

## Aspectos de un Sistema de Visión Artificial

Este apunte tiene como objetivo presentar brevemente los pasos fundamentales de un proyecto de visión artificial.

La visión artificial es una técnica basada en la adquisición de imágenes, generalmente en dos dimensiones, para luego procesarlas digitalmente mediante algún tipo de CPU (computadora, microcontrolador, DSP, etc), con el fin de extraer y medir determinadas propiedades de la imágenes adquiridas. Se trata, por tanto, de una tecnología que combina las computadoras con las cámaras de video para adquirir, analizar e interpretar imágenes de una forma equivalente a la inspección visual humana.

Actualmente se aplica en diversos procesos científicos y militares, extendiéndose su uso además, en un amplio rango de sectores industriales para la automatización de tareas anteriormente reservadas para la inspección visual humana. Este tipo de incremento de utilización a nivel industrial es debido a que se trata de una tecnología especialmente útil en labores de inspección o supervisión, siendo cuantitativamente más objetivos y consistentes que la inspección humana. Un segundo aspecto complementario de estas técnicas es la automatización derivada de esta inspección y la posibilidad de actuar sobre el proceso de fabricación o inspección modificando parámetros de la máquina de producción. Las técnicas de visión artificial, como demuestra su gran uso en la industria, son particularmente apropiadas para la realización de trabajos visuales altamente repetitivos que sean fatigosos o difíciles de realizar para un operario, especialmente cuando este trabajo es ineficiente o costoso en términos económicos o temporales.

Un sistema de visión artificial se compone básicamente de los siguientes elementos:

- Fuente de luz: es un aspecto de vital importancia ya que debe proporcionar unas condiciones de iluminación uniformes e independientes del entorno, facilitando además, si es posible, la extracción de los rasgos de interés para una determinada aplicación.
- Sensor de imagen: es el encargado de recoger las características del objeto bajo estudio.
- Tarjeta de captura o adquisición de imágenes: es la interfaz entre el sensor y la computadora o módulo de proceso que permite al mismo disponer de la información capturada por el sensor de imagen.
- Algoritmos de análisis de imagen: es la parte inteligente del sistema. Su misión consiste en aplicar las transformaciones necesarias y extracciones de información de las imágenes capturadas, con el fin de obtener los resultados para los que haya sido diseñado.
- Computadora o módulo de proceso: es el sistema que analiza las imágenes recibidas por el sensor para extraer la información de interés en cada uno de los casos implementando y ejecutando los algoritmos diseñados para la obtención de los objetivos.

- Sistema de respuesta en tiempo real: con la información extraída, los sistemas de visión artificial pueden tomar decisiones que afecten al sistema productivo con el fin de mejorar la calidad global de producción.

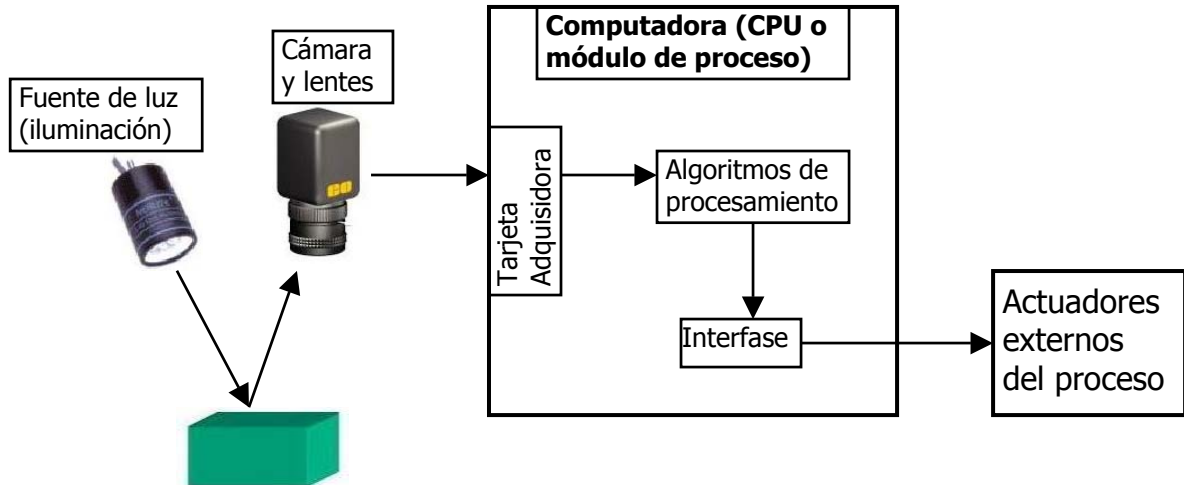


Figura 1: Elementos de un sistema de visión artificial.

## Fuentes de luz

La fuente de luz es un factor de vital importancia en los sistemas de visión artificial y afectan de forma crucial a los algoritmos de visión que se vayan a utilizar bajo esas condiciones. Una iluminación arbitraria del entorno no suele ser aceptable ya que muy probablemente se obtengan imágenes con un bajo contraste, reflexiones especulares, además de sombras y detalles espurios. Los objetivos de una iluminación adecuada son, básicamente, independizar las condiciones del entorno y resaltar los rasgos de interés de una determinada aplicación.

### Tipos de iluminación

A continuación sólo se enuncian los diferentes tipos de iluminación. En el apunte “**Tipos de Iluminación.pdf**” se pueden encontrar los detalles acerca de las técnicas de iluminación más comunes y generalmente usadas.

- Iluminación posterior (backlight).
- Iluminación Frontal Oblicua y Direccional.
- Iluminación Frontal Axial (difusa).
- Iluminación de Día Nublado (cloudy day illumination, CDI).
- Iluminación de Campo Oscuro (dark field).
- Array de luces.
- Iluminación de campo claro (o brillante).

## Sensores de imagen

Los sensores de imagen son componentes sensibles a la luz que modifican su señal eléctrica en función de la intensidad luminosa que perciben. La tecnología más habitual en este tipo de sensores es el **CCD** (*charge coupled devices* o dispositivos de acoplamiento de carga) donde se integra en un mismo chip los elementos fotosensibles y el conjunto de puertas lógicas y circuitería de control asociada. En éstos, la señal eléctrica que transmiten los fotodiodos es función de la intensidad luminosa que reciben, su espectro, y el tiempo de integración (tiempo durante el cual los fotodiodos son sensibles a la luz incidente). Otra tecnología que está empezando a extenderse son los sensores **CMOS** (*complementary metal oxide semiconductor*) dada las ventajas de éstos sobre los CCD, y la reducción de precios de estos dispositivos. En cuanto al *rango dinámico* (rango de amplitud entre los valores máximos y mínimos que un elemento puede producir, medido en decibeles) se pasa de los 70dB de los sensores CCD a los 120dB de los sensores CMOS, valor más cercano a los 200dB del ojo humano, lo que facilita la autoadaptación en el propio chip al brillo existente en el entorno. El sensor CMOS puede amplificar de forma individual la respuesta de los píxeles en función de sus respectivas condiciones de iluminación. El acceso multipíxel permite configurar individualmente distintos intervalos de tiempo de integración, mejorando la calidad de imagen mediante la selección del mejor tiempo de exposición de cada píxel.

### Resolución

Existen diferentes arquitecturas de sensores. En primer lugar están los *sensores lineales*. En éstos, el sensor es una línea de fotodiodos. Esta arquitectura permite la utilización de sensores de 1x1024, 1x2048, 1x4096, e incluso 1x6000 píxeles, lo que la hace muy adecuada para trabajar con altas resoluciones sobre superficies en movimiento. Para condiciones de iluminación muy exigentes o velocidades de trabajo muy altas existe la posibilidad del uso de sensores **TDI** (*time delay integrated*). Esta tecnología consiste en el uso de varias líneas de captura sobre la misma línea del objeto, con el fin de sumar su carga y obtener así una mayor sensibilidad. En segundo lugar están los *sensores de área*. Estos alcanzan resoluciones habituales de 1024x1024, aunque existen en el mercado algunas casas que disponen de cámaras especiales con resoluciones de hasta 3072x2048. En este caso existen tecnologías de adquisición de imágenes, entrelazada y no entrelazada. El método entrelazado captura las líneas pares e impares que forman una imagen en instantes de tiempo diferentes. La tecnología de no entrelazado (*progressive scan*) captura todas las líneas en el mismo instante de tiempo. Es más costoso económicamente, pero indispensable para trabajar con objetos en movimiento.

### Cuantización

Independientemente de la arquitectura utilizada, existen otros aspectos importantes en estos sensores. La cuantización (conversión analógica-digital) determina el número de bits usados para representar la información capturada. Por ejemplo, usando un sistema blanco y negro de 8 bits tenemos 256 niveles diferentes mientras que, usando un sistema de 10 bits, obtendríamos 1024 niveles de gris diferentes, lo que permite una mayor definición.

La evolución de las capacidades y características técnicas de estos dispositivos ha seguido una tendencia muy favorable, potenciado en parte por el mayor uso industrial de estos sistemas. Por tanto, en el mercado existen hoy en día numerosos dispositivos para a-

bordar cualquier tipo de aplicación, ya sea en blanco y negro o color, a las velocidades requeridas en los procesos de producción.

Entre los avances más significativos, además de la utilización de puertos de comunicación cada vez más potentes y fiables, está una capacidad de proceso más o menos relevante e integrada en la propia cámara. Por ejemplo, la cámara lineal OPSIS 5150ALC dispone de FPGA (*Field-programmable gate array*, matriz de puertas programable por campo) programable y de un microprocesador Motorola Power PC con sistema operativo en tiempo real VxWorks que permite realizar procesamiento de imágenes. De esta forma se puede transmitir la imagen, la imagen preprocesada, o los resultados del procesamiento de la imagen a través de los puertos de comunicación de que dispone la cámara, haciendo innecesario el uso de tarjetas de captura.

## Tarjetas de captura (o adquisidora)

Las tarjetas de captura de imagen permiten transferir la imagen de la cámara a la memoria de la computadora con el fin de que ésta pueda realizar el procesamiento adecuado a las imágenes.

Aspectos importantes en la adquisidora: velocidad de transmisión, el formato de los datos, la profundidad de píxel, la capacidad de captura por disparo, la capacidad de preprocesado de la imagen, la velocidad de transferencia de la imagen de la memoria de la tarjeta a la memoria de la computadora, el reset asíncrono o la capacidad de controlar parámetros de la cámara en tiempo real. Evidentemente, todos los datos relativos a velocidades de transmisión son especialmente importantes para poder trabajar en tiempo real. Este aspecto ha mejorado notablemente en los últimos años. Otro parámetro que ha evolucionado muy favorablemente es el formato de los datos. Con la estandarización definitiva de los formatos EIA-644 o LVDS, y del IEEE 1394, se consigue no sólo aumentar la velocidad de transferencia, sino que también se hace con una mayor inmunidad al ruido y con longitudes de cableado aceptadas mucho mayores (de 2-10 metros a 25 metros de distancia cámara – tarjeta). La capacidad de sincronizar la captura de las imágenes con una señal externa también se ha mejorado con el *reset asíncrono*, lo que habilita el inicio de la captura sin tener que esperar el siguiente *vertical blank* (es decir, sin necesidad de finalizar completamente la captura de anterior). La capacidad de controlar la cámara con salidas RS-232 de la tarjeta facilita la variación de parámetros de la cámara, fundamentalmente la ganancia de los canales RGB, o de ópticas motorizadas en tiempo real.

Un aspecto relativamente reciente es la capacidad de poder trabajar con *multitap*. Existen tarjetas capaces de trabajar con varias cámaras a la vez (hasta 6 cámaras de 8 bits por píxel), o usar los canales de que dispone para trabajar con cámaras multitap que transmiten a la vez varios píxeles por cada uno de los canales. Por ejemplo, a 40 MHz, una tarjeta puede recibir la información por un canal de 8 bits a 40 Mbytes/seg mientras que si usa dos canales, la información la puede recibir a 80Mbytes/seg, incrementándose de forma proporcional al número de canales la velocidad de transmisión.

En cuanto a la capacidad de proceso de las tarjetas, actualmente sólo existe la posibilidad de incorporar un Procesador Digital de Señal (DSP), que permite realizar determinados algoritmos de análisis de imagen en la propia memoria de la tarjeta, acelerando la im-

plementación de los algoritmos de visión. El desarrollo de módulos de proceso altamente especializados ha sido abandonado debido a la evidencia de la evolución superior en prestaciones de los procesadores convencionales, que hacen que el costo de desarrollo de las placas altamente especializadas no sea rentable para las casas fabricantes de la misma, al ser éstos rápidamente superados por los constantes incrementos en velocidad y mejoras internas de los procesadores convencionales.

Otra de las recientes ventajas en cuanto a las tarjetas de captura de imágenes es la integración de la tarjeta de captura, la visualización, y el módulo de proceso (CPU) en un solo elemento. Estos equipos se proporcionan en un chasis adecuado para trabajar en entornos industriales.

## Algoritmos de análisis de imagen y reconocimiento de formas

Los algoritmos relacionados con visión artificial son muy variados y abarcan numerosas técnicas y objetivos. Los pasos más comunes en que consiste el análisis o procesamiento de imagen se muestran en la figura 2.

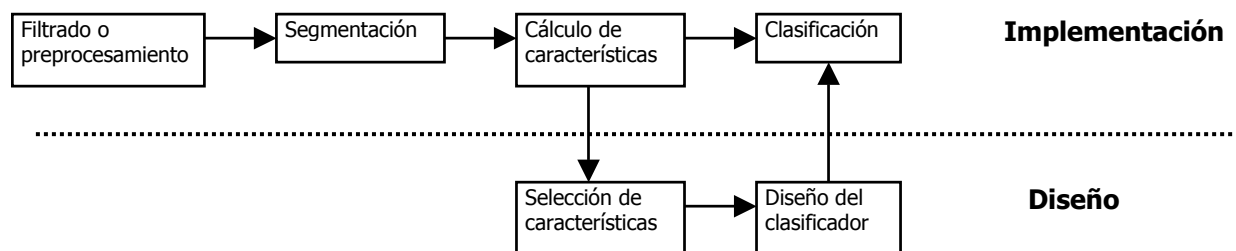


Figura 2: Etapas de un proceso de reconocimiento de formas.

### Filtrado o Pre-procesamiento

Cuando se adquiere una imagen mediante cualquier sistema de captura, por lo general esta no es directamente utilizable por el sistema de visión. La aparición de variaciones en intensidad debidas al ruido, por deficiencias en la iluminación, o la obtención de imágenes de bajo contraste, hace necesario un preprocesamiento de la imagen con el objetivo fundamental de corregir estos problemas, además de aplicar aquellas transformaciones a la imagen que acentúen las características que se deseen extraer de las mismas, de manera que se facilite las operaciones de las etapas posteriores. A este campo pertenecen las técnicas de transformaciones geométricas, las basadas en el histograma, el filtrado espacial (apunte ***"Filtrado Espacial.pdf"***) y el filtrado frecuencial (apunte ***"Filtrado en el dominio de la frecuencia.pdf"***). Aunque existen líneas de investigación abiertas, no es el campo más activo en ese sentido, y las técnicas existentes son bastante útiles para la mayoría de las aplicaciones.

Algunas de las técnicas más habituales son:

- Conversión de los niveles de gris: su objetivo es la mejora de la calidad de la imagen.

- Transformaciones geométricas: los objetivos de ésta son la corrección de la perspectiva y la reconstrucción tridimensional de los objetos de la escena.
- Transformación del histograma: las transformaciones del histograma pueden facilitar la segmentación de objetos de la imagen, aunque habitualmente sólo sirve para modificar el contraste de la imagen y el rango dinámico de los niveles de gris.
- Filtrado espacial y frecuencial: dado que pueden suponer una importante modificación de la imagen original, generalmente sólo se consideran filtrados aquellas operaciones realizadas en el dominio espacial o frecuencial que suponen transformaciones pequeñas de la imagen original.

### **Segmentación o aislamiento de los objetos de interés**

Cuando ya se dispone de la imagen capturada y filtrada, es necesario aislar o separar los objetos de interés de la escena. Se pretende por tanto dividir una imagen en diferentes regiones, o dicho de otra forma, detectar automáticamente los bordes entre los elementos o regiones. Las operaciones de segmentación de una escena dependen de la propia escena y de la información que se busque dentro de la imagen. Por ejemplo, en fotografía aérea se utiliza para separar regiones pobladas, agrícolas, bosques, aguas, etc. En imagen médica se usa para aislar zonas enfermas de los órganos, o para detectar y contar microorganismos en microscopia. Industrialmente, se usa como primer paso para la medición de piezas, para determinar la correcta disposición del etiquetado de los envases (aislar el envase, aislar la etiqueta, determinar si la relación entre estos dos objetos es correcta), para la medición de irregularidades en superficies homogéneas, para la detección de cambios de texturas en las que el sistema ha de determinar si es un defecto o no lo es, etc.

Las técnicas básicas de segmentación se pueden dividir en tres grupos:

- Aplicación de umbrales de niveles de gris (apunte **“Segmentación por umbralización - Método de Otsu.pdf”**).
- Agrupación por rasgos comunes.
- Extracción de bordes.

La segmentación de imagen es una temática actual de investigación y se han definido numerosos métodos diseñados para aplicaciones particulares. No obstante, no hay métodos que sean útiles en una amplia variedad de situaciones, por lo que *cada problema necesita la adaptación o desarrollo de nuevas técnicas*.

### **Extracción y selección de características**

Con los objetos a estudiar ya definidos, es necesario extraer una serie de medidas que los caractericen adecuadamente, en forma de vector de características.

Como norma general, para la elección de características se intenta que:

- Sean discriminantes, es decir, tengan valores numéricos diferentes para clases diferentes.
- Sean fiables, es decir, tengan valores *parecidos* dentro de una misma clase.



- Estén incorreladas, es decir, obtener la misma información con el mínimo número de características.
- Se calculen en un tiempo aceptable, de manera que sea utilizable en problemas de tiempo real.

Dado que las características relevantes son a menudo desconocidas a priori, habitualmente se introducen muchas características candidatas para intentar representar mejor las clases, aún cuando muchas de ellas puedan ser parcial o completamente redundantes o irrelevantes. No obstante, existen unas técnicas para mejorar las prestaciones del conjunto de características escogido inicialmente. Éstas están íntimamente ligadas al reconocimiento de formas en su etapa de diseño del clasificador.

### **Reconocimiento de formas e inteligencia artificial**

Normalmente, la visión artificial se utiliza para tomar decisiones a partir de la información proporcionada por el sistema de adquisición de imágenes y las transformaciones y operaciones realizadas con ellas. La información extraída se puede considerar como un vector que recoge las características o rasgos diferenciadores de la imagen analizada. En el caso de que se trate de una aplicación de medición, este vector recoge todas aquellas medidas que se deseaban obtener. En aplicaciones de inspección, y sobre todo, en aplicaciones de clasificación, este vector es el conjunto de datos con los que ha de trabajar un reconocedor o clasificador encargado de extraer las conclusiones posibles a partir del vector de entrada. Para el diseño del clasificador es necesaria una etapa de selección de características y una etapa de aprendizaje o entrenamiento. Generalmente, se usa el propio clasificador como evaluador del conjunto de características en prueba dado que el método que proporciona la mayor fiabilidad, aunque sea el más costoso temporalmente y de menor generalidad o capacidad de aplicación del clasificador a otros problemas. Respecto a los métodos clásicos de búsqueda secuencial, actualmente uno de los campos de investigación abiertos es el uso de algoritmos genéticos para la selección en paralelo de características.

En el reconocimiento de formas aplicado a la visión artificial se utilizan técnicas de reconocimiento geométrico de formas, como el aprendizaje supervisado (se conoce la clase a la que pertenece cada vector) en condiciones estadísticas o algoritmos de clasificación no supervisados o *clustering* y, además, las redes neuronales, siendo estas últimas especialmente interesantes por su capacidad de aprendizaje adaptativo. Existen numerosos algoritmos para cada una de estas técnicas y se investiga ampliamente el desarrollo de nuevos algoritmos e implementaciones hardware de éstos. El problema fundamental de estas técnica, es que cada una de ellas suele ser la más adecuada para las características de un tipo específico de problema no conocidas a priori, lo que dificulta la selección de la técnica con la que se desea abordar inicialmente el problema.

### **CPU o módulo de proceso**

Se trata del sistema encargado de recibir y analizar, mediante los algoritmos adecuados, las imágenes a la velocidad necesaria para poder interactuar con el entorno en tiem-

po real. Es, pues, la parte del sistema encargada de implementar las funciones y objetivos para los que se haya diseñado el sistema.

Los altos requerimientos computacionales de esta tecnología no se pueden especificar de una forma genérica. Un ejemplo demostrativo de estos requerimientos es el del reconocimiento automático de matrículas mediante una cámara de 8 bits y una resolución de 640x480 (307.200 píxeles). En esta aplicación se necesita una etapa de filtrado, una binarización que convierta la imagen de 256 niveles de gris a sólo 2, una de segmentación que aísle cada caracter de la matrícula y una de clasificación. Las operaciones necesarias se resumen en la siguiente tabla:

|                      | Comparaciones | Sumas y restas | Multiplicaciones | Divisiones | Op. en coma flotante | Total            |
|----------------------|---------------|----------------|------------------|------------|----------------------|------------------|
| <b>Filtrado</b>      | 0             | 2.457.600      | 0                | 307.200    | 0                    | 2.764.800        |
| <b>Binarización</b>  | 307.200       | 0              | 0                | 0          | 0                    | 307.200          |
| <b>Segmentación</b>  | 1.229.310     | 1.510          | 0                | 1.000      | 0                    | 1.231.820        |
| <b>Clasificación</b> | 0             | 3.304          | 129.857          | 0          | 417                  | 133.578          |
| <b>Total</b>         | 1.536.510     | 2.462.414      | 129.857          | 308.200    | 417                  | <b>4.437.398</b> |

Si se tiene en cuenta que el sistema ha de trabajar en tiempo real, capturando para ello 30 imágenes por segundo, se necesita una capacidad de cálculo de 133.121.940 instrucciones por segundo (133 MIPS). Teniendo en cuenta que el ejemplo comentado es una aplicación relativamente simple, podemos comprender los altos requerimientos computacionales para el uso de estas técnicas.

Quizás sea este el campo que más rápidamente ha evolucionado, debido al uso generalizado de este tipo de sistemas en todos los campos. La rápida evolución de los procesadores de propósito general ha desbancado la utilización de procesadores especializados (más caros y de una evolución más lenta por la demanda del mercado), excepto para casos muy concretos. Por tanto, la potencia de los procesadores actuales, junto con las arquitecturas hardware y software que soportan los *sistemas multiprocesadores*, permiten el desarrollo de algoritmos complejos con tiempo de respuesta muy cortos y dentro de los requerimientos de las aplicaciones en tiempo real.

Entre las mejoras más destacadas podemos mencionar:

- Pipeline: capacidad de separar la ejecución de cada instrucción en varias etapas, lo que permite que varias instrucciones se ejecuten en paralelo. El primer procesador Pentium dispone de 5 etapas mientras que el Pentium IV dispone de 20.
- Tecnologías de ejecución dinámica: éstas permiten una ejecución especulativa superando la secuenciación lineal de las instrucciones de búsqueda y ejecución de instrucción. El procesador Pentium contiene un buffer de 40 instrucciones, mientras que el Pentium IV tiene un buffer de 126 instrucciones y un algoritmo mejorado para la predicción de saltos.
- Juegos de instrucciones SIMD (Single Instruction, Multiple Data): las 57 instrucciones MMX de 64 bits que aparecieron con los primeros Pentium MMX permitían un alto paralelismo en la ejecución de instrucciones con tipos de datos multimedia, como son, por ejemplo, las imágenes utilizadas en los sistemas de visión. Esto se traduce en importantísimas mejoras en los tiempos de ejecución de determinadas o-



peraciones muy utilizadas en los algoritmos de procesamiento de imágenes, como las convoluciones o las transformadas de Fourier. Los procesadores Pentium IV disponen del juego de instrucciones de 128 bits SSE2 con 144 instrucciones nuevas.

- Cachés de 2º nivel: estas cachés permiten disponer de memoria para instrucciones y datos dentro de la propia pastilla del procesador, lo que evita los ciclos de acceso a la memoria convencional, incapaz de proporcionar los datos a las velocidades que los procesadores pueden consumirlas. El procesador Pentium II dispone de 512Kb de caché L2 para código y datos, y un bus de 64 bits para conseguir altas transferencias de datos entre la caché y el procesador.
- Unidades de ejecución en coma flotante FPU avanzadas: soporte del formato IEEE 754 de 32, 64 y 80 bits.

No obstante, para aplicaciones de muy elevados requerimientos computacionales y estrictos requerimientos de tiempo, la existencia de hardware especializado o, principalmente, la utilización de otras arquitecturas paralelas que hagan uso de DSP o sistemas FPGA aseguran la viabilidad en este sentido de cualquier aplicación de visión artificial, al permitir la utilización de sistemas multiprocesadores con los DSP en paralelo que se necesiten en la aplicación.

Paralelamente al hardware utilizado, se han desarrollado librerías software que implementan numerosas funciones utilizadas habitualmente en esta tecnología, lo que permite reducir el ciclo de desarrollo de las aplicaciones y facilita la obtención e implementación de algoritmos complejos con tiempos de ejecución adecuadamente bajos.

## Sistemas de respuesta

Se trata de automatismos que responden electromecánicamente con el fin de corregir o evitar, por ejemplo, en los sistemas de producción, las causas generadoras de los problemas de detección. Otra de las funciones de estos sistemas es la generación de estadísticas e informes de del proceso al que están asociados, lo que permite disponer de una información global de los procesos para facilitar la toma de decisiones.

En cuanto a los dispositivos electromecánicos utilizables en los sistemas de visión, destacan los variadores de frecuencia para el control de motores y la comunicación entre las PC's y los autómatas programables que gestionan los distintos dispositivos mediante redes de campo como "*Profibus*" o "*WorldFIP*", y más recientemente mediante "*Industrial Ethernet*", y protocolos *Wireless*. Estos buses permiten la integración de los sistemas de visión y sus acciones de respuesta dentro de la estructura CIM de las empresas, lo que supone un ahorro económico en los casos en que ya se disponga de esta estructura.

## Ejemplo de un Sistema de Visión: Control de Blisters

En esta sección se presentará un breve resumen del proyecto final de carrera de la egresada de la Universidad Nacional de Quilmes, Iris Herrero, de manera que puedan verse los aspectos de un sistema de visión en una aplicación probada y aprobada.

El proyecto consiste en el desarrollo de un sistema de visión para la inspección y supervisión de blisters con base de PVC o aluminio. Tal inspección debe ser capaz de determinar la ausencia de comprimidos en los alvéolos, la presencia de comprimidos fuera de los alvéolos, alvéolos sobrecargados, presencia de cuerpos extraños, comprimidos partidos, comprimidos con desmoronamientos, y comprimidos manchados. Todo el análisis necesario debe tomar como máximo **1 segundo** por cada paso de la blistera. Además, el proyecto cuenta con una interfase hombre-máquina que permite configurar ciertos parámetros de la inspección (nivel de rigurosidad), proporcionar información del error detectado y generar estadísticas de producción. No obstante, nos concentraremos en lo relacionado a la inspección y supervisión.

### El sistema

El sistema consiste básicamente en la integración de un sensor, una CPU, un lente, un software y una estructura de iluminación y montaje.

Para la implementación del proyecto se definió la utilización de la tecnología de la Smart-camera mvBlueLYNX, de la librería mvIMPACT-SDK de Matriz Vision para el procesamiento y análisis de imágenes, y de la librería GNU-PicoGUI para la programación de la interfase hombre-máquina.

En particular, la Smart-camera mvBlueLYNX integra el sensor con el CPU, obteniendo una potente herramienta de desarrollo para una amplia gama de aplicaciones. Su tamaño reducido, sumado al hecho de no necesitar una PC dedicada para que realice el procesamiento, permite que el sistema sea altamente portable y adaptable a los diferentes modelos de blisteras.

### Sensor, tarjeta de captura y CPU

Por lo dicho en el párrafo anterior, se eligió la Smart-camera mvBlueLYNX como elemento sensor, que además integra la tarjeta de captura (necesaria para la digitalización de las imágenes) y el módulo de proceso o CPU. Puede encontrarse información detallada de esta cámara en la siguiente página web:

<http://www.matrix-vision.com/manuals/mvBlueLYNX/main.html>

### Fuente de luz. Tipo de Iluminación. Lentes

Para resaltar los detalles de interés en los comprimidos, se optó por luz blanca difusa como iluminación. Para ello, se utilizaron como fuentes de luz arrays de LEDs blancos. La técnica de iluminación frontal se logró mediante una combinación de difusión y reflexión.

La fuente de luz, el difusor y la cámara se integraron mediante una estructura (campana) rígida. Para favorecer la difusión de la luz aportada por los LEDs, la estructura se pintó interiormente de color blanco mate.

La primera estructura de iluminación desarrollada fue complementada con un lente de 4mm instalado en la Smart-camera. Luego de observar las imágenes obtenidas con esta estructura, se observó que ésta no brindaba una iluminación uniforme: era más intensa en el centro de las imágenes. Por este motivo se debió rediseñar la estructura eliminando las terminaciones en ángulos rectos. En esta segunda estructura se observó una mejora en la distribución de luz.

La utilización de un lente de 4mm generó una importante deformación geométrica de la imagen. Para reducir este efecto no deseado, se decidió su reemplazo por un lente de 8mm. El cambio de lente implicó que la segunda estructura deba ser rediseñada para una nueva distancia cámara-objeto de 42.7cm, dando paso a la creación de la tercera estructura de iluminación.

En las imágenes captadas usando la tercera estructura se observó que las deformaciones geométricas se redujeron notablemente. Esto fue sumamente beneficioso para lograr mayor simplicidad y robustez en la etapa de procesamiento.

A pesar de haberse obtenido unas imágenes muy adecuadas para el procesamiento, esta tercera estructura debió ser rediseñada una vez más. Este cambio se debió a que, por cuestiones operativas y de espacio, el tamaño de la estructura (que se ubica sobre la mesa de amasado de la blistera) debió ser reducido.

El diseño de la estructura definitiva tuvo en cuenta un sistema de montaje robusto para la cámara, que la aisle de los cambios de luz del entorno y la resguarde mecánicamente.

### **Software. Algoritmos de análisis de imagen**

El software está integrado por una lógica de control y análisis y una interfase hombre-máquina.

Para el funcionamiento del sistema, el software requiere de la determinación de varios parámetros. Por ello, se configuró de modo que como primera instancia, luego del arranque automático, se carguen automáticamente dichos parámetros desde un archivo de configuración almacenado en el disco de la Smart-camera.

Como se mencionó, para el procesamiento y análisis de las imágenes se eligieron las funciones contenidas en la librería mvIMPACT-SDK de Matrix Vision.

En la figura 3, 4 y 5 se pueden ver diferentes diagramas de flujo que muestran esquemáticamente la forma en que trabaja el software.

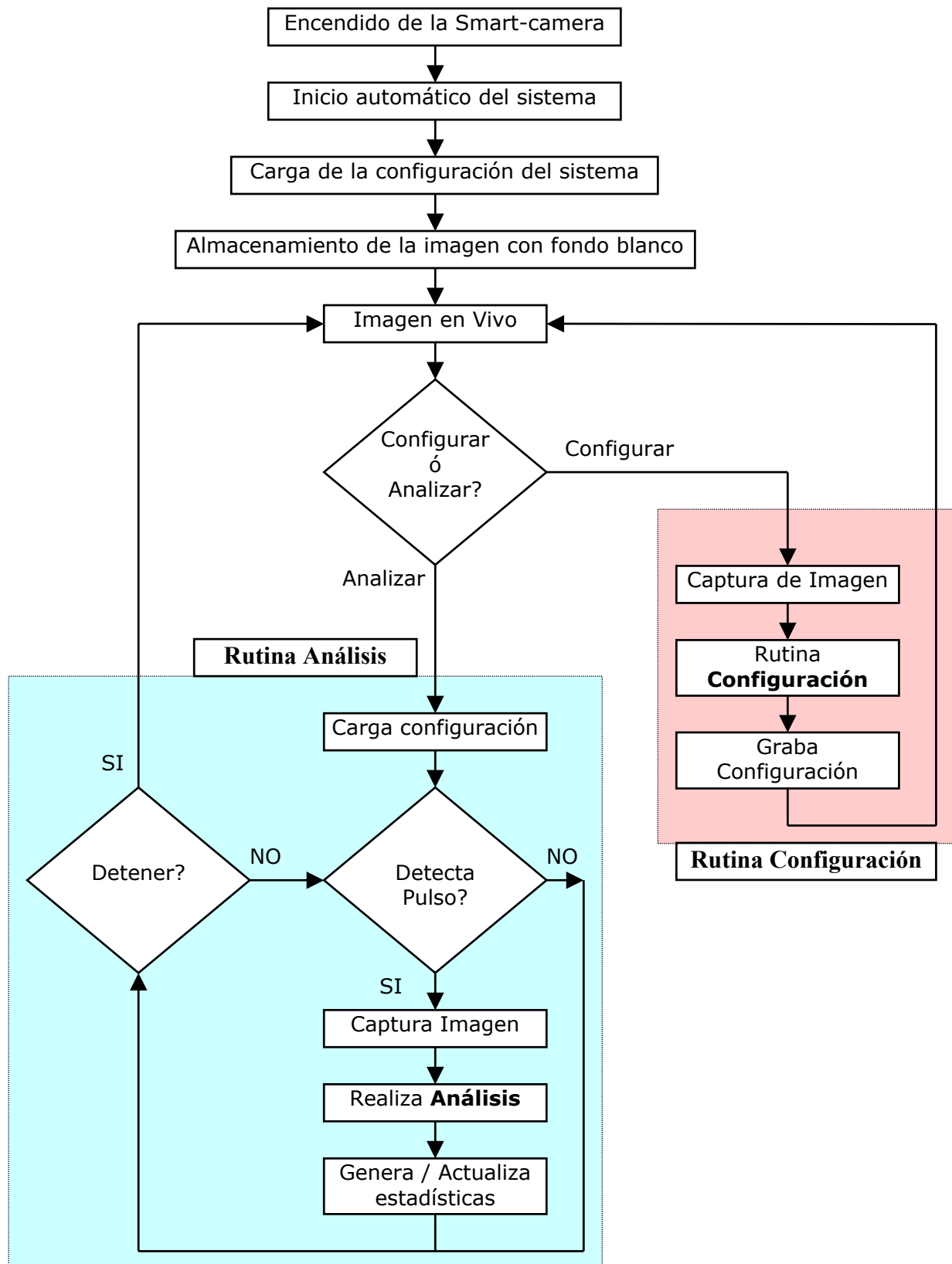


Figura 3: Diagrama de estados.

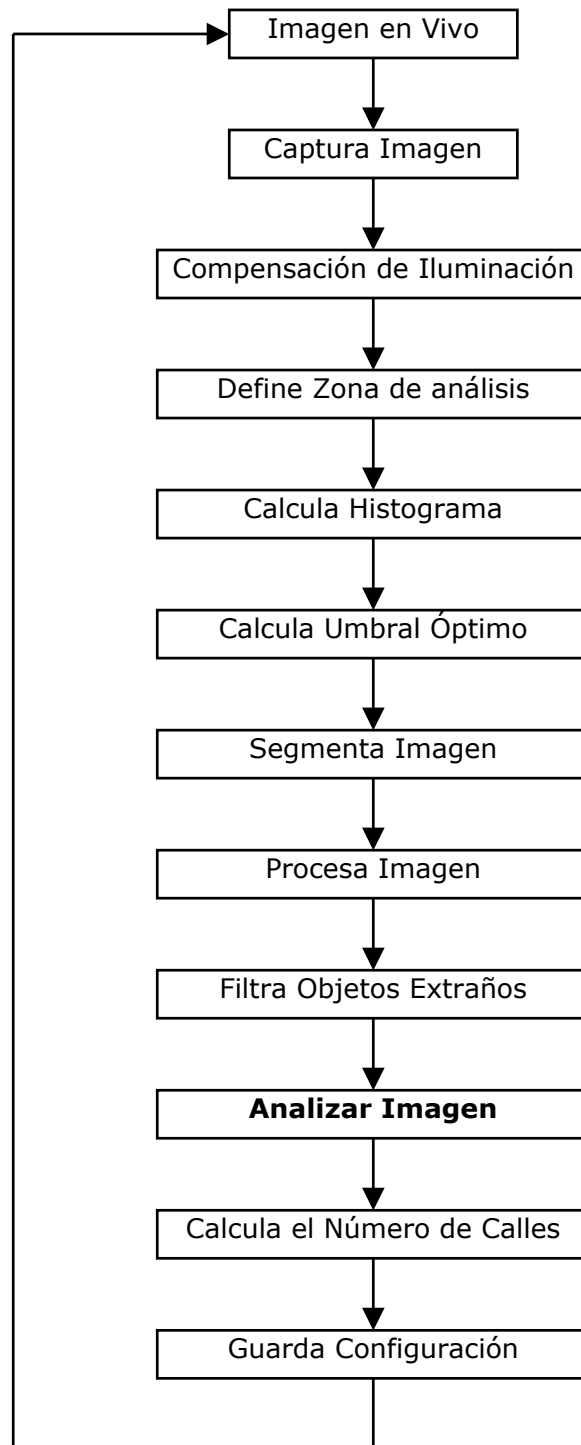


Figura 4: Diagrama de la **Rutina Configuración**.

A continuación se explica brevemente algunos de los pasos presentados en los diagramas.

**Compensación de Iluminación:** debido a la forma rectangular de la estructura de iluminación del sistema, se debió proceder a realizar una compensación por software de la luminosidad de la imagen. La compensación de la intensidad de los píxeles de la imagen se realiza en base a la distribución de la intensidad detectada en una imagen con fondo blanco. Esta imagen es capturada por única vez al instalar el sistema y se encuentra almacenada en el disco de la Smart-camera.

**Define Zona de Análisis:** esta función tiene como objetivo delimitar la zona de análisis con el fin de descartar posible contaminación de la imagen, excluyendo zonas que exceden los límites de interés.

**Calcula Histograma:** permite determinar en qué rango de intensidades se encuentra centralizada la mayor cantidad de píxeles.

**Cálculo Umbral Óptimo:** esta función permite el cálculo óptimo del umbral de una imagen o histograma, usando el *método de Otsu*.

**Segmentación de la Imagen:** usando el umbral óptimo, se realiza la segmentación de la imagen. Si el objeto y el fondo difieren en el nivel de gris, se pueden identificar los píxeles que pertenecen al objeto mediante el proceso de segmentado. Para segmentar la imagen se utiliza la binarización: se etiqueta con cero los píxeles cuyo nivel de gris sean menores al valor del umbral óptimo y con 1 a los píxeles con nivel de gris mayor.

**Proceso Imagen:** la imagen utilizada para obtener la configuración es la imagen a tomar como patrón. De ésta, se obtienen los valores que determinarán qué comprimido cumple las especificaciones y cuáles no. Por este motivo se procesa esta imagen de modo de obtener los datos más exactos. Con este fin, en esta etapa se aplican dos funciones a los efectos de eliminar las fallas de iluminación u otros factores que afecten la definición de la imagen a establecer como patrón. Éstas son:

- 1) Transformada de Morfología ***Open***: esta función se aplica sobre la imagen segmentada (binarizada) y sirve para suavizar los contornos de los comprimidos, separando así objetos que se encuentren vinculados por pequeñas uniones.
- 2) Filtrado de Objetos por Área: esta función permite remover objetos basándose en su área. De este modo podemos filtrar (eliminar) todos los objetos de una imagen cuya área sea menor que un determinado valor.

**Analizar Imagen:** Esta etapa es la más importante de la rutina, dado que previo procesado de la imagen, se obtienen los parámetros esenciales del patrón. Estos componen la configuración que se utilizará como base para el control de la blisteadada de este producto.

Los parámetros que se utilizaron para la creación de la configuración fueron:

1. Área (**AREA**);
2. Factor de forma (**ROUDNESS**);
3. Coordenadas del rectángulo perimetral (**EXTREME\_BOX**).



En base a las características de las fallas que se pretenden detectar, se comprobó empíricamente que este conjunto de parámetros otorgaba una gran confiabilidad requiriendo un tiempo de procesamiento reducido.

Definición de los parámetros:

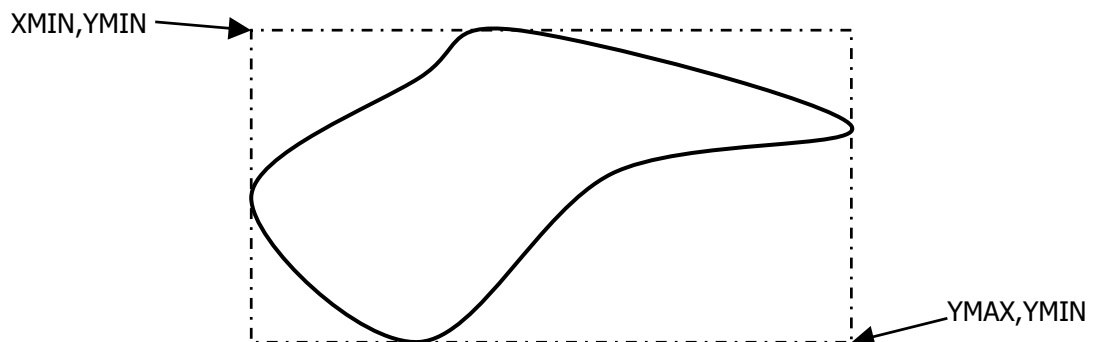
Área (AREA): es el número de píxeles blancos en un objeto. El número de píxeles es multiplicado por el valor del área del píxel. Los agujeros no contribuyen para el área.

Factor de forma (ROUDNESS): este valor es derivado del perímetro y el área de un objeto. Este parámetro tiene un rango de valores de 0 a 1. Un objeto circular tiene el máximo valor del rango (1) y para objetos irregulares o alargados el parámetro tiende al mínimo valor del rango (0). Su ecuación es:

$$Roudness = \frac{4\pi \cdot Area}{Perimetro^2}$$

Coordenadas del rectángulo perimetral (EXTREME BOX): especifica la máxima y la mínima coordenada de un objeto (bounding box). Las coordenadas son retornadas en dos puntos en las unidades especificadas: el punto superior izquierdo (XMIN,YMIN) y el punto inferior derecho (XMAX,YMAX). De modo que:

$$\begin{cases} ancho = X_{Max} - X_{Min} + 1 \\ alto = Y_{Max} - Y_{Min} + 1 \end{cases}$$



**Guardo Configuración:** los parámetros extraídos del análisis de la imagen patrón son guardados en un archivo en el disco de la cámara con extensión 'txt'. Los datos que se guardan son los siguientes: número de alvéolos; área mínima y máxima; factor de forma mínimo y máximo; valor del umbral; flag de descarte por comprimido fallado, ausente y manchado; tipo de fondo; coordenadas del rectángulo perimetral de cada alvéolo; número de calles; y posición de las calles.

El valor del umbral se calcula en forma automática y es un parámetro altamente afectado por las variaciones en la iluminación. Por este motivo, es muy importante que el sistema de iluminación implementado permita un aislamiento adecuado del sistema a las variaciones externas de la luz, y que a su vez otorgue una iluminación constante y uniforme sobre la zona de análisis.

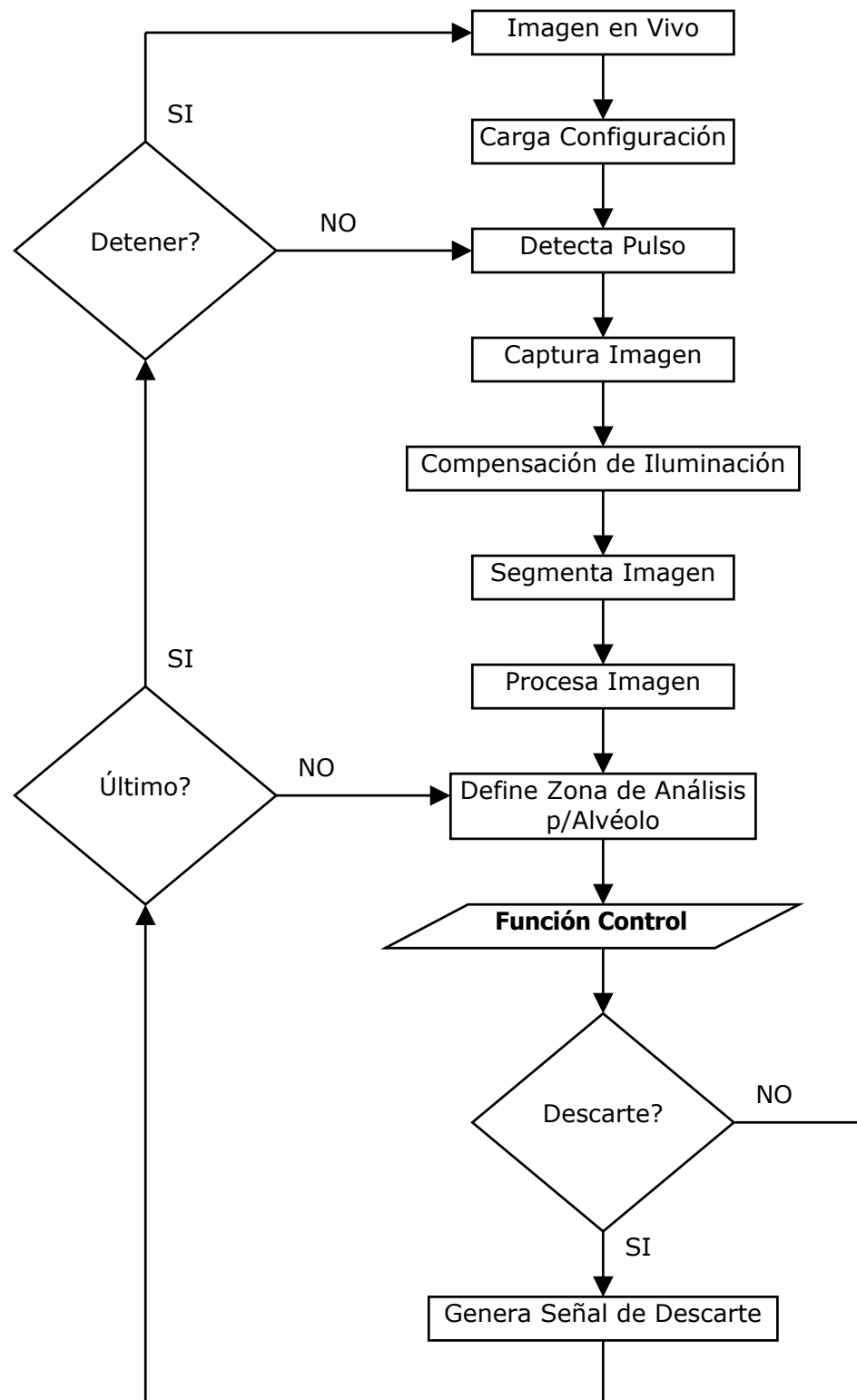


Figura 5: Diagrama de flujo de la **Rutina Analizar**.

A continuación se describen algunos de los pasos de la **Rutina Analizar**.

**Carga Configuración:** carga en la memoria de la Smart-camera los parámetros guardados en el archivo de configuración (en el disco de la cámara), correspondientes al tipo de blisteado a realizarse.

**Detecta Pulso:** una vez cargada la configuración, el sistema se queda a la espera de la señal digital que envía el sensor a la Smart-camera para realizar la captura y procesamiento de la imagen a analizar.

**Compensación de Iluminación:** Idem a la **Rutina Configuración**.

**Segmenta Imagen:** se realiza usando el valor de umbral calculado en base a la imagen patrón. Este valor es cargado del archivo de configuración.

**Proceso Imagen:** a diferencia de la rutina configuración, en esta etapa no se realiza el filtrado de objetos por área mínima. Si realizáramos ese filtrado en este paso, podríamos perder información importante para realizar la supervisión del correcto estado de los comprimidos de los blisters.

**Define Zona de Análisis por Alvéolo:** en esta etapa se define una subimagen por alvéolo. El número total de alvéolos y sus correspondientes coordenadas son parámetros que se cargan del archivo de configuración. Las regiones definidas por estas coordenadas corresponden a los rectángulos externos de cada alvéolo de la imagen patrón con la que se realizó la configuración. Se trabaja con una subimagen para optimizar la búsqueda y minimizar el tiempo de requerido por el sistema para ejecutar la rutina de análisis.

**Función Control:** la subimagen definida en la etapa anterior es pasada como parámetro a la **Subrutina Control**, donde se realiza un análisis y se determina el correcto estado del comprimido de ese alvéolo.

**Genera Señal de Descarte:** en caso de producirse la detección de un comprimido defectuoso, la subrutina retorna un valor que indica que se produjo una falla en el comprimido de ese alvéolo, por lo que se genera en forma automática la señal de descarte de ese blister.

## **Fases del Proyecto**

A continuación se enunciarán los pasos que constituyeron el desarrollo del sistema de visión (ordenados temporalmente):

- Análisis de las características técnicas (alcances y limitaciones) de la Smart-camera mvBlueLINX.
- Puesta en funcionamiento de la Smart-camera mvBlueLYNX: Upgrade del software, conexionado, configuración de interfases.
- Configuración del puerto serial RS 232 requeridos para la comunicación e interacción entre la Smart-camera mvBlueLYNX y la PC mediante el protocolo de comunicación TCP.

- Configuración de la interfase Ethernet LAN para establecer la comunicación entre la Smart-camera mvBlueLYNX y una PC mediante el protocolo FTP (para enviar a la Smart-camera los programas compilados en la PC de desarrollo).
- Instalación, configuración y actualización de los programas y librerías requeridas para el desarrollo y compilación de los programas en la PC.
- Estudio de la librería mvIMPACT-SDK.
- Programación en el entorno de desarrollo Visual C++ de las rutinas de configuración y análisis mediante el uso de las funciones de análisis y procesamiento de imágenes definidas en la librería mvIMPACT-SDK.
- Prueba y ajuste del funcionamiento de las rutinas de configuración y análisis en la PC mediante el uso de la función *import* de la librería mvIMPACT-SDK, la cual permite la carga de imágenes guardadas en el disco.
- Estudio de la librería GNU-PicoGUI mediante el análisis de ejemplos y el manual de referencia API.
- Desarrollo de la interfase hombre-máquina mediante el uso de la librería GNU-PicoGUI.
- Acople a las rutinas de configuración y análisis de la interfase hombre-máquina.
- Pruebas y ajustes del funcionamiento del nuevo software integrado en la Smart-camera mvBlueLYNX.
- Diseño y armado de la primera estructura de iluminación.
- Cálculo de lente acorde a las dimensiones de la estructura de iluminación y la zona de inspección definida.
- Modificación del software para la captura de imágenes por detección de la señal digital emitida por el sensor.
- Montaje de la mesa de ensayos para el sistema integrado por la Smart-camera mvBlueLYNX, el software, la estructura de iluminación y el lente.
- Realización de pruebas y ajustes del sistema integrado en la mesa de ensayos.
- Instalación y ensayo del prototipo en la planta.
- Procesamiento de los resultados del ensayo.
- Ajustes en el diseño de la estructura de iluminación.
- Segundo ensayo en la planta.
- Ajustes en el diseño y tamaño de la estructura de iluminación.
- Cálculo del lente adecuado a la nueva estructura de iluminación.
- Tercer ensayo en planta.
- Rediseño de la estructura de iluminación para una mejor integración mecánica a la línea de blisteado, e inclusión de un mecanismo para el montaje de la Smart-camera mvBlueLYNX.
- Adaptación del software a la nueva estructura de iluminación.
- Pruebas y ajustes del sistema final.
- Realización de prueba final en la mesa de ensayos con el sistema terminado.

### Resultados obtenidos

- Durante las pruebas de funcionamiento, el sistema presentó una muy buena respuesta en la detección de comprimidos ausentes, manchados y calados.

- Se presentaron algunos casos de no detección de comprimidos partidos, dado que la detección de éstos depende fuertemente del valor del umbral que calcula el sistema para la segmentación de la imagen. Otro factor que afecta a la detección de estos casos es la holgura (o la falta de la misma) existente entre el comprimido y el alvéolo. Se debe tener en cuenta que los casos no detectados por el sistema automático tampoco son fácilmente detectables por el operador humano.
- La forma rectangular de la estructura de iluminación (con su distribución no uniforme de la luz) dificultó la determinación de un umbral único adecuado para la correcta segmentación de toda la imagen. Por ello, se debió proceder a compensar por software el efecto de la iluminación a fin de evitar que el cálculo automático del umbral se vea afectado.
- El tiempo requerido por la Smart-camera mvBlueLYNX para realizar un análisis completo fue de aproximadamente **1.5 segundos**, siendo el tiempo máximo admisible para el análisis de **1 segundo**. En este punto se propone como solución posible la adquisición de una Smart-camera mvBlueLYNX de mayor capacidad de procesamiento.

A continuación se presenta un ejemplo de la puesta a prueba del sistema con una tanda de blisteado en el que se forzaron diferentes casos de comprimidos defectuosos.

|                                |            |
|--------------------------------|------------|
| Total comprimidos analizados   | <b>640</b> |
| Total comprimidos OK           | <b>621</b> |
| Total comprimidos NoOK         | <b>19</b>  |
| Total comprimidos ausentes     | <b>6</b>   |
| Total comprimidos fallados     | <b>10</b>  |
| Total comprimidos manchados    | <b>2</b>   |
| Total comprimidos superpuestos | <b>1</b>   |

Durante esta prueba, el sistema presentó tan sólo 2 falencias en la detección de un total de 640 comprimidos analizados (Tasa de Error: 0,31%).

## Referencias

1) AITEX – Instituto Tecnológico Textil, “Análisis de la tecnología de visión artificial aplicada al sector textil”.

[http://www.aitex.es/estudios/Vision\\_Artificial1.pdf](http://www.aitex.es/estudios/Vision_Artificial1.pdf)

2) Iris Herrero, “Control de Blisters mediante Visión Artificial”, Informe del Proyecto Final de la carrera de Ingeniería en Automatización y Control Industrial de la Universidad Nacional de Quilmes.