

# Detección y seguimiento de vehículos en carretera utilizando técnicas de tratamiento digital de imágenes

L.M. BERGASA / M. MAZO / L. BOQUETE / R. BAREA / M.A. SOTELO / R. REDONDO

Departamento de Electrónica  
Universidad de Alcalá

En este artículo se presentan los primeros resultados de un trabajo desarrollado por los autores, cuyo objetivo final es realizar el seguimiento de múltiples vehículos (de diferentes categorías) en una secuencia de imágenes reales de una carretera con un único sentido de circulación, así como contarlos y clasificarlos según su categoría (turismos, camiones, etc.). Todo ello utilizando una cámara fija y técnicas de tratamiento digital de imágenes.

Obtener información sobre la trayectoria de los vehículos que circulan por una carretera, así como contarlos y clasificarlos, puede resultar de interés en el control de tráfico o simplemente para conocer la densidad total o de cada una de sus categorías (turismo, camión, etc.). La aplicación de técnicas de tratamiento digital de imágenes para abordar este problema plantea la resolución de diferentes aspectos tales como la detección de las zonas de la imagen en las que hay movimiento, obtención de los contornos de los objetos que se mueven, obtención de las trayectorias que sigue cada uno de los móviles, etc. Y todo ello teniendo presente que las condiciones de iluminación, en entornos reales, son cambiantes.

En esta primera fase del trabajo se han ensayado diferentes tipos de algoritmos y se han realizado numerosas pruebas, siempre con imágenes captadas en un tramo de carretera de un único sentido y con diferentes condiciones ambientales. Los algoritmos utilizados han sido en general simples ya que el objetivo fundamental era tener una idea clara del nivel de dificultad y conocer en detalle los problemas que se plantean y la viabilidad de utilizar, con éxito para este fin, la visión artificial. En nuestro caso el sistema desarrollado responde al diagrama de bloques mostrado en la figura 1, siendo los resultados obtenidos altamente satisfactorios.

La función asignada a cada uno de los bloques es:

*Detector de objetos en movimiento:* se encarga de localizar las zonas en las que hay movimiento, utilizando para ello técnicas de resta de imágenes.

*Detector de la Imagen de Referencia:* se encarga de captar imágenes en las que no exista movimiento y tomar ésta como imagen de referencia, actualizándola continuamente para que el sistema sea robusto ante cambios de iluminación.

*Predictor de posición:* su función consiste en predecir la posición de cada móvil entre dos imágenes consecutivas.

*Cálculo de trayectorias:* este bloque se encarga de obtener la posición, dentro del plano imagen, de los objetos en movimiento y a partir de ellos

establecer la trayectoria que sigue cada uno.

*Clasificador de objetos:* su función consiste en detectar si un objeto ha seguido una trayectoria durante N imágenes consecutivas, y en caso afirmativo se le cataloga como móvil y se clasifica en moto, coche o camión, según su tamaño.

Dado que el sistema debe funcionar en tiempo real (teniendo en cuenta la velocidad máxima de un automóvil), un aspecto a tener presente será el tiempo total de ejecución de los diferentes algoritmos que intervienen en el

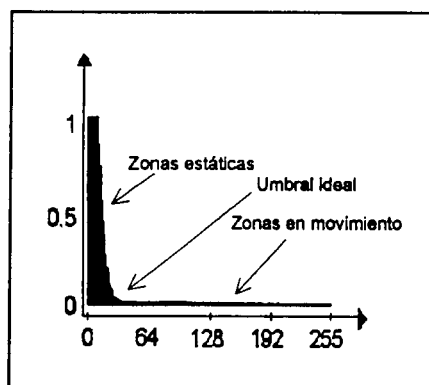


Figura 2. Histograma de la imagen diferencia,  $M(x,y,t)$ .

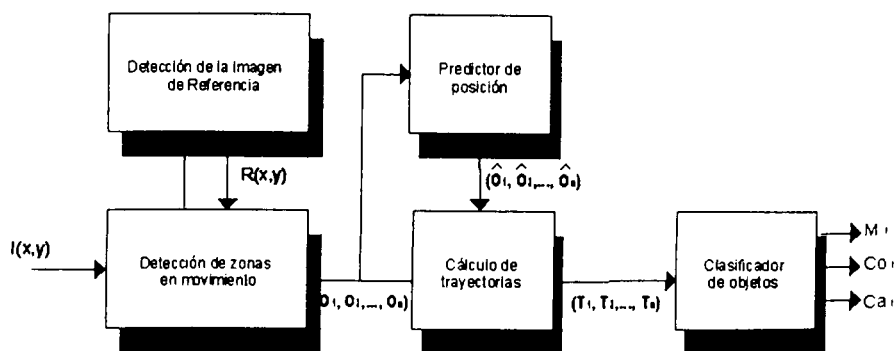


Figura 1. Diagrama de bloques del sistema desarrollado.

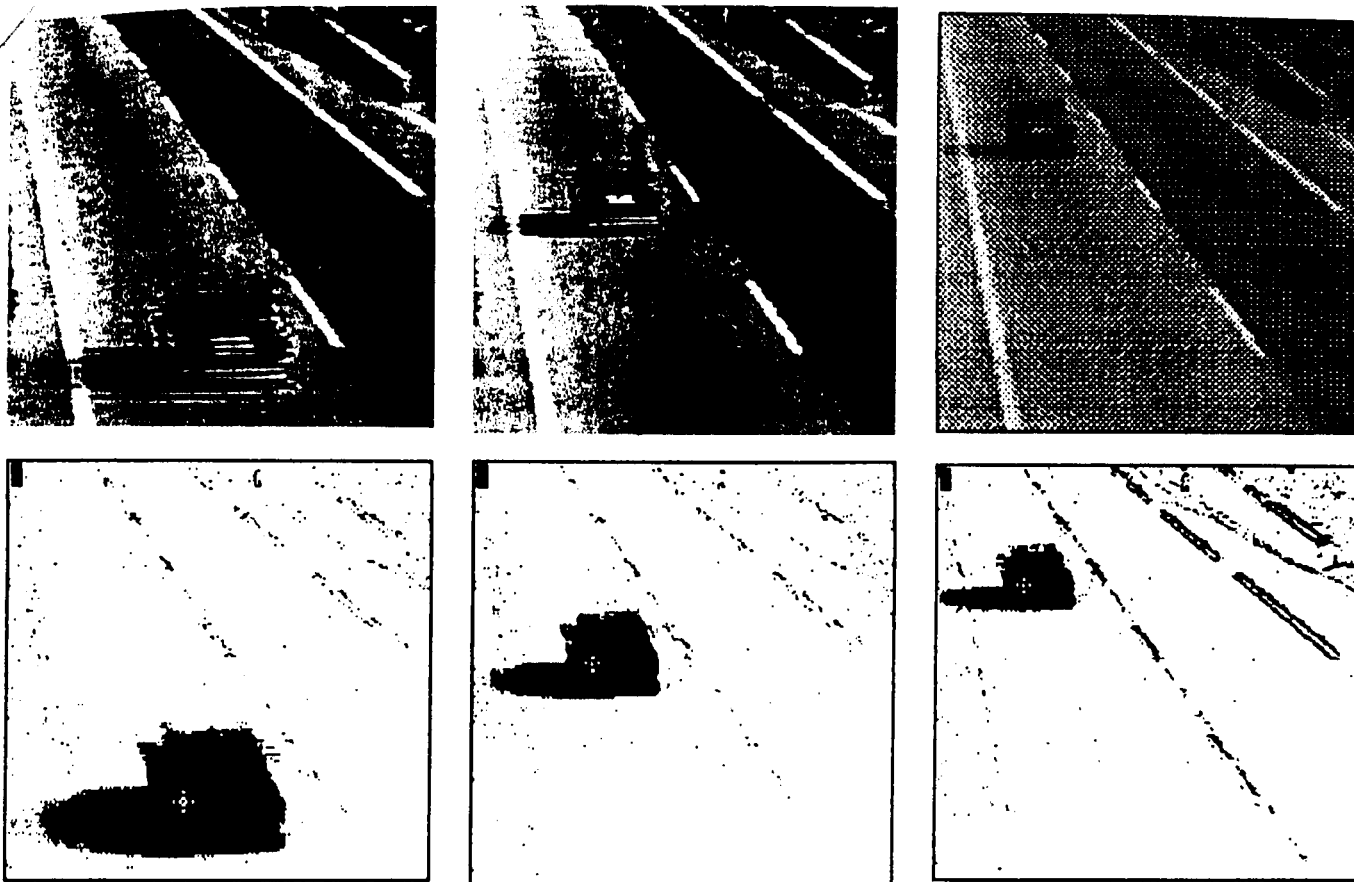


Figura 3. Ejemplo de detección de un turismo con color similar al de fondo.

proceso, y para ello se deberá tener presente el hardware que se utilice. En nuestro caso el equipo utilizado en las pruebas realizadas está formado por: cámara CCD de  $512 \times 512$  pixels, ordenador Pentium 90 MHz con 16 Mb de RAM, tarjeta digitalizadora Matrox IM-640 (2 Mb de memoria). El lenguaje de programación empleado ha sido C para Windows.

A continuación se realiza una descripción de los diferentes bloques que constituyen este primer sistema desarrollado (figura 1).

### Obtención de la Imagen de Referencia

Por imagen de referencia,  $R(x,y)$ , o fondo entendemos aquella que no tiene zonas en movimiento, y constituye la base para determinar si existen o no objetos en movimiento dentro de la escena bajo estudio.

En nuestro caso se ha diseñado una estrategia que permite actualizar cada cierto tiempo esta imagen de referencia. La idea básica consiste, fundamentalmente, en detectar si en la imagen actual no hay zonas en movimiento. Así, si a la imagen actual se le denomina  $I(x,y,t)$ , a la precedente  $I(x,y,t-1)$

y a la diferencia, en un instante  $t$  entre ambas  $D(x,y,t)$ , entonces:

$$\text{Si } D(x,y,t-i) = I(x,y,t-i) - I(x,y,t-1-i) = 0, i = 0,1,\dots,P = I(x,y,t) = R(x,y,t)$$

Esto es, para que una imagen sea candidata a ser  $R(x,y,t)$  debe cumplirse que  $D(x,y,t)$  sea cero y la imagen candidata se convertirá en imagen de referencia cuando esta circunstancia se produzca  $P+1$  veces consecutivas.

Todo esto hace que el algoritmo de detección de los móviles sea más robusto frente a los cambios de iluminación y las sombras creadas por los objetos estáticos. En nuestro caso se ha fijado  $P = 2$ , ya que se ha considerado una velocidad media de los vehículos de 100 km/h y dado que el tiempo de procesado entre imágenes consecutivas es de 200 ms, se evita que cuando comience a aparecer algún móvil por un extremo de la imagen éste, por su pequeño tamaño, no sea considerado como objeto en movimiento (más adelante se verá que todo objeto ha de tener un tamaño mínimo para evitar que sea confundido con ruido).

### Detección de zonas en movimiento. Método de la diferencia doble

Si bien existen diferentes alternativas para detectar objetos en movimiento cuando el fondo es estático (diferencia de niveles de gris, coincidencia de bordes, etc.) todas ellas se basan, fundamentalmente, en la diferencia de imágenes. Con ello se consigue resaltar las zonas dentro de la imagen donde se han producido movimientos. Este algoritmo funciona correctamente en la mayor parte de los casos, pero cuando el color de los móviles es parecido al del fondo, no consigue detectar movimiento. Para solucionar este problema se ha implementado un método que hemos denominado como «diferencias doble», que consigue magnificar la intensidad de todas aquellas zonas que se mueven, lo que facilita la detección de las zonas en movimiento. El método de diferencias dobles viene dada por:

$$M(x,y,t) = | I(x,y,t-1) - R(x,y,t) | + | I(x,y,t) - R(x,y,t) |$$

Por tanto,  $M(x,y,t)$  será una imagen en la que resaltarán las zonas en movimiento (imagen diferencia). Las

nas en las que no exista tal aparecencia en color negro ( $M(x,y,t) = 0$ ) y en caso contrario su nivel de gris será función de la diferencia entre el objeto y el fondo. Evidentemente, si el objeto que se mueve es del mismo color que el fondo (caso poco probable si se trata de vehículos), éste no podrá ser detectado.

### Umbralización automática y segmentación

Teóricamente la imagen  $M(x,y,t)$  sólo debería tomar valores distintos de cero en aquellas zonas donde hay movimiento. En la práctica, al trabajar en entornos reales, la cámara se ve sometida a vibraciones debidas al viento, que originan pequeñas diferencias entre las imágenes captadas y la referencia. Este hecho, unido a otros tipos de ruido que aparecen en estos sistemas, hacen que la suposición teórica no se cumpla. En efecto, si se analiza el histograma de la imagen diferencia,  $M(x,y,t)$  (en la figura 2 se muestra un ejemplo real), se puede comprobar que existen pixels con niveles de gris próximos a cero en zonas donde no hay móviles. Para reducir este efecto del ruido se binariza la imagen  $M(x,y,t)$ , utilizando para ello un umbral  $T$  (calculado automáticamente) obtenido a partir de su histograma. Ésto permite obtener una imagen segmentada en dos zonas: estáticas y en movimiento.

En nuestro caso, el umbral  $T$  óptimo se obtiene calculando el punto del histograma, previa suavización de éste, donde el cambio de pendiente es máxima (punto donde la pendiente pasa de tener unos valores altos a otros próximos a cero).

Una vez binarizada la imagen, se obtiene el centroide ( $x_g, y_g$ ) de los objetos en movimiento, utilizando para ello los momentos de orden 1 ( $m_{10}$  y  $m_{01}$ ) y de orden 0 ( $m_{00}$ ) que, junto al tamaño de los objetos, son las únicas características de interés en esta aplicación. El centroide es la característica que se utiliza para realizar el seguimiento y el tamaño sirve para realizar la clasificación de los vehículos.

$$\bar{x}_g = \frac{m_{10}}{m_{00}} \quad \bar{y}_g = \frac{m_{01}}{m_{00}}$$

### Seguimiento de los móviles

En este apartado se describe la estrategia utilizada para, una vez segmentada la imagen, establecer la correspondencia en imágenes consecutivas entre los diferentes objetos móviles

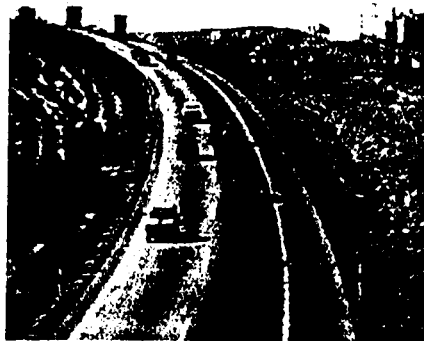


Figura 4. Ejemplo de imagen con múltiples vehículos

(múltiples objetos en movimiento), y determinar la trayectoria seguida por cada uno de ellos.

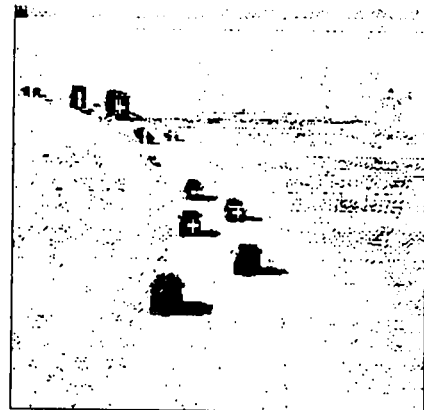
La correspondencia entre objetos en imágenes consecutivas se basa en la predicción de la posición que ocupará cada uno de ellos, dado que la información que se obtiene de los objetos en movimiento es solamente su centroide. Con ello se consigue, además, resolver el problema derivado del tiempo de captación y procesado de imagen, que hace que la posición obtenida al segmentar la misma sea la que tenía el móvil un tiempo antes (tiempo de captación y procesado).

Si no se compensaran estos desfases, este fenómeno se traduciría en la aparición de errores de seguimiento del móvil. Por todo ello, es necesario un sistema de predicción capaz de identificar la trayectoria del móvil para estimar la posición y, de este modo, establecer la correspondencia entre objetos y evitar que se produzcan errores de seguimiento. En nuestro caso, una vez detectados y calculados los centroides en la imagen  $I(x,y,t)$ , se inicializa el módulo de predicción y se obtienen los centroides en la imagen  $I(x,y,t+1)$ .

A partir de esta información se realiza la búsqueda, en el entorno de estos valores, para determinar las nuevas posiciones.

### Predicción de trayectoria

Para la predicción de trayectoria se ha supuesto que las trayectorias seguidas por los móviles serán aproximadamente lineales (no es probable que un automóvil cambie significativamente su trayectoria en tiempos del orden de los ms.). Pero dado que la relación entre las coordenadas de un punto de la escena (3D) y las del plano imagen (2D) es no lineal, el algoritmo de predic-



ción deberá ser capaz de seguir trayectorias no lineales.

En nuestro caso, el algoritmo de predicción de posición (expresada en coordenadas cartesianas) utilizado responde a las ecuaciones:

$$\hat{x}[t+1] = X[t] + \omega_x \cdot \Delta_x(t) + k_x \cdot e_x(t)$$

$$\hat{y}[t+1] = Y[t] + \omega_y \cdot \Delta_y[t] + k_y \cdot e_y(t)$$

donde  $\hat{x}(t+1)$  e  $\hat{y}(t+1)$  representan las coordenadas estimadas,  $X(t)$  e  $Y(t)$  vienen dadas por:

$$X(t) = \frac{1}{m+1} \sum_{i=0}^m x(t-i);$$

$$Y(t) = \frac{1}{m+1} \sum_{i=0}^m y(t-i)$$

siendo  $m$  la profundidad de memoria del modelo y  $x(t-i)$  e  $y(t-i)$  son las coordenadas del centroide de cada uno de los móviles, obtenidas del procesamiento de la imagen.

$\Delta_x[t]$  y  $\Delta_y[t]$  son matrices columna de  $m \times 1$ , dadas por:

$$\Delta_x[t] = \begin{bmatrix} x(t) - x(t-1) \\ x(t-1) - x(t-2) \\ x(t-2) - x(t-3) \\ \vdots \\ x(t-m) - x(t-m-1) \end{bmatrix};$$

$$\Delta_y[t] = \begin{bmatrix} y(t) - y(t-1) \\ y(t-1) - y(t-2) \\ y(t-2) - y(t-3) \\ \vdots \\ y(t-m) - y(t-m-1) \end{bmatrix}$$

$\omega_x$  y  $\omega_y$  vienen dadas por:

$$\omega_x = [\omega_{x0} \omega_{x1} \omega_{x2} \dots \omega_{xm}]$$

$$\omega_y = [\omega_{y0} \omega_{y1} \omega_{y2} \dots \omega_{ym}]$$

y son las matrices de parámetros a identificar. En nuestro caso se ha utilizado descenso del gradiente como método de identificación y actualización de los coeficientes de  $\omega_x$  y  $\omega_y$ .

En cuanto a  $e_x(t)$  y  $e_y(t)$  vendrán dadas por:

$$e_x(t) = \hat{x}(t) - x(t)$$

$$e_y(t) = \hat{y}(t) - y(t)$$

En cuanto a  $k_x$  y  $k_y$  han sido ajustadas de forma experimental.

### Algoritmo de seguimiento

Este algoritmo permite, a partir de la información proporcionada por el predictor de trayectorias, obtener la trayectoria de cada uno de los objetos en movimiento.

Dado que en la imagen segmentada pueden aparecer objetos de pequeño tamaño (cuyo origen se debe a los efectos del ruido) y objetos extraños que se presenten esporádicamente en la escena, se rechazan todos los objetos cuyo tamaño no supere un determinado número de pixels, que es fijado por el usuario en el momento de inicializar el sistema (y que, como es lógico, será función de las características de la óptica, ubicación de cámara, etc.) y aquellos cuya presencia en escena no supere un determinado número de frames consecutivas.

El algoritmo comienza detectando el número de objetos que hay en movimiento. Cada uno de estos objetos se identifica por:  $O_j[x_j, y_j]$ ,  $0 \leq j \leq n$ , siendo  $n$  el número de objetos y  $(x_j, y_j)$  las coordenadas del centroide de cada uno de ellos (ver figura 1), obtenidas a partir del procesamiento de la imagen. Del resultado de la predicción se obtiene:  $\hat{O}_h[\hat{x}_h, \hat{y}_h]$ , siendo  $\hat{x}_h, \hat{y}_h$  las coordenadas estimadas. La correspondencia entre los objetos estimados  $\hat{O}_h[\hat{x}_h, \hat{y}_h]$  y los obtenidos en la última imagen,  $O_j[x_j, y_j]$ , se hace a partir de la distancia euclídea de los centroides. Es decir, para cada uno de los objetos estimados  $\hat{O}_h$  se calculan las distancias a los centroides de los objetos de la imagen,  $O_j$

$$\text{desde } h = 1, \dots, n: \quad d_{hj} =$$

$$= \sqrt{(x_j - \hat{x}_h)^2 + (y_j - \hat{y}_h)^2}$$

$$1 \leq j \leq n$$

y:  $\hat{O}_h[\hat{x}_h, \hat{y}_h]$  se corresponde con  $O_j[x_j, y_j]$  si  $d_{hj}$  es la mínima.

Dado que pueden aparecer y desaparecer objetos móviles, hay que prever estas situaciones. En el caso de que aparezcan nuevos objetos, basta con crear nuevas trayectorias. Si desaparecen objetos, se puede deber a dos causas: objetos que han salido del campo de visión y/o se han ocultado detrás de otro. Para solucionar estos casos se predicen las trayectorias que siguen los objetos desaparecidos en función del tipo de movimiento que venían realizando. Si después de un número determinado de imágenes no aparece el objeto, se elimina y si vuelve a la escena porque estaba oculto, entonces se le asigna de nuevo la trayectoria obtenida del predictor.

También se ha fijado una distancia máxima entre la posición y la predicción, para que no se confundan objetos nuevos con los que desaparecen de la escena.

### Conteo de móviles

Como ya se ha comentado, uno de los objetivos era contar el número de vehículos por categorías (moto, turismo o camión) que circulan por el tramo de carretera captado por la cámara. Para ello, lo primero que se hace es determinar el sentido de circulación de cada uno de los vehículos (con ello se evita que se cuente dos veces un mismo vehículo). Este proceso se realiza automáticamente en la configuración inicial del programa. El método utilizado para averiguar el sentido de movimiento, consiste en "seguir" a varios vehículos y a partir de ellos determinar la dirección y el sentido hacia donde se mueven.

Por otra parte, para que un objeto se considere como vehículo, se debe detectar  $r$  veces consecutivas siguiendo una determinada trayectoria y debe cruzar la línea horizontal central de la imagen, dado que el eje de la cámara está alineado con el de la carretera. Una vez detectado un vehículo, se podrá clasificar en función de su tamaño en moto, turismo o camión.

Como es lógico, la ubicación (similar a las de la DGT) y óptica de la cámara, hace que el tamaño de los vehículos en el plano imagen cambie con el tiempo, aumentando o disminuyen-

do (dependiendo del sentido de circulación). Una forma fácil de hacer la clasificación de los vehículos, sin tener en cuenta parámetros de perspectiva, es medir todos los objetos en un mismo punto de la imagen. En nuestro caso se ha elegido de forma experimental, tras diversas pruebas, la zona central de la imagen. Para ello, una vez ubicada la cámara, se debe hacer una calibración inicial (off-line) que permita posteriormente (on-line) clasificar, por tamaño, los vehículos. Se ha incluido, en la configuración inicial, una opción de calibración que permite obtener los patrones (utilizando como única característica el área) del universo de trabajo (moto, turismo, camión). Una vez definidos estos patrones la clasificación se ha realizado, en esta primera fase del trabajo, utilizando un clasificador Bayesiano.

### Resultados

En este apartado se van a presentar los resultados obtenidos en las diferentes pruebas realizadas con imágenes reales. Para ello se llevó a cabo la grabación de imágenes (en formato VHS) a distintas horas del día (y por tanto en diferentes condiciones de iluminación natural y de densidad de tráfico) en una autopista.

En la figura 3 se muestra una secuencia de tres imágenes, con un solo turismo que se aleja de la cámara, y las correspondientes imágenes de diferencias, donde el signo + indica el centroide del móvil y en la esquina superior izquierda se presenta el número de vehículos detectados en cada imagen. Se puede observar que las coordenadas del centroide corresponden a las del objeto formado por el vehículo más su propia sombra (esto no plantea problemas importantes en el método y además su efecto puede ser evitado utilizando información en color).

En la figura 4 se muestra otra imagen analizada, en la que se puede apreciar la detección de múltiples vehículos de distintos tipos.

El problema más importante de este método es el solapamiento de vehículos. Cuando dos o más móviles van muy cerca entre sí (figura 5), se detectarán como uno solo, si bien en el momento en que se separen se les considerará como varios. Normalmente estos solapamientos suelen ocurrir en los adelantamientos, o cuando hay una elevada densidad de circulación. Las posibles soluciones a este problema son varias: utilización de diversas cámaras

...ndolas de forma adecuada, em-  
...o de plantillas deformables en la  
...ntificación, utilización de imágenes  
...n color de manera que se introduzca  
...ste como parámetro para realizar seg-  
...mentación, etc.

Se puede concluir, por tanto, indi-  
cando que el empleo de la visión arti-  
ficial en la detección y seguimiento de  
vehículos es factible.

Con este primer trabajo se han obte-  
nido unos resultados con unos erro-  
res del orden de magnitud de los exis-  
tentes en los sistemas de bandas de  
presión utilizados actualmente en las  
carreteras, en entornos con densidades  
de tráfico no muy elevadas (el error  
del sistema aumenta cuando los  
vehículos van en caravana, ya que  
los problemas de solapamiento son  
más acusados) y con imágenes de  
día.

Los problemas aquí presentados se-  
rán abordados en una segunda fase del  
proyecto; además, cabe destacar que  
esta técnica aporta ventajas adiciona-  
les sobre los sistemas tradicionales

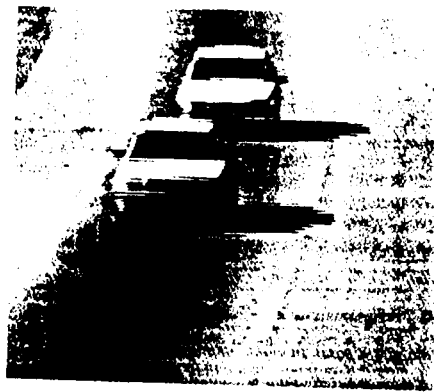


Figura 5. Ejemplo de imagen con solapamiento de dos vehículos.

como son: la detección de vehículos  
circulando en sentido contrario, detec-  
ción de accidentes, control de densi-  
dad de tráfico, distinción de los dife-  
rentes tipos de vehículos, monitoriza-  
ción de zonas de interés, etc. □

#### Bibliografía

1) R. Redondo García. Detección de obje-  
tos móviles mediante Visión Artificial.

2) L. Baumela Molina. Contribución al  
seguimiento de móviles con cabezas robó-  
ticas y visión artificial, 1995.

3) M.V. Correia. Motion Estimation Fra-  
mework. (VI Spanish Symposium on pat-  
tern recognition and image analysis). Pag.  
605-614. 1995.

4) E. Diaz, J. Domingo. Detection and  
model fitting of moving vehicles in planar  
curves using morphological techniques. (VI  
Spanish symposium on pattern recognition  
and image analysis). Pag. 685-691. 1995.

## REDE/LIBROS

### La tecnología del haz de electrones y sus aplicaciones

La obra tiene como principal objetivo dar a conocer los principios de funcionamiento y sus aplicaciones, como son soldadura, tratamientos térmicos, recubrimientos, perforación y fusión. A lo largo de sus páginas se presentan las principales características de esta tecnología, así como los equipos industriales existentes.

160 págs. Autores: F. Santamaría, I. San José  
Libro nº 654. Precio: 1.990 pts. (inc. IVA)

### Acumuladores electroquímicos

Hace una descripción de los acumuladores primarios y secundarios y analiza los sistemas avanzados para las nuevas aplicaciones, vehículos eléctricos, sistemas de alimentación ininterrumpida, acumuladores solares, etc. Su objetivo es que el profesional disponga de los conocimientos y datos necesarios para elegir el sistema más adecuado a cada tipo de aplicación.

304 págs. Autor: José Fullea  
Libro nº 655. Precio: 1.990 pts. (inc. IVA)

### Hornos Industriales de resistencias

Estudia de forma sistemática y extensa las características y funcionamiento de las resistencias de calentamiento en hornos industriales, los materiales de aislamiento, intercambio térmico entre las resistencias y la carga con los métodos de cálculo correspondientes. Se describen los tipos más frecuentes, los equipos de regulación y control para concluir con el estudio de los balances energéticos y de explotación.

334 págs. Autor: Julio Astigarraga  
Libro nº 656. Precio: 2.730 pts. (inc. IVA)

### Hornos Industriales de inducción

Estudia, en primer lugar, los principios físicos, eléctricos y térmicos de la aplicación de la inducción al calenta-

miento de materiales conductores, señalando las propiedades y características correspondientes, principalmente de los metales. Describe los equipos utilizados en frecuencia de red y en media frecuencia, y las aplicaciones principales clasificadas en hornos de fusión y mantenimiento de metales, calentadores de inducción para deformación plástica y equipos de inducción para tratamientos térmicos.

232 págs. Autor: Julio Astigarraga  
Libro nº 657. Precio: 2.335 pts. (inc. IVA)

### Arranque industrial de motores asíncronos

Describe la problemática general del arranque de accionamiento conjunto; analiza los modos de arranque clásicos y los arrancadores electrónicos estáticos. Dedicamos un capítulo a la influencia de la alimentación. También describe los componentes comunes a todos los modos de arranque. Hay un capítulo para el estudio de los problemas térmicos y mecánicos, y varios dedicados a especificaciones, aplicaciones y selección de modos de arranque.

324 págs. Autor: J. M. Merino-Azcárraga  
Libro nº 659. Precio: 2.750 pts. (inc. IVA)

### Hornos de alta frecuencia y microondas

Presenta los principios físicos del calentamiento por alta frecuencia en cargas homogéneas y heterogéneas, así como en microondas, con el concepto de profundidad de penetración y definiendo en ambos casos las características de funcionamiento. Describe a continuación los equipos electrónicos y mecánicos, tanto en alta frecuencia como en microondas. Se dedica especial atención a los aplicadores de carga, con los diferentes tipos que normalmente se utilizan.

140 págs. Autor: Julio Astigarraga  
Libro nº 660. Precio: 2.980 pts. (inc. IVA)

Para sus pedidos utilice el Boletín de la página 58