



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

(45) 공고일자 2022년04월01일  
(11) 등록번호 10-2380540  
(24) 등록일자 2022년03월25일

(51) 국제특허분류(Int. Cl.)  
H04N 21/439 (2011.01) H04N 21/81 (2011.01)  
H04N 21/845 (2011.01)  
(52) CPC특허분류  
H04N 21/4394 (2013.01)  
H04N 21/8106 (2013.01)  
(21) 출원번호 10-2020-0117628  
(22) 출원일자 2020년09월14일  
심사청구일자 2020년09월14일  
(65) 공개번호 10-2022-0035635  
(43) 공개일자 2022년03월22일  
(56) 선행기술조사문헌  
KR100774708 B1\*  
KR1020140024433 A\*  
KR1020190051265 A\*  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
네이버 주식회사  
경기도 성남시 분당구 불정로 6, 그린팩토리 (정자동)  
(72) 발명자  
박종은  
경기도 성남시 분당구 불정로 6(정자동, 그린팩토리)  
김대황  
경기도 성남시 분당구 불정로 6(정자동, 그린팩토리)  
(뒷면에 계속)  
(74) 대리인  
양성보

전체 청구항 수 : 총 20 항

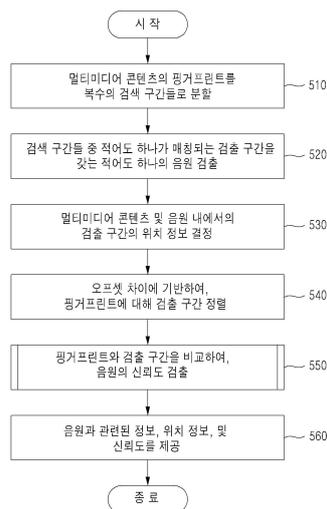
심사관 : 구영희

**(54) 발명의 명칭 음원을 검출하기 위한 전자 장치 및 그의 동작 방법**

**(57) 요약**

다양한 실시예들은 멀티미디어 콘텐츠에 사용된 음원을 검출하기 위한 전자 장치 및 그의 동작 방법에 관한 것으로, 멀티미디어 콘텐츠의 핑거프린트를 미리 설정된 시간 간격에 따라 복수의 검색 구간들로 분할하고, 검색 구간들 중 적어도 하나가 매칭되는 검출 구간을 갖는 적어도 하나의 음원을 검출하고, 멀티미디어 콘텐츠 내에서의 검출 구간의 시간 위치 및 음원 내에서의 검출 구간의 시간 위치를 나타내는 위치 정보를 결정하고, 음원과 관련된 정보 및 위치 정보를 제공하도록 구성될 수 있다.

**대표도 - 도5**



(52) CPC특허분류

*H04N 21/845* (2013.01)

(72) 발명자

**서동우**

경기도 성남시 분당구 불정로 6(정자동, 그린팩토리)

**김동환**

경기도 성남시 분당구 불정로 6(정자동, 그린팩토리)

**전지수**

경기도 성남시 분당구 불정로 6(정자동, 그린팩토리)

**이장희**

경기도 성남시 분당구 불정로 6(정자동, 그린팩토리)

## 명세서

### 청구범위

#### 청구항 1

전자 장치의 동작 방법에 있어서,

멀티미디어 콘텐츠의 핑거프린트를 미리 설정된 시간 간격에 따라 복수의 검색 구간들로 분할하는 단계;

상기 검색 구간들 중 적어도 하나가 매칭되는 검출 구간을 갖는 적어도 하나의 음원을 검출하는 단계;

상기 멀티미디어 콘텐츠 내에서의 상기 검출 구간의 시간 위치 및 상기 음원 내에서의 상기 검출 구간의 시간 위치를 나타내는 위치 정보를 결정하는 단계; 및

상기 음원과 관련된 정보 및 상기 위치 정보를 사용자에게 제공하는 단계를 포함하는, 방법.

#### 청구항 2

제 1 항에 있어서,

상기 멀티미디어 콘텐츠의 시작점으로부터의 시간 오프셋과 상기 음원의 시작점으로부터의 시간 오프셋 사이의 오프셋 차이에 기반하여, 상기 핑거프린트에 대해 상기 검출 구간을 정렬시키는 단계; 및

상기 핑거프린트와 상기 검출 구간을 비교하여, 상기 음원의 신뢰도를 검출하는 단계

를 더 포함하는, 방법.

#### 청구항 3

제 2 항에 있어서,

상기 음원과 관련된 정보 및 상기 위치 정보를 제공하는 단계는,

상기 음원과 관련된 정보 및 상기 위치 정보와 함께, 상기 신뢰도를 제공하는 단계

를 포함하는, 방법.

#### 청구항 4

제 2 항에 있어서,

상기 신뢰도를 검출하는 단계는,

상기 핑거프린트와 상기 검출 구간의 비교 연산을 통해, 비교 구간을 생성하는 단계;

상기 비교 구간을 복수의 비트 구간들로 분할하는 단계;

상기 비트 구간들의 비트 에러율들을 각각 계산하는 단계;

상기 비트 에러율들을 상기 비트 구간들의 점수들로 각각 변환하는 단계; 및

상기 점수들의 합으로부터 상기 신뢰도를 검출하는 단계

를 포함하는, 방법.

#### 청구항 5

제 4 항에 있어서,

상기 점수들로 변환하는 단계는,

상기 비트 에러율들 중 적어도 하나가 임계값을 초과하면, 상기 적어도 하나의 비트 에러율에 대응하여 0을 점수로 부여하는 단계; 및

상기 비트 에러율들 중 나머지가 상기 임계값 이하이면, 상기 나머지의 비트 에러율에 대응하여 계산되는 점수를 부여하는 단계

를 포함하는, 방법.

#### 청구항 6

제 1 항에 있어서,

상기 음원과 관련된 정보 및 상기 위치 정보를 제공하는 단계는,

상기 멀티미디어 콘텐츠가 복수의 음원들과 연관되면, 상기 음원과 관련된 정보 및 상기 위치 정보를 리스트로 제공하는 단계

를 포함하는, 방법.

#### 청구항 7

제 1 항에 있어서,

상기 멀티미디어 콘텐츠는,

영상 데이터 또는 오디오 데이터 중 적어도 하나로 이루어지는,

방법.

#### 청구항 8

제 1 항에 있어서,

상기 음원과 관련된 정보가 선택되면, 상기 음원을 제공하는 단계; 또는

상기 음원과 관련된 정보가 선택되면, 상기 음원과 연관된 다른 멀티미디어 콘텐츠를 제공하는 단계

중 적어도 하나를 더 포함하는, 방법.

#### 청구항 9

제 1 항에 있어서,

상기 검색 구간들 중 적어도 하나는,

복수의 음원들에 매칭되는, 방법.

#### 청구항 10

제 1 항 내지 제 9 항 중 어느 한 항의 동작 방법을 상기 전자 장치에 실행시키기 위해 비-일시적인 컴퓨터 판독 가능한 기록 매체에 저장되는 컴퓨터 프로그램.

#### 청구항 11

제 1 항 내지 제 9 항 중 어느 한 항의 동작 방법을 상기 전자 장치에 실행시키기 위한 프로그램이 기록되어 있는 비-일시적인 컴퓨터 판독 가능한 기록 매체.

#### 청구항 12

전자 장치에 있어서,

메모리; 및

상기 메모리와 연결되고, 상기 메모리에 저장된 적어도 하나의 명령을 실행하도록 구성된 프로세서를 포함하고,

상기 프로세서는,

멀티미디어 콘텐츠의 핑거프린트를 미리 설정된 시간 간격에 따라 복수의 검색 구간들로 분할하고,

상기 검색 구간들 중 적어도 하나가 매칭되는 검출 구간을 갖는 적어도 하나의 음원을 검출하고,

상기 멀티미디어 콘텐츠 내에서의 상기 검출 구간의 시간 위치 및 상기 음원 내에서의 상기 검출 구간의 시간 위치를 나타내는 위치 정보를 결정하고,

상기 음원과 관련된 정보 및 상기 위치 정보를 사용자에게 제공하도록 구성되는,

장치.

#### 청구항 13

제 12 항에 있어서,

상기 프로세서는,

상기 멀티미디어 콘텐츠의 시작점으로부터의 시간 오프셋과 상기 음원의 시작점으로부터의 시간 오프셋 사이의 오프셋 차이에 기반하여, 상기 핑거프린트에 대해 상기 검출 구간을 정렬시키고,

상기 핑거프린트와 상기 검출 구간을 비교하여, 상기 음원의 신뢰도를 검출하도록 구성되는,

장치.

#### 청구항 14

제 12 항에 있어서,

상기 프로세서는,

상기 검출 구간에 대응하는 상기 음원의 신뢰도를 검출하고,

상기 음원과 관련된 정보 및 상기 위치 정보와 함께, 상기 신뢰도를 제공하도록 구성되는,

장치.

#### 청구항 15

제 14 항에 있어서,

상기 프로세서는,

상기 핑거프린트와 상기 검출 구간의 비교 연산을 통해, 비교 구간을 생성하고,  
 상기 비교 구간을 복수의 비트 구간들로 분할하고,  
 상기 비트 구간들의 비트 에러율들을 각각 계산하고,  
 상기 비트 에러율들을 상기 비트 구간들의 점수들로 각각 변환하고,  
 상기 점수들의 합으로부터 상기 신뢰도를 검출하도록 구성되는,  
 장치.

**청구항 16**

제 15 항에 있어서,  
 상기 프로세서는,  
 상기 비트 에러율들 중 적어도 하나가 임계값을 초과하면, 상기 적어도 하나의 비트 에러율에 대응하여 0을 점수로 부여하고,  
 상기 비트 에러율들 중 나머지가 상기 임계값 이하이면, 상기 나머지의 비트 에러율에 대응하여 계산되는 점수를 부여하도록 구성되는,  
 장치.

**청구항 17**

제 12 항에 있어서,  
 상기 프로세서는,  
 상기 멀티미디어 콘텐츠가 복수의 음원들과 연관되면, 상기 음원과 관련된 정보 및 상기 위치 정보를 리스트로 제공하도록 구성되는,  
 장치.

**청구항 18**

제 12 항에 있어서,  
 상기 멀티미디어 콘텐츠는,  
 영상 데이터 또는 오디오 데이터 중 적어도 하나로 이루어지는,  
 장치.

**청구항 19**

제 12 항에 있어서,  
 상기 프로세서는,  
 상기 음원과 관련된 정보가 선택되면, 상기 음원 또는 상기 음원과 연관된 다른 멀티미디어 콘텐츠 중 적어도 하나를 제공하도록 구성되는,  
 장치.

**청구항 20**

제 12 항에 있어서,  
 상기 검색 구간들 중 적어도 하나는,  
 복수의 음원들에 매칭되는,  
 장치.

**발명의 설명**

**기술 분야**

[0001] 다양한 실시예들은 멀티미디어 콘텐츠(multimedia content)에 사용된 적어도 하나의 음원(audio source)을 검출하기 위한 전자 장치 및 그의 동작 방법에 관한 것이다.

**배경 기술**

[0002] 음원 검출 기술은 멀티미디어 콘텐츠에 사용된 음원을 검출하는 기술이다. 일반적으로, 서버에는, 복수의 음원들이 등록되며, 음원들의 핑거프린트(finger print)들이 각각 저장되어 있다. 이러한 서버는 음원 검출 기술을 통해, 멀티미디어 콘텐츠의 핑거프린트를 기반으로, 등록된 음원들로부터 멀티미디어 콘텐츠에 사용된 음원을 검출한다. 이를 통해, 서버는 음원에 대한 정보와 음원 내에서 멀티미디어 콘텐츠에 사용된 부분의 시작 위치를 제공한다.

[0003] 그러나, 상기와 같은 서버에서, 멀티미디어 콘텐츠에 사용된 음원을 검출하기 위한 동작 성능이 낮은 문제점이 있다. 구체적으로, 서버가 멀티미디어 콘텐츠의 전체의 핑거프린트를 등록된 음원들의 핑거프린트들과 비교해야 하기 때문에, 서버의 연산량이 증가되어, 서버의 동작 효율성이 낮다. 그리고, 서버가 음원 내에서 멀티미디어 콘텐츠에 사용된 부분을 정확하게 검출하는 데 어려움이 있다.

**발명의 내용**

**해결하려는 과제**

[0004] 다양한 실시예들은, 멀티미디어 콘텐츠에 사용된 적어도 하나의 음원을 효율적으로 검출할 수 있는 전자 장치 및 그의 동작 방법을 제공한다.

[0005] 다양한 실시예들은, 멀티미디어 콘텐츠의 핑거프린트를 부분적으로 이용하여, 음원 내에서 멀티미디어 콘텐츠에 사용된 부분을 효율적으로 검출할 수 있는 전자 장치 및 그의 동작 방법을 제공한다.

[0006] 다양한 실시예들은, 음원 내에서 멀티미디어 콘텐츠에 사용된 부분에 대해 음원 내에서의 시간 위치뿐 아니라 멀티미디어 콘텐츠 내에서의 시간 위치를 검출함으로써, 음원 내에서 멀티미디어 콘텐츠에 사용된 부분을 보다 정확하게 특정할 수 있는 전자 장치 및 그의 동작 방법을 제공한다.

[0007] 다양한 실시예들은, 음원 내에서 멀티미디어 콘텐츠에 사용된 부분의 음원 내에서의 시간 위치와 멀티미디어 콘텐츠 내에서의 시간 위치에 기반하여, 음원에 대한 신뢰도를 검출할 수 있는 전자 장치 및 그의 동작 방법을 제공한다.

**과제의 해결 수단**

[0008] 다양한 실시예들에 따른 전자 장치의 동작 방법은, 멀티미디어 콘텐츠의 핑거프린트를 미리 설정된 시간 간격에 따라 복수의 검색 구간들로 분할하는 단계, 상기 검색 구간들 중 적어도 하나가 매칭되는 검출 구간을 갖는 적어도 하나의 음원을 검출하는 단계, 상기 멀티미디어 콘텐츠 내에서의 상기 검출 구간의 시간 위치 및 상기 음원 내에서의 상기 검출 구간의 시간 위치를 나타내는 위치 정보를 결정하는 단계, 및 상기 음원과 관련된 정보 및 상기 위치 정보를 제공하는 단계를 포함할 수 있다.

[0009] 다양한 실시예들에 따른 컴퓨터 프로그램은, 상기 동작 방법을 상기 전자 장치에 실행시키기 위해 비-일시적인 컴퓨터 판독 가능한 기록 매체에 저장될 수 있다.

[0010] 다양한 실시예들에 따른 비-일시적인 컴퓨터 판독 가능한 기록 매체는, 상기 동작 방법을 상기 전자 장치에 실행

행시킴을 위한 프로그램이 기록되어 있다.

[0011] 다양한 실시예들에 따른 전자 장치는, 메모리, 및 상기 메모리와 연결되고, 상기 메모리에 저장된 적어도 하나의 명령을 실행하도록 구성된 프로세서를 포함하고, 상기 프로세서는, 멀티미디어 콘텐츠의 핑거프린트를 미리 설정된 시간 간격에 따라 복수의 검색 구간들로 분할하고, 상기 검색 구간들 중 적어도 하나가 매칭되는 검색 구간을 갖는 적어도 하나의 음원을 검출하고, 상기 멀티미디어 콘텐츠 내에서의 상기 검색 구간의 시간 위치 및 상기 음원 내에서의 상기 검색 구간의 시간 위치를 나타내는 위치 정보를 결정하고, 상기 음원과 관련된 정보 및 상기 위치 정보를 제공하도록 구성될 수 있다.

**발명의 효과**

[0012] 다양한 실시예들에 따르면, 전자 장치는 멀티미디어 콘텐츠에 사용된 적어도 하나의 음원을 효율적으로 검출할 수 있다. 구체적으로, 전자 장치는 멀티미디어 콘텐츠의 핑거프린트에서 검색 구간들 중 하나로부터 시간 범위를 확장시키면서, 음원 내에서 멀티미디어 콘텐츠에 매칭되는 검색 구간을 효율적으로 검출할 수 있다. 그리고, 전자 장치는 음원 내에서의 검색 구간의 시간 위치뿐만 아니라 멀티미디어 콘텐츠 내에서의 검색 구간의 시간 위치를 검출함으로써, 음원 및 멀티미디어 콘텐츠 내에서 검색 구간을 보다 정확하게 특정할 수 있다. 아울러, 전자 장치는 검색 구간에 대한 멀티미디어 콘텐츠의 시작점으로부터 시간 오프셋과 음원의 시작점으로부터의 시간 오프셋 사이의 오프셋 차이에 기반하여 멀티미디어 콘텐츠와 음원을 비교함으로써, 음원에 대한 신뢰도를 검출할 수 있다. 이를 통해, 전자 장치는 사용자를 위해, 음원과 관련된 정보와 위치 정보뿐만 아니라, 신뢰도를 제공할 수 있다.

**도면의 간단한 설명**

[0013] 도 1은 다양한 실시예들에 따른 전자 장치를 도시하는 도면이다.  
 도 2, 도 3a, 도 3b, 도 3c, 및 도 3d는 도 1의 프로세서의 동작 특징을 예시적으로 설명하기 위한 도면들이다.  
 도 4는 도 1의 프로세서를 세부적으로 도시하는 도면이다.  
 도 5는 다양한 실시예들에 따른 전자 장치의 동작 방법을 도시하는 도면이다.  
 도 6은 도 5의 음원의 신뢰도 검출 단계를 세부적으로 도시하는 도면이다.  
 도 7, 도 8, 도 9, 도 10, 도 11, 도 12, 도 13, 도 14, 도 15, 도 16, 및 도 17은 다양한 실시예들에 따른 전자 장치의 동작 방법을 예시적으로 설명하기 위한 도면들이다.

**발명을 실시하기 위한 구체적인 내용**

[0014] 이하, 본 문서의 다양한 실시예들이 첨부된 도면을 참조하여 설명된다.

[0016] 도 1은 다양한 실시예들에 따른 전자 장치(100)를 도시하는 도면이다. 도 2, 도 3a, 도 3b, 도 3c, 및 도 3d는 도 1의 프로세서(160)의 동작 특징을 예시적으로 설명하기 위한 도면들이다. 도 4는 도 1의 프로세서(160)를 세부적으로 도시하는 도면이다.

[0017] 도 1을 참조하면, 다양한 실시예들에 따른 전자 장치(100)는 연결 단자(110), 통신 모듈(120), 입력 모듈(130), 출력 모듈(140), 메모리(150), 또는 프로세서(160) 중 적어도 어느 하나를 포함할 수 있다. 어떤 실시예에서, 전자 장치(100)의 구성 요소들 중 적어도 어느 하나가 생략될 수 있으며, 적어도 하나의 다른 구성 요소가 추가될 수 있다. 어떤 실시예에서, 전자 장치(100)의 구성 요소들 중 적어도 어느 두 개가 하나의 통합된 회로로 구현될 수 있다. 예를 들면, 전자 장치(100)는 서버(server), 스마트폰(smart phone), 휴대폰, 내비게이션, 컴퓨터, 노트북, 디지털방송용 단말, PDA(personal digital assistants), PMP(portable multimedia player), 태블릿 PC, 게임 콘솔(game console), 웨어러블 디바이스(wearable device), IoT(internet of things) 디바이스, 가전 기기, 의료 기기, 또는 로봇(robot) 중 적어도 어느 하나를 포함할 수 있다.

[0018] 연결 단자(110)는 전자 장치(100)에서 외부 장치(102)와 물리적으로 연결될 수 있다. 예를 들면, 외부 장치(102)는 다른 전자 장치를 포함할 수 있다. 이를 위해, 연결 단자(110)는 적어도 하나의 커넥터를 포함할 수 있다. 예를 들면, 커넥터는 HDMI 커넥터, USB 커넥터, SD 카드 커넥터, 또는 오디오 커넥터 중 적어도 어느 하나를 포함할 수 있다.

[0019] 통신 모듈(120)은 전자 장치(100)에서 외부 장치(102, 104)와 통신을 수행할 수 있다. 통신 모듈(120)은 전자

장치(100)와 외부 장치(102, 104) 간 통신 채널을 수립하고, 통신 채널을 통해, 외부 장치(102, 104)와 통신을 수행할 수 있다. 여기서, 외부 장치(102, 104)는 위성, 기지국, 또는 다른 전자 장치 중 적어도 어느 하나를 포함할 수 있다. 통신 모듈(120)은 유선 통신 모듈 또는 무선 통신 모듈 중 적어도 어느 하나를 포함할 수 있다. 유선 통신 모듈은 연결 단자(102)를 통해, 외부 장치(102)와 유선으로 연결되어, 유선으로 통신할 수 있다. 무선 통신 모듈은 근거리 통신 모듈 또는 원거리 통신 모듈 중 적어도 어느 하나를 포함할 수 있다. 근거리 통신 모듈은 외부 장치(102)와 근거리 통신 방식으로 통신할 수 있다. 예를 들면, 근거리 통신 방식은, 블루투스(Bluetooth), 와이파이 다이렉트(WiFi direct), 또는 적외선 통신(IrDA; infrared data association) 중 적어도 어느 하나를 포함할 수 있다. 원거리 통신 모듈은 외부 장치(104)와 원거리 통신 방식으로 통신할 수 있다. 여기서, 원거리 통신 모듈은 네트워크(190)를 통해 외부 장치(104)와 통신할 수 있다. 예를 들면, 네트워크(190)는 셀룰러 네트워크, 인터넷, 또는 LAN(local area network)이나 WAN(wide area network)과 같은 컴퓨터 네트워크 중 적어도 어느 하나를 포함할 수 있다.

[0020] 입력 모듈(130)은 전자 장치(100)의 적어도 하나의 구성 요소에 사용될 신호를 입력할 수 있다. 입력 모듈(130)은, 사용자가 전자 장치(100)에 직접적으로 신호를 입력하도록 구성되는 입력 장치, 주변 환경을 감지하여 신호를 발생하도록 구성되는 센서 장치, 또는 영상을 촬영하여, 영상 데이터를 생성하도록 구성되는 카메라 모듈 중 적어도 어느 하나를 포함할 수 있다. 예를 들면, 입력 장치는 마이크론(microphone), 마우스(mouse), 또는 키보드(keyboard) 중 적어도 어느 하나를 포함할 수 있다. 어떤 실시예에서, 센서 장치는 터치를 감지하도록 설정된 터치 회로(touch circuitry) 또는 터치에 의해 발생하는 힘의 세기를 측정하도록 설정된 센서 회로 중 적어도 어느 하나를 포함할 수 있다.

[0021] 출력 모듈(140)은 정보를 출력할 수 있다. 출력 모듈(140)은 정보를 시각적으로 표시하도록 구성되는 표시 모듈 또는 정보를 청각적으로 재생하도록 구성되는 오디오 모듈 중 적어도 하나를 포함할 수 있다. 예를 들면, 표시 모듈은 디스플레이, 홀로그램 장치, 또는 프로젝터 중 적어도 어느 하나를 포함할 수 있다. 일 예로, 표시 모듈은 입력 모듈(130)의 터치 회로 또는 센서 회로 중 적어도 어느 하나와 조립되어, 터치 스크린으로 구현될 수 있다. 예를 들면, 오디오 모듈은 스피커 또는 리시버 중 적어도 어느 하나를 포함할 수 있다.

[0022] 메모리(150)는 전자 장치(100)의 적어도 하나의 구성 요소에 의해 사용되는 다양한 데이터를 저장할 수 있다. 예를 들면, 메모리(150)는 휘발성 메모리 또는 비휘발성 메모리 중 적어도 어느 하나를 포함할 수 있다. 데이터는 적어도 하나의 프로그램 및 이와 관련된 입력 데이터 또는 출력 데이터를 포함할 수 있다. 프로그램은 메모리(150)에 적어도 하나의 명령을 포함하는 소프트웨어로서 저장될 수 있으며, 예컨대 운영 체제, 미들 웨어, 또는 어플리케이션 중 적어도 어느 하나를 포함할 수 있다.

[0023] 프로세서(160)는 메모리(150)의 프로그램을 실행하여, 전자 장치(100)의 적어도 하나의 구성 요소를 제어할 수 있다. 이를 통해, 프로세서(160)는 데이터 처리 또는 연산을 수행할 수 있다. 이 때 프로세서(160)는 메모리(150)에 저장된 명령을 실행할 수 있다. 프로세서(160)는 멀티미디어 콘텐츠(multimedia content)에 사용된 적어도 하나의 음원(audio source)에 대한 검출을 시도할 수 있다. 여기서, 멀티미디어 콘텐츠는 영상 데이터 또는 오디오 데이터 중 적어도 하나로 이루어질 수 있다. 일 예로, 멀티미디어 콘텐츠는 영상 데이터와 오디오 데이터로 이루어지며, 뮤직 비디오, 네트워크를 통해 공유되는 동영상 등을 포함할 수 있다. 다른 예로, 멀티미디어 콘텐츠는 오디오 데이터로 이루어지며, 팟캐스트, 방송국 등에서 생성될 수 있다. 그리고, 도 2에 도시된 바와 같이, 멀티미디어 콘텐츠의 오디오 데이터에는, 적어도 하나의 음원이 사용될 수 있으며, 각 음원의 적어도 일부가 사용될 수 있다.

[0024] 다양한 실시예들에 따르면, 프로세서(160)는 멀티미디어 콘텐츠에 대응하여, 음원의 적어도 일부를 검출 구간(200)으로서 검출할 수 있다. 이 때 검출 구간(200)에 대한 오프셋 차이가 정의될 수 있다. 오프셋 차이( $\Delta T_m - \Delta T_a$ )는 멀티미디어 콘텐츠의 시작점( $T_{m0}$ )으로부터 검출 구간(200)의 시작점( $T_{d0}$ )까지의 시간 오프셋( $\Delta T_m$ )과 음원의 시작점( $T_{a0}$ )으로부터 검출 구간(200)의 시작점( $T_{d0}$ )까지의 시간 오프셋( $\Delta T_a$ ) 사이의 차이를 나타낼 수 있다. 여기서, 오프셋 차이( $\Delta T_m - \Delta T_a$ )는 한 값으로 정의될 수 있으며, 일정 범위 내의 값들로 정의될 수도 있다. 일 예로, 오프셋 차이( $\Delta T_m - \Delta T_a$ )는 멀티미디어 콘텐츠의 시작점( $T_{m0}$ )으로부터의 시간 오프셋( $\Delta T_m$ )과 음원의 시작점( $T_{a0}$ )으로부터의 시간 오프셋( $\Delta T_a$ ) 사이의 차이를 중심으로 하는 범위 내의 값들로 정의될 수 있다. 오프셋 차이( $\Delta T_m - \Delta T_a$ )가 일정 범위 내의 값들로 정의되는 경우, 동일한 음원에 대한 다양한 재생 속도들이 고려될 수 있다.

[0025] 제 1 예로, 도 3a에 도시된 바와 같이, 검출 구간(200)은 음원의 전체 영역일 수 있으며, 멀티미디어 콘텐츠의 일부 영역에 사용될 수 있다. 여기서, 음원의 시작점( $T_{a0}$ )으로부터 검출 구간(200)의 시작점( $T_{d0}$ )까지의 시간 오프셋( $\Delta T_a$ )은 0이므로, 오프셋 차이( $\Delta T_m - \Delta T_a$ )는 멀티미디어 콘텐츠의 시작점( $T_{m0}$ )으로부터 검출 구간(200)의 시작점( $T_{d0}$ )까지의 시간 오프셋( $\Delta T_m$ )으로 결정될 수 있다(오프셋 차이 =  $\Delta T_m$ ). 제 2 예로, 도 3b에 도시된 바와 같이, 검출 구간(200)은 음원의 일부 영역일 수 있으며, 멀티미디어 콘텐츠의 일부 영역에 사용될 수 있다. 여기서, 멀티미디어 콘텐츠의 시작점( $T_{m0}$ )으로부터 검출 구간(200)의 시작점( $T_{d0}$ )까지의 시간 오프셋( $\Delta T_m$ )이므로, 오프셋 차이( $\Delta T_m - \Delta T_a$ )는 음원의 시작점( $T_{a0}$ )으로부터 검출 구간(200)의 시작점( $T_{d0}$ )까지의 시간 오프셋( $\Delta T_a$ )으로부터 결정될 수 있다(오프셋 차이 =  $-\Delta T_a$ ). 제 3 예로, 도 3c에 도시된 바와 같이, 검출 구간(200)은 음원의 일부 영역일 수 있으며, 멀티미디어 콘텐츠의 일부 영역에 사용될 수 있다. 여기서, 음원의 시작점( $T_{a0}$ )으로부터 검출 구간(200)의 시작점( $T_{d0}$ )까지의 시간 오프셋( $\Delta T_a$ )은 0이므로, 오프셋 차이( $\Delta T_m - \Delta T_a$ )는 멀티미디어 콘텐츠의 시작점( $T_{m0}$ )으로부터 검출 구간(200)의 시작점( $T_{d0}$ )까지의 시간 오프셋( $\Delta T_m$ )으로 결정될 수 있다(오프셋 차이 =  $\Delta T_m$ ). 제 4 예로, 도 3d에 도시된 바와 같이, 검출 구간(200)은 음원의 일부 영역일 수 있으며, 멀티미디어 콘텐츠의 전체 영역에 사용될 수 있다. 여기서, 멀티미디어 콘텐츠의 시작점( $T_{m0}$ )으로부터 검출 구간(200)의 시작점( $T_{d0}$ )까지의 시간 오프셋( $\Delta T_m$ )이므로, 오프셋 차이( $\Delta T_m - \Delta T_a$ )는 음원의 시작점( $T_{a0}$ )으로부터 검출 구간(200)의 시작점( $T_{d0}$ )까지의 시간 오프셋( $\Delta T_a$ )으로부터 결정될 수 있다(오프셋 차이 =  $-\Delta T_a$ ).

[0026] 다양한 실시예들에 따르면, 프로세서(160)는 오프셋 차이( $\Delta T_m - \Delta T_a$ )를 이용하여, 음원의 신뢰도(confidence)를 검출할 수 있다. 신뢰도는 검출된 음원이 멀티미디어 콘텐츠에 사용된 것인지에 대한 정확도를 나타내는 것으로, 신뢰도가 높을수록, 정확도가 높을 수 있다. 구체적으로, 프로세서(160)는 오프셋 차이( $\Delta T_m - \Delta T_a$ )에 기반하여, 멀티미디어 콘텐츠에 대해 검출 구간(200)을 정렬시킬 수 있다. 그리고, 프로세서(160)는 멀티미디어 콘텐츠와 검출 구간(200)을 비교하여, 음원의 신뢰도를 검출할 수 있다. 일 실시예에 따르면, 프로세서(160)는 비트 연산을 통해, 음원의 신뢰도를 검출할 수 있다. 프로세서(160)는 멀티미디어 콘텐츠의 핑거프린트와 검출 구간(200)의 핑거프린트의 비교 연산을 통해, 검출 구간(200)의 비트 에러율(bit error rate; BER)들을 계산할 수 있다. 프로세서(160)는 비트 에러율들을 점수(score)들로 각각 변환할 수 있다. 프로세서(160)는 미리 정해진 스코어 함수(score function)를 이용하여, 비트 에러율들을 점수들로 각각 변환할 수 있다. 프로세서(160)는 점수들의 합으로부터 신뢰도를 검출할 수 있다. 프로세서(160)는 미리 정해진 컨피던스 함수(confidence function)를 이용하여, 점수들의 합으로부터 신뢰도를 검출할 수 있다.

[0027] 다양한 실시예들에 따르면, 프로세서(160)는, 도 4에 도시된 바와 같이 API(application programming interface)(461), 프로세스-API(process-API)(463), 제어부(465), 콘텐츠 획득부(467), 핑거프린트부(469), 매칭부(471), 비교부(473), 또는 클러스터링부(475) 중 적어도 하나를 포함할 수 있다. 어떤 실시예에서, 프로세서(160)의 구성 요소들 중 적어도 어느 하나가 생략될 수 있으며, 적어도 하나의 다른 구성 요소가 추가될 수 있다. 어떤 실시예에서, 프로세서(160)의 구성 요소들 중 적어도 어느 두 개가 하나의 통합된 회로로 구현될 수 있다.

[0028] API(461)는 사용자의 요청을 검출할 수 있다. 프로세스-API(463)는 사용자의 요청에 기반하여, 명령어를 생성할 수 있다. 제어부(465)는 프로세서(160)의 구성 요소들 중 적어도 하나를 제어할 수 있다. 이 때 제어부(465)는 프로세서(160)의 구성 요소들 중 적어도 두 개를 위한 중개 역할을 수행할 수 있으며, 프로세서(160)의 구성 요소들 중 적어도 하나를 위한 작업을 수행할 수 있다. 콘텐츠 획득부(467)는 명령어에 기반하여, 멀티미디어 콘텐츠를 획득할 수 있다. 핑거프린트부(469)는 멀티미디어 콘텐츠의 핑거프린트를 획득할 수 있다. 이 때 핑거프린트부(469)는 멀티미디어 콘텐츠의 오디오 데이터로부터 핑거프린트를 직접적으로 추출할 수 있다. 매칭부(471)는 멀티미디어 콘텐츠의 핑거프린트에 기반하여, 적어도 하나의 음원을 검출할 수 있다. 이 때 메모리(150)에는, 복수의 음원들이 미리 등록되어 있으며, 등록된 음원들의 핑거프린트들이 각각 저장되어 있을 수 있다. 매칭부(471)는 멀티미디어 콘텐츠의 핑거프린트와 등록된 음원들의 핑거프린트들을 매칭시킴으로써, 등록된 음원들의 핑거프린트들 중 적어도 하나를 검출할 수 있다. 비교부(473)는 멀티미디어 콘텐츠의 핑거프린트와 검출된 음원의 핑거프린트를 비교하여, 검출된 음원의 신뢰도를 검출할 수 있다. 클러스터링부(475)는 검출된 음원을 기반으로, 멀티미디어 콘텐츠에 대한 비교 대상 또는 멀티미디어 콘텐츠와의 비교 결과 중 적어도 하나를

검출된 음원과 동일하거나 유사한 음원을 포괄하도록 확장시킬 수 있다. 구체적으로, 클러스터링부(475)는 검출된 음원과 동일하거나 유사한 음원의 정보를 획득하여, 멀티미디어 콘텐츠에 대한 비교 대상을 검출된 음원과 동일하거나 유사한 음원으로 확장시킬 수 있다. 한편, 클러스터링부(475)는 비교부(473)의 비교 결과에 기반하여, 검출된 음원과 동일하거나 유사한 음원을 취합할 수 있다.

[0030] 도 5는 다양한 실시예들에 따른 전자 장치(100)의 동작 방법을 도시하는 도면이다. 도 6은 도 5의 음원의 신뢰도 검출 단계(550 단계)를 세부적으로 도시하는 도면이다. 도 7, 도 8, 도 9, 도 10, 도 11, 도 12, 도 13, 도 14, 도 15, 도 16, 및 도 17은 다양한 실시예들에 따른 전자 장치(100)의 동작 방법을 예시적으로 설명하기 위한 도면들이다.

[0031] 도 5를 참조하면, 전자 장치(100)는 510 단계에서 멀티미디어 콘텐츠의 핑거프린트(710)를 복수의 검색 구간(720)들로 분할할 수 있다. 여기서, 멀티미디어 콘텐츠는 영상 데이터 또는 오디오 데이터 중 적어도 하나로 이루어질 수 있다. 일 예로, 멀티미디어 콘텐츠는 영상 데이터와 오디오 데이터로 이루어지며, 뮤직 비디오, 네트워크를 통해 공유되는 동영상 등을 포함할 수 있다. 다른 예로, 멀티미디어 콘텐츠는 오디오 데이터로 이루어지며, 팟캐스트, 방송국 등에서 생성될 수 있다. 그리고, 오디오 데이터에는, 적어도 하나의 음원이 사용될 수 있으며, 각 음원의 적어도 일부가 포함될 수 있다. 프로세서(160)는 멀티미디어 콘텐츠의 핑거프린트(710)를 획득할 수 있다. 일 실시예에 따르면, 프로세서(160)는 멀티미디어 콘텐츠의 오디오 데이터로부터 핑거프린트(710)를 직접적으로 추출할 수 있다. 예를 들면, 사용자에 의해 멀티미디어 콘텐츠가 선택되면, 프로세서(160)는 멀티미디어 콘텐츠의 오디오 데이터로부터 핑거프린트(710)를 추출할 수 있다. 다른 실시예에 따르면, 프로세서(160)는 외부 장치(102, 104)로부터 멀티미디어 콘텐츠의 핑거프린트(710)를 쿼리로서 수신할 수 있다. 여기서, 핑거프린트는 오디오 데이터에 대한 시간에 따른 주파수 분포를 나타낼 수 있다. 프로세서(160)는, 도 7에 도시된 바와 같이 멀티미디어의 핑거프린트(710)를 미리 설정된 시간 간격에 따라 복수의 검색 구간(720)들로 분할할 수 있다. 일 예로, 시간 간격은 수 초의 범위 내에서 정해질 수 있다.

[0032] 전자 장치(100)는 520 단계에서 검색 구간(720)들 중 적어도 하나가 매칭되는 검출 구간(1010)을 갖는 적어도 하나의 음원을 검출할 수 있다. 이 때 메모리(150)에는, 복수의 음원들이 미리 등록되어 있으며, 등록된 음원들의 핑거프린트(910)들이 각각 저장되어 있을 수 있다. 프로세서(160)는, 도 8에 도시된 바와 같이 검색 구간(720)들의 각각을 등록된 음원들의 핑거프린트(910)들과 비교할 수 있다. 이를 통해, 프로세서(160)는 검색 구간(720)들 중 하나에 기반하여, 등록된 음원들의 핑거프린트(910)들 중 적어도 하나를 검출할 수 있다. 이 때 프로세서(160)는, 도 9에 도시된 바와 같이 검색 구간(720)들 중 하나로부터 시간 범위를 확장시키면서, 멀티미디어 콘텐츠의 핑거프린트(710) 및 검출된 음원의 핑거프린트(910)를 비교할 수 있다. 이에 따라, 프로세서(160)는, 도 10에 도시된 바와 같이 검출된 음원의 핑거프린트(910)에서, 검색 구간(720)들 중 적어도 하나가 매칭되는 검출 구간(1010)을 검출할 수 있다.

[0033] 전자 장치(100)는 530 단계에서 멀티미디어 콘텐츠 및 검출된 음원 내에서의 검출 구간(1010)의 위치 정보를 결정할 수 있다. 위치 정보는 멀티미디어 콘텐츠의 핑거프린트(710) 내에서의 검출 구간(1010)의 시간 위치 및 검출된 음원의 핑거프린트(910) 내에서의 검출 구간(1010)의 시간 위치를 나타낼 수 있다. 프로세서(160)는 검출 구간(1010)의 위치 정보에 기반하여, 검출 구간(101)에 대한 오프셋 차이( $\Delta T_m - \Delta T_a$ )를 검출할 수 있다. 오프셋 차이( $\Delta T_m - \Delta T_a$ )는 멀티미디어 콘텐츠의 핑거프린트(710)의 시작점( $T_{m0}$ )으로부터 검출 구간(1010)의 시작점( $T_{a0}$ )까지의 시간 오프셋( $\Delta T_m$ )과 검출된 음원의 핑거프린트(910)의 시작점( $T_{a0}$ )으로부터 검출 구간(1010)의 시작점( $T_{a0}$ )까지의 시간 오프셋( $\Delta T_a$ ) 사이의 차이를 나타낼 수 있다.

[0034] 전자 장치(100)는 540 단계에서 검출 구간(1010)에 대한 오프셋 차이( $\Delta T_m - \Delta T_a$ )에 기반하여, 멀티미디어 콘텐츠의 핑거프린트(710)에 대해 검출 구간(1010)을 정렬시킬 수 있다. 프로세서(160)는, 도 10에 도시된 바와 같이 검출 구간(1010)과 검출 구간(1010)에 매칭된 적어도 하나의 검색 구간(720)이 서로에 대응되도록, 검출 구간(1010)을 정렬시킬 수 있다. 프로세서(160)는 적어도 하나의 검색 구간(720)의 시작점에 검출 구간(1010)의 시작점을 정렬시킬 수 있다.

[0035] 전자 장치(100)는 550 단계에서 멀티미디어 콘텐츠의 핑거프린트(710)와 검출 구간(1010)을 비교하여, 검출된 음원의 신뢰도를 검출할 수 있다. 신뢰도는 검출된 음원이 멀티미디어 콘텐츠에 사용된 것인지에 대한 정확도를 나타내는 것으로, 신뢰도가 높을수록, 정확도가 높을 수 있다. 프로세서(160)는 검출 구간(1010)과 검출 구간(1010)에 매칭된 적어도 하나의 검색 구간(720)을 비교하여, 검출된 음원의 신뢰도를 검출할 수 있다. 일 실시예에 따르면, 프로세서(160)는 적어도 하나의 검색 구간(720)에 대한 검출 구간(1010)의 비트 연산을 통해,

검출된 음원의 신뢰도를 검출할 수 있다. 이에 대해, 도 6을 참조하여, 보다 상세하게 후술될 것이다.

- [0036] 도 6을 참조하면, 전자 장치(100)는 651 단계에서 멀티미디어 콘텐츠의 핑거프린트(710)와 검출 구간(1010)의 비교 연산을 통해, 비교 구간(1110)을 생성할 수 있다. 일 예로, 비교 연산은 배타적 논리합(XOR)을 포함할 수 있다. 프로세서(160)는 검출 구간(1010)과 검출 구간(1010)에 매칭된 적어도 하나의 검색 구간(720)의 비교 연산을 통해, 도 11에 도시된 바와 같이 비교 구간(1110)을 생성할 수 있다.
- [0037] 전자 장치(100)는 653 단계에서 비교 구간(1110)을 복수의 비트 구간(1210)들로 분할할 수 있다. 프로세서(160)는, 도 12에 도시된 바와 같이 비교 구간(1110)을 미리 설정된 시간 간격에 따라 복수의 비트 구간(1210)들로 분할할 수 있다. 일 예로, 시간 간격은 약 1 초일 수 있다.
- [0038] 전자 장치(100)는 655 단계에서 비트 구간(1210)들의 비트 에러율들을 각각 계산할 수 있다. 프로세서(160)는 비트 구간(1210)들의 각각에 대해 연속되는 비트들로 계산하고, 비트 구간(1210)들의 각각의 비트들로부터 비트 에러율을 계산할 수 있다. 여기서, 각 비트 에러율은 0과 1 사이의 값으로 표현되며, 비트 에러율이 낮을수록 유사성이 높을 수 있다. 유사성은 검출 구간(1010)과 검출 구간(1010)에 매칭된 적어도 하나의 검색 구간(720) 사이의 유사성을 나타낼 수 있다. 즉, 비트 에러율이 0이라는 것은, 검출 구간(1010)과 적어도 하나의 검색 구간(720)이 동일함을 의미할 수 있다. 일 예로, 도 13에 도시된 바와 같이 비트 구간(1210)들의 비트 에러율들이 계산된 경우, 이는 멀티미디어 콘텐츠에서 513 초에서부터 551 초까지에 검출된 음원의 검출 구간(1010)이 사용되었음을 나타낼 수 있다. 다른 예로, 도 14에 도시된 바와 같이 비트 구간(1210)들의 비트 에러율들이 계산된 경우, 이는 멀티미디어 콘텐츠에서 복수의 음원들의 검출 구간(1010)들이 사용되었으며, 동일한 시간 범위에도 복수의 음원들의 검출 구간(1010)들이 사용되었음을 나타낼 수 있다. 이러한 경우, 프로세서(160)는 검출 구간(1010)들의 비트 에러율들 사이에서 가장 높은 비트 에러율들을 추출할 수 있다.
- [0039] 전자 장치(100)는 657 단계에서 비트 에러율들을 비트 구간(1210)들의 점수들로 각각 변환할 수 있다. 프로세서(160)는 미리 정해진 스코어 함수를 이용하여, 비트 에러율들을 점수들로 각각 변환할 수 있다. 여기서, 도 15에 도시된 바와 같이, 낮은 비트 에러율일수록, 높은 점수로 변환되고, 높은 비트 에러율일수록, 낮은 점수로 변환될 수 있다. 이 때 비트 에러율들 중 적어도 하나가 임계값을 초과하면, 프로세서(160)는 적어도 하나의 비트 에러율에 대응하여 0을 점수로 부여할 수 있다. 한편, 비트 에러율들 중 나머지가 임계값 이하이면, 프로세서(160)는 나머지의 비트 에러율에 기반하여 점수를 각각 계산하고, 나머지의 비트 에러율에 대응하여 계산된 점수를 각각 부여할 수 있다. 예를 들면, 스코어 함수는 하기 [수학식 1]과 같이 표현될 수 있다.

**수학식 1**

$$y = \max(0.0, 10 \cdot ((1-x)^2 - (1.0 - \text{임계값})^2))$$

- [0040]
- [0041] 여기서, x는 비트 에러율을 나타내고, y는 점수를 나타내고, 임계값은 0을 초과하고 0.5 이하일 수 있으며, 일 예로 0.35 이상이고 0.45 이하일 수 있다. 임계값이 0에 가까운 값으로 설정될수록, 음원에 대한 오검출 가능성이 낮은 한편, 노이즈를 갖는 음원에 대한 검출 가능성이 높을 수 있다. 이에 반해, 임계값이 0.5에 가까운 값으로 설정될수록, 노이즈를 갖는 음원에 대한 검출 가능성이 낮은 한편, 음원에 대한 오검출 가능성이 높을 수 있다.
- [0042] 전자 장치(100)는 659 단계에서 점수들의 합으로부터 검출된 음원에 대한 신뢰도를 검출할 수 있다. 신뢰도는 검출된 음원이 멀티미디어 콘텐츠에 사용된 것인 지에 대한 정확도를 나타내는 것으로, 신뢰도가 높을수록, 정확도가 높을 수 있다. 프로세서(160)는 미리 정해진 컨피던스 함수를 이용하여, 점수들의 합으로부터 신뢰도를 검출할 수 있다. 여기서, 도 16에 도시된 바와 같이, 신뢰도는 0과 1 사이의 값으로 표현될 수 있다. 점수들의 합이 일정 범위 내에 있을 때, 점수들의 합이 신뢰도에 크게 영향을 미침으로써, 점수들의 합이 클수록, 신뢰도가 현저하게 높을 수 있다. 한편, 점수들의 합이 일정 범위 밖에 있을 때, 점수들의 합이 신뢰도에 미치는 영향이 감소될 수 있다. 예를 들면, 컨피던스 함수는 하기 [수학식 2]와 같이 표현될 수 있다.

수학식 2

$$y = \tanh(\text{가중치}x)$$

[0043]

[0044]

[0045]

[0046]

[0047]

[0048]

[0050]

여기서,  $x$ 는 점수들의 합을 나타내고,  $y$ 는 신뢰도를 나타내고, 가중치는 스코어 함수의 임계값 또는 후술되는 신뢰도에 대한 기준값 중 적어도 하나에 따라 결정될 수 있으며, 예컨대 0.1 이상이고 0.2 이하일 수 있다.

이 후, 전자 장치(100)는 도 5로 리턴하여, 560 단계로 진행할 수 있다.

다시 도 5를 참조하면, 전자 장치(100)는 560 단계에서 검출된 음원과 관련된 정보, 위치 정보, 및 신뢰도를 제공할 수 있다. 음원과 관련된 정보는 음원의 식별자, 명칭, 또는 아티스트 중 적어도 하나를 포함할 수 있다. 위치 정보는 멀티미디어 콘텐츠의 핑거프린트(710) 내에서의 검출 구간(1010)의 시간 위치 및 검출된 음원의 핑거프린트(910) 내에서의 검출 구간(1010)의 시간 위치를 나타낼 수 있다. 프로세서(160)는, 도 17에 도시된 바와 같이 멀티미디어 콘텐츠에 대응하여 검출된 음원과 관련된 정보, 위치 정보, 및 신뢰도를 제공할 수 있다. 여기서, 멀티미디어 콘텐츠로부터 복수의 음원들이 검출된 경우, 프로세서(160)는 음원들의 리스트로서, 검출된 음원과 관련된 정보, 위치 정보, 및 신뢰도를 제공할 수 있다. 일 예로, 프로세서(160)는 검출된 음원의 신뢰도와 관계 없이, 검출된 음원의 관련된 정보, 위치 정보, 및 신뢰도를 제공할 수 있다. 다른 예로, 검출된 음원의 신뢰도가 기준값 이상이면, 프로세서(160)가 검출된 음원의 관련된 정보, 위치 정보, 및 신뢰도를 제공할 수 있다. 바꿔 말하면, 검출된 음원의 신뢰도가 기준값 미만이면, 프로세서(160)는 검출된 음원의 관련된 정보, 위치 정보, 및 신뢰도를 제공하지 않을 수 있다. 프로세서(160)는 외부 장치(102, 104)의 쿼리에 대한 응답으로서, 검출된 음원과 관련된 정보, 위치 정보, 및 신뢰도를 제공할 수 있다. 일 실시예에 따르면, 프로세서(160)는 외부 장치(102, 104)로 검출된 음원과 관련된 정보, 위치 정보, 및 신뢰도를 송신할 수 있다. 다른 실시예에 따르면, 프로세서(160)는 출력 모듈(140)을 통해, 검출된 음원과 관련된 정보, 위치 정보, 및 신뢰도를 직접적으로 출력할 수 있다.

다양한 실시예들에 따르면, 사용자는 멀티미디어 콘텐츠에 사용된 음원을 확인하고, 이를 다양하게 활용할 수 있다. 일 예로, 멀티미디어 콘텐츠가 방송이나 공연의 동영상인 경우, 사용자는 멀티미디어 콘텐츠에 사용된 음원에 기반하여, 멀티미디어 콘텐츠의 큐시트(cue sheet)를 획득할 수 있다. 다른 예로, 사용자는 멀티미디어 콘텐츠에 사용된 음원의 저작권 보호 또는 저작권 정산을 위해 활용할 수 있다.

다양한 실시예들에 따르면, 검출된 음원의 관련된 정보, 위치 정보, 및 신뢰도를 제공한 후에, 전자 장치(100)는 검출된 음원과 연관된 다양한 서비스들을 제공할 수 있다. 일 실시예에 따르면, 프로세서(160)는 외부 장치(102, 104)로 검출된 음원을 제공할 수 있다. 외부 장치(102, 104)에 의해 검출된 음원의 관련된 정보가 선택되면, 프로세서(160)가 외부 장치(102, 104)로 검출된 음원을 제공할 수 있다. 다른 실시예에 따르면, 프로세서(160)는 검출된 음원과 연관된 다른 멀티미디어 콘텐츠를 제공할 수 있다. 외부 장치(102, 104)에 의해 검출된 음원의 관련된 정보가 선택되면, 프로세서(160)는 검출된 음원과 관련된 정보에 기반하여, 다른 멀티미디어 콘텐츠를 검색하고, 외부 장치(102, 104)로 검색된 멀티미디어 콘텐츠를 제공할 수 있다. 또 다른 실시예에 따르면, 프로세서(160)는 검출된 음원과 연관된 부가 정보를 제공할 수 있다. 외부 장치(102, 104)에 의해 검출된 음원의 관련된 정보가 선택되면, 프로세서(160)는 검출된 음원과 관련된 정보에 기반하여, 예컨대 뉴스, 소셜 네트워크 서비스(social network service; SNS) 등을 통해 부가 정보를 검색하고, 외부 장치(102, 104)로 검색된 부가 정보를 제공할 수 있다.

다양한 실시예들에 따르면, 전자 장치(100)는 멀티미디어 콘텐츠에 사용된 적어도 하나의 음원을 효율적으로 검출할 수 있다. 구체적으로, 전자 장치(100)는 멀티미디어 콘텐츠의 핑거프린트(710)에서 검색 구간(720)들 중 하나로부터 시간 범위를 확장시키면서, 음원 내에서 멀티미디어 콘텐츠에 매칭되는 검출 구간(1010)을 효율적으로 검출할 수 있다. 그리고, 전자 장치(100)는 음원 내에서의 검출 구간(1010)의 시간 위치뿐만 아니라 멀티미디어 콘텐츠 내에서의 검출 구간(1010)의 시간 위치를 검출함으로써, 음원 및 멀티미디어 콘텐츠 내에서 검출 구간(1010)을 보다 정확하게 특정할 수 있다. 아울러, 전자 장치(100)는 검출 구간(1010)에 대한 멀티미디어 콘텐츠의 시작점( $T_{m0}$ )으로부터 시간 오프셋( $\Delta T_m$ )과 음원의 시작점( $T_{a0}$ )으로부터의 시간 오프셋( $\Delta T_a$ ) 사이의 오프셋 차이( $\Delta T_m - \Delta T_a$ )에 기반하여 멀티미디어 콘텐츠와 음원을 비교함으로써, 음원에 대한 신뢰도를 검출할 수 있다. 이를 통해, 전자 장치(100)는 사용자를 위해, 음원과 관련된 정보와 위치 정보뿐만 아니라, 신뢰도를 제공

할 수 있다.

- [0052] 다양한 실시예들에 따른 전자 장치(100)의 동작 방법은, 멀티미디어 콘텐츠의 핑거프린트(710)를 미리 설정된 시간 간격에 따라 복수의 검색 구간(720)들로 분할하는 단계(510 단계), 검색 구간(720)들 중 적어도 하나가 매칭되는 검출 구간(1010)을 갖는 적어도 하나의 음원을 검출하는 단계(520 단계), 멀티미디어 콘텐츠 내에서의 검출 구간(1010)의 시간 위치 및 음원 내에서의 검출 구간(1010)의 시간 위치를 나타내는 위치 정보를 결정하는 단계(530 단계), 및 음원과 관련된 정보 및 위치 정보를 제공하는 단계(560 단계)를 포함할 수 있다.
- [0053] 다양한 실시예들에 따르면, 전자 장치(100)의 동작 방법은, 멀티미디어 콘텐츠의 시작점( $T_{m0}$ )으로부터의 시간 오프셋( $\Delta T_m$ )과 음원의 시작점( $T_{a0}$ )으로부터의 시간 오프셋( $\Delta T_a$ ) 사이의 오프셋 차이( $\Delta T_m - \Delta T_a$ )에 기반하여, 핑거프린트(710)에 대해 검출 구간(1010)을 정렬시키는 단계(540 단계), 및 핑거프린트(710)와 검출 구간(1010)을 비교하여, 음원의 신뢰도를 검출하는 단계(550 단계)를 더 포함할 수 있다.
- [0054] 다양한 실시예들에 따르면, 음원과 관련된 정보 및 위치 정보를 제공하는 단계(560 단계)는, 음원과 관련된 정보 및 위치 정보와 함께, 신뢰도를 제공하는 단계를 포함할 수 있다.
- [0055] 다양한 실시예들에 따르면, 신뢰도를 검출하는 단계(550 단계)는, 핑거프린트(710)와 검출 구간(1010)의 비교 연산을 통해, 비교 구간(1110)을 생성하는 단계(651 단계), 비교 구간(1110)을 복수의 비트 구간(1210)들로 분할하는 단계(653 단계), 비트 구간(1210)들의 비트 에러율들을 각각 계산하는 단계(655 단계), 비트 에러율들을 비트 구간(1210)들의 점수들로 각각 변환하는 단계(657 단계), 및 점수들의 합으로부터 신뢰도를 검출하는 단계(659 단계)를 포함할 수 있다.
- [0056] 다양한 실시예들에 따르면, 점수들로 변환하는 단계(657 단계)는, 비트 에러율들 중 적어도 하나가 임계값을 초과하면, 적어도 하나의 비트 에러율에 대응하여 0을 점수로 부여하는 단계, 및 비트 에러율들 중 나머지가 임계값 이하이면, 나머지의 비트 에러율에 대응하여 계산되는 점수를 부여하는 단계를 포함할 수 있다.
- [0057] 다양한 실시예들에 따르면, 음원과 관련된 정보 및 위치 정보를 제공하는 단계(560 단계)는, 멀티미디어 콘텐츠가 복수의 음원들과 연관되면, 음원과 관련된 정보 및 위치 정보를 리스트로 제공하는 단계를 포함할 수 있다.
- [0058] 다양한 실시예들에 따르면, 멀티미디어 콘텐츠는, 영상 데이터 또는 오디오 데이터 중 적어도 하나로 이루어질 수 있다.
- [0059] 다양한 실시예들에 따르면, 전자 장치(100)의 동작 방법은, 음원과 관련된 정보가 선택되면, 음원을 제공하는 단계, 또는 음원과 관련된 정보가 선택되면, 음원과 연관된 다른 멀티미디어 콘텐츠를 제공하는 단계 중 적어도 하나를 더 포함할 수 있다.
- [0060] 다양한 실시예들에 따르면, 검색 구간(720)들 중 적어도 하나는, 복수의 음원들에 매칭될 수 있다.
- [0062] 다양한 실시예들에 따른 전자 장치(100)는, 메모리(150), 및 메모리(150)와 연결되고, 메모리(150)에 저장된 적어도 하나의 명령을 실행하도록 구성된 프로세서(160)를 포함할 수 있다.
- [0063] 다양한 실시예들에 따르면, 프로세서(160)는, 멀티미디어 콘텐츠의 핑거프린트(710)를 미리 설정된 시간 간격에 따라 복수의 검색 구간(720)들로 분할하고, 검색 구간(720)들 중 적어도 하나가 매칭되는 검출 구간(1010)을 갖는 적어도 하나의 음원을 검출하고, 멀티미디어 콘텐츠 내에서의 검출 구간(1010)의 시간 위치 및 음원 내에서의 검출 구간(1010)의 시간 위치를 나타내는 위치 정보를 결정하고, 음원과 관련된 정보 및 위치 정보를 제공하도록 구성될 수 있다.
- [0064] 다양한 실시예들에 따르면, 프로세서(160)는, 멀티미디어 콘텐츠의 시작점( $T_{m0}$ )으로부터의 시간 오프셋( $\Delta T_m$ )과 음원의 시작점( $T_{a0}$ )으로부터의 시간 오프셋( $\Delta T_a$ ) 사이의 오프셋 차이( $\Delta T_m - \Delta T_a$ )에 기반하여, 핑거프린트(710)에 대해 검출 구간(1010)을 정렬시키고, 핑거프린트(710)와 검출 구간(1010)을 비교하여, 음원의 신뢰도를 검출하도록 구성될 수 있다.
- [0065] 다양한 실시예들에 따르면, 프로세서(160)는, 검출 구간(1010)에 대응하는 음원의 신뢰도를 검출하고, 음원과 관련된 정보 및 위치 정보와 함께, 신뢰도를 제공하도록 구성될 수 있다.
- [0066] 다양한 실시예들에 따르면, 프로세서(160)는, 핑거프린트(710)와 검출 구간(1010)의 비교 연산을 통해, 비교 구간(1110)을 생성하고, 비교 구간(1110)을 복수의 비트 구간(1210)들로 분할하고, 비트 구간(1210)들의 비트 에러율들을 각각 계산하고, 비트 에러율들을 비트 구간(1210)들의 점수들로 각각 변환하고, 점수들의 합으로부터

신뢰도를 검출하도록 구성될 수 있다.

- [0067] 다양한 실시예들에 따르면, 프로세서(160)는, 비트 에러율들 중 적어도 하나가 임계값을 초과하면, 적어도 하나의 비트 에러율에 대응하여 0을 점수로 부여하고, 비트 에러율들 중 나머지가 임계값 이하이면, 나머지의 비트 에러율에 대응하여 계산되는 점수를 부여하도록 구성될 수 있다.
- [0068] 다양한 실시예들에 따르면, 프로세서(160)는, 멀티미디어 콘텐츠가 복수의 음원들과 연관되면, 음원과 관련된 정보 및 위치 정보를 리스트로 제공하도록 구성될 수 있다.
- [0069] 다양한 실시예들에 따르면, 멀티미디어 콘텐츠는, 영상 데이터 또는 오디오 데이터 중 적어도 하나로 이루어질 수 있다.
- [0070] 다양한 실시예들에 따르면, 프로세서(160)는, 음원과 관련된 정보가 선택되면, 음원 또는 음원과 연관된 다른 멀티미디어 콘텐츠 중 적어도 하나를 제공하도록 구성될 수 있다.
- [0071] 다양한 실시예들에 따르면, 검색 구간(720)들 중 적어도 하나는, 복수의 음원들에 매칭될 수 있다.
- [0073] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 프로세서, 컨트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 어플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.
- [0074] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 컴퓨터 저장 매체 또는 장치에 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.
- [0075] 다양한 실시예들에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 이 때 매체는 컴퓨터로 실행 가능한 프로그램을 계속 저장하거나, 실행 또는 다운로드를 위해 임시 저장하는 것일 수도 있다. 그리고, 매체는 단일 또는 수 개의 하드웨어가 결합된 형태의 다양한 기록수단 또는 저장수단일 수 있는데, 어떤 컴퓨터 시스템에 직접 접속되는 매체에 한정되지 않고, 네트워크 상에 분산 존재하는 것일 수도 있다. 매체의 예시로는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM 및 DVD와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical medium), 및 ROM, RAM, 플래시 메모리 등을 포함하여 프로그램 명령어가 저장되도록 구성된 것이 있을 수 있다. 또한, 다른 매체의 예시로, 어플리케이션을 유통하는 앱 스토어나 기타 다양한 소프트웨어를 공급 내지 유통하는 사이트, 서버 등에서 관리하는 기록매체 내지 저장매체도 들 수 있다.
- [0077] 본 문서의 다양한 실시예들 및 이에 사용된 용어들은 본 문서에 기재된 기술을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 해당 실시 예의 다양한 변경, 균등물, 및/또는 대체물을 포함하는 것으로 이해되어야 한다. 도면의 설명과 관련하여, 유사한 구성 요소에 대해서는 유사한 참조 부호가 사용될 수 있다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함할 수 있다. 본 문서에서, "A 또는 B", "A 및/또는 B 중 적어도 하나", "A, B 또는 C" 또는 "A, B 및/또는 C 중 적어도 하나" 등의 표현은 함께 나열된 항목들의 모든 가능한 조합을 포함할 수 있다. "제 1", "제 2", "첫째" 또는 "둘째" 등의 표현들은 해당 구성 요소들을, 순서 또는 중요도에 상관없이 수식할 수 있고, 한 구성 요소를 다른 구성 요소와 구분하기 위해 사용될 뿐 해당 구성 요소들을 한정하지 않는다. 어떤(예: 제 1) 구성 요소가 다른(예: 제 2) 구성 요소에 "(기능적으로 또는

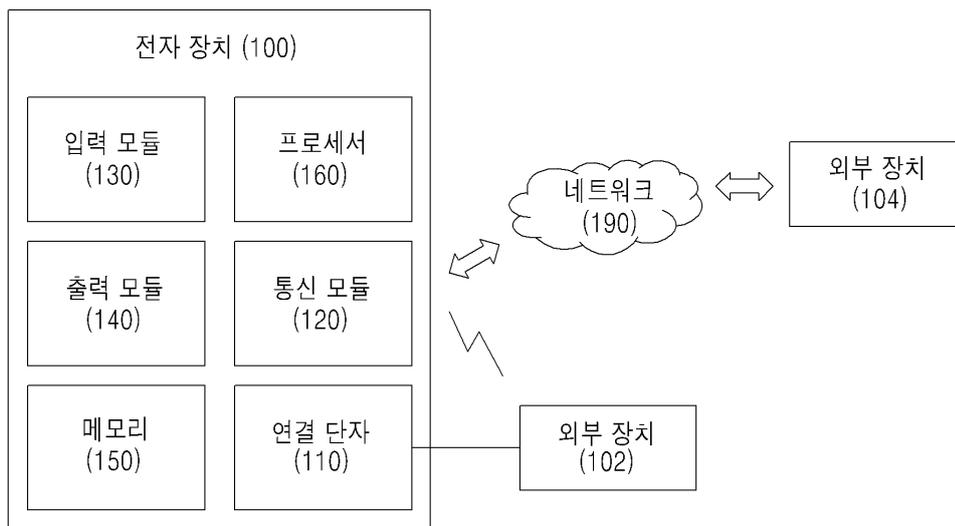
통신적으로) 연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 상기 어떤 구성 요소가 상기 다른 구성 요소에 직접적으로 연결되거나, 다른 구성 요소(예: 제 3 구성 요소)를 통하여 연결될 수 있다.

[0078] 본 문서에서 사용된 용어 "모듈"은 하드웨어, 소프트웨어 또는 펌웨어로 구성된 유닛을 포함하며, 예를 들면, 로직, 논리 블록, 부품, 또는 회로 등의 용어와 상호 호환적으로 사용될 수 있다. 모듈은, 일체로 구성된 부품 또는 하나 또는 그 이상의 기능을 수행하는 최소 단위 또는 그 일부가 될 수 있다. 예를 들면, 모듈은 ASIC(application-specific integrated circuit)으로 구성될 수 있다.

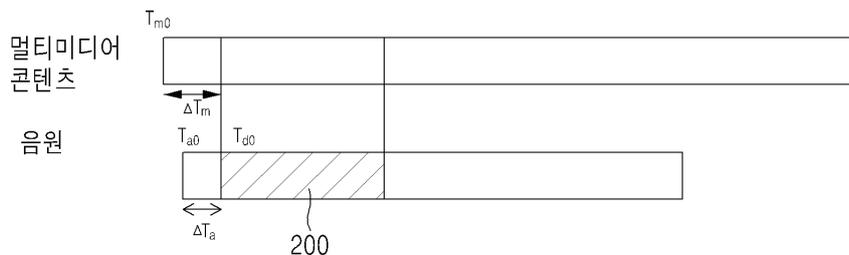
[0079] 다양한 실시예들에 따르면, 기술한 구성 요소들의 각각의 구성 요소(예: 모듈 또는 프로그램)는 단수 또는 복수의 개체를 포함할 수 있다. 다양한 실시예들에 따르면, 전술한 해당 구성 요소들 중 하나 이상의 구성 요소들 또는 단계들이 생략되거나, 또는 하나 이상의 다른 구성 요소들 또는 단계들이 추가될 수 있다. 대체적으로 또는 추가적으로, 복수의 구성 요소들(예: 모듈 또는 프로그램)은 하나의 구성 요소로 통합될 수 있다. 이런 경우, 통합된 구성 요소는 복수의 구성 요소들 각각의 구성 요소의 하나 이상의 기능들을 통합 이전에 복수의 구성 요소들 중 해당 구성 요소에 의해 수행되는 것과 동일 또는 유사하게 수행할 수 있다. 다양한 실시예들에 따르면, 모듈, 프로그램 또는 다른 구성 요소에 의해 수행되는 단계들은 순차적으로, 병렬적으로, 반복적으로, 또는 휴리스틱하게 실행되거나, 단계들 중 하나 이상이 다른 순서로 실행되거나, 생략되거나, 또는 하나 이상의 다른 단계들이 추가될 수 있다.

도면

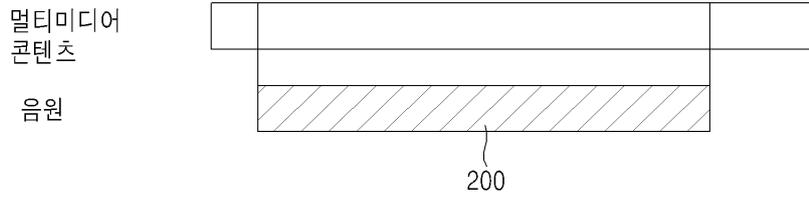
도면1



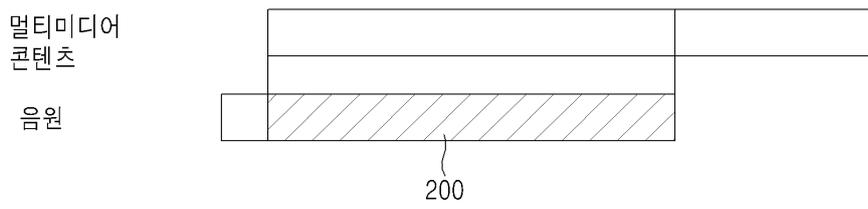
도면2



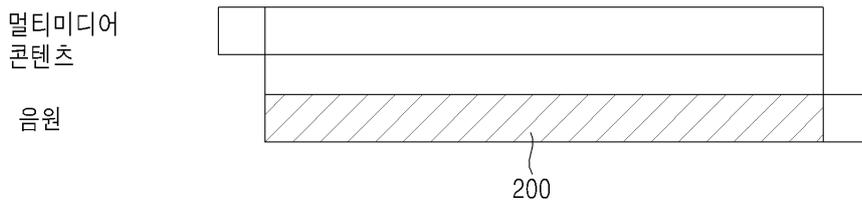
도면3a



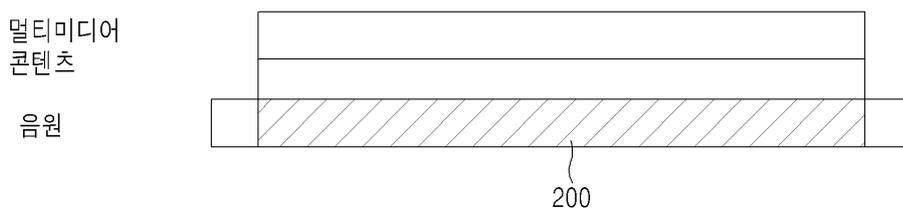
도면3b



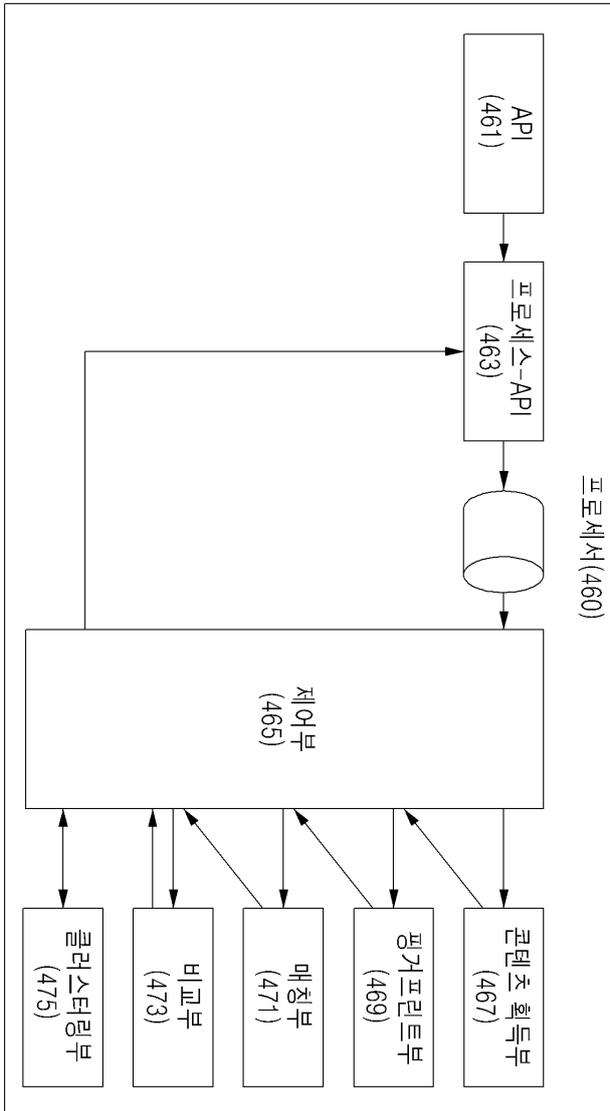
도면3c



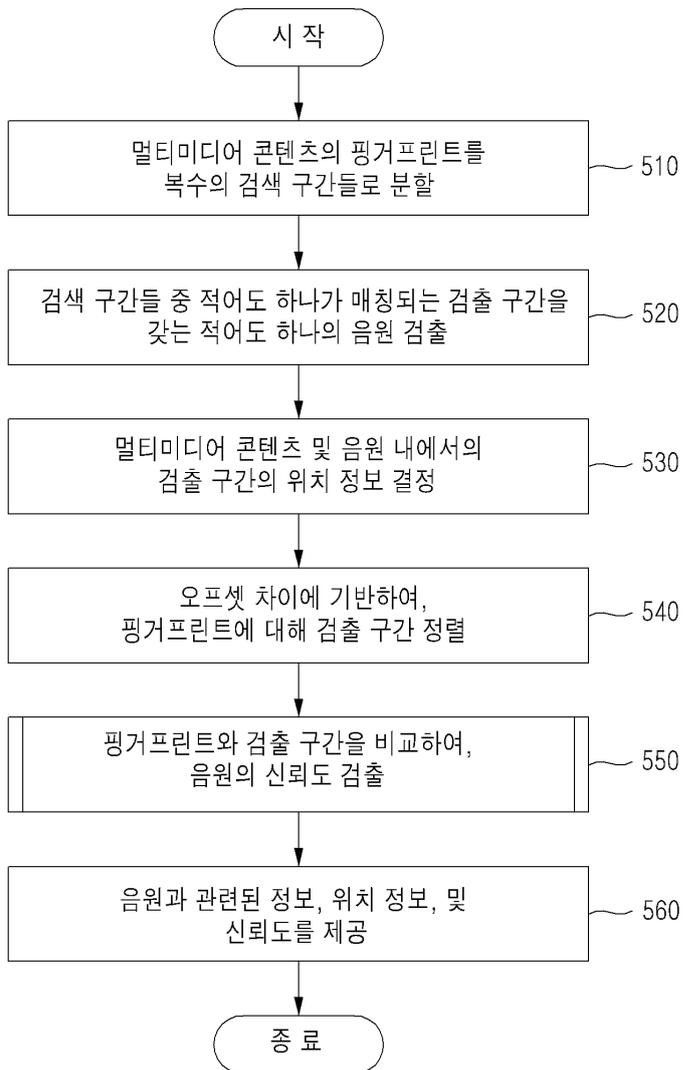
도면3d



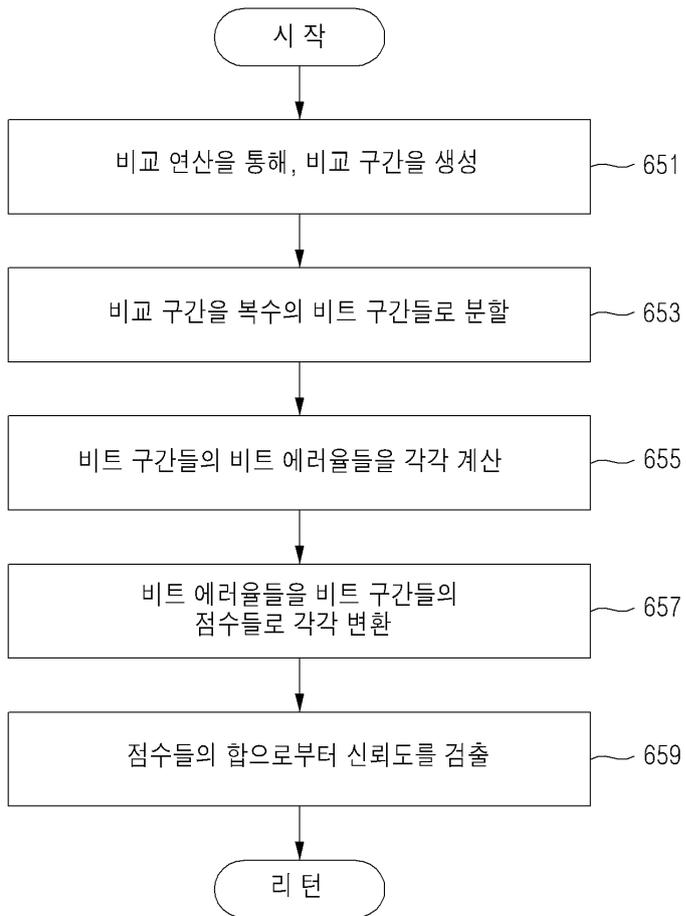
도면4



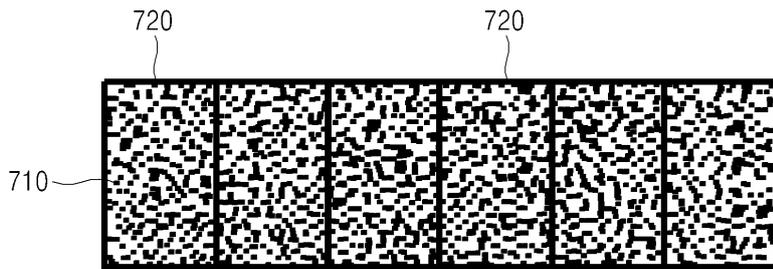
도면5



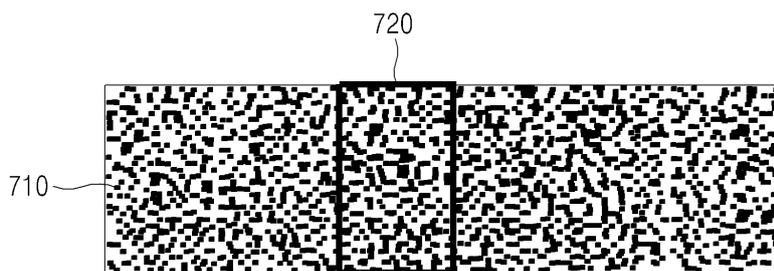
도면6



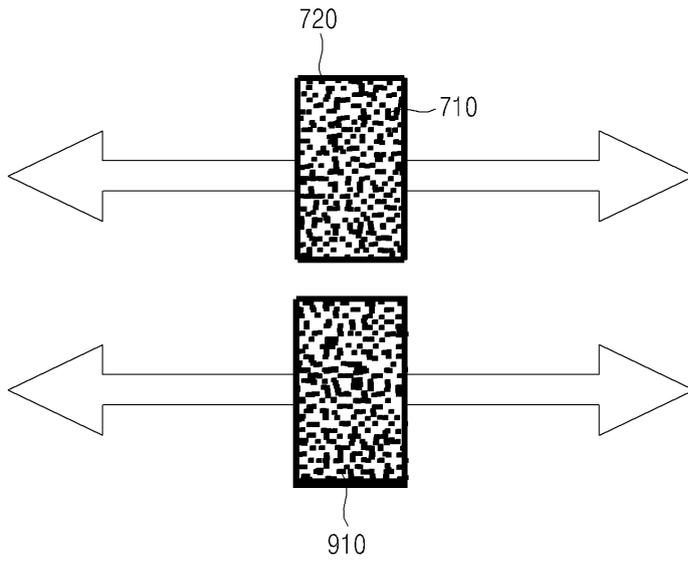
도면7



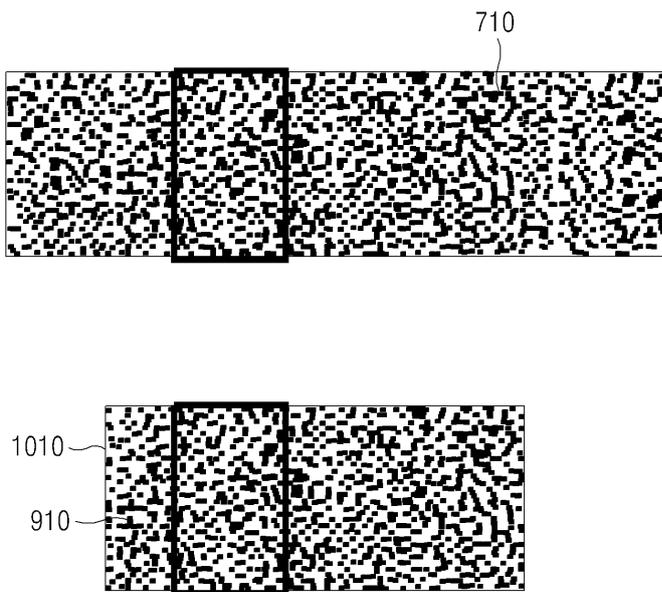
도면8



도면9



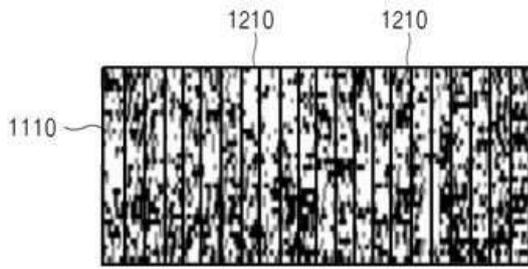
도면10



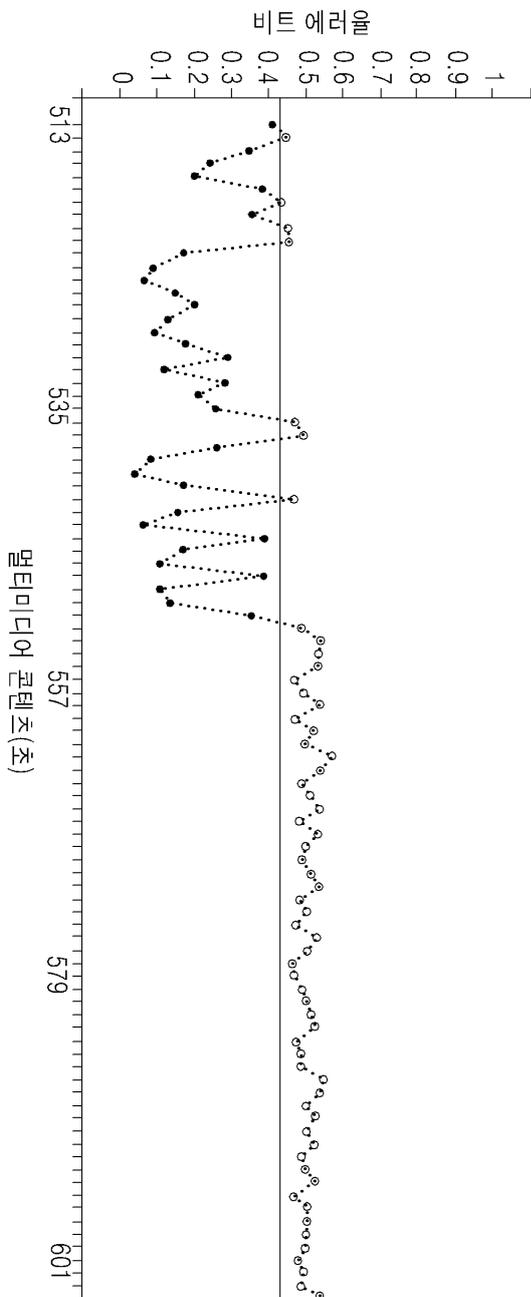
도면11



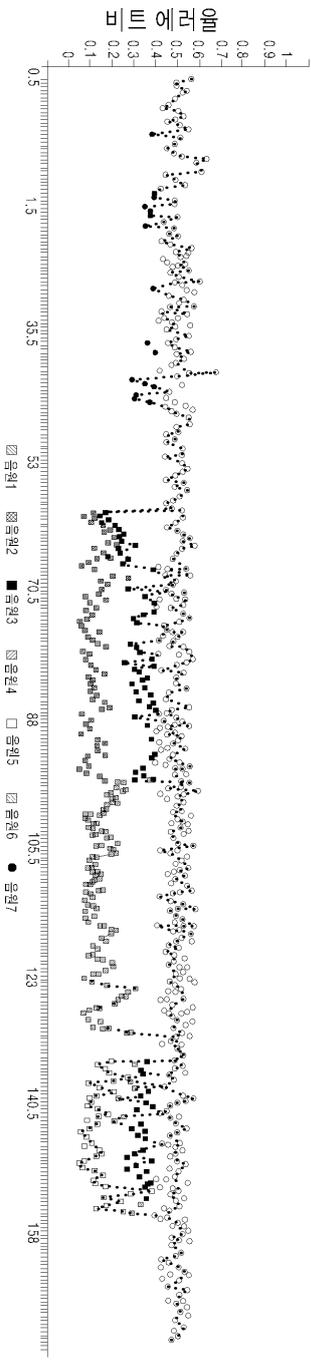
도면12



도면13

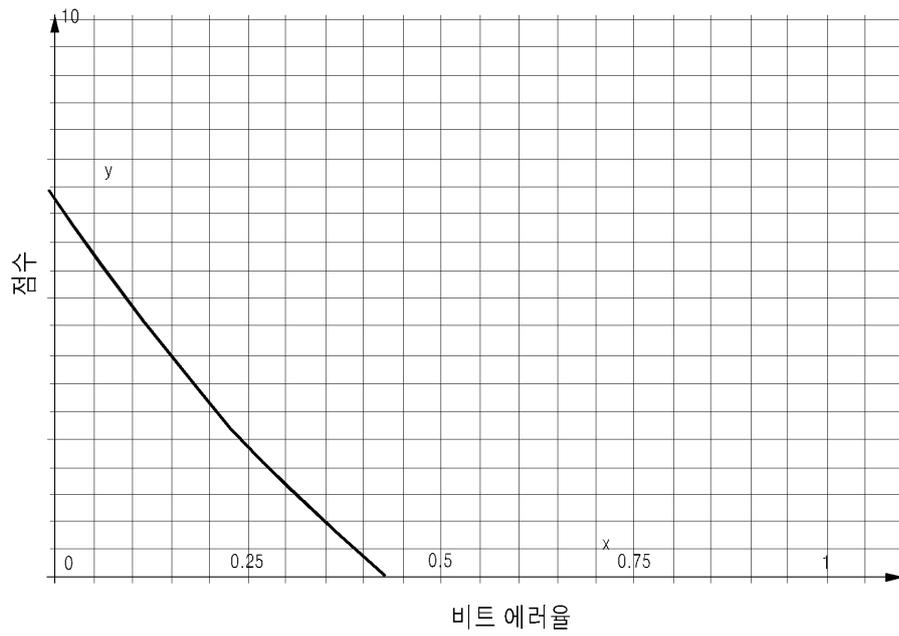


도면14

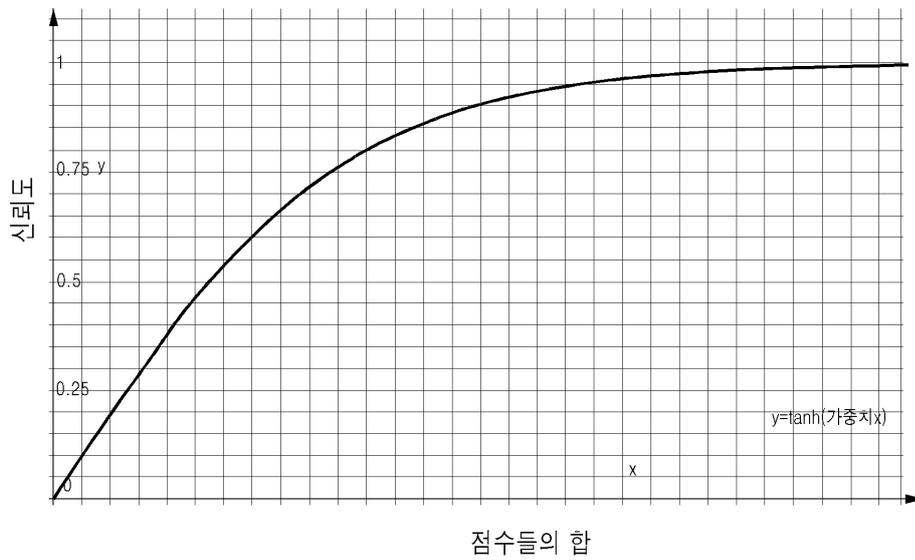


멀티미디어 콘텐츠(초)

도면15



도면16



도면17

| # | Video Range         | Music Range         | Music meta       | Confidence |
|---|---------------------|---------------------|------------------|------------|
| 1 | 00:00:27 ~ 00:00:35 | 00:02:33 ~ 00:02:41 | 곡 'A' - 아티스트 '가' | 0.749075   |
| 2 | 00:00:47 ~ 00:00:55 | 00:02:53 ~ 00:03:01 | 곡 'A' - 아티스트 '가' | 0.335756   |
| 3 | 00:06:45 ~ 00:07:20 | 00:00:09 ~ 00:00:44 | 곡 'B' - 아티스트 '나' | 0.999999   |
| 4 | 00:08:33 ~ 00:09:11 | 00:00:00 ~ 00:00:38 | 곡 'C' - 아티스트 '다' | 0.999999   |
| 5 | 00:12:03 ~ 00:12:21 | 00:02:34 ~ 00:02:52 | 곡 'B' - 아티스트 '나' | 0.999999   |