

Un enfoque para la detección y el diagnóstico de fallas en la instrumentación de un proceso usando reconocimiento de patrones en el dominio wavelet

An approach for fault detection and diagnosis in the instrumentation process using patterns recognition in the wavelet domain

Guillén, Marcos^{1*}; Paredes, José² y Camacho, Oscar²

¹Escuela de Ingeniería Mecánica, Facultad de Ingeniería, ULA

²Escuela de Ingeniería Eléctrica, Facultad de Ingeniería, ULA

Mérida 5101, Venezuela

marcosg@ula.ve

Recibido: 05-06-2009

Revisado: 14-04-2010

Resumen

En este trabajo se presenta un enfoque para detectar y diagnosticar fallas en la instrumentación de un proceso usando la Transformada Wavelet. El método propuesto es usado para detectar y diagnosticar fallas en la instrumentación de un proceso químico (válvulas y transmisores), pero puede ser extendido a otros tipos de procesos. Se presenta un esquema basado en el análisis en línea de la señal de salida del transmisor, utilizando para este fin un análisis de la señal de falla (salida del transmisor) en el dominio wavelet, tanto la detección como el diagnóstico de la falla se realizan en el mencionado dominio. Un algoritmo de extracción de características seguido de un algoritmo de reconocimiento de patrones realizado sobre los coeficientes wavelet permite detectar y clasificar las fallas. El método propuesto tiene capacidad de detectar y diagnosticar cuatro diferentes tipos de fallas de medida en un instrumento, específicamente: falla en el span, en la zona muerta del span, en la calibración de cero y en la zona muerta de cero. Además, habiendo definido el tipo de falla y el componente que está fallando, se desarrollan dos métodos para estimar el porcentaje de desajuste que se está presentando en la válvula o transmisor.

Palabras clave: Fallas, válvulas, transmisores, reconocimiento de patrones, transformada wavelet.

Abstract

This paper presents an approach for fault detection and diagnosis based on the wavelet transform. The proposed method is used to detect faults in chemical process instrumentation (valves and transmitters), but it can be extended to other kinds of processes. The proposed approach applies the wavelet transform on the online-monitored signal to map it into the wavelet domain. A feature extraction algorithm followed by a pattern recognition algorithm performed on the wavelet coefficients allows us to detect and classify faults present in the process instrumentation (valve or transmitter). Furthermore, having defined the faulty component and the kind of fault, two methods to estimate the percentage of breakdown are developed. Thus, the approach not only detects and classifies the fault but also yields an estimate of how serious the problem is.

Key words: Fault, valve, transmitters, pattern recognition, wavelet transform.

1 Introducción

La detección y el diagnóstico de fallas en los procesos de ingeniería son de gran importancia, ya que una temprana

detección de la ocurrencia de una falla podría evitar el deterioro de la producción, degradación de la eficiencia de los procesos, daños mayores en la maquinaria implicada en el proceso y operaciones inseguras de la planta que podrían

poner en riesgo al personal que participa en las operaciones del proceso. El diagnóstico rápido y correcto del componente que está fallando facilita la toma de decisiones, las acciones correctivas y por supuesto las reparaciones que sea necesario realizar (Gertler, 1998).

Los métodos para la detección y diagnóstico de fallas se pueden clasificar en general dentro de dos grupos, los que utilizan un modelo matemático de la planta y los que no utilizan un modelo matemático. Los métodos basados en modelo utilizan una representación matemática del proceso basada en ecuaciones que describen el comportamiento del proceso. Los métodos libres de modelo no hacen uso de modelos matemáticos de la planta (Gertler, 1998). Estos métodos usan redundancias físicas, sensores especiales, chequeo de límites y análisis espectral. El análisis espectral, es un método utilizado frecuentemente a nivel industrial y es el que se usa en este trabajo. Muchas de las variables de la planta presentan una dinámica de cambios lentos a los cuales se les asocia representaciones frecuenciales de baja frecuencia. Dichas representaciones, exhiben espectros que caracterizan el comportamiento de las variables en condiciones normales de operación; alguna desviación de este espectro es considerada como una indicación de anomalía (falla) (Gertler, 1998). Así las fallas pueden ser estudiadas realizando un análisis frecuencial de dichas señales (fallas). Las herramientas utilizadas en este caso son herramientas matemáticas tales como la Transformada de Fourier (TF) y la Transformada de Fourier en tiempo limitado (STFT). Sin embargo, la TF no es adecuada para analizar fallas presentes en la instrumentación de un proceso ya que estas presentan características no estacionarias, transitorios y/o cambios abruptos que exhiben patrones de frecuencia variables en el tiempo. Adicionalmente, se pierde información temporal relacionada a la localización en el tiempo de la falla. Surge entonces, la necesidad de analizar los fenómenos tanto **localizados en tiempo como en frecuencia**, es decir, fenómenos que requieren información conjunta de ambos dominios.

En este trabajo se propone un método basado en la Transformada Wavelet para detectar y diagnosticar fallas que se puedan presentar en los instrumentos de un proceso químico concretamente válvulas y transmisores. Una aplicación de la TW en el contexto de fallas es desarrollada en (Gaouda, 2002), en este trabajo se describe un método novedoso para detectar y diagnosticar cambios abruptos en sensores. El uso de la TW permite localizar con precisión las características de una señal en ambos dominios, identificándose el instante de ocurrencia de una salida anormal del proceso usando análisis multiresolución de la señal. Después que este instante es identificado se extraen las características más relevantes de las señales de fallas en el dominio Wavelet permitiendo así clasificar las fallas usando técnicas de reconocimiento de patrones. Adicionalmente, la TW permite eliminar el ruido de una señal sin apreciable degradación de la misma (Taiwen, 2003), este aspecto es importante para nuestra aplicación ya que las señales producidas

por las fallas en los instrumentos contienen un cierto ruido de fondo que hace que se solapen las fallas y, por lo tanto, sea más difícil su clasificación.

2 Marco teórico

2.1 Fallas en un instrumento

Un instrumento se considera libre de falla si la relación entre los valores reales de la variable comprendidos dentro del campo de medida y los valores de lectura del instrumento es lineal o aproximadamente lineal de acuerdo a las especificaciones del fabricante (Tarantino, 2000; Szigeti, 1998). En la Fig. 1(a) la recta denotada con el número (1) expresa esta relación. El modelado matemático que describe las fallas en los instrumentos (válvulas, transmisores, sensores) se puede representar a través de la ecuación diferencial (1).

$$\tau_v \frac{dq(t)}{dt} + q(t) = \left(\frac{Q_{\max} - Q_{\min}}{m_{\max} - m_{\min}} \right) (m(t) - m_{\min}) + Q_{\min} \quad (1)$$

donde Q_{\max} y Q_{\min} son los flujos máximos y mínimos permitidos para las señales de control m_{\max} y m_{\min} respectivamente, τ_v es la constante de tiempo del actuador de la válvula. Conforme se vayan presentando desviaciones de estos parámetros se tendrán desviaciones de la recta valor real vs. valor medido de un instrumento específico lo cual se traduce en condiciones de fallas de medida del instrumento. Así, las fallas de medida del instrumento pueden ser de cuatro tipos: Falla en el span producida por desviaciones del parámetro Q_{\max} , Falla en la zona Muerta del span ocasionado por desviaciones del parámetro m_{\max} , Falla en la calibración de cero originada por descalibración del parámetro Q_{\min} , y Falla en la zona Muerta de cero causada por desviaciones del parámetro m_{\min} (Guillen, 2004). Las Figs. 1a, 1b, 1c, 1d ilustran el efecto de estas fallas sobre la curva de calibración del instrumento y las Figs. 1e, 1f, 1g, 1h el efecto sobre la señal a la salida del transmisor. Aunque (1) describe el comportamiento para una falla en una válvula, las fallas en los transmisores siguen un modelado similar.

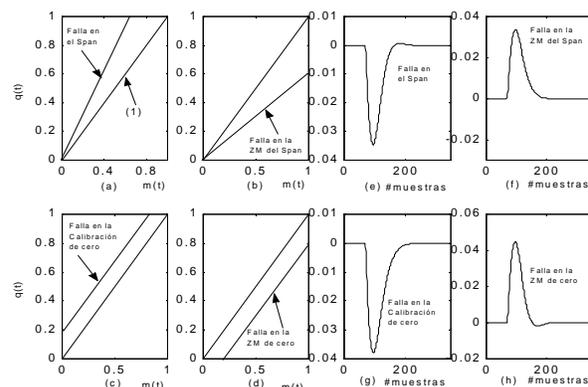


Fig. 1. Fallas en el actuador de la válvula

2.2 La transformada discreta wavelet (TDW)

Considere la señal $x(t) \in L^2(\mathbb{R})$, donde L^2 representa el espacio de todas las funciones integrables al cuadrado (es decir si $x(t) \in L^2 \rightarrow \int_{-\infty}^{+\infty} x(t)^2 dt < \infty$), $x(t)$ puede ser escrita como una serie en expansión en términos de las funciones de escala $\varphi_k(t)$ y wavelet $\psi_{j,k}(t)$ (Burrus, 1989).

$$x(t) = \sum_k c(k)\varphi_k(t) + \sum_k \sum_j d(j,k)\psi_{j,k}(t) \quad (2)$$

donde φ_k y $\psi_{j,k}$ son familias de funciones ortogonales generadas al desplazar y escalar las funciones básicas de escalamiento $\varphi(t)$ y wavelet $\psi(t)$, es decir $\varphi_k = \varphi(t-k)$ y $\psi_{j,k} = 2^{j/2}\psi(2^j t - k)$. Cuando el índice k cambia de posición, la wavelet se mueve sobre el eje del tiempo permitiendo la localización en tiempo del evento. Cuando el índice j cambia, la forma de la wavelet cambia en escala, es decir, se expande o comprime en tiempo, permitiendo así representar detalles de la señal a diferentes escalas. En la ecuación 2, $c(k)$ denota los coeficientes de aproximación y representan una medida del grado de similitud de la señal $\varphi(t)$ localizada en el instante k con la señal $x(t)$. Igualmente, $d(j,k)$ denotan los coeficientes de detalles y representan la correlación entre la señal $x(t)$ y la wavelet a la escala j y en la posición temporal k . Esta similitud es expresada matemáticamente como:

$$c(k) = \int_{-\infty}^{+\infty} x(t)\varphi(t-k)dt \quad (3)$$

$$d(j,k) = \int_{-\infty}^{+\infty} x(t)\psi(2^j t - k)2^{j/2}dt$$

dado que k y j pueden tomar cualquier valor real, la representación de la señal $x(t)$ en el dominio wavelet ($c(k)$ y $d(j,k)$) es redundante. Así como la señal $x(t)$ puede representarse por su versión muestreada, $x(n)$, sin que se pierda información de la señal continua original, los coeficientes wavelet pueden igualmente ser muestreados sin que se pierda información, dando origen así a la transformada discreta wavelet (TDW). Mallat (Mallat, 1989) desarrollo un algoritmo que permite calcular la transformada discreta wavelet a partir de operaciones de filtrado paso alto y paso bajo, donde los coeficientes c y d se obtienen como sigue:

$$c(k) = \sum_m h(m-2k)x(m) \text{ Coeficientes de aproximación} \quad (4)$$

$$d(k) = \sum_m g(m-2k)x(m) \text{ Coeficientes de detalles} \quad (5)$$

donde $h(n)$ y $g(n)$ son los coeficientes de los filtros paso bajo y paso alto respectivamente relacionados con las fun-

ciones escalar y wavelet (Mallat, 1989). Así la TDW analiza la señal $x(n)$ a diferentes resoluciones, descomponiendo la señal en dos partes, una que representa la información de aproximación (baja frecuencias) y otra parte que representa la información de detalles de la señal $x(n)$ (alta frecuencias).

2.3 Detección de la condición de falla

Sin pérdida de generalidad, se puede asumir que la salida del transmisor ilustrada en la Fig. 2 viene dada por

$$y(t) = x(t) + \eta(t) + (1_i)^T \mathbf{F}(t, p) \quad (6)$$

donde $y(t)$ es la señal de salida del transmisor, $x(t)$ es la señal de referencia o set point, $\eta(t)$ representa ruido de fondo el cual se asumirá como ruido blanco Gaussiano con media cero y varianza σ^2 . El termino $(1_i)^T \mathbf{F}(t, p)$ representa uno de los posibles desajustes (fallas o perturbaciones) que se pueden presentar en la planta, donde $(1_i)^T$ es un vector n -dimensional cuyas componentes son ceros excepto posiblemente la i -enésima componente la cual representa la i -enésima condición de desajuste. El termino $\mathbf{F}(t, p)$ es una función vectorial que constituye el patrón de las múltiples fallas. El parámetro "p" representa el porcentaje de desajuste que se está presentando en la planta, $\mathbf{F}(t, p)$ puede ser escrita como:

$$\mathbf{F}(t, p) = [F_1(t, p), F_2(t, p), F_3(t, p), \dots, F_n(t, p)]^T \quad (7)$$

Los desajustes (fallas) se manifiestan como cambios en la señal $y(t)$ y por consiguiente como cambios en el proceso de medición del transmisor (Jian, 2003). Cuando dichos cambios se presentan generan un cambio en la señal de salida del error $e(t)$, definido como:

$$e(t) = y(t) - x(t) \quad (8)$$

teniendo un valor de aproximadamente cero¹ cuando la variable controlada $y(t)$ coincide con el valor de referencia $x(t)$, es decir, cuando el proceso se encuentra en condiciones normales de operación (no existe condición de falla). Sin embargo, si existe un desajuste (falla o perturbación) la variable controlada $y(t)$ se aleja del valor de referencia y el error quedaría como sigue:

$$e(t) = \eta(t) + (1_i)^T \mathbf{F}(t, p) \quad (9)$$

por lo tanto realizando un monitoreo de la señal del error se detectan los cambios en la señal $y(t)$ y, por ende, se pueden

¹ No es cero debido al ruido de fondo que se está considerando.

detectar posibles desajustes, los cuales podrían estar siendo provocados por una posible falla o una perturbación. Así, la detección del cambio en $y(t)$ puede ser pensado como un problema de regulación, donde por un monitoreo constante de la señal de error, éste dará una indicación del desajuste en el proceso bajo estudio. En este trabajo la Integral del Valor Absoluto del Error (*IAE*) es usado como un indicador de la condición de falla. El *IAE* se define como:

$$IAE = \int_{t_0}^t |e(t)| dt \quad (10)$$

donde la integral se extiende desde el momento en que ocurre la perturbación (t_0) hasta un tiempo posterior (t), el cual debe ser lo suficientemente grande para cubrir el tiempo de duración de la falla. El valor de *IAE* para condiciones normales de operación debería ser cero, sin embargo debido al ruido de fondo que contiene la señal $y(t)$ dicho valor es diferente de cero y dependiente del nivel de ruido de fondo. Con el fin de dar una mayor robustez al proceso de detección de la falla, en este trabajo se realiza la estimación del *IAE* en el dominio wavelet, es decir, se aprovecha la capacidad que tiene la transformada discreta wavelet (TDW) para suprimir el ruido que contamina la señal en estudio y así poder tener un *IAE* fijo para el caso en que el proceso se encuentre en estado estacionario evitando así, las variaciones del *IAE* debido al ruido de fondo. Expresando el error $e(t)$ en el dominio wavelet, el *IAE* se reduce a:

$$IAE = \int_{t_0}^t |e(t)| dt = \int_{t_0}^t \left| \sum_k c(k) \varphi_k(t) + \sum_k \sum_j d(j,k) \psi_{j,k}(t) \right| dt \quad (11)$$

Así, despreciando los coeficientes de detalles (componentes ruidosas de la señal) y tomando sólo los coeficientes de aproximación, entonces el *IAE* quedaría como:

$$IAE = \int_{t_0}^t \left| \sum_k c(k) \varphi_k(t) \right| dt \quad (12)$$

El *IAE* en condiciones normales de operación, estimado en el dominio wavelet, presenta un valor muy cercano a cero debido a que la variable controlada se encuentra en el umbral de referencia. Si el índice *IAE* supera cierto valor umbral, la variable controlada se está alejando de la referencia, lo cual indica que algo no deseado está ocurriendo en el proceso, por lo tanto el sistema presenta un desajuste originado por alguna perturbación o falla. Note que en (12) el cálculo del *IAE* toma en cuenta sólo los coeficientes de aproximación del análisis multiresolución, en la práctica dicho cálculo se simplifica aproximando el $IAE = \sum |c(k)|$. Así, se presenta una condición de falla si $\sum |c(k)|$ es mayor o igual a un cierto valor umbral el cual es determinado empíricamente por medio de los históricos de la planta.

3 Descripción del método propuesto para el diagnóstico de la falla

Definido el método que permite detectar la existencia o no de la condición de falla se procede a realizar el diagnóstico de la misma, llevando a cabo las siguientes tareas:

- Aplicación de la Transformada Discreta Wavelet (TDW) sobre la señal $y(t)$.
- Extracción de los parámetros que caracterizan la falla.
- Reconocimiento y Clasificación.
- Determinación del porcentaje de desajuste.

La Fig. 2 ilustra este procedimiento. La wavelet usada en esta aplicación es db4 y se realizan 6 niveles de descomposición.

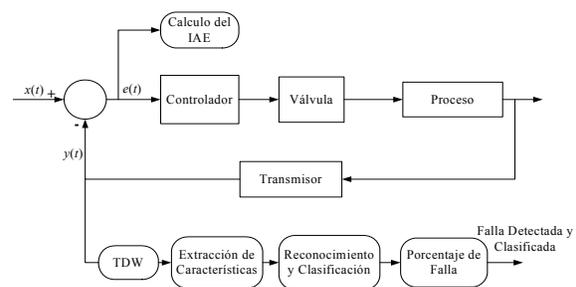


Fig. 2. Tareas implicadas en el diagnóstico de la falla

3.1 Extracción de características

La extracción de características consiste en extraer de la señal de falla o perturbación, representada en el dominio wavelet, características relevantes que permitan diferenciar una señal de falla de cierta clase de otra señal de falla, por ejemplo diferenciar una falla en el actuador de la válvula, de una falla el transmisor. El vector característico selecciona aquellos coeficientes que aporten un porcentaje de energía significativo del total de la energía de la señal y se desprecian aquellos donde se encuentre concentrada la mayor cantidad de ruido. El ruido se encuentra almacenado en mayor grado en los coeficientes de más alta frecuencia; estos coeficientes son $cd1$ y $cd2$, y va decreciendo conforme disminuye la frecuencia, específicamente desde $cd3$ hasta $cd6$. Resultados empíricos mostraron que el mejor vector característico para esta aplicación en particular resulta: $V_f = [ca6 \ cd3 \ cd4 \ cd5 \ cd6]$, pues permite obtener resultados satisfactorios en la clasificación de las fallas manteniendo baja la complejidad y el tiempo de procesamiento. En (Guillén, 2004) se detalla el procedimiento seguido para la selección del mejor vector característico.

3.2 Generación de patrones

Es necesario generar patrones para cada una de los tipos de fallas que se puedan presentar en la instrumentación. Estos patrones son los que van a permitir reconocer la falla

que se esté presentando. Los vectores característicos de dichos patrones de fallas son almacenados en una base de datos. Para cada falla el porcentaje de desajuste se vario desde 7.5%, a 17.5% en intervalos de 2.5%.

3.3 Reconocimiento y clasificación

El reconocimiento y clasificación de la falla se puede tratar con técnicas asociadas al reconocimiento de patrones. En la literatura pueden encontrarse distintas expresiones para medir el grado de similitud pero las más empleadas son medidas de distancias, y entre estas unas de las más utilizadas son la distancia Euclideana y la distancia entre vecinos más cercanos mejor conocida como K vecinos más cercanos (Fu, 1982).

3.3.1 Clasificador basado en la distancia Euclideana

Un patrón desconocido V_f (falla cualquiera) puede considerarse que pertenece a una cierta clase de falla, si y solo si, la distancia entre V_f y un prototipo V_{f_v} (falla en el actuador de la válvula) es menor que la distancia entre V_f y otro prototipo, por ejemplo V_{f_T} (falla de trasmisor). La distancia Euclideana se calcula como sigue (Gaouda, 2002):

$$d_E = \sqrt{[(V_{f_v} - V_f)(V_{f_v} - V_f)^T]} = \left(\sum (V_{f_v} - V_f)^2 \right)^{1/2} \quad (13)$$

En este trabajo se calcula la distancia Euclideana entre el vector característico de la falla desconocida y cada uno de los vectores característicos de los centroides de las clases de fallas que se encuentran en la base de datos, definiéndose el centroide como el valor promedio de vectores característicos pertenecientes a la misma clase.

3.3.2 K vecinos más cercanos

Sea R_1, R_2, \dots, R_j los j grupos de vectores característicos asociados con las clases S_1, S_2, \dots, S_j , respectivamente, es decir, el grupo R_1 pertenece a la clase S_1 , R_2 pertenece a la clase S_2 y así sucesivamente. También denótense como $R_j^{(k)}$ a cada vector característicos perteneciente a cada uno de los grupos R_j , es decir, $R_j^{(k)} \in R_j$ con $k = 1, \dots, u_j$, donde u_j es el número de vectores característicos en el grupo R_j . Se define la distancia entre un patrón de entrada desconocido V_f y la clase S_j como:

$$d(V_f, S_j) = \min_{k=1, \dots, u_j} \left[(V_f - R_j^{(k)})(V_f - R_j^{(k)})^T \right]^{1/2} \quad (14)$$

donde la distancia entre V_f y S_j , es la distancia más pequeña entre V_f y cada vector en R_j , la ecuación (14) arroja así j distancias. El clasificador asignara el patrón desconocido a la clase a la cual pertenezca el vector $R_j^{(k)}$ para el cual se consiguió la mínima distancia (Fu, 1982).

4 Simulación y resultados

El proceso sobre el cual se realiza el estudio es un tanque de precalentamiento cuyas especificaciones están dadas en (Camacho, 2000).

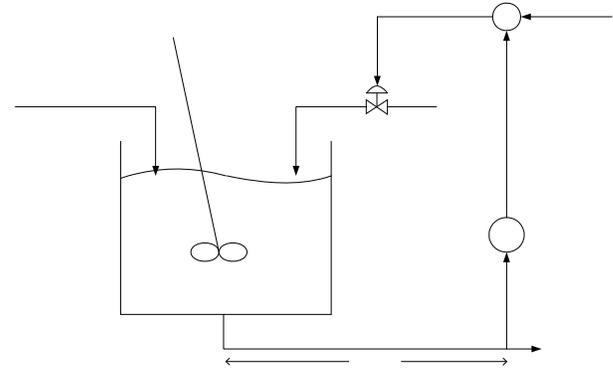


Fig. 3. Tanque de mezclado

4.1 Descripción del proceso industrial

El proceso consiste en un tanque que recibe un flujo caliente, $Q_1(t)$, y otro flujo frío, $Q_2(t)$ los cuales son mezclados para obtener una temperatura $T_3(t)$ deseada a la salida. La variable manipulada es el flujo de agua fría Q_2 , para mantener la temperatura de salida T_3 en el nivel establecido. La temperatura de salida es medida a 38,1 m del tanque aguas abajo. La Fig. 3 muestra el proceso sobre el cual se realizaron las simulaciones. Se utilizó un controlador proporcional-integral (PI), cuya sintonización se realizó por el método de Dahlin. Se determinó que la wavelet más adecuada para esta aplicación es la wavelet Daubechies más concretamente la db4 y se realizó una descomposición a 6 niveles. En (Guillen, 2004) se detalla el procedimiento para determinar la wavelet más apropiada y el número de niveles de descomposición.

4.2 Resultados experimentales

Se generó un grupo de 5000 señales las cuales se dividieron en cinco grupos de 1000. Cada grupo de 1000 representa un tipo de falla. Se trabaja con diferentes niveles de ruido, específicamente $\sigma^2=0.025$ y 0.05 . Estando el proceso en condiciones normales de operación, se obtuvo que el valor umbral para el IAE, para esta aplicación, es igual a 7. El comportamiento de los dos clasificadores, tanto por distancia Euclideana como K vecinos más cercanos utilizando las especificaciones anteriormente descritas es mostrado en la Tabla 1. Igualmente, se calculó la exactitud o porcentaje de exactitud de clasificación obtenido para cada uno de los clasificadores. El porcentaje de exactitud del clasificador se calculó de acuerdo con la siguiente expresión (Gaouda, 2002):

$$exactitud = \frac{N_{correctas}}{N_{Total}} * 100 \quad (15)$$

donde $N_{correctas}$ es el número de clasificaciones correctas y N_{Total} es el número total de realizaciones del experimento incluyendo las clasificaciones erróneas.

Tabla 1. Matriz de clasificación usando distancia Euclídeana. **C1v** (falla en el *span* del actuador de la válvula), **C3v** (falla calibración de cero del actuador de la válvula), **C2t** (falla ZM del *span* del transmisor), **C4t** (falla ZM de cero del transmisor), **P1** (perturbación en el flujo Q_1).
3.a. $\sigma^2 = 0.025$

	Fallas	C1v	C3v	C2t	C4t	P1	Exactitud (%)
Distancia Euclídeana	C1v	877	3	0	0	87	
	C3v	0	997	0	0	0	
	C2t	0	0	1000	0	0	
	C4t	0	0	0	1000	0	
	P1	123	0	0	0	913	95.74
K vecinos más cercano	C1v	882	0	0	0	89	
	C3v	1	1000	0	0	0	
	C2t	0	0	999	0	0	
	C4t	0	0	1	1000	0	
	P1	117	0	0	0	911	95.84

3.b. $\sigma^2 = 0.05$.

	Fallas	C1v	C3v	C2t	C4t	P1	Exactitud (%)
Distancia Euclídeana	C1v	789	37	0	0	134	
	C3v	4	960	0	0	0	
	C2t	0	0	999	0	0	
	C4t	0	0	1	1000	0	
	P1	207	3	0	0	866	92.28
K vecinos más cercano	C1v	791	20	0	0	126	
	C3v	7	978	0	0	1	
	C2t	0	0	983	0	0	
	C4t	0	0	17	1000	0	
	P1	202	2	0	0	873	92.5

Tabla 2. Porcentaje de desajuste originado por las fallas

Patrón de Falla	% error usando Distancia Euclídeana		% error K vecinos más cercano con la media de las distancias, k=2.		% error K vecinos más cercano con la media ponderada de las distancias, k=2.	
	$\sigma^2=0.025$	$\sigma^2=0.05$	$\sigma^2=0.025$	$\sigma^2=0.05$	$\sigma^2=0.025$	$\sigma^2=0.05$
C1v	37.3	40.2	8.1	9.2	6.4	8.3
C3v	25.6	26.9	5.9	7.2	6	6
C2t	24.2	25.5	6.9	8.7	7.4	8.5
C4t	53.0	48.4	5.8	6.1	5.6	6.1
P1	37.5	43.6	8.31	9.28	6.65	7.12

Cada una de las fallas de la Tabla 1 se hace pasar por el clasificador 1000 veces para cada nivel de ruido. Cada vez que la señal pasa por el clasificador se calcula la distancia Euclídeana del vector característico de la falla desconocida con cada una de las clases de falla y se determina cuantas veces se clasifica correctamente e incorrectamente. En la matriz de clasificación (Tabla 1) el número de clasificaciones correctas se encuentran a lo largo de la diagonal principal y fuera de la diagonal las clasificaciones incorrectas. Así, por ejemplo, para un nivel de ruido de 0.05, de las 1000 veces que se hace pasar la falla generada en la válvula con un 13% de desajuste (C1v), 78.9% de las veces se clasificó correctamente, 20.7% de las veces se clasificó incorrectamente ya que se clasificó como una perturbación (P1), finalmente 4% se clasificó como falla tipo C3v.

En la Tabla 1 también, se puede observar que, en cada uno de los clasificadores la exactitud del mismo se ve afectada

negativamente por el nivel de ruido que está presente en las fallas, sobre todo en aquellas que presentan mayor similitud, específicamente entre las fallas en el *span* del actuador de la válvula (C1v) y las perturbaciones (P1).

4.3 Determinación del porcentaje de fallas

Una vez identificadas las fallas se procedió a estimar el porcentaje de desajuste que produce cada una de estas fallas y perturbaciones sobre los instrumentos del proceso. Para estimar la robustez del clasificador estimando el porcentaje de desajuste, se procedió a realizar simulaciones de fallas desde 7.5% hasta 17.5% de desajuste, incrementando el porcentaje cada 0.5% obteniéndose un total de 21 fallas por clase. Luego se calculó el error promedio para cada clase de falla y perturbación. Estos resultados se muestran en la Tabla 2, donde en el método basado en K vecinos más cercano

se selecciono un valor de $k=2$. Dicha tabla muestra que el error promedio aumenta conforme aumenta el nivel de ruido. Sin embargo, el método basado en K vecinos más cercano es afectado en menor grado que para el caso del clasificador por distancia Euclideana.

Para el caso de distancia Euclideana (ver Tabla 2) el porcentaje de desajuste presentado tanto para fallas como para perturbaciones, se observa que el error promedio aumenta conforme aumenta el nivel de ruido, también se observa que el menor error se produce para el caso de C2t, este resultado es prácticamente independientemente del nivel de ruido.

Para el caso de la media de las distancias Euclidianas usando dos valores de k (ver Tabla 2) se observa que el porcentaje de error es afectado nuevamente por el nivel de ruido. Sin embargo, es afectado en menor grado que para el caso del clasificador por distancia Euclideana. También se puede observar que el menor error se comete para el caso de C4t. El valor de k es un parámetro que influye sobre la estimación del error ya que el error de estimación disminuye para casi todos los casos, esto es debido a que cuando k crece el valor calculado por la estimación se aleja del valor verdadero, ya que esa tercera distancia está muy alejada del valor más cercano al valor verdadero de desajuste.

Para el caso de la estimación del porcentaje de desajuste estimado a través del uso de la media ponderada. Se observa que el menor error se comete para C4t, igualmente el error es afectado por el nivel de ruido presente en la señal.

En general puede concluirse que el clasificador por distancia Euclideana rinde resultados similares a los encontrados usando el clasificador por K vecinos más cercanos, note esto al comparar las Tablas 1 y 2. Sin embargo, en la determinación del porcentaje de desajuste el método basado en K vecinos más cercanos produce mejores resultados a los encontrados con el método basado en la distancia Euclideana.

4.4 Comparación entre el clasificador por Distancia Euclideana y K-vecino más cercano

Para realizar la comparación entre estos dos clasificadores se compararon los porcentajes de exactitud de cada una de las matrices de clasificación los cuales se observan en la Tabla 3.

Tabla 3. Comparación entre el clasificador por Distancia Euclideana y K-vecino más cercano

Clasificador	%exactitud	
	$\sigma^2=0.025$	$\sigma^2=0.05$
Distancia Euclideana	95.74	92.28
K-vecino más cercano	95.84	92.5
Diferencia	0.1	0.22

De acuerdo con los resultados de la Tabla 3 los dos clasificadores presentan prácticamente la misma exactitud para clasificar las fallas, las diferencias que existen entre el porcentaje de exactitud generado por un método u otro son

muy pequeñas. Igualmente, estos resultados son presentados en los siguientes gráficos de barras.

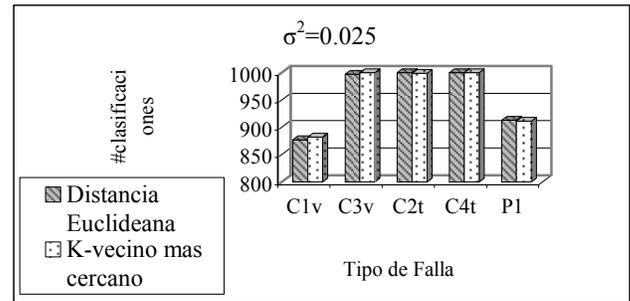


Fig.4 Comparación de los dos clasificadores para un $\sigma^2=0.025$

En la Fig. 4, se observa que el clasificador por K -vecino más cercano es más eficiente a la hora de clasificar la falla C3v ya que de las 1000 veces que la falla pasa por el clasificador, este la clasifica 1000 veces correctamente, igualmente para el caso de C1v la cual es clasificada correctamente en un 88.2% comparado con un 87.7% del clasificador por distancia Euclideana (se sugiere ver también la Tabla 1). El resto de las fallas las clasifica prácticamente con el mismo porcentaje de exactitud.

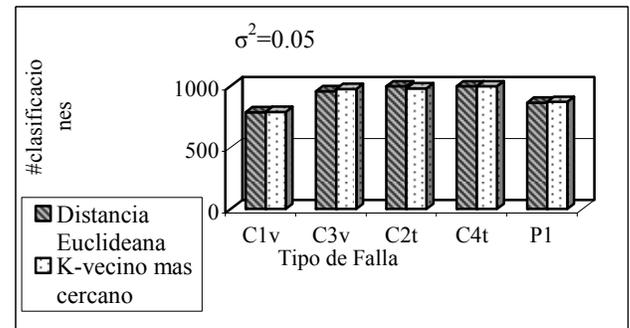


Fig. 5. Comparación de los dos clasificadores para un $\sigma^2=0.05$

En cuanto a la Fig. 5 correspondiente a un nivel de ruido de 0.05, la diferencia más resaltante se observa en la clasificación de la falla C3v, donde el clasificador por K -vecino más cercano la clasifica correctamente en un 97.8%, comparado con un 96% para el clasificador por distancia Euclideana (se sugiere ver también la Tabla 1). Contrariamente, ocurre con la falla C2t donde el clasificador por distancia Euclideana la clasifica correctamente en un 99.9% comparado con 98.3% del otro clasificador. Para el resto de las fallas el porcentaje de clasificaciones es muy similar.

5 Conclusiones

Los dos clasificadores implementados en este trabajo (Distancia Euclideana y K vecinos más cercano) permiten el diagnóstico y clasificación de la falla que se está presentando como una de las distintas clases de fallas que se encuentran almacenadas en la base de datos, bien sea falla en el actuador de la válvula, falla de transmisor o perturbación. Ambos clasificadores presentan porcentajes de exactitud aceptables y las dife-

rencias entre estos dos clasificadores son muy pequeñas. Adicionalmente, el método propuesto tiene la ventaja de estimar el porcentaje de desajuste que se está presentado con ambos clasificadores. Sin embargo; la estimación de dicho porcentaje es realizada de una forma más eficiente por el clasificador K vecinos más cercano, pues, éste presenta errores menores que el clasificador por Distancia Euclídeana. El uso de otros clasificadores, como por ejemplo maquinas de soporte vectorial o redes neuronales, pudieran mejorar el desempeño en el proceso de clasificación con, posiblemente, mayor complejidad. La metodología aquí establecida, en cuanto al uso de la técnica de pre-procesamiento wavelet así como la escogencia del vector característico puede ser extendida usando estos clasificadores.

Referencias

- Burrus C, 1989, Introduction to Wavelets and Wavelet Transforms, Prentice Hall, Upper Saddle River, New Jersey 07458, ISBN 0-13-489600-9.
- Camacho O, Smith C.A., 2000, Sliding mode control: an approach to regulate nonlinear chemical processes, ISA Transactions 39 pp. 205-218.
- Fu KS, 1982, Applications of Pattern Recognition, Boca de Ratón, Florida, pp. 5-6, pp. 25-26.
- Gaouda AM, Kanoum SH, Salama, MMA and. Chikhani AY 2002, Pattern "Recognition Applications for Power System Disturbance Classification, IEEE Transactions on Power Delivery, Vol. 17, NO. 3.
- Gertler Janos J, 1998, Fault Detection and Diagnosis in Engineering Systems, cap1, New York: Marcel Dekker.
- Guillén M, Leonel, 2004, Detección y Diagnóstico de Fallas utilizando la Transformada Wavelet, Tesis de Grado para optar al título de Magister Scientiae dirigida por José L. Paredes, Postgrado en Automatización e Instrumentación, Facultad de Ingeniería, Universidad de Los Andes Mérida.
- Jiang T Bingzhen C, 2003, Application of steady-state detection method based on wavelet transform, Computers and Chemical Engineering 27 pp. 569-578.
- Mallat G, 1989, A theory for multiresolution signal decomposition: The wavelet representation, IEEE Transactions on Pattern Recognition and Machine Intelligence. Vol. 11.pp. 674-693.
- Szigeti F, and Tarantino R, 1998, Augmented Space Method for Fault Detection and Isolation Filter. Poster Workshop in Petrochemical Processes Fault Detection. Instituto Francés del Petróleo. Lyon.
- Tarantino R and Szigeti F, 2000, Generalized Luenberger Observer-Based-Fault Detection Filter Design: An Industrial Application, Control Engineering Practice, Vol. 8, Issue 6. pp. 665-671.
- Zhang JQ and Yong Y, 2001, A Wavelet-Based Approach to Abrupt Fault Detection and Diagnosis of Sensors, IEEE Transactions on Instrumentation and Measurement, Vol. 50, NO. 5, OCTOBER.