



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년11월02일
(11) 등록번호 10-2172608
(24) 등록일자 2020년10월27일

(51) 국제특허분류(Int. Cl.)
HO4M 9/08 (2006.01)

(52) CPC특허분류
HO4M 9/082 (2013.01)

(21) 출원번호 10-2019-0051223

(22) 출원일자 2019년05월02일
심사청구일자 2019년05월02일

(56) 선행기술조사문헌
KR1020000076309 A*

Kunarak, Sunisa. "ADAPTIVE NON-LINEAR NETWORK FILTER ESTIMATION ERROR FOR STEREO ECHO CANCELLATION IN HOME THEATRE 9.1 SURROUND SOUND SYSTEM." International Journal of GEOMATE PP.17-22 (2018.09.) 1부.*

KR101934636 B1

KR1020020055806 A

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

남서울대학교 산학협력단

충청남도 천안시 서북구 성환읍 대학로 91, 남서울대학교내

(72) 발명자

이행우

충청남도 천안시 서북구 성거읍 천흥2길 30-17

(74) 대리인

김건수

전체 청구항 수 : 총 5 항

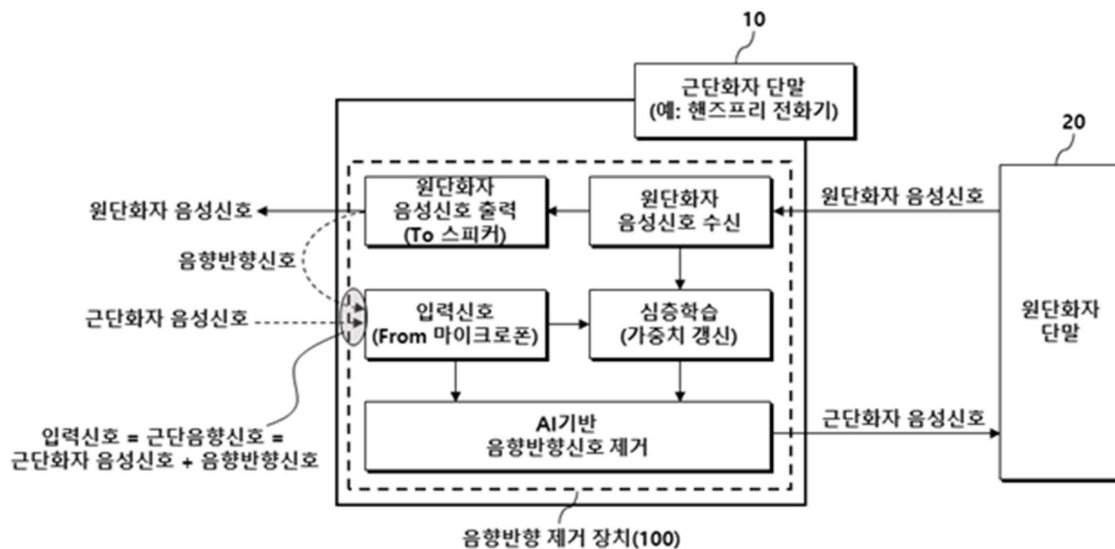
심사관 : 구영희

(54) 발명의 명칭 동시통화에 강인한 심층학습 기반 음향반향 제거 장치 및 그 방법

(57) 요약

본 발명은 동시통화에 강인한 심층학습 기반 음향반향 제거 장치 및 그 방법에 관한 것으로, 심층학습을 포함하는 기계학습 알고리즘과 원단화자의 음성신호를 이용하여 음향반향 신호를 제거하기 위한 심층학습 네트워크의 가중치를 실시간으로 갱신할 수 있도록 함으로써, 실제 근단음향신호에 포함된 음향반향신호를 실시간으로 제거하여 원단화자에게 송신할 수 있도록 함으로서, 동시통화구간에서도 고품질의 음성통화를 수행할 수 있도록 하는 동시통화에 강인한 심층학습 기반 음향반향 제거 장치 및 그 방법에 관한 것이다.

대표도 - 도1



명세서

청구범위

청구항 1

수신단을 통해 수신한 원단화자의 음성신호를 심층학습하여, 상기 수신한 원단화자의 음성신호에 의해 야기되는 음향반향신호를 제거하기 위한 심층학습 네트워크의 가중치를 실시간으로 갱신하는 심층학습부; 및

상기 심층학습 네트워크를 토대로 마이크로폰으로 입력되는 근단화자의 음성신호와 상기 음향반향신호를 포함하는 근단음향신호로부터 상기 음향반향신호를 실시간으로 제거함으로써, 상기 수신한 원단화자의 음성신호가 상기 원단화자로 되돌아가는 것을 방지하는 AI기반 음향반향신호 제거부;를 포함하며,

상기 심층학습부는,

상기 원단화자의 음성신호를 이용하여 상기 심층학습을 수행함으로써, 동시통화 또는 단일통화 구간에 상관없이 상기 심층학습 네트워크의 가중치를 지속적으로 갱신하는 것을 특징으로 하는 동시통화에 강인한 심층학습 기반 음향반향 제거 장치.

청구항 2

삭제

청구항 3

청구항 1에 있어서,

상기 심층학습부는,

상기 원단화자의 음성신호를 심층학습하도록 NLMS(normalized least mean square) 알고리즘 및 오류 역전파 알고리즘을 통해 상기 심층학습 네트워크의 각 레이어에 대한 가중치와 상기 각 레이어를 구성하는 각 노드에 대한 출력값을 실시간으로 갱신함으로써, 상기 심층학습 네트워크를 통해 학습한 원단화자의 음성신호가 마이크로폰을 통해 입력된 근단음향신호에 포함된 음향반향신호에 수렴하도록 하는 것을 특징으로 하는 동시통화에 강인한 심층학습 기반 음향반향 제거 장치.

청구항 4

청구항 1에 있어서,

상기 AI기반 음향반향신호 제거부는,

상기 학습한 심층학습 네트워크에 원단화자의 음성신호를 미리 설정한 소정 개수의 지연요소에 순차적으로 입력하여, 상기 마이크로폰을 통해서 입력된 근단음향신호에 포함된 음향반향신호를 제거하여 출력하는 것을 특징으로 하는 동시통화에 강인한 심층학습 기반 음향반향 제거 장치.

청구항 5

수신단을 통해 수신한 원단화자의 음성신호를 심층학습하여, 상기 수신한 원단화자의 음성신호에 의해 야기되는 음향반향신호를 제거하기 위한 심층학습 네트워크의 가중치를 실시간으로 갱신하는 심층학습 단계; 및

상기 심층학습 네트워크를 토대로 마이크로폰으로 입력되는 근단화자의 음성신호와 상기 음향반향신호를 포함하는 근단음향신호로부터 상기 음향반향신호를 실시간으로 제거함으로써, 상기 수신한 원단화자의 음성신호가 상기 원단화자로 되돌아가는 것을 방지하는 AI기반 음향반향신호 제거 단계;를 포함하며,

상기 심층학습 단계는,

상기 원단화자의 음성신호를 이용하여 상기 심층학습을 수행함으로써, 동시통화 또는 단일통화 구간에 상관없이 상기 심층학습 네트워크의 가중치를 지속적으로 갱신하는 것을 특징으로 하는 동시통화에 강인한 심층학습 기반 음향반향 제거 방법.

청구항 6

청구항 5에 있어서,

상기 심층학습 단계는,

상기 원단화자의 음성신호를 심층학습하도록 NLMS(normalized least mean square) 알고리즘 및 오류 역전파 알고리즘을 통해 상기 심층학습 네트워크의 각 레이어에 대한 가중치와 상기 각 레이어를 구성하는 각 노드에 대한 출력값을 실시간으로 갱신함으로써, 상기 심층학습 네트워크를 통해 학습한 원단화자의 음성신호가 상기 마이크로폰을 통해 입력된 근단음향신호에 포함된 음향반향신호에 수렴하도록 하는 것을 포함하며,

상기 AI기반 음향반향신호 제거 단계는,

상기 학습한 심층학습 네트워크에 원단화자의 음성신호를 미리 설정한 소정 개수의 지연요소에 순차적으로 입력하여, 상기 마이크로폰을 통해서 입력된 근단음향신호에 포함된 음향반향신호를 제거하여 출력하는 것을 특징으로 하는 동시통화에 강인한 심층학습 기반 음향반향 제거 방법.

발명의 설명

기술 분야

[0001] 본 발명은 동시통화에 강인한 심층학습 기반 음향반향 제거 장치 및 그 방법에 관한 것으로, 더욱 상세하게는 심층학습을 포함하는 기계학습 알고리즘과 원단화자의 음성신호를 이용하여 음향반향 신호를 제거하기 위한 심층학습 네트워크의 가중치를 실시간으로 갱신할 수 있도록 함으로써, 실제 근단음향신호에 포함된 음향반향신호를 실시간으로 제거하여 원단화자에게 근단화자 음성신호를 송신할 수 있도록 함으로써, 동시통화구간에서도 고품질의 음성통화를 수행할 수 있도록 하는 동시통화에 강인한 심층학습 기반 음향반향 제거 장치 및 그 방법에 관한 것이다.

배경 기술

[0002] 일반적으로 음향반향 제거 장치(acoustic echo canceller, AEC)는 핸드프리 전화기나 스마트폰과 같은 음성통화 수단을 이용하여 영상통화 등과 같이 복수의 사용자가 동시통화를 수행하기 위한 통신시스템에서, 상기 음성통화수단에 구비되는 스피커로부터 출력되는 음성신호가 직간접적으로 상기 음성통화수단의 마이크로폰으로 재입력되는 음향반향신호를 제거하는 디바이스를 의미한다.

[0003] 이러한 음향반향신호는, 원단화자로부터 수신되는 음성신호(즉, 수신신호)가 상기 통화수단에 구비되는 스피커(예: 라우드 스피커)를 통해 방사(즉, 출력)된 후, 상기 마이크로폰에서 근단화자의 음성신호(즉, 송신신호)와 결합됨으로써, 원단화자가 상기 근단화자의 음성을 청취할 때 명료한 청취를 방해받게 된다.

[0004] 즉, 상기 음향반향신호가 발생하는 경우, 상기 원단화자는 상기 통신시스템의 왕복시간만큼 지연된 자신의 음성을 듣게 되거나, 근단화자의 음성을 명확하게 인식하지 못하는 문제점이 발생될 수 있다.

[0005] 이러한 문제점을 해결하기 위해 종래의 음향반향 제거 장치는, 유한 임펄스응답(finite impulse response, FIR)필터를 사용하여 상기 스피커와 상기 마이크로폰 사이의 음향 임펄스응답을 적응적으로 수렴하도록 함으로써, 상기 마이크로 입력되는 음성신호로부터 음향반향신호를 제거하도록 구현된다.

[0006] 그러나 종래의 유한 임펄스 응답필터를 사용한 음향반향 제거 장치는, 원단화자의 음성신호만 존재하는 경우에만 정상적으로 동작하고, 근단화자의 음성신호도 함께 존재하는 동시통화 구간에서는 음향반향신호를 제거하는 능력이 급격하게 저하되는 현상이 발생하는 문제점이 있었다.

[0007] 이때, 근단화자의 음성신호의 존재는 상기 유한 임펄스 응답필터의 적응 알고리즘의 수렴을 심각하게 저하시키며, 심지어 상기 유한 임펄스 응답필터에 대한 필터계수가 발산하는 요인이 되어 명료한 근단화자의 음성신호를 원단화자에게 송신할 수 없는 문제점이 발생된다.

[0008] 이러한 문제점을 해결하기 위해 최근에는 동시통화상태를 검출하여 반향제거기의 유한 임펄스 응답필터의 계수를 갱신하는 것을 중지시켜, 상기 계수의 발산을 방지할 수 있도록 하는 방법이 개발되어 상용화되고 있다.

[0009] 그러나 동시통화상태를 검출하여 유한 임펄스 응답필터의 계수의 갱신을 중지시키는 상기 방법은, 상기 동시통화상태를 검출하는데 장시간 소요되어 상기 동시통화상태가 검출되기 전에 상기 유한 임펄스 응답필터의 계수가

발산하는 문제점이 있어 정확하게 음향반향신호를 제거하여 명료한 음성신호를 전송하는데 그 한계가 있다.

- [0010] 이에 따라 본 발명에서는, 일반적인 원단화자의 음성신호와 심층학습을 이용하여 음향반향신호를 정확하게 예측하기 위한 심층학습 네트워크의 가중치 및 해당 심층학습 네트워크의 각 레이어를 구성하는 복수의 노드에 대한 출력값을 실시간으로 갱신함으로써, 동시통화상태에서도, 실제 마이크로폰으로 입력되는 근단화자의 음성신호에 포함된 음향반향신호를 정확하고 효과적으로 제거할 수 있도록 하는 방안을 제안하고자 한다.
- [0011] 즉, 본 발명은, 상기 심층학습을 통해 동시통화상태를 감지하기 위한 별도의 감지회로를 구비하지 않고서도 마이크로폰으로 입력되는 입력신호로부터 음향반향신호를 효율적으로 제거하여 근단화자의 명료한 음성신호를 상기 원단화자로 송신함으로써, 상기 원단화자가 상기 근단화자의 음성을 정확하게 인식할 수 있도록 하는 것이다.
- [0012] 다음으로 본 발명의 기술분야에 존재하는 선행기술에 대하여 간단하게 설명하고, 이어서 본 발명이 상기 선행기술에 비해서 차별적으로 이루고자 하는 기술적 사항에 대해서 기술하고자 한다.
- [0013] 먼저 한국등록특허 제1083710호(2011.11.09.)는 적응 필터의 수렴을 이용한 음향반향 제거 장치 및 방법에 관한 것으로, 원단화자의 음신신호와 근단화자의 음성신호가 동시에 입력되는지에 대한 여부를 판단하여 동시통화 구간인지 또는 단일 통화 구간인지를 검출하여, 상기 검출한 결과에 따라 서로 다른 필터 계수를 이용하여 음향반향 제거를 수행하는 적응 필터의 수렴을 이용한 음향반향 제거 장치 및 방법에 관한 것이다.
- [0014] 즉, 선행기술은, 동시통화 구간인지에 대한 여부를 검출하여 동시통화 구간인 경우에는 이전에 저장된 적응 필터의 계수를 이용하여 상기 음향반향 신호를 제거하고, 단일통화 구간인 경우에는 상기 적응 필터의 계수를 갱신하여 상기 갱신한 적응 필터의 계수를 이용하여 상기 음향반향 신호를 제거하는 것이다.
- [0015] 그러나 상기 선행기술은 동시통화 구간에 대한 여부를 검출하여 상기 검출한 결과에 따라 적응 필터의 계수를 달리 적용하는 것으로 기재되어 있으나, 동시통화상태 검출 및 유한 임펄스 응답필터를 사용하는 종래의 음향반향신호 제거 장치와 동일한 구성을 가지고 있기 때문에 동시통화상태가 검출되기 전에 상기 적응 필터의 계수가 발산될 수 있는 문제점을 내포하고 있다.
- [0016] 반면에 본 발명은, 동시통화상태에 상관없이 원단화자의 음성신호를 이용하여 음향반향 신호를 제거하기 위한 심층학습 네트워크의 가중치를 갱신하기 위한 심층학습을 지속적으로 수행함으로써, 마이크로폰으로부터 입력되는 입력신호로부터 상기 원단화자의 음성신호에 대한 음향반향신호를 실시간으로 제거하여 명료한 음성신호를 원단화자에게 제공할 수 있도록 하는 것으로, 상기 선행기술은 이러한 본 발명의 기술적 특징을 기재하거나 시사하고 있지 않다.
- [0017] 또한 한국등록특허 제0545832호(2006.01.18.)는 간섭신호에 강인한 음향반향 제거장치에 관한 것으로, 복수의 마이크 입력 신호간의 레벨 차, 지연 차 및 상관도를 계산하여 공간 정보를 생성하고, 상기 생성한 공간 정보를 토대로 동시 통화 구간을 검출하여, 동시통화 구간인 경우 최근에 사용된 적응 필터의 계수를 이용하여 원화자(즉, 원단화자)의 음성신호로부터 음향반향 추정 신호를 생성하여 마이크로 입력되는 입력신호에서 상기 추정된 음향반향 추정 신호를 감산하는 간섭신호에 강인한 음향반향 제거장치에 관한 것이다.
- [0018] 즉, 상기 선행기술은, 동시통화 구간 여부에 따라 최근에 사용한 적응 필터의 계수를 이용하여 음향반향신호를 제거하는 것으로, 원단화자의 음성신호를 이용하여 음향반향신호를 제거하기 위한 심층학습네트워크에 대한 심층학습을 수행하는 것이 아니며, 상기 심층학습을 통해 실제 마이크로폰으로 입력되는 입력신호에 포함된 음향반향신호를 정확하고 효율적으로 제거할 수 있도록 하는 기술적 구성을 전혀 기재하고 있지 않다. 따라서 상기 선행기술과 본 발명은 서로 상이한 기술적 특징을 가진 것이 분명하다.

발명의 내용

해결하려는 과제

- [0019] 본 발명은 상기와 같은 문제점을 해결하기 위해 창안된 것으로서, 심층학습 알고리즘을 이용하여 음향반향신호를 제거하기 위한 심층학습 네트워크의 가중치를 지속적으로 학습하여, 마이크로폰으로 입력되는 근단화자의 음성신호에 포함된 음향반향신호를 정확하게 예측하여 실시간으로 제거함으로써, 근단화자의 명료한 음성신호를 원단화자에게 제공할 수 있도록 하는 동시통화에 강인한 심층학습 기반 음향반향 제거 장치 및 그 방법을 제공하는 것을 그 목적으로 한다.
- [0020] 또한 본 발명은, 일반적인 원단화자의 음성신호를 심층학습하여, 심층학습한 결과에 따라 상기 마이크로폰으로

입력되는 입력신호에 포함된 음향반향신호를 제거할 수 있도록 함으로써, 동시통화상태를 검출하기 위한 별도의 검출수단을 구비하지 않고서도 상기 음향반향신호를 실시간으로 제거할 수 있도록 하는 동시통화에 강인한 심층 학습 기반 음향반향 제거 장치 및 그 방법을 제공하는 것을 또 다른 목적으로 한다.

[0021] 즉, 본 발명은, 적응필터 및 동시통화여부를 검출하기 위한 검출수단을 이용한 종래의 음향반향 제거 장치와 달리, 심층학습 알고리즘과 상기 원단화자의 음성신호를 이용하여 음향반향신호를 제거하기 위한 심층학습 네트워크의 가중치를 지속적으로 갱신함으로써, 동시통화상태에 상관없이 효율적으로 상기 음향반향신호를 제거할 수 있도록 하는 것이다.

과제의 해결 수단

[0022] 본 발명의 일 실시예에 따른 수신단으로 입력되는 원단화자의 음성신호를 심층학습하여, 상기 원단화자의 음성신호에 의해 야기되는 음향반향신호를 제거하기 위한 심층학습 네트워크의 가중치를 실시간으로 갱신하는 심층 학습부 및 상기 심층학습 네트워크를 통해 근단음향 입력신호에 포함된 음향반향신호를 실시간으로 제거하는 AI 기반 음향반향신호 제거부를 포함하는 것을 특징으로 한다.

[0023] 또한 상기 심층학습부는, 상기 원단화자의 음성신호를 이용하여 상기 심층학습을 수행함으로써, 동시통화 또는 단일통화 구간에 상관없이 상기 심층학습 네트워크의 가중치를 갱신하는 것을 특징으로 한다.

[0024] 또한 상기 심층학습부는, 상기 원단화자의 음성신호를 심층학습하도록 NLMS(normalized least mean square) 알고리즘 및 오류 역전파 알고리즘을 통해 상기 심층학습 네트워크의 각 레이어에 대한 가중치와 상기 각 레이어를 구성하는 각 노드에 대한 출력값을 실시간으로 갱신함으로써, 상기 심층학습 네트워크를 통해 학습한 원단화자의 음성신호가 마이크로폰을 통해 입력된 근단음향 입력신호에 포함된 음향반향신호에 수렴하도록 하는 것을 특징으로 한다.

[0025] 또한 상기 AI기반 음향반향신호 제거부는, 상기 학습한 심층학습 네트워크에 원단화자의 음성신호를 미리 설정한 소정 개수의 지연요소에 순차적으로 입력하여, 마이크로폰을 통해서 입력된 근단음향 입력신호에 포함된 음향반향신호를 제거하여 출력하도록 하는 것을 특징으로 한다.

[0026] 아울러 본 발명의 일 실시예에 따른 동시통화에 강인한 심층학습 기반 음향반향 제거 방법은, 수신단으로 입력되는 원단화자의 음성신호를 심층학습하여, 상기 원단화자의 음성신호에 의해 야기되는 음향반향신호를 제거하기 위한 심층학습 네트워크의 가중치를 실시간으로 갱신하는 심층학습 단계 및 상기 심층학습 네트워크를 통해 근단음향 입력신호에 포함된 음향반향신호를 실시간으로 제거한 후 원단화자에게 음성신호를 출력하는 AI기반 음향반향신호 제거 단계를 포함하는 것을 특징으로 한다.

[0027] 또한 상기 심층학습 단계는, 상기 원단화자의 음성신호를 심층학습하도록 NLMS(normalized least mean square) 알고리즘 및 오류 역전파 알고리즘을 통해 상기 심층학습 네트워크의 각 레이어에 대한 가중치와 상기 각 레이어를 구성하는 각 노드에 대한 출력값을 실시간으로 갱신함으로써, 상기 심층학습 네트워크를 통해 학습한 원단화자의 음성신호가 마이크로폰을 통해 입력된 근단음향 입력신호에 포함된 음향반향신호에 수렴하도록 하는 것을 포함하며, 상기 AI기반 음향반향신호 제거 단계는, 상기 학습한 심층학습 네트워크에 원단화자의 음성신호를 미리 설정한 소정 개수의 지연요소에 순차적으로 입력하여, 마이크로폰을 통해서 입력된 근단음향 입력신호에 포함된 음향반향신호를 제거하여 출력하도록 하는 것을 특징으로 한다.

발명의 효과

[0028] 이상에서와 같이 본 발명의 동시통화에 강인한 심층학습 기반 음향반향 제거 장치 및 그 방법은 원단화자의 음성신호를 학습하여 음향반향신호를 제거하기 위한 심층학습 네트워크의 가중치를 지속적으로 갱신함으로써, 상기 학습한 결과로 음성통화상황에 관계없이 마이크로폰으로 입력되는 입력신호인 근단음향신호에 포함된 음향반향신호를 실시간으로 정확하게 제거하여, 근단화자의 명료한 음성신호를 원단화자에게 제공함으로써, 고품질의 음성통화를 수행할 수 있도록 하는 효과가 있다.

도면의 간단한 설명

[0029] 도 1은 본 발명의 일 실시예에 따른 동시통화에 강인한 심층학습 기반 음향반향 제거 장치 및 그 방법을 설명하기 위해 나타낸 개념도이다.

도 2는 본 발명의 일 실시예에 따른 동시통화에 강인한 심층학습 기반 음향반향 제거 장치의 구성을 나타낸 블

록도이다.

도 3은 본 발명의 일 실시예에 따른 AI기반 신경망필터의 동작을 설명하기 위해 나타난 도면이다.

도 4는 본 발명의 일 실시예에 따른 심층학습 네트워크의 구조와 심층학습과정을 설명하기 위해 나타난 도면이다.

도 5는 본 발명의 일 실시예에 따른 동시통화에 강인한 심층학습 기반 음향반향 제거 장치를 통해 음향반향신호를 제거하는 절차를 나타낸 흐름도이다.

도 6은 종래기술에 따른 유한 임펄스 응답필터를 이용한 음향반향제거 장치에서, 동시통화시 근단화자신호, 음향반향에 대한 ERLE 및 평균오차신호를 나타낸 도면이다.

도 7은 본 발명의 일 실시예에 따른 동시통화에 강인한 심층학습 기반 음향반향 제거 장치에서, 동시통화시 근단화자신호, 음향반향에 대한 ERLE 및 평균오차신호를 나타낸 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0030] 이하, 첨부한 도면을 참조하여 본 발명의 동시통화에 강인한 심층학습 기반 음향반향 제거 장치 및 그 방법에 대한 바람직한 실시예를 상세히 설명한다. 각 도면에 제시된 동일한 참조부호는 동일한 부재를 나타낸다. 또한 본 발명의 실시예들에 대해서 특정한 구조적 내지 기능적 설명들은 단지 본 발명에 따른 실시예를 설명하기 위한 목적으로 예시된 것으로, 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는 것이 바람직하다.
- [0031] 도 1은 본 발명의 일 실시예에 따른 동시통화에 강인한 심층학습 기반 음향반향 제거 장치 및 그 방법을 설명하기 위해 나타난 개념도이다.
- [0032] 도 1에 도시한 바와 같이, 본 발명의 일 실시예에 따른 동시통화에 강인한 심층학습 기반 음향반향 제거 장치(이하 음향반향 제거 장치라 칭함)(100)는, 사용자 단말(10 및 20)에 구비되어 근단화자와 원단화자간의 음성통화시에 발생할 수 있는 음향반향신호를 실시간으로 제거하여 명료하고 깨끗한 품질로 상기 음성통화를 수행할 수 있도록 하는 기능을 지원한다.
- [0033] 한편 상기 근단화자 및 원단화자가 각각 구비한 근단화자 단말(10) 및 상기 원단화자 단말(20)은, 핸드프리 전화기, 헤드셋 전화기, 스마트폰 등과 같이 화상통화 시스템 또는 음성통화 시스템에 이용되는 유무선 통신단말을 의미한다.
- [0034] 즉, 본 발명의 음향반향 제거 장치(100)는, 상기 근단화자 및 원단화자가 각각 구비한 사용자 단말에 탑재되어 마이크로폰으로 입력되는 입력신호(즉, 근단음향신호)로부터 상기 원단화자의 음성신호에 의한 음향반향신호를 제거하고, 상기 음향반향신호를 제거한 근단화자의 음성신호를 원단화자에게 전송할 수 있도록 함으로써, 상기 근단화자 및 원단화자간의 명료한 음성통화를 수행할 수 있도록 하는 것이다.
- [0035] 한편 본 발명에서의 원단화자는, 특정 사용자를 기준으로 상기 특정 사용자의 원격에 위치하여 상기 특정 사용자와 음성통화를 수행하는 사용자를 의미하며, 상기 근단화자는, 상기 원단화자와 음성통화를 수행하는 상기 특정 사용자를 의미한다.
- [0036] 또한 상기 음향반향 제거 장치(100)는, 근단화자가 근단화자 단말(10)을 통해 원격에 위치하는 원단화자와 통화를 수행할 때, 상기 마이크로폰의 입력신호로부터 원단화자의 음성신호에 의해 발생한 음향반향신호를 제거함으로써, 명료한 근단화자의 음성신호를 상기 원단화자에게 제공할 수 있도록 학습을 수행한다.
- [0037] 상기 학습은, 딥러닝과 같은 심층학습을 의미하는 것으로, 음향반향 제거 장치(100)의 스피커를 통해서 출력되는 원단화자의 음성신호를 심층학습하여, 마이크로폰의 입력신호(즉, 근단음향신호)에 포함된 음향반향신호를 예측하여 제거하기 위한 심층학습 네트워크의 가중치를 갱신하기 위해 수행된다.
- [0038] 이때, 상기 음향반향신호는 상기 수신한 원단화자의 음성신호가 스피커를 통해 출력되어 다시 마이크로폰으로 재입력되는 신호를 의미한다.
- [0039] 또한 상기 음향반향 제거 장치(100)는, 상기 원단화자의 음성신호가 입력되는 경우에, 상기 입력되는 원단화자

의 자연스러운 음성신호를 이용하여 상기 심층학습을 수행하기 때문에, 상기 근단화자 및 원단화자에 대한 동시통화 여부에 상관없이 상기 심층학습 네트워크의 가중치를 갱신하기 위한 심층학습을 실시간으로 수행할 수 있다.

- [0040] 또한 상기 음향반향 제거 장치(100)는, 상기 음성통화가 수행되고 상기 마이크로폰을 통해 마이크로폰 입력신호가 입력되는 경우, 상기 마이크로폰 입력신호를 상기 심층학습 네트워크에 적용하여, 해당 입력신호에 포함된 음향반향신호를 예측하여 상기 입력신호에 포함된 음향반향신호를 실시간으로 제거하여 고품질의 음성통화를 수행할 수 있도록 한다.
- [0041] 한편, 상기 마이크로폰을 통해 입력되는 입력신호는, 근단화자의 음성신호, 상기 스피커를 통해 출력되는 원단화자의 음성신호에 의한 음향반향신호 또는 이들의 조합을 포함할 수 있다.
- [0042] 즉, 상기 음향반향 제거 장치(100)는 근단화자와 원단화자간의 음성통화 중에, 원단화자의 사용자 단말(20)로부터 입력되는 원단화자의 자연스러운 음성신호를 이용하여 상기 심층학습 네트워크의 가중치를 갱신하기 위한 심층학습을 실시간으로 적응적으로 수행하는 것이다.
- [0043] 이때, 상기 마이크로폰으로 입력되는 입력신호에, 음향반향신호만 존재하는 경우, 상기 심층학습 네트워크를 이용하여 상기 원단화자의 음성신호에 의한 음향반향신호를 예측하여, 상기 입력신호로부터 상기 예측한 음향반향신호를 제거함으로써, 네트워크의 지연에 따라 원단화자가 자신의 음성신호를 다시 듣게 되는 것을 방지할 수 있다.
- [0044] 또한 상기 마이크로폰으로 입력되는 입력신호가 음향반향신호와 근단화자의 음성신호를 포함하는 경우에도, 실시간으로 가중치가 갱신되는 상기 심층학습 네트워크를 통해 상기 원단화자의 음성신호에 의한 음향반향신호를 예측하게 되므로, 상기 근단화자의 음성신호와 상기 음향반향신호가 포함된 상기 마이크로폰 입력신호로부터 상기 예측한 음향반향신호를 제거하여 상기 원단화자로 송신할 수 있도록 함으로써, 상기 원단화자가 근단화자의 음성을 명료하게 인식할 수 있도록 한다.
- [0045] 즉, 본 발명은 근단화자와 원단화자 사이의 통화상태가 동시통화 구간이든, 일방통화 구간이든 상관없이 상기 마이크로폰의 입력신호에 포함된 음향반향신호를 제거할 수 있도록 함으로써, 고품질의 통화음질로 상기 음성통화를 수행할 수 있도록 하는 것이다.
- [0046] 따라서 본 발명은, 유한 임펄스 응답필터(즉, FIR 필터)를 사용하는 종래의 음향반향 제거 장치에서 동시통화 상태에서 발생될 수 있는 문제점을 극복하고, 동시통화 여부를 검출하기 위한 검출수단을 구비하지 않고서도 마이크로폰 입력신호로부터 음향반향신호를 실시간으로 제거함으로써, 고품질의 음성통화를 제공할 수 있도록 한다.
- [0047] 한편, 상기 심층학습 네트워크에 대한 심층학습과 상기 심층학습 네트워크를 이용하여 마이크로폰 입력신호에서 음향반향신호를 제거하는 것은 본 발명의 핵심적인 기술적 특징으로 도 3 및 도 4를 참조하여 상세히 설명하도록 한다.
- [0048] 또한 상기 음향반향신호를 제거하는 것은 다음과 같은 과정을 통해 수행된다.
- [0049] 우선, 상기 음향반향 제거 장치(100)는, 마이크로폰으로부터 입력되는 입력신호를 디지털 신호로 변환하고, 상기 실시간으로 가중치가 갱신되는 심층학습 네트워크에 상기 디지털 신호로 변환한 상기 입력신호를 적용한다.
- [0050] 이때, 상기 심층학습 네트워크의 출력은, 상기 입력신호에 포함된 음향반향신호의 추정치이며, 상기 입력신호는 근단화자의 음성신호, 상기 음향반향신호 또는 이들의 조합을 포함한다.
- [0051] 즉, 상기 심층학습 네트워크는 상기 입력신호에 포함된 음향반향신호를 예측하여 출력하는 것이다.
- [0052] 다음으로, 상기 음향반향 제거 장치(100)는, 상기 수신한 입력신호에서 상기 예측한 음향반향신호를 제거하며, 상기 마이크로폰 입력신호에서 상기 음향반향신호를 제거한 근단화자 음성신호를 원단화자 단말로 전송한다. 상기 제거하는 과정을 AI기반으로 수행하는 것이 본 발명의 특징이다.
- [0053] 이를 통해, 상기 음향반향 제거 장치(100)는, 상기 마이크로폰을 통해 입력되는 입력신호에 포함된 음향반향신호를 제거하여 출력함으로써, 상기 음성통화를 수행하는 근단화자 및 원단화자 사이에서 고품질의 음성을 상호간에 인식할 수 있도록 한다.
- [0054] 도 2는 본 발명의 일 실시예에 따른 동시통화에 강인한 심층학습 기반 음향반향 제거 장치의 구성을 나타낸 블

록도이다.

- [0055] 도 2에 도시한 바와 같이, 본 발명의 일 실시예에 따른 동시통화에 강인한 심층학습 기반 음향 제거 장치(100)는, 원단화자와의 음성통화시에 근단화자의 음성신호를 포함하는 입력신호를 마이크로폰으로부터 입력받아 디지털신호로 변환하여 AI기반 신경망필터에 제공하는 근단음향신호 입력부(110), 상기 음성통화시에 스피커를 통해서 출력되는 원단화자 음성신호를 이용하여 심층학습 네트워크에 대한 심층학습을 수행하고, 상기 심층학습 네트워크를 이용하여 상기 근단음향신호 입력부(110)를 통해 디지털 신호로 변환된 입력신호로부터 음향반향신호를 실시간으로 제거하는 AI기반 신경망필터(120), 상기 원단화자 단말로부터 원단화자의 음성신호를 수신하여 심층학습 네트워크에서 사용하도록 제공하고 상기 수신한 원단화자의 음성신호를 스피커를 통해서 출력하도록 하는 음성신호 수신부(130), 상기 마이크로폰으로부터 입력된 입력신호에서 음향반향신호를 제거한 근단화자 음성신호를 아날로그 신호로 변환하여 상기 원단화자로 송신하는 음성신호 송신부(140) 및 상기 입력신호, 원단화자 음성신호, 심층학습을 수행한 결과 생성되는 가중치 등을 저장하는 메모리(150)를 포함하여 구성된다.
- [0056] 상기 근단음향신호 입력부(110)는, 상기 마이크로폰으로부터 수신한 상기 입력신호를 디지털 신호로 변환하여 상기 AI기반 신경망필터(120)로 제공하는 기능을 수행한다.
- [0057] 상기 입력신호에는 원단화자에 의한 음성신호만 존재하는 단일통화 구간에서는 상기 원단화자의 음성신호에 의한 음향반향신호만이 포함될 수 있다. 또한 상기 근단화자에 의한 음성신호만이 존재하는 단일통화 구간 또는 상기 원단화자와 근단화자의 음성신호가 동시에 존재하는 동시통화 구간에서는 상기 근단화자의 음성신호 및 상기 음향반향신호가 포함될 수 있다.
- [0058] 또한 AI기반 신경망필터(120)는, 상기 원단화자의 음성신호가 수신되는 경우, 상기 수신되는 원단화자의 음성신호를 이용하여, 상기 원단화자의 음성신호에 의해 발생하는 음향반향신호를 예측하여 제거하기 위한 심층학습 네트워크의 가중치를 실시간으로 갱신하도록 심층학습을 수행하는 심층학습부(121) 및 상기 심층학습 네트워크를 이용하여 상기 근단음향신호 입력부(110)로부터 입력되는 입력신호에 포함된 음향반향신호를 제거하는 AI기반 음향반향신호 제거부(122)를 포함하여 구성된다.
- [0059] 상기 심층학습부(121)는, 상기 음성통화시에 원단화자의 음성신호를 음성신호 수신부(130)를 통해서 수신하고, 스피커를 통해서 출력되는 신호, 즉 원단화자의 음성신호를 심층학습하여 상기 심층학습 네트워크의 가중치를 갱신하는 기능을 수행한다.
- [0060] 또한 상기 심층학습부(121)는, 상기 스피커를 통해서 출력되는 원단화자의 음성신호만을 이용하여 상기 심층학습을 수행하기 때문에 상기 음성통화상태가 단일통화 구간이거나, 동시통화구간에 상관없이 상기 심층학습 네트워크의 가중치를 실시간으로 갱신할 수 있다.
- [0061] 즉, 상기 심층학습부(121)는, 근단화자가 상기 근단화자 단말(10)을 통해 원단화자와 음성통화를 수행할 때, 상기 원단화자로부터 원단화자의 음성신호가 입력되는 경우 상기 원단화자의 음성신호를 심층학습하여 상기 음향반향신호를 제거하기 위한 심층학습 네트워크의 가중치를 실시간으로 갱신함으로써, 실제 마이크로폰에서 입력되는 입력신호로부터 상기 음향반향신호를 적응적으로 제거할 수 있도록 하는 것이다.
- [0062] 한편 상기 심층학습과 상기 심층학습 네트워크는 도 4를 참조하여 상세히 설명하도록 한다.
- [0063] 또한 상기 AI기반 음향반향신호 제거부(122)는 상기 심층학습부(121)를 통해 실시간으로 가중치가 갱신되는 상기 심층학습 네트워크를 이용하여 상기 마이크로폰에서 입력되는 실제 입력신호로부터 음향반향신호를 제거함으로써, 명료한 근단화자의 음성신호를 원단화자에게 송신할 수 있도록 하는 기능을 수행한다.
- [0064] 이때, 상기 심층학습 네트워크의 입력은, 상기 음성신호 수신부(130)을 통해서 수신하고 스피커를 통해서 출력되는 원단화자의 음성신호이며, 상기 심층학습 네트워크의 출력은 상기 원단화자의 음성신호에 대한 음향반향신호의 추정치(즉, 예측한 음향반향신호)가 된다.
- [0065] 상기 심층학습 네트워크는, 상기 실시간 심층학습을 통해 실제 마이크로폰에서 입력되는 입력신호에 포함되어 있는 음향반향신호를 예측하도록 학습(즉, 원단화자의 음성신호에 의한 음향반향신호를 학습)된다.
- [0066] 즉, 상기 심층학습부(121)는, 상기 심층학습을 통해 상기 심층학습 네트워크의 가중치를 실시간으로 갱신하여, 상기 심층학습 네트워크가 상기 실제 입력신호로부터 상기 음향반향신호를 예측할 수 있도록 함으로서, 상기 실제 입력신호에 포함된 음향반향신호를 효과적으로 제거하여 명료한 근단화자의 음성신호를 상기 원단화자에게 제공할 수 있도록 한다.

- [0067] 이때, 상기 음향반향신호를 제거한 근단화자의 음성신호는, 디지털 신호이므로, 상기 음성신호 송신부(140)는, 상기 음향반향신호를 제거한 근단화자의 음성신호를 아날로그 신호로 변환하여 원단화자로 송신하는 기능을 수행한다.
- [0068] 또한 상기 메모리(150)는, 상기 심층학습을 통해 갱신한 심층학습 네트워크의 가중치와 상기 심층학습 네트워크를 구성하는 각 레이어의 노드값을 포함하여 상기 음향반향 제거 장치(100)의 동작에 필요한 펌웨어 또는 프로그램을 포함하여 다양한 데이터를 저장한다.
- [0069] 또한 상기 음향반향 제거 장치(100)는, 상기 음향반향신호를 제거하기 위한 상기 각 구성부분의 동작에 필요한 전력을 상기 근단화자 단말(10)로부터 공급받도록 구현될 수 있다.
- [0070] 도 3은 본 발명의 일 실시예에 따른 AI기반 신경망필터의 동작을 설명하기 위해 나타낸 도면이다.
- [0071] 도 3에 도시한 바와 같이, 본 일 실시예에 따른 AI기반 신경망필터(120)는, 마이크로폰을 통해 입력되는 근단화자의 음성신호를 포함하는 근단음향신호인 입력신호로부터 원단화자의 음성신호에 의해 발생된 음향반향신호를 적응적으로 제거하여 명료한 근단화자의 음성신호를 원단화자 단말(20)로 송신하기 위한 것으로, 상기 원단화자의 음성신호가 입력되는 경우 상기 원단화자의 음성신호를 심층학습하는 심층학습부(121) 및 상기 심층학습한 결과를 이용하여 실제 마이크로폰에서 입력되는 근단음향 입력신호로부터 음향반향신호를 실시간으로 제거하는 AI기반 음향반향신호 제거부(122)를 포함하여 구성된다.
- [0072] 또한 상기 심층학습부(121)는, 근단화자가 자신이 구비한 근단화자 단말(10)을 통해 음성통화를 수행하는 경우, 수신단을 통해 수신되는 원단화자의 음성신호(x_i)를 이용하여 상기 심층학습 네트워크의 가중치를 갱신하는 심층학습을 실시간으로 수행하여, 상기 학습한 결과에 따라 실시간으로 갱신되는 심층학습 네트워크의 가중치를 상기 메모리(150)에 저장하는 기능을 수행한다.
- [0073] 상기 심층학습부(121)는, 원단화자로부터 수신되어 스피커로 출력되는 원단화자 음성신호를 상기 심층학습 네트워크에 입력하여, 상기 원단화자 음성신호를 심층학습함으로써, 상기 심층학습 네트워크의 가중치를 갱신할 수 있도록 한다.
- [0074] 즉, 상기 심층학습부(121)는, 종래의 유한 임펄스 응답필터에서 음향반향신호를 제거하기 위해 사용되는 백색잡음을 이용하는 것이 아니라, 상기 원단화자로부터 입력되는 자연스러운 음성신호를 이용하여 상기 심층학습을 수행하게 되는 것이다.
- [0075] 또한 상기 심층학습부(121)는, 상기 원단화자 음성신호를 상기 심층학습 네트워크의 입력레이어에 순차적으로 입력하며, 상기 입력한 원단화자 음성신호를 심층학습하여, 상기 원단화자 음성신호가 스피커를 통해 출력되어 발생될 수 있는 음향반향신호를 예측하기 위해, 상기 심층학습 네트워크를 구성하는 각 레이어의 가중치를 실시간으로 갱신한다.
- [0076] 이때, 상기 심층학습부(121)는, NLMS(normalized least mean square)알고리즘과 역진파(back propagation) 알고리즘을 이용하여 상기 심층학습 네트워크를 구성하는 각 레이어의 가중치를 갱신할 수 있도록 한다.
- [0077] 또한 상기 심층학습부(121)는, 상기 원단화자의 음성신호를 실시간으로 심층학습하여 상기 심층학습 네트워크의 가중치를 갱신할 수 있도록 함으로써, 상기 가중치를 갱신한 심층학습 네트워크를 통해 상기 원단화자의 음성신호로 야기되는 음향반향신호를 정확하게 예측하여 마이크로폰으로부터 입력되는 근단음향 입력신호($S_k + e_k$)에 포함된 음향반향신호를 제거할 수 있도록 한다.
- [0078] 상기 AI기반 음향반향신호 제거부(122)는, 상기 심층학습부(121)를 통해 실시간으로 가중치가 갱신되는 심층학습 네트워크를 토대로 상기 원단화자의 음성신호에 의한 음향반향신호경로 응답에 따라 상기 음향반향신호를 예측하여 상기 마이크로폰에서 입력되는 근단음향 입력신호에 포함된 음향반향신호를 제거할 수 있도록 하는 기능을 수행한다.
- [0079] 이때, 상기 AI기반 음향반향신호 제거부(122)는, 상기 마이크로폰을 통해 입력된 근단음향 입력신호가 입력되는 경우, 상기 가중치를 갱신한 심층학습 네트워크를 이용하여 현재 원단화자의 음성신호로부터 야기되는 음향반향신호를 예측하고, 상기 입력된 근단음향 입력신호로부터 상기 예측한 음향반향신호를 제거하여 오차값을 계산한다.
- [0080] 이후, 상기 AI기반 음향반향신호 제거부(122)는, 상기 계산한 오차값을 상기 심층학습 네트워크를 구성하는 각

레이어의 각 노드에 적용한다.

- [0081] 즉, 상기 AI기반 신경망 필터(120)는, 상기 심층학습을 실시간으로 수행하여, 상기 근단화자가 원단화자와 음성통화를 수행하는 경우, 상기 심층학습 네트워크를 통해 예측되는 음향반향신호가 실제 마이크로폰으로부터 입력되는 입력신호에 포함된 음향반향신호에 효과적으로 수렴할 수 있도록 함으로서, 상기 근단화자와 원단화자에게 고품질의 통화음질을 제공할 수 있도록 하는 것이다.
- [0082] 또한 상기 AI기반 음향신호 제거부(122)는, 상기 마이크로폰을 통해 외부로부터 근단음향 입력신호가 입력되는 경우, 상기 실시간으로 상기 가중치를 갱신한 심층학습 네트워크를 이용하여 상기 근단음향 입력신호에 포함된 음향반향신호를 예측함으로서, 상기 실제 입력되는 근단음향 입력신호로부터 상기 예측한 음향반향신호를 제거한다.
- [0083] 또한 상기 음향반향 제거 장치(100)는, 상기 음향반향신호를 제거한 근단화자 음성신호를 음성신호 송신부(140)를 통해 아날로그 신호로 변환하여 송신단을 통해 상기 변환한 아날로그 신호를 송신함으로서, 상기 원단화자가 근단화자의 명료한 음성신호를 청취할 수 있도록 한다.
- [0084] 이때, 상기 음향반향 제거 장치(100)는, 상기 음향반향신호를 제거한 상기 근단화자 음성신호를 증폭한 후, 상기 아날로그 신호로 변환하여 원단화자로 송신할 수 있음은 상술한 바와 같다.
- [0085] 상기에서 설명한 바와 같이, 본 발명의 AI기반 적응 신경망필터(130)는, 원단화자로부터 수신되는 원단화자 음성신호를 실시간으로 심층학습하여, 실제 입력되는 근단음향신호에 포함된 음향반향신호를 실시간으로 제거하여 상기 원단화자로 송신할 수 있도록 함으로서, 상기 원단화자와 근단화자가 고품질의 음성통화를 수행할 수 있도록 한다.
- [0086] 한편, 도 3에는, AI기반 신경망필터(120)는 상기 심층학습을 수행하는 심층학습부(121) 및 음향반향신호를 제거하는 AI기반 음향반향신호 제거부(122)로 분리되어 도시되어 있으나, 상기 AI기반 신경망필터(120)에서 상기 심층학습 및 음향반향신호 제거를 통합하여 수행되도록 구현될 수도 있다.
- [0087] 도 4는 본 발명의 일 실시예에 따른 심층학습 네트워크의 구조와 심층학습과정을 설명하기 위해 나타낸 도면이다.
- [0088] 도 4에 도시한 바와 같이, 본 발명의 일 실시예에 따른 심층학습 네트워크는 복수의 입력노드 a_i (즉, a_{i1} 내지 a_{i8})로 구성되는 입력레이어, 복수의 히든노드 a_j (즉, a_{j1} 내지 a_{j3})로 구성되는 히든레이어 및 하나의 출력노드 a_k (즉, a_{k1}) 구성되는 출력레이어를 포함하여 구성된다.
- [0089] 상기 복수의 입력노드는, 미리 설정한 개수의 지연요소로 구성되며, 상기 원단화자의 음성신호가 수신되는 경우, 미리 설정한 시간으로 지연시킨 상기 수신한 원단화자의 음성신호를 FIFO 방식으로 순차적으로 입력받는 기능을 수행한다.
- [0090] 한편, 상기 입력레이어의 입력노드와 상기 히든 레이어의 히든노드는 특정 가중치 w_i (즉, w_{i1} 내지 w_{i12})를 가지고 각각 연결되며, 상기 히든 레이어의 히든노드와 상기 출력 레이어의 출력노드 또한 특정 가중치 w_j (즉, w_{j1} 내지 w_{j3})를 가지고 각각 연결된다.
- [0091] 한편 도 4에 도시한 것과 같이, 입력노드, 히든노드, 출력노드는 영문자 a에 영문자(i, j, k) 및 숫자로 구성된 아래첨자로 구성되어 있는 것으로 표현되어 있으며, 상기 도 4에 도시한 가중치 또한 동일한 방식으로 표현된다.
- [0092] 또한 도 4에는, 상기 심층학습 네트워크가 하나의 히든레이어를 포함하는 것으로 도시되어 있으나 복수의 히든레이어로 포함하여 보다 정확한 음향반향신호를 예측할 수 있도록 구현될 수 있다.
- [0093] 또한 상기 가중치를 갱신하는 과정은 상기 수신단을 통해 수신되는 원단화자의 음성신호를 학습함으로서 갱신되며 상기 히든레이어 및 출력레이어의 각 노드의 출력값(즉, 각 노드의 노드값)은 다음의 [수학식 1]에 따라 계산된다.

[0094] [수학식 1]

$$a_j = \sum_{i=1}^l w_{ij} a_i$$

$$a_k = \sum_{j=1}^m w_{jk} a_j$$

[0095]

[0096] 여기서, a_j 는 상기 히든레이어를 구성하는 각각의 히든노드의 출력값을 의미하며, 상기 a_k 는 상기 출력레이어를 구성하는 출력노드의 출력값을 의미하고, 상기 w 는 상기 입력노드와 히든노드 사이의 가중치 및 상기 히든노드와 출력노드 사이의 가중치를 각각 의미한다. 또한 상기 출력노드 a_k 의 출력은 예측한 음향반향신호인 \hat{e}_k 가 된다.

[0097] 이때, 상기 마이크로폰으로부터 입력되는 근단음향 입력신호가 근단화자의 음성신호(s_k)와 상기 원단화자의 음성신호에 의해 야기된 음향반향신호(e_k)를 포함하여 구성되어 있다고 가정한다면, 상기 예측한 음향반향신호를 상

기 입력신호로부터 제거하면 이에 대한 오차값(r_k)은 $r_k = s_k + e_k - \hat{e}_k$ 로 계산될 수 있다.

[0098] 또한 상기 각 레이어의 가중치는, NLMS 알고리즘을 이용하여 다음의 [수학식 2]를 통해 갱신된다.

[0099] [수학식 2]

$$w_j(t+1) = w_j(t) + \eta r_k \frac{a_j}{E\{a_{jm}\}}$$

$$w_i(t+1) = w_i(t) + \eta r_k \frac{a_i}{E\{a_{il}\}} \sum_{j=1}^m w_j$$

[0100]

[0101] 여기서, η 는 학습속도를 결정하는 파라미터를 의미하며, 상기 가중치 w_i 는 오류 역전과 알고리즘을 이용하여 상기 계산한 오차값(r_k)이 가중치 w_j 를 거쳐 전달되어 업데이트된다.

[0102] 한편 상기 심층학습 네트워크가 복수의 히든레이어로 구성되는 경우, 상기 히든레이어에 대한 가중치 또한 상기 각 레이어의 가중치를 갱신하는 동일한 과정을 통해 갱신된다.

[0103] 또한 상기 학습속도는 상기 가중치가 무한대로 증가해버리는 오버 슈팅(over shooting)문제 또는 상기 가중치가 최적의 값으로 갱신되기 전에 상기 학습이 종료되는 문제가 발생되지 않도록 최적의 값으로 미리 설정될 수 있다.

[0104] 이때, 상기 AI기반 신경망필터(120)는, 상기 오차값을 가지는 오차신호를 상기 심층학습 네트워크의 입력레이어에 순차적으로 입력하여 상기 심층학습 네트워크의 가중치와 상기 각 히든레이어 및 출력레이어의 각 노드값(즉 출력값)을 갱신함으로써, 상기 심층학습 네트워크를 통해 예측되는 음향반향신호가 실제 입력되는 마이크로폰(110)에 포함된 음향반향신호에 정확하게 수렴할 수 있도록 한다.

[0105] 도 5는 본 발명의 일 실시예에 따른 동시통화에 강인한 심층학습 기반 음향반향 제거 장치를 통해 음향반향신호를 제거하는 절차를 나타낸 흐름도이다.

[0106] 도 5에 도시한 바와 같이, 본 발명의 일 실시예에 따른 동시통화에 강인한 심층학습 기반 음향반향 제거 장치를 통해 음향반향신호를 제거하는 절차는 우선, 상기 근단화자가 자신이 소지한 근단화자 단말(10)을 통해 원단화자와 음성통화를 수행할 때, 상기 음향반향 제거 장치(100)는, 상기 음향반향 제거 장치(100)의 수신단에서 상기 원단화자의 음성신호가 수신되는 경우(S110), 상기 수신한 원단화자의 음성신호를 이용하여, 상기 원단화자 음성신호에 의해 야기되는 음향반향신호를 예측하여 제거하기 위한 심층학습 네트워크의 가중치를 갱신하기 위

한 심층학습을 수행한다(S120).

- [0107] 상기 심층학습은 근단화자와 원단화자가 서로 음성통화를 수행하지 않을 경우에도 미리 학습을 진행하여 심층학습 네트워크의 가중치를 저장해 둘 수 있다.
- [0108] 상기 심층학습 네트워크는, 수신단으로부터 수신되는 원단화자의 자연스러운 음성신호를 입력으로 하여, 상기 원단화자의 음성신호에 의해 발생될 수 있는 음향반향신호를 예측하여 출력하는 기능을 수행한다.
- [0109] 또한 상기 음향반향 제거 장치(100)는, 상기 AI기반 신경망필터(120)와 상기 원단화자의 음성신호를 이용하여 상기 심층학습을 수행함으로써, 상기 통화상태가 동시통화 구간이거나, 일반통화 구간인 경우에 상관없이 상기 심층학습 네트워크의 가중치를 갱신하기 위한 심층학습을 중단 없이 실시간으로 수행하여 실제 마이크로폰을 통해서 입력되는 근단음향 입력신호에 포함된 음향반향신호를 정확하게 제거할 수 있도록 한다.
- [0110] 이를 통해 상기 음향반향 제거 장치(100)는 상기 동시통화 여부를 검출하기 위한 별도의 검출수단을 구비하지 않고서도 상기 음향반향신호를 제거하여 고품질의 음성통화를 수행할 수 있도록 한다.
- [0111] 다음으로, 상기 마이크로폰을 통해 외부로부터 근단음향 입력신호가 입력되는 경우, 상기 음향반향 제거 장치(100)는 상기 심층학습 네트워크를 이용하여 상기 입력신호에 포함된 음향반향신호를 제거한다(S130).
- [0112] 다음으로, 상기 음향반향 제거 장치(100)는, 상기 근단음향 입력신호로부터 상기 음향반향신호를 제거한 최종적인 근단화자 음성신호를 상기 송신단을 통해 송신(S140)함으로써, 고품질의 통화음질로 상기 음성통화를 수행할 수 있도록 한다.
- [0113] 한편, 상기 입력신호가 음향반향신호로만 구성되어 있는 경우(즉, 근단화자가 말을 하지 않는 경우)에도, 상기 음향반향 제거 장치(100)는 해당 입력신호로부터 상기 음향반향신호를 실시간으로 제거함으로써, 상기 원단화자의 음성신호가 상기 원단화자로 되돌아가는 것을 방지할 수 있다.
- [0114] 또한 상기 입력신호가 근단화자의 음성신호 및 음향반향신호를 포함하고 있는 경우에도, 상기 음향반향 제거 장치(100)는 해당 입력신호로부터 상기 원단화자의 음성신호에 의해 발생된 상기 음향반향신호를 실시간으로 제거함으로써, 상기 원단화자가 해당 근단화자의 음성을 정확하게 인식할 수 있도록 한다.
- [0115] 다음으로, 현재 수행하고 있는 음성통화가 종료되는 경우, 상기 음향반향 제거 장치(100)의 동작도 종료되며, 상기 음성통화가 유지되는 경우에는 상기 S110 단계 내지 S140단계를 반복적으로 수행하여 고품질의 음성통화를 상기 근단화자 및 원단화자에게 제공할 수 있도록 한다(S150).
- [0116] 한편 상기 심층학습 및 상기 음향반향신호를 제거하는 과정이 순차적으로 단계별로 기재되어 있으나, 상기 심층학습 및 상기 음향반향신호를 동시에 실시간으로 수행될 수 있음은 상술한 바와 같다.
- [0117] 이하에서는 도 6 및 도 7을 참조하여 유한 임펄스 응답필터를 이용한 종래 기술의 음향반향제거 장치의 성능과 본 발명의 동시통화에 강인한 심층학습 기반 음향반향 제거 장치의 성능을 비교하도록 한다.
- [0118] 도 6은 종래기술에 따른 유한 임펄스 응답필터를 이용한 음향반향제거 장치에서, 동시통화시 근단화자신호, 음향반향에 대한 ERLE 및 평균오차신호를 나타낸 도면이며, 도 7은 본 발명의 일 실시예에 따른 동시통화에 강인한 심층학습 기반 음향반향 제거 장치에서, 동시통화시 근단화자신호, 음향반향에 대한 ERLE 및 평균오차신호를 나타낸 도면이다.
- [0119] 도 6에 도시한 바와 같이, 유한 임펄스 응답필터를 사용한 종래의 음향반향 제거 장치는, 근단화자의 음성신호와 원단화자의 음성신호가 동시에 존재하는 동시통화 상태에서, 상기 유한 임펄스 응답필터에 대한 필터계수가 발산하는 문제점으로 인해 ERLE(echo reduction loss enhancement)가 매우 낮은 값(단위: dB)으로 유지됨을 알 수 있다.
- [0120] 또한 마이크로폰을 통해서 입력되는 근단음향 입력신호에서 상기 음향반향신호를 제거하여 계산한 오차값을 나타내는 그래프에 있어서도, 상기 오차값이 매우 크고 불규칙한 것을 알 수 있다.
- [0121] 즉, 상기 종래기술에 따른 종래기술에 따른 유한 임펄스 응답필터를 이용한 음향반향제거 장치는, 동시통화시에 상기 원단화자의 음성신호로 인해 발생하는 음향반향신호를 효율적으로 제거하지 못하는 것을 알 수 있다.
- [0122] 결국 종래기술에 따른 음향반향제거 장치는, 상기 음향반향신호에 의해 왜곡된 근단화자의 음성신호를 상기 원단화자에게 제공함으로써, 상기 원단화자가 상기 근단화자의 음성을 정확하게 인식하지 못하는 문제점을 내포하고 있는 것이다.

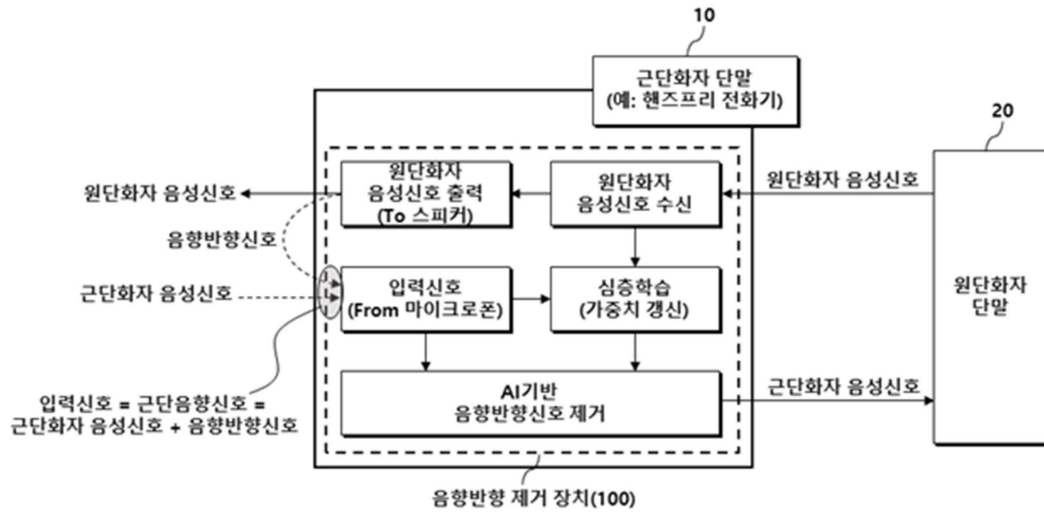
- [0123] 이에 반하여 도 7에 도시한 바와 같이, 본 발명의 일 실시예에 따른 음향반향 제거 장치(100)는, 동시통화 상태에서 ERLE가 매우 높은 값(단위: dB)으로 증가하는 것을 알 수 있다.
- [0124] 즉, 본 발명의 음향반향 제거 장치(100)는, 수신단으로 입력되는 원단화자의 음성신호를 이용하여 상기 원단화자의 음성신호로 인한 음향반향신호를 예측하기 위한 심층학습을 수행하기 때문에 동시통화 구간에서도 안정적으로 상기 음향반향신호를 제거할 수 있음을 알 수 있다.
- [0125] 또한 마이크로폰을 통해서 입력되는 입력신호에서 상기 음향반향신호를 제거한 후의 평균오차값(mean error) 또한 매우 적은 것을 알 수 있다.
- [0126] 이러한 결과로 볼 때, 본 발명의 음향반향 제거 장치(100)의 심층학습 네트워크를 통해 원단화자의 음성신호에 의해 야기되는 음향반향신호를 예측할 때, 상기 예측한 음향반향신호가 실제 마이크로폰을 통해서 입력되는 입력신호에 포함된 음향반향신호에 동일하게 수렴되는 것을 알 수 있다.
- [0127] 따라서 본 발명의 음향반향 제거 장치(100)는, 동시통화 구간에서도 상기 음향반향신호를 실시간으로 제거하여 근단화자의 명료한 음성신호를 상기 원단화자에게 제공할 수 있다.
- [0128] 이상에서 설명한 바와 같이, 본 발명의 일 실시예에 따른 동시통화에 강인한 심층학습 기반 음향반향 제거 장치 및 그 방법은, 음성통화시 원단화자의 음성신호를 심층학습하여 심층학습 네트워크의 가중치를 실시간으로 갱신함으로써, 실제 입력되는 입력신호에 포함된 상기 원단화자의 음성신호에 의한 음향반향신호를 정확하게 예측하여 제거할 수 있도록 함으로써, 명료한 음성통화를 수행할 수 있도록 하는 효과가 있다.
- [0129] 상기에서는 본 발명에 따른 바람직한 실시예를 위주로 상술하였으나, 본 발명의 기술적 사상은 이에 한정되는 것은 아니며 본 발명의 각 구성요소는 동일한 목적 및 효과의 달성을 위하여 본 발명의 기술적 범위 내에서 변경 또는 수정될 수 있을 것이다.
- [0130] 또한, 이상에서는 본 발명의 바람직한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 다양한 변형 실시가 가능한 것은 물론이고, 이러한 변형 실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어서는 안 될 것이다.

부호의 설명

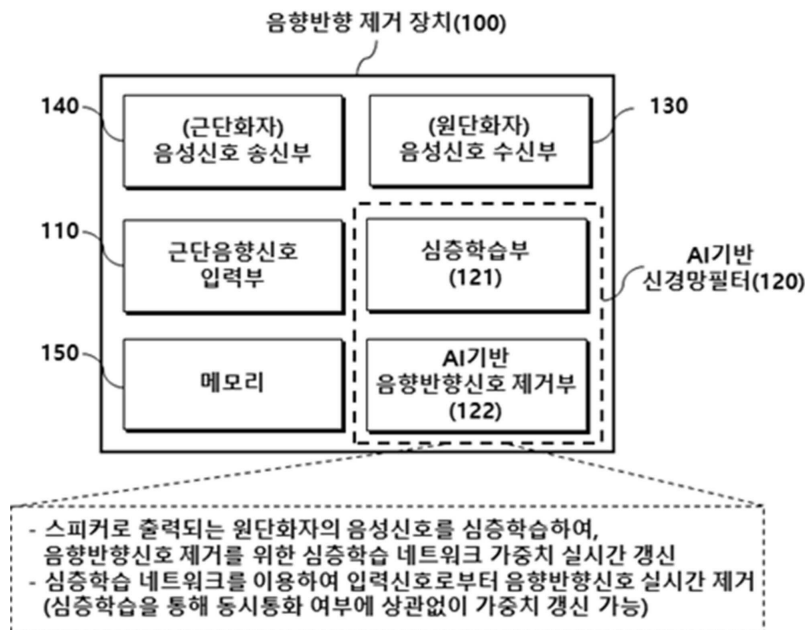
- | | | |
|--------|-----------------------|------------------|
| [0131] | 10 : 근단화자 단말 | 20 : 원단화자 단말 |
| | 100 : 음향반향 제거 장치 | 110 : 근단음향신호 입력부 |
| | 120 : AI기반 신경망필터 | 121 : 심층학습부 |
| | 122 : AI기반 음향반향신호 제거부 | 130 : 음성신호 수신부 |
| | 140 : 음성신호 송신부 | 150 : 메모리 |

도면

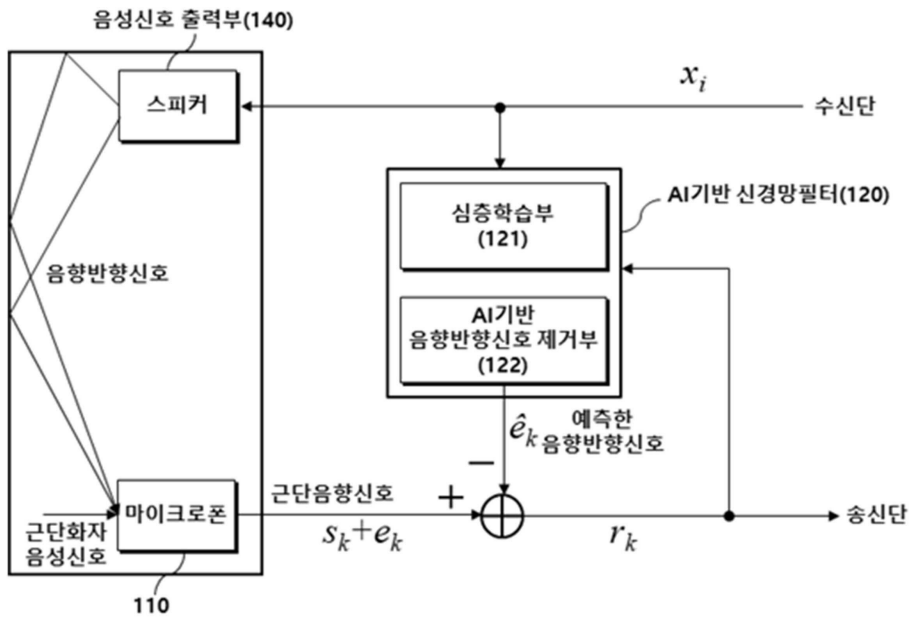
도면1



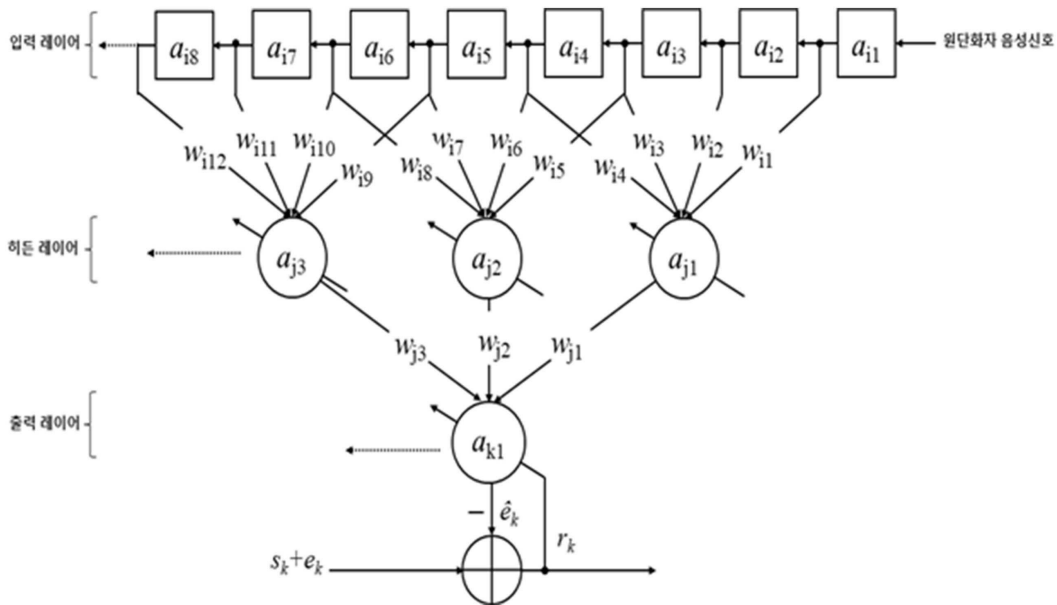
도면2



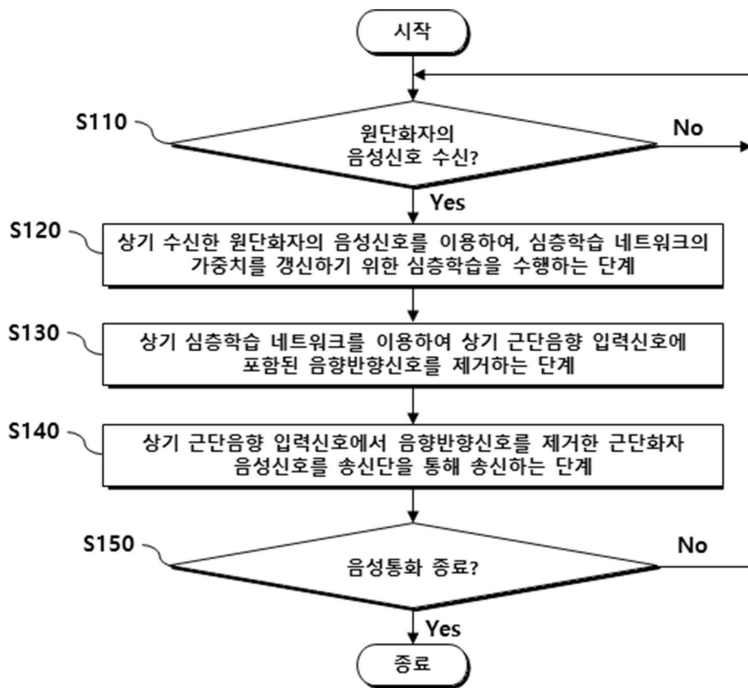
도면3



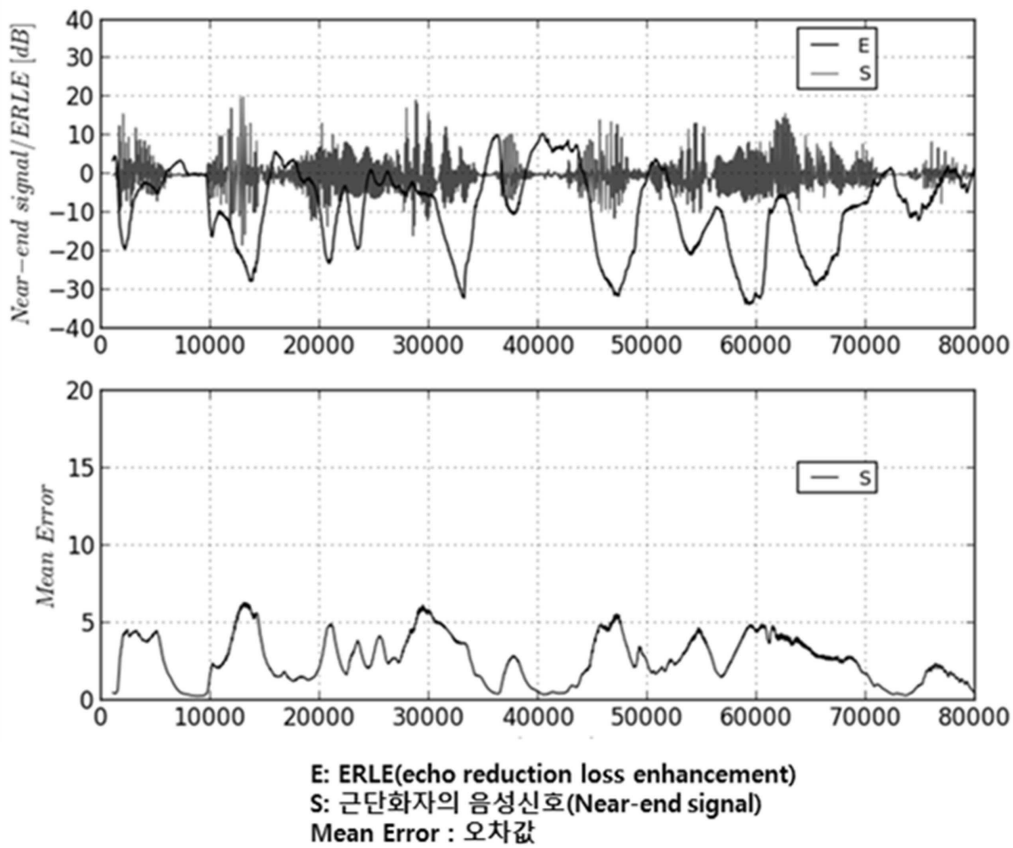
도면4



도면5



도면6



도면7

