



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2016년02월05일

(11) 등록번호 10-1592425

(24) 등록일자 2016년02월01일

(51) 국제특허분류(Int. Cl.)

G10L 21/0208 (2013.01) G10L 15/28 (2006.01)

(21) 출원번호 10-2014-0127928

(22) 출원일자 2014년09월24일

심사청구일자 2014년09월24일

(56) 선행기술조사문헌

Alexander Krueger et al., 'Model-based feature enhancement for reverberant speech recognition', IEEE Trans. on Audio, Speech, and Language Processing, Vol.18, No.7, pp.1692~1707, September 2010.*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

현대자동차주식회사

서울특별시 서초구 현릉로 12 (양재동)

(72) 발명자

김비호

서울특별시 서초구 강남대로8길 40-35, 101호 (양재동, 효성하이츠)

(74) 대리인

유미특허법인

전체 청구항 수 : 총 22 항

심사관 : 정성운

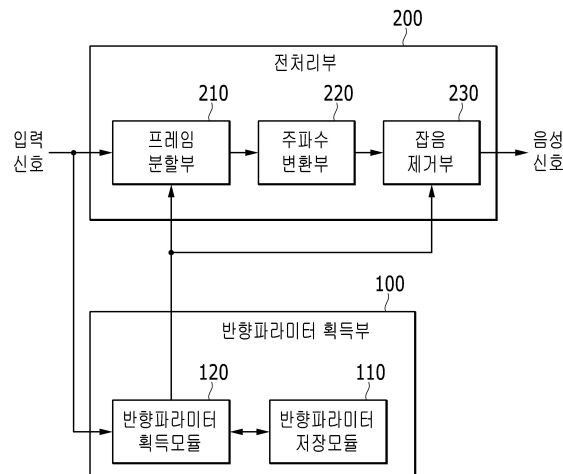
(54) 발명의 명칭 음성 전처리 장치, 음성 인식 장치 및 그 방법

(57) 요약

음성 전처리 장치, 음성 인식 장치 및 그 방법에 관한 것이다.

음성 전처리 장치는, 제1반향파라미터를 획득하는 반향파라미터 획득부, 그리고 입력신호에 잡음 제거 알고리즘을 적용함으로써, 상기 입력신호로부터 잡음성분을 제거하여 출력하며, 상기 제1반향파라미터가 기 설정된 제2반향파라미터와 다른 경우, 상기 제1반향파라미터를 토대로 상기 잡음 제거 알고리즘의 파라미터를 갱신하는 전처리부를 포함할 수 있다.

대표도 - 도2



명세서

청구범위

청구항 1

좌석 센서를 통해 동승자의 동승위치를 감지하고, 기 설정된 복수의 후보 반향파라미터 중 상기 동승자의 동승 위치에 대응하는 후보 반향파라미터를 이용하여 제1반향파라미터를 획득하는 반향파라미터 획득부, 그리고

상기 제1반향파라미터가 기 설정된 제2반향파라미터와 다른 경우 상기 제1반향파라미터를 토대로 잡음 제거 알고리즘의 파라미터를 갱신하고, 상기 잡음 제거 알고리즘을 이용하여 입력신호로부터 잡음성분을 제거하여 출력하는 전처리부를 포함하는 음성 전처리 장치.

청구항 2

삭제

청구항 3

제1항에 있어서,

상기 반향파라미터 획득부는,

상기 동승자의 동승위치에 대응하는 후보 반향파라미터를 상기 제1반향파라미터로 획득하는 음성 전처리 장치.

청구항 4

제1항에 있어서,

상기 반향파라미터 획득부는,

상기 복수의 후보 반향파라미터 중 상기 동승자의 동승위치에 대응하는 후보 반향파라미터를 제3반향파라미터로 획득하고, 상기 입력신호의 음의 방향 분산을 토대로 제4반향파라미터를 획득하며, 상기 제3반향파라미터와 상기 제4반향파라미터의 평균값을 상기 제1반향파라미터로 획득하는 음성 전처리 장치.

청구항 5

제1항에 있어서,

상기 전처리부는,

윈도우 함수를 이용하여 상기 입력신호를 프레임 단위로 분할하는 프레임 분할부,

프레임 단위로 분할된 상기 입력신호를 주파수 영역의 신호로 변환하는 주파수 변환부, 그리고

주파수 영역으로 변환된 상기 입력신호로부터 잡음성분을 제거하는 잡음 제거부를 포함하는 음성 전처리 장치.

청구항 6

제5항에 있어서,

상기 파라미터는 상기 윈도우 함수의 윈도우 사이즈를 포함하는 음성 전처리 장치.

청구항 7

제5항에 있어서,

상기 파라미터는 이득보정계수를 포함하며,

상기 잡음 제거부는, 상기 이득보정계수를 토대로 상기 입력신호로부터 잡음성분을 제거하는 정도를 조절하는 음성 전처리 장치.

청구항 8

제5항에 있어서,

상기 파라미터는 상기 잡음 제거 알고리즘의 반복횟수를 포함하는 음성 전처리 장치.

청구항 9

좌측 센서를 통해 동승자의 동승위치를 감지하고, 기 설정된 복수의 후보 반향파라미터 중 상기 동승자의 동승 위치에 대응하는 후보 반향파라미터를 이용하여 제1반향파라미터를 획득하는 반향파라미터 획득부,

상기 제1반향파라미터가 기 설정된 제2반향파라미터와 다른 경우 상기 제1반향파라미터를 토대로 잡음 제거 알고리즘의 파라미터를 갱신하고, 상기 잡음 제거 알고리즘을 이용하여 입력신호로부터 잡음성분을 제거하여 출력하는 전처리부,

상기 전처리부에서 출력되는 신호로부터 음성신호의 특징패턴을 검출하는 특징패턴 검출부, 그리고

상기 특징패턴을 토대로 음성인식결과를 출력하는 음성 인식부를 포함하는 음성 인식 장치.

청구항 10

제9항에 있어서,

상기 전처리부는,

윈도우 함수를 이용하여 상기 입력신호를 프레임 단위로 분할하는 프레임 분할부,

프레임 단위로 분할된 상기 입력신호를 주파수 영역의 신호로 변환하는 주파수 변환부, 그리고

주파수 영역으로 변환된 상기 입력신호로부터 잡음성분을 제거하는 잡음 제거부를 포함하는 음성 인식 장치.

청구항 11

제10항에 있어서,

상기 파라미터는 상기 윈도우 함수의 윈도우 사이즈를 포함하는 음성 인식 장치.

청구항 12

제10항에 있어서,

상기 파라미터는 이득보정계수를 포함하며,

상기 잡음 제거부는, 상기 이득보정계수를 토대로 상기 입력신호로부터 잡음성분을 제거하는 정도를 조절하는 음성 인식 장치.

청구항 13

제10항에 있어서,

상기 파라미터는 상기 잡음 제거 알고리즘의 반복횟수를 포함하는 음성 인식 장치.

청구항 14

삭제

청구항 15

제9항에 있어서,

상기 반향파라미터 획득부는,

상기 복수의 후보 반향파라미터 중 상기 동승자의 동승위치에 대응하는 후보 반향파라미터를 상기 제1반향파라미터로 획득하는 음성 인식 장치.

청구항 16

제9항에 있어서,

상기 반향파라미터 획득부는,

상기 복수의 후보 반향파라미터 중 상기 동승자의 동승위치에 대응하는 후보 반향파라미터를 제3반향파라미터로 획득하고, 상기 입력신호의 음의 방향 분산을 토대로 제4반향파라미터를 획득하며, 상기 제3반향파라미터와 상기 제4반향파라미터의 평균값을 상기 제1반향파라미터로 획득하는 음성 인식 장치.

청구항 17

좌석 센서를 통해 동승자의 동승위치를 감지하는 단계,

서로 다른 동승위치에 대응하는 복수의 후보 반향파라미터 중 상기 동승자의 동승위치에 대응하는 후보 반향파라미터를 이용하여 제1반향파라미터를 획득하는 단계,

상기 제1반향파라미터가 기 설정된 제2반향파라미터와 다른 경우, 상기 제1반향파라미터를 토대로 잡음 제거 알고리즘의 파라미터를 갱신하는 단계,

상기 잡음 제거 알고리즘을 이용하여 입력신호로부터 잡음성분을 제거하는 전처리 단계,

상기 전처리 단계를 통해 잡음성분이 제거된 상기 입력신호로부터 음성신호의 특징패턴을 검출하는 단계, 그리고

상기 특징패턴을 토대로 음성인식결과를 출력하는 단계

를 포함하는 음성 인식 장치의 음성 인식 방법.

청구항 18

제17항에 있어서,

상기 전처리 단계는,

윈도우 함수를 이용하여 상기 입력신호를 프레임 단위로 분할하는 단계,

프레임 단위로 분할된 상기 입력신호를 주파수 영역의 신호로 변환하는 단계, 그리고

주파수 영역으로 변환된 상기 입력신호로부터 잡음성분을 제거하는 단계를 포함하는 음성 인식 장치의 음성 인식 방법.

청구항 19

제18항에 있어서,

상기 갱신하는 단계는,

상기 제1반향파라미터에 따라서 상기 윈도우 함수의 윈도우 사이즈를 가변하는 단계를 포함하는 음성 인식 장치의 음성 인식 방법.

청구항 20

제18항에 있어서,

상기 갱신하는 단계는,

상기 제1반향파라미터에 따라서 이득보정계수를 갱신하는 단계를 포함하며,

상기 잡음성분을 제거하는 단계는,

상기 이득보정계수를 토대로 상기 입력신호로부터 잡음성분을 제거하는 정도를 조절하는 단계를 포함하는 음성 인식 장치의 음성 인식 방법.

청구항 21

제18항에 있어서,

상기 갱신하는 단계는,

상기 제1반향파라미터에 따라서 상기 잡음 제거 알고리즘의 반복횟수를 갱신하는 단계를 포함하는 음성 인식 장치의 음성 인식 방법.

청구항 22

삭제

청구항 23

제18항에 있어서,

상기 획득하는 단계는,

상기 동승자의 동승위치에 대응하는 후보 반향파라미터를 상기 제1반향파라미터로 획득하는 단계를 포함하는 음성 인식 장치의 음성 인식 방법.

청구항 24

제18항에 있어서,

상기 획득하는 단계는,

상기 동승자의 동승위치에 대응하는 후보 반향파라미터를 제3반향파라미터로 획득하는 단계,

상기 입력신호의 음의 방향 분산을 토대로 제4반향파라미터를 획득하는 단계, 그리고

상기 제3 및 제4반향파라미터의 평균값을 상기 제1반향파라미터로 획득하는 단계를 포함하는 음성 인식 장치의 음성 인식 방법.

청구항 25

삭제

청구항 26

제17항 내지 제21항, 제23항 및 제24항 중 어느 한 항의 방법을 실행시키기 위해 기록매체에 저장된 프로그램.

발명의 설명

기술 분야

[0001] 본 발명은 음성 전처리 장치, 음성 인식 장치 및 그 방법에 관한 것이다.

배경 기술

[0002] 인간-장치간 인터페이스(Man Machine Interface, MMI)는 컴퓨터 등의 기계장치와 그것을 이용하는 사용자 간의 인터페이스로, 사용자의 시각, 청각, 촉각을 사용하는 모든 인터페이스를 포함한다.

[0003] 최근 주행 중 운전자의 주의 분산을 최소화하고 편의성을 증대시키기 위해, 차량 내 MMI 수단으로 음성을 사용하기 위한 연구가 활발하다.

[0004] 그러나, 차량 특성 상 잡음 환경에서 음성 인식이 이루어져 음성 인식률이 떨어지는 단점으로 인해, 차량 내 극히 일부에서만 음성 인식이 사용되고 있다.

[0005] 따라서, 차량 내에서 음성 인식을 사용하는 부분을 확대하기 위해서는 차량 내 소음 환경에서 음성 인식률을 높이기 위한 방안이 필요하다.

[0006] 일반적으로 음성 인식을 수행하기 위해서는 다음의 처리과정이 필요하다.

[0007] 우선, 아날로그 신호인 음성신호가 입력되면, 디지털 신호처리를 위해 이를 적당한 주파수로 샘플링한다. 그리고, 샘플링된 데이터를 분석 단위인 프레임으로 나눈 후, 각 프레임에 대한 주파수 분석을 통해 원하는 특징패턴을 추출하고, 추출된 특징패턴과 기준 음성의 특징 패턴과의 일치성을 비교하여 음성 인식 결과를 출력한다.

[0008] 음성 인식률을 높이기 위한 방안 중 하나로, 음성 인식을 수행하기 전에 음성신호에 더해진 잡음을 제거하는 전처리 방안이 제안되었다.

[0009] 음성 인식 시 비교 대상이 되는 기준음성은 그 데이터베이스 구축 시 잡음이 없는 환경에서 녹음된 것이다. 따라서, 실제 입력되는 음성신호로부터 잡음을 제거하지 않은 상태로 특징패턴을 추출할 경우, 기준 음성에서 추출한 특징패턴과 차이가 존재할 수 밖에 없으며, 이러한 차이로 인해 음성 인식률이 떨어지는 문제가 있다.

[0010] 한편, 기존의 음성신호 전처리 과정에서는, 잡음 제거에 필요한 파라미터들을 차량의 출고 시 상태에 맞춰 고정된 상태로 사용하였다. 이와 같이, 고정된 파라미터들을 사용하여 전처리를 수행할 경우, 구조 변화, 동승자 탑승 등과 같이 차량 환경이 크게 변화하는 경우, 잡음 제거 성능이 떨어질 뿐만 아니라 심한 경우 음성 신호에 대한 왜곡을 가져와 음성 인식 성능을 크게 떨어뜨리는 문제가 발생할 수 있다.

발명의 내용

해결하려는 과제

[0011] 본 발명의 실시 예를 통해 해결하려는 과제는, 차량 내 환경 변화를 능동적으로 반영하여 효과적으로 잡음을 제거하기 위한 음성 전처리 방법과, 이를 토대로 음성 인식을 수행하는 음성 전처리 장치, 음성 인식 장치 및 그 방법을 제공하는 것이다.

과제의 해결 수단

- [0012] 상기 과제를 해결하기 위한 본 발명의 일 실시 예에 따른 음성 전처리 장치는, 제1반향파라미터를 획득하는 반향파라미터 획득부, 그리고 입력신호에 잡음 제거 알고리즘을 적용함으로써, 상기 입력신호로부터 잡음성분을 제거하여 출력하며, 상기 제1반향파라미터가 기 설정된 제2반향파라미터와 다른 경우, 상기 제1반향파라미터를 토대로 상기 잡음 제거 알고리즘의 파라미터를 갱신하는 전처리부를 포함할 수 있다.
- [0013] 또한, 본 발명의 일 실시 예에 따른 음성 인식 장치는, 제1반향파라미터를 획득하는 반향파라미터 획득부, 입력신호에 잡음 제거 알고리즘을 적용함으로써, 상기 입력신호로부터 잡음성분을 제거하여 출력하며, 상기 제1반향파라미터가 기 설정된 제2반향파라미터와 다른 경우, 상기 제1반향파라미터를 토대로 상기 잡음 제거 알고리즘의 파라미터를 갱신하는 전처리부, 상기 전처리부에서 출력되는 신호로부터 음성신호의 특징패턴을 검출하는 특징패턴 검출부, 그리고 상기 특징패턴을 토대로 음성인식결과를 출력하는 음성 인식부를 포함할 수 있다.
- [0014] 또한, 본 발명의 일 실시 예에 따른 음성 인식 장치의 음성 인식방법은, 제1반향파라미터를 획득하는 단계, 상기 제1반향파라미터가 기 설정된 제2반향파라미터와 다른 경우, 상기 제1반향파라미터에 따라서 잡음 제거 알고리즘의 파라미터를 갱신하는 단계, 그리고 입력신호에 상기 잡음 제거 알고리즘을 적용하여, 상기 입력신호로부터 잡음성분을 제거하는 전처리 단계를 포함할 수 있다.

발명의 효과

- [0015] 본 발명의 실시 예들에 따르면, 동승자 탑승 등 차량 내 변화하는 환경에 따라서 적응적으로 반향파라미터를 설정함으로써, 전처리 과정을 통한 잡음제거 성능을 향상시키는 효과가 있다.

도면의 간단한 설명

- [0016] 도 1은 본 발명의 일 실시 예에 따른 음성 인식 장치를 개략적으로 도시한 구조도이다.
- 도 2는 본 발명의 일 실시 예에 따른 음성 전처리부를 개략적으로 도시한 구조도다.
- 도 3은 본 발명의 일 실시 예에 따른 음성 인식 장치에서의 음성 인식 방법을 도시한 흐름도이다.
- 도 4는 본 발명의 일 실시 예에 따른 음성 인식 장치에서의 전처리 방법을 도시한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0017] 이하, 첨부한 도면을 참고로 하여 본 발명의 실시 예들에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시 예들에 한정되지 않는다.
- [0018] 본 발명의 실시 예를 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 동일 또는 유사한 구성요소에 대해서는 동일한 참조 부호를 붙이도록 한다.
- [0019] 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라, 그 중간에 다른 소자를 사이에 두고 "전기적으로 연결"되어 있는 경우도 포함한다. 또한 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.
- [0020] 이하, 필요한 도면들을 참조하여 본 발명의 일 실시 예에 따른 음성 전처리 장치, 음성 인식 장치 및 그 방법에 대해 설명하기로 한다.
- [0021] 도 1은 본 발명의 일 실시 예에 따른 음성 인식 장치를 개략적으로 도시한 구조도이다. 또한, 도 2는 본 발명의 일 실시 예에 따른 음성 전처리 장치를 개략적으로 도시한 구조도다.
- [0022] 도 1을 참조하면, 본 발명의 일 실시 예에 따른 음성 인식 장치는 반향파라미터(Reverberation parameter) 획득

부(100), 전처리부(200), 특징패턴 검출부(300), 음성 인식부(400) 등을 포함할 수 있다.

- [0023] 반향파라미터 획득부(100)는 차량 내 환경에 따라서 적응적으로 반향파라미터를 획득하고, 이를 전처리부(200)로 출력할 수 있다.
- [0024] 본 발명의 실시 예에서, 반향파라미터는 반향시간(Reverberation Time)을 나타내는 파라미터로서, 음원으로부터 출력된 음향신호의 에너지가 처음(출력 시점)의 100만분의 1, 음압으로는 1000분의 1, 즉, 60dB로 감쇠하기까지의 시간을 나타낸다. 아래에서는, 반향파라미터를 'RT60'로 표기하기도 한다.
- [0025] 도 2를 참조하면, 반향파라미터 획득부(100)는, 반향파라미터 저장모듈(110), 반향파라미터 획득모듈(120) 등을 포함할 수 있다.
- [0026] 반향파라미터 저장모듈(110)은 차량 출고 시 디폴트로 설정된 디폴트 반향파라미터를 저장할 수 있다.
- [0027] 또한, 반향파라미터 저장모듈(110)은 동승자의 동승 위치 별로 각각 대응하는 후보 반향파라미터를 저장할 수 있다. 후보 반향파라미터는, 차량 출고 시 미리 설정된 반향파라미터로서, 각 좌석에서의 동승자 유무에 따라서 미리 학습된 반향파라미터이다.
- [0028] 반향파라미터 획득모듈(120)은 지속적으로 반향파라미터를 획득한다.
- [0029] 반향파라미터 획득모듈(120)은 시간 영역(time domain)에서 입력신호를 샘플링하고, 샘플링 값의 음의 방향 분산(negative side variance)을 토대로 반향파라미터를 추정할 수 있다.
- [0030] 아래에서는, 수학적식 1 및 2를 참조하여 입력신호의 음의 방향 분산을 이용한 반향파라미터 산출 방법에 대하여 상세하게 설명하기로 한다.
- [0031] 음성신호는 스파스(sparse)한 라플라시안(laplacian) 분포를 따르는 경향이 있다. 또한, 음성신호의 왜도(skewness)는 음성신호의 음의 방향 및 양의 방향 분산(positive side variance)을 따르는 경향이 있다.
- [0032] 한편, 음성신호에서 양의 방향 분산은, 주로 음성 감쇠율 분포(speech decay rate distribution)를 따르며, 음성신호의 왜도는 음의 방향 분산에 비해 좀 더 화자 의존성이 크게 나타난다.
- [0033] 따라서, 본 발명의 일 실시 예에서는, 입력신호의 음의 방향 분산을 이용하여 반향파라미터를 추정한다.
- [0034] 반향파라미터를 추정하기 위해, 반향파라미터 획득모듈(120)은 우선 입력신호의 음의 방향 분산 ($f_x^-(\lambda)$)을 아래의 수학적식1과 같이 획득할 수 있다.

수학적식 1

$$f_x^-(\lambda) = \begin{cases} f_x(\lambda) & \text{for } \lambda \leq 0 \\ f_x(-\lambda) & \text{if } \lambda > 0. \end{cases}$$

- [0035]
- [0036] 위 수학적식 1에서, λ 는 음성 감쇠율(decay rate)을 나타내는 파라미터이다.
- [0037] 위 수학적식 1을 참조하면, 음의 방향 분산($f_x^-(\lambda)$)은, 입력신호의 오리지널 분포(original distribution)($f_x(\lambda)$)에서의 음의 방향 분포와 동일한 음의 방향 분포를 가지면서, 대칭 분포를 보이는 분산으로 획득될 수 있다.
- [0038] 전술한 바와 같이, 음의 방향 분산이 획득되면, 반향파라미터 획득모듈(120)은 아래의 수학적식 2와 같이, 반향파라미터(RT60)를 획득할 수 있다.

수학식 2

$$RT60 = 3 \ln \frac{-20}{\lambda_n} \text{ where } \lambda_n = \gamma_2 (\sigma_{x^-}^2)^2 + \gamma_1 \sigma_{x^-}^2 + \gamma_0$$

[0039]

[0040]

위 수학식2에서, $\gamma_2, \gamma_1, \gamma_0$ 는 복수의 반향파라미터(RT60)에 해당하는 믹싱 시뮬레이션 모듈을 2차 다항식으로 근사한 파라미터이고, $\sigma_{x^-}^2$ 는 위 수학식 1을 통해 획득한 음의 방향 분산을 나타낸다.

[0041]

반향파라미터 획득모듈(120)은 동승자의 동승 위치에 따라서 반향파라미터를 획득할 수도 있다.

[0042]

이 경우, 반향파라미터 획득모듈(120)은 차량 내 각 좌석에 설치된 복수의 좌석 센서를 포함하며, 좌석 센서를 통해 동승자의 위치를 검출할 수 있다. 그리고, 좌석 센서를 통해 검출한 동승자의 동승위치에 대응하는 후보 반향파라미터를 반향파라미터 저장모듈(110)로부터 읽어와, 이를 반향파라미터로 획득할 수 있다.

[0043]

좌석 센서는 감압 센서를 포함할 수 있다. 이 경우, 반향파라미터 획득모듈(120)은 감압 센서를 통해 차량 내 좌석에 가해지는 압력 변화를 감지하고, 이를 토대로 동승자의 동승 여부를 감지할 수 있다.

[0044]

반향파라미터 획득모듈(120)은 입력신호의 음의 방향 분산을 이용하여 추정한 반향파라미터와, 동승자의 동승 위치에 따라서 획득한 반향파라미터를 모두 사용하여 반향파라미터를 획득할 수도 있다. 이 경우, 반향파라미터 획득모듈(120)은 입력신호의 음의 방향 분산을 이용하여 추정한 반향파라미터와, 동승자의 동승 위치에 따라서 획득한 반향파라미터의 평균 값을 최종 반향파라미터로 획득할 수 있다.

[0045]

전처리부(200)는 잡음 제거 알고리즘을 적용하여 입력신호로부터 잡음성분을 제거하고, 입력신호로부터 잡음성분이 제거되고 남은 신호를 음성신호로 출력한다.

[0046]

또한, 전처리부(200)는 반향파라미터 획득부(100)에서 출력되는 반향파라미터를 기 설정된 반향파라미터와 비교하고, 두 반향파라미터가 다른 경우, 반향파라미터 획득부(100)에서 출력되는 반향파라미터에 따라서 잡음제거 알고리즘에 사용되는 파라미터를 가변할 수 있다. 여기서, 반향파라미터 획득부(100)에서 출력되는 반향파라미터와 비교 대상이 되는 파라미터는, 차량 출고 시 디폴트(default)로 설정된 반향파라미터 또는 현재 적용 중인 반향파라미터일 수 있다.

[0047]

또한, 반향 파라미터에 의해 가변되는 잡음제거 알고리즘의 파라미터는, 윈도우(window) 사이즈, 이득보정계수, 잡음제거 알고리즘의 반복횟수 등을 포함할 수 있다.

[0048]

도 2를 참조하면, 전처리부(200)는, 프레임 분할부(210), 주파수 변환부(220), 잡음 제거부(230) 등을 포함할 수 있다.

[0049]

프레임 분할부(210)는 연속적으로 입력되는 입력신호에 윈도우 함수를 적용하는 윈도우잉(windowing)을 수행하여, 입력신호를 프레임(frame) 단위로 분할한다.

[0050]

윈도우잉은, 연속적으로 입력되는 신호를 프레임 단위로 나누기 위해서 입력신호에 윈도우 함수를 적용하는 것을 의미하며, 윈도우 함수로는 구형(rectangular) 윈도우, 해밍(Hamming) 윈도우, 해닝(Hanning) 윈도우 등이 사용될 수 있다.

[0051]

프레임 분할부(210)는 음성신호를 복원 시 프레임의 양 끝에서 신호의 불연속을 방지하거나 주파수 스펙트럼의 왜곡을 막기 위해 윈도우잉을 수행할 수 있다.

[0052]

한편, 프레임 분할부(210)에서 윈도우잉 시 사용하는 윈도우의 사이즈는 잡음 제거 알고리즘의 성능 및 계산량에 큰 영향을 미치는 파라미터이며, 반향파라미터와 상관성을 가진다. 예를 들어, 반향이 길수록 윈도우 사이즈를 크게 설정하고, 반향이 짧을수록 윈도우 사이즈를 작게 설정하는 것이 잡음 제거 알고리즘을 성능을 향상시킬 수 있다.

[0053]

따라서, 본 발명의 실시 예에서 프레임 분할부(210)는 반향파라미터 획득부(100)를 통해 획득한 반향파라미터에

따라서, 윈도우 시 적용하는 윈도우 사이즈를 가변할 수 있다. 예를 들어, 반향파라미터(RT60)가 0.1인 경우에는 윈도우 사이즈는 128탭(tap)으로 설정하고, 반향파라미터(RT60)가 0.3인 경우에는 윈도우 사이즈를 256탭으로 설정할 수 있다.

[0054] 주파수 변환부(220)는 프레임 분할부(210)에서 복수의 프레임으로 분할된 입력신호가 입력되면, 고속푸리에변환(Fast Fourier Transform, FFT)을 적용하여 이를 주파수 영역(frequency domain)의 신호로 변환한다.

[0055] 잡음 제거부(230)는 주파수 변환부(220)에 의해 주파수 변환된 입력신호로부터 주파수 대역 별 잡음성분을 추정하고, 이를 토대로 주파수 신호로 변환된 입력신호에서 잡음성분을 제거할 수 있다. 또한, 입력신호로부터 잡음성분을 제거하고 남은 음성신호를 특징패턴 검출부(300)로 출력한다.

[0056] 한편, 잡음 제거부(230)에서 입력신호로부터 잡음성분을 제거하는 과정은, 음성신호의 왜곡을 유발할 수 있다. 특히, 반향의 정도가 큰 환경일수록, 잡음성분을 제거함에 따라 발생하는 음성신호의 왜곡이 더 커져, 잡음 제거 알고리즘의 신뢰성을 저하시키는 문제가 있다.

[0057] 따라서, 본 발명의 실시 예에서 잡음 제거부(230)는 추정된 잡음성분에 이득보정계수를 적용하고, 이득보정계수를 조절하여 입력신호로부터 잡음성분을 제거하는 정도를 조절할 수 있다. 또한, 잡음 제거부(230)는 반향파라미터 획득부(100)를 통해 획득한 반향파라미터에 따라서 잡음성분에 적용하는 이득보정계수를 조절할 수 있다.

[0058] 전처리부(200)는, 잡음 제거 알고리즘에 사용되는 파라미터가 최적의 성능을 내는 값으로 수렴하도록 전술한 잡음 제거 알고리즘을 반복적으로 수행할 수 있다. 한편, 잡음 제거 알고리즘에 사용되는 파라미터가 최적의 값으로 수렴하는데 걸리는 시간은, 반향 정도에 따라서 다를 수 있다. 예를 들어, 반향이 적은 환경일수록 잡음 제거 알고리즘의 파라미터들의 수렴속도가 빨라지며, 반향이 많은 환경일수록 잡음 제거 알고리즘의 파라미터들의 수렴속도는 느려질 수 있다.

[0059] 따라서, 본 발명의 실시 예에서 잡음 제거부(230)는, 반향파라미터 획득부(100)를 통해 획득한 반향파라미터에 따라서 잡음 제거 알고리즘의 반복 횟수를 결정할 수 있다. 예를 들어, 반향파라미터(RT60)가 0.1인 경우에는 반복 횟수를 최대 30번으로 설정하고, 반향파라미터(RT60)가 0.3인 경우에는 반복 횟수를 최대 50번으로 결정할 수 있다.

[0060] 잡음 제거부(230)는, 전술한 과정을 통해 입력신호로부터 잡음성분을 제거하여 획득한 음성신호를 특징패턴 검출부(300)로 출력한다. 여기서, 잡음 제거부(230)는 주파수 역변환을 통해 음성신호를 시간 영역의 신호로 변환하여 특징패턴 검출부(300)로 출력할 수 있다.

[0061] 특징패턴 검출부(300)는 전처리부(200)에 의해 잡음성분이 제거된 후의 음성신호가 입력되면, 이를 분석하여 시작점(start point) 및 끝점(end point)을 가지는 음성구간을 추출한다. 또한, 추출한 음성구간의 신호를 분석하여 음성신호의 특징패턴을 검출한다.

[0062] 음성 인식부(400)는 특징패턴 검출부(300)를 통해 검출된 특징패턴을 기 설정된 기준음성과 비교하고, 비교 결과를 토대로 음성을 인식한다.

[0063] 도 3은 본 발명의 일 실시 예에 따른 음성 인식 장치에서의 음성 인식 방법을 도시한 흐름도이다. 또한, 도 4는 본 발명의 일 실시 예에 따른 음성 전처리 장치에서의 전처리 방법을 도시한 흐름도이다.

[0064] 도 3을 참조하면, 입력신호가 입력되면, 전처리부(200)는 입력신호에 잡음 제거 알고리즘을 적용하는 전처리 과정을 수행한다(S100).

[0065] 이하, 도 4를 참조하여, 상기 S100 단계의 전처리 과정을 상세하게 설명한다.

[0066] 도 4를 참조하면, 반향파라미터 획득부(100)는 잡음 제거 알고리즘에 사용되는 파라미터들의 갱신 여부를 결정하기 위해, 차량 내 환경에 따라서 적응적으로 반향파라미터를 획득한다(S200).

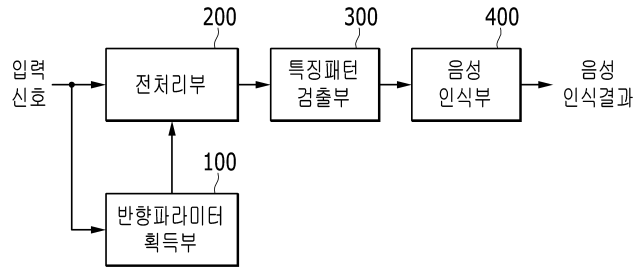
[0067] 상기 S200 단계에서, 반향파라미터 획득부(100)는, 입력신호의 음의 방향 분산을 이용하여 반향파라미터를 추정하거나, 동승자의 동승위치에 따라서 반향파라미터를 선택할 수 있다.

[0068] 반향파라미터가 획득되면, 전처리부(200)는 이를 기 설정된 반향파라미터와 비교한다(S210). 여기서, 비교 대상이 되는 기 설정된 반향파라미터는, 차량 출고 시 디폴트로 설정된 반향파라미터, 또는 현재 잡음 제거 알고리즘에 적용 중인 반향파라미터를 포함할 수 있다.

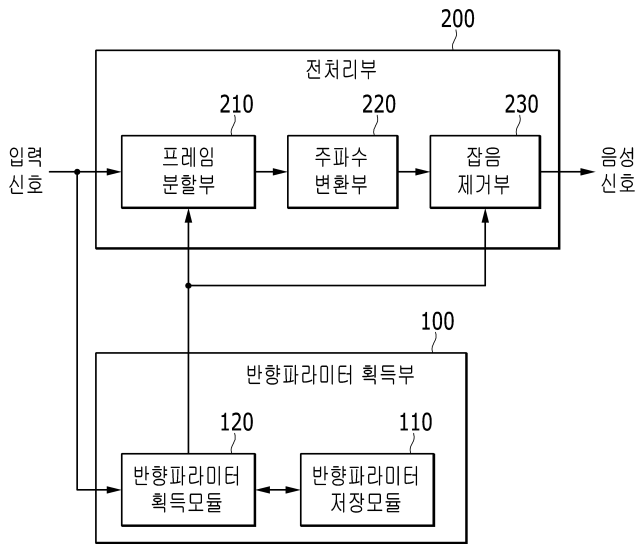
- [0069] 상기 S210 단계에서, 두 반향파라미터가 서로 다른 경우, 전처리부(200)는 상기 S200 단계에서 획득한 반향파라미터를 토대로, 잡음 제거 알고리즘의 파라미터들을 갱신하고(S220), 갱신된 파라미터들을 이용하여 잡음 제거 알고리즘을 수행한다.
- [0070] 즉, 전처리부(200)는 입력신호에 윈도우 함수를 적용하여, 입력신호를 프레임 단위로 분할하고(S230), 주파수 영역의 신호로 변환한 뒤(S240), 주파수 대역 별로 잡음성분을 제거하여 출력(S250)하는 잡음 제거 알고리즘을 수행한다.
- [0071] 상기 S220 단계에서, 반향파라미터를 토대로 갱신되는 파라미터는, 상기 S230 단계에서 사용되는 윈도우 함수의 윈도우 사이즈, 상기 S250 단계에서 잡음성분에 적용되는 이득보정계수, 잡음 제거 알고리즘의 반복횟수 등을 포함할 수 있다.
- [0072] 상기 S210 단계에서, 두 반향파라미터가 동일한 경우, 전처리부(200)는 파라미터 갱신 없이 상기 S230 단계 내지 S250 단계의 잡음 제거 알고리즘을 수행한다.
- [0073] 다시, 도 3을 보면, 전처리부(200)에 의해 입력신호에 포함된 잡음성분이 제거된 후 음성신호가 출력되면, 특징 패턴 검출부(300)는 이를 분석하여 음성신호의 특징패턴을 검출한다(S110).
- [0074] 음성 인식부(400)는 특징패턴 검출부(300)를 통해 검출된 특징패턴이 입력되면, 기를 기 설정된 기준 음성의 특징 패턴들과 비교하고, 비교 결과를 토대로 음성인식결과를 출력한다(S120).
- [0075] 전술한 바에 따르면, 본 발명의 실시 예에서는 차량 내 환경 변화에 따라서 적응적으로 반향파라미터를 획득하고, 이를 토대로 전처리 과정에서 잡음 제거 알고리즘에 사용되는 파라미터들을 갱신한다.
- [0076] 따라서, 차량 출고 시 잡음 제거 알고리즘의 파라미터들이 고정되는 기존의 방식에 비해, 잡음 제거 성능을 향상시킬 수 있으며, 차량 내 환경 변화에 따라 잡음 제거 과정에서 음성신호가 왜곡되는 것을 방지할 수 있다.
- [0077] 본 발명의 실시 예에 의한 음성 전처리 방법 및 음성 인식 방법은 소프트웨어를 통해 실행될 수 있다. 소프트웨어로 실행될 때, 본 발명의 구성 수단들은 필요한 작업을 실행하는 코드 세그먼트들이다. 프로그램 또는 코드 세그먼트들은 프로세서 판독 가능 매체에 저장되거나 전송 매체 또는 통신망에서 반송파와 결합된 컴퓨터 데이터 신호에 의하여 전송될 수 있다.
- [0078] 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록 장치를 포함한다. 컴퓨터가 읽을 수 있는 기록 장치의 예로는, ROM, RAM, CD-ROM, DVD_ROM, DVD_RAM, 자기 테이프, 플로피 디스크, 하드 디스크, 광 데이터 저장장치 등이 있다. 또한, 컴퓨터로 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 장치에 분산되어 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다.
- [0079] 지금까지 참조한 도면과 기재된 발명의 상세한 설명은 단지 본 발명의 예시적인 것으로서, 이는 단지 본 발명을 설명하기 위한 목적에서 사용된 것이지 의미 한정이나 특허청구범위에 기재된 본 발명의 범위를 제한하기 위하여 사용된 것은 아니다. 그러므로 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 용이하게 선택하여 대체할 수 있다. 또한 당업자는 본 명세서에서 설명된 구성요소 중 일부를 성능의 열화 없이 생략하거나 성능을 개선하기 위해 구성요소를 추가할 수 있다. 뿐만 아니라, 당업자는 공정 환경이나 장비에 따라 본 명세서에서 설명한 방법 단계의 순서를 변경할 수도 있다. 따라서 본 발명의 범위는 설명된 실시형태가 아니라 특허청구범위 및 그 균등물에 의해 결정되어야 한다.

도면

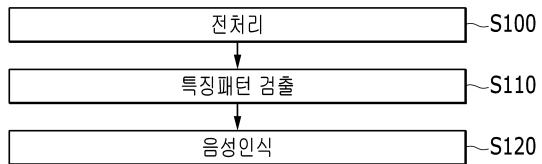
도면1



도면2



도면3



도면4

