



(51) 국제특허분류:

G10L 15/16 (2006.01)

G06T 1/00 (2006.01)

G10L 15/22 (2006.01)

G06T 7/00 (2006.01)

(21) 국제출원번호:

PCT/KR2018/007410

(22) 국제출원일:

2018년 6월 29일 (29.06.2018)

(25) 출원언어:

한국어

(26) 공개언어:

한국어

(30) 우선권정보:

10-2017-0109461 2017년 8월 29일 (29.08.2017) KR

(71) 출원인: 서울대학교 산학협력단 (SEOUL NATIONAL UNIVERSITY R&DB FOUNDATION) [KR/KR]; 08826 서울시 관악구 관악로 1(신림동), Seoul (KR).

(72) 발명자: 장병탁 (ZHANG, Byoung-Tak); 06603 서울시 서초구 서초중앙로24길 33, 103동 501호, Seoul (KR). 한철호 (HAN, Cheolho); 08826 서울시 관악구 관악로 1, 서울대학교 918동 522호, Seoul (KR). 허유정 (HEO, Yu-Jung); 14306 경기도 광명시 하안로 284, 1213동 1108호, Gyeonggi-do (KR). 강우영 (KANG, Wooyoung); 08783 서울시 관악구 죽고개로 9, 711호, Seoul (KR). 전재현 (JUN, Jaehyun); 06603 서울시 서초구 서초중앙로24길 43, 103동 1203호, Seoul (KR).

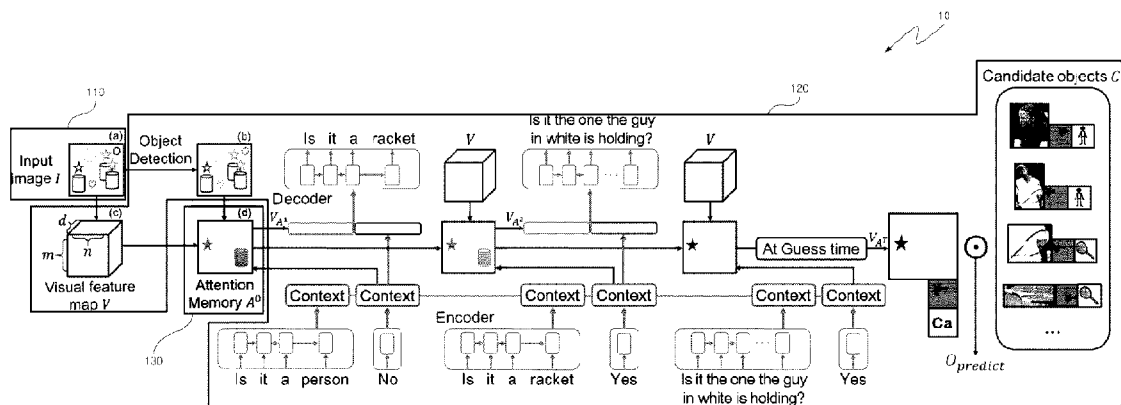
(74) 대리인: 특허법인 아이스퀘어 (ISQUARE PATENT & LAW FIRM); 06174 서울시 강남구 영동대로82길 11, 5층(대치동, 도원빌딩), Seoul (KR).

(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(54) Title: ATTENTIVE MEMORY METHOD AND SYSTEM FOR LOCATING OBJECT THROUGH VISUAL DIALOGUE

(54) 발명의 명칭: 시각 대화를 통해 객체의 위치를 알아내기 위한 주의 기억 방법 및 시스템



(57) Abstract: The present invention presents an attentive memory method and system for locating an object through visual dialogue. The attentive memory system for identifying an object in an image includes: a control unit which generates a question for identifying a preset object in the image, deduces an answer to the generated question, and identifies the preset object in the image on the basis of the question and answer; and memory for storing the image, wherein the control unit can generate the question by taking into account information about objects contained in the image, and update the information about the objects on the basis of the answer.

(57) 요약서: 시각 대화를 통해 객체의 위치를 알아내기 위한 주의 기억 방법 및 시스템을 제시하며, 이미지 상의 객체를 식별하는 주의기억시스템에 있어서, 상기 이미지 상에서 기 설정된 객체를 식별하기 위한 질문을 생성하고, 생성된 질문에 대해 답변을 도출하고, 상기 질문 및 상기 답변에 기초하여 상기 이미지에서 기 설정된 객체를 식별하는 제어부 및 상기 이미지를 저장하는 메모리를 포함하며, 상기 제어부는, 상기 이미지에 포함된 객체들에 대한 정보를 반영하여 상기 질문을 생성하고, 상기 답변에 기초하여 상기 객체들에 대한 정보를 갱신할 수 있다.

규칙 4.17에 의한 선언서:

- 신규성을 해치지 아니하는 개시 또는 신규성 상실의 예
외에 관한 선언 (규칙 4.17(v))

공개:

- 국제조사보고서와 함께 (조약 제21조(3))

명세서

발명의 명칭: 시각 대화를 통해 객체의 위치를 알아내기 위한 주의 기억 방법 및 시스템

기술분야

- [1] 본 명세서에서 개시되는 실시예들은 시각 대화를 통해 객체의 위치를 알아내기 위한 주의 기억 방법 및 시스템에 관한 것으로, 보다 상세하게는 인공지능으로 구현된 질문자 및 답변자 사이의 대화를 통해 이미지상의 객체를 식별하는 방법에서 이미지상의 객체의 위치정보를 저장하는 주의 기억을 이용한 효율적인 질문을 생성하도록 하는 주의 기억 방법 및 시스템에 관한 것이다.
- [2] 2018년 과제번호 및 사사표기
- [3] 1-1. 과제고유번호: 1711065242
- [4] 1-2. 사사 국문표기: 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW컴퓨팅산업원천기술개발사업(SW스타랩)의 연구결과로 수행되었음" (IITP-2018-1711065242)
- [5] 1-3. 사사 영문표기: This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the SW Starlab support program (IITP-2018-1711065242) supervised by the IITP(Institute for Information & communications Technology Promotion)
- [6] 1-4. 기여율: 1/2
- [7]
- [8] 2-1. 과제고유번호: 1415155958
- [9] 2-2. 사사 국문표기: 이 연구는 2018년도 산업통상자원부 및 산업기술평가관리원(KEIT) 연구비 지원에 의한 연구임('1415155958')
- [10] 2-3. 사사 영문표기: This work was supported by the Technology Innovation Program (or Industrial Strategic Technology Development Program) (1415155958 , A robot intelligence software framework as an open and self-growing integration foundation of intelligence and knowledge for personal service robots) funded By the Ministry of Trade & Energy(MOTIE, Korea)
- [11] 2-4. 기여율: 1/2

배경기술

- [12] 뉴럴네트워크는 생물 신경계의 고도의 정보처리기구를 공학적으로 모방해서 입력과 출력을 상호간에 상세히 관련지어 복잡한 제어를 하는 정보처리기술로 스위치나 각 센서 등에서 신호를 보내는 입력 레이어, 그 정보를 바탕으로 입력과 출력의 우선순위를 정하면서 상호관계를 조정하는 은닉 레이어, 이것을 바탕으로 필요한 제어량을 산출해서 출력하는 출력 레이어 3종류의 뉴론(신경세포) 모델이 복수로 얹혀서 네트워크를 구성하고 있다.

- [13] 이러한 뉴럴네트워크는 이미지상의 객체를 식별하는데 이용되고 있으며, 객체를 효율적으로 인식하기 위한 다양한 기술에 응용되고 있다.
- [14] 이와 같이 뉴럴네트워크를 이용한 이미지의 객체를 인식하는 방법 중 스무고개와 같이 질문자와 답변자 역할을 하는 두 개의 플레이어가 질문과 답변을 반복함으로써 이미지에서 기 설정된 객체를 식별하는 방법이 있다.
- [15] 즉, 질문자는 주어진 이미지에 기 설정된 객체를 예측하기 위한 제 1 질문을 생성하여 답변자에게 제공하고, 질문에 대한 답변자의 답변과 제 1 질문을 기초로 제 2 질문을 생성하는 과정을 반복하여 이미지상의 객체 중 기 설정된 객체를 선택한다.
- [16] 하지만, 질문자는 질문을 생성할 때, 자신이 한 질문과 답변자의 답변만을 기초로 질문을 생성함에 따라 객체를 식별하는데 중요한 정보를 포함하는 효율적인 질문을 생성하지 못한다는 문제점이 있다.
- [17] 관련하여 선행기술 문헌인 한국특허공개번호 제10-2017-0033398호에서는 발화 인식 시스템을 위해 신경망을 구현하는 것에 관련된 기법으로 발화 디코더에 의해 필요로 되는 출력만이 신경망 또는 근사계산 기법을 통하여 제공되도록, 근사계산된 스킵 프레임에 이용한 프레임 스킵핑 및 수요거리를 계산할 뿐, 상술된 바와 같이 대화 시스템에서 효율적인 질문을 생성하지 못한다.
- [18] 따라서 상술된 문제점을 해결하기 위한 기술이 필요하게 되었다.
- [19] 한편, 전술한 배경기술은 발명자가 본 발명의 도출을 위해 보유하고 있었거나, 본 발명의 도출 과정에서 습득한 기술 정보로서, 반드시 본 발명의 출원 전에 일반 공중에게 공개된 공지기술이라 할 수는 없다.

발명의 상세한 설명

기술적 과제

- [20] 본 명세서에서 개시되는 실시예들은, 시각 대화를 통해 객체의 위치를 알아내기 위한 주의 기억 방법 및 시스템을 제시하는데 목적이 있다.
- [21] 또한, 본 명세서에서 개시되는 실시예들은, 주의 기억에 저장된 이미지상의 객체의 위치를 이용하여 질문을 생성하는 주의 기억 방법 및 시스템을 제시하는데 목적이 있다.
- [22] 아울러 본 명세서에서 개시되는 실시예들은, 대화를 통해 주의 기억에 저장된 이미상의 객체의 위치를 갱신함으로써 효율적인 질문을 생성하는 주의 기억 방법 및 시스템을 제시하는 데 목적이 있다.

과제 해결 수단

- [23] 상술한 기술적 과제를 달성하기 위한 기술적 해결방법은, 이미지상의 객체를 식별하기 위해 질문을 생성하고 생성된 질문에 대한 답변에 기초하여 이미지상의 객체를 식별한다.

발명의 효과

- [24] 전술한 과제 해결 수단 중 어느 하나에 의하면, 시각 대화를 통해 객체의 위치를 알아내기 위한 주의 기억 방법 및 시스템을 제시할 수 있다.
- [25] 또한, 전술한 과제 해결 수단 중 어느 하나에 의하면, 주의 기억에 저장된 이미지상의 객체의 위치를 이용하여 질문을 생성하는 주의 기억 방법 및 시스템을 제시할 수 있다.
- [26] 아울러 본 명세서에서 개시되는 실시예들은, 대화를 통해 주의 기억에 저장된 이미지 상의 객체의 위치를 갱신함으로써 효율적인 질문을 생성하는 주의 기억 방법 및 시스템을 제시할 수 있다.
- [27] 개시되는 실시예들에서 얻을 수 있는 효과는 이상에서 언급한 효과들로 제한되지 않으며, 언급하지 않은 또 다른 효과들은 아래의 기재로부터 개시되는 실시예들이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

- [28] 도 1은 일 실시예에 따른 주의기억시스템을 도시한 블록도이다.
- [29] 도 2는 일 실시예에 따른 주의기억방법을 설명하기 위한 순서도이다.

발명의 실시를 위한 최선의 형태

- [30] 상술한 기술적 과제를 달성하기 위한 기술적 수단으로서, 일 실시예에 따르면, 이미지 상의 객체를 식별하는 주의기억시스템에 있어서, 상기 이미지 상에서 기 설정된 객체를 식별하기 위한 질문을 생성하고, 생성된 질문에 대해 답변을 도출하고, 상기 질문 및 상기 답변에 기초하여 상기 이미지에서 기 설정된 객체를 식별하는 제어부 및 상기 이미지를 저장하는 메모리를 포함하며, 상기 제어부는, 상기 이미지에 포함된 객체들에 대한 정보를 반영하여 상기 질문을 생성하고, 상기 답변에 기초하여 상기 객체들에 대한 정보를 갱신할 수 있다.
- [31] 다른 실시예에 따르면, 주의기억시스템이 이미지 상의 객체를 식별하는 주의기억방법에 있어서, 상기 이미지를 입력받는 단계, 상기 이미지에 포함된 객체들에 대한 정보를 반영하여 기 설정된 객체를 식별하기 위한 질문을 생성하고, 생성된 질문에 대한 답변을 도출하는 단계 및 적어도 하나의 질문 및 상기 질문 및 상기 답변에 기초하여 상기 이미지에서 기 설정된 객체를 식별하는 단계를 포함하고, 상기 주의기억방법은, 상기 답변에 기초하여 상기 객체들에 대한 정보를 갱신하는 단계를 포함할 수 있다.
- [32] 다른 실시예에 따르면, 주의기억방법을 수행하는 프로그램이 기록된 컴퓨터 판독이 가능한 기록매체로서, 상기 이미지를 입력받는 단계, 상기 이미지에 포함된 객체들에 대한 정보를 반영하여 기 설정된 객체를 식별하기 위한 질문을 생성하고, 생성된 질문에 대한 답변을 도출하는 단계 및 적어도 하나의 질문 및 상기 질문 및 상기 답변에 기초하여 상기 이미지에서 기 설정된 객체를 식별하는 단계를 포함하고, 상기 주의기억방법은, 상기 답변에 기초하여 상기 객체들에 대한 정보를 갱신하는 단계를 포함할 수 있다.

- [33] 다른 실시예에 따르면, 주의기억시스템에 의해 수행되며, 주의기억방법을 수행하기 위해 기록매체에 저장된 컴퓨터프로그램으로서, 상기 이미지를 입력받는 단계, 상기 이미지에 포함된 객체들에 대한 정보를 반영하여 기 설정된 객체를 식별하기 위한 질문을 생성하고, 생성된 질문에 대한 답변을 도출하는 단계 및 적어도 하나의 질문 및 상기 질문 및 상기 답변에 기초하여 상기 이미지에서 기 설정된 객체를 식별하는 단계를 포함하고, 상기 주의기억방법은, 상기 답변에 기초하여 상기 객체들에 대한 정보를 갱신하는 단계를 포함할 수 있다.

발명의 실시를 위한 형태

- [34] 아래에서는 첨부한 도면을 참조하여 다양한 실시예들을 상세히 설명한다. 아래에서 설명되는 실시예들은 여러 가지 상이한 형태로 변형되어 실시될 수도 있다. 실시예들의 특징을 보다 명확히 설명하기 위하여, 이하의 실시예들이 속하는 기술분야에서 통상의 지식을 가진 자에게 널리 알려져 있는 사항들에 관해서 자세한 설명은 생략하였다. 그리고, 도면에서 실시예들의 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.
- [35] 명세서 전체에서, 어떤 구성이 다른 구성과 "연결"되어 있다고 할 때, 이는 '직접적으로 연결'되어 있는 경우뿐 아니라, '그 중간에 다른 구성을 사이에 두고 연결'되어 있는 경우도 포함한다. 또한, 어떤 구성이 어떤 구성을 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한, 그 외 다른 구성을 제외하는 것이 아니라 다른 구성들을 더 포함할 수도 있음을 의미한다.
- [36] 이하 첨부된 도면을 참고하여 실시예들을 상세히 설명하기로 한다.
- [37] 다만 이를 설명하기에 앞서, 아래에서 사용되는 용어들의 의미를 먼저 정의한다.
- [38] 이하에서 '시각적특징맵(visual feature map)'은 입력된 이미지로부터 이미지에 포함된 객체에 대한 특징을 나타내는 값의 배열로 예를 들어, 이미지에 포함된 객체의 모서리에 대한 정보를 추출한 값일 수 있다.
- [39] 그리고 '주의기억특징맵'은 이미지상에서 기 설정된 객체로 선택될 수 있는 후보 객체에 대한 특징을 나타내는 값으로 시각적특징맵과 이미지에서 식별된 적어도 하나의 객체에 대한 위치정보에 기초하여 생성된 정보이다.
- [40] 위에 정의한 용어 이외에 설명이 필요한 용어는 아래에서 각각 따로 설명한다.
- [41] 도 1은 일 실시예에 따른 주의기억시스템(10)을 설명하기 위한 블록도이다.
- [42] 주의기억시스템(10)은 이미지상의 기 설정된 객체에 대한 정보를 알고 있는 답변자와 이미지에서 적어도 하나의 객체 중 기 설정된 객체에 대한 정보를 획득하기 위해 질문을 생성하는 질문자 그리고 질문자와 답변자 간의 대화에 기초하여 이미지 상의 적어도 하나의 객체 중 기 설정된 객체를 선택하는 추측자 역할을 수행하여, 질문자와 답변자 간의 대화를 통해 획득된 정보를 기초로

추측자가 기 설정된 객체를 선택합니다.

- [43] 이러한 주의기억시스템(10)는 서버-클라이언트 시스템으로 구현될 수 있으며, 클라이언트는 전자단말기에 설치되어 서버와 통신할 수 있다.
- [44] 이때, 전자단말기는 네트워크(N)를 통해 원격지의 서버에 접속하거나, 타 단말 및 서버와 연결 가능한 컴퓨터나 휴대용 단말기, 텔레비전, 웨어러블 디바이스(Wearable Device) 등으로 구현될 수 있다. 여기서, 컴퓨터는 예를 들어, 웹 브라우저(WEB Browser)가 탑재된 노트북, 데스크톱(desktop), 랩톱(laptop)등을 포함하고, 휴대용 단말기는 예를 들어, 휴대성과 이동성이 보장되는 무선 통신 장치로서, PCS(Personal Communication System), PDC(Personal Digital Cellular), PHS(Personal Handyphone System), PDA(Personal Digital Assistant), GSM(Global System for Mobile communications), IMT(International Mobile Telecommunication)-2000, CDMA(Code Division Multiple Access)-2000, W-CDMA(W-Code Division Multiple Access), Wibro(Wireless Broadband Internet), 스마트폰(Smart Phone), 모바일 WiMAX(Mobile Worldwide Interoperability for Microwave Access) 등과 같은 모든 종류의 핸드헬드(Handheld) 기반의 무선 통신 장치를 포함할 수 있다. 또한, 텔레비전은 IPTV(Internet Protocol Television), 인터넷 TV(Internet Television), 지상파 TV, 케이블 TV 등을 포함할 수 있다. 나아가 웨어러블 디바이스는 예를 들어, 시계, 안경, 액세서리, 의복, 신발 등 인체에 직접 착용 가능한 타입의 정보처리장치로서, 직접 또는 다른 정보처리장치를 통해 네트워크를 경유하여 원격지의 서버에 접속하거나 타 단말과 연결될 수 있다.
- [45] 그리고 서버는 전자단말기와 네트워크(N)를 통해 통신이 가능한 컴퓨터로 구현될 수 있으며, 데이터를 저장할 수 있는 저장장치가 포함될 수 있거나 또는 제 3의 서버(미도시)를 통해 데이터를 저장할 수도 있다.
- [46] 이와 같이 본 발명의 일실시예에 따른, 주의기억시스템(10)은 입출력부(110), 제어부(120), 통신부(130) 및 메모리(130)를 포함할 수 있다.
- [47] 입출력부(110)는 사용자로부터 입력을 수신하기 위한 입력부와, 작업의 수행 결과 또는 주의기억시스템(10)의 상태 등의 정보를 표시하기 위한 출력부를 포함할 수 있다. 예를 들어, 입출력부(110)는 사용자 입력을 수신하는 조작 패널(operation panel) 및 화면을 표시하는 디스플레이 패널(display panel) 등을 포함할 수 있다.
- [48] 구체적으로, 입력부는 키보드, 물리 버튼, 터치 스크린, 카메라 또는 마이크 등과 같이 다양한 형태의 사용자 입력을 수신할 수 있는 장치들을 포함할 수 있다. 또한, 출력부는 디스플레이 패널 또는 스피커 등을 포함할 수 있다. 다만, 이에 한정되지 않고 입출력부(110)는 다양한 입출력을 지원하는 구성을 포함할 수 있다.
- [49] 제어부(120)는 주의기억시스템(10)의 전체적인 동작을 제어하며, CPU 등과 같은 프로세서를 포함할 수 있다. 제어부(120)는 입출력부(110)를 통해 수신한

사용자 입력에 대응되는 동작을 수행하도록 주의기억시스템(10)에 포함된 다른 구성들을 제어할 수 있다.

- [50] 예를 들어, 제어부(120)는 메모리(130)에 저장된 프로그램을 실행시키거나, 메모리(130)에 저장된 파일을 읽어오거나, 새로운 파일을 메모리(130)에 저장할 수도 있다.
- [51] 이러한 제어부(120)는 입출력부(110)를 통해 입력된 이미지에 대해 이미지에 포함된 적어도 하나의 객체를 식별할 수 있다.
- [52] 예를 들어, 제어부(120)는 이미지에 대해 YOLO(You only look once)라는 객체검출(Object detection)모델을 통해 적어도 하나의 객체를 식별할 수 있다.
- [53] 그리고 제어부(120)는 식별된 적어도 하나의 객체 각각의 위치를 추출할 수 있고, 추출된 적어도 하나의 객체 각각의 위치정보를 기초로 후술할 메모리(130)상의 주의기억(Attention memory)을 초기화할 수 있다.
- [54] 예를 들어, 제어부(120)는 메모리(130)에 일정한 영역을 주의기억으로 할당할 수 있고, 할당된 주의기억에 이미지에서 식별된 적어도 하나의 객체 각각에 대해 이미지상에서 상대적인 위치에 대한 위치정보를 추출하여 저장할 수 있다.
- [55] 이후, 제어부(120)는 이미지에 포함된 적어도 하나의 객체에 대한 특징을 포함하는 시각적특징맵(visual feature map)을 추출할 수 있다.
- [56] 예를 들어, 제어부(120)는 CAEs(Convolutional AutoEncoders)를 통해 이미지로부터 시각적특징맵을 추출할 수 있다.
- [57] 그리고 제어부(120)는 이미지 상에서 기 설정된 객체를 식별하기 위한 질문을 생성할 수 있다.
- [58] 이를 위해 제어부(120)는 실시예에 따라 질문을 구성할 수 있도록 적어도 하나의 단어를 기 저장하거나 또는 제 3의 서버를 통해 적어도 하나의 단어를 획득할 수 있다.
- [59] 우선 제어부(120)는 대화를 시작하기 위해 시작 질문을 생성할 수 있고, 시작 질문에 대한 답변을 수행할 수 있다.
- [60] 예를 들어, 제어부(120)는 대화를 시작하기 위해 “사람입니까?”라는 질문을 랜덤하게 생성할 수 있고, 생성된 질문에 대해 “아니오”라는 답변을 수행할 수 있다.
- [61] 그리고 제어부(120)는 시작 질문에 대한 답변과 시각적특징맵을 기초로 주의기억에 저장된 주의기억정보를 갱신할 수 있다.
- [62] 예를 들어, 제어부(120)는 “사람입니까?”라는 질문에 대한 답변 “아니오”를 기초로 주의기억에 저장된 주의기억정보인 이미지상의 적어도 하나의 객체의 위치 중 이미지상의 사람에 대한 정보를 삭제할 수 있다.
- [63] 이후, 제어부(120)는 주의기억정보와 시각적특징맵을 기초로 주의기억특징맵(feature map with attention memory value)을 생성할 수 있고, 생성된 주의기억특징맵과 질문에 대한 답변을 기초로 질문을 생성할 수 있다.
- [64] 즉, 제어부(120)는 이미지상의 객체에 대한 특징 정보를 포함하는

시각적특징맵과 후보 객체에 대한 정보를 포함하는 주의기억정보를 기초로 이미지상의 적어도 하나의 객체 중 후보 객체와 관련된 질문을 생성할 수 있다.

[65] 예를 들어, 제어부(120)는 주의기억정보에서 “사람”을 제외한 나머지 “라켓, 공”에 대한 위치를 기초로 시각적특징맵에서 “라켓, 공”에 대한 시각적특징맵을 추출하여 주의기억특징맵을 생성할 수 있고, 생성된 주의기억특징맵을 기초로 “라켓” 또는 “공”과 관련된 질문을 생성할 수 있다.

[66] 그리고 제어부(120)는 새롭게 생성된 질문에 대한 답변을 도출할 수 있고, 도출된 답변과 시각적특징맵을 기초로 주의기억정보를 갱신하여 후보 객체의 수를 줄일 수 있다.

[67] 이후, 제어부(120)는 상술된 바와 같이 갱신된 주의기억정보와 시각적특징맵을 이용하여 후보 객체의 수가 줄어든 주의기억특징맵을 생성하여 질문을 생성할 수 있고, 생성된 질문에 대한 답변을 도출하여 주의기억정보를 갱신하는 과정을 반복할 수 있다.

[68] 그리고 제어부(120)는 주의기억에 저장된 주의기억정보에 기초하여 이미지상의 적어도 하나의 객체 중 기 설정된 객체를 선택할 수 있다.

[69] 예를 들어, 제어부(120)는 주의기억정보, 이미지의 공간 정보 및 카테고리별 예시 객체와 비교하여 기 설정된 객체를 선택할 수 있다.

[70] 메모리(130)에는 파일, 어플리케이션 및 프로그램 등과 같은 다양한 종류의 데이터가 설치 및 저장될 수 있다. 제어부(120)는 메모리(130)에 저장된 데이터에 접근하여 이를 이용하거나, 또는 새로운 데이터를 메모리(130)에 저장할 수도 있다. 또한, 제어부(120)는 메모리(130)에 설치된 프로그램을 실행할 수도 있다. 도 1을 참조하면, 메모리(130)에는 주의기억방법을 수행하기 위한 프로그램이 설치될 수 있다.

[71] 이러한 메모리(130)는 입출력부(110)를 통해 입력된 이미지를 저장할 수 있다.

[72] 그리고 메모리(130)는 제어부(120)에 의해 일시적으로 정보를 저장할 수 있는 주의기억이 할당될 수 있으며, 할당된 주의기억에 제어부(120)에 의해 계산된 주의기억정보가 저장될 수 있다.

[73] 도 2 는 일 실시예에 따른 주의기억방법을 설명하기 위한 순서도이다.

[74] 도 2 에 도시된 실시예에 따른 주의기억방법은 도 1 에 도시된 주의기억시스템(10)에서 시계열적으로 처리되는 단계들을 포함한다. 따라서, 이하에서 생략된 내용이라고 하더라도 도 1 에 도시된 주의기억시스템(10)에 관하여 이상에서 기술한 내용은 도 2 에 도시된 실시예에 따른 주의기억방법에도 적용될 수 있다.

[75] 우선, 주의기억시스템(10)은 이미지를 획득할 수 있다(S2001).

[76] 즉, 주의기억시스템(10)은 기 설정된 객체가 포함된 이미지를 입력받을 수 있으며, 입력된 이미지를 저장할 수 있다. 이때 이미지에는 적어도 하나의 객체가 포함될 수 있다.

[77] 그리고 주의기억시스템(10)은 식별된 적어도 하나의 객체 각각의 위치를

추출할 수 있다.

- [78] 예를 들어, 주의기억시스템(10)은 S2001단계에서 획득된 이미지에 대해 YOLO 객체검출방법을 적용할 수 있고, 이미지를 적어도 하나 이상의 영역으로 나누어 객체로 예측되는 위치에 바운딩 박스(bounding box)를 표시할 수 있고, 예측되는 확률에 따라 가중치를 적용한 바운딩 박스를 통해 실시간으로 이미지로부터 적어도 하나의 객체를 검출할 수 있다.
- [79] 이후, 주의기억시스템(10)은 추출된 적어도 하나의 객체 각각의 위치를 기초로 메모리에 할당된 주의기억의 주의기억정보를 초기화할 수 있다(S2002).
- [80] 그리고 주의기억시스템(10)은 이미지에 포함된 적어도 하나의 객체에 대한 특징을 포함하는 시각적특징맵(visual feature map)을 추출할 수 있다.
- [81] 실시예에 따라, 주의기억시스템(10)은 CAEs를 통해 이미지상의 객체에 대한 특징 정보를 추출할 수 있다.
- [82] 예를 들어, 주의기억시스템(10)은 객체의 경계선을 추출할 수 있는 필터를 콘볼루션 계산을 통해 이미지에 적용하여 이미지에 포함되는 객체의 경계선에 대한 값을 나타내는 시각적특징맵을 추출할 수 있다.
- [83] 그리고 주의기억시스템(10)은 상기 이미지 상에서 기 설정된 객체를 식별하기 위한 질문을 생성하고, 생성된 질문에 대한 답변을 도출할 수 있다(S2003).
- [84] 이를 위해, 주의기억시스템(10)은 시각적특징맵 및 주의기억에 저장된 주의기억정보에 기초하여 주의기억특징맵을 생성할 수 있다.
- [85] 즉, 주의기억시스템(10)은 이미지 상에서 객체의 경계선에 대한 정보와 이미지 상에서 검출된 객체의 위치에 기초한 정보를 이용하여 이미지에 포함된 적어도 하나의 객체 즉, 기 설정된 객체로 선택될 수 있는 후보 객체에 대한 특징인 주의기억특징맵을 생성할 수 있다.
- [86] 그리고 주의기억시스템(10)은 기존 질문 및 답변에 대한 정보인 콘텍스트 벡터 그리고 주의기억특징맵을 기초로 질문을 생성할 수 있다.
- [87] 예를 들어, 주의기억시스템(10)은 LSTM(Long Short-Term memory) 그리고 HRED(Hierarchical Recurrent Encoder Decoder)를 통해 기존 질문과 답변을 기초로 인코더(Encoder)에서 출력된 값을 콘텍스트 층(Context Layer)에 입력하여 도출된 값과 주의기억특징맵을 디코더(Decoder)에 입력하여 현재 상태에서의 자연어 질문을 생성할 수 있다.
- [88] 그리고 주의기억시스템(10)은 생성된 질문에 대해 답변을 도출할 수 있다.
- [89] 예를 들어, 주의기억시스템(10)은 “자전거야” 라는 질문에 대해 MLP(Multi-Layer Perceptron)를 통해 “예/아니오” 중 어느 하나의 답변을 도출할 수 있다.
- [90] 이후, 주의기억시스템(10)은 기존 질문과 답변 그리고 시각적특징맵을 기초로 주의기억에 저장된 주의기억정보를 갱신할 수 있다(S2004).
- [91] 즉, 주의기억시스템(10)은 S2003 단계에서 인코더에서 출력된 값과 시각적특징맵을 기초로 주의기억에 저장된 주의기억정보를 갱신할 수 있다.

- [92] 예를 들어, 주의기억시스템(10)은 “자전거야?”라는 질문에 대한 답변인 “아니”를 포함하는 콘텍스트 벡터와 시각적특징맵을 기초로 주의기억정보에서 ‘자전거’에 대한 정보를 제거할 수 있다.
- [93] 그리고 주의기억시스템(10)은 S2003 단계 내지 S2004 단계를 반복하여 기 설정된 횟수만큼 질문을 생성하여 답변을 도출할 수 있고, 기 설정된 질문 수에 도달하였는지 판단할 수 있다(S2005).
- [94] 이때, 주의기억시스템(10)은 질문과 답변을 반복하는 과정에서 주의기억정보를 갱신할 수 있고, 갱신된 주의기억정보를 기초로 효과적인 질문을 생성할 수 있다.
- [95] 이후, 질문과 답변이 완료되면 주의기억시스템(10)은 주의기억에 저장된 주의기억정보 및 객체를 나누는 카테고리 별로 제공되는 예시 객체를 기초로 이미지 상에서 기 설정된 객체를 선택할 수 있다(S2006).
- [96] 예를 들어, 주의기억시스템(10)은 기 설정된 객체에 근접한 후보 객체에 대한 정보인 주의기억정보, 이미지 상의 객체의 공간 정보 및 객체의 카테고리과 비교하여 이미지 상의 후보 객체 중 기 설정된 객체를 선택할 수 있다.
- [97] 이상의 실시예들에서 사용되는 '~부'라는 용어는 소프트웨어 또는 FPGA(field programmable gate array) 또는 ASIC 와 같은 하드웨어 구성요소를 의미하며, '~부'는 어떤 역할들을 수행한다. 그렇지만 '~부'는 소프트웨어 또는 하드웨어에 한정되는 의미는 아니다. '~부'는 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 따라서, 일 예로서 '~부'는 소프트웨어 구성요소들, 객체지향 소프트웨어 구성요소들, 클래스 구성요소들 및 태스크 구성요소들과 같은 구성요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램특허 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로코드, 회로, 데이터, 데이터베이스, 데이터 구조들, 테이블들, 어레이들, 및 변수들을 포함한다.
- [98] 구성요소들과 '~부'들 안에서 제공되는 기능은 더 작은 수의 구성요소들 및 '~부'들로 결합되거나 추가적인 구성요소들과 '~부'들로부터 분리될 수 있다.
- [99] 뿐만 아니라, 구성요소들 및 '~부'들은 디바이스 또는 보안 멀티미디어카드 내의 하나 또는 그 이상의 CPU 들을 재생시키도록 구현될 수도 있다.
- [100] 도 2 를 통해 설명된 실시예에 따른 주의기억방법은 컴퓨터에 의해 실행 가능한 명령어 및 데이터를 저장하는, 컴퓨터로 판독 가능한 매체의 형태로도 구현될 수 있다. 이때, 명령어 및 데이터는 프로그램 코드의 형태로 저장될 수 있으며, 프로세서에 의해 실행되었을 때, 소정의 프로그램 모듈을 생성하여 소정의 동작을 수행할 수 있다. 또한, 컴퓨터로 판독 가능한 매체는 컴퓨터에 의해 액세스될 수 있는 임의의 가용 매체일 수 있고, 휘발성 및 비휘발성 매체, 분리형 및 비분리형 매체를 모두 포함한다. 또한, 컴퓨터로 판독 가능한 매체는 컴퓨터 기록 매체일 수 있는데, 컴퓨터 기록 매체는 컴퓨터 판독 가능 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 포함할 수 있다. 예를 들어, 컴퓨터 기록 매체는 HDD 및 SSD 등과 같은 마그네틱 저장 매체, CD, DVD 및 블루레이 디스크 등과 같은 광학적 기록 매체, 또는 네트워크를 통해 접근 가능한 서버에 포함되는 메모리일 수 있다.

[101] 또한 도 2를 통해 설명된 실시예에 따른 주의기억방법은 컴퓨터에 의해 실행 가능한 명령어를 포함하는 컴퓨터 프로그램(또는 컴퓨터 프로그램 제품)으로 구현될 수도 있다. 컴퓨터 프로그램은 프로세서에 의해 처리되는 프로그래밍 가능한 기계 명령어를 포함하고, 고레벨 프로그래밍 언어(High-level Programming Language), 객체 지향 프로그래밍 언어(Object-oriented Programming Language), 어셈블리 언어 또는 기계 언어 등으로 구현될 수 있다. 또한 컴퓨터 프로그램은 유형의 컴퓨터 판독가능 기록매체(예를 들어, 메모리, 하드디스크, 자기/광학 매체 또는 SSD(Solid-State Drive) 등)에 기록될 수 있다.

[102] 따라서 도 2를 통해 설명된 실시예에 따른 주의기억방법은 상술한 바와 같은 컴퓨터 프로그램이 컴퓨팅 장치에 의해 실행됨으로써 구현될 수 있다. 컴퓨팅 장치는 프로세서와, 메모리와, 저장 장치와, 메모리 및 고속 확장포트에 접속하고 있는 고속 인터페이스와, 저속 버스와 저장 장치에 접속하고 있는 저속 인터페이스 중 적어도 일부를 포함할 수 있다. 이러한 성분들 각각은 다양한 버스를 이용하여 서로 접속되어 있으며, 공통 머더보드에 탑재되거나 다른 적절한 방식으로 장착될 수 있다.

[103] 여기서 프로세서는 컴퓨팅 장치 내에서 명령어를 처리할 수 있는데, 이런 명령어로는, 예컨대 고속 인터페이스에 접속된 디스플레이처럼 외부 입력, 출력 장치상에 GUI(Graphic User Interface)를 제공하기 위한 그래픽 정보를 표시하기 위해 메모리나 저장 장치에 저장된 명령어를 들 수 있다. 다른 실시예로서, 다수의 프로세서 및(또는) 다수의 버스가 적절히 다수의 메모리 및 메모리 형태와 함께 이용될 수 있다. 또한 프로세서는 독립적인 다수의 아날로그 및(또는) 디지털 프로세서를 포함하는 칩들이 이루는 칩셋으로 구현될 수 있다.

[104] 또한 메모리는 컴퓨팅 장치 내에서 정보를 저장한다. 일례로, 메모리는 휘발성 메모리 유닛 또는 그들의 집합으로 구성될 수 있다. 다른 예로, 메모리는 비휘발성 메모리 유닛 또는 그들의 집합으로 구성될 수 있다. 또한 메모리는 예컨대, 자기 혹은 광 디스크와 같이 다른 형태의 컴퓨터 판독 가능한 매체일 수도 있다.

[105] 그리고 저장장치는 컴퓨팅 장치에게 대용량의 저장공간을 제공할 수 있다. 저장 장치는 컴퓨터 판독 가능한 매체이거나 이런 매체를 포함하는 구성일 수 있으며, 예를 들어 SAN(Storage Area Network) 내의 장치들이나 다른 구성도 포함할 수 있고, 플로피 디스크 장치, 하드 디스크 장치, 광 디스크 장치, 혹은 테이프 장치, 플래시 메모리, 그와 유사한 다른 반도체 메모리 장치 혹은 장치 어레이일 수 있다.

[106] 상술된 실시예들은 예시를 위한 것이며, 상술된 실시예들이 속하는 기술분야의

통상의 지식을 가진 자는 상술된 실시예들이 갖는 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 상술된 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.

- [107] 본 명세서를 통해 보호 받고자 하는 범위는 상기 상세한 설명보다는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태를 포함하는 것으로 해석되어야 한다.

청구범위

- [청구항 1] 이미지 상의 객체를 식별하는 주의기억시스템에 있어서,
상기 이미지 상에서 기 설정된 객체를 식별하기 위한 질문을 생성하고,
생성된 질문에 대해 답변을 도출하고, 상기 질문 및 상기 답변에 기초하여
상기 이미지에서 기 설정된 객체를 식별하는 제어부; 및
상기 이미지를 저장하는 메모리를 포함하며,
상기 제어부는,
상기 이미지에 포함된 객체들에 대한 정보를 반영하여 상기 질문을
생성하고, 상기 답변에 기초하여 상기 객체들에 대한 정보를 갱신하는,
주의기억시스템.
- [청구항 2] 제 1 항에 있어서,
상기 제어부는,
상기 이미지 상의 적어도 하나의 객체를 식별하고, 식별된 적어도 하나의
객체 각각의 위치를 추출하고, 추출된 적어도 하나의 객체 각각의
위치정보를 기초로 상기 메모리 상의 주의기억을 초기화하는,
주의기억시스템.
- [청구항 3] 제 2 항에 있어서,
상기 제어부는,
상기 이미지에 포함된 적어도 하나의 객체에 대한 특징을 포함하는
시각적특징맵을 추출하는, 주의기억시스템.
- [청구항 4] 제 3 항에 있어서,
상기 제어부는,
상기 시각적특징맵 및 상기 주의기억에 저장되는 상기 객체들에 대한
정보인 주의기억정보에 기초하여 주의기억특징맵을 생성하고, 생성된
주의기억특징맵을 기초로 상기 질문을 생성하는, 주의기억시스템.
- [청구항 5] 제 4 항에 있어서,
상기 제어부는,
상기 답변 및 상기 시각적특징맵을 기초로 상기 주의기억정보를
갱신하는, 주의기억시스템.
- [청구항 6] 제 5 항에 있어서,
상기 제어부는,
상기 주의기억정보에 기초하여, 상기 이미지에서 상기 기 설정된 객체를
선택하는, 주의기억시스템.
- [청구항 7] 주의기억시스템이 이미지 상의 객체를 식별하는 주의기억방법에 있어서,
상기 이미지를 입력받는 단계;
상기 이미지에 포함된 객체들에 대한 정보를 반영하여 기 설정된 객체를
식별하기 위한 질문을 생성하고, 생성된 질문에 대한 답변을 도출하는

단계; 및

적어도 하나의 질문 및 상기 질문 및 상기 답변에 기초하여 상기 이미지에서 기 설정된 객체를 식별하는 단계를 포함하고,

상기 주의기억방법은,

상기 답변에 기초하여 상기 객체들에 대한 정보를 갱신하는 단계를 포함하는, 주의기억방법.

[청구항 8]

제 7 항에 있어서,

상기 이미지를 입력받는 단계는,

상기 이미지 상의 적어도 하나의 객체를 식별하는 단계; 및

식별된 적어도 하나의 객체 각각의 위치를 추출하고, 추출된 적어도 하나의 객체 각각의 위치정보를 기초로 메모리 상의 주의기억을 초기화하는 단계를 포함하는, 주의기억방법.

[청구항 9]

제 8 항에 있어서,

상기 이미지를 입력받는 단계는,

상기 이미지에 포함된 적어도 하나의 객체에 대한 특징을 포함하는 시각적특징맵을 추출하는 단계를 포함하는, 주의기억방법.

[청구항 10]

제 9 항에 있어서,

상기 주 질문을 생성하는 단계는,

상기 시각적특징맵 및 상기 주의기억에 저장되는 상기 객체들에 대한 정보인 주의기억정보에 기초하여 주의기억특징맵을 생성하는 단계; 및 생성된 주의기억특징맵을 기초로 상기 질문을 생성하는 단계를 포함하는, 주의기억방법.

[청구항 11]

제 10 항에 있어서,

상기 주의기억방법은,

상기 답변 및 상기 시각적특징맵을 기초로 상기 주의기억정보를 갱신하는 단계를 더 포함하는, 주의기억방법.

[청구항 12]

제 11 항에 있어서,

상기 기 설정된 객체를 식별하는 단계는,

상기 주의기억정보에 기초하여, 상기 이미지 상에서 상기 기 설정된 객체를 선택하는 단계를 더 포함하는, 주의기억방법.

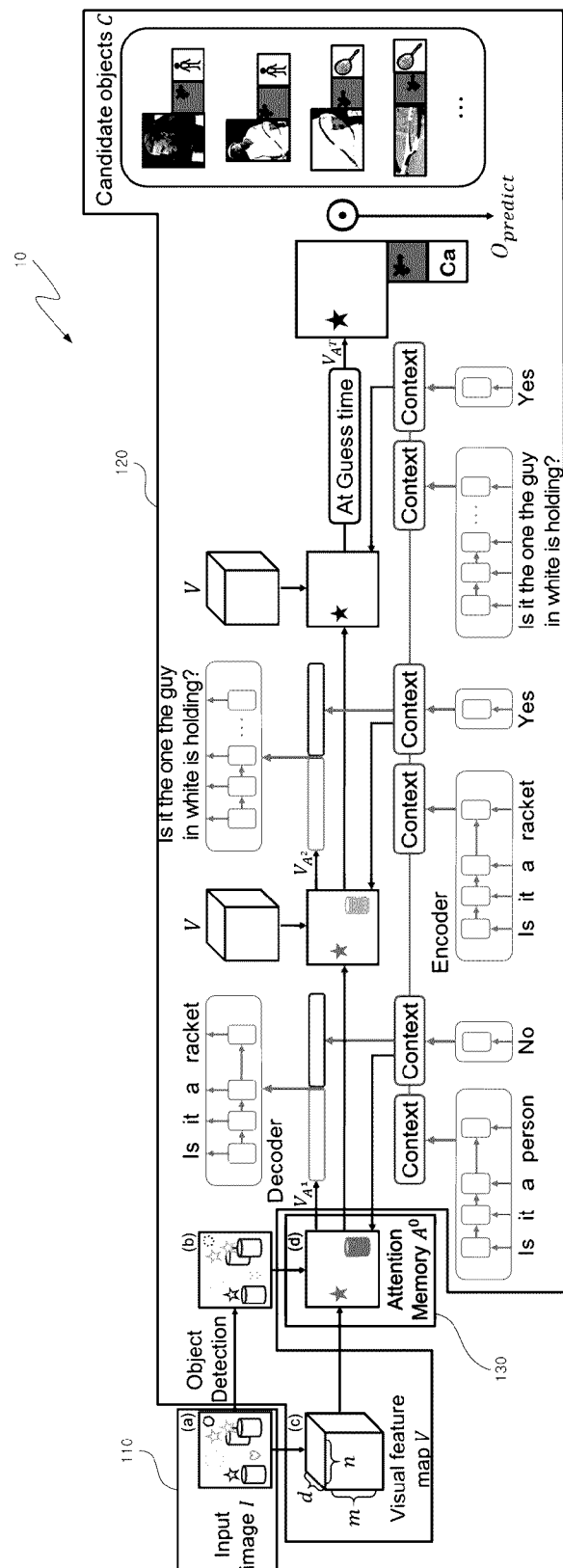
[청구항 13]

제 8 항에 기재된 방법을 수행하는 프로그램이 기록된 컴퓨터 판독 가능한 기록 매체.

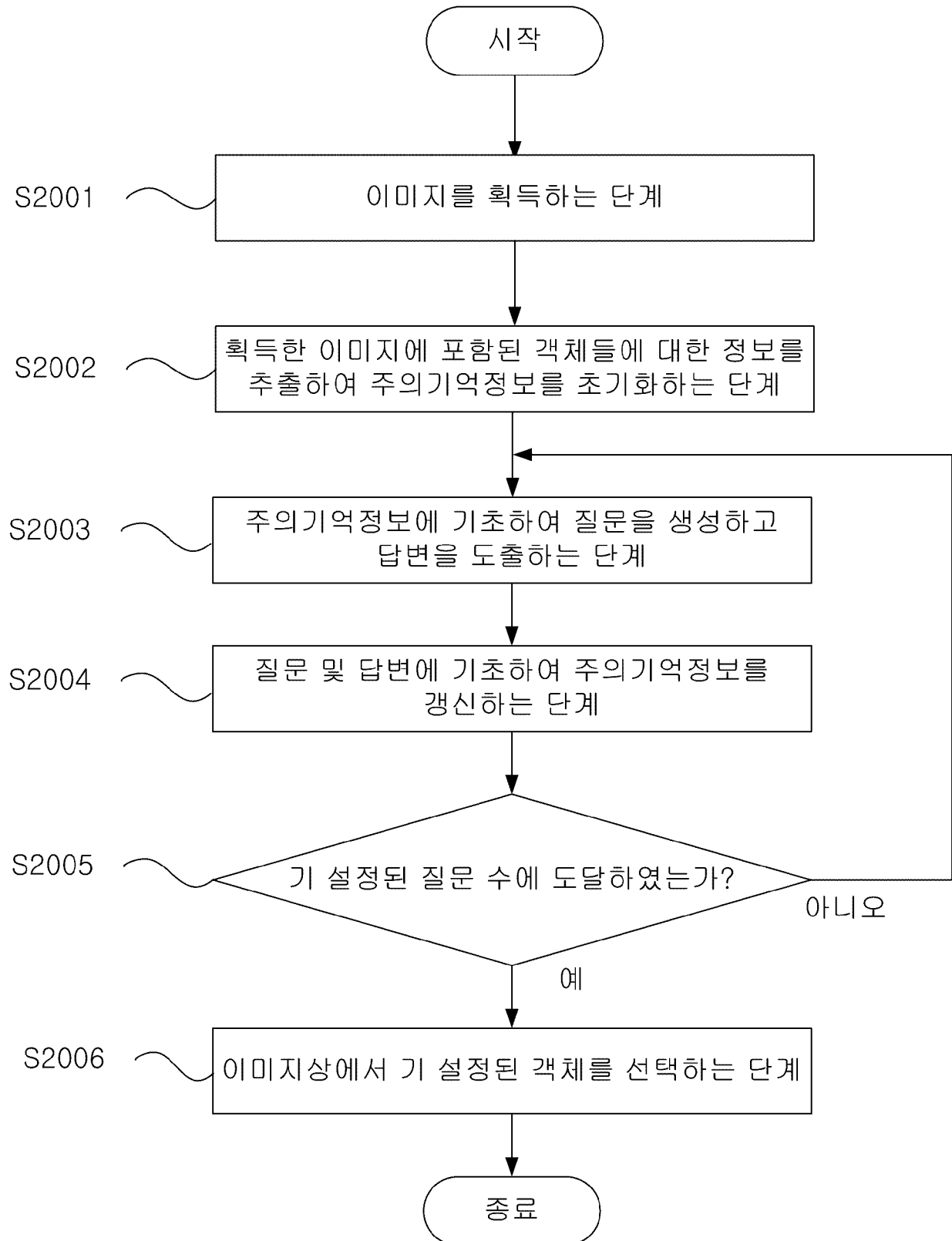
[청구항 14]

주의기억시스템에 의해 수행되며, 제 8 항에 기재된 방법을 수행하기 위해 매체에 저장된 컴퓨터 프로그램.

[도1]



[도2]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2018/007410

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/16(2006.01)i, G10L 15/22(2006.01)i, G06T 1/00(2006.01)i, G06T 7/00(2006.01)i

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L 15/16; G06F 3/01; G06F 9/06; G06K 9/62; G06K 9/66; G06N 3/10; G06N 7/06; G10L 15/22; G06T 1/00; G06T 7/00

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

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

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

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

eKOMPASS (KIPO internal) & Keywords: sight, object, position, question, answer, attention memory

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DE VRIES, Harm et al., "GuessWhat?! Visual Object Discovery through Multi-modal Dialogue", arXiv.org, arXiv:1611.08481v1 [cs.AI], 23 November 2016, Retrieved from < https://arxiv.org/abs/1611.08481v1 >. See pages 1-2, 5, 7-8; and figure 7.	1,7
Y		2-6,8-14
Y	XIONG, Caiming et al., "Dynamic Memory Networks for Visual and Textual Question Answering", arXiv.org, arXiv: 1603.01417v1 [cs.NE], 04 March 2016, Retrieved from < https://arxiv.org/abs/1603.01417 >. See pages 1-2 and figure 1.	2-6,8-14
A	WO 2017-062645 A1 (EVOLV TECHNOLOGIES, INC.) 13 April 2017 See paragraphs [00138]-[00168] and figures 1-4.	1-14
A	KR 10-2007-0118038 A (SONY CORPORATION) 13 December 2007 See paragraphs [0151]-[0218] and figures 4-5.	1-14
A	US 2017-0124432 A1 (BAIDU USA, LLC.) 04 May 2017 See paragraphs [0023]-[0064] and figures 2-3.	1-14



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

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

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

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

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

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

"T"

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

"X"

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

"Y"

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

"&"

document member of the same patent family

Date of the actual completion of the international search

29 OCTOBER 2018 (29.10.2018)

Date of mailing of the international search report

29 OCTOBER 2018 (29.10.2018)

Name and mailing address of the ISA/KR



Korean Intellectual Property Office
Government Complex Daejeon Building 4, 189, Cheongsu-ro, Seo-gu,
Daejeon, 35208, Republic of Korea

Facsimile No. +82-42-481-8578

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2018/007410

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HAN, Cheolho et al., "Attention Memory for Locating an Object through Visual Dialogue", VQA Challenge Workshop at CVPR 2017, Honolulu, Hawaii, USA, pp. 1-3, 26 July 2017 See the entire document. ※ The above document is a document declaring non-prejudicial lack of novelty by the applicant.	1-14

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2018/007410

Patent document cited in search report	Publication date	Patent family member	Publication date
WO 2017-062645 A1	13/04/2017	US 2017-098161 A1	06/04/2017
		US 2017-098162 A1	06/04/2017
		US 2017-098172 A1	06/04/2017
		US 2017-099200 A1	06/04/2017
		WO 2017-062610 A1	13/04/2017
		WO 2017-062623 A1	13/04/2017
		WO 2017-062635 A1	13/04/2017
KR 10-2007-0118038 A	13/12/2007	CN 101086741 A	12/12/2007
		CN 101086741 B	21/03/2012
		CN 101655867 A	24/02/2010
		CN 101655867 B	14/11/2012
		EP 1865426 A2	12/12/2007
		EP 1865426 A3	02/05/2012
		EP 1865426 B1	07/03/2018
		JP 04175390 B2	05/11/2008
		JP 2007-328675 A	20/12/2007
		US 2008-0010060 A1	10/01/2008
		US 7945439 B2	17/05/2011
US 2017-0124432 A1	04/05/2017	CN 106649542 A	10/05/2017
		EP 3166049 A1	10/05/2017
		JP 06351689 B2	04/07/2018
		JP 2017-091525 A	25/05/2017
		KR 10-1865102 B1	07/06/2018
		KR 10-2018-0038937 A	17/04/2018
		US 9965705 B2	08/05/2018

A. 발명이 속하는 기술분류(국제특허분류(IPC)) G10L 15/16(2006.01)i, G10L 15/22(2006.01)i, G06T 1/00(2006.01)i, G06T 7/00(2006.01)i																							
B. 조사된 분야 조사된 최소문헌(국제특허분류를 기재) G10L 15/16; G06F 3/01; G06F 9/06; G06K 9/62; G06K 9/66; G06N 3/10; G06N 7/06; G10L 15/22; G06T 1/00; G06T 7/00 조사된 기술분야에 속하는 최소문헌 이외의 문헌 한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우)) eKOMPASS(특허청 내부 검색시스템) & 키워드: 시각, 객체, 위치, 질문, 답, 주의기억																							
C. 관련 문헌 <table border="1"> <thead> <tr> <th>카테고리*</th> <th>인용문헌명 및 관련 구절(해당하는 경우)의 기재</th> <th>관련 청구항</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>HARM DE VRIES 등, 'GuessWhat?! Visual object discovery through multi-modal dialogue', arXiv.org, arXiv:1611.08481v1 [cs.AI], 2016.11.23 <https://arxiv.org/abs/1611.08481v1>에서 검색됨. 페이지 1-2, 5, 7-8; 및 도면 7 참조.</td> <td>1,7</td> </tr> <tr> <td>Y</td> <td></td> <td>2-6,8-14</td> </tr> <tr> <td>Y</td> <td>CAIMING XIONG 등, 'Dynamic Memory Networks for Visual and Textual Question Answering', arXiv.org, arXiv:1603.01417v1 [cs.NE], 2016.03.04 <https://arxiv.org/abs/1603.01417>에서 검색됨. 페이지 1-2 및 도면 1 참조.</td> <td>2-6,8-14</td> </tr> <tr> <td>A</td> <td>WO 2017-062645 A1 (EVLV TECHNOLOGIES, INC.) 2017.04.13 단락 [00138]-[00168] 및 도면 1-4 참조.</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>KR 10-2007-0118038 A (소니 가부시키 가이샤) 2007.12.13 단락 [0151]-[0218] 및 도면 4-5 참조.</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>US 2017-0124432 A1 (BAIDU USA, LLC) 2017.05.04 단락 [0023]-[0064] 및 도면 2-3 참조.</td> <td>1-14</td> </tr> </tbody> </table>			카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항	X	HARM DE VRIES 등, 'GuessWhat?! Visual object discovery through multi-modal dialogue', arXiv.org, arXiv:1611.08481v1 [cs.AI], 2016.11.23 <https://arxiv.org/abs/1611.08481v1>에서 검색됨. 페이지 1-2, 5, 7-8; 및 도면 7 참조.	1,7	Y		2-6,8-14	Y	CAIMING XIONG 등, 'Dynamic Memory Networks for Visual and Textual Question Answering', arXiv.org, arXiv:1603.01417v1 [cs.NE], 2016.03.04 <https://arxiv.org/abs/1603.01417>에서 검색됨. 페이지 1-2 및 도면 1 참조.	2-6,8-14	A	WO 2017-062645 A1 (EVLV TECHNOLOGIES, INC.) 2017.04.13 단락 [00138]-[00168] 및 도면 1-4 참조.	1-14	A	KR 10-2007-0118038 A (소니 가부시키 가이샤) 2007.12.13 단락 [0151]-[0218] 및 도면 4-5 참조.	1-14	A	US 2017-0124432 A1 (BAIDU USA, LLC) 2017.05.04 단락 [0023]-[0064] 및 도면 2-3 참조.	1-14
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항																					
X	HARM DE VRIES 등, 'GuessWhat?! Visual object discovery through multi-modal dialogue', arXiv.org, arXiv:1611.08481v1 [cs.AI], 2016.11.23 <https://arxiv.org/abs/1611.08481v1>에서 검색됨. 페이지 1-2, 5, 7-8; 및 도면 7 참조.	1,7																					
Y		2-6,8-14																					
Y	CAIMING XIONG 등, 'Dynamic Memory Networks for Visual and Textual Question Answering', arXiv.org, arXiv:1603.01417v1 [cs.NE], 2016.03.04 <https://arxiv.org/abs/1603.01417>에서 검색됨. 페이지 1-2 및 도면 1 참조.	2-6,8-14																					
A	WO 2017-062645 A1 (EVLV TECHNOLOGIES, INC.) 2017.04.13 단락 [00138]-[00168] 및 도면 1-4 참조.	1-14																					
A	KR 10-2007-0118038 A (소니 가부시키 가이샤) 2007.12.13 단락 [0151]-[0218] 및 도면 4-5 참조.	1-14																					
A	US 2017-0124432 A1 (BAIDU USA, LLC) 2017.05.04 단락 [0023]-[0064] 및 도면 2-3 참조.	1-14																					
☒ 추가 문헌이 C(계속)에 기재되어 있습니다. ☒ 대응특허에 관한 별지를 참조하십시오.																							
* 인용된 문헌의 특별 카테고리: “A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌 “E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌 “L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌 “O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌 “P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌 “T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌 “X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신구성 또는 진보성이 없는 것으로 본다. “Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다. “&” 동일한 대응특허문헌에 속하는 문헌																							
국제조사의 실제 완료일 2018년 10월 29일 (29.10.2018)		국제조사보고서 발송일 2018년 10월 29일 (29.10.2018)																					
ISA/KR의 명칭 및 우편주소  대한민국 특허청 (35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사) 팩스 번호 +82-42-481-8578		심사관 노지명 전화번호 +82-42-481-8528 																					

C (계속). 관련 문헌		
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
X	CHEOLHO HAN 등, `Attention Memory for Locating an Object through Visual Dialogue`, VQA Challenge Workshop at CVPR 2017, Honolulu, Hawaii, USA, pp. 1-3, 2017.07.26 전체 문서 참조. ※상기 문헌은 출원인이 신규성 상실의 예외로서 선언한 공지문헌임.	1-14

국제조사보고서
대응특허에 관한 정보

국제출원번호

PCT/KR2018/007410

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
WO 2017-062645 A1	2017/04/13	US 2017-098161 A1 US 2017-098162 A1 US 2017-098172 A1 US 2017-099200 A1 WO 2017-062610 A1 WO 2017-062623 A1 WO 2017-062635 A1	2017/04/06 2017/04/06 2017/04/06 2017/04/06 2017/04/13 2017/04/13 2017/04/13
KR 10-2007-0118038 A	2007/12/13	CN 101086741 A CN 101086741 B CN 101655867 A CN 101655867 B EP 1865426 A2 EP 1865426 A3 EP 1865426 B1 JP 04175390 B2 JP 2007-328675 A US 2008-0010060 A1 US 7945439 B2	2007/12/12 2012/03/21 2010/02/24 2012/11/14 2007/12/12 2012/05/02 2018/03/07 2008/11/05 2007/12/20 2008/01/10 2011/05/17
US 2017-0124432 A1	2017/05/04	CN 106649542 A EP 3166049 A1 JP 06351689 B2 JP 2017-091525 A KR 10-1865102 B1 KR 10-2018-0038937 A US 9965705 B2	2017/05/10 2017/05/10 2018/07/04 2017/05/25 2018/06/07 2018/04/17 2018/05/08