

Reconocimiento de Caracteres en Imagen Digital, una Aplicación al Reconocimiento Automático de Placas de Vehículos

Ricardo Rolón Alonso

Facultad Politécnica, Universidad Nacional del Este
Avda. Monseñor Rodríguez c/ Los Sauces
Ciudad del Este, Paraguay
ricardo@tecnoserv.com.py

y

Vicente Domínguez Torres

Facultad Politécnica, Universidad Nacional del Este
Avda. Monseñor Rodríguez c/ Los Sauces
Ciudad del Este, Paraguay
chikidom@politec.une.edu.py

Resumen

Este trabajo presenta los resultados obtenidos y las propuestas de perfeccionamiento del desarrollo de un sistema de visión artificial, capaz de localizar y reconocer el contenido de una placa de licencia de un vehículo para su gestión en una base de datos, dándole la capacidad de percepción próxima a la de los seres humanos para reconocer objetos; en este caso son caracteres y números presentes en una imagen.

Palabras clave: OCR, procesamiento de imágenes, binarización, reconocimiento

1. Introducción

Un sistema automático de análisis de imagen, también llamado sistema de visión artificial, tiene como objetivo adquirir, procesar, analizar e interpretar un determinado tipo de imagen para una aplicación concreta. Generalmente, este sistema posee aplicaciones bien específicas, lo que permite que presente una alta performance para una determinada clase de imagen [3].

La identificación automática de vehículos es una importante aplicación de sistemas de visión artificial. Además del control y monitoreo del tráfico, un sistema de identificación automática de vehículos posee otras aplicaciones, como:

- medición y planeamiento de flujo de tráfico.
- reconocimiento de vehículos en situación irregular.
- control de peajes y estacionamientos.
- identificación de vehículos infractores.

El sistema fue proyectado para trabajar con imágenes que poseen las siguientes características:

- Placas con 6 caracteres.
- Fondo claro y caracteres oscuros.
- Imágenes de frente de los vehículos.
- Distribución de luz homogénea y heterogénea sobre la superficie de la placa

El aspecto general de este trabajo puede ser el de

una recopilación de diferentes Proyectos concretos en el campo de “Análisis de Imágenes” aplicado al reconocimiento de caracteres en una imagen.



Figura 1: Imagen frontal del vehículo.

2. Metodología

Todos los algoritmos de Reconocimiento Óptico de Caracteres persiguen la finalidad de reconocer un texto para tratarlo posteriormente.

El sistema desarrollado es un clasificador de imágenes. Tal sistema comprende, por lo general, una etapa inicial en que la imagen de partida es procesada para resaltar alguna de sus características más interesantes; una segunda etapa, que realiza medidas sobre la imagen procesada y produce un vector de características; y una tercera etapa, en que el vector de características de la imagen se compara con vectores correspondientes a distintas clases, y se asigna a la clase que se le parezca más, según algún criterio acordado.

2.1-Adquisición

El primer paso consiste en obtener una representación digital de la imagen de interés dentro del ordenador. A partir de ese momento será éste el que permita realizar las operaciones pertinentes sobre dicha imagen para lograr el objetivo propuesto.

Las imágenes de los vehículos se han obtenido fotografiándolos a una distancia y altura constante de su frontal, y almacenándolas en el formato gráfico JPEG.

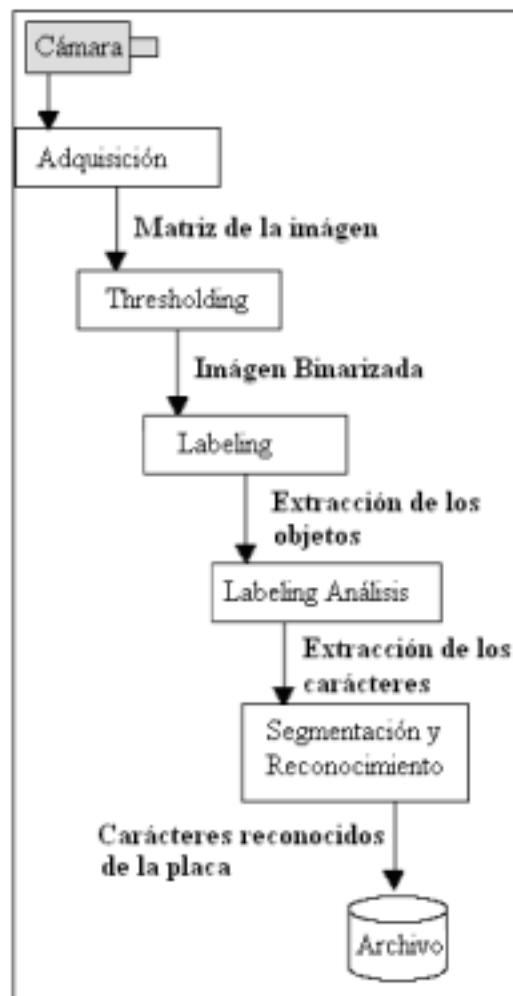


Figura 2: Esquema básico de reconocimiento.

2.2-Procesamiento de la Imagen

En esta etapa se procesa la imagen con el objetivo de hacer de la misma una fuente de información, lo más precisa y confiable, dentro de lo que será el proceso de reconocimiento. Se divide en los siguientes puntos.

Binarización

La mayor parte de los algoritmos para el reconocimiento de la escritura están enunciados a partir de imágenes binarias, por lo que se hace conveniente el paso de una imagen en niveles de gris (o color) a una binaria: esto permite, además, reducir el volumen de los datos que serán procesados.

Se ha utilizado el algoritmo de Bernsen para binarizar la imagen, ya que éste es un método adaptable que ajusta de acuerdo con las variaciones de los píxeles de alto contraste. La figura de abajo muestra el resultado de la binarización.



Figura 3: Thresholding específico de la aplicación.

Extracción de componentes conectados.

La segmentación es una parte del procesamiento digital de imágenes que consiste en la detección de objetos o regiones de interés en ellas [2]. El éxito del proceso depende del paso anterior.

Una vez finalizado el algoritmo de thresholding, la imagen binaria resultante posee una característica fundamental: los píxeles negros corresponderán a todos aquellos elementos que podrían ser objeto de interés, y el blanco, al fondo de la imagen.

El algoritmo de segmentación se basa en la utilización de la técnica de etiquetado de componentes conexas para enumerar la cantidad de objetos en la imagen, de forma a extraer el área de los caracteres en la placa del vehículo.

La figura 4 muestra el resultado del algoritmo de etiquetado de las componentes conexas, cada objeto etiquetado aparece con un color distinto.



Figura 4: Imagen binaria a la que se aplica el etiquetado de conexas.

Análisis de los componentes etiquetados (Labeling Analysis)

Los componentes fueron etiquetados en el paso anterior; ahora se debe analizar cada componente para extraer los caracteres. Cada componente conexas posee sus propias características geométricas similares establecidas por la placa del vehículo.

Algunas características de los caracteres de la placa:

- Están separados en 2 grupos: tres letras y tres números, es decir, contiene seis objetos alineados sobre una misma posición.
- Las alturas son similares entre sí, no así el ancho, generalmente cuando se encuentra el número 1 (uno).
- La distancia entre los caracteres del mismo grupo es constante.
- El algoritmo seguido para localizar el área en que están presentes los caracteres es el siguiente:
- Las características que se utilizan para analizar cada conexas son la altura y el ancho.
- Si el ancho es mayor que el alto, se rechaza, porque un carácter de la placa siempre el alto es mayor que el ancho.
- Se ordenan las conexas en función a su anchura para agrupar las componentes conexas.
- Se establece un valor de corte para dividir cada grupo de componentes caracterizados por su altura.
- Se calcula el promedio de la altura de cada grupo y la cantidad de componentes de cada grupo.
- Se compara la cantidad de componentes mayor o igual a cuatro y menor o igual a seis, para seleccionar el grupo que puede contener los caracteres de la placa.



Figura 5 – Caracteres extraídos después del análisis de cada conexas.

2.2-Reconocimiento de caracteres.

Clasificador de mínima distancia

Se ha utilizado el clasificador de mínima distancia al reconocimiento de caracteres alfanuméricos, es decir, el sistema deberá interpretar una imagen entre los números 0-9 y las letras A hasta Z. Para lo cual se debe haber obtenido de alguna manera la digitalización de la imagen y su representación en una matriz de ceros y unos, donde un cero va a significar un blanco(Background) y un uno, un negro.

Descriptores

El primer problema que surge es encontrar descriptores efectivos para cada imagen, esto es, que describan con bastante exactitud el patrón de la imagen correspondiente, y que además, permitan calcular con facilidad la distancia de ese patrón a cualquier otro patrón de descriptores.

Se estima que un buen patrón para describir una imagen binaria con exactitud, y considerándola como una matriz de unos y ceros, es el número de unos que tiene cada columna; es decir, si la primera columna de esta matriz tiene un número 'n' de unos, el primer valor del vector será 'n'. Siguiendo con este barrido vertical de la matriz (de izquierda a derecha) se obtiene un vector con unos valores (descriptores) que describen esa imagen binaria de una manera bastante fidedigna [1].

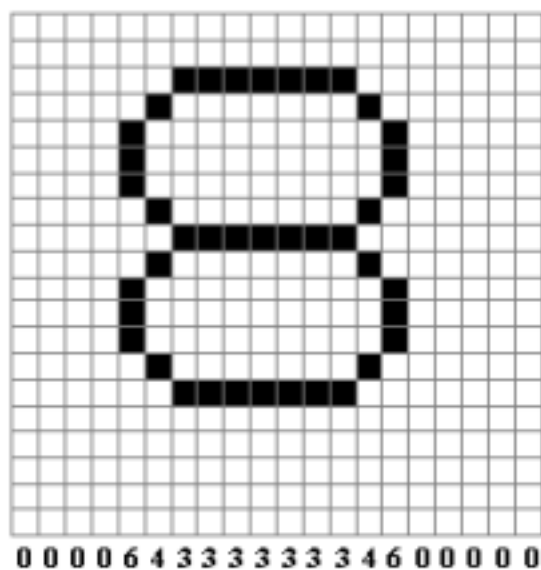


Figura 6: Matriz de descriptor.

Imagen binaria representada por el número de unos (negros) que tiene cada columna. Aunque es difícil imaginar dos imágenes binarias distintas que tengan dos vectores iguales siguiendo este método, esto se puede observar en la siguiente figura.

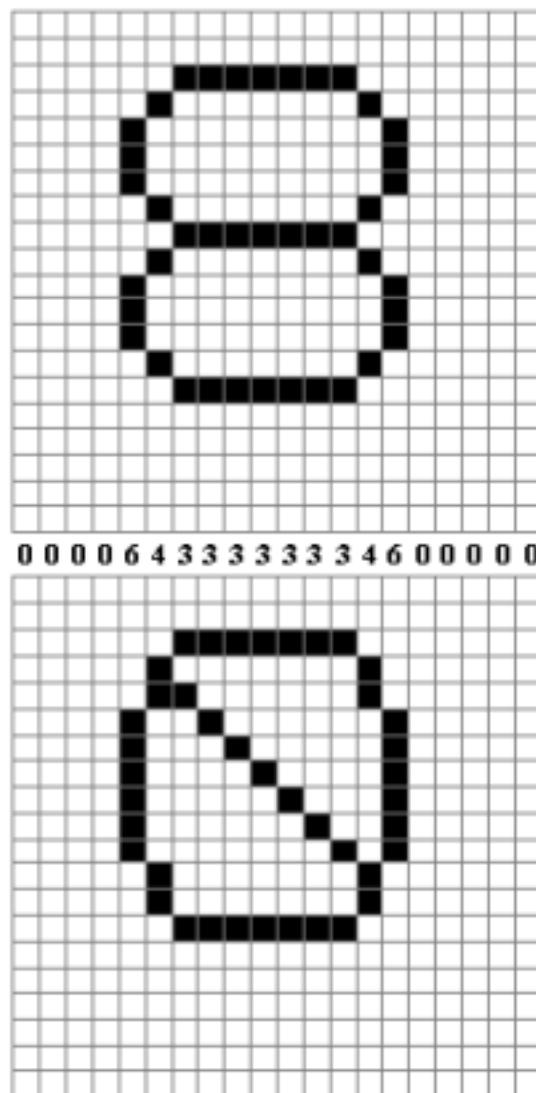


Figura 7: Dos imágenes distintas con el mismo patrón.

A esta secuencia de valores se le añaden otros valores correspondientes a un barrido horizontal de la matriz, es decir, que los siguientes valores serán los números de unos que tiene cada fila. De esta forma queda muy poca posibilidad de repetirse el mismo patrón para imágenes distintas.

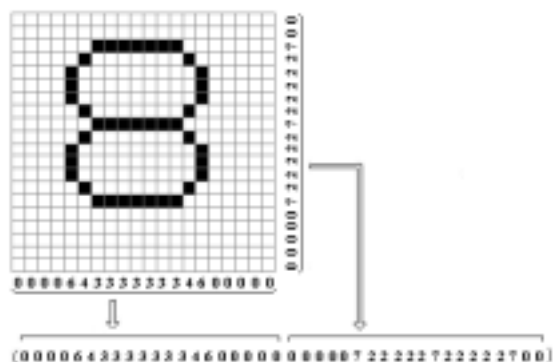


Figura 8: Patrón vectorial que representa la imagen

Los patrones vectoriales se representan de la siguiente forma: $X=[x_1 \ x_2 \ x_3 \ ... \ x_n]$ donde cada componente x_i representa el i -ésimo descriptor, y n es el número de descriptores [1].

Decisión y reconocimiento

Después de la extracción de características se aplica una técnica de decisión o clasificación. La técnica de decisión consiste en la comparación de las características de la imagen de entrada con las de un conjunto de imágenes almacenadas para las cuales se conoce el carácter correcto (conjunto de entrenamiento).

Puede establecerse medidas de similaridad o disimilaridad entre los elementos tomados de dos en dos. La **distancia euclidiana** d entre dos puntos es la norma de diferencia de dos elementos [1].

$$D=\|x-c\| \quad (4.1)$$

Si D es la menor distancia, entonces se asigna X a la clase c . Esto es, en esta formulación, la menor distancia implica el mejor emparejamiento o correspondencia [1].

Aquellos que no realizan una comparación explícita sino que aplican cierta función al patrón de entrada (función de decisión). Estos decisores requieren de una etapa de entrenamiento previo en la cual se presentan al decisor todos los patrones del conjunto de entrenamiento. Esta operación se realiza una y otra vez y sirve para ir ajustando (mediante algún algoritmo de aprendizaje) los parámetros de la función de decisión.



Figura 9: Carácter reconocido de la placa.

2.3-Entrenamiento

Tanto para el entrenamiento como para el test se necesita una base de datos de letras y otra de números, esta base de datos se ha formado extrayendo de las imágenes con diferentes umbrales de conversión a blanco y negro.

Primero el sistema computacional debe ser entrenado para reconocer patrones en los datos. El entrenamiento es el proceso de definir los criterios por los cuales serán reconocidos los patrones.

La figura 10, que representa a la letra "A", muestra su correspondiente descriptor que lo identifica como tal, en la cual será entrenado para su posterior reconocimiento.

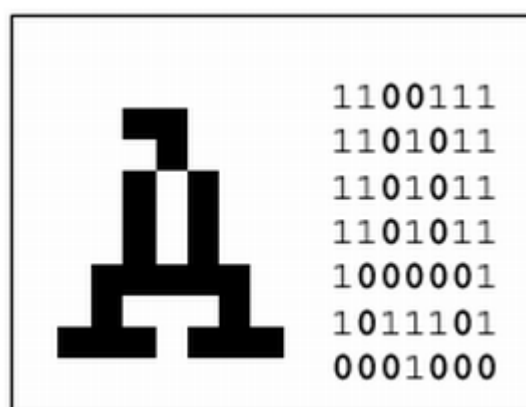


Figura 10: El mapa de bits que representa a la letra "A".

3. Resultados obtenidos

Para la realización del trabajo fue imprescindible la obtención de una gran cantidad de fotografías

digitales, pues a través de ellas fue posible el estudio y desarrollo del modelo computacional propuesto. Se obtuvo un conjunto de imágenes correspondientes a 250 placas de vehículos estacionados.

Un primer análisis del conjunto de imágenes fue realizado para verificar la cantidad de fotografías capaces de ser identificadas por el ser humano. De las 250 imágenes, el 20% no pudieron ser identificadas, por diversos problemas, tales como: imagen de vehículo completamente oscura, con excesiva luminosidad en la placa, con caracteres mal definidos en la placa o con bajo grado de resolución.

En resumen, la evaluación del funcionamiento global del sistema fue realizada con base en las 200 (doscientas) fotos restantes, consideradas imágenes válidas.

Para estimar los resultados obtenidos se tomó como parámetro evaluar dos módulos del sistema: el proceso de segmentación encargado de extraer los caracteres de la imagen de la placa y el proceso de reconocimiento; por otro lado se ha evaluado el tiempo necesario para realizar el reconocimiento de los caracteres de una placa.

En la siguiente tabla se expresan los resultados obtenidos tras el procesamiento de las muestras:

Procesos Evaluados	Letras	Números
Reconocimiento	98.5 %	95 %
Segmentación	99.3%	

Tabla 1: Resultado de las pruebas.

En el test del reconocedor sobre el total de las muestras válidas se ha reconocido perfectamente en ciento noventa y siete placas los primeros tres dígitos (98,5 %) correspondientes a las letras; y los otros 3 dígitos reconocidos en un total de ciento noventa placas (95 %).

El porcentaje de acierto de la localización de los caracteres de las placas fue superior a 98%, constituidos por 3 letras y 3 números. Los errores se debieron a la baja calidad de la imagen procesada, a la falta de luminosidad aliada a la existencia de ruidos generados por el ambiente, o a otro elemento o factor externo incidente en la adquisición de la imagen.

Tiempo de respuesta del sistema

Para analizar el tiempo de respuesta de la aplicación se han utilizado cuatro ordenadores, de tal forma a medir las diferencias dependiendo de la capacidad de procesamiento de cada uno:

Procesador	Memoria RAM	Tiempo de respuesta
Pentium MMX 233	80 Mb	7.3 segundos
AMD 900MHz	128 Mb	4.1 segundos
AMD 1300MHz	128 Mb	3.6 segundos
Pentium4 2.4GHz	256 Mb	3.0 segundos

Tabla 2 : Tiempo de respuesta del sistema

Como se muestra en la tabla de arriba, el tiempo de respuesta en diferentes ordenadores, para realizar el reconocimiento de las placas, funciona en un tiempo bastante aceptable para las necesidades de este tipo de aplicación.

Análisis de los resultados obtenidos

Las principales causas de la falla en el reconocimiento fueron las sombras y el alto contraste en la imagen capturada. La figura 11 muestra una imagen capturada con falta de luminosidad, que afecta al reconocimiento de los caracteres.



Figura 11: Error de binarización

También se han observado errores de segmentación debido a caracteres unidos, ya sea por ruidos en la fase de adquisición, o por alguna suciedad de la placa en el momento de la captura, o debido a algún tornillo situado entre los caracteres, el que hace que los mismos se unan y dificulta la

fase de etiquetado, con la consecuencia de una mala segmentación, tal como muestra la figura 12.



Figura 12 : Error de segmentación.

Otro problema puede constituir la existencia de caracteres externos dentro de la imagen, es decir, objetos como adhesivos o alguna característica de la marca del vehículo, localizados en zonas cercanas a la de la placa, que pueden ocasionar alguna falla en la localización de los caracteres.

Los errores de reconocimiento se deben a la dificultad de clasificar adecuadamente un carácter, principalmente porque cierta distorsión del mismo puede provocar una confusión con otro de estructura similar. La figura 13 muestra un ejemplo de este inconveniente.



Figura 13: Error de reconocimiento.

Basados en los resultados obtenidos se puede considerar un nivel de acierto muy satisfactorio, ya que se ha percibido un bajo margen de error en el proceso de segmentación, considerada ésta una falla muy frecuente en variadas técnicas de reconocimiento.

Se puede concluir que el sistema propuesto funciona acertadamente y podría ser empleado en diversos sistemas que requieran detección automática de placas.

6.1-Revisión de lo realizado

Los principales problemas detectados fueron a causa de la influencia directa de factores externos al sistema; la baja calidad de la imagen capturada puede afectar al funcionamiento general de la aplicación. Otros factores incidentes constituyen el posicionamiento correcto de la placa dentro de la imagen, así como el estado de conservación de la placa y la luminosidad en la adquisición. Este inconveniente evidencia la necesidad de algún control sobre la luminosidad local en la adquisición de la imagen.

6.2-Principales logros

Respecto a la aplicación, aunque ésta podría haberse realizado de manera mucho más eficiente e incluso mejorada para futuras ampliaciones, sin embargo, se considera que el objetivo del trabajo, vale decir, desarrollar un software para la interpretación de caracteres aplicado al reconocimiento de placas de vehículos, fue alcanzado.

REFERENCIAS BIBLIOGRAFICAS

- [1] Universidad de Sevilla – España “Reconocedor por Mínima Distancia”,
<http://www.sav.us.es/formaciononline/asignaturas/asigpid/apartados/textos/recursos/rmd/documentacion.zip>, Noviembre-2003
- [2] Universidad de Sevilla-España - “Generic OCR”,
www.sav.us.es/formaciononline/asignaturas/asigpid/textos/proyecto1/G-OCR2.ppt,
Noviembre-2003
- [3] Tatiane Campos, Sergio Bampi, Altamiro Susin - Instituto de Informática – UFRGS
“Sistema de Identificação de Placas por Processamento Automático de Imagens”
www.iberchip.org/VII/cdnav_pdf_49
Mayo-2004