

(19)대한민국특허청(KR)  
(12) 공개특허공보(A)

(51) Int. Cl.<sup>7</sup>  
G06F 3/16

(11) 공개번호 10-2005-0074920  
(43) 공개일자 2005년07월19일

(21) 출원번호 10-2005-0003281  
(22) 출원일자 2005년01월13일

(30) 우선권주장 JP-P-2004-00006456 2004년01월14일 일본(JP)

(71) 출원인 소니 가부시끼 가이샤  
일본국 도쿄도 시나가와쿠 기타시나가와 6쵸메 7반 35고

(72) 발명자 다나카이즈루  
일본 도쿄도 시나가와쿠 기따시나가와 6쵸메 7-35 소니 가부시끼 가이샤 내  
이이다겐이찌  
일본 도쿄도 시나가와쿠 기따시나가와 6쵸메 7-35 소니 가부시끼 가이샤 내  
미하라사또시  
일본 도쿄도 시나가와쿠 기따시나가와 6쵸메 7-35 소니 가부시끼 가이샤 내  
야마다에이찌  
일본 도쿄도 시나가와쿠 기따시나가와 6쵸메 7-35 소니 가부시끼 가이샤 내

(74) 대리인 장수길  
이중희  
구영창

심사청구 : 없음

(54) 음성 신호 처리 장치 및 음성 신호 처리 방법

요약

유저의 손을 번거롭게 하지 않고, 처리 대상의 음성 데이터 중의 목적으로 하는 부분을 신속하게 찾아내어 이용할 수 있도록 한다. 음성 특징 해석부(143)에 의해 처리 대상의 음성 신호의 변화점을 자동적으로 검출하고, 그 변화점의 해당 음성 신호 상의 위치를 나타내는 변화점 정보를 CPU(101)가 취득하여, 이 변화점 정보를 데이터 기억 장치(111)에 기록한다. CPU(101)에 의해, 키 조작부(121)를 통하여 접수한 유저로부터의 지시 입력에 따른 변화점 정보를 특정하고, 그 특정한 변화점 정보에 대응하는 음성 데이터에 위치 결정이 행하여지며, 그 위치로부터 처리 대상의 음성 데이터의 재생 등의 처리를 행할 수 있도록 한다.

대표도

도 1

색인어

음성 데이터, 음성 특징 해석부, 변화점 정보, 음성 신호, 데이터 기억 장치, CPU

명세서

도면의 간단한 설명

도 1은 본 발명이 적용되어 구성된 기록 재생 장치의 일례를 설명하기 위한 블록도.

도 2는, 도 1에 도시한 기록 재생 장치에서, 짐음하여 녹음하는 음성 신호의 변화점에 마크를 붙이도록 하는 처리의 개요를 설명하기 위한 도면.

도 3은 녹음된 음성 신호의 재생 시에 행해지는 마크에의 위치 결정 동작을 설명하기 위한 도면으로, 조작에 따라 변화하는 LCD(135)의 표시 정보의 변화를 나타내는 도면.

도 4는 도 1에 도시한 기록 재생 장치에서의 녹음 처리를 설명하기 위한 흐름도.

도 5는 도 1에 도시한 기록 재생 장치에서의 재생 처리를 설명하기 위한 흐름도.

도 6은 도 1에 도시한 구성을 갖는 기록 재생 장치의 외부 기억 장치(111)의 기억 영역에 형성되는 음성 데이터베이스의 일례를 설명하기 위한 도면.

도 7은 도 1에 도시한 구성을 갖는 기록 재생 장치에서 행해지는 짐음한 음성 신호에 마크를 붙이는 처리의 개요를 설명하기 위한 도면.

도 8은 녹음된 음성 신호의 재생 시에 행해지는 마크에의 위치 결정 동작을 설명하기 위한 도면으로, 조작에 따라 변화하는 LCD(135)의 표시 정보의 변화를 나타내는 도면.

도 9는, 녹음 처리 종료 후에, 녹음한 음성 신호의 변화점에 마크를 붙이도록 하는 처리를 행하는 경우의 처리를 설명하기 위한 흐름도.

도 10은, 도 1에 도시한 기록 재생 장치로부터 퍼스널 컴퓨터에 전송된 데이터에 따라, 표시 장치의 표시 화면에의 변화점 정보의 표시예를 설명하기 위한 도면.

도 11은, 도 1에 도시한 기록 재생 장치로부터 퍼스널 컴퓨터에 전송된 데이터에 따라, 표시 장치의 표시 화면에의 변화점 정보의 표시예를 설명하기 위한 도면.

도 12는 본 발명이 적용되어 구성된 기록 재생 장치의 다른 예를 설명하기 위한 블록도.

도 13은 마이크로폰(131(1), 131(2))과 음성 신호 처리부(136)의 일례를 설명하기 위한 도면.

도 14는 마이크로폰(131(1), 131(2))과 음성 신호 처리부(136)의 다른 예를 설명하기 위한 도면.

도 15는, 녹음 처리 종료 후에, 녹음된 음성 신호의 변화점에 마크를 붙이도록 하는 처리를 행하는 경우의 처리를 설명하기 위한 도면.

도 16은 화자 위치 데이터베이스의 일례를 설명하기 위한 도면.

도 17은, 마이크로폰으로부터의 출력 신호에 따라 발원자의 위치를 특정함으로써, 발원자 자신을 특정하는 방식의 다른 예를 설명하기 위한 도면.

도 18은 본 발명이 적용되어 구성된 기록 재생 장치의 다른 예를 설명하기 위한 블록도.

<도면의 주요 부분에 대한 부호의 설명>

101 : CPU

102 : ROM

103 : RAM

104 : CPU 버스

110 : 파일 처리부

111 : 데이터 기억 장치

120 : 입력 처리부

121 : 키 조작부

131 : 마이크로폰

- 132 : A/D 컨버터
- 133 : 스피커
- 134 : D/A 컨버터
- 135 : LCD
- 141 : 데이터 압축 처리부
- 142 : 데이터 신장 처리부
- 143 : 음성 특징 해석부
- 144 : 통신 I/F
- 145 : 접속 단자
- 131(1), 131(2) : 마이크로폰
- 136 : 음성 신호 처리부

**발명의 상세한 설명**

**발명의 목적**

**발명이 속하는 기술 및 그 분야의 종래기술**

본 발명은, 예를 들면 IC(Integrated Circuit) 레코더, MD(Mini Disc) 레코더, 퍼스널 컴퓨터 등의 음성 신호를 처리하는 여러 가지 장치 및 이 장치에서 이용되는 방법에 관한 것이다.

예를 들면, 후술되는 특허 문헌 1에 개시되어 있는 바와 같이, 녹음된 음성 데이터에 대하여 음성 인식을 행하고, 이를 텍스트 데이터로 변환하여, 자동적으로 회의록을 작성하도록 하는 회의록 작성 장치가 제안되어 있다. 이러한 기술을 이용함으로써, 사람의 손을 통하지 않고, 회의의 회의록을 신속하게 작성하는 것이 가능하게 된다. 그러나, 녹음한 모든 음성 데이터에 기초를 둔 회의록을 작성할 필요도 없고, 중요한 부분만의 회의록을 작성하고자 하는 경우도 있다. 이 때문에, 녹음한 음성 데이터로부터 목적으로 하는 부분을 찾아낼 필요가 생긴다.

예를 들면, IC 레코더나 MD 레코더 등을 이용하여 장시간의 회의 등의 모습을 녹음한 경우, 기록된 음성 데이터로부터 듣고 싶은 장소를 찾아내기 위해서는, 그 음성 데이터를 재생하고, 재생 음성을 청취하도록 해야 한다. 물론, 앞으로 감기나 되감기 등의 기능을 이용하여, 목적으로 하는 부분을 찾도록 하는 것도 가능하지만, 수고나 시간이 걸리는 경우가 많다. 이 때문에, 「검색을 용이하게 하기 위한 안표」를, 녹음한 음성 데이터에 매립하도록 하는(부가하도록 하는) 것이 가능하도록 한 기능을 구비한 녹음 장치가 제공되어 있다. 예를 들면, MD 레코더 등에서는 트랙 마크를 부가하는 기능으로서 실현되어 있다.

[특허 문헌 1] 일본특허공개 평2-206825호 공보

**발명이 이루고자 하는 기술적 과제**

그런데, 상술한 바와 같이, 「검색을 용이하게 하기 위한 안표」를 음성 데이터에 부가하도록 하는 기능은 유저의 수동 조작에 의해 이용할 수 있는 것으로, 유저의 조작이 없으면 안표를 부가할 수 없다. 따라서, 녹음 중에 중요하다고 판단한 부분에 안표를 붙이는 조작을 행한다고 생각하고 있어도, 회의에 집중하고 있는 경우 등에서는, 안표를 붙이기 위한 조작을 잊게 되는 경우도 있다고 생각된다.

또한, 주목하여야 할 발언 부분에 안표를 붙인다고 하더라도, 안표를 매립하는 조작은 주목하여야 할 발언을 들었을 때에 행해지기 때문에, 안표는 주목 발언의 뒤에 기록된다. 그 때문에, 유저가 주목 발언을 듣기 위해서는, 재생 위치를 안표 부분으로 이동시킨 후, 약간 앞으로 되돌아가는 조작을 해야 한다. 희망하는 장소의 앞으로 지나치게 진행하거나 지나치게 되돌아가거나 하여, 이 조작을 반복해야만 한다는 것은 유저에게 대단히 번거롭고 스트레스가 쌓이는 작업이다.

또한, 안표가 붙어 있는 장소가 어떤 내용인지는 들어볼 때까지 알 수 없다. 들어 보아서, 목적의 장소가 아니면, 다음 안표로 이동시키는 조작을 목적의 장소에 도달할 때까지 반복해야 하고, 이것도 또한 수고로운 작업이다. 이와 같이, 「검색을 용이하게 하기 위한 안표」를 음성 데이터에 부가하도록 하는 기능은 편리한 것이지만, 유저의 조작이 불안한 경우 등에서는, 음성 데이터의 목적으로 하는 부분에 안표를 붙이는 기능을 충분히 기능시킬 수 없다.

이상의 것을 감안하여, 본 발명은, 유저의 손을 번거롭게 하지 않고, 처리 대상의 음성 신호 중의 목적으로 하는 부분을 신속하게 찾아내어 이용할 수 있도록 하는 장치 및 방법을 제공하는 것을 목적으로 한다.

### 발명의 구성 및 작용

상기 과제를 해결하기 위해서, 제1 발명의 음성 신호 처리 장치는,

처리 대상의 음성 신호에 기초하여, 해당 음성 신호의 화자의 변화를 소정의 처리 단위마다 검출하는 검출 수단과,

상기 검출 수단에 의해 화자가 변화하였다고 검출된 상기 음성 신호 상의 위치를 나타내는 변화점 정보를 취득하는 취득 수단과,

상기 취득 수단에 의해 취득된 상기 변화점 정보를 보유하는 보유 수단

을 포함하는 것을 특징으로 한다.

이 제1 발명의 음성 신호 처리 장치는, 검출 수단에 의해 처리 대상의 음성 신호의 변화점이 자동적으로 검출되고, 그 변화점의 해당 음성 신호 상의 위치를 나타내는 변화점 정보가 취득 수단에 의해 취득된다. 이 변화점 정보는 보유 수단에 의해 보유된다. 이와 같이, 변화점의 위치 정보인 변화점 정보를 보유하는 것은, 처리 대상의 음성 신호의 변화점에 대하여, 마크를 붙이는 것과 동의이다.

이와 같이 하여 검출되어 보유되는 변화점 정보를 이용하여, 변화점 정보에 대응하는 음성 신호에의 위치 결정이 가능하게 되어, 그 위치로부터 처리 대상의 음성 신호의 재생 등의 처리를 행할 수 있게 된다. 이에 의해, 유저는 자기의 손을 번거롭게 하지 않고, 음성 신호의 변화점에 자동 부여되는 마크를 기준으로 하여, 처리 대상의 음성 신호로부터 목적으로 하는 부분의 음성 신호를 신속하게 검색할 수 있게 된다.

또한, 제2 발명의 음성 신호 처리 장치는, 제1 발명의 음성 신호 처리 장치로서,

상기 검출 수단은, 상기 처리 단위마다 상기 음성 신호의 특징을 추출하고, 추출한 상기 음성 신호의 특징에 기초하여, 화음성 이외의 부분으로부터 화음성 부분으로의 변화점 및 화음성 부분의 화자의 변화점을 검출 가능한 것인 것을 특징으로 한다.

이 제2 발명의 음성 신호 처리 장치에 따르면, 검출 수단은, 처리 대상의 음성 신호에 대하여, 미리 결정된 처리 단위마다 그 특징을 검출하고, 먼저 검출한 특징과의 비교를 행하는 등의 처리를 행함으로써, 무음 부분이나 잡음 부분으로부터의 화음성 부분으로의 변화점이나, 화음성 부분이라도 화자가 변화한 부분의 변화점을 검출할 수 있게 된다.

이에 의해, 적어도, 화자가 변화한 부분에는 마크를 부여할 수 있게 된다. 그리고, 화자의 변화점을 기준으로 하여, 목적으로 하는 음성 데이터 부분을 신속하게 검색할 수 있게 된다.

또한, 제3 발명의 음성 신호 처리 장치는, 제2 발명의 음성 신호 처리 장치로서,

1명 이상의 화자의 화음성의 특징을 나타내는 특징 정보와 상기 화자의 식별 정보를 대응지어 기억 보유하는 기억 수단과,

상기 검출 수단에 의해 추출된 상기 음성 신호의 상기 특징과, 상기 기억 수단에 기억 보유되어 있는 상기 특징 정보를 비교하여, 화자를 특정하는 특정 수단

을 포함하고,

상기 보유 수단은, 상기 변화점 정보와, 상기 특정 수단에서 특정된 화자의 상기 식별 정보를 관련지어 보유하는 것을 특징으로 한다.

이 제3 발명의 음성 신호 처리 장치에 따르면, 화자의 화음성의 특징 정보와 화자의 식별 정보가 대응지어 기억 수단에 기억되어 있다. 특정 수단에 의해, 검출 수단으로부터의 처리 대상의 음성 데이터의 특징 정보와, 기억 수단의 특징 정보를 비교함으로써, 변화점에서의 화자가 특정되고, 변화점과 화자의 식별 정보가 보유 수단에 보유된다.

이에 의해, 보유 수단에 보유된 정보에 기초하여, 특정한 화자의 발언 부분만을 재생하거나 추출하거나 할 수 있음과 함께, 각 변화점에서의 화자가 누구인지에 따라, 목적으로 하는 음성 데이터 부분의 검색을 행할 수 있게 된다.

또한, 제4 발명의 음성 신호 처리 장치는, 제2 발명의 음성 신호 처리 장치로서,

복수의 마이크로폰의 각각에 대응하는 복수의 음성 채널의 음성 신호를 해석하여 화자 위치를 검출하는 제2 검출 수단을 포함하고,

상기 취득 수단은, 상기 제2 검출 수단에 의해 검출된 화자 위치의 변화도 고려하여, 상기 변화점을 특정하고, 특정한 변화점에 대응하는 상기 변화점 정보를 취득하는 것을 특징으로 한다.

이 제4 발명의 음성 신호 처리 장치에 따르면, 제2 검출 수단에 의해, 각 음성 채널의 음성 신호를 해석함으로써, 화자의 위치(화자 위치)가 검출되게 되고, 이에 기초하여, 처리 대상의 음성 신호의 변화점이 검출된다. 그리고, 취득 수단에 의해, 검출 수단으로부터의 변화점과 제2 검출 수단에 의해 검출되는 변화점의 양쪽이 이용되어, 실제로 이용하게 되는 변화점이 특정되고, 그 특정된 변화점의 위치를 나타내는 변화점 정보가 취득된다.

이에 의해, 제2 검출 수단에 의해 검출되는 변화점도 고려하여, 음성 신호에서의 변화점을 보다 정확하고 또한 확실하게 검출하여, 목적으로 하는 음성 데이터 부분의 검색을 행할 수 있게 된다.

또한, 제5 발명의 음성 신호 처리 장치는, 제3 발명의 음성 신호 처리 장치로서,

복수의 마이크론의 각각에 대응하는 복수의 음성 채널의 음성 신호에 따라 정해지는 화자 위치와, 상기 화자 위치의 화자의 식별 정보를 대응지어 기억 보유하는 화자 정보 기억 수단과,

상기 복수의 음성 채널의 음성 신호를 해석하여 얻어지는 상기 화자 위치에 따른 화자의 상기 식별 정보를 상기 화자 정보 기억 수단으로부터 취득하는 화자 정보 취득 수단

을 포함하고,

상기 특정 수단은, 상기 화자 정보 취득 수단에 의해 취득된 화자의 상기 식별 정보도 고려하여, 상기 화자를 특정하는 것을 특징으로 한다.

또한, 제5 발명의 음성 신호 처리 장치에 따르면, 각 음성 채널에 대응하는 마이크론에 따라 정해지는 화자 위치와, 그 화자 위치에 위치하게 되는 화자의 식별 정보가 화자 정보 기억 수단에 의해 기억 보유되어 있다. 구체예를 설명하면, 제1 마이크론에 가장 가까운 위치(화자 위치)에 있는 화자는 A씨, 제2 마이크론에 가장 가까운 위치(화자 위치)에 있는 화자는 B씨와 같이, 각 마이크론의 위치에 따라 결정되는 각 화자의 위치가 각 마이크론에 따라(각 마이크론의 배치 위치에 따라) 정해진다. 이 때문에, 예를 들면 어느 마이크론의 음성 채널의 음성 데이터의 레벨이 가장 높은가에 따라, 어느 마이크론의 근처에 있는 화자가 이야기하고 있는지를 식별할 수 있게 된다.

그리고, 화자 정보 취득 수단에 의해서, 각 음성 채널의 음성 데이터가 해석되고, 상술된 바와 같이, 어느 음성 채널의 마이크론을 통하여 주로 음성이 집중되었는가에 따라, 화자 위치를 특정하고, 그 화자 위치에 위치하는 화자를 특정할 수 있게 된다. 이와 같이 하여 취득한 정보도 이용하여, 특정 수단에 의해, 변화점에서의 화자가 특정되게 된다. 이에 의해, 화자의 특정 정밀도를 향상시키고, 정확한 정보를 이용하여, 해당 처리 대상의 음성 데이터로부터 목적으로 하는 부분을 검색할 수 있게 된다.

또한, 제6 발명의 음성 신호 처리 장치는, 제3 발명의 음성 신호 처리 장치로서,

상기 기억 수단에는, 각 식별 정보에 대응하는 화자에 관련되는 정보가 각 식별 정보에 대응지어 기억되어 있고,

상기 음성 신호에 대한 변화점의 위치와 상기 화자에 관련되는 정보를 표시하는 표시 정보 처리 수단을 포함하는 것을 특징으로 한다.

이 제6 발명의 음성 신호 처리 장치에 따르면, 기억 수단에는, 각 식별 정보에 대응하는 화자에 관련되는 정보, 예를 들면 얼굴 사진 데이터, 아이콘 데이터, 마크 화상 데이터, 애니메이션 화상 데이터 등의 여러 가지 화상 데이터나 그래픽스 데이터 등이, 각 식별 정보에 대응지어 기억 보유되도록 되어 있다. 그리고, 표시 정보 처리 수단에 의해서, 변화점의 위치와 화자에 관련되는 정보가 표시되게 된다.

이에 의해, 유저는, 처리 대상의 음성 데이터에 대하여, 시각을 통하여 각 화자의 발언 부분을 알 수 있게 되어, 처리 대상의 음성 데이터 중의 목적으로 하는 부분을 신속하게 찾아낼 수 있게 된다.

또한, 제7 발명의 음성 신호 처리 장치는, 제1 발명의 음성 신호 처리 장치로서,

상기 검출 수단은, 서로 다른 마이크론에 의해 집중되는 각 음성 채널의 음성 신호를 해석함으로써 얻어지는 화자 위치에 기초하여, 상기 화자의 변화를 검출하는 것인 것을 특징으로 한다.

이 제7 발명의 음성 신호 처리 장치에 따르면, 각 음성 신호 채널의 음성 신호를 해석함으로써, 화자의 위치(화자 위치)가 특정되게 되고, 그 화자 위치의 전환점이 변화점으로서 검출되게 된다.

이에 의해, 복수개 존재하는 각 음성 채널의 음성 신호를 해석함으로써, 처리 대상의 음성 신호의 변화점을 간단하고 또한 정확하게 검출하여, 화자가 변화한 부분에는 마크를 부여할 수 있게 된다. 그리고, 화자의 변화점을 기준으로 하여, 목적으로 하는 음성 데이터 부분을 신속하게 검색할 수 있게 된다.

또한, 제8 발명의 음성 신호 처리 장치는, 제7 발명의 음성 신호 처리 장치로서,

상기 보유 수단은, 상기 변화점 정보와, 상기 검출 수단에 의해 검출되는 화자 위치를 나타내는 정보를 관련지어 보유하는 것을 특징으로 한다.

이 제8 발명의 음성 신호 처리 장치에 따르면, 보유 수단에 보유되는 정보를 유저에게 제공할 수 있게 된다. 이에 의해, 어느 변화점에서, 어느 위치의 화자가 발언하고 있었는지를 파악할 수 있게 되고, 이에 기초하여, 처리 대상의 음성 데이터로부터 목적으로 하는 부분의 검색을 행할 수 있게 된다.

또한, 제9 발명의 음성 신호 처리 장치는, 제7 발명의 음성 신호 처리 장치로서,

복수의 마이크로폰의 각각에 대응하는 복수의 음성 채널의 음성 신호에 따라 정해지는 화자 위치와, 상기 화자 위치의 화자의 식별 정보를 대응지어 기억 보유하는 화자 정보 기억 수단과,

상기 복수의 음성 채널의 각각의 음성 신호를 해석하여 얻어지는 상기 화자 위치에 따른 화자의 상기 식별 정보를 상기 화자 정보 기억 수단으로부터 취득하는 화자 정보 취득 수단

을 포함하고,

상기 보유 수단은, 상기 변화점 정보와, 상기 화자 정보 취득 수단에 의해 취득된 화자의 상기 식별 정보를 관련지어 보유하는 것을 특징으로 한다.

이 제9 발명의 음성 신호 처리 장치에 따르면, 화자 정보 기억 수단에, 마이크로폰의 위치에 따라 정해지는 화자 위치와, 화자 위치의 화자의 식별 정보가 대응지어 기억되어 있고, 화자 정보 취득 수단에 의해서, 각 음성 채널의 음성 신호가 해석되어 화자 위치가 특정되고, 그 화자 위치에 위치하는 화자의 식별 정보가 변화점 정보와 대응지어 보유 수단에 의해 보유된다.

이에 의해, 각 변화점에서의 화자를 특정할 수 있게 되고, 이를 유저에게 제공할 수 있게 되어, 처리 대상의 음성 데이터로부터 목적으로 하는 부분의 검색을 간단하고 또한 정확하게 행할 수 있게 된다.

또한, 제10 발명의 음성 신호 처리 장치는, 제9 발명의 음성 신호 처리 장치로서,

상기 화자 정보 기억 수단에는, 각 식별 정보에 대응하는 화자에 관련되는 정보가 각 식별 정보에 대응지어 기억되어 있고,

상기 음성 신호에 대한 변화점의 위치와 상기 화자에 관련되는 정보를 표시하는 표시 정보 처리 수단을 포함하는 것을 특징으로 한다.

이 제10 발명의 음성 신호 처리 장치에 따르면, 화자 정보 기억 수단에는, 각 식별 정보에 대응하는 화자에 관련되는 정보, 예를 들면 얼굴 사진 데이터, 아이콘 데이터, 마크 화상 데이터, 애니메이션 화상 데이터 등의 여러 가지 화상 데이터나 그래픽스 데이터 등이, 각 식별 정보에 대응지어 기억 보유되도록 되어 있다. 그리고, 표시 정보 처리 수단에 의해서, 변화점의 위치와 화자에 관련되는 정보가 표시되게 된다.

이에 의해, 유저는, 처리 대상의 음성 데이터에 대하여, 시각을 통하여 각 화자의 발언 부분을 알 수 있게 되어, 처리 대상의 음성 데이터 중의 목적으로 하는 부분을 신속하게 찾아낼 수 있게 된다.

<실시예>

이하, 도면을 참조하면서, 본 발명에 따른 장치, 방법, 프로그램의 일 실시예에 대하여 설명한다. 이하에 설명하는 실시예에서는, 본 발명을 음성 신호의 기록 재생 장치인 IC 레코더에 적용한 경우를 예로 하여 설명한다.

[제1 실시예]

[IC 레코더의 구성과 동작의 개요]

도 1은 본 제1 실시예의 기록 재생 장치인 IC 레코더를 설명하기 위한 블록도이다. 도 1에 도시하는 바와 같이, 본 실시예의 IC 레코더는, CPU(Central Processing Unit)(101), 프로그램이나 각종 데이터가 기억되어 있는 ROM(Read Only Memory)(102), 주로 작업 영역으로서 이용되는 RAM(Random Access Memory)(103)이 CPU 버스(104)를 통하여 접속되고, 마이크로 컴퓨터의 구성으로 이루어진 제어부(100)를 구비하고 있다. 또한, RAM(103)은, 후술하는 바와 같이, 압축 데이터 영역(103(1))과 PCM(Pulse Code Modulation) 영역(103(2))이 형성되어 있다.

제어부(100)에는, 파일 처리부(110)를 통하여 데이터 기억 장치(111)가 접속되고, 입력 처리부(120)를 통하여 키 조작부(121)가 접속되어 있다. 또한, 제어부(100)에는, 아날로그/디지털 컨버터(이하, A/D 컨버터라고 약칭함)(132)를 통하여 마이크로폰(131)이 접속되고, 디지털/아날로그 컨버터(이하, D/A 컨버터라고 약칭함)(134)를 통하여 스피커(133)가 접속되어 있다. 또한, 제어부(100)에는 LCD(Liquid Crystal Display)(135)가 접속되어 있다. 또한, 본 실시예에서, LCD(135)는 LCD 컨트롤러의 기능도 구비한 것이다.

또한, 제어부(100)에는, 데이터 압축 처리부(141), 데이터 신장 처리부(142), 음성 특징 해석부(143), 통신 인터페이스(이하, 통신 I/F라고 약칭함)(144)가 접속되어 있다. 도 1에서, 2중선으로 나타낸 데이터 압축 처리부(141), 데이터 신장 처리부(142), 음성 특징 해석부(143)는, 제어부(100)의 CPU(101)에서 실행되는 소프트웨어(프로그램)에 의해서도 그 기능을 실현할 수 있는 것이다.

또한, 본 실시예에서, 통신 인터페이스(144)는, 예를 들면 USB(Universal Serial Bus)나 IEEE(Institute of Electrical and Electronics Engineers) 1394 등의 디지털 인터페이스로, 접속 단자(145)에 접속되는 퍼스널 컴퓨터, 디지털 카메라 등의 여러 가지 전자 기기와의 사이에서 데이터의 송수신을 행할 수 있는 것이다.

본 제1 실시예의 IC 레코더는, 키 조작부(121)의 REC 키(녹음 키)(211)가 누름 조작되면, CPU(101)가 각 부를 제어하여 녹음 처리를 행한다. 이 경우, 마이크로폰(131)에서 집음되고, A/D 컨버터(132)에서 디지털 변환됨과 함께, 데이터 압축 처리부(141)의 기능에 의해 데이터 압축된 음성 신호가, 파일 변환부(110)를 통하여 데이터 기억 장치(111)의 소정의 기억 영역에 기록된다.

본 제1 실시예의 데이터 기억 장치(111)는 플래시 메모리 혹은 플래시 메모리를 이용한 메모리 카드로, 후술하는 바와 같이, 데이터베이스 영역(111(1))과 음성 파일(111(2))이 마련된 것이다.

녹음 처리 시에서, 본 제1 실시예의 IC 레코더는, 음성 특징 해석부(143)의 기능에 의해, 집음하여 녹음(기록)하는 음성 신호에 대하여, 소정의 처리 단위마다 특징 해석을 행하고, 특징이 변화한 것을 검출한 경우에, 그 특징이 변화한 시점에 마크(안표)를 붙이도록 되어 있다. 그리고, 이 마크를 이용하여, 녹음한 음성 신호로부터 목적으로 하는 음성 신호 부분의 검색을 신속하게 행할 수 있게 하고 있다.

도 2는 집음하여 녹음하는 음성 신호의 변화점에 마크를 붙이도록 하는 처리의 개요를 설명하기 위한 도면이다. 본 제1 실시예의 IC 레코더에서는, 상술한 바와 같이, 마이크로폰(131)에 의해 집음된 음성 신호에 대하여, 소정의 처리 단위마다 특징 해석을 행한다.

그리고, 직전의 특징 해석 결과와 비교함으로써, 무음 부분이나 잡음 부분으로부터 화음성 부분으로 변화한 변화점, 혹은 화음성 부분이더라도, 화자가 변화한 변화점을 검출하고, 해당 음성 신호 상의 변화점의 위치(시간)를 특정한다. 그리고, 그 특정한 위치를 변화점 정보(마크 정보)로서 데이터 기억 장치(111)에 기억하여 두게 한다. 이와 같이, 음성 신호 상의 변화점의 위치를 나타내는 변화점 정보를 보유하는 것이, 집음하여 녹음하는 음성 신호에 대하여 마크를 붙이는 것으로 된다.

구체적으로 설명하면, 도 2에 도시하는 바와 같이, 회의의 모습을 녹음하도록 한 경우, 녹음 개시로부터 10초 후에, A씨가 발언을 시작하였다고 하자. 이 경우, A씨의 발언의 개시 전은, 무음 혹은 웅성거림이나 의자를 빼는 소리, 테이블에 무엇인가가 닿는 소리 등, 명료한 화음성과는 상이한 소위 잡음 등의 무의미한 음성이 집음되어 있고, A씨가 발언을 시작하고, 그 화음성이 집음됨으로써, 집음한 음성 신호의 특징 해석 결과는, A씨가 발언을 시작하기 전과는 분명히 다른 것으로 된다.

이 집음하여 녹음하는 음성 신호의 변화점을 음성 특징 해석부(143)에서 검출하고, 그 변화점의 음성 신호 상의 위치를 특정(취득)하고, 이 특정한 변화점 정보(음성 신호 상의 특정한 위치 정보)를 도 2에서의 마크 MK1으로서 데이터 기억 장치(111)에 기억 보유한다. 또한, 도 2에서는, 녹음 개시로부터의 경과 시간을 변화점 정보로서 기억 보유하도록 하고 있는 경우의 예를 도시하고 있다.

그리고, A씨의 발언이 종료한 후, 조금 간격을 두고, B씨가 발언을 시작하였다고 하자. 이 B씨의 발언 개시의 직전도 무음 혹은 잡음이다. 이 경우에도, B씨가 발언을 시작하고, 그 화음성이 집음됨으로써, 집음한 음성 신호의 특징 해석 결과는, B씨가 발언을 시작하기 전과는 분명히 상이한 것으로 되고, 도 2에서, 마크 MK2가 나타내는 바와 같이, B씨의 발언의 개시 부분에 마크를 붙이도록, 변화점 정보(마크 MK2)를 데이터 기억 장치(111)에 기억 보유한다.

또한, B씨의 발언 도중에 C씨가 끼어든 경우도 발생한다. 이 경우에는, B씨의 이야기하는 소리와 C씨의 이야기하는 소리에서는, 상이하게 되어 있으므로, 집음한 음성 신호의 해석 결과도 상이한 것으로 되고, 도 2에서, 마크 MK3가 나타내는 바와 같이, C씨의 발언의 개시 부분에 마크를 붙이도록, 변화점 정보(마크 MK3)를 데이터 기억 장치(111)에 기억 보유한다.

이와 같이, 본 실시예의 IC 레코더는, 녹음 처리 시에, 집음한 음성 신호의 특징 해석을 행하고, 특징이 변화한 음성 신호 상의 위치를 기억 보유함으로써, 음성 신호의 특징이 변화한 시점에 마크를 붙이도록 할 수 있게 한 것이다.

또한, 도 2에서, 마크 MK1, MK2, MK3에서, '그 외'라고 하는 란이 나타내는 바와 같이, 예를 들면 발언 부분을 음성 인식하여 텍스트 데이터로 변환함으로써, 그 텍스트 데이터를 관련지어 기억 보유하도록 하거나, 그 밖의 관련 정보를 함께 기억 보유시키도록 하거나 할 수도 있게 되어 있다.

그리고, 본 제1 실시예의 IC 레코더는, 키 조작부(121)의 PLAY 키(재생 키)(212)가 누름 조작되면, CPU(101)가 각 부를 제어하여 재생 처리를 행한다. 즉, 데이터 압축되어 데이터 기억 장치(111)의 소정의 기억 영역에 기억되어 있는 녹음된 음성 신호(디지털 음성 신호)가 파일 처리부(110)를 통하여 판독되고, 이것이 데이터 신장 처리부(142)의 기능에 의해 신장 처리되어, 데이터 압축 전의 원래의 디지털 음성 신호로 복원된다. 이 복원된 디지털 음성 신호가 D/A 컨버터(134)에서 아날로그 음성 신호로 변환되고, 이것이 스피커(133)에 공급되어, 녹음되어 재생하도록 된 음성 신호에 따른 음성이 방음된다.

이 재생 처리 시에, 본 제1 IC 레코더에서는, 키 조작부(121)의 NEXT 키(다음 마크에의 위치 결정을 지시하는 키)(214)나 PREV 키(앞 마크에의 위치 결정을 지시하는 키)(215)가 조작된 경우에, 이에 따라, 재생 위치를 재빠르게 마크가 부여된 위치에 위치 결정하고, 그곳으로부터 재생을 행할 수 있게 되어 있다.

도 3은 녹음한 음성 신호의 재생 시에 행해지는 마크가 나타내는 음성 신호 상의 위치에의 위치 결정 동작을 설명하기 위한 도면으로, 조작에 따라 변화하는 LCD(135)의 표시 정보의 변화를 도시하는 도면이다. 도 3에 도시하는 바와 같이, PLAY 키(211)가 누름 조작되면, 상술한 바와 같이, CPU(101)는 각 부를 제어하여, 지시된 녹음 음성 신호의 선두로부터 재생을 개시한다.

그리고, A씨의 발언 부분에서는, 도 2를 이용하여 설명한 바와 같이, 녹음 처리 시에 붙여진(기억 보유된) 마크 MK1에 기초하여, 도 3의 A에 나타내는 바와 같이, A씨의 발언의 개시 시각이 표시됨과 함께, 이것이 녹음 개시로부터 최초로 붙인 마크인 것을 나타내는 SEQ-No.1이라는 표시가 이루어진다.

재생이 속행되어, B씨의 발언 부분의 재생이 개시되면, 도 3의 B에 나타내는 바와 같이, B씨의 발언의 개시 시각이 표시됨과 함께, 이것이 녹음 개시로부터 2번째로 붙인 마크인 것을 나타내는 SEQ-No.2라는 표시가 이루어진다. 이후, PREV 키(215)가 누름 조작되면, CPU(101)는, 도 3의 C에 나타내는 바와 같이, 개시 시각이 선두로부터 10초 후(0분 10초 후)의 마크 MK1이 나타내는 A씨의 발언의 개시 부분에 재생 위치를 위치 결정하고, 그곳으로부터 재생을 재개한다.

이 후, NEXT 키가 누름 조작되면, CPU(101)는, 도 3의 D에 나타내는 바와 같이, 개시 시각이 선두로부터 1분 25초 후의 마크 MK2가 나타내는 B씨의 발언의 개시 부분에 재생 위치를 위치 결정하고, 그곳으로부터 재생을 재개한다. 또한, NEXT 키가 누름 조작되면, CPU(101)는, 도 3의 E에 나타내는 바와 같이, 개시 시각이 선두로부터 2분 30초 후의 마크 MK3가 나타내는 C씨의 발언의 개시 부분에 재생 위치를 위치 결정하고, 그곳으로부터 재생을 재개한다.

이와 같이, 본 실시예의 IC 레코더는, 녹음 처리 시에, 집음한 음성 신호의 특징 해석을 자동적으로 행하고, 특징의 변화점에 마크를 붙이도록 함과 함께, 재생 처리 시에서는, NEXT 키(214), PREV 키(215)를 조합함으로써, 붙여진 마크가 나타내는 녹음된 음성 신호 상의 위치에 재생 위치를 재빠르게 위치 결정하고, 그곳으로부터 재생을 행하도록 할 수 있는 것이다.

이에 의해서, 유저는, 재빠르게 목적으로 하는 화자(발언자)의 발언 부분에 재생 위치를 위치 결정하여, 녹음한 음성 신호를 재생하여 청취할 수 있으므로, 목적으로 하는 발언 부분의 회의록을 신속하게 작성할 수 있다.

또한, 여기서는, 설명을 간단히 하기 위해서, 변화점 정보로서, 녹음 개시 시점으로부터의 시각 정보를 이용하도록 하였지만, 이에 한정하는 것이 아니고, 녹음된 음성 신호의 데이터 기억 장치(111)의 기록 매체 상의 어드레스를 변화점 정보로서 이용할 수 있다.

[IC 레코더의 동작의 상세에 대하여]

다음으로, 도 4, 도 5의 흐름도를 참조하면서, 본 제1 실시예의 IC 레코더에서의 녹음 처리와 재생 처리에 대하여 상세하게 설명한다.

[녹음 처리에 대하여]

우선, 녹음 처리에 대하여 설명한다. 도 4는 본 제1 실시예의 IC 레코더에서 행해지는 녹음 처리를 설명하기 위한 흐름도이다. 도 4에 도시하는 처리는 CPU(101)가 각 부를 제어함으로써 행해지는 처리이다.

본 제1 실시예의 IC 레코더는 전원이 투입된 상태이며, 동작하지 않을 때에는 유저로부터의 조작 입력 대기로 된다(단계 S101). 유저가 조작부(121)에 있는 조작 키를 누르면, 입력 처리부(120)가 이를 검지하여 CPU(101)에 통지하기 때문에, CPU(101)는, 접수된 조작 입력이 REC 키(211)의 누름 조작인지의 여부를 판단한다(단계 S102).

단계 S102의 판단 처리에서, 접수된 조작 입력은 REC 키(211)의 누름 조작이 아니라고 판단하였을 때에는, CPU(101)는 유저에 의해 조작된 키에 따른 처리, 예를 들면 PLAY 키(212)에 따른 재생 처리, NEXT 키(214)에 따른 다음 마크에의 위치 결정 처리, PREV 키(215)에 따른 하나 전의 마크에의 위치 결정 처리 등을 행하게 된다(단계 S103). 물론, 앞으로 감기 처리나 되감기 처리 등을 행할 수도 있도록 되어 있다.

단계 S102의 판단 처리에서, REC 키가 눌러졌다고 판단한 경우에는, CPU(101)는 파일 처리부(110)에 파일 기록 처리를 행하도록 지시를 내리고, 이에 따라, 파일 처리부(110)는 데이터 기록 장치(111)에 음성 파일(111(2))을 작성한다(단계 S104).

그리고, CPU(101)는 키 조작부(121)의 STOP 키(정지 키)(213)가 누름 조작되었는지의 여부를 판단한다(단계 S105). 단계 S105의 판단 처리에서, STOP 키(213)가 조작되었다고 판단하였을 때에는, 후술하는 바와 같이, 소정의 종료 처리를 행하고(단계 S114), 이 도 4에 도시하는 처리를 종료한다.

단계 S105의 판단 처리에서, STOP 키(213)는 조작되어 있지 않다고 판단하였을 때에는, CPU(101)는 A/D 컨버터(132)에 마이크로폰(131)을 통하여 입력되는 아날로그 음성 신호를 디지털 음성 신호로 변환할 것을 지시하여, 집음 음성의 디지털 변환을 행하게 한다(단계 S106).



이에 따라, A/D 컨버터(132)는 마이크로폰(131)을 통하여 입력되는 아날로그 음성 신호를 일정 주기마다(소정의 처리 단위마다) 변환한 디지털 음성 신호를, RAM(103)의 PCM 데이터 영역(103(2))에 기입하고, 기입한 것을 CPU(101)에 통지한다(단계 S107).

이를 받아, CPU(101)는, 데이터 압축 처리부(141)에 대하여, RAM(104)의 PCM 데이터 영역(103(2))에 저장한 디지털 음성 신호(PCM 데이터)를 데이터 압축하도록 지시한다(단계 S108). 이에 따라, 데이터 압축 처리부(141)는 RAM(103)의 PCM 데이터 영역(103(2))의 디지털 음성 신호를 압축 처리하고, 압축한 디지털 음성 신호를 RAM(103)의 압축 데이터 영역(103(1))에 기입한다(단계 S109).

그리고, CPU(101)는, 파일 처리부(110)에 대하여, RAM(103)의 압축 데이터 영역(103(1))의 압축된 디지털 음성 신호를 데이터 기억 장치(111)에 작성한 음성 파일(111(2))에 기입할 것을 지시하고, 이에 따라, 파일 처리부(110)에 의해, RAM(103)의 압축 데이터 영역의 압축된 디지털 음성 신호가, 데이터 기억 장치(111)의 음성 파일(111(2))에 기입된다(단계 S110).

파일 처리부(110)는, 압축된 디지털 음성 신호의 음성 파일(111(2))에의 기입을 종료하면, 이를 CPU(101)에 통지하기 때문에, CPU(101)는, 음성 특징 해석부(143)에 대하여, RAM(103)의 PCM 데이터 영역(103(2))에 앞서 기록된 디지털 음성 신호의 특징 해석을 지시하여, 음성 특징 해석부(143)에 의해, RAM(103)의 PCM 데이터 영역(103(2))의 디지털 음성 신호의 특징을 추출한다(단계 S111).

또한, 음성 특징 해석부(143)에서 행해지는 디지털 음성 신호의 특징 해석(특징 추출) 처리는, 성문 분석, 화속 분석, 사이 멈춤 분석, 음성의 강약 분석 등의 여러 가지 방법을 이용하는 것이 가능하다. 여기서는 설명을 간단히 하기 위해서, 본 제1 실시예의 IC 레코더의 음성 특징 해석부(143)는, 성문 분석을 함으로써, 해석 대상의 디지털 음성 신호의 특징을 추출하는 것으로서 설명한다.

그리고, 음성 특징 해석부(143)는, 금회 추출한 음성의 특징(성문 데이터)과, 과거에 추출한 음성의 성문 데이터를 비교하여, 입력된 음성 신호로부터 추출한 특징이, 지금까지의 음성의 특징으로부터 변화하였는지의 여부를 판단하고, 그 판단 결과를 CPU(101)에 대하여 통지하기 때문에, 이에 기초하여, CPU(101)는 짐음 음성의 특징이 변화하였는지의 여부를 판단한다(단계 S112).

단계 S112의 판단 처리에서, 변화가 없었다고 판단하였을 때에는, CPU(101)는 단계 S105로부터의 처리를 반복하고, 다음 주기(다음 처리 단위)의 음성 신호에 대해서도 상술한 단계 S105부터 단계 S112까지의 처리를 행하게 한다.

단계 S112의 판단 처리에서, 변화가 있었다고 판단하였을 때에는, CPU(101)는 「화자가 전환되었다」고 판단하고, 파일 처리부(110)에 대하여, 처리 대상의 음성 신호 상의 음성의 특징의 변화점에 마크를 붙일 것을 지시한다(단계 S113). 이에 따라, 파일 처리부(110)는, 데이터 기록 장치(111) 상의 데이터베이스 영역(111(1))에 해당 음성 파일(111(2))에 관한 정보로서, 음성의 특징에 변화가 있던 장소를 나타내는 정보로서, 해당 음성 파일(111(2))의 선두로부터의 시각 정보 혹은 기록 위치에 대응하는 어드레스 정보를 기입한다. 이 경우, 음성 파일과 음성의 특징에 변화가 있던 장소를 나타내는 정보는 대응지어 기억된다.

이 단계 S113의 처리 후, CPU(101)는 단계 S105로부터의 처리를 반복하고, 다음 주기(다음 처리 단위)의 음성 신호에 대해서도 상술한 단계 S105부터 단계 S112까지의 처리를 행하게 한다.

그리고, 단계 S105의 판단 처리에서, 유저가 STOP 키(213)를 누름 조작하였다고 판단하였을 때에는, CPU(101)는 파일 처리부(110)에 대하여 데이터 기억 장치(111)의 음성 파일(111(2))에의 데이터의 기입의 정지를 지시하고, 데이터 압축 처리부(141)에 대하여 압축 처리의 정지를 지시하고, A/D 컨버터(132)에 대하여 디지털 신호로의 변환의 정지를 지시하는 등의 소정의 종료 처리를 행하고(단계 S114), 이 도 4에 도시하는 처리를 종료한다.

또한, 음성 특징 해석부(143)에서 행해지는 음성의 특징이 변화하였는지의 여부를 판단은, 과거에 추출한 음성의 특징 데이터(성문 데이터)를 보유해 두고, 이것과 새롭게 추출한 특징 데이터(성문 데이터)를 비교함으로써 행한다. 이 경우, 직전의 1개의 특징 데이터와만 비교하는 것만으로 되는 것이면, 과거의 특징 데이터는 항상 직전의 1개만을 보유해 두면 된다. 그러나, 정밀도를 향상시키기 위해, 과거의 2개 이상의 특징 데이터와 비교하고, 2개 이상의 차이가 생긴 경우에 특징이 변화하였다고 판단하도록 하는 경우에는, 2개 이상의 과거의 특징 데이터를 보유해 둘 필요가 있다.

이와 같이, 본 제1 실시예의 IC 레코더는, 짐음하여 녹음하는 음성 신호의 특징 해석을 행하고, 그 짐음 음성 신호의 특징의 변화점을 검출하여, 그 변화점에 상당하는 짐음 음성 신호 상의 위치에 마크를 붙이도록 할 수 있는 것이다.

[재생 처리에 대하여]

다음으로, 재생 처리에 대하여 설명한다. 도 5는 본 제1 실시예의 IC 레코더에서 행해지는 재생 처리를 설명하기 위한 흐름도이다. 도 5에 도시하는 처리는 CPU(101)가 각 부를 제어함으로써 행해지는 처리이다.

본 제1 실시예의 IC 레코더의 재생 처리에서는, 도 4를 이용하여 설명한 바와 같이, 녹음 처리 시에 붙여지는 짐음 음성(짐음하여 녹음하는 음성)의 특징의 변화점에 붙여진 마크를 이용하여, 녹음된 음성 신호로부터 신속하게 목적으로 하는 음성 신호 부분을 검출할 수 있도록 하고 있다.

본 제1 실시예의 IC 레코더는 전원이 투입된 상태에 있고, 동작하지 않을 때에는 유저로부터의 조작 입력 대기로 된다(단계 S201). 유저가 조작부(121)에 있는 조작 키를 누르면, 입력 처리부(120)가 이를 검지하여 CPU(101)에 통지하기 때문에, CPU(101)는, 접수한 조작 입력이 PLAY 키(212)의 누름 조작인지의 여부를 판단한다(단계 S202).

단계 S202의 판단 처리에서, 접수한 조작 입력이, PLAY 키(212)의 누름 조작이 아니라고 판단하였을 때에는, CPU(101)는 유저에 의해 조작된 키에 따른 처리, 예를 들면 REC 키(212)에 따른 녹음 처리, NEXT 키(124)에 따른 다음 마크에의 위치 결정 처리, PREV 키(215)에 따른 하나 전의 마크에의 위치 결정 처리 등을 행하게 된다(단계 S203). 물론, 앞으로 감기 처리나 되감기 처리 등을 행할 수도 있도록 되어 있다.

단계 S202의 판단 처리에서, 접수한 조작 입력이, PLAY 키의 누름 조작이라고 판단하였을 때에는, CPU(101)는 파일 처리부(110)에 데이터 기록 장치(111) 상의 음성 파일(111(2))의 판독을 지시한다(단계 S204). 그리고, CPU(101)는 키 조작부(121)의 STOP 키(정지 키)(213)가 누름 조작되었는지의 여부를 판단한다(단계 S205).

단계 S205의 판단 처리에서, STOP 키(213)가 조작되었다고 판단하였을 때에는, 후술하는 바와 같이, 소정의 종료 처리를 행하고(단계 S219), 이 도 5에 도시하는 처리를 종료하게 된다.

단계 S205의 판단 처리에서, STOP 키(213)가 조작되어 있지 않다고 판단하였을 때에는, CPU(101)는 파일 처리부(110)를 제어하여, 데이터 기억 장치(111)의 음성 파일(111(2))에 기억되어 있는 압축된 디지털 음성 신호를 시스템에서 규정된 소정의 처리 단위의 양만큼 판독하여 RAM(103)의 압축 데이터 영역(103(1))에 기입하게 한다(단계 S206).

기입이 종료되면, 이것이 CPU(101)에 통지되기 때문에, CPU(101)는, 데이터 신장 처리부(142)에 대하여, RAM(103)의 압축 데이터 영역(103(1))의 압축된 디지털 음성 신호의 신장 처리를 행할 것을 지시하여, 데이터 신장 처리부(142)에 의해, 압축된 디지털 음성 신호의 신장 처리를 행하고, RAM(103)의 PCM 데이터 영역(103(2))에 기입하게 한다(단계 S207).

기입이 종료되면, 이것이 CPU(101)에 통지되기 때문에, CPU(101)는, D/A 컨버터(134)에 대하여, RAM(103)의 PCM 데이터 영역(103(2))에 저장된 디지털 음성 신호(신장된 디지털 음성 신호)를 아날로그 음성 신호로 변환하여 스피커(133)에 공급하도록 제어한다.

이에 따라, 데이터 기억 장치(111)의 음성 파일(111(2))에 기억 보유되어 있는 디지털 음성 신호에 따른 음성이 스피커(133)로부터 방음되게 된다. 그리고, D/A 컨버터(134)는 D/A 변환한 아날로그 음성 신호를 출력한 것을 CPU(101)에 통지하기 때문에, CPU(101)는 키 조작부(121)의 조작 키가 조작되었는지의 여부를 판단한다(단계 S209).

단계 S209의 판단 처리에서, 조작 키는 조작되어 있지 않다고 판단하였을 때에는 단계 S205로부터의 처리를 반복하고, 데이터 기억 장치(111)의 음성 파일(111(2))의 디지털 음성 신호의 재생을 속행한다.

단계 S209의 판단 처리에서, 조작 키가 조작되었다고 판단하였을 때에는, CPU(101)는 조작된 키가 PREV 키(215)인지의 여부를 판단한다(단계 S210). 단계 S210의 판단 처리에서, PREV 키(215)가 조작되었다고 판단하였을 때에는, CPU(101)는 파일 처리부(110)에 대하여 음성 파일(111(2))로부터의 디지털 음성 신호의 판독의 정지를 지시하고, 데이터 신장 처리부(142)에 대하여 신장 처리의 정지를 지시하고, D/A 컨버터(134)에 대하여 아날로그 신호로의 변환의 정지를 지시한다(단계 S211).

다음으로, CPU(101)는, 파일 처리부(110)에, 데이터 기억 장치(111)의 데이터베이스 영역(111(1))으로부터 현재 재생하고 있는 위치 직전의 마크의 정보(변화점 정보)를 판독하도록 지시하여, 그 판독된 마크의 정보에 의해 지시되는 음성 신호 상의 위치에 재생 위치를 위치 결정하고, 그곳으로부터 재생을 개시하도록 하고(단계 S212), 도 3을 이용하여 설명한 바와 같이, 위치 결정에 이용한 마크의 정보에 따른 재생 위치 정보를 표시하고(단계 S213), 단계 S205로부터의 처리를 반복한다.

단계 S210의 판단 처리에서, 조작된 키는 PREV 키(215)가 아니라고 판단되었을 때에는, CPU(101)는 조작된 키가 NEXT 키(214)인지의 여부를 판단한다(단계 S214). 단계 S214의 판단 처리에서, NEXT 키(214)가 조작되었다고 판단하였을 때에는, CPU(101)는 파일 처리부(110)를 제어하여, 음성 파일(111(2))로부터의 디지털 음성 신호의 판독의 정지를, 데이터 신장 처리부(142)에 신장 처리의 정지를, D/A 컨버터(134)에 아날로그 신호로의 변환의 정지를 각각 지시한다(단계 S215).

다음으로, CPU(101)는, 파일 처리부(110)에, 데이터 기억 장치(111)의 데이터베이스 영역(111(1))으로부터 현재 재생하고 있는 위치 직후의 마크의 정보(변화점 정보)를 판독하도록 지시하여, 그 판독된 마크의 정보에 의해 지시되는 음성 신호 상의 위치에 재생 위치를 위치 결정하고, 그곳으로부터 재생을 개시하도록 하며(단계 S216), 도 3을 이용하여 설명한 바와 같이, 위치 결정에 이용한 마크의 정보에 따른 재생 위치 정보를 표시하고(단계 S217), 단계 S205로부터의 처리를 반복한다.

단계 S214의 판단 처리에서, 조작된 키는 NEXT 키(214)가 아니라고 판단되었을 때에는, CPU(101)는 조작된 키에 따른 처리, 예를 들면 앞으로 감기, 되감기 등의 조작된 키에 따른 처리를 행하고, 단계 S205로부터의 처리를 반복한다.

이와 같이, 녹음 시에 IC 레코더가 음성의 특징에 변화가 있던 것을 화자의 전환이라고 판단하고, 그 위치에 마크를 자동적으로 붙임으로써, 유저는 재생 시에 PREV 키(215), NEXT 키(214)를 누름 조작하여, 간단하게 각 발언의 선두 위치를 호출하는 것이 가능하게 되어, 회의록 작성 시에, 임의의 발언을 반복하여 재생시키는 것이나, 중요한 발언을 찾아낼 때의 수고가 대폭 삭감될 수 있다. 즉, 녹음된 음성 신호 중에서, 목적으로 하는 음성 신호 부분을 신속하게 검색할 수 있다.

또한, 집음 음성의 특징의 변화점은 자동 검출되어, 그 변화점에의 마크의 부여도 자동적으로 행해지기 때문에, 변화점에의 마크의 부여에 관하여, 유저의 손을 번거롭게 하는 일은 일절 없다.

[제1 실시예의 변형예]

그런데, 회의의 모습을 녹음하고, 이 녹음에 기초하여 회의록을 작성하는 경우, 누가 어디에서 발언을 하였는지를, 녹음 음성을 재생하지 않고 알 수 있다면, 보다 편리하다. 따라서, 본 변형예의 IC 레코더는, 회의에의 출석자의 음성의 특징 해석 결과인 성문 데이터를, 각 출석자를 식별하기 위한 심볼과 대응지어 기억시켜 둬으로써, 화자를 특정할 수 있는 마크를 붙이도록 한 것이다.

본 변형예의 IC 레코더는 도 1에 도시한 제1 실시예의 IC 레코더와 마찬가지로 구성되는 것이다. 그러나, 본 변형예의 IC 레코더의, 예를 들면 외부 기억 장치(111)나 RAM(103)의 기억 영역에는, 회의의 출석자에 대한 음성 특징 데이터베이스를 형성하도록 한 것이다. 또한, 이하의 설명에서는, 음성 특징 데이터베이스는 외부 기억 장치(111)에 형성하는 것으로서 설명한다.

도 6은 본 변형예의 IC 레코더의 외부 기억 장치(111)의 기억 영역에 형성되는 음성 데이터베이스의 일례를 설명하기 위한 도면이다. 도 6에 도시하는 바와 같이, 본 예의 음성 데이터베이스는, 회의의 출석자를 식별하기 위한 식별자(예를 들면, 등록 순에 따른 시퀀스 넘버 등)와, 회의의 출석자의 이름, 회의의 출석자의 음성의 특징 해석 결과인 성문 데이터, 회의의 출석자의 얼굴 사진 등의 화상 데이터, 회의의 출석자의 각각에 할당된 아이콘 데이터, 그 외, 텍스트 데이터 등으로 이루어진 것이다.

성문 데이터, 화상 데이터, 아이콘 데이터, 그 밖의 데이터의 각각은, 파일의 형식으로 외부 기억 장치(111)에 기억되어 있고, 이는 회의의 각 출석자의 식별자를 키 정보(대응화 정보)로 하여 기억 보유되어 있다. 또한, 특징 해석 결과인 성문 데이터는, 회의에 앞서서, 회의의 출석자의 음성을 집음하고, 특징 해석을 행함으로써, 미리 얻도록 한 것이다.

즉, 본 예의 IC 레코더는 음성 데이터베이스 작성 모드를 갖는 것이다. 그리고, 음성 데이터베이스 작성 모드가 선택된 경우에는, 회의의 출석자의 음성을 집음하고, 이 집음 음성의 특징 해석을 음성 특징 해석부(143)에서 행하여 성문 데이터를 얻고, 이 성문 데이터를 시퀀스 넘버 등의 식별자와 대응지어 외부 기억 장치(111)의 기억 영역에 기억할 수 있는 것이다.

식별자와 성문 데이터 이외의 정보인 이름, 화상 데이터, 아이콘 데이터 등의 정보는, 접속 단자(145)에 접속되는, 예를 들면 퍼스널 컴퓨터 등을 통하여, 본 예의 IC 레코더에 공급되고, 도 6에 도시한 바와 같이, 대응하는 식별자, 성문 데이터와 관련지어 기억 보유되게 된다. 물론, 이름 등은, IC 레코더의 키 조작부(121)에 설치되어 있는 조작 키를 조작하여 입력하는 것도 가능하다. 또한, 화상 데이터는, 접속 단자(145)에 접속되는 디지털 카메라로부터 취득되는 것도 가능하다.

그리고, 본 예의 IC 레코더도 또한, 도 1, 도 2 및 도 4를 이용하여 설명한 바와 같이, 집음 음성의 특징 해석을 행하여 성문 데이터의 변화점을 검출하고, 그 변화점에 대응하는 음성 신호 상의 위치에 마크를 자동적으로 부여해 가는 것이지만, 변화점을 검출한 경우에, 최신의 집음 음성의 성문 데이터와 음성 데이터베이스의 성문 데이터와의 매칭을 행하여, 성문 데이터가 일치한 회의의 출석자의 식별자를 부여하는 마크에 포함시키도록 하고 있다.

도 7은 본 변형예의 IC 레코더에서 행해지는 집음하여 녹음하는 음성 신호에 마크를 붙이는 처리의 개요를 설명하기 위한 도면이다. 기본적으로, 마크를 붙이는 처리는 도 2를 이용하여 설명한 경우와 마찬가지로 행해진다. 그러나, 마크에는 화자의 식별자가 부가된다.

도 7에 도시하는 바와 같이, 회의의 모습을 녹음하도록 한 경우, 녹음 개시로부터 10초 후에, A씨가 발언을 시작하였다고 하자. 이 경우, A씨의 발언의 개시 전은 무음 혹은 웅성거림이나 의자를 빼는 소리, 테이블에 무엇인가가 닿는 소리 등, 명료한 화음성과는 상이한 소위 잡음 등의 무의미한 음성이 집음되어 있기 때문에, 집음한 음성 신호의 특징 해석 결과는, A씨가 발언을 시작하기 전과는 분명히 상이한 것으로 된다. 이 변화점의 음성 신호 상의 위치를 특정(취득)하고, 이 특정한 변화점 정보를 도 7에서의 마크 MK1으로서 기억 보유한다.

이 경우에, 최신의 성문 데이터와 음성 데이터베이스의 성문 데이터와의 매칭을 행하여, 일치하는 성문 데이터에 대응하는 화자(회의의 출석자)의 식별자를 마크 MK1에 포함시키도록 한다. 또한, 도 7에서도, 녹음 개시로부터의 경과 시간을 변화점 정보로서 기억 보유하고 있는 경우를 도시하고 있다.

그리고, A씨의 발언이 종료된 후, 조금 간격을 두고, B씨가 발언을 시작하였다고 하자. 이 B씨의 발언 직전도 무음 혹은 잡음이었다고 한다. 이 경우에도, B씨가 발언을 시작하고, 그것이 집음됨으로써, 집음한 음성 신호의 특징 해석 결과는, B씨가 발언을 시작하기 전과는 분명히 상이한 것으로 되고, 도 7에서, 마크 MK2가 나타내는 바와 같이, B씨의 발언의 개시 부분에 마크를 붙이도록, 변화점 정보(마크 MK2)를 기억 보유한다.

이 경우에도, 최신의 성문 데이터와 음성 데이터베이스의 성문 데이터와의 매칭을 행하여, 일치하는 성문 데이터에 대응하는 화자(회의의 출석자)의 식별자를 마크 MK2에 포함시키도록 한다.

또한, B씨의 발언 도중에 C씨가 끼어든 경우도 발생하지만, 이 경우에는, B씨의 이야기하는 소리와 C씨의 이야기하는 소리에서는, 상이하게 되어 있으므로, 집음한 음성 신호의 해석 결과도 상이한 것으로 되고, 도 7에서, 마크 MK3가 나타내는 바와 같이, C씨의 발언의 개시 부분에 마크를 붙이도록, 변화점 정보(마크 MK3)를 기억 보유한다.

이 경우에도, 최신의 성문 데이터와 음성 데이터베이스의 성문 데이터와의 매칭을 행하여, 일치하는 성문 데이터에 대응하는 화자(회의의 출석자)의 식별자를 마크 MK3에 포함시키도록 한다.

이와 같이 함으로써, 녹음한 음성 신호의 어느 부분이 누구의 발언 부분인지를 특정할 수 있게 되어, 예를 들면, A씨의 발언 부분만을 재생하도록 하여 A씨의 발언의 요지를 정리하는 등의 것이 간단히 할 수 있게 된다.

또한, 본 변형예의 각 마크의 그 밖의 정보는, 예를 들면 짐음 음성의 음성 인식을 행하여, 짐음 음성을 텍스트 데이터로 변환하고, 이 텍스트 데이터를 그 밖의 정보로서 파일 형식(텍스트 데이터 파일)으로 기억 보유하도록 하고 있다. 이 텍스트 데이터를 이용함으로써, 회의록이나 발언의 요약을 신속하게 작성할 수 있게 된다.

그리고, 본 변형예의 IC 레코더에서도, 도 1, 도 3, 도 5를 이용하여 설명한 경우와 마찬가지로 하여, 녹음 음성의 재생을 행할 수 있게 된다. 그리고, 본 변형예의 IC 레코더의 경우에는, 녹음 음성에서의 각 발언자의 발언 부분의 녹음 음성을 재생하지 않고 특정할 수 있게 된다.

도 8은 녹음한 음성 신호의 재생 시에 행해지는 마크에의 위치 결정 동작을 설명하기 위한 도면으로, 조작에 따라 변화하는 LCD(135)의 표시 정보의 변화를 나타내는 도면이다. 도 8에 도시하는 바와 같이, PLAY 키(211)가 누름 조작되면, 상술한 바와 같이, CPU(101)는 각 부를 제어하여, 지시된 녹음 음성 신호의 선두로부터 재생을 개시하도록 한다.

그리고, A씨의 발언 부분에서는, 도 7을 이용하여 설명한 바와 같이, 녹음 처리 시에 붙여진(기억 보유된) 마크 MK1에 기초하여, 도 8의 A에 나타내는 바와 같이, A씨에 대한, 발언의 개시 시각 D(1), 화자의 화상 데이터에 따른 얼굴 사진 D(2), 화자의 이름 D(3), 발언의 최초 부분의 텍스트 데이터 D(4)가 표시됨과 함께, 재생 중 표시 D(5)가 표시된다.

그리고, 재생이 속행되어, B씨의 발언 부분의 재생이 개시되면, 녹음 시에 붙여진 마크 MK2에 기초하여, 도 8의 B에 나타내는 바와 같이, B씨에 대한, 발언의 개시 시각 D(1), 화자의 화상 데이터에 따른 얼굴 사진 D(2), 화자의 이름 D(3), 발언의 최초 부분의 텍스트 데이터 D(4)가 표시됨과 함께, 재생 중 표시 D(5)가 표시된다.

이 후, PREV 키(215)가 누름 조작되면, CPU(101)는, 도 8의 C에 나타내는 바와 같이, 개시 시각이 선두로부터 10초 후(0분 10초 후)의 마크 MK1이 나타내는 A씨의 발언의 개시 부분에 재생 위치를 위치 결정하고, 그곳으로부터 재생을 개시하도록 한다. 이 경우에는, 도 8의 A의 경우와 마찬가지로, A씨에 대한, 발언의 개시 시각 D(1), 화자의 화상 데이터에 따른 얼굴 사진 D(2), 화자의 이름 D(3), 발언의 최초 부분의 텍스트 데이터 D(4)가 표시됨과 함께, 재생 중 표시 D(5)가 표시된다.

이 후, NEXT 키가 누름 조작되면, CPU(101)는, 도 8의 D에 나타내는 바와 같이, 개시 시각이 선두로부터 1분 25초 후의 마크 MK2가 나타내는 B씨의 발언의 개시 부분에 재생 위치를 위치 결정하고, 그곳으로부터 재생을 개시하도록 한다. 이 경우에는, 도 8의 B의 경우와 마찬가지로, B씨에 대한, 발언의 개시 시각 D(1), 화자의 화상 데이터에 따른 얼굴 사진 D(2), 화자의 이름 D(3), 발언의 최초 부분의 텍스트 데이터 D(4)가 표시됨과 함께, 재생 중 표시 D(5)가 표시된다.

또한, NEXT 키가 누름 조작되면, CPU(101)는, 도 8의 E에 나타내는 바와 같이, 개시 시각이 선두로부터 2분 30초 후의 마크 MK3가 나타내는 C씨의 발언의 개시 부분에 재생 위치를 위치 결정하고, 그곳으로부터 재생을 개시하도록 한다. 이 경우에는, C씨에 대한, 발언의 개시 시각 D(1), 화자의 화상 데이터에 따른 얼굴 사진 D(2), 화자의 이름 D(3), 발언의 최초 부분의 텍스트 데이터 D(4)가 표시됨과 함께, 재생 중 표시 D(5)가 표시된다.

또한, 본 변형예에서, 예를 들면 A씨의 발언 부분을 재생 중에 NEXT 키 또는 PREV 키를 재빠르게 2회 누르면, 다음으로 A씨의 발언 부분이 출현하는 부분, 또는 이것 이전에 A씨의 발언 부분이 출현한 부분에 재생 위치를 위치 결정하고, 그곳으로부터 재생을 개시하는 모드를 부가하여도 된다. 즉, 이 조작을 반복함으로써, A씨의 발언 부분만을 찾거나, 혹은 거슬러 올라가 재생시킬 수 있다. 물론, NEXT 키나 PREV 키가 아니라, 이 모드를 명시적으로 나타내는 조작 키를 설치하여도 되며, 그 경우에는 자동적으로 차례차례로 A씨의 발언 부분이 재생되도록 한다.

이와 같이, 본 변형예의 IC 레코더는, 녹음 처리 시에, 짐음한 음성 신호의 특징 해석을 자동적으로 행하여, 특징의 변화점에 마크를 부여하도록 함과 함께, 재생 처리 시에는, NEXT 키(214), PREV 키(215)를 조작함으로써, 부여된 마크가 나타내는 녹음된 음성 신호 상의 위치에 재생 위치를 재빠르게 위치 결정하고, 그곳으로부터 재생을 행하도록 할 수 있는 것이다.

또한, 녹음된 음성 신호의 변화점에서는, 누구의 발언 부분인지를, 화자의 이름의 표시나 얼굴 사진의 표시에 의해 명확하게 나타낼 수 있기 때문에, 목적으로 하는 화자의 발언 부분을 신속하게 검색할 수 있음과 함께, 특정한 화자의 발언 부분만을 재생하도록 하는 등의 것을 간단하게 할 수 있다. 물론, 화자를 특정하기 위한 정보로서, 각 화자에 고유한 아이콘 데이터에 따른 아이콘을 표시하도록 하여도 된다. 또한, 발언의 최초 부분의 텍스트 데이터를 표시할 수도 있기 때문에, 목적으로 하는 발언 부분인지의 여부를 판단할 때에 쓸모있게 할 수 있다.

그리고, 본 변형예의 IC 레코더의 유저는, 재생 시의 표시 정보도 이용하여, 목적으로 하는 사람의 발언 부분에 재생 위치를 신속하게 위치 결정하여, 녹음한 음성 신호를 재생하여 청취할 수 있으므로, 목적으로 하는 발언 부분의 회의록을 신속하게 작성할 수 있다.

바꾸어 말하면, 녹음 후에 녹음 음성 신호를 일일이 재생하지 않고, 어디에 누구의 발언이 있는지를 시각적으로 파악할 수가 있어, 특정한 화자의 발언을 간단하게 찾아내는 것이 가능하게 된다. 심볼에는, 문자열이나 기호 외에, 화자의 얼굴 사진 등, 화자를 보다 특정하기 쉽게 할 수 있도록 하는 정보가 이용될 수 있기 때문에, 검색성이 향상된다.

또한, 음성의 특징이 미등록된 화자(등록 완료되어도 IC 레코더가 식별할 수 없었던 경우)의 발언에는, 미등록 화자인 것을 의미하는 심볼을 대응지어 둬으로써, 그 부분을 찾아내기 쉽게 할 수 있다. 이 경우, 회의록 작성자는, 미등록 화자의 발언 부분을 재생하여, 그것이 누구인지를 판단하면 된다.

미등록 화자가 누구인지를 알아낸 때에는, 그것이 등록 완료된 화자이었으면, 그 화자에 대응지어진 심볼을 마크로서 다시 붙일 수 있도록 할 수도 있다. 또한, 미등록의 화자인 경우에는, 화자의 신규 등록 조작을 행할 수 있도록 할 수도 있다. 음성의 특징은 녹음 음성으로부터 추출하고, 대응지는 심볼은 IC 레코더에 미리 등록 완료된 기호나 문자열 입력, IC 레코더에 카메라 촬영 기능이 있으면 촬영한 화상, 또는 외부 기기로부터 취득한 화상 데이터 등을 이용한다.

또한, 본 변형예의 IC 레코더의 녹음 처리는 도 4를 이용하여 설명한 녹음 처리와 마찬가지로 행해지지만, 단계 S113의 화자 전환의 마크 MK1, MK2, MK3, ...를 부여하는 처리에서, 음성 데이터베이스의 성문 데이터와의 매칭을 행하여, 해당하는 화자의 식별자가 부가되게 된다. 또한, 해당하는 성문 데이터가 없던 경우에는, 해당 없음을 나타내는 마크가 부여되게 된다.

또한, 본 변형예의 IC 레코더의 재생 처리는 도 5를 이용하여 설명한 재생 처리와 마찬가지로 행해지지만, 단계 S213, 단계 S217의 재생 위치 정보의 표시 처리에서, 화자의 얼굴 사진이나 성명, 발언 내용의 텍스트 데이터 등이 표시되도록 되게 된다.

또한, 본 변형예의 IC 레코더의 경우에도, 변화점 정보로서, 녹음 개시 시점으로부터의 시각을 이용하도록 하였지만, 이에 한정하는 것이 아니고, 녹음된 음성 신호의 데이터 기억 장치(111)의 기록 매체 상의 어드레스를 변화점 정보로서 이용하도록 하여도 된다.

[마크 부여 처리의 실행 타이밍에 대하여]

상술한 제1 실시예의 IC 레코더, 제1 실시예의 변형예의 IC 레코더에서는, 녹음 처리 시에 짐음 음성의 변화점을 검출하고, 그 변화점에 대응하는 음성 신호 상의 위치에 마크를 붙이도록 하였지만, 이에 한정하는 것이 아니다. 녹음 처리 종료 후에, 마크를 붙이도록 할 수 있다. 즉, 재생 처리 시에 마크를 붙이도록 하거나, 혹은 마크 부여 처리만을 행하도록 하거나 하는 것이 가능하다.

도 9는, 녹음 처리 종료 후에, 녹음한 음성 신호의 변화점에 마크를 붙이도록 하는 처리를 설명하기 위한 흐름도이다. 즉, 도 9에 도시하는 처리는, 재생 처리 시에 녹음 음성의 변화점에 마크를 붙이도록 하는 경우, 혹은 녹음 음성의 변화점에 대하여 마크 부여 처리만을 독립적으로 행하는 경우에 행해지는 것이다. 이 도 9에 도시하는 처리도 또한, IC 레코더의 CPU(101)가 각 부를 제어함으로써 행해지는 처리이다.

우선, CPU(101)는, 파일 처리부(104)를 제어하여, 데이터 기억 장치(111)의 음성 파일에 데이터 압축되어 기억되어 있는 녹음 음성 신호를 소정 단위분마다 관독하고(단계 S301), 모든 녹음 음성 신호의 관독을 종료하여 있는지의 여부를 판단한다(단계 S302).

단계 S302의 판단 처리에서, 모든 녹음 음성 신호가 관독되어 있지 않다고 판단하였을 때에는, CPU(101)는, 데이터 신장 처리부(142)를 제어하여, 데이터 압축되어 있는 녹음 음성 신호의 신장 처리를 행한다(단계 S303). 이후, CPU(101)가 음성 특징 해석부(143)를 제어하여, 신장한 음성 신호의 특징 해석을 행하여 성문 데이터를 얻고, 먼저 취득한 성문 데이터와 비교함으로써, 녹음 음성 신호의 특징이 변화하였는지의 여부를 판단한다(단계 S305).

단계 S305의 판단 처리에서, 녹음 음성 신호의 특징은 변화하여 있지 않다고 판단하였을 때에는 단계 S301로부터의 처리를 반복하도록 한다. 또한, 단계 S305의 판단 처리에서, 녹음 음성 신호의 특징이 변화하였다고 판단하였을 때에는, CPU(101)는 「화자가 전환되었다」고 판단하고, 파일 처리부(110)에, 음성의 특징에 변화가 있던 장소에 마크를 부가할 것을 지시한다(단계 S306).

이에 따라, 파일 처리부(110)는, 데이터 기록 장치(111) 상의 데이터베이스 영역(111(1))에 해당 음성 파일(111(2))에 관한 정보로서, 음성의 특징에 변화가 있던 장소를 나타내는 정보로서, 파일의 선두로부터의 시각 정보 혹은 기록 위치에 대응하는 어드레스 정보를 기입한다. 이 경우, 음성 파일과 음성의 특징에 변화가 있던 장소를 나타내는 정보는 대응지어 기억된다.

이 단계 S306의 처리 후, CPU(101)는 단계 S301로부터의 처리를 반복하고, 다음 주기(다음 처리 단위)의 음성 신호에 대해서도 마찬가지로의 처리를 행한다. 그리고, 단계 S302의 판단 처리에서, 모든 녹음 음성 신호에 대하여 관독이 종료하였다고 판단하였을 때에는 소정의 종료 처리를 행하고(단계 S307), 이 도 9에 도시하는 처리를 종료한다.

이에 의해, 녹음 처리 후에, 재생 처리 시에 녹음 음성의 변화점을 검출하여 해당 녹음 음성 신호에 대하여 마크를 부여하도록 하거나, 혹은 녹음 음성에 대하여 마크 부여 처리만을 독립적으로 행하도록 하거나 할 수 있다. 재생 처리 시에서, 마크의 부여를 행하는 경우에는, 도 9에 도시한 단계 S303에서 신장 처리된 음성 신호를 D/A 변환하고, D/A 변환 후의 아날로그 음성 신호를 스피커(133)에 공급하도록 하면 된다.

이와 같이, 녹음 후에 녹음 음성 신호의 특징의 변화점에 대하여 마크를 부여하도록 함으로써, 녹음 시의 처리의 부하와 소비 전력을 경감하는 것을 기대할 수 있다. 또한, 사용자가 모든 녹음에서 자동 마크 부여를 희망하지 않는 경우도 있다. 녹음 시의 자동 마크 부여 기능을 온/오프 설정을 할 수 있도록 하여도 된다. 그리고, 사용자가 오프로 설정한 채로 녹음하여 둔 경우에, 나중에 마크 부여가 필요하게 된 경우에는, 상술한 바와 같이 하여, 녹음 처리 후에도, 녹음 음성 신호에 대하여 마크 부여가 가능하기 때문에, 매우 편리하다.

또한, 상술한 바와 같이, 녹음된 음성 신호에 대한 마크 부여가 가능하기 때문에, 녹음 기능을 갖지 않지만 신호 처리 기능을 구비한 기기에서의 적용이 가능하게 된다. 예를 들면, 퍼스널 컴퓨터의 애플리케이션 소프트웨어에 본 발명을 적용하는 것도 가능하다. 즉, 음성 녹음 기기로 녹음된 음성 신호를 퍼스널 컴퓨터에 전송하고, 이 퍼스널 컴퓨터 상에서 동작하는 상술한 신호 처리 애플리케이션 소프트웨어에 의해, 마크 부여를 하는 것이 가능하다.

또한, 본 발명을 적용한 기기에서 작성한 데이터를 네트워크 등을 통하여 공유함으로써, 이 데이터로부터 회의록을 작성하지 않고, 이 데이터 그 자체를 회의록으로서 이용하는 것도 가능하게 된다.

따라서, 본 발명은, 녹음 기기 뿐만 아니라, 신호 처리가 가능한 여러 가지 전자 기기에 적용 가능하며, 이미 녹음 완료한 음성 신호이어도 본 발명을 적용한 전자 기기에서 처리함으로써 마찬가지로의 결과를 얻을 수 있다. 즉, 회의록의 작성을 효율적으로 행할 수 있게 된다.

또한, 상술한 바와 같이, 도 1을 이용하여 설명한 제1 실시예의 IC 레코더는 통신 I/F(144)를 구비하고 있어 퍼스널 컴퓨터 등의 전자 기기에 접속 가능하다. 따라서, 상술한 제1 실시예의 IC 레코더로 녹음됨과 함께, 변화점에 마크가 붙도록 된 음성 신호(디지털 음성 신호)를 퍼스널 컴퓨터에 전송하도록 하면, 퍼스널 컴퓨터의 큰 표시 화면의 표시 장치를 통하여 상세 정보를 보다 많이 표시하여, 목적으로 하는 발언자의 발언 부분을 신속하게 검색할 수 있다.

도 10, 도 11은 상술한 제1 실시예의 IC 레코더로부터 퍼스널 컴퓨터에 전송된 녹음 음성 신호, 부여된 변화점 정보(마크 정보)에 기초하여, 퍼스널 컴퓨터에 접속된 표시 장치(200)의 표시 화면에의 변화점 정보의 표시예를 설명하기 위한 도면이다.

도 10의 경우에는, 녹음 음성 신호에 대응하는 시간대 표시(201)와, 그 시간대 표시(201)의 해당 위치에, 마크 표시(변화점 표시) MK1, MK2, MK3, MK4, ...를 표시하도록 한다. 이와 같이 하면, 복수의 변화점의 위치를 일견에 인식할 수 있다. 그리고, 예를 들면 마우스 등의 포인팅 디바이스를 이용하여, 목적으로 하는 마크 표시에 커서를 위치 결정하여 클릭함으로써, 그 위치로부터 녹음 음성의 재생을 행하도록 하는 것 등을 할 수 있게 된다.

또한, 도 11의 경우에는 도 8에 도시한 표시를 표시 장치(200)의 표시 화면에 복수개 동시에 표시하도록 한 것으로, 화자의 얼굴 사진(211(1), 211(2), 211(3), ...)이나 발언 내용에 따른 텍스트 데이터(212(1), 212(2), 212(3), ...)를 표시하여, 목적으로 하는 화자의 발언 부분을 신속하게 검색하는 등의 것을 할 수 있게 된다. 또한, 퍼스널 컴퓨터의 기능을 이용하여, 타이틀 표시(210)를 행하도록 하는 것도 할 수 있다.

또한, 도 11의 표시예의 경우, 좌측의 「00」, 「01」, 「02」, 「03」, ...은 녹음 음성의 선두로부터의 시간을 나타내는 것이다. 물론, 도 8에 도시한 바와 같은 표시를 복수개 행하도록 하는 등, 여러 가지 표시 양태의 실현이 가능하다.

그리고, 발언(녹음 음성)과 그 발언자를 식별하는 정보(심볼)가 대응지어진 데이터를 퍼스널 컴퓨터 등의 표시부가 큰 기기에 전송하면, 음성 데이터로부터 문장을 작성하지 않더라도 회의록을 작성할 수 있다. 즉, 본 발명을 적용한 IC 레코더로 녹음한 데이터 그 자체가 회의록으로 되어 있게 된다.

또한, 그 데이터를 Web 페이지로 공개하여, Web 브라우저로 열람할 수 있도록 하는 플러그인(plugin)과 같은 소프트웨어를 마련하면, 네트워크를 통하여 회의록을 공유하는 것이 가능하게 된다. 이에 따라, 정보의 공유 즉, 정보를 공개하기까지의 수고와 시간이, 본 발명을 이용함으로써, 대폭 삭감될 수 있다.

[제2 실시예]

[IC 레코더의 구성과 동작의 개요]

도 12는 본 제2 실시예의 기록 재생 장치인 IC 레코더를 설명하기 위한 블록도이다. 본 제2 실시예의 IC 레코더는, 2개의 마이크로폰(131(1), 131(2))과, 이들 2개의 마이크로폰(131(1), 131(2))으로부터의 음성 신호를 처리하는 음성 신호 처리부(136)를 구비하는 점을 제외하면, 도 1에 도시한 제1 실시예의 IC 레코더와 마찬가지로 구성되는 것이다. 이 때문에, 본 제2 실시예의 IC 레코더에서, 도 1에 도시한 제1 실시예의 IC 레코더와 마찬가지로 구성되는 부분에는 동일한 참조 부호를 붙이고, 그 부분의 상세한 설명에 대해서는 생략하는 것으로 한다.

그리고, 본 제2 실시예의 IC 레코더에서는, 2개의 마이크로폰(131(1), 131(2))의 각각으로부터의 집음 음성 신호를 음성 신호 처리부(136)에서 처리함으로써, 화자의 위치(음원의 위치)를 특정하도록 하며, 이것도 고려하여 집음한 음성 신호의 변화점(화자의 변화점)을 특정할 수 있도록 한 것이다. 즉, 음성 해석의 결과로 얻어지는 성문 데이터를 이용한 집음 음성 신호의 변화점 검출을 행하는 경우의 보조 정보로서, 2개의 마이크로폰의 집음 음성에 기초하는, 화자의 위치도 이용하여, 보다 정확하게, 변화점이나 화자를 특정할 수 있도록 한 것이다.

도 13은 마이크로폰(131(1), 131(2))과 음성 신호 처리부(136)의 구성예를 설명하기 위한 도면이다. 이 도 13에 도시하는 예의 경우, 2개의 마이크로폰(131(1), 131(2))의 각각은, 도 13에 이들의 지향 특성을 도시한 바와 같이, 모두 단일 지향성의 것이다. 그리고, 마이크로폰(131(1), 131(2))은 주 지향 방향이 역 방향으로 되도록 등정합하게 근접 배치되어 있다. 이에 의해, 마이크로폰(131(1))은 발언자 A의 음성을 양호하게 집음하고, 마이크로폰(131(2))은 발언자 B의 음성을 양호하게 집음할 수 있게 된다.

그리고, 음성 신호 처리부(136)는, 도 13에 도시한 바와 같이, 가산기(1361)와 콤퍼레이터(비교기)(1362)와 A/D 컨버터(1363)를 구비한 것이다. 그리고, 마이크로폰(131(1), 131(2))의 각각 집음된 음성 신호는 가산기(1361)와 콤퍼레이터(1362)에 공급된다.

가산기(1361)는 마이크로폰(131(1))으로부터의 집음 음성 신호와 마이크로폰(131(2))으로부터의 집음 음성 신호를 가산하고, 가산 후의 음성 신호를 A/D 컨버터(1363)에 공급한다. 마이크로폰(131(1))으로부터의 집음 음성과 마이크로폰(131(2))으로부터의 집음 음성의 가산 신호는 다음의 수학적 식 1과 같이 표현할 수 있고, 무지향성 마이크로 집음한 것과 동일하게 되는 것을 알 수 있다.

수학식 1

$$((1 + \cos\theta)/2) + ((1 - \cos\theta)/2) = 1$$

또한, 콤퍼레이터(1362)는 마이크로폰(131(1))으로부터의 집음 음성 신호와 마이크로폰(131(2))으로부터의 집음 음성 신호를 비교한다. 그리고, 콤퍼레이터(1362)는 마이크로폰(131(1))으로부터의 집음 음성 신호의 레벨쪽이 크면, 발원자 A가 주로 발언하고 있다고 판단하고, 값이 「1(하이 레벨)」로 되는 화자 판별 신호를 제어부(100)에 공급한다. 또한, 콤퍼레이터(1362)는 마이크로폰(131(2))으로부터의 집음 음성 신호의 레벨쪽이 크면, 발원자 B가 주로 발언하고 있다고 판단하고, 값이 「0(로우 레벨)」으로 되는 화자 판별 신호를 제어부(100)에 공급한다.

이에 의해, 마이크로폰(131(1))으로부터의 집음 음성 신호와 마이크로폰(131(2))으로부터의 집음 음성 신호에 기초하여, 화자의 위치를 특정하게 하여, 발원자 A의 발원인지 발원자 B의 발원인지를 판별할 수 있도록 하고 있다.

또한, 3명째의 발원자 C가, 마이크로폰(131(1), 131(2))의 주 지향 방향과 교차하는 방향(도 13에서, 발원자 A, 발원자 B를 각각 경사 전방에 보는 위치(도 13의 가로 방향))으로부터 발언한 경우에는, 마이크로폰(131(1), 131(2))으로부터의 집음 음성의 출력 레벨은 거의 같게 된다.

이러한 위치에 있는 발원자 C에 대해서도 대응하는 경우에는, 콤퍼레이터(1362)에서의 임계값을 2개 설정하고, 레벨 차가  $\pm V_{th}$  이내이면 가로 방향에 있는 발원자 C에 의한 발원이라고 판단하고, 레벨 차가  $+V_{th}$ 보다 크면 발원자 A이고, 레벨 차가  $-V_{th}$ 보다 작으면 발원자 B라고 판단하도록 하여도 된다.

그리고, 마이크로폰(131(1))의 지향 방향에 위치하는 발원자, 마이크로폰(131(2))의 지향 방향에 위치하는 발원자, 마이크로폰(131(1), 131(2))의 지향 방향과 교차하는 방향에 위치하는 발원자의 각각이 누구인지를 파악해 줌으로써, 발원자(화자)가 누구인지를 식별할 수 있게 된다. 따라서, 집음 음성의 특징 해석의 결과로 얻어지는 성문 데이터에 의한 변화점 검출 외에, 마이크로폰의 집음 음성의 레벨도 고려함으로써, 발원자의 특징을 보다 정확하게 행하도록 할 수 있다.

[마이크로폰과 음성 신호 처리부의 다른 예]

또한, 마이크로폰(131(1), 131(2))과 음성 신호 처리부(136)는 도 14에 도시하는 바와 같이 구성할 수도 있다. 즉, 도 14는 마이크로폰(131(1), 131(2))과 음성 신호 처리부(136)의 다른 구성예를 설명하기 위한 도면이다. 이 도 14에 도시하는 예의 경우, 2개의 마이크로폰(131(1), 131(2))의 각각은 도 14에 이들의 지향 특성을 도시한 바와 같이, 모두 무지향성의 것이다. 마이크로폰(131(1), 131(2))은, 예를 들면 1cm 정도 이격하여 근접 배치하도록 한다.

또한, 도 14에 도시한 바와 같이, 본 예의 음성 신호 처리부(136)는 가산기(1361), A/D 컨버터(1363), 감산기(1364), 위상 비교기(1365)를 구비한 것이다. 그리고, 마이크로폰(131(1), 131(2))의 각각으로부터의 집음 음성 신호는 가산기(1361)와 감산기(1364)의 각각에 공급된다.

여기서, 가산기(1361)로부터의 가산 출력 신호는 무지향성 마이크 출력과 등가이고, 감산기(1364)로부터의 감산 출력은 양지향성(8자형 지향성) 마이크 출력과 등가이다. 양지향성 마이크는, 그 음파의 입사 방향에 따라 출력의 위상이 정상 또는 역상으로 된다. 따라서, 가산기(1361)로부터의 가산 출력(무지향성 출력)과 감산기(1364)로부터의 감산 출력 사이에서 위상 비교기(1365)에 의해 위상 비교를 행함으로써 감산기(1364)로부터의 감산 출력의 극성을 판단하여, 발원자를 특정할 수 있다.

즉, 감산기(1364)로부터의 감산 출력의 극성이 정상인 경우에는 발원자 A의 발언을 집음하고 있고, 감산기(1364)로부터의 감산 출력의 극성이 역상인 경우에는 발원자 B의 발언을 집음하고 있다고 판단할 수 있다.

또한, 도 13을 이용하여 설명한 경우와 마찬가지로, 발원자 A, 발원자 B의 각각을 경사 전방에 보는 위치(도 14의 가로 방향)에 위치하는 발원자 C의 발언도 판단하고자 하는 경우에는, 해당 발원자 C의 발언을 집음한 음성 신호의 감산 출력은 그 레벨이 작게 된다. 따라서, 가산기(1361)로부터의 가산 출력과 감산기(1364)로부터의 감산 출력과의 레벨을 체크함으로써, 발원자 C의 발언도 인식하는 것이 가능하게 된다.

또한, 도 14에 도시한 음성 신호 처리부(136)의 경우에는 가산기(1361)를 이용하도록 하였다. 그러나, 가산기(1361)는 필수적인 구성 요소가 아니다. 예를 들면, 마이크로폰(131(1) 및 131(2)) 중 어느 한쪽의 출력 신호를 A/D 컨버터(1363)와 위상 비교기(1365)에 공급하도록 하여도 된다.

이와 같이, 도 13, 도 14는, 녹음 처리 시에, 2개의 마이크로폰(131(1), 131(2))의 집음 음성의 레벨이나 극성을 이용하여, 발원자의 위치를 특정할 수 있도록 하고 있다. 그리고, 이 특정 결과도 고려함으로써, 집음 음성의 변화점의 검출 및 발원자의 특징을 정밀도 좋게 행할 수 있도록 하고 있다.

그리고, 도 13, 도 14를 이용한 방식은, 녹음 처리 시 뿐만 아니라, 재생 처리 시에 녹음 음성에 대하여 마크를 부여하는 경우나, 녹음 음성에 대하여 마크 부여 처리만을 독립적으로 행하는 경우에도 이용할 수 있다.

예를 들면, 도 13을 이용하여 설명한 방식을 녹음 처리 후에 이용하려고 하는 경우에는, 도 15의 A에 도시하는 바와 같이, 단일 지향성의 마이크로폰(131(1), 131(2))의 각각으로부터의 집음 음성 신호를 2 채널 스테레오 녹음한다. 그리고,

도 15의 B에 도시하는 바와 같이, 재생 시, 혹은 마크 부여 처리를 독립적으로 행하는 경우에, 외부 기억 장치(111)로부터 판독되는 데이터 압축된 2 채널의 음성 신호의 각각을 신장 처리하고, 신장 처리 후의 2 채널의 음성 신호를 도 13에 도시한 콤퍼레이터(1362)와 마찬가지로의 기능을 갖는 콤퍼레이터에 입력한다.

이에 의해, 마이크로폰(131(1))의 집음 음성 신호가 주로 이용되었는지, 마이크로폰(131(2))의 집음 음성 신호가 주로 이용되었는지를 판별할 수가 있고, 이 판별 결과와 미리 파악되어 있는 각 마이크로폰에 대한 발언자의 위치에 기초하여, 발언자를 특정할 수 있다.

또한, 도 14를 이용하여 설명한 방식을 녹음 처리 후에 이용하고자 하는 경우에서도 마찬가지로, 마이크로폰(131(1), 131(2))으로부터의 출력 신호를 2 채널 스테레오 녹음하고, 재생 시나 마크 부여 처리를 독립적으로 행하는 경우에, 도 14에 도시한 음성 신호 처리부(136)와 마찬가지로의 처리를 행함으로써, 발언자를 특정할 수 있다.

그리고, 마이크로폰(131(1), 131(2))으로부터의 출력 신호를 이용한 발언자의 특정 처리를 행하는 경우에 미리 준비하는 마이크로폰(131(1), 131(2))의 각각에 대한 발언자의 위치 정보는, 예를 들면 도 16에 도시하는 화자 위치 데이터베이스와 같이 하여, IC 레코더에 기억 보유되어 두도록 하면 된다.

도 16은 화자 위치 데이터베이스의 일례를 설명하기 위한 도면이다. 본 예의 화자 위치 데이터베이스는, IC 레코더의 음성 신호 처리부(136)로부터의 식별 결과에 대응하는 화자 식별 신호와, 각 화자 식별 신호에 대응하는 마이크로폰의 식별 정보와, 각 마이크로폰을 주로 이용하는 발언자 후보의 식별자(화자 식별자)로 이루어지는 것이다. 또한, 도 16에 도시한 바와 같이, 1개의 마이크로폰에 대하여, 화자 식별자는 복수개 등록할 수 있도록 하고 있다.

이 도 16에 도시하는 바와 같은 화자 위치 데이터베이스는, 바람직하게는, 회의가 시작하기 전에 작성해 둔다. 일반적으로, 회의에의 출석자나 각 출석자의 석순은 미리 결정되어 있는 경우가 많기 때문에, IC 레코더의 설치 위치를 고려하여, 회의가 시작되기 전에 화자 위치 데이터베이스를 작성하는 것이 가능하다.

또한, 회의에의 출석자의 급한 변경이나, 회의 중에 좌석이 변경된 경우에는, 예를 들면 마이크로폰의 집음 음성에 따른 발언자의 인식은 행하지 않도록 하고, 음성 해석 처리에 의해 얻은 성문 데이터에 의한 변화점의 검출만을 행하도록 하거나, 혹은, 녹음 처리 후에, 화자 위치 데이터베이스를 조정하여 정확한 것으로 하여 녹음 음성에 대하여 마크의 재부여를 행하도록 하거나 할 수도 있다.

이 도 16에 도시하는 바와 같은 화자 위치 데이터베이스를 이용함으로써, 화자 위치를 특정하고, 그 위치의 화자 자체도 특정할 수 있게 된다.

또한, 본 제2 실시예에서는, 2개의 마이크로폰(131(1), 131(2))을 이용하고, 발언자도 2명 또는 3명이 있는 경우를 예로 하여 설명하였지만, 이에 한정하는 것이 아니다. 이용하는 마이크로폰을 많게 함으로써, 보다 많은 발언자의 식별을 행하는 것이 가능하다.

또한, 마이크로폰으로부터의 출력 신호에 따라 발언자의 위치를 특정함으로써, 발언자 자신을 특정하는 방식으로서, 도 13, 도 14를 이용하여 설명한 방식에 한정하는 것이 아니다. 예를 들면, 근접 4점법(closely located four point microphone method)이나 근접 3점법 등을 이용하도록 할 수도 있다.

근접 4점법은, 도 17의 A에 도시하는 바와 같이, 반드시 1개의 마이크로폰이 동일 평면 내에 있지 않도록 하여 근접 배치되는 4개의 마이크로폰 M0, M1, M2, M3로 집음된 음성 신호의 시간 구조가 근소한 차이에 주목하여, 단시간 상관(Short-Time Correlation) 혹은 음향 인텐시티(Acoustic Intensity) 등의 방법에 의해, 음원의 위치나 크기 등의 공간 정보를 산출하도록 하는 방법이다. 이와 같이, 적어도 4개의 마이크로폰을 이용함으로써, 발언자의 위치를 성격으로 특정하고, 그 발언자의 위치(좌석 위치)에 따라 발언자를 특정하는 것도 가능하다.

또한, 발언자가 거의 수평면 내에 위치한다고 한정하여 지장이 없는 경우에는, 근접 배치되는 마이크로폰의 배치 관계는 도 17의 B에 도시하는 바와 같이 수평면 내의 3개이어도 된다.

또한, 도 17의 A, B에 도시한 바와 같이, 각 마이크로폰의 배치 관계는 직교 관계로 되지 않아도 된다. 도 17의 B에 도시한 근접 3점법의 경우에는 3개의 마이크로폰이 예를 들면 정삼각형의 정점에 배치되도록 하는 위치 관계로 되도록 하여도 된다.

## [제2 실시예의 변형예]

상술한 제2 실시예의 IC 레코더에서는, 음성 해석의 결과로 얻어지는 성문 데이터를 이용한 집음 음성 신호의 변화점 검출을 행하는 경우에, 2개의 마이크로폰의 집음 음성에 기초하는, 주로 이용되고 있는 마이크로폰의 판별 결과도 고려함으로써, 음성 신호의 변화점의 검출을 보다 정밀도 좋게 행하도록 하였다. 그러나, 이에 한정하는 것이 아니다.

예를 들면, 도 18에 도시하는 바와 같이, 2개의 마이크로폰(131(1), 131(2))과 음성 신호 처리부(136)는 구비하지만, 음성 특징 해석부(143)를 구비하지 않도록 한 IC 레코더를 구성하는 것도 가능하다. 즉, 도 18의 IC 레코더는, 음성 특징 해석부(143)를 구비하지 않는 점을 제외하면, 도 12에 도시한 제2 실시예의 IC 레코더와 마찬가지로 구성되는 것이다.

그리고, 2개의 마이크로폰(131(1), 131(2))의 집음 음성에 기초하는, 주로 이용되고 있는 마이크로폰의 판별 결과에만 기초하여, 화자의 변화점을 검출하고, 그 변화점에 따른 음성 신호 상의 대응하는 위치에 마크를 붙이도록 할 수도 있다. 이와 같이 한 경우에는, 음성 특징 해석 처리를 행할 필요가 없기 때문에, CPU(101)에 관한 부하를 경감할 수 있다.



또한, 상술한 실시예에서는, 처리 대상의 음성 신호의 변화점에 마크를 붙이도록 하였지만, 변화점이더라도, 화음성으로의 변화점에만 마크를 붙이도록 함으로써, 보다 효율적으로 검색을 행하도록 할 수 있다. 예를 들면, 처리 대상의 음성 신호의 신호 레벨이나 성문 데이터 등에 기초하여, 화음성과 그 이외의 잡음 등의 불필요 부분을 명확하게 인식하도록 하여, 화음성의 개시점에만 마크를 붙이도록 할 수도 있다.

또한, 성문 데이터나 음성 신호의 주파수의 특징 데이터 등에 기초하여, 화자가 남성인지 여성인지를 판별하여, 변화점에서의 화자의 성별을 통지하도록 하는 것도 가능하다.

또한, 상술한 바와 같이 붙여지는 마크 정보에 기초하여, 검색만을 행하는 용도로 하는 검색 모드나, 붙여진 마크의 위치를 변경하거나, 삭제하거나, 추가하거나 하는 마크 편집 모드나, 또한, 붙여진 마크에 따라 지정 가능한 화자의 발언 부분만, 예를 들면 A씨의 발언 부분만을 재생하도록 하는 특수 재생 모드 등을 설치하도록 할 수도 있다. 이들 각 모드의 실현은, CPU(101)에서 실행하는 프로그램을 추가하는 것만으로 비교적 간단히 실현 가능하다.

또한, 도 6에 도시한 음성 특징 데이터베이스의 성문 데이터를, 변화점의 검출에 이용한 성문 데이터에 의해 갱신하여, 정밀도가 높은 음성 특징 데이터베이스로 하는 등, 데이터베이스의 갱신 기능을 갖출 수 있도록 하여도 된다. 예를 들면, 성문 데이터의 비교 처리에서 불일치하여도, 실제로 그 발언자의 음성 특징 데이터베이스가 존재하고 있는 경우에는, 그 발언자의 데이터베이스의 성문 데이터를 새롭게 취득한 성문 데이터로 변경하도록 할 수 있다.

또한, 성문 데이터의 비교 처리에서 일치하여도, 실제로는 다른 화자의 성문 데이터와 일치하는 것과 같은 경우에는, 그 다른 화자의 성문 데이터를 비교 처리에 이용하지 않도록 설정하는 등의 것도 할 수 있다.

또한, 성문 데이터가 복수의 화자의 성문 데이터와 일치하는 것과 같은 경우에는, 올바른 화자와만 일치하도록, 이용하는 성문 데이터에 우선순위를 부여하도록 하여도 물론 된다.

또한, 마크를 붙이는 위치는, 발언의 개시점 뿐만 아니라, 종료점에 붙이도록 하여도 되고, 그 외, 개시점으로부터 수초 후 또는 수초 전 등의 유저 개개인의 편리성을 고려하여 변경할 수 있도록 하는 것도 가능하다.

또한, 음성 신호의 특징 해석은, 상술한 바와 같이 성문 해석 뿐만 아니라, 여러 가지 방법 중 하나 이상을 이용함으로써, 정밀도가 높은 해석 데이터를 얻도록 할 수 있다.

또한, 상술한 제2 실시예에서는, 주로 2개의 마이크로폰을 이용한 경우를 예로 하여 설명하였지만, 이에 한정하는 것이 아니다. 마이크로폰의 수는 2개 이상이면, 몇 개이어도 되고, 이들 복수의 마이크로폰의 개개의 집음 음성의 신호 레벨, 극성, 또한, 집음까지의 지연 시간 등의 여러 가지 파라미터를 이용하여, 화자의 위치를 특정하고, 그 위치에 따른 화자 자신을 특정할 수 있도록 할 수 있다.

또한, 상술한 제1, 제2 실시예에서는, 음성 신호의 기록 재생 장치인 IC 레코더에 본 발명을 적용한 경우를 예로 하여 설명하였지만, 이에 한정하는 것이 아니다. 예를 들면, 하드디스크 드라이브나 MD 등의 광자기 디스크, DVD 등의 광 디스크 등의 기록 매체를 이용하는 기록 장치, 재생 장치, 기록 재생 장치에 본 발명을 적용할 수 있다. 즉, 여러 가지 기록 매체를 이용하는 기록 장치, 재생 장치, 기록 재생 장치에 본 발명을 적용하는 것이 가능하다.

[소프트웨어에 의한 실현]

또한, 상술한 실시예의 IC 레코더의 음성 특징 해석부(143), 음성 신호 처리부(136) 등의 각 처리부의 기능을 실현하도록 함과 함께, 각 기능을 유기적으로 결합하도록 하는 프로그램을 작성하고, 이 프로그램을 CPU(101)에서 실행하도록 함으로써, 본 발명을 실현할 수 있다. 즉, 도 4, 도 5의 흐름도에 도시한 처리를 행하는 프로그램을 작성하고, 이를 CPU(101)에서 실행시킴으로써, 본 발명을 실현할 수 있다.

또한, 상술한 실시예와 마찬가지로, 예를 들면 음성 특징 해석부(143)의 기능을 실현한 프로그램이 인스톨된 퍼스널 컴퓨터에, 녹음기로 녹음한 음성 데이터를 취득하여, 화자의 전환을 검출시키는 것도 가능하다.

**발명의 효과**

본 발명에 따르면, 장시간의 회의를 녹음하여도 화자가 전환될 때마다 전환 마크(안표)가 자동적으로 부가되게 되기 때문에, 회의록을 작성할 때에, 발언의 검색성을 향상시켜, 목적으로 하는 화자의 발언 부분을 반복하여 재생하는 등의 일이 간단하고 또한 신속하게 행하여지게 된다.

또한, 변화점에서의 화자를 식별하고, 그 식별한 화자를 나타내는 정보와 음성 데이터의 변화점을 대응지어 관리할 수 있으므로, 음성 데이터를 재생하지 않고, 특정한 화자의 발언 부분을 간단하고 또한 신속하게 찾아낼 수 있다.

또한, 지금까지 회의록 작성자의 기억에 의지하고 있던 부분을 배제하고, 수고와 시간이 걸렸던 회의록 작성 작업의 효율을 향상시킬 수 있다. 또한, 회의록 자체의 작성을 생략하고, 검색성이 높은, 음성 데이터의 형식의 회의록으로서 녹음 데이터를 이용하도록 할 수 있다.

**(57) 청구의 범위**

**청구항 1.**

음성 신호 처리 장치에 있어서,  
처리 대상의 음성 신호에 기초하여, 해당 음성 신호의 화자의 변화를 소정의 처리 단위마다 검출하는 검출 수단과,  
상기 검출 수단에 의해 화자가 변화하였다고 검출된 상기 음성 신호 상의 위치를 나타내는 변화점 정보를 취득하는 취득 수단과,  
상기 취득 수단에 의해 취득된 상기 변화점 정보를 보유하는 보유 수단  
을 포함하는 것을 특징으로 하는 음성 신호 처리 장치.

## 청구항 2.

제1항에 있어서,  
상기 검출 수단은, 상기 처리 단위마다 상기 음성 신호의 특징을 추출하고, 추출한 상기 음성 신호의 특징에 기초하여, 화음성 이외의 부분으로부터 화음성 부분으로의 변화점 및 화음성 부분의 화자의 변화점을 검출 가능한 것을 특징으로 하는 음성 신호 처리 장치.

## 청구항 3.

제2항에 있어서,  
1명 이상의 화자의 화음성의 특징을 나타내는 특징 정보와 상기 화자의 식별 정보를 대응지어 기억 보유하는 기억 수단과,  
상기 검출 수단에 의해 추출된 상기 음성 신호의 상기 특징과 상기 기억 수단에 기억 보유되어 있는 상기 특징 정보를 비교하여, 화자를 특정하는 특정 수단  
을 포함하고,  
상기 보유 수단은, 상기 변화점 정보와, 상기 특정 수단에서 특정된 화자의 상기 식별 정보를 관련지어 보유하는 것을 특징으로 하는 음성 신호 처리 장치.

## 청구항 4.

제2항에 있어서,  
복수의 마이크로폰의 각각에 대응하는 복수의 음성 채널의 음성 신호를 해석하여 화자 위치를 검출하는 제2 검출 수단  
을 포함하고,  
상기 취득 수단은, 상기 제2 검출 수단에 의해 검출된 화자 위치의 변화도 고려하여, 상기 변화점을 특정하고, 특정한 변화점에 대응하는 상기 변화점 정보를 취득하는 것을 특징으로 하는 음성 신호 처리 장치.

## 청구항 5.

제3항에 있어서,  
복수의 마이크로폰의 각각에 대응하는 복수의 음성 채널의 음성 신호에 따라 정해지는 화자 위치와, 상기 화자 위치의 화자의 식별 정보를 대응지어 기억 보유하는 화자 정보 기억 수단과,  
상기 복수의 음성 채널의 음성 신호를 해석하여 얻어지는 상기 화자 위치에 따른 화자의 상기 식별 정보를 상기 화자 정보 기억 수단으로부터 취득하는 화자 정보 취득 수단  
을 포함하고,

상기 특정 수단은, 상기 화자 정보 취득 수단에 의해 취득된 화자의 상기 식별 정보도 고려하여, 상기 화자를 특정하는 것을 특징으로 하는 음성 신호 처리 장치.

#### 청구항 6.

제3항에 있어서,

상기 기억 수단에는, 각 식별 정보에 대응하는 화자에 관련되는 정보가 각 식별 정보에 대응지어 기억되어 있고, 상기 음성 신호에 대한 변화점의 위치와 상기 화자에 관련되는 정보를 표시하는 표시 정보 처리 수단을 포함하는 것을 특징으로 하는 음성 신호 처리 장치.

#### 청구항 7.

제1항에 있어서,

상기 검출 수단은, 서로 다른 마이크로폰에 의해 집음되는 각 음성 채널의 음성 신호를 해석함으로써 얻어지는 화자 위치에 기초하여, 상기 화자의 변화를 검출하는 것인 것을 특징으로 하는 음성 신호 처리 장치.

#### 청구항 8.

제7항에 있어서,

상기 보유 수단은, 상기 변화점 정보와, 상기 검출 수단에 의해 검출되는 화자 위치를 나타내는 정보를 관련지어 보유하는 것을 특징으로 하는 음성 신호 처리 장치.

#### 청구항 9.

제7항에 있어서,

복수의 마이크로폰의 각각에 대응하는 복수의 음성 채널의 음성 신호에 따라 정해지는 화자 위치와, 상기 화자 위치의 화자의 식별 정보를 대응지어 기억 보유하는 화자 정보 기억 수단과,

상기 복수의 음성 채널의 각각의 음성 신호를 해석하여 얻어지는 상기 화자 위치에 따른 화자의 상기 식별 정보를 상기 화자 정보 기억 수단으로부터 취득하는 화자 정보 취득 수단을

을 포함하고,

상기 보유 수단은, 상기 변화점 정보와, 상기 화자 정보 취득 수단에 의해 취득된 화자의 상기 식별 정보를 관련지어 보유하는 것을 특징으로 하는 음성 신호 처리 장치.

#### 청구항 10.

제9항에 있어서,

상기 화자 정보 기억 수단에는, 각 식별 정보에 대응하는 화자에 관련되는 정보가 각 식별 정보에 대응지어 기억되어 있고,

상기 음성 신호에 대한 변화점의 위치와 상기 화자에 관련되는 정보를 표시하는 표시 정보 처리 수단을 포함하는 것을 특징으로 하는 음성 신호 처리 장치.

#### 청구항 11.

음성 신호 처리 방법에 있어서,  
처리 대상의 음성 신호에 기초하여, 해당 음성 신호의 화자의 변화를 소정의 처리 단위마다 검출하는 검출 단계와,  
상기 검출 단계에서, 화자가 변화하였다고 검출한 상기 음성 신호 상의 위치를 나타내는 변화점 정보를 취득하는 취득 단계와,  
상기 취득 단계에서 취득한 상기 변화점 정보를 기록 매체에 저장하는 저장 단계  
를 포함하는 것을 특징으로 하는 음성 신호 처리 방법.

## 청구항 12.

제11항에 있어서,  
상기 검출 단계에서는, 상기 처리 단위마다 상기 음성 신호의 특징을 추출하고, 추출한 상기 음성 신호의 특징에 기초하여, 화음성 이외의 부분으로부터 화음성 부분으로의 변화점 및 화음성 부분의 화자의 변화점을 검출하는 것을 특징으로 하는 음성 신호 처리 방법.

## 청구항 13.

제12항에 있어서,  
상기 검출 단계에서 추출된 상기 음성 신호의 상기 특징과 1명 이상의 화자의 화음성의 특징을 나타내는 특징 정보와 상기 화자의 식별 정보가 대응지어 기억되어 있는 기록 매체의 상기 특징 정보를 비교하여, 화자를 특정하는 특정 단계  
를 포함하고,  
상기 저장 단계에서는, 상기 변화점 정보와, 상기 특정 단계에서 특정한 화자의 상기 식별 정보를 관련지어 상기 기록 매체에 저장하는 것을 특징으로 하는 음성 신호 처리 방법.

## 청구항 14.

제12항에 있어서,  
복수의 마이크로폰의 각각에 대응하는 복수의 음성 채널의 음성 신호를 해석하여 화자 위치를 검출하는 제2 검출 단계  
를 포함하고,  
상기 취득 단계에서는, 상기 제2 검출 단계에서 검출한 화자 위치의 변화도 고려하여, 상기 변화점을 특정하고, 특정한 변화점에 대응하는 상기 변화점 정보를 취득하는 것을 특징으로 하는 음성 신호 처리 방법.

## 청구항 15.

제13항에 있어서,  
복수의 마이크로폰의 각각에 대응하는 복수의 음성 채널의 음성 신호에 따라 정해지는 화자 위치와, 상기 화자 위치의 화자의 식별 정보를 대응지어 미리 화자 정보 기억 수단에서 기억 보유하는 화자 정보 기억 단계  
를 설정하고,  
상기 복수의 음성 채널의 음성 신호를 해석하여 얻어지는 화자 위치에 따른 화자의 상기 식별 정보를 상기 화자 정보 기억 수단으로부터 취득하는 화자 정보 취득 단계  
를 포함하며,

상기 특정 단계에서는, 상기 화자 정보 취득 단계에서 취득한 화자의 상기 식별 정보도 고려하여, 상기 화자를 특정하는 것을 특징으로 하는 음성 신호 처리 방법.

### 청구항 16.

제13항에 있어서,

상기 기억 수단에는, 각 식별 정보에 대응하는 화자에 관련되는 정보가 각 식별 정보에 대응지어 기억되어 있고, 상기 음성 신호에 대한 변화점의 위치와 상기 화자에 관련되는 정보를 표시하는 표시 정보 처리 단계를 포함하는 것을 특징으로 하는 음성 신호 처리 방법.

### 청구항 17.

제11항에 있어서,

상기 검출 단계에서는, 서로 다른 마이크로폰에 의해 집음되는 각 음성 채널의 음성 신호를 해석함으로써 얻어지는 화자 위치에 기초하여, 상기 변화점을 검출하는 것을 특징으로 하는 음성 신호 처리 방법.

### 청구항 18.

제17항에 있어서,

상기 저장 단계에서는, 상기 변화점 정보와, 상기 검출 단계에서 검출되는 화자 위치를 나타내는 정보를 관련지어 저장하는 것을 특징으로 하는 음성 신호 처리 방법.

### 청구항 19.

제17항에 있어서,

복수의 마이크로폰의 각각에 대응하는 복수의 음성 채널의 음성 신호에 따라 정해지는 화자 위치와, 상기 화자 위치의 화자의 식별 정보를 대응하여 미리 화자 정보 기억 수단에 기억 보유하는 화자 정보 기억 단계를

를 설정하고,

상기 복수의 음성 채널의 음성 신호를 해석하여 얻어지는 화자 위치에 따른 화자의 상기 식별 정보를 상기 화자 정보 기억 수단으로부터 취득하는 화자 정보 취득 단계를

를 포함하며,

상기 저장 단계에서는, 상기 변화점 정보와, 상기 화자 정보 취득 단계에서 취득되는 화자의 상기 식별 정보를 관련지어 저장하는 것을 특징으로 하는 음성 신호 처리 방법.

### 청구항 20.

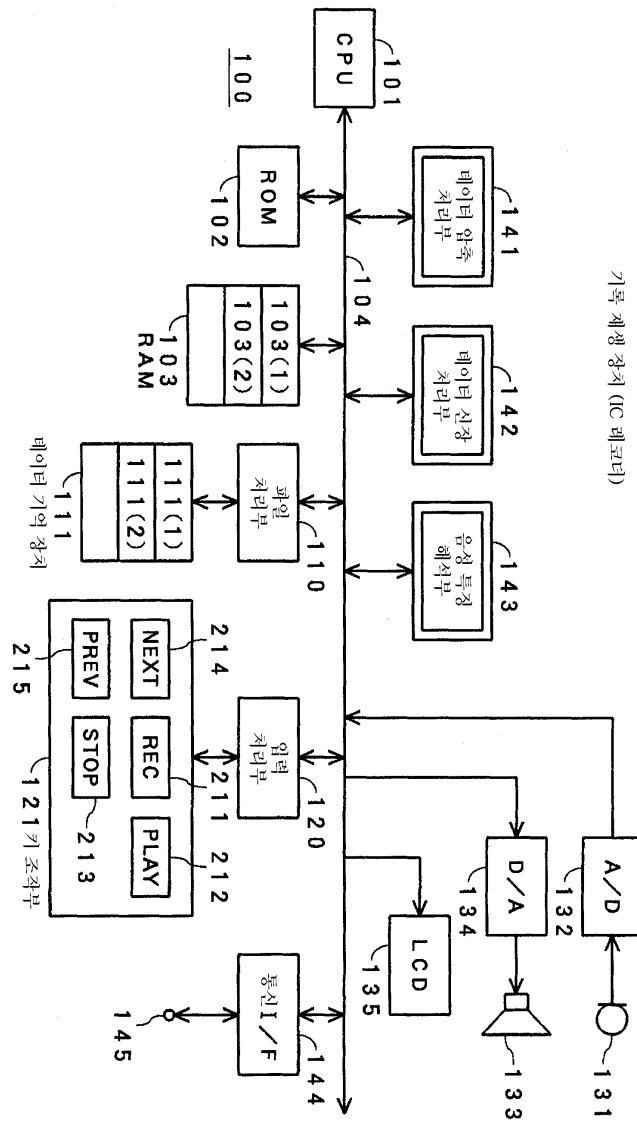
제19항에 있어서,

상기 화자 정보 기억 수단에는, 각 식별 정보에 대응하는 화자에 관련되는 정보가 각 식별 정보에 대응지어 기억되어 있고,

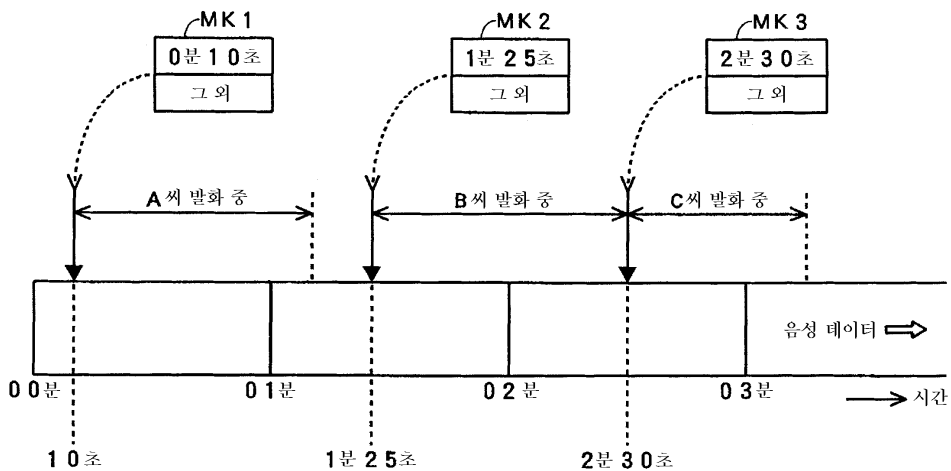
상기 음성 신호에 대한 변화점의 위치와 상기 화자에 관련되는 정보를 표시하는 표시 정보 처리 단계를 포함하는 것을 특징으로 하는 음성 신호 처리 방법.

도면

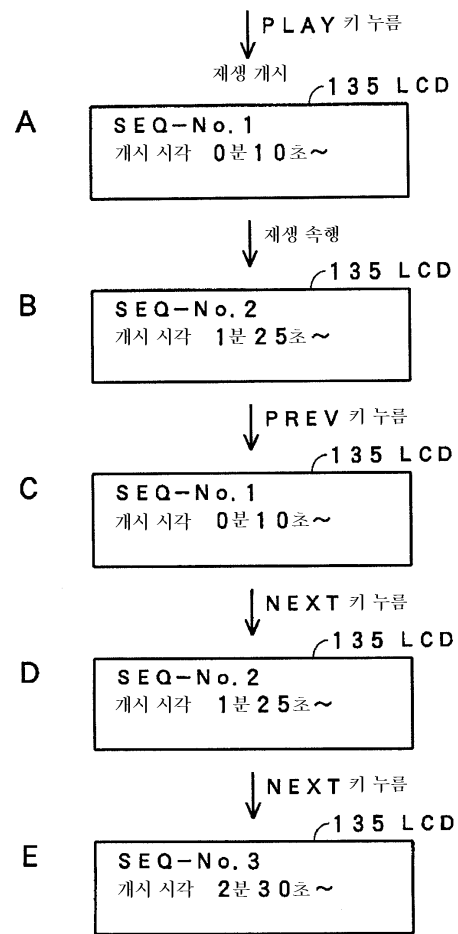
도면1



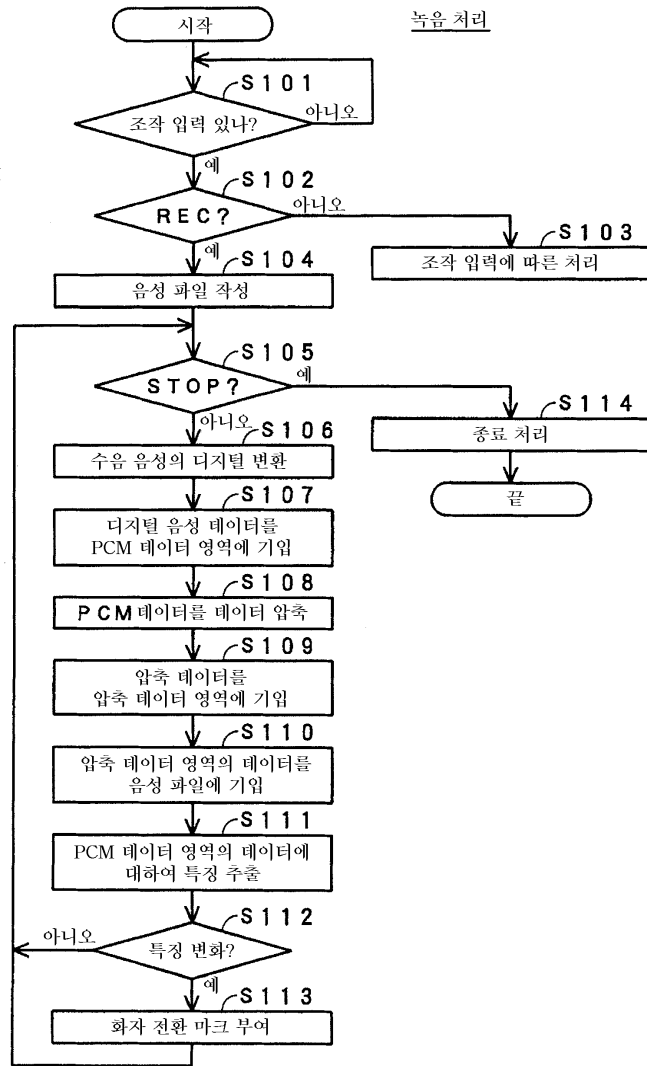
도면2



도면3

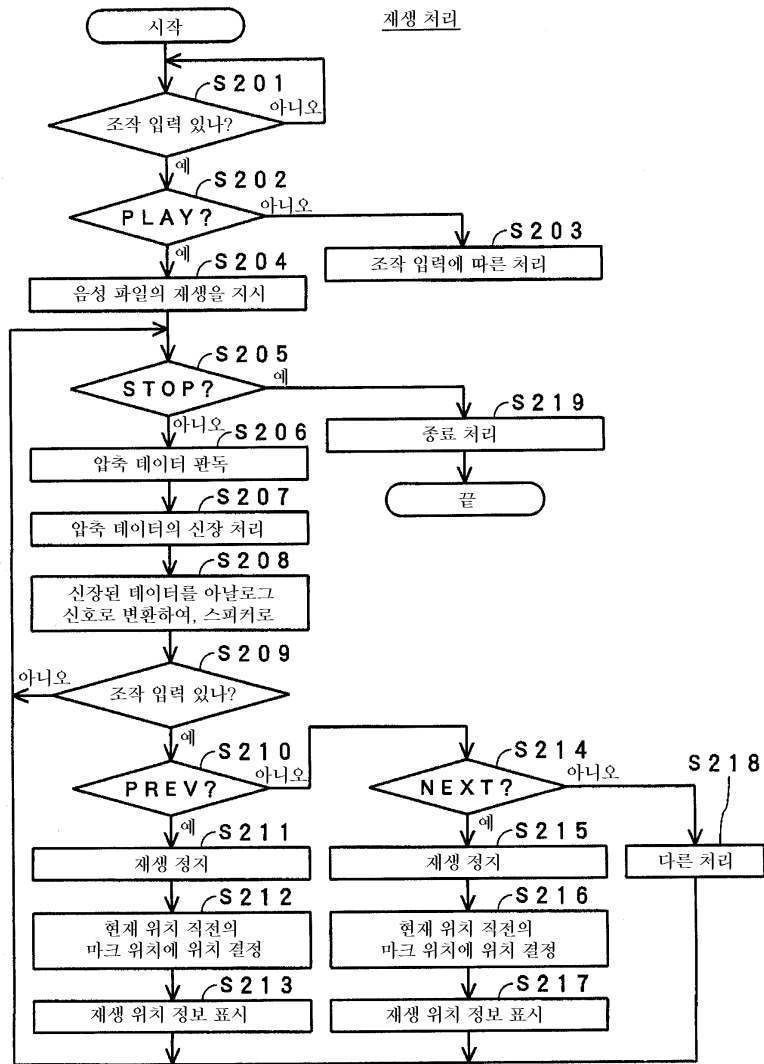


도면4





도면5

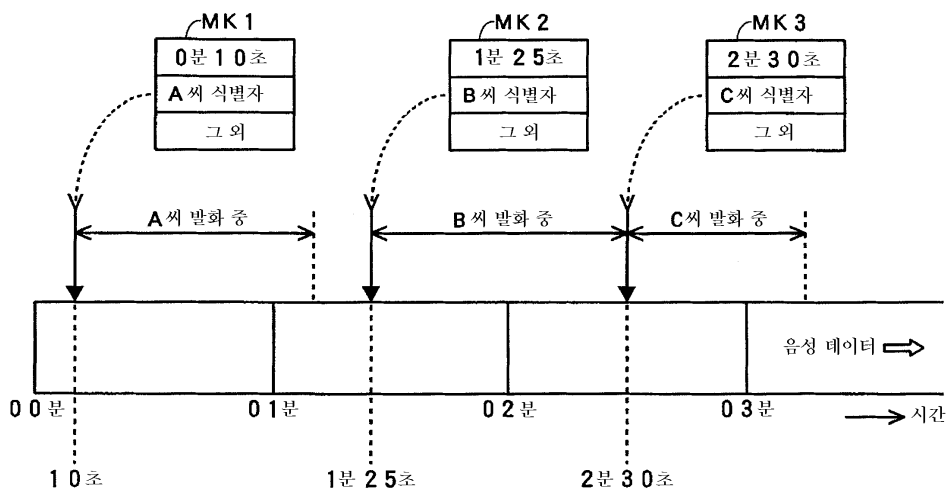


도면6

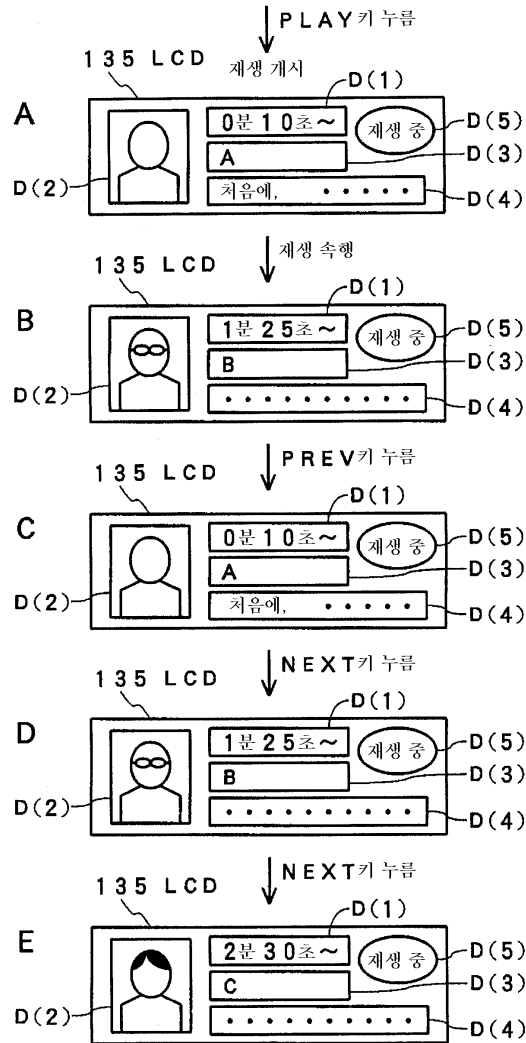
음성 특징 데이터베이스

| 식별자 | 이름  | 특정 해석 결과(성문데이터) | 화상 데이터 | 아이콘 데이터 | 그 외 |
|-----|-----|-----------------|--------|---------|-----|
| 01  | A   | 해석 결과 데이터 FA    | A. jpg | A. gif  |     |
| 02  | B   | 해석 결과 데이터 FB    | B. jpg | B. gif  |     |
| 03  | C   | 해석 결과 데이터 FC    | C. jpg | C. gif  |     |
| ... | ... | ...             | ...    | ...     | ... |

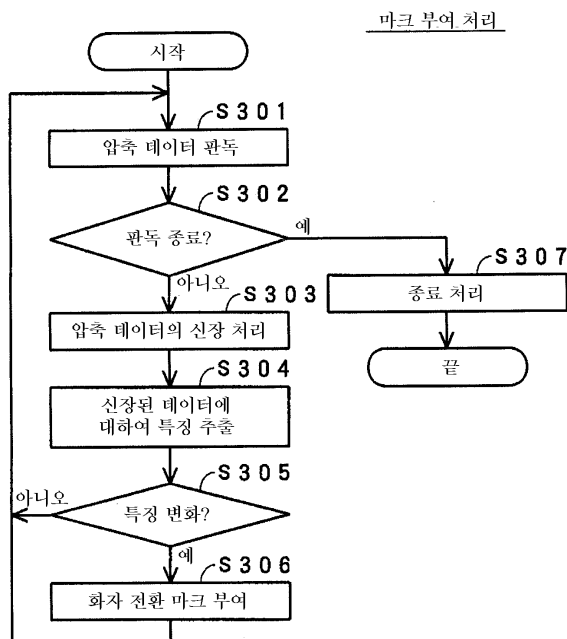
도면7



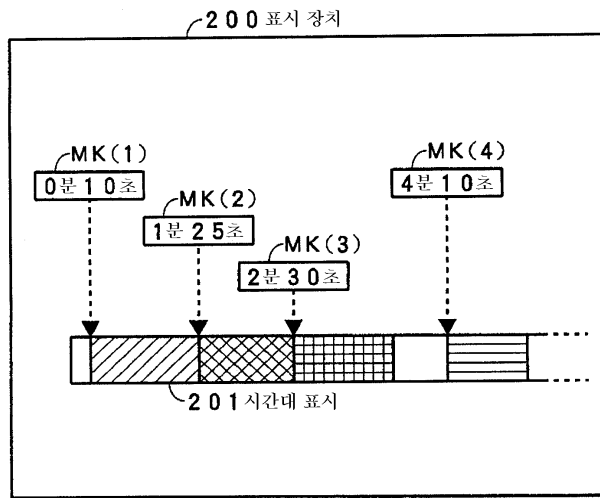
도면8



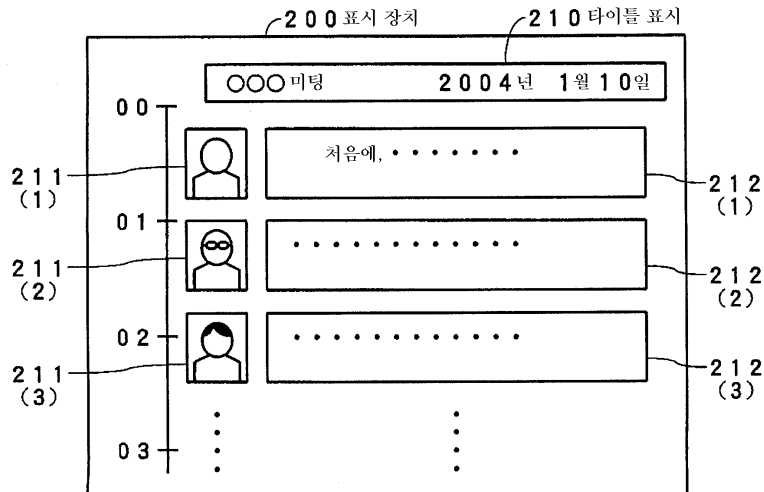
도면9



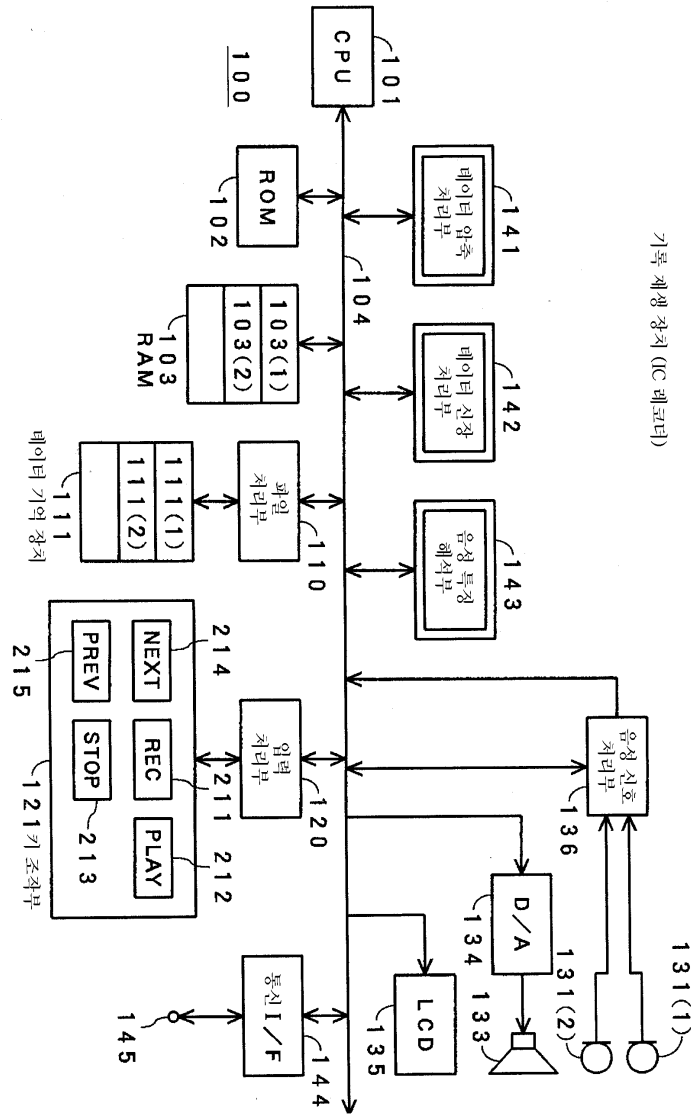
도면10



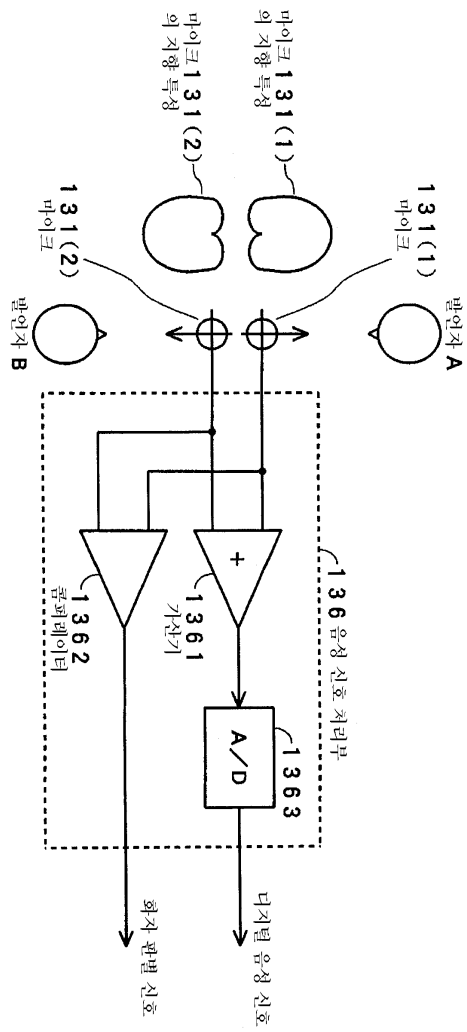
도면11



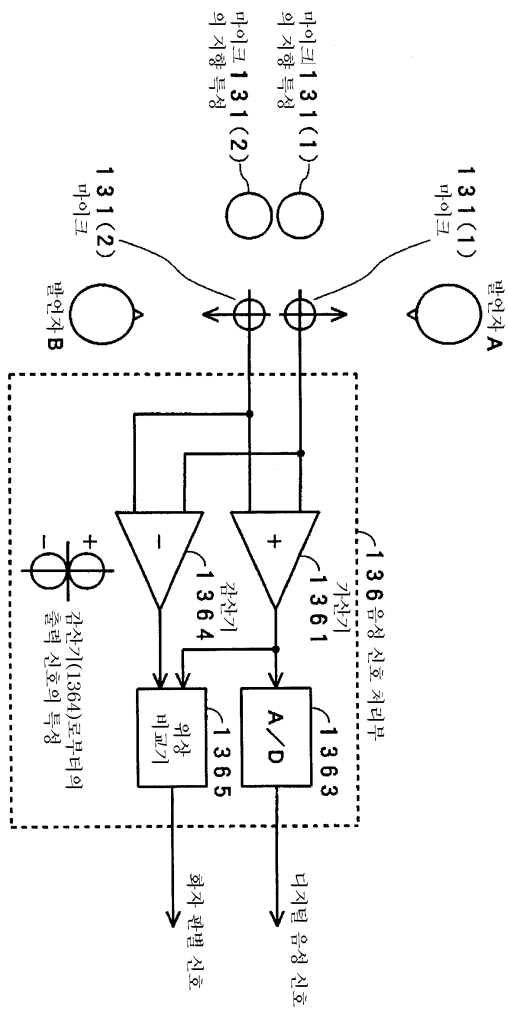
도면12



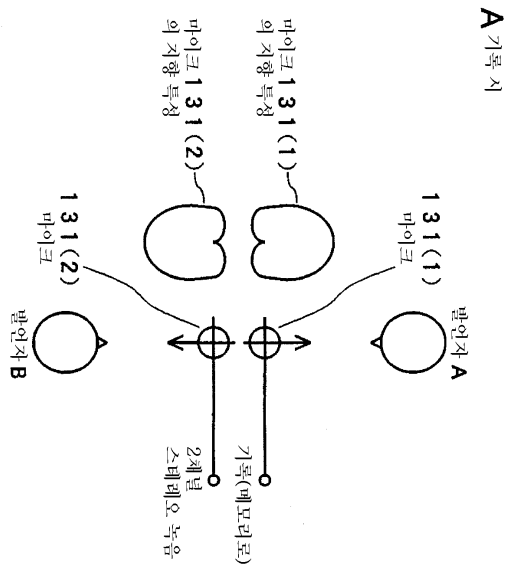
도면13



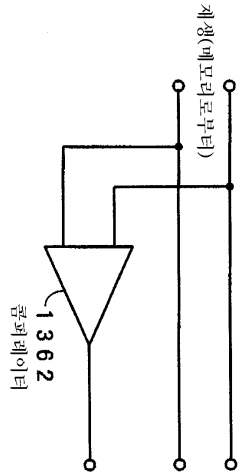
도면14



도면15



B 재생 시



도면16

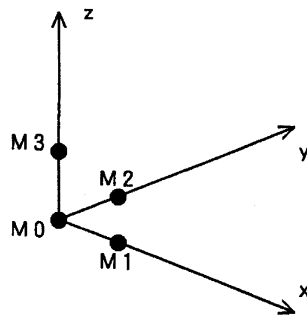
화자 위치 데이터베이스(미이크로폰 위치를 기준)

| 화자 식별 신호 | 미이크로폰            | 화자 식별자 1 | 화자 식별자 2 | 화자 식별자 3 | ..... |
|----------|------------------|----------|----------|----------|-------|
| 1 (+)    | 미이크로폰 1 (131(1)) | 화자 A 1   | 화자 A 2   | 화자 A 3   | ..... |
| 0 (-)    | 미이크로폰 2 (131(2)) | 화자 B 1   | 화자 B 2   | 화자 B 3   | ..... |

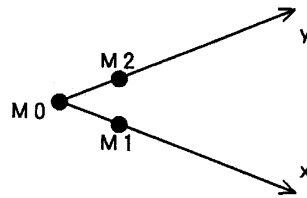


도면17

A 근점 4점법



B 근점 3점법



도면18

