

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국



(43) 국제공개일
2015년 5월 14일 (14.05.2015) WIPO | PCT

(10) 국제공개번호

WO 2015/068947 A1

(51) 국제특허분류:

G10L 15/28 (2013.01) G10L 15/08 (2006.01)

(21) 국제출원번호:

PCT/KR2014/008706

(22) 국제출원일:

2014년 9월 18일 (18.09.2014)

(25) 출원언어:

한국어

(26) 공개언어:

한국어

(30) 우선권정보:

10-2013-0134246 2013년 11월 6일 (06.11.2013) KR

(71) 출원인: 주식회사 시스트란인터내셔널 (SYSTRAN INTERNATIONAL CO., LTD.) [KR/KR]; 135-855 서울시 강남구 양재천로 163, 5층, Seoul (KR).

(72) 발명자: 지창진 (JI, Chang Jin); 156-080 서울시 동작구 동작대로 39길 46-2, B동 403호, Seoul (KR).

(74) 대리인: 유미특허법인 (YOU ME PATENT AND LAW FIRM); 135-912 서울시 강남구 테헤란로 115, Seoul (KR).

(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO,

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

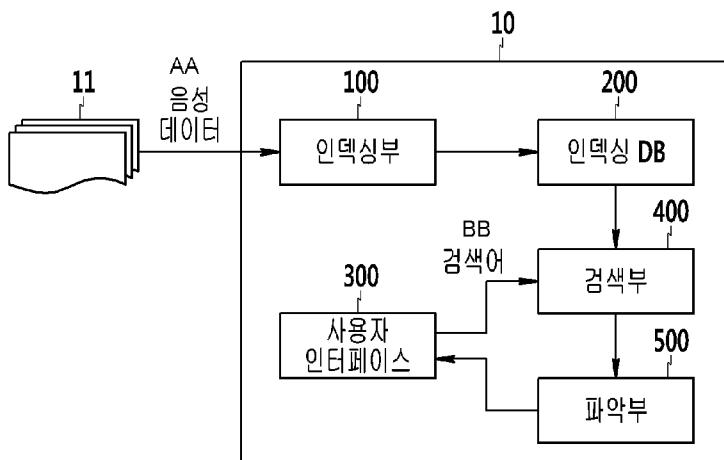
(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS,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

공개:

— 국제조사보고서와 함께 (조약 제 21 조(3))

(54) Title: SYSTEM FOR ANALYZING SPEECH CONTENT ON BASIS OF EXTRACTION OF KEYWORDS FROM RECORDED VOICE DATA, INDEXING METHOD USING SYSTEM AND METHOD FOR ANALYZING SPEECH CONTENT

(54) 발명의 명칭: 녹취된 음성 데이터에 대한 핵심어 추출 기반 발화 내용 파악 시스템과, 이 시스템을 이용한 인덱싱 방법 및 발화 내용 파악 방법



100 ... Indexing unit

200 ... Indexing DB

300 ... User interface

400 ... Searching unit

500 ... Analyzing unit

AA ... Voice data

BB ... Search word

the voice data.

(57) 요약서:

(57) Abstract: Disclosed are a system for analyzing speech content on the basis of the extraction of keywords from recorded voice data, an indexing method using the system and a method for analyzing speech content. An indexing unit of the system receives the voice data and forms a phoneme lattice by performing a phoneme-based voice recognition by frame unit, generates indexing information partitioned for a time limit frame comprising the plurality of frames and then stores the indexing information in an indexing database, wherein the partitioned indexing information includes the phoneme lattices formed for each time limit frame. A searching unit searches a phoneme string matching a search word through a phoneme-based comparison on the partitioned indexing information stored in the indexing database by using, as the search word, the keyword inputted from a user, and finds out a voice portion corresponding to the search word through a precise acoustical analysis on the matching phoneme string. An analyzing unit analyzes a subject word through the search result searched by the searching unit and then outputs the subject word to the user so as to enable the user to understand the speech content of

[다음 쪽 계속]



녹취된 음성 데이터에 대한 핵심어 추출 기반 발화 내용 파악 시스템과, 이 시스템을 이용한 인텍싱 방법 및 발화 내용 파악 방법이 개시된다. 이 시스템의 인텍싱부는 음성 데이터를 입력받아서 프레임 단위로 음소 기준의 음성 인식을 수행하여 음소 격자를 형성하고, 복수의 프레임으로 구성되는 제한 시간의 프레임에 대해 분할된 인텍싱 정보 – 여기서 분할된 인텍싱 정보는 제한 시간의 프레임별로 형성되는 음소 격자를 포함함 – 를 생성하여 인텍싱 데이터베이스에 저장한다. 검색부는 사용자로부터 입력되는 핵심어를 검색어로 하여 인텍싱 데이터베이스에 저장된 분할된 인텍싱 정보에 대해 음소 기준의 비교를 통해 상기 검색어와 일치하는 음소열을 검색하고 일치하는 음소열에 대해 정밀한 음향학적 분석을 통해 검색어에 해당하는 음성부분을 찾아내고, 파악부는 상기 검색부에 의해 검색되는 검색 결과를 통해 주제어를 파악하여 상기 음성 데이터의 발화 내용을 파악할 수 있도록 사용자에게 출력한다.

명세서

발명의 명칭: 녹취된 음성 데이터에 대한 핵심어 추출 기반 발화 내용 파악 시스템과, 이 시스템을 이용한 인덱싱 방법 및 발화 내용 파악 방법

기술분야

[1] 본 발명은 녹취된 음성 데이터에 대한 핵심어 추출 기반 발화 내용 파악 시스템과, 이 시스템을 이용한 인덱싱 방법 및 발화 내용 파악 방법에 관한 것이다.

배경기술

[2] 최근 업체들의 고객 만족 서비스를 위해 고객과의 대화를 녹취하거나 또는 전자 상거래를 위해 전화 통화 내용을 녹취한 후 녹취된 음성 데이터로부터 음성 인식을 수행하여 전체적인 발화 내용을 파악하고자 하는 녹취 대상 검색의 필요성이 대두되고 있다.

[3] 그러나, 고객의 녹취 음성의 빅데이터를 처리할 수 있는 기본 가공 프로그램이나 서비스가 개발되지 못하였고, 이러한 빅데이터를 수작업으로 청취하기에는 인력이나 시간이 많이 소요되기 때문에 이러한 녹취 대상 검색이 제대로 수행되지 못하고 있는 실정이다.

[4] 또한, 음성 인식 기술에 기반하여 녹취 대상 검색을 수행하고자 하는 시도가 있으나, 이러한 음성 인식 기술은 애플의 '쉬리'나 구글의 '어시스턴트'처럼 사용자의 실시간 입력을 음성 인식 엔진을 통해 텍스트로 출력하고, 이러한 텍스트 출력에 대해 인공지능적 접근법을 통해 의미를 추출하는 인공지능 음성 인식 서비스를 지향하고 있다. 따라서, 녹취된 음성 데이터가 정확하게 텍스트로 변환되지 않는 경우에는 녹취 대상 검색이 정확하게 이루어질 수가 없어 녹취된 발화 내용을 정확하게 파악하기가 어렵다는 문제점이 있다.

발명의 상세한 설명

기술적 과제

[5] 본 발명이 이루고자 하는 기술적 과제는 녹취된 음성 데이터에 대해 음소 기준의 음성 인식을 통해 인덱싱된 데이터를 저장하고 이를 통해 핵심어 기반으로 발화 내용을 파악함으로써 발화 내용 파악이 제대로 수행되며 간단하고 빨라지는 발화 내용 파악 시스템과, 이 시스템을 이용한 인덱싱 방법 및 발화 내용 파악 방법을 제공하는 것이다.

과제 해결 수단

[6] 본 발명의 한 특징에 따른 발화 내용 파악 시스템은, 음성 데이터를 입력받아서 프레임 단위로 음소 기준의 음성 인식을 수행하여 음소 격자를 형성하고, 복수의 프레임으로 구성되는 제한 시간의 프레임에 대해

분할된 인덱싱 정보—여기서 분할된 인덱싱 정보는 제한 시간의 프레임별로 형성되는 음소 격자를 포함함—를 생성하는 인덱싱부; 상기 인덱싱부에 의해 생성되는 분할된 인덱싱 정보를 분할된 인덱싱 정보별로 인덱싱 가능하도록 각각 저장하는 인덱싱 데이터베이스; 사용자로부터 입력되는 핵심어를 검색어로 하여 인덱싱 데이터베이스에 저장된 분할된 인덱싱 정보에 대해 음소 기준의 비교를 통해 상기 검색어와 일치하는 음소열을 검색하는 검색부; 및 상기 검색부에 의해 검색되는 검색 결과를 통해 주제어를 파악하여 상기 음성 데이터의 발화 내용을 파악할 수 있도록 사용자에게 출력하는 파악부를 포함한다.

- [8] 여기서, 상기 인덱싱부는, 프레임 단위의 음성 데이터로부터 특정 벡터를 추출하는 특징 벡터 추출부; 상기 특징 벡터 추출부에 의해 추출되는 특징 벡터를 사용하여 프레임 동기에 기준한 음소 인식을 수행하여 대응되는 음소열을 생성하는 음소 인식부; 상기 음소 인식부에 의해 생성되는 음소열을 입력받아서 프레임 단위의 시간적 흐름에 따라 음소 인식의 후보군들을 생성하는 후보군 형성부; 상기 후보군 형성부에 의해 생성되는 음소열 후보군으로부터 역시간적인 연산을 수행하여 1개의 음소열 후보군을 선택하여 대응되는 음소 격자를 형성하는 음소 격자 형성부; 및 상기 특징 벡터 추출부, 상기 음소 인식부, 상기 후보군 형성부 및 상기 음소 격자 형성부를 제어하여, 음성 데이터 전체에 대해 상기 제한 시간별로 또한 상기 제한 시간 내에서 프레임 단위별로 음소 단위의 격자가 형성되도록 제어를 수행하고, 이렇게 형성되는 음소 격자를 각 제한 시간별로 인덱싱될 수 있도록 제한 시간별로 분할된 인덱싱 정보로써 상기 인덱싱 데이터베이스에 저장되도록 제어를 수행하는 인덱싱 제어부를 포함한다.
- [9] 또한, 상기 인덱싱 제어부는, 음성 데이터에서 음성이 검출되는지의 여부를 나타내는 음성 검출기; 음성 데이터에 대해 음소 인식이 수행되는 음소의 시간적 위치를 카운트하는 시간 카운터; 상기 음성 검출기에 의해 음성 구간이 검출되는 때에 상기 시간 카운터로부터 해당 음성 구간이 검출되는 때부터 시간을 카운트하여 상기 제한 시간을 카운트하는 시간 제한기; 및 상기 시간 제한기에 의해 카운트되는 제한 시간 내에 상기 음성 검출기에 의해 검출되는 유효한 음성 구간에 대해 프레임 단위로 음소 인식을 수행하여 음소 격자를 형성하여 분할된 인덱싱 정보로써 상기 인덱싱 데이터베이스에 저장하는 제어를 수행하는 동작 제어기를 포함한다.
- [10] 또한, 상기 동작 제어기는 상기 음성 검출기에 의해 검출되는 유효한 음성 구간에 대해 상기 시간 제한기에 의해 카운트되는 상기 제한 시간이 도파하는 경우 이전 음성 구간 중 특정 시간 또는 특정 프레임의 음성 구간을 중첩하여 해당 프레임부터 새로운 제한 시간에 대응되는 음성 구간으로써 음소 인식을 수행하도록 제어를 수행하는 것을 특징으로 한다.
- [11] 또한, 상기 검색부는, 사용자로부터 입력되는 검색어에 대해 이미 처리된 검색

결과를 찾아주는 검색결과 히스토리 검출부; 검색어에 대응되는 음소 단위의 발음열을 생성하는 발음열 생성기; 상기 검색어와 상기 검색어에 대응되는 복수의 문맥 키워드를 저장하는 검색어 데이터베이스; 상기 발음열 생성기에서 생성되는 발음열을 사용하여 상기 인덱싱 데이터베이스에 저장되어 있는 분할된 인덱싱 정보에 대해 일치하는 음소열을 검색하여 1차 검색 결과 후보의 음성구간을 선정하는 동적 정합 처리기; 및 상기 동적 정합 처리기에서 선정되는 1차 검색 결과 후보의 음성구간에 대해서 음향학적 모델을 통해 일치 여부를 판단하여 하나의 음성구간을 결정하고, 결정되는 음성구간과 이 음성구간에 관련된 정보를 상기 검색 결과 히스토리 데이터베이스에 저장함과 동시에 상기 파악부로 출력하는 검증기를 포함한다.

[12] 또한, 상기 동적 정합 처리기는 상기 분할된 인덱싱 정보의 음소열과 상기 발음열의 일치 여부는 동적 정합 알고리즘을 통해 판단되며, 상기 동적 정합 알고리즘을 통한 판단시 그 정합도가 임계값 이상일 경우 일치하는 것으로 판단하는 것을 특징으로 한다.

[13] 또한, 상기 검증기는 후보가 된 음성 구간에 대해 검색어의 음소열에 대해 트라이폰(tri-phone) 모델 기준으로 음소 모델의 상태 정보로 음성구간을 음소열에 따른 프레임 단위로 할당을 하고, 트라이폰 모델에 대한 관측 확률값과 모노폰(mono-phone) 모델에 대한 관측 확률값의 비율에 대한 누적값을 구한 후 정규화하여 신뢰도 값을 산출한 후, 정규화된 신뢰도 값에 따라 상기 후보가 된 음성 구간에 대해 최종적으로 검색된 결과물로 출력할 것인가를 판단하는데 활용된다는 것을 특징으로 한다.

[14] 또한, 상기 검증기는 상기 하나의 음성구간이 결정되면, 상기 하나의 음성구간을 기준으로 전후 일정 시간 범위 내에서 상기 검색 데이터베이스에서 상기 검색어에 대응하여 추출되는 문맥 키워드의 발음열에 일치하는 음소열이 있는지를 추가 검색하는 것을 특징으로 한다.

[15] 또한, 상기 검색 결과 히스토리 데이터베이스에 저장되는 상기 음성구간에 관련된 정보는 상기 하나의 음소열이 포함된 파일명, 음성 데이터에서의 시작 위치 및 끝 위치, 상기 검색어, 검색된 구간에 대한 정규화된 신뢰도 값, 일치하는 문맥 키워드, 발성화자 성별인 것을 특징으로 한다.

[16] 또한, 상기 파악부는, 검색어와 이에 대응되는 문맥 키워드별로 주제어가 설정되어 있는 주제어 데이터베이스; 상기 검색부에서 출력되는 검색 결과 정보 중에서 검색어와 문맥 키워드를 추출하여 상기 주제어 데이터베이스를 통해 대응되는 주제어를 검색하는 주제어 파악부; 및 상기 주제어 파악부로부터 검색 결과 정보와 주제어를 전달받아서 사용자에게 표시하는 출력부를 포함한다.

[17] 또한, 상기 문맥 키워드는 동일한 의미를 나타내는 복수의 단어가 카테고리 방식으로 설정되어 있는 것을 특징으로 한다.

[18] 또한, 상기 검색부는 상기 검색어에 대해 동일한 의미를 가지는 별칭에 대해서는 발음열을 생성하여 동시에 검색을 수행하는 것을 특징으로 한다.

- [19] 또한, 상기 음소 인식부는 음소 단위로 비터비(Viterbi) 알고리즘 및 토큰 패싱(Token passing) 알고리즘을 수행하여 대응되는 음소열을 생성하는 것을 특징으로 한다.
- [20] 또한, 상기 음소 격자 형성부는 해당 음소열에 대해 음소의 시작점과 끝점, 그리고 지속시간 등의 정보를 함께 형성하는 것을 특징으로 한다.
- [21] 또한, 상기 분할된 인덱싱 정보는 프레임 개수, 음소 개수, 특징 벡터, 음소의 상태별 관측 확률값, 타임 스탬프(time stamp), 음소열, 각 음소의 지속 시간을 포함한다.
- [22] 본 발명의 다른 특징에 따른 인덱싱 방법은,
- [23] 발화 내용 파악 시스템이 음성 데이터에 대해 음소 기준의 음성 인식을 수행하여 인덱싱 가능하도록 저장하는 인덱싱 방법으로서, 음성 데이터를 입력받는 단계; 입력되는 음성 데이터의 음소들에 대해 프레임 단위의 음소 인식에 따른 음소 격자를 형성하는 단계; 및 미리 설정된 제한 시간 단위로 형성되는 상기 음소 격자를 분할된 인덱싱 정보로써 인덱싱 데이터베이스에 저장하는 단계를 포함하며, 상기 음소 격자를 형성하는 단계가 음성 데이터에서 음성 검출에 의해 유용한 음성 구간에 대해서만 수행되도록 제어되는 것을 특징으로 한다.
- [24] 여기서, 상기 음소 격자를 형성하는 단계에서, 상기 제한 시간 단위로 음소 격자를 형성하는 경우 이전 음성 구간 중 특정 시간 또는 특정 프레임의 음성 구간을 중첩하여 다음 음성 구간에 대한 음소 인식을 수행하여 불연속에 따른 정보 손실을 방지하며 음소 격자를 형성하는 것을 특징으로 한다.
- [25] 또한, 상기 음소 격자를 형성하는 단계는, 프레임 단위의 음성 데이터로부터 추출되는 특징 벡터를 사용하여 프레임 동기에 기준한 음소 인식에 의해 생성되는 다수의 음소열 후보군에 대해 역시간적인 연산을 수행하여 최종 선택되는 하나의 음소열에 대응되는 음소 격자를 형성하는 것을 특징으로 한다.
- [26] 본 발명의 또 다른 특징에 따른 발화 내용 파악 방법은,
- [27] 발화 내용 파악 시스템이 음성 데이터의 발화 내용을 파악하는 방법으로서, 사용자로부터 검색어를 입력받는 단계; 상기 검색어에 대응되는 음소 기준의 발음열을 생성하는 단계; 상기 발음열을 사용하여 인덱싱 데이터베이스—여기서 인덱싱 데이터베이스는 상기 음성 데이터에 대해 프레임 단위로 음소 기준의 음성 인식이 수행되어 형성되는 음소 격자가 제한 시간의 복수의 프레임별로 분할된 인덱싱 정보로써 저장되어 있음—에 저장되어 있는 분할된 인덱싱 정보에 대해 일치하는 음소열을 검색하여 1차 후보의 음성구간을 선정하는 단계; 상기 1차 후보의 음성구간에 대해서 음향학적 모델을 통해 일치 여부를 판단하여 하나의 검색된 음성구간을 결정하는 단계; 상기 하나의 검색된 음성구간을 기준으로 전후 일정 시간 범위 내에서 상기 검색어에 대응되는 문맥 키워드의 발음열에 일치하는 음소열이 있는지를 추가 검색하는 단계; 및 상기 검색어와 상기 문맥 키워드를 통해 상기 음성 데이터에 대한 주제어를 파악하여

사용자에게 제공하는 단계를 포함한다.

[28] 또한, 상기 주제어를 파악하여 사용자에게 제공하는 단계에서, 상기 주제어의 파악은 검색어와 이에 대응되는 문맥 키워드별로 주제어가 설정되어 있는 주제어 데이터베이스를 통해 수행되는 것을 특징으로 한다.

[29] 또한, 상기 검색어는 복수 개이며, 복수 개의 검색어가 연산 논리자에 의해 사용되어 상기 복수 개의 검색어의 조합에 의해 검색이 수행되는 것을 특징으로 한다.

발명의 효과

[30] 본 발명에 따르면, 검색어로 입력되는 핵심어에 대해서만 음성 데이터에서 음소 기준으로 추출하면 되므로 실제 서비스에 사용되는 음성 데이터의 전체 단어를 파악할 필요가 없어 검색 속도가 빠르고 정확도가 높으며, 미등록어를 지속적으로 반영하여야 하는 일반적인 대용량 연속어 음성 인식 엔진을 활용하는 일명 받아쓰기(Dictation) 프로그램을 녹취 검색에 적용하는 것보다 실녹취 데이터에 대해 성능이 우수하며 서비스 시스템의 유지 보수 비용 측면에서도 획기적이다.

[31] 또한, 다수의 검색어의 조합에 의한 검색 결과 및 주제어 검색 결과를 통해 음성 파일의 발화 내용을 파악할 수 있으므로 음성 파일에 대한 발화 내용 파악 시간이 줄어들고, 그 결과로써 고객에 대한 응답 처리가 신속해질 수 있다.

도면의 간단한 설명

[32] 도 1은 본 발명의 실시예에 따른 핵심어 추출 기반 발화 내용 파악 시스템의 개략적인 구성을 도시한 도면이다.

[33] 도 2는 도 1에 도시된 인덱싱 DB의 일례를 도시한 도면이다.

[34] 도 3은 도 1에 도시된 인덱싱부의 구체적인 구성을 도시한 도면이다.

[35] 도 4는 도 3에 도시된 인덱싱 제어부의 구체적인 구성을 도시한 도면이다.

[36] 도 5는 도 1에 도시된 검색부의 구체적인 구성을 도시한 도면이다.

[37] 도 6은 도 5에 도시된 동적 정합 처리기에서 수행되는 동적 정합 알고리즘의 기본적인 동작을 나타내는 개략도이다.

[38] 도 7은 도 5에 도시된 동적 정합 처리기에서 수행되는 동적 정합 알고리즘의 일례를 도시한 도면이다.

[39] 도 8은 도 5에 도시된 동적 정합 처리기가 인덱싱 데이터(B)에 대해 검색어(C)를 이동시키면서 검색(Shifting & Searching)을 수행하는 동적 정합 처리를 도시한 도면이다.

[40] 도 9는 도 1에 도시된 파악부의 구체적인 구성을 도시한 도면이다.

[41] 도 10은 도 9에 도시된 주제어 DB의 일례를 도시한 도면이다.

[42] 도 11은 도 9에 도시된 출력부가 사용자에게 표시하는 화면의 예를 도시한 도면이다.

[43] 도 12는 본 발명의 실시예에 따른 음소 기준의 인덱싱 방법의 흐름도이다.

- [44] 도 13은 도 12에 도시된 음소 격자 형성 단계의 구체적인 흐름도이다.
- [45] 도 14는 본 발명의 실시예에 따른 핵심어 추출 기반 발화 내용 파악 방법의 흐름도이다.

발명의 실시를 위한 최선의 형태

- [46] 아래에서는 첨부한 도면을 참고로 하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 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.
- [47] 명세서 전체에서, 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. 또한, 명세서에 기재된 "...부", "...기", "모듈" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.
- [48] 도 1은 본 발명의 실시예에 따른 핵심어 추출 기반 발화 내용 파악 시스템의 개략적인 구성을 도시한 도면이다.
- [49] 도 1에 도시된 바와 같이, 본 발명의 실시예에 따른 핵심어 추출 기반 발화 내용 파악 시스템(10)은 인덱싱부(100), 인덱싱 데이터베이스(DB)(200), 사용자 인터페이스(300), 검색부(400) 및 파악부(500)를 포함한다.
- [50] 인덱싱부(100)는 녹취 서버(도시되지 않음)에 저장되어 있는 음성 파일(11)의 음성 데이터를 입력받아서 음소 기준의 음성 인식을 수행하여 음소 격자를 형성한다. 이러한 음소 기준의 음성 인식은 음성 구간 시간 제한 단위로 수행되어 대응되는 음소 격자를 형성하며, 음성 구간 시간 제한 단위 내의 음성 데이터에 대해서는 또한 프레임 단위로 음소 인식이 수행되어 대응되는 음소 격자가 형성된다. 인덱싱부(100)에 의해 음성 구간 시간 제한 단위로 생성되는 음소 격자는 분할된 인덱싱 정보로써 인덱싱 DB(200)에 저장된다.
- [51] 인덱싱 DB(200)는 인덱싱부(100)에 의해 생성되는 인덱싱 정보를 저장한다. 인덱싱 정보는 음성 구간 시간 제한 단위로 생성되는 음소 격자를 포함하는 분할된 인덱싱 정보이다. 이러한 인덱싱 정보에는 프레임 개수, 음소 개수, 특징 벡터, 음소의 상태별 관측 확률값, 타임 스탬프(time stamp), 음소열, 각 음소의 지속 시간 등이 포함된다. 여기서, 각 음소의 지속 시간은 음성 파일(11)에서의 지속 시간으로, 예를 들어 음성 파일에서의 시작 위치와 끝 위치 및 시작 위치와 끝 위치를 통해 산출되는 지속 시간일 수 있다.
- [52] 도 2는 도 1에 도시된 인덱싱 DB(200)의 일례를 도시한 도면이다.
- [53] 도 2에 도시된 바와 같이, 인덱싱 DB(200)는 N(N은 1 이상의 자연수임)개의

분할된 인덱싱 정보(200-1, 200-2, ..., 200-N)가 하나의 파일 단위로 저장된다. 파일 단위로 헤더에는 분할된 인덱싱 부분의 개수, 인덱싱부의 버전 번호, 원 음성 파일의 크기, 특징 벡터의 차수 등 정보가 기록된다. 또한, 분할된 인덱싱 정보는 프레임의 개수 및 음소열의 개수를 포함하는 헤더정보와 실질 인덱싱된 정보를 포함한다. 실질적인 인덱싱 정보는 특징 벡터, 타임스탬프(Time Stamp), 음소열, 음소열 지속정보, 음소열에 대한 관측 확률값 등이 기록된다.

- [54] 계속해서, 사용자 인터페이스(300)는 사용자와의 상호 작용을 위한 장치로서, 일반적인 사용자 인터페이스 구성을 사용하여 구현 가능하다. 예를 들어, 사용자 인터페이스(300)는 사용자로부터의 입력을 받기 위한 키보드 및 마우스 장치, 사용자에게 화면을 통한 표시를 제공하기 위한 디스플레이 장치, 소리 입출력을 위한 마이크 및 스피커 등을 포함할 수 있다.
- [55] 검색부(400)는 사용자 인터페이스(300)를 통해 사용자로부터 입력되는 핵심어인 검색어를 음소로 변환하여 인덱싱 DB(200)에 저장되어 있는 분할된 인덱싱 정보(200-1, 200-2, ..., 200-N)에 대해 음소 기준의 비교를 통해 검색어와 일치하는 음소열을 검색한 후 검색 결과를 생성한다. 여기서, 검색부(400)는 다양한 방식을 통해 검색을 수행할 수 있으며, 본 발명의 실시예에서는 인식 알고리즘의 대표적으로 알려져 있는 동적 정합 알고리즘(Dynamic Time Warping:DTW)을 사용하여 검색을 수행한다. 이러한 동적 정합 알고리즘은 대표 패턴과 주어진 입력 패턴을 비교하여 유사성을 판별하는 방법으로써 이미 잘 알려져 있으므로 여기에서는 구체적인 설명을 생략한다.
- [56] 또한, 검색부(400)는 둘 이상의 핵심어를 사용하여 검색을 수행할 수 있다. 이 경우, 검색부(400)는 사용되는 복수 개의 핵심어와 인덱싱 DB(200)에 저장되어 있는 분할된 인덱싱 정보(200-1, 200-2, ..., 200-N)를 동시에 비교하여 검색을 수행할 수 있다.
- [57] 파악부(500)는 검색부(400)에 의해 생성되는 검색 결과를 사용하여 음성 파일(11)에 녹취된 음성 데이터에 대응되는 사용자의 발화 내용을 파악하여 검색 결과와 함께 사용자 인터페이스(300)를 통해 출력한다. 이러한 발화 내용의 예로, 홈쇼핑 녹취의 경우에는 고객의 주문 의사 여부를 판단하는 내용, 상담원의 친절도나 특정 상품에 대한 적절한 홍보 여부를 판단하는 내용, 특정 단어의 누설을 판단하는 내용 또는 군사 보안 내용 등이 있다.
- [58] 도 3은 도 1에 도시된 인덱싱부(100)의 구체적인 구성을 도시한 도면이다.
- [59] 도 3에 도시된 바와 같이, 인덱싱부(100)는 버퍼부(110), 특징 벡터 추출부(120), 음소 인식부(130), 후보군 형성부(140), 음소 격자 형성부(150) 및 인덱싱 제어부(160)를 포함한다.
- [60] 버퍼부(110)는 사용자 인터페이스(300)를 통해 사용자에 의해 지정된 음성 파일(11)의 음성 데이터를 입력받아서 임시로 저장하는 다수의 버퍼(도시되지 않음)를 포함한다.
- [61] 특징 벡터 추출부(120)는 버퍼부(110)로부터 저장되어 출력되는 프레임 단위의

음성 데이터로부터 특징 벡터를 추출한다. 이러한 특징 벡터 추출은 음성 인식 분야에서 잘 알려져 있으므로 여기에서는 구체적인 설명을 생략한다.

- [62] 음소 인식부(130)는 특징 벡터 추출부(120)에 의해 추출되는 특징 벡터를 사용하여 프레임 동기에 기준한 음소 인식을 수행한다. 이러한 음소 인식부(130)는 음소 단위로 비터비(Viterbi) 알고리즘 및 토큰 패싱(Token passing) 알고리즘을 수행하여 대응되는 음소열을 생성한다. 상기한 비터비 알고리즘이나 토큰 패싱 알고리즘에 대해서도 잘 알려져 있으므로 여기에서는 구체적인 설명을 생략한다.
- [63] 후보군 형성부(140)는 음소 인식부(130)에 의해 생성되는 음소열을 입력받아서 프레임 단위의 시간적 흐름에 따라 음소 인식의 후보군들을 생성하며, 이러한 후보군들은 최소 N개의 음소열 후보군이 유지되도록 제어된다.
- [64] 음소 격자 형성부(150)는 후보군 형성부(140)에 의해 형성되는 N개의 음소열 후보군으로부터 역시 간적인 연산을 수행하여 1개의 최적의 음소열 후보군을 선택하여 대응되는 음소 격자를 형성한다. 이 때 음소 격자 형성부(150)는 해당 음소열에 대해 음소의 시작점과 끝점, 그리고 지속시간 등의 정보를 함께 형성한다. 여기서, 음소의 시작점 및 끝점은 음소가 시작되는 지점의 음성 데이터에서의 시간 정보와 음소가 끝나는 지점의 음성 데이터에서의 시간 정보이다. 음소 격자 형성부(150)에 의해 형성되는 음소 격자는 분할된 인덱싱 정보에 포함되어 인덱싱 DB(200)로 출력된다.
- [65] 인덱싱 제어부(160)는 버퍼부(110), 특징 벡터 추출부(120), 음소 인식부(130), 후보군 형성부(140) 및 음소 격자 형성부(150)를 제어하여, 음성 데이터 전체에 대해 제한된 시간별로 또한 제한된 시간 내에서 프레임 단위별로 음소 단위의 격자가 형성되도록 제어를 수행하고, 이렇게 형성되는 음소 단위의 격자, 즉 음소 격자를 각 제한된 시간별로 인덱싱 될 수 있도록 제한된 시간별로 분할된 인덱싱 정보로써 인덱싱 DB(200)에 저장될 수 있도록 제어를 수행한다. 예를 들어, 인덱싱 제어부(160)는 10초라는 제한 시간동안 음성 구간이 지속되면 이 제한 시간에 대응되는 프레임 단위의 음소 격자를 형성하여 인덱싱 DB(200)에 저장한다. 만약 음성 구간이 제한 시간인 10초 후에도 지속되는 경우에는 제한 시간인 10초에서 강제 음성 끝점을 검출한 후 특정 시간, 예를 들어 3초 동안 이전 시간의 음성 구간을 중첩하여 다음의 제한 시간의 분할된 인덱싱 정보에 해당하는 음소 격자 형성 동작을 수행한다.
- [66] 이하, 인덱싱 제어부(160)에 대해 구체적으로 설명한다.
- [67] 도 4는 도 3에 도시된 인덱싱 제어부(160)의 구체적인 구성을 도시한 도면이다.
- [68] 도 4에 도시된 바와 같이, 인덱싱 제어부(160)는 음성 검출기(161), 시간 카운터(162), 시간 제한기(163) 및 동작 제어기(164)를 포함한다.
- [69] 음성 검출기(161)는 버퍼부(110)에 저장된 음성 데이터로부터 유효한 음성 구간을 검출한다. 즉, 음성 검출기(161)는 검출되는 음성 데이터가 현재 음성 구간인지 아닌지를 판단할 수 있도록 한다. 이것은, 녹취 파일에 항상 음성만

존재하는 것이 아닐 수 있기 때문에, 인덱싱 데이터의 크기를 줄이기 위해 유효한 음성 구간에 대해서만 인덱싱을 수행하여 그 결과 정보를 저장하기 위해서이다.

- [70] 시간 카운터(162)는 버퍼부(110)에 저장된 음성 데이터에 대해 음소 인식이 수행되는 음소의 시간적 위치를 카운트한다. 이 시간 카운터(162)는 동작 제어기(164)에 의해 음소 인식이 시작되는 때에 카운트를 시작하고, 현재 인식되는 음소의 시간을 출력할 수 있다.
- [71] 시간 제한기(163)는 음성 검출기(161)에 의해 음성 구간이 검출되는 때에 시간 카운터(162)로부터 해당 음성 구간이 검출되는 때부터 시간을 카운트하여 미리 설정된 제한 시간, 상기 예를 참조하면 10초의 시간을 카운트한다. 만약 시간 제한기(163)가 10초의 시간을 모두 카운트하여 도과되는 때에 제한 시간을 도과함을 동작 제어기(164)로 알리게 된다. 또는, 동작 제어기(164)의 요구에 따라 현재 카운트되는 시간이 제한 시간이 도과하였는지 아닌지의 여부를 알려준다.
- [72] 동작 제어기(164)는 시간 제한기(163)에 의해 카운트되는 제한 시간 내에 음성 검출기(161)에 의해 검출되는 유효한 음성 구간에 대해 프레임 단위로 음소 인식을 수행하여 음소 격자를 형성하여 분할된 인덱싱 정보로써 인덱싱 DB(200)에 저장한다. 동작 제어기(164)는 음성 검출기(161)에 의해 검출되는 유효한 음성 구간에 대해 시간 제한기(163)에 의해 제한 시간이 도과하는 경우 이전 음성 구간 중 특정 시간 또는 특정 프레임의 음성 구간을 중첩하여 해당 프레임부터 새로운 제한 시간에 대응되는 음성 구간으로써 음소 인식을 수행하도록 제어를 수행한다.
- [73] 다음, 상기한 검색부(400)에 대해 구체적으로 설명한다.
- [74] 도 5는 도 1에 도시된 검색부(400)의 구체적인 구성을 도시한 도면이다.
- [75] 도 5에 도시된 바와 같이, 검색부(400)는 발음열 생성기(420), 검색 DB(410), 동적 정합 처리기(430), 검증기(440) 및 검색 결과 히스토리 DB(450)를 포함한다.
- [76] 발음열 생성기(420)는 사용자 인터페이스(300)를 통해 사용자로부터 입력되는 핵심어(keyword), 즉 검색어를 입력받아서, 음소 단위의 발음열을 생성한다. 이 때, 검색어는 일반적인 철자 기준의 단어이지만, 반드시 철자일 필요는 없으며 발음열 생성시 발음 현상을 제대로 반영하지 못할 수도 있으므로 소리나는 대로 기입되어도 좋다. 발음열 생성의 예로써, 검색어가 '송파구'인 경우 발음열 생성기(420)는 이에 대해 's o ng p a g u" 또는 "ㅅ ㅗ ㅇ ㅍㅏ ㄱ ㅏ" 등으로 발음열을 생성할 수 있다.
- [77] 검색어 DB(410)는 검색어와 검색어에 대응되는 복수의 문맥 키워드를 저장한다. 여기서, 문맥 키워드는 대응되는 검색어에 보조되는 단어로써, 검색어와 함께 사용되어 음성 데이터의 발화 내용을 파악할 수 있다. 이러한 문맥 키워드는 사용자에 또는 본 발명의 실시예에 따른 핵심어 추출 기반 발화 내용 파악 시스템(10)의 관리자에 의해 미리 설정되는 것으로 가정한다. 예를

들면, 가전 제품 콜센터 상담 녹취록일 경우, 검색어가 '냉장고'인 경우 이러한 '냉장고'에 대해 문맥 키워드로 '김치', '와인', '고장', '수리' 등이 설정될 수 있다. 이 때, 검색어 DB(410)에 저장되어 있는 문맥 키워드는 일반 철자 기준의 단어와 함께 대응되는 음소 단위의 발음열이 미리 생성되어 함께 저장되어 있는 것으로 가정한다.

- [78] 동적 정합 처리기(430)는 발음열 생성기(420)로부터 검색어와 검색어에 대응되어 생성되는 발음열을 전달받고, 해당 발음열을 사용하여 인덱싱 DB(200)에 저장되어 있는 분할된 인덱싱 정보에 대해 일치하는 음소열을 검색하여 1차 후보의 음소구간을 출력한다. 여기서, 분할된 인덱싱 정보에 대해 발음열의 일치 여부는 동적 정합 알고리즘을 통해 판단하여 그 정합도가 임계값 이상일 경우를 의미한다.
- [79] 이하, 동적 정합 처리기(430)에 의해 수행되는 동적 정합 알고리즘에 대해 구체적으로 설명한다.
- [80] 도 6은 도 5에 도시된 동적 정합 처리기(430)에서 수행되는 동적 정합 알고리즘의 기본적인 동작을 나타내는 개략도이다.
- [81] 도 6을 참조하면, 가로축은 인덱싱 DB(200)에 저장되어 있는 분할된 인덱싱 정보의 음소열 중 일부를 나타내고, 세로축은 검색어를 음소열로 변환한 발음열을 나타낸다. 여기서, 검색어는 영어로 'SPEECH'인 경우를 나타낸다.
- [82] 도 6에서는 복수의 음소열 중에서 검색어인 'SPEECH'에 매칭되는 음소열을 나타낸 것이다.
- [83] 기본적으로, 음소 격자에 의해 결정되는 음소열은 그 정확도가 60%정도이다. 따라서, 음소 단위로 틀린 음소가 삽입되기도 하고, 나타나지 않을 수도 있고, 또한 다른 음소로 인식될 수도 있다. 이에, 시간적으로는 맞는 부분이 일찍 나타날 수도 있고 늦게 나타날 수도 있다. 따라서, 본 발명의 실시예에서는 도 6에 도시된 바와 같이 최적으로 매칭될 수 있는 경로 방식을 동적 정합 알고리즘을 통해 산출한다.
- [84] 도 7은 도 5에 도시된 동적 정합 처리기(430)에서 수행되는 동적 정합 알고리즘의 일 예를 도시한 도면이다.
- [85] 도 7을 참조하면, 실제 발성되어 저장된 음성 데이터 중에서 도 7에서 사용되는 일부 실제 발성 내용(A)은 "서울 특별시 송파구 송파동"이다. 이러한 음성 데이터에 대해 인덱싱부(100)가 음소 기준의 인덱싱을 수행하여 인덱싱 DB(200)에 저장한 음소열로써 도 7에서 사용되는 인덱싱 데이터(B)는 "ㅅ ㅗ ㄹ ㅍ ㅡ ㅌ ㅋ ㅎ ㅠ ㅅ ㅗ ㅂ ㅏ ㅅ ㅗ ㅂ ㅏ ㄷ ㅜ ㅇ"이다. 그리고, 검색어가 "송파구"이므로 발음열 생성기(420)에 의해 생성되는 발음열(C)은 상기한 바와 같이 "ㅅ ㅗ ㅇ ㅍ ㅏ ㄱ ㅜ"이다.
- [86] 따라서, 동적 정합 처리기(430)는 검색어의 발음열(C)인 "ㅅ ㅗ ㅇ ㅍ ㅏ ㄱ ㅜ"를 기준으로 인덱싱 데이터(B)인 "ㅅ ㅗ ㄹ ㅍ ㅡ ㅌ ㅋ ㅎ ㅠ ㅅ ㅗ ㅂ ㅏ ㅅ ㅗ ㅂ ㅏ ㄷ ㅜ ㅇ"에 대해 도 6에서와 같은 동적 정합 알고리즘을 음소 단위로

수행한다. 여기서, 동적 정합 처리기(430)는 도 8에 도시된 바와 같이, 인덱싱 데이터(B)에 대해 검색어(C)를 이동시키면서 검색(Shifting & Searching)을 수행함으로써 동적 정합 알고리즘을 수행한다.

[87] 도 7을 다시 참조하면, 예를 들어, 인덱싱 데이터(B) 중에서 P1의 위치에서 도 6에 도시된 바와 같이, 검색어(A)와 인덱싱 데이터(B)의 동적 정합 알고리즘 수행에 의해 정합도가 산출되며, 예를 들어 P1의 위치에서의 정합도는 35점이 된다.

[88] 한편, P2의 위치의 인덱싱 데이터(B)에 대한 검색어(C)의 동적 정합 알고리즘 수행에 따른 정합도는 75점이 산출된다. 여기서, P2의 위치에서 인덱싱 데이터(B)의 'ㅅ'과 '그'는 검색어(C)의 'ㅅ'과 '그'와 매칭되어 높은 점수로 반영되지만, 검색어(C)의 'ㅇ'은 누락되어 패널티가 적용되고, 검색어(C)의 'ㅍ'은 인덱싱 데이터(B)의 'ㅂ'으로 오인식되었지만 음가상 비슷하여 어느 정도 높은 점수로 반영되며, 검색어(C)의 'ㄱ'은 누락되어 패널티가 적용되고, 검색어(C)의 'ㅌ'은 인덱싱 데이터(B)의 'ㅌ'과 비슷한 음ガ를 가지므로 어느 정도 높은 점수를 가져서, 전체적으로 음소 글자 매칭을 통해 75점이라는 높은 점수, 실제로는 임계값으로 설정된 60점 이상이어서 1차 후보의 음소구간으로 선정된다.

[89] 마찬가지로, 상기한 바와 같은 동작이 P3 위치의 인덱싱 데이터(B)에 대해서도 수행되어 정합도가 80점으로 산출되어 역시 1차 후보로 선정된다. 여기서, 실질적으로 발성은 /송파동/이지만, 음소인식의 부정확성 때문에 'ㅅ 그 ㅂ ㄷ ㅌ'가 되어, 'ㅅ 그 ㅇ ㅍ ㅏ ㄱ ㅌ'과 'ㅅ 그 ㅂ ㄴ ㄷ ㅌ'는 문자열 비교에서 높은 점수를 가질 수 있다. 이와 같이 검색어(C)와 다소 상이한 인덱싱 데이터(B)도 1차 후보가 되지만, 추후 설명되는 검증기(440)에 의해 1차 후보 중에서 보다 정확한 매칭이 이루어짐으로 인해 P3 위치의 인덱싱 데이터(B)는 제거될 것이다.

[90] 다음, 검증기(440)는 동적 정합 처리기(430)에서 선정되는 1차 후보의 음소구간에 대해서 음향학적 모델을 통해 상세하게 일치 여부를 판단하여 최적의 하나의 음성구간을 결정하고, 결정되는 음성구간과 이 음성구간에 관련된 각종의 정보를 검색 결과 히스토리 DB(450)에 저장함과 동시에 파악부(500)로 출력한다. 여기서, 검증기(440)는 후보가 된 음성 구간에 대해 검색어의 음소열에 대해 트라이폰(tri-phone) 모델 기준으로 음소 모델의 상태 정보로 음성구간을 음소열에 따른 프레임 단위로 할당을 하고, 트라이폰 모델에 대한 판측 확률값과 모노폰(mono-phone) 모델에 대한 판측 확률값의 비율에 대한 누적값을 구한 후 0~100점 사이로 정규화하여 정규화된 신뢰도 값을 산출한 후, 정규화된 신뢰도 값이 가장 높은 후보를 최적의 후보로 결정하게 된다. 상기한 트라이폰 모델이나 모노폰 모델과 같은 음향학적 모델을 사용하는 내용에 대해서는 이미 잘 알려져 있으므로 여기에서는 구체적인 설명을 생략한다.

[91] 또한, 검증기(440)는 최적의 후보가 결정되면, 최적의 후보에 해당하는 음성구간을 기준으로 전후 일정 시간 범위 내에서 검색 DB(410)에서 검색어에

대응하여 추출되는 문맥 키워드의 발음열에 일치하는 음소열이 있는지를 추가 검색한다. 여기서, 일정 시간 범위는 전체적으로 제한 시간의 범위, 상기 예에서 10초의 이내로 설정되는 것이 바람직하다.

- [92] 한편, 검증기(440)가 검색 결과로써 출력하는 정보는 검색된 음소열이 포함된 파일명, 음성 데이터에서의 시작 위치 및 끝 위치, 검색어, 검색된 구간에 대한 정규화된 신뢰도 값, 일치하는 문맥 키워드, 발성화자 성별(남자 또는 여자) 등을 포함한다. 여기서, 발성화자 성별 정보는 음향학적으로 남성 모델과 여성 모델 각각에 대해 음성 신호의 유사도가 큰 쪽으로 판별하여 추출할 수 있다.
- [93] 검색 결과 히스토리 DB(450)는 검증기(440)를 통해 출력되는 검색 결과를 검색어와 함께 저장하며, 사용자 인터페이스를 통해 검색어가 입력되면 해당 검색어에 대해 검색 수행 이력이 존재하면 중복 검색 수행을 방지하고 해당 검색어에 대해 저장된 검색 결과를 파악부(500)로 전달한다.
- [94] 다음, 파악부(500)에 대해 구체적으로 설명한다.
- [95] 도 9는 도 1에 도시된 파악부(500)의 구체적인 구성을 도시한 도면이다.
- [96] 도 9에 도시된 바와 같이, 파악부(500)는 주제어 DB(510), 주제어 파악부(520) 및 출력부(530)를 포함한다.
- [97] 주제어 DB(510)는 예를 들어, 도 10에 도시된 바와 같이, 검색어와 문맥 키워드별로 대응되는 주제어가 설정되어 있다. 즉, 하나의 검색어에 복수의 문맥 키워드가 설정되어 있고, 이러한 복수의 문맥 키워드별로 각각의 주제어가 미리 설정되어 있다.
- [98] 주제어 파악부(520)는 검색부(400)에서 출력되는 검색 결과 정보 중에서 검색어와 문맥 키워드를 추출하여 주제어 DB(510)를 통해 대응되는 주제어를 검색하여 검색되는 주제어를 출력한다. 예를 들어, 검색어가 '냉장고'이고, 일치되는 문맥 키워드가 '김치'와 '안되요'인 경우 주제어 파악부(520)는 '김치 냉장고의 고장'으로 주제어를 파악하게 된다. 따라서, 여기에서의 주제어가 음성 데이터의 발화 내용에 해당한다고 볼 수 있다.
- [99] 출력부(530)는 주제어 파악부(520)로부터 검색 결과 정보와 주제어 정보를 전달받아서 사용자에게 제공하는 검색 화면을 사용자 인터페이스(300)를 통해 출력한다. 출력부(530)가 사용자 인터페이스(300)를 통해 사용자에게 표시하는 화면의 예가 도 11에 도시되어 있다. 도 11에서는 고객에 의해 실제 발성된 내용은 "아 저 뭐더라 ... 김치 냉장고가 ... 안되셔요"이고, 사용자가 제공하는 검색어는 '냉장고'이며, 검색어에 대응되어 일치되는 문맥 키워드는 '김치'와 '안되요'임을 알 수 있다. 여기서, 고객은 '안되셔요'를 발생하였으나, 일치되는 문맥 키워드가 '안되요'로 나타나는 것은 검색어 DB(410)에 유사 키워드를 설정할 때 카테고리 방식으로 유사 단어를 설정함으로써 이루어질 수 있다. 예를 들어, 문맥 키워드가 '안되요'인 경우 이에 대해 '안되요'는 물론, '안되셔요', '안되네요', '안됩니다' 등의 유사한 단어들을 카테고리로 묶어서 검색될 수 있도록 한다. 이 뿐만 아니라, 검색어의 경우에도 별칭 검색을 지원하여,

사용자가 '휴대폰'을 검색어로 입력하는 경우 검색부(400)는 '핸드폰', '스마트폰', '갤럭시폰', 'G2폰' 등의 별칭에 대해서도 동시에 검색을 수행하도록 구성된다. 나아가, 명사어뿐만 아니라 동사의 경우에도 별칭 검색이 가능하다. 예를 들어, 사용자가 '동의하십니까'를 검색어로 입력하더라도 검색부(400)는 '동의하십니까'는 물론 '동의하세요' 등의 별칭에 대해서도 동시에 검색을 수행하도록 구성된다.

- [100] 또한, 도 11에서는 검색 결과로써 표준화된 신뢰도값, 전체 원 음원 음성파일 상에서의 시작점과 끝점의 시간이 표시되고, 검색어 '냉장고'와 문맥 키워드 '김치'와 '안되요'에 의해 파악되는 주제어가 '가전고장'인 것으로 파악된 결과가 함께 표시됨을 알 수 있다.
- [101] 이하, 도면을 참조하여 본 발명의 실시예에 따른 음소 기준의 인덱싱 방법에 대해 설명한다.
- [102] 도 12는 본 발명의 실시예에 따른 음소 기준의 인덱싱 방법의 흐름도이다.
- [103] 본 발명의 실시예에 따른 인덱싱 방법에서는 고객 등의 음성 녹취가 수행되어 녹취된 음성 파일(11)이 생성되면, 바로 또는 추후 사용자의 지시에 따라 음성 파일(11)에 대해 인덱싱 방법이 수행될 수 있다.
- [104] 도 12를 참조하면, 인덱싱될 음성 파일(11)이 지정되면 지정된 음성 파일(11)에 저장된 음성 데이터가 인덱싱부(100)로 입력되어 버퍼부(110)에 저장된다(S100).
- [105] 그러면, 인덱싱 제어부(160)는 버퍼부(110)에 저장된 음성 데이터로부터 음성의 시작점이 검출되는지를 판단한다(S110).
- [106] 만약 음성의 시작점이 검출되지 않으면 음성의 시작점이 검출될 때까지 음성 데이터상에서 프레임 이동(S120)을 통해 지속적으로 프레임 단위로 음성 데이터를 판독하게 된다.
- [107] 상기 단계(S110)에서 음성의 시작점이 검출되면, 음성 데이터에서 유효한 음성 구간이 처음 시작되는 것이므로, 인덱싱 제어부(160)는 시간 카운팅을 시작한 후(S130), 유효한 음성 구간의 음소들에 대해 프레임 단위의 음소 인식에 따른 음소 격자를 형성한다(S140). 이와 같이 프레임 단위의 음소 인식에 따라 음소 격자를 형성하는 단계에 대해서는 뒤에서 구체적으로 설명한다.
- [108] 다음, 인덱싱 제어부(160)는 미리 설정된 제한 시간, 예를 들면 10초가 도과되었는지를 판단하고(S150), 만약 미리 설정된 제한 시간이 도과되었으면 강제로 음성의 끝점 검출이 완료된 것과 동일하게 처리하여 분할된 인덱싱 정보를 인덱싱 DB(200)에 저장한다(S160). 여기서, 분할된 인덱싱 정보는 상기 단계(S140)에서 형성되는 음소 격자와 다수의 정보를 포함한다. 이러한 다수의 정보에 대해서는 상기에서 설명하였으므로, 여기에서는 구체적인 설명을 생략한다. 그 이후, 강제 음성 끝점 처리에 따른 음성 정보의 손실을 보전하기 위해 프레임 중첩을 수행한다(S170). 예를 들면 3초 정도의 음성구간을 되돌아가 인덱싱 과정을 수행한다.
- [109] 한편, 상기 단계(S150)에서 제한 시간이 도과되지 않은 것으로 판단되면,

인덱싱 제어부(160)는 분할된 인덱싱 정보 저장을 위해 음성 데이터에 대해 음성의 끝점이 검출되는지를 판단하고(S180), 만약 음성의 끝점이 검출되면 분할된 인덱싱 정보를 저장하고(S190), 다음 음성 데이터에 대해서 프레임 이동하여(S120) 새로운 시작점을 찾는 과정(S110)이 반복 수행된다. 한편, 상기 단계(S180)에서 음성의 끝점이 검출되지 않는 경우, 음성 파일(11)의 끝이어서 인덱싱을 수행할 음성 데이터가 더 이상 없는지를 판단하고(S200), 만약 음성 파일(11)의 끝이 아니어서 인덱싱을 수행할 음성 데이터가 더 남아 있는 것으로 판단되면 다음 프레임에 대한 처리를 위해 프레임 이동을 수행한 후(S210) 이동된 프레임의 음소들에 대해 음소 격자를 형성하는 상기 단계(S130)가 반복 수행된다.

- [110] 이와 같이, 계속되는 음성 데이터에 대해 인덱싱을 수행하다가 상기 단계(S200)에서 음성 파일(11)의 끝이어서 인덱싱을 수행할 음성 데이터가 없는 것으로 판단되면 인덱싱 제어부(160)는 음성이 시작되었으나 제한 시간이 도과한 것도 아니고 또한 음성 끝점이 검출된 것이 아니면서 음성 파일(11)의 끝이 되었으므로 정보 손실을 막기 위해 현재까지 인덱싱 처리된 정보를 분할된 인덱싱 정보로써 저장한 후(S220) 음성 파일(11)에 대한 인덱싱 수행을 종료한다.
- [111] 도 13은 도 12에 도시된 음소 격자 형성 단계(S140)의 구체적인 흐름도이다.
- [112] 도 13을 참조하면, 인덱싱 제어부(160)는 버퍼부(110)에 저장되어 출력되는 프레임 단위의 음성 데이터로부터 특징 벡터를 추출한 후(S131), 추출되는 특징 벡터를 사용하여 프레임 동기에 기준한 음소 인식을 수행한다(S132).
- [113] 계속해서, 인덱싱 제어부(160)는 음소 인식에 의해 생성되는 음소열을 프레임 단위의 시간적 흐름에 따라 최소 N개의 음소열 후보군을 형성하고(S133), 형성되는 N개의 음소열 후보군으로부터 역시 시간적인 연산을 수행하여 1개의 최적의 음소열 후보군을 선택하여 대응되는 음소 격자를 형성한다(S134).
- [114] 다음, 도면을 참조하여 본 발명의 실시예에 따라 음소 기준으로 인덱싱된 정보에 대해 핵심어 추출 기반에 의거하여 발화 내용을 파악하는 방법에 대해 설명한다.
- [115] 도 14는 본 발명의 실시예에 따른 핵심어 추출 기반 발화 내용 파악 방법의 흐름도이다.
- [116] 본 발명의 실시예에 따른 핵심어 추출 기반 발화 내용 파악 방법에서는 사용자가 발화 내용을 파악하기 위한 음성 파일(11)이 선택되면 선택된 음성 파일(11)에 대해 인덱싱이 수행된 후의 인덱싱 정보가 저장된 인덱싱 DB(200) 파일에 대해 발화 내용 파악이 수행되는 것으로 가정하여 설명한다.
- [117] 도 14를 참조하면, 사용자 인터페이스(300)를 통해 사용자로부터 검색을 위한 핵심어가 검색어로써 입력되면(S300), 검색부(400)는 입력되는 검색어에 대해 음소 단위의 발음열을 생성한다(S310).
- [118] 그 후, 검색부(400)는 음소 단위의 발음열을 사용하여 인덱싱 DB(200)에 저장되어 있는 분할된 인덱싱 정보에 대해 동적 정합 알고리즘을 사용하여

- 일치하는 음소열을 검색하여 1차 후보의 음소열로써 출력한다(S320).
- [119] 상기한 단계(S320)의 동적 정합 처리는 인덱싱 DB(200)에 저장되어 있는 분할된 인덱싱 정보의 끝까지 반복 수행된다(S330).
- [120] 그 후, 검색부(400)는 동적 정합 처리에 의해 선정되는 1차 후보의 음소열에 대해서 음향학적 모델을 통해 상세하게 일치 여부를 판단하여 최적의 하나의 음성구간을 결정하는 검증 처리를 수행하고, 결정되는 음성구간과 이 음성구간에 관련된 각종의 정보를 검색 DB(410)에 저장함과 동시에 파악부(500)로 출력한다(S340). 이 때, 검색부(400)는 최적의 후보가 결정되면, 최적의 후보에 해당하는 음성구간을 기준으로 전후 일정 시간 범위 내에서 검색 DB(410)에서 검색어에 대응하여 추출되는 문맥 키워드의 발음열에 일치하는 음소열이 있는지를 추가 검색을 수행한다.
- [121] 다음, 파악부(500)는 검색부(400)에 의해 검증된 결과를 통해 검색어와 문맥 키워드를 추출하여 주제어 DB(510)를 통해 대응되는 주제어를 파악한다(S350).
- [122] 그리고, 파악부(500)는 최종적으로 검색어에 의해 검색이 수행된 검색 결과와 함께 파악된 주제어를 사용자 인터페이스(300)를 통해 사용자가 음성 파일(11)의 발화 내용을 파악할 수 있도록 사용자에게 표시 출력한다(S360).
- [123] 따라서, 사용자는 자신이 지정한 음성 파일(11)에 대해 핵심어를 통한 검색을 수행한 결과를 통해 음성 파일(11)의 발화 내용을 파악할 수 있게 된다.
- [124] 한편, 상기에서는 검색부(400)가 사용자 인터페이스(300)를 통해 사용자로부터 입력되는 하나의 검색어에 기초하여 검색을 수행하는 것에 대해서만 설명하였으나, 본 발명은 이에 한정되지 않고 둘 이상의 검색어를 입력받아서 각각에 대한 검색을 수행하고, 각각의 검색 결과를 사용자에게 표시할 수 있다. 특히, 둘 이상의 검색어에 대해 연산 논리자의 사용이 가능하도록 하여 둘 이상의 검색어의 조합에 의한 주제어 파악이 보다 구체적으로 실행되도록 함으로써 발화 내용을 보다 구체적으로 파악할 수 있다.
- [125] 또한, 검색부(400)의 검색 결과가 검색 결과 히스토리 DB(450)에 검색 히스토리로 저장되어 있어 이러한 검색 히스토리를 통계 자료로써 활용함으로써 추후 광고나 홍보시 사용할 수 있을 뿐만 아니라 이미 검색된 검색어에 대해 중복 검색이 수행되지 않도록 할 수 있다.
- [126] 이와 같이, 본 발명의 실시예에서는 검색어로 입력되는 핵심어에 대해서만 음성 데이터에서 음소 기준으로 추출하면 되므로 실제 서비스에 사용되는 음성 데이터의 전체 단어를 파악할 필요가 없어 검색 속도가 빠르고 정확도가 높으며, 미등록어를 지속적으로 반영하여야만 정상적인 인식 과정을 수행할 수 있는 대용량 연속어 인식 방식보다 실제 녹취 데이터에 대해 성능이 우수하며 유지 보수 비용을 줄일 수 있다.
- [127] 또한, 다수의 검색어의 조합에 의한 검색 결과 및 주제어 검색 결과를 통해 음성 파일의 발화 내용을 파악할 수 있으므로 음성 파일에 대한 발화 내용 파악 시간이 줄어들고, 그 결과로써 고객에 대한 응답 처리가 신속해질 수 있게 된다.

[128] 이상에서 본 발명의 실시예에 대하여 상세하게 설명하였지만 본 발명의 권리범위는 이에 한정되는 것은 아니고 다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한 당업자의 여러 변형 및 개량 형태 또한 본 발명의 권리범위에 속하는 것이다.

[129]

청구범위

[청구항 1]

음성 데이터를 입력받아서 프레임 단위로 음소 기준의 음성 인식을 수행하여 음소 격자를 형성하고, 복수의 프레임으로 구성되는 제한 시간의 프레임에 대해 분할된 인덱싱 정보—여기서 분할된 인덱싱 정보는 제한 시간의 프레임별로 형성되는 음소 격자를 포함함—를 생성하는 인덱싱부;
 상기 인덱싱부에 의해 생성되는 분할된 인덱싱 정보를 분할된 인덱싱 정보별로 인덱싱 가능하도록 각각 저장하는 인덱싱 데이터베이스;
 사용자로부터 입력되는 핵심어를 검색어로 하여 인덱싱 데이터베이스에 저장된 분할된 인덱싱 정보에 대해 음소 기준의 비교를 통해 상기 검색어와 일치하는 음소열을 검색하는 검색부; 및
 상기 검색부에 의해 검색되는 검색 결과를 통해 주제어를 파악하여 상기 음성 데이터의 발화 내용을 파악할 수 있도록 사용자에게 출력하는 파악부
 를 포함하는 발화 내용 파악 시스템.

[청구항 2]

제1항에 있어서,
 상기 인덱싱부는,
 프레임 단위의 음성 데이터로부터 특징 벡터를 추출하는 특징 벡터 추출부;
 상기 특징 벡터 추출부에 의해 추출되는 특징 벡터를 사용하여 프레임 동기에 기준한 음소 인식을 수행하여 대응되는 음소열을 생성하는 음소 인식부;
 상기 음소 인식부에 의해 생성되는 음소열을 입력받아서 프레임 단위의 시간적 흐름에 따라 음소 인식의 후보군들을 생성하는 후보군 형성부;
 상기 후보군 형성부에 의해 생성되는 음소열 후보군으로부터 역시간적인 연산을 수행하여 1개의 음소열 후보군을 선택하여 대응되는 음소 격자를 형성하는 음소 격자 형성부; 및
 상기 특징 벡터 추출부, 상기 음소 인식부, 상기 후보군 형성부 및 상기 음소 격자 형성부를 제어하여, 음성 데이터 전체에 대해 상기 제한 시간별로 또한 상기 제한 시간 내에서 프레임 단위별로 음소 단위의 격자가 형성되도록 제어를 수행하고, 이렇게 형성되는 음소 격자를 각 제한 시간별로 인덱싱될 수 있도록 제한 시간별로 분할된 인덱싱 정보로써 상기 인덱싱 데이터베이스에 저장되도록 제어를 수행하는 인덱싱 제어부

를 포함하는 발화 내용 파악 시스템.

[청구항 3]

제2항에 있어서,

상기 인덱싱 제어부는,

음성 데이터에서 음성이 검출되는 지의 여부를 나타내는 음성 검출기;

음성 데이터에 대해 음소 인식이 수행되는 음소의 시간적 위치를 카운트하는 시간 카운터;

상기 음성 검출기에 의해 음성 구간이 검출되는 때에 상기 시간 카운터로부터 해당 음성 구간이 검출되는 때부터 시간을

카운트하여 상기 제한 시간을 카운트하는 시간 제한기; 및

상기 시간 제한기에 의해 카운트되는 제한 시간 내에 상기 음성 검출기에 의해 검출되는 유효한 음성 구간에 대해 프레임 단위로 음소 인식을 수행하여 음소 격자를 형성하여 분할된 인덱싱 정보로써 상기 인덱싱 데이터베이스에 저장하는 제어를 수행하는 동작 제어기

를 포함하는 발화 내용 파악 시스템.

[청구항 4]

제3항에 있어서,

상기 동작 제어기는 상기 음성 검출기에 의해 검출되는 유효한 음성 구간에 대해 상기 시간 제한기에 의해 카운트되는 상기 제한 시간이 도과하는 경우 이전 음성 구간 중 특정 시간 또는 특정 프레임의 음성 구간을 중첩하여 해당 프레임부터 새로운 제한 시간에 대응되는 음성 구간으로써 음소 인식을 수행하도록 제어를 수행하는 것을 특징으로 하는 발화 내용 파악 시스템.

[청구항 5]

제4항에 있어서,

상기 검색부는,

상기 검색부에 의해 검색된 검색 결과를 저장하고, 사용자로부터 입력되는 검색어에 대해 이미 처리된 검색 결과가 있는 경우 이를 상기 파악부로 전달하는 검색결과 히스토리 데이터베이스;

사용자로부터 입력되는 검색어에 대응되는 음소 단위의 발음열을 생성하는 발음열 생성기;

상기 검색어와 상기 검색어에 대응되는 복수의 문맥 키워드를 저장하는 검색어 데이터베이스;

상기 발음열 생성기에서 생성되는 발음열을 사용하여 상기 인덱싱 데이터베이스에 저장되어 있는 분할된 인덱싱 정보에 대해 일치하는 음소열을 검색하여 1차 후보의 음성구간을 선정하는 동적 정합 처리기; 및

상기 동적 정합 처리기에서 선정되는 1차 후보의 음성구간에 대해서 음향학적 모델을 통해 일치 여부를 판단하여 하나의

음성구간을 결정하고, 결정되는 음성구간과 이 음성구간에 관련된 정보를 상기 검색결과 히스토리 데이터베이스에 저장함과 동시에 상기 파악부로 출력하는 검증기
를 포함하는 발화 내용 파악 시스템.

[청구항 6]

제5항에 있어서,

상기 동적 정합 처리기는 상기 분할된 인덱싱 정보의 음소열과 상기 발음열의 일치 여부는 동적 정합 알고리즘을 통해 판단되며, 상기 동적 정합 알고리즘을 통한 판단시 그 정합도가 임계값 이상일 경우 일치하는 것으로 판단하는 것을 특징으로 하는 발화 내용 파악 시스템.

[청구항 7]

제5항에 있어서,

상기 검증기는 후보가 된 음성 구간에 대해 검색어의 음소열에 대해 트라이폰(tri-phone) 모델 기준으로 음소 모델의 상태 정보로 음성구간을 음소열에 따른 프레임 단위로 할당을 하고, 트라이폰 모델에 대한 관측 확률값과 모노폰(mono-phone) 모델에 대한 관측 확률값의 비율에 대한 누적값을 구한 후 정규화하여 신뢰도 값을 산출한 후, 정규화된 신뢰도 값에 따라 상기 음성구간에 대해 최종적으로 검색된 결과물로 출력할 것인가를 판단하는 것을 특징으로 하는 발화 내용 파악 시스템.

[청구항 8]

제5항에 있어서,

상기 검증기는 상기 하나의 음성구간이 결정되면, 상기 하나의 음성구간을 기준으로 전후 일정 시간 범위 내에서 상기 검색 데이터베이스에서 상기 검색어에 대응하여 추출되는 문맥 키워드의 발음열에 일치하는 음소열이 있는지를 추가 검색하는 것을 특징으로 하는 발화 내용 파악 시스템.

[청구항 9]

제5항에 있어서,

상기 검색결과 히스토리 데이터베이스에 저장되는 상기 음성구간에 관련된 정보는 상기 하나의 음성구간이 포함된 파일명, 음성 데이터에서의 시작 위치 및 끝 위치, 상기 검색어, 검색된 구간에 대한 정규화된 신뢰도 값, 일치하는 문맥 키워드, 발성화자 성별인 것을 특징으로 하는 발화 내용 파악 시스템.

[청구항 10]

제5항에 있어서,

상기 파악부는,

검색어와 이에 대응되는 문맥 키워드별로 주제어가 설정되어 있는 주제어 데이터베이스;

상기 검색부에서 출력되는 검색 결과 정보 중에서 검색어와 문맥 키워드를 추출하여 상기 주제어 데이터베이스를 통해 대응되는 주제어를 검색하는 주제어 파악부; 및

- 상기 주제어 파악부로부터 검색 결과 정보와 주제어를 전달받아서 사용자에게 표시하는 출력부
를 포함하는 발화 내용 파악 시스템.
- [청구항 11] 제5항에 있어서,
상기 문맥 키워드는 동일한 의미를 나타내는 복수의 단어가 카테고리 방식으로 설정되어 있는 것을 특징으로 하는 발화 내용 파악 시스템.
- [청구항 12] 제5항에 있어서,
상기 검색부는 상기 검색어에 대해 동일한 의미를 가지는 별칭에 대해서는 발음열을 생성하여 동시에 검색을 수행하는 것을 특징으로 하는 발화 내용 파악 시스템.
- [청구항 13] 제2항에 있어서,
상기 음소 인식부는 음소 단위로 비터비(Viterbi) 알고리즘 및 토큰 패싱(Token passing) 알고리즘을 수행하여 대응되는 음소열을 생성하는 것을 특징으로 하는 발화 내용 파악 시스템.
- [청구항 14] 제2항에 있어서,
상기 음소 격자 형성부는 해당 음소열에 대해 음소의 시작점과 끝점, 그리고 지속시간 등의 정보를 함께 형성하는 것을 특징으로 하는 발화 내용 파악 시스템.
- [청구항 15] 제1항에 있어서,
상기 분할된 인덱싱 정보는 프레임 개수, 음소 개수, 특징 벡터, 음소의 상태별 관측 확률값, 타임 스탬프(time stamp), 음소열, 각 음소의 지속 시간을 포함하는 발화 내용 파악 시스템.
- [청구항 16] 발화 내용 파악 시스템이 음성 데이터에 대해 음소 기준의 음성 인식을 수행하여 인덱싱 가능하도록 저장하는 인덱싱 방법에 있어서,
음성 데이터를 입력받는 단계;
입력되는 음성 데이터의 음소들에 대해 프레임 단위의 음소 인식에 따른 음소 격자를 형성하는 단계; 및
미리 설정된 제한 시간 단위로 형성되는 상기 음소 격자를 분할된 인덱싱 정보로써 인덱싱 데이터베이스에 저장하는 단계를 포함하며,
상기 음소 격자를 형성하는 단계가 음성 데이터에서 음성 검출에 의해 유용한 음성 구간에 대해서만 수행되도록 제어되는 것을 특징으로 하는 인덱싱 방법.
- [청구항 17] 제16항에 있어서,
상기 음소 격자를 형성하는 단계에서, 상기 제한 시간 단위로 음소 격자를 형성하는 경우 이전 음성 구간 중 특정 시간 또는 특정

프레임의 음성 구간을 중첩하여 다음 음성 구간에 대한 음소 인식을 수행하여 불연속에 따른 정보 손실을 방지하며 음소 격자를 형성하는 것을 특징으로 하는 인덱싱 방법.

[청구항 18]

제16항에 있어서,

상기 음소 격자를 형성하는 단계는,

프레임 단위의 음성 데이터로부터 추출되는 특징 벡터를 사용하여 프레임 동기에 기준한 음소 인식에 의해 생성되는 다수의 음소열 후보군에 대해 역시간적인 연산을 수행하여 최종 선택되는 하나의 음성구간에 대응되는 음소 격자를 형성하는 것을 특징으로 하는 인덱싱 방법.

[청구항 19]

발화 내용 파악 시스템이 음성 데이터의 발화 내용을 파악하는 방법에 있어서,

사용자로부터 검색어를 입력받는 단계;

상기 검색어에 대응되는 음소 기준의 발음열을 생성하는 단계;

상기 발음열을 사용하여 인덱싱 데이터베이스—여기서 인덱싱 데이터베이스는 상기 음성 데이터에 대해 프레임 단위로 음소 기준의 음성 인식이 수행되어 형성되는 음소 격자가 제한 시간의 복수의 프레임별로 분할된 인덱싱 정보로써 저장되어 있음—에 저장되어 있는 분할된 인덱싱 정보에 대해 일치하는 음소열을 검색하여 1차 후보의 음성구간을 선정하는 단계;

상기 1차 후보의 음성구간에 대해서 음향학적 모델을 통해 일치 여부를 판단하여 하나의 음성구간을 결정하는 단계;

상기 하나의 음성구간을 기준으로 전후 일정 시간 범위 내에서 상기 검색어에 대응되는 문맥 키워드의 발음열에 일치하는 음소열이 있는지를 추가 검색하는 단계; 및

상기 검색어와 상기 문맥 키워드를 통해 상기 음성 데이터에 대한 주제어를 파악하여 사용자에게 제공하는 단계
를 포함하는 발화 내용 파악 방법.

[청구항 20]

제19항에 있어서,

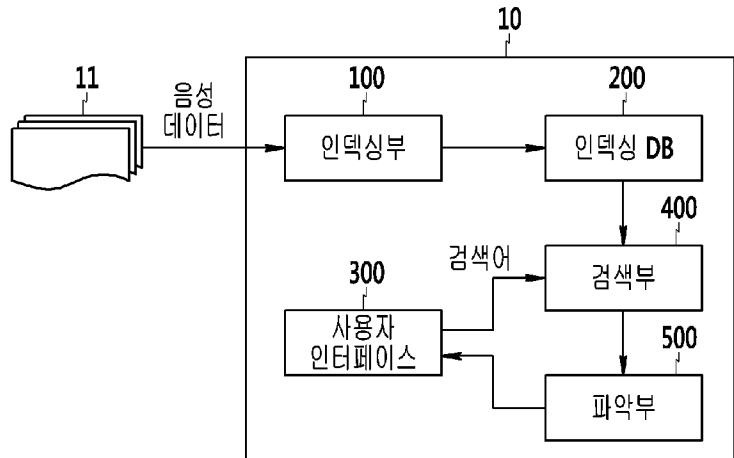
상기 주제어를 파악하여 사용자에게 제공하는 단계에서, 상기 주제어의 파악은 검색어와 이에 대응되는 문맥 키워드별로 주제어가 설정되어 있는 주제어 데이터베이스를 통해 수행되는 것을 특징으로 하는 발화 내용 파악 방법.

[청구항 21]

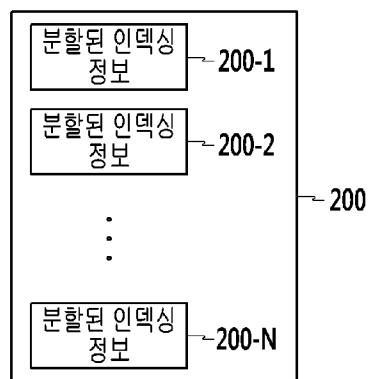
제19항에 있어서,

상기 검색어는 복수 개이며, 복수 개의 검색어가 연산 논리자에 의해 사용되어 상기 복수 개의 검색어의 조합에 의해 검색이 수행되는 것을 특징으로 하는 발화 내용 파악 방법.

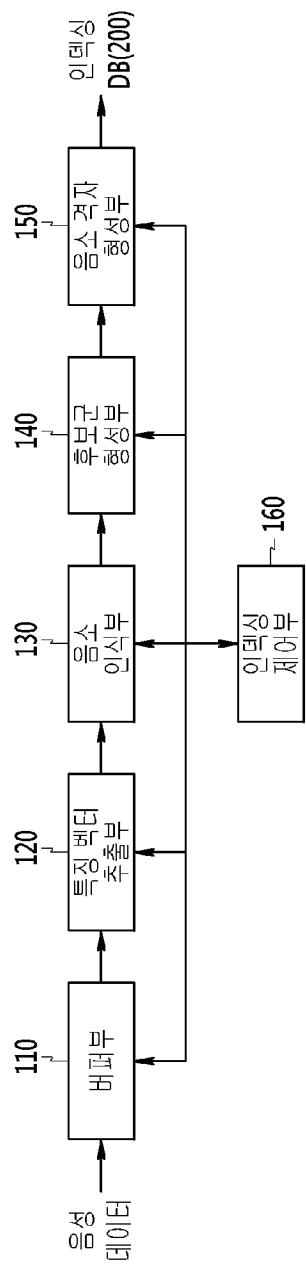
[Fig. 1]



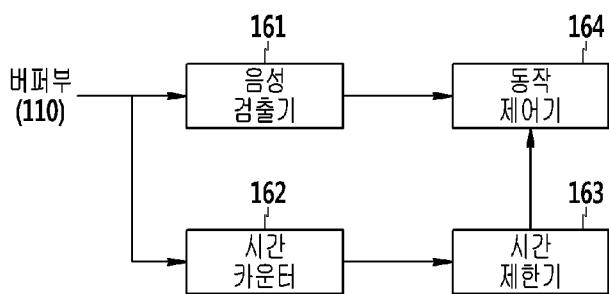
[Fig. 2]



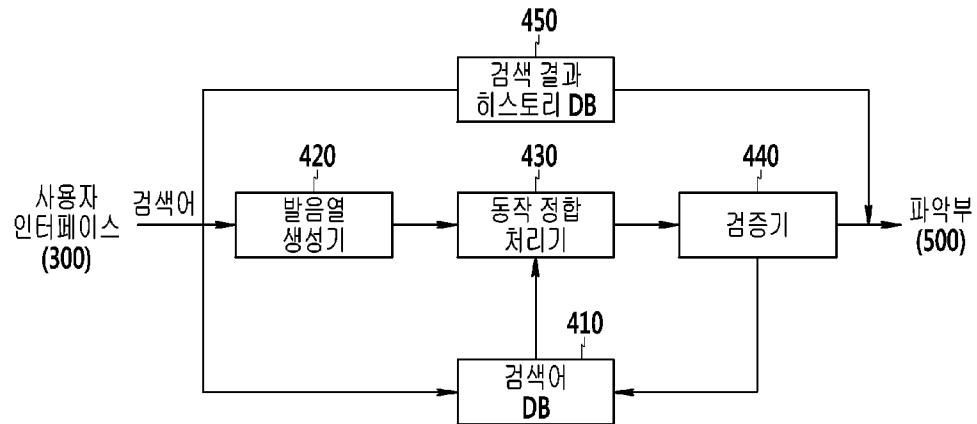
[Fig. 3]



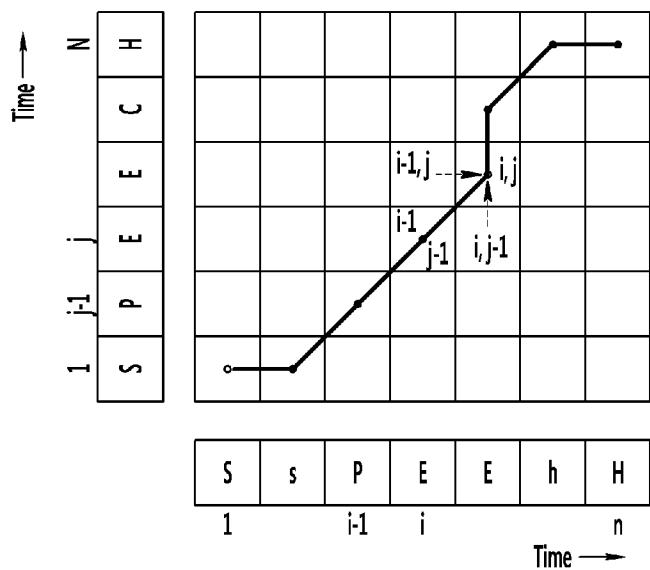
[Fig. 4]



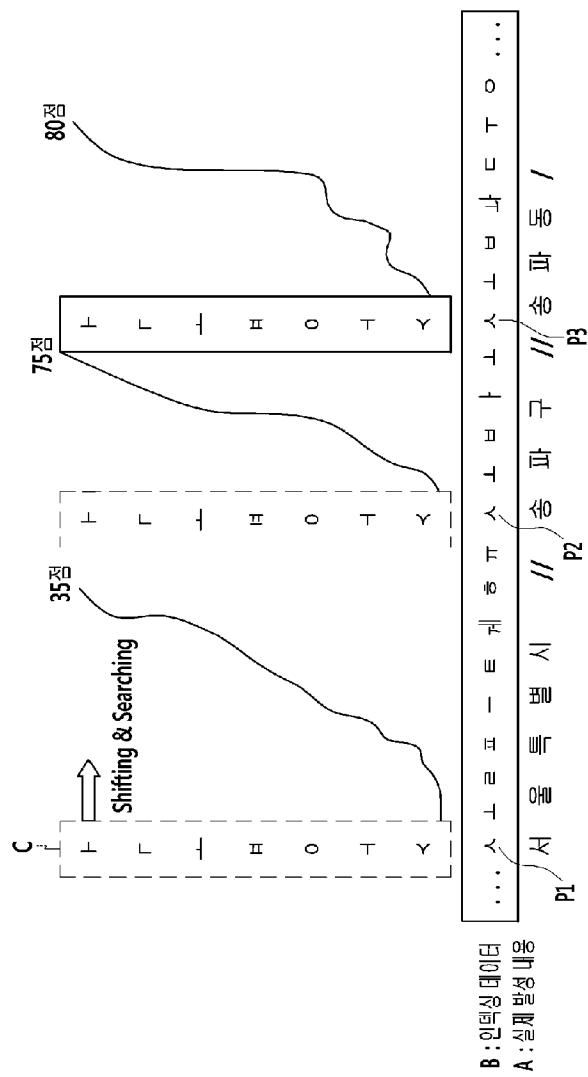
[Fig. 5]



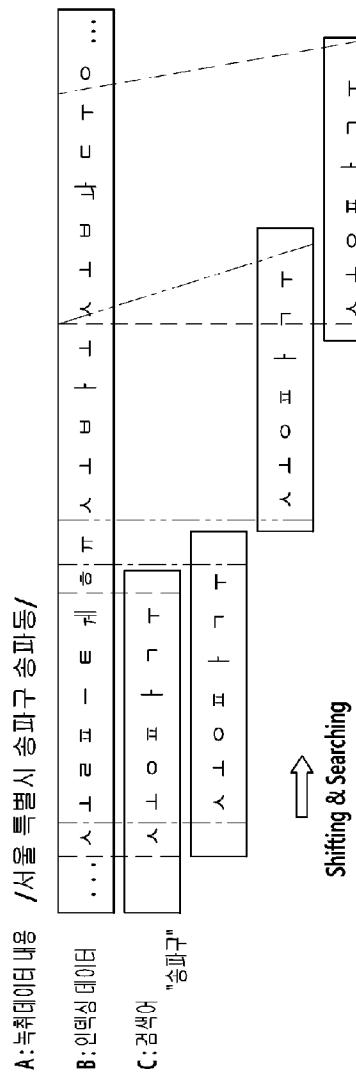
[Fig. 6]



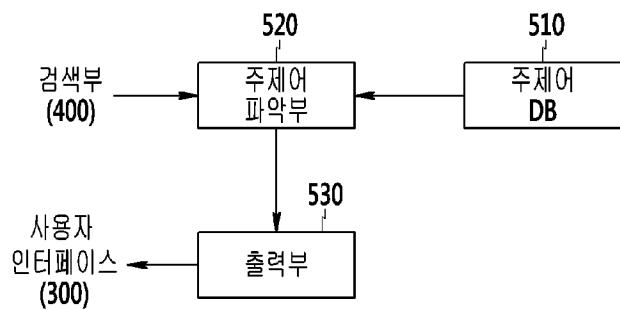
[Fig. 7]



[Fig. 8]



[Fig. 9]

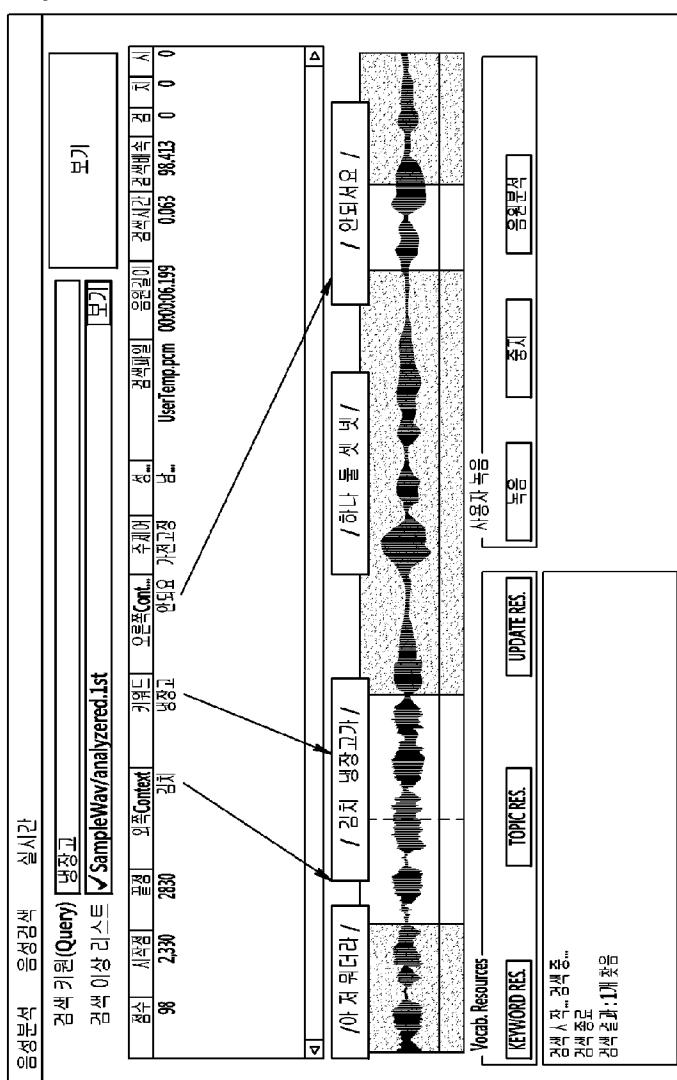


[Fig. 10]

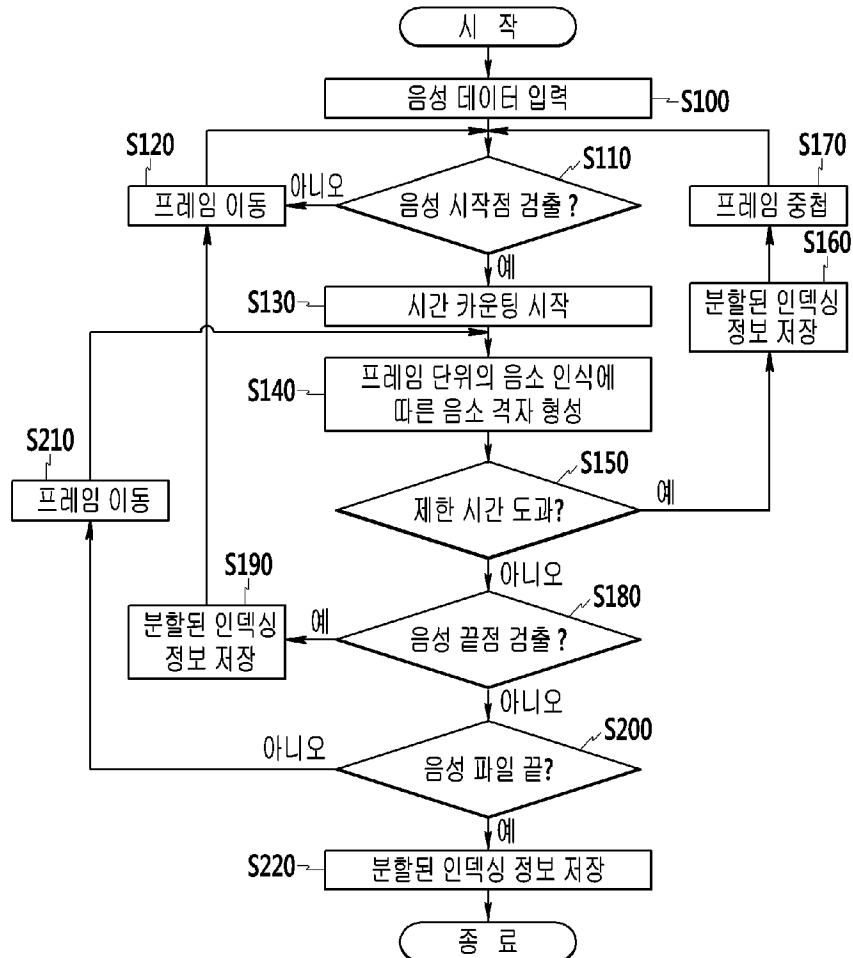
검색어	문맥 키워드	주제어
검색어1	문맥 키워드1 문맥 키워드2 ⋮ ⋮	주제어1 주제어2 ⋮ ⋮
검색어2	문맥 키워드3 문맥 키워드4 문맥 키워드5 ⋮ ⋮	주제어3 주제어4 주제어5 ⋮ ⋮
⋮	⋮	⋮

510

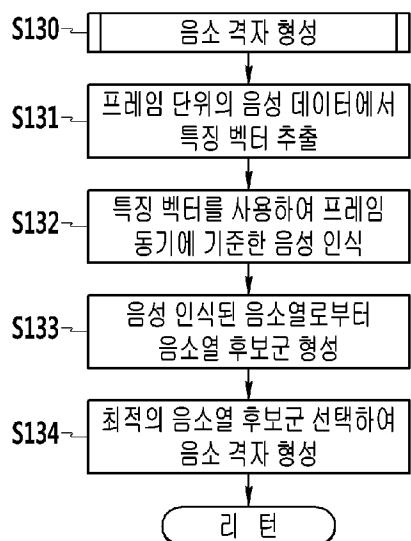
[Fig. 11]



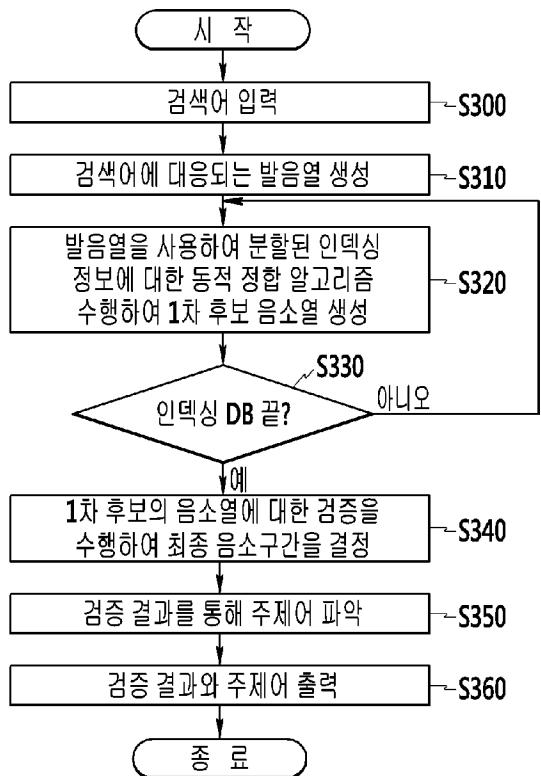
[Fig. 12]



[Fig. 13]



[Fig. 14]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2014/008706

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/28(2006.01)i, G10L 15/08(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 15/28; G10L 11/00; G06F 17/30; G10L 15/26; G10L 15/30; G10L 15/02; G06F 17/27; G10L 15/08

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: "subject*", "meaning*", "voice*", "recognition*", "understanding*", "recognition*", "search"

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KR 10-2010-0120977 A (POSTECH ACADEMY-INDUSTRY FOUNDATION) 17 November 2010	1-2,13-16,18
A	Abstract, claims 1, 3, paragraphs [0035]-[0037], figure 1.	3-12,17,19-21
Y	KR 10-2009-0085673 A (MOTOROLA, INC.) 07 August 2009 Abstract, claim 1, figure 2.	1-2,13-16,18
A		3-12,17,19-21
A	JP 2000-259628 A (ATR INTERPRETING TELECOMUNICATIONS RES. LAB.) 22 September 2000 Abstract, claim 1.	1-21
A	JP 2011-070192 A (RICOH CO., LTD.) 07 April 2011 Abstract, claim 1, figure 1.	1-21



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	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"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
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search	Date of mailing of the international search report
26 NOVEMBER 2014 (26.11.2014)	27 NOVEMBER 2014 (27.11.2014)
Name and mailing address of the ISA/KR  Korean Intellectual Property Office Government Complex-Daejeon, 189 Seonsa-ro, Daejeon 302-701, Republic of Korea Facsimile No. 82-42-472-7140	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/KR2014/008706

Patent document cited in search report	Publication date	Patent family member	Publication date
KR 10-2010-0120977 A	17/11/2010	KR 10-1069534 B1	30/09/2011
KR 10-2009-0085673 A	07/08/2009	CN 101558442 A EP 2092514 A2 US 2008-0130699 A1 WO 2008-115285 A2 WO 2008-115285 A3	14/10/2009 26/08/2009 05/06/2008 25/09/2008 18/12/2008
JP 2000-259628 A	22/09/2000	JP 3016779 B1	06/03/2000
JP 2011-070192 A	07/04/2011	CN 102023995 A CN 102023995 B EP 2306345 A2 EP 2306345 A3 US 2011-0071833 A1 US 8504367 B2	20/04/2011 30/01/2013 06/04/2011 13/07/2011 24/03/2011 06/08/2013

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G10L 15/28(2006.01)i, G10L 15/08(2006.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G10L 15/28; G10L 11/00; G06F 17/30; G10L 15/26; G10L 15/30; G10L 15/02; G06F 17/27; G10L 15/08

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: "주제*", "의미*", "음성*", "인식*", "이해*", "파악*", "검색*"

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y A	KR 10-2010-0120977 A (포항공과대학교 산학협력단) 2010.11.17 초록, 청구항 1,3, 식별번호 [0035]~[0037], 도면 1.	1-2, 13-16, 18 3-12, 17, 19-21
Y A	KR 10-2009-0085673 A (모토로라 인코포레이티드) 2009.08.07 초록, 청구항 1, 도면 2.	1-2, 13-16, 18 3-12, 17, 19-21
A	JP 2000-259628 A (ATR INTERPRETING TELECOMMUNICATIONS RES. LAB.) 2000.09.22 초록, 청구항 1.	1-21
A	JP 2011-070192 A (RICOH CO., LTD.) 2011.04.07 초록, 청구항 1, 도면 1.	1-21

 추가 문헌이 C(계속)에 기재되어 있습니다. 대응특허에 관한 별지를 참조하십시오.

* 인용된 문헌의 특별 카테고리:

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

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

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

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

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

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

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

“&” 동일한 대응특허문헌에 속하는 문헌

“P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌

국제조사의 실제 완료일

2014년 11월 26일 (26.11.2014)

국제조사보고서 발송일

2014년 11월 27일 (27.11.2014)

ISA/KR의 명칭 및 우편주소

대한민국 특허청

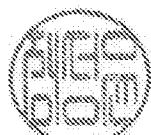
(302-701) 대전광역시 서구 청사로 189,
4동 (둔산동, 정부대전청사)

팩스 번호 +82-42-472-7140

심사관

정성윤

전화번호 +82-42-481-8322



국제조사보고서에서
인용된 특허문현

공개일

대응특허문현

공개일

KR 10-2010-0120977 A	2010/11/17	KR 10-1069534 B1	2011/09/30
KR 10-2009-0085673 A	2009/08/07	CN 101558442 A EP 2092514 A2 US 2008-0130699 A1 WO 2008-115285 A2 WO 2008-115285 A3	2009/10/14 2009/08/26 2008/06/05 2008/09/25 2008/12/18
JP 2000-259628 A	2000/09/22	JP 3016779 B1	2000/03/06
JP 2011-070192 A	2011/04/07	CN 102023995 A CN 102023995 B EP 2306345 A2 EP 2306345 A3 US 2011-0071833 A1 US 8504367 B2	2011/04/20 2013/01/30 2011/04/06 2011/07/13 2011/03/24 2013/08/06