Removing Text and Inpainting Irregular Regions in Digital Images within Spatial and Frequency Domains

Zainab A. Abdul Kareem[1,*], and Ahmed K. Al-Jaberi[2]

1. Department of Mathematics, Open Educational College, The General Directorate of Basrah Governorate Education, Basra, Iraq.

2. Department of Mathematics, College of Education for Pure Science, University of Basrah, Basra, Iraq

*Corresponding authors E-mail: almosawizainab27@gmail.com

https://doi.org/10.29072/basjs.20250103

ARTICLE INFO

Keywords

Image Inpainting; Frequency;Wavelet Transform(Dwt); Bertalmio; Transport.

ABSTRACT

Techniques for inpainting missing irregularly shaped regions and removing missing text in digital images have been introduced in both the frequency and spatial domains. Comparison of methodologies employed in the two domains Techniques that function directly on pixels in the spatial domain are straightforward and effective for real-time applications; nevertheless, they may lack the flexibility to simultaneously enhance all aspects of the image. Mathematical transformations, such as the Fourier transform and the discrete wave transform (DWT), are employed, with the discrete wave transform being a crucial instrument, particularly when utilizing Haar wavelets.

Received 13 Dec 2024; Received in revised form 30 Mar 2025; Accepted 19 Apr 2025, Published 30 Apr 2025

1. Introduction

Image inpainting is a technique that seeks to restore the damaged section of an image in order to rebuild it and provide a high-quality semantic approximation of the original image[1]. Image inpainting became an important subject of research with wide fields of applications such as medical imaging, augmented reality, art, computer-assisted restoration, image coding, and transmission[2]. The disadvantage of the image, a technique of restoring the lost or damaged parts of the images, uses various algorithms ranging from methods based on diffusion to in - depth learning approaches [3]. Its applications in digital art and catering have had a significant impact on creativity, improving technological capacities in the visual media. Image inpainting domain is classified into two categories: spatial domain as well as Frequency wavelet domain [4]. The inpainting techniques directly work to recover the missing areas in images that are defined in the spatial domain [5,6]. The spatial domain method works directly on pixels in the image. The spatial domain techniques have many advantages such as it is less complex and easy to understand, thus they are used in real-time applications, but they have some disadvantages such as they do not provide adequate robustness. The methods are easy to computation and robust but it cannot able to enhance all the parts of the image in the same time very well and difficult to automate the image enhancement [7]. Generally, the spatial domain techniques are more efficient computationally and require fewer processing resources to implement [4]. Wavelet-based approaches, in particular, allow for localized frequency analysis across multiple resolutions, making them highly effective for tasks like denoising, compression, and image inpainting. Wavelet transforms have become essential in modern image processing due to their flexibility in analyzing both frequency and spatial details of images[8].The intensity matrix representation of an image refers to its spatial domain, where each entry represents the digitization of the reflected light at the related position in the image. The frequency domain is a powerful methodology centered around a particular Fourier transform mathematical tool, discrete wavelet transforms (DWT), as well discrete cosine transforms (DCT). Image inpainting task of images in the frequency domain is got by computing the Fourier transform (FT) of the image to be enhanced and multiplying the result by a filter and taking the inverse transform to make the enhanced image [9]. The mathematical functions of frequency domain methods work by applying the frequency transform, such as (FT), Discrete Wavelet Transform (DWT), and Discrete Cosine Transform (DCT) [7]. Wavelet Transform (WT) and (FT) are the most commonly used frequency transforms. The discrete FT and WT were used to

analyze any digital image into different frequency values. For both cases, transformed coefficients in the high-frequency value correspond to edges and other image discontinuities in an image, whereas the coefficients in the low-frequency values are associated with smooth areas. In 2012, the authors [10] found a technique for image inpainting by using wavelet decomposition. It had three steps, called image decomposition, structure, and texture inpainting. In the first step, the original image is decomposed into four parts using the WT where the four parts are used as one section for texture inpainting and three sections for the structure inpainting. The DWT is decomposing a signal into low and high-frequency sub-bands and each of them could be transformed repeatedly, providing multiple resolution representations of the signal at various spatial scales and various frequency ranges. The DWT of any signal is a representation of the signal of a family of orthonormal wavelet bases got from a single wavelet function named a mother wavelet [11]. However, a 2-dimensional DWT was used to decompose an image into different frequency components at different resolution scales. which can be decomposed using many wavelet decomposition filters such as the Haar wavelet filter, Daubechies wavelet, ..., and so on. Haar wavelet is the first known wavelet proposed by Hungarian mathematician Alfred Haar in 1910 [11] .The Haar wavelet transform allows for a compact representation of image details, which is crucial for the preservation and reconstruction of fine textures in image [12] .and is the simplest wave transform.

Original Haar filter can be defined by:

$$h(t) = \begin{cases} 1 & 0 \le t < 0.5 \\ -1 & 0.5 \le t < 1 \\ 0 & otherwise \end{cases}$$

At first, DWT works in the horizontal direction across the rows of the image into low and high-frequency sub-bands. The low-frequency sub-bands referrers to the approximation sub-band that represents low frequencies, whereas the high sub-band, named the details sub-band represents the high frequencies in the horizontal direction. However, each one of the sub-bands can have been vertically decomposed into two low and high sub-bands. Thus, the image is decomposing into 4-subbands: low-frequency sub-band (LL) and high-frequency sub-bands (LH, HL, and HH). The LL sub-band is low frequencies in horizontal and vertical directions while the LH, HL, and HH sub-bands are high frequencies toward the vertical direction,

horizontal direction as well in diagonal direction, respectively [6]. The study compares the Bertalmio model's spatial and frequency domain performance in digital image inpainting, showing that the spatial domain is better. It also tests the model on real photos to determine its efficacy with wavelet transform (DWT) and missing area size.

2 Transport Model (Bertalmio)

The term inpainting for computer science was first introduced by Bertalmio et al. in the year 2000[13]. In the process of inpainting, the missing region will be filled with information from the region around the gap. The curves of equal intensity (isophotes) arriving at the boundary are propagated inwards.Bertalmio et al. [13] .Were modified an approach based on partial differential equations (PDE)is designed to imitate the techniques of museum artists in restoring degraded artwork and paintings.

2.1 Mathematical Model

The inputs required for the algorithm in the Bertalmio model are just masks and images. The region to be inpainted is determined by the mask while the image is subjected to a noise reduction process. Then, the image enters inpainting loops, and the enhancement and processing are done within the boundaries of the missing area Ω. The original image is described as I_O and the area that has to be inpainted is denoted as Ω while the border of the region $\delta\Omega$ is shown in Fig.1.

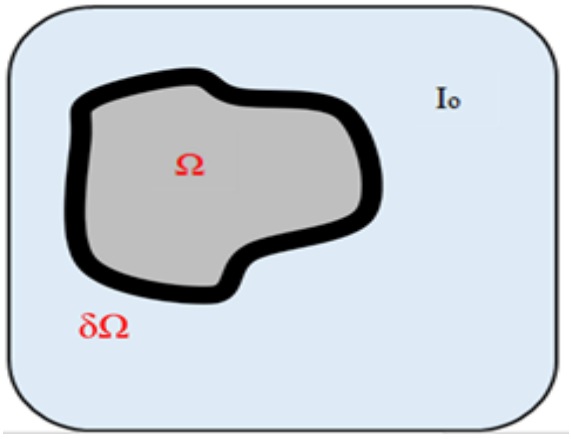

Figure 1: Image inpainting-based PDE algorithm

The Bertalmio model is an iterative method that works by creating a series of images I_1, I_2, I_3, … thus, the new image represents the enhanced image from the previous one. This process will

continue until a certain number of iterations or tiles of the algorithm converges. The process described by Bertalmio et al. [13] was supposed I is an image and (i,j) is the position of pixel then, the inpainting process for this pixel is given by:

$$I^{n+1}(i,j) = I^{n}(i,j) + \Delta t\, {I_t}^{n}(i,j) \ , \forall (i,j) \in \Omega \qquad 2$$

Where n is inpainting time. Every time a new value of pixel $I^{n+1}(i,j)$ is calculated depending on the old value of pixel $I^{n}(i,j)$.When the new value of pixel represents the value of old pixel plus the rate of improvement Δt .The improvement of pixel (i,j) can be calculated by:

$$I_t^n(i,j) = \left(\overrightarrow{\delta L^n}(i,j)\right).\frac{\vec{N}(i,j,n)}{|N(i,j,n)|} \mid \nabla I^n(i,j) \mid \qquad 3$$

The values for each iteration are represented by $I_t^n(i,j)$ can be calculated using equation 2. The value of 2D smoothness estimation (L)at pixel (i,j) is given by:

$$L^n(i,j) = I_{xx}^n(i,j) + I_{yy}^n(i,j) \qquad 4$$

Where $\overrightarrow{\delta L^n}(i,j)$ is measure the change of the information $L^n(i,j)$[2] along $\vec{N}$ from the equation:

$$\overrightarrow{\delta L^n}(i,j) = \left(\mathrm{L^n(i+1,j) - L^n(i-1,j), L^n(i,j+1) - L^n(i,j-1)}\right) \qquad 5$$

When $\vec{N}$ is propagation direction of L^n then $\vec{N}/\mid \overrightarrow{N} \mid$ is isophote direction that can be calculated through the calculation of the gradient ∇I as:

$$\nabla I^n(i,j) = (I_x^n(i,j), I_y^n(i,j)) \qquad 6$$

∇I gives the direction of the largest spatial change and to obtain the smallest spatial change should be calculating the orthogonal to this vector which is given by:

$$\nabla^{\perp} I^n(i,j) = (-I_y^n(i,j)\ , I_x^n(i,j)) \qquad 7$$

Now, the normalized version of the equation can be given by:

$$\frac{\vec{N}(i,j,n)}{|\vec{N}(i,j,n)|} = \frac{\left(-I_y^n(i,j)\ , I_x^n(i,j)\right)}{\sqrt{(I_x^n(i,j)^2)+(I_y^n(i,j)^2)}} \qquad 8$$

The β^n is projection of $\overrightarrow{\delta L^n}$ onto the normalized vector $\vec{N}$ is given by:

$$\beta^n(i,j)=\overrightarrow{\delta L^n}(i,j).\frac{\vec{N}(i,j,n)}{|\vec{N}(i,j,n)|} \quad 9$$

$|\nabla I^n(i,j))|$ represented the slope-limited version of the norm of the gradient of the image that makes the algorithm more stable and given by:

$$|\nabla I^n(i,j)| = \begin{cases} \sqrt{(I^n_{xbm})^2+(I^n_{xfM})^2+(I^n_{ybm})^2+(I^n_{yfM})^2}, when\ \beta^n > 0 \\ \sqrt{(I^n_{xbM})^2+(I^n_{xfm})^2+(I^n_{ybM})^2+(I^n_{yfm})^2}, when\ ^{\beta^n} < 0 \end{cases} \quad 10$$

The sub-indexes b and f denote the backward and forward differences respectively, while the sub-indexes m denote the minimum while M denoted the maximum between the derivative and zero. This equation is used to prevent too small and too big numbers. The finite difference method has been used to solve the Bertalmio equation and these numerical results will recover the damaged regions in the image.

3. Experimental Results of Bertalmio model (transport)

The behavior of the transport model was studied in both spatial and frequency domains, and its ability was tested to restore the missing regions in color images. This technique has been used to remove the texts and scratches in the damaged natural images. Every color image is considered a set of three images and the method is applied independently for each one. For instance, for inpainting the traditional (R, G, B) color images using the transport model, the image can be decomposed into three scalars (grayscale) images I_R, I_G, I_B, and the inpainting perform on each one of the three images separately. Then, the inpainted color image is obtained by combining the resulting inpainting of I_R, I_G, and I_B. Bertalmio mod Berkeley database in spatial and frequency domains[14] The transport model is applied on four masks with different sizes and types of missing regions such as text and scratches but the size of the mask image should be the same size as the original image. Figure 2 shows the four types of masks that were applied el is applied to 100 natural images selected from the to the images chosen from the Berkeley database[14]. Figure 3 shows the four types of masks that were applied to the images chosen from the Berkeley database[14]. Different quality measures are used to assess the quality such as MSE, PSNR, and SSIM are calculated to determine the quality of the inpainted images.

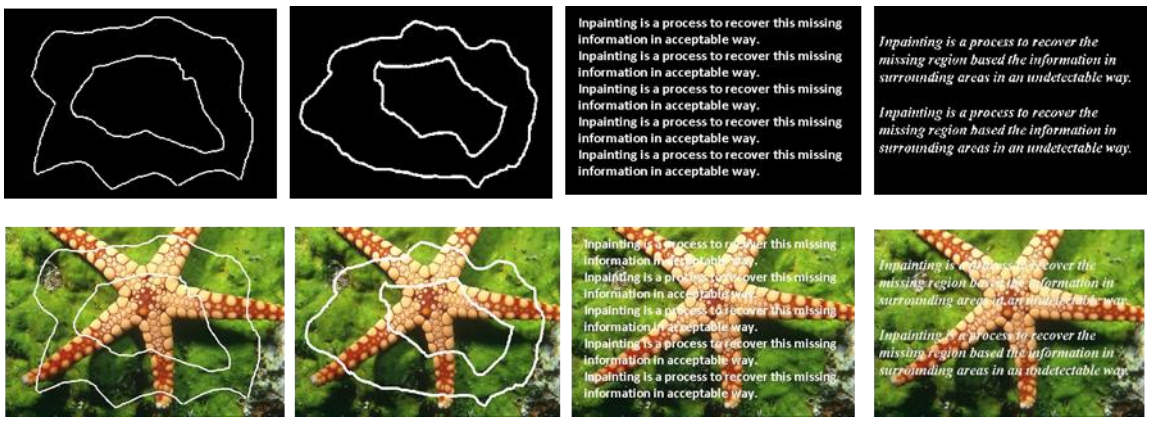

Figure 2: Types of mask and damaged images, first row is binary masked images and second row is the color damaged image

Figure 3: Inpainting based transport equation in spatial domain, column 1 original images, column 2 masked images, column 3 inpainted images.

3.1 Inpainting-based Transport model in the spatial domain

The damaged area is restored by applying the transport model in a special domain according to the following steps:

1. After reading the original (clear) image, the mask will be added to it.
2. Removing scratches and texts from the image that was added using the optimal masks that were added by the user and inpainted area.
3. Calculating the quality of inpainted images using quality assessment measurements. The inpainting process in the special domain is shown in Figure 4.

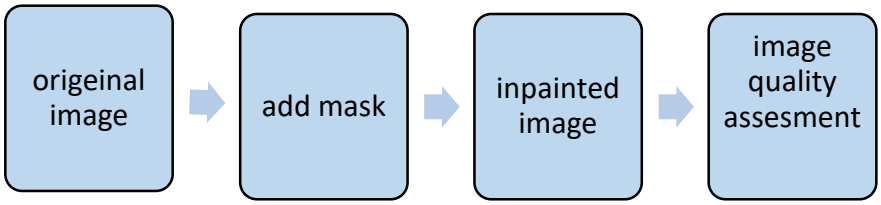

Figure 4: Flow chart of inpainting process in spatial domain

3.2 Inpainting-Based Transport Model in The Frequency Domain

To apply the transport model, the images must first be transformed mathematically by using Discrete Wavelet Transform (DWT) to convert from the spatial domain to the frequency domain. Based on the Haar wavelet function, the original image with the damaged image is segmented into low- and high-frequency components. Every image is decomposed into four sub-bands in the first level of wavelet and on the second level there are seven sub- bands wavelet. The inpainted image in the frequency domain can be represented by the algorithm as follows:

1. The user manually selects the area of the color image.
2. The Haar wavelet function is used to decompose the original image and the damaged image into low-and high-frequency components for 4 sub-bands in the first level of the wavelet for 7 sub-bands in the second level of the wavelet.
3. The transport model has applied to inpaint the missing region in every one of these sub-bands.
4. The inverse wavelet transform is used for the composition of the inpainted image.

Figure 5 shows the flow chart of image inpainting using (WT) in the frequency domain.

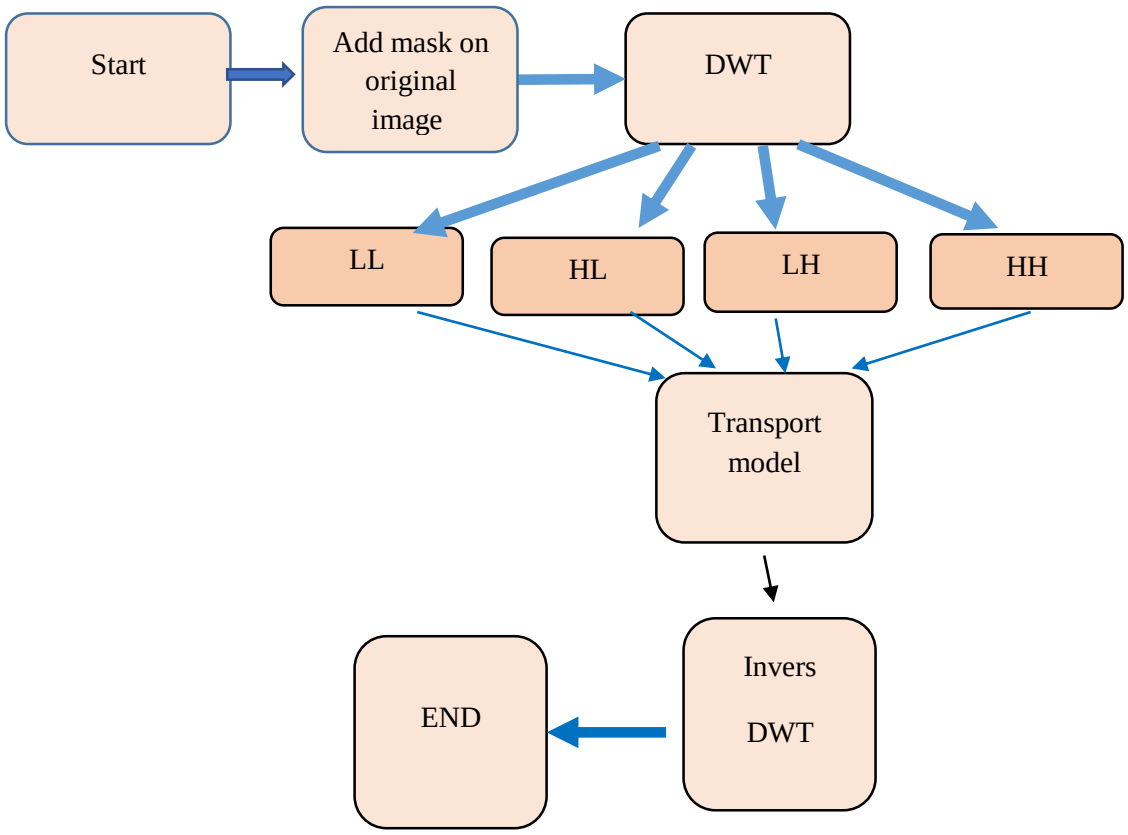

Figure 5.: Flow chart of inpainting process using wavelet transform in frequency domain

Also, the inpainted images recovered in the wavelet domain applied on two levels are shown in Figure 6. Visually, the inpainted images appeared good removing text and scratches. The efficacy of this method for recovering missing regions has been studied by using the first four masks on the set of natural images.

Figure 6: Inpainting based transport equation in frequency domain, column 1 original images, column 2 masked images, column 3 inpainted images in wavelet level 1, column 4 inpainted images in wavelet level 2

The quality of the inpainted areas was verified through statistical measurements such as MSE, PSNR, and SSIM. Therefore, the difference between the inpainted areas and the analogous areas in the source images in each domain was calculated and listed in Table 1. The obtained results confirmed that when the transport equation is applied in the spatial domain, the value of MSE is less than the wavelet domain with level 1 and level 2 for all masks on original images. For example, the MSE value was 103.585 in the spatial domain, while the value was 129.894 in wavelet level 1 and 165.128 in frequency wavelet level 2. Furthermore, the PSNR and SSIM values appeared better when the equation is used in the spatial domain than in both level 1 and level 2 wavelet domain. Also, when mask 2 (thick scratches) is used, the MSE value was 57.394 when applying the transport equation in the spatial domain increased to 74.569 and 97.054 in the level 1 and level 2 domains, respectively.

Moreover, the PSNR values when the model is applied in the spatial domain are larger than those when it is applied in both levels frequency domain. PSNR was 30.948 in the spatial domain while 29.729 in the level 1 wavelet and 28.591in the level 2 wavelet. The same behavior is got when mask 3 (text mask) is used in the spatial domain, where the value of MSE is 69.730 while in the level 1 wavelet domain was 96.257 and 131.996 in the level 2 wavelet. The values of the MSE and PSSR confirmed the transport model better when applied in the spatial domain than when applied in both level 1 and 2 wavelet domains. The MSE and PSSR values confirmed the transport equation better when applied in the spatial domain than when applied in both level 1 and 2 wavelet domains in addition, SSIM values that got with using all masks using the spatial domain are better values when using than frequency wavelet domain for level 1 and level 2

Table 1: Average values for 100 images using transport equation in spatial domain and frequency domain

Mask	Domain	MSE	PSNR	SSIM	TIME (s)
Mask1	Spatial	103.585	28.416	0.990	110.238
	W_level 1	129.894	27.413	0.987	115.204
	W_level 2	165.128	26.304	0.982	205.829
Mask2	Spatial	57.394	30.948	0.992	98.592
	W_level 1	74.569	29.729	0.989	160.441
	W_level 2	97.054	28.591	0.984	199.566
Mask3	Spatial	69.730	30.107	0.984	101.567
	W_level 1	96.257	28.628	0.976	193.816
	W_level 2	131.996	27.180	0.961	208.751
Mask4	Spatial	103.051	28.400	0.983	178.451
	W_level 1	137.455	27.032	0.975	206.012
	W_level 2	160.108	26.123	0.970	223.016

Figure 6: Performance of transport model in spatial domain using four masks at different iterations.

Table 2 listed the results of the effect of changing the attrition. Mask 1 (thin scratch) appeared best results of the measurements when the value of iteration was 200 and the processing time of inpainting was 510.364 sec while the MSE, PSNR, and SSIM values are 90.751, 28.599, and 0.99, respectively. The obtained results are the same for mask 2 (thick scratch) where the iteration value

of 200 gives the best results for the quality values of MSE (46.892), PSNR (31.471), and SSIM (0.993). Mask 3 (thin writing text with attrition of 200 gives. MSE of 51.825, PSNR of 31.03, and SSIM of 0.987. The values of MSE, PSNR, and SSIM are 83.204, 28.965, and 0.985, respectively for mask 4(thick writing text) when the iteration value was 200. Figure 7 shows the the effect of atteration on the MSE, PSNR and SSIM values for various maskes that used.

Table 2: Values of MSE, PSNR, SSIM, and time using different iterations for four types of masks

Masks	Iteration	MSE	PSNR	SSIM	Time(s)
Mask1	10	121.981	27.355	0.989	20.387
	50	93.991	28.528	0.991	96.591
	100	91.123	28.580	0.990	399.215
	150	90.769	28.598	0.991	464.348
	200	90.751	28.599	0.991	510.364
Mask 2	10	61.324	30.377	0.992	20.307
	50	57.843	30.563	0.991	93.372
	100	52.252	30.999	0.991	252.672
	150	49.033	31.279	0.992	367.628
	200	46.892	31.471	0.993	566.132
Mask 3	10	58.808	30.511	0.988	20.597
	50	54.385	30.839	0.989	100.269
	100	51.955	31.015	0.987	243.141
	150	51.860	31.023	0.987	393.928
	200	51.825	31.030	0.987	501.364
Mask 4	10	89.657	28.665	0.984	19.896
	50	84.574	28.907	0.987	94.896
	100	83.644	28.942	0.985	247.835
	150	83.291	28.960	0.985	457.972
	200	83.204	28.965	0.985	501.639

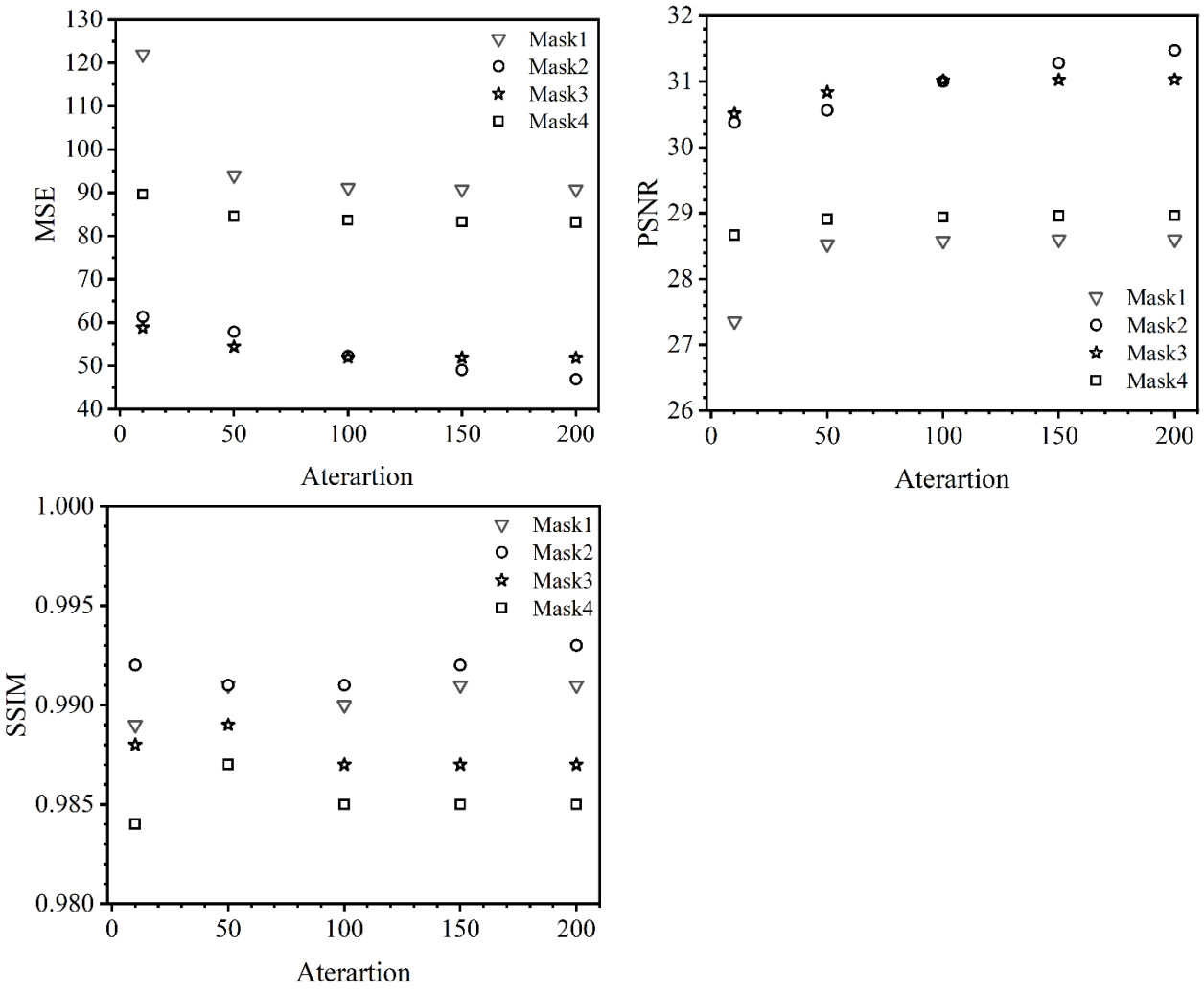

Figure 7: The MSE, PSNR and SSIM values vs. ateration for various maskes that used.

Conclusions

Bertalmio's approach demonstrates superior performance in the spatial domain, particularly for small and localized locations. This superiority is primarily due to the improved PSNR and SSIM coefficients, including MSE, as well as reduced execution time. In the frequency domain (DWT), restoration quality is negatively affected by poor detail resolution in complex features where complexity is evident. DWT is neutral in addressing defects or missing regions. Although DWT offers advantages in multi-resolution image analysis, its results are insufficient for fine-grained work. Therefore, it is the ideal choice when precision and speed are required for restoration. We recommend using DWT only in cases where optimization across broad frequency ranges is required.

References

[1] Z. A. A. Kareem, A. K. Al-Jaberi, A Novel Image Inpainting Technique Based on Isotropic Diffusion, Bas. J. Sci., 40 (2022) 289–305, Doi:10.29072/basjs.20220203

[2] Z. A. A. Kareem, A. K. Al-jaberi, A New Model for Image Inpainting for Different Color Spaces, Al-Rafidain Journal of Computer Sciences and Mathematics (RJCM), 17 (2023) 39-48.

[3] J. Jam, C. Kendrick, K. Walker, V. Drouard, J. G.-S. Hsu, M. H. Yap, A comprehensive review of past and present image inpainting methods, Comput. Vis. image Underst., 203(2021) 103-147, https://doi.org/10.1016/j.cviu.2020.103147

[4] F. Li and T. Zeng, A new algorithm framework for image inpainting in transform domain, SIAM J. Imaging Sci., 9(2016) 24–51, https://doi.org/10.1137/15M1013890

[5] A. Bansal, R. Bajpai, J. P. Saini, Simulation of image enhancement techniques using Matlab, in First Asia International Conference on Modelling & Simulation ,7(2007)296–301, https://doi.org/10.1109/AMS.2007.92

[6] R. C. Gonzalez , R. E. Woods, Digital Image Processing, New Jersey, prentice Hall, (2002)

[7] S. Khidse , M. Nagori, A comparative study of image enhancement techniques, Int. J. Comput. Appl, 81(2013) 28–32, https://doi.org/10.5120/14201-2421

[8] M. Esfehani et al., Forensic Gender Determination by Using Mandibular Morphometric Indices an Iranian Population: A Panoramic Radiographic Cross-Sectional Study, J. Imaging, 9 (2023) 40,https://www.mdpi.com/2313-433X/9/2/40

[9] C. Solomon ,T. Breckon, Fundamentals of Digital Image Processing: A practical approach with examples in Matlab. John Wiley & Sons, (2011).

[10] H. Zhang, S. Dai, Image inpainting based on wavelet decomposition, Procedia Eng., 29(2012) 3674–3678, https://doi.org/10.1016/j.proeng.2012.01.419

[11] L. Debnath , F. A. Shah, Lecture notes on wavelet transforms. Springer, (2017). https://doi.org/10.1007/978-981-10-2104-3

[12] H. Wu, F. He, Y. Duan, X. Yan, B. Fan, Haar-wavelet based texture inpainting for human pose transfer, Inf. Process. Manag., 61,(2024) 103-612, https://doi.org/10.1016/j.ipm.2023.103612

[13] M. Bertalmio, G. Sapiro, V. Caselles, C. Ballester, Image inpainting, in Proceedings of the 27th annual conference on Computer graphics and interactive techniques, (2000) 417–424, https://doi.org/10.1145/344779.344972

[14] P. Arbelaez, C. Fowlkes, and D. Martin, The Berkeley segmentation dataset and benchmark," Comput. Sci. Dep. Berkeley Univ. Available http//www. eecs. berkeley. edu/Research/Projects/CS/vision/bsds, (2007).

إزالة النصوص وطلاء المناطق الغير منتضمة في الصور الرقمية في المجالين المكاني والترددي

زينب علي عبد الكريم [1,*]واحمد كاظم الجابري

1 قسم الرياضيات الكلية التربوية المفتوحة مديرية تربية البصرة..

2قسم الرياضيات كلية التربية للعلوم الصرفة جامعة البصرة

المستخلص

تتناول هذه الدراسة طرق ترميم المناطق المفقودة أو غير المنتظمة الشكل في الصور الرقمية، بالإضافة إلى ازالة النصوص التالفة باستخدام تقنيات مطبقة في كل من المجال المكاني والمجال الترددي. يتميز المجال المكاني بالبساطة وسهولة التنفيذ، مما يجعله مناسبًا لتطبيقات المعالجة الفورية، إلا أنه قد يفتقر إلى المرونة الكافية لمعالجة جميع تفاصيل الصورة بدقة متزامنة. في المقابل، يعتمد المجال الترددي على تحويلات رياضية مثل تحويل فوريه وتحويل المويجات المتقطع(DWT) ، والذي يُعد من الأدوات الأساسية في هذا السياق، لا سيما عند استخدام مويجات .Haar تهدف المقارنة بين الطريقتين إلى توضيح مزايا وقيود كل منهما، وتحديد الخيار الأنسب حسب طبيعة الصورة والمنطقة التالفة.