1. **Etapas del procesamiento digital de imágenes.**

El procesamiento digital de imágenes reúne un vasto número de conceptos relacionados desde la etapa física, pasando por el algoritmo de procesamiento y el desarrollo de teoría para llegar a la solución del problema propuesto. Por lo anterior, se identifican 5 etapas consecuentes que integran una tarea de visión artificial.

La etapa inicial es la adquisición de la imagen digital. Se emplean dispositivos de captura de imágenes con capacidad de digitalizar la señal producida y se implementan metodologías de adquisición. Para el presente proyecto, se usará un banco de imágenes desarrollado en el 2019 que detalla una metodología rigurosa en la captura de imágenes (Alfaro Mejía, 2019).

Después de la consecución de la imagen, el pre - procesamiento mejora las condiciones de la imagen con el fin de filtrar el ruido. Seguido, la segmentación tiene como objetivo dividir dicha imagen en partes u objetos que la conforman destacando las partes de estudio, este paso es de vital importancia en el proceso; una buena segmentación facilitara la solución del problema, de lo contrario, dificultara el desarrollo de las etapas posteriores. Es necesario obtener los objetos o rasgos de interés de la imagen segmentada, para ello es preciso seleccionar un método para identificar los datos de interés, este se encarga de extraer facciones que producen información relevante de carácter cuantitativo otorgando rasgos básicos para la selección de clases u objetos. El reconocimiento y la interpretación es la etapa epilogo; este proceso brinda etiqueta y significado a un objeto mediante la información resultante de los determinadores de objetos. (EDMANS, 2006)



Ilustración Diagrama de etapas del procesamiento de imágenes. Fuente: propia.

* 1. **Etapa 1: Captura.**
     1. Adquisición de imagen.

La visión artificial se compone de sensor de imagen, iluminación, proceso de comunicaciones y procesamiento digital de imagen. Para la adquisición de imágenes se resalta el factor de iluminación que permite destacar las características del objetivo, la cámara captura la imagen y transforma la luz en señal digital que luego se podrá procesar mediante algoritmos que revisa la imagen.

Los componentes físicos (hardware) son productos comerciales de fácil adquisición, es decir, Los sistemas de procesamiento de imágenes se pueden construir a partir de estos productos, creando un solo módulo integrado de visión artificial. (Cognex, 2016)

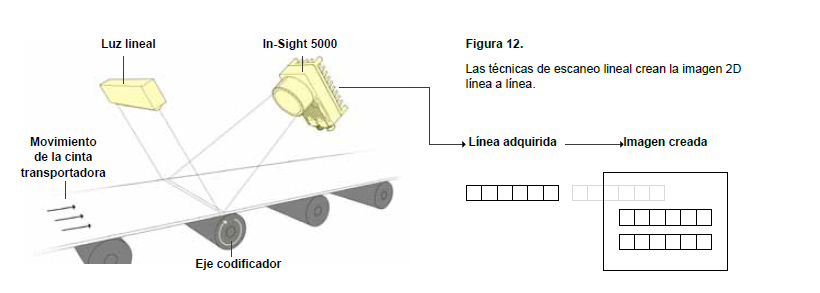


Ilustración Ejemplo de sistema integrado de visión artificial. Fuente: (Cognex, 2016)

* 1. **Etapa 2: Pre – Procesamiento.**
     1. Filtro promediador.

Perteneciendo a la categoría de los filtros espaciales lineales. El filtro de la media o promediador determinado por la ecuación 1. Construye del cálculo de la media aritmética de los pixeles que están en una máscara en un punto determinado de la imagen. En dicha máscara de tamaño con centro en . El filtro sustituye el valor de por el promedio de la máscara.

Ecuación Filtro promediador

* + 1. Filtro Gaussiano.

Es un operador 2D implementado para eliminar ruido y suavizar bordes, similar al filtro de la media, pero su representación es a través de una campana de gauss. La idea es utilizar la distribución gaussiana bidimensional como función de una desviación puntual. El grado de suavizado se establece, según se define en la ecuación 2, mediante la desviación estándar (sigma) en proporción directa al suavizado de la imagen, es decir, a mayor valor de la desviación estándar, mayor suavizado de la imagen.

La diferencia con el filtro de la media radica en los pesos que tienen los pixeles centrales y adyacentes. El promediador les otorga igual peso a todos los pixeles, en cambio, la salida del filtro gaussiano pondera con más peso al pixel central. (Aguirre Dobernack, 2013)

Ecuación Filtro Gaussiano.

* + 1. **Filtro de mediana.**

Hace parte de la categoría de los filtros no lineales. En este caso, no opera los valores que están en la máscara con un promedio ponderado. El filtro de mediana es el filtro no lineal más conocido, este operador crea nuevos pixeles a partir de calcular la mediana del conjunto de pixeles que constituyen la máscara tal como se expresa en la ecuación 3.

Ecuación Filtro de mediana.

Donde es el valor del filtro de mediana recorriendo la imagen pixel a pixel calculado con los pixeles vecinos. En la ilustración 3 se representa un ejemplo de este filtro en una ventada de 3x3.



Ilustración Ejemplo de filtro mediana en una ventana 3x3. (Fuente propia)

* + 1. **Filtro Bilateral.**

Es un filtro que suaviza la imagen conservando los bodes según la combinación no lineal de lo pixeles adyacentes. Tiene en cuanta los criterios geométricos y fotoeléctricos, es decir, utiliza las distancias espaciales y sus diferentes intensidades.

Ecuación Determinación de filtro bilateral.

Ecuación Asignación de peso Wp.

Ecuación El coeficiente de peso w (i, j, k, l).

Ecuación Filtro bilateral normalizado después de calcular los pesos.

* 1. **Etapa 3: Segmentación.**
     1. **Filtro laplaciano.**

Este operador laplaciano se construye en la segunda derivada de la función . Cuando en la primera derivada el valor es un máximo, su resultado es cero. (Serrano, 2003) Su representación matemática se ilustra en la ecuación 8.

Ecuación filtro laplaciano.

Si se toma una celda de 3 x 3, se aplica la primera derivada a cada píxel:

Se observa que y que entonces: . Se obtiene el operador laplaciano resalta en cambio de intensidad en la celda.

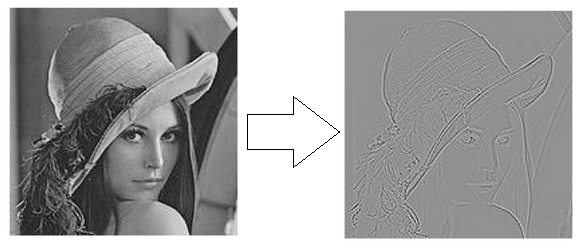


Ilustración En la izquierda, la foto origina. A la derecha, laplaciano de la imagen. Fuente: (Serrano, 2003)

* + 1. **Detector de bordes: Canny.**

El algoritmo detector de bordes Canny se basa en 3 aspectos; detección que no elimina bordes importantes y suprime falsos bordes, la distancia entre la localización de posición real y la localización del borde es mínima y la respuesta que integra múltiples resultados a un solo borde.

Los detectores de bordes que usan la primera derivada como el algoritmo de Canny, ilustrada en la ecuación 9, establecen valores en ceros a las regiones que no varían en intensidad y, cuando hay transición de intensidad, sufre una gran variación en la primera derivada. (Valverde Rebaza, 2007)

Ecuación definición del vector bidimensional mediante la primera derivada.

Para emplear el algoritmo detector de bordes de Canny, se deben seguir 3 pasos; obtener el gradiente que se define como magnitud y dirección, supresión no máxima para lograr un borde de un pixel de ancho e histéresis de umbral para reducir la aparición de contornos falsos. (Valverde Rebaza, 2007)

El gradiente está definido por las ecuaciones 10 y 11. Donde el valor absoluto de es la magnitud y indica la variación de la dirección:

Ecuación magnitud del vector bidimensional G

Ecuación dirección del vector bidimensional G

La supresión no máxima toma los valores de salida de magnitud y dirección del gradiente en el vector considerando 4 direcciones en el eje (. Localiza la dirección que más se le aproxima al gradiente del vector que resulta perpendicular al borde. Si la magnitud del vector es más pequeño que al menos uno de sus vecinos, se asigna el valor de cero a ese pixel.

La histéresis de umbra toma la imagen obtenida por la supresión no máxima que suele tener máximos locales por ruido. Se puede determinar un umbra para eliminar los pixeles por debajo de determinado intervalo; donde La imagen se explora en todos los puntos siguiendo un orden, tal que , este criterio sigue una cadena de máximos perpendicular al borde y enlista los puntos que . Este procedimiento minimiza el obtener contornos falsos. (Valverde Rebaza, 2007)

* + 1. **Transformada de Hough.**

Esta herramienta permite detectar linear y curvas en una imagen. La transformadas de Hough es fuerte al ruido. Para esta técnica se obtiene la imagen binaria el los pixeles que conforman la frontera. El objetivo de esta herramienta es detectar los puntos pertenecientes a una recta que existan en la imagen.

Ecuación Ecuación de la recta polar.

En la ecuación 12, se define la ecuación de la recta polar que realiza la transformación de coordenadas en puntos de referencia en el plano polar También se debe crear una serie de celdas acumuladas identificadas como y , representado en la ilustración 5. Posteriormente, se evalúa la ecuación de la recta por cada uno de los puntos en la imagen , donde el un número alto de intercepciones en las celdas acumuladas indican que dicho punto pertenece a una recta. (Gómez Trejos & Guerrero Guzmán, 2016)



Ilustración Representación el plano polar de las celdas acumuladas. Fuente: propia.

* 1. **Etapa 4: Identificación de objetos.**
     1. HOG (Histograma orientado a gradientes).

El HOG, es un descriptor que detecta la distribución de gradientes alojados en una imagen. Plantea que la apariencia y forma de un objeto se pueden describir usando un histograma construido por las direcciones de los bordes del objeto.

Para desarrollar el descriptor es necesario hallar en primera instancia el gradiente. En la práctica, el gradiente es el cambio de dirección en la intensidad de una imagen que consta de magnitud y dirección. Este vector se calcula mediante la diferencia de intensidades de un píxel con sus vecinos en horizontal y vertical, se desarrolla en las siguientes ecuaciones:

Ecuación cálculo del gradiente.

En la construcción del histograma es necesario definir celdas en la imagen. Las celdas son divisiones cuadradas fijas entre 6 y 8 pixeles que se despliegan por toda imagen. También se debe seleccionar el intervalo de las orientaciones, usualmente, se consideran 8 intervalos de 0° a 180°. Normalmente no se toma el rango de 0° a 360° porque se emplea el gradiente sin signo. (Lara-Lévano, 2019)

El cálculo de determinador se hace mediante bloques de celdas normalizados. Los bloques son agrupación de celdas generados para combatir la invarianza de aspectos de la imagen. La normalización se obtiene del vector que contiene los histogramas encadenados de un bloque, a esto se desarrolla tal vector dividiéndolo sobre una norma expresado en la ecuación 14. (Lara-Lévano, 2019)

Ecuación Normalización del vector de la imagen.

En , donde k es la norma y es una constante que interviene para que el gradiente no sea igual a cero. La norma del vector se encuentra con la siguiente ecuación: (Lara-Lévano, 2019)

Ecuación Norma del vector

* + 1. SIFT (Scale Invariant Feature Transform).

La función de transformación de escala invariable (SIFT), presentada por David Lowe en 1999, es un método de detección de puntos característicos de una imagen que seguido se describen utilizando HOG. El método presenta en la localización y descripción una invarianza con respecto a la orientación, posición y escala. SIFT fue determinado para imágenes en escala de grises, por lo tanto, es definido mediante su vector de 128 elementos, obteniendo su posición en coordenadas, escala y orientación dominante de la imagen y la región alrededor de dicho punto.

Para emplear descriptores SIFT en una imagen se debe obtener los puntos característicos, posteriormente, es necesario calcular el vector descriptor para cada punto de interés a partir de la información de los píxeles que lo rodean. En consecuencia, el algoritmo consta de obtención de los puntos característicos y descripción de la región que rodea cada punto de interés.

La obtención de los puntos característicos consiste en detectar las regiones con diferencias de gradiente significativas en ambos lados del punto de interés. Para ello se estudia y compara la diferencia gaussiana como posible método para obtener los puntos característicos (Ilustración 2). Finalmente se buscan pequeñas regiones con un nivel de intensidad estable y alrededor de las cuales se producen una variación significativa del gradiente en más de una dirección. (Alegre Gutierrez, 2016)

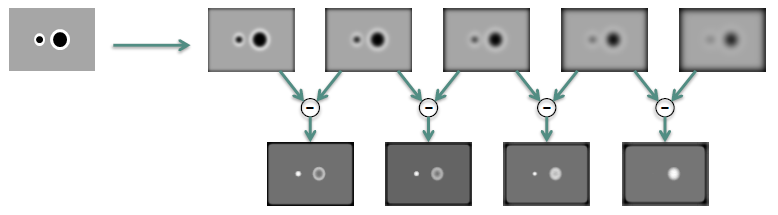


Ilustración Empleo de diferencia gaussiana para obtener puntos de interés. Fuente: (Valveny, 2021)

Para obtener el descriptor de la región que rodea cada punto de interés, es imperante calcular un histograma de direcciones del gradiente local. Por lo tanto, se obtiene un descriptor invariante a la escala al estabilizar el tamaño de los píxeles que rodena el punto de interés. También es invariante a la rotación porque se extrae la orientación dominante del gradiente en el vecindario del punto característico lo que da la orientación de la rejilla para calcular el histograma.

El espacio escala en SIFT de cada imagen se construye mediante convolución con diferentes Gaussianas, se crean varios espacios suavizando el tamaño de la imagen original. Estos espacios toman el nombre de **octava** y se obtienen eliminando una de cada 2 filas y columnas sobre la imagen con respecto a la octava anterior, de esta forma se reduce las dimensiones a la mitad. El proceso tiene como nombre; obtención de puntos característicos a partir de los extremos del espacio escala, generados de la diferencia gaussiana dentro de una pirámide de diferencia gaussiana. (Alegre Gutierrez, 2016)

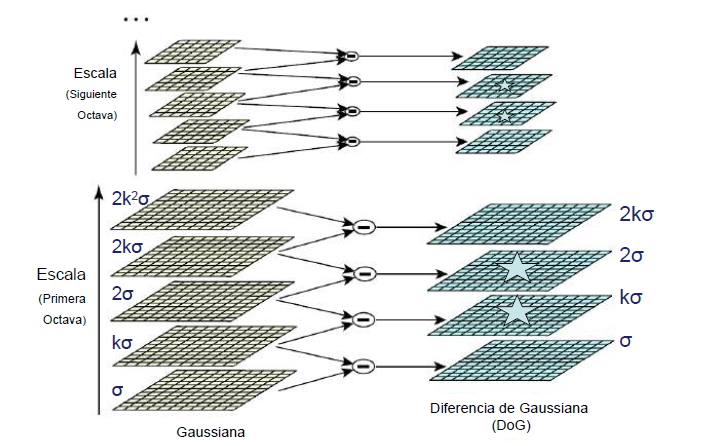


Ilustración Diagrama ilustrativo de la pirámide de las diferencias gaussianas compuestas de espacios - escalas. Fuente: (Alegre Gutierrez, 2016).

* + 1. SURF (Speeded Up Robust Features).

El SURF, es otro detector y descriptor de puntos característicos muy popular en la visión artificial. El algoritmo tiene alto rendimiento en los puntos de interés, tomando como base el SIFT, tiene en común la particularidad que son invariantes al escalado, orientación e iluminación de la imagen. La diferencia radica en que el SURF mejora la velocidad de procesamiento de su antecesor reduciendo la complejidad del cálculo y la dimensión de vectores utilizados. (Roos Hoefgeest Toribio, 2017)

La localización de los puntos de interés se realiza hallando el determinante de la matriz Hessiana la cual genera una aproximación para seleccionar la escala y la ubicación de los puntos.

Ecuación Matriz Hessiana aplicado al punto y escala de una imagen.

En la ecuación 16, es el punto en una imagen y es la matriz Hessiana entre y la escala . , es filtro Gaussiano de segundo orden aplicado a dicha imagen en el punto . Esto se repite para , . El determinante de define el cambio local alrededor del punto .

Es pueden implementar filtros Gaussianos de segundo orden con mejores resultados mediante el cálculo de imágenes integrales. Las integrales de una imagen se obtienen de la siguiente forma (García Merino, 2016):

Ecuación imagen integral.

Siendo la imagen . Se requieren 4 puntos de evaluación para obtener la suma de la imagen original en un rectángulo T:

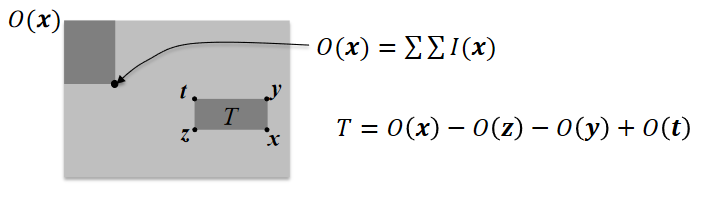


Ilustración Ejemplo de uso de imagen integral. Fuente: (Valveny, 2021)

Empleando los filtro Gaussianos de segundo orden y el concepto de imagen integral, se deja de lado el escalado de la imagen y los bucles de procesamiento a cada nivel de la piramide, como en SIFT, para analizar la imagen escalando el tamaño del filtro en la imagen original. El punto de inicio puede ser un filtro de 9 X 9 con escalado de 1.2 y, gradualmente, se aplican capas más grandes de filtro. Al subir una octava se duplica el tamaño del filtro. (Roos Hoefgeest Toribio, 2017)

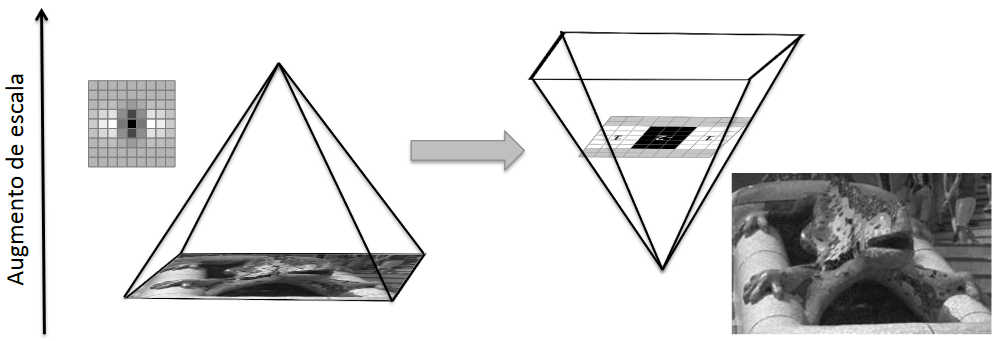


Ilustración Comparación de escalado entre SIFT (derecha) y SURF (izquierda). Fuente: (Valveny, 2021).

* 1. **Etapa 5: Reconocimiento.**
     1. **Clasificador Bayesiano.**

El clasificador bayesiano se fundamenta en el teorema de bayes, expresado en la ecuación 18, sobre el calculo de la probabilidad posterior de la clase.

Ecuación Teorema de Bayes.

Dónde C es la clase y el conjunto de atributos se indica con Entonces se puede escribir según la ecuación 19.

Ecuación Teorema de Bayes indicando el número de atributos.

Para aplicar la clasicación bayesiana, es necesaria las probabilidades; a priori dada por y la verosimilitud dada por . Este conjunto de datos se debe estimar para que el clasificador aprenda los parametros de clasificación. Además, se considera a como invariante para las distintas clases con el interes de maximizar la probabilidad de cada clase. Entonces el clasificador bayesiano se puede expresar como la ecuación 20. (Hernández Baez, 2016)

Ecuación Clasificador Bayesiano.

* + 1. **Máquina de soporte vectorial.**

La SVM (Support Vector Machines) es un clasificador supervisado capaz de constituir una frontera alrededor del dominio de los datos de aprendizaje. Están basadas en la idea de minimizar el riesgo estructural. La SVM ha demostrado gran desempeño en trabajar imágenes comparadas con las redes neuronales. (Betancourt, 2005)

Para generar la clasificación se empieza mapeando un conjunto de datos de entrenamiento representados por las muestras de la clase 1 y para la clase 2 formando un hiperplano que separe los margen entre clases.

|  |  |
| --- | --- |
| Ilustración 10 Clasificador lineal a partir de los vectores de soporte. Fuente: Propia | Ilustración 11 Clasificador no lineal a partir de los vectores de soporte. Fuente: Propia |

En la ilustración 10, el hiperplano solución dado por y los hiperplanos de soporte vectorial , resultantes de y respectivamente. Entre el hiperplano positivo y negativo se genera un margen cuya maximización se resuelve con optimización cuadratica mediante multiplicadores de Lagrange observado en la ecuación 22.

|  |  |
| --- | --- |
| Ecuación 21 Hiperplano solución lineal de SVM. | Ecuación 22 Hiperplano solución lineal maximizado con multiplicadores de Lagrange |

El clasificador no es linealmente separable cuando elementos de la clase 1 trasgreden el espacio de la clase 2 y viceversa como se evidencia en la ilustración 11. Para este caso se introducen variables , lo cual modifica el hiperplano de soporte vectorial por . También se intruduce un kernel que transforma el espacios obteniendo la funcion de decisión de la ecuación23. (Valveny, 2021)

Ecuación 23 Hiperplano solución no lineal maximizado con multiplicadores de Lagrange .

# Referencias

Aguirre Dobernack, N. (2013). Implementación de un sistema de detección de señales de tráfico mediante visión artificial basado en FPGA. Sevilla, España: Universidad de Sevilla.

Alegre Gutierrez, E. F. (2016). *SIFT (Scale Invariant Feature Transform).* Léon, España: Grupo de Visión del comité Español de Automática.

Alfaro Mejía, E. (2019). Detección y determinación de severidad de fallas en paneles solares a partir de imágenes áereas termográficas. Cali, Valle del Cauca, Colombia: Universidad del Valle.

Betancourt, G. (2005). *Las maquinas de soporte vectorial.* Pereira, Colombia: Universidad tecnológica de Pereira.

Cognex, C. (2016). Introdicción a la visión artificial. *Guía para la autimatización de procesos y mejoras de calidad*, 11-14.

EDMANS, G. d. (2006). *Ténicas y algoritmos básicos de visión artificial.* Londoño, España: Universidad la Rioja.

García Merino, J. (2016). *Sistema avanzado de detección de obstáculos y navegación autónoma para vigilancia y protección basado en flota de vehículos aéreos no tripulados.* Málaga, España: Universidad de Málaga.

Gómez Trejos, D., & Guerrero Guzmán, A. (2016). *Estudio y análisis de ténicas para procesamiento digital de imágenes.* pereira, Colombia.: Universidad tecnológica de pereira.

Hernández Baez, I. (2016). *Clasificador Bayesiano Ingenuo en RapidMiner.* Puebla, México: Benemérita Universidad Autónoma de Puebla.

Lara-Lévano, E. (2019). *Sistema de reconocimiento de gestos faciales captados a través de cámaras para analizar el nivel de satisfacción de clientes en restaurantes.* Lima, Perú: Universidad de Lima.

Roos Hoefgeest Toribio, S. (2017). *Reconocimiento de objetos mediante visión por computador para ayudas a invidentes.* Oviedo, España: Universidad de Oviedo.

Serrano, J. V. (2003). *Visión por computador.* Madrid: Dykinson.

Valveny, E. (2021). INTRODUCCIÓN A LA CLASIFICACIÓN DE IMÁGENES. *Clasificación de imágenes.* Barcelona, España.: Universitat Autònoma de Barcelona.

Valverde Rebaza, J. (2007). *Detección de bordes mediante el algoritmo de canny.* Trujillo, Perú: Universidad nacional de Trujillo.