

Superresolución de imágenes basado en la transformada wavelet bidimensional Image super resolution through two-dimensional wavelet transform

R. J. Osorno-Ortiz ^a, V. Ponomaryov ^a, R. Reyes-Reyes ^{a,*}, C. Cruz-Ramos ^a

^a Instituto Politécnico Nacional, ESIME Culhuacan, Av. Sta. Ana 1000, San Fco. Culhuacan, 04440, CDMX, México.

Resumen

La superresolución (SR) es una técnica diseñada para aumentar la resolución espacial de una imagen digital de baja resolución (LR). A diferencia de los algoritmos basados en métodos de interpolación como método principal, los cuales en su mayoría tienden a deformar los bordes o producen imágenes con bordes irregulares; en este artículo se propone un algoritmo que preserva los bordes de la imagen original mediante los métodos de interpolación entre las sub-bandas de detalles obtenidas por medio de la Transformada Discreta Wavelet (DWT). Se realiza la descomposición wavelet utilizando tres diferentes familias: Daubechies, Symlet y Coiflet. Finalmente, todas las imágenes de sub-banda interpoladas se combinan para la generación de la imagen SR. Los resultados obtenidos demuestran un buen desempeño de acuerdo con las métricas objetivas: tiempo de ejecución, SSIM y PSNR (1.669 seg., 0.8908 y 30.61 dB respectivamente para imágenes con una superresolución de 4080x2712) y en términos subjetivos medidos por medio de la percepción visual humana de diferentes imágenes.

Palabras Clave: Transformada Discreta Wavelet (DWT), interpolación, extracción de bordes, superresolución.

Abstract

Super resolution (SR) is a technique aimed at enhancing the spatial resolution of low-resolution digital images (LR). In contrast to interpolation-based algorithms that often introduce distortions or irregular borders, the algorithm proposed in this article preserves the edges of the original image by interpolating the detail sub-bands derived from the Discrete Wavelet Transform (DWT). The wavelet decomposition was conducted using three different families: Daubechies, Symlet, and Coiflet. Subsequently, all the interpolated sub-band images are combined to generate the SR image. The results obtained demonstrate excellent performance in terms of objective metrics, including execution time, SSIM, and PSNR (1.669 sec., 0.8908, and 30.61 dB, respectively, for 4080x2712 super resolution images). Moreover, subjective metrics based on human visual perception also indicate favorable outcomes across various images.

Keywords: Discrete Wavelet Transform (DWT), interpolation, edge extraction, super resolution.

1. Introducción

Hoy en día, las imágenes capturadas desde sensores ópticos, radares, cámaras de vehículos aéreos no tripulados (VANT) y cámaras de teledetección tienen características físicas (tamaño de los sensores de luz o del lente) que limitan la resolución con la que son capturadas. Debido a sus componentes físicos y los altos costos de los sensores, se ha vuelto necesario mejorar su precisión, empleando métodos y algoritmos que ofrezcan una mejora en la resolución de las imágenes de dichos dispositivos.

Así surge la necesidad de desarrollar diversas propuestas para la mejora de la resolución de una imagen digital. Existen propuestas basadas en métodos de superresolución que

realizan técnicas de interpolación que calculan los valores de los píxeles cercanos a un punto de aproximación. El inconveniente que presentan los métodos basados en interpolaciones es que tienden a deformar los bordes o producen una imagen SR con bordes irregulares.

Para la mejora de la resolución en imágenes digitales se han desarrollado métodos que implementan el uso de la Transformación Wavelet bidimensional (2D-DWT). La cual se obtiene aplicando la descomposición Wavelet en una dimensión (1D-DWT) a lo largo de las filas de la imagen y posteriormente a lo largo de las columnas. Esta operación genera imágenes de sub-banda que contienen información de los detalles de las altas y bajas frecuencias en diferentes direcciones: aproximación (LL), detalle horizontal (LH),

*Autor para la correspondencia: rreyesre@ipn.mx

Correo electrónico: rosomoo1600@alumno.ipn.mx (Raúl Jair Osorno-Ortiz), vponomar@ipn.mx (Volodymyr Ponomaryov), rreyesre@ipn.mx (Rogelio Reyes-Reyes), ccruzra@ipn.mx (Clara Cruz-Ramos).

detalle vertical (HL) y diagonal (HH). Los componentes frecuenciales de las imágenes de sub-banda abarcan el espectro de frecuencia completo de la imagen original.

Una propuesta es diseñada por (Chavez et al., 2012) que propone utilizar las imágenes de sub-banda para interpolarlas y añadir los valores obtenidos mediante el método de extracción de bordes Sobel; para aproximar los valores de las imágenes obtenidas de la DWT, utilizan una imagen de diferencia entre los valores de la imagen LR y los coeficientes LL de las sub-bandas. El trabajo hecho por (Bhatt et al., 2016) propone utilizar la DWT en conjunto con la Transformada Wavelet Estacionaria (SWT) para una mejor reconstrucción de una imagen SR, ya que, a diferencia de otras transformadas Wavelet, la SWT conserva la energía de la señal en todas las escalas y resoluciones. Por último, el diseño propuesto por (Soumya et al., 2016) implementa el uso de la Transformada Wavelet de Elevación (LWT) en conjunto con la DWT; debido a que la LWT suele ser más rápida que la DWT, ya que la LWT utiliza filtros de lifting para reducir el número de cálculos requeridos, sin embargo, no se utiliza la información extraída de la sub-banda LL que otorga una mejor reconstrucción de la imagen SR.

Las principales contribuciones de este trabajo se resumen a continuación: (1) imágenes con diversas características de dos diferentes bases de datos fueron probadas y evaluadas mediante el sistema propuesto obteniendo resultados satisfactorios, lo que demuestra la mejora en la resolución de imágenes de diferente naturaleza; (2) el algoritmo propuesto no requiere de una etapa final de interpolación a diferencia de algunas propuestas en el estado del arte, demostrando no tener pérdidas significativas en los componentes de altas frecuencias de la imagen, preservando los bordes con detalles finos.

En este trabajo se propone el desarrollo de un algoritmo utilizando técnicas de interpolación aplicadas a las imágenes de sub-banda de la DWT en la imagen LR. Posteriormente las imágenes resultantes son combinadas e interpoladas con la información extraída del proceso de extracción de bordes para finalmente aplicar la Transformación Wavelet Discreta Inversa (IDWT) y reconstruir la imagen SR.

2. Método propuesto

El algoritmo propuesto incluye dos etapas como se muestra en la Figura 1: descomposición Wavelet y etapa de extracción de bordes. El método propuesto ofrece conservar la información de las altas frecuencias de la imagen a través del método de extracción de bordes y a las imágenes de sub-banda de altas y bajas frecuencias de la DWT.

A. Descomposición Wavelet

La figura 1 muestra el diagrama a bloques del sistema propuesto para la superresolución de la imagen, la cual consta de dos fases, la primera consiste en la extracción de los coeficientes por medio de la obtención de las imágenes de sub-banda de la Transformada Wavelet Discreta bidimensional y la segunda en la extracción de los bordes mediante el método de Canny. La descomposición Wavelet genera cuatro imágenes de sub-banda a partir de la imagen LR que representan: coeficientes de aproximación (LL), detalles horizontales (LH), detalles verticales (HL) y detalles diagonales (HH).

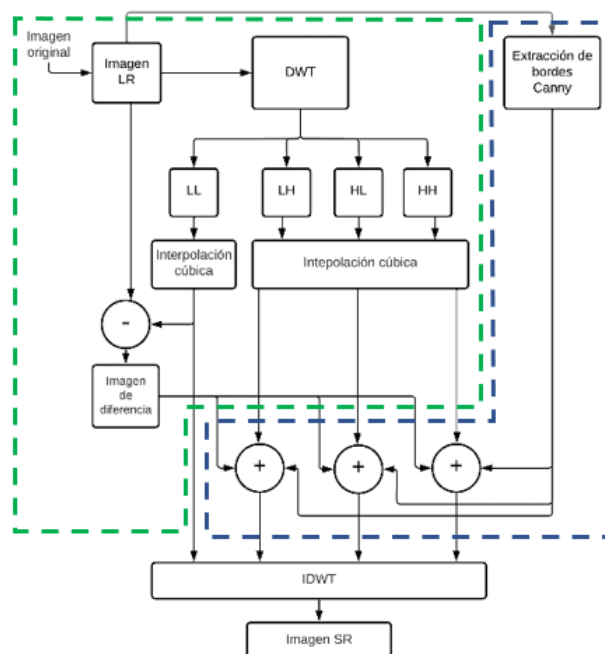


Figura 1. Diagrama a bloques del proceso de superresolución

Posteriormente, se realiza una interpolación cúbica a la imagen LL para obtener las mismas dimensiones de la imagen LR y realizar el cálculo de una imagen de diferencia que se añadirá a las imágenes LH, HL y HH interpoladas con el método de interpolación cúbica.

B. Extracción de bordes

El método de extracción de bordes de Canny detecta bordes o curvas en una imagen digital cuando el brillo cambia abruptamente en la dirección x y y , posteriormente combina esta información para obtener una imagen en escala de grises que muestra los bordes de la imagen original. La imagen resultante se añade a las imágenes interpoladas con la información de la imagen de diferencia para de esta manera resaltar los bordes; finalmente, se emplean las cuatro imágenes de detalle de las sub-bandas interpoladas en la operación IDWT para obtener la imagen reconstruida de SR.

3. Métricas de desempeño

Para evaluar objetivamente el rendimiento del algoritmo propuesto de superresolución, se emplean los siguientes criterios: relación señal-ruido pico (PSNR) y el índice de similitud estructural perceptiva (SSIM) calculados entre la imagen SR original y la imagen SR obtenida a partir de la imagen LR con el método propuesto basado en (Wang et al., 2004). El PSNR se obtiene mediante (1):

$$PSNR = 10 \cdot \log_{10} \frac{(255)^2}{MSE}. \quad (1)$$

Las métricas estándar como PSNR utilizadas en el pasado pueden arrojar resultados erróneos al analizar los métodos propuestos; por lo tanto, métricas como SSIM son útiles ya que ofrecen una coincidencia mayor relacionada a la subjetividad de la percepción humana, y deben utilizarse para caracterizar

el rendimiento del algoritmo. La métrica de los valores del SSIM se obtienen mediante (2):

$$SSIM(y, \hat{y}) = [l(y, \hat{y})] \cdot [c(y, \hat{y})] \cdot [s(y, \hat{y})], \quad (2)$$

donde los parámetros l , c y s se calculan usando (3)-(5) para cada imagen:

$$l(y, \hat{y}) = \frac{2\mu_y\mu_{\hat{y}} + C_1}{\mu_y^2 + \mu_{\hat{y}}^2 + C_1}, \quad (3)$$

$$c(y, \hat{y}) = \frac{2\sigma_y\sigma_{\hat{y}} + C_2}{\sigma_y^2 + \sigma_{\hat{y}}^2 + C_2}, \quad (4)$$

$$s(y, \hat{y}) = \frac{\sigma_{y\hat{y}} + C_3}{\sigma_y\sigma_{\hat{y}} + C_3}, \quad (5)$$

donde \hat{y} es la imagen SR reconstruida e y es la imagen original (SR); μ y σ^2 son el valor de la media y la varianza de las muestras para las imágenes \hat{y} o y respectivamente, y $\sigma_{y\hat{y}}$ es la varianza cruzada entre las muestras de las imágenes \hat{y} o y . l representa la similitud de luminancia, c caracteriza la similitud de contraste y s es la similitud estructural de una imagen. La justificación de los índices del SSIM se pueden encontrar en (Wang et al., 2004). Las constantes C_1 , C_2 , C_3 son usadas para estabilizar la métrica en el caso en que las medias y las varianzas se vuelven muy pequeñas, y comúnmente $C_1 = C_2 = C_3 = 1$.

La idea clave detrás del índice SSIM es reconocer que las imágenes naturales están muy estructuradas y que la medida de la correlación estructural entre la imagen original (sin corromper) y la imagen filtrada es muy importante, y define de manera general la calidad visual. Además, el índice SSIM mide la calidad localmente y es capaz de capturar disimilitudes mejor, a diferencia de medidas de calidad global como el PSNR. A pesar de que emplear (3)-(5) es más complejo que el PSNR, es analíticamente más manipulable.

Debido a la dificultad de definir los criterios de error que deben utilizarse para garantizar una reconstrucción precisa, se utiliza también un análisis subjetivo basado en la percepción visual humana. Una comparación visual subjetiva de las imágenes proporciona otra forma de evaluar la información sobre la distorsión espacial de los diferentes detalles obtenidos mediante los métodos de interpolación, extracción de bordes y la DWT.

4. Resultados experimentales

El sistema se desarrolló en el IDE de Matlab@2022b, utilizando una computadora con un procesador Intel® Core i5-11260H, memoria RAM de 8.00GB y un Sistema Operativo Windows 10 Home Single Language de 64 bits. La implementación secuencial del algoritmo se llevó a cabo empleando el CPU antes mencionado.

Las imágenes digitales que se utilizaron para evaluar el sistema propuesto se obtuvieron de dos bases de datos; la primera DIV2K (Agustsson et al., 2017) la cual contiene un total de 800 imágenes de alta calidad y, la segunda SIPI (Tamrakar et al., 2005) la cual contiene un total de 108

imágenes. En la figura 2 se muestran como ejemplo algunas de las imágenes de estas bases de datos utilizadas para evaluar el algoritmo mediante las métricas SSIM y PSNR.

Las imágenes originales fueron escaladas a diferentes dimensiones en las imágenes LR.

Las imágenes mostradas a continuación tienen una resolución de: Flores (2040x1536), Mariposa (4080x2712), Lena (1024x1024), Cabra (4080x2712), Cisne (4080x2712) y Colibrí (4080x2712).

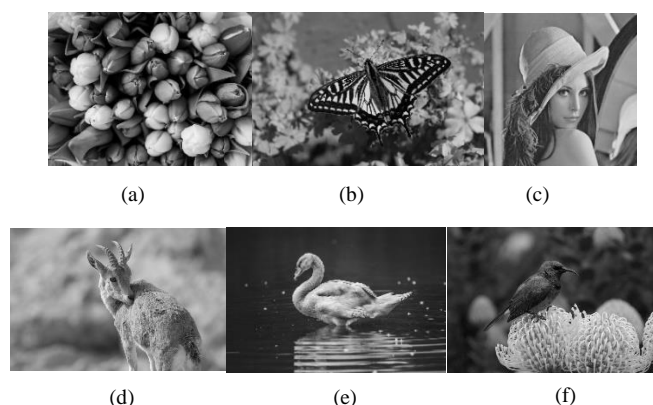


Figura 2. Imágenes originales utilizadas: (a) Flores, (b) Mariposa, (c) Lena, (d) Cabra, (e) Cisne y (f) Colibrí.

Para evaluar el algoritmo propuesto también se calcula el tiempo de ejecución, el SSIM y el PSNR de las imágenes, en adición se presentan las imágenes SR obtenidas para realizar una evaluación subjetiva (percepción visual humana) propuesta por (Pappas et al., 2000).

La compresión aplicada a la imagen original fue de 0.25 a lo largo de las filas y columnas, por lo que se tiene una imagen resultante de 1/16 del tamaño como imagen LR y de acuerdo con las figuras 3, 4 y 5 se observa que la imagen SR se redimensiona a dieciséis veces el tamaño de la imagen LR, obteniendo como resultado una imagen del mismo tamaño que la imagen original. Finalmente se realiza la evaluación objetiva del algoritmo empleando las métricas SSIM y PSNR, para ambas imágenes (original y SR).

En la Tabla 1 y 2 se observan las métricas de tiempo de ejecución, SSIM y PSNR para la evaluación de las imágenes propuestas: Mariposa, Lena, Flores, Cabra, Cisne y Colibrí empleando diferentes familias Wavelet: Daubechies, Symlet y Coiflet, mostrando las 5 wavelets que ofrecen mejores resultados para cada familia.

Se demuestra que las imágenes SR resultantes tienen una alta calidad de acuerdo con las métricas objetivas SSIM y PSNR obtenidas, 0.918 y 32.76dB respectivamente. De igual forma se puede apreciar que mientras más pequeña sea la imagen LR de entrada al sistema, la imagen SR perderá calidad al ser generada, como se puede observar en los valores SSIM y PSNR de la imagen Lena (0.59 y 17.8) debido a la reducida cantidad de información que puede obtenerse con respecto a las altas y bajas frecuencias de la imagen original.

Tabla 1. Valores de tiempo de ejecución, PSNR y SSIM de la imagen Mariposa, Lena y Flores utilizando las familias wavelet: Daubechies, Symlet, Coiflet con superresolución aumentada a 16 veces su tamaño original.

	Mariposa				Lena				Flores			
Familia	Función wavelet	Tiempo de ejecución (s)	SSIM	PSNR (dB)	Función wavelet	Tiempo de ejecución (s)	SSIM	PSNR (dB)	Función wavelet	Tiempo de ejecución (s)	SSIM	PSNR (dB)
Daubechies	db1	2.40	0.843	25.39	db1	0.55	0.694	23.89	db1	1.06	0.734	24.24
	db2	1.42	0.906	30.07	db2	0.29	0.764	27.43	db2	0.60	0.847	28.89
	db3	1.44	0.903	28.56	db3	0.24	0.755	26.17	db3	0.38	0.836	27.29
	db4	1.27	0.876	24.33	db4	0.21	0.701	22.62	db4	0.42	0.768	23.03
	db5	1.35	0.846	21.42	db5	0.27	0.640	20.02	db5	0.42	0.694	20.09
Symlet	sym2	1.47	0.906	30.07	sym2	0.28	0.764	27.43	sym2	0.28	0.764	27.43
	sym7	1.30	0.909	30.49	sym7	0.22	0.766	27.47	sym7	0.22	0.766	27.47
	sym16	1.29	0.912	29.33	sym12	0.25	0.773	26.57	sym12	0.25	0.773	26.57
	sym28	4.61	0.913	29.56	sym14	0.29	0.774	26.90	sym14	0.29	0.774	26.90
	sym30	8.16	0.913	29.63	sym24	2.89	0.776	27.21	sym24	2.89	0.776	27.21
Coiflet	coif1	2.93	0.904	28.83	coif1	0.23	0.759	26.55	coif1	0.46	0.840	27.66
	coif2	2.27	0.906	29.24	coif2	0.15	0.758	26.62	coif2	0.40	0.842	27.99
	coif3	2.22	0.870	23.89	coif3	0.14	0.681	22.14	coif3	0.42	0.752	22.57
	coif4	2.16	0.837	20.70	coif4	0.12	0.622	19.36	coif4	0.37	0.670	19.36
	coif5	2.23	0.814	18.84	coif5	0.13	0.599	17.80	coif5	0.41	0.617	17.50

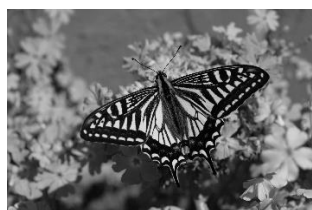


Imagen original (4080x2712)

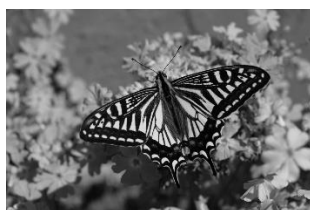


Imagen LR (1020x678)

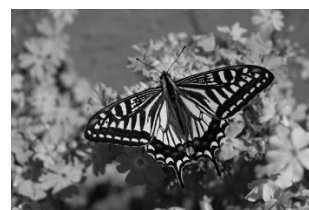


Imagen SR (4080x2712)

Figura 3. Resultados de la percepción visual para Mariposa.

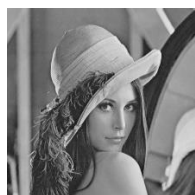


Imagen original (1024x1024)

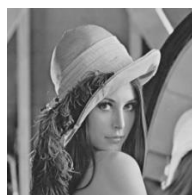


Imagen LR (256x256)

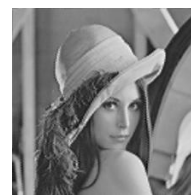


Imagen SR (1024x1024)

Figura 4. Resultados de la percepción visual para Lena.

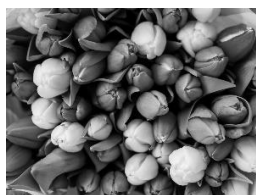


Imagen original (2040x1536)



Imagen LR (510x384)



Imagen SR (2040x1536)

Figura 5. Resultados de la percepción visual para Flores.

Tabla 2. Valores de tiempo de ejecución, PSNR y SSIM de la imagen Cabra, Cisne y Colibrí utilizando las familias wavelet: Daubechies, Symlet, Coiflet con superresolución aumentada a 16 veces su tamaño original.

	Cabra				Cisne				Colibrí			
Familia	Función wavelet	Tiempo de ejecución (s)	SSIM	PSNR (dB)	Función wavelet	Tiempo de ejecución (s)	SSIM	PSNR (dB)	Función wavelet	Tiempo de ejecución (s)	SSIM	PSNR (dB)
Daubechies	db1	2.90	0.874	31.20	db1	1.43	0.868	28.06	db1	1.40	0.860	29.58
	db2	1.58	0.899	33.75	db2	1.40	0.909	32.03	db2	1.43	0.907	34.34
	db3	1.34	0.899	33.25	db3	1.36	0.910	30.79	db3	1.28	0.908	32.59
	db4	1.28	0.885	30.65	db4	1.28	0.887	26.89	db4	1.25	0.884	28.26
	db5	1.40	0.868	28.46	db5	1.29	0.856	24.00	db5	1.29	0.855	25.27
Symlet	sym11	1.34	0.902	33.90	sym7	1.37	0.915	32.31	sym11	1.29	0.917	34.54
	sym19	1.35	0.904	34.20	sym11	1.28	0.915	32.23	sym14	1.32	0.917	33.45
	sym23	1.49	0.904	34.19	sym14	1.30	0.917	31.44	sym15	1.31	0.919	35.17
	sym24	2.08	0.903	33.71	sym19	1.33	0.918	32.76	sym23	1.63	0.920	35.21
	sym30	8.33	0.903	33.67	sym28	4.87	0.918	31.69	sym24	2.20	0.919	33.93
Coiflet	coif1	1.52	0.897	33.19	coif1	1.49	0.907	31.03	coif1	1.43	0.905	33.09
	coif2	1.34	0.900	33.47	coif2	1.55	0.912	31.32	coif2	1.47	0.912	33.34
	coif3	1.31	0.878	30.17	coif3	1.38	0.878	26.38	coif3	1.30	0.878	27.81
	coif4	1.31	0.863	27.85	coif4	1.35	0.843	23.22	coif4	1.30	0.846	24.52
	coif5	1.34	0.859	26.51	coif5	1.31	0.822	21.36	coif5	1.31	0.826	22.64



Imagen original (4080x2712)



Imagen LR (1020x678)



Imagen SR (4080x2712)

Figura 6. Resultados de la percepción visual para Cabra.



Imagen original (4080x2712)



Imagen LR (1020x678)



Imagen SR (4080x2712)

Figura 7. Resultados de la percepción visual para Cisne.



Imagen original (4080x2712)



Imagen LR (1020x678)



Imagen SR (4080x2712)

Figura 8. Resultados de la percepción visual para Colibrí.

Para demostrar el rendimiento del algoritmo, en adición, con la evaluación objetiva presentada en las Tablas 1 y 2, se muestran las figuras 3 a 8 para poder realizar una evaluación subjetiva de los resultados obtenidos utilizando la percepción visual de cada individuo, mostrando que no existen diferencias significativas visualmente entre las imágenes LR y SR obtenidas.

La diferencia en la calidad y cantidad de detalles presentes en las 6 imágenes analizadas producen una variación en las métricas objetivas SSIM y PSNR; debido a que el análisis de la DWT se aplica directamente a los detalles de una imagen en diferentes direcciones (horizontal, vertical y diagonal). Tomando en consideración que una imagen puede tener mayor o menormente presentes estos detalles, no es posible definir una única familia de funciones wavelet o una función wavelet específica que funcione adecuadamente para cualquier tipo de imagen; por consiguiente, se recomienda realizar una evaluación con las diferentes familias de funciones wavelet para seleccionar la más adecuada al tipo de imagen de entrada.

5. Conclusiones

El método propuesto en el presente artículo que implementa el uso de la DWT en conjunto con un método de extracción de bordes proporciona una mejora significativa en la calidad de la imagen SR sin requerir de una etapa final de interpolación, preservando los bordes de una mejor manera comparado con algoritmos similares y con los basados únicamente en métodos de interpolación. Adicionalmente, el sistema demostró un rendimiento robusto para diversos tipos de imágenes con diferentes características, las cuales se obtuvieron de dos bases de datos diferentes dando como resultado valores promedio en tiempo de ejecución, SSIM y PSNR de 1.669 seg., 0.8908 y 30.61 dB respectivamente, para imágenes con una superresolución de 4080x2712.

Como trabajo a futuro, se pretende implementar y evaluar el método propuesto en archivos de video, así como realizar una paralelización del algoritmo para reducir los tiempos de ejecución.

Agradecimientos

Los autores agradecen al Instituto Politécnico Nacional (IPN-México), a la Comisión de Operación y Fomento de Actividades Académicas (COFAA-IPN) y al Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT), por el apoyo brindado para la realización de este trabajo.

Referencias

- Tamrakar, A. & Ortega, A. (2005). *Base de datos de imágenes USC-SIPI*. <https://sipi.usc.edu/database/>
- Bhatt, U., Singh, A., Bhaduria, H. S., & Kumar, M. (2016). Image super resolution based on discrete and stationary wavelet transform using canny edge extraction and non local mean. In 2016 International Conference on Inventive Computation Technologies (ICICT) (Vol. 3, pp. 1-5). IEEE. <https://doi.org/10.1109/inventive.2016.7830216>.
- Chavez-Roman, H., Ponomaryov, V., & Peralta-Fabi, R. (2012). Image super resolution using interpolation and edge extraction in wavelet transform space. In 2012 9th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE) (pp. 1-6). IEEE. <https://doi.org/10.1109/ICEEE.2012.6421202>.
- Agustsson, E., & Timofte, R. (2017). Ntire 2017 challenge on single image super-resolution: Dataset and study. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 126-135).
- Pappas, T. N., Safranek, R. J., & Chen, J. (2000). Perceptual criteria for image quality evaluation. *Handbook of image and video processing*, 110.
- Sowmya, K. (2016). Single image super resolution with wavelet domain transformation and sparse representation. *International Journal of Innovative Research in Computer Science & Technology (IJIRCST)*, 4.
- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4). <https://doi.org/600-612.10.1109/TIP.2003.819861>