
IDENTIFICACION AUTOMATICA DE MATRICULAS DE AUTOMOVILES MEDIANTE VISION ARTIFICIAL Y REDES NEURONALES

L. BOQUETE, M. MAZO, J.A. JIMENEZ, R. BAREA, L.M. BERGASA y M.A. SOTELO
Univ. de Alcalá, Esc. Politécnica, Depto. de Electrónica, 28801 Alcalá de Henares, Madrid - España
(e-mail: boquete@depeca.alcala.es)

RESUMEN

En este artículo se describe un sistema de identificación automática de matrículas de vehículos, utilizando procesamiento digital de imágenes en color y redes neuronales. Con dicho sistema se consigue una alta fiabilidad en la identificación, con pequeños tiempos de respuesta y bajo costo de instalación. Los algoritmos de tratamiento digital de la imagen funcionan correctamente sin necesidad de utilizar un sistema especial de iluminación. El reconocimiento de los dígitos se realiza mediante una red neuronal que siendo entrenada con ejemplos contaminados por ruido, consigue una correcta identificación aún en casos difíciles.

AUTOMATIC IDENTIFICATION OF VEHICLE NUMBER PLATES USING ARTIFICIAL VISION AND NEURAL NETWORKS

ABSTRACT

A system for automatic vehicle number plate identification using color digital imaging and neural network is described in this paper. This system offers a high identification reliability, low response times and low set-up cost. The algorithms for digital imaging work correctly without using special lighting systems. Figure recognition is made using a neural network which, having been trained with noise-contaminated examples, accurately identifies figures even under adverse conditions.

Keywords: artificial vision, digital imaging, pattern recognition, neural networks.

INTRODUCCIÓN

El reconocimiento e identificación de las placas de vehículos puede ser utilizado en una gran variedad de aplicaciones: tarificación en garajes o autopistas, detección de infractores del código de tráfico, seguridad, etc. En el presente trabajo se detallan las características de un sistema de identificación de placas, el cual se caracteriza porque puede funcionar en condiciones de luz cambiantes, no necesitando ningún sistema especial de iluminación y por otro lado, no es necesario un exacto posicionamiento del vehículo a identificar. El proceso de lectura de la matrícula se divide en dos partes: captura y análisis de la imagen (con el fin de segmentar los diferentes dígitos) e identificación de dichos dígitos para reconocer la correspondiente matrícula.

Es evidente que los sistemas de reconocimiento de vehículos van a ir introduciéndose en los próximos años, con el objetivo de llevar un control más detallado del tráfico que se mueve por el entramado de las redes viarias. A modo de ejemplo se puede citar el que recientemente se ha introducido en algunas ciudades europeas un sistema similar al utilizado en los accesos a parkings, mediante una tarjeta que identifica al usuario del vehículo, permitiendo el acceso a zonas con tráfico restringido, normalmente las zonas céntricas, muy congestionadas por el tráfico o la contaminación (Savatier, 1996). Sin duda estos sistemas son poco flexibles, siendo necesario buscar otras alternativas que den solución a estos problemas, garantizando un alto grado de fiabilidad. Una solución a este problema consistiría en reconocer su placa de matrícula, ya que ésta constituye una importante fuente de información sobre el vehículo. En este sentido se ha desarrollado este proyecto de investigación y para ello se plantearon dos objetivos claros: en primer lugar, aislar la placa del resto del vehículo y en segundo lugar reconocer los caracteres que hay impresos en ella. Todo ello se ha abordado dentro del campo de la Visión Artificial, contando con un sistema que permita detectar los vehículos que se quieren reconocer.

En esta línea existen en la literatura técnica diversas referencias a trabajos donde se ha abordado este problema. Así en (Arrègui y Mitxelena, 1990) se describe un sistema que utiliza iluminación tipo flash y para detectar la posición de la matrícula se apoya, fundamentalmente, en operadores morfológicos, basándose el reconocimiento de los dígitos en un OCR comercial. En (Fuertes, 1995) se describe un método que calcula los ejes de simetría de la imagen, obteniendo características intrínsecas e invariables para cada vehículo. Cowell (1995) estudia un método sintáctico para el reconocimiento de los dígitos, suponiendo que éstos han sido previamente segmentados; esta técnica se basa en la utilización de un diccionario de las primitivas que forman un dígito y mediante el análisis de

las mismas, se consigue su identificación.

El objetivo que se planteaba era diseñar un sistema que fuera capaz de identificar matrículas de vehículos, asegurando un alto grado de fiabilidad y tiempos de respuesta pequeños. Para ello se pensó en utilizar información en color para segmentar los dígitos de la matrícula y utilizar para el reconocimiento de los mismos una red neuronal. Como condicionantes más importantes del sistema se imponían: tiempo de respuesta inferior a 2 s., equipo con un coste reducido y una fiabilidad en el reconocimiento superior al 95%.

Para garantizar un alto grado de fiabilidad se han realizado todas las pruebas con imágenes captadas en condiciones naturales de iluminación. Así se han tomado imágenes a diferentes horas del día con el fin de tener diferentes condiciones de iluminación ambiental. Hay que tener presente que las condiciones de iluminación del entorno y el ruido de imagen que siempre se superpone en una placa de matrícula son los problemas más importantes que se encuentran en cualquier sistema encaminado a identificar matrículas de vehículos. Todo ello es lo que nos ha llevado a trabajar con imágenes en color, ya que éstas nos permiten disponer de una mayor información y por tanto de mayores posibilidades de garantía para una buena fiabilidad.

ANÁLISIS DE LA IMAGEN

La captura de las imágenes se realiza sin que el sistema de iluminación deba cumplir unas condiciones especialmente exigentes, estando el vehículo parado y visualizando su parte delantera, a una distancia que puede estar comprendida en el margen de 1 a 1,5 metros y el ángulo de orientación del vehículo con respecto a la cámara puede ser de $\pm 15^\circ$, lo que supone cierta tolerancia a los errores de posicionamiento.

En la Fig.1 se muestra una imagen original, en formato RGB. En este formato se tiene la posibilidad de realizar conversiones de color, con el fin de trabajar en el formato que sea más apropiado



Fig. 1: Imagen original



Fig. 2: Resultado de la conversión de color

para el objetivo de interés. Los espacios de color más utilizados son RGB y el formato HSI (color-saturación-intensidad). Las relaciones entre ambos espacios son las indicadas en las ecuaciones (1):

$$\begin{bmatrix} I \\ V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & 0 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}; \quad (1)$$

$$H = \text{tg}^{-1}\left(\frac{V_2}{V_1}\right); \quad S = \sqrt{V_1^2 + V_2^2};$$

Siendo la relación inversa la indicada en (2):

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} \frac{4}{3} & \frac{-2\sqrt{6}}{9} & \frac{\sqrt{6}}{3} \\ \frac{2}{3} & \frac{\sqrt{6}}{9} & \frac{-\sqrt{6}}{9} \\ 1 & \frac{\sqrt{6}}{3} & 0 \end{bmatrix} \cdot \begin{bmatrix} I \\ V_1 \\ V_2 \end{bmatrix}; \quad (2)$$

$$V_1 = S \cdot \cos(H); \quad V_2 = S \cdot \sin(H);$$

Experimentando diversas transformaciones, se ha llegado a la conclusión de que los mejores resultados, aquéllos que resaltan mejor la placa de la matrícula sobre el resto de la imagen, se consiguen después de aplicar, a la imagen original, dos veces consecutivas las transformaciones dadas por las ecuaciones (1) y trabajando con la componente H. Esta mejora se puede apreciar claramente en el ejemplo mostrado en la Fig. 2., en donde se observa como se reduce el número de tonalidades y a su vez se puede comprobar que existe una clara identificación de los dígitos que constituyen la matrícula. Otra ventaja de este método, es que una vez obtenida la componente H, el



Fig. 3: Resultado de la ecualización

resto del procesamiento se puede hacer sobre un buffer de 8 bits de profundidad, no siendo necesario trabajar con imágenes de 24 bits, lo cual tiene una importante influencia en la velocidad de procesamiento de los algoritmos y en la capacidad de memoria necesaria.

El siguiente paso en el preprocesamiento de la imagen consiste en realizar una ecualización del histograma mediante una función de distribución de Rayleigh. Las expresiones de la función densidad de probabilidad y de la correspondiente función de transferencia son las dadas por las ecuaciones (3) y (4), respectivamente (Pratt, 1991).

$$f_r(y) = \frac{1}{y^2} \cdot e^{-\frac{y^2}{2}}; \quad (3)$$

$$y(x) = \sqrt{2} \cdot y^2 \cdot \ln\left(\frac{1}{1 - \sum_{k=0}^x h(k)}\right); \quad (4)$$

Siendo $h(k)$ el histograma normalizado de la imagen de entrada (componente H). Eligiendo un coeficiente (γ) adecuado para la ecualización, se obtienen resultados como los mostrados en la Fig. 3, donde se puede observar una considerable reducción del ruido sobre la imagen.

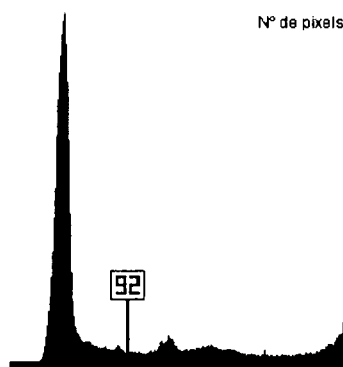


Fig. 4: Histograma de la imagen ecualizada



Fig. 5: Imagen umbralizada

UMBRALIZACIÓN Y SEGMENTACIÓN

La umbralización se ha realizado eligiendo el correspondiente umbral de forma dinámica, en función de las características del histograma de la imagen (Fig. 4). Para elegir el umbral óptimo se supone que el histograma es bimodal; esto es, existe un claro agrupamiento de los niveles de grises de los pixels correspondientes a los dígitos de la matrícula, mientras que el resto tienen unos niveles de grises muy diferentes. Aprovechando esta circunstancia, el umbral se puede tomar como el punto medio entre los dos máximos principales del histograma de la imagen previamente ecualizada. En la Fig. 4 se indica con una marca el nivel seleccionado automáticamente. Debemos indicar que para realizar la búsqueda de los dos máximos del histograma, previamente se debe realizar una suavización del mismo, con el fin de no obtener falsos resultados.

Los resultados de la umbralización de las imágenes se muestran en la Fig. 5 (en la umbralización se ha asignado a los pixels con nivel de gris inferior al umbral el nivel 255 - blanco- y al resto 0 -negro-).

El siguiente paso en el análisis de la imagen consiste en eliminar aquellos objetos (zonas blancas) cuyas dimensiones máximas, tanto en la dirección horizontal como vertical, no están dentro de los márgenes que cabe esperar de las dimensiones de los dígitos, considerando que las imágenes se pueden captar a distancias de 1 a 1,5 metros. De este modo, solamente se consideran para su posterior análisis aquellos objetos que tienen dimensiones similares a las de los dígitos (Fig. 6).

Dado que en la fase anterior pueden aparecer objetos que no son dígitos (marca del automóvil, pegatinas, manchas, etc.), se realiza un pro-

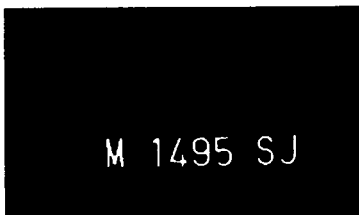


Fig. 6: Placa aislada



Fig. 7: Dígitos con ruido.



Fig. 8: Dígitos limpios.

ceso mediante el cual solamente se tienen en consideración aquellos objetos alineados en sentido horizontal, y que se agrupan en la parte central de la imagen (que es donde existe mayor probabilidad de tener la matrícula, por disposición geométrica de la cámara y vehículo), descartando por tanto elementos aislados dentro de la imagen.

Realizadas estas fases, se procede a aislar cada uno de los dígitos que forman la matrícula. Para ello, sobre cada uno de los objetos que quedan en la imagen, se calcula el rectángulo mínimo que lo encierra y el centro de gravedad. Con estos datos, es fácil aislar entre sí los diferentes dígitos, considerando que deben de estar alineados horizontalmente y con unas dimensiones predeterminadas dentro de ciertos márgenes (Fig. 7).

Por último, y con el fin de mejorar la calidad de cada uno de los dígitos (por ejemplo, eliminar muescas sobre el contorno o huecos oscuros en su interior), se realiza un suavizado binario, que evalúa una función booleana sobre un entorno de vecindad centrado sobre un pixel "p", cuyos vecinos más próximos nombrados de izquierda a derecha, y de arriba hacia abajo son respectivamente: a, b, c, d, p, e, f, g, h. El nivel del pixel "p" se modifica aplicando consecutivamente las expresiones de las ecuaciones (5) y (6), (Fu, 1988).

$$B_1 = p + b \cdot g \cdot (d + e) + d \cdot e \cdot (b + g); \quad (5)$$

$$B_2 = p \cdot [(a + b + d) \cdot (e + g + h) + (b + c + e) \cdot (d + f + g)]; \quad (6)$$

En la Fig. 8 se muestra un ejemplo del resultado de este proceso, donde se observa que se han obtenido unos dígitos con una calidad aceptable.

RECONOCIMIENTO DE DÍGITOS

Una vez realizada una correcta segmentación de los dígitos que forman la matrícula, se procede a la identificación de los mismos. Ésta se puede realizar mediante diversas técnicas (momentos invariantes, descriptores de Fourier, OCR, métodos gramáticos, redes neuronales). En nuestro caso se ha optado por la utilización de redes neuronales. Las razones fundamentales que nos

han llevado a ello son:

- Es posible enseñar a la red una parte de su universo de trabajo en las mismas condiciones en las que debe funcionar el sistema.
- Se pueden reconocer dígitos contaminados por ruido.
- Una vez entrenada la red, el tiempo de reconocimiento es despreciable, en comparación con el de técnicas basadas en el cálculo de momentos, descriptores de Fourier, etc.

La red neuronal utilizada ha sido programada en C++, con el compilador Borlandc- v. 3.1, aplicando tecnologías de programación orientada a objetos.

Entrenamiento de la red neuronal

La red utilizada es del tipo "feed-forward", formada por una capa de entrada o de activación, una capa oculta y una capa de salida, existiendo solamente conexiones desde la salida de una neurona hasta las entradas de las neuronas de la siguiente capa.

Una dificultad en la utilización de las redes neuronales es la ausencia de un método a priori que determine la arquitectura adecuada para resolver un determinado problema. Normalmente se procede a realizar pruebas con diferentes estructuras, con el fin de elegir aquella con la cual se obtienen los mejores resultados.

En este caso la función de transferencia de cada una de las neuronas es la función sigmoideal (ecuación 7), excepto en las neuronas de la capa de salida, en las cuales se ha implementado una función lineal, aumentando así la velocidad de convergencia en el entrenamiento de los pesos.

$$f(x) = \frac{1}{1+e^{-x}}; \quad (7)$$

El vector salida viene dado por la ecuación (8):

$$y = f(w.x); \quad (8)$$

El ajuste de los pesos de interconexión se realiza por el algoritmo de "backpropagation", que es el más utilizado en el entrenamiento de las redes neuronales. Se trata de un método de descenso por el gradiente, donde se consigue minimizar el error cuadrático dado por la salida de la red y el valor deseado para cierta entrada. Para N muestras de vectores de entrada, el error se define como (9):



Fig. 9: Ventanas para la codificación de los dígitos

$$E^2 = \frac{1}{2} \sum_{k=0}^{N-1} (y_{ki} - y_{kd})^2; \quad (9)$$

En donde y_{ki} es la salida dada por la red ante la entrada k e y_{kd} es la salida deseada para esa misma entrada. Partiendo de valores aleatorios en los pesos, éstos se actualizan de forma iterativa según la siguiente expresión (10) (Freeman y Skapura, 1993):

$$\Delta w(t) = -\alpha \frac{\partial E}{\partial w(t)} + \beta \Delta w(t-1); \quad (10)$$

La información de entrada se codifica de la siguiente forma: se obtiene el rectángulo mínimo que encierra cada dígito, considerando que están alineados verticalmente, y seleccionando por tanto las coordenadas de los píxel extremos, y dicho rectángulo se divide en 14x16 regiones, cada una de ellas puede tomar un valor alto o bajo, dependiendo del porcentaje de píxels de esa región que tenga nivel alto. Nótese que la región se adapta a las características del dígito a reconocer, reduciendo el ruido frente a la utilización de una ventana con dimensiones constantes para todos los dígitos. De este modo se consigue una cierta invarianza a la homotecia, de ahí que el sistema desarrollado funcione a diversas distancias.

La codificación de las salidas se realiza teniendo en cuenta que deben reconocerse 37 dígitos (27 letras + 10 números)- se ha considerado la duplicidad entre ciertos números y letras: "0" y "o", "1" e "l" y "5" y "S". La red se entrena de tal forma que solamente el nivel de una de las salidas se pone a nivel alto, precisamente aquella que corresponde con el dígito de entrada (de este modo se consigue mayor inmunidad al ruido). Como medida de seguridad, y con el fin de evitar falsos reconocimientos de dígitos, para que una salida se considere un nivel alto, ésta debe ser mayor que 0,75. En caso contrario, se considera que no existe ninguna salida dominante y se da una señal indicando que no es posible identificar dicho dígito.

En cada momento, la red neuronal realiza el reconocimiento de un dígito, que le son presentados de forma ordenada de izquierda a derecha, mediante una subrutina que maneja las ventanas para su codificación. Esta subrutina supervisa el funcionamiento de la red, de tal modo que es posi-

ble saber previamente si se intenta reconocer un número o una letra de la matrícula: por ejemplo, se sabe que los cuatro dígitos centrales corresponden con números y los de los extremos son letras.

Los parámetros α y β toman el valor de 0.2 y 0.01 respectivamente. La red se entrena con 8 muestras de cada uno de los dígitos, y se finaliza el entrenamiento cuando se cumple para la expresión 9 que:

$$E^2 < 0.001 \quad (11)$$

PRUEBAS Y VALIDACIÓN DEL MÉTODO

Con el fin de comprobar la fiabilidad del sistema diseñado, se ha realizado un test con imágenes de automóviles que no han sido utilizadas en el entrenamiento de la red neuronal. Sobre una muestra de 50 casos, se ha comprobado una fiabilidad del 98%, reconociendo incluso placas sucias o con abollamientos. El tiempo medio de análisis de una imagen se puede establecer en 1,8 sg. Las principales características del equipo utilizado son: cámara en color SONY (CMA-999/999P), tarjeta de procesamiento de imágenes MATROX-640 y ordenador personal PC-486.

CONCLUSIONES

Se ha diseñado un sistema automático para la lectura de las placas de matrículas de vehículos. Trabajando con las características de color de la imagen, es posible conseguir una alta fiabilidad en la lectura, aún en condiciones de luz

cambiantes, todo ello sin utilizar un sistema especial de iluminación. Debido a que se hace un entrenamiento de los dígitos de matrículas en condiciones reales, el sistema funciona correctamente con matrículas sucias, o cuando el automóvil no está centrado en la escena.

REFERENCIAS

- Arregui M.A. y J. A. Mitxelena, Automatic recognition of plate numbers. 14th Spanish Symposium: Pattern recognition and image analysis, España - (1990).
- Cowel J.R., Sybctactic pattern recognizer for vehicle identification numbers. Image and Vision Computing. 13 (1) (1995).
- Freeman, J.A. y D. M. Skapura, Redes neuronales. Algoritmos, aplicaciones y técnicas de programación. Ed. Addison-Wesley. (1993)
- Fu K.S., R. C. González, Lee C.S.G. Robótica: Control, detección, visión e inteligencia. Ed. McGraw-Hill. (1988).
- Fuertes, A simmetry finding as a mean to identify particular cars in a parking. VI Spanish Symposium on Pattern Recognition and Image Analysis. Córdoba, España, (1995).
- Pratt, W. K. Digital Image Processing. John Wiley & Sons Inc. (1991).
- Savatier F. , Barcelona Untangles Traffic. A busy city finds an RF solution to regulation traffic in its overcrowded center. Id Systems European Edition. (4), 1 (1996).