De-fencing and Multi-Focus Fusion using Markov Random Field and Image Inpainting

HANNAH ADEEL1, M MOHSIN RIAZ2, AND SYED SOHAIB ALI3
1Department of Electrical & Computer Engineering, COMSATS University, Islamabad, Pakistan (e-mail: sp16.pce-002@isbstudent.comsats.edu.pk)
2Centre for Advanced Studies in Telecommunication, COMSATS University, Islamabad, Pakistan (e-mail: mohsin.riaz@comsats.edu.pk)
3School of Science & Engineering, Habib University, Karachi, Pakistan (e-mail: sohaib.ali@sse.habib.edu.pk)

Corresponding author: M Mohsin Riaz (e-mail: mohsin.riaz@comsats.edu.pk).

ABSTRACT Multi-focus image fusion aims at combining source information from differently focused images. Fusion of multi-focus images has great applications in machine vision. The paper focuses on removal of fence occlusions in multi-focus images. The proposed model extracts fence occlusion map using salient image features, and refined by morphological operators. Binary operators, and inpainting methods are used for fence removal and restoration. The proposed model estimates the fence area using statistical characteristics in focused regions. Similarly, binary filtering is used to perform thinning of enlarged areas for optimised restoration. The proposed model employs guided filtering for consistency verification. Fusion and restoration results are compared using several (reference and no-reference based) image quality metrics. Simulations show that proposed scheme achieves better results (visually and quantitatively) as compared to existing state-of-the-art techniques.

INDEX TERMS Multi-focus, image fusion, de-fencing, inpainting, fence removal

I. INTRODUCTION Restoration of fence occluded images has emerged as an interesting research area. It is desirable to remove unwanted objects and distractors from images in a variety of ways. The scenes captured at places like zoo, parks, windows, or balconies are normally occluded with fences. Various occlusion removal and restoration schemes can be classified as intensity, transform-domain, and deep learning based methods. Generally these methods require multiple images or frames for fence removal and restoration. Therefore, fusion of fenced images in itself is a challenging task.

In this paper, a two stage framework is proposed for fence removal task from multi-focus images. The salient image features in focused regions are used to extract a fence occlusion map. The raw map is refined using guided and morphological filters. The fence is removed using binary operators, and inpainting is then used for restoration. The fusion weight map is refined by employing guided filtering for consistency verification. Proposed fusion model for multi-focus fenced images is shown in figure 1. The proposed scheme is compared with latest schemes for fusion. A step wise illustration of proposed scheme is represented in figure 2. The results of proposed scheme are much better visually and quantitatively as compared to existing state of the art techniques. Several image quality metrics based on entropy, phase congruency, and properties of human visual systems, are used for fusion evaluation. Similarly, reference and no-reference based metrics are used for evaluation of restored images.

The rest of the paper proceeds as follows: Section 2 presents a brief survey of multi-focus fusion, and fence removal techniques. The proposed scheme, along with its mathematical formulations, is presented in section 3. Simulation results and comparisons are discussed in section 4. The final section then concludes the paper while suggesting some future work.

II. RELATED WORK This section presents various state-of-the-art schemes for multi-focus fusion, and de-fencing. A systematic literature review is also reported in table 1. Xiong et al. [1], proposed deep learning based image inpainting scheme. The scheme uses deepcut [38], and Sobel operator for detection of in-
TABLE 1: A systematic literature review of multi-focus fusion and de-fencing schemes.

| Schemes                          | Task                     | Techniques                                                                 | Refinement methods                                                                 | Issues                                |
|---------------------------------|--------------------------|---------------------------------------------------------------------------|------------------------------------------------------------------------------------|---------------------------------------|
| Deep learning based methods     | Restoration              | deepcut and Sobel operator                                                 | generator descriptor                                                               | large training data                   |
| Model based methods [3], [4], [5], [6], [7] | De-fencing              | k-neighbours, subtraction, super-pixel segmentation, multiple-thresholds  | color similarity, morphological operators, hybrid inpainting, Criminisi inpainting [8] | images with and without flash, accurate mask segmentation, manual fence selection |
| Model based methods [9], [10]   | Restoration              | homographic matrices, dictionary based on k-means labeling,                | exemplar based inpainting                                                           | manual initialization                 |
| Deep learning based methods [11], [12], [13], [14], [15], [16], [17], [18], [19], [20] | De-fencing              | adversarial, structural, and perceptual information, residual learning, gradient distribution, image matting, graph cut, wavelet thresholding | Criminisi inpainting, conditional adversarial network, Optical flow estimation, split Bregman optimization, exemplar based inpainting | multiple GPUs, training data, blobs artifacts |
| Sparse Representation (SR) based methods [21], [22], [23], [24], [25] | Fusion                   | multi-scale decomposition, dictionary-learning, nuclear norm regularizer, morphologies constraints, adaptive fusion rules, fractional differential coefficients, geometric sparse coefficients | overcomplete dictionary, patch based clustering, single dictionary learning | time efficiency, separate fusion and noise removal tasks, information loss due to channel-wise processing |
| Model based methods [26], [27], [28], [29], [30], [31], [32], [33] | Fusion                   | unified low rank, NSCT, visual uniqueness, saliency measures, structure tensors, focus region detection, spatial-frequency, median filtering, anisotropic diffusion filters | super-pixel clustering, object consistency, guided filtering, residual removal, joint bilateral filter, morphological filtering | computational complexity, gray scale image fusion |
| Deep learning based methods [34], [35], [36], [37] | Fusion                   | SSIM, encoder features, K-means clustering, NSCT, Coupled-Neural-Ps        | consistency verification                                                            | photo realistic fusion, blocking artifacts, post processing |

In [11], image de-fencing technique uses graph cut and dilation to detect, and better represent the fence region irrespective of shape and color. In case of noisy input, wavelet thresholding and convolutional neural networks (CNN) based image de-noising, is performed. However, the scheme sometimes produces blobs in resultant image. In [12], single stage image de-fencing network based on adversarial, structural and perceptual information is proposed. The scheme provides promising results, however, requires multiple GPUs and training data. In [13], regression based method is proposed for fence detection to overcome the limitations of classification methods. Gaussian smoothing based on neighbouring pixels is used to inpaint behind the fence portion. The scheme uses empirical settings, and sometimes fails to detect complete contours; and generator-discriminator setting for prediction of contour completion. In [3], subtraction based on gray-scale binarization is applied on multi-focus auxiliary images to obtain initial mask for image inpainting. Color similarity is then used to remove false regions for refined map. However, the availability of auxiliary images (with and without flash), is a challenging task. In [9], an array of frames is used to remove an intrusive object from the reference (middle) frame. It uses homographic matrices to find the macro-blocks of reference image in other frames. Yu et al. [4], proposed iterative clustering based super-pixel segmentation method for reticular shield detection and removal. The scheme achieves better results than other algorithms, however, sometimes suffers from color and thin structural issues.

TABLE 2: Frequently used acronyms and notations, with their meanings.

| Notation | Meaning |
|----------|---------|
| CF       | Curvature-filtering based fusion          |
| GFDF     | Guided-filtering and focus-region detection fusion |
| CNN      | Convolutional neural network              |
| MISP     | Multimodal image seamless fusion          |
| NSCT     | Nonsubsampled contourlet transform        |
| RR       | Residual removal                          |
| HIM      | Hybrid inpainting method                  |
| FRPI     | Fence removal based on pyramid inpainting |
| $Q_{TE}$ | Tsallis entropy                           |
| $Q_{NCIE}$ | Nonlinear correlation information entropy |
| $Q_X$    | Xydeas metric                            |
| $Q_P$    | Phase congruency metric                   |
| $Q_C$    | Cvejc metric                             |
| $Q_Y$    | Yang’s metric                            |
| $Q_{CV}$ | Chen-Vashey metric                       |
| $Q_{CB}$ | Chen-Blum metric                         |
| $Q_{FMI}$ | Normalized feature mutual information metric |
| $Q_{SSIM}$ | Structural similarity index measure     |
| $Q_{NAE}$ | Normalized Absolute Error                 |
| $Q_{MSE}$ | Mean squared error                       |
| $Q_{BRISQUE}$ | Blind/referenceless image spatial quality evaluator |
| $Q_{NIQE}$ | Natural image quality evaluator          |
| $Q_{PIQE}$ | Perceptual image quality evaluator        |
irregular or unique shaped fences. Deep learning based image inpainting method [2], uses a pyramid based loss function (based on adversarial, style and $L_1$ information) to generate multi-scale structural information simultaneously.

In [5], multiple threshold values are used to extract multicolored fence based on intensity component, morphological operations, and hybrid inpainting scheme. In [10], exemplar-based image inpainting along with dictionary of salient structures (based on k-means labeling) is used to restore the damaged areas. The algorithm shows good performance, however, it needs manual initialization for detection of damaged areas. In [14], disparity map of stereo images is computed and refined by matting technique to obtain the fence mask. The optical flow estimate (while blurring the fence regions) is used in split Bregman optimization method while considering total variation as a regularization constraint. In [15], a modified convolutional neural network (CNN) based deep learning technique uses multiple stereo frames for fence detection. This scheme shows good performance, however, sometimes it fails to restore fences due to inaccurate motion estimation. Liu et al. [16], proposed fence removal as a layer decomposition and reconstruction problem based on deep learning. Coarsest level feature representation and motion vectors are used to reconstruct the background image using residual learning network. In [17], long short-term memory cells based supervised recurrent network with asymmetric loss function is employed to detect fences.

In [21], multi-scale decomposition, and sparse representation are used to jointly perform noise-removal and fusion. The fusion model achieves promising results for noisy images, however, the scheme can be further improved by employing hybrid features. Li et al. [22], proposed dictionary-learning fusion framework based on nuclear norm regularizer and morphological constraints. The scheme preserves fine scale details while avoiding noise, however, fusion and restoration tasks, are performed separately. In [23], all-in-focus resultant image is obtained using multi-scale sparse representation to extract salient information. The scheme defines adaptive fusion rules to decide in-focus, out-of-focus, and boundary regions. However, the scheme sometimes suffers from information loss due to channel-wise processing in color images. Yu et al. [24], used fractional differential coefficients and gradient histograms to extract salient image features. The scheme learns overcomplete dictionary using patch based clustering to transfer structural information in all-in-focus images. Similarly, Tan et al. [25], used geometric sparse coefficients (obtained from single dictionary image) to generate focus regions, and preserve important information in resultant images. The scheme also improves time complexity as overcomplete dictionary training is not required.

Khalid et al. [19], proposed supervised learning algorithm based on gradients distribution and edge orientations to classify fence pixels. Du et al. [18], employed fully connected CNN and temporal refinement to remove fence. Optical flow based motion estimation method [14] is then used for information fusion. Farid et al. [7], formulated the Gaussian distribution of the fence based on k-neighbours of an initially chosen sample. To effectively classify each pixel as a fence or non-fence pixel, Bayesian rule along with connected component method is used. Criminisi et al [8] method is then used at down-sampled level for filling the textural information. However, the scheme requires manual selection of fence pixels carefully. Gupta et al. [20], proposed image de-fencing framework based on cascaded adversarial networks. First network outputs the binary mask, while, the second network generates fence free image. In [26], median and anisotropic diffusion filtering are performed on high resolution images for focus map extraction. The maximum intensity rule and filtering-based difference images are then used to obtain fusion weight map.

Curvature filtering (CF) [27], method defines focus detection criterion based on spatial frequency, and local characteristics. The scheme efficiently extracts salient features in focused regions, and shows promising results. Qiu et al. [28], used guided filtering and focus region detection for fusion of multi-focus images. The scheme shows robust
Deep learning based technique [34], uses convolutional layered architecture, binary segmentation, and morphological filtering to obtain focus maps. The weighted-sum fusion strategy shows good quality results, however, sometimes suffers from inaccuracies which can be minimized by using different training models. In [29], bilateral filtering is used to refine the coarse weight maps obtained from saliency maps based on image gradients. The spatial domain method seamlessly fuses images from different modalities. Recently, Li et al. [30], proposed residual removal based multifocus fusion using multiscale focus correction, and non-subsampled contourlet transform (NSCT). The hybrid framework based on structure tensor shows promising results, and is time efficient. In [6], subtraction and Gaussian filtering are used to extract fences in multifocus images. Image inpainting methods are then used to restore the fence occluded regions. However, the scheme sometimes suffers from improper focus map detection.

In [31], focus maps are obtained using sum-modified Laplacian which are further refined by block consistency verification and guided filtering. Neighbour distance filtering and decision maps are then used to reconstruct an all in focus image. Recently, Peng et al. [35], proposed coupled neural-P architecture and NSCT based fusion scheme for multi-focus images. The scheme, however, requires an additional focus measure step in contrast to other CNN based methods. Yang et al. [32], used NSCT decomposition, and consistency verification with guided filter, in their image fusion framework. The fusion map for frequency components is obtained from visual uniqueness and saliency measures. Wang et al. [36], used K-means clustering and NSCT to fuse frequency components of different segmented regions. Zhang et al. [33], proposed super clustering of varying sizes to deal flat and rich in details regions to minimize jagged artifacts. The scheme depicts good results for color images, however, shows lower performance on gray scale images. Zhang et al. [37], used polarization imagery to train deep CNN in an unsupervised manner without ground truth images. The framework combines the loss functions based on SSIM and encoder features with a suitable parameter. In nutshell, deep learning based methods provide better accuracy, however, they have complex architecture and require large training data as compared to classical methods.
III. PROPOSED SCHEME

Let $A$ and $B$ represent two multi-focus input images with dimensions $X \times Y \times \gamma$; where $X$ designates the width, $Y$ represents the height, and $\gamma$ shows the number of channels in an image. Image $A$ has fence in focus whereas image $B$, has background in focus.

Since the focused regions contain more high frequency components as compared to de-focused areas, therefore, high frequency components of normalized intensity images are computed. Let $A_g$ and $B_g$ depict the normalized intensity information contained in source images, defined as,

$$
A_g(x, y) = \frac{1}{\gamma} \sum_{\gamma} A(x, y, \gamma)
$$

$$
B_g(x, y) = \frac{1}{\gamma} \sum_{\gamma} B(x, y, \gamma)
$$

Moreover, let $\overline{A}$ and $\overline{B}$ represent smoothed images as,

$$
\overline{A} = A_g * h_r
$$

$$
\overline{B} = B_g * h_r
$$
where $h_r$ represents the disk operator with radius $r \in (1, 2)$, that makes it more suitable to extract high frequency components in the fence region.

The difference images, based on intensity and corresponding blurred images, can roughly detect the focus regions in source images as,

\[
A_s(x, y) = |A_g(x, y) - A(x, y)| \\
B_s(x, y) = |B_g(x, y) - B(x, y)|
\tag{3}
\]

To exclude the false edges from estimated salient features through high frequency components, range threshold binarization is employed as,

\[
A_b(x, y) = \begin{cases} 
1, & \text{if } n_1 < \tilde{A}(x, y) < n_2 \\
0, & \text{otherwise}
\end{cases}
\tag{4}
\]

The limits pertaining to fence region components, $n_1$ and $n_2$, are derived from intensity histogram; while the smoothed version of difference image is defined as,

\[
\tilde{A} = A_s * h_s.
\tag{5}
\]
The binary map $A_b$ provides an initial estimate for the fence region with some gaps specifically in the crossing areas of the fence regions. The fence region estimation is then refined by performing the guided filtering on the binary map $A_b$, given by,

$$
VOLUME 4, 2016
$$
The binary eroded mask \( \omega \) masked image as,

\[
M_I(x, y) = M_A(x, y) - M_B(x, y)
\]

(7)

Here \( M_B \) depicts the background activity extracted from \( B_\lambda \), by performing guided filtering with \( B_\lambda \) as guidance image. Further refined fence map is represented by,

\[
\omega(x, y) = 1 - \zeta((1 - \delta(M_I(x, y))))
\]

(8)

where \( \zeta(\cdot) \) and \( \delta(\cdot) \) represent morphological closing and small objects removing operations. The disk size in the closing operation is optimized based on blur metric [48].

The fence region weight map is then used to generate initial fused image, represented as,

\[
I = \omega \odot A + (1 - \omega) \odot B
\]

(9)

The final fusion map is obtained by employing guided filtering for consistency verification performed on \( \omega \), as follows,

\[
W = \xi(I, \omega)
\]

(10)

based on guided image, \( I \).

The fused image is generated by using guided filtered based final weight map as,

\[
F = W \odot A + (1 - W) \odot B
\]

(11)

To remove fence, the initial weight map \( \omega \) is eroded as,

\[
\omega^a = (\omega \odot se)
\]

(12)

where \( \odot \) indicates erosion and \( se \) is structural element.

The erosion reduces the fence area for better inpainting. The binary eroded mask \( \omega^a \) is then used to prepare the masked image as,

\[
C(x, y) = \begin{cases} B(x, y), & \text{if } \omega^a(x, y) = 1 \\ 0, & \text{otherwise} \end{cases}
\]

(13)

Next, fence removal task is achieved by using such an inpainting method that uses local information to fill the empty regions while avoiding artifacts and distortions. Figure 8 shows major steps followed to perform the de-fencing task. Markov random field (MRF) based learnt network [49], using minimum mean squared error (MSE) estimation effectively restores the missing areas [50]. The generative properties of MRF use flexible Gaussian scale mixtures along with noisy and smoothed versions to reconstruct the missing regions. The network has learnt statistics of natural images that make it more suitable for restoration of natural scene images. The generation process can be represented as,

\[
R = \Psi(I, C, (\omega^a))
\]

(14)

Here, \( \Psi \) presents the inpainting method, that uses learnt model \( I \) to fill the masked image \( C \), while \( R \) represents the restored image. To overcome the difference in the filled region, and making a more pleasant look, the restored image is sharpened as,

\[
S = \chi(R, se)
\]

(15)

where \( \chi(\cdot) \) represents the sharpening operation with structuring element \( se \). The restored and sharpened image \( S \), is then used to generate final de-fenced image \( D \).

\[
D = (1 - \omega^a) \odot F + \omega^a \odot S)
\]

(16)

IV. RESULTS & DISCUSSION

Different fence masks are superimposed over clean scenes to simulate real time fenced images. Multi focus fence occluded data are simulated by inducing Gaussian blur in clear scenes, and extracted fences. The clear fence superimposed are used as in restoration, while clean images are used as reference [13].

The proposed and existing schemes are compared visually and quantitatively using several metrics like Tsallis entropy \( Q_{TE} \) [39], NCIE \( Q_{NCIE} \) [40], Xydeas Metric \( Q_x \) [41], Phase Congruency \( Q_P \) [42], Cvejic metric \( Q_C \) [43], Yang’s \( Q_Y \) [44] Chen-Varshney \( Q_{CV} \) [45], Chen-Blum \( Q_{CB} \) [46], and normalized feature mutual information \( Q_{FMI} \) [47].

Figure 3 presents fusion results of proposed and existing schemes on a ‘field’ image (case-1). The scene contains fence, field, and sky regions. Most existing methods have shown overall good performance. CNN [34], and MISF [29], have shown better visual results in comparison to GFDF.
[28], NSCT-RR [30], and CF [27]. Still, existing schemes produce different unwanted artifacts. For instance, CF [27], contains partially focused fence. MISF [29], suffers from blurring effect near the lower part of the image. Similarly, a small portion of fence is thinned (in the central part of output image) by CNN [34]. However, the proposed scheme has properly focused the fence and background parts.

Figure 4 presents fusion results for ‘fuel-station’ (case-2). The scene contains a lot of edge and color information. Fused results contain both background and foreground (fence) information. However, output images generated by CNN [34], and NSCT-RR [30], suffer from blur along the sides. On careful observation, it can be seen that fence information is not properly used by MISF [29], GFDF [28], and CF [27]. While most schemes under study have shown limited performance, proposed scheme has produced better results by proper extraction and fusion of source information.

Figure 5 shows fusion results for the pool (case-3). First multi-focus input has foreground fence in focus, while second input has background in focus. It can be seen that most of the schemes have properly used the background information. However, the foreground fence is not properly fused, and results suffer from blurring artifacts. For instance, bottom (center) and top (right) parts, in GFDF [28], output are blurred. Similar observations can be made for outputs generated by CNN [34], MISF [29], and NSCT-RR [30]. CNN [34], and MISF [29] also suffer from blurring artifact along the fence in tree regions. CF [27], has properly used background information, however, fence information is not properly fused. NSCT-RR [30], has some unclear areas particular along sides, boundary regions as well as suffer from broken fence. While most of the existing schemes suffer from artifacts, the proposed scheme has successfully recovered the details in background as well as fence region.

Figure 6 shows fusion results for ‘deer’ image. The multi-focus pair contains thin occluded foreground. In this case, existing schemes have shown limited fusion performance. For instance, the output generated by CF [27], is nearly similar to background-focused input (figure 6(b)), while ignoring the foreground net. On the other hand, outputs generated by CNN [34], MISF [29], GFDF [28], and NSCT-RR [30] are close in appearance to fence-in-focus input (figure 6(a)). Further, some part of foreground fence is blurred in CNN [34]. The proposed scheme generates pleasant fusion result where both parts, foreground and background, are entirely focused.

Figure 7 shows fusion results for ‘canal’ image. Fusion results of CF [27], CNN [34], MISF [29], and NSCT-RR [30], suffer from blurring artifacts. Furthermore, in CF [27], entirely blurred fence output indicates that foreground information is least used. Proposed scheme along with GFDF [28], (figure 6(d)), show acceptable outputs.

Table 4 presents a time complexity analysis of the existing and proposed fusion methods. Simulations are carried out using machine having CPU (specifications: Intel(R) Core(TM) i7-6500U CPU @ 2.50GHz-2.60 GHz), and 8GB RAM. The mean run time (based on images having size 407 × 271), show that proposed scheme is time efficient in comparison to most of the existing fusion methods.
Proposed de-fencing method is compared to hybrid inpainting method (HIM) [7], and fence removal based on pyramid inpainting (FRPI) [6]. The multi-focus image pairs are used in case of [6]. However, in case of [7], fenced image is used as input. Here, five cases are presented to show the superiority of our proposed algorithm.

Figure 9 presents the de-fencing case called 'cow'. Here, the cow, grass and sky region are occluded with fence. Restored image by HIM [7] has successfully restored sky region and parts of grass. However fence in the lower part is not fully removed and parts of cow’s legs are merged with grass. Restored parts by FRPI [6] are better as compared to HIM [7], however, it suffers from fence residues. The proposed scheme has shown better restoration results as fence is fully removed as well as the inpainted regions are least affected.

Figure 10 presents image de-fencing case called 'home'. Here, fence covers body, clothes, leaves, shower and other parts in background. Restored image by HIM [7] has removed most of the fence however, regions use improper information to fill in colors as shower is hardly restored. Moreover, it can be observed that left part of the image is distorted with information from parts of woman’s shirt and hand.

Apart from fence region distortions, FRPI [6], has fully removed the fence. In most regions, suitable background information is used in filling the fence region. However, some parts of shower, leaves, and cap, suffer from color distortions. The proposed scheme yields better results as fence is successfully removed, fence regions are properly filled with local information, and there are no distortions otherwise.

Figure 11 presents image de-fencing case called 'veg'. Here, fence mainly occludes vegetation along with wooden crates and stones. Outputs generated by HIM [7], and FRPI [6], have successfully removed most of the fence. Blurring effect is evident in HIM output [7], specifically in the lower left corner of the image. Some noise distortions can be observed in the Chon [6]. The proposed scheme successfully removed the fence as well as fence regions are properly filled without suffering from other distortions.

Figure 12 shows the multi-focus images, along with fence super imposed image, and de-fencing results for the 'bird' case. It can be observed that resultant images by HIM [7], and Chon [6], are not properly de-fenced. For example, fence parts (with gray shade) in HIM [7], recovered image are not completely removed. On the other hand, background fence area is not properly filled in Chon [6]. It can be observed that the de-fenced image generated by the proposed scheme is better compared to existing schemes.

Figure 13 illustrates de-fenceing results for the 'lawn' case. In this case also, Chon [6], suffers from improper inpainting results in the extracted fence regions. Similarly, on careful observation, distortions are evident in the upper parts of the recovered image by HIM [7]. The proposed scheme shows better results in comparison to existing methods.

In summary, HIM [7] suffers from poor fence detection resulting in distortions in parts even not covered by the fence.
FRPI [6], better detects the fence thus avoids such problems, however, fence map must be refined to reduce residual noise. The proposed scheme not only better detects the fence regions, but also improves the fence map to avoid distortions in fence free regions, and then followed by inpainting method showing better results.

Image quality features like contrast, luminance, and structure, are inherently related to their statistical characteristics. Popularly known image quality testers [51], used for proposed scheme evaluation are briefly described here. Mean squared error ($Q_{MSE}$), measures average squared-error between original and recovered image pixels. Normalized Absolute Error ($Q_{NAE}$) [52] computes mean absolute error between images, while structural similarity index measure ($Q_{SSIM}$) [53], is based on locally used luminance and structural information. Blind/referenceless image spatial quality evaluator ($Q_{BRISQUE}$) [54], predicts differential score by using feature model trained on natural scene images with distortions like blurring and noise. Natural image quality evaluator ($Q_{NIQE}$) [55], computes difference between image features and features extracted by statistical model based on natural scenes. Perceptual image quality evaluator ($Q_{PIQE}$)
is inversely proportional to image quality as it evaluates blocking artifacts in distorted blocks using local contrast information. Here $Q_{MSE}$, $Q_{NLE}$, and $Q_{SSIM}$ are full reference while $Q_{BRISQUE}$, $Q_{NIQE}$, and $Q_{PIQE}$ are no-reference quality metrics. Ideal values for $Q_{MSE}$, $Q_{NLE}$, and $Q_{SSIM}$ are 0, $\infty$, and 1 respectively; while $Q_{BRISQUE}$, and $Q_{PIQE}$ take values in the range [0, 100]. Lower scores for $Q_{MSE}$, $Q_{NLE}$, $Q_{BRISQUE}$, $Q_{NIQE}$, and $Q_{PIQE}$ while higher scores for $Q_{SSIM}$, indicate better quality images.

Scores presented in table 5 support the proposed method in comparison to state-of-the-art methods [7], [6].

V. CONCLUSION

This study focuses on restoration of fence occluded multi-focus images. The proposed framework uses salient image features in focused regions to extract fence occlusion maps. The raw map is refined using guided and morphological filtering. The fusion map is refined by employing guided filtering for consistency verification. The fence is removed using binary operators, and inpainting is then used for restoration. The results of proposed scheme are much better, visually and quantitatively, as compared to existing state of the art techniques.

It should be mentioned that filtering parameters require fine tuning while working with different types of fences. Hence, optimal restoration needs fence map refinement, and thinning operations with careful parameter selection. This proposed model uses multi-focus scenes without skewing artifacts. Deep learning methods can be incorporated for better extraction of depth segmentation maps from multifocus fence occluded images. Similarly, the work can be extended to videos by exploiting frames similarity and projective geometry.

**REFERENCES**

[1] W. Xiong, J. Yu, Z. Lin, J. Yang, X. Lu, C. Barnes, and J. Luo, “Foreground-aware image inpainting,” in Proceedings of the IEEE CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 5840–5848.

[2] J. Yang, Z. Qi, and Y. Shi, “Learning to incorporate structure knowledge for image inpainting,” in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, no. 07, 2020, pp. 12605–12612.

[3] Y. Li, Y. Wang, and Y. Piao, “Extraction of thin occlusions from digital images,” in Selected Papers of the Chinese Society for Optical Engineering Conferences held October and November 2016, vol. 10255. International Society for Optics and Photonics, 2017, p. 1025540.

[4] Z. Yu, L. Cai, W. Jin, and T. Jin, “Reticular shield detection algorithm using superpixel segmentation,” in 2020 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC). IEEE, 2020, pp. 1–6.

[5] M. Varalakshamamma and T. Venkateswarlu, “Detection and restoration of image from multi-color fence occlusions,” Pattern Recognition and Image Analysis, vol. 29, no. 3, pp. 546–558, 2019.

[6] C. Lam Ka, “Fence removal from images,” University of Macau, Hong-kong, 2020. [Online]. Available: https://github.com/fyiatrian/GroupPfyp.git

[7] M. S. Farid, A. Mahmood, and M. Grangetto, “Image de-fencing framework with hybrid inpainting algorithm,” Signal, Image and Video Processing, vol. 10, no. 7, pp. 1193–1201, 2016.

[8] A. Criminisi, P. Pérez, and K. Toyama, “Region filling and object removal by exemplar-based image inpainting,” IEEE Transactions on Image processing, vol. 13, no. 9, pp. 1200–1212, 2004.

[9] R. Balanescu, A. Sterca, and I. Badarinza, “Removal of unwanted objects from still photographs,” in SoftCOM, 2020, pp. 1–6.

[10] L. Calatroni, M. d’Autume, R. Hocking, S. Panayotova, S. Parisotto, P. Ricciardi, and C.-B. Schönlieb, “Unveiling the invisible: mathematical methods for restoring and interpreting illuminated manuscripts,” Heritage science, vol. 6, no. 1, pp. 1–21, 2018.

[11] M. Varalakshamamma and T. Venkateswarlu, “Restoration of noisy and noiseless fence occlusion images,” Pattern Recognition and Image Analysis, vol. 30, no. 3, pp. 428–439, 2020.
TABLE 5: Quantitative performance evaluation for latest and proposed schemes.

| Case  | Scheme | $Q_{MSE}$ [51] | $Q_{NAE}$ [52] | $Q_{SSIM}$ [53] | $Q_{BRISQUE}$ [54] | $Q_{NIQE}$ [55] | $Q_{PIQE}$ [56] |
|-------|--------|---------------|---------------|---------------|----------------|----------------|---------------|
| Case-1 | HIM [7] | 552.3021 | 0.0642 | 0.8733 | 36.4286 | 3.8262 | 33.7980 |
|        | FRPI [6] | 315.1843 | 0.0440 | 0.9249 | 30.2057 | 3.4534 | 27.3827 |
|        | Proposed | 143.0999 | 0.0275 | 0.9545 | 19.4145 | 3.3967 | 23.9660 |
| Case-2 | HIM [7] | 1123.9737 | 0.1354 | 0.7777 | 36.2970 | 3.9395 | 51.6555 |
|        | FRPI [6] | 255.1539 | 0.0483 | 0.9213 | 37.9356 | 3.5291 | 41.4184 |
|        | Proposed | 126.7220 | 0.0311 | 0.9520 | 33.4107 | 3.3607 | 37.7252 |
| Case-3 | HIM [7] | 668.2809 | 0.1164 | 0.8062 | 47.6456 | 4.8882 | 45.2240 |
|        | FRPI [6] | 416.9340 | 0.0899 | 0.8731 | 42.3933 | 4.7738 | 31.8811 |
|        | Proposed | 244.4553 | 0.0717 | 0.9005 | 33.3716 | 4.8464 | 27.5886 |
| Case-4 | HIM [7] | 83.7274 | 0.0369 | 0.9620 | 21.4050 | 3.1698 | 40.8037 |
|        | FRPI [6] | 184.0263 | 0.0870 | 0.9230 | 33.0699 | 3.7766 | 48.4080 |
|        | Proposed | 51.1898 | 0.0429 | 0.9788 | 23.1879 | 2.9895 | 40.1341 |
| Case-5 | HIM [7] | 3044.5415 | 0.1780 | 0.7634 | 49.6324 | 5.0675 | 46.6395 |
|        | FRPI [6] | 684.1267 | 0.0949 | 0.8653 | 42.6184 | 4.3942 | 62.7985 |
|        | Proposed | 149.7147 | 0.0362 | 0.9350 | 37.3351 | 3.6109 | 47.9230 |

[12] D. Gupta, S. Jain, U. Tripathi, P. Chattopadhyay, and L. Wang, “Fully automated image de-fencing using conditional generative adversarial networks,” arXiv preprint arXiv:1908.06837, 2019.

[13] T. Matou and M. Ikheara, “Single-image fence removal using deep convolutional neural network,” IEEE Access, vol. 8, pp. 38846–38854, 2019.

[22] H. Li, Y. Wang, Z. Yang, R. Wang, X. Li, and D. Tao, “Discriminative dictionary learning-based multi-component decomposition for detail-preserving noisy image fusion,” IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 4, pp. 94–109, 2020.

[37] J. Zhang, J. Shao, J. Chen, D. Yang, and B. Liang, “Polarization image representation for multi-focus image fusion,” Pattern Recognition, vol. 79, pp. 107752, 2021.

[47] L. Yu, Z. Zeng, H. Wang, and W. Pedrycz, “Fractional-order differentiation based on k-means clustering and nsct,” in Seventh Symposium on Novel Photoelectronic Detection Technology and Applications, vol. 11763. In Proceedings of the International Symposium on Photoelectronic Detection Technology and Applications, vol. 11763. International Society for Optics and Photonics, 2021, p. 1176393.

[49] J. Zhang, J. Shao, J. Chen, D. Yang, and B. Liang, “Polarization image fusion with self-learned fusion strategy,” Pattern Recognition, vol. 118, p. 108045, 2021.

[51] W. Tan, H. Zhou, S. Rong, K. Qian, and Y. Yu, “Fusion of multi-focus images via a gaussian curvature filter and synthetic focusing degree criterion,” Applied optics, vol. 57, no. 35, pp. 10.0992–10.101, 2018.

[52] X. Qiu, M. Li, L. Zhang, and Y. Yuan, “Guided filter-based multi-focus image fusion through focus region detection,” Signal Processing: Image Communication, vol. 72, pp. 35–46, 2019.

[53] K. Zhan, L. Kong, B. Liu, and Y. He, “Multimodal image seamless fusion,” Journal of Electronic Imaging, vol. 28, no. 2, p. 023027, 2019.

[54] X. Li, F. Zhou, H. Tan, Y. Chen, and W. Zuo, “Multi-focus image fusion based on nonsubsampled contourlet transform and residual removal,” Signal Processing, vol. 184, p. 108062, 2021.

[55] F. Zhou, X. Li, J. Li, R. Wang, and H. Tan, “Multifocus image fusion based on fast guided filter and focus pixels detection,” IEEE Access, vol. 7, pp. 50780–50796, 2019.

[56] B. Yang, Y. Sun, and Y. Li, “Image fusion with structural saliency measure and content adaptive consistency verification,” Journal of Electronic Imaging, vol. 29, no. 1, p. 013014, 2020.

[57] Q. Zhang, F. Wang, Y. Luo, and J. Han, “Exploring a unified low rank representation for multi-focus image fusion,” Pattern Recognition, vol. 113, p. 107752, 2021.

[58] Y. Liu, X. Chen, H. Peng, and Z. Wang, “Multi-focus image fusion with a deep convolutional neural network,” Information Fusion, vol. 36, pp. 191–207, 2017.

[59] H. Peng, B. Li, Q. Yang, and J. Wang, “Multi-focus image fusion approach based on ssd cost function,” Pattern Recognition, vol. 113, p. 107752, 2021.

[60] X. Wang, W. Peng, and X. Feng, “Research on image fusion algorithm based on k-means clustering and nsct,” in Seventh Symposium on Novel Photoelectron Detection Technology and Applications, vol. 11763. International Society for Optics and Photonics, 2021, p. 1176393.

[61] J. Zhang, J. Shao, J. Chen, D. Yang, and B. Liang, “Polarization image fusion with self-learned fusion strategy,” Pattern Recognition, vol. 118, p. 108045, 2021.
International Journal of Innovative Computing, Information and Control, vol. 3, no. 6, pp. 1433–1447, 2007.

[43] N. Čvejic, D. Bull, and C. Canagarajah, “Metric for multimodal image sensor fusion,” Electronics Letters, vol. 43, no. 2, pp. 95–96, 2007.

[44] C. Yang, J.-Q. Zhang, X.-R. Wang, and X. Liu, “A novel similarity based quality metric for image fusion,” Information Fusion, vol. 9, no. 2, pp. 156–160, 2008.

[45] H. Chen and P. K. Varshney, “A human perception inspired quality metric for image fusion based on regional information,” Information fusion, vol. 8, no. 2, pp. 193–207, 2007.

[46] Y. Chen and R. S. Blum, “A new automated quality assessment algorithm for image fusion,” Image and vision computing, vol. 27, no. 10, pp. 1421–1432, 2009.

[47] M. B. A. Haghhighat, A. Aghagolzadeh, and H. Seyedarabi, “A non-reference image fusion metric based on mutual information of image features,” Computers & Electrical Engineering, vol. 37, no. 5, pp. 744–756, 2011.

[48] F. Crete, T. Dolmiere, P. Ladret, and M. Nicolas, “The blur effect: perception and estimation with a new no-reference perceptual blur metric,” in Human vision and electronic imaging XII, vol. 6492. International Society for Optics and Photonics, 2007, p. 64920I.

[49] M. Hu, “Image inpainting,” https://github.com/MihawkHu/Image_Inpainting/tree/master/Markov_Random_Field, retrieved August 20, 2021.

[50] U. Schmidt, Q. Gao, and S. Roth, “A generative perspective on mrfs in low-level vision,” in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2010, pp. 1751–1758.

[51] Matlab, “Image Quality Metrics,” https://www.mathworks.com/help/images/image-quality-metrics.html, retrieved August 20, 2021.

[52] S. Muthulakshmi, C. S. Dash, and S. Prabaharan, “Memristor augmented approximate adders and subtractors for image processing applications: An approach,” AEU-International Journal of Electronics and Communications, vol. 91, pp. 91–102, 2018.

[53] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” IEEE transactions on image processing, vol. 13, no. 4, pp. 600–612, 2004.

[54] A. Mittal, A. K. Moorthy, and A. C. Bovik, “No-reference image quality assessment in the spatial domain,” IEEE Transactions on image processing, vol. 21, no. 12, pp. 4695–4708, 2012.

[55] A. Mittal, R. Soundararajan, and A. C. Bovik, “Making a “completely blind” image quality analyzer,” IEEE Signal processing letters, vol. 20, no. 3, pp. 209–212, 2012.

[56] N. Venkatanath, D. Praneeth, M. C. Bh, S. S. Channappayya, and S. S. Medasani, “Blind image quality evaluation using perception based features,” in 2015 Twenty First National Conference on Communications (NCC). IEEE, 2015, pp. 1–6.