



**UNED**

**UNIVERSIDAD NACIONAL DE EDUCACIÓN A DISTANCIA**

**MÁSTER EN INGENIERÍA DE SISTEMAS Y DE CONTROL**

**TRABAJO FIN DE MÁSTER**



**SEÑALES MULTICANAL:  
ANÁLISIS Y FILTRADO DE  
ARTEFACTOS**

Presentado por:

**Luis Alberto Ramon Surutusa**

Director:

**D. Ángel R. Mur Güerri**

Tutoras:

**Prof. Raquel Dormido y Natividad Duro**

Curso académico 2015 / 2016

## ÍNDICE

---

ÍNDICE.....	2
INTRODUCCIÓN .....	3
1. OBJETIVOS .....	1
2. MOTIVACIÓN Y ORIGEN .....	2
2.1. EEG .....	2
2.2. Artefactos.....	3
2.2.1 Artefactos del entorno .....	3
2.2.2 Artefactos biológicos.....	4
3. ICA y DETECT .....	6
3.1. Separación ciega de fuentes. ....	6
3.2. Ambigüedad de ICA.....	9
3.3. Independencia estadística.....	9
3.4. Algoritmo FastICA.....	9
3.5. Algoritmo JADE.....	11
3.6. Aplicación de ICA en EEG.....	11
3.7. DETECT – Detección e identificación de eventos .....	12
3.7.1. Aplicación de DETECT para etiquetar eventos en EEG.....	13
4. DESARROLLO Y RESULTADOS DEL PROYECTO.....	14
4.1. SOLUCIÓN PROPUESTA .....	14
4.2. VALIDACIÓN EXPERIMENTAL Y RESULTADOS .....	16
4.2.1 Caso 1: separación de señales simuladas mediante algoritmo FastICA.....	16
4.2.2 Caso 2: filtrado de “artefacto” en señales simuladas mediante FastICA.....	18
4.2.3 Caso 3: filtrado de “artefacto” en señales simuladas mediante JADE. ....	21
4.2.4 Caso 4: Medición de calidad del filtrado de ambos algoritmos por comparativa de sus Errores Cuadráticos Medios (EQM) a partir de una señal de control. ....	23
4.2.5 Caso 5: Filtrado automático mediante DETECT e ICA .....	24
4.2.6 Caso 6: Filtrado automático con supresión de varias CI .....	29
4.2.7 Caso 7: Evolución del filtrado automático con supresión de varias CI.....	35
4.2.8 Caso 8: Filtrado iterativo para artefactos persistentes.....	45
5. CONCLUSIONES Y LÍNEAS FUTURAS .....	53
BIBLIOGRAFÍA.....	55

## INTRODUCCIÓN

Se entiende por sistema un grupo de objetos que se combinan para lograr un determinado objetivo. Una señal es un suceso que sirve para iniciar una acción. Los conceptos de señales y sistemas están presentes en disciplinas tan diversas como sismología, comunicaciones, acústica, sistemas de generación y distribución de energía, ingeniería biomédica, etc. En estos campos la naturaleza física de las señales y de los sistemas pueden tener matices diferentes, pero todos ellos presentan características básicas comunes:

- Las señales son funciones de una o más variables independientes, y contienen información sobre la naturaleza o comportamiento de algún fenómeno.
- Los sistemas responden a señales dadas produciendo otras señales.

Normalmente los sistemas vienen descritos mediante los conceptos de señal y de respuesta (véase la Figura 1).

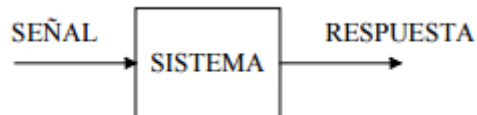


Figura 1- Diagrama general de un Sistema

Cuando en lugar de trabajar sobre una señal concreta, se trabaja en aplicaciones de un fenómeno físico donde las señales son generadas por fuentes múltiples o múltiples sensores, se habla de **señales multicanal**. Las fuentes o sensores múltiples generan señales escalares. Aunque tales señales no son magnitudes vectoriales desde un punto de vista físico, pueden ser tratadas como componentes de un vector por conveniencia notacional y matemática. Un ejemplo de señal multicanal es la salida de un electrocardiograma que tenga tres electrodos (sensores), que se puede representar como un vector  $S_3(t)$  de la forma:

$$S_3(t) = \begin{bmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \end{bmatrix}$$

Por otro lado, otro aspecto común a las señales observadas en un entorno real de cualquier tipo (ya sean señales simples o multicanal) es la aparición de **alteraciones no deseadas que modifican la señal observada y generan diferentes efectos sobre la misma**. Estas alteraciones pueden ser atenuaciones, que reducen la intensidad de la señal o de otros tipos:

- **Distorsión.** Es la alteración de la señal debido a la respuesta imperfecta del sistema a dicha señal. En el diseño de sistemas se suele considerar algún criterio de minimización de este efecto. La distorsión desaparece cuando la señal deja de aplicarse.
- **Interferencia.** Es la contaminación por señales extrañas, generalmente artificiales y de forma similar a la de la señal. El problema de interferencia se soluciona eliminando la señal interferente o su fuente.
- **Ruido.** Por ruido se entiende cualquier señal aleatoria o impredecible de tipo eléctrico, originada en forma natural dentro o fuera del sistema.

De forma general, cuando observamos un cierto sistema se pueden definir estas alteraciones como **artefactos**, es decir, **señales observadas que no se encuentran presentes en la fuente original si no que se deben a alteraciones externas a dicha fuente.**

Dependiendo de la aplicación de las señales observadas, estos artefactos pueden producir problemas de varios tipos. Así por ejemplo, en la interpretación médica puede generar confusión a la hora de diagnosticar una cierta patología o, de forma más general, pueden enmascarar información importante que llevaría a la detección de una cierta patología. Por otro lado, en sistemas de reconocimiento de voz pueden impedir el correcto reconocimiento de la fuente.

El presente proyecto tiene como objetivo definir un sistema que pueda operar en el ámbito de las señales multicanal para identificar y suprimir estos artefactos de forma automática. Este sistema se basará en 2 tecnologías principales:

- Un sistema de aprendizaje supervisado, capaz de detectar y clasificar eventos en señales temporales.
- El análisis de Componentes Independientes (ACI) (en inglés ICA), un método estadístico orientado a dividir una señal en las fuentes que originan las señales multicanal observadas.

Con el fin de ejemplificar la propuesta y los avances realizados, se particulariza al tratamiento de señales EEG, capturadas mediante la observación de 72 sensores dispuestos sobre el cuero cabelludo de los sujetos observados.

## 1. OBJETIVOS

El objetivo general de este trabajo es proponer una metodología de supresión de artefactos en señales multicanal de cualquier tipo. Es decir, se quiere plantear una solución que sea capaz de eliminar dichos artefactos sin requerir un conocimiento específico sobre la naturaleza de las señales a tratar ni del problema que introduce el artefacto en las señales observadas. Para ello se plantea como base fundamental de este trabajo la complementariedad entre un sistema inteligente para la detección de artefactos presentes en la señal a tratar y la aplicación de técnicas de separación de componentes independientes.

Bajo este reto se definen los siguientes objetivos específicos:

- Establecer criterios generales en cuanto al tamaño de la ventana de trabajo al operar sobre un artefacto concreto para aplicar ICA.
- Escoger un algoritmo ICA apropiado para este propósito de filtrado general.
- Definir parámetros genéricos de referencia en las señales en base a los cuales se puedan establecer criterios de selección y eliminación de componentes independientes que se consideren fuentes de ruido.
- Detectar la posición de las fuentes de ruido en base a su cercanía a los sensores de observación de la señal.
- Analizar la aplicación de un sistema de detección de eventos en señales multicanal.

Aunque se pretende una solución general para señales multicanal de cualquier naturaleza, en este proyecto nos centraremos en el análisis de señales EEG. Por ese motivo se introduce a continuación la naturaleza de estas señales multicanal y su problemática en cuanto a la aparición de artefactos.

## 2. MOTIVACIÓN Y ORIGEN

### 2.1. EEG

La electroencefalografía (EEG) es el registro y evaluación de los potenciales eléctricos generados por el cerebro y obtenidos en la superficie del cuero cabelludo [1]. La actividad eléctrica cerebral es consecuencia de las corrientes iónicas generadas por los diversos procesos bioquímicos a nivel celular. La electroencefalografía es el resultado de la suma temporal y espacial de los potenciales postsinápticos procedentes de las neuronas piramidales orientadas verticalmente en la corteza cerebral.

La actividad del EEG tiene una amplitud pequeña, del orden de 5  $\mu\text{V}$  a 200  $\mu\text{V}$ , lo que implica que estas señales son difíciles de registrar y pueden distorsionarse con facilidad. Además, la actividad eléctrica cerebral debe involucrar de forma sincrónica una zona de corteza relativamente extensa (aproximadamente 6  $\text{cm}^2$ ) para que los potenciales cerebrales puedan registrarse en el cuero cabelludo. Al igual que en otros ámbitos de la ciencia y la tecnología, la metodología digital ha revolucionado la práctica médica y el EEG, ya que permite la medición cuantificada y precisa de los distintos aspectos de amplitud, latencia de las puntas epileptiformes, caracterización de los componentes de frecuencia y la detección automática de patrones.

La actividad electroencefalográfica se divide, según la frecuencia de los potenciales, en cuatro grupos que se designan con letras griegas:  $\alpha$  (8 a 13 ciclos por segundo o Hz),  $\beta$  (mayores de 13 Hz),  $\theta$  (4 a 7 Hz) y  $\delta$  (bajo 4 Hz) (Figura 2). En el EEG de superficie estas ondas tienen un tamaño que se ubica en el nivel de los microvoltios ( $\mu\text{V}$ ); su amplitud es baja si la tensión no excede de 20  $\mu\text{V}$ , mediana si oscila entre 20 y 50  $\mu\text{V}$  y grande si es de 50  $\mu\text{V}$ . En general las frecuencias más rápidas tienen menor amplitud.

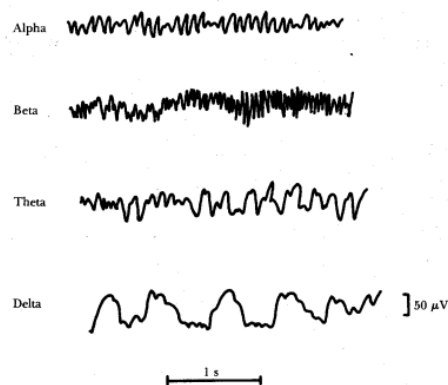


Figura 2 - Ejemplos de ritmos EEG normales

Para la medición del EEG se utilizan electrodos, siendo estos piezas de un metal buen conductor de la corriente eléctrica que se aplican o insertan sobre el cuero cabelludo o el cerebro con el fin de registrar los potenciales eléctricos neuronales. Sus características y diseño dependen fundamentalmente de qué localización adoptan para dicho fin. Existen dos

tipos de electrodos según la ubicación dónde se sitúan: superficiales o extracraneales, aplicables sobre el cuero cabelludo o en la base del cráneo, y neuroquirúrgicos o intracelulares insertados en el tejido nervioso, pudiendo ser éstos corticales o intracorticales. Para los experimentos en neurociencia cognitiva se utilizan generalmente electrodos superficiales, de fácil y rápida colocación e indoloros y, sobre todo, con una impedancia interelectrónica (resistencia al paso de la corriente alterna) no superior a  $5K\Omega$ . La Figura 3 muestra un ejemplo de distribución de los electrodos sobre el cuero cabelludo.

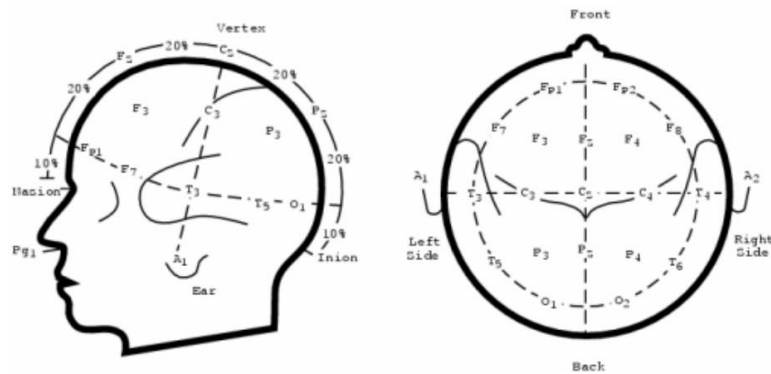


Figura 3 - Sistema 10-20 internacional de distribución de electrodos sobre la superficie de la cabeza

## 2.2. Artefactos

Uno de los problemas en la medición y análisis de EEG reside en la introducción de ruidos y señales no deseadas denominados artefactos que distorsionan las señales cerebrales.

La señal EEG puede verse afectada por señales fisiológicas del propio sujeto o señales externas a él. Así por ejemplo, a pesar de solicitar al sujeto que durante el registro permanezca lo más relajado e inmóvil posible, existen cambios de postura, movimientos de cuello o cabeza que se inmiscuyen en el señal eléctrica registrada. Incluso, si el sujeto consigue mitigar cualquier voluntad de movimiento, las señales fisiológicas o biopotenciales, debidas al parpadeo o a la actividad cardiaca, resultan inevitables.

### 2.2.1 Artefactos del entorno

En esta posible clasificación de los artefactos presentes en un EEG, agrupamos como artefactos de entorno aquellos que no se deben a fuentes de ruido internas al propio sujeto observado si no de los elementos del contexto a la propia lectura del EEG. Entre las principales fuentes de ruido del entorno pueden incluirse:

Corriente alterna: Provocado por la corriente eléctrica de 50Hz.

Mal contacto de electrodos: Provoca inestabilidad en las impedancias, dando lugar a ondas muy agudas, que confluyen en el trazado en la región correspondiente al electrodo mal colocado.

Marcapasos: Da lugar a ondas agudas, rítmicas, cuando la derivación electrocardiográfica del EEG se sitúa cerca del marcapasos.

### 2.2.2 Artefactos biológicos

Este grupo de artefactos están relacionados con el propio sujeto observado, y se deben a las señales eléctricas producidas por el cuerpo que se introducen en la lectura de los electrodos desplegados sobre su cabeza:

Artefacto electromiográfico: Es el artefacto biológico más frecuente. Secundario a la contracción de los músculos del cuero cabelludo, frontal, facial, masetero, y cuello. La contracción de maseteros, pueden hacer un registro indescifrable.

Artefacto de movimientos musculares involuntarios: presentes en pacientes críticos producidos por la existencia de mioclonías, temblor o actividad tónico-clónica.

Artefactos de parpadeo: De elevada amplitud, se transmite a regiones frontales (véase Figura 4).



Figura 4 - Ejemplo de artefactos producidos por parpadeo

Artefacto glosocinético: El movimiento de la lengua, provoca un artefacto que puede interpretarse como un brote de actividad lenta generalizado, de máxima amplitud en regiones centrales.

Artefacto glosofaríngeo: El movimiento realizado al tragar, activa la lengua, musculatura faríngea, y esfenopalatina, que se transmite como brotes de actividad lenta en regiones centrales.

Sudor: Genera importantes oscilaciones muy lentas en todos los canales, sobre todo en los frontales.

Artefacto Arterial: Los electrodos situados cerca de las arterias temporales, o en arteriolas, generan una onda con morfología de punta pulsátil, rítmica.

Artefacto electrocardiográfico: El potencial de acción generado por el corazón, se transmite hasta el cuero cabelludo, dando lugar a puntas rítmicas, que pueden ser confundidas con crisis comiciales.

Balístocardiográfico: El propio movimiento del corazón transmite fluctuaciones que se registran en cuero cabelludo.

### 3. ICA y DETECT

Para la supresión de artefactos existen múltiples estrategias y técnicas basadas en la propia naturaleza de las señales a tratar y en las características que se conocen de los propios artefactos. Así por ejemplo, se aplican técnicas de filtrado de baja frecuencia para eliminar el ruido introducido por la red eléctrica a través de los equipos de medición (siempre que la naturaleza de la señal lo permita en caso de no presentar información importante en esa banda) o soluciones complejas basadas en el conocimiento que se tiene de las señales a tratar.

Sin embargo, dado el planteamiento abierto de esta propuesta, donde se pretende abordar el problema sin un conocimiento previo de la naturaleza de las señales, sus características, ni las de los artefactos a eliminar, el antecedente más relevante se encuentra en la aplicación del análisis de componentes independientes (ICA de sus siglas en inglés), método que propone recuperar las fuentes de una señal estadísticamente independientes a partir de la observación de mezclas de dichas señales (como es el caso de las señales multicanal). Esta técnica es de utilidad en diversas aplicaciones, entre las que se encuentran los sistemas automáticos de reconocimiento de voz, el procesamiento de imágenes, las telecomunicaciones y la bioingeniería.

#### 3.1. Separación ciega de fuentes.

ICA nace en la década de los 90 a través de diferentes propuestas para resolver el problema de la separación ciega de fuentes (Blind source separation) [2][3]. Un ejemplo habitual en la literatura sobre ICA plantea un problema con una habitación en la que hay varias personas hablando simultáneamente y en la que se cuenta con varios micrófonos que graban a todas las personas mezcladas desde diferentes posiciones. Se trata del denominado *Cocktail-Party problema* (véase Figura 5).

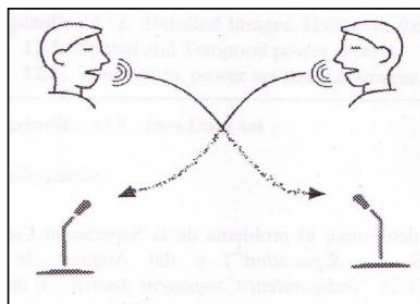


Figura 5 - Cocktail-Party Problem

Si despreciamos los ecos y reverberaciones, tendríamos en cada micrófono una combinación lineal de cada señal de voz, es decir, la intensidad de cada señal dependerá de la distancia entre cada fuente y sensor. De esta manera, podemos modelar matemáticamente esta situación como:

$$\begin{aligned} x_1(t) &= a_{11}s_1(t) + a_{12}s_2(t) + a_{13}s_3(t) \\ x_2(t) &= a_{21}s_1(t) + a_{22}s_2(t) + a_{23}s_3(t) \\ x_3(t) &= a_{31}s_1(t) + a_{32}s_2(t) + a_{33}s_3(t) \end{aligned} \quad (3-1)$$

Donde:

- $x_1(t)$ ,  $x_2(t)$  y  $x_3(t)$  son las grabaciones en cada instante  $t$  tomadas por los micrófonos, las cuales llamaremos mezclas  $x_i(t)$ ;
- $s_1(t)$ ,  $s_2(t)$  y  $s_3(t)$  son las señales provenientes de cada individuo que llamaremos fuentes  $s_i(t)$
- Los coeficientes  $a_{ij}$  con  $i, j = 1, 2, 3$ , se denominan parámetros de mezclado y representan la contribución de cada fuente a cada mezcla  $x_i(t)$ . El valor de estos parámetros dependerá de las distancias entre los individuos y los sensores.

Para generalizar el problema no limitaremos las señales emitidas por las fuentes solo a voces sino a cualquiera señal que puede sufrir los efectos de mezclado. De esta manera no sólo mediremos con micrófonos, sino con cualquier otro tipo de sensor (como es el caso del EEG en el que se particulariza esta memoria).

Podemos simplificar la ecuación  $x(t) = a_{21}s_1(t) + a_{22}s_2(t) + a_{23}s_3(t)$  (3-1) utilizando notación matricial, de manera que el proceso de mezclado lo expresamos como:

$$x(t) = As(t) \quad (3-2)$$

donde

$$x(t) = [x_1(t), x_2(t), \dots, x_N(t)]^T,$$

$$s(t) = [s_1(t), s_2(t), \dots, s_M(t)]^T$$

$$A = \begin{pmatrix} a_{11} & \dots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \dots & a_{MN} \end{pmatrix}$$

$M$  y  $N$  representan el número de fuentes y de mezclas respectivamente. Si los parámetros de mezclado  $a_{ij}$  fueran conocidos, el sistema de ecuaciones lineales sería resuelto invirtiendo el sistema de la siguiente forma

$$s(t) = A^{-1}x(t) \quad (3-3)$$

Sin embargo, estos parámetros no se conocen, ni tampoco las fuentes. Es el desconocimiento de las fuentes y de los parámetros de mezclado lo que se conoce con el término “*blind*” o ciego (en español).

Este es el punto de partida de ICA, que **establece tres hipótesis principales** para su correcta aplicación:

- Que las fuentes son estadísticamente independientes entre sí y no tienen distribución

gaussiana.

- Que se dispone de un número de señales mezcladas igual o superior al número de componentes independientes que se desean identificar.
- Que la mezcla de señales es lineal (existen aproximaciones posteriores de ICA para mezclados no lineales, pero este ámbito no se contempla en este trabajo).

En este contexto, para resolver el problema ICA plantea una transformación lineal  $\mathbf{W}$ , tal que:

$$u(t) = \mathbf{W}x(t) \quad (3-4)$$

Donde  $u(t) = [u_1(t), \dots, u_M(t)]^T$  es una estimación de  $s(t)$  y

$$\mathbf{W} = \begin{pmatrix} w_{11} & \dots & w_{1N} \\ \vdots & \ddots & \vdots \\ w_{M1} & \dots & w_{MN} \end{pmatrix} \quad \text{es una matriz de pesos}$$

Para estimar  $\mathbf{W}$ , existen diferentes algoritmos implementados en la literatura basados en diferentes reglas de búsqueda de independencia. Algunos de los criterios utilizados son:

- Maximización de la No gaussianidad.
- Estimación de la máxima verosimilitud.
- Análisis no lineal de componentes principales.
- Métodos tensoriales.
- Minimización de la mutua información.
- Maximización de la información.

Todos estos criterios están relacionados entre sí y su operación puede ser expresada de la siguiente forma:

#### **Método ICA = función objetivo + algoritmo de optimización**

La función objetivo es alguna regla o función que mide la independencia y que se busca maximizar o minimizar bajo algún tipo de restricción (empleando los criterios anteriores), mientras que el algoritmo de optimización es algún método como gradiente descendiente, natural o Newton para buscar esta maximización o minimización.

Todos estos algoritmos buscan la independencia estadística entre las salidas, y su elección depende de las propiedades estadísticas de la función objetivo y de la velocidad de convergencia, requerimientos de memoria, estabilidad numérica, etc.

### 3.2. Ambigüedad de ICA

$u(t)$  es una estimación de  $s(t)$  que proporciona formas de onda que contienen la información relevante de las fuentes. Sin embargo, ICA no puede garantizar que  $W = A^{-1}$ . Esto es debido a que:

- Las varianzas (energías) de las componentes independientes (CI) no se pueden determinar. La razón se debe a que tanto  $A$  como  $s(t)$  son desconocidas. Cualquier escalar que multiplique a  $s(t)$  podría ser siempre cancelado al dividir la correspondiente columna de  $A$  por el mismo escalar.
- El orden de las CI no se puede determinar. De nuevo, dado que tanto  $s$  como  $A$  son desconocidos, la matriz de permutación  $P$  y su inversa pueden ser sustituidas en el modelo de manera que  $x = AP^{-1}Ps$ . Los elementos de  $Ps$  son las variables independientes originales  $s$  pero en otro orden. La matriz  $AP^{-1}$  es una nueva matriz de mezclado, que puede ser resuelta con algoritmos de ICA [4].

### 3.3. Independencia estadística

Conviene resaltar aquí la importancia del término independiente. Como se ha comentado anteriormente, la primera hipótesis de ICA es que las fuentes a identificar provienen de sistemas distintos que emiten señales independientes.

En el lenguaje coloquial, a dos señales que no están correlacionadas se les suele llamar independientes. Desde el punto de vista matemático, la afirmación anterior no es cierta: independencia y correlación son propiedades estadísticas diferentes.

**En concreto, la independencia estadística es una condición más restrictiva que la ausencia de correlación, ya que dos señales independientes necesariamente estarán descorrelacionadas pero dos señales descorrelacionadas no tiene porqué ser independientes.**

Esta diferencia conceptual se detalla en diferentes documentos como [5], pero en resumen, cuando dos señales  $x$  e  $y$  son estadísticamente independientes, la densidad de probabilidad conjunta  $\rho_{xy}(x,y)$  viene dada por el producto de las densidades de probabilidad marginales de cada una de las señales  $\rho_x(x)$  y  $\rho_y(y)$ .

$$\rho_{xy}(x,y) = \rho_x(x)\rho_y(y) \quad (3-5)$$

### 3.4. Algoritmo FastICA

El algoritmo **FastICA** fue desarrollado en el Laboratorio de Información y Ciencias Computacionales de la Universidad Tecnológica de Helsinki, por Hugo Gvert, Jarmo Hurri, Jaakko Srel y Aapo Hyvarinen. Este algoritmo busca la independencia basándose en la medida de la no gaussianidad. Según el teorema de límite central [6] si mezclamos dos o más señales independientes, la mezcla resultante será más gaussiana que las señales originales. Recordando el Cocktail-Party problem, los micrófonos grababan mezclas de señales de voz. Por

lo tanto, según este teorema las mezclas son más gaussianas que las señales originales. El algoritmo **FastICA** busca por tanto revertir este proceso de mezclado. Recordando la ecuación  $u(t) = Wx(t)$  (3-4):

$$u(t) = Wx(t)$$

**FastICA** busca que  $u(t)$  sea lo menos gaussiano posible cambiando los pesos  $W$ . Al hacerlo menos gaussiano logramos encontrar los componentes independientes, lo cuales son las estimaciones de nuestras fuentes. Para ello se utiliza una función que mide la no gaussianidad a partir de la **Negentropía**, que se puede considerar como una medida de distancia entre la densidad de probabilidad de una variable gaussiana de referencia y la densidad de la variable a que estamos analizando. La Negentropía se define como [7]:

$$J(u) = H(u_{gauss}) - H(u) \quad (3-6)$$

Donde  $H(u_{gauss})$  es la entropía de un vector gaussiano con igual varianza que el vector  $u$  y  $H(u)$  es la entropía de  $u$ . La Negentropía es siempre positiva, y es cero sólo cuando  $u$  tiene una distribución gaussiana.

En **FastICA** se toma la Negentropía de  $u$  y se verifica su distancia con respecto a la variable gaussiana. El objetivo es encontrar unos pesos que lleven al máximo de esta distancia (que maximicen la Negentropía). Los pesos encontrados son los que conducen a los componentes independientes y a nuestras estimaciones de las fuentes.

Sin entrar en detalle al proceso del algoritmo, cabe destacar que el algoritmo **FastICA** no estima todos los componentes a la vez si no que lo hace secuencialmente. En la Figura 6 se muestra el esquema de búsqueda de las componentes. En este ejemplo tenemos tres mezclas. Estas mezclas sufren un preproceso (de blanqueado y centrado de las señales, hecho de aplicación común en análisis ICA) y son multiplicadas por un vector de pesos para producir las estimaciones  $u$ . Este vector es utilizado para buscar el máximo de la Negentropía y generar unos nuevos pesos. Este proceso se repite hasta encontrar los pesos óptimos.

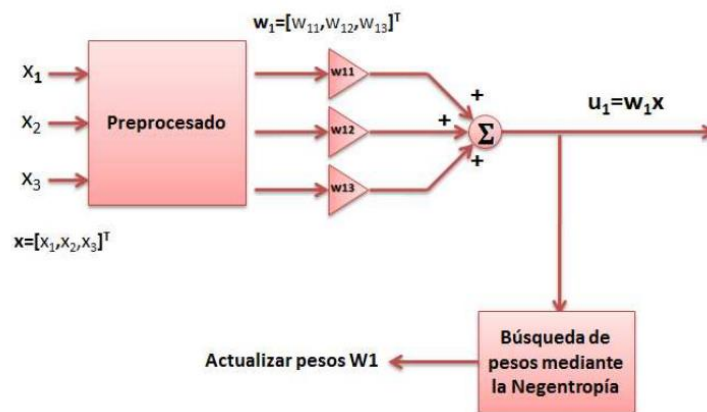


Figura 6 - Proceso de FastICA para la búsqueda de una componente independiente

A partir de ahí, para encontrar los siguientes componentes, se utiliza el mismo proceso pero, para asegurarse de que no coinciden con el peso previamente encontrado, se aplica un proceso de ortogonalización. Este proceso asegura que el siguiente vector de pesos sea ortogonal al encontrado previamente y de esta forma diferente e independiente. De esta manera el proceso descrito en el esquema se repite hasta encontrar todos pesos.

### 3.5. Algoritmo JADE

---

JADE (Joint Approximate Diagonalization of Eigen-matrices) [8], basa su proceso en la combinación de métodos de descomposición en autovalores (o eigen-valores), criterios de optimización introducidos en [9] y la diagonalización conjunta de matrices.

A diferencia de FastICA, JADE no realiza una estimación del número de CI en el conjunto de señales observadas, si no que permite identificar el nº de CI en que se desean separar. Dado este hecho diferencial, se introduce en este trabajo para realizar comprobaciones sobre la flexibilidad que puede aportar (principalmente en relación a la carga computacional).

### 3.6. Aplicación de ICA en EEG

---

En [10], se plantea el uso de ICA en la separación de artefactos en EEG. En dicho artículo se muestra cómo muchos de los métodos existentes que plantean eliminar artefactos de los registros de EEG suprimen a la vez información relevante de la señal a analizar, y cómo esto conlleva en muchas ocasiones a descartar analizar las secciones de EEG con artefactos.

El experimento se basa en muestras de EEG recogidas de sujetos de referencia con actividad cerebral normal y sujetos autistas, y demuestra la efectividad de ICA a la hora de detectar y eliminar los artefactos que se introducen en las señales EEG observadas. Este trabajo considera que, aunque los mecanismos neurales que generan EEG no son totalmente conocidos, los supuestos del algoritmo de ICA son generalmente compatibles con un modelo ampliamente asumido en que los datos del EEG registrados mediante múltiples sensores del cuero cabelludo son una suma lineal de las activaciones generadas por las fuentes neuronales y las fuentes de distintos artefactos. Por tanto, concluye que ICA parece ser un método de aplicación general y eficaz para la eliminación de una amplia variedad de artefactos en registros de EEG, dado que las fuentes de ruido son generalmente independientes de las fuentes de actividad cerebral.

ICA muestra varias ventajas en comparación con otros métodos de eliminación de artefactos:

- Se trata de un algoritmo computacionalmente eficiente y con requisitos computacionales no demasiado elevados incluso con grandes conjuntos de datos de EEG.
- Separa simultáneamente el EEG y los artefactos sin depender de canales de referencia "limpios" para cada tipo de artefacto (práctica común para la eliminación de fuentes de ruido en el EEG provenientes del músculo cardíaco y de los parpadeos). Esto evita además los problemas de regresión que generan estos métodos con canal de referencia.

- No se requieren análisis adicionales para diferentes tipos de artefacto.
- El filtrado por ICA preserva más completa la señal que otros mecanismos más agresivos.
- Concluyen también que el método debería ser válido para otras fuentes multicanal biomédicas (lo que se alinea con el propósito abierto de este proyecto).

Sin embargo, se destacan también algunos limitantes importantes:

- Como se ha mencionado anteriormente, ICA puede descomponer, a lo sumo, tantas fuentes independientes como canales de datos se capturen. El número de fuentes que contribuyen a las lecturas en un EEG captado en el cuero cabelludo es desconocido, pero de forma general se considera que las fuentes físicas cerebrales que generan la actividad son mayores que el número de electrodos disponibles en la captura de un EEG.
- ICA se basa en el análisis estadístico de datos, por lo tanto, sus resultados no serán significativos si la cantidad de datos es insuficiente. Esto conduce a pensar que lo mejor es utilizar todos los datos disponibles para estimar de forma fiable las CI de las señales observadas. Sin embargo, esto solo es cierto cuando las fuentes de los artefactos y la actividad cerebral son estacionarias a través del tiempo, y el número total de estas fuentes es menor que el número de canales de datos.

Aunque en general sí hay razones para creer que las fuentes cerebrales y de los artefactos permanecen estacionarias con el tiempo, el propio trabajo demuestra que en determinados experimentos la actividad cerebral y las fuentes de los artefactos pueden llegar hacerse dependientes, lo que uniría sus componentes y conllevaría la supresión de componentes del EEG al eliminar la componente que incluye el artefacto. De forma experimental se establecen periodos de 10 segundos como el intervalo que optimiza los resultados del experimento.

- Otra limitación del método propuesto es que la eliminación de artefactos requiere una inspección visual de los componentes independientes para determinar aquellas que se desean eliminar. Esto consume tiempo y no es deseable para la eliminación de artefactos rutinarios en la captura de EEG.

### 3.7. DETECT – Detección e identificación de eventos

---

DETECT es una toolbox para Matlab<sup>1</sup> que incluye métodos para la detección y el etiquetado de eventos en series temporales. Se trata de un sistema de inteligencia artificial supervisada que debe alimentarse con ejemplos de eventos etiquetados con los que crear un clasificador o modelo que será el encargado de analizar futuras señales para detectar dichos eventos.

---

<sup>1</sup> <https://github.com/VisLab/detect>

Una vez creado el modelo, la toolbox ofrece otros métodos con los que analizar esas nuevas señales y devuelve etiquetados los eventos detectados (categoría e intervalo de tiempo en el que aparece). También incluye métodos que facilitan la construcción de ejemplos etiquetados y funciones para la construcción de modelos, así como conjuntos de datos de muestra para la realización de experimentos.

### **3.7.1. Aplicación de DETECT para etiquetar eventos en EEG**

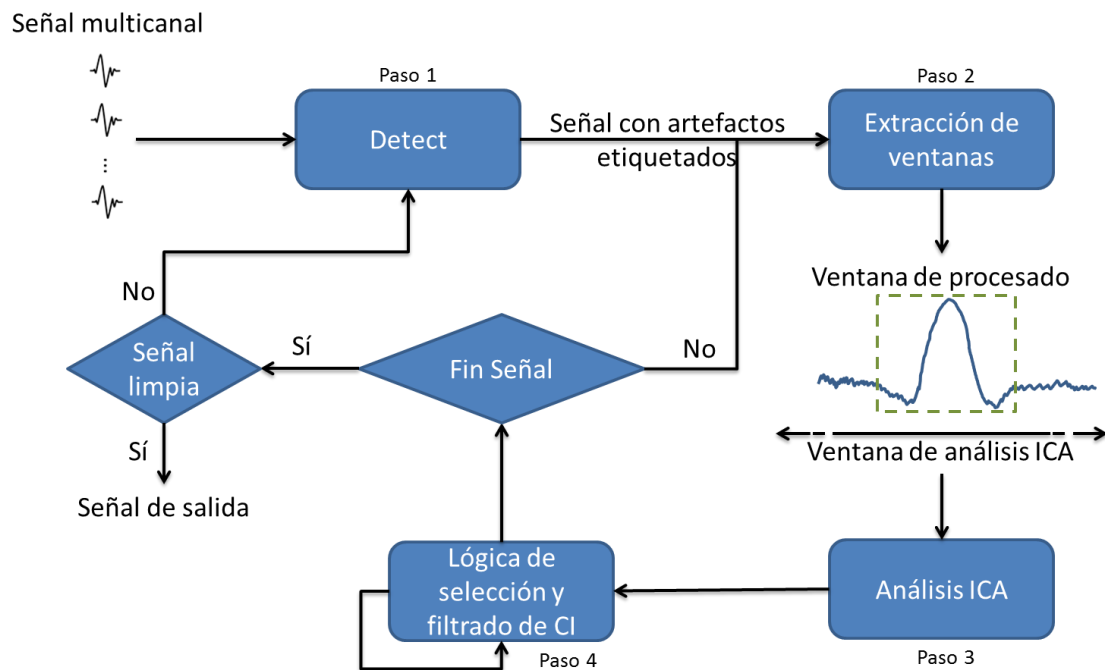
En [11] se valida de forma experimental la aplicación de este conjunto de herramientas en la detección e identificación de artefactos en señales EEG multicanal. Esta publicación muestra el análisis realizado para la detección de estos artefactos y los resultados del mismo al compararlos con la detección manual por parte de 3 expertos en la materia, con los que se alcanzó entorno a un 84% de coincidencias.

## 4. DESARROLLO Y RESULTADOS DEL PROYECTO

### 4.1. SOLUCIÓN PROPUESTA

Como se ha planteado, el objetivo de este trabajo es proponer una metodología de supresión de artefactos en señales multicanal de cualquier tipo, sin requerir para ello conocimiento específico de la naturaleza de la señal o de sus artefactos. Esta solución se basa en la combinación de las técnicas ICA y DETECT descritas en la sección anterior.

Este proceso se compone de varios pasos que se pueden contemplar en el siguiente diagrama.



A continuación se describe brevemente el contenido de cada uno de estos pasos del proceso, así como los aspectos experimentados en este trabajo. El detalle de la implementación de estos bloques de trabajo y las validaciones experimentales realizadas se detallan en la sección 4.2 a lo largo de diferentes casos de estudio planteados:

- **El primer paso** consiste en un análisis de la señal multicanal de entrada con las herramientas de la toolbox DETECT. En la experimentación que se plantea en este trabajo, se entrena un clasificador supervisado para la detección de diferentes artefactos sobre un tramo de EEG (tanto las señales de entrenamiento como la de test es ofrecida dentro de la toolbox de DETECT como ejemplo).

A la salida se obtiene la señal multicanal e información sobre los artefactos presentes y sus intervalos temporales de aparición dentro de la señal.

- **En el paso 2** se plantean diferentes opciones de ventanas de trabajo para el procesado

de las señales. Se debe considerar la ventana de trabajo sobre la que realizar el análisis ICA y la ventana de trabajo sobre la que aplicar el filtrado (se valora si filtrar cada artefacto de manera independiente o realizar procesados conjuntos en determinados casos, así como la anchura de la ventana de procesado sobre la que aplicar el filtrado).

- **En el paso 3** se aplica el análisis ICA para separar las componentes principales. En este trabajo se han probado 2 implementaciones disponibles en la literatura, escogiéndose una de ellas para las pruebas globales de la solución.
- **En el 4º paso** se deben identificar de forma automática las componentes independientes (CI) relacionadas con el artefacto para su eliminación. A lo largo de los casos de estudio se plantean diferentes criterios de selección y filtrado de componentes.

A modo de resumen cabe destacar que las variantes propuestas para la identificación automática de las componentes independientes asociadas a la fuente de ruido están basadas en la comparación energética de la señal. Se asume que, en presencia de un artefacto, la energía de la señal en el intervalo de ruido aumentará, y que será la fuente de ruido la que más energía aporte a la señal en ese tramo. Si bien es cierto que en alguna observación la fuente de ruido puede generar una interferencia destructiva que reduzca la energía de la señal observada en un determinado canal, de forma general los canales presentarán una mayor energía (recordemos que trabajamos con señales multicanal).

Los criterios y variantes planteadas para este procesado se detallan en la siguiente sección, en diferentes casos de estudio, junto con los resultados obtenidos.

Los pasos 2 a 4 se realizan de forma iterativa hasta terminar el procesado de la señal. Además, se valora también la posibilidad de aplicar la solución de forma iterativa en el caso de que, tras el proceso anterior, persistan posibles artefactos (dado que quizá ciertos artefactos enmascarasen otros menores).

En la siguiente sección se muestra, haciendo uso de distintos casos de estudio, la metodología propuesta y las principales características de cada uno de los pasos anteriores.

## 4.2. VALIDACIÓN EXPERIMENTAL Y RESULTADOS

Para la validación experimental de la solución propuesta, se ha procedido a implementar la solución en el entorno de simulación de Matlab®. Se cuenta con la toolbox DETECT<sup>2</sup>, así como con 2 algoritmos ICA de los implementados en literatura: FastICA<sup>3</sup> y JADE<sup>4</sup>.

### 4.2.1 Caso de estudio 1: separación de señales simuladas mediante algoritmo FastICA

Como primer paso, se definen una serie de señales artificiales fácilmente distinguibles con la intención de realizar una primera valoración de los algoritmos ICA.

Se generan 3 señales de 10 segundos cada una muestreadas a 500 muestras/sg. Las señales son las siguientes:

- Señal senoidal de amplitud 1u, frec 10 Hz y SNR 20
- Señal sierra de amplitud 1.3u, frec 7.5 Hz y SNR 18
- Señal cuadrada de amplitud 0.7u, frec 4 Hz y SNR 21

El ruido adicionado a todas las señales es ruido blanco gaussiano aleatorio. La Figura 7 muestra dichas señales y su histograma simulado.

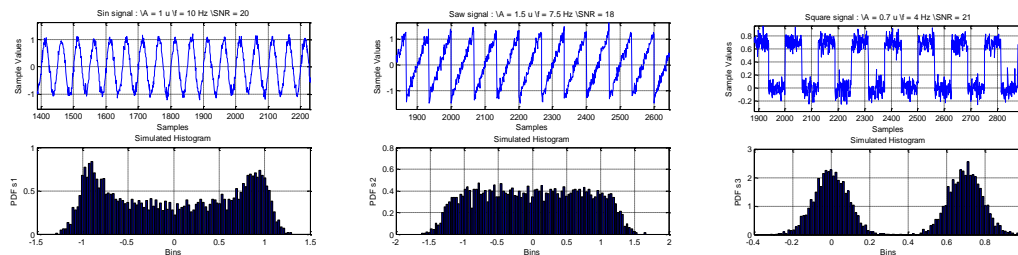


Figura 7 - De izda. a dcha.: detalle de las señales senoidal, sierra y cuadrada

Para el experimento, se generan aleatoriamente 5 combinaciones lineales de las señales anteriores, obteniéndose los resultados que se muestran en la Figura 8.

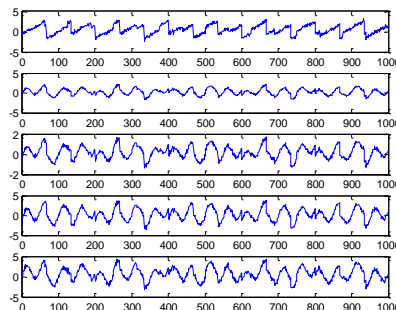


Figura 8 - Extracto de las señales "observadas" resultantes de la mezcla lineal de las fuentes

<sup>2</sup> DETECT: <https://github.com/VisLab/detect>

<sup>3</sup> FastICA: <http://research.ics.aalto.fi/ica/fastica/>

<sup>4</sup> JADE: <http://perso.telecom-paristech.fr/~cardoso/guidesepsou.html>

Por tanto, el experimento en este punto presenta las condiciones básicas para la separación de componentes principales mediante el algoritmo ICA:

- Se trata de señales estadísticamente independientes.
- Sus distribuciones son no gaussianas (como puede observarse de los histogramas simulados). De hecho, la única componente gaussiana presente en el medio observado corresponde al ruido blanco que afecta a las señales, siendo esta una componente que no se pretende separar.
- El medio en el que han sido mezcladas es lineal con un retardo de propagación despreciable, existiendo al menos tantas señales observadas (mezclas) como fuentes (componentes a separar). De hecho en el ejemplo definido se observan 5 señales en las que se mezclan las originadas por tan solo 3 fuentes.

Como medida de la no-gaussianidad de las señales de fuente, a continuación se recoge el nivel de curtosis (basado en el cuarto momento con respecto a la media) de las 3 señales:

```
>> kurtosis(s1)
ans =
    1.5598

>> kurtosis(s2)
ans =
    1.8692

>> kurtosis(s3)
ans =
    1.2514
```

Por lo tanto, se puede observar que se trata de señales platocúrticas.

Con este entorno de partida, en teoría se podría aplicar el algoritmo FastICA con la intención de obtener las 3 señales origen a partir de las observadas en el medio. Por tanto se ejecuta dicho algoritmo obteniendo las 3 componentes principales a la salida que se muestran en la Figura 9.

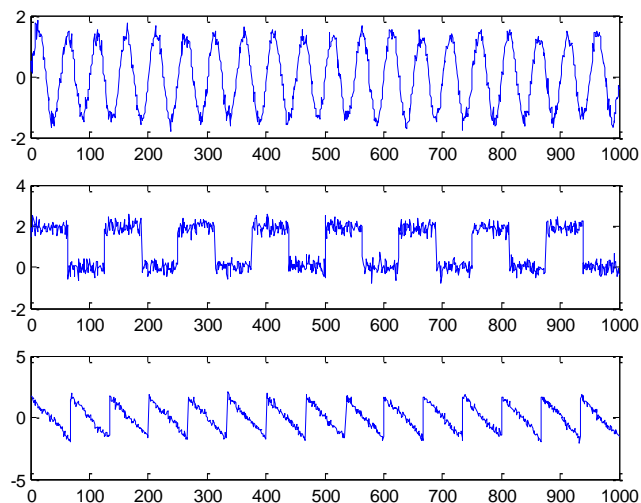


Figura 9 - Detalle los primeros segundos de las fuentes recuperadas por ICA

Como puede observarse, el algoritmo separa las mezclas de entrada en una aproximación bastante fidedigna de las 3 señales originales, pero mostrando la ambigüedad esperada en cuanto a la energía de las señales y su orden de salida.

#### 4.2.2 Caso de estudio 2: filtrado de “artefacto” en señales simuladas mediante algoritmo FastICA

Como siguiente paso se crea un artefacto artificial correspondiente a un pulso de intensidad elevada en el torno al segundo 7 de la muestra (véase Figura 10).

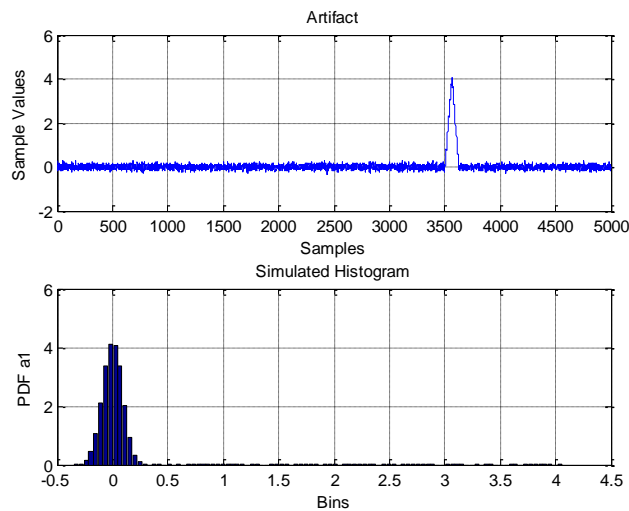


Figura 10 - Artefacto simulado

Si se combinan linealmente las señales fuente con este artefacto se obtienen las señales de la Figura 11.

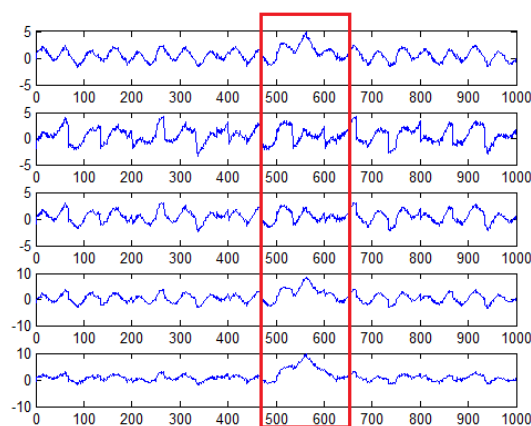


Figura 11 - Artefacto simulado mezclado en las señales observadas

A simple vista puede observarse el efecto del artefacto en las señales mezcladas en la zona resaltada, elevando su amplitud y distorsionando por tanto la señal observada.

Aplicando de nuevo el algoritmo FastICA a estas 5 mezclas de 4 señales se obtienen las componentes principales que aparecen en la Figura 12.

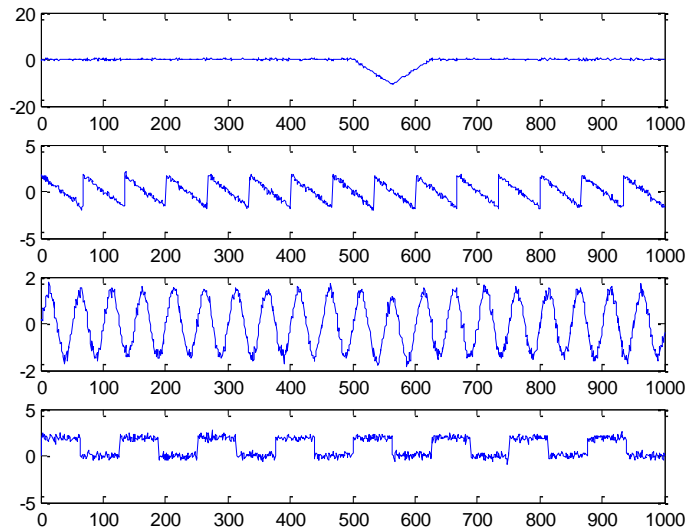


Figura 12 - Componentes principales del experimento con artefacto simulado

Como puede observarse, en esta ocasión ICA devuelve tanto las 3 señales originales como el artefacto simulado. En esta ocasión el artefacto queda recogido en la 1ª componente. Para eliminar el artefacto de las señales observadas es por tanto necesario eliminar esta componente.

En este punto es importante atender a la ambigüedad en cuanto a la energía original de las señales. No es posible sustraer directamente la componente separada de las señales observadas, si no que se debe calcular la proyección de la componente en cada una de ellas.

Recordando la fórmula general de ICA tenemos:

$$S = W * X$$

Donde  $S$  es la matriz de señales fuente,  $X$  la matriz de señales observadas y  $W$  la matriz de transformación del medio. Es decir, si se supone que se trata de la adquisición de un EEG, se tendría

$$X = \begin{matrix} & t1 & t2 & t3 & t4 & \dots \\ \text{electrodo 1} & & & & & \\ \text{electrodo 2} & & & & & \\ \text{electrodo 3} & & & & & \\ \text{electrodo 4} & & & & & \\ \dots & & & & & \end{matrix}$$

y  $S$  sería la matriz que recoge en cada fila cada una de las fuentes que emiten las señales observadas en los electrodos.

Por tanto, de esta función se puede interpretar cómo cada una de las fuentes tiene una proyección en cada uno de los electrodos que está regulada por la matriz de transformación  $W$ .

Una vez detectada una componente que corresponde a una fuente de ruido, el proceso para su eliminación pasa por calcular la proyección de esta fuente en las señales observadas para, entonces sí, sustraerla directamente.

Para ello se hace necesario calcular la inversa de la matriz de transformación  $W$  que denominaremos  $A$ :

$$X = A * S$$

A partir de esta información se puede calcular la proyección de la componente que recoge el artefacto en las señales observadas, que en este caso corresponde con la 1ª componente. La siguiente figura muestra la proyección de este artefacto en cada una de las 5 señales observadas:

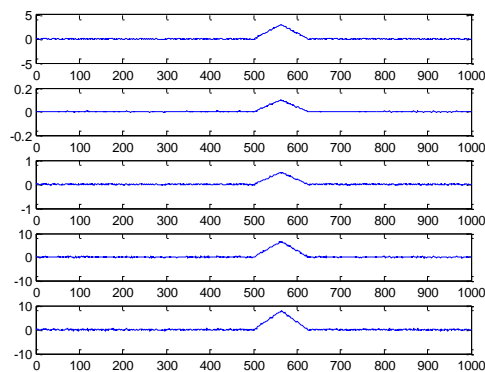


Figura 13 - Proyección del artefacto en las señales observadas

Por tanto, sustrayendo estas proyecciones de las señales observadas, se obtienen las señales sin el ruido introducido por el artefacto.

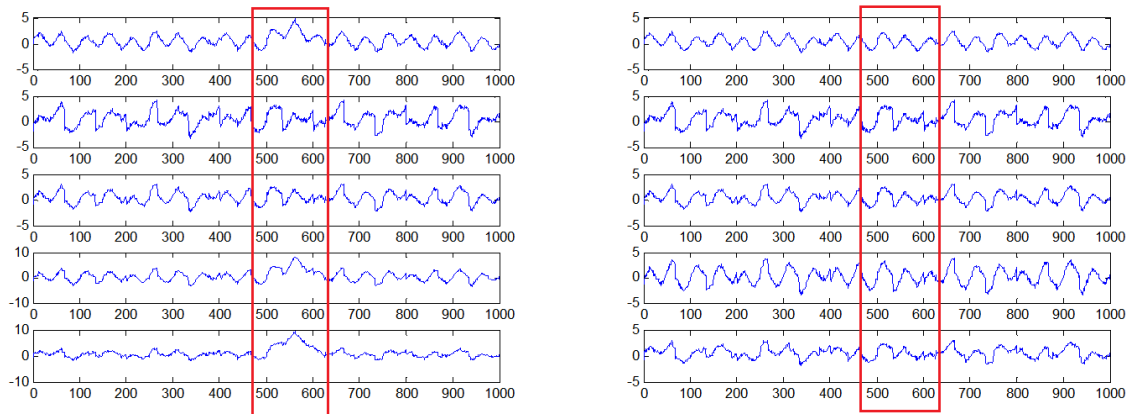


Figura 14 - Izda. señales originales; Dcha. señales con el artefacto filtrado

### 4.2.3 Caso de estudio 3: filtrado de “artefacto” en señales simuladas mediante algoritmo JADE.

El experimento anterior puede realizarse mediante el algoritmo JADE indicando al mismo la separación de 4 componentes y los resultados son muy similares.

Sin embargo, resulta interesante comprobar el efecto de una mala elección en el número de componentes independientes

En la Figura 15 se muestran las CI que genera el algoritmo en función del número N de componentes indicado:

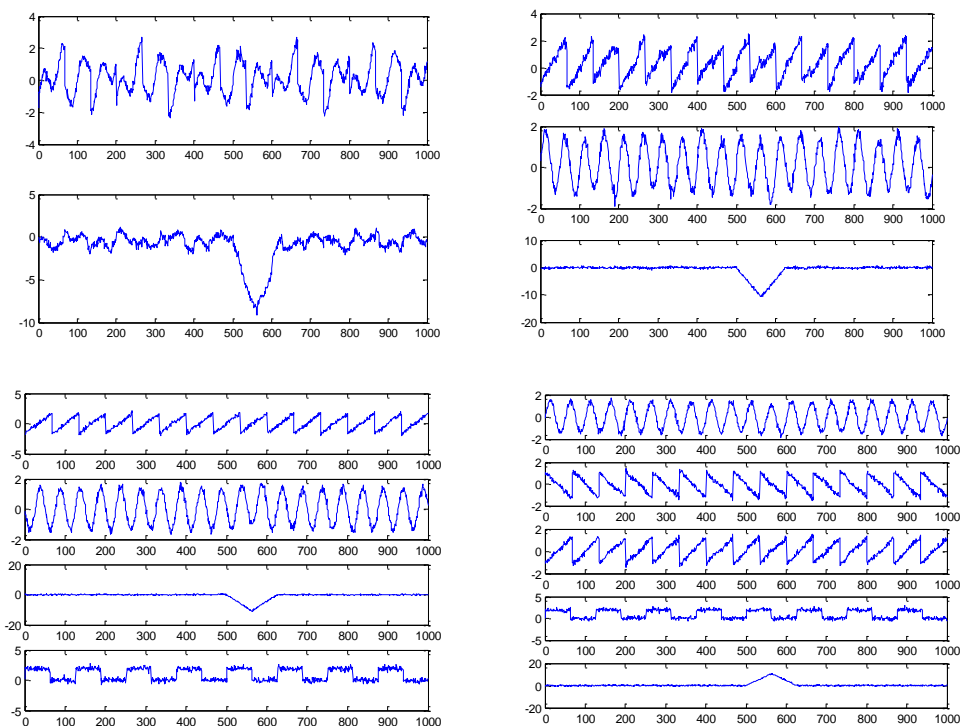


Figura 15 - N = 2, 3, 4 y 5 respectivamente

Como puede observarse, una incorrecta selección en el nº de CI puede llevar a resultados indeseados si, como en el caso N=2, el artefacto a eliminar se mezcla en una componente con información de las fuentes de señal que se desean preservar.

Para ver sus efectos en el proceso de filtrado, se repite el experimento con las mismas señales simuladas y se crean 5 mezclas aleatorias. Al ser un experimento controlado, se conoce que el número óptimo de Componentes Independientes (NOCI) es 4, pero si se utiliza JADE con el fin de obtener 3 componentes el resultado es el que se muestra en la Figura 16.

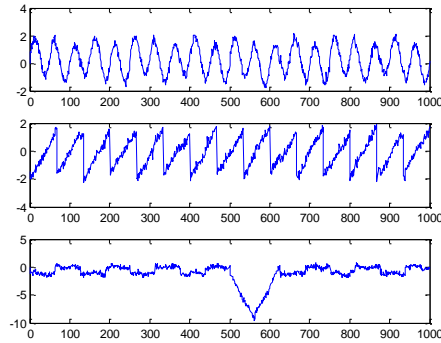


Figura 16 - CI obtenidas

A simple vista la 3ª CI parece la suma de la señal cuadrada y el artefacto, pero dado que incluye al artefacto se calcula su proyección sobre las señales observadas y se elimina (véase Figura 17).

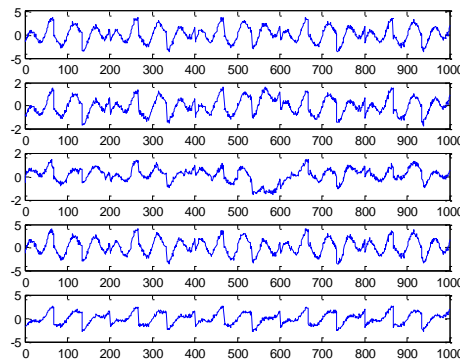


Figura 17 - Señal resultante tras el filtrado

Si ahora aplicamos FastICA sobre este conjunto de señales, el algoritmo devuelve las 3 componentes independientes que se muestran en la Figura 18.

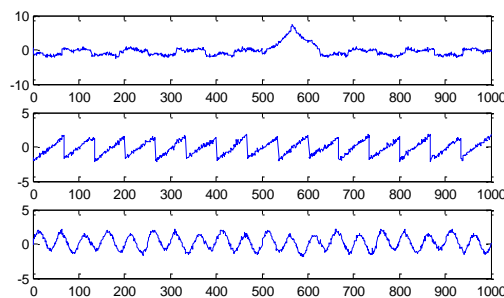


Figura 18 - CI devueltas por FastICA

Viendo lo anterior, se puede concluir que:

- La CI 3 sí combinaba información de las fuentes cuadrada y artefacto.
- Tras filtrar esta componente ya no se consigue obtener ninguna de las 2 de forma separada.
- El filtrado deja restos de la señal y de hecho quedan relacionadas estadísticamente, no siendo ya separables en diferentes CI.

#### 4.2.4 Caso de estudio 4: Medición de calidad del filtrado de ambos algoritmos por comparativa de sus Errores Cuadráticos Medios (EQM) a partir de una señal de control.

En este experimento, se pretende medir la efectividad de ambos algoritmos a la hora de realizar el filtrado. El experimento es el mismo que en el caso de estudio 2, con la diferencia de que en este caso se generan también las mezclas lineales de las 5 señales observadas sin el artefacto. De este modo, se cuenta con una señal de referencia para evaluar la calidad del filtrado llevado a cabo en los casos de estudio 1 y 2.

Para valorar la calidad de la señal a la salida se calcula el Error Cuadrático Medio (EQM) entre la señal resultante y la señal de referencia.

Aplicando FastICA, este EQM resulta:

$EQM_{total} =$

$0.6171$

En el caso de JADE con 4 CI, se obtiene lo siguiente:

$EQM_{total2} =$

$0.6170$

En ambos casos el resultado es bastante similar en calidad, siendo la diferencia generada principalmente por la presencia de ruido blanco en las señales de entrada. La Figura 19 muestra la diferencia entre la señal de salida tras el filtrado con FastICA y la señal de referencia (muy similar a lo que se observa en el caso de JADE).

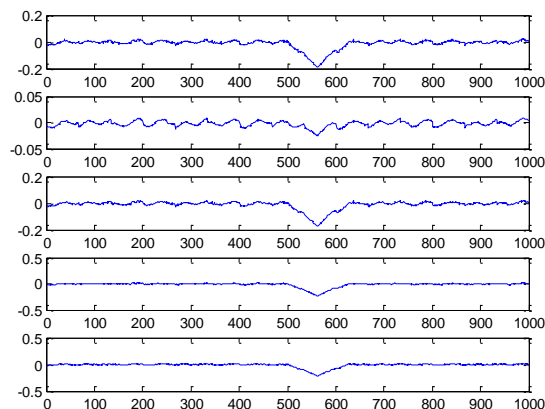


Figura 19 - Error introducido en el proceso de filtrado

En este caso, sí resulta importante destacar de nuevo el efecto de una mala selección en el número de componentes independientes durante el empleo de JADE.

- Para  $N=2$ , obtenemos un  $EQM_{total} = 35.1599$
- Para  $N=3$ , obtenemos un  $EQM_{total} = 17.2573$

Dado que no se conocen a priori el nº de componentes independientes de las señales multicanal sobre las que eliminar los artefactos, se opta por continuar en adelante con la implementación FastICA por su capacidad de automatizar el nº de CI óptimo.

#### 4.2.5 Caso de estudio 5: Filtrado automático mediante DETECT e ICA

En este caso de estudio se comienza a integrar DETECT e ICA para tratar de definir un método óptimo con el que filtrar los artefactos que sea capaz de detectar DETECT mediante separación de los mismos por ICA, detección automática de la componente relacionada con el artefacto, y supresión de la misma de las señales observadas. Para realizar el experimento, se utiliza el conjunto de señales EEG reales recogidas en la propia librería de DETECT para mostrar su funcionamiento.

En este primer caso de estudio de filtrado la separación automática de las CI que se consideran ruido y aquellas que se consideran actividad cerebral se basa en 2 hipótesis principales:

- Que los artefactos a eliminar los generarán las fuentes que mayor energía aportan a las señales observadas en los sensores.
- Que la fuente que genera el artefacto será separada en una única componente independiente durante la ejecución de ICA.

Para poder cumplir con la segunda hipótesis, el análisis de la señal para el filtrado se centra en cada artefacto de forma independiente, barriendo la señal completa para realizar el mismo proceso en cada uno de los artefactos detectados por DETECT.

El experimento se implementa en los siguientes bloques.

##### **Paso 1 – Uso de DETECT**

El primer paso, que corresponde con el 1er paso de la solución propuesta en la sección 4.1., consiste en procesar la señal a limpiar mediante DETECT. Como se ha descrito en la sección 3, DETECT ofrece un conjunto de herramientas para entrenar un modelo capaz de clasificar la presencia de hasta 6 tipos de artefactos analizando en diferentes tramos un EEG.

En este experimento, se genera el modelo a partir de las señales EEG de entrenamiento proporcionadas por la *toolbox* de DETECT y se analiza la señal de test aportada también en la misma *toolbox*.

```
%Cargamos las señales de entrenamiento y etiquetas de artefactos
load DETECT/data/training;
load DETECT/data/labels;
load DETECT/data/testing;

% generamos el modelo
model = getModel(training, labels);

%Analizamos la señal de test para detectar artefactos
results = labelData(testing, model, 256, .125); % label testing data
```

Además se aplica la función de umbral de DETECT para que clasifique como “nada” (segmentos de señal limpia, sin artefactos), aquellos tramos en los que se considera que puede haber un artefacto de algún tipo pero con un nivel de certeza por debajo del 50% y sea la segunda opción de mayor certeza el hecho de que no hay nada.

```
%Aplicamos el umbral del 50% de certeza  
results1 = thresholdPolicy(results, 'None', .5);
```

La Figura 20 muestra la representación del EEG facilitada por la toolbox DETECT, en la que se resalta en este caso un artefacto producido por el parpadeo del ojo.

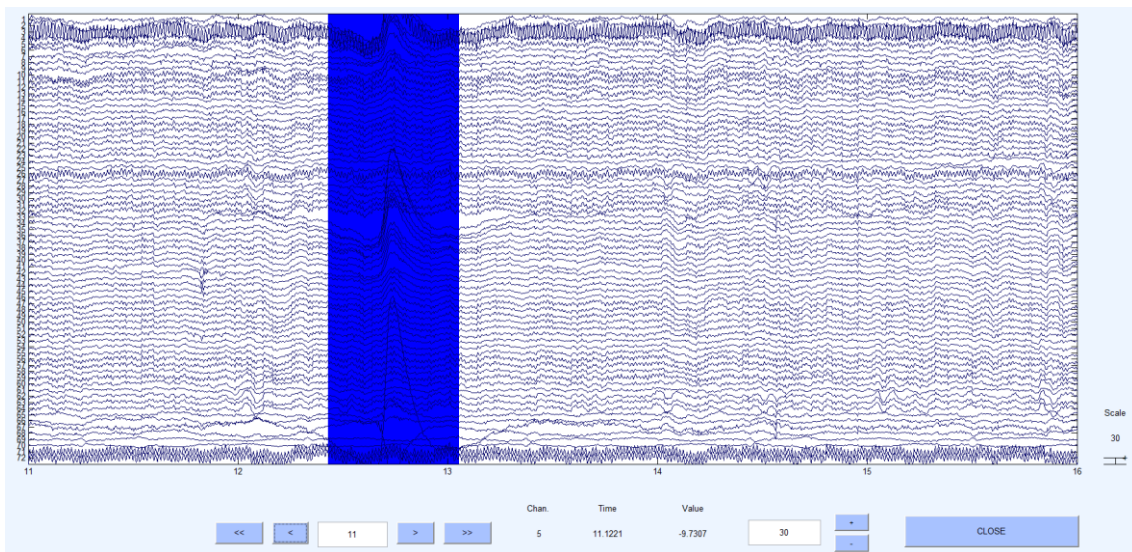


Figura 20 - Representación de EEG y artefactos ofrecida por DETECT

Además DETECT nos proporciona información sobre los artefactos detectados en la señal, y sus posiciones en la misma (véase Figura 21).

```

labelSet =
'Eye Blink'          [ 12.4336] [ 13.0547]
'Eye Left Movement' [ 16.3086] [ 17.0547]
'Eye Blink'          [ 19.8086] [ 20.3047]
'Eye Up Movement'   [ 20.3086] [ 20.4297]
'Eye Blink'          [ 25.8086] [ 26.4297]
'Eye Blink'          [ 40.9336] [ 41.5547]
'Eye Blink'          [119.0586] [119.5547]
'Eye Blink'          [173.4336] [174.0547]
'Eye Blink'          [175.1836] [175.9297]
'Eye Up Movement'   [178.5586] [178.8047]
'Eye Left Movement' [178.9336] [179.0547]
'Eye Blink'          [179.0586] [179.5547]
'Eye Up Movement'   [179.5586] [180.9297]
'Eye Left Movement' [180.9336] [181.0547]
'Jaw Movement'      [181.0586] [181.5547]
'Jaw Clench'         [181.5586] [181.9297]
'Jaw Movement'      [181.9336] [182.0547]
'Eye Left Movement' [182.0586] [182.1797]
'Eye Blink'          [184.5586] [185.1797]
'Eye Blink'          [274.1836] [274.8047]
'Eye Blink'          [285.1836] [285.6797]
'Eye Blink'          [296.9336] [297.4297]
'Eye Blink'          [350.0586] [350.6797]
'Eye Blink'          [355.5586] [356.1797]
'Eye Blink'          [361.9336] [362.6797]
'Eye Up Movement'   [362.6836] [362.8047]
'Eye Blink'          [362.8086] [363.1797]
'Eye Up Movement'   [363.1836] [363.3047]
'Eye Left Movement' [363.3086] [363.6797]
'Eye Up Movement'   [363.6836] [364.3047]
'Eye Blink'          [446.8086] [447.5547]

```

Figura 21 - Artefactos detectados: Tipo; Inicio (seg); Fin (seg)

### **Paso 2 – Extracción de ventanas**

Como se ha planteado, el filtrado automático se centrará en un análisis independiente de cada artefacto detectado. En este caso se comienza con el parpadeo (*'Eye Blink'*) detectado entre los segundos 12.43 - 13.05. Por tanto, se extrae para el análisis una ventana de la EEG observada en cada uno de los 72 sensores disponibles.

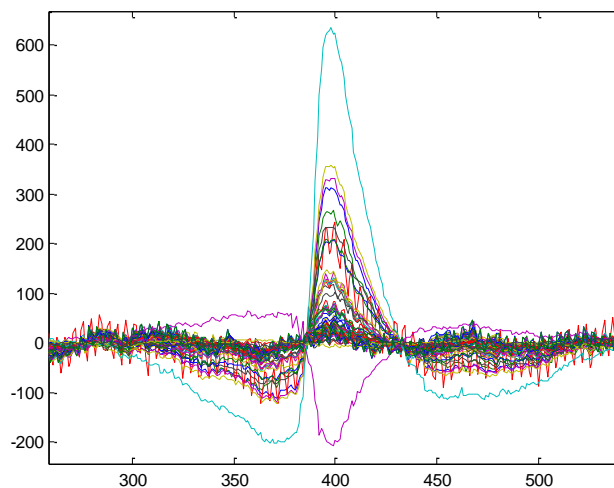


Figura 22 - Superposición de la EEG observada en los 72 sensores en el momento del parpadeo

El tamaño óptimo de esta ventana y sus efectos serán considerados en un posterior caso de estudio. Por el momento se asume que para obtener información relevante de contexto a la hora de aplicar ICA se ha de escoger una ventana de duración igual a 10 veces la duración del

artefacto estudiado en cada caso. En cuanto a la ventana de procesado, se trabaja de forma independiente con cada artefacto ciñendo la ventana al intervalo en el que DETECT indica que se encuentra dicho artefacto.

Para clarificar el proceso, se denomina al conjunto de señales observadas  $SO$  que se compone de las señales observadas en cada uno de los 72 electrodos ( $SO_1 \dots SO_{72}$ )

### **Paso 3 – Análisis ICA**

A continuación, se separan las componentes independientes de la ventana bajo análisis que contiene el artefacto. Para ello se aplica FastICA y se obtienen un total de 69 componentes independientes (que denominamos  $CI = CI_1 \dots CI_{69}$ ). A partir de este punto se pasa a trabajar con una ventana de la señal que se ajuste exactamente a la longitud del artefacto a eliminar.

### **Paso 4 – Selección y filtrado de CI**

Como se ha planteado, se asume en este caso que una de las 69 CI contendrá a la fuente que genera el artefacto, siendo las otras 68 las fuentes cerebrales a preservar. Por tanto, se pretende detectar la  $CI_{art}$  que genera el artefacto.

Por otro lado, se asume que la  $CI_{art}$  será la que más energía aporte a las observaciones resultantes. Por tanto, el siguiente paso es calcular las señales que se obtienen al filtrar cada una de las componentes de las señales observadas. Es decir, si llamamos  $P_1 \dots P_{69}$  a las proyecciones de cada CI en las señales observadas, calcularemos:

$$SF_1 = SO - P_1$$

$$SF_2 = SO - P_1$$

...

$$SF_{69} = SO - P_{69}$$

Hay que tener en cuenta que cada una de las  $SF_x$  contendrá las 72 señales correspondientes a los 72 sensores con los que se capturó el EEG, pero en cada una de ellas se habrá eliminado el aporte de 1 de las 69 componentes independientes calculadas.

Para calcular cuál de las señales  $SF_x$  resultantes ha perdido más energía y por tanto comprobar en cuál se ha eliminado el artefacto, generamos 72 señales escalón (una por cada electrodo) en el intervalo del artefacto, cada una con amplitud igual al máximo valor absoluto leído en cada sensor (véase Figura 23).

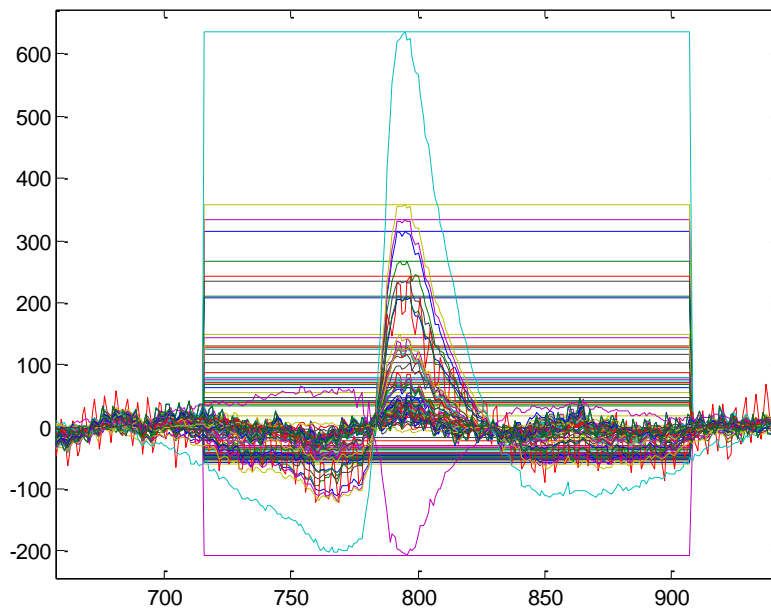


Figura 23 - Señal observada en los 72 sensores y escalones asociados

A continuación se calcula el EQM entre cada escalón y el electrodo asociado en cada una de las señales filtradas ( $SF_x$ ). Calculando el EQM total de cada señal filtrada como la suma de los EQM de todos sus electrodos, consideraremos que la señal sin artefacto será aquella que maximice el EQM (al haber desaparecido la CI que más energía aportaba).

La Figura 24 muestra el resultado. Puede observarse que efectivamente la amplitud de la señal filtrada se reduce de forma considerable, por lo que estamos eliminando la fuente que mayor energía genera. Sin embargo, se puede observar que el efecto en la limpieza del EEG no es ideal.

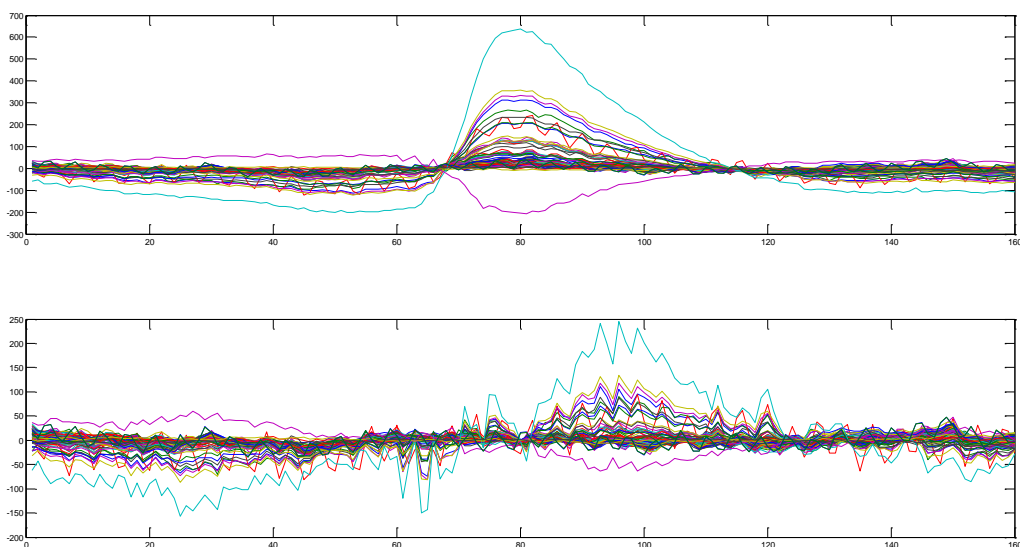


Figura 24 - Arriba: parpadeo observado en el EEG, Abajo: señal filtrada

### Continuación

Se deben repetir los pasos 2 a 4 hasta analizar y filtrar todos los intervalos de EEG en los que DETECT ha identificado un artefacto.

#### 4.2.6 Caso de estudio 6: Filtrado automático con supresión de varias CI

Partiendo del experimento anterior, en el que el nivel de filtrado no es el perseguido, se plantea en este caso la posibilidad de realizar una aproximación iterativa ya que se observa que parte del ruido se mantiene presente, por lo que se considera que existen otras CI que lo contienen.

Se repiten los Pasos 1, 2 y 3. Por tanto, como en el caso de estudio anterior, se usa DETECT para identificar zonas con artefacto en nuestra señal EEG, se selecciona una ventana en torno al artefacto que contenga 10 veces más muestras que las que componen el artefacto en sí, y se aplica FastICA para separar sus componentes principales. Llegado ese momento se restringe la ventana de trabajo a la zona en la que se encuentra el artefacto (según DETECT).

#### Paso 4

Manteniendo la nomenclatura anterior, tendremos:

- El EEG compuesto por las señales observadas SO en los 72 electrodos ( $SO_1 \dots SO_{72}$ ).
- Las Componentes Independientes Principales calculadas por ICA, de las que en este experimento se obtienen 66 ( $CI = CI_1 \dots CI_{66}$ ).

Se calcula el resultado de filtrar cada una de estas componentes de la señal original (calculando sus proyecciones en los 72 electrodos).

$$SF_1 = SO - P_1$$

$$SF_2 = SO - P_1$$

...

$$SF_{69} = SO - P_{69}$$

La diferencia en este experimento radica en la metodología para seleccionar las CI que incluyen el artefacto. En este caso la selección se basa de nuevo en el hecho de que serán las componentes que más energía al tramo de señal en el que DETECT identifica el ruido, pero en este caso se identifican aplicando otra técnica.

Se define la energía de la señal discreta  $x$  de  $n$  muestras como:

$$E \equiv \sum_{n=-\infty}^{\infty} |x(n)|^2$$

- Se calcula la energía de cada uno de los electrodos de la señal original:  $En_{SO_1}$ ,  $En_{SO_2}$  ...  $En_{SO_{72}}$ .
- Se calcula la energía de cada uno de los electrodos en cada una de las señales filtradas:
  - $En_{SF_{1-1}}$ ,  $En_{SF_{1-2}}$  ...  $En_{SF_{1-72}}$
  - $En_{SF_{2-1}}$ ,  $En_{SF_{2-2}}$  ...  $En_{SF_{2-72}}$
  - ...
  - $En_{SF_{66-1}}$ ,  $En_{SF_{66-2}}$  ...  $En_{SF_{66-72}}$
- Se calcula el cociente entre la energía de cada electrodo en cada una de las señales filtradas y el de la señal original:
  - $En_{SO_1}/En_{SF_{1-1}}$ ,  $En_{SO_2}/En_{SF_{1-2}}$  ...  $En_{SO_{72}}/En_{SF_{1-72}}$
  - ...
  - $En_{SO_1}/En_{SF_{66-1}}$ ,  $En_{SO_2}/En_{SF_{66-2}}$  ...  $En_{SO_{72}}/En_{SF_{66-72}}$

De este modo, seleccionando el cociente de mayor valor, detectaremos cuál es el electrodo que recoge una mayor variación de la energía tras el filtrado y se supone:

- Que se trata del electrodo más cercano a la fuente de ruido (lo que puede permitir, según su colocación, corroborar el ensayo)
- Que la componente filtrada en ese caso es la que incluye la mayor parte de la energía del artefacto.

A modo de ejemplo, se aplica este proceso sobre el primer artefacto de la señal EGG de test utilizada también en el caso anterior, que se trata de un parpadeo. El algoritmo implementado para este experimento identifica el electrodo 37 como el más cercano a la fuente del artefacto, y la 1ª Componente Independiente como aquella que más energía aporta a la señal observada (y por tanto la que incluye el artefacto).

La Figura 25 muestra la lectura de dicho electrodo en la señal original:

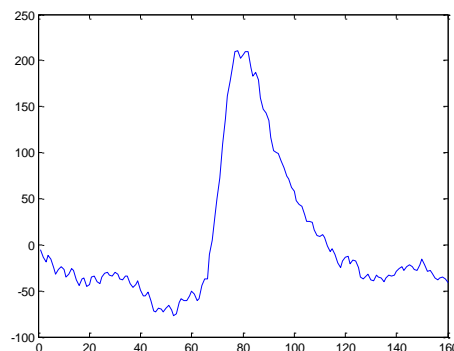


Figura 25 - Lectura del electrodo 37

Y la Figura 26 recoge la proyección de la CI 1 en dicho electrodo (es decir, la señal que vamos a filtrar en ese electrodo):

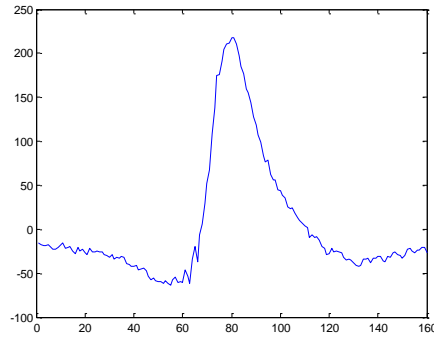


Figura 26 - Contribución de la CI 1 a la señal captada en el electrodo 37

Se puede apreciar que son bastante similares, aunque en la CI conseguimos disminuir el rizado que parece corresponder con una fuente del EEG. Si se elimina esta proyección de la señal original obtenida en el electrodo se obtiene la señal de la Figura 27.

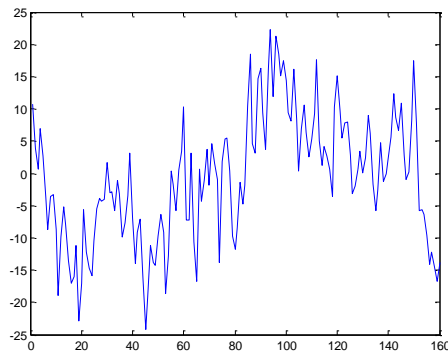


Figura 27 - Señal ya filtrada en el electrodo 37

El cálculo de la energía de esta señal (como suma de los cuadrados de sus muestras) da un total de  $1.5408e+04$ .

Para la comparación, se calcula la energía en el entorno cercano al artefacto en ese mismo electrodo. En este caso el artefacto tiene una longitud de 160 muestras, por lo que se seleccionan otras tantas del entorno cercano eliminando el artefacto en sí (véase Figura 28).

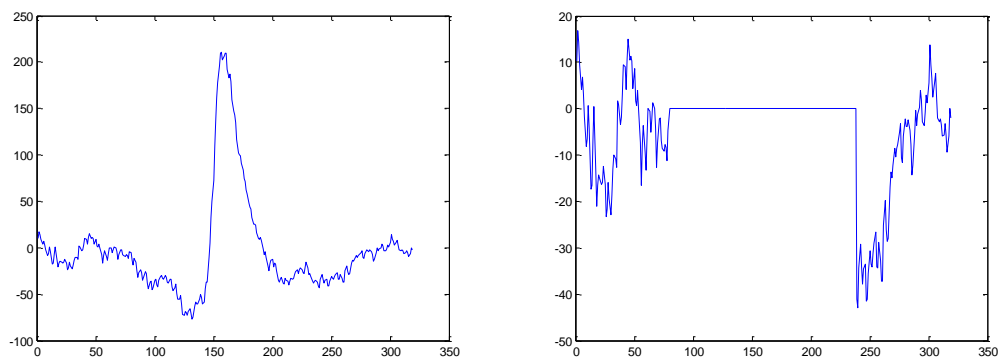


Figura 28 - Izda. Señal con contexto; Dcha. Contexto para el cálculo de la energía

En este caso la energía del contexto suma  $4.3333e+04$ , por tanto, la energía de referencia es mayor que la del segmento filtrado, y no se siguen eliminando CI. En la Figura 29, la secuencia de imágenes muestra, de izquierda a derecha:

- La señal original observada en los 72 sensores (superpuestos).
- La proyección de la CI seleccionada en los 72 sensores.
- La proyección de la CI en el electrodo identificado como el más cercano a la fuente.

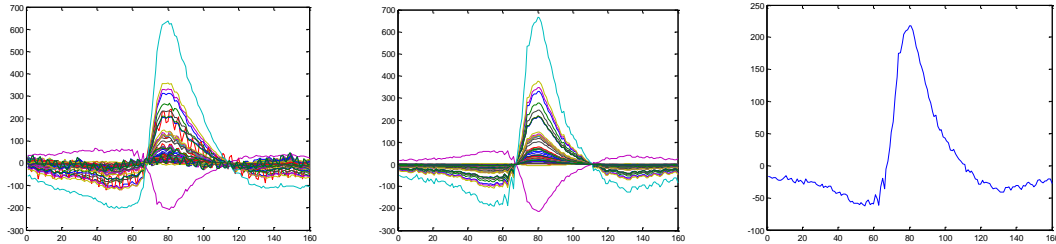


Figura 29 - Secuencia de identificación de CI y electrodo que más energía aportan a la señal

Llegados a este punto, se filtra la CI identificada de la señal original. La Figura 30 muestra el efecto de este filtrado:

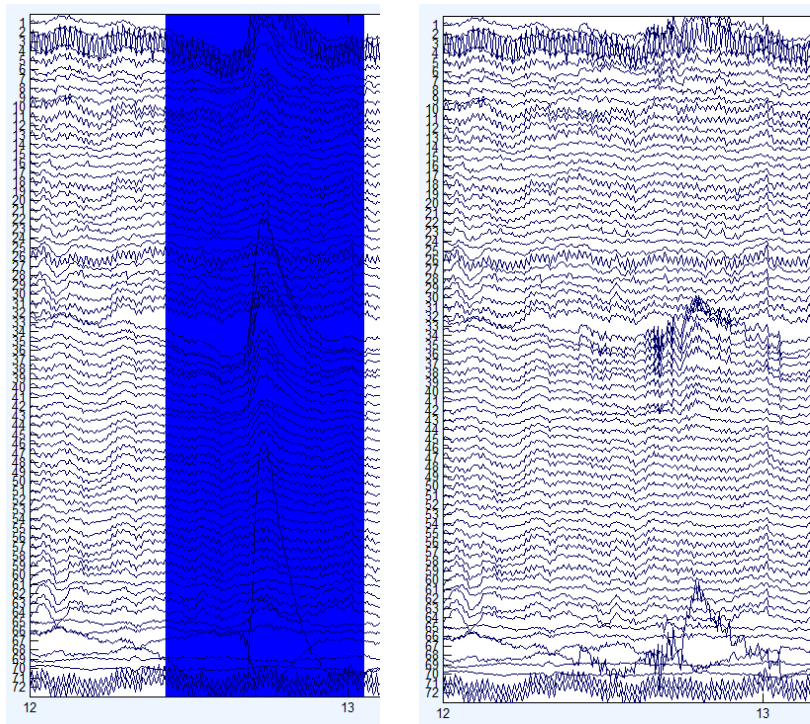


Figura 30 - Izda. Señal original. Dcha. Señal con la CI filtrada

#### **Paso 4 - Continuación**

Como se ha mencionado al comienzo del caso, se plantea aquí una **aproximación iterativa a la solución**. Es decir, se baraja la posibilidad de que la componente principal filtrada no contenga la totalidad del artefacto. Por tanto, se pretende analizar una característica de la señal EEG que nos permita seguir identificando artefacto en la misma.

Se escoge para ello de nuevo el nivel de energía en el electrodo identificado como el más cercano a la fuente de ruido. Para ello:

- Se calcula la energía de la señal observada en el electrodo identificado en el paso anterior, pero sobre la señal filtrada. Es decir, en ejemplo actual,  $En_{SF1-37}$ .
- Se calcula una energía de referencia ( $e_{ref}$ ) del entorno cercano al artefacto en ese mismo electrodo. En este caso tomamos un total de las mismas muestras del artefacto entre las disponibles justo antes y después del artefacto (este intervalo podría minimizarse normalizando la energía por muestra).

A continuación, mientras la energía medida en el electrodo 37 dentro del artefacto sea mayor que la energía del entorno del artefacto, se repite el paso 4 eliminando la siguiente CI que más energía aporte a la señal. En este caso, damos un 5% de tolerancia de diferencia entre ambas para parar la iteración (es decir, iteramos mientras  $En_{SF1-37} > e_{ref} * 1.05$ ).

Aquí cabe destacar que las CI independientes no se recalculan, si no que se parte del conjunto anterior, y que aquellas que ya han sido filtradas no se tienen en cuenta.

En el ejemplo anterior, tenemos que:

$e_{sf} =$

1.6089e+04

$e_{ref} =$

4.3333e+04

Por tanto no se elimina ninguna componente adicional.

### **Finalización del método**

El proceso anterior se debe repetir en cada uno de los artefactos identificador por DETECT en la señal EEG. Las imágenes de la Figura 31 muestran una progresión en la que la energía de referencia sí es menor que la energía de la zona del artefacto, y por tanto se realiza la aproximación iterativa filtrando una segunda componente en un artefacto introducido por un parpadeo ('Eye blink')

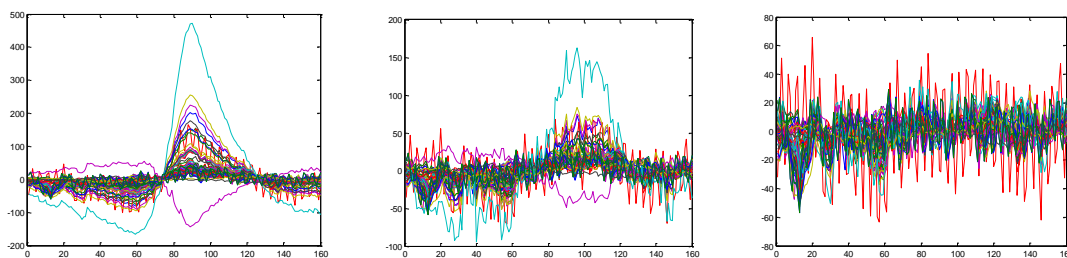


Figura 31 - De Izda a Dcha. Señal original. Componente n1 filtrada. Componente n2 filtrada

### Análisis de resultados

Revisando el efecto del procesado propuesto, se observan varios aspectos no deseables:

- Muchos de los artefactos siguen siendo identificados como tal por DETECT una vez procesado.
- En los extremos de los tramos de señal donde antes había artefactos, se observan ahora escalones producidos por el efecto del filtrado. Todo apunta a que la ventana en la que se incluye el artefacto es mayor que la indicada por DETECT (Figura 32).
- Existen artefactos consecutivos en los que la energía de referencia en el contexto del artefacto tratado se está tomando sobre otros artefactos colindantes, lo que puede generar errores en la elección de los electrodos de referencia y las CI a eliminar (Figura 33).
- El análisis ICA se aplica en ventanas fijas de 10 segundos, independientemente de los tipos de evento que se den dentro de dicha ventana. Para optimizar el análisis ICA es importante maximizar la información que se le ofrece al algoritmo, pero evitando que, por la tipología de las fuentes, se acaban vinculando estadísticamente una fuente de ruido y una de señal (ej. un parpadeo en un EEG que se produce al recibir un estímulo brillante para conocer la respuesta del cerebro).
- En determinados artefactos se observa que se toma como electrodo de referencia señales muy energéticas pero que, efectivamente, proceden de fuentes cerebrales (dado que se observa además como la observación en dicho electrodo se mantiene constante antes y después del tramo identificado como artefacto).
- En otros artefactos procesados, como el de la Figura 30, se observa como el filtrado no es del todo apropiado y quedan CI relacionadas con el ruido.

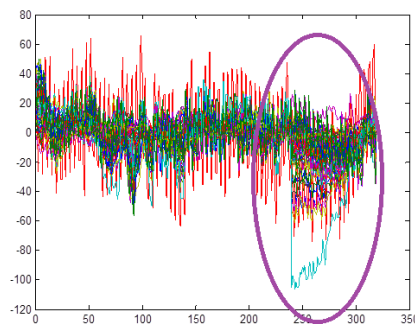


Figura 32 - Detalle del escalón producido en el margen de un artefacto procesado

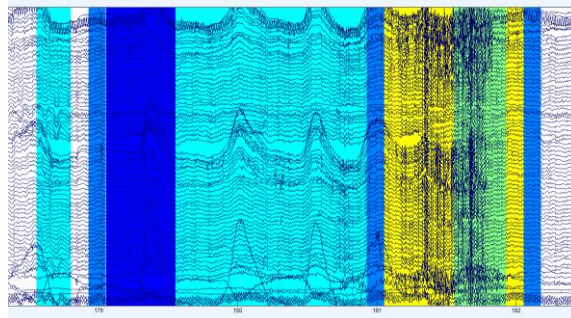


Figura 33 - Sección con varios artefactos consecutivos

En base a estos resultados se propone una evolución del método que se presenta en el siguiente caso.

#### 4.2.7 Caso de estudio 7: Evolución del filtrado automático con supresión de varias CI

Con el fin de evitar el escalón que se observa en los extremos de las ventanas procesadas. Se realiza una observación de las mismas y se concluye que DETECT no recoge con total precisión la ventana en la que se muestra todo el artefacto si no el tramo de sus características principales. Mediante observación se estima que es necesario ampliar 0.7 segundos esta ventana en ambos lados (véase Figura 34).

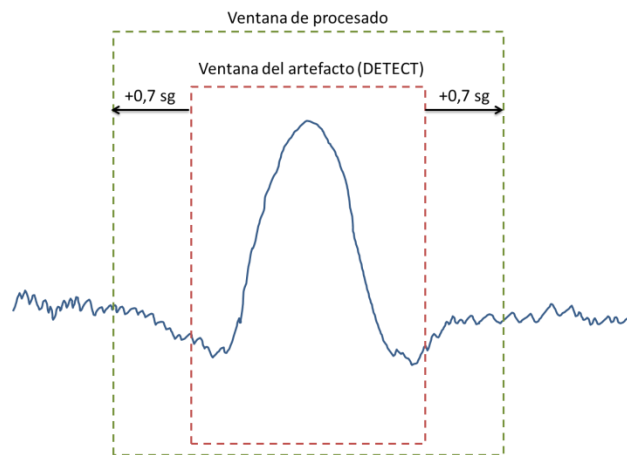


Figura 34 - Ventana de procesado de señal

Por otro lado, para evitar trabajar con electrodos en los que se capta una señal muy energética pero en los que pueda observarse que esta energía se da también antes y después del artefacto, se plantea un nuevo método de selección de electrodo y CI:

- Se calcula la energía de la señal que se está procesando (en la ventana de procesado) en cada electrodo (región A de la Figura 35).
- Se calcula la energía del contexto de dicho artefacto en cada electrodo (se plantean 2 opciones que corresponden con la región B y C de la Figura 35).
- Electrodo a electrodo se compara si la energía de contexto es menor que la energía en el tramo identificado como artefacto (en este punto aplicamos la tolerancia general del 5% ya comentada). De ser así, consideramos que en dicho electrodo se está observando el ruido.

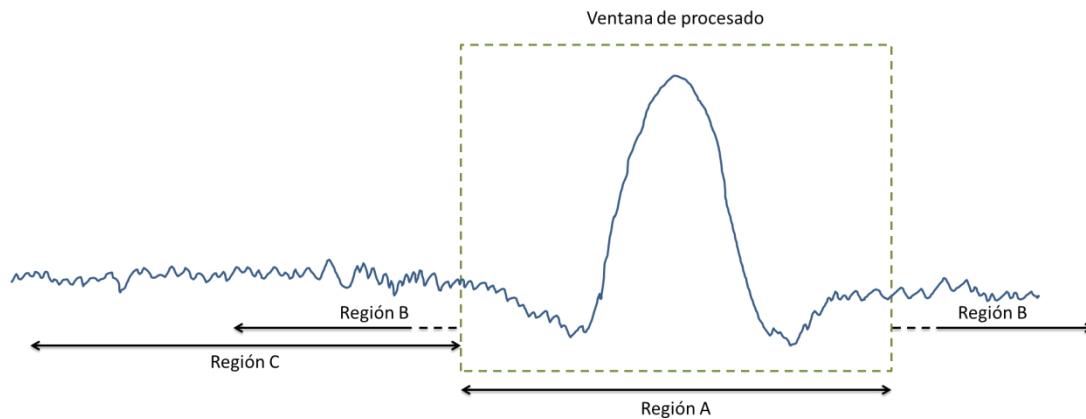


Figura 35 - Regiones para el cálculo de la energía

Una vez identificados los electrodos en los que se aprecia el ruido, se procede a identificar las CI que lo aportan. Para ello, se aplica el cálculo del caso anterior:

- Se calcula el cociente entre la energía de cada electrodo en el artefacto (de la señal que estamos tratando) y la energía en cada electrodo tras eliminar cada una de las CI calculadas. Es decir, se obtiene un valor por cada par electrodo/CI.
- Se toma el valor máximo de los anteriores. El valor máximo en este cociente indicará la CI que más energía aportaba. Este valor nos indicará por tanto una CI y un electrodo.
- En caso de que el electrodo indicado sea uno de los candidatos revisados en el paso anterior, tendremos la CI a eliminar. De lo contrario descartaríamos el valor y continuaríamos buscando el siguiente cociente máximo.

Finalizado este proceso, tendremos seleccionada la CI a filtrar y el electrodo más cercano a la fuente de ruido (lo que nos permite también posicionar físicamente a la fuente de ruido). Una vez eliminada la CI, se realiza un proceso iterativo para comprobar si existen más CI asociadas a la fuente de ruido. Para ello, de nuevo como en el caso de estudio 6:

- Se calcula la energía de la señal filtrada en la ventana de procesamiento del artefacto y en el electrodo seleccionado.
- Se calcula la energía de contexto en el mismo electrodo.
- Mientras la primera sea mayor a la segunda (de nuevo se puede aplicar un cierto margen), se calcula el cociente entre la energía de la señal que se está procesando en el electrodo en cuestión y dicha energía tras eliminar las CI restantes usando de nuevo el valor máximo como indicador de la CI a eliminar.

Como se ha planteado, la comparativa para descartar electrodos compara la energía en la región A con la energía en la región B (antes y después del artefacto) o la región C (exclusivamente antes del artefacto). Estas opciones surgen de 2 variantes de cara a resolver otro de los problemas detectados: el tratamiento de artefactos con otros colindantes que pueden verse afectados por estos que les rodean.

Se considera de nuevo la situación de la Figura 36. Dado que cada color representa un tipo de artefacto en la señal, y se encuentran todos colindando en una ventana de menos de 2 segundos, en el momento de procesar el artefacto incluido en la región A indicada la ventana

de procesado abarcará a otros. Además, a la hora de comprar energías ya sea tomando la región B o C estaremos comparando la energía tomando como referencia tramos de EEG con otros artefactos.

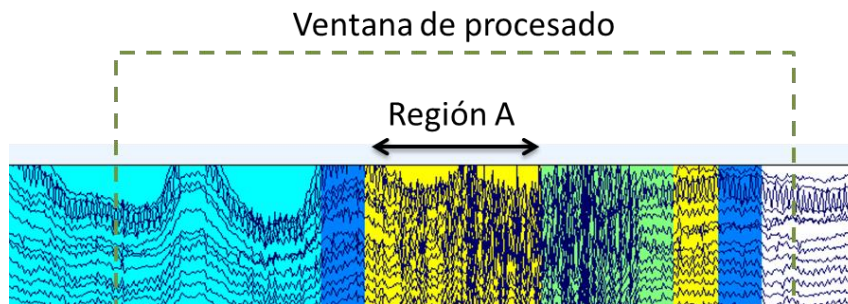


Figura 36 - Artefactos colindantes en un tramo de un EEG

En este punto se plantean 2 opciones:

- Opción 1. Agrupar los artefactos colindantes y tratarlos como un solo artefacto, caso en el que utilizaremos la región B para la comparativa energética (de forma que se tenga una media de la energía antes y después de la zona con ruido).
- Opción 2. Tratar los artefactos de forma independiente, caso en el que tomaremos la región C ya que consideraremos que a la izquierda del artefacto en cuestión no hay ruido, o si lo había lo hemos procesado.

La decisión anterior también va vinculada a la ventana que se procesa para realizar en análisis ICA. Como se ha planteado, ventanas más grandes de trabajo dan mejores resultados pero generan situaciones en las que, si se analizan secciones con varios artefactos de la misma naturaleza, estos pueden vincularse de algún modo a las fuentes de señal que pretendemos preservar y por tanto se perderán al filtrar las fuentes de ruido.

La opción 1 permite maximizar la ventana ICA al espacio disponible entre artefactos. Al unificar bajo un solo artefacto todos aquellos sucesos que se producen en un breve intervalo de tiempo, podemos extender la ventana ICA entre el bloque anterior y el siguiente.

En la opción 2 no es posible, por lo que se debe mantener una ventana ICA general que se mantiene, como hasta ahora, en 10 segundos de señal.

### *Opción 1*

En este experimento se amplía la ventana del artefacto indicada por DETECT en 0.7 segundos y se vinculan todos los artefactos que son colindantes. Es decir, que aquellos artefactos que, al extender la ventana 0.7 segundos antes y después, se solapan, se consideran un solo artefacto. Además se aplica el cálculo de energía en la región B, por lo que se tiene en cuenta la ventana de cálculo de dicha energía. Si el artefacto anterior o siguiente se introduce en la región B también se unifican.

En la Figura 37 se puede observar un ejemplo en el que el artefacto 1 y su artefacto anterior no se consideran solapados, pero sí el artefacto 1 y el posterior (que pasarán a considerarse un

solo artefacto). Sin embargo, esta fusión se hace de forma iterativa, por lo que en una segunda vuelta, dado que la región B tienen la mitad de muestras que la región A antes y después de esta, una vez el artefacto 1 se integre al posterior su región B crecerá y se solapará con el artefacto anterior, por lo que finalmente acabarán los 3 unidos.

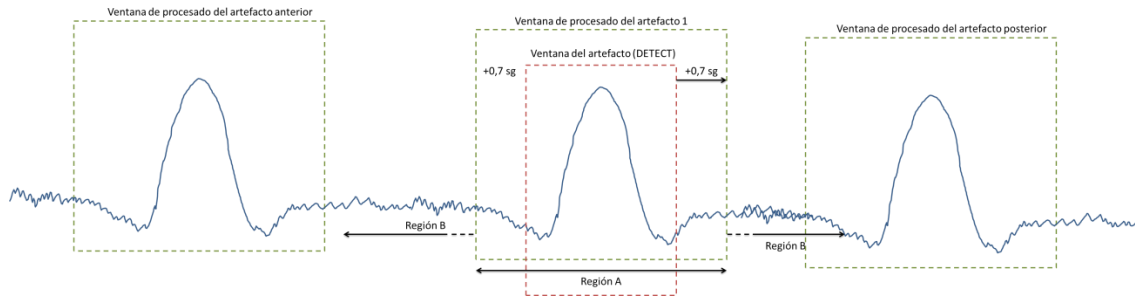


Figura 37 - Solapamiento de artefactos

En el listado de artefactos de la Figura 38 se ve cuáles se unen dentro de la señal de test que se utiliza en el experimento (hay que tener en cuenta que la etiqueta del tipo de artefacto queda meramente de referencia):

labelSet =			labelSet =		
'Eye Blink'	[ 12.4336]	[ 13.0547]	'Eye Blink'	[ 11.7336]	[ 13.7547]
'Eye Left Movement'	[ 16.3086]	[ 17.0547]	'Eye Left Movement'	[ 15.6086]	[ 17.7547]
'Eye Blink'	[ 19.8086]	[ 20.3047]	'Eye Up Movement'	[ 19.1086]	[ 21.1297]
'Eye Up Movement'	[ 20.3086]	[ 20.4297]	'Eye Blink'	[ 25.1086]	[ 27.1297]
'Eye Blink'	[ 25.8086]	[ 26.4297]	'Eye Blink'	[ 40.2336]	[ 42.2547]
'Eye Blink'	[ 40.9336]	[ 41.5547]	'Eye Blink'	[118.3586]	[120.2547]
'Eye Blink'	[119.0586]	[119.5547]	'Eye Blink'	[172.7336]	[185.8797]
'Eye Blink'	[173.4336]	[174.0547]	'Eye Blink'	[273.4836]	[275.5047]
'Eye Blink'	[175.1836]	[175.9297]	'Eye Blink'	[284.4836]	[286.3797]
'Eye Up Movement'	[178.5586]	[178.8047]	'Eye Blink'	[296.2336]	[298.1297]
'Eye Left Movement'	[178.9336]	[179.0547]	'Eye Blink'	[349.3586]	[351.3797]
'Eye Blink'	[179.0586]	[179.5547]	'Eye Blink'	[354.8586]	[356.8797]
'Eye Up Movement'	[179.5586]	[180.9297]	'Eye Up Movement'	[361.2336]	[365.0047]
'Eye Left Movement'	[180.9336]	[181.0547]	'Eye Blink'	[446.1086]	[448.2547]
'Jaw Movement'	[181.0586]	[181.5547]			
'Jaw Clench'	[181.5586]	[181.9297]			
'Jaw Movement'	[181.9336]	[182.0547]			
'Eye Left Movement'	[182.0586]	[182.1797]			
'Eye Blink'	[184.5586]	[185.1797]			
'Eye Blink'	[274.1836]	[274.8047]			
'Eye Blink'	[285.1836]	[285.6797]			
'Eye Blink'	[296.9336]	[297.4297]			
'Eye Blink'	[350.0586]	[350.6797]			
'Eye Blink'	[355.5586]	[356.1797]			
'Eye Blink'	[361.9336]	[362.6797]			
'Eye Up Movement'	[362.6836]	[362.8047]			
'Eye Blink'	[362.8086]	[363.1797]			
'Eye Up Movement'	[363.1836]	[363.3047]			
'Eye Left Movement'	[363.3086]	[363.6797]			
'Eye Up Movement'	[363.6836]	[364.3047]			
'Eye Blink'	[446.8086]	[447.5547]			

Figura 38 - Izda. Lista de artefactos original. Dcha. Lista artefactos extendidos y unificados

Dado el posicionamiento de los artefactos agrupados, la ventana para el análisis ICA variará entre unos pocos segundos (por encima de 5) y varias decenas. A continuación se detalla el proceso de trabajo con el primer artefacto

### Artefacto 1 – Eye Blink

Trabajaremos con el primer artefacto (parpadeo) mostrado por DETECT, el mostrado en la Figura 39.

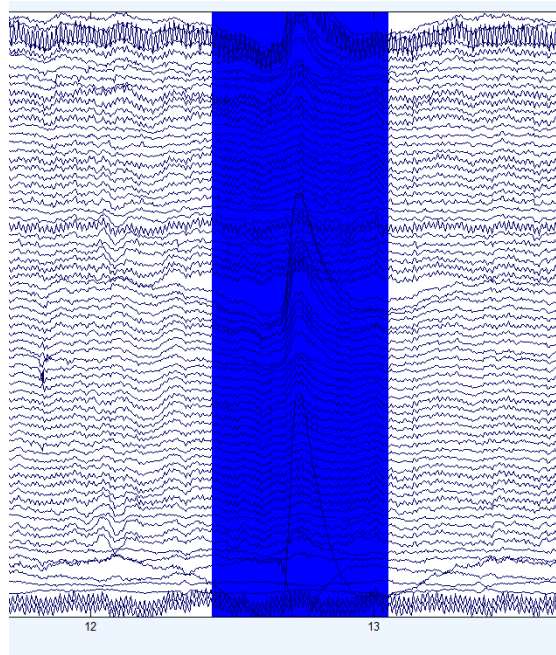
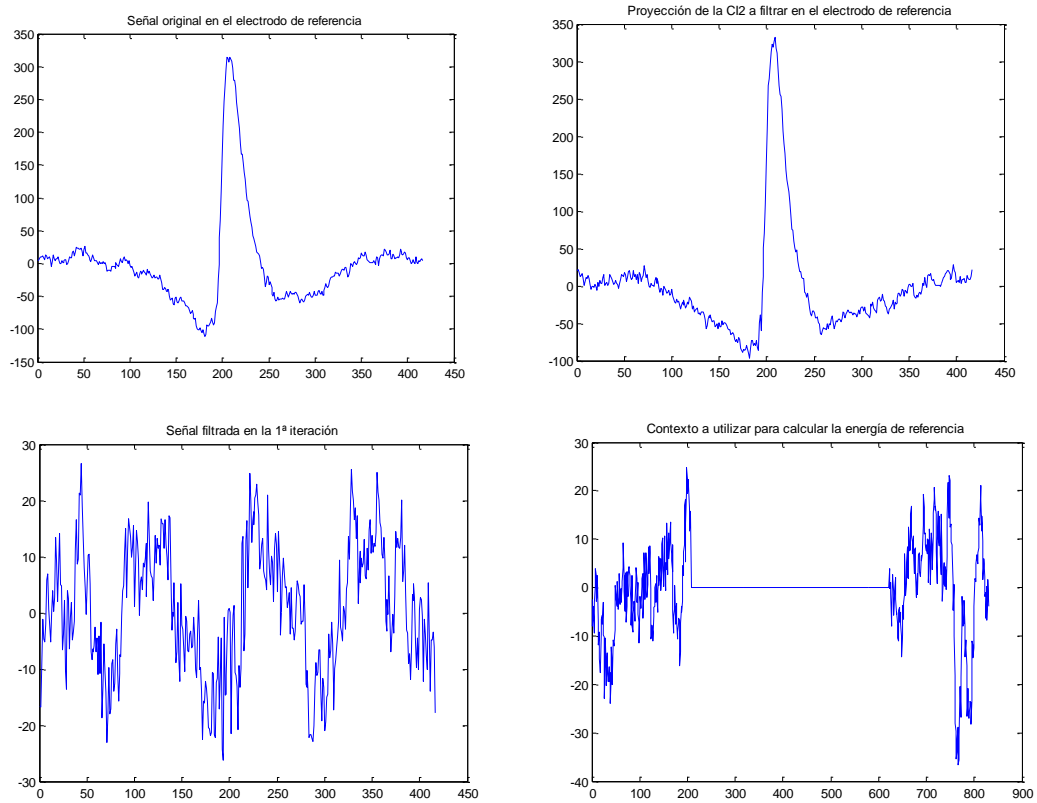


Figura 39 - Parpadeo en la señal original

El algoritmo arroja los siguientes resultados:



Comparando las energías, no se realiza una segunda iteración. La Figura 40 muestra el resultado del proceso.

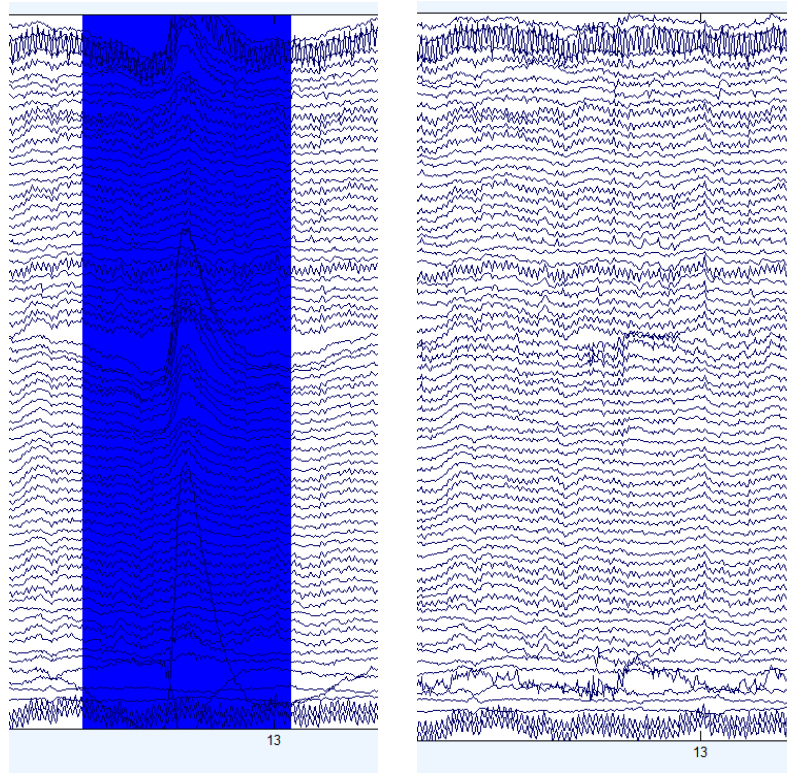


Figura 40 - Izda. Señal original; Dcha. Señal filtrada

Como puede observarse, en este caso la calidad del filtrado aumenta de forma considerable con respecto al caso anterior.

### **Artefacto 3**

Se considera ahora el artefacto agrupado en la 3ª posición, que recoge un parpadeo y una movimiento hacia arriba del ojo consecutivos (Figura 41).

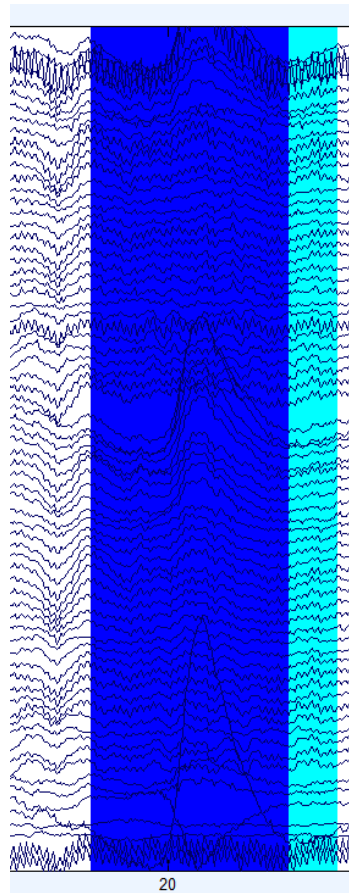
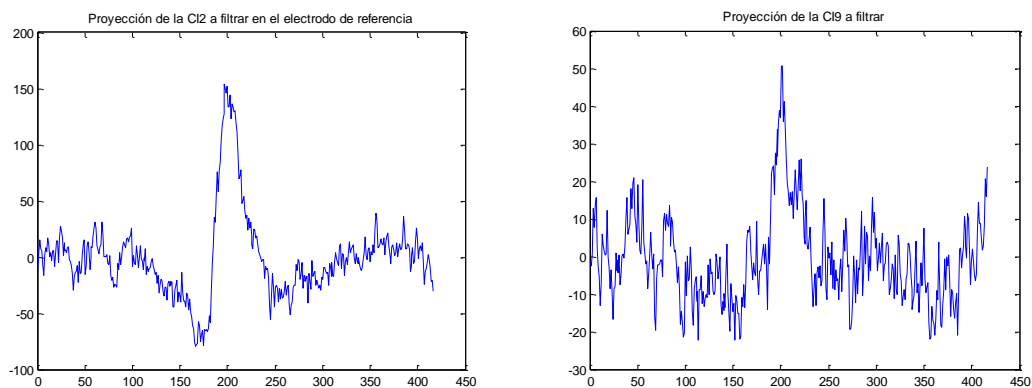
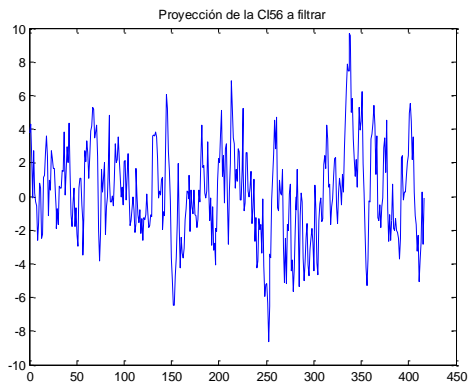
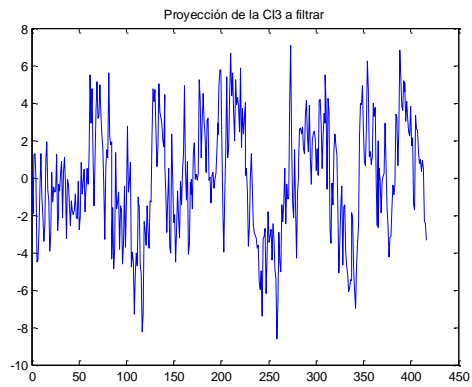
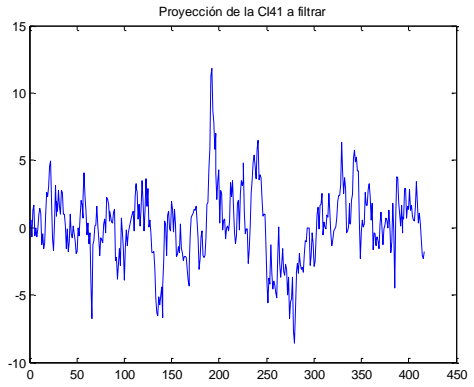
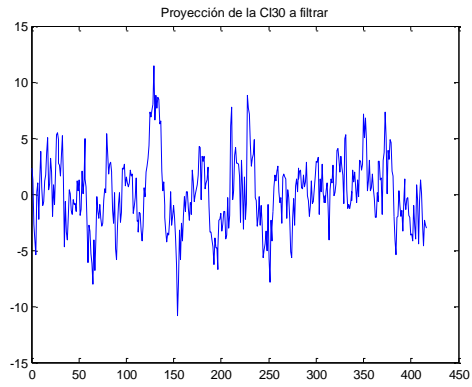
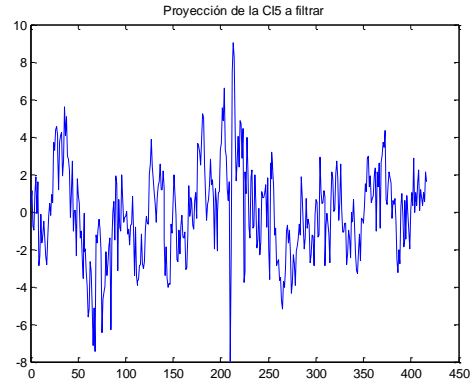
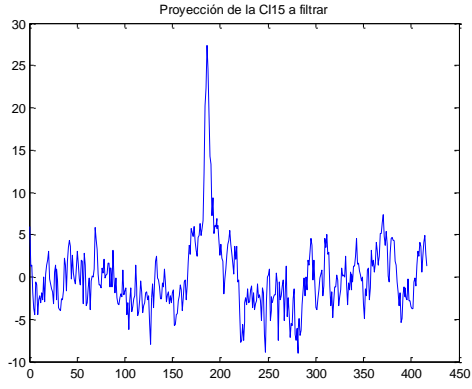


Figura 41 - Artefactos colindantes en la señal original.

En este caso, tras el filtrado de la primera componente identificada como ruido, sí existe mayor energía en el tramo del artefacto que en su contexto por lo que sí se considera que existen más CI relacionadas con las fuentes de ruido (lo que concuerda con el hecho de que existen, de hecho, diferentes tipos de ruido en el mismo tramo). En total en el proceso se eliminan las 12 CI que se muestra en la Figura 42.





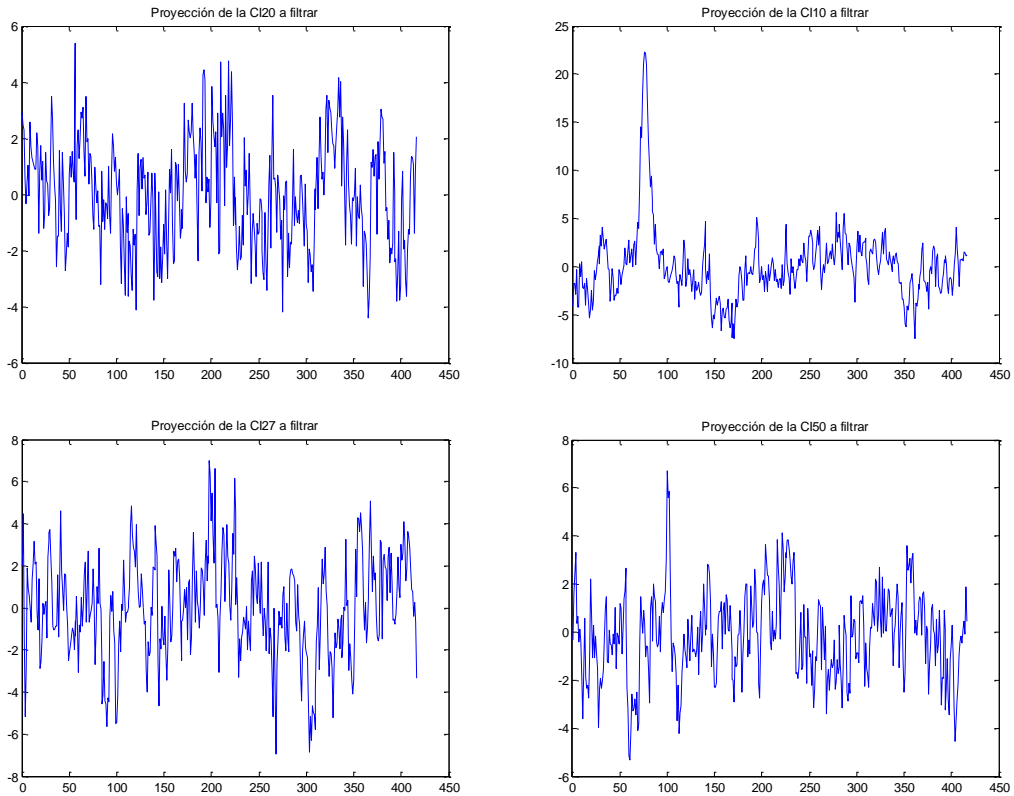


Figura 42 - Diferentes CI eliminadas de la señal

Finalmente, tras el filtrado de las 12 CI anteriores, se obtiene la siguiente señal de salida de la Figura 43, que DETECT considera libre de artefactos.

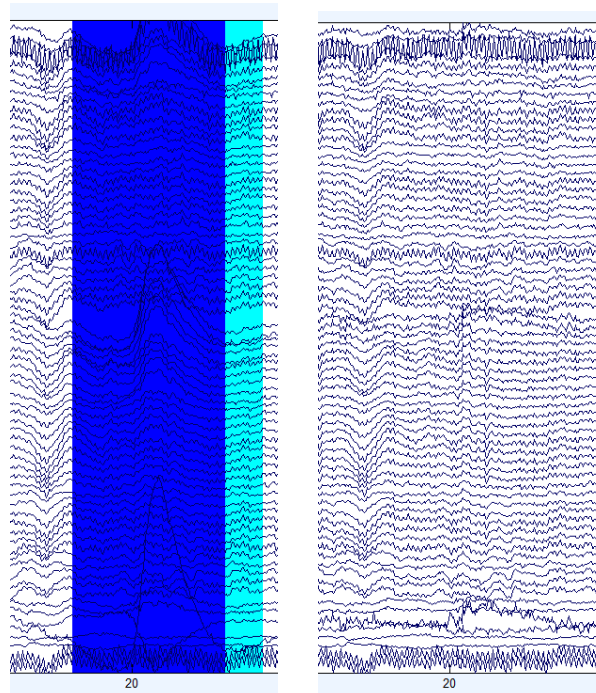


Figura 43 - Izda. Artefacto 3 en la señal original; Dcha. Señal filtrada

### Resto de artefactos

Si aplicamos el método a todos los artefactos, el cuadro de la Figura 44 recoge lo que sucede a cada uno de los artefactos detectados originalmente (tabla de la izquierda), los artefactos agrupados que se analizan (tabla central) y los artefactos que se detectan tras el procesado (tabla derecha).

labelSet =		labelSet =		labelSet =	
'Eye Blink'	[ 12.4936] [ 19.6547]	'Eye Blink'	[ 11.7336] [ 13.7547]	'Eye Left Movement'	[ 16.3086] [ 16.8047]
'Eye Left Movement'	[ 16.3086] [ 17.0547]	'Eye Left Movement'	[ 15.6086] [ 17.7547]	'Eye Left Movement'	[ 179.0586] [ 179.1797]
'Eye Blink'	[ 19.6066] [ 20.9047]	'Eye Up Movement'	[ 19.1086] [ 21.1297]	'Eyebrow Movement'	[ 180.8086] [ 181.0547]
'Eye Up Movement'	[ 20.9086] [ 20.4297]	'Eye Blink'	[ 25.1086] [ 27.1297]	'Jaw Clench'	[ 181.0586] [ 182.0547]
'Eye Blink'	[ 25.8086] [ 26.4297]	'Eye Blink'	[ 40.2336] [ 42.2547]	'Eye Left Movement'	[ 362.9336] [ 363.1797]
'Eye Blink'	[ 40.9336] [ 41.5547]	'Eye Blink'	[ 118.3586] [ 120.2547]	'Jaw Clench'	[ 363.1836] [ 363.3047]
'Eye Blink'	[ 119.0586] [ 119.5547]	'Eye Blink'	[ 172.7336] [ 185.8797]	'Eye Left Movement'	[ 363.1836] [ 363.3047]
'Eye Blink'	[ 179.1836] [ 179.6547]	'Eye Blink'	[ 273.4836] [ 275.5047]	'Jaw Clench'	[ 363.1836] [ 363.3047]
'Eye Blink'	[ 179.1836] [ 179.6547]	'Eye Blink'	[ 284.4836] [ 286.3797]	'Eye Left Movement'	[ 363.3086] [ 364.0547]
'Eye Up Movement'	[ 179.5586] [ 179.8047]	'Eye Blink'	[ 296.2336] [ 298.1297]		
'Eye Left Movement'	[ 179.9336] [ 179.6547]	'Eye Blink'	[ 349.3586] [ 351.3797]		
'Eye Blink'	[ 179.0586] [ 179.5547]	'Eye Blink'	[ 354.8586] [ 356.8797]		
'Eye Up Movement'	[ 179.5586] [ 180.9297]	'Eye Up Movement'	[ 361.2336] [ 365.0047]		
'Eye Left Movement'	[ 180.9336] [ 181.0547]	'Eye Blink'	[ 446.1086] [ 448.2547]		
'Jaw Movement'	[ 181.0586] [ 181.5547]				
'Jaw Clench'	[ 181.5586] [ 181.9297]				
'Jaw Movement'	[ 181.9336] [ 182.0547]				
'Eye Left Movement'	[ 182.0586] [ 182.1797]				
'Eye Blink'	[ 184.5586] [ 185.1797]				
'Eye Blink'	[ 274.1836] [ 274.6547]				
'Eye Blink'	[ 285.1836] [ 285.6547]				
'Eye Blink'	[ 296.9336] [ 297.4297]				
'Eye Blink'	[ 350.0586] [ 350.6297]				
'Eye Blink'	[ 355.5586] [ 356.1797]				
'Eye Blink'	[ 361.9336] [ 362.6297]				
'Eye Up Movement'	[ 362.6296] [ 362.8047]				
'Eye Blink'	[ 362.8086] [ 363.1797]				
'Eye Up Movement'	[ 363.1836] [ 363.3047]				
'Eye Left Movement'	[ 363.3086] [ 363.6797]				
'Eye Up Movement'	[ 363.6836] [ 364.3047]				
'Eye Blink'	[ 446.8086] [ 447.5547]				

Figura 44 – Resultado del procesado de los artefactos

En este punto, se consiguen eliminar el 74,19% de los artefactos iniciales. Además, se observa que se elimina el 100% de los artefactos de parpadeo, con gran aporte energético, y como en algunos casos un artefacto da paso a otro en la misma ventana temporal o similar.

De hecho, de los más de 480 segundos de EEG en los que aparecen en total 15.1289 sg de artefactos, el ruido remanente se reduce a 4 intervalos de 0.5 segundos, 0.1, 1.2 y 1.1 segundos aproximadamente, es decir, menos de un 20% del ruido inicial. La Figura 45 muestra el resultado en la zona de mayor ruido persistente.

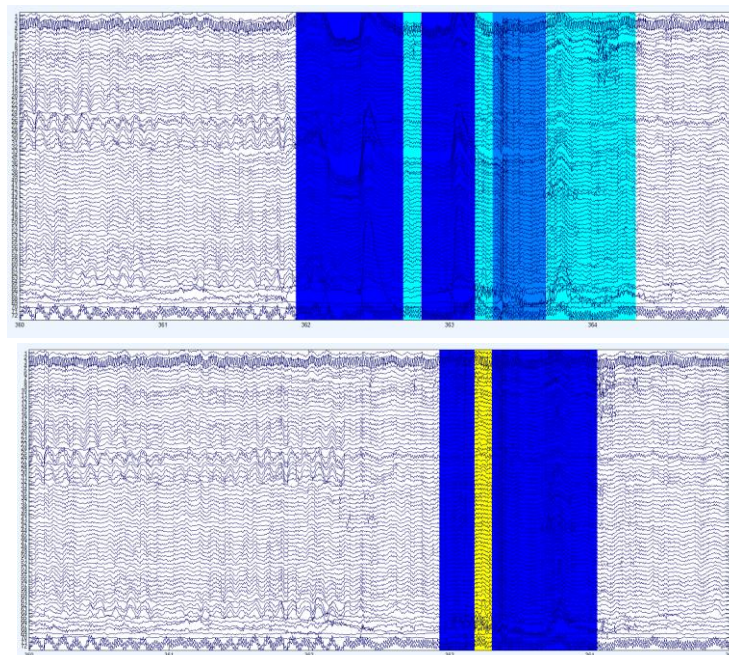


Figura 45 - Zona de ruido persistente

## Opción 2

En esta variante el proceso es el mismo salvo en 3 aspectos:

- Los artefactos se tratan de forma independiente, incluso si son colindantes o cuando la ampliación de la ventana de procesamiento los solapa.
- Por tanto, dado que no hay seguridad de que a la derecha de nuestro artefacto no haya otro ruido, la comparativa con energía de contexto se realiza a la izquierda del artefacto (región C de la Figura 35).
- Dado que la ventana para aplicar ICA no puede extenderse entre artefactos (porque habría ocasiones en las que solo se contaría con una décimas de segundo en caso de artefactos breves rodeados de otros artefactos), se utiliza siempre una ventana de 10 segundos centrada en el propio artefacto.

La Figura 46 muestra la comparación entre los resultados de un método y otro:

labelSet =			labelSet =		
'Eye Left Movement'	[ 16.3086]	[ 16.8047]	'Eye Left Movement'	[ 16.3086]	[ 16.4297]
'Eye Left Movement'	[179.0586]	[179.1797]	'Eye Left Movement'	[ 16.5586]	[ 16.8047]
'Eyebrow Movement'	[180.8086]	[181.0547]	'Eye Left Movement'	[ 40.6836]	[ 40.8047]
'Jaw Clench'	[181.0586]	[182.0547]	'Eye Left Movement'	[173.6836]	[173.8047]
'Eye Left Movement'	[362.9336]	[363.1797]	'Eye Left Movement'	[174.0586]	[174.1797]
'Jaw Clench'	[363.1836]	[363.3047]	'Eye Left Movement'	[175.1836]	[175.3047]
'Eye Left Movement'	[363.3086]	[364.0547]	'Eye Blink'	[175.3086]	[176.0547]
			'Eye Left Movement'	[180.3086]	[180.4297]
			'Eyebrow Movement'	[180.4336]	[180.9297]
			'Jaw Clench'	[180.9336]	[181.8047]
			'Jaw Movement'	[181.8086]	[182.0547]
			'Eye Left Movement'	[182.0586]	[182.6797]
			'Eye Left Movement'	[274.1836]	[274.5547]
			'Eye Left Movement'	[349.8086]	[350.1797]
			'Eye Left Movement'	[350.4336]	[350.5547]
			'Eye Left Movement'	[354.6836]	[354.8047]
			'Eyebrow Movement'	[362.8086]	[362.9297]
			'Jaw Clench'	[362.9336]	[363.3047]
			'Eyebrow Movement'	[363.4336]	[363.5547]
			'Eye Left Movement'	[363.5586]	[364.0547]

Figura 46 - izqda. Resultados Opción 1; Dcha. Resultados Opción 2

A priori se presenta más efectivo el tratamiento conjunto de artefactos, principalmente por la flexibilidad y maximización de la ventana para aplicar ICA.

### 4.2.8 Caso de estudio 8: Filtrado iterativo para artefactos persistentes

A raíz de los resultados del caso anterior, cabe preguntarse si tras un primer filtrado en el que se obtienen a la salida unos artefactos ligeramente modificados, la realidad puede ser que existen artefactos que quedan ocultos por otros de mayor energía. De ser el caso, un procesamiento iterativo de la salida del primer filtrado podría llevar a limpiar por completo la señal de entrada.

Se plantea esta metodología con las 2 opciones implementadas en el caso de estudio 7.

### Opción 1

Para comenzar el experimento, vamos a centrarnos en el intervalo 150 – 200 segundos de la señal original, que pasa a ser el tramo 0-50, y que incluye los artefactos etiquetados por DETECT que se muestran en la Figura 47.

```
labelSet =
    'Eye Blink'           [18.4336]    [19.0547]
    'Eye Blink'           [20.1836]    [20.9297]
    'Eye Up Movement'    [23.5586]    [23.8047]
    'Eye Left Movement'  [23.9336]    [24.0547]
    'Eye Blink'           [24.0586]    [24.5547]
    'Eye Up Movement'    [24.5586]    [25.9297]
    'Eye Left Movement'  [25.9336]    [26.0547]
    'Jaw Movement'       [26.0586]    [26.5547]
    'Jaw Clench'         [26.5586]    [26.9297]
    'Jaw Movement'       [26.9336]    [27.0547]
    'Eye Left Movement'  [27.0586]    [27.1797]
    'Eye Blink'           [29.5586]    [30.1797]
```

Figura 47 - Artefactos en el intervalo de estudio

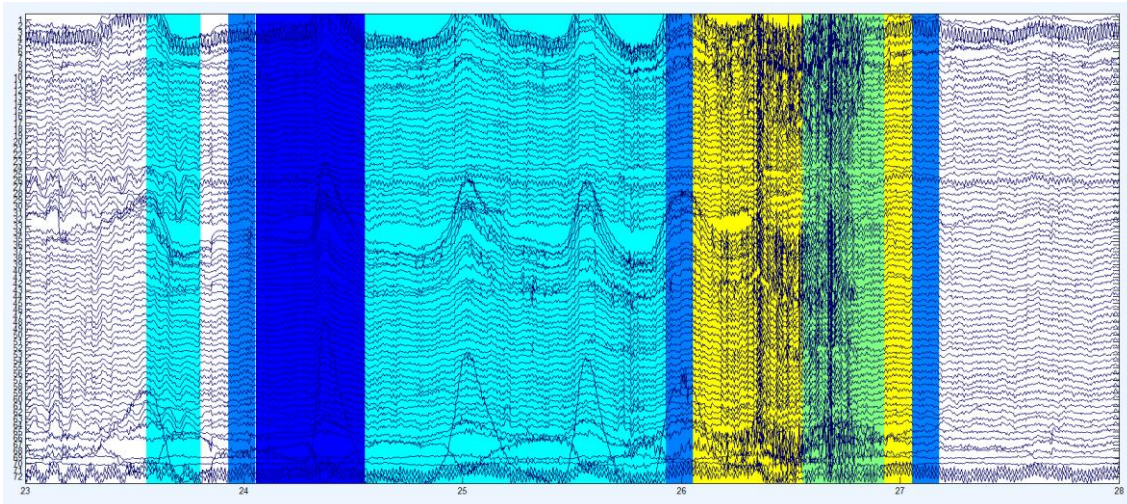


Figura 48 - Intervalo donde se concentran la mayor parte de los artefactos persistentes

Mantenemos el método anterior, estableciendo de nuevo la tolerancia del 5% (para las comparativas de niveles de energía) y un margen en las ventanas de 0.7 segundos a cada lado (para asegurarnos de abarcar los artefactos completos).

En cada iteración vamos a comprobar para cada artefacto qué electrodo se selecciona en cada artefacto tratado, y vamos a comprobar si se procesa de alguna manera y cómo evoluciona. Dado que los artefactos irán desapareciendo, transformándose o cambiando su ventana, se analizan los artefactos en bloque.

#### **1 iteración**

Al tratarse de artefactos colindantes, se procesan todos en un único bloque. El sistema elige el electrodo 65 y comienza a eliminar CI (8 en total). A la salida se mantienen presentes los siguientes artefactos:

'Eye Left Movement'	[24.0586]	[24.1797]
'Eye Left Movement'	[28.9336]	[29.4297]
'Eyebrow Movement'	[29.5586]	[29.8047]
'Jaw Clench'	[29.8086]	[30.3047]
'Eyebrow Movement'	[30.4336]	[30.6797]
'Jaw Clench'	[30.8086]	[30.9297]
'Jaw Movement'	[30.9336]	[31.3047]
'Jaw Clench'	[31.3086]	[31.8047]
'Jaw Movement'	[31.8086]	[32.0547]

Figura 49 - Artefactos persistentes tras la 1ª iteración

Vemos que la mayor parte de los artefactos en el intervalo 23-28 segundos desaparecen, pero surge una mayor concentración de los mismos en el intervalo 28-33.

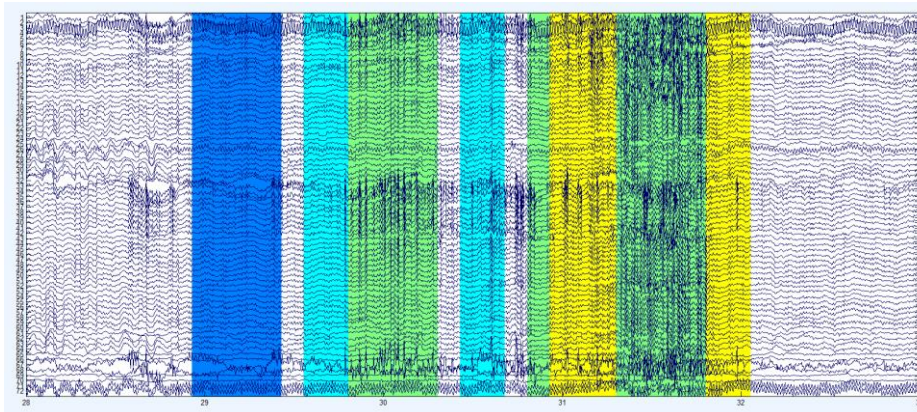


Figura 50 - Artefactos en la ventana 28-33 segundos tras la 1ª iteración

### Tras 9 iteraciones

Siguen existiendo artefactos. Algunos sí desaparecen, pero otros siguen extendiéndose ligeramente:

'Eye Left Movement'	[29.9336]	[30.0547]
'Eyebrow Movement'	[30.5586]	[30.8047]
'Eye Left Movement'	[30.9336]	[31.0547]
'Jaw Movement'	[31.0586]	[31.3047]
'Jaw Clench'	[31.3086]	[31.8047]
'Jaw Movement'	[31.8086]	[32.0547]
'Eye Left Movement'	[32.0586]	[32.3047]
'Eye Left Movement'	[32.4336]	[32.5547]

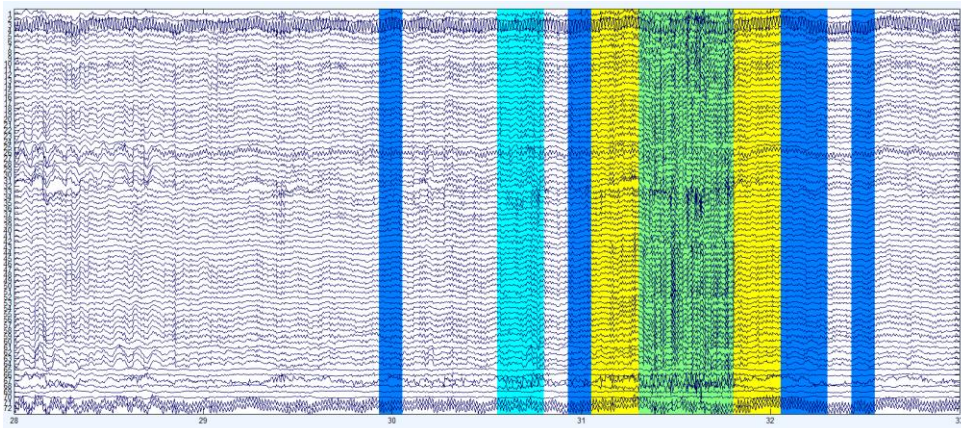


Figura 51 - Artefactos persistentes tras la 9ª iteración

La siguiente matriz indica los electrodos tomados como referencia para cada iteración (columna) y artefacto (fila).

```

65   66   44   69   68   66   24   66   3
0    15   9    0    0    0    0    0    0

```

Se puede concluir que en las iteraciones 2 y 3 los artefactos que restan se encuentran lo suficientemente separados para tratarlos de forma independiente. También se observa como el algoritmo procesa de forma insistente el electrodo 66 (lo que puede indicar un problema en el criterio de parada en la selección de CIs a eliminar).

### Tras 72 iteraciones

```

labelSet =
'Eye Left Movement' [29.0586] [29.3047]
'Eye Left Movement' [29.5586] [29.8047]
'Jaw Movement' [29.8086] [30.1797]
'Jaw Clench' [30.1836] [30.3047]
'Eyebrow Movement' [30.3086] [30.8047]
'Jaw Movement' [30.8086] [31.1797]
'Jaw Clench' [31.1836] [31.9297]
'Jaw Movement' [31.9336] [32.0547]
'Eye Left Movement' [32.0586] [32.8047]

```

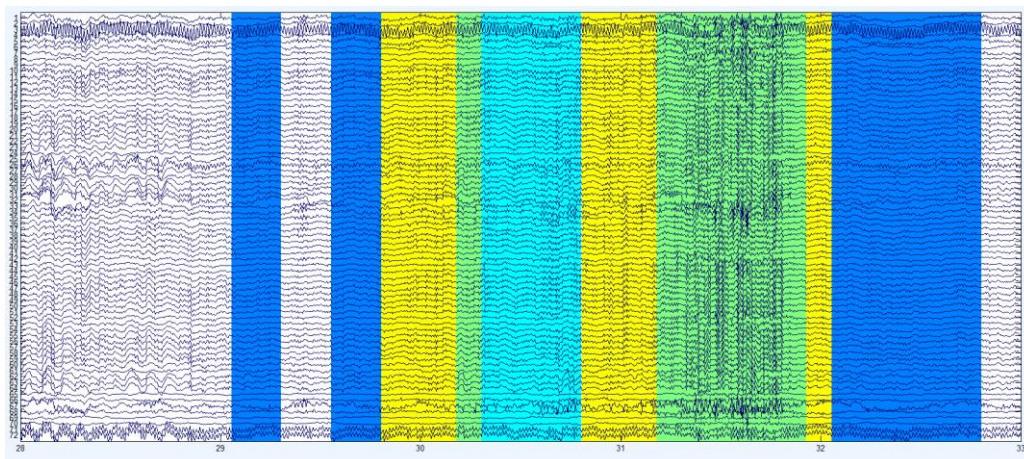


Figura 52 - Artefactos persistentes tras la 72ª iteración

Tras un número cuantioso de iteraciones, se obtienen los resultados de la Figura 54. A continuación se muestran los electrodos sobre los que se trabaja en cada iteración:

```

Columns 1 through 22
65   66   44   69   68   66   24   66   3   49   8   24   66   8   70   70   0   0   0   0   0
0    15   9    0    0    0    0    0    0    0    0    0    0    0    0    0    0   0   0   0   0

Columns 23 through 44
0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0

Columns 45 through 66
0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0

Columns 67 through 71
0    0    0    0    0
0    0    0    0    0

```

Como se puede apreciar, existe evolución en la señal pero los artefactos no solo no desaparecen si no que aumentan paulatinamente su extensión (ya que el procesado va afectando al entorno de los artefactos). Además, desde la iteración nº 16 deja de procesarse la señal ya que no existe ningún electrodo que cumpla las condiciones para seleccionarlo como referencia a la hora de comenzar a decidir las CI a filtrar.

### Opción 2

Centramos de nuevo el experimento en el mismo intervalo (150 – 200 segundos) de la señal original. De nuevo, aplicamos ICA sobre ventanas de 10 segundos y trabajamos los artefactos uno a uno en su intervalo ampliado (0.7 segundos a cada lado).

En este caso, tras las primera 9 iteraciones, se obtiene la siguiente matriz en la que cada fila corresponde a 1 artefacto, mientras que cada columna a una iteración. Hay que tener en cuenta que:

- Si un artefacto desaparece y por tanto pasamos de tener de los 12 iniciales a una cifra inferior, la matriz seguirá rellenando las filas de la iteración con 0's hasta las 12 filas.
- También se identificará el electrodo como 0 cuando al procesar un artefacto no se detecte ningún electrodo que cumpla las condiciones para su selección (que la energía en la región A sea mayor que en C).

		Iteraciones								
		→								
Artefactos ↓	65	34	67	67	67	67	68	67	67	
	7	68	59	67	0	71	0	0	0	
	2	69	71	60	61	72	3	3	15	
	36	69	69	0	61	53	0	0	3	
	36	10	7	14	6	71	71	45	0	
	67	7	0	16	0	6	46	45	0	
	65	0	0	0	0	0	32	17	53	
	15	0	0	0	0	0	0	0	15	
	9	0	0	0	0	0	0	0	71	
	65	0	0	0	0	0	0	0	0	
	42	0	0	0	0	0	0	0	0	
	67	0	0	0	0	0	0	0	0	

Figura 53 - Matriz de electrodo escogidos en cada iteración

De la Figura 53 se puede concluir:

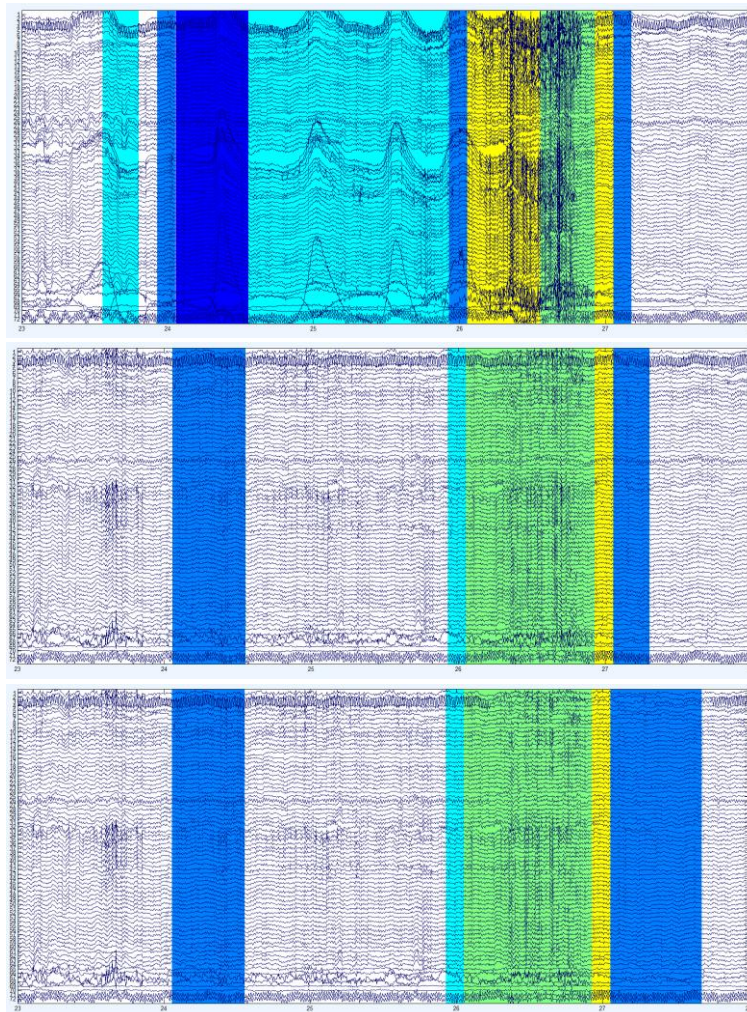
- Que tras 9 iteraciones siguen existiendo artefactos presentes
- Que en todos los puntos indicados en rojo, existe un artefacto que no se trata al no haber ningún electrodo que cumpla las condiciones de filtrado.

Tras las 9 primeras iteraciones, siguen presentes los siguientes artefactos:

'Eye Left Movement'	[23.3586]	[25.5047]
'Eyebrow Movement'	[24.1086]	[25.7547]
'Jaw Clench'	[24.3586]	[26.0047]
'Eyebrow Movement'	[25.1086]	[26.6297]
'Jaw Clench'	[25.2336]	[27.6297]
'Jaw Movement'	[26.2336]	[27.7547]
'Eyebrow Movement'	[26.3586]	[27.8797]
'Eye Left Movement'	[26.4836]	[28.1297]
'Eyebrow Movement'	[26.7336]	[28.6297]

Figura 54 - Artefactos persistentes

Las siguientes imágenes muestran la evolución de los artefactos en estas primeras iteraciones entre los segundos 23 y 28 (donde se concentran los artefactos persistentes):



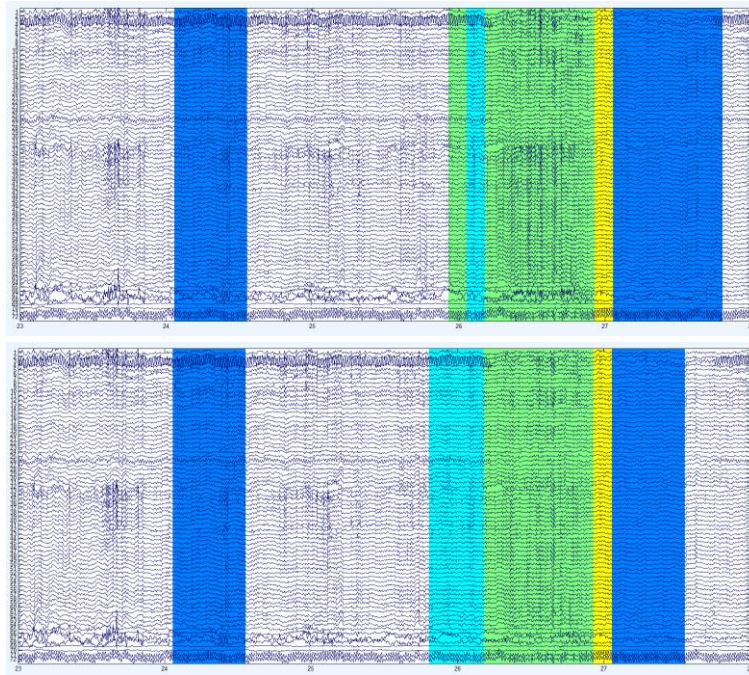
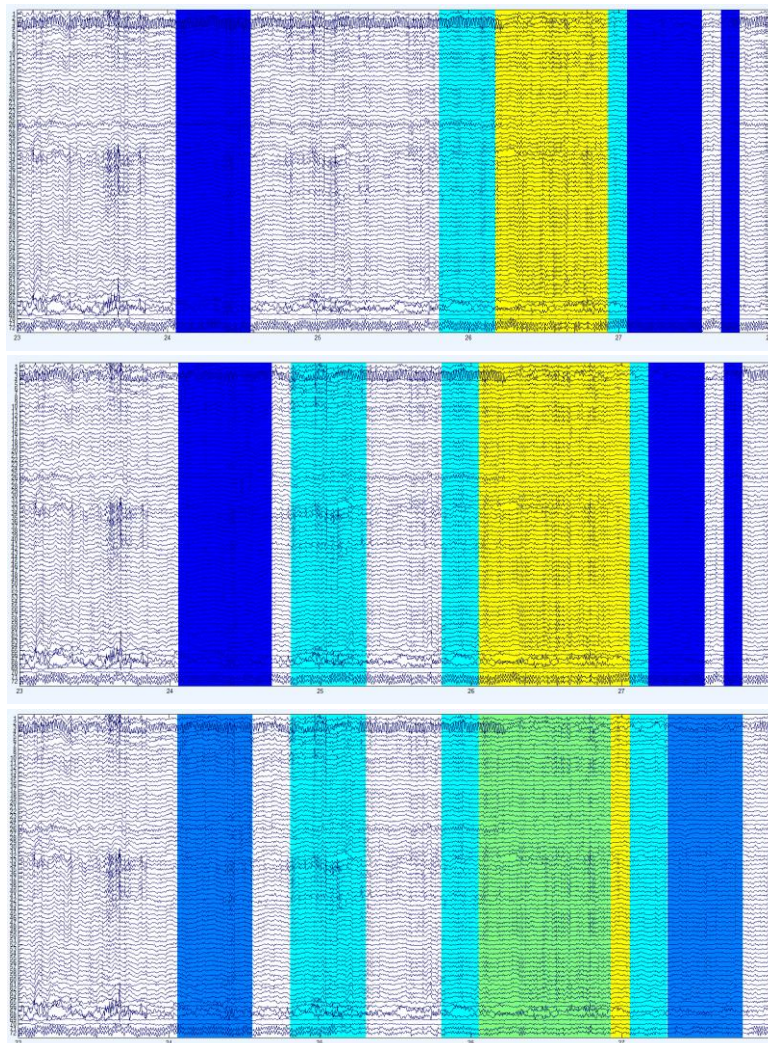


Figura 55 - Señal original y tras las iteraciones de la 1 a la 4



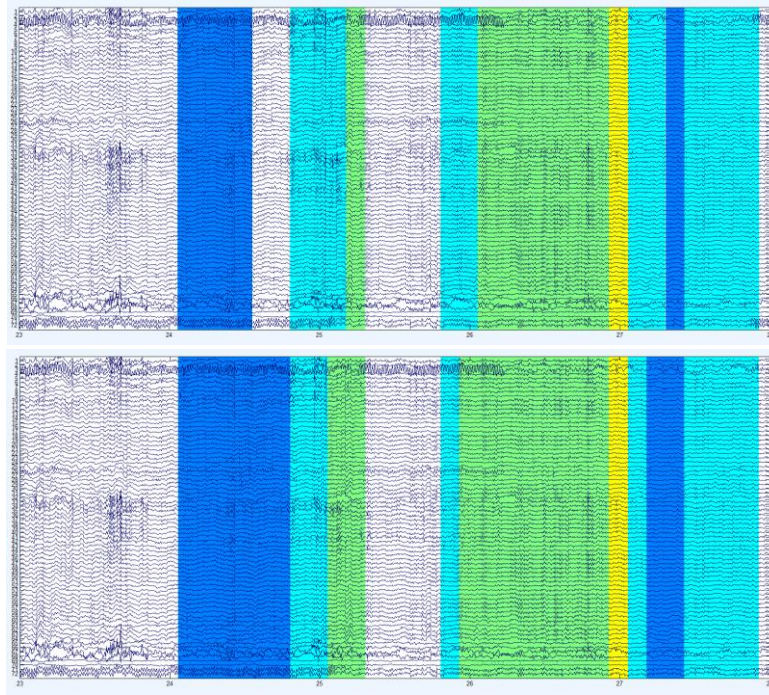


Figura 56 - Señal tras las iteraciones de la 5 a la 9

Repitiendo el proceso para múltiples iteraciones lo que se comprueba es que los artefactos llega un momento que comienzan a extenderse. Esto es debido a que el criterio de filtrado comienza a eliminar componentes de actividad cerebral, afectando al entorno cercano de las zonas tratadas y llevando a DETECT a considerar artefactos en zonas antes despejadas.

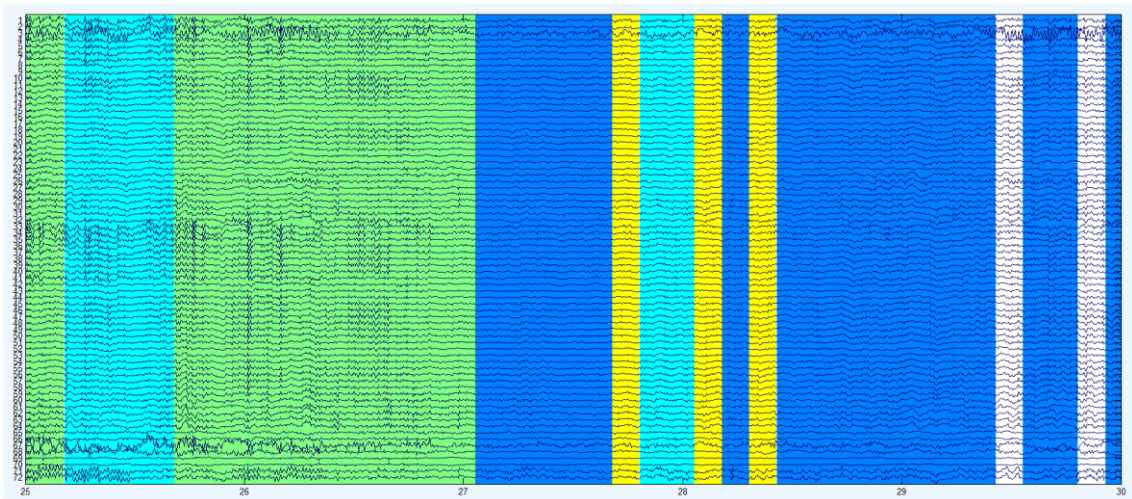


Figura 57 - Resultado tras 69 iteraciones

## 5. CONCLUSIONES Y LÍNEAS FUTURAS

Tras el análisis de los resultados de los experimentos de validación de la solución propuesta, se concluye:

- Que la implementación FastICA es apropiada para el sistema propuesto, ya que además de separar las componentes calcula de forma automática el nº óptima de estas, lo que facilita la aplicación del sistema propuesto a otros campos con señales multicanal más allá del EEG estudiado en esta memoria.
- Que la aplicación ICA muestra mejores resultados al operar sobre ventanas mayores. Dado que el riesgo al ampliar esta ventana es que vincule fuentes de ruido con fuentes de interés, la casuística para elegir la ventana estará muy relacionada con la tipología de señal adquirida, la naturaleza de las fuentes de ruido y el propósito del análisis (ya que las señales y ruidos pueden verse vinculados estadísticamente en determinadas condiciones). Como solución generalista se propone un método por el que se unen ruidos colindantes o muy cercanos para contemplarlos como un solo artefacto y se extiende la ventana del análisis ICA hasta el bloque anterior y siguiente. De este modo se amplía la ventana y se evita artefactos que se produzcan a lo largo de un intervalo relevante y que, por tanto, puedan mostrar cierta dependencia de las fuentes.
- Que DETECT es una herramienta muy útil para identificar y caracterizar los artefactos a filtrar en cualquier tipo de sistema multicanal, pero que muestra una carencia importante en cuanto a la precisión de la ventana señalada para cada artefacto. Se comprueba que, en este caso, una ampliación de la ventana de 0.7 segundos en ambos lados es suficiente para abarcar todo el artefacto.
- Que los criterios de selección y eliminación de componentes independientes son muy robustos para artefactos con gran energía. En los experimentos planteados el sistema es capaz de eliminar el 100% de los artefactos introducidos por un parpadeo (los que mayor energía presentan en los ruidos analizados)
- Que estos criterios han sido capaces de eliminar más del 80% de los tramos de ruido identificados por DETECT, por lo que resultan muy efectivos para la eliminación de ruido de forma automática.
- Que el criterio de selección permite además identificar la fuente de ruido en base al sensor escogido como referencia. Esto permite posicionar físicamente la fuente de ruido, lo que puede ser útil para comprender dichos ruidos o eliminarlos.

Estas conclusiones, así como los resultados anteriores, abren nuevas líneas de trabajo para depurar el planteamiento realizado. Entre estas líneas cabe destacar:

- El ensayo del método propuesto con un sistema de detección de eventos diferente a DETECT. Se plantea el uso de un sistema no supervisado para la detección de eventos (lo que generalizaría aún más la solución) como el propuesto en [12].
- Contrastar con un especialista en EEG el comportamiento del sistema propuesto en cuanto a la conservación de las fuentes cerebrales de interés, para confirmar que solo se eliminan fuentes de ruido.
- Analizar las limitaciones en el filtrado del 20% restante de ruido para constatar si se trata de falsos positivos por parte del sistema de detección de eventos o, de lo

contrario, comprobar características comunes a los ruidos persistentes con los que definir criterios adicionales para su eliminación sin afectar a las señales de interés.

## BIBLIOGRAFÍA

---

1. Delamónica E. "Electroencefalografía". Ed. El Ateneo 1984.
2. Jutten, C. and Héroult, J. (1991). "Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture". *Signal Processing*, 24:1-10.
3. Pierre Comon (1994) "Independent component analysis, A new concept?".
4. A. Hyvärinen. "Survey on Independent Component Analysis". *Neural Computing Surveys* 2:94--128, 1999.
5. Gerard Escudero, Raúl Benítez, David Masip Rodó, Samir Kanaan. "Inteligencia artificial avanzada". 2014
6. S. M. Ross, "Distributions of sampling statistics," in *Probability and statistics for engineers and scientists*, E. A. PRESS, Ed., III ed, 2004, pp. 204-210.
7. A. Hyvärinen, et al., "ICA by Maximization of Nongaussianity," in *Independent Component Analysis*, I. JOHN WILEY & SONS, Ed., ed, 2001, pp. 165-202.
8. Jean-Francois Cardoso, Antoine Souloumiac , Telecom Paris. "BLIND BEAMFORMING FOR NON GAUSSIAN SIGNALS" 1993
9. M. Gaeta and J.-L. Lacoume, "Source separation without a priori knowledge: the maximum likelihood solution," in *Proc. EUSIPCO*, pp. 621-624, 1990.
10. Tzyy-ping Jung, Scott Makeig, Colin Humphries, Te-Won Lee, Martin J. Mckeown, Vicente Iragui, and Terrence J. Sejnowski. "Removing electroencephalographic artifacts by blind source separation". 2000
11. Vernon Lawhern<sup>1</sup> \*, W. David Hairston<sup>2</sup> , Kay Robbins. "DETECT: A MATLAB Toolbox for Event Detection and Identification in Time Series, with Applications to Artifact Detection in EEG Signals". 2013
12. Angel Mur, Raquel Dormido, Jesús Vega, Sebastian Dormido-Canto y Natividad Duro. "Unsupervised event detection and classification of multichannel signals". *Expert Systems with applications*, 2016, vol. 54, pp. 294-303.