HALO-FREE DESIGN FOR RETINEX BASED REAL-TIME VIDEO ENHANCEMENT SYSTEM

Yuecheng Li and Hong Zhang
Image Processing Center
Beihang University
Beijing, China
email: yclee0202@gmail.com

ABSTRACT
Retinex becomes a prevailing theory in the field of image enhancement. In the literature, efforts on developing better illumination estimator have been made to alleviate the traditional halo artifacts existing on abrupt intensity changes. FPGA or DSP based real-time realization of Retinex algorithm is also studied. However, it is found that those real-time systems often suffer from a different kind of halo artifacts caused by inter-frame motion. That is because the algorithm is often realized with pipeline structure. In this paper, a complete halo-free design for Retinex based real-time video enhancement system is proposed. Separable bilateral filter, which is further optimized with a three-scale combined spatial Gaussian filter, is adopted to alleviate traditional halo artifacts. And an improved pipeline structure which adopts a motion estimation technique is proposed to restrain the motion caused halo artifacts. The overall design is analyzed and simulated. Experiments verify that the proposed design is able to achieve the goal of halo-free.

KEY WORDS
Retinex, application, halo artifacts, parallel processing, motion estimation.

1 Introduction
Landmark research of the human visual system by Land [1] showed that human visual system is able to practically recognize and match colors under a wide range of different illuminations, a property that is commonly referred to as the Color Constancy phenomenon. By this, the Retinex methodology was motivated and has been widely used as one of the well-known image enhancement schemes for dynamic range compression, color consistency, lightness rendition, face recognition, de-hazing and shadow detection and removal.

In Retinex algorithms, a given image \( S \) is decomposed into two different images, the reflectance image \( R \) and the illumination image \( L \), such that \( S(x, y) = R(x, y) \cdot L(x, y) \) at each pixel \( (x, y) \) in the image domain. Recovering the illumination image directly from a given image is a mathematically ill-posed problem [2], and algorithms proposed in the literature for its solution vary in their way of overcoming this limitation. Single-scale Retinex (SSR), Multi-scale Retinex (MSR) and Multi-scale Retinex with color restoration (MSRCR) based on center/surround mode are proposed by Jobson et al [3], where illumination is estimated by calculating convolutions with Gaussian kernels on the assumption that the reflectance image corresponds to the sharp details in the image while the illumination image is expected to be spatially smooth. Because of the isotropy of Gaussian filters, edges (abrupt intensity changes) are not able to be preserved when illumination is estimated. It results in halo artifacts in the reflectance which is one of the main drawbacks of Retinex based algorithms, especially for greyscale image enhancement. To prevent halo artifacts in the center/surround mode, bilateral filtering (BF) [4] and mean-shift [5] are employed for their nature of edge-preserving. A variation mode based Retinex is presented by Kimmel et al. [6]. Shen and Hwang [7] modify the traditional variation approach combining the traditional envelope requirement and propose a new robust envelope with a gradient dependent weighting function to estimate the illumination. Although halo artifacts are effectively alleviated by those approaches, computational complexity and resources requirement are increased.

Real-time realization of Retinex algorithm for video enhancement is also studied. NASA first implements the SSR and MSR algorithms with Digital Signal Processors (DSP) [8] and establishes its real-time Enhanced Vision System (EVS). Tsutsui et al. propose the VLSI architecture for real-time Retinex video image enhancement based on variation mode [9]. Those implementations share one consensus: a pipeline structure, where current frames reflectance is derived from previous frames illumination, can save frame buffer and system delay. Unless it is in a video enhancement system for static scene where inter-frame motion can be ignored, motion caused mismatch between previous frames illumination and current frame will cause calculation error and then produce a different kind of halo artifacts in reflectance which is termed as motion caused halo artifacts. Its visual appearance is also annoying for operators as will be presented later.

Our work is to design a FPGA based video enhancement system for outdoor scenes under natural sunlight and weather condition, such as residential surveillance. Thus, input images suffer from non-uniform light producing high dynamic range and haze resulting in low contrast.
2 BF Based Retinex and Motion Caused Halo Artifacts

In this section, Retinex based on BF is firstly introduced where separable BF and optimization on multi-scale spatial Gaussian filter are adopted to reduce the computational complexity. Then the motion caused halo artifacts under traditional pipeline structure are analyzed and illustrated.

2.1 Retinex Based on Optimized Separable Bilateral Filter

BF, which is first presented by Aurich and Weule and later described and named by Tomasi and Manduchi [10], is a nonlinear filtering approach computed as a weighted average of each pixels surrounding. The ability of edge preservation makes BF suitable to alleviate halo artifacts existing in traditional SSR and MSR where spatial Gaussian filter functions as the illumination estimator.

Supposing \( I \) is a given gray-scale image and \( p = (p_x, p_y) \) the 2-D pixel position. Supposing \( L \) is the illumination image and \( R \) the reflectance image, BF based Retinex model in logarithmic form can be expressed as:

\[
\log R = \log[I/L] = \log I - \log[BF[I]].
\]

The BF, donated by \( BF[\cdot] \), is defined by:

\[
BF[I_q] = \frac{1}{W_p} \sum_{q \in \Omega} G_s(||p - q||)G_r(||I_p - I_q||)I_q
\]

(2)

where \( \Omega \) is a neighboring domain and \( W_p \) is a normalization factor:

\[
W_p = \sum_{q \in \Omega} G_s(||p - q||)G_r(||I_p - I_q||)
\]

(3)

with \( G_s \) being a spatial Gaussian that decreases the influence of distant pixels, and \( G_r \) a range Gaussian that decreases the influence of pixels \( q \) with an intensity value different from \( I_p \).

The computation is complex and heavy even for a single scale BF. To speed up the BF for application, separable BF was introduced in [11]. The input image is filtered in one dimension and the intermediate result is filtered again in the other dimension. The separable implementation is not only fast but also approximates the true BF reasonably well. Even in the worst case of a 45°-tilt step edge, the separable filter reduces noise and preserves edge very well [11]. Consequently, the separable BF is adopted in our system.

However, a single scale for illumination estimator is not sufficient to produce a result image with sufficient details and visually pleasing appearance. Similar with MSR, multi-scale is necessary to maintain the trade-off of high frequency component (details) and low frequency component (fidelity) for output images [4]. To reduce the computation increased by adopting multi-scale when performing separable BF, only the spatial Gaussian filter is applied on three scales while the range Gaussian filter is conducted on a proper single scale for preserving edges. Furthermore, s-scale combining method is used to optimize the calculation of the spatial Gaussian filter on three scales:

\[
G'_{ms} = (G_{sh} + G_{sm} + G_{sl})/3
\]

(4)

where \( G_{sh}, G_{sm}, \) and \( G_{sl} \) are three spatial Gaussian filters with large, medium, and small scale factors.
The result of combined multi-scale spatial Gaussian filter is illustrated as Figure 2. With the combined multi-scale spatial Gaussian filter, $G'_{ms}$, the optimized separable BF, $BF_{os}$, is constructed in 1-D form as:

$$BF_{os}[I_{qi}] = \frac{1}{W'_{pi}} \sum_{q_i \in \Omega'} G'_{ms}(\| p_i - q_i \|) G'_{r}(\| I_{pi} - I_{qi} \|) I_{qi}$$

(5)

where $i \in \{x, y\}$, $\Omega'$ is a 1-D neighbouring domain, $G'_{r}$ is a single scale range Gaussian filter in 1-D, and $W'_{pi}$ is a normalization factor:

$$W'_{pi} = \sum_{q_i \in \Omega'} G'_{ms}(\| p_i - q_i \|) G'_{r}(\| I_{pi} - I_{qi} \|).$$

(6)

Then, the final expression of Retinex algorithm based on optimized separable BF is:

$$logR = logI - log\{BF_{os}[BF_{os}[I]_y]_y\}.$$  

(7)

And as shown in Figure 3, although single scale range Gaussian filter is used with the optimized separable BF, the halo artifacts are restrained effectively compared to traditional MSR using Gaussian filter as the estimator.

### 2.2 Motion Caused Halo Artifacts

Figure 4 shows the commonly adopted pipeline structure for illumination estimation and reflectance calculation when Retinex algorithm is implemented as in [8] and [9]. Huge frame buffer for input image which inevitably increases system delay is not occupied and the illumination estimation and reflectance calculation can be processed in parallel. Initially, this structure is also applied by our video enhancement system to implement the optimized separable BF based Retinex algorithm.

BF preserves edge for illumination estimation and thus alleviates traditional halo artifacts. However, because a pipeline structure is used for illumination estimation and reflectance calculation, it may fail to satisfy the requirement that the expected reflectance is extracted from the source image with exactly its corresponding illumination. The motion in video will inevitably produce spatial mismatch between previous frames illumination and current source frame. With spatial shift, while obtaining the reflectance, pixels of the source frame are not processed with the illumination value at their corresponding positions. This error produces the motion caused halo artifacts, especially on the edges in the output reflectance.

To illustrate how the mismatch causes halo artifacts, three neighboring frames are taken as example as shown in Figure 5. There is inter-frame motion only on vertical direction among frame $\sharp 1$, frame $\sharp 2$, and frame $\sharp 3$. Those three frames are all enhanced with the illumination of frame $\sharp 2$. From Figure 5, the motion caused halo artifacts can be seen around the edges between building and sky in frame $\sharp 1$ and $\sharp 3$, and motion caused halo artifacts do not exist in frame $\sharp 2$ for it is enhanced with the illumination exactly extracted from it.
3 Improved Pipeline Structure Based on Motion Estimation

As analyzed in section 2, the movement between previous frame and current frame must be taken into consideration when applying Retinex algorithm on real-time system with pipeline structure for moving scenes. For our application, the outdoor surveillance system are positioned high on the building and scene movement is supposed to be inevitable but slow. Conventional ME is adopted to supplement the mismatch among sequences caused by inter-frame motion. Consequently, a modified pipeline structure based on traditional ME for Retinex based video enhancement system is proposed as shown in Figure 6.

ME is defined as searching the best motion vector to represent the displacement from previous frame to current frame. However, conventional ME does not take inter-frame brightness variations into consideration, which causes inefficient accuracy for sequences involving brightness variations. In [12], images are proposed to be transformed into the scaled Retinex domain before performing ME, motion compensation (MC) and prediction error coding because the image contents are generally illumination free in the scaled Retinex domain and are suitable for motion estimation and compensation using conventional motion estimator. Similarly, in the proposed structure, the accuracy of traditional ME algorithm can be improved because the inter-frame brightness normalization is done in Retinex domain. In return, MC can be performed to correct the mismatch between illumination of previous frame and source image of current frame, and can consequently restrain the motion caused halo artifacts.

As the ME is designed to be done for each neighboring frame and it is assumed that there are no moving objects in the scene, most of the inter-frame motion can be treated as global translation. We choose gray-scale projection algorithm as the ME method. The gray-scale projection is a traditional effective algorithm used in electronic image stabilization [13] [14]. It can generate high precision results for inter-frame translation and is easy to realize.

With gray-scale projection algorithm, each reflectance \( R_n \) is firstly project into two independent 1-D curves, the column projection curve and the row projection curve:

\[
\text{Gray}_n(j) = \sum_i R_n(i, j) \quad (8)
\]

\[
\text{Gray}_n(i) = \sum_j R_n(i, j) \quad (9)
\]

where the \( \text{Gray}_n(j) \) and \( \text{Gray}_n(i) \) correspond to the gray projection value of column \( j \) and row \( i \) respectively.

Then cross correlation for the projection value of \( R_n \) is calculated with the projection value of its previous frames reflectance \( R_{n-1} \) on columns and rows to find the motion vectors. With them, MC for the illumination \( L_n \) can be made to correct its spatial mismatch with \( I_{n+1} \) so that the motion caused halo artifacts on \( R_{n+1} \) are alleviated. Those simple and effective ME and MC procedures can definitely be realized and integrated into the overall parallel pipeline structure.

4 Experiments and Discussion

For software simulation in MATLAB, a set of images captured from our real-time video enhancement system are selected to illustrate the effectiveness of gray-scale projection algorithm and verify the alleviation of the motion caused halo artifacts with proposed structure for optimized separable BF based Retinex algorithm.

Totally, 52 images of different scenes under various lightness conditions are used to illustrate the effectiveness of gray-scale projection algorithm. Samples of them
are shown in Figure 7. For simulating inter-frame ME, both the source images and their reflectance are manually translated with vertical 10 pixels and horizontal 20 pixels respectively. Then the gray-scale projection algorithm is performed with the original images and their translated images to estimate the preset inter-frame motion. Table 1 lists the average accuracy for ME with source images and reflectance images. Compared to being conducted directly on source images, gray-scale projection algorithm results in greater average accuracy for ME while being conducted on the reflectance images. Of course, the accuracy decreases when the movement increases. It matches the conclusion in [12]. Consequently, it is rational to use the traditional gray-scale projection algorithm for the inter-frame ME with reflectance images.

Three sequences from a real scene are presented in Figure 8 for the comparison between the traditional pipeline structure, which is without motion estimation and compensation, and the proposed pipeline structure, where ME and MC are integrated. As shown in Figure 8, it demonstrates the efficiency of the proposed pipeline structure to alleviate the motion caused halo artifacts. The halo artifacts caused by motion exist in the output images (left column) which are the results under traditional pipeline structure without ME and MC. Buildings in the far and characters on the top of building even get ghost shadows. And the intensity of the sky is incorrectly dark. The appearance is visually annoying. However, under the proposed pipeline structure with ME and MC (right column), the pixels on current source frame is matched with those in its previous illumination and thus be enhanced properly.

But on the boundary, mismatch still exists as the same as other applications with global MC. In our system, we just duplicate the mismatched pixels on the boundary of the source frame to the output reflectance. As the inter-frame motion is supposed within a small range and the central pixels usually denote the region of interest (ROI), acceptable performance is obtained compared with the output image under traditional pipeline structure where the motion caused halo artifacts unexpectedly exist in the ROI.

### Table 1

Average accuracy of ME under simulation

<table>
<thead>
<tr>
<th></th>
<th>Source images</th>
<th>Reflectance images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizontal</td>
<td>68%</td>
<td>77%</td>
</tr>
<tr>
<td>Vertical</td>
<td>69%</td>
<td>94%</td>
</tr>
</tbody>
</table>

## 5 Conclusion

In this paper, a halo-free design is presented for Retinex based video enhancement system. Firstly, separable BF is optimized with scale combing strategy and is used to restrain the traditional halo artifacts existing in MSR. Then, the motion caused halo artifacts are analyzed in pipeline structure based video enhancement system. To alleviate this new kind of halo artifacts, a modified system structure with additional motion estimation and compensation is proposed. It takes advantage of the illumination-free reflectance to perform the ME and in return correct the mismatch among frames with MC. Simulation experiments verified that the proposed design is effective in restraining both of the traditional halo artifacts and the motion caused halo artifacts for moving scenes. In the future, the proposed halo-free design will be implemented on our FPGA based real-time video enhancement system with further necessary optimization.
References


