文章编号：1001-9014(2022)05-0930-11

DOI:10.11972/j.issn.1001-9014.2022.05.018

Infrared image dehazing based on hierarchical subdivision superpixels and information integrity prior

LI Wei-Hua[1,2,3], LI Fan-Ming[1,3*], MIAO Zhuang[1,2,3], TAN Chang[1,2,3], MU Jing[1,2,3]

(1. CAS Key Laboratory of Infrared System Detection and Imaging Technology, Shanghai Institute of Technical Physics, Shanghai 200083, China;
2. University of Chinese Academy of Sciences, Beijing 100049, China;
3. Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China)

Abstract: Hazy weather degrades the contrast and visual quality of infrared imaging systems due to the presence of suspended particles. Most existing dehazing methods focus on enhancing global contrast or exploit a local grid transmission estimation strategy on images, which may lead to loss of information, halo artifacts and distortion in sky region. To address these problems, a novel single image dehazing model based on superpixel structure decomposition and information integrity protection is proposed. In this model, based on the local structure information, the image is first adaptively divided into multiple objective regions using a hierarchical superpixel algorithm to eliminate halo artifacts. Meanwhile, to avoid the error estimate caused by the local highlighted targets, a modified quadtree subdivision based on superpixel blocks is applied to obtain the global atmospheric light. Furthermore, a combined constraint is used to optimize the transmission map by minimizing the loss of information. Compared with state-of-the-art methods in terms of qualitative and quantitative analysis, experiments on real-world hazy infrared images demonstrate the efficacy of the proposed method in both contrast and visibility.

Key words: infrared image dehazing, physical model restoration, superpixels segmentation, combined constraint, enhance visibility

基于分层细分超像素和信息完整性先验的红外图像去雾

李伟华[1,2,3], 李范鸣[1,3*], 苗 壮[1,2,3], 谭 畅[1,2,3], 穆 靖[1,2,3]

(1. 中国科学院红外探测与成像技术重点实验室,上海 200083;
2. 中国科学院大学,北京 100049;
3. 中国科学院上海技术物理研究所,上海 200083)

摘要:雾霾天气由于悬浮粒子的存在,降低了红外成像系统的对比度和视觉质量。现有的去雾方法大多侧重于增强全局对比度或在图像中采取局部网格透射率估计策略,容易导致图像信息丢失、光晕伪影以及天空区域失真的现象。针对这些问题,本文提出了一种基于超像素结构分解以及信息完整性保护的单幅图像去雾方法。在该模型中,首先基于局部结构信息,设计了分层超像素算法自适应地将图像分割成多个目标区域以消除光晕伪影。同时为了避免局部高亮目标造成估计误差,采用基于超像素块的改进四叉树细分方法获得全局大气光值。在此基础上,利用组合约束通过最小化信息损失实现透射率图最优化。在真实红外雾天图像上的实验表明,在定性和定量方面与现有的经典算法相比,本文所提方法在对比度和可见性方面具有优越性。

关 键 词:红外图像去雾;物理模型重建;超像素分割;联合约束;可见性增强

中图分类号:TP391.41 **文献标识码**:A

Received date: 2021-08-18, **revised date**: 2022-05-30 **收稿日期**:2021-08-18,**修回日期**:2022-05-30

Foundation items:Supported by Shanghai Key Laboratory of Criminal Scene Evidence funded Foundation(2017xcwzk08)

Biography:LI Wei-Hua (1994-), male, Shandong, PhD. Research area involves image processing and pattern recognition. E-mail:lwh232127@163.com

*Corresponding author:E-mail: lfmjws@163.com

Introduction

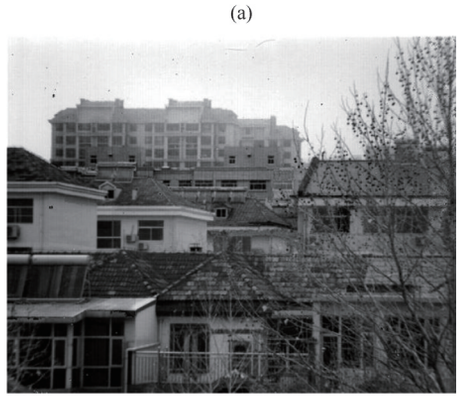
Owing to the effect of reflection and scattering of light by suspension particles, fog and haze are common atmospheric conditions that reduce the perception of the imaging system and result in low contrast, local blur, and narrow dynamic range of the imaging system [1-2]. In the field of surveillance, infrared (IR) systems with better penetration have more advantages over visible light systems, but are also affected by haze. Continuous haze increases the difficulty of observation and identification, and seriously affects the reliability and environmental adaptability of imaging systems [3]. With the development of computer technology, image dehazing algorithms have received significant attention.

Image haze removal is a challenging problem because the degree of image degradation is affected by the concentration of suspended particles and the distance from target to detector, both of which are difficult to be obtained directly from the images [4]. In recent years, researchers have proposed a variety of dehazing methods, which can be divided into two main categories: multiple information fusion, and single-frame image processing[5]. The first category requires additional information to assist with image restoration. Schechner *et al*. estimated the degree of polarization and atmospheric light from two polarized images taken at orthogonal polarization angles[6], and then calculated the physical model parameters. Narasimhan et al. obtained the depth structure by analyzing the changes in the intensities of more than two images under various weather conditions [7]. Owing to the stronger near-IR transmittance, Kudo *et al*. fused IR image with visible image in the same scene to remove haze from visible images, but it was not effective for thick fog [8]. The application of these methods is limited because of the additional hardware and complex operations. Recently, mainstream schemes based on single-image restoration have two development directions: image enhancement, and image reconstruction.

Image enhancement methods optimize visual quality by adjusting image contrast through digital image processing techniques without considering the physical factors of image degradation. Xu *et al*. applied contrast limited adaptive histogram equalization (CLAHE) to enhance the contrast of foggy images [9]. Direct manipulation of the histogram can easily be over-enhanced and thus reduce image authenticity. Multi-scale retinex (MSR) theory is an effective method in the field of image dehazing, but the incident component is estimated by Gaussian convolution, which leads to halo artifacts around the edges[10]. Filtering approaches, such as homomorphic filters[11] and high-boost filters[12], can enhance the edge information and achieve better real-time performance. Bo et al. reconstructed foggy images using wavelet decomposition and guided filtering [13]. Such methods are simple and fast, but often tend to fail when the scene changes dramatically. Moreover, the processing of details is non-ideal.

In the field of image restoration, atmospheric scattering model is the theoretical basic for reconstruction, and its optimum parameters are estimated by increasing the priori information. The dark channel prior (DCP)is a classical and effective method proposed by He *et al*. [14], but halo artifacts and block effects are introduced into the reconstructed images. Tan *et al*. [15] and Ancuti *et al*. [2] proposed improved algorithms based on the DCP to revise the transmission map. Tan utilized Markov random fields to avoid underestimating the transmission, while Ancuti fused multi-scale image patches to refine the reconstruction process in which local information is fully used. The color attenuation prior proposed by Zhu *et al*. [16], the detail prior proposed by Li *et al*. [17], and gradient channel prior proposed by Singh *et al*. [18] were also applied to estimate the thickness of the haze. Compared with DCP, these methods have better performance, but fail to completely eliminate halo artifacts and deal with highlighted regions.

(a)



(b)

Fig. 1 Sample result of our proposed method, (a) original real hazy infrared image, (b) reconstructed image obtained by our technique
图1 所提方法结果对比,(a)原始真实雾天红外图像,(b)去雾后重建图像

Currently, the vast majority of image dehazing algorithms are targeted at multispectral color images and underwater images, and IR images only provide additional auxiliary information for dehazing. Haze weather also affects the visual quality and subsequent processing of IR imaging systems, as shown in Fig. 1(a), but there is little reference to deal with the degradation of IR images. The process of single-wave IR image dehazing is more difficult due to noise and less information. In this paper, we propose a novel IR image dehazing method based on

physical model that utilizes superpixel segmentation and upper-lower boundary constraints to avoid halo artifacts and distortions in the sky region. The resulting technique achieves excellent reconstruction and better visual recovery, as shown in Fig. 1(b). The main contributions of our article can be summarized as following aspects:

● Local grid windows are replaced by superpixels to eliminate halo artifacts. To produce content-sensitive superpixels, a hierarchical subdivision superpixel-splitting algorithm is proposed, that texture information is added to the segmentation process to guide the re-segmentation of complex regions. Hierarchical subdivision based on local texture information and regional mergers can guarantee consistency of local information.

● A modified superpixel-based quadtree subdivision is proposed to obtain the airlight value. This method can ensure the accuracy of airlight value, while solving the drawback of artificially preset thresholds. Compared to other methods, our method is more robust for local highlighted targets.

● A combined upper-lower boundary constraint based on information integrity prior is proposed for IR images to calculate the transmission map, which improves visual quality and solves the distortion in the sky region. The reconstructed image can be inverted by a reasonable estimation of the scattering model parameters.

The remainder of this paper is organized as follows. In Sect. 1, we briefly introduce the physical model of the optic energy attenuation process in a foggy atmosphere. Sect. 2 describes the proposed method in detail. Sect. 3 presents the experimental results of our algorithm and several existing methods. Finally, a summary is presented in Sect. 4.

1 Background

1.1 Atmospheric scattering model of image formation

Fog and haze are common atmospheric conditions and typical aerosol particles, especially in winter. The target energy received by the photoelectric system along the line of sight consists of two parts: target radiation that is attenuated by atmospheric absorption and scattering, and path radiation that is superimposed on the target and background. Path radiation increases the background noise detected by the system and reduces the contrast between the target and background, which caused distant targets to become grayish white [19]. McCartney's atmospheric scattering model [20] has been proven to be a reasonable approximation for light propagation in the atmosphere, defined as follows:

$$I(x) = J(x) \cdot t(x) + A \cdot (1 - t(x)) \quad , \quad (1)$$

where $I(x)$ and $J(x)$ are the degraded image and the scene radiance, respectively; A is global atmospheric light at infinity, which is independent of the position of the target; and $t(x)$ is the atmospheric transmission along the path from the target to the detector. Referring to Eq. 1, the first term, $J(x) \cdot t(x)$, is known as the direct transmission, which represents the attenuation in the medium; the second term, $A \cdot (1 - t(x))$, represents the impact of airlight, which is the main cause of low contrast and poor visibility.

The aerosol particle radius in hazy weather is concentrated in 0.5~10 μm, and the scattering efficiency of the particles K is related to the relative size between the scattered particle radius r and the wavelength of the incident light λ, their relationship is shown in Fig. 2[21]. When the radius r is approximately equal to the wavelength λ, the maximum value of K is close to 3.8, and the scattering ability is strongest[22]. From the above analysis, it can be concluded that hazy weather also has a significant effect on IR images. We assume that the suspended particulates in the atmosphere are spherical and uniformly distributed, and according to the Lambert-Beer law and Mie scattering theory, $t(x)$ is approximated as follows:

$$t(x) = e^{-\beta d(x)} \quad , \quad (2)$$

$$\beta = \alpha + \gamma \quad , \quad (3)$$

where β is the extinction coefficient, consisting of the absorption coefficient α and the scattering coefficient γ, and is related to the particle radius and aerosol particle concentration; and d is the length of the optical path between the target and the receiver. When the observed distance and atmospheric status remain stationary, the value of t is constant. In other words, t is negatively related to d.

Referring to Eq. 1, to obtain J, it is necessary to utilize additional information and constraints to infer t and A. In short, the haze-free IR image J can be derived from Eq. 4 with the reasonable parameters:

$$J = \frac{I(x) - A}{t(x)} + A \quad . \quad (4)$$

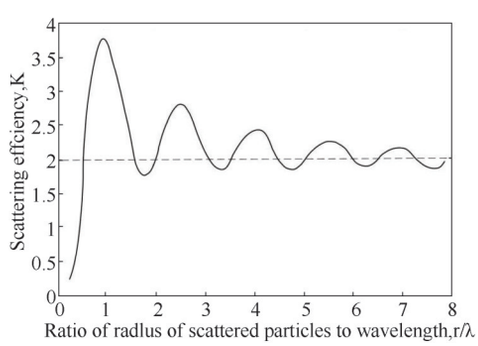


Fig. 2 The relationship between scattering efficiency and the ratio of the wavelength to the radius of scattered particles. As the particle radius increases, the scattering efficiency eventually converges about 2 after a slight oscillation

图2 散射效率与波长和粒子半径之比的关系。随着粒子半径的增大,散射效率最终收敛到2左右

1.2 Simple linear iterative clustering

Achanta *et al.* proposed the concept of simple linear iterative clustering (SLIC), which optimizes k-means clustering [23]. It has obvious advantages over other superpixel segmentation algorithms, in that SLIC limits the search range to a space proportional to the patch size, which dramatically reduces the number of distance calculations. The complexity of SLIC is linearly related to the

total number of pixels N and is not dependent on the number of blocks k. Therefore, SLIC is characterized by fast processing speed and low computational complexity; more importantly, it can better identify boundaries.

SLIC utilizes a multidimensional feature vector to calculate the similarity between pixel pairs. The desired number of superpixels k and the compactness m are the only two parameters that need to be specified. For color images in the CIELAB color space, each image is transformed into a five-dimensional feature space $V = [l, a, b, x, y]^{\mathrm{T}}$, which is expanded from the color space and coordinate space. The initial cluster centers C_i are grid sampled at a constant interval $S = \sqrt{N/k}$ on the image. To avoid the overlap of C_i and edge pixels on the image, center point C_i is moved to the pixel with the lowest gradient in the 3×3 pixel neighborhood. D is defined as the distance in multidimensional space, which describes the similarity of pixels and cluster centers.

$$D = \sqrt{(d_{lab})^2 + (\frac{d_s}{S})^2 \cdot m^2} \quad , \quad (5)$$

$$d_{lab} = \sqrt{(l_p - l_q)^2 + (a_p - a_q)^2 + (b_p - b_q)^2} \quad , \quad (6)$$

$$d_s = \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2} \quad , \quad (7)$$

where d_{lab} and d_s represent the luminance proximity and spatial proximity, respectively; p and q are two pixels within the bounded range; $[l, a, b]$ and (x, y) indicate the color of pixels in the CIELAB color space and coordinate, respectively. Compactness m is relevant to the boundary preservation ability. When m is large, it tends to produce regular superpixels. More details can be found in Ref. [23]. After several iterations of the centers and blocks, all pixels were divided into corresponding blocks to minimize D.

2 Image dehazing method

In this section, we provide a detailed description of our dehazing method. First, the foggy image is divided into several parts using a hierarchical subdivision superpixel segmentation algorithm. Then, a modified quadtree method is proposed to automatically obtain accurate global atmospheric light A. Next, a reasonable transmission map t is inferred based on information integrity prior to improve visual quality and avoid information loss. Finally, the haze-free image J can be deduced from Eq. 4. Fig. 3 shows a flowchart of the proposed method.

2.1 Superpixels structure decomposition

In Ref. [14] and Ref. [24], the foggy image was divided into small blocks using fixed-size rectangular patches in a localized area to analyze atmospheric transmission. However, halo artifacts and block effects are inevitable, regardless of the size of the patches chosen. Although some techniques (e. g. guided filters and soft matting) can mitigate this phenomenon, they require a large amount of computation and memory consumption. Referring to Eq. 2, the value of transmission t is correlated with the distance d; it is difficult to ensure that an identical depth is suitable for each pixel in a fixed segmentation block. For example, when the rectangular block is located at the junction of two targets with different depths, a fixed transmission t applied to the reconstructive process of this block may lead to some pixels over-enhancement (t is less than the optimal value) or under-enhancement (t is more than the optimal value).

A superpixel block is a set of adjacent pixels that have analogous colors, similar brightness, or related structures. Compared with rectangular patch segmentation, the superpixel segmentation guarantees the same depth of pixels in the local block. Therefore, superpixels reduce the probability of over-segmentation and under-segmentation, which is the essential reason for halo artifacts. Compared with color images, IR images only carry luminance information, and the classical SLIC algorithm is not applicable. We transform Eq. 6 as a luminosity function for IR images, and the spatial distance D' can be expressed as:

$$D' = \sqrt{{d_l}^2 + (\frac{d_s}{S})^2 \cdot m^2} \quad , \quad (8)$$

$$d_l = \sqrt{(l_p - l_q)^2} \quad , \quad (9)$$

where d_l represents the luminance proximity, and l is the luminance value. Due to the non-uniformity of IR images, stripe noise adversely affects results of segmentation. Bilateral filter is widely used in infrared image processing [25] and is effective in reducing image noise while maintaining sharp edges [26]. Therefore, bilateral filter is applied as a pre-processing step before segmentation.

An image usually contains both smooth and complex regions, and the size of superpixels is determined by the

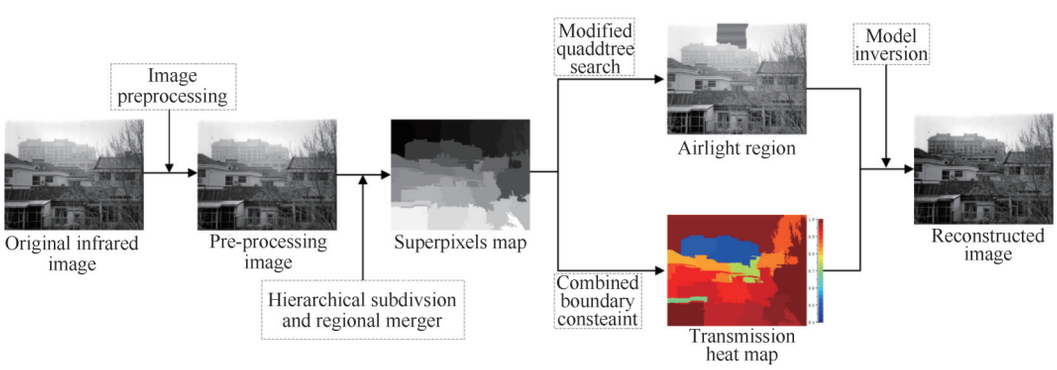


Fig. 3 The flowchart of the proposed algorithm
图3 算法流程图

expected number k. However, the value of k is only specified in the initialization phase, which causes the segmentation process not distinguishing the image content. When the value of k is small, the size of the superpixel is larger, and it tends to be under-segmented in texture-rich regions. Under-segmentation results in the superpixel easily containing targets of different depths. When the value of k is large, it is easy to produce over-segmentation which reduces the correlation of uniform depth targets. To solve this problem, we develop a hierarchical subdivision SLIC based on image texture and regional merger. The principle of our modified SLIC method is based on the analysis of the image: a small value of k should be used for simple texture regions, such as the sky; conversely, a large value of k should be used for texture-rich regions to ensure adequate segmentation. To achieve this, we propose a hierarchical segmentation method. We use k_1 to obtain a series of rough superpixel blocks with large size; the sky region is often guaranteed to be segmented accurately with k_1. At this point, the superpixels from sky region tend to be simple in texture and do not need to be segmented again; we need to segment the richly-textured superpixel blocks again to ensure consistent depth within the superpixels. The entropy value of image is widely used to characterize image texture information [27-28], which can be defined as:

$$en = -\sum_{i=0}^{H} P(i)\log[P(i)] \quad , \quad (10)$$

where H is the number of gray-scale interval; and $P(i)$ is the probability of each gray interval obtained by the histogram statistic. When en of superpixel block p_i is greater than threshold T, p_i needs to be segmented. The number of sub-superpixels k_2 is set to further segment p_i. After the two-level segmentation, we obtained a series of finer sub-superpixels. In the process of fine segmentation, local over-segmentation could occur because it does not focus on image texture similarity, which is more likely to result in block effects. To reduce the block effect, a regional merger is applied to combine the over-segmented fine sub-superpixels. The region adjacency graph (RAG) centered on each superpixel block is constructed as shown in Fig. 4. The RAG, as an undirected graph, provides a spatial view for the segmented image, where each superpixel p_i is regarded as a nodes v_i. We let $G = (V, E)$ be the RAG for superpixel p_i, where $V = \{v_1, v_2, \cdots, v_n\}$ is a set of nodes and $E = \{e_1, e_2, \cdots, e_n\}$ is the corresponding edge between adjacent nodes. Each edge has a weight w_i that represents the similarity between two adjacent nodes. The luminance average $\bar{I}$ and texture entropy en are utilized to achieve a reasonable representation of the regional similarity as follows:

$$w_i = aw_l + bw_{en} \quad , \quad (11)$$

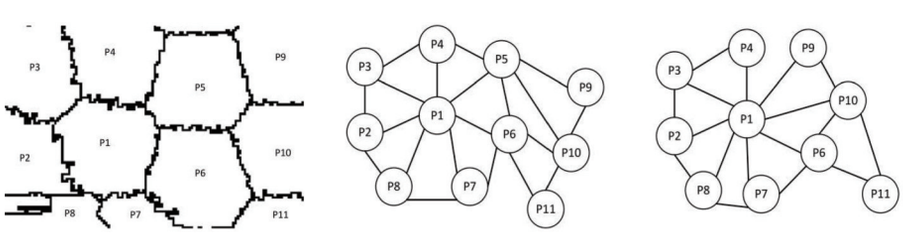
$$w_l = \frac{\min(\bar{I}(p_i), \bar{I}(p_j))}{\max(\bar{I}(p_i), \bar{I}(p_j))} \quad , \quad (12)$$

$$w_{en} = \frac{\min(en(p_i), en(p_j))}{\max(en(p_i), en(p_j))} \quad , \quad (13)$$

where p_i and p_j are adjacent nodes; a and b are two weight factors ($a + b = 1$). When the content of the IR image is simple, the luminance average $\bar{I}$ has a higher influence than the texture entropy en, so that we can set a greater than b. When w_i is greater than the thresholdw_T, p_i and p_j are highly similar, thus performing the merge operation. In the following experiments, we set w_T=0.8, a=0.5, and b=0.5. Each superpixel needs to be searched, and the merger process of each layer is iterated until convergence; then a final superpixel map is obtained, which has the desired characteristics.

2.2 Quadtree method based on superpixel blocks

The global atmospheric light A describes the ambient illumination in the scene, which should be estimated from the most haze-thick region in the image, as proposed by Narasimhan [29]. Generally, the infinite region of the sky to obtain A is precise. In most existing methods, the top 0.1% of the brightest pixels from the dark channel obtained by DCP are selected to extract the value of A [30-31]. Overly bright objects, such as local artificial light sources, are easily to be mixed in candidate pixels leading to erroneous estimate. Besides, DCP is not suitable for IR images because it is derived from the statistics of multispectral color images. The local highlighted region usually occupies a small portion of the image, and the influence of highlighted region can be reduced by the surrounding pixels when calculating the average value. Therefore, in order to accurately locate the sky regions, a hierarchical searching method based on quadtree subdivision is normally adopted, as shown in Fig. 5(b) [32-34]. The process divides the image into four equal parts and selects the sub-block with the largest average value as the new block, and then repeats until the size of the eventual block is less than a preset threshold. The size of the threshold affects the selection of the final region and is



(a) (b) (c)

Fig. 4 Example of RAG in local region (a) A local region in fine segmented layer, (b) the corresponding RAG,(c) the next RAG after merge of P1 and P5 if the weight w is greater than w_T

图4 局部窗口RAG图示例 (a)精细分割图中局部窗口,(b)局部窗口相应的RAG图,(c)当w大于阈值w_T时,合并P1、P5后下一层RAG图

less robust to images with a small sky region. When the threshold is large, it is difficult to ensure that the final block contains only the sky region. In this work, inspired by quadtree subdivision, a modified quadtree method based on superpixels is proposed to obtain airlight$A$, as shown in Fig. 5(a).

Our proposed method regards the rough superpixel p_i, obtained by the first segmentation, as the basic element, and the routine quartered region is replaced by a combination of superpixels. The detailed description is as follows. First, the initial segmentation is executed to generate rough superpixels, and each superpixel is labelled as L_{p_i}.

$$L_{p_i} = i, \quad i = 1,2,\cdots,k \qquad (14)$$

where k is the total number of superpixel p_i. Each pixel x in the image is tagged with the label L_{p_i} of the corresponding superpixel.

$$L_{(x)} = L_{p_i}, \quad x \in p_i \qquad (15)$$

where $L_{(x)}$ represents the label of pixel x. Next, the image is divided into four parts of equal size, and the labels that have appeared in the largest average part are counted as φ. A new sub-block Ω_n is constructed with superpixels corresponding to the labels in φ.

$$\Omega_n = \left\{ \sum p_i \middle| L_{p_i} \in \varphi \right\} \qquad (16)$$

where Ω_n is the new sub-block formed by the nth iteration. Finally, the sub-block Ω_n will perform a new round of quadtree subdivision. This process is repeated until each pixel within the quarter part with the maximum average belongs to a same label. The average luminance value of the final block Ω_n' is an accurate estimation of the airlight A.

$$A = \frac{\sum l(x)}{M}, x \in \Omega_n' \qquad (17)$$

where $l(x)$ represents the luminance value of the pixel; and M is the total number of pixels in region Ω_n'. When the sub-block to be segmented contains fewer than four superpixels, the selected quarter part with the maximum average may contain all the labels that appear in the sub-block. In this situation, the iterative process falls into partial circulation. To obtain the airlight automatically, the occurrence of the first circulation is set as the termination criterion for the iteration. The average values of each superpixel in the sub-block were calculated, and the maximum value was selected as the airlight A. This modified method without human intervention can reduce the impact of local noise and blind pixels on the estimation of the airlight. For images without sky regions, the method is also robust in finding the most blurred region in the image.

2.3 Estimation of the transmission map based on information integrity prior

After obtaining the atmospheric light A, the quality of the haze-free image J depends on the value of the transmission t, referring to Eq. 4. Assuming that t^* is the optimal transmission for a superpixel, the dehazed image will remain a little haze with values of $t > t^*$. Conversely, the effects of low brightness, poor vision, and loss of detail will be brought into the reconstructed image with values of $t < t^*$. The relationship between transmission t and image J is shown in Fig. 6. The original image in Fig. 6 is captured from a short-wave infrared camera with resolution of 640*512. The frame rare, integration time and focal length are 50 Hz, 2 ms and 50 mm, respectively.

For most existing method based on image restoration, the reconstruction of the sky regions and highlighted target regions are more or less distorted because a small value of transmission is estimated in these regions, which overly magnifies the differences between pixels.

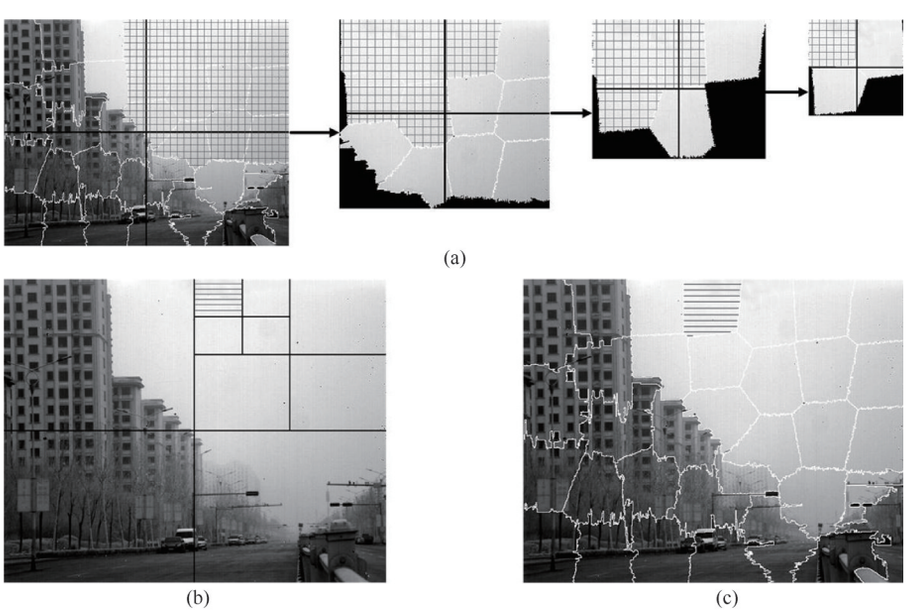


Fig. 5 Searching method to obtain A. Marked region is the next sub-block (a) process of modified quadtree method based on superpixels, (b) result and process of traditional quadtree method, (c) result of our method

图5 大气光值A搜索方法 (a)基于超像素的改进四叉树搜索方法流程图,(b)传统四叉树搜索方法流程与结果,(c)所提方法流程与结果

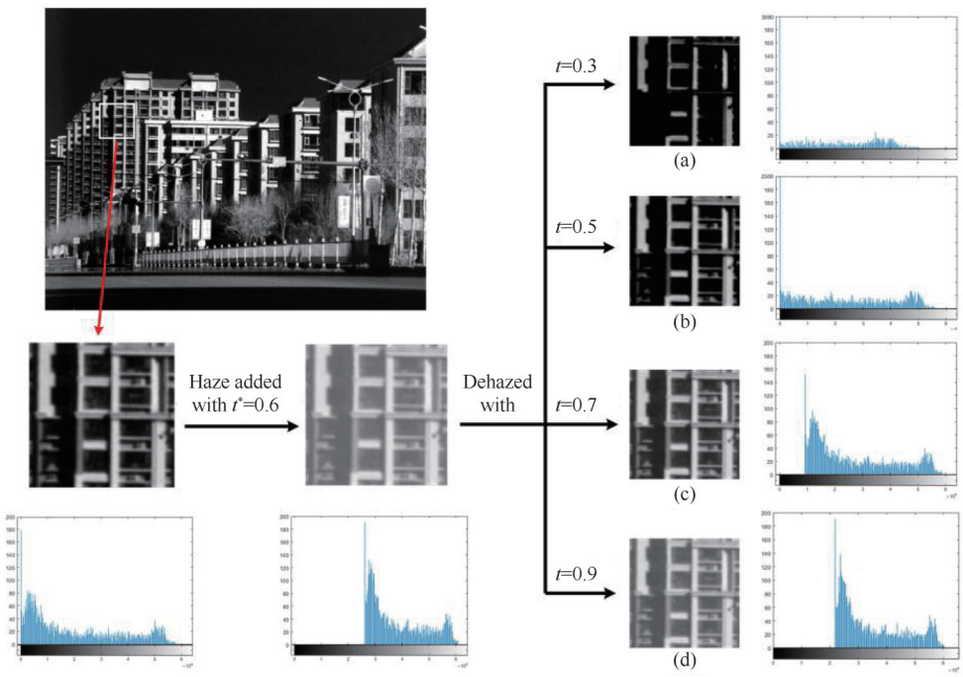


Fig. 6 Relationship between transmission and reconstructed image. Haze is added to a block with t = 0.6. Dehazing the artificial block with t is (a) 0.3, (b) 0.5, (c) 0.7, (d) 0.9, and the histogram distribution for each block.
图6 透射率与重建图像关系仿真图。使用透射率t=0.6添加雾霾衰减，选定透射率t为(a)0.3、(b)0.5、(c)0.7、(d)0.9复原雾霾后图像及相应直方图分布

For a target with obvious features, taking too large or small value of transmission can lead to poor results such as Fig. 6(a) and Fig. 6(d). Although it is an effective way to obtain accurate transmission map through statistical priors (e. g. DCP and color attenuation prior), single-wave IR images carry less information, and multi-band statistics are not applicable. In terms of the quality of dehazing images, the reconstructed IR images should have better visual quality, and the highlighted targets cannot be erased. Therefore, the goal of image reconstruction is to improve visual quality while maintaining information integrity for subsequent recognition tasks.

The mean squared error (MSE) contrast, C_{MSE}, which represents the variance of the pixel values and has been widely applied to evaluate the contrast characteristic of images[23], is given by

$$C_{\mathrm{MSE}} = \sum_{x=1}^{N} \frac{(I(x) - \bar{I})^2}{N} \quad , \quad (18)$$

where $\bar{I}$ is the average luminance value of image; and N represents the total number of pixels in the image. Human vision has a stronger perception for high-contrast images, while foggy images reduce the visual perception due to the narrow dynamic range. In general, the value of C_{MSE} for clear natural images is larger because of large luminance value dispersion; relatively, the value is smaller for foggy images. Under extreme and ideal conditions, the value of C_{MSE} decreases continuously as the fog concentrates, and eventually converges to zero. To improve the visibility of the reconstructed image, the selected value of transmission t should optimize the image contrast, which means increasing the value of C_{MSE}. From Eq. 4 and Eq. 18, C_{MSE} for each reconstructed superpixel is expressed as follows:

$$C_{\mathrm{MSE\_re}} = \sum_{x=1}^{N} \frac{(I(x) - \bar{I})^2}{t^2 N} \quad , \quad (19)$$

where t is the transmission for each superpixel. Note from Eq. 19 that the value of $C_{\mathrm{MSE\_re}}$ is a monotonically decreasing function of the transmission; therefore, a small value of t can be selected to increase the value of $C_{\mathrm{MSE\_re}}$. Mathematically, the value of $C_{\mathrm{MSE\_re}}$ continues to increase as t decreases. But for 16-bit IR images, pixel values cannot exceed the valid range of data bits. Some pixel values are truncated to 0 or 65 535 because of underflow or overflow. The truncation of these pixels causes the degradation of visual quality, while the truncated pixel information is erased. Fig. 7 shows the mapping of the input values to output values.

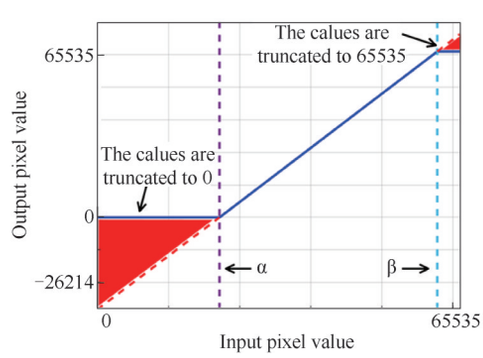


Fig. 7 The input values of [0,65535] are mapped to the output with a small t. The information in the red region is lost because the values are truncated
图7 利用较小的t将输入范围[0,65535]映射到输出范围. 红色区域信息因发生数据截断而丢失

As shown in Fig. 7, the input values in $[0,\alpha]$, and $[\beta, 65\,535]$ are truncated to 0 or 65 535. The atmospheric transmission t determines that only the input values in range $[\alpha,\beta]$ can occupy the full dynamic range of $[0, 65\,535]$ after reconstruction, which means that the truncated pixel values lead to information loss in the red region in Fig. 7 and decreases the details in the haze-free image as shown in Fig. 6(a). The valid input range is inversely proportional to slope $1/t$. For simplicity, the squared sum of the truncated values is used to represent the loss cost of information, and is expressed as:

$$C_{\mathrm{LOSS}} = \sum_{p=1}^{N}\{\min[0,J(x)]^2 + \max[0,J(x) - 65\,535]^2\}, \quad (20)$$

where $J(x)$ denotes the luminance value of each pixel in the reconstructed image. To maintain the integrity of the image information, the selected value of t should not result in any loss of information. Therefore, Eq. 20 must satisfy the following constraint conditions:

$$\begin{cases} \min[0,J(x)] = 0 \\ \max[0,J(x) - 65\,535] = 0 \end{cases}. \quad (21)$$

Referring to Eq. 4 and Eq. 21, a group of constraint conditions is derived.

$$\begin{cases} t \geqslant 1 - \dfrac{\min[I(x)]}{A} \\ t \geqslant \dfrac{\max[I(x)] - A}{65\,535 - A} \end{cases}. \quad (22)$$

Accordingly, the constraint of transmission can be inferred:

$$t \geqslant \max\left\{1 - \frac{\min[I(x)]}{A}, \frac{\max[I(x)] - A}{65\,535 - A}\right\}. \quad (23)$$

Based on the analysis that the value of C_{MSE} decreases as t increases, the minimum value that satisfies Eq. 23 is selected as the optimal transmission. It signifies that a high-contrast superpixel is obtained without loss of information, we call it information integrity prior. The transmission of each superpixel can be calculated by

$$t = \max\left\{1 - \omega \cdot \frac{\min[I(x)]}{A}, 1 + \omega \cdot \frac{\max[I(x)] - 65\,535}{65\,535 - A}\right\}, \quad (24)$$

where ω is a constant used to control the degree of dehazing. As ω decreases, more fog is retained, and we set it to 0.95 in the following experiments. To further reduce the effect of non-uniformity and blind pixels, we take the average value of the top 1.5% of the largest and smallest pixels as the maximum value and minimum value. The first term in Eq. 24 uses a similar constraint as dark channel prior in He's DCP algorithm [14]. The DCP is considered to be an effective prior, but fails for some objects whose values are brighter than the atmospheric light. In contrast, Eq. 24 employs an additional constraint to prevent overflow of the highlighted region.

3 Experimental results

In this section, to assess the performance of our proposed method, we test it in real self-built infrared dataset and compare it with CLAHE [9], MSR [10], the method of Bo *et al.* [13], and the method of Zheng *et al.* [27]. Common natural mutation and slow-varying scenes are selected to exhibit in Fig. 8 and Fig. 9. All test images were captured from the same IR camera mentioned in Sect. 2.3, whose spectral response range is 0.9~1.7 μm. We implemented the proposed technique in MATLAB 2018a and tested it on an i5 CPU with 16 GB of RAM.

3.1 Visual assessment

Fig. 8 and Fig. 9 respectively show the visual comparison of CLAHE, MSR, Bo *et al.*, Zheng *et al.* and our proposed method under mutation scene and slowly-varying scene, which are the most common scenes in the city. In Fig. 8(a), the boundary between buildings of different depths is more obvious in the original IR image and the visual quality of building at different distance has significant differences. As a comparison, Fig. 9(a) shows a scene where the distance changes slowly and the boundary of the degradation degree is not clear.

The CLAHE method can effectively enhance the global contrast of IR images by equalizing the image statistical histogram. The foggy region, whose pixel values are relatively concentrated in the histogram, has better visual performance after reconstruction; however, over-enhancement has occurred in the regions with slight degeneration such as trees and roofs, as shown in Fig. 8

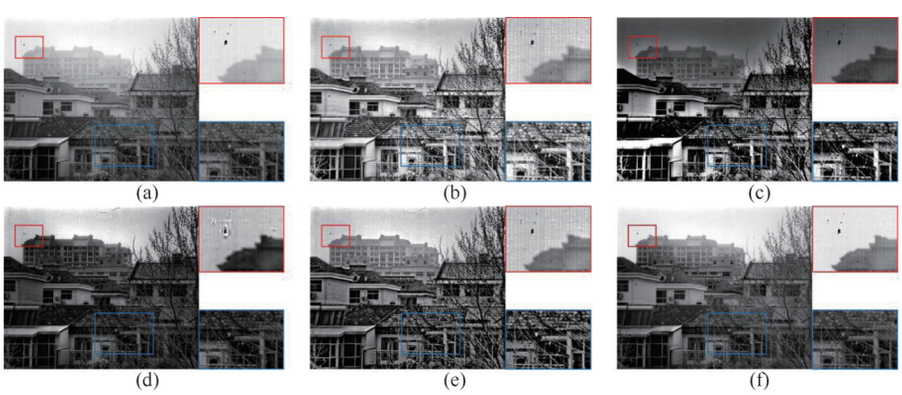


Fig. 8 Visual comparisons in mutation scene (a) original image, (b) CLAHE, (c) MSR, (d) Bo *et al.*, (e) Zheng *et al.*(f) proposed technique. The zoomed-in details are shown on the right side of the picture
图8 突变场景中所提方法对比 (a)原图,(b)CLAHE,(c)MSR,(d)Bo,(e)Zheng,(f)本文方法

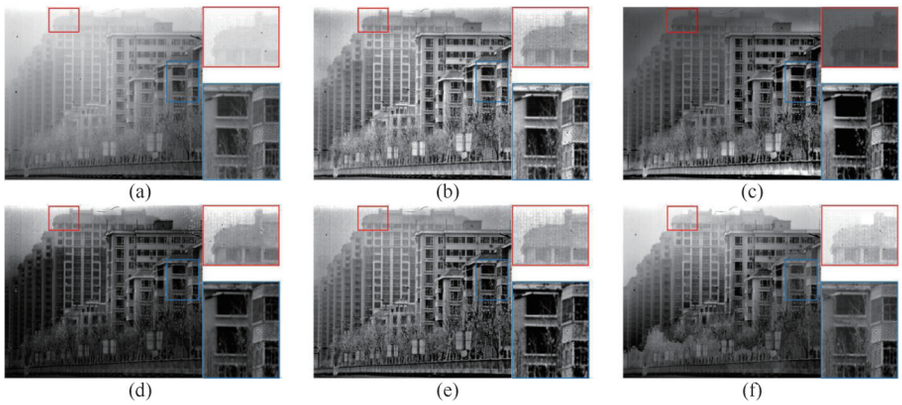


Fig. 9 Visual comparisons in slowly-varying scene (a) original image, (b) CLAHE, (c) MSR, (d) Bo *et al*., (e) Zheng *et al*.,(f) proposed technique. The zoomed-in details are shown on the right side of the picture

图9 缓变场景中所提方法对比 (a)原图,(b)CLAHE,(c)MSR,(d)Bo等,(e)Zheng等,(f)本文方法。放大后细节图于右侧展示

(b) and Fig. 9(b). Because the local information of the image is not considered, the over-enhancement phenomenon cannot be effectively suppressed. Compared with Fig. 8(b) and Fig. 9(b), the over-enhancement phenomenon is diminished in Fig. 8(c) and Fig. 9(c). Although multi-scale Gaussian surround space constants ∂_i (set to 10, 50, and 250 in this study) are used to solve the incident component, Gaussian convolution will result in halo artifacts at the edges with large differences in brightness. Besides, reconstructed image detail loss occurs in low-luminance areas with slow gradient changes, such as windows. In Fig. 8(d) and Fig. 9(d), the method of Bo et al. achieves better results than above methods, particularly in foggy regions, but the overall luminance value of the image is reduced. In Bo's method, a detail layer, which is amplified by the scene depth proportion, is added to the wavelet filtered output images, and the enhancement coefficient is set to a higher value in the far regions. Therefore, the effect of blind pixels is the most severe in this method, especially in the sky area. Over-enhancement is also present in the reconstructed images obtained by the method of Zheng *et al*., as shown in Fig. 8(e) and Fig. 9(e). The principle of their algorithm is multi-exposure image fusion (MEF) based on an adaptive patch structure. Owing to the selection of adaptive image block size and MEF for all corresponding image blocks in the underexposed image sequences, a large computational volume is inevitable, leading to a long running time. Overall, the methods of Bo et al. and Zheng *et al*. produced better results than the CLAHE and MSR methods. All of the above methods result in poor reconstruction of the sky region, as shown in the red boxes in Figs. 8 and 9, because the sky region has non-uniform noise and is over-amplified. By applying the pre-processing step for noise removal mentioned in Sect. 2.1, we test it with the CLAHE algorithm, as shown in Fig. 10. The stripe noises are effectively suppressed, but the image details are blurred due to the smoothing effect. Therefore, it is difficult to achieve satisfactory results by simply suppressing image noises. Our proposed method can reduce the halo artifacts being brought to the reconstructed image because it ensures that the transmission of pixels in the same superpixel is uniform. Observing the results in Fig. 8(f) and Fig. 9(f), the results of our method is clearer and more natural, while the sky region is not distorted because of the additional constraint.

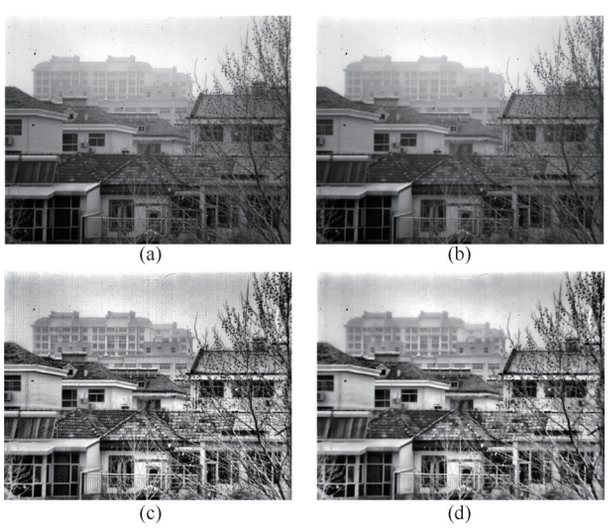


Fig. 10 Example of comparison on pre-processing before CLAHE. (a) original infrared image,(b) Pre-processing image by bilateral filtering, (c) result of CLAHE in (a), (d) result of CLAHE in (b)

图10 CLAHE算法采用图像预处理后实验图 (a)原始红外图像,(b)滤波后效果图,(c)图a处理效果图,(d)图b处理效果图

3.2 Quantitative analysis

Although subjective visual assessment is an effective method to evaluate the reconstructed quality, mainstream full reference and no-reference metrics are also calculated to further illustrate the performance. In the objective evaluation, the common quality assessment parameters were calculated: peak signal-to-noise ratio (PSNR), structural similarity index metric (SSIM)[35], blind assessment indicators e and $\bar{r}$[36], image visibility measurement (IVM)[37], and MSE. In addition, we defined a near-scene structural similarity index metric (NSSSIM). In the IR imaging system, when the imaging dis-

tance is small, the haze effect is negligible. Compared with visible images, clear near-scene regions with high contrast are more likely to exist in the IR images. For these near-scene regions, the transmission is a large value close to 1, and these regions are easily problem areas due to over-enhancement. NS-SSIM can more accurately assess the degree of over-enhancement. The more severe the over-enhancement and deviation from authenticity, the smaller the value. The blue rectangular boxes in Figs. 8 and 9 are the manually selected near-scene regions, which do not require excessive processing. A quantitative comparison is presented in Table I.

Table1 Scores for PSNR, SSIM, e, $\bar{r}$, IVM, MSE, NS-SSIM for all compared algorithms in Fig. 8 and Fig. 9
表 1 图 8 和图 9 中 PSNR,SSIM,e,$\bar{r}$,IVM,MSE 和 NS-SSIM 客观评价指标

		PSNR	SSIM	e	$\bar{r}$	IVM	MSE ($\times 10^8$)	NS-SSIM
Fig. 8	CLAHE	13.54	0.46	9.32	**2.52**	9.44	3.19	0.59
	MSR	13.12	0.42	2.71	1.83	9.19	2.67	0.52
	Bo's	14.27	0.58	3.47	1.24	9.64	3.60	0.82
	Zheng's	15.05	0.53	**11.38**	2.26	**11.06**	3.66	0.63
	Proposed	**15.23**	**0.66**	4.06	1.76	9.53	**3.89**	**0.96**
Fig. 9	CLAHE	12.72	0.48	**20.66**	**2.84**	10.01	2.00	0.74
	MSE	12.13	0.40	13.33	1.84	9.37	1.16	0.56
	Bo's	13.36	0.50	14.20	1.40	10.07	1.97	0.73
	Zheng's	13.82	**0.61**	19.46	2.66	**10.78**	2.32	0.62
	Proposed	**14.08**	0.57	15.25	2.06	10.32	**3.07**	**0.91**

The bolded score is the best score.

PSNR and SSIM are fully reference-based metrics. If the PSNR is larger, the distortion of the reconstructed image will be smaller. The range of the SSIM score is 0 to 1, and a larger value closer to 1 means that the reconstructed image has better perception and realism. Except for the slightly lower SSIM in Fig. 9, the proposed method achieves the highest score. Therefore, our method has better performance, and the method of Zheng et al. is second. Both of e and $\bar{r}$ assess the enhanced degree of image visibility using the enhanced degree of image edges. The edges of the foggy images should be enhanced for ideal reconstructed images. IVM also uses the enhancement of the image edges to illustrate the performance. Among these three indices, our method is better than MSR and that of Bo et al., but has lower scores than CLAHE and that of Zheng et al. This is because the present of over-enhancement in CLAHE and Zheng et al., which makes the image edges more visible. Our method obtains the highest score for the MSE and NS-SSIM metrics; those of Bo et al. and Zheng et al. also have better performance. This shows that our images have better contrast and authenticity, and the results are consistent with those of the subjective evaluation.

In summary, our method performs better in quantitative comparisons, which is consistent with the qualitative results. The results of both qualitative and quantitative analyses verify that our proposed method is effective in terms of contrast, visibility, and especially in the avoidance of over-enhancement.

4 Conclusion

This paper presented an effective framework for haze removal from IR images. Our strategy optimizes the contrast of hazy images by ensuring regional similarity and information integrity. Hierarchical subdivision superpixel segmentation ensures regional similarity to reduce the impact of halo effects, and a reasonable transmission map can be estimated by maintaining the information integrity. The atmospheric light can be obtained automatically because of our hierarchical search method based on superpixels. Compared with advanced methods, our approach is more natural without over-enhancement. Furthermore, the distortion problem of the sky region can be solved. In the future, we plan to combine our method with deep learning to explore the feasibility of extrapolating the target depth.

References

[1] Hu B R, Wu W J, Dai M Y, *et al*. Study on property of infrared obscure of artificial fog[J]. *J. Infrared Millim. Waves*, 2006, **25**(2): 131-134.
[2] Ancuti C, Ancuti C O, Vleeschouwer C D, *et al*. Day and night-time dehazing by local airlight estimation[J]. *IEEE Trans. Image Process*, 2020, **29**: 6264-6275.
[3] Xu Y, Wen J, Fei L K, *et al*. Review of video and image defogging algorithms and related studies on image restoration and enhancement [J]. *IEEE Access*, 2016, **4**: 165-188.
[4] Wu Q, Zhang J, Ren W, *et al*. Accurate transmission estimation for removing haze and noise from a single image[J]. *IEEE Trans. Image Process*, 2020, **29**: 2583-2597.
[5] Wang P, Fan Q, Zhang Y, *et al*. A novel dehazing method for color fidelity and contrast enhancement on mobile devices[J]. *IEEE Trans. Consum. Electron*, 2019, **65**(1): 47-56.
[6] Schechner Y Y, Narasimhan S G, Nayar S K. Polarization-based vision through haze[J], *Appl. Opt.*, 2003, **42**(3): 511-525.
[7] Narasimhan S G, Nayar S K. Vision and the atmosphere[J]. *Int. J. Comput. Vis*,2002, **48**(3): 233-254.
[8] Kudo Y, Kubota A. Image dehazing method by fusing weighted near-infrared image[C]. Proceedings of the International Workshop on Advanced Image Technology (IWAIT), IEEE, 2018: 1-2.
[9] Xu Z, Liu X, Chen X. Fog removal from video sequences using contrast limited adaptive histogram equalization[C]. Proceedings of the International Conference on Computational Intelligence and Software Engineering(CCISE), IEEE, 2009: 1-2.
[10] Li J. Application of image enhancement method for digital images based on Retinex theory[J]. *Optik*, 2013, **124**(23): 5986-5988.
[11] Seow M J, Asari V K. Ratio rule and homomorphic filter for enhancement of digital color image [J]. *Neurocomputing*, 2006, **69**: 954-958.
[12] Gonzalez R C, Woods R E. Digital image processing[M]. Prentice-Hall, State of New Jersey, 2007.
[13] Jiang B, Zhang W, Zhao J, *et al*. Gray-scale image dehazing guided by scene depth information[J]. *Math. Probl. Eng.*, 2016, **10**: 1-10.
[14] He K, Sun J, Tang X. Single image haze removal using dark channel prior[J]. *IEEE Trans. Pattern Anal. Mach. Intell*, 2011, **33**: 2341-2353.
[15] Tan Y, Wang G. Image haze removal based on superpixels and markov random field[J]. *IEEE Access*, 2020, **8**: 60728-60736.
[16] Zhu Q, Mai J, Shao L. A fast single image haze removal algorithm using color attenuation prior[J]. *IEEE Trans. Image Process*, 2015, **24**: 3522-3533.
[17] Li J, Zhang H, Yuan D, *et al*. Single image dehazing using the change of detail prior[J]. *Neurocomputing*, 2015, **156**: 1-11.
[18] Singh D, Kumar V, Kaur M. Single image dehazing using gradient

channel prior[J]. *Appl. Intell*, 2019, **49**: 4276-4293.
[19] Zhang Y, Chen Y, Fu X, *et al*. The research on the effect of atmospheric transmittance for the measuring accuracy of infrared thermal imager[J]. *Infrared Phys. Technol*, 2016, **77**: 375-381.
[20] McCartney E J. Optics of the atmosphere: scattering by molecules and particles[J]. *Phy. Bull*, 1997, **521**: 28.
[21] Zhou S. Introduction to advanced infrared photoelectric engineering [M]. Science Press, (周世椿. 高级红外光电工程导论. 科学出版社), 2014: 67-78. ISBN: 978-7-03-041280-5.
[22] Huang P, Sheng X. A number of scattering characteristics of atmospheric aerosol particles in troposphere and stratosphere[J]. *J. Meteorol. Sci*, 1996, **16**(3): 233-239.
[23] Achanta R, Shaji A, Smith K, *et al*. SLIC superpixels compared to state-of-the-art superpixel methods[J]. *IEEE Trans. Pattern Anal. Mach. Intell*, 2012, **34**: 2274-2281.
[24] Kim J H, Jang W D, Sim J Y, *et al*. Optimized contrast enhancement for real-time image and video dehazing[J], J. *Vis. Commun. Image Represent*, 2013, **24**: 410-425.
[25] Zhu H L, Zhang X Z, Chen X, *et al*. Dim small targets detection based on horizontal-vertical multi-scale grayscale difference weighted bilateral filtering[J]. *J. Infrared Millim. Waves*, 2020, **39**(4): 513-522
[26] Fattal R, Agrawala M, Rusinkiewicz S. Multiscale shape and detail enhancement from multi-light image collections [J]. *ACM Trans. Graph*, 2007, **26**(3): 511-519.
[27] Zheng M, Qi G, Zhu Z, *et al*. Image dehazing by an artificial image fusion method based on adaptive structure decomposition[J]. *IEEE Sens*. J, 2020, 20: 8062-8072.
[28] Zhang J, Wang P, Gong F, *et al*. Content-based superpixel segmentation and matching using its region feature descriptors [J]. *IEICE Trans. Inf. Syst*, 2020, **103**: 1888-1900.
[29] Narasimhan S G, Nayar S K. Contrast restoration of weather degraded images [J]. *IEEE Trans. Pattern Anal. Mach. Intell*, 2003, **25**(6): 713-724.
[30] Li L, Sang H, Zhou G, *et al*. Instant haze removal from a single image[J]. *Infrared Phys. Technol*, 2017, **83**: 156-163.
[31] Long J, Shi Z, Tang W, *et al*. Single remote sensing image dehazing [J]. *IEEE Geosci. Remote Sens. Lett.*, 2014, **11**: 59-63.
[32] Baig N, Riaz M M, Ghafoor A, *et al*. Image dehazing using quadtree decomposition and entropy-based contextual regularization[J]. *IEEE Signal Process. Lett*, 2016, **23**: 853-857.
[33] Shin J, Kim M, Paik J, *et al*. Radiance-reflectance combined optimization and structure-guided ℓ_0 - *norm* for single image dehazing [J]. *IEEE Trans. Multimed*, 2020, **22**: 30-44.
[34] Qin B, Huang Z, Zeng F, *et al*. Fast single image dehazing with domain transformation-based edge-preserving filter and weighted quadtree subdivision[C]. Proceedings of the IEEE International Conference on Image Processing (ICIP), IEEE, 2015: 4233-4237.
[35] Li Y, Chen X. A coarse-to-fine two-stage attentive network for haze removal of remote sensing images[J]. *IEEE Geoscience Remote Sensing Letters*, 2020, **18**(10):1-5.
[36] Hautiere N, Tarel J P, Aubert D, *et al*. Blind contrast enhancement assessment by gradient ratioing at visible edges[J]. *Image Anal. Stereol*, 2011, **27**: 87-95.
[37] Yu X, Xiao C, Deng M, *et al*. A classification algorithm to distinguish image as haze or non-haze[C]. Proceedings of the International Conference on Image and Graphics (ICIG), IEEE, 2011: 286-289.