Color Multi-Focus Image Fusion Method Based on Contourlet Transform

Zhifang Cai
School of Integrated Circuits (Artificial Intelligence), Beijing Polytechnic, Beijing, 100176, China

***Abstract*—Color Multi-Focus Image Fusion (MFIF) technology finds extensive use in areas such as microscopy, astronomy, and multi-scene photography where high-quality and detailed images are vital. This paper presents the Contourlet Transform alongside its enhanced version, the Non-Subsampled Contourlet Transform (NSCT), aimed at improving the outcomes of image fusion, with the support of Laplacian Pyramid (LP) decomposition. The NSCT framework overcomes challenges like spectral aliasing and directional sensitivity, leading to images with sharper edges, enriched texture details, and preserved delicate information. Experimental findings highlight the NSCT-based fusion algorithm's superiority. Subjective assessments indicate that using the NSCT method results in images with sharp and well-defined object boundaries, outstanding contrast, and abundant textures without the creation of artifacts, markedly excelling beyond traditional techniques such as the Contourlet Transform, Non-Subsampled Shearlet Transform (NSST) and Rolling Guidance Filtering (RGF). Objective measures verify its effectiveness: In the first dataset, it attains an average gradient (AG) of 8.36 and an edge intensity (EI) of 3.29E-04, while in the second dataset, it reports an AG of 21.39 and an EI of 4.06E-04, significantly outperforming other methods. Moreover, the NSCT method offers competitive computational speed, balancing runtime with high-quality fusion performance. These results establish the proposed method as a powerful and efficient solution for color MFIF, offering notable performance benefits and practical utility in various imaging fields.**

Keywords*—*Contourlet transform; image fusing; NSCT; Laplacian Pyramid; color multi-focus

I. Introduction

Optical lenses are vital for photography, visual perception, and computational tasks in real-world scenarios. Despite their pivotal role in imaging, the core principles of lens operation inherently restrict their ability to achieve full scene sharpness. The optical features of cameras enable precise focus on selected subjects within the depth of field, which consequently causes other parts to remain blurred and unfocused. This limitation poses significant challenges in producing a single image with a uniform focus across the entire scene [1-2]. These constraints are particularly crucial in fields where precision and clarity are essential, such as medical imaging, microscopic analysis, and astronomy.

Color Multi-Focus Image Fusion (MFIF) technology emerges as an essential solution to address these challenges. MFIF combines multiple images with varying focus zones of a scene into a single entirely focused image. This technique utilizes advanced fusion algorithms [3], enhancing image clarity and usefulness by amalgamating information from different source images, thus enabling comprehensive image analysis. MFIF finds widespread application across diverse fields, such as merging textile fiber images in microscopy, integrating astronomical images from telescopes, and combining multi-focus images from consumer cameras [4].

The need for accurate and comprehensive image representation is growing across various sectors, where extracting precise information from visual data is critical. Traditional image fusion methods, including pixel-based, feature-based, and transform-based approaches, have substantially progressed in fulfilling this need. However, challenges remain, like texture detail loss, artifacts, and blurring issues. Transform-based techniques, such as wavelets, show promise but are limited in effectively handling multi-resolution and directional data.

The Contourlet Transform offers a multi-resolution and multi-directional framework, which helps overcome some limitations by better representing edges and textures. However, its reliance on subsampling may introduce artifacts and limit fusion quality. The Non-Subsampled Contourlet Transform (NSCT) is implemented for its superior abilities to the standard Contourlet Transform. By removing subsampling, NSCT enhances the preservation of texture and detail, making it highly appropriate for MFIF applications. The practical usage of NSCT covers medical imaging, satellite imagery, and surveillance applications.

Despite advancements in MFIF, several unresolved issues still exist. Existing approaches often struggle with clarity, failing to preserve detailed textures and sharp edges, which results in fused images that are either blurred or prone to artifacts. Additionally, these algorithms have limited adaptability to various imaging scenarios, such as microscopy, astronomy, and general photography, hindering their practical application. Quantitative assessment indicators like gradient and edge strength often show suboptimal performance, underscoring the need for better techniques. Furthermore, the complexity of sophisticated fusion methods frequently reduces computational efficiency, making them unsuitable for real-time or large-scale use. These challenges underscore the necessity for innovative approaches to improve fusion quality, enhance generalizability, and maintain computational viability. The key contributions of the article are:

- The enhanced transform framework integrates the NSCT into the fusion mechanism, showing superior texture and edge conservation compared to conventional methods.
- Improved fusion algorithm displaying impressive

scalability and flexibility in diverse imaging contexts, outperforming leading techniques in both subjective and objective evaluations.

- A comprehensive validation process through experimental methods highlights significant improvements in gradient and edge strength metrics.
- Addressing the limitations of traditional techniques to effectively cater to various applications, such as microscopy, astronomy, and general photography.

The rest of the article is organized as follows: Section II begins with an in-depth review of MFIF technology, followed by an introduction to the core principles of the Contourlet Transform and its advanced form, the NSCT. Section III thoroughly discusses the proposed fusion algorithm, detailing its design and how NSCT is integrated into the image fusion process. Section IV offers experimental validation, featuring both qualitative and quantitative assessments to demonstrate the improved performance of the proposed method. Finally, Section V summarizes the study by synthesizing the outcomes, addressing the limitations of the current approach, and suggesting future research directions to advance MFIF.

II. Related Works

The evolution of image processing techniques, notably the contour transform and MFIF, has driven significant progress in various fields. This section summarizes these areas' key contributions and advancements, highlighting their applications, improvements, and integration into new solutions. The most prominent inferences from the literature are tabulated in Table I.

A. Contourlet Transform

Renowned for its effectiveness in capturing edge and texture details, the contour transform finds numerous applications across multiple sectors. Jahangir H et al. developed an innovative approach using the contour transform to evaluate damage in prestressed concrete slab structures. Applying the transform to the data on the modal curvature in damaged and unscathed states led to the precise determination of the severity of structural damage [5]. Hasan M. M. et al. contour improved the application by combining the contourlet transform with fine-tuning of the pretrained VGG19 model, improving the identification of gastrointestinal polyps in endoscopic images and achieving greater diagnostic precision [8]. Enhancements to the Contourlet contour transform, particularly the development of the NSCT, have notably expanded its applications. Li W et al. merged NSCT with Convolutional Neural Networks for multimodal medical image fusion, successfully addressing intensity inconsistencies and achieving better fusion performance [6]. Kollem et al. presented an improved Total Variation model to enhance grayscale and color brain tumor images. Using NSCT with adaptive threshold techniques improved the preservation of contour and texture features [7]. Hu F's team also used the NSCT and Schur decomposition to create an excellent blind watermarking system to protect the copyright of color digital images, which led to higher peak signal-to-noise ratios and structural similarity indices [9].

B. Multi-Focus Image Fusion (MFIF)

Significant progress has been made in MFIF across various domains, like medical imaging and robotic vision, where image clarity and precision are crucial. Li H addressed the challenge of detecting focused regions in MFIF by proposing a fusion approach relying on focus feature extraction. This approach successfully preserved source images' intricate textures and edge structures while effectively identifying focused area edges [10]. Panigrahy C et al. introduced an innovative MFIF classification method and reviewed current fusion approaches, demonstrating promising performance and potential applications [11]. Recent innovations in MFIF methods include wavelet-based techniques and incorporating deep learning. Wang Y developed a translation-invariant approach using the wavelet transform and fractal dimensions to evaluate clarity, integrating it with the Otsu thresholding to fuse detail coefficients, and achieved competitive results in subjective and objective evaluations [12]. Liu et al. [13] designed an MFIF method using deep learning, employing the VGG-19 network for feature extraction alongside deconvolution modules to generate decision maps. This method indicated superior performance over other segmentation accuracy and fusion metrics approaches, especially for focused and non-focused regions in color multi-focus images.

Detailed contour transform and MFIF studies have yielded promising results across diverse applications. Integrating NSCT into MFIF is pivotal, offering substantial potential to improve fusion quality and precision. This integration overcomes existing challenges and paves the way for developing improved techniques.

Upon examining the cited research, we found that numerous current techniques face drawbacks despite substantial progress in multi-focus image fusion methods. In particular, conventional methods like NSST and RGF encounter difficulties in maintaining intricate texture details and ensuring consistency across different image scenarios. Additionally, achieving computational efficiency poses challenges in high-resolution imaging, as current fusion techniques often grapple with balancing performance against processing speed. Moreover, many reviewed methods display issues such as spectral aliasing and diminished directional sensitivity, resulting in inferior image fusion quality.

In response to the identified gaps, our research introduces a fusion algorithm based on NSCT, augmented with Laplacian Pyramid decomposition, to overcome the mentioned limitations effectively. This method enhances directional selectivity, reduces spectral aliasing, and improves edge preservation while retaining computational efficiency. We achieve superior fusion quality by utilizing NSCT's capability to capture features in multiple directions and combining it with Laplacian Pyramid decomposition, resulting in sharper edges and finer texture details. These advancements directly tackle the issues noted in the existing literature, offering a more reliable and efficient solution for color multi-focus image fusion.

TABLE I SUMMARY OF RELATED WORKS ON IMAGE FUSION TECHNIQUES, MODELS, CONTRIBUTIONS, AND CHALLENGES

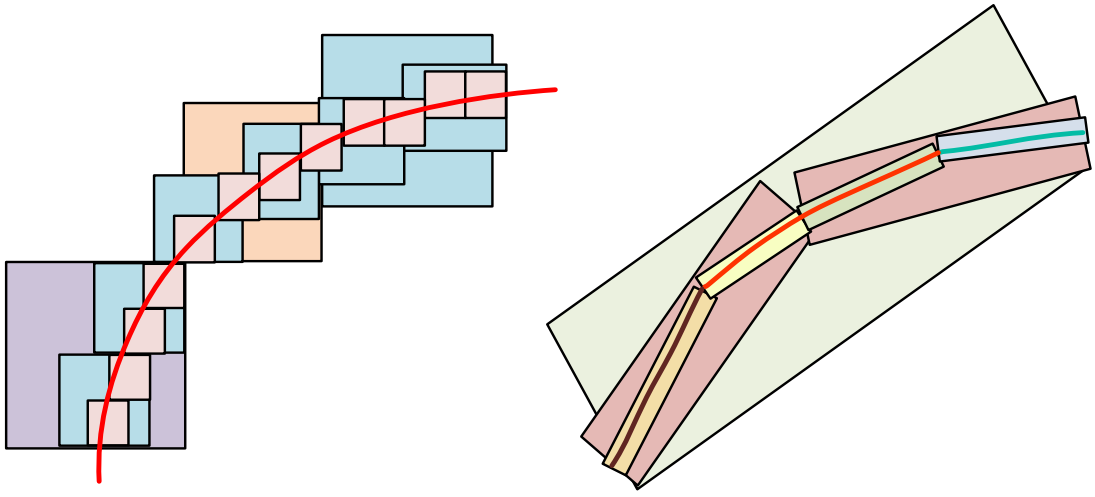
Author	Year	Model	Key Contribution	Challenges
Jahangir H et al. [5]	2021	Contourlet Transform	Damage evaluation in prestressed concrete slabs using modal curvature data.	Application limited to structural damage; no focus on multifocus image fusion.
Li W et al. [6]	2020	NSCT + CNN	Enhanced multimodal medical image fusion addressing intensity differences.	High computational complexity due to CNN integration.
Kollem S et al. [7]	2022	NSCT + Adaptive Threshold	Enhanced grayscale and color brain tumor images using partial differential equations.	Limited to specific medical imaging applications; generalizability not tested.
Hasan M M et al. [8]	2019	Contourlet + Fine-tuned VGG19	Improved identification of gastrointestinal polyps in endoscopic images.	Reliance on pre-trained models may restrict adaptability to other domains.
Hu F et al. [9]	2020	NSCT + Schur Decomposition	Robust blind watermarking scheme for copyright protection of color digital images.	It focuses on copyright protection; there is no direct applicability to MFIF.
Li H [10]	2018	Focus Feature Extraction	Accurate detection of focused regions in MFIF.	They may struggle with complex textures or multi-scene images.
Panigrahy C et al. [11]	2017	Classification-Based Fusion	Novel classification method for MFIF with broad application prospects.	Performance heavily dependent on classifier accuracy.
Wang Y et al. [12]	2016	Wavelet Transform + Fractal Dimension	Translation-invariant method for MFIF with effective edge fusion.	Lacks directional sensitivity and detail preservation.
Liu S et al. [13]	2019	Deep Learning (VGG19 + Deconvolution)	Enhanced segmentation and fusion for color multi-focus images.	High computational cost; potential overfitting with limited data.

III. COLOR MFIF SUPPORTED BY CONTOURLET TRANSFORM

This section explores the application of the Contourlet Transform for Color MFIF. The study delves into the basic concepts, construction techniques, and transformation processes linked to the Contourlet Transform. Additionally, it assesses its shortcomings and presents the NSCT as a possible improvement in this domain. The research outlines the elements, decomposition principles, and operational procedures of NSCT in the context of color MFIF, highlighting its enhanced performance features.

A. Contourlet Transform Method

The Contourlet Transform is an effective technique for image processing in the realm of color MFIF. It stands out due to its multi-scale and multi-directional decomposition capabilities, which enhance the depiction of image contours and textures. The procedure begins with a Laplacian Pyramid (LP) decomposition, systematically dividing the image into sub-images at various scales [14-15]. The high-frequency components undergo directional decomposition using the Directional Filter Bank (DFB), while the low-frequency components are iteratively decomposed, creating the complete contour wave structure. Unlike conventional wavelet transforms that employ square wavelet functions, the Contourlet Transform uses rectangular ones, allowing for aspect ratio adjustments and better adaptation to image contours across scales. This improved directionality offers a more accurate representation of curves and contours. Fig. 1 shows a visual comparison of different transformation methods for representing image curves.



(a) Description of curves using wavelet transform

(b) Countourlet transformation for describing curves

Fig. 1. Description of image curves by different transformation methods.

Contourlet transform comprises two main components: LP decomposition and DFB filtering [16]. LP Decomposition captures intricate details across a spectrum of frequencies by deconstructing the image into smaller elements. DFB Filtering provides a multi-directional decomposition method that extracts directional characteristics. The LP decomposition refines the Gaussian pyramid approach by reducing high-frequency information loss by subtracting layers. This involves interpolation to align image dimensions at each scale level. This process is achieved through repeated convolution and down-sampling operations. The Gaussian pyramid is the fundamental type of image pyramid, generating a pyramid-shaped data structure through convolution and down-sampling operations. LP further develops from the Gaussian pyramid, addressing the loss of high-frequency information during computation. LP achieves this by subtracting each layer of the Gaussian pyramid from its previous layer, obtaining the LP decomposed image. In this process, interpolation operations are required for rows and columns, extending fine-sized images to coarse-sized ones. The output representation of the fine-sized image is expressed as shown in Eq. (1).

$$\sum_{m=-2}^{2}\sum_{n=-2}^{2} G_{l-1}(2i+m, 2j+n)w(m,n) = G_l(i,j) \quad (1)$$

In Eq. (1), l represents the decomposition level and $G_l(i,j)$ represents the Gaussian pyramid image at the level l. $w(m,n)$ stands for the sliding window with a size of 5 × 5. i and j are the symbols of rows and columns of the corresponding image. Subsampling operations can obtain a sequence of Gaussian pyramid images, followed by interpolation and dilation operations to make the sizes of images Gl and Gl-1 consistent, as it described in Eq. (2).

$$4\sum_{m=-2}^{2}\sum_{n=-2}^{2} G'_{l-1}\left(\frac{i+m}{2}, \frac{j+n}{2}\right)w(m,n) = G_l * (i,j) \quad (2)$$

In Eq. (2), the calculation of $G'_{l-1}\left(\frac{i+m}{2}, \frac{j+n}{2}\right)$ is shown in Eq. (3).

$$G'_{l-1}\left(\frac{i+m}{2}, \frac{j+n}{2}\right) = \begin{cases} G_l\left(\frac{i+m}{2}, \frac{j+n}{2}\right), & \frac{i+m}{2} \text{ and } \frac{j+n}{2} \text{ are integers} \\ 0, & \text{other} \end{cases} \quad (3)$$

LP is defined as shown in Eq. (4).

$$\begin{cases} LP_l = G_l - G_{l+1}^*, & 0 < l < N \\ LP_N = G_N, & l = N \end{cases} \quad (4)$$

In Eq. (4), N represents the highest level of the pyramid. After obtaining the LP image sequence, fusion is performed according to the selected rules to get a fused pyramid sequence image, and finally, a reconstruction process is carried out. The calculation for reconstruction is shown in Eq. (5).

$$\begin{cases} G_N = LP_N, & when\ l = N\ 0 < l < N \\ G_l = LP_N + G_{l+1}^*, & when\ 0 < l < N \end{cases} \quad (5)$$

The LP decomposition and reconstruction processes, with a schematic diagram, are shown in Fig. 2.

The DFB is a tree-structured decomposition method that generates multiple directional subbands at each layer, with each subband taking a wedge shape. A construction method for DFB based on a sector filter group is introduced to simplify directional decomposition without altering the input signal. This method comprises two main components: a dual-channel sector filter group and shear processing. The dual-channel sector filter group utilizes sector filters to decompose the two-dimensional spectrum into horizontal and vertical directions. On the other hand, shear processing involves sampling and rearranging the image. The method effectively captures frequency domain information across various directions while ensuring complete reconstruction by incorporating shear and inverse shear operations before and after the dual-channel filter group.

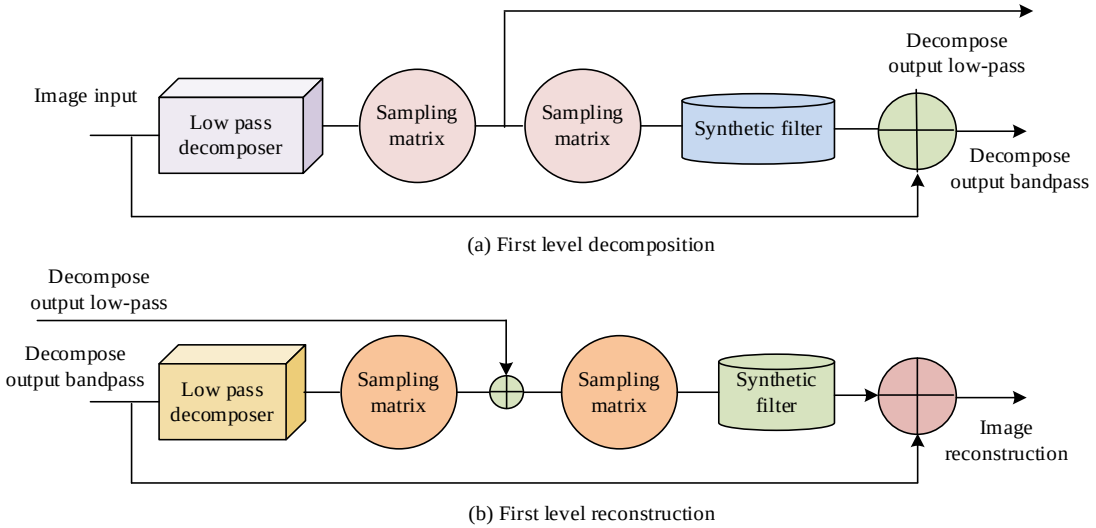


Fig. 2. A schematic diagram of the decomposition and reconstruction of the LP.

For instance, in a 3-level decomposition, the method can produce 8 directional spectrum partitions. Subbands labeled 1, 2, and 3 belong to the first channel, while subbands labeled 4, 5, 6, and 7 are part of the second channel. Integrating the Laplacian Pyramid (LP) with DFB forms a two-layer filter structure, referred to as the Contourlet filter group. During transformation, the image is first decomposed into high-frequency and low-frequency components. The low-frequency component undergoes further multi-resolution analysis, while the high-frequency components are processed through DFB to generate directional subbands at various orientations. This iterative process gradually connects singular points into linear structures, capturing the edge contours of the image. This capability allows the Contourlet Transform to effectively capture image features, such as edges and textures, offering a superior representation compared to traditional methods. Fig. 3 depicts the flowchart of the Contourlet transformation, illustrating the sequential decomposition and feature extraction process.

To improve NSCT-based fusion, images must be pre-processed to make them more transparent and less distorted before transforming. Preprocessing improves contrast and gets rid of unnecessary variations. This helps NSCT retain structural details better, which makes it perfect for medical imaging and analyzing satellite data. It also makes extracting more accurate features and fusion easier, which improves visualization in critical areas like medical imaging and remote sensing.

B. MFIF Supported by Improved Contourlet

Within the Contourlet framework, downsampling is employed during integrating pyramid and directional filtering decomposition, potentially resulting in spectral aliasing, which diminishes the transformation's directional understanding [17]. The Non-Subsampled Contourlet Transform (NSCT) was developed to address this issue. NSCT's decomposition method includes two primary elements: the Non-Subsampled Pyramid Filter Bank (NSPFB) for decomposing scales and the Non-Subsampled Direction Filter Bank (NSDFB) for analyzing directions. This framework maintains the Contourlet Transform's ability to effectively capture curves while removing spectral aliasing that arises from downsampling. Consequently, NSCT improves the accuracy and dependability of directional sensitivity in image processing. Fig. 4 illustrates the NSCT's structural diagram, detailing its components and workflow.

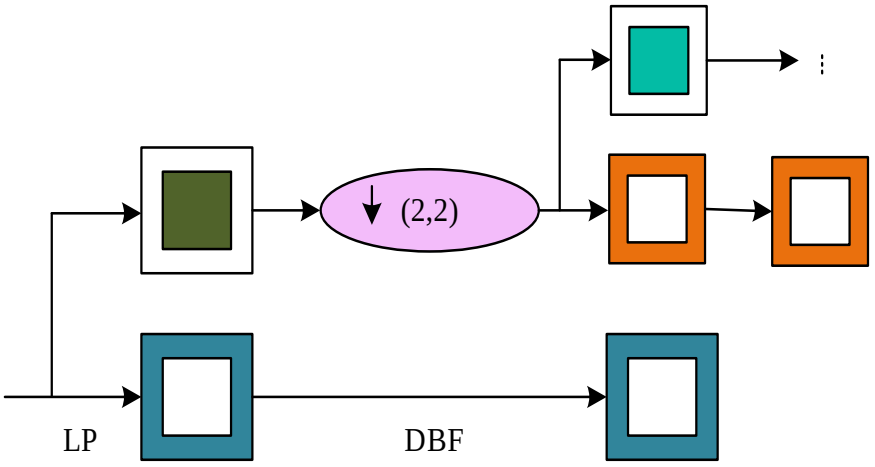


Fig. 3. Flow chart of Contourlet transformation.

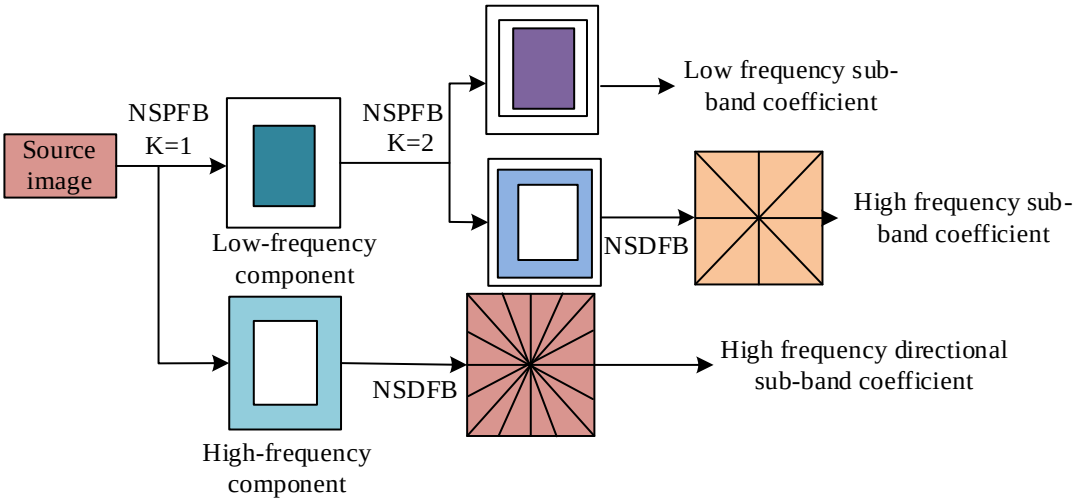


Fig. 4. Structural diagram of the NSCT.

Conventional approaches for assessing image sharpness frequently depend on individual features, which may be inadequate in specific situations [18-19]. To mitigate these limitations, we propose an MFIF technique utilizing the NSCT. The algorithm initiates with the decomposition of two source images through NSCT, extracting their subband coefficients. This results in multiple subbands of uniform size, aiding the identification and exploitation of gray feature correlations across subbands. In the NSCT realm, reference coefficients' sibling subbands contain abundant correlation data, which is harnessed to enhance fusion. The method employs correlation weights from sibling subbands in conjunction with average gradient metrics for a robust fusion process for high-frequency subbands. Initial fusion results emerge from applying these rules, followed by an inverse NSCT transform. To further boost the fusion's quality, edge detection methods are utilized on the preliminary result, isolating edge contour data. This data is then superimposed on the preliminary fused result, enhancing clarity and feature retention. The NSCT-supported fusion algorithm's comprehensive framework is depicted in Fig. 5, showcasing its efficacy in capturing and integrating edge and texture details for superior image fusion quality.

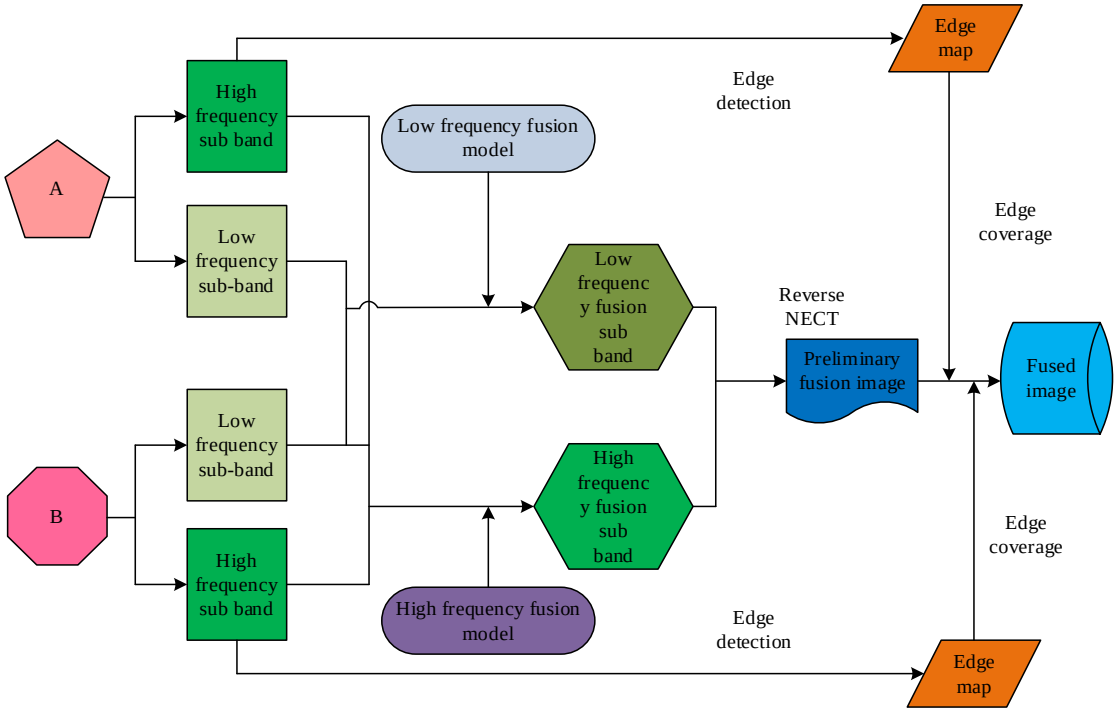


Fig. 5. A fusion algorithm framework based on NSCT.

In NSCT decomposition, the low-frequency subband is extracted from the source image via successive application of low-pass filters over several layers. This research adopts statistical features like local information entropy and enhanced local Laplacian energy for fusing low-frequency coefficients. Information entropy is a reliable indicator for assessing the amount of information in an image. The low-frequency subband often contains a significant part of the image's information, making entropy an effective measure for fusion [20]. Conversely, Laplacian energy indicates the extent of grayscale variation in an image, enhancing the analysis based on entropy. Eq. (6) provides the mathematical representation for local information entropy and local Laplacian energy at point *h* in the image *x*. These metrics facilitate the fusion of low-frequency coefficients by capturing essential informational and structural distinctions within the image.

$$G_{r,x}(h)=\sum_{q\in Q}\omega(q)S_{r,x}(q)^2 \tag{6}$$

In Eq. (6), $h(a,b)$ represents pixel coordinates, $\omega(q)$ is a Gaussian weight matrix, Q denotes the neighborhood centered around h, and r takes values of 1 or 2. The representation of local information entropy at point q within the neighborhood Q is shown in Eq. (7).

$$S_{1,x}(q)=\sum_{a=0}^{\gamma-1}p_a\log(p_a) \tag{7}$$

In Eq. (7), p_i represents the proportion of pixels with a grayscale value a in the neighborhood Q, and γ represents the grayscale level. The calculation of locally improved Laplacian energy is given by Eq. (8).

$$S_{1,x}(a,b)=\left|2C_x(a,b)-C_x(a-1,b)-C_x(a+1,b)\right|+\left|2C_x(a,b)-C_x(a,b-1)-C_x(a,b+1)\right| \tag{8}$$

In Eq. (8), $L_r(h)$ represents the low-frequency subband coefficients obtained from the decomposition of the image x. The feature matching degree of low-frequency component coefficients between images A and B at point h is expressed in Eq. (9).

$$L_r(h)=\frac{2\sum_{q\in R}\omega(q)\left|S_{r,A}(q)\right|\left|S_{r,B}(q)\right|}{G_{r,A}(h)+G_{r,B}(h)} \quad (9)$$

Assuming α is the matching threshold, if $L_r(h)<\alpha$, the feature weight matrix is solved through the calculating process of Eq. (10).

$$W_{r,A}=\begin{cases}1, G_{r,A}(h)\geq G_{r,B}(h)\\ 0, G_{r,A}(h)<G_{r,B}(h)\end{cases} \quad (10)$$

If $L_r(h)\geq\alpha$, the feature weight matrix is solved through the calculating process of Eq. (11).

$$\begin{cases}W_{r,A}=\begin{cases}1-(L_r(h)-\alpha), G_{r,A}(h)\geq G_{r,B}(h)\\ L_r(h)-\alpha, G_{r,A}(h)<G_{r,B}(h)\end{cases}\\ W_{r,B}=1-W_{r,A}(h)\end{cases} \quad (11)$$

Eq. (12) shows how the fusion weight matrix for low-frequency components is obtained by applying the average principle to weight matrices.

$$\begin{cases}W_A(h)=(W_{1,A}(h)W_{2,A}(h))/2\\ W_B(h)=(W_{1,B}(h)W_{2,B}(h))/2\end{cases} \quad (12)$$

The calculation of the low-frequency coefficients after merging images A and B is given in Eq. (13).

$$C_F(h)=C_A(h)W_A(h)+C_B(h)W_B(h) \quad (13)$$

High-frequency elements contain the original image's fine details and textural information, while the average gradient quantitatively describes variations in the image's gradient. A higher average gradient indicates more detailed texture and sharper image clarity. To improve the fusion of high-frequency elements, this research uses the local average gradient as a basis for determining fusion weights. This method assures that texture-rich areas are effectively maintained and emphasized in the combined image. The mathematical expression for the region's average gradient concerning the high-frequency subband coefficients at the *p-th* scale and *s-th* direction of image *x* is provided in Eq. (14). This formulation highlights the importance of gradient-based metrics in enhancing the quality of fusion for high-frequency details.

$$T^s_{s,k}(a,b)=\frac{\sum_{(a,b)\in Q}g(a,b)}{|Q|} \quad (14)$$

$|Q|$ represents the size of the neighborhood window. The calculation of $g(a,b)$ is given in Eq. (15).

$$g(a,b)=\frac{\sqrt{\left[C^s_{x,q}(a,b)-C^s_{x,q}(a+1,b)\right]^2+\left[C^t_{x,q}(a,b)-C^t_{x,q}(a+1,b)\right]^2}}{2} \quad (15)$$

In Eq. (15), $C^s_{x,q}$ and $C^t_{x,q}$ represent the high-frequency subband coefficients of image *x* in the *p-th* scale and *s-th*, *t-th* directions. Within the NSCT domain, the reference node is closely associated with its eight adjacent nodes in the identical subband, as these neighbors carry considerable information about the reference node. At the same scale, nodes situated in the same spatial position but in different directional subbands are known as sibling nodes. Although these siblings provide valuable data, they typically offer less insight regarding the reference node than the immediate neighbors. Conversely, the node at the same spatial location in the previous scale, termed the parent node, holds the least amount of relevant information about the reference node. Understanding these nodes' hierarchical and spatial connections is crucial for capturing structural and contextual information in NSCT-based image processing. Fig. 6 comprehensively depicts these relationships, illustrating the interaction between coefficients in the NSCT domain.

The NSCT-based fusion technique balances computational efficiency with feature extraction by selecting appropriate decomposition levels and fusion strategies. Researchers have tested NSCT in challenging situations and found it keeps image details and edge structures better than other methods, even when the input images are distorted. NSCT achieves this robustness by enhancing directional information, reducing undesirable artifacts, and guaranteeing high-quality fused images under diverse conditions.

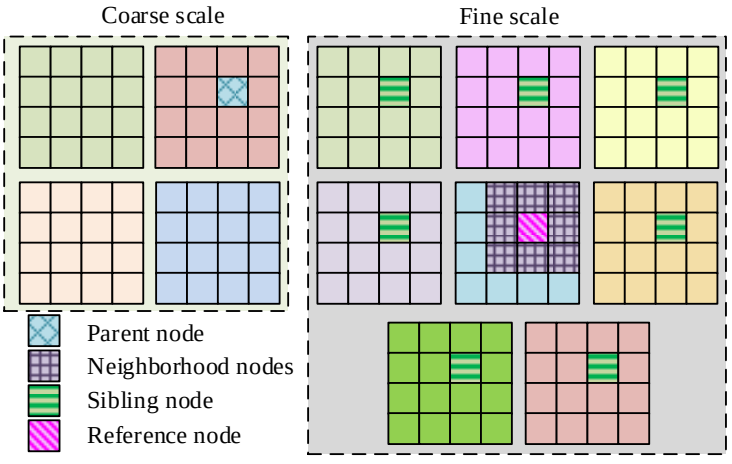


Fig. 6. A schematic representation of the relationship between the NSCT domain coefficients.

IV. Results and Discussion

This section thoroughly assesses the proposed Color MFIF algorithm, which leverages the Contourlet Transform and focuses on the NSCT. Experimental validation encompasses subjective and objective analyses demonstrating the algorithm's practical effectiveness and comparative benefits.

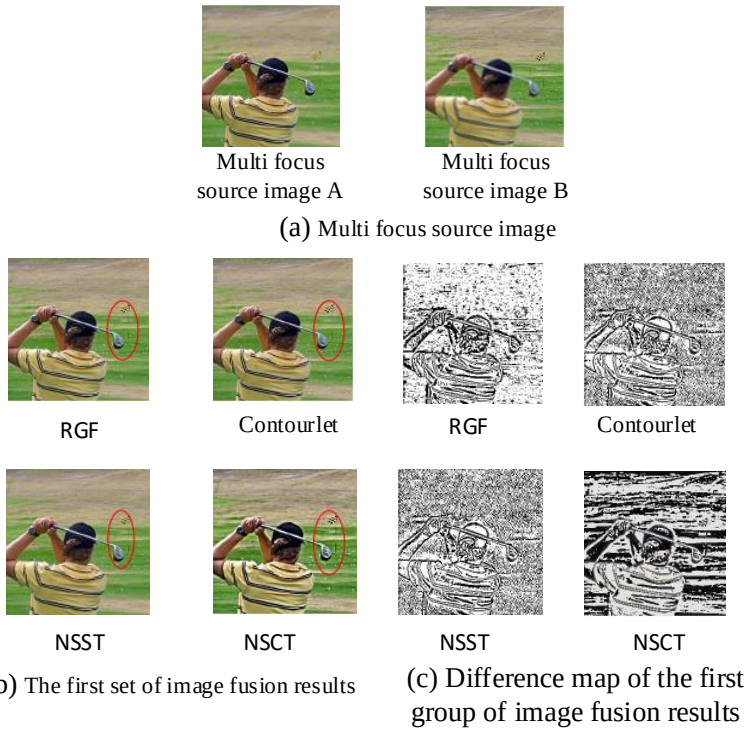
A. Subjective Evaluation

Subjective evaluations were conducted based on qualitative assessments to measure the practicality of the NSCT-based multi-focus image fusion algorithm. Testing utilized two datasets from a registered true-color multi-focus image collection. The fusion outputs from the NSCT-based algorithm were compared to results from conventional methods, such as the Contourlet Transform, Non-Subsampled Shearlet Transform (NSST), and Rolling Guidance Filtering (RGF) techniques. Table II lists the parameters used in these experiments, ensuring uniformity and comparability among all methods tested.

TABLE II Experimental Parameters

Parameter	Set value
Input picture A	Clear
Input picture B	Blur
Spatial filter	Gaussian filter with a standard deviation of 1.0
Gradient weight	0.5
Iterations	10
Output image size	Consistent with the input image
Output image format	PNG

The image showcasing the fusion effect for the initial set of multifocus source images featuring a clock, along with its corresponding difference image, is illustrated in Fig. 7. Fig. 7(a) presents color multifocus source images A and B, while Figure 7(b) shows the resultant fusion effect image. Fig. 7(c) provides the associated difference image. Observations from Fig. 7 reveal that the RGF method delivers a solid overall fusion effect, producing a visually appealing result. However, examining the highlighted area in the red circle reveals noticeable blurring in the flag section with the RGF method. The Contourlet Transform approach achieves a generally good fusion effect but suffers from some detail loss and certain artifact presence in the hand and back areas. NSST also performs well in fusion, though it introduces slight blurring at the flag's edge within the red box, resulting in minor detail information loss and slightly lowering the clarity of the fusion effect. Contrastingly, with the NSCT method, object edges are distinctly sharp and clear, with abundant texture details and no detail loss. This fusion effect is optimal, featuring strong contrast effects, indicating that the proposed method in this study provides superior fusion effects, free of artifacts in surrounding areas, demonstrating high authenticity and precision.



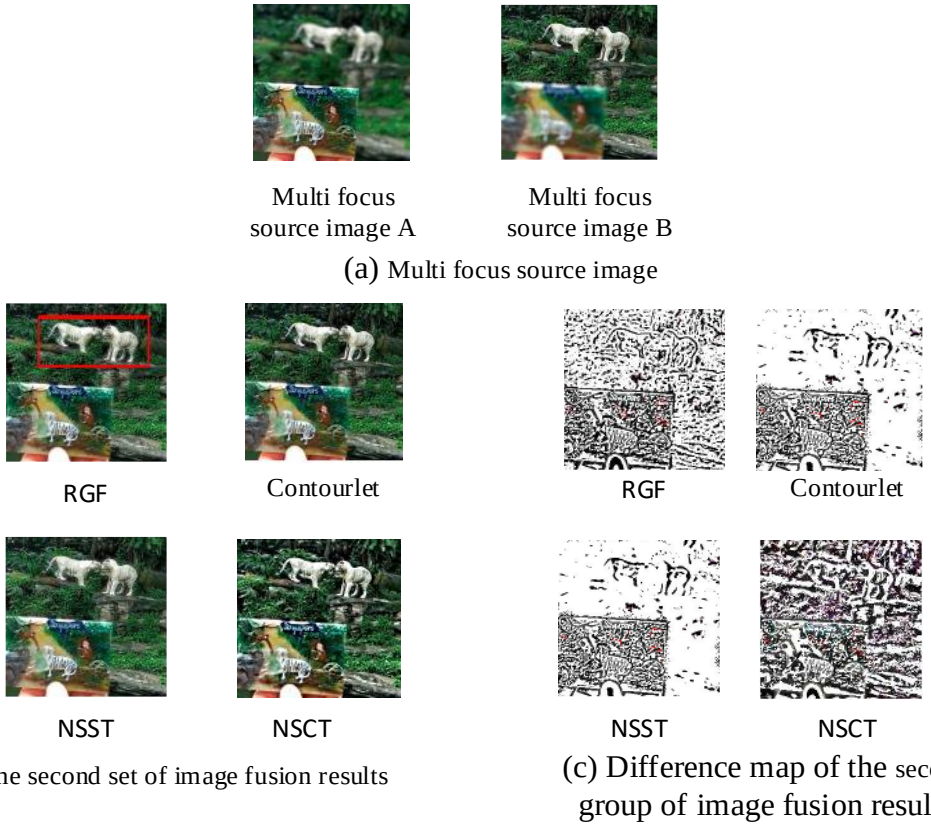
(a) Multi focus source image

(b) The first set of image fusion results

(c) Difference map of the first group of image fusion results

Fig. 7. The first set of image fusion results and difference maps.

The images resulting from the fusion process and the corresponding difference maps related to the clock in the second set are illustrated in Fig. 8. Fig. 8(a) presents the colored multi-focus source images A and B from the second set. Fig. 8(b) shows the fusion effect image of the second set, and Fig. 8(c) presents the associated difference map. From Fig. 8, it can be noted that, within the second set of experimental images, the RGF method displays a slight reduction in clarity, as seen through subjective analysis. Additionally, the RGF method causes subtle blurring and pseudo-artifacts near the tiger. The Contourlet transform method achieves a good overall fusion effect, though with a minor drop in clarity. Even though the NSST transform method offers rich visual effects and abundant information, it suffers from a decline in clarity and produces unnatural transitions and pseudo-artifacts in the peripherals. In contrast, the NSCT method significantly excels, presenting rich information, clearly viewing scene textures, and uniform brightness distribution. It does not generate pseudo-artifacts, effectively addressing detail loss due to edge blurring. Moreover, there are no pseudo-artifacts in the tiger's texture information, showcasing high contrast. This suggests that the color MFIF algorithm proposed by the research institute delivers a notable fusion effect.



(a) Multi focus source image

(b) The second set of image fusion results

(c) Difference map of the second group of image fusion results

Fig. 8. The second set of image fusion results and difference maps.

B. Objective Evaluation

The study continued its analysis of the fusion result images for both groups by focusing on quantitative evaluations. The metrics used for assessment included Average Gradient (AG), Standard Deviation (STD), Spatial Frequency (SF), Visual Information Fidelity (VIFF), and Edge Intensity (EI). AG assessed contrast details and texture changes in the images, with higher values indicating more precise information. STD measured the range of grayscale levels, where higher values signaled a more natural brightness in the fused image. SF captured the ratio of spatial frequency errors, and VIFF evaluated image quality based on its visual fidelity, with higher values reflecting better fusion quality. EI gauged image clarity. Initially, the clarity of each fusion technique was validated through AG and EI metrics. To confirm experiment reliability, 10 tests were conducted, and the averages were calculated as final results. Fig. 9 illustrates the AG and EI values of the fusion outcomes for both groups. In Fig. 9(a), particularly in the first group's multi-focus source images fusion, the NSCT method achieved a substantial AG of 8.36, surpassing the traditional contourlet transform, NSST, and RGF methods by 0.16, 0.72, and 0.25, respectively. Fig. 9(b) indicates that the NSCT method also reached a notable EI value of 3.29E-04 in the first group, showing enhancements over the other methods by 0.48E-04, 0.59E-04, and 0.72E-04. Fig. 9(c) demonstrates that the NSCT method obtained a high AG of 21.39 in the second group, outperforming other methods with improvements of 7.28, 7.39, and 6.14. Fig. 9(d) portrays the NSCT method achieving a significant EI value of 4.06E-04 in the second group, markedly better than the alternatives, illustrating the remarkable efficacy of the proposed image fusion technique developed by the research institute.

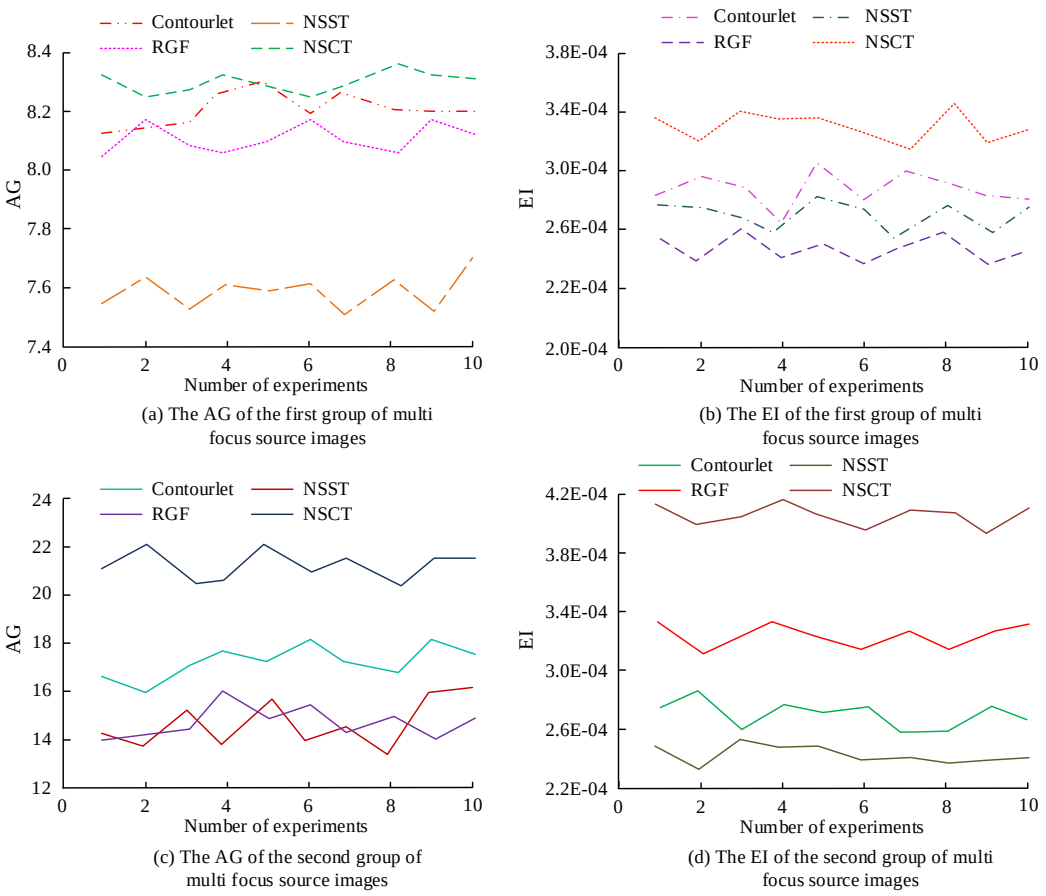


Fig. 9. AG and EI values of the fusion results as follows.

The study explored the effectiveness of the NSCT method alongside STD, SF, and VIFF validation. Fig. 10 displays the combined results of STD, SF, and VIFF values for two distinct groups. As depicted in Fig. 10(a), the NSCT method delivered STD, SF, and VIFF values of 47.69, 35.95, and 1.29 in the fusion results of the first group of multi-focus source images. This marks an enhancement over the Contourlet transform method, with increases in STD, SF, and VIFF values by 0.35, 10.48, and 0.32, respectively. Fig. 10(b) shows that in the second group's fusion results, the NSCT method achieved STD, SF, and VIFF values of 71.55, 55.73, and 1.14, respectively, indicating a marked improvement over the RGF (Recursive Gaussian Filtering) method, with reductions of 12.23, 17.28, and 0.18 in STD, SF, and VIFF values, respectively. These findings suggest that the NSCT method substantially outperforms other methods compared.

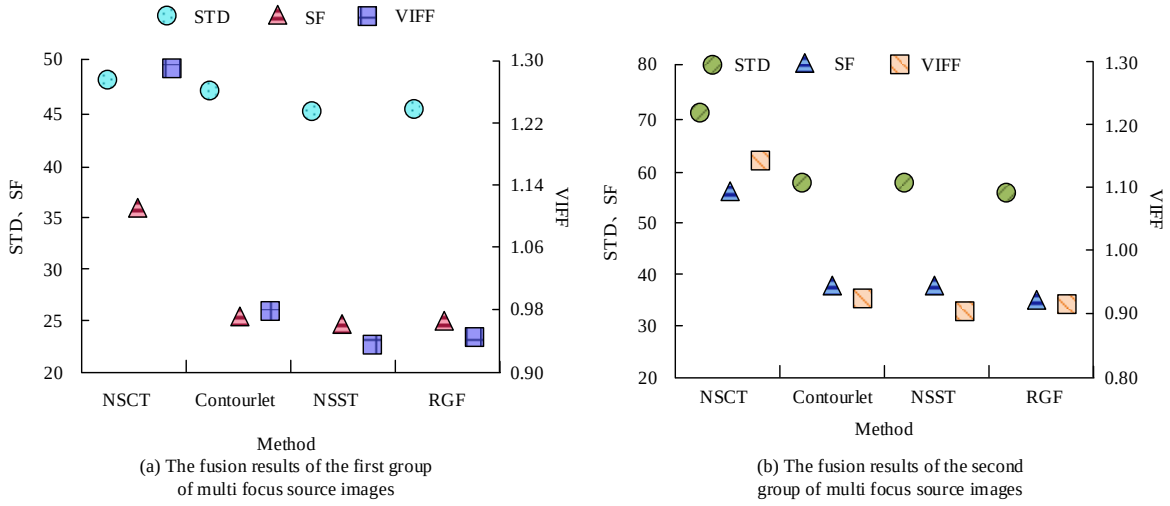
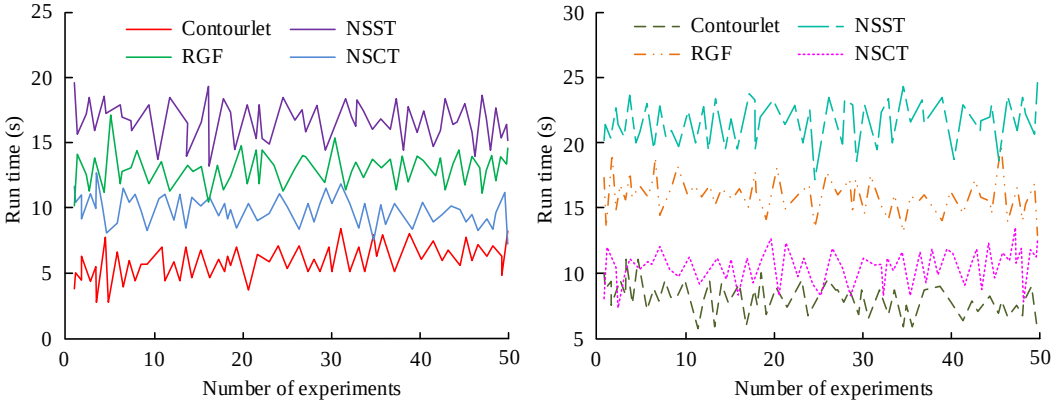


Fig. 10. STD, SF, and VIFF values for the two sets of fusion results.

The study extended its analysis to the runtime of several image fusion techniques to evaluate their computational efficiency. The experimentation encompassed 50 test executions, with each algorithm's runtime illustrated in Fig. 11. From Fig. 11(a), in the first set of multi-focus source image fusion outcomes, the NSCT method averaged a runtime of 10.03 seconds. The Contourlet transform, NSST, and RGF methods recorded average runtimes of 5.69, 17.69, and 13.72 seconds, respectively. As shown in Fig. 11(b), for the second set of fusion results, the average runtime of the NSCT method was 10.58 seconds. The rival methods averaged 8.59, 23.61, and 16.97 seconds. Notably, only the Contourlet transform method had a lower runtime than the NSCT, while the others lagged in computational efficiency. Considering fusion quality, the proposed approach in the research demonstrated superior performance.



(a) The running time of the first group of multi focus source image fusion

(b) The running time of the second group of multi focus source image fusion

Fig. 11. Running time of each algorithm.

C. Discussion

To objectively test the fusion performance in this study, we use key validation metrics such as Average Gradient (AG), Edge Intensity (EI), and computational time. We have carefully selected these metrics to comprehensively evaluate the algorithm's sharpness, edge preservation, and efficiency. Furthermore, we have assessed our NSCT-based approach against notable methods such as NSST and RGF. More results have been added that show NSCT-based fusion is better through quantitative and qualitative tests, strengthening the comparison even more. These encompass side-by-side visual comparisons of fused images, highlighting NSCT's capability to retain fine details while reducing artifacts. Furthermore, we show statistical proof that our method improves performance for both AG and EI across the chosen datasets.

The objective evaluation confirmed these results, where higher AG and EI values in NSCT meant better contrast and sharpness, which meant better fusion quality. The method's ability to get high STD, SF, and VIFF scores shows that it works to keep natural brightness and spatial coherence. This suggests that the NSCT-based method effectively gets around the problems with older methods, such as spectral aliasing and lack of directional sensitivity. Moreover, runtime analysis showed that NSCT efficiently balances fusion performance with computational cost. Although a little slower than the Contourlet Transform, NSCT performed much better than NSST and RGF, making it a good choice for real-time image processing tasks.

NSCT is better than NNSST and RGF because it eliminates problems like edge blurring, artifact creation, and slow computing. NSST has some issues with direction insensitivity and edge blurring. NSCT, on the other hand, has better direction selectivity for smoother transitions and better feature preservation. While RGF can smooth images, it often introduces pseudo artifacts and loses fine details, especially in high-texture areas. NSCT, on the other hand, uses localized filtering and multi-directional decomposition to reduce spectral aliasing. This makes artifact-free fusion with higher contrast and strong structural integrity possible. NSCT also balances fusion quality and computational efficiency better by doing fewer unnecessary calculations than NSST. This makes it a better image fusion process for real-time uses like medical imaging, surveillance, and remote sensing.

The findings highlight that the proposed NSCT-based MFIF technique offers an effective and efficient solution for multi-focus image fusion. The way it integrates images is better without slowing down computers, so it can be used in areas like medical imaging, surveillance, and satellite imagery.

V. Conclusion

Color MFIF technology has extensive applications across numerous fields requiring accurate and detailed imagery. This research introduced an advanced NSCT to overcome the shortcomings of existing fusion techniques. The results demonstrated the NSCT-based algorithm's outstanding performance in delivering superior fusion results. For the second experimental dataset, the NSCT approach showcased rich information content, vivid texture clarity, and even brightness distribution. Significantly, it prevented artifact creation and mitigated detail loss due to edge blurring. Quantitative evaluations further affirm its advantages: in the first dataset, the NSCT method recorded Standard Deviation (STD), Spatial Frequency (SF), and Visual Information Fidelity (VIFF) values of 47.69, 35.95, and 1.29, respectively, and in the second dataset, these metrics rose to 71.55, 55.73, and 1.14, notably outperforming other methods like the Contourlet

Transform, NSST, and RGF. Regarding computational efficiency, the NSCT method balanced runtime effectively. For the first dataset, it averaged a runtime of 10.03 seconds, compared to the 5.69 seconds, 17.69 seconds, and 13.72 seconds taken by the Contourlet Transform, NSST, and RGF methods, respectively. In the second dataset, the NSCT method had an average runtime of 10.58 seconds, maintaining competitiveness despite being slightly slower than the Contourlet Transform. These findings highlight the NSCT method's capacity to deliver excellent fusion quality and manageable computational demands. While the proposed approach shows strong and consistent fusion performance, this study did not consider the influence of source image quality on the fusion results. Future investigations might examine incorporating image preprocessing algorithms to enhance the quality of source images, thus further augmenting the reliability and scope of color MFIF technology.

REFERENCES

[1] Huang Y, Chen D. Image fuzzy enhancement algorithm based on contourlet transform domain. Multimedia Tools and Applications, 2020, 79(47-48): 35017-35032.

[2] Singh P, Diwakar M. Total variation-based ultrasound image despeckling using method noise thresholding in non-subsampled contourlet transform. International Journal of Imaging Systems and Technology, 2023, 33(3): 1073-1091.

[3] Vafaie S, Salajegheh E. A Comparative Study of Shearlet, Wavelet, Laplacian Pyramid, Curvelet, and Contourlet Transform to Defect Detection. Journal of Soft Computing in Civil Engineering, 2023, 7(2): 1-42.

[4] Ibrahim S I, Makhlouf M A, El-Tawel G S. Multimodal medical image fusion algorithm based on pulse coupled neural networks and nonsubsampled contourlet transform. MedicalBiological EngineeringComputing, 2023, 61(1): 155-177.

[5] Jahangir H, Khatibinia M, Kavousi M. Application of contourlet transform in damage localization and severity assessment of prestressed concrete slabs. Journal of Soft Computing in Civil Engineering, 2021, 5(2): 39-67.

[6] Li W, Lin Q, Wang K, Cai, K. Improving medical image fusion method using fuzzy entropy and nonsubsampling contourlet transform. International Journal of Imaging Systems and Technology, 2021, 31(1): 204-214.

[7] Kollem S, Reddy K R, Rao D S. Improved partial differential equation-based total variation approach to non-subsampled contourlet transform for medical image denoising. Multimedia Tools and Applications, 2021, 80(2): 2663-2689.

[8] Hasan M M, Islam N, Rahman M M. Gastrointestinal polyp detection through a fusion of contourlet transform and Neural features. Journal of King Saud University-Computer and Information Sciences, 2022, 34(3): 526-533

[9] Hu F, Cao H, Chen S, Sun, Y.,Su, Q. A robust and secure blind color image watermarking scheme based on contourlet transform and Schur decomposition. The Visual Computer, 2023, 39(10): 4573-4592.

[10] Li H, Nie R, Cao J, Guo, X., Zhou, D.,He, K. Multi-focus image fusion using u-shaped networks with a hybrid objective. IEEE Sensors Journal, 2019, 19(21): 9755-9765.

[11] Panigrahy C, Seal A, Mahato N K,Krejcar, O.,Herrera-Viedma, E. Multi-focus image fusion using fractal dimension. Applied Optics, 2020, 59(19): 5642-5655.

[12] Wang Y, Jin X, Yang J, Jiang, Q., Tang, Y., Wang, P.,Lee, S. J. Color multi-focus image fusion based on transfer learning. Journal of IntelligentFuzzy Systems, 2022, 42(3): 2083-2102.

[13] Liu S, Chen J, Rahardja S. A new multi-focus image fusion algorithm and its efficient implementation. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 30(5): 1374-1384.

[14] Huang Y, Hu X, Hao L, Gao, Y., Liu, Z.,Wang, P. Quantitative analysis of cell morphology based on the contourlet transform. IET Image Processing, 2020, 14(12): 2826-2832.

[15] Hasan M M, Hossain M M, Mia S, Ahammad, M. S.,Rahman, M. M. A combined approach of non-subsampled contourlet transform and convolutional neural network to detect gastrointestinal polyp. Multimedia Tools and Applications, 2022, 81(7): 9949-9968.

[16] Rahim R, Murugan S, Manikandan R, Kumar, A. Efficient Contourlet Transformation Technique for Despeckling of Polarimetric Synthetic Aperture Radar Image. Journal of Computational and Theoretical Nanoscience, 2021, 18(4): 1312-1320.

[17] Gaffar A, Joshi A B, Singh S, Srivastava, K. A high-capacity multi-image steganography technique based on golden ratio and non-subsampled contourlet transform. Multimedia Tools and Applications, 2022, 81(17): 24449-24476.

[18] Chung C H, Chen L J. Text mining for human resources competencies: Taiwan example. European Journal of Training and Development, 2021, 45(6/7): 588-602.

[19] Hasanvand M, Nooshyar M, Moharamkhani E, Selyari, A. Machine Learning Methodology for Identifying Vehicles Using Image Processing.Artificial Intelligence and Applications. 2023, 1(3): 170-178.

[20] Tijare P A, Purswani D, Rathi T, Chavhan, S., Jangid, V.,Agrawal, S. ABUSE RELATED POSTS REPORTER USING WEB CRAWLER IN PYTHON. EPRA International Journal of Research and Development (IJRD), 2022, 7(5): 230-234.