

**Universidad de las Ciencias Informáticas
Facultad 2**



“Propuesta de un filtro para la eliminación del ruido causado por la iluminación en el sistema de visión de un robot móvil.”

**Trabajo de Diploma para optar por el título de
Ingeniero en Ciencias Informáticas**

Autores:

Yaimara Márquez Álvarez
Licel Salazar Reyes

Tutor:

Ing. Sergio Orlando Boudy González

Co-Tutor:

Ing. Daimara Martinez Borrell

**Ciudad de la Habana
Junio 2009**

Declaración de Autoría

Declaramos ser los únicos autores de la presente tesis y reconocemos a la Universidad de las Ciencias Informáticas los derechos patrimoniales de la misma, con carácter exclusivo.

Para que así conste firmo la presente a los ____ días del mes de _____ del año _____.

Yaimara Márquez Alvarez

Licel Salazar Reyes

Ing. Sergio Orlando Boudy González

Ing. Daimara Martínez Borrel

Datos del tutor

Sergio Orlando Boudy González

Graduado de Ingeniería Informática. Graduado en la Universidad de las Ciencias Informáticas durante el año 2008. Profesor de la asignatura Máquinas Computadoras de la Universidad de Ciencias Informáticas (UCI). Investiga en los temas de visión por computadora, reconocimiento de patrones, robótica, Inteligencia Artificial aplicada al control y desarrollo de hardware para la automatización.

Correo electrónico: soboudy@uci.cu

Daimara Martínez Borrel

Graduada de Ingeniería Informática. Graduado en la Universidad de las Ciencias Informáticas durante el año 2007. Profesora de la asignatura de Teleinformática de la Universidad de Ciencias Informáticas (UCI).

Correo electrónico: dmartínez@uci.cu

Pensamiento



*Cuando se es joven, se crea. Cuando se es inteligente, se produce.
No se adapta, se innova: la medianía copia; la originalidad se atreve.*
José Martí

AGRADECIMIENTOS

Agradezco a todas las personas que de una forma u otra han contribuido a la realización de este Trabajo de Diploma. A mi familia por su apoyo incondicional. A mis amigos. En especial quiero agradecer a nuestro tutor Ing.Sergio Orlando Boudy González y a nuestra co-tutora Ing.Daimara Martinez Borrell por las horas de entrega y dedicación.

A todos... muchas gracias

Yaimara Márquez Álvarez.

Quiero agradecer primeramente a mis abuelos Francisca y Pastor, ya que de algún modo esta tesis es fruto de sus esfuerzos y sacrificios, por haber sido el mejor ejemplo de vida que un hijo puede esperar. Gracias a su amor he alcanzado mis metas. Mis logros son suyos.

A mis padres Margarita y Emilio por su amor, por su apoyo, confianza, por los consejos brindados en el transcurso de mi vida.

A mis queridas hermanas por su inmenso cariño, por su alegría, por toda la confianza que depositan en mí, por ser las mejores amigas.

A mi familia en general, mi tío, tías, primos, a mi abuelo Emilio.

A mi abuela Beni que aunque no está presente físicamente siempre está con nosotros.

A Alexei Aguilera Valiente por su apoyo incondicional, por su preocupación, porque me proporcionó felicidad y tranquilidad. Gracias por hacerme tan feliz todo este tiempo y por conocer contigo un sentimiento que no había experimentado.

A Andrés por su preocupación y por su apoyo.

A mi compañera de tesis Yaimara por ser mi mano derecha en este trabajo.

A mis tutores Sergio y Daimara por el apoyo brindado y los consejos para llevar a cabo el desarrollo de este trabajo.

A Bessy por su amistad y todo el cariño.

A todos los profes por las enseñanzas brindadas.

A nuestras secretarias Nilsa y Tania que me han ayudado en todo lo que pueden.

A los compañeros que encontré en esta universidad, en especial Ronaldo, Alexander, Yordan, Denise, Yenli, Yara, Yailin, Halena, Yeisis, Arianna, a mis compañeras de cuarto, Carmen, Dilma y Maelis.

A mis amigas Yisel, Maily, Sheila, por su gran amistad.

A la Revolución y a nuestro Fidel por permitirme tener una educación como la que he recibido, y por estudiar en un centro tan importante como este.

Licel Salazar Reyes.

DEDICATORIA

En especial quisiera dedicarle este trabajo a mi mamá, por estar presentes en cada momento importante de mi vida, por haber extendido su mano ante cada obstáculo, por ser mi razón de ser, mi guía, mi hermana y por sobre todas las cosas mi mejor amiga.

A mis abuelos por haber depositado tanta confianza en mí, por ser abuelos y padres que cada día me ayudan a ser una mejor persona, por estar ahí siempre que los necesito.

A mi padre por sus buenos consejos y esas largas conversaciones del día a día que me levantan el ánimo y me ayudan a ver la vida de una forma diferente.

A mi hermana Yuliet, que aunque no le he dedicado mucha atención en los últimos tiempos es mi preciado tesoro.

A mi prima Yaima por su compañía y su apoyo sin medida.

A mis tíos por su ayuda incondicional.

A todos mis familiares que cada día se preocupan por mis estudios y mi salud.

A todos mis amigos(as) en especial a Yanelis, Akemi, Yusneidys y Daines María por hacer más llevadera mi vida como estudiante.

A mi compañera de tesis por ayudarme a realizar este trabajo tan importante en mi vida.

A nuestro querido compañero Fidel Castro por permitirme estudiar en esta universidad, donde he logrado cumplir parte de mis metas.

A todas aquellas personas que confiaron en mi, espero haber estado a la altura de sus deseos, y a todos aquellos que no creyeron que llegaría a la cima, solo decirles que aquí estoy.

Yaimara Márquez Álvarez.

Este trabajo tan importante quiero dedicarlo a:

A mis abuelos Francisca y Pastor por dar todo lo que tienen para lograr mis éxitos y metas.

A mis padres por su apoyo incondicional.

A Lidice y Lilitiana por ser mis mejores amigas y las mejores hermanas del mundo.

A mi tío, mis tías y mi abuelo Emilio por el apoyo y la ayuda que me brindan.

A la memoria de mi abuelita Beni que siempre esta en mi corazón.

A Alexei por ayudarme a tener fuerzas para realizar este trabajo, y por hacer que saliera adelante, principalmente en este curso que ha sido duro para mí, por su apoyo, por ser un amigo y mucho más.

A mi compañera de tesis por ayudarme a realizar un trabajo tan importante como este.

A mis tutores por la ayuda brindada.

A mis amigas Tita, Maily y Sheila.

Licel Salazar Reyes.

RESUMEN

En el presente trabajo se realizó un estudio sobre los filtros existentes en algunas de las etapas del procesamiento digital de imágenes el cual emplea técnicas que se aplican a las imágenes con el objetivo de mejorar el aspecto de estas. En el transcurso de la investigación se pudieron resaltar características ventajosas y desventajosas de estos filtros, permitiendo conocer cual de ellos resultaría más efectivo al ser aplicado a imágenes captadas por algún dispositivo óptico. Se propone un filtro para la eliminación del ruido causado por la iluminación en el sistema de visión del robot móvil Ricimaf – SjyV, este robot fue creado por un grupo de profesores de la UCI con la ayuda de especialistas del Instituto de Cibernética, Matemática y Física. Se propone el mejoramiento de su sistema de visión mediante la incorporación de los métodos de filtrado.

PALABRAS CLAVES

Filtro, ruido, sistema de visión, robot móvil.

Índice

Declaración de Autoría.....	I
AGRADECIMIENTOS	IV
DEDICATORIA.....	V
RESUMEN.....	VII
CAPITULO 1: FUNDAMENTACION TEORICA.....	7
1.1 Introducción	7
1.2 Estado del Arte.....	7
1.2.1 La robótica en la actualidad.....	10
1.2.2 Límites de la robótica actual.....	11
1.2.3 Grupo de Investigación de Automática Aplicada	12
1.2.5 Departamento de Control Automático del Instituto de Cibernética, Matemática y Física (ICIMAF).....	12
1.3 Principales conceptos.	13
1.3.1 Robótica.....	13
2.2.1 Robot	13
1.3.3 Método.....	14
1.3.4 Píxel.....	14
1.3.5 Imagen	14
1.3.7 Tratamiento de imágenes.....	14
1.3.8 Visión por computadora	16
1.3.9 Visión Artificial	17
1.4 Conclusiones	17
CAPITULO 2: ESTUDIO DE FILTROS EN LAS DISTINTAS ETAPAS DE PROCESAMIENTO DE IMÁGENES.....	18
2.1 Introducción	18
2.2 Procesamiento de Imágenes Digitales	18
2.2.1 Captura de la imagen.....	19
2.2.1.1 Espacios de color.....	21
2.2.2 Filtrado o pre –procesado de imágenes.....	24
2.2.3 Segmentación.....	44
2.2.3.1 Segmentación basada en características	46
2.2.3.2 Segmentación basada en transiciones	49
2.2.3.3 Segmentación basada en homogeneidad.....	60
2.3 Conclusiones	60
CAPITULO 3: PROPUESTA DE FILTRADO	61
3.1 Introducción	61
3.2 Deficiencias detectadas en el sistema de visión antiguo	61
3.3 Sistema de visión propuesto	63
3.3.1 Captura.....	63
3.3.2 Filtrado.....	66

3.3.3 Segmentación	70
3.4 Conclusiones	75
Conclusiones Generales	76
Recomendaciones	77
Referencias Bibliográficas.....	78
Anexos	80
Glosario de términos.....	84

INTRODUCCION

Uno de los sentidos más importantes de los seres humanos y los animales superiores es la visión. Los ojos proporcionan mucha información sobre el entorno. Hay varios factores que combinados, determinan el tipo de visión que tienen los distintos animales. Aunque diferentes, todos los ojos tienen algo en común: células que reaccionan ante la luz. Tanto en los humanos como en los animales, las imágenes captadas por los ojos son procesadas por el cerebro. Los ojos más sencillos no hacen más que detectar si los alrededores están iluminados u oscuros y los más complejos sirven para proporcionar el sentido de la vista y están formados por muchas facetas simples que dan una imagenⁱ píxelada; no imágenes múltiples, como a menudo se cree (del Risco Garces, Izquierdo Perez, & Alonso de Leon, 2007).

La evolución ha separado en dos tendencias la forma en que se ha ido complicando los órganos visuales a lo largo de las especies. Nosotros, al igual que el resto de los vertebrados en general, tenemos un ojo en cámara. Eso significa que un único ojo, compuesto por muchas células, da una imagen completa que envía al sistema nervioso. Forma una cámara hueca, cóncava, estando los receptores de luz tapizando el interior. Los rayos luminosos entran a partir de una zona transparente que hace un “orificio”, de forma que dependiendo de dónde viene la luz, estimula a unos receptores o a otros. Esa es la idea de la “cámara oscura”, un ingenioso y simple mecanismo por el que obtenemos una imagen proyectada e invertida. El sistema de lentes potencia y afina este mecanismo en el que la luz de una localización concreta del espacio estimule al menos un número posible de receptores, de forma que obtengamos una imagen nítida.

Con el paso de los años y el avance de la ciencia y tecnología se ha querido adaptar el sistema de visión humano a máquinas capaces de desempeñar tareas realizadas por el ser humano o que requieren del uso de inteligencia, en este caso específico los robots móviles. Anteriormente en contraste a lo antes dicho el uso de cámaras en los robots móviles aún no se había extendido masivamente. Esto se debió en parte a la complejidad conceptual de extraer la información visual del flujo de imágenes y en parte al alto coste computacional que conllevaba el tratamiento de imágenes. A estos problemas se les ha ido dando solución con el paso de los años, siendo creadas máquinas capaces de visualizar y procesar imágenes de manera rápida y menos costosa.

La visión estereoscópica o binocular en el ser humano permite capturar dos imágenes en las retinas y posteriormente ser fusionadas en el cerebro para dar una sensación de profundidad. Por lo tanto si se toman dos imágenes a diferente paralaje es posible tener una mayor información de la escena permitiendo al cerebro discernir entre objetos localizados a diferentes profundidades así como obtener una perspectiva de la escena. Este proceso ha sido implementado, haciendo uso de diversas técnicas de procesamiento digital de imágenes y visión por computador. Actualmente es aplicado en diversas tareas en el campo de la robótica. Las aplicaciones incluyen manipulación guiada visualmente dado que la visión estéreo es clave en la detección de posiciones y formas.

Tradicionalmente el campo de la visión artificial ha evolucionado aparte de la robótica móvil y la generación de comportamiento, concentrándose en problemas como la segmentación, el reconocimiento de formas, el filtrado de bordes, distintas transformadas y filtros, etc. Dentro de los sistemas deliberativos clásicos, la misión de la visión es la reconstrucción “objetiva” y etiquetada del entorno (J.Nilsson, 1969). Dentro de los sistemas basados en comportamientos, la visión se inserta como un sensor más (Horswill, 1997) sobre el que se hacen cálculos específicos, en el marco de una percepción subjetiva y orientada a tarea. Algunos usos típicos de la visión en los robots móviles incluyen el seguimiento visual (Garcia Alegre, Bustos, & Guinea, October 1995.), el reconocimiento de caras (Arkin, Fujita, & Hasegawa, 2003) la auto localización y la navegación por flujo óptico (Duchon, Haren, & Pack Kaelbling, 1998). La mayoría de ellos utilizan un análisis bidimensional de imágenes.

Entre las capacidades visuales que posee el ser humano se encuentra la de poder seguir objetos que tenemos delante. Aunque lo hagamos de forma inconsciente, mediante esta capacidad podemos realizar múltiples actividades cotidianas, como leer, caminar, conducir, etc. El seguimiento visual está estrechamente imbricado con muchas de las tareas que realizamos. La introducción de estas capacidades en los sistemas visuales artificiales es una de las áreas de investigación en la visión por computador y la robótica de hoy en día. El procesamiento de caras y reconocimiento de rostros familiares compromete la actividad de múltiples regiones cerebrales.

En el caso de los humanos estos atributos son sensorados por el cerebro, en el caso del sistema de visión de los robots existen métodos donde dichos atributos se guardan en un fichero, utilizados luego en el programa de dicha máquina.

Al menos dos principales aspectos justifican el uso de la visión para proveer de capacidad sensorial a un robot móvil: la amplia disponibilidad de cámaras CCDⁱⁱ baratas y la gran cantidad de información del entorno donde se mueve el robot en una sola imagen (en comparación con las medidas de los sensores de ultrasonidos o láser). Actualmente, la visión generalmente resulta en una buena combinación balanceada de bajo ruido, alta cantidad e información y bajo costo. Sin embargo, la extracción de información relevante en tiempo real a partir de una sola imagen o un conjunto de imágenes es una tarea bastante dura, lo que sigue siendo un gran obstáculo todavía para el uso intensivo de la visión en robots móviles. El flujo óptico obtenido a partir de una secuencia de imágenes consecutivas es una técnica que ha sido adoptada para extraer información de importancia del entorno del robot móvil.

La extracción de información de una imagen siempre ha sido un tema de suma importancia en el campo de la Física por el fácil manejo de una imagen, su bajo costo y en los últimos años, por el avance tecnológico de la informática. El campo de aplicaciones del análisis de imágenes es muy amplio cubriendo áreas como la medicina, agricultura y medio ambiente, robótica, sensoramiento remotoⁱⁱⁱ, biología y otros. Una aplicación importante del análisis de imágenes es la segmentación de una imagen que consiste en obtener una descripción concreta en términos de sus bordes y/o regiones. Esto significa identificar las características que describen tanto los contornos (detección de bordes) como las regiones donde ellos se concentran (extracción de regiones). Estas dos aproximaciones son duales pero sus algoritmos de identificación son muy diferentes y desgraciadamente la mayoría de veces lleva a resultados diferentes de segmentación. Los métodos clásicos de segmentación de imágenes usualmente asumen los bordes como los extremos locales del gradiente de los niveles de gris, los cuales introducen dificultades en su detección debido al ruido introducido por la cámara, el muestreo o los mismos objetos^{iv} que componen la imagen. Estas dificultades de segmentación se genera básicamente por el hecho de usar operadores matemáticos definidos en el dominio de las funciones continuas y extrapoladas a funciones discretas de una imagen. En el presente trabajo se investigan propuestas que posibiliten la eliminación del ruido que causa la iluminación en el sistema visual del robots móvil Ricimaf-SjyV a la hora de procesar las imágenes, mediante el estudio de filtros existentes. Logrando así mejoras en el sistema de visión del mismo.

La investigación se ha desarrollado para el robot móvil Ricimaf-SjyV creado por un conjunto de profesores de nuestra universidad y con la ayuda de especialistas del ICIMAF^V. El cual presenta limitaciones en su sistema de visión, debido a que cuenta con una única cámara y le es imposible determinar la forma de un mismo objeto bajo la influencia de diferentes focos de iluminación. Las imágenes son tomadas desde aproximadamente 15cm de altura. La cámara está ubicada en la parte frontal del robot, con posibilidades de inclinación hacia ambos lados, así como arriba y abajo. Los elementos situados en el campo de visión poseen diferentes colores y formas. Se cuenta con una cámara de baja resolución, donde los colores de las imágenes captadas pueden ser diferentes a los que esperamos obtener, por lo que resulta conveniente calibrar la cámara para obtener mejores resultados.

A raíz de las limitaciones presentadas por el robot móvil Ricimaf-SjyV en su sistema de visión, y en aras de darle solución a estas, en el desarrollo de este trabajo se siguieron las recomendaciones hechas en la tesis “Desarrollo de una aplicación para el control inteligente de un robot manipulador”. Donde recomiendan continuar el desarrollo de la investigación con el propósito de mejorar dicho sistema. Refinando específicamente la etapa de segmentación de la imagen, mediante la aplicación de un método basado en análisis discriminante (método de Otsu) que permitiría al sistema tener un mayor grado de independencia en lo relativo a la iluminación del entorno de trabajo.

La iluminación de una escena o entorno de trabajo tiene que realizarse de una forma correcta dada la importancia que tiene, ésta permite distinguir claramente los detalles de un objeto. Una mala iluminación en el área de trabajo puede producir fenómenos tales como la creación de las sombras y los reflejos. Las sombras son inconvenientes para detectar correctamente la forma de los objetos, sobre todo cuando estos son oscuros, por eso se trata de evitar el efecto de las sombras.

Los sistemas de visión basados en una única cámara adolecen de indeterminación en cuanto a la distancia a la que se encuentran los objetos detectados, los colores y el brillo que inciden sobre ellos. Dicha ambigüedad de profundidad se puede reducir añadiendo información al sistema sobre el tamaño de los objetos. El procesamiento digital de imágenes (PDI) es una de las ramas de la ciencia que ha experimentado gran desarrollo en la última década, debido fundamentalmente a dos factores: el avance vertiginoso que se ha originado en el hardware, en particular, los procesadores especializados, los arreglos de procesadores, cámaras digitales y el otro factor es la importancia que esta técnica ha adquirido en la vida moderna. En la actualidad muchos

autores han determinado incluir dentro del concepto de visión por computadora (VC)^{vi} al PDI incluyendo el análisis o interpretación de imágenes. El análisis de imágenes por computadora es un procedimiento que tiene como objetivo procesar, identificar e interpretar automática o de forma semiautomática los aspectos más relevantes presentes en imágenes que han sido previamente digitalizadas. Esta fusión se justifica con el desarrollo de la robótica y los sistemas de control, en los cuales es de suma importancia el análisis o interpretación de las escenas.

En general la visión por computadora, y así aparece en muchos sistemas prácticos, se lleva a cabo acorde a los pasos siguientes:

- 1- Adquisición de la imagen.
- 2- Filtrado o preprocesamiento.
- 3- Segmentación o posprocesamiento.
- 4- Reconocimiento o descripción.

El sistema de visión por computadora de un robot móvil está sujeto a la incidencia de agentes externos. La iluminación es uno de los que más ruido provoca a la hora de realizar el procesamiento de las imágenes. Por tal motivo se hace necesario realizar un estudio para seleccionar un método de filtrado que sea capaz de brindar un grado de robustez al sistema de visión en lo relativo a las variaciones en la iluminación dentro del área de trabajo^{vii}.

Al aplicar filtros para el procesamiento digital de imágenes es posible resaltar características que permiten al usuario tener una mayor información acerca de lo que se está captando la cámara.

Puede reconocerse así el siguiente **problema científico**:

¿Qué método de filtrado sería capaz de brindar un grado de robustez al sistema de visión del robot móvil Ricimaf-SjyV en lo relativo a las variaciones en la iluminación dentro de su área de trabajo?

Para dar respuesta al problema científico se plantea como **objeto de estudio** los filtros utilizados para el sistema de visión de un robot móvil, donde el **campo de acción** está enmarcado en los filtros para la visión de robots móviles donde la iluminación causa ruido.

Esta investigación tiene como **objetivo general**: Proponer un filtro para la eliminación del ruido causado por la iluminación en el sistema de visión del robot móvil Ricimaf-SjyV.

Para el logro del objetivo general se plantean **objetivos específicos**:

- Estudio de filtros existentes.
- Selección de filtros que satisfagan las necesidades planteadas para una posterior propuesta.

El desarrollo de este trabajo estará organizado de la siguiente manera:

Capítulo 1: Fundamentación Teórica. Se realiza un estudio del estado del arte y se introducen los principales conceptos relacionados con robótica, robot, método, píxel, histograma, tratamiento de imágenes, visión por computadora o visión artificial y sistemas de visión de un robot móvil. Además, se analizan las tendencias actuales para el desarrollo del trabajo.

Capítulo 2: Se expondrá el uso más práctico de las funciones comunes en preprocesado. Haciendo énfasis en filtros existentes, dentro de la etapa de preprocesamiento y segmentación. Se seleccionan aquellos filtros que más utilidad puede tener en las aplicaciones de visión computacional.

Capítulo 3: Se propone un nuevo filtro compuesto por varios filtros existentes ya estudiados, los cuales en conjunto pudieran ayudar a solucionar el problema del ruido causado por la iluminación en el sistema de visión del robot móvil Ricimaf-SjyV. Los filtros escogidos para conformar el filtro propuesto constan de características resaltadas con respecto a los otros filtros analizados.

CAPITULO 1: FUNDAMENTACION TEORICA.

1.1 Introducción

Este capítulo consta de dos epígrafes: Estado del Arte y los principales conceptos que están relacionados con la robótica, robot, método, píxel, histograma, tratamiento de imágenes, visión por computadora o visión artificial y sistemas de visión de un robot móvil.

1.2 Estado del Arte

Es hasta el momento cuando el decreciente costo del poder de computación acoplado a la capacidad de una mayor producción ha impulsado la introducción de mecanismos automáticos en las fábricas, pero además, han sido introducidos en otros campos como son: la medicina, la exploración y la biomedicina. El campo de procesamiento digital de imágenes en estéreo es un área de Ciencias de la Computación relativamente nueva e inexplorada, de reciente estudio e implementación ha sido muy escasa; la visión por computadora puede llegar a ser tan poderosa como para extenderse más allá del simple reconocimiento, por medio de ésta se pueden construir descripciones explícitas y con el significado de los objetos físicos de una imagen. El rango de aplicaciones de la visión por computadora ha sufrido en los últimos años un incremento considerable.

En los últimos años se ha concentrado un gran esfuerzo hacia las investigaciones de predicción de trayectorias y reconocimiento de ambiente en robótica y son muchas las universidades que manejan áreas de investigación orientadas en esta dirección por ejemplo:

En la universidad de Purdue dentro del laboratorio de visión Robótica "Robot Visión Laboratory" se trabaja en un robot conocido como Peter, en su primera versión (el ahora conocido old_peter) en el control fue usado un Motorola 68010-12Hz. con VxWorks y como elemento de adquisición de datos fueron usados dos cámaras en blanco y negro(fijas) y un sensor ultrasonido (Amador Gonz, 2008).

Otro trabajo desarrollado por esta misma universidad, es uno basado en redes neuronales que permiten mantener al robot centrado en la mitad de su trayectoria y genera secuencias de comandos cuando el robots necesita una corrección en su curso para evadir obstáculos. Únicamente es necesario proporcionarle al sistema modelos

Capítulo 1: Fundamentación Teórica

topológicos del espacio interior, los modelos de esa forma son fácilmente creados. De hecho el sistema se ayuda de lógica difusa para un mejor control del mismo. En el procesamiento de la imagen es usada la detección de bordes, el filtrado y la segmentación basada en la diferencia de grises. (Amador Gonz, 2008).

Una última mención acerca de esta universidad, la debemos situar en 1989 cuando realizaron el proyecto PSEIKI, que realmente fue ejercicio de la operación de máquina usando interpretación de imágenes, lo importante de este sistema, es que no es necesario diseñar un sistema de navegación basado en la visión sin hacer una interpretación completa de la imagen, necesaria para posicionar al robot (Amador Gonz, 2008).

En la universidad de Utah se ha desarrollado el robot Egor, este robot compañero de laboratorio cuenta con un subsistema de proximidad, está bien equipado con 24 sensores, 8 sensores infrarrojos, una cámara de video Panasonic en blanco y negro conectada a través de una tarjeta captadora de video Win/TV y como plataforma de procesamiento se encuentra una PC Notebook. En julio de 1994, Egor obtuvo la sexta posición en visión durante la “AAAI Mobile Robotic” en Seattle, Washington. (Amador Gonz, 2008).

Muchas otras universidades importantes en el extranjero han desarrollado y continúan con trabajos de investigación en el área de la visión y robótica, entre estas universidades se encuentra: la Universidad de Rochester y la Universidad de Western Australia.

El Instituto Tecnológico de Massachusetts (MIT) es una de las universidades con mayor camino y experiencia en cuanto autómatas se refiere, ellos han construido una gran variedad de robots. MIT trabaja en lo que son los algoritmos de visión con módulos de procesadores de información visual que es un hardware especializado para realizar operaciones de convolución de manera muy rápida (Amador Gonz, 2008).

En el campo del procesamiento digital de imágenes en dos dimensiones, se puede citar el trabajo realizado en la Universidad Iberoamericana de Puebla, en donde se estudian varios conceptos sobre el análisis digital de imágenes aplicado al desplazamiento de objetos. Específicamente se trabaja con el reconocimiento de contornos de objetos y en base a la posición del objeto detectado. Para el análisis de la imagen se usa

Capítulo 1: Fundamentación Teórica

primeramente la segmentación para realzar la región de interés, una vez hecho esto se obtiene el contorno mediante el operador Sobel del método de Gradiente, un proceso intermedio implementado fue el adelgazamiento de los bordes mediante el algoritmo MAT (Medial Axis Transform) y el análisis de formas fue implementado mediante la transformada Hough (Amador Gonz, 2008).

Otra universidad que se encuentra realizando trabajos en el área del procesamiento de imágenes y reconocimiento de objetos a partir de imágenes 2D es la Universidad de Cambridge, se encuentra realizando investigaciones en tópicos como son el reconocimiento de rostros y la segmentación de imágenes.

En Cuba se viene trabajando este tema en varias universidades y centros de investigación, con propósitos más o menos específicos. Aplicado a la robótica, solamente en la Universidad Central "Marta Abreu" de las Villas se han venido realizando una serie de trabajos para dos y tres dimensiones espaciales.

Aunque se han descrito numerosas técnicas de filtrado, existen muchas más que han aparecido últimamente o que van apareciendo día a día. Se describen brevemente algunas de ella :

El filtro Wiener se utiliza para mejorar imágenes borrosas aplicando una transformación de la imagen y su inversa para obtener la función de transferencia del filtro. Lo que se busca es minimizar una función de error del cuadro de la resta de una señal bidimensional aleatoria con una imagen estimada formada con la imagen con ruido convulcionada con un filtro lineal.

El método de interpolación Spline que trata de ajustar una curva a los píxeles observados y deducir de ella el ruido a partir de su segunda derivada.

Filtros adaptativos que usan métodos estadísticos para actuar en diferentes zonas de la imagen de forma distinta.

Existe una combinación muy usada, donde se usa el filtro Wiener adaptándose a cada parte de la imagen de esta forma se realiza un filtrado preservándose los bordes y otras partes de altas frecuencias de las imágenes.

Como se ve, las distintas técnicas de filtrado son numerosas y existe un campo muy amplio de la aplicación de éstas al tratamiento de imagen.

1.2.1 La robótica en la actualidad

Los robots son usados hoy en día para llevar a cabo tareas sucias, peligrosas, difíciles, repetitivas o agotadoras para los humanos. Esto usualmente toma la forma de un robot industrial usado en las líneas de producción. Otras aplicaciones incluyen la limpieza de residuos tóxicos, exploración espacial, minería, búsqueda y rescate de personas y localización de minas terrestres. La manufactura continúa siendo el principal mercado donde los robots son utilizados. En particular, robots manipuladores (similares en capacidad de movimiento a un brazo humano) son los más usados comúnmente. Las aplicaciones incluyen soldado, pintado y carga de maquinaria. La Industria automotriz ha tomado gran ventaja de esta nueva tecnología donde los robots han sido programados para reemplazar el trabajo de los humanos en muchas tareas repetitivas. Existe una gran esperanza, especialmente en Japón, de que el cuidado del hogar para la población de edad avanzada pueda ser llevado a cabo por robots. (Urdiales Garcia, 2007).

Recientemente, se ha logrado un gran avance en los robots dedicados a la medicina, con dos compañías en particular, Computer Motion e Intuitive Surgical, que han recibido la aprobación regulatoria en América del Norte, Europa y Asia para que sus robots sean utilizados en procedimientos de cirugía invasiva mínima. Desde la compra de Computer Motion (creador del robot Zeus) por Intuitive Surgical, se han desarrollado ya dos modelos de robot daVinci por esta última. En la actualidad, existen más de 800 robots quirúrgicos daVinci en el mundo, con aplicaciones en Urología, Ginecología, Cirugía general, Cirugía Pediátrica, Cirugía Torácica y Cirugía Cardíaca. La automatización de laboratorios también es un área en crecimiento. Aquí, los robots son utilizados para transportar muestras biológicas o químicas entre instrumentos tales como incubadoras, manejadores de líquidos y lectores. Otros lugares donde los robots están reemplazando a los humanos son la exploración del fondo oceánico y exploración espacial. Para esas tareas, robots de tipo artrópodo son generalmente utilizados. Mark W. Tilden del Laboratorio Nacional de los Álamos se especializa en robots económicos de piernas dobladas pero no empalmadas, mientras que otros buscan crear la réplica de las piernas totalmente empalmadas de los cangrejos. (Urdiales Garcia, 2007).

Capítulo 1: Fundamentación Teórica

Robots experimentales y otros ejemplos que explotan el biomimetismo también están en fases previas. Se espera que los así llamados nanomotores^{viii} simplifiquen drásticamente el poder de locomoción, mientras que la estabilización en vuelo parece haber sido mejorada substancialmente por giroscopios extremadamente pequeños. Un impulsor muy significativo de este tipo de trabajo es el desarrollar equipos de espionaje militar. (Urdiales Garcia, 2007).

Dean Kamen, fundador de FIRST (Foundation for the Inspiration and Recognition of Science and Technology) y de la Sociedad Americana de Ingenieros Mecánicos (ASME), ha creado una Competencia Robótica multinacional que reúne a profesionales y jóvenes para resolver un problema de diseño de ingeniería de una manera competitiva. En 2003 contó a más de 20.000 estudiantes en más de 800 equipos en 24 competencias. Los equipos vienen de Canadá, Brasil, Reino Unido y Estados Unidos. A diferencia de las competencias de los robots de lucha sumo, que tienen lugar regularmente en algunos lugares, o las competencias de “Battlebots” transmitidas por televisión, estas competencias incluyen la creación de un robot. (Urdiales Garcia, 2007).

Los robots parecen estar abaratándose y empequeñeciéndose en tamaño, todo relacionado con la miniaturización de los componentes electrónicos que se utilizan para controlarlos. También, muchos robots son diseñados en simuladores mucho antes de que sean construidos e interactúen con ambientes físicos reales. Un buen ejemplo de esto es el equipo Spiritual Machine, un equipo de 5 robots totalmente desarrollado en un ambiente virtual con el fin de jugar al fútbol en la liga mundial de la F.I.R.A (Federation of International Robosoccer Association) (Urdiales Garcia, 2007).

1.2.2 Límites de la robótica actual

Michael Knasel, director del Centro de Aplicaciones Robóticas de Science Application Inc., explicó en la exposición Robots-8 en Detroit (1984) que los robots deberán evolucionar durante cinco generaciones. Las dos primeras ya se habían manifestado en los ochenta e incluían la gestión de tareas repetitivas, pero estaban muy limitadas en cuanto a movimiento se refiere. La tercera generación incluiría visión artificial, la cuarta movilidad avanzada en exteriores e interiores y la quinta entraría en el dominio de la inteligencia artificial. Lamentablemente, si bien se han cumplido algunas de estas expectativas, aún no se han conseguido modelos lo suficientemente efectivos.

Engelberger aseguraba que si bien los sistemas sensoriales por visión y tacto, se limitan a aplicaciones muy básicas y específicas con bajo nivel de detalle debido a la enorme carga computacional inherente al procesado de imagen (Urdiales Garcia, 2007).

1.2.3 Grupo de Investigación de Automática Aplicada

En nuestra universidad se viene realizando un grupo de trabajos como parte de la línea de desarrollo de hardware, robótica y control inteligente, del Grupo de Investigación de Automática Aplicada. Este grupo tiene como misión definir la estrategia investigativa, asesorar metodológicamente las investigaciones y obtener resultados investigativos en los temas de automática vinculados a los proyectos productivos y al desarrollo tecnológico de la universidad en esa rama, con el fin de garantizar la relación investigación-desarrollo a través de la organización de las investigaciones y de la socialización y aplicación de los resultados obtenidos (Boudy Gonzales & Leyva Fonseca, 2008).

1.2.4. Departamento de Telecomunicación y Electrónica.

En la Universidad Central “Marta Abreu” de las Villas se han venido realizando una serie de trabajos para dos y tres dimensiones espaciales, en el Departamento de Telecomunicación y Electrónica perteneciente a esta universidad, el cual se destaca en el procesamiento digital de imágenes se describe un algoritmo de reducción del ruido aleatorio gaussiano por medio de la transformada de Wavelet. Se comparan los resultados obtenidos a partir de un conjunto de filtros diseñados por esta vía con el objetivo de seleccionar aquellos que brinden mejores resultados, a partir de la evaluación de métricas cuantitativas de la calidad de imagen (Cardoso, 2008).

1.2.5 Departamento de Control Automático del Instituto de Cibernética, Matemática y Física (ICIMAF).

El Instituto de Cibernética, Matemática y Física (ICIMAF) fundado en 1964, es un Instituto de Investigaciones de alto nivel teórico y aplicadas en ramas de la Cibernética, la Matemática y la Física.

A finales de 1963 se formó un colectivo de investigación en Matemática. Este grupo de trabajo estaría formado también por especialistas en Cibernética Técnica. Con

Capítulo 1: Fundamentación Teórica

posterioridad se fueron ampliando las líneas de investigación hasta conformar las que hoy existen en la institución.

El Instituto con la denominación actual surge en 1968 al fundirse dos instituciones que fueron creadas en todo este proceso de diversificación de las líneas de trabajo. Instituto de Matemática, Cibernética y Computación (IMACC). Instituto de Investigaciones Técnicas Fundamentales (ININTEF). El actual ICIMAF integra a especialistas de alto nivel científico y un conjunto de técnicos de sobrada experiencia, que abordan en sus temáticas de trabajo, tanto aspectos teóricos como prácticos en ramas básicas del conocimiento a saber: la Matemática, la Física Teórica, la Ultrasónica, el Control Automático y las Ciencias de la Computación en general.

Tiene como misión desarrollar investigación de carácter teórico y aplicado en las áreas de Cibernética, Matemática y Física como institución perteneciente al Ministerio de Ciencias, Tecnología y Medio Ambiente.

Desde su fundación hasta la fecha, el ICIMAF ha desplegado una amplia labor en los campos del saber de su competencia. Hay muchos resultados en investigaciones teóricas y aplicadas que prestigian la institución (ICIMAF).

1.3 Principales conceptos.

1.3.1 Robótica

Es una ciencia o rama de la tecnología, que estudia el diseño y construcción de máquinas capaces de desempeñar tareas realizadas por el ser humano o que requieren del uso de inteligencia. Las ciencias y tecnologías de las que deriva podrían ser: el álgebra, los autómatas programables, las máquinas de estados, la mecánica o la informática.

2.2.1 Robot

Es un dispositivo generalmente mecánico, que desempeña tareas automáticamente, ya sea de acuerdo a supervisión humana directa, a través de un programa predefinido o siguiendo un conjunto de reglas generales, utilizando técnicas de inteligencia artificial. Generalmente estas tareas reemplazan, asemejan o extienden el trabajo humano, como ensamble en manufactura, manipulación de objetos pesados o peligrosos, trabajo en el

espacio.

Los procesos de razonamiento seleccionan las acciones que se deben tomar para realizar cierta tarea encomendada. La habilidad de razonamiento permite el acoplamiento natural entre las habilidades de percepción y acción.

1.3.3 Método

Es una palabra que proviene del término griego *methodos* (“camino” o “vía”) y que se refiere al medio utilizado para llegar a una solución requerida. Su significado original señala el camino que conduce a obtener resultados.

1.3.4 Píxel

Es la unidad más pequeña de una imagen digital, la mínima resolución, el detalle más pequeño. Es esa celdilla que compone el mosaico de una imagen. Así funcionan las televisiones y los monitores, así se recogen y capturan las imágenes digitales en píxeles.

1.3.5 Imagen

Es una representación óptica de un objeto iluminado por una fuente radiante.

1.3.6 Histograma

Básicamente el histograma es la representación gráfica de las frecuencias relativas con las que aparecen los distintos niveles de grises en una imagen determinada. En general se representa como un gráfico de barras, en el que las abscisas son los distintos niveles de grises de la imagen y las ordenadas la cantidad de veces con la que cada color aparece en la imagen. El histograma proporciona información sobre el brillo, contraste y colores de la imagen, y puede ser utilizado para ajustar estos parámetros, eliminar ciertas tonalidades molestas, etc. El histograma de una imagen se puede definir como una función discreta que representa el número de píxeles en la imagen en función de los niveles de intensidad.

1.3.7 Tratamiento de imágenes

El procesamiento digital de imágenes (PDI) es el conjunto de técnicas que se aplican a las imágenes digitales con el objetivo de mejorar el aspecto de las imágenes y hacer

Capítulo 1: Fundamentación Teórica

más evidentes en ellas ciertos detalles que se desean hacer notar. El procesamiento de las imágenes se puede hacer por medio de métodos ópticos, o bien por medio de métodos digitales, en una computadora.

El procesamiento digital de imágenes se divide en cuatro etapas fundamentales. (Boudy Gonzales & Leyva Fonseca, 2008)(Ver Figura1)

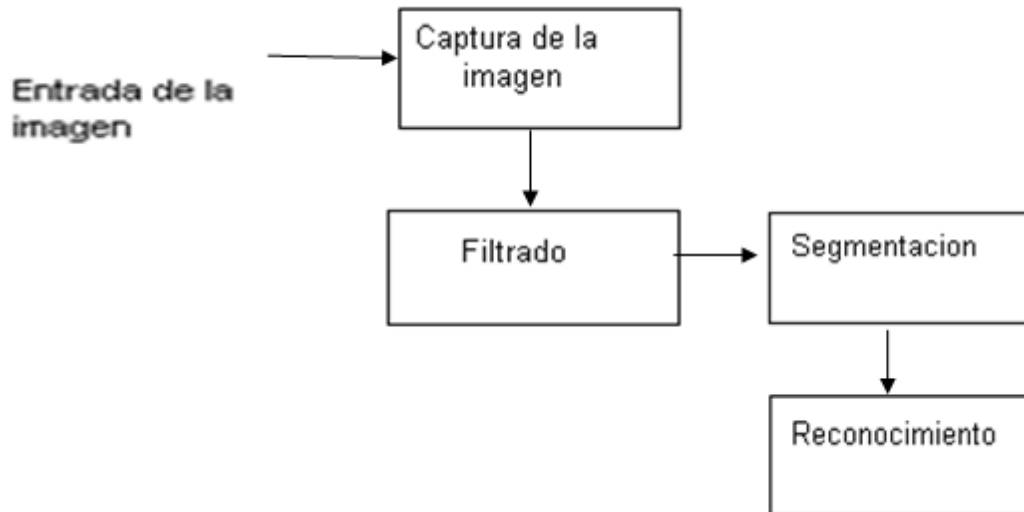


Fig 1 Etapas del procesamiento digital de imágenes.

1. **Captura:** es donde se obtiene la imagen a procesar desde un dispositivo óptico conectado a la computadora.
2. **Filtrado:** en esta etapa se transforma la imagen a un formato más adecuado para facilitar posteriores procedimientos en el análisis de la escena (eliminación de información no relevante, aumento del contraste, aumento o disminución del rango dinámico, entre otros).
3. **Segmentación:** partición de la imagen en un conjunto de regiones no solapadas, homogéneas con respecto a algún criterio. Esta constituye la fase de mayor importancia antes de llevarse a cabo los pasos siguientes del análisis de la imagen. De su calidad depende en gran medida el resultado final de su procedimiento.
4. **Reconocimiento:** tiene como objetivo principal detectar a partir de determinadas características conocidas, los patrones visuales presentes en una imagen. Un patrón no es más que una descripción estructural o cuantitativa de un objeto o de alguna otra entidad de interés en la imagen.

En este trabajo la etapa de filtrado es la fundamental, ya que lo que se quiere es obtener un filtro capaz de eliminar los problemas existentes.

1.3.8 Visión por computadora

Los humanos obtenemos la mayor parte de nuestra entrada sensorial a través de nuestro sistema visual, y se ha hecho un enorme esfuerzo para mejorar de forma artificial este sentido. Lentes, binoculares, telescopios, radares, sensores infrarrojos, todos funcionan para mejorar nuestra vista del mundo y del universo.

La visión por computadora tiene como objetivo duplicar el efecto de la visión humana percibiendo y entendiendo una imagen por medios electrónicos. Utiliza los resultados y métodos de las matemáticas, reconocimiento de patrones, inteligencia artificial, ciencias de la computación, electrónica y otras disciplinas científicas para conseguir su objetivo. (Ver Figura 2)

Dificultades:

- Condiciones de iluminación variable y no controlada (con sombras)
- Objetos complejos o difíciles de describir
- Objetos ocultos (escenas de interiores o con objetos no rígidos)

Proyección en perspectiva es una transformación de muchos a uno:

- Diferentes escenas pueden producir imágenes idénticas.
 - Pueden contener ruido:
- Luz ambiental.
- Otros factores.

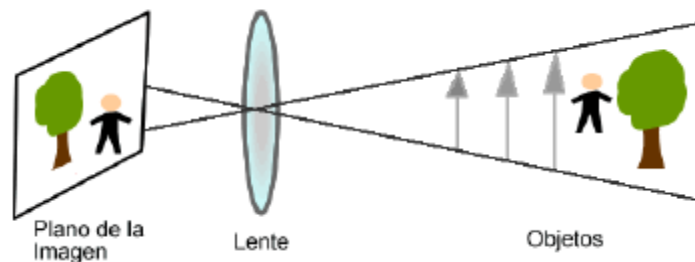


Fig 2 Objeto visto por un lente.

1.3.9 Visión Artificial

El término visión artificial o visión por computadora se refiere a la capacidad de una máquina computadora de reconocer su entorno a partir de dispositivos que le permitan recolectar datos referentes al mismo, y de algoritmos capaces de interpretarlos convenientemente. Entre los más utilizados está el reconocimiento de patrones en imágenes digitales.

1.4 Conclusiones

En este capítulo se estudia lo referente al estado del arte y a los conceptos fundamentales para el conocimiento previo del tema de procesamiento de imágenes para el sistema de visión de un robot móvil.

CAPITULO 2: ESTUDIO DE FILTROS EN LAS DISTINTAS ETAPAS DE PROCESAMIENTO DE IMÁGENES.

2.1 Introducción

En este capítulo se expondrán funciones comunes en el preprocesado de imágenes digitales. Haciendo énfasis en filtros existentes, dentro de la etapa de preprocesamiento y segmentación. De los filtros vistos se han seleccionado aquellos que más utilidad puede tener en las aplicaciones de visión computacional. La descripción de estos filtros no pretende ser muy exhaustiva, si no, que se basará en el uso práctico y aplicaciones para la eliminación de ruido.

2.2 Procesamiento de Imágenes Digitales

Visión artificial o visión de máquina no es más que la adquisición, análisis y preparación automática de imágenes digitales, con el fin de extraer información relevante del entorno.

El procesamiento de las imágenes, en aras de facilitar el trabajo y mejorar la eficiencia de la aplicación, se ha dividido en 4 fases, en este capítulo se hace referencia solo a las tres primeras etapas, se profundiza en la segunda y tercera etapa, la de filtrado y segmentación, sin dejar de mencionar y explicar brevemente la etapa de captura de la imagen (Ver Figura 3).

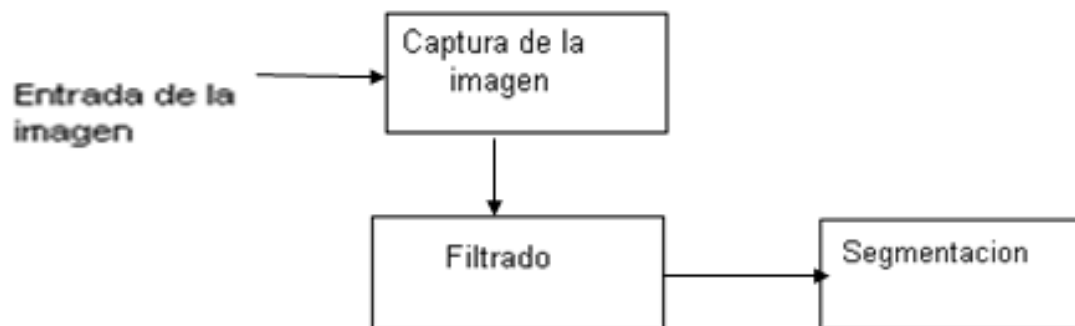


Fig 3 Etapas 1,2,3 de Procesamiento de imágenes.

2.2.1 Captura de la imagen

Se obtiene una imagen desde una cámara digital (webcam), estando la cámara, al igual que el ojo humano, formada por un lente convergente que proyecta la imagen sobre una superficie sensible a la luz denominada sensor de imagen. (Ver Figura 4).

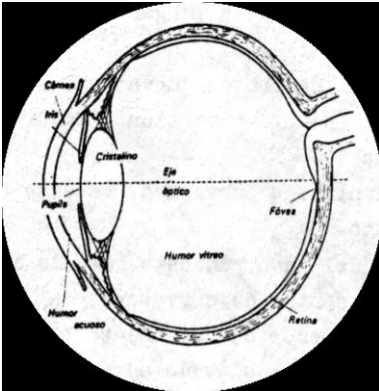


Fig 4 Ojo humano.

Sensores digitales de imagen

Los sensores digitales de imagen están formados por una serie de elementos foto sensores (sensibles a la luz) que modifican su señal eléctrica según la intensidad luminosa que reciben, lo que permite la captura de los puntos que conforman la imagen. Estos sensores suelen estar configurados en forma matricial de modo que proporcionan una imagen bidimensional.

Formación y Representación de la Imagen

Una imagen es una representación visual de un objeto iluminado por una fuente radiante. Las que se perciben en las actividades visuales cotidianas provienen normalmente de la luz reflejada por los objetos. La naturaleza básica de una imagen, representada por $f(x, y)$ elevado a 2 está caracterizada por dos componentes: la cantidad de luz incidente que proviene de la fuente de la escena contemplada y la cantidad de luz reflejada por los objetos de la escena. Dichas componentes reciben el nombre de iluminación y reflectancia^x.

Capítulo 2: Estudio de filtros en las distintas etapas del PDI

En el proceso de formación de la imagen intervienen los siguientes elementos: el objeto, la fuente radiante y el sistema de formación de la imagen, que consiste básicamente en un sistema óptico, un sensor y un digitalizador.

Un aspecto importante en la etapa de adquisición de la imagen son los dispositivos a utilizar, los cuales están en dependencia a los objetivos que se persigan sobre la imagen a analizar. Dentro de los dispositivos más importantes que conforman el sistema de adquisición se encuentran:

- **Sensores:** Dispositivos que censan las radiaciones electromagnéticas, las cuales pueden estar en el espectro visible, ultravioleta, de rayos x.
- **Captadores:** Dispositivos utilizados una vez que la imagen fue censada. Los más usados son las cámaras de televisión, las cámaras fotográficas (en la actualidad las digitales juegan un papel importante), los exploradores (scanner). Todos ellos pueden captar la información en nivel de gris o en colores.
- **Digitalizadores:** Dispositivos que transforman las imágenes captadas analógicamente en digitales.

Mejoramiento de una imagen

El objetivo fundamental del mejoramiento digital de imágenes (MDI) es, como su nombre lo indica mejorar la apariencia de una escena para la vista humana, o detectar información adicional acerca de un objeto que no es observado con claridad en la imagen original.

¿Cuándo se dice que una imagen está mejorada?

El término MDI tiene un carácter netamente subjetivo, en el sentido de que dos o más observadores pueden tener criterios diferentes para una misma imagen supuestamente mejorada. En efecto, lo que para el observador puede ser mejoramiento para otro puede ser ruido.

En el MDI cabe distinguir, entre otro, los aspectos siguientes: mejoramiento del contraste (MC), mejoramiento de bordes o contornos (MB), acentuación de determinadas frecuencias espaciales, etc.

2.2.1.1 Espacios de color

Un espacio de color es un modelo para la representación matemática de un conjunto de colores. Se les llama espacio de color porque son representaciones en 2 dimensiones. Cada punto dentro del espacio corresponde a un color. Aunque existen multitud de modelos de color, según su finalidad, el RGB está entre los modelos típicos empleados en procesamiento digital.

RGB

La descripción RGB (del inglés Red, Green, Blue; “rojo, verde, azul”) de un color hace referencia a la composición del color en términos de la intensidad de los colores primarios con que se forma: el rojo, el verde y el azul. 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 luz primarios. Indicar que el modelo de color RGB no define por sí mismo lo que significa exactamente rojo, verde o azul, razón por la cual los mismos valores RGB pueden mostrar colores notablemente diferentes en distintos dispositivos que usen este modelo de color. Aunque utilicen un mismo modelo de color, sus espacios de color pueden variar considerablemente (Ver Figura 5).



Fig 5 Mezcla aditiva de colores.

Muchos dispositivos modernos permiten realizar el registro en color, de tal forma que para cada zona de la imagen se genera un píxel con tres componentes, la representación común es mediante un vector formado por una combinación de colores básicos utilizados en la electrónica de video.

Capítulo 2: Estudio de filtros en las distintas etapas del PDI

Definiremos en el espacio RGB un color “gris” aquel que tiene sus tres componentes iguales de tal manera que todo gris tendrá la forma: $Z = (z, z, z)$.

Es decir, la proporción de cada uno de los colores básicos es la misma. Por tanto podemos afirmar que los grises se ubican en la recta que va del origen (0, 0,0) al punto (255, 255,255).

Se denomina resolución al número de píxeles por pulgadas utilizados en el proceso de adquisición, ésta se indica en dpi (dots per inch = puntos por pulgada). Mientras mayor es la resolución mayor es la fidelidad de la imagen, es decir, se obtienen más detalles de ella. Los valores de la resolución varían de 75 dpi hasta algunos miles (7200 dpi). Generalmente este es un parámetro del dispositivo de adquisición que es posible definir durante el proceso de registro.

El modelo HSI representa los valores de matiz, saturación e intensidad. La matiz representa la fuerza del color, la saturación representa el grado con el que el color se diluye en luz blanca, y la intensidad representa la característica de ser prácticamente independiente de la iluminación, y se obtiene como la suma de los valores RGB ($I = R + G + B$).

Así, mientras que una imagen RGB presenta los colores que deben ser directamente representados en el monitor, una imagen HSI resulta apropiada para obtener información independientemente de las condiciones de iluminación.

Hardware de Color

El espacio de color usado por el robot es YCbCr, el cual se puede observar en la figura a continuación. **Y** es la componente de luminancia y **Cb** y **Cr** de crominancia.^x YCbCr es una forma común de codificar el espacio RGB. Cualquier color RGB puede transformarse en YCbCr y viceversa. Todos los filtros deben establecerse definiendo los límites superiores e inferiores de las tres componentes de color. Cada píxel con todos sus valores Y, Cb y Cr dentro de los límites de un filtro F es asignado a dicho filtro. Un píxel puede pertenecer a uno, ninguno e incluso a varios filtros (Ver Figura 6).

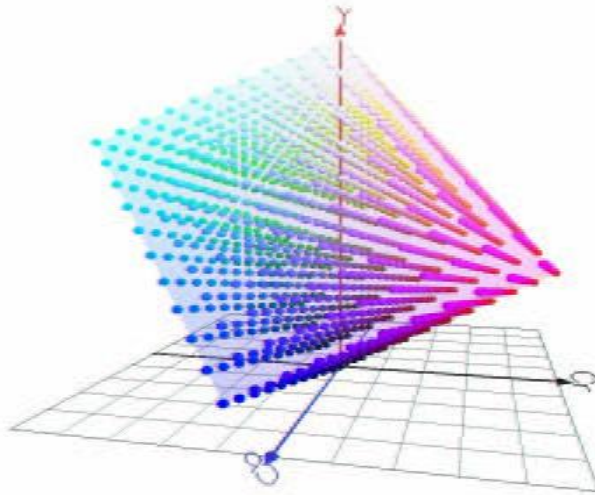


Fig 6 Espacio de color Y Cb Cr.

Se pueden establecer tres parámetros de la cámara:

- Balance de Blancos (Modo Interior, Exterior o Fluorescente)
- Ganancia de la cámara (Baja, Media o Alta)
- Velocidad de obturación (Lenta, Media o Rápida)

La ganancia de la cámara suele establecerse con el valor alto y no se obtienen cambios importantes usando los diferentes valores. La velocidad de obturación es un parámetro importante y su valor óptimo depende del brillo o luminosidad del entorno. Con una luminosidad de entorno baja o media, la velocidad de obturación debe ser rápida, con una elevada luminosidad la velocidad debe reducirse a un valor medio, y solamente con una iluminación extremadamente alta, la velocidad debe ser baja. A mayor velocidad de obturación, las imágenes obtenidas por la cámara son más sensibles a brillos o valores altos de luminosidad. Los tres parámetros de balance de blancos son los adecuados para los entornos de exterior, interior y habitaciones de interior iluminadas con luces fluorescentes. El modo de balance de blancos exterior obtiene imágenes azuladas algo alejadas de las imágenes reales, por lo que no ha sido usado. Tanto el balance de blancos de interior como fluorescente obtienen buenos resultados para entornos de interior, pero seleccionamos el balance de blancos fluorescente por obtener mejores resultados dentro del entorno seleccionado. Todos los experimentos fueron realizados con un valor de ganancia alto y con balance de blancos de cámara fluorescente.

Capítulo 2: Estudio de filtros en las distintas etapas del PDI

2.2.2 Filtrado o pre –procesado de imágenes

El objetivo principal de la etapa de filtrado o pre-procesamiento es una mejora de los datos de la imagen suprimiendo los efectos de ruido e incrementando las características deseadas para un exitoso procesamiento.

Toda imagen que se adquiere por medios ópticos, electroópticos o electrónicos, sufre en cierta medida de efectos de degradación que se manifiestan en forma de ruido, pérdida de definición y fidelidad de la imagen. La degradación viene provocada por el ruido de los sensores de captura, imprecisiones en el enfoque de la cámara, movimiento de la misma o perturbaciones aleatorias. Los mecanismos que tratan de contrarrestar estos efectos se incluyen dentro de la etapa de pre-procesado, recibiendo el nombre de **operaciones de restauración**.

Generalmente el pre-procesado o filtrado pretende reparar en la imagen los desperfectos producidos o no eliminados por el hardware: deformación de ésta, ruido introducido, poco o mucho contraste o brillo, falta de ecualización apropiada, etc. Los mecanismos de pre-procesado permiten modificar la imagen para eliminar ruido, mejorar la intensidad o el contraste; procesos que intentan mejorar el resultado final de la imagen capturada. Por supuesto con estos algoritmos necesitan una gran cantidad de tiempo de filtrado, es lógico que lo mejor sería utilizar un conveniente hardware que los evite.

En un proceso de visión artificial estos algoritmos tienen que ser utilizados lo menos posibles, un uso excesivo de ellos repercutirá en el tiempo de proceso total e indicará que la calibración, iluminación y selección de los elementos de la etapa de adquisición no ha sido la adecuada.

A parte de la degradación de la imagen, existen características de la imagen que en muchos casos es conveniente mejorar. Así muchas veces se quiere mejorar el contraste, brillo, niveles de grises, eliminar brillo, aumentar los bordes, mejorar las texturas, etc. Todas estas técnicas vienen a denominarse **operaciones de mejora de la imagen**.

La parte de filtrado de un sistema de visión artificial está formada por estos dos tipos de operaciones, es decir, todo el tratamiento de corrección y de mejora de la imagen que facilite el pre-procesado de las siguientes etapas.

Capítulo 2: Estudio de filtros en las distintas etapas del PDI

En resumen el Pre-procesamiento o Filtrado, como bien le podríamos llamar es un método para resaltar o suprimir, de forma selectiva información contenida en una imagen a diferentes escalas espaciales, para destacar algunos elementos de la imagen, o también para ocultar valores anómalos. Este caso en ocasiones es crucial para lograr los objetivos propuestos sobre la imagen. El rol fundamental de este paso es la transformación más adecuada de la imagen en aras de facilitar los posteriores procedimientos en el análisis de la escena. Esta etapa comprende varios tipos de procedimientos como son el mejoramiento y restauración.

Entre las técnicas más comunes usadas en la etapa de mejoramiento están:

- Eliminación o atenuación del ruido.
- Aumento de contraste.
- Disminución o aumento del rango dinámico.
- Detección de realce de los bordes, entre otros.

El proceso de filtrado consiste en la aplicación a cada uno de los píxeles de la imagen de una matriz de filtrado de tamaño $N \times N$ (generalmente de 3×3 aunque puede ser mayor) compuesta por números enteros y que genera un nuevo valor mediante una función del valor original y los de los píxeles circundantes.

Una diferencia importante entre las técnicas de filtrado y las de visualización, es que el filtrado modifica necesariamente la imagen ya que pretenden extraer información cuantitativa que debe almacenarse como una nueva capa en lugar de como una paleta de colores.

Los filtros más utilizados son los de paso bajo (suavizan la imagen), de paso alto (aumentan el contraste), los filtros direccionales (detectan en la imagen estructuras que siguen una determinada dirección) y los de detección de bordes (permiten identificar y aislar objetos con propiedades homogéneas dentro de la imagen).

2.2.2.1 Ruido



Fig 7 Ejemplo de imagen sin Ruido.

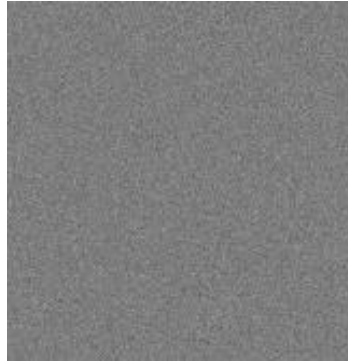


Fig 8 Ejemplo de Ruido



Fig 9 Ejemplo de imagen con Ruido

Se denota por ruido cualquier entidad en las imágenes, datos o resultados intermedios que no son interesantes para la computación que se pretende llevar a cabo, es decir, cualquier valor de un píxel de una imagen que no se corresponde exactamente con la realidad. Cuando se adquiere una imagen digital está contaminada por ruido. El ruido se debe, la mayoría de las veces al equipo electrónico utilizado en la captación de las imágenes (ruido de cuantificación de la imagen, efecto de niebla en la imagen, etc.) y al ruido añadido en los tramos de transmisión (posibles interferencias o errores al transmitir los bits de información) (Ver Figura 7, 8, 9).

Vamos a distinguir dos clases diferentes de ruido:

Ruido Gaussiano

Se caracteriza por tener un espectro de energía constante para todas las frecuencias. Cuando se presenta este problema, el valor exacto de cualquier píxel es diferente cada vez que se captura la misma imagen. Este efecto, suma o resta un determinado valor al nivel de gris real y es independiente de los valores que toma la imagen.

Capítulo 2: Estudio de filtros en las distintas etapas del PDI

Ruido Impulsivo

Se caracteriza por la aparición de píxeles con valores arbitrarios normalmente detectables porque se diferencian mucho de sus vecinos más próximos. La distribución viene dada por:

$$g(x,y) = \begin{cases} 0 & \text{si } r(x,y) < p/2 & \text{(VII)} \\ L - 1 & \text{si } p/2 \leq r(x,y) < p & \text{(VIII)} \\ \langle td \rangle f(x,y) & \text{si } r(x,y) \geq p & \text{(IX)} \end{cases}$$

Donde $r(x, y)$ es un número aleatorio con distribución uniforme en $(0,1)$ y p es la probabilidad de ocurrencia del ruido aleatorio, es decir, el porcentaje de puntos de la imagen que se verán afectados por el ruido impulsivo del total de puntos de la imagen.

El ruido tipo sal y pimienta (Salt and Pepper) es un tipo especial de ruido impulsivo, que presenta un punteado característico que lo hace fácilmente identificable y que se elimina mediante el empleo del filtro de mediana.

El ruido Gaussiano tiene un efecto general en toda la imagen, es decir, la intensidad de cada píxel de la imagen se ve alterada en cierta medida con respecto a la intensidad en la imagen original. Por el contrario, se observa que el ruido impulsivo tiene un efecto más extremo sobre un subconjunto del total de píxeles de la imagen. Un tanto por ciento de los píxeles de la imagen toma arbitrariamente el valor extremo 0 o 255.

Una forma de eliminar el ruido de una imagen es mediante los métodos de filtrado, en este trabajo nos vamos a centrar en el ruido en el sistema de visión de un robot móvil causado por la iluminación que es lo que se quiere eliminar.

2.2.2.2 Filtrado

Tal como se ha visto en el apartado anterior, existen diversos tipos de ruidos que se producen por múltiples circunstancias: los sensores, ruidos eléctricos, perturbaciones en el medio de transición, efectos térmicos, campos electromagnéticos, etc. La mayoría de las implementaciones de los filtros se utilizan en dos dominios: el dominio espacial y el dominio frecuencial.

Capítulo 2: Estudio de filtros en las distintas etapas del PDI

Los métodos basados en dominio espacial hacen referencia a la manipulación directa de la luminancia de los píxeles, mientras que los basados en dominio frecuencial modifican indirectamente la luminancia de cada píxel en la imagen.

Técnicas de filtrado para la eliminación de ruido

Se trata de métodos para resaltar o suprimir, de forma selectiva, información contenida en una imagen a diferentes escalas espaciales, para destacar algunos elementos de la imagen, o también para ocultar valores anómalos.

Los filtros se pueden dividir en:

Filtro de color y Post-Procesamiento

El filtrado hardware de color es una vía rápida para dividir la imagen original en imágenes filtradas por color, pero esta propuesta posee algunos problemas. El filtrado se realiza de forma determinista, sin tener en cuenta los píxeles de alrededor. Por tanto, parece necesario un nuevo procesamiento que tenga en cuenta los píxeles cercanos, pero que mantenga unos bajos requisitos computacionales. Para llevar a cabo este procesamiento necesitamos realizar dos pasos, de forma que el primero realiza el filtrado de forma determinista, y el segundo procesa cada píxel usando los píxeles circundantes. Este paso elimina los píxeles que han superado un filtro pero que no poseen ningún píxel alrededor que también lo haya superado, y añade píxeles no seleccionados por el filtro pero circundados por una gran cantidad de píxeles que si lo han superado. El procesamiento seleccionado examina cada píxel y los píxeles cercanos usando una rejilla cuadrada de tamaños 1, 2 y 3. Una rejilla de tamaños igual a 2 (5 x 5) se muestra en la figura a continuación (Ver Figura10) donde las casillas sombreadas se corresponden con píxeles que han superado el filtro. Cada píxel posee un número, el cual es la diferencia absoluta entre el píxel actual y el píxel central que esta siendo examinado. Esta distancia absoluta se usa para obtener un valor asociado a cada píxel utilizando los píxeles cercanos.

Capítulo 2: Estudio de filtros en las distintas etapas del PDI

4	3	2	3	4
3	2	1	2	3
2	1	0	1	2
3	2	1	2	3
4	3	2	3	4

Fig 10 Rejilla 5 x 5.

Para obtener un valor que nos indique si un píxel está rodeado o no por una gran cantidad de píxeles se define el valor V . Una vez somos capaces de obtener un valor V por píxel, debemos definir un límite superior que determine cuando añadir un píxel que no haya superado un filtro y un límite inferior que decida cuando eliminar el píxel de la muestra de píxeles que han superado el filtro. Los límites deben ser establecidos con cautela, y sus valores pueden ser altamente restrictivos con el objetivo de evitar añadir o eliminar píxeles. Si los límites se configuran correctamente, de forma que eliminen correctamente píxeles ruidosos, los filtros de color se pueden establecer de forma menos restrictiva, con el objetivo de facilitar la detección de objetos.

El uso de este planteamiento nos permite obtener unas imágenes resultado que no estén tan influenciadas por el ruido como las originales. La figura 11 muestra el proceso completo de filtrado para el color azul, donde solamente los píxeles entre los límites superiores e inferiores de los filtros pasan el primer filtrado, pero que tras el paso adicional usando una rejilla de 3x3, se obtiene la imagen inferior derecha, la cual no posee píxeles que realmente no pertenecen a la portería azul (Ver Figura 11).

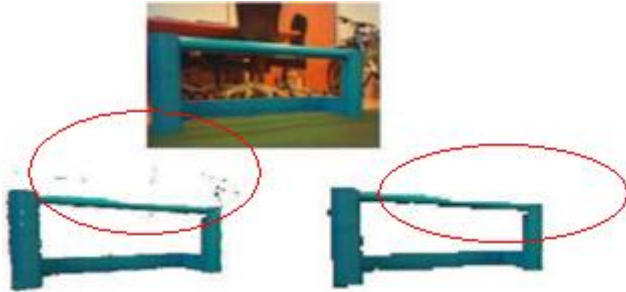


Fig 11 Terreno de fútbol visto por un robot.

Este filtro presenta algunos problemas ya que el filtrado se realiza de forma determinista, sin tener en cuenta los píxeles alrededor de la imagen. Este razonamiento determinista no es el adecuado para entornos ruidosos, dentro de los cuales el valor obtenido de los colores puede estar distorsionado y algunos píxeles que no deberían pasar los filtros los superarían y viceversa.

Filtro basado en Correlación

El objetivo principal de este algoritmo es seleccionar como modelo patrón, un objeto presente en una primera imagen en escala de grises de la escena, para posteriormente, por medio de correlación, realizar un seguimiento del móvil en las imágenes subsecuentes.

Para tal fin, se implementó un algoritmo, que para la correlación tiene en cuenta al modelo no solamente en la orientación original sino también, en posiciones rotadas, con lo cual el modelo solo se almacenará en memoria una vez. El algoritmo utiliza un valor de aceptación para tomar o rechazar un resultado, el cual está muy ligado a las condiciones de iluminación.

Inicialmente se lee una imagen en escala de grises que ha sido pre-procesada. Seguidamente se pregunta si el modelo ya está almacenado en memoria o no. Si la respuesta es afirmativa se ejecuta la función de correlación para buscar dicho modelo en la imagen adquirida, de lo contrario se visualiza la imagen con el propósito que el usuario pueda seleccionar por medio del clic izquierdo del ratón el centro del modelo. Finalmente si se ha seleccionado satisfactoriamente un modelo, éste se guarda, se habilita la función de correlación angular, la cual tiene en cuenta los movimientos

Capítulo 2: Estudio de filtros en las distintas etapas del PDI

rotacionales del móvil y se fija un valor umbral con el cual se decide si un resultado es aceptable o no.

Filtro de convolución

Se denomina convolución a una función, que de forma lineal y continua, transforma una señal de entrada en una nueva señal de salida. La función de convolución se expresa por el símbolo $*$.

Una parte muy importante de los procesos de mejora se basan en la aplicación de funciones, cuyo resultado depende únicamente de los niveles de intensidad de cada píxel de la imagen, pero no de la posición dentro de la imagen (position-invariant), que permiten resaltar determinados elementos de la imagen (como resaltar los contornos de los objetos), mejorar el enfoque o reducir el ruido de fondo.

Estas funciones, a las que por analogía con la teoría de señales se las denomina comúnmente "filtros", se basan en el proceso de convolución y en las propiedades de la convolución respecto la transformación de Fourier, lo que simplifica el cálculo matemático y los tiempos de computación. Por esto, a muchos de estos filtros, se les solía denominar según su acción en el espacio de las frecuencias (filtros de baja o alta frecuencia, filtros direccionales). Sin embargo, cada vez es más frecuente, en los sistemas de proceso de imagen, adjudicarles nombres más de acuerdo con el resultado visual que producen sobre la imagen "filtrada" (filtros de enfoque, desenfoque).

Un filtro de convolución, para una imagen digital, en el espacio real (X, Y) , puede representarse como una matriz cuadrada o rectangular (matriz de convolución), de dimensiones (M, N) mucho más pequeñas que las de la imagen. La matriz de convolución se desplaza sobre la imagen de tal forma que el elemento central de la matriz de convolución coincida con cada uno de los píxeles de la imagen. En cada posición, se multiplica el valor de cada píxel de la imagen, que coincide en posición con un elemento de la matriz de convolución, por el valor de éste. El píxel de la imagen, que coincide con el elemento central de la matriz de convolución, es substituido por la suma de los productos.

Capítulo 2: Estudio de filtros en las distintas etapas del PDI

1/9 1/9 1/9

1/9 1/9 1/9

Así por ejemplo, si tenemos la matriz anterior de convolución: el resultado de su aplicación será sustituir cada píxel de la imagen por el promedio de dicho píxel con los ocho píxeles de su inmediato alrededor.

1/36	1/36	1/36	1/36	1/36
1/36	2/36	2/36	2/36	1/36
1/36	2/36	4/36	2/36	1/36
1/36	2/36	2/36	2/36	1/36
1/36	1/36	1/36	1/36	1/36

En otro ejemplo, la matriz produciría también un efecto de promedio, pero ponderando el valor de cada elemento respecto a la distancia con el elemento central. La única condición que deben cumplir estos filtros es que la suma de todos los valores debe ser uno. También puede atribuírsele unas propiedades de interpolación con efecto ponderado de la distancia. El efecto de estos filtros puede experimentarse utilizando el editor de filtros y la función filtro usuario de los programas MIP y Tecnomip de D.I.S. Tiene como problema un alto coste computacional, debido al gran número de cálculo matemático y los tiempos de computación que necesita para llegar a un resultado.

Filtro de bajas frecuencias (paso bajo) o filtro de desenfoque y reducción de ruido de fondo o también denominado filtro de interpolación.

Los denominados filtros de bajas frecuencias o paso bajo son aquellos, que en el dominio de las frecuencias, dejan pasar solamente las frecuencias bajas (los elementos más comunes y que ocupan mayor superficie en la imagen) y eliminan las altas frecuencias (los elementos esporádicos y elementos poco frecuentes).

Este efecto, en el espacio real de la imagen, suele traducirse por la eliminación de ruidos o pequeños objetos esporádicos y por una difusión del borde de los objetos

Capítulo 2: Estudio de filtros en las distintas etapas del PDI

(desenfoque). Aunque en algunos casos puede resultar en una notable mejora de la imagen.

Actualmente, con los nuevos sistemas de captura digital de imágenes es poco frecuente tener necesidad de eliminar ruidos de fondo (sólo se producen en algunos dispositivos electrónicos de captura muy lenta y generación de imagen por fluorescencia o fosforescencia por ejemplo microscopios electrónicos de transmisión o en sistemas de captura a grandes distancias con interferencias). Sin embargo este tipo de filtros presenta importantes aplicaciones en la mejora de imágenes ampliadas por métodos digitales o analógicos (fotografía convencional).

Véase por ejemplo el efecto de aplicar por dos veces consecutivas un filtro de paso bajo de 3x3 (definido en la página anterior) a una imagen digitalizada con un escáner de poca resolución y mucho aumento. Este tipo de filtrado es el equivalente de una interpolación bicúbica (Ver Figuras 12, 13, 14).



Fig 12 Imagen original.



Fig 13 Imagen digitalizada.



Fig 14 Resultado después de filtrar.

En la **Figura12** la imagen original. **Figura13** la imagen fue digitalizada demasiado grande con un escáner de poca resolución. Se ven las líneas del escaneado, ruido en el fondo negro y un efecto pixelado en el degradado de color de la parte inferior derecha. **Figura14** resultado después de filtrar. La imagen se nota algo desenfocada, pero en parte se reducen algunos de los defectos del proceso de digitalización.

Otro ejemplo de aplicación de un filtro de paso bajo podría ser el caso de una imagen que se ha ampliado mucho, con lo que se ha producido un efecto de píxelado en todos sus bordes. En este caso es interesante aplicar un filtro de paso bajo o promedio, con ponderación de distancia como el anteriormente. Ver Figura 15,16



Fig 15 Imagen excesivamente ampliada. Se observa el efecto de píxelado provocado por la ampliación sin interpolación ni corrección de ningún tipo.



Fig 16 Imagen anterior después de aplicarle un filtro de tipo paso bajo modificado o también llamado de interpolación ponderada.

Ejemplos del empleo de filtros de baja frecuencia para eliminar el efecto muaré o el efecto de pixelado. Cada vez es más frecuente almacenar de forma digital imágenes que previamente han sido impresas con técnicas de medias tintas ("half toning"). La reproducción electrónica de este tipo de imágenes, previo escaneado, suele producir un efecto de interferencias de colores que se produce en impresión cuando las tramas no están correctamente inclinadas. Este defecto puede ser corregido utilizando filtro de reducción de ruido como puede ser un filtro de paso bajo o un filtro de tipo gaussiano. Sin embargo estos filtros producen un efecto de desenfoque que puede corregirse mediante un filtro de realce de contornos o mediante distintas técnicas de máscaras borrosas ("unsharp masking").

Ejemplo de la eliminación de la trama en imágenes digitalizadas de fotografías impresas (Ver Figura 17 ,18).

Al digitalizar con cierta resolución, fotografías impresas en libros, revistas y periódicos, aparece la trama de impresión a modo de ruido de fondo de la imagen. La trama de impresión es ruido periódico y direccional, que puede eliminarse mediante un filtro de frecuencias en el espacio transformado de Fourier. Sin embargo, técnicas de máscaras borrosas ("unsharp masking"), utilizando filtros de tipo gaussiano y laplaciano logran eliminar el efecto de la trama por difusión de los bordes y reconstruir el enfoque.

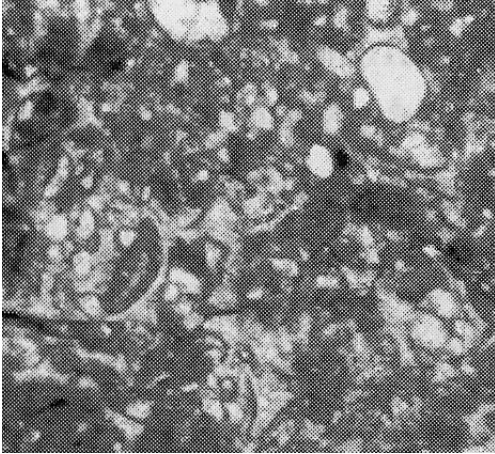


Fig 17 Imagen antes.

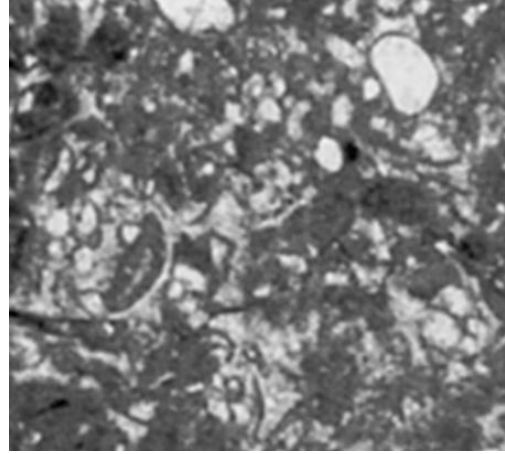


Fig 18 Imagen despues.

Este filtro tiene el inconveniente de destrozarse las altas frecuencias (bordes) y que picos muy grandes de ruido no son eliminados. Solo es utilizado para eliminar ruidos o pequeños objetos esporádicos y por una difusión del borde de los objetos, produciendo así un efecto de desenfoque, necesitando de otros filtros que solucionen esta dificultad.

Filtro de paso bajo

Su objetivo es suavizar la imagen, son útiles cuando se supone que la imagen tiene gran cantidad de ruido y se quiere eliminar. También pueden utilizarse para resaltar la información correspondiente a una determinada escala (tamaño de la matriz de filtrado), entre los filtros que componen la gran familia de filtros de paso bajo vamos a hacer referencia al filtro gaussiano.

- **Filtros gaussianos.** Consiste en la mezcla de los colores de una imagen para conseguir un efecto de desenfoque. Presenta desventajas ya que puede producir la pérdida de detalles, el aumento de borrosidad, y la disminución de la nitidez. Esta es una explicación muy general, y por supuesto el algoritmo es más complejo que esto. Simulan una distribución gaussiana bivalente. El valor máximo aparece en el píxel central y disminuye hacia los extremos de manera directamente proporcional a la disminución del parámetro de desviación típica S . El resultado será un conjunto de valores entre 0 y 1. Para transformar la matriz a una matriz de números enteros se divide toda la matriz por el menor de los valores obtenidos. La ecuación para calcularla es:

$$g(x, y) = e^{-\frac{x^2+y^2}{2*s^2}}$$

$$G(x, y) = \frac{g(x, y)}{\min_{x,y}(g(x, y))}$$

1	4	7	4	1
4	20	33	20	4
7	33	55	33	7
4	20	33	20	4
1	4	7	4	1

Tabla 1 Filtro Gaussiano con s=1, r=2.

Donde:

S: Desviación típica

r: Atenuación

0	0	0
0	1	0
0	0	0

DIV 1

Tabla 2 Matriz de filtrado identidad.

Son una manera de obtener filtros de tipo genérico. Pueden ser útiles, por ejemplo, cuando se asume que la respuesta espectral de un píxel es función de la reflectividad de los píxeles vecinos atenuada en función de la distancia. El alcance de esta atenuación (r) viene marcado

Capítulo 2: Estudio de filtros en las distintas etapas del PDI

por el tamaño de la ventana de filtrado ($w = 2r + 1$) que debe especificarse previamente. En la tabla# 1 aparece un ejemplo.

En primer lugar, consideraremos la imagen como una matriz de puntos (píxeles), con unos valores determinados de rojo, verde y azul. A partir de aquí, los valores de cada punto de la nueva imagen, serán una combinación de los puntos cercanos. De esta forma, si un punto es de color rojo, y cercano a él hay puntos negros, este rojo se oscurecerá (y los negros se enrojecerán) (Ver Figura 19).



Fig 19 Filtro Gaussiano.

Pero para determinar cuáles son los píxeles que son cercanos, y los que no, se debe especificar al filtro un radio de acción. Con esto, si por ejemplo indicamos un radio de 10 píxeles, cada píxel de la nueva imagen se verá modificado por todos los píxeles a menos de 10 unidades de distancia. Pero no todos los píxeles modifican con la misma “fuerza”, cuanto más cerca estén del píxel que se está modificando, más valor tendrán en el cálculo del nuevo color.

La función de la distribución Gaussiana va a ayudarnos a adjudicar cuánto aporta cada píxel, en función de lo cerca o lejos que se encuentre. Esta función tiene la forma donde $(x-b)$ es la distancia entre el píxel que está siendo modificado y el píxel del que se está tomando el valor, a y c valen 1, y $f(x)$ resulta un número Real entre 0 y 1 que representa el porcentaje de influencia.

En resumen, lo que está claro es que hay que mezclar los píxeles entre sí, y que usaremos un radio de influencia en que cuanto más lejos está un píxel de otro, menos le afecta. Después de analizar el problema, y aplicar una fórmula estadística, podemos concluir que el valor de cada píxel es:

La suma de todos los píxeles dentro del radio R , multiplicados cada uno por

$$f(x) = ae^{-(x-b)^2/c^2} \quad (\text{tal y como se ha descrito antes}), \text{ y todo esto, dividido entre la suma}$$

de todos los valores de la fórmula $f(x) = ae^{-(x-b)^2/c^2}$ en los puntos que intervienen.

Capítulo 2: Estudio de filtros en las distintas etapas del PDI

Todo eso es sólo para calcular uno de los tres componentes (Rojo, Verde o Azul) de un sólo píxel (Ver Figura 20).

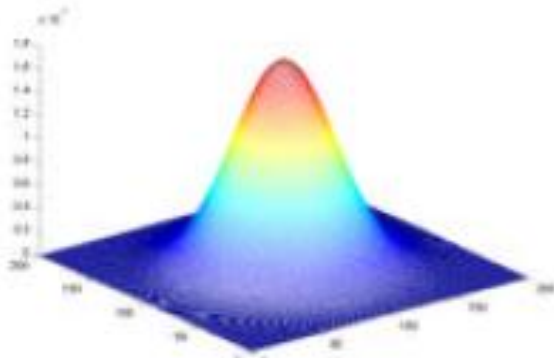


Fig 20 Histograma.

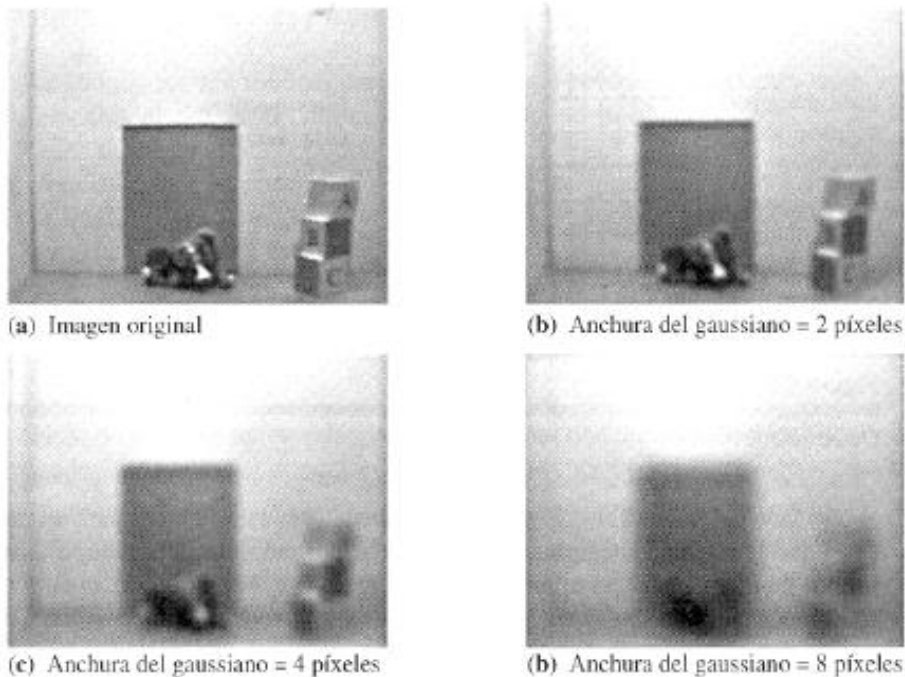


Fig 21 Imagen suavizada por filtro Gaussiano.

Son filtros paso-bajos que se utilizan fundamentalmente para reducir el ruido de alta frecuencia que se produce en una imagen.

- **Filtro de la media.** Consiste en sustituir cada píxel por la media aritmética de los puntos que tiene alrededor (incluido el mismo). Se toma una ventana de $N \times N$ puntos, donde el punto a sustituir es el central, se toma la media de la suma de los valores de los píxeles

Capítulo 2: Estudio de filtros en las distintas etapas del PDI

presentes en la ventana y se sustituye el píxel en la nueva imagen por el valor obtenido. El efecto del filtro aumenta a medida que lo hace el tamaño de la ventana y los resultados que produce son equivalentes a una difuminación^{xi} de la imagen, produciendo un suavizado de los gradientes en toda la imagen. Es de gran utilidad para eliminar ruido producido por un muestreo deficiente o una transición ruidosa (Ver Tabla 3).

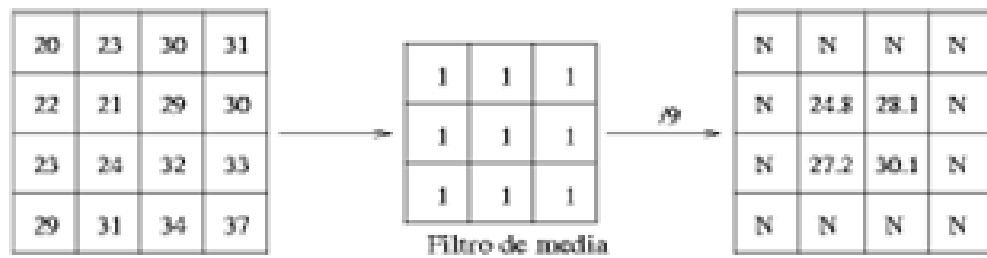


Tabla 3 Filtro de media.

- **Filtro de media ponderada.** Los elementos de la matriz de filtrado no son todos 1 sino que se da más peso a uno de ellos (generalmente el central) para obtener un resultado más parecido a la imagen original y evitar que aparezca borrosa (Ver Tabla 4).

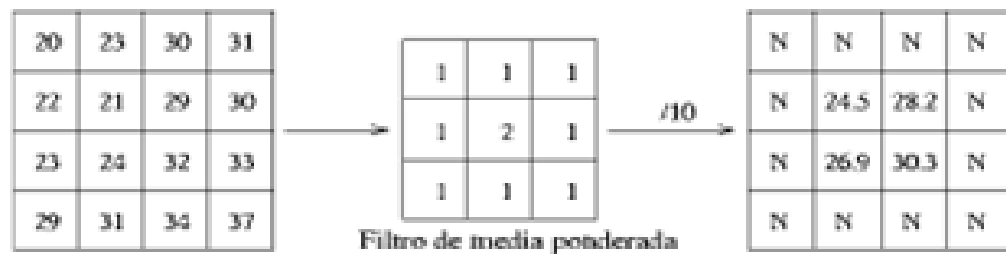


Tabla 4 Filtro de media ponderada.

- **Filtro de la mediana.** Es un filtro parecido al de mediana. Consiste en obtener la lista de todos los valores de los píxeles de la ventana, ordenarlos y coger el del medio (mediana), es decir, el valor que tenga el mismo número de valores superiores e inferiores a él dentro de la ventana. Este filtro al igual que el de mediana es muy utilizado para eliminar ruido impulsivo de una imagen. Este filtro presenta la ventaja de no eliminar los bordes como lo haría un filtro de media, aunque resulta computacionalmente costoso debido a que requiere la ordenación de los valores de los puntos de la vecindad para calcular la mediana y así sustituir el valor de cada píxel (Ver Tabla 5, Fig. 22).

20	23	30	31
22	21	29	30
23	24	32	33
29	31	34	37

Filtro de mediana

N	N	N	N
N	23	30	N
N	29	31	N
N	N	N	N

Tabla 5 Filtro de mediana.

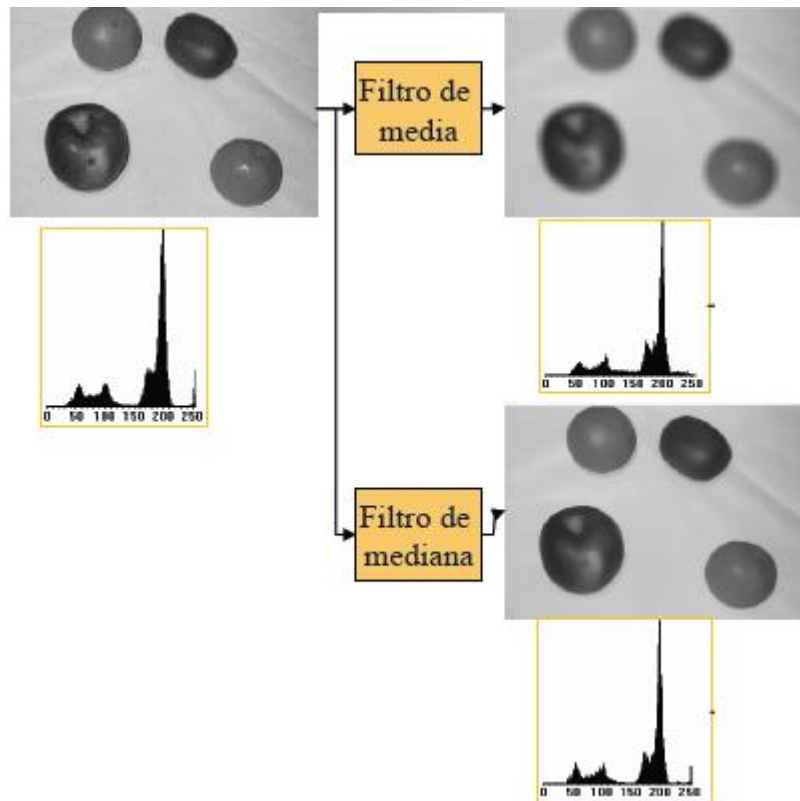


Fig 22 Tipos de filtros y Mecanismos de aplicación.

- **Filtros adaptativos.** Son considerablemente más complejos ya que los coeficientes de ponderación se re-calculan para cada uno de los píxeles en función del histograma. Se han utilizado con gran éxito filtros adaptativos para eliminar el ruido de las imágenes de radar y para detectar con un solo filtro, diferentes elementos.
- **Filtros de moda.** La moda de un conjunto de valores se define como el valor que más se repite dentro de ellos. Por lo tanto el filtro de moda consiste en calcular el valor más repetido dentro de todos los píxeles de una ventana.

Capítulo 2: Estudio de filtros en las distintas etapas del PDI

Los efectos que se producen dependen en gran manera del tipo de imagen que se filtre. Si está formada por texturas que tienen gran cantidad de puntos dispares (impares), el resultado será bastante significativo y variable, ya que en cada punto los valores serán muy dispares (impares). Si en cambio la imagen tiene zonas homogéneas este filtro las homogeneizara cada vez más.

Filtro de paso alto

Su objetivo es resaltar las zonas de mayor variabilidad eliminando lo que sería la componente media, precisamente la que detectan los filtros de paso bajo. Por otra parte la respuesta de cada píxel está contaminada por la de los píxeles vecinos ya que la radiación reflejada por un píxel se reparte hacia los píxeles vecinos. Los filtros de paso alto consiguen también eliminar en parte esta contaminación. Existen diversos métodos:

- **Sustracción de la media.** Si se considera que un filtro de paso bajo sirve para resaltar componentes a gran escala eliminando la variabilidad local, si a la imagen original se le resta el resultado de pasarle un filtro de paso bajo se consigue resaltar esa variabilidad local, la matriz de filtrado de este filtro menos media puede calcularse directamente restando a la matriz de filtrado identidad.
- **Filtros basados en las derivadas.** La derivada de una función $y = f(x)$ es el incremento de (y) para cada incremento infinitesimal de x . En el caso de Modelo Digital de Elevaciones (MDE) la derivada es la pendiente. La segunda derivada es la derivada de la derivada, en el caso de un MDE nos da información acerca de la forma (ladera recta, cóncava o convexa, valle, informar de cómo son los cambios, más o menos bruscos, que se producen entre píxeles contiguos. A continuación aparecen las ecuaciones de las derivadas respecto a x e y , las segundas derivadas respecto a x e y , la derivada compuesta y la derivada segunda compuesta denominada también laplaciana.

-1	-1	-1
-1	8	-1
-1	-1	-1

 DIV 9

Tabla 6 Filtro menos media.

0	1	0
1	-4	1
0	1	0

DIV=1

Tabla 7 Filtro laplaciano.

$$\Delta \quad xf(i,j) = f(i,j) - f(i-1,j)$$

$$\Delta \quad yf(i,j) = f(i,j) - f(i,j-1)$$

$$\Delta \quad x2f(i,j) = \Delta \quad xf(i+1,j) - \Delta \quad xf(i,j) = f(i+1,j) + f(i-1,j) - 2f(i,j)$$

$$\Delta \quad y2f(i,j) = \Delta \quad yf(i,j+1) - \Delta \quad yf(i,j) = f(i,j+1) + f(i,j-1) - 2f(i,j)$$

$$\Delta \quad xyf(i,j) = \text{sqr} \quad \Delta \quad xf(i,j)^2 \quad \Delta \quad yf(i,j)^2$$

$$\Delta \quad xy2f(i,j) = [f(i+1,j) + f(i-1,j) + f(i,j+1) + f(i,j-1)] - 4f(i,j)$$

- **Filtro laplaciano.** Se recomienda para el realce de rasgos lineales en entornos urbanos. Otra opción para resaltar los elementos de mayor variabilidad es restar a la imagen original la obtenida mediante un filtrado laplaciano. El filtro laplaciano se basa en crear una máscara que destaque los píxeles (a través del aumento de su nivel de gris) cuya variación, con respecto a su vecindad, es significativa.

0	-1	0
-1	5	-1
0	-1	0

DIV=1

Tabla 8 Filtro menos-laplaciano.

$$\begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & -2 & 1 \\ \hline -1 & -1 & -1 \\ \hline \end{array} \text{DIV}=1 \quad \begin{array}{|c|c|c|} \hline -1 & 1 & 1 \\ \hline -1 & -2 & 1 \\ \hline -1 & 1 & 1 \\ \hline \end{array} \text{DIV}=1$$

Tabla 9 Filtros Norte y Este.

Filtros Espaciales

Este tipo de filtros suele consistir en recorrer toda la imagen, píxel a píxel, y realizar alguna operación aritmética con un número concreto de píxeles vecinos. Al conjunto de estos vecinos se les denomina ventana, la cual podrá ser de 3x3, 5x5.

En la (Fig. 23) puede ver la representación de una ventana de 3x3 y de 5x5, en la que los píxeles vecinos van del 1 al 9 y del 1 al 25 respectivamente siendo el píxel central sobre el que se aplica el algoritmo. La mayoría de los algoritmos suelen ser combinaciones lineales de los píxeles de la ventana. Para representar un filtro se suele dar el vector o matriz de convolución o núcleo que representa la combinación lineal entre ellos.

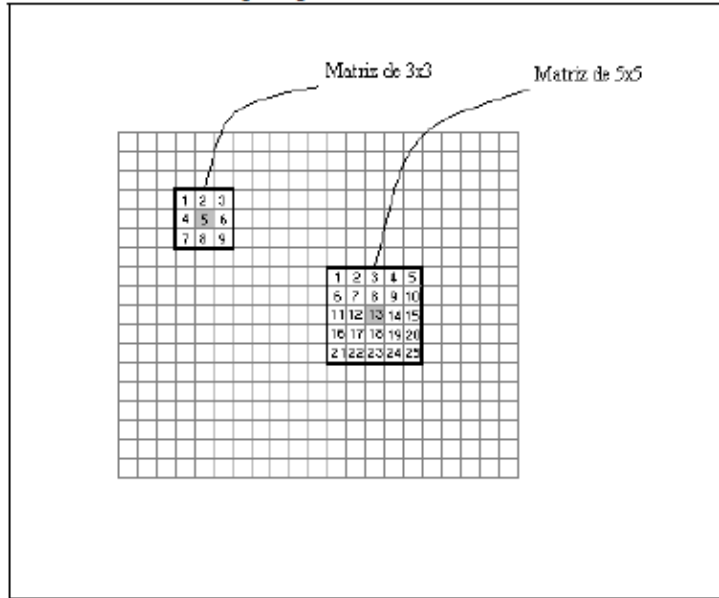


Fig 23 Representación de dos tipos de ventana para el filtrado de imágenes.

Estos filtros hacen uso de las matrices de convolución, las cuales como antes se mencionó tiene un alto coste computacional.

2.2.3 Segmentación

En las técnicas de procesamiento de las imágenes cada píxel era transformado de un valor a otro. Estos cambios, mayoritariamente, se realizan para facilitar la partición de la imagen en áreas de píxeles con significado. En esta nueva fase, se trata de agrupar los píxeles, por algún criterio de homogeneidad, para particionar la escena en regiones de interés. Estas áreas deben de tener algún significado físico. Por tanto, la segmentación de una imagen es un proceso de extracción de objetos de interés insertados en la escena capturada. La agrupación de los píxeles se hace a razón de que sus vecinos sean similares en criterios como de luminancia, color, bordes, texturas, movimientos. Una vez que la imagen ha sido particionada, la unidad dejará de ser el píxel para ser la agrupación de píxeles que constituye el objeto. La imagen estará definida por un conjunto de objetos, habiendo pasado de un nivel bajo a otro más elaborado o nivel medio visual. La información estará preparada para el reconocimiento e interpretación de la imagen.

Para la segmentación de las imágenes se usan tres conceptos básicos:

- Similitud: los píxeles agrupados del objeto deben ser similares respecto a algún criterio (nivel de gris, color, borde, textura,...).
- Conectividad: los objetos corresponden a áreas de píxeles con conectividad. Las particiones corresponden con regiones continuas de píxeles.
- Discontinuidad: los objetos tienen formas geométricas que definen unos contornos. Estos bordes delimitan unos objetos de otros.

En la práctica, la imposición de estas condiciones sobre la estrategia de segmentación resulta casi imposible. Respecto a la similitud, la aparición del ruido presente en todas las imágenes, la falta de iluminación uniforme sobre el escenario, la creación de sombras de unos objetos sobre otros, produce que algo que parecía sencillo de definir como es alguna regla sobre similitud, resulte impracticable de acotar. Todas ellas fracasan sobre escenas más o menos complejas.

La ocultación parcial de un objeto físico capturado hace fallar la conectividad del objeto virtual. En el proceso de segmentación se puede producir varias regiones de píxeles agrupados que provienen del mismo objeto físico. No sólo debido a las propias oclusiones de éste, sino también a los propios cambios debido a la textura del objeto, la variación de la iluminación de la escena o al reflejo de otros objetos físicos adyacentes. Por último y tal cual se comenta a continuación, la detección de los contornos físicos suele estar plagadas de errores y de discontinuidades en los bordes. Resulta extraordinariamente difícil obtener los

Capítulo 2: Estudio de filtros en las distintas etapas del PDI

contornos cerrados, inmunes al ruido y sin desplazamiento entre el contorno real y el obtenido. La conclusión es obvia, la segmentación es un proceso altamente complejo. Es una labor de investigación actual y en las siguientes décadas venideras. En el paradigma actual se combinan las técnicas basadas en la localización de regiones uniformes con los contornos de los objetos.



Fig 24 Basadas en la localización de regiones uniformes con los contornos de los objetos.

No obstante, para la actual Visión Artificial todavía existe mucho recorrido para las aplicaciones industriales. Las citadas dificultades de la fase de segmentación son resueltas con una elección esmerada en la formación de la imagen, eligiendo una iluminación adecuada y simplificando la escena. Como se ha comentado varias veces, esta técnica resulta factible para el análisis de imágenes simples y repetitivas (Ver Fig. 24).

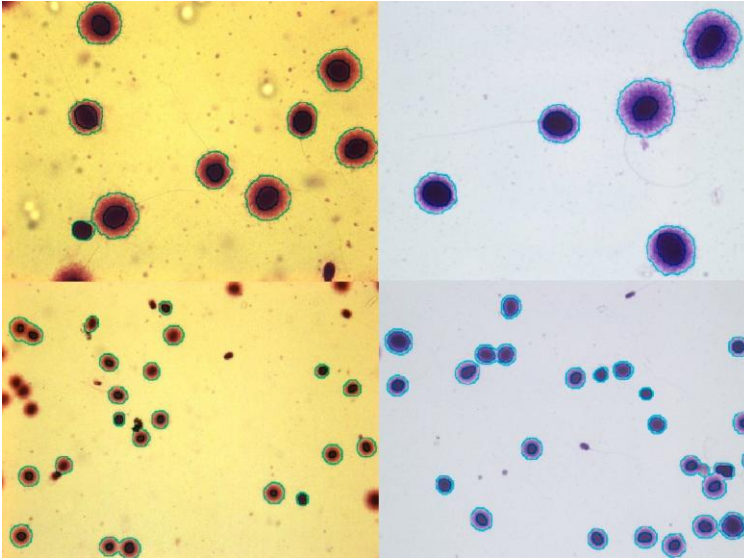


Fig 25 Imagen segmentada.

Existen varios tipos de segmentación, a continuación se mencionan algunos de ellos, explicando algunas de sus características brevemente.

Tipos de segmentación analizados en el trabajo

- Segmentación basada en características
 1. Segmentación por niveles de gris.
 2. Segmentación de imágenes en color.
 3. Segmentación por texturas.

- Segmentación basada en transiciones
 1. Detección de bordes.
 2. Segmentación basada en modelos.
 3. Transformada de Hough.

- Segmentación basada en homogeneidad

2.2.3.1 Segmentación basada en características

Se asigna cada píxel a una región en función de características locales de la imagen en el píxel y (posiblemente) en su vecindad (Ver Figura 26).

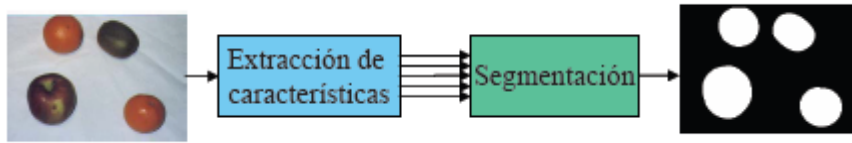


Fig 26 Segmentación basada en características.

Segmentación por niveles de gris

- ❖ Utiliza como característica solamente la luminancia de cada píxel.
 - Útil si distintos objetos se corresponden con niveles de gris diferentes.
- ❖ El nivel de gris determina alguna propiedad física en una imagen específica de una aplicación (Rayos X, RMN, bandas en satélite, etc.)
 - Técnicas de operaciones puntuales (ej.: doble umbralización)
 - Técnicas de clasificación basada en amplitud de una imagen binaria (Ver Figuras 27 y 28).



Fig 27 Imagen original.

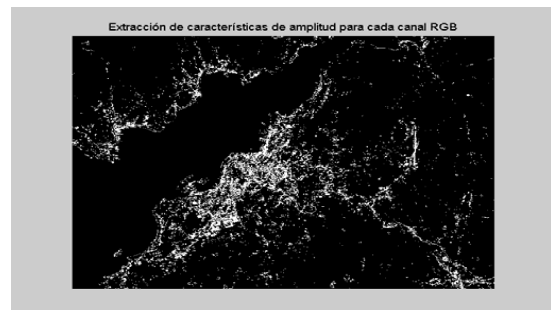


Fig 28 Imagen binarizada.

- En el caso ideal en que el objeto posea un rango estrecho de niveles de gris frente a un fondo uniforme, podemos establecer un nivel de gris intermedio (umbral) para separar objeto y fondo.
- Para separar por umbral, es útil recurrir al histograma (Ver Figura 29).

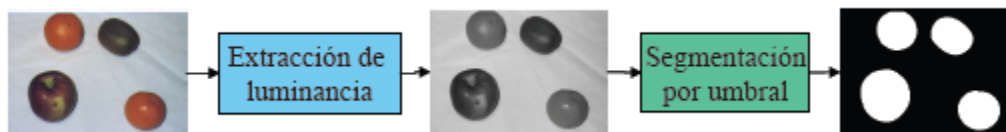


Fig 29 Segmentación por umbral.

Segmentación de imágenes en color

Es posible segmentar una imagen en color utilizando umbrales. Ver Figura 30

Para ello, se presentan dos alternativas:

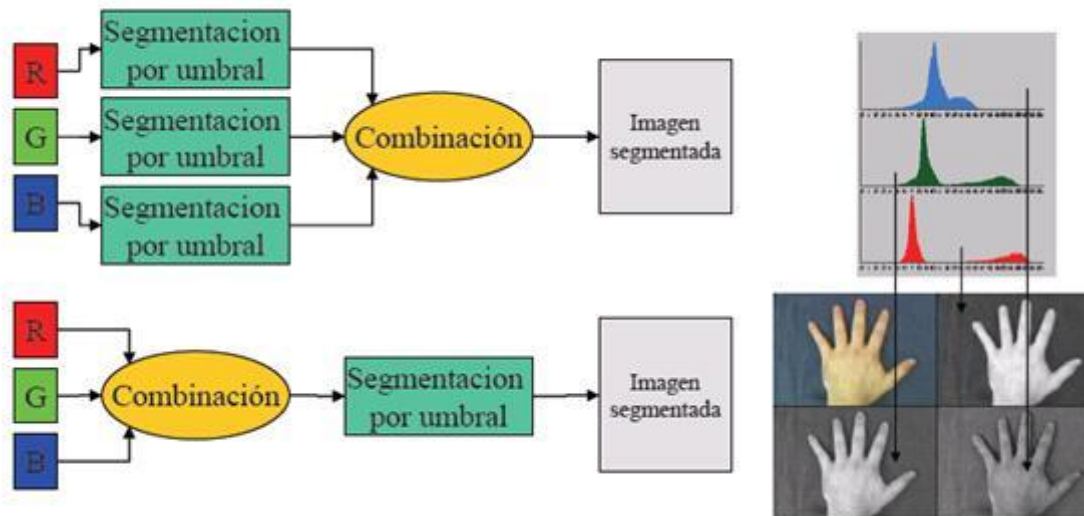


Fig 30

La exposición de este capítulo versa sobre las distintas técnicas de Segmentación y filtrado, clasificadas en:

- Contornos de los objetos: trata de localizar las fronteras de los objetos. Hace referencia a los aspectos geométricos de los objetos.
- Búsqueda de regiones homogéneas: agrupa los píxeles por que cumple algún criterio de similitud y tienen conectividad entre ellos. Está unido a las características radiométricas del objeto.

Filtros de texturas

¿Que es una textura?

Es una región macroscópica estructurada de la imagen que presenta propiedades locales constantes, lentamente variables o aproximadamente periódicas. Este orden local radica en la presencia repetida y no casual de partes elementales que tienen dimensión similar. Depende de la escala a la que se observa la imagen: la textura de un píxel no está definida, hay que estudiar la vecindad de un grupo de píxeles. Dependiendo del tamaño de vecindad elegido la textura puede ser diferente.

- **Filtro Trilinear:** Es una técnica para corregir la distorsión que sufren las texturas aplicadas a los objetos en la distancia.
- **Anti-aliasing:** Por anti-aliasing se conoce al efecto de filtrar la imagen para suavizar los bordes, y disimular los bordes de los polígonos, consiguiendo una apariencia mucho más realista.
- **Supersampling:** Es una de las formas de anti-aliasing. Consiste en escalar la imagen multiplicando la resolución de los ejes X e Y por un número entero, y renderizarla ^{xi}a mayor tamaño del original. Tras realizar este proceso, se pasa al tamaño original mediante un filtro de reducción, por lo que varios píxeles de la imagen renderizada a mayor resolución se utilizan para calcular el valor de los píxeles de la resolución final. De este modo se consiguen transiciones más suaves. Una de las ventajas del supersampling es que realiza un filtrado consiguiendo que las texturas tengan mayor definición. Permite aumentar la calidad de las texturas con un descenso de rendimiento muy inferior.

La ventaja del filtro anti-aliasing sobre el supersampling es que conseguimos una calidad de imagen muy similar a la del supersampling, pero con un rendimiento muy superior. Esto hace que en las tarjetas modernas se esté abandonando el supersampling debido a que la mejora en calidad visual no compensa la pérdida de rendimiento que conlleva esta técnica.

2.2.3.2 Segmentación basada en transiciones

Técnicas basadas en los bordes

Para la localización de las fronteras de los objetos en la escena se emplearán las técnicas de detección de los bordes. Sin embargo, esta etapa no es definitiva para poder segmentar los objetos presentes en la imagen. La presencia de ruido, el efecto de las sombras, la falta de iluminación uniforme y un largo etcétera de causas, produce que los contornos no sean del todo continuos y cerrados sobre los objetos. Se requiere otra etapa de post-procesamiento. Esta nueva fase emplea los resultados de la detección de bordes para elaborar las fronteras de los objetos. Se trata de agrupar los píxeles etiquetados como bordes, de la etapa anterior (de detección de bordes), empleando la propiedad de conectividad. Para que un píxel etiquetado como borde se defina como píxel frontera de un objeto se necesitará que otros píxeles bordes tengan similar dirección y módulo del gradiente. Dos píxeles serán considerados pertenecientes a

una misma frontera si presentan alguna condición de conectividad y las diferencias entre sus gradientes no superan un determinado umbral.

$$\begin{aligned} \|G_1 - G_2\| &< T_M \\ |\vartheta_1 - \vartheta_2| &< T_A \end{aligned}$$

Donde $|G_i|$ y ϑ_i se ha denotado el módulo y argumento del gradiente del píxel i , respectivamente. Evidentemente, en la construcción de la frontera, aparecerá ruido, píxeles que constituyen esquinas de los objetos y otras causas que producen discontinuidades en la localización de la frontera del objeto. Hay varios planteamientos para su decantación. Algunos autores emplean técnicas de crecimiento de regiones, las cuales se verán más adelante, pero que básicamente se trata de partir de un píxel borde semilla e ir agrupando píxeles con conectividad siempre y cuando cumplan algún criterio de homogeneidad, por ejemplo, el expresado en la anterior ecuación. Otros autores emplean un marco de trabajo basado en la optimización. Realizan un grafo de los posibles caminos que pueda llevar la frontera del objeto y mediante una función de costo, eligen la mejor solución. Un planteamiento alternativo es la agrupación de los píxeles que cumplen un cierto criterio en una primitiva más elaborada, como puede ser un pequeño segmento con orientación. Posteriormente, se procede a encadenar estas primitivas dando paso a la delimitación de los objetos.

Un campo muy activo en esta materia son los contornos activos. Se trata de emplear una curva cerrada que vaya adaptándose dinámicamente hasta alcanzar la frontera del objeto. Ver Figura 31.

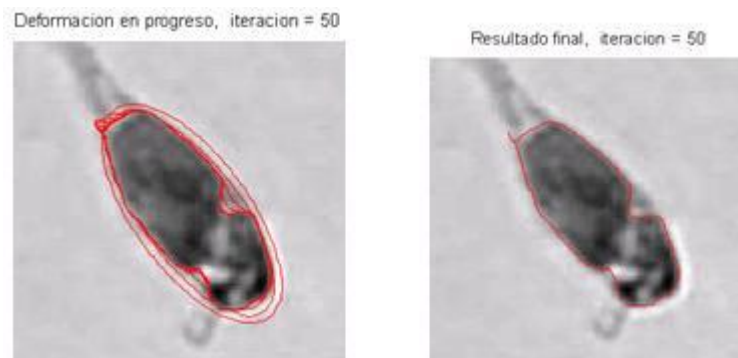


Fig 31

Muchas aplicaciones se basan en las localizaciones de líneas rectas o de otras formas básicas como círculos o elipses, ya que se consideran elementos fundamentales en la percepción de entornos estructurados o semi-estructurados. Se basa en el hecho de que las escenas artificiales, construidas por el hombre, están plagadas de contornos rectos o de objetos con formas circulares o elípticas. En este sentido, se aplican las transformadas de Hough para la localización de estas fronteras pero las transformadas de Hough, tienen un alto coste computacional, devuelve el modelo de una recta infinita y no de un segmento, lo que significa que no se sabe donde empieza y acaba la recta.

Filtros realzadores de contorno

Son utilizados para hacer que los bordes sean más marcados. **¿Qué es un borde?**

Cambio brusco en la intensidad. Realmente actúan los puntos que más se diferencien de los vecinos aumentando por lo tanto las altas frecuencias. En este caso el ruido de alta frecuencia aumenta.

Las matrices más utilizadas son:

$$h_i = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad h_j = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 5 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad h_k = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 5 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

Detección de Bordes

- Al detectar los bordes se puede transformar la imagen en un dibujo a base de líneas.
- Los bordes están relacionados con propiedades relevantes de los objetos de la escena.
- Comparar con los patrones de contorno característicos de los objetos que pueden aparecer en la imagen.

Filtro o Algoritmo de Canny

El algoritmo de Canny es usado para detectar todos los bordes existentes en una imagen. Este algoritmo está considerado como uno de los mejores métodos de detección de contornos mediante el empleo de máscaras de convolución y basado en la primera derivada^{xiii}. Los puntos de contorno son como zonas de píxeles en las que existe un cambio brusco de nivel de gris. En el tratamiento de imágenes, se trabaja con

píxeles, y en un ambiente discreto, es así que en el algoritmo de Canny se utilizan máscaras, las cuales representan aproximaciones en diferencias finitas.

El algoritmo de Canny basa su funcionamiento en tres criterios

- Un criterio de detección expresa el hecho de evitar la eliminación de bordes importantes y no suministrar falsos bordes.
- El criterio de localización establece que la distancia entre la posición real y la localizada del borde se debe minimizar.
- El criterio de una respuesta que integre las respuestas múltiples correspondientes a un único borde.

Uno de los métodos relacionados con la detección de bordes es el uso de la primera derivada, la que es usada por que toma el valor de cero en todas las regiones donde no varía la intensidad y tiene un valor constante en toda la transición de intensidad. Por tanto un cambio de intensidad se manifiesta como un cambio brusco en la primera derivada, característica que es usada para detectar un borde, y en la que se basa el algoritmo de Canny.

El algoritmo de Canny consiste en tres grandes pasos:

- Obtención del gradiente: en este paso se calcula la magnitud y orientación del vector gradiente en cada píxel.
- Supresión no máxima: en este paso se logra el adelgazamiento del ancho de los bordes, obtenidos con el gradiente, hasta lograr bordes de un píxel de ancho.
- Histéresis de umbral: en este paso se aplica una función de histéresis basada en dos umbrales; con este proceso se pretende reducir la posibilidad de aparición de contornos falsos.

El algoritmo de Canny, tiene como principal ventaja su gran adaptabilidad para poder ser aplicado a diversos tipos de imágenes, además de no disminuir su performance ante la presencia de ruido en la imagen original. Sin embargo, algunas de las desventajas que pude identificar al implementar este algoritmo se encuentran en el suavizado, puesto que, si aumentamos la desviación estándar de la mascarilla logramos reducir el ruido pero difuminamos los bordes y perdemos calidad al momento de calcular la localización.

Transformada de Hough

Este algoritmo trata de detectar formas geométricas sencillas en una imagen. En su procesamiento emplea el conjunto total de la imagen, haciéndolo robusto ante la presencia del ruido o discontinuidades de las etapas previas. De hecho, para su ejecución requiere de una imagen binarizada en la que se han seleccionado previamente los bordes. Por tanto, la entrada al algoritmo de Hough es una imagen en la que se ha hecho la detección de los bordes. Hough, desde una perspectiva global, intentará extraer primitivas de más alto nivel como son líneas, circunferencias, elipses o cualquier tipo de curva parametrizada o no. El mayor inconveniente es su alto coste computacional. Otro inconveniente de la transformada de Hough es que devuelve el modelo de una recta infinita y no de un segmento. No se sabe donde empieza y acaba la recta.

Filtro basado en Gradientes

Estos métodos proveen una solución al problema de estimación de movimiento, a partir de la observación de los cambios en el brillo de una secuencia de imágenes. Aprovechando este fundamento, se pueden detectar los bordes de un elemento en movimiento en cualquier escena dinámica. En la estimación del movimiento por medio de métodos basados en gradientes uno de los principales problemas es la presencia de elementos no móviles que pueden estar en la escena y que interfieren con la correcta detección de los bordes de los elementos en movimiento.

Filtros basados en histograma.

Utilidad del histograma

Es la base para el desarrollo de numerosas técnicas de procesamiento de imágenes. El histograma es el gráfico de las frecuencias de aparición de cada nivel de gris, o de los niveles de cada canal de un modelo de color de la imagen, y facilita información para realizar múltiples operaciones de reasignación de los valores de cada píxel de la imagen.

Realce de imágenes

- Provee estadísticas de la imagen.

- Segmentación: determinación de umbrales para binarización.
 - Herramienta fácil de implementar.
 - Modificación global de los niveles de gris de la imagen.
- Definir, para cada nivel de gris i de la imagen original, un nuevo valor j .

Transformaciones de histograma

Las principales transformaciones del histograma son:

- **Negativo o vídeo inverso:** cambia valores oscuros por claros y viceversa, de tal modo que $g(x, y) = 255 - f(x, y)$.

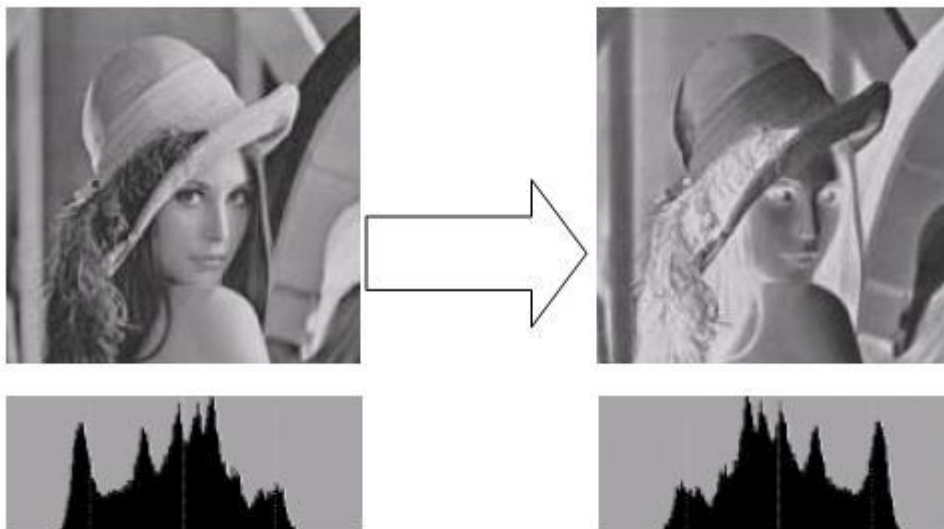


Fig 32 Imagen gamma: resalta los oscuros y satura los claros, mejorando la percepción en imágenes oscuras, o a la inversa, $(g(x,y) = f(x,y)^\gamma)$ según el valor del exponente.

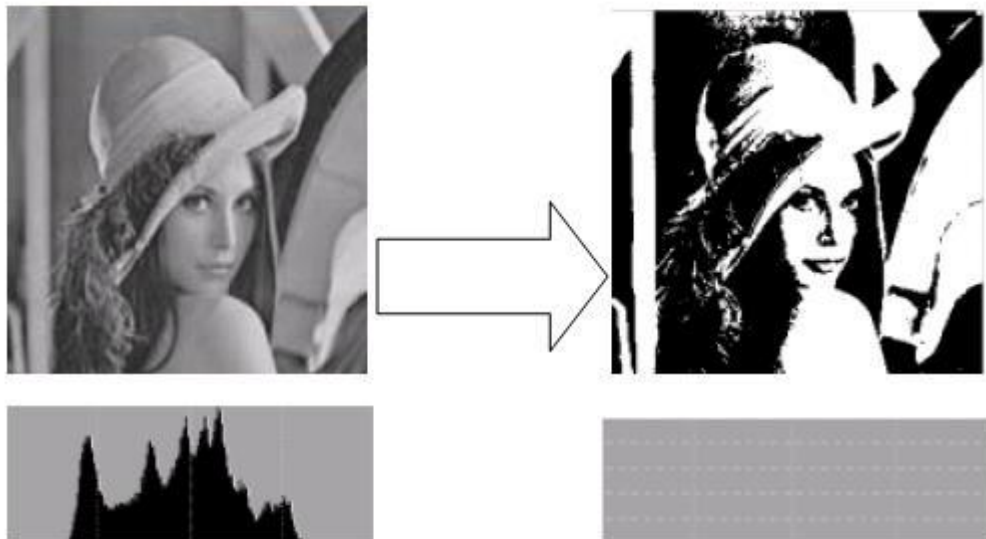


Fig 33

- **Pseudocolor:** se asigna un mapa de colores a una imagen en escala de grises, o se reasigna el mapa de colores de una imagen a color (color falso), con lo que se logra apreciar detalles de la imagen que debido a la limitación del ojo humano (solo detecta unos 20 niveles de gris) pasarían desapercibidos.
- **Estiramiento del histograma:** se emplea para mejorar el contraste de las imágenes, dado que distribuye un intervalo del histograma de la imagen de modo que ocupe todo el rango de valores, de tal modo que $g(x, y) = 255 * (f(x, y) - a) / (b-a)$, donde $[a, b]$ es el intervalo de la imagen que será extendido.

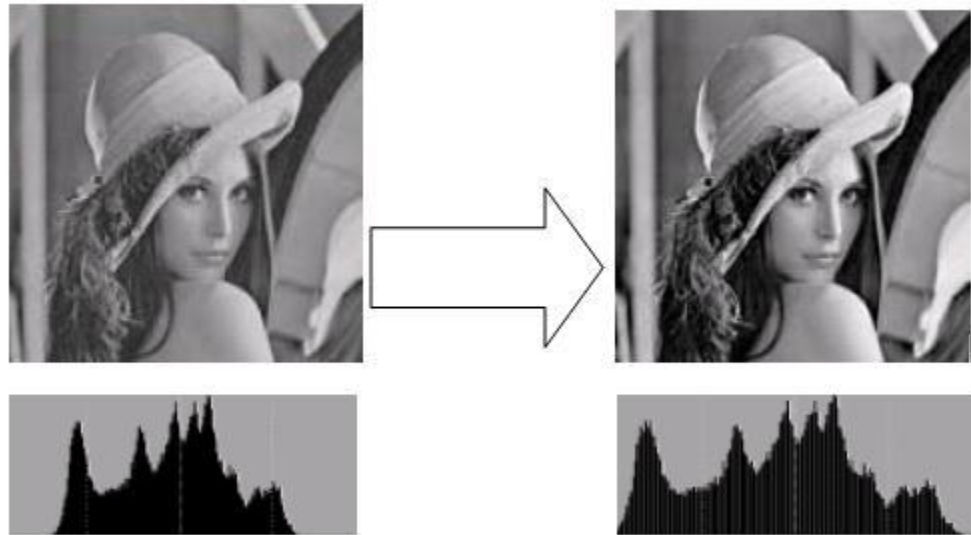


Fig 34 Estiramiento del histograma.

- **Ecuación del histograma:** intenta igualar la frecuencia de aparición de todos los niveles de gris, mejorando el contraste de la imagen, aunque a veces puede provocar la aparición de ruido donde hay una rampa en el histograma acumulado.

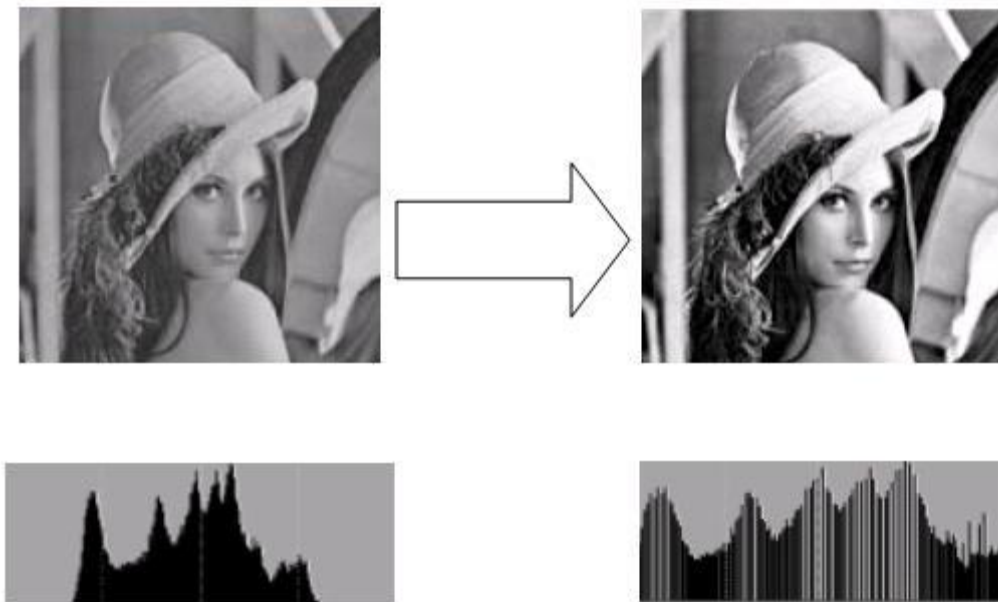


Fig 35 Ecuación del histograma.

Segmentación por histograma

La segmentación por histograma es una técnica global que se basa en considerar una definición clara de los objetos respecto del fondo, con una iluminación relativamente uniforme en toda la escena. Se supone que el histograma está constituido por dos picos y un valle. Al segmentar empleando el histograma, se denota que este proceder no considera las relaciones de vecindad de los píxeles. El umbral estará definido justamente en el valle.

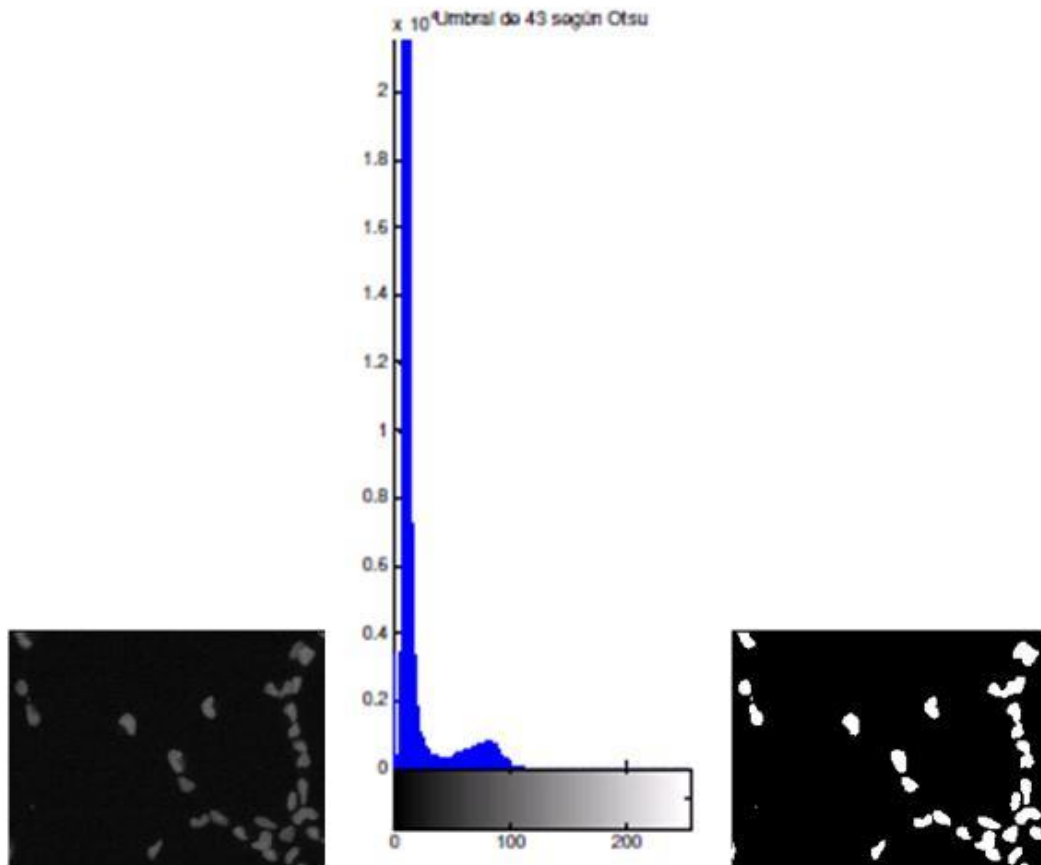


Fig 36 Técnica de segmentación mediante análisis del histograma.

Esta técnica es extensible a un histograma con N picos y $N-1$ valles, presentando varios umbrales. Sin embargo, a mayor número de umbrales en la imagen, menos fiables serán los resultados. La umbralización por histograma presenta ciertas limitaciones:

- Es difícil identificar correctamente los mínimos del histograma.
- Se tienen problemas cuando las regiones varían suavemente su nivel (por ejemplo, el efecto de las sombras).
- Se aplica sólo cuando hay pocas regiones.
- No se pueden distinguir regiones separadas de niveles similares de gris.
- No considera la conectividad de los píxeles. Debido a este inconveniente, se suele emplear a continuación técnicas de post-procesado (se tratarán más adelante).

Segmentación por Umbralización

La umbralización es una técnica de segmentación ampliamente utilizada en las aplicaciones industriales. Se emplea cuando hay una clara diferencia entre los objetos a extraer respecto al fondo de la escena. Los principios que rigen son la similitud entre los píxeles pertenecientes a un objeto y sus diferencias respecto al resto. Por tanto, la escena debe caracterizarse por un fondo uniforme y por objetos parecidos. Al aplicar un umbral, la imagen de niveles de grises quedará binarizada; etiquetando con '1' los píxeles correspondientes al objeto y con '0' aquellos que son del fondo. Por ejemplo, si los objetos son claros respecto del fondo, se aplicará:

$$g(x,y) = \begin{cases} 1 & f(x,y) > T \\ 0 & f(x,y) \leq T \end{cases}$$

$f(x, y)$ es la función que retorna el nivel de gris del píxel (x, y) , $g(x,y)$ será la función de la imagen binarizada y T es el umbral. En el caso de que los objetos sean oscuros respecto del fondo, la asignación sería a la inversa:

$$g(x,y) = \begin{cases} 1 & f(x,y) < T \\ 0 & f(x,y) \geq T \end{cases}$$

Si el umbral sólo depende de $f(x, y)$ se dice que es un umbral global; en el caso de que además dependa de $p(x, y)$, por ejemplo, el valor medio del entorno de vecindad, el umbral es denominado local y si depende también de la posición del píxel, se denominará dinámico. El problema es encontrar el umbral; operación nada sencilla ya que las imágenes están contaminadas con el ruido. Para acotar el estudio, sólo se centrará en las técnicas globales.

Selección óptima del umbral

Hay diferentes estrategias para la selección del umbral óptimo; entre ellas destacan aquellas cuyos fundamentos se basan en:

- Ajuste Gaussiano.
- Minimización de la varianza inter-clase.
- Entropía del histograma.
- Análisis de la concavidad.
- Métodos basados en momentos.

Un resumen detallado de los distintos algoritmos se puede encontrar en revistas especializadas. De entre todas las propuestas se va a destacar el umbral de Otsu, por su alta utilización en las aplicaciones de Visión Artificial.

2.2.3.3 Segmentación basada en homogeneidad

Segmentación de regiones por agrupación de píxeles

El método de crecimiento de regiones por agrupación de píxeles realiza la agrupación de regiones de modo que cada una de ellas se clasifica como objeto o como fondo, en dependencia de un umbral previamente definido. Este método es muy bueno debido a que no es muy complicado de programar y es bastante rápido, pero tiene como inconveniente que solo puede detectar regiones homogéneas (objetos de un solo color), bien definidos (que no exista un objeto cerca del otro) y trabaja contra un umbral manual, el cual es molesto ya que cada vez que cambia la iluminación en el área de trabajo hay que ajustarlo, esto trae consigo que determinados objetos de colores claros no se detecten.

2.3 Conclusiones

En este capítulo se han expuesto y descrito diferentes métodos de filtrado que fueron estudiados. En el transcurso de la investigación llegamos a la conclusión de que estos no son convenientes para la solución de la problemática planteada debido a que con ellos no se obtienen los resultados esperados. Estos no aportan mejoras al sistema de visión actual que presenta el robot, por lo tanto no formaran parte de nuestra propuesta.

CAPITULO 3: PROPUESTA DE FILTRADO

3.1 Introducción

Este capítulo tiene como objetivo proponer un filtro capaz de solucionar el problema del ruido causado por la iluminación en el sistema de visión de un robot móvil, para ello se describen métodos de filtrado a utilizar para la creación de un nuevo filtro.

El objetivo de este trabajo es demostrar que existen filtros que al trabajar en conjunto son capaces de solucionar el problema del ruido causado por la iluminación en el sistema de visión de un robot móvil.

Para un mejor rendimiento y entendimiento se dividió el PDI en 4 etapas. Como se había planteado anteriormente, este trabajo se centra en mejorar las etapas de filtrado y segmentación, las cuales tienen un papel importante dentro del proceso.

3.2 Deficiencias detectadas en el sistema de visión antiguo

El sistema de visión antiguo es bastante simple debido a que en la etapa de filtrado solo filtra utilizando escala de gris, por lo que solo elimina información irrelevante (se abordará con mayor profundidad en el epígrafe 3.3). Para la etapa de segmentación fue utilizado un método basado en la formación de regiones conocido como crecimiento de regiones por agrupación de píxeles. El mismo realiza la agrupación de regiones de modo que cada una de ellas se clasifica como objeto o como fondo, en dependencia de un umbral previamente definido. Este método es muy bueno debido a que no es muy complicado de programar y es bastante rápido, pero tiene como inconveniente que solo puede detectar regiones homogéneas (objetos de un solo color), bien definidos (que no exista un objeto cerca del otro) y trabaja contra un umbral manual, el cual es molesto ya que cada vez que cambia la iluminación en el área de trabajo hay que ajustarlo, esto trae consigo que determinados objetos de colores claros no se detecten. Ver figuras 37, 38.

Se puede contemplar la segmentación como un proceso que divide a R en n subregiones, R_1, R_2, \dots, R_n , de tal forma que:

$$a) \bigcup_{i=1}^n R_i = R,$$

- b) R_i es una región conexa, $i = 1, 2, \dots, n$,
- c) $R_i \cap R_j = \phi$ para todo $i \neq j$,
- d) $P(R_i) = \text{Verdadero}$ para $i = 1, 2, \dots, n$,
- e) $P(R_i \cup R_j) = \text{FALSO}$ para $i \neq j$

Donde R es la representación completa de una imagen, $P(R_i)$ es una propiedad de los puntos del conjunto R_i y ϕ es el conjunto vacío

Siendo R_i una región de tamaño N :

$$m = \frac{1}{N} \sum_p$$

entonces se dice que la región es homogénea si:

$\max |f(p) - m| < T$, donde T es un umbral y m es la media. (Boudy González & Leyva Fonseca, 2008).

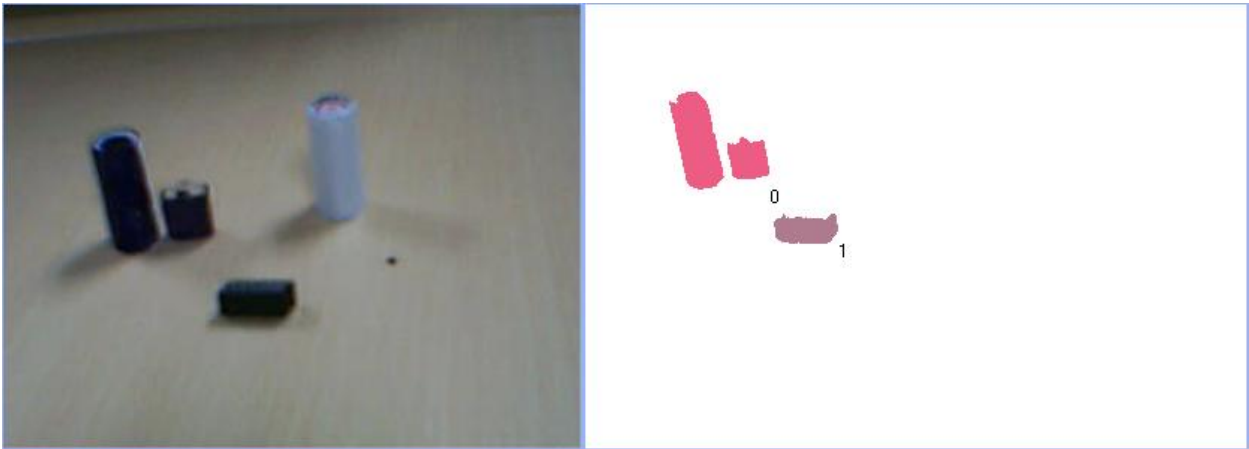


Fig 37 Imagen obtenida por el sistema de vision antiguo.

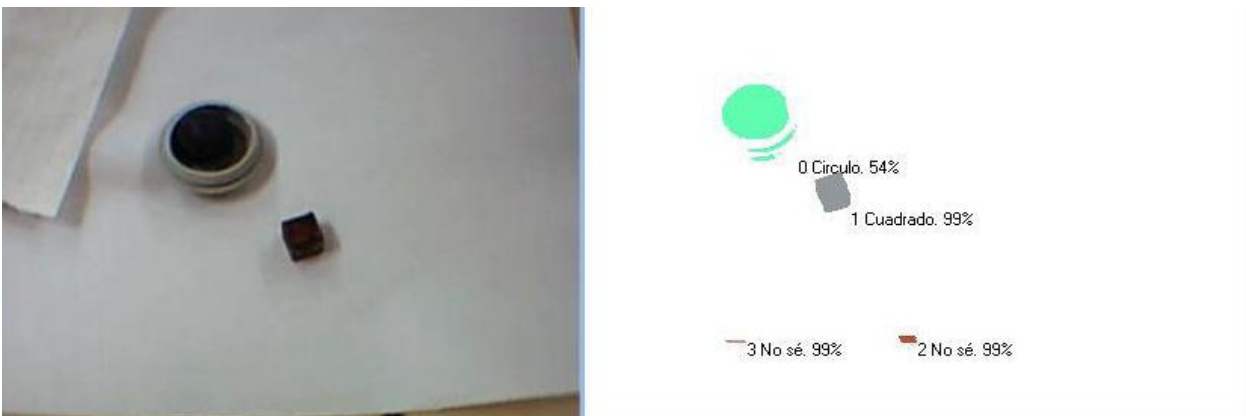


Fig 38 Imagen final obtenida por el sistema de vision antiguo.

3.3 Sistema de visión propuesto

3.3.1 Captura

Se obtiene una imagen desde una cámara digital (webcam).

Calibrar cámara

Para dotar de visión al robot móvil es preciso familiarizar al mismo con su entorno, en términos prácticos, esto no es más que calibrar la cámara. Para este fin se le muestra una hoja de papel con dos puntos del

sistema de referencia de la cámara, situados a una distancia conocida. En la figura 39 se muestra una representación del espacio de visión del robot (o sea el sistema de referencia de la cámara), donde C es el contorno que encierra el área de visión de la cámara, R es el punto de giro de la base del robot, y P1 y P2 son dos puntos conocidos alineados dentro del contorno C (Ver Figura 39). (Boudy Gonzales & Leyva Fonseca, 2008).

$$d = \quad (I)$$

$$M = \sqrt{\quad} \quad (II)$$

$$x_R = \quad (III)$$

$$y_R = \quad (IV)$$

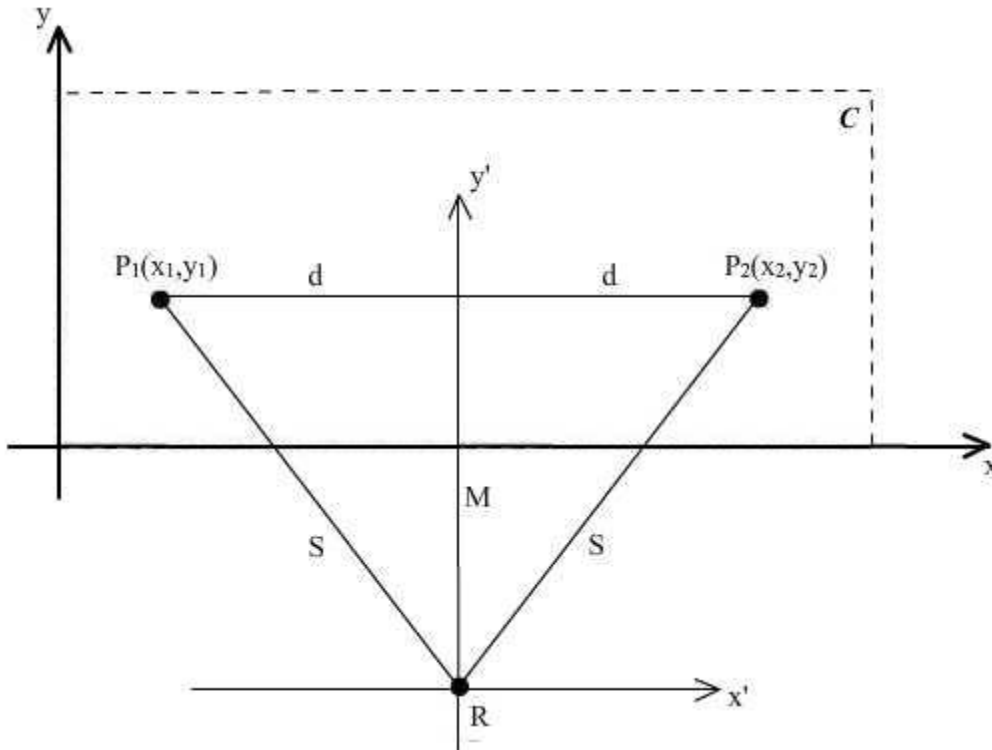


Fig 39 Representación de los sistemas de referencia de la cámara y el robot respectivamente. P1 y P2 son los puntos de calibración de la cámara.

Es necesario sacar la relación de píxeles por centímetro, para convertir la posición del objeto de las unidades de la cámara a unidades físicas reales.

$$D = x_2 \quad (V)$$

$$N_{p \ x \ cm} = \frac{D_p}{D_{cm}} \quad (VI)$$

Donde:

$N_{p \ x \ cm}$: cantidad de píxeles por cm

D : distancia entre dos puntos

D_p : distancia expresada en píxeles

D cm: distancia expresada en cm

3.3.2 Filtrado

Escala de grises

Un aspecto importante en el procesamiento digital de imágenes es la resolución espacial en tonos de grises en que la imagen va a ser subdividida. Esta resolución está en dependencia del tipo de digitalizador a utilizar, el cual a la vez está condicionado por la calidad de la imagen entrada, por las condiciones de iluminación y por los objetivos del procesamiento.

Definicion1

La resolución espacial de una imagen se define como la cantidad de filas y columnas en que es dividida una escena. (Rodriguez Morales, 2000)

Definicion2

La resolución en tonos de grises de una imagen se define como la cantidad de niveles de grises en que es dividida una escena. (Rodriguez Morales, 2000)

Un aspecto a tener en cuenta es que en una escena que tenga cambios ligeros en la iluminación es importante tener una fina cuantificación de los niveles de grises, es decir, la cuantificación de los niveles de grises debe ser alta. Ejemplos típicos de resoluciones son: 256X256X8, 512X512X8, 1024X1024X8, donde los primeros números significan la resolución espacial y se le llaman píxeles, mientras que el último número es la resolución en tonos de grises en base 2 que es conocida como el número de bits. Por ejemplo, 2^8 significa en primer lugar una resolución de 8 bits, o sea, 256 niveles de grises, donde 0 es el más oscuro y 255 es el valor más claro. Así existe $2^9, 2^{10}, 2^{11}, 2^{12}$. O sea, en la medida que aumente el exponente aumenta la resolución en los niveles de grises.

En computación una escala de grises es una escala empleada en la imagen digital en la que el valor de cada píxel posee un valor equivalente a una graduación de gris. Las imágenes representadas de este tipo

están compuestas de sombras de grises, que van desde el negro más profundo variando gradualmente en intensidad de grises hasta llegar al blanco.



Fig 40 Fotografía en escala de grises, en ocasiones se suele denominar en "blanco y negro".

La escala de grises utiliza matices de negro para representar objetos. Los objetos en escala de grises tienen un valor de brillo comprendido entre el 0% (negro) y el 100% (blanco). Las imágenes producidas con escáner de blanco y negro o de escala de grises se visualizan normalmente en escala de grises. La escala de grises también permite convertir ilustraciones en color en ilustraciones en blanco y negro de alta calidad. En este caso, Adobe Ilustrador elimina toda la información del color de la ilustración original. Los niveles de gris (matices) de los objetos convertidos representan la luminosidad de los originales. Al convertir objetos de escala de grises a RGB, se asigna a los valores de color de cada objeto los valores de gris de los objetos originales.

Cuando se convierte o se toma una foto en escala de grises se puede representar un conjunto de colores en un tono de gris, o incluso poner cada color en una intensidad. Las escalas de grises son diferentes de las fotografías en blanco y negro, en las que los colores se codifican en blanco o en negro; la escala de grises ofrece una gama de tonalidades de gris entre ambos.

En la mayoría de los contextos de fotografía (excepto en fotografía digital) se menciona como fotografía en "blanco y negro" una imagen que técnicamente es una escala de grises; en estos casos en los que se denomina "blanco y negro" se habla de monocromático aunque en algunos contextos de fotografía digital se expresa como sinónimo de escala de grises, mientras que en otros es sinónimo de "blanco y negro".

Las imágenes en escala de grises, emplean 8 bits para representar cada píxel lo que sólo permite una escala con 256 intensidades (o escalas de gris); es decir, 2 valores posibles para cada bit (0 y 1) elevado a 8 bits que se emplean para representar cada píxel, nos da 256 tonos de color diferentes que pueden representarse en una imagen en escala de grises.

Ejemplo



Fig 41 La misma foto en tres estados pictóricos o modos de color.

La foto de la izquierda es la foto original con todos sus colores representados. La foto del centro es la versión en escala de grises de la fotografía de la izquierda; todos los colores contenidos son negros blancos o una graduación entre los dos (es decir grises con diferente tonalidad). Por último, la fotografía de la derecha es en blanco y negro (monocromo), donde se representan los colores sin escala intermedia.

Monocromo: algo monocromático (del griego mono (uno) y croma (una, superficie o color de la piel)) tiene un solo color. El término monocromo puede referirse a:

En informática, monocromo tiene dos significados: puede indicar que se tiene un solo color que está encendido o apagado o también que tiene tonos intermedios, por lo que tiene la misma ambigüedad que blanco y negro. Para una imagen el término monocromo es esencialmente blanco y negro, pero a veces se prefiere el término monocromo para indicar que puede ser en realidad "blanco y verde", "verde y negro", etc.

Un monitor monocromo sólo dispone de un color, normalmente blanco o verde o color ámbar anaranjado, y normalmente también tonos intermedios (grises o tonos verdes más o menos

oscuros, respectivamente). Monocromático es algo de un solo color y se usa principalmente en las tarjetas de vídeo.

Monocromo se refiere a la radiación del haz del CRT ((Tubo de Rayos Catódicos)) al chocar contra la pantalla; por tanto debido a las diferencias de frecuencias que puede generar, pinta degradación de los colores comúnmente llamados colores intermedios. Es incorrecto decir que una fotografía, imagen o pantalla es Blanco y Negro cuando pueden apreciarse sobras o colores intermedios, pues esto es una imagen en escala de grises.

- Al aplicar el algoritmo se obtiene una imagen de 8 bits y 256 colores con el siguiente formato:
- $W \times H \times 8$
- $$I_f = \sum_{i=0}^W \sum_{j=0}^H (0.299P_r + 0.587P_g + 0.114P_b)$$
- Donde:
- W: ancho de la imagen en píxeles.
- H: largo de la imagen en píxeles.
- P_c: Valor del píxel, (r g b).
- I_f: Imagen final.



Fig 42 Imagen original.

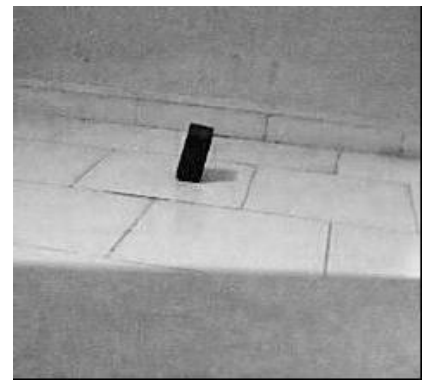


Fig 43 Imagen filtrada por escala de grises

3.3.3 Segmentación

Filtro de textura Anti-aliasing

Anti-aliasing

¿En qué consiste?

Por aliasing se conoce a un fallo en la representación de los gráficos debido a que la resolución final es finita. Existen diferentes tipos de aliasing visual, pero normalmente nos referimos al aliasing geométrico, efecto que consiste en la presencia de dientes de sierra en los bordes de los polígonos. El efecto de aliasing da a las imágenes una apariencia tosca. Por anti-aliasing se conoce al efecto de filtrar la imagen para suavizar los bordes, y disimular los bordes de los polígonos, consiguiendo una apariencia mucho más realista.

Este filtro trata de eliminar el efecto de escalonado que aparece en las líneas rectas que no son horizontales o verticales, tal efecto se conoce como aliasing. Es importante remarcar que este efecto no afecta a la geometría del objeto, si no a la forma en que éste se muestra por pantalla, este efecto se debe a las limitaciones que presenta el dispositivo de salida, es decir, la pantalla. Este filtro elimina el ruido conservando los detalles de la textura. Trata cada píxel de manera distinta dependiendo de los píxeles contiguos.

Técnicas de Anti-aliasing

Existen diversas técnicas de anti-aliasing, describiremos a continuación alguna de ellas:

- **Anti-aliasing de contornos:** Solamente se aplica en los contornos de los polígonos, no altera los objetos situados en los contornos planos. Genera una imagen detallada y poco borrosa.
- **FSAA:** (Full Screen Anti Aliasing) También conocido como supersampling, consiste en calcular la imagen a una resolución superior (2x ó 4x) y posteriormente mostrar la imagen por pantalla. Esto aumenta considerablemente la intensidad de los cálculos (se multiplican por 16). Este tipo de filtrado ha sido utilizado por Nvidia en algunos de sus modelos.
- **HRAA:** (High Resolution Anti Aliasing) En lugar de calcular la imagen a una resolución mayor, utiliza una técnica conocida como Quincunx, que calcula el color de un píxel haciendo un promedio del valor de los píxeles que le rodean. De este modo se obtiene un resultado semejante al anterior pero se disminuye la cantidad de cálculo.

- **PFSAA:** (Pseudo Full Screen Anti Aliasing) Este filtro se aprovecha de las características de una pantalla de TV, dado que ha sido desarrollado para evitar el aliasing de la PS2. En lugar de actuar sobre los píxeles antes de que estos se envíen al dispositivo de salida se actúa sobre la propia señal mediante HW.

La ventaja es que conseguimos una calidad de imagen con un rendimiento muy superior.



Fig 44 Imagen con aliasing. Se pueden apreciar claramente aristas en las varas del camión.



Fig 45 La escena anterior después de aplicar un filtro de anti-aliasing (FSAA 4x).

Método Otsu

Los métodos de filtrado de valor umbral son un grupo de algoritmos cuya finalidad es la de segmentación de gráficos, es decir separar los objetos de una imagen que nos interesen del resto. Con la ayuda de los métodos de segmentación 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. Entre los métodos de segmentación de valor umbral se hace énfasis sobre el método de Otsu. Los métodos de valor umbral pertenecen a los métodos más antiguos de tratamiento de imágenes digitales. El método de Otsu, el cual es explicado a continuación, fue publicado en el año 1979 por Nobuyuki Otsu. No obstante existen publicaciones aún más antiguas sobre este tema. Los métodos de valor umbral son

especialmente sencillos lo cual permite implementarlos rápidamente y obtener resultados con relativamente poco esfuerzo lo cual ha contribuido a su aceptación.



Fig 46 Imagen original.



Fig 47 Imagen utilizando el algoritmo de Otsu.

En la visión artificial y procesamiento de imágenes, el método de Otsu (un umbral de imagen mejorado), se usa para realizar automáticamente histograma basado en la forma de imagen umbral o la reducción de graylevel (nivel de grises) de una imagen a una imagen binaria. El método de filtrado asume que la imagen contiene dos tipos de píxeles (por ejemplo, primer plano y fondo) y luego calcula el umbral óptimo de las dos clases que separa a fin de que su propagación combinado (intra-clase de la varianza) sea mínima.

El umbral es un simple pero efectivo método para separar los objetos del fondo. Comúnmente es utilizado el método de Otsu para mejorar la imagen de segmentación, éste es más simple y más fácil de aplicar. La segmentación de imágenes es uno de los difíciles problemas de investigación en la visión y la industria de reconocimiento de imágenes. Los experimentos indican que este algoritmo de segmentación tiene ventajas en tiempo real y ciertas habilidades anti-ruido, ya que el objetivo se puede extraer con mayor precisión. Por lo tanto, en la fase de reconocimiento el objetivo será simple y fiable.

Otsu es un método de búsqueda exhaustiva para el umbral que minimiza la varianza intra-clase. Se define como una suma ponderada de las diferencias de las dos clases.

Este método se basa en la suposición de que la función de densidad del fondo,

C_f , y de los objeto, C_o , tienen modelo gaussiano, $N(\mu_f, \sigma_f^2)$ y $N(\mu_o, \sigma_o^2)$. Cada grupo estará formado por los niveles de grises fijados por el umbral, T :

$$C_f = \{0,1,2,\dots,T\} \quad C_o = \{T+1,T+2,\dots,I-1\}$$

El umbral debe de minimizar la suma ponderada de cada una de las varianzas de las clases, C_f y C_o , ya que se supone que conforme la suma de las dos normales se aproxime más al histograma real, las desviaciones serán menores.

Para determinar los coeficientes se toman las probabilidades de cada una de las clases. Considerando un valor fijo de umbral, T , las probabilidades de cada categoría serán

:

$$P_{C_f} = \sum_{i=0}^T p_i \quad P_{C_o} = \sum_{i=T+1}^{I-1} p_i$$

Donde p_i es la probabilidad de la intensidad i en la imagen. Las medias y varianzas de cada grupo corresponderán a:

$$\begin{aligned} \mu_{C_f} &= \frac{1}{P_{C_f}} \sum_{i=0}^T i \cdot p_i & \mu_{C_o} &= \frac{1}{P_{C_o}} \sum_{i=T+1}^{I-1} i \cdot p_i \\ \sigma_{C_f}^2 &= \frac{1}{P_{C_f}} \sum_{i=0}^T (i - \mu_{C_f})^2 \cdot p_i & \sigma_{C_o}^2 &= \frac{1}{P_{C_o}} \sum_{i=T+1}^{I-1} (i - \mu_{C_o})^2 \cdot p_i \end{aligned}$$

Siendo entonces la varianza ponderada:

$$\sigma_p^2 = P_{C_f} \cdot \sigma_{C_f}^2 + P_{C_o} \cdot \sigma_{C_o}^2$$

Se recorre todo el rango de niveles de gris, de T igual 0 a $I-1$, calculándose la varianza ponderada y se elige el umbral, T , que minimiza este valor.

El filtrado de imágenes juega un papel importante en el análisis de imágenes y en el sistema de visión por computador. Entre todas las técnicas de filtrado, los métodos de umbral automático son ampliamente utilizados debido a sus ventajas de fácil aplicación y ahorro de tiempo. Otsu es un método de umbral utilizado con frecuencia en diversos campos. En dos dimensiones (2D) el método Otsu se comporta bien en el filtrado de imágenes de baja relación señal-ruido de una dimensión (1D). Pero da resultados sólo cuando el número de píxeles en cada una de ellas están muy cerca una de otra. Si bien hay diferentes métodos para hallar un umbral, la mayoría de ellos no dan buenos resultados cuando se trabaja con imágenes del mundo real debido a la presencia de ruido, histogramas planos o una iluminación inadecuada. Por el contrario, el método de Otsu fue uno de los mejores métodos de selección de umbral para imágenes del

mundo real. Sin embargo, como anticipamos, este método usa una búsqueda exhaustiva para evaluar el criterio para maximizar la varianza entre clases. A medida que el número de clases de una imagen aumenta, el método de Otsu necesita mucho más tiempo para seleccionar un umbral multinivel adecuado. Para determinar el umbral de una imagen eficientemente, se propone hacer uso de los filtros antes mencionados en este capítulo. El uso de estos filtros podría disminuir considerablemente el tiempo de cálculo.

La importancia del método de Otsu radica en que es automático, es decir, no necesita supervisión humana ni información previa de la imagen antes de su procesamiento.

Fusión de métodos de filtrados propuestos

El método de empleo de los filtros propuestos después de haber sido capturada la imagen por el sistema de visión del robot para el cual se hace la investigación, sería primeramente en la etapa de filtrado aplicar a esta imagen el filtro por escala de grises, ya que este permite quedarse con la tercera parte de la imagen o sea solo 8 bits de 24 bits que por lo general tienen las imágenes a color, se obtiene así una imagen de un solo color gris, reduciendo de este modo en gran parte el tiempo de procesamiento que podría demorar la computadora. En la etapa de segmentación aplicar el filtro de textura Anti-aliasing para eliminar el efecto escalonado que pueda existir en la imagen. Por último en esta misma etapa aplicar el método de Otsu para obtener así el umbral de la imagen eficientemente. Aunque este método no es perfecto, es robusto ante la presencia de ruido en la imagen y ante la iluminación inadecuada este es el ruido que se trata de eliminar ya que afecta el sistema de visión del robot. Por lo tanto después de aplicados estos filtros se espera obtener una imagen lo más libre de ruido posible y lista para la fase de reconocimiento.

Dado que no se puede comprobar en este apunte la eficiencia y rapidez del método Otsu para calcular el umbral óptimo, tendremos que confiar en el gran uso que se hace de él en la industria a la hora de binarizar imágenes reales en tiempo real y que están inmersas en todo tipo de perturbaciones. Se puede extender su utilidad para aquellos casos en que la iluminación no es uniforme o es en general inadecuada.

3.4 Conclusiones

Esta investigación arroja como resultado que el filtrado por escala de grises y crecimiento de regiones por agrupación de píxeles no es suficiente debido a los errores que introducen. Por lo que la implantación de esta propuesta supone una mejora ya que propone utilizar parte de los filtros ya implementados (escala de grises) y otros nuevos como Filtro de textura Anti-aliasing y Otsu.

Conclusiones Generales

Después de haber realizado una vasta investigación de todos los filtros para la eliminación del ruido, se propone en este trabajo un método de filtrado combinando la robustez y capacidad de los filtros Anti-aliasing y Otsu para solucionar el problema del ruido en el sistema de visión del robot móvil para el cual va dirigida la investigación, cuyo punto fuerte es mejorar el sistema de visión en sus etapas de filtrado y segmentación. El filtro de textura Anti-aliasing permite filtrar la imagen para suavizar los bordes consiguiendo una apariencia mucho más realista. El método de filtrado Otsu nos permite obtener un umbral de imagen mejorado. La combinación de ambos filtros es simple es posible aplicarlos no solo en este sistema, si no, en cualquier otro proyecto que requiera de las técnicas de análisis y procesamiento de imágenes. El estudio de estos filtros dará una idea más clara acerca de cuál es mejor, para qué propósito y cuál se recomendaría para implementarlo en una aplicación real.

Recomendaciones

En aras de mejorar el sistema de visión luego de haber realizado el estudio de la investigación y la propuesta de un filtro que posibilita el mejoramiento de este, se recomienda la implementación del filtro propuesto.

Referencias Bibliográficas

1. Andrew P. Duchon, W. H. (1998). *Ecological robotics. Adaptive Behavior . Special Issue on Biologically Inspired Models of Spatial Navigation.*
2. Artal, C. G. (2003). *Seguimiento Visual Artificial.* Retrieved from <http://dialnet.unirioja.es/servlet/articulo?codigo=1973168>
3. Garcia, C. U. *Introduccion a la Robotica .*
4. Horswill, I. (Abril Septiembre 1997). *Visual architecture and cognitive architecture. Journal of Experimental and Theoretical AI, 9(2-3):277-292.*
5. M.C. García-Alegre, P. B. (October 1995). *Complex behaviour generation on autonomous robots: a case study. In Proceedings of the IEEE International Conference on Systems, Man and Cybernetics.* Vancouver (Canada).
6. N.J.Nilsson. (1969). *A mobile automaton:an application of artificial intelligence Techniques. In Proceedings of the 1st Internatinal Joint Conference on Artificial Intelgence IJCAI.* Washington,(USA),1969.
7. R, F. L. (2009). *Procesamiento de caras:Bases Neurológicas, trastornos y evaluación.* Retrieved from <http://neurologia.rediris.es/congreso-1/conferencias/neuropsicologia-2-2.html>
8. Rodríguez Morales, R. (2000). *Monografía del curso de procesamiento de imágenes.* Grupo de Procesamiento Digital de Señales. Instituto de Cibernética, Matemática y Física (ICIMAF).
9. Ronald C. Arkin, M. F. (2003). *An Tethological and emotional basis human-robot interaction. Robotics and Autonomous Systems (Vol. 42).*
10. Sergio Orlando, B. G., & Milagro Leonor, L. F. (2008). *Desarrollo de una aplicación para el control inteligente de un robot manipulador.* Ciudad de La Habana.

Bibliografía General

Gonzalo Pajares, Jesús M. de la Cruz. (2001). *Visión por Computador. Imágenes digitales y aplicaciones*. Ra-Ma,

Sezgin Sankur y B. (2004). "Estudio sobre las técnicas de imagen umbral y cuantitativos de evaluación del desempeño". *Imágenes Electrónicas de Diario* 13 (1): 146-165. Doi: 10.1117/1.1631315.

Otsu (1979). "Un método de selección del umbral de nivel de gris histogramas". *IEEE Trans. Sys., Hombre., Cyber.* 9: 62-66. Doi: 10.1109/TSMC.1979.4310076.

Liao, Ping-Sung ; Tse-Sheng, Chen & Pau-Choo, Chung. (2001). "Un algoritmo rápido para multinivel Thresholding". *J. Inf. Sci.. Eng.* 17 (5): 713-727.

Canny. (1986). *A Computational Approach to Edge Detection*, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(6), pp. 679-698.

Alvarez Rubiños, Roman. (Septiembre 2004). *Procesado de la imagen*.

Rodríguez Araújo, Jorge. (Marzo 2009). *Visión Artificial –Introducción a los sistemas de visión artificial Version 1.4* .

Escolano, F; Colomina, O & Cazorla, M. A. (2005-2006). *Técnicas de Inteligencia Artificial – Restauración de Imágenes Version 1*.

Luis Alba, Jose & Rey, Juan Carlos. (Marzo 2006). *Métodos de análisis de imágenes* .

Lopez Rial, Miguel Angel & Lopez Labarta, Rafael. *Filtros de texturas*. Huelker. (Marzo 21, 2007). *Selección Salmon. Conceptos de robotica y robot*.

Anexos

A continuación se mostrará una función hecha en un script de MatLab que calcula el umbral óptimo para binarizar una imagen utilizando el método de Otsu.

```

% Calculo del umbral optimo segun Otsu
sigmaB2max = 0;
sigmaB2 = 0;
mut = mu(256);
umbralOptimo = 0;
% Intensidad media total de la imagen

for t=1:256
    omegal = omega(t);
    omega2 = 1 - omegal;
    if omegal~=0 & omega2~=0
        mu1 = mu(t) / omegal;
        mu2 = (mut-mu(t)) / omega2;

        sigmaB2 = omegal * (mu1-mut)^2 + omega2 * (mu2 - mut)^2;
        if sigmaB2 > sigmaB2max
            sigmaB2max = sigmaB2;
            umbralOptimo = t-1;
        end
    end
end

% Calculo de la distribucion de probabilidad de los niveles de gris
P=zeros(1,256);
for i=1:f*c
    P(Imagen(i)+1) = P(Imagen(i)+1) + 1;
end
P = P / (f*c);

% Calculo del momento acumulado de orden cero y del de primer orden
omega = zeros(1,256);
omega(1) = P(1);
mu = zeros(1,256);
for i=2:256
    omega(i) = omega(i-1) + P(i);
    mu(i) = mu(i-1) + (i-1)*P(i);
end

```



Fig 48 Imagen original que se desea binarizar.

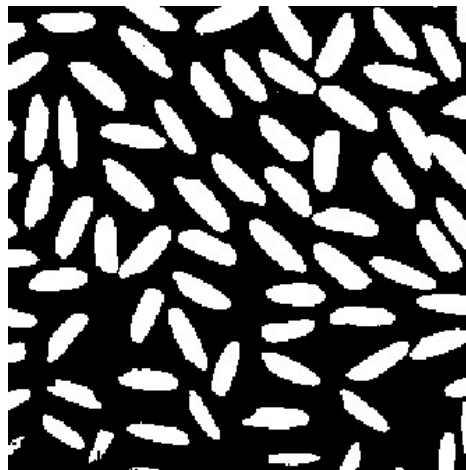


Fig 49 Imagen resultante de binarizar la original usando un umbral arbitrario.

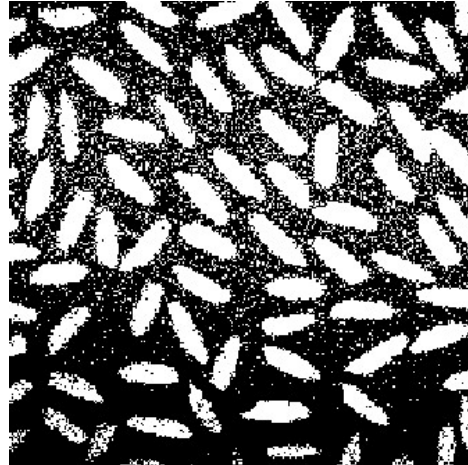


Fig 50 Imagen resultante de binarizar la imagen con ruido gaussiano usando el mismo umbral arbitrario.

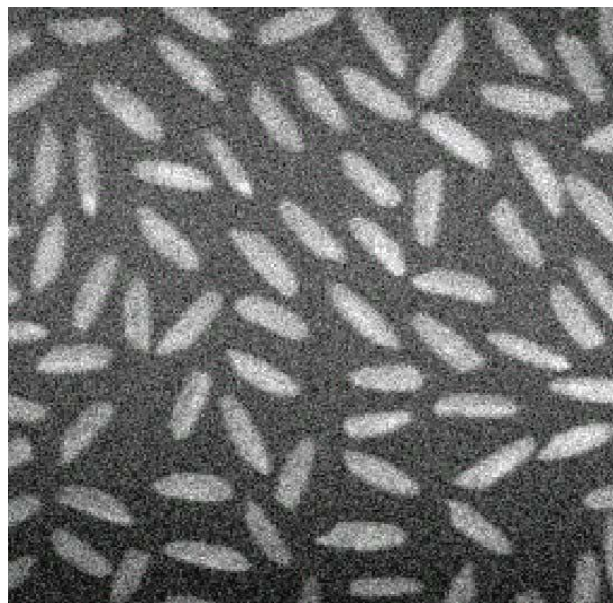


Fig 51 Imagen con ruido gaussiano de media cero y varianza 20 que se desea binarizar.

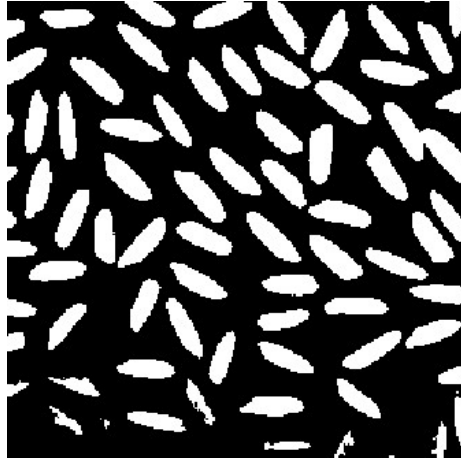


Fig 52 Imagen resultante de binarizar la original usando un umbral calculado con el método de Otsu.

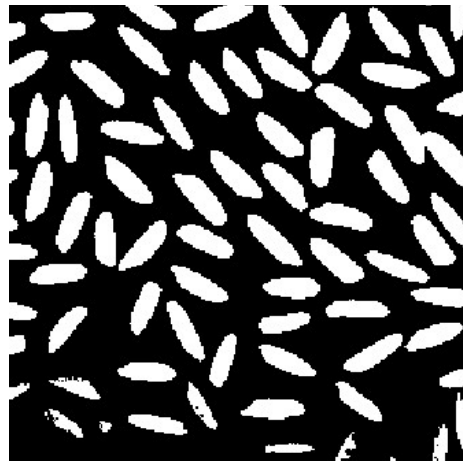


Fig 53 Imagen resultante de binarizar la imagen con ruido gaussiano mediante Otsu.

Glosario de términos

ⁱ **Imagen:** Llámese imagen a una representación gráfica plasmada en una superficie y que es copia de otra original localizada a menudo distante de la imagen a una cierta distancia pero sin ser esta copia necesariamente de iguales dimensiones ni idéntica orientación que el original.

ⁱⁱ **Cámaras CCD:** Las cámaras CCD son dispositivos electrónicos muy sensibles, ideados para captar la luz y formar una imagen a partir de ella, CCD significa “Dispositivo de Carga Acoplada”. Las cualidades operativas de la cámara CCD, tales como su bajo costo, facilidad de operación, durabilidad, tamaño, poco consumo de energía y alta resolución, la han convertido en la cámara estándar para los sistemas de procesamiento de imágenes y sistemas de visión.

ⁱⁱⁱ **Sensoramiento Remoto:** Es una Ciencia (y en cierta forma un ARTE), de adquirir información de la superficie terrestre sin tener un contacto directo con ella.

^{iv} **Objeto:** Agrupación de píxeles dentro de la imagen y su posición.

^v **ICIMAF:** Instituto de Cibernética, Matemática y Física.

^{vi} **Visión por Computadora:** La visión por computadora es una rama de la inteligencia artificial que tiene por objetivo modelar matemáticamente los procesos de percepción visual en los seres vivos y generar programas que permitan simular estas capacidades visuales por computadora.

^{vii} **Área de trabajo:** Espacio donde se encontrarán ubicados los objetos.

^{viii} **Nanomotores:** Dispositivo capaz de convertir energía en movimiento.

^{ix} **Reflectancia:** La reflectancia es la proporción de luz reflejada por una superficie y se determina comparando los lúmenes que inciden en ella (iluminancia) con los que refleja (luminancia).

^x **Crominancia:**La crominancia es el componente que contiene la información sobre el color de una señal de vídeo.

^{xi} **Difuminar:** Disminuir la intensidad de un color. Perdida de la claridad y exactitud. de los objeto, las líneas y colores de una imagen.

^{xii} **Renderizar:**Es la acción de asignar y calcular todas las propiedades de un objeto antes de mostrarlo en pantalla.

^{xiii} **Primera derivada:**1-Cuando la derivada es positiva la función crece.2-Cuando la derivada es negativa la función decrece.3-Cuando la derivada es cero la función tiene un máximo o un mínimo.