



UNIVERSIDAD DE COLIMA

FACULTAD DE INGENIERÍA ELECTROMECÁNICA

Ingeniería de software

REGIONES DE INTERÉS (KEYPOINTS)

Procesamiento de imágenes

Karla Karina Ramírez Márquez

4D

Puntos Clave de Entrenamiento Con Tamaño



Puntos Clave de Entrenamiento Sin Tamaño



Introducción:

Las zonas de interés, también conocidas como keypoints o puntos característicos, son pequeñas áreas de una imagen que se consideran relevantes para la identificación y descripción de esta. Estos puntos se caracterizan por tener una alta probabilidad de ser parte de un objeto o elemento importante en la escena.

En este reporte, se presenta una investigación individual sobre las técnicas para detectar zonas de interés en imágenes usando la biblioteca OpenCV. Se estudiaron las siguientes técnicas:

- **Si necesitas la máxima precisión y robustez:** Elige SIFT, aunque ten en cuenta su costo computacional.
- **Si necesitas velocidad y eficiencia:** Opta por SURF u ORB, considerando el balance entre velocidad y precisión necesario para tu tarea.
- **Si la velocidad es crítica y solo necesitas detectar esquinas:** FAST puede ser una buena opción.
- **Si necesitas descriptores extremadamente rápidos y compactos:** BRIEF puede ser útil, pero ten en cuenta su menor poder distintivo.

1. Scale Invariant Feature Transform (SIFT)

La Transformación Invariante a Escala (SIFT) es un algoritmo de visión artificial utilizado para detectar y describir puntos clave en imágenes. Fue desarrollado por David Lowe en 1999 y es conocido por su robustez a la escala, rotación, cambios de iluminación y parcialmente invariante a la distorsión afín. Esto convierte a SIFT en una herramienta poderosa para el reconocimiento de objetos, la comparación de imágenes y otras aplicaciones en visión artificial.

Aquí hay un resumen del algoritmo SIFT:

1. **Detección de extremos en el espacio-escala:** El algoritmo primero construye una representación del espacio-escala de la imagen. Esto se logra convolucionando la imagen con filtros gaussianos a diferentes escalas. Luego, identifica puntos clave potenciales al encontrar ubicaciones en la imagen que son extremos (máximos o mínimos) tanto en escala como en espacio.
2. **Localización de puntos clave:** Una vez que se identifican los puntos clave potenciales, se refina su ubicación y escala exactas. Esto implica ajustar un modelo a la información del gradiente de la imagen local alrededor del punto clave.
3. **Asignación de orientación:** SIFT asigna una orientación a cada punto clave. Esto es importante para lograr la invariancia a la rotación. La orientación se determina en función de las direcciones del gradiente de la imagen local alrededor del punto clave.
4. **Descriptor de punto clave:** Finalmente, SIFT calcula un descriptor para cada punto clave. Este descriptor captura la distribución de los gradientes locales alrededor del punto clave y se utiliza para emparejar puntos clave entre imágenes. El descriptor es un vector de 128 elementos, que representa la información del gradiente en diferentes posiciones y orientaciones alrededor del punto clave.

Ventajas de SIFT:

- **Robustez:** SIFT es altamente robusto a la escala, rotación, cambios de iluminación y parcialmente a distorsiones afines. Esto lo hace adecuado para aplicaciones del mundo real donde las imágenes pueden capturarse bajo diferentes condiciones.

- **Puntos clave distintivos:** SIFT detecta puntos clave distintivos que son informativos y confiables para la comparación.
- **Alta precisión:** SIFT puede lograr una alta precisión en tareas de reconocimiento de objetos y comparación de imágenes.

Desventajas de SIFT:

- **Costo computacional:** SIFT es computacionalmente costoso en comparación con otros detectores de puntos clave como FAST o ORB. Esto puede ser una limitación para aplicaciones en tiempo real.
- **No es perfecto:** Si bien es robusto, SIFT no es perfecto y puede verse engañado por cambios extremos de punto de vista o imágenes muy borrosas.

Aplicaciones de SIFT:

- **Reconocimiento de objetos:** SIFT se usa ampliamente para tareas de reconocimiento de objetos. Al emparejar puntos clave entre una imagen y una base de datos de objetos, SIFT puede identificar la presencia de objetos específicos en la imagen.
- **Comparación de imágenes:** SIFT se utiliza para emparejar puntos clave entre dos imágenes para encontrar puntos correspondientes. Esto puede ser útil para tareas como la composición de imágenes (crear una panorámica a partir de varias imágenes) o la estructura a partir del movimiento (reconstrucción 3D a partir de varias imágenes).
- **Visión robótica:** SIFT se puede utilizar en aplicaciones de visión robótica para ayudar a los robots a reconocer objetos en su entorno.

2. Speeded Up Robust Features (SURF)

Speeded Up Robust Features (SURF) es un algoritmo de visión artificial diseñado para la detección y descripción de puntos clave en imágenes. Se inspira en el algoritmo SIFT (Scale Invariant Feature Transform) pero ofrece una alternativa más rápida y computacionalmente eficiente.

Aquí te presento un resumen de SURF:

Funcionamiento:

1. **Detección de puntos clave en múltiples escalas:** Similar a SIFT, SURF construye una representación en el espacio-escala de la imagen. Utiliza la convolución con filtros gaussianos en diferentes escalas para lograr esto. En cada escala, se identifican posibles puntos clave como regiones con alto valor absoluto del determinante de la matriz Hessiana. La matriz

Hessiana mide la curvatura local de la intensidad de la imagen y ayuda a encontrar regiones con cambios significativos de intensidad.

2. **Refinamiento de la localización:** Una vez identificados los posibles puntos clave, se aplica un proceso de refinamiento para determinar su ubicación exacta. Esto involucra la interpolación para obtener una mejor precisión subpíxel.
3. **Asignación de orientación:** SURF asigna una orientación a cada punto clave. En lugar de analizar el gradiente de la imagen como en SIFT, SURF utiliza la distribución de la respuesta wavelet de Haar alrededor del punto clave para determinar la orientación dominante. Las respuestas wavelet de Haar capturan información local de la imagen y son eficientes de calcular.
4. **Descripción del punto clave:** Para describir la región alrededor de cada punto clave, SURF utiliza descriptores basados en la suma de respuestas wavelet de Haar. Se calculan las respuestas wavelet en una ventana alrededor del punto clave y en la orientación asignada. La suma de estas respuestas se organiza en un vector que forma el descriptor del punto clave.

Ventajas de SURF:

- **Velocidad:** Comparado con SIFT, SURF es significativamente más rápido, lo que lo hace adecuado para aplicaciones en tiempo real.
- **Robustez:** SURF mantiene un buen nivel de robustez a la escala, rotación, iluminación y distorsiones geométricas.
- **Eficiencia computacional:** SURF requiere menos recursos computacionales que SIFT, haciéndolo viable para dispositivos con menor potencia.

Desventajas de SURF:

- **Menos preciso:** En comparación con SIFT, SURF puede ser un poco menos preciso en la detección y descripción de puntos clave.
- **Sensible al ruido:** SURF puede ser más sensible al ruido en la imagen que SIFT.

Aplicaciones de SURF:

- **Reconocimiento de objetos:** SURF se puede utilizar para reconocer objetos en imágenes al comparar descriptores de puntos clave entre la imagen y una base de datos de objetos.

- **Rastreo de objetos:** SURF es útil para rastrear objetos en movimiento a través de secuencias de video al seguir el desplazamiento de puntos clave a lo largo del tiempo.
- **Panorama de imágenes:** SURF se puede utilizar para crear panoramas a partir de múltiples imágenes mediante la identificación de puntos clave coincidentes en las imágenes adyacentes.
- **Aplicaciones móviles:** Debido a su velocidad y eficiencia computacional, SURF es adecuado para aplicaciones de visión artificial en dispositivos móviles.

3. Features from Accelerated Segment Test (FAST)

FAST (por sus siglas en inglés *Features from Accelerated Segment Test*) es un algoritmo de visión artificial utilizado para la detección rápida de esquinas en imágenes. A diferencia de SIFT y SURF que buscan puntos clave generales, FAST se enfoca específicamente en identificar esquinas, las cuales son regiones de la imagen con cambio brusco de intensidad en dos direcciones. Su principal fortaleza es su velocidad computacional, lo que lo hace ideal para aplicaciones en tiempo real con limitaciones de recursos.

Aquí te presento un resumen del funcionamiento de FAST:

1. **Selección de píxel:** El algoritmo comienza por seleccionar un píxel de interés en la imagen.
2. **Círculo de prueba:** Alrededor del píxel seleccionado, se define un círculo con un radio preestablecido.
3. **Prueba de segmento:** FAST utiliza una serie de píxeles distribuidos estratégicamente sobre el círculo para evaluar rápidamente si el píxel central es una esquina. La prueba analiza la intensidad de estos píxeles en relación con la intensidad del píxel central.
4. **Criterio de esquina:** FAST define un conjunto de criterios basados en la intensidad relativa de los píxeles de prueba. Si una cantidad suficiente de píxeles de prueba cumplen estos criterios (por ejemplo, todos más brillantes o todos más oscuros que el píxel central con una cierta diferencia de intensidad), entonces se considera que el píxel central es una esquina.
5. **Reducción de candidatos:** La prueba de segmento inicial descarta rápidamente regiones sin esquinas. Si el píxel central no pasa la prueba inicial, no se realiza ningún análisis adicional.

6. **Verificación opcional:** Para mayor precisión, se puede incorporar una etapa de verificación adicional que analiza la intensidad de todos los píxeles en el círculo para confirmar la presencia de una esquina.

Ventajas de FAST:

- **Velocidad computacional:** FAST es uno de los detectores de esquinas más rápidos disponibles. Su diseño basado en la prueba de segmento permite un análisis rápido de posibles esquinas en la imagen.
- **Sencillez de implementación:** El algoritmo FAST es relativamente simple de entender e implementar, lo que lo hace atractivo para aplicaciones de desarrollo rápido.
- **Adecuado para tiempo real:** Debido a su velocidad, FAST es una buena opción para tareas de visión artificial que requieren procesamiento en tiempo real, como el rastreo de objetos en video.

Desventajas de FAST:

- **Menos preciso:** Comparado con otros detectores de puntos clave como SIFT y SURF, FAST puede ser menos preciso en la detección de esquinas. Puede omitir algunas esquinas reales o detectar falsas esquinas debido a ruido en la imagen.
- **Sensible al ruido:** La simplicidad del algoritmo lo hace más susceptible al ruido en la imagen, lo que puede generar falsos positivos.
- **Limitado a esquinas:** FAST solo detecta esquinas y no otros tipos de puntos clave que podrían ser relevantes para la tarea.

Aplicaciones de FAST:

- **Rastreo de objetos:** Debido a su velocidad, FAST se puede utilizar para rastrear objetos en movimiento en video al identificar y seguir las esquinas de los objetos a lo largo del tiempo.
- **Estimación de movimiento:** La detección de esquinas por FAST puede ser útil para estimar el movimiento de la cámara o de objetos en la escena.
- **Aplicaciones de realidad aumentada:** La velocidad de FAST lo hace adecuado para tareas de realidad aumentada que requieren procesamiento en tiempo real, como la superposición de objetos virtuales sobre imágenes del mundo real.

4. Binary Robust Independent Elementary Features (BRIF)

Binary Robust Independent Elementary Features (BRIF)

BRIF (por sus siglas en inglés *Binary Robust Independent Elementary Features*) es un algoritmo de visión artificial utilizado para describir puntos clave en imágenes. A diferencia de SIFT y SURF que generan descriptores ricos en información pero computacionalmente costosos, BRIF se enfoca en la creación de descriptores binarios cortos y eficientes.

Aquí te presento un resumen de BRIF:

Funcionamiento:

1. **Parche de imagen:** BRIF toma como entrada un parche de imagen centrado en el punto clave que se desea describir.
2. **** pares de píxeles:**** El algoritmo define un conjunto de pares de píxeles aleatorios dentro del parche de la imagen.
3. **Comparación de intensidades:** Para cada par de píxeles, se compara la intensidad del primer píxel con la del segundo. Se genera un bit en el descriptor final (0 si el primer píxel es más oscuro, 1 si es más claro).
4. **Independencia:** Los pares de píxeles se seleccionan de forma independiente para garantizar que el descriptor final no sea sensible a pequeños cambios en la posición del punto clave.
5. **Descriptor binario:** La concatenación de todos los bits generados a partir de las comparaciones de pares de píxeles forma el descriptor binario final. El tamaño del descriptor depende del número de pares de píxeles utilizados.

Ventajas de BRIF:

- **Velocidad computacional:** BRIF es un algoritmo extremadamente rápido para generar descriptores de puntos clave. Su simplicidad basada en comparaciones binarias permite un procesamiento muy eficiente.
- **Bajo consumo de memoria:** Los descriptores binarios de BRIF son compactos, lo que requiere menos memoria para almacenar y transmitir en comparación con descriptores más complejos.
- **Robustez:** A pesar de su simplicidad, BRIF puede ser sorprendentemente robusto a la escala, rotación y cambios moderados de iluminación.

Desventajas de BRIF:

- **Menos distintivo:** Los descriptores binarios cortos de BRIEF pueden ser menos distintivos que los descriptores ricos en información de algoritmos como SIFT y SURF. Esto puede afectar la precisión en tareas de comparación de descriptores.
- **Sensible a ruido:** BRIEF puede ser sensible al ruido en la imagen, ya que las comparaciones de intensidad binarias pueden verse afectadas por variaciones aleatorias en la intensidad de los píxeles.

Aplicaciones de BRIEF:

- **Aplicaciones en tiempo real:** Debido a su velocidad computacional, BRIEF es adecuado para tareas de visión artificial en tiempo real donde la eficiencia es crítica.
- **Combinación con otros detectores:** BRIEF se suele combinar con detectores de puntos clave más precisos como FAST para aprovechar la velocidad de BRIEF en la generación de descriptores.
- **Aplicaciones móviles:** El bajo consumo de memoria y la velocidad de BRIEF lo hacen viable para su uso en dispositivos móviles con recursos limitados.

5. Oriented FAST and Rotated BRIEF (ORB)

ORB (por sus siglas en inglés *Oriented FAST and Rotated BRIEF*) es un algoritmo de visión artificial que combina las ventajas de dos algoritmos populares: FAST y BRIEF. Se utiliza para detectar y describir puntos clave en imágenes de manera rápida, eficiente y robusta a cambios como escala, rotación e iluminación.

Funcionamiento:

1. **Detección de puntos clave con FAST:** ORB utiliza el algoritmo FAST para detectar posibles puntos clave en la imagen. FAST es rápido y eficiente en la identificación de esquinas, que son regiones de la imagen con cambio brusco de intensidad.
2. **Asignación de orientación:** A diferencia de FAST, que no proporciona orientación, ORB calcula la orientación para cada punto clave. Esto se logra analizando el gradiente de la intensidad de la imagen en un entorno del punto clave.
3. **Descripción con BRIEF:** Para describir cada punto clave, ORB emplea el algoritmo BRIEF. BRIEF genera descriptores binarios cortos y eficientes comparando la intensidad de pares de píxeles aleatorios dentro de un parche de imagen centrado en el punto clave.

4. **Adaptación a la orientación:** Un aspecto clave de ORB es que adapta el patrón de pares de píxeles de BRIEF a la orientación asignada en el paso anterior. Esto garantiza que el descriptor capture información relevante independientemente de la rotación del punto clave en la imagen.

Ventajas de ORB:

- **Velocidad computacional:** Al combinar FAST y BRIEF, ORB hereda la rapidez de ambos algoritmos, siendo significativamente más rápido que SIFT y SURF.
- **Robustez:** ORB mantiene un buen nivel de robustez a escala, rotación, iluminación y algunas distorsiones geométricas.
- **Eficiencia:** ORB es computacionalmente eficiente y tiene un bajo consumo de memoria, gracias a los descriptores binarios cortos de BRIEF.

Desventajas de ORB:

- **Menos distintivo que SIFT:** Si bien es robusto, ORB puede ser un poco menos distintivo que SIFT en la descripción de puntos clave. Esto puede afectar la precisión en tareas de comparación de descriptores muy complejas.
- **Sensible al ruido:** Al igual que BRIEF, ORB puede ser sensible al ruido en la imagen, ya que las comparaciones de intensidad binarias pueden verse afectadas por variaciones aleatorias.

Aplicaciones de ORB:

- **Aplicaciones en tiempo real:** Debido a su velocidad, ORB es adecuado para tareas de visión artificial en tiempo real como el rastreo de objetos en video o la realidad aumentada.
- **Reconocimiento de objetos:** ORB se puede utilizar para reconocer objetos en imágenes al comparar descriptores de puntos clave entre la imagen y una base de datos de objetos.
- **Robótica:** La eficiencia y robustez de ORB lo hacen viable para aplicaciones de robótica que requieren detección y descripción de puntos clave en tiempo real.

Conclusiones:

Las técnicas para detectar zonas de interés en imágenes son una parte fundamental del procesamiento de imágenes y la visión artificial. Cada técnica tiene sus propias ventajas y desventajas, y la elección de la técnica adecuada depende de la aplicación específica.

En general, SIFT y SURF son buenas opciones para aplicaciones donde la precisión es importante, mientras que FAST, BRIEF y ORB son buenas opciones para aplicaciones donde la velocidad es crítica.

Fuentes:

- OpenCV documentation: <https://docs.opencv.org/4.x/index.html>
- SIFT paper: <https://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf>
- SURF paper: <https://www.instagram.com/paper.surf/>
- FAST paper: <https://www.linkedin.com/company/fast-paper>
- BRIEF paper: <https://rasmussen.libanswers.com/faq/313397>
- ORB paper: <https://ieeexplore.ieee.org/document/6126544>