, 및 제1 스크립트 정보에 따라 행동을 수행할 수 있다. 이를 위해, 제2 사용자 단말(130)은 마이크, 입력 키 등을 포함하는 사용자 입력부, 스피커 등의 음성 출력부, 화면 등의 디스플레이부, 및 프로세서 등을 포함할 수 있다. 사용자 입력부는, 사용자로부터 정보를 입력받기 위한 것으로서, 사용자 입력부를 통해 정보가 입력될 수 있다. 이러한, 사용자 입력부는 하드웨어식 물리 키(예를 들어, 사용자 단말의 전면, 후면 및 측면 중 적어도 하나에 위치하는 버튼, 돔 스위치 (dome switch), 조그 휠, 조그 스위치 등) 및 소프트웨어식 터치 키를 포함할 수 있다. 일 예로서, 터치 키는, 소프트웨어적인 처리를 통해 터치스크린 타입의 디스플레이부 상에 표시되는 가상 키(virtual key), 소프트 키(soft key) 또는 비주얼 키(visual key)로 이루어지거나, 터치스크린 이외의 부분에 배치되는 터치 키(touch key)로 이루어질 수 있다. 한편, 가상 키 또는 비주얼 키는, 다양한 형태를 가지면서 터치스크린 상에 표시되는 것이 가능하며, 예를 들어, 그래픽(graphic), 텍스트(text), 아이콘(icon), 비디오(video) 또는 이들의 조합으로 이루어질 수 있다. 일 실시예에서, 사용자 입력부는 태블릿 패드로 구현되어 사용자가 전자 펜으로 필사함으로써 생성된 텍스트 입력 정보를 수신할 수 있다. 디스플레이부는 터치 센서와 상호 레이어 구조를 이루거나 일체형으로 형성됨으로써, 터치 스크린을 구현할 수 있다. 이러한 터치 스크린은, 서버(200)와 사용자 단말 사이의 입력 인터페이스를 제공하는 사용자 입력부로서 기능함과 동시에, 서버(200)와 사용자 단말 간에 출력 인터페이스를 제공할 수 있다. 디스플레이부는 서버(200) 및/또는 사용자 단말에서 처리되는 정보를 표시(출력)한다. 예를 들어, 디스플레이부는 서버(200)에서 구동되는 응용 프로그램(일 예로, 어플리케이션)의 실행화면 정보, 또는 이러한 실행화면 정보에 따른 UI(User Interface), GUI(Graphic User Interface)를 표시할 수 있다.

- [0040] 통신망(140)은 유선 통신이나 무선통신과 같은 통신 양태를 가리지 않고 구성될 수 있으며, 근거리 통신망(LAN; Local Area Network), 도시권 통신망(MAN; Metropolitan Area Network), 광역 통신망(WAN; Wide Area Network) 등 다양한 통신망으로 구성될 수 있다. 예를 들면, 통신망(140)은 무선 데이터 통신망으로서, 와이파이(WiFi) 통신, 와이파이 다이렉트(WiFi-Direct) 통신, 롱텀 에볼루션(LTE, Long Term Evolution) 통신, 5G 통신, 블루투스 통신(저전력 블루투스(BLE; Bluetooth Low Energy) 통신 포함), 적외선 통신, 초음파 통신 등과 같은 종래의 통신 방법을 적어도 그 일부분에 있어서 구현하는 것일 수 있다. 다른 예를 들면, 통신망(140)은 광 통신망으로서, 라이파이(LiFi, Light Fidelity) 등과 같은 종래의 통신 방법을 적어도 그 일부분에 있어서 구현하는 것일 수 있다.

- [0041] 도 2는 도 1의 서버의 구성을 나타낸 도면이다.

- [0042] 도 2를 참조하면, 서버(200)는 프로세서(210), 통신부(220), 메모리(230), 및 버스(240)를 포함할 수 있다. 프로세서(210), 통신부(220), 및 메모리(230)는 버스(240)를 통해 데이터를 송수신할 수 있다. 서버(200)는 입출력부(입력부, 출력부), 사용자 인터페이스 등의 구성을 더 포함할 수 있다. 서버(200)는 서버 장치를 포함하여 서버로 동작하고, 웹 또는 애플리케이션을 통해 서버에 접속한 사용자에게 자리 배치 프로그램, 서비스를 제공할 수도 있다.

- [0043] 프로세서(210)는 서버(200) 내 구성들의 제어, 작동을 담당하며, 메모리(230)에 저장되어 있는 각종 명령어, 알고리즘, 인공지능 모델을 이용하여 학생과 교사들의 반 배정을 진행할 수 있다. 프로세서(210)는 서버(200) 내의 구성요소들의 동작을 제어하기 위한 알고리즘 또는 알고리즘을 재현한 프로그램에 대한 데이터를 저장하는 메모리(230), 및 메모리(230)에 저장된 데이터를 이용하여 전송한 동작을 수행하는 적어도 하나의 프로세서(210)로 구현될 수 있다. 이때, 메모리(230)와 프로세서(210)는 각각 별개의 칩으로 구현될 수 있다. 또는, 메모리(230)와 프로세서(210)는 단일 칩으로 구현될 수도 있다. 또한, 프로세서(210)는 이하의 도면에서 설명되는

본 개시에 따른 다양한 실시 예들을 서버(200) 상에서 구현하기 위하여, 위에서 살펴본 구성요소들을 중 어느 하나 또는 복수를 조합하여 제어할 수 있다.

- [0044] 일부 실시예들에서, 프로세서(210)는, 가상 공간 상에서 교육자(선생님) 및 학생들의 대화를 스크립트로 생성하고, 각각의 발화량 및 스크립트에 대한 분석을 수행할 수 있다. 학생의 발화량이 교육자의 발화량 보다 높다면, 프로세서(210)는, 효과적인 학습으로 판단할 수 있다. 교육자의 스크립트의 수준과 학생의 스크립트의 수준을 분석, 예를 들어 자주 사용되는 단어, 단어의 난이도, 감정표현, 설명표현 등을 분석함으로써, 학습자가 가상 공간 상에서 체험 및 게임을 하듯 교육을 하면서도, 이에 대한 분석/피드백을 정밀하게 수행할 수 있다.
- [0045] 일부 실시예들에서, 프로세서(210)는, 어플리케이션을 제1 사용자 단말(120) 및 제2 사용자 단말(130)에 제공할 수 있다. 프로세서(210)는, 제1 음성 정보를 기초로 제1 스크립트 정보를 생성하고, 제2 음성 정보를 기초로 제2 스크립트 정보를 생성하고, 제1 스크립트 정보를 기초로 제1 사용자 단말에 대한 제1 발화량 정보를 생성하고, 제2 스크립트 정보를 기초로 제2 사용자 단말에 대한 제2 발화량 정보를 생성하고, 제2 발화량 정보의 제2 발화량이 제1 발화량 정보의 제1 발화량보다 크거나 같은 경우에, 제2 스크립트 정보를 기초로 학습자의 언어 학습에 대한 숙련도를 평가할 수 있다.
- [0046] 일부 실시예들에서, 프로세서(210)는, 거대하고 복잡한 월드에 약 3개 내지 4개의 표식에서 적절한 표현에 대한 학습이 이루어지도록, 다양한 월드를 생성할 수 있다.
- [0047] 일부 실시예들에서, 가상 공간 내의 월드에 놓인 원어민, 고급 학습자, 중급 학습자, 및 초급 학습자 순으로 그 단어 종류의 수 및/또는 단어 종류의 빈도수가 다르게 나타날 수 있다. 프로세서(210)는, 가상 공간 내의 월드에 놓인 원어민, 고급 학습자, 중급 학습자, 및 초급 학습자에 따라 다양한 오브젝트들로 구성되는 월드를 구현할 수 있다.
- [0048] 일부 실시예들에서, 프로세서(210)는 제2 스크립트 정보에서 서로 다른 종류의 단어의 개수 및 빈도수를 측정하고, 단어의 개수 및 단어의 빈도수를 기초로 학습자의 언어 학습에 대한 숙련도를 평가할 수 있다. 단어의 개수 및 단어의 빈도수가 높을수록 숙련도는 높게 평가될 수 있다. 숙련도는 초급, 중급, 고급, 또는 원어민 등으로 이산적으로 분류될 수 있다. 하지만, 이에 한정되는 것은 아니다.
- [0049] 일부 실시예들에서, 프로세서(210)는 제2 스크립트 정보에서 기준에 발생한 언어 표현과 다른 언어 표현의 발생량, 단어의 개수 및 단어의 빈도수가 높을수록, 학습자의 언어 학습에 대한 숙련도를 높게 평가할 수 있다.
- [0050] 일부 실시예들에서, 프로세서(210)는 제2 발화량이 제1 발화량보다 작은 경우에, 제1 스크립트 정보에 대응되는 문장에 포함되는 단어들 각각에 매칭되는 오브젝트들을 포함하는 월드를 가상 공간 내에서 구현할 수 있다. 예를 들면, "헨젤과 그레텔이 빵 부스러기로 가야할 길을 잃지 않았던 것처럼, 당신은 path를 따라 가서 길을 잃지 않을 수 있습니다" 라는 문장이 주어진 월드에서, 해당 월드의 제1 오브젝트는, 경로(path)이고, 제2 오브젝트들은 경로 상에 놓인 빵 부스러기들이며, 제3 오브젝트들은 경로에 위치하는 헨젤 캐릭터와 그레텔 캐릭터일 수 있다. 이는 학습자의 참여도를 높일 수 있다.
- [0051] 일부 실시예들에서, 프로세서(210)는 학습자의 취학 연령 및 숙련도에 따른 복수의 월드들을 설정하고, 숙련도가 증가할수록 더 많은 단어들 각각에 매칭되는 오브젝트들을 포함하는 월드를 설정할 수 있다. 예를 들면, 초급자의 경우, 프로세서(210)는 걷기, 물건 쌓기, 달리기 등의 단순한 오브젝트들로 구성된 월드를 생성할 수 있다. 예를 들면, 중급자의 경우, 프로세서(210)는 동물에게 먹이 주기, 물건 만들기 등의 초급자의 월드보다 복잡한 오브젝트들로 구성된 월드를 생성할 수 있다. 예를 들면, 고급자의 경우, 프로세서(210)는 수학이나 과학 등의 학문, 미션 등의 초급자 및 중급자 각각의 월드보다 더 복잡한 오브젝트들로 구성된 월드를 생성할 수 있다.
- [0052] 통신부(220)는 외부 장치와 통신을 가능하게 하는 하나 이상의 구성 요소를 포함할 수 있으며, 예를 들어, 방송 수신 모듈, 유선통신 모듈, 무선통신 모듈, 근거리 통신 모듈, 위치정보 모듈 중 적어도 하나를 포함할 수 있다. 유선 통신 모듈은, 지역 통신(Local Area Network; LAN) 모듈, 광역 통신(Wide Area Network; WAN) 모듈 또는 부가가치 통신(Value Added Network; VAN) 모듈 등 다양한 유선 통신 모듈뿐만 아니라, USB(Universal Serial Bus), HDMI(High Definition Multimedia Interface), DVI(Digital Visual Interface), RS-232(recommended standard232), 전력선 통신, 또는 POTS(plain old telephone service) 등 다양한 케이블 통신 모듈을 포함할 수 있다. 무선 통신 모듈은 와이파이(Wifi) 모듈, 와이브로(Wireless broadband) 모듈 외에도, GSM(global System for Mobile Communication), CDMA(Code Division Multiple Access), WCDMA(Wideband Code Division Multiple Access), UMTS(universal mobile telecommunications system), TDMA(Time Division

Multiple Access), LTE(Long Term Evolution), 4G, 5G, 6G 등 다양한 무선 통신 방식을 지원하는 무선 통신 모듈을 포함할 수 있다. 무선 통신 모듈은 신호를 송신하는 안테나 및 송신기(Transmitter)를 포함하는 무선 통신 인터페이스를 포함할 수 있다. 또한, 무선 통신 모듈은 프로세서(210)의 제어에 따라 무선 통신 인터페이스를 통해 프로세서(210)로부터 출력된 디지털 제어 신호를 아날로그 형태의 무선 신호로 변조하는 신호 변환 모듈을 더 포함할 수 있다. 근거리 통신 모듈은 근거리 통신(Short range communication)을 위한 것으로서, 블루투스(Bluetooth™), RFID(Radio Frequency Identification), 적외선 통신(Infrared Data Association; IrDA), UWB(Ultra Wideband), ZigBee, NFC(Near Field Communication), Wi-Fi(Wireless-Fidelity), Wi-Fi Direct, Wireless USB(Wireless Universal Serial Bus) 기술 중 적어도 하나를 이용하여, 근거리 통신을 지원할 수 있다.

[0053] 메모리(230)는 서버(200)의 다양한 기능을 지원하는 데이터와, 프로세서(210)의 동작을 위한 프로그램을 저장할 수 있고, 입/출력되는 데이터들(예를 들어, 음악 파일, 정지영상, 동영상 등)을 저장할 수 있고, 서버(200)에서 구동되는 다수의 응용 프로그램(application program 또는 애플리케이션(application)), 서버(200)의 동작을 위한 데이터들, 명령어들을 저장할 수 있다. 이러한 응용 프로그램 중 적어도 일부는, 무선 통신을 통해 외부 서버로부터 다운로드 될 수 있다. 이러한, 메모리(230)는 플래시 메모리 타입(flash memory type), 하드디스크 타입(hard disk type), SSD 타입(Solid State Disk type), HDD 타입(Hard Disk Drive type), 멀티미디어 카드 마이크로 타입(multimedia card micro type), 카드 타입의 메모리(예를 들어 SD 또는 XD 메모리(230) 등), 램(random access memory; RAM), SRAM(static random access memory), 롬(read-only memory; ROM), EEPROM(electrically erasable programmable read-only memory), PROM(programmable read-only memory), 자기 메모리, 자기 디스크 및 광디스크 중 적어도 하나의 타입의 저장매체를 포함할 수 있다.

[0054] 본 개시에 따른 인공지능과 관련된 기능은 프로세서(210)와 메모리(230)를 통해 동작된다. 프로세서(210)는 하나 또는 복수의 프로세서(210)들로 구성될 수 있다. 이때, 하나 또는 복수의 프로세서(210)들은 CPU, AP, DSP(Digital Signal Processor) 등과 같은 범용 프로세서(210), GPU, VPU(Vision Processing Unit)와 같은 그래픽 전용 프로세서(210) 또는 NPU와 같은 인공지능 전용 프로세서(210)일 수 있다. 하나 또는 복수의 프로세서(210)들은, 메모리(230)에 저장된 기 정의된 동작 규칙 또는 인공지능 모델에 따라, 입력 데이터를 처리하도록 제어한다. 또는, 하나 또는 복수의 프로세서(210)들이 인공지능 전용 프로세서(210)인 경우, 인공지능 전용 프로세서(210)는, 특정 인공지능 모델의 처리에 특화된 하드웨어 구조로 설계될 수 있다. 예를 들면, 프로세서(210)는, MCU(microcontroller unit), 팬 제어 액추에이터, APU(Accelerated Processing Unit) 등을 포함할 수 있다.

[0055] 기 정의된 동작 규칙 또는 인공지능 모델은 학습을 통해 만들어진 것을 특징으로 한다. 여기서, 학습을 통해 만들어진다는 것은, 기본 인공지능 모델이 학습 알고리즘에 의하여 다수의 학습 데이터들을 이용하여 학습됨으로써, 원하는 특성(또는, 목적)을 수행하도록 설정된 기 정의된 동작 규칙 또는 인공지능 모델이 만들어짐을 의미한다. 이러한 학습은 본 개시에 따른 인공지능이 수행되는 기기 자체에서 이루어질 수도 있고, 별도의 서버 및/또는 시스템을 통해 이루어 질 수도 있다. 학습 알고리즘의 예로는, 지도형 학습(supervised learning), 비지도형 학습(unsupervised learning), 준지도형 학습(semi-supervised learning) 또는 강화 학습(reinforcement learning)이 있으나, 전술한 예에 한정되지 않는다.

[0056] 인공지능 모델은, 복수의 신경망 레이어들로 구성될 수 있다. 복수의 신경망 레이어들 각각은 복수의 가중치들(weight values)을 갖고 있으며, 이전(previous) 레이어의 연산 결과와 복수의 가중치들 간의 연산을 통해 신경망 연산을 수행한다. 복수의 신경망 레이어들이 갖고 있는 복수의 가중치들은 인공지능 모델의 학습 결과에 의해 최적화될 수 있다. 예를 들어, 학습 과정 동안 인공지능 모델에서 획득한 로스(loss) 값 또는 코스트(cost) 값이 감소 또는 최소화되도록 복수의 가중치들이 갱신될 수 있다. 인공 신경망은 심층 신경망(DNN: Deep Neural Network)를 포함할 수 있으며, 예를 들어, CNN (Convolutional Neural Network), DNN (Deep Neural Network), RNN (Recurrent Neural Network), RBM (Restricted Boltzmann Machine), DBN (Deep Belief Network), BRDNN(Bidirectional Recurrent Deep Neural Network) 또는 심층 Q-네트워크 (Deep Q-Networks) 등이 있으나, 전술한 예에 한정되지 않는다.

[0057] 본 개시의 예시적인 실시예에 따르면, 프로세서(210)는 인공지능을 구현할 수 있다. 인공지능이란 사람의 신경 세포(biological neuron)를 모사하여 기계가 학습하도록 하는 인공신경망(Artificial Neural Network) 기반의 기계 학습법을 의미한다. 인공지능의 방법론에는 학습 방식에 따라 훈련데이터로서 입력데이터와 출력데이터가 같이 제공됨으로써 문제(입력데이터)의 해답(출력데이터)이 정해져 있는 지도학습(supervised learning), 및 출력데이터 없이 입력데이터만 제공되어 문제(입력데이터)의 해답(출력데이터)이 정해지지 않는 비지도학습

(unsupervised learning), 및 현재의 상태(State)에서 어떤 행동(Action)을 취할 때마다 외부 환경에서 보상(Reward)이 주어지는데, 이러한 보상을 최대화하는 방향으로 학습을 진행하는 강화학습(reinforcement learning)으로 구분될 수 있다. 또한, 인공지능의 방법론은 학습 모델의 구조인 아키텍처에 따라 구분될 수도 있는데, 널리 이용되는 딥러닝 기술의 아키텍처는, 합성곱신경망(CNN; Convolutional Neural Network), 순환신경망(RNN; Recurrent Neural Network), 트랜스포머(Transformer), 생성적 대립 신경망(GAN; generative adversarial networks) 등으로 구분될 수 있다.

[0058] 서버(200)는 인공지능 모델을 포함할 수 있다. 인공지능 모델은 하나의 인공지능 모델일 수 있고, 복수의 인공지능 모델로 구현될 수도 있다. 인공지능 모델은 뉴럴 네트워크(또는 인공 신경망)로 구성될 수 있으며, 기계학습과 인지과학에서 생물학의 신경을 모방한 통계학적 학습 알고리즘을 포함할 수 있다. 뉴럴 네트워크는 시냅스의 결합으로 네트워크를 형성한 인공 뉴런(노드)이 학습을 통해 시냅스의 결합 세기를 변화시켜, 문제 해결 능력을 가지는 모델 전반을 의미할 수 있다. 뉴럴 네트워크의 뉴런은 가중치 또는 바이어스의 조합을 포함할 수 있다. 뉴럴 네트워크는 하나 이상의 뉴런 또는 노드로 구성된 하나 이상의 레이어(layer)를 포함할 수 있다. 예시적으로, 장치는 input layer, hidden layer, output layer를 포함할 수 있다. 장치를 구성하는 뉴럴 네트워크는 뉴런의 가중치를 학습을 통해 변화시킴으로써 임의의 입력(input)으로부터 예측하고자 하는 결과(output)를 추론할 수 있다.

[0059] 프로세서(210)는 뉴럴 네트워크를 생성하거나, 뉴럴 네트워크를 훈련(train, 또는 학습(learn)하거나, 수신되는 입력 데이터를 기초로 연산을 수행하고, 수행 결과를 기초로 정보 신호(information signal)를 생성하거나, 뉴럴 네트워크를 재훈련(retrain)할 수 있다. 뉴럴 네트워크의 모델들은 GoogleNet, AlexNet, VGG Network 등과 같은 CNN(Convolution Neural Network), R-CNN(Region with Convolution Neural Network), RPN(Region Proposal Network), RNN(Recurrent Neural Network), S-DNN(Stacking-based deep Neural Network), S-SDNN(State-Space Dynamic Neural Network), Deconvolution Network, DBN(Deep Belief Network), RBM(Restrcted Boltzman Machine), Fully Convolutional Network, LSTM(Long Short-Term Memory) Network, Classification Network 등 다양한 종류의 모델들을 포함할 수 있으나 이에 제한되지는 않는다. 프로세서(210)는 뉴럴 네트워크의 모델들에 따른 연산을 수행하기 위한 하나 이상의 프로세서(210)를 포함할 수 있다. 예를 들어 뉴럴 네트워크는 심층 뉴럴 네트워크 (Deep Neural Network)를 포함할 수 있다.

[0060] 뉴럴 네트워크는 CNN(Convolutional Neural Network), RNN(Recurrent Neural Network), 퍼셉트론(perceptron), 다층 퍼셉트론(multilayer perceptron), FF(Feed Forward), RBF(Radial Basis Network), DFF(Deep Feed Forward), LSTM(Long Short Term Memory), GRU(Gated Recurrent Unit), AE(Auto Encoder), VAE(Variational Auto Encoder), DAE(Denoising Auto Encoder), SAE(Sparse Auto Encoder), MC(Markov Chain), HN(Hopfield Network), BM(Boltzmann Machine), RBM(Restricted Boltzmann Machine), DBN(Depp Belief Network), DCN(Deep Convolutional Network), DN(Deconvolutional Network), DCIGN(Deep Convolutional Inverse Graphics Network), GAN(Generative Adversarial Network), LSM(Liquid State Machine), ELM(Extreme Learning Machine), ESN(Echo State Network), DRN(Deep Residual Network), DNC(Differentiable Neural Computer), NTM(Neural Turning Machine), CN(Capsule Network), KN(Kohonen Network) 및 AN(Attention Network)를 포함할 수 있으나 이에 한정되는 것이 아닌 임의의 뉴럴 네트워크를 포함할 수 있음은 통상의 기술자가 이해할 것이다.

[0061] 본 개시의 예시적인 실시예에 따르면, 프로세서(210)는 GoogleNet, AlexNet, VGG Network 등과 같은 CNN(Convolution Neural Network), R-CNN(Region with Convolution Neural Network), RPN(Region Proposal Network), RNN(Recurrent Neural Network), S-DNN(Stacking-based deep Neural Network), S-SDNN(State-Space Dynamic Neural Network), Deconvolution Network, DBN(Deep Belief Network), RBM(Restrcted Boltzman Machine), Fully Convolutional Network, LSTM(Long Short-Term Memory) Network, Classification Network, Generative Modeling, eXplainable AI, Continual AI, Representation Learning, AI for Material Design, 자연어 처리를 위한 BERT, SP-BERT, MRC/QA, Text Analysis, Dialog System, GPT-3, GPT-4, 비전 처리를 위한 Visual Analytics, Visual Understanding, Video Synthesis, ResNet 데이터 지능을 위한 Anomaly Detection, Prediction, Time-Series Forecasting, Optimization, Recommendation, Data Creation 등 다양한 인공지능 구조 및 알고리즘을 이용할 수 있으며, 이에 제한되지 않는다.

[0062] 프로세서(210)는 전술한 인공지능을 구현할 수 있다. 인공지능의 방법론은 전술한 바와 같이, 지도학습, 및 비지도학습, 및 강화학습으로 구분될 수 있다. 딥러닝 기술의 아키텍처는, CNN, RNN, 트랜스포머(Transformer), GAN 등으로 구분될 수 있다. 인공지능 모델은 하나 이상의 인공지능 모델일 수 있다. 프로세서(210)는 전술한

바와 같이 뉴럴 네트워크를 생성하거나, 뉴럴 네트워크를 훈련(train, 또는 학습(learn))하거나, 수신되는 입력 데이터를 기초로 연산을 수행하고, 수행 결과를 기초로 정보 신호(information signal)를 생성하거나, 뉴럴 네트워크를 재훈련(retrain)할 수 있다. 뉴럴 네트워크는 CNN, RNN, 퍼셉트론, 다층 퍼셉트론 등을 포함할 수 있으나 이에 한정되는 것이 아닌 임의의 뉴럴 네트워크를 포함할 수 있음은 통상의 기술자가 이해할 것이다. 프로세서(210)는, 전술한 바와 같이, GoogleNet, AlexNet, VGG Network 등과 같은 CNN, R-CNN, 자연어 처리를 위한 BERT, SP-BERT, MRC/QA, Text Analysis, Dialog System, GPT-3, GPT-4, 비전 처리를 위한 Visual Analytics, Visual Understanding, Video Synthesis, ResNet 데이터 지능을 위한 Anomaly Detection, Prediction, Time-Series Forecasting, Optimization, Recommendation, Data Creation 등 다양한 인공지능 구조 및 알고리즘을 이용할 수 있으며, 이에 제한되지 않는다. CNN은 영상의 각 영역에 대해 복수의 필터를 적용하여 특징 지도 (Feature Map)를 만들어 내는 컨볼루션 층(Convolution Layer)과 특징 지도를 공간적으로 통합함으로써, 위치나 회전의 변화에 불변하는 특징을 추출할 수 있도록 하는 통합층(Pooling Layer)을 번갈아 수차례 반복하는 구조로 형성될 수 있다. 이를 통해, 점, 선, 면 등의 낮은 수준의 특징에서부터 복잡하고 의미 있는 높은 수준의 특징까지 다양한 수준의 특징을 추출해낼 수 있다. 컨볼루션 층은 입력 영상의 각 패치에 대하여 필터와 국지 수용장(Local Receptive Field)의 내적에 비선형 활성화 함수(Activation Function)를 취함으로써 특징 지도 (Feature Map)를 구할 수 있다. 다른 네트워크 구조와 비교하여, CNN은 희소한 연결성 (Sparse Connectivity)과 공유된 가중치(Shared Weights)를 가진 필터를 사용하는 특징을 가질 수 있다. 이러한 연결구조는 학습할 모수의 개수를 줄여주고, 역전파 알고리즘을 통한 학습을 효율적으로 만들어 결과적으로 예측 성능을 향상시킬 수 있다. 이와 같이, 컨볼루션 층과 통합 층의 반복을 통해 최종적으로 추출된 특징은 다중 신경망(MLP: Multi-Layer Perceptron)이나 서포트 벡터 머신(SVM: Support Vector Machine)과 같은 분류 모델이 완전 연결 층(Fully-connected Layer)의 형태로 결합되어 압축모델 학습 및 예측에 사용될 수 있다. 한편, 인공지능 기반의 문제 개발 모델은, 딥 러닝 기반으로 학습된 인공지능 모델을 의미할 수 있으며, 일 예로, CNN(Convolutional Neural Network)을 이용하여 학습된 모델을 의미할 수도 있다. 또한, 인공지능 기반의 문제 개발 모델은, Natural Language Processing(NLP), Random Forest (RF), Support Vector Machine (SVC), eXtra Gradient Boost (XGB), Decision Tree (DC), Knearest Neighbors (KNN), Gaussian Naive Bayes (GNB), Stochastic Gradient Descent (SGD), Linear Discriminant Analysis (LDA), Ridge, Lasso 및 Elastic net 중 적어도 하나의 알고리즘을 포함할 수도 있다.

- [0063] 도 3은 본 개시에 따른 방법을 나타낸 순서도이다.
- [0064] 도 3을 참조하면, 도 3의 방법은, 스크립트 정보 생성 단계(S100), 발화량 정보 생성 단계(S200), 발화량 비교 단계(S300), 및 학습자 언어 학습 숙련도 평가 단계(S400)를 포함할 수 있다.
- [0065] 스크립트 정보 생성 단계(S100)에서, 프로세서(210)는 제1 음성 정보를 기초로 제1 스크립트 정보를 생성하고, 제2 음성 정보를 기초로 제2 스크립트 정보를 생성할 수 있다.
- [0066] 발화량 정보 생성 단계(S200)에서, 프로세서(210)는 제1 스크립트 정보를 기초로 제1 사용자 단말을 이용해 소리 내어 말을 하는 교육자의 언어 행위의 총량을 나타내는 제1 발화량 정보를 생성하고, 제2 스크립트 정보를 기초로 제2 사용자 단말을 이용해 소리 내어 말을 하는 학습자의 언어 행위의 총량을 나타내는 제2 발화량 정보를 생성할 수 있다.
- [0067] 발화량 비교 단계(S300)에서, 프로세서(210)는 제1 발화량 정보의 제1 발화량과 제2 발화량 정보의 제2 발화량을 비교할 수 있다. 만약, 프로세서(210)는, 제2 발화량이 제1 발화량보다 크거나 같은 경우에, 학습이 원활하게 진행되고 있는 것으로 판단하여 다음 단계를 수행할 수 있다. 만약, 제2 발화량이 제1 발화량보다 작은 경우에, 제2 발화량이 제1 발화량 이상이 되도록, 프로세서(210)는 학습자의 참여를 유도하도록 다양한 월드를 구성할 수 있다.
- [0068] 학습자 언어 학습 숙련도 평가 단계(S400)에서, 프로세서(210)는 제2 스크립트 정보를 기초로 학습자의 언어 학습에 대한 숙련도를 평가할 수 있다. 예를 들면, 프로세서(210)는 제2 스크립트 정보에서 서로 다른 종류의 단어의 개수 및 빈도수를 측정하고, 제2 스크립트 정보에서 기존에 발생한 언어 표현과 다른 언어 표현의 발생량을 측정할 수 있다. 그리고, 프로세서(210)는 단어의 개수 및 단어의 빈도수, 언어 표현의 발생량을 이용하여 숙련도를 초급, 중급, 고급, 또는 원어민으로 평가할 수 있다.
- [0069] 도 4는 본 개시의 예시적인 실시예들을 실행하기 위한 웹화면을 예시적으로 나타낸 도면이고, 도 5 및 도 6은 본 개시의 예시적인 실시예를 실행한 화면을 예시적으로 나타낸 도면들이다.

[0070] 도 4를 참조하면, 본 개시의 서버(200)에서 제공하는 어플리케이션은, PLanguage라는 명칭으로 제공될 수 있다. 본 개시의 서버(200)에서 제공하는 어플리케이션은, 모국어 취득 환경 플랫폼으로서, 학습자가 교육자 캐릭터의 행동을 보고 교육자가 말하는 영어 표현을 익히고, 학습자 캐릭터가 행동을 하면 교육자가 그 행동에 대한 영어 문장을 음성으로 말하며, 학습자가 영어로 감정을 느끼고 교감할 수 있도록 월드를 구현할 수 있다.

[0071] 학습자가 교육자 캐릭터의 행동을 보고 교육자가 말하는 영어 표현을 익히는 학습의 경우, 도 5 및 도 6을 참조하면, 교육자 캐릭터(예, GD)와 학습자 캐릭터(예, Thor)가 가상 공간 상에서 복수의 양들을 포함하는 월드(예, 양떼 목장)에서 만나고, 교육자 캐릭터와 학습자 캐릭터가 양떼 목장 안으로 진입하고, 교육자 캐릭터가 학습자 캐릭터에게 가상의 아이템(예, Shear)을 제공할 수 있다. 학습자 캐릭터는 가상의 아이템의 의미를 모를 수 있다. 하지만, 교육자 캐릭터가 가상 공간에서 양털을 깎는 것을 보여줄 수 있다. 이 경우, 학습자 캐릭터를 조종하는 학습자는 가상의 아이템의 의미와 활용 표현을 습득할 수 있다. 교육자 캐릭터를 조종하는 교육자는 가상의 아이템의 의미와 관련된 활용 표현을 음성으로 제공하면, 해당 음성에 대응되는 스크립트 정보가 시각적으로 표시될 수 있다.

[0072] 학습자 캐릭터가 행동을 하면 교육자가 그 행동에 대한 영어 문장을 음성으로 말하는 학습의 경우, 교육자 캐릭터(예, GD)와 학습자 캐릭터(예, Thor)가 가상 공간 상에서 월드(예, 창고)에서 만나고, 학습자 캐릭터는 교육자 캐릭터의 말을 듣고 창고의 상자 안에 있는 오브젝트(예, 사다리)를 꺼낼 수 있다. 교육자 캐릭터는 학습자 캐릭터에게 사다리를 이용하여 벽을 올라가자고 할 수 있다. 교육자 캐릭터는 사다리를 벽에 설치하고 학습자 캐릭터에게도 해 보라고 권유할 수 있다. 학습자 캐릭터도 교육자 캐릭터를 따라 사다리를 벽에 설치할 수 있다. 이때, 교육자 캐릭터는 학습자 캐릭터의 행동(예, 사다리를 벽에 설치하는 행동)에 대한 영어 표현을 음성으로 제공하고, 이에 따라 스크립트 정보와 음성 정보가 시각적 및 청각적으로 출력될 수 있다.

[0073] 학습자가 영어로 감정을 느끼고 교감할 수 있는 학습의 경우, 교육자 캐릭터(예, GD)와 학습자 캐릭터(예, Thor)가 가상 공간 상에서 가파른 경사를 포함하는 월드(예, 산)에서 만나고, 학습자 캐릭터가 교육자 캐릭터의 행동을 따라서 가파른 경사길을 오를 수 있다. 가상 공간에서 가상의 해가 지고 날이 어두워지며 비가 올 수 있다. 이때, 가상의 오브젝트(예, 좀비 캐릭터)가 등장할 수 있다. 주변 분위기에 학습자 캐릭터와 교육자 캐릭터가 등골이 오싹해지는 느낌을 경험할 수 있다. 교육자 캐릭터가 학습자 캐릭터에게 "you feel nervous"라고 질문할 수 있다. 감정과 상태 표현은 사전적 의미를 외우는 것이 아니라, 직접 경험하여 체득해야 한다. 학습자 캐릭터는 직접 느낀 감정을 영어로 표현할 수 있다.

[0076] 이상 설명된 본 개시에 따른 실시예는 다양한 컴퓨터 구성요소를 통하여 실행될 수 있는 프로그램 명령어의 형태로 구현되어 컴퓨터 판독 가능한 기록 매체에 기록될 수 있다. 컴퓨터 판독 가능한 기록 매체는 프로그램 명령어, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 컴퓨터 판독 가능한 기록 매체에 기록되는 프로그램 명령어는 본 개시를 위하여 특별히 설계되고 구성된 것이거나 컴퓨터 소프트웨어 분야의 당업자에게 공지되어 사용 가능한 것일 수 있다. 컴퓨터 판독 가능한 기록 매체의 예에는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM 및 DVD와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical medium), 및 ROM, RAM, 플래시 메모리 등과 같은, 프로그램 명령어를 저장하고 실행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령어의 예에는, 컴파일러에 의하여 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용하여 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드도 포함된다. 하드웨어 장치는 본 개시에 따른 처리를 수행하기 위하여 하나 이상의 소프트웨어 모듈로 변경될 수 있으며, 그 역도 마찬가지이다.

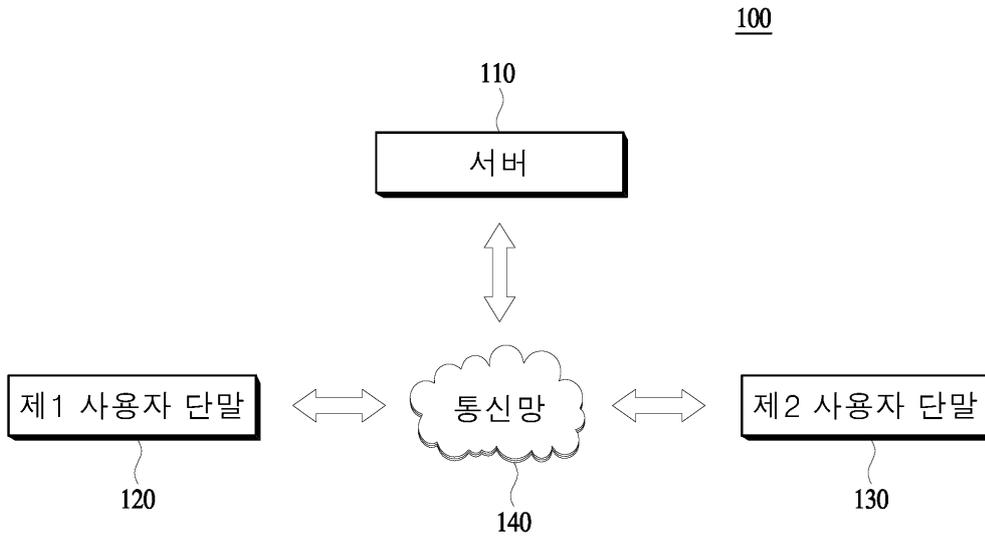
[0077] 본 개시가 속하는 기술분야에서 통상의 지식을 가진 자는 본 개시의 기술적 사상이나 필수적인 특징을 변경하지 않고도, 개시된 실시예들과 다른 형태로 본 개시가 실시될 수 있음을 이해할 것이다. 개시된 실시예들은 예시적인 것이며, 한정적으로 해석되어서는 안 된다.

부호의 설명

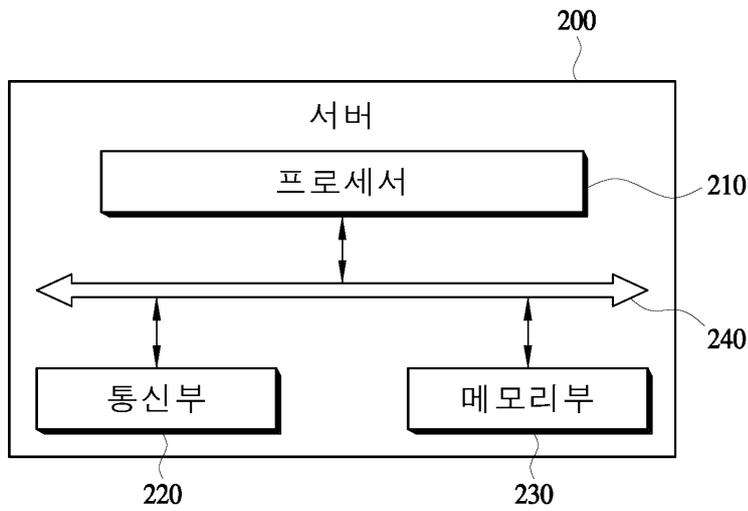
- [0078] 100: 시스템
- 110: 서버
- 120: 제1 사용자 단말
- 130: 제2 사용자 단말

도면

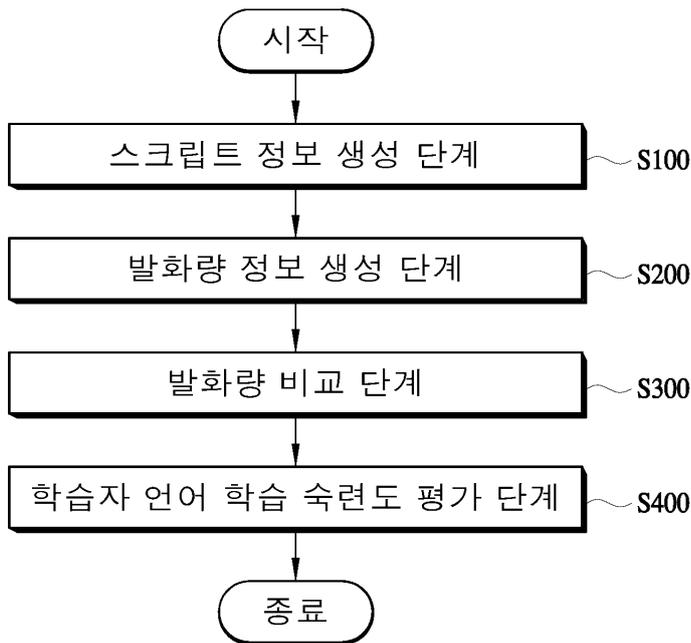
도면1



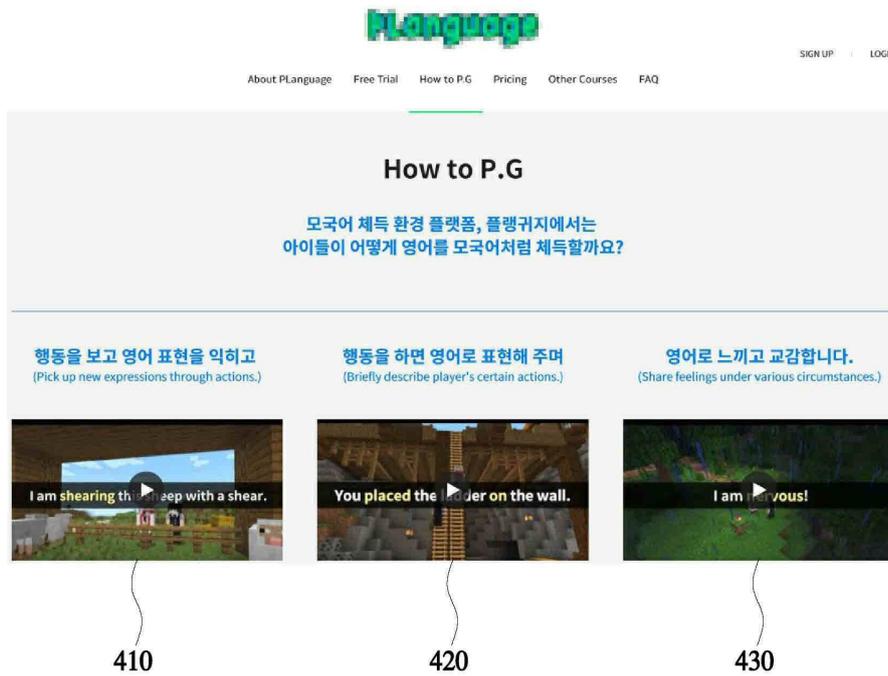
도면2



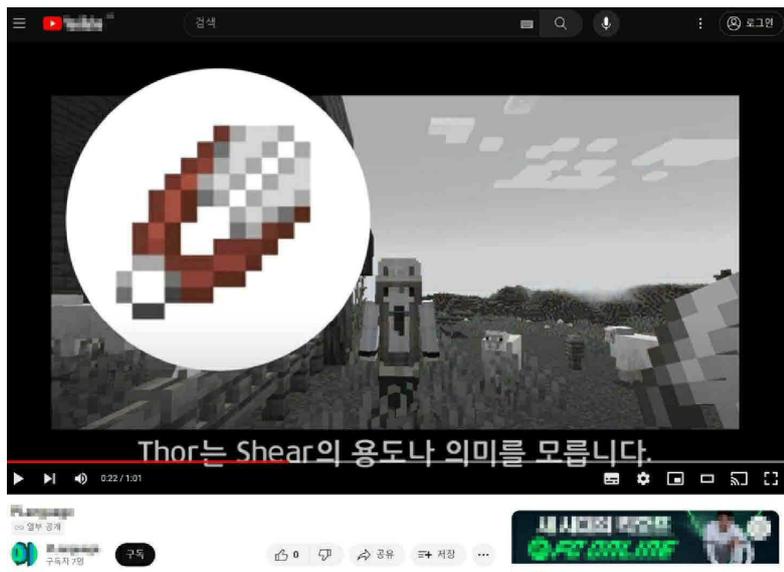
도면3



도면4



도면5



도면6

