

# Clasificación automática de obstáculos empleando escáner láser y visión por computador

Aurelio Ponz. Intelligent Systems Lab, Universidad Carlos III de Madrid, aurelio.ponz@uc3m.es  
Fernando Garcia. Intelligent Systems Lab, Universidad Carlos III de Madrid, fegarcia@ing.uc3m.es  
David Martin. Intelligent Systems Lab, Universidad Carlos III de Madrid, dmgomez@ing.uc3m.es  
Arturo de la Escalera. Intelligent Systems Lab, Universidad Carlos III de Madrid, escalera@ing.uc3m.es  
Jose Maria Armingol. Intelligent Systems Lab, Universidad Carlos III de Madrid, armingol@ing.uc3m.es

## Resumen

*Muchos sistemas enmarcados en el estado del arte del campo de los Sistemas Avanzados de Asistencia a la Conducción (ADAS) y de la conducción autónoma emplean fusión sensorial con el fin de conseguir detección y clasificación de obstáculos fiable en cualquier condición meteorológica y de iluminación. La fusión entre escáner láser y cámara se usa habitualmente en aplicaciones ADAS para mitigar las limitaciones inherentes a cada uno de los sensores empleados. En el sistema presentado se emplean algunas técnicas novedosas para alineamiento de datos y se aplican técnicas de Inteligencia Artificial (IA) en el tratamiento de las nubes de puntos para mejorar la fiabilidad de la clasificación de obstáculos. En este documento se presentan nuevos enfoques para la obtención de clusters en nubes de puntos dispersas, maximizando la información obtenida desde escáneres láser de baja resolución. Tras la mejora de la detección de clusters, se emplean técnicas de IA para clasificar el obstáculo no solo empleando visión por computador, sino también con información del láser. La fusión de la información obtenida desde ambos sensores, con la adición de la capacidad de clasificación del láser, mejoran la fiabilidad del sistema.*

**Palabras clave:** Visión por computador, LIDAR, Sistemas avanzados de asistencia a la conducción, Fusión sensorial.

## 1. INTRODUCCIÓN

Los accidentes de tráfico se encuentran entre los riesgos evitables más importantes para la vida humana. Alrededor de 1,2 millones de personas mueren, y otros 50 millones sufren secuelas cada año como consecuencia de accidentes de tráfico. Los sistemas ADAS pueden reducir la cantidad y la gravedad de los accidentes de tráfico mediante el uso de tecnologías de visión por computador (VC) e IA. Los ADAS emplean comúnmente escáneres láser y cámaras para detectar y clasificar obstáculos en la vía. Los sensores presentados son complementarios, ya que la capacidad del láser para

detectar obstáculos independientemente de la calidad de la iluminación y de seleccionar Regiones de Interés (ROI) en la imagen en la que se clasificará mediante VC, mejora de forma importante la velocidad y la precisión en esta fase del proceso.

El presente trabajo ha sido desarrollado usando la plataforma de investigación Intelligent Vehicle based on Visual Information 2.0 (IVVI 2.0) del Laboratorio de Sistemas Inteligentes (LSI) de la Universidad Carlos III de Madrid. Esta plataforma está detallada en profundidad en [11] y se puede ver en la figura 1.

El artículo está dividido en las siguientes secciones: La sección 2 presenta el estado del arte en el contexto científico relacionado. La sección 3 muestra una descripción general del sistema. La sección 4 describe el método de obtención de clusters en las nubes de puntos (PC), como parte inicial del proceso de detección de obstáculos. La sección 5 presenta el proceso de alineamiento de datos, esencial para la correcta conversión entre coordenadas de los sistemas de referencia del láser y de la cámara. La sección 6 muestra la estrategia para la clasificación de los obstáculos detectados con las nubes de puntos e imágenes asociadas, empleando Support Vector Machines (SVM). La sección 7 presenta los resultados obtenidos, y, para concluir, la sección 8 muestra las conclusiones obtenidas del presente trabajo.

## 2. ESTADO DEL ARTE

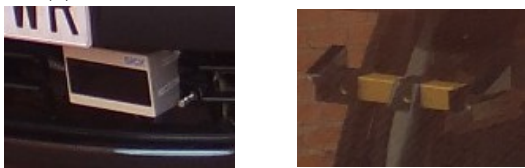
El trabajo descrito en el presente artículo abarca diversos campos con interesantes estados del arte. En cuanto al proceso de alineamiento automático y desatendido de datos en el sistema presentado, [8] propone un método de calibración empleando un patrón de tablero de ajedrez; [14] propone la calibración automática entre cámara y láser basándose en la reconstrucción mediante la nube de puntos de la superficie de la carretera. Otras aproximaciones como [8] y [7] proyectan las características en un plano 2D para minimizar las distancias entre las características en los diferentes sensores. [10] Presenta un sistema de calibración basado en un modelo CAD para conversión entre

sensores. Una aproximación similar, basada en un modelo triangular, es presentada en [1], mientras que en [3] se emplea un modelo circular.

Una vez que los diferentes sistemas han sido alineados, el siguiente paso es la detección y clasificación de los obstáculos presentes en la escena. Los sistemas de fusión sensorial pueden ser divididos en esquemas centralizados y descentralizados. Los primeros obtienen una única detección basándose en la información obtenida de los diferentes sensores. Algunos ejemplos de esquema descentralizado se pueden encontrar en [12] y [13] con diferentes algoritmos para combinar las características de la VC y del escáner láser, como Naive Bayes, GMMC, NN, FLDA. Los sistemas descentralizados implementan la detección y la clasificación de forma independiente para cada sensor, y una fase posterior de fusión combina las detecciones de acuerdo con el grado de certeza asignado a cada sensor. Otros trabajos defienden la fusión a alto nivel basándose en características multidimensionales para el escáner láser y en Histograma de Gradientes Orientados (HOG) para VC. El trabajo [5] ofrece detección de peatones basada en un modelo de las piernas del peatón para el láser, y en características HOG para VC, buscando la detección distribuida de peatones y la evaluación del peligro. El mismo autor hace uso de técnicas avanzadas de fusión (Joint Probabilistic Data Association Filter) para mejorar la detección de peatones descentralizada empleando láser y VC.



(a) Plataforma de investigación IVVI 2.0



(b) Láser Sick LD-MRS (c) Cámara trinocular XB3

Figura 1: Sensores de la plataforma de investigación IVVI 2.0

### 3. DESCRIPCIÓN GENERAL

El presente trabajo emplea fusión sensorial entre escáner láser y VC para la detección y clasificación de obstáculos en aplicaciones de automoción.

Este trabajo está incluido en el proyecto IVVI 2.0 (ver figura 1). IVVI 2.0 es la segunda plataforma para investigación y desarrollo del Laboratorio de Sistemas Inteligentes, de la Universidad Carlos III de Madrid.

En la presente aplicación se ha empleado un sensor láser SICK LD-MRS de 4 planos y una cámara trinocular PoinGrey Bumblebee XB3. El láser efectúa una detección inicial de obstáculos y otra fase posterior de clasificación, mientras que la capacidad estereoscópica de la cámara se usa para representar la nube de puntos del plano de la carretera y poder con ello estimar los parámetros extrínsecos entre los sensores. Posteriormente, la cámara es empleada para capturar la imagen y posteriormente tratarla mediante VC.

El láser genera una nube de puntos de la que el sistema extrae los obstáculos en forma de agrupamientos de puntos, llamados clusters. Estos clusters son usados tanto para obtener la ROI en la imagen como para efectuar la clasificación del obstáculo. La ROI extraída de la imagen es procesada para clasificar el obstáculo representado en ella, empleando técnicas de IA aplicadas a la VC. El último paso del proceso fusiona la información de cámara y láser para llevar a cabo la clasificación final del obstáculo basándose en técnicas de Aprendizaje Automático (ML). Para el entrenamiento y puesta a punto de los clasificadores SVM para imágenes y para PC se ha empleado una base de datos de obstáculos etiquetados como imágenes y como nubes de puntos.

### 4. EXTRACCIÓN DE CLUSTER DE LA NUBE DE PUNTOS PARA DETECCIÓN MEDIANTE LÁSER

El primer paso en el sistema propuesto es la detección de obstáculos empleando nubes de puntos obtenidas del láser. Al no verse afectado por las condiciones de iluminación, se trata del sensor más fiable en nuestro sistema. El sensor láser de cuatro planos proporciona una nube de puntos que representa la escena situada frente al vehículo. Los obstáculos son parte de esta escena, y pueden ser localizados como concentraciones locales de puntos en la nube, que pueden ser categorizados matemáticamente como clusters.

Con el objetivo de obtener de la nube de puntos la

mayor cantidad posible de información posible, se han estudiado diversas técnicas de extracción de clusters. Es importante remarcar que, debido a limitaciones propias del sensor, los obstáculos a detectar estarán representados por una pequeña cantidad de puntos de la nube, habitualmente desde cuatro puntos hasta no muchos más de cincuenta, dependiendo de la distancia desde el sensor hasta el obstáculo. La mayor parte de las estrategias de extracción de clusters disponibles han sido diseñadas para nubes de puntos muy pobladas, habitualmente obtenidas desde escáneres láser multicapa de alta resolución, o cámaras estéreo, y no se adaptan bien a las nubes de puntos obtenidas en nuestro sistema, obtenidas en exteriores, dispersas y que ofrecen información muy limitada.

El láser SICK LD-MRS ofrece diferentes frecuencias de escaneo con diversas resoluciones angulares. La frecuencia menor, de 12,5 Hz, permite una resolución angular de  $0,125^\circ$  en el frente del vehículo,  $0,25^\circ$  entre los  $10^\circ$  y los  $30^\circ$ , y  $0,5^\circ$  entre los  $30^\circ$  y los  $50^\circ$  ( $60^\circ$  si se trata del lado derecho de la escena), como se ilustra en la figura 2. Esta configuración incrementa la capacidad de detección enfrente del vehículo, donde los obstáculos tienden a estar situados a mayor distancia. Para aplicaciones de automoción, resoluciones menores en los laterales de la escena son aceptables, ya que los obstáculos de interés suelen estar situados más cerca que en el frontal, y serán representados por suficientes puntos incluso con resoluciones menores.

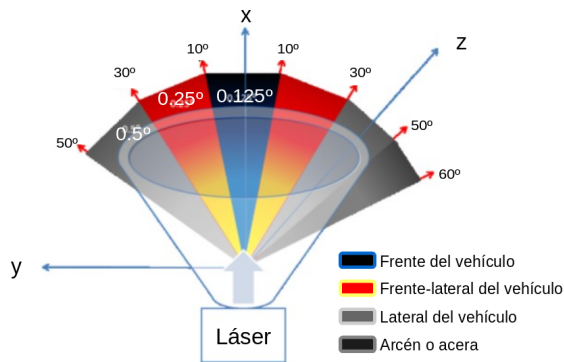


Figura 2: Resolución angular variable en el láser

Las distancias entre los puntos detectados en el cuadro 1 explican la necesidad de diseñar una estrategia de adaptación del umbral de pertenencia a un cluster dependiendo de la distancia entre el obstáculo y el sensor, con el fin de obtener clusters lo más poblados posible. El significado de los valores en el cuadro 1 es el siguiente [15]:

- $Y_S$  es el ancho del punto medido
- $Y_G1$  es la distancia entre puntos en un plano

de medición

- $Y_G2$  es la distancia entre puntos entre dos pulsos del láser
- $X_{layer}$  es la altura del punto medido

Cuadro 1: Distancias entre puntos medidos con una resolución angular de  $0.125$  grados

Distancia (m)	$Y_S$	$Y_G1$	$Y_G2$	$X_{layer}$
10	0.014	0.029	0.007	0.139
25	0.035	0.074	0.019	0.349
50	0.069	0.148	0.039	0.698
100	0.139	0.296	0.078	1.396

Las siguientes secciones explicarán las diferentes técnicas desarrolladas para obtener una detección automática avanzada de clusters en las nubes de puntos.

#### 4.1. Distancia Euclídea adaptada y clusters geoméricamente condicionados

En el enfoque presentado se ha empleado como criterio para formación de clusters la existencia de una distancia euclídea máxima entre los pares de puntos pertenecientes al cluster, si bien esta distancia será modulada en función de la distancia del sensor hasta el obstáculo, condiciones geométricas, máximo número de puntos permitidos en un cluster, etc.

Adicionalmente, ciertos parámetros del proceso de clustering, como la distancia máxima entre puntos en un cluster, se ven modificados de acuerdo con las formas detectadas en la zona de la nube de puntos cercana al cluster en formación, con el fin de mejorar las detecciones de obstáculos oblicuos. Se ha probado también una estrategia alternativa empleando la distancia de Mahalanobis como criterio de formación de los cluster. Este método tiende a obtener clusters compactos, ignorando los puntos adicionales que representarían obstáculos posicionados oblicuamente. Esta estrategia ha sido descartada puesto que nuestro sistema suministra clusters pequeños, por lo que se requieren estrategias que incrementen la información obtenida.

La aproximación presentada define los clusters como un conjunto de puntos separados entre sí una cierta distancia, variable en función de diversos parámetros, a los que se añaden puntos que no cumplen los requisitos de distancia pero sí ciertas restricciones geométricas respecto al cluster al que se añaden, como por ejemplo pertenecer a la misma recta 3D que algunos de los puntos del cluster.



La estrategia de clustering se define como la adición iterativa de puntos a un cluster ejecutando los siguientes pasos:

El primer punto de la nube es considerado como el primer punto en el cluster.

La distancia euclídea de todos los puntos restantes en la nube con el punto inicial es comparada con el valor  $ClusterTh$  siguiendo la ecuación 1,

$$ClusterTh = BaseTh + DistCorr(x)$$

$$DistCorr(x) = \sqrt{(x * \tan(\alpha_y))^2 + (x * \tan(\alpha_z))^2}$$

$$\begin{aligned} \text{if } |\arctan\left(\frac{y}{x}\right)| < 2\pi \frac{10}{360} \\ \text{then } \alpha_y &= 2\pi \frac{0,125}{360} \\ \text{if } 2\pi \frac{10}{360} \leq |\arctan\left(\frac{y}{x}\right)| < 2\pi \frac{30}{360} \\ \text{then } \alpha_y &= 2\pi \frac{0,25}{360} \\ \text{if } 2\pi \frac{30}{360} \leq |\arctan\left(\frac{y}{x}\right)| < 2\pi \frac{60}{360} \\ \text{then } \alpha_y &= 2\pi \frac{0,5}{360} \end{aligned} \quad (1)$$

$x, y, z$  son coordenadas de puntos.

Por restricciones del sensor,  $\alpha_z$  siempre es 0,8.

donde  $BaseTh$  es un parámetro cuyo valor ha sido determinado experimentalmente.  $DistCorr(x)$  es una función de la coordenada  $x$  (distancia hasta el obstáculo) que asegura que la distancia euclídea empleada nunca es inferior a la distancia mínima posible entre los puntos del láser a esa distancia, según se ve en la ecuación 1, y dependiendo de las diferentes resoluciones angulares mostradas en la figura 2.  $DistCorr(x)$  se calcula como la mínima distancia posible entre dos puntos consecutivos en las coordenadas  $z$  (eje vertical) e  $y$  (eje horizontal perpendicular al sentido de la marcha).  $\alpha_y$  Representa el ángulo entre dos lecturas consecutivas del láser en horizontal (eje  $y$ ), y  $\alpha_z$  es el ángulo entre dos lecturas consecutivas del láser en vertical (eje  $z$ ).

Todos los puntos en la nube se comprueban para su inclusión en el cluster. La misma iteración se lleva a cabo para cada uno de los puntos añadidos al cluster, hasta que todas las combinaciones posibles de comprobaciones han sido agotadas. En ese momento, los puntos cercanos al cluster pero no incluidos en él por no cumplir la restricción de

distancia, se incluyen en una nube temporal junto con el cluster obtenido. En esta nube temporal se ejecuta el algoritmo Random sample consensus (RANSAC) para búsqueda de rectas 3D. Si las rectas obtenidas contienen un determinado número mínimo de puntos pertenecientes al cluster original, y además puntos no pertenecientes al cluster original, asumimos que estos puntos deben ser añadidos al cluster original puesto que representan el mismo obstáculo. Esta técnica se ha mostrado útil en obstáculos oblicuos cuyos puntos se distancian uniformemente entre ellos según aumenta la distancia al sensor.

La figura 3 muestra el resultado del algoritmo. Los puntos rojos y los verdes forman parte del cluster original. Los puntos azules se encuentran cerca del cluster pero no forman parte de él. Las líneas azules son líneas 3D encontradas por RANSAC en el cluster original, y si alguno de los puntos cercanos al cluster pertenece también a las líneas azules, es añadido al cluster extendido, como sucede en este caso con los puntos azules. Tras completar la extracción del cluster, éste es comparado con los parámetros  $ClusterTolerance$  para la máxima anchura horizontal del cluster en metros, y  $maxClusterSize$  y  $minClusterSize$  para el máximo y mínimo número de puntos, respectivamente. Estos parámetros son también función de la distancia hasta el obstáculo.

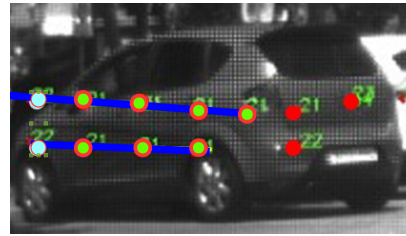


Figura 3: Cluster extendido. Los puntos azules se añaden al cluster porque comparten una línea con puntos del cluster

El objetivo de la estrategia es obtener clusters lo más poblados posible, teniendo en cuenta que se usa un láser multicapa de baja resolución. La distancia umbral de clusterización debe ser adaptada en función de la distancia  $x$  al obstáculo, puesto que la distancia mínima entre puntos consecutivos del láser crece con  $x$ . Debido a limitaciones de la construcción del láser, la distancia mínima detectada en  $y$  y en  $z$  en puntos consecutivos sería mayor que el umbral inicial si no se adaptara siguiendo la ecuación 1.

#### 4.2. Detección del suelo y eliminación en la nube de puntos

Como se mostrará en la sección 5, el sistema permite calcular el plano correspondiente a la superficie de la carretera, por lo que es posible eliminar de la nube de puntos aquéllos que pertenezcan a la carretera e ignorarlos de cara al cálculo de los clusters.



Figura 4: Los puntos en el plano de la carretera cumplen las restricciones para ser considerados cluster, pero son descartados

#### 5. Alineamiento de datos

El sistema presentado se basa en fusión sensorial entre diversos sensores, que a su vez se basan en distintos fenómenos físicos. Cada uno de estos sensores poseen su propio sistema de referencia, por lo que se debe calcular los parámetros extrínsecos de conversión entre los sistemas de referencia de los sensores para proceder al alineamiento de los datos.

Aplicando el algoritmo M-estimator-SAmple-Consensus (MSAC) [18], aplicado a la detección de un plano en el espacio, es posible generar un vector  $[a, b, c, d]$  que defina el plano más poblado en la nube de puntos, que podemos asumir que coincidirá con el plano de la carretera.

$$\pi_{(x)} : ax_c + by_c + cz_c + d = 0 \quad (2)$$

$$\pi_{(x)} : \vec{n} \cdot \vec{p} = h \quad (3)$$

La ecuación 2 se puede escribir en su forma Hessiana como 3, donde  $\vec{n}$  es el vector normal al plano de la carretera, y la relación entre este vector y los ángulos de rotación de cámara y láser se pueden calcular como se muestra en el trabajo [19].

De la ecuación 3 deducimos que el vector  $\vec{n}$  es normal al plano  $\pi_{(x)}$  encontrado, puesto que la proyección en  $\vec{n}$  de cualquier punto localizado en el plano genera siempre una distancia fija. Esta distancia es mínima desde el plano al origen de coordenadas de la nube de puntos, por lo que se trata de la altura  $h$  del sensor.

Una vez que todos los parámetros extrínsecos de calibración entre los sensores roll, pitch, yaw y  $x, y, z$  han sido calculados, el sistema es capaz de trasladar las coordenadas del láser a coordenadas de la cámara, de forma que se pueden emplear técnicas de VC para clasificar los obstáculos en la imagen.

La conversión entre coordenadas del láser y de la cámara se puede conseguir empleando la ecuación 4

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + T$$

$$R = \begin{bmatrix} \cos(\delta) & 0 & \sin(\delta) \\ 0 & 1 & 0 \\ -\sin(\delta) & 0 & \cos(\delta) \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 1 \\ 0 & \cos(\phi) & -\sin(\phi) \\ 0 & \sin(\phi) & \cos(\phi) \end{bmatrix} \cdot \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

$$T = \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix}$$

donde  $T$  representa el vector de traslación y  $R$  la matriz de rotación entre sensores.



Figura 5: Detección de obstáculo basado en cálculo de cluster. Rectángulo azul oscuro es el cluster, rectángulo azul claro es la ROI para VC

## 6. Clasificación de obstáculos usando fusión sensorial entre láser e imagen

La clasificación de obstáculos se puede obtener en este trabajo empleando fusión sensorial o utilizando uno solo de los sensores. Se ha desarrollado un entorno de desarrollo que permite la comparación directa de resultados y mejoras rápidas en los entrenamientos.

### 6.1. Clasificación SVM

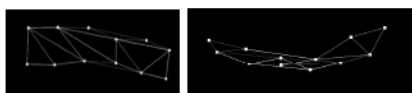
La clasificación se consigue con la implementación del algoritmo SVM de la librería de VC de OpenCV. El algoritmo SVM fue desarrollado por Vapnik & Cortés [2] y es ampliamente usado en aprendizaje por computador como método de clasificación. En el presente trabajo se ha empleado para efectuar los entrenamientos una base de datos de imágenes etiquetadas manualmente, además de otra base de datos de clusters, obtenidos ambos de una captura ejecutada con la plataforma IVVI 2.0. Tras el proceso de entrenamiento, el clasificador SVM obtenido es almacenado y empleado posteriormente para la clasificación de imágenes y clusters, como se ve en la figura 7.

### 6.2. Vector de características del láser

Los clusters detectados en la nube de puntos generada por el láser se emplean no solo para determinar una región de interés (ROI) en la imagen, en la que posteriormente se efectúa una clasificación empleando técnicas de IA, sino también se pueden usar para efectuar clasificación de obstáculos sin soporte de imagen [12].

Los clusters se convierten en una estructura mallada empleando la triangulación de Delaunay con el fin de reconstruir la forma del objeto y de extraer características relevantes, como se ve en la figura 6.

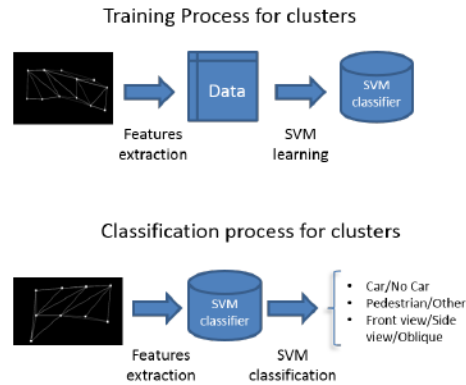
Estos obstáculos son detectados por el sistema en forma de clusters, y tienen características apropiadas para efectuar un entrenamiento SVM siguiendo el proceso descrito en la figura 7a. Los clusters obtenidos en las secuencias de prueba son almacenados y etiquetados de acuerdo con las imágenes asociadas de la cámara.



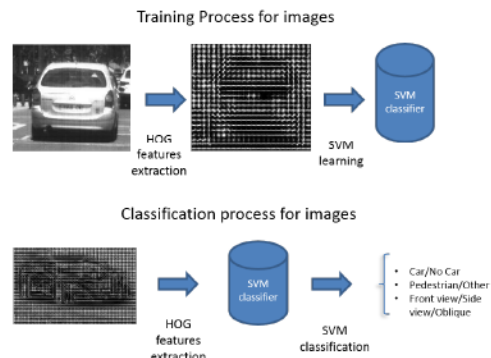
(a) Vista frontal (b) Vista cenital

Figura 6: Representación mallada de un cluster. Los triángulos representan la superficie del objeto.

Previos trabajos como [12] consideran nubes de puntos 2D para clasificación, mientras que el presente trabajo extrae las características desde una nube de puntos 3D, en un esfuerzo por maximizar la información disponible. Algunas de las características consideradas para los clusters son la concentración global de los puntos, la concentración de los puntos en cada uno de los planos, la planicidad, la esfericidad, la cubicidad, la triangularidad o la desviación típica de los puntos, todo ello normalizado y calculado respecto al centroide del cluster.



(a) Proceso de aprendizaje SVM para clusters: Entrenamiento y clasificación.



(b) Proceso de aprendizaje SVM para imágenes: Entrenamiento y clasificación.

Figura 7: Proceso de aprendizaje SVM

### 6.3. Vector de características para visión por computador

Los obstáculos detectados en la nube de puntos del láser se usan para determinar ROIs en las imágenes, susceptibles de ser clasificadas empleando clasificadores SVM. Empleando las ROI detectadas se ha generado una base de datos de imágenes etiquetadas. En un paso posterior, se extraen de las imágenes las características HOG y se efectúa un entrenamiento SVM siguiendo el proceso mostrado en la figura 7b, con el fin de obtener un clasificador SVM [17].



### 6.4. Fusión de la información

El proceso de fusión sensorial se produce de la forma siguiente: El láser genera una nube de puntos en la que se detectan los obstáculos en forma de clusters. Estos clusters se emplean para una clasificación inicial del obstáculo y para la generación de una ROI en la imagen. Las ROI obtenidas son clasificadas empleando técnicas de IA aplicadas a la VC. En un último paso, la información obtenida de los procesos de clasificación basados en PC y de los basados en imagen son fusionados para obtener una clasificación consolidada.

## 7. Resultados

Se ha llevado a cabo un estudio de relevancia de las características consideradas para la clasificación SVM de los clusters, usando un conjunto de entrenamiento de 14.000 clusters representando peatones y 8.400 clusters de otros obstáculos. Sólo las características que diferencien nítidamente entre muestras positivas y negativas serán usadas para el entrenamiento. La figura 8 muestra algunas características con buena discriminación entre muestras positivas y negativas.

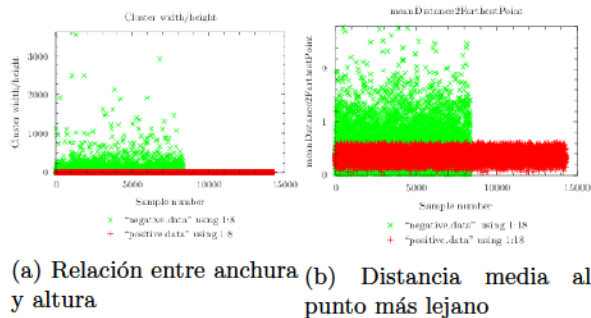
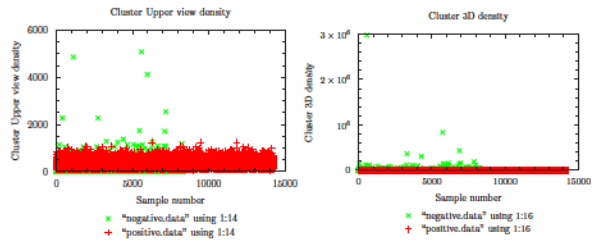


Figura 8: Distribución de los valores de características que diferencian correctamente muestras positivas de negativas

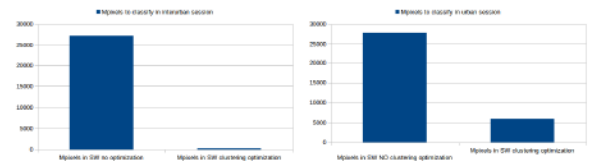
La figura 9 representa las estadísticas para características que no discriminan correctamente entre muestras positivas y negativas. La figura 9a indica que la característica *Densidad del cluster desde vista superior* no discrimina correctamente positivos y negativos, ya que gran parte de los valores coinciden para muestras positivas y negativas.



(a) Densidad desde vista superior (b) Densidad 3D superior

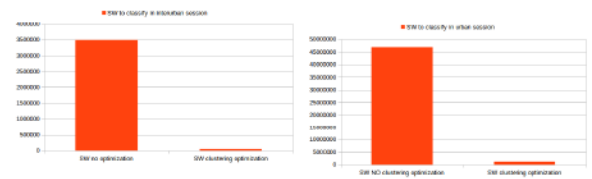
Figura 9: Distribución de los valores de características que no diferencian correctamente muestras positivas de negativas

El uso de fusión sensorial entre láser y cámara permite reducir considerablemente el esfuerzo computacional de clasificación por imagen, al reducir las áreas a clasificar desde toda la extensión de las imágenes a sólo la parte de ellas que presentan un obstáculo, conociendo además el tamaño esperado del mismo. En la figura 10 se compara los Mpixels a clasificar en una sesión de conducción según se emplee o no fusión sensorial.



(a) Mpixels a clasificar en una sesión interurbana (b) Mpixels a clasificar en una sesión urbana

Figura 10: Mejora de prestaciones de la ventana deslizante empleando fusión sensorial



(a) Número de ventanas a clasificar en una sesión interurbana (b) Número de ventanas a clasificar en una sesión urbana

Figura 11: Mejora de prestaciones de la ventana deslizante empleando fusión sensorial

## 8. Conclusión

El objetivo del presente trabajo es desarrollar un sistema de fusión sensorial para láser y cámara que aproveche las fortalezas de cada uno de los sensores y mitigue sus debilidades. Las estadísticas obtenidas demuestran que se ha obtenido un

sistema fiable y robusto capaz de detectar y clasificar obstáculos tanto aprovechando la fusión sensorial como empleando cada uno de los sensores por separado en caso necesario.

### Agradecimientos

El presente trabajo está financiado por el gobierno de España a través de los proyectos CICYT (TRA2013-48314-C3-1-R) y (TRA2011-29454-C03-02).

### Referencias

- [1] Debattisti, S, Mazzei, L & Panciroli, M 2013. Automated extrinsic laser and camera intercalibration using triangular targets. Intelligent Vehicles Symposium (IV), 2013 IEEE, 2013, pp. 696-701.
- [2] Cortes, C & Vapnik, V 1995, Support vector network, Machine Learning, vol. 20, pp. 1-25.
- [3] Fremont, V & Bonnifait, P 2008. Extrinsic calibration between a multi-layer lidar and a camera. 2008 IEEE Int. Conf. Multisens. Fusion Integr. Intell. Syst., 2008.
- [4] García, F, Jiménez, F, Naranjo, JE, Zato, JG, Aparicio, F, Armingol, JM & de la Escalera, A. 2012. Environment perception based on LIDAR sensors for real road applications.
- [5] García, F, García, J, Ponz, A, de la Escalera, A & Armingol, JM 2014. Context Aided Pedestrian Detection for Danger Estimation Based on Laser Scanner and Computer Vision. Expert Systems With Applications, Vol: 41 (15), pp.6646-6661.
- [6] Kaempchen, N, Buehler, M & Dietmayer, K 2005. Feature-level fusion for free-form object tracking using laser scanner and video. IEEE Proceedings Intelligent Vehicles Symposium 2005, pp. 453-458, 2005.
- [7] Kwak, K, Huber, DF, Badino, H & Kanade, T. 2011 .Extrinsic calibration of a single line scanning lidar and a camera. IEEE/RSJ Int. Conf. Intell. Robot. Syst., pp. 3283-3289, 2011.
- [8] Li, Y, Ruichek, Y & Cappelle, D 2011. 3D triangulation based extrinsic calibration between a stereo vision system and a LIDAR. 14th Int. IEEE Conf. Intell. Transp. Syst., pp. 797- 802, 2011.
- [9] Li, Y, Liu, Y, Dong, L, Cai, X 2007. An algorithm for extrinsic parameters calibration of a camera and a laser range finder using line features, IEEE/RSJ Int. Conf. Intell. Robot. Syst.
- [10] Lisca, G, Jeong, PJP & Nedeveschi, S 2010. Automatic one step extrinsic calibration of a multi layer laser scanner relative to a stereo camera. Intell. Comput. Commun. Process. (ICCP), 2010 IEEE Int. Conf., 2010.
- [11] Martín, D, García, F, Musleh, B, Olmeda, D, Marín, P, Ponz, A, Rodríguez, CH, Al-Kaff, A, de la Escalera, A & Armingol, JM 2014. IVVI 2.0: An intelligent vehicle based on computational perception. Expert Systems with Applications 41.
- [12] Premebida, C, Ludwig, O & Nunes, U 2009. LIDAR and Vision-Based Pedestrian Detection System. Journal of Field Robotics, vol. 26, no. Iv, pp. 696-711, 2009.
- [13] Premebida, C, Ludwig, O, Silva, M & Nunes, U 2010. A Cascade Classifier applied in Pedestrian Detection using Laser and Image-based Features. Transportation, pp. 1153-1159, 2010.
- [14] Rodríguez-Garavito, CH, Ponz, A, García, F, Martín, D, de la Escalera, A & Armingol, JM 2014. Automatic Laser And Camera Extrinsic Calibration for Data Fusion Using Road Plane.
- [15] SICK, 2009. LD-MRS manual. SICK AG Waldkirch, Reute, Germany.
- [16] WHO, 2009. Global status report on road safety. Time for action. WHO library cataloguing-in-publication data, World Health Organization 2009, ISBN 978-9-241563-84-0, Geneva, Switzerland.
- [17] Zezhi C, Pears, N, Freeman, M, & Austin, J 2009. Road vehicle classification using support vector machines. Intelligent Computing and Intelligent Systems, 2009. ICIS 2009. IEEE International Conference On , 4 214-218. doi:10.1109/ICICISYS.2009.5357707
- [18] Philip H. S. Torr and Andrew Zisserman. MLESAC: A New Robust Estimator with Application to Estimating Image Geometry. Computer Vision and Image Understanding, 78(1):138-156, 2000.
- [19] Cesar H. Rodriguez Garavito, Aurelio Ponz, Fernando Garcia, David Martin, Arturo de la Escalera, and Jose M. Armingol. Automatic Laser And Camera Extrinsic Calibration for Data Fusion Using Road Plane. In Proc. IEEE International Conference on Information Fusion (FUSION),2014.