



- (51) **국제특허분류:**
G10L 19/00 (2013.01)

(21) **국제출원번호:** PCT/KR2012/010785

(22) **국제출원일:** 2012년 12월 12일 (12.12.2012)

(25) **출원언어:** 한국어

(26) **공개언어:** 한국어

(30) **우선권정보:**
10-2012-0097345 2012년 9월 3일 (03.09.2012) KR

(71) **출원인:** 경희대학교 산학협력단 (UNIVERSITY-INDUSTRY COOPERATION GROUP OF KYUNG HEE UNIVERSITY) [KR/KR]; 446-701 경기도 용인시 기흥구 서천동 덕영대로 1732, Gyeonggi-do (KR).

(72) **발명자:** 이승룡 (LEE, Sung-Young); 463-500 경기도 성남시 분당구 구미동 삼환빌라 102동 102호, Gyeonggi-do (KR). 트바브이 (LE, Ba Vui); 446-701 경기도 용인시 기흥구 서천동 경희대학교 국제캠퍼스 전자정보대학 351호, Gyeonggi-do (KR).

(74) **대리인:** 특허법인 신지 (SINJI PATENT FIRM); 135-080 서울시 강남구 역삼동 648-1 BYC 빌딩 7층, 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, 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, PA, 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.

(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).

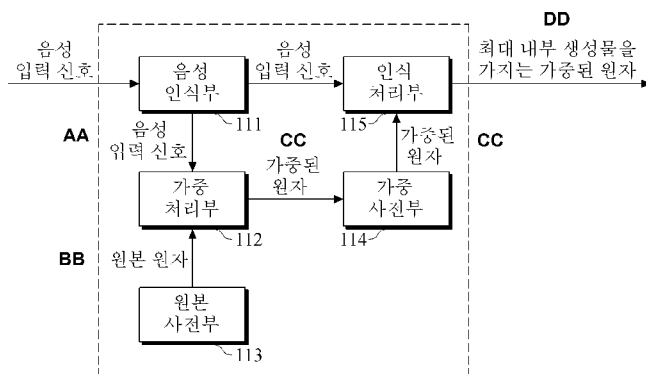
공개:
— 국제조사보고서와 함께 (조약 제 21 조(3))

공개:

— 국제조사보고서와 함께 (조약 제 21 조(3))

(54) Title: EMOTIVE SPEECH RECOGNITION APPARATUS AND METHOD

(54) 발명의 명칭 : 감정 음성 인식장치 및 방법



111 ... Speech recognition unit
112 ... Weight processing unit
113 ... Original dictionary unit
114 ... Weighted dictionary unit
115 ... Recognition processing unit
AA ... Speech input signal
BB ... Original atom
CC ... Weighted atom
DD ... Weighted atom having maximum internal product

(57) Abstract: The present invention relates to a technology for an emotive speech recognition apparatus and to a method for recognizing the emotional and physical state of a user from the speech of said user. A health management apparatus using a mobile terminal includes: a weight processing unit which generates and delivers a weighted atom by adding a weight to a received original atom using a spectrogram estimated on the basis of a speech input signal received from a speech recognition unit; and a recognition processing unit which generates a new residue based on the received original atom and on an allocated residue, and which, when the generated residue is smaller than a preset threshold value, determines the weighted atom as an output atom maximizing internal generation.

(57) **요약서:** 본 발명은 사용자의 음성으로부터 사용자의 감정 및 신체 상태를 인식하기 위한 감정 음성 인식장치 및 방법에 대한 기술이다. 본 발명에 따른 모바일 단말을 이용한 건강관리 장치는 음성 인식부로부터 수신된 음성 입력 신호에 기초하여 추정된 스펙트로그램을 이용하여 수신된 원본 원자에 가중치를 부가하여 가중된 원자를 생성하여 전달하는 가중 처리부 및 수신된 가중된 원자와 매칭된 레지듀에 기초하여 새로운 레지듀를 생성하고, 생성된 매칭된 레지듀가 미리 설정된 문턱값(Threshold)보다 작

으면 상기 가중된 원자를 상기 내부 생성을 최대화 시키는 출력 원자로 결정하는 인식 처리부를 포함한다.

명세서

발명의 명칭: 감정 음성 인식장치 및 방법

기술분야

- [1] 본 발명은 사용자의 음성으로부터 사용자의 감정을 인식할 수 있는 감정 음성 인식장치 및 그 방법에 관한 것이다.

배경기술

- [2] 음성을 통해 기계장치나 전자장치에 의사 또는 제어 명령을 전달하기 위한 음성 인식 기술은 더 편리한 서비스를 원하는 사용자들의 요구에 발맞추어 빠르게 발전하고 있다. 특히, 기술이 점차 발전하면서 사용자들은 일반적인 음성의 내용이나 명령만을 인식할 수 있는 음성 인식 기술에서 벗어나 사용자의 음성에 기초하여 사용자의 감정 상태까지 인식할 수 있는 감정 음성 인식 기술을 요구하고 있다. 감정은 사람의 행동이나 의도를 파악하는데 매우 중요한 역할을 담당하기 때문에, 사람과 기계장치 사이에 상호작용을 증가시킬 수 있다. 이를 통해, 종래의 단순히 명령에만 동작하는 것이 아니라, 사용자의 상황에 따라 적합한 서비스를 제공해주는 장치의 개발 등에 적용할 수 있다.
- [3] 감정 음성 인식 기술은 말하는 사람의 음성을 인식하고, 인식된 음성으로부터 말하는 사람의 감정 또는 신체적 상태를 인식할 수 있는 음성 처리 응용 방법의 하나이다. 음성으로부터 사용자의 감정을 인식하는 방법은 일반적으로 특징 추출(Feature Extraction) 방법을 사용하고 있다. 그리고 음성으로부터 인식된 감정의 정확도는 추출된 특징에 따라 달라지게 된다.
- [4] 현재 일반적으로 사용되는 특징 추출 방법은 음성의 피치(Pitch), 낭독속도(Speech Rate), 강도(Intensity), MFCC(Mel-Frequency Cepstral Coefficients) 및 LPCC(Linear Prediction Cepstral Coefficients) 등을 이용하여 음성으로부터 감정 상태를 인식한다. 상술한 특징 추출 방법에 이용되는 요소들은 시간과 주파수의 서로 다른 영역(Domain)에서 음성 신호로부터 추출된다. 이러한 요소들을 시간 도메인에서 주파수 도메인으로 또는 그 반대 방향으로 변환하기 위해 고속 푸리에 변환(Fast Fourier Transform, FFT)이 일반적으로 사용된다. 하지만 고속 푸리에 변환의 경우, 시간에 국한된 함수의 표면만을 제공한다.
- [5] 이러한 고속 푸리에 변환의 한계를 극복하고자 대응 추적(Matching Pursuit) 알고리즘을 사용한다. 대응 추적은 원본 신호를 미리 설정된 원자(Atom)들의 사전(Dictionary)으로 분해하는 반복 알고리즘이다. 대한민국 공개특허 제10-2011-0097304호에는 대응 추적을 이용하여 심음을 분석하는 기술에 대해 기재되어 있다. 하지만 종래의 대응 추적은 원자의 사전이 거대하기 때문에 매우 많은 내적 연산을 필요로 한다. 그래서 모든 원자들을 위한 검색은 많은 시간과 노력을 필요로 한다.

발명의 상세한 설명

기술적 과제

- [6] 본 발명에서 해결하고자 하는 과제는 사용자의 음성으로부터 사용자의 감정 및 신체 상태를 인식하기 위한 기술로서, 특히 연산량을 감소시켜 더욱 빠른 결과를 도출할 수 있는 알고리즘을 적용한 감정 음성 인식장치 및 방법을 제공하는 것이다.

과제 해결 수단

- [7] 본 발명에 따른 모바일 단말을 이용한 건강관리 장치는 사용자로부터 음성 입력 신호를 수신 받아 전달하는 음성 인식부, 다수의 원본 원자를 저장하고 있으며, 수신된 요청에 따라 원본 원자를 전달하는 원본 사전부, 음성 인식부로부터 수신된 음성 입력 신호에 기초하여 스펙트로그램(Spectrogram)을 추정하고, 추정된 스펙트로그램을 이용하여 수신된 원본 원자에 가중치를 부가하여 가중된 원자를 생성하여 전달하는 가중 처리부, 상기 가중 처리부로부터 수신된 가중된 원자를 저장하며, 수신된 요청에 따라 가중된 원자를 전달하는 가중 사전부 및 음성 인식부로부터 수신된 음성 입력 신호를 레지듀(Residue)로 배정하고, 가중 사전부에 기초하여 원자 선택 기준을 수정하여 가중 사전부에 가중된 원자를 요청하고, 요청에 따라 수신된 가중된 원자와 배정된 레지듀에 기초하여 새로운 레지듀를 생성하고, 생성된 배정된 레지듀가 미리 설정된 문턱값(Threshold)보다 작으면 상기 가중된 원자를 상기 내부 생성을 최대화 시키는 출력 원자로 결정하는 인식 처리부를 포함한다.

발명의 효과

- [8] 본 발명에 따른 감정 음성 인식장치 및 방법을 통해 단순히 명령이나 내용만을 인식하던 음성 인식과 달리 사용자의 감정 및 신체상태를 인식할 수 있기 때문에, 사용자의 상황에 따라 적합한 서비스를 제공받을 수 있다.

도면의 간단한 설명

- [9] 도 1은 본 발명에 따른 감정 음성 인식 장치의 일 실시예의 구성도이다.
 [10] 도 2는 본 발명의 일 실시예에 따른 대응 추적 알고리즘의 흐름도이다.
 [11] 도 3은 본 발명의 일 실시예에 따른 가중된 사전을 생성하는 방법을 나타내는 흐름도이다.
 [12] 도 4는 본 발명의 일 실시예에 따른 원자의 선택기준을 수정하는 방법을 나타내는 흐름도이다.
 [13] 도 5는 본 발명의 일 실시예에 따른 선택된 원자와 입력 신호의 스펙트로그램 사이의 관계를 나타내는 일례의 도면이다.
 [14] 도 6은 본 발명의 일 실시예에 따른 인식을 위한 대응 추적 알고리즘의 선택된 원자에 기반한 특징 추출 방법을 나타내는 흐름도이다.

발명의 실시를 위한 최선의 형태

- [15] 이하, 첨부된 도면들을 참조하여 본 발명의 실시예를 상세하게 설명한다. 본

명세서에서 사용되는 용어는 실시예에서의 기능 및 효과를 고려하여 선택된 용어들로서, 그 용어의 의미는 사용자 또는 운용자의 의도 또는 업계의 관례 등에 따라 달라질 수 있다. 따라서 후술하는 실시예들에서 사용된 용어의 의미는, 본 명세서에 구체적으로 명시된 경우에는 명시된 정의에 따르며, 구체적으로 명시하지 않는 경우, 당업자들이 일반적으로 인식하는 의미로 해석되어야 할 것이다.

- [16] 도 1은 본 발명에 따른 감정 음성 인식 장치의 일 실시예의 구성도이다.
- [17] 도 1을 참조하면, 본 발명에 따른 감정 음성 인식 장치(100)는 음성 인식부(111), 가중 처리부(112), 원본 사전부(113), 가중 사전부(114) 및 인식 처리부(115)를 포함한다.
- [18] 음성 인식부(111)는 사용자로부터 음성 입력 신호를 인식하여 수신 받고, 수신된 음성 입력 신호를 가중 처리부(112) 및 인식 처리부(115)로 전달한다.
- [19] 가중 처리부(112)는 단시간 푸리에 변환(Short Time Fourier Transform)을 이용하여 음성 인식부(111)로부터 수신된 음성 입력 신호의 스펙트로그램(Spectrogram)을 추정한다. 그리고 추정된 스펙트로그램의 값을 0부터 1까지로 정규화한다. 모든 스펙트로그램의 값은 0부터 1까지 정규화가 가능하기 때문에, 원자들의 가중치 또한 0부터 1까지로 정규화 될 수 있다.
- [20] 가중 처리부(112)는 원본 사전부(113)로부터 원본 원자를 전달 받는다. 그리고 시간과 중앙 주파수의 최 근사값을 가지는 스펙트로그램의 대응값을 찾는다. 이를 통해 가중치 값을 대응하는 원본 원자에 할당하여 가중된 원자를 생성한다. 다음으로 가중 처리부(112)는 생성된 가중된 원자를 내림차순으로 분류하고 가중 사전부(114)로 전달한다.
- [21] 원본 사전부(113)는 가중되지 않은 기존의 원자를 저장하고 있으며 가중 처리부(112)의 요청에 따라 저장된 원자를 전달한다.
- [22] 가중 사전부(114)는 가중 처리부(112)로부터 전달된 가중된 원자를 저장한다. 그리고 인식 처리부(115)의 요청에 따라 저장된 가중된 원자를 전달한다.
- [23] 인식 처리부(115)는 음성 인식부(111)로부터 수신된 음성 입력 신호를 초기 레지듀(residue) 또는 현재 레지듀로 배정한다. 그리고 배정된 현재 레지듀와 가중 사전부(114)에 기초하여 원자 선택 기준을 수정한다. 생성된 스펙트로그램 가중치에 의해 가중 사전부가 생성되면, 음성 입력 신호에 대한 최적의 대응 원자를 사전에서 검색하기 위한 기준이 변경된다. 가중된 값의 내림차순에 따라 가중된 원자를 선택하고, 선택된 원자와 현재의 레지듀 사이의 내부 생성물을 계산한다. 그리고 생성된 내부 생성물이 감소하며, 문턱값보다 작은 경우 해당 원자를 내부 생성을 최대화 시킨 출력 원자로 결정한다.
- [24] 도 2는 본 발명의 일 실시예에 따른 대응 추적 알고리즘의 흐름도이다.
- [25] 도 2를 참조하면 본 발명의 일 실시예에 따른 대응 추적 알고리즘은 먼저, 사용자로부터 음성 입력 신호를 전달받는다(201). 그리고 수신된 음성 입력 신호를 초기 레지듀(residue)로 배정한다(202). 수신된 입력 신호를 수학적 1과

같이 초기 레지듀로 초기화 하여 배정한다.

[26] 수학식 1

$$Rf_1 = f(t)$$

[27]

$$Rf_1 = f(t)$$

[28] 수학식 1에서 R은 함수 f의 레지듀이고,

$$Rf_1$$

는 초기 레지듀이고,

$$f(t)$$

는 입력 신호를 나타낸다.

[29] 다음으로 수신된 음성 입력 신호로부터 스펙트로그램 가중치(Spectrogram Weighting)를 생성한다(203). 일반적인 대응 추적 알고리즘은 사전으로부터 새로운 원자를 선택하는데 있어서, 최대 값을 가지는 내부 생성물 전체를 검색해야 한다. 하지만 충분한 데이터를 계산하기 위해 사건의 크기는 충분히 커야 하므로 새로운 원자를 선택하는데 많은 연산량과 시간을 필요로 한다. 이러한 문제를 해결하기 위해 스펙트로그램 가중치를 이용하여 연산량을 줄일 수 있다. 생성된 스펙트로그램 가중치를 이용하여 가중된 사전을 생성한다(204). 가중치에 의해 가중된 사전은 모든 원자들이 동일한 가중치를 가지지 않는다. 따라서 더 큰 가중치를 가지는 원자들은 출력 원자로 선택될 수 가능성이 높게 된다. 203단계 및 204단계에 대한 구체적인 설명은 후술하는 도 3에서 설명하도록 한다.

[30] 다음으로 초기 레지듀와 가중된 사전에 기초하여 원자 선택 기준을 수정한다(205). 생성된 스펙트로그램 가중치에 의해 가중된 사전이 생성되면, 입력 신호에 대한 최적의 대응 원자를 사전에서 검색하여 선택하기 위한 기준들이 변경된다. 변경된 기준에 따라 가중된 사전에서 가중된 원자를 선택한다. 초기 레지듀와 가중된 사전에 기초하여 원자 선택 기준을 수정하는 방법은 후술하는 도 4에서 설명하도록 한다.

[31] 다음으로 새로운 레지듀를 계산한다(206). 수학식 2와 같이, 현재의 레지듀는 내부 생성물에 의해 제거되어 새로운 레지듀를 생성한다.

[32] 수학식 2

$$Rf_{n+1} = Rf_n - a_n g_m$$

[33] 수학식 2에서

$$Rf_n$$

는 현재의 레지듀(또는 매정된 레지듀)이고, n 은 원자의 색인(index)이고,

$$a_n$$

은 가중치 인자(Weight Factor)이고,

$$g_m$$

은 원자이고,

$$a_n g_m$$

는 내부 생성물이며 가중된 원자이고,

$$Rf_{n+1}$$

는 새로운 레지듀이다.

- [34] 그리고 현재의 레지듀가 문턱값(Threshold)보다 더 커지는 값을 가지는지 여부를 확인한다(207).
- [35] 만약 원자들의 수가 현재의 레지듀가 문턱값보다 더 큰 값을 가지도록 하는 값을 가진다면, 처리과정을 종료한다(208). 처리과정 종료에 따른 출력은 대응하는 계수를 가진 가중된 원자의 그룹이다.
- [36] 만약 원자들의 수가 현재의 레지듀가 문턱값보다 더 큰 값을 가지도록 하는 값을 가지지 못한다면, 원자 선택 기준을 다시 수정한다(209). 만약 원자들의 수가 현재의 레지듀가 문턱값보다 더 큰 값을 가지도록 하는 값을 가지지 못한다면, 원자 선택 기준을 수정하는 205 단계부터 처리 과정을 다시 수행한다.
- [37] 도 3은 본 발명의 일 실시예에 따른 가중된 사전을 생성하는 방법을 나타내는 흐름도이다.
- [38] 도 3을 참조하면, 본 발명의 일 실시예에 따른 가중된 사전을 생성하는 방법은 먼저, 단시간 푸리에 변환(Short Time Fourier Transform)을 이용하여 수신된 음성 입력 신호에 스펙트로그램을 추정한다(301).
- [39] 스펙트로그램 값을 0부터 1까지 정규화한다(302). 모든 스펙트로그램의 값은 0부터 1까지로 정규화되기 때문에, 원자들의 가중치 또한 0부터 1까지로 정규화될 수 있다.
- [40] 원본 사전으로부터 각각의 원자를 선택한다(303). 그리고 시간과 중앙 주파수의 최 근사값을 가지는 스펙트로그램의 대응값을 찾는다(304). 원본 사전에서 각각의 원자는 주파수축 및 시간축에 의해 결정된 정규화된 스펙트로그램의 최 근접값을 위한 검색에 의해 가중된다.
- [41] 가중치 값을 대응 원자에 할당한다(305). 새로운 사전, 즉 가중된 사전을 생성하기 위해 각각의 원자는 가중치 계수를 나타내는 부가정보가 부착(tag)된다.
- [42] 가중치 값의 내림차순으로 원자를 분류한다(306). 가중치 계수가 부착된

- 새로운 가중된 원자들은 가중치의 내림차순으로 분류된다. 그리고 가중치의 내림차순으로 분류된 가중된 원자들이 저장되어 가중된 사전을 생성한다(307).
- [43] 도 4는 본 발명의 일 실시예에 따른 원자의 선택기준을 수정하는 방법을 나타내는 흐름도이다.
- [44] 도 4를 참조하면, 본 발명의 일 실시예에 따른 원자의 선택기준을 수정하는 방법은 먼저, 내림차순으로 분류된 가중된 사전에서 원자를 선택한다(401). 가중된 사전에서 가중치의 내림차순에 따라 분류된 원자들 중에서 원자를 선택한다. 그리고 현재의 레지듀의 내부 생성물(레지듀)과 선택된 원자를 계산한다(402).
- [45] 다음으로 현재의 레지듀의 내부 생성물과 선택된 원자가 감소하고 문턱값보다 작은지 여부를 확인한다(403). 만약 현재의 레지듀의 내부 생성물과 선택된 원자가 감소하고 문턱값보다 작다면, 선택된 원자를 내부 생성물을 최대화하는 출력 원자로 결정한다(404). 만약 만약 현재의 레지듀의 내부 생성물과 선택된 원자가 감소하고 문턱값보다 작지 않다면, 가중치의 내림차순에서 다시 원자를 선택한다(405).
- [46] 본 발명에서 선택된 원자와 입력 신호의 스펙트로그램을 가지는 한 그룹의 원자들에 대한 대응 추적의 출력에서는 원자들은 스펙트로그램의 더 어두운 영역에 대응하는 높은 에너지 영역에 할당된다. 반면에, 신호의 스펙트로그램은 단시간 푸리에 변환(Short-Time Fourier Transform, STFT)을 이용하여 쉽게 계산할 수 있다. 이에 따라, 사전에 가중치를 두는 스펙트로그램을 사용할 수 있게 된다. 그래서 더 큰 가중치를 가지는 원자들은 상대적으로 적은 가중치를 원자들에 비해서 출력 원자로 선택될 가능성이 높아지게 된다. 이처럼 사전의 원자에 가중치를 적용하여 선택 빈도에 차이를 둬으로써 전체 연산량이 감소하고 연산에 필요한 시간이 감소하게 된다. 이는 전체 대응 추적 알고리즘의 성능을 향상시킬 수 있다.
- [47] 도 5는 본 발명의 일 실시예에 따른 인식을 위한 대응 추적 알고리즘의 선택된 원자에 기반한 특징 추출 방법을 나타내는 흐름도이다.
- [48] 도 5를 참조하면, 본 발명의 일 실시예에 따른 인식을 위한 대응 추적 알고리즘의 선택된 원자에 기반한 특징 추출 방법은 먼저, 대응 추적을 이용하여 입력 신호를 원자로 분해한다(501). 대응 추적을 통해 입력 신호를 하나의 미리 설정된 원자들의 시간-주파수 표현으로 분해한다. 그리고 시간 값의 히스토그램(Histogram)을 계산한다(502). 시간 특징들은 시간 영역의 추정 히스토그램에 의해 추출되고, 이 히스토그램은 입력 신호의 포락선(Envelope)의 고유성과 같다.
- [49] 다음으로 중첩 윈도우(Overlapping Windows)를 히스토그램을 분해한다(503). 중첩 윈도우 기술은 히스토그램을 작은 프레임의 윈도우로 나눈다. 그리고 각각의 윈도우에 대응하는 원자들의 수를 확인한다(504). 다음으로 각각의 윈도우에 대응하는 원자들의 자세의 평균과 편차를 계산한다(505). 윈도우에

- 포함된 모든 원자들의 자세의 평균과 편차는 계산되어 특징 벡터로 공식화된다.
- [50] 다음으로 주파수 값의 히스토그램을 계산한다(506). 주파수 특징들은 주파수 영역의 추정 히스토그램에 의해 추출된다. 그리고 주파수 영역을 서브 밴드(Sub-Band)로 분할한다(507). 주파수 값의 히스토그램은 둘 이상의 서브 밴드 주파수로 나누어 진다. 다음으로 각각의 서브 밴드에 대응하는 원자들의 수를 확인한다(508).
- [51] 각각의 서브 밴드에 대응하는 원자들의 자세의 평균과 편차를 계산한다(509). 각각의 서브 밴드 주파수에 있어서, 동일한 서브 밴드의 자세의 평균과 편차는 계산되어 특징 벡터로 공식화된다.
- [52] 다음으로 SVM 분류를 이용하여 원자들을 다른 클래스 라벨(Class Label)처럼 분류한다(510). SVM 분류(Support Vector Machine Classification)는 분류 알고리즘의 하나로서 결정트리, 신경망 등의 분류 알고리즘에 비해 일반화 및 분류율이 뛰어난 분류 방법이다. SVM 분류는 지지 벡터 머신이라고도 불린다. SVM 분류는 지도 학습에서 사용되는 방법으로, 주어진 자료에 대해서 그 자료들을 분리하는 초평면 중에서, 가장 거리가 먼 초평면을 찾는 방법이다.
- [53] 이상 바람직한 실시 예를 들어 본 발명을 상세하게 설명하였으나, 본 발명은 전술한 실시 예에 한정되지 않고, 본 발명의 기술적 사상의 범위 내에서 당 분야에서 통상의 지식을 가진 자에 의하여 여러 가지 변형이 가능하다.

청구범위

- [청구항 1] 사용자로부터 음성 입력 신호를 수신 받아 전달하는 음성 인식부; 다수의 원본 원자를 저장하고 있으며, 수신된 요청에 따라 상기 원본 원자를 전달하는 원본 사전부; 상기 음성 인식부로부터 수신된 상기 음성 입력 신호에 기초하여 스펙트로그램(Spectrogram)을 추정하고, 추정된 스펙트로그램을 이용하여 수신된 상기 원본 원자에 가중치를 부가하여 가중된 원자를 생성하여 전달하는 가중 처리부; 상기 가중 처리부로부터 수신된 가중된 원자를 저장하며, 수신된 요청에 따라 상기 가중된 원자를 전달하는 가중 사전부; 및 상기 음성 인식부로부터 수신된 상기 음성 입력 신호를 레지듀(Residue)로 배정하고, 상기 가중 사전부에 기초하여 원자 선택 기준을 수정하여 상기 가중 사전부에 가중된 원자를 요청하고, 상기 요청에 따라 수신된 가중된 원자와 상기 배정된 레지듀에 기초하여 새로운 레지듀를 생성하고, 상기 생성된 배정된 레지듀가 미리 설정된 문턱값(Threshold)보다 작으면 상기 가중된 원자를 상기 내부 생성을 최대화 시키는 출력 원자로 결정하는 인식 처리부; 를 포함하는 것을 특징으로 하는 감정 음성 인식 장치.
- [청구항 2] 제 1항에 있어서, 상기 가중 처리부는 단시간 푸리에 변환(Short Time Fourier Transform)을 이용하여 상기 음성 입력 신호의 스펙트로그램을 추정하는 것을 특징으로 하는 감정 음성 인식 장치.
- [청구항 3] 제 1항에 있어서, 상기 스펙트로그램 및 상기 가중치는 0부터 1까지로 정규화 되는 것을 특징으로 하는 감정 음성 인식 장치.
- [청구항 4] 제 1항에 있어서, 상기 인식 처리부는, 상기 배정된 레지듀가 상기 미리 설정된 문턱값(Threshold)보다 크면 상기 가중 사전부에 새로운 가중된 원자를 요청하고, 상기 새로운 레지듀와 상기 새로운 가중된 원자를 이용하여 다른 새로운 레지듀를 생성하는 것을 특징으로 하는 감정 음성 인식 장치.
- [청구항 5] 제 1항에 있어서, 상기 새로운 레지듀는,

$$Rf_{n+1} = Rf_n - \alpha_n g_{\gamma n} \quad \text{에 의해 산출되며,}$$

상기 Rf_{n+1} 은 상기 새로운 레지듀이고, 상기 Rf_n 는 상기
 배정된 레지듀이고, a_n 은 상기 가중치이고, g_m 은 상기
 원본 원자이고, $a_n g_m$ 는 가중된 원자인 것을 특징으로 하는

감정 음성 인식 장치.

[청구항 6]

제 1항에 따른 감정 음성 인식 장치의 감정 음성 인식 방법에
 있어서,

사용자로부터 음성 입력 신호를 수신 받는 단계;

상기 수신된 음성 입력 신호를 레지듀로 배정하는 단계;

상기 수신된 음성 입력 신호에 기초하여 가중된 원자를 생성하는
 단계;

상기 배정된 레지듀와 상기 가중된 사전에 기초하여 원자 선택
 기준을 수정하는 단계;

상기 배정된 레지듀와 상기 가중된 원자를 이용하여 새로운
 레지듀를 생성하는 단계;

상기 배정된 레지듀를 문턱값(Threshold)과 비교하는 단계;

상기 배정된 레지듀가 문턱값보다 더 커지면, 상기 가중된 원자를
 출력하는 단계;

를 포함하는 것을 특징으로 하는 감정 음성 인식 방법.

[청구항 7]

제 6항에 있어서,

상기 수신된 음성 입력 신호에 기초하여 가중된 원자를 생성하는
 단계는,

상기 수신된 음성 입력 신호에 기초하여 스펙트로그램을 추정하는
 단계;

상기 스펙트로그램의 값을 정규화하는 단계;

상기 스펙트로그램의 시간과 주파수에서 미리 저장된 원본 원자와
 대응하는 최근사값을 찾는 단계;

상기 원본 원자와 대응하는 상기 스펙트로그램의 값을 가중치로
 하고, 상기 가중치 값을 상기 원본 원자에 할당하여 가중된 원자를
 생성하는 단계; 및

상기 가중된 원자를 내림차순으로 분류하는 단계;

를 더 포함하는 것을 특징으로 하는 감정 음성 인식 방법.

[청구항 8]

대응 추적(Matching Pursuit) 알고리즘을 이용하여 음성 입력
 신호를 원자로 분해하는 단계;

상기 원자에 기초하여 시간 영역에 대한 상기 원자의 자세의 평균
 및 편차를 계산하는 단계;

상기 원자에 기초하여 주파수 영역에 대한 상기 원자의 자세의 평균 및 편차를 계산하는 단계;
 상기 시간 영역에 대한 상기 원자의 평균 및 편차와 상기 주파수 영역에 대한 상기 원자의 평균 및 편차를 SVM 분류를 통해 분류하는 단계;
 를 포함하는 것을 특징으로 하는 대응 추적 알고리즘의 선택된 원자에 기반한 특징 추출 방법.

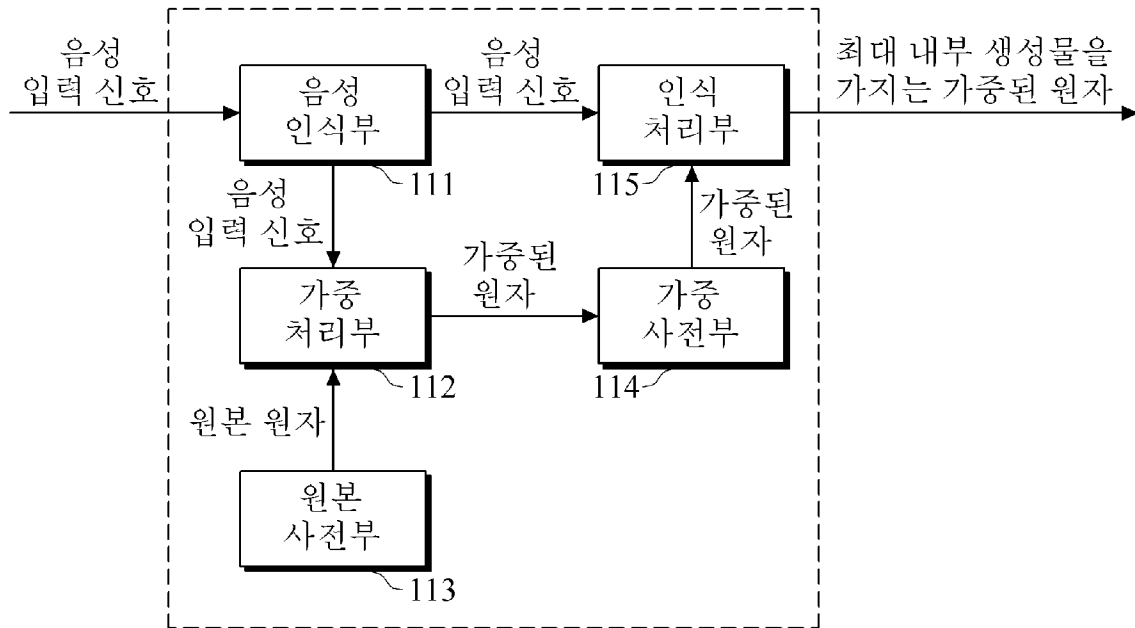
[청구항 9]

제 8항에 있어서,
 상기 원자에 기초하여 시간 영역에 대한 상기 원자의 자세의 평균 및 편차를 계산하는 단계는,
 상기 원자에 기초하여 시간 값의 히스토그램(Histogram)을 계산하는 단계;
 중첩 윈도우(Overlapping Windows)를 이용하여 상기 히스토그램을 분해하는 단계;
 상기 중첩 윈도우 각각에 대응하는 원자들의 수를 확인하는 단계;
 및
 상기 중첩 윈도우 각각에 대응하는 원자들의 자세의 평균과 편차를 계산하는 단계;
 를 더 포함하는 것을 특징으로 하는 대응 추적 알고리즘의 선택된 원자에 기반한 특징 추출 방법.

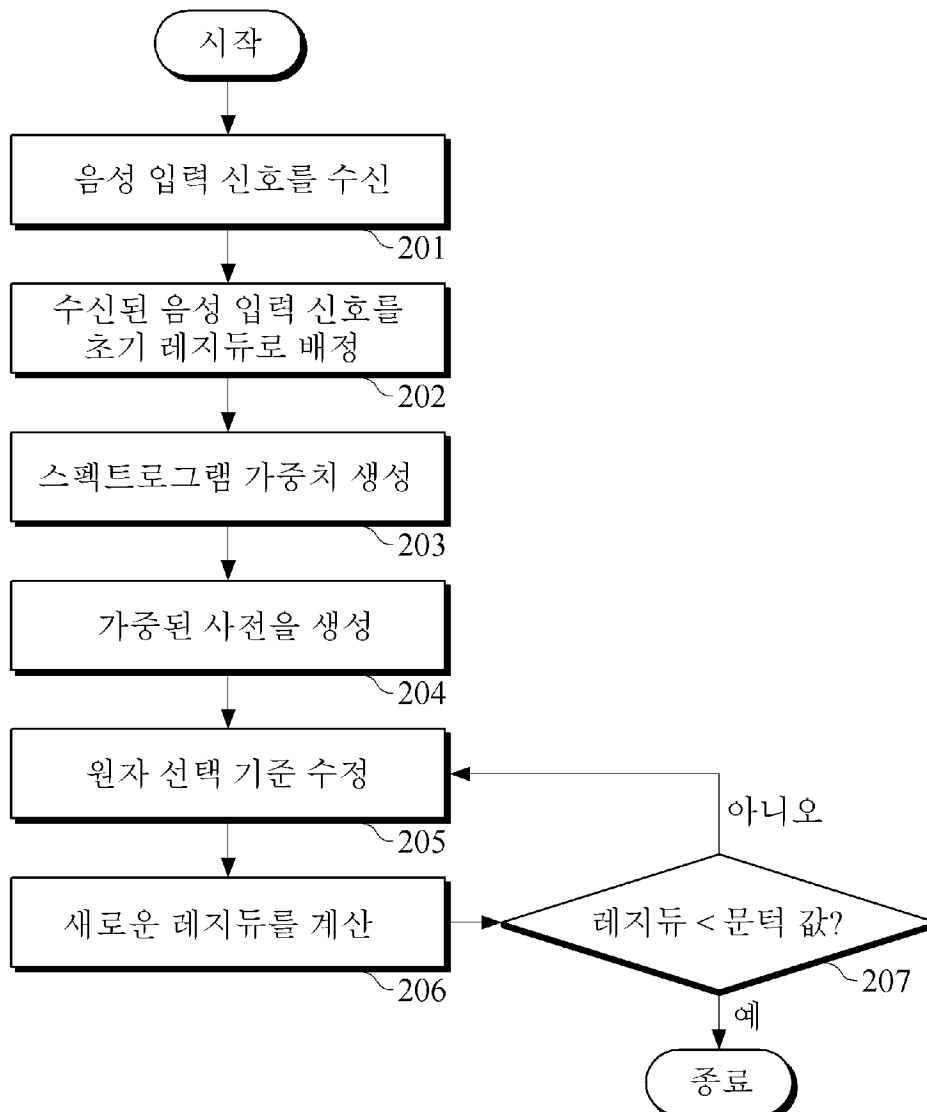
[청구항 10]

제 8항에 있어서,
 상기 원자에 기초하여 주파수 영역에 대한 상기 원자의 자세의 평균 및 편차를 계산하는 단계는,
 상기 원자에 기초하여 주파수 값의 히스토그램(Histogram)을 계산하는 단계;
 서브 밴드(Sub-Band)를 이용하여 상기 히스토그램을 분해하는 단계;
 상기 서브 밴드 각각에 대응하는 원자들의 수를 확인하는 단계; 및
 상기 서브 밴드 각각에 대응하는 원자들의 자세의 평균과 편차를 계산하는 단계;
 를 더 포함하는 것을 특징으로 하는 대응 추적 알고리즘의 선택된 원자에 기반한 특징 추출 방법.

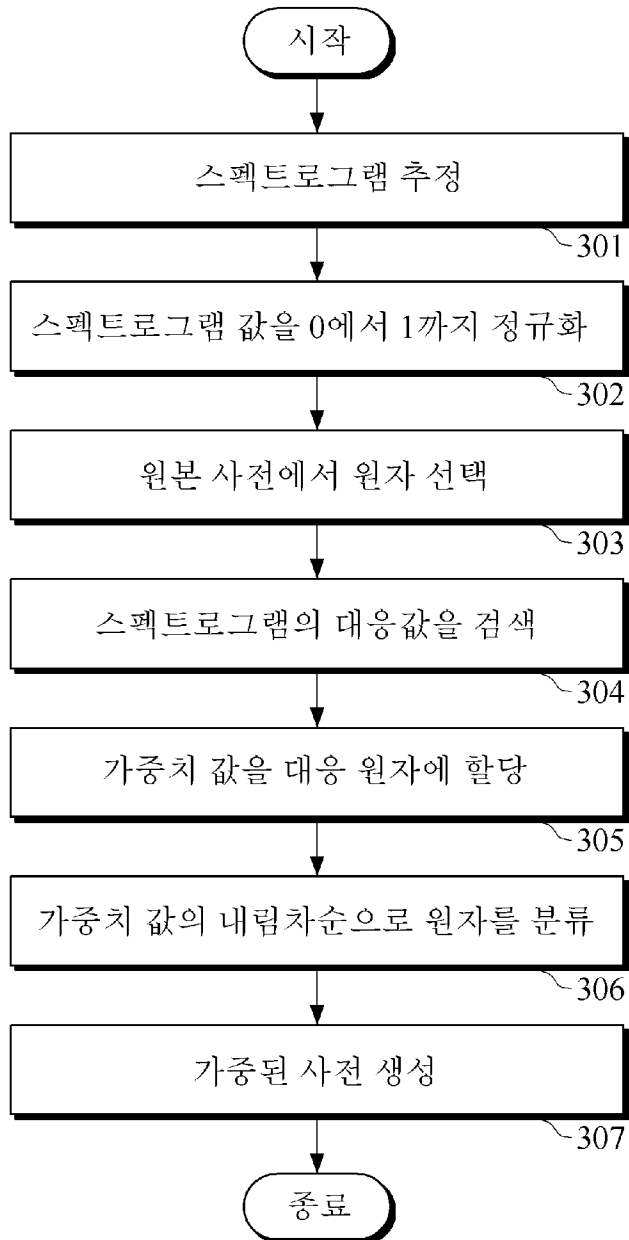
[Fig. 1]



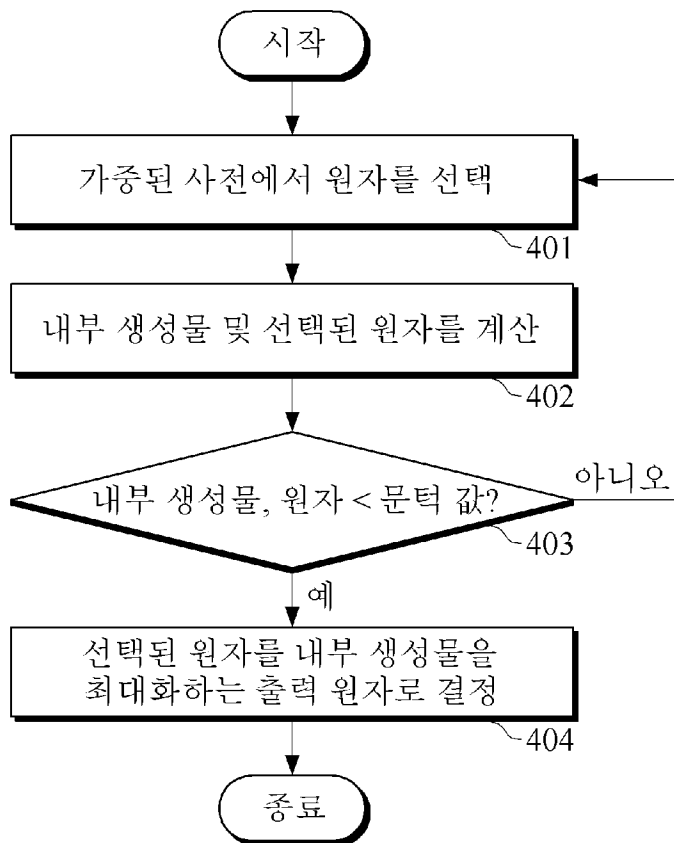
[Fig. 2]



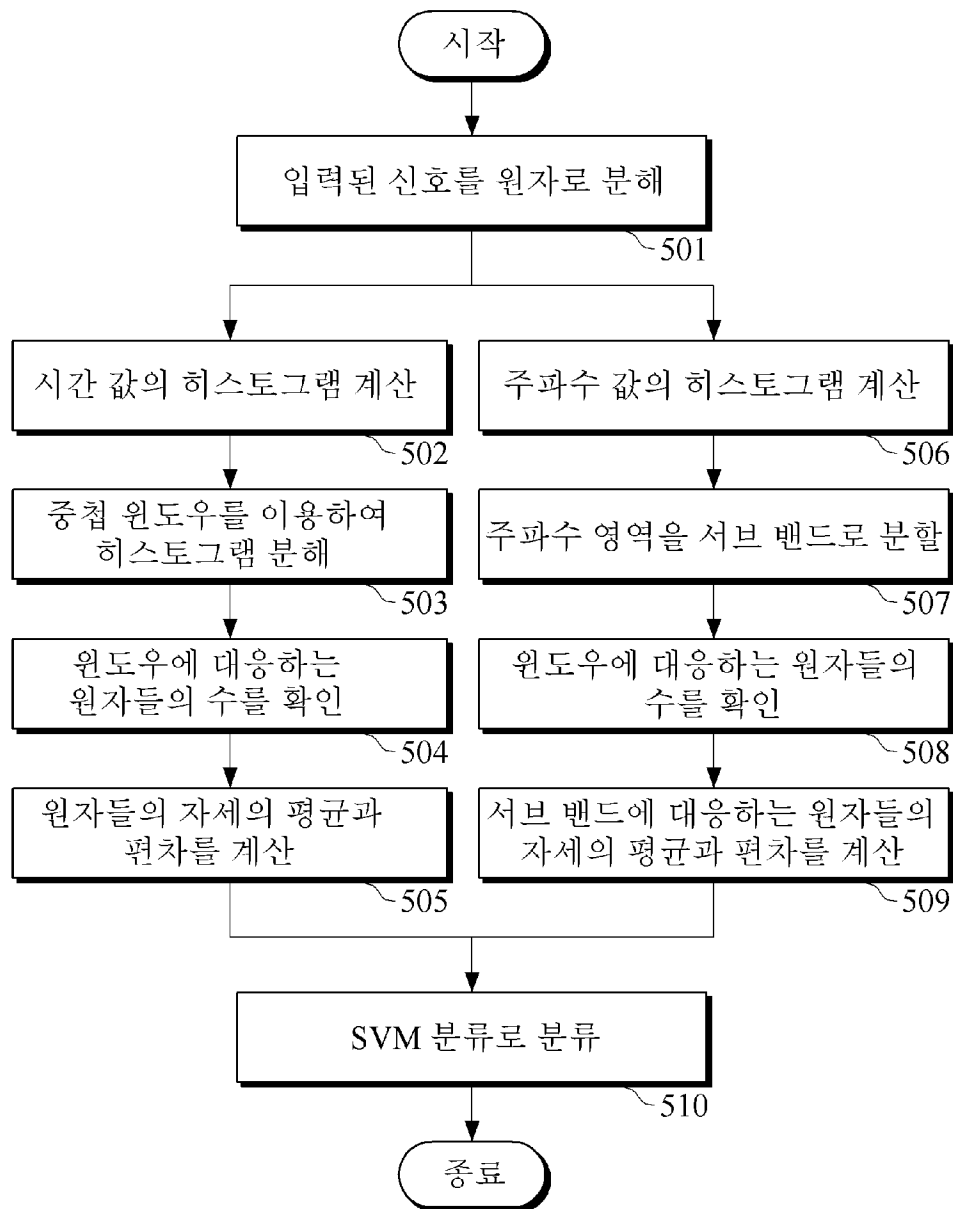
[Fig. 3]



[Fig. 4]



[Fig. 5]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2012/010785**A. CLASSIFICATION OF SUBJECT MATTER****G10L 19/00(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 19/00; G10L 25/63; G10L 15/00; G10L 15/02; G10L 25/09; G10L 15/28; 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: voice recognition, emotion, dictionary, spectrogram, residue, weight value

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KR 10-1006049 B1 (KANG, Jung Hwan) 06 January 2011 See abstract; paragraphs 52-56; and figure 2.	1-10
A	KR 10-2009-0120640 A (RESEARCH & BUSINESS FOUNDATION SUNGKYUNKWAN UNIVERSITY) 25 November 2009 See abstract; paragraphs 32-34, 51; figures 1, 2; and claim 7.	1-10
A	JP 2010-197644 A (GIFU UNIV. et al.) 09 September 2010 See abstract; paragraphs 24-27; figure 2; and claim 1.	1-10
A	KR 10-2009-0063202 A (POSTECH ACADEMY-INDUSTRY FOUNDATION) 17 June 2009 See abstract; paragraphs 24-26; figure 2; and claim 1.	1-10
A	KR 10-2008-0086791 A (LG ELECTRONICS INC.) 26 September 2008 See abstract; paragraphs 28-35; and figures 2, 3.	1-10

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:	"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
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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	

Date of the actual completion of the international search

15 MAY 2013 (15.05.2013)

Date of mailing of the international search report

16 MAY 2013 (16.05.2013)

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/KR2012/010785

Patent document cited in search report	Publication date	Patent family member	Publication date
KR 10-1006049 B1	06.01.2011	KR 10-2010-0042482 A	26.04.2010
KR 10-2009-0120640 A	25.11.2009	KR 10-0937101 B1	15.01.2010
JP 2010-197644 A	09.09.2010	NONE	
KR 10-2009-0063202 A	17.06.2009	NONE	
KR 10-2008-0086791 A	26.09.2008	NONE	

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G10L 19/00(2006.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G10L 19/00; G10L 25/63; G10L 15/00; G10L 15/02; G10L 25/09; G10L 15/28; 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: 음성인식, 감정, 사전, 스펙트로그램, 레지듀, 가중치

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
A	KR 10-1006049 B1 (강정환) 2011.01.06 요약; 단락 52-56; 및 도면 2 참조.	1-10
A	KR 10-2009-0120640 A (성균관대학교산학협력단) 2009.11.25 요약; 단락 32-34, 51; 도면 1, 2; 및 청구항 7 참조.	1-10
A	JP 2010-197644 A (GIFU UNIV. 외 1명) 2010.09.09 요약; 단락 24-27; 도면 2; 및 청구항 1 참조.	1-10
A	KR 10-2009-0063202 A (포항공과대학교 산학협력단) 2009.06.17 요약; 단락 24-26; 도면 2; 및 청구항 1 참조.	1-10
A	KR 10-2008-0086791 A (엘지전자 주식회사) 2008.09.26 요약; 단락 28-35; 및 도면 2, 3 참조.	1-10

☐ 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

2013년 05월 15일 (15.05.2013)

국제조사보고서 발송일

2013년 05월 16일 (16.05.2013)

ISA/KR의 명칭 및 우편주소

 대한민국 특허청
(302-701) 대전광역시 서구 청사로 189,
4동 (둔산동, 정부대전청사)
팩스 번호 82-42-472-7140

심사관

김도원

전화번호 82-42-481-5560



국제조사보고서에서
인용된 특허문헌

공개일

대응특허문헌

공개일

KR 10-1006049 B1

2011.01.06

KR 10-2010-0042482 A

2010.04.26

KR 10-2009-0120640 A

2009.11.25

KR 10-0937101 B1

2010.01.15

JP 2010-197644 A

2010.09.09

없음

KR 10-2009-0063202 A

2009.06.17

없음

KR 10-2008-0086791 A

2008.09.26

없음