



**(19) 대한민국특허청(KR)**  
**(12) 등록특허공보(B1)**

(45) 공고일자 2008년05월27일  
(11) 등록번호 10-0832534  
(24) 등록일자 2008년05월20일

(51) Int. Cl.

G10L 15/00 (2006.01) G10L 15/02 (2006.01)

G10L 13/02 (2006.01) G10L 15/22 (2006.01)

(21) 출원번호 10-2006-0095049

(22) 출원일자 2006년09월28일

심사청구일자 2006년09월28일

(65) 공개번호 10-2008-0029238

(43) 공개일자 2008년04월03일

(56) 선행기술조사문헌

10-2005-63886\*

10-2006-16987

JP2002156991 A\*

\*는 심사관에 의하여 인용된 문헌

(73) 특허권자

한국전자통신연구원

대전 유성구 가정동 161번지

(72) 발명자

김복원

대전 중구 목동 목양마을아파트 114동 204호

이강우

대전 유성구 어은동 한빛아파트 136-1506호

(뒷면에 계속)

(74) 대리인

특허법인 씨엔에스·로고스

전체 청구항 수 : 총 15 항

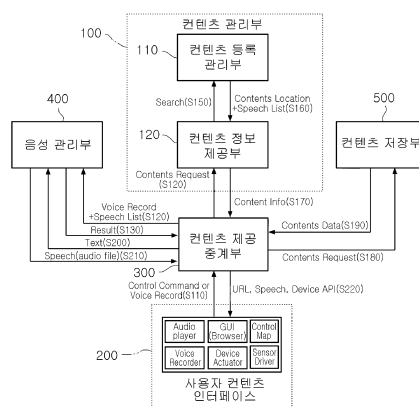
심사관 : 경연정

**(54) 음성 인식을 통한 콘텐츠 정보 서비스 제공 장치 및 그방법**

**(57) 요약**

본 발명은 음성 인식을 통한 콘텐츠 정보 서비스 제공 장치 및 그 방법에 관한 것으로, 본 발명의 콘텐츠 정보 서비스 제공 장치는, 음성 인식을 포함하는 복수의 사용자 입출력 인터페이스를 통해 수집되는 해당 콘텐츠의 제공 요청정보를 입력 받고, 상기 입력된 요청정보에 대응하여 입력되는 상기 해당 콘텐츠를 상기 사용자 입출력 인터페이스를 통해 사용자에게 제공하는 사용자 콘텐츠 인터페이스; 상기 사용자 콘텐츠 인터페이스에 입력된 콘텐츠 제공 요청정보에 대응하는 콘텐츠 관련 정보를 요청하고, 이에 대응하여 제공되는 콘텐츠 관련 정보에 대응하는 콘텐츠를 요청하며, 이에 대응하여 제공되는 콘텐츠를 상기 사용자 콘텐츠 인터페이스에 제공하는 콘텐츠 제공 중계부; 제공 가능한 상기 콘텐츠에 대한 콘텐츠 관련 정보를 등록 및 관리하고, 상기 콘텐츠 제공 중계부의 콘텐츠 관련 정보 요청에 대응하여 대응하는 콘텐츠 관련정보를 제공하는 콘텐츠 정보 관리부; 및 상기 콘텐츠 관련 정보에 대응하여 제공 가능한 복수의 콘텐츠들을 저장 및 관리하고, 상기 콘텐츠 제공 중계부가 요청한 상기 콘텐츠 관련 정보에 대응하는 콘텐츠를 검출하여 제공하는 콘텐츠 저장부를 포함한다.

**대표도 - 도1**



(72) 발명자

**서영호**

광주 광산구 운남동 주공아파트 508동 403호

**김민영**

충북 충주시 이류면 매현리 481

**김연준**

대전 서구 탄방동 641번지 메이저타운 907호

**김현**

대전 유성구 신성동 삼성한울아파트 103-702호

**조영조**

경기 성남시 분당구 이매동 아름마을건영아파트

104동 201호

## 특허청구의 범위

### 청구항 1

음성 인식을 포함하는 복수의 사용자 입력력 인터페이스를 통해 수집되는 해당 콘텐츠의 제공 요청정보를 입력 받고, 상기 입력된 요청정보에 대응하여 입력되는 상기 해당 콘텐츠를 상기 사용자 입력력 인터페이스를 통해 사용자에게 제공하는 사용자 콘텐츠 인터페이스;

상기 사용자 콘텐츠 인터페이스에 입력된 콘텐츠 제공 요청정보에 대응하는 콘텐츠 관련 정보를 요청하고, 이에 대응하여 제공되는 콘텐츠 관련 정보에 대응하는 콘텐츠를 요청하며, 이에 대응하여 제공되는 콘텐츠를 상기 사용자 콘텐츠 인터페이스에 제공하는 콘텐츠 제공 중계부;

제공 가능한 상기 콘텐츠에 대한 콘텐츠 관련 정보를 등록 및 관리하고, 상기 콘텐츠 제공 중계부의 콘텐츠 관련 정보 요청에 대응하여 대응하는 콘텐츠 관련정보를 제공하는 콘텐츠 정보 관리부;

상기 콘텐츠 관련 정보에 대응하여 제공 가능한 복수의 콘텐츠들을 저장 및 관리하고, 상기 콘텐츠 제공 중계부가 요청한 상기 콘텐츠 관련 정보에 대응하는 콘텐츠를 검출하여 제공하는 콘텐츠 저장부; 및

상기 사용자 콘텐츠 인터페이스로부터 상기 콘텐츠 제공 중계부에 입력되는 상기 콘텐츠의 제공 요청정보가 유효한 정보인지를 판별하고, 상기 콘텐츠 저장부로부터 상기 콘텐츠 제공 중계부에 입력되는 상기 콘텐츠의 텍스트 데이터를 음성 데이터로 변환하며, 상기 판별 및 변환한 결과 정보를 상기 콘텐츠 제공 중계부에 제공하는 음성 관리부를 포함하는 것을 특징으로 하는 콘텐츠 정보 서비스 제공 장치.

### 청구항 2

제 1항에 있어서,

상기 사용자 콘텐츠 인터페이스는,

상기 해당 콘텐츠의 제공 요청정보에 대응하는 명령을 생성하여 상기 콘텐츠 제공 중계부에 제공하는 콘텐츠제공 요청명령 생성수단; 및

상기 콘텐츠제공 요청명령 생성수단에 의해 제공된 상기 명령에 대응하여 상기 콘텐츠 제공 중계부로부터 제공되는 상기 콘텐츠를 상기 사용자에게 제공하는 콘텐츠 출력수단을 포함하는 것을 특징으로 하는 콘텐츠 정보 서비스 제공 장치.

### 청구항 3

제 2항에 있어서,

상기 콘텐츠제공 요청명령 생성수단은, 음성 녹음 수단, 'GUI'수단, 및 'Sensor Driver' 수단 중 적어도 어느 하나를 통해 상기 해당 콘텐츠의 제공 요청정보를 수집하는 것을 특징으로 하는 콘텐츠 정보 서비스 제공 장치.

### 청구항 4

제 3항에 있어서,

상기 음성 녹음 수단은 상기 해당 콘텐츠의 제공 요청정보를 음성 녹음하여 상기 콘텐츠제공 요청명령을 생성하는 것을 특징으로 하는 콘텐츠 정보 서비스 제공 장치.

### 청구항 5

제 3항에 있어서,

상기 'GUI'수단 및 'Sensor Driver' 수단은 상기 해당 콘텐츠의 제공 요청정보를 미리 정의된 제어 지도(control map) 테이블에 기초하여 음성 명령 심벌로 맵핑하는 것을 특징으로 하는 콘텐츠 정보 서비스 제공 장치.

### 청구항 6

제 2항에 있어서,

상기 콘텐츠 출력수단은,

상기 콘텐츠 제공 중계부로부터 제공된 콘텐츠를 오디오 데이터로 출력하는 오디오 출력 수단;

상기 콘텐츠를 GUI 데이터로 화면에 표시하는 GUI 표시 수단; 및

상기 콘텐츠에 대응하여 구비된 임의의 하드웨어장치를 구동시키는 장치 액추에이터(Device Actuator) 중 적어도 하나를 포함하는 것을 특징으로 하는 콘텐츠 정보 서비스 제공 장치.

#### 청구항 7

제 6항에 있어서,

상기 GUI 데이터는 이미지 파일 및 HTML문서 중 어느 하나의 형식을 갖는 것을 특징으로 하는 콘텐츠 정보 서비스 제공 장치.

#### 청구항 8

제 1항에 있어서,

상기 콘텐츠 정보 관리부는,

상기 콘텐츠 저장부를 통해 제공 가능한 콘텐츠에 대한 콘텐츠 관련 정보를 등록 및 관리하는 콘텐츠 등록 관리부; 및

상기 콘텐츠 제공 중계부로부터 콘텐츠 관련 정보 요청이 있으면, 상기 콘텐츠 등록 관리부로부터 상기 요청한 콘텐츠에 대응하는 콘텐츠 관련정보를 검출하여 상기 콘텐츠 제공 중계부에 제공하는 콘텐츠 정보 제공부를 포함하는 것을 특징으로 하는 콘텐츠 정보 서비스 제공 장치.

#### 청구항 9

삭제

#### 청구항 10

복수의 사용자 입출력 인터페이스를 통해 해당 콘텐츠의 제공 요청정보를 입력 받는 단계;

상기 입력된 콘텐츠 제공 요청정보에 대응하는 콘텐츠 관련 정보를 요청하는 단계;

상기 입력된 콘텐츠 제공 요청정보가 유효한 정보인지를 판별하는 단계;

상기 콘텐츠 관련 정보 요청에 대응하는 콘텐츠 관련정보를 제공하는 단계;

상기 제공된 콘텐츠 관련정보에 대응하는 콘텐츠를 요청하는 단계;

상기 요청한 콘텐츠 관련 정보에 대응하는 콘텐츠를 검출하는 단계; 및

상기 검출된 콘텐츠를 사용자에게 출력하는 단계를 포함하는 것을 특징으로 하는 콘텐츠 정보 서비스 제공 방법.

#### 청구항 11

제 10항에 있어서,

상기 해당 콘텐츠의 제공 요청정보를 입력 받는 단계에서는, 음성 녹음, 'GUI', 및 'Sensor' 중 적어도 어느 하나를 통해 상기 해당 콘텐츠의 제공 요청정보를 수집하는 것을 특징으로 하는 콘텐츠 정보 서비스 제공 방법.

#### 청구항 12

제 11항에 있어서,

상기 콘텐츠 관련 정보를 요청 단계에서는, 상기 해당 콘텐츠의 제공 요청정보를 상기 음성 녹음을 통해 상기 콘텐츠제공 요청명령을 생성하는 것을 특징으로 하는 콘텐츠 정보 서비스 제공 방법.

#### 청구항 13

제 11항에 있어서,

상기 콘텐츠 관련 정보를 요청 단계에서는, 상기 'GUI' 및 'Sensor'를 통해 수집된 상기 해당 콘텐츠의 제공 요청정보를 미리 정의된 제어 지도(control map) 테이블에 기초하여 음성 명령 심벌로 맵핑하는 것을 특징으로 하는 콘텐츠 정보 서비스 제공 방법.

**청구항 14**

제 10항에 있어서,

상기 제공되는 콘텐츠를 사용자에게 출력하는 단계에서는,

상기 제공된 콘텐츠를 오디오 데이터로 출력, 상기 제공된 콘텐츠를 GUI 데이터로 화면에 표시, 및 상기 제공된 콘텐츠에 대응하여 구비된 임의의 하드웨어장치를 구동시키는 것 중 적어도 어느 하나를 통해 상기 콘텐츠를 출력하는 것을 특징으로 하는 콘텐츠 정보 서비스 제공 방법.

**청구항 15**

제 10항에 있어서,

상기 콘텐츠 관련정보 요청단계에서는 상기 유효한 정보에 대한 상기 콘텐츠 관련정보를 요청하는 것을 특징으로 하는 콘텐츠 정보 서비스 제공 방법.

**청구항 16**

제 10항 또는 제 15항에 있어서,

상기 콘텐츠 검출 단계 후, 상기 검출된 콘텐츠의 텍스트 데이터를 음성 데이터로 변환하는 단계를 더 포함하며,

상기 콘텐츠 출력 단계에서는 상기 변환된 음성 데이터로 상기 콘텐츠를 상기 사용자에게 출력하는 것을 특징으로 하는 콘텐츠 정보 서비스 제공 방법.

**명세서**

**발명의 상세한 설명**

**발명의 목적**

**발명이 속하는 기술 및 그 분야의 종래기술**

- <11> 본 발명은 인터넷 상의 콘텐츠 서비스 제공 장치 및 그 방법에 관한 것으로서, 보다 상세하게는, 요청된 정보에 대해 음성 인식 인터페이스를 이용하여 요청된 정보에 대응하는 해당 콘텐츠 정보를 인터넷상에서 검색하여 사용자 인터페이스를 통해 사용자에게 제공할 수 있는 콘텐츠 서비스 제공 장치 및 그 방법에 관한 것이다.
- <12> 현재 인터넷상에 존재하는 수많은 정보에 대해, 사용자들은 인터넷 통신이 가능한 개인 컴퓨터나 유무선 통신 단말 등을 이용하여 해당하는 정보에 접근할 수 있다. 인터넷을 통해 접근한 정보들이 해당 단말의 화면을 통해 표시되면, 사용자는 그 정보를 직접 읽고 판단하여 단순한 마우스 클릭이나 키 조작으로 화면 상에서 다시 새로운 명령을 입력한다. 이러한 방법은 단순히 해당 단말의 화면 표시 장치에 의존하여 해당 정보를 검색하는 방법이다.
- <13> 이러한 방법 이외에 음성을 이용하여 해당 정보를 획득할 수 있는 할 수 있는 방법이 있다. 종래의 음성을 이용한 정보 검색 방법은 'Voice XML(eXtensible Markup Language)'을 활용하여 사용자가 직접 음성 인식 리스트를 정의하고 각각의 음성 명령에 대응되는 동작을 수행할 수 있다.
- <14> 다른 방법으로는, 'HTML'을 'Voice XML'로 자동으로 변환되게 하여 사용하는 경우도 있는데, 이러한 경우에는 'HTML'에 포함된 제어(Control) 부분을 음성 인식 리스트로 추출한다. 즉, 이 방법은 'HTML'에 포함되어 있는 하이퍼링크(Hyper Link)에 대해서 심벌을 추출하고, 추출된 하이퍼링크 심벌을 음성 인식 리스트로 활용을 한다. 이와 비슷하게 일반적인 어플리케이션의 GUI(Graphic User Interface) 컴포넌트의 정보를 해석하여 제어 정보를 추출한 후, 제어 정보에 해당하는 음성인식 리스트를 작성하는 방법도 있다. 여기서 제어 정보는

'HTML'의 하이퍼링크와 같은 용도로 이용된다.

- <15> 이와 같이, 현재까지 콘텐츠 제공자들로부터 자신이 원하는 콘텐츠 서비스를 제공받고자 하는 경우, 사용자들은 컴퓨터나 통신 단말 등을 이용하여 화면을 통해서 표시되는 해당 정보들을 확인한 후 특정 정보의 표시 명령을 입력하여 자신이 원하는 정보를 확인하는 방법을 이용한다.
- <16> 그런데, 최근 들어 디지털 홈 네트워크 환경은 지능형 로봇뿐만 아니라 가전 기기들에 대해서도 확대 적용되고 있는 실정이다. 홈 네트워크 연동 가능한 기기들은 사용자와 상호작용하기 위해 음성 인터페이스 또는 지능형 리모콘 등을 이용한다. 사용자가 로봇이 제공하는 음성 인터페이스를 사용하여 원하는 정보를 제공받고 명령을 내린다고 가정할 때, 기존의 음성 명령을 통한 정보제공 방법은 매우 불편할 수 있다. 왜냐하면 기존의 음성 인터페이스 방법은 화면을 기반으로 하는 정보 제공 상황에서 단순히 음성을 통한 명령이 가능하도록 화면 변화에 대한 음성인식 리스트만을 업데이트하기 때문이다. 즉, 기존의 방법의 경우 'GUI'의 보조적인 역할로 음성을 인터페이스를 활용하기 때문에, 사용자와 정보제공 서비스간의 다양하고 즉흥적인 상호 작용을 할 수 없는 문제점이 있다.
- <17> 예를 들어, 사용자가 요리 정보를 요청하고 홈 네트워크 환경의 로봇이 요리에 대한 정보 서비스를 제공한다고 하였을 때, 사용자는 요리 순서가 음성으로 전달되는 중에 요리 재료에 대한 질문이 필요한 경우가 발생할 수도 있다. 이에 따라, 사용자는 이전 단계를 다시 반복하라는 명령 등을 제공할 수도 있으며, 심지어 전혀 엉뚱한 날씨 정보의 제공을 요청할 수도 있다.

**발명이 이루고자 하는 기술적 과제**

- <18> 상기와 같은 문제점을 해결하기 위한 본 발명의 제1 목적은, 음성 인식을 통한 콘텐츠 정보 서비스 제공 장치 및 그 방법을 제공하는 데 있다.
- <19> 본 발명의 제2 목적은, 음성 인식 기능이 포함된 인식 가능한 복수의 사용자 입출력 인터페이스를 통해 해당 콘텐츠 제공 요청을 인식하고 이에 대응하는 콘텐츠 정보를 상호 작용을 하면서 사용자에게 제공할 수 있는 음성 인식을 통한 콘텐츠 정보 서비스 제공 장치 및 그 방법을 제공하는 데 있다.

**발명의 구성 및 작용**

- <20> 상기와 같은 목적을 달성하기 위한 본 발명의 실시예에 따른 음성 인식을 통한 콘텐츠 정보 서비스 제공 장치는, 음성 인식을 포함하는 복수의 사용자 입출력 인터페이스를 통해 수집되는 해당 콘텐츠의 제공 요청정보를 입력 받고, 상기 입력된 요청정보에 대응하여 입력되는 상기 해당 콘텐츠를 상기 사용자 입출력 인터페이스를 통해 사용자에게 제공하는 사용자 콘텐츠 인터페이스; 상기 사용자 콘텐츠 인터페이스에 입력된 콘텐츠 제공 요청정보에 대응하는 콘텐츠 관련 정보를 요청하고, 이에 대응하여 제공되는 콘텐츠 관련 정보에 대응하는 콘텐츠를 요청하며, 이에 대응하여 제공되는 콘텐츠를 상기 사용자 콘텐츠 인터페이스에 제공하는 콘텐츠 제공 중계부; 제공 가능한 상기 콘텐츠에 대한 콘텐츠 관련 정보를 등록 및 관리하고, 상기 콘텐츠 제공 중계부의 콘텐츠 관련 정보 요청에 대응하여 대응하는 콘텐츠 관련정보를 제공하는 콘텐츠 정보 관리부; 및 상기 콘텐츠 관련 정보에 대응하여 제공 가능한 복수의 콘텐츠들을 저장 및 관리하고, 상기 콘텐츠 제공 중계부가 요청한 상기 콘텐츠 관련 정보에 대응하는 콘텐츠를 검출하여 제공하는 콘텐츠 저장부를 포함한다.
- <21> 바람직하게는, 상기 사용자 콘텐츠 인터페이스는, 상기 해당 콘텐츠의 제공 요청정보에 대응하는 명령을 생성하여 상기 콘텐츠 제공 중계부에 제공하는 콘텐츠제공 요청명령 생성수단; 및 상기 콘텐츠제공 요청명령 생성수단에 의해 제공된 상기 명령에 대응하여 상기 콘텐츠 제공 중계부로부터 제공되는 상기 콘텐츠를 상기 사용자에게 제공하는 콘텐츠 출력수단을 포함한다.
- <22> 본 실시예에서 상기 콘텐츠제공 요청명령 생성수단은, 음성 녹음 수단, 'GUI'수단, 및 'Sensor Driver' 수단 중 적어도 어느 하나를 통해 상기 해당 콘텐츠의 제공 요청정보를 수집한다.
- <23> 상기 음성 녹음 수단은 상기 해당 콘텐츠의 제공 요청정보를 음성 녹음하여 상기 콘텐츠제공 요청명령을 생성한다.
- <24> 상기 'GUI'수단 및 'Sensor Driver' 수단은 상기 해당 콘텐츠의 제공 요청정보를 미리 정의된 제어 지도(control map) 테이블에 기초하여 음성 명령 심벌로 맵핑한다.
- <25> 본 실시예에서 상기 콘텐츠 출력수단은, 상기 콘텐츠 제공 중계부로부터 제공된 콘텐츠를 오디오 데이터로 출력하는 오디오 출력 수단; 상기 콘텐츠를 GUI 데이터로 화면에 표시하는 GUI 표시 수단; 및 상기 콘텐츠에 대응하

여 구비된 임의의 하드웨어장치를 구동시키는 장치 액추에이터(Device Actuator)를 포함한다.

- <26> 상기 GUI 데이터는 이미지 파일 또는 HTML문서 형식을 갖는다.
- <27> 상기 콘텐츠 정보 관리부는, 상기 콘텐츠 저장부를 통해 제공 가능한 콘텐츠에 대한 콘텐츠 관련 정보를 등록 및 관리하는 콘텐츠 등록 관리부; 및 상기 콘텐츠 제공 중계부로부터 콘텐츠 관련 정보 요청이 있으면, 상기 콘텐츠 등록 관리부로부터 상기 요청한 콘텐츠에 대응하는 콘텐츠 관련정보를 검출하여 상기 콘텐츠 제공 중계부에 제공하는 콘텐츠 정보 제공부를 포함한다.
- <28> 본 실시예의 음성 인식을 통한 콘텐츠 정보 서비스 제공 장치는, 상기 사용자 콘텐츠 인터페이스로부터 상기 콘텐츠 제공 중계부에 입력되는 상기 해당 콘텐츠의 제공 요청정보가 유효한 정보인지를 판별하고, 상기 콘텐츠 저장부로부터 상기 콘텐츠 제공 중계부에 입력되는 상기 콘텐츠의 텍스트 데이터를 음성 데이터로 변환하며, 상기 판별 및 변환한 결과 정보를 상기 콘텐츠 제공 중계부에 제공하는 음성 관리부를 더 포함한다.
- <29> 한편, 상기와 같은 목적을 달성하기 위한 본 발명의 실시예에 따른 음성 인식을 통한 콘텐츠 정보 서비스 제공 방법은, 복수의 사용자 입출력 인터페이스를 통해 해당 콘텐츠의 제공 요청정보를 입력 받는 단계; 상기 입력된 콘텐츠 제공 요청정보에 대응하는 콘텐츠 관련 정보를 요청하는 단계; 상기 콘텐츠 관련 정보 요청에 대응하는 콘텐츠 관련정보를 제공하는 단계; 상기 제공된 콘텐츠 관련정보에 대응하는 콘텐츠를 요청하는 단계; 상기 요청한 콘텐츠 관련 정보에 대응하는 콘텐츠를 검출하는 단계; 및 상기 검출된 콘텐츠를 사용자에게 출력하는 단계를 포함한다.
- <30> 상기 콘텐츠 관련 정보 요청 단계에서는, 상기 해당 콘텐츠의 제공 요청정보를 상기 음성 녹음을 통해 상기 콘텐츠제공 요청명령을 생성한다.
- <31> 상기 콘텐츠 관련 정보 요청 단계에서는, 상기 'GUI' 및 'Sensor'를 통해 수집된 상기 해당 콘텐츠의 제공 요청정보를 미리 정의된 제어 지도(control map) 테이블에 기초하여 음성 명령 심별로 맵핑한다.
- <32> 상기 제공되는 콘텐츠를 사용자에게 출력하는 단계에서는, 상기 제공된 콘텐츠를 오디오 데이터로 출력, 상기 제공된 콘텐츠를 GUI 데이터로 화면에 표시, 및 상기 제공된 콘텐츠에 대응하여 구비된 임의의 하드웨어장치를 구동시키는 것 중 어느 하나를 통해 상기 콘텐츠를 출력한다.
- <33> 본 실시예의 콘텐츠 정보 서비스 제공 방법은, 상기 콘텐츠 관련 정보 요청 단계 후, 상기 입력되는 해당 콘텐츠의 제공 요청정보가 유효한 정보인지를 판별하는 단계를 더 포함한다. 이에 의해, 상기 콘텐츠 관련정보 요청단계에서는 상기 유효한 정보에 대한 상기 콘텐츠 관련정보를 요청한다.
- <34> 본 실시예의 콘텐츠 정보 서비스 제공 방법은, 상기 콘텐츠 검출 단계 후, 상기 검출된 콘텐츠의 텍스트 데이터를 음성 데이터로 변환하는 단계를 더 포함한다. 이에 의해, 상기 콘텐츠 출력 단계에서는 상기 변환된 음성 데이터로 상기 콘텐츠를 상기 사용자에게 출력한다.
- <35> 본 발명에 따르면, 음성 인식 및/또는 입출력 가능한 사용자 인터페이스를 통해 사용자로부터 입력되는 음성 또는 문자 형태의 해당 콘텐츠 제공 요청정보를 인식하고 이에 대응하는 콘텐츠 정보를 검색하여 사용자에게 제공 가능한 인터페이스를 통해 제공함으로써, 디지털 홈 또는 차량에 설치되는 다양한 통신 기기들이 표시장치의 구비 유무에 상관없이 사용자에게 입출 가능한 사용자 인터페이스를 통해 보다 용이하게 해당 콘텐츠 정보 서비스를 제공할 수 있다.
- <36> 이하, 본 발명의 바람직한 실시예들을 첨부한 도면을 참조하여 상세히 설명한다. 도면들 중 동일한 구성요소들은 가능한 한 어느 곳에서든지 동일한 부호들로 나타내고 있음에 유의해야 한다. 또한 본 발명의 요지를 불필요하게 흐릴 수 있는 공지 기능 및 구성에 대한 상세한 설명은 생략한다.
- <37> 도 1은 본 발명의 바람직한 일 실시예에 따른 음성 인식을 통한 콘텐츠 서비스 제공 장치를 도시한 블록도이다.
- <38> 도시된 바와 같이, 음성 인식을 통한 콘텐츠 서비스 제공 장치는, 콘텐츠 관리부(100), 사용자 콘텐츠 인터페이스(200), 콘텐츠 제공 중계부(300), 음성 관리부(400), 및 콘텐츠 저장부(500)를 포함한다.
- <39> 콘텐츠 관리부(100)는 제공 가능한 콘텐츠에 대한 정보를 관리 및 제공한다. 이를 위해, 콘텐츠 관리부(100)는 콘텐츠 등록 관리부(110) 및 콘텐츠 정보 제공부(120)를 포함한다.
- <40> 콘텐츠 등록 관리부(110)는 콘텐츠 정보에 접근하는데 필요한 제어 정보 및 콘텐츠의 주소정보를 저장 및 관리



한다. 이를 위해, 콘텐츠 등록 관리부(110)는 콘텐츠 아이템 테이블(Content Item Table)(도 2), 콘텐츠 아이템 제어 테이블(Content Item Control Table)(도 3), 및 콘텐츠 데이터 테이블(Content Data Table)(도 4)을 저장 및 관리하고, 이들 각각에 대한 검색 인터페이스를 제공한다. 즉, 콘텐츠 등록 관리부(110)는 콘텐츠 정보의 검색 명령(Search)이 입력되면, 이에 대응하는 콘텐츠 제어정보(Contents Location + Speech List)를 검색하여 콘텐츠 정보 제공부(120)에게 제공한다.

- <41> 콘텐츠 정보 제공부(120)는 콘텐츠 제공 중계부(300)로부터 콘텐츠 요청정보(Contents info Request)가 입력되면, 콘텐츠 요청정보에 대응하여 저장된 콘텐츠의 제어 정보(Contents Location + Speech List)를 콘텐츠 등록 관리부(110)로부터 검색한다. 콘텐츠 정보 제공부(120)는 검색한 콘텐츠 제어정보(Contents Location + Speech List)에 대응하는 콘텐츠 정보(Content Info)를 생성하여 콘텐츠 제공 중계부(300)로 전송한다.
- <42> 사용자 콘텐츠 인터페이스(200)는 사용자로부터 해당 콘텐츠의 제공을 요청하는 명령(Control Command or Voice Record)을 입력 받아 콘텐츠 제공 중계부(300)에게 제공한다. 또한 사용자 콘텐츠 인터페이스(200)는 콘텐츠 제공 중계부(300)로부터 상기 요청 명령에 대응하여 제공되는 콘텐츠(URL, Speech, Device API)를 사용자 인터페이스를 통해 콘텐츠 요청자에게 제공한다.
- <43> 사용자 콘텐츠 인터페이스(200)는 본 실시예에 따라 핸드폰, 와이브로 단말, 로봇, 디지털 홈 디바이스등과 같은 네트워크로 연결가능 한 모든 장치에 탑재될 수 있다. 사용자 콘텐츠 인터페이스(200)는 사용자에게 콘텐츠를 음성, API(Application Program Interface) 및 표시 장치가 존재하는 경우 'GUI'를 통해서 콘텐츠를 사용자에게 제공할 수 있다. 사용자 콘텐츠 인터페이스(200)는 사용자로부터 입력된 명령을 콘텐츠 제공 중계부(300)에 전송하기 위해서 사용자의 목소리를 녹음한 음성 녹음(Voice Record)한다. 또는 사용자 콘텐츠 인터페이스(200)는 입력되는 명령에 대해, 'GUI'나 'Device 장치'조작을 통해 제어 명령(Control Command)을 생성한다.
- <44> 콘텐츠 제공 중계부(300)는 사용자 콘텐츠 인터페이스(200)로부터 전달된 제어 명령 또는 음성 녹음 정보(Voice Record)를 입력 받는다. 제어 명령을 입력 받은 경우, 콘텐츠 제공 중계부(300)는 콘텐츠 관리부(100)의 콘텐츠 정보 제공부(120)에게 콘텐츠 정보 요구정보(Content info Request)를 전달한다. 음성 녹음 정보(Voice Record)를 입력 받은 경우, 콘텐츠 제공 중계부(300)는 음성 녹음과 클라이언트 세션 테이블(Client Session Table)에 저장되어 있는 유효한 음성 명령리스트(Voice Record + Speech List)를 음성 관리부(400)에 전달하여 그 결과(result)를 제공 받는다. 여기서 음성 관리부(400)는 입력되는 음성 녹음 정보 및 유효 음성 명령리스트에 대해, 음성 인식 기법(Automatic Speech Recognition: ASR)을 이용하여 인식 가능한 정보로 인식한다.
- <45> 만약 음성 인식 결과가 유효한 정보인 경우, 콘텐츠 제공 중계부(300)는 콘텐츠 정보 제공부(120)에게 콘텐츠 정보 요구정보(Content info Request)를 생성하여 전달한다. 반면 음성 인식 결과가 유효하지 않은 정보인 경우, 콘텐츠 제공 중계부(300)는 이를 무시하거나 사용자 콘텐츠 인터페이스(200)에게 에러 이벤트를 전달한다.
- <46> 콘텐츠 제공 중계부(300)는 콘텐츠 정보 제공부(120)로부터 콘텐츠 정보 요구정보(Content info Request)대환 응답으로 콘텐츠 정보(Content Info)를 제공 받는다. 이에 따라, 콘텐츠 제공 중계부(300)는 콘텐츠 저장부(500)에 콘텐츠 제공을 요청(Contents Request)하여, 저장된 콘텐츠(Contents Data)를 제공 받는다. 이때 가져온 콘텐츠가 텍스트(Text) 정보인 경우, 콘텐츠 제공 중계부(300)는 음성 관리부(400)의 TTS(Text to Speech) 기능을 이용하여 음성파일(Speech(audio file))로 변환한 후, 'GUI'와 'Device API' 정보와 함께 사용자 콘텐츠 인터페이스(200)에 제공한다. 이에 따라, 사용자 콘텐츠 인터페이스(200)는 콘텐츠 제공 중계부(300)로부터 제공 받은 콘텐츠(URL, Speech, Device API)를 사용자 인터페이스를 통해 사용자에게 출력 제공한다.
- <47> 도 1에서 본 발명의 실시예에 따라 콘텐츠를 요청하고 이에 대응하여 콘텐츠를 제공하는 과정을 단계별로 신호의 흐름에 따라 순차적으로 표시(S110 내지 S220)하고 있다. 이러한 과정은 각 블록의 기술적 특징과 함께 개시하고 있으므로, 이에 대한 중복 설명은 생략한다.
- <48> 도 2는 본 발명의 실시예에 따른 콘텐츠 아이템 테이블(Content Item Table)의 스키마 구조를 나타낸 도면이다.
- <49> 도시된 바와 같이, 콘텐츠 아이템 테이블(1100)은 콘텐츠의 구성 단위들에 대한 정보를 저장한다. 이 정보들은 콘텐츠의 개념 체계를 표현하는 카테고리(Category)(1110), 콘텐츠 아이템의 식별정보(ID)(1120), 같은 콘텐츠 아이템 식별정보에 포함된 구성 요소들을 구별하는 SEQ(1130), 현재 콘텐츠 아이템이 클라이언트에게 전달되고 있을 때 사용자로부터 입력받을 수 있는 음성 명령 인식 리스트(Speech List)(1140), 실제 콘텐츠가 저장된 콘



텐츠 데이터 테이블의 레코드 중 하나를 참조하는 콘텐츠 링크(Content Link)(1150) 정보로 구성된다.

- <50> 여기서 음성 명령 인식 리스트(1140)에 저장되는 것은 제어 심볼 리스트(Control Symbol List)(1160) 정보이다. 또한 콘텐츠 아이템 테이블(1100)에 저장된 콘텐츠 아이템 중에 같은 식별정보(1120)를 가지는 아이템은 같은 콘텐츠에 속한다.
- <51> 도 3은 본 발명의 실시예에 따른 콘텐츠 아이템 제어 테이블(Content Item Control Table)의 스키마 구조를 나타낸 도면이다.
- <52> 도시된 바와 같이, 콘텐츠 아이템 제어 테이블(1200)은 도 2의 콘텐츠 이이템 테이블(1100)의 속성인 음성 명령 인식리스트(1140)에 정의된 제어 심볼 리스트(Control Symbol List)(1160)에 표현된 여러 제어에 대한 액션 정보를 저장한다.
- <53> 도시된 콘텐츠 아이템 제어 테이블(1200)은 현재 콘텐츠 식별정보(Current Content Id)(1210), 현재 콘텐츠(Current Seq)(1220), 제어 심볼(Control Symbol)(1230), 액션 콘텐츠 식별정보(Action Content Id)(1240), 및 현재 액션 콘텐츠 정보(Action Content Seq)(1250)를 포함한다.
- <54> 현재 콘텐츠 식별정보(Current Content Id)(1210) 및 현재 콘텐츠(Current Seq)(1220)는 콘텐츠 아이템 테이블(1100)의 레코드를 참조하는 키로 사용된다.
- <55> 제어 심볼(Control Symbol)(1230)는 사용자에게 의해서 선택될 수 있는 특정 명령에 대한 심볼을 의미한다. 이 심볼은 콘텐츠 아이템 테이블(1100)의 음식 명령인식 리스트(1140)의 제어 심볼 리스트(1160) 내에 반드시 미리 정의되어 있어야 한다.
- <56> 액션 콘텐츠 식별정보(Action Content Id)(1240) 및 현재 액션 콘텐츠 정보(Action Content Seq)(1250)는 클라이언트로부터 입력된 제어 심볼(1230)에 대해서 반응하는 콘텐츠를 의미한다. 액션 콘텐츠 식별정보(1240)는 콘텐츠 아이템 식별정보(1120)를 참조하고, 현재 액션 콘텐츠 정보(1250)는 'SEQ'(1130)를 참조한다.
- <57> 도 4는 본 발명의 실시예에 따른 콘텐츠 데이터 테이블(Content Data Table)의 스키마 구조를 나타낸 도면이다.
- <58> 도시된 바와 같이, 콘텐츠 데이터 테이블(Content Data Table)(1300)은 실제 콘텐츠가 존재하는 물리적인 서버로부터, 콘텐츠를 가져 올 수 있는 콘텐츠에 대한 접근 경로를 명시한다.
- <59> 이에 따라, 콘텐츠는 텍스트(Text)(1310), 'GUI'(1320), 및 '장치(Device) API'(1330)로 구성될 수 있다.
- <60> 텍스트(Text)(1310) 속성은 콘텐츠 아이템 테이블(1100)에 저장된 하나의 레코드에 대한 설명이다. 이것은 콘텐츠 제공 중계부(300)에서 음성 관리부(400)를 통해 음성 파일로 변환되어, 사용자 콘텐츠 인터페이스(200)에서 사용자에게 음성으로 전달된다. 여기서, 텍스트(Text)(1310) 속성에는 텍스트 값이 아닌 텍스트 값을 얻을 수 있는 주소값(URL)이 저장될 수도 있다.
- <61> 'GUI'(1320) 속성은 사용자 콘텐츠 인터페이스(200)의 표시 장치에서 표시되는 이미지 정보에 대한 링크를 표현한다.
- <62> '장치(Device) API'(1330) 속성은 콘텐츠가 전달되는 클라이언트 디바이스가 스피커와 표시 장치 외에 로봇과 같이 제스처나 동작 등을 할 수 있는 경우, 콘텐츠를 실행하면서 동시에 실행되는 클라이언트의 행위를 지정한다.
- <63> 도 5는 본 발명의 실시예에 따른 콘텐츠 정보 요구(Content info Request) 정보의 데이터 구조 예를 도시한 도면이다.
- <64> 도시된 바와 같이, 콘텐츠 정보 요구(Content info Request) 정보(1400)는 도 1의 콘텐츠 제공 중계부(300)에 의해서 생성된다. 콘텐츠 제공 중계부(300)는 사용자 콘텐츠 인터페이스(200)로부터 음성 녹음 정보(Voice Record)(도 8의 1700)를 입력 받는다. 이때 콘텐츠 제공 중계부(300)는 음성 녹음 정보(Voice Record)(1700)의 음성 속성값(Voice)(도 8의 1720)과 클라이언트 세션 식별정보(Client Session ID)(도 8의 1710), 및 이에 대응하는 클라이언트 세션 테이블(Client Session Table)(도 7의 1600)의 스피치 리스트(Speech List)(도 7의 1640) 속성값을 음성관리부(400)에 전달한다.
- <65> 콘텐츠 제공 중계부(300)는 전송한 정보들에 대한 유효 여부 결과 정보를 음성관리부(400)로부터 제공 받는다. 이에 따라, 녹음된 음성이 유효한 명령인 경우, 콘텐츠 제공 중계부(300)는 클라이언트 세션 테이블(Client Session Table)(1600)과 유효한 명령어를 이용하여 콘텐츠 정보 요구 정보(Content info Request)(1400)를 생

성하고, 이것을 콘텐츠 정보 제공부(120)에 전달한다.

- <66> 한편, 사용자 콘텐츠 인터페이스(200)로부터 제어 명령(Control Command)(도 9의 1800)을 입력 받았을 경우, 콘텐츠 제공 중계부(300)는 클라이언트 세션 테이블(Client Session Table)(1600)에서 해당 클라이언트 세션 식별 정보(Client Session ID)(1610) 값을 가지는 음성 녹음 정보과 제어 명령(Control Command)(1800)의 명령(Command)(1820) 속성값을 이용하여, 콘텐츠 정보 요구 정보(Content info Request)(1400)를 생성하고 이것을 콘텐츠 정보 중계부(120)에 전달한다.
- <67> 도 6은 본 발명의 실시예에 따른 콘텐츠 정보(Content Info)의 데이터 구조 예를 도시한 도면이다.
- <68> 도시된 바와 같이, 콘텐츠 정보(Content Info)(1500)는 콘텐츠 정보를 전달할 사용자 콘텐츠 인터페이스(200)의 종류를 식별하는 클라이언트 세션 식별정보(Client Session ID)(1510), 사용자가 원하는 콘텐츠를 구별하는 콘텐츠 식별정보(Content ID)(1520)와 그 구성 요소 SEQ(1530), 현재 콘텐츠의 요소가 실행 중일 때 유효한 명령 리스트(Speech List)(1540), 콘텐츠 데이터 테이블(Content Data Table)(1300)에 저장된 속성 값으로부터 구체적인 해당 값을 추출한 텍스트 정보(Text)(1550), 그래픽 사용자 인터페이스(GUI)(1560), 및 장치(Device) API(1570)를 포함하여 구성된다.
- <69> 본 발명의 실시예에 따른 콘텐츠 정보(Content Info)(1500)의 생성과정은 아래와 같다. 콘텐츠 정보 제공부(120)는 입력된 콘텐츠 정보 요구정보(Content info Request)(1400)에 포함된 콘텐츠 식별정보(Content ID)(1420), SEQ(1430), 명령(Command)(1440) 3가지 값들과 동일한 속성값을 포함하는 음성 녹음 정보를 콘텐츠 아이템 제어 테이블(Content Item Control Table)(1200)에서 찾는다.
- <70> 이후, 콘텐츠 정보 제공부(120)는 녹음 정보에 포함된 액션 콘텐츠 식별정보(Action Content ID)(1240)와 액션 콘텐츠(Action Content) Seq(1250)의 속성값과 동일한 값을 가지는 식별정보(ID)(1120)와 SEQ(1130)이 포함된 콘텐츠 아이템 테이블(Content Item Table)(1100)의 음성 녹음 정보를 찾고 이 녹음 정보의 콘텐츠 링크(Content Link)(1150)의 속성을 참조하는 콘텐츠 데이터 테이블(Content Data Table)(1300)의 음성녹음정보를 찾는다. 이에 따라, 콘텐츠 정보 제공부(120)는 검색된 콘텐츠 아이템 테이블(Content Item Table)(1100) 음성 녹음 식별정보(ID)(1120), SEQ(1120), 스피치 리스트(Speech List)(1140) 속성값은 콘텐츠 정보(Content Info)(1500)의 콘텐츠 식별정보(Content ID)(1520), SEQ(1530), 스피치 리스트(Speech List)(1540)에 각각 입력하고 검색된 콘텐츠 데이터 테이블(Content Data Table)(1300) 음성녹음의 텍스트(Text)(1310), GUI(1320), 장치(Device) API(1330) 속성값은 콘텐츠 정보(Content Info)(1500)의 텍스트(Text)(1550), GUI(11056), 장치(Device) API(1570)에 각각 입력한다.
- <71> 도 7은 본 발명의 실시예에 따른 콘텐츠 세션 테이블(Content Session Table)의 스키마 구조 예를 도시한 도면이다.
- <72> 도시된 바와 같이, 콘텐츠 세션 테이블(Content Session Table)(1600)은 콘텐츠 제공 중계부(300)에 의해 생성, 갱신, 삭제 등의 연산이 이루어진다. 콘텐츠 세션 테이블(Content Session Table)(1600)의 속성들은 사용자 콘텐츠 인터페이스(200)에서 실행중인 세션을 구별하는 식별자인 클라이언트 세션 식별정보(Client Session ID)(1610), 실행 세션이 사용하는 콘텐츠 식별정보(Content ID)(1620), 콘텐츠를 구성하는 하나의 요소를 가리키는 SEQ(1630), 그리고 현재 콘텐츠와 구성 인자에서 유효한 명령어 리스트를 저장하는 스피치 리스트(Speech List)(1640)로 구성된다.
- <73> 하나의 클라이언트 세션 식별정보(Client Session ID)(1610)는 세션이 진행되는 동안 계속 업데이트 되다가, 세션이 끝나면 삭제된다. 클라이언트 세션 테이블(1600)의 어떤 음성녹음정보의 생성과 업데이트를 위한 값들은 콘텐츠 정보 제공부(120)로부터 입력된 콘텐츠 정보(Content Info)(1500)에 포함된 콘텐츠 식별정보(Content ID)(1520), SEQ(1530) 그리고 스피치 리스트(Speech List)(1540)의 값과 사용자 콘텐츠 인터페이스(200)로부터 입력된 음성녹음정보(Voice Record)(1700)나 제어 명령(Control Command)(1800)에 포함된 클라이언트 세션 식별정보(Client Session ID)(1710 또는 1810)의 값을 추출하여 그 값을 해당 속성에 저장한다.
- <74> 도 8은 본 발명의 실시예에 따른 음성 녹음정보(Voice Record)의 스키마 구조 예를 도시한 도면이다.
- <75> 도시된 바와 같이, 음성 녹음정보(Voice Record)(1700)는 사용자가 음성 명령을 입력할 때 사용자 콘텐츠 인터페이스(200)에 포함된 음성 녹음부(Voice Recorder)에 의해서 음성이 녹음되면, 음성 녹음(Voice Record)(1720) 속성값으로 채우고 현재 클라이언트 세션 식별정보(Client Session ID)(1710)와 함께 콘텐츠 제공 중계부(300)로 전송된다.

- <76> 도 9는 본 발명의 실시예에 따른 제어 명령(Control Command)의 스키마 구조 예를 도시한 도면이다.
- <77> 도시된 바와 같이, 제어 명령(Control Command)(1800)은 사용자 콘텐츠 인터페이스(200)의 'GUI'나 'Sensor Driver'로부터 이벤트가 입력이 되면, 제어지도인 'Control Map'을 참조하여 현재 상태와 입력 이벤트에 대한 명령 심벌을 명령(Command)(1820)의 속성값으로 채우고 클라이언트 세션 식별정보(Client Session ID)(1810)와 함께 콘텐츠 제공 중계부(300)로 전송된다.
- <78> 도 10은 본 발명의 실시예에 따른 제어지도(Control Map)의 스키마 구조 예를 도시한 도면이다.
- <79> 도시된 바와 같이, 제어지도(Control Map)(1900)는 사용자 콘텐츠 인터페이스(200)의 상태 조건을 논리식으로 표현한 'Pre Condition(1920)'과 'GUI'나 'Sensor Driver'로부터 입력되는 컨트롤 이벤트를 의미하는 'Physical Control'(1930) 그리고 'Pre Condition(1920)'과 'Physical Control(1930)'이 만족될 때 사용자가 의도하는 명령 심벌을 저장하는 'Speech Command(1940)'로 구성된다.
- <80> 이 음성 녹음 값들은 사용자 콘텐츠 인터페이스(200)가 제공할 수 있는 하드웨어의 특징과 콘텐츠 정보 서비스의 종류를 고려하여 사용자 콘텐츠 인터페이스(200)가 실행되기 이전에 미리 값이 설정되어 있어야 한다. 또한, 콘텐츠나 'GUI' 등이 갱신되면 제어 지도(Control Map)(1900)도 갱신되어야 한다.
- <81> 이상에서는 본 발명에서 특정의 바람직한 실시예에 대하여 도시하고 또한 설명하였다. 그러나 본 발명은 상술한 실시예에 한정되지 아니하며, 특허 청구의 범위에서 첨부하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 누구든지 다양한 변형 및 균등한 타 실시가 가능할 것이다. 따라서 본 발명의 진정한 기술적 보호범위는 첨부한 특허청구범위에 의해서만 정해져야 할 것이다.

**발명의 효과**

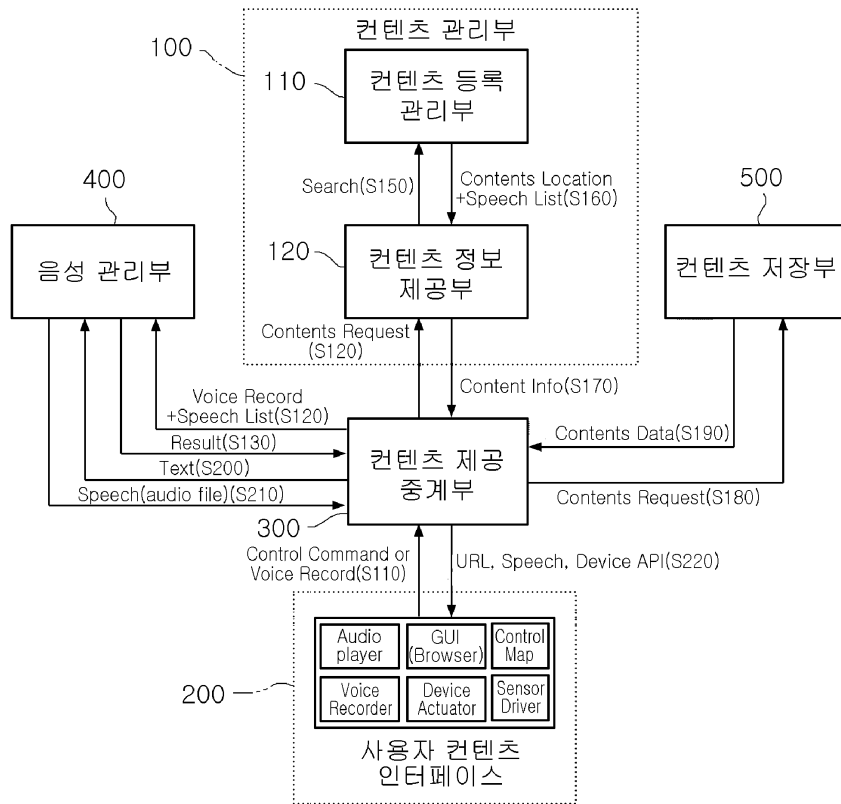
- <82> 상술한 본 발명에 따르면, 음성 인식 및/또는 입출력 가능한 사용자 인터페이스를 통해 사용자로부터 입력되는 음성 또는 문자 형태의 해당 콘텐츠 제공 요청정보를 인식하고 이에 대응하는 콘텐츠 정보를 검색하여 사용자에게 제공 가능한 인터페이스를 통해 제공함으로써, 디지털 홈 또는 차량에 설치되는 다양한 통신 기기들이 표시 장치의 구비 유무에 상관없이 사용자에게 입출 가능한 사용자 인터페이스를 통해 보다 용이하게 해당 콘텐츠 정보 서비스를 제공할 수 있다.

**도면의 간단한 설명**

- <1> 도 1은 본 발명의 바람직한 일 실시예에 따른 음성 인식을 통한 콘텐츠 서비스 제공 장치를 도시한 블록도,
- <2> 도 2는 본 발명의 실시예에 따른 콘텐츠 아이템 테이블(Content Item Table)의 스키마 구조를 나타낸 도면,
- <3> 도 3은 본 발명의 실시예에 따른 콘텐츠 아이템 제어 테이블(Content Item Control Table)의 스키마 구조를 나타낸 도면,
- <4> 도 4는 본 발명의 실시예에 따른 콘텐츠 데이터 테이블(Content Data Table)의 스키마 구조를 나타낸 도면,
- <5> 도 5는 본 발명의 실시예에 따른 콘텐츠 정보 요구(Content info Request) 정보의 데이터 구조 예를 도시한 도면,
- <6> 도 6은 본 발명의 실시예에 따른 콘텐츠 정보(Content Info)의 데이터 구조 예를 도시한 도면,
- <7> 도 7은 본 발명의 실시예에 따른 콘텐츠 세션 테이블(Content Session Table)의 스키마 구조 예를 도시한 도면,
- <8> 도 8은 본 발명의 실시예에 따른 음성 녹음정보(Voice Record)의 스키마 구조 예를 도시한 도면,
- <9> 도 9는 본 발명의 실시예에 따른 제어 명령(Control Command)의 스키마 구조 예를 도시한 도면, 그리고
- <10> 도 10은 본 발명의 실시예에 따른 제어지도(Control Map)의 스키마 구조 예를 도시한 도면이다.

도면

도면1



Category (1110)	ID (1120)	SEQ (1130)	Speech List (1140)	Content_link (1150)
Recipe/kimbap	1001	1	{다음:next},(반복:repeat),(정지:stop)}	300
Recipe/kimbap	1001	2	{다음:next},(이전:pre),(반복:repeat),(정지:stop)}	301
Recipe/kimbap	1001	3		302
Recipe/kimbap	1001	4		303
Recipe/kimbap	1001	5		304

1100

1160

도면2

도면3

Current Content_id (1210)	Current_Seq (1220)	Control symbol (1230)	Action Content_id (1240)	Action Content_Seq (1250)
1001	1	next	1001	2
1001	1	repeat		
1001	1	stop		

1200



도면4

id	Text (Service invocation, API) (1310)	GUI (HTML, IMG) (1320)	Device API (1330)
300	“맛있는 쿡밤 요리에 대해서 말씀 드리겠습니다.”	Img:=receipe.getFirsImage(“쿡밤”); or <a href="http://www.kth.com/recipe/쿡밤1.html">http://www.kth.com/recipe/쿡밤1.html</a>	
301	“쿡 100g, 불린 쌀 3컵, 다시 국물 31/2컵...”	Img:=receipe.getingImage(“쿡밤”); <a href="http://www.kth.com/recipe/쿡밤2.html">http://www.kth.com/recipe/쿡밤2.html</a>	
302	은행, 밤, 대추는 손질하여 먹기 좋은 크기로 준비한다.	<a href="http://www.kth.com/recipe/쿡밤3.html">http://www.kth.com/recipe/쿡밤3.html</a>	
303	불린 쌀과 1, 2를 합하여 다시마 우린 물을 넣어 밥을 짓는다.	<a href="http://www.kth.com/recipe/쿡밤4.html">http://www.kth.com/recipe/쿡밤4.html</a>	
304	양념장 재료를 골고루 섞어 양념장을 만들어 놓는다.	<a href="http://www.kth.com/recipe/쿡밤5.html">http://www.kth.com/recipe/쿡밤5.html</a>	

1300

도면5

Client Session ID (1410)	Content ID (1420)	SEQ (1430)	Command (1440)
UVID_1231	1001	1	next

1400

도면6

Client Session ID (1510)	Content ID (1520)	SEQ (1530)	Speech List (1540)	Text (1550)	GUI (1560)	Device API (1570)
UUID_1234	1001	2	{(다음:next),(이전:pre), 반복:repeat),정지(stop)}	글 100g, 불린 쌀 3컵, 다시죽물 31/2컵... or <a href="http://www.kth.com/recipe/쿠폰1.html">http://www.kth.com/ recipe/쿠폰1.html</a>	<a href="http://www.kth.com/recipe/쿠폰1.html">http://www.kth.com/ recipe/ 쿠폰1.html</a>	MoveHead(15)

1500

도면7

Client Session ID (1610)	Content ID (1620)	SEQ (1630)	Speech List (1640)
UUID_1234	1001	1	{{다음:next},{이전:pre},{반복:repeat},{정지:stop}}
UUID_1234	1131	2	{{날씨:weather},{뉴스:news},{일정:schedule}}

1600

도면8

Client Session ID (1710)	Voice (1720)
UVID_1234	Byte{Voice Data}

1700

도면9

Client Session ID (1810)	Command (1820)
UVID_1234	next

도면10

Num (1910)	Precondition (1920)	Physical Control (1930)	Speech Command (1940)
1	Content('Recipe') ^ GLLON ^ Speaking	Botton Pushed(bt1)	next
2	Content('Recipe') ^ GLLON ^ Speaking	Botton Pushed(bt2)	repeat
3	Content('Recipe') ^ GLLON ^ Speaking	Botton Pushed(bt3)	stop
4	Content(' Weather ') ^ Ready to Speak	Sensor Touched(s1)	start