

Jornadas de Automática

Robustez visual frente a degradaciones para exploración planetaria

Laura Castro-Lara^a, David Rodríguez-Martínez^{a,*}, Carlos Pérez-del-Pulgar^a

^aLaboratorio de Robótica Espacial, Instituto Universitario de Investigación en Ingeniería Mecatrónica y Sistemas Ciberfísicos, Universidad de Málaga.

To cite this article: Castro-Lara, L., Rodríguez-Martínez, D., Pérez-del-Pulgar, C.. 2026. Visual Robustness under Image Degradations for Planetary Exploration. Jornadas de Automática, 47.
<https://doi.org/10.17979/ja-cea.2026.47.13804>

Resumen

La exploración robótica espacial requiere sistemas de percepción visual fiables en entornos donde las condiciones de adquisición pueden degradar severamente la información visual disponible. Este trabajo presenta una evaluación sistemática de extractores de características visuales, clásicos y basados en modelos de aprendizaje, bajo tres degradaciones representativas de escenarios de exploración planetaria con sensado restringido: reducción de resolución espacial, cuantización radiométrica y degradación fotoelectrónica asociada a baja iluminación. La evaluación se realiza sobre secuencias de navegación en entornos planetarios y mediante métricas que incluyen la consistencia geométrica, robustez descriptiva y coste computacional. Los resultados muestran que no existe un extractor universalmente robusto, ya que el rendimiento depende estrechamente del tipo de degradación: la reducción espacial limita las correspondencias útiles, la cuantización conserva parcialmente la geometría pero reduce la discriminabilidad, y la baja iluminación afecta a la estabilidad de las correspondencias. El estudio proporciona criterios para seleccionar extractores visuales en misiones robóticas de exploración planetaria con restricciones de sensado.

Palabras clave: Percepción y sensado, robótica aérea y espacial, navegación, programación y visión de robots.

Visual Robustness under Image Degradations for Planetary Exploration

Abstract

Robotic space exploration demands reliable visual perception systems capable of operating in environments where acquisition conditions can severely degrade the available visual information. This work presents a systematic evaluation of visual feature extractors — both classical and learning-based — under three degradations representative of planetary exploration scenarios with constrained sensing: spatial resolution reduction, radiometric quantization, and photoelectronic degradation associated with low illumination. The evaluation is conducted on navigation sequences in planetary environments using metrics encompassing geometric consistency, descriptive robustness, and computational cost. Results show that no universally robust extractor exists, as performance is closely tied to the type of degradation: spatial reduction limits the number of useful correspondences, radiometric quantization partially preserves geometry while reducing discriminability, and low illumination undermines correspondence stability. This study provides actionable criteria for selecting visual feature extractors in robotic planetary exploration missions operating under sensing constraints.

Keywords: Perception and sensing, aerial and space robotics, robot navigation, programming and vision

1. Introducción

La percepción visual es esencial para la autonomía de robots en misiones de exploración espacial, donde tareas como la odometría, la localización, el mapeo y la reconstrucción

ambiental dependen de correspondencias fiables entre imágenes (Scaramuzza and Fraundorfer, 2011). En estos sistemas, los detectores y descriptores locales de características siguen desempeñando un papel fundamental, ya que proporcionan

*Autor para correspondencia: david.rm@uma.es

puntos de referencia visuales sobre los que estimar relaciones geométricas entre vistas.

No obstante, las imágenes adquiridas en escenarios planetarios pueden diferir significativamente de las condiciones RGB convencionales comúnmente utilizadas para el desarrollo de los modelos de detección y descripción de características. La iluminación puede ser limitada, las sombras y los cambios de contraste pueden ser extremos, la textura puede ser escasa y las cámaras usadas para navegación suelen estar sujetas a restricciones de resolución, precisión radiométrica, energía, cómputo o ancho de banda (Gerdes et al., 2024; Meyer et al., 2021; Rodríguez-Martínez et al., 2026; Vayugundla et al., 2018). Bajo estas condiciones, no está claro qué extractores mantienen correspondencias útiles ni cómo se degradan sus mecanismos de detección y descripción en base a estas limitaciones y las características propias del entorno de exploración.

Este problema es especialmente relevante porque los extractores existentes responden de forma diferente ante la pérdida de información visual. Los métodos clásicos, como SIFT (Lowe, 2004), ORB (Rublee et al., 2011) y AKAZE (Fernández Alcantarilla, 2011), dependen de estructuras locales bien definidas y estados de iluminación constantes, mientras que los métodos basados en aprendizaje, como SuperPoint (DeTone et al., 2018), R2D2 (Revaud et al., 2019) y D2-Net (Dusmanu et al., 2019), incorporan representaciones aprendidas que pueden aportar mayor robustez, pero que, al no haber sido entrenados específicamente bajo estas condiciones de degradación, presentan un comportamiento menos predecible.

Este trabajo evalúa de forma sistemática extractores clásicos y aprendidos bajo tres regímenes controlados de degradación: reducción de resolución espacial, cuantización radiométrica y degradación fotoelectrónica en baja iluminación. El objetivo es determinar qué métodos siguen siendo útiles en condiciones de sensado restringido y analizar cómo cada degradación afecta a la cantidad, estabilidad y consistencia geométrica de las correspondencias.

2. Estado del arte

Los métodos de características locales se han utilizado ampliamente en visión por computador y robótica para establecer correspondencias entre imágenes. Los enfoques clásicos, como SIFT (Lowe, 2004), SURF (Bay et al., 2006), ORB (Rublee et al., 2011), BRIEF (Calonder et al., 2010), BRISK (Leutenegger et al., 2011), KAZE (Fernández Alcantarilla et al., 2012) y AKAZE (Fernández Alcantarilla, 2011), siguen normalmente un esquema de detección y descripción basado en estructuras locales como esquinas, regiones de contraste o gradientes, ofreciendo distintos compromisos entre robustez, distintividad y eficiencia computacional.

En los últimos años, los métodos basados en aprendizaje han surgido como alternativa a los extractores diseñados manualmente. Modelos como SuperPoint (DeTone et al., 2018), D2-Net (Dusmanu et al., 2019) y R2D2 (Revaud et al., 2019) aprenden una o varias etapas del proceso de extracción para mejorar la repetibilidad y la fiabilidad de las correspondencias.

La robustez de estos métodos se ha evaluado habitualmente bajo cambios geométricos y fotométricos, incluyendo variaciones de escala, rotación, punto de vista, iluminación o desenfoque (Tareen and Saleem, 2018; Karami et al., 2015; Khan Tareen and Raza, 2023; Salau and Jain, 2019). En el contexto de SLAM visual, Gil et al. (Gil et al., 2010) analizaron comparativamente distintos detectores de puntos de interés y descriptores locales, prestando especial atención a la repetibilidad de los puntos y a la estabilidad de los descriptores ante cambios de punto de vista. Este tipo de variaciones es inherente a la navegación robótica, ya que el robot observa la misma escena desde posiciones y orientaciones diferentes durante su desplazamiento; en exploración planetaria, a esta variabilidad se suman además las restricciones radiométricas propias del entorno y del sensor.

Protocolos como HPatches (Balntas et al., 2017) han permitido comparar extractores y descriptores de forma controlada, pero muchas evaluaciones se basan en imágenes bien formadas y con suficiente información visual. En robótica espacial, la adquisición de imágenes puede diferir de estas condiciones. Así, un método que funciona bien con imágenes convencionales puede generar correspondencias inestables, poco discriminativas o geoméricamente inconsistentes en escenarios de operación más restrictivos.

Por tanto, la selección de extractores para estos escenarios requiere analizar el compromiso entre disponibilidad de correspondencias, consistencia geométrica, robustez descriptiva y eficiencia computacional bajo restricciones radiométricas reales.

3. Metodología

La evaluación analiza cómo distintas degradaciones visuales afectan al comportamiento de extractores de características locales. Para ello, se estudia el rendimiento de cada método sobre pares de imágenes extraídas de una serie de conjuntos de datos representativos de entornos planetarios.

3.1. Datasets

La evaluación utiliza secuencias de imágenes RGB de navegación robótica en entornos exteriores y similares a los que se encontrarían en planetas áridos. Se seleccionan trayectorias representativas de cuatro conjuntos públicos: *BASE-PROD* (Gerdes et al., 2024), *MADMAX* (Meyer et al., 2021), *SPICE-HL3* (Rodríguez-Martínez et al., 2026) y *ETNA* (Vayugundla et al., 2018). En conjunto, estas secuencias cubren terrenos semiáridos, análogos lunares, análogos marcianos y escenarios volcánicos, con diferentes condiciones de textura, contraste e iluminación, tal como se detalla en la Tabla 1.

3.2. Procedimiento de evaluación

El procedimiento de evaluación, mostrado en la Figura 1, consta de cuatro etapas: generación de imágenes degradadas, extracción de características, emparejamiento con verificación geométrica mediante RANSAC (Fischler and Bolles, 1981) y cálculo de métricas. Para aislar el efecto de cada degradación, todos los extractores se ejecutan con la misma configuración en todos los niveles evaluados.

Tabla 1: Secuencias de navegación utilizadas en la evaluación.

Seq.	Dataset	Resolución	Frames	Sensor	Propiedades
Secuencia 1	BASEPROD	640x360	500	Realsense D435i RGB-D	Ilum. balanceada, textura moderada, giros abruptos
Secuencia 2	SPICE-HL3	640x360	130	StereoLabs ZED2	Sombras duras, textura rica, trayectoria recta
Secuencia 3	MADMAX	722x540	300	AlliedVision Mako G-319	Ilum. irregular, textura escasa, curvas moderadas
Secuencia 4	MADMAX	722x540	250	AlliedVision Mako G-319	Sombras, textura escasa, giros abruptos y rápidos
Secuencia 5	SPICE-HL3	640x360	600	StereoLabs ZED2	Baja ilum., textura rica, curvas moderadas
Secuencia 6	ETNA	1021x762	500	AlliedVision Mako G-125	Ilum. balanceada, textura moderada, curvas moderadas

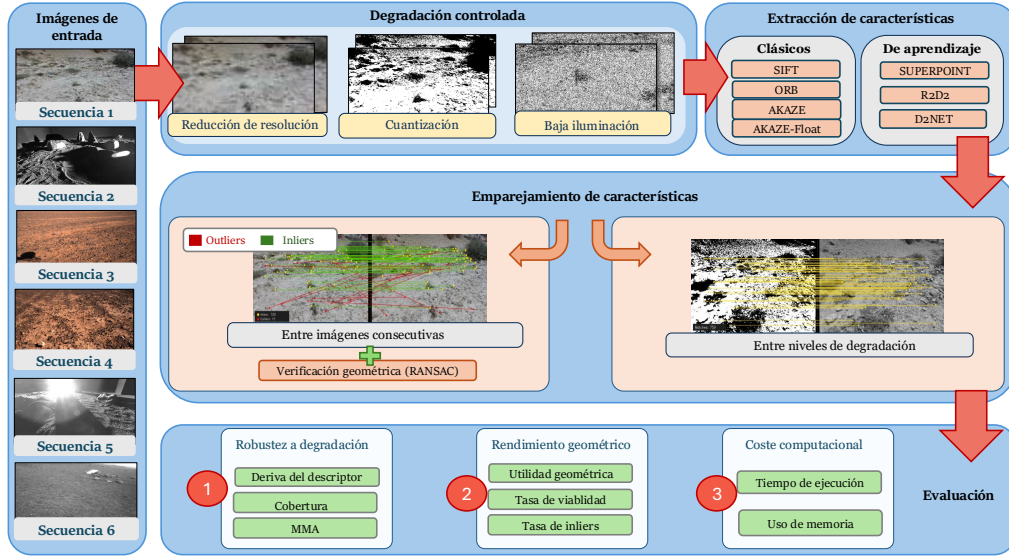


Figura 1: Procedimiento unificado de evaluación para el análisis de robustez a nivel de características, aplicado a extractores evaluados.

Se emplean dos esquemas de emparejamiento complementarios: el *emparejamiento temporal*, entre fotogramas consecutivos I_t e I_{t+1} , para evaluar la utilidad geométrica de las correspondencias; y el *emparejamiento entre representaciones*, entre una imagen de referencia I_t y su versión degradada $\tilde{I}_t$, para medir repetibilidad, estabilidad descriptiva y precisión espacial bajo degradación.

3.3. Degradaciones controladas

- **Reducción de resolución espacial.** La imagen se submuestra progresivamente para simular una pérdida de detalle espacial. Se evalúan cuatro niveles de escala (s):

$$s \in \{1, 1/2, 1/4, 1/8\},$$

- **Cuantización radiométrica.** Las intensidades se cuantizan uniformemente a un número reducido de bits por pixel. Se evalúan cuatro niveles de cuantización (b):

$$b \in \{8, 4, 2, 1\},$$

- **Degradación fotoelectrónica de baja iluminación.** Se introduce un modelo simplificado de adquisición bajo baja señal. La imagen normalizada $I \in [0, 1]$ se convierte en un número esperado relativo de fotones mediante:

$$\lambda = \gamma I,$$

donde γ controla el presupuesto fotónico relativo de la simulación. El conteo observado k , es decir, el número

de fotones detectados por píxel, se simula como:

$$k \sim \text{Poisson}(\lambda).$$

Posteriormente se añaden corriente oscura, ruido de patrón fijo (FPN), ruido de lectura y cuantización ADC.

Los niveles empleados para la degradación fotoelectrónica de baja iluminación se muestran en la Tabla 2.

3.4. Métricas

3.4.1. Métricas geométricas

Las métricas geométricas se calculan sobre pares consecutivos de imágenes:

- **Tasa de inliers, R_{in} :** fracción de emparejamientos aceptados como geoméricamente consistentes tras RANSAC.
- **Tasa de viabilidad, V_r :** fracción de pares de imágenes que superan un umbral mínimo de inliers.
- **Utilidad geométrica, G :** producto del número de inliers por su dispersión espacial, aproximada como $\sqrt{\sigma_x^2 + \sigma_y^2}$.

Tabla 2: Niveles de la degradación fotoelectrónica de baja iluminación. El parámetro γ indica el presupuesto fotónico relativo

Nivel	γ	Ruido de lectura	FPN	Corriente oscura	ADC
limpio	1000	0	0	0	8 bits
leve	500	0,0005	0,005	0,1	8 bits
moderado	150	0,001	0,010	0,5	4 bits
severo	10	0,003	0,020	1,0	1 bit

3.4.2. Métricas de robustez frente a degradación

Las métricas de robustez se calculan comparando una imagen original I_t con su versión degradada $\tilde{I}_t$:

- **Cobertura, C_d** : fracción de puntos de la imagen de referencia recuperados bajo degradación.
- **Deriva del descriptor, D_d** : distancia media ℓ_2 entre descriptores correspondientes.
- **Precisión media de emparejamiento, $MMA@t$** : fracción de emparejamientos cuyo error de localización es inferior a un umbral t píxeles.

3.4.3. Métricas computacionales

- **Tiempo de ejecución, T_{exec}** : tiempo medio requerido para extraer características y descriptores.
- **Uso de memoria, M_{mem}** : memoria consumida durante la ejecución del extractor.

Las mediciones se realizaron en un portátil con CPU Intel Core i7-11800H de 11ª generación, 8 núcleos/16 hilos, 15 GiB de RAM y GPU NVIDIA GeForce RTX 3050 Ti Mobile, usando Ubuntu 24.04.3 LTS y Python 3.12.3.

4. Resultados

4.1. Análisis geométrico

La Figura 2 muestra la utilidad geométrica, la tasa de viabilidad y la tasa de inliers para las tres degradaciones. La reducción de resolución espacial produce el deterioro geométrico más marcado: al perderse textura fina y estructuras locales, la utilidad geométrica cae bruscamente y la tasa de viabilidad disminuye en los niveles más agresivos. Sin embargo, la tasa de inliers se mantiene relativamente alta en algunos casos, lo que indica que el problema principal no es sólo la presencia de correspondencias erróneas, sino la falta de suficientes inliers bien distribuidos. Los casos en que no es posible calcular la métrica por correspondencias insuficientes se indican con $\times$.

La cuantización radiométrica afecta menos al rendimiento geométrico, ya que conserva la estructura espacial de la imagen. La tasa de viabilidad permanece cercana a uno y la tasa de inliers sólo disminuye ligeramente. La degradación fotoelectrónica de baja iluminación presenta una caída progresiva de la utilidad geométrica desde *limpio* hasta *severo*; en el nivel más severo, la tasa de inliers cae de forma notable, indicando que muchas correspondencias emparejadas dejan de ser geoméricamente consistentes. En conjunto, R2D2 muestra el comportamiento geométrico más estable, mientras que ORB se mantiene como una alternativa competitiva, especialmente en condiciones moderadas.

4.2. Robustez frente a degradación

La Figura 3 resume la deriva del descriptor, la cobertura y la MMA al comparar la imagen de referencia con su versión degradada. Bajo reducción espacial, la cobertura y la MMA disminuyen rápidamente, mientras que la deriva aumenta. Esto confirma que la pérdida de detalle espacial afecta simultáneamente a la repetibilidad de los puntos, la estabilidad del descriptor y la precisión del emparejamiento.

La cuantización radiométrica presenta un comportamiento distinto: en el nivel de 1 bit, la cobertura aumenta para algunos extractores, especialmente AKAZE y AKAZE-float, debido a estructuras de contraste artificialmente repetibles. Sin embargo, la MMA disminuye, lo que indica que esos puntos repetibles no siempre generan descriptores discriminativos. En baja iluminación, la cobertura es muy baja desde los niveles *leve* y *moderado*, y en el nivel *severo* apenas se recuperan correspondencias válidas. Este resultado muestra que la degradación de sensor afecta de forma más severa a la estabilidad referencia-degradada que la cuantización directa.

4.3. Coste computacional

La Figura 4 muestra el tiempo de ejecución y uso de memoria durante la ejecución. La reducción de resolución es la única degradación que reduce claramente el coste computacional, ya que se procesan menos píxeles y se extraen menos características. En cambio, la cuantización radiométrica, a niveles bajos de cuantización se detectan más características, lo que introduce un ligero aumento en el coste computacional.

La degradación fotoelectrónica de baja iluminación muestra un coste más irregular, especialmente en métodos aprendidos, debido a respuestas menos estables o más numerosas bajo ruido. R2D2 ofrece el mejor rendimiento geométrico global, pero con mayor tiempo de ejecución. D2-Net queda penalizado por su alto consumo de memoria. Por el contrario, ORB, AKAZE y AKAZE-float mantienen costes bajos y estables, lo que los convierte en alternativas relevantes para plataformas de recursos limitados.

5. Conclusiones

Este trabajo ha evaluado la robustez de extractores de características locales bajo condiciones representativas de sensor restringido para percepción robótica espacial. Los resultados muestran que las degradaciones analizadas no afectan de forma equivalente al proceso de correspondencia: la reducción espacial limita la detección de estructuras útiles, la cuantización reduce la discriminabilidad descriptiva y la baja iluminación provoca una mayor inestabilidad en las correspondencias.

R2D2 presenta la mayor robustez geométrica global, aunque con un coste computacional elevado. ORB ofrece un equilibrio favorable entre rendimiento y eficiencia, mientras que

Cross-degradation comparison — geometric metrics

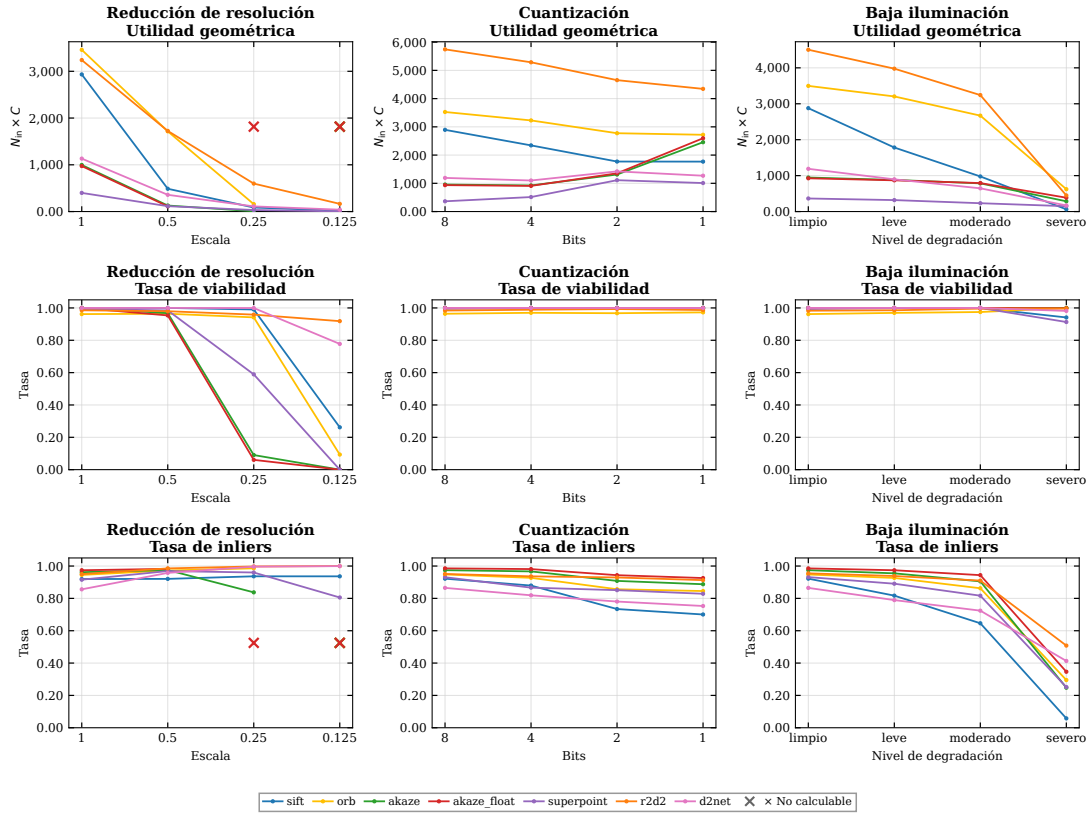


Figura 2: Rendimiento geométrico bajo degradaciones controladas: reducción de resolución espacial, cuantización radiométrica y degradación fotoelectrónica de baja iluminación. Se muestran la utilidad geométrica, la tasa de viabilidad y la tasa de inliers.

Cross-degradation comparison — robustness metrics

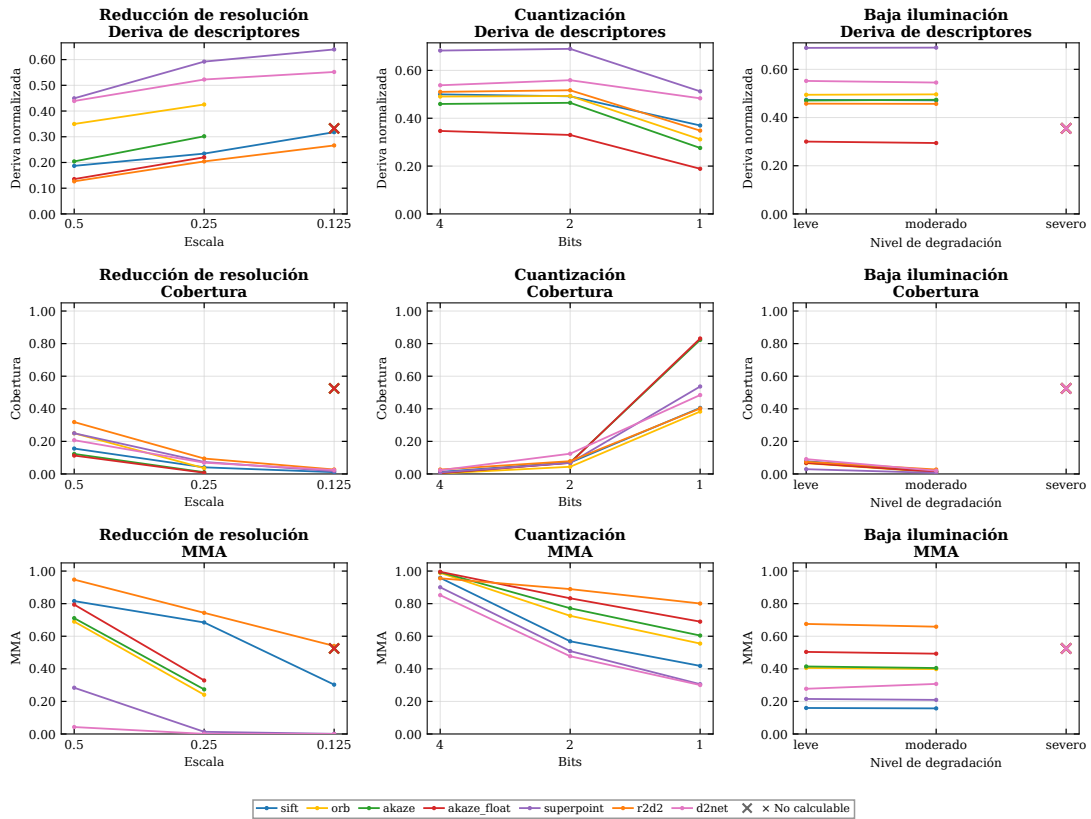


Figura 3: Robustez de las características frente a degradación. Se muestran la deriva del descriptor, la cobertura y la MMA para la reducción de resolución espacial, la cuantización radiométrica y la degradación fotoelectrónica de baja iluminación.

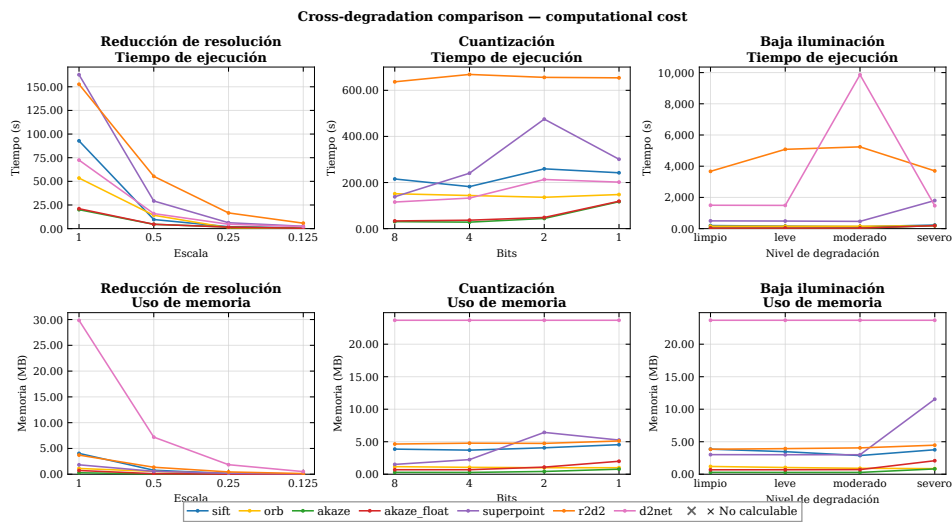


Figura 4: Coste computacional bajo degradaciones controladas. Se muestran el tiempo de ejecución y el uso de memoria para cada extractor.

AKAZE y AKAZE-float siguen siendo alternativas ligeras y competitivas en ciertos escenarios. Estos resultados sugieren que la elección del extractor no debe basarse únicamente en su rendimiento medio, sino en el tipo de degradación dominante y en las restricciones de la plataforma robótica.

Agradecimientos

Este trabajo se ha realizado en el marco del proyecto INSIGHT (PID2024-160373OB-C21), financiado por el Ministerio de Ciencia, Innovación y Universidades, la Agencia Estatal de Investigación y el Fondo Europeo de Desarrollo Regional (FEDER, UE).

Referencias

- Baltas, V., Lenc, K., Vedaldi, A., Mikolajczyk, K., 2017. HPatches: A benchmark and evaluation of handcrafted and learned local descriptors. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- Bay, H., Tuytelaars, T., Van Gool, L., 2006. SURF: Speeded up robust features. In: Computer Vision – ECCV. Vol. 3951. pp. 404–417. DOI: 10.1007/11744023_32
- Calonder, M., Lepetit, V., Strecha, C., Fua, P., 2010. BRIEF: Binary robust independent elementary features. In: Eur. Conf. Comput. Vis. Vol. 6314. pp. 778–792. DOI: 10.1007/978-3-642-15561-1_56
- DeTone, D., Malisiewicz, T., Rabinovich, A., 2018. Superpoint: Self-supervised interest point detection and description. In: 2018 IEEE/CVF CVPR Workshops. pp. 337–33712. DOI: 10.1109/CVPRW.2018.00060
- Dusmanu, M., Rocco, I., Pajdla, T., Pollefeys, M., Sivic, J., Torii, A., Sattler, T., 2019. D2-net: A trainable cnn for joint detection and description of local features. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 8092–8101.
- Fernández Alcantarilla, P., 2011. Fast explicit diffusion for accelerated features in nonlinear scale spaces. In: IEEE TPAMI. pp. 1281–1298. DOI: 10.5244/C.27.13
- Fernández Alcantarilla, P., Bartoli, A., Davison, A., 2012. KAZE features. In: European Conference on Computer Vision (ECCV). pp. 214–227. DOI: 10.1007/978-3-642-33783-3_16
- Fischler, M. A., Bolles, R. C., Jun, 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 24 (6), 381–395. DOI: 10.1145/358669.358692
- Gerdes, L., Wiese, T., Castilla-Arquillo, R., Bielenberg, L., Azkarate, M., Leblond, H., Wilting, F., Cortés, J., Bernal, A., Palanco, S., Perez-del Pulgar, C., 2024. BASEPROD: The bardenas semi-desert planetary rover dataset. Scientific Data 11, 1–15. DOI: 10.1038/s41597-024-03881-1
- Gil, A., Mozos, O. M., Ballesta, M., Reinoso, O., 2010. A comparative evaluation of interest point detectors and local descriptors for visual slam. Machine Vision and Applications 21 (6), 905–920. DOI: 10.1007/s00138-009-0195-x
- Karami, E., Prasad, S., Shehata, M., 2015. Image matching using SIFT, SURF, BRIEF and ORB: performance comparison for distorted images. DOI: 10.1109/iCoMET57998.2023.10099250
- Khan Tareen, S. A., Raza, R. H., 2023. Potential of SIFT, SURF, KAZE, AKAZE, ORB, BRISK, AGAST, and 7 more algorithms for matching extremely variant image pairs. In: 4th International Conference on Computing, Mathematics and Engineering Technologies (iCoMET). pp. 1–6. DOI: 10.1109/iCoMET57998.2023.10099250
- Leutenegger, S., Chli, M., Siegwart, R., 2011. BRISK: Binary robust invariant scalable keypoints. In: IEEE International Conference on Computer Vision (ICCV). pp. 2548–2555. DOI: 10.1109/ICCV.2011.6126542
- Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60 (2), 91–110. DOI: 10.1023/B:VISI.0000029664.99615.94
- Meyer, L., Smíšek, M., Fontan Villacampa, A., Oliva Maza, L., Medina, D., Schuster, M. J., Steidle, F., Vayugundla, M., Müller, M. G., Rebele, B., Wedler, A., Triebel, R., 2021. The madmax data set for visual-inertial rover navigation on mars. Journal of Field Robotics 38 (6), 833–853. DOI: https://doi.org/10.1002/rob.22016
- Revaud, J., De Souza, C., Humenberger, M., Weinzaepfel, P., 2019. R2D2: Reliable and repeatable detector and descriptor. In: Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., Garnett, R. (Eds.), Advances in Neural Information Processing Systems. Vol. 32.
- Rodríguez-Martínez, D., van der Meer, D., Song, J., Bera, A., Pérez-del Pulgar, C. J., Olivares-Mendez, M. A., 2026. SPICE-HL3: Single-photon, inertial, and stereo camera dataset for exploration of high-latitude lunar landscapes. Scientific Data 13 (1). DOI: 10.1038/s41597-026-06668-8
- Rublee, E., Rabaud, V., Konolige, K., Bradski, G., 2011. ORB: An efficient alternative to SIFT or SURF. In: International Conference on Computer Vision. pp. 2564–2571. DOI: 10.1109/ICCV.2011.6126544
- Salau, A. O., Jain, S., 2019. Feature extraction: A survey of the types, techniques, applications. In: International Conference on Signal Processing and Communication (ICSC). pp. 158–164. DOI: 10.1109/ICSC45622.2019.8938371
- Scaramuzza, D., Fraundorfer, F., 2011. Visual odometry [tutorial]. IEEE Robot. Automat. Mag. 18, 80–92. DOI: 10.1109/MRA.2011.943233
- Tareen, S. A. K., Saleem, Z., 2018. A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK. In: International Conference on Computing, Mathematics and Engineering Technologies (iCoMET). pp. 1–10. DOI: 10.1109/ICOMET.2018.8346440
- Vayugundla, M., Steidle, F., Smisek, M., Schuster, M. J., Bussmann, K., Wedler, A., 2018. Datasets of long range navigation experiments in a moon analogue environment on mount etna. In: ISR 2018; 50th International Symposium on Robotics. pp. 1–7.