

(19)대한민국특허청(KR)  
(12) 공개특허공보(A)

(51) Int. Cl.<sup>7</sup> (11) 공개번호 10-2005-0115857  
G10L 21/02 (43) 공개일자 2005년12월08일

(21) 출원번호 10-2005-7010611  
(22) 출원일자 2005년06월10일  
    번역문 제출일자 2005년06월10일  
(86) 국제출원번호 PCT/US2003/039593 (87) 국제공개번호 WO 2004/053839  
    국제출원일자 2003년12월11일                      국제공개일자 2004년06월24일

(30) 우선권주장 60/432,691 2002년12월11일 미국(US)  
60/502,253 2003년09월12일 미국(US)

(71) 출원인 소프트맥스 인코퍼레이티드  
미국 캘리포니아주 라호야 씨에이 92128 스위트 455 라호야 빌리지 드라이브 4180

(72) 발명자 비세르 에릭  
미국 캘리포니아주 샌디에고 리전츠 로드 #437 8465  
이태원  
미국 캘리포니아주 샌디에고 아로요 그랜드 로드 7257

(74) 대리인 김명신  
김중호  
박장규  
김국현  
김민철

심사청구 : 없음

(54) 안정성 강제하에서 독립 성분 분석을 사용하여 음향을처리하는 시스템 및 방법

요약

오디오 신호의 혼합을 원하는 오디오 신호(430)(예컨대, 음향) 및 노이즈 신호(440)으로 분리하는 시스템 및 방법이 개시된다. 마이크로폰(310, 320)은 혼합된 오디오 신호를 수신하도록 배치되고, 독립 성분 분석(ICA)은 안정성 강제를 사용하여 사운드 혼합을 처리한다(212). ICA 처리(508)는 타깃 사운드 신호(430)을 식별 및 분리하기 위해 원하는 음향 신호의 소정의 특징을 사용한다. 필터 계수는 학습 규칙에 적용되고 필터 웨이트 업데이트 역학은 안정적인 분리된 ICA 신호 결과에 대한 수렴을 지원하기 위해 안정화된다. 분리된 신호는 전처리(214) 및 후처리(220,230) 기술 및 정보를 사용하여 노이즈 효과를 추가로 제거하기 위해 주변 처리될 수 있다. 제안된 시스템은 오디오 통신 하드웨어 환경의 DSP 유닛 또는 CPU 상의 실행을 위해 설계되고 쉽게 적응가능하다.

대표도

도 4

명세서

## 기술분야

본 발명은 오디오 신호 처리 시스템 및 방법에 관한 것으로서, 상세하게는 어쿠스틱 환경에서 음향의 품질을 강화하는 시스템 및 방법에 관한 것이다.

## 배경기술

음향 신호 처리는 일상적인 통신의 많은 영역, 특히 노이즈가 많은 영역에서 중요하다. 실 생활에서의 노이즈는 실 생활에서 에코 및 잔향과 함께 다수의 사운드로 퍼지는 외관상으로 단일의 음원 노이즈를 포함하는 다수의 음원으로 증만하다. 분리되거나 격리되지 않는다면, 배경 노이즈로부터 원하는 노이즈를 추출하는 것은 어렵다. 배경 노이즈는 각 신호로부터 생성된 에코, 반사, 및 잔향 뿐만 아니라, 일반 환경에 의해 생성된 다수의 노이즈 신호, 다른 사람의 배경 대화에 의해 생성된 신호를 포함할 수 있다. 사용자가 종종 잡음이 있는 환경에서 이야기하는 통신에서, 사용자의 음향 신호를 배경 노이즈로부터 분리하는 것이 바람직하다. 셀룰라 폰, 스피커폰, 헤드셋, 보청기, 무선 전화기, 텔레콘퍼런스, CB 라디오, 워키-토키, 컴퓨터 전화통화 애플리케이션, 컴퓨터 및 자동차 음성 명령 애플리케이션, 및 다른 핸드-프리 애플리케이션, 인터콤, 마이크로폰 시스템 등과 같은 음향 통신 매체는 원하는 음향 신호를 배경 노이즈로부터 분리하기 위해서 음향 신호 처리를 이용할 수 있다.

많은 방법은 배경 노이즈 신호로부터 원하는 사운드 신호를 분리하기 위해 창작된다. 종래 기술인 노이즈 필터는 화이트 노이즈 신호와 같은 소정의 특징으로 신호를 식별하고, 입력 신호로부터 이러한 신호를 제거한다. 이러한 방법들은 사운드 신호를 실시간 처리하는데 충분히 단순하고 빠르지만, 다른 사운드 환경에 쉽게 적용될 수 없고, 분해되기 위해 시도된 음향 신호는 실질적인 열화로 귀착될 수 있다. 노이즈 특성에 대한 소정의 가정은 지나치게 포괄적(overinclusive)이거나 덜 포괄적(underinclusive)일 수 있다. 결과적으로, 사람의 음향의 일부분들은 이러한 방법들에 의해 "노이즈"로 간주될 수 있고 그러므로 출력 음향 신호에서 제거될 수 있는 반면, 음악 또는 대화와 같은 배경 노이즈의 일부분들은 이러한 방법들에 의해 노이즈가 아닌 것으로 간주될 수 있고 그러므로 출력 음향 신호에 포함될 수 있다.

독립 성분 분석(ICA : Independent Component Analysis)과 같이 더욱 최근에 개발된 다른 방법은 배경 노이즈로부터 음향 신호의 분리를 위해 상대적으로 날카롭고 유연한 수단을 제공한다. 예컨대, PCT 공보 WO 00/41441은 출력 오디오 신호에서 노이즈를 감소시키는 입력 오디오 신호를 처리하는 특정 ICA 기술을 사용하는 것을 개시한다. ICA는 서로로부터 아마도 독립된 혼합 음원 신호(성분)를 분리하는 기술이다. 이 단순화된 형태에서, 독립 성분 분석은 예컨대, 분리된 신호를 생성하기 위해서 혼합 신호로 행렬을 배가하는 것과 같이, 혼합 신호에 대해 웨이트(weight)의 "비혼합(un-mixing)" 행렬을 연산시킨다. 웨이트는 초기 값으로 할당되고, 그다음 정보의 과잉을 최소화하기 위해서 신호의 결합 엔트로피를 최대화하도록 조절된다. 이 웨이트-조절 및 엔트로피-증가 처리는 신호의 정보 과잉이 최소로 감소할 때까지 반복된다. 이러한 기술은 각 신호의 음원 상에서 정보를 필요로하지 않기 때문에, "블라인드 음원 분리" 방법(BSS : Blind Source Separation)으로 알려져 있다. 블라인드 분리 문제는 다수의 독립 음원으로부터 생기는 혼합 신호를 분리하는 아이디어에 관련된다.

ICA의 가장 이른 논의 중의 하나는 추가적인 연구를 낳은 미국 특허 제5,706,402호의 Tony Bell에 의해서다. 현재 많은 다른 ICA 기술 및 알고리즘이 존재한다. 가장 널리 사용된 알고리즘 및 기술의 요약은 ICA에 관해 여기의 책 및 참조문헌에서 알 수 있다(예컨대, Te-Won Lee, Independent Component Analysis: Theory and Applications, Kluwer Academic Publishers, Boston, 1998년 9월, Hyvarinen 등, Independent Component Analysis, 제 1 판(Wiley-Interscience, 2001년 5월 18일); Mark Girolami, Self-Organizing Neural Networks: Independent Component Analysis and Blind Source Separation(Perspectives in Natural Computing)(Springer Verlag, 1999년 9월); 및 Mark Girolami(편집자), Advances in Independent Component Analysis(Perspectives in Natural Computing)(Springer Verlag, 2000년 8월)). 단일값 분해 알고리즘은 Simon Haykin에 의해 Adaptive Filter Theory(제 3 판, Prentice-Hall(NJ), 1996년)에 개시되었다.

많은 인기있는 ICA 알고리즘은 10년전에 존재만 했던 그들의 중대한 변형에 의해 진화된 숫자를 포함하면서 그들의 성능을 최적화하기 위해서 발전되었다. 예컨대, A.J. Bell 및 TJ Sejnowski, Natural Computation 7:1129-1159(1995), 및 Bell, A.J. 미국 특허 제5,706,402호에 설명된 연구는 일반적으로 그 특허된 형태로 사용되지 않는다. 대신에, 그 성능을 최적화하기 위해서, 이 알고리즘은 다수의 다른 존재에 의해 약간의 재특징화를 거친다. 이러한 한 변화는 Amari, Cichocki, Yang(1996)에 설명된 "자연 그레디언트(natural gradient)"의 사용을 포함한다. 다른 유명한 ICA 알고리즘은 큐물런트(cumulant)와 같은 고차 통계량을 계산하는 방법을 포함한다(Cardoso, 1992; Comon, 1994; Hyvaerinen 및 Oja, 1997).

그러나, 많은 알려진 ICA 알고리즘들은 룸 반사에 기인하는 것들과 같은 어쿠스틱 에코를 본질적으로 포함하는 실제 환경에서 기록되는 신호를 효과적으로 분리할 수 없다. 지금까지 언급된 방법들은 음원 신호의 선형적으로 고정된 혼합으로부터 생기는 신호의 분리에 한정되는 것이 강조되었다. 직접적인 경로 신호 및 그들의 에코의 혼합 신호의 합계로부터 생기는 현상은 잔향으로 칭해지고 인공적인 음향 강화 및 인식 시스템에서 중대한 이슈를 제기한다. 현재, ICA 알고리즘은 이러한 시간-지연되고 에코되어, 효과적인 실시간 사용을 방해하는 신호를 분리시킬 수 있는 긴 필터를 포함하는 것이 필요하다.

도 1은 종래의 ICA 신호 분리 시스템(100)의 일 실시예를 도시한다. 이러한 종래의 시스템에서, 신경망으로서 동작하는 필터의 네트워크는 필터 네트워크로 입력되는 임의의 개수의 혼합 신호로부터 개별적인 신호를 분해하기 위해 동작한다. 도 1에 도시된 바와 같이, 시스템(100)은 입력 신호  $X_1$  및  $X_2$  를 수신하는 두개의 입력 채널(110 및 120)을 포함한다. 신호  $X_1$  에 대해, ICA 다이렉트 필터  $W_1$  및 ICA 크로스 필터  $C_2$ 가 적용된다. 신호  $X_2$  에 대해, ICA 다이렉트 필터  $W_2$  및 ICA 크로스 필터  $C_1$ 이 적용된다. 다이렉트 필터  $W_1$  및  $W_2$  는 다이렉트 조정을 위해 통신한다. 크로스 필터는 그들 각각의 필터링된 신호와 다이렉트 필터에 의해 필터링된 신호를 합치는 피드백 필터이다. ICA 필터의 수렴 이후에, 발생된 출력 신호  $U_1$  및  $U_2$ 는 분리된 신호를 나타낸다.

Torkkola 등에 의한 미국 특허 제5,675,659호는 지연되고 필터링된 음원의 블라인드 분리를 위한 방법 및 장치를 제안한다. Torkkola는 Bell의 특허와 같은 정지 계수 대신에 비혼합 필터를 사용하는 것을 제외하고는 분리된 출력의 엔트로피를 최대화하는 ICA 시스템을 제안한다. 그러나, 결합 엔트로피를 계산하기 위해 및 크로스 필터 웨이트를 조절하기 위해서 Torkkola에 설명된 ICA 계산은 음향 신호와 같은 시간에 따라 변화하는 입력 에너지를 갖는 입력 신호의 존재하에서 숫자 상으로 안정되지 않으며 잔향 인공물을 분리된 출력 신호에 삽입한다. 그러므로 제안된 필터링 스킴은 실제 음향 신호의 안정적이고 지각적으로 받아들일 수 있는 블라인드 음원 분리를 실행하지 않는다.

통상의 ICA 실행은 또한 신호의 결합 엔트로피를 반복적으로 계산하기 위해 및 필터의 웨이트를 조절하기 위해 전력을 실질적으로 계산하는 것이 필요할 때 추가적인 장애물에 직면한다. 많은 ICA 실행은 또한 피드백 필터의 다수의 순회 및 필터의 직접적인 상호관계를 필요로 한다. 결과적으로, 실시간으로 음향의 ICA 필터링을 실행하고 다수의 혼합된 음원 신호를 분리하기 위해서 다수의 마이크로폰을 사용하는 것은 어렵다. 공간적으로 배치된 위치로부터 발생하는 음원의 경우에, 비혼합 필터 계수는 합리적인 양의 필터 탭 및 마이크로폰의 기록으로 계산될 수 있다. 그러나 음원 신호가 진동, 바람에 의한 노이즈 또는 배경 대화로부터 발생한 배경 노이즈와 같은 공간 내에 배치된다면, 마이크로폰 위치에서 기록된 신호는 매우 길고 복잡한 필터 구조 또는 매우 많은 개수의 마이크로폰 중의 어느 하나를 필요로 하는 많은 다른 방향으로부터 방출된다. 임의의 실제 시스템은 전력 및 하드웨어 복잡성을 처리하는 것이 제한되기 때문에, 추가적인 처리 접근은 실시간 음향 신호 강화에 대해 강력한 방법론을 제공하기 위해서는 논의된 ICA 필터 구조를 보충해야만 한다. 이러한 시스템의 계산적인 복잡성은 셀룰라 폰, PDA(Personal Digital Assistants), 오디오 감시 장치, 라디오 등과 같은 작은 소비자 장치의 처리 전력과 호환되어야 한다.

실시간으로 배경 노이즈로부터 음향 신호를 분리할 수 있고 실질적인 계산 전원이 필요치 않으나, 계속해서 상대적으로 정확한 결과를 만들며 다른 환경에 유연하게 적용될 수 있는 단순화된 음향 처리 방법이 바람직하다.

### 발명의 상세한 설명

본 발명은 잡음이 있는 어쿠스틱 환경에서 적어도 하나의 음향 신호와 같이 원하는 오디오 신호를 식별하고 분리하는데 유용한 음향 처리 시스템 및 방법에 관한 것이다. 음향 처리는 무선 모바일 폰, 헤드셋, 또는 셀룰라 폰과 같은 적어도 두개 이상의 마이크로폰을 갖는 장치 상에서 동작한다. 적어도 두개의 마이크로폰은 화자(speaker)로부터의 음향과 같이 타겟으로부터 원하는 신호를 수신하는 장치의 하우징 상에 위치한다. 마이크로폰은 타겟 사용자의 음향을 수신하나, 또한 노이즈, 다른 음원으로부터의 음향, 잔향, 에코, 및 다른 바람직하지 않은 어쿠스틱 신호를 수신한다. 적어도 두개의 마이크로폰은 원하는 타겟 음향 및 다른 원하지 않는 어쿠스틱 정보의 혼합을 포함하는 오디오 신호를 수신한다. 마이크로폰으로부터의 혼합 신호는 변형된 ICA(독립 성분 분석) 처리를 사용하여 처리된다.

본 발명의 일 측면은 각각 오디오 신호의 하나 또는 조합을 포함하는 입력 신호의 적어도 두개의 채널, 및 두개의 개선된 독립 성분 분석 크로스 필터를 포함하는 음향 분리 시스템에 관한 것이다. 입력 신호의 두 채널은 비선형적인 유계 함수(bounded function)를 갖는 바람직하게는 무한 임펄스 응답 필터인 크로스 필터에 의해 필터링된다. 비선형 유계 함수는

예컨대 입력값에 기초하여 양 또는 음의 값 중 어느 하나를 출력으로서 돌려보내는 부호(sign) 함수와 같이, 빠르게 계산될 수 있는 소정 최대값 및 최소값을 갖는 비선형 함수이다. 신호의 반복된 피드백을 따르면, 하나의 채널이 실질적으로 원하는 오디오 신호를 포함하고 다른 채널이 실질적으로 노이즈 신호를 포함하는 출력 신호의 두개의 채널이 생성된다.

본 발명의 일 측면은 오디오 신호를 원하는 음향 신호 및 노이즈 신호로 분리하는 시스템 및 방법에 관한 것이다. 원하는 음향 신호 및 노이즈 신호의 조합인 입력 신호는 적어도 두개의 채널로부터 수신된다. 동일 개수의 독립 성분 분석 크로스 필터가 사용된다. 제 1 채널로부터의 신호는 제 1 크로스 필터에 의해 필터링되고 제 2 채널 상에 증대된 신호를 형성하기 위해 제 2 채널로부터의 신호와 결합된다. 제 2 채널 상의 증대된 신호는 제 2 크로스 필터에 의해 필터링되고 제 1 필터 상의 증대된 신호를 형성하기 위해 제 1 채널로부터의 신호와 결합된다. 제 1 채널 상의 증대된 신호는 제 1 크로스 필터에 의해 추가로 필터링될 수 있다. 필터링 및 결합 처리는 신호의 두개의 채널 사이에서 정보의 중복을 줄이기 위해서 반복된다. 생성된 출력 신호의 두개의 채널은 음향 신호가 지배적인 일 채널 및 비음향 신호가 지배적인 일 채널을 나타낸다. 스펙트럼 차감법(spectral subtraction), 비너 필터링(Wiener filtering), 디-노이징 및 음향 피쳐 추출과 같은 추가적인 음향 강화 방법은 음향 품질을 추가로 개선하기 위해서 실행될 수 있다.

본 발명의 다른 측면은 피드백 필터링 스킴의 설계에서 안정화 요소를 포함하는 것에 관한 것이다. 하나의 안정화 예에서, 필터 웨이트 적응 규칙은 웨이트 적응 역학(dynamics)이 피드백 구조의 전체적인 안정화 요구와 보조를 맞추는 방식으로 설계된다. 이전의 접근과는 다르게, 전체 시스템 성능은 분리된 출력의 원하는 엔트로피 최대화를 지향할 뿐만 아니라 보다 현실적인 목적에 부합하기 위해서 안정성 강제를 고려한다. 이러한 목적은 안정성 강제 하에서 최대 가능성 원칙으로서 더 잘 설명된다. 최대 가능성 평가에서 이러한 안정성 강제는 음원 신호의 시간적인 특징을 모델링하는 것에 상응한다. 엔트로피 최대화 접근에서 신호 음원은 i.i.d.(independently, identically drawn) 랜덤 변수로서 가정된다. 그러나, 사운드 및 음향 신호와 같은 실제 신호는 랜덤 신호가 아니라 시간에서 상호관계를 갖고 주파수가 평탄하다. 이는 상응하는 본래의 ICA 필터 계수 학습 규칙으로 귀결된다.

다른 안정화 예에서, 이러한 학습 규칙은 기록된 입력 진폭에 직접적으로 의존하기 때문에, 입력 채널은 필터 웨이트 적응 속도를 강제하기 위해서 적응형 스케일링 인자에 의해 줄어든다. 스케일링 인자는 순환식(recursive equation)으로부터 판정되고 채널 입력 에너지의 함수이다. 따라서 차후 ICA 필터 동작의 엔트로피 최대화에 관련되지 않는다. 또한 ICA 필터 구조의 적응 성질은 필터 계수가 아주 빨리 조절되거나 또는 진동 동작을 나타낸다면, 분리된 출력 신호가 잔향 인공물을 포함하는 것을 의미한다. 따라서 학습된 필터 웨이트는 잔향 효과를 피하기 위해서 시간 및 주파수 도메인에서 평탄해져야 한다. 이 평탄화 동작이 필터 학습 처리의 속도를 떨어뜨리기 때문에, 이 강화된 음향 명료성 설계 측면은 전체 시스템 성능에 대해 추가적인 안정화 효과를 갖는다.

계산 자원 및 마이크로폰의 개수에서의 한계에 이를 수 있는 공간적으로 분포된 배경 노이즈의 블라인드 음원 분리의 성능을 증가시키기 위해서, ICA 계산된 입력 및 출력은 각각 전처리되거나 또는 후처리될 수 있다. 예컨대, 본 발명의 한 선택적인 실시예는 이들 방법이 상기 처리된 신호에 관한 시간 또는 스펙트럼 정보를 단독적으로 이용하고, 이같은 ICA 필터링 유닛을 보충할 수 있기 때문에, 음성 활동 감지 및 적응형 비너 필터링을 포함하려고 한다.

본 발명의 마지막 측면은 필터 피드백 구조의 계산 정밀도 및 능력 문제에 관한 것이다. 유한 비트 정밀도 계산 환경(통상 16비트 또는 32비트)에서, 필터링 동작은 계수 양자화 에러를 필터링하는 것이다. 이들은 일반적으로 약화된 수렴 성능 및 전체 시스템 안정성으로 귀결된다. 양자화 효과는 후처리된 ICA 출력이 ICA 필터 구조로 대신 피드백되게끔 크로스 필터 길이를 제한함으로써 및 본래의 피드백 구조를 변경함으로써 제어될 수 있다. 유한 정밀도 환경에서의 입력 에너지의 축소는 안정화 관점으로부터 뿐만 아니라 계산된 숫자값의 유한한 범위에도 필수적이다. 비록 유한한 정밀도 환경에서의 성능이 신뢰할 만하고 조절가능하지만, 제안된 음향 처리 스킴은 바람직하게는 부동 소수점(floating point) 정밀도 환경에서 실행되어야 한다. 최종적으로 계산적인 강제 하에서의 실행은 필터 길이의 적절한 선택 및 필터 계수의 업데이트 주파수를 튜닝함으로써 실행된다. 실제로 ICA 필터 구조의 계산적인 복잡성은 이러한 후자 변수의 직접적인 함수이다.

다른 측면 및 실시예가 도면에 도시되고 이하의 상세한 설명에서 설명되며, 또한 청구의 범위의 기술적 사상에 의해 정의될 것이다.

### 도면의 간단한 설명

도 1은 종래의 ICA 신호 분리 시스템의 블록도를 나타내고,

도 2는 본 발명에 따른 음향 분리 시스템의 일 실시예의 블록도이며,

- 도 3은 본 발명에 따른 개선된 ICA 처리 서브-모듈의 일 실시예의 블록도이고,  
 도 4는 본 발명에 따른 개선된 ICA 음향 분리 처리의 일 실시예의 블록도이며,  
 도 5는 본 발명에 따른 음향 처리 방법의 흐름도이고,  
 도 6은 본 발명에 따른 음향 디-노이징 처리의 흐름도이며,  
 도 7은 본 발명에 따른 음향 피쳐 추출 처리의 흐름도이고,  
 도 8은 본 발명에 따른 음향 처리의 결합의 예들을 나타내는 표이며,  
 도 9는 본 발명에 따른 음향 분리 시스템을 갖는 셀룰라 폰의 일 실시예의 블록도이고,  
 도 10은 본 발명에 따른 음향 분리 시스템을 갖는 셀룰라 폰의 다른 실시예의 블록도이다.

### 실시예

음향 분리 시스템의 바람직한 실시예가 도면을 참조하여 이하에서 설명된다. 제한된 계산 능력으로 실시간 처리를 하기 위해서, 시스템은 단순하고 계산하기 쉬운 유계 함수를 갖는 크로스 필터의 개선된 ICA 처리 서브-모듈을 사용한다. 통상적인 접근과 비교하면, 이러한 단순화된 ICA 방법은 계산 능력 요구를 감소시키고 비-음향 신호로부터 음향 신호를 성공적으로 분리한다.

### 음향 분리 시스템 개관

도 2는 음향 분리 시스템(200)의 일 실시예를 나타낸다. 시스템(200)은 음향 강화 모듈(210), 선택적인 음향 디-노이징 모듈(220), 및 선택적인 음향 피쳐 추출 모듈(230)을 포함한다. 음향 강화 모듈(210)은 개선된 ICA 처리 서브-모듈(212) 및 선택적인 후처리 서브-모듈(214)을 포함한다. 개선된 ICA 처리 서브-모듈(212)은 상대적으로 낮은 계산 능력으로 실시간 음향 분리를 달성하기 위해서 단순화되고 개선된 ICA 처리를 사용한다. 실시간 음향 분리가 필요하지 않는 적용에서, 개선된 ICA 처리는 계산 능력 상의 요구량을 추가로 줄일 수 있다. 여기서 사용된 것과 같이, 용어 ICA 및 BSS는 교환가능하고 근사법을 통해 직접적으로 또는 간접적으로 상호 정보의 수학적 공식을 최소화 또는 최대화하는 방법을 지칭하며, 시간 및 주파수 도메인에 기초한 시간 지연 무상관화(decorrelation)와 같은 무상관화 방법 또는 임의의 다른 2차 또는 그 이상의 통계에 기초한 무상관화 방법을 포함한다.

여기에서 사용된 것과 같이, "모듈" 또는 "서브-모듈"은 임의의 방법, 장치, 설비, 유닛 또는 소프트웨어, 하드웨어 또는 펌웨어 형태의 컴퓨터 명령을 포함하는 컴퓨터로 판독가능한 데이터 저장 매체를 지칭할 수 있다. 동일한 기능을 실행하기 위해서 다수의 모듈 또는 시스템은 하나의 모듈 또는 시스템으로 결합될 수 있고, 하나의 모듈 또는 시스템은 다수의 모듈 또는 시스템으로 분리될 수 있는 것으로 알려져 있다. 셀룰라 폰 적용에 관한 바람직한 실시예에서, 개선된 ICA 처리 서브-모듈(212)은 단독으로 또는 다른 모듈과 조합하여 셀룰라 폰 내에 위치한 마이크로프로세서 칩에 설치된다. 소프트웨어 또는 다른 컴퓨터로 실행가능한 명령에서 실행되는 경우, 본 발명의 구성요소들은 본질적으로 예를 들면, 루틴, 프로그램, 객체, 성분, 데이터 구조 등으로 필요한 작업을 실행하는 코드 세그먼트(code segment)이다. 프로그램 또는 코드 세그먼트는 프로세서가 판독가능한 매체에 저장될 수 있거나 또는 전송 매체 또는 통신 링크를 통해 반송파에 실려진 컴퓨터 데이터 신호에 의해 전송될 수 있다. "프로세서가 판독가능한 매체"는 정보를 저장할 수 있거나 전송할 수 있는 임의의 매체를 포함할 수 있고, 휘발성, 비휘발성, 삭제가능한 및 삭제 불가능한 매체를 포함한다. 프로세서가 판독가능한 매체의 예들은 전자 회로, 반도체 메모리 장치, ROM, 플래쉬 메모리, 삭제가능한 ROM(EROM), 플로피 디스켓 또는 다른 자기 저장장치, CD-ROM/DVD 또는 다른 광 저장장치, 하드 디스크, 광섬유 매체, RF 링크, 또는 원하는 정보를 저장하는데 사용될 수 있고 액세스될 수 있는 임의의 다른 매체를 포함한다. 컴퓨터 데이터 신호는 전자 네트워크 채널, 광섬유, 공기, 전자기, RF 링크 등과 같은 전송 매체를 통해 전파될 수 있는 임의의 신호를 포함할 수 있다. 코드 세그먼트는 인터넷, 인트라넷, 등과 같은 컴퓨터 네트워크를 통해 다운로드될 수 있다. 어떠한 경우에도, 본 발명은 이러한 실시예들에 의해 한정되는 것으로 해석되어서는 안된다.

음향 분리 시스템(200)은 하나 이상의 음향 강화 모듈(210), 음향 디-노이징 모듈(220), 및 음향 피쳐 추출 모듈(230)의 다양한 조합을 포함할 수 있다. 음향 분리 시스템(200)은 또한 이하에서 설명되는 하나 이상의 음향 인식 모듈(도시 안됨)

을 포함할 수 있다. 각각의 모듈은 독립 시스템 또는 큰 시스템의 일부로서 단독으로 사용될 수 있다. 이하에서 설명된 것과 같이, 음향 분리 시스템은 특정 기능을 제어하기 위해서 음향 입력을 수신하는 전자장치에 통합되는 것이 바람직하나, 그렇지 않으면 배경 노이즈로부터 원하는 노이즈의 분리를 필요로 한다. 많은 적용은 다수의 방향으로부터 발생하는 배경 사운드로부터 깨끗한 원하는 사운드를 강화 또는 분리하는 것을 필요로 한다. 이러한 적용은 음성 인식 및 감지, 음향 강화 및 분리, 음성-동작 제어 등과 같은 성능을 포함하는 전자 또는 컴퓨터 장치와 같은 휴먼-머신 인터페이스를 포함한다. 본 발명인 음향 분리 시스템에 의해 요구되는 감소된 처리 능력 때문에, 제한된 처리 능력 만을 제공하는 장치에 적절하다.

개선된 ICA 처리

도 3은 개선된 ICA 또는 BSS 처리 서브-모듈(212)의 일 실시예(300)를 나타낸다. 입력 신호  $X_1$  및  $X_2$ 는 각각 채널(310 및 320)로부터 수신된다. 통상, 이러한 신호의 각각은 적어도 하나의 마이크로폰으로부터 발생할 수 있으나, 다른 음원이 사용될 수 있는 것으로 이해될 수 있다. 크로스 필터  $W_1$  및  $W_2$ 는 분리된 신호(U1)의 채널(330) 및 분리된 신호(U2)의 채널(340)을 생성하기 위해서 입력 신호의 각각에 적용된다. 채널(330)(음향 채널)은 원하는 신호를 지배적으로 포함하고 채널(340)(노이즈 채널)은 노이즈 신호를 지배적으로 포함한다. 비록 용어 "음향 신호" 및 "노이즈 채널"이 사용되었지만, 용어 "음향" 및 "노이즈"는 바람직한 것으로 대체될 수 있으며, 예컨대 하나의 음향 및/또는 노이즈가 다른 음향들 및/또는 노이즈들에 대해 바람직한 것일 수 있다. 추가로, 이 방법은 또한 두개 이상의 음원으로부터 혼합 노이즈 신호를 분리하는데 사용될 수 있다.

무한 임펄스 응답 필터는 바람직하게는 개선된 ICA 처리에 사용된다. 무한 임펄스 응답 필터는 그 출력 신호가 입력 신호의 적어도 일부로서 상기 필터로 피드백되는 필터이다. 유한 임펄스 응답 필터는 그 출력 신호가 입력으로서 피드백되지 않는 필터이다. 크로스 필터  $W_{21}$  및  $W_{12}$ 는 긴 주기의 시간 지연을 캡처하기 위해서 시간에 대해 희박하게 분포된 계수를 가질 수 있다. 가장 단순화된 형태에서, 크로스 필터  $W_{21}$  및  $W_{12}$ 는 필터당 단지 하나의 필터 계수를 갖는 이득 인자이며, 예컨대 출력 신호와 피드백 입력 신호 사이의 시간 지연에 대한 지연 이득 인자 및 입력 신호를 증폭하는 진폭 이득 인자이다. 다른 형태에서, 크로스 필터는 각각 수십, 수백 또는 수천개의 필터 계수를 가질 수 있다. 이하에서 설명되는 것과 같이, 출력 신호  $U_1$  및  $U_2$ 는 후처리 서브-모듈, 디-노이징 모듈 또는 음향 피쳐 추출 모듈에 의해 추가로 처리될 수 있다.

비록 ICA 학습 규칙이 블라인드 음원 분리를 달성하기 위해 명백하게 유도되었지만, 어쿠스틱 환경에서 음향 처리에 대한 실질적인 실행은 필터링 스킴의 불안정한 행동을 야기할 수 있다. 이 시스템의 안정성을 보장하기 위해서,  $W_{12}$  및 유사한  $W_{21}$ 의 적용 역학은 우선 안정적이어야 한다. 이와 같은 시스템에 대한 이득 마진은 일반적인 의미에서 낮으므로, 변동하는 음성 신호에서와 같은 입력 이득의 증가는 불안정을 야기하고 따라서 웨이트 계수가 급격히 증가한다. 음향 신호는 일반적으로 제로 평균을 갖는 희박한 분포를 나타내기 때문에, 부호함수는 시간에서 빈번히 진동하고 불안정한 행동에 기여하게 될 것이다. 최종적으로 큰 학습 파라미터는 빠른 수렴을 위해 바람직하고, 큰 입력 이득이 시스템을 더욱 불안정하게 만들기 때문에 안정성과 성능 사이에 상충이 내재되어 있다. 알려진 학습 규칙은 불안정성으로 이끌 뿐만 아니라, 비선형 부호 함수 때문에 진동하는 경향이 있고, 특히 안정성 한계에 접근하는 경우 필터링된 출력 신호  $Y_1[t]$  및  $Y_2[t]$ 의 잔향을 야기한다. 이러한 문제를 역점을 두어 다루기 위해,  $W_{12}$  및  $W_{21}$ 에 대한 적용 규칙은 안정화되는 것을 필요로 한다. 필터 계수에 대한 학습 규칙이 안정적이라면, 광범위한 분석적이고 경험적인 연구는 시스템이 BIBO(Bounded Input Bounded Output)에서 안정적임을 보여주었다. 따라서 전체 처리 스킴의 최종에 해당하는 목적은 안정성 강제 하에서 잡음이 있는 음향 신호의 블라인드 음원 분리일 것이다.

그러므로 안정성을 보장하기 위한 주요한 방법은 도 3에 도시된 것과 같이 입력을 적절하게 스케일링 하는 것이다. 이러한 기초 구조에서 스케일링 인자  $sc\_fact$ 는 진입하는 입력 신호 특징에 기초하여 변경된다. 예컨대, 입력이 너무 높다면, 이는  $sc\_fact$ 에서 증가를 야기할 것이고, 따라서 입력 진폭을 감소시킬 것이다. 성능과 안정성 사이에는 타협이 존재한다.  $sc\_fact$ 에 의한 입력의 축소는 감소된 분리 성능을 야기하는 SNR을 감소시킨다. 이와 같이 입력은 안정성을 보장하기 위해 필요한 정도로 스케일링되어야 한다. 모든 샘플에서 웨이트 계수의 단기의 변화를 처리하는 필터 구조를 동작시키고, 이에 의해 연관된 잔향을 피함으로써 크로스 필터에 대한 추가적인 안정화가 얻어질 수 있다. 이 적용 규칙 필터는 시간 도메인의 평탄화로서 보여질 수 있다. 추가적인 필터 평탄화가 이웃하는 주파수 저장소로 수렴된 분리 필터의 코히런스(coherence)를 강화하기 위해서 주파수 도메인에서 실행될 수 있다. 이는 일반적으로 K-탭 필터를 길이 L로 제로 탭핑하고, 그다음 역변환이 따르는 증가된 시간 지원으로 이 필터를 푸리에 변환함으로써 행해질 수 있다. 필터는 사각형 시간 도메인 창으로 효과적으로 창이 만들어지기 때문에, 주파수 도메인에서 싱크(sinc) 함수에 의해 상응하게 평탄화된다. 이 주파수 도메인 평탄화는 코히런트 솔루션에 적용된 필터 계수를 주기적으로 재초기화하여 일정한 시간 간격으로 실행될 수 있다.

다음의 수학적식은 크기 t인 각 시간 샘플창에 사용될 수 있고 시간 변수인 k를 갖는 비선형 유계 함수의 예이다.

$$U_1(t) = X_1(t) + W_{12}(t) \otimes X_2(t)$$

$$U_2(t) = X_2(t) + W_{21}(t) \otimes X_1(t)$$

$$Y_1 = \text{sign}(U_1)$$

$$Y_2 = \text{sign}(U_2)$$

$$\Delta W_{12k} = -f(Y_1) \times U_2(t-k)$$

$$\Delta W_{21k} = -f(Y_2) \times U_1(t-k)$$

함수 f(x)는 비선형 유계 함수, 즉 소정의 최대값 및 소정의 최소값을 갖는 비선형 함수이다. 바람직하게는, f(x)는 변수 x의 부호에 의존하여 최대값 또는 최소값에 빠르게 접근하는 비선형 유계 함수이다. 예컨대, 상기 수학적식 3 및 수학적식 4는 단순한 유계 함수로서 부호 함수를 사용한다. 부호 함수 f(x)는 x가 양인지 음인지 여부에 따라서 1 또는 -1의 2진 값을 갖는 함수이다. 예가되는 비선형 유계 함수는 이하의 수학적식을 포함하나, 이에 한정되지 않는다.

$$f(x) = \text{sign}(x) = \begin{cases} 1 & | \ x > 0 \\ -1 & | \ x \leq 0 \end{cases}$$

$$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

$$f(x) = \text{simple}(x) = \begin{cases} 1 & | \ x \geq \epsilon \\ x/\epsilon & | \ -\epsilon > x > \epsilon \\ -1 & | \ x \leq -\epsilon \end{cases}$$

이러한 규칙은 필요한 계산을 수행하기 위해 부동 소수점 정밀도가 이용 가능한 것을 가정한다. 비록 부동 소수점 정밀도가 바람직하지만, 더욱 구체적으로는 최소화된 계산 처리 능력을 갖는 장치에 적용되기 때문에 고정점 계산이 또한 사용될 수 있다. 고정점 계산을 사용할 수 있음에도 불구하고, 최적의 ICA 해법으로의 수렴은 더욱 어렵다. 실제로 ICA 알고리즘은 간섭하는 음원이 삭제되어야 하는 원칙에 기초한다. 거의 동일한 숫자가 감해지는(또는 매우 다른 숫자가 더해지는) 상황에서 고정점 계산의 어떤 부정확도 때문에, ICA 알고리즘은 최적 이하의 수렴 성질을 나타낼 수 있다.

분리 성능에 영향을 끼칠 수 있는 다른 인자는 필터 계수 양자화 에러 효과이다. 제한된 필터 계수 해상도 때문에, 필터 계수의 적응은 특정점에서 점차적인 부가 분리 개선 및 이에 따른 수렴 성질의 판정에 있어서 고려를 양산할 것이다. 양자화 에러 효과는 다수의 인자에 의존하나 주로 사용된 필터 길이 및 비트 해상도의 함수이다. 이전에 언급된 입력 스케일링 문제는 숫자 상의 오버플로를 방지하는 유한 정밀도 계산에서 또한 필수적이다. 필터링 처리에 관련된 컨볼루션(convolution)은 이용 가능한 해상도 범위 보다 큰 숫자까지 잠재적으로 더해질 수 있기 때문에, 스케일링 인자는 필터 입력이 이러한 형태가 발생하는 것을 방지하기 위해서 충분히 작아야 하는 것을 보장해야 한다.

#### 다중-채널 개선된 ICA 처리

개선된 ICA 처리 서브-모듈(212)은 마이크로폰과 같은 적어도 두개의 오디오 입력 채널로부터 입력 신호를 수신한다. 오디오 입력 채널의 개수는 최소 개수인 두 채널을 넘어 증가될 수 있다. 입력 채널의 개수가 증가함에 따라서, 음향 분리 품질은 일반적으로 입력 채널의 개수가 오디오 신호 음원의 개수와 같아지는 정도까지 개선될 수 있다. 예컨대, 입력 오디오

신호의 음원이 화자, 배경 화자, 배경 음원, 및 멀리 떨어진 도로의 노이즈 및 바람의 노이즈에 의해 생성된 일반적인 배경 노이즈를 포함한다면, 4-채널 음향 분리 시스템은 보통 2-채널 시스템보다 성능이 우수하다. 물론, 더 많은 입력 채널이 사용될 때, 더 많은 필터 및 더 많은 계산 능력이 요구된다.

개선된 ICA 처리 서브-모듈 및 처리는 입력 신호의 두개 채널 보다 많은 것을 분리하기 위해 사용될 수 있다. 예컨대, 셀룰라폰 적용에 있어서, 일 채널은 실질적으로 원하는 음향 신호를 포함할 수 있고, 다른 채널은 실질적으로 하나의 노이즈 음원으로부터의 노이즈 신호를 포함할 수 있으며, 및 다른 채널은 실질적으로 다른 노이즈 음원으로부터의 오디오 신호를 포함할 수 있다. 예컨대, 다수의 사용자 환경에서, 일 채널은 한 타겟 사용자의 지배적인 음향을 포함할 수 있으며, 반면에 다른 채널은 다른 타겟 사용자의 지배적인 음향을 포함할 수 있다. 세번째 채널은 노이즈를 포함할 수 있고, 두개의 음향 채널의 추가 처리에 유용할 수 있다. 추가적인 음향 또는 타겟 채널은 유용할 수 있는 것으로 여겨질 것이다.

비록 일부의 적용이 원하는 음향 신호의 하나의 음원에만 관련되지만, 다른 적용에서 원하는 음향 신호에 대해 다수의 음원이 존재할 수 있다. 예컨대, 텔레콘퍼런스 적용 또는 오디오 감시 적용은 배경 노이즈로부터 및 서로로부터 다수의 화자의 음향 신호를 분리하는 것을 필요로 할 수 있다. 개선된 ICA 처리는 배경 노이즈로부터 음향 신호의 한 음원을 분리할 뿐만 아니라 한 화자의 음향 신호를 다른 화자의 음향 신호로부터 분리하는데 사용될 수 있다.

### 주변 처리

본 발명의 방법 및 시스템의 성능을 효과적이고 강하게 증가시키기 위해서, 다양한 주변 처리 기술이 입력 및 출력 신호에 다양한 등급으로 적용될 수 있다. 여기서 명백하게 설명된 방법 및 시스템을 보충하는 후처리 기술 뿐만 아니라 전처리 기술은 오디오 혼합에 적용된 블라인드 음원 분리 기술의 성능을 강화할 것이다. 예컨대, 후처리 기술은 바람직하지 않은 출력 또는 분리되지 않은 입력을 이용하여 원하는 신호의 품질을 개선하는데 사용될 수 있다. 유사하게, 전처리 기술 또는 정보는 여기에 설명된 방법 및 시스템을 보충하기 위해 믹싱 시나리오의 조건을 개선함으로써 오디오 혼합에 적용된 블라인드 음원 분리 기술의 실행을 강화할 수 있다.

개선된 ICA 처리는 사운드 신호를 적어도 두개의 채널, 예컨대 노이즈 신호에 대한 한 채널(노이즈 채널) 및 원하는 음향 신호에 대한 한 채널(음향 채널)로 분리한다. 도 4에 도시된 바와 같이, 채널(430)은 음향 채널이고 채널(440)은 노이즈 채널이다. 음향 채널이 바람직하지 않은 수준의 노이즈 신호를 포함하고 노이즈 채널이 여전히 약간의 음향 신호를 포함하는 것은 꽤 가능성이 있다. 예컨대, 두개 이상의 의미있는 음원 및 단지 두개의 마이크로폰이 존재하거나, 또는 두개의 마이크로폰이 음원이 멀리 떨어져 위치한 것을 제외하고 함께 가까이 위치한다면, 개선된 ICA 처리는 단독으로 노이즈로부터 원하는 음향을 항상 충분히 분리할 수 없을 것이다. 그러므로 처리된 신호는 배경 노이즈의 잔여 수준을 제거하고 및/또는 음향 신호의 품질을 추가로 개선하기 위해서 후처리되는 것이 필요할 수 있다. 이는 예컨대 단일 또는 다중 채널 음향 강화 알고리즘을 통해 분리된 ICA 출력을 제공함으로써 얻어진다. 음성 활동 검출기로 검출된 비음향 시간 간격으로부터 산정된 노이즈 스펙트럼을 갖는 비너 필터(Wiener filter)는 긴 시간 지원으로 배경 노이즈에 의해 약해진 신호에 대한 좋은 SNR을 얻는데 사용된다. 또한, 유계 함수는 결합 엔트로피 계산에 대한 단지 단순화된 근사법이고, 신호의 정보 중복을 항상 완전히 줄일 수 없다. 그러므로, 신호가 개선된 ICA 처리를 사용하여 분리된 이후에, 후처리는 음향 신호의 품질을 추가로 개선하기 위해 실행될 수 있다.

분리된 노이즈 신호 채널은 버려질 수 있으나, 다른 목적을 위해 사용될 수 있다. 음향 채널의 잔여 노이즈 신호가 노이즈 채널의 노이즈 신호로서 유사한 신호 부호를 갖는다는 합리적인 가정에 기초하여, 부호가 노이즈 채널 신호의 부호와 유사한 원하는 음향 채널에서의 신호는 후처리 유닛에서 필터링되어야 한다. 예컨대, 스펙트럼 차감 기술은 후처리를 실행하기 위해서 사용될 수 있다. 노이즈 채널에서 신호의 부호가 식별된다. 노이즈 특징에 대한 소정의 가정에 의존하는 종래의 노이즈 필터와 비교하여, 후처리는 특별한 환경의 노이즈 부호를 분석하고 특별한 환경을 나타내는 노이즈 신호를 제거하기 때문에 더욱 융통성이 있다. 그러므로 이는 노이즈 제거에서 지나치게 포괄적이거나 또는 덜 포괄적인 것이 더 작게 가능할 것이다. 비너 필터링 및 칼만(Kalman) 필터링과 같은 다른 필터링 기술은 또한 후처리를 실행하는데 사용될 수 있다. ICA 필터 해법은 실재 해법의 한계 사이클로 단지 수렴될 것이기 때문에, 필터 계수는 좋은 분리 성능으로 귀결되는 것을 제외하고 계속해서 적용될 것이다. 약간의 계수는 그 해상도 한계로 이동되었다. 그러므로 원하는 화자 신호를 포함하는 ICA 출력의 후처리된 버전은 수렴 한계 사이클이 극복되고 ICA 알고리즘을 동요시키지 않게끔 도 4에 의해 도시된 IIR 피드백 구조를 통해 피드백된다. 이 과정의 유익한 부산물은 수렴이 상당히 가속된다는 것이다.

디-노이즈, 음향 피쳐 추출과 같은 다른 처리들은 음향 신호의 품질을 추가로 개선하기 위해서 음향 강화와 함께 사용될 수 있다. 음향 인식 적용은 음향 강화 처리에 의해 분리된 음향 신호를 이용할 수 있다. 노이즈로부터 실질적으로 분리된 음향 신호와 함께, Hidden Markov Model chains과 같은 방법에 기초한 음향 인식 엔진, 신경망 학습 및 지원 벡터 머신은 더 큰 정확성을 가지고 작동할 수 있다.

지금부터 도 5를 참조하여, 음향 처리의 흐름도를 설명한다. 방법(500)은 예컨대 휴대용 무선 모바일 폰, 전화 헤드셋과 같은 음향 장치, 또는 핸즈-프리 자동차 키트에서 사용될 수 있다. 방법(500)은 다른 음향 장치에서 사용될 수 있고, DSP 프로세서, 일반적인 계산 프로세서, 마이크로프로세서, 게이트 어레이, 또는 다른 계산 장치에서 실행될 수 있는 것으로 여겨진다. 사용에서, 방법(500)은 사운드 신호(502)의 형태로 어쿠스틱 신호를 수신한다. 이러한 사운드 신호(502)는 많은 음원으로부터 발생할 수 있고, 타겟 사용자로부터의 음향, 근처의 다른 사람으로부터의 음향, 노이즈, 잔향, 에코, 반사, 및 다른 바람직하지 않은 사운드를 포함할 수 있다. 비록 단일 타겟 음향 신호를 식별하고 분리하는 방법(500)이 도시되었지만, 방법(500)은 추가적인 타겟 사운드 신호를 식별하고 분리하기 위해 변형될 수 있다.

또한, 다양한 전처리 기술 또는 정보가 사전 지식을 사용하고, 입력 신호 및 조건에서 발산하는 정보 또는 특징을 최대화하며, 믹싱 시나리오의 조건을 개선하는 등과 같은 혼합 오디오 신호의 처리 및 분리를 개선하거나 또는 촉진하는데 사용될 수 있다. 예컨대, 분리된 ICA 사운드 채널의 출력 순서는 일반적으로 미리 알려지지 않으므로, 추가 채널 선택 단계(510)는 반복하는 방식으로 원하는 화자에 관한 사전 지식(501)에 기초하여 분리된 채널의 콘텐츠를 처리한다. 원하는 화자의 음향의 특징을 식별하기 위해 사용된 표준(504)은 공간 또는 시간 피쳐, 에너지, 볼륨, 주파수 콘텐츠, 제로 크로싱 비율 또는 분리 처리에 대해 병렬로 계산된 화자 의존 및 독립 음향 인식 스코어에 기초할 수 있으나 이에 한정되지 않는다. 예컨대, 표준(504)은 특별한 명령 예컨대 "일어나시오(wake up)"와 같은 강제된 어휘에 응답하기 위해 구성될 수 있다. 다른 예에서, 음향 장치는 자동차의 전방 운전자 위치와 같은 특별한 위치 또는 방향으로부터 방출하는 사운드 신호에 대해 응답할 수 있다. 이러한 방식에서 핸즈-프리 자동차 키트는 승객 및 라디오로부터의 음향을 무시하면서 운전자로부터의 음향에만 응답하기 위해 구성될 수 있다. 선택적으로, 믹싱 시나리오의 조건은 예컨대 공간, 시간, 에너지, 스펙트럼 등의 변형 및 조작에 의한 것과 같이 입력 신호의 특징을 변형 또는 조작하는 것에 의해 개선될 수 있다.

일부 음향 장치에서, 마이크로폰은 음원, 배경 노이즈로부터의 소정의 거리에 기초하거나 또는 다른 마이크로폰과 관련되어 일관성 있게 배치되거나, 또는 예컨대 지향성 마이크로폰과 같은 입력 신호를 조절하는 특정한 특징을 갖는다. 블록(506)에 나타난 바와 같이, 두개의 마이크로폰은 공간적으로 분리될 수 있고 음향 장치의 하우징에 배치될 수 있다. 예컨대, 전화 헤드셋은 통상적으로 마이크로폰이 화자의 입에서 대략 1 인치 내에 있어서, 화자의 음성이 일반적으로 마이크로폰에 가장 가까운 음원이 되게끔 조절된다. 유사한 방식으로, 핸드헬드 무선 폰, 헤드셋, 또는 접히는 마이크로폰은 일반적으로 타겟 화자의 입까지 상당히 알려진 거리를 갖는다. 마이크로폰으로부터 타겟 음원까지의 거리는 알려지기 때문에, 이 거리는 타겟 음향 신호를 식별하기 위한 특징으로 사용될 수 있다. 예컨대, 처리(510)는 2인치 이하로 떨어진 것으로부터 발생하고 남성의 음성을 표시하는 주파수 성분을 갖는 사운드 신호만을 선택할 수 있다. 두개의 마이크로폰 셋업이 사용되는 경우에, 마이크로폰은 원하는 화자의 입에 가깝게 배열된다. 이러한 셋업은 원하는 화자의 음성 신호를 하나의 분리된 ICA 채널로 분리되는 것을 허용하여 단지 노이즈만을 포함하는 나머지 분리된 출력 채널은 원하는 화자의 채널의 연속적인 후처리를 위한 노이즈 레퍼런스로서 사용될 수 있게끔 한다.

두개 이상의 마이크로폰이 사용되는 기록 시나리오에서, 두 채널 ICA 알고리즘은 두개의 채널 시나리오에 대해 일찌기 설명된 것과 같은 유사한 방식으로,  $N*(N-1)$  ICA 크로스 필터를 갖는  $N$ -채널(마이크로폰) 알고리즘으로 확장된다. 후자는  $N$ 개의 기록 채널 중에서 선택하기 위해서 [ad2]에 나타난 채널 선택 과정을 따라 음원 할당 목적으로 사용되며, 최적의 두개의 채널 조합은 두개의 채널 ICA 알고리즘에서 그다음 처리되어 원하는 화자를 분리한다.  $N$ -채널 ICA 분리로부터 생기는 모든 종류의 정보 음원은 유사하지만 한정되지 않으며, 학습된 ICA 크로스 필터 계수 뿐만 아니라 기록된 입력으로부터 분리된 출력 음원까지의 상대적인 에너지 변화는 이 목적에 이용된다.

공간적으로 분리된 마이크로폰 각각은 원하는 타겟 사운드와 약간의 노이즈 및 잔향 음원의 혼합물인 신호를 수신한다. 혼합 사운드 신호(507 및 509)는 분리를 위해 ISA 처리(508)에서 수신된다. 식별 처리(510)를 사용하여 타겟 음향 신호를 식별한 이후에, ICA 처리(508)는 혼합 사운드를 원하는 음향 신호 및 노이즈 신호로 분리한다. ICA 처리는 예컨대 웨이트 인자를 추가로 다듬고 설정하기 위해 노이즈 신호를 사용하는 것과 같이, 음향 신호를 추가로 처리(512)하기 위해 노이즈 신호를 사용할 수 있다. 또한 이하에서 추가로 설명되는 것과 같이, 노이즈 신호는 또한 음향 신호로부터 노이즈 콘텐츠를 추가로 제거하는 추가 필터링(514) 또는 처리에 의해 사용될 수 있다.

## 디-노이징

도 6은 디-노이징 처리의 일 실시예를 나타내는 흐름도이다. 셀룰라 폰 적용에서, 디-노이징은 모든 방향에서 생기는 바람 노이즈와 같이 공간적으로 할당되지 않는 노이즈 음원을 분리하는데 최적이다. 디-노이징 기술은 또한 고정된 주파수로 노이즈 신호를 제거하는데 사용될 수 있다. 처리는 시작 블록(600)에서 블록(610)으로 진행된다. 블록(610)에서, 처리는 음향 신호  $x$ 의 블록을 수신한다. 처리는 시스템이 음원 계수를 다음의 수학적식을 사용하여 바람직하게 계산하는 블록(620)으로 진행된다.

$$S_i = \sum_j w_{ij} * x_j$$

상기 수학적식에서,  $w_{ij}$  는 ICA 웨이트 행렬을 나타낸다. 미국 특허 제5,706,402호에 설명된 ICA 방법 또는 미국 특허 제 6,424,960호에서 설명된 ICA 방법은 디-노이징 처리에서 사용될 수 있다. 그다음 처리는 블록(630), 블록(640), 또는 블록(650)으로 진행된다. 블록(630, 640 및 650)은 선택적인 실시예를 나타낸다. 블록(630)에서, 처리는 신호  $s_i$ 의 전력에 기초하여 다수의 의미있는 음원 계수를 선택한다. 블록(640)에서, 중요치 않은 계수를 제거하기 위해서 최대 가능성 있는 축소 함수를 계산된 음원 계수에 적용한다. 블록(650)에서, 처리는 각 시간 샘플  $t$ 에 대한 기초 기능 중의 하나로 음향 신호  $x$ 를 필터링한다.

처리는 블록(630, 640, 또는 650)으로부터 다음의 수학적식을 사용하여 바람직하게 음향 신호를 재구성하는 블록(660)으로 진행된다.

$$x_{new} = \sum_j a_{ij} * s_{j,shrinked}$$

상기 수학적식에서,  $a_{ij}$  는 웨이트 인자로 진입 신호를 필터링함으로써 생성된 연습중인 신호를 나타낸다. 디-노이징 처리는 이같이 노이즈를 제거하고 재구성된 음향 신호  $x_{new}$ 를 생성한다. 양호한 디-노이징 결과는 노이즈 음원에 관한 정보가 이용가능한 경우에 얻어진다. 개선된 ICA 처리에 관하여 상기에서 설명된 것과 같이, 노이즈 채널에서의 신호의 부호는 음향 채널의 신호로부터 노이즈를 제거하는 디-노이징 처리에 의해 사용될 수 있다. 처리는 블록(660)으로부터 종료 블록(670)으로 진행된다.

#### 음향 피쳐 추출

도 7은 ICA를 사용하는 음향 피쳐 추출 처리의 일 실시예를 나타낸다. 처리는 시작 블록(700)에서부터 음향 신호  $x$ 를 수신하는 블록(710)으로 시작한다. 도 9를 참조하여 이하에서 설명된 것과 같이, 음향 신호  $x$ 는 입력 음향 신호, 음향 강화에 의해 처리된 신호, 디-노이징에 의해 처리된 신호, 또는 음향 강화 및 디-노이징에 의해 처리된 신호가 될 수 있다.

도 7로 돌아가 참조하면, 처리는 블록(710)으로부터 수학적식 10에 의해 상기에서 설명된 것과 같이 공식  $s_{ij,new} = W * x_{ij}$ 을 참조하여 음원 계수를 계산하는 블록(720)으로 진행된다. 그다음 처리는 수신된 음향 신호가 기초 함수로 분해되는 블록(730)으로 진행된다. 처리는 블록(730)으로부터 계산된 음원 계수가 피쳐 벡터로서 사용되는 블록(740)으로 진행된다. 예컨대, 계산된 계수  $s_{ij,new}$  또는  $2 \log s_{ij,new}$ 가 계산 피쳐 벡터에서 사용된다. 그다음 처리는 종료 블록(750)으로 진행된다.

추출된 음향 피쳐는 음향을 인식하거나 또는 다른 오디오 신호로부터 인식가능한 음향을 구분하기 위해 사용될 수 있다. 추출된 음향 피쳐는 홀로 사용될 수 있거나 또는 셉스트럴(cepstral) 피쳐(MFCC)와 결합하여 사용될 수 있다. 추출된 음향 피쳐는 또한 예컨대 다수의 화자의 음향 신호로부터 개별 화자를 식별하기 위해서, 또는 남성 또는 여성 화자로부터의 음향과 같은 특정 분류에 속하는 음향 신호를 식별하기 위해서와 같이 화자를 식별하기 위해 사용될 수 있다. 추출된 음향 피쳐는 또한 음향 신호를 검출하는 분류 알고리즘에 의해 사용될 수 있다. 예컨대, 최대 가능성 있는 계산은 의문의 신호가 인간의 음향 신호인 가능성을 판정하는데 사용될 수 있다.

추출된 음향 피쳐는 또한 텍스트의 컴퓨터 음독을 생성하는 텍스트 투 스피치(text-to-speech)에 적용될 수 있다. 텍스트 투 스피치 시스템은 음향 신호의 방대한 데이터베이스를 사용한다. 하나의 도전은 음소(phoneme)의 양호한 대표적인 데이터 베이스를 얻는 것이다. 종래의 시스템은 음향 데이터를 음소 데이터 베이스로 분류하기 위해서 셉스트럴 피쳐를 사용한다. 음향 신호를 기초 함수로 분해함으로써, 개선된 음향 피쳐 추출 방법은 음향을 음소 세그먼트로 더 잘 분류할 수 있고 그러므로 더 좋은 데이터 베이스를 생성할 수 있으며, 따라서 텍스트 투 스피치 시스템에 대해 더 좋은 음향 품질을 허용할 수 있다.

음향 피쳐 추출 처리의 일 실시예에서, 일 세트의 기초 함수는 음향을 인식하기 위해 모든 음향 신호에 대해 사용된다. 다른 실시예에서, 일 세트의 기초 함수는 각 화자를 인식하기 위해 각 화자에 대해 사용된다. 이는 텔레콘퍼런스로서 복수-화자 적용에 특히 유리할 수 있다. 또한 다른 실시예에서, 일 세트의 기초 함수는 각 분류를 인식하기 위해 일 분류의 화자

에 대해 사용된다. 예컨대, 일 세트의 기초 함수는 남자 화자에 대해 사용되고 다른 세트는 여자 화자에 대해 사용된다. 미국 특허 제6,424,960호는 다른 분류의 음성을 인식하기 위해 ICA 혼합 모델을 사용하여 설명한다. 이러한 모델은 다른 화자 또는 다른 성별의 화자의 음향 신호를 인식하기 위해 사용될 수 있다.

음향 인식

음향 인식 적용은 개선된 ICA 처리에 의해 분리된 음향 신호를 이용할 수 있다. 음향 신호가 실질적으로 노이즈로부터 분리되면, 음향 인식 적용은 더욱 큰 정확도로 동작할 수 있다. 이러한 Hidden Markov Model, 신경망 학습 및 지원 벡터 머신과 같은 방법이 음향 인식 적용에 사용될 수 있다. 상기한 바와 같이, 두개의-마이크로폰 배열에서, 개선된 ICA 처리는 입력 신호를 원하는 음향 신호 및 약간의 노이즈 신호의 일 음향 채널, 및 노이즈 신호 및 약간의 음향 신호의 일 노이즈 채널로 분리한다.

노이즈 환경에서 음향 인식 정확도를 개선하기 위해서, 노이즈 레퍼런스 신호에 기초하여 음향 신호로부터 노이즈를 제거하기 위해서 정확한 노이즈 레퍼런스 신호를 갖는 것이 바람직하다. 예컨대, 음향 스펙트럼 차감법을 사용하여 실질적인 일 채널의 음향 신호로부터 노이즈 레퍼런스 신호의 특징을 갖는 신호를 분리한다. 그러므로, 큰 노이즈 환경에 대한 바람직한 음향 인식 시스템에서, 시스템은 신호의 음향 채널 및 노이즈 채널을 수신하고 노이즈 레퍼런스 신호를 식별한다.

처리 조합

음향 피쳐 추출의 특정 실시예에서, 디-노이징 및 음향 인식 프로세스는 음향 강화 처리에 따라 설명되었다. 모든 처리가 함께 사용되는 것을 필요로 하지 않는다는 것을 주목할 필요가 있다. 표 800의 좌측 열은 신호의 형태가 기재되고 우측 열은 신호의 해당 형태를 처리하는 바람직한 처리가 기재된다.

줄 810에 나타난 일 배열에서, 입력 신호는 음향 강화를 사용하여 먼저 처리되고, 그다음 음향 디-노이징을 사용하여 처리되며, 그다음 음향 피쳐 추출을 사용하여 처리된다. 이러한 세개의 처리의 조합은 입력 신호가 큰 노이즈 및 경쟁 음원을 포함하는 경우 잘 동작한다. 큰 노이즈는 예컨대, 다양한 형태의 노이즈가 다른 방향으로부터 발생하나 일 형태의 노이즈는 특히 크지 않은 길거리 위의 복수의 음원으로부터 발생한 상대적으로 낮은 진폭의 신호를 참조한다. 경쟁 음원은 예컨대 운전자가 카 폰으로 대화하는 경우 높은 볼륨으로 튜닝된 자동차 라디오와 같이 원하는 음향 신호와 경쟁하는 하나 또는 몇몇의 음원으로부터 높은 진폭의 신호를 지칭한다. 줄 820에 나타난 다른 배열에서, 입력 신호는 음향 강화를 사용하여 먼저 처리되고 그다음 음향 피쳐 추출을 사용하여 처리된다. 음향 디-노이징 처리는 생략된다. 음향 강화 및 음향 피쳐 추출 처리의 조합은 본래의 신호가 경쟁 음원을 포함하고 큰 노이즈를 포함하지 않는 경우에 잘 동작한다.

줄 830에 나타난 또 다른 배열에서, 입력 신호는 음향 디-노이징을 사용하여 먼저 처리되고 그다음 음향 피쳐 추출을 사용하여 처리된다. 음향 강화 처리는 생략된다. 음향 디-노이징 및 음향 피쳐 추출 처리의 조합은 입력 신호가 큰 노이즈를 포함하고 경쟁 음원을 포함하지 않는 경우 잘 동작한다. 이 처리는 특정 큰 노이즈 또는 경쟁 음원을 포함하지 않는 상대적으로 깨끗한 음향에 대한 좋은 결과에 도달하기에 충분하다. 물론, 표 800은 단지 예들의 일 리스트이고 다른 실시예들이 사용될 수 있다. 예컨대, 음향 강화, 음향 디-노이징 및 음향 피쳐 추출 처리의 모두는 그들 형태와 상관없이 신호를 처리하는데 적용될 수 있다.

셀룰라 폰 적용

도 9는 셀룰라 폰 장치의 일 실시예를 도시한다. 셀룰라 폰 장치(900)는 사운드 신호를 기록하는 두개의 마이크로폰(910 및 920), 및 배경 노이즈로부터 원하는 음향 신호를 분리하기 위해 상기 기록된 신호를 처리하는 음향 분리 시스템(200)을 포함한다. 음향 분리 시스템(200)은 채널(930 및 940) 상으로 분리된 신호를 생성하기 위해 크로스 필터를 상기 기록된 신호에 적용하는 개선된 ICA 처리 서브-모듈을 적어도 포함한다. 그다음 분리된 원하는 음향 신호는 전송기(950)에 의해 유선 전화 또는 다른 셀룰라 폰과 같은 오디오 신호 수신 장치로 전송된다.

분리된 노이즈 신호는 버려질 수 있으나 다른 목적을 위해 또한 사용될 수 있다. 분리된 노이즈 신호는 환경 특징을 판정하고 그에 맞게 셀룰라 폰 파라미터를 조절하기 위해 사용될 수 있다. 예컨대, 노이즈 신호는 화자의 환경의 노이즈 레벨을 판정하기 위해 사용될 수 있다. 그다음 셀룰라 폰은 화자가 높은 노이즈 레벨을 갖는 환경에 있는 경우 마이크로폰의 볼륨을 증가시킨다. 상기한 바와 같이, 노이즈 신호는 또한 분리된 음향 신호로부터 잔여 노이즈를 추가로 제거하기 위한 레퍼런스 신호로서 사용될 수 있다.

쉬운 도시를 위해, 배터리, 디스플레이 패널 등과 같은 다른 셀룰라 폰 부속은 도 9에서 생략되었다. 아날로그-디지털 변환, 변조와 관련되거나 또는 주파수분할다중접속(FDMA : Frequency Division Multiple Access), 시간분할다중접속(TDMA : Time Division Multiple Access) 또는 채널분할다중접속(CDMA : Channel Division Multiple Access) 등을 가능케 하는 셀룰라 폰 신호 처리 단계는 또한 쉬운 도시를 위해 생략되었다.

비록 도 9가 두개의 마이크로폰을 도시하였지만, 두개 이상의 마이크로폰이 사용될 수 있다. 현존하는 제조 기술은 대략 10센트 은화, 핀의 머리 또는 그보다 작은 크기인 마이크로폰을 생산할 수 있으며, 다수의 마이크로폰이 장치(900) 상에 놓여질 수 있다.

일 실시예에서, 셀룰라 폰에서 실행되는 일반적인 에코-제거 처리는 개선된 ICA 서브-모듈에 의해 실행되는 처리와 같은 ICA 처리로 대체될 수 있다.

오디오 신호 음원은 일반적으로 서로 떨어져 있기 때문에, 마이크로폰은 바람직하게는 셀룰라 폰 상에서 어쿠스틱적으로 떨어져서 배치된다. 예컨대, 하나의 마이크로폰은 셀룰라 폰의 전방면 상에 배치될 수 있고 반면 다른 마이크로폰은 셀룰라 폰의 후방면 상에 배치될 수 있다. 하나의 마이크로폰은 셀룰라 폰의 정상 또는 좌측면 근처에 배치될 수 있고 반면 다른 마이크로폰은 셀룰라 폰의 바닥 또는 우측면 근처에 배치될 수 있다. 두개의 마이크로폰은 셀룰라 폰의 헤드셋의 다른 위치 상에 배치될 수 있다. 일 실시예에서, 두개의 마이크로폰은 헤드셋 상에 배치되고 두개 이상의 마이크로폰은 셀룰라 폰 핸드헬드 유닛 상에 배치된다. 그러므로 두개의 마이크로폰은 사용자가 핸드헬드 유닛 또는 헤드셋을 사용하는가의 여부와 상관없이 사용자의 음향을 기록할 수 있다.

비록 개선된 ICA 처리를 갖는 셀룰라 폰이 예로서 설명되었지만, 전자 설비, 유선 전화, 스피커폰, 무선 전화, 텔레콘퍼런스, CB 라디오, 워키-토키, 컴퓨터 전화통화 애플리케이션, 컴퓨터 및 자동차 음향 인식 애플리케이션, 감시 장치, 인터콤 등에 대한 음성 명령과 같은 다른 음향 통신 매체는 또한 다른 신호로부터 원하는 음향 신호를 분리하기 위해 개선된 ICA 처리를 이용한다.

도 10은 셀룰라 폰 장치의 다른 실시예를 도시한다. 셀룰라 폰 장치(1000)는 다른 셀룰라 폰과 같은 다른 통신 장치로부터 사운드 신호를 수신하는 두개의 채널(1010 및 1020)을 포함한다. 채널(1010 및 1020)은 두개의 마이크로폰에 의해 기록된 동일한 대화의 사운드 신호를 수신한다. 두개 이상의 유닛은 두개 이상의 입력 신호를 수신하기 위해 사용될 수 있다. 장치(1000)는 또한 배경 노이즈로부터 원하는 음향 신호를 분리하기 위해 수신된 신호를 처리하는 음향 분리 시스템(200)을 포함한다. 그다음 분리된 원하는 음향 신호는 셀룰라 폰 사용자의 귀에 도달하도록 증폭기(1030)에 의해 증폭된다. 수신 셀룰라 폰 상에 음향 분리 시스템(200)을 배치함으로써, 수신 셀룰라 폰의 사용자는 전송 셀룰라 폰이 음향 분리 시스템(200)을 구비하지 않았더라도 높은 품질의 음향을 들을 수 있다. 그러나, 이는 전송 셀룰라 폰 상의 두개의 마이크로폰에 의해 기록된 대화 신호의 두 채널을 수신하는 것을 필요로 한다.

쉬운 도시를 위해, 배터리, 디스플레이 패널 등과 같은 다른 셀룰라 폰 부속은 도 10에서 생략되었다. 아날로그-디지털 변환, 복조와 관련되거나 또는 FDMA, TDMA, CDMA 등을 가능케 하는 셀룰라 폰 신호 처리 단계는 또한 쉬운 도시를 위해 생략되었다.

본 발명의 특정 측면, 장점 및 신규한 피처가 여기에서 설명되었다. 물론, 이러한 측면, 장점 또는 피처 모두는 반드시 본 발명의 임의의 특별한 실시예에 의해 반드시 구체화되는 것은 아니라는 것으로 이해되어야 한다. 여기에서 논의된 실시예는 본 발명의 예로서 제공되었으며, 추가, 변경 및 조절될 수 있다. 예컨대, 비록 수학식 7, 8, 및 9가 비선형 유계 함수의 예를 나타내었지만, 비선형 유계 함수는 이러한 예에 한정되는 것이 아니라 소정 최대값 및 최소값을 갖는 임의의 비선형 함수를 포함할 수 있다. 그러므로, 본 발명의 기술적 사상은 다음의 청구의 범위에 의해 한정되어야 한다.

#### 참조문헌

Hyvaerinen, A., Karhunen, J, Oja, E. Independent component analysis. John Wiley & Sons, Inc. 2001

Te-Won Lee, Independent Component Analysis: Theory and Applications, Kluwer Academic Publishers, Boston, September 1998

Mark Girolami, Self-Organizing Neural Networks: Independent Component Analysis and Blind Source Separation. In Perspectives in Neural Computing, Springer Verlag, September 1999

Mark Girolami(Editor), Advances in Independent Component Analysis. In Perspectives in Neural Computing,, Springer Verlag, August 2000

Simon Haykin, Adaptive Filter Theory, Third Edition, Prentice-Hall(NJ), 1996.

Bell, A., Sejnowski, T., Neural Computation 7:1129-1159, 1995

Amari, S., Cichocki, A., Yang, H., A New Learning Algorithm for Blind Signal Separation, In: Advances in Neural Information Processing System 8, Editors D. Touretzky, M. Mozer, and M. Hasselmo, pp. 757-763, MIT Press, Cambridge MA, 1996

Cardoso, J.-F., Iterative techniques for blind source separation using only fourth order cumulants In Proc. EUSIPCO, pages 739-742, 1992.

Comon, P., Independent component analysis

## (57) 청구의 범위

### 청구항 1.

어쿠스틱 환경에서 원하는 음향 신호를 분리하는 방법에 있어서,

복수의 입력 신호를 수신하는 단계와,

안정성 강제 하에서 독립 성분 분석(ICA : Independent Component Analysis) 또는 블라인드 음원 분리(BSS : Blind Source Separation)를 사용하여 상기 수신된 입력 신호를 처리하는 단계와,

상기 수신된 입력 신호를 하나 이상의 원하는 오디오 신호 및 하나 이상의 노이즈 신호로 분리하는 단계를

포함하고,

상기 입력 신호는 상기 원하는 음향 신호 및 다른 어쿠스틱 신호에 대해 응답하여 생성되는 것을 특징으로 하는 음향 신호의 분리 방법.

### 청구항 2.

제 1 항에 있어서,

상기 원하는 오디오 신호 중의 하나는 상기 원하는 음향 신호인 것을 특징으로 하는 음향 신호의 분리 방법.

### 청구항 3.

제 1 항에 있어서,

상기 ICA 또는 BSS 처리는 근사법을 통해 직접 또는 간접적으로 상호 정보의 수학적인 형식을 최소화 또는 최대화하는 단계를 포함하는 것을 특징으로 하는 음향 신호의 분리 방법.

### 청구항 4.

제 1 항에 있어서,

ICA 웨이트 적응 역학과 보조를 맞추므로써 상기 ICA 처리를 안정화하는 단계를 추가로 포함하는 것을 특징으로 하는 음향 신호의 분리 방법.

#### 청구항 5.

제 1 항에 있어서,

웨이트 적응 속도를 강제하는 적응형 스케일링 인자를 사용하여 ICA 입력을 스케일링함으로써 상기 ICA 처리를 안정화하는 단계를 추가로 포함하는 것을 특징으로 하는 음향 신호의 분리 방법.

#### 청구항 6.

제 1 항에 있어서,

잔향 효과를 피하기 위해 시간 도메인 및 주파수 도메인에서 학습된 필터 웨이트를 필터링함으로써 상기 ICA 처리를 안정화하는 단계를 추가로 포함하는 것을 특징으로 하는 음향 신호의 분리 방법.

#### 청구항 7.

제 1 항에 있어서,

주변 처리 기술은 다양한 등급으로 상기 입력 및 분리된 신호에 적용되는 것을 특징으로 하는 음향 신호의 분리 방법.

#### 청구항 8.

제 1 항에 있어서,

상기 분리의 성능을 강화하기 위해 전처리 기술 또는 정보를 이용하는 단계를 추가로 포함하는 것을 특징으로 하는 음향 신호의 분리 방법.

#### 청구항 9.

제 8 항에 있어서,

상기 입력 신호에 적용되는 믹싱 시나리오의 조건을 개선하는 단계를 추가로 포함하는 것을 특징으로 하는 음향 신호의 분리 방법.

#### 청구항 10.

제 2 항에 있어서,

상기 분리된 원하는 음향 신호를 포함하는 채널을 식별하기 위해 상기 원하는 음향 신호의 특징적인 정보를 이용하는 단계를 추가로 포함하는 것을 특징으로 하는 음향 신호의 분리 방법.

### 청구항 11.

제 10 항에 있어서,

상기 특징적인 정보는 공간, 스펙트럼 또는 시간 정보인 것을 특징으로 하는 음향 신호의 분리 방법.

### 청구항 12.

제 1 항에 있어서,

후처리 기술은 적어도 하나의 상기 노이즈 신호 또는 적어도 하나의 상기 입력 신호를 이용하여 상기 원하는 신호의 품질을 개선하는데 사용되는 것을 특징으로 하는 음향 신호의 분리 방법.

### 청구항 13.

제 12 항에 있어서,

상기 원하는 음향 신호를 추가로 분리 및 강화하기 위해 상기 분리된 노이즈 신호를 사용하는 단계를 추가로 포함하는 것을 특징으로 하는 음향 신호의 분리 방법.

### 청구항 14.

제 13 항에 있어서,

상기 분리된 노이즈 신호 사용 단계는

노이즈 필터에 대한 상기 노이즈 스펙트럼을 평가하기 위해 상기 노이즈 신호를 사용하는 단계를 포함하는 것을 특징으로 하는 음향 신호의 분리 방법.

### 청구항 15.

제 1 항에 있어서,

적어도 두개의 마이크로폰을 공간을 두어 배치하는 단계와,

각각의 마이크로폰에서 상기 입력 신호의 하나를 생성하는 단계를

추가로 포함하는 것을 특징으로 하는 음향 신호의 분리 방법.

### 청구항 16.

제 15 항에 있어서,

상기 공간을 두어 배치하는 단계는 상기 마이크로폰을 대략 1mm ~ 1m 공간을 두어 배치하는 단계를 포함하는 것을 특징으로 하는 음향 신호의 분리 방법.

### 청구항 17.

제 15 항에 있어서,

상기 공간을 두어 배치하는 단계는 전화 수신기, 헤드셋, 또는 핸즈프리 키트 상의 상기 마이크로폰을 공간을 두어 배치하는 단계를 포함하는 것을 특징으로 하는 음향 신호의 분리 방법.

### 청구항 18.

제 15 항에 있어서,

상기 ICA 처리는

제 1 출력 채널 및 제 2 입력 채널에 연결된 제 1 적응 독립 성분 분석(ICA) 필터와,

제 1 입력 채널 및 제 2 출력 채널에 연결된 제 2 적응 독립 성분 분석(ICA) 필터를 포함하고,

상기 제 1 필터는 상기 노이즈 신호 채널에 대한 비선형 유계 부호 함수의 적용과 관련되는 순환 학습 규칙에 의해 적응되며,

상기 제 2 필터는 상기 원하는 음향 신호 채널에 대한 비선형 유계 부호 함수의 적용과 관련되는 순환 학습 규칙에 의해 적응되고,

상기 제 1 필터 및 상기 제 2 필터는 상기 원하는 음향 신호를 생성하기 위해 반복적으로 적용되는 것을 특징으로 하는 음향 신호의 분리 방법.

### 청구항 19.

제 18 항에 있어서,

(a) 상기 제 1 적응 독립 성분 분석 필터에 의해 순환하여 필터링된 상기 원하는 음향 채널은 패드백되어 상기 제 2 마이크로폰으로부터 상기 입력 채널로 부가되고, 이로써 상기 노이즈 신호 채널을 생성하고,

(b) 상기 제 2 적응 독립 성분 분석 필터에 의해 순환하여 필터링된 상기 노이즈 신호 채널은 패드백되어 상기 제 1 마이크로폰으로부터 상기 입력 채널로 부가되어, 상기 원하는 음향 신호 채널을 생성하는 것을 특징으로 하는 음향 신호의 분리 방법.

### 청구항 20.

제 19 항에 있어서,

상기 입력 채널 신호는 진입 신호의 에너지의 함수로서 순환식으로부터 계산된 적응형 스케일링 인자에 의해 줄어드는 (scaled down) 것을 특징으로 하는 음향 신호의 분리 방법.

### 청구항 21.

제 18 항에 있어서,

상기 제 1 적응 ICA 크로스 필터에 대한 상기 필터 웨이트 학습 규칙은 때를 맞추어 상기 필터 계수를 평탄화함으로써 안정화되고, 상기 제 2 적응 ICA 크로스 필터에 대한 상기 필터 웨이트 학습 규칙은 때를 맞추어 상기 필터 계수를 평탄화함으로써 안정화되는 것을 특징으로 하는 음향 신호의 분리 방법.

## 청구항 22.

제 18 항에 있어서,

상기 제 1 적응 ICA 크로스 필터의 웨이트는 상기 주파수 도메인에서 필터링되고, 상기 제 2 적응 ICA 크로스 필터의 웨이트는 상기 주파수 도메인에서 필터링되는 것을 특징으로 하는 음향 신호의 분리 방법.

## 청구항 23.

제 18 항에 있어서,

음성 활동 감지를 포함하는 단일 또는 복수의 채널 음향 강화 모듈을 적용하는 상기 원하는 음향 신호에 연결된 후처리 모듈을 추가로 포함하고,

상기 후처리 출력은 상기 입력 채널로 피드백되지 않는 특징으로 하는 음향 신호의 분리 방법.

## 청구항 24.

제 18 항에 있어서,

상기 ICA 처리는 필터 계수가 복수의 샘플링 순간에서 업데이트되고 다양한 사이즈의 필터 길이가 계산 능력을 이용가능하게 맞추는데 사용되는 것을 제외하고 상기 적응형 ICA 크로스 필터가 상기 모든 샘플링 순간에 적용되는 고정된 포인트 정밀도 환경에서 실행되는 것을 특징으로 하는 음향 신호의 분리 방법.

## 청구항 25.

제 18 항에 있어서,

상기 노이즈 신호를 사용하여 상기 원하는 음향 신호를 후처리하는 단계를 추가로 포함하고,

상기 후처리 모듈은 상기 노이즈 신호에 기초하여 상기 원하는 음향 신호에 스펙트럼 차감법을 적용하는 것을 특징으로 하는 음향 신호의 분리 방법.

## 청구항 26.

제 18 항에 있어서,

상기 노이즈 신호를 사용하여 상기 원하는 음향 신호를 후처리하는 단계를 추가로 포함하고,

상기 후처리 모듈은 상기 노이즈 신호에 기초하여 상기 원하는 음향 신호에 비너(Wiener) 필터링을 적용하는 것을 특징으로 하는 음향 신호의 분리 방법.

### 청구항 27.

제 18 항에 있어서,

제 3 채널로부터 제 3 세트의 오디오 입력 신호를 수신하는 단계와

제 3 필터를 사용하여 비선형 유계 함수를 진입 신호에 적용하는 단계를

추가로 포함하는 것을 특징으로 하는 음향 신호의 분리 방법.

### 청구항 28.

음향 장치에 있어서,

어쿠스틱 사운드 신호를 수신하며 공간을 두어 배치된 적어도 두개의 마이크로폰과,

상기 마이크로폰과 연결된 ICA 또는 BSS 프로세서를

포함하고,

상기 마이크로폰은 음향 음원으로부터 예정된 거리를 갖고,

상기 프로세서 동작은

상기 두개의 마이크로폰으로부터 사운드 신호를 수신하는 단계와,

안정성 강제 하의 상기 사운드 신호를 적어도 하나의 원하는 음향 신호 라인 및 적어도 하나의 노이즈 신호 라인으로 분리하는 단계를

포함하는 것을 특징으로 하는 음향 장치.

### 청구항 29.

제 28 항에 있어서,

상기 노이즈 라인 및 상기 원하는 음향 신호 라인에 연결된 후처리 필터를 추가로 포함하는 것을 특징으로 하는 음향 장치.

### 청구항 30.

제 28 항에 있어서,

상기 마이크로폰은 대략 1mm ~ 1m 공간을 두어 배치되는 것을 특징으로 하는 음향 장치.

**청구항 31.**

제 30 항에 있어서,

각 마이크로폰에서 수신된 상기 어쿠스틱 사운드 신호를 후처리하는 단계를 추가로 포함하는 것을 특징으로 하는 방법.

**청구항 32.**

제 28 항에 있어서,

상기 마이크로폰 중의 하나는 장치 하우징의 표면 상에 있고 상기 다른 마이크로폰은 상기 장치 하우징의 다른 표면 상에 있는 것을 특징으로 하는 음향 장치.

**청구항 33.**

제 28 항에 있어서,

상기 음향 장치는 무선 폰인 것을 특징으로 하는 음향 장치.

**청구항 34.**

제 28 항에 있어서,

상기 음향 장치는 무선 폰인 것을 특징으로 하는 음향 장치.

**청구항 35.**

제 28 항에 있어서,

상기 음향 장치는 핸즈 프리 자동차 키트인 것을 특징으로 하는 음향 장치.

**청구항 36.**

제 28 항에 있어서,

상기 음향 장치는 헤드셋인 것을 특징으로 하는 음향 장치.

**청구항 37.**

제 28 항에 있어서,

상기 음향 장치는 PDA인 것을 특징으로 하는 음향 장치.

**청구항 38.**

제 28 항에 있어서,

상기 음향 장치는 핸드헬드 바코드 스캐닝 장치인 것을 특징으로 하는 음향 장치.

### 청구항 39.

어쿠스틱 환경에서 원하는 음향 신호를 분리하는 시스템에 있어서,

하나 이상의 어쿠스틱 신호를 각각 수신하는 복수의 입력 채널과,

적어도 하나의 ICA 또는 BSS 필터와,

분리된 신호를 전송하는 복수의 출력 채널을

포함하고,

상기 필터는 안정성 강제 하의 상기 수신된 신호를 하나 이상의 원하는 오디오 신호 및 하나 이상의 노이즈 신호로 분리하는 것을 특징으로 하는 음향 신호의 분리 시스템.

### 청구항 40.

제 39 항에 있어서,

상기 원하는 오디오 신호는 상기 복수의 어쿠스틱 신호에서 수신된 음향 신호인 것을 특징으로 하는 음향 신호의 분리 시스템.

### 청구항 41.

제 39 항에 있어서,

상기 필터는 근사법을 통해 직접 또는 간접적으로 상호 정보의 수학적 형식을 조절하는 것을 특징으로 하는 음향 신호의 분리 시스템.

### 청구항 42.

제 39 항에 있어서,

상기 필터는 ICA 웨이트 적응 역학과 보조를 맞추으로써 상기 ICA 처리를 안정화하는 것을 특징으로 하는 음향 신호의 분리 시스템.

### 청구항 43.

제 39 항에 있어서,

상기 필터는 웨이트 적응 속도를 강제하는 적응형 스케일링 인자를 사용하여 ICA 입력을 스케일링함으로써 상기 ICA 처리를 안정화하는 것을 특징으로 하는 음향 신호의 분리 시스템.

**청구항 44.**

제 39 항에 있어서,

간향 효과를 피하기 위해 시간 도메인 및 주파수 도메인에서 학습된 필터 웨이트를 필터링함으로써 상기 ICA 처리를 안정화하는 것을 특징으로 하는 음향 신호의 분리 시스템.

**청구항 45.**

제 39 항에 있어서,

상기 입력 및/또는 출력 신호에 적용되는 하나 이상의 주변 처리 필터를 추가로 포함하는 것을 특징으로 하는 음향 신호의 분리 시스템.

**청구항 46.**

제 45 항에 있어서,

하나 이상의 전처리 필터를 추가로 포함하는 것을 특징으로 하는 음향 신호의 분리 시스템.

**청구항 47.**

제 45 항에 있어서,

하나 이상의 후처리 필터를 추가로 포함하는 것을 특징으로 하는 음향 신호의 분리 시스템.

**청구항 48.**

제 39 항에 있어서,

상기 입력 채널에 연결된 하나 이상의 마이크론을 추가로 포함하는 것을 특징으로 하는 음향 신호의 분리 시스템.

**청구항 49.**

제 48 항에 있어서,

대략 1mm ~ 1m 공간을 두어 각각 배치되는 두개 이상의 마이크론을 포함하는 것을 특징으로 하는 음향 신호의 분리 시스템.

**청구항 50.**

제 39 항에 있어서,

상기 시스템은 핸드헬드 장치 상에 있는 것을 특징으로 하는 음향 신호의 분리 시스템.

### 청구항 51.

제 39 항에 있어서,

상기 필터는

제 1 출력 채널 및 제 2 입력 채널에 연결된 제 1 적응 독립 성분 분석(ICA) 필터와,

제 1 출력 채널 및 제 2 입력 채널에 연결된 제 2 적응 독립 성분 분석(ICA) 필터를 포함하고,

상기 제 1 필터는 상기 노이즈 신호 채널에 대한 비선형 유계 부호 함수의 적용과 관련되는 순환 학습 규칙에 의해 적응되며,

상기 제 2 필터는 상기 원하는 음향 신호 채널에 대한 비선형 유계 부호 함수의 적용과 관련되는 순환 학습 규칙에 의해 적응되고,

상기 제 1 필터 및 상기 제 2 필터는 상기 원하는 음향 신호를 생성하기 위해 반복적으로 적용되는 것을 특징으로 하는 음향 신호의 분리 시스템.

### 청구항 52.

음향 신호를 분리하는 시스템에 있어서,

일 세트의 신호 생성기와,

상기 혼합된 신호를 각각 수신하는 프로세서와,

상기 음향 신호를 수신하는 음향 가능화 유닛을

포함하고,

상기 각 신호 생성기는 상기 음향 신호 및 다른 어쿠스틱 신호의 혼합을 나타내는 혼합 신호를 생성하기 위해 배치되며,

상기 프로세서 동작은

안정성 강제 하의 독립 성분 분석(ICA) 또는 블라인드 음원 분리(BSS)를 사용하여 혼합된 신호의 상기 세트를 처리하는 단계와,

상기 혼합된 신호를 상기 음향 신호 및 적어도 하나의 노이즈 신호로 분리하는 단계를 추가로 포함하는 것을 특징으로 하는 음향 신호의 분리 시스템.

### 청구항 53.

제 52 항에 있어서,

상기 신호 생성기는 어쿠스틱 트랜스듀서인 것을 특징으로 하는 음향 신호의 분리 시스템.

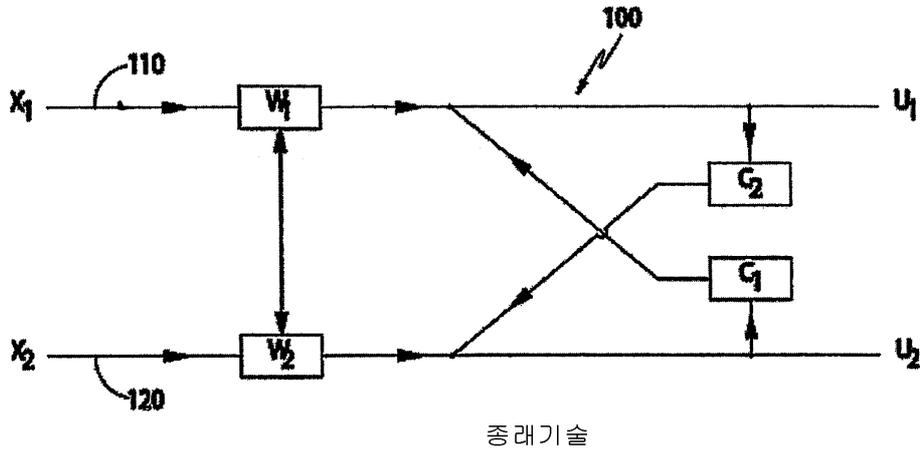
### 청구항 54.

제 53 항에 있어서,

상기 어쿠스틱 트랜스듀서는 인간 음향 주파수 범위 내의 어쿠스틱 신호를 수신하는 마이크로폰인 것을 특징으로 하는 음향 신호의 분리 시스템.

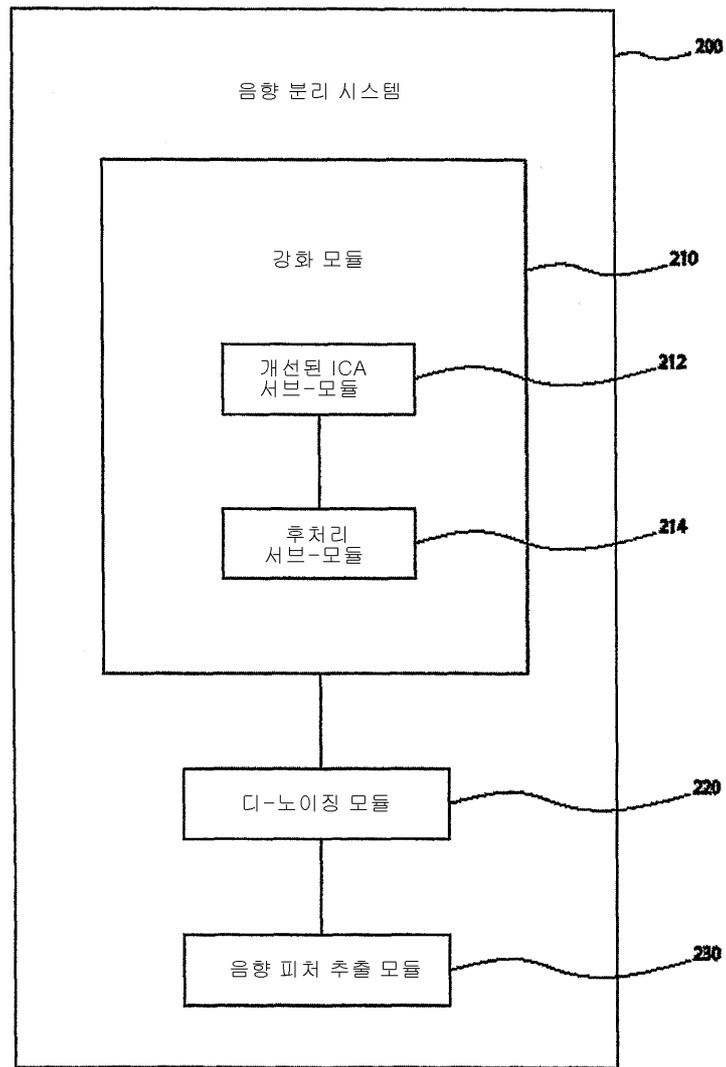
도면

도면1

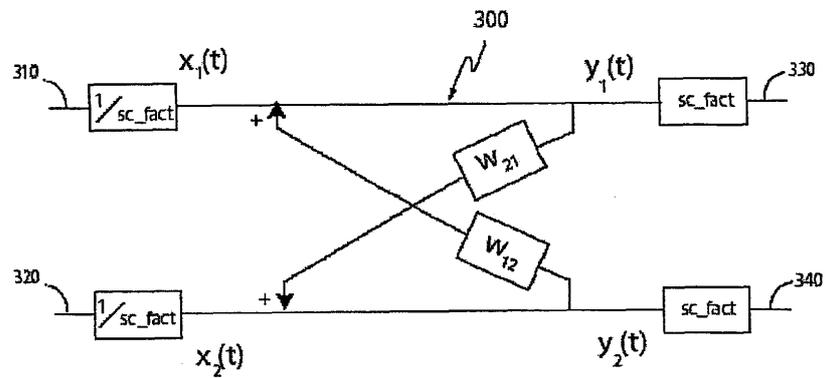


종래기술

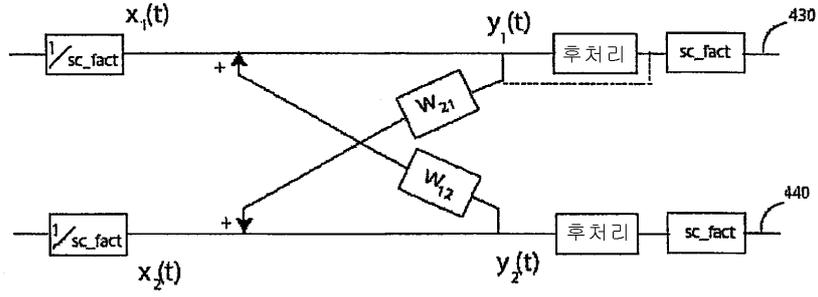
도면2



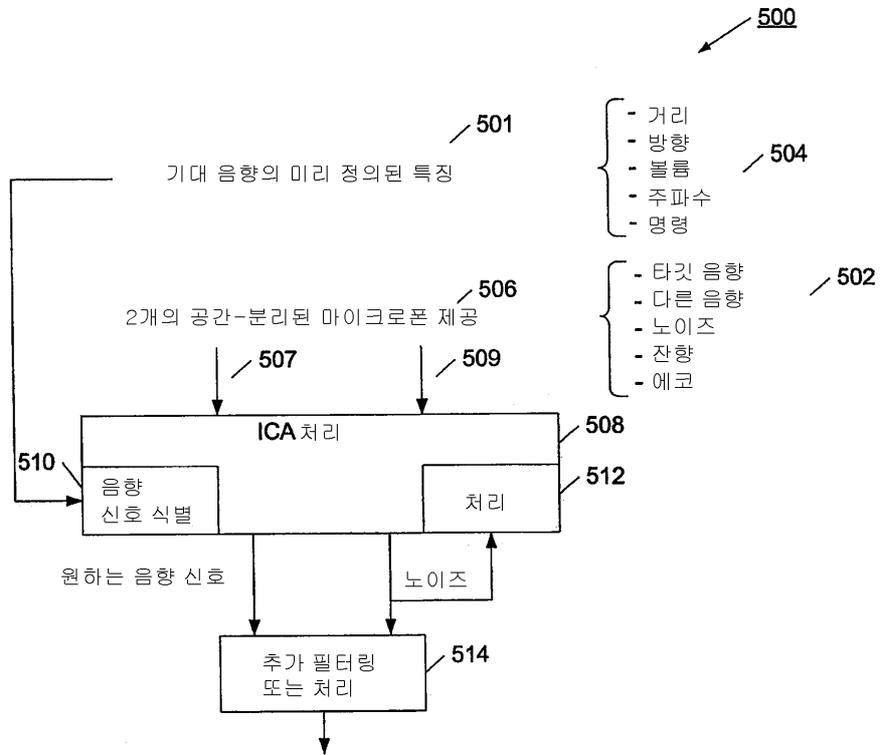
도면3



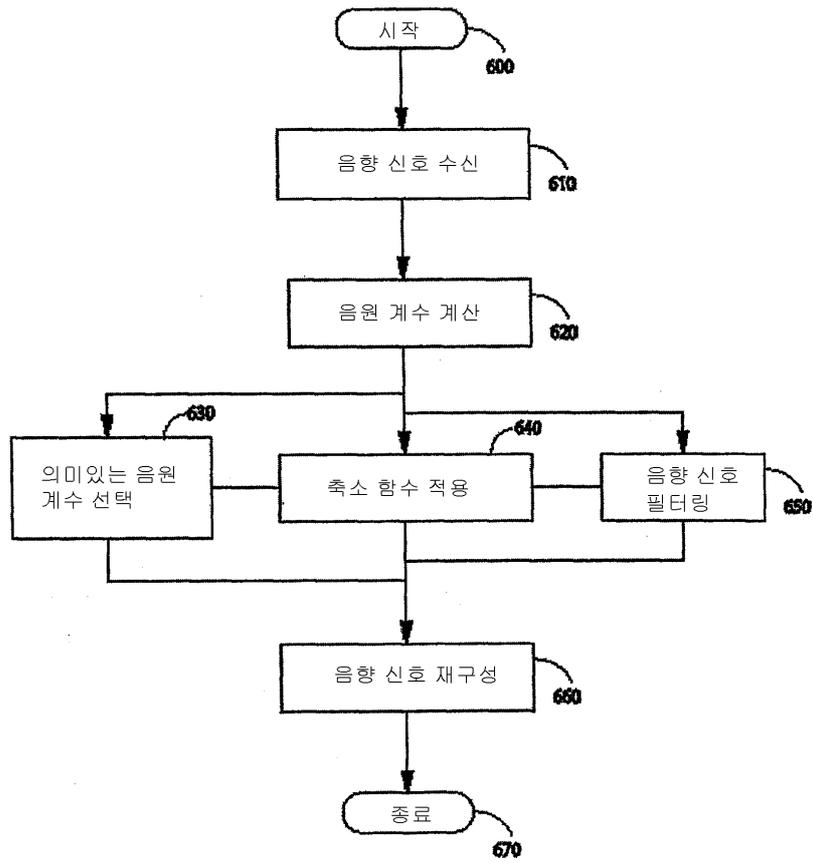
도면4



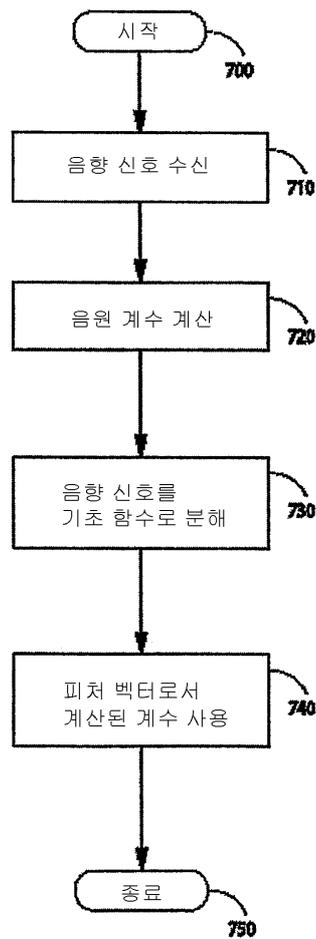
도면5



도면6



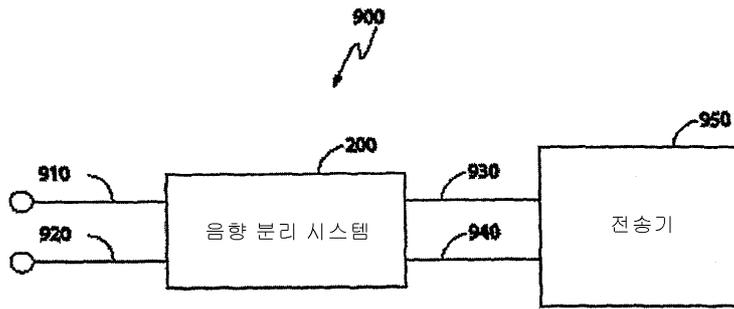
도면7



도면8

신호 형태	바람직한 처리
810 큰 노이즈 + 경쟁 음원	음향 강화 + 음향 디-노이즈 + 음향 피치 추출
820 경쟁 음원	음향 강화 + 음향 피치 추출
830 큰 노이즈	음향 디-노이즈 + 음향 피치 추출
840 상대적으로 깨끗한 음향	음향 피치 추출

도면9



도면10

