



Universidad
Carlos III de Madrid

Departamento de Teoría de la Señal y Comunicaciones

PROYECTO FIN DE CARRERA

**PROCESADO DE SEÑALES
EEG PARA UN INTERFAZ
CEREBRO-MÁQUINA (BCI)**

Autor: Javier de la Torre Abaitua

Tutor: Jose Miguel Leiva Murillo

Leganés, Octubre de 2012



Título: Procesado de señales EEG para un interfaz cerebro-máquina (BCI)

Autor: Javier de la Torre Abaitua

Director: Jose Miguel Leiva Murillo

EL TRIBUNAL

Presidente: Matilde Sánchez Fernández

Vocal: Emilio Parrado Hernández

Secretario: Cristina de Dios Fernández

Realizado el acto de defensa y lectura del Proyecto Fin de Carrera el día 29 de Octubre de 2012 en Leganés, en la Escuela Politécnica Superior de la Universidad Carlos III de Madrid, acuerda otorgarle la CALIFICACIÓN de

VOCAL

SECRETARIO

PRESIDENTE

*Si guardas en tu puesto la cabeza tranquila,
cuando todo a tu lado es cabeza perdida.
Si tienes en ti mismo una fe que te niegan
y no desprecias nunca las dudas que ellos tengan.
Si esperas en tu puesto, sin fatiga en la espera.
Si engañado, no engañas.
Si no buscas más odio, que el odio que te tengan.
Si eres bueno, y no finges ser mejor de lo que eres.
Si al hablar no exageras, lo que sabes y quieres.
Si sueñas y los sueños no te hacen su esclavo.
Si piensas y rechazas lo que piensas en vano.
Si alcanzas el TRIUNFO ó llega tu DERROTA,
y a los dos impostores les tratas de igual forma.
Si logras que se sepa la verdad que has hablado,
a pesar del sofisma del Orbe encanallado.
Si vuelves al comienzo de la obra perdida,
aunque esta obra sea la de toda tu vida.
Si arriesgas de un golpe y lleno de alegría,
tus ganancias de siempre a la suerte de un día,
y pierdes, y te lanzas de nuevo a la pelea,
sin decir nada a nadie lo que eres, ni lo que eras.
Si logras que los nervios y el corazón te asistan,
aún después de su fuga, en tu cuerpo en fatiga,
y se agarren contigo, cuando no quede nada,
porque tú lo deseas, lo quieres y mandas.
Si hablas con el pueblo, y guardas la virtud.
Si marchas junto a Reyes, con tu paso y tu luz.
Si nadie que te hiera, llega a hacerte la herida.
Si todos te reclaman, y ninguno te precisa.
Si llenas el minuto inolvidable y cierto,
de sesenta segundos, que te llevan al cielo.
TODO lo de esta Tierra será de tu dominio,
Y mucho más aún ...
¡ Serás un HOMBRE, hijo mío !*

Si - Rudyard Kipling

Agradecimientos

A mi tutor Jose Miguel, por tu infinita paciencia, por haber estado siempre dispuesto a aclarar mis dudas y ayudarme, y por todo lo que he aprendido de ti.

A mis padres, por haberme dado siempre vuestro apoyo y cariño y haber tenido siempre paciencia y comprensión, por haberme enseñado la virtud del esfuerzo y haber hecho de mi la persona que soy. A mi hermano Gonzalo, por haber estado siempre ahí y haberme ayudado siempre que se lo he pedido.

A Beatriz, por todo, por estar ahí, por convertir los días malos en días buenos, por haberme sabido escuchar y apoyar en todo momento y regalarme siempre una sonrisa. Por todo lo que hemos vivido y sobre todo, por hacerme feliz.

A mis amigos de toda la vida, Fres, Javi, Kike, Dani, Banda, etc. Por poder contar siempre con vosotros y haberme demostrado que por mucho tiempo que pase hay cosas que no cambian nunca.

A mis amigos de la universidad, Álvaro, Jhony, María, Marian, Dani, Estefanía, Javi, Alfonso, Isma, Diana, Marta, etc., con los que tanto tiempo y esfuerzo he compartido, y que habéis hecho que todo haya sido mucho más ameno y divertido. Y en especial a Barco, Jose, Jaime y Rober, por todo lo que hemos vivido durante todos estos años de universidad.

Resumen

En la actualidad, muchos bio-ingenieros están centrados en la comprensión de las propiedades eléctricas del cerebro humano. De hecho, en los últimos años se han sucedido las noticias acerca de interfaces cerebro máquina que permitan manejar ordenadores con la mente, tanto en el campo biomédico como en el militar. Dentro de estas propiedades eléctricas podemos encontrar los potenciales evocados P300, señales cerebrales que se presentan ante estímulos visuales inesperados. Estas señales serán el punto de partida de este proyecto fin de carrera, con el que se busca diseñar un decodificador de caracteres. El usuario final al que va dirigido serán aquellas personas que presenten una incapacidad de comunicación con el exterior y que no puedan usar otro tipo de interfaces que requieran voz o capacidades motoras. El sistema empleado es el denominado *Visual P300 speller*, que permite evocar en el sujeto los potenciales P300 mediante una interfaz gráfica y grabar su electroencefalograma. Dicha interfaz consistirá en una matriz 6x6 de caracteres, que se irá iluminando aleatoriamente por filas o columnas, aunque existen otras alternativas como la iluminación por patrones establecidos, durante un determinado periodo de tiempo, mientras que el sujeto focaliza su atención en una letra. Sin más que ir grabando la actividad cerebral del paciente (EEG) y manteniendo un registro de los estímulos que se le han presentado, seremos capaces de predecir la letra que el usuario quiere transmitir. Para ello se dividirá el registro EEG, en segmentos de tiempo, llamados épocas, con los que, tras aplicar el algoritmo generativo que se va a desarrollar, obtendremos unas plantillas, de eventos objetivos y no objetivos que nos permitirán generar la respuesta cerebral que debería tener el sujeto ante una colección de estímulos, para finalmente usar un decodificador MAP y obtener la letra transmitida. Para ello, se ha empleado una base de datos con 6 sujetos donde cada uno fue sometido a 64 experimentos formados por 72 eventos cada uno. Así, se ha podido demostrar que la tasa de acierto se encuentra cercana al 100%.

Palabras clave: Interfaz BCI, *Visual P300 speller*, algoritmo generativo, decodificador MAP, electroencefalograma.

Abstract

Currently, bio-engineers are trying to understand the electrical properties of the human brain. News about BCI control of computers in both the biomedical and the military industries has appeared in the media. Within those electrical properties we find the P300 evoked potentials, which are the response to unexpected visual stimuli. These signals will be the starting point of this final project, which seeks to design a character decoder. The final users will be people who have some kind of communication difficulty and who can't use interfaces based on voice or movement. The employed system is the Visual P300 speller, which stimulates P300 potential by using a graphical interface and records the subject's electroencephalogram. This interface will have a 6x6 letter grid, which will be lit randomly by rows or columns during a fixed time, while the subject is focusing his attention on a letter. By recording his EEG and keeping a record of the presented stimulus, we will be able to predict the desired letter. In order to do that, the EEG will be divided into time segments, or epochs, which will be used, once they are processed by our generative model, to obtain two different templates, one for target events and another one for non target events. These templates will allow us to generate the cerebral response to a collection of input stimuli in order to use a MAP decoder to determine the transmitted letter. We have used a database with the records of 6 persons each of whom was subjected to 64 experiments of 72 events each. We were able to demonstrate that it is possible to achieve almost 100% success.

Keywords: BCI Interface, Visual P300 speller, generative algorithm, MAP decoder, electroencephalogram

Contenido

Capítulo 1.....	1
Introducción y objetivos	1
1.1 Introducción.....	1
1.2 Objetivos.....	3
1.3 Fases del desarrollo	3
1.4 Medios empleados	4
1.5 Estructura de la memoria	4
Capítulo 2.....	5
Conceptos básicos	5
2.1 Introducción.....	5
2.2 Teorema de Bayes	5
2.3 Decisor MAP	7
2.4 Validación Cruzada	8
2.5 Sistema lineal.....	8
2.6 Sistema con y sin memoria	9
2.7 Matriz Toeplitz.....	9
2.8 Teorema central del límite	10
2.9 Correlación.....	10
2.10 Covarianza	11
2.11 Teorema de muestreo de Nyquist.....	11
2.12 Distancia de Hamming.....	12
Capítulo 3.....	13
Procesado de señales EEG	13
3.1 Interfaz cerebro-máquina (BCI)	13
3.2 Electroencefalograma (EEG)	15
3.2.1 Estado mental del sujeto	18
3.2.2 Electrodo para la medición del EEG	19
3.3 La señal P300 y el paradigma <i>Oddball</i>	20
3.3.1 Señal P300	20
3.3.2 Paradigma <i>Oddball</i>	24
3.4 El sistema <i>visual speller</i>	26

3.4.1	Introducción al sistema <i>visual speller</i>	26
3.4.2	El sistema <i>visual speller</i> empleado	32
3.4.3	Solape y efectos refractarios en el <i>speller</i>	37
3.4.4	Distancia de Hamming y el <i>Visual Speller</i>	41
3.5	Pre-procesado de las señales	42
3.5.1	Filtrado de las señales.....	42
3.5.2	Análisis de componentes principales	44
3.5.3	Introducción.....	44
3.5.4	Descripción del método.....	46
Capítulo 4.....		49
Modelado probabilístico de las señales		49
4.1	Modelos discriminativos vs generativos.....	49
4.2	Modelo generativo de señales cerebrales.....	53
4.3	Estimación MAP	57
4.4	Caracterización del ruido	59
4.5	Estimación MAP de señales decorreladas	60
4.6	Resumen del modelo	61
Capítulo 5.....		63
Experimentos y resultados		63
5.1	Introducción.....	63
5.2	Datos y medios empleados.....	63
5.3	Obtención del conjunto de datos	66
5.4	Indicadores de prestaciones	67
5.5	Resultados parciales	68
5.6	Resultados	76
5.6.1	Resultados de los experimentos Flash.....	76
5.6.2	Resultados experimentos Flip.....	79
5.6.3	Comparativa entre experimentos Flash y Flip	82
5.6.4	Comparativa entre modelo generativo y discriminativos	83
5.7	Mejoras y nuevas aportaciones.....	85
5.7.1	Introducción.....	85
5.7.2	Reducción de la dimensión del problema	85
5.7.3	Adición de canales retardados	89
Capítulo 6.....		97

Conclusiones y futuras líneas de trabajo	97
6.1 Introducción.....	97
6.2 Conclusiones	97
6.3 Futuras líneas de trabajo	99
Capítulo 7.....	101
Presupuesto.....	101
7.1 Presupuesto del proyecto.....	101
Capítulo 8.....	105
Glosario.....	105
Capítulo 9.....	107
Referencias	107
9.1 Introducción.....	107
Capítulo 10.....	111
Anexo	111
10.1 Sistema internacional 10-20: partes del Cerebro	111
10.2 Minimizar el error mediante LS	113

Índice de figuras

Figura 1: Representación sistema lineal.....	9
Figura 2: Clasificación técnicas captura actividad cerebral	15
Figura 3: Ejemplo electroencefalograma (EEG).....	16
Figura 4: Representación neurona.....	17
Figura 5: Representación de los potenciales evocados	24
Figura 6: Esquema del sistema Visual P300 speller	26
Figura 7: Esquema tx - canal ruidoso - rx	28
Figura 8: Matriz de caracteres.....	32
Figura 9: Matriz de caracteres y <i>Codebook</i>	33
Figura 10: Unidades del sistema speller.....	34
Figura 11: Representación TTI y TNI.....	39
Figura 12: Ejemplo estímulo Flash y Flip.....	40
Figura 13: Filtro paso Banda.....	42
Figura 14: Reducción redundancia temporal. a) Espectro de la señal y del filtro. b) Señal muestreada a frecuencia original. c) Señal muestreada a la nueva frecuencia.....	43
Figura 15: Distribución probabilidad conjunta de x_1 y x_2	45
Figura 16: Distribución probabilidad conjunta de x_1 y x_2 decorreladas	46
Figura 17: Ejemplo modelo gráfico	51
Figura 18: Modelo determinista de la respuesta cerebral ante un estímulo	53
Figura 19: Generación de la señal cerebral observada con ruido.....	54
Figura 20: Generación de las matrices S_t y S_n	55
Figura 21: Esquema general para la obtención de la letra estimada	65
Figura 22: Evolución de los datos	68
Figura 23: Señal cerebral antes y después del filtro paso banda.....	69
Figura 24: Plantillas de evento no objetivo.....	70
Figura 25: Plantillas para eventos objetivo	71
Figura 26: Blanqueamiento de la señal (entrenamiento).....	72

Figura 27: Matriz de covarianza de la señal antes y después de ser blanqueada (entrenamiento)	72
Figura 28: Blanqueamiento de la señal (test)	73
Figura 29: Matriz de covarianzas de la señal antes y después de ser blanqueada (test) ...	74
Figura 30: Señal cerebral generada	75
Figura 31: Probabilidad acierto 1 (Flash) (Flash)	76
Figura 32: Probabilidad acierto 2 (Flash)	76
Figura 33: Probabilidad acierto 3 (Flash)	77
Figura 34: Probabilidad acierto 4 (Flash)	77
Figura 35: Probabilidad acierto 5 (Flash)	77
Figura 36: Probabilidad acierto 5 (Flash) ...	77
Figura 37: Probabilidad acierto promedio (Flash)	78
Figura 38: Probabilidad acierto 1 (Flip)	79
Figura 39: Probabilidad acierto 2 (Flip) ...	79
Figura 40: Probabilidad acierto 3 (Flip)	80
Figura 41: Probabilidad acierto 4 (Flip)	80
Figura 42: Probabilidad acierto 5 (Flip)	80
Figura 43: Probabilidad acierto 6 (Flip)	80
Figura 44: Probabilidad acierto promedio (Flip)	81
Figura 45: Probabilidad acierto promedio (Flash vs Flip)	82
Figura 46: Probabilidad acierto promedio (Modelos discriminativos)	83
Figura 47: Probabilidad acierto promedio (Generativo vs Discriminativos)	84
Figura 48: Evolución de la probabilidad de acierto en función del número de canales usados	86
Figura 49: Evolución de la probabilidad de acierto	88
Figura 50: Probabilidad acierto 1 retardo retardos	89
Figura 51: Probabilidad acierto 2 retardos	89
Figura 52: Probabilidad acierto 3 retardos retardos	90
Figura 53: Probabilidad acierto 4 retardos	90
Figura 54: Evolución de la probabilidad acierto (1 retardo)	91
Figura 55: Evolución de la probabilidad acierto (2 retardos)	92
Figura 56: Evolución de la probabilidad acierto (3 retardos)	93
Figura 57: Evolución de la probabilidad acierto (4 retardos)	94
Figura 58: Probabilidad de acierto vs número de retardos (72 bits)	95
Figura 59: Presupuesto del proyecto	102
Figura 60: Diagrama de Gantt del proyecto	103
Figura 61: Posiciones de referencia en el sistema internacional 10-20	112
Figura 62: Partes principales del cerebro	112
Figura 63: Posiciones de referencia del casco de malla utilizado para las capturas muestrales durante los experimentos	113

Índice de tablas

Tabla 1: Ritmos cerebrales	18
----------------------------------	----

Capítulo 1

Introducción y objetivos

1.1 Introducción

En la actualidad los gobiernos de todo el mundo están creando normativas y leyes que permitan a personas discapacitadas tener mayores facilidades de integración en la vida cotidiana. Un ejemplo de esto lo encontramos en la Accesibilidad Universal, definiéndose esta como la condición que deben cumplir los entornos, procesos, bienes, productos y servicios, así como objetos o instrumentos, herramientas y dispositivos, para ser comprensibles, utilizables y practicables por todas las personas en condiciones de seguridad y comodidad y de la forma más autónoma y natural posible (Artículo 2, Ley 51/2003, 2 de diciembre, de Igualdad de oportunidades, No discriminación y Accesibilidad Universal de las Personas con Discapacidad).

El hecho de poder trabajar en un proyecto por el cual, por medio de la ingeniería se puede mejorar la calidad de vida, ha sido el elemento clave a la hora de querer realizar este proyecto. Aunque todavía quede mucho por avanzar en este campo, toda contribución que pueda hacerse, por pequeña que sea, supone un paso en la esperanza de muchas personas de poder vivir más integrados en la sociedad.

Para ello, partiendo de los estudios de Farwell-Donchin [1] y muy especialmente en el trabajo de S. Martens [2] se buscará desarrollar un interfaz BCI que consiga establecer una comunicación entre el sujeto y un ordenador, donde la respuesta cerebral modulada se verá alterada por el ruido presente en los electrodos y el cableado para finalmente,

CAPITULO 1: INTRODUCCIÓN Y OBJETIVOS

aplicar una decodificación MAP. De esta forma se consigue establecer un paralelismo entre el problema que se presenta y un canal ruidoso como los estudiados en la carrera de Ingeniería de Telecomunicaciones.

Así, este proyecto trata de facilitar a personas que no tienen medio de comunicación con el exterior, una vía con la que ser capaces de transmitir una secuencia de letras o caracteres y poder, por lo tanto, transmitir cualquier tipo de mensaje. Para ello se pretende emplear las propiedades eléctricas del cerebro ante estímulos visuales poco frecuentes. La idea principal consiste en codificar las diferentes letras que se muestran al sujeto mediante 1s y 0s, según si han formado parte o no de la columna o fila que ha sido iluminada en un determinado instante de tiempo. Gracias a esta codificación podemos definir cada una de las letras unívocamente mediante una palabra código y plantear el sistema como un decodificador para un canal con ruido. Llegados a este punto se abren dos vías con las que resolver el problema. Modelos discriminativos o modelos generativos. En este proyecto se pretende abordar el problema empleando para ello un modelo generativo junto con un decisor MAP que nos permita predecir la letra transmitida. La elección del modelo generativo en vez de uno discriminativo viene motivada por tres puntos principales:

- Simplicidad conceptual debido a que el sistema se puede reducir a elegir la señal, de entre un conjunto de señales, más parecida a una señal de referencia.
- Mayor interpretabilidad física, ya que lo que se busca es representar la señal cerebral lo más parecida a la respuesta que presenta el cerebro.
- Menor complejidad computacional respecto a los modelos discriminativos, en especial cuando se emplean máquinas de vector soporte ya que estas requieren de un mayor número de datos para entrenamiento.

Por último, se pretende demostrar que los modelos generativos pueden ofrecer los mismos resultados que los mejores métodos discriminativos. Por ello, se desarrollará el modelo generativo para posteriormente analizar qué posibilidades de mejora existen e implementarlas.

1.2 Objetivos

El objetivo fundamental de este proyecto es conseguir implementar un modelo generativo que permita la estimación de una letra, a partir de la lectura y procesado de señales EEG utilizando para ello el sistema *Visual P300 speller*. En base a ese objetivo principal, se proponen los siguientes objetivos parciales:

- Estudio e implementación del modelo generativo a desarrollar.
- Maximizar la tasa de acierto del sistema.
- Estudio del efecto del tipo de estímulo en la tasa de acierto.
- Estudio de alternativas que reduzcan el elevado coste computacional.
- Estudio del efecto, en la tasa de acierto, de la inclusión de canales extra con versiones retardadas de las señales.

1.3 Fases del desarrollo

Las fases del desarrollo del proyecto han sido las siguientes:

- Estudio y análisis de las dos vías de desarrollo del sistema, modelos generativos y modelos discriminativos.
- Estudio de las interfaces BCI, sistema *Visual P300 speller*, y paradigma *Oddball*.
- Desarrollo e implementación del modelo generativo.
- Simulación y pruebas del modelo generativo.
- Desarrollo e implementación de dos nuevas aportaciones.
 - Minimizar la dimensión del problema
 - Incrementar el número de canales

1.4 Medios empleados

Para el desarrollo de este proyecto se empleó exclusivamente el siguiente programa:

- Matlab 7.4.0(R2007a)

Para llevar a cabo ciertas simulaciones se emplearon los ordenadores del departamento Teoría de la Señal y Comunicaciones, debido a las altas exigencias de memoria requeridas. Para ello se estableció una VPN con la que trabajar a distancia.

1.5 Estructura de la memoria

Para facilitar la lectura de la memoria, se incluye a continuación un breve resumen de cada capítulo.

En el capítulo 2 se pretende dar una explicación de ciertos elementos básicos que se emplearán a lo largo de la memoria, para intentar que un lector no experto en la materia pueda seguir su lectura con la mayor comprensión posible.

En el capítulo 3 se explicará en qué consiste el sistema *Visual P300 speller* y el pre-procesado a realizar al conjunto de datos disponibles. Se realizará un estudio del estado del arte y se presentarán sus principales propiedades. Para ello, se entrará en detalle en las propiedades y los tipos de ondas cerebrales y cómo estas pueden ser usadas para llevar a cabo la tarea de transmisión deseada. Finalmente se entrará en detalle del pre-procesado, que realizará un filtrado de las señales y su posterior blanqueamiento.

En el capítulo 4 se explicarán los dos modelos que permiten resolver el problema, explicando brevemente el modelo discriminativo. Se desarrollará el modelo generativo que se ha implementado, explicando las características de las señales cerebrales asumidas para enfrentar el problema. Será en este capítulo donde se explique cómo se obtienen las plantillas cerebrales y cómo se decodifica la información para obtener una estimación de la letra transmitida.

Finalmente, se presentarán los resultados, las conclusiones extraídas del proyecto, las aportaciones y las líneas futuras de investigación, así como el presupuesto que ha sido necesario.

Capítulo 2

Conceptos básicos

2.1 Introducción

En este capítulo se pretende presentar al lector todos los conceptos básicos necesarios para poder seguir con mayor facilidad la descripción del método desarrollado, con el fin de que el texto resulte ameno para todo tipo de lectores, tanto expertos como aquellos que no tengan conocimiento alguno de conceptos de tratamiento de señales. Como referencia para explicar estos conceptos, se han empleado tanto apuntes de las diferentes asignaturas estudiadas en la carrera como los libros [23] y [24].

2.2 Teorema de Bayes

El teorema de Bayes permite calcular las probabilidades de un suceso A condicionadas a otro B, a partir de las probabilidades de B condicionadas a A.

Así, el teorema afirma que, dados dos sucesos $A \in \Omega$ y $B \in \Omega$, donde Ω se define como un espacio muestral, es decir, el conjunto de todos los resultados posibles del experimento aleatorio,

CAPITULO 2: CONCEPTOS BÁSICOS

$$\Pr\{A|B\} = \frac{\Pr\{B|A\}P\{A\}}{\Pr\{B\}}$$

Algunas extensiones del teorema de Bayes resultan inmediatas, siendo p una distribución de probabilidad y f una función de distribución de probabilidad (fdp):

- Dadas dos variables discretas X e Y

$$p\{x|y\} = \frac{p\{y|x\}p\{x\}}{p\{y\}}$$

- Si X es continua e Y discreta

$$f\{x|y\} = \frac{p\{y|x\}f\{x\}}{p\{y\}}$$

- Si X e Y son continuas

$$f\{x|y\} = \frac{f\{y|x\}f\{x\}}{f\{y\}}$$

El teorema de Bayes es consecuencia directa de la definición de probabilidad condicionada, y su demostración resulta sencilla. Pese a su simplicidad, constituye una de las piedras angulares de la teoría de la probabilidad y tiene una enorme utilidad práctica. La razón es que con frecuencia Y es una variable observable, la cual representa el efecto de una causa asociada a la magnitud X que no es directamente observable, y se dispone de un modelo probabilístico de la forma en la que X causa Y : en definitiva, $p_{y|x}$ es conocida. Dado que Y es observable y X no, en la práctica se plantea la necesidad de calcular $p_{x|y}$, problema resuelto por el teorema de Bayes.

2.3 Decisor MAP

El decisor MAP, máximo a posteriori, debe su nombre a la denominación de $p(b_j|q_0)$ como probabilidades a posteriori por representar la probabilidad de los símbolos una vez realizada la transmisión, y en contraposición a las probabilidades $p(b_j)$ denominadas a priori por representar la probabilidad de los símbolos antes de realizar la transmisión.

Definiendo p como una distribución de probabilidad y f una función de distribución de probabilidad (fdp):

La determinación de las probabilidades a posteriori puede realizarse mediante la regla de Bayes,

$$p\{b_j|q_0\} = \frac{f\{q_0|b_j\}p\{b_j\}}{f\{q_0\}}$$

donde, teniendo en cuenta que $B = b_j$ implica que $A = a_j$ y viceversa, establecemos la siguiente igualdad:

$$f\{q_0|b_j\} = f\{q_0|a_j\}$$

Considerando que la salida del decisor será el símbolo que cumpla

$$p(b_i|q_0) > p(b_j|q_0)$$

el criterio MAP se reduce a encontrar el símbolo b_i que cumple

$$\frac{f\{q_0|a_i\}p\{b_i\}}{f\{q_0\}} > \frac{f\{q_0|a_j\}p\{b_j\}}{f\{q_0\}}$$

Que, como $f(q_0)$ es una cantidad no negativa e independiente del símbolo transmitido, es equivalente

$$f\{q_0|a_i\}p\{b_i\} > f\{q_0|a_j\}p\{b_j\}$$

Quedando así definido el decisor MAP. Por último indicar que el decisor ML sería igual salvo que considera los símbolos equiprobables, es decir, $p(b_j) = p(b_i)$.

2.4 Validación Cruzada

La validación cruzada (*cross-validation*) consiste en dividir una muestra de datos en subconjuntos de tal modo que el análisis se realiza inicialmente en uno de ellos, mientras los otros subconjuntos son usados en la confirmación y validación del análisis inicial.

La validación simple consiste en dividir en dos conjuntos complementarios los datos de la muestra, usar uno de ellos para construir el modelo (conjunto de entrenamiento) y usar el otro para medir el ratio de error del modelo construido (conjunto de test).

La validación cruzada aplica k-veces la validación simple, dividiendo en k conjuntos la muestra (*k-fold cross validation*). Así, en cada iteración se construirá y evaluará un modelo, usando uno de los conjuntos como conjunto de entrenamiento y el resto como conjunto de test. Posteriormente, obteniendo la media aritmética de los ratios de error obtenidos, conseguiremos el ratio de error para la muestra final. En cuanto a la elección del valor k dependerá del tamaño y características de la muestra.

2.5 Sistema lineal

Un sistema lineal, en tiempo continuo o en tiempo discreto, es aquel que posee la propiedad de superposición.

“Si una entrada consiste en la suma ponderada de varias señales, entonces la salida será simplemente la superposición (o suma ponderada) de las respuestas del sistema a cada una de estas señales” Señales y Sistemas Alan V. Oppenheim pág. 53

Por lo tanto, matemáticamente, si consideramos $y_1(t)$ la respuesta del sistema continua a una entrada $x_1(t)$, el sistema se considera lineal si:

- La respuesta ante una entrada, $x_1(t) + x_2(t)$, será $y_1(t) + y_2(t)$
- La respuesta ante una entrada, $ax_1(t)$, será $ay_1(t)$, siendo a una constante compleja cualquiera.

La primera propiedad se denomina propiedad de aditividad mientras que la segunda como propiedad de escalamiento y homogeneidad donde ambas propiedades pueden emplearse de igual manera para señales continuas.

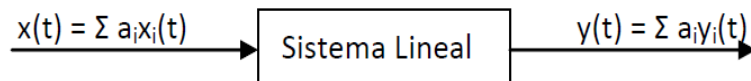


Figura 1: Representación sistema lineal

2.6 Sistema con y sin memoria

Se dice que un canal es sin memoria si para cada valor de la variable independiente en un tiempo dado depende solamente de la entrada en ese mismo tiempo. Por otro lado, si hay una dependencia entre valores de tiempo anteriores, se dice que el canal es con memoria.

De forma general, el concepto de memoria en un sistema corresponde a la presencia de un mecanismo en el sistema que mantiene o almacena información sobre los valores de entrada en instantes diferentes del tiempo actual.

Atendiendo a la matemática del problema podemos describir el canal con memoria y sin memoria con las siguientes expresiones:

- Con memoria: $y(t) = f(x(t), x(t-1), x(t-2), \dots, x(t-i))$
- Sin memoria: $y(t) = f(x(t))$

2.7 Matriz Toeplitz

Se denomina matriz Toeplitz a aquella matriz cuadrada con todas sus diagonales de izquierda a derecha paralelas numéricamente.

$$T = \begin{pmatrix} a & b & c & d & k \\ f & a & b & c & d \\ g & f & a & b & c \\ h & g & f & a & b \\ j & h & g & f & a \end{pmatrix} \quad \forall a_{i,j} \in T \rightarrow a_{i,j} = a_{i+1,j+1}$$

2.8 Teorema central del límite

Dado un conjunto de variable aleatorias, X_1, X_2, \dots, X_N , independientes e idénticamente distribuidas (es decir, con la misma fdp), de media μ y varianza σ^2 , su suma $S_N = \sum_{i=1}^N X_i$ es una variable aleatoria de media $N\mu$ y varianza $N\sigma^2$. El teorema central del límite demuestra que, si las variables X_i , son continuas, la fdp de su suma tiende a una gaussiana, de modo que, para valores grandes de N , es válida la aproximación

$$f_{S_N}(s) \approx \frac{1}{(N2\pi\sigma^2)^{1/2}} \exp\left(-\frac{(s - N\mu)^2}{2N\sigma^2}\right)$$

Si las variables X_i tienen diferentes distribuciones, de medias μ_i y varianzas σ_i^2 , se obtendría S_N como una variable de media $\mu = \sum_{i=1}^N \mu_i$ y varianza $\sigma^2 = \sum_{i=1}^N \sigma_i^2$. En esta situación, el teorema también se cumple siempre que las varianzas σ_i^2 sean comparables (σ_i^2 / σ^2 tiende a 1 cuando N crece).

2.9 Correlación

La correlación estadística es una medida de la fuerza y la dirección de una relación lineal entre dos variables aleatorias. Se considera que dichas variables están correlacionadas cuando los valores de una de ellas varían sistemáticamente con respecto a los valores homónimos de la otra. La expresión de la correlación es la siguiente

$$R_{XY}(t_1, t_2) = E\{X(t_1)Y^*(t_2)\}$$

La relación entre dos variables queda representada mediante la línea de mejor ajuste, trazada a partir de la nube de puntos. Los componentes principales de una línea de ajuste y, por lo tanto, de una correlación son:

- Fuerza: mide el grado en que la línea representa a la nube de puntos. Si la nube es estrecha y alargada, se representa por una línea recta, lo que indica una relación fuerte; si la nube tiene tendencia elíptica o circular, la relación es débil.
- Sentido: mide las variaciones de Y con respecto a X será positiva cuando al crecer los valores de X también lo hacen los de Y y negativa cuando al crecer los valores de X decrecen los de Y .

2.10 Covarianza

La covarianza estadística es una medida de dispersión conjunta de dos variables estadísticas. Es decir, la covarianza indica mediante un valor numérico la relación existente entre las dos variables estadísticas y cuya expresión es la siguiente

$$S_{XY} = Cov(X, Y) = \sum_{i=1}^N \frac{(x_i - \bar{X})(y_i - \bar{Y})}{N} = E((X - E(X))(Y - E(Y))^T)$$

La interpretación de la covarianza está en función del valor obtenido

- $S_{XY} > 0 \rightarrow$ Relación lineal directa: a valores elevados de X corresponden valores elevados de Y.
- $S_{XY} = 0 \rightarrow$ Variables incorreladas: no existe relación lineal entre X e Y.
- $S_{XY} < 0 \rightarrow$ Relación lineal inversa: a valores elevados de X corresponden valores bajos de Y.

Las propiedades más destacadas de la covarianza son las siguientes

- $S_{XY} = S_{YX}$
- $S_{XX} = S_X^2$
- $Cov(c + aX, d + bY) = ab * Cov(X, Y)$

2.11 Teorema de muestreo de Nyquist

Sea $x(t)$ una señal de banda limitada con $X(j\omega) = 0$ para $|\omega| > \omega_M$ y ω_s la frecuencia de muestreo deseada. Entonces $x(t)$ se determina unívocamente mediante sus muestras $x(nT)$, $n = 0, \pm 1, \pm 2, \dots$, si

$$\omega_s > 2\omega_M$$

donde

$$\omega_s = \frac{2\pi}{T}$$

Dadas estas muestras, podemos construir $x(t)$ generando un tren de impulsos periódicos en el cual los impulsos sucesivos tengan amplitudes que correspondan a valores de muestras sucesivas. Este tren de impulsos es entonces procesado a través de un filtro

paso bajo ideal con ganancia T cuya frecuencia de corte sea mayor que w_M y menos que $w_s - w_M$. La señal de salida resultante será exactamente igual a $x(t)$.

2.12 Distancia de Hamming

Se define la distancia de Hamming (d_H) como el número de dígitos en que difieren dos vectores o palabras código. Así, podemos definir la mínima distancia de Hamming como la menor distancia entre dos palabras código de entre todas las palabras código pertenecientes a un *codebook*:

$$d_{\min} = \min_{i,j/i \neq j} d_H(c_i, c_j)$$

Gracias a esta definición se pueden emplear códigos Hamming, que permitan detectar y corregir errores, añadiendo redundancia, en función del valor de la distancia mínima de Hamming.

Capítulo 3

Procesado de señales EEG

3.1 Interfaz cerebro-máquina (BCI)

En este capítulo se pretende dar una visión global sobre todos los elementos que conforman el sistema *visual P300 speller*, así como la explicación de todos aquellos procesos que se dan tanto a nivel biológico (señal P300, proceso neuronal, etc.) como a nivel computacional (procesado de señal). Lo que se pretende explicar en este capítulo es la procedencia de los datos y el análisis previo de los mismos antes de aplicarles el modelo generativo que se pretende estudiar en este Proyecto Fin de Carrera.

Una interfaz cerebro-máquina (BCI) consiste en un canal de control y comunicación electrónico, el cual no emplea los nervios periféricos ni los músculos, que conforman las salidas naturales del cerebro, sino que transforma directamente los procesos cerebrales en acciones.

Así, las BCI analizan la actividad cerebral del sujeto paciente, permitiendo detectar los diferentes estados mentales del mismo en tiempo real. A partir de estos estados se puede, posteriormente, transformar en órdenes operativas. En este caso el estudio que se está realizando permitirá determinar qué letra, de una matriz de caracteres, elige el sujeto sin necesidad de que este deba realizar ningún movimiento. No debemos olvidar

CAPITULO 3: PROCESADO DE SEÑALES EEG

que el concepto de BCI es mucho más amplio y que en este estudio realmente se está abordando una parte del mismo. En concreto, un sistema *visual-speller*.

Esta posibilidad de enviar órdenes mentales directas a una máquina se sustenta en una peculiaridad del cerebro: codifica de manera distribuida cada una de nuestras capacidades, lo mismo motoras (agitar una mano) como cognitivas (el lenguaje); es decir, cada región cerebral participa en más de una tarea mental (ver Anexo 10), al tiempo que cada tarea activa diversas regiones con intensidad y ritmos sutilmente diferentes.

Existen dos modalidades en que una interfaz registra la actividad cerebral atendiendo a la técnica empleada para recoger los datos: invasivas y no invasivas. Las interfaces invasivas utilizan micro electrodos implantados en el cerebro que miden la actividad de las neuronas, individualmente consideradas. Por su parte, la mayoría de las interfaces no invasivas se basan en el electroencefalograma (EEG), que indica la actividad eléctrica del cerebro medida con electrodos superficiales colocados sobre el cuero cabelludo. Refleja la actividad síncrona de miles de neuronas corticales, por lo que nos proporciona información detallada sobre la actividad local de pequeñas regiones cerebrales. Se caracteriza, además, por señales de pequeñísima amplitud y de gran variabilidad en el tiempo así como por tener una SNR pobre.

Por otro lado, si tenemos en cuenta la naturaleza del patrón que genera la actividad cerebral, encontramos otros dos tipos, endógenas y exógenas.

- Endógenas: asociadas a algún tipo de proceso cognitivo, como cálculo mental o la memoria, la actividad cerebral se lleva a cabo sin ningún tipo de estimulación sensorial, es decir, no se requiere ningún tipo de estimulación externa. La limitación de este tipo de BCIs viene en el número de acciones que puede transcribir ya que, en principio, no es posible modular toda la información que el cerebro produce. Un ejemplo de este tipo de BCIs es el descrito por Birbaumer [6], el TTD (*thought translation device*). Este tipo de sistema utiliza los potenciales corticales lentos (PCL) que se encuentran situados en la banda de más baja de frecuencia (0.1-2 HZ) y que son generados por la corteza cerebral. Este estudio ha demostrado que las personas pueden aprender a modular sus PCL para mover objetos dentro de una pantalla. Este tipo de BCIs resulta ideal para personas sin mucha movilidad como las afectadas por la esclerosis múltiple, aunque por el momento requiere de largos procesos de entrenamiento (meses) con tasa de error bastante altas.

- Exógenas: en este caso será completamente necesario la existencia de una estimulación externa que produzca respuestas nerviosas, que serán moduladas, en el sujeto. Como ejemplo de este tipo de estimulación está el BCI basado en el P300 y el uso de un paradigma de tipo *Oddball* [1], que será el sistema ha estudiar en este proyecto fin de carrera. Este paradigma se basa en el reconocimiento, de entre una serie de estímulos idénticos o no relevantes, entre aquél que es distinto o de interés, mediante impulsos luminosos en una pantalla. Los BCIs que actualmente obtienen mejores resultados son aquellos que utilizan las respuestas exógenas.

3.2 Electroencefalograma (EEG)

En primer lugar, resulta relevante explicar el por qué de la elección del electroencefalograma como método de medida de la actividad cerebral. Atendiendo a la Figura 2 [25], que muestra una clasificación de las diferentes técnicas de captura de la actividad cerebral en función de la resolución espacio-temporal que requieren, se puede observar cómo el EEG es de los menos invasivos. Además, ofrece las mejores prestaciones en relación calidad-precio y su fabricación es sencilla.

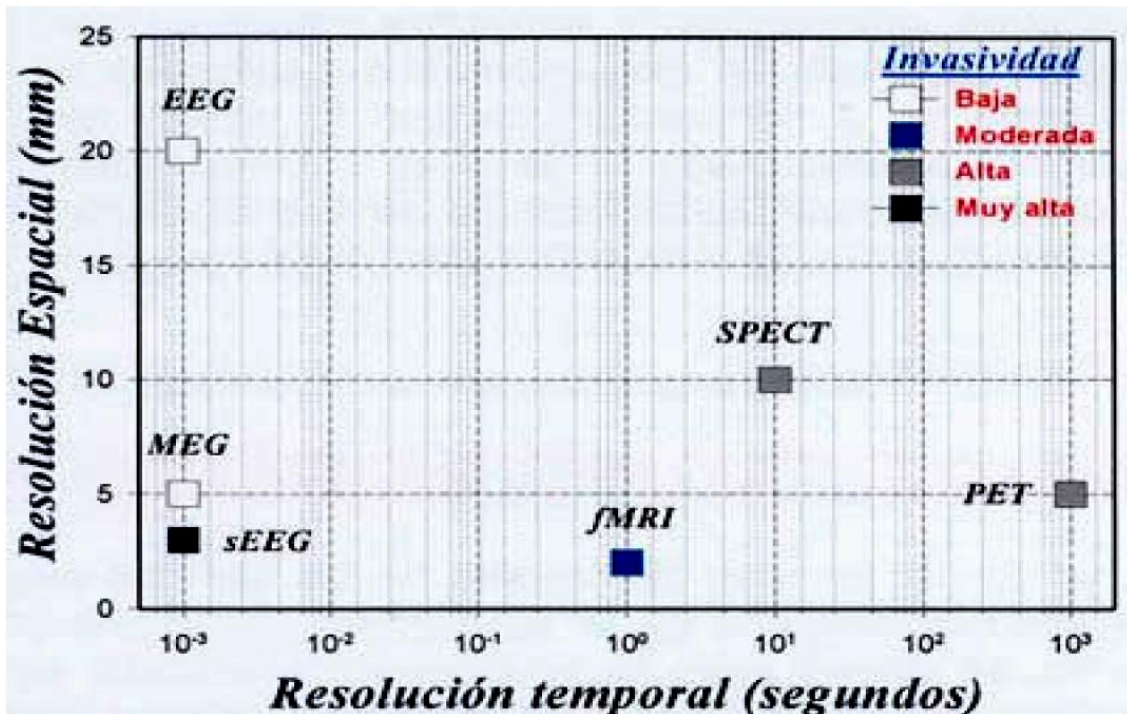


Figura 2: Clasificación técnicas captura actividad cerebral

CAPITULO 3: PROCESADO DE SEÑALES EEG

La electroencefalografía es uno de los métodos electro-fisiológicos más antiguos de la ciencia moderna, datando de 1924 de la mano de Hans Berger. En 1929, él mismo fue capaz de constatar la existencia de los ritmos alfa y beta en sus análisis de los EEGs. El EEG se basa en las corrientes de naturaleza iónica presentes en la corteza cerebral, que son el producto de la actividad cerebral y pueden ser capturadas con unos electrodos colocados en el exterior del cráneo. Previamente, estas corrientes iónicas deben ser convertidas a eléctricas, condición necesaria para que los electrodos metálicos puedan transportar la corriente hasta el amplificador de instrumentación. Para maximizar la eficiencia de la transducción iónica a eléctrica, se suele usar un electrolito en contacto con la piel y un electrodo de oro, plata o algún derivado químico. La señal eléctrica recogida, se amplifica y se representa en formas de ondas a lo largo del tiempo (ver Figura 3).



Figura 3: Ejemplo electroencefalograma (EEG)

El EEG es el encargado de leer el sistema nervioso y a partir de él, poder realizar el análisis pertinente. El sistema nervioso es un conjunto de tejidos dentro de nuestro cuerpo, encargados de captar y procesar rápidamente las señales internas y externas, tomando el control y coordinación sobre los demás órganos, para así, lograr una oportuna y eficaz interacción con el medio ambiente cambiante. La unidad básica del sistema nervioso es la neurona, la cual tiene la capacidad de comunicarse eléctricamente con otras células, sean éstas nerviosas o no. La información viaja entre neuronas por

CAPITULO 3: PROCESADO DE SEÑALES EEG

medio de impulsos eléctricos que se transmiten de unas neuronas a otras. Estos impulsos, se reciben de otras neuronas en las dendritas y pasan a través de la neurona hasta ser conducidas por el axón a los terminales de salida, los cuales pueden conectarse con otra neurona, fibras musculares o glándulas (ver Figura 4).

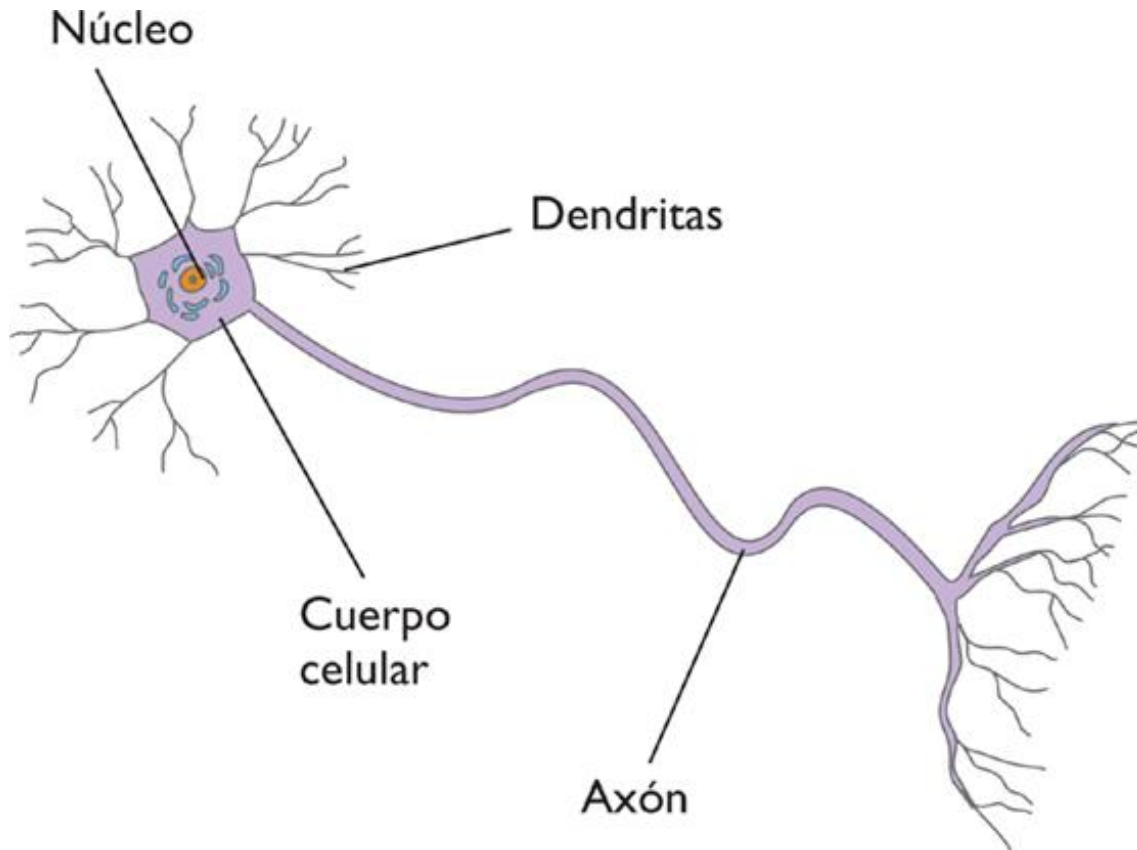


Figura 4: Representación neurona

Cuando se produce un estímulo externo, el sistema nervioso actúa de la siguiente manera. El estímulo es recibido en alguna región sensorial capturando alguna información, la cual es transportada por el sistema nervioso (a través de las neuronas) hasta una componente integradora en donde se analiza. Esta componente elabora la respuesta, que es conducida a través de las neuronas hacia fibras musculares (respuesta motora) o hacia glándulas (secreción glandular). Hay que tener en cuenta que la actividad cerebral es producida por un número muy elevado de neuronas (aproximadamente cien mil millones en un cuerpo humano medio) y cada una de las tareas que nuestro cuerpo puede realizar provoca una actividad cerebral con forma e intensidad diferentes, además de localizarse en distintas zonas del sistema nervioso.

3.2.1 Estado mental del sujeto

El concepto de estado mental del sujeto está fuertemente ligado al tipo de onda cerebral predominante en un momento determinado, como muestra la Tabla 1.

Ritmos	Situación mental a la que corresponde la onda cerebral
Delta	Con una frecuencia de 0.5-3 Hz, surgen principalmente en el sueño profundo y muy raras veces se pueden experimentar estando despierto. No suelen ser de interés en los BCIs
Theta	Con una frecuencia de 4-7 Hz, se producen durante el sueño o en meditación profunda, entrenamiento autógeno, por la práctica del yoga o cualquier otra actividad que trabaje la relajación físico-mental. Las características de este estado son: memoria plástica, mayor capacidad de aprendizaje, fantasía, imaginación e inspiración creativa. Se localiza en zona parietal-temporal.
Alfa	Tienen una frecuencia de 8–12 Hz y están asociadas con estados de relajación. Se registran especialmente momentos antes de dormirse. Está situado sobre las regiones posteriores de la cabeza y sobre las zonas occipitales. La amplitud es variable pero generalmente no es mayor de 50 microvoltios. En adultos se observa al cerrar los ojos o en estados de relajación y poca actividad mental. La amplitud disminuye durante la atención, especialmente visual, y durante el esfuerzo mental.
Beta	Originan un campo electromagnético con una frecuencia comprendida entre 13 y 30 Hz. Se registran cuando los sentidos se hallan volcados hacia el exterior y en plena actividad mental. Su amplitud es algo menor que las de ritmo alfa, localizándose en zona parietal y Frontal. Son más acusadas cuando tenemos el estado alfa bloqueado o en actividad mental intensa como estados de máxima alerta. A veces va asociada con el ritmo mu.
Mu	De frecuencia y amplitud similar al ritmo alfa. Generado por el área sensor motriz, situado en la zona central superior (posiciones C3, Cz y C4 del Sistema Internacional 10-20 descrito en el Anexo 10.1). Está presente en estados de reposo o concentración y su amplitud disminuye durante la realización de movimientos o su imaginación motora.

Tabla 1: Ritmos cerebrales

Aunque siempre predominen unas ondas frente a otras, todas están presentes durante el desarrollo de una actividad con diferente intensidad, según sea el tipo de tarea que se desarrolla. En nuestro caso, estarán presentes con mayor intensidad los ritmos Beta y Mu debido a la actividad mental del sujeto y se traducirán como ruido presente en los potenciales evocados P300. Esta idea mezclada con el concepto de cerebro modular, sugiere que para hacer un análisis adecuado capaz de describir alguna característica o función del cerebro, no se deben centrar los registros del EEG en unas determinadas zonas del cerebro exclusivamente, u otorgar un peso mayor a unos canales que a otros. En este proyecto fin de carrera se tendrá en cuenta la información de todos los canales

CAPITULO 3: PROCESADO DE SEÑALES EEG

(decorrelando la información de dichos canales) y el mayor rango posible del espectro de frecuencias.

La forma de las señales que el EEG recoge, está asociado a dos factores fundamentales:

- Estado mental en el que se encuentra el sujeto
- Actividad sensorial que se experimenta en el momento de la toma de los datos.

En la Tabla 1 se muestra la relación entre el tipo de excitación en la que se encuentra el cuerpo y la banda de frecuencia en el que se produce, junto con algunas de sus características más relevantes.

3.2.2 Electrodo para la medición del EEG

Los neurólogos creen que el origen de las tensiones que describe el EEG está en las células piramidales de la corteza cerebral [19]. Cada célula constituye un diminuto dipolo eléctrico, cuya polaridad depende de que la entrada en la red a la célula sea una inhibición o una excitación. La captación de las señales se realizará de diferentes maneras, desde la toma de datos en el cuero cabelludo, en la base craneal, con el cerebro expuesto o en localizaciones cerebrales profundas. Debemos capturar estas corrientes pero sin olvidarnos de la baja relación señal a ruido (SNR), ya que nuestros sensores también deben de ser alimentados aunque sea en el orden de los milivoltios. Esto supone una señal al menos 1000 veces mayor en amplitud que las medidas (del orden de microvoltios). Además, todas estas tensiones deberían ser iguales en todos los electrodos para que el amplificador diferencial no dé problemas

Los mejores materiales de electrodos sencillos y electrodos de pasta son sinterizados por plata y cloruro de plata (AgCl). La parte activa de estos electrodos se compone completamente de plata y cloruro de plata. Por lo tanto, la abrasión de la superficie, como ocurre siempre durante el uso normal, no deteriora las propiedades eléctricas de los electrodos y el efecto galvánico se reduce, hecho que resulta crucial para reducir el ruido térmico. Si se emplean electrodos de muy baja tensión galvánica se produce poco ruido de tensión en frecuencias inferiores a 1 Hz, por lo que los ruidos galvánicos y térmicos se reducen de manera importante, resultando muy beneficioso para mejorar la precisión de los registros. Existen por lo tanto diferentes tipos de sensores:

- **Electrodos en casco de malla**

Los electrodos están repartidos a lo largo de un casco elástico. Este tipo es el más usado para pruebas de larga duración debido a la comodidad que presenta con respecto a otro tipo de sensores. En el Anexo 10.1, Figura 63, tenemos un esquema del casco que ha sido empleado para las capturas muestrales de los datos con los que se ha trabajado en este proyecto fin de carrera, con las posiciones de todos los posibles canales a utilizar para las tomas de las muestras.

- **Electrodos de contacto**

Son pequeños tubos de plata unidos a soportes de plástico. En el extremo se coloca una almohadilla que se humedece con una solución conductora. Se sujetan al cráneo con bandas elásticas y se conectan con pinzas. Son sencillos de colocar, pero incómodos para el paciente, por lo que no permiten registros de larga duración.

- **Electrodos subdurales**

Se utilizan para registrar la actividad eléctrica directamente desde el cerebro. Los electrodos subdurales se implantan en quirófano bajo anestesia general. Estos electrodos permiten además la estimulación eléctrica de las neuronas para activar y estudiar ciertas partes del cerebro. Resultan incómodos y pueden ocasionar dolores de cabeza posteriores a su extracción, aunque se consigue mejorar la SNR notablemente.

3.3 La señal P300 y el paradigma *Oddball*

3.3.1 Señal P300

La onda P300 (EP300, P3) es un potencial relacionado con eventos que puede ser registrado mediante electroencefalografía como una deflexión positiva de voltaje con una latencia de unos 300 ms en el EEG [Referencia WEB 1].

Aunque la señal del EEG se adquiere con más fuerza en torno a los electrodos parietales, se ha sugerido que existen también interacciones entre las regiones frontales y temporales, así como con algunas áreas del encéfalo profundo. Se piensa que la onda p300 está compuesta por dos ondas secundarias conocidas como "señales P3a y P3b". Estos componentes responden individualmente a diferentes estímulos y se ha sugerido que la onda P3a "se origina en los mecanismos de atención frontal dirigidos por estímulo durante el procesamiento de tareas, mientras que la P3b se origina en la actividad parietal-temporal asociada con la atención y parece estar relacionada con el consiguiente procesamiento en la memoria." Algunas veces se denomina potencial evocado "objetivo (*target*)" a la P3b, mientras que la P3a sería el potencial "no objetivo (*non-target*)".

La señal P300 es un registro agregado de una gran cantidad de neuronas. Aunque es una técnica típicamente no invasiva, se pueden tomar muestras de partes de la señal más directamente de ciertas partes del cerebro mediante electrodos (de ahí la P300 temporal medial o MTL-P300). Esta metodología permite aislar y registrar localmente un área sin

CAPITULO 3: PROCESADO DE SEÑALES EEG

el ruido de otras señales adquiridas a través de electrodos sobre la piel del cráneo. En la práctica, la forma de onda de la P300 se debe evocar utilizando un estímulo provisto por alguna de las modalidades sensoriales. Un procedimiento típico es el paradigma *odd-ball*, en el que un estímulo diana se presenta entre estímulos de fondo más frecuentes. También puede usarse un estímulo distractor para asegurarse de que la respuesta se debe a la diana en lugar de a un cambio en el patrón de fondo. El paradigma *odd-ball* clásico ha sufrido muchas variaciones, pero al final la mayor parte de los protocolos utilizados para evocar la P300 implica alguna forma de hacer consciente una situación o una toma de decisiones. Se necesita ejercitar la atención en estos protocolos. No se han observado sujetos que efectúen un control fino sobre sus P300.

Como con cualquier estímulo evocado la P300 se produce mediante corrientes eléctricas en el cerebro. Los distintos potenciales inhibidores y excitadores postsinápticos de múltiples neuronas crean estas corrientes y de ese modo se puede discutir el origen de un potencial con respecto a los efectos de los neurotransmisores que provocan los potencial postsináptico y los dominios anatómicos en los que funcionan. La onda se genera directamente como resultado de potenciales postsinápticos a través de las vías glutamatérgicas, en la que los receptores de tipo NMDA desempeñan un papel importante. También modulan la actividad P300 influencias acetilcolinérgicas y GABAérgicas, actuando el GABA (ácido gamma-aminobutírico principal neurotransmisor inhibitorio cerebral) mediante la producción de potenciales inhibitorios post-sinápticos (PIPS) que tienden a disminuir la amplitud de onda e incrementar la latencia y la acetilcolina actuando como modulador con efectos opuestos al GABA. La Norepinefrina, dopamina y serotonina han sido implicadas en la modulación de la P300, pero los resultados son inconsistentes y estas influencias pueden ser menores.

Existe alguna controversia sobre los sustratos anatómicos de la P300. Entre las candidatas estarían estructuras profundas y cercanas del cerebro en el sistema límbico, como la amígdala, hipocampo y giro parahipocampal, así como regiones más dispersas (p.ej. las cortezas parietales superior y posterior, el giro cingulado y la corteza temporoparietal). Los registros intracraneales han dado credibilidad a la teoría de contribuciones ampliamente distribuidas. Una hipótesis vincula la activación de la P300 al locus ceruleus un centro noradrenérgico del tallo cerebral que se encuentra en el puente de Varolio. Desde esta perspectiva el papel del circuito noradrenérgico es potenciar los estímulos significativos para la toma de decisiones ejecutivas. Los orígenes de las ondas componentes (P3a and P3b) aún no están claras. Los estudios mediante fMRI sugieren que estos componentes se pueden localizar en diferentes regiones, en el que la P3a es principalmente un fenómeno de las cortezas frontales e insulares y la P3b se origina en las regiones parietal inferior y temporal. Se ha propuesto que las diferentes regiones también son lugares de procesamiento diferentes para los estímulos "diana" y "distractor" en el paradigma *odd-ball*. Los investigadores también están divididos en cuanto a si los potenciales evocados en general se generan como un modo inducido por estímulo (las poblaciones neuronales reaccionan disparándose o no ante el estímulo presentado) o en un modo de "reiniciado de fase" (las poblaciones

CAPITULO 3: PROCESADO DE SEÑALES EEG

neuronales reaccionan enfocando sus patrones de disparo en una fase particular basado en un estímulo). Una evidencia reciente que se basa en los estudios sobre P300 sugiere que ambos métodos probablemente contribuyen a la formación de un potencial evocado. Así pues, la P300 no se distingue por la dinámica de fondo, pero tampoco está completamente explicado por ella. Por tanto, la P300 no está provocada solamente por la reiniciación de la actividad oscilatoria preexistente, sino que más bien está afectada por ella. Ambas explicaciones encajan con la explicación más aceptada para el comportamiento a grandes rasgos de la P300, la hipótesis de actualización de contexto. Se piensa que la P300 representa en este modelo el cambio físico que subyace al hecho de que las probabilidades ligadas a ciertos resultados se refrescan a la luz de la información contextual más reciente para el sujeto.

Se ha implicado a muchos factores en la modulación de la P300. Los trastornos neurológicos muestran típicamente una reducción de la amplitud de la P300 en su conjunto, pero más allá de esto están los cambios que acompañan a la esquizofrenia. Las investigaciones que miden la P300 de la esquizofrenia comparándola con la de controles sanos han mostrado una marcada reducción en la amplitud de P300 sugiriendo un desarreglo del procesamiento de la información controlada. Estudios recientes que intentan desarrollar las bases para esta reducción han identificado un polimorfismo puntual en un gen que codifica la catecol-O-metiltransferasa (COMT) una enzima crítica en la eliminación de la dopamina en las regiones corticales del cerebro.

Los descensos en la amplitud e incrementos en el tiempo de latencia también se observan en el envejecimiento normal. Se acepta generalmente que los cambios en la amplitud de la P300 están relacionados con incrementos o bajadas en la intensidad, energía requerida o nivel de excitación ligada a una tarea específica. Los cambios en la latencia han sido peor caracterizados pero aún dan una correlación fuerte con el tiempo de procesamiento necesario para el desempeño de tareas. Las medidas de la onda P300 también pueden variar dependiendo del tiempo de medición. Se han señalado variaciones significativas de la amplitud y latencia basadas en el ritmo diurno de los sujetos. Estas variaciones precisan de gran cuidado con el uso del P300 como medida diagnóstica. No se sabe si estos cambios están directa o indirectamente relacionados con los de alerta/atención, que están bien documentados en su desarrollo a lo largo del día.

Un punto de partida interesante sería medir las P300 cuando se piensa que se va alterar la toma de decisiones conscientes. Se piensa que los estados hipnóticos y de sueño proporcionan buenas vías para la explicación de este campo. Las personas que son más receptivas a la hipnosis muestran cambios en la amplitud de los potenciales evocados entre los estados normales y los estados llamados de "alucinación obstructiva positivo/negativa". Esto parece estar en línea con la idea de que la P300 está regulada para transmitir decisiones a la conciencia y por ellos parecen sugerir la plausibilidad de una conciencia alterada en los estados hipnóticos.

CAPITULO 3: PROCESADO DE SEÑALES EEG

El alcoholismo también tiende a correlacionarse con una reducción de la amplitud de la P300. Se pueden efectuar comparaciones en las características de la P300 entre sujetos alcohólicos y los que tienen lesiones en el lóbulo frontal, sugiriendo una afectación a largo plazo de las funciones frontales relacionadas con el control ejecutivo. Se requieren más investigaciones para asegurar si este tipo de cambio en la P300 es específico del alcoholismo en lugar de una característica de la conducta adictiva en general. Las adiciones relacionadas con el tabaco han sido a menudo relacionadas con la conducta alcohólica y esto es una variable de confusión que aún no ha sido eliminada.

También se ha visto que los subcomponentes de la P300 varían con el género, mostrando las mujeres una amplitud y latencia mayores que los varones. No obstante existe una cantidad significativa de variación entre géneros en ambos subcomponentes y en la misma P300. Gran parte de esto puede deberse a las sutiles diferencias en los equipos de medición y en las técnicas (electrodos, lugares de registro, preparación de la cabellera). Otros investigadores han intentado probar que los daños traumáticos sutiles en la cabeza son típicamente seguidos en las P300 que duran más tiempo que la recuperación conductual. En un estudio que empleaba estudiantes universitarios sanos que habían experimentado un daño cefálico suave en el pasado, los sujetos afectados no revelaban diferencias con los controles en test de memoria y atención aunque aún mostraban amplitudes reducidas y latencias incrementadas en el marco de un paradigma *odd-ball*. El efecto del daño en la cabeza no se había investigado en ensayos previos y destaca como un posible candidato como variable de confusión para los resultados. La estimulación de secuencias más complicadas que una simple rutina *odd-ball* muestran variaciones en las características P300. Por ejemplo, cuando se asignan elecciones aleatorias a situaciones percibidas como beneficios o pérdidas las P300 demuestran la variabilidad de la magnitud recompensa/pérdidas. Se ha mostrado que las características de la P300 son independientes de la valencia de un resultado (de si representa pérdida o ganancia). Se piensa que los cambios en la P300 están relacionados con la evaluación y la reevaluación por parte del sujeto de opciones seleccionadas o no seleccionadas.

Por tanto, la respuesta en frecuencia de la P300 estará compuesta por la superposición de los diferentes ritmos cerebrales (ver Tabla 1). La componente predominante de la P300 se encuentra en torno a los 2.5 Hz (ritmo delta). Existe otra componente relevante en la banda de theta, 6 Hz que aparece en la ventana temporal 250-500 ms. Los ritmos alfa y gamma también estarán presentes aunque en menor medida. La Figura 5 muestra los potenciales evocados para los estímulos frecuentes, en rojo, y para los raros o poco frecuentes, en azul.

Debido a las diferencias en las respuestas cerebrales de cada sujeto, los experimentos han de realizarse de forma independiente. Es decir, hay que entrenar el sistema por separado para cada uno de los sujetos.

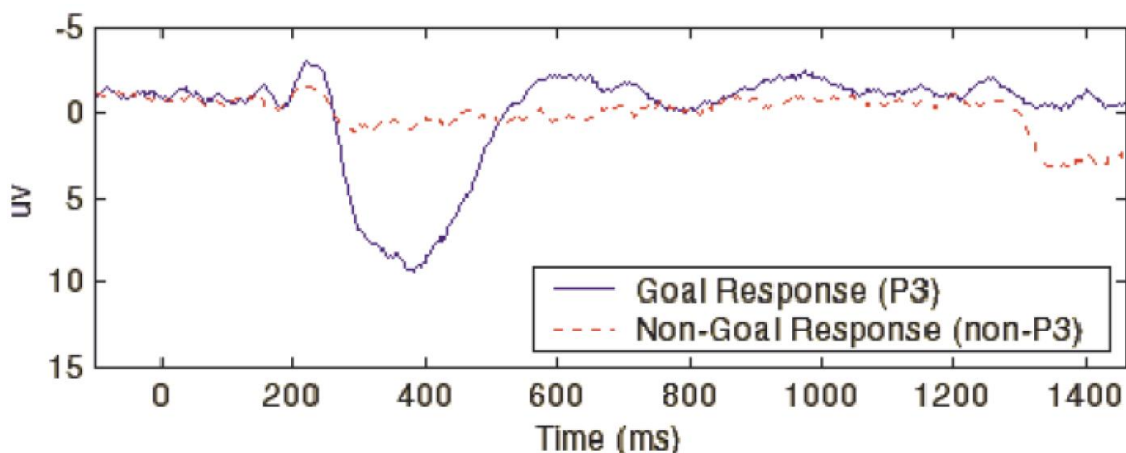


Figura 5: Representación de los potenciales evocados

3.3.2 Paradigma *Oddball*

El paradigma *Oddball* se basa en las demostraciones, realizadas en numerosos estudios, que han investigado la aparición de las señales cerebrales ERP. Es decir, discriminando en dos categorías diferentes: “probables” o “improbables” una serie de elementos, se pide al sujeto de estudio que determine a cuál de esas dos categorías pertenece un evento. Si a esta situación le añadimos el hecho de que una de estas categorías ocurre con muy baja frecuencia, “improbables”, se comprueba que los elementos pertenecientes a dicha categoría obtendrán un ERP con un componente positivo y una latencia de aproximadamente unos 300 ms, el cual será etiquetado como señal P300.

Esta señal, P300, muestra variaciones en su amplitud proporcionales a la relevancia de los eventos obtenidos e inversamente con la probabilidad del estímulo. Este hecho es de gran relevancia ya que supone que el sujeto no tenga que indicar mediante voz, pulsando un botón, etc., sobre la ocurrencia de un evento sino que mediante algo tan sencillo como enumerar mentalmente el número de eventos acontecidos (siendo esta tarea incluso innecesaria ya que el cerebro responde ante un evento objetivo, pero para asegurar la concentración del sujeto se le pide que lleve la cuenta), se consigue crear una comunicación, no oral, gracias a la aparición de las señales P300 que indican la existencia de un evento poco frecuente. Gracias a esta singular característica se abre la puerta, por medio del paradigma *Oddball*, a la realización de “prótesis mentales” que permitan la comunicación a personas que debido a un accidente o a una causa de nacimiento, se encuentren incapacitados de movimiento y habla, siendo este el objetivo que persigue la realización de este proyecto fin de carrera.

Esta prótesis mental permitiría al sujeto comunicarse mediante la elección continua entre una serie de elementos. En nuestro caso, estos elementos serían las 26 letras, aunque se debe indicar que existen otros modelos que contemplan espacios, puntuaciones, etc., que conforman nuestro alfabeto, mientras que la comunicación se

CAPITULO 3: PROCESADO DE SEÑALES EEG

llevaría a cabo mediante las señales P300 que irían apareciendo a lo largo del tiempo que requiere el experimento. Por el contrario, resulta evidente pensar que este sistema no permitiría una comunicación fluida ya que el sistema sería lento. Para intentar mejorar este hecho y acelerar el proceso, el sistema que se emplea recurre al uso de una matriz de caracteres que contemple que los eventos críticos serán presentados con una tasa elevada de ocurrencia. Esto puede suponer una degradación de las capacidades del sistema debido a los efectos refractarios de las neuronas, las cuales requieren de un tiempo de reposo de aproximadamente un segundo entre dos eventos objetivos, mientras que el tiempo entre eventos empleado es inferior a un segundo. Esto supondrá que en el caso en que se produzcan dos eventos objetivos consecutivos, la señal P300 del segundo estará más degradada y su amplitud será considerablemente de menor nivel que la de la señal P300 anterior.

Cabe destacar la importancia que este sistema tiene ya que permite a los sujetos con una parálisis completa poder comunicarse. Si bien existen otros medios que permiten la comunicación del paciente por medio de pequeños movimientos de manos, dedos o cabeza, gracias a esta prótesis mental no será necesaria ninguna interacción, en cuanto a movimiento se refiere, por parte del paciente. Así, el sistema que se presenta permite al sujeto “pulsar mentalmente”, metafóricamente hablando, un botón, mediante la focalización de su atención ante una serie de eventos que produzcan un estímulo en sus señales cerebrales.

3.4 El sistema *visual speller*

3.4.1 Introducción al sistema *visual speller*

La investigación en interfaces cerebro-máquina (BCI) ha suscitado la atención de los investigadores desde los años setenta a partir de las publicaciones de la Universidad de California Los Ángeles. Una BCI permite a una persona controlar un dispositivo ajeno a él usando para ello las señales procedentes de su cerebro, sin necesidad de activar nervios periféricos o músculos. Surge así un amplio abanico de aplicaciones que pasan desde su uso en PC y controladores de juegos a neuroprótesis para pacientes que sufren parálisis. Se ha demostrado que pacientes que se encontraban en una situación de bloqueo, en la cual la mayoría de funciones cognitivas se mantienen intactas a pesar de la completa parálisis en la que se encuentran, pueden aprender a comunicarse por medio de una interfaz que interprete las señales eléctricas provenientes del cerebro. La Figura 6 muestra un esquema del sistema *visual speller*.

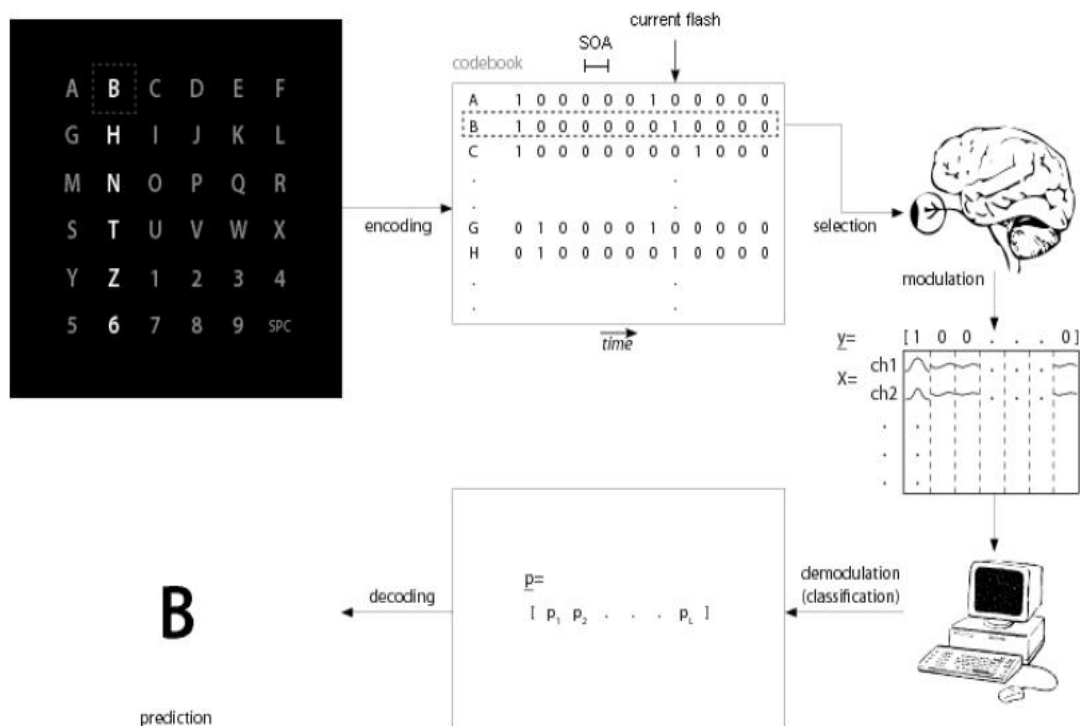


Figura 6: Esquema del sistema Visual P300 speller

CAPITULO 3: PROCESADO DE SEÑALES EEG

En general, el usuario intenta transmitir un mensaje mediante cambios en los diagramas de la actividad cerebral. El inicio de estos cambios puede ser desencadenado por el BCI o por el usuario. Es por ello que el BCI pueda ser considerado síncrono o asíncrono. La actividad cerebral es recogida y almacenada en formato digital mediante un dispositivo grabador. Una vez obtenida la grabación y basándose en ella, el BCI estima el mensaje que fue transmitido por el usuario. Este proceso de decodificación suele realizarse mediante el uso de software e involucra una etapa de pre-procesado y el empleo de modelos generativos o discriminativos que permitan estimar las letras transmitidas.

El *visual-speller* desarrollado por Farwell y Donchin es un claro ejemplo de un BCI síncrono. Este BCI emplea una pantalla de ordenador en la cual se muestra una matriz de letras (llamaremos letras a todos los caracteres que aparecen en dicha matriz ya sean letras, números, signos de puntuación u otro carácter cualquiera). En cada prueba se producen eventos estímulo sobre las diferentes letras. Así, el usuario puede seleccionar una letra sin más que focalizar su atención sobre una de las letras y llevar la cuenta de cuántos eventos estímulo se han producido sobre ella. Los eventos estímulo producidos sobre la letra elegida por el usuario serán denominados como eventos objetivo. Realizando una tarea mental, contar, después de cada evento objetivo, el usuario evoca una señal de atención-modulada en su encefalograma (EEG) como el potencial evento-relacionado P300. Por lo tanto, detectando las señales de atención-modulada se puede deducir cuando tuvieron lugar los eventos objetivo y por tanto cuál fue la letra elegida por el usuario. La información acerca de las letras elegidas también se conoce como la información oculta en las señales cerebrales, siendo esta información la que necesariamente debe ser obtenida de las señales cerebrales observadas. Esta tarea no resulta sencilla debido a la pequeña relación señal a ruido (SNR) de las señales cerebrales. Se considera ruido al EEG de segundo plano, movimiento de los ojos y las medidas de ruido del equipo de grabación.

En este proyecto fin de carrera se tomará como referencia el punto de vista de Hill [21], en el cual se presenta el sistema *visual-speller* como un canal de comunicación ruidoso. Este punto de vista resulta de mucha ayuda tanto para el diseño como para la decodificación del sistema *visual-speller*.

La teoría de la codificación es una parte importante de la teoría de la información que trata con el diseño de métodos de transmisión eficientes y precisos. La transmisión involucra un transmisor, un canal de comunicación y finalmente un receptor. El transmisor envía cierta información a través del canal, que puede verse corrompido por el ruido. Finalmente, el receptor recibe dicha información corrompida por el ruido proveniente del canal (ver Figura 7).

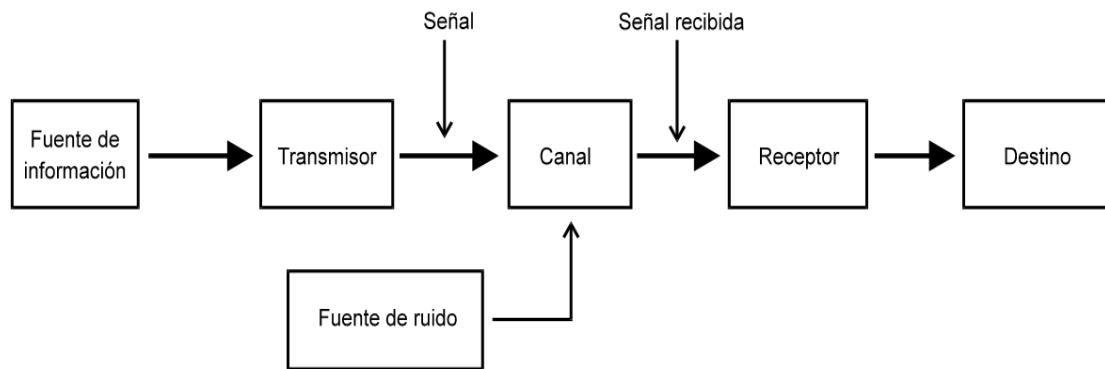


Figura 7: Esquema tx - canal ruidoso - rx

Parece lógico pensar que si el ruido es muy elevado, la transmisión de la información a través del canal se vuelve inviable. Sin embargo, Shannon demostró que pese a que el canal se encuentre contaminado por el ruido, se puede conseguir comunicar información que se encuentre cerca de estar libre de error para una tasa máxima a través del canal, conocida como capacidad del canal, cantidad que se encuentra ligada a la SNR. Esto involucra la codificación de la información en palabras código, en las cuales, la información se presenta de una forma redundante y permite por lo tanto una decodificación de lo recibido, presumiblemente corrupto, obteniendo una palabra a la salida. El conjunto de palabras código es conocido como código o *codebook*.

La redundancia en las palabras códigos se puede conseguir incrementando la mínima distancia Hamming del *codebook*, por ejemplo, el mínimo de números de entradas que difieren entre dos palabras código cualesquiera, como se explicó en el capítulo 2. Para ayudar a la comprensión de esta idea habría que imaginar un *codebook* en el que todas las palabras código tuviesen una distancia elevada. Luego, incluso si algunas de las entradas o bits en la cadena de bits recibida ha sido corrompida por el ruido y se ha vuelto errónea (por ejemplo, un 0 se convierte en un 1), la salida seguirá teniendo una distancia mucho más cercana a la palabra código transmitida antes que con cualquier otra palabra código. Eligiendo la palabra código con la mínima distancia a la salida permitirá obtener la palabra código correcta pese a que algunos de los bits se hayan visto afectados por el ruido. Consecuentemente, la mínima distancia de Hamming del *codebook* es una medida de la capacidad de corrección del *codebook*. La mínima distancia Hamming es una buena medida de la calidad del *codebook* si el canal es sin memoria (ver capítulo 2), por ejemplo, cuando la distribución de probabilidad de un bit recibido en un determinado momento de tiempo sólo depende del bit transmitido en ese mismo instante de tiempo y no en los bits transmitidos con anterioridad.

En la teoría de codificación existen diferentes modelos de decodificación como máximo a posteriori (MAP), decodificación de máxima verosimilitud y decodificación por vecinos más cercanos. Si el espacio de las palabras código es elevado, podrían aplicarse

CAPITULO 3: PROCESADO DE SEÑALES EEG

técnicas de decodificación más eficientes como *min-sum* o el algoritmo *forward-backward*.

Ahora, considerando el *visual-speller*, el conjunto de información es una letra de las presentes en la matriz de letras. El usuario puede comunicar una letra usando su cerebro como un canal de comunicación y el amplificador EEG como el receptor. Comunicar una letra por el simple hecho de pensar en ella podría producir señales cerebrales con una SNR muy baja por lo que la capacidad y la tasa de comunicación del canal correspondiente serían también muy bajas. En su lugar, las letras están asociadas a unas secuencias de eventos estímulo. La secuencia de eventos estímulo en las letras puede ser representada por palabras código binarias. Por ejemplo, un 1 en una palabra código en particular hace referencia a un evento estímulo en una letra mientras que un 0 hace referencia a un evento estímulo en el que la letra estuvo involucrada. Mediante la observación de los eventos estímulo de una letra de la matriz en particular, el usuario comunica implícitamente la información acerca de su palabra código. Partiendo del hecho de que la salida del canal es la señal cerebral observada la decodificación consiste en estimar la palabra código de dicha señal.

Gracias a esta codificación, podemos asemejar de nuevo, el sistema *visual speller*, a un canal de comunicaciones, pudiendo seguir dos caminos diferentes en el receptor, la decodificación dura o la blanda. Si se opta por emplear la decodificación dura, esta se realizará, en primer lugar, a nivel de bit para posteriormente decodificar la palabra código entera. Por otro lado, y siendo la decodificación empleada en el modelo generativo, si se opta por la decodificación blanda, se decodificará a nivel de palabra código, empleando un decisor MAP (Capítulo 2.3) que elija la palabra código que más se asemeje a una dada.

Si tenemos en cuenta el *visual speller* original de Farwell y Donchin se puede comprobar cómo propusieron algunas aproximaciones discriminativas, aunque sin un marco de trabajo probabilístico. Estas involucran el área, el análisis discriminante paso a paso, *peak-picking* y las medidas de la covarianza (ver capítulo 2) en un canal EEG. A continuación se muestra cómo lo hicieron.

CAPITULO 3: PROCESADO DE SEÑALES EEG

1. Análisis discriminante paso a paso (SWDA).

SWDA (*stepwise discriminant analysis*) es un procedimiento de clasificación. Este método obtiene una puntuación que mide la “distancia” entre cada época y la media del grupo de pruebas para incluir una P300. Esta puntuación se obtiene mediante la aplicación de una función discriminante de los datos de cada época. Dicha función fue desarrollada en base al conjunto de entrenamiento de las pruebas guardadas mientras el sujeto focalizaba su atención en las dos primeras letras (B, R). Las restante ERP proveyeron el conjunto de análisis. Se usó por tanto, el conjunto de entrenamiento para calcular la función discriminante con la que distinguir entre las subpruebas asistidas, es decir, los 600 ms siguientes a la iluminación de una columna o fila que contiene la casilla asistida, y las subpruebas no asistidas, los 600 ms siguientes a la iluminación de una fila o columna que no contenga la casilla asistida.

2. *Peak-picking*.

La amplitud de la P300 fue calculada como la diferencia entre el menor punto negativo previo a la ventana P300, siendo ésta el intervalo de tiempo donde la media de la forma de onda asistida en el conjunto de entrenamiento para cada sujeto fue positiva, y el mayor punto positivo en dicha ventana. La ventana P300 suele tener un rango típicamente entre 220 y 500 ms.

3. Área.

El área de la P300 fue calculada como la suma de los puntos de datos en la ventana P300, la cual fue descrita anteriormente.

4. Covarianza.

Se calculó una plantilla para la P300 como la media de los sub ensayos asistidos en el conjunto de entrenamiento para cada sujeto. La puntuación P300 en el conjunto de análisis fue hallada mediante el cálculo de la covarianza de cada sub ensayo con dicha plantilla. Así, la covarianza se calculó usando todos los puntos en la época de 600 ms.

Un método similar fue el propuesto por Sutter para un diferente tipo de sistema de speller basado en potenciales evocados (VEP – *Visual Evoked Potentials*). Su método estimó la respuesta cerebral a una secuencia de eventos estímulo. Este método también considera sólo un canal EEG aunque otros canales EEG contengan también información relevante.

CAPITULO 3: PROCESADO DE SEÑALES EEG

Más tarde, se lograron buenos resultados mediante la mejora de la aproximación discriminativa, método que trata la decodificación como si fuera un problema de clasificación per-bit [3], [10]. Cabe remarcar que los clasificadores de uso más frecuente como pueden ser las máquinas de vectores soporte (SVM) o el análisis discriminativo paso a paso (SWDA) no están diseñadas para ofrecer las medidas probabilísticas adecuadas, resultando, además, poco evidente en la combinación de las salidas del clasificador para los diferentes eventos estímulo a lo largo del tiempo. Sin embargo, tomando el producto interno de las palabras código y de las salidas del clasificador y tomando posteriormente la palabra código cuyo resultado sea el mayor, parece ofrecer resultados satisfactorios. Aunque las comparaciones entre la realización de los enfoques discriminativos frente a las aproximaciones generativas son escasos, estos parecen mostrar que las aproximación discriminativa supera a la generativa. No obstante, en la memoria de este proyecto fin de carrera se realizará una comparativa entre los resultados obtenidos por un modelo discriminativo y el generativo desarrollado en este proyecto.

Por el contrario y pese a la actual popularidad de las aproximaciones discriminativas, hay que considerar que el modelo generativo vale realmente la pena, ya que requieren de menos datos de entrenamiento para aproximar su error al del los métodos discriminativos [4], [5]. Una explicación posible sería que el modelo generativo restringe el espacio de posibles modelos y elimina el sobre-entrenamiento gracias a la incorporación de un conocimiento previo acerca de los datos, mientras que en la aproximación discriminativa esta regulación del espacio modelo se obtiene generalmente, mediante una compleja validación cruzada (ver capítulo 2).

En los sistemas BCI resulta de gran importancia la rápida convergencia relativa al método de clasificación, ya que la tarea de adquisición de datos de entrenamiento consume mucho tiempo y resulta tediosa. De hecho, si el BCI va a ser usado en sujetos, la velocidad de clasificación de la curva de aprendizaje es incluso más estricta debido a la reducida capacidad de atención del sujeto [6], [7].

Por otro lado, debido al hecho de que la realización de la decodificación indica cómo de bien es capaz de representar la realidad el modelo generativo, una comparación entre los diferentes modelos generativos permite un mejor entendimiento de cómo las señales cerebrales son generadas. Es probable que esta información pueda ayudar a la mejora de los paradigmas BCI, consiguiendo unas mayores relaciones señal a ruido de las señales cerebrales.

3.4.2 El sistema *visual speller* empleado

En primer lugar, como se ha indicado anteriormente, se muestra por pantalla una matriz de 6x6, la cual contiene las letras del abecedario seguidas de los números y un último carácter, como se puede observar en la Figura 8.



A	B	C	D	E	F
G	H	I	J	K	L
M	N	O	P	Q	R
S	T	U	V	W	X
Y	Z	1	2	3	4
5	6	7	8	9	_

Figura 8: Matriz de caracteres

La Figura 8 muestra también un patrón de iluminación donde las letras resaltadas son escogidas tras definir un *codebook* determinado por su distancia de Hamming

Gracias a esta matriz podemos definir dos diferentes tipos de eventos: eventos estímulo o eventos relevantes.

- Llamamos evento estímulo a aquellos ocurridos durante las pruebas realizadas consistentes en la iluminación, tanto de una columna como de una fila (también existen experimentos donde la iluminación de las letras es aleatoria o que incluso se rotan los caracteres).
- Consecuencia de esto será, por tanto, el sujeto quien determine qué eventos son relevantes, ya que en cada instante intentará comunicar el contenido de una celda de la matriz focalizando su atención en ella.

Existen por tanto 12 posibles eventos (6 eventos columna y 6 eventos fila) donde tan sólo dos de ellos serán relevantes (como se indicó anteriormente, estos eventos tienen probabilidad reducida, es decir, ocurren con muy baja frecuencia) correspondientes a la columna y fila que contenga el carácter que se desea transmitir.

Por tanto, cualquier flash que contenga la celda en la cual el sujeto se está concentrando generará una señal P300. En cada flash se medirá la amplitud de la señal P300 para posteriormente poder determinar la celda elegida, como aquella celda que corresponda a la intersección de la fila y la columna que obtengan una señal P300 cuya amplitud sea mayor.

Por tanto, se considera el cerebro como un canal de comunicaciones, el cual codifica un mensaje de entrada como actividad cerebral. Nuestro objetivo será por tanto decodificar

CAPITULO 3: PROCESADO DE SEÑALES EEG

dicho mensaje mediante funciones probabilísticas, donde denominaremos como x a la actividad cerebral e y al estado mental.

Para ello existen dos alternativas: el modelado generativo y la aproximación discriminativa, que se explicarán en el capítulo 4.

A medida que se van presentando los diferentes estímulos controlados al sujeto de estudio, se irá grabando su encefalograma (EEG). Si presentamos los eventos estímulo como una cadena de bits, referida a la letra elegida de entre las existentes en la matriz, podemos considerar la cadena como una palabra código en un canal de comunicaciones ruidoso (ver Figura 7).

Si nos fijamos ahora en las palabras código obtenidas, podemos comprobar cómo aquellas posiciones de la misma en donde tengamos un bit igual a 1, se corresponden con los eventos estímulo asociados a una letra elegida que han participado a lo largo del experimento, mientras que el resto de bits valdrán 0. Gracias a este sistema podremos representar cada letra de la matriz por medio de palabras código y considerar todo el conjunto de palabras código como el *codebook*. La importancia de este *codebook* reside en los valores que toman los bits que conforman sus columnas, ya que cada uno de ellos representa un instante temporal, o lo que es lo mismo, un evento estímulo. Así, comprobando en cada columna que bits están a 1 sabremos que letras han tomado parte en qué evento estímulo.

Como ejemplo se muestra la Figura 9, donde, tomando como referencia el evento estímulo 6 (columna 6), se puede comprobar que los caracteres A, C, H, J, M, Q, U, X, Z, 2, 5, 7, 9 han sido activados en dicho evento estímulo (la Figura no muestra el *codebook* completo).

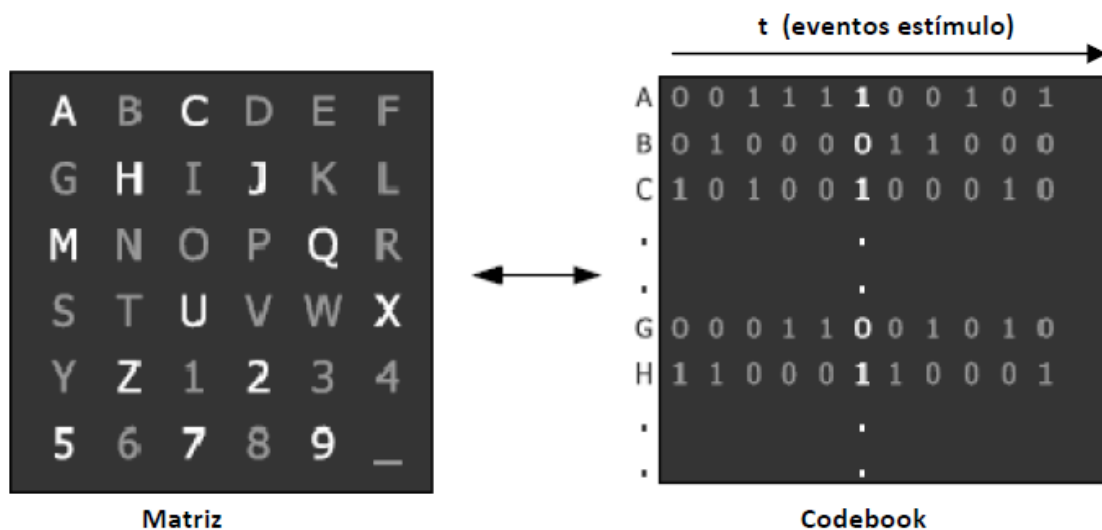


Figura 9: Matriz de caracteres y Codebook

CAPITULO 3: PROCESADO DE SEÑALES EEG

Como se dijo anteriormente, el sujeto de estudio transmitirá la información acerca de la palabra código de la letra que quiera comunicar, es decir, la palabra código objetivo. Por tanto, cualquier entrada cuyo valor valga 1 en esta palabra código objetivo significa que el evento será considerado como evento objetivo, mientras que los ceros serán eventos no objetivos. Mediante las diferentes respuestas cerebrales tanto de los eventos objetivo como los no objetivo sobre una letra, conseguiremos que el sujeto transmita la información de su palabra código mediante la cuenta de los eventos objetivos, el desestimar los no objetivos y el uso de las señales P300.

En definitiva, se pueden distinguir tres unidades en el sistema *visual speller*: codificación, modulación y decodificación, como se refleja en la Figura 10.

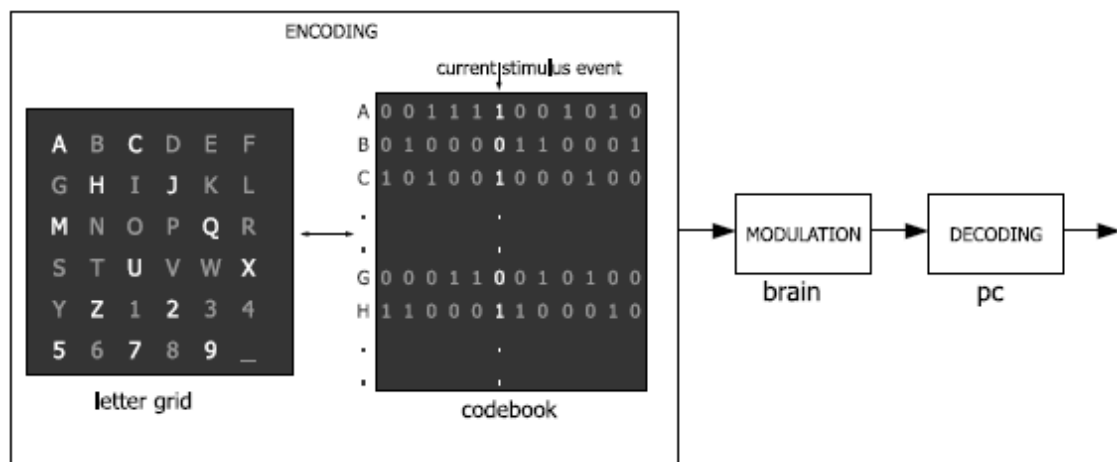


Figura 10: Unidades del sistema speller

- El codificador describe de qué forma cada letra es codificada como una secuencia de eventos estímulo a lo largo del tiempo. Estos eventos estímulo por letra pueden ser, por lo tanto, considerados como palabras código. Estas serán cadenas de bits de longitud N para cada una de las entradas correspondientes a un evento estímulo. De esta forma un bit con valor 1 indica que la letra ha participado en el evento estímulo y 0 si no. Podemos definir el peso del *codebook* como el número de 1's existentes en la palabra código. Por otro lado, se define distancia Hamming, entre dos palabras código de misma longitud, como el número de posiciones donde las dos palabras código toman un valor diferente. El conjunto completo de las palabras código para todas las letras de la matriz se denomina *codebook*, cuyas columnas representan un evento estímulo en un determinado momento.

- En el proceso de modulación, el cual sólo resulta relevante para el modelo generativo (Figura 10), los eventos estímulo de las letras elegidas por el sujeto son convertidos en unas señales cerebrales de atención moduladas. Consideramos un 1 en la correspondiente palabra código a un evento objetivo donde la señal cerebral evocada es la respuesta objetivo. De manera similar, un 0 en la correspondiente palabra código es un evento no objetivo y la señal cerebral evocada será la respuesta no objetivo. Produciendo diferentes respuestas cerebrales para los eventos objetivo y los no objetivo de una letra en particular, se consigue que el sujeto transmita la información acerca de la palabra código correspondiente a dicha letra. Una estrategia usada habitualmente (aunque la simple focalización en la letra deseada bastaría pero así se puede evitar faltas de atención) para el sujeto es llevar la cuenta de los eventos objetivo e ignorar los eventos no objetivo. Las épocas (tiempo posterior) resultantes siguientes a un evento objetivo contendrán las componentes de atención moduladas como señales P300 ERP con mayores amplitudes que las épocas siguientes a un evento no objetivo. El intervalo de tiempo entre el comienzo de un evento estímulo y el comienzo del siguiente evento se denomina intervalo de tiempo SOA (“*stimulus onset asynchrony*”). Por otro lado el intervalo de tiempo entre el comienzo de un evento objetivo y el comienzo del anterior evento objetivo se denomina intervalo TTI (“*target to target interval*”), mientras que el intervalo entre el comienzo de un evento no objetivo y el comienzo de un evento objetivo anterior es conocido como intervalo TNI (“*target to non target interval*”).
- La decodificación consiste en la obtención de la letra que fue elegida por el sujeto basándonos en las señales cerebrales medidas. Aplicando una estimación MAP (máximo a posteriori) podemos obtener la palabra código más probable \hat{c} dada la señal cerebral medida b :

$$\hat{c} = \underset{c \in C}{\operatorname{argmax}} p(c|b)$$

Donde c denota el *codebook* y $p(c|b)$ la probabilidad de la palabra código dada la señal cerebral observada b . La decodificación estándar en el *visual P300 speller* consiste en pasar un segmento del EEG después de cada evento estímulo a un clasificador. Este clasificador, el cual se emplea en el modelo discriminativo, es entrenado en respuestas objetivo y no objetivo con todos los bits de un conjunto de entrenamiento y asigna al valor de la salida del clasificador mayor que un determinado umbral para las respuestas objetivo y menor que dicho umbral para las repuestas no objetivo. Es una práctica común obtener la letra correspondiente a la fila y la columna cuya suma de los valores de salida del clasificador para todos los eventos estímulo sea mayor, o, siendo equivalente, la letra para la cual el producto interno ($\langle c, k \rangle$) de su palabra código con el vector de salidas del clasificador $k = [k_1 \ k_2 \ \dots \ k_N]$ es mayor. Esto significa que se elegirá la letra para la cual su palabra código satisface la siguiente expresión

$$\hat{c} = \underset{c \in C}{\operatorname{argm\acute{a}x}} \langle c, k \rangle$$

Resulta relevante, ya que es un hecho que afecta tanto a los modelos generativos como los discriminativos, indicar que las palabras código forman las filas de un codebook binario, donde las posición (i,j) de esta matriz, indica, cuando el valor es 1, que la letra correspondiente a la fila i en el instante j fue iluminada, mientras que un 0 indica lo contrario. Así, el código de fila-columna estándar, que será exactamente una fila o una columna iluminada en un momento de tiempo determinado, será denotado como RC.

Por último, tenemos que tener en cuenta la existencia de dos errores. Debido a la pobre relación señal a ruido de las ERP's, que obstaculiza la clasificación de las épocas, se puede producir que el bit de salida sea diferente que el bit que se transmitió (error de decodificación). El segundo error a tener en cuenta viene producido, por ejemplo, por la falta de atención del sujeto ante un evento estímulo, quedando este evento perdido. Este hecho produce que la cadena transmitida pueda diferir de la correspondiente fila del codebook y se denomina error de modulación.

3.4.3 Solape y efectos refractarios en el *speller*

Como se ha explicado, el *visual P300 speller* es una interfaz cerebro-máquina que permite a los usuarios deletrear palabras focalizando su atención en la letras, pertenecientes a una matriz de letras, mostradas en una pantalla de ordenador, donde la letra objetivo es aquella letra que el sujeto desea comunicar. La intensificación de la columna o la fila en la cual se encuentra dicha letra se denomina evento objetivo dentro de una secuencia de estímulos. Por otro lado, la intensificación del resto de columnas o filas en las que no se encuentra la letra objetivo se denomina evento no objetivo. Llamaremos época al periodo de tiempo durante el cual se espera se obtenga una respuesta evocada en el EEG en respuesta a un evento estímulo. El intervalo de tiempo existente entre el comienzo de un evento estímulo y el comienzo del siguiente evento estímulo se conoce como *stimulus onset asynchrony* (SOA). El intervalo entre el comienzo de un evento objetivo y el comienzo del evento objetivo anterior se conoce como *target-to-target interval* (TTI). De una forma similar, se puede presentar el *target-to-non-target interval* (TNI) que denota el intervalo entre el comienzo de un evento objetivo y uno no objetivo. Para cada hipotética letra objetivo, se puede representar sus eventos estímulo como una cadena de bits binaria, formada por 0 y 1, que denotan los intervalos en los que los eventos estímulo ocurrieron o no ocurrieron en relación a una letra objetivo determinada. El codebook es la colección de las diferentes cadenas pertenecientes a cada una de las letras de la matriz de letras. Ya que la respuesta cerebral del sujeto será diferente en función del tipo de evento que se le presente, objetivo o no objetivo, se consigue que el sujeto transmita implícitamente la información de la cadena de bits. Para obtener dichas respuestas cerebrales diferenciales el sujeto contará cada vez que se produzca un evento objetivo mientras que ignorará los eventos no objetivo. Las épocas resultantes siguientes a un evento objetivo contendrán componentes de atención-moduladas tales como los potenciales P300 relacionados a un evento (ERP) cuya amplitud será mucho más grande que la de las épocas siguientes a un evento no objetivo.

Se han realizado pocos intentos para la mejorar la configuración del sistema del *speller* desde su introducción en 1988. Existen estudios que han evaluado el rendimiento del sistema empleando matrices de letras con más caracteres. Se descubrió que la amplitud de la P300 era mayor en una matriz de letras mayor y la explicación a este hecho fue la diferente probabilidad de evento objetivo (menos probables al existir más caracteres). Siendo la probabilidad de objetivo la proporción de eventos objetivos dentro de la secuencia de eventos estímulo. También se han llevado a cabo estudios tomando diferentes valores de SOA. Así, Farwel y Donchin [1] demostraron que un SOA de 0.500 sg ofrecía mayores amplitudes de la P300 y una mayor precisión en la predicción de letras que con un SOA de 0.125 sg para un número de eventos estímulo fijado. Allison y Pineda [2] consiguieron demostrar la misma relación entre SOA y amplitud de la P300. Por otro lado, Sellers [2] comparó un SOA de 0.175 sg con uno de 0.350 sg y no encontró ninguna diferencia en la amplitud de la P300 ni en la precisión de

predicción de letras. Un estudio innovador llevado a cabo por Allison [2] involucró la comparación de diferentes patrones de flash en la matriz de letras. Compararon el patrón estándar, donde las columnas o las filas son iluminadas, con un patrón de estímulo múltiple, donde la mitad de las letras de la matriz se iluminan al mismo tiempo en cualquiera de los distintos eventos estímulo. La idea principal por la que emplear un nuevo patrón de iluminación era el buscar un *codebook* con buenas propiedades en cuanto a distancia Hamming se trata y a partir de dicho *codebook*, establecer el patrón de iluminación.

Su estudio permitió descubrir que, y resultando lo opuesto a lo esperado, para tasas de probabilidad de eventos objetivo mayores (33% - 50% de los eventos estímulo fueron eventos objetivo, codificados como 1s) bajo las condiciones de estímulo múltiple, la amplitud de la P300 no incrementó en función del SOA empleado. Se especuló que una tarea de mayor complejidad pudo ser la causa de la obtención de una mayor amplitud de la P300 en los experimentos de estímulo múltiple. También se descubrió que los sujetos cometieron un error cercano al 4% a la hora de contabilizar los eventos objetivo en el sistema estándar de iluminación de filas o columnas con un SOA de 0.125 sg. Un error a la hora de contabilizar los eventos objetivo indica, que algunos de estos eventos se pierden, obteniendo falsos negativos, que eventos no objetivos se consideraron como eventos objetivos, falsos positivos, o que el sujeto pierde la cuenta. Al usar los estímulos múltiples el error incremento hasta un 19% incluso cuando la probabilidad de objetivo y el SOA eran la misma que en el estímulo estándar. Se concluyó que este incremento de error venía por el hecho de perder la cuenta ya que no había efectos en la amplitud de la P300. En este caso, el rendimiento del sistema no debería verse afectado por el error al contabilizar los eventos objetivos.

Estos estudios demuestran que una mejor comprensión de los efectos neurofisiológicos y psicofisiológicos en el sistema *speller* podría resultar ventajosa. En los *visual-speller*, suelen emplearse valores de SOA bajos de aproximadamente 0.2 sg para conseguir tasas de transferencia de información elevadas. Sin embargo, la longitud de una época se establece entre los 0.6 sg y 1.0 sg, lo que conlleva al solapamiento de épocas. Surge entonces la pregunta de cómo el uso de valores de SOA bajos en los sistemas *visual-speller* introducen solape y efectos refractarios en los componentes ERP de larga latencia. Por ejemplo, una P300 que aparece aproximadamente después de 0.3 sg de producirse un flash objetivo, se seguirá manifestando en la época siguiente a la época objetivo. La evidencia de los efectos refractarios en las ERP viene de estudios auditivos donde se usaron eventos objetivos y no objetivos, que demostraron que la amplitud de la P300 decae mientras que la latencia de la P300 incrementa si el TTI se reduce de 8 a 2 sg. La misma relación entre TTI y amplitud de la P300 fue descrita en 1980 incluso para valores menores de TTIs en un estudio usando sólo objetivos. Los resultados de este estudio indicaron un periodo refractario de la P300 de aproximadamente 0.9 sg. Aunque en media el TTI está cercano al segundo en la configuración del *speller* estándar, de hecho el TTI de cada época individual varía extensamente ya que el orden de flash de columna o fila es aleatorio (ver Figura 11).

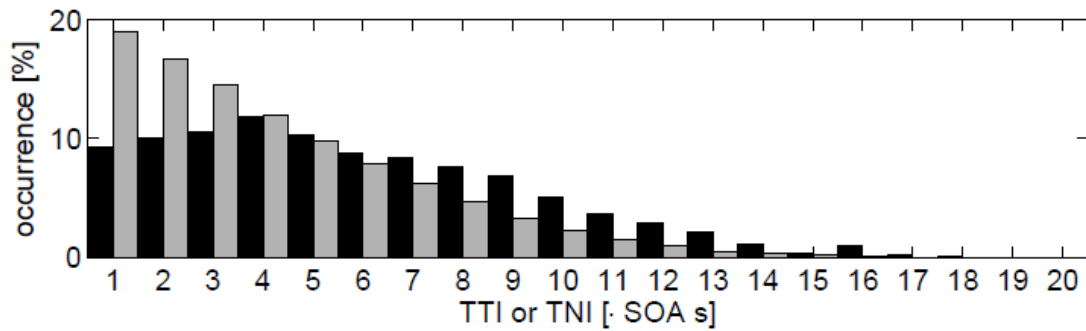


Figura 11: Representación TTI y TNI

Por ejemplo, para una matriz 6x6, el TTI estará dentro del rango entre $1 \cdot \text{SOA}$ (aproximadamente 0.2 sg) y $2 \cdot \text{SOA}$ (cerca de 0.4 sg). En consecuencia, los efectos refractarios pueden jugar un papel importante en los eventos objetivos en el *visual P300 speller*. En este proyecto fin de carrera se tendrán en cuenta estos efectos y se trabajará con dos tipos de estímulos diferentes, el estándar (Flash) de columnas y filas y un nuevo estímulo denominado estímulo Flip donde cada letra de la matriz de letras se sitúa dentro en un rectángulo gris con orientación vertical u horizontal. En este caso, el estímulo consistirá en la rotación de 90° de dicho rectángulo, entre un *frame* de video y el siguiente (ver Figura 12). Este tipo de estímulos, Flip, crean la sensación de movimiento mientras que el modelo estándar Flash no. Curiosamente, los efectos refractarios de las ERPs destacan menos o incluso se ausentan empleando el estímulo tipo Flip.

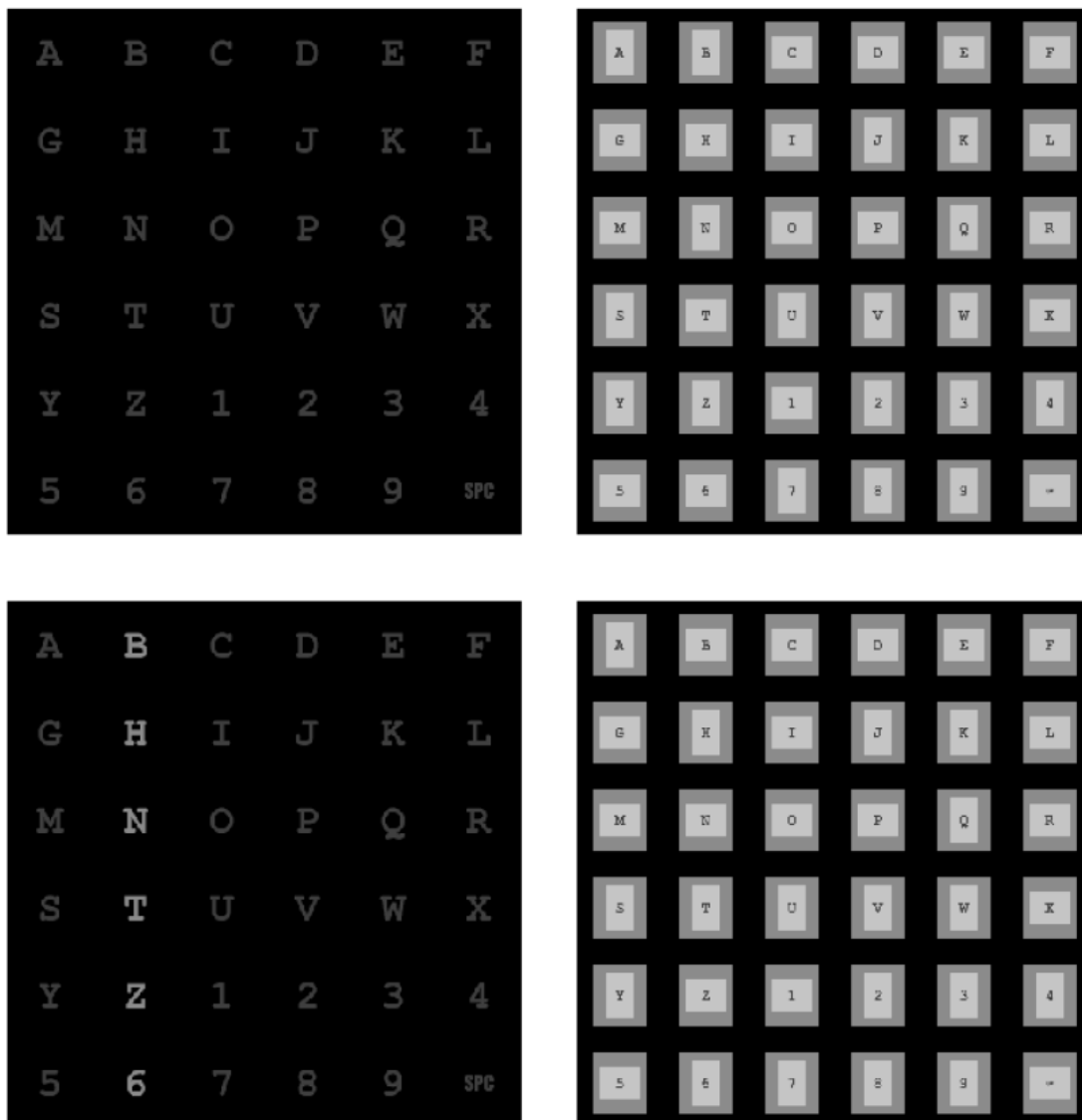


Figura 12: Ejemplo estímulo Flash y Flip

3.4.4 Distancia de Hamming y el *Visual Speller*

Aplicando la teoría de codificación, se sabe que se puede detectar y corregir transmisiones y decodificar errores simplemente añadiendo redundancia a la cadena de bits transmitida. La distancia de Hamming d , es el número de posiciones de bit en que difieren dos filas en el *codebook*. Así, la distancia mínima de Hamming, d_{\min} , de todas las parejas de palabras código está relacionada con las habilidades de corrección de errores del código mediante la siguiente expresión

$$e = (d_{\min} - 1)/2$$

donde “e” es el número máximo de errores que un código puede garantizar que es capaz de corregir.

En los modelos generativo y discriminativo, la redundancia se añade mediante la repetición de la iluminación de todas las filas o columnas R veces. Esto nos lleva a $d = 4R$ entre dos letras que no se encuentran en la misma fila o columna y a una $d_{\min} = 2R$ entre dos letras que se encuentran en la misma fila o columna. El código RC resulta pobre en cuanto a mínima distancia de Hamming, ya que para codificar 36 letras diferentes en 12 bits, una $d_{\min} = 4$ es posible y la distancia mínima realizable d_{\min} incrementa supra-linealmente con la longitud total del código, L .

Es importante indicar que un código con mejor d_{\min} no tiene, necesariamente, que funcionar mejor, ya que influye de gran manera la forma en que se muestran las letras al sujeto. Como se indicó en anteriores apartados, a la hora de elegir cómo iluminar los diferentes caracteres presentes en la matriz, se pueden seguir diferentes estrategias. Si se opta por seguir la estrategia relacionada con la distancia Hamming, se diseñará un *codebook* con la mejor distancia de Hamming, pensando en las capacidades de detección y corrección de errores que este podría tener, y a partir de dicho *codebook*, establecer el patrón de iluminación. La otra opción, será establecer un patrón de iluminación por filas o columnas y a partir de él obtener el *codebook* resultante.

Será mejor siempre que las letras se muestren por la iluminación de columnas o filas que cuando se muestren mediante la iluminación de letras aleatorias y la razón por la que esto ocurre es porque el cerebro presentará una respuesta cerebral P300 de mayor amplitud cuando tenga que realizar menos tareas y más uniforme sean los estímulos que se le presenten. Así, un patrón columna/fila será más uniforme que uno donde la iluminación de caracteres sea aleatoria.

3.5 Pre-procesado de las señales

3.5.1 Filtrado de las señales

Para poder trabajar con las señales EEG, obtenidas del paciente, procedentes de los 58 canales se debe realizar un pre-procesado. La razón para ello, es la existencia de un componente común en los canales EEG denominada CAR (*Common average reference*), así como para conseguir eliminar el máximo ruido posible que afecta a las señales.

Para ello, se filtrarán los 58 canales mediante un filtro FIR de Bartlett-Hanning de orden 1000, tomando como frecuencias de corte 0.5 y 10 Hz, como muestra la Figura 13.

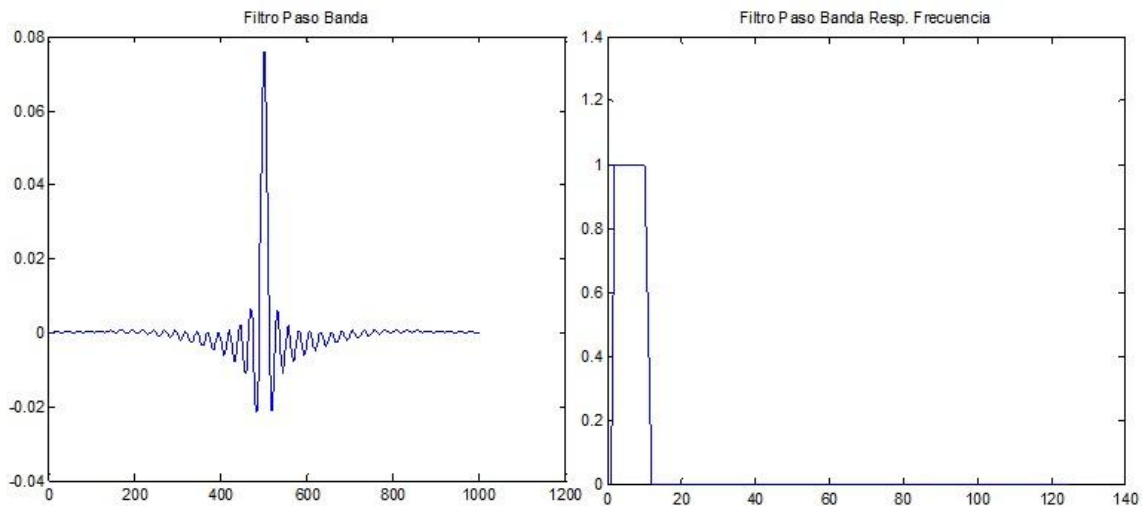


Figura 13: Filtro paso Banda

Resulta importante remarcar tres conceptos sobre la elección del filtro. En primer lugar, debido a que las componentes principales del espectro de las señales P300 se encuentran entre los 2.5 y los 6 Hz, la elección como frecuencias de corte superior, 0.5 Hz, e inferior, 10 Hz, nos permite mantener dichas componentes principales.

Por otro lado, gracias al filtrado conseguimos eliminar el ruido existente en altas frecuencias y corregir derivas en baja frecuencias.

Por último, debido a que posteriormente se realizará un sub-muestro a una frecuencia de 25 Hz, aseguramos que se cumple el teorema de muestreo de Nyquist (ver capítulo 2).

Conseguiremos así, realizar un primer filtrado (temporal) que nos permitirá reducir la redundancia existente en la señal (Figura 14) y por tanto la carga computacional.

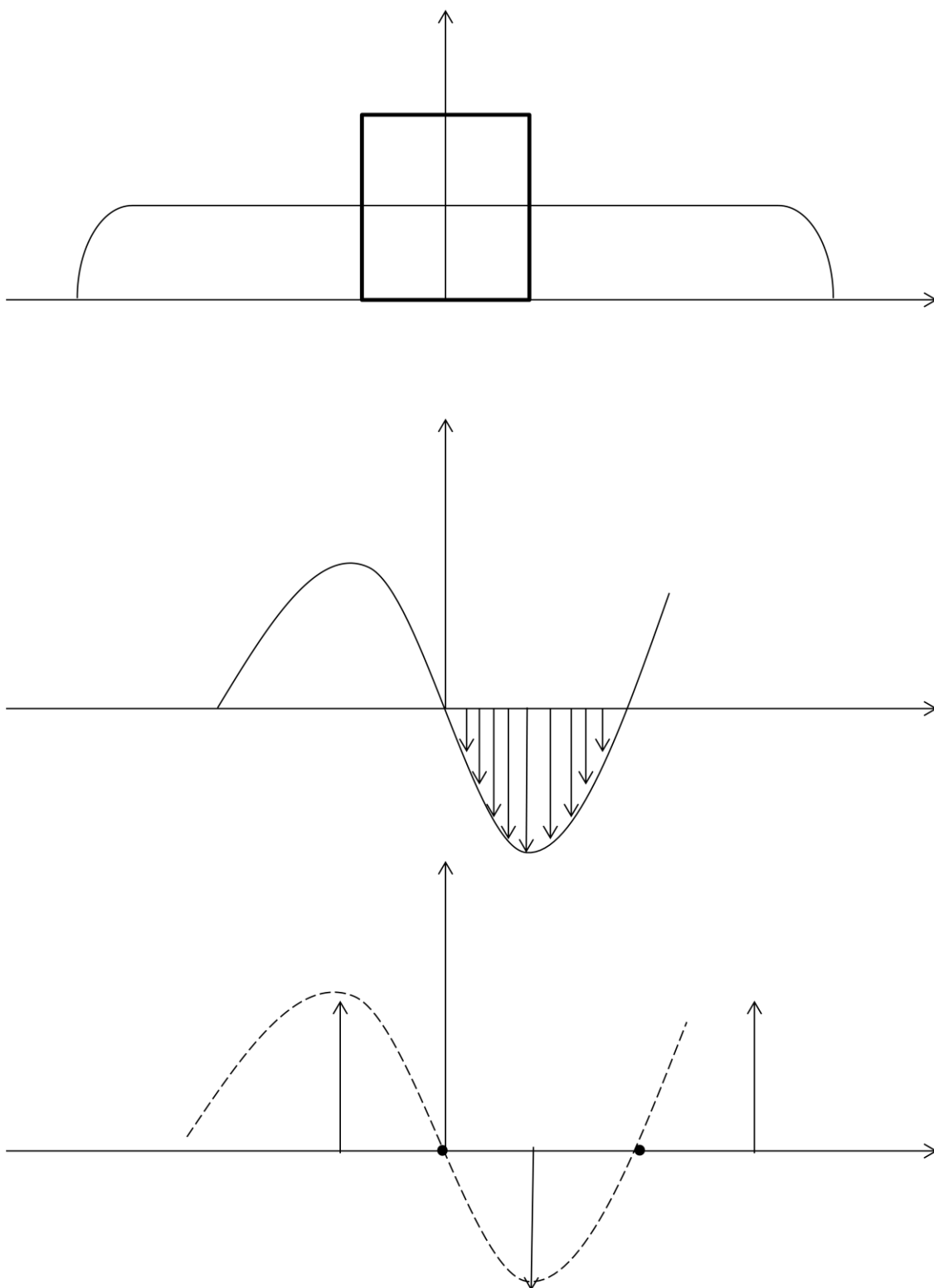


Figura 14: Reducción redundancia temporal. a) Espectro de la señal y del filtro. b) Señal muestreada a frecuencia original. c) Señal muestreada a la nueva frecuencia.

La Figura 14 se explica de la siguiente forma. Al conocer que las componentes principales de las señales P300 se encuentran a bajas frecuencias, podremos filtrar la

señal y obtener otra cuya frecuencia sea mucho menor. El muestrear esta nueva señal con la frecuencia de muestreo original nos hace tener muestras muy próximas en tiempo entre sí, lo que supone una gran correlación entre ellas y existencia de redundancia innecesaria. Pero gracias a que la nueva señal tiene una frecuencia menor que la original podemos escoger una frecuencia de muestreo menor (siempre cumpliendo Nyquist (ver capítulo 2)) que permita que eliminemos muestras redundantes.

Posteriormente al filtrado de las señales se procederá a la creación de los conjuntos de entrenamiento y de test a partir de los 64 experimentos diferentes a los que fue sometido el paciente. Para ello se empleó una validación cruzada mediante la cual se asigna un experimento al conjunto de test y los 63 restantes al conjunto de entrenamiento.

3.5.2 Análisis de componentes principales

El análisis de componentes principales, o PCA (*Principal Component Analysis*) permite decorrelar o blanquear un conjunto de señales y reducir las dimensiones del mismo. En nuestro caso no estamos interesados en reducir la dimensión, sino en conseguir decorrelar las señales de los 58 canales EEG entre sí.

Gracias al algoritmo PCA conseguiremos por lo tanto realizar un segundo filtrado (el primero fue el filtrado temporal), que nos permita eliminar la redundancia espacial de las señales. Esta redundancia aparece por la cercanía de algunos de los electrodos entre sí (ver Anexo 10.1), lo que supone que sus señales estén fuertemente correladas.

3.5.3 Introducción

El análisis de componentes principales (PCA), también llamado transformada de Karhunen-Loève, es una técnica estadística para el tratamiento de datos de forma compacta, que elimina la redundancia de los datos originales. Se trata por tanto de conseguir una representación de los datos lo mejor posible. Las técnicas PCA encuentran numerosas aplicaciones en diversas materias que requieren el tratamiento de datos, especialmente en compresión.

Dado un vector aleatorio x con n componentes, del cual conocemos una muestra $x(1)$, $x(2)$, ..., $x(T)$, el objetivo de PCA consiste en tratar de eliminar la redundancia existente entre cada uno de sus componentes, de forma que se obtenga un nuevo vector y con m componentes, tal que $m \leq n$. Gracias a esto podemos conseguir un vector de menor o igual dimensión, sin redundancia entre sus componentes.

En PCA, no se tiene en cuenta el tipo de función de densidad que poseen los vectores. Sólo supondremos que se pueden estimar los estadísticos de primer y segundo orden de la muestra.

CAPITULO 3: PROCESADO DE SEÑALES EEG

PCA trata de buscar un nuevo sistema de ejes coordenados formado por una base ortogonal, para representar al vector x de forma que se maximice la varianza de las proyecciones del vector en dicho sistema. Así conseguimos que las componentes de x en el nuevo sistema no estén correladas.

Para ilustrar el funcionamiento de PCA, se muestra en la Figura 15, la distribución de probabilidad conjunta de dos variables aleatorias x_1 y x_2 , obtenidas cada una de ellas a partir de la combinación lineal de otras dos variables aleatorias s_1 y s_2 independientes entre sí.

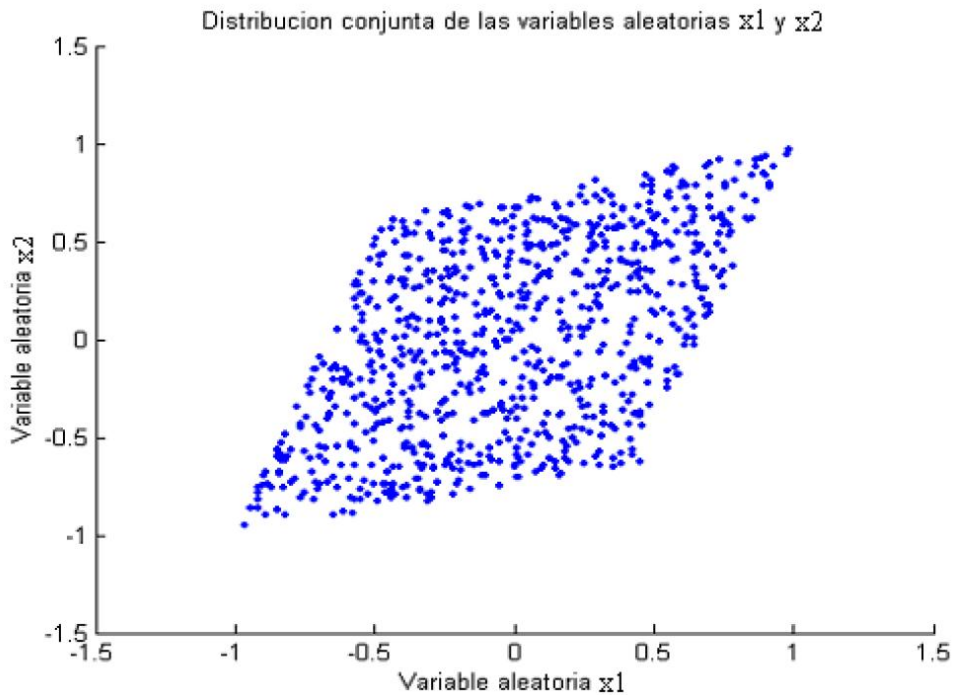


Figura 15: Distribución probabilidad conjunta de x_1 y x_2

Si a las variables aleatorias x_1 y x_2 se les realiza un procesamiento PCA, las nuevas variables que se obtienen no estarán correladas entre sí y tendrá la distribución de probabilidad conjunta que se muestra en la Figura 16.

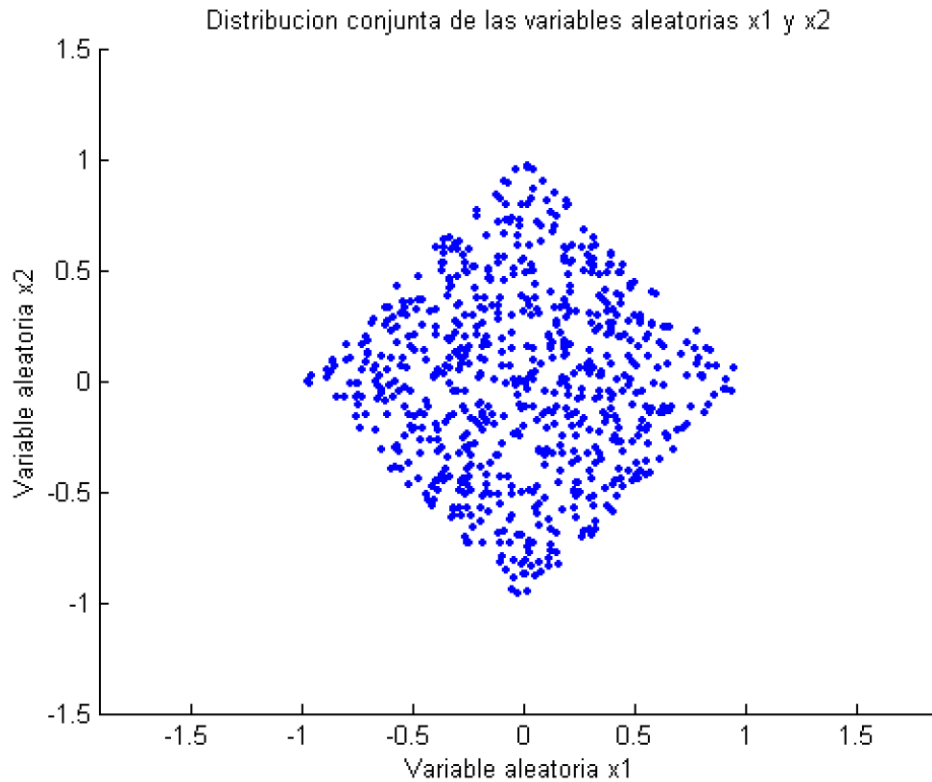


Figura 16: Distribución probabilidad conjunta de x1 y x2 decorreladas

Puede observarse cómo las nuevas variables obtenidas estarán decorreladas entre sí y su función de distribución conjunta estará contenida en un cuadrado.

3.5.4 Descripción del método

En PCA se quiere buscar las componentes principales de las observaciones, las cuales consisten en los autovalores de la matriz de covarianzas de estas. Para poder aplicar el método, debemos centrar el vector x sin más que restarle su propia media,

$$x = x - E\{x\}.$$

Representaremos con y , la nueva representación de los datos. Este nuevo vector tendrá una dimensión m , cumpliéndose que $m \leq n$, siendo n la dimensión de los datos originales.

La nueva representación de los datos tras aplicar las técnicas PCA vendrá dada por

$$y = W^T x.$$

Donde W es una matriz cuyas filas son vectores de dimensión n , tales que maximizan la varianza entre las componentes del vector y . Para dicha matriz sus columnas serán ortogonales entre sí, es decir

CAPITULO 3: PROCESADO DE SEÑALES EEG

$$W_i^T W_j = \delta_{ij}.$$

Puede demostrarse que las columnas de la matriz W coinciden con una base ortogonal expandida a partir de los autovectores $\lambda_1, \lambda_2, \dots, \lambda_m$. La mayor parte de la información estará contenida en el primer autovector e_1 . La cantidad de información contenida en cada autovector (e_j) será mayor cuanto mayor sea su autovalor (λ_j) asociado, de forma que en las primeras componentes se concentrará la mayor parte de la información. Al valor $y_1 = e_1^T x$ se le denomina componente principal de x , y es la que contiene más información de todas.

En el caso en que $m < n$, se producirá un error en la representación de los datos. Como ventaja, se estará comprimiendo la información contenida en el vector inicial. El error cuadrático mínimo que se puede cometer se dará al tomar una base ortogonal expandida a partir de los m autovectores e_1, e_2, \dots, e_m de la matriz de covarianzas. Dicho error viene dado por

$$J_{MSE}^{PCA} = d_i$$

Si $m = n$, como es el caso de estudio de este proyecto fin de carrera, no se cometerá error alguno en la representación de los datos en la nueva base formada por los autovectores de la matriz de covarianzas de los datos. Como se comentó anteriormente, lo que conseguiremos en esta situación será decorrelar o blanquear los datos.

Capítulo 4

Modelado probabilístico de las señales

4.1 Modelos discriminativos vs generativos

Como se indicó en el capítulo 3, a la hora de solucionar el problema presentado mediante el sistema *visual speller*, se puede seguir dos caminos diferentes: el modelo generativo y la aproximación discriminativa.

- El **modelo generativo** se basa en la obtención de la distribución de las señales cerebrales a partir del estado mental dado $f(x|y)$, donde $x \in X$ e $y \in Y$, o también mediante el uso de la probabilidad conjunta $p(x,y)$. Gracias a esto podemos obtener un modelo, el cual nos muestra cómo generar las señales cerebrales para todos los posibles estados mentales, para luego posteriormente aplicar el teorema de Bayes (ver capítulo 2), que nos permitirá alternar las variables x e y de forma que obtengamos $p(y|x)$ o $p(y,x)$. Gracias a este cambio podremos aplicar posteriormente una decodificación a posteriori o MAP, para poder encontrar, finalmente, el estado más probable ante una señal cerebral

observada. Para mejorar las estimaciones, cualquier conocimiento previo acerca de las señales cerebrales puede ser incorporado en el modelo generativo. Además, la realización de la decodificación nos da una indicación de cómo de realista resulta nuestra modelación de la parte de la señal cerebral involucrada en el estado mental codificado. Así, el modelo generativo permite “muestrear”, es decir, permite generar unas muestras de acuerdo con $f(x|y)$ y obtener las plantillas para la respuesta ante un evento objetivo y uno no objetivo.

- Por otro lado tenemos la **aproximación discriminativa**, mediante la cual se realiza la estimación de la probabilidad $p(y|x)$ directamente, sin tener en cuenta en ningún momento, el modelado de las señales cerebrales. Esta aproximación será más sencilla siempre y cuando la función $f(x|y)$ resulte compleja. Este sistema se basa en el uso de clasificadores. Así, podemos considerar la aproximación discriminativa como un modelo que consiste en un sistema entrenado, que ante una muestra x , obtiene como respuesta una probabilidad $p(y|x)$.

Centrándonos ahora en el problema estadístico referente al *visual-speller*, donde nuestra tarea es adivinar la letra de interés, mediante la traducción de la información acerca de la palabra código proveniente del EEG del sujeto. Denotaremos la palabra código como c y el conglomerado de canales de señales cerebrales (señales multi-canal) observado como b . Empleando una decodificación mediante el sistema MAP, como se indicó anteriormente, hallaremos la palabra código más probable de entre un conjunto de palabras códigos posibles, las cuales vienen dadas por las señales cerebrales. Es decir:

$$\hat{c} = \operatorname{argm\acute{a}x}_{c \in \mathcal{C}} p(c|b)$$

Donde \mathcal{C} es el *codebook* usado. Ya que $p(c|b) = \frac{p(c,b)}{p(b)}$ y considerando $p(b)$ independiente de las palabras código, la anterior expresión resulta equivalente a la siguiente:

$$\hat{c} = \operatorname{argm\acute{a}x}_{c \in \mathcal{C}} p(c, b)$$

En la realidad, tanto el aprendizaje del modelo discriminativo $p(c|b)$ en la primera expresión, como el modelo generativo $p(c,b)$ en la segunda, resulta una tarea complicada. Esto es debido a que la variable b , multi-canal observado, es evaluada continuamente y tiene una gran dimensión, mientras que la variable c , palabra código, tiene tanta dimensión como longitud el *codebook*. Además, para conocer $p(c|b)$ o $p(b|c)$

CAPITULO 4: MODELADO PROBABILÍSTICO DE LAS SEÑALES

necesitaríamos un conjunto de entrenamiento extremadamente grande, lo cual sería inviable.

Ya que el modelo generativo se explicará con más detalle a continuación, conviene dar un breve resumen acerca de los modelos discriminativos, recordando que estos no son el objeto de estudio de este proyecto, pero que resulta conveniente dar una explicación ya que se utilizarán valores obtenidos con modelos discriminativos para comparar nuestros resultados. Para ello, haremos uso de la referencias de S. Martens [2] y Jeremy Hill [21].

La tesis de S. Martens presenta un modelo gráfico (ver Figura 17) con el que obtener un modelo discriminativo (t no hace referencia a las plantillas objetivo, t , del modelo generativo). No es objeto de este proyecto profundizar en el modelado matemático que esto conlleva sino de dar una visión global al modelo discriminativo.

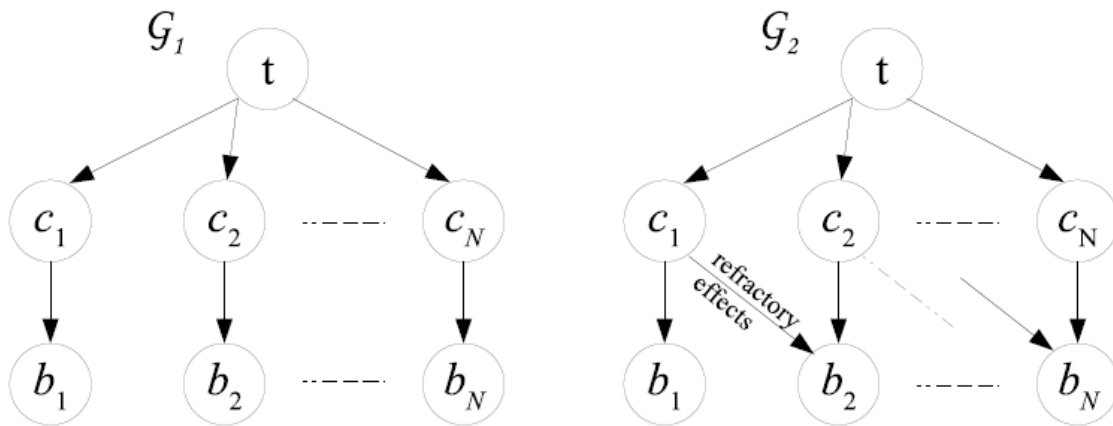


Figura 17: Ejemplo modelo gráfico

De esta forma, la aproximación discriminativa se puede dividir en dos partes. La primera, donde se decodifica a nivel de bit mediante el uso de SVM (*support vector machine*) u otra función discriminativa, para posteriormente realizar una decodificación a nivel de palabra. La SVM nos permitirá obtener $p(c_j|b_j)$ y a partir de ella podremos obtener $p(c|b)$, aplicando Bayes una primera vez

$$p(c|b) = \frac{p(b|c) * p(c)}{p(b)}$$

donde $p(b)$ puede extraerse por ser independiente, y asumiendo independencia estadística entre las señales, como se demuestra con la expresión

$$p(c, b) = p(c) \prod_{j=1}^N p(b_j|c_j)$$

CAPITULO 4: MODELADO PROBABILÍSTICO DE LAS SEÑALES

Finalmente, se puede estimar la palabra código sin más que aplicar una segunda vez el teorema de Bayes.

$$\hat{c} = \underset{c \in C}{\operatorname{argmax}} p(c) \prod_{j=1}^N \frac{p(c_j|b_j)}{p(c_j)}$$

donde las probabilidades $p(c_j)$ son conocidas debido al patrón de iluminación del sistema empleado (por filas o columnas). Al tener una matriz 6x6 y un patrón de iluminación por filas y/o columnas, podemos asegurar que la probabilidad de iluminarse la letra elegida por el sujeto, evento objetivo (1), será igual a $p(c_1) = 1/6$, mientras que la probabilidad de un evento no objetivo (0), será igual a $p(c_0) = 5/6$.

Volviendo al caso que nos afecta, gracias a la consideración de independencia en el proceso de generación de las señales cerebrales, se puede simplificar el aprendizaje de la probabilidad conjunta $p(c,b)$ en la expresión $\hat{c} = \operatorname{argmax} p(c,b)$ notablemente. Es decir, si asumimos que las señales cerebrales obtenidas en un evento estímulo b_j en particular, son sólo dependientes del bit c_j entrante y no de ninguno de los otros eventos estímulo ni en ninguna de las señales cerebrales pertenecientes a los otros eventos estímulo. Será en este caso cuando la probabilidad conjunta $p(c,b)$ pueda ser factorizada mediante el productorio de las probabilidades condicionales $p(b_j|c_j)$ y la probabilidad marginal $p(c_j)$, quedando finalmente:

$$p(c,b) = \prod_j p(b_j|c_j)p(c_j)$$

Puede observarse cómo el dimensionamiento del problema de aprendizaje se ha visto reducido ya que tanto b_j como c_j tienen menor dimensión que b y c respectivamente. Resulta interesante ver cómo, de nuevo, podemos volver a elegir entre la aproximación generativa, la cual se basa en el aprendizaje de $p(b_j|c_j)$ y la aproximación discriminativa, que recurre al aprendizaje de $p(c_j|b_j)$. La aproximación posterior convierte el problema en un problema de clasificación por bit. Es por ello que se puede obtener una decodificación MAP mediante la expresión de la probabilidad aprendida $p(c_j|b_j)$ como $p(b_j|c_j)$ gracias al teorema de Bayes (ver Capítulo 2), estimando la probabilidad condicionada $p(c,b)$ y posteriormente decodificando y obteniendo \hat{c} como se mostró en la expresión $\hat{c} = \operatorname{argmax} p(c,b)$.

4.2 Modelo generativo de señales cerebrales

El principal interés en realizar un modelo generativo en vez de uno discriminativo es, que los modelos generativos se pueden asemejar a un canal de comunicaciones ruidoso. De esta forma, podemos considerar el cerebro como un sistema que modula una señal en función de la entrada que se le presente, donde dichas entradas serán los eventos visuales objetivo o no objetivo que serán codificados en sistema binario. Posteriormente, el ruido será añadido por el medio empleado para adquirir las señales EEG, electrodos y cables. Así, nuestro modelo generativo se encargará de generar unas plantillas para la señal modulada según se presente un evento objetivo o no objetivo, para finalmente y siguiendo el modelo MAP, poder decodificar la palabra código transmitida y estimar por lo tanto la letra transmitida.

Por ello, será motivo de estudio el conseguir maximizar la tasa de acierto con el modelo desarrollado. Con ese fin, y partiendo de las propiedades conocidas de las señales cerebrales, se presentará el modelo generativo implementado con el que probaremos el sistema.

Teniendo en cuenta los estudios relativos a psicofisiología, que postulan que las señales cerebrales ERP poseen unas fuertes características deterministas [8], [9], podemos modelar la respuesta que ofrece el cerebro a un evento estímulo como un sistema determinista. Dicho sistema consta de los sistemas o matrices M_n y M_t , los cuales mapean o asignan una palabra código entrante c_j a una respuesta cerebral multidimensional h_j como se observa en la Figura 18, donde puede comprobarse cómo realmente la respuesta que se acaba midiendo en una versión ruidosa de h_j .

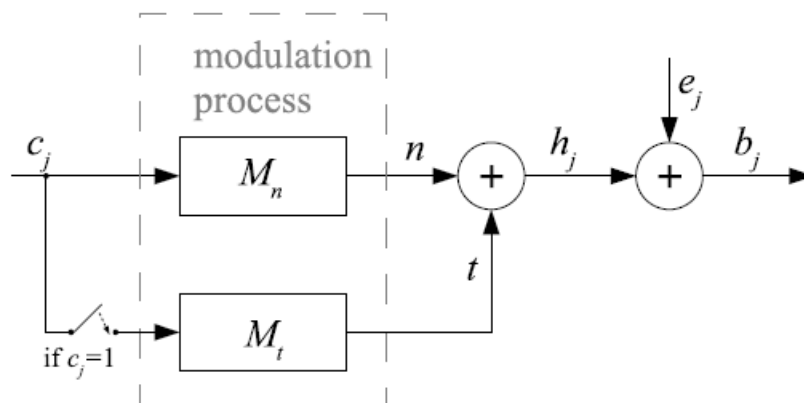


Figura 18: Modelo determinista de la respuesta cerebral ante un estímulo

CAPITULO 4: MODELADO PROBABILÍSTICO DE LAS SEÑALES

Cada palabra código c_j puede tomar un valor perteneciente al conjunto $\{0,1\}$, mientras que la respuesta cerebral h_j , es una variable evaluada continuamente con unas dimensiones $[N_s \times N_{ch}]$, siendo N_s el número de muestras temporales y N_{ch} el número de canales.

Se considera que cada evento estímulo c_j evoca una respuesta sensorial n , cuya duración es de SOA segundos, siendo este valor independiente del valor de c_j . SOA representa la asincronía de inicio del estímulo y se define como el intervalo de tiempo entre el comienzo de un evento estímulo y el comienzo del siguiente evento estímulo. La respuesta sensorial, denotada como n , tiene una dimensión de $[N_n \times N_{ch}]$ donde N_n queda definido como $N_n = [SOA * fs]$ muestras, donde el operador $[\cdot]$ indica el entero más cercano, ya sea superior o inferior. Además, si el evento estímulo fuera un evento objetivo, por ejemplo $c_j = 1$, una modulación objetivo deseada, t , es evocada durante 0.6 sg. Esta modulación t tiene unas dimensiones de $[N_t \times N_{ch}]$, donde N_t se define como $N_t = [0.6 * fs]$ muestras. Siguiendo el esquema presentado en la Figura 18 podemos escribir

$$h_j = \begin{cases} n, & c_j = 0 \\ t + n, & c_j = 1 \end{cases}$$

donde podemos conseguir las dimensiones correctas sin más que completando con ceros a "n", aunque no sería necesario ya que la plantilla suele ser de menor tamaño y existe solapamiento.

Como puede comprobarse observando la Figura 18 anterior, h_j no puede ser observada directamente. Sin embargo, podemos definir una época b_j , como un segmento del EEG en el cual esperamos capturar la respuesta cerebral oculta h_j , junto con un ruido no relacionado con la tarea, denominado e_j . En esta situación, el ruido representa la EEG de segundo plano, los aparatos y el ruido de medida. Así, pasaremos a denominar h_j como la respuesta cerebral oculta, mientras que b_j será la respuesta cerebral observada (Figura 19).

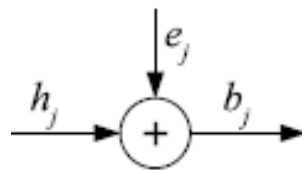


Figura 19: Generación de la señal cerebral observada con ruido

Si consideramos que el proceso de modulación es equivalente a un sistema lineal (ver capítulo 2), en el sentido en que el principio de superposición se mantiene, la respuesta cerebral oculta h referente a una secuencia de eventos estímulo c , puede ser construida sencillamente, mediante la suma de las respuestas ocultas h_j correspondientes a los eventos estímulo por separado c_j , en los instantes de tiempo donde los eventos estímulo j tienen lugar. Esta operación puede ser expresada mediante la convolución o mediante

CAPITULO 4: MODELADO PROBABILÍSTICO DE LAS SEÑALES

el producto de matrices, siendo estas matrices Toeplitz (ver capítulo 2). Considerando el producto de matrices Toeplitz, podemos obtener h de la siguiente manera:

$$h = S_t^T t + S_n^T n$$

Las matrices S_t y S_n se construyen de la siguiente manera. Partiendo de una palabra código c , de longitud N , donde la entrada que tome valor 1 quiere decir que es un evento objetivo, mientras que una entrada cuyo valor sea 0 indicará un evento no objetivo, podemos construir una palabra código ampliada, cuyas dimensiones serán $[1 \times N_s]$ con una frecuencia de muestreo f_s , simplemente mediante el relleno con ceros entre las diferentes entradas de la palabra código original (muestreada a nivel de evento). Así, se puede establecer una matriz Toeplitz S_t de dimensiones $[N_t \times N_s]$ donde la primera fila será la palabra código ampliada anteriormente obtenida. Las entradas vacías serán consideradas o establecidas a cero. Por otro lado, la matriz S_n se obtiene de una forma similar, sólo que partiendo de un vector de unos de dimensiones $[1 \times N]$. De nuevo, se rellena con ceros dicho vector, con una frecuencia de muestreo f_s y posteriormente se obtiene la matriz Toeplitz de la misma forma que antes. Cabe remarcar que el solapamiento proveniente de señales cerebrales de larga latencia con un SOA corto, [2] está explícitamente modelado en la expresión anterior al incluir versiones retardadas de la palabra código. En otras palabras, conseguimos situar todos los elementos objetivo y no objetivo uno seguido de otro para poder trabajar con ellos de una forma coherente.

En la Figura 20 se muestra gráficamente el proceso anteriormente descrito para una mejor comprensión del mismo.

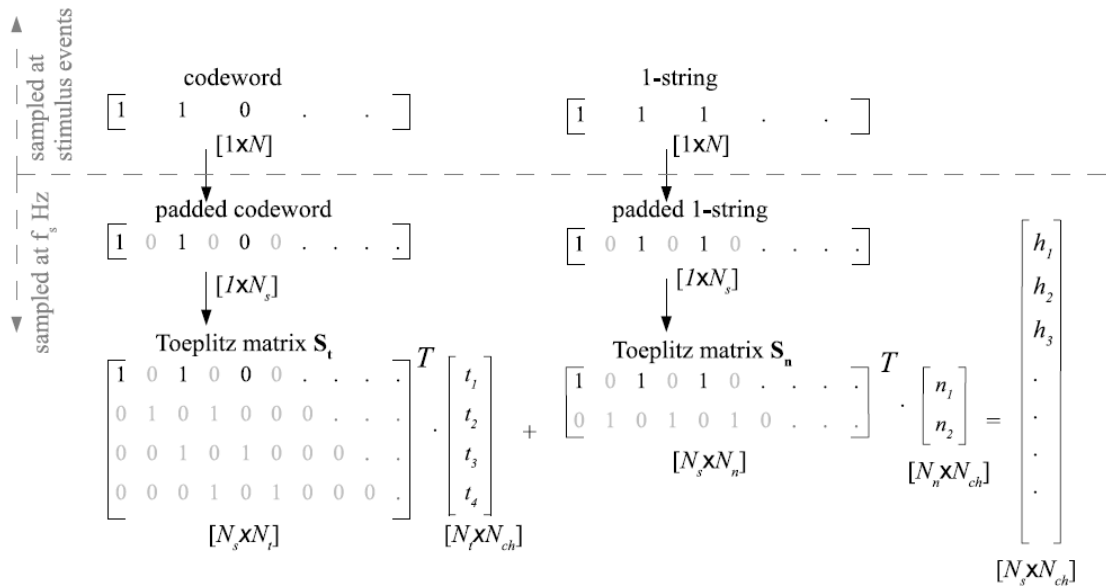


Figura 20: Generación de las matrices S_t y S_n

CAPITULO 4: MODELADO PROBABILÍSTICO DE LAS SEÑALES

Dado un ERP con ruido aditivo, es sabido, gracias al teorema central del límite (ver capítulo 2), que aplicar un procedimiento de promedio mejora la relación señal a ruido proporcionalmente al número de épocas empleadas para promediar. Por lo tanto, si e (ruido aditivo) está incorrelado respecto a h , entonces la media de las épocas EEG observadas b_j para $c_j = 1$ y para $c_j = 0$ en un conjunto de entrenamiento, es estimada para $t+n$ y n respectivamente. Dicho en otras palabras, podemos estimar t y n mediante la solución aportada por el algoritmo LS (*least-square*/mínimos-cuadrados), el cual nos permite minimizar el error $\varepsilon = E[b-h]^2$ (Anexo 10.2) en un conjunto de entrenamiento, donde h se obtiene mediante la expresión anterior, $h = S_t^T t + S_n^T n$. Para lograr obtener la solución del algoritmo LS para t y n hay que resolver el siguiente sistema de ecuaciones lineales:

$$\begin{bmatrix} S_t S_t^T & S_t S_n^T \\ S_n S_t^T & S_n S_n^T \end{bmatrix} \begin{bmatrix} t \\ n \end{bmatrix} = \begin{bmatrix} S_t b \\ S_n b \end{bmatrix}$$

Debemos tener en cuenta el hecho de que un desalineamiento entre la señal cerebral oculta construida (h_j) y la señal cerebral observada (b_j) puede suponer un impacto significativo a la hora de decodificar. Por ello, resulta de gran importancia tener el instante temporal en el que comienza cada evento estímulo lo más preciso que sea posible y llevar a cabo la construcción de la señal cerebral oculta con una frecuencia de muestreo lo suficientemente alta. Tomando como referencia el artículo [12] presente en la bibliografía usada, podemos determinar que una frecuencia de 250 Hz es adecuada, ya que se obtienen errores de redondeo lo suficientemente pequeños.

4.3 Estimación MAP

La estimación MAP nos permite decidir cuál es la palabra código \hat{c} más probable, dada una observación b . Sólo existe una única respuesta cerebral oculta para cada una de las diferentes palabras código pertenecientes al *codebook*. En este caso, podemos buscar de forma equivalente la respuesta cerebral oculta más probable \hat{h} dada una señal cerebral observada b de la siguiente manera:

$$\hat{h} = \underset{h}{\operatorname{argm\acute{a}x}} p(h|b)$$

Aplicando el teorema de Bayes, podemos establecer $p(h|b) = p(b|h)p(h)/p(b)$. A la hora de llevar a cabo la estimación MAP podemos ignorar el término $p(b)$ ya que este es independiente de h :

$$\hat{h} = \underset{h}{\operatorname{argm\acute{a}x}} [p(b|h)p(h)]$$

Debemos asumir que las señales cerebrales observadas son, en realidad, una versión ruidosa de las respuesta cerebrales ocultas, es decir, $b = h+e$. En este caso, la distribución $p(b|h)$ es exactamente la distribución del ruido $p(e)$. Consideremos ahora h como un vector en un espacio vectorial de dimensión L . Por otro lado, podemos considerar, razonablemente, que el ruido es blanco en la banda de frecuencia que va desde 0 a 15 Hz aproximadamente (se aplicará un filtro paso bajo para conseguirlo), que es Gaussiano y cuya media es cero. Bajo esta suposición tendremos que el ruido en cada canal, tiene una distribución esférica a través de los diferentes componentes del espacio vectorial antes mencionado. (La caracterización del ruido se explica en el próximo apartado).

Así, para el canal k , la distribución del ruido $p(e_k) = p(b_k|h_k)$ será

$$p(b_k|h_k) = \frac{1}{(2\pi\sigma_k^2)^{L/2}} \exp\left(-\frac{\|b_k-h_k\|^2}{2\sigma_k^2}\right)$$

Asumamos ahora que decorrelamos el ruido en los canales mediante una transformación lineal W . El ruido decorrelado e_w resultante es, de hecho, independiente en los canales y con varianza σ_w^2 , con lo que podemos escribir la distribución $p(e_w) = p(b_w|h_w)$ como el siguiente producto:

$$p(b^w|h^w) = \prod_k \frac{1}{(2\pi\sigma_w^2)^{L/2}} \exp\left(-\frac{\|b_k^w-h_k^w\|^2}{2\sigma_w^2}\right)$$

Resulta que el criterio MAP anteriormente descrito es equivalente al siguiente criterio MAP

$$\hat{h} = \underset{h}{\operatorname{argm\acute{a}x}} [p(b^w|h^w)p(h)] \quad (\text{ya que } p(b^w|h^w) = p(b|h) + \text{cte})$$

CAPITULO 4: MODELADO PROBABILÍSTICO DE LAS SEÑALES

Usando la propiedad del logaritmo de ser monótonico, podemos realizar la decisión MAP sobre la probabilidad logarítmica (convertimos el productorio en sumatorio)

$$\hat{h} = \operatorname{argmáx}_h [\ln [p(b^w|h^w)p(h)]]$$

$$\hat{h} = \operatorname{argmáx}_h \sum_k \left[\ln q - \frac{\|b_k^w - h_k^w\|^2}{2\sigma_w^2} \right] + \ln p(h)$$

donde se define $q = (2\pi\sigma_w^2)^{-L/2}$.

Podemos establecer σ_w^2 a 1 e ignorar la variable q ya que este término es independiente de h_w . Resultando la decisión MAP como sigue:

$$\hat{h} = \operatorname{argmín} [\sum_k [\|b_k^w - h_k^w\|^2] - 2 \ln p(h)]$$

Si las letras se extraen de forma uniforme o equiprobable del conjunto de letras que componen la matriz de letras, entonces todas las letras resultarán equiprobables y la estimación MAP se convertirá en una estimación ML (*maximum likelihood*/máxima probabilidad). En ese caso la decisión se llevará a cabo de la siguiente manera:

$$\hat{h} = \operatorname{argmáx}_h p(b|h) = \operatorname{argmín}_h [\sum_k [\|b_k^w - h_k^w\|^2]]$$

La regla de decodificación correspondiente a la expresión anterior nos indica que seleccionamos la respuesta cerebral para la cual la suma de los valores al cuadrado de la diferencia entre la señal cerebral blanqueada observada (b_k^w) y la señal cerebral oculta transformada (h_k^w) sobre todos los componentes k, es mínima. Conseguimos por lo tanto una estimación MAP en virtud de una serie de suposiciones acerca del ruido y en virtud de letras equiprobables.

4.4 Caracterización del ruido

Como se indicó anteriormente, para realizar la decisión MAP necesitamos, previamente, asumir una serie de hechos acerca de la señal y del ruido.

La señal h_k correspondiente al canal k , tiene una característica paso bajo cuya frecuencia de corte resulta ser f_0 y permite una representación sobre una base ortonormal de señales unitarias, $\{\phi_l(t)\}$ con $l = 1, 2, \dots, L$. Este hecho nos permite referirnos indistintamente a h_k como una señal o como un vector en un espacio vectorial de Hilbert L -dimensional.

Por otro lado, el ruido e_k tiene media cero y sigue una distribución Gaussiana. También debemos asumir que el ruido es blanco en el margen de frecuencias paso bajo de la señal, con una densidad espectral de potencia constante e igual a $\eta_k/2$.

Para todo ello necesitamos caracterizar estadísticamente el ruido en el espacio vectorial de la señal definido anteriormente. Esto se puede conseguir mediante la aplicación de la teoría de sistemas lineales [13]. Tomemos $\Phi_l(f)$ como cada uno de los elementos de la base de la señal, expresado en el dominio de la frecuencia. En este caso, la varianza del ruido σ_{kl}^2 en cada dimensión del espacio vectorial de la señal, viene dado por la proyección del ruido en cada elemento de la base:

$$\int_{-f_0}^{f_0} \frac{\eta_k}{2} |\phi(f)|^2 df = \eta_k \cdot f_0$$

Esta ecuación demuestra que la varianza del ruido es idéntica para cada componente de la base. Además, una propiedad de los filtros lineales es, que una entrada Gaussiana supone una salida también Gaussiana. Por lo tanto, para la hipótesis establecida acerca del ruido, la distribución en el espacio vectorial se define mediante la Gaussiana esférica, cuya matriz de covarianzas viene dada por $C = \sigma_k^2 I$. Finalmente, podemos expresar la distribución para cada canal como

$$e_k \sim \frac{1}{(2\pi\sigma_k^2)^{L/2}} \exp\left(-\frac{\|b_k - h_k\|^2}{2\sigma_k^2}\right)$$

4.5 Estimación MAP de señales decorreladas

Asumamos que el ruido en diferentes canales puede ser decorrelado por medio de una transformación lineal W , en el sentido en que el ruido está ahora sin correlación (ver capítulo 2) a través de los canales y tiene la misma varianza $\sigma_{W,k}^2$. La matriz de transformación W puede ser calculada sobre una parte de la señal, en la cual no se haya presentado ningún estímulo al sujeto, o sobre la señal completa, si el ruido es mucho mayor que la señal. Debido a que decorrelación supone independencia estadística para las variables aleatorias Gaussianas, podemos describir la distribución conjunta del ruido para todos los canales $e^2 \sim p(h^W + e^W | h^W) = p(b^W | h^W)$ como

$$\begin{aligned} p(b^W | h^W) &= \prod_k \frac{1}{(2\pi\sigma_{w,k}^2)^{L/2}} \exp\left(-\frac{\|b_k^w - h_k^w\|^2}{2\sigma_{w,k}^2}\right) \\ &= \prod_k \frac{1}{(2\pi\sigma_w^2)^{L/2}} \exp\left(-\frac{\|b_k^w - h_k^w\|^2}{2\sigma_w^2}\right) \end{aligned}$$

Para ello, se hace uso de la propiedad de transformación lineal sin pérdidas de las distribuciones de probabilidad: $p(AT x) = (\det A)^{-1} p(x)$. Luego, sin más que aplicar el criterio MAP podemos expresarlo como sigue:

$$\begin{aligned} \hat{h} &= \operatorname{argm\acute{a}x} [p(W^{-1}b^w | W^{-1}h^w)p(h)] \\ &= \operatorname{argm\acute{a}x} [(\det W)p(b^w | h^w)p(h)] \\ &= \operatorname{argm\acute{a}x} [p(b^w | h^w)p(h)] \end{aligned}$$

Este hecho implica que realizar una decisión tipo MAP en señales decorreladas, resulta equivalente a realizarla sobre las señales originales.

4.6 Resumen del modelo

MODELO GENERATIVO

Input: Conjunto de observaciones de entrenamiento b_{train} y palabras código elegidas.

Filtrado paso banda de b_{train} almacenado como b

IF promediado

 División de b_{train} en épocas b_j

 Promediado de todas las b_j para cada $c_j = 0$ almacenado como n

 Promediado de todas las b_j para cada $c_j = 1$, restar n y almacenar como t

ELSEIF algoritmo LS

 Construir matrices S_t y S_n

 Construir plantillas objetivo y no objetivo t y n por medio de

$$\begin{pmatrix} S_t S_t^T & S_t S_n^T \\ S_n S_t^T & S_n S_n^T \end{pmatrix} \begin{pmatrix} t \\ n \end{pmatrix} = \begin{pmatrix} S_t b \\ S_n b \end{pmatrix}$$

END

DECODIFICACIÓN

Input: Observación de test b_{test}

Filtrado paso banda de b_{test} almacenado como b

Blanqueado de b mediante matriz blanqueadora W , guardado como b^W

FOR cada palabra código c del *codebook*

 Construir matrices S_t y S_n

 Construir respuesta oculta h usando t y n mediante $h = S_t T_t + S_n T_n$

 Blanquear h por medio de W , guardado como h^W

 FOR cada canal k

 Calculo de distancia $\|b_k^W - h_k^W\|$

 END

END

CAPITULO 4: MODELADO PROBABILÍSTICO DE LAS SEÑALES

Aplicar regla de decodificación

$$\hat{h} = \underset{h}{\operatorname{argmín}} \left[\sum_k [\|b_k^w - h_k^w\|^2] - 2 \ln p(h) \right]$$

Predecir la letra correspondiente a \hat{h}

Capítulo 5

Experimentos y resultados

5.1 Introducción

En este capítulo se quiere mostrar los diferentes experimentos llevados a cabo y los resultados obtenidos en los mismos. Para ello, primero se explicarán las características de los datos y los medios empleados para realizar los experimentos. Posteriormente se mostrarán los resultados parciales, tales como la evolución de la señal, la matriz blanqueadora, etc., después, los resultados obtenidos en términos de probabilidad de acierto, tras aplicar el modelo generativo desarrollado para, finalmente, presentar las mejoras y nuevas aportaciones realizadas.

5.2 Datos y medios empleados

Una vez explicado el algoritmo empleado y sus características conviene analizar los datos que se han utilizado y los medios de los que se ha dispuesto.

CAPITULO 5: EXPERIMENTOS Y RESULTADOS

- Los datos utilizados corresponden a 64 experimentos realizados a 6 sujetos diferentes, sin tener en cuenta sexo ni edad, con el sistema *Visual P300 Speller*, donde se codificó la información empleando 72 bits por palabra código.
- Se emplearon dos maneras de resaltar los caracteres de la matriz de caracteres. El sistema Flash, con el cual se iluminan unos caracteres, de una columna o fila, en un determinado momento, o el sistema Flip, donde también se rotan dichos caracteres. Se realizó por lo tanto, el mismo experimento para estos dos sistemas para poder compararlos hasta que punto influye la forma en la que se presentan los eventos estímulo al sujeto.
- *Visual Speller* y *RC code*: La frecuencia de iluminación es aproximadamente de 6 filas o columnas por segundo, iluminando todos los caracteres de manera aleatoria, dos veces por cada 12 flashes (o flips). El *RC code* realizará 12 iluminaciones por letra, lo que en tiempo equivaldrá aproximadamente a 12 segundos.
- Se realizará el mismo experimento para diferentes longitudes de palabra código, con el fin de demostrar que el hecho de incluir redundancia supone una mejora en las prestaciones en términos de probabilidad de acierto.
- Las diferentes longitudes de palabra código empleadas son {12, 24, 36, 48, 60, 72}. Puede observarse que las longitudes empleadas son múltiplos de 12 ya que como se ha indicado anteriormente, cada 12 iluminaciones se consigue resaltar todos los caracteres.
- Lenguaje de programación: Matlab 7.4.0(R2007a)
- Se intentó implementar el código en Java, pero se desechó la idea por lo complicado, comparado con Matlab, del manejo y operaciones con matrices en Java.
- Procesador utilizado: Intel® Core™ i7 CPU Q720 @1.60GHz 1.60GHz
- Memoria RAM disponible: 4 GB
- Para la segunda de las mejoras el procesador y memoria RAM empleados fueron insuficientes, por lo que se solicitó autorización y se empleó los ordenadores del departamento de Teoría de la Señal y Comunicaciones de la universidad Carlos III de Madrid, estableciendo una conexión VPN entre casa y la universidad.
- Se realizaron los experimentos mediante un LOO (*leave one out*).
- El decisor utilizado es tipo MAP.
- La Figura 21 muestra el esquema a seguir para obtener la letra estimada a partir de un experimento.

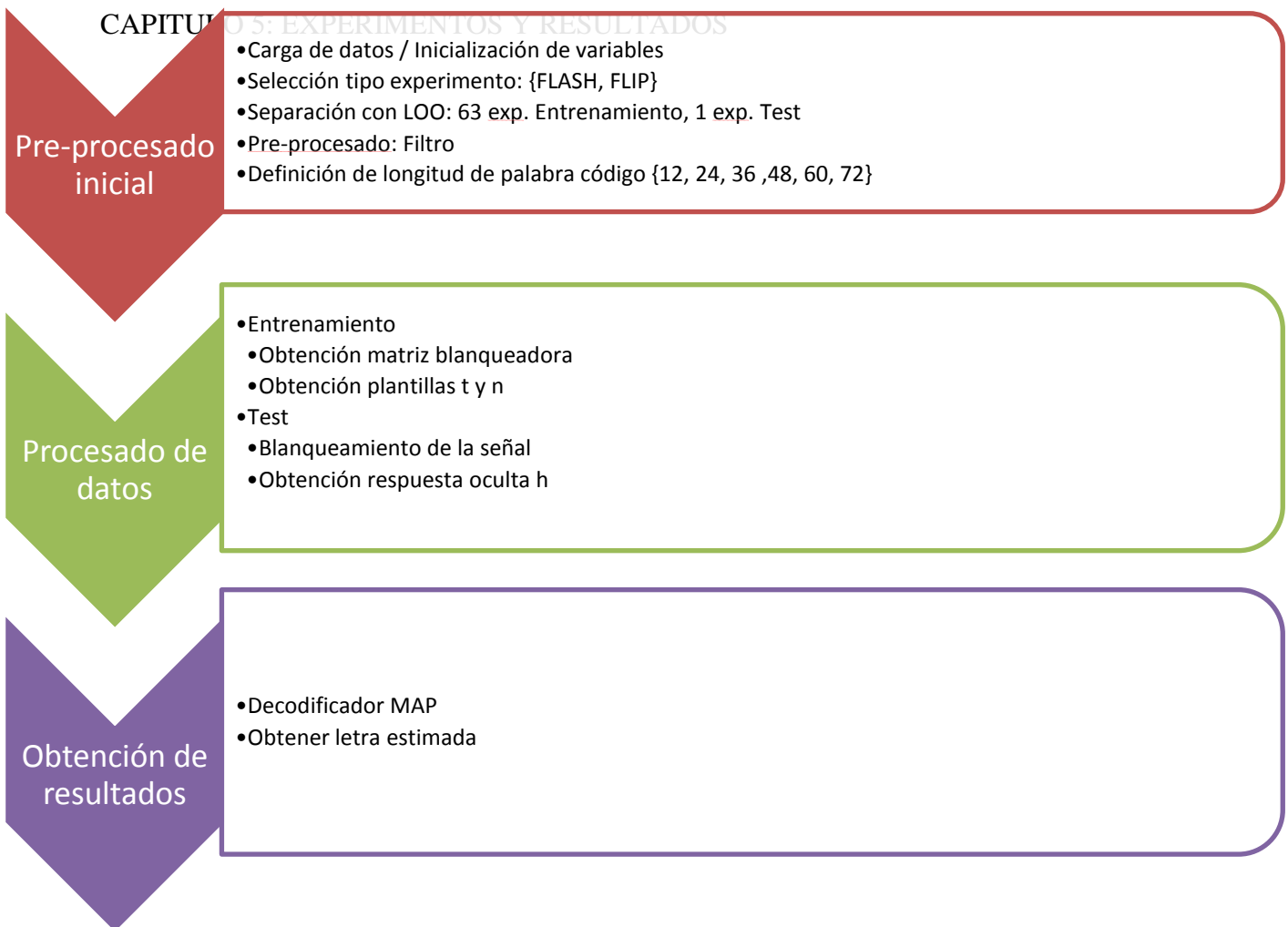


Figura 21: Esquema general para la obtención de la letra estimada

De la Figura 21 se observa que se puede dividir el proceso en tres bloques fundamentales:

1. Pre-procesado inicial: Encargado de obtener y adecuar los datos disponibles para su procesado.
2. Procesado de los datos mediante el modelo generativo desarrollado.
3. Decodificación de letra.

Por último recordar que el sistema se prueba mediante *cross-validation* (LOO) con lo que la decodificación se realizará sólo para un experimento cada vez, es decir, para cada sujeto se tendrá 64 estimaciones de letra distintas, una por cada experimento. Esto implica realizar 64 *cross-validation* por cada longitud de palabra código y sujeto.

5.3 Obtención del conjunto de datos

Se empleó una configuración estándar del visual-speller con una matriz de caracteres alfanuméricos de 6x6, presentada por medio de un monitor LCD en un despacho con el máximo silencio posible. Cada letra tiene un estado de baja o alta intensidad. La alta intensidad se mantiene durante 0.033 sg, Esta configuración corresponde al modelo de estímulo Flash. Por otro lado, también se realizó una configuración para el modelo de estímulo Flip. La frecuencia de refresco del monitor es de 60 Hz mientras que el valor de SOA es de 0.167 sg.

Las señales EEG se midieron mediante el empleo de un sistema *QuickAmp (Brain Products GbmH)* en combinación con un casco de electrodos (ver Figura 63) (*Electro-Cal International, Inc.*). El equipo se configuró para medir 58 canales de EGG, uno horizontal EOG en el ojo izquierdo, uno bipolar vertical de señal EOG, y un canal de señal de sincronización de un sensor de luz unido a la pantalla, todo muestreado a 250 Hz. Todas las impedancias de los electrodos se encuentran por debajo de los 10 k Ω .

Seis sujetos, indicados como {I, II, III, IV, V, VI}, 5 de ellos masculinos y uno femenino, se incluyeron en el estudio, estando cada uno de ellos dentro de la franja de edad de los 20 a los 40 años. Los sujetos tenían poco o ninguna experiencia con el *visual-speller* estándar (de 0 a 3 sesiones previas). Todos los sujetos participaron en ambos experimentos, Flip y Flash. Las sesiones se dividieron en 16 bloques de copia-deletreo, cada bloque conteniendo 20 pruebas (letras a deletrear) tanto de estímulos Flash como Flip. En el modo copia-deletreo, el sujeto tuvo que fijarse en letras decididas de antemano por lo que las letras objetivo eran conocidas en todo momento. Durante esta parte, los sujetos no recibieron ninguna información acerca del rendimiento en la predicción de letras. Se usaron diferentes métodos de codificación en las letras. La distribución de los valores de TTI y de TNI fue la misma tanto para los estímulos Flash como para los Flip. En total, cada sujeto deletreó 64 letras para cada uno de los dos tipos de estímulo. Se considera sub-prueba a cada parte de una prueba durante la cual cada columna o fila de la matriz de letras se iluminará o rotará una vez. Estas sub-pruebas tienen una duración de $12 * SOA = 2.0$ sg. Cada prueba está compuesta por 6 sub-pruebas, por lo que se requieren 12 sg para completar una letra. Los tipos de estímulo se alternaron entre bloques para prevenir lo máximo posible la influencia debida a la fatiga tanto en los estímulos tipo Flash como Flip. Cada prueba comenzó con un recuadro rojo sobre una de las letras (elegido aleatoriamente), el cual indicaba al sujeto en que letra debía focalizar su atención. Esta pista se mantuvo aproximadamente durante un segundo y se suprimió un segundo antes del comienzo de la secuencia de estímulos. Se pidió a los pacientes que llevaran la cuenta de los eventos estímulo sobre el objetivo elegido y que minimizaran el parpadeo, el movimiento y el tragar durante la secuencia

La base de datos empleada se obtuvo del *Max Plank Institut for Biological Cybernetics*, *Departamento Empirical Inference*, *BCI group* liderado por Jeremy Hill.

5.4 Indicadores de prestaciones

En todo sistema se debe establecer los indicadores que midan las prestaciones de manera objetiva.

- Acierto de letra: Se define como la probabilidad de acierto de los caracteres

$$P_{character}(\%) = \frac{\text{Número letras acertados}}{\text{Número letras totales}} * 100$$

- Carga computacional:
 - Tiempo de procesado de datos: Tiempo necesario en llevar a cabo la estimación de una letra.
 - Necesidades de memoria RAM: Requisitos de memoria debido al manejo de matrices de elevadas dimensiones.

5.5 Resultados parciales

Antes de entrar en detalle con los resultados que se han obtenido tras aplicar el modelo generativo implementado, conviene mostrar la evolución de las señales cerebrales a lo largo del proceso, en primer lugar para conseguir una mejor comprensión del funcionamiento del mismo y en segundo lugar para demostrar que efectivamente, los resultados que se obtienen son coherentes con lo descrito anteriormente.

Para ello conviene seguir la evolución de la señal según muestra la Figura 22

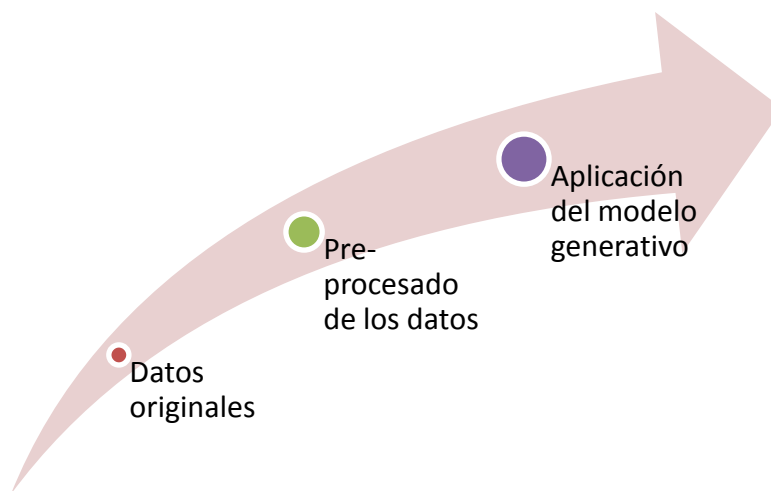


Figura 22: Evolución de los datos

Los resultados parciales mostrados a continuación han sido obtenidos escogiendo de forma arbitraria al sujeto, la longitud de palabra código, el experimento empleado como test y el tipo de resalte de letras, ya que este hecho no influye sobremanera en dichos resultados. Así, los datos provienen del sujeto 1, tomando los 72 eventos (longitud de palabra código), resalte tipo Flash y empleando como experimento de test el quinto a la hora de realizar la *cross-validation*.

CAPITULO 5: EXPERIMENTOS Y RESULTADOS

En primer lugar se muestra la señal cerebral original, en este caso la relativa al canal 37 de los 58 posibles. La Figura 23 muestra la señal antes y después de ser filtrada.

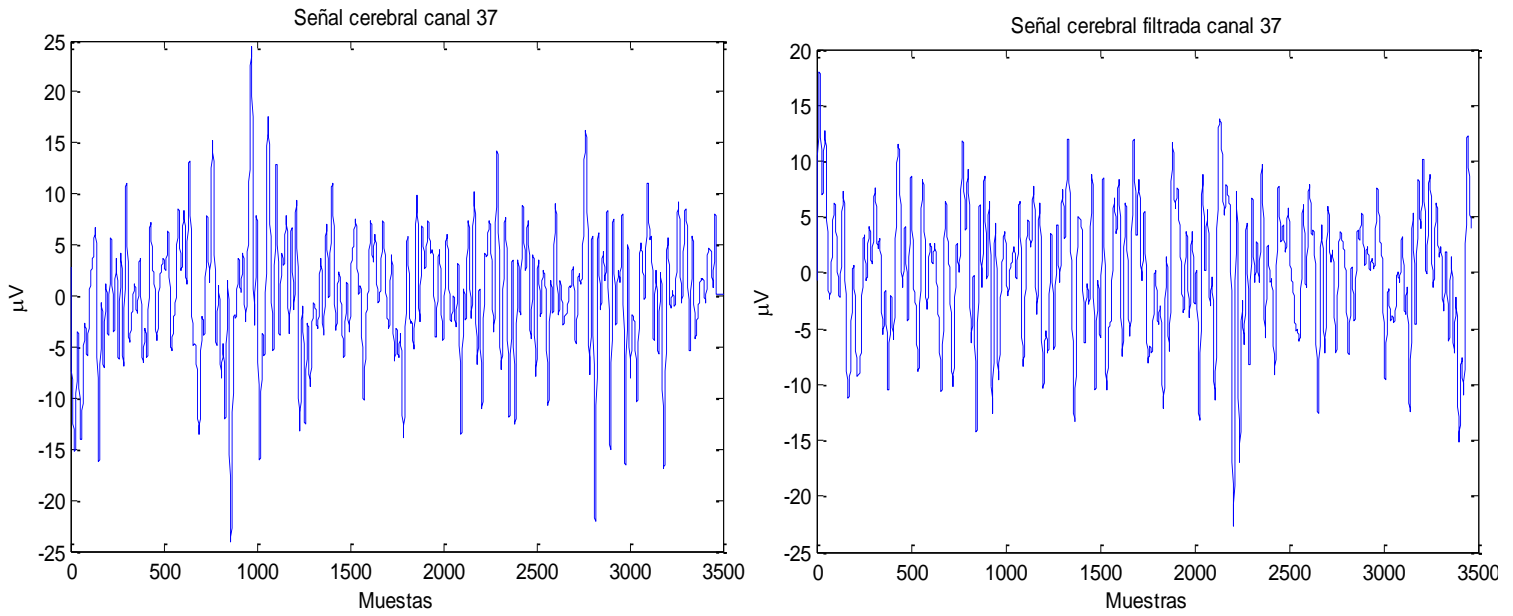


Figura 23: Señal cerebral antes y después del filtro paso banda

CAPITULO 5: EXPERIMENTOS Y RESULTADOS

Como se comentó, las señales EEG se caracterizan por tener una pobre relación señal a ruido, por eso aplicar un filtrado resulta tan importante a la hora de eliminar todo el ruido existente fuera de la banda de interés.

Una vez obtenidos los datos el siguiente paso es obtener las plantillas t (eventos objetivo) y n (eventos no objetivo) con las que posteriormente construiremos nuestra representación de la señal cerebral. Así, la Figura 24 muestra la plantilla de eventos no objetivos para todos los canales, mientras que la Figura 25 hace lo correspondiente para los eventos objetivo.

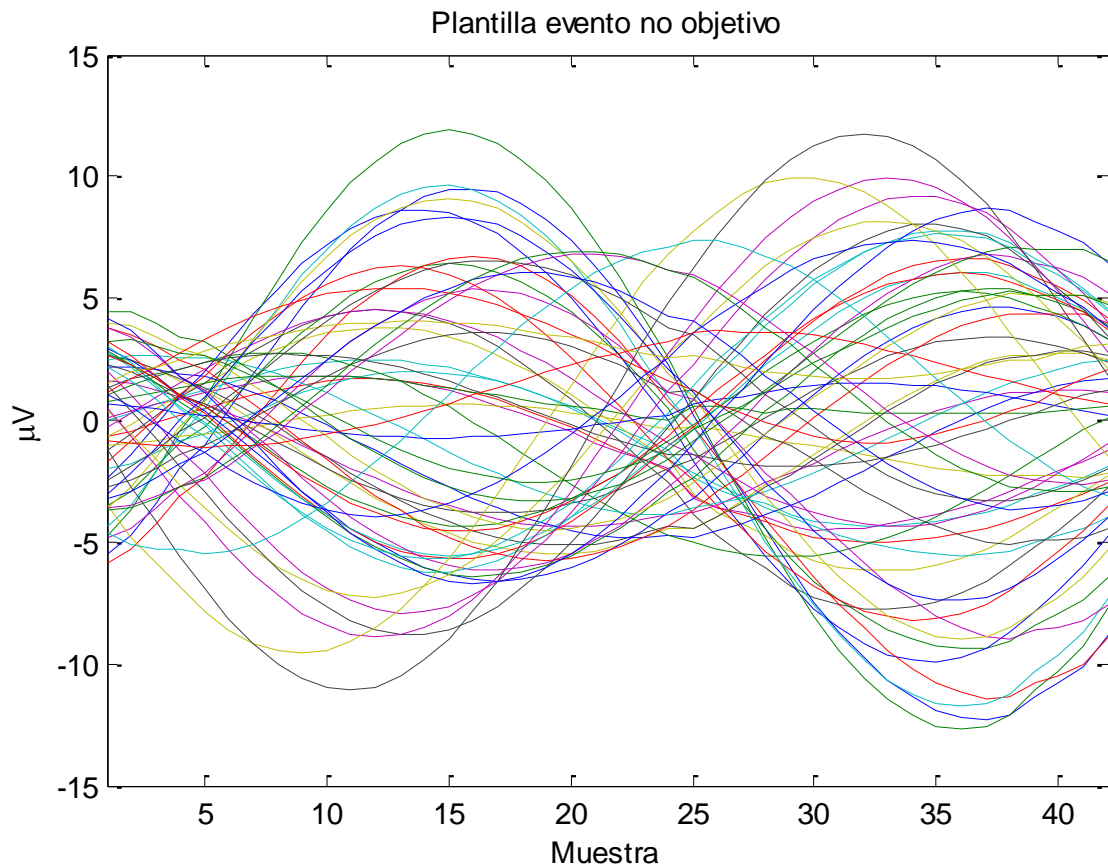


Figura 24: Plantillas de evento no objetivo

Puede observarse como la señal cerebral en estado de relajación, es decir, al no presentarse un evento objetivo, es similar en forma a una señal tipo senoidal.

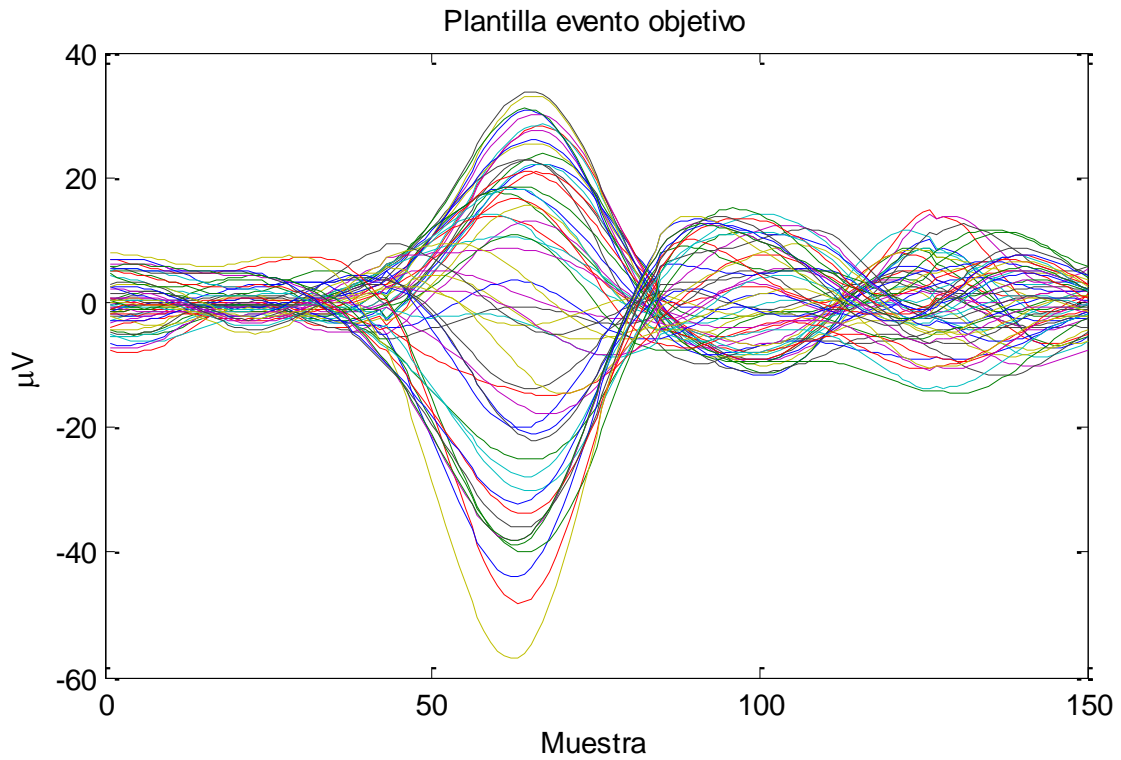


Figura 25: Plantillas para eventos objetivo

Por otro lado, puede observarse que al presentarse un evento objetivo, lo que nos encontramos es una forma de onda que muestra un pico. Este pico indica que la respuesta cerebral es de tipo P300, que aparece al producirse un evento objetivo.

El otro elemento clave que se obtiene en la fase de entrenamiento es la matriz de blanqueamiento ya que permitirá decorrelar las señales cerebrales. Para entender visualmente lo que esto significa, hay que fijarse en la Figura 26.

CAPITULO 5: EXPERIMENTOS Y RESULTADOS

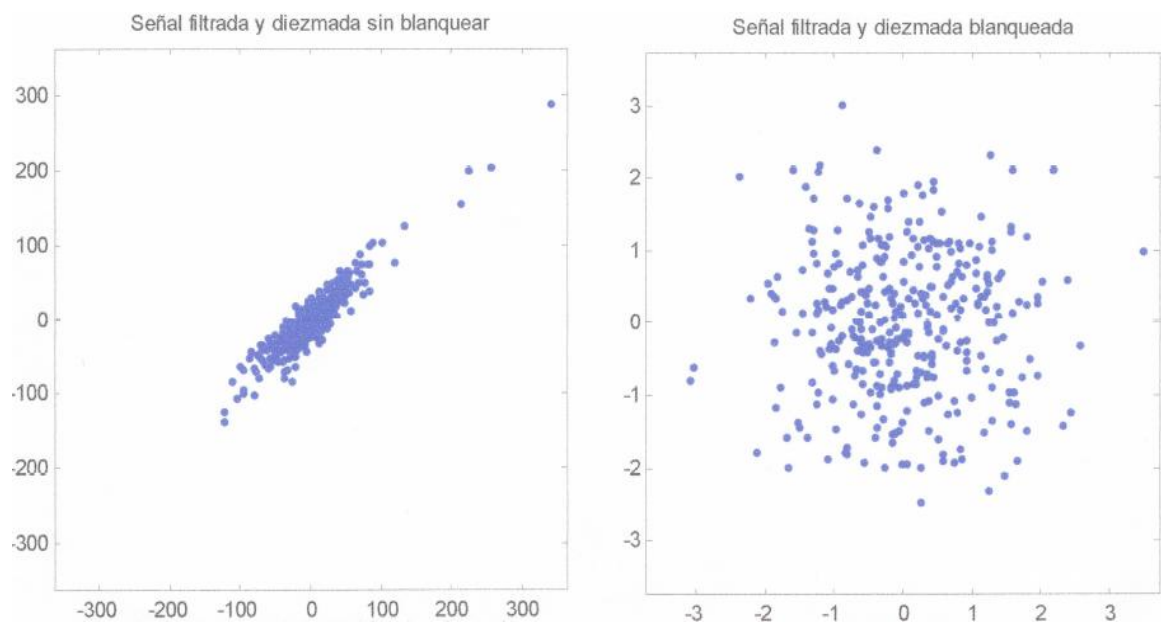


Figura 26: Blanqueamiento de la señal (entrenamiento)

Esta figura muestra como la señal fuertemente correlada al ser representada consigo misma mantiene un modelo lineal (relación fuerte, ver Capítulo 2) en el espacio debido a la correlación entre las señales de los diferentes canales, mientras que al blanquear dichas señales lo que conseguimos es una forma esférica (relación débil), con dispersión, que indica que las señales de los diferentes canales han perdido dicha correlación.

Otra forma de ver el blanqueamiento de la señal es mediante la representación de su matriz de covarianzas. Así, la Figura 27 (con el número de canal en los ejes), muestra las matrices de covarianzas de la señal, antes y después de haber sido blanqueada.

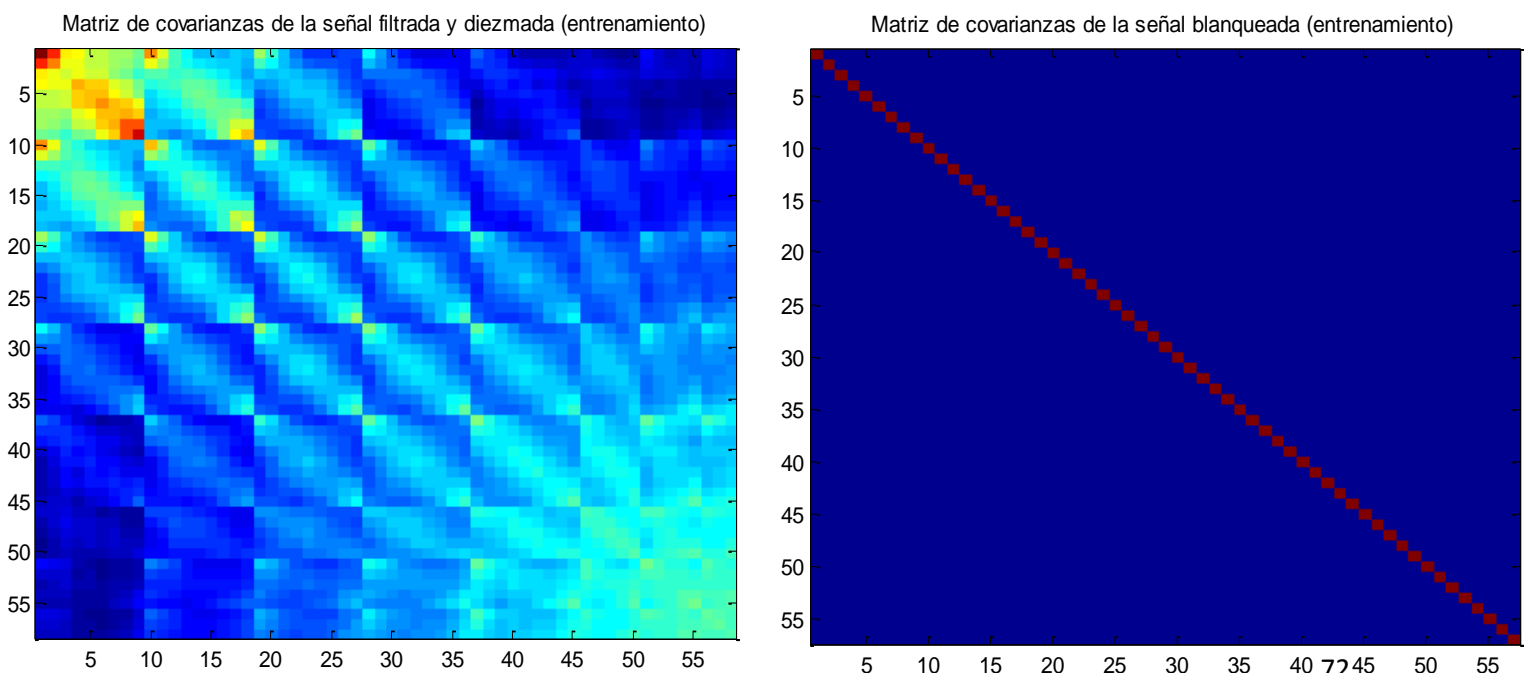


Figura 27: Matriz de covarianza de la señal antes y después de ser blanqueada (entrenamiento)

CAPITULO 5: EXPERIMENTOS Y RESULTADOS

Como puede observarse, antes de blanquear la señal, esta presenta una correlación entre los diferentes canales, hecho que queda reflejado ya que cuanto más tienda un color hacia el rojo, mayor será la correlación entre ambos canales. También se observa que canales (electrodos) cercanos entre sí, muestran mayor correlación que canales más alejados entre sí. Por el contrario, si observamos la matriz de covarianzas de la señal blanqueada, se observa que no existe correlación entre los distintos canales, ya que el azul oscuro que se observa representa covarianza 0 entre dos canales.

Una vez terminada la fase de entrenamiento pasamos a la fase de test. Aquí, será interesante ver el comportamiento de la señal tras pasar por el filtro blanqueador ya que ahora, al no haber sido diseñado especialmente para estos datos, deberíamos observar un blanqueamiento parcial, en el sentido en que este no será del todo perfecto como lo era con los datos de entrenamiento.

Así, observando la Figura 28, podemos ver la señal de test antes y después de ser blanqueada.

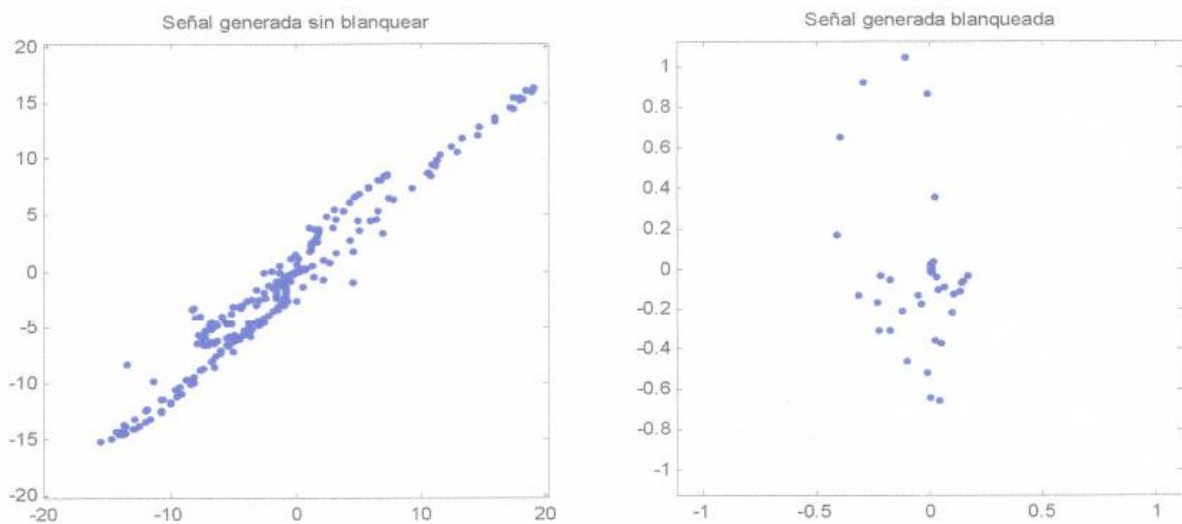


Figura 28: Blanqueamiento de la señal (test)

En este ocasión podemos observar como la señal sin blanquear presenta la misma forma lineal debido a la fuerte correlación entre canales, mientras que una vez ha sido blanqueada la forma es relativamente esférica aunque sigue presentado una baja correlación entre algunos canales.

Si observamos ahora sus matrices de covarianzas, Figura 29, podemos comprobar que en este caso la decorrelación no es perfecta, pero si lo suficientemente buena como para considerar que los canales están casi decorrelados entre sí. Como se indicó anteriormente, esto era de esperar ya que la matriz blanqueadora se definió en el proceso de entrenamiento.

CAPITULO 5: EXPERIMENTOS Y RESULTADOS

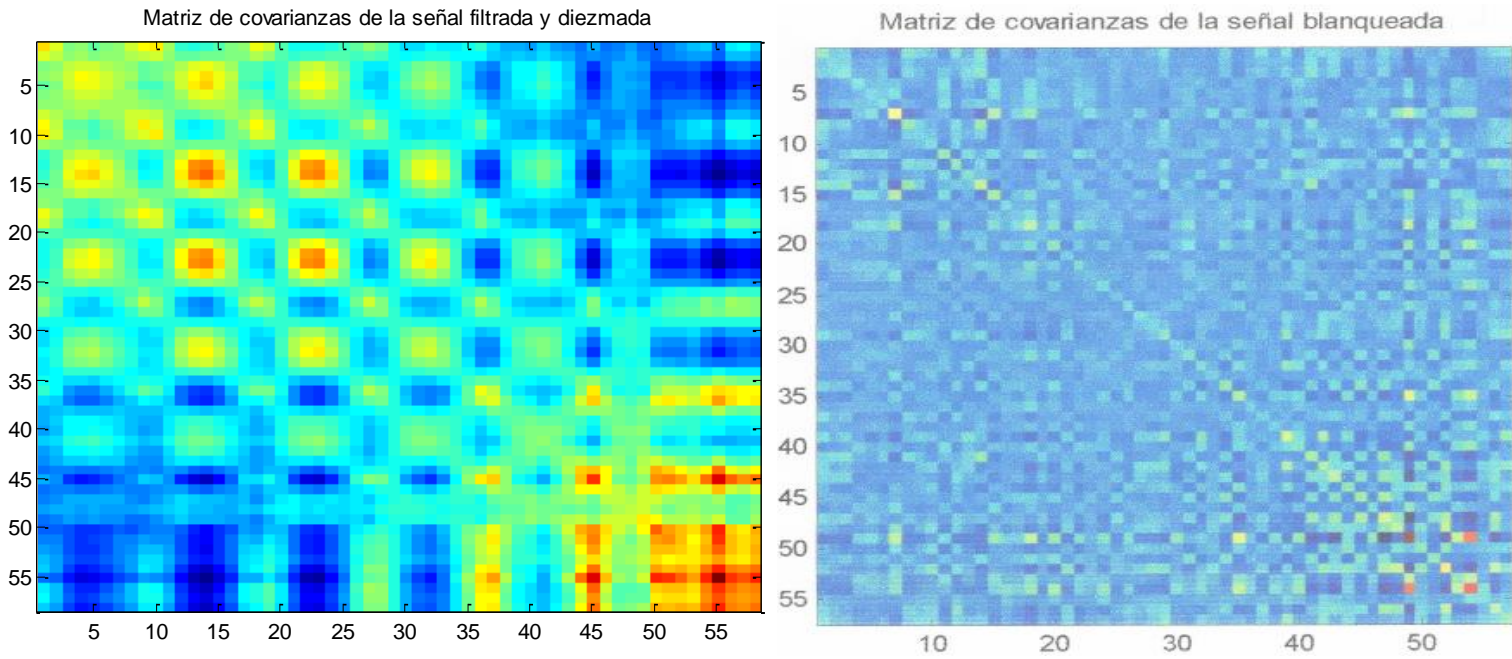


Figura 29: Matriz de covarianzas de la señal antes y después de ser blanqueada (test)

Finalmente, el último resultado parcial de la fase de test será comprobar la forma de onda que obtenemos al aplicar las plantillas obtenidas en la fase de entrenamiento cuando queremos estimar una letra.

Así, la Figura 30 muestra dicha forma de onda una vez filtrada.

CAPITULO 5: EXPERIMENTOS Y RESULTADOS

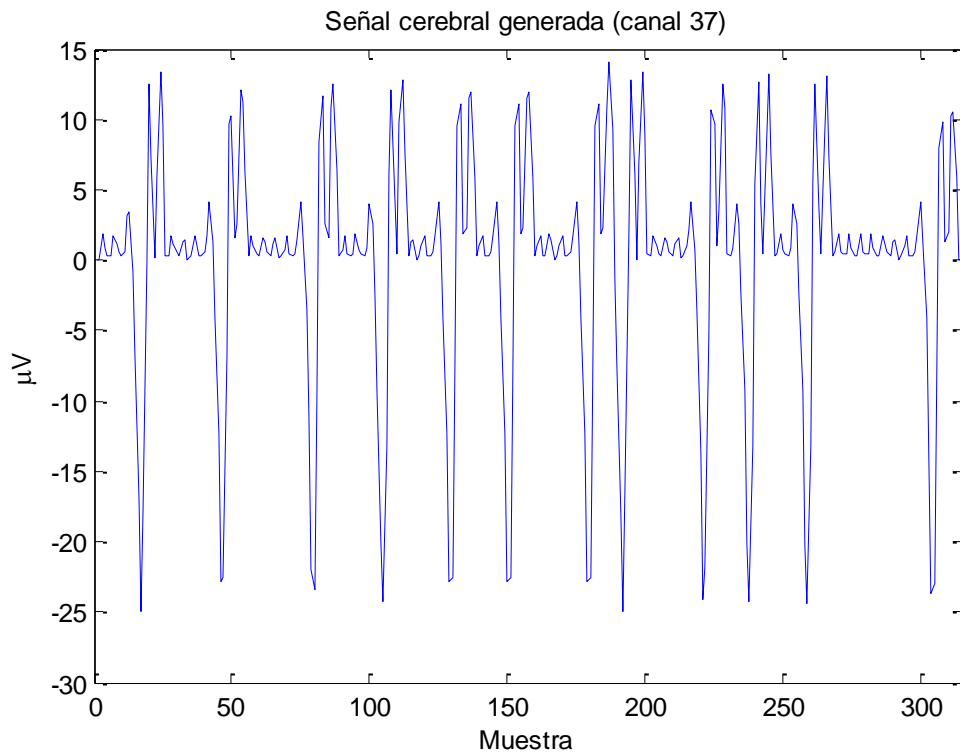


Figura 30: Señal cerebral generada

Puede observarse como resulta evidente los momentos en los que se produce un evento objetivo y se representa el potencial evocado P300.

5.6 Resultados

5.6.1 Resultados de los experimentos Flash

Como se indicó anteriormente, los experimentos tipo Flash son aquellos en los que el resalte de las letras se produce únicamente mediante la iluminación de una fila o columna por evento.

Los resultados se presentarán mediante gráficos en forma de diagrama de barras que muestran la probabilidad de acierto por sujeto y en función de la longitud de palabra código empleada. Así, las figuras 31 – 36, representan dichas probabilidades de acierto.

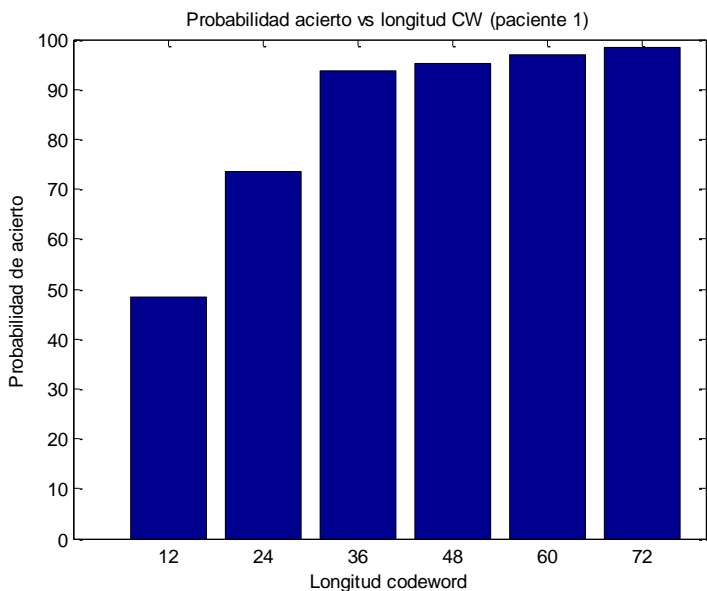


Figura 31: Probabilidad acierto 1 (Flash)

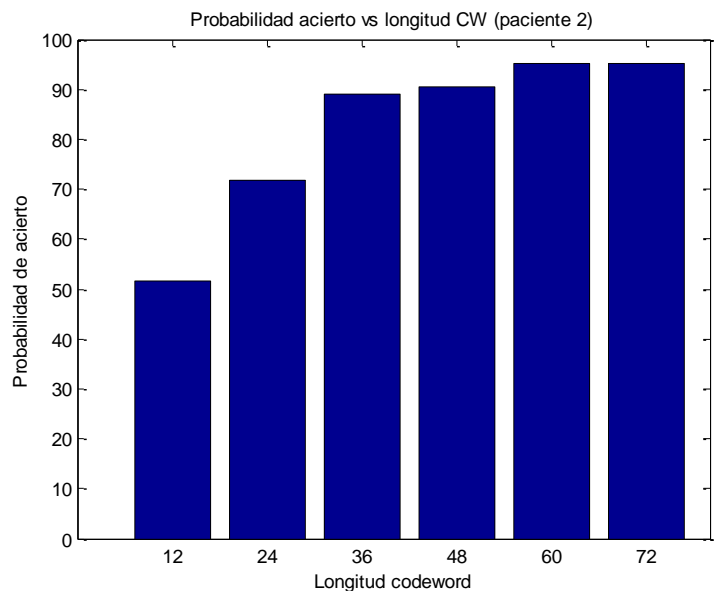


Figura 32: Probabilidad acierto 2 (Flash)

CAPITULO 5: EXPERIMENTOS Y RESULTADOS

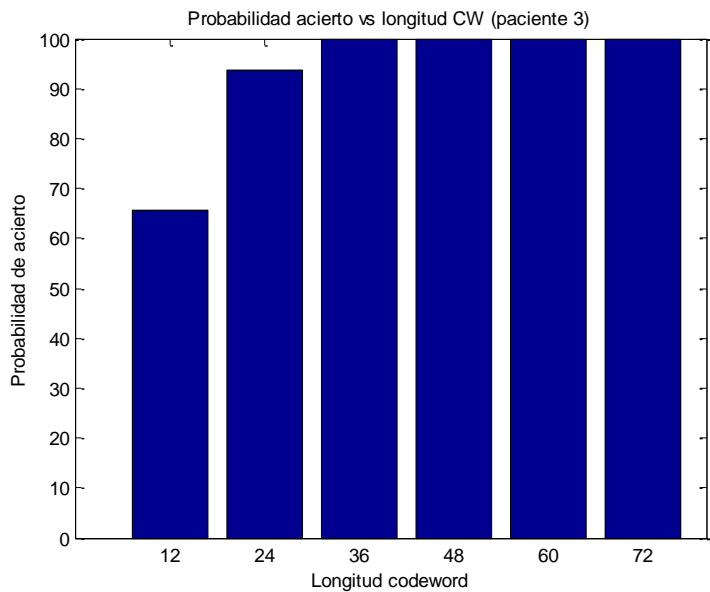


Figura 33: Probabilidad acierto 3 (Flash)

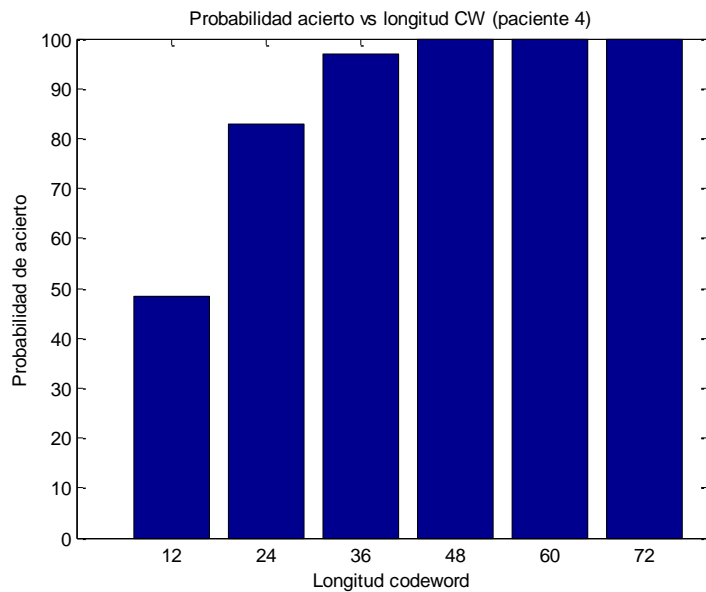


Figura 34: Probabilidad acierto 4 (Flash)

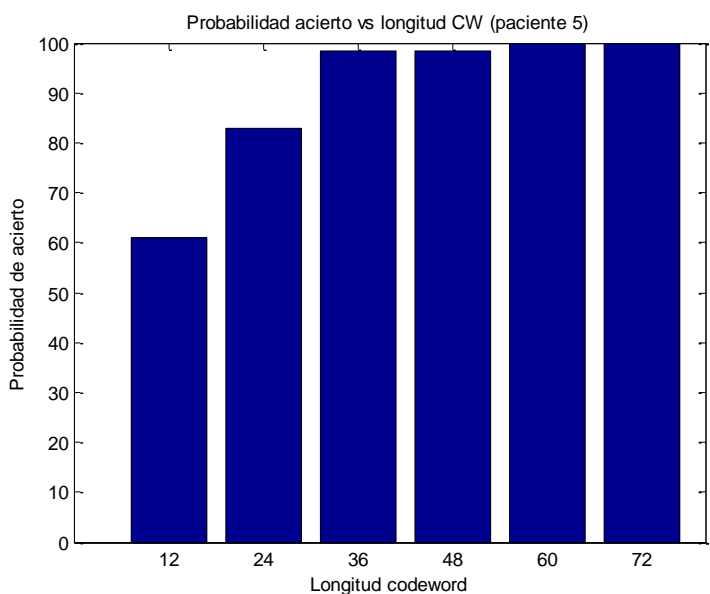


Figura 35: Probabilidad acierto 5 (Flash)

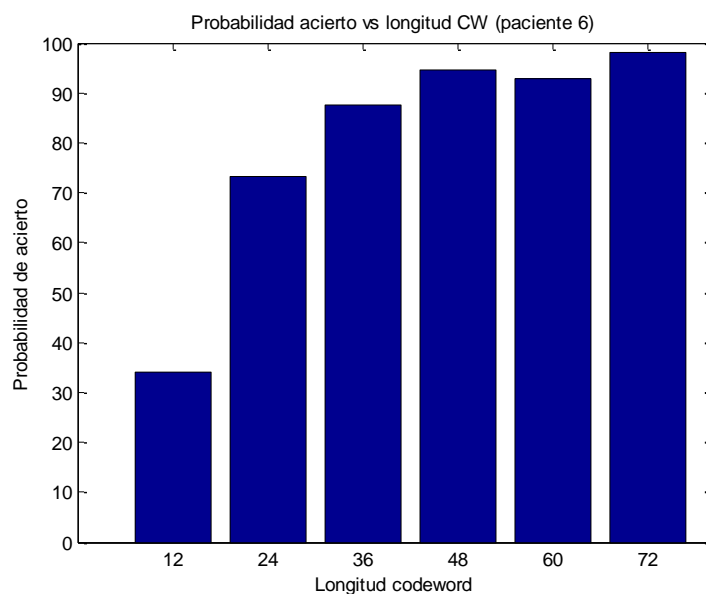


Figura 36: Probabilidad acierto 5 (Flash)

A continuación se muestra la figura más importante de todas, Figura 37, que representa la probabilidad promedio de acierto en los experimentos tipo Flash. También representa la desviación típica de los resultados obtenidos ya que será un reflejo de lo bien o mal que trabaja nuestro sistema.

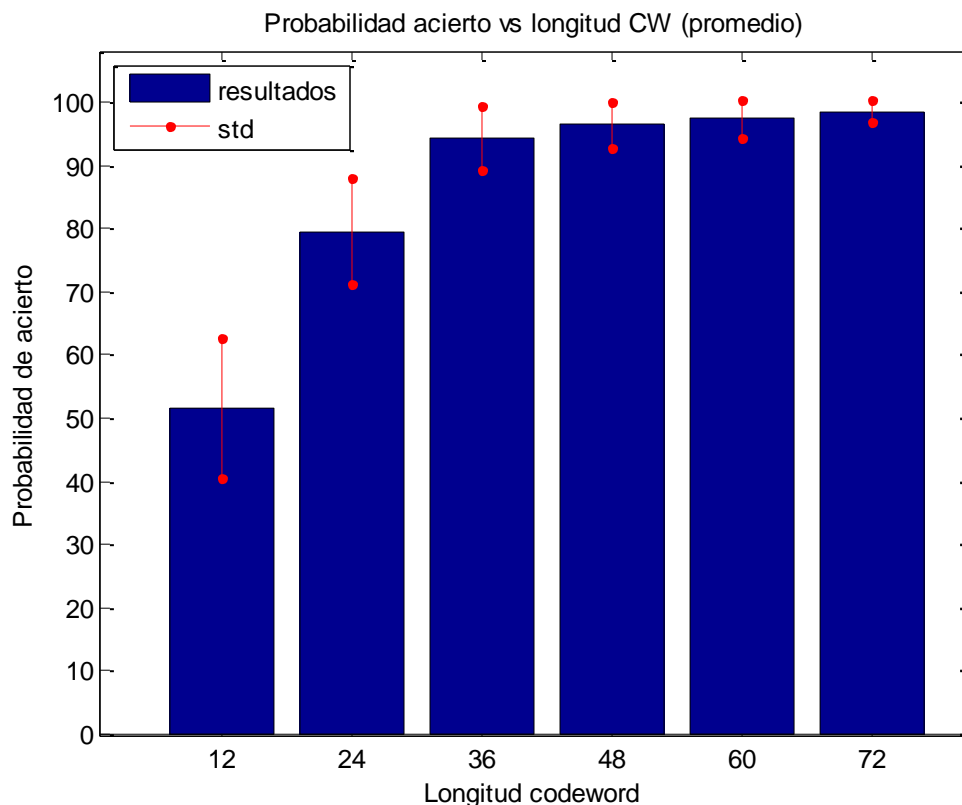


Figura 37: Probabilidad acierto promedio (Flash)

Es importante recordar que los valores de longitud de palabra código vienen determinados por la longitud mínima de palabra código, 12. El resto de valores son debidos a la repetición de un patrón de 12 bits para analizar si se mejora la probabilidad de acierto gracias a esta redundancia.

A la vista de estos resultados resulta importante comentar varios aspectos.

En primer lugar, queda patente el hecho de que aplicar una mayor redundancia supone obtener mayor probabilidad de acierto, sin olvidar también el incremento en coste computacional que esto supone, debido al manejo de matrices de mayor tamaño y el aumento del tiempo de procesado.

Los sujetos muestran tasas de acierto semejantes ante una misma longitud de palabra código, lo que nos indica que el sistema funciona bien para diferentes personas, teniendo en cuenta que la falta de atención de estas puede empeorar los resultados.

Por último, observando la desviación típica obtenida en el promedio, se pueden destacar dos hechos. El primero, que los resultados para los distintos sujetos se encuentran cerca de la media (como se ha indicado anteriormente) y el segundo y más importante, que a medida que aumenta la redundancia, la desviación típica se reduce en gran medida, lo que indica que, pese al coste computacional que la redundancia pueda generar,

obtendremos mejores resultados y con mayor probabilidad de acierto si realizamos los experimentos a un número mayor de personas.

5.6.2 Resultados experimentos Flip

Como se indicó anteriormente, los experimentos tipo Flip son aquellos en los que el resalte de las letras se produce mediante la iluminación de una fila o columna por evento y la rotación de las letras que esta abarque.

Los resultados se presentarán mediante gráficos en forma de diagrama de barras que muestran la probabilidad de acierto por sujeto y en función de la longitud de palabra código empleada. Así, las figuras 38 – 43, representan dichas probabilidades de acierto.

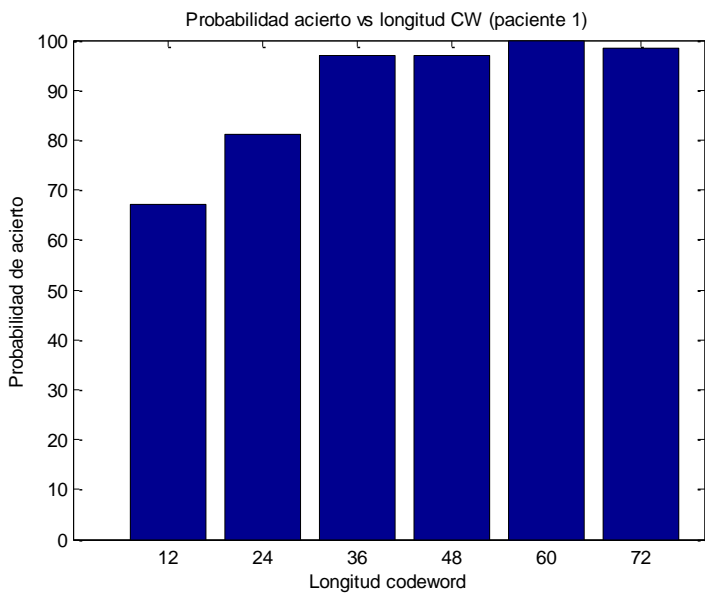


Figura 38: Probabilidad acierto 1 (Flip)

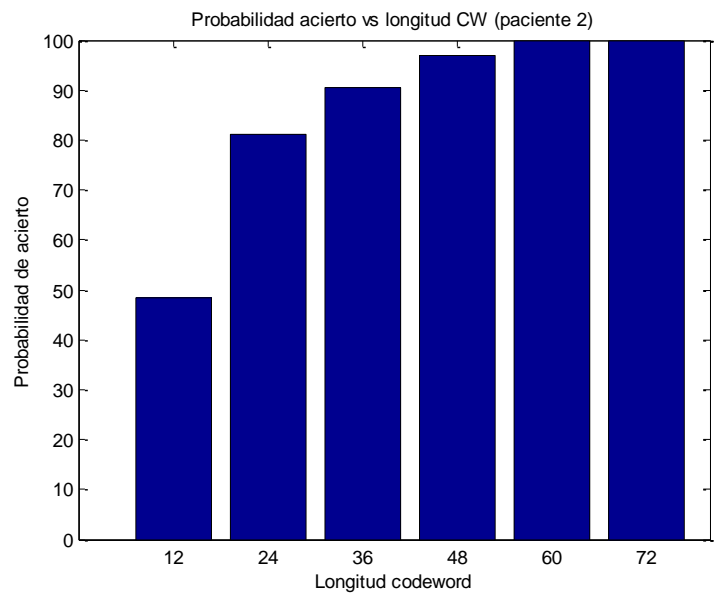


Figura 39: Probabilidad acierto 2 (Flip)

CAPITULO 5: EXPERIMENTOS Y RESULTADOS

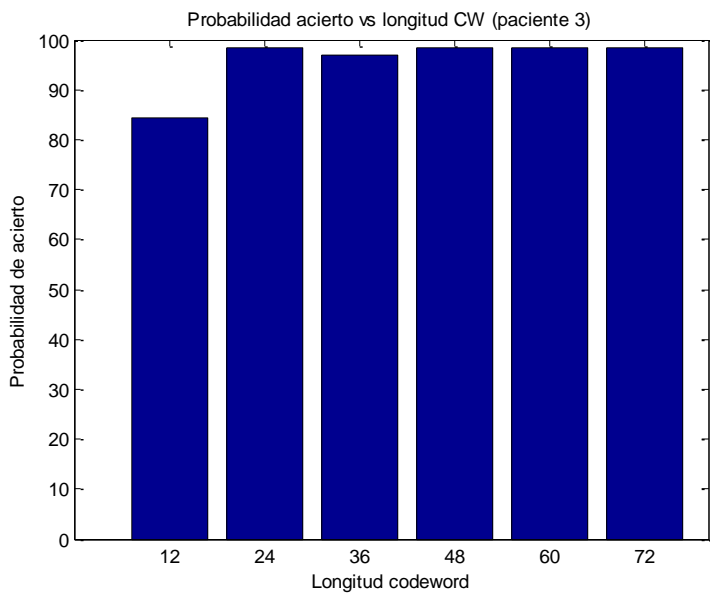


Figura 40: Probabilidad acierto 3 (Flip)

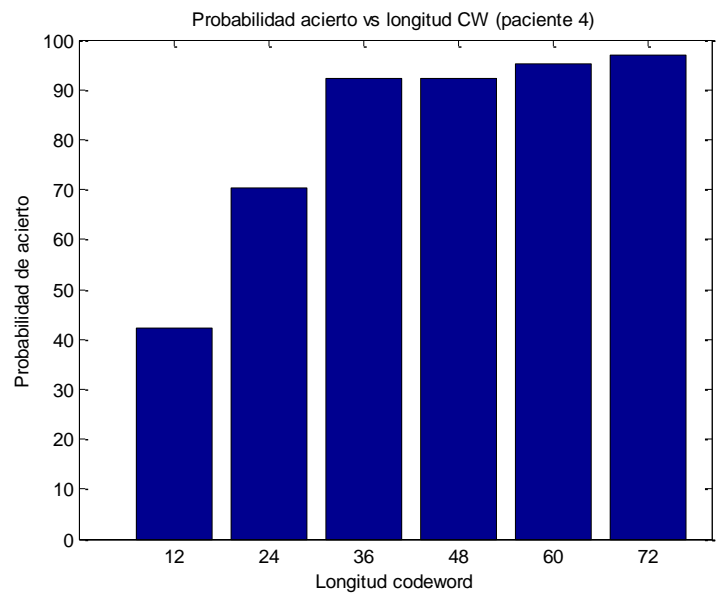


Figura 41: Probabilidad acierto 4 (Flip)

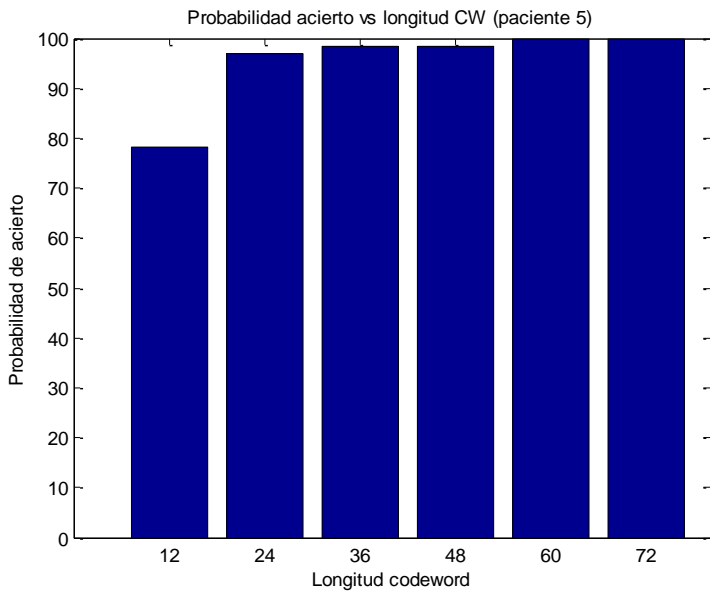


Figura 42: Probabilidad acierto 5 (Flip)

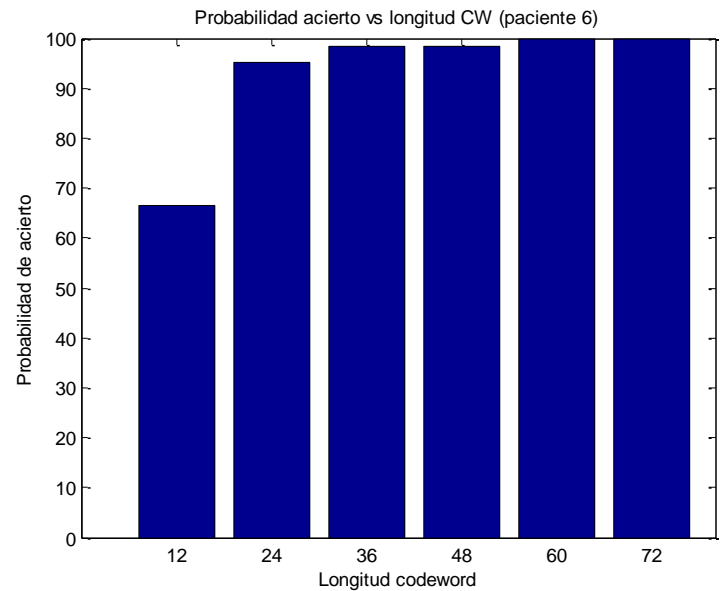


Figura 43: Probabilidad acierto 6 (Flip)

A continuación se muestra la figura más importante de todas, Figura 44, que representa la probabilidad promedio de acierto en los experimentos tipo Flip. También representa la desviación típica de los resultados obtenidos ya que será un reflejo de lo bien o mal que trabaja nuestro sistema.

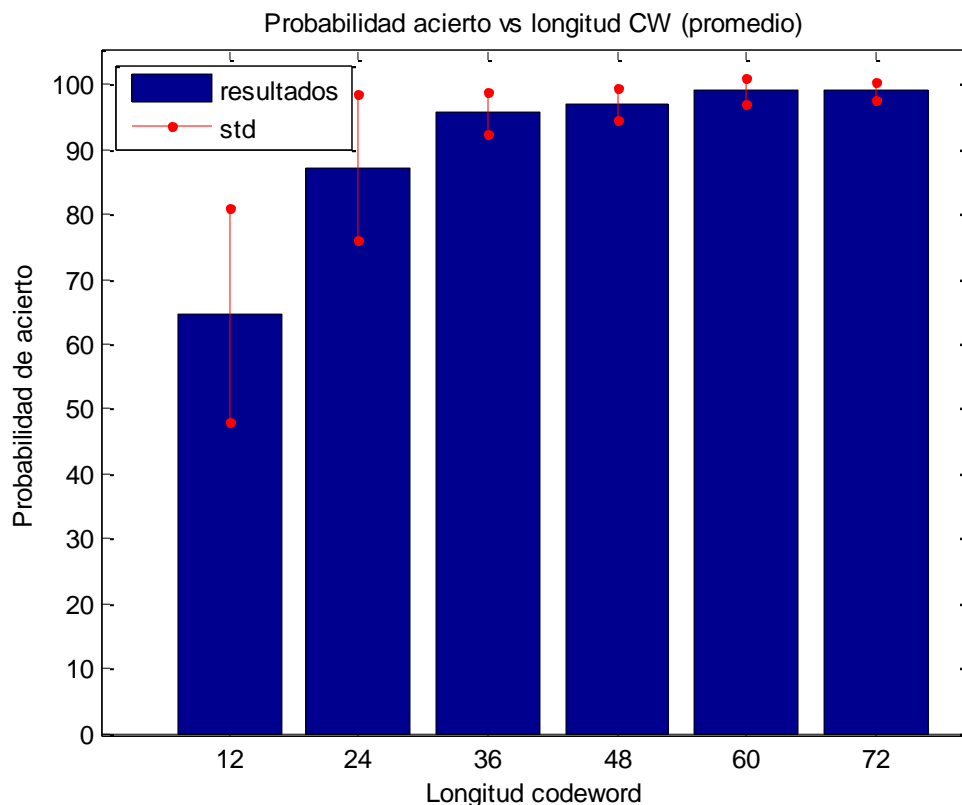


Figura 44: Probabilidad acierto promedio (Flip)

De nuevo es importante recordar que los valores de longitud de palabra código vienen determinados por la longitud mínima de palabra código, 12. El resto de valores son debidos a la repetición de un patrón de 12 bits para analizar si se mejora la probabilidad de acierto gracias a esta redundancia

A la vista de estos resultados resulta importante comentar varios aspectos.

En primer lugar, queda patente el hecho de que aplicar una mayor redundancia supone obtener mayor probabilidad de acierto, sin olvidar también el incremento en coste computacional que esto supone, debido al manejo de matrices de mayor tamaño y el aumento del tiempo de procesado.

Los sujetos muestran tasas de acierto semejantes ante una misma longitud de palabra código, lo que nos indica que el sistema funciona bien para diferentes personas, teniendo en cuenta que la falta de atención de estas puede empeorar los resultados.

Por último, observando la desviación típica obtenida en el promedio, se pueden destacar dos hechos. El primero, que los resultados para los distintos sujetos se encuentran cerca de la media (como se ha indicado anteriormente) y el segundo y más importante, que a medida que aumenta la redundancia, la desviación típica se reduce en gran medida, lo que indica que pese al coste computacional que la redundancia pueda generar,

obtendremos mejores resultados y con mayor probabilidad de acierto si realizamos los experimentos a un número mayor de personas.

5.6.3 Comparativa entre experimentos Flash y Flip

Una vez obtenido los resultados para los dos tipos de experimentos, conviene realizar una comparativa entre ambos y analizar las causas de las diferencias existentes.

Así, la Figura 45 muestra los valores promedios de ambos experimentos.

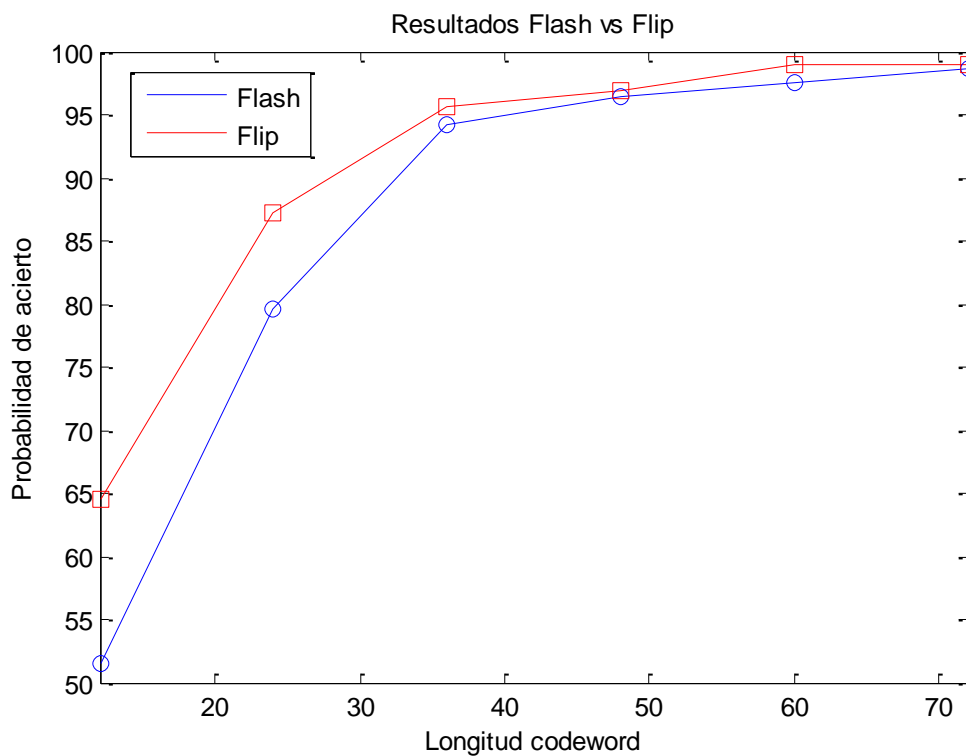


Figura 45: Probabilidad acierto promedio (Flash vs Flip)

Puede observarse cómo los experimentos tipo Flip muestran mejores resultados, no sólo a nivel de probabilidad de acierto, sino también al presentar unas desviaciones típicas menores (figuras 37 y 44).

Esto se debe al hecho de rotar las letras aparte de iluminarlas, en los experimentos tipo Flip, ya que el cerebro presenta una repuesta cerebral más fuerte al haber dos estímulos en uno, además de conseguir que el sujeto perciba mejor los estímulos al producirse varios efectos en la letra.

No obstante, las mayores diferencias se dan con la mínima redundancia, mientras que a medida que esta aumenta, ambos experimentos se aproximan al 100% de acierto.

5.6.4 Comparativa entre modelo generativo y discriminativos

Tomando como referencia el PFC de Pedro Rosillo Magariño [20], que empleó modelos discriminativos, se puede realizar una comparativa entre ambos modelos. La comparativa se realizará con los experimentos tipo Flip, ya que fue con estos con los que se realizó dicho PFC.

La Figura 46 representa la probabilidad de acierto media para el problema resuelto mediante modelos discriminativos. (Figura obtenida del PFC Métodos discriminativos para la decodificación del *Visual P300 Speller* - Página 83 - Figura 50: Gráfico conjunto con las tasas de acierto medio de letra del sistema global en los tres tipos de modelos discriminativos)

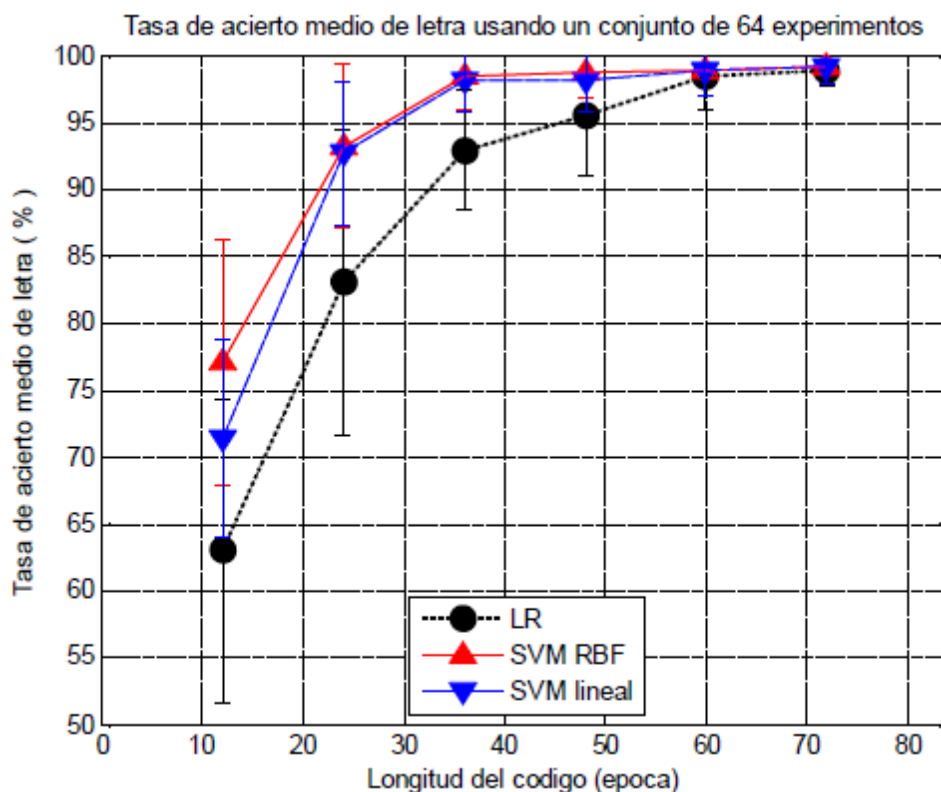


Figura 46: Probabilidad acierto promedio (Modelos discriminativos)

CAPITULO 5: EXPERIMENTOS Y RESULTADOS

La Figura 47 muestra la comparativa entre los modelos discriminativos y el modelo generativo.

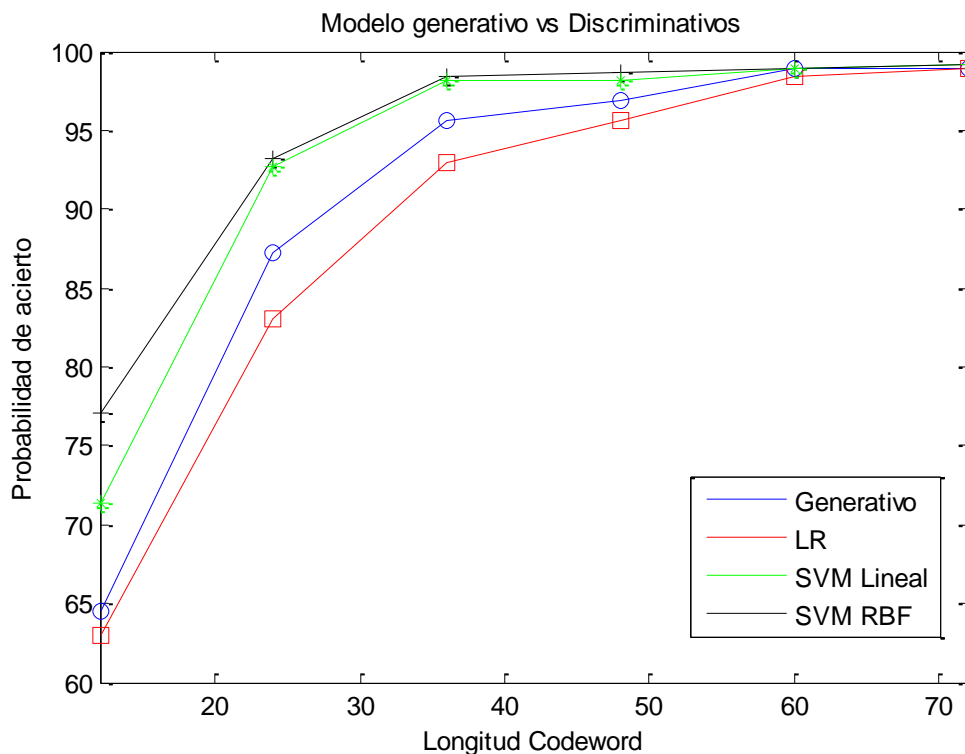


Figura 47: Probabilidad acierto promedio (Generativo vs Discriminativos)

Observando la Figura 47, se puede comprobar cómo las máquinas de vectores soporte ofrecen mejores prestaciones para longitudes de palabra código menor, mientras que el modelo generativo supera notablemente al algoritmo discriminativo LR. Se puede observar que, a la hora de introducir redundancia, el modelo generativo consigue las mismas prestaciones que cualquier modelo discriminativo.

Este hecho es muy relevante ya que la principal idea de usar modelos generativos en vez de discriminativos, es que los modelos generativos resultan más sencillos de entender conceptualmente y requieren menor carga computacional que los discriminativos. Se puede apreciar que a partir de una longitud de palabra código de 60 bits, el uso de un modelo generativo presenta las mismas capacidades que las SVM, por lo que, y debido a su mayor simplicidad, conviene recomendar el uso de algoritmos generativos en vez de discriminativos.

5.7 Mejoras y nuevas aportaciones

5.7.1 Introducción

En este proyecto se presentan dos mejoras o nuevas aportaciones con el objetivo de mejorar las prestaciones del modelo generativo desarrollado. En un primer lugar se propondrá el uso de las propiedades del algoritmo PCA para reducir la dimensión del problema y mejorar la carga computacional. Mientras que una segunda vía de estudio será añadir más canales con versiones retardadas de la señal, tomando como conjunto de señales la original y la misma con una serie de retardos, para comprobar si se produce una mejora en la probabilidad de acierto, aun pagando un precio en coste computacional.

5.7.2 Reducción de la dimensión del problema

A la hora de obtener los resultados anteriores se comprobó que el tiempo de procesamiento era muy elevado, debido a las múltiples repeticiones de pruebas (sistema validado mediante *cross-validation*), y que las necesidades de memoria eran elevadas debido al tamaño de las matrices con las que se opera. Por ello, una primera idea para mejorar el sistema, fue intentar reducir dicha dimensión sin que ello afectase a los resultados obtenidos. Para ello se tuvo en cuenta las propiedades del algoritmo empleado para obtener la matriz de blanqueado (Análisis de Componentes Principales – PCA (Ver Capítulo 3)). Anteriormente sólo estábamos interesados en blanquear la señal, pero ahora podemos utilizar esta matriz de blanqueado W , para blanquear la señal y para reducir su dimensión. Como se explicó en el capítulo 3, al reducir la dimensión del problema mediante el algoritmo PCA incurriremos en un error que puede empeorar las prestaciones.

Los motivos por el cual haremos uso de estas características de la matriz de blanqueado son, reducir la carga computacional y eliminar las componentes ruidosas ya que la información relevante estará contenida en pocos canales.

Teniendo esto en cuenta, se decidió realizar un estudio, considerando todos los sujetos y empleando los experimentos tipo Flip (por ofrecer mejores prestaciones), consistente en disminuir la dimensión de 58 canales iniciales a x , donde x irá de 1 a 58 progresivamente.

La Figura 48 muestra la evolución de la probabilidad de acierto en función del número de canales y de la longitud de palabra código empleados.

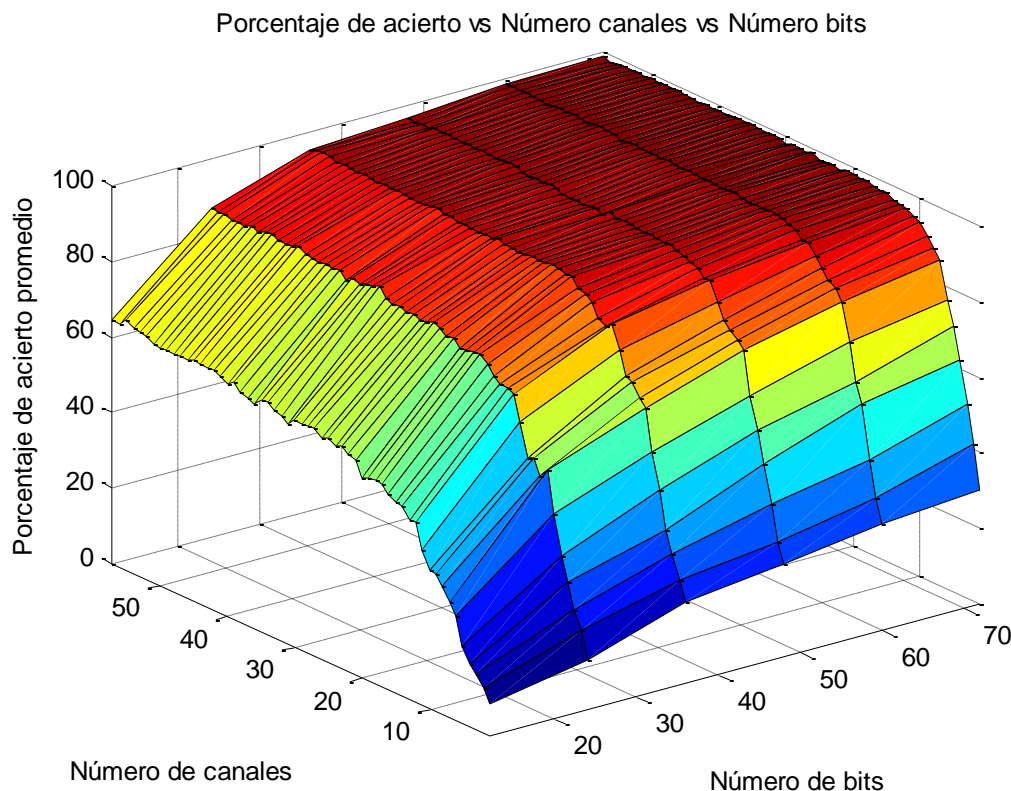


Figura 48: Evolución de la probabilidad de acierto en función del número de canales usados

Se observan dos hechos. En primer lugar, y como era de esperar a la vista de los resultados anteriores, al aumentar la redundancia incrementando la longitud de palabra código, mejoran las prestaciones. En segundo lugar, se aprecia que reducir demasiado la dimensión del problema conlleva a obtener resultados peores (del 100% en el mejor caso al 30% en el peor).

Pero resulta interesante observar que se puede reducir la dimensión del problema drásticamente (casi a la mitad) sin perder apenas prestaciones. La Figura 49 muestra la misma gráfica mediante diagramas de barras, 6 diagramas de barra según la longitud de la palabra código, para cada número de canales usados.

A tenor de los resultados obtenidos, podemos decir que reducir la dimensión de 58 a 40-45 cuando usamos menos redundancia nos proporciona los mismos resultados que si usamos todos los canales. Pero lo realmente interesante viene al analizar los resultados para la máxima redundancia (72 bits de longitud de palabra código) ya que aquí podemos reducir la dimensión hasta usar solamente 26 canales, menos de la mitad del experimento original, lo que reducirá drásticamente la carga computacional y el tiempo de ejecución, y mantener una probabilidad de acierto superior al 97%.

CAPITULO 5: EXPERIMENTOS Y RESULTADOS

Como nota informativa la diferencia entre usar un canal o 58 canales se traduce en 45 minutos por experimento.

Porcentaje de acierto promedio vs número de canales usados

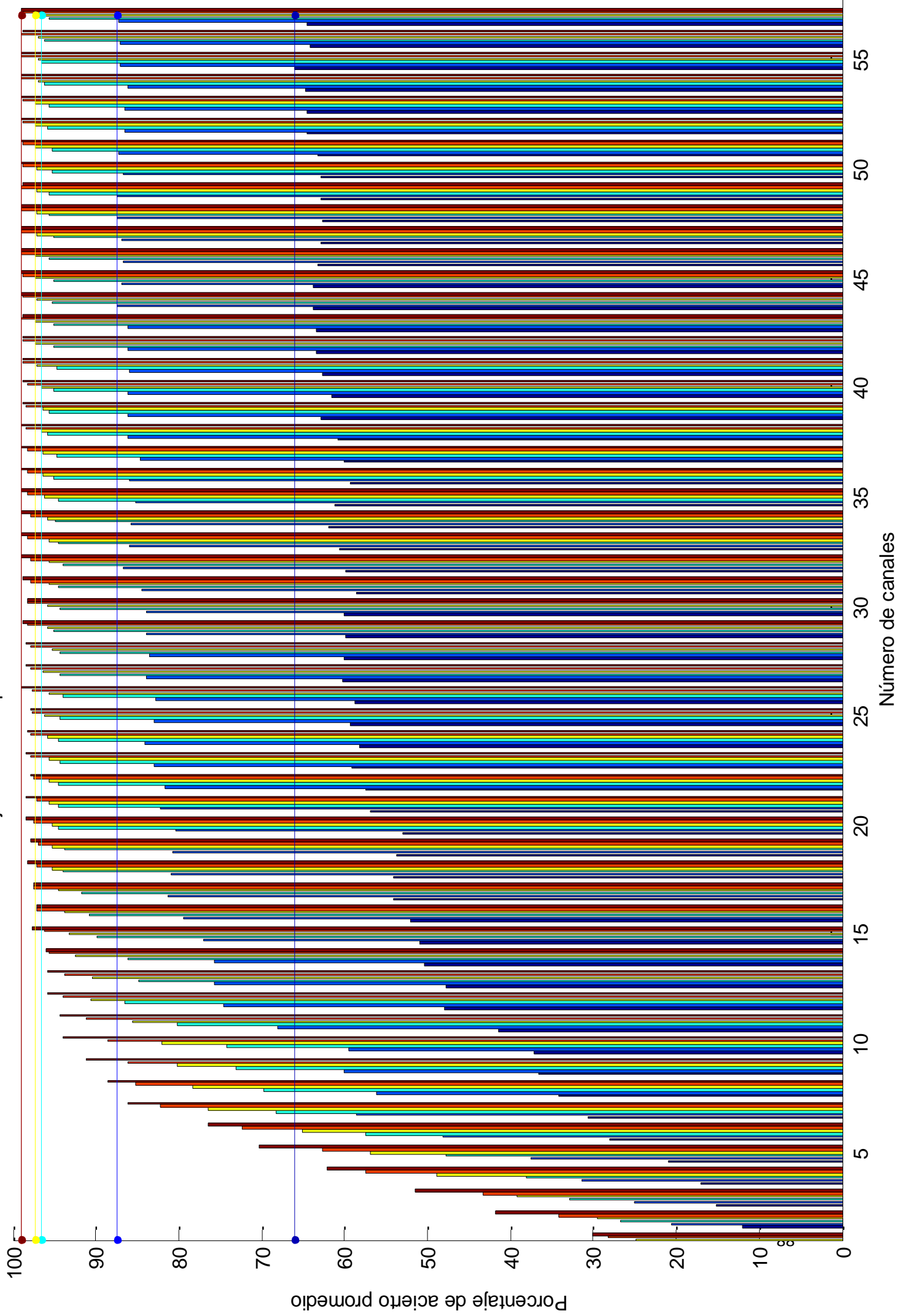


Figura 49: Evolución de la probabilidad de acierto

5.7.3 Adición de canales retardados

En esta ocasión lo que se busca es mejorar las prestaciones del sistema añadiendo versiones retardadas de los canales y aplicar a la vez la mejora previa de reducción de la dimensión del problema. Para ello se realizará un estudio donde a la señal original se le añadirá la misma señal pero retardándola un número x de muestras, donde x irá de 1 a 4.

La idea de añadir versiones retardadas de los canales [26], [27], es eliminar la dependencia temporal que pueda existir entre las diferentes partes del cerebro. Es decir, se considera que las señales cerebrales en una determinada posición del cerebro y en un determinado instante de tiempo influirán en otras regiones en instantes posteriores.

En esta ocasión la carga computacional fue demasiado elevada por lo que se tuvo que recurrir a los ordenadores del departamento de Teoría de la Señal y Comunicaciones de la universidad Carlos III de Madrid para poder llevar a cabo los experimentos. Finalmente se paró en 4 muestras el número de retardos, tras 2 meses simulando, porque se consideró suficiente para poder comprobar las mejoras, si las hubiera.

Las figuras 50 – 53 muestran los resultados obtenidos según el número de retardos empleados y en función del número de canales y longitud de palabra código usados.

Porcentaje acierto vs Número canales vs Número bits (1 retardo)

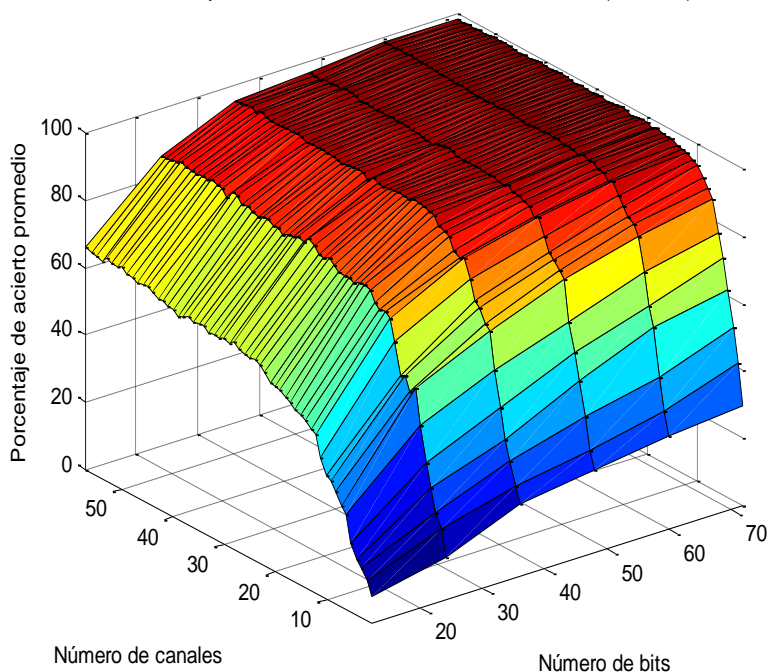


Figura 50: Probabilidad acierto 1 retardo

Porcentaje acierto vs Número canales vs Número bits (2 retardos)

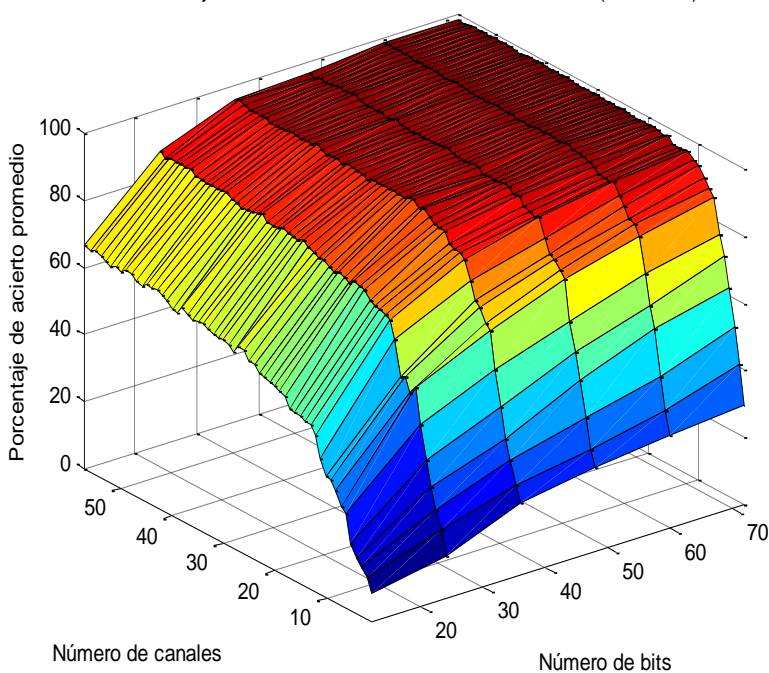


Figura 51: Probabilidad acierto 2 retardos

CAPITULO 5: EXPERIMENTOS Y RESULTADOS

Porcentaje acierto vs Número canales vs Número bits (3 retardos)

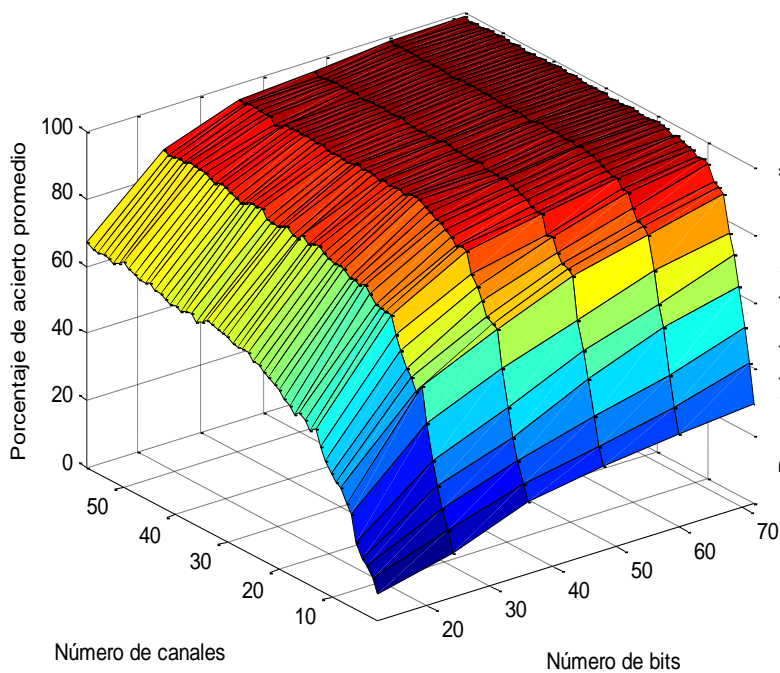


Figura 52: Probabilidad acierto 3 retardos

Porcentaje acierto vs Número canales vs Número bits (4 retardos)

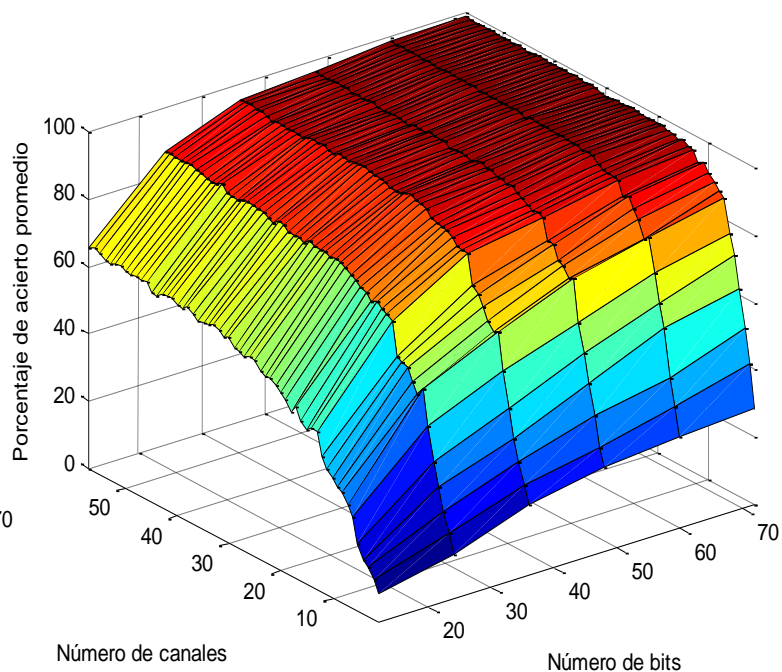


Figura 53: Probabilidad acierto 4 retardos

Observando las figuras resulta casi imperceptible notar alguna mejoría, y si bien podría observarse alguna pequeña mejora, esta no es lo suficientemente buena como para tomar en consideración la adición de nuevos canales mediante el uso de la señal retardada, debido al enorme coste computacional que acarrea consigo. También hay que indicar que quizá con un número elevado de retardos las prestaciones pudieran mejorar considerablemente y se deja como objeto de estudio futuro este hecho, siempre que se controle el coste computacional que conlleva.

De nuevo, las figuras 54 – 57, muestran los mismos resultados representados mediante diagramas de barras e indicando el número óptimo de canales en función de la longitud de palabra código, adecuado para cada uno de los retardos.

Cabe destacar que al añadir las versiones retardadas, necesitamos quedarnos con un mayor número de canales para obtener la máxima probabilidad de acierto.

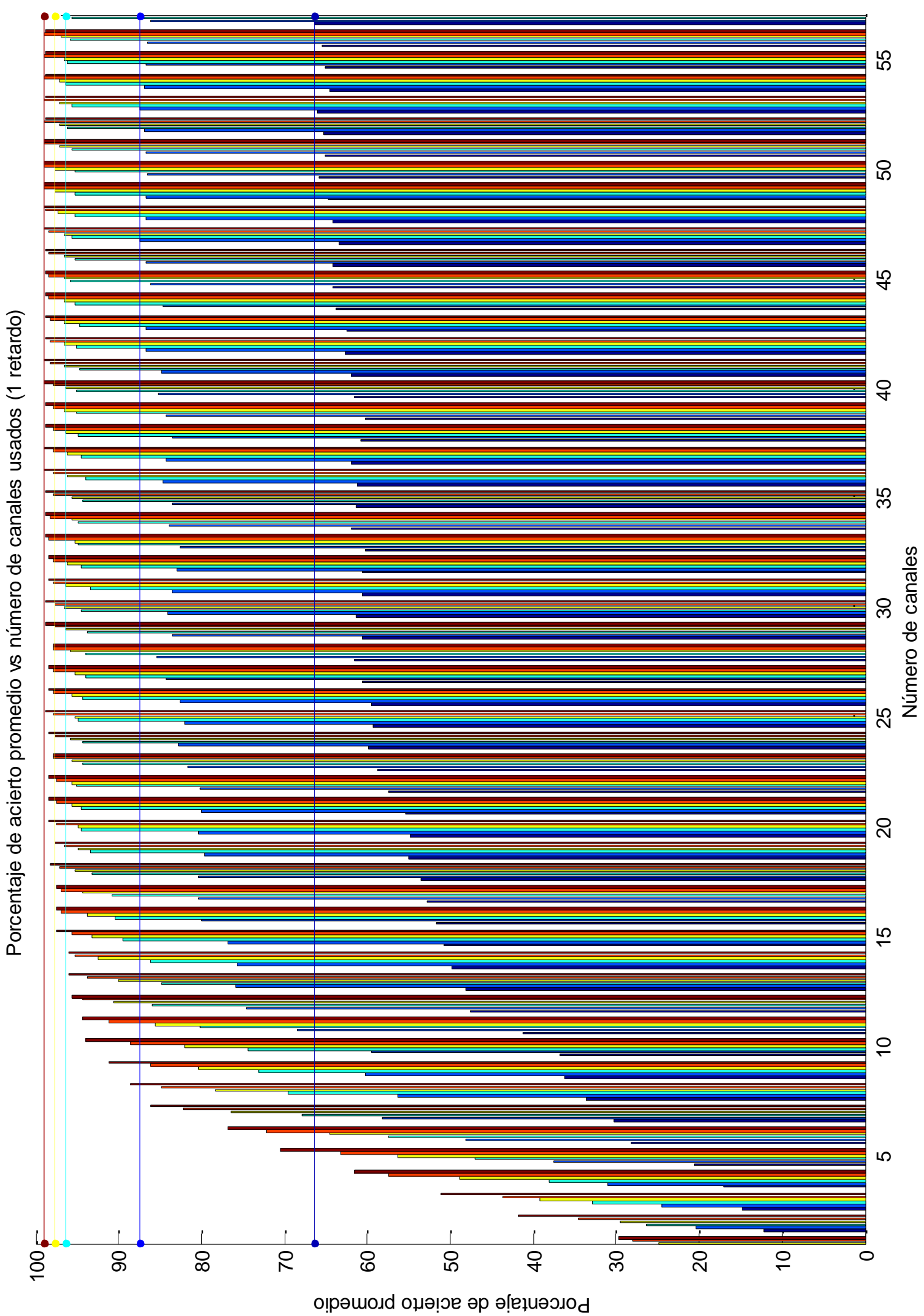


Figura 54: Evolución de la probabilidad acierto (1 retardo)

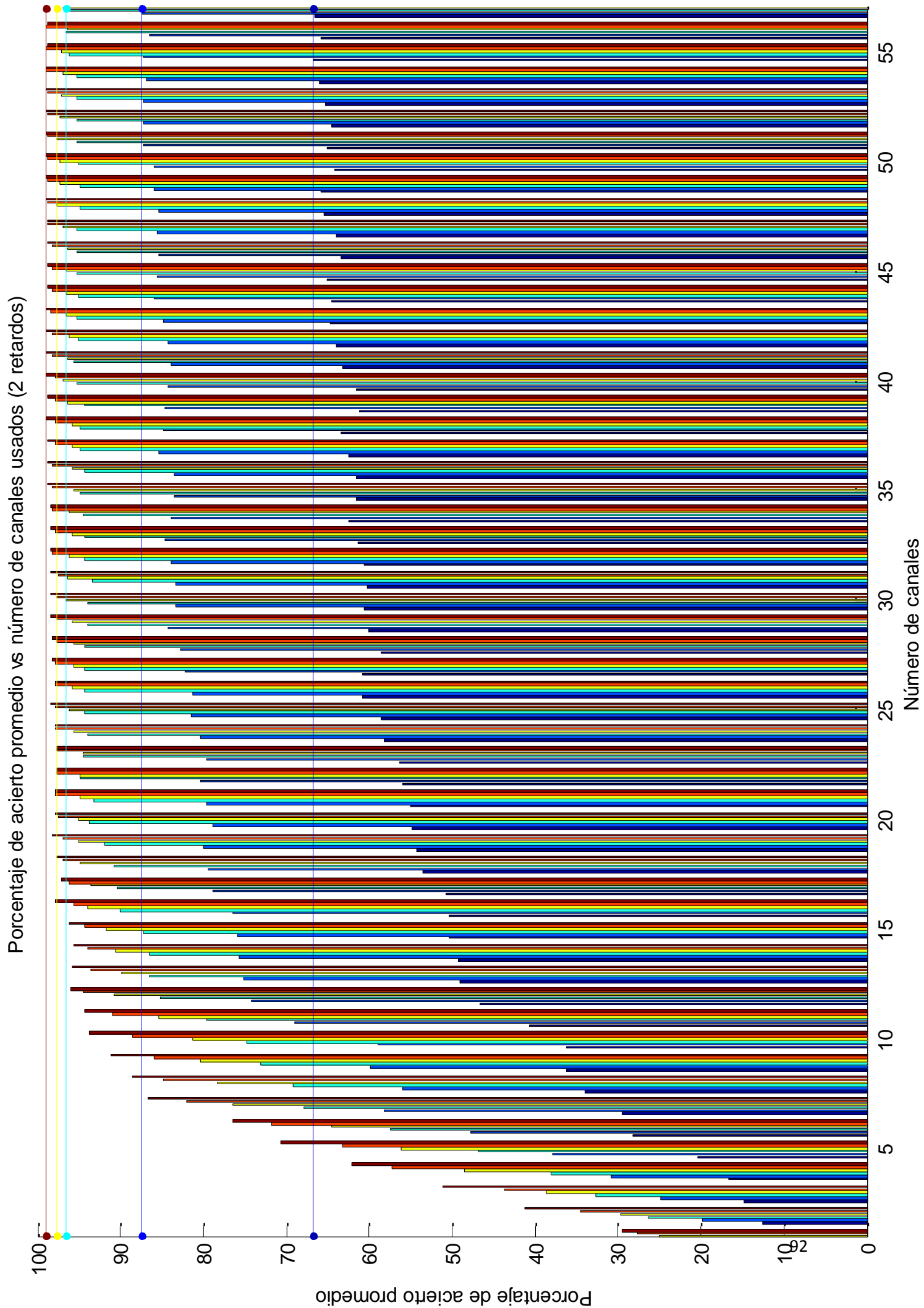


Figura 55: Evolución de la probabilidad acierto (2 retardos)

Porcentaje de acierto promedio vs número de canales usados (3 retardos)

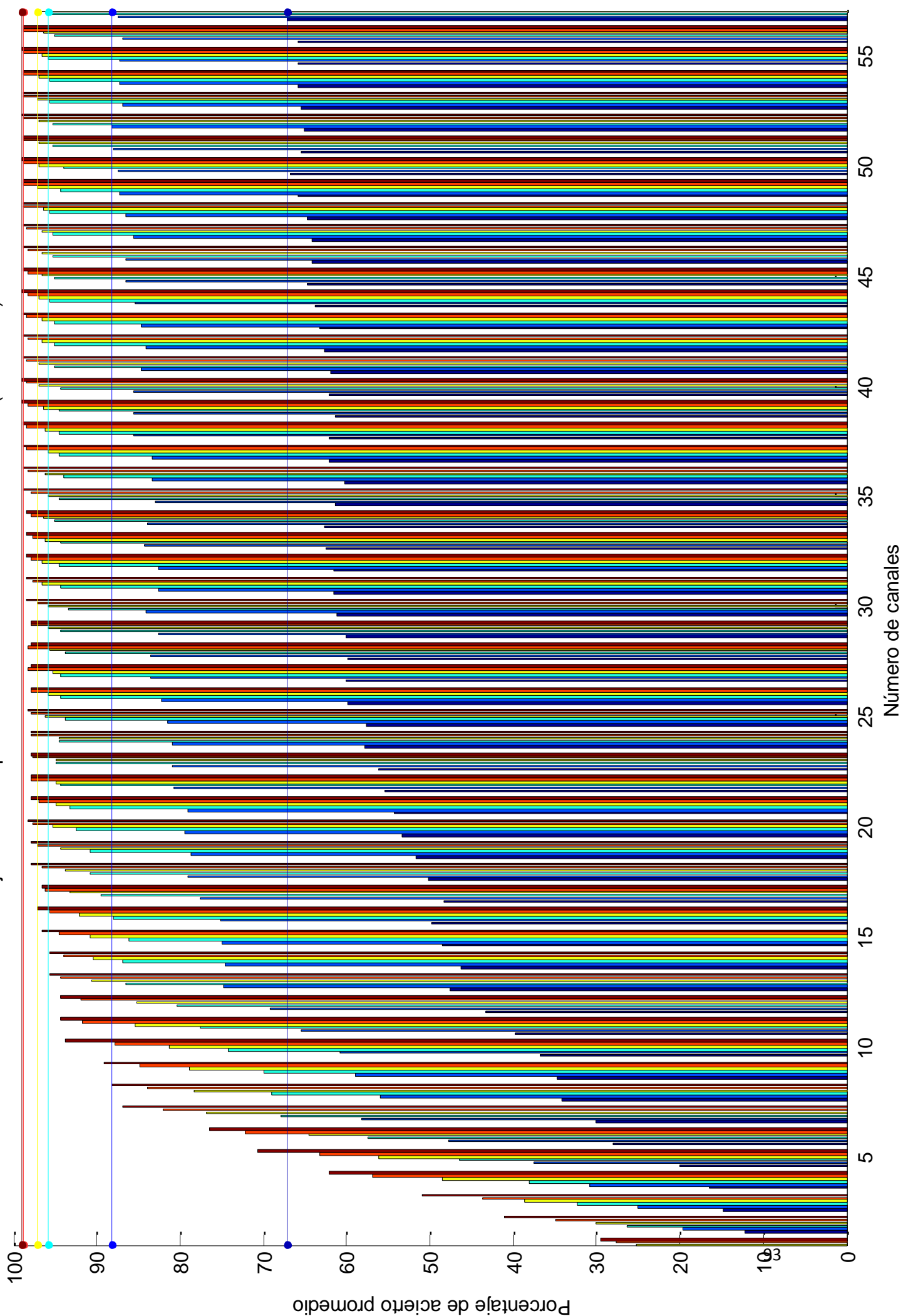


Figura 56: Evolución de la probabilidad acierto (3 retardos)

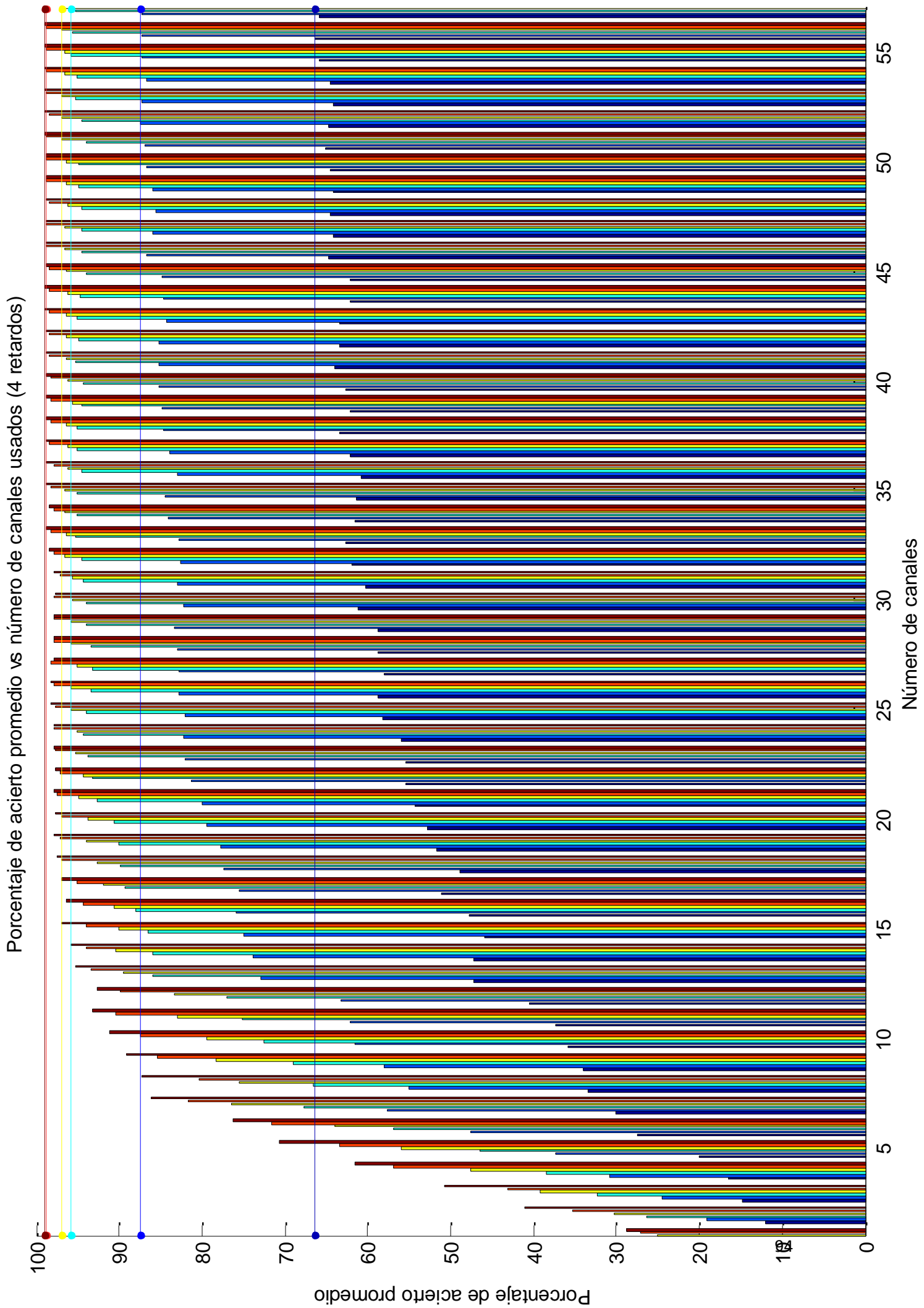


Figura 57: Evolución de la probabilidad acierto (4 retardos)

CAPITULO 5: EXPERIMENTOS Y RESULTADOS

Para una mejor comparativa de estos resultados, se presenta la Figura 58, que muestra las probabilidades de acierto empleando 72 bits, para los diferentes retardos.

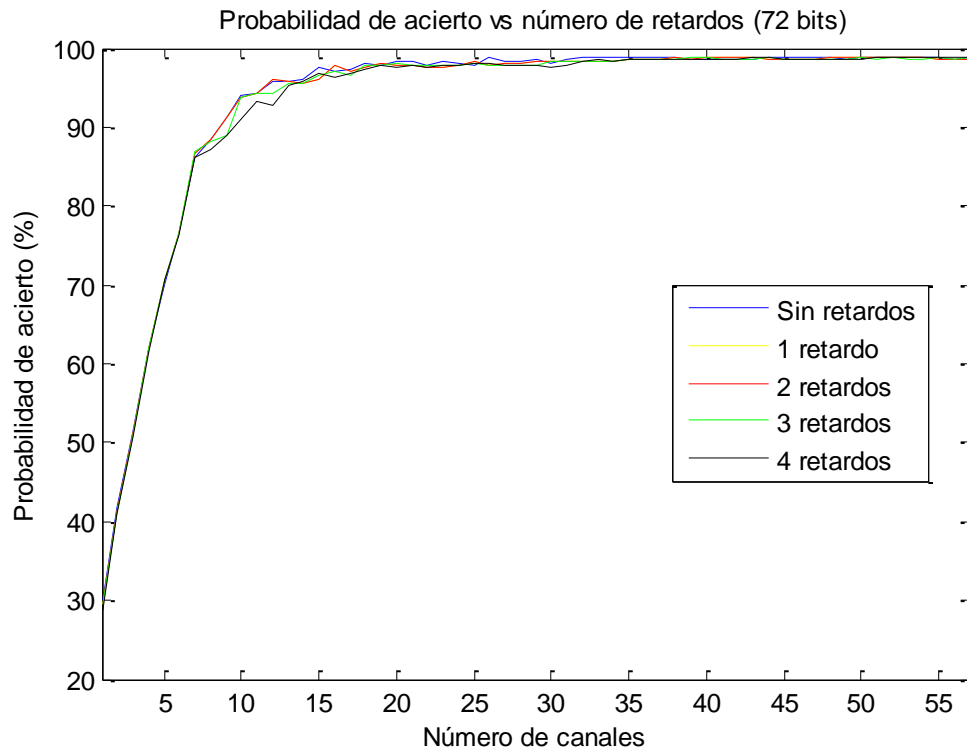


Figura 58: Probabilidad de acierto vs número de retardos (72 bits)

Como puede observarse en la figura, de existir una mejora, esta se dará para un elevado número de canales y será mínima, mientras que para un número reducido de canales puede comprobarse que añadir canales mediante la señal retardada empeora las prestaciones (curva de aprendizaje más lenta).

CAPITULO 5: EXPERIMENTOS Y RESULTADOS

Capítulo 6

Conclusiones y futuras líneas de trabajo

6.1 Introducción

El objetivo de este capítulo es extraer las conclusiones más importantes del proyecto fin de carrera realizado, así como las posibles líneas de trabajo futuras que se han encontrado o han abierto al realizarlo.

6.2 Conclusiones

A lo largo de este proyecto se ha desarrollado un modelo generativo con el fin de conseguir que personas con capacidad de movimiento limitada, pudieran transmitir letras sin más que fijando su atención en la letra deseada dentro de una matriz de letras. Para ello hemos tenido que realizar un pre-procesado de la señal debido al bajo nivel de señal a ruido que presentan las señales EEG, para posteriormente aplicar un algoritmo

CAPITULO 6: CONCLUSIONES Y FUTURAS LÍNEAS DE TRABAJO

generativo que permitiese, mediante un decisor MAP, estimar la letra que el sujeto quiere comunicar.

Por todo ello, hay que tener en cuenta determinados factores que permitan obtener las mejores prestaciones del sistema:

- Eliminar todo el ruido que tienen nuestros registros, mediante un filtrado paso banda que nos permita resaltar las características deseadas de la señal proveniente del EEG.
- La importancia de tener una referencia temporal que nos permita conocer en qué momento determinado se produjo un estímulo.
- Cumplir el teorema de Nyquist para poder llevar a cabo un diezmado de la señal que nos permita reducir la dimensión del problema y eliminar la redundancia innecesaria de la información.
- Decorrelar la información de los 58 canales del EEG en vez de dar más valor a unos sensores que a otros. Este cambio conceptual es una de las claves de este sistema porque pasamos de tener un cerebro modular, donde las actividades definen, qué sensores contienen más información (actividad visual, actividad cognitiva, actividad motora, etc.) a un cerebro donde la información está distribuida en todos los módulos funcionales que nos definen las posibles actividades cerebrales, pudiendo aportar datos útiles, zonas que en principio no parecen propicias para encontrar los potenciales P300. Gracias a este cambio de paradigma seremos capaces de proporcionar al clasificador, una época más detallada con cierta información sutil que tiene lugar en algunas zonas del cerebro en el momento que estimulamos con impulsos lumínicos sorprendentes.
- Destacar la importancia que tiene en el modelo generativo el incluir toda la información conocida sobre las señales a procesar, para conseguir un mejor modelado de las plantillas de eventos objetivo (t) y no objetivo (n).
- La importancia a la hora de elegir la forma en la que se destacan las letras, ya que se ha demostrado que el modelo Flip presenta mejores prestaciones debido a que se producen más cambios, que el cerebro detecta.
- Debido a que el sujeto debe llevar la cuenta de los eventos estímulo, resulta crucial la atención de este a la hora de llevar a cabo los experimentos.
- La comparativa con modelos discriminativos demuestra que el modelo generativo puede presentar las mismas prestaciones.
- Los costes computacionales para llevar a cabo los experimentos resultan elevados debido a la gran cantidad de datos a manejar.

A la vista de las dos nuevas aportaciones en el campo realizadas en este proyecto, se pueden destacar varios puntos:

- El hecho de emplear el algoritmo PCA no sólo para decorrelar sino también para reducir la dimensión de la señal, supone un ahorro en el coste computacional que debe ser tomado en consideración.
- Por otro lado, el añadir canales extra usando la señal retardada x muestras, no parece valer la pena ya que no se consigue una mejora suficiente en las prestaciones y se paga para ello un precio muy elevado en términos de coste computacional. Además, esta adición de señales retardadas, hace que la curva de

aprendizaje se ralentice, lo que supone eliminar la mejora encontrada al reducir la dimensión del problema mediante el algoritmo PCA.

- Se descarta el uso de canales con señales retardadas ya que en términos globales empeora el sistema.

6.3 Futuras líneas de trabajo

Las futuras líneas de trabajo que se pueden desarrollar para este trabajo son las siguientes:

- Modificar la matriz de letras que se muestra para que en vez de deletrear letras, se puedan llevar a cabo acciones mediante el manejo de aparatos electrónicos (mover en cualquier dirección, girar, subir, bajar, pulsar un botón, etc.)
- Estudiar los tipos de estímulo que presentar al sujeto ya que se ha demostrado que cuanto más varíe la letra elegida, mejores prestaciones se obtienen.
- Empleo de *error-correcting P300 speller* [22].
- Estudiar diferentes tipos de algoritmos PCA que permitan reducir más la dimensión del problema sin perder la calidad de las prestaciones.
- Aunque se haya descartado, confirmar la posible mejora con un elevado número de canales al añadir señales con mayores retardos.
- Conseguir un código más óptimo o en un lenguaje más potente que Matlab. Se intentó implementar en java para usar programación concurrente pero se desechó la idea por lo complicado del manejo de matrices frente a Matlab.
- Estudio de diferentes filtrados que permitan reducir más aun si cabe el nivel de ruido en los datos.

CAPITULO 6: CONCLUSIONES Y FUTURAS LÍNEAS DE TRABAJO

Capítulo 7

Presupuesto

7.1 Presupuesto del proyecto

El proyecto se ha realizado en tres grandes etapas. En la primera se investigó el estado del arte de los temas de los que trata este trabajo. A continuación, se desarrolló el modelo generativo y se realizaron las pruebas del sistema oportunas. Finalmente, se implementaron las mejoras y nuevos aportes y se probaron.

El desglose de costes de personal, costes de equipo y otros costes directos, se muestra a continuación en la Figura 59.

CAPITULO 7: PRESUPUESTO



UNIVERSIDAD CARLOS III DE MADRID
Escuela Politécnica Superior

PRESUPUESTO DE PROYECTO

1.- Autor: Javier de la Torre Abaitua

2.- Departamento: TSC

3.- Descripción del Proyecto:

- Título: PROCESADO DE SEÑALES EEG PARA UN INTERFAZ CEREBRO-MÁQUINA (BCI)

- Duración (meses) **12**

Tasa de costes Indirectos: **20%**

4.- Presupuesto total del Proyecto (valores en Euros):

Euros

5.- Desglose presupuestario (costes directos)

PERSONAL

Apellidos y nombre	N.I.F. (no rellenar - solo a título informativo)	Categoría	Dedicación (meses) ^{a)}	(hombres)	Coste hombre mes	Coste (Euro)	Firma de conformidad
de la Torre Abaitua Javier		Ingeniero		6,7	2.694,39	18.052,41	
Leiva Murillo Jose Miguel		Ingeniero Senior		0,7	4.289,54	3.002,68	
Técnico Laboratorio		Ingeniero		0,1372	2.694,39	369,67	
						0,00	
						0,00	
Hombres mes 7,5372						Total	21.424,76

^{a)} 1 Hombre mes = 131,25 horas. Máximo anual de dedicación de 12 hombres mes (1575 horas)
Máximo anual para PDI de la Universidad Carlos III de Madrid de 8,8 hombres mes (1.155 horas)

EQUIPOS

Descripción	Coste (Euro)	% Uso dedicado proyecto	Dedicación (meses)	Periodo de depreciación	Coste imputable ^{d)}
Casco Electrodo	770,00	100	0	60	1,28
Ordenador Personal	1.000,00	100	8	60	133,33
Ordenador TSC	1.000,00	100	3	60	50,00
Programas Simulación	2.000,00	100	12	60	400,00
		100		60	0,00
					0,00
Total					584,62

^{d)} Fórmula de cálculo de la Amortización:

$$\frac{A}{B} \times C \times D$$

A = nº de meses desde la fecha de facturación en que el equipo es utilizado

B = periodo de depreciación (60 meses)

C = coste del equipo (sin IVA)

D = % del uso que se dedica al proyecto (habitualmente 100%)

SUBCONTRATACIÓN DE TAREAS

Descripción	Empresa	Coste imputable
Total		0,00

OTROS COSTES DIRECTOS DEL PROYECTO^{e)}

Descripción	Empresa	Costes imputable
Médico para medir EEG	Hospital	630,00
Total		630,00

^{e)} Este capítulo de gastos incluye todos los gastos no contemplados en los conceptos anteriores, por ejemplo: fungible, viajes y dietas, otros,...

6.- Resumen de costes

Presupuesto Costes Totales	Presupuesto Costes Totales
Personal	21.425
Amortización	585
Subcontratación de tareas	0
Costes de funcionamiento	630
Costes Indirectos	4.528
Total	27.167

Figura 59: Presupuesto del proyecto

CAPITULO 7: PRESUPUESTO

Por lo tanto, podemos concluir que el presupuesto de este proyecto asciende a 27.167 €. Finalmente, se muestra el diagrama de Gantt en la Figura 60 con las tareas realizadas.



Figura 60: Diagrama de Gantt del proyecto

CAPITULO 7: PRESUPUESTO

Capítulo 8

Glosario

BCI	Brain Computer Interface
MAP	Maximum a Posteriori
LS	Least Square
PCA	Principal Component Analysis
LOO	Leave One Out
RC	Row-Column
SOA	Stimulus Onset Asynchronies
TTI	Target-to-Target Interval
TNI	Target-to-Non target Interval
EEG	Electroencephalogram
SVM	Support Vector Machine

CAPITULO 8: GLOSARIO

Capítulo 9

Referencias

9.1 Introducción

[1] Farwell L A and Donchin E 1988 Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials *Electroencephalogr. Clin. Neurophysiol.* 70 510–23

[2] Martens S M M, Hill N J, Farquhar J and Schölkopf B 2009 Overlap and refractory effects in a brain–computer interface speller based on the visual P300 event-related potential *J. Neural Eng.* 6 026003

[3] Rivet B, Souloumiac A, Attina V and Gibert G 2009 xDAWN algorithm to enhance evoked potentials: application to brain–computer interface *IEEE Trans. Biomed. Eng.* 56 2035–43

[4] Ng A Y and Jordan M I 2001 On discriminative versus generative classifiers: a comparison of logistic regression and naive Bayes *NIPS* 841–8

[5] Rubinstein Y D and Hastie T 1997 Discriminative versus informative learning *Proc. 3rd Int. Conf. Knowledge Discovery and Data Mining* pp 49–53

CAPITULO 9: REFERENCIAS

- [6] Birbaumer N 2006 Brain–computer interface research: coming of age *Clin. Neurophysiol.* 117 479–83
- [7] Dobkin B H 2007 Brain–computer interface technology as a tool to augment plasticity and outcomes for neurological rehabilitation *J. Physiol.* 579 637–42
- [8] Sutton S, Braren M, Zubin J and John E R 1965 Evoked-potential correlates of stimulus uncertainty *Science* 150 1187–8
- [9] Rugg M D and Coles M G H 1996 *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition* (Oxford: Oxford University Press)
- [10] Kaper M, Meinicke P, Grossekhoefer U, Lingner T and Ritter H 2004 BCI competition 2003–data set IIb: support vector machines for the P300 speller paradigm *IEEE Trans. Biomed. Eng.* 51 1073–6
- [11] Kay S M 1993 *Fundamentals of Statistical Signal Processing, Vol I (Estimation Theory)* (Englewood Cliffs, NJ: Prentice Hall) p 9
- [12] S M M Martens and J M Leiva 18 February 2010 A generative model approach for decoding in the visual event-related potential-based brain–computer interface speller Empirical Inference Department, Max Planck Institute for Biological Cybernetics, Tübingen, Germany
- [13] Proakis J G 2008 *Digital Communications 5th edn* (New York: McGraw-Hill)
- [14] V.N. Vapnik. 1998 *Statistical Learning Theory*. John Wiley & Sons, Inc.
- [15] L. González Diciembre 2000. Teoría del aprendizaje estadístico de la regresión. Máquinas de regresión de vector base. Trabajo interno del departamento de economía aplicada I, Facultad de Ciencias Económicas y Empresariales, Universidad de Sevilla..
- [16] L. González and J.M. Alba Julio 2001. Similitud entre sucesos. Terceras Jornadas de Trabajo sobre Metodologías Cualitativas Aplicadas a los Sistemas Socioeconómicos.
- [17] C. Angulo Abril 2001. Aprendizaje con máquinas núcleos en entornos de multclasificación. Tesis doctoral, Universidad Politécnica de Cataluña.
- [18] Alan V. Oppenheim, Alan S. Willsky, S. Hamid Nawad *Señales y Sistemas* segunda edición.
- [19] Nuevas Tecnologías, Revista PocketBite, N°7, Jun 2009, páginas 20-28
- [20] Pedro Jonathan Rosillo Magariño Octubre 2011 Métodos discriminativos para la decodificación del Visual P300 Spelle.
- [21] Jeremy Hill, Suzanne Martens Effects of Stimulus and error of Error Correcting Code Design on BCI Speller Performance

CAPITULO 9: REFERENCIAS

[22] Felix Bießmann Julio 2007, Thesis – Error Correction Codes for the P300 Visual Speller.

[23] Antonio Artés Rodríguez , Comunicaciones digitales

[24] Alan V. Oppenheim, Señales y Sistemas – Segunda Edición

[25] Muñoz-Céspedes , Aplicación de la neuroimagen funcional al estudio de la rehabilitación neuropsicológica

[26] Steven Lemmm, Spatio-Spectral Filters for Improving the Classification of Single Trial EEG

[27] Guido Dornhege, Combined Optimizarion of Spatial and Temporal Filters for Improving Brain-Computer Interfacing

WEB

[1] <http://es.wikipedia.org/wiki/P300>

[2] www.it.uc3m.es/jvillena/irc Apuntes IRC Julio Villena UCIIIIM 02/11/2011 20:18

[3] http://issuu.com/samaelarchangel/docs/pocket_bite_no_7_ok

CAPITULO 9: REFERENCIAS

Capítulo 10

Anexo

10.1 Sistema internacional 10-20: partes del Cerebro

El sistema internacional de referencia 10-20 se diseñó para homogenizar la colocación de los electrodos al usar un EEG. La Figura 61 junto con la Figura 63 (representación del casco de malla empleado para realizar las capturas), muestra la concordancia con la zona específica del cerebro a la que está asociado cada sensor del sistema 10-20. La Figura 62 describe las principales partes de las que se compone el cerebro.

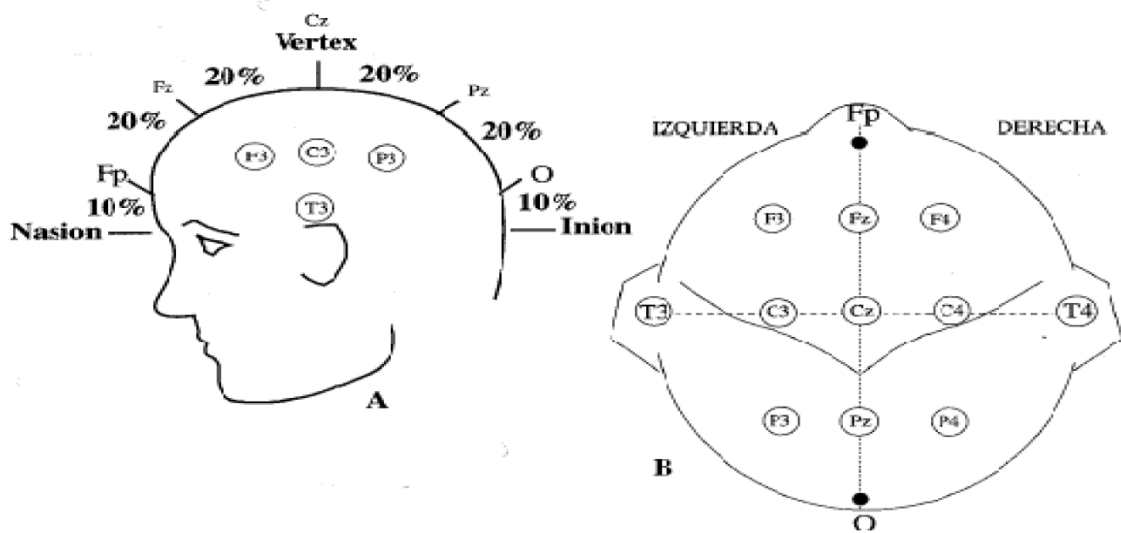


Figura 61: Posiciones de referencia en el sistema internacional 10-20

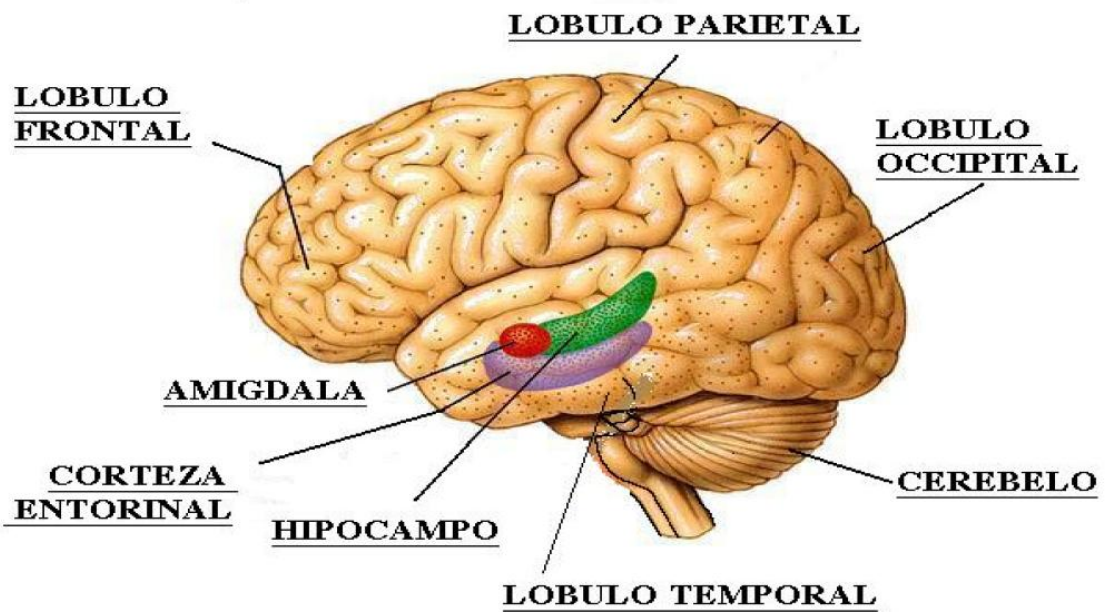


Figura 62: Partes principales del cerebro

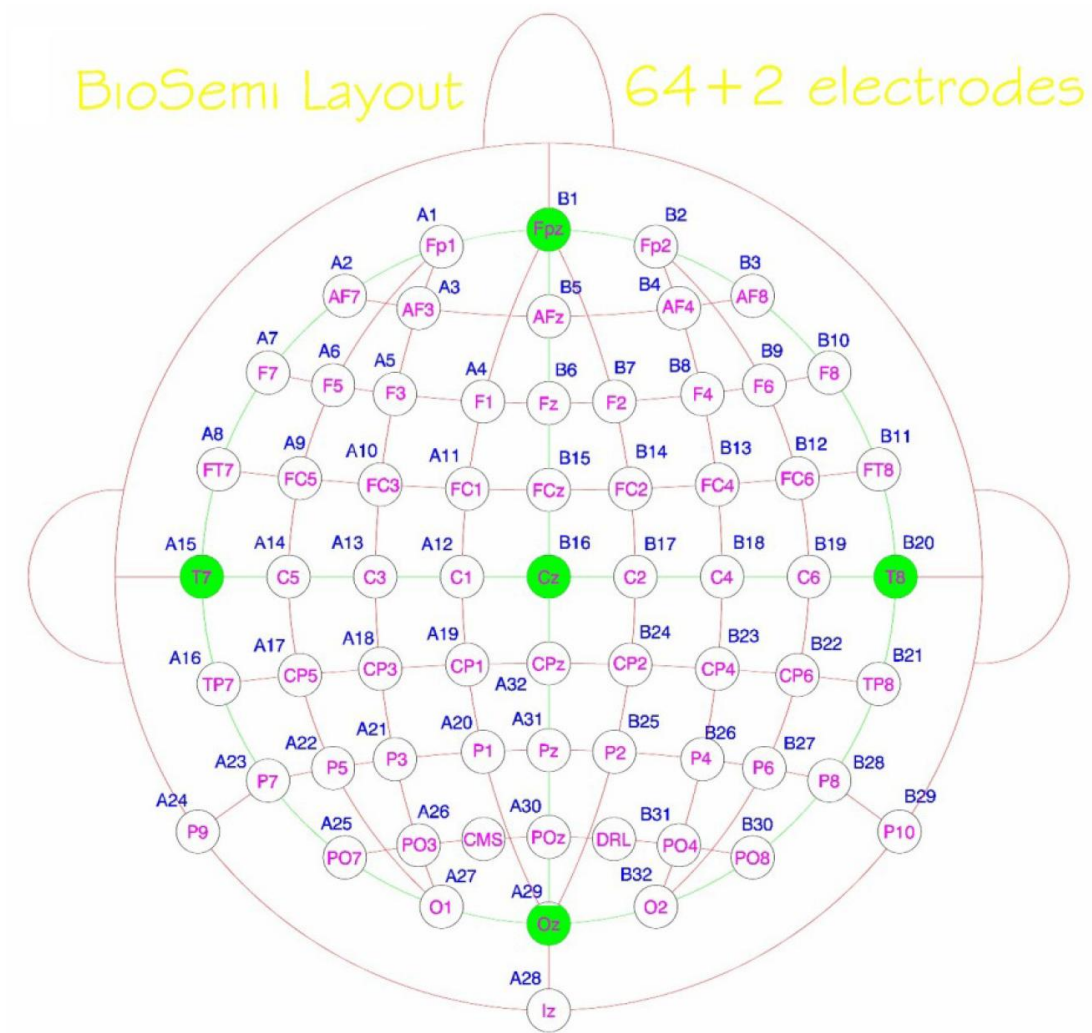


Figura 63: Posiciones de referencia del casco de malla utilizado para las capturas muestrales durante los experimentos

10.2 Minimizar el error mediante LS

Queremos minimizar el error mediante el algoritmo LS, para ello debemos derivar e igual a 0 el error $\varepsilon = E[b-h]^2$ en un conjunto de entrenamiento, donde h se obtiene mediante la expresión anterior, $h = S_t T_t + S_n T_n$

$$\frac{\partial \varepsilon}{\partial t} = 2 * E[b - h] * \frac{\partial h}{\partial t} = 2 * E[b - (S_t^T t + S_n^T n)] * S_t = 0$$

$$S_t b = (S_t^T t + S_n^T n) S_t$$

Haciendo lo mismo para n

$$S_n b = (S_t^T t + S_n^T n) S_n$$

CAPITULO 10: ANEXO

Con lo que se obtiene

$$\begin{pmatrix} S_t S_t^T & S_t S_n^T \\ S_n S_t^T & S_n S_n^T \end{pmatrix} \begin{pmatrix} t \\ n \end{pmatrix} = \begin{pmatrix} S_t b \\ S_n b \end{pmatrix}$$