Infrared imaging enhancement through local window-based saliency extraction with spatial weight

Tong Li1,2 | Jufeng Zhao1,2 | Xiaohui Wu2 | Haifeng Mao2 | Guangmang Cui1,2

1 Zhejiang Provincial Key Lab of Equipment Electronics, Hangzhou Dianzi University, Hangzhou, China
2 School of Electronics and Information, Hangzhou Dianzi University, Hangzhou, China

Correspondence
Jufeng Zhao, No. 1158, 2nd Street, Jianggan District, Hangzhou City, Zhejiang Province, China
Email: dabaozjf@hdu.edu.cn

Abstract
Infrared image enhancement is an effective way to solve contrast reduction or details degradation in infrared imagery. An infrared enhancement approach based on local saliency extraction is proposed here. First, saliency maps are extracted within a local window by combining spatial weight. Second, with the change of the window size, potential targets and details in different sizes can be extracted. Considering window sizes as the scales, the saliency maps are obtained and infrared images are enhanced at different scales, and finally, multi-scale fusion is used to achieve the enhancement. Eight popular infrared enhancement approaches are introduced for comparison. Subjective qualitative observation experiments show that our strategy based on local saliency analysis and multi-scale fusion can well extract potential targets and areas of varying size from the source images with different sizes, to obtain a good enhancement effect. Meanwhile, we introduce six objective evaluation methods to measure the results, and the evaluation data to prove the effectiveness of the proposed algorithm. The experiments also indicate the real-time processing capability of the proposed method. The proposed method is finally well applied in the hardware system and shows its good performance.

1 | INTRODUCTION

Because of the characteristics of non-active imaging, infrared imaging is widely applied in modern military and civilian fields [1, 2]. However, infrared imaging also has disadvantages, such as low target-background contrast, blurred or degraded target edges and details. Therefore, the infrared image enhancement method is a very effective way to improve image quality.

The research on infrared image enhancement is a hot topic, which can be roughly divided into spatial domain enhancement and frequency domain enhancement [3–5]. Because of fast implementation, histogram equalization (HE)-based methods have been successfully applied. HE methods usually rescale the dynamic range and grayscale histogram distribution [6, 7]. However, the HE approaches would lead to weak enhancement, loss of detailed information, and amplification of image noise. Considering the above disadvantages of HE, some improved algorithms are designed, including enhancement method using an adjacent-blocks-based modification for local HE [8], contrast limited adaptive histogram equalization (CLAHE) [9], adaptive gamma correction based on cumulative HE approach [10] and adaptive histogram equalization and brightness correction (AHPBC) [11]. To enhance the contrast and details, Ashiba et al. improve and propose three ideas for infrared enhancement [12]: The first approach makes use of gamma correction with histogram matching, the second one depends on hybrid gamma correction with CLAHE, the third method utilises a trilateral enhancement which includes three stages, segmentation, enhancement and sharpening. Those methods would well highlight the target areas, but they would create some negative effects, such as enhanced background and amplified noise. For the problem of low dynamic range, researchers propose a local edge-preserving filter-based method (LDPF) [13], which could well highlight the details. In LDPF, the authors use entropy-guided gamma correction to adjust the base layer, but gamma correction may amplify some noise and make the image distorted. Moreover, to improve the visual effects of degradation, a water surface infrared image enhancement method based on

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. IET Image Processing published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.
radiative transfer theory is proposed \cite{14}. Some multi-level or multi-scale-based spatial domain enhancement methods like the multi-scale top-hat transform (MSTH) method \cite{15}, multi-level image enhancement \cite{16} and sub-band decomposition \cite{17}, all can produce results with good visual effect. However, these algorithms may produce an unsatisfactory result, amplifying the grey levels of the target area and introducing noise. And those multi-level or multi-scale-based methods work relative slowly. Moreover, image enhancement approaches based on convolutional neural network (CNN) and deep learning are studied recently \cite{18,19,20}. Inspired by Retinex theory, Zhang and his partners build a simple yet effective network for kindling the darkness (KinD) \cite{18}, and it performs well in contrast and detail enhancement. An image enhancement using zero-reference deep curve estimation (Zero-DCE) is proposed \cite{20}, which formulates enhancement as a task of image-specific curve estimation with a deep network. Kuang et al. mention a single infrared image enhancement using a deep CNN \cite{19}, which produces a result with good contrast. The power of CNN and deep learning is high, but they need accurate, proper and enough data to train. The data in our infrared enhancement may not satisfy these demands, the deep learning-based method needs to be improved with more data in real applications.

The visual effect is one of the most important factors for infrared enhancement. Studying from the human visual mechanisms, saliency analysis can reflect the ability of objects in the scene to attract visual attention. Thus, visual saliency is introduced into infrared enhancement. Visual saliency is a research hotspot in the field of computer vision in recent years \cite{21}. Now, many studies show that this theory has been applied well in the field of infrared enhancement and further studies, including object detection \cite{22}, image fusion and image enhancement \cite{23}.

Here, we propose a local saliency extraction-based infrared image enhancement method. The first main contribution and novelty is local-window with spatial weight-based grey distance. The second main contribution and novelty is local-window with spatial weight-based grey distance and local window with spatial weight.

2.1 Local window-based saliency extraction

Since HVS is sensitive to image contrast, researchers present a histogram-based contrast (HC) method to get a saliency map based on colour statistics of the original image \cite{25}, and they operate the algorithm in \(L^*a^*b^*\) space. Inspired by this work, we prefer to use pixel value distance in infrared images to achieve saliency extraction.

The saliency value \(V(i,j)\) at pixel \((i,j)\) is defined as:

\[
V(i,j) = \sum_{\forall (x,y) \in f} \left| f_{ij} - f_{xy} \right| \sum_{\forall (x,y) \in f} \left| f_{ij} - f_{xy} \right|
\]

where \(f\) denotes the original image, and \((x,y)\) is an arbitrary pixel within \(f\). \(f_{ij}\) and \(f_{xy}\) represent the pixel value or grey value at the corresponding position or pixel, \(\forall (x,y) \in f\) means any arbitrary pixel \((x,y)\) in \(f\). So, according to Equation (1), each pixel \((x,y)\) that satisfies the condition \(\forall (x,y) \in f\) should do \(f_{ij} = f_{xy}\), to achieve \(\sum_{\forall (x,y) \in f} \left| f_{ij} - f_{xy} \right| \sum_{\forall (x,y) \in f} \left| f_{ij} - f_{xy} \right| = 0\).

There are potential targets with different sizes, so we intend to highlight features of different targets, especially those with small sizes. Since global-based saliency extraction determined by Equation (1) cannot be adjusted for different sizes, this kind of approach would be out of action sometimes. We expect to design a local-based strategy. In addition, considering the spatial correlation between neighbouring pixels, pixel-to-pixel correlations need to be taken into account when calculating pixel differences. Thus, we impose a spatial weight matrix \(W\) on local-based saliency extraction. The spatial relationship can be found in Figure 1, and Equation (1) is rewritten as follows:

\[
V(i,j) = \left| W((i,j), (x,y)) \sum_{(i,j)} \left[ W((i,j), (x,y)) \sum_{(i,j)} \left[ f_{ij} - f_{xy} \right]^2 \right] \right|^2
\]

where \(U\) denotes a local window with a size of \(M\times M\), whose centre is \((i,j)\). And \(\Omega\) denotes a local disk-like region, whose radius is \(r\). Meanwhile, the centre is still \((i,j)\). \(\Omega \in U, \Omega\setminus U\) means the area \(U\) except \(\Omega\), \(f_{ij}\) denotes the average grey value within a local region \(\Omega\). In Equation (2), \(W((i,j), (x,y))\) represents the spatial weight for the position \((x,y)\), which is determined by the spatial distance between \(\Omega\) and \((x,y)\). Since the spatial correlation between neighbouring pixels, we consider the small weight in neighbours and large in distant. That means the weight of the nearest pixel should have the smallest weight. And for an arbitrary pixel \((x,y)\) in Figure 1(b) outside \(\Omega\), the corresponding \(W((i,j), (x,y))\) is defined as:

\[
W((i,j), (x,y)) = \frac{1}{\sum_{(i,j)} \left[ f_{ij} - f_{xy} \right]^2}
\]
where $\sigma^2$ is the dilatation factor, and $D_2(\Omega, x,y)$ denotes the spatial distance between $\Omega$ and pixel $(x, y)$ using numbers of pixels. From the equation, we could find that the closer between $(x,y)$ and $\Omega$, the smaller $W(\Omega, x,y)$ would be, and that means the smaller influence $f_{ij}$ affects $f_\Omega$ which includes $f_j$.

For the region, the weight of the nearest pixel should have the smallest weight. The nearest pixel is just one pixel away. This pixel $(x,y)$ satisfies $D_2(\Omega, x,y) = 1$. Little influence means small $W$. The maximum is 1, decrease one or two orders of magnitude is enough to satisfy small $W$ to achieve little influence. So, 1 decrease one or two orders of magnitude is 0.1, and 1 decrease two orders of magnitude is 0.01. We need $W = 0.01–0.1$. Finally, we select the middle value 0.05, which is determined as the weight for the nearest pixel from $\Omega$ in $U$. Thus, here $W = 0.05$, and $\sigma^2$ satisfy the following formula:

$$e^{(1/\sigma^2)/2\pi\sigma^2=0.05}$$

Thus, $\sigma^2$ is fixed, $\sigma^2 = 4.0697$.

The computation is operated within an $M \times M$ local window as shown in Figure 1. The saliency value $V(i,j)$ is calculated as follows:

- Extract an $M \times M$ local image patch $U$, $(i,j)$ is the centre pixel. $U$ is segmented from image $f$ as shown in Figure 1(a).
- In the centre of $U$, extract a disk-like region $\Omega$ with radius $r$.
- Using Equation (2), the saliency value of centre pixel is calculated for patch $U$, the value is treated as $V(i,j)$.

Thus, following the above three steps, $V(i,j)$ is obtained. To get the whole map $V$, we need to move the local window pixel-by-pixel. Proper size $M$ and $r$ would highlight salient objects, regions and pixels. And $M$ is an odd number. Finally, the function for $V$ could be shortly rewritten as $\text{func}(\cdot)$:

$$V = \text{func}(f, M, r)$$

The final saliency map or matrix $V$ needs to be normalized.

### 2.2 Analysis for local calculation

From Section 2.1, we have mentioned the local calculation and spatial weight. What is the role of the above two designs? We will briefly analyze this problem.

Figure 2 shows examples of spatial weight matrices, which are generated by Equation (3) and (4) with different $M$ and $r$. Figure 2(a)–(c) are spatial weight matrices with local window $(M=35, r=1; M=15, r=1; M=15, r=7)$, $\alpha=1$, respectively. We can find that large $M$ and small $r$ can create a proper spatial weight matrix, just like (a).

Examples of saliency maps are shown in Figure 3, which denotes our method and global idea [25], respectively. (a) denotes an original infrared image. (b)–(d) are saliency maps with local window $(M=35, r=1; M=15, r=1; M=15, r=7)$, $\alpha=1$, respectively. And (e) is the global method’s result. Those characteristics with different sizes can be well extracted with proper parameters under the local strategy, which performs well in the next image enhancement. And $r=1$ is the most proper for our enhancement. We will further discuss this in Section 3. Thus, Equation (5) is further improved:

$$V = \text{func}(f, M)$$

### 2.3 Enhancement

We can generate a saliency map $V$ by utilizing the above theory. And in our design, those areas and pixels with relatively salient values should be highlighted. With the selection of the proper size of the local window, potential targets and details with different sizes could be well extracted. Thus, we expect a
multi-window design to produce multi-saliency maps to achieve enhancement for the potential targets and details with different sizes.

Different target regions can be highlighted at different scales, which are determined by the selection of local window \( M \times M \). For these multiple windows, we assume \( n \) scales, and a local window at each scale \( k \) (\( 1 \leq k \leq n \)) is determined as follows:

\[
M_k = M_0 + k \times \Delta m
\] (7)

where \( M_0 \) represents the initial size or basic size, and \( \Delta m \) is the incremental size step. Thus, using multiple saliency analysis, the \( k \)th (\( 1 \leq k \leq n \)) saliency map could be produced according to Equation (6):

\[
V_k = \text{func} \left( f, M_k \right)
\] (8)

\( E_k \) denotes the intermediate result with the potential target region to be enhanced at the \( k \)th scale, and \( E_k \) is defined by:

\[
E_k = f \circ V_k
\] (9)

where \( \circ \) denotes component-wise multiplication operator. With changing of \( M_k \), potential target regions with different sizes are extracted for our enhancement. Furthermore, the intermediate result with enhanced target regions of multi-scale \( E_k \) can be obtained as follow:

\[
E_i = \max (E_1, E_2, ..., E_n)
\] (10)

Making use of Equation (10), we aim to enlarge the contrast of those potential target areas, and meanwhile, the general background information should also be kept. Finally, a fusion strategy is designed:

\[
E = \alpha_1 \times f + \alpha_2 \times E_i
\] (11)

where \( E \) denotes the final result we need, \( f \) is our original infrared image, \( E_i \) is from Equation (10). \( \alpha_1 \) and \( \alpha_2 \) are adjusting factors. Smaller \( \alpha_1 \) and larger \( \alpha_2 \) will produce a result with brighter or blacker salient regions, which is particularly useful for further target detection and recognition. But Smaller \( \alpha_1 \) and larger \( \alpha_2 \) will lead to the poor visual effect of background regions. So, the two factors should be balanced. Usually, \( \alpha_1 = \alpha_2 = 0.5 \) can balance the enhancement and visual effect.

### 2.4 Implementation

In Figure 4, we have shown a flow diagram of the proposed approach. Our method mainly contains three steps as follows:
Step 1. Generate Saliency maps for \( n \) scales according to Equation (8), referring to Section 2.1.

Step 2. Apply Equation (9) to produce images for enhancement of potential target regions. And obtain enhanced target regions of multi-scale utilizing Equation (10).

Step 3. Achieve enhancement with Equation (11).

In our method, the parameters need to be set including local window \( M \) (initial \( M_0 \), scales \( n \), and the incremental size step \( \Delta m \)) and \( r \).

3 | EXPERIMENTAL RESULTS AND ANALYSIS

We operate MATLAB R2017a on the personal computer with a dual-core CPU(2.2GHz) for the experiment.

The parameters \( M \) and \( r \) should be set. The size of the window \( M \) is closely related to the locality of the target in the infrared image. A very small \( M \) cannot highlight the differences between centre pixels and surrounding pixels, while a very large \( M \) cannot help analyze local regions. Moreover, since we need a central pixel for the local window, the window size \( M \) should be an odd number. The parameter \( r \) mirrors the difference between the pixel values within \( r \) radius and those of the rest. If \( r \) becomes larger, the saliency map will be more blurred. A blurred saliency map is good for large objects, while a less blurred one is more suitable for small targets in the infrared image. According to our experience, we set local window \( M \) (initial \( M_0 = 15 \), scales \( n = 3 \) and the incremental size step \( \Delta m = 10 \)) and \( r = 1 \). We will discuss them in Section 4.

To prove the effectiveness of the proposed algorithm, eight popular infrared enhancement approaches are selected for comparison, including, HE [6], radiative transfer and edge weight analysis (RTEWA) [14], CLAHE [9], multi-scale top-hat transform (MSTH) [15], adaptive histogram partition and brightness correction approach (AHPBC) [11], local edge preserving filter method (LDPF) [13], and two deep learning-based methods Kind [18] and Zero-DCE [20]. Because of the lack of enhanced versions of low-contrast infrared images, usually visible images are used for training for deep learning-based methods. In this paper, we just use the original data of the two approaches as they provide the codes and training process.

Meanwhile, to verify the nine algorithms, six images are adopted just as shown in Figure 5, which are collected from open-source datasets, especially from [14, 15, 26]. Those images contain different potential targets of different sizes. We have tried our best to adjust the parameters for those six methods to get the best performance. These potential targets include buildings, bridge, ships, vehicles and persons, etc. Figures 6–7 show the corresponding results.

3.1 | Subjective qualitative observation

HVS is one of the most valid methods to evaluate the quality of enhanced images. Thus, we analyze subjective qualitative observation of results produced by nine algorithms for six images.

As shown in Figure 5, there are six original images, which are collected from open-source datasets, especially from [14, 15, 26].

Figures 6 and 7 show source infrared image examples and the enhanced results. Every row represents different methods, and each column denotes different image examples. In Figure 6, (a0) (b0) (c0) are three source images named buildings on the other side, bridge and traffic, and road and buildings. In Figure 7, (a0) (b0) (c0) are three source images named a ship, a yacht, and two yachts. In both two figures, (a\(i\)) (b\(i\)) (c\(i\)) \( (i = 1, 2, \ldots, 9) \) are the corresponding results using RTEWA method, HE method, CLAHE method, MSTH method, AHPBC method, LDPF method, Kind method, Zero-DCE method, and the proposed algorithm, respectively.

Figure 6(a0) is the original image. There are some buildings by the river and a cargo ship on the river. Figures 6(a1)–(a9) show the enhanced images that corresponding to Figure 6(a0) using nine methods. According to these results, (a1) looks well with appropriate contrast, but with unclear details. And there are some shadows in (a1). In (a2) and (a3), the methods of HE and CLAHE leads to worse visual effects, as some regions seem over-enhanced. (a4) is better than (a2) and (a3), which means
MSTH performs better than HE and CLAHE. But MSTH fails to enhance some buildings. The result of AHPBC produces a good cargo ship but over enhanced buildings and LDPF make the image be with low contrast, and the details are greatly smoothed. (a8) produces a result with poor contrast. And (a7) seems better than (a8), but it has produced background noise. The result of (a9) looks the best, as the proposed algorithm makes a good balance between enhancement of potential targets and the suppression of the background.

In Figure 6(b0), potential targets consist of bridges, vehicles, and buildings. And there exist both small and large target regions with rich details. The enhanced results are shown in (b1)–(b9). In Figures 6(b2) and (b5), HE and AHPBC algorithms have over enhanced some areas. According to (b6), LDPF’s result indicates that the details are greatly smoothed with low contrast. Meanwhile, the approach of CLAHE leads to a worse visual effect with noise amplification. Images (b1) and (b7) look better, the two have improved contrast for details. From the image (b9), the proposed method also performs well, enhancing potential targets and improving the contrast.

Figure 6(c0) has several potential targets, which include four people and details of buildings. The enhanced results are shown in Figures 6(c1)–(c9). In this figure, small targets are more prominent, especially the four people. The methods of HE and CLAHE produce much noise, leading to uncomfortable visual effects in (c2) and (c3). (c4) is better than (c2) and (c3), but the road has grey distortion, and some areas are overhighlighted. In (c6), the contrast looks low and the details are greatly smoothed. (c8) produces a result with poor contrast. Image (c1) looks better than (c5) and (c6), while the contrast is unsatisfactory enough. Image (c7) also produces a good result, but with some background noises. For our algorithm, (c9) keeps more details than (c1) and (c7), and enhances target regions better. The results of the proposed method achieve the best visual effects.

Figure 7(a0) contains a ship with abundant details. And Figures 7(a1)–(a9) show the results of all nine methods. The HE and CLAHE algorithms both have enhanced potential target areas. However, they still created a lot of noises, especially (a2) is over enhanced. According to image (a4), the MSTH method well enhanced the ship, but it also produced a distorted background such as sea spray. The result of LDPF in (a6) has low contrast. The background of (a1) seems too dim. The results of (a5) and (a9) have better visual effects. (a8) seems a little low contrast. (a7) also looks very good with good contrast and abundant details.

The potential target in Figure 7(b0) is a yacht, which is a small target compared to the previous image. But according to nine methods of results in Figures 7(b1)–(b9), the difference is even starker. For methods of HE, CLAHE, and LDPF, images (b2) (b3) and (b6) enhance target uncomfortably, we even cannot find the yacht. In image (b4), the yacht is enhanced greatly, while the background is fuzzy. And in (b5) and (b7), it produced a lot of noises. (b8) creates a low contrast result. (b9) is better than (b1) by contrast. The proposed method preserves more details and achieves a better visual effect than the others methods.

The potential targets of Figure 7(c0) are two small assault boats. And the two targets look too vague to distinguish in the original image. So the targets need to be greatly enhanced. All results of the seven algorithms are shown and compared in Figures 7(c1)–(c9). In (c2), the HE method creates an over-enhanced result. According to (c3) and (c6), the results of CLAHE and LDPF cannot highlight the targets enough. In (c4) and (c7), the background has been amplified. (c1), (c5), and (c9) perform better than the others. The contrast of (c9) looks the best and strikes a good balance between potential target enhancement and background suppression.

3.2 Further subjective experiment

In addition to the above intuitive visual performance assessment, a subjective evaluation experiment is needed to evaluate the enhancement effect. Learning from Ref. [27], our
**FIGURE 6** Source infrared image examples and the enhanced results. Every row represents different methods, and each column denotes different image examples. (a0) (b0) (c0) are three source images named buildings on the other side, bridge and traffic, and road and buildings. (a1) (b1) (c1) (i = 1, 2…, 9) are the corresponding results using RTEWA method, HE method, CLAHE method, MSTH method, AHPBC method, LDPF method, KinD method, Zero-DCE method, and the proposed algorithm, respectively.

**FIGURE 7** Source infrared image examples and the enhanced results. Every row represents different methods, and each column denotes different image examples. (a0) (b0) (c0) are three source images named a ship, a yacht, and two yachts. (a1) (b1) (c1) (i = 1, 2…, 9) are the corresponding results using RTEWA method, HE method, CLAHE method, MSTH method, AHPBC method, LDPF method, KinD method, Zero-DCE method, and the proposed algorithm, respectively.
subjective metric includes four factors: Description of infrared image (DIR), Contrast (Con), Edge Feature (EF) and Background Noise (BN). DIR denotes the similarity of characteristic information between the result and source images. Con means the contrast level of the result, and we consider the contrast of the whole image as well as local regions. EF represents the sharpness of edges and contours of objects. EF could be noticed as target details enhancement. BN could mirror the negative facts of the enhancement method, especially the noise level. We set four levels to describe the worst to the best for these four factors to evaluate the enhanced results. '1' means very bad, '2' means not good, '3' means acceptable, but it is not very good, '4' means very good. For every person, the source image and results produced by all methods are shown together to give a score based on the above four factors. So, we could get an average score of each factor for each enhanced image.

We invite a total of 50 people in our college to join in our research. Our subjective results for Figures 6 and 7 are given as following Tables 1 and 2. Table 1 lists subjective evaluation for Figure 6(a)(b)(c) (i = 1, 2,...,9), and Table 2 shows subjective evaluation for Figure 7(a)(b)(c) (i = 1, 2,...,9). From our design, every figure could get a score between 1 and 4. A higher score means better subjective assessment.

Learning from Table 1, we can analyze the four factors. For the ‘DIR’ factor, HE and the proposed method are better. Our approach performs best in ‘Con’ comparison. In the ‘EF’ factor, RTEWA and our method seem better. And the proposed algorithm produces the best results in the ‘BN’ comparison. Overall, the proposed method is the best in subjective evaluation.

According to Table 2, we can also analyze the four factors. For the ‘DIR’ factor, the proposed method performs best. RTEWA and our approach are outstanding in the ‘Con’ comparison. In the ‘EF’ factor, AHPBC and our method get better scores. In the ‘BN’ comparison, the proposed method, RTEWA, and MSTH are the top three. In general, the proposed approach performs best in subjective evaluation.

### Table 1
| Methods  | Figures/ Figure 6 | DIR | Con | EF | BN |
|----------|-------------------|-----|-----|----|----|
| RTEWA (a1) | 3.2 | 2.7 | 3.4 | 3.1 |
| (b1) | 2.5 | 2.2 | 3.5 | 3.2 |
| (c1) | 3.1 | 2.4 | 3.1 | 3.4 |
| HE (a2) | 3.5 | 3.1 | 2.5 | 1.5 |
| (b2) | 3.4 | 2.5 | 2.5 | 1.6 |
| (c2) | 3.4 | 2.5 | 2.1 | 1.4 |
| CLAHE (a3) | 2.4 | 2.2 | 3.7 | 2.0 |
| (b3) | 2.6 | 2.1 | 3.0 | 2.5 |
| (c3) | 1.9 | 1.9 | 2.5 | 2.2 |
| MSTH (a4) | 2.1 | 2.5 | 3.5 | 2.9 |
| (b4) | 2.2 | 2.5 | 3.0 | 3.1 |
| (c4) | 2.3 | 2.1 | 2.5 | 2.3 |
| AHPBC (a5) | 2.5 | 3.9 | 3.5 | 2.5 |
| (b5) | 3.1 | 2.7 | 3.1 | 3.0 |
| (c5) | 3.0 | 2.6 | 2.9 | 2.8 |
| ADPF (a6) | 3.1 | 2.2 | 2.7 | 3.7 |
| (b6) | 2.5 | 2.2 | 2.5 | 3.4 |
| (c6) | 2.4 | 2.1 | 2.4 | 3.6 |
| KinD (a7) | 2.6 | 2.1 | 2.6 | 2.2 |
| (b7) | 2.1 | 1.5 | 2.7 | 2.3 |
| (c7) | 2.2 | 2.1 | 2.3 | 2.3 |
| Zero-DCE (a8) | 1.4 | 1.5 | 1.5 | 2.6 |
| (b8) | 1.1 | 1.6 | 2.2 | 2.2 |
| (c8) | 1.7 | 1.8 | 1.7 | 2.1 |
| Proposed (a9) | 3.1 | 3.1 | 3.2 | 3.9 |
| (b9) | 3.5 | 3.2 | 3.5 | 3.8 |
| (c9) | 3.1 | 3.3 | 3.2 | 3.6 |

### Table 2
| Methods  | Figure 7 | DIR | Con | EF | BN |
|----------|----------|-----|-----|----|----|
| RTEWA (a1) | 3.4 | 3.6 | 3.5 | 3.4 |
| (b1) | 2.5 | 3.5 | 2.1 | 3.6 |
| (c1) | 2.4 | 2.2 | 3.1 | 2.4 |
| HE (a2) | 3.5 | 2.1 | 3.0 | 2.3 |
| (b2) | 2.4 | 1.6 | 1.5 | 1.1 |
| (c2) | 3.1 | 2.2 | 2.1 | 1.2 |
| CLAHE (a3) | 2.5 | 2.3 | 2.5 | 3.4 |
| (b3) | 1.5 | 1.5 | 1.2 | 2.5 |
| (c3) | 1.4 | 1.5 | 1.5 | 2.5 |
| MSTH (a4) | 2.5 | 3.5 | 3.1 | 3.5 |
| (b4) | 2.1 | 2.1 | 1.5 | 2.5 |
| (c4) | 2.2 | 2.1 | 2.5 | 2.2 |
| AHPBC (a5) | 2.4 | 3.5 | 3.5 | 3.1 |
| (b5) | 2.1 | 2.5 | 3.0 | 3.5 |
| (c5) | 1.5 | 2.2 | 2.5 | 2.5 |
| ADPF (a6) | 2.5 | 2.5 | 3.0 | 3.5 |
| (b6) | 2.0 | 1.5 | 2.2 | 3.1 |
| (c6) | 1.5 | 1.5 | 2.1 | 3.0 |
| KinD (a7) | 2.6 | 3.1 | 2.5 | 2.5 |
| (b7) | 1.5 | 2.0 | 2.5 | 1.5 |
| (c7) | 2.3 | 2.1 | 2.5 | 1.5 |
| Zero-DCE (a8) | 2.5 | 2.2 | 2.5 | 2.5 |
| (b8) | 1.3 | 1.1 | 1.2 | 2.1 |
| (c8) | 1.2 | 1.2 | 1.3 | 1.9 |
| Proposed (a9) | 3.5 | 3.5 | 3.5 | 3.5 |
| (b9) | 3.1 | 3.2 | 2.5 | 3.5 |
| (c9) | 3.3 | 2.6 | 3.2 | 3.5 |
3.3 Objective evaluation

We introduce six objective evaluation methods, including linear index of fuzziness (LIF) [28], measure of enhancement (EME) [29], information entropy (Entropy) [30], the lightness order error (LOE) [31], natural image quality evaluator (NIQE) [32] and tone-mapped image quality index (TMQI) [33]. These could be used for objective evaluation for infrared enhancement.

Next, the derivation process and definition of each parameter are introduced in detail. LIF metric γ is defined as follows:

\[
γ(f) = \frac{2}{WH} \sum_{x=1}^{W} \sum_{y=1}^{H} \min \{ p(x,y)(1 - p(x,y)) \} \tag{12}
\]

\[
p(x,y) = \sin \left( \frac{\pi}{2} \left( 1 - \frac{f(x,y)}{\max(f)} \right) \right) \tag{13}
\]

where \(f(x,y)\) denotes the grey value at pixel \((x, y)\), \(\min()\), and \(\max()\) are operators to calculate the minimum and maximum value, respectively. In Equation (12), \(W\) and \(H\) represent the width and height of the image (unit: pixel), respectively. The smaller the value, the image contains more information, thus the measure the information the image contains. When the entropy ill-problem.

The measure of enhancement (EME) evaluates the dynamic range, which indicates a better quality of the enhanced image.

\[
EME = \frac{1}{N} \sum_{q=1}^{N} 20 \ln \frac{f_{\text{max},B(q)}}{f_{\text{min},B(q)} + \xi} \tag{14}
\]

where \(f_{\text{max},B(q)}\) and \(f_{\text{min},B(p,q)}\) represents the maximum and minimum of \(B(q)\), respectively. \(\xi\) is a tiny constant to avoid ill-problem.

The third metric is called information entropy, which usually measures the information the image contains. When the entropy value is larger, the image contains more information, thus the image is better. The entropy is defined by:

\[
Entropy = -\sum_{i=0}^{L-1} p_f(i) \log_2 (p_f(i)) \tag{15}
\]

where \(L\) represents \(f\)’s grey level, \(i\) is grey value. For example, if \(f\) is an 8-bit-depth image, the grey level \(L = 2^8 = 256\). \(p_f(i)\) means the probability for \(i\) in the whole image \(f\).

The LOE metric is based on the lightness order error between the original image and its enhanced version. The metric is defined as:

\[
LOE = \frac{1}{WH} \sum_{x=1}^{W} \sum_{y=1}^{H} RD(x,y) \tag{16}
\]

where \(W\) and \(H\) represent width and height of the image (unit: pixel), \((x, y)\) denotes pixel coordinate. And RD represents a difference matrix of the lightness. More definition could be referred to Ref. [31]. The smaller the LOE value is, the better the lightness order as well as the enhanced result.

The NIQE evaluates the quality of the image by measuring the distance between the quality-aware natural scene statistic (NSS) feature model and the multivariate Gaussian (MVG) fit the features. The expression of NIQE is complex, and more details could be referred to Ref. [32]. A smaller NIQE value means better quality of the enhanced image.

By combing a multiscale signal fidelity measure and a naturalness measure, the TMQI is defined as:

\[
Q = \alpha S^2 + (1 - \alpha) N^2 \tag{17}
\]

where \(\alpha \in [0, 1]\) is used to adjust the relative importance of the two terms, and the sensitivity is determined by \(\alpha \beta\). \(N\) is the statistical naturalness measure, and \(S\) is the overall structural fidelity [33].

These six objectives metrics are all utilized to evaluate all results of the above six images processed by nine algorithms. Since there is lots of data, we have to list four tables. Tables 3 and 5 show the calculated former three metrics (LIF, EME, and entropy) for Figures 6 and 7, respectively. Tables 4 and 6 show the calculated latter three metrics (LOE, NIQE, and TMQI) for Figures 6 and 7, respectively.

In Table 3, smaller LIF, larger EME, and larger entropy denote the better image quality, respectively. Learning from Table 3, we can analyze and compare the proposed method with other ones. The proposed method gets the smallest LIF value, and it indicates that our algorithm is the best. For EME, the proposed method is the second best. HE method looks like the better results in the EME metric. But through the observation of subjective evaluation, HE creates overenhanced results. Finally, the entropy metric indicates that the proposed method achieves almost the second-best performance by comparison, while KinD wins the best performance.

In Table 4, smaller LOE, smaller NIQE, and larger TMQI indicate better image quality, respectively. For LOE and NIQE, HE seems the best. We think the overenhanced image affects these two metrics. LDOP also performs well. The proposed method wins the third best. In TMQI comparison, our method performs the best in this objective evaluation.

According to the results of six metrics for Figure 6 in Tables 3 and 4, each method has its own advantages, but the proposed one obtains the overall best.

In Table 5, smaller LIF, larger EME, and larger entropy denote the better image quality, respectively. Learning from Table 3, the proposed method and RTEWA are the top two in LIF comparison. For EME, the Zero-DCE method seems the best, and many methods create results worse than the source image. The proposed method performs not well enough. In entropy comparison, Zero-DCE performs the best, and our method could be listed in the middle.
### TABLE 3
The calculated former three metrics for Figure 6, smaller LIF, larger EME, and larger entropy indicate the better image quality, respectively.

| Image Metric | Figure 6(a0)–(a9) | Figure 6(b0)–(b9) | Figure 6(c0)–(c9) |
|--------------|------------------|------------------|------------------|
|              | LIF | EME | Entropy | LIF | EME | Entropy | LIF | EME | Entropy |
| Source image | 0.6184 | 6.6149 | 5.623 | 0.8213 | 4.142 | 5.3409 | 0.7892 | 6.1692 | 4.8043 |
| RTEWA        | 0.653 | 8.2148 | 6.4001 | 0.1008 | 20.15 | 6.2433 | 0.5245 | 13.699 | 5.7396 |
| HE           | 0.3924 | 58.063 | 5.389 | 0.3923 | 73.43 | 5.1831 | 0.3939 | 72.2317 | 4.6255 |
| CLAHE        | 0.5492 | 14.026 | 6.3187 | 0.6265 | 11.41 | 6.2183 | 0.7306 | 9.2005 | 6.1103 |
| MSTH         | 0.2456 | 10.441 | 5.9177 | 0.6931 | 6.367 | 5.6358 | 0.459 | 15.7623 | 4.8847 |
| AHPBC        | 0.3403 | 15.162 | 5.6174 | 0.3844 | 12.88 | 5.3378 | 0.6993 | 20.2702 | 4.8116 |
| LDPF         | 0.6231 | 5.2085 | 5.9085 | 0.5151 | 3.416 | 4.5119 | 0.8521 | 5.9611 | 5.23 |
| KinD         | 0.1990 | 19.767 | 6.5509 | 0.1306 | 22.35 | 6.2879 | 0.3214 | 18.5455 | 5.9314 |

### TABLE 4
The calculated latter three metrics for Figure 6, smaller LOE, smaller NIQE, and larger TMQI indicate the better image quality, respectively.

| Image Metric | Figure 6(a0)–(a9) | Figure 6(b0)–(b9) | Figure 6(c0)–(c9) |
|--------------|------------------|------------------|------------------|
|              | LOE | NIQE | TMQI | LOE | NIQE | TMQI | LOE | NIQE | TMQI |
| Source image | – | 6.3771 | – | – | 7.6329 | – | – | 39.7993 | – |
| RTEWA        | 487.8785 | 6.0262 | 0.7399 | 271.724 | 5.303 | 0.802 | 87.476 | 39.7997 | 0.8549 |
| HE           | 57.502 | 4.8370 | 0.7883 | 68.857 | 5.8409 | 0.7818 | 702.8603 | 39.8021 | 0.8074 |
| CLAHE        | 773.2338 | 5.2426 | 0.7654 | 803.632 | 5.7834 | 0.7818 | 702.8603 | 39.8015 | 0.8074 |
| MSTH         | 101.2647 | 5.4187 | 0.7871 | 75.302 | 5.6369 | 0.7818 | 92.49 | 39.8011 | 0.8052 |
| AHPBC        | 179.9173 | 4.6132 | 0.7872 | 220.224 | 5.325 | 0.7494 | 142.4 | 39.8011 | 0.8005 |
| LDPF         | 200.8074 | 4.5320 | 0.7839 | 160.664 | 4.8317 | 0.8064 | 111.3887 | 39.8011 | 0.8552 |
| KinD         | 464.616 | 6.6361 | 0.7636 | 331.942 | 7.9413 | 0.795 | 348.726 | 39.8 | 0.8334 |
| Zero-DCE     | 75.5827 | 4.8947 | 0.8003 | 40.236 | 5.5361 | 0.8203 | 63.4173 | 39.8004 | 0.8556 |

### TABLE 5
The calculated former three metrics for Figure 7, smaller LIF, larger EME, and larger entropy indicate the better image quality, respectively.

| Image Metric | Figure 7(a0)–(a9) | Figure 7(b0)–(b9) | Figure 7(c0)–(c9) |
|--------------|------------------|------------------|------------------|
|              | LIF | EME | Entropy | LIF | EME | Entropy | LIF | EME | Entropy |
| Source image | 0.5285 | 6.5134 | 5.7124 | 0.6901 | 2.011 | 3.8989 | 0.5944 | 1.8563 | 4.1945 |
| RTEWA        | 0.1525 | 26.236 | 5.8848 | 0.0493 | 25.21 | 5.457 | 0.2532 | 13.5635 | 6.7055 |
| HE           | 0.5136 | 68.595 | 5.3493 | 0.409 | 81.59 | 3.8592 | 0.5753 | 44.0436 | 7.8501 |
| CLAHE        | 0.5987 | 10.953 | 6.5076 | 0.7291 | 6.491 | 5.1854 | 0.7811 | 6.0548 | 4.918 |
| MSTH         | 0.306 | 10.316 | 5.5338 | 0.669 | 3.674 | 4.2963 | 0.398 | 6.5195 | 5.0871 |
| AHPBC        | 0.4512 | 9.9796 | 5.7068 | 0.3731 | 11.21 | 3.9049 | 0.5109 | 11.0196 | 4.2045 |
| LDPF         | 0.4759 | 2.6753 | 4.7462 | 0.825 | 2.044 | 4.1212 | 0.869 | 1.7709 | 3.8174 |
| KinD         | 0.5323 | 11.669 | 6.4174 | 0.6043 | 15.22 | 6.8841 | 0.8107 | 9.4302 | 6.4106 |
| Zero-DCE     | 0.8171 | 3.3258 | 5.1208 | 0.3727 | 1.137 | 3.0994 | 0.3373 | 1.1029 | 3.3313 |
| Proposed     | 0.2092 | 19.371 | 5.8621 | 0.047 | 23.12 | 5.094 | 0.1161 | 14.7259 | 5.2632 |
In Table 6, smaller LOE, smaller NIQE, and larger TMQI indicate better image quality, respectively. For LOE, HE and the proposed method perform the best. In the NIQE comparison, HE looks the best, and our method could be listed in the middle. For the TMQI metric, our algorithm is the best one.

According to the results of six metrics for Figure 7 in Tables 5 and 6, each method has its advantages. Overall, the proposed method and Zero-DCE are the best two.

Because of too much data, we will sometimes misread these tables. Learning from Ref. [34], we consider box plots for all images used, and then performance measure results can be observed better. The numerical results have been given visually in Figure 8. In this figure, (a)–(f) are LIF, EME, entropy, LOE, NIQE, and TMQI, respectively. The source image could be used as a reference. When we compare the results for each method, the source image will be ignored.

As shown in Figure 8(a), the smallest LIF value belongs to the proposed method. Meanwhile, the proposed method also has the smallest mean value of LIF. It indicates the best performance of our approach. Figure 8(b) shows the box plot on the EME of the test images. The Zero-DCE method obtains the best performance. And the proposed method ranks seven in this category. The entropy box plot has been shown in Figure 8(c). In Figure 8(c), KinD obtains the highest value, and the proposed method wins the third highest among other enhancement methods. LOE measures have been shown in Figure 8(d). MSTH and the proposed method have the smallest LOE, which mirrors the two that have the best performance. The NIQE box plot has been shown in Figure 8(e). In Figure 8(e), HE and AHPBC seem the best, but the proposed method, KinD, MSTH, and CLAHE are very close to them. Figure 8(f) shows a box plot of TMQI. The proposed method, RTEWA, HE, MSTH, and LDPF have the largest value of TMQI. Furthermore, according to the mean value of TMQI, the proposed approach ranks one. Overall, the box plot proves the relative superiority of the proposed method.

### 3.4 Efficiency

For efficiency comparison, we also have made a table to analyze different methods. From Table 7, the HE and CLAHE are the first and second most efficient methods. They could rapidly finish the enhancement. The proposed method ranks three among these nine algorithms. The RTEWA method is close to the proposed method. The MSTH, LDPF, and KinD run slowly, and Zero-DCE and AHPBC are much slower. Take Figure 6(a) (i = 1, 2, …, 9) for example, we consider the computational time of the proposed method as one unit of time, so all the data of this column should be divided by 0.03657. Thus, we obtain RTEWA (1.012), HE (0.05), CLAHE (0.30), MSTH (40.10), AHPBC (150.65), LDPF (51.79), KinD (47.25), Zero-DCE (68.76), ours (1.00). From the above comparison, we could find the differences in efficiency between different methods, especially the difference in the order of magnitude. By comparison, our method is relatively fast. The current speed can be used in a real system with algorithm acceleration.

### 4 Discussion

#### 4.1 Parameters

From Section 2.1, we have mentioned the local calculation and spatial weight. According to Section 2.4, to implement our method, we need to select several important parameters, including the size of the local window $M$ and $r$.

Since the window area needs to have a central pixel, the window size $M$ is generally an odd number. The size of the window $M$ is closely related to the locality of the target in the infrared image: If the window is too small, it cannot quantitatively reflect the differences of the pixels around the centre, and if the window is too large, it cannot reflect the information characteristics of the local area. The local difference value $r$ refers to the difference between the pixel values in the $r$ radius.
FIGURE 8  Box plots of performance measure results. (a) LIF, (b) EME, (c) entropy, (d) LOE, (e) NIQE, (f) TMQI
TABLE 7  The computational time corresponding to Figures 6–7. Unit: Second(s). \(i = 1, 2, \ldots, 9\)

| Figure/size | Figure 6(a)/380×236 | Figure 6(b)/380×236 | Figure 6(c)/123×103 | Figure 7(a)/255×203 | Figure 7(b)/250×204 | Figure 7(c)/258×216 |
|-------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| RTEWA       | 0.03728              | 0.02721              | 0.01155              | 0.01253              | 0.02454              | 0.02976              |
| HE          | 0.00174              | 0.00146              | 0.00747              | 0.0016               | 0.00095              | 0.00225              |
| CLAHE       | 0.01111              | 0.02437              | 0.06273              | 0.01135              | 0.01014              | 0.01977              |
| MSTH        | 1.4668               | 1.57803              | 0.35874              | 0.90351              | 0.91154              | 0.90716              |
| AHPBC       | 5.50934              | 5.44219              | 0.39154              | 2.46415              | 2.13271              | 3.50435              |
| LDPF        | 1.89399              | 1.94729              | 0.32528              | 1.1747               | 1.24216              | 1.2010               |
| KinD        | 1.72810              | 1.2660               | 0.21400              | 0.78000              | 0.75400              | 0.80300              |
| Zero-DCE    | 2.51490              | 2.51850              | 0.36600              | 1.47710              | 1.45610              | 1.56810              |
| Proposed    | 0.03657              | 0.03894              | 0.01187              | 0.02522              | 0.02522              | 0.02679              |

FIGURE 9  Saliency maps with different \(r (r = 1, 3, 5, 7, 9 \text{ and } 11), M = 35\)

and the rest of the pixel values. The larger the difference value is, the more unable to accurately highlight the information characteristics of the local area. According to our experimental experience, when \(r\) becomes larger, the saliency map will be ambiguous, and we find the best \(r = 1 \text{ or } 2\). As shown in Figure 9(a) is an original image, (b)–(g) are based on different \(r = 1, 3, 5, 7, 9 \text{ and } 11\) when \(M = 35\). From the results, it can be concluded that the value of \(r\) becomes larger, the saliency map becomes more blurred, the more unable to accurately highlight the information characteristics of the local area, and the more blurred the enhancement of small targets becomes, the outline of the characters in Figures 9(b)–(g) are getting blurred. We will discuss the effect on the infrared image results when \(r\) changes later.

When \(M\) or \(r\) changes, the infrared image result will also change. As shown in Figure 10, we use six images for experiments and use the objective metric LIF to evaluate the enhanced results. The relationship between LIF and \(M\) is shown in Figure 11(a). Since a smaller LIF indicates better image quality, the optimal \(M = 35\). When \(M\) becomes larger, the external image enhancement result becomes worse. The objective measure LIF is used to evaluate the results. The relationship between LIF and \(r\) is shown in Figure 11(b). Because a smaller LIF indicates better image quality, the optimal \(r = 1 \text{ or } 2\). When \(r\) is large, the external image enhancement result will be worse.

From the above analysis, we set local window \(M\) (initial \(M_0 = 15\), scales \(s = 3\) and the incremental size step \(\Delta m = 10\), so \(M = 15, 25, 35\) and \(r = 1\).

4.2  Further attempts in a hardware system

To further attempt our algorithm, we are expected to apply it on the hardware system. This method is coded and operated on FPGA (Field Programmable Gate Array) platform. Figure 12 shows the setup of the whole system, including
infrared camera (infrared lens, camera component), FPGA (FPGA-based processing circuit), data transmission, and output display. The camera is a long wave infrared (LWIR) camera (infrared lens and camera component), which is bought from Zhejiang Zhao Sheng Technology CO., LTD (http://www.suncti.com/). This uncooled 384×288 IRFPA camera’s detector pixel size is 17 um, and it works at 8–14 um and captures 30 frames per second. In our setup, the focal length is 10 mm. This camera only needs two dry batteries to work.

The FPGA could fast acquire the frames with the help of the company. The main work is the code optimization on FPGA. Beforehand, we have stored some spatial weight matrixes for the local window to accelerate. Meanwhile, the scale number $n$ is set below 3. And some approximate is done, such as the value of $\pi$ and $\sigma^2$.

Until now, this algorithm could achieve about 20–30 frames per second in this system. In our application, it meets the requirement of real time. But, there is potential to accelerate our method for further application, especially for high-resolution cameras. Moreover, the whole system can be further assembled and integrated to be a separate module.

5 | CONCLUSION

In this paper, we have designed a local saliency extraction-based infrared enhancement algorithm. Saliency analysis is done within a local window with spatial weight. And those potential targets and details with different sizes can be all enhanced as the size of the local window changes. Utilizing
multi-window saliency extraction, we achieve enhancement through fusion at different scales. Eight popular infrared enhancement approaches are introduced for comparison. The subjective qualitative observation indicates the results of the proposed method achieve the best visual effects, meanwhile, the subjective experiment also proves it. Six objective evaluation methods are selected to measure the results and a mass of evaluation data proves the effectiveness of the proposed algorithm. Meanwhile, a box plot is introduced to observe and compare easily, the graphs of six objective evaluation methods demonstrate the relative superiority of the proposed method. For efficiency comparison, the proposed method ranks three among nine approaches. The parameters are also discussed, and we have given an illustration that indicates proper $M$ and small $r$ should be adopted. The experiments and discussions demonstrate that local saliency-based strategy could well extract feature information with different sizes of source images to obtain a good enhanced result.

Since our method can obtain a good enhanced result with one single image, the algorithm is suitable for single image processing and application, such as further target detection, scene surveillance, and other relative fields, which would be our next try.

One of the key problems is the automation of the algorithm, especially how to select the size of the local window. In the future, the size of the local window is supposed to be self-adaptive by automatically analyzing image content.

We have tested the proposed algorithm in the hardware system. But it is limited to apply in high-speed imaging systems. Thus, how to accelerate our method will be also focused on.

**ACKNOWLEDGEMENT**

We thank the reviewers for helping us in improving this paper. We also thank Bob and Peter for helping us in improving the paper. This work is supported by Zhejiang Provincial Key Lab of Equipment Electronics.

**ORCID**

Jufeng Zhao https://orcid.org/0000-0002-4491-5566

**REFERENCES**

1. Fan, J., Yang, J.: Development trends of infrared imaging detecting technology. Infrared Laser Eng. 41(12), 3145–3153 (2012)
2. Francesco, C., et al.: Recent advances in active infrared thermography for non-destructive testing of aerospace components. Sensors 18(2), 609 (2018)
3. Xu, F., et al.: Detail enhancement of blurred infrared images based on frequency extrapolation. Infrared. Phys. Technol. 76, 560–568 (2016)
4. Sheng, Z., et al.: Image enhancement based on wavelet transformation and pseudo-color coding with phase-modulated image density processing. Infrared. Phys. Technol. 58, 56–63 (2013)
5. Qi, Y., et al.: Novel infrared image enhancement technology based on the frequency compensation approach. Infrared. Phys. Technol. 76, 521–529 (2016)
6. Stark, J.A.: Adaptive image contrast enhancement using generalizations of histogram equalization. IEEE Trans. Image Process. 9(5), 889–896 (2000)
7. Wong, C.Y., et al.: Histogram equalization and optimal profile compression based approach for colour image enhancement. J. Visual Commun. Image Represent. 38, 802–813 (2016)
8. Wang, Y., Pan, Z.: Image contrast enhancement using adjacent-blocks-based modification for local histogram equalization. Infrared. Phys. Technol. 86, 59–65 (2017)
9. Mohan, S., Ravishankar, M.: Modified contrast limited adaptive histogram equalization based on local contrast enhancement for mammogram images. Mobile Commun. Power Eng. Commun. Comput. and Inf. Sci. 296, 397–403 (2013)
10. Huang, Z., et al.: Adaptive gamma correction based on cumulative histogram for enhancing near-infrared images. Infrared. Phys. Technol. 79, 205–215 (2016)
11. Wán, M., et al.: Infrared image enhancement using adaptive histogram partition and brightness correction. Remote Sens. 10(5), 682 (2018)
12. Ashiba, H.I., Ashiba, M.I.: Super-efficient enhancement algorithm for infrared night vision imaging system. Multimed. Tools Appl. 80, 1–27 (2020)
13. Song, Q., et al.: High dynamic range infrared images detail enhancement based on local edge preserving filter. Infrared. Phys. Technol. 77, 464–473 (2016)
14. Li, Q., et al.: Infrared enhancement for water surface imaging based on theory of radiative transfer and edge weight analysis. IEEE Access 7, 175051–175061 (2019)
15. Bai, X., et al.: Image enhancement using multi scale image features extracted by top-hat transform. Opt. Laser Technol. 44(2), 328–336 (2012)
16. Babatunde, I.G., et al.: A multi-level model for fingerprint image enhancement. In: Proceedings of the International Multi Conference of Engineers and Computer Scientists, (2012)
17. Abi-Abdallah, D., et al.: Cardiac and respiratory MRI gating using combined wavelet sub-band decomposition and adaptive filtering. Ann. Biomed. Eng. 35(5), 733–743 (2007)
18. Zhang, Y., et al.: Kindling the darkness: A practical low-light image enhancer. In: Proceedings of the 27th ACM International Conference on Multimedia, ACM, Nice, France (2019)
19. Kuang, X., et al.: Single infrared image enhancement using a deep convolutional neural network. Neurocomputing 332, 119–128 (2019)
20. Guo, C., et al.: Zero-reference deep curve depth estimation for low-light image enhancement. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020
21. Cong, R., et al.: Review of visual saliency detection with comprehensive information. IEEE Trans. Circuits Syst. Video Technol. 29(10), 2941–2959 (2018)
22. Dong, L., et al.: Robust infrared maritime target detection based on visual attention and spatiotemporal filtering. IEEE Trans. Geosci. Remote Sens. 55(5), 3037–3050 (2017)
23. Wan, M., et al.: Infrared small target enhancement: Grey level mapping based on improved sigmoid transformation and saliency histogram. J. Mod. Opt. 65(10), 1161–1179 (2018)
24. Koch, C., Ullman, S.: Shifts in selective visual attention: Towards the underlying neural circuitry. Hum. Neurobiol. 4(4), 219–227 (1985)
25. Cheng, M., et al.: Global contrast based salient region detection. IEEE Trans. Pattern Anal. Mach. Intell. 37(3), 569–582 (2014)
26. Zhao, J., et al.: Infrared image enhancement through saliency feature analysis based on multi-scale decomposition. Infrared. Phys. Technol. 62, 86–93 (2014).
27. Zhao, J., et al.: Multi-window visual saliency extraction for fusion of visible and infrared images. Infrared. Phys. Technol. 76, 295–302 (2016).
28. Kaufmann, A., Swanson, D.L., Introduction to the Theory of Fuzzy Subsets. vol. 6, pp. 29–30. : Academic Press, New York, (1975)
29. Agaian, S.S., et al.: Transform coefficient histogram-based image enhancement algorithms using contrast entropy. IEEE Trans. Image Process. 16, 741–758 (2007)
30. Roberts, J.W., et al.: Assessment of image fusion procedures using entropy, image quality, and multispectral classification. J. Appl. Remote Sens. 2(1), 023522 (2008)
31. Wang, S., et al.: Naturalness preserved enhancement algorithm for non-uniform illumination images. IEEE Trans. Image Process. 22(9), 3538–3548 (2013)
32. Mittal, A., et al.: Making a 'completely blind' image quality analyzer. IEEE Signal Process Lett. 20(3), 209–212 (2013)
33. Yeganeh, H., Wang, Z.: Objective quality assessment of tone-mapped images. IEEE Trans. Image Process. 22(2), 657–667 (2013)
34. Katrcoğlu, F.: Colour image enhancement with brightness preservation and edge sharpening using a heat conduction matrix. IET Image Proc. 14(13), 3202–3214 (2020)

How to cite this article: Li T, Zhao J, Wu X, Mao H, Cui G. Infrared imaging enhancement through local window-based saliency extraction with spatial weight. IET Image Process. 2021;1–16. https://doi.org/10.1049/ipr2.12276.