Abstract—This paper is concerned with the enhancement of backlit images by compensating for abnormal illumination conditions. The underexposed (backlit) or/and overexposed regions in a backlit image are identified by a soft binary segmentation process that is driven by a Gaussian mixture model. Optimal tone-mapping is performed on the under- and over-exposed regions separately to improve the visual quality. Experimental results demonstrate the efficacy of the proposed restoration method and its advantages over existing image enhancement algorithms in perceptual quality.

I. INTRODUCTION

A common cause of image quality degradation is back-lighting. When photographs are taken under backlit conditions, the backlit foreground objects can be severely underexposed; on the other hand the background objects directly exposed to the light source(s) can be overexposed. Such polarized light exposures can overwhelm the autoexposure mechanism of cameras and lead to poor image quality.

Recently, high-dynamic-range (HDR) imaging techniques [1] have been adopted to compensate for ill illumination conditions described above. In HDR imaging, cameras take two or more photos in succession with different exposure settings; these photos are combined to generate an image of much improved visual quality. However, HDR cannot handle object or/and camera motions nor varying illuminations, as it works under the assumption that the exposure is the only variable that affects image quality; furthermore, HDR can only be used for image acquisition not for image restoration, such as restoring poor quality images taken under poor lighting conditions.

A common approach to restoring backlit images is single-image enhancement. Contrast enhancement methods, such as contrast limited adaptive histogram equalization (CLAHE) [2] and the Retinex algorithm [3], can perform better than HE, but they suffer from objectionable artifacts such as halos.

In this paper, we present a novel single-image restoration algorithm to rectify improper, unbalanced exposures in backlit images. The proposed algorithm first performs a soft binary segmentation of under- and over-exposed regions, called U-regions and O-regions, respectively; then it applies different different tone mapping functions to these two types of regions and fuses the results. The tone mapping function for each-type of regions is determined by solving a constrained optimization problem; the objective is to maximize the contrast while keeping tone reproduction distortions and chromaticity errors under control. The output contrast of either type of regions is automatically adjusted in the optimization process so that no user-specified parameters are required. The segmentation of the U-type and O-type regions is probabilistic (fuzzy decision) to prevent unnatural boundary effects caused by different tone mappings. Fig. 1 shows an example presenting the flow of the proposed method.

The paper is organized as follows. Section II details a soft binary classifier based on a Gaussian mixture model (GMM) for intensity distributions; the classifier segments an input image into U- and O-regions. In section III we discuss the enhancement algorithm applied on either U-region and O-regions. Section IV demonstrates the performance of the proposed method versus several existing image contrast enhancement methods.

II. SOFT SEGMENTATION OF BACKLIT IMAGES

In order to segment a backlit image $I(x, y)$ into U-region(s) and O-region(s), we first separate the luminance signal $L(x, y)$ from $I(x, y)$ based on the following image formation model

$$I(x, y) = L(x, y) \cdot R(x, y)$$

$$\Rightarrow \log I(x, y) = \log L(x, y) + \log R(x, y)$$

(1)

where $R(x, y)$ is the surface reflectance at $(x, y)$ that depends on the geometric and physical properties of the object surface, independent of $L(x, y)$. As in most natural scenes, the
luminance signal $L(x, y)$ is low-pass whereas $R(x, y)$ tends to be high-pass, we apply the non-linear homomorphic filtering technique [4] to obtain an estimated luminance image $L(x, y)$.

One way of forming the $U$-region(s) and $O$-region(s) in $I(x, y)$ is binary thresholding of the histogram of the luminance image $L(x, y)$ into low and high intensity ranges. But better results can be obtained by a two-component Gaussian mixture modeling (GMM) technique, with the two components corresponding to the intensity distributions of the $U$-region(s) and $O$-region(s).

Specifically, denote by $p(c)$ the three-dimensional distribution of colors $c$ in the CIELab space, a perceptually uniform color space that has a linear relationship with human perception of colors [4]; $p(c)$ is modeled by a 2-component GMM, namely the linear combination of two Gaussian distributions:

$$p(c) = \sum_{i} P(w_i)p(c|w_i), i \in \{O, U\}$$

where $p(c|w_i) \sim N(\mu_i, \Sigma_i)$ is the PDF of the Gaussian component $w_i$ (including the $U$-region component $w_U$ and the $O$-region component $w_O$) specified by mean $\mu_i$ and covariance matrix $\Sigma_i$, and $P(w_i)$ is the prior probability of the pixels generated from $w_i$. The GMM model is constructed by estimating the unknown parameters $\theta = \{P(w_i), \mu_i, \Sigma_i\}|_{i \in \{O, U\}}$ via maximizing the likelihood function $L$ with respect to the luminance signal $L(x, y)$ and $\theta$:

$$\hat{\theta} = \arg \max_{\theta} L(L; \theta)$$

$$= \arg \max_{\theta} \prod_{(x,y)} p(L(x,y); \theta).$$

We apply the expectation-maximization (EM), a classical iterative algorithm that guarantees to increase the likelihood function $L$ through iterations until convergence [5], to solve the problem above.

The proposed fuzzy binary classifier $\phi(c)$, which is a likelihood that color $c$ falls in a $U$-region, is then given by

$$\phi(c) = P(w_U)p(c|w_U).$$

Here $\phi(c)$ ranges between 0 (if color $c$ is generated solely by the $O$-component in GMM) and 1 (if color $c$ is generated solely by the $U$-component in GMM). The larger $\phi(c)$ is, the more likely $c$ belongs to the $U$-region.

Finally, the segmentation results are refined by the guided filter [6] so that they align closely with object boundaries in the original image. Although the GMM is built on the statistics of the luminance signal only without considering image semantics, it is effective to separate the $U$-regions (the backlit objects) from the $O$-regions (the frontlit objects).

The above described process of homomorphic filtering and GMM modeling is explained in Fig. 1, where Fig. 1 (b) is the homomorphic filter-generated luminance image component of Fig. 1 (a); Fig. 1 (c) is the GMM-generated fuzzy binary segmentation of the luminance image Fig. 1 (b).

### III. Regionwise Enhancement

Based on the fuzzy binary segmentation of a backlit image $I(x, y)$ into the $U$-regions and $O$-regions, we can enhance $I(x, y)$ by applying different tone mapping functions $T_U$ and $T_O$ on the $U$-regions and $O$-regions, respectively, in resemblance to local histogram equalization [2]. For backlit images the intensity histogram of the $U$-regions differs significantly from that of the $O$-regions, denoted by $p_U$ and $p_O$; $p_U$ is much biased towards the dark range, whereas $p_O$ towards the bright range, each having a relatively small dynamic range. Therefore, image details in both $U$-regions and $O$-regions can be enhanced by separate tone mappings $T_U(\cdot)$ and $T_O(\cdot)$ that stretch the dynamic range of $p_U$ and $p_O$, respectively.

The proposed segmentation-based tone mapping scheme can be viewed as an object-adaptive image enhancement technique. The new technique does not suffer from blocking artifacts as in the block-based CLAHE technique. However, to prevent possible artifacts near the boundary between a $U$-region and an $O$-region due to different tone mappings $T_U(\cdot)$ and $T_O(\cdot)$, we apply both tone mappings $T_U(l)$ and $T_O(l)$ to a pixel on or near the boundary, where $l$ is the pixel intensity value, and then fuse the results by the following weighting scheme

$$T(l) = \phi(l)T_U(l) + (1 - \phi(l))T_O(l),$$

the weight $\phi(l)$ is an estimated probability that the pixel of intensity value $l$ is in a $U$-region.

Next we discuss how to optimize the tone mapping function $T_i, i \in \{U, O\}$, given the histograms $p_U$ and $p_O$. Denote by $p = \{p_1, \ldots, p_N\}$ either $p_U$ or $p_O$, for $N$ existing gray levels $1 = \{l_1, \ldots, l_N\}$, $p_k$ being the probability of the gray level $l_k, 1 \leq k \leq N$.

In this work we take the OCTM approach of tone mapping [7], and pose the construction of $T_i$ as the following
constrained optimization problem
\[
\max_{T_i} G(p, T_i(l)) \\
\text{s.t. } D(p, T_i(l)) \leq \tau
\]  \tag{6}

where the objective function \(G(p, T_i(l))\) is a global contrast gain metric; \(D(p, T_i(l))\) is the tone distortion caused by \(T_i\), to be bounded by a user-specified parameter \(\tau\); \(\Theta\) is an upper bound imposed on \(T_i\) for the purpose of controlling chromaticity errors in tone mapping. The development and details of the above enhancement method can be found in a recent paper by Li and Wu [8].

For easier implementation, we write \(T_i\) as:
\[
T_i(l_k) = \sum_{0 \leq j \leq k} s_j, i = 0, \cdots, N - 1
\]  \tag{7}

where \(s_j\) is defined as the increment in output intensity versus one unit step up in input intensity \(l_j\). \(s_j\) must be nonnegative so as not to violate the order of input luminance. In practice we optimize \(T_i\) regarding \(s_j\) and then map it into \(T_i(l)\). Next, we derive the definitions for \(G\) and \(D\) with respect to \(s_j\).

We adopt Weber’s contrast \(\Delta I/I\) [4], based on Weber’s Law that the just-noticable difference in intensity is proportional to the background intensity, as the contrast metric where \(I\) is the background intensity and \(\Delta I\) is the intensity difference of the interested pixels. In a histogram-based context, we get the relative contrast gain \(g\) regarding \(T_i\) at gray level \(l_j\) as:
\[
g(l_j) = s_j/T_i(l_j). \tag{8}
\]

For all pixels at gray level \(l_j\), the background intensity \(I\) should be the corresponding output intensity \(T_i(l_j)\), and the difference \(\Delta I\) can be approximated by \(s_j\) according to its physical meaning. Then we define the global contrast gain \(G\) by the probability-weighed average of \(g(l_j)\):
\[
G(p, T_i(l)) = G(p, s) = \sum_{j=0}^{N-1} \frac{p_j s_j}{T_i(l_j)}. \tag{9}
\]

As to tone distortion \(D(p, T_i(l))\), the definition goes as follows [8]:
\[
D(p, T_i(l)) = D(p, s) = \max_{0 \leq j, k < N} \{T_i(l_j) = T_i(l_k); p_j > 0, p_k > 0\}, \tag{10}
\]

which can be interpreted as the maximum number of adjacent intensity levels that can be mapped into a single one. The larger it is, the more distortions will be observed on details and gradients. To implement \(D(p, T(l)) \leq \tau\), an \(s\)-dependent definition can also be derived:
\[
\sum_{k \leq j \leq k+d} s_j \geq 1, 0 \leq k < N - d \tag{11}
\]

meaning that \(s_j\) cannot be 0 for more than \(d\) continuous input gray levels, where \(d\) is substantially an upper bound of \(D\). Empirically we set \(d\) as 5 in our application.

Upon the base of (9) and (11), (6) can be written as the following optimization problem regarding an \(N\)-dimensional vector \(s\):
\[
\max_{T_i} \sum_{j=0}^{N-1} p_j s_j/T_i(l_j). \\
\text{s.t. } 0 \leq s_j \leq u
\]
\[
\sum_{k \leq j \leq k+d} s_j \geq 1, 0 \leq k < N - d
\]
\[
T_i(l_k) \leq \Theta(l_k), 1 \leq k \leq N,
\]

where the second constraint keeps the output dynamic range constant at \(N\), and \(u\) is the upper limit of \(s_j\) as a much too large \(s_j\) will result in overenhancement. The problem has an efficient dynamic programming solution.

In the application, we adjust \(u\) for the both regions respectively so as to achieve enhancement with a proper strength, given by \(u_i = 0.75 L_i/r_i\), where \(r_i\) is the dynamic range of the pixels belonging to region(s) \(O (\phi < 0.5)\) or \(U (\phi > 0.5)\). Finally, the mapping function \(T_i\) obtained from optimal \(T_U\) and \(T_O\) using (5), is utilized on the image.

IV. EXPERIMENTAL RESULTS

In this section, we compare the performance of the proposed backlit image enhancement method with those of other existing methods, including the classical HE, CLAHE [2], multiscale Retinex algorithm (MSR), together with Li and Wu’s enhancement algorithm [8] applied on the whole image without segmentation. HE and [8] are representatives of the global tone mapping approach; the other two are the approach of local context-based enhancement.

The comparison study is conducted on a set of images from the Internet that consist of various backlighting scenes. Two test images and the corresponding results produced by the aforementioned algorithms are shown in Fig. 2 and 3.

As clearly revealed by the figures, the traditional methods do not produce satisfying enhancement results in backlighting conditions. Global enhancement methods cannot satisfactorily rectify the underexposure and overexposure problems that coexist in backlit images via a single tone mapping function. For examples, HE overexposes the foreground in Fig. 2 severely, but underexposes the foreground in Fig. 3. Li and Wu’s method suffers from the same problem, for instances the overexposed sky in Fig. 2(c) and Fig. 3(c)) and the underexposed person in Fig. 2(c), though it is reported to have superior performance on most natural images versus HE and its variants [8]. On the other hand, the context-based methods, although being able to enhance details, are susceptible to visible artifacts, such as the overenhanced sky and clouds by CLAHE (Fig. 2(d) and Fig. 3(d)) and the halo artifacts in Fig. 2(e), Fig. 3(d) and (e). In contrast, the proposed method enhances both backlit and non-backlit regions without introducing noticeable side effects.
Fig. 2. Enhancement results on a backlit image captured in the sunset. (a) Original; (b) output by classical HE; (c) output by Li and Wu’s method [8]; (d) output by CLAHE [2]; (e) output by MSR [3]; (f) output by the proposed method.

Fig. 3. Enhancement results on a backlit image in which the woman stands in front of the sunlight. (a) Original; (b) output by classical HE; (c) output by Li and Wu’s method [8]; (d) output by CLAHE [2]; (e) output by MSR [3]; (f) output by the proposed method.
V. Conclusion

We proposed a method to rectify and enhance backlit images. First, underexposed (backlit) regions are segmented out of the input image. Then, the image is enhanced by applying two different tone mappings, one to underexposed regions and the other to the rest of the image. The segmentation process is driven by a two-components Gaussian mixture model. On backlit images, the proposed method achieves much superior visual quality to existing image enhancement methods.

REFERENCES


