



UNIVERSIDAD DE JAÉN
ESCUELA POLITÉCNICA SUPERIOR DE LINARES

Trabajo Fin de Grado

DESARROLLO DE HERRAMIENTA DE PROCESADO MORFOLÓGICO PARA MEJORA DE IMÁGENES

Alumno: Alejandro Moya Muñoz

Tutor: Prof. D. Raúl Mata Campos

Depto.: Ingeniería de telecomunicación

Contenido

1. Resumen.....	5
1.1 Abstract	6
2. Introducción	7
2.1 La imagen	7
2.1.1 ¿Qué es la imagen?	7
2.1.2 Tipos de imágenes.....	8
2.1.3 Tipos de imágenes digitales.	9
2.1.4 Clasificación en función del color.....	16
2.2 Vídeo	21
2.2.1 Definición	21
2.2.2 Características	21
2.2.3 Principales formatos.....	24
3. Objetivo	27
4. Estado del arte	29
4.1 ANPR.....	29
5. Materiales y métodos	37
5.1 Procesamiento morfológico.....	37
5.1.1 Morfología	37
5.1.2 Dilatación.....	37
5.1.3 Erosión.....	41
5.1.4 Apertura	46
5.1.5 Cierre	48
5.1.6 Transformaciones Top-Hat y Bottom-Hat	50
5.1.7 Elementos estructurantes.	51
5.2 Filtros detección de bordes.....	51
5.2.1 Introducción.	51
5.2.2 Operadores basados en la primera derivada (Gradiente).....	52
5.3 Conceptos matemáticos.....	60
5.3.1 Correlación de Pearson	60
5.3.2 Umbral óptimo.	64
6. Desarrollo	65
6.1 Modo imagen	65

6.1.1 Localización de forma automática	65
6.1.2 Procesado mediante recorte de la imagen	76
6.2 Modo vídeo	79
6.2.1 Procesado a través de imágenes por dispositivo externo.....	79
6.2.2 Análisis de video.....	83
7. Interfaz gráfica de la aplicación.....	86
7.1 Interfaz principal	88
7.2 Interfaz para procesado de imágenes estáticas	89
7.3 Interfaz para procesado de vídeo.....	100
7.4 Barra de progreso.....	106
8. Resultados y discusión	109
8.1 Resultados en función del tiempo de procesado.....	109
8.2 Resultados en función del porcentaje de imágenes	117
8.3 Relación: tiempo de procesado aciertos.	123
9. Conclusiones.....	124
10. Líneas futuras	126
Anexo I.....	127
Anexo II.....	129
Anexo III.....	130
Referencias bibliográficas	132

1. Resumen

La presente memoria tiene por objeto orientar sobre el funcionamiento y desarrollo de un sistema de detección de matriculas de vehículos mediante la manipulación de imágenes. Este proyecto, como muchos otros, no pretende establecer una nueva base teórica, sino que se apoya en una serie conceptos abiertamente aceptados, a partir de los cuales se intenta avanzar hacia los objetivos propuestos.

La función principal de la aplicación que se desea realizar es la localización y análisis de matrículas tanto en imágenes estáticas como vídeos haciendo uso de técnicas de procesado morfológico para la obtención de los dígitos que componen la matrícula de un coche.

El sistema que se llevará a cabo, para el caso de imágenes, devolverá los caracteres alfanuméricos que componen la matrícula del coche que aparezca en la imagen.

Por otro lado se define otra función para el caso de análisis de vídeos. El sistema, además de dar al usuario una lista con las matrículas de los vehículos que aparecen en el vídeo, también es capaz de dar un número exacto que indique la cantidad de coches que se presentan en el vídeo.

1.1 Abstract

This report aims to provide guidance on the operation and development of a system for detecting vehicle license plates using image manipulation. This project, like many others, is not intended to establish a new theoretical basis, but is based on a series of openly accepted concepts, from which you try to move towards their objectives.

The main function of the application you want to make, is the location and analysis of enrollments in both still images and videos, using morphological processing techniques for obtaining the digits that make up the license plate of a car.

The system will be implemented, in the case of images, by returning the alphanumeric characters that comprise the registration of the car that appears in the picture.

Furthermore, another function for the case of video analysis is defined. The system also gives the user a list of vehicle registration numbers, that appear in the video, is also able to give an exact number indicating the number of cars that appear in the video.

2. Introducción

En este capítulo se realizará una breve introducción sobre los diferentes tipos de imágenes que existen (en función del formato y como representan la información). Además se explicarán brevemente algunos conceptos básicos y características sobre el vídeo.

2.1 La imagen

2.1.1 ¿Qué es la imagen?

Una imagen es una reproducción de una persona o cosa y la formación de la imagen es el proceso de reproducción. Por lo tanto, las imágenes son representaciones de objetos que son captados a través de su energía radiante, es decir, la luz.

De esta definición se puede observar que para la formación de una imagen se requiere de una fuente radiante, un objeto y un sistema de formación. Existe una gran variedad de fuentes radiantes como son las fuentes de luz blanca, sistemas de láser, tubos de rayos X, fuentes térmicas e incluso fuentes de ondas acústicas.

La naturaleza de la fuente radiante tiene gran influencia en la estructura del sistema de formación. Los sistemas de formación pueden ser biológicos (visión humana y animal), fotoquímicos (cámaras fotográficas) o fotoelectrónicos (cámaras de televisión). Debido a esta diversidad es difícil construir un modelo general de formación de imágenes. La Figura 2.1 muestra un modelo básico general que puede ser usado en procesamiento digital de imágenes y aplicaciones de visión computarizada.



Figura 2.1 Modelo de un sistema de procesamiento digital de una imagen

En la figura: f = distribución de energía radiante reflejada en el objeto.

b = distribución de energía radiante de la imagen.

i = intensidad de luz transformada en corriente eléctrica (detector fotoeléctrico).

g = imagen digitalizada.

Una estrecha banda de radiación electromagnética puede estimular los fotoreceptores biológicos y los sensores de imágenes artificiales. Generalmente la luz visible tiene una longitud de onda λ que oscila entre 380 nm (ultravioleta) y 700 nm (infrarrojo). Sin embargo, existen sensores de imagen que pueden salirse del rango de la luz visible. Las fuentes de radiación visible producen luz de cualquier longitud de onda dentro del rango del espectro visible, mientras que las fuentes de luz blanca producen radiación en todo el espectro visible.

2.1.2 Tipos de imágenes.

Las imágenes se pueden presentar en dos formas, digitales o analógicas.

Las imágenes digitales están formadas por píxeles a diferencia de las analógicas que son de “grano” sales de plata (halogenuro de plata). El grano fotográfico es irregular en forma y tamaño, con formas redondeadas o en forma de judía que cubren el espacio con aleatoriedad, mientras que los píxeles son cuadrados, regulares y absolutamente iguales.

Las imágenes formadas por el grano adquieren una calidad especial, una textura muy rica visualmente que se añade a la imagen en cambio las imágenes con píxeles toman esa trama fría y matemática, la textura regular y sus cuadros de color liso, así el grano lo podemos aprovechar para determinadas calidades y los píxeles conseguir que sean tan

pequeños que dejen de apreciarse.

Grano y pixel son pues dos formantes de la imagen de naturaleza opuesta. Orgánico frente a regular, cálido frente a frío, desordenado frente a repetible. Son conceptos opuestos que se enfrentan en cada una de sus características.

La imagen digital en lugar de almacenarse la información en forma de procesos químicos se almacena como electricidad. En vez de almacenarse la información con valores continuos se almacena con valores discretos, reduciendo la riqueza de la información. Son todos factores que diferencian notablemente ambos sistemas, pero recordamos que sólo se diferencian en la fase de almacenamiento de la información, en lo concerniente al soporte y su capacidad para recoger y almacenar esa información.

En el siguiente ejemplo podemos apreciar la diferencia entre ‘grano’ y pixel:

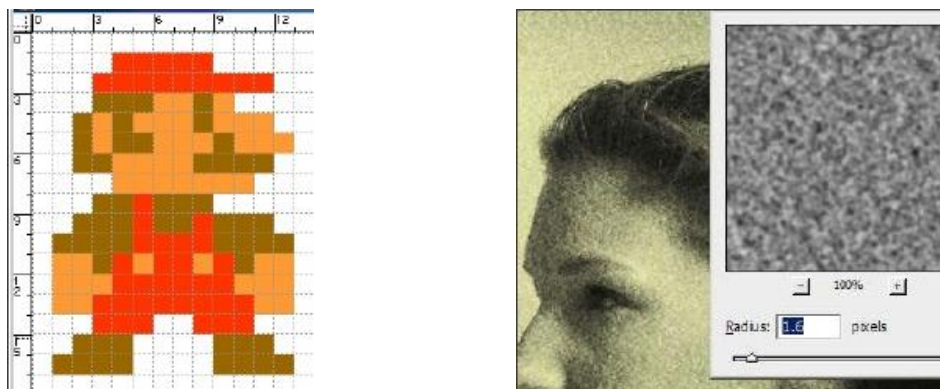


Figura 2.2 Diferencia analógica y digital

Para el propósito del proyecto nos centraremos exclusivamente en imágenes digitales.

2.1.3 Tipos de imágenes digitales.

A grandes rasgos, hay dos tipos de imágenes digitales:

- Imágenes vectoriales

- Imágenes de mapa de bits.

Las imágenes vectoriales son imágenes constituidas por objetos geométricos autónomos (líneas, curvas, polígonos,...), definidos por ciertas funciones matemáticas (vectores) que determinan sus características (forma, color, posición,...)

Las imágenes de mapa de bits están formadas por una serie de puntos (píxeles), cada uno de los cuales contiene información de color y luminosidad. Salvando la diferencia, podemos compararla con un mosaico y sus teselas.

Las imágenes vectoriales se crean con programas de diseño o dibujo vectorial (Adobe Illustrator, Corel Draw, Inkscape...) y suelen usarse en dibujos, rótulos, logotipos... Su principal ventaja es que una imagen puede ampliarse sin sufrir el efecto de “*pixelado*” que tienen las imágenes de mapa de bits al aumentarse.

Las imágenes de mapa de bits presentan una mayor gama de colores y de tonos que las vectoriales, por lo que son el tipo de imágenes usado en fotografía y, se crean con las cámaras de fotos, los escáneres y con programas de edición de imagen y dibujo (Adobe Photoshop, Gimp, etc.) Las imágenes mapa de bits generan archivos que ocupen mucha más memoria (bytes) que las imágenes vectoriales.

Para poder reproducirse o utilizarse en un ordenador u otros dispositivos las imágenes vectoriales y de mapa de bits se guardan en archivos o ficheros (conjunto de datos que se almacenan en algún medio –disco duro, DVD, lápiz de memoria...–) Cada archivo gráfico, se identifica además de por su nombre, por su extensión, que indica el tipo o formato de que se trata.

Algunos formatos de imagen vectorial son: AI (Adobe Illustrator), CDR (Corel Draw), DXF. (Autodesk), EMF, EPS, ODG (Open Office Draw), SVG (Inkscape), SWF (Adobe flash), WMF (Microsoft).

2.1.3.1 Formatos imagen de mapa de bits.

Para efectuar una elección adecuada de un formato de imagen, conviene valorar previamente:

- El contenido de la imagen: foto, dibujo, gráfico, logotipo...
- La calidad que se desea obtener en función de su destino: publicación en la web, impresión en impresora doméstica, impresión profesional,...
- El tamaño que tendrá el archivo resultante.

Algunos formatos de mapa de bits son los siguientes:

BMP (Bitmap)

Este es un Formato de mapa de bits desarrollado por Microsoft para utilizarlo en Windows 3.X y OS/2 Presentaron Manager. BMP provee un formato común de archivo para varias aplicaciones básicas de Windows. Es un formato estándar para Windows y es el tipo de archivo creado por el programa Paintbrush.

El formato BMP es un registro de los bits reales en un archivo, por lo tanto no puede ser fácilmente ajustado si se cambia la resolución de la pantalla. Los archivos BMP son rápidos para ser cargados y desplegados en Windows.

- La imagen se forma mediante una parrilla de píxeles.
- El formato BMP no sufre pérdidas de calidad y por tanto resulta adecuado para guardar imágenes que se desean manipular posteriormente.
- Ventaja: Guarda gran cantidad de información de la imagen.
- Inconveniente: El archivo tiene un tamaño muy grande.

GIF (Graphics Interchange Format).

El formato GIF, en general, puede comprimir una imagen en un archivo de menor tamaño que cualquier otro formato popular de archivos gráficos. GIF fue desarrollado por CompuServe Information Service, en parte para crear un formato de imágenes que pueda ser utilizado en cualquier computador. El objetivo fundamental al crear GIF fue el de permitir observar imágenes en línea desde el servidor de una red, debido a lo cual ganó

rápidamente respaldo de una gran cantidad de fabricantes de software de procesamiento de imágenes. GIF no asume ningún tipo particular de computador o adaptador gráfico, el software es el responsable de hacer que la imagen aparezca correctamente en la pantalla.

- Ha sido diseñado específicamente para comprimir imágenes digitales.
- Reduce la paleta de colores a 256 colores como máximo (profundidad de color de 8 bits).
- Admite gamas de menor número de colores y esto permite optimizar el tamaño del archivo que contiene la imagen.
- Ventaja: Es un formato idóneo para publicar dibujos en la web.
- Inconveniente: No es recomendable para fotografías de cierta calidad ni originales ya que el color real o verdadero utiliza una paleta de más de 256 colores.

JPG-JPEG (Joint Photographic Experts Group)

Fue desarrollado por la Unión del Grupo de Expertos Fotográficos. Es además, uno de los métodos de compresión con riesgo de pérdida de información más populares.

El archivo .JPG está compuesto por datos altamente comprimidos, inclusive se descartan algunos datos cuyo número puede ser controlado por el usuario. Este nuevo formato reduce significativamente el espacio requerido para almacenar características de calidad y claridad de las imágenes.

- Admite una paleta de hasta 16 millones de colores.
- La compresión JPEG puede suponer cierta pérdida de calidad en la imagen. En la mayoría de los casos esta pérdida se puede asumir porque permite reducir el tamaño del archivo y su visualización es aceptable.
- Cada vez que se modifica y guarda un archivo JPEG, se puede perder algo de su calidad si se define cierto factor de compresión.
- Las cámaras digitales suelen almacenar directamente las imágenes en formato JPEG con máxima calidad y sin compresión.
- Ventaja: Es ideal para publicar fotografías en la web siempre y cuando se configuren adecuadamente dimensiones y compresión.
- Inconveniente: Si se define un factor de compresión se pierde calidad. Por este motivo no es recomendable para archivar originales.

TIF-TIFF (Tagged Image File Format)

El formato TIFF es un formato de gráficos antiguo, que permite almacenar imágenes de mapas de bits (trama) muy grandes (más de 4 GB comprimidos) pero perdiendo calidad y sin considerar las plataformas o periféricos utilizados (*Mapa de bits independiente del dispositivo*, reconocido como *DIB*).

El formato TIFF permite almacenar imágenes en blanco y negro, en colores verdaderos (hasta 32 bits por píxel) y también indexar imágenes utilizando una paleta.

- Almacena imágenes de una calidad excelente.
- Utiliza cualquier profundidad de color de 1 a 32 bits.
- Es el formato ideal para editar o imprimir una imagen.
- Ventaja: Es ideal para archivar archivos originales.
- Inconveniente: Produce archivos muy grandes.

PNG (Portable Network Graphic)

Es un formato gráfico basado en un algoritmo de compresión sin pérdida para bitmaps no sujeto a patentes. Este formato fue desarrollado en buena parte para solventar las deficiencias del formato GIF y permite almacenar imágenes con una mayor profundidad de contraste y otros importantes datos.

- Tiene una tasa de compresión superior al formato GIF (+10%)
- Admite la posibilidad de emplear un número de colores superior a los 256 que impone el GIF.

Comparación de los cinco formatos expuestos anteriormente:



Figura 2.3 BMP GIF JPG TIFF PNG

2.1.3.2 Formato imágenes vectoriales.

Existen varios formatos para este tipo de imágenes. A continuación se expondrán los más importantes y utilizados.

EPS

Un Archivo EPS está formado por dos partes: Una PostScript (que es la que se imprimirá) y otra parte de previsualización (que se mostrará en la pantalla). Esa zona de previsualización se hace sólo para dar una idea de qué contiene el EPS: Es de baja resolución, con pocos colores e, incluso, puede no existir.

WMF

WMF (Windows MetaFile Format) es un metaformato de 16 bits de los sistemas operativos Windows, siendo un estándar de intercambio de gráficos entre las diferentes aplicaciones Microsoft.

WMF es un formato vectorial (aunque no basado en curvas de Bézier) y escalable, que funciona copiando en un archivo los comandos para realizar la imagen en cuestión, ahorrando con ello una cantidad considerable de espacio. Teóricamente puede almacenar cualquier elemento gráfico, ya sean imágenes bitmap, textos o gráficos vectoriales complejos.

EMF

EMF (Enhanced MetaFile) es un metaformato gráfico vectorial de 32 bits, reconocido por casi todas las aplicaciones de diseño gráfico y compatible con los sistemas operativos Windows.

Junto a las características propias de los formatos vectoriales presenta la ventaja adicional de que sus ficheros pueden ser creados rápidamente, ya que lo que se encola en la impresora son comandos de dibujo, con lo que se puede evitar la sobrecarga en el caso de impresión remota de ficheros gráficos.

Además, este formato es más eficiente porque genera un archivo relativamente pequeño y genérico, que es compatible con todas las impresoras.

Como desventaja, los archivos de formato EMF no contienen la misma cantidad de detalles que los de otros tipos de formatos gráficos vectoriales, como los archivos DWF.

CDR

Es un formato vectorial, pero admite la inclusión de elementos de mapa de bits (integrados o vinculados a ficheros externos), pudiendo llevar además cabecera de previsualización (thumbnail). Es uno de los formatos con más posibilidades con respecto al color, a la calidad de los diseños y al manejo de fuentes, pudiendo contener los textos trazados o con fuentes incluidas.

Una de las principales desventajas de este formato es su falta de compatibilidad con el resto de aplicaciones gráficas (solo funciona para Corel Draw), al ser éstas incapaces de almacenar imágenes bajo este formato.

2.1.4 Clasificación en función del color

Las imágenes se representarán en el proyecto en tres formas diferentes:

- RGB.
- Escala de grises.
- Blanco y negro.
- Indexada.

2.1.4.1 Imagen RGB

RGB (Red, Green y Blue) es la composición del color en términos de la intensidad de los colores primarios de la luz.

RGB es un modelo de color basado en la síntesis aditiva, con el que es posible representar un color mediante la mezcla por adición de los tres colores de luz primarios. El modelo de color RGB no define por sí mismo lo que significa exactamente rojo, verde o azul, por lo que los mismos valores RGB pueden mostrar colores notablemente diferentes en diferentes dispositivos que usen este modelo de color. Aunque utilicen un mismo modelo de color, sus espacios de color pueden variar considerablemente.

Los ojos humanos tienen dos tipos de células sensibles a la luz o fotorreceptores: los bastones y los conos. Estos últimos son los encargados de aportar la información de color.

Para saber cómo es percibido un color, hay que tener en cuenta que existen tres tipos de conos con respuestas frecuenciales diferentes, y que tienen máxima sensibilidad a los colores que forman la terna RGB. Aunque los conos, que reciben información del verde y el rojo, tienen una curva de sensibilidad similar, la respuesta al color azul es una veintava (1/20) parte de la respuesta a los otros dos colores. Este hecho lo aprovechan algunos sistemas de codificación de imagen y vídeo, como el JPEG o el MPEG, "perdiendo" de manera consciente más información de la componente azul, ya que el ser humano no percibe esta pérdida.

La sensación de color se puede definir como la respuesta de cada una de las curvas de sensibilidad al espectro radiado por el objeto observado. De esta manera, obtenemos tres respuestas diferentes, una por cada color.

El hecho de que la sensación de color se obtenga de este modo, hace que dos objetos observados, radiando un espectro diferente, puedan producir la misma sensación. Y en esta limitación de la visión humana se basa el modelo de síntesis del color, mediante el cual podemos obtener a partir de estímulos visuales estudiados y con una mezcla de los tres colores primarios, el color de un objeto con un espectro determinado.

Para indicar con qué proporción se mezcla cada color, se asigna un valor a cada uno de los colores primarios, de manera que el valor 0 significa que no interviene en la mezcla y, a medida que ese valor aumenta, se entiende que aporta más intensidad a la mezcla. Aunque el intervalo de valores podría ser cualquiera (valores reales entre 0 y 1, valores enteros entre 0 y 37, etcétera), es frecuente que cada color primario se codifique con un byte (8 bits). Así, de manera usual, la intensidad de cada una de las componentes se mide según una escala que va del 0 al 255.

Por lo tanto, el rojo se obtiene con (255, 0, 0), el verde con (0, 255, 0) y el azul con (0,0,255), obteniendo, en cada caso un color resultante monocromático. La ausencia de color —lo que se conoce como color negro— se obtiene cuando las tres componentes son 0, (0, 0, 0). La combinación de dos colores a nivel 255 con un tercero en nivel 0 da lugar a tres colores intermedios. De esta forma el amarillo es (255, 255, 0), el cyan (0, 255, 255) y el magenta (255,0, 255). Obviamente, el color blanco se forma con los tres colores primarios a su máximo nivel (255, 255, 255).

El conjunto de todos los colores se puede representar en forma de cubo. Cada color es un punto de la superficie o del interior de éste. La escala de grises estaría situada en la diagonal que une al color blanco con el negro.

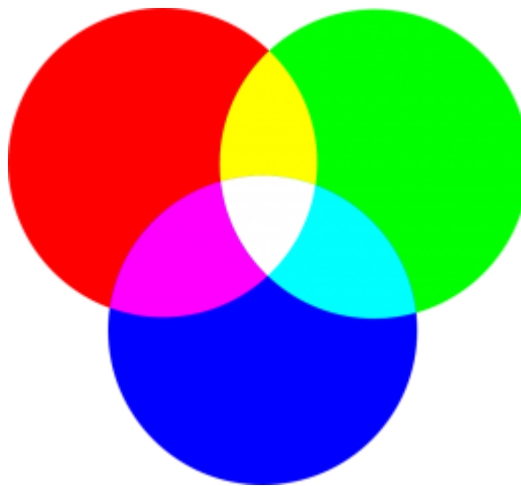


Figura 2.4 Modelo RGB

2.1.4.2 Imagen escala de grises.

Este modo utiliza hasta 256 tonos de gris. Cada píxel de una imagen en escala de grises tiene un valor de brillo comprendido entre 0 (negro) y 255 (blanco). Los valores de la escala de grises también se pueden medir como porcentajes de cobertura de la tinta negra (0% es igual a blanco, 100% a negro). Normalmente, las imágenes generadas con escáneres en blanco y negro o escala de grises se visualizan en modo Escala de grises.

Aunque Escala de grises es un modo de color estándar, puede variar el rango exacto de los grises representados, dependiendo de las condiciones de impresión.

2.1.4.3 Imagen en blanco y negro.

Son imágenes monocromáticas. Da la sensación de estar compuesto por un sólo color, aunque pueda tener diversos grados e intensidades: De lo más claro (usualmente blanco) al más intenso (el color al 100% de su intensidad). En este caso solo blanco y negro.

Una imagen en la que para crear el motivo se ha usado un único pigmento o fuente luminosa y las variaciones tonales se han hecho variando la cantidad de pigmento o luz es monocromática o monocroma.

Así ocurre con la fotografía en blanco y negro, con la televisión en blanco y negro (la única forma de "dibujar" era variar la intensidad luminosa de cada punto de la pantalla), etc.

Una luz que no dé la sensación de ser blanca y que no varíe en su composición espectral de forma perceptible también es monocromática. Al no ser neutra ni tender a ello y no variar, es lo que se llama una "luz de color" monocromática (una bombilla roja, por ejemplo). Si parpadea sigue siendo monocromática (su composición no varía perceptiblemente, sólo su intensidad).

Una imagen en blanco y negro consiste en un conjunto de píxeles cuyos valores solo son blanco o negro. Consiste en obtener una imagen con los dos valores límite de la escala de grises. Suele ser llamada de forma abreviada imagen B/W (del inglés, black and white).

2.4.4 Imagen indexada.

Esta es una forma práctica de representar imágenes en color. Una imagen indexada almacena una imagen como dos matrices. La primera matriz tiene el mismo tamaño que la imagen y un número para cada píxel. La segunda matriz se denomina mapa de color y su tamaño puede diferir del de la imagen. Los números de la primera matriz son una instrucción de qué color usar según el mapa de color.

Este modo reduce la gama cromática de las imágenes a un máximo de 256 colores, con lo que el tamaño del archivo también disminuye. Al cambiar una imagen de modo RGB a la modalidad de color indexado, el programa reduce los tonos presentes en la misma, con lo que la imagen pierde calidad si tiene mucha riqueza cromática.

Es el modo propio del formato GIF, un tipo de archivo de imagen destinado a internet. Su pequeño tamaño en disco, así como su posibilidad de animación y soporte de transparencias lo hacen ideal para la web. En imágenes con degradados o muchos colores este no presenta mucha calidad, pues el GIF y el modo indexado) no la ofrece.

Aunque la paleta de colores es limitada, el modo Color indexado puede reducir el tamaño de archivo manteniendo la calidad visual necesaria para presentaciones multimedia y usos similares

2.4.5 Comparativa de formatos de color.

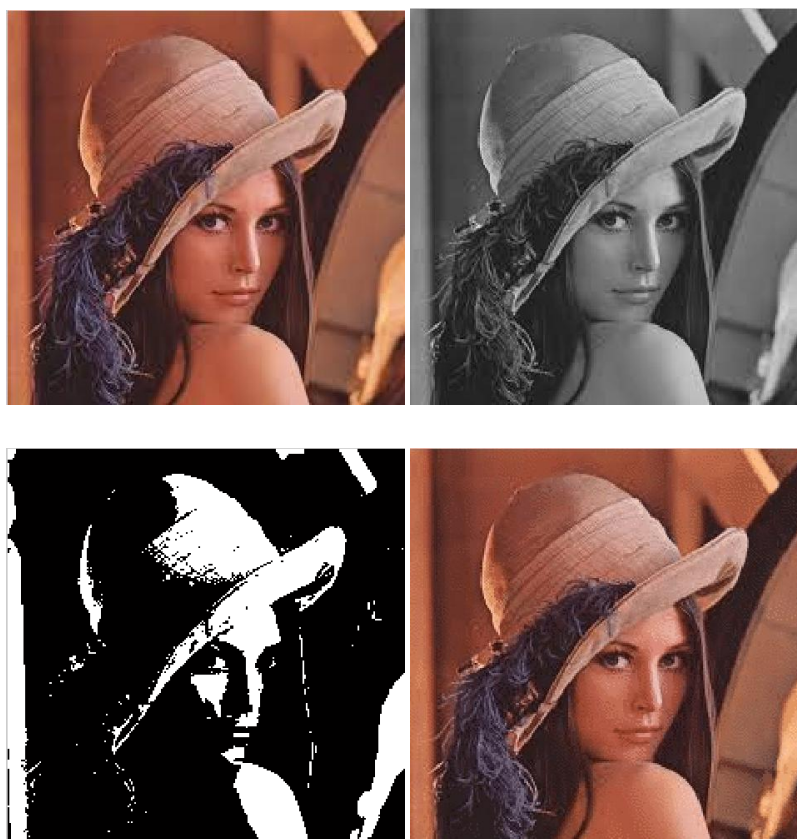


Figura 7.2.5 RGB. Escala de grises. Blanco y negro. Indexada

2.2 Vídeo

2.2.1 Definición

El video es una tecnología utilizada para capturar, grabar, procesar, transmitir y reproducir una secuencia de imágenes representativas de una escena que se encuentra en movimiento. Inicialmente la señal de vídeo está formada por un número de líneas agrupadas en varios cuadros y estos a la vez divididos en dos campos portan la información de luz y color de la imagen. El número de líneas, de cuadros y la forma de portar la información del color depende del estándar de televisión concreto.

2.2.2 Características

- **Número de fotogramas por segundo**

Velocidad de carga de las imágenes: Un video resulta de la exposición imágenes o fotogramas uno detrás de otro. Un parámetro de la calidad del video es el número de fotogramas por segundo que muestra durante su reproducción. Este valor oscila entre 15 y 30. El número de imágenes por unidad de tiempo de video, para viejas cámaras mecánicas cargas de seis a ocho fotogramas por segundo (fps) o 120 imágenes por segundo o más para las nuevas cámaras profesionales. El cine es más lento con una velocidad de 24fps, lo que complica un poco el proceso de transferir una película de cine a video. Para lograr la ilusión de una imagen en movimiento, la velocidad mínima de carga de las imágenes es de unas quince imágenes por segundo, sin embargo el ojo humano puede distinguir movimiento mucho más fluida por encima de los 48 fotogramas por segundo.

- **Sistemas de barrido**

Entrelazado

La exploración entrelazada 2/1, característica de los sistemas de televisión PAL, NTSC y SECAM así como de algunos otros desarrollados posteriormente, consiente en analizar cada cuadro (frame) de la imagen en dos semicuadros iguales denominados campos (field), de forma que las líneas resultantes estén imbricadas entre sí alternadamente por superposición. Uno de los campos contiene las líneas pares, se le denomina "campo par", mientras que el otro contiene la impares, se le denomina "campo impar" al comienzo de cada uno de ellos se sitúa el sincronismo vertical. Hay un desfase de media línea entre

un campo y otro para que así el campo par explore la franja de imagen que dejó libre el campo impar. La exploración entrelazada de un cuadro de dos campos exige que el número de líneas del cuadro sea impar para que la línea de transición de un campo al otro sea divisible en dos mitades.

Las especificaciones abreviadas de la resolución de vídeo a menudo incluyen una *i* para indicar entrelazado. Por ejemplo, el formato de vídeo PAL es a menudo especificado como 576i50, donde 576 indica la línea vertical de resolución, *i* indica entrelazado, y el 50 indica 50 cuadros (la mitad de imágenes) por segundo.

Progresivo

En los sistemas de barrido progresivo, en cada período de refresco se actualizan todas las líneas de exploración.

Un procedimiento conocido como desentrelazado puede ser utilizado para transformar el flujo entrelazado, como el analógico, el de DVD, o satélite, para ser procesado por los dispositivos de barrido progresivo, como el que se establece en los televisores TFT, los proyectores y los paneles de plasma.

- **Resolución del vídeo**

El tamaño de una imagen de vídeo se mide en píxeles para vídeo digital, o en líneas de barrido horizontal y vertical para vídeo analógico. En el dominio digital, la televisión de definición estándar (SDTV) se especifica como $720/704/640 \times 480i60$ para NTSC y $768/720 \times 576i50$ para resolución PAL o SECAM.

- **Relación de aspecto**

La relación de aspecto se expresa por la anchura de la pantalla en relación a la altura. El formato estándar hasta el momento en que se comenzó con la estandarización de la televisión de Alta resolución tenía una relación de aspecto de 4/3. El adoptado es de 16/9. La compatibilidad entre ambas relaciones de aspecto se puede realizar de diferentes formas.

Una imagen de 4/3 que se vaya a ver en una pantalla de 16/9 puede presentarse de tres formas diferentes:

- Con barras negra verticales a cada lado (*letterbox*).
- Agrandando la imagen hasta que ocupe toda la pantalla horizontalmente.

- Deformando la imagen para adaptarla al formato de la pantalla.

Una imagen de 16/9 que se vaya a ver en una pantalla de 4/3, de forma similar, tiene tres formas de verse:

- Con barras horizontales arriba y abajo de la imagen (*letterbox*). Se ve toda la imagen pero se pierde tamaño de pantalla.
- Agrandando la imagen hasta ocupar toda la pantalla verticalmente, perdiéndose las partes laterales la imagen.
- Deformando la imagen para adaptarla a la relación de aspecto de la pantalla.

- **Espacio de color y bits por píxel**

El número de colores distintos que pueden ser representados por un píxel depende del número de bits por píxel (bpp). Una forma de reducir el número de bits por píxel en vídeo digital se puede realizar por submuestreo de croma (por ejemplo, 4:2:2, 4:1:1, 4:2:0).

- **Método de compresión de vídeo (sólo digital)**

Se usa una amplia variedad de métodos para comprimir secuencias de vídeo. Los datos de vídeo contienen redundancia temporal, espacial y espectral. En términos generales, se reduce la redundancia espacial registrando diferencias entre las partes de una misma imagen (frame); esta tarea es conocida como compresión intraframe y está estrechamente relacionada con la compresión de imágenes. Así mismo, la redundancia temporal puede ser reducida registrando diferencias entre imágenes (frames); esta tarea es conocida como compresión interframe e incluye la compensación de movimiento y otras técnicas.

- **Tasa de bits (sólo digital)**

La tasa de bits es una medida de la tasa de información contenida en un flujo o secuencia de video. La unidad en la que se mide es bits por segundo (bit/s o bps) o también Megabits por segundo (Mbit/s o Mbps). Una mayor tasa de bits permite mejor calidad de vídeo. La VBR (Variable Bit Rate – Tase de Bits Variable) es una estrategia para maximizar la calidad visual del vídeo y minimizar la tasa de bits. En las escenas con movimiento rápido, la tasa variable de bits usa más bits que los que usaría en escenas con movimiento lento pero de duración similar logrando una calidad visual consistente.

2.2.3 Principales formatos

Los formatos más populares y con más uso en la actualidad son los siguientes:

MPEG-4

Utilizando la extensión ".mp4", incluye en su interior audio M4A y video M4V, y pueden ser creados a partir de archivos en formato AAC, Apple Lossless y MP3, entre otros.

.mp4: Contiene audio, video y otros contenidos avanzados, tales como transmisión de flujos sobre Internet, puede contener video del tipo MPEG-4, MPEG-2 y MPEG-1, audio MPEG-4 AAC, MP3, MP2, MPEG-1, MPEG-2 y CELP.

.m4a: Sólo audio.

.m4p: Tipo de archivo protegido FairPlay de Apple.

.m4v: Sólo video.

.3gp y .3g2: Es la extensión utilizada en la telefonía móvil, para almacenar audio y video.

DivX

El DivX ha logrado un éxito rotundo en los últimos años, y esto fundamentalmente se debe a que combina la compresión de audio digital en formato MP3, y brinda una muy buena calidad de imagen. Se trata de un códec que permite codificar contenido audiovisual. Es decir que permite una compresión de tamaño de archivos realmente alta, manteniendo cierta calidad con la fuente de origen.

AVI

El formato AVI posee soporte para múltiples flujos de audio y video simultáneo, por lo cual la información que es almacenada en este contenedor pasa previamente por un proceso de entrelazado, haciendo posible que cada uno de los fragmentos del archivo con extensión ".avi" incluya la información necesaria para poder reproducir una serie de fotogramas en conjunto con el audio que lo debe acompañar.

Múltiples flujos de datos, esto significa que ofrece la posibilidad, por ejemplo, de almacenar una película junto con varias pistas de sonido, es decir conteniendo diversas bandas sonoras en distintos idiomas.

WMV

Es un grupo de algoritmos de compresión que pertenecen a la compañía Microsoft, y que fue desarrollado por la empresa para que fuera parte integrante del framework denominado Windows Media. el video WMV puede estar dentro de algún contenedor y de esta manera unido a su respectivo audio en WMA. Por lo general este contenedor es el llamado ASF, que incluye audio y video. También suelen ser frecuentemente utilizados los formatos MPEG-2 y MPEG-4, ya que éstos permiten la difusión de material a través de streaming sin mayores inconvenientes.

RealVideo

Si bien la calidad que ofrecen los archivos de video de RealVideo, que podemos identificar con las extensiones ".rv", ".ram", ".rm", ".rmvb", no suelen ser de una buena calidad, lo cierto es que en los últimos años ha sido utilizado por los usuarios de Internet que buscaban un tipo de formato de gran tasa de compresión para compartir material multimedia en la red.

En principio, es importante mencionar que RealVideo es el formato de video propietario de RealMedia, y que junto con con el formato de audio RealAudio, ha sido desarrollado por la compañía RealNetworks.

Debido a sus características, sobre todo al hecho de que permite generar archivos de tamaño realmente pequeño, el RealVideo y el RealAudio son formatos que suelen ser utilizados para aplicaciones de Streaming.

Debido a que el RealVideo comprime el material audiovisual, se producen ciertas pérdidas que perjudican la calidad del contenido. Sin embargo, en la actualidad, y dependiendo de las configuraciones de codificación, el RealVideo permite alcanzar calidades similares a las que se logran con codificaciones realizadas en MPEG-4.

Flash Video

El formato contenedor llamado Flash Video no es otra cosa que el conocido FLV, el cual ha sido desarrollado por la compañía Adobe con el fin de ofrecer un formato adecuado para realizar transmisiones de video a través de Internet, para lo cual se utiliza Adobe Flash Player.

Una de las características más destacables de este formato reside precisamente en que los contenidos audiovisuales en FLV permiten su incrustación dentro de archivos del

tipo SWF, por lo cual una gran cantidad de sitios en los que se difunden videos se han volcado a utilizar este formato, como es el caso de comunidades y páginas webs tales como YouTube, Google Video, Yahoo! Video y MySpace, entre otros.

La mayor ventaja de este formato es que un video FLV puede ser visualizado en casi todos los sistemas operativos existentes en la actualidad.

3GP para móviles

El formato más utilizado por la mayoría de los teléfonos celulares de la actualidad suele ser el llamado 3GP, el cual consiste en ser un contenedor en el cual es posible almacenar información de diversos medios, tanto de video como de audio. Para lograr su cometido, es decir ofrecer archivos de video con audio realmente pequeños en tamaño, utiliza los códecs MPEG-4 o H.263 para video, y los formatos AMR-NB o AAC-LC para el audio.

Matroska

En principio, debemos destacar que Matroska se trata de un tipo de contenedor estándar abierto. Su mayor característica reside en el hecho de ser un fichero informático que permite contener en sí mismo una cantidad ilimitada de contenido, como video, audio, imágenes, e incluso archivos de texto especial como lo son los subtítulos.

Los archivos Matroska pueden identificarse por sus extensiones, las cuales pueden ser ".mkv" para video que contiene además audio y archivos de subtítulos, ".mka" en aquellos que sólo contiene audio, y ".mks" para subtítulos.

3. Objetivo

El propósito principal de la realización del proyecto es la localización y procesamiento de matrículas de coches a través de operaciones morfológicas. El sistema diseñado consta de varias formas de trabajo. Las diferentes modalidades se pueden dividir en dos tipos: procesamiento de imágenes (estáticas) y manipulación de un vídeo (dinámico) junto con el procesamiento de los frames que contiene dicho vídeo.

Se pretende que el sistema sea capaz de realizar las siguientes tareas:

- Análisis de una imagen individual: se basa en conseguir que la aplicación tenga la capacidad para procesar de forma automática una imagen seleccionada y conseguir obtener el resultado deseado.
- Análisis de una imagen individual mediante recorte: alternativa a la tarea presentada anteriormente. Se pretende crear una herramienta que permita la selección solo de la parte de la imagen que interese, esa área deberá ser donde se sitúa la matrícula. Una vez recortada, esa sección se procesará para la obtención de los dígitos de la placa. Esta alternativa será útil para imágenes con luminosidad homogénea y por tanto con dificultad para la localización de forma automática de la matrícula en la imagen.
- Análisis de un vídeo: procesamiento de un vídeo previamente capturado mediante un dispositivo externo. Se realizará un procesamiento exhaustivo de la totalidad del vídeo, cargado por el usuario, a través de los frames que lo componen. Posteriormente, una vez finalizado el estudio del vídeo, se obtendría el número de coches que aparecen en el vídeo y sus correspondientes matrículas.
- Análisis a través de dispositivo externo conectado al ordenador: este modo de trabajo de la aplicación constará de dos submodos:

- Un submodo permitirá capturar una imagen en tiempo real y procesarla para obtener la información de la matrícula que se pretenda obtener.
- Otro submodo dará la opción de analizar las imágenes que se capturan en tiempo real.

La aplicación podría ser usada en múltiples ámbitos de seguridad y control de tráfico de vehículos. Entre las diferentes posibilidades de uso que permitirá la aplicación se destacan a continuación algunas de ellas:

- Control de entrada y salida de coches en un parking.
- Control del número de coches en un parking.
- Control de tráfico en las zonas de las ciudades en las que solo está permitido el tránsito de vehículos durante un periodo de tiempo.
- Análisis del tráfico a través de una vía.
- Flujo de vehículos en un determinado transito de tiempo.
- Acceso a vehículos autorizados a un recinto.

4. Estado del arte

Es necesario hacer un repaso del estado actual de las tecnologías de las que se hace uso, así como de los conceptos sobre los que se apoya este proyecto.

Mediante la ejecución de este proyecto fin de grado se ha llevado a cabo un método para paliar algunas carencias de los sistemas actuales de procesamiento de matrículas.

Se expone a continuación una breve descripción sobre el estado del arte actual. Se desarrolla el funcionamiento y características sobre el modelo actual de detección de matrículas denominado ANPR.

4.1 ANPR

El reconocimiento automático de matrículas (Automatic Number Plate Recognition o ANPR en inglés) es un método de vigilancia en masa que utiliza reconocimiento óptico de caracteres en imágenes para leer las matrículas de los vehículos. En 2005, los sistemas pueden escanear las matrículas con una frecuencia aproximada de una por segundo en vehículos con velocidades de hasta 160 km/h. Pueden utilizar el circuito cerrado de televisión existente o radares, o unas diseñadas específicamente para dicha tarea. Son utilizadas por las diversas fuerzas de policía y como método de recaudación electrónica de peaje en las autopistas de pago, y para vigilar la actividad del tránsito como una luz roja en una intersección.

El ANPR se puede utilizar para almacenar las imágenes capturadas por las cámaras fotográficas, así como el texto de la matrícula, y algunas se pueden configurar para almacenar una fotografía del conductor. Estos sistemas a menudo utilizan iluminación infrarroja para hacer posible que la cámara pueda tomar fotografías en cualquier momento del día. En al menos una versión de cámara fotográfica para la supervisión de intersecciones se incluye un flash de gran alcance, que sirve para iluminar la escena y hacer que el infractor se dé cuenta de su error. La tecnología ANPR tiende a ser específica para una región, debido a la variación entre matrículas de un lugar a otro.

La preocupación por estos sistemas se ha centrado en los miedos de privacidad de los movimientos del gobierno de los ciudadanos de rastreo, identificación errónea, las altas tasas de error, y el aumento de gasto público.

ANPR se inventó en 1976 en el Departamento de Desarrollo de la Policía Científica en el Reino Unido. Sistemas prototipos estaban trabajando en 1979, y los contratos se dejaron de producir los sistemas industriales, por primera vez en EMI Electrónica, y luego en sistemas de reconocimiento de Informática en Wokingham, Reino Unido. Los primeros sistemas de prueba fueron desplegados en la carretera A1 y en el túnel de Dartford. La primera detención a través de la detección de un auto robado fue hecha en 1981.

ANPR utiliza el reconocimiento óptico de caracteres en imágenes tomadas por las cámaras. Cuando las placas de matriculación de vehículos holandeses cambiaron a un estilo diferente en el año 2002, uno de los cambios realizados fue a la fuente, la introducción de pequeñas lagunas en algunas cartas para hacerlas más claras y por tanto más legible a dichos sistemas. Algunos acuerdos de placas usan variaciones en tamaños de fuente y los sistemas de posicionamiento ANPR debe ser capaz de hacer frente a esas diferencias para ser realmente eficaz. Los sistemas más complicados pueden hacer frente a las variantes internacionales, aunque muchos programas están diseñados de forma individual para cada país.

Las cámaras utilizadas pueden incluir la aplicación de la regla de carretera existente o cámaras de televisión de circuito cerrado, así como las unidades móviles, que por lo general están unidos a vehículos. Algunos sistemas utilizan cámaras infrarrojas para tener una imagen más clara de las placas.

ANPR es a veces conocido por varios otros términos:

- El reconocimiento automático de matrículas
- Identificación automática de vehículos
- Reconocimiento de matrículas de coches
- Reconocimiento de matrículas
- Conferencia Automatique de placas d'immatriculation

ANPR en sistemas móviles

Los recientes avances en la tecnología han tenido los sistemas de reconocimiento de matrículas de aplicaciones fijas a las móviles. Componentes de escala reducida en los puntos de precio más rentables han llevado a un número récord de despliegues de las fuerzas del orden de todo el mundo. Cámaras más pequeñas con la capacidad de leer las matrículas a altas velocidades, junto con los procesadores más pequeños, más duraderos que encajan en los troncos de los vehículos de la policía, permiten a la policía para patrullar a diario con el beneficio de la matrícula de lectura en tiempo real, cuando pueden interceptar inmediatamente.

A pesar de su eficacia, existen retos significativos relacionados con ANPRs móviles. Uno de los mayores es que el procesador y las cámaras deben trabajar lo suficientemente rápido como para dar cabida a las velocidades relativas de más de 100 kilómetros por hora, un escenario probable en el caso de tráfico. Este equipo también tiene que ser muy eficiente, ya que la fuente de alimentación es la batería del vehículo, y el equipo debe ser pequeño para minimizar el espacio que requiere.

Velocidad relativa es sólo una cuestión que afecta a la capacidad de la cámara para leer realmente una matrícula. Algoritmos deben ser capaces de compensar todas las variables que pueden afectar a la capacidad de la ANPR para producir una lectura precisa, tales como la hora del día, tiempo y ángulos entre las cámaras y las placas de matrícula. Longitudes de onda de iluminación de un sistema también puede tener un impacto directo en la resolución y la precisión de una lectura en estas condiciones.

Instalación de cámaras ANPR en vehículos policiales requiere una cuidadosa consideración de la yuxtaposición de las cámaras a las placas que se van a leer. Usando el número correcto de las cámaras y posicionar con precisión para obtener resultados óptimos puede ser un reto, debido a las diversas misiones y entornos actuales. Patrulla de la carretera requiere cámaras con visión de futuro que se repitan durante varios carriles y son capaces de leer las matrículas a velocidades muy altas. Ciudad patrulla necesita menor alcance, cámaras de distancia focal más bajos para capturar las placas de los coches aparcados. Los estacionamientos con coches aparcados perpendicularmente a menudo requieren una cámara especializada con una distancia focal muy corta. La mayoría de

sistemas avanzados técnicamente son flexibles y se pueden configurar con un número de cámaras que van de uno a cuatro que se puede reposicionar fácilmente según sea necesario. Unidos con placas de sólo trasera tienen un reto adicional ya que una cámara con visión de futuro es ineficaz con el tráfico entrante. En este caso una cámara se puede girar hacia atrás.

Algoritmos.

Hay seis algoritmos principales que el software requiere para la identificación de una placa de matrícula:

- Localización Plate - responsable de encontrar y aislar el plato en la imagen.
- Placa de orientación y el apresto - compensa la inclinación de la placa y ajusta las dimensiones para el tamaño requerido.
- Normalización - ajusta el brillo y el contraste de la imagen.
- Segmentación de caracteres - encuentra los caracteres individuales en las placas.
- Reconocimiento óptico de caracteres.
- Sintáctico/geométrico análisis - verificación caracteres y posiciones en contra de las normas específicas de cada país.

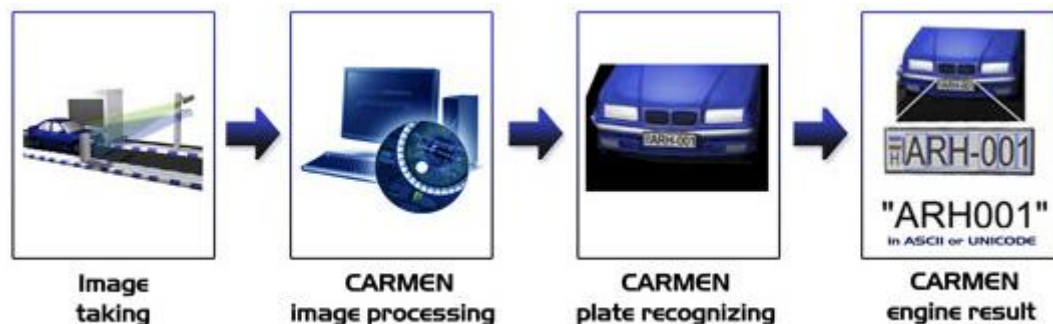


Figura 4.1. Proceso localización ANPR

La complejidad de cada una de estas subsecciones del programa determina la precisión del sistema. Durante la tercera fase, algunos sistemas utilizan técnicas de detección de

bordes para aumentar la diferencia de imagen entre las letras y la placa de respaldo. Un filtro de mediana también se puede utilizar para reducir el ruido visual en la imagen.

Dificultades.

Hay una serie de posibles dificultades que el software debe ser capaz de hacer frente. Estos incluyen:

- Mala resolución de la imagen, por lo general debido a que la placa es demasiado lejos pero a veces como resultado del uso de una cámara de baja calidad.
- Imágenes borrosas, en particular la falta de definición de movimiento.
- La falta de iluminación y de bajo contraste debido a la sobreexposición, reflexión o sombras.
- Un objeto oculta la placa, muy a menudo una barra de remolque, o la suciedad en la placa.
- Una fuente diferente, popular por placas de vanidad.

Técnicas de elusión.

La falta de coordinación entre los países o estados. Dos coches de diferentes países o estados pueden tener el mismo número pero diferente diseño de la placa.

Mientras que algunos de estos problemas pueden ser corregidos dentro del software, se deja principalmente para el lado del hardware del sistema para encontrar soluciones a estas dificultades. El aumento de la altura de la cámara puede evitar problemas con los objetos que oscurecen la placa pero introduce y aumenta otros problemas, tales como el ajuste para el aumento de inclinación de la placa.

En algunos automóviles, barras de remolque pueden oscurecer uno o dos caracteres de la matrícula. Bicicletas en bastidores de bicicletas también pueden ocultar el número de matrícula, aunque en algunos países y jurisdicciones, como Victoria, Australia, "placas de bicicletas" Se supone que para el montaje. Algunos sistemas de pequeña escala permiten algunos errores en la matrícula. Cuando se utiliza para dar a vehículos específicos de acceso a una zona de barricadas, se puede hacer la decisión de tener un índice de error aceptable de un carácter. Esto es debido a que la probabilidad de que un vehículo no autorizado que tiene una placa de licencia tales similar se observa como bastante pequeña.

Sin embargo, este nivel de imprecisión no sería aceptable en la mayoría de las aplicaciones de un sistema ANPR.

Controversia

La introducción de sistemas ANPR ha dado lugar a temores de error en la identificación y el fomento de la vigilancia de 1984 de estilo. En Estados Unidos, algunos, como Gregg Easterbrook se oponen a lo que ellos llaman "máquinas que emiten infracciones y entradas de luz roja" como el comienzo de una pendiente resbaladiza hacia un sistema de justicia automatizado:

"Una máquina clasifica a una persona como a un delincuente, y no se puede enfrentar a su acusador, porque no hay acusador ... ¿puede ser conveniente establecer un principio de que cuando un equipo dice que hiciste algo ilegal, que se presume culpable?"

Críticas similares se han planteado en otros países. Easterbrook también argumenta que esta tecnología se emplea para maximizar los ingresos para el Estado, en lugar de promover la seguridad. El sistema de vigilancia electrónica produce billetes que en los EE.UU. se encuentran a menudo en exceso de \$ 100, y es prácticamente imposible que un ciudadano pueda impugnar ante el tribunal sin la ayuda de un abogado. Los ingresos generados por estas máquinas son compartidos generosamente con la empresa privada que construye y opera ellos, creando un fuerte incentivo para modificar el sistema para generar tantas entradas como sea posible.

Los sistemas más antiguos habían sido notablemente poco fiable, en el Reino Unido, lo que ha sido conocido por llevar a los cargos que se realizan incorrectamente con el propietario del vehículo tiene que pagar 10 para ser emitido con la prueba del delito. Las mejoras en la tecnología han disminuido drásticamente las tasas de error, pero falsas acusaciones son todavía suficientemente frecuentes como para ser un problema.

Tal vez el mejor conocido incidente que implica el abuso de una base de datos ANPR en América del Norte es el caso de Edmonton Sun reportero Diotte Kerry en 2004 - Diotte escribió un artículo crítico con el uso policial Edmonton de cámaras de tráfico para la mejora de los ingresos, y en represalia se añadió a una base de datos de ANPR de "conductores de alto riesgo" en un intento de controlar sus hábitos y crear una oportunidad para arrestarlo. El jefe de la policía y varios oficiales fueron despedidos como resultado, y la Oficina del Comisionado de Privacidad de Canadá expresaron preocupación pública

sobre el "uso de la policía cada vez mayor de la tecnología para espiar a los automovilistas."

Otras preocupaciones incluyen el almacenamiento de información que podría ser utilizada para identificar a las personas y los detalles de sus hábitos de conducción y la vida cotidiana, contraviniendo la Ley de Protección de Datos, junto con una legislación similar. Las leyes en el Reino Unido son estrictas para cualquier sistema que utiliza imágenes de CCTV y se puede identificar a los individuos.

Otro motivo de preocupación es la seguridad de los datos una vez que se extraen, a raíz del descubrimiento de los registros de vigilancia de la policía perdidos en una alcantarilla.

También hay un caso en el Reino Unido por haber dicho que el uso de cámaras ANPR es ilegal en virtud del Reglamento de la Ley de Poderes de Investigación 2000. Existe la infracción, según algunos, en el hecho de que ANPR se utiliza para controlar las actividades de los ciudadanos respetuosos de la ley y trata a todos como a los presuntos criminales destinados a ser objeto de reconocimiento en el marco del acto. La propia policía se ha conocido para referirse al sistema de ANPR como un "movimiento de base de datos 24/7 el tráfico", que es una desviación de la finalidad prevista de los vehículos involucrados en la identificación de las actividades delictivas.

La Associated Press informó en agosto de 2011 que el Departamento de Policía de Nueva York de coches y de placa de equipo de rastreo adquirido con fondos federales HIDTA fueron utilizados para espiar a los musulmanes en las mezquitas, y para rastrear los números de matrícula de los fieles. La policía en vehículos sin identificación equipados con lectores de placas electrónicas podrían bajar a la calle y catalogar automáticamente las placas de cada uno estacionado cerca de la mezquita, acumulando una base de datos secreta que se distribuye entre los oficiales y se utiliza para los musulmanes en el perfil público.

Otros usos

Sistemas ANPR también se pueden usar para/por:

- Sección de control, para medir la velocidad media del vehículo en distancias más largas.

- Pasos fronterizos
- Embargos Automóviles
- Estaciones de servicio para iniciar la sesión cuando un automovilista aleja sin pagar por su combustible.
- Una herramienta de marketing para registrar patrones de uso
- Carteles de publicidad dirigida.
- Sistemas de gestión de tráfico, que determinan el flujo de tráfico utilizando el tiempo que tarda vehículos para pasar dos sitios ANPR
- Los análisis de los hábitos de transporte para planificar el transporte
- Conduzca a través del reconocimiento del cliente, reconocer automáticamente a los clientes en función de su matrícula y ofrecerles los productos que han pedido la última vez que utilizan el servicio, mejorar el servicio al cliente.
- Para ayudar a los sistemas de gestión de visitantes en el reconocimiento de los vehículos de los huéspedes.
- La policía y los auxiliares de policía
- Empresas de parking.
- Hoteles.

Medir el rendimiento del sistema ANPR

Un artículo de 2008 en el estacionamiento Tendencia Internacional discutió una disparidad en las tasas reclamaba contra el reconocimiento de matrículas experimentado leer, con los fabricantes afirman que sus motores de reconocimiento pueden informar correctamente el 98% de las veces, aunque los clientes experimentan sólo el 90% del éxito del 94%, incluso con nuevos equipos en perfectas condiciones. Los primeros sistemas eran los informes, sólo el 60% a 80% fiable. Verdadero tasa de error del sistema es el producto de sus tasas de error del subsistema; ligeros aumentos en las tasas de error del subsistema pueden producir reducciones dramáticas de tasas de lectura. Los efectos de los factores que interfieren en el mundo real de velocidad de lectura no están expresados ni probados por los fabricantes de manera uniforme. El artículo afirma que "hay una necesidad de la industria a adoptar un protocolo de medición de rendimiento estándar para que los clientes potenciales evaluar la mejor opción para sus necesidades particulares.

5. Materiales y métodos

5.1 Procesamiento morfológico.

5.1.1 Morfología

La morfología matemática es una herramienta para extraer componentes de una imagen que sean útiles en la representación y descripción de la forma de una región, tales como contornos, esqueletos y cerco convexo. También son de interés las técnicas de morfológicas para el pre y postprocesado, tales como el filtrado morfológico, la reducción y el recortado.

El lenguaje de la morfología matemática es la teoría de conjuntos. Los conjuntos en la morfología matemática representan las formas de los objetos. Por ejemplo, el conjunto de todos los píxeles negros de una imagen binaria es una descripción completa de ella. En imágenes binarias, los conjuntos son miembros del espacio bidimensional entero Z^2 , donde cada elemento de un conjunto es vector bidimensional cuyas coordenadas (x,y) son las de un píxel negro (por convención) de una imagen. Las imágenes digitales en escala de grises se pueden representar como conjuntos cuyos componentes están en Z^3 . En este caso, dos componentes hacen referencia a la posición (x,y) y la tercera componente se corresponde con el valor de intensidad discreta del píxel. Los conjuntos de espacios de dimensiones mayores pueden contener otros atributos de la imagen, tales como color y componentes variables con el tiempo.

A continuación se explicarán los diferentes tipos de operaciones morfológicas utilizadas para el desarrollo de la aplicación.

5.1.2 Dilatación.

5.1.2.1 Definición.

La dilatación es la transformación morfológica que combina dos vectores utilizando la suma. La dilatación binaria fue usada por Minkowski (en la literatura recibe el nombre de suma de Minkowski), Para caracterizar la medida integral de ciertos conjuntos abiertos

(dispersos). La operación de dilatación para el procesamiento de imágenes fue empleada como una operación de suavizamiento por varios investigadores contemporáneos en el procesamiento de imágenes. Matheron y Serra exploraron a la dilatación como un operador de imágenes usado para la extracción de la forma y la estimación de los parámetros de la imagen. Todas estas aplicaciones iniciales trataron con imágenes binarias únicamente.

5.1.2.2 Conceptos matemáticos.

Sean A y B conjuntos en un espacio N -dimensional, con elementos a y b respectivamente, $a = (a_1 \dots a_N)$ y $b = (b_1 \dots b_N)$, siendo vectores que indican las coordenadas de los elementos, luego la dilatación de A por B es el conjunto de todos los posibles vectores suma de pares de elementos, uno proveniente de A y otro proveniente de B .

La dilatación de la imagen binaria A por el elemento estructurante B (para los conjuntos en Z^2 y $\emptyset$ representando al conjunto vacío) está definida por:

$$A \oplus B = \{x | (B)_x \cap A \neq \emptyset\} \quad (5.1)$$

Por tanto el proceso de dilatación consiste en obtener la reflexión de B sobre su origen y después cambiar esta reflexión por x . La dilatación de A por B es entonces el conjunto de todos los desplazamientos x tales que B y A se solapen en al menos un elemento distinto de cero. Basándose en esta interpretación, se puede definir la dilatación como:

$$A \oplus B = \{x | [(B)_x \cap A] \subseteq A\} \quad (5.2)$$

Al elemento B se le conoce como elemento de estructura de la dilatación, al igual que en otras operaciones morfológicas.

La ecuación anterior no es la única que define la dilatación, sin embargo, tiene la ventaja distintiva sobre otras formulaciones de ser más intuitiva cuando se toma el elemento de estructura B como una máscara de convolución. Aunque la dilatación se basa en operaciones de conjuntos y la convolución en operaciones aritméticas, el proceso básico de ‘mover’ el elemento estructurante B respecto a su origen y desplazarlo sucesivamente

después de tal forma que se deslice sobre el conjunto (imagen) A es análogo al proceso de convolución.

La dilatación por elementos estructurales de disco, corresponde a un crecimiento isotrópico o a un algoritmo de expansión común al procesamiento de imágenes binarias. La dilatación por pequeños cuadrados (3 X 3) es una operación de vecinos, fácilmente implementada por una arquitectura de arreglos conectados adyacentemente (grillas), este procesamiento de imágenes es conocido comúnmente como "relleno", "expansión" o "crecimiento".

- Ejemplo de dilatación de objeto:

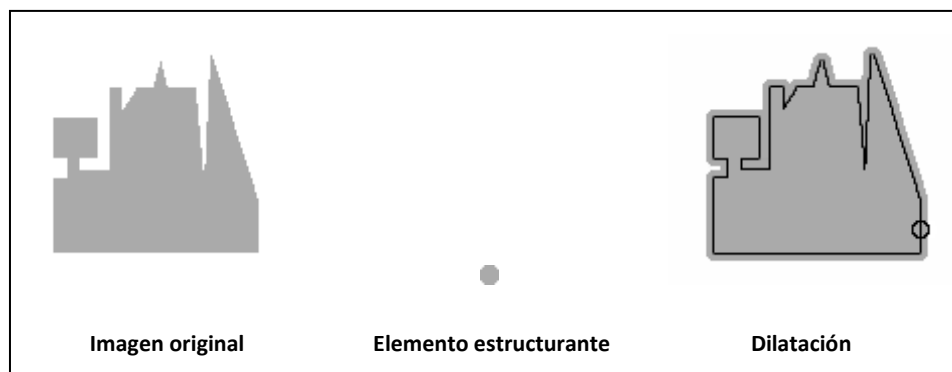


Figura 5.1. Dilatación 1

- Ejemplo de dilatación observando los píxeles:

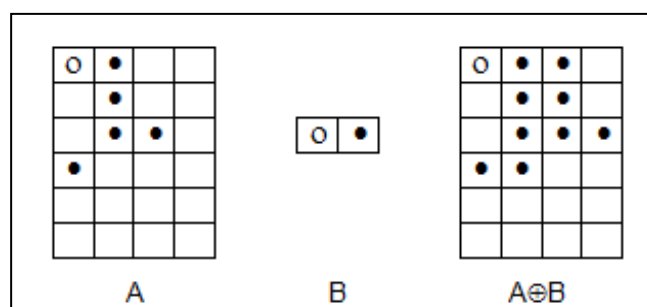


Figura 5.2 Dilatación. Original, elemento estructurante y resultado. Fuente:

<http://www.monografias.com/trabajos94/proceso-segmentacion-automatica-lesiones-patologicas-imagenes-mamografias/proceso-segmentacion-automatica-lesiones-patologicas-imagenes-mamografias.shtml>

5.1.2.3 Propiedades de la dilatación.

- 1) Se puede calcular la dilatación como la unión de traslaciones de A por los elementos de B. Se cumple que: $A \oplus B = \bigcup_{b \in B} (A)_b$
- 2) Conmutativa: En la dilatación morfológica, la función de los conjuntos A y B es similar, es decir que la operación de dilatación es conmutativa porque la suma es conmutativa. Dicha propiedad puede ser expresada de la siguiente forma:

$$A \oplus B = B \oplus A \quad (5.3)$$

- 3) Invarianza a la translación: La dilatación por el trasladado de un elemento estructural es el trasladado de la dilatación:

$$A \oplus B_t = (A \oplus B)_t \quad (5.4)$$

Corolario: Se aplica a dilataciones implementadas por la regla de la cadena. Este corolario establece que la traslación de uno de los elementos estructurantes en la descomposición de la dilatación traslada la imagen dilatada en la misma forma.

$$A \oplus B_1 \oplus \dots \oplus (B_n)_x \oplus \dots \oplus B_N = (A \oplus B_1 \oplus \dots \oplus B_n \oplus \dots \oplus B_N)_x \quad (5.5)$$

- 4) Distributiva: Respecto a la unión. El orden de la unión de imágenes y la dilatación pueden ser intercambiados. La dilatación de la unión de dos imágenes es igual a la unión de las dilataciones de esas imágenes. Esto se representa así:

$$A \oplus (B \cup C) = (A \oplus B) \cup (A \oplus C) \quad (5.6)$$

- 5) Asociativa: Los procesadores de imágenes por vecinos interconectados pueden implementar algunas dilataciones por elementos estructurales más grandes que el tamaño de los vecinos, realizando dilataciones iterativas con una secuencia de

elementos estructurales vecinos. Si una imagen A va a dilatarse por un elemento estructurante D , el cual a su vez puede ser expresado como la dilatación de B por C , entonces $A \oplus D$ puede calcularse como:

$$A \oplus D = A \oplus (B \oplus C) = (A \oplus B) \oplus C \quad (5.7)$$

La forma $(A \oplus B) \oplus C$, representa un considerable ahorro en cuanto al número de operaciones a ser efectuadas cuando A es la imagen y $B \oplus C$ es el elemento estructural. Este ahorro se produce debido a que una dilatación bruta forzada de $B \oplus C$ puede tomar unas N^2 operaciones, mientras que si primero se dilata A por B y luego se dilata el resultado por C , esto tornaría apenas $2N$ operaciones, donde N es el número de elementos en B y en C .

La propiedad asociativa es comúnmente conocida como la “regla de la cadena” para dilataciones. Es importante señalar que las transformaciones de dilatación que pueden ser realizadas como una cadena de dilataciones no es posible hacerlas como una cadena de operaciones entre vecinos.

- 6) Creciente: La dilatación es creciente, quiere decir que las relaciones de continencia se mantienen a través de la dilatación :

$$\text{si } A \subseteq C \text{ entonces } A \oplus B \subseteq C \oplus B \quad (5.8)$$

5.1.3 Erosión

5.1.3.1 Definición.

La erosión, junto con la dilatación, son las bases de la mayoría de operaciones morfológicas. Es la operación morfológica dual de la dilatación.

La erosión es la transformación morfológica que combina dos conjuntos usando el concepto de inclusión, usando el vector resta de los elementos del conjunto.

Se concibe normalmente como una reducción de la imagen original.

5.1.3.2 Conceptos matemáticos.

La erosión de la imagen binaria A por el elemento estructurante B (para los conjuntos en Z^2) está definida por:

$$A \ominus B = \{x | (B)_x \subseteq A\}, \text{ donde } B_x \text{ es la traslación de } B \text{ por el vector } x, \text{ esto es, } B_x = \{b+x | b \in B\}, \quad \forall x \in E.$$

Tengamos en cuenta que, para la condición $B_x \subseteq A$, sólo consideramos los píxeles negros de A y B .

Esta igualdad dice que la erosión de A por B es el conjunto de todos los puntos x tales que B , trasladado por x , está contenido en A .

- Ejemplo de erosión de una imagen con un círculo como elemento estructurante:

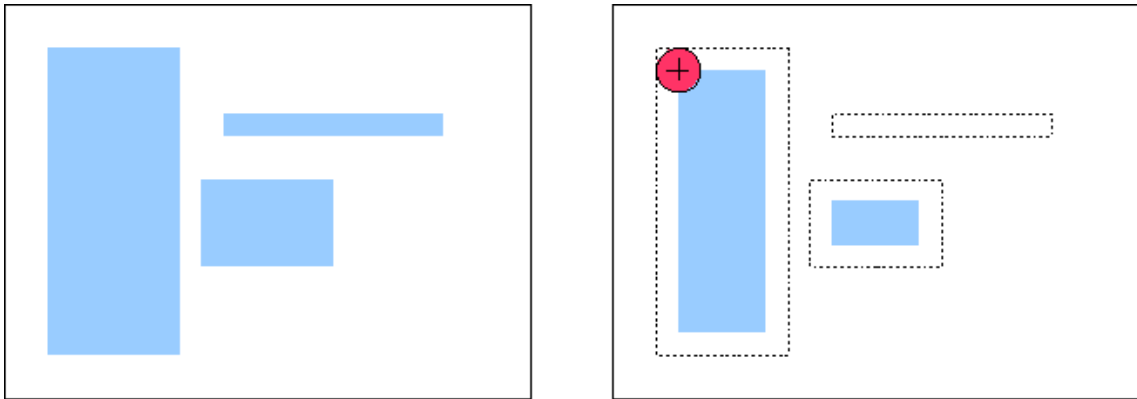


Figura 5.3 Erosión. Original y resultado con elemento estructurante circular. Fuente:

[http://es.wikipedia.org/wiki/Erosi3n_\(morfolog3a\)](http://es.wikipedia.org/wiki/Erosi3n_(morfolog3a))

- Ejemplo de erosión observando los píxeles:

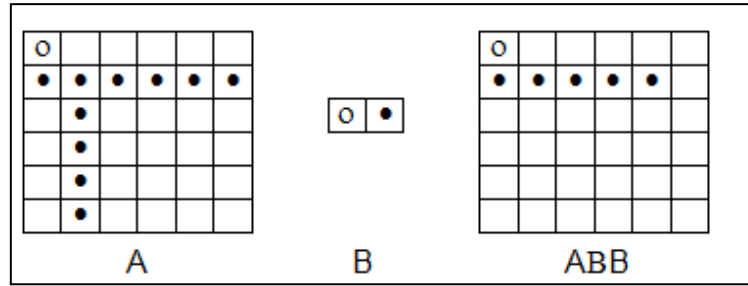


Figura 5.4 Erosión. Original, elemento estructurante y resultado. Fuente:

<http://www.monografias.com/trabajos94/proceso-segmentacion-automatica-lesiones-patologicas-imagenes-mamografias/proceso-segmentacion-automatica-lesiones-patologicas-imagenes-mamografias.shtml>

Teorema de dualidad entre la erosión y la dilatación.- Es posible afirmar la existencia de una operación dual a la dilatación, por el mismo hecho de ser esta una transformación morfológica. La erosión es esta operación dual a la dilatación, teniéndose la siguiente relación:

$$(A \ominus B)^C = A^C \oplus B \quad (5.9)$$

Demostración:

$x \in (A \ominus B)^C$ si y solo si $x \in A \ominus B$.

$x \in A \ominus B$ si y solo si existe un $b \in B$ tal que $x + b \in A$.

Existe un $b \in B$ tal que $x + b \in A^C$ si y solo si existe un $b \in B$ tal que $x \in (A^C)_{-b}$.

Existe un $b \in B$ tal que $x \in (A^C)_{-b}$, si y solo si $x \in \bigcup_{b \in B} (A^C)_{-b}$.

Ahora, $x \in \bigcup_{b \in B} (A^C)_{-b}$ si y solo si $x \in \bigcup_{b \in B} (A^C)_{-b}$.

Y, finalmente $x \in \bigcup_{b \in B} (A^C)_{-b}$ si y solo si $x \in A^C \oplus B$.

La dualidad de la dilatación y la erosión emplea tanto negación lógica como geométrica debido a los roles diferentes que tienen la imagen y el elemento estructural en una expresión empleando estos operadores morfológicos.

Se puede notar de lo anterior, la dualidad entre la suma y la resta de Minkowski:

$$A \oplus B = (A^C \ominus B)^C \quad (5.10)$$

$$A \ominus B = (A^C \oplus B)^C \quad (5.11)$$

5.1.3.3 Propiedades de la erosión:

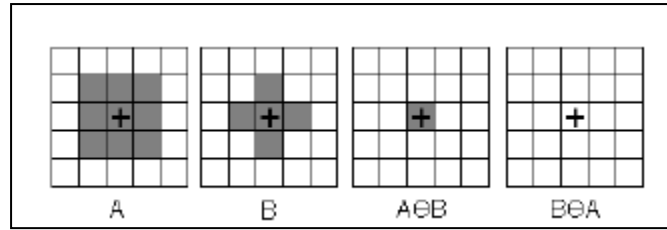


Figura 5.5 Erosión. Fuente: <http://alojamientos.us.es/gtocom/pid/tema5-1.pdf>

1) La erosión no es conmutativa.

2) Propiedad distributiva:

2.1) *Con intersección de imágenes.-* Se observa que la erosión difiere sensiblemente de la dilatación, esto se realza analizando las propiedades algebraicas de ambas operaciones, especialmente las relacionadas con operaciones sencillas de conjuntos como la intersección o unión. La erosión de la intersección de dos imágenes es igual a la intersección de sus erosiones. Sean A y B las imágenes y C el conjunto simétrico de un elemento estructural arbitrario, así:

$$(A \cap B) \ominus C = (A \ominus C) \cap (B \ominus C) \quad (5.12)$$

Esta relación suele tomarse como base para probar el conocimiento local de la transformación de dilatación y se considera como su expresión correspondiente para la erosión.

2.2) *Con unión de imágenes.*- A diferencia de la dilatación, la erosión no es conmutativa, por lo tanto, la erosión de una unión de imágenes por un elemento estructural, es diferente a la erosión de una imagen por un elemento estructural compuesto por la unión de dos elementos estructurales, esto, considerando en ambos casos los mismos conjuntos de referencia que se han estado utilizando, esto es: A, B, C. Consecuentemente, la erosión de una imagen por la unión de dos elementos estructurales es igual a la intersección de las erosiones de la imagen por cada uno de los elementos estructurales, así:

$$A \ominus (B \cup C) = (A \ominus B) \cap (A \ominus C) \quad (5.13)$$

La utilidad práctica de esta propiedad radica en que muestra la forma de calcular la erosión con elementos estructurales que solo pueden ser descompuestos como la unión de elementos estructurales individuales.

- 3) La erosión es creciente: Si la imagen A está contenida en la imagen B, entonces la erosión de A estará contenida en la erosión de B.

$$\text{Si } A \subseteq B \text{ entonces } A \ominus C \subseteq B \ominus C \quad (5.14)$$

- 4) Además la erosión por un elemento estructural mayor produce un resultado menor:

$$\text{Si } K \subseteq L, \text{ entonces } A \ominus L \subseteq A \ominus K \quad (5.15)$$

- 5) Distributiva entre la erosión y dilatación (iteración): Con respecto a la descomposición de elementos estructurales, una regla de la cadena para la erosión se verifica cuando el elemento estructural se puede descomponer mediante dilatación. Esta relación es tan importante como la regla de cadena de la dilatación ya que permite calcular una gran erosión como dos pequeñas erosiones sucesivas. Considerando la imagen A y los elementos estructurales B y C, se tiene lo siguiente:

$$A \ominus (B \oplus C) = (A \ominus B) \ominus C \quad (5.16)$$

- 6) Antiextensividad: La transformación de erosión es concebida popularmente como una contracción de la imagen original. El conjunto erosionado se ve como un conjunto contenido en la imagen original (antiextensividad). Sin embargo, la transformación de erosión será necesariamente antiextensiva solo si el origen pertenece al elemento estructural, esto es:

$$\text{Si } 0 \in B, \text{ entonces } A \ominus B \subseteq A \quad (5.17)$$

5.1.4 Apertura

Como se ha visto la dilatación expande la imagen y la erosión la contrae. Los dos siguientes puntos desarrollan dos importantes procesos morfológicos: apertura y cierre.

En la práctica, la erosión y dilatación se utilizan en pares, la erosión seguida de dilatación o viceversa. En ambos casos, el resultado de aplicarlos iterativamente es la eliminación de detalle específico de la imagen menor al tamaño del elemento estructurante, sin una distorsión geométrica global de las características no suprimidas.

Efectuar una apertura (erosión-dilatación) por un elemento estructurante generalmente suaviza el contorno, rompe istmos estrechos, elimina pequeñas islas y agudiza picos.

La apertura de un conjunto A por un elemento de estructura B, representada por $A \circ B$, se define como:

$$A \circ B = (A \ominus B) \oplus B \quad (5.18)$$

que nos dice que la apertura de A por B es simplemente la erosión de A por B, seguida por una dilatación del resultado de B.

- Propiedades:

- 1) Antiextensiva: Que la transformación de apertura sea antiextensiva, quiere decir que la apertura de A por un elemento estructural B está necesariamente contenida en A sin importar si el origen pertenece o no a B. La antiextensividad de la apertura se expresa de la siguiente manera:

$$A \circ B \subset A \quad (5.19)$$

- 2) Idempotencia: Esta propiedad expresa que si se vuelve a aplicar una apertura a una imagen previamente abierta (con el mismo elemento estructural), la imagen abierta no mostrará cambio alguno. Considerando la imagen A y el elemento estructural B, se tiene que:

$$A \circ B = (A \circ B) \circ B. \quad (5.20)$$

- 3) Creciente: Esta propiedad se obtiene directamente del hecho de que la dilatación es creciente. Si se consideran dos imágenes X_1 y X_2 y un elemento estructural B, se puede escribir esta propiedad así:

$$\text{Si } X_1 \subseteq X_2 \rightarrow (X_1 \circ B) \subseteq (X_2 \circ B) \quad (5.21)$$

- Ejemplo de apertura:

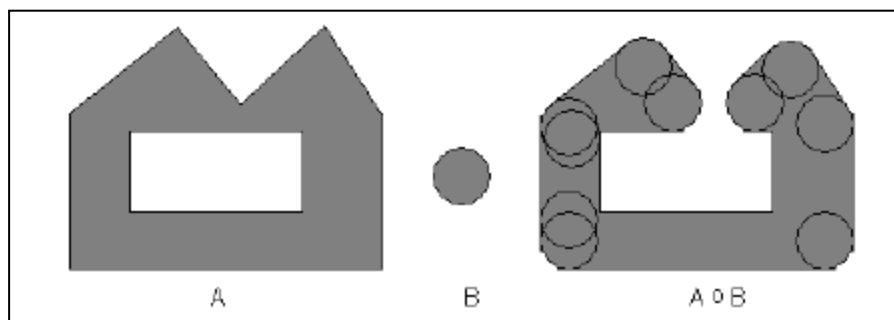


Figura 5.6 Apertura. Original, elemento estructurante y resultado. Fuente:

<http://alojamientos.us.es/gtocom/pid/tema5-1.pdf>

En este caso el elemento estructurante tiene forma de disco. Se puede apreciar como al realizar la apertura se suavizan los contornos, se rompen uniones estrechas entre pastes de conjuntos y se eliminan salientes estrechos.

- En el siguiente ejemplo podemos apreciar la apertura observando los pixeles del saliente que son eliminados.

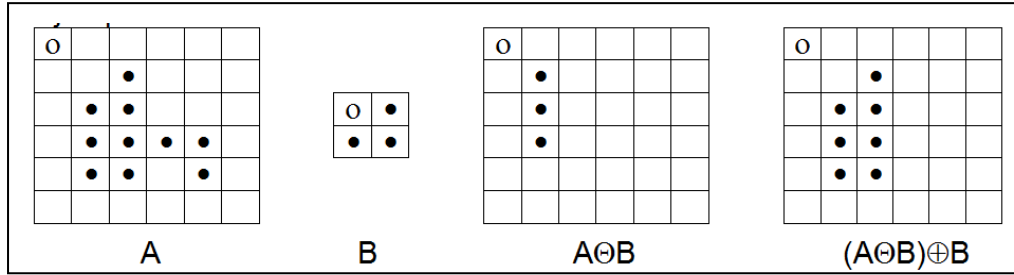


Figura 5.7 Apertura. Original, elemento estructurante, erosión y dilatación. Fuente:

<http://www.monografias.com/trabajos94/proceso-segmentacion-automatica-lesiones-patologicas-imagenes-mamografias/proceso-segmentacion-automatica-lesiones-patologicas-imagenes-mamografias.shtml>

5.1.5 Cierre

El cierre (dilatación-erosión) es otra operación morfológica importante, ya que tiende a suavizar secciones de contornos pero, en oposición a la apertura, generalmente fusiona separaciones estrechas y entrantes delgados y profundos, elimina pequeños huecos y rellena agujeros del contorno.

El cierre del conjunto de A por el elemento de estructura B, representado por $A \cdot B$, se define como:

$$A \cdot B = (A \oplus B) \ominus B \quad (5.22)$$

que nos dice que el cierre de A por B es simplemente la dilatación de A por B, seguida por la erosión del resultado por B.

Se cumple la dualidad entre apertura y cierre, es decir,

$$(A \cdot B)^C = A^C \circ B \quad (5.23)$$

- Propiedades

- 1) Extensividad: Que la transformación de cierre sea extensiva, quiere decir que el cierre de A por un elemento estructural B contiene a A sin importar si B contiene o no al origen. La antiextensividad de la apertura se expresa así:

$$A \subseteq A \cdot B \quad (5.24)$$

- 2) Idempotencia: De manera similar a lo que ocurre con la apertura, si se vuelve a aplicar un cierre a una imagen previamente cerrada (con el mismo elemento estructural), la imagen cerrada no mostrará cambio alguno. Considerando la imagen A y el elemento estructural B, se tiene que:

$$A \cdot B = (A \cdot B) \cdot B \quad (5.25)$$

- 3) Creciente: Esta propiedad se obtiene directamente del hecho de que la erosión es creciente. Si se consideran dos imágenes X_1 y X_2 y un elemento estructural B, se puede escribir esta propiedad así:

$$\text{Si } X_1 \subseteq X_2 \rightarrow (X_1 \cdot B) \subseteq (X_2 \cdot B) \quad (5.26)$$

- Ejemplo de cierre:

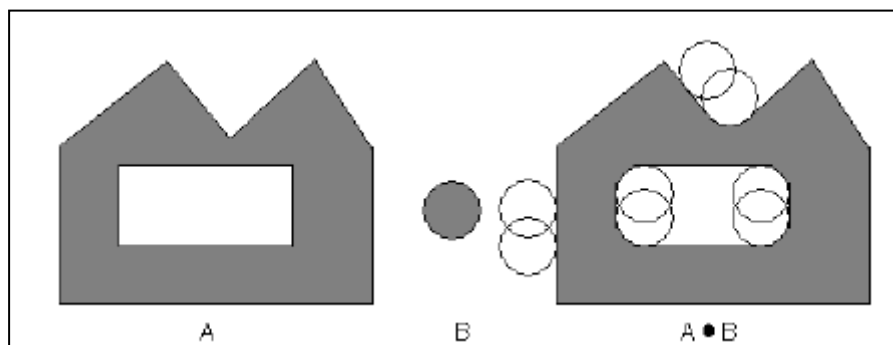


Figura 5.8 Cierre. Original, elemento estructurante y resultado. Fuente:

<http://alojamientos.us.es/gtocom/pid/tema5-1.pdf>

Si tomamos un disco como elemento estructural, el cierre tiende a suavizar las secciones de contornos pero en sentido inverso: una separaciones estrechas, elimina golfos estrechos y huecos.

- Ejemplo de cierre donde podemos observar el proceso en píxeles:

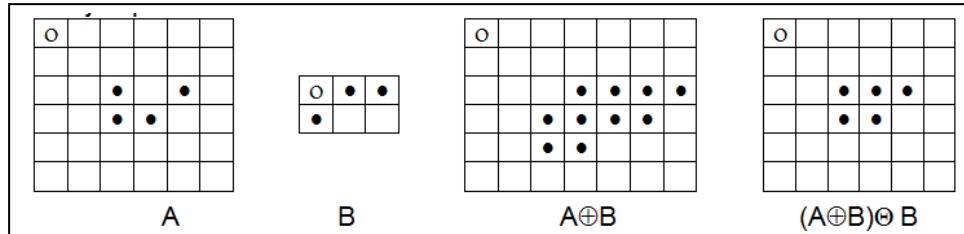


Figura 5.9 Cierre. Original, elemento estructurante, dilatación y erosión. Fuente:

<http://www.monografias.com/trabajos94/proceso-segmentacion-automatica-lesiones-patologicas-imagenes-mamografias/proceso-segmentacion-automatica-lesiones-patologicas-imagenes-mamografias.shtml>

5.1.6 Transformaciones Top-Hat y Bottom-Hat

Las transformaciones Top-hat consisten en descubrir aquellas estructuras de la imagen que han sido eliminadas en el filtrado de apertura o cierre. Una operación entre la imagen original y el filtrado aumenta considerablemente el contraste de las zonas eliminadas. Esta operación se define como un residuo entre la identidad y la apertura o bien entre el cierre y la identidad

Se definen a partir de apertura y cierre utilizando substracciones de imágenes:

$$T_{hat} = A - (A \circ B) \quad (5.27)$$

$$B_{hat} = (A \cdot B) - A \quad (5.28)$$

Donde A es la imagen y B el elemento estructural.

La transformación Top-Hat es útil para resaltar detalles claros en la presencia de sombras (pequeños detalles brillantes). También sirve para corregir efectos de iluminación no uniforme en la imagen.

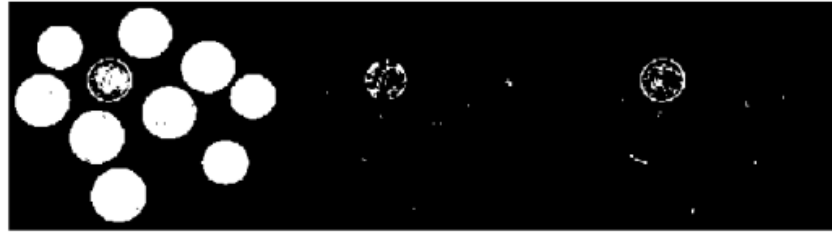


Figura 5.10 a) Imagen b)Top-hat: cierre e imagen c)Top-hat: imagen y apertura

Si se quieren obtener pequeños objetos oscuros sobre un fondo blanco (claro) se hará uso de la transformación Bottom-Hat.

5.1.7 Elementos estructurantes.

El objetivo de las transformaciones morfológicas es la extracción de estructuras geométricas en los conjuntos sobre los que se opera, mediante la utilización de otro conjunto de forma conocida, al que se le denomina elemento estructurante. El tamaño y forma del elemento estructurante se elige, a priori, de acuerdo con la morfología sobre la que va a interseccionar y en función de la obtención de formas que se desea extraer.

En la figura adjunta aparecen algunos tipos de elementos estructurantes empleados en el procesamiento morfológico.



Figura 5.11 Elementos estructurantes típicos

5.2 Filtros detección de bordes

5.2.1 Introducción.

La detección de bordes es una herramienta fundamental en el procesamiento de imágenes y en visión por computadora, particularmente en las áreas de detección y

extracción de características, que tiene como objetivo la identificación de puntos en una imagen digital en la que el brillo de la imagen cambia drásticamente o, más formalmente, tiene discontinuidades. Es un proceso en el análisis digital de imágenes que detecta los cambios en la intensidad de luz. Estos cambios se pueden usar para determinar la profundidad, tamaño, orientación y propiedades de la superficie dentro de una muestra o pieza de trabajo.

La mayoría de las técnicas para detectar bordes emplean operadores locales basados en distintas aproximaciones discretas de la primera y segunda derivada de los niveles de grises de la imagen.

5.2.2 Operadores basados en la primera derivada (Gradiente).

La derivada de una señal continua proporciona las variaciones locales con respecto a la variable, de forma que el valor de la derivada es mayor cuanto más rápidas son estas variaciones.

En el caso de funciones bidimensionales $f(x,y)$, la derivada es un vector que apunta en la dirección de la máxima variación de $f(x,y)$ y cuyo módulo es proporcional a dicha variación.

El gradiente de una función continua $f(x,y)$ es un vector y su magnitud expresa el ratio de cambio por unidad de distancia en la dirección del vector. El objetivo de los operadores gradiente es detectar cambios en los niveles de gris que tienen lugar en zonas o regiones reducidas. El gradiente y la magnitud se expresan por:

$$\frac{\partial f(x,y)}{\partial x} = \Delta x = \frac{f(x+d_x,y)-f(x,y)}{d_x} \quad (5.29)$$

$$\frac{\partial f(x,y)}{\partial y} = \Delta y = \frac{f(x,d_y+y)-f(x,y)}{d_y} \quad (5.30)$$

dx y dy son las distancias en las direcciones x e y respectivamente, en términos de número de pixels entre dos puntos.

A la magnitud del gradiente se la conoce como fuerza del borde y junto con la orientación del gradiente constituyen datos cruciales en la extracción de este tipo de características.

En orden de detectar la presencia de una discontinuidad en el gradiente, debemos calcular el cambio en el gradiente en el punto (i,j). Esto se puede hacer referenciando la medida aportada por la magnitud del gradiente y su dirección.

$$M = \sqrt{\Delta^2_x + \Delta^2_y} \quad (5.31)$$

$$\theta = \tan^{-1}\left(\frac{\Delta_y}{\Delta_x}\right) \quad (5.32)$$

En el caso bidimensional discreto, las distintas aproximaciones del operador gradiente se basan en diferencias entre los niveles de grises de la imagen. La derivada parcial $f_x(x,y)$ (gradiente de fila $G_F(i,j)$) puede aproximarse por la diferencia de píxeles adyacentes de la misma fila.

En orden de detectar la presencia de una discontinuidad en el gradiente, debemos calcular el cambio en el gradiente en el punto (i,j). Esto se puede hacer referenciando la medida aportada por la magnitud del gradiente y su dirección.

$$\Delta_x = f(i + 1, j) - f(i, j) \quad (5.33)$$

$$\Delta_y = f(i, j + 1) - f(i, j) \quad (5.34)$$

Para la implementación y computación del gradiente se utilizan máscaras o filtros que representan o equivalen a dichas ecuaciones.

Métodos basados en la primera derivada: Operador Gradiente. Constituyen los métodos con más proliferación dentro de la comunidad del análisis de imagen y la visión computacional. Se fundamentan en que un borde existe si hay una discontinuidad en la función de intensidad de la imagen, es decir, si la derivada de los valores de intensidad de

la imagen es un máximo.

5.2.2.1 Operador Robert

Las máscaras utilizadas en este operador son:

Tabla 5.1 Gradiente fila

0	0	0
0	0	1
0	-1	0

Tabla 5.2 Gradiente columna

-1	0	0
0	1	0
0	0	0

Obtiene buena respuesta ante bordes diagonales. Ofrece buenas prestaciones en cuanto a localización. El gran inconveniente de este operador es su extremada sensibilidad al ruido y por tanto tiene pobres cualidades de detección.



Figura 5.12 Ejemplo de uso operador Robert

5.2.2.2 Operador Prewitt

El operador de Prewitt, técnicamente, es un operador de diferenciación discreta, calcular una aproximación de la función gradiente de intensidad de la imagen. En cada punto de la imagen, el resultado del operador Prewitt es o bien el vector de gradiente correspondiente o la norma de este vector. El operador de Prewitt se basa en la convolución de la imagen con un filtro pequeño, separable, y valor entero en direcciones horizontales y verticales y por lo tanto es relativamente económico en términos de cálculos. Por otra parte, la aproximación del gradiente que se produce es relativamente crudo, en particular, para las variaciones de altas frecuencias en la imagen.

Las máscaras utilizadas en este operador son:

Tabla 5.3 Gradiente fila

1	0	-1
1	0	-1
1	-1	-1

Tabla 5.4 Gradiente columna

-1	-1	-1
0	1	0
1	1	1



Figura 5.13. Ejemplo de uso operador Prewitt

5.2.2.3 Operador Sobel

El operador Sobel calcula el gradiente de la intensidad de una imagen en cada punto (píxel). Así, para cada punto, este operador da la magnitud del mayor cambio posible, la dirección de éste y el sentido desde oscuro a claro. El resultado muestra cómo de abruptamente o suavemente cambia una imagen en cada punto analizado y, en consecuencia, cuán probable es que éste represente un borde en la imagen y, también, la orientación a la que tiende ese borde. En la práctica, el cálculo de la magnitud - probabilidad de un borde- es más fiable y sencillo de interpretar que el cálculo de la dirección y sentido.

Matemáticamente, el gradiente de una función de dos variables (en este caso, la función de intensidad de la imagen) para cada punto es un vector bidimensional cuyos componentes están dados por las primeras derivadas de las direcciones verticales y horizontales. Para cada punto de la imagen, el vector gradiente apunta en dirección del incremento máximo posible de la intensidad, y la magnitud del vector gradiente corresponde a la cantidad de cambio de la intensidad en esa dirección.

Las máscaras utilizadas en este operador son:

Tabla 5.5 Gradiente fila

1	0	-1
2	0	-2
1	-1	-1

Tabla 5.6 Gradiente columna

-1	-2	-1
0	1	0
1	2	1



Figura 5.14. Ejemplo de uso operador Sobel

5.2.2.4 Operador Canny.

Para que un detector de bordes pueda ser considerado óptimo debe cumplir los siguientes puntos:

- Buena detección- el algoritmo debe marcar el mayor número real en los bordes de la imagen como sea posible.
- Buena localización- los bordes de marca deben estar lo más cerca posible del borde de la imagen real.
- Respuesta mínima - El borde de una imagen sólo debe ser marcado una vez, y siempre que sea posible, el ruido de la imagen no debe crear falsos bordes.

Para satisfacer estos requisitos Canny utiliza el cálculo de variaciones - una técnica que encuentra la función que optimiza un funcional indicado. La función óptima en el algoritmo de Canny es descrito por la suma de cuatro términos exponenciales, pero se puede aproximar por la primera derivada de una gaussiana.

El método de Canny se basa en tres criterios principales:

- El criterio de detección, que expresa el hecho de evitar la eliminación de bordes importantes así como no suministrar falsos bordes.

- El criterio de localización, que establece que la distancia entre la posición real y la posición localizada para el borde debe ser minimizada.
- El criterio de respuesta única, que establece la necesidad de que el detector retorne un único punto por cada punto de borde verdadero. Esto implica que el detector no debe encontrar múltiples píxeles de borde donde solo existe uno.



Figura 5.15 Ejemplo de uso operador Canny

5.2.2.5 Comparativa

Operadores Gradiente	Ventajas	Desventajas
Roberts	- Buena respuesta en bordes horizontales y verticales. - Buena localización. - Simpleza y rapidez de cálculo.	- Mala respuesta en bordes diagonales. - Sensible al ruido. - Empleo de máscaras pequeñas. - No da información acerca de la orientación del borde. - Anchura del borde de varios píxeles
Sobel	- Buena respuesta en bordes horizontales y verticales. - Diversidad de tamaños en las máscaras. - Poco sensible al ruido.	- Mala respuesta en bordes diagonales. - Lentitud de cálculo. - No da información acerca de la orientación del borde. - Anchura del borde de varios píxeles
Prewitt Prewitt	- Buena respuesta en bordes horizontales y verticales. - Poco sensible al ruido. - Proporciona la magnitud y dirección del borde.	- Mala respuesta en bordes diagonales. - Lentitud de cálculo. - Anchura del borde de varios píxeles.
Canny	- Uso de probabilidad para encontrar la tasa de error. - La localización y la respuesta. - Mejorar la relación señal a ruido. - Mejor detección especialmente en condiciones de ruido	- Los cálculos complejos. - Falso zero crossing. - Consume tiempo.

5.3 Conceptos matemáticos

5.3.1 Correlación de Pearson

La correlación entre dos variables refleja el grado en que las puntuaciones están asociadas. La formulación clásica, conocida como correlación producto momento de Pearson, se simboliza por la letra griega rho cuando ha sido calculada en la población. Si se obtiene sobre una muestra, se designa por la letra " r_{xy} ". Este tipo de estadístico puede utilizarse para medir el grado de relación de dos variables si ambas utilizan una escala de medida a nivel de intervalo/razón (variables cuantitativas).

El coeficiente de correlación de Pearson, pensado para variables cuantitativas (escala mínima de intervalo), es un índice que mide el grado de covariación entre distintas variables relacionadas linealmente. Adviértase que decimos "variables relacionadas linealmente". Esto significa que puede haber variables fuertemente relacionadas, pero no de forma lineal, en cuyo caso no proceder a aplicarse la correlación de Pearson.

El coeficiente de correlación de Pearson es un índice de fácil ejecución e, igualmente, de fácil interpretación. Digamos, en primera instancia, que sus valores absolutos oscilan entre 0 y 1. Esto es, si tenemos dos variables X e Y, y definimos el coeficiente de correlación de Pearson entre estas dos variables como r_{xy} entonces:

$$0 \leq r_{xy} \leq 1 \quad (5.35)$$

Se ha especificado los términos "valores absolutos" ya que en realidad si se contempla el signo el coeficiente de correlación de Pearson oscila entre -1 y $+1$. No obstante ha de indicarse que la magnitud de la relación viene especificada por el valor numérico del coeficiente, reflejando el signo la dirección de tal valor. En este sentido, tan fuerte es una relación de $+1$ como de -1 . En el primer caso la relación es perfecta positiva y en el segundo perfecta negativa.

Una correlación de $+1$ significa que existe una relación lineal directa perfecta (positiva) entre las dos variables. Es decir, las puntuaciones bajas de la primera variable

(X) se asocian con las puntuaciones bajas de la segunda variable (Y), mientras las puntuaciones altas de X se asocian con los valores altos de la variable Y.

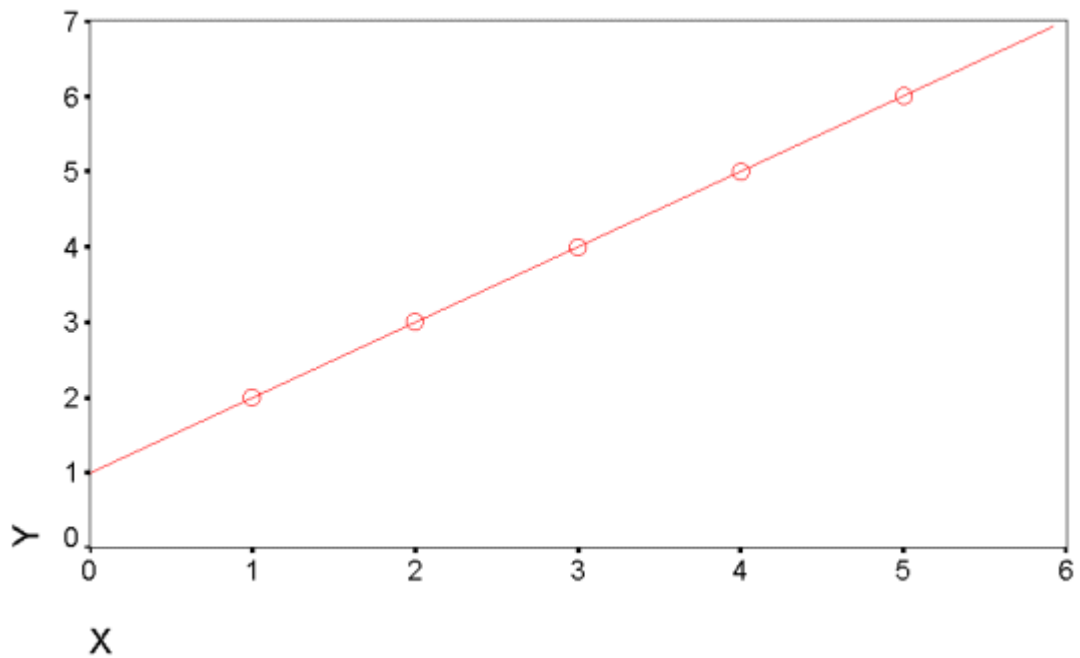


Figura 5.16. Correlación valor 1

Una correlación de -1 significa que existe una relación lineal inversa perfecta (negativa) entre las dos variables. Lo que significa que las puntuaciones bajas en X se asocian con los valores altos en Y, mientras las puntuaciones altas en X se asocian con los valores bajos en Y. Igual que en el caso anterior esto sucede para relaciones funcionales exactas, propio de las ciencias físicas.

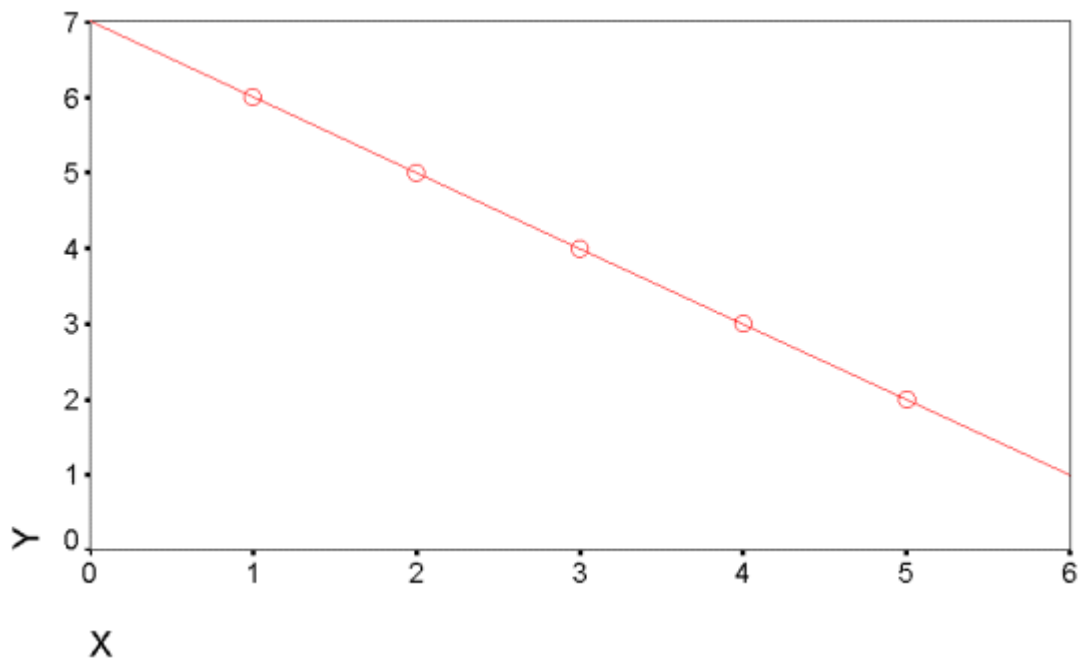


Figura 5.17 Correlación valor -1

Una correlación de 0 se interpreta como la no existencia de una relación lineal entre las dos variables estudiadas.

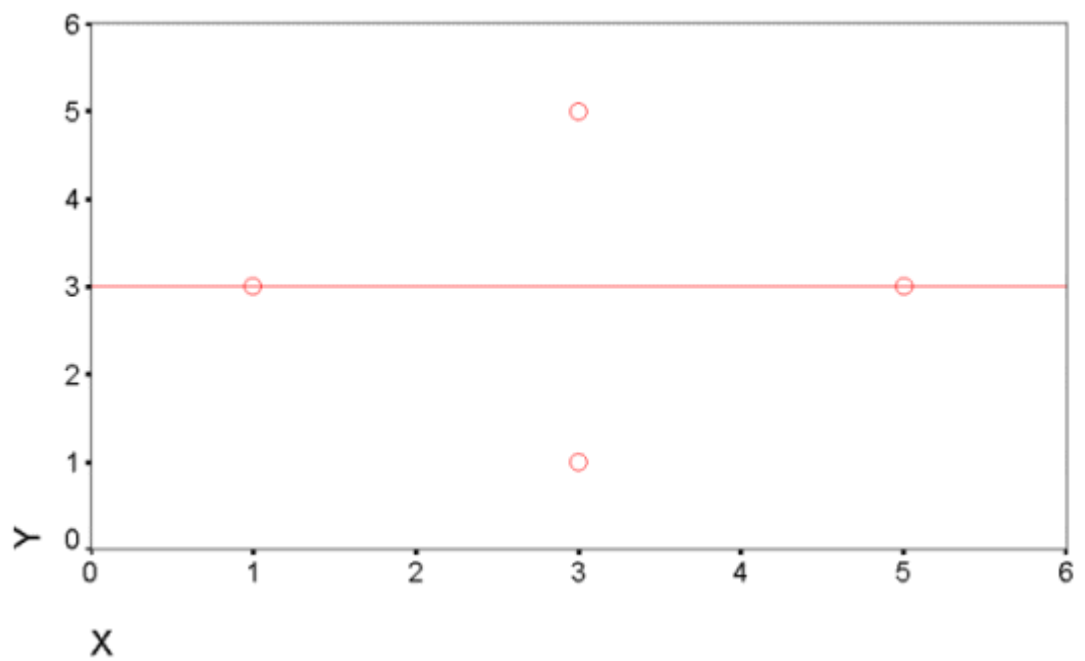


Figura 5.18 Correlación valor 0

5.3.1.1 Formulación.

El coeficiente de correlación de Pearson viene definido por la siguiente expresión:

$$r_{xy} = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}} \quad (5.36)$$

Donde $x = X - \bar{X}$ e $y = Y - \bar{Y}$

Esto es, el coeficiente de correlación de Pearson hace referencia a la media de los productos cruzados de las puntuaciones estandarizadas de X y de Y. Esta fórmula reúne algunas propiedades que la hacen preferible a otras. A operar con puntuaciones estandarizadas es un índice libre de escala de medida. Por otro lado, su valor oscila, como ya se ha indicado, en términos absolutos, entre 0 y 1.

Téngase en cuenta que las puntuaciones estandarizadas muestran, precisamente, la posición en desviaciones tipo de un individuo respecto a su media. Reflejan la medida en que dicho individuo se separa de la media. En este sentido, supongamos que para cada individuo tomamos dos medidas en X e Y. La correlación entre estas dos variables será N perfecta positiva cuando cada individuo manifieste la misma superioridad o inferioridad en cada una de ellas. Esto se cumple cuando su posición relativa sea la misma, es decir, cuando sus puntuaciones tipo sean iguales ($Z_x = Z_y$). En este caso la formula de la correlación se transforma en:

$$r_{xy} = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}} = \frac{\sum xx}{\sqrt{(\sum x^2)(\sum x^2)}} = 1 \quad (5.37)$$

Otra forma matemática para definir la correlación viene dada por la relación entre la covarianza y las desviaciones típicas de las dos variables:

$$r_{xy} = \frac{\sum S_{xy}}{S_x S_y} \quad (5.38)$$

5.3.2 Umbral óptimo.

Los métodos del valor umbral son un grupo de algoritmos cuya finalidad es segmentar gráficos rasterizados, es decir separar los objetos de una imagen que nos interesen del resto. Con la ayuda de los métodos de valor umbral en las situaciones más sencillas se puede decidir qué píxeles conforman los objetos que buscamos y qué píxeles son sólo el entorno de estos objetos..

Como con todos los métodos de segmentación se trata de asignar cada píxel a un cierto grupo, llamado comúnmente "segmento". La imagen que se debe segmentar, como cualquier gráfico rasterizado, está compuesta por valores numéricos (uno o más valores de color para cada píxel). La pertenencia de un píxel a un cierto segmento se decide mediante la comparación de su nivel de gris (u otro valor unidimensional) con un cierto valor umbral. El nivel de gris de un píxel equivale a su nivel de luminosidad; el resto de la información sobre el color no se tiene en cuenta. Dado que esta comparación de valores se realiza individualmente para cada píxel, al método del valor umbral se le considera un método de segmentación orientado a píxeles.

El modo de obtener el valor de este umbral óptimo se define como:

$$Sum_Q = Sum_Q + L \quad (5.39)$$

$$Number_Q = Number_Q + 1 \quad (5.40)$$

$$Mean_Q = Sum_Q / Number_Q \quad (5.41)$$

$$U = (Mean_0 Number_1 + Mean_1 Number_0) / (Number_0 + Number_1) \quad (5.42)$$

Siendo los distintos estados: claro y oscuro, en función de si se localizan en la imagen como blanco o negro.

6. Desarrollo

En este apartado se explicará de forma detallada cada una de las decisiones tomadas para la elaboración de la aplicación. A continuación, se irán enumerando los pasos seguidos para un correcto funcionamiento y consecución del fin último del software.

El desarrollo de esta aplicación, basada en procesos morfológicos, se puede resumir en dos pasos:

- Localización de la matrícula.
- Procesado y análisis de los dígitos de la matrícula.

Este apartado se dividirá en función de los modos de trabajo de los que dispone la aplicación diseñada.

Las herramientas necesarias para poder hacer uso de esta aplicación se muestran en el Anexo I.

Las características que deben cumplir tanto las imágenes como los videos y los sistemas de captura se exponen en el Anexo II.

6.1 Modo imagen

Dentro de este modo se presentan dos formas de obtener los resultados que se persiguen con el proyecto, puede ser automático o mediante recorte de una imagen. A continuación se comenta el funcionamiento de cada uno de estos modos.

6.1.1 Localización de forma automática

Este modo consiste en detectar de forma automática la matrícula, sin necesidad de la interacción del usuario. El siguiente esquema muestra los pasos a seguir para obtener la solución deseada.

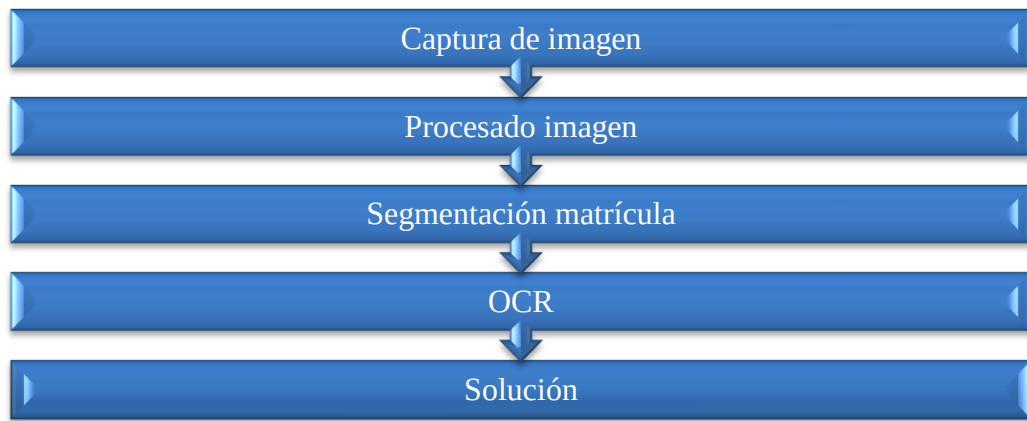


Figura 6.1 Proceso Imagen Automático

1. Captura de imagen: Para comenzar se debe estar a disposición de poder suministrar una imagen que ilustre una matrícula de coche. Esa imagen es cargada y enviada a la función *Matricula.m* la cual se encargará de su procesado. Una vez está la imagen disponible para ser usada, se pasará al siguiente paso.



Figura 6.2 Imagen original

2. Procesado de imagen: En este punto se realizará una serie de transformaciones a la imagen para poder realizar una correcta segmentación del área que contiene la matrícula. Para ello, la primera transformación que sufrirá la imagen será un cambio de formato, pasando de ser una imagen a color (RGB) a una en escala de grises. Este cambio se realiza para poder trabajar con la iluminación de cada píxel y por tanto poder obtener un umbral adecuado para su procesado.



Figura 6.3 Imagen en escala de grises

Una vez con la imagen en escala de grises se debe adecuar su tamaño al que ha sido diseñado el software. Este tamaño es de un máximo de 4012784 píxeles, correspondiente a una resolución de 768x1280 ya que como se verá en el apartado de segmentación un tamaño mayor no resultaría adecuado para obtener resultados lógicos.

El procesado previo de la imagen se daría por finalizado con estas dos transformaciones.

3. Segmentación de la matrícula: Esta es la fase más importante del proceso, ya que es la encargada de localizar el área donde se sitúa la matrícula. Para comenzar, con la intención de eliminar niveles de grises oscuros que puedan afectar posteriormente a la obtención de un umbral óptimo, se transforman píxeles de colores grises oscuros a color negro.



Figura 6.4 Imagen en escala de grises con eliminación de grises oscuros

Se puede apreciar comparando la Figura 6.3 y 6.4 como se ha acentuado la zona donde se localiza la matrícula al modificar disminuyendo el valor digital de píxeles grises oscuros.

Seguidamente se produce un recorte automático de la imagen con el objeto de eliminar de esta detalles e información que no sea de utilidad para el objeto del programa.



Figura 6.5 Recorte automático de la imagen

Se aprecia claramente como se ha eliminado en su mayoría la zona de la imagen cuya información es irrelevante para la finalidad de la aplicación.

A continuación, se realiza un filtro de detección de bordes del tipo Sobel (definido en el apartado '*Métodos y materiales*'), ya que una de las características común en todas las matrículas es el contraste entre los dígitos y el fondo de la placa. Esa cualidad es de gran utilidad para la detección del borde de los dígitos y de la matrícula con el resto del vehículo.

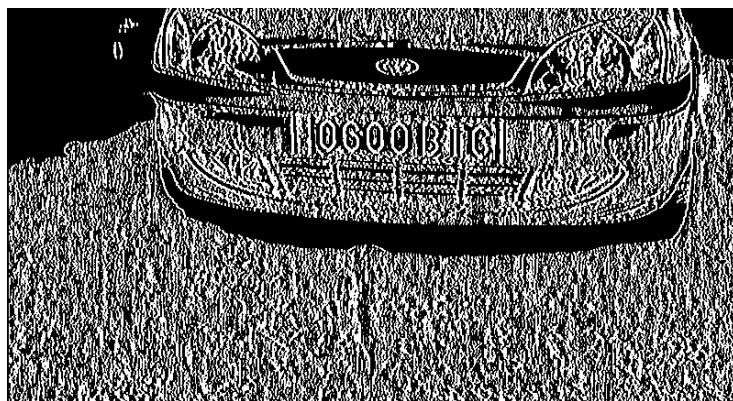


Figura 6.6 Detección de bordes de Sobel.

A partir de este resultado se procede al cálculo del umbral óptimo (definido en el apartado ‘*Materiales y métodos*’) para poder localizar solo los dígitos de la matrícula. Una vez calculado el umbral se aplica a la imagen, dando como resultado la siguiente figura:



Figura 6.7 Imagen umbralizada

Llegado a este punto se procede a realizar una serie de operaciones morfológicas con el objeto de intentar quedarse solo con el área donde se sitúan los dígitos. En primer lugar, se eliminarán todas las áreas de color blanco con un tamaño menor a 50 píxeles. Seguidamente, se realiza un cierre con un elemento estructurante de línea horizontal, de tal forma que se consigan unir los dígitos en una sola área:



Figura 6.8 Proceso de cierre.

Una vez realizado este paso, se hará uso de una apertura con un elemento estructurante de línea vertical, para poder eliminar las áreas con poca altura, y tres dilataciones para expandir las áreas restantes. El resultado es el siguiente:

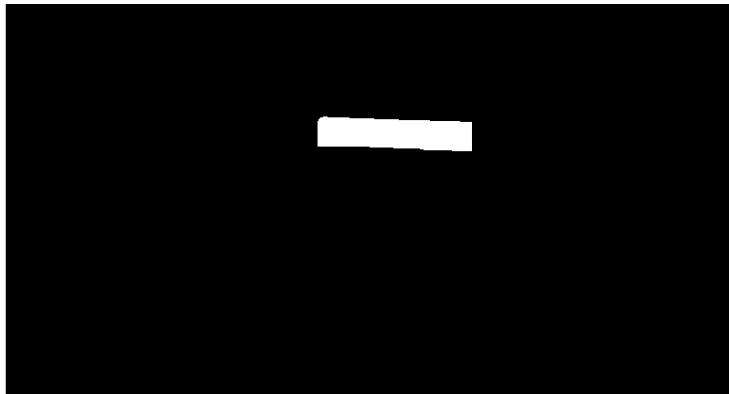


Figura 6.9 Resultado obtenido tras las operaciones morfológicas.

El sistema buscará áreas que contengan al menos 12100 píxeles blancos y que no superen los 80000. Una vez el sistema ha localizado el área deseada, se hace un recorte automáticamente de esa zona en la imagen original, perteneciente a la Figura 6.5:



Figura 6.10 Corte área

Se realiza una segunda umbralización (denominada umbral2), esta vez en el recorte (Figura 6.10) para poder localizar de forma más precisa los dígitos.



Figura 6.11 Umbralización matrícula

Con esta umbralización el sistema puede intentar localizar áreas de píxeles blancos en las que el número de estos sea superior a 4500. Con esa sección localizada se procederá de nuevo a un recorte automático para así poder eliminar las zonas oscuras de los bordes. En este caso en particular detectaría solo una zona.



Figura 6.12 Segundo recorte

Comparando la Figura 6.10 y 6.12 se puede apreciar las zonas que han sido eliminadas debido a que no contienen información de interés, como por ejemplo los bordes de la matrícula.

Siguiendo con la segmentación, se vuelve a realizar una umbralización (denominado umbral3), esta vez a un valor de 100 en la escala de niveles de grises. Obteniendo el siguiente resultado:



Figura 6.13 Tercera umbralización

Una vez se tiene la matricula umbralizada, se puede proceder a una localización de todos los dígitos que la componen. Esta tarea se realiza buscando en la imagen cualquier área de píxeles blancos, ya que son estas áreas las que contendrán los diferentes dígitos. Para ser considerada un área posible dígito debe tener al menos un 0.1% de tamaño sobre el total del último recorte. En la siguiente Figura 7. se puede apreciar las distintas áreas localizadas por el sistema.



Figura 6.14 Localización de dígitos

Una vez tenemos las áreas localizadas, se puede dar por concluido la segmentación de la imagen. Con estos resultados, se podrá realizar el proceso OCR que se explica a continuación.

4. OCR: El Reconocimiento Óptico de Caracteres es un proceso dirigido a la digitalización de textos, los cuales identifican automáticamente a partir de una imagen símbolos o caracteres que pertenecen a un determinado alfabeto, para luego almacenarlos en forma de datos.

Con los dígitos localizados, el siguiente paso consiste en el procesado de cada uno de ellos. Para esta tarea se realizará una correlación (véase capítulo '*Materiales y métodos*') del dígito a analizar y cada uno de los dígitos almacenados en la base de datos de imágenes con formas alfanuméricas, de tal forma que el valor, ya sea número o letra, de mayor correlación con el dígito de la imagen será el valor que se busca.

Realizando este proceso con todas las áreas encontradas se podrá obtener el resultado deseado.

Con este proceso repetitivo finalizado se procede a averiguar si la matrícula obtenida se podría corresponder con la composición de una matrícula auténtica. Por tanto, en este momento, se produce una comparación de la matrícula obtenida con el conjunto de distintos modelos de matrículas que existen. Esos modelos son (L=Letra y N=Número):

- LNNNNL
- NNNNLLL
- LNNNNLL
- LLNNNNL
- LLNNNNLL

En el caso de que no coincidiese con ninguno de los modelos anteriormente enumerados, se procedería a buscar el modelo de mayor semejanza con la matrícula obtenida. A partir de ese modelo se buscaría que dígito localizado es erróneo y si el error es un número o una letra. Con estos datos, el programa está en disposición de poder subsanar el error y conseguir la matrícula correcta.

Una vez la matrícula ha sido verificada se daría por finalizado este punto denominado OCR.

5. Solución: La solución que se pretendía debe de haberse obtenido gracias a los pasos comentados anteriormente, pero puede que hayan surgido errores durante el proceso. A continuación se comentarán los errores más comunes y las distintas soluciones implementadas para poder corregirlos.

Error al seleccionar umbral: Los umbrales usados para la segunda y la tercera umbralización están predefinidos, por lo que una matrícula que se sitúe en un ambiente de poca luz o contenga algún reflejo podría no dar toda la información para su correcta captura y procesamiento de sus dígitos. El siguiente ejemplo muestra una imagen y la matrícula del vehículo una vez procesada justo antes de llegar al proceso OCR:



Figura 6.15 Error de umbral

Se puede apreciar como la umbralización elimina los dígitos de la matrícula. Para corregir esta desviación hacia el objetivo, es necesario disminuir de forma progresiva tanto el umbral2 como el umbral3 para poder llegar a visualizar y capturar todos los dígitos de la matrícula. Con los nuevos umbrales, se repite todo el proceso desde la aparición del umbral2.

El resultado para este ejemplo se presenta a continuación:

Primera reducción de los umbrales:



Figura 6.16 Umbral2 – 5 y umbral3 – 20

Segunda reducción de los umbrales:



Figura 6.17 Umbral2 – 10 y umbral3 – 40

En este momento el sistema está en condiciones de poder realizar el paso OCR sin problemas para obtener los resultados buscados.

Error al seleccionar la primera área del proceso: Podría darse la situación en la cual el programa pudiese seleccionar entre más de un área con más de 12100 píxeles. En este caso si el programa seleccionase un área errónea (suelen ser focos, llantas o lugares donde se produzca un gran contraste), una vez comience con el proceso descrito en las páginas anteriores y descubra que no existe ninguna matrícula en tal área, automáticamente sería desechada para dar comienzo al análisis de la siguiente área. Dentro de este modo para solucionar un error se sitúa la corrección del umbral también, de tal forma que en el caso de encontrar el área correcta, el programa pueda variar el umbral para obtener el objetivo deseado.

Se muestra a continuación un ejemplo del error en cuestión y de su solución:



Figura 6.18 Imagen original error área



Figura 6.19 Áreas localizadas

En este caso el sistema detectaría tanto el foco como la matrícula. La primera área que analiza es el foco (debido a que se seleccionan las áreas de izquierda a derecha en la imagen) sin obtener resultado alguno, por lo que pasaría al área siguiente. De esta forma detectaría de forma correcta los dígitos que se perseguían.

6.1.2 Procesado mediante recorte de la imagen

Este modo de funcionamiento requiere de la interacción del usuario para realizar una selección de la zona en la que se sitúa la matrícula. Es un modo de funcionamiento auxiliar, diseñado para los casos en los que por algún motivo el modo automático no tenga la capacidad suficiente para obtener los resultados que se persiguen, que no son otros que la localización y correcto procesado de la matrícula.

El esquema de funcionamiento es muy parecido al realizado en el apartado anterior con algunas modificaciones:

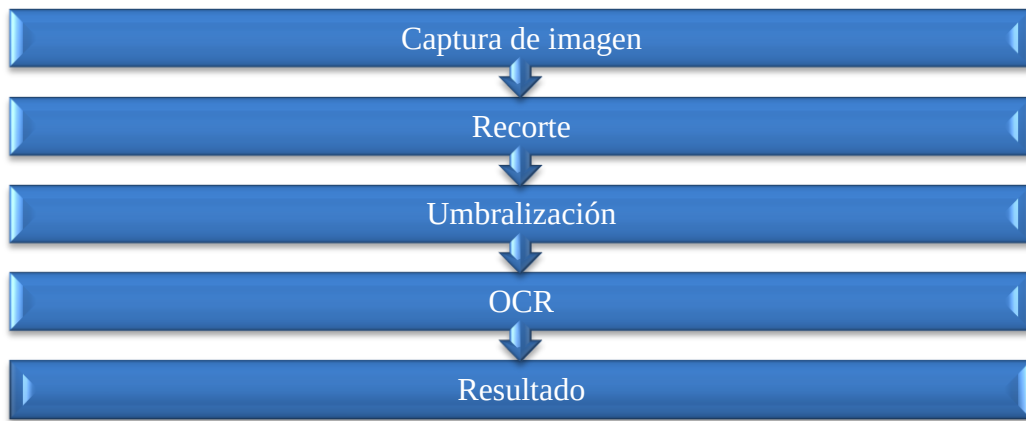


Figura 6.20 Esquema procesado mediante recorte.

1. Captura de imagen: Al igual que en el apartado anteriormente descrito, se debe de estar en posesión de la imagen que se desea procesar. Para pasar al siguiente punto, la imagen ha de ser cargada en el programa.
2. Recorte: Con la imagen cargada en el software, se está en posición de que el usuario interaccione con el sistema para seleccionar de forma manual la zona de la imagen que desea analizar. Ese recorte se transformará, al igual que en el caso anterior, de RGB a escala de grises y será enviado a '*Matricula_Recorte.m*' donde será procesado. Las siguientes imágenes muestran como es la imagen original y el recorte que se ha producido una vez transformado a escala de grises:



Figura 6.21 Imagen original para recorte



Figura 6.22 Recorte en escala de grises

3. Umbralización: En este modo de trabajo no se necesita de segmentación de la imagen para localizar la matrícula, ya que esta ha sido recortada por el usuario. Por tanto, se pasaría al paso de la umbralización.
- El sistema con el recorte cargado comenzará aplicando un umbral a la imagen:



Figura 6.23 Umbral escala de grises

Tras esta umbralización, se procede a realizar una dilatación seguida de una erosión con el objeto de poder eliminar pequeños detalles que sean de interés y pudiesen confundir al programa a la hora de localizar áreas.

4. OCR: El sistema comenzará a analizar todas las áreas de píxeles blancos que existan en la imagen. Cada una de esas áreas será correlada con una base de imágenes de caracteres alfanuméricos, se producirá el mismo proceso que en el modo automático.

Cuando el sistema termina de analizar cada una de las áreas, comprobará que la matrícula adquirida se corresponde con un modelo correcto de matrícula de los descritos anteriormente. Para el caso en el que el resultado no sea exactamente igual a uno de los modelos, el sistema intenta subsanar ese error corrigiendo los dígitos que hayan fallado.

5. Resultado: Una vez se ha llegado a este punto, si el sistema ha obtenido la matrícula que se buscaba se acabará el proceso. En caso contrario, volverá al punto 3, pero con una reducción del umbral y comenzará el proceso de nuevo desde esa tarea.

6.2 Modo vídeo

En el apartado anterior se explicó la forma de trabajo de la aplicación para obtener el resultado perseguido a partir de imágenes capturadas de antemano por el usuario. En este caso, el sistema se encargará de procesar vídeos para obtener las matrículas mediante dos formas de trabajar: dispositivo externo y vídeo grabado por el usuario.

6.2.1 Procesado a través de imágenes por dispositivo externo

Este modo de trabajo permite el uso de un dispositivo de captura de imágenes para el procesado y análisis de estas en directo o captura y procesado posterior. En primer lugar, se detallará el funcionamiento del modo en directo y a continuación el modo captura y procesado.

6.2.1.1 Procesado en tiempo real

La forma de trabajo que se explica en este apartado consiste en la captura de imágenes en directo del exterior a través de un dispositivo y su análisis en tiempo real.

Para comenzar, se expone el esquema de los pasos a seguir para su correcto funcionamiento y seguidamente se detalla cada uno de esos pasos.

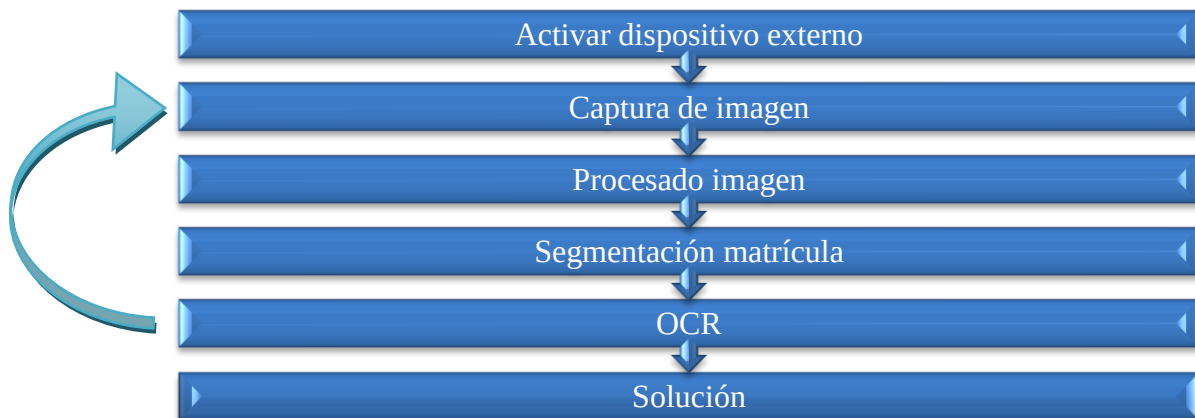


Figura 6.24 Esquema video modo directo

1. Activar dispositivo: El primer paso que se debe realizar es tener el capturador de imágenes disponible para la captura de imágenes. La activación del dispositivo alimentará a la aplicación de ilustraciones para poder analizar.
2. Captura imagen: Una vez se ha activado el dispositivo, el sistema comienza a procesar el primer frame que captura. El sistema no volverá a adquirir otro frame hasta que termine de procesar el actual.
3. Procesado imagen: La imagen obtenida a través del dispositivo se convertirá de RGB a escala de grises para poder trabajar con niveles de grises, ya que resultará un procesado con mayor facilidad. Esta imagen en ‘grises’ se envía a la función ‘*Matricula_video.m*’ donde se segmentará y se procesará. Esta función se comenta en los dos puntos siguientes.
4. Segmentación: Este paso es muy parecido al que se producía para imágenes estáticas en modo automático. La segmentación comienza con una disminución

del valor de grises oscuros a negro para optimizar el umbral óptimo que se obtendrá.

A continuación, un filtro Sobel (véase '*Materiales y métodos*') se encargará de localizar los bordes de la imagen, y por tanto de localizar los bordes de los dígitos. Seguidamente se calcula el umbral óptimo para la imagen, de tal forma que se localice la placa del coche. Se aplica ese umbral a la imagen y se procede a realizar una serie de operaciones morfológicas: un cierre, una apertura y tres dilataciones, todas ellas con un elemento estructurante de línea, para que, al igual que en el modo imagen automática, se puedan conectar los dígitos entre sí para formar una sola área (véase Figura 6.8 y 6.9).

El sistema buscará zonas de al menos 12100 píxeles, en caso de no encontrar área con tal característica se producirán una serie de dilataciones con la intención de aumentar el número de píxeles de una área. Con una zona localizada, se procede a realizarse un recorte quedándose el sistema solo con esa zona de la imagen. A ese recorte se le realiza un umbral para destacar los dígitos, en el caso de que existieran, y se buscarían áreas con 4500 píxeles como mínimo. Una vez se tiene segmentada la matrícula, se puede pasar al siguiente punto del esquema, proceso OCR.

5. OCR: Este punto es el mismo que se producía en los modos de análisis de imágenes estáticas (véase punto 4 del apartado 6.1.1).
6. Solución: Con el proceso de segmentación y localización terminado, en este punto se tiene la solución deseada y el sistema pasará automáticamente al punto 1 del esquema de nuevo para procesar el siguiente frame. En el caso de no haber localizado la matrícula, el sistema hará uso de los métodos de solución de errores desarrollados en el apartado 6.1.1 de este capítulo. Para terminar el proceso el usuario debe indicar al sistema que desea finalizar el análisis en tiempo real de imágenes.

El sistema puede devolver varias soluciones diferentes, tantas como coches sean capturados durante el tiempo que dure el proceso. Además de los dígitos de las placas capturadas, el sistema facilita el número de vehículos capturados.

6.2.1.2 Procesado mediante captura

En este modo la aplicación necesita de la actuación del usuario. El usuario ordena al sistema que capture un frame durante un momento determinado mientras se obtienen imágenes a través de un dispositivo externo. Esa imagen será almacenada y se podrá procesar en el instante en el que el usuario lo solicite. La siguiente Figura 7.muestra el esquema de funcionamiento del sistema para esta tarea:



Figura 6.25 Esquema video, captura y procesado

Si se compara con el esquema de modo de video un directo (véase Figura 6.24), se puede apreciar que los pasos son los mismos. La única diferencia entre ambos es que el sistema que se está exponiendo solo captura la imagen cuando se lo solicite el usuario, mientras que el anterior analizaba de forma continua las imágenes adquiridas a través del dispositivo externo.

Por tanto, la única tarea que varía respecto al modo anterior es ‘*Captura imagen*’, ya que en este caso no captura el primer frame, sino que el usuario interactúa con el sistema para adquirir la imagen que desee en tiempo real a través del dispositivo externo. Con la imagen ya almacenada, el usuario está en posición de solicitar su procesado. Para llevar este punto a cabo, la imagen será modificada de RGB a escala de grises y enviada a

‘*Matricula_video.m*’, donde se tratará para intentar obtener el resultado buscado. Los pasos denominados ‘*Procesado imagen*’, ‘*Segmentación*’ y ‘*OCR*’ son los mismos que en el modo anterior (véase apartado 6.2.1.1).

El sistema, al finalizar, también dispone de las soluciones tomadas para los errores en umbral y localización de áreas, que fueron definidos en el apartado 6.1.1.

Este modo devuelve una solución única, ya que solo se analiza una imagen, un frame de forma individual.

6.2.2 Análisis de video

Como última función de la aplicación, se desarrolla en las siguientes líneas una función capaz de explorar un vídeo grabado y devolver al usuario una lista con las matrículas encontradas y el número de coches que han pasado por el vídeo.

Este modo de funcionamiento sigue el siguiente esquema de tareas, las cuales serán desarrolladas a continuación:

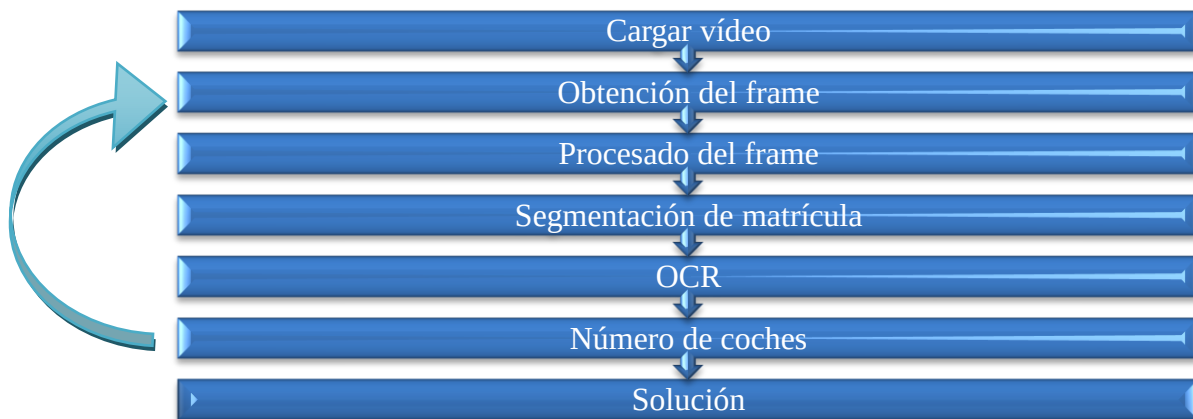


Figura 6.26 Esquema video grabado

1. Cargar vídeo: En primer lugar, será necesario que el usuario importe al software un vídeo, el cual habrá sido grabado de antemano. El vídeo cargado se mandará a la función denominada ‘*Calcular_coches.m*’ donde comenzará el proceso.

2. Obtención de frame: Con el vídeo en la función '*Calcular_coches.m*' comenzará el sistema a analizarlo, adquiriendo uno frame de cada cinco del vídeo, es decir, comenzará procesando el frame 1, seguidamente el frame 6, el frame 11, el frame 16 ... y así hasta el frame último del vídeo. El hecho de procesar frames de cinco en cinco es porque en los intermedios la información será de poca relevancia, ya que esa información habrá sido capturado en el primer frame o será capturada en el quinto frame siguiente, debido a que un coche grabado aparecerá el vídeo más de 5 frames.
3. Procesado del frame: La imagen del vídeo a analizar se transformará, al igual que en el resto de los casos, de RGB a escala de grises. Seguidamente, se envía a '*Matricula_video*', donde seguirá el proceso definido para los dos modos de trabajo anteriores.
4. Segmentación: La forma de obtener la zona de la imagen donde se sitúa la matrícula se obtiene realizando la misma segmentación que se describió en el punto 6.2.1.1 de este capítulo.
5. OCR: El procesado basado en correlación dígitos y obtención de los valores adecuados, junto con la comprobación de modelo correcto de matrícula es idéntica a la definida en el punto 6.1.1.
6. Número de coches: El sistema compara la matrícula última que se adquirió y la compara dígito a dígito con la matrícula que acaba de obtener. Llegado a este punto pueden suceder dos acciones debidas a esa comparación: el sistema contabilizará los dígitos que son idénticos en ambas series de valores alfanuméricos, si este contador es mayor a 4 se desecha la matrícula anterior y el sistema se queda con la más reciente. En el caso que existan menos de 4 similitudes, el sistema se quedará en posesión tanto de la anterior como de la más reciente. Por cada matrícula que almacena nueva, aumenta en uno el contador de número de coches que se han localizado.

Al finalizar la comparación descrita anteriormente, la aplicación volverá a la tarea '*Obtención de frame*' para procesar el quinto frame siguiente y comenzar de nuevo el proceso de búsqueda.

Cuando se termina de analizar el último frame del vídeo se puede dar por concluido el proceso. Para finalizar se pasa a la tarea final.

7. Solución: Ahora el proceso ha finalizado, ya que se ha llegado al último frame y se ha procesado. La aplicación devuelve como solución una lista con el número total de coches obtenidos y las matrículas de los distintos vehículos.

7. Interfaz gráfica de la aplicación

En los sistemas informáticos, la relación humano-computadora se realiza por medio de la interfaz, que se podría definir como mediador. Cuando existen dos sistemas cualesquiera que se deben comunicar entre ellos la interfaz será el mecanismo, el entorno o la herramienta que hará posible dicha comunicación.

Se pueden definir básicamente dos tipos de interfaces:

- La interfaz física: un ratón y un teclado que sirven para introducir y manipular datos en nuestro ordenador.
- La interfaz virtual o interfaz gráfica (GUI) que permite, mediante iconos (cursor + objetos gráficos metafóricos), interactuar con los elementos gráficos convirtiendo al ser humano en usuario de la aplicación.

Estas dos mediaciones son relaciones del tipo entrada de datos (*input*). Al igual que se tiene una entrada, se necesita algo que facilite la salida de datos (*output*), para esto se tiene, por ejemplo, la pantalla del ordenador, donde se visualizan estas interfaces gráficas, o la impresora, donde se imprimen los datos.

En definitiva GUI es una interfaz de usuario en la que una persona interactúa con la información digital a través de un entorno gráfico de simulación. Este sistema de interacción con los datos se denomina WYSIWYG (*What you see is what you get*, ‘lo que ves es lo que obtienes’), y en él, los objetos, iconos (representación visual) de la interfaz gráfica, se comportan como metáforas de la acción y las tareas que el usuario debe realizar. Estas relaciones también se denominan interfaces objetos-acción (*object-action-interface*, OAI).

Para diseñar una interfaz es necesario pasar por cuatro etapas:

1. Análisis de requerimientos del producto, análisis de las tareas. Conocimiento del usuario. Generación de posibles metáforas y análisis del tipo de diálogo. Revisión.

2. Generación de prototipos virtuales (*layouts*) o físicos para investigar desde lo general hasta el detalle. Desarrollo de la aplicación, del sitio o del sistema.
3. Planificación (desarrollo del plan, definición de las medidas, selección de participantes, formación de observadores, preparación de los materiales). Test (prueba piloto, test con usuarios).
4. Conclusión (análisis de los datos, elaboración del informe, resultados y recomendaciones).

En el siguiente esquema se pueden apreciar los diferentes módulos en los que se divide el proceso de una interfaz gráfica:

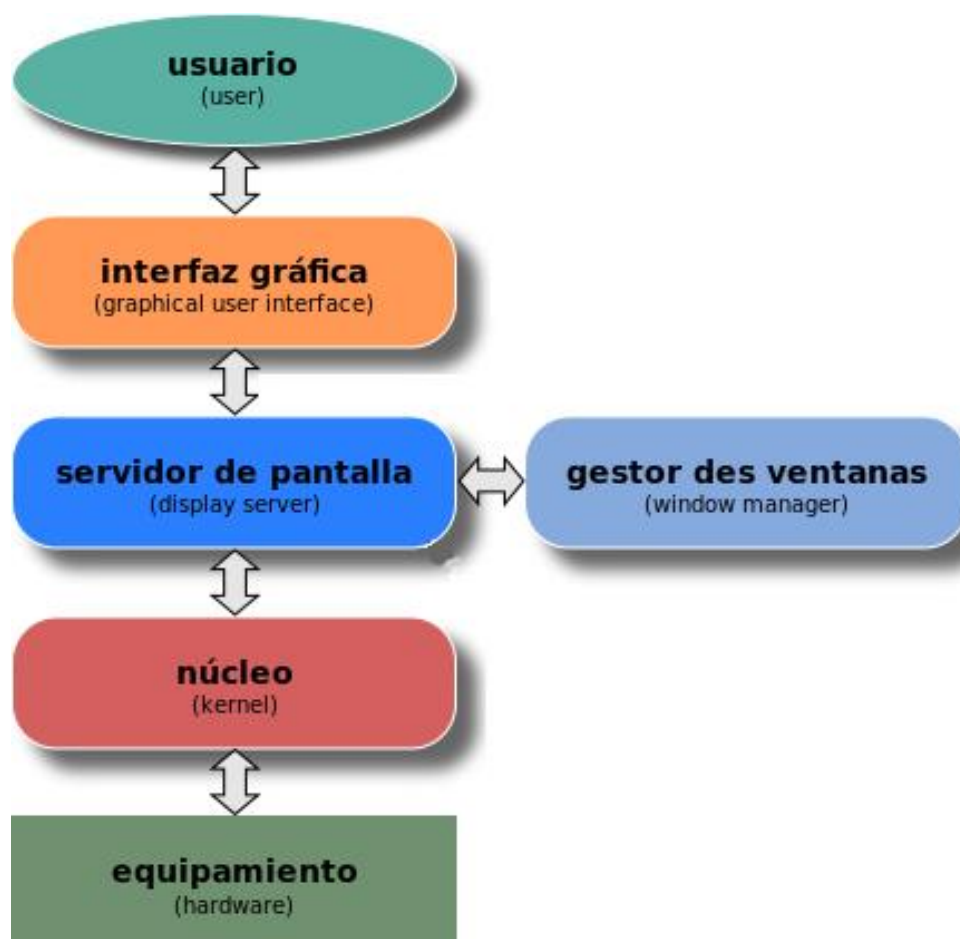


Figura 7.1. Esquema función interfaz gráfica

Comienza con el usuario, el cual será el encargado de seleccionar las diferentes funcionalidades que desea ejecutar en el programa. A continuación, la interfaz gráfica que aparece en la pantalla mandará esa sentencia seleccionada por el usuario al núcleo y este, correspondientemente, al hardware oportuno para la operación.

Existirá un gestor de ventanas para poder navegar entre las distintas ventanas de las que disponga el software.

7.1 Interfaz principal

La interfaz principal es la puerta para el inicio de la aplicación, es la portada de esta. Su función principal es dar paso al software. Consiste en una interfaz sencilla, para no confundir al usuario.

Está formada por tres partes claramente diferenciadas:

- Información. Texto donde se indica el autor del proyecto y la universidad que ha supervisado la ejecución de este.
- Imagen. Icono de la universidad.
- Pulsadores. Para ejecutar procesos de la aplicación.

Consta de dos botones para interactuar, se obvian los pulsadores propios de la ventana, como son el de minimizar (-) y el de cerrar (x):

- Comenzar. Pulsando este botón la interfaz se dirige a la aplicación de procesado de imágenes.
- Salir. Como su nombre indica es para salir. Si se desea abandonar la aplicación, pulsando ese botón se cerraría la ventana y el proceso.

A continuación se muestra una imagen de la interfaz principal.



Figura 7.2 Interfaz Principal

7.2 Interfaz para procesado de imágenes estáticas

La función de esta interfaz es la de ejecutar la parte de la aplicación correspondiente a procesar imágenes estáticas suministradas por el usuario. Es más compleja que la anterior debido a que ahora contiene muchas funcionalidades y herramientas para el tratamiento de la imagen (recordar que la anterior era solo una pequeña portada de introducción a la aplicación).

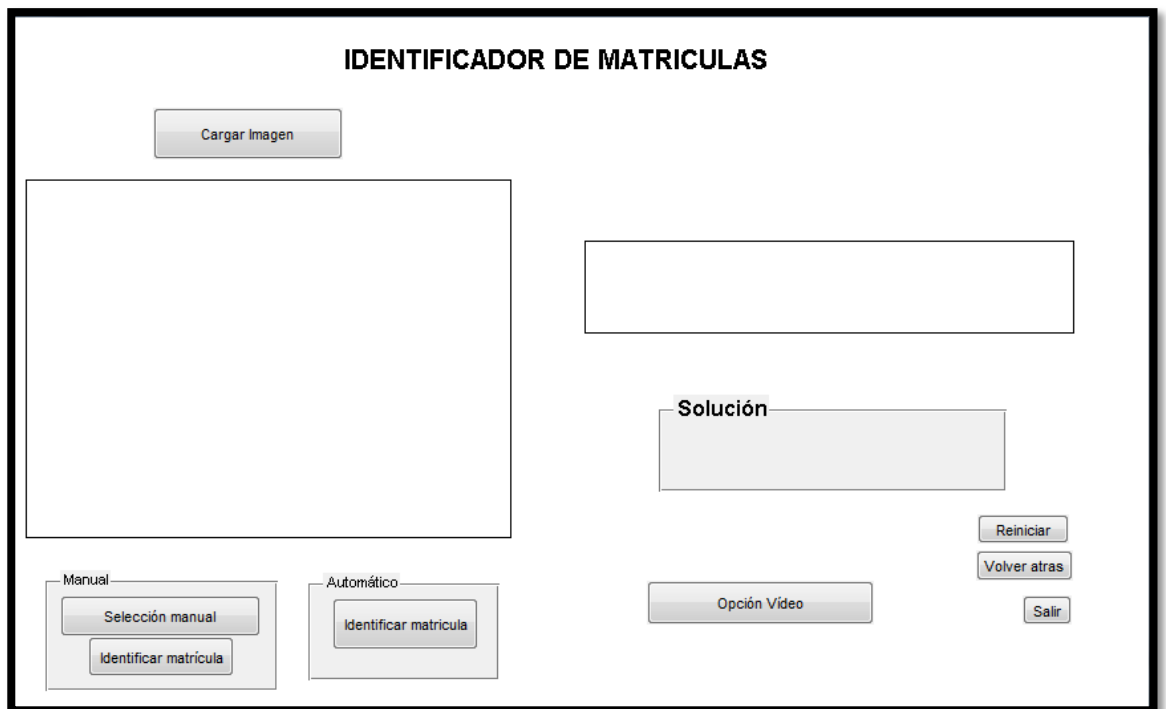


Figura 7.3 Interfaz Imagen

A continuación se enumerarán cada uno de los objetos de la interfaz y se comentarán todos los pasos para la correcta utilización de esta:

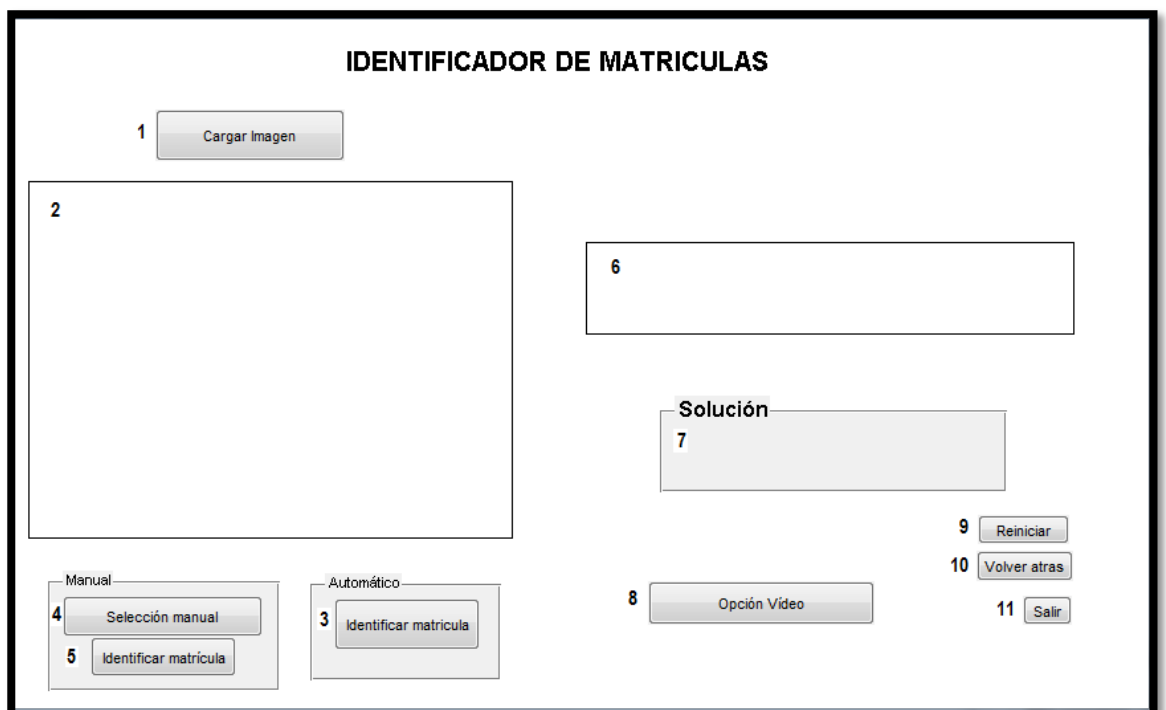


Figura 7.4 Interfaz Imagen Numerada

1. Cargar imagen.
2. Pantalla imagen.
3. Identificar matrícula.
4. Selección manual.
5. Identificar matrícula.
6. Pantalla matrícula.
7. Solución.
8. Opción vídeo.
9. Reiniciar.
10. Volver atrás.
11. Salir.

Para comenzar el uso de esta ventana de la interfaz, en primer lugar se pulsa el botón 1, llamado '*Cargar imagen*'. De esta forma, como su propio nombre indica, se podrá seleccionar la imagen que se desea procesar. Al pulsar el pulsador '*Cargar imagen*' se abrirá la siguiente ventana:

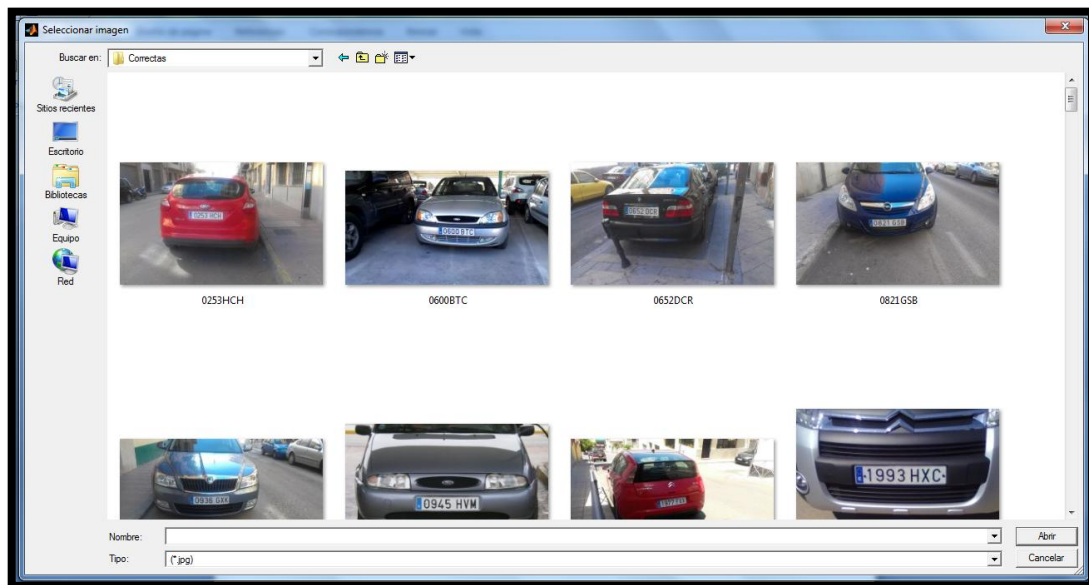


Figura 7.5 Explorador imágenes

La Figura 7.5 muestra un explorador de imágenes. A través de él se selecciona la imagen que se desea analizar. El explorador solo muestra los archivos con formato '*.jpg*',

que son para los que ha sido diseñada la exploración de imágenes. Una vez se ha elegido la imagen del vehículo para procesar se pulsa el botón '*Abrir*'.

Realizado este paso, ya se tiene la imagen montada en la interfaz. Será mostrada en '*Pantalla imagen*' (ver Figura 7.6).



Figura 7.6 Imagen en '*Pantalla imagen*'

Una vez se tiene la imagen deseada en pantalla, se puede realizar el sistema consta de dos tipos de procesado: automático o manual.

1. Procesado de forma automática. Para activar esta función se debe pulsar en el cuadro denominado '*Automático*' el botón '*Identificar matrícula*' (Botón 3 en la Figura 7.4). La siguiente imagen muestra tal botón señalado por una elipse azul.



Figura 7.7 Identificar matricula

Una vez pulsado el botón, la interfaz manda la orden al sistema de procesar esa imagen. El resultado de esta acción se mostrará en '*Pantalla matrícula*' (cuadro 4 de la Figura 7.4). Se ilustrará en el cuadro la imagen de la matrícula capturada en blanco y negro. A continuación, y sin necesidad de que el usuario interactúe, se mostrará la matrícula en el cuadro '*Solución*' (cuadro 5 de la Figura 7.4) en el formato de texto denominado 'MS Sans Serif, en tamaño de fuente 22 y en negrita.



Figura 7.8 Pantalla matrícula



Figura 7.8 Imagen Solución

Puede que por algún motivo el sistema no sea capaz de reconocer la matrícula o se produzca algún error durante el procesado de la imagen. Para el caso de no localización de la matrícula, en el cuadro ‘*Solución*’ aparecerá el mensaje ‘*NO LOCALIZADA*’. Para intentar paliar estas carencias, se diseñó el seleccionado de forma manual.



Figura 7.9 No localizada

2. Procesado de forma manual. Para cualquier imagen en la cual no se pueda identificar de forma automática la matrícula. Consiste en realizar un recorte de forma manual de la matrícula sobre la imagen elegida. Para este proceso se siguen los siguientes pasos:

Para comenzar, partiendo de haber seleccionado la imagen a procesar (véase Figura 7.5 y 7.6), se pulsa en el menú ‘*Manual*’ el botón ‘*Selección manual*’ (en la Figura 7.4 botón 4). En este momento, se está en disposición de poder ‘cortar’ y seleccionar el fragmento de imagen que se desea examinar. Pinchando sobre la imagen con el botón izquierdo del ratón y arrastrando, se crea un rectángulo dentro del cual situaremos la matrícula del automóvil. Para finalizar el recorte, se debe

pinchar dos veces con el ratón dentro del rectángulo que se ha formado. El área seleccionada se mostrará en ‘*Pantalla matrícula*’ (cuadro 6 de la Figura 7.4).

Con el recorte listo para ser procesado, se procede a activar el botón ‘*Identificar matrícula*’ del cuadro denominado ‘*Manual*’ (botón 5 en la Figura 7.4), con el cual se inicializará el proceso de identificación.

Tras el análisis de esa nueva imagen se obtendrán los dígitos de la matrícula que anteriormente no se localizaron por distintos factores.

Este proceso se muestra paso a paso en las figuras que siguen a continuación, desde la Figura 7.10 hasta la Figura 7.13.

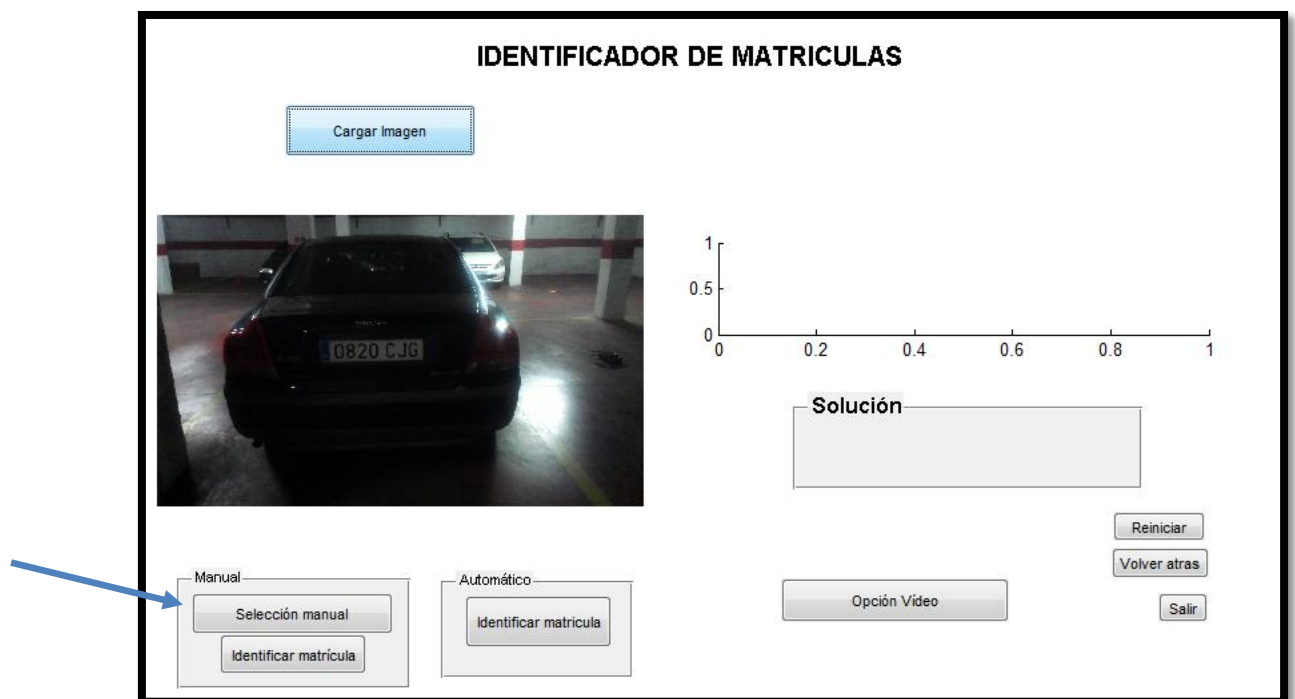


Figura 7.10 Selección manual

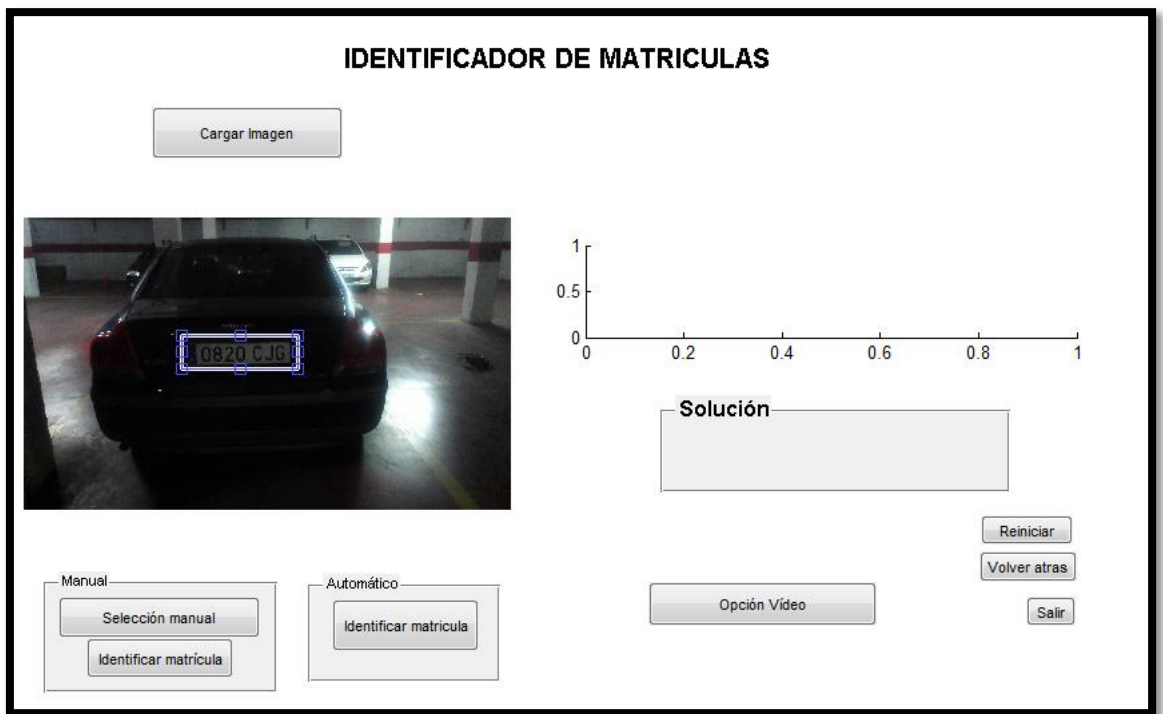


Figura 7.11 Selección recorte



Figura 7.12 Recorte

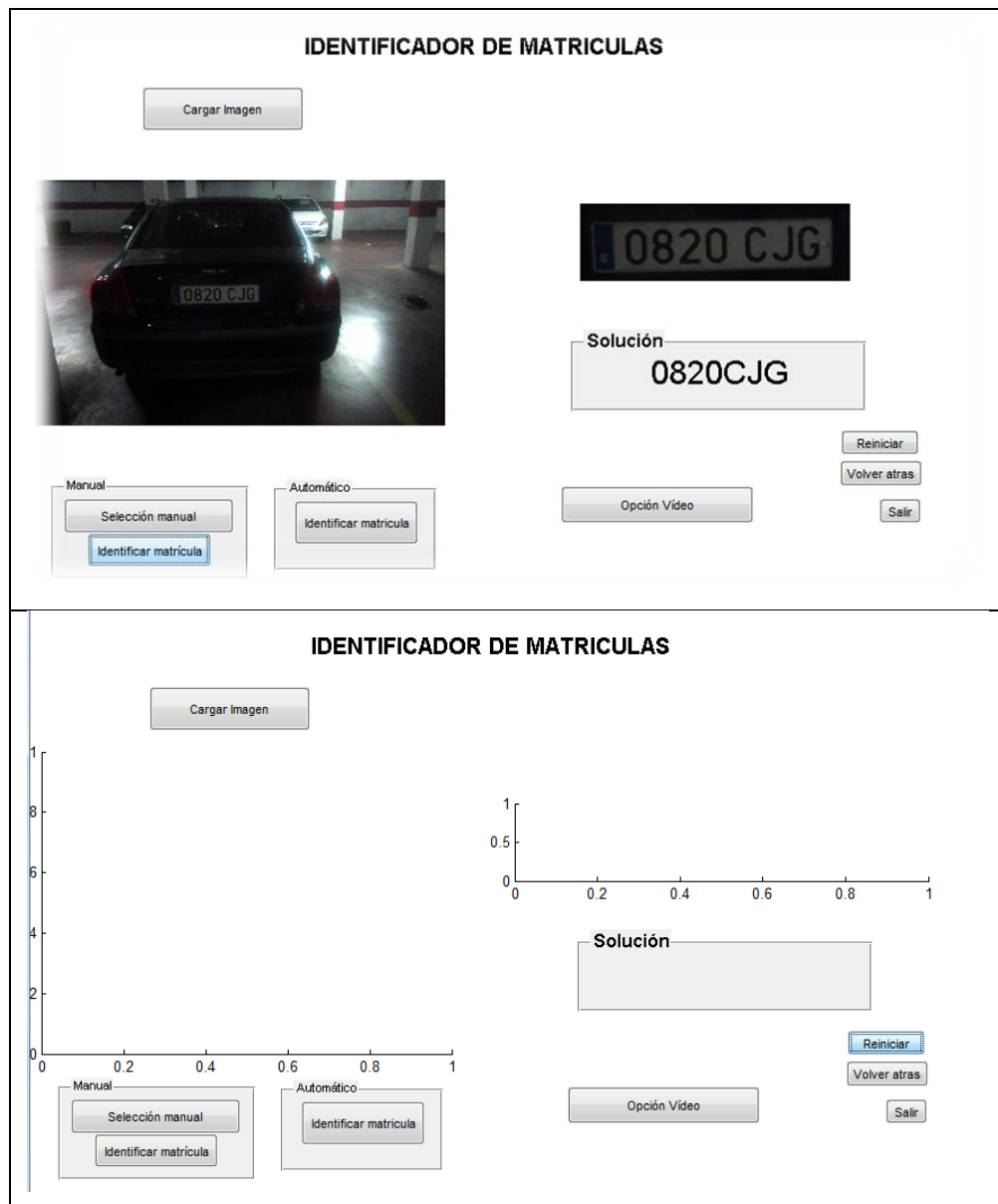


Figura 7.13 Solución recorte

Ahora se puede observar como resultado que para una imagen que no se procesaba de forma automática, de forma manual se obtiene el resultado deseado (compárese Figura 7.9 y Figura 7.13).

Existen 3 botones para funciones básicas en la interfaz de imágenes. Todos se pueden usar tanto cuando el programa ha funcionado tanto en modo manual o como en automático.

- Reiniciar. Pulsador muy útil si se quiere volver a dejar la interfaz como al comienzo de su utilización. A continuación se expone un ejemplo del uso de este botón (número 9 en la Figura 7.4):



- Volver atrás. Asignado como botón 10 en la Figura 7.4, sirve para volver a la interfaz principal, descrita en el punto anterior.
- Salir. Pulsador 11 en la Figura 7.4, con el se permite cerrar el programa. Su función es idéntica al aspa (x) que aparece en la parte superior derecha de la ventana del programa.

Para finalizar la descripción del funcionamiento de esta interfaz, solo queda describir la función del pulsador llamado '*Opción vídeo*'. Por medio de este botón se puede acceder a la interfaz para análisis de vídeo, la cual será detallada en el siguiente punto.

7.3 Interfaz para procesado de vídeo

El objetivo de este punto es poder facilitar la comprensión del funcionamiento y utilidades de la interfaz de la aplicación que se encarga del análisis de vídeos.

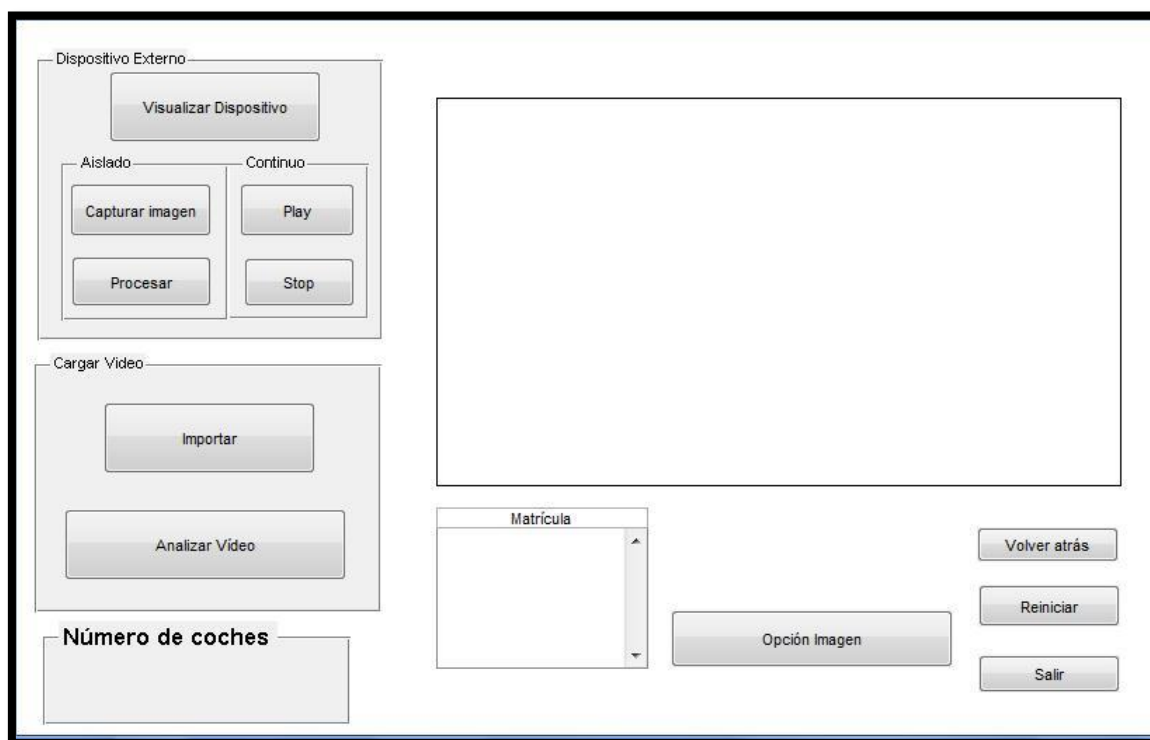


Figura 7.14 Interfaz video

La interfaz consta de dos modos de funcionamiento: Dispositivo externo y Cargar Video. En las páginas que siguen se detallará el funcionamiento de cada uno de los modos.

En primer lugar se numerará cada uno de los botones y cuadros de diálogo para una mejor guía cuando se nombre a cada uno de ellos.

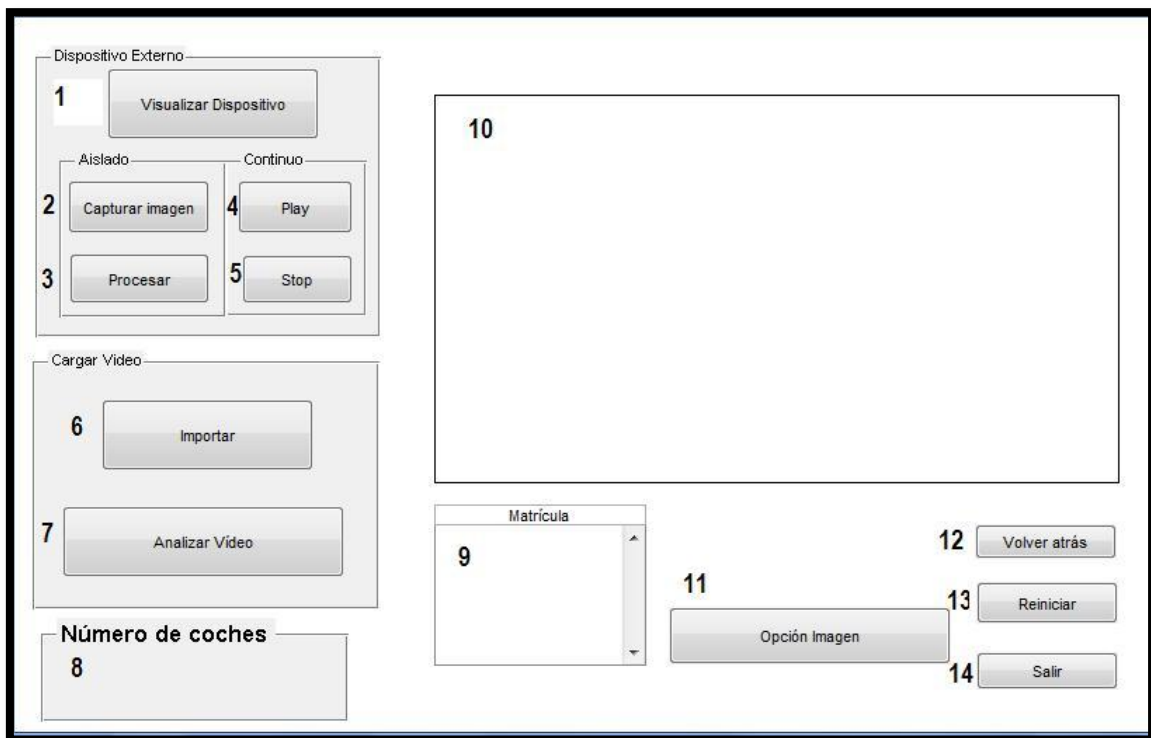


Figura 7.15 Numeración Interfaz Vídeo

1. Visualizar Dispositivo.
2. Capturar imagen.
3. Procesar.
4. Play.
5. Stop.
6. Importar.
7. Analizar Vídeo.
8. Diálogo Número de coches.
9. Diálogo Matrícula.
10. Pantalla vídeo.
11. Opción imagen.
12. Volver atrás.
13. Reiniciar.
14. Salir.

OPCIÓN DISPOSITIVO EXTERNO.

Para hacer uso de este modo se debe tener conectado al ordenador algún dispositivo que permita la entrada de video en directo. El dispositivo externo conectado puede ser una web cam, cámara de vídeo, teléfono móvil...

Una vez se tiene conectado el dispositivo con el cual se obtendrá la imagen en directo, se puede iniciar su visualización por medio del botón '*Visualizar Dispositivo*'. Habiendo hecho clic sobre el pulsador, aparece en '*Pantalla vídeo*' la imagen que captura el ordenador a través del dispositivo externo.



Figura 7.16 Ejemplo uso dispositivo externo

Con la imagen del vídeo en la interfaz, se puede proceder al siguiente paso: la captura de la imagen que se desea, que se produce a través del modo '*Aislado*'.

En cualquier momento, durante la visión del vídeo, el usuario puede interactuar y obtener una imagen que le interese. Este proceso se obtiene pulsando '*Capturar imagen*'. Para asegurarse de que la captura se obtuvo y almacenó satisfactoriamente aparecerá un cuadro de diálogo con el mensaje siguiente: Capturada.



Figura 7.17 Capturada

Una vez capturada la imagen el software continuará mostrando la señal de vídeo por si el usuario deseara realizar una segunda captura.

Con la imagen que se ha obtenido se llega al siguiente punto, el procesado. Accionando el botón '*Procesar*' (número 3 de la Figura 7.15), el sistema analizará la imagen capturada del vídeo como una imagen estática, pudiendo de esta forma obtener la matrícula del automóvil que se deseaba. El resultado de este proceso se mostrará en la sección de diálogo '*Matrícula*' (número 9 en la Figura 7.15), en el cual aparecerán los dígitos de la matrícula. En el caso de no poderse analizar la matricula en este diálogo aparecerá el mensaje 'NO LOCALIZADA'.

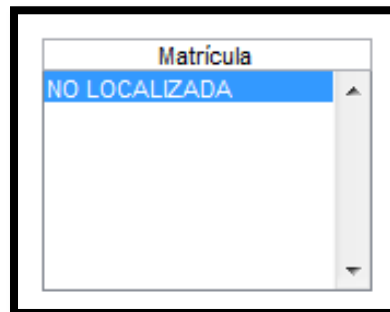


Figura 7.18 No localizada aislado

La segunda opción disponible haciendo uso de un dispositivo de captura de imágenes en directo es el modo '*Continuo*'. Este modo consta de dos pulsadores, '*Play*' y '*Stop*'. La utilidad de este modo se basa en poder analizar un vídeo en el momento en el que se está capturando (semejante a funcionamiento en tiempo real).

En primer lugar, para su uso, se debe pulsar el botón '*Play*' (numerado 4 en Figura 7.15), con esto se consigue que el software capture el primer frame y empiece su

procesado. Una vez termina de procesar el primer frame y de obtener, en el caso de existir, la matrícula, el sistema capturará el frame actual en ese momento y comenzará un nuevo análisis. Este proceso será repetitivo hasta que el usuario pulse el botón denominado 'Stop', es en ese momento cuando el sistema deja de procesar en modo continuo. Cada una de las matrículas localizadas en este proceso, en todos los frames que se analicen, aparecerán en el cuadro de diálogo 9 de la Figura 7.15, denominado 'Matrículas'.

OPCIÓN CARGAR VÍDEO.

En esta alternativa del programa se puede analizar un video grabado con anterioridad. Para ello, en primer lugar se cargará el vídeo que se desea analizar. Debe ser un vídeo en formato '.avi', ya que es el tipo de formato para el que ha sido diseñado el software. Por tanto, en el cuadro 'Cargar Video' se selecciona la opción 'Importar' (botón 6 de la Figura 7.15). Al pulsar se abrirá un explorador de archivos, al igual que sucedía en la opción de procesamiento de imágenes estáticas. Una vez localizado el video del que se desea obtener la información, se actúa sobre el botón 'Abrir' del explorador. Ahora el programa tiene el video cargado y está en disposición de empezar el proceso cuando el usuario lo solicite. Las siguientes capturas muestran los pasos realizados hasta llegar a este punto.

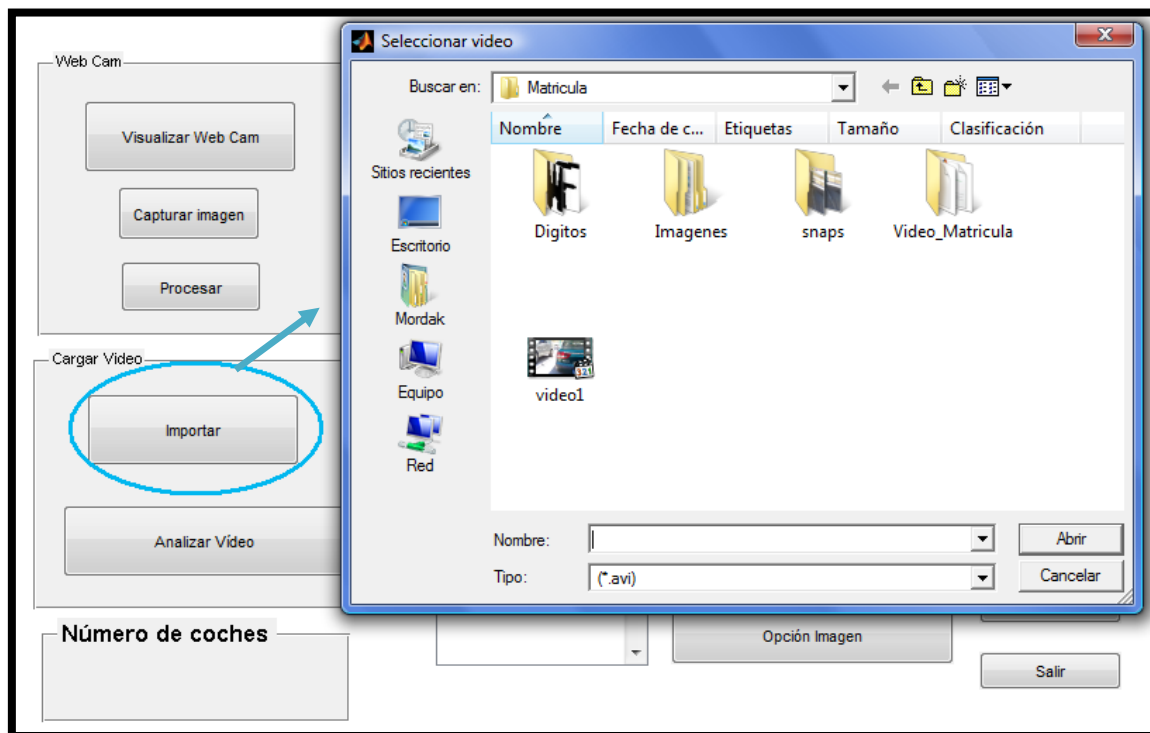


Figura 7.19 Importar vídeo y explorador de archivos

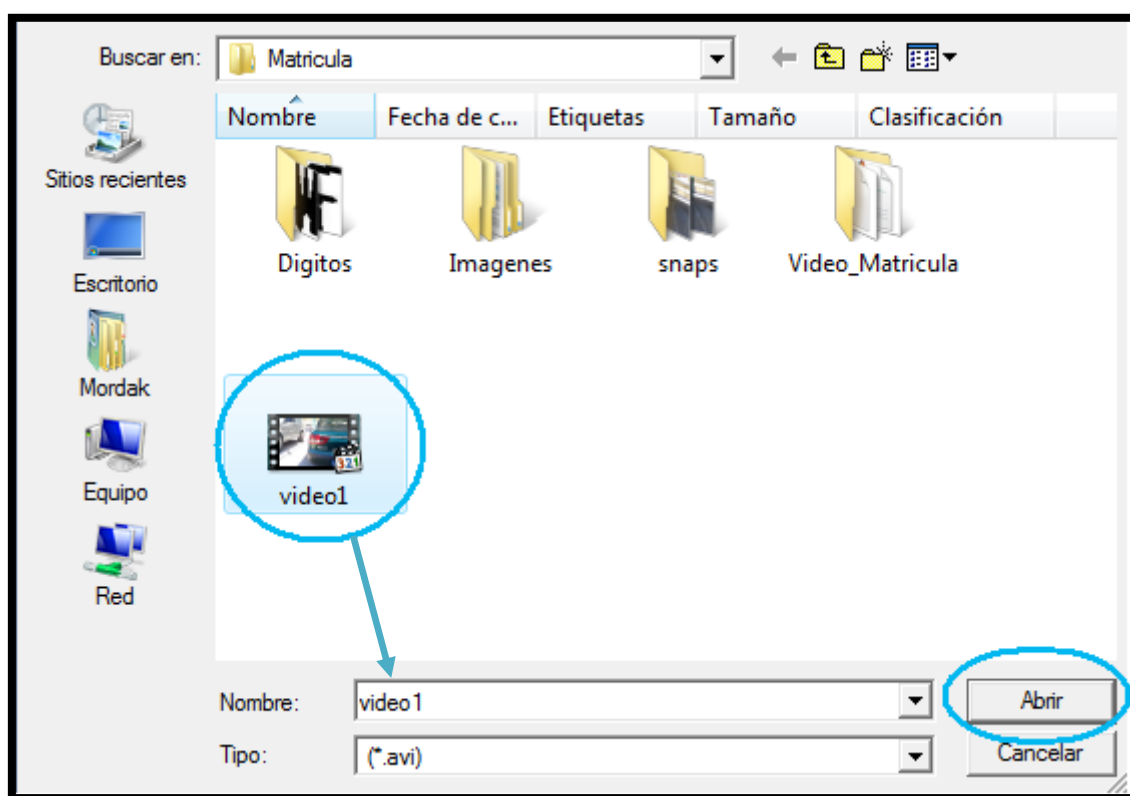


Figura 7.20 Selección del vídeo

Con el vídeo cargado y a disposición del software, se está en disposición de presionar el botón '*Analizar Video*' con el que comienza el proceso de localización de cada una de las matrículas y conteo del número total de vehículos que aparecen en el vídeo. A continuación se muestra un ejemplo del resultado de este proceso:

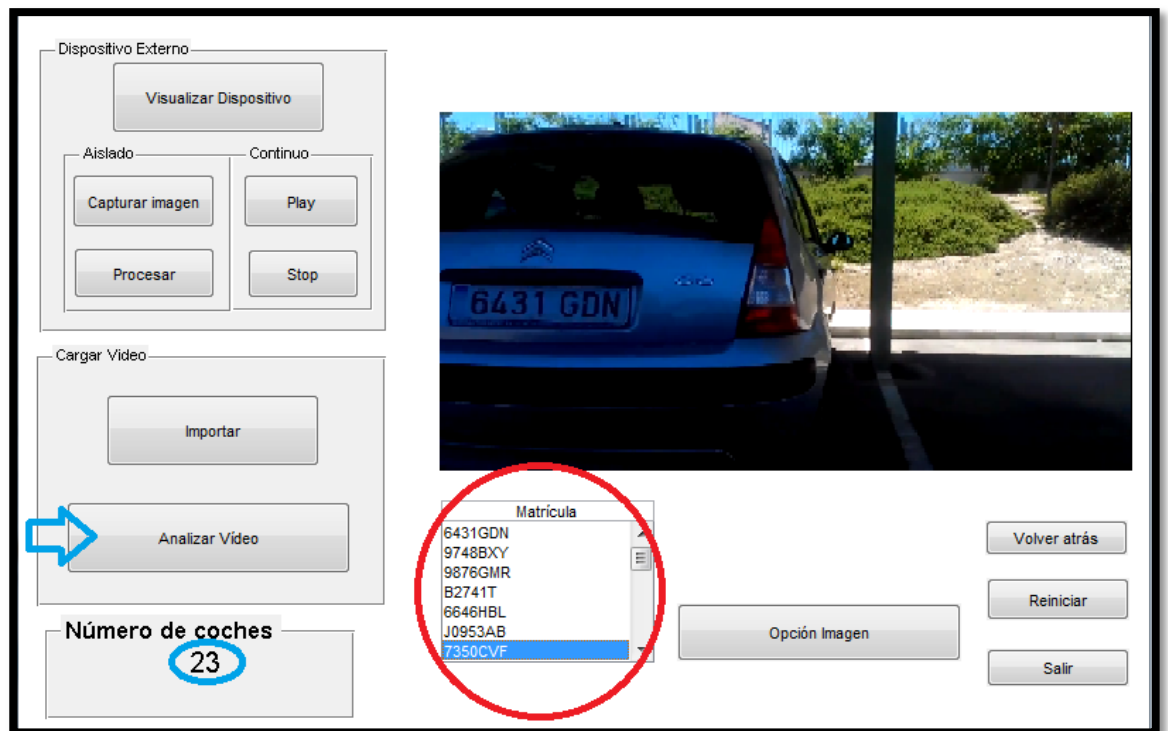


Figura 7.22 Final análisis video

En la Figura 7. anterior se pueden distinguir tres zonas señaladas. La flecha azul indica el pulsador que hace falta accionar para que comience el procesado, el número en el interior de la circunferencia azul enumera la cantidad de coches capturados y por último, el círculo rojo encierra las diferentes matrículas obtenidas durante el proceso.

7.4 Barra de progreso

Elemento de la interfaz gráfica que permite mostrar de forma gráfica el estado de avance de una tarea o proceso. Está acompañado de información en forma de texto para medir con exactitud el avance. Para cada una de las diferentes tareas que es capaz de realizar la aplicación existe una barra de progreso. Los mensajes que aparecen en ella se adecuan al proceso que se está realizando en el momento.

Mensajes para procesos de la interfaz de imágenes estáticas en modo automático:

- 'Preparando imagen...'.
- 'Calculando umbral...'.
- 'Proceso morfológico...'.
- 'Localizando área...'.

- 'Umbralizando área...'.
- 'Procesando morfológico del area...'.
- 'Localizando dígitos...'.
- 'Correlando dígitos...'.
- 'Comprobación validez matrícula...'.
- 'Modificación umbral y comprobación de matrícula...'.
- 'Modificación área y comprobación de matrícula...'.
- 'Proceso finalizado'.

Mensajes para procesos de la interfaz de imágenes estáticas en modo recorte:

- 'Iniciando análisis del recorte...'.
- 'Umbralizando recorte ...'.
- 'Localizando dígitos ...'.
- 'Correlando dígitos ...'.
- 'Comprobación de matrícula ...'.

Mensajes para proceso de la interfaz de vídeo en tiempo real:

- 'Iniciando dispositivo...'.
- 'Comenzando análisis de una nueva imagen...'.
- 'Preparando imagen...'.
- 'Localizando área...'.
- 'Procesando posibles dígitos...'.
- 'Correlación de dígitos y validez de matrícula...'.
- 'Modificación umbral...'.
- 'Modificación área...'.
- 'Finalizando proceso'.

Mensajes para capturar imagen a través de la interfaz de vídeo:

- 'Iniciando dispositivo...'.
- 'Capturando imagen...'.

- 'Almacenando imagen...'

Mensajes para procesar un vídeo de la interfaz de vídeo:

- 'Preparando vídeo...'
- 'Procesando frame número '...'.
- 'Preparando imagen...'
- 'Localizando área...'
- 'Procesando posibles dígitos...'
- 'Correlación de dígitos y validez de matrícula...'
- 'Modificación umbral...'
- 'Modificación área...'
- 'Comparación de las matrículas obtenidas'
- 'Análisis finalizado'

El aspecto de la barra se muestra en la siguiente imagen

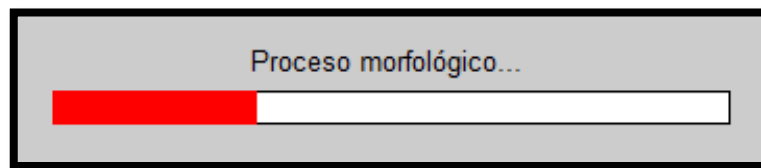


Figura 7.23 Barra de estado

Una vez la barra roja llega al final el proceso habrá finalizado.

8. Resultados y discusión

Esta sección se dedicará a analizar los distintos resultados obtenidos con la base de imágenes analizada y comentar cada uno de ellos. Se divide en resultados en tiempo y en porcentaje de imágenes.

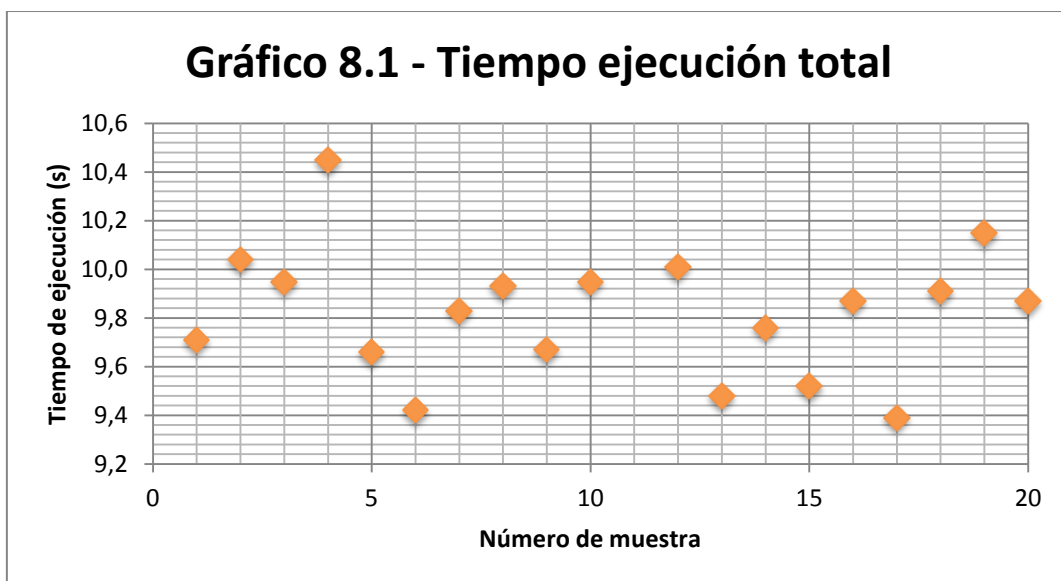
8.1 Resultados en función del tiempo de procesado.

Se comenzará analizando el tiempo invertido en el proceso total de análisis de imágenes estáticas de modo automático. Posteriormente se analizará paso por paso cada uno de las tareas que se realizan hasta finalizar el proceso, de tal forma se podrán descubrir los puntos críticos del proceso.

La siguiente tabla (8.1) muestra los valores numéricos que se han obtenido realizando el proceso completo del modo automático a 20 imágenes de muestra. Además, se adjunta gráfico con tales datos.

Imagen 1	Imagen 2	Imagen 3	Imagen 4	Imagen 5	Imagen 6	Imagen 7	Imagen 8	Imagen 9	Imagen 10
9.71	10.04	9.95	10.45	9.66	9.42	9.83	9.93	9.67	9.95
Imagen 11	Imagen 12	Imagen 13	Imagen 14	Imagen 15	Imagen 16	Imagen 17	Imagen 18	Imagen 19	Imagen 20
10.83	10.01	9.48	9.76	9.52	9.87	9.39	9.91	10.05	9.87

Tabla 8.1

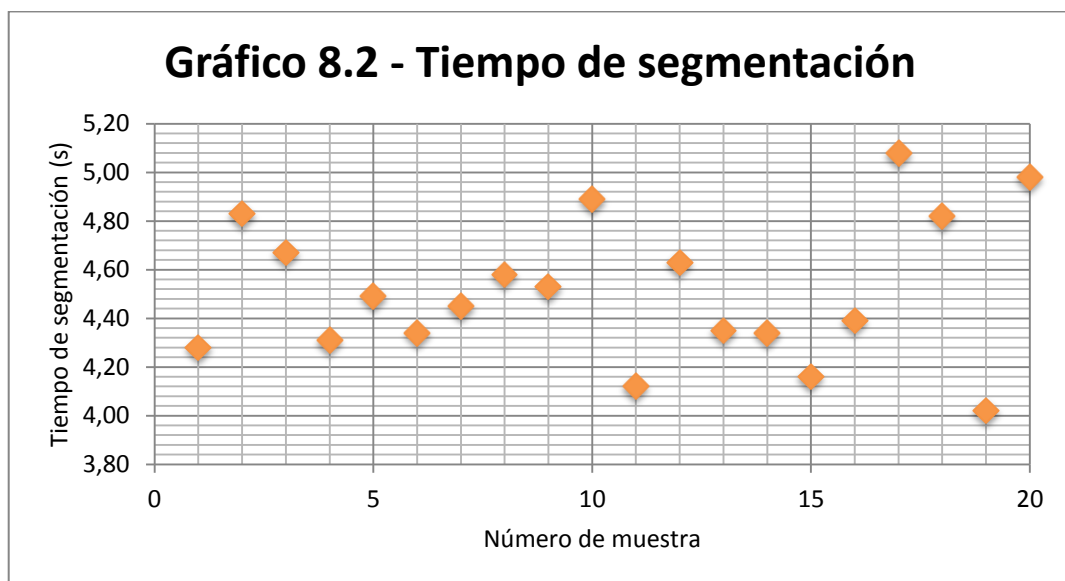


Observado los resultados obtenidos, se puede decir que el proceso necesita un tiempo de ejecución que se sitúa entre los 9 y 10 segundos aproximadamente. Por otro lado, existe un valor cercano a 11 segundos debido una mayor cantidad de áreas correladas por el sistema.

A continuación se desarrollan cada una de las diferentes tareas con las mismas 20 imágenes del estudio anterior. Se comenzará analizando el tiempo invertido por el programa en la segmentación de la matrícula, obteniendo los siguientes resultados:

Imagen 1	Imagen 2	Imagen 3	Imagen 4	Imagen 5	Imagen 6	Imagen 7	Imagen 8	Imagen 9	Imagen 10
4.28	4.83	4.67	4.31	4.49	4.34	4.45	4.58	4.53	4.89
Imagen 11	Imagen 12	Imagen 13	Imagen 14	Imagen 15	Imagen 16	Imagen 17	Imagen 18	Imagen 19	Imagen 20
4.12	4.63	4.35	4.34	4.16	4.39	5.08	4.82	4.02	4.98

Tabla 8.2



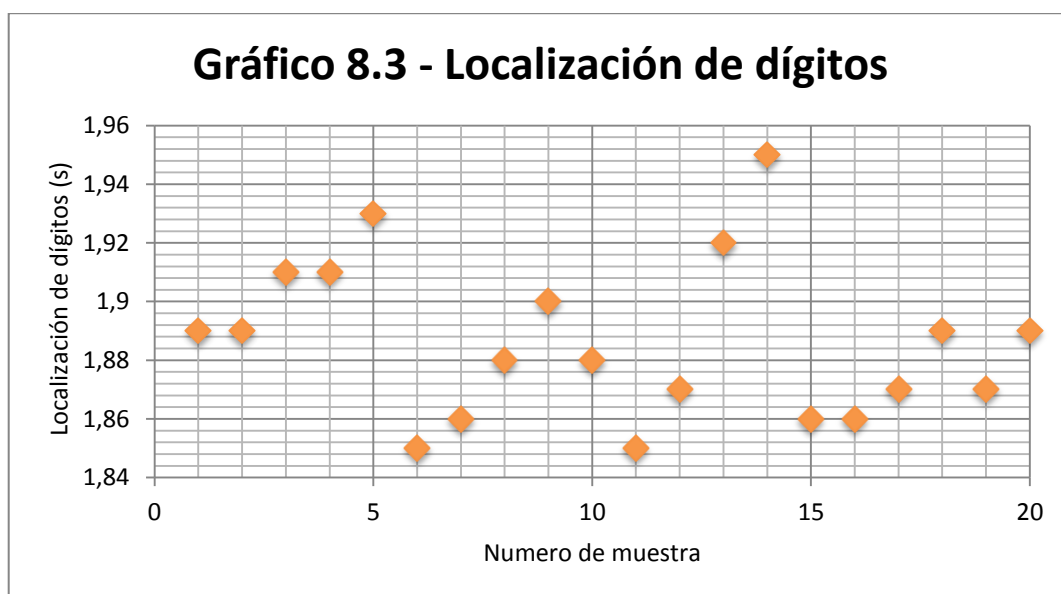
Analizando los resultados conseguidos tras la segmentación de la matrícula de 20 imágenes (gráfico 8.2 y tabla 8.2), se puede mencionar que los tiempos de ejecución se disponen entre 4 y 5 segundos. Este tiempo equivale aproximadamente a la mitad del tiempo total de ejecución visto en la tabla 8.1. Se podría decir que esta fase del proceso afecta en gran parte al tiempo de finalización del proceso total.

Tras analizar la segmentación de la placa de la imagen, se presentan los resultados conseguidos en la fase de localización de los diferentes dígitos. Esta fase, se recuerda que es la inmediatamente posterior a la segmentación de la matrícula.

La tabla 8.3 y el gráfico 8.3 muestran los tiempos empleados por el sistema en localizar los diferentes dígitos:

Imagen 1	Imagen 2	Imagen 3	Imagen 4	Imagen 5	Imagen 6	Imagen 7	Imagen 8	Imagen 9	Imagen 10
1,89	1,89	1,91	1,91	1,93	1,85	1,86	1,88	1,90	1,88
Imagen 11	Imagen 12	Imagen 13	Imagen 14	Imagen 15	Imagen 16	Imagen 17	Imagen 18	Imagen 19	Imagen 20
1,85	1,87	1,92	1,95	1,86	1,86	1,87	1,89	1,87	1,89

Tabla 8.3



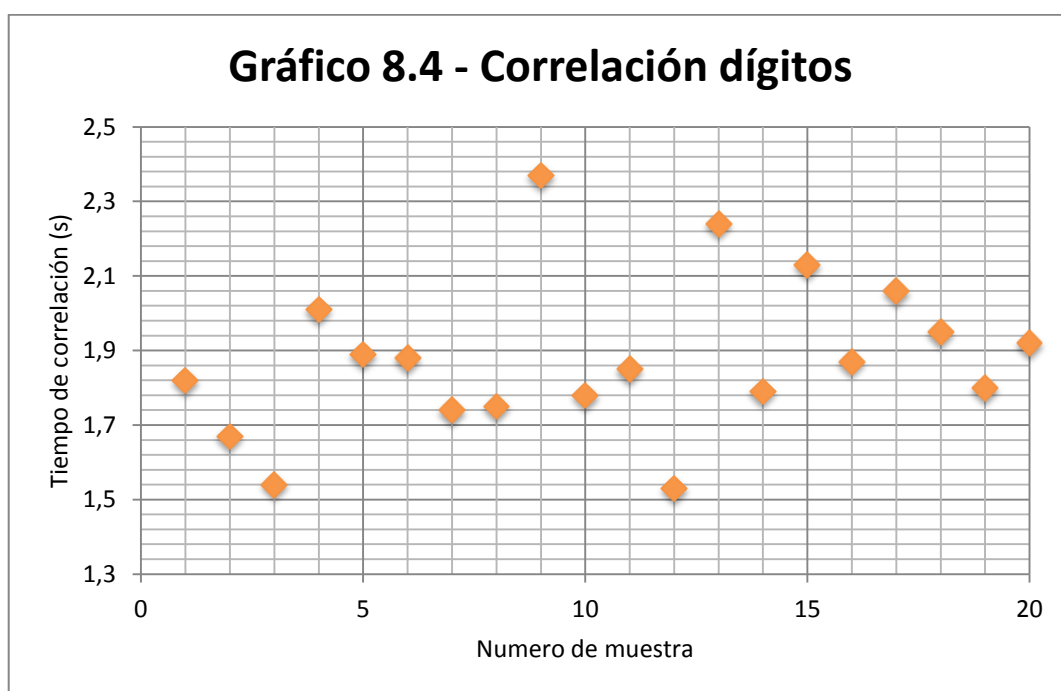
Todos los valores logrados mediante estas 20 muestras son muy parecidos, variando solo algunas centésimas de segundo, hecho que no es muy relevante.

Tras la localización de las diferentes áreas que componen la placa, el siguiente punto a analizar es la correlación de todos los dígitos con la base de imágenes alfanumérica.

Los resultados se pueden observar en la tabla 8.4 y el gráfico 8.4.

Imagen 1	Imagen 2	Imagen 3	Imagen 4	Imagen 5	Imagen 6	Imagen 7	Imagen 8	Imagen 9	Imagen 10
1.82	1.67	1.54	2.01	1.89	1.88	1.74	1.75	2.37	1.78
Imagen 11	Imagen 12	Imagen 13	Imagen 14	Imagen 15	Imagen 16	Imagen 17	Imagen 18	Imagen 19	Imagen 20
1.85	1.53	2.24	1.79	2.13	1.87	2.06	1.95	1.80	1.92

Tabla 8.4

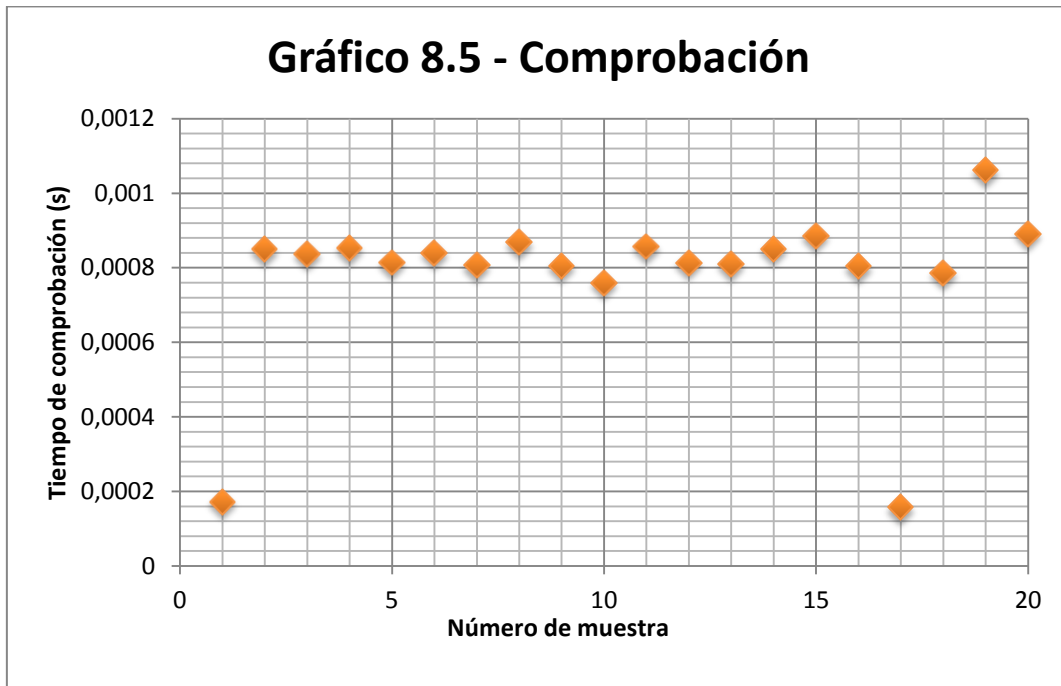


Los valores de tiempo, para la fase de correlación, se sitúan entre los 1 y 2 segundos. Solo existen 3 valores de las 20 muestras cuyo tiempo en segundos esté por encima de los 2 segundos al menos una décima de segundo. Esas excepciones son producto de un mayor número de correlaciones, debido a una cantidad superior de áreas, ya que aún estando fuera de la normalidad, son valores muy cercanos a los idóneos.

El último punto para el procesado de la imagen es la comprobación de la matrícula. Consiste en comparar la matrícula obtenida con los modelos actuales de diferentes composiciones alfanuméricas de dígitos. Este proceso es simplemente una consulta de los distintos tipos de matrícula y una comparación con la adquirida, por lo que los tiempos, como se verá en la tabla 8.5 no influyen en el tiempo total.

Tabla 8.5

Imagen 1	Imagen 2	Imagen 3	Imagen 4	Imagen 5	Imagen 6	Imagen 7	Imagen 8	Imagen 9	Imagen 10
0.000172	0.000849	0.000837	0.000853	0.000814	0.000840	0.000806	0.000868	0.000805	0.000760
Imagen 11	Imagen 12	Imagen 13	Imagen 14	Imagen 15	Imagen 16	Imagen 17	Imagen 18	Imagen 19	Imagen 20
0.000857	0.000813	0.00081	0.000850	0.000885	0.000804	0.000157	0.000786	0.001061	0.000891



Como se mencionó anteriormente, los valores de tiempo son prácticamente 0 segundos, por lo que este punto no es crítico en el desarrollo de la aplicación. Existen un par de valores muy por debajo de la media, pero al ser todos los valores tan pequeños no supone un problema.

El tiempo medio de comprobación es de 0.0007759 segundos

Como resultado de las diferentes comprobaciones, en cuanto a tiempos de ejecución desarrolladas, se puede llegar a la conclusión de que el punto crítico en el proceso es la segmentación de la matrícula. Esta tarea equivale prácticamente a la mitad del tiempo total de procesado que necesita el sistema para obtener los resultados perseguidos.

Los siguientes puntos que se analizarán serán los casos en los que el sistema se encuentre con un problema para terminar el proceso, pudiendo ser necesitado un cambio de umbral o una modificación del área que se está procesando.

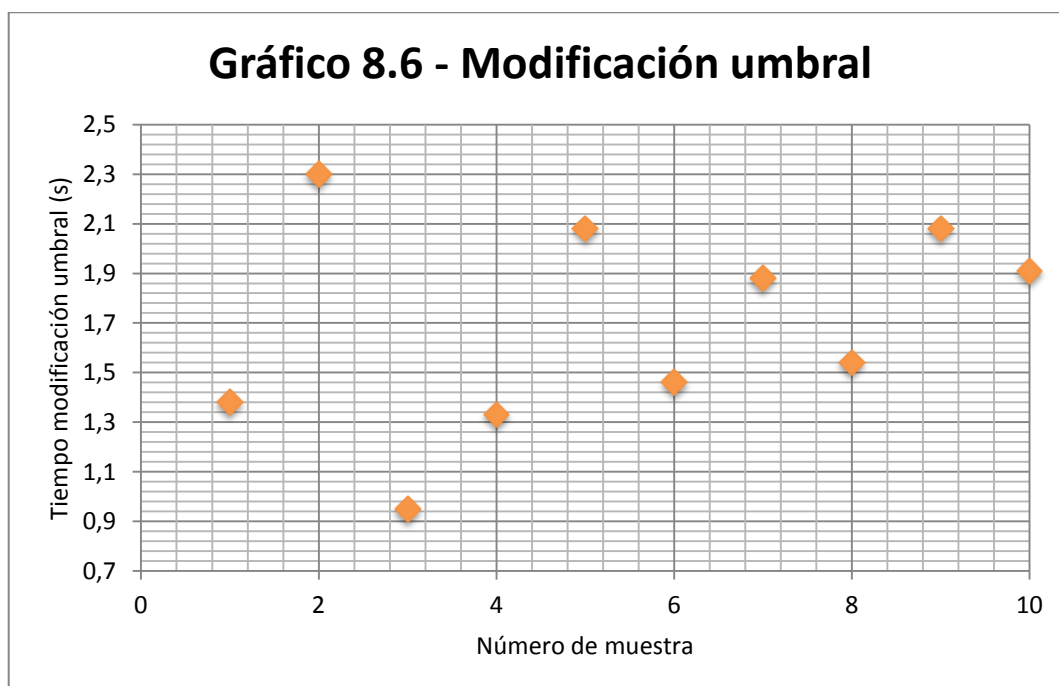
Lógicamente, estos procesos de solución de problemas supondrán un incremento en el tiempo de ejecución medio total, ya que se necesitará ese tiempo añadido para implementar las distintas operaciones que subsanen los inconvenientes mencionados.

Se comenzará desarrollando el tiempo invertido en el caso de necesitar uno o más cambios de umbral, para ello se dispondrá de varias imágenes de muestra para poder comparar resultados.

Los resultados que se muestran en la tabla 8.6 se producen para 10 imágenes, en las cuales se necesita de más de un cambio de umbral. Los tiempos que se presentan se corresponden con el tiempo invertido en el primer cambio de umbral, en el cual no se localizan los dígitos.

Imagen 1	Imagen 2	Imagen 3	Imagen 4	Imagen 5
1.38	2.3	0.95	1.33	2.08
Imagen 6	Imagen 7	Imagen 8	1.91	Imagen 10
1.46	1.88	1.54	2.08	1.91

Tabla 8.6

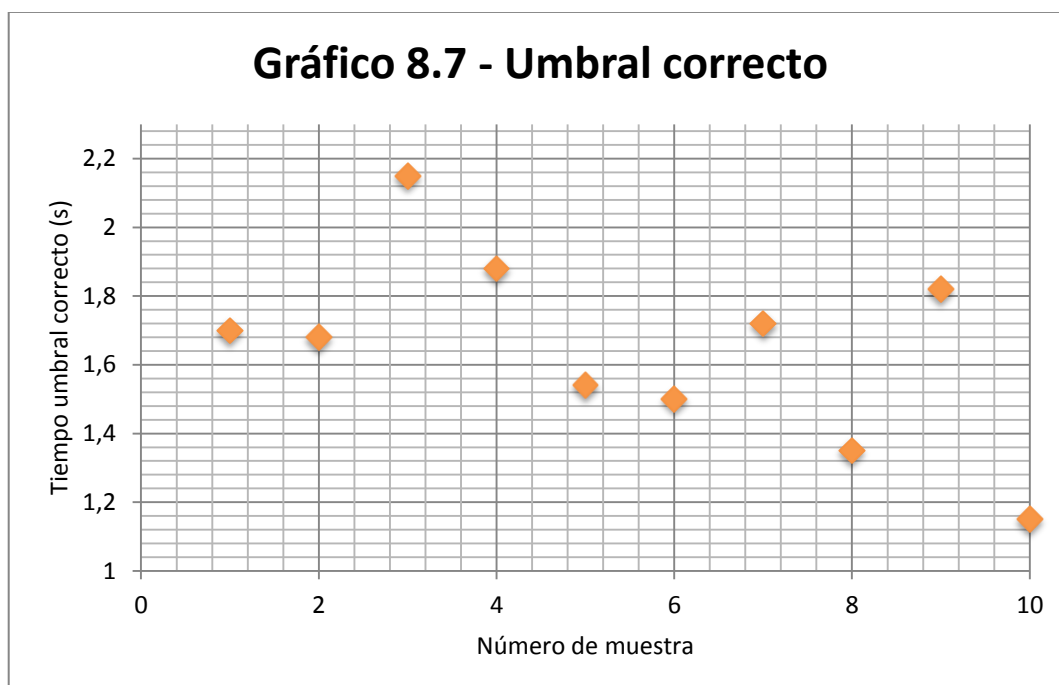


La diferencia de tiempos que se produce es debida a la cantidad de áreas que el sistema debe analizar al modificar el umbral, a mayor número de áreas mayor tiempo.

Ahora se expondrá el tiempo implementado por el software para el cambio de umbral en el que se obtiene la matrícula buscada.

Imagen 1	Imagen 2	Imagen 3	Imagen 4	Imagen 5
1.7	1.68	2.15	1.88	1.54
Imagen 6	Imagen 7	Imagen 8	Imagen 9	Imagen 10
1.5	1.72	1.35	1.82	1.15

Tabla 8.7



Las variaciones de tiempo en las diferentes muestras se deben a la cantidad de áreas (localizadas como dígitos) que se deben procesar y correlar.

El tiempo total del programa para conseguir los resultados perseguidos para este tipo de imágenes con errores en el umbral se compone por: la suma de las distintas iteraciones de cambio de umbral sin resultado, más la suma del proceso en el umbral correcto y el tiempo total (similar a los de la tabla 8.1).

Se expone a continuación un ejemplo de imagen en la que se producen 2 iteraciones de cambio de umbral:

Primera modificación de umbral: 1.38 segundos.

Segunda modificación de umbral: 1.52 segundos.

Tiempo total del procesado: 14.03 segundos.

Tiempo añadido por el error: 2.9 segundos.

Tiempo sin las modificaciones de umbral: 11.13 segundos, tiempo que se asemeja al tiempo medio de análisis obtenido en la tabla 8.1.

En relación con el segundo problema que puede aparecer, basado en el cambio de área, se expone en la siguiente tabla los valores (de tiempo total de proceso) de un conjunto de muestras sobre las cuales se ha producido al menos una modificación de la primera área seleccionada.

Imagen 1	Imagen 2	Imagen 3	Imagen 4	Imagen 5
39.7	45.67	60.51	42.18	49.85

Tabla 8.8



La diferencia de tiempos entre las diferentes muestras se corresponde al diferente número de ocasiones en las que el programa ha cambiado de área.

Son valores muy elevados debido a dos factores:

- El primero factor se debe a que la aplicación está programada para cambiar de umbral antes de modificar el área, por lo que para llegar a cambiar de área, se debe intentar corregir el cambio de umbral con anterioridad. Este hecho implica un aumento del tiempo total de ejecución.
- El segundo factor es producido de forma directa por el hecho de modificar el área y comenzar el proceso de nuevo (incluido las 5 iteraciones de cambio de umbral por cada área implicada en el proceso).

En lo referido al vídeo, en el ámbito del tiempo invertido, se debe comentar que un vídeo está formado por imágenes y por tanto el tiempo empleado en analizar un vídeo vendrá directamente determinado por el número de fotogramas que lo componen.

Los siguientes datos muestran el tiempo empleado en analizar un vídeo de 100 frames: 2269.204 segundos totales, que equivale a 22.69 segundos/frame.

8.2 Resultados en función del porcentaje de imágenes

Imágenes

Este apartado del capítulo se dedicará a comentar distintos porcentajes de imágenes de la base de datos basados en diferentes principios.

Se comenzará exponiendo, en general, el porcentaje de imágenes de las cuales se obtiene de forma correcta el objetivo del proyecto.

Total de imágenes	Correctas	Porcentaje
150	130	86.6 %

Tabla 8.9

El sistema es capaz de obtener el resultado de forma correcta de 130 fotos de un total de 150. Esto hace un porcentaje del 86.6%, un porcentaje bastante aceptable y fiable.

Las imágenes erróneas se pueden dividir en dos tipos de problemas. El primero de ellos es el error al localizar y correlar alguno de los dígitos de la matrícula. Este hecho se puede producir por alguna mancha en la placa, debida por ejemplo a barro, o por haberse desgastado por el tiempo. La siguiente imagen muestra una placa afectada por el paso del tiempo:



Figura 8.1 Matrícula desgastada

Se aprecia como la 'S' final contiene un leve desconchón que produce que sea irreconocible por el sistema.

Imágenes en las que falla algún dígito:

Total imágenes erróneas	Error dígito	Porcentaje
20	8	40 %

Tabla 8.10

El porcentaje sobre las imágenes erróneas es de un 40 %, siendo un 5.33% sobre el total de imágenes almacenadas en la base de datos.

El otro problema, que provoca que la aplicación no pueda llegar su fin último, es debido a no poder segmentar la zona de la matrícula, ya sea por situarse la imagen en un

ambiente oscuro o por otros factores como por ejemplo una sombra sobre los dígitos. La figura 8.2 y 8.3 muestran dos ejemplos:



Figura 8.2 Problema de oscuridad

En la figura 8.2, debido a los tonos oscuros que posee la foto y al pequeño reflejo sobre la matrícula, el sistema no es capaz de reconocer la matrícula y sus dígitos.



Figura 8.3 Problema de sombra

En la figura 8.3, los dígitos de la matrícula están bajo una sombra, la aplicación no es capaz de separarlo. Tras las distintas modificaciones de umbral, el programa da el proceso por concluido sin resultado.

A este tipo de imágenes se corresponde el resto de imágenes erróneas. La siguiente tabla muestra su porcentaje correspondiente.

Total imágenes erróneas	Otros errores	Porcentaje
20	12	60 %

Tabla 8.11

El porcentaje sobre las imágenes erróneas es de un 60 %, siendo un 8% sobre el total de imágenes almacenadas en la base de datos.

El modo recorte permite poder obtener la solución de parte de esas imágenes, las cuales no se podían procesar por el modo automático. La siguiente tabla ofrece el porcentaje de imágenes erróneas procesadas correctamente mediante recorte.

Total imágenes erróneas	Solución mediante recorte	Porcentaje
20	15	75 %

Tabla 8.12

Por tanto, para finalizar, si se unen los resultados positivos de ambos sistemas (automático y recorte) se obtienen un total de resultados positivos del 96.6%.

Total imágenes	Total solución correcta	Porcentaje
150	145	96.6%

Tabla 8.13

El siguiente gráfico se ofrece una visión de los porcentajes según el resultado obtenido al procesar las imágenes:

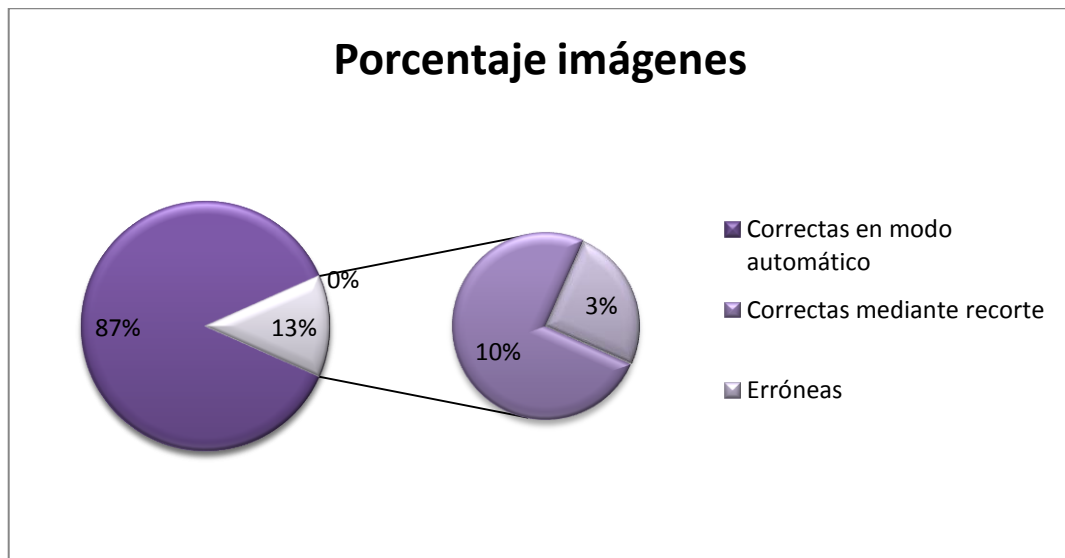


Gráfico 8.9 Porcentajes de imágenes correctamente procesadas

Videos

En el apartado r relacionado con las estadísticas obtenidas del modo vídeo, se destacará el porcentaje de coches localizados y la cantidad de matrículas correctas.

La siguiente tabla muestra el porcentaje de coches localizados sobre un vídeo del total:

Total	Localizados	Porcentaje
26	24	92.3%

Tabla 8.14



Gráfico 8.10 - Número de coches

En la tabla 8.15 se expone el total de matrículas que han sido localizadas y procesadas de forma correcta, es decir, obteniendo todos los dígitos correspondientes a la placa:

Total	Correctas	Porcentaje
26	22	84.61 %

Tabla 8.15

Comparando la tabla 8.14 y 8.15 se observa que el porcentaje de matriculas correctas es menor a la cantidad de coches capturados. Este suceso es consecuencia del error en la localización y correlación de algún dígito de la placa.

8.3 Relación: tiempo de procesamiento aciertos.

Existe un vínculo entre el tiempo que el sistema tarda en finalizar todo el procedimiento y la probabilidad de que se obtengan los resultados perseguidos.

Según se haga uso de más tiempo para finalizar el método de localización y reconocimiento de matrículas, las probabilidades de obtener con éxito el fin del proyecto se ven disminuidas.

Sin embargo, también se puede dar el caso de terminar todo el proceso en un corto plazo de tiempo (unos 10 segundos) y no haber conseguido el propósito de la aplicación. Este hecho se puede dar en el caso en el que el sistema localice una única área en la segmentación de la imagen, y esa área no se corresponda con la que contiene la matrícula. También puede haberse segmentado de forma correcta la matrícula, pero que en esta exista algún dígito que no está bien definido, ya sea por desgaste, un reflejo, una sombra...

9. Conclusiones

El presente capítulo ha sido destinado a desplegar las distintas resoluciones a las que se ha llegado tras los datos presentados en el apartado 8, denominado '*Resultados*'.

Una vez se han analizado los distintos resultados, se pueden exponer las conclusiones obtenidas.

En lo respectivo al análisis de imágenes estáticas, se han presentado unas estadísticas bastante positivas. El hecho de poder resolver de forma automática, sin necesidad de la interacción del usuario, hasta un 87% de las imágenes de la base de datos es un gran porcentaje.

La nota negativa reside en el tiempo de procesado que necesita el sistema para conseguir los resultados perseguidos. Este tiempo se podría ver reducido si el dispositivo de procesado estuviese en posesión de un hardware de mayor potencia que el ordenador sobre el que se han realizado las pruebas (consultar Anexo I).

Además, cualquier fotograma sobre el que se necesite una modificación del umbral, verá como se produce un incremento en su tiempo de ejecución, superando la media.

En adición, si una imagen necesita el análisis de distintas áreas, el tiempo de ejecución podría aproximarse a 1 minuto si se realizasen hasta 4 cambios de zona de análisis.

Por tanto, el principal problema que plantea el software es el tiempo que requiere para procesar la imagen en el caso que necesite alguna operación extra. Se define operación extra como variación de umbral y cambio de área.

El sistema auxiliar de recorte permite la resolución de la mayoría de las imágenes, cuyos resultados en el modo automático eran erróneos. Gracias a este sistema, el usuario está en disposición de poder obtener un resultado correcto, ya que el modo automático junto al modo recorte consiguen un porcentaje de garantías, alrededor de un 97% de acierto.

Por lo que la aplicación permite localizar la mayor parte de las matrículas que aparezcan en imágenes que sigan las características definidas en el Anexo II.

En lo referente al análisis de los resultados obtenidos de la sección de vídeos, cabe destacar el gran porcentaje de acierto en localización de coches con sus correspondientes matrículas. Este proceso podría ser interesante para el control de vehículos en un parking, gracias a su elevado porcentaje de acierto. Las matrículas que no fueron localizadas durante el estudio del vídeo, no se percibieron por reflejos en la placa o por estar en defectuoso estado.

Al igual que en el análisis de imágenes, el tiempo de procesado de un vídeo se vería reducido si se hiciera uso de un ordenador con mayor potencia que el usado para las diferentes pruebas realizadas (consultar características en Anexo I).

10. Líneas futuras

Se comentan en este capítulo diferentes métodos para mejorar el sistema. Las mejoras que se proponen están relacionadas con el tiempo de ejecución para que se realice el procesado y con el tipo de imágenes que se pueden analizar.

Respecto al primer punto, una línea futura puede consistir en intentar reducir los tiempos de segmentación de matrícula y localización de dígitos, intentando eliminar parte de los bucles que estas fases contienen. De esta forma, el sistema podría ser más eficiente en el aspecto temporal.

Un cambio en el tiempo de segmentación y localización de dígitos en imágenes se podría implementar en la sección de vídeo, reduciendo también los tiempos de análisis de los distintos frames. En consecuencia, se reduciría el tiempo total de proceso de un vídeo completo.

El segundo problema a resolver como línea futura sería las restricciones definidas en el Anexo II, referidas a las características necesarias para vídeos e imágenes. Una mejora debería ser el poder procesar matrículas inclinadas o en ambientes oscuros de forma automática, sin necesidad del modo recorte.

Anexo I

Este anexo contiene las herramientas necesarias para poder hacer uso de la aplicación descrita en la presente memoria.

Software Matlab.

Matlab es el programa elegido para el desarrollo del proyecto. Se trata de una herramienta de una gran potencia y con gran utilidad para el procesado de este tipo de trabajos basados en imágenes, ya que contiene definidas e implementadas numerosas funciones para el tratado de imágenes. Esas funciones hacen alusión a procesos morfológicos como erosión, dilatación, apertura... además de ser un software con una gran capacidad para cálculos matemáticos.

Entre sus prestaciones básicas se hallan: la manipulación de matrices, la representación de datos y funciones, la implementación de algoritmos, la creación de interfaces de usuario (GUI) y la comunicación con programas en otros lenguajes y con otros dispositivos hardware. El paquete MATLAB dispone de dos herramientas adicionales que expanden sus prestaciones, a saber, Simulink (plataforma de simulación multidominio) y GUIDE (editor de interfaces de usuario - GUI). Además, se pueden ampliar las capacidades de MATLAB con las cajas de herramientas (toolboxes); y las de Simulink con los paquetes de bloques (blocksets).

Ordenador.

Dispositivo electrónico sobre el que se ejecutará la aplicación. El desarrollo del proyecto se ha producido en un ordenador con las siguientes características que determinan el rendimiento del software:

- Procesador: Intel Core i5-450M (2.4 GHz, 3MB L3 cache).
- Memoria RAM: 4 GB DDR3.
- Sistema operativo: Windows 7.

Un ordenador con mayor potencia, desde el punto de vista del hardware, obtendría mejores resultados en cuanto al rendimiento expuesto en el capítulo 8 de la presente memoria.

Dispositivo para captura de imágenes.

Puede ser cualquier dispositivo que permita la captura de imágenes para su posterior procesado. Debido a que la aplicación está preparada para procesar imágenes de video en tiempo real, es recomendable disponer de algún dispositivo que proporcione imágenes externas de video en directo, como puede ser una webcam propia de un ordenador o una cámara de vídeo.

Se recomienda que el dispositivo disponga de una resolución aceptable, definiendo este término como una resolución con valores que rondan los 1280x768. Para una resolución menor el sistema podría tener problemas para conseguir su objetivo.

Anexo II

Este anexo tiene como objeto orientar al usuario sobre las características óptimas que debería poseer la imagen o el video que se quiere procesar.

Se comenzará exponiendo las necesidades para un correcto funcionamiento del sistema en modo de análisis de imagen estática. La imagen estática debe poseer al menos una resolución en píxeles mínima de 440x330, es decir, 145200 píxeles. Este límite es debido a que para tamaños inferiores la imagen, al ser procesada, se vería ‘pixelada’ y por tanto no se llevaría a cabo una correcta localización de la matrícula. El tamaño óptimo de funcionamiento es de 1280x768, aunque para tamaños superiores no existiría ningún problema. Además de la necesidad de un tamaño en píxeles adecuado, la imagen de la matrícula no debe estar muy inclinada, debido a que el sistema ha sido diseñado para fotografías con placas prácticamente rectas.

Otra de las necesidades para obtener resultados coherentes es la posición de la imagen respecto a la fotografía, es decir, la matrícula no debe estar situada en los bordes de la captura, en el caso de estar situada cerca de uno de los límites, el programa al recortar la imagen (véase apartado 6.1.1) eliminaría también parte de la imagen con información relevante.

En cuanto al formato de la imagen, no es un factor demasiado decisivo en cuanto a resultados, aunque se ha desarrollado la aplicación para hacer uso de imágenes en formato jpg, ya que es un tipo de imagen con poco peso y de fácil manipulación.

En lo referente a características de vídeo, se ha de mencionar que el sistema se ha creado para manipular vídeos en formato .avi. Además la resolución del vídeo debe de ser al menos de 1280x768 píxeles como las imágenes estáticas. De esta forma se tendrá una calidad buena de vídeo para el objetivo final del programa.

En cuanto al número de frames por segundo se recomienda una capacidad de 30 fotogramas/segundo mínima.

El tiempo de duración del vídeo solo será un problema si el ordenador que lo procese no tiene suficiente capacidad y no quede suficiente memoria para terminar de analizarlo.

Anexo III

Para poder llevar a cabo la correlación, el sistema está en posesión de un conjunto de imágenes que representan cada uno de los caracteres que pueden aparecer en una matrícula. Las siguientes tablas muestra el número de prototipos para cada número y letra:

Número	Cantidad de prototipos distintos
0	6
1	2
2	1
3	3
4	2
5	4
6	4
7	2
8	3
9	2

Letra	Cantidad de prototipos distintos
A	2
B	5
C	3
D	3
E	1
F	2
G	1
H	1
I	1
J	1
K	1
L	1
M	3
N	1
P	1
R	1
S	2
T	1
U	1
V	1
W	3
X	1
Y	1
Z	1

Existen dígitos repetidos varias veces debido a que esos valores producen conflictos y confusión al programa. Por lo que, a mayor número de muestras distintas, mayor probabilidad de acertar el carácter que se está analizando.

Referencias bibliográficas

<https://sites.google.com/site/ticvalcarcel/optimizacion-de-imagenes-para-internet/tipos-de-imagenes-y-formatos>

<http://www.ite.educacion.es/formacion/materiales/107/cd/imagen/imagen0105.html>

<http://www.buenastareas.com/ensayos/Imagen-Analogica-e-Imagen-Digital/561948.html>

http://gusgsm.com/eps_es_malo

<http://es.wikipedia.org/wiki/RGB>

http://www.ecured.cu/index.php/Modelo_RGB

<http://es.kioskea.net/contents/724-el-formato-tif7>

<http://es.wikipedia.org/wiki/V%C3%ADdeo>

<http://www.ite.educacion.es/formacion/materiales/107/cd/video/video0101.html>

<http://www.informatica-hoy.com.ar/multimedia/Todos-los-formatos-de-video.php>

http://centrodeartigos.com/articulos-utiles/article_107677.html

<http://www.anpr.net/>

http://www.psico.uniovi.es/dpto_psicologia/metodos/tutor.6/fcope.html

http://es.wikipedia.org/wiki/M%C3%A1todo_del_valor_umbral

<http://mathforum.org/kb/message.jspa?messageID=6815312>

<http://alojamientos.us.es/gtocom/pid/tema5-1.pdf>

John C. Russ, *Image Processing Handbook*, CRC Press, 2011.

William K Pratt, *Digital Image Processing*, John Wiley & Sons Inc, 2007.

Rafael González, Richard Woods and Steve Eddins, *Tratamiento Digital de Imágenes*, Ediciones Díaz de Santos S.A., 1996.

Matlab Images Toolbox.