

	(19) 대한민국특허청(KR) (12) 공개특허공보(A)	(11) 공개번호 10-2012-0133038 (43) 공개일자 2012년12월10일
(51) 국제특허분류(Int. Cl.) G10L 19/00 (2006.01)	(71) 출원인 서강대학교산학협력단 서울특별시 마포구 백범로 35 (신수동, 서강대학교)	
(21) 출원번호 10-2011-0051512 (22) 출원일자 2011년05월30일 심사청구일자 2011년05월30일	(72) 발명자 박형민 서울특별시 서초구 잠원동 56-3 잠원한신아파트 5동 903호 오명우 서울특별시 마포구 승문길 106, 103동 1207호 (염리동, 상록아파트)	
	(74) 대리인 이지연	

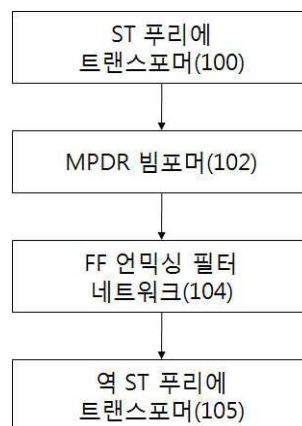
전체 청구항 수 : 총 6 항

(54) 발명의 명칭 피드 포워드 네트워크를 이용하는 독립 벡터 분석에 따른 블라인드 소스 분리 방법 및 장치

(57) 요약

본 발명의 블라인드 소스 분리 장치는, 혼합신호들을 FD의 혼합신호들로 변환하는 ST(Short-Time) 푸리에 트랜스포밍하는 ST(Short-Time) 푸리에 트랜스포머; 상기 FD의 혼합신호들을 소스신호들로 분리하는 FF(Feed Forward) 언믹싱 필터 네트워크; 상기 FD의 분리된 소스 신호들을 TD로 되돌리는 역 ST(Short-Time) 푸리에 트랜스포머; 를 포함하는 것을 특징으로 한다.

대표도 - 도1



이 발명을 지원한 국가연구개발사업

과제고유번호 2010-0027797

부처명 교육과학기술부

연구사업명 중견연구자지원사업

연구과제명 모바일 단말기를 위한 지능형 음성대화 인터페이스 개발

주관기관 서강대학교 산학협력단

연구기간 2010.09.01 ~ 2011.08.31

특허청구의 범위

청구항 1

블라인드 소스 분리 장치에 있어서,

TD(Time-Domain)의 혼합신호들을 FD(Frequency-Domain)의 혼합신호들로 변환하는 ST(Short-Time) 푸리에 트랜스포머;

상기 FD의 혼합신호들을 수학적 식 8에 따라 FD의 소스신호들로 분리하는 FF 언믹싱 필터 네트워크;

상기 분리된 FD의 소스신호들을 TD의 소스신호들로 변환하는 역 ST 푸리에 트랜스포머;를 구비하는 것을 특징으로 하는 블라인드 소스 분리 장치.

수학적 식 8

$$\hat{\mathbf{s}}(k, n) = \sum_{m=0}^U \mathbf{W}(k, m) \hat{\mathbf{x}}(k, n - m)$$

상기 수학적 식 8에서 상기 $\hat{\mathbf{s}}(k, n)$ 는 소스신호들의 ST(short time) FD 형태를 나타내며, 상기 $\mathbf{W}(k, m)$ 은 언믹싱 필터 계수 행렬(unmixing filter coefficient matrix)을 나타내며, 상기 k 는 주파수 빈을 나타내고, 상기 n 은 프레임을 나타내며, 상기 m 은 언믹싱 필터 계수 행렬을 입력 신호에 컨볼루션(convolution)하기 위해 입력신호를 쉬프팅하는 프레임 숫자이며, 상기 U 는 언믹싱 필터 계수 행렬의 프레임 단위의 길이를 나타내며, $\hat{\mathbf{x}}(k, n)$ 는 FD의 혼합신호들임.

청구항 2

제1항에 있어서,

상기 ST 푸리에 트랜스포머로부터 FD의 혼합신호들을 제공받아 수학적 식 9에 따라 전처리하여 상기 $\hat{\mathbf{x}}(k, n)$ 을 생성하여 상기 FF 언믹싱 필터 네트워크로 제공하는 MPDR 빔포머;를 더 구비함을 특징으로 하는 블라인드 소스 분리 장치.

수학적 식 9

$$\hat{\mathbf{x}}(k, n) = (\mathbf{D}^H(k) [\mathbf{R}(k) + \lambda \mathbf{I}]^{-1} \mathbf{D}(k))^{-1} \mathbf{D}^H(k) [\mathbf{R}(k) + \lambda \mathbf{I}]^{-1} \mathbf{x}(k, n),$$

상기 수학적 식 9에서 $\hat{\mathbf{x}}(k, n)$ 는 MPDR 빔 포머의 출력이며, 상기 $\mathbf{D}(k)$ 는 소스들에 대한 스티어링 벡터(steering vector)로 구성되는 행렬 $[\mathbf{d}_1(k) \cdots \mathbf{d}_L(k)]$ 을 나타내며, 이는 소스신호들과 마이크 간의 방향에 따른 각도정보에 의해 변화하는 복소평면에서의 페이즈 값과 거리에 의한 신호의 크기 감쇄값을 곱한 형태이며, 상기 $\mathbf{D}^H(k)$ 는 상기 $\mathbf{D}(k)$ 의 켤레 전치 벡터이고, 상기 $\mathbf{I}$ 는 단위행렬이고, 상기 $\mathbf{R}(k)$ 는 입력 스펙트럼의 공분산 행렬(input spectral covariance matrix)을 나타내고, 상기 λ 는 $\mathbf{R}(k)$ 의 역행렬 값을 제한하기 위한 상수(small positive constant)임.

청구항 3

제2항에 있어서,

상기 언믹싱 필터 계수 행렬 $\mathbf{W}(k, m)$ 은 수학적 식 10에 따라 학습됨을 특징으로 하는 블라인드 소스 분리 장치.

수학식 10

$$\Delta \mathbf{W}(k, m) \propto - \sum_{r=0}^{\infty} \{ \text{off-diag}(\varphi^{(k)}(\hat{\mathbf{t}}(n-U)) \hat{\mathbf{s}}^H(k, n-U-m+r)) \\ + \beta(\hat{\mathbf{s}}(k, n-U) - \hat{\mathbf{x}}(k, n-3U/2)) \hat{\mathbf{s}}^H(k, n-U-m+r) \} \mathbf{W}(k, r),$$

상기 수학식 10에서, 상기 off-diag(?)은 대각 성분이 0인 행렬을 나타내며, β 는 MDP의 가중치(weight constant)이며, r 은 언믹싱 필터 계수 행렬을 입력 신호에 콘볼루션(convolution) 하기 위해 입력 신호를 쉬프팅하는 프레임 숫자임.

청구항 4

블라인드 소스 분리 방법에 있어서,

TD의 혼합신호들을 FD의 혼합신호들로 변환하는 단계;

상기 FD의 혼합신호들을 수학식 11에 따라 FD의 소스신호들로 분리하는 단계;

상기 분리된 FD의 소스신호들을 TD의 소스신호들로 변환하는 단계;를 구비하는 것을 특징으로 하는 블라인드 소스 분리 방법.

수학식 11

$$\hat{\mathbf{s}}(k, n) = \sum_{m=0}^U \mathbf{W}(k, m) \hat{\mathbf{x}}(k, n-m)$$

상기 수학식 11에서 상기 $\hat{\mathbf{s}}(k, n)$ 는 소스신호들의 ST(short time) FD 형태를 나타내며, 상기 $\mathbf{W}(k, m)$ 은 언믹싱 필터 계수 행렬(unmixing filter coefficient matrix)을 나타내며, 상기 k 는 주파수 빈을 나타내고, 상기 n 은 프레임을 나타내며, 상기 m 은 언믹싱 필터 계수 행렬을 입력 신호에 콘볼루션(convolution)하기 위해 입력신호를 쉬프팅하는 프레임 숫자이며, 상기 U 는 언믹싱 필터 계수 행렬의 프레임 단위의 길이를 나타내며, $\hat{\mathbf{x}}(k, n)$ 는 FD의 혼합신호들임.

청구항 5

제4항에 있어서,

상기 FD의 혼합신호들을 제공받아 수학식 12에 따라 전처리하여 상기 $\hat{\mathbf{x}}(k, n)$ 를 생성하는 단계;를 더 구비함을 특징으로 하는 블라인드 소스 분리 방법.

수학식 12

$$\hat{\mathbf{x}}(k, n) = (\mathbf{D}^H(k)[\mathbf{R}(k) + \lambda \mathbf{I}]^{-1} \mathbf{D}(k))^{-1} \mathbf{D}^H(k)[\mathbf{R}(k) + \lambda \mathbf{I}]^{-1} \mathbf{x}(k, n),$$

상기 수학식 12에서 $\hat{\mathbf{x}}(k, n)$ 는 MPDR 빔 포머의 출력이며, 상기 $\mathbf{D}(k)$ 는 소스들에 대한 스티어링 벡터(steering vector)로 구성되는 행렬 $[\mathbf{d}_1(k) \cdots \mathbf{d}_L(k)]$ 을 나타내며, 이는 소스신호들과 마이크 간의 방향에 따른 각도정보에 의해 변화하는 복소평면에서의 페이즈 값과 거리에 의한 신호의 크기 감쇄값을 곱한 형태이며, 상기 $\mathbf{D}^H(k)$ 는 상기 $\mathbf{D}(k)$ 의 켤레 전치 벡터이고, 상기 $\mathbf{I}$ 는 단위행렬이고, 상기 $\mathbf{R}(k)$ 는 입력 스펙트럼의 공분산 행렬(input spectral covariance matrix)을 나타내고, 상기 λ 는 $\mathbf{R}(k)$ 의 역행렬 값을 제한하기 위한 상수(small positive constant)임.

청구항 6

제5항에 있어서,

상기 언믹싱 필터 계수 행렬 $W(k, m)$ 은 수학식 13에 따라 학습됨을 특징으로 하는 블라인드 소스 분리 방법.

수학식 13

$$\Delta W(k, m) \propto - \sum_{r=0}^{\infty} \{ \text{off-diag}(\varphi^{(k)}(\hat{\mathbf{t}}(n-U))\hat{\mathbf{s}}^H(k, n-U-m+r)) \\ + \beta(\hat{\mathbf{s}}(k, n-U) - \hat{\mathbf{x}}(k, n-3U/2))\hat{\mathbf{s}}^H(k, n-U-m+r) \} W(k, r),$$

상기 수학식 13에서, 상기 off-diag(?)은 대각 성분이 0인 행렬을 나타내며, β 는 MDP의 가중치(weight constant)이며, r 은 언믹싱 필터 계수 행렬을 입력 신호에 컨볼루션(convolution) 하기 위해 입력 신호를 쉬프팅하는 프레임 숫자임.

명세서

기술분야

[0001] 본 발명은 블라인드 소스 분리 기술에 관한 것으로, 더욱 상세하게는 피드 포워드 네트워크를 이용하는 독립 벡터 분석에 따른 블라인드 소스 분리 방법 및 장치에 관한 것이다.

배경기술

[0002] 블라인드 소스 분리(Blind Source Separation; BSS)란 혼합 환경 및 소스 신호를 알지 못한 상태에서 혼합된 혼합 신호들로부터 소스 신호들을 분리하는 것으로, 음향 및 영상 처리, 생체 신호 분석 및 통신 등의 다양한 분야에 적용되고 있다.

[0003] 상기한 블라인드 신호 분리를 위한 방법 중 하나인 독립 성분 분석(Independent Component Analysis; ICA) 기술은 혼합 신호들의 고차 통계량(higher-order statistics)을 이용하여 소스 신호들을 분리하는 것으로, 혼합 신호들을 시간 도메인(Time-Domain; TD) 또는 주파수 도메인(Frequency-Domain; FD)에서 디콘볼루션하여 처리하였다. 이러한 FD 방식을 채용하는 ICA 기술은 시간 TD 방식이 내재하고 있는 많은 계산과 느린 컨버전스 문제를 해소하기는 하였으나, 소스 신호들의 순서와 크기에 대한 모호성(Permutation and Scale Ambiguity)이 존재하는 문제를 내재하고 있었다. 이러한 ICA 기술에 대해서는 Hyvarinen, A., Karhunen, J., and Oja, E.: Independent Component Analysis(John Wiley & Sons, 2001)에 기술되어 있다.

[0004] 상기한 ICA 기술의 순서에 대한 모호성 문제는 독립 벡터 분석(Independent Vector Analysis; IVA) 기술에 의해 해소될 수 있다. 즉 상기 IVA 기술은 주파수 빈 각각에 대해 독립적인 ICA 기술과 달리 주파수 빈 사이에 서로 의존성을 가지도록 하여 소스 신호들의 순서에 대한 모호성을 해소하여 블라인드 신호 분리 성능을 향상시켰다. 이러한 IVA 기술에 대해서는 Kim, T., Attias, H. T., Lee, S.-Y., and Lee, T.-W.: Blind source separation exploiting higher-order frequency dependencies, *IEEE Trans. Audio, Speech, and Language Processing*, 2007, 15, pp. 70-79에 기술되어 있다. 또한, 크기에 대한 모호성 문제는 최소 왜곡 원리(Minimal Distortion Principle; MDP) 기술에 의해 해결할 수 있다. 즉 혼합 전의 소스 신호의 크기는 알 수 없으므로 혼합 신호에 가장 유사한 크기의 소스 분리 결과 신호를 구한다. 이러한 최소 왜곡 원리 기술에 대해서는 Matsuoka, K., and Nakashima, S.: Minimal distortion principle for blind source separation, Int. Workshop on ICA and BSS, 2001, pp. 722-727에 기술되어 있다.

[0005] 상기한 IVA 기술은 일반적인 ICA 기술과 같이 각 주파수 빈에 대해 단순 가중치 언믹싱 행렬로 컨볼루션을 이행하여 소스 신호들을 분리하였으며, 이는 TD 접근 방식의 컨볼루션이 FD 접근 방식에서 단순 곱으로 대체되기 때 문이었다. 그러나 이러한 대체는 프레임 길이가 믹싱 프로세스의 반향을 감당하기에 적당할 때에만 유효하다는 한계가 있었다. 이는 Kim, L.-H., Tashev, I., and Acero, A.: Reverberated speech signal separation based on regularized subband feedforward ICA and instantaneous direction of arrival, ICASSP, 2010, pp. 2678-2681에 기술되어 있다.

[0006] 그러나 실제 환경에서의 음향 반향은 단순 가중치 언믹싱 행렬이 처리할 수 있는 프레임 길이에 비해 긴 경우가 많다.

[0007] 이에따라 Kim, L.-H., Tashev, I., and Acero, A.: Reverberated speech signal separation based on

regularized subband feedforward ICA and instantaneous direction of arrival, ICASSP, 2010, pp. 2678-2681에서는, 높은 반향 환경에서 소스 신호들을 분리하기 위해 피드 포워드 언믹싱 필터 구조(FeedForward unmixing filter structure)를 채용하는 FD 방식 ICA 기술을 제안하였다. 이 제안된 FF ICA(FeedForward ICA) 기술은 순서 및 크기의 모호성없이 원하는 출력 채널로 분리된 소스 신호를 제공하기 위해 소스들의 공간 정보를 적용하고, 공간 정보에 따르는 첫번째 스테이지의 출력에 가까운 출력 신호들을 가능한 얻을 수 있도록 부가적인 제약조건 텀을 포함하는 학습 규칙을 가지고 언믹싱 필터를 갱신하였다.

[0008] 상기한 FF ICA 기술은 반향이 긴 혼합 신호들에 적용할 수 있었으나, 여전히 소스 신호들의 주파수간 독립성에 따른 문제를 가지고 있었다. 부가적으로 적용된 엔트로피 최대화 학습 규칙은 분리된 소스신호들의 프레임간 상관성을 떨어뜨리는 부작용에 의해 음질이 손상되었다. 또한 FD 방식 ICA 기술에 바탕을 두고 있으므로 순서 모호성 문제가 직접적으로 해결되었다고 볼 수 없었고 다만 첫 스테이지의 빔 포밍 출력과 학습 규칙의 부가적인 제약조건 텀이 채용된 휴리스틱한 기술을 통해 순서 바뀔의 가능성을 줄였을 뿐이었다.

[0009] 이에 상기 반향이 긴 혼합 신호를 분리할 때의 FF ICA 기술의 한계와 IVA의 문제점을 극복하기 위한 기술의 개발이 절실하게 요망되었다.

발명의 내용

해결하려는 과제

[0010] 본 발명은 독립 성분 분석이 아닌 독립 벡터 분석을 사용하여 반향이 긴 혼합 신호들의 분리에서 휴리스틱한 기술 없이도 주파수간 독립성에 따른 문제를 해소할 수 있게 함은 물론이며 프레임 길이의 한계를 해소할 수 있는 피드 포워드 네트워크를 이용하는 블라인드 소스 분리 방법 및 장치를 제공하는 것을 그 목적으로 한다.

[0011] 또한 본 발명의 다른 목적은 선택적인 전처리 단계로서 널 포밍 제한을 포함하는 최소 파워 왜곡 응답(Minimum Power Distortionless Response;MPDR) 빔 포머를 적용한 피드 포워드 네트워크를 이용하는 독립 벡터 분석에 따른 블라인드 소스 분리 방법 및 장치를 제공하는 것이다.

[0012] 또한 본 발명의 또 다른 목적은 분리된 소스 신호들에서의 프레임간 화이트닝 효과와 스케일링 불확정성을 피하기 위해 확장된 넌-홀로노믹 제한(extended non-holonomic constraint)과 최소 왜곡 원리를 적용한 피드 포워드 네트워크를 이용하는 독립 벡터 분석에 따른 블라인드 소스 분리 방법 및 장치를 제공하는 것이다.

과제의 해결 수단

[0013] 상기한 목적을 달성하기 위한 본 발명의 블라인드 소스 분리 장치는, 혼합신호들을 FD의 혼합신호들로 변환하는 ST(Short-Time) 푸리에 트랜스포밍하는 ST(Short-Time) 푸리에 트랜스포머; 상기 FD의 혼합신호들을 소스신호들로 분리하는 FF(Feed Forward) 언믹싱 필터 네트워크; 상기 FD의 분리된 소스 신호들을 TD로 되돌리는 역 ST(Short-Time) 푸리에 트랜스포머; 를 포함하는 것을 특징으로 한다.

발명의 효과

[0014] 상기한 본 발명은 독립 성분 분석을 확장한 독립 벡터 분석을 사용하였고, 프레임 길이의 한계를 해소할 수 있는 피드 포워드 네트워크를 이용하였기 때문에 반향이 긴 혼합 신호들을 분리할 때에 휴리스틱한 기술 없이도 주파수간 독립성에 따른 문제를 해소할 수 있게 함은 물론이며 프레임 길이의 한계를 해소할 수 있는 효과가 있다.

[0015] 또한 본 발명은 선택적인 전처리 단계로서 널 포밍 제한을 포함한 최소 파워 왜곡 응답(Minimum Power Distortionless Response;MPDR) 빔 포머를 적용하여 향상된 소스 신호를 초기값으로 사용할 수 있다.

[0016] 또한 본 발명은 확장된 넌-홀로노믹 제한(extended non-holonomic constraint)과 최소 왜곡 원리를 적용하여, 언믹스트 소스 신호들에서의 프레임간 화이트닝 효과와 스케일링 불확정성을 피할 수 있는 효과가 있다.

도면의 간단한 설명

[0017] 도 1은 본 발명의 바람직한 실시예에 따른 블라인드 소스 분리 장치의 구성도.

도 2는 본 발명의 바람직한 실시예에 따른 블라인드 소스 분리 성능을 측정하기 위해 소스 및 마이크 설치예를 도시한 도면.

도 3은 본 발명에 따른 블라인드 소스 분리 방법과 종래의 블라인드 소스 분리 방법의 블라인드 소스 분리 성능 분석 결과표를 도시한 도면.

발명을 실시하기 위한 구체적인 내용

- [0018] 본 발명의 바람직한 실시예에 따른 블라인드 소스 분리 장치의 구성을 도 1을 참조하여 설명한다.
- [0019] 상기 블라인드 소스 분리 장치는 ST(Short-time) 푸리에 트랜스포머(100)와 MPDR(Minimum Power Distortionless Response) 빔 포머(102)와 FF(Feed Forward) 언믹싱 필터 네트워크(104)와 역 ST(Short-time) 푸리에 트랜스포머(105)로 구성된다.
- [0020] 상기 ST 푸리에 트랜스포머(100)는 TD의 혼합신호들을 입력받아 FD의 혼합신호들로 변환한다.
- [0021] 상기 FD의 혼합신호들은 수학적 식 1에 따라 나타낼 수 있다.

수학적 식 1

- [0022]
$$\mathbf{x}(k, n) = [x_1(k, n) \cdots x_L(k, n)]^T$$
- [0023] 상기 수학적 식 1에서 $\mathbf{x}(k, n)$ 은 혼합신호들을 나타내고, k 는 주파수 빈을 나타내고, n 은 프레임을 나타내고, L 은 소스신호들 및 혼합신호들의 수를 나타낸다.
- [0024] 상기 FD의 혼합신호들은 소스신호들의 공간정보에 따르는 널 포밍 제한(null forming constraints)을 포함한 MPDR(Minimum Power Distortionless Response) 빔 포머(102)에 입력된다.
- [0025] 상기 MPDR 빔 포머(102)는 수학적 식 2에 따라 상기 FD의 혼합신호들의 성분(component)들의 간섭을 억제하여 출력한다.

수학적 식 2

- [0026]
$$\hat{\mathbf{x}}(k, n) = [\hat{x}_1(k, n) \cdots \hat{x}_L(k, n)]^T$$
- [0027]
$$\hat{\mathbf{x}}(k, n) = (\mathbf{D}^H(k)[\mathbf{R}(k) + \lambda \mathbf{I}]^{-1} \mathbf{D}(k))^{-1} \mathbf{D}^H(k)[\mathbf{R}(k) + \lambda \mathbf{I}]^{-1} \mathbf{x}(k, n),$$
- [0028] 상기 수학적 식 2에서 $\hat{\mathbf{x}}(k, n)$ 은 MPDR 빔 포머(102)의 출력이며, 상기 $\mathbf{D}(k)$ 는 소스신호들에 대한 스티어링 벡터(steering vector)로 구성되는 행렬 $[\mathbf{d}_1(k) \cdots \mathbf{d}_L(k)]$ 을 나타내며, 이는 소스신호들과 마이크 사이의 방향에 따른 각도정보에 의해 변화하는 복소평면에서의 페이즈 값, 그리고 거리에 의한 신호의 크기 감쇄값을 곱한 형태로 표현된다. 상기 $\mathbf{D}^H(k)$ 는 상기 $\mathbf{D}(k)$ 의 켤레 전치 벡터, 즉 켤레전치 벡터이다. 그리고 상기 $\mathbf{I}$ 는 단위행렬이다. 그리고 상기 $\mathbf{R}(k)$ 는 입력 스펙트럼의 공분산 행렬(input spectral covariance matrix)을 나타내고, 상기 λ 는 $\mathbf{R}(k)$ 의 역행렬 값이 매우 커지는 것을 피하기 위한 작은 양의 상수(small positive constant)이다. 상기 수학적 식 2는 입력 스펙트럼의 공분산 행렬과 소스신호들의 방향정보가 담긴 스티어링 벡터(steering vector)로 구성되는 행렬을 이용하여 관심 소스신호의 방향 외의 신호인 간섭 신호에 대한 파워를 최소화함으로써, 간섭 신호 대비 관심 소스 신호를 강화시킨다.
- [0029] 상기 혼합신호들에 대한 전처리가 완료되면, 상기 전처리된 혼합신호들은 FF 언믹싱 필터 네트워크(104)로 제공된다.
- [0030] 상기 FF 언믹싱 필터 네트워크(104)는 수학적 식 3에 따라 전처리된 혼합신호들을 입력받아 소스신호들로 분리한다.

수학식 3

$$\hat{\mathbf{s}}(k, n) = \sum_{m=0}^U \mathbf{W}(k, m) \hat{\mathbf{x}}(k, n - m)$$

[0031]

[0032] 상기 수학식 3에서 상기 $\hat{\mathbf{s}}(k, n)$ 는 소스신호들의 ST(short time) FD 형태를 나타내며, 상기 $\mathbf{W}(k, m)$ 은 언믹싱 필터 계수 행렬(unmixing filter coefficient matrix)을 나타낸다. 상기 k 는 주파수 빈을 나타내고, 상기 n 은 프레임을 나타내며, m 은 언믹싱 필터 계수 행렬을 입력 신호에 콘볼루션(convolution) 하기 위해 입력 신호를 쉬프팅하는 프레임 숫자이며, U 는 언믹싱 필터 계수 행렬의 프레임 단위의 길이를 나타낸다.

[0033] 한편 독립성의 측정으로 추정된 소스신호들 간의 결합 확률 밀도(joint probability density function) $p(\hat{\mathbf{t}}_1(n) \cdots \hat{\mathbf{t}}_L(n))$ 과 추정된 소스신호들에 대한 확률 밀도 모델의 곱 $\prod_{i=1}^L q(\hat{\mathbf{t}}_i(n))$ 사이의 Kullback-Leibler 발산(Kullback-Leibler divergence)을 적용한다. 여기서, $\hat{\mathbf{t}}_i(n) = [\hat{s}_i(1, n) \cdots \hat{s}_i(K, n)]$ 으로 표현하며, K 는 주파수 빈의 수이다. 그리고 L 은 소스신호들 및 혼합신호들의 수를 나타낸다. 따라서 $\hat{\mathbf{t}}_i(n) = [\hat{s}_i(1, n) \cdots \hat{s}_i(K, n)]$ 은 추정된 소스신호들의 주파수별 각 성분들을 모은 벡터가 된다.

[0034] 상기 소스신호들간 독립성 최대화 기반 FF 언믹싱 네트워크 추정을 위한 비용 함수(cost function) J 는 수학식 4와 같다.

$$J = - \sum_{k=1}^K \log |\det \mathbf{W}(k, 0)| - \sum_{i=1}^L E\{\log q(\hat{\mathbf{t}}_i(n))\}$$

[0035]

[0036] 상기 수학식 4에서 k 는 주파수 빈을 나타내고, n 은 프레임을 나타내며, L 은 소스 신호들 및 혼합 신호들의 수를 나타낸다.

[0037] 상기한 비용함수 J 를 이용하여 자연 경사(natural gradient) 기반 확률 최소 경사법(stochastic gradient descent algorithm)에 의해 FF 언믹싱 네트워크를 추정하면 수학식 5와 같다.

수학식 5

$$\Delta \mathbf{W}(k, m) \propto \mathbf{W}(k, m) - \varphi^{(k)}(\hat{\mathbf{t}}(n - U)) \sum_{r=0}^U \hat{\mathbf{s}}^H(k, n - U - m + r) \mathbf{W}(k, r)$$

[0038]

[0039] 상기 $\varphi^{(k)}(\hat{\mathbf{t}}(n - U))$ 은 소스신호에 대한 다변수 스코어 함수를 말하고, 이것은 비인과성(noncausality)을 피하기 위한 U -샘플 지연을 포함하고 있다. 상기 r 은 언믹싱 필터 계수 행렬을 입력 신호에 콘볼루션(convolution) 하기 위해 입력 신호를 쉬프팅하는 프레임 숫자이며, $\hat{\mathbf{s}}^H(k, n - U - m + r)$ 은 소스신호에서 FF 언믹싱 네트워크에 의해 지연된 샘플값의 행렬의 켈레전치 행렬이다.

[0040] 다변수 스코어 함수에서 $\varphi^{(k)}(\hat{\mathbf{t}}(n)) = [\varphi^{(k)}(\hat{\mathbf{t}}_1(n)) \cdots \varphi^{(k)}(\hat{\mathbf{t}}_L(n))]^T$ 이고,

$$\varphi^{(k)}(\hat{\mathbf{t}}_i(n)) = -\frac{\partial \log q(\hat{\mathbf{t}}_i(n))}{\partial \hat{s}_i(k,n)} = \frac{\hat{s}_i(k,n)}{\sqrt{\sum_{k=1}^K |\hat{s}_i(k,n)|^2}}.$$
로 표현된다.

[0041] 그리고 엔트로피 최대화(entropy maximization)에 따른 언믹싱 소스 신호들의 프레임간 상관 제거 부작용을 막기 위해, 확장된 넌-홀로노믹 제한(extended non-holonomic constraint)와 MDP(Minimal Distortion Principle)을 적용하며, 이는 수학식 6에 따른다.

수학식 6

[0042]
$$\Delta \mathbf{W}(k, m) \propto - \sum_{r=0}^{\infty} \{ \text{off-diag}(\varphi^{(k)}(\hat{\mathbf{t}}(n-U)) \hat{\mathbf{s}}^H(k, n-U-m+r)) \\ + \beta(\hat{\mathbf{s}}(k, n-U) - \hat{\mathbf{x}}(k, n-3U/2)) \hat{\mathbf{s}}^H(k, n-U-m+r) \} \mathbf{W}(k, r),$$

[0043] 상기 수학식 6에서, 상기 off-diag(?)은 대각 성분이 0인 행렬을 나타내며, β 는 작은 양의 가중치(small positive weighting constant)이다. 또한 상기 r 은 언믹싱 필터 계수 행렬을 입력 신호에 콘볼루션(convolution) 하기 위해 입력 신호를 쉬프팅하는 프레임 숫자이다. 그러므로 $\beta(\hat{\mathbf{s}}(k, n-U) - \hat{\mathbf{x}}(k, n-3U/2))$ 은 추정된 소스 신호에서 마이크 입력신호의 FF 언믹싱 필터의 절반 길이의 딜레이 값을 뺀 값이고, 이를 통해 소스신호의 추정을 위한 학습의 왜곡을 최소화할 수 있다.

[0044] 또한
$$\Delta \mathbf{W}(k, m) \propto - \sum_{r=0}^{\infty} \{ \text{off-diag}(\varphi^{(k)}(\hat{\mathbf{t}}(n-U)) \hat{\mathbf{s}}^H(k, n-U-m+r)) \\ + \beta(\hat{\mathbf{s}}(k, n-U) - \hat{\mathbf{x}}(k, n-3U/2)) \hat{\mathbf{s}}^H(k, n-U-m+r) \} \mathbf{W}(k, r),$$
은 다변수 스코어 함수의 대각 성분들을 학습에 반영하지 않음으로써 학습시 소스신호의 추정에서 불필요한 크기 변화를 방지한다.

[0045] 한편 믹싱 시스템은 넌 미니멈 페이즈(non-minimum phase) 특성을 가질 수 있고 MDP에 의해 $\beta(\hat{\mathbf{s}}(k, n-U) - \hat{\mathbf{x}}(k, n-3U/2))$ 와 $\beta(\hat{\mathbf{s}}(k, n-U) - \hat{\mathbf{x}}(k, n-3U/2))$ 의 차이가 최소화되므로, FF 언믹싱 필터 계수들은 모든 주파수 빈들에서 $\mathbf{W}(k, U/2)$ 의 대각 요소만 1이 되고 나머지 요소는 모두 0이 되도록 초기화된다.

[0046] 이제 본 발명의 바람직한 실시예에 따른 블라인드 소스 분리 방법과 일반적인 FD 방식 ICA, IVA 및 FF ICA 사이의 SIR(signal-to-interference ratio)을 비교하며, 이 SIR 산술식은 수학식 7에 따른다.

수학식 7

[0047]
$$\text{SIR(dB)} = \frac{1}{L} \sum_{i=1}^L 10 \log_{10} \left(\frac{\text{target signal energy in the } i\text{-th output}}{\text{interference signal energy in the } i\text{-th output}} \right)$$

[0048] 이러한 본 발명의 바람직한 실시예에 따른 블라인드 소스 분리 방법과 일반적인 FD 방식 ICA, IVA 및 FF ICA 사이의 비교를 위하여 본 출원인은 세가지의 상이한 환경에서 두 개의 소스 신호를 혼합하고, 두 개의 마이크를 통해 두 개의 혼합 신호들을 획득하여 소스 신호들을 분리하여 그에 대한 SIR를 검출하였다.

[0049] 도 2는 상기 세가지의 상이한 환경을 도시한 것으로, 제1환경은 제1 및 제2소스(204,206)가 제1 및 제2마이크(200,202)에 대해 방향이 서로 다르며 이격 거리가 가깝게 위치한 경우이고, 제2환경은 제3 및 제4소스(208,210)가 제1 및 제2마이크(200,202)에 대해 방향이 비슷하며 이격 거리가 가깝게 위치한 경우이고, 제3환경은 제5 및 제6소스(212,214)가 제1 및 제2마이크(200,202)에 대해 방향이 비슷하며 이격 거리가 멀게 위치한 경

우이다. 그리고 상기 제1 내지 제6소스(204~214)로 출력되는 소스 신호들은 TIMIT 데이터베이스에서 두 남자 화자와 두 여자 화자에 의해 말해진 문장들을 연결하여 구성하였다. 그리고 각 소스 신호는 16kHz의 샘플링 레이트로 샘플링되며 8-s 길이이다. 상기 혼합 신호의 SIR은 -2.5dB에서 2.7dB이다. 단구간 푸리에 분석에 있어 한 프레임의 윈도우 길이는 FF 언믹싱 필터 네트워크가 적용된 알고리즘들(FF ICA, prop. FF IVA)은 512샘플, 그렇지 않은 알고리즘들(conv. ICA, conv. IVA)은 2048샘플을 사용했다. 그리고 각 방법의 최적 스택 크기는 다양한 실험으로부터 확정된다.

[0050] 상기 세가지 환경에 대한 실험 분석 결과를 도 3에 도시한 표 1 내지 표 3에 나타내었다.

[0051] 먼저 제1환경에 대한 실험 분석 결과를 표 1을 참조하여 설명한다. 상기 표 1은 반향시간 RT_{60} 을 0.2초, 0.4초, 0.6초로 변경하면서 본 발명에 따른 블라인드 소스 분리 방법(prop. FF IVA)과 기존의 블라인드 소스 분리 방법들(conv. ICA, conv. IVA, FF ICA)에 따라 소스 신호들을 분리하고, 분리된 소스 신호들에 대한 SIR을 측정하여 기록한 것이다. 상기 표 1을 참조하면, 본 발명에 따른 블라인드 소스 분리 방법(prop. FF IVA)은 기존의 블라인드 소스 분리 방법들(conv. ICA, conv. IVA, FF ICA)보다 SIR이 향상되었으며, 특히 반향시간이 긴 경우에도 SIR이 9.16dB로 높게 검출되었다.

[0052] 그리고 제2환경에 대한 실험 분석 결과를 표 2를 참조하여 설명한다. 상기 표 2는 반향시간 RT_{60} 을 0.2초, 0.4초, 0.6초로 변경하면서 본 발명에 따른 블라인드 소스 분리 방법(prop. FF IVA)과 기존의 블라인드 소스 분리 방법들(conv. ICA, conv. IVA, FF ICA)에 따라 소스 신호들을 분리하고, 분리된 소스 신호들에 대한 SIR을 측정하여 기록한 것이다. 상기 표 2를 참조하면, 본 발명에 따른 블라인드 소스 분리 방법(prop. FF IVA)은 기존의 블라인드 소스 분리 방법들(conv. ICA, conv. IVA, FF ICA)보다 SIR이 향상되었으며, 특히 반향시간이 긴 경우에도 SIR이 8.44dB로 높게 검출되었다.

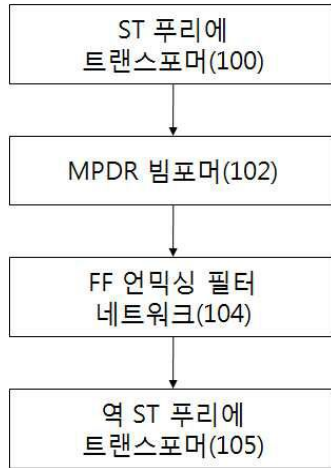
[0053] 그리고 제3환경에 대한 실험 분석 결과를 표 3을 참조하여 설명한다. 상기 표 3은 반향시간 RT_{60} 을 0.2초, 0.4초, 0.6초로 변경하면서 본 발명에 따른 블라인드 소스 분리 방법(prop. FF IVA)과 기존의 블라인드 소스 분리 방법들(conv. ICA, conv. IVA, FF ICA)에 따라 소스 신호들로 분리하고, 분리된 소스 신호들에 대한 SIR을 측정하여 기록한 것이다. 상기 표 3을 참조하면, 본 발명에 따른 블라인드 소스 분리 방법(prop. FF IVA)은 기존의 블라인드 소스 분리 방법들(conv. ICA, conv. IVA, FF ICA)보다 SIR이 향상되었으며, 특히 반향시간이 긴 경우에도 SIR이 7.58dB로 높게 검출되었다.

부호의 설명

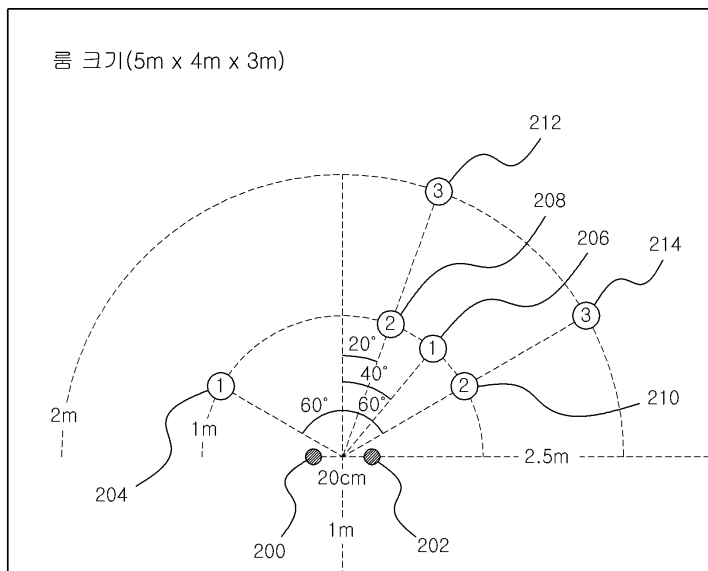
- [0054]
- 100 : ST 푸리에 트랜스포머
 - 102 : MPDR 빔 포머
 - 104 : FF 언믹싱 필터 네트워크
 - 105 : 역 ST 푸리에 트랜스포머

도면

도면1



도면2



도면3

표 1(dB)

RT_{60}	conv. ICA	conv. IVA	FF ICA	prop. FF IVA
0.2초	17.97	17.44	21.28	22.10
0.4초	10.56	11.27	11.87	12.79
0.6초	7.87	8.68	6.59	9.16

표 2(dB)

RT_{60}	conv. ICA	conv. IVA	FF ICA	prop. FF IVA
0.2초	16.70	16.04	21.59	22.26
0.4초	5.87	6.66	10.77	12.68
0.6초	3.48	4.02	6.81	8.44

표 3(dB)

RT_{60}	conv. ICA	conv. IVA	FF ICA	prop. FF IVA
0.2초	10.08	10.49	13.75	14.08
0.4초	3.87	4.80	7.10	9.23
0.6초	2.57	3.52	4.41	7.58