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 6 etapas consecuentes que integran una tarea de visión artificial.

La etapa inicial es la adquisición de la imagen digital. Para lo cual, se emplean dispositivos de captura de imágenes con capacidad de digitalizar la señal producida y se implementan metodologías de adquisición. 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, 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 traducir los datos obtenidos apropiadamente para los sistemas computarizados y determinar si se usa la representación por frontera; centrando las características de la forma externa o por región completa; inclinando el objetivo hacia las propiedades internas. Para ello es preciso seleccionar un método para la extracción de los datos de interés. La selección de rasgos 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. El proceso brinda etiqueta y significado a un objeto mediante la información resultante de los descriptores. (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)

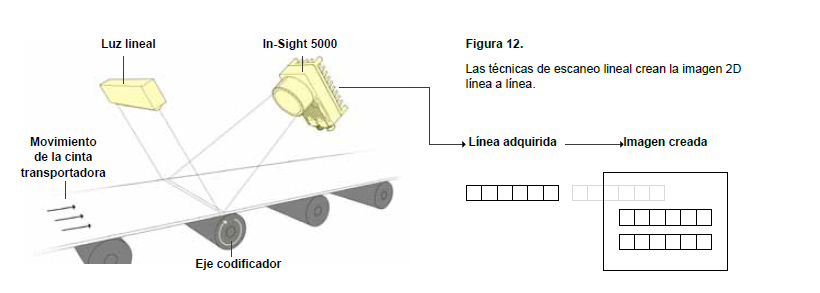


Ilustración Ejemplo de sistema integrado de visión artificial. Fuente: (Cognex, 2016)

* 1. **Etapa 2: Pre – Procesamiento.**
     1. Filtro de la laplaciana.

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 es:

Ecuación filtro laplaciano.

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

Tenemos que y que entonces: . Obtenemos que el operador laplaciano resalta en cambio de intensidad en la celda.

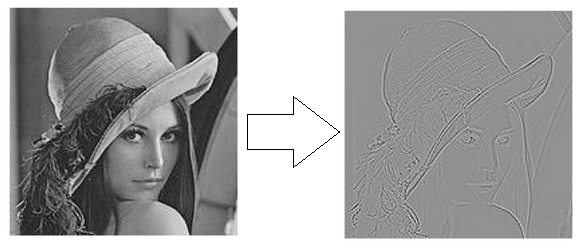


Ilustración En la izquierda, la foto origina. A la derecha, laplaciana de la imagen. Fuente: (Serrano, 2003)

* 1. **Etapa 3: Segmentación.**
  2. **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. (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)

Ecuación Ecuación de filtro Gaussiano.

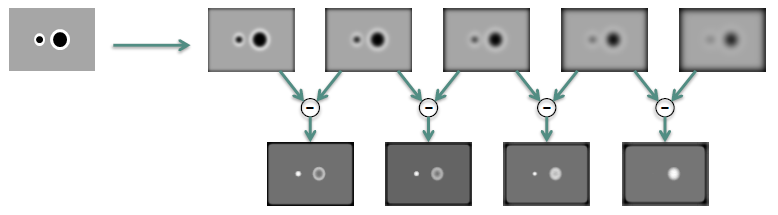


Ilustración Empleo de diferencia gaussiana para obtener puntos de interés. Fuente: (Valveny, 2021)

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)

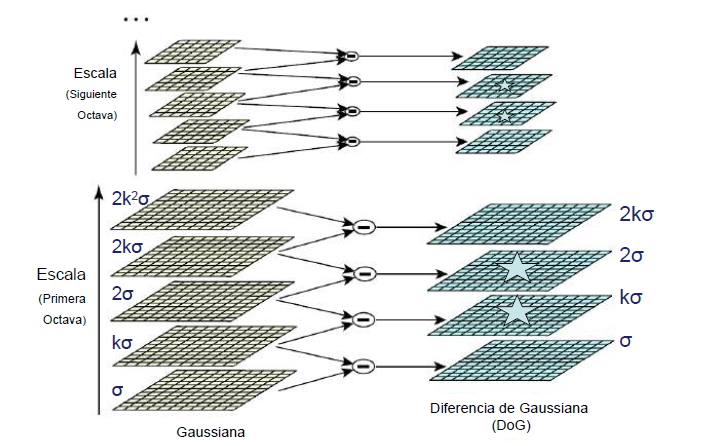


Ilustración Diagrama ilustrativo de la pirámide de las diferencias gaussianas compuestas de espacios - escalas. Fuente: (Alegre Gutierrez, 2016).

* + 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 2, 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:

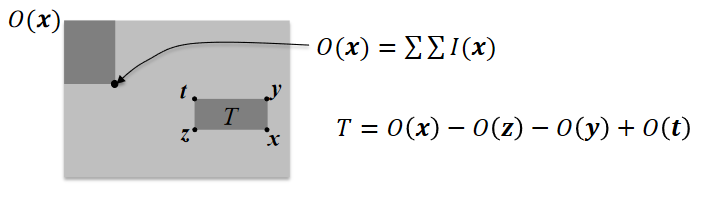


Ilustración Ejemplo de uso de imagen integral. Fuente: (Valveny, 2021)

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)

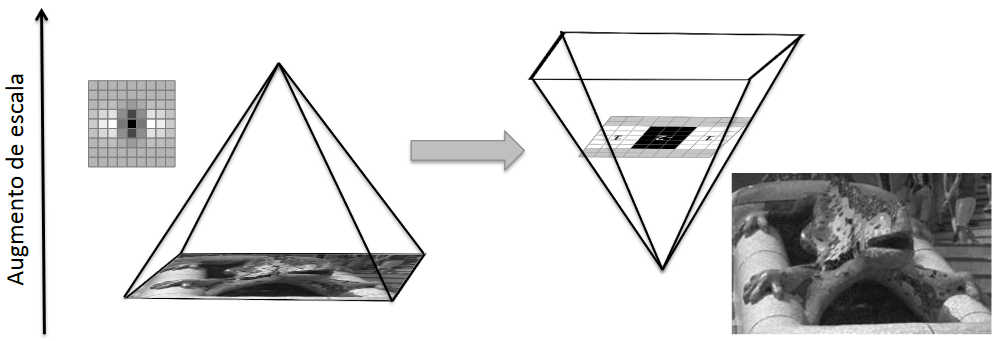


Ilustración Comparación de escalado entre SIFT (derecha) y SURF (izquierda). Fuente: (Valveny, 2021).

* 1. **Etapa 5: Reconocimiento.**
  2. **Conclusión.**