

Blind Identification of Noisy Non-Stationary Sources Using a Binary Mask

J. Villavicencio, J. Álvarez, and L. Márquez

Abstract—In most functional blind source separation (BSS) applications, the observations contain additive noise that limits the performance of most existing BSS algorithms, especially in the case where the noise is not modeled by a random process (e.g., electromagnetic power supply noise). In this paper, we describe an algorithm where a new cost function is fed with a frequency profile of the noise. In this way, the coefficients of the separating matrix are identified without the bias introduced by the noise. The proposed cost function is based on a frequency domain binary mask and the coherence function. The binary mask selects the frequencies where the signal-to-noise ratio (SNR) is relatively high, while the coherence is minimized to obtain the inverse system. Moreover, any frequency noise profile, whether given a priori or estimated, can be applied to the binary mask to achieve the identification of the inverse matrix. Computer simulations show that the proposed algorithm exhibits better performance under different SNR scenarios compared to methods developed previously.

Index Terms—Blind source separation, Noise sources, Second order statistics, Binary mask.

I. INTRODUCCIÓN

EL problema de separación ciega de fuentes (BSS) consiste en recuperar señales de fuentes ocultas que han sido mezcladas por un sistema desconocido. Las fuentes, que tienen la propiedad de ser estadísticamente independientes, pueden ser recuperadas o extraídas por los métodos de Análisis de Componentes Independientes (ICA) o por los enfoques de Estadísticas de Segundo Orden (SOS). ICA utiliza estadísticas de orden superior (HOS) como criterio de optimización para lograr la separación de las fuentes independientes e idénticamente distribuidas (iid), pero las cuales forzosamente deben tener una distribución de probabilidad no gaussiana [1]. Los métodos SOS suponen que las fuentes, probablemente gaussianas pero coloreadas, pueden mostrar alguna estructura temporal que les permita caracterizarse por una o varias matrices de covarianza retardadas en el tiempo, lo que conlleva a utilizar métodos de identificación basados en técnicas de eigendescomposición [2].

El problema de BSS es de gran interés porque tiene una amplia gama de aplicaciones que van desde aplicaciones biomédicas hasta audio [3]–[6]. Dentro del área biomédica existen aplicaciones que requieren del procesamiento de señales obtenidas de electroencefalogramas (EEG) [7], [8], electrocardiogramas (ECG) [9], [10], electromiogramas (EMG), etc.

J. R. Villavicencio, Departamento de Ingeniería Eléctrica y Electrónica, Instituto Tecnológico de Sonora, (e-mail: jesus.villavicencio@itson.edu.mx; rigoberto.villavicencio@gmail.com).

Jq. Álvarez and J. L.A. Márquez are with Departamento de Electrónica y Telecomunicaciones, Centro de Investigación Científica y de Educación Superior de Ensenada (e-mail: jqalvar@cicese.edu.mx; lmarquez@cicese.edu.mx).

Todas ellas, son señales compuestas que pueden ser generadas por decenas (en el caso de las EMG) y hasta millones (en el caso de las EEG) de fuentes biofísicas [11]. El interés de la BSS es reducir la redundancia en las señales fisiológicas de manera que se pueda observar, estudiar o encontrar patrones que correspondan a comportamientos cognitivos [12], o con el fin de facilitar la obtención de diagnósticos en evaluaciones médicas preoperatorias [13], por mencionar algunas.

En [13] se realizó una comparación entre varios métodos de BSS en conjunto con un método basado en ondeletas para la supresión de artefactos causados por la estimulación eléctrica en la corteza cerebral durante la captura del EEG en pacientes con epilepsia. En [12] se analizaron los EEG de niños con diferentes habilidades matemáticas para encontrar patrones en el EEG que tengan relación con el nivel de proceso de aprendizaje del infante. En el estudio realizado en [12] se utilizó el algoritmo desarrollado en [5] para preprocesar las EEG adquiridas. Básicamente, el algoritmo se encarga de separar las fuentes utilizando la BSS y remover artefactos/ruido mediante un método no lineal. Estos ejemplos, no son una tarea sencilla para el BSS debido a que el número de fuentes es mucho mayor que el número de sensores utilizados en el problema, y esto debilita las suposiciones sobre las cuales la metodología de la BSS tiene fundamento.

Al igual que en el ejemplo anterior, existen muchas aplicaciones de la BSS en el mundo real, donde el ruido añadido puede deberse a factores que no pueden controlarse dentro de las observaciones, tal como el ruido ambiental, los fenómenos electromagnéticos, ruido de cuantización del convertidor A/D, etcétera. Estos eventos, los cuales no necesariamente podrían ser modelados por ruido blanco gaussiano, debilitan la suposición sobre el modelo del problema de BSS. Esto es, el modelo de la BSS podría reescribirse como un sistema de ecuaciones indeterminadas donde el número de fuentes, que ahora incluye al ruido, es mayor que el número de sensores. Este caso puede llevar a realizar una estimación sesgada del sistema inverso degradando el desempeño considerablemente de las metodologías BSS. Sin embargo, los resultados de [5], que consisten en un comparativo entre métodos de BSS para separar las fuentes biofísicas, y un método (no lineal) basado en ondeletas para eliminar ruido y artefactos, mostraron que es preferible implementar primero la BSS para tener una estimación ruidosa de las fuentes y después aplicar el algoritmo de eliminación de ruido/artefactos. Realizar los pasos en orden opuesto, es decir, primero eliminar (de manera no lineal) el ruido de las observaciones para posteriormente tratar de separar las fuentes conduce a una disminución en el rendimiento de la extracción de las señales por el sesgo en la

estimación del sistema inverso, tal como se puede ver en los resultados experimentales de [5]. Debido a esto, es primordial la estimación del sistema inverso sin el sesgo que pueda ser introducido por el ruido aditivo. Por lo tanto, la idea principal del método propuesto en este artículo es obtener un perfil del ruido en el dominio de la frecuencia, para después seleccionar los componentes frecuenciales de las observaciones que tienen una SNR alta para el cálculo de la función de costo. De esta manera, considerando los *bins* de frecuencia donde el ruido no es fuerte, podemos recuperar la matriz inversa, y por consiguiente extraer las fuentes. Las principales contribuciones de este trabajo son las siguientes.

- Proponemos una nueva función de costo la cual es una modificación de la *Magnitude-Squared Coherence* (MSC). La nueva función de costo usa una máscara binaria en el dominio de la frecuencia (FBM) que mejora la estimación de la matriz inversa en condiciones ruidosas (ver Sección 3). En el Apéndice, mostramos por qué la función de coherencia no puede separar las mezclas con ruido, y discutimos la justificación de la función de costo propuesta.
- Proponemos un método para crear una máscara binaria en el dominio de frecuencia para la función de costo propuesta (Sección 3). El umbral del nivel de ruido para la creación de la FBM es seleccionado por λ . Para ilustrar la efectividad del enfoque propuesto, realizamos treinta simulaciones (Sección 4).
- La FBM propuesta en la función de costo puede generarse a priori, si se dispone de información del ancho de banda de interés de las señales originales, o se puede estimar mediante métodos de estimación del ruido comúnmente utilizados en *speech enhancement*. Una discusión de esto se realiza en la Sección 5.

La MSC ha sido considerada previamente en el problema de BSS [14]–[16]; sin embargo, la suma de la MSC no puede ser utilizada *per se* como una función de costo en la BSS de fuentes ruidosas (ver Apéndice). Además, el uso de una BM en el contexto del BSS no es nuevo. En [17] se utilizó una BM en el problema de BSS de mezclas convolutivas en el que se recuperaron múltiples fuentes a partir de dos observaciones anecoicas. Los resultados mostrados por [17] sugieren que la extracción de múltiples fuentes a partir de dos observaciones puede lograrse si y sólo si las representaciones de tiempo-frecuencia (T-F) de todas las fuentes no se superponen. Sin embargo, en el presente trabajo, esta suposición no se cumple. En otro trabajo de señales biomédicas [10], se muestra una aplicación del presente algoritmo para la extracción de pulso cardíaco fetal a partir de ECG cuando el número de sensores en la zona abdominal es limitado.

II. FORMULACIÓN DEL PROBLEMA

La BSS de mezclas instantáneas tiene muchas aplicaciones [5], [17]–[19], en donde el proceso de mezclado puede ser aproximado por el siguiente modelo:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \quad t = 1, 2, \dots \quad (1)$$

donde el vector de observación, $\mathbf{x}(t) \in \mathbb{R}^n$, se obtiene de las salidas de múltiples sensores que capturaron las mezclas de las

fuentes $\mathbf{s}(t) \in \mathbb{R}^m$, contaminadas por ruido blanco gaussiano $\mathbf{n}(t) \in \mathbb{R}^n$. En este caso particular, $\mathbf{A} \in \mathbb{R}^{n \times m}$ es la matriz que modela el proceso de mezclado de las fuentes, y t es el índice de tiempo. Por simplicidad, en el resto del artículo se supondrá igual número de fuentes y sensores ($n = m$).

El objetivo de la BSS es encontrar una matriz \mathbf{W} de dimensión $m \times m$ a partir de los datos observados con la mínima o nula influencia del ruido aditivo $\mathbf{n}(t)$. Las fuentes estimadas están descritas por la expresión:

$$\hat{\mathbf{s}}(t) = \mathbf{W}\mathbf{x}(t) = \mathbf{W}\mathbf{A}\mathbf{s}(t) + \mathbf{W}\mathbf{n}(t), \quad t = 1, 2, \dots \quad (2)$$

donde $\hat{\mathbf{s}}(t) \in \mathbb{R}^m$ es una estimación ruidosa de las fuentes $\mathbf{s}(t)$, y \mathbf{W} es la matriz de separación o el sistema inverso. Es bien conocido en el contexto de la BSS que \mathbf{W} puede ser descompuesta como $\mathbf{W} = \mathbf{D}\mathbf{P}\mathbf{A}^{-1}$, donde \mathbf{P} es una matriz de permutación, y \mathbf{D} es una matriz diagonal. Las matrices \mathbf{P} y \mathbf{D} se agregan para tener en cuenta las ambigüedades de escalamiento, signo y permutación de la BSS.

La ecuación (2) puede ser reescrita de la siguiente forma:

$$\hat{\mathbf{s}}(t) = \mathbf{W}\mathbf{x}(t) = \mathbf{W}\mathbf{A}\mathbf{s}(t) + \mathbf{W}\mathbf{n}(t), \quad t = 1, 2, \dots \quad (3)$$

donde $\mathbf{G} = \mathbf{W}\mathbf{A}$ es la matriz diagonal global que combina el mezclado y separación del sistema. La idea principal del método propuesto es usar el rango de frecuencias o los *bins* de frecuencia que no son afectados por el ruido para lograr una estimación de la matriz \mathbf{W} .

En el resto del artículo se consideran las siguientes suposiciones:

- A1. \mathbf{A} es una matriz de rango pleno m .
- A2. Las fuentes $\mathbf{s}(t)$ son no estacionarias, tienen media cero, están mutuamente descorrelacionadas, y están correlacionadas en el tiempo.
- A3. $\mathbf{n}(t)$ es el vector de ruido blanco gaussiano, el cual es estadísticamente independiente de las fuentes.
- A4. Existen algunas frecuencias de las observaciones en las que el ruido es relativamente débil en comparación con las energías que proporcionan las fuentes, es decir, existe f_j tal que, $P_{s_1 s_1}(f_j), \dots, P_{s_m s_m}(f_j) \gg P_{nn}(f_j)$, donde $P_{s_1 s_1}(f_j), \dots, P_{s_m s_m}(f_j)$ and $P_{nn}(f_j)$ son las densidades espectrales de potencia (PSD) de las m fuentes y del ruido.

III. MÉTODO

La función de costo propuesta para encontrar la matriz \mathbf{W} está basada en la MSC. La MSC es una función normalizada en el dominio de frecuencia que tiene muchas aplicaciones útiles, y está dada por:

$$MSC = |\gamma_{\hat{s}_1 \hat{s}_2}(f)|^2 = \frac{|P_{\hat{s}_1 \hat{s}_2}(f)|^2}{P_{\hat{s}_1 \hat{s}_1}(f)P_{\hat{s}_2 \hat{s}_2}(f)}, \quad (4)$$

donde $P_{\hat{s}_1 \hat{s}_2}(f)$ es la densidad espectral de potencia cruzada entre dos señales $\hat{s}_1(t)$ y $\hat{s}_2(t)$. $P_{\hat{s}_1 \hat{s}_1}(f)$ y $P_{\hat{s}_2 \hat{s}_2}(f)$ son las respectivas densidades espectrales de potencia de las señales $\hat{s}_1(t)$ y $\hat{s}_2(t)$. La coherencia está estrechamente relacionada con la estimación de la SNR, ver [20]. La MSC es una medida de la relación lineal entre dos señales, cuyo valor está entre 0 y 1. Un valor de 0 para un bin de frecuencia f significa que

las dos señales están descorrelacionadas para esa frecuencia en particular [21], mientras que un valor de 1 indica que las dos señales están totalmente correlacionadas para ese bin de frecuencia en particular. Por lo tanto, la coherencia es mínima en señales descorrelacionadas. Como se muestra en [14], se puede crear una función de costo para más de dos fuentes si utilizamos la suma de la MSC entre todas las combinaciones posibles (l, k) de las señales $s_{l,k}$, donde $1 \leq l < k \leq m$, $l < k$ y m el número de fuentes.

A. Máscara Binaria de Frecuencia y Selección del Criterio Local para la Estimación de Ruido

La MSC *per se* no es adecuada para la BSS de mezclas ruidosas instantáneas (ver Apéndice A). Hemos añadido una FMB que permite encontrar la matriz \mathbf{W} sin influencia del ruido blanco gaussiano aditivo. La idea de mejorar la estimación del sistema inverso proviene de algunas obras de reconocimiento de voz [22], [23], donde se hace la suposición de que el ruido aditivo enmascara algunas partes de la representación T-F de la señal vocal y deja las otras partes sin afectar. Los resultados en [23] han demostrado que se puede lograr una mejora significativa en el reconocimiento de la voz bajo condiciones ruidosas cuando sólo se emplean los coeficientes fiables de la representación T-F. La selección de una unidad T-F fiable se basa en el umbral del SNR a posteriori que está estrechamente relacionado con la SNR a priori, ver [23]. Usando la misma idea que la descrita en el párrafo anterior, proponemos una máscara binaria (BM) en el dominio de la frecuencia, denotada por $\varphi(f)$. La BM se puede simplificar a una función en el dominio de la frecuencia porque el efecto de promediación sobre el ruido blanco hace que el espectro de potencia tenga una distribución espectral uniforme que conduce a una potencia constante sobre todas las frecuencias, esto es, $P_{nn}(f) = c$, donde c es una constante. Entonces, suponemos que cualquier señal que supere el umbral del ruido pertenecerá a la potencia de las fuentes no-estacionarias. Debido a esto, se propone una estimación aproximada del espectro de potencia del ruido calculando la media de la PSD del vector de observación, denotada por $\mu_{P_{xx}}$, donde $\mu_{P_{xx}} \geq \mu_{P_{nn}}$. El umbral para la creación de la máscara binaria es seleccionado por la siguiente expresión:

$$\psi = \mu_{P_{xx}} + \lambda \sigma_{P_{xx}}, \quad (5)$$

donde $\mu_{P_{xx}}$ y $\sigma_{P_{xx}}$ son la media y la desviación estándar del periodograma de las observaciones, respectivamente. λ es una constante que se utiliza para ajustar el nivel del umbral.

De este modo, el umbral se basa en la desviación estándar de la estimación aproximada del ruido. Se propone controlar el nivel de umbral a través del parámetro. La FBM propuesta está dada por:

$$\varphi(f) = \begin{cases} 1 & \text{if } P_{xx}(f) \geq \psi \\ 0 & \text{cualquier otro caso.} \end{cases} \quad (6)$$

La función de costo propuesta, es una combinación de coherencia y una máscara binaria está dada por:

$$\mathbf{J} = \sum_{l,k} \sum_{f=0}^p |\gamma_{\hat{s}_l \hat{s}_k}(f)|^2 \varphi(f), \quad (7)$$

Algoritmo 1 nMSC-BSS

Require: Vector de observación \mathbf{x}

Ensure: Fuentes estimadas $\hat{\mathbf{s}}$

- 1: Remueve la media de las observaciones.
- 2: Calcula la máscara binaria, $\varphi(f)$.
- 3: $\mathbf{W} = \mathbf{I}$
- 4: **procedure** RECOCIDO SIMULADO
- 5: **while** no se alcance el número máximo de iteraciones **do**
- 6: Genera una solución, \mathbf{W} , según un programa de temperatura.
- 7: Genera $\hat{\mathbf{s}}$, ecuación (2).
- 8: Calcula la función de costo \mathbf{J} , ecuación (7).
- 9: La solución se acepta o rechaza de acuerdo a un programa de temperatura (función de programación de temperatura de Matlab; temperatureboltz).
- 10: **end while**
- 11: **end procedure**
- 12: \mathbf{W} is the best solution found by the simulated annealing.
- 13: Estimate sources $\hat{\mathbf{s}}$, eq. (2).

donde p es el *bin* de la frecuencia de Nyquist, $\varphi(f)$ es una BM en función de la frecuencia y es estimada de acuerdo con (6). Finalmente, la función de costo está representada por la ecuación (7), la cual es optimizada mediante el algoritmo recocido simulado [24]. El pseudo- algoritmo de la metodología propuesta se muestra en el Algoritmo 1.

IV. RESULTADOS EXPERIMENTALES

En esta sección presentamos los resultados obtenidos con la metodología propuesta. En la subsección 4.1 se muestra un ejemplo de los resultados obtenidos con el algoritmo propuesto en una sola ejecución. En la subsección 4.2 se muestra una comparación de la metodología propuesta con algunos algoritmos SOS BSS existentes, SOBI [25], SOBI-RO [26], y SOBI-BPF [27] (los métodos que utilizan funciones de costo basadas en estadísticas de orden superior no son adecuados para extraer fuentes contaminadas con ruido blanco gaussiano). El método propuesto fue implementado en el toolbox ICALAB [28]. El índice de razón señal-a-interferencia (SIR) [29] se utilizó como medida de desempeño.

$$SIR = -20 \log \|g_i - u_i\|, \quad (8)$$

donde g_i es el i -ésimo renglón de la matriz global normalizada \mathbf{G} , y u_i es el i -ésimo renglón de la matriz identidad. Una SIR superior a 16 dB indica una buena capacidad de separación [29]. La función de coherencia se calculó mediante el método *Welch* con un 50 % de solapamiento entre segmentos, y los segmentos solapados se multiplicaron por la función de *Hamming*. Todas las pruebas se realizaron usando $n_d = 8$ segmentos (número de ventanas) para estimar la coherencia. n_d es también el número de segmentos promediados utilizados en el cálculo de $P_{xx}(f)$ mediante el método de *Welch*. Definimos la longitud de cada segmento de tiempo promediado como $p = 2\alpha$ donde $\alpha = \text{floor}(\log_2(\beta))$, y $\beta = \text{length}(x(t))/n_d$.

El ruido blanco gaussiano se agregó a las señales observadas en una relación de 0, 10 y 20 dB. El nivel de ruido está definido como:

$$SNR_i = 10 \log_{10} \frac{\text{var}(x_i)}{\text{var}(n_i)}, \quad (9)$$

donde $\text{var}(x_i)$ es la varianza de la observación del i -ésimo sensor sin el ruido blanco gaussiano aditivo, y $\text{var}(n_i)$ es la varianza del ruido blanco gaussiano aditivo en el i -ésimo sensor, donde $i = 1, 2, \dots, m$. En los experimentos se usaron los siguientes parámetros en la función `Recocido Simulado` (`simulannealbnd`) de Matlab; `TemperatureFcn`: `@temperatureboltz`; `InitialTemperature`: 0.15; `ReannealInterval`: 500.

A. Ejemplo Ilustrativo

Considere un conjunto de tres sensores que reciben una mezcla de tres señales independientes desconocidas no estacionarias con media cero. Estas tres señales provienen de la biblioteca de audio de Harvard, $s_1(t)$ (hombre 1: ‘*The child crawled into the dense grass*’), $s_2(t)$ (hombre 2: ‘*The horse trotted around the field at a brisk pace*’), y $s_3(t)$ (hombre 3: ‘*Paint the sockets in the wall dull green*’). Las señales se muestran en la Fig. 1. La matriz de mezclado \mathbf{A} creada de manera aleatoria está dada por

$$\mathbf{A} = \begin{bmatrix} 0.9311 & 0.2070 & 0.6351 \\ 0.1718 & 0.0710 & 0.8198 \\ 0.5451 & 0.9372 & 0.0021 \end{bmatrix}. \quad (10)$$

El ruido blanco gaussiano se agregó a la mezcla a una razón de 10dB. Las mezclas con ruido se muestran en la Fig. 1. Después de 9000 iteraciones el algoritmo propuesto con $\lambda = 1$ logró obtener la siguiente matriz inversa:

$$\mathbf{W} = \begin{bmatrix} 0.7884 & -0.6040 & -0.1170 \\ -0.1671 & 0.9852 & -0.0389 \\ -0.5157 & 0.3967 & 0.7594 \end{bmatrix}. \quad (11)$$

La SIR obtenida en este ejemplo es de 40.90dB. La estimación ruidosa de las fuentes se muestra en la parte inferior de la Fig. 1.

Es importante notar que $\hat{s}(t)$ es una estimación ruidosa de las fuentes. Si se requiere una estimación sin ruido de las fuentes extraídas, posteriormente se deben aplicar otros métodos enfocados en la eliminación del ruido. Realizar estos pasos en la dirección opuesta, es decir, primero eliminar el ruido para posteriormente tratar de separar las fuentes, conduce a una disminución en el rendimiento de la extracción de las señales, tal como se muestra en los experimentos de [5]. Podemos prever que la razón principal de la reducción del rendimiento de separación se debe a que se utilizaron métodos que eliminan el ruido de forma no lineal, debilitando así las suposiciones hechas sobre el modelo de BSS. En este trabajo no se analizarán métodos de postprocesamiento para la eliminación de ruido en las fuentes recuperadas.

B. Experimentos

Se hicieron algunos experimentos para ilustrar la efectividad de la metodología propuesta con diferentes niveles del umbral de la Máscara Binaria, $\lambda = (0, 0.5, 1.5, 2, 2.5, 3)$, en varias

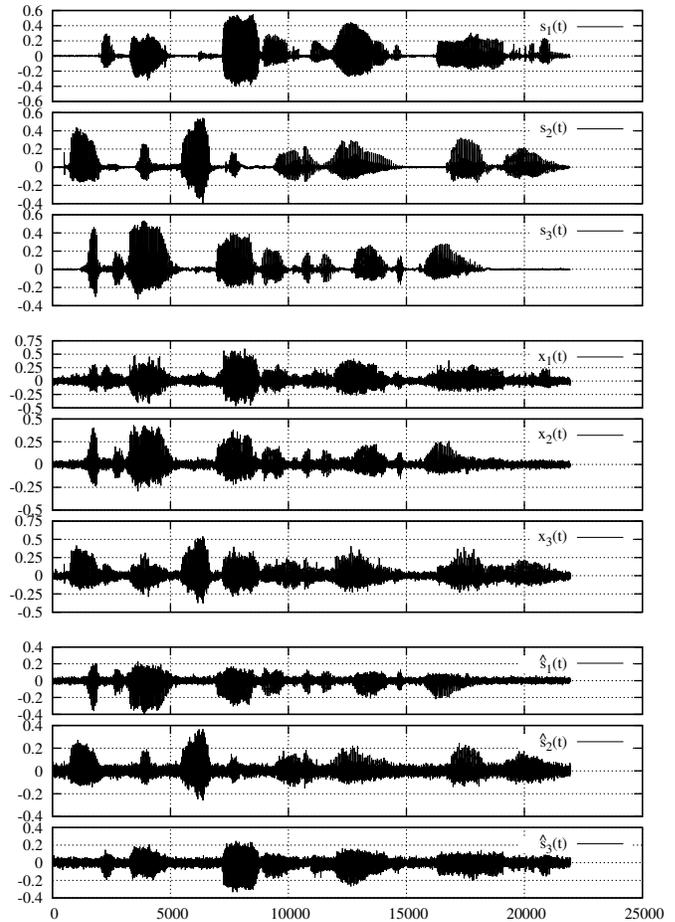


Fig. 1. De arriba a abajo: Fuentes originales $s_1(t)$, $s_2(t)$, and $s_3(t)$. Observaciones $x_1(t)$, $x_2(t)$, y $x_3(t)$. Señales recuperada $\hat{s}_1(t) \approx -s_3(t)$, $\hat{s}_2(t) \approx s_2(t)$, y $\hat{s}_3(t) \approx -s_1(t)$.

condiciones de SNR. La Tabla 1 muestra el SIR promedio obtenido en 30 pruebas. En cada prueba, la matriz de mezclado $\mathbf{A} \in \mathbb{R}^{3 \times 3}$ fue generada aleatoriamente. Las mezclas se obtuvieron mezclando las mismas fuentes utilizadas en el Ejemplo Ilustrativo. La longitud de cada fuente consistió en 21937 muestras a una frecuencia de muestreo de 8000Hz. A partir de los resultados, vemos que el método propuesto muestra el rendimiento SIR más alto en todos los entornos de SNR comparado con las otras metodologías SOS, excepto en 0 dB SNR. Los SIR más altos están marcados en negrita en la Tabla 1. En la prueba de 0 dB, la metodología propuesta tiene el segundo mejor SIR con menos de medio dB de magnitud del mejor rendimiento (SOBI-RO).

V. DISCUSIÓN

Los resultados muestran que una FBM y la MSC pueden utilizarse eficazmente para recuperar fuentes en mezclas con ruido. De la subsección Experimentos concluimos que la selección de un umbral $\lambda \in [0, 1]$ es un buen compromiso en el rendimiento SIR en ambientes de ruido moderado y alto. El rango de frecuencias seleccionadas por la FBM tendrá un SNR relativamente alto; por lo tanto, el sesgo introducido por el vector de ruido será insignificante.

TABLA I
COMPARACIÓN DEL RENDIMIENTO DE LOS RESULTADOS PARA TRES FUENTES BAJO DIFERENTES ESCENARIOS DE SNR.

Algorithm	SNR			
	0dB SIR (dB)	10dB SIR (dB)	20dB SIR (dB)	∞ SIR (dB)
SOBI	2.6	2.11	13.5	21.81
SOBI-RO	19.76	26.69	26.16	26.32
SOBI-BPF	13.26	20.14	20.78	20.19
nMSC-BSS $\varphi(f)=1$	4.57	7.17	20.13	34.95
nMSC-BSS $\lambda = 0$	12.79	23.84	33.21	34.09
nMSC-BSS $\lambda = 0.5$	14.68	27.32	30.79	32.90
nMSC-BSS $\lambda = 1$	16.68	26.98	25.73	26.54
nMSC-BSS $\lambda = 1.5$	17.84	27.62	28.45	28.92
nMSC-BSS $\lambda = 2$	18.62	25.06	24.42	23.88
nMSC-BSS $\lambda = 2.5$	18.60	25.37	24.78	24.90
nMSC-BSS $\lambda = 3$	19.33	25.17	23.20	24.58

Elijiendo un umbral con un valor bajo, $\psi \leq \mu_{P_{xx}}(f)$ conducirá a la recuperación de una matriz inversa sesgada debido a la influencia del ruido. En el caso contrario, tener un umbral alto, puede omitir información útil de algunas de las fuentes, especialmente en niveles altos de ruido, lo que lleva a la recuperación de algunas de las fuentes con mayor potencia. El valor umbral más bajo posible es $\psi = 0$, en este caso la FBM es $\varphi(f) = 1 \forall f$.

La estimación aproximada que proponemos en este trabajo del ruido blanco gaussiano está sobreestimada. Sin embargo, esto representa una excelente opción para la estimación del ruido. Además, este método se puede ampliar para trabajar con otros tipos de ruido, como lo reportados en los experimentos de *Computational Auditory Scene Analysis* (CASA), ya que siempre es posible una estimación de PSD del ruido. En este caso, prevemos que el umbral debería fijarse para cada unidad de frecuencia. Finalmente, mencionamos algunos inconvenientes del algoritmo propuesto. La metodología propuesta se basa en un algoritmo de optimización heurística conocido por su largo tiempo de procesamiento. Sin embargo, hay algunas soluciones que pueden implementarse fácilmente. Por ejemplo, la implementación del algoritmo en una estructura de programación paralela junto con un alto poder de procesamiento de las computadoras actuales puede aminorar esta desventaja considerablemente. También es importante considerar otra limitación importante de los métodos de la BSS en general. En los casos en que la matriz de mezclado está mal condicionada (matriz casi singular), lo cual puede ocurrir, por ejemplo, cuando los sensores se colocan demasiado cerca, los métodos BSS disponibles tienen un rendimiento deficiente. Éste sigue siendo un problema que debe considerarse para los métodos de separación de BSS, incluido el que presentamos en este artículo y, por lo tanto, presenta una oportunidad de investigación para el futuro.

VI. CONCLUSIONES

En este artículo se presenta una solución al problema de la separación ciega de fuentes de mezclas instantáneas con ruido. Para lograr la estimación del sistema inverso sin el sesgo introducido por el ruido, se propone una función de costo basada en la función de coherencia y una máscara binaria. Los resultados de esta investigación apoyan la idea de que

el uso de una máscara binaria es un enfoque que puede ser utilizado en el problema de BSS. Al implementar estimadores de ruido para crear la máscara binaria se espera que este método funcione para más tipos de ruido y no sólo para el ruido blanco gaussiano.

APÉNDICE A

En esta sección se justifica la afirmación de que la suma de la MSC no puede usarse para separar las fuentes de ruido por si sola. A continuación, explicamos la función de costo propuesta, que es una modificación de la suma de la MSC. Analizando la MSC entre las fuentes estimadas con ruido para $m = 2$, dadas por:

$$\begin{aligned}\hat{s}_1(t) &= g_{11}s_1(t) + g_{12}s_2(t) + w_{11}n_1(t) + w_{12}n_2(t) \\ \hat{s}_2(t) &= g_{21}s_1(t) + g_{22}s_2(t) + w_{21}n_1(t) + w_{22}n_2(t),\end{aligned}\quad (12)$$

donde ambos $n_1(t)$ y $n_2(t)$ son ruido blanco gaussiano, g_{ij} y w_{ij} son los elementos de las matrices \mathbf{G} , y \mathbf{W} , respectivamente; calculando $P_{\hat{s}_1\hat{s}_1}(f)$, $P_{\hat{s}_2\hat{s}_2}(f)$, $P_{\hat{s}_1\hat{s}_2}(f)$ y sustituyéndolos en (4), y teniendo en cuenta que las PSD de las fuentes de ruido gaussiano blanco son iguales ($P_{n_1n_1}(f) = P_{n_2n_2}(f) = P_{nn}(f)$), y ($P_{n_1n_2}(f) = 0$), entonces la MSC produce la siguiente ecuación:

$$|\gamma_{\hat{s}_1\hat{s}_2}(f)|^2 = \frac{B_1}{B_2}, \quad (13)$$

$$B_1 := (g_{11}g_{21})^2 P_{s_1s_1}^2(f) + (g_{12}g_{22})^2 P_{s_2s_2}^2(f) + [w_{11}w_{21} + w_{12}w_{22}]^2 P_{nn}^2(f),$$

$$B_2 := [g_{11}^2 P_{s_1s_1}(f) + g_{12}^2 P_{s_2s_2}(f) + [w_{11}^2 + w_{12}^2] P_{nn}(f)][g_{21}^2 P_{s_1s_1}(f) + g_{22}^2 P_{s_2s_2}(f) + [w_{21}^2 + w_{22}^2] P_{nn}(f)].$$

La ecuación (13) implica que no se puede construir una función de costo para recuperar las fuentes porque, aunque \mathbf{G} sea diagonal, la estimación de la MSC no es cero. El término de ruido en el numerador de la ecuación introduce un sesgo en la estimación de la MSC y no permite encontrar un mínimo en el que \mathbf{G} sea una matriz diagonal. Para resolver este problema se necesitan lograr dos metas:

- 1) Hacer \mathbf{G} una matriz diagonal
- 2) Hacer la función de costo $\mathbf{J} = 0$.

A continuación, justificamos la función de costo propuesta utilizando un enfoque similar en el dominio de la frecuencia seguido por [21], [30]. Sustituyendo la coherencia entre las observaciones ruidosas (12) en la función de costo propuesta (7) tenemos:

$$|\gamma_{\hat{s}_1\hat{s}_2}(f)|^2 = \frac{B_3}{B_4}, \quad (14)$$

$$B_3 := \{(g_{11}g_{21})^2 P_{s_1s_1}^2(f) + (g_{12}g_{22})^2 P_{s_2s_2}^2(f) + [w_{11}w_{21} + w_{12}w_{22}]^2 P_{nn}^2(f)\} \varphi(f),$$

$$B_4 := [g_{11}^2 P_{s_1 s_1}(f) + g_{12}^2 P_{s_2 s_2}(f) + [w_{11}^2 + w_{12}^2] P_{nn}(f)] [g_{21}^2 P_{s_1 s_1}(f) + g_{22}^2 P_{s_2 s_2}(f) + [w_{21}^2 + w_{22}^2] P_{nn}(f)].$$

Supongamos que, debido a lo no estacionario de las fuentes, hay algunas f_j donde $P_{s_1 s_1}(f_j), P_{s_2 s_2}(f_j) \gg P_{nn}(f_j)$, es decir, hay algunas frecuencias donde el ruido es relativamente débil. La máscara binaria mantiene estas frecuencias. Por el contrario, las frecuencias donde el ruido gaussiano blanco aditivo es fuerte son eliminadas por la máscara binaria, $\varphi(f)$. Bajo estas circunstancias, se puede reducir el término de ruido en el numerador de (7). De esta manera, la ecuación (13) conduce a una función de costo que no está sesgada por el ruido:

$$|\gamma_{\hat{s}_1 \hat{s}_2}(f)|^2 \approx \frac{B_5}{B_6}, \quad (15)$$

$$B_5 := \{(g_{11}g_{21})^2 P_{s_1 s_1}^2(f) + (g_{12}g_{22})^2 P_{s_2 s_2}^2(f)\},$$

$$B_6 := [g_{11}^2 P_{s_1 s_1}(f) + g_{12}^2 P_{s_2 s_2}(f) + [w_{11}^2 + w_{12}^2] P_{nn}(f)] [g_{21}^2 P_{s_1 s_1}(f) + g_{22}^2 P_{s_2 s_2}(f) + [w_{21}^2 + w_{22}^2] P_{nn}(f)].$$

De este modo, la coherencia es mínima (cero) cuando $\mathbf{G} = \mathbf{W}\mathbf{A}$ es una matriz diagonal. El cumplimiento de estas dos condiciones implica que podemos estimar la matriz de separación \mathbf{W} .

AGRADECIMIENTOS

Los autores agradecen a PROFAPI 2017_2262.

REFERENCIAS

- [1] A. Hyvarinen, *Fast and robust fixed-point algorithms for independent component analysis*, Neural Networks, IEEE Transactions on, vol. 10, no. 3, pp. 626-634, May, 1999.
- [2] L. Tong, R. Liu, V. Soon, and Y. Huang, *Indeterminacy and identifiability of blind identification*, Circuits and Systems, IEEE Transactions on, vol. 38, no. 5, pp. 499-509, May, 1991.
- [3] P. Parathai and W. Woo, *Single-channel blind separation using L1 sparse complex non-negative matrix factorization for acoustic signals*, The Journal of the Acoustical Society of America, vol. 137, no. 1, pp. EL124-EL129, 2015.
- [4] L. Bai, B. Gao, G. Tian, W. Woo, and Y. Cheng, *Spatial and time patterns extraction of eddy current pulsed thermography using blind source separation*, Sensors Journal, IEEE, vol. 13, no. 6, pp. 20942101, 2013.
- [5] R. Vázquez, H. Perez, R. Ranta, V. Dorr, and L. Maillard, *Blind source separation, wavelet denoising and discriminant analysis for EEG artefacts and noise cancelling*, Biomedical Signal Processing and Control, Elsevier, vol. 7, no. 4, pp. 389-400, 2012.
- [6] M. Souden, S. Araki, K. Kinoshita, T. Nakatani, and H. Sawada, *A multichannel MMSE-based framework for speech source separation and noise reduction*, Audio, Speech, and Language Processing, IEEE Transactions on, vol. 21, no. 9, pp. 113-128, 2013.
- [7] L. da Silva Junior, F. Cesar, F. Rocha, and C. Thomaz, *A Combined Eye-tracking and EEG Analysis on Chess Moves*, IEEE Latin America Transactions, vol.16, no. 5, pp. 1288-1297, 2018.
- [8] F. Rocha, R. de Magalhaes Junior, T. Horta, F. Cesar, D. D'Israel, W. da Silva, and C. Thomaz, *EEG acquisition and processing for cognitive brain mapping during chess problem solving*, IEEE Latin America Transactions, vol. 14, no. 3, pp. 1129-1134, 2016.
- [9] A. Martinez, F. Pfrimer, M. Costa, and A. Nakano, *How to Develop a Single Channel Electrocardiograph with a Low Budget*, IEEE Latin America Transactions, vol. 16, no. 4, pp. 10571063, 2018.
- [10] J. Villavicencio, Jq. Álvarez, L. Márquez, O. Peñaloza, R. Borbón, and J. Barraza, *Separación de Electrocardiograma Materno para la Extracción de Electrocardiograma Fetal Basado en Separación Ciega de Fuentes*, Congreso Nacional de Control Automático, 2017.
- [11] L. Mesin, A. Holobar and R. Merletti, *Blind Source Separation: Application to Biomedical Signals*, 2018.
- [12] A. González, F. Gómez, R. Romo, H. Vélez, R. Salido, A. Espinoza and A. Martínez, *Dataset on the EEG time-frequency representation in children with different levels of mathematical achievement*, Data in brief, vol.21, pp. 1071-1075, 2018.
- [13] A. Bhattacharyya, R. Ranta, S. Le Cam, V. Dorr, L. Tyvaert, S. Colnat and R. Pachori, *A multi-channel approach for cortical stimulation artefact suppression in depth EEG signals using time-frequency and spatial filtering*, IEEE Transactions on Biomedical Engineering, 2018.
- [14] C. Fancourt and L. Parra, *The Coherence Function In Blind Source Separation Of Convolutional Mixtures Of Non-Stationary Signals*, In Proc. NNSP, pp. 303-312, 2001.
- [15] T. Oku and A. Sano, *Nonlinear blind source separation using coherence function*, SICE 2003 Annual Conference, vol. 3, pp. 2550-2560, Ago, 2003.
- [16] J. Villavicencio, L. Márquez, and Jq. Álvarez, *A heuristic approach for blind source separation of instant mixtures*, Computación y Sistemas, vol. 18, no. 4, pp. 719-730, 2014.
- [17] O. Yilmaz and S. Rickard, *Blind separation of speech mixtures via timefrequency masking*, IEEE Transactions on signal processing, vol. 52, no. 7, pp. 1830-1847, 2004.
- [18] D. Arnaud, P. Jason, J. Onton, R. Oostenveld, and S. Makeig, *Independent EEG sources are dipolar*, PloS one, vol. 7, no. 2, pp. e30135, 2012.
- [19] Y. Sevim and A. Atasoy, *Performance evaluation of nonparametric ICA algorithm for fetal ECG extraction*, Turkish Journal of Electrical Engineering & Computer Sciences, vol. 19, no. 4, pp. 657-666, 2011.
- [20] M. Foster, N. Guinzy, *The coefficient of coherence: its estimation and use in geophysical data processing*, Geophysics, vol. 32, no. 4, pp. 602-616, 1967.
- [21] G. Carter, *Estimation of the Magnitude-Squared Coherence Function (Spectrum)*, Naval Underwater Systems Center Newport RI, New London, Rep. AD0743945, 1972.
- [22] U. Kjems, J. Boldt, M. S. Pederson, T. Lunner, and D. Wang, *Role of mask pattern in intelligibility of ideal binary-masked noisy speech*, The Journal of the Acoustical Society of America, vol. 126, no. 3, pp. 14151426, 2009.
- [23] P. Renevey and S. Drygajlo, *Detection of reliable features for speech recognition in noisy conditions using a statistical criterion*, in Proceedings of Workshop on CRAC, pp. 71-74, 2001.
- [24] S. Kirkpatrick, C. Gelatt, and M. Vecchi, *Optimization by Simulated Annealing*, Science Vol. 220, no. 4598, pp. 671-680, 1983, doi: 10.1126/science.220.4598.671.
- [25] A. Belouchrani, K. Meraim, J. Cardoso, and E. Moulines, *A blind source separation technique using second-order statistics*, Signal Processing, IEEE Transactions on, vol. 45, no. 2, pp. 434-444, Feb, 1997.
- [26] S. Choi and A. Cichocki, *Blind separation of nonstationary sources in noisy mixtures*, Electronics Letters, vol. 36, no. 9, pp. 848-849, Abr, 2000.
- [27] R. Gharieb and A. Cichocki, *Second-order statistics based blind source separation using a bank of subband filters*, Digital Signal Processing, vol. 13, no. 2, pp. 252-274, 2003.
- [28] A. Cichocki, S. Amari, and K. Siwek, *ICALAB Tolboxes*, [Online]. Available <http://www.bsp.brain.riken.jp/ICALAB>.
- [29] A. Cichocki and S. Amari, *Adaptive blind signal and image processing: learning algorithms and applications*, 1 ed., John Wiley & Sons, 2002.
- [30] E. Weinstein, M. Feder, and A. Oppenheim, *Multi-channel signal separation by decorrelation*, IEEE transactions on Speech and Audio Processing, vol. 1, no. 4, pp. 405-413, 1993.



J.R. Villavicencio Obtuvo el grado de Ingeniero en Electrónica por la Universidad Autónoma de Baja California (México) en 2001. En el 2005 recibió el grado de Maestro en Ciencias por el Centro de Investigación Científica y Educación Superior de Ensenada (CICESE). En el año 2014 recibió el grado de Doctor en Ciencias por el CICESE. Actualmente es Profesor Investigador Auxiliar del Instituto Tecnológico de Sonora. Sus intereses de investigación son el procesamiento digital de señales y el aprendizaje máquina.



Jq. Álvarez Obtuvo el grado de ingeniería en Electrónica y telecomunicaciones en el Instituto Politécnico Nacional (IPN) de México (1975); el grado de Maestría en Ciencias en Ingeniería Eléctrica por el Centro de Investigación y de Estudios Avanzados (CINVESTAV) del IPN (1976), y Doctor en Ingeniería en control automático por la National Polytechnique Institute of Grenoble, Francia (1979). Actualmente es profesor titular del Departamento de Electrónica y Telecomunicaciones del Centro de Investigaciones Científicas y Estudios

Avanzados de Ensenada, México (CICESE). Es editor asociado de algunas revistas científicas, miembro del consejo de editores asociados de varios congresos y de otros eventos científicos patrocinados por la IFAC. El profesor Álvarez ha sido miembro del Sistema Nacional de Investigaciones Científicas de México desde 1984, nivel III en la actualidad. Además, es miembro de la Academia Mexicana de Ciencias, la Academia Mexicana de Ingeniería, y la Asociación Mexicana de Control Automático. Sus intereses de investigación se centran en los campos de control no lineal, sistemas discontinuos, control del caos y sincronización.



L. A. Márquez Obtuvo el grado de Ingeniero en electrónica en la Universidad Autónoma de Baja California (México) en 1992, Maestro en Ciencias en el área de electrónica y telecomunicación: CICESE 1994 (México), Doctorado en Ciencias: Université de Nantes/Ecole Centrale de Nantes, 2000 (Francia). Dos veces investigador invitado en UTIA Lab (Prague, CZ), varias veces en IRCCyN (Nantes, Francia), Un pos-doctorado de 6 meses en la Universidad de Pretoria. Trabaja en el Departamento de Electrónica y Telecomunicaciones en el Centro de

Investigación del CICESE. Sus intereses de investigación incluyen: sistemas no lineales, sistemas con retardo temporal, sistemas de álgebra computacional, procesamiento digital de señales y sistemas mecánicos.