

A Kernel based approach for classification of electromagnetic interference signals

Ender Luzardo¹, José L. Paredes² y Jaime Ramírez²

¹Postgrado en Automatización e Instrumentación. ²Escuela de Ingeniería Eléctrica.

Universidad de los Andes. Mérida 5101, Venezuela.

Tel: (0274) 240-2903. Fax: (0274) 240 2890. paredesj@ula.ve

Abstract

This paper introduces Electromagnetic Interference signal classification methods for signals obtained on ribbon cables with different crosstalk configurations. The proposed method comprises two stages. The first one is a preprocessing stage that applies either Principal Components Analysis (PCA), Kernel Principal Components Analysis (KPCA) or Independent Components Analysis (ICA) to reduce the data dimension and, at the same time, to obtain the most relevant information from the raw data. The second stage, the classification one, uses Support Vector Machine (SVM) to classify the kind of electromagnetic coupling. We compare the performance of the different classification structures obtained by combining a pre-processing method with SVM, namely PCA+SVM, KPCA+SVM, ICA+SVM as well as SVM in the time domain.

Key words: Electromagnetic interference, principal components analysis, Kernel principal components analysis, independent components analysis, support vector machine.

Clasificación de señales de interferencia electromagnética basada en Kernel

Resumen

En este artículo se propone una herramienta de clasificación de señales de interferencia electromagnética sobre cables planos con distintas configuraciones prácticas de acoplamiento. La técnica propuesta consiste de dos etapas. La primera etapa es un preprocesamiento basado en análisis de componentes principales (PCA), análisis de componentes principales basado en *Kernel* (KPCA) o análisis de componentes independientes (ICA) que intenta reducir la dimensión de los datos de análisis y extraer a su vez la información más relevante; para la segunda etapa de clasificación desarrollada usando máquinas de soporte vectorial (SVM). En este trabajo se realiza un estudio comparativo de los distintos métodos de preprocesamiento y clasificación PCA+SVM, KPCA+SVM, ICA+SVM y SVM en el dominio temporal. También se estudia el desempeño de cada una de las técnicas ante ciertas configuraciones prácticas de acoplamiento por inducción.

Palabras clave: Interferencia electromagnética, análisis de componente principales, análisis de componentes principales basado en Kernel, análisis de componentes independientes, máquinas de soporte vectorial.

1. Introducción

Es bien conocido que los conductores captan con relativa facilidad emisiones de campos

electromagnéticos provenientes de sistemas electrónicos muy cercanos. Estas emisiones de campos electromagnéticos inducen corrientes y voltajes que se pueden mezclar con señales que

llevan información de interés y que actúan como señales de ruido de entrada en cualquier dispositivo receptor, causando severas e inexplicables fallas si los niveles de interferencia exceden el margen de ruido permitidos por el receptor.

En el análisis de señales de Interferencia Electromagnética (*EMI*) la transformada *Wavelet* ha sido recientemente utilizada para el diagnóstico de fallas en líneas de transmisión [1] y para identificar si el acoplamiento entre conductores es más inductivo o capacitivo [2]. Las Redes Neuronales Artificiales, por su parte, han sido utilizadas en la identificación de fuente de campos electromagnéticos sobre configuraciones de tarjetas de circuito impreso [3].

En este trabajo se propone la aplicación de una técnica basada en máquinas de soporte vectorial (SVM, *Support Vector Machine*) [4, 5] para clasificación de señales *EMI* de tipo inductivo o capacitivo en conjunto con técnicas de preprocesamiento basados en análisis de componentes principales (*PCA*, *Principal Component Analysis*) [6-8], análisis de componentes principales basado en *Kernel* (*KPCA*, *Kernel Principal Component Analysis*) [9, 10] y análisis de componentes independientes (*ICA*, *Independent Component Analysis*) [8, 11] para extracción de características y reducción de dimensionalidad [12, 13] de las señales.

Esencialmente, el método propuesto de clasificación de señales de Interferencia Electromagnética realiza un preprocesamiento de señales basado en extracción de características relevantes y reducción de dimensión de la señal perturbadora y posteriormente clasifica el tipo de acoplamiento de Interferencia Electromagnética. Cabe mencionar que el uso de técnicas de clasificación basadas en SVM en el dominio creado por las transformadas PCA, ICA o KPCA no es nuevo y ha sido recientemente reportado en Déniz y col. [14] para reconocimiento de rostros, detección de objetos [15] y extracción de características [16].

2. Interferencia Electromagnética por Diafonía

En situaciones prácticas donde varios conductores comparten la misma trayectoria desde la instrumentación ubicada en campo hasta la

sala de control, o en sistemas que utilizan múltiples conductores, como por ejemplo el cable plano usado en equipos de computación, existen conductores activos transportando pulsos digitales o transiciones lógicas a alta velocidad. Dichas transiciones inducen señales en conductores no activos cercanos debido al acoplamiento electromagnético por diafonía. A fin de simplificar el complejo problema de acoplamiento por diafonía se acostumbra a modelar los conductores como líneas de transmisión con tres conductores paralelos de longitud \mathcal{L} , tal como se muestra en la Figura 1, donde una fuente de voltaje $V_s(t)$ con su respectiva resistencia R_s se conecta a una resistencia de carga R_L por medio de un conductor denominado generador (conductor agresor) G y un conductor de referencia O . En esta figura, un segundo conductor denominado receptor (conductor víctima) R interconecta dos terminales resistivos R_{NE} y R_{FE} cerrando el lazo a través del mismo conductor de referencia. Ambos conductores, G y R , comparten el mismo retorno.

Por el circuito generador circula una corriente I_G a lo largo del conductor G y se presenta un voltaje V_G entre el conductor agresor y la referencia. La corriente y el voltaje asociado al circuito generador crean campos electromagnéticos que se acoplan al circuito receptor induciendo voltajes $V_{NE}(t)$ y $V_{FE}(t)$ en los terminales R_{NE} y R_{FE} del circuito receptor [17].

Debido al efecto de las tensiones y corrientes inducidas se define el acoplamiento por diafonía como la superposición de dos componentes, una debida a la inductancia mutua entre los dos circuitos (acoplamiento inductivo) y la otra debida a la capacitancia mutua entre los dos circuitos (acoplamiento capacitivo), así el voltaje en los extremos del circuito receptor viene dado por [17]:

$$V_{NE}(t) = \left(\frac{R_{NE}}{R_{NE} + R_{FE}} \frac{1}{R_S + R_L} L_m + \frac{R_{NE} R_{FE}}{R_{NE} + R_{FE}} \frac{R_L}{R_S + R_L} C_m \right) \frac{dV_s(t)}{dt} = \left(M_{NE}^{Ind} + M_{NE}^{Cap} \right) \frac{dV_s(t)}{dt}, \quad (1)$$

donde L_m y C_m son la inductancia mutua total y capacitancia mutua total respectivamente, que se obtienen al multiplicar sus respectivos pará-

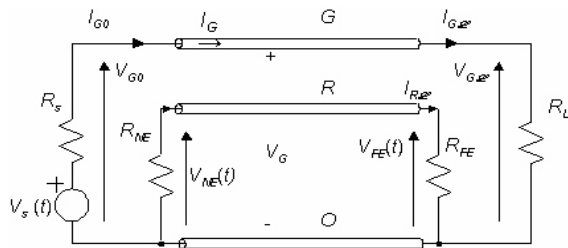


Figura 1. Línea de transmisión de tres conductores.

metros por unidad de longitud por la longitud de la línea de transmisión, y $M_{NE} = M_{NE}^{Ind} + M_{NE}^{Cap}$ es definido como el coeficiente de transferencia de acoplamiento electromagnético.

Del circuito de la Figura 1 y de la Ecuación (1) se espera que la componente del acoplamiento inductivo domine a la componente de acoplamiento capacitivo para bajas impedancias de carga R_L (altas corrientes, bajos voltajes) y que la componente de acoplamiento capacitivo domine a la componente de acoplamiento inductivo para altas impedancias de carga R_L (bajas corrientes, altos voltajes).

En general, especificaciones que permitan caracterizar líneas de transmisión y por ende permitan definir el tipo de acoplamiento inductivo o capacitivo presente en la línea son poco conocidos, por tal motivo se hace necesario recurrir a técnicas de análisis de señales que permitan determinar el tipo de acoplamiento predominante para posteriormente usar la técnica apropiada de mitigación de EMI.

3. Extracción de Características y Reducción de Dimensionalidad de las Señales

En ambientes electromagnéticos donde el funcionamiento de equipos electrónicos producen efectos perturbadores sobre otros equipos, es necesario realizar una correcta y oportuna identificación de las fuentes de interferencia electromagnética. Dicha identificación requiere en principio de un preprocesamiento basado en extracción de características de las señales en estudio, un mecanismo de reducción de dimensionalidad y finalmente un reconocimiento de patrones que

permita identificar el tipo de acoplamiento de interferencia electromagnética.

Entre las técnicas empleadas en este trabajo para preprocesamiento de los datos de acoplamiento por diafonía se encuentran: Análisis de componentes principales (PCA), análisis de componentes principales basado en Kernel (KPCA) y análisis de componentes independientes (ICA), mientras que para la clasificación de los tipos de acoplamiento de EMI las máquinas de soporte vectorial (SVM) son usadas debido a que han dado resultados superiores a las redes neuronales en problemas complejos de clasificación [18].

3.1. Análisis de componentes principales (PCA)

Dado el conjunto de vectores de entrada $X = [x_1, x_2, x_3, \dots, x_N] \mid x_i \in R^n$ definido a partir de las señales acopladas de forma electromagnética al circuito receptor de la Figura 1, PCA transforma linealmente cada uno de los vectores de entrada en nuevos vectores denominados *componentes principales* [7], de la forma

$$Y = U^T X, \tag{2}$$

donde U es una matriz ortogonal de dimensión $n \times n$ formada por los autovectores de la matriz de covarianza muestral $C = \frac{1}{N} \sum_{i=1}^N x_i x_i^T$. En la Ecuación (2), $Y = [y_1, y_2, y_3, \dots, y_N]$ con $y_i \in R^n$ son los nuevos vectores, caracterizados por tener componentes descorrelacionadas.

En general, PCA resuelve el problema de autovalores

$$\lambda_i u_i = C u_i \quad \forall i = 1, 2, \dots, n, \tag{3}$$

donde λ_i es uno de los autovalores de C y $u_i \in R^n$ es el correspondiente auto vector. Debido al carácter unitario de la transformada, los autovalores obtenidos representan la energía de las señales de entrada de interferencia electromagnética y en general son ordenados en forma creciente

$$\lambda_n < \lambda_{n-1} < \dots < \lambda_3 < \lambda_2 < \lambda_1, \tag{4}$$

donde $\lambda_i = 100\%$ de la energía de las señales originales.

PCA tiene la propiedad de compactar energía en pocas componentes, de esta manera, la idea central en PCA radica en reducir la dimensión del conjunto de vectores de entrada minimizando la pérdida de información. Por consiguiente, de $Y = [y_1, y_2, y_3, \dots, y_N]$ se toman solo los m Componentes Principales, donde m se escoge apropiadamente ($1 \leq m \leq N$) usando criterios como: retener componentes principales que expliquen más varianza que cualquier variable original tipificada, retener un número fijo de componentes principales manteniendo un cierto porcentaje de energía de la señal original [6] o determinar un número de componentes tal que se produzcan un error mínimo en un proceso de entrenamiento o validación.

3.2. Análisis de componentes principales no lineal basado en Kernel (KPCA, Kernel Principal Component Analysis)

El objetivo del método PCA no lineal es realizar una transformación de los vectores de entrada x_i a un espacio característico de alta dimensión Φ

$$\Phi: R^n \rightarrow R^h \mid z = \Phi(x_i), \quad (5)$$

con $h > n$, para luego determinar PCA lineal en el nuevo espacio característico [9]. Así, el interés del análisis de componentes principales basado en Kernel es calcular PCA lineal en un dominio diferente al dominio original de entrada donde las variables o características están relacionadas no linealmente con el espacio de entrada por medio de una herramienta matemática denominada Kernel [10]. En general, la función Kernel es una herramienta útil para calcular el producto punto $\Phi(x_i) \cdot \Phi(x_j)$ en espacios de alta dimensión, con lo cual la transformación de x_i a $\Phi(x_i)$ queda implícita en dicha función.

De esta manera, el problema de KPCA se traduce en resolver el problema de autovalores de la matriz Kernel

$$K = k(x_i, x_j) \quad \forall i = 1, 2, \dots, N \quad j = 1, 2, \dots, N \quad (6)$$

y realizar la transformada PCA con los autovectores obtenidos a partir de K . La Tabla 1 especifica algunos ejemplos de Kernel comúnmente usados [4].

Similar a PCA, en KPCA también se puede reducir la dimensión de los datos de entrada, en este caso, considerando solamente los autovalores más elevados y por ende los correspondientes autovectores ya que el criterio de energía utilizado en PCA para reducir dimensión no es válido en KPCA, dado que dicha transformación, por su carácter no lineal, es no unitaria.

3.3. Análisis de componentes independientes (ICA, Independent Component Analysis)

El análisis de componentes independientes fue desarrollado inicialmente para separación ciega de componentes (*blind source separation - BSS*) y más recientemente ha sido utilizado, al igual que PCA y KPCA, para extracción de características [13]. ICA es una técnica que recobra un conjunto de señales independientes S a partir del conjunto de señales medidas X .

$$S = WX, \quad (7)$$

donde $S = [s_1, s_2, s_3, \dots, s_N]$, $s_i \in R^n$ y X la matriz de datos de entrada.

Esta técnica no produce componentes descorrelacionadas, como es el caso de PCA, sino crea componentes con independencia estadística. El análisis de componentes independientes equivale a forzar que las componentes resultantes del proceso de transformación no presenten distribución gaussiana. Las componentes independientes son calculadas optimizando medidas de no gaussianidad por medio de métodos como *Curtosis*, *Negentropía* o *aproximación de la Negentropía* [11].

En general, la información mutua es usada como criterio para estimar s_i como medida natu-

Tabla 1
Funciones Kernel más comunes

Kernel	Lineal	Polinomial	Sigmoide	RBF	ERBF	Multicuadrática inversa
$k(x_i, x_j)$	$\langle x_i \cdot x_j \rangle$	$(\langle x_i \cdot x_j \rangle + \theta)^d$	$\tanh(c \langle x_i \cdot x_j \rangle + \theta)$	$\exp \frac{-\ x_i - x_j\ ^2}{2\sigma^2}$	$\exp \frac{-\ x_i - x_j\ }{2\sigma^2}$	$(\ x_i - x_j\ ^2 + c^2)^{-1/2}$

ral de independencia entre variables aleatorias [8]. Así, minimizar la información mutua entre componentes corresponde a maximizar su negentropía. Sin embargo maximizar la negentropía directamente es complejo de conseguir, por tanto se recurre a maximizar la aproximación de la negentropía, la cual está definida como:

$$J(w) \approx \sum_{i=1}^N k_i [\text{media}(G_i(s_i)) - \text{media}(G_i(r))]^2 \quad | \quad r \in R^n \tag{8}$$

donde $G(s_i)$ es alguna función no cuadrática, r es un vector cuyas componentes sigue una distribución gaussiana de media cero y varianza unitaria, k_i es un valor constante y w_i la i -ésima fila de la matriz de separación W .

El algoritmo desarrollado en Oursland y col. [11] y usado en este trabajo, maximiza $J(w)$ estimando w_i por

$$w_{n+1} = \frac{\text{media}(x \cdot G'(w_n^T x))}{\text{media}(x \cdot G''(w_n^T x)) w_n} \tag{9}$$

hasta que $w^T w$ converja a 1. En la Ecuación (9), G' y G'' son la primera y segunda derivada de la función G respectivamente. Previo a maximizar la aproximación de la negentropía es necesario que los vectores de entrada sean blanqueados, es decir que sus nuevas componentes sean descorrelacionadas y su varianza sea igual a la unidad. El hecho de descorrelacionar componentes implica usar *PCA*. Al usar *PCA* se pueden reducir la dimensión de los datos en *ICA* utilizando los mismos criterios de energía.

4. Clasificación de Señales mediante Máquinas de Soporte Vectorial

4.1. Máquinas de soporte vectorial (SVM, Support Vector Machine)

Las máquinas de soporte vectorial son máquinas de aprendizaje supervisado basados en la teoría de aprendizaje usados en reconocimiento de patrones y en regresión lineal. En este trabajo, SVM considera el problema de clasificación binaria dado por un conjunto de datos de entrenamiento de la forma

$$(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N) \quad | \quad x_i \in R^n, y_i \in \{+1, -1\} \tag{10}$$

donde los x_i son patrones n -dimensional debidamente etiquetados y pertenecientes a la clase y_i . El objetivo es construir un clasificador binario o encontrar una función de decisión, a partir de las muestras de entrenamiento, que genere pequeños errores de clasificación para muestras desconocidas.

4.2. Máquinas de soporte vectorial. Caso linealmente separable

Las SVMs están basadas en las clases de hiperplanos separadores de la forma

$$w^T \cdot x + b = 0 \quad | \quad w \in R^n, b \in R, \tag{11}$$

donde w denota los vectores normales a cada hiperplano y suele denominarse vector de ponderación y b define el umbral de decisión y suele denominarse sesgo (*bias*).

El problema de clasificación se reduce así a un problema de optimización dado por

$$\text{minimizar } \frac{1}{2} \|w\|^2 \tag{12}$$

$$\text{sujeto a: } y_i (w^T \cdot x_i + b) = 1 \quad i = 1, 2, \dots, N, \tag{13}$$

donde $\|w\|$ denota la norma euclidiana de w . Así, el hiperplano óptimo puede ser calculado resolviendo un problema de optimización convexo el cual minimiza una función cuadrática bajo restricciones de desigualdades lineales. La minimización de la Ecuación (12) sujeto a las restricciones de la Ecuación (13) se resuelve con la introducción de multiplicadores de *Lagrange* α_i

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^N \alpha_i (y_i ((w^T \cdot x_i) + b) - 1), \tag{14}$$

donde $\alpha_i \in R$ y $L(w, b, \alpha)$ es minimizada con respecto a las variables principales w y b , y maximizada con respecto a la variable dual α_i . Por consiguiente, el vector w es una combinación lineal de los vectores de entrenamiento, es decir, la solución es un resultado único globalmente optimizado, que presenta la siguiente propiedad

$$w = \sum_{i=1}^N \alpha_i y_i x_i, \quad (15)$$

así para los $\alpha_i > 0$ los correspondientes x_i son llamados vectores de soporte. Otra propiedad de hiperplano clasificadores que necesita ser enfatizada es que tanto el problema de optimización (usado para encontrar el hiperplano óptimo) como la función de decisión (usado para clasificación de vectores) pueden ser expresados en forma dual, por medio de productos punto entre vectores, es decir [5]:

$$f(x) = \text{sgn} \left(\sum_{i=1}^N y_i \alpha_i \langle x^T \cdot x_i \rangle + b \right), \quad (16)$$

donde $\text{sgn}(\cdot)$ es la función signo.

4.3. Máquinas de soporte vectorial.

Caso linealmente no separable

Para el caso linealmente no separable, SVM ejecuta una transformación no lineal de los vectores de entrada x_i a un espacio característico de mayor dimensión, donde dicha transformación es determinada por la función *Kernel*.

Por consiguiente, la función de decisión no lineal que generaliza la clasificación a problemas no lineales puede ser expresada en términos de la función *Kernel* mediante

$$f(x) = \text{sgn} \left(\sum_{i=1}^N y_i \alpha_i k(x, x_i) + b \right), \quad (17)$$

donde $k(x, x_i)$ es la función *Kernel*. De acuerdo a los diferentes problemas de clasificación, las funciones *Kernel* de la Tabla 1 pueden ser seleccionadas para obtener resultados óptimos de clasificación.

4.4. SVM con margen blando (*soft margin SVM*)

En la práctica un hiperplano separador óptimo puede no existir, es decir, no siempre es posible encontrar una transformación de los datos que permita separarlos linealmente, por ejemplo altos niveles de ruido pueden causar solapamiento de las clases, ya sea en el espacio de entrada o en el espacio característico inducido por alguna función *Kernel*.

La idea de SVM con margen blando es introducir variables de holgura (ξ), que permitan a la SVM poder modelar de forma mucho más robusta. La formulación habitual de la SVM con margen blando [7] puede ser expresada a partir de:

$$\text{minimizar } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i$$

sujeto a:

$$y_i ((w \cdot x_i) + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \quad i = 1, 2, \dots, N \quad (18)$$

donde C es un parámetro de aprendizaje, estrictamente positivo, que determina la holgura del margen blando y que establece las variables de holgura diferente de cero.

Similarmente, el problema de clasificación se reduce a resolver un problema de optimización (solución de la Ecuación (18)) donde el objetivo es determinar w , b y ξ_i (con la adicional restricción de $0 < \xi_i < C$) para un conjunto de datos de entrenamiento x_1, x_2, \dots, x_N . Posteriormente, para un vector de entrada desconocido (patrón a clasificar) el clasificador se reduce finalmente a evaluar la Ecuación (17).

Brevemente revisadas algunas técnicas usadas como preprocesamiento para extracción de características y reducción de dimensionalidad así como de clasificación, se define, en la siguiente sección, configuraciones prácticas para adquirir señales de estudio de Interferencia Electromagnética, luego se procede a la aplicación de cada una de las técnicas de preprocesamiento y clasificación antes descritas sobre datos reales adquiridos.

5. Señales de Estudio de Interferencia Electromagnética

En esta investigación, las señales de estudio se obtienen por medio de configuraciones físicas de líneas de transmisión de tres conductores construidas sobre cables planos comúnmente usados en equipos de computación. La Figura 2 muestra las configuraciones prácticas comúnmente utilizadas, donde el conductor (1) es el generador de interferencia y el conductor (2) es el receptor. Ambos conductores comparten un retorno común, R , (tierra). Las señales de interferencia son las que se manifiestan en el conductor (2) y son objeto de estudio posterior.

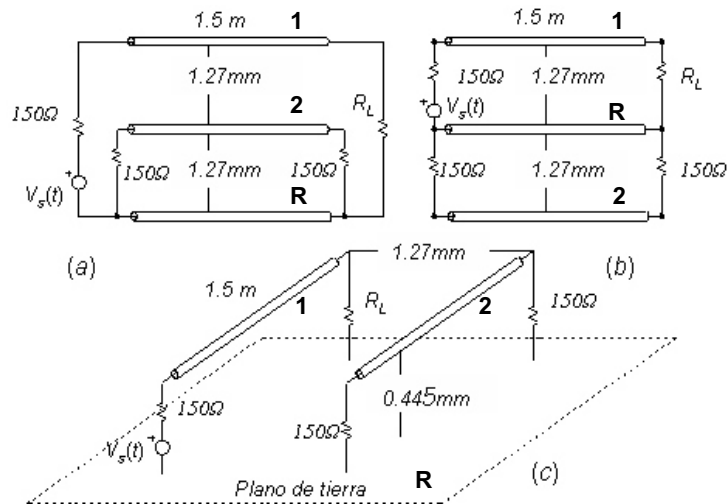


Figura 2. Configuraciones prácticas para una línea de transmisión de tres conductores.

La Figura 2(a) muestra la configuración más desfavorable de acoplamiento de una línea de transmisión de tres conductores. Aquí se utiliza el conductor de retorno, R, alejado del conductor generador de interferencia, observándose un acoplamiento capacitivo mayor debido al acercamiento entre los conductores, y acoplamiento inductivo debido al área elevada del lazo de corriente agresor [17]. La Figura 2(b) muestra una configuración que es usada para atenuar acoplamiento por diafonía, al separar el conductor generador de interferencia y el conductor receptor por medio del conductor de retorno. Aquí se ha disminuido el área del lazo de corriente agresor, disminuyéndose así el acoplamiento inductivo. La Figura 2(c) muestra la configuración más usada en la práctica para atenuación de voltajes acoplados electromagnéticamente. Esta configuración utiliza como retorno, un plano de tierra, diferente de las configuraciones 2(a) y 2(b) que utilizan un conductor como retorno. Dicho plano de tierra ofrece una impedancia mucho menor que los casos anteriores y como consecuencia una disminución de los respectivos acoplamientos [17].

Normalmente se hace necesario computar valores de parámetros por unidad de longitud sobre configuraciones prácticas de líneas de transmisión que permitan definir el tipo de acoplamiento predominante inductivo o capacitivo. En general, dichos parámetros no son todos conocidos con seguridad, y por tanto, no se puede establecer con certeza el tipo de acoplamiento que

predomina. Hyötyniemi [18] presenta valores de parámetros por unidad de longitud para cada configuración de la Figura 2, así como valores de inductancia mutua total y capacitancia mutua total para una línea de transmisión de 1,5 m de longitud.

Regresando a la Ecuación (1), del coeficiente de transferencia de acoplamiento $M_{NE} = M_{NE}^{Ind} + M_{NE}^{Cap}$ se define, a efectos prácticos, una relación de acoplamiento, cr , como el cociente entre el coeficiente de acoplamiento debido al efecto inductivo y el coeficiente de acoplamiento debido al efecto capacitivo $\left(cr = \frac{M_{NE}^{Ind}}{M_{NE}^{Cap}} \right)$. Esta rela-

ción de acoplamiento cr permite establecer un criterio propio para la selección de valores de la resistencia de carga que determine la contribución de cada componente de acoplamiento. Dicho criterio es establecido empíricamente para valores de resistencia de carga donde la relación de acoplamiento cr es mayor que 3.0 o menor que 1/3. Esto permite separar ambas clases de acoplamiento, facilitando posteriormente el proceso de entrenamiento de los clasificadores basado en Kernel (SVM).

Asimismo, en Hyötyniemi [18] se puede encontrar una tabla de valores de resistencia de carga a usar en cada una de las configuraciones prácticas, para generar cada uno de los tipos de acoplamiento de diafonía sobre la línea de transmisión en la adquisición de los datos.

6. Generación de Señales Diafonia para Identificación y Clasificación de Señales EMI

En este trabajo, el análisis y clasificación de señales de Interferencia Electromagnética, es hecho a señales reales adquiridas en el dominio temporal mediante un osciloscopio digital de alta resolución.

La Figura 3 muestra la conexión práctica para la adquisición de los datos concerniente a la configuración *a*, *b* y *c* de la Figura 2. En dicha conexión se considera principalmente el cable plano como línea de transmisión y un tren de pulsos como fuente de excitación $V_s(t)$ con niveles de tensión de 0 y 5 V y frecuencia de trabajo de 1 MHz.

Las Figuras 4(a) y 4(b) muestran las señales de acoplamiento por diafonia de tipo predominantemente inductivo y predominantemente capacitivo, respectivamente, capturadas a partir de la configuración de la Figura 2(a). Estas figuras muestran la gran similitud que tienen dichas señales resultando difícil, por no decir imposible, clasificar determinadas muestras mediante simple inspección visual.

La Figura 5 muestra la densidad espectral de frecuencias de cada una de las señales de la Figura 4, obtenido usando la función *psd* de Matlab. Nuevamente se observa que ambas señales relacionadas con los dos tipos de acoplamiento (inductivo y capacitivo) presentan similares componentes de frecuencia.



Figura 3. Conexión para la adquisición de datos.

El conjunto de datos de diafonia, para aplicar técnicas de identificación y clasificación, consiste de un total de 720 vectores para entrenamiento y 400 observaciones para validación y 400 para prueba. Los vectores de entrenamiento, validación y prueba se generan a partir de señales capturadas en el dominio temporal. Inicialmente se registran 60 señales sobre una ventana fija de adquisición, 30 por cada clase. De las 30 señales, se seleccionan 20 aleatoriamente para entrenamiento y validación y las 10 restantes para prueba. Cada señal capturada es una secuencia de 1000 puntos. De cada señal se escogen aleatoriamente grupos de vectores de 250 muestras de longitud¹ evitando así que la configuración del dispositivo de captura genere siempre la misma secuencia fija y por tanto se aleje de

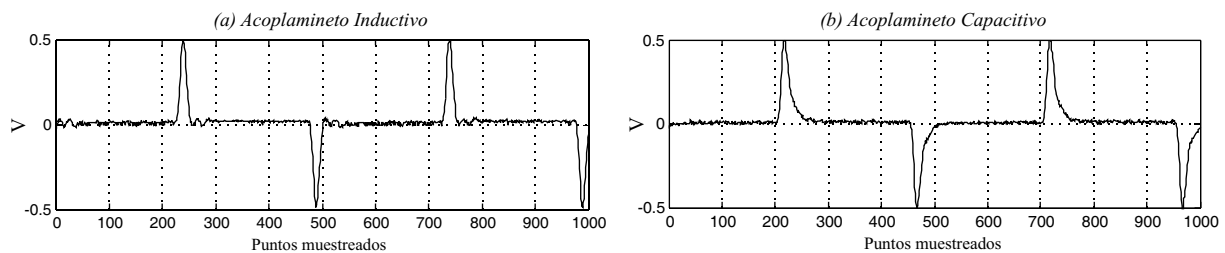


Figura 4. Señales reales de acoplamiento muestreadas a 500 MHz. Configuración Figura 2(a).

1 La longitud mínima de cada observación de entrada (vector de entrenamiento) es fijada mediante ensayo y error en 250 muestras. En general, longitudes menores que 250 no capturan completamente el comportamiento dinámico de los datos y hacen que todos los clasificadores fallen mientras que longitudes mayores a 250 muestras presentan información original redundante que conduce a un elevado tiempo de cómputo.

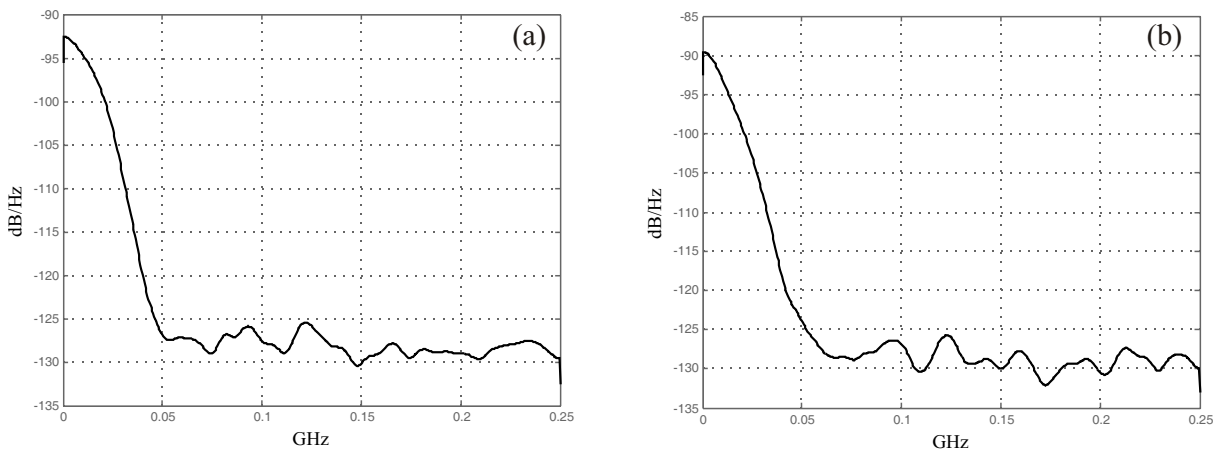


Figura 5. Espectro de frecuencia de las señales de la Figura 4.

una situación más realista de operación. El total de vectores para entrenamiento se forma por la unión de los grupos de vectores obtenidos por cada señal.

El entrenamiento del clasificador se lleva a cabo inicialmente con 80 vectores de entrenamiento y posteriormente se va incrementando hasta alcanzar un total de 720.

En la etapa de preprocesamiento (extracción de características y reducción de dimensión), el algoritmo *PCA* se desarrolla en código *Matlab*[®] según Hyötyniemi [19]. El algoritmo *KPCA* también se desarrolla en *Matlab*[®] a partir de Schölkopf y col. [9, 10]. La función *Kernel* utilizada en el algoritmo *KPCA* es la función de base radial (*RBF*) de la Tabla 1 ya que permite obtener el mejor rendimiento. El mejor valor de σ , seleccionado mediante ensayo y error es 1. La implementación del preprocesamiento *ICA* se define en base al algoritmo *my_ica* desarrollado en Oursland y col. [11] el cual estima la no gaussianidad a través de la aproximación de la *Negentropía*.

En la implementación de las técnicas *PCA*, *KPCA* e *ICA*, uno de los problemas es encontrar el número óptimo de componentes principales. El criterio utilizado para seleccionar componentes principales esta basado en el criterio de energía retenida. Por ejemplo, Antonini y Orlandi [2] toman el 80% de la energía retenida por los coeficientes de descomposición Wavelet Packet que luego sirve de entrada al clasificador. En este trabajo, las componentes principales seleccionadas para *PCA* e *ICA* quedaron determinadas por los

autovalores que retienen el 95% de la energía total de la señal de entrada. Dicho procedimiento fue establecido mediante ensayo y error, desde un total de componentes (100% de la energía) hasta un valor (95%) en el cual no se producen diferencia notables sobre el error de clasificación y se reduce apreciablemente la dimensión de los datos de entrada de 250 a 80 y, por ende, el tiempo de cómputo.

Para el caso de *KPCA*, las componentes en el espacio característico fueron seleccionadas mediante prueba y error, desde un total de componentes hasta sólo aquellas componentes cuyo autovalor en el espacio característico fuese mayor a la unidad, resultando así una reducción de dimensión de los vectores de 250 a 180.

En la etapa de clasificación, el algoritmo de *SVM* es implementado en *Matlab*[®] según Gunn [5]. La función *RBF* es usada como función *Kernel* para la *SVM*. Los valores óptimos de los parámetros de ajuste fueron determinados mediante ensayo y error obteniéndose $\sigma = 1$ y $C = 1$.

7. Análisis y Discusión de Resultados

Definido cada uno de los parámetros de ajuste de las técnicas de preprocesamiento y clasificación, se aplican cada una de las técnicas descritas sobre datos adquiridos usando las configuraciones prácticas de la Figura 2.

A fin de comparar el desempeño de las múltiples técnicas de preprocesamiento, los resulta-

dos obtenidos en el proceso de clasificación mediante *SVM* usando los vectores de entrada en el dominio temporal (*datos crudos + SVM*), usando luego componentes independientes como método de extracción de características (*ICA + SVM*), seguidamente, componentes principales (*PCA + SVM*) y posteriormente, extracción de características con *Kernel PCA (KPCA + SVM)* son mostrados en la Figura 6. Las Figuras 6(a) y 6(b) muestran gráficas de error porcentual de clasificación, en el proceso de validación y prueba respectivamente, en función del número de vectores de entrenamiento para la configuración de la Figura 2(a). En la Figura 6 se observa como el preprocesamiento *PCA* (extracción de características y reducción de dimensionalidad) produce mejores resultados en la clasificación con pocas muestras de entrenamiento. Se observa además la potencialidad del clasificador ya que al realizar clasificación directamente sobre los datos de entrada, denotado en las gráficas como *SVM(t)*, genera un error de clasificación bastante aproximado a la clasificación usando *PCA + SVM*, pero con la desventaja de requerir un tiempo de entrenamiento relativamente alto debido a la dimensión de los datos ($n = 250$).

En la Figura 6 también se observa una mejor clasificación en el procedimiento *PCA + SVM* que en el procedimiento *ICA + SVM*. Se observa cómo la clasificación *KPCA + SVM* es poco confiable para pocas muestras de entrenamiento pero mejora sus resultados de clasificación a medida que aumentan las muestras de entrenamiento, a

tal punto de obtener una perfecta clasificación. El análisis *KPCA* presenta la desventaja de utilizar mayor tiempo de cálculo. Sin embargo, similar al procesamiento *PCA+SVM*, se obtienen clasificación perfecta aproximadamente para la misma cantidad de vectores de entrenamiento.

Para mostrar el desempeño de clasificación de cada una de las técnicas, se adquieren datos de las disposiciones prácticas de la Figura 2(b y c). La Figura 7 presenta el error porcentual de clasificación para los vectores de prueba en función del número de muestras de entrenamiento para las disposiciones prácticas de la Figura 2(b y c).

Sobre datos provenientes de la disposición de la Figura 2(b), la Figura 7(a) muestra nuevamente como el preprocesamiento *PCA* como entrada al clasificador *SVM* genera mejores resultados de clasificación comparado con el preprocesamiento *ICA* a medida que aumentan los vectores de entrenamiento. Análogo a la clasificación hecha para la configuración de la Figura 2(a), *KPCA+SVM* también clasifica perfectamente empleando aproximadamente la misma cantidad de vectores de entrenamiento, pero empleando un elevado tiempo de cómputo.

Con los resultados obtenidos se determina que con sólo remover la información estadística de segundo orden [6, 8] (relacionada con la matriz de covarianza, *PCA*) es suficiente para poder establecer una buena clasificación sobre configuraciones que usan un conductor como retorno. Además, aun cuando *ICA* maneja mayor or-

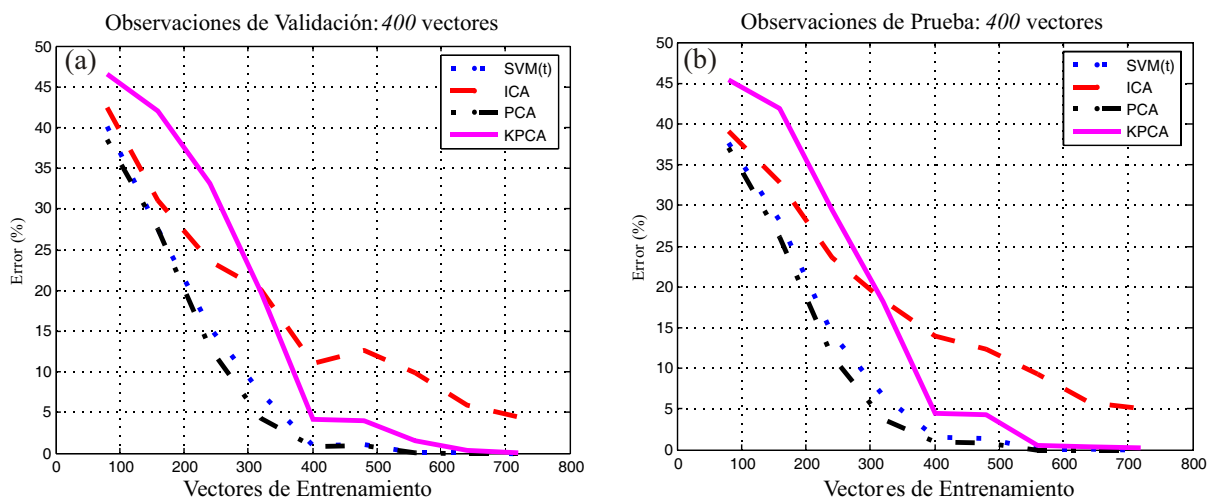


Figura 6. Error de prueba y validación. Configuración 2(a). Observaciones muestreadas a 500 MHz.

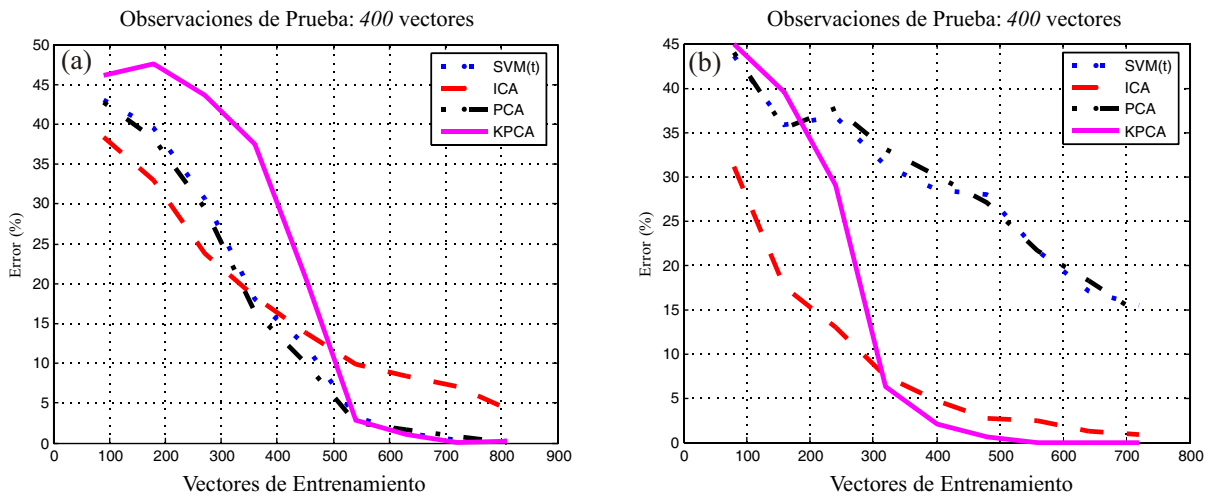


Figura 7. Error de prueba. (a) Con configuración 2(b). (b) Con configuración 2(c). Muestra de 500 MHz.

den de información de los datos que *PCA* a través de la *Negentropía* [8], el hecho de estimar no gaussianidad distorsiona los datos que sirven de entrada a la *SVM* desmejorando considerablemente la clasificación. Por su parte, el análisis *KPCA* maneja mayor orden de información [9, 10] que *PCA* e *ICA* y además remueve correlación de los datos en un espacio de mayor dimensión; por tanto es esperado que el procesamiento realice buena clasificación cuando se proporcionan suficientes datos de entrenamiento.

La Figura 7(b) muestra los resultados de clasificación para señales reales del tipo línea de transmisión con plano de referencia (véase Figura 2(c)). En esta figura se observa como cambia el rendimiento de algunos clasificadores. En este tipo de configuración se generan datos que no presentan una gran autocorrelación temporal tanto en el acoplamiento de tipo predominantemente inductivo como en el de capacitivo, ya que el procesamiento *PCA+SVM* no produce un rendimiento satisfactorio, por el contrario, genera un error de clasificación alrededor del 15% para 720 vectores de entrenamiento. Asimismo, el clasificador como tal, sin previo procesamiento de datos (*datos crudos + SVM*), también genera un error de clasificación de señales de acoplamiento de diafonía relativamente altos, aproximadamente del 15%. Para la configuración con plano de referencia, se establecen los datos como estadísticamente independientes ya que el error de clasificación generado por el procesamiento *ICA + SVM* disminuye significativamente. Para un to-

tal de 720 muestras, el procesamiento *ICA + SVM* nunca alcanza la clasificación perfecta pero mejora considerablemente la clasificación en comparación con el procesamiento *PCA + SVM*. Similar a la clasificación de señales sobre líneas de transmisión con conductor de referencia, el procesamiento *KPCA + SVM* mantiene su comportamiento propio, es decir, sigue siendo poco confiable para pocas muestras en el entrenamiento pero mejora notablemente la clasificación, a tal punto de obtener clasificación perfecta cuando el número de vectores de entrenamiento supera los 550. Esta cantidad de muestras es menor, inclusive, que las requeridas por *PCA+SVM* en el proceso de clasificación sobre las configuraciones que usan conductor como referencia.

Con el análisis hecho, se establece que el único tratamiento de señales confiable que genera resultados de clasificación seguros, independiente de la correlación temporal que presenten los datos de acoplamiento o de la configuración práctica que se defina, es el procesamiento *KPCA+SVM*. Este resultado es de esperar pues, como se mencionó anteriormente, *KPCA* maneja mayor orden de información que *PCA* e *ICA*.

Cabe mencionar que el procedimiento de adquisición de señales sobre cada una de las configuraciones de la Figura 2 fue realizado en repetidas ocasiones a diferentes condiciones de temperatura y humedad y a frecuencias de muestreo del equipo de adquisición de 200 MHz y 500 MHz, obteniéndose resultados similares.

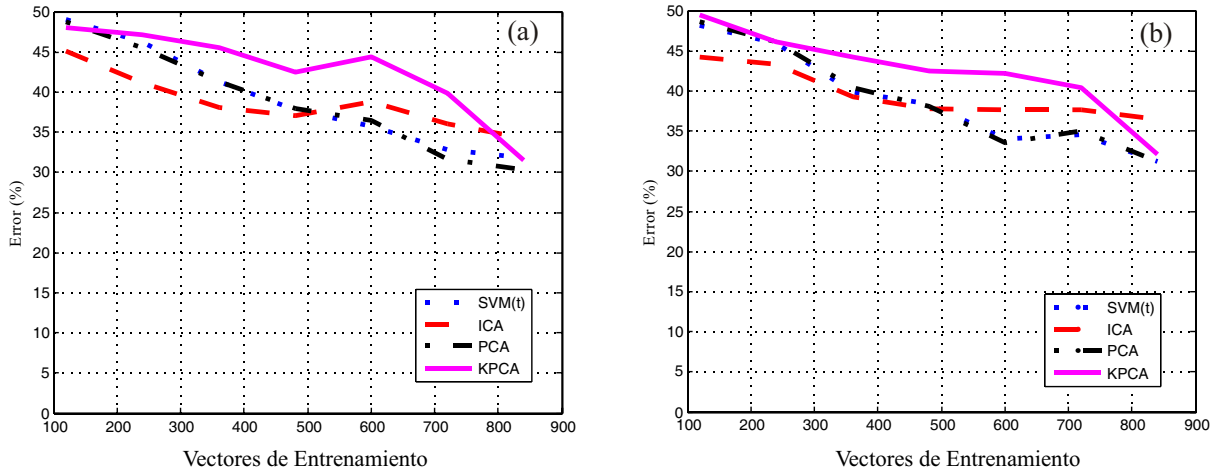


Figura 8. (a) Error de validación y (b) Error de prueba en la generalización del clasificador.

Para concluir, se realiza un análisis de la capacidad de generalización de las diferentes técnicas de preprocesamiento y clasificación. La idea se desarrolla estableciendo un universo de vectores de entrada para el proceso de entrenamiento formado por datos provenientes de cada una de las configuraciones de la Figura 2; posteriormente se desea observar el desempeño del clasificador $KPCA+SVM$ ante datos desconocidos, derivados de cualquiera de las disposiciones prácticas utilizadas en el entrenamiento. En la Figura 8 se muestra el error de validación y prueba para la generalización del clasificador. En esta figura se observa cómo disminuye el desempeño de clasificación con cada una de las técnicas de preprocesamiento. Para esta última prueba, ya no se obtiene una correcta clasificación utilizando $KPCA + SVM$; sino errores de clasificación alrededor del 30%, pero se observa una disminución consistente del error porcentual de clasificación al incrementarse el número de muestras de entrenamiento, repercutiendo así en el tiempo utilizado durante el entrenamiento. Se observa así, que el entrenar con todas las configuraciones a la vez no conduce a resultados aceptables.

8. Conclusiones

En este trabajo se ha investigado el problema de clasificación de señales de interferencia electromagnética por acoplamiento de diafonía de tipo inductivo y de tipo capacitivo. En el análisis de señales EMI se observa que el único trata-

miento de señales que genera buenos resultados de clasificación acerca del tipo de acoplamiento inductivo o capacitivo e independiente de la configuración práctica es el procesamiento $KPCA + SVM$. La razón de estos resultados se atribuye al hecho de que el preprocesamiento $KPCA$ puede explorar más alto orden de información que ICA (*Negentropia*) y que PCA (información de segundo orden - matriz de covarianza), como entrada al SVM . La desventaja que presenta es el elevado tiempo de cómputo utilizado para el entrenamiento.

9. Agradecimientos

Parte de esta investigación ha sido realizada gracias al apoyo del Laboratorio de Investigación y Desarrollo en Automatización e Instrumentación (LABIDAI) de la Escuela de Ingeniería Eléctrica, de la Universidad de Los Andes, y ha sido parcialmente financiada por el Consejo de Desarrollo Científico, Humanístico y Tecnológico (CDCHT-ULA) bajo el proyecto I-876-05-02-EM y el Fondo Nacional de Ciencia, Tecnología e Investigación (FONACIT) bajo el proyecto G-2005000342.

Referencias

1. C Buccella, et al. "Diagnosing Transmission Line Termination Faults by Means of Wavelet Based Crosstalk Signature Recognition" IEEE Transactions on Component and Pack-

- aging Technology, vol. 23, N°. 1, March 2000. pp 165-170.
2. Giulio Antonini, Antonio Orlandi. "Wavelet Packet-Based EMI Signal Processing and Source Identification. IEEE Transactions on Electromagnetic Compatibility, vol. 43, N° 2, May 2001. pp 140-148.
 3. M. D'Amore, A. Moriello, M.S. Sarto "A Neural Network Approach for Identification of EM Field Sources: Analysis of PCB Configurations". IEEE International Symposium on Electromagnetic Compatibility. Vol 2. N° 24-28 August 1998. pp 664-669
 4. Schölkopf and A.J. Smola. "A short introduction to learning with kernels". In S. Mendelson and A. J. Smola, editors, Machine Learning, Proceedings of the Summer School, Australian National University. Springer 2003. pp 41-64.
 5. Steve R. Gunn. "Support Vector Machines for Classification and Regression". Technical Report. May 1998. <http://www.kernel-machines.org/> Consultada: Enero - Marzo de 2004.
 6. Antonio Morillas. "Introducción Al Análisis De Datos Multivariantes" <http://campusvirtual.uma.es/morillas/PresentacionAM.pdf>. Consultada: Septiembre de 2004.
 7. José Hernández Orallo, Ma José Ramírez Quintana, Cesar Ferri Ramírez. "Introducción a la Minería de Datos". Pearson Educación, S.A. Madrid 2004.
 8. Aapo Hyvärinen, Juha Karhunen, Erkki Oja. "Independent Component Analysis". John Wiley & Sons, 2001.
 9. Bernhard Schölkopf, Alexander Smola, Klaus-Robert Müller. "Nonlinear Component Analysis as a Kernel Eigenvalue Problem". Neural Computation. Vol. 10, N° 5. 1998. pp 1299-1319.
 10. Schölkopf, A. Smola, and K.-R. Müller "Kernel principal component analysis". In B. Schölkopf, C. J. C. Burges, and A. J. Smola, editors, Advances in Kernel Methods—Support Vector Learning. MIT Press, Cambridge, MA, 1999. pp 327-352.
 11. Alan Oursland, Judah De Paula, Nasim Mahmood. "Cases studies of Independent Component Analysis". <http://www.oursland.net/tutorials/ica/> Consultada: Abril de 2004.
 12. Jeft Fortuna, David Capson. "Improved support vector classification using PCA and ICA feature space modification" Pattern Recognition. N° 37. 2004. pp 1117-1129
 13. L.J. Cao et al. "A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine" Neurocomputing. N° 55. 2003. pp 321-336
 14. O. Déniz, M. Castrillón y M. Hernández. "Face recognition using independent component analysis and support vector machines" Pattern Recognition Letters. Vol. 24. N° 13. Sept. 2003. pp. 2153-2157.
 15. R. Santiago-Mozos, J.M. Leiva-Murillo, F. Perez-Cruz, A. Artes-Rodriguez, "Supervised-PCA and SVM Classifiers for Object Detection in Infrared Images", IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS'03), 2003. p. 122.
 16. Cao, L.J. y Chong, W.K. Feature extraction in support vector machine: a comparison of PCA, XPCA and ICA Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP '02. Vol. 2, Nov. 2002. pp1001-1005.
 17. Clayton R. Paul. "Introduction to Electromagnetic Compatibility". Wiley Series in Microwave and Optical Engineering. First Edition - April 1992.
 18. Ender Luzardo. "Clasificación de Señales de Interferencia Electromagnética Basada en Kernel" Proyecto de Postgrado. Postgrado en Automatización e Instrumentación. ULA. Octubre de 2004.
 19. Heikki Hyötyniemi. "Regression Toolbox for Matlab". http://saato014.hut.fi/hyotyniemi/publications/01_report125/RegrToolbox/. Version 1.0. Feb. 20, 2001. Consultada: Mayo de 2004.

Recibido el 05 de Diciembre de 2005

En forma revisada el 12 de Febrero de 2007