

NUEVOS AVANCES EN LA APLICACIÓN DE RECONOCIMIENTO DE PATRONES Y ALGORITMOS DE SEPARACIÓN DE FUENTES A LA DETECCIÓN Y CLASIFICACIÓN AUTOMÁTICA DE RUIDO DE TRÁFICO

Fernando J. Mato-Méndez, Manuel A. Sobreira-Seoane

E.T.S. de Telecomunicación, Campus de la Universidad de Vigo. Universidad de Vigo. España.
(fmato@tsc.uvigo.es, msobre@gts.uvigo.es)

Resumen

El método común CNOSSOS-EU (Common NOise ASSEsment MethOdS), en fase aún de elaboración, persigue la unificación de resultados en la elaboración de mapas estratégicos de ruido dentro de la UE. En el caso concreto de ruido de tráfico, el método requiere ser alimentado con información relativa a la velocidad media de la vía e intensidad de tráfico, por categorías de vehículos y sentido de circulación. Por otro lado, los requisitos de aplicación impiden que esta información pueda ser entregada por un sólo dispositivo mediante la aplicación de soluciones clásicas. Es por este motivo que los autores de esta investigación se propusieron hace un tiempo la obtención de un sistema portable, capaz de entregar dicha información de forma compacta y en tiempo real. Su obtención, a partir de la señal de audio adquirida, supone sin embargo una tarea de elevada complejidad técnica. El trabajo que aquí se presenta aborda los aspectos más relevantes en la evolución del sistema, tras lograr importantes avances en materia de separación de fuentes y reconocimiento de patrones.

Palabras-clave: Ruido de Tráfico, CNOSSOS-EU, Clasificación, Reconocimiento de Patrones, Separación Ciega de Fuentes, BSS, BSS-PCA, Separación Ciega de Fuentes Implícitas, BISS, BISS-PCA.

Abstract

The method CNOSSOS-EU, which is under development, tries to give a common framework for the calculation of strategic noise map in the European Union. When traffic noise is considered, information on the hourly traffic intensity and average speed of the road is needed. This information must be classified according to the vehicles categorization specified in the method. This information is obtained using several sources and devices. The authors have proposed and develop a portable system which allows to collect real time information of the speed and number of vehicles, classified in agreement with the CNOSSOS-EU method, using the audio signal recorded with two microphones. In this paper the most relevant advances obtained in the last few years are presented.

Keywords: Traffic Noise, CNOSSOS-EU, Classification, Pattern Recognition, Blind Source Separation, BSS, BSS-PCA, Blind Implicit Source Separation, BISS, BISS-PCA.

PACS no. 43.60.Ac, 43.60.Bf, 43.60.Cg, 43.60.Fg, 43.60.Gk, 43.60.Hj, 43.60.Jn, 43.60.Lq, 43.60.Uv, 43.60.Vx, 43.60.Wy.

1 Introducción

La elaboración de mapas de ruido requiere el conocimiento previo de las características de las fuentes emisoras. En el caso de ruido de tráfico, la recolección de datos necesarios implica disponer de la intensidad media horaria de los vehículos, clasificados en al menos dos clases: pesados y ligeros. La propia evaluación mediante medidas conforme a la Norma ISO 1996-2: 2007 [1] requiere clasificar en estas dos clases los vehículos contados durante el período de medida. El método de cálculo europeo CNOSSOS-EU [2], actualmente en preparación, amplía los requisitos de la Norma extendiendo a cuatro categorías (vehículos ligeros, vehículos medio pesados, vehículos pesados y vehículos de dos ruedas) el modelo de clasificación. A este respecto, los sistemas disponibles para llevar a cabo esta recolección de datos no satisfacen las necesidades del método debido a diversos motivos que han sido analizados en detalle en [3]. En el caso concreto de la adquisición de la señal mediante un array de sensores acústicos, existen dos importantes problemas que no son abordados por el estado del arte: (1) En primer lugar, y como característica común, los sistemas propuestos proceden a la detección del evento mediante la localización del valor máximo de energía de la señal. En este sentido, la mayor parte de los sistemas relacionados abordan la tarea de clasificación sobre pasos aislados, presentando entonces un importante handicap de aplicabilidad real. En el caso de pasos combinados, el proceso habitual de segmentación seleccionando un tramo de señal entorno a dicho valor, presenta el inconveniente de la detección de falsos positivos (identificación de más de un evento en vehículos pesados de varios ejes) y la imposibilidad de detección de eventos próximos entre sí. (2) En condiciones reales, la tarea de clasificación a partir de la señal adquirida se ve dificultada por la existencia de información ajena al evento a clasificar. Ésta viene determinada por el solapamiento energético provocado por el paso combinado de vehículos y por reflexiones indeseadas asociadas a las características de la vía.

El sistema desarrollado en [3] resuelve los problemas anteriormente descritos, presentando actualmente una precisión de clasificación cercana al 100% en condiciones reales de funcionamiento. A diferencia de los sistemas existentes, además, responde de forma integral a las necesidades del método de cálculo CNOSSOS-EU. Su integración en un sistema portátil, dotado de un sencillo array de dos sensores, posibilita la recolección compacta durante el intervalo de medición de toda la información demandada. Se presenta aquí, desde sus inicios a partir de un primer estudio desarrollado por nuestro grupo de investigación [4], la evolución en su diseño [5-9]. Este estudio preliminar [4] aborda el problema de reconocimiento de patrones sobre tres categorías de vehículos (motos, coches y camiones) aplicando atributos de uso habitual en procesado de voz. Aparte de los aspectos de segmentación descritos, el problema de reconocimiento de patrones presenta una fuerte dependencia entre la base de datos utilizada para entrenar al clasificador, los atributos extraídos para caracterizar al evento, su selección, el algoritmo de clasificación propuesto y el procedimiento de evaluación de prestaciones. En el diseño del sistema desarrollado en [3] se han estudiado estas dependencias con objeto de optimizar cada una de las fases involucradas en la tarea de clasificación. Asimismo, este estudio nos ha permitido determinar el esquema de categorización apropiado para el manejo de las clases especificadas en [2], al presentar éstas una importante incertidumbre en las fronteras de separación.

El reto más complejo, sin embargo, ha sido conseguir la combinación óptima de técnicas de *Separación Ciega de Fuentes* (BSS) y reconocimiento de patrones. En primer lugar, la teoría desarrollada en BSS está destinada a la extracción de un conjunto de variables que coinciden con las fuentes presentes en la mezcla [10]. Ésta se caracteriza en general por una elevada relación S/N y por encontrarse la señal de ruido incorrelada con las variables de interés. En este sentido, su aplicación

aquí a la señal de ruido de tráfico puede ser considerada un trabajo pionero, como ha sido reconocido expresamente en diferentes trabajos [11,12]. Desde una perspectiva más general, la aplicación de BSS a una mezcla de señales de ruido no había sido abordada hasta ahora. El problema de clasificación de ruido de tráfico, además, difiere sustancialmente del entorno de aplicación de BSS, al aparecer las variables de interés como función implícita de las fuentes manejadas por el modelo. En esta dirección, la investigación llevada a cabo en [3] da lugar a la *Teoría de Separación de Fuentes Implícitas* (BISS) que es presentada en [10]. Tras una caracterización completa del nuevo problema BSS a resolver, en esta publicación se propone el método BISS-PCA para su solución. La posibilidad de aplicación de esta nueva teoría a otros ámbitos de investigación supone una contribución relevante al estado del arte.

2 Planteamiento del problema

Para nutrir de información al método CNOSSOS-EU se establece la necesidad de obtener datos acerca de la intensidad de tráfico de la vía y de su velocidad media. Se hace mención expresa a que el usuario proceda a la estimación de toda esta información en caso de no estar disponible. Por un lado, en el caso de la velocidad se debe utilizar el valor real, salvo que se exceda el límite legal de la vía. Por otro lado, la información acerca de la intensidad de tráfico debe ser entregada por sentido de circulación y categoría de vehículo. Para ello, se decide en [2] que la vía sea dividida en tramos homogéneos y que se disponga de los valores requeridos para cada uno de los tres períodos de evaluación (día, tarde, noche). Como se verá a continuación, la señal adquirida resulta de un proceso de combinación de información asociada a distintos eventos, dificultando con ello la tarea de clasificación.

2.1 Detección de eventos

Para nutrir al método de cálculo de información relativa a la intensidad de tráfico de la vía, se requiere proceder al conteo de los distintos eventos presentes en la señal adquirida, entregando el resultado de acuerdo a la categorización preestablecida en [2].

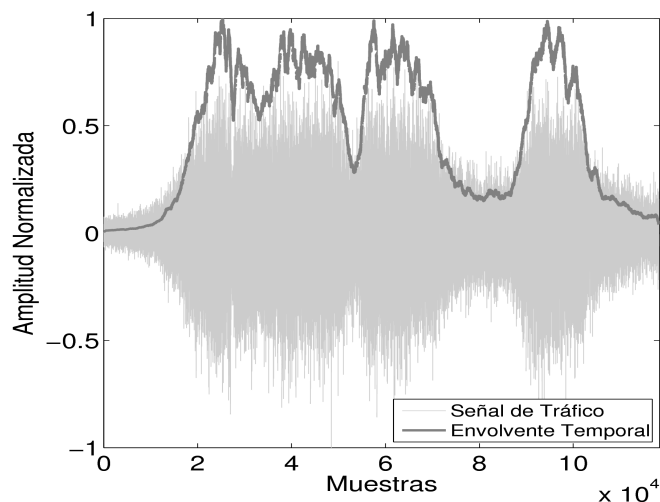


Figura 1 – Evolución de la envolvente temporal de la señal de ruido de tráfico.

El método de segmentación a emplear, en primer lugar, debe ser capaz de detectar con precisión los distintos eventos presentes y de entregar su información asociada al sistema de clasificación. Para pasos combinados, esta tarea se presenta compleja como se puede extraer de la Figura 1 para la evolución de la envolvente temporal de la señal adquirida. La existencia de máximos de energía de valor similar, como se puede observar, se traduce en una elevada incertidumbre en la selección del punto central del intervalo de segmentación. Se extrae, por otro lado, la dificultad en el establecimiento de los extremos de segmentación. En este sentido, se puede apreciar, como el uso del umbral asociado al ruido de fondo, solución habitual, no resulta suficiente.

2.2 Clasificación de eventos

En la Figura 2 se describen de forma genérica las etapas necesarias para abordar un problema de reconocimiento de patrones [13]. Entre éstas, se da cobertura a dos tareas íntimamente relacionadas y que constituyen el corazón del proceso: por un lado, encontrar el conjunto de características que mejor identifiquen a los diferentes candidatos (procesado de características) y, por otro, pensar en cómo cuantificar dicha similitud (clasificación).

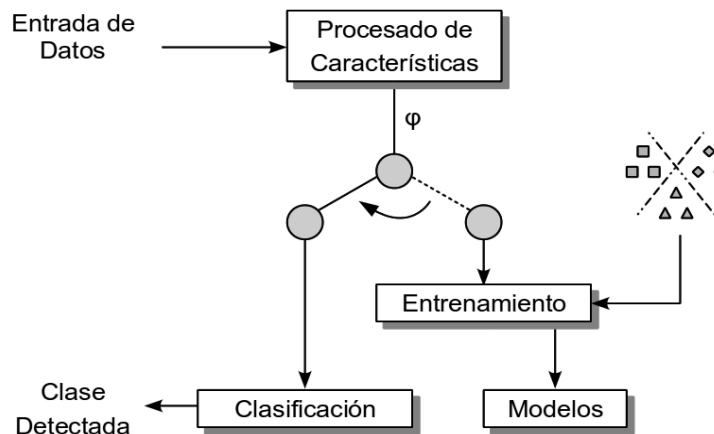


Figura 2 – Esquema de clasificación genérico.

En un problema de clasificación genérico se persigue asignar un patrón desconocido φ a una determinada clase C_i , perteneciente a un conjunto C de c clases determinadas previamente. La condición de partida es que cada uno de los patrones a clasificar debe estar representado por un único vector de características, y sólo puede pertenecer a una clase. Así, dicho patrón puede ser representado de forma unívoca por $\varphi = [\varphi_1, \varphi_2, \dots, \varphi_d]^T$, donde d indica el número de características consideradas y, por lo tanto, la dimensionalidad del problema de clasificación. La labor de clasificación comienza con un proceso de aprendizaje en el que se construye un modelo para el conjunto de clases definido, mediante la aplicación de una secuencia de entrenamiento. Este modelo se construye en la *fase de entrenamiento* a partir de las características seleccionadas en la etapa anterior. Es importante tener en

cuenta que uno de los puntos críticos para la robustez del sistema se encuentra en la capacidad del clasificador para responder a cambios en las condiciones de operación. Será determinante, en primer lugar, el nivel de inmunidad que presenten ante estos cambios las características seleccionadas. El tamaño de la secuencia de entrenamiento, por otro lado, debe ser calibrado de modo que se evite el sobreajuste de los datos, es decir, teniendo en mente que una secuencia de entrenamiento larga resultará en complejas fronteras de decisión que no generalizarán bien los nuevos patrones a clasificar. A continuación, las bondades del sistema son evaluadas mediante el uso de un conjunto de patrones de la base de datos, independientes de los extraídos para la construcción de la secuencia de entrenamiento. Esta fase, diferenciada de la fase de entrenamiento, se conoce con el nombre de *fase de test*.

La señal de ruido asociada a un vehículo se caracteriza por una combinación de componentes deterministas con ruido de banda ancha de naturaleza estocástica, pudiendo ser considerada cuasiestacionaria para pequeños intervalos de tiempo (inferiores a 180 ms) [14]. Bajo estas condiciones, los atributos que caracterizan a un vehículo al paso pueden ser considerados invariantes en el tiempo, viniendo determinados principalmente por cuatro fuentes que originan su señal asociada y que son estudiadas en profundidad en [3]: ruido de propulsión, ruido de rodadura, ruido aerodinámico y ruido de expulsión. Esta señal asociada a cada paso aislado se ve alterada, a su vez, por otro tipo de factores: patrón de directividad, efecto suelo, distancia de la fuente al receptor, meteorología (presencia de hidrometeoros muy en particular), efecto Doppler, reverberación y estado del vehículo. Finalmente, la concurrencia espacio-temporal de estas señales da lugar a la señal de ruido de tráfico resultante, combinándose de modo interferente en su construcción. Para proceder a clasificar cada uno de los eventos presentes en la señal de la Figura 1, aparecen entonces tres problemas encadenados y que están asociados al proceso de segmentación aplicado habitualmente: (1) Aún procediendo a una segmentación manual óptima de la señal, el fuerte solapamiento de energía asociada a los vehículos involucrados provocará un notable desvío del vector de características extraído respecto de los valores mostrados para su paso aislado. (2) El entrenamiento sobre un conjunto de datos extraído desde una base de datos de pasos aislados, lo más habitual, presenta unas fronteras de separación interclase para las que el vector de características asociado a los eventos de la Figura 1 presenta importantes desviaciones. (3) El entrenamiento sobre un conjunto de datos *in situ* muestra unas fronteras de separación con una elevada incertidumbre, asociada principalmente a la variación de los resultados en relación al tipo de vehículos involucrados en los pasos combinados utilizados.

2.3 Enfoque de la solución

Por lo visto hasta ahora, se extrae en primer lugar que en condiciones reales de operación los procesos de detección y segmentación usuales impiden la obtención de un sistema de clasificación verosímil. Desde una red de sensores acústicos, como se ha visto, la señal de ruido de tráfico presenta un importante inconveniente, al componerse de una combinación de información asociada a pasos combinados de vehículos con la señal de ruido de fondo asociada a las características de la vía. El empleo aquí de técnicas tradicionales de segmentación impide afrontar con garantías el proceso de clasificación. La variabilidad interclase e intraclase mostradas desde esta perspectiva ha llevado a algunos autores [15] a cuestionar la posibilidad de encontrar un sistema de clasificación común a todas las clases. Por otra parte, el método de cálculo CNOSSOS-EU, a través de la guía recogida en [16], asigna una desviación máxima de 0.5 dB en el error asociado al conteo de vehículos. En este sentido, garantizar el umbral de precisión establecido por el método complica más la situación.

Para resolver los problemas descritos, se trazó la estrategia de incorporación de técnicas de separación de fuentes con las hipótesis de poder aumentar la precisión en la detección y de poder extraer de forma eficiente la información a entregar al sistema de clasificación. La evolución secuencial de la solución se describe a lo largo de las secciones siguientes. Ésta toma como punto de partida los resultados obtenidos en un trabajo preliminar llevado a cabo por nuestro grupo de investigación [4]. En este sistema previo, la señal de ruido de tráfico es adquirida a través del mismo array de sensores utilizado por el sistema actual. Tras la segmentación de la señal, el proceso de extracción se lleva a cabo mediante un conjunto de características de uso habitual en procesado de voz [17]: *Zero Crossing Rate* (ZCR), *Spectral Centroid* (SC), *Spectral Roll-off Point* (SR), *Spectral Band Energy Ratio* (SBER) y *Mel-Frequency Cepstral Coefficients* (MFCC). La extracción de la característica SBER se divide en 4 bandas y se consideran solamente los 5 primeros coeficientes MFCC. A continuación, se evalúan las prestaciones del conjunto de características en tareas de clasificación tras proponer el uso de dos algoritmos de clasificación: *k-Nearest Neighbour* (k-NN) [13] y *Fisher Linear Discriminant* (FLD) [18]. Para este último, se establecen dos estrategias de decisión (*One-vs-All* y *One-vs-One*). La base de datos utilizada para las fases de entrenamiento y test se compone de tres categorías de vehículos: motos, coches y camiones. Tras la evaluación de características con sendos algoritmos, se procede a su selección estableciendo la composición del vector de características final: {SR, SBER₃, SBER₄, MFCC}. Para este vector se obtienen los mejores resultados de clasificación: 3-NN (89.93 %) y FLD One-vs-One (88.18 %). Los resultados obtenidos, pese a ser prometedores, no son concluyentes debido a que la base de datos utilizada para su evaluación resulta insuficiente.

3 Separabilidad

El problema BSS fue planteado en [19], por la necesidad de separar un conjunto de señales generadas por el sistema nervioso central [20,21]. La extracción de estas señales desde la mezcla adquirida se caracteriza como ciega, al carecer de información acerca de la composición de fuentes generadoras, de su caracterización y del modo en que son mezcladas. En el contexto de ruido de tráfico, la señal adquirida se compone de una combinación de diversas fuentes que no pueden ser observadas de forma explícita. En condiciones reales, además, el proceso de mezcla es desconocido. Desde un enfoque matemático, sin embargo, existe solución al problema planteado siempre que las señales de entrada al proceso sean estadísticamente independientes. Para modelar matemáticamente la mezcla, la señal adquirida a través de la red de sensores puede ser expresada como

$$x(n) = H(s(n) + \varepsilon(n)), \quad (1)$$

siendo H una función dependiente de las fuentes que generan la señal, del canal y del array de sensores utilizado. La señal ε , por su parte, se corresponde con ruido gaussiano aditivo generado por el canal y la red de sensores, independiente de la señal s . Para el caso de ruido de tráfico, la expresión de la Ecuación (1) puede ser reescrita mediante un modelo de mezcla convolutivo. Esto es,

$$x(n) = [H(z)]s(n) + \varepsilon(n), \quad (2)$$

siendo $[H(z)]$ la transformada Z del filtro FIR que modela la mezcla.

3.1 Mezclas sintéticas

En una primera fase, se procedió a abordar un estudio de los beneficios a obtener con la incorporación de un proceso de separación de fuentes previo a la etapa de extracción, cuyos resultados se encuentran publicados en [5-6]. El estudio se lleva a cabo aplicando *Análisis de Componentes Independientes* (ICA), mediante el algoritmo en punto fijo Fast-ICA [22]. El algoritmo se aplica sobre mezclas sintéticas obtenidas a partir de pasos aislados de la base de datos utilizada en [4]. Para este caso el problema puede ser considerado lineal, permitiendo aproximar la Ecuación (1) por

$$x(n) = \mathbf{H} \cdot s(n). \quad (3)$$

El problema a resolver consiste en obtener $\hat{\mathbf{X}} \approx \mathbf{A} \cdot \mathbf{S}$, siendo $\hat{\mathbf{X}}$ el vector de fuentes y \mathbf{S} el vector de observaciones. La solución pasa por encontrar la matriz de separación \mathbf{B} , una estimación de la inversa de la matriz de mezcla \mathbf{A} . Para ello, el proceso se aborda en dos pasos: en una primera etapa, se procede a la proyección ortogonal de las mezclas de entrada mediante un proceso de decorrelación. Esta etapa simplifica entonces la solución a una rotación de los datos. La matriz de separación puede ser factorizada entonces como $\mathbf{B} = \mathbf{R} \cdot \mathbf{W}$, siendo \mathbf{W} una matriz de blanqueo y \mathbf{R} una matriz de rotación. El proceso de blanqueo se comienza sustrayendo de las muestras de entrada su media. El proceso concluye tras la aplicación de un proceso de ortonormalización sobre las muestras centradas a través de la *Descomposición en Valores Singulares* (SVD) [23]. Los resultados aplicados sobre mezclas sintéticas obtenidas desde la base de datos utilizada en [4] resultan positivos. De ellos se extraen además dos importantes conclusiones, a su vez relacionadas entre sí: (1) En primer lugar, la posibilidad de mejorar la separabilidad interclase mediante la aplicación previa de sustracción espectral de ruido. Esta posibilidad es validada en [7] mediante métricas estadísticas (curtosis y negentropía), poniendo de manifiesto la existencia de un enmascaramiento de información discriminante más agudizado en baja frecuencia. (2) En segundo lugar, la posibilidad de reducir la incertidumbre en las fronteras de separación interclase mediante un aumento del número de clases consideradas. Esta hipótesis es reforzada por los resultados obtenidos en [7], al deducirse que las fronteras de separación manejadas en [4] no resultan apropiadas.

3.2 Mezclas reales

Para analizar la contribución del vector de características propuesto en [4], se procedió en una segunda fase a comparar los resultados de clasificación con el uso de un vector de características extraído directamente del espectro de la señal. Debido a que este vector trabaja en un espacio de elevada dimensión, se abordó su reducción mediante el empleo de *Análisis de Componentes Principales* (PCA), de uso habitual en reconocimiento de patrones [13]. La clasificación se llevó a cabo mediante el residuo obtenido tras la proyección del vector de características sobre las direcciones de máxima varianza, para lo que se adaptó el algoritmo sugerido en [24]. Los resultados de este nuevo estudio se encuentran publicados en [8] y de ellos se extrae, en primer lugar, una importante reducción de prestaciones al repetir los experimentos llevados a cabo en [4] sobre mezclas sintéticas de pasos combinados. A excepción del clasificador 3-NN, para el que los resultados son comparables a los obtenidos por el clasificador basado en PCA, para el resto de clasificadores la precisión final se degrada de forma notable. Se comprueba, sin embargo, la supremacía del nuevo sistema de clasificación propuesto respecto a 3-NN al ser aplicado sobre mezclas reales. Se aborda entonces un estudio, cuyos resultados se encuentran publicados en [10], del que se extraen las limitaciones del uso de ICA para abordar el proceso de separación. Para el nuevo enfoque de la solución se sustituye el modelo de mezcla de la Ecuación (3) por el de la Ecuación (2).

3.3 Estrategia de diseño

Las conclusiones obtenidas a través de las investigaciones anteriores nos indujeron a trazar la siguiente estrategia: (1) Establecer la categorización óptima para la base de datos a obtener, (2) optimizar el proceso de separación en un dominio espectro-temporal, buscando al mismo tiempo abordar en este dominio el proceso de sustracción espectral [25], y (3) abordar un estudio acerca de la dependencia entre los procesos de reducción de dimensionalidad y de separación de fuentes. Para ello, con las conclusiones extraídas, se procedió a la construcción de una nueva base de datos, mejor dimensionada y con mayor resolución que la utilizada en [4]. Aunque las categorías manejadas en CNOSSOS-EU son cuatro, se propone un nuevo esquema de categorización según la Tabla 1. Con objeto de eliminar la incertidumbre asociada al etiquetado *online* utilizado en [4], de uso habitual por otra parte, para la construcción de la nueva base de datos se propuso un etiquetado *offline* mediante la adquisición de información lateral adicional. Ésta se obtuvo mediante la incorporación de información de vídeo sincronizada con la información adquirida mediante el array de sensores. Se demuestra entonces como este nuevo esquema, manejado internamente por el sistema de clasificación, conduce a una mayor separabilidad de las características extraídas.

Tabla 1 – Esquema de categorización usado internamente por el sistema de clasificación.

<i>Código</i>	<i>Descripción</i>
M1	Ciclomotores.
M2	Motocicletas, con o sin sidecar, triciclos, cuatriciclos, ...
L1	Vehículos ligeros con masa inferior a 3500 kg, o con 7 plazas como máximo.
L2	Resto de vehículos ligeros: furgonetas ligeras, SUV, pickup, ...
P1	Camiones ligeros, furgonetas pesadas y autobuses ligeros.
P2	Autobuses pesados, camiones pesados y vehículos especiales.
E	Vehículos eléctricos y vehículos híbridos en modo eléctrico.

Por otro lado, la optimización del proceso de separación se consigue mediante el uso del algoritmo de clasificación BSS-PCA propuesto en [9]. El éxito de su funcionamiento reside en su concepción, consiguiendo optimizar la extracción de información discriminante en un subespacio de dimensión reducida. Se consigue alcanzar una precisión de 94.83%, mejorando de forma drástica los resultados de clasificación anteriores. En este sentido, el algoritmo BSS-PCA permite reducir la incertidumbre asociada por el método CNOSSOS-EU al proceso de conteo, situando dicha desviación en 0.2 dB.

4 Teoría BISS de Separación Ciega de Fuentes

Las componentes deterministas de la señal de ruido de tráfico se originan en distintos procesos (combustión, escape de gases, paso de rueda, ...) asociados a los vehículos involucrados en la mezcla. Para un paso aislado, la información contenida en estas líneas espectrales queda alterada con energía producida por la interacción con el resto de la estructura (chasis, suspensión, frenos, ...), por las turbulencias creadas en los procesos de combustión y escape, y por el movimiento del vehículo (ruido aerodinámico). Este proceso conlleva un importante enmascaramiento de información, siendo crítico a velocidades altas donde el ruido de rodadura es predominante. La consideración de otros aspectos (patrón de directividad del vehículo, estado de mantenimiento/edad, conservación de la vía, efecto suelo, efecto Doppler, tipo de pavimento, distancia desde las fuentes a la red de sensores, condiciones atmosféricas y reflexiones de la señal sobre diferentes superficies cercanas a la vía), así

como el enmascaramiento producido por la combinación interferente de las señales asociadas a los vehículos presentes en la mezcla, complican de forma notable la situación. Como consecuencia, el problema BSS asociado se vuelve extremadamente complejo de resolver.

Desde la perspectiva de clasificación, el objetivo es extraer el vector de características del evento a clasificar para el que su señal asociada permanece oculta en la mezcla. En esta situación, y para el evento que se quiere clasificar, la señal recuperada mediante la aplicación de la teoría BSS en su forma clásica contiene información asociada al resto de eventos que contribuyen a la mezcla. Como alternativa, se propone en [10] abordar el proceso de extracción de características mediante la búsqueda de un subespacio de proyección que maximice la información discriminante. Para el *i-ésimo* evento a clasificar, ésta puede ser expresada a través del vector de observaciones como

$$s_{i,r}^i(n) = \Gamma_i(x(n)). \quad (4)$$

De este modo, el proceso de extracción consiste en encontrar el vector $\phi_i = [\phi_{i1}, \dots, \phi_{id}]^T$, pudiendo ser obtenido mediante la función ϕ como

$$\phi(s_{i,r}^i(n)) \xrightarrow{\phi} (\phi_{i1}(n), \dots, \phi_{id}(n)). \quad (5)$$

Desde la señal adquirida, el problema a resolver consiste entonces en encontrar la solución a la ecuación

$$\phi_i(n) = \phi(s_i(n)) = \phi(\Gamma_i(x(n))) = \phi(\Gamma_i([H(z)]s(n) + \varepsilon(n))). \quad (6)$$

Este nuevo problema no ha sido abordado hasta ahora por la teoría BSS. Encontrar su solución se vuelve una meta particularmente compleja de alcanzar cuando las fuentes presentan una naturaleza ruidosa, como ocurre para la señal de ruido de tráfico adquirida. Para abordarla, la investigación llevada a cabo en [3] conduce a una nueva teoría de *Separación Ciega de Fuentes Implícitas* (BISS) que se presenta en [10]. En esta publicación se introduce una caracterización completa del problema descrito, para el que el problema BSS en su forma clásica debe ser reformulado. Se presenta entonces el método BISS-PCA, cuya versión embebida es utilizada en [9] demostrando su robustez en tareas de clasificación.

5 Prestaciones del sistema actual

El sistema desarrollado en [3] permite, desde la versión publicada en [9], proporcionar la recolección completa de datos demandada por el método de cálculo CNOSSOS-EU. Esta recolección puede llevarse a cabo desde un mismo dispositivo portable dotado de un array incorporado de dos micrófonos, permitiendo adquirir durante el intervalo de medida toda la información requerida. Facilita, por un lado, la estimación de la velocidad instantánea, necesaria para el cálculo de la velocidad media de la vía, entregando ésta por sentido de circulación. Permite, por otro lado, entregar la intensidad de tráfico por sentido de circulación y categoría de vehículo. El resto de información a registrar puede ser introducida de forma manual. Desde el sistema publicado en [9], no obstante, se han introducido mejoras en diferentes direcciones que se encuentran en proceso de publicación: por un lado, los métodos actuales de estimación de la velocidad instantánea y de clasificación permiten al

método de cálculo CNOSSOS-EU manejar más de dos ejes viarios. Por otro lado, se han incrementado los resultados de clasificación, mostrando actualmente una precisión próxima al 100%.

6 Conclusiones

Los actuales métodos de cálculo destinados a la predicción de niveles de ruido necesitan ser alimentados con información precisa acerca de diferentes variables. Los requisitos de inmediatez, ubicuidad y coste hacen que la adquisición de esta información a través de sensores acústicos ofrezca un importante valor añadido. Para el caso del método de cálculo CNOSSOS-EU, actualmente en discusión, el sistema desarrollado en [3] permite la recolección compacta y en tiempo real de toda la información requerida mediante un sistema portable dotado de un sencillo array de dos micrófonos. Para su obtención, la tarea de clasificación se vuelve particularmente compleja atendiendo al esquema de categorías manejado y al nivel de precisión exigido. El enfoque dado en [3] al problema de clasificación de ruido de tráfico ha puesto de manifiesto diferentes problemas no abordados por el estado del arte. A este respecto, su solución es necesaria para la obtención de un sistema de clasificación verosímil capaz de operar en condiciones reales de funcionamiento. Como se ha visto, el problema parte del proceso de segmentación de la señal adquirida. Por un lado, los trabajos relacionados no proporcionan una solución óptima para esta etapa, limitándose al establecimiento de unos umbrales (determinados muchas veces de modo manual), que imposibilitan un conteo preciso de vehículos en tráfico real multipaso. Por otro lado, las soluciones aportadas por el estado del arte trabajan con señales de ruido de tráfico con suficiente separación entre pasos de vehículos, muchas veces limitadas a pasos aislados obtenidos de una base de datos. El tráfico real, sin embargo, es multipaso, con un elevado nivel de solapamiento en vías principales. En este sentido, y como ha sido mostrado al lector, resulta obligado abordar el problema de la presencia de información ajena al evento a clasificar.

Para resolver los problemas descritos, se propone por vez primera en el estado del arte la incorporación de técnicas de separación ciega de fuentes al problema de clasificación de ruido de tráfico. Se ha visto sobre el sistema de partida propuesto en [4] cómo el modo de aplicación de las mismas guarda una estrecha dependencia con las distintas etapas del sistema de clasificación. Se demuestra, en primer lugar, que la aplicación de técnicas ICA sobre mezclas sintéticas mejora la separación interclase [5-8]. El uso de este tipo de mezclas resulta lícito, al tratarse de un problema de clasificación, y no de separación de señales. Sobre mezclas reales, sin embargo, la información extraída no resulta óptima para ser manejada por el algoritmo de clasificación [8]. Se obtiene, por otro lado, que una mejor definición de clases manejada internamente por el sistema de clasificación redundante en una mejor separación del vector de características [9]. Se demuestra, por último, cómo proceder a la selección de las características que componen este vector limita los beneficios a obtener con la aplicación previa de técnicas de separación ciega de fuentes [10]. El principal problema recae en cómo la teoría BSS maneja la información discriminante, presentando una importante limitación en su aplicación sobre señales de naturaleza ruidosa. Enfrentar este problema, no abordado hasta ahora, ha supuesto uno de los mayores esfuerzos del trabajo de investigación desenvuelto en [3]. Como resultado, esta investigación da lugar a la *Teoría de Separación Ciega de Fuentes Implícitas* (BISS) que se presenta en [10]. Desde esta nueva teoría, la solución a este problema deriva en el método BISS-PCA, que manejado de modo embebido por el sistema presentado en [9] permite alcanzar una elevada precisión de clasificación (94.83%). Los resultados obtenidos en esta dirección son aplicables a otros campos de investigación, suponiendo por ello una de las contribuciones más relevantes al estado del arte del trabajo de investigación llevado a cabo.

Por último, la precisión alcanzada por el sistema presentado en [9] ha sido recientemente mejorada, encontrándose próxima al 100% en el momento actual. Este sistema maneja como solución común al conjunto de clases manejado por el método CNOSSOS-EU, un sólo vector de características de dimensión reducida y un sólo algoritmo de clasificación. Se demuestra con ello, al contrario de lo manifestado en [15] por sus autores, que un sistema así es posible. Su obtención, como ha podido observar el lector, ha requerido sin embargo abordar aspectos del problema de clasificación no tratados hasta ahora. Por otro lado, las mejoras introducidas en los métodos de estimación de la velocidad instantánea y de clasificación permiten al método CNOSSOS-EU manejar más de dos ejes viarios. Esperamos que estos últimos avances de investigación se encuentren publicados en breve, invitando desde este momento al lector a su crítica lectura.

Agradecimientos

Este trabajo ha sido financiado parcialmente por el MCYT de España, ref.: TEC2009-14414-C03-03, dentro del proyecto de investigación “*Combinación de Técnicas de Separación de Fuentes Multicanal para la Reducción de Ruido*”.

Referencias

- [1] International Organization for Standardization. ISO 1996-2: 2007. *Description, Measurement and Assessment of Environmental Noise. Part 2: Determination of Environmental Noise Levels*. Switzerland; 2007.
- [2] *Draft JRC Reference Report I on Common NOise ASSESSMENT MethOdS in EU (CNOSSOS-EU)*.
- [3] Fernando J. Mato-Méndez. *Aplicación de Técnicas de Separación de Fuentes y Reconocimiento de Patrones al Conteo y Clasificación de Ruido de Tráfico*. Ph. D. Thesis, Universidad de Vigo; 2012.
- [4] M. A. Sobreira, A. R. Molaes, J. L. Alba Castro. *Automatic Classification of Traffic Noise*. In Proceedings of the International Congress on Acoustics, volume 123, pp 6221-6226. París; 2008.
- [5] F. J. Mato-Méndez. *Automatic Segmentation of Traffic Noise*. In Proceedings of the International Congress on Acoustics, volume 123, pages 5867–5872. París; 2008.
- [6] F. J. Mato-Méndez, M. A. Sobreira-Seoane. *Análisis de Componentes Independientes en Separación de Fuentes de Ruido de Tráfico*. In Proceedings of the International European Acoustics Symposium, Congreso Ibérico de Acústica-Tecniacústica 2008. Coimbra; 2008.
- [7] F. J. Mato-Méndez, M. A. Sobreira-Seoane. *Sustracción Espectral de Ruido en Separación Ciega de Fuentes de Ruido de Tráfico*. In Proceedings of the International European Acoustics Symposium, Congreso Ibérico de Acústica-Tecniacústica 2008. Coimbra; 2008.
- [8] F. J. Mato-Méndez, M. A. Sobreira-Seoane. *Detecting Multiple Simultaneous Vehicles Pass-by by Sound Source Separation Techniques*. In Proceedings of INTERNOISE, Lisbon, Portugal; June 2010.
- [9] F. J. Mato-Méndez, M. A. Sobreira-Seoane. *Blind Separation to Improve Classification of Traffic Noise*. Applied Acoustics, vol. 72, no. 8 (Special Issue on Noise Mapping), pages 590–598; 2011. (Invited paper).

- [10] F. J. Mato-Méndez, M. A. Sobreira-Seoane. Independent Component Analysis for Audio and Biosignal Applications. Chapter: *Blind Implicit Source Separation. A new concept in BSS Theory*. InTech, Wien, Austria; 2012. ISBN 980-953-307-197-3. (Invited chapter).
- [11] L. J. Díaz. *Modelo Computacional para el Análisis Avanzado de Señales de Ruidos Ambientales*. M. S. Thesis, Instituto Politécnico Nacional de México; 2011.
- [12] M. G. López. *Clasificador Difuso de Señales Acústicas Ambientales basado en Análisis de Componentes Independientes*, M. S. Thesis, Instituto Politécnico Nacional de México; 2012.
- [13] K. Fukunaga. *Introduction to Statistical Pattern Recognition*. Second Ed., Academic Press Professional, Inc; 1990.
- [14] V. Cevher, R. Chellappa and J. H. McClellan. *Vehicle Speed Estimation using Acoustic Wave Patterns*, IEEE Transactions on Signal Processing 57(1): 30–47; 2009.
- [15] B. Defréville, P. Roy, C. Rosin and F. Pachet. *Automatic Recognition of Urban Sound Sources*. In *Proceedings of the AES 120th Convention, Paris, France; 2006*.
- [16] *Good Practice Guide for Strategic Noise Mapping and the Production of Associated Data on Noise Exposure*. Position paper. version 2. Technical report, European Commission Working Group Assessment of Exposure to Noise; 2007.
- [17] G. Tzanetakis and P. Cook. *Musical Genre Classification of Audio Signals*. IEEE Transactions on Speech and Audio Processing 10(5): 293–302; 2002.
- [18] R. A. Fisher. *The Use of Multiple Measurements in Taxonomic Problems*. Annals of Eugenics 7: 179–188; 1936.
- [19] J. Herault, C. Jutten and B. Ans. *Détection de Grandeurs Primitives dans un Message Composite par une Architecture de Calcul Neuromimétique en Apprentissage Non Supervisé*. In *Proceedings of the X GRETSI Symposium on Signal and Image Processing, Nice, France, pp. 1017–1022; 1985*.
- [20] B. Ans, J. C. Gilhodes and J. Herault. *Simulation de Réseaux Neuronaux (Sirene). II. Hypothèse de Décodage du Message de Mouvement Porté par les Afférences Fusoriales IA et II par un Mécanisme de Plasticité Synaptique*. Comptes Rendus de l'Académie des Sciences Paris-Serie III (Sciences Naturelles), Vol. 297, pp. 419–422; 1983.
- [21] J. Herault and B. Ans. *Neural Network with Modifiable Synapses: Decoding of Composite Sensory Messages under Unsupervised and Permanent Learning*. Comptes Rendus de l'Académie des Sciences Paris-Serie III (Sciences Naturelles), Vol. 299, pp. 525–528.; 1984.
- [22] A. Hyvärinen. *Fast and Robust Fixed-Point Algorithms for Independent Component Analysis*. IEEE Transactions on Neural Networks 10(3): 626–634; 1999.
- [23] C. Eckart and G. Young. *The Approximation of One Matrix by Another of Lower Rank*. Psychometrika 1: 211–218; 1936.
- [24] H. Wu, M. Siegel and P. Khosla. *Vehicle Sound Signature Recognition by Frequency Vector Principal Component Analysis*. IEEE Transactions on Instrumentation and Measurement 48(1): 1005–1009; 1999.
- [25] S. Rickard, R. Balan and J. Rosca. *Real-Time Time-Frequency based Blind Source Separation*. *Proceedings of the International Conference on Independent Component Analysis and Blind Source Separation, San Diego, USA, pp. 651–656; 2001*.