



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2009년02월12일
(11) 등록번호 10-0883712
(24) 등록일자 2009년02월06일

(51) Int. Cl.
H04R 3/00 (2006.01) H04R 1/40 (2006.01)
(21) 출원번호 10-2007-0077162
(22) 출원일자 2007년07월31일
심사청구일자 2007년07월31일
(65) 공개번호 10-2008-0013734
(43) 공개일자 2008년02월13일
(30) 우선권주장
JP-P-2006-00217293 2006년08월09일 일본(JP)
JP-P-2007-00033911 2007년02월14일 일본(JP)
(56) 선행기술조사문헌
JP2004012151 A
JP2003337164 A
JP2005077205 A
KR1020050110790 A

(73) 특허권자
후지쯔 가부시끼가이샤
일본국 가나가와켄 가와사키시 나카하라꾸 가미고
다나카 4초메 1-1
(72) 발명자
하야카와 쇼지
일본 가나가와켄 가와사키시 나카하라꾸 가미코다
나카 4쵸메 1-1후지쯔 가부시끼가이샤 내
(74) 대리인
이중희, 장수길

전체 청구항 수 : 총 8 항

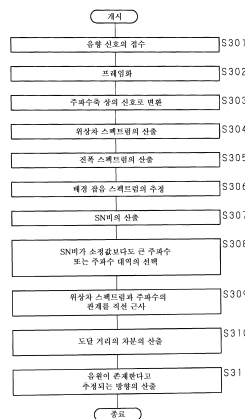
심사관 : 성백두

(54) 음원 방향 추정 방법, 및 음원 방향 추정 장치

(57) 요약

마이크로폰으로부터의 음 입력 중에 주위 잡음이 존재하는 경우라도, 음원의 존재 방향을 고정밀도로 추정할 수 있다. 복수 방향에 존재하는 음원으로부터의 음향 신호가 복수 채널의 입력으로서 접수되고(S301), 주파수축 상의 신호로 변환 된다(S303). 변환된 주파수축 상의 신호의 위상 성분이 동일 주파수마다 산출되고, 복수 채널간의 위상차분이 산출된다(S304). 한편, 변환된 주파수축 상의 신호의 진폭 성분이 산출되고(S305), 산출된 진폭 성분으로부터 잡음 성분이 추정된다(S306). 진폭 성분 및 잡음 성분에 기초하여 주파수마다의 SN비가 산출되고(S307), SN비가 소정값보다도 큰 주파수가 선택된다(S308). 선택된 주파수의 위상차분에 기초하여 도달 거리의 차분이 산출되고(S310), 목적으로 하는 음원이 존재한다고 추정되는 방향이 산출된다(S311).

대표도 - 도3



특허청구의 범위

청구항 1

복수 방향에 존재하는 음원으로부터의 음향 신호를 복수 채널의 입력으로서 입력받는 음향 신호 입력부에 입력된 음향 신호가 존재하는 방향을 추정하는 음원 방향 추정 방법에 있어서,

상기 음향 신호 입력부에 의해 입력된 복수 채널의 입력을 접수하여, 채널마다의 시간축 상의 신호로 변환하는 스텝과,

시간축 상의 각 채널의 신호를 주파수축 상의 신호로 변환하는 스텝과,

변환된 주파수축 상의 각 채널의 신호의 위상 성분을 동일 주파수마다 산출하는 스텝과,

동일 주파수마다 산출된 각 채널의 신호의 위상 성분을 이용하여, 복수 채널간의 위상차분을 산출하는 스텝과,

변환된 주파수축 상의 신호의 진폭 성분을 산출하는 스텝과,

산출된 진폭 성분으로부터 잡음 성분을 추정하는 스텝과,

산출된 진폭 성분 및 추정된 잡음 성분에 기초하여 주파수마다의 신호 대 잡음비를 산출하는 스텝과,

신호 대 잡음비가 소정값보다도 큰 주파수를 추출하는 스텝과,

추출된 주파수의 산출된 위상차분에 기초하여, 목적으로 하는 음원으로부터의 음향 신호의 도달 거리의 차분을 산출하는 스텝과,

산출된 도달 거리의 차분에 기초하여, 목적으로 하는 음원이 존재하는 방향을 추정하는 스텝

을 포함하는 것을 특징으로 하는 음원 방향 추정 방법.

청구항 2

제1항에 있어서,

상기 주파수를 추출하는 스텝은, 신호 대 잡음비가 소정값보다도 큰 주파수를 산출된 신호 대 잡음비의 내림차순으로 소정수 선택하여 추출하는 것을 특징으로 하는 음원 방향 추정 방법.

청구항 3

복수 방향에 존재하는 음원으로부터의 음향 신호를 복수 채널의 입력으로서 입력받는 음향 신호 입력부에 입력된 음향 신호의 음원이 존재하는 방향을 추정하는 음원 방향 추정 방법에 있어서,

상기 음향 신호 입력부에 의해 입력된 복수 채널의 입력을 접수하여, 채널마다의 시간축 상의 샘플링 신호로 변환하는 스텝과,

시간축 상의 각 샘플링 신호를 주파수축 상의 신호로 채널마다 변환하는 스텝과,

변환된 주파수축 상의 각 채널의 신호의 위상 성분을 동일 주파수마다 산출하는 스텝과,

동일 주파수마다 산출된 각 채널의 신호의 위상 성분을 이용하여, 복수 채널간의 위상차분을 산출하는 스텝과,

소정의 샘플링 시점에서 변환된 주파수축 상의 신호의 진폭 성분을 산출하는 스텝과,

산출된 진폭 성분으로부터 잡음 성분을 추정하는 스텝과,

산출된 진폭 성분 및 추정된 잡음 성분에 기초하여 주파수마다의 신호 대 잡음비를 산출하는 스텝과,

산출된 신호 대 잡음비 및 과거의 샘플링 시점에서의 위상차분의 산출 결과에 기초하여, 샘플링 시점에서의 위상차분의 산출 결과를 보정하는 스텝과,

보정 후의 위상차분에 기초하여, 목적으로 하는 음원으로부터의 음향 신호의 도달 거리의 차분을 산출하는 스텝과,

산출된 도달 거리의 차분에 기초하여, 목적으로 하는 음원이 존재하는 방향을 추정하는 스텝

을 포함하는 것을 특징으로 하는 음원 방향 추정 방법.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서,

접수한 음향 신호 입력 중의 음성을 나타내는 구간인 음성 구간을 특정하는 스텝을 더 포함하고,

상기 주파수축 상의 신호로 변환하는 스텝은, 상기 음성 구간을 특정하는 스텝에서 특정된 음성 구간의 신호만을 주파수축 상의 신호로 변환하는 것을 특징으로 하는 음원 방향 추정 방법.

청구항 5

복수 방향에 존재하는 음원으로부터의 음향 신호를 복수 채널의 입력으로서 입력받는 음향 신호 입력 수단에 입력된 음향 신호가 존재하는 방향을 추정하는 음원 방향 추정 장치에 있어서,

상기 음향 신호 입력 수단에 의해 입력된 복수 채널의 음향 신호를 접수하여, 채널마다의 시간축 상의 신호로 변환하는 음향 신호 접수 수단과,

상기 음향 신호 접수 수단에 의해 변환된 시간축 상의 각 신호를 주파수축 상의 신호로 채널마다 변환하는 신호 변환 수단과,

상기 신호 변환 수단에 의해 변환된 주파수축 상의 각 채널의 신호의 위상 성분을 동일 주파수마다 산출하는 위상 성분 산출 수단과,

상기 위상 성분 산출 수단에 의해 동일 주파수마다 산출된 각 채널의 신호의 위상 성분을 이용하여, 복수 채널 간의 위상차분을 산출하는 위상차분 산출 수단과,

상기 신호 변환 수단에 의해 변환된 주파수축 상의 신호의 진폭 성분을 산출하는 진폭 성분 산출 수단과,

상기 진폭 성분 산출 수단에 의해 산출된 진폭 성분으로부터 잡음 성분을 추정하는 잡음 성분 추정 수단과,

상기 진폭 성분 산출 수단에 의해 산출된 진폭 성분 및 상기 잡음 성분 추정 수단에 의해 추정된 잡음 성분에 기초하여 주파수마다의 신호 대 잡음비를 산출하는 신호 대 잡음비 산출 수단과,

상기 신호 대 잡음비 산출 수단에 의해 산출된 신호 대 잡음비가 소정값보다도 큰 주파수를 추출하는 주파수 추출 수단과,

상기 주파수 추출 수단에 의해 추출된 주파수의 상기 위상차분 산출 수단에 의해 산출된 위상차분에 기초하여, 목적으로 하는 음원으로부터의 음향 신호의 도달 거리의 차분을 산출하는 도달 거리 차분 산출 수단과,

상기 도달 거리 차분 산출 수단에 의해 산출된 도달 거리의 차분에 기초하여, 목적으로 하는 음원이 존재하는 방향을 추정하는 음원 방향 추정 수단

을 구비한 것을 특징으로 하는 음원 방향 추정 장치.

청구항 6

제5항에 있어서,

상기 주파수 추출 수단은, 상기 신호 대 잡음비 산출 수단에 의해 산출된 신호 대 잡음비가 소정값보다도 큰 주파수를, 산출된 신호 대 잡음비의 내림차순으로 소정수 선택하여 추출하도록 하고 있는 것을 특징으로 하는 음원 방향 추정 장치.

청구항 7

복수 방향에 존재하는 음원으로부터의 음향 신호를 복수 채널의 입력으로서 입력받는 음향 신호 입력 수단에 입력된 음향 신호의 음원이 존재하는 방향을 추정하는 음원 방향 추정 장치에 있어서,

상기 음향 신호 입력 수단에 의해 입력된 복수 채널의 음향 신호를 접수하여, 채널마다의 시간축 상의 샘플링 신호로 변환하는 음향 신호 접수 수단과,

상기 음향 신호 접수 수단에 의해 변환된 시간축 상의 각 샘플링 신호를 주파수축 상의 신호로 채널마다 변환하

는 신호 변환 수단과,

상기 신호 변환 수단에 의해 변환된 주파수축 상의 각 채널의 신호의 위상 성분을 동일 주파수마다 산출하는 위상 성분 산출 수단과,

상기 위상 성분 산출 수단에 의해 동일 주파수마다 산출된 각 채널의 신호의 위상 성분을 이용하여, 복수 채널 간의 위상차분을 산출하는 위상차분 산출 수단과,

상기 신호 변환 수단에 의해 소정의 샘플링 시점에서 변환된 주파수축 상의 신호의 진폭 성분을 산출하는 진폭 성분 산출 수단과,

상기 진폭 성분 산출 수단에 의해 산출된 진폭 성분으로부터 잡음 성분을 추정하는 잡음 성분 추정 수단과,

상기 진폭 성분 산출 수단에 의해 산출된 진폭 성분 및 상기 잡음 성분 추정 수단에 의해 추정된 잡음 성분에 기초하여 주파수마다의 신호 대 잡음비를 산출하는 신호 대 잡음비 산출 수단과,

상기 신호 대 잡음비 산출 수단에 의해 산출된 신호 대 잡음비 및 과거의 샘플링 시점에서의 위상차분의 산출 결과에 기초하여, 샘플링 시점에서의 위상차분의 산출 결과를 보정하는 보정 수단과,

상기 보정 수단에 의한 보정 후의 상기 위상차분 산출 수단에 의해 산출된 위상차분에 기초하여, 목적으로 하는 음원으로부터의 음향 신호의 도달 거리의 차분을 산출하는 도달 거리 차분 산출 수단과,

상기 도달 거리 차분 산출 수단에 의해 산출된 도달 거리의 차분에 기초하여, 목적으로 하는 음원이 존재하는 방향을 추정하는 음원 방향 추정 수단

을 구비한 것을 특징으로 하는 음원 방향 추정 장치.

청구항 8

제5항 내지 제7항 중 어느 한 항에 있어서,

상기 음향 신호 접수 수단에 의해 접수된 음향 신호 입력 중의 음성을 나타내는 구간인 음성 구간을 특징하는 음성 구간 특정 수단을 더 구비하고,

상기 신호 변환 수단은, 상기 음성 구간 특정 수단에 의해 특정된 음성 구간의 신호만을 주파수축 상의 신호로 변환하도록 하고 있는 것을 특징으로 하는 음원 방향 추정 장치.

명세서

발명의 상세한 설명

기술분야

- <1> 본 발명은, 복수의 마이크로폰을 이용하여, 주위 잡음이 존재하는 경우라도, 음원으로부터의 음 입력의 도래 방향을 고정밀도로 추정할 수 있는 음원 방향 추정 방법, 및 음원 방향 추정 장치에 관한 것이다.

배경기술

- <2> 최근의 컴퓨터 기술의 진전에 의해, 대량의 연산 처리를 필요로 하는 음향 신호 처리라도 실용적인 처리 속도로 실행할 수 있도록 되어 있었다. 이와 같은 사정으로부터, 복수의 마이크로폰을 이용한 멀티 채널의 음향 처리 기능의 실용화가 기대되고 있다. 그 일례가, 음향 신호의 도래 방향을 추정하는 음원 방향 추정 처리이다. 음원 방향 추정 처리는, 복수의 마이크로폰을 설치해 놓고, 목적으로 하는 음원으로부터의 음향 신호가 2개의 마이크로폰에 도달하였을 때의 지연 시간을 구하고, 마이크로폰간의 도달 거리의 차 및 마이크로폰의 설치 간격에 기초하여, 음원으로부터의 음향 신호의 도래 방향을 추정하는 처리이다.
- <3> 종래의 음원 방향 추정 처리는, 예를 들면 2개의 마이크로폰으로부터 입력된 신호간의 상호 상관을 산출하고, 상호 상관이 최대로 되는 시간에서의 2개의 신호간의 지연 시간을 산출한다. 산출된 지연 시간에, 상온의 공기 중에서의 음의 전파 속도인 약 340m/s(온도에 따라서 변화됨)를 승산함으로써 도달 거리차가 구해지므로, 마이크로폰의 설치 간격으로부터 삼각법에 따라서 음향 신호의 도래 방향이 산출된다.

- <4> 또한, 특허 문헌1에 개시되어 있는 바와 같이, 2개의 마이크로폰으로부터 입력된 음향 신호의 주파수마다의 위상차 스펙트럼을 산출하고, 주파수 베이스에 직선 근사한 경우의 위상차 스펙트럼의 기울기에 기초하여, 음원으로부터의 음향 신호의 도래 방향을 산출하는 것도 가능하다.
- <5> <종래기술의 문헌 정보>
- <6> [특허 문헌1] 일본 특개 2003-337164호 공보

발명의 내용

해결 하고자하는 과제

- <7> 전술한 종래의 음원 방향 추정 방법에서는, 잡음이 중첩된 경우에는 상호 상관이 최대가 되는 시간을 특정하는 것 자체가 곤란하다. 이것은, 음원으로부터의 음향 신호의 도래 방향을 올바르게 특정하는 것이 곤란하게 된다고 하는 문제점을 초래한다. 또한, 특허 문헌 1에 개시되어 있는 방법이라도, 위상차 스펙트럼을 산출할 때에, 잡음이 중첩되어 있는 경우에는 위상차 스펙트럼이 심하게 변동되므로, 위상차 스펙트럼의 기울기를 정확하게 구할 수 없다고 하는 문제점이 있었다.
- <8> 본 발명은, 이상과 같은 사정을 감안하여 이루어진 것으로, 마이크로폰의 주변에 주위 잡음이 존재하는 경우라도, 목적으로 하는 음원으로부터의 음향 신호의 도래 방향을 고정밀도로 추정할 수 있는 음원 방향 추정 장치, 및 음원 방향 추정 방법을 제공하는 것을 목적으로 한다.

과제 해결수단

- <9> 상기 목적을 달성하기 위해 제1 발명에 따른 음원 방향 추정 방법은, 복수 방향에 존재하는 음원으로부터의 음향 신호를 복수 채널의 입력으로서 입력받는 음향 신호 입력부에 입력된 음향 신호가 존재하는 방향을 추정하는 음원 방향 추정 방법으로서, 상기 음향 신호 입력부에 의해 입력된 복수 채널의 입력을 접수하여, 채널마다의 시간축 상의 신호로 변환하는 스텝과, 시간축 상의 각 채널의 신호를 주파수축 상의 신호로 변환하는 스텝과, 변환된 주파수축 상의 각 채널의 신호의 위상 성분을 동일 주파수마다 산출하는 스텝과, 동일 주파수마다 산출된 각 채널의 신호의 위상 성분을 이용하여, 복수 채널간의 위상차분을 산출하는 스텝과, 변환된 주파수축 상의 신호의 진폭 성분을 산출하는 스텝과, 산출된 진폭 성분으로부터 잡음 성분을 추정하는 스텝과, 산출된 진폭 성분 및 추정된 잡음 성분에 기초하여 주파수마다의 신호 대 잡음비를 산출하는 스텝과, 신호 대 잡음비가 소정값보다도 큰 주파수를 추출하는 스텝과, 추출된 산출된 위상차분에 기초하여, 목적으로 하는 음원으로부터의 음향 신호의 도달 거리의 차분을 산출하는 스텝과, 산출된 도달 거리의 차분에 기초하여, 목적으로 하는 음원이 존재하는 방향을 추정하는 스텝을 포함하는 것을 특징으로 한다.
- <10> 또한, 제1 발명에 따른 음원 방향 추정 장치는, 복수 방향에 존재하는 음원으로부터의 음향 신호를 복수 채널의 입력으로서 입력받는 음향 신호 입력 수단에 입력된 음향 신호가 존재하는 방향을 추정하는 음원 방향 추정 장치로서, 상기 음향 신호 입력 수단에 의해 입력된 복수 채널의 음향 신호를 접수하여, 채널마다의 시간축 상의 신호로 변환하는 음향 신호 접수 수단과, 상기 음향 신호 접수 수단에 의해 변환된 시간축 상의 각 신호를 주파수축 상의 신호로 채널마다 변환하는 신호 변환 수단과, 상기 신호 변환 수단에 의해 변환된 주파수축 상의 각 채널의 신호의 위상 성분을 동일 주파수마다 산출하는 위상 성분 산출 수단과, 상기 위상 성분 산출 수단에 의해 동일 주파수마다 산출된 각 채널의 신호의 위상 성분을 이용하여, 복수 채널간의 위상차분을 산출하는 위상차분 산출 수단과, 상기 신호 변환 수단에 의해 변환된 주파수축 상의 신호의 진폭 성분을 산출하는 진폭 성분 산출 수단과, 상기 진폭 성분 산출 수단에 의해 산출된 진폭 성분으로부터 잡음 성분을 추정하는 잡음 성분 추정 수단과, 상기 진폭 성분 산출 수단에 의해 산출된 진폭 성분 및 상기 잡음 성분 추정 수단에 의해 추정된 잡음 성분에 기초하여 주파수마다의 신호 대 잡음비를 산출하는 신호 대 잡음비 산출 수단과, 상기 신호 대 잡음비 산출 수단에 의해 산출된 신호 대 잡음비가 소정값보다도 큰 주파수를 추출하는 주파수 추출 수단과, 상기 주파수 추출 수단에 의해 추출된 주파수의 상기 위상차분 산출 수단에 의해 산출된 위상차분에 기초하여, 목적으로 하는 음원으로부터의 음향 신호의 도달 거리의 차분을 산출하는 도달 거리 차분 산출 수단과, 상기 도달 거리 차분 산출 수단에 의해 산출된 도달 거리의 차분에 기초하여, 목적으로 하는 음원이 존재하는 방향을 추정하는 음원 방향 추정 수단을 구비한 것을 특징으로 한다.
- <11> 또한, 제2 발명에 따른 음원 방향 추정 방법은, 제1 발명에서, 상기 주파수를 추출하는 스텝은, 신호 대 잡음비가 소정값보다도 큰 주파수를 산출된 신호 대 잡음비의 내림차순으로 소정수 선택하여 추출하는 것을 특징으로

한다.

- <12> 또한, 제2 발명에 따른 음원 방향 추정 장치는, 제1 발명에서, 상기 주파수 추출 수단은, 상기 신호 대 잡음비 산출 수단에 의해 산출된 신호 대 잡음비가 소정값보다도 큰 주파수를, 산출된 신호 대 잡음비의 내림차순으로 소정수 선택하여 추출하도록 하고 있는 것을 특징으로 한다.
- <13> 또한, 제3 발명에 따른 음원 방향 추정 방법은, 복수 방향에 존재하는 음원으로부터의 음향 신호를 복수 채널의 입력으로서 입력받는 음향 신호 입력부에 입력된 음향 신호의 음원이 존재하는 방향을 추정하는 음원 방향 추정 방법으로서, 상기 음향 신호 입력부에 의해 입력된 복수 채널의 입력을 접수하여, 채널마다의 시간축 상의 샘플링 신호로 변환하는 스텝과, 시간축 상의 각 샘플링 신호를 주파수축 상의 신호로 채널마다 변환하는 스텝과, 변환된 주파수축 상의 각 채널의 신호의 위상 성분을 동일 주파수마다 산출하는 스텝과, 동일 주파수마다 산출된 각 채널의 신호의 위상 성분을 이용하여, 복수 채널간의 위상차분을 산출하는 스텝과, 소정의 샘플링 시점에서 변환된 주파수축 상의 신호의 진폭 성분을 산출하는 스텝과, 산출된 진폭 성분으로부터 잡음 성분을 추정하는 스텝과, 산출된 진폭 성분 및 추정된 잡음 성분에 기초하여 주파수마다의 신호 대 잡음비를 산출하는 스텝과, 산출된 신호 대 잡음비 및 과거의 샘플링 시점에서의 위상차분의 산출 결과에 기초하여, 샘플링 시점에서의 위상차분의 산출 결과를 보정하는 스텝과, 보정 후의 위상차분에 기초하여, 목적으로 하는 음원으로부터의 음향 신호의 도달 거리의 차분을 산출하는 스텝과, 산출된 도달 거리의 차분에 기초하여, 목적으로 하는 음원이 존재하는 방향을 추정하는 스텝을 포함하는 것을 특징으로 한다.
- <14> 또한, 제3 발명에 따른 음원 방향 추정 장치는, 복수 방향에 존재하는 음원으로부터의 음향 신호를 복수 채널의 입력으로서 입력받는 음향 신호 입력 수단에 입력된 음향 신호의 음원이 존재하는 방향을 추정하는 음원 방향 추정 장치로서, 상기 음향 신호 입력 수단에 의해 입력된 복수 채널의 음향 신호를 접수하여, 채널마다의 시간축 상의 샘플링 신호로 변환하는 음향 신호 접수 수단과, 상기 음향 신호 접수 수단에 의해 변환된 시간축 상의 각 샘플링 신호를 주파수축 상의 신호로 채널마다 변환하는 신호 변환 수단과, 상기 신호 변환 수단에 의해 변환된 주파수축 상의 각 채널의 신호의 위상 성분을 동일 주파수마다 산출하는 위상 성분 산출 수단과, 상기 위상 성분 산출 수단에 의해 동일 주파수마다 산출된 각 채널의 신호의 위상 성분을 이용하여, 복수 채널간의 위상차분을 산출하는 위상차분 산출 수단과, 상기 신호 변환 수단에 의해 소정의 샘플링 시점에서 변환된 주파수축 상의 신호의 진폭 성분을 산출하는 진폭 성분 산출 수단과, 상기 진폭 성분 산출 수단에 의해 산출된 진폭 성분으로부터 잡음 성분을 추정하는 잡음 성분 추정 수단과, 상기 진폭 성분 산출 수단에 의해 산출된 진폭 성분 및 상기 잡음 성분 추정 수단에 의해 추정된 잡음 성분에 기초하여 주파수마다의 신호 대 잡음비를 산출하는 신호 대 잡음비 산출 수단과, 상기 신호 대 잡음비 산출 수단에 의해 산출된 신호 대 잡음비 및 과거의 샘플링 시점에서의 위상차분의 산출 결과에 기초하여, 샘플링 시점에서의 위상차분의 산출 결과를 보정하는 보정 수단과, 상기 보정 수단에 의한 보정 후의 상기 위상차분 산출 수단에 의해 산출된 위상차분에 기초하여, 목적으로 하는 음원으로부터의 음향 신호의 도달 거리의 차분을 산출하는 도달 거리 차분 산출 수단과, 상기 도달 거리 차분 산출 수단에 의해 산출된 도달 거리의 차분에 기초하여, 목적으로 하는 음원이 존재하는 방향을 추정하는 음원 방향 추정 수단을 구비한 것을 특징으로 한다.
- <15> 또한, 제4 발명에 따른 음원 방향 추정 방법은, 제1 내지 제3 발명 중 어느 하나에서, 접수한 음향 신호 입력 중의 음성을 나타내는 구간인 음성 구간을 특징하는 스텝을 더 포함하고, 상기 주파수축 상의 신호로 변환하는 스텝은, 상기 음성 구간을 특징하는 스텝에서 특정된 음성 구간의 신호만을 주파수축 상의 신호로 변환하는 것을 특징으로 한다.
- <16> 또한, 제4 발명에 따른 음원 방향 추정 장치는, 제1 내지 제3 발명 중 어느 하나에서, 상기 음향 신호 접수 수단에 의해 접수된 음향 신호 입력 중의 음성을 나타내는 구간인 음성 구간을 특징하는 음성 구간 특정 수단을 더 구비하고, 상기 신호 변환 수단은, 상기 음성 구간 특정 수단에 의해 특정된 음성 구간의 신호만을 주파수축 상의 신호로 변환하도록 하고 있는 것을 특징으로 한다.
- <17> 제1 발명, 및 제5 발명에서는, 복수 방향에 존재하는 음원으로부터의 음향 신호가 복수 채널의 입력으로서 접수되어, 채널마다의 시간축 상의 신호로 변환된다. 또한, 시간축 상의 각 채널의 신호가 주파수축 상의 신호로 변환되고, 변환된 주파수축 상의 각 채널의 신호의 위상 성분이 이용됨으로써, 복수 채널간의 위상차분이 주파수마다 산출된다. 산출된 위상차분(이하, 위상차 스펙트럼이라고 함)에 기초하여, 목적으로 하는 음원으로부터의 음 입력의 도달 거리의 차분이 산출되고, 산출된 도달 거리의 차분에 기초하여, 음원이 존재하는 방향이 추정된다. 한편, 변환된 주파수축 상의 신호의 진폭 성분이 산출되고, 산출된 진폭 성분으로부터 배경 잡음 성분이 추정된다. 산출된 진폭 성분 및 추정된 배경 잡음 성분에 기초하여 주파수마다의 신호 대 잡음비가 산출된

다. 그리고, 신호 대 잡음비가 소정값보다도 큰 주파수가 추출되고, 추출된 주파수의 위상차분에 기초하여 도달 거리의 차분이 산출된다. 이 결과, 입력된 음향 신호의 진폭 성분, 소위 진폭 스펙트럼과, 추정된 배경 잡음 성분, 소위 배경 잡음 스펙트럼에 기초하여 주파수마다의 신호 대 잡음비(SN비 : Signal-to-Noise ratio)가 구해지므로, 신호 대 잡음비가 큰 주파수에서의 위상차분만이 이용됨으로써, 보다 정확한 도달 거리의 차분을 구할 수 있다. 따라서, 정밀도가 높은 도달 거리의 차분에 기초하여 음향 신호의 입사각, 즉 음원이 존재하는 방향을 고정밀도로 추정하는 것이 가능하게 된다.

<18> 제2 발명에서는, 신호 대 잡음비가 소정값보다도 큰 주파수가 신호 대 잡음비의 내림차순으로 소정수 선택되어 추출된다. 이 결과, 잡음 성분의 영향 정도가 작은 주파수가 샘플링되어 도달 거리의 차분이 산출되므로, 도달 거리의 차분의 산출 결과가 크게 변동되지 않는다. 따라서, 보다 고정밀도로 음향 신호의 입사각, 즉 목적으로 하는 음원이 존재하는 방향을 추정하는 것이 가능하게 된다.

<19> 제3 발명, 및 제6 발명에서는, 복수 방향에 존재하는 음원으로부터의 음향 신호가 복수 채널의 입력으로서 접수되어, 채널마다의 시간축 상의 샘플링 신호로 변환되고, 시간축 상의 각 샘플링 신호가 주파수축 상의 신호로 채널마다 변환된다. 변환된 주파수축 상의 각 채널의 신호의 위상 성분이 이용됨으로써, 복수 채널간의 위상차분이 주파수마다 산출된다. 산출된 위상차분에 기초하여, 목적으로 하는 음원으로부터의 음 입력의 도달 거리의 차분이 산출되고, 산출된 도달 거리의 차분에 기초하여, 목적으로 하는 음원이 존재하는 방향이 추정된다. 소정의 샘플링 시점에서 변환된 주파수축 상의 신호의 진폭 성분이 산출되고, 산출된 진폭 성분으로부터 배경 잡음 성분이 추정된다. 산출된 진폭 성분 및 추정된 배경 잡음 성분에 기초하여 주파수마다의 신호 대 잡음비가 산출된다. 그리고, 산출된 신호 대 잡음비와 과거의 샘플링 시점에서의 위상차분의 산출 결과에 기초하여, 샘플링 시점에서의 위상차분의 산출 결과가 보정되고, 보정 후의 위상차분에 기초하여 도달 거리의 차분이 산출된다. 이 결과, 과거의 샘플링 시점에서의 신호 대 잡음비가 큰 주파수에서의 위상차분의 정보를 반영시킨 위상차 스펙트럼을 얻을 수 있다. 이 때문에, 배경 잡음의 상태, 목적으로 하는 음원으로부터 발하여지는 음향 신호의 내용의 변화 등에 의해 위상차분이 크게 변동되지 않는다. 따라서, 보다 정밀도가 높고 안정된 도달 거리의 차분에 기초하여 음향 신호의 입사각, 즉 목적으로 하는 음원이 존재하는 방향을 고정밀도로 추정하는 것이 가능하게 된다.

<20> 제4 발명에서는, 접수된 음향 신호 중의 음성을 나타내는 구간인 음성 구간이 특정되고, 특정된 음성 구간의 신호만이 주파수축 상의 신호로 변환된다. 이 결과, 음성을 발하는 음원이 존재하는 방향을 고정밀도로 추정하는 것이 가능하게 된다.

효 과

<21> 제1 발명 및 제5 발명에 따르면, 입력된 음향 신호의 진폭 성분, 소위 진폭 스펙트럼과, 추정된 배경 잡음 스펙트럼에 기초하여 주파수마다의 신호 대 잡음비(SN비)가 구해지고, 신호 대 잡음비가 큰 주파수에서의 위상차분(위상차 스펙트럼)만을 이용함으로써, 보다 정확한 도달 거리의 차분을 구할 수 있다. 따라서, 정밀도가 높은 도달 거리의 차분에 기초하여 음향 신호의 입사각, 즉 음원이 존재하는 방향을 고정밀도로 추정하는 것이 가능하게 된다.

<22> 제2 발명에 따르면, 잡음 성분의 영향의 정도가 작은 주파수를 우선적으로 선택함으로써 도달 거리의 차분이 산출되므로, 도달 거리의 차분의 산출 결과가 크게 변동되지 않는다. 따라서, 보다 고정밀도로 음향 신호의 입사각, 즉 목적으로 하는 음원이 존재하는 방향을 고정밀도로 추정하는 것이 가능하게 된다.

<23> 제3 발명 및 제6 발명에 따르면, 도달 거리의 차분을 구하기 위해서 위상차분(위상차 스펙트럼)을 산출하는 경우에, 과거의 샘플링 시점에서 산출된 위상차분 에 기초하여, 새롭게 산출된 위상차분을 순차적으로 보정할 수 있다. 보정된 위상차 스펙트럼에는, 과거의 샘플링 시점에서의 신호 대 잡음비가 큰 주파수에서의 위상차분의 정보도 반영되어 있으므로, 배경 잡음의 상태, 목적으로 하는 음원으로부터 발하여지는 음향 신호의 내용의 변화 등에 의해 위상차분이 크게 변동되지 않는다. 따라서, 보다 정밀도가 높고 안정된 도달 거리의 차분에 기초하여 음향 신호의 입사각, 즉 목적으로 하는 음원이 존재하는 방향을 고정밀도로 추정하는 것이 가능하게 된다.

<24> 제4 발명에 따르면, 음성을 발하는 음원, 예를 들면 인간이 존재하는 방향을 고정밀도로 추정하는 것이 가능하게 된다.

발명의 실시를 위한 구체적인 내용

- <25> 이하, 본 발명을 그 실시 형태를 나타내는 도면에 기초하여 상세히 설명한다. 본 실시 형태에서는, 처리 대상의 음향 신호가 주로 인간이 발하는 음성인 경우에 대해 설명한다.
- <26> (실시 형태1)
- <27> 도 1은 본 발명의 실시 형태1에 따른 음원 방향 추정 장치(1)를 구현화하는 범용 컴퓨터의 구성을 도시하는 블록도이다.
- <28> 본 발명의 실시 형태1에 따른 음원 방향 추정 장치(1)로서 동작하는 범용 컴퓨터는, 적어도 CPU, DSP 등의 연산 처리부(11), ROM(12), RAM(13), 외부의 컴퓨터와의 사이에서 데이터 통신 가능한 통신 인터페이스부(14), 음성 입력을 접수하는 복수의 음성 입력부(15, 15, ...), 음성을 출력하는 음성 출력부(16)를 구비하고 있다. 음성 출력부(16)는 통신망(2)을 통해서 데이터 통신 가능한 통신 단말 장치(3, 3, ...)의 음성 입력부(31)로부터 입력된 음성을 출력한다. 또한, 통신 단말 장치(3, 3, ...)의 음성 출력부(32)로부터는 잡음을 억제한 음성이 출력된다.
- <29> 연산 처리부(11)는 내부 버스(17)를 통해서 음원 방향 추정 장치(1)의 전술한 바와 같은 하드웨어 각 부와 접속되어 있다. 연산 처리부(11)는, 전술한 하드웨어 각 부를 제어함과 함께, ROM(12)에 기억되어 있는 처리 프로그램, 예를 들면 주파수축 상의 신호의 진폭 성분을 산출하는 프로그램, 산출된 진폭 성분으로부터 잡음 성분을 추정하는 프로그램, 산출된 진폭 성분 및 추정된 잡음 성분에 기초하여 주파수마다의 신호 대 잡음비(Signal-to-Noise ratio : SN비)를 산출하는 프로그램, SN비가 소정값보다도 큰 주파수를 추출하는 프로그램, 추출된 주파수의 위상차분(이하, 위상차 스펙트럼이라고 함)에 기초하여 도달 거리의 차분을 산출하는 프로그램, 도달 거리의 차분에 기초하여 음원의 방향을 추정하는 프로그램 등에 따라서 다양한 소프트웨어적 기능을 실행한다.
- <30> ROM(12)은, 플래시 메모리 등으로 구성되어 있으며, 범용 컴퓨터를 음원 방향 추정 장치(1)로서 기능시키기 위해서 필요한 전술한 바와 같은 처리 프로그램 및 처리 프로그램이 참조하는 수치 정보를 기억하고 있다. RAM(13)은, SRAM 등으로 구성되어 있으며, 프로그램의 실행 시에 발생하는 일시적인 데이터를 기억한다. 통신 인터페이스부(14)는, 외부의 컴퓨터로부터의 전술한 프로그램의 다운로드, 통신망(2)을 통해서 통신 단말 장치(3, 3, ...)에의 출력 신호의 송신, 및 입력된 음향 신호의 수신 등을 행한다.
- <31> 음성 입력부(15, 15, ...)는, 구체적으로는, 각각 음 입력을 접수하는 마이크로폰이며, 음원의 방향을 특정하기 위해 복수의 마이크로폰, 증폭기, 및 A/D 변환기 등으로 구성되어 있다. 음성 출력부(16)는 스피커 등의 출력 장치이다. 또한, 설명의 편의상, 도 1에는 음성 입력부(15) 및 음성 출력부(16)가 음원 방향 추정 장치(1)에 내장되어 있는 것처럼 도시되어 있다. 그러나, 실제로는 음성 입력부(15) 및 음성 출력부(16)가 인터페이스를 통해서 범용 컴퓨터에 접속됨으로써 음원 방향 추정 장치(1)가 구성되어 있다.
- <32> 도 2는 본 발명의 실시 형태1에 따른 음원 방향 추정 장치(1)의 연산 처리부(11)가 전술한 바와 같은 처리 프로그램을 실행함으로써 실현되는 기능을 도시하는 블록도이다. 또한, 도 2에 도시되어 있는 예에서는, 2개의 음성 입력부(15, 15)가 모두 1개의 마이크로폰인 경우에 대해 설명한다.
- <33> 도 2에 도시하는 바와 같이, 본 발명의 실시 형태1에 따른 음원 방향 추정 장치(1)는, 처리 프로그램이 실행된 경우에 실현되는 기능 블록으로서, 적어도 음성 접수부(음향 신호 접수 수단)(201), 신호 변환부(신호 변환 수단)(202), 위상차 스펙트럼 산출부(위상차분 산출 수단)(203), 진폭 스펙트럼 산출부(진폭 성분 산출 수단)(204), 배경 잡음 추정부(잡음 성분 추정 수단)(205), SN비 산출부(신호 대 잡음비 산출 수단)(206), 위상차 스펙트럼 선택부(주파수 추출 수단)(207), 도달 거리차 산출부(도달 거리 차분 산출 수단)(208), 및 음원 방향 추정부(음원 방향 추정 수단)(209)를 구비하고 있다.
- <34> 음성 접수부(201)는 음원인 인간이 발하는 음성을 2개의 마이크로폰으로부터 음 입력으로서 각각 접수한다. 본 실시 형태에서는, 입력1 및 입력2가 각각 마이크로폰인 음성 입력부(15, 15)를 통해서 접수된다.
- <35> 신호 변환부(202)는, 입력된 음성에 대해서, 시간축 상의 신호를 주파수축 상의 신호, 즉 복소 스펙트럼 $IN1(f)$, $IN2(f)$ 로 변환한다. 여기서 f 는 주파수(radian)를 나타내고 있다. 신호 변환부(202)에서는, 예를 들면 푸리에 변환과 같은 시간-주파수 변환 처리가 실행된다. 본 실시 형태1에서는, 푸리에 변환과 같은 시간-주파수 변환 처리에 의해, 입력된 음성이 스펙트럼 $IN1(f)$, $IN2(f)$ 로 변환된다.
- <36> 위상차 스펙트럼 산출부(203)는, 주파수 변환된 스펙트럼 $IN1(f)$, $IN2(f)$ 에 기초하여 위상 스펙트럼을 산출하고, 산출된 위상 스펙트럼간의 위상차분인 위상차 스펙트럼 $DIFF_PHASE(f)$ 를 주파수마다 산출한다. 또한, 스펙트럼 $IN1(f)$, $IN2(f)$ 각각의 위상 스펙트럼을 구하는 것이 아니라, $IN1(f)/IN2(f)$ 의 위상 성분을 구

함으로써 위상차 스펙트럼 DIFF_PHASE(f)를 구해도 된다. 여기서, 진폭 스펙트럼 산출부(204)는, 어느 한쪽, 예를 들면 도 2에 도시하는 예에서는 입력1의 입력 신호 스펙트럼 IN1(f)의 진폭 성분인 진폭 스펙트럼 |IN1(f)|을 산출한다. 어느 쪽의 진폭 스펙트럼을 산출할지는 특별히 한정되는 것은 아니다. 진폭 스펙트럼 |IN1(f)|과 |IN2(f)|를 산출하여, 큰 쪽의 값을 선택하여도 된다.

- <37> 또한, 실시 형태1에서는, 푸리에 변환된 스펙트럼에서의 주파수마다 진폭 스펙트럼 |IN1(f)|을 산출하는 구성을 채용하고 있다. 그러나, 실시 형태1에서는, 대역 분할을 행하여, 특정한 중심 주파수와 간격으로 분할된 분할 대역 내에서 진폭 스펙트럼 |IN1(f)|의 대표값을 구하는 구성을 채용해도 된다. 그 경우의 대표값은, 분할 대역 내에서의 진폭 스펙트럼 |IN1(f)|의 평균값이어도 되고, 최대값이어도 된다. 또한, 대역 분할된 후의 진폭 스펙트럼의 대표값은 |IN1(n)|로 된다. 여기서, n은 분할된 대역의 인덱스를 나타내고 있다.
- <38> 배경 잡음 추정부(205)는, 진폭 스펙트럼 |IN1(f)|에 기초하여 배경 잡음 스펙트럼 |NOISE1(f)|을 추정한다. 배경 잡음 스펙트럼 |NOISE1(f)|의 추정 방법은 특별히 한정되는 것은 아니다. 음성 인식에서의 음성 구간 검출 처리, 또는 휴대 전화기 등에서 이용되고 있는 노이즈 캔슬러 처리에서 행해지는 배경 잡음 추정 처리 등과 같은 이미 공지인 방법을 이용하는 것이 가능하다. 바꾸어 말하면, 배경 잡음의 스펙트럼을 추정하는 방법이면 어떠한 방법이라도 이용 가능하다. 또한, 전술한 바와 같이, 진폭 스펙트럼이 대역 분할되어 있는 경우에는, 분할 대역마다 배경 잡음 스펙트럼 |NOISE1(n)|을 추정하면 된다. 여기서, n은 분할된 대역의 인덱스를 나타내고 있다.
- <39> SN비 산출부(206)는, 진폭 스펙트럼 산출부(204)에서 산출된 진폭 스펙트럼 |IN1(f)|과, 배경 잡음 추정부(205)에서 추정된 배경 잡음 스펙트럼 |NOISE1(f)|의 비율을 산출함으로써, SN비 SNR(f)을 산출한다. SN비 SNR(f)은 하기 수학적 식 1에 의해 산출된다. 또한, 진폭 스펙트럼이 대역 분할되어 있는 경우에는, 분할 대역마다 SNR(n)을 산출하면 된다. 여기서, n은 분할된 대역의 인덱스를 나타내고 있다.

수학적 식 1

<40>
$$SNR(f) = 20.0 \times \log_{10} (| IN 1 (f) | / | NOISE 1 (f) |) \quad \dots (1)$$

- <41> 위상차 스펙트럼 선택부(207)는, 소정값보다도 큰 SN비가 SN비 산출부(206)에서 산출된 주파수 또는 주파수 대역을 추출하고, 추출된 주파수에 대응하는 위상차 스펙트럼 또는 추출된 주파수 대역 내의 위상차 스펙트럼을 선택한다.
- <42> 도달 거리차 산출부(208)는, 선택된 위상차 스펙트럼과 주파수 f의 관계를 직선 근사한 함수를 구한다. 이 함수에 기초하여 도달 거리차 산출부(208)는, 음원과 양 음성 입력부(15, 15) 각각의 사이의 거리의 차, 즉 음성이 양 음성 입력부(15, 15)에 각각 도달할 때까지의 거리차 D를 산출한다.
- <43> 음원 방향 추정부(209)는, 도달 거리차 산출부(208)가 산출한 거리차 D와, 양 음성 입력부(15, 15)의 설치 간격 L을 이용하여 음성 입력의 입사각 θ , 즉 음원인 인간이 존재한다고 추정되는 방향을 나타내는 각도 θ 를 산출한다.
- <44> 이하, 본 발명의 실시 형태1에 따른 음원 방향 추정 장치(1)의 연산 처리부(11)가 실행하는 처리 수순에 대해서 설명한다. 도 3은 본 발명의 실시 형태1에 따른 음원 방향 추정 장치(1)의 연산 처리부(11)가 실행하는 처리 수순을 설명하는 플로우차트이다.
- <45> 음원 방향 추정 장치(1)의 연산 처리부(11)는 우선, 음성 입력부(15, 15)로부터 음향 신호(아날로그 신호)를 접수한다(스텝 S301). 연산 처리부(11)는, 접수한 음향 신호를 A/D 변환한 후, 얻어진 샘플 신호를 소정의 시간 단위로 프레임화한다(스텝 S302). 이 때, 안정된 스펙트럼을 구하기 위해서, 프레임화된 샘플 신호에 대하여 해밍 창(hamming window), 해닝 창(hanning window) 등의 시간창이 곱해진다. 프레임화의 단위는, 샘플링 주파수, 어플리케이션의 종류 등에 의해 결정된다. 예를 들면, 10ms~20ms씩 오버랩시키면서 20ms~40ms 단위로 프레임화가 행해지고, 프레임마다 이하의 처리가 실행된다.
- <46> 연산 처리부(11)는, 프레임 단위로 시간축 상의 신호를 주파수축 상의 신호, 즉 스펙트럼 IN1(f), IN2(f)로 변환한다(스텝 S303). 여기서 f는 주파수(radian)를 나타내고 있다. 연산 처리부(11)는, 예를 들면 푸리에 변환과 같은 시간-주파수 변환 처리를 실행한다. 본 실시 형태1에서는, 연산 처리부(11)는, 푸리에 변환과 같은 시간-주파수 변환 처리에 의해, 프레임 단위의 시간축 상의 신호를 스펙트럼 IN1(f), IN2(f)로 변환한다.
- <47> 다음으로, 연산 처리부(11)는, 주파수 변환된 스펙트럼 IN1(f), IN2(f)의 실부 및 허부를 이용하여 위상 스펙트

럼을 산출하고, 산출된 위상 스펙트럼간의 위상차분인 위상차 스펙트럼 DIFF_PHASE(f)를 주파수마다 산출한다(스텝 S304).

<48> 한편, 연산 처리부(11)는, 입력1의 입력 신호 스펙트럼 IN1(f)의 진폭 성분인 진폭 스펙트럼 |IN1(f)|을 산출한다(스텝 S305).

<49> 단, 입력1의 입력 신호 스펙트럼 IN1(f)에 대해서 진폭 스펙트럼을 산출하는 것에 한정될 필요는 없다. 그 밖에 예를 들면, 입력2의 입력 신호 스펙트럼 IN2(f)에 대해서 진폭 스펙트럼을 산출해도 되고, 양 입력1, 2의 진폭 스펙트럼의 평균값 또는 최대값 등을 진폭 스펙트럼의 대표값으로서 산출해도 된다. 여기서는 푸리에 변환된 스펙트럼에서의 주파수마다 진폭 스펙트럼 |IN1(f)|을 산출하는 구성을 채용하고 있지만, 대역 분할을 행하여, 특정한 중심 주파수와 간격으로 분할된 분할 대역 내에서 진폭 스펙트럼 |IN1(f)|의 대표값을 산출하는 구성을 채용해도 된다. 또한, 대표값은, 분할 대역 내에서의 진폭 스펙트럼 |IN1(f)|의 평균값이어도 되고, 최대값이어도 된다. 또한, 진폭 스펙트럼을 산출하는 구성에 한정될 필요는 없으며, 예를 들면 파워 스펙트럼을 산출하는 구성이어도 된다. 이 경우의 SN비 SNR(f)은 하기 수학적 식 2에 의해 산출된다.

수학적 식 2

$$SNR(f) = 10.0 \times \log_{10} (| IN 1 (f) | ^ 2 / | NOISE 1 (f) | ^ 2) \dots (2)$$

<50> 연산 처리부(11)는, 산출된 진폭 스펙트럼 |IN1(f)|에 기초하여 잡음 구간을 추정하고, 추정된 잡음 구간의 진폭 스펙트럼 |IN1(f)|에 기초하여 배경 잡음 스펙트럼 |NOISE1(f)|을 추정한다(스텝 S306).

<52> 단, 잡음 구간의 추정 방법은 특별히 한정될 필요는 없다. 배경 잡음 스펙트럼 |NOISE1(f)|을 추정하는 방법에 대해서는, 예를 들면 그 밖에, 음성 인식에서의 음성 구간 검출 처리, 또는 휴대 전화기 등에서 이용되고 있는 노이즈 캔슬러 처리에서 행해지는 배경 잡음 추정 처리 등과 같은 이미 공지인 방법을 이용하는 것이 가능하다. 바꾸어 말하면, 배경 잡음의 스펙트럼을 추정하는 방법이면 어떠한 방법이라도 이용 가능하다. 예를 들면, 전체 주파수 대역에서의 파워 정보를 이용하여 배경 잡음 레벨을 추정하고, 추정된 배경 잡음 레벨에 기초하여 음성/잡음을 판정하기 위한 임계값을 구함으로써 음성/잡음 판정을 행하는 것이 가능하다. 이 결과, 잡음으로 판정된 경우에는, 그 때의 진폭 스펙트럼 |IN1(f)|을 이용하여 배경 잡음 스펙트럼 |NOISE1(f)|을 보정함으로써, 배경 잡음 스펙트럼 |NOISE1(f)|을 추정하는 것이 일반적이다.

<53> 연산 처리부(11)는, 수학적 식 1(파워 스펙트럼의 경우에는 수학적 식 2)에 따라서 주파수 또는 주파수 대역마다의 SN비 SNR(f)을 산출한다(스텝 S307). 연산 처리부(11)는, 산출된 SN비가 소정값보다도 큰 주파수 또는 주파수 대역을 선택한다(스텝 S308). 소정값의 결정 방법에 따라서, 선택되는 주파수 또는 주파수 대역을 변동시킬 수 있다. 예를 들면, 인접하는 주파수 또는 주파수 대역간에서 SN비의 비교를 행하여, SN비가 보다 큰 주파수 또는 주파수 대역을 순차적으로 RAM(13)에 기억시키면서 선택해 감으로써, SN비가 최대인 주파수 또는 주파수 대역을 선택할 수 있다. 또한, SN비가 큰 순으로 상위 N(N은 자연수)개를 선택하여도 된다.

<54> 연산 처리부(11)는, 1 또는 복수의 선택된 주파수 또는 주파수 대역에 대응하는 위상차 스펙트럼 DIFF_PHASE(f)에 기초하여, 위상차 스펙트럼 DIFF_PHASE(f)와 주파수 f의 관계를 직선 근사한다(스텝 S309). 이 결과, SN비가 큰 주파수 또는 주파수 대역에서의 위상차 스펙트럼 DIFF_PHASE(f)의 신뢰도가 높은 것을 이용할 수 있다. 이에 의해, 위상차 스펙트럼 DIFF_PHASE(f)와 주파수 f의 비례 관계의 추정 정밀도를 높일 수 있다.

<55> 도 4의 (a), (b) 및 (c)는 SN비가 소정값보다도 큰 주파수 또는 주파수 대역을 선택한 경우의 위상차 스펙트럼의 보정 방법을 도시하는 모식도이다.

<56> 도 4의 (a)는 주파수 또는 주파수 대역에 대응하는 위상차 스펙트럼 DIFF_PHASE(f)를 도시하고 있다. 통상은 배경 잡음이 중첩되어 있으므로, 일정한 관계를 발견하는 것은 곤란한 상태로 되어 있다.

<57> 도 4의 (b)는 주파수 또는 주파수 대역 내에 있는 SN비 SNR(f)을 도시하고 있다. 구체적으로는, 도 4의 (b)에서 이중 동그라미로 표시하는 부분이, SN비가 소정값보다도 큰 주파수 또는 주파수 대역을 나타내고 있다. 따라서, 도 4의 (b)에 도시하는 바와 같은 SN비가 소정값보다도 큰 주파수 또는 주파수 대역을 선택함으로써, 선택된 주파수 또는 주파수 대역에 대응하는 위상차 스펙트럼 DIFF_PHASE(f)는 도 4의 (a)에서 이중 동그라미로 표시하는 부분으로 된다. 도 4의 (a)에 도시하는 바와 같이 선택된 위상차 스펙트럼 DIFF_PHASE(f)를 직선 근사함으로써, 위상차 스펙트럼 DIFF_PHASE(f)와 주파수 f 사이에는, 도 4의 (c)에 도시하는 바와 같은 비례 관계가 존재하는 것을 알 수 있다.

<58> 따라서, 연산 처리부(11)는, 나이키스트 주파수 F와, 나이키스트 주파수 F에서의 직선 근사된 위상차 스펙트럼 DIFF_PHASE(π)의 값, 즉 도 4의 (c)에서의 R과, 음속 c를 이용하여, 하기 수학적 식 3에 따라서 음원으로부터의 음 입력의 도달 거리의 차분 D를 산출한다(스텝 S310). 또한, 나이키스트 주파수는 샘플링 주파수의 절반의 값이며, 도 4의 (a), (b) 및 (c)에서는 π 이다. 구체적으로는, 샘플링 주파수가 8kHz인 경우에는 나이키스트 주파수는 4kHz로 된다.

<59> 또한, 도 4의 (c)에는, 선택된 위상차 스펙트럼 DIFF_PHASE(f)를 원점을 통과하는 직선으로 근사한 근사 직선이 도시되어 있다. 그러나, 음성 입력부(15, 15, ...)로서의 마이크로폰 각각의 특성이 상위하는 경우에는 위상차 스펙트럼에 전체 대역에 걸쳐 바이어스가 걸릴 가능성이 있다. 그와 같은 경우에는, 근사 직선의 주파수 0에 대응하는 값, 즉 근사 직선의 절편의 값을 고려하여 나이키스트 주파수에서의 위상차의 값 R을 보정함으로써 근사 직선을 구하는 것도 가능하다.

수학적 식 3

$$D = (R \times c) / (F \times 2 \pi) \quad \dots (3)$$

<60>

<61> 연산 처리부(11)는, 산출된 도달 거리의 차분 D를 이용하여, 음 입력의 입사각 θ , 즉 음원이 존재한다고 추정되는 방향을 나타내는 각도 θ 를 산출한다(스텝 S311). 도 5는 음원이 존재한다고 추정되는 방향을 나타내는 각도 θ 를 산출하는 방법의 원리를 도시하는 모식도이다.

<62> 도 5에 도시하는 바와 같이, 2개의 음성 입력부(15, 15)는 간격 L만큼 이격하여 설치되어 있다. 이 경우, 음원으로부터의 음 입력의 도달 거리의 차분 D와, 2개의 음성 입력부(15, 15) 사이의 간격 L과의 사이에는, 「 $\sin \theta = (D/L)$ 」의 관계가 있다. 따라서, 음원이 존재한다고 추정되는 방향을 나타내는 각도 θ 는 하기 수학적 식 4에 의해 구할 수 있다.

수학적 식 4

$$\theta = \sin^{-1} (D/L) \quad \dots (4)$$

<63>

<64> 또한, SN비가 큰 순으로 N개의 주파수 또는 주파수 대역이 선택된 경우에도, 전술한 바와 같이, 상위 N개의 위상차 스펙트럼을 이용하여 직선 근사한다. 이 밖에, 나이키스트 주파수 F에서의 직선 근사된 위상차 스펙트럼 DIFF_PHASE(F)의 값 R은 이용하지 않고, 선택된 주파수 f에서의 위상차 스펙트럼 r(=DIFF_PHASE(f))의 값을 이용하여, 수학적 식 3의 F 및 R을 각각 f 및 r로 치환하여, 선택된 주파수마다 도달 거리의 차분 D를 산출하고, 산출된 차분 D의 평균값을 이용하여 음원이 존재한다고 추정되는 방향을 나타내는 각도 θ 를 산출하는 것도 가능하다. 물론, 이와 같은 방법에 한정될 필요는 없다. 예를 들면, SN비에 따른 가중치 부여를 행하여 도달 거리의 차분 D의 대표값을 산출함으로써, 음원이 존재한다고 추정되는 방향을 나타내는 각도 θ 를 산출하여도 된다.

<65> 또한, 음성을 발하는 인간이 존재하는 방향을 추정하는 경우에는, 음 입력이 인간이 발한 음성을 나타내는 음성 구간인지의 여부를 판단하고, 음성 구간이라고 판단된 경우에만 전술한 처리를 실행함으로써, 음원이 존재한다고 추정되는 방향을 나타내는 각도 θ 를 산출하여도 된다.

<66> 또한, SN비가 소정값보다도 크다고 판단된 경우라도, 어플리케이션의 사용 상태, 사용 조건 등을 감안하여, 상정되어 있지 않은 위상차인 경우에는, 대응하는 주파수 또는 주파수 대역을 선택 대상으로부터 제외하는 것이 바람직하다. 예를 들면 휴대 전화기와 같이 정면 방향으로부터 발화하는 것이 상정되어 있는 기기에 본 실시 형태1에 따른 음원 방향 추정 장치(1)를 적용하는 경우, 정면을 0도로 하여 음원이 존재한다고 추정되는 방향 θ 가, $\theta < -90$ 도 또는 $90 < \theta$ 인 것으로 산출된 경우에는 상정 외인 것으로 판단된다.

<67> 또한, SN비가 소정값보다도 크다고 판단된 경우라도, 어플리케이션의 사용 상태, 사용 조건 등을 감안하여, 목적으로 하는 음원의 방향을 추정하기 위해서는 바람직하지 않은 주파수 또는 주파수 대역을 선택 대상으로부터 제외하는 것이 바람직하다. 예를 들면 목적으로 하는 음원이 인간이 발하는 음성인 경우에는, 100Hz 이하의 주파수에는 음성 신호가 존재하지 않는다. 따라서, 100Hz 이하의 선택 대상으로부터 제외할 수 있다.

<68> 이상과 같이, 본 실시 형태1에 따른 음원 방향 추정 장치(1)는, 입력된 음향 신호의 진폭 성분, 소위 진폭 스펙트럼과, 추정된 배경 잡음 스펙트럼에 기초하여 주파수 또는 주파수 대역마다의 SN비를 구하고, SN비가 큰 주파수에서의 위상차분(위상차 스펙트럼)을 이용함으로써, 보다 정확한 도달 거리의 차분 D를 구할 수 있다. 따라서, 정밀도가 높은 도달 거리의 차분 D에 기초하여 음향 신호의 입사각, 즉 목적으로 하는 음원(본 실시 형태1

에서는 인간이 존재한다고 추정되는 방향을 나타내는 각도 θ 를 고정밀도로 산출하는 것이 가능하게 된다.

- <69> (실시 형태2)
- <70> 이하, 본 발명의 실시 형태2에 따른 음원 방향 추정 장치(1)를, 도면을 참조하면서 상세하게 설명한다. 본 발명의 실시 형태2에 따른 음원 방향 추정 장치(1)로서 동작하는 범용 컴퓨터의 구성은, 실시 형태1과 마찬가지로의 구성이므로, 도 1에 도시하는 블록도를 참조하는 것으로 하고 상세한 설명을 생략한다. 본 실시 형태2는, 프레임 단위로의 위상차 스펙트럼의 산출 결과를 기억해 놓고, 기억되어 있는 전회의 위상차 스펙트럼 및 산출 대상 프레임에서의 SN비에 기초하여, 산출 대상 프레임에서의 위상차 스펙트럼을 수시 보정하는 구성을 채용하고 있는 점에서 실시 형태1과 상이하다.
- <71> 도 6은 본 발명의 실시 형태2에 따른 음원 방향 추정 장치(1)의 연산 처리부(11)가 처리 프로그램을 실행함으로써 실현되는 기능을 도시하는 블록도이다. 또한, 도 6에 도시되어 있는 예에서는, 실시 형태1과 마찬가지로, 음성 입력부(15, 15)를 2개의 마이크로폰으로 구성한 경우에 대해 설명한다.
- <72> 도 6에 도시하는 바와 같이, 본 발명의 실시 형태2에 따른 음원 방향 추정 장치(1)는, 처리 프로그램이 실행된 경우에 실현되는 기능 블록으로서, 적어도 음성접수부(음향 신호 접수부)(201), 신호 변환부(신호 변환 수단)(202), 위상차 스펙트럼 산출부(위상차분 산출 수단)(203), 진폭 스펙트럼 산출부(진폭 성분 산출 수단)(204), 배경 잡음 추정부(잡음 성분 추정 수단)(205), SN비 산출부(신호 대 잡음비 산출 수단)(206), 위상차 스펙트럼 보정부(보정 수단)(210), 도달 거리차 산출부(도달 거리 차분 산출 수단)(208), 및 음원 방향 추정부(음원 방향 추정 수단)(209)를 구비하고 있다.
- <73> 음성 접수부(201)는 음원인 인간이 발하는 음성 입력을 2개의 마이크로폰으로부터 접수한다. 본 실시 형태에서는, 입력1 및 입력2가 각각 마이크로폰인 음성 입력부(15, 15)를 통해서 접수된다.
- <74> 신호 변환부(202)는, 입력된 음성에 대해서, 시간축 상의 신호를 주파수축 상의 신호, 즉 복소 스펙트럼 $IN1(f)$, $IN2(f)$ 로 변환한다. 여기서 f 는 주파수(radian)를 나타내고 있다. 신호 변환부(202)에서는, 예를 들면 푸리에 변환과 같은 시간-주파수 변환 처리가 실행된다. 본 실시 형태2에서는, 푸리에 변환과 같은 시간-주파수 변환 처리에 의해, 입력된 음성이 스펙트럼 $IN1(f)$, $IN2(f)$ 로 변환된다.
- <75> 또한, 음성 입력부(15, 15)에서 접수한 입력 신호는, A/D 변환된 후, 얻어진 샘플 신호가 소정의 시간 단위로 프레임화된다. 이 때, 안정된 스펙트럼을 구하기 위해서, 프레임화된 샘플 신호에 대하여 해밍 창(hamming window),海宁 창(hanning window) 등의 시간창이 곱해진다. 프레임화의 단위는, 샘플링 주파수, 어플리케이션의 종류 등에 의해 결정된다. 예를 들면, 10ms~20ms씩 오버랩시키면서 20ms~40ms 단위로 프레임화가 행해지고, 프레임마다 이하의 처리가 실행된다.
- <76> 위상차 스펙트럼 산출부(203)는, 주파수 변환된 스펙트럼 $IN1(f)$, $IN2(f)$ 에 기초하여 프레임 단위로 위상 스펙트럼을 산출하고, 산출된 위상 스펙트럼간의 위상차분인 위상차 스펙트럼 $DIFF_PHASE(f)$ 를 프레임 단위로 산출한다. 여기서, 진폭 스펙트럼 산출부(204)는, 어느 한쪽, 예를 들면 도 6에 도시하는 예에서는 입력1의 입력 신호 스펙트럼 $IN1(f)$ 의 진폭 성분인 진폭 스펙트럼 $|IN1(f)|$ 을 산출한다. 어느 쪽의 진폭 스펙트럼을 산출할지는 특별히 한정되는 것은 아니다. 진폭 스펙트럼 $|IN1(f)|$ 과 $|IN2(f)|$ 를 산출하여, 양자의 평균값을 선택해도 되고, 큰 쪽의 값을 선택해도 된다.
- <77> 배경 잡음 추정부(205)는, 진폭 스펙트럼 $|IN1(f)|$ 에 기초하여 배경 잡음 스펙트럼 $|NOISE1(f)|$ 을 추정한다. 배경 잡음 스펙트럼 $|NOISE1(f)|$ 의 추정 방법은 특별히 한정되는 것은 아니다. 음성 인식에서의 음성 구간 검출 처리, 또는 휴대 전화기 등에서 이용되고 있는 노이즈 캔슬러 처리에서 행해지는 배경 잡음 추정 처리 등과 같은 이미 공지인 방법을 이용하는 것이 가능하다. 바꾸어 말하면, 배경 잡음의 스펙트럼을 추정하는 방법이면 어떠한 방법이라도 이용 가능하다.
- <78> SN비 산출부(206)는, 진폭 스펙트럼 산출부(204)에서 산출된 진폭 스펙트럼 $|IN1(f)|$ 과, 배경 잡음 추정부(205)에서 추정된 배경 잡음 스펙트럼 $|NOISE1(f)|$ 의 비율을 산출함으로써, SN비 $SNR(f)$ 을 산출한다.
- <79> 위상차 스펙트럼 보정부(210)는, SN비 산출부(206)에서 산출된 SN비와 위상차 스펙트럼 보정부(210)에서 보정된 후에 RAM(13)에 기억되어 있는 전회의 샘플링 시점에서 산출된 위상차 스펙트럼 $DIFF_PHASE_{t-1}(f)$ 에 기초하여, 다음 샘플링 시점, 즉 현재의 샘플링 시점에서 산출된 위상차 스펙트럼 $DIFF_PHASE_t(f)$ 를 보정한다. 현재의 샘플링 시점에서는, SN비 및 위상차 스펙트럼 $DIFF_PHASE_t(f)$ 가 전회까지와 마찬가지로 하여 산출된 후, SN비에 따라서 설정되어 있는 보정 계수 α ($0 \leq \alpha \leq 1$)를 사용하여 하기 수학적 식 5에 따라서, 현재의 샘플링 시점에서의

프레임의 위상차 스펙트럼 DIFF_PHASEt(f)가 산출된다.

<80> 또한, 보정 계수 α에 대해서는 상세는 후술하지만, 예를 들면, SN비에 따른 값이 처리 프로그램이 참조하는 수치 정보로서 각 프로그램과 함께 ROM(12)에 기억되어 있다.

수학식 5

$$DIFF_PHASEt(f) = \alpha \times DIFF_PHASEt(f) + (1 - \alpha) \times DIFF_PHASEt-1(f) \dots (5)$$

<81>

<82> 도달 거리차 산출부(208)는, 보정된 위상차 스펙트럼과 주파수 f의 관계를 직선 근사한 함수를 구한다. 이 함수에 기초하여, 도달 거리차 산출부(208)는, 음원과 양 음성 입력부(15, 15) 각각과의 사이의 거리의 차, 즉 음성이 양 음성 입력부(15, 15)에 각각 도달할 때까지의 거리차 D를 산출한다.

<83> 음원 방향 추정부(209)는, 거리차 D와, 양 음성 입력부(15, 15)의 설치 간격 L을 이용하여, 음 입력의 입사각 θ, 즉 음원인 인간이 존재한다고 추정되는 방향을 나타내는 각도 θ를 산출한다.

<84> 이하, 본 발명의 실시 형태2에 따른 음원 방향 추정 장치(1)의 연산 처리부(11)가 실행하는 처리 수순에 대해서 설명한다. 도 7 및 도 8은, 본 발명의 실시 형태2에 따른 음원 방향 추정 장치(1)의 연산 처리부(11)가 실행하는 처리 수순을 설명하는 플로우차트이다.

<85> 음원 방향 추정 장치(1)의 연산 처리부(11)는 우선, 음성 입력부(15, 15)로부터 음향 신호(아날로그 신호)를 접수한다(스텝 S701). 연산 처리부(11)는, 접수한 음향 신호를 A/D 변환한 후, 얻어진 샘플 신호를 소정의 시간 단위로 프레임화한다(스텝 S702). 이 때, 안정된 스펙트럼을 구하기 위해서, 프레임화된 샘플 신호에 대하여 해밍 창(hamming window), 해닝 창(hanning window) 등의 시간창이 곱해진다. 프레임화의 단위는, 샘플링 주파수, 어플리케이션의 종류 등에 의해 결정된다. 예를 들면, 10ms~20ms씩 오버랩시키면서 20ms~40ms 단위로 프레임화가 행해지고, 프레임마다 이하의 처리가 실행된다.

<86> 연산 처리부(11)는, 프레임 단위로 시간축 상의 신호를 주파수축 상의 신호, 즉 스펙트럼 IN1(f), IN2(f)로 변환한다(스텝 S703). 여기서 f는 주파수(radian) 또는 샘플링 시의 일정한 폭을 갖는 주파수 대역을 나타내고 있다. 연산 처리부(11)는, 예를 들면 푸리에 변환과 같은 시간-주파수 변환 처리를 실행한다. 본 실시 형태2에서는, 연산 처리부(11)는, 푸리에 변환과 같은 시간-주파수 변환 처리에 의해, 프레임 단위의 시간축 상의 신호를 스펙트럼 IN1(f), IN2(f)로 변환한다.

<87> 다음으로, 연산 처리부(11)는, 주파수 변환된 스펙트럼 IN1(f), IN2(f)의 실부 및 허부를 이용하여 위상 스펙트럼을 산출하고, 산출된 위상 스펙트럼간의 위상차분인 위상차 스펙트럼 DIFF_PHASEt(f)를 주파수 또는 주파수 대역마다 산출한다(스텝 S704).

<88> 한편, 연산 처리부(11)는, 입력1의 입력 신호 스펙트럼 IN1(f)의 진폭 성분인 진폭 스펙트럼 |IN1(f)|을 산출한다(스텝 S705).

<89> 단, 입력1의 입력 신호 스펙트럼 IN1(f)에 대해서 진폭 스펙트럼을 산출하는 것에 한정될 필요는 없다. 그 밖에 예를 들면, 입력2의 입력 신호 스펙트럼 IN2(f)에 대해서 진폭 스펙트럼을 산출해도 되고, 양 입력1, 2의 진폭 스펙트럼의 평균값 또는 최대값 등을 진폭 스펙트럼의 대표값으로서 산출해도 된다. 또한, 진폭 스펙트럼을 산출하는 구성에 한정될 필요는 없으며, 예를 들면 파워 스펙트럼을 산출하는 구성이어도 된다.

<90> 연산 처리부(11)는, 산출된 진폭 스펙트럼 |IN1(f)|에 기초하여 잡음 구간을 추정하고, 추정된 잡음 구간의 진폭 스펙트럼 |IN1(f)|에 기초하여 배경 잡음 스펙트럼 |NOISE1(f)|을 추정한다(스텝 S706).

<91> 단, 잡음 구간의 추정 방법은 특별히 한정될 필요는 없다. 배경 잡음 스펙트럼 |NOISE1(f)|을 추정하는 방법에 대해서는, 예를 들면 그 밖에, 전체 주파수 대역에서의 파워 정보를 이용하여 배경 잡음 레벨을 추정하고, 추정된 배경 잡음 레벨에 기초하여 음성/잡음을 판정하기 위한 임계값을 구함으로써 음성/잡음 판정을 행하는 것이 가능하다. 이 결과, 잡음으로 판정된 경우에는, 그 때의 진폭 스펙트럼 |IN1(f)|을 이용하여 배경 잡음 스펙트럼 |NOISE1(f)|을 보정함으로써, 배경 잡음 스펙트럼 |NOISE1(f)|을 추정하는 것 등과 같은, 배경 잡음 스펙트럼을 추정하는 방법이면 어떠한 방법을 이용하여도 된다.

<92> 연산 처리부(11)는, 전술한 수학식 1에 따라서 주파수 또는 주파수 대역마다의 SN비 SNR(f)을 산출한다(스텝 S707). 다음으로, 연산 처리부(11)는, RAM(13)에 전회의 샘플링 시점에서의 위상차 스펙트럼 DIFF_PHASEt-

1(f)이 기억되어 있는지의 여부를 판단한다(스텝 S708).

- <93> 연산 처리부(11)는, 전회의 샘플링 시점에서의 위상차 스펙트럼 DIFF_PHASEt-1(f)이 기억되어 있다고 판단한 경우(스텝 S708 : 예), 산출된 샘플링 시점(현재의 샘플링 시점)에서의 SN비에 따른 보정 계수 α 를 ROM(12)으로부터 읽어낸다(스텝 S710). 또한, SN비와 보정 계수 α 의 관계를 나타내는 함수를 프로그램에 짜 넣어 놓고, 계산에 의해 보정 계수 α 를 구해도 된다.
- <94> 도 9는 SN비에 따른 보정 계수 α 의 일례를 도시하는 그래프이다. 도 9에 도시하는 예에서는, SN비가 0(제로)인 경우에 보정 계수 α 가 0(제로)으로 설정되어 있다. 이것은, 산출된 SN비가 0(제로)인 경우에는, 전술한 수학적 식 5로부터 이해되는 바와 같이, 산출된 위상차 스펙트럼 DIFF_PHASEt(f)는 이용하지 않고, 전회의 위상차 스펙트럼 DIFF_PHASEt-1(f)을 현재의 위상차 스펙트럼으로서 이용함으로써 후속의 처리가 행해지는 것을 의미하고 있다. 이하, SN비가 커짐에 따라서 보정 계수 α 는 단조 증가하도록 설정되어 있다. SN비가 20dB 이상인 영역에서는, 보정 계수 α 는 1보다도 작은 최대값 α_{max} 로 고정되어 있다. 여기서, 보정 계수 α 의 최대값 α_{max} 를 1보다도 작은 값으로 설정하고 있는 이유는, SN비가 높은 잡음이 돌발적으로 발생한 경우에, 위상차 스펙트럼 DIFF_PHASEt(f)의 값이 그 잡음의 위상차 스펙트럼으로 100% 치환되는 것을 방지하기 위해서이다.
- <95> 연산 처리부(11)는, SN비에 따라서 ROM(12)으로부터 읽어내어진 보정 계수 α 를 이용하여, 전술한 수학적 식 5에 따라서 위상차 스펙트럼 DIFF_PHASEt(f)를 보정한다(스텝 S711). 이 후, 연산 처리부(11)는, RAM(13)에 기억되어 있는 전회의 샘플링 시점에서의 보정 후의 위상차 스펙트럼 DIFF_PHASEt-1(f)을, 현재의 샘플링 시점에서의 보정 후의 위상차 스펙트럼 DIFF_PHASEt(f)로 갱신하여 기억한다(스텝 S712).
- <96> 연산 처리부(11)는, 전회의 샘플링 시점에서의 위상차 스펙트럼 DIFF_PHASEt-1(f)이 기억되어 있지 않다고 판단한 경우(스텝 S708 : 아니오), 현재의 샘플링 시점에서의 위상차 스펙트럼 DIFF_PHASEt(f)를 이용할지의 여부를 판단한다(스텝 S717). 현재의 샘플링 시점에서의 위상차 스펙트럼 DIFF_PHASEt(f)를 이용할지의 여부의 판단 기준으로서, 전체 주파수 대역의 SN비, 음성/잡음 판정의 결과 등과 같은, 목적으로 하는 음원으로부터 음향 신호가 발하여 지고 있는지(인간이 발생하고 있는지)의 여부를 판단 기준이 이용된다.
- <97> 한편, 연산 처리부(11)는, 현재의 샘플링 시점에서의 위상차 스펙트럼 DIFF_PHASEt(f)를 이용하지 않는, 즉 음원으로부터 음향 신호가 발하여 지고 있을 가능성이 낮다고 판단한 경우(스텝 S717 : 아니오), 미리 정해져 있는 위상차 스펙트럼의 초기값을 현재의 샘플링 시점에서의 위상차 스펙트럼으로 한다(스텝 S718). 이 경우, 위상차 스펙트럼의 초기값은 예를 들면 전체 주파수에 걸쳐 0(제로)으로 설정된다. 그러나, 이 스텝 S718에서의 설정은 이 값(즉, 제로)으로 한정될 필요는 없다.
- <98> 다음으로, 연산 처리부(11)는, 위상차 스펙트럼의 초기값을 현재의 샘플링 시점에서의 위상차 스펙트럼으로서 RAM(13)에 기억하고(스텝 S719), 처리를 스텝 S713으로 진행시킨다.
- <99> 연산 처리부(11)는, 현재의 샘플링 시점에서의 위상차 스펙트럼 DIFF_PHASEt(f)를 이용하는, 즉 음원으로부터 음향 신호가 발하여 지고 있을 가능성이 높다고 판단한 경우(스텝 S717 : 예), 현재의 샘플링 시점에서의 위상차 스펙트럼 DIFF_PHASEt(f)를 RAM(13)에 기억하고(스텝 S720), 처리를 스텝 S713으로 진행시킨다.
- <100> 다음으로 연산 처리부(11)는, 스텝 S712, S719, S720 중 어느 하나에서 기억된 위상차 스펙트럼 DIFF_PHASE(f)에 기초하여, 위상차 스펙트럼 DIFF_PHASE(f)와 주파수 f의 관계를 직선 근사한다(스텝 S713). 이 결과, 보정 후의 위상차 스펙트럼에 기초하여 직선 근사한 경우에는, 현재의 샘플링 시점뿐만 아니라, 과거의 샘플링 시점에서 SN비가 컸었던(즉, 신뢰도가 높았던) 주파수 또는 주파수 대역에서의 위상차분의 정보를 반영하고 있는 위상차 스펙트럼 DIFF_PHASE(f)를 이용할 수 있다. 이에 의해, 위상차 스펙트럼 DIFF_PHASE(f)와 주파수 f의 비례 관계의 추정 정밀도를 높일 수 있다.
- <101> 연산 처리부(11)는, 나이키스트 주파수 F에서의 직선 근사된 위상차 스펙트럼 DIFF_PHASE(F)의 값 R을 이용하여, 전술한 수학적 식 3에 따라서, 음원으로부터의 음향 신호의 도달 거리의 차분 D를 산출한다(스텝 S714). 단, 나이키스트 주파수 F에서의 직선 근사된 위상차 스펙트럼 DIFF_PHASE(F)의 값 R을 이용하지 않고, 임의의 주파수 f에서의 위상차 스펙트럼 r(=DIFF_PHASE(f))의 값을 이용하였다고 해도, 수학적 식 3의 F 및 R을 f 및 r로 각각 치환함으로써, 도달 거리의 차분 D를 구할 수 있다. 그리고 연산 처리부(11)는, 산출된 도달 거리의 차분 D를 이용하여, 음향 신호의 입사각 θ , 즉 음원(인간)이 존재한다고 추정되는 방향을 나타내는 각도 θ 를 산출한다(스텝 S715).
- <102> 또한, 음성을 발하는 인간이 존재하는 방향을 추정하는 경우에는, 음 입력이 인간이 발한 음성을 나타내는 음성 구간인지의 여부를 판단하고, 음성 구간이라고 판단된 경우에만 전술한 처리를 실행함으로써, 음원이 존재한다

고 추정되는 방향을 나타내는 각도 θ 를 산출하여도 된다.

- <103> 또한, SN비가 소정값보다도 크다고 판단된 경우라도, 어플리케이션의 사용 상태, 사용 조건 등을 감안하여, 상정되어 있지 않은 위상차인 경우에는, 대응하는 주파수 또는 주파수 대역을 현재의 샘플링 시점에서의 위상차 스펙트럼의 보정 대상으로부터 제외하는 것이 바람직하다. 예를 들면 휴대 전화기와 같이 정면 방향으로부터 발화하는 것이 상정되어 있는 기기에 본 실시 형태2에 따른 음원 방향 추정 장치(1)를 적용하는 경우, 정면을 0도로 하여 음원이 존재한다고 추정되는 방향 θ 가, $\theta < -90$ 도 또는 $90 < \theta$ 인 것으로 산출된 경우에는 상정 외인 것으로 판단된다. 이 경우, 현재의 샘플링 시점에서의 위상차 스펙트럼을 이용하지 않고 전회까지 산출된 위상차 스펙트럼이 이용된다.
- <104> 또한, SN비가 소정값보다도 크다고 판단된 경우라도, 어플리케이션의 사용 상태, 사용 조건 등을 감안하여, 목적으로 하는 음원의 방향을 추정하기 위해서는 바람직하지 않은 주파수 또는 주파수 대역을 선택 대상으로부터 제외하는 것이 바람직하다. 예를 들면 목적으로 하는 음원이 인간이 발하는 음성인 경우에는, 100Hz 이하의 주파수에는 음성 신호가 존재하지 않는다. 따라서, 100Hz 이하의 보정 대상으로부터 제외할 수 있다.
- <105> 이상과 같이 본 실시 형태2에 따른 음원 방향 추정 장치(1)는, SN비가 큰 주파수 또는 주파수 대역에서의 위상차 스펙트럼을 산출하는 경우에, 전회의 샘플링 시점에서 산출된 위상차 스펙트럼보다도 샘플링 시점(현재의 샘플링 시점)에서의 위상차 스펙트럼쪽에 무게를 두고 보정하고, SN비가 작은 경우에는 전회의 위상차 스펙트럼쪽에 무게를 두고 보정한다. 이와 같이 함으로써, 새롭게 산출된 위상차 스펙트럼을 순차적으로 보정할 수 있다. 보정된 위상차 스펙트럼에는, 과거의 샘플링 시점에서의 SN비가 큰 주파수에서의 위상차분의 정보도 반영되어 있다. 따라서, 배경 잡음의 상태, 목적으로 하는 음원으로부터 발하여지는 음향 신호의 내용의 변화 등에 영향 받아 위상차 스펙트럼이 크게 변동되지 않는다. 따라서, 보다 정밀도가 높은 안정된 도달 거리의 차분 D에 기초하여 음향 신호의 입사각, 즉 목적으로 하는 음원이 존재한다고 추정되는 방향을 나타내는 각도 θ 를 고정밀도로 산출하는 것이 가능하게 된다. 또한, 목적으로 하는 음원이 존재한다고 추정되는 방향을 나타내는 각도 θ 의 산출 방법은 전술한 도달 거리의 차분 D를 이용한 방법에 한정되는 것이 아니라, 마찬가지로 추정 가능한 방법이면 다양한 변형이 존재하는 것은 물론이다.

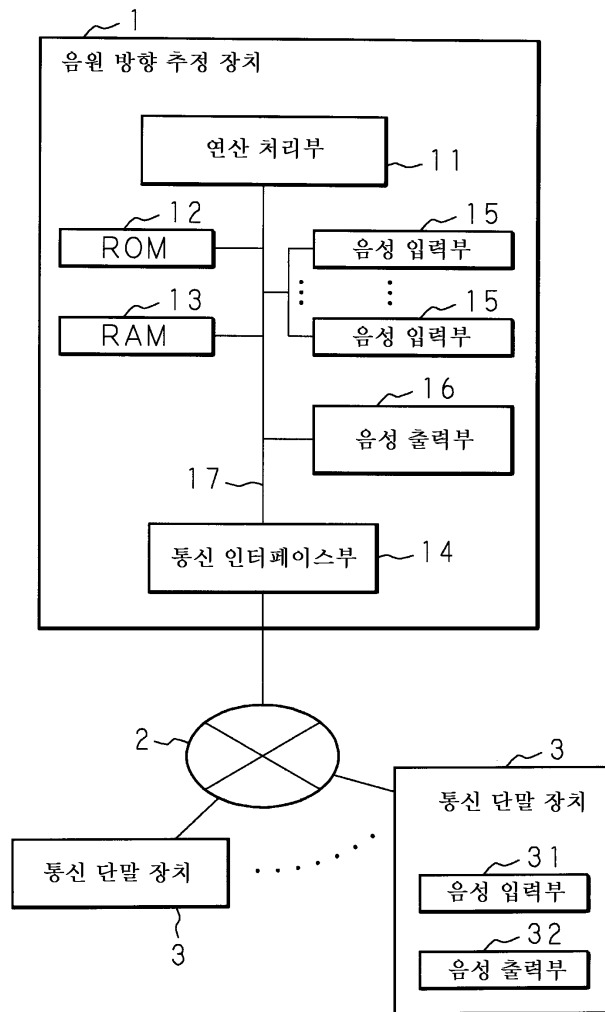
도면의 간단한 설명

- <106> 도 1은 본 발명의 실시 형태1에 따른 음원 방향 추정 장치를 구현화하는 범용 컴퓨터의 구성을 도시하는 블록도.
- <107> 도 2는 본 발명의 실시 형태1에 따른 음원 방향 추정 장치의 연산 처리부가 처리 프로그램을 실행함으로써 실현되는 기능을 도시하는 블록도.
- <108> 도 3은 본 발명의 실시 형태1에 따른 음원 방향 추정 장치의 연산 처리부의 처리 수순을 설명하는 플로우차트.
- <109> 도 4의 (a), (b) 및 (c)는 SN비가 소정값보다도 큰 주파수 또는 주파수 대역을 선택한 경우의, 위상차 스펙트럼의 보정 방법을 도시하는 모식도.
- <110> 도 5는 음원이 존재한다고 추정되는 방향을 나타내는 각도를 산출하는 방법의 원리를 도시하는 모식도.
- <111> 도 6은 본 발명의 실시 형태2에 따른 음원 방향 추정 장치의 연산 처리부가 처리 프로그램을 실행함으로써 실현되는 기능을 도시하는 블록도.
- <112> 도 7은 본 발명의 실시 형태2에 따른 음원 방향 추정 장치의 연산 처리부의 처리 수순을 설명하는 플로우차트.
- <113> 도 8a 및 도 8b는 본 발명의 실시 형태2에 따른 음원 방향 추정 장치의 연산 처리부의 처리 수순을 설명하는 플로우차트.
- <114> 도 9는 SN비에 따른 보정 계수의 일례를 도시하는 그래프.
- <115> <도면의 주요 부분에 대한 부호의 설명>
- <116> 1 : 음원 방향 추정 장치
- <117> 11 : 연산 처리부
- <118> 12 : ROM
- <119> 13 : RAM

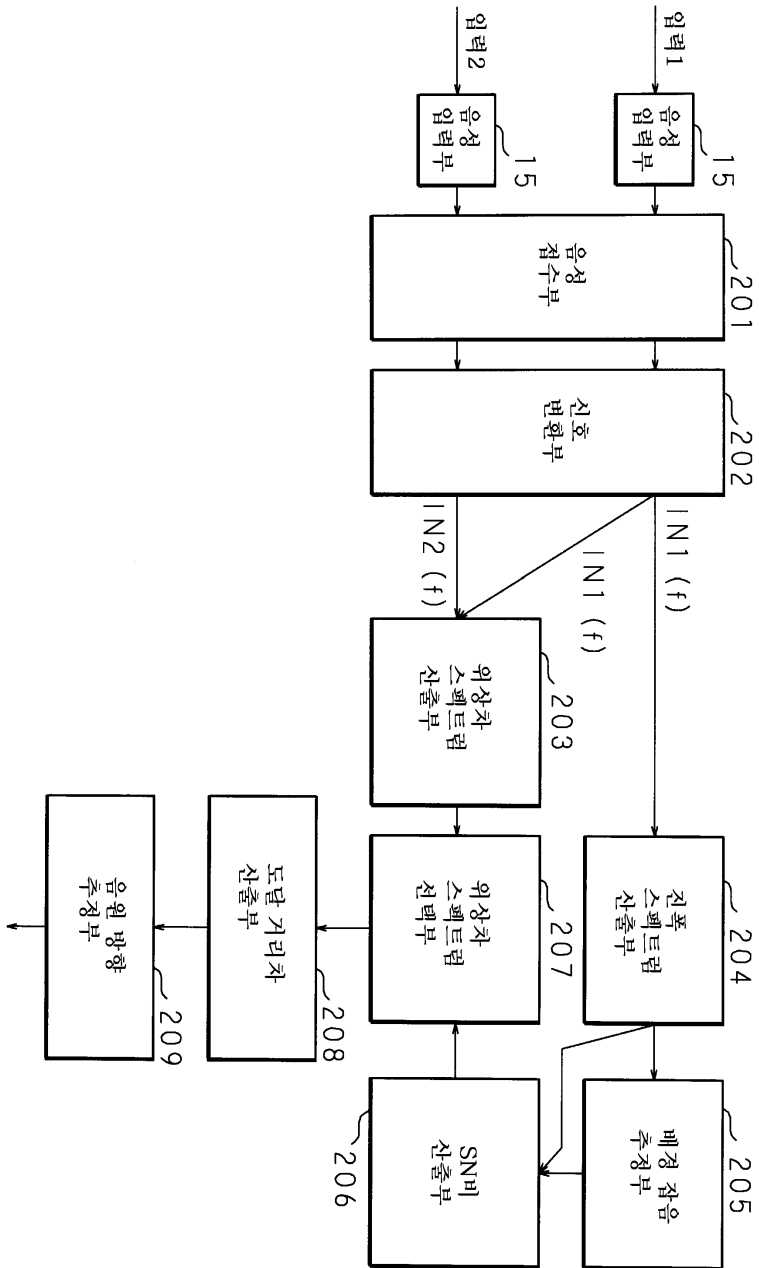
- <120> 14 : 통신 인터페이스부
- <121> 15 : 음성 입력부
- <122> 16 : 음성 출력부
- <123> 17 : 내부 버스
- <124> 201 : 음성 접수부
- <125> 202 : 신호 변환부
- <126> 203 : 위상차 스펙트럼 산출부
- <127> 204 : 진폭 스펙트럼 산출부
- <128> 205 : 배경 잡음 추정부
- <129> 206 : SN비 산출부
- <130> 207 : 위상차 스펙트럼 선택부
- <131> 208 : 도달 거리차 산출부
- <132> 209 : 음원 방향 추정부
- <133> 210 : 위상차 스펙트럼 보정부

도면

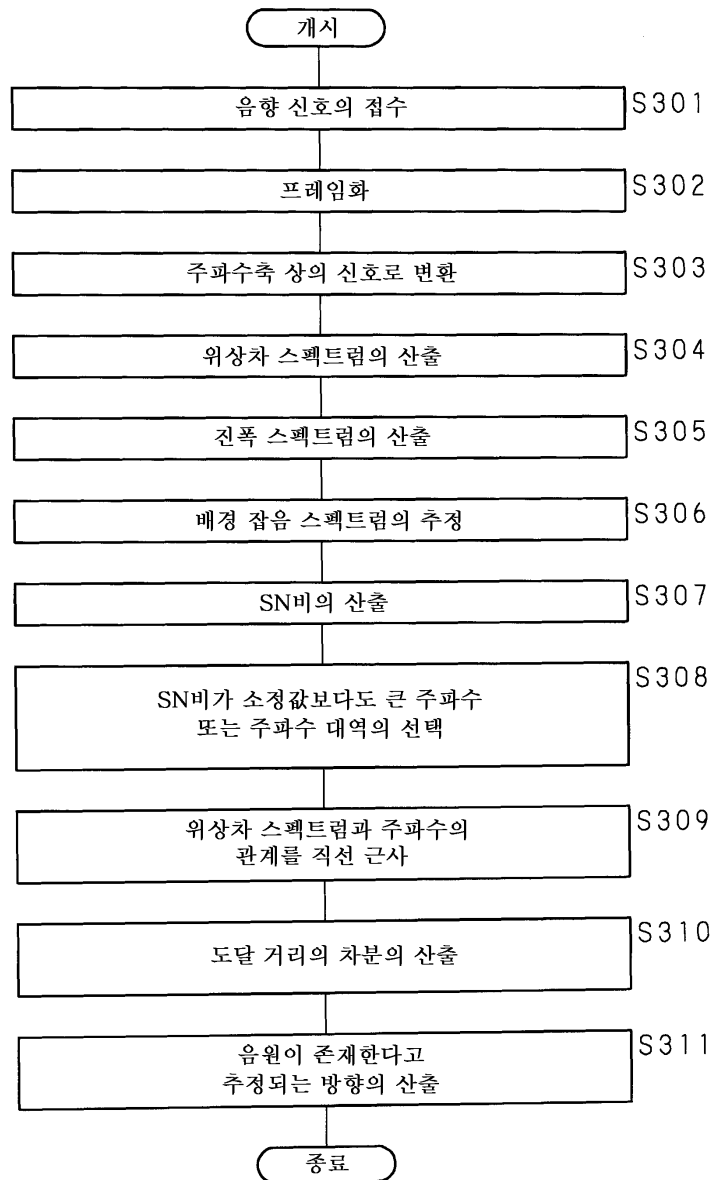
도면1



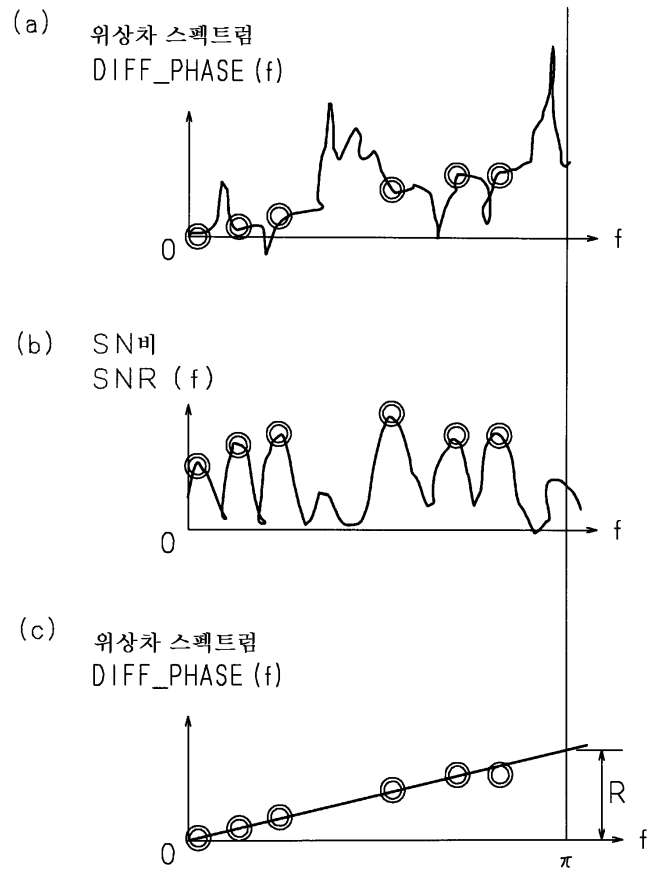
도면2



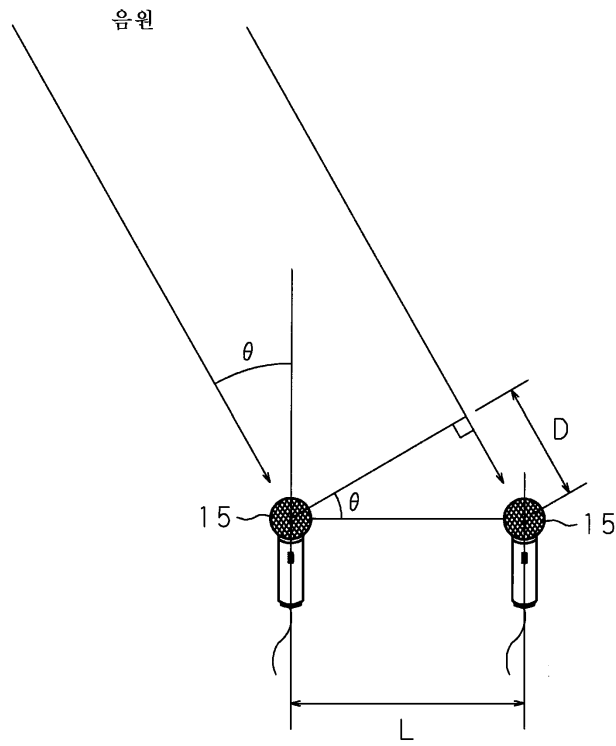
도면3



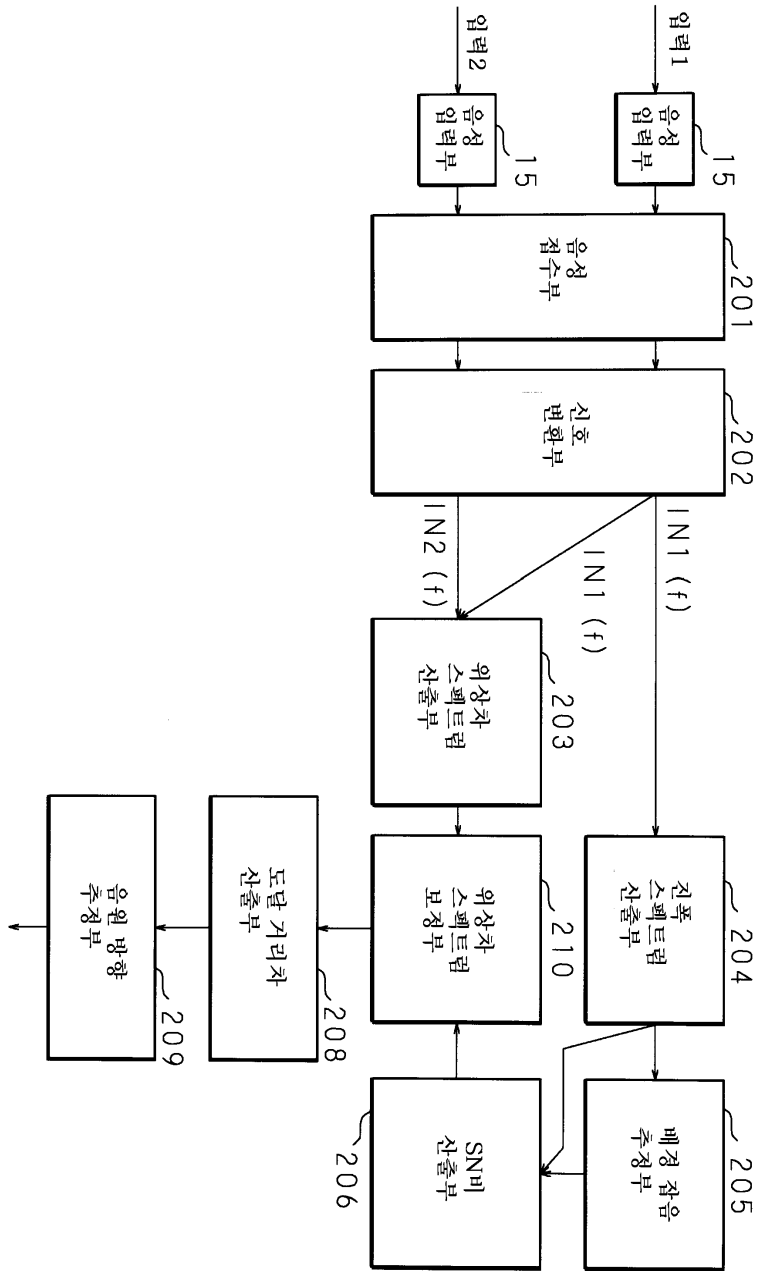
도면4



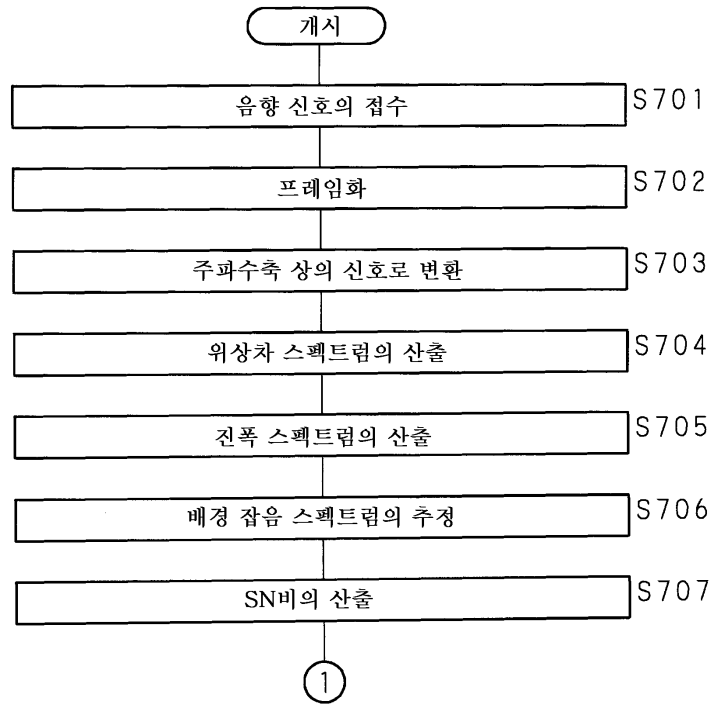
도면5



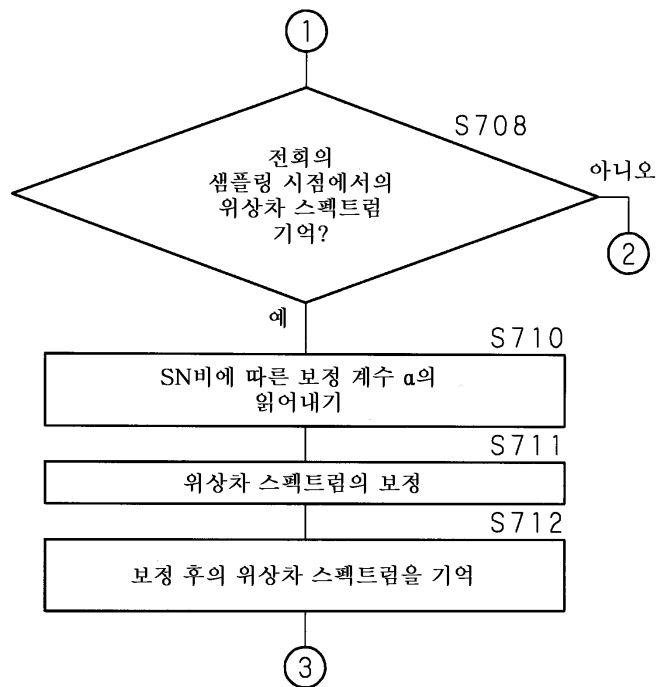
도면6



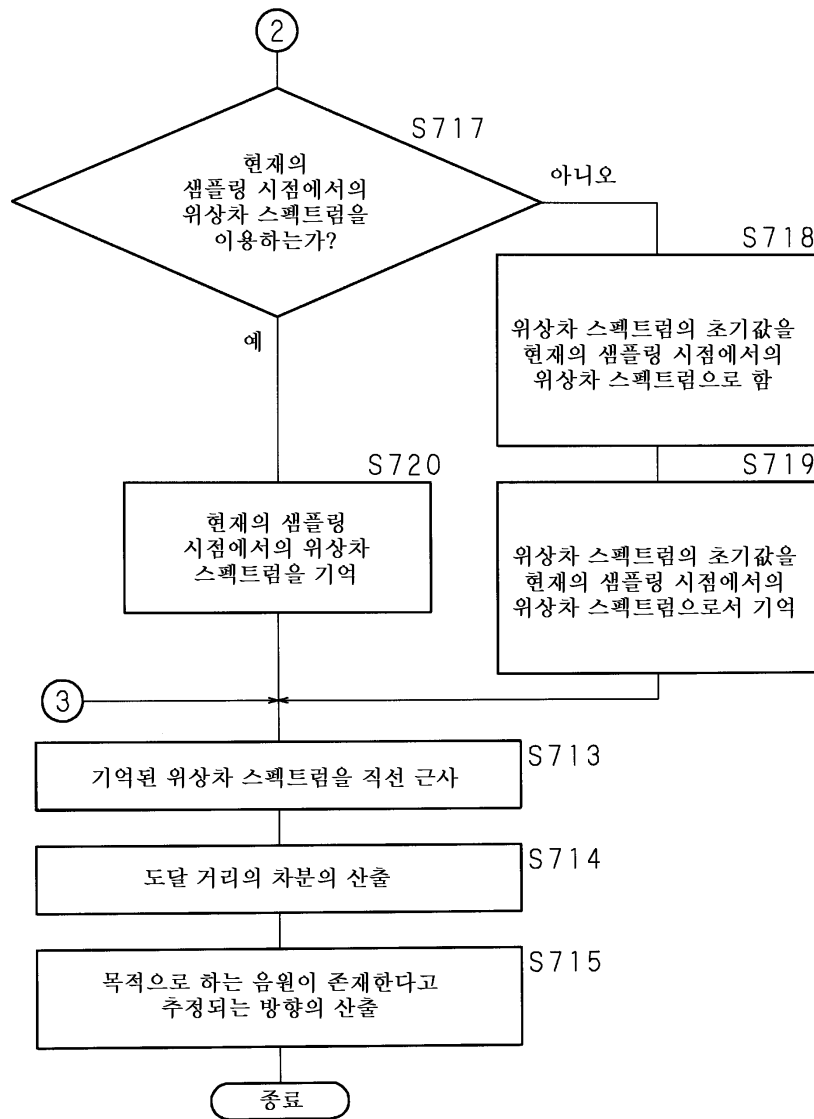
도면7



도면8a



도면8b



도면9

