

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2012년 12월 6일 (06.12.2012)



(10) 국제공개번호
WO 2012/165929 A2

- (51) 국제특허분류: G06F 17/30 (2006.01) G10L 15/26 (2006.01)
- (21) 국제출원번호: PCT/KR2012/004405
- (22) 국제출원일: 2012년 6월 4일 (04.06.2012)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보: 10-2011-0053400 2011년 6월 2일 (02.06.2011) KR
- (71) 출원인 (US 을(를) 제외한 모든 지정국에 대하여): 포항 공과대학교 산학협력단 (POSTECH ACADEMY-INDUSTRY FOUNDATION) [KR/KR]; 경상북도 포항시 남구 효자동 산 31번지, 790-784 Gyeongsangbuk-do (KR).
- (72) 발명자; 겸
- (75) 발명자/출원인 (US 에 한하여): 이근배 (LEE, Geun Bae) [KR/KR]; 경북 포항시 남구 지곡동 교수아파트 9동 1003호, 790-751 Gyeongbuk (KR). 김석환 (KIM, Seok Hwan) [KR/KR]; 경북 포항시 남구 지곡동 포항 공과대학교 기숙사 19동 318호, 790-390 Gyeongbuk

(KR). 김경덕 (KIM, Kyung Duk) [KR/KR]; 경북 포항시 남구 효자동 포항공과대학교 공학 2동 202호, 790-784 Gyeongbuk (KR). 이동현 (LEE, Dong Hyeon) [KR/KR]; 경상남도 남해군 고현면 이어리 221-1, 668-883 Gyeongsangnam-do (KR). 최준휘 (CHOI, Jun Hwi) [KR/KR]; 서울 은평구 신사동 354번지 라이프 시티 아파트 102동 904호, 122-080 Seoul (KR).

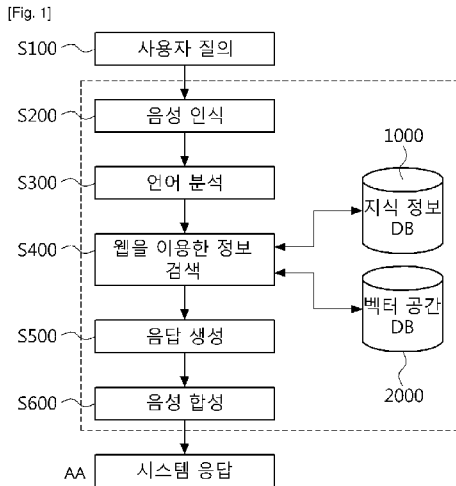
(74) 대리인: 특허법인 이상 (E-SANG PATENT & TRADE-MARK LAW FIRM); 서울 서초구 양재동 82-2 우도빌딩 3층, 137-890 Seoul (KR).

(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, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, 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.

[다음 쪽 계속]

(54) Title: METHOD FOR SEARCHING FOR INFORMATION USING THE WEB AND METHOD FOR VOICE CONVERSATION USING SAME

(54) 발명의 명칭: 웹을 이용한 정보 검색 방법 및 이를 사용하는 음성 대화 방법



AA ... Response from a system
 S100 ... User query
 S200 ... Voice recognition
 S300 ... Language analysis
 S400 ... Information search using the Web
 S500 ... Generate a response
 S600 ... Voice synthesis
 1000 ... Knowledge information DB
 2000 ... Vector space DB

(57) Abstract: According to the present invention, a method for searching for information using the Web and a method for voice conversation using same involve generating a basic word vector for at least one of a provided user query and a language analysis result; using the basic word vector, searching a vector space database for a vector space corresponding to the basic word vector; searching, if a similarity between the basic word vector and the found vector space is lower than a preset reference, a vector space database for a vector space corresponding to an expanded word vector using the expanded word vector generated from the result of the Web search performed using at least one of the user query and the language analysis result; and searching for knowledge information on the basis of the vector space found in the basic searching step or the expanded searching step. Thus, improved research results may be provided in response to the user query.

(57) 요약서: 웹을 이용한 정보 검색 방법 및 이를 사용하는 음성 대화 방법은 제공된 사용자 질의 및 언어 분석 결과 중 적어도 하나에 대한 기본 단어 벡터를 생성하고, 기본 단어 벡터를 이용하여 벡터 공간 데이터베이스에서 기본 단어 벡터에 대응되는 벡터 공간을 검색한 후, 기본 단어 벡터와 검색된 벡터 공간과의 유사도가 미리 설정된 기준 이하인 경우, 사용자 질의 및 언어 분석 결과 중 적어도 하나를 이용하여 수행된 웹 검색 결과로부터, 생성한 확장 단어 벡터를 이용하여 벡터 공간 데이터베이스에서 확장 단어 벡터에 대응되는 벡터 공간을 검색하고, 기본 검색 단계 또는 확장 검색 단계에서 검색된 벡터 공간에 기초하여 지식 정보를 검색한다. 따라서, 사용자 질의에 대하여 보다 나은 검색 결과를 제공할 수 있다.

WO 2012/165929 A2



(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, 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, ML, MR, NE, SN, TD, TG).

공개:

- 국제조사보고서 없이 공개하며 보고서 접수 후 이를 별도 공개함 (규칙 48.2(g))

명세서

발명의 명칭: 웹을 이용한 정보 검색 방법 및 이를 사용하는 음성 대화 방법

기술분야

- [1] 본 발명은 검색 방법에 관한 것으로, 더욱 상세하게는 웹을 이용하여 정보 검색을 보다 효과적으로 수행할 수 있는 웹을 이용한 정보 검색 방법 및 이를 사용하는 음성 대화 방법에 관한 것이다.

배경기술

- [2] 정보 검색은 특정 지식 정보를 바탕으로 사용자의 질의에 대해 적절한 응답을 내어주는 것이다. 예를 들어, 스마트 TV 시스템에서 사용자가 특정 프로그램을 찾는 질의를 했을 때, 정보 검색을 통해 TV 프로그램 일정을 담고 있는 지식 정보 데이터베이스에 접근하여 사용자의 질의에 대한 응답을 내어 줄 수 있다. 그러나, 기존의 검색 방법에서 사용되는 지식 정보 데이터베이스는 주로 관계형 데이터베이스 형태로 저장되며, 이 경우에 사용자가 특정 정보에 대해 데이터베이스에 저장된 형태로 발화하지 않은 경우 적절한 항목을 찾아낼 수 없는 문제가 있다. 또한, 의미적인 추론이 필요한 복잡한 형태의 질의인 경우에는 적절한 응답을 나타내지 못하는 경향이 있다. 예를 들어 스마트 TV에서 '박지성 경기 보자'라는 사용자 질의에 대해 TV 스케줄에 저장된 '아시안컵'이라는 검색 결과를 내지 못한다. 왜냐하면 '아시안컵' 항목에는 박지성과 관련된 정보가 포함되어 있지 않기 때문이다. 이를 제대로 수행하려면 박지성이 아시안컵에 참석하고 있다는 정보가 추가로 필요하다. 이런 의미적인 부분을 고려하기 위한 접근 방법으로 온톨로지를 활용하는 경우가 있으나, 온톨로지를 구축하고 관리하는 것은 매우 힘든 작업이며 지식 정보와 관련된 질의를 추론하기 위한 수많은 온톨로지를 구축하는 것은 불가능하다. 또한, 실시간으로 변화하는 정보를 반영하는 것에도 한계가 있다.

발명의 상세한 설명

기술적 과제

- [3] 상기와 같은 문제점을 해결하기 위한 본 발명의 목적은, 사용자 질의에 대해 웹을 이용하여 사용자 질의에 대한 보다 향상된 응답을 제공할 수 있는 웹을 이용한 정보 검색 방법을 제공하는데 있다.
- [4] 상기와 같은 문제점을 해결하기 위한 본 발명의 다른 목적은, 음성 형태의 사용자 질의에 대해 웹을 이용하여 사용자 질의에 대한 보다 향상된 응답을 음성 형태로 제공할 수 있는 웹을 이용한 정보 검색 방법을 사용하는 음성 대화 방법을 제공하는데 있다.
- [5] 본 발명의 기술적 과제들은 이상에서 언급한 기술적 과제로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 아래의 기재로부터 당업자에게

명확하게 이해될 수 있을 것이다.

과제 해결 수단

[6] 상기한 본 발명의 목적을 달성하기 위한 본 발명의 일 측면에 따른 웹을 이용한 정보 검색 방법은, 정보 검색 시스템의 웹을 이용한 정보 검색 방법에 있어서, 제공된 사용자 질의 및 언어 분석 결과 중 적어도 하나에 대한 기본 단어 백터를 생성하고, 상기 기본 단어 백터를 이용하여 백터 공간 데이터베이스에서 상기 기본 단어 백터에 대응되는 백터 공간을 검색하는 기본 검색 단계와, 상기 기본 단어 백터와 검색된 상기 백터 공간과의 유사도를 판단하는 단계와, 상기 유사도가 미리 설정된 기준 이하인 경우, 상기 사용자 질의 및 상기 언어 분석 결과 중 적어도 하나를 이용하여 수행된 웹 검색 결과로부터, 확장 단어 백터를 생성하고, 생성된 상기 확장 단어 백터를 이용하여 상기 백터 공간 데이터베이스에서 상기 확장 단어 백터에 대응되는 백터 공간을 검색하는 확장 검색 단계와, 상기 기본 검색 단계 또는 상기 확장 검색 단계에서 검색된 상기 백터 공간에 기초하여 지식 정보를 검색하는 단계를 포함한다.

[7] 상기한 본 발명의 다른 목적을 달성하기 위한 본 발명의 일 측면에 따른 웹을 이용한 정보 검색 방법을 사용하는 음성 대화 방법은 정보 검색 시스템의 웹을 이용한 정보 검색 방법을 사용하는 음성 대화 방법에 있어서, 제공된 음성 형태의 사용자 질의를 텍스트 형태로 변환하는 음성 인식 단계와, 상기 텍스트 형태로 변환된 사용자 질의에 언어 분석을 수행하는 언어 분석 단계와, 변환된 상기 사용자 질의 및 상기 언어 분석 결과 중 적어도 하나에 대한 기본 단어 백터를 생성하고, 상기 기본 단어 백터를 이용하여 백터 공간 데이터베이스에서 상기 기본 단어 백터에 대응되는 백터 공간을 검색한 후, 상기 기본 단어 백터와 검색된 상기 백터 공간과의 유사도가 미리 설정된 기준 이하인 경우, 상기 사용자 질의 및 상기 언어 분석 결과 중 적어도 하나를 이용하여 수행된 웹 검색 결과로부터, 생성한 확장 단어 백터를 이용하여 상기 백터 공간 데이터베이스에서 상기 확장 단어 백터에 대응되는 백터 공간을 검색하고, 상기 기본 단어 백터 또는 상기 확장 단어 백터에 대응되어 검색된 상기 백터 공간에 기초하여 지식 정보를 검색하는 웹을 이용한 정보 검색 단계와, 검색된 상기 지식 정보를 상기 사용자의 질의에 대응되는 응답으로 변환하는 응답 생성 단계와, 상기 응답을 음성 형태로 변환하는 음성 합성 단계를 포함한다.

발명의 효과

[8] 상기와 같은 본 발명에 따른 웹을 이용한 정보 검색 방법 및 이를 사용하는 음성 대화 방법은 정보 검색을 위해 웹을 활용하여 지식 정보 및 사용자 질의에 관한 정보를 확장함으로써 사용자 질의에 대하여 보다 나은 검색 결과를 제공할 수 있다.

[9] 또한, 정보 확장이 웹을 통하여 수행되므로 사람의 개입이 필요 없고, 기존의 백터 공간 모델을 이용할 수 있고, 최근에 성행하는 Q&A나 소셜 네트워크 정보

등 수많은 정보들이 실시간으로 반영되므로 추론 및 추천 기능의 수행이 가능한 장점이 있다.

도면의 간단한 설명

- [10] 도 1은 본 발명의 일 실시예에 따른 웹을 이용한 정보 검색 방법을 나타내는 흐름도이다.
- [11] 도 2는 본 발명의 일 실시예에 따른 웹을 이용한 정보 검색 방법 중 기본 검색 과정을 나타내는 흐름도이다.
- [12] 도 3은 본 발명의 일 실시예에 따른 웹을 이용한 정보 검색 방법 중 확장 검색 과정을 나타내는 흐름도이다.
- [13] 도 4는 본 발명의 일 실시예에 따른 웹을 이용한 정보 검색 방법 중 벡터 공간 데이터베이스를 형성 과정을 나타내는 흐름도이다.
- [14] 도 5는 본 발명의 일 실시예에 따른 웹을 이용한 정보 검색 방법을 사용한 음성 대화 방법을 나타내는 흐름도이다.

발명의 실시를 위한 형태

- [15] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였다.
- [16] 제1, 제2, A, B 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.
- [17] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다.
- [18] 본 출원에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품

또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

- [19] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [20] 이하, 본 발명의 실시예에서, 웹을 이용한 정보 검색 방법 및 이를 사용한 음성 대화 방법은 적어도 하나의 서버를 포함하는 시스템으로 구현될 수 있고, 상기 시스템에는 적어도 하나의 서버와 다수의 사용자 단말이 포함될 수 있다.
- [21] 상기 적어도 하나의 서버와 다수의 사용자 단말은 직접적으로 연결될 수도 있고, 유선 또는 무선 네트워크를 통하여 상호 연결될 수 있다. 또한, 상기 서버는 웹 서버(Web Server) 등이 될 수 있고, 상기 사용자 단말은 PMP(Portable Multimedia Player), 노트북, DTV, 스마트폰(Smart Phone) 등 상기 서버와 통신이 가능하고, 정보 처리 기능을 구비한 다양한 단말이 될 수 있다.
- [22] 이하, 본 발명의 실시예에서는 설명의 편의를 위하여 상기 웹을 이용한 정보 검색 방법 및 이를 사용한 음성 대화 방법을 수행하는 시스템을 정보 검색 시스템이라 지칭한다.
- [23]
- [24] 이하, 본 발명에 따른 바람직한 실시예를 첨부된 도면을 참조하여 상세하게 설명한다.
- [25] 도 1은 본 발명의 일 실시예에 따른 웹을 이용한 정보 검색 방법을 사용한 음성 대화 방법을 나타내는 흐름도이다.
- [26] 도 1을 참조하면, 웹을 이용한 정보 검색 방법을 사용한 음성 대화 방법은 음성 형태의 사용자 질의를 제공받는 단계(S100)와, 상기 음성 형태의 사용자 질의를 텍스트 형태로 변환하는 단계(S200)와, 상기 텍스트 형태로 변환된 사용자 질의에 언어 분석을 수행하는 단계(S300)와, 상기 사용자 질의 및 상기 언어 분석 결과 중 적어도 하나를 제공받아 정보 검색을 하는 웹을 이용한 정보 검색 단계(S400)와, 상기 웹을 이용한 정보 검색 단계에서 검색된 결과를 상기 사용자 질의에 대응되는 응답으로 변환하여 응답을 생성하는 응답 생성 단계(S500)와, 상기 응답을 음성 형태로 변환하는 음성 합성 단계(S600)를 포함한다. 상기 응답은 상기 사용자 질의에 대응되는 응답을 의미할 수 있다.
- [27] 상기 웹을 이용한 정보 검색 방법을 사용한 음성 대화 방법은 사용자가 음성 형태로 질의를 발화하고, 상기 음성 형태의 사용자 질의를 텍스트 형태의 사용자 질의로 변환하고, 상기 텍스트 형태로 변환된 사용자 질의에 대하여 언어 분석을 수행하며, 상기 언어 분석을 수행한 결과인 언어 분석 결과 및 텍스트 형태의

상기 사용자 질의 중 적어도 하나를 이용하여 웹 검색을 수행하여 상기 사용자 질의에 대한 적절한 응답을 검색한다. 상기 검색된 결과를 상기 사용자 질의에 대응되는 응답으로 변환하고, 상기 응답을 음성 형태의 응답으로 변환하여 사용자 질의에 대한 응답을 제공하는 기능을 수행한다. 상기 웹을 이용한 정보 검색 단계(S400)에 대하여 이하에서 구체적으로 설명한다.

- [28] 도 2는 본 발명의 일 실시예에 따른 웹을 이용한 정보 검색 방법을 나타내는 흐름도이고, 도 3은 본 발명의 일 실시예에 따른 웹을 이용한 정보 검색 방법 중 기본 검색 과정을 나타내는 흐름도이고, 도 4는 본 발명의 일 실시예에 따른 웹을 이용한 정보 검색 방법 중 확장 검색 과정을 나타내는 흐름도이고, 도 5는 본 발명의 일 실시예에 따른 웹을 이용한 정보 검색 방법 중 벡터 공간 데이터베이스를 형성 과정을 나타내는 흐름도이다.
- [29] 도 2 내지 도 5를 참조하면, 웹을 이용한 정보 검색 방법(S400)은 사용자 질의 및 사용자 질의에 대한 언어 분석 결과 중 적어도 하나를 제공받아 지식 정보 데이터베이스(1000)에 저장된 지식 정보를 검색하여 사용자 질의에 대응되는 응답을 제공할 수 있다.
- [30] 상기 언어 분석은 형태소 분석, 구문 분석, 의미 분석 등의 언어 분석 기법을 의미할 수 있다.
- [31] 상기 언어 분석 결과는 화행, 주행 및 구조적 정보가 포함될 수 있다. 상기 화행(Dialog Act)은 반드시 문형과 일치하지는 않으나, 일반적으로 서술문, 의문문, 요청문 등의 문형 정보를 나타내며, 상기 주행(Main Action)은 예를 들어, TV 켜기, TV 끄기, 프로그램 찾기 등의 사용자가 원하는 행위를 나타내는 의미적인 정보를 의미하며, 상기 구조 정보는 예를 들어, 장르, 프로그램명, 채널명, 시작시간, 출연진 등의 사용자의 질의에 나타나는 개체명 정보를 의미할 수 있다. 또한, 상기 사용자 질의는 텍스트 형태의 사용자 발화 또는 음성 형태로 입력된 사용자 발화를 텍스트 형태로 변환된 것을 의미할 수 있다.
- [32] 도 2를 참조하면, 웹을 이용한 정보 검색 방법(S400)은 사용자 질의 및 언어 분석 결과 중 적어도 하나를 제공 받는다(S410).
- [33] 제공받은 상기 사용자 질의 및 상기 언어 분석 결과 중 적어도 하나에 대한 기본 단어 벡터를 생성하고, 상기 기본 단어 벡터를 이용하여 상기 벡터 공간 데이터베이스(2000)에서 상기 기본 단어 벡터에 대응되는 벡터 공간을 검색하는 기본 검색을 수행한다(S420).
- [34] 상기 벡터 공간 데이터베이스(2000)에서 상기 기본 단어 벡터에 대응되는 검색 결과 유무 또는 기본 검색 점수가 미리 설정된 기준 점수 이하인지를 판단한다(S430). 상기 기준 점수는 시스템의 설계에 따라 다양하게 설정될 수 있다.
- [35] 판단 결과 상기 벡터 공간 데이터베이스(2000)에서 상기 기본 단어 벡터에 대응되는 검색 결과가 없거나, 상기 기본 검색 점수가 상기 기준 점수 이하인 경우에 확장 검색을 수행한다(S440).

- [36] 상기 확장 검색 단계(S440)에서는 먼저 상기 사용자 질의 및 상기 언어 분석 결과 중 적어도 하나와 관련된 검색 키워드를 생성하여 웹 검색을 수행한 후, 검색된 웹 문서에 대한 확장 단어 벡터를 생성하고, 상기 생성된 확장 단어 벡터를 이용하여 벡터 공간 데이터베이스(2000)에서 상기 확장 단어 벡터에 대응되는 벡터 공간을 검색한다.
- [37] 상기 기본 검색 단계 또는 상기 확장 검색 단계에서 검색된 상기 벡터 공간에 대응되는 상기 지식 정보 데이터베이스에 저장된 지식 정보를 검색 결과로 출력한다.
- [38] 또한, 웹을 이용한 정보 검색 방법(S400)은 지식 정보 데이터베이스(1000)에 저장된 지식 정보에 대한 벡터 공간(Vector Space)을 생성하여 벡터 공간 데이터베이스(2000)에 저장하는 단계(S450)를 더 포함하여 구성될 수 있다.
- [39] 여기서, 상기 벡터 공간을 생성하여 벡터 공간 데이터베이스에 저장하는 단계는 기본 검색 또는 확장 검색 단계 이전에 수행되는 것이 바람직하나, 여기에 한정되는 것은 아니며 상기 단계 S410 내지 S440의 순서와 무관하게 수행될 수도 있다.
- [40] 상기 기본 단어 벡터는 상기 사용자 질의 및 상기 사용자 질의의 언어 분석 결과 중 적어도 하나에 대한 벡터 공간을 의미하고, 상기 확장 단어 벡터는 상기 사용자 질의 및 상기 사용자 질의의 언어 분석 결과 중 적어도 하나와 관련된 검색 키워드를 생성하여 웹 검색을 통하여 검색된 웹 문서에 포함된 단어들에 대한 벡터 공간을 의미한다.
- [41] 또한, 상기 단어는 하나의 단어, 키워드, 구(句) 등이 될 수 있으며, 상기 단어의 의미는 응용 분야에 따라 달라질 수 있다.
- [42] 상기 사용자 질의 및 언어 분석 결과를 제공받는 단계(S410)는 텍스트 형태의 사용자 질의와, 상술한 화행, 주행 및 구조 정보가 포함될 수 있는 언어 분석 결과를 제공받는 것을 의미할 수 있다.
- [43] 도 3을 참조하면, 상기 기본 검색 단계(S420)는 제공받은 상기 사용자 질의 및 상기 언어 분석 결과 중 적어도 하나에 대한 기본 단어 벡터를 생성하고(S421), 상기 기본 단어 벡터를 이용하여 상기 벡터 공간 데이터베이스(2000)에서 상기 기본 단어 벡터에 대응되는 벡터 공간을 검색할 수 있다(S422). 상기 기본 단어 벡터는 상기 사용자 질의 및 상기 언어 분석 결과 중 적어도 하나를 사용하여 형성된 벡터 공간을 의미할 수 있다.
- [44] 상기 판단하는 단계(S430)는 상기 벡터 공간 데이터베이스(2000)에 저장된 벡터 공간 중에 상기 기본 단어 벡터와 대응되는 벡터 공간의 유무 또는 기본 검색 점수가 미리 설정된 기준 점수 이하인지를 판단하는 기능을 수행할 수 있다. 또한, 상기 판단하는 단계(S430)는 확장 검색 단계가 웹을 이용한 검색을 수행하므로 응답시간이 상기 기본 검색 단계보다 늦기 때문에 상기 기본 검색 단계를 통하여 처리할 수 있는 경우는 상기 기본 검색 단계만을 수행할 수 있도록 판단하는 기능을 수행할 수 있다. 상기 기본 검색 점수는 상기 기본 단어

벡터와 상기 벡터 공간 데이터베이스(2000)에 저장된 벡터 공간과의 대응되는 정도를 수치화한 것을 의미할 수 있고, 상기 기본 검색 점수는 코사인 유사도(Cosine Similarity) 등의 다양한 점수 계산 방법을 통해 산출될 수 있다. 또한, 상기 코사인 유사도는 수학적 식 1을 통해 산출될 수 있다.

[45] 수학적 식 1

$$S = \frac{A \cdot B}{\|A\| \|B\|}$$

[46] 상기의 수학적 식 1에서, S는 코사인 유사도, A와 B는 벡터 공간, \|A\|와 \|B\|는 벡터 공간의 절대값(크기)을 의미한다.

[47] 또한, 상기 언어 분석 결과에서 나타난 기본 단어 벡터의 구조 정보와 상기 벡터 공간 데이터베이스(2000)에 저장된 벡터 공간의 구조 정보가 동일한지 여부에 따라 차등된 가중치를 부여하여 대응되는 정도를 더 높일 수 있다.

[48] 예를 들어, 사용자의 질의가 "아시안컵 축구 좀 보자", 사용자 질의에 대한 언어 분석 결과가 "TITLE: 아시안컵 축구"인 경우에 기본 단어 벡터는 <(TITLE: 아시안컵, 0.5)>, <(TITLE: 축구, 0.5)>로 생성되고, 상기 벡터 공간 데이터베이스(2000)에 저장된 벡터 공간의 각 필드가 <(TITLE: 아시안컵, 0.2)>, <(CATEGORY: 축구, 0.3)>인 경우에 코사인 유사도를 사용하면, 기본 검색 점수가 $(0.5 \times 0.2 + 0.5 \times 0.3) / (0.707 + 0.361) = 0.234$ 로 계산되고, 만약, 구조 정보에 대한 가중치를 구조 정보가 같은 경우에는 0.7, 다른 경우에는 0.3으로 설정한 경우에 상기 구조 정보를 고려한 기본 검색 점수는 $(0.5 \times 0.2 \times 0.7 + 0.5 \times 0.3 \times 0.3) / (0.707 + 0.361) = 0.108$ 로 계산된다.

[49] 상기 확장 검색 단계(S440)는 벡터 공간 데이터베이스(2000)에서 상기 기본 단어 벡터와 대응되는 벡터 공간이 없거나, 상기 기본 검색 점수가 미리 설정된 기준 점수 이하인 경우, 상기 사용자 질의 및 상기 언어 분석 결과 중 적어도 하나와 관련된 검색 키워드를 생성(S441)하여, 웹 검색을 수행하고(S443), 검색된 웹 문서에 대한 확장 단어 벡터를 생성하고(S444), 상기 확장 단어 벡터를 이용하여 벡터 공간 데이터베이스(2000)에 저장된 벡터 공간 중에 상기 확장 단어 벡터와 대응되는 벡터 공간을 검색할 수 있다(S445). 상기 검색 키워드는 상기 사용자 질의 및 상기 언어 분석 결과 중 적어도 하나의 중요한 필드를 상기 검색 키워드로 사용할 수 있으며, 상기 중요한 필드의 결정은 사용자가 가장 많이 찾고 대표성이 있는 필드일 수 있으며, 휴리스틱(Heuristic)하게 결정될 수 있다. 여기서, 상기 사용자 질의를 검색 키워드의 생성에 고려하는 이유는 중요한 단어임에도 언어 분석 결과에 나타나지 않을 수 있는 경우를 고려하기 위한 것이다. 예를 들어, 사용자 질의가 "재미있는 영화 좀 보자"인 경우에 "재미있는"은 중요한 단어임에도 언어 분석 결과에 나타나지 않을 수 있기 때문이다.

[50] 상기 검색 키워드를 사용하여 웹 검색을 수행하고(S443), 검색된 웹 문서에

대한 확장 단어 벡터를 생성할 수 있다(S444). 상기 확장 단어 벡터는 검색된 웹 문서에서 각각의 단어에 대해 벡터 공간을 형성하는 것을 의미할 수 있다. 상기 확장 단어 벡터를 사용하여 상기 벡터 공간 데이터베이스(2000)에 저장된 벡터 공간 중에 상기 확장 단어 벡터와 대응되는 벡터 공간을 검색할 수 있다. 상기 대응되는 벡터 공간인지의 판단은 확장 검색 점수가 미리 설정된 기준 점수 이상인 경우에 대응되는 벡터 공간으로 판단할 수 있으며, 상기 확장 검색 점수가 높은 순으로 상기 검색된 벡터 공간에 대응되는 정보를 정렬시킬 수 있다.

[51] 상기 확장 검색 점수는 상기 확장 단어 벡터와 상기 벡터 공간 데이터베이스에 저장된 벡터 공간과의 유사도로 산출될 수 있으며, 상기 확장 단어 벡터와 상기 벡터 공간 데이터베이스에 저장된 벡터 공간의 형성 방식에 따라 가중치를 차등적으로 설정할 수 있다.

[52] 구체적으로, 상기 확장 검색 점수는 수학식 2를 통하여 산출될 수 있다.

[53] 수학식 2

$$\text{score}(q, t_i) = \sum_{k \in \{1,2\}} \alpha_k \cdot f(\vec{q}_k, \vec{t}_i) + \sum_{k \in \{1,2\}} \beta_k \cdot f(\vec{q}_k^w, \vec{t}_i^w)$$

[54] 상기 수학식 2에서

\vec{q}_1

는 웹 검색 전의 언어 분석 결과로부터 생성한 기본 단어 벡터,

\vec{q}_2

는 웹 검색 전의 사용자 질의로부터 생성한 기본 단어 벡터,

\vec{q}_1^w

는 웹 검색 후의 언어 분석 결과로부터 생성한 확장 단어 벡터,

\vec{q}_2^w

는 웹 검색 후의 사용자 질의로부터 생성한 확장 단어 벡터를 의미한다.

[55] 또한,

\vec{t}_i^z

는 벡터 공간 기본 데이터베이스(2100)에 저장된 지식 정보 데이터베이스(1000)에 저장된 지식 정보의 각 필드에 대한 벡터 공간,

\vec{t}_i^w

는 벡터 공간 확장 데이터베이스(2200)에서 검색된 문서의 각 필드에 대한 단어 벡터를 의미한다.

[56] 또한,

α_1

는 언어 분석 결과로부터 생성된 확장 단어 벡터가 벡터 공간 기본 데이터베이스(2100)에 저장된 벡터 공간에 대응되는 경우의 가중치이고,

α_2

는 사용자 질의로부터 생성된 확장 단어 벡터가 벡터 공간 기본 데이터베이스(2100)에 저장된 벡터 공간에 대응되는 경우의 가중치이고,

β_1

는 언어 분석 결과로부터 생성된 확장 단어 벡터가 벡터 공간 확장 데이터베이스(2200)에 저장된 벡터 공간에 대응되는 경우의 가중치이고,

β_2

는 사용자 질의로부터 생성된 확장 단어 벡터가 벡터 공간 확장 데이터베이스(2200)에 저장된 벡터 공간에 대응되는 경우의 가중치이고, 상기 가중치들은 확장 단어 벡터와 벡터 공간 데이터베이스의 벡터 공간의 형성 방식에 따라 변경될 수 있고, 차등적으로 설정될 수 있다.

[57] 예를 들어, 웹 검색을 거치지 않은 벡터 공간 기본 데이터베이스(2100)에 저장된 벡터 공간의 신뢰도가 웹 검색을 거친 벡터 공간 확장 데이터베이스(2200)에 저장된 벡터 공간의 신뢰도보다 높다고 볼 수 있기 때문에 α_1, α_2 가 β_1, β_2 보다 큰 값으로 가중치를 설정할 수 있다.

[58] 또한, 언어 분석 결과에서의 키워드가 사용자 질의에서의 키워드보다 중요하다고 볼 수 있기 때문에 α_1 는 α_2, β_1 는 β_2 보다 큰 값으로 설정될 수 있다.

[59] 상기 벡터 공간을 생성하여 벡터 공간 데이터베이스(2000)에 저장하는 단계(S450)는 지식 정보 데이터베이스(1000)에 저장된 지식 정보를 벡터 공간으로 형성한 기본 벡터 공간을 벡터 공간 기본 데이터베이스(2100)에 저장하는 단계(S451)와, 지식 정보 데이터베이스(1000)에 저장된 지식 정보에서 키워드를 추출하고, 상기 추출된 키워드로 웹 검색을 수행(S452)하여 검색된 웹 문서로부터 확장 벡터 공간을 형성하고, 상기 확장 벡터 공간을 벡터 공간 확장 데이터베이스(2200)에 저장하는 단계(S453)를 포함하여 구성될 수 있다.

[60] 상기 기본 벡터 공간은 지식 정보 데이터베이스(1000)에 저장된 문서로부터 각각의 단어들에 대해 미리 저장하여 생성된 벡터 공간(Vector Space), 구조 정보를 포함할 수 있다.

[61] 예를 들어, 지식 정보 데이터베이스(1000)의 엔트리 ID 10번에 "박지성 맨체스터 유나이티드 경기"라는 문서가 저장된 경우에 <박지성, 10, 0.3>, <맨체스터, 10, 0.2>, <유나이티드, 10, 0.1>, <경기, 10, 0.1>의 기본 벡터 공간이 벡터 공간 기본 데이터베이스(2100)에 저장될 수 있다. 여기서, "박지성"은

단어를 의미하고, "10"은 단어가 저장된 엔트리 ID를 의미하고, "0.3"은 단어에 대한 가중치를 의미할 수 있다. 상기 구조 정보는 단어가 속한 필드 정보를 의미할 수 있고, 상기 구조 정보는 예를 들어, 지식 정보 데이터베이스(1000)에서는 컬럼(column) 명이 될 수 있고, 사용자 질의에 대한 언어 분석 결과에서는 개체명이 될 수 있다. 또한, 상기 구조 정보는 상기 기본 검색 단계(S420)에서 사용되는 기본 검색 점수에 추가 정보로서 사용될 수 있다.

- [62] 예를 들어, 지식 정보 데이터베이스(1000)에서 "출연진" 정보와 언어 분석 결과에서 "출연진"으로 검색된 경우에 보다 높은 가중치를 줄 수 있다.
- [63] 상기 지식 정보 데이터베이스(1000)에 저장된 지식 정보의 각 단어에 대한 상기 가중치는 단어 빈도(TF: Term Frequency)와 역 문서 빈도(IDF: Inverse Document Frequency)의 곱 연산인 수학적 식 3을 통하여 산출될 수 있다.

[64] 수학적 식 3

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}}, \quad idf_i = \log \frac{|D|}{|\{d_j : t_i \in d_j\}|}, \quad tfidf_{ij} = tf_{i,j} \cdot idf_i$$

[65] 상기 수학적 식 3에서 단어 빈도(

$tf_{i,j}$

)는 문서(

d_j

)에서 특정 단어가 나타난 개수(

$n_{i,j}$

)를 문서(

d_j

)의 전체 단어 개수(

$$\sum_k n_{k,j}$$

)로 나누어 계산되며, 역 문서 빈도(

idf_i

)는 전체 문서의 개수(

$|D|$

)를 단어가 나타난 문서 개수로 나눈 것(

$$|\{d_j : t_i \in d_j\}|$$

)의 로그 값으로 계산된다.

[66] 상기 벡터 공간 확장 데이터베이스(2200)는 상기 지식 정보

데이터베이스(1000)에 저장된 지식 정보에서 키워드를 추출하고, 상기 추출된 키워드를 사용하여 웹을 통하여 검색을 수행하고(S452), 검색된 문서에서 각각의 단어에 대해 확장 벡터 공간을 형성할 수 있다(S453).

- [67] 상기 키워드로는 지식 정보 데이터베이스(1000)에 저장된 지식 정보의 중요한 필드를 사용할 수 있으며, 상기 중요한 필드의 결정은 사용자가 가장 많이 찾고, 대표성이 있는 필드일 수 있으며, 휴리스틱(Heuristic)하게 결정될 수 있다.
- [68] 예를 들어, TV 일정 도메인의 경우 프로그램 제목, 도서 검색의 경우 책의 제목이 중요한 필드로서 키워드로 사용될 수 있다. 웹 문서는 지식 정보 데이터베이스(1000)에 저장된 지식 정보에서 부족한 정보를 포함할 수 있어 상기 지식 정보를 보충할 수 있다. 예를 들어, 지식 정보 데이터베이스에 저장된 지식 정보인 "아시안컵"에 관한 정보는 웹 검색을 통하여 "조광래", "박지성" 등의 추가 정보를 획득하여 확장될 수 있다.
- [69] 상기에서는 본 발명의 바람직한 실시예를 참조하여 설명하였지만, 해당 기술 분야의 숙련된 당업자는 하기의 특허 청구의 범위에 기재된 본 발명의 사상 및 영역으로부터 벗어나지 않는 범위 내에서 본 발명을 다양하게 수정 및 변경시킬 수 있음을 이해할 수 있을 것이다.

[70]

[71]

청구범위

- [청구항 1] 정보 검색 시스템의 웹을 이용한 정보 검색 방법에 있어서,
 제공된 사용자 질의 및 언어 분석 결과 중 적어도 하나에 대한 기본
 단어 백터를 생성하고, 상기 기본 단어 백터를 이용하여 백터 공간
 데이터베이스에서 상기 기본 단어 백터에 대응되는 백터 공간을
 검색하는 기본 검색 단계;
 상기 기본 단어 백터와 검색된 상기 백터 공간과의 유사도를
 판단하는 단계;
 상기 유사도가 미리 설정된 기준 이하인 경우, 상기 사용자 질의 및
 상기 언어 분석 결과 중 적어도 하나를 이용하여 수행된 웹 검색
 결과로부터, 확장 단어 백터를 생성하고, 생성된 상기 확장 단어
 백터를 이용하여 상기 백터 공간 데이터베이스에서 상기 확장
 단어 백터에 대응되는 백터 공간을 검색하는 확장 검색 단계; 및
 상기 기본 검색 단계 또는 상기 확장 검색 단계에서 검색된 상기
 백터 공간에 기초하여 지식 정보를 검색하는 단계를 포함하는
 웹을 이용한 정보 검색 방법.
- [청구항 2] 제1항에 있어서,
 상기 기본 단어 백터, 상기 확장 단어 백터 및 상기 백터 공간
 데이터베이스에 저장된 백터 공간은 단어, 가중치 정보 필드 및
 구조 정보 필드 중 적어도 하나를 포함하는 것을 특징으로 하는
 웹을 이용한 정보 검색 방법.
- [청구항 3] 제1항에 있어서,
 상기 백터 공간 데이터베이스는,
 상기 지식 정보를 백터 공간으로 형성한 백터 공간 기본
 데이터베이스; 및
 상기 지식 정보를 이용하여 수행된 웹 검색 결과로부터 백터
 공간을 형성한 백터 공간 확장 데이터베이스를 포함하는 것을
 특징으로 하는 웹을 이용한 정보 검색 방법.
- [청구항 4] 제1항에 있어서,
 상기 지식 정보를 검색하는 단계는,
 상기 확장 단어 백터와 상기 검색된 적어도 하나의 백터 공간과의
 확장 검색 점수를 산출하고, 상기 확장 검색 점수에 따라 정렬하는
 단계를 포함하는 것을 특징으로 하는 웹을 이용한 정보 검색 방법.
- [청구항 5] 제4항에 있어서,
 상기 확장 검색 점수의 산출시에,
 상기 언어 분석 결과 또는 상기 사용자 질의로부터 생성되는 상기
 확장 단어 백터 및 백터 공간 기본 데이터베이스와 백터 공간 확장

데이터베이스에 저장된 상기 벡터 공간에 따라 가중치를 차등적으로 설정하는 것을 특징으로 하는 웹을 이용한 정보 검색 방법.

[청구항 6]

제5항에 있어서,

상기 가중치는,

상기 언어 분석 결과로부터 생성된 확장 단어 벡터가 상기 벡터 공간 기본 데이터베이스에 저장된 벡터 공간에 대응되는 경우 및 상기 사용자 질의로부터 생성된 확장 단어 벡터가 상기 벡터 공간 기본 데이터베이스에 저장된 벡터 공간에 대응되는 경우가, 상기 언어 분석 결과로부터 생성된 확장 단어 벡터가 벡터 공간 확장 데이터베이스에 저장된 벡터 공간에 대응되는 경우 및 상기 사용자 질의로부터 생성된 확장 단어 벡터가 벡터 공간 확장 데이터베이스에 저장된 벡터 공간에 대응되는 경우보다 큰 값으로 설정되는 것을 특징으로 하는 웹을 이용한 정보 검색 방법.

[청구항 7]

제5항에 있어서,

상기 가중치는,

상기 언어 분석 결과로부터 생성된 확장 단어 벡터가 상기 벡터 공간 기본 데이터베이스에 저장된 벡터 공간에 대응되는 경우에는 상기 사용자 질의로부터 생성된 확장 단어 벡터가 상기 벡터 공간 기본 데이터베이스에 저장된 벡터 공간에 대응되는 경우보다 큰 값으로 설정되고, 상기 언어 분석 결과로부터 생성된 확장 단어 벡터가 벡터 공간 확장 데이터베이스에 저장된 벡터 공간에 대응되는 경우에는 상기 사용자 질의로부터 생성된 확장 단어 벡터가 벡터 공간 확장 데이터베이스에 저장된 벡터 공간에 대응되는 경우보다 큰 값으로 설정되는 것을 특징으로 하는 웹을 이용한 정보 검색 방법.

[청구항 8]

정보 검색 시스템의 웹을 이용한 정보 검색 방법을 사용하는 음성 대화 방법에 있어서,

제공된 음성 형태의 사용자 질의를 텍스트 형태로 변환하는 음성 인식 단계;

상기 텍스트 형태로 변환된 사용자 질의에 언어 분석을 수행하는 언어 분석 단계;

변환된 상기 사용자 질의 및 상기 언어 분석 결과 중 적어도 하나에 대한 기본 단어 벡터를 생성하고, 상기 기본 단어 벡터를 이용하여 벡터 공간 데이터베이스에서 상기 기본 단어 벡터에 대응되는 벡터 공간을 검색한 후, 상기 기본 단어 벡터와 검색된 상기 벡터 공간과의 유사도가 미리 설정된 기준 이하인 경우, 상기 사용자 질의 및 상기 언어 분석 결과 중 적어도 하나를 이용하여 수행된 웹

검색 결과로부터, 생성한 확장 단어 벡터를 이용하여 상기 벡터 공간 데이터베이스에서 상기 확장 단어 벡터에 대응되는 벡터 공간을 검색하고, 상기 기본 단어 벡터 또는 상기 확장 단어 벡터에 대응되어 검색된 상기 벡터 공간에 기초하여 지식 정보를 검색하는 웹을 이용한 정보 검색 단계;
 검색된 상기 지식 정보를 상기 사용자의 질의에 대응되는 응답으로 변환하는 응답 생성 단계; 및
 상기 응답을 음성 형태로 변환하는 음성 합성 단계를 포함하는 웹을 이용한 정보 검색 방법을 사용한 음성 대화 방법.

[청구항 9]

제8항에 있어서,
 상기 기본 단어 벡터, 상기 확장 단어 벡터 및 상기 벡터 공간 데이터베이스에 저장된 벡터 공간은 단어, 가중치 정보 필드 및 구조 정보 필드 중 적어도 하나를 포함하는 것을 특징으로 하는 웹을 이용한 정보 검색 방법을 사용한 음성 대화 방법.

[청구항 10]

제8항에 있어서,
 상기 벡터 공간 데이터베이스는,
 상기 지식 정보를 벡터 공간으로 형성한 벡터 공간 기본 데이터베이스; 및
 상기 지식 정보를 이용하여 수행된 웹 검색 결과로부터 벡터 공간을 형성한 벡터 공간 확장 데이터베이스를 포함하는 것을 특징으로 하는 웹을 이용한 정보 검색 방법을 사용한 음성 대화 방법.

[청구항 11]

제8항에 있어서,
 상기 지식 정보를 검색하는 단계는,
 상기 확장 단어 벡터와 상기 검색된 적어도 하나의 벡터 공간과의 확장 검색 점수를 산출하고, 상기 확장 검색 점수에 따라 정렬하는 단계를 포함하는 것을 특징으로 하는 웹을 이용한 정보 검색 방법을 사용한 음성 대화 방법.

[청구항 12]

제11항에 있어서,
 상기 확장 검색 점수의 산출시에,
 상기 언어 분석 결과 또는 상기 사용자 질의로부터 생성되는 상기 확장 단어 벡터 및 벡터 공간 기본 데이터베이스와 벡터 공간 확장 데이터베이스에 저장된 상기 벡터 공간에 따라 가중치를 차등적으로 설정하는 것을 특징으로 하는 웹을 이용한 정보 검색 방법을 사용한 음성 대화 방법.

[청구항 13]

제12항에 있어서,
 상기 가중치는,
 상기 언어 분석 결과로부터 생성된 확장 단어 벡터가 상기 벡터

공간 기본 데이터베이스에 저장된 벡터 공간에 대응되는 경우 및 상기 사용자 질의로부터 생성된 확장 단어 벡터가 상기 벡터 공간 기본 데이터베이스에 저장된 벡터 공간에 대응되는 경우가, 상기 언어 분석 결과로부터 생성된 확장 단어 벡터가 벡터 공간 확장 데이터베이스에 저장된 벡터 공간에 대응되는 경우 및 상기 사용자 질의로부터 생성된 확장 단어 벡터가 벡터 공간 확장 데이터베이스에 저장된 벡터 공간에 대응되는 경우 보다 큰 값으로 설정되는 것을 특징으로 하는 웹을 이용한 정보 검색 방법을 사용한 음성 대화 방법.

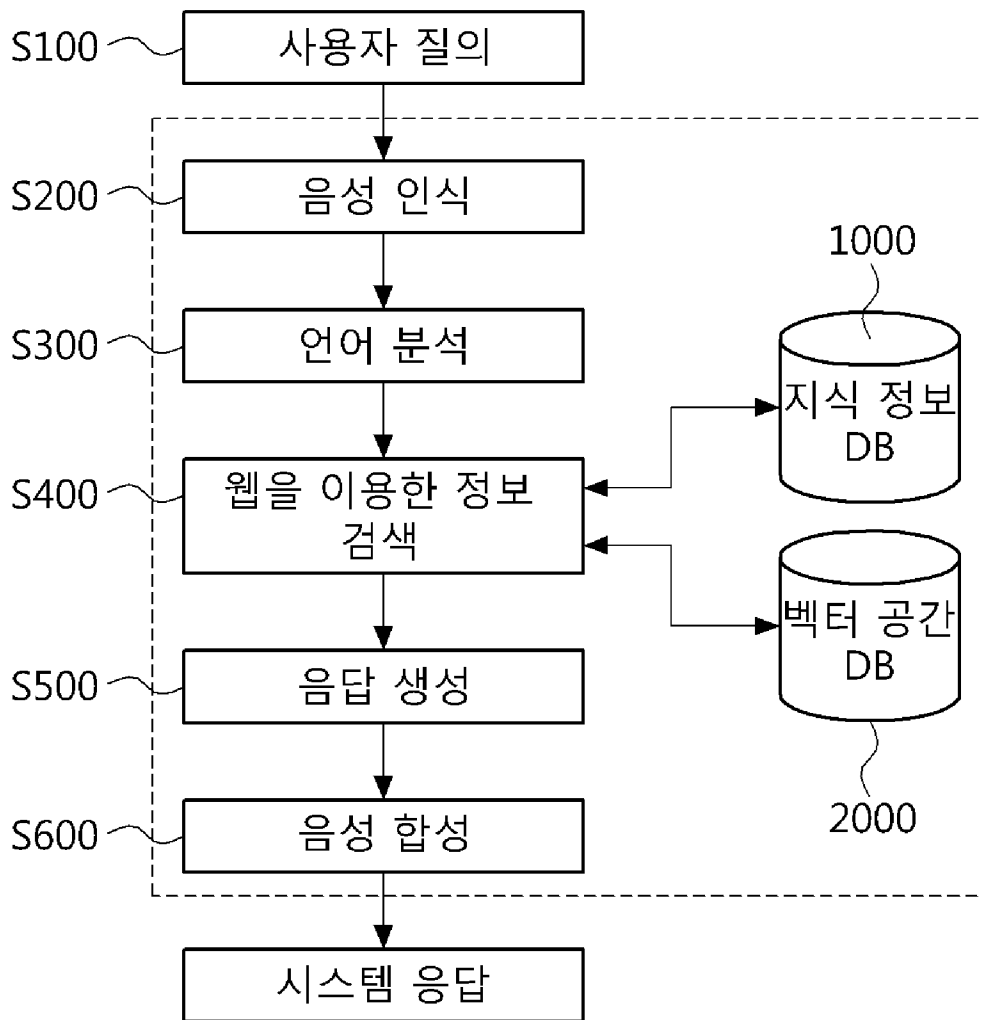
[청구항 14]

제12항에 있어서,

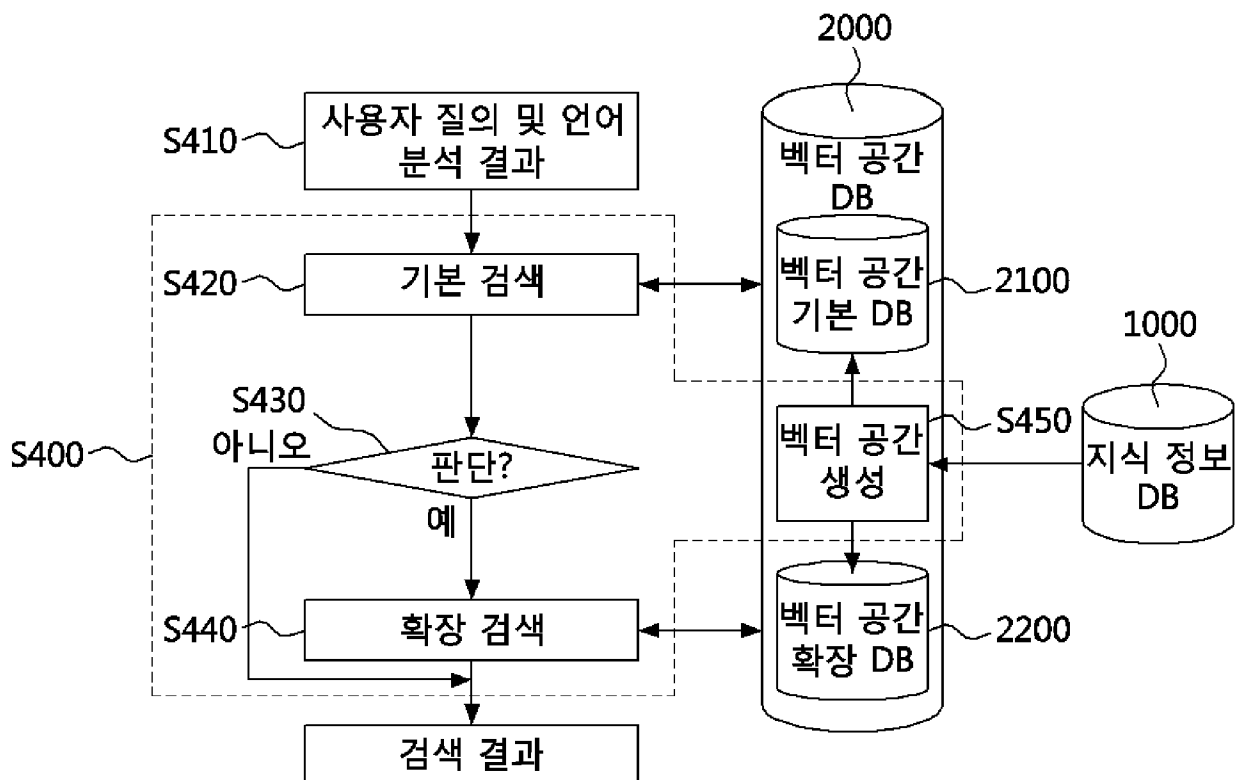
상기 가중치는,

상기 언어 분석 결과로부터 생성된 확장 단어 벡터가 상기 벡터 공간 기본 데이터베이스에 저장된 벡터 공간에 대응되는 경우에는 상기 사용자 질의로부터 생성된 확장 단어 벡터가 상기 벡터 공간 기본 데이터베이스에 저장된 벡터 공간에 대응되는 경우보다 큰 값으로 설정되고, 상기 언어 분석 결과로부터 생성된 확장 단어 벡터가 벡터 공간 확장 데이터베이스에 저장된 벡터 공간에 대응되는 경우에는 상기 사용자 질의로부터 생성된 확장 단어 벡터가 벡터 공간 확장 데이터베이스에 저장된 벡터 공간에 대응되는 경우보다 큰 값으로 설정되는 것을 특징으로 하는 웹을 이용한 정보 검색 방법을 사용한 음성 대화 방법.

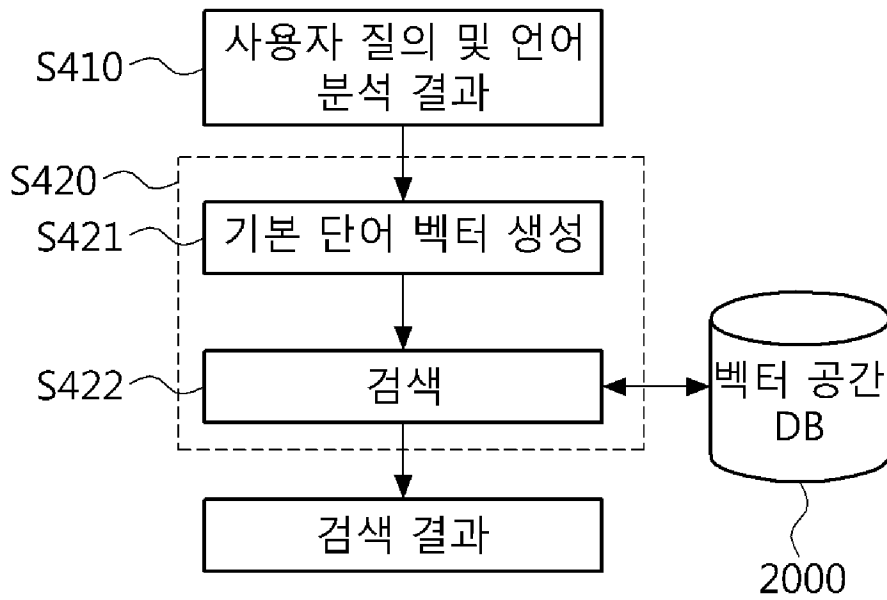
[Fig. 1]



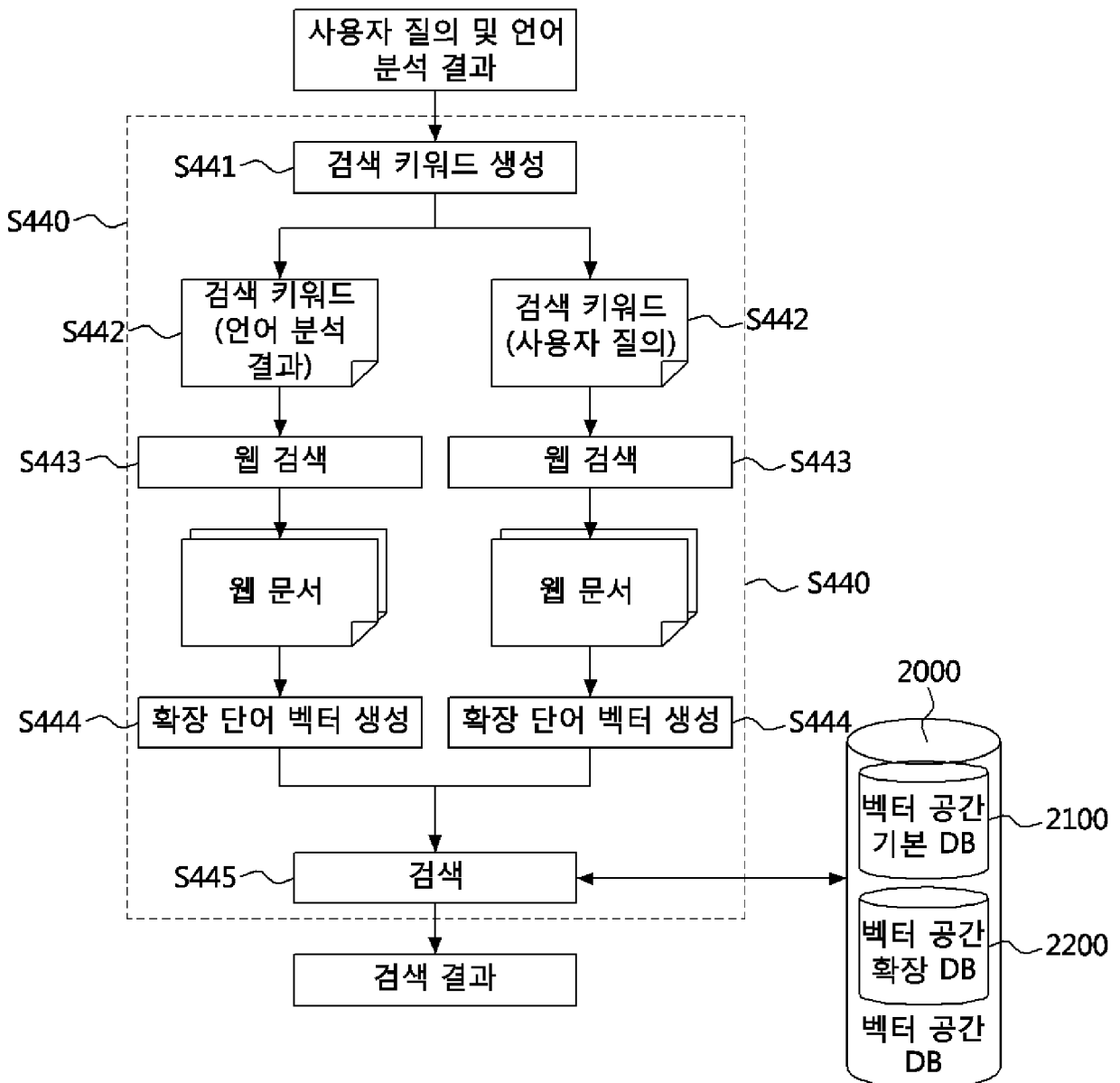
[Fig. 2]



[Fig. 3]



[Fig. 4]



[Fig. 5]

