CONTRAST-ACCUMULATED HISTOGRAM EQUALIZATION FOR IMAGE ENHANCEMENT

Xiaomeng Wu, Xinhao Liu, Kaoru Hiramatsu, and Kunio Kashino

NTT Communication Science Laboratories, NTT Corporation

ABSTRACT

Among image enhancement methods, histogram equalization (HE) has received the most attention because of its intuitive implementation quality, high efficiency, and the monotonicity of its intensity mapping function. However, HE is indiscriminate and overemphasizes the contrast around intensities with large pixel populations but little visual importance. To address this issue, we propose an HE-based method that adaptively controls the contrast gain according to the potential visual importance of intensities and pixels. Observing that in natural scenes image details are usually hidden in darker regions that have noticeable local differences, we formulate the potential visual importance on the basis of the multi-resolution, dark-pass filtered gradients in the image. Experiments show that our method is highly discriminating in terms of noises and trivial image gradients, and it guarantees great global contrast preservation.

Index Terms — Contrast enhancement, dark-pass filtered gradient, histogram equalization, image enhancement

1. INTRODUCTION

The objective of image enhancement is to bring out hidden image details and to increase the contrast within an image with a low dynamic range. This technique is critical to many vision tasks because many computer vision algorithms are primarily designed to accommodate high-quality inputs. Most existing methods of image enhancement can be categorized into one of three groups: histogram equalization, frequency domain methods, and decomposition-based methods.

In the history of image enhancement, histogram equalization (HE) [1] has received the most attention because of its intuitive implementation quality and high efficiency. It aims at deriving an intensity mapping function such that the entropy of the distribution of output intensities can be maximized. Despite its reasonable usefulness, HE is wholly indiscriminate: intensities with large pixel populations are expanded to a larger range even if they are of little visual importance. This issue has been addressed in various ways. CLAHE [2] computes multiple histograms, each corresponding to a distinct section of the image, and uses them to redistribute the intensities. Partition-based HE [3, 4] first partitions the histogram into multiple sub-histograms, e.g., by using local extrema [3] or a Gaussian mixture model [4], and then independently equalizes the sub-histograms. However, the effectiveness of these methods remains situational. Recently, a few HE algorithms have been developed to incorporate spatial information in density estimation. Arici et al. [5] proposed several histogram modification strategies, one of which completely excludes smooth regions from histogram construction. Lee et al. [6] assumed that the intensity pairs that occur frequently in the input should be emphasized more in the output. However, this method may overemphasize the contrast between a signal and a noise that frequently co-occur in an image.

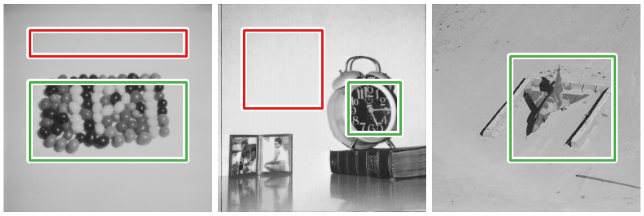
Frequency domain methods [7, 8] transform the signal into, e.g., a gradient domain [7] or a DCT domain [8], to flexibly enhance the contrast in easily controlled ways that are difficult to achieve with conventional image domain techniques. However, they are usually computationally complex because of transformations between the image and frequency domains. In Retinex theory [9], a captured image is taken to be a combination of reflectance and illumination components: the former constitutes the sharp details in the image, while the latter is spatially smooth. There has been some research [10–16] on reflectance as the desired recovery, which can be obtained by estimating and removing the illumination component. Most of these methods do not preserve the trend of global contrast, and so may lead to over-enhanced brightness. Moreover, a graying-out effect may occur, whereby the scene tends to change to middle gray. Unsharp masking algorithms [17, 18] decompose an image into high-frequency and low-frequency components, and realize contrast enhancement by processing the two components individually. They enable global and local contrast enhancement simultaneously, but were reported [19] to be sensitive to parameter settings. In Wang et al.'s method [19], the illumination is estimated through a bright-pass filter, compressed with a bi-log transformation, and recombined with the reflectance to reconstruct the desired recovery. This method enables contrast enhancement with naturalness preserved, but in practice it preserves the trend of global contrast only to a limited extent.

In this study, we focus our attention on HE because of its high efficiency and ability to preserve global contrast. We propose a novel method, called Contrast-ACcumulated Histogram Equalization (CACHE), to address the indiscriminateness issue. CACHE adaptively controls the contrast gain according to the potential visual importance of intensities and pixels. Observing that human vision is more interested in the visual information hidden in darker regions that are of sufficient spatial difference, we formulate the potential visual importance on the basis of the multi-resolution, dark-pass filtered gradients of the image. This formulation leads us to a modified HE, which is highly discriminating in terms of noises and trivial image gradients. CACHE provides great naturalness and consistency as regards image enhancement under diverse illumination conditions, and guarantees greater preservation of global contrast than decomposition-based methods [16, 19].

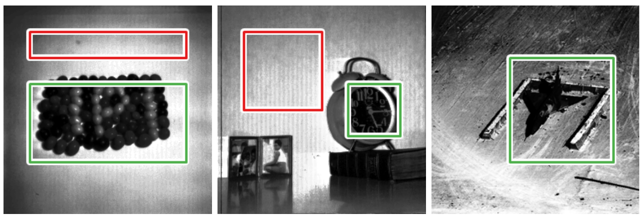
2. CONTRAST-ACCUMULATED HISTOGRAM EQUALIZATION

2.1. Motivation

Consider a discrete grayscale image $\mathbf{A} = \{a(x, y)\}$, with a size of $H \times W$ pixels, where $0 \leq a(x, y) < K$. K is the total number of intensities (typically 256). Let n_k be the number of occurrences of an intensity k. The probability of an occurrence of a pixel of intensity k in $\mathbf{A}$ is



(a) Input



(b) Enhanced images obtained using HE

Fig. 1. Examples of over and under-enhancement caused by HE [1]. Regions of little visual importance are outlined in red, while visually important regions are outlined in green.

$$p_a(k) = \frac{n_k}{n}, \tag{1}$$

where n is the total number of pixels in $\mathbf{A}$ and $0 \leq k < K$. Let us also define the cumulative distribution function (CDF) corresponding to p_a as $P_a(k)$.

We would like to create a transformation of the form $b = T(a)$ to produce a new image $\mathbf{B} = \{b(x, y)\}$. In HE [1], it is assumed that the highest quality images are obtained when they contain the highest possible amount of information, indicating a flat histogram of $\mathbf{B}$. Such an image would have a linearized CDF across the dynamic range. Let $P_b(\cdot)$ denote a CDF corresponding to the normalized histogram of $\mathbf{B}$. HE chooses the transformation T by

$$\begin{aligned} T(k) &= \arg\min_T |P_b(T(k)) - P_a(k)| \\ &= P_b^{-1}(P_a(k)) \\ &= (K-1)P_a(k). \end{aligned} \tag{2}$$

From Eq. 2, the increment in output intensity versus a unit step up in input intensity $k-1, 0 < k < K$, can be easily seen to be

$$\begin{aligned} \Delta T(k) &= T(k) - T(k-1) \\ &= (K-1)p_a(k). \end{aligned} \tag{3}$$

Equation 3 indicates that the increment of intensities, i.e., the contrast gain, is proportional to the probability of the corresponding intensity in the input. This property is usually inconsistent with human vision and is indiscriminate: it may increase the contrast of background noises that have large pixel populations, while reducing the number of usable signals with fewer pixels, as shown in Fig. 1.

To address this issue, we aim at giving an advantage to enhancing the contrast of intensities that are more likely to contain usable signals while attenuating the contrast gain for intensities of little visual importance. To this end, we define the potential visual importance by investigating the spatial information of the image, and incorporate it into density estimation such that the objective described above can be easily accomplished by equalizing the reformulated histogram. We describe the proposed method, CACHE, in detail below.

2.2. CACHE

In contrast to Eq. 1, which depends only on pixel populations, we reformulate the probability of an intensity k in $\mathbf{A}$ by

$$\hat{p}_a(k) = \frac{\sum_x \sum_y \Phi(x, y)\delta(a(x, y), k)}{\sum_x \sum_y \Phi(x, y)}, \tag{4}$$

where Φ is an appropriate spatially variant function expressing the potential visual importance of each pixel and δ is the Kronecker delta. Equation 4 indicates that each pixel contributes to the density estimation adaptively, and $\hat{p}_a(k)$ can be understood as the potential visual importance of an intensity given $\mathbf{A}$. Equalizing this equation instead of Eq. 1 naturally ensures that the increment $\Delta\hat{T}(k)$ of k is proportional to its expected importance. CACHE thus replaces Eq. 1 with Eq. 4 from HE. At stake is how to appropriately define the spatially variant function Φ.

We observe that: 1) potentially important pixels usually have noticeable local differences, i.e., differences from their neighbors (spatial distinctiveness prior) and 2) in natural scenes image details are more likely to lie in darker regions (dark prior). Evidence of the first observation is shown in Fig. 1, where smooth regions are of much less visual importance. The second observation is inherited from the first-order redundancy of natural scenes [20, 21]. In natural scenes brighter pixels usually correspond to light sources or the presence of sky, which are visually less important. Meanwhile, image details are usually darker due to poor lighting conditions, e.g., back-lighting, low-key lighting, underexposure.

Given these key assumptions, we propose formulating the potential visual importance of each pixel by using its dark-pass filtered gradients. Because natural images contain details on multiple scales, we employ a multi-resolution scheme to reliably detect the importance of all significant image details.

We begin by constructing an image pyramid $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_L$, where L is the number of levels and $\mathbf{A}_L$ is a representation of $\mathbf{A}$ at the coarsest level. $\mathbf{A}_l$ at a coarser level can be obtained by downsampling $\mathbf{A}_{l-1}$ by a factor of two with bicubic interpolation. To ensure that the computed gradients do not vary with image resolution, we consider a canonical resolution, which has the same aspect ratio as $\mathbf{A}$ and has a constant short side length S. $\mathbf{A}$ is resized to this canonical resolution, and the resized image is denoted $\mathbf{A}_1$. All the processes required for obtaining the mapping function $\hat{T}$ are executed on the basis of $\mathbf{A}_1$ instead of the original $\mathbf{A}$. With a properly selected S, this strategy also provides a beneficial side effect in terms of efficiency, especially when dealing with large images.

Let $q = (x, y)$ denote the coordinates of a pixel and $\mathcal{N}(q)$ a set of neighboring coordinates of q. At each level l, we compute the dark-pass filtered gradients by

$$\varphi_l(q) = -\sum_{q' \in \mathcal{N}(q)} \min\left(\frac{a_l(q) - a_l(q')}{K-1}, 0\right). \tag{5}$$

Equation 5 attenuates the contribution of pixels in smooth regions to density estimation while amplifying the contribution of spatially distinctive pixels, on condition that the pixel $a_l(q)$ is darker than its neighbors $a_l(q')$. It incorporates both the spatial distinctiveness prior and the dark prior discussed above. The function Φ used in Eq. 4 can thus be computed by merging the gradients φ with their geometric mean:

$$\Phi(x, y) = \left(\prod_{l=1}^{L} \max(U(\varphi_l(x, y)), \varepsilon)\right)^{1/L}, \tag{6}$$

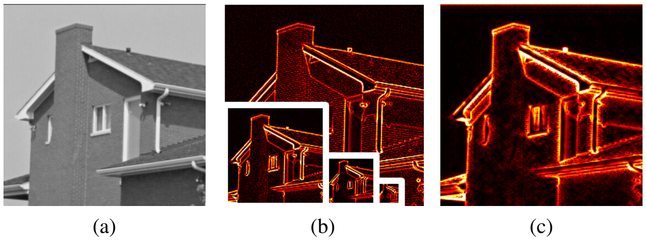


Fig. 2. Dark-pass filtered gradients of input image in (a). Gradients at each level of a four-level Gaussian pyramid are shown in (b). Their geometric mean is shown in (c).

where $U(\cdot)$ is an upsampling operator with a factor 2^{l-1} and ε is a small positive number that is used to avoid negative gradients due to interpolation. The geometric mean actually puts more weight on the gradients at denser levels, which are usually much sparser than the gradients at coarser levels. The geometric mean is adopted instead of the arithmetic mean because in most cases the gradients at denser levels better characterize the spatial distinctiveness of pixels.

Once the potential visual importance Φ has been computed, it is substituted into Eq. 4, which is equalized to obtain the mapping function $\hat{T}$. This function is applied to the original $\mathbf{A}$, and so the equalized image $\mathbf{B}$ has the same resolution as $\mathbf{A}$. In our implementation, we empirically set $L = 4$, $S = 256$, and $\varepsilon = 0.001$. We adopted the von Neumann neighborhood as a set of neighboring coordinates $\mathcal{N}(q)$.

Figure 2 shows the multi-resolution gradients and their geometric mean computed from a four-level Gaussian pyramid. The dark-pass filtered gradients can actually be extended to bright-pass filtered and regular gradients by replacing the negative minimum in Eq. 5 with a positive maximum and with an absolute function, respectively. Figure 3 shows the way in which contrast enhancement, especially for meaningful image details hidden in darker regions, is affected by these different types of gradients.

2.3. Color Image Enhancement

To accommodate color images, we adopt a method that has been widely used in previous studies [7, 22]. Specifically, the color components of the enhanced image are computed by

$$\mathbf{C}_b = \left(\frac{\mathbf{C}_a}{\mathbf{A}}\right)^{\alpha} \mathbf{B}, \tag{7}$$

where $\mathbf{C}$ denotes each RGB component and all the operations are element-wise operations. $\mathbf{A}$ ($\mathbf{C}_a$) and $\mathbf{B}$ ($\mathbf{C}_b$) denote the grayscale (color) component before and after enhancement, respectively. The exponent α (typically 1) controls the color saturation of $\mathbf{C}_b$.

3. EXPERIMENTS

A dataset comprising more than 650 test images, including an SIPI image database [23], an Image Compression benchmark [24], BSDS500 [25], and 95 images obtained from previous studies [16, 19], was used to evaluate and compare CACHE with the literature. We compared CACHE with HE [1], BPDFHE [3], LIME [16], and NPEA [19]. All the implementations of the compared methods were provided by the original authors and are in MATLAB. The experiments were conducted on a machine running CentOS 6.3 with a 160GB memory and a 3.0GHz CPU.

Table 1. Quantitative assessment. The higher the statistics, the better the quality for DE, EME, and PD, which is in contrast to ABME. The best two results per line are italicized.

	Input	HE	BPDFHE	LIME	NPEA	CACHE
DE	*7.2*	6.5	7.1	7.1	*7.2*	*7.3*
EME	15.7	*37.7*	19.5	13.0	14.4	*30.8*
PD	27.9	*42.6*	31.6	27.6	26.0	*42.1*
AMBE	–	25.4	*5.1*	71.8	*22.6*	33.7

3.1. Qualitative Assessment

Some contrast enhancement results for color images are shown in Fig. 4. Not surprisingly, HE produced undesirable effects, e.g., the over-enhancement of background noises in the first line and visible image gradients in the last line. BPDFHE was primarily designed to preserve the mean brightness of the input, resulting in woefully inadequate contrast enhancement when dealing with low-key lighting images. It also produced unnatural gradients as shown in the second line. LIME worked reasonably well for improving the quality of low-key lighting images, but induced the graying-out effect, e.g., in the hair and the black business suit regions in the last line. Because this method eliminates the illumination completely, it results in a loss of global contrast and sometimes causes the brightness to increase excessively. In comparison, NPEA aims at preserving the global contrast within the image. Because of the illumination compression based on a bi-log transformation, this method is sometimes disadvantageous as regards producing satisfactory brightness for low-key lighting images, as shown in the first two lines. The graying-out effect also occurred in the last example. In general, CACHE showed consistently superior or comparable quality for all the examples in Fig. 4 (see the supplementary material for more examples).

3.2. Quantitative Assessment

We assessed the contrast enhancement performance objectively using four quality metrics:

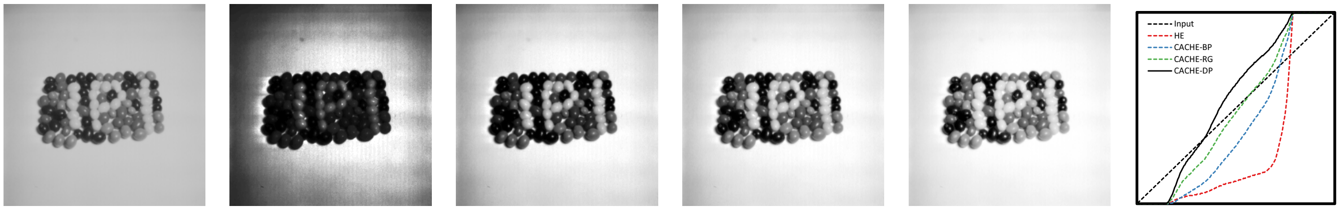
Discrete entropy (DE) [26]. DE measures the amount of information in an image: a high DE indicates that the image contains more variations and is assumed to have better visual quality.

Measure of enhancement (EME) [27]. EME approximates the average local contrast in an image: a high EME is assumed to indicate greater effectiveness in contrast enhancement.

PixDist (PD) [28]. PD computes the average intensity difference over all the pixel pairs in an image: it yields a high score, indicating better quality, when the pixel intensities are uniformly distributed and do not concentrate at particular intensities or at particular spatial positions.

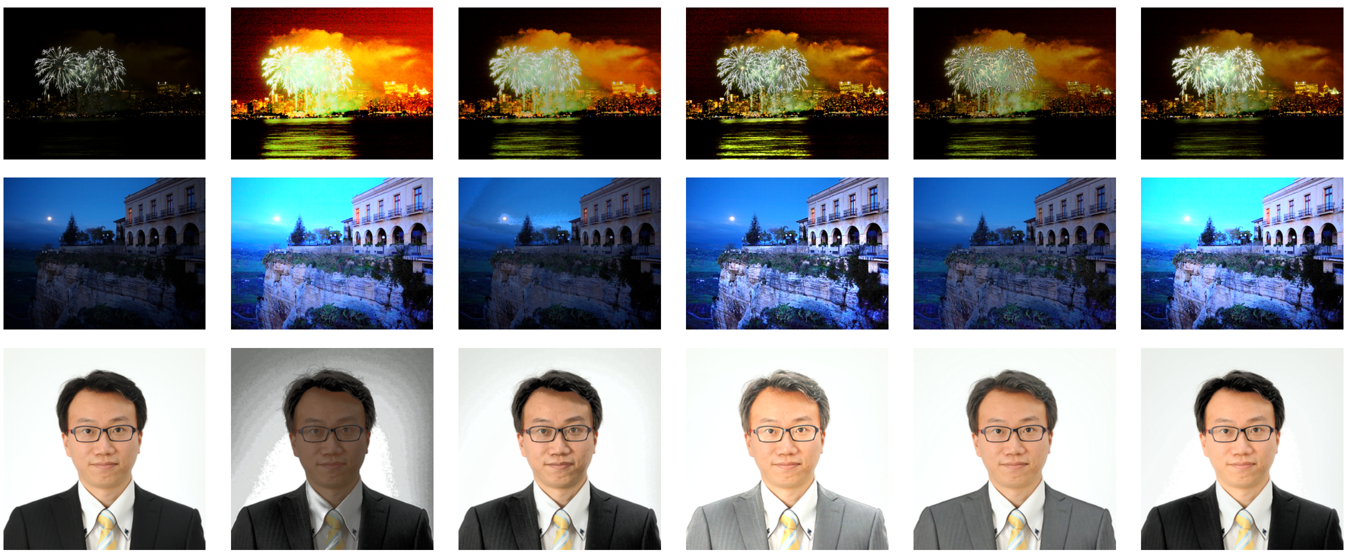
Absolute mean brightness error (AMBE) [29]. AMBE measures the absolute difference between input and output intensity means: a lower score implies better brightness preservation.

Table 1 compares the values measured with different methods averaged over the 500 test images in BSDS500. In general, CACHE achieved the highest DE and the second highest EME and PD among all the methods. These results demonstrate the consistent effectiveness of CACHE in terms of image enhancement from different perspectives. As for AMBE, BPDFHE obtained the best score, as might be expected. CACHE did not show a high score because it does not impose any brightness preserving constraint on image enhancement.



(a) Input (b) HE [1] (c) CACHE-BP (d) CACHE-RG (e) CACHE-DP (f) Mapping functions

Fig. 3. Grayscale image enhancement. "CACHE-BP", "CACHE-RG", and "CACHE-DP" indicate the proposed method based on bright-pass filtered, regular, and dark-pass filtered gradients, respectively.



(a) Input (b) HE [1] (c) BPDFHE [3] (d) LIME [16] (e) NPEA [19] (f) CACHE

Fig. 4. Color image enhancement for low-key lighting and regular images (see the supplementary material for more examples).

Table 2. Processing time in milliseconds for the first line and in seconds for the second and third lines. The best two results per line are italicized.

Resolution	HE	BPDFHE	LIME	NPEA	CACHE
481×321	*10.9*	*17.8*	176.0	6,730.2	55.0
$2,268 \times 1,512$	*.3*	*.4*	2.9	139.7	*.4*
$7,216 \times 5,412$	*1.5*	3.4	36.9	1,619.7	*1.6*

3.3. Efficiency

Table 2 compares the time required by different methods for processing inputs with different resolutions. In theory, CACHE should always be slower than HE because of its additional computation of dark-pass filtered gradients. LIME, NEPA, and CACHE all require operations in the spatial domain, which is usually computationally more expensive than operations in the intensity domain. Therefore, in most cases these methods are slower than HE and BPDFHE. When processing a 481×321 image, the processing time of CACHE is less than a third that of LIME and less than a hundredth that of NPEA. As discussed in Section 2.2, a multi-resolution scheme based on a canonical resolution provides a beneficial side effect in terms of efficiency. When processing a very large image of $7,216 \times 5,412$ pixels, the time required by CACHE is much the same as HE. In this case, it only took 1.6 seconds, which is less than a twentieth of the processing time needed by LIME.

4. CONCLUSION

In this study, we proposed a novel contrast enhancement method, called CACHE, which derives a desirable relationship between the potential visual importance of pixels and the contrast gain of intensities in an enhanced image. CACHE is based on a spatial distinctiveness prior and a dark prior: the former helps avoid both excessive contrast enhancement for trivial visual elements and the compression of meaningful image details; the latter supports the emphasis of the image details hidden in darker regions. Performance comparisons with state-of-the-art contrast enhancement methods consistently demonstrated the great effectiveness of CACHE from different perspectives. CACHE only required 1.6 seconds when processing a very large image of $7,216 \times 5,412$ pixels, which is much less than the time required by its competitors [16, 19]. In the future, we shall investigate the complementarity of CACHE with local contrast enhancement techniques based on the Retinex theory.

5. REFERENCES

[1] Tinku Acharya and Ajoy K. Ray, *Image Processing - Principles and Applications*, Wiley-Interscience, 2005. 1, 2, 3, 4

[2] Karel Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems IV*, Paul S. Heckbert, Ed., pp. 474–485. Academic Press Professional, Inc., 1994. 1

[3] Debdoot Sheet, Hrushikesh Garud, Amit Suveer, Manjunatha Mahadevappa, and Jyotirmoy Chatterjee, "Brightness preserving dynamic fuzzy histogram equalization," *IEEE Trans. Consumer Electronics*, vol. 56, no. 4, pp. 2475–2480, 2010. 1, 3, 4

[4] Turgay Çelik and Tardi Tjahjadi, "Automatic image equalization and contrast enhancement using Gaussian mixture modeling," *IEEE Trans. Image Processing*, vol. 21, no. 1, pp. 145–156, 2012. 1

[5] Tarik Arici, Salih Dikbas, and Yucel Altunbasak, "A histogram modification framework and its application for image contrast enhancement," *IEEE Trans. Image Processing*, vol. 18, no. 9, pp. 1921–1935, 2009. 1

[6] Chulwoo Lee, Chul Lee, and Chang-Su Kim, "Contrast enhancement based on layered difference representation of 2D histograms," *IEEE Trans. Image Processing*, vol. 22, no. 12, pp. 5372–5384, 2013. 1

[7] Raanan Fattal, Dani Lischinski, and Michael Werman, "Gradient domain high dynamic range compression," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 249–256, 2002. 1, 3

[8] Jayanta Mukherjee and Sanjit K. Mitra, "Enhancement of color images by scaling the DCT coefficients," *IEEE Trans. Image Processing*, vol. 17, no. 10, pp. 1783–1794, 2008. 1

[9] E. H. Land, "Lightness and Retinex theory," *J. Opt. Soc. Am.*, vol. 61, pp. 1–11, 1971. 1

[10] Daniel J. Jobson, Zia-ur Rahman, and Glenn A. Woodell, "Properties and performance of a center/surround Retinex," *IEEE Trans. Image Processing*, vol. 6, no. 3, pp. 451–462, 1997. 1

[11] Daniel J. Jobson, Zia-ur Rahman, and Glenn A. Woodell, "A multiscale Retinex for bridging the gap between color images and the human observation of scenes," *IEEE Trans. Image Processing*, vol. 6, no. 7, pp. 965–976, 1997. 1

[12] Giovanni Ramponi, "A cubic unsharp masking technique for contrast enhancement," *Signal Processing*, vol. 67, no. 2, pp. 211–222, 1998. 1

[13] Frédo Durand and Julie Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 257–266, 2002. 1

[14] Michael Elad, "Retinex by two bilateral filters," in *Scale-Space*, 2005, pp. 217–229. 1

[15] Zeev Farbman, Raanan Fattal, Dani Lischinski, and Richard Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," *ACM Trans. Graph.*, vol. 27, no. 3, 2008. 1

[16] Xiaojie Guo, Yu Li, and Haibin Ling, "LIME: Low-light image enhancement via illumination map estimation," *IEEE Trans. Image Processing*, 2016. 1, 3, 4

[17] Andrea Polesel, Giovanni Ramponi, and V. John Mathews, "Image enhancement via adaptive unsharp masking," *IEEE Trans. Image Processing*, vol. 9, no. 3, pp. 505–510, 2000. 1

[18] Guang Deng, "A generalized unsharp masking algorithm," *IEEE Trans. Image Processing*, vol. 20, no. 5, pp. 1249–1261, 2011. 1

[19] Shuhang Wang, Jin Zheng, Hai-Miao Hu, and Bo Li, "Naturalness preserved enhancement algorithm for non-uniform illumination images," *IEEE Trans. Image Processing*, vol. 22, no. 9, pp. 3538–3548, 2013. 1, 3, 4

[20] Nuala Brady and David J. Field, "Local contrast in natural images: Normalisation and coding efficiency," *Perception*, vol. 29, no. 9, pp. 1041–1055, 2000. 2

[21] Tania Pouli, Douglas W. Cunningham, and Erik Reinhard, "A survey of image statistics relevant to computer graphics," *Comput. Graph. Forum*, vol. 30, no. 6, pp. 1761–1788, 2011. 2

[22] Christophe Schlick, "Quantization techniques for visualization of high dynamic range pictures," in *Photorealistic Rendering Techniques*, Georgios Sakas, Stefan Müller, and Peter Shirley, Eds., pp. 7–20. Springer Berlin Heidelberg, 1995. 3

[23] `http://sipi.usc.edu/database/`. 3

[24] `http://imagecompression.info/`. 3

[25] Pablo Arbelaez, Michael Maire, Charless C. Fowlkes, and Jitendra Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, 2011. 3

[26] Claude E. Shannon, "A mathematical theory of communication," *Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3–55, 2001. 3

[27] Sos S. Agaian, Blair Silver, and Karen A. Panetta, "Transform coefficient histogram-based image enhancement algorithms using contrast entropy," *IEEE Trans. Image Processing*, vol. 16, no. 3, pp. 741–758, 2007. 3

[28] ZhiYu Chen, Besma R. Abidi, David L. Page, and Mongi A. Abidi, "Gray-level grouping (GLG): an automatic method for optimized image contrast enhancement-part I: the basic method," *IEEE Trans. Image Processing*, vol. 15, no. 8, pp. 2290–2302, 2006. 3

[29] Soong-Der Chen and Abd. Rahman Ramli, "Minimum mean brightness error bi-histogram equalization in contrast enhancement," *IEEE Trans. Consumer Electronics*, vol. 49, no. 4, pp. 1310–1319, 2003. 3