

Detección y clasificación de defectos en estructuras tipo tubería mediante reconocimiento de patrones.

Gustavo Adolfo Moreno Beltrán¹, Jhonatan Camacho Navarro², Rodolfo Villamizar Mejia³

Universidad Industrial de Santander, Bucaramanga, Santander, Colombia.

³Universidad Politécnica de Cataluña, Barcelona, España.

Artículo recibido en mes 08 de 2016; artículo aceptado en mes 08 de 2016

Resumen

En este documento se presentan los resultados de la implementación de máquinas de soporte vectorial en la tarea de detección de fugas en una tubería de acero carbono, usando los datos provenientes de sensores piezoeléctricos y teniendo en cuenta condiciones de temperatura variable.

¹Ingeniero Electrónico, Universidad Industrial de Santander. Estudiante De Maestría en Ingeniería Electrónica- Investigador del grupo: CEMOS. Universidad Industrial de Santander de la ciudad de Bucaramanga (Colombia): Dirección Carrera 27 calle 9. PBX: (57) (7) 6344000 Ext:1295. Correo electrónico institucional: gustavo.moreno2@correo.uis.edu.co

²Ingeniero Electrónico, Universidad Industrial de Santander. Maestría En Ingeniería Electrónica, Universidad Industrial de Santander. Candidato a Doctor- investigador del grupo: CoDaLab. Universidad Politécnica de Cataluña (Barcelona, España). Docente Investigador grupo GPS universitaria de investigación y desarrollo UDI. Correo electrónico institucional: jcamacho2@udi.edu.co.

³Ingeniero Electrónico, Ingeniero Electricista, Doctor en Tecnologías de la Información. Universidad Industrial de Santander. Profesor Asociado, investigador grupo CEMOS – UIS. Correo electrónico institucional: rovillam@uis.edu.co.

Se plantean dos métodos para la creación de un modelo nominal o de estado sin daños de la estructura, en el primero en donde se utilizan los datos de la estructura simplemente a temperatura ambiente y en el otro método incluyendo información del estado nominal a las diferentes temperaturas de operación de la estructura. Y se realiza la comparación de la implementación de máquinas de soporte vectorial para cada uno de los modelos.

Finalmente se realiza la comparación del desempeño del uso de un kernel gaussiano y uno lineal en el entrenamiento de las máquinas de soporte vectorial para ambos modelos

Palabras clave: Análisis de componentes principales, Maquinas de Soporte Vectorial, Monitorización de Salud Estructural, detección de fugas en tuberías.

Resumen en inglés.

This document presents the results of implementing support vector machines for damage detection in a steel pipe section. Data are collected from piezoelectric devices and considering variable temperature conditions. Two procedures were used to obtain a baseline model which represent structural undamaged states. The first one consist on using only environmental temperature as reference performance, and the second is based on an augmented data matrix containing information about operational ranges of temperature. Support vector machines are implemented for both cases. A comparison of the performance for linear and Gaussian kernels is developed in order to evaluate the effectiveness of support vector machines for pipeline damage detection.

Introducción

La detección temprana de fallas en tuberías permite evitar desde pérdidas económicas hasta catástrofes que pueden acabar con pérdidas humanas. Dichas fallas pueden ocurrir por diferentes causas tales como fugas, agrietamientos, envejecimiento y corrosión, entre otros. Aunque actualmente los métodos de evaluación pueden ser usados para tareas de inspección, dichos sistemas no consideran un monitoreo continuo y la obtención de un diagnóstico requiere detener la operación normal de la tubería. La Monitorización de salud estructural (SHM) busca una detección temprana de la falla en una estructura, mediante el desarrollo de

sistemas de detección y prevención temprana de la falla en estructuras de manera no invasiva y así reducir desde pérdidas económicas hasta el fallecimiento de personas.

En este proyecto se busca utilizar un método de reconocimiento de patrones para detectar si existe una fuga en una tubería o se encuentra en su estado normal de operación, mediante el uso de dispositivos piezoeléctricos de manera no invasiva para enviar y registrar las ondas tipo Lamb que viajan a lo largo de la estructura, con estas señales registradas se realiza la extracción de características para identificar los diferentes estados de la estructura y seguido a esto se realiza el entrenamiento de una máquina de soporte vectorial (SVM) la cual determina si existe o no una fuga en la tubería. Además, se realizó la validación cruzada de los datos para determinar la eficiencia de la SVM en la detección de la anomalía en la estructura.

Detección de fugas usando máquinas de soporte vectorial

Dado que no es tan sencillo identificar a simple vista en las señales registradas por los dispositivos piezoeléctricos se ve la necesidad de utilizar métodos basados en el reconocimiento de patrones para realizar la tarea de clasificación entre los estados sin daño y con daño de la estructura.

Metodología

adquisición de datos.

Se realizaron experimentos por medio de piezo-diagnos⁴, donde se almacenaron los datos de las señales de los piezoeléctricos por medio de un PicoScope, también se registró la temperatura por medio de un prototipo de desarrollo Arduino.

⁴ Piezo-diagnos: se define como una técnica de SHM que se basa en el uso de ondas guiadas obtenidas al instrumentar una estructura con dispositivos piezoeléctricos PZT, aprovechando su dualidad actuador-sensor. Así, el diagnóstico del estado actual de una estructura se realiza mediante el análisis estadístico de las características de propagación de la onda guiada a través de la estructura

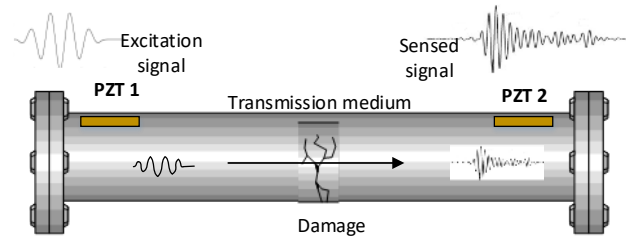


Figura 1. Ilustración del comportamiento de la onda guiada sobre una estructura, el PZT1 representa el actuador quien se encarga de convertir la señal eléctrica generada en el PicoScope en una señal mecánica que viaja a través de la tubería y es recibida de manera inversa en PZT2.

organización de los datos.

Mediante MATLAB se realizó la lectura de los datos generados por el PicoScope mediante un código en Matlab se organizan los datos en una matriz como se muestra en la *Figura 2*.

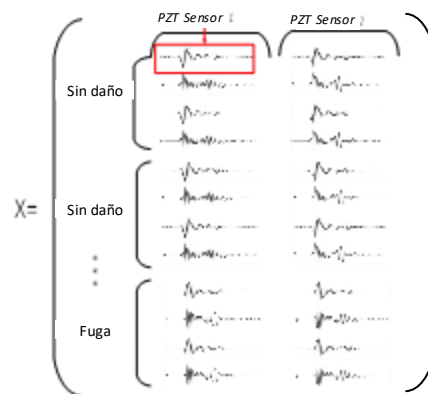


Figura 2. Creación de matriz de estados de la estructura, se organizan los datos obtenidos en la etapa de adquisición de datos con el fin de tener una estructura estándar como la mostrada anteriormente para facilitar el procesamiento de los datos.

PCA-Modelo de línea base.

Después de organizados los datos se toman los del estado sin daños y se realiza el cálculo de los componentes principales donde se obtiene una reducción en la dimensionalidad, en la *Figura 3* se puede observar un resumen de este proceso, los valores de los vectores propios

asociados a la mayor varianza de los datos son los que se utilizan para realizar la proyección de los nuevos datos sobre el nuevo espacio.

El número de componentes utilizados es el que acumula el 99% de la varianza de los datos.

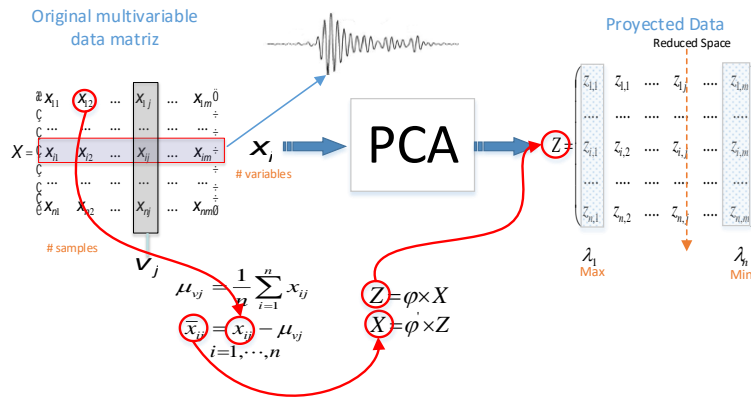


Figura 3. Valores y Vectores propios por medio de los v

Calculo índices T² y Q

Índice Hotelling T²: El índice T² es la distancia en el espacio *r-dimensional* definido por los componentes principales respecto al origen, calculado usando la ecuación (6).

$$T^2 = \sum_{k=1}^r \frac{t_k^2}{\lambda_k} = Z^T \lambda^{-1} Z \quad (6)$$

En la ecuación (6), el índice T² se calcula mediante la sumatoria de los cuadrados de cada elemento t_k del vector de SCORES o proyecciones normalizado por su respectiva varianza λ_k o valor propio.

Índice Q-estadístico: El índice Q corresponde al error cuadrático expresado en la ecuación (7).

$$Q = \sum_k (e_k)^2 \quad (7)$$

Donde e_k son los elementos del vector de error residual obtenido luego de reconstruir la medición como se expresa en la ecuación (8).

$$e = X - \phi'Z \quad (8)$$

La razón principal del uso de los scores y de los índices PCA en el algoritmo, se debe a que facilitan la visualización en 2D de los diferentes tipos de daños, y que facilitan la tarea de clasificación de defectos.

entrenamiento y validación de la SVM (Maquina de Soporte Vectorial)

Después de calcular los valores de los índices T^2 y Q , teniendo en cuenta la temperatura y la etiqueta de cada uno de los estados de la estructura, dado que estos datos quedan organizados en paquetes por cada uno de los estados se realiza una mezcla conservando las etiquetas.

Luego de tener los datos mezclados se dividen en 5 grupos se realiza el respectivo entrenamiento y validación cruzada.

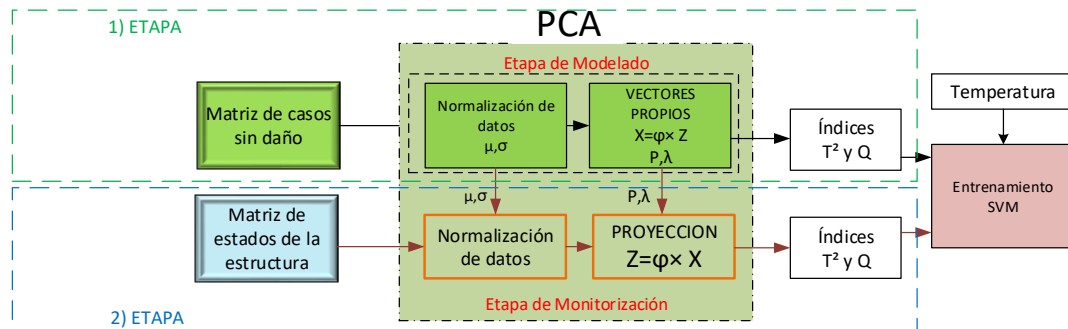


Figura 4. Resumen metodológico del procesamiento de datos, en la ilustración se muestra un diagrama de bloques donde se puede ver el procedimiento llevado a cabo para la detección de fugas

Procedimiento y Resultados

Por medio un algoritmo realizado en Matlab se llevó a cabo la organización de los datos de los sensores para facilitar el procesamiento de los mismos, después de esto se realizó el respectivo cálculo de los índices T^2 y Q , se creó una matriz de datos donde se organizaron estas características, además se incluyó el valor de la temperatura de cada uno de los registros.

En la Figura 5 se observa en los tiempos y las diferentes temperaturas a las cuales se realizó el registro de los datos de los sensores.

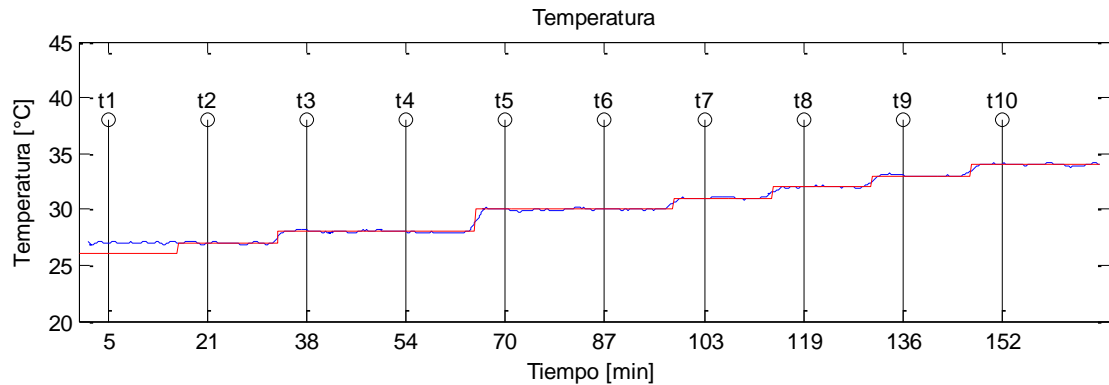


Figura 5. Registro de datos a diferentes temperaturas. Se presenta la gráfica de los sensores de temperatura ubicados cerca a los sensores y los tiempos en los cuales se realizó el registro de los piezoeléctricos

Estimación de la cantidad de componentes principales utilizados y cálculo de los estadísticos T2 y Q

En la Figura 6 se puede observar el comportamiento de los valores propios de la matriz de covarianza de los datos correspondientes a las señales del estado sin daños de la estructura. Y se determina el número de componentes principales que se van a utilizar en este caso el criterio es el del 90% de la varianza de los datos.

NOMBRE NOMBRE APELLIDO APELLIDO, NOMBRE NOMBRE APELLIDO APELLIDO
Efectos de la ejercitación gestual mediante sensores faciales en la expresión de la ira.

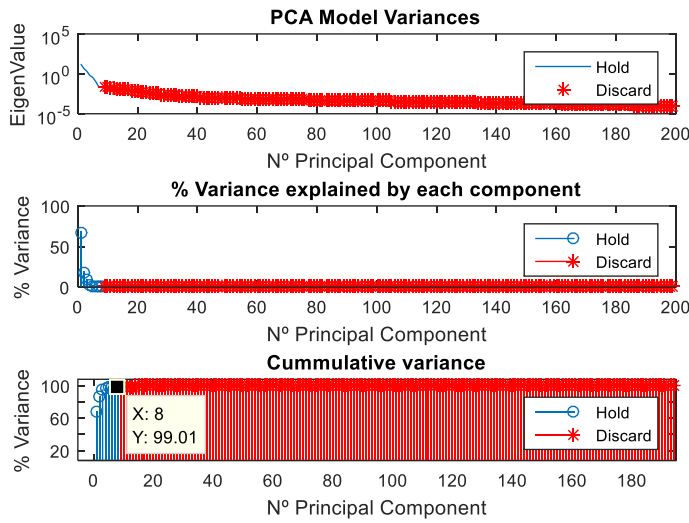


Figura 6 Graficas valores propios y varianza acumulada

Para calcular Q se calculó el error que existe entre los datos originales y su proyección sobre los vectores propios elegidos. Y el T^2 es la distancia en el espacio r -dimensional definido por los componentes principales respecto al origen. Seguido de esto se realizó el cálculo de los índices. Dado que se tienen 3 características T^2 , Q y la temperatura se realizó un gráfico en 3 dimensiones para observar el comportamiento de los datos como se podrá observar en las figuras mostradas más adelante.

A continuación, se presentará el paralelo entre 2 métodos utilizados para calcular los índices T^2 y Q :

En el primer método se realizó el cálculo del modelo PCA teniendo en cuenta solo el estado sin daños de la estructura a temperatura ambiente. Y en el segundo método se calculó el modelo PCA teniendo en cuenta los Datos a temperatura mínima y los datos a la temperatura máxima de operación.

En las figuras presentadas la representación de la izquierda corresponde al **método 1** y a la derecha al **método 2**:

NOMBRE NOMBRE APELLIDO APELLIDO, NOMBRE NOMBRE APELLIDO APELLIDO
 Efectos de la ejercitación gestual mediante sensores faciales en la expresión de la ira.

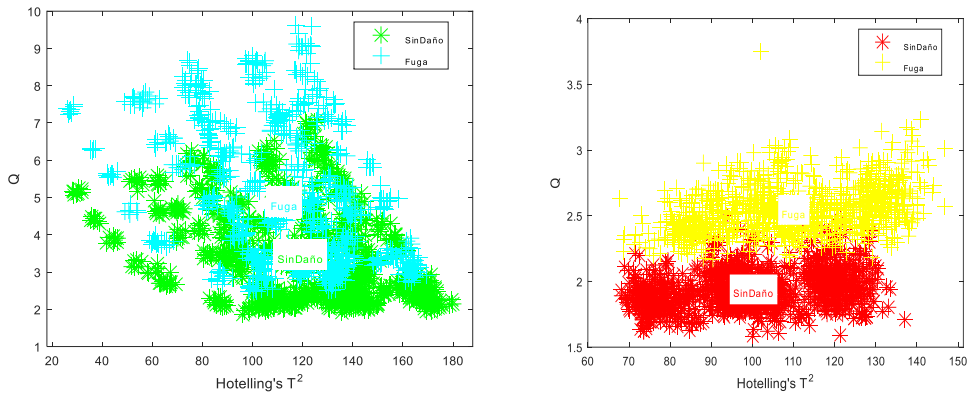


Figura 7. Grafica de los índices Q vs T²

Como se observa en la Figura 7 el método 1 presenta un solapamiento de las señales de los estados sin daño y con daño de la tubería, por otro lado, el método 2 ofrece una separabilidad de los datos mayor la en la cual se podría llegar a pensar que un clasificador simple como un umbral a lo largo de eje Q brindaría un buen desempeño.

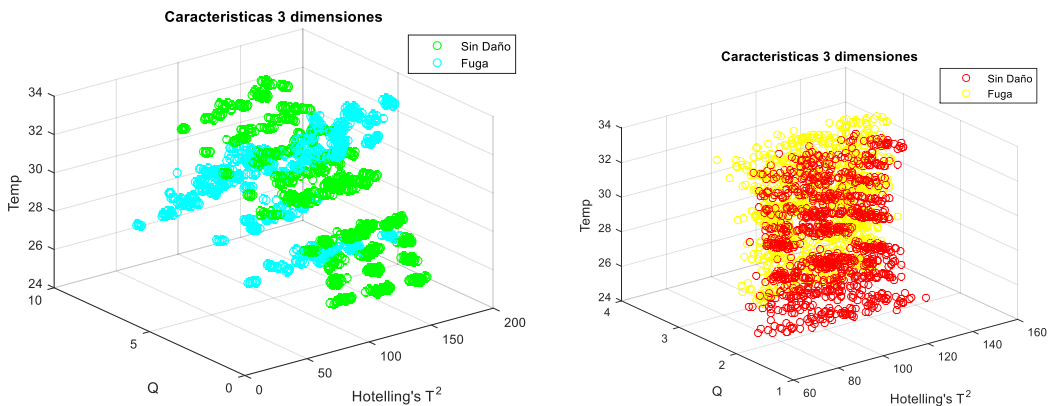


Figura 8. Grafica Q vs T² y Temperatura de registro

En la Figura 1 se observa que al graficar en 3 dimensiones la características Q, T² y en el eje z la temperatura a la que se encontraba la estructura en la toma de datos se puede observar que mejora la separabilidad entre los estados con daño y sin daño de la estructura.

Se realizó el entrenamiento una máquina de soporte vectorial usando un kernel gaussiano tanto en el método 1 con las características Q, T² e incluyendo la condición de temperatura a

la cual se encontraba la estructura en el momento del registro de los datos. Y en el método 2 solo con los índices Q, T².

Después de tener las SVM de cada uno de los dos métodos se realizó su respectiva validación cruzada para obtener el valor del Accuracy en cada uno de los 5 grupos usados para la validación, donde se realizó una mezcla de las posiciones de los datos para que no quedaran de manera organizada y así tener mayor variación en los datos tanto de entrenamiento como de validación.

En las Figura 9 y Figura 10 se muestran los datos de cada una de las 2 clases y sus respectivos vectores de soporte encerrados en círculos negros. Tanto en 2 dimensiones como en 3 dimensiones.

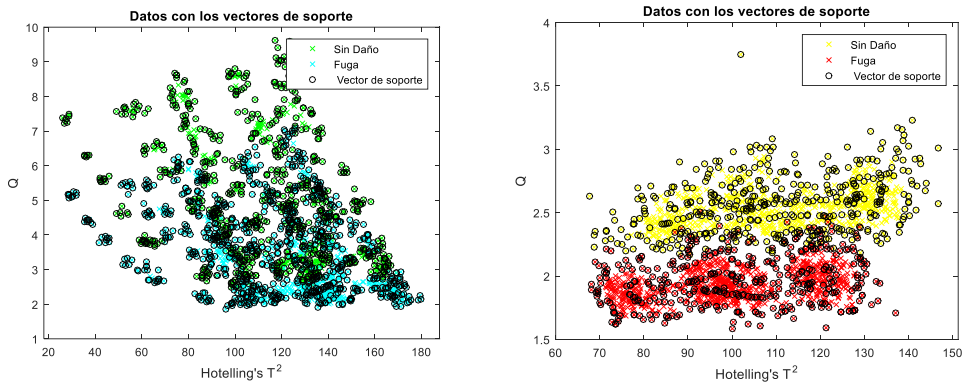


Figura 9. Grafica Q vs T² y Vectores de Soporte

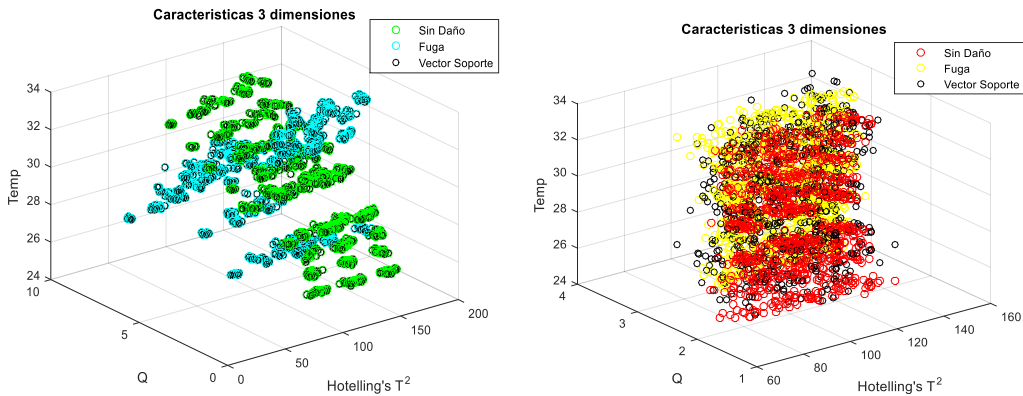


Figura 10 Grafica Q vs T² vs Temperatura de registro y los vectores de soporte

En la Figura 11 se presentan los datos con una x y su respectiva clase además de ello se presentan los resultados del clasificador con un círculo que encierra cada uno de los datos y el color define la clase a la cual se está asignando dicho dato por el clasificador.

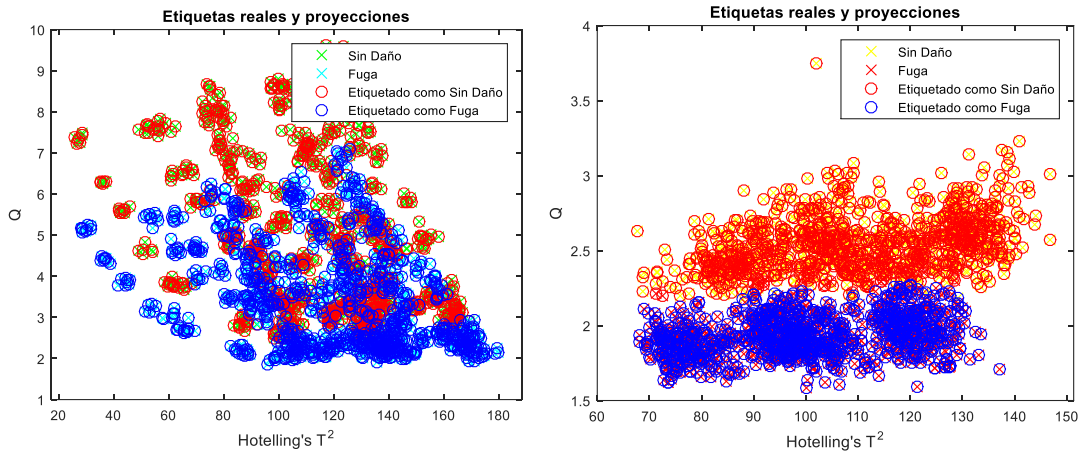


Figura 11 Grafica de clasificación de los datos

A continuación, se presentan los valores de Accuracy para cada uno de los grupos y el valor del promedio de Accuracy, para los dos clasificadores entrenados.

Vector de Accuracy para los grupos:	Vector de Accuracy para los grupos:
99.2857 98.0952 98.3333 99.0476 98.5714	97.8571 98.3333 97.8571 97.1429 96.1905
Accuracy Promedio de los grupos:	Accuracy Promedio de los grupos:
98.6667	97.4762

Debido a la distribución de los datos del método 2 se decidió realizar el entrenamiento con un kernel lineal y comparar el desempeño.

NOMBRE NOMBRE APELLIDO APELLIDO, NOMBRE NOMBRE APELLIDO APELLIDO
 Efectos de la ejercitación gestual mediante sensores faciales en la expresión de la ira.

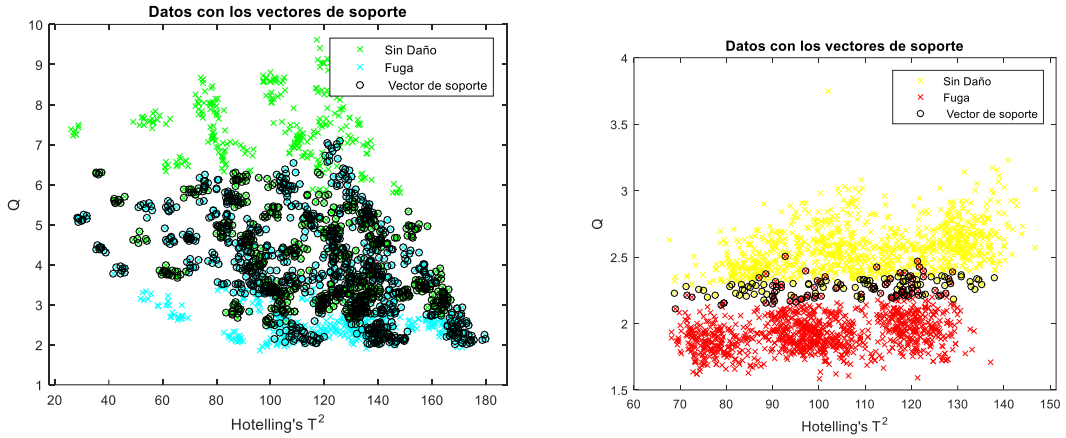


Figura 12. Grafica Q vs T² y Vectores de Soporte

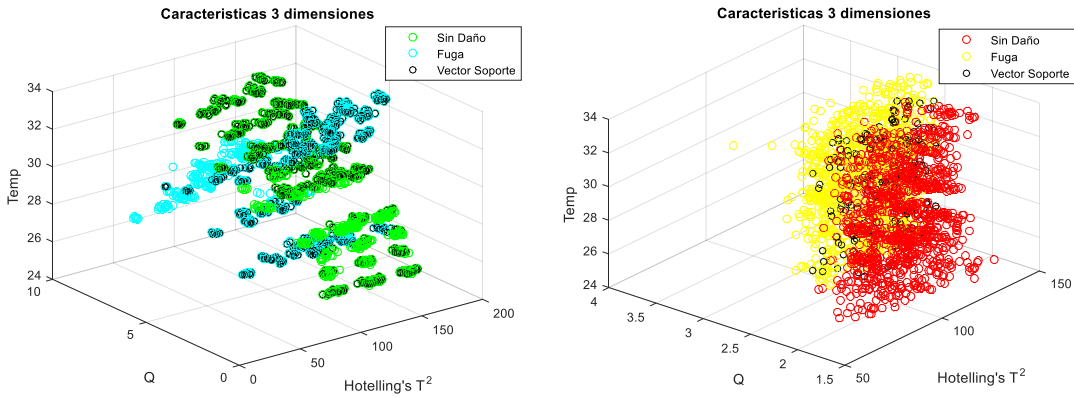


Figura 13. Grafica Q vs T² vs Temperatura de registro y los vectores de soporte

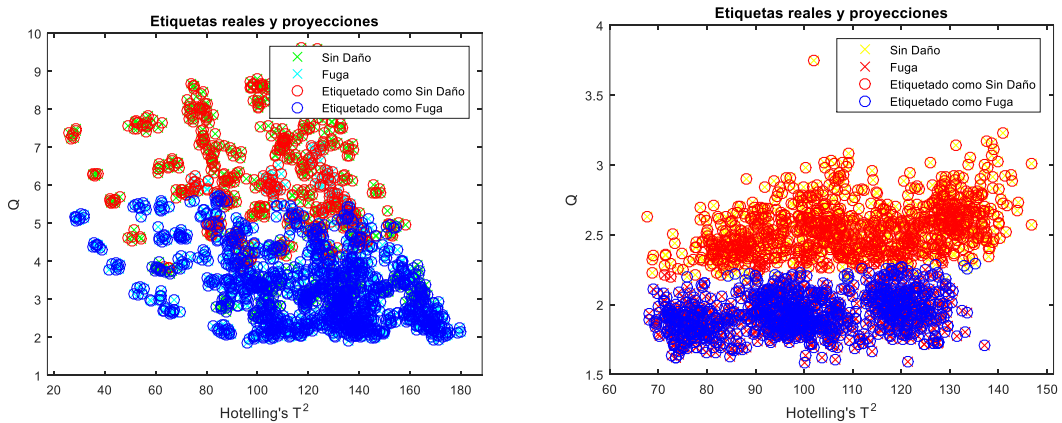


Figura 14. Grafica de clasificación de los datos

NOMBRE NOMBRE APELLIDO APELLIDO, NOMBRE NOMBRE APELLIDO APELLIDO
Efectos de la ejercitación gestual mediante sensores faciales en la expresión de la ira.

A continuación, se presentan los valores de Accuracy para cada uno de los grupos y el valor del promedio de Accuracy, para los dos clasificadores entrenados.

Vector de Accuracy para los grupos: 72.3810 70.9524 71.6667 75.2381 71.1905	Vector de Accuracy para los grupos: 98.0952 97.3810 98.3333 97.3810 95.2381
Accuracy Promedio de los grupos: 72.2857	Accuracy Promedio de los grupos: 97.2857

Comentarios

Se pudo observar que por medio una SVM como clasificador de fugas se puede identificar una fuga en la estructura por cualquiera de los dos métodos implementados en la creación del modelo base para la extracción de las características utilizada donde se utilizaron los índices estadísticos T^2 y Q para obtener la información obtenida por medio de las señales provenientes de los sensores piezoeléctricos. Además se puede observar que un Kernel gaussiano en la implementación de la SVM permite una mayor precisión en ambos modelos ya que el lineal en el caso del segundo modelo presenta un porcentaje de acierto alrededor del 75%.

Agradecimientos

Esta investigación hace parte del proyecto “Monitorización y detección de defectos en estructuras usando algoritmos expertos embebidos” financiado por COLCIENCIAS, donde el algoritmo usado ha sido estudiado y mejorado por los grupos de investigación CEMOS de la UIS (Bucaramanga, Colombia) y de CoDALab de la UPC (Barcelona, España).

Referencias

Tibaduiza, D. A., Mujica, L. E. and Rodellar, J. (2013), Damage classification in structural health monitoring using principal component analysis and self-organizing maps. *Struct. Control Health Monit.*, 20: 1303–1316. doi:10.1002/stc.1540

L.E. Mujica, J. Rodellar, A. Fernandez, and A. Guemes. (2010.), Q-statistic and t2-statistic pca-based for damage assessment in structures. *Structural Health Monitoring*, online version:1–15.

A. & C. C. E. Raghavan, (2008). Effects of elevated temperature on guided-wave structural health monitoring.,» *Journal of Intelligent Material Systems and Structures*, pp. 1383-1398.

I+D Revista de Investigaciones ISSN 22561676 Volumen 1 Número 1 Año 01 Enero-Junio 2013 pp.xx-xx

Camacho-Navarro, J., Ruiz, M., Mujica, L.E., Perez-Gamboa, O., Villamizar, R. (2015) Pipe leaks detection under varying environmental conditions by using a data driven approach. A: Panamerican Conference of Nondestructive Testing and International Congress Of Welding and NDT. "6th Panamerican Conference of Nondestructive Testing and 8th International Congress Of Welding and NDT". Cartagena.

Giraldo, D. F., Dyke, S. J., & Caicedo, J. M. (2006). Damage detection accommodating varying environmental conditions. *Structural Health Monitoring*, 5(2), 155-172.

Sohn, H. (2007). Effects of environmental and operational variability on structural health monitoring. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 365(1851), 539-560.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Machine Learning (Vol. 4). springer New York.

Webb, A. R. (2002) *Nonlinear Discriminant Analysis – Kernel Methods*, in *Statistical Pattern Recognition*, Second Edition, John Wiley & Sons, Ltd, Chichester, UK. doi: 10.1002/0470854774.ch5