Optical Engineering

SPIEDigitalLibrary.org/oe

Display and detail enhancement for high-dynamic-range infrared images

Chao Zuo
Qian Chen
Ning Liu
Jianle Ren
Xiubao Sui

Display and detail enhancement for high-dynamic-range infrared images

Chao Zuo
Qian Chen
Ning Liu
Jianle Ren
Xiubao Sui
Nanjing University of Science and Technology
National Defense Key Laboratory of Optoelectronic Engineering Nanjing
Jiangsu Province, China, 210094
E-mail: surpasszuo@163.com

Abstract. Dynamic range reduction and detail enhancement are two important issues for effectively displaying high-dynamic-range images acquired by thermal camera systems. They must be performed in such a way that the high dynamic range image signal output from sensors is compressed in a pleasing manner for display on lower dynamic range monitors without reducing the perceptibility of small details. In this paper, a new method of display and detail enhancement for high dynamic range infrared images is presented. This method effectively maps the raw acquired infrared image to 8-bit domain based on the same architecture of bilateral filter and dynamic range partitioning approach. It includes three main steps: First, a bilateral filter is applied to separate the input image into the base component and detail component. Second, refine the base and detail layer using an adaptive Gaussian filter to avoid unwanted artifacts. Then the base layer is projected to the display range and the detail layer is enhanced using an adaptive gain control approach. Finally, the two parts are recombined and quantized to 8-bit domain. The strength of the proposed method lies in its ability to avoid unwanted artifacts and adaptability in different scenarios. Its great performance is validated by the experimental results tested with two real infrared imagers. [DOI: 10.1117/1.3659698]

Subject terms: contrast enhancement; detail enhancement; infrared images; high dynamic range.

Paper 110809R received Jul. 11, 2011; revised manuscript received Oct. 12, 2011; accepted for publication Oct. 21, 2011; published online Nov. 18, 2011.

1 Introduction

Modern high-quality infrared cameras are able to produce images that have a wide dynamic range. Raw sensor data are often acquired at the 12- to 14-bit level that typically exceeds the 8-bit sensitivity of a state-of-the-art display device. Furthermore, a human observer can distinguish only about 128 levels of gray (7-bit) in an image.[1] Hence, a procedure aimed at reducing the data range must take place in the processing stage to enable the display to work with data from the detector. This procedure must accomplish two goals: reduce the dynamic range of the input image into a low range one that is acceptable for the display system and do this in such a manner that the output image is pleasing to the human observer.

Contrast enhancement for infrared images has been widely investigated by many researchers, and a number of useful techniques have been proposed in the literature. However, most of these methods[2–6] operate on the low dynamic range image (typically 8-bit) and only consider the case of still images. It is clear that compression and enhancement on the raw infrared data with a high dynamic range (HDR) may be a more meaningful and challenging task since the raw data from the analog-to-digital converter contains the most complete information and details. Sophisticated techniques are required to adapt the raw signal to the display, maintaining, and possibly improving objects' visibility and image contrast.

To our knowledge, automatic gain control (AGC) and histogram equalization (HE)-based methods[1,7–9] are the most widely used display methods in infrared imaging systems. The AGC method removes extreme values and linearly maps the dynamic range onto an 8-bit domain. Histogram equalization is achieved by normalizing the intensity distribution using its cumulative distribution function so that the result image may have a uniform distribution of intensity. But it typically increases contrast in the dominating temperature/irradiance range and may suffer from some drawbacks such as overenhancement, increase in the noise level, lost in details, and washed-out effect in some almost homogeneous area.[7,8] To overcome these problems, plateau histogram equalization[9] has been proposed to display infrared images by suppressing the enhancement of homogeneous regions with a plateau threshold value. In addition to these global histogram equalization methods (we call these techniques global histogram equalization methods since the whole image uses only one gray-level mapping function), local histogram equalization techniques can also be applied to display wide dynamic range infrared images. Adaptive histogram equalization (AHE)[7] is capable of improving an image's local contrast, bringing out more detail in the image by computing the histogram of a local window centered at a given pixel to determine the mapping for that pixel. Meanwhile, it also can produce significant noise. A generalization of AHE, contrast limited AHE (CLAHE),[8] has more flexibility in choosing the local histogram mapping function. By selecting the clipping level of the histogram, undesired noise amplification can be reduced. In summation, these HE-based methods could compress the dynamic

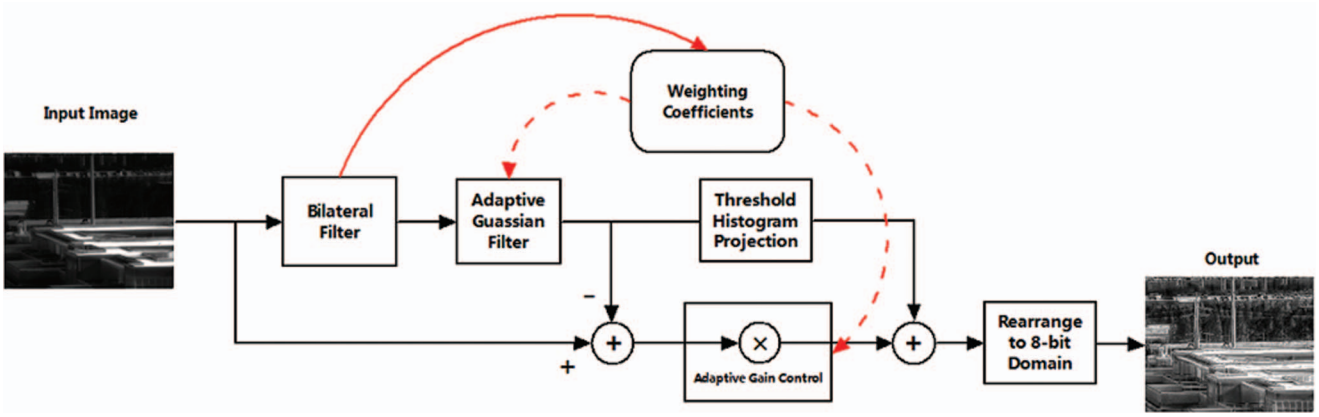


Fig. 1 Block diagram of the proposed algorithm. The curved arrow symbols shown in red denote the parameters of the adaptive Gaussian filter and the gain for detail layer are determined by the weighting coefficients extracted from the bilateral filter.

range of the raw images more effectively than AGC. However, they lack flexibility in manipulating small details of the raw input images since they are based only on histogram information.

Considering these problems, advanced methods[10–12] combining dynamic range compression and details enhancement are proposed to improve the visualization of infrared images. The balanced CLAHE and contrast enhancement (BCCE)[10] and bilateral filter and dynamic range partitioning (BF&DRP)[11] are two methods for visualizing high dynamic range infrared images proposed by Branchitta et al. More recently, Karali et al.[12] developed a technique that increases the contrast of the sea-surface based on local frequency cues. In the BF&DRP method, bilateral filtering[13] is adopted to separate the detail layer from the base component, and then the two components are handled respectively. It has been reported to be superior to the existing methods including BCCE.

Normally the BF&DRP method performs well, but as stated in Ref. 11, a careful tuning of many parameters is needed for the specific scenario so as to obtain the best visualization performance. In addition, we found in our experiments that it sometimes produces severe gradient reversal artifacts and highlights the noise in flat regions. So its performance is scene dependent, which may be unsuitable for digital video camera applications. It is therefore very desirable to have an efficient display technique for infrared cameras that enhance the global contrast and perceptibility of details without highlighting the noise and introducing unwanted artifacts. For this reason, a new method of display and detail enhancement for high dynamic range infrared images is presented. This method can be viewed as an enhanced version of BF&DRP since it is based on the similar framework as that of the BF&DRP method. But it includes the following improvements: First, an adaptive Gaussian filter is introduced as post-processing for bilateral filtering. Second, the base component is compressed using modified histogram projection with a threshold. Finally, detail layer is enhanced using an adaptive gain control process. The aforesaid algorithms could not only improve the method's performance and adaptability in different scenes, but also prevent noise amplification and unwanted artifacts effectively.

2 Principle of the Proposed Algorithm

Inspired by the successful applications of bilateral filter in tone reproduction for high dynamic range images in the visible spectral region,[10,13] Branchitta et al.[11] use a BF&DRP framework to visualize high dynamic range infrared images. The framework relies on bilateral filtering to separate a base component, which contains the large amplitude variations and must be compressed, from a detail component, which must be expanded because it contains the small signal variations related to fine texture. The two components are processed independently and finally recombined. The visualization method for infrared image proposed here uses the similar framework. The intact principle scheme of the proposed algorithm is clearly illustrated in Fig. 1.

The overall process is very simple: first, a bilateral filter[13] is applied to the input frames $\boldsymbol{I}_{in}$ to obtain the image $\boldsymbol{I}_{bf}$.

$$I_{bf}(i,j) = \frac{1}{k(i,j)} \sum_{(i',j')\in S_{i,j}} g_s(i-i', j-j') g_r [I_{in}(i,j) - I_{in}(i',j')] I_{in}(i',j'), \tag{1}$$

where $k(i,j)$ is a normalization term:

$$k(i,j) = \sum_{(i',j')\in S_{i,j}} g_s(i-i', j-j') g_r [I_{in}(i,j) - I_{in}(i',j')]. \tag{2}$$

The notation $(i',j') \in S_{i,j}$ denotes the fact that (i',j') and (i,j) are neighborhooding pixels. Usually g_s is chosen as a normalized Gaussian kernel that sums to one. Also, we use a Gaussian for g_r in the intensity domain. Thus the total mask of the weights $k(i,j)$ is obtained from the product of two different masks in the spatial and intensities domains and its value is between 0 and 1. σ_s and σ_r are two standard deviation parameters defining the extension of the two Gaussian kernels. σ_s determines the size of the considered neighborhood and should be proportionate to the image size. Here we choose 2.5% of the image diagonal. The choice of the range parameter σ_r is more crucial since it indicates the "minimum" amplitude of an edge. If the variation amplitude is less than σ_r, it will be smoothed by the bilateral filter and

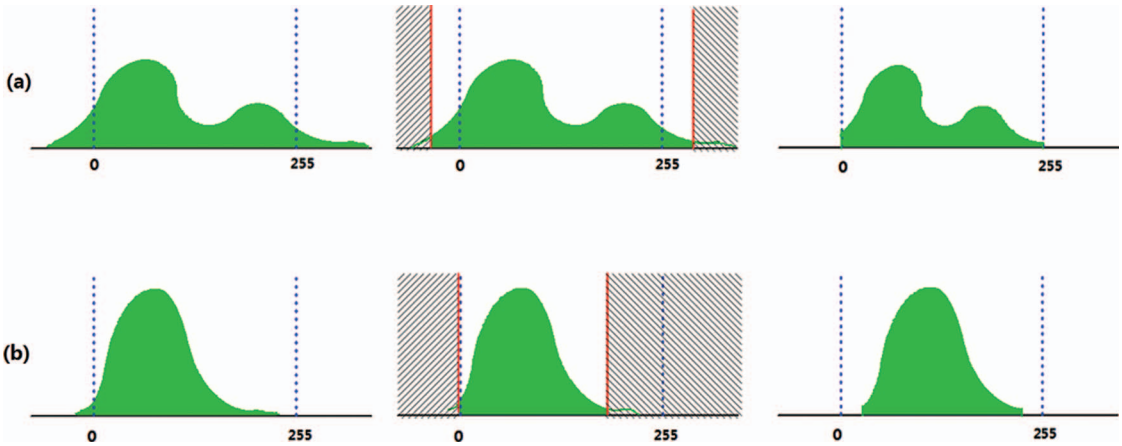


Fig. 2 The two steps of the operator $\lfloor \bullet \rfloor$ illustrated with plots of two histograms. Step 1: We located the 1% minimal and maximal intensity of the input image $\boldsymbol{I}_{BP} + \boldsymbol{I}_{DP}$ and the leading and trailing tails of the histogram are zeroed. (Middle column) Step 2: if the dynamic range of the tail-less histogram is larger than 256, it will be compressed linearly to the display range (a). Otherwise it will only be shifted to the middle of the display range (b) (right column).

the variation will leak into the detail layer. On the other hand, if the oscillation is sharper than σ_r, it will be less altered by the filter. So here we choose σ_r as 20% of the maximum distinguishable gray scales for the human observer (typically about 25). We have noticed that this choice performed consistently well by combining with our compression and detail gain control approach in all our experiments.

After the preliminary separation by a bilateral filter, the image $\boldsymbol{I}_{bf}$ needs post-processing since the bilateral filter may introduce annoying gradient reversal artifacts in detail decomposition. This part will be discussed at length in Sec. 3. The corrected $\boldsymbol{I}_{bf}$, called base layer, is represented by $\boldsymbol{I}_B$. After that, the detail component $\boldsymbol{I}_D$ is merely obtained as

$$I_D(i, j) = I_{in}(i, j) - I_B(i, j). \tag{3}$$

The two components are then processed separately. The base layer is mapped into the proper range using modified threshold histogram projection, and the detail image is enhanced using the adaptive gain control method. These approaches will be presented in Secs. 3 and 4, respectively. The processed base and detail components are represented by $\boldsymbol{I}_{BP}$ and $\boldsymbol{I}_{DP}$. Finally, the two components are recombined and rearranged to get the final 8-bit output $\boldsymbol{I}_{\text{out}}$:

$$\boldsymbol{I}_{\text{out}} = \lfloor \boldsymbol{I}_{BP} + \boldsymbol{I}_{DP} \rfloor . \tag{4}$$

The operator $\lfloor \bullet \rfloor$ means rearranging the image according to its span (dynamic range between the two gray scales that correspond to the locations in the corresponding cumulative density are equal to 0.01 and 0.99) and an overview of the process can be illustrated in Fig. 2. First, the "tails" of the histogram are eliminated, forming a new histogram as described in the middle column of Fig. 2. By removing the tails of the histogram, outlier pixels can be forced into saturation, increasing the contrast in the output image. If the span of the image $\boldsymbol{I}_{BP} + \boldsymbol{I}_{DP}$ is wider than the display range (typically 0 to 255), the tail-less histogram will be mapped linearly to the display range. If the span is narrower than the display range, the tail-less histogram will be shifted to the middle of the display range.

3 Removal of Gradient Reversal Artifacts

The bilateral filter can smooth small fluctuations in intensity while preserving strong edges. While this filter is effective in many situations, it may have unwanted gradient reversal artifacts[14,15] near edges. The reason is that the bilateral mechanism is closely related to a robust iterative procedure (i.e., the mean shift) which achieves edge-preserving filtering by searching for local modes in the joint spatial-range domain. One iteration of the bilateral filter amounts to converge to the local mode.[16] However, when a pixel around an edge has few similar pixels around it, the Gaussian weighted average is quite unstable. The clustering effect forces the blur edges to become sharper and some reversed gradients leak into the detail layers.

Below, we illustrate the reversal artifacts using a synthetic one-dimensional (1D) signal (Fig. 3) and a real two-dimensional (2D) image (Fig. 4). The detail signal in Fig. 3(c) and the detail layer in Fig. 4(b) are yielded by subtracting the bilateral filtered signal/image from their originals. The oversharpened edge causes large fluctuations in the detail layers, which are shown in Fig. 3(c). Also, it is evident that in Fig. 4(b) there are small false edges and blooming at high contrast regions, such as the perimeter of the flower beds. When the detail layer is enhanced and then added back to the base image, some annoying artifacts can be perceived inevitably.

Some researchers attempted to address the inherit shortcomings of the bilateral filter. Durand and Dorsey[14] introduced a fixing strategy using a linear interpolation between the base layer and its smooth version, but they admitted that this method is not so satisfying. Bae et al.[15] addressed this by directly constraining the gradient of the decomposition to prevent reversal, but solving the Poisson equation is quite time consuming. Here we tackle this problem using adaptive Gaussian filtering. Our basic observation is that the gradient reversal artifacts always appear near strong edges. Besides, these edges after bilateral filtering are always sharper than the original. If we blur the sharper edges so that they could get closer to the original, the blurred version is preferable as a base layer since the oversharpened edges are diffused and the gradient reversal artifacts can be eliminated. On the other

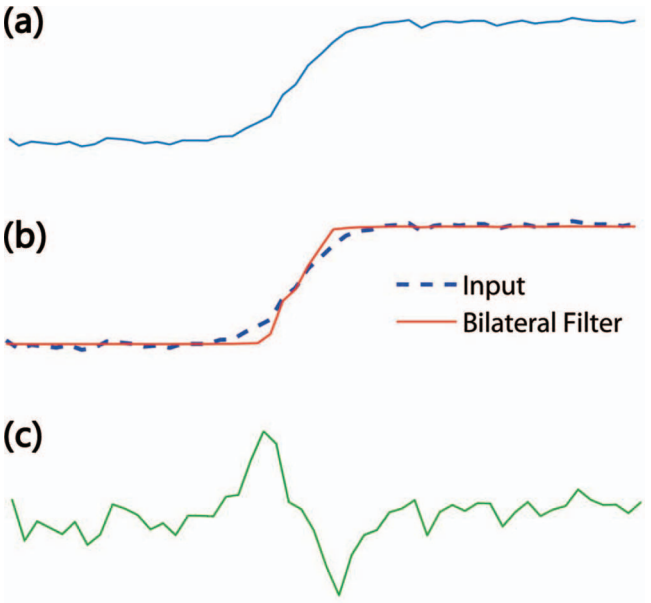


Fig. 3 1D illustration for gradient reversal artifacts. The original signal is shown in blue (a), and the bilateral filtered signal is in red (b). The detail signal (green) yielded by the bilateral filter is amplified (c). The oversharpening causes ringing in the detail layer.

hand, we also hope that other parts in the result of the bilateral filter will be less altered since these regions are less affected by the gradient reversal artifacts. Both of the two aspects inspire us to adopt an adaptive smoothing as post-processing for bilateral filtering.

Gaussian filtering has been intensively used in image processing and computer vision. Mathematically, applying a Gaussian filter to an image is the same as convolving the image with a Gaussian kernel. Since the Fourier transform of a Gaussian is another Gaussian, Gaussian filtering is thus a low pass filter and has the effect of reducing the image's high-frequency components, i.e., smoothing sharp edges. However, the Gaussian filter with a fixed filter standard deviation cannot be used for smoothing the oversharpened edges since it may not fit all regions of the image and even distorts those parts where the details are already properly extracted by a bilateral filter. There are two basic methods to solve the problem. The first method is to process the image using different filters with different standard deviations, and the final result is constructed by selectively blending among those blurred versions. The second method, which we adopted here, is to make the filter standard deviation adapt to the local characteristics of an image.

The adaptive Gaussian filtering is based on the strategy that different parts of an image should be smoothed differently, depending on the degree of oversharpening and the type of edges. The goal of our method is to construct a modified base layer $\boldsymbol{I}_B$ by properly smoothing the base layer $\boldsymbol{I}_{bf}$ to make it more closely with the original input image $\boldsymbol{I}_{\text{in}}$. First, we calculate the weighted difference between the output of bilateral filter and the raw image

$$E(i,j) = k(i,j)[I_{in}(i,j) - I_{bf}(i,j)]. \tag{5}$$

The weighting coefficient $k(i,j)$ is the normalization factor in Eq. (2). It is a direct indicator that whether a certain pixel belongs to the unstable region around strong edges. Thus the difference image $\boldsymbol{E}$ can be regarded as the unwanted distortion caused by bilateral filtering. To fix these errors with adaptive Gaussian smoothing, the filter standard deviation should be properly selected so that the smoothed version is an optimum approximation to the original input image. In other words, the error resulting from the smoothing process should be as close to the difference image $\boldsymbol{E}$. Hodson et al.[17] has shown that a Gaussian filtering of a signal $F(x)$, denoted as $F_g(x)$, can be expressed as

$$F_g(x) = F(x) + \frac{F''(x)}{2}\sigma^2 + \cdots + \frac{F^{(2m)}(x)}{\Pi_{p=1}^{m} 2p}\sigma^{2m} + \cdots, \tag{6}$$

where $m = 1, 2, \ldots, F''(x)$ is the second derivative of $F(x)$. $F^{(2m)}(x)$is the $2m$'th order derivatives of $F(x)$. σ is the standard deviation parameter of the Gaussian filter. Omitting the higher order terms, Eq. (6) can be approximated by

$$F_g(x) \approx F(x) + \frac{F''(x)}{2}\sigma^2. \tag{7}$$

This result can be easily extended to a 2D image since the Gaussian filter is linear and isotropic.[18] So the relationship between the Gaussian smoothed image $\boldsymbol{I}_g$ and its original $\boldsymbol{I}$ can be represented as

$$I_g(i,j) \cong I(i,j) + \frac{\nabla^2 I(i,j)}{2}\sigma^2, \tag{8}$$

where $\nabla^2 I(i,j)$ is the second derivative of $I(i,j)$ and can be approximated as

$$\nabla^2 I(i,j) \approx I(i+1,j) + I(i-1,j) + I(i,j+1) + I(i,j-1) - 4I(i,j). \tag{9}$$

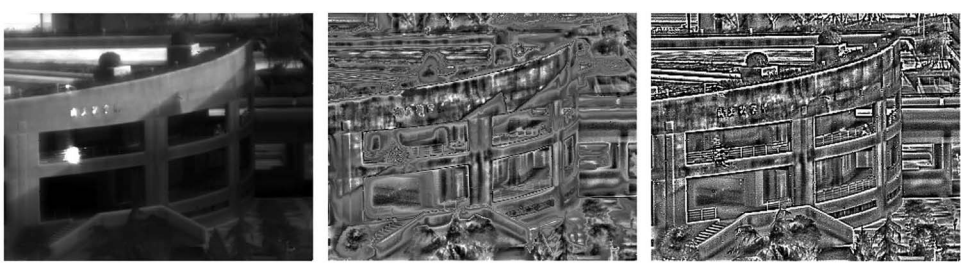


Fig. 4 Gradient reversal artifacts of bilateral filter for a 2D image. (a) Input HDR infrared image displayed by linear scaling (AGC). (b) The detail layer yielded by a bilateral filter. (c) Corrected detail layer by adaptive Gaussian filtering.

Since we use a Gaussian filter to approach the real edges, the difference between $\boldsymbol{I}_{bf}$ and its Gaussian smoothed version should be almost equal to the difference image $\boldsymbol{E}$. Thus, the standard deviation of the adaptive filter should be

$$\sigma(i, j) = \sqrt{\frac{2E(i, j)}{\nabla^2 I_{bf}(i, j)}}. \tag{10}$$

The base layer $\boldsymbol{I}_B$ is the result of an adaptive Gaussian filter of $\boldsymbol{I}_{bf}$. Then the detail layer $\boldsymbol{I}_B$ is obtained accordingly through Eq. (3). In Fig. 4(c), it can be seen that when an adaptive Gaussian filter is used, the quality of the detail layer is noticeably improved and the gradient reversal artifacts go away.

4 Histogram Projection for Base Component

The base component is usually characterized by piecewise smoothness and wide dynamic range. Simple methods can be applied to dynamic range compression of the base layer without considering the problems of losing details and noise amplification. In Ref. 11, the base layer is compressed simply with a gamma curve, allowing a small percentage with extreme pixel values to saturate. But it cannot be guaranteed that the base image has the property that the histogram has a quasiuniform distribution in its dynamic range. Here the histogram projection method is modified to make it more suitable for our purposes.

In the histogram projection method,[1,9] the output range is assigned equally to each valid gray level present, regardless of how many pixels occupy that level. If one specific gray level is occupied by one or more pixels, this level is considered valid. To accomplish this, one needs only to compute an occupancy (binary) histogram $h(x)$ from which the corresponding cumulative distribution function $b(x)$ now represents the fraction of occupied levels at or below the level x. $b(x)$ rises from 0 to 1 in discrete and original values at level x are now mapped into $D \cdot b(x)$, where D is the total range of the output (typically D equals 255).

The histogram projection could effectively reproduce the shape of the raw signal histogram aside from the omission of the unoccupied levels. However, under conditions of scenes with wide dynamic ranges, it tends to produce too little contrast where it is needed since the output range is filled by gray levels occupied by a very small number of pixels. So we improve it as follows: First, the original histogram of $\boldsymbol{I}_B$ is binarized using a threshold T:

$$H(x) = \begin{cases} 0 & \text{for } n_x < T \\ 1 & \text{for } n_x \geq T \end{cases}, \tag{11}$$

where n_x denotes the number of pixels that is resident within the histogram bin for gray level x. The purpose of threshold T is to improve the overall contrast. It is also very useful for reducing the influence of small outliers that could dramatically change the global intensity of the display. When n_x is larger than the threshold T, it means the gray level x can be found frequently in the original image and should be preserved in the output. In our experiments, we have seen that the best results are obtained when T is chosen as 0.1% of the total number of pixels. The cumulative distribution function of $H(x)$ is used as the gray scales transform functions to assign the new intensity values to the input image. The cumulative distribution function $B(x)$ is defined as

$$B(x) = \begin{cases} 0 & \text{for } x = 0 \\ \dfrac{\sum_{y=0}^{x-1} H(y)}{n_{\text{valid}}} & \text{other} \end{cases}. \tag{12}$$

Using this transform function, level x is now mapped into $R \cdot B(x)$ and the maximum range of output image R is

$$R = \min(n_{\text{valid}}, D), \tag{13}$$

where n_{valid} denote the total number of the valid gray levels and $\min(\bullet)$ operator is used to select the minimal value. In this manner, the base image can be mapped into the 8-bit domain properly. Note that we do not simply map the original into the output dynamic range D since the dynamic range of the input is not necessarily higher than the display range. It is quite possible when the camera faces a uniform background, perhaps a uniform wall, or the cloudless sky. Given these possible scenarios, we prefer to keep the dynamic range at its original level.

5 Adaptive Gain Control for Detail Image

As mentioned before, the detail component must be expanded because it contains the small signal variations related to fine structures. However, it also contains most of the noise in the original image signal. If the detail component is magnified directly, it will give rise to an unacceptable amplification of noise in the homogeneous region of the image. In cases where the human observer is the receiver of the output image, the properties of the visual system should be incorporated into the enhancement algorithm in order to obtain visually optimal results.

Psychophysical experiments confirm that noise in flat regions of the image will give rise to spurious or texture to the observer and that at sharp transitions in image intensity the contrast sensitivity of the human visual system decreases with the sharpness of the transition. This masking effect of the human visual system results in lower noise visibility in the complex background. Based on this information, Anderson et al.[19] first defined the noise masking function at coordinate (i, j) as a measure of spatial detail. Then, they performed subjective tests and obtained visibility function $f(i, j)$ at coordinate (i, j) which express the relationship between the visibility of noise and masking function. Katsaggelos et al.[20] adopted the local variance as noise masking function $M(i, j)$ to measure the spatial detail. Then they defined the noise visibility function as:

$$f(i, j) = \frac{1}{M(i, j) \cdot \theta + 1}, \tag{14}$$

where θ is a tuning parameter. The visibility function is normalized and takes the values between 0 and 1. It is clear from Eq. (14) that for the areas with high spatial activity, the visibility function goes to zero and noise is almost unperceivable; while for flat areas the visibility function goes to 1 and noise is in complete exposure. Obviously, in our cases, the normalization term $k(i, j)$ of the bilateral filter is a good indicator of spatial detail. For simplicity, we could use $k(i, j)^{-1} - 1$ the as noise masking function and set θ to 1. We hope that in the smooth region of the image, the gain of the details should be set low to avoid the noise

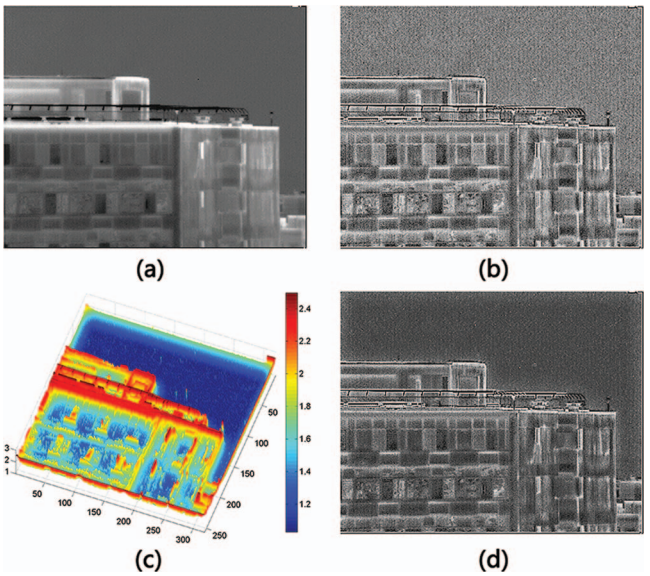


Fig. 5 Adaptive gain control for detail layer. (a) Input HDR infrared image displayed by linear scaling (AGC). (b) The detail layer of (a) without adaptive gain control. (c) Gain matrix for the detail layer shown in (b) ($G_{\min} = 1, G_{\max} = 2.5$). Note that larger values will amplify more the subtleties in detail layer. (d) The detail layer after adaptive gain control. (b) and (d) are scaled to a same range for better display.

overamplification. On the contrary, in the high spatial activity area of the image, the gain of the details should be set high to enhance the image's visual effect.

Assuming the gain of the final details of the images ranges from $G_{\min}$ to $G_{\max}$, according to the noise visibility function, it should be $f(i,j) \to 0, G(i,j) \to G_{\max}$ and $f(i,j) \to 1, G(i,j) \to G_{\min}$. For simplicity in this implementation, we use linear mapping to obtain the detail gain for each pixel as follows:

$$G(i,\,j) = G_{\min} + [1 - f(i,\,j)]\,(G_{\max} - G_{\min}) = G_{\min} + [1 - k(i,\,j)]\,(G_{\max} - G_{\min})\,. \tag{15}$$

The final form is quite simple. In general, in order to prevent the loss of details and avoid noise amplification, $G_{\min}$ can be set as 1. $G_{\max}$ can be selected as needed, but the value must be greater than or equal to $G_{\min}$. A larger $G_{\max}$ allows one to generate more pronounced or even exaggerated details. In our experiments, we have noticed that a value of $G_{\max} = 2.5$ yields satisfactory results. Figure 5 gives an example to demonstrate the effect of the adaptive gain control for an image of a building with a clean background. It can be noted that the values of gains of the sky region are near 1, which prevents amplification of noise in that region, while the building region has larger gains to ensure better enhancement. The detail image $I_D(i, j)$ is multiplied by the corresponding gain coefficient $G(i, j)$ and combined with the compressed base layer using Eq. (4) to get the final result.

Table 2 Mean RMSC results for the four test sequences.

Algorithm	Sequence I	Sequence II	Sequence III	Sequence IV
AGC	45.70	21.32	29.12	41.61
Histogram equalization	72.26	73.84	69.32	71.72
BF&DRP	51.32	34.77	35.67	40.81
Proposed	59.17	66.13	49.72	44.12

6 Experimental Results

In order to validate the proposed approach in different scenarios, we acquired four sets of real infrared data using two infrared cameras. A brief description of the test sequences are summarized in Table 1.

The proposed method is compared with three visualization techniques, i.e., AGC, classic histogram equalization, and the BF&DRP. For the BF&DRP, we use the parameter set as proposed in Ref. 11. In our method we have $G_{\min} = 1$ and $G_{\max} = 2.5$ for all the experiments. The metric used to quantitatively evaluate the enhancement effect of different methods is given by the root-mean-square contrast (RMSC),[21] which is defined as

$$\mathrm{RMSC} = \sqrt{\frac{1}{M \cdot N} \sum_{i,j} (I(i,\,j) - \bar{I})^2}, \tag{16}$$

where $\bar{I}$ is the average intensity of all pixel values in the image. M and N are the image's rows and columns, respectively. Note the RMSC is measured by calculating the average difference between the pixel values and the mean of the image. It reflects the amount of ambiguity in a picture that is measured, and larger values usually indicate better enhancement. The results of the RMSC over each sequence are presented in Table 2. From there, it can be seen that the histogram equalization and the proposed method has higher RMSC in the mean sense.

A performance measurement like RMSC can help, but it does not necessarily indicate good visual effects. Therefore, the video sequences must be watched to perform a subjective evaluation (Video 1, Video 2, Video 3, and Video 4). From

Table 1 Description of test image sequences.

Test sequence	IR cameras	Capture time	Weather	Image characteristics
Dataset 1	320×256 3–5 μm HgCdTe IRFPA	11 a.m.	Very hot	Outdoor building surveillance
Dataset 2	320×256 3–5 μm HgCdTe IRFPA	8 p.m.	—	Human indoor and presence of hot objects
Dataset 3	320×256 8–14 μm HgCdTe IRFPA	10 a.m.	Fine	Outdoor scenery
Dataset 4	320×256 8–14 μm HgCdTe IRFPA	1 p.m.	Cloudy	Aircraft tracking sequences against clear sky background

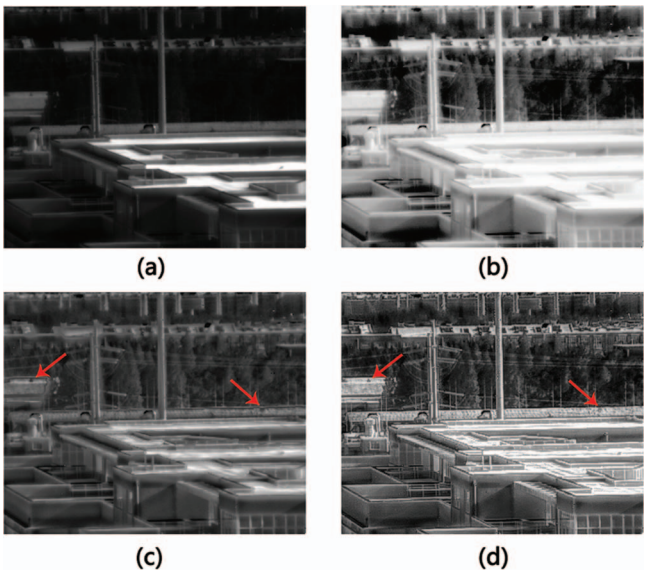


Fig. 6 (Video 1). Comparison results: (a) AGC, (b) histogram equalization, (c) BF&DRP, and (d) proposed method. (QuickTime, 8.70 MB) [URL: http://dx.doi.org/10.1117/1.3659698.1]

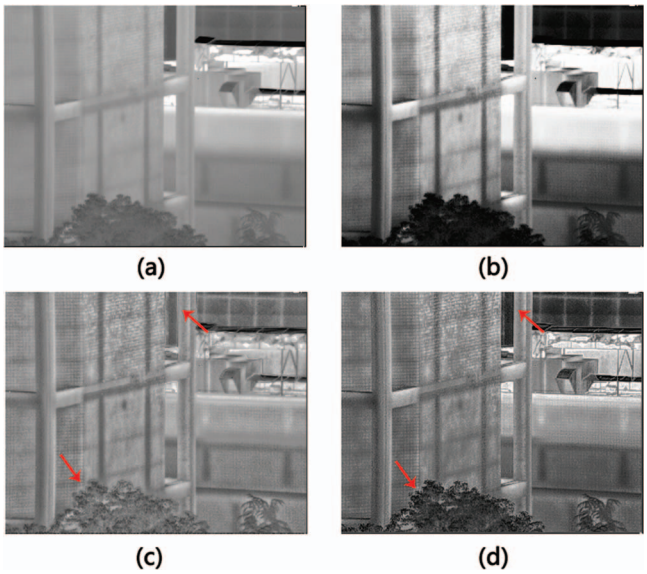


Fig. 8 (Video 3) Comparison results: (a) AGC, (b) histogram equalization, (c) BF&DRP, and (d) proposed method. (QuickTime, 8.81 MB) [URL: http://dx.doi.org/10.1117/1.3659698.3]

these video sequences, it is very noticeable that the proposed method provides the best visual effect and detail enhancement over all four sequences. Besides, it effectively avoids undesirable artifacts and amplification of the noise. Nevertheless, the histogram equalization method underperforms in our subjective evaluation. Frame samples are presented in Figs. 6–9.

The first sample image is chosen as the representative of images with a complex scene and high dynamic range. In the AGC's output, the roof of the building is very bright while the trees are very dark and cannot be readily seen. This is because the raw signals of bright roofs caused by sunshine reflection are much larger than those of the background (the difference is about 2500 digital levels). AGC maps the original dynamic range to the display range linearly, so the large brightness difference is transferred to the output image proportionally. Observing the resulting image of histogram equalization, it can be seen that the overall contrast is improved. However, it also results in loss of details and saturation in the bright regions. BF&DRP gives a more satisfactory result, but it suffers gradient reversal artifacts around some edges, while our method is free of this problem. Furthermore, our algorithm shows a better enhancement in contrast and tiny details.

Figure 7 illustrates a more challenging situation: a side view representation of a human face with an extremely hot object (a soldering iron) in the scene. The dynamic range of the raw signal surpasses 8000 and the digital level of the soldering iron is about 6500 more than those of the human

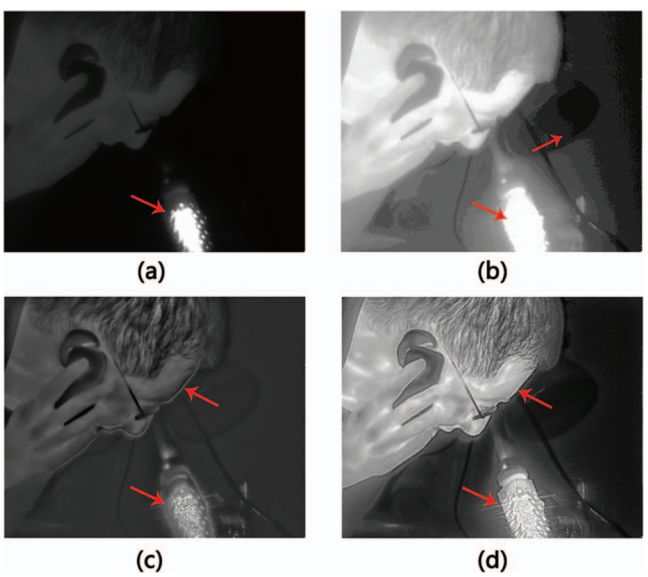


Fig. 7 (Video 2) Comparison results: (a) AGC, (b) histogram equalization, (c) BF&DRP, and (d) proposed method. (QuickTime, 8.65 MB) [URL: http://dx.doi.org/10.1117/1.3659698.2]

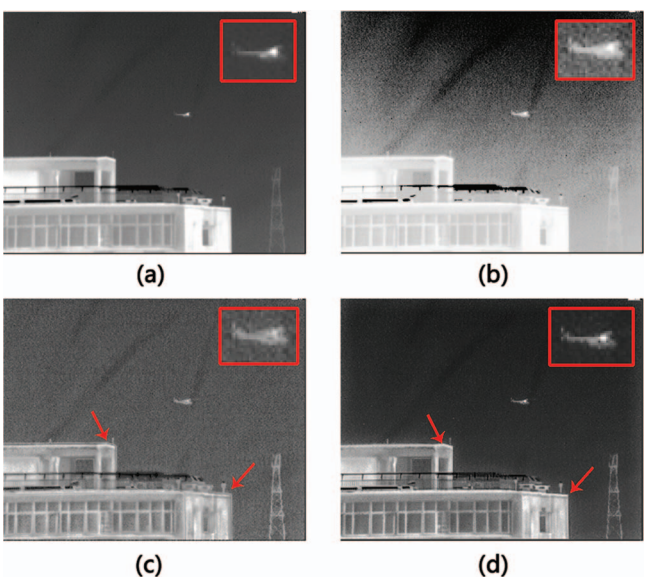


Fig. 9 (Video 4) Comparison results: (a) AGC, (b) histogram equalization, (c) BF&DRP, and (d) proposed method. (QuickTime, 8.66 MB) [URL: http://dx.doi.org/10.1117/1.3659698.4]

face and about 7500 more than the background. Note that in the output of AGC, the background is completely dark and the face is rather faint. Histogram equalization tends to overenhance the background, inducing some stair-step artifacts. Besides, the soldering iron regions are all saturated. The details of human and soldering iron are improved by BF&DRP, but the result is not suitable for observation by the human eye since the image is too dark and without strong contrast. The poor contrast is probably caused by inappropriate compression of the base layer. Because the pixel number of the soldering iron region is much larger than the saturation threshold, compression using a gamma curve may be severely affected by these extreme values, while in Fig. 7(d) the details of the human face and soldering iron are greatly enhanced with the overall contrast improved. And it avoids the gradient reversal artifacts that are noticeable in the result of BF&DRP.

Figure 8 shows a scene characterized by rich details. The AGC version Fig. 8(a) could be considered the most "plain,'' in that it fails to reproduce the details of the wall and bush. Figure 8(b) is the result image by histogram equalization. The wall on the right is too bright while the background window and the tree are too dark. Also, the details are not improved. The overall contrast of BF&DRP's result is a resemblance to that of the AGC, while the details show a great improvement. But the fake contours caused by gradient reversal make one uncomfortable. A result from the proposed method indicates that not only the details are enhanced but also the contrast of the scene is significantly improved. The bricks on the wall at the leaves of the bush are the clearest to recognize.

Figure 9 depicts a helicopter hovering in the sky with a simple and a little noisy background. We have zoomed in the helicopter regions of each image to draw a better comparison. AGC controls the noise level very well, but the object of interest, which is the helicopter, appears to be a little weak and blurry. It is obvious from Fig. 9(b) that background and noise is overenhanced by histogram equalization, resulting in unpleasant artifacts in the sky regions. BF&DRP enhance the image details as well as noise. The gradient reversal artifacts affect the visual effect once again and change the shape of the helicopter body. Our method gives the most satisfactory result: the helicopter is well enhanced, and we could clearly identify its shape; the details and contrast of the building region and iron tower are greatly improved. Further, our method shows a great ability of noise suppression.

7 Conclusions

In this paper, we have presented a new display technique for high dynamic range infrared images. The technique, similar to the recently proposed BF&DRP, is based on separating the acquired image into two components: base part and detail part using a bilateral filter, then processing each part independently. To avoid unwanted gradient reversal artifacts, the base layer is modified by adaptive Gaussian filtering. Then the base layer is projected to the display range, and the detail layer is added back after an adaptive gain control process.

The performance of the proposed method is tested with a mid-wavelength infrared camera and a long-wavelength infrared camera with a default setting of the parameters. The experimental results validate that our method can effectively reproduce the acquired signal in different scenarios, improve the overall contrast, enhance the target and image details, avoid halos and gradient reversal artifacts, and suppress background noise.

Our proposed method has a very good application prospect among such applications as IR inspection in design, test, and manufacturing, chemical imaging, night vision for driver assistance, surveillance in security, target signature measurement, and tracking. But it also shares some limitations with most nonlinear visualization methods: it may not work for some particular conditions such as temperature measurement where the radiometric accuracy is needed.

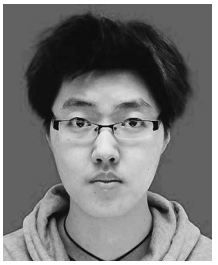
Acknowledgments

This project was supported by the Research and Innovation Plan for Graduate Students of Jiangsu Higher Education Institutions, China (Grant No. CXZZ11_0237) and China Postdoctoral Science Foundation (Grant No. 20110491424).

References

1. J. Silverman, "Display and enhancement of infrared images," in *International Conference on Image Processing and its Applications*, pp. 345–348 (1992).
2. T. H. Yu, Q. M. Li, and J. M. Dai, "New enhancement of infrared image based on human visual system," *Chin. Opt. Lett.* **7**(3), 206–209 (2009).
3. X. Zhou, R. A. Liu, and F. Chen, "Infrared and visible image fusion enhancement technology based on multi-scale directional analysis," *Proceedings of the 2009 2nd International Congress on Image and Signal Processing*, Vols. 1–9, pp. 4035–4037 (2009).
4. R. Lai, Y. T. Yang, B. J. Wang, and H. X. Zhou, "A quantitative measure based infrared image enhancement algorithm using plateau histogram," *Opt. Commun.* **283**(21), 4283–4288 (2010).
5. X. B. Bai, X. Z., F. G. Zhou, and B. D. Xue, "Infrared image enhancement through contrast enhancement by using multiscale new top-hat transform," *Infrared Phys. Technol.* **54**(2), 61–69 (2011).
6. C. L. Lin and C. L., "An approach to adaptive infrared image enhancement for long-range surveillance," *Infrared Phys. Technol.* **54**(2), 84–91 (2011).
7. S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. T. H. Romeny, and J. B. Zimmerman, "Adaptive histogram equalization and its variations," *Comput. Vis. Graph. Image Process.* **39**(3), 355–368 (1987).
8. K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems IV*, pp. 474–485, Academic Press Professional, Inc., San Diego (1994).
9. V. E. Vickers, "Plateau equalization algorithm for real-time display of high-quality infrared imagery," *Opt. Eng.* **35**, 1921 (1996).
10. F. Branchitta, M. Diani, G. Corsini, and A. Porta, "Dynamic-range compression and contrast enhancement in infrared imaging systems," *Opt. Eng.* **47**(7), 076401 (2008).
11. F. Branchitta, M. Diani, G. Corsini, and M. Romagnoli, "New technique for the visualization of high dynamic range infrared images," *Opt. Eng.* **48**(9), 096401 (2009).
12. T. Aytac, A. O. Karali, and O. E. Okman, "Adaptive enhancement of sea-surface targets in infrared images based on local frequency cues," *J. Opt. Soc. Am. A* **27**(3), 509–517 (2010).
13. C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proceedings of the Sixth International Conference on Computer Vision*, p. 839, IEEE Computer Society (1998).
14. F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," *ACM Trans. Graphics* **21**(3), 257–266 (2002).
15. S. M. Bae, S. Paris, and F. E. Durand, "Two-scale tone management for photographic look," *ACM Trans. Graphics* **25**(3), 637–645 (2006).
16. J. Van de Weijer and R. Van den Boomgaard, "Local mode filtering," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, , Vol. 422, pp. II-428–II-433 (2001).
17. E. Hodson, D. Thayer, and C. Franklin, "Adaptive Gaussian filtering and local frequency estimates using local curvature analysis," *IEEE Trans. Acoust., Speech, Signal Process.* **29**(4), 854–859 (1981).
18. G. Deng and L. W. Cahill, "An adaptive Gaussian filter for noise reduction and edge detection," in *Nuclear Science Symposium and Medical Imaging Conference*, IEEE Conference Record, Vol. 1613, pp. 1615–1619 (1993).
19. G. L. Anderson and A. N. Netravali, "Image restoration based on a subjective criterion," *IEEE Trans. Syst. Man Cybern.* **6**(12), 845–853 (1976).
20. A. K. Katsaggelos, J. Biemond, R. W. Schafer, and R. M. Mersereau, "A regularized iterative image restoration algorithm," *IEEE Trans. Signal Process.* **39**(4), 914–929 (1991).

21. E. Peli, "Contrast in complex images," *J. Opt. Soc. Am. A* **7**(10), 2032–2040 (1990).



Chao Zuo received his BS degree from Nanjing University of Science and Technology, China, in 2009. He is currently pursuing his PhD degree in the School of Electronic Engineering and Optoelectronic Techniques, Nanjing University of Science and Technology, China. His research interests include signal and image processing, algorithms for infrared spectral sensors and imagers, digital holography, and three-dimensional shape measurement. He is a student member of IEEE and SPIE.

Qian Chen received his BS, MS, and PhD degrees from the School of Electronic Engineering and Optoelectronic Techniques, Nanjing University of Science and Technology. He is currently a professor and a dean of the Department of Optoelectronic Technology, Nanjing University of Science and Technology. He has led many research projects and authored more than 100 journal papers. His works have covered different topics, such as real-time digital image processing, optoelectronic imaging, electro-optical displaying technology, optoelectronic signal processing, and transmission. He is a member of SPIE and SID.

Ning Liu received his BS degree from Nanjing University of Science and Technology, China, in 2007. He is a PhD candidate of Nanjing University of Science and Technology, Institute of Optoelectronics and Electric Techniques. He is interested in imaging theory of thermal detector and the ultralow noise proxy technology.

Jianle Ren received his BS degree from Nanjing University of Science and Technology, in 2009. He is a PhD candidate in optical engineering from the School of Electronic Engineering and Optoelectronic Techniques, Nanjing University of Science and Technology. He is interested in infrared image processing and image fusion.

Xiubao Sui received his PhD degree in optical engineering from the School of Electronic Engineering and Optoelectronic Techniques, Nanjing University of Science and Technology. His research interests include the driver of infrared focal plane arrays and the theory research of infrared detectors.