REGION-BASED BACKLIGHT COMPENSATION ALGORITHM FOR IMAGES AND VIDEOS

Dae-Young Hyun*, Jun-Hee Heu†, Chang-Su Kim‡, and Sang-Uk Lee*

*School of Electrical Engineering and Computer Sciences, Seoul National University, Seoul, Korea
†Laboratory of Computational Computer Vision, Georgia Institute of Technology University, Atlanta, USA
‡School of Electrical Engineering, Korea University, Seoul, Korea

ABSTRACT

An algorithm for compensating the effects of backlight in images and videos is proposed in this work. We first determine the region of interest (ROI) to compensate mainly using a saliency map, which is based on darkness, skin color, color and texture prominence features. We then compute the compensating offset value for each pixel. Initial offset values are derived to improve the brightness and the contrast of the ROI, and also to provide temporally consistent output frames in case of the video compensation. Finally, we obtain the final offset values by minimizing an energy function, consisting of a data term and a smoothness term. Simulation results show that the proposed algorithm improves the picture qualities of backlit images and videos efficiently.

Index Terms— Image enhancement, video enhancement, backlight compensation, saliency, and motion reliability.

1. INTRODUCTION

In spite of recent popularity of digital cameras and camcorders, it is still not easy to take high quality images. Especially, when an object is illuminated from the back, it is difficult to acquire a good photograph with proper amounts of illumination on both the foreground object and the background. It is hence essential to develop an efficient algorithm to compensate backlight in images and videos.

Various contrast enhancement techniques have been proposed to improve backlit images. Some approaches, including histogram equalization (HE) [1], global contrast enhancement, and block-based local contrast enhancement, attempt to make the histogram of pixel values as uniform as possible to enhance the contrast. These approaches are effective when the input histogram has a narrow range, but they may amplify noise components as well. Gamma correction [2] increases the contrast of dark regions by sacrificing the contrast of bright regions. It may lose the details of bright regions.

The retinex approach [3] estimates the light source of an input image, and eliminates the effects of the light for a clearer image. It, however, causes severe artifacts near strong edges of uniform regions. The multiple-scale retinex (MSR) algorithm combines several different single-scale retinex results to provide an output image. However, it is difficult to tune the parameters in MSR.

These general contrast enhancement techniques, however, insufficiently improve the contrast of backlit objects, since they process input images globally. An alternative approach is to divide backlit objects from the background and enhance the objects separately according to the level of lighting. Haruki and Kikuchi [4] divided an image into several regions, and added a compensation value to all pixels in each region. In [5], Lin and Huang utilized the fuzzy c-means algorithm to extract backlit objects and compute their compensation values. But, in [4, 5], unnatural errors can occur around object edges, since backlit objects are simply adjusted by offset values, while the intensities of the background are preserved. In [6], Lin and Liu defined backlight factors using the histograms of bright and dark areas, and used them to determine the curvature of the intensity transformation function. These algorithms, however, are proposed for images only, and are not suitable for video enhancement. If these algorithms are applied to each frame of a video sequence independently, the resulting sequence may contain flickering artifacts since temporal coherence cannot be maintained.

In this work, we propose a backlight compensation algorithm for images and videos. First, we define an energy function, which consists of a data term and a smoothness term. To define the data term, we employ a saliency map that describes the ROI to be compensated mainly. For the image compensation, the compensating offset values are computed to improve the brightness and the contrast of the ROI. For the video compensation, they are computed by considering temporal consistency as well. The smoothness term enforces that neighboring pixels should have similar offset values. By minimizing the energy function, we obtain a high quality output, which enhances the quality of ROI, while preserving the details of the background without artifacts.

The paper is organized as follows. Section 2 describes the proposed compensation algorithm, and Section 3 presents
experimental results. Finally, we discuss our approach and draw conclusions in Section 4.

2. PROPOSED ALGORITHM

Given an input image, the proposed algorithm computes a saliency map to formulate an energy function, and minimizes it to obtain the compensating offset value for each pixel. Let us first describe the image compensation scheme, and then extend it to the video compensation scheme.

2.1. Image Compensation

We enhance a backlit image by adding an offset value \( O_p \) to the luminance \( Y(p) \) of each pixel \( p \). Specifically,

\[
\hat{Y}(p) = Y(p) + O_p. \tag{1}
\]

To compute the offset, we first determine the saliency map, which represents the likelihood that each pixel belongs to a backlit object. Fig. 1 (a) shows an example image, and Fig. 1 (b) is the corresponding saliency map, in which a darker region or human skin and has different color and texture. The saliency map is computed using four features: the luminance prominency \( Y_p(p) \), the skin color indicator \( H(p) \), and the color prominency \( C(p) \) and the texture prominence \( T(p) \). Each feature has a value within \([0, 1]\). \( Y_p(p) = 0 \) when \( p \) has the brightest intensity, whereas \( Y_p(p) = 1 \) when \( p \) has the darkest intensity. The skin color indicator \( H(p) \) has a large value, when \( p \) is on human skin [7]. \( C(p) = 1 \) and \( T(p) = 1 \), when \( p \) has quite different color and texture from its neighbors, respectively. \( C(p) \) and \( T(p) \) are calculated by Itti et al.’s algorithm [8]. Then, we define the saliency of pixel \( p \) by

\[
s(p) = \frac{1}{4} (Y_p(p) + H(p) + C(p) + T(p)). \tag{2}
\]

In other words, \( s(p) \) has a high value, when \( p \) belongs to a dark region or human skin and has different color and texture from its neighbors.

After generating the saliency map, we obtain the average intensity \( Y_m \) of salient pixels by

\[
Y_m = \frac{\sum_{p \in I} s(p) Y(p)}{\sum_{p \in I} s(p)}, \tag{3}
\]

where \( I \) denotes the set of pixels in the image. We intend to adjust the average intensity of salient pixels to 128, and enhance the contrast of salient pixels. To this end, we use a piecewise quadratic polynomial of ‘S’ shape to obtain the initial desired intensity \( \hat{Y}(p) \), given by

\[
\hat{Y}(p) = \begin{cases} 
\frac{[Y(p)+(128-Y_m)]^2}{128} & \text{if } Y(p) \leq 128, \\
\frac{[Y(p)+(255-Y_m)]^2}{128} + 255, & \text{otherwise.}
\end{cases}
\tag{4}
\]

Then, the initial offset \( \hat{O}_p \) is given by

\[
\hat{O}_p = s(p) \left( \hat{Y}(p) - Y(p) \right). \tag{5}
\]

Note that, in Eq. (5), we weight the initial offset by the saliency value \( s(p) \) in order to change the intensity of a salient pixel but to preserve that of a non-salient pixel.

Using the initial offsets, we obtain the optimal offsets that minimize the energy function \( E \), which is given by

\[
E = \sum_{p \in I} E_1(O_p) + \lambda \sum_{(p, q) \in N} E_2(O_p, O_q), \tag{6}
\]

where \( N \) denotes the set of pairs of neighboring pixels, and \( \lambda \) is a weighting factor. In this work, \( \lambda \) is fixed to 5. The data term \( E_1 \) indicates that the final offsets should be close to the initial offsets, and is given by

\[
E_1(O_p) = (O_p - \hat{O}_p)^2. \tag{7}
\]

The second term \( E_2 \) measures the smoothness of neighboring offsets, given by

\[
E_2(O_p, O_q) = \exp \left( -\frac{(Y(p) - Y(q))^2}{2\sigma^2} \right) (O_p - O_q)^2,
\tag{8}
\]

where \( \sigma^2 \) is the variance of the luminance intensities in the image. As two neighboring pixels have similar intensities, their offset differences are multiplied by a bigger weight in the energy function. This smoothness term \( E_2 \) helps to provide more reliable enhancement results without unnatural artifacts.

Finally, the energy function in Eq. (6) is minimized using the graph-cut algorithm, and the intensity of each pixel is updated using Eq. (1). Fig. 1 (c) shows an enhancement example. We see that the proposed algorithm effectively improves the brightness and the contrast of the main object.

2.2. Video Compensation

In the video compensation, we enhance the first frame using the image compensation scheme in the last subsection, and process the rest frames sequentially. Let \( O_p^t \) denote the compensating offset for pixel \( p \) in frame \( t \), and let \( m(p) \) denote the motion vector of \( p \) from frame \( t \) to frame \( t-1 \). Then, we transfer the offset \( O_p^{t-1} \) in the previous frame \( t-1 \) to pixel \( p \) in the current frame \( t \), so that we can obtain temporally consistent outputs. We estimate the motion vector \( m(p) \)
using the intensities \( Y^t \) and \( Y^{t-1} \) of frame \( t \) and frame \( t-1 \) by employing a block matching algorithm. Then, instead of Eq. (7), we employ the following data term

\[
E_1(O'_p) = r(p)(O'_p - O^{t-1}_{p+m(p)})^2 + (1 - r(p))(O'_p - \tilde{O}'_p)^2,
\]

where \( r(p) \) denotes the reliability for the transferred offset \( O^{t-1}_{p+m(p)} \).

It is an ill-posed problem to estimate true motion vectors, and the block matching algorithm may provide unreliable motion vectors, which can cause severe artifacts in resultant images. Fig. 2 (a) shows an example, in which the offset values in the previous frame are directly copied to the current frame using estimated motion vectors. Severe artifacts are observed along the object boundaries and the background building. To alleviate these artifacts, we define the reliability \( r(p) \) in Eq. (9) as follows.

\[
r(p) = \exp \left( -\frac{r_1(p) + r_2(p)}{2\sigma^2} \right), \quad (10)
\]

\[
r_1(p) = \left[ Y^t(p) - Y^{t-1}(p + m(p)) \right]^2, \quad (11)
\]

\[
r_2(p) = \frac{1}{|N(p)|} \sum_{q \in N(p)} \left( O^{t-1}_{p+m(p)} - O^{t-1}_{q+m(q)} \right)^2, \quad (12)
\]

where \( N(p) \) is the set of neighboring pixels of \( p \). Specifically, the reliability becomes small, if the intensity difference between \( Y^t(p) \) and \( Y^{t-1}(p + m(p)) \) is large or if the transferred offset \( O^{t-1}_{p+m(p)} \) is different from the offset \( O^{t-1}_{q+m(q)} \) for a neighboring pixel \( q \).

Then, as in the image compensation, we obtain the offsets for the current frame by minimizing the energy function using the graph-cut algorithm. Finally, we compensate the current frame by adding the offsets to the input intensities.

Fig. 2 (b) illustrates the reliabilities of the transferred offsets, where a brighter pixel depicts a higher reliability. Using these reliabilities, we obtain the compensated result in Fig. 2 (c). We see that the proposed algorithm reduces the artifacts in Fig. 2 (a) effectively, and provides a significantly higher quality output.

3. EXPERIMENT RESULTS

We evaluate the performance of the proposed algorithm on various test images and videos. For comparison, the compensation results of the gamma correction [2], HE, the adaptive HE [1], MSR, and (f) proposed algorithm.

Fig. 3 compares the results on the “Smile” image. As shown in Fig. 3 (b), the gamma correction reduces the visibility of the background, since it increases the contrast of dark regions by compressing that of bright regions. As shown in Figs. 3 (c) and (d), the HE-based methods cannot sufficiently improve the brightness of the woman, although the adaptive HE provides a better result than HE. In Fig. 3 (e), MSR yields a low contrast result, since it estimates light sources incorrectly and excessively increases the intensity values of dark areas. On the other hand, in Fig. 3 (f), the proposed algorithm produces a naturally compensated output with balanced contrast stretching for both the subject and the background.

Fig. 4 compares the compensation results of the 1st frame and 2nd frame in the “Lee” sequence. The conventional algorithms process each frame independently. Fig. 4 (b) provides a hazy background, since the gamma correction reduces the contrast of a bright region. Although HE increases the intensities of dark areas effectively, it yields undesirable artifacts in the building and the tree in Fig. 4 (c). In contrast, using the saliency map, the proposed algorithm enhances the luminance of the person more clearly while preserving the details of the background as shown in Fig. 4 (d).

Fig. 5 shows the results on the 4th frame and the 5th frame in the “Sheep” sequence. The adaptive HE cannot preserve the temporal consistency between adjacent frames as shown.
Fig. 4. The compensation of the “Lee” sequence. From top to bottom, the 1st frame, the 2nd frame and the enlarged part of the 2nd frame: (a) original image, (b) gamma correction, (c) HE, and (d) proposed algorithm.

in Fig. 5 (b). MSR provides contour artifacts on the floor behind the doll because of incorrect light source estimation in Fig. 5 (c). On the other hand, the proposed algorithm provides temporally consistent, faithful enhancement results as shown in Fig. 5 (d).

4. CONCLUSIONS

In this work, we proposed an efficient compensation algorithm for backlit images and videos. The proposed algorithm first defines the ROI with the saliency map to mainly compensate, and then determines initial luminance offsets. In the image compensation, the offsets are computed based on the intensity transformation function of ‘S’ shape and the saliency map. On the other hand, in the video compensation, the offsets in the previous frames are transferred and used for computing the offsets in the current frame as well. Simulation results demonstrate that the proposed algorithm increases the brightness and the contrast of backlit objects in images and videos more efficiently than the conventional algorithms.

5. REFERENCES


