

On the evaluation of illumination compensation algorithms

Vassilios Vonikakis¹ · Rigas Kouskouridas² ·
Antonios Gasteratos³ 

Received: 3 December 2016 / Revised: 20 April 2017 / Accepted: 28 April 2017
© Springer Science+Business Media New York 2017

Abstract This paper presents a comparison framework for algorithms that can diminish the effects of illumination in images. Its main objective is to reveal the positive and negative characteristics of such algorithms, allowing researchers to select the most appropriate one for their target application. The proposed framework utilizes artificial illumination degradations on real images, which are then processed by the tested algorithms. The results are evaluated by an ensemble of performance metrics, highlighting the various characteristics of the algorithms across a range of different image attributes. The proposed framework represents a useful tool for the selection of illumination compensation algorithms due to a) its quantitative nature, b) its multifaceted analysis and c) its easy reproducibility. The validity of the proposed framework is tested by applying it to the enhancement results of four illumination compensation algorithms, which are used as preprocessing in two classic computer vision applications. The improvements brought about by the algorithms are in accordance with the predictions of the proposed framework.

Keywords Illumination compensation · Image enhancement · Exposure correction · Comparison framework

✉ Vassilios Vonikakis
bbonik@adsc.com.sg

Rigas Kouskouridas
rkouskou@gmail.com

Antonios Gasteratos
agaster@pme.duth.gr

¹ Advanced Digital Sciences Center, 1 Fusionopolis Way, #08-10 Connexis North Tower, Singapore 138632, Singapore

² WIREWAX, 2-4 Whitfield Street, London W1T 2RB, UK

³ Department of Production, Management Engineering, Democritus University of Thrace, Vassilis Sofias 12, GR-67100 Xanthi, Greece

1 Introduction

Illumination is among the most important factors affecting the performance of the majority of computer vision and image processing algorithms. Non-diffusive or directional illumination may shade or highlight scene objects, resulting in High Dynamic Range (HDR) imaging conditions, which affect negatively the quality of captured images. Under such conditions, the dynamic range of the scene exceeds the dynamic range of the camera and, as a result, no particular exposure can sufficiently capture the radiances of the scene. Longer exposure times will effectively capture details in the shadows, but will lead to overexposure of the brighter image regions. On the other hand, short exposure times will adequately capture the bright image regions, but will result in underexposed shadowed image regions. Both underexposed and overexposed regions pose a significant challenge to computer vision algorithms, since any image characteristic, such as edges, colors or local features, become a lot harder to detect. This is clearly evident in Fig. 1a, which depicts two different exposures of an artificial HDR scene, along with their detected Canny edges. The correct extraction of edges, in both the left and the right part of the scene, becomes impossible for any of the exposures.

To overcome the above illumination problems, four different approaches are available. The straight-forward approach is to use HDR capturing equipment, instead of common, Low Dynamic Range (LDR), cameras. In such cases, the camera sensor has adequate dynamic range to capture the radiance of the scene, providing additional visual details in shadowed and very bright image regions. The cost, however, of HDR capturing and display equipment is still prohibiting, as compared to LDR systems. A second approach to tackle this problem is to combine multiple images of the same scene, captured by an LDR camera, using different exposures [1]. Specialized algorithms are designed to combine the sequence of LDR images into one HDR image containing the radiance of the scene [3, 11]. This technique, however, requires complicated image registration approaches to overcome the ghosting artifacts of moving objects. The simplest approach to the problem, both in terms of cost and computational needs, is the use of controlled illumination. Many computer vision systems are designed to perform in environments with diffusive illumination, in order to minimize HDR imaging conditions. However, this approach can be employed mainly in indoor environments, where the illumination can be controlled. Besides, imaging systems that have to operate outdoors will, inevitably, come across under/overexposed regions, diminishing the quality of the captured images. If HDR equipment is not employed in a vision system

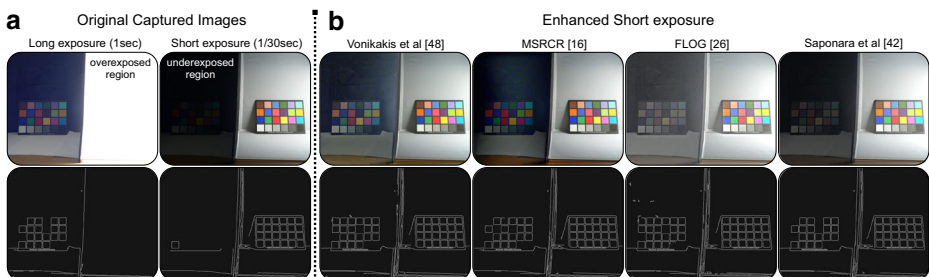


Fig. 1 **a** Effects of uneven illumination on the extraction of Canny edges (*bottom row*). **b** Improvements introduced by various illumination compensation algorithms

operating outdoors, the only method left to deal with the problem of non-uniform illumination, is the use of a preprocessing illumination compensation technique, which can minimize the effects of shadows or bright regions on the captured images.

Several algorithms used for illumination compensation have been presented in the literature, coming from many different research fields. Many of them utilize properties of the Human Visual System (HVS) and attempt to compute the *appearance* of the scene, as perceived by human observers. The Retinex family of algorithms [4, 10, 16–18, 25, 26, 38, 46] represents a significant part in this category, along with other HVS-inspired ones, such as ACE [42], RACE [39] and Vonikakis et al. [49]. Although the primary objective of the aforementioned algorithms is to compute the appearance of the scene, their illumination compensation characteristics have established them as a preprocessing step in many other classical imaging applications, such as image retrieval [7], stereo [54], general image enhancement [40], shadow removal [9], real-time video enhancement [15, 43] and the correspondence problem [53].

Many existing image enhancement algorithms exhibit illumination compensation characteristics, but have no relevant connection with the HVS. Some algorithms in this category include the Fused Logarithmic Transform (FLOG) [27], the Schlick algorithm [44], wavelet-based algorithms [6], image enhancement in the DCT domain [22], histogram-based transformations [19, 24, 57], illumination invariant color spaces [8], genetic algorithms [41] and, intrinsic images [32, 45]. Furthermore, apart from static images, illumination compensation techniques have also been applied in video sequences [12, 23, 29, 30].

All the aforementioned methods have different characteristics or even objectives. However, they can all diminish the effects of uneven illumination, by enhancing the underexposed and, some of them, the overexposed image regions. This makes them ideal for general image enhancement and preprocessing for computer vision applications. Figure 1b depicts the enhancement results of various illumination compensation algorithms, applied to the Short Exposure image of Fig. 1a, along with their extracted Canny edges. All the enhancement results exhibit superior edge information, as compared to the extracted edges of the original Short Exposure image. The diversity of illumination compensation algorithms, along with their potential to improve imaging applications, give rise to a significant problem; how to choose the most appropriate one for a particular imaging or computer vision application.

The comparison methods, which have been used so far for the evaluation of illumination compensation algorithms, can be grouped into four major categories: visual evaluation, psychophysical experiments, image quality metrics (IQM) and improvement of specific vision tasks. Visual evaluation represents the simplest of all these forms of comparison, since the original and the enhanced images are placed side by side and the human observer evaluates the quality of the results. This approach is, by its nature, subjective and difficult to draw accurate conclusions about the individual characteristics of the algorithms. Psychophysical experiments are studies in which human observers are asked to quantitatively evaluate specific characteristics of the enhanced images, such as contrast, brightness, naturalness, colorfulness etc., or to rank the results of a number of enhancement algorithms [5, 21, 28, 56]. Unlike the visual evaluation, this approach delivers quantitative results, based on the evaluation statistics of many observers, although it still remains subjective to a certain degree. This is so because the evaluation results are not immune to factors such as the educational background, the age or the gender of the observers. Furthermore, this comparison technique is very hard to reproduce, since it requires the coordination of many observers, calibration of viewing conditions and statistical analysis of the results. IQM are measures

used for the evaluation of perceived characteristics of imaging systems or of image processing techniques [2]. Several IQM have been introduced in recent years for various image characteristics, such as contrast [48], colorfulness [14], naturalness [55] etc. Many of them are based on psychophysical studies [14, 48, 55], while others on purely mathematical analysis [2] and natural image statistics [36]. Generally, IQMs either calculate particular image characteristics or predict the distortion introduced into the image by processing algorithms. This, however, is not adequate for evaluating illumination compensation algorithms, with the specific purpose of preprocessing in computer vision applications. On the other hand, this is the objective of the fourth category of evaluation, which attempts to assess the improvements that the algorithms bring about to specific vision tasks, such as face recognition [47] or stereo matching [37]. This approach usually relies on benchmark datasets and performance metrics, and thus, quantitatively estimates how much the preprocessing algorithms improve the task been solved. Undoubtedly, this approach delivers by far the most objective results among all the existing comparison approaches. However, it also exhibits some limitations. First, the fact that the tested algorithms are used in very specific high-level computer vision tasks makes it difficult to draw general conclusions regarding their preprocessing performance to other high-level tasks. For example, an illumination compensation algorithm, exhibiting good results for face recognition, does not necessarily mean that it will exhibit equally good results for stereo matching, simply because the two tasks may rely on different low level attributes. Second, this approach essentially ranks the illumination compensation algorithms according to a single characteristic (i.e. the improvements brought about to the specific computer vision task), without actually providing detailed information regarding their behavior. More specifically, no information is provided as to how an algorithm's performance changes with different degrees of shadows or highlights, or to what extent an algorithm affects the correctly exposed image regions, in the attempt to render the under/overexposed ones. This kind of information may provide important clues to researchers who want to choose an illumination compensation algorithm. According to the authors' knowledge, no existing evaluation scheme has been designed in order to highlight the overall behavior of such algorithms.

This paper attempts to resolve this issue. It introduces a new comparison framework for illumination compensation algorithms. The proposed approach is an extension of the work presented in [51]. It utilizes similar artificial illumination degradations, and extends their quantitative interpretation by introducing an additional step of statistical meta-analysis. This meta-analysis includes an ensemble of evaluation statistics, *quantifying* important characteristics of illumination compensation algorithms, such as enhancement of various degrees of under/overexposure, preservation of correctly-exposed regions and consistency of performance. In order to test the validity of the estimations of the proposed framework, four illumination compensation algorithms are used as preprocessing in two classic computer vision tasks. The improvements that the illumination compensation algorithms introduce to the final results of these computer vision methods are found to be in agreement with the conclusions derived by the proposed framework. Consequently, the proposed framework can be used as a tool, assisting researchers on the selection the most appropriate illumination compensation algorithm for their task.

The paper is organized as follows: Section 2 presents a detailed description of the proposed comparison framework. Section 3 gives an extensive example for the evaluation of four illumination compensation algorithms, using the proposed framework. Section 4 provides the necessary tests, for evaluating the validity of the proposed framework, using real scenes under uniform and non-uniform illumination. Finally, concluding remarks and discussion are presented in Section 5.

2 Comparison framework

The objectives of the proposed framework are summarized as follows:

1. To provide quantitative measures for the evaluation of illumination compensation algorithms, without the need of human observers.
2. To be easily reproducible.
3. To include the most important types of illumination degradations.
4. To highlight the positive and negative characteristics of the algorithms, which provide strong indications about its preprocessing potential for computer vision applications.

In order to satisfy the first objective, the framework excludes psychophysics from the evaluation of the enhanced images. The second objective is satisfied by using computer generated test images, which can be easily used as benchmarks by other researchers. The framework is intended to be used for algorithms, which compensate for illumination degradations. In order to meet the third objective, a set of computer-generated test images is compiled, comprising the most common illumination degradations that can be encountered in non-controlled environments, i.e. uniform/non-uniform illumination and under/overexposure. Finally, a statistical analysis for 7 algorithm characteristics is performed, concerning their enhancement performance on the set of test images, in order to satisfy the fourth objective.

2.1 Constructing the test images

All the comparison tests, performed by the framework, are carried out using a set of test images, generated by applying artificial degradations on well-exposed real images. The real images, which are also used as Ground Truth (GT), can depict any type of scene, as long as the illumination is *uniform* and the image is *correctly-exposed*. For demonstration purposes, we use the *Lena* image as GT, as it features an overall good exposure without any under/overexposed regions. In practice, however, for a thorough comparison between illumination compensation algorithms, more than one GT image should be used. Figure 2 depicts the flow chart of the proposed comparison framework.

The artificial degradations applied to the GT image, are generated by approximating two kinds of illumination (uniform and non-uniform), along with two kinds of degradations (underexposed and overexposed regions). In the case of uniform illumination, the degradations are applied to the whole GT image, whereas, for non-uniform illumination, the

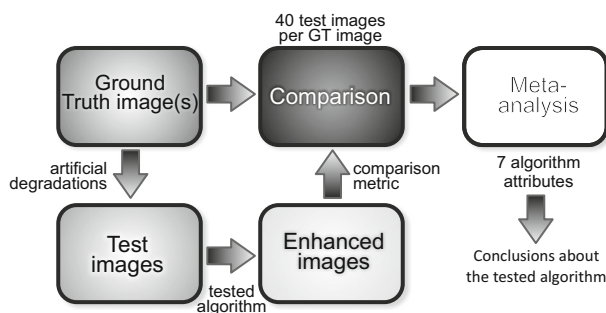


Fig. 2 Flowchart of the proposed comparison framework

degradations are applied only to a specific part of it, leaving the other part *intact*. Underexposure degradations are simulated using the following equation, which is based on the multiplicative relation between reflectance and illumination.

$$UE_{ij} = (1 - IL_{ij}) \times GT_{ij} \quad (1)$$

where UE is the underexposed-degraded test image, IL is a function that determines the type of illumination at each spatial position, and (i, j) are the pixel coordinates.

Similarly to the underexposed regions, overexposure degradations are simulated using the following equation.

$$OE_{ij} = B - [(1 - IL_{ij}) \times (B - GT_{ij})] \quad (2)$$

where OE is the overexposed-degraded test image, B is the maximum possible value of the GT image (usually 255), IL is a function that determines the type of illumination and (i, j) are the pixel coordinates.

For both underexposed and overexposed-degraded test images, we have identified three types of illumination, as indicated by the following equations.

$$IL_{ij}^{const} = IL_{max} \quad (3)$$

$$IL_{ij}^{step} = u\left(j - \frac{imx}{2}\right) \cdot IL_{max} \quad (4)$$

$$IL_{ij}^{grad} = \frac{j}{imx} \cdot IL_{max} \quad (5)$$

where IL^{const} corresponds to a uniform (constant) illumination, IL^{step} corresponds to a sharp illumination transition (step) and IL^{grad} is an illumination gradient. IL_{max} is the strength of illumination, with $IL_{max} \in [0, 1]$, $u(\cdot)$ the unitary step function and imx the width of the GT image. When $IL_{max} = 0$, underexposed or overexposed image regions disappear and the degraded test image is equal to the GT image. When $IL_{max} = 1$, the degradation strength is maximum, resulting to the complete loss of any visual information. For all the intermediate values of IL_{max} , the strength of the degradation varies linearly between these two extremes.

In our implementation we focused on the uniform illumination of (3) and the non-uniform (step) illumination of (4). We preferred the step illumination to the gradient of (5) since the former poses a grater challenge for illumination compensation algorithms, by triggering the appearance of possible *halo artifacts* or *gradient reversals* in the region of the sharp illumination transition. Consequently, the use of (3) may expose important limitations in the performance of illumination compensation algorithms. In order to analyze various degrees of illumination, test images with 10 different under/overexposure strengths are generated, for both uniform and non-uniform illumination. This essentially means that $IL_{max} \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95\}$. Consequently, 40 different test images ($T_1 - T_{40}$) are generated from each GT image, a subset of which is depicted in Fig. 3.

- $T_1 - T_{10}$: Uniform underexposure (IL^{const}).
- $T_{11} - T_{20}$: Uniform overexposure (IL^{const}).
- $T_{21} - T_{30}$: Non-uniform underexposure (IL^{step}).
- $T_{31} - T_{40}$: Non-uniform overexposure (IL^{step}).

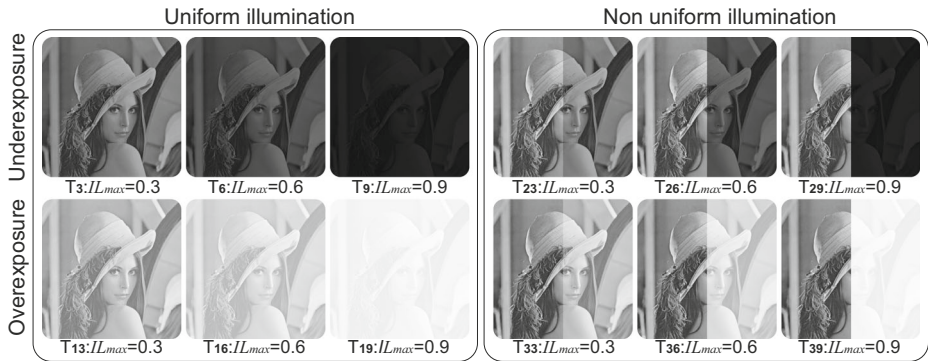


Fig. 3 Selected examples from the 40 test images, generated from the Lena image

2.2 Performance metric

The enhanced 40 test images are tested one by one against their corresponding GT image, in the way the flow chart of Fig. 2 depicts. A ‘theoretically perfect’ illumination compensation algorithm, would output an image identical to the GT image. The departure of the algorithm output from the GT image is measured using normalized Root Mean Square Error, which is the performance metric upon which, all the evaluations of the proposed framework are based. The formula of this metric is given in the following equation:

$$m_k^s = 1 - \frac{1}{B} \sqrt{\frac{\sum_{i \in S} \sum_{j \in S} (GT_{ij} - E(T_k)_{ij})^2}{N_S}} \quad (6)$$

$\forall (i, j) \in S = \{S_D, S_{ND}, S_W\}$. m_k^S is the performance metric applied to the S region of test image T_k , taking values in the interval $[0, 1]$. N_S is the number of pixels belonging to region S , B is the maximum possible value of the GT image (usually 255), $E(T_k)$ is the test image T_k , enhanced by the evaluated algorithm and (i, j) the pixel coordinates. Set S comprises three different image regions, on which the performance metric can be applied: S_D , which is the *degraded* region of the image, S_{ND} , which is the *non-degraded* region of the image and S_W , which is the *whole* size of the image. According to (6), the ‘theoretically perfect’ enhancement algorithm would exhibit a value of $m_k^S = 1$. Any departure from this, will lead to a lower value.

Depending on the image region on which the performance metric is applied, different insights can eventually be drawn about the characteristics of an illumination compensation algorithm. These insights can be summarized in the following 6 points:

- Improvement of uniformly underexposed regions (T_1 - T_{10} , S_W).
- Improvement of uniformly overexposed regions (T_{11} - T_{20} , S_W).
- Improvement of non-uniformly underexposed regions (T_{21} - T_{30} , S_D).
- Improvement of non-uniformly overexposed regions (T_{31} - T_{40} , S_D).
- Preservation of intact regions for underexposure (T_{21} - T_{30} , S_{ND}).
- Preservation of intact regions for overexposure (T_{31} - T_{40} , S_{ND}).

The last two cases are very important since a particular algorithm may perform very well in enhancing, for example, the underexposed regions of an image, yet, it may also affect

negatively the correctly exposed ones. The latter is an unwanted side-effect, as it lowers the quality of output images and passes undetected when the metric is applied globally, to the whole image size (S_W). Algorithms which can enhance under/overexposed areas, without affecting the correctly exposed ones, will exhibit better performance as preprocessing for other applications. To our knowledge, no other comparison approach has taken into consideration this attribute.

2.3 Meta-analysis

A statistical meta-analysis is applied on the performance results of the previous section. The main objective of this analysis is to bring out the positive and negative characteristics of the evaluated algorithms. Seven different algorithm characteristics are analyzed at this stage. These algorithm characteristics are presented in Table 1, along with their abbreviations and their mathematical formulae.

“Improvement of Uniform Underexposure” (IUU) conveys the capacity of an algorithm to enhance various degrees of *globally underexposed* images. It is essentially the average of an algorithm’s performance in all the test images with different degrees of uniform underexposure (T_1 - T_{10}). IUU assumes values in the interval [0, 1], with the ‘theoretically perfect’ algorithm having a value of 1.

“Improvement of Uniform Overexposure” (IUO) conveys the capacity of an algorithm to enhance various degrees of *globally overexposed* images. It is essentially the average performance of an algorithm across all test images with uniform overexposure (T_{11} - T_{20}). IUO assumes values in the interval [0, 1], with the ‘theoretically perfect’ algorithm having a value of 1.

“Constancy of performance in Uniform Illumination” (CUI) conveys how *consistent* the performance of an algorithm is, when enhancing various degrees of uniform illumination degradations. It is based on the standard deviations of the m_k^S metrics for test images T_1 - T_{20} , with a range of possible values in the interval [0, 1]. An algorithm that exhibits exactly the same enhancement performance for all the degradation levels of uniform under/overexposure, will have a CUI of 1. High CUI values indicate that an algorithm gives very predictable results under different degrees of uniform illumination, exhibiting an enhancement performance very close to the values of IUU and IUO. On the other hand,

Table 1 Metrics employed by the proposed comparison framework

	Algorithm characteristic	Abbreviation & formula
1	Improvement of uniform underexposure	$IUU = \text{mean}[m_k^{S_W}]_{k=1}^{10}$
2	Improvement of uniform overexposure	$IUO = \text{mean}[m_k^{S_W}]_{k=11}^{20}$
3	Constancy of performance in uniform illumination	$CUI = 1 - \text{std}[m_k^{S_W}]_{k=1}^{10} - \text{std}[m_k^{S_W}]_{k=11}^{20}$
4	Improvement of non-uniform underexposure	$INUU = \text{mean}[m_k^{S_D}]_{k=21}^{30}$
5	Improvement of non-uniform overexposure	$INUO = \text{mean}[m_k^{S_D}]_{k=31}^{40}$
6	Preservation of well-exposedness in non-uniform illumination	$PNU = \text{mean}[m_k^{S_{ND}}]_{k=21}^{40}$
7	Constancy of performance in non-uniform illumination	$CNUI = 1 - \text{std}[m_k^{S_D}]_{k=21}^{30} - \text{std}[m_k^{S_D}]_{k=31}^{40}$

low CUI values indicate *large variability* on the enhancement results of an algorithm, under different degrees of uniform illumination.

“Improvement of Non-Uniform Underexposure” (INUU) conveys the capacity of an algorithm to enhance various degrees of *underexposed regions*. It is essentially the average of the algorithm performance, when considering the enhancement on the degraded region (S_D) of test images T_{21} - T_{30} . INUU also assumes values in the interval $[0, 1]$, with 1 indicating the ‘theoretically perfect’ algorithm performance.

“Improvement only of Non-Uniform Overexposure” (INUO) conveys the capacity of an algorithm to enhance various degrees of *overexposed regions*. It is essentially the average of the algorithm performance when considering the enhancement on the degraded region (S_D) of test images T_{31} - T_{40} . The INUO acquires values in the interval $[0, 1]$, with 1 indicating the ‘theoretically perfect’ algorithm performance.

“Preservation of well-exposedness, in Non-Uniform illumination” (PNU) conveys the capacity of an algorithm to leave *intact* the well-exposed image regions, while enhancing various degradations in non-uniform illumination. It is essentially the average of the algorithm performance when applied to the non-degraded region (S_{ND}) of test images T_{21} - T_{40} . PNU takes values in the interval $[0, 1]$. An algorithm that has no effect on the well-exposed image regions, will exhibit a PNU value of 1.

“Constancy of performance in Non-Uniform Illumination” (CNUI) conveys how *consistent* the performance of an algorithm is, under various degrees of non-uniform illumination degradations. Similarly to CUI, it is based on the standard deviations of the m_k^S metrics, for test images T_{21} - T_{40} . It also assumes values in the interval $[0, 1]$. An algorithm that exhibits exactly the same performance in all non-uniform degradations, will have a CNUI value of 1. High CNUI values indicate an algorithm with very *predictable* results, very close to the values of INUU and INUO. On the other hand, low CNUI values indicate large performance variability, under non-uniform illumination.

All the above parameters can provide useful insights about the overall behavior of an illumination compensation algorithm. For example, if an algorithm exhibits low constancy measures (CUI, CNUI), there is strong indication that any improvements introduced will span a great range and will be rather unpredictable. Similarly, if an algorithm exhibits a low preservation characteristic (PNU), there is a strong indication that it will negatively affect the correctly-exposed image regions. This in effect could negate any improvement introduced by the illumination compensation algorithm.

3 Applying the framework

In order to demonstrate the proposed comparison framework, four different illumination compensation algorithms are evaluated. The algorithms used in this comparative study, along with other characteristics, are depicted in Table 2. Vassilios Vonikakis’ (VV) algorithm [49] is a center-surround image enhancement algorithm, based on the Human Visual System. Multi-Scale Retinex with Color Restoration (MSRCR) [16] and Saponara Retinex (SR) [43] are two different versions of the Retinex algorithm. MSRCR employs a multi-scale approach for the calculation of the surround, while SR utilizes a diffusion-based filter. Finally, the Fused Logarithmic Transform (FLOG) is an improvement of the classic logarithmic mapping, employing also a multi-scale pyramid [27].

Figure 4 depicts the intermediate results of the proposed framework. The six graphs of Fig. 4 show how the performance of algorithms vary across different degrees of degradations. Although some conclusions can be drawn solely by observing these graphs, it is

Table 2 Algorithms used in the comparison

	Algorithm	Ref.	Implementation	Parameters
1	Vonikakis et al. (VV)	[49]	Orasis software (ver.1.0) ^a	Default
2	Saponara Retinex (SR)	[43]	Manual implementation	b=4, g=0.5, a=0.97, K=0.5, d=0.00001
3	Multi-Scale Retinex with Color Restoration (MSRCR)	[16]	PhotoFlair software ^b	Default
4	Fused logarithmic transform (FLOG)	[27]	Manual implementation	p=1, q=1, g1=1, g2=2.5

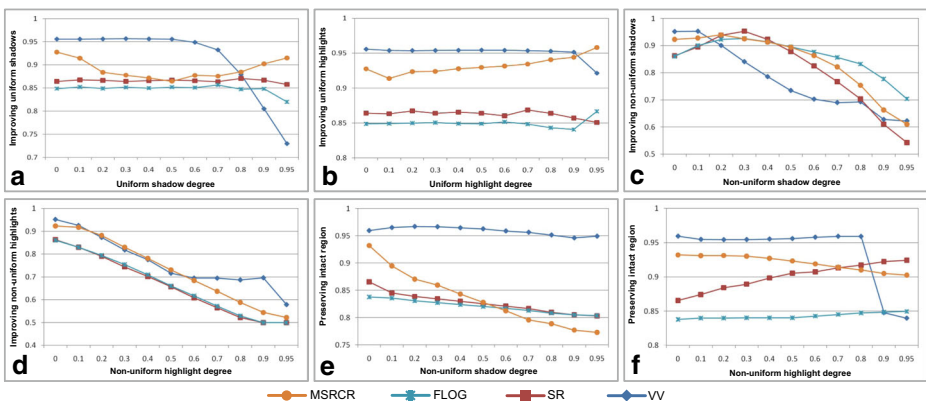
^a<https://sites.google.com/site/vonikakis/software-code>

^b<http://www.trueview.com>

difficult to extract accurate conclusions regarding the overall behavior of an algorithm. The results of the statistical meta-analysis (Fig. 5) offer such *quantitative overview*.

For uniform illumination, VV is ranked first in improving uniform underexposure, closely followed by MSRCR. These two algorithms exhibit also approximately the same performance for improving uniform overexposure. An important feature, however, is the fact that VV exhibits the lowest CUI measure, indicating a more *fluctuating* performance, for various degrees of degradations. The remaining two algorithms perform almost identically in these categories. Regarding the constancy in performance, SR exhibits very good predictability in its results, followed closely by FLOG.

For non-uniform illumination, the best underexposure improvement is exhibited by FLOG, followed closely by MSRCR. Regarding the improvement of non-uniform overexposed regions, VV exhibits the best performance, with MSRCR following next, while SR and FLOG exhibit the lowest performance in this category. FLOG and VV are seen to exhibit the best constancy in performance, while the other two algorithms (SR and MSRCR) are ranked low, exhibiting greater fluctuations for different degrees of degradations. Finally, regarding the preservation of correctly-exposed image regions, VV outperforms by far all the other algorithms. MSRCR and SR show a similar performance, while FLOG is the algorithm likely to have the most detrimental effect on correctly exposed regions.

**Fig. 4** Intermediate results of the proposed framework

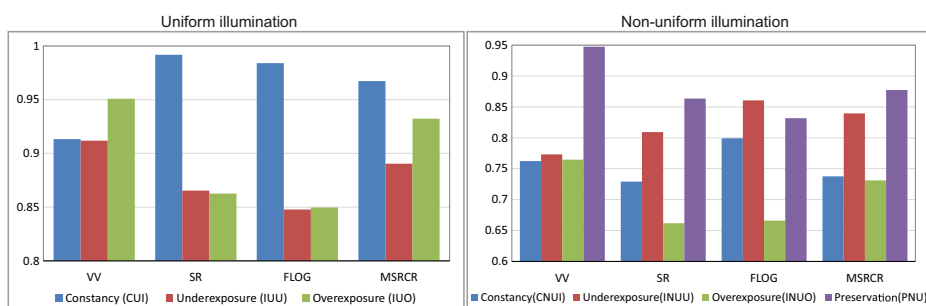


Fig. 5 Final results of the proposed comparison framework for uniform and non-uniform illumination

3.1 Detailed performance insights

Using the evaluation results of Fig. 5, important conclusions can be derived about the performance of the four algorithms.

The most prominent characteristic of VV is that it targets enhancement specifically to the degraded image regions, affecting minimally the correctly exposed ones. This might be important for both aesthetic reasons, as well as in many computer vision applications, because it ensures the preservation of visual information in the non-degraded regions. However, such a good characteristic comes at a cost; the enhancement of non-uniform underexposed regions, compared to other algorithms, is only modest. Apart from that, VV is good in enhancing uniformly underexposed or overexposed images with moderate degradation strength, while exhibiting fairly predictable results. As a conclusion, the VV algorithm exhibits moderate illumination compensation characteristics, with its strongest point being the overall appearance of the image, since it does not affect the correctly-exposed regions. When it comes to computer vision applications, VV should be given preference only when the strength of the illumination degradations is moderate.

MSRCR is not ranked first in any evaluation category; although, it consistently exhibits high performance. It is generally very good in enhancing any kind of underexposure, both uniform and non-uniform. Regarding overexposure and overall predictability of results, MSRCR should only be given preference when the lighting conditions are uniform, since it exhibits moderate only performance in non-uniform illumination. An important feature is that the preservation of the correctly exposed image regions is satisfactory. As a conclusion, MSRCR is an algorithm which can reliably be used for illumination compensation. It could be successfully used both for aesthetic correction of images, as well as preprocessing for other computer vision algorithms.

Similarly to MSRCR, SR is not ranked first in any evaluation category exhibiting, in general, a rather moderate performance. In the case of uniform illumination, SR could be a good choice of enhancement algorithm, due to the high predictability of its results. However, when illumination is non-uniform, the consistency in its performance decreases significantly. In such illumination conditions, it exhibits good enhancement in underexposed regions, very low enhancement of overexposed ones, while preserving well the correctly-exposed image regions. Similarly to MSRCR, SR exhibits a potential for both aesthetic and preprocessing use.

FLOG is an algorithm which exhibits a superior performance in enhancing underexposure in non-uniform illumination conditions. However, this attribute comes at a cost since, at the same time, it compromises the preservation of the correctly-exposed regions.

In this sense, FLOG exhibits an *opposite* performance compared to VV, which preserves almost perfectly the correctly-exposed regions, although at the expense of the underexposed ones. Additionally, FLOG is not to be recommended for enhancing overexposure under non-uniform illumination conditions. On the contrary, it exhibits moderate enhancement performance for uniform illumination, with highly predictable results. For these reasons, FLOG should be given preference mostly in preprocessing, rather than in the aesthetic enhancement of images.

A general conclusion that can be drawn from the previous analysis is that, there is a clear trade-off between enhancement improvements and preservation of correctly-exposed image regions. Large enhancement improvements seems to come at the expense of deteriorating image regions that do not require any enhancement. This, perhaps, highlights the need for more *selective* enhancement algorithms which can introduce stronger enhancements only in image regions that require it.

4 Evaluating the framework

In order to evaluate the usefulness of the proposed framework, a study is conducted utilizing *non-synthetic* images, captured under real-life conditions. The objective is to use illumination compensation algorithms as *preprocessing* on two classic low-level computer vision techniques. The improvements introduced by the evaluated algorithms are then compared to the theoretical predictions of the framework, in order to test whether there is alignment between the two.

4.1 SIFT

The first of the two selected low-level computer vision techniques is the Scale Invariant Feature Transform (SIFT) [31]. SIFT implements a detector and a descriptor, which are part of many existing higher-level computer vision applications. Although SIFT is scale and rotation invariant, illumination invariance is not its strongest point, compared to other similar methods [33–35]. In particular, under/overexposure may lower the magnitude of local image gradient, forcing many valid scale-space extrema to reside below the global detection threshold of the SIFT detector [50]. As a result, the number of extracted keypoints is indirectly affected by illumination. For this reason, detection and matching of SIFT keypoints is a reasonable application for testing the ability of illumination compensation algorithms to improve illumination invariance.

Twenty scenes containing various objects with different kinds of surfaces were used in this experiment. Two different images were captured per scene. The first image was captured under Strong Uniform Illumination (SUI), resulting in a well-exposed image free of any under/overexposed regions. SUI images serve as a GT, since they represent a case of almost ‘ideal’ illumination conditions. The second image of each scene was purposely captured under non-ideal illumination conditions, suffering from various degrees of underexposure. Two types of underexposure were used; one resulting from Low Uniform Illumination (LUI) and one resulting from Non-Uniform (directional) Illumination (NUI). LUI images are globally underexposed, whereas NUI images have local patches of underexposed regions and correctly-exposed ones. Figure 6 depicts two of these scenes, along with their enhanced versions and their extracted keypoints SIFT.

The main objective was to test how the use of an illumination compensation algorithm affects the performance of SIFT. This requires analyzing SIFT performance *before* the use

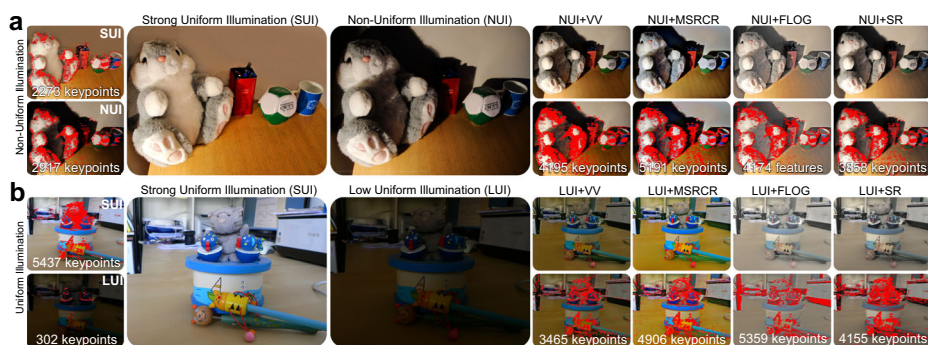


Fig. 6 Two of the real scenes used in the evaluation procedure, along with their enhanced versions and the number of detected keypoints by the SIFT detector

of the illumination compensation algorithm (baseline) and *after* it. Preferably, the use of such algorithms should increase the number of correct matches between a well-exposed image of a scene (SUI) and a version of it that is affected by illumination (LUI/NUI).

Figure 7 depicts this experimental procedure. SIFT is applied on the original image couples SUI-LUI and SUI-NUI, detecting features and counting the total number of matching points between them. This serves as the *baseline* for every scene, providing an indication about the performance of SIFT *without* the contribution of any illumination compensation algorithm. Next, the evaluated algorithm is applied on the images of the scene. SIFT is again applied to the enhanced images, in order to find the total number of matching points between the enhanced pairs (SUI-LUI and SUI-NUI). Finally, the ratio between the total matching points of the enhanced pairs over the total matching points of the baseline, gives an indication about how the enhancement illumination algorithm affects the efficiency of SIFT for the particular scene. If the use of the enhancement algorithm results in an increased number of correct matching points, the algorithm has a positive impact. In such cases, the comparison ratio is greater than unity. Conversely, if the enhancement algorithm results in a lower number of correct matching points, then the comparison ratio is lower than unity and the enhancement algorithm has a negative impact.

Figure 8a shows the average number of detected keypoints for each algorithm under *uniform* illumination, both for SUI and LUI images. As expected, the number of detected keypoints in SUI images is always higher than those in LUI images. This is due to the inadequate exposure of LUI images. The results of Fig. 8a clearly indicate that MSRCR outperforms all other algorithms in the number of detected keypoints in LUI images. Figure 8b shows the average descriptor efficiency; total matching keypoints between the enhanced SUI-LUI images, over the total matching keypoints between the original SUI-LUI images, under uniform illumination conditions.

In order to correctly evaluate these results, it is essential to make an assessment about the *strength* of the underexposure (uniform or non-uniform) that were induced by the illumination conditions. This is important, since the efficiency of every illumination compensation algorithm changes under different illumination levels (see Fig. 4). The assessment of the relative strength of underexposure, induced to the LUI and NUI images, is performed as follows.

$$ESD_j = 1 - \frac{\bar{L}_j}{\bar{L}_j^{SUI}} \quad (7)$$

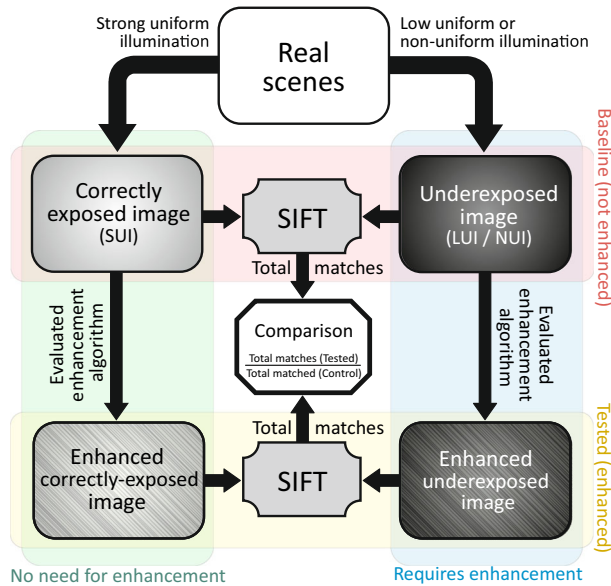


Fig. 7 Flowchart of the procedure for evaluating the impact of illumination compensation algorithms on SIFT

where ESD_j is the Estimated Shadow Degree for scene j , $\bar{L}_j^{SUI}$ is the mean luminance value of the SUI image and $\bar{L}_j$ is the mean luminance value of the illumination-affected image (either LUI or NUI) for scene j . Both $\bar{L}_j^{SUI}$ and $\bar{L}_j$ are average pixel values on the luminance channel of the image depicting scene j . Since LUI images are uniformly illuminated, $\bar{L}_j$ represents the mean luminance value of the *entire* image, whereas for NUI images, it represents the mean luminance value of the underexposed regions. The range of the ESD measure is in the interval $[0, 1]$, with values near 1 indicating very strong underexposure (either global or local).

The average ESD value of the LUI images is found to be 0.87. This means that the expected behavior of the enhancement algorithms, for this particular underexposure strength, should be predicted by the proposed framework in Fig. 4a, for an underexposure strength of 0.9. In this case, the ranking predicted by the proposed framework is: $MSRCR > SR > FLOG > VV$. Figure 8b indicates that the improvement introduced by the algorithms on SIFT, is similar to the ranking predicted by the framework. MSRCR exhibits the best ratio among the four algorithms (≈ 3.25). SR exhibits the second best improvement ratio, at approximately 3. Finally, FLOG and VV follow, with the latter exhibiting the lowest ratio of all, improving only by a factor of 2 the number of matching keypoints.

Similarly, Fig. 8c and d demonstrate the results of the experiment for non-uniform (directional) illumination conditions. Figure 8c shows the average number of detected keypoints for every algorithm. According to these results, the number of detected keypoints in SUI images is always *lower* than the respective one in NUI images. This seems counter-intuitive, since the SIFT detector usually extracts fewer keypoints in dark regions. However, the increased number of keypoints in these cases is a result of the *additional edges* that the local

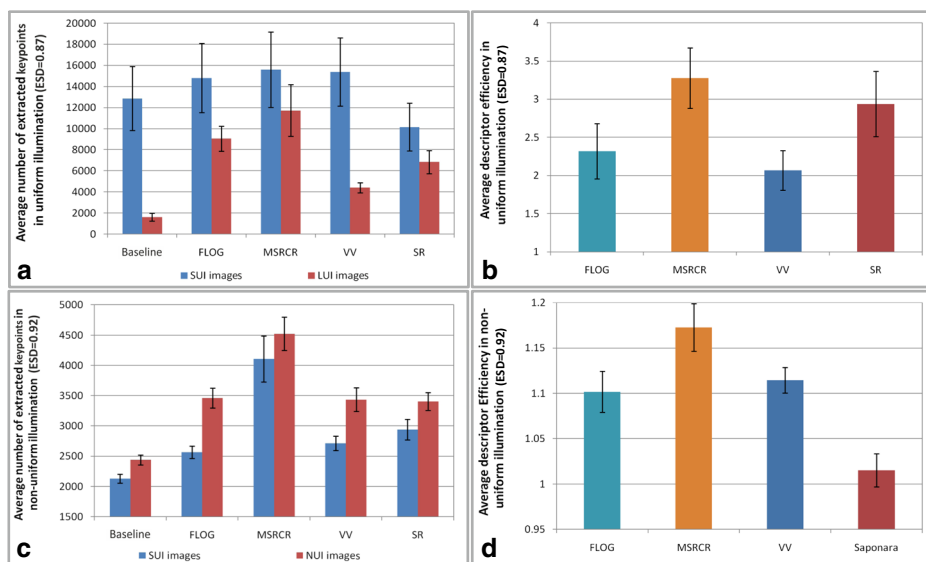


Fig. 8 **a** Average number of extracted keypoints in uniform illumination (LUI images). **b** Average descriptor efficiency in uniform illumination (LUI images). **c** Average number of extracted keypoints in non-uniform illumination (NUI images). **d** Average descriptor efficiency in non-uniform illumination (NUI images)

shadows introduce to the NUI images. This is clearly shown also in Fig. 6a, in which, the detected keypoints (marked by red) are concentrated around shadow edges.

As before, it is important to make an assessment for the underexposure degree of the NUI images. The average ESD value was found to be 0.92. Consequently, the predictions of the framework for an underexposure degree of 0.9 are used (Fig. 4a). The predicted ranking of the framework is: FLOG>MSRCR>VV>SR. Figure 8d indicates a similar, but not identical, ranking. That is due to a discrepancy regarding the prediction of FLOG algorithm. More specifically, the results of the real-scene experiment indicate a ranking of MSRCR>FLOG>VV>SR. Nevertheless, it is clear that the difference in the efficiencies of FLOG and MSRCR is rather small, not constituting a significant deviation from the predictions of the framework. Furthermore, the ranking of the other two algorithms (VV and SR), is in accordance with the pattern predicted by the proposed framework.

4.2 Real images and Harris' corner detector

To further evaluate the proposed framework, a similar experiment is performed using the Harris corner detector [13]. The images used in this test are the ones depicted in Fig. 9a, featuring two identical *ColorCheckers* on two very differently illuminated image regions. These images are part of an HDR workshop presented during the last CREATE (Colour Research for European Advanced Technology Employment) meeting.¹ This particular scene

¹<http://www.create.uwe.ac.uk/>.

is selected because ColorCheckers, apart from their obvious use in color measurements, can also be used as test-targets for corner detection, comprising 96 well-defined corners ($24 \text{ squares} \times 4 \text{ corners}$). The four tested algorithms are used for enhancing the six exposure of Fig 9a. Harris corner detector is then applied on the enhanced images and the correctly extracted corners are counted. “No Preprocessing” refers to the application of the Harris corner detector on the six exposures *without* the use of an illumination compensation algorithm, forming the *baseline* for this experiment. It should be mentioned that only the corners within the surface of the ColorChecker are counted. Furthermore, any corners extracted at the bottom of the ColorChecker, formed by the words: “GretagMacbeth™ColorChecker Color Rendition Chart”, are also omitted from the counting, due to the ambiguity of their position.

One obvious conclusion derived from Fig. 9d, is that all four illumination compensation algorithms improve the total number of correctly extracted corners in all test images. Figure 9b depicts the number of correctly extracted corners in the bright region of the scene. In the first exposure (ESD=0.75), the bright region is totally overexposed, resulting in a complete loss of any visual information. Inevitably, none of the test algorithms assist in

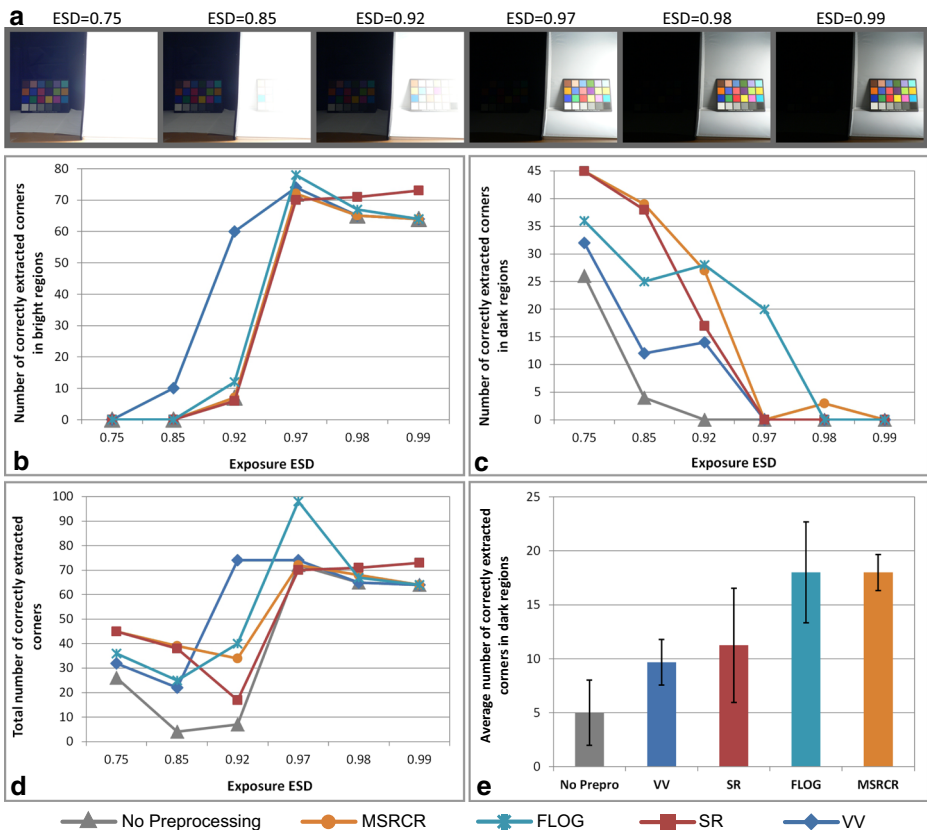


Fig. 9 Results of the Harris corner extraction experiment. **a** The 6 images used in the experiment, along with their ESD. **b** Number of correctly extracted corners in the bright region of the scene. **c** Number of correctly extracted corners in the dark region of the scene. **d** Total number of correctly extracted corners. **e** Average number of correctly extracted corners in the dark region of the scene, and their standard errors

corner extraction in this case. In the following two exposures ($ESD = 0.85$ and 0.92), the bright region is still overexposed. However, there is more visual information, compared to the first exposure. In these two cases, VV exhibits the best performance, improving the bright region of the scene and resulting in the highest number of correctly extracted corners. The rest of the algorithms exhibit similar or slightly better performance over the baseline. In the other 3 exposures ($ESD = 0.97, 0.98$ and 0.99) all the tested algorithms have approximately similar performance. These results are in agreement with the predictions of the proposed framework (Fig. 4d), depicting the predicted performance of algorithms for over-exposure. More specifically, for degrees $0.7\text{--}0.95$, the framework indicates that the best performance is exhibited by VV. This is exactly the case in Fig. 9b, for exposures 0.85 and 0.92 , which are within the range $0.7\text{--}0.95$. For an ESD of 0.95 and greater, the proposed framework correctly indicates that the performance of algorithms tends to be similar, something which is confirmed in the exposures $ESD = 0.97, 0.98$ and 0.99 of Fig. 9b.

Figure 9c depicts the number of correctly extracted corners in the dark region of the scene. According to this graph, MSRCR and SR exhibit initially the best performance. However, their performance gradually decreases, and for high $ESDs$, FLOG is ranked first. VV exhibits always the lowest performance. A similar trend is also indicated by the proposed framework, in Fig. 4c.

Finally, Fig. 9e depicts the average number of correctly extracted corners in the dark regions of the six exposures. Again, this graph is in agreement with the predicted INUU measure of the framework; the highest performance is exhibited by FLOG and MSRCR, followed by SR and last by VV. This indicates that the proposed framework can highlight the general performance characteristics of algorithms, predicting their performance tendency across various degrees of illumination.

5 Conclusions

A new comparison framework for the evaluation of illumination compensation algorithms is presented in this paper. The framework attempts to highlight the positive and negative characteristics of algorithms, which would ultimately provide a yardstick, assisting researchers on choosing the right algorithm for their application. The proposed framework does not involve human observers or any perceptual IQM; instead it utilizes computer generated synthetic images, with artificial illumination degradations of various degrees. The improvements introduced by any algorithm are measured by a set of metrics, quantifying performance for a range of characteristics. The proposed approach is used to examine the performance of four illumination compensation algorithms, providing indications about their positive and negative characteristics. The usefulness of the framework's predictions is evaluated by two experiments involving real scenes. These experiments demonstrate that the improvements introduced by the algorithms are in accordance with the predictions of the framework. Consequently, the proposed framework can be used as a benchmarking tool for evaluating illumination compensation algorithms, highlighting their suitability for preprocessing in other imaging/vision applications.

In the context of multimedia, the proposed approach can provide a valuable tool for highlighting the best potential solutions for the improvement of image or video-based applications. For example, it can assist with the selection of image enhancement algorithms, for the aesthetic improvement of images, in multimedia systems with personal photo-collections [52]. It can narrow down candidate illumination compensation algorithms for improving the performance of image retrieval systems [8, 20]. Finally, it may help

researchers with the selection of image enhancement algorithms for the improvement of the aesthetic quality of videos [29] or increasing the robustness of video-based computer vision tasks [23].

Acknowledgements This study is partially supported by the research grant for the Human-Centered Cyber-physical Systems Programme at the Advanced Digital Sciences Center from Singapore's Agency for Science, Technology and Research (A*STAR).

References

1. Aggarwal M, Ahuja N (2004) Split aperture imaging for high dynamic range. *Int J Comput Vis* 58(1): 7–17
2. Avciabas I, Sankur B, Sayood K (2002) Statistical evaluation of image quality measures. *J Electron Imaging* 11(2):206–223. doi:[10.1117/1.1455011](https://doi.org/10.1117/1.1455011)
3. Battiatto S, Castorina A, Mancuso M (2003) High dynamic range imaging for digital still camera: an overview. *J Electron Imaging* 12:459–469. doi:[10.1117/1.1580829](https://doi.org/10.1117/1.1580829)
4. Bertalmio M, Caselles V, Provenzi E (2009) Issues about retinex theory and contrast enhancement. *Int J Comput Vis* 83(1):101–119. doi:[10.1007/s11263-009-0221-5](https://doi.org/10.1007/s11263-009-0221-5)
5. Čadík M, Wimmer M, Neumann L, Artusi A (2006) Image attributes and quality for evaluation of tone mapping operators. In: *Proceedings of pacific graphics 2006* (14th pacific conference on computer graphics and applications), pp 35–44. National Taiwan University Press
6. Cao W, Che R, Ye D (2008) An illumination-independent edge detection and fuzzy enhancement algorithm based on wavelet transform for non-uniform weak illumination images. *Pattern Recogn Lett* 29(3):192–199
7. Ciocca G, Marini D, Rizzi A, Schettini R, Zuffi S (2003) Retinex preprocessing of uncalibrated images for color-based image retrieval. *J Electronic Imaging* 12(1):161–172
8. Dubey SR, Singh SK, Singh RK (2015) A multi-channel based illumination compensation mechanism for brightness invariant image retrieval. *Multimed Tools Appl* 74(24):11,223–11,253. doi:[10.1007/s11042-014-2226-5](https://doi.org/10.1007/s11042-014-2226-5)
9. Finlayson GD, Hordley SD, Drew MS (2002) Removing shadows from images using retinex. In: *The tenth color imaging conference: color science and engineering systems, technologies, applications, CIC 2002*, November 12, 2002, Scottsdale, Arizona, USA, pp. 73–79. IST - The Society for Imaging Science and Technology. <http://www.imaging.org/store/epub.cfm?abstrid=8402>
10. Funt B, Ciurea F, McCann J (2004) Retinex in matlab. *J Electron Imaging* 13(1):48–57. doi:[10.1117/1.1636761](https://doi.org/10.1117/1.1636761)
11. Goshtasby AA (2005) Fusion of multi-exposure images. *Image Vis Comput* 23(6):611–618
12. Han Y, Zhang Z (2014) An efficient estimation method for intensity factor of illumination changes. *Multimed Tools Appl* 72(3):2619–2632. doi:[10.1007/s11042-013-1521-x](https://doi.org/10.1007/s11042-013-1521-x)
13. Harris C, Stephens M (1988) A combined corner and edge detector. In: *Proceedings of fourth alvey vision conference*, pp 147–151
14. Hasler D, Suesstrunk SE (2003) Measuring colorfulness in natural images. In: Rogowitz BE, Pappas TN (eds) *Human vision and electronic imaging VIII*, vol 5007, pp 87–95. doi:[10.1117/12.477378](https://doi.org/10.1117/12.477378)
15. Iakovidou C, Vonikakis V, Andreadis I (2008) Fpga implementation of a real-time biologically inspired image enhancement algorithm. *J Real-Time Image Proc* 3(4):269–287
16. Jobson D, Rahman Zu, Woodell G (1997) A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Trans Image Process* 6(7):965–976. doi:[10.1109/83.597272](https://doi.org/10.1109/83.597272)
17. Jobson DJ, Rahman Z, Woodell GA (1997) Properties and performance of a center/surround retinex. *IEEE Trans Image Process* 6(3):451–462. doi:[10.1109/83.557356](https://doi.org/10.1109/83.557356)
18. Kimmel R, Elad M, Shaked D, Keshet I, Sobel I (2003) A variational framework for retinex. *Int J Comput Vis* 52(1):7–23. doi:[10.1023/A:1022314423998](https://doi.org/10.1023/A:1022314423998)
19. Kong TL, Isa NAM (2016) Enhancer-based contrast enhancement technique for non-uniform illumination and low-contrast images. *Multimed Tools Appl* 1–22. doi:[10.1007/s11042-016-3787-2](https://doi.org/10.1007/s11042-016-3787-2)
20. Konstantinidis K, Vonikakis V, Panitsidis G, Andreadis I (2011) A center-surround histogram for content-based image retrieval. *Pattern Anal Appl* 14(3):251–260. doi:[10.1007/s10044-011-0217-y](https://doi.org/10.1007/s10044-011-0217-y)

21. Kuang J, Yamaguchi H, Johnson GM, Fairchild MD (2004) Testing hdr image rendering algorithms. *Color and Imaging Conference* 2004(1):315–320
22. Kuo CM, Yang NC, Liu CS, Tseng PY, Chang CK (2016) An effective and flexible image enhancement algorithm in compressed domain. *Multimed Tools Appl* 75(2):1177–1200. doi:[10.1007/s11042-014-2363-x](https://doi.org/10.1007/s11042-014-2363-x)
23. Kushwaha AKS, Srivastava R (2016) Automatic moving object segmentation methods under varying illumination conditions for video data: comparative study, and an improved method. *Multimed Tools Appl* 75(23):16,209–16,264. doi:[10.1007/s11042-015-2927-4](https://doi.org/10.1007/s11042-015-2927-4)
24. Lai YR, Tsai PC, Yao CY, Ruan SJ (2015) Improved local histogram equalization with gradient-based weighting process for edge preservation. *Multimed Tools Appl* 1–29. doi:[10.1007/s11042-015-3147-7](https://doi.org/10.1007/s11042-015-3147-7)
25. LAND EH (1964) The retinex. *Am Sci* 52(2):247–264
26. Land EH (1986) An alternative technique for the computation of the designator in the retinex theory of color vision *Proc Natl Acad Sci USA*
27. Le HS, Li H (2008) Fused logarithmic transform for contrast enhancement. *Electron Lett* 44(1): 19–20
28. Ledda P, Chalmers A, Troscianko T, Seetzen H (2005) Evaluation of tone mapping operators using a high dynamic range display. In: *ACM SIGGRAPH 2005 Papers*, SIGGRAPH '05. ACM, New York, NY, USA, pp 640–648. doi:[10.1145/1186822.1073242](https://doi.org/10.1145/1186822.1073242)
29. Lin GS, Ji XW (2016) Video quality enhancement based on visual attention model and multi-level exposure correction. *Multimed Tools Appl* 75(16):9903–9925. doi:[10.1007/s11042-015-2777-0](https://doi.org/10.1007/s11042-015-2777-0)
30. Lin TL, Thakur US, Chou CC, Chen SL (2016) Hole filling using multiple frames and iterative texture synthesis with illumination compensation. *Multimed Tools Appl* 75(4):1899–1921. doi:[10.1007/s11042-014-2379-2](https://doi.org/10.1007/s11042-014-2379-2)
31. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 60(2):91–110. doi:[10.1023/B:VISI.0000029664.99615.94](https://doi.org/10.1023/B:VISI.0000029664.99615.94)
32. Matsushita Y, Nishino K, Ikeuchi K, Sakauchi M (2004) Illumination normalization with time-dependent intrinsic images for video surveillance. *IEEE Trans Pattern Anal Mach Intell* 26(10):1336–1347. doi:[10.1109/TPAMI.2004.86](https://doi.org/10.1109/TPAMI.2004.86)
33. Mikolajczyk K, Schmid C (2004) Scale & affine invariant interest point detectors. *Int J Comput Vis* 60(1):63–86. doi:[10.1023/B:VISI.0000027790.02288.f2](https://doi.org/10.1023/B:VISI.0000027790.02288.f2)
34. Mikolajczyk K, Schmid C (2005) A performance evaluation of local descriptors. *IEEE Trans Pattern Anal Mach Intell* 27(10):1615–1630. doi:[10.1109/TPAMI.2005.188](https://doi.org/10.1109/TPAMI.2005.188)
35. Mikolajczyk K, Tuytelaars T, Schmid C, Zisserman A, Matas J, Schaffalitzky F, Kadir T, Gool LV (2005) A comparison of affine region detectors. *Int J Comput Vis* 65(1):43–72. doi:[10.1007/s11263-005-3848-x](https://doi.org/10.1007/s11263-005-3848-x)
36. Moorthy AK, Bovik AC (2011) Blind image quality assessment: from natural scene statistics to perceptual quality. *IEEE Trans Image Process* 20(12):3350–3364. doi:[10.1109/TIP.2011.2147325](https://doi.org/10.1109/TIP.2011.2147325)
37. Nalpanitidis L, Gasteratos A (2010) Stereo vision for robotic applications in the presence of non-ideal lighting conditions. *Image Vis Comput* 28(6):940–951. doi:[10.1016/j.imavis.2009.11.011](https://doi.org/10.1016/j.imavis.2009.11.011)
38. Provenzi E, Fierro M, Rizzi A, Carli LD, Gadia D, Marini D (2007) Random spray retinex: a new retinex implementation to investigate the local properties of the model. *IEEE Trans Image Process* 16(1):162–171. doi:[10.1109/TIP.2006.884946](https://doi.org/10.1109/TIP.2006.884946)
39. Provenzi E, Gatta C, Fierro M, Rizzi A (2008) A spatially variant white-patch and gray-world method for color image enhancement driven by local contrast. *IEEE Trans Pattern Anal Mach Intell* 30(10):1757–1770. doi:[10.1109/TPAMI.2007.70827](https://doi.org/10.1109/TPAMI.2007.70827)
40. Rahman Zu, Jobson DJ, Woodell GA (2004) Retinex processing for automatic image enhancement. *J Electron Imaging* 13(1):100–110. doi:[10.1117/1.1636183](https://doi.org/10.1117/1.1636183)
41. Rao Y, Hou L, Wang Z, Chen L (2014) Illumination-based nighttime video contrast enhancement using genetic algorithm. *Multimedia Tools and Applications* 70(3):2235–2254. doi:[10.1007/s11042-012-1226-6](https://doi.org/10.1007/s11042-012-1226-6)
42. Rizzi A, Gatta C, Marini D (2003) A new algorithm for unsupervised global and local color correction. *Pattern Recogn Lett* 24(11):1663–1677
43. Saponara S, Fanucci L, Marsi S, Ramponi G (2007) Algorithmic and architectural design for real-time and power-efficient retinex image/video processing. *J Real-Time Image Proc* 1(4):267–283
44. Schlick C (1994) Quantization techniques for visualization of high dynamic range pictures. Springer, Heidelberg, pp 7–20
45. Shen J, Yang X, Jia Y, Li X (2011) Intrinsic images using optimization. In: *2011 IEEE conference on computer vision and pattern recognition (CVPR)*, pp 3481–3487. doi:[10.1109/CVPR.2011.5995507](https://doi.org/10.1109/CVPR.2011.5995507)

46. Sobol R (2004) Improving the retinex algorithm for rendering wide dynamic range photographs. *J Electron Imaging* 13(1):65–74. doi:[10.1117/1.1636762](https://doi.org/10.1117/1.1636762)
47. del Solar JR, Quinteros J (2008) Illumination compensation and normalization in eigenspace-based face recognition: a comparative study of different pre-processing approaches. *Pattern Recogn Lett* 29(14):1966–1979. doi:[10.1016/j.patrec.2008.06.015](https://doi.org/10.1016/j.patrec.2008.06.015)
48. Tenenbaum JM (1971) Accommodation in computer vision. Ph.D. thesis, Stanford, CA, USA. AAI7119769
49. Vonikakis V, Andreadis I, Gasteratos A (2008) Fast centre-surround contrast modification. *IET Image Process* 2(1):19–34
50. Vonikakis V, Chrysostomou D, Kouskouridas R, Gasteratos A (2013) A biologically inspired scale-space for illumination invariant feature detection. *Meas Sci Technol* 24(7):074–024. <http://stacks.iop.org/0957-0233/24/i=7/a=074024>
51. Vonikakis V, Kouskouridas R, Gasteratos A (2013) A comparison framework for the evaluation of illumination compensation algorithms. In: 2013 IEEE international conference on imaging systems and techniques (IST), pp 264–268. doi:[10.1109/IST.2013.6729703](https://doi.org/10.1109/IST.2013.6729703)
52. Vonikakis V, Winkler S (2012) Emotion-based sequence of family photos. In: Proceedings of the 20th ACM international conference on multimedia, MM '12. ACM, New York, NY, USA, pp 1371–1372. doi:[10.1145/2393347.2396490](https://doi.org/10.1145/2393347.2396490)
53. Werner F, Maire F, Sitté J (2009) Topological slam using fast vision techniques. Springer-Verlag, Berlin, Heidelberg
54. Xiong W, Funt B (2009) Stereo retinex. *Image Vis Comput* 27(1–2):178–188
55. Yendrikhovski SN, Blommaert FJJ, de Ridder H (1998) Perceptually optimal color reproduction. doi:[10.1117/12.320117](https://doi.org/10.1117/12.320117)
56. Yoshida A, Blanz V, Myszkowski K, Seidel HP (2005) Perceptual evaluation of tone mapping operators with real-world scenes. doi:[10.1117/12.587782](https://doi.org/10.1117/12.587782)
57. Zhou S, Zhang F, Siddique MA (2015) Range limited peak-separate fuzzy histogram equalization for image contrast enhancement. *Multimed Tools Appl* 74(17):6827–6847. doi:[10.1007/s11042-014-1931-4](https://doi.org/10.1007/s11042-014-1931-4)



Vassilios Vonikakis received the degree in Electrical and Computer Engineering, as well as the PhD degree, from the Democritus University of Thrace, Greece. He currently works as a Research Scientist at the University of Illinois' Advanced Digital Sciences Center (ADSC) in Singapore. His research interests include Computer Vision, Affective Computing and Image Enhancement.



Rigas Kouskouridas is a computer vision/machine learning scientist at WIREWAX since September 2016. He has been a Postdoctoral Research Associate at the Computer Vision & Learning Lab at the Department of Electrical and Electronic Engineering of Imperial College London. He obtained his PhD in Computer Vision and Machine Learning in 2013 from the Department of Production Engineering and Management. He also holds a M.Eng. from the same Department (2006). His research interests include: Machine learning for vision with particular interest in object recognition, detection and pose estimation.



Antonios Gasteratos is a Full Professor of Mechatronics and Computer Vision at Democritus University of Thrace, Head of the Department of Production and Management Engineering and Director of the Laboratory of Robotics and Automation. He holds a B.Eng. and a PhD in Electrical and Computer Engineering, DUTH, (1994 and 1999, respectively). He has published over 160 papers in peer reviewed journals and international conferences. He is Assoc. Editor at the Electronics Letters and the International Journal of Optomechatronics. He is also reviewer of projects supported by the EC and of many international journals in the field of Computer Vision and Robotics. Antonios Gasteratos has been a member of programme committees of international conferences and chairman and co-chairman of international conferences and workshops.