



UNIVERSIDAD NACIONAL DE INGENIERÍA
Facultad de Ingeniería Industrial y de Sistemas

“KAZE Features”

Maestría en Inteligencia Artificial

Tarea 2 – Computer Vision

Fonseca Rodríguez, Christian

2025-I

Introducción

KAZE es un algoritmo de detección y descripción de características locales en imágenes que introdujo por primera vez el uso de espacios de escala no lineales basados en difusión para preservar los contornos naturales de los objetos. A diferencia de los métodos que emplean el espacio de escala gaussiano (SIFT, SURF), KAZE aplica un filtro de difusión no lineal para realizar un suavizado adaptativo, reduciendo el ruido sin atenuar excesivamente los detalles importantes de la imagen. Esto se traduce en una mayor precisión de localización y distintividad de los puntos clave [1].

Filtro de difusión y Perona–Malik

El corazón de KAZE es la PDE de Perona–Malik para suavizado adaptativo.

$$\frac{\partial L}{\partial t} = \operatorname{div}(c(x, y, t) \nabla L)$$

donde

- $L(x, y, t)$ es la imagen suavizada al "tiempo" (escala) t ,
- ∇L su gradiente espacial,
- div la divergencia,
- y $c(x, y, t)$ la **conductancia** que modula el flujo de difusión según el contraste local.

Esta formulación atenúa el ruido sin difuminar excesivamente los bordes .

Filtro de difusión y Perona–Malik

Perona y Malik proponen que la conductancia c dependa del gradiente local de una versión suavizada de la imagen L . Definen primero:

$$c(x, y, t) = g(|\nabla L_\sigma(x, y, t)|),$$

donde ∇L_σ es el gradiente de L pre-suavizada con un gaussiano de desviación σ .

Dos funciones clásicas de conductancia son:

$$g_1(s) = \exp\left(-\frac{s^2}{k^2}\right), \quad g_2(s) = \frac{1}{1 + \frac{s^2}{k^2}},$$

con $s = |\nabla L_\sigma|$ y k el parámetro de contraste que regula la sensibilidad al borde .

Filtro de difusión y Perona–Malik

Con $k=5.10$ la función decae tan rápido que incluso los gradientes más pequeños se toman como bordes, reteniendo ruido y texturas finas; al aumentar a $k=12.75$ se difuminan detalles muy débiles, conservando aún muchos contornos; con $k=38.25$ solo los contornos anatómicos más marcados impiden la difusión; y con $k=89.25$ la conductancia varía tan lentamente que casi toda la imagen se suaviza, dejando únicamente los bordes de contraste extremo.

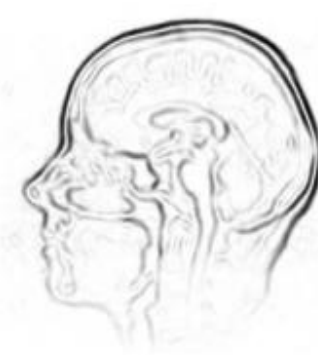
- Un k pequeño = sensibilidad alta: retienes hasta el más mínimo detalle, pero corres el riesgo de amplificar ruido.
- Un k grande = sensibilidad baja: sólo preservas contornos muy marcados, eliminando detalles finos y ruido.



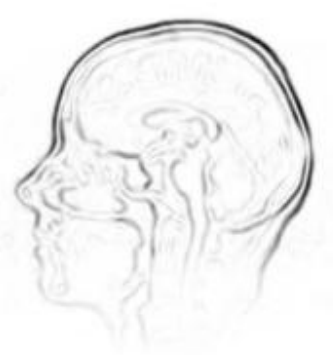
$k = 5.10$



$k = 12.75$



$k = 38.25$



$k = 89.25$

Filtro de difusión y Perona–Malik

Para una difusión más selectiva, Weickert definió

$$g_3(s) = \begin{cases} 1, & s^2 = 0, \\ 1 - \exp(-3.315 / (\frac{s}{k})^8), & s^2 > 0, \end{cases}$$

que promueve un suavizado intrarregional más fuerte que el interregional .

Integramos la ecuación de difusión no lineal

$$\frac{\partial L}{\partial t} = \operatorname{div}(c(x, y, t) \nabla L)$$

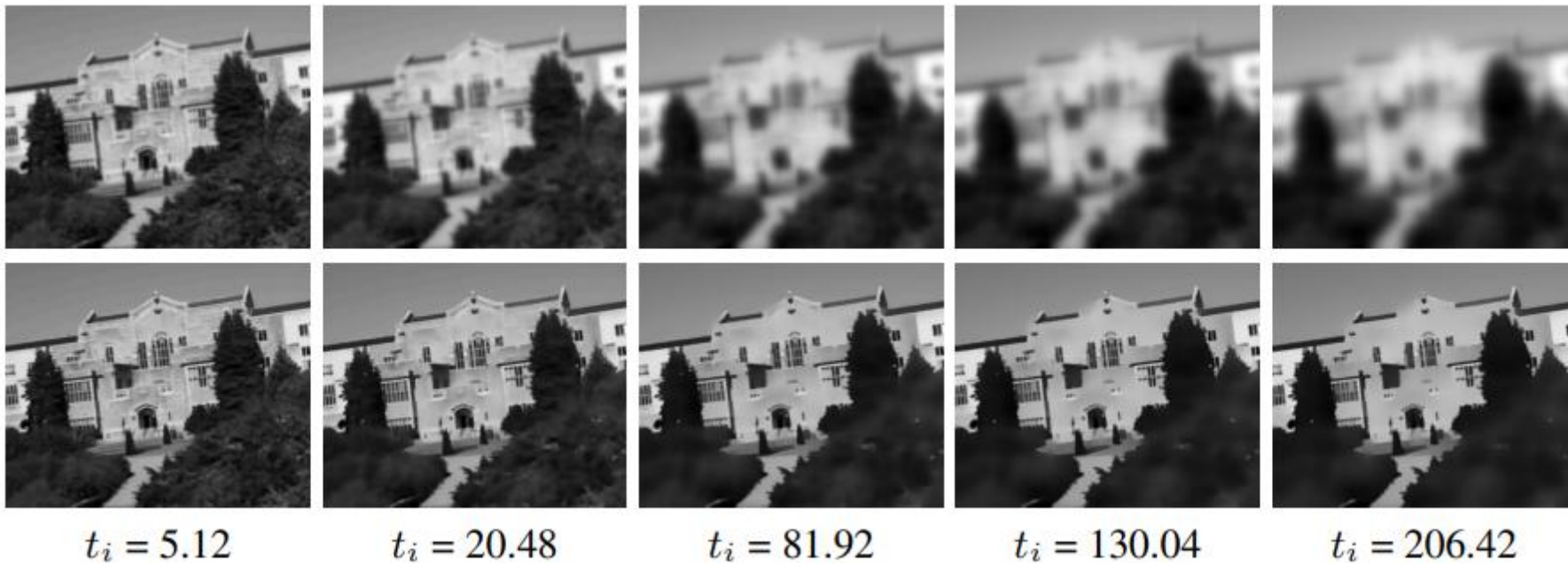
para varios valores de t , generando así el espacio de escala no lineal, donde el suavizado es adaptativo según el contenido local.

Construcción del espacio de escala no lineal

El espacio de escala no lineal en KAZE se construye resolviendo ecuaciones en derivadas parciales (PDE) de difusión con conductancia variable. Se emplea la técnica de Additive Operator Splitting (AOS) para garantizar estabilidad y eficiencia frente a tamaños de paso arbitrarios. Cada paso de difusión requiere resolver un sistema tridiagonal, que se realiza eficientemente mediante el algoritmo de Thomas. Este proceso crea un conjunto de imágenes suavizadas donde el grado de difuminado se adapta localmente a la presencia de bordes y texturas.

Construcción del espacio de escala no lineal

Aquí vemos la comparación entre el espacio de escala lineal y el no lineal con la función de conductividad g_3 para distintos tiempos de evolución t_i (cada t_i equivale a la varianza σ_i^2 en el caso gaussiano y al tiempo de integración de la ecuación de difusión en el no lineal). En la primera fila, la difusión gaussiana difumina la imagen de forma uniforme al crecer t_i , perdiendo bordes y detalles. En la segunda fila, la difusión no lineal suaviza regiones homogéneas pero conserva los contornos anatómicos fuertes, manteniendo los límites nítidos incluso para valores altos de t_i .



Detección de Keypoints

Una vez construido el espacio de escala no lineal, KAZE detecta keypoints buscando máximos locales de la respuesta de la escala-normalizada del Hessiano (o aproximaciones eficientes del mismo) a través de todas las escalas. Esto asegura invarianza ante cambios de escala y ofrece una alta repetibilidad incluso en condiciones de transformaciones geométricas y fotométricas. La combinación de la difusión no lineal y la detección basada en el Hessiano mejora la precisión de la localización de los puntos de interés.

$$R(x, y, t) = t^2 \det(\mathbf{H}L(x, y, t)) = t^2 (L_{xx}(x, y, t) L_{yy}(x, y, t) - L_{xy}(x, y, t)^2)$$

donde

- $L(x, y, t)$ es la imagen suavizada en la escala t ,
- L_{xx}, L_{yy}, L_{xy} son las derivadas segundas espaciales de L ,
- y el factor t^2 normaliza la respuesta para que sea invariante a cambios de escala.

Cálculo de descriptores

KAZE extrae un vector de características muestreando gradientes en un parche circular adaptado a la escala y orientación de cada punto clave:

1. Se centra un parche con radio proporcional a la escala t ; si `upright=False`, se rota según la orientación dominante.
2. El parche se divide en 4×4 subregiones; en cada una se registra un histograma de gradientes con 8 direcciones.
3. La concatenación de los 16 histogramas da un vector de 128 dimensiones (extended) o 64 dimensiones (standard).
4. Ese vector se normaliza (norma L2), se aplica un recorte suave y se vuelve a normalizar, lo que refuerza la invariancia a contraste y brillo.

De este modo se obtiene un descriptor robusto a escala, rotación y variaciones fotométricas.

Invarianzas

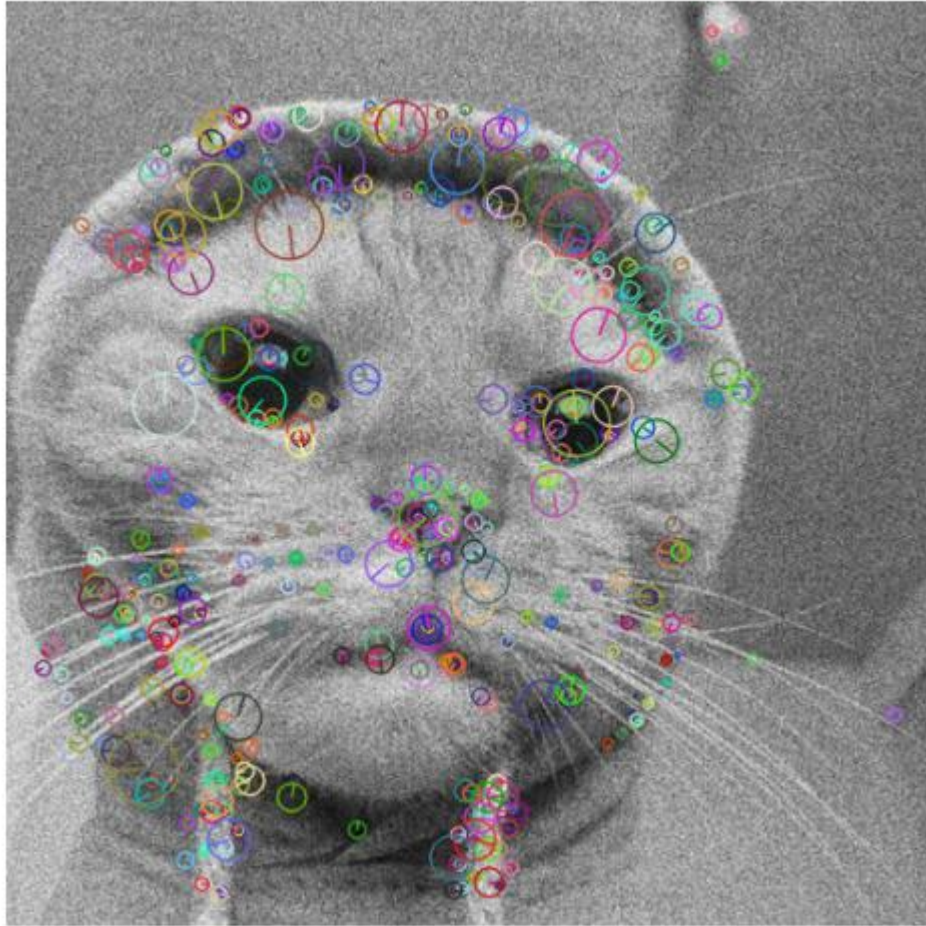
- **Escala:** Gracias al espacio de escala no lineal y a la normalización, KAZE detecta el mismo punto clave independientemente de cuánto se amplíe o reduzca la imagen.
- **Rotación:** Cuando se estima la orientación dominante de cada keypoint (`upright=False`), el descriptor es totalmente invariante a giros. Si se activa `upright=True`, pierdes esa invariancia a cambio de velocidad.
- **Iluminación y Fotométrica:** Al basarse en gradientes locales y no en valores absolutos de intensidad, KAZE es muy robusto frente a cambios de brillo, contraste y cierto nivel de ruido.
- **Afín:** Aunque no está diseñado explícitamente para invariancia afín, su combinación de detección multiescala y descripción orientada le confiere cierta tolerancia a estiramientos y compresiones geométricas leves.
- **Perspectiva y Visión de Punto de Vista:** KAZE no corrige transformaciones de perspectiva pronunciadas (cambios fuertes de ángulo de cámara), aunque pequeños cambios de punto de vista sí suelen conservar la repetibilidad de los keypoints.

Usos

- Cuando se necesite máxima precisión en la localización de puntos clave y preservación de bordes. Su suavizado adaptativo mantiene los contornos nítidos incluso en presencia de ruido o texturas finas, algo que SIFT/SURF (filtro gaussiano) o BRISK/ORB (binarios) pierden.
- Al trabajar con escenas muy texturizadas o con bajo contraste. La difusión no lineal refuerza las estructuras relevantes y descarta el ruido, mejorando la repetibilidad en imágenes médicas, satelitales o industriales.
- Se requiera invarianza completa a escala y a rotación (opcional). El Hessiano escala-normalizado garantiza detección estable al cambiar el zoom, y si se usa `upright=False` además estimas orientación para cubrir giros.
- Se pueda sacrificar algo de velocidad por calidad. KAZE es más lento que ORB o FAST+BRIEF, pero A-KAZE (su variante acelerada) reduce el tiempo sin perder gran parte de la precisión.

Ejemplos

Puntos clave detectados con KAZE



Puntos clave detectados con KAZE

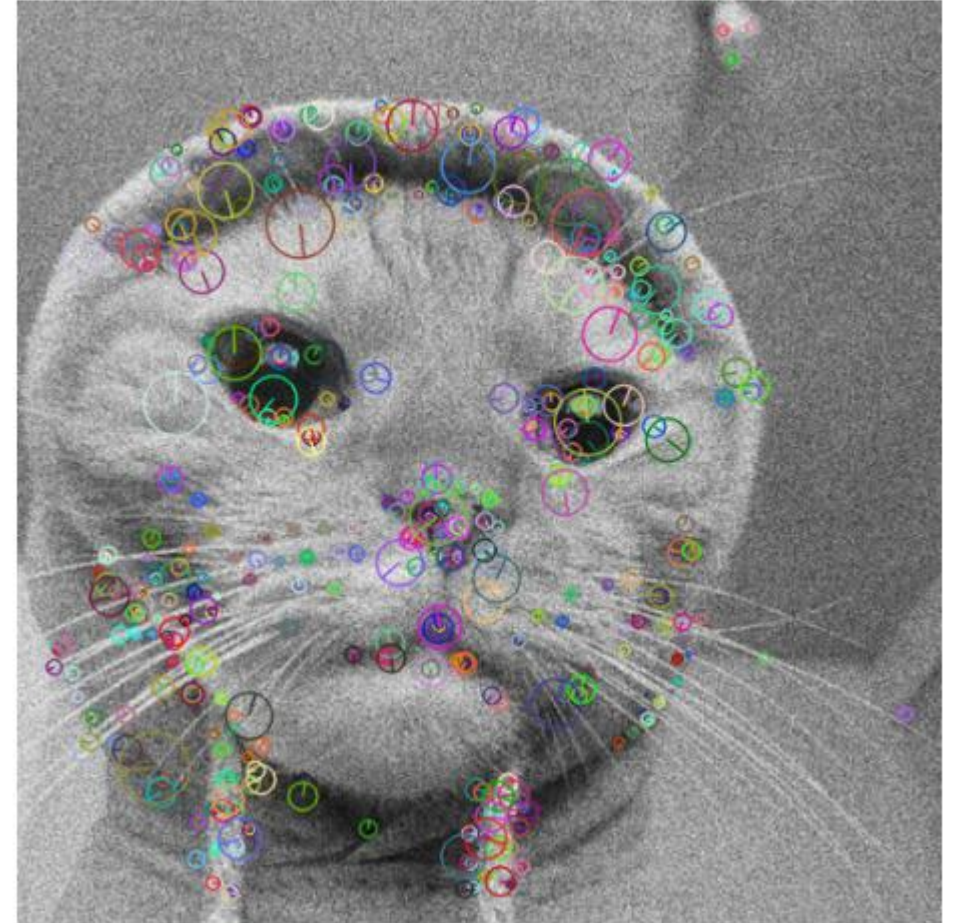


Descripción

Ubicación: KAZE escanea todo el espacio de escala no lineal buscando máximos locales de la respuesta del Hessiano. Esos máximos corresponden a puntos donde la intensidad cambia bruscamente en más de una dirección (esquinas, cruces de bordes, texturas pronunciadas). Por eso hay muchos puntos alrededor de los ojos, el hocico y los bigotes: son zonas de alto contraste y estructura compleja.

Tamaño (escala): Cada círculo tiene un radio proporcional a la escala t en la que se detectó el punto clave. Los puntos más “grandes” corresponden a estructuras detectadas en escalas gruesas (objetos de mayor tamaño), mientras que los más pequeños capturan detalles finos.

Puntos clave detectados con KAZE



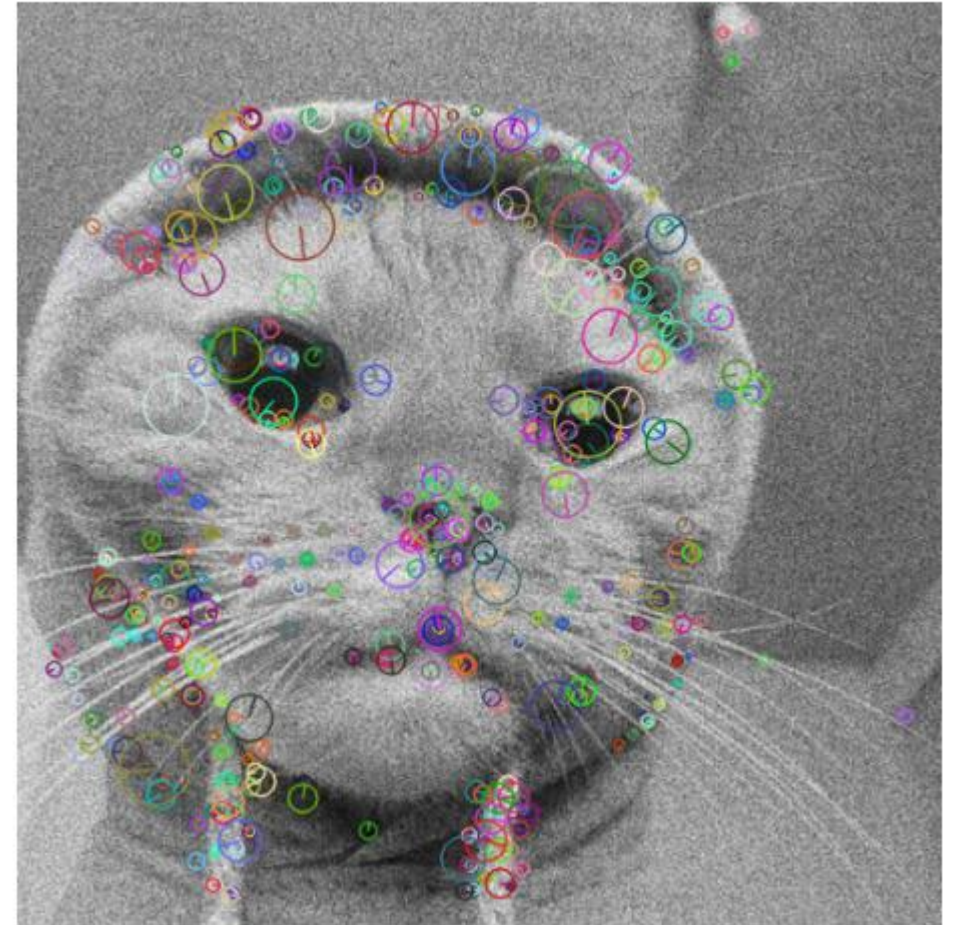
Descripción

Orientación: La línea que aparece dentro de cada círculo indica la orientación dominante del gradiente local, calculada para hacer el descriptor invariante a rotaciones (a menos que se use `upright=True`, que omitiría este paso).

Descriptor: Una vez fijada la localización, escala y orientación, KAZE extrae un vector de características (descriptor) muestreando gradientes en una vecindad adaptada al punto. Ese descriptor es el “perfil” de cómo varía la intensidad alrededor del keypoint, y sirve para compararlo con puntos de otra imagen.

Colores: OpenCV asigna colores aleatorios a cada keypoint solo para que visualmente se puedan distinguir.

Puntos clave detectados con KAZE



BIBLIOGRAFÍA

[1] Alcantarilla, P. F., Bartoli, A., & Davison, A. J. (2012). KAZE features. In Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part VI 12 (pp. 214-227). Springer Berlin Heidelberg.