OPTIMAL EXPOSURE VALUE SHIFT IN ACQUISITION OF HDR IMAGES

Tomislav Kartalov   Zoran Ivanovski   Ljupcho Panovski
kartalov@feit.ukim.edu.mk   mars@feit.ukim.edu.mk   panovski@feit.ukim.edu.mk
Faculty of Electrical Engeneering and Information Technologies, Skopje, Macedonia

ABSTRACT

This paper addresses the problem of choosing the optimal exposure value shift in the process of acquisition of input overexposed and underexposed images that are used to produce the output high dynamic range image. The analysis performed in the paper is based on mathematical background of the photometric properties of the digital still cameras, defined by the international standards established by CIE. The paper shows that in order to obtain input images suitable for the HDR process, for greyscale images represented by 8 bits per pixel, relative exposure difference should be at least 2 exposure points over and under the calculated optimal exposure for the observed scene.

Keywords: HDR, exposure value, CIE Lab

I. INTRODUCTION

The main objective of the modern electronic devices for digital image and video acquisition is to represent a captured scene as realistic, and as identical to the real eye-observed scene, as possible. However, the photometric quantities existent in the real scene, are differently interpreted by the electronic devices and the human observer, and this causes, in some cases, the visual appearance of the digital media to be far from visual appearances of the original scene. This especially applies to the real world scenes that have very high ratio between maximum and minimum intensity of the light in the scene, e.g. high dynamic range. Neither image acquisition device, nor display device is manufactured today, which can reproduce the intensity dynamic range that may exist in the real scene, or even the perceived luminance dynamic range of the human eye (brain). This often results in loss of details in the digital reproduction, and overall unpleasant image to the viewer. The "details problem" can be addressed by applying signal processing techniques that make use of multiple snapshots of the scene of interest. Using different light sensitivity settings (different exposure values) of the capturing device in each of the captured digital images, various segments of the whole light intensity interval that exist in the real scene can be captured. Images taken with longer exposures will reveal the darker objects, while the brighter objects in the scene will be shown on images with shorter exposure times. Combining these multiple snapshots into single understandable image can produce digital representation of the real scene that would be able of displaying on standard display devices, and yet represents broad interval of the real world light intensity. As pointed above, display devices have lower dynamic range than real world scene, so some dynamic range compression has to take place in the process.

Digital image and video community have put a lot of effort into optimal solution of the problem of different exposure image fusion, utilizing various concepts and methodologies [1] - [3]. All these proposed solutions use several input images obtained with different exposures, to produce single output image which will contain all the useful parts from the input images. However, all these algorithms assume perfectly spatially aligned input images. Every effort to implement this in practice will meet severe problems in the reality, because the time needed for acquisition of all input images is long enough to put very hard constraints on the types of the recorded scenes. The end user would be limited only to fixed scenes and a camera with tripod, complicating the whole "point and shoot" process. Implementation of any of these algorithms to a mobile phone with camera would be impossible, as a result of the fact that these devices are by definition handheld. Also, the great amount of data that should be processed as a result of capturing multiple images in a short time would appear as problem, taking into account the limited processing power and memory resources of the mobile devices.

The first step towards bringing the automated HDR algorithms closer to the implementation in real consumer equipment is to reduce the number of input images that are acquired for HDR processing. This is possible following the fact that the majority of the information necessary for producing the output HDR image is contained in only two input images, darker objects in the image with higher exposure, brighter objects in the image with lower exposure, and middle bright objects in both images. As shown in Fig.1, the human observer can easy reconstruct the recorded scene based only on two chosen images, while other four images carry redundant information.

Figure 1: Input images acquired with EV= - $^6/_3$; - $^4/_3$; - $^2/_3$; $^2/_3$; $^4/_3$; $^6/_3$ (respectively from left to right)

All the objects in the scene in Fig.1, edges, colors, textures, can be extracted from only two input images. The benefits of reducing the number of input images are:

- Lower amount of data that should be processed in the HDR algorithm.
- Less time needed for acquisition of the input images, reducing the limitation for shooting only static scenes and using only cameras with tripod.
- Almost every part of the image is extracted from only one, or maximum two input images. This helps with the consistency of the moving objects in the scene, which usually create a problem when their parts are extracted from multiple input images, because those objects could have different location in each of the input images.

However, taking only two input images from the scene also bears a certain risk of losing useful information if the two input images are not optimally chosen. If the Exposure Value (EV) difference between the two images is too low, there can be great deal of redundancy – the same information will be present in both images. If the EV difference is too high, there is a risk for missing the proper exposure range for some elements of the scene. They might be white-saturated in the overexposed image, and in the same time black-saturated in the underexposed image.

In this paper, we perform an analysis of the photometric properties of the digital camera (standalone, or built in a mobile phone), in order to obtain a relation between the light intensity in the scene, the exposure metering of the camera, and the luminance in the digital photography. That relation should help in deriving a conclusion about the optimal settings of the camera that would utilize a maximum range of the light intensity present in the scene.

The paper is organized as follows. In Section 2 the basic photometric definitions are given. Section 3 represents the actual analysis and the mathematical background of the paper, and in Section 4 few results are shown for visual confirmation of the obtained quantities. The last two Sections enclose the conclusions and the references.

II. Definitions

The process of digital image acquisition is usually consisted of three parts, metering, capturing, and post processing.

A. The Metering Part

The metering part's goal is to provide the camera (or the photographer) with accurate information about the light intensity and distribution in the recorded scene. Usually a separate part of the photographic equipment is responsible for the metering part, sometimes built in the camera, sometimes present as standalone, handheld light meter. The metering equipment responds to the physical – radiometric quantities of the light present in the scene, as the power of the electromagnetic radiance, however, all the measured quantities are calculated and displayed in their respective photometric counterparts, like light intensity expressed in candelas, or scene luminance expressed in candelas per square meter (or nits). Those quantities describe the scene as it is experienced by the human observer, in order to make solid starting point for capturing and later reconstructing the display of the scene as accurately as possible to the viewer. In that process, for the human observer to agree that certain scene is displayed properly, it is not necessary to reconstruct the exact amount of luminance present in the objects from the scene, but rather it is enough to rebuild the proper relative relations between the objects. If the distribution of the dark and bright objects from the scene is accurate in the reconstruction, the observer would accept the reconstruction regardless of the absolute level of the luminance in it. Even in the case when the brightest object in the reconstruction carries less absolute light intensity than the darkest object in the real scene, the reconstruction would still be accurate.

With that phenomenon taken into consideration, in order to obtain accurate reconstruction of the real scene, the camera should capture only the relative distribution of the luminance from the scene. However, all imaging devices, as stated above, have capability of capturing luminance range far shorter than the range that can occur in the reality. This means that certain parts of the luminance range will not be displayed in the reconstruction. In order to minimize those parts, or place them in the irrelevant interval of the range, it is important to know the average luminance in the recorded scene. Statistically, to obtain the best utilization of the luminance range of the scene, the luminance range of the camera should be placed symmetrically around the average luminance of the scene, e.g. the camera should consider the average luminance of the scene as middle point between the white (maximum light) and the black (absence of light). Only in that way the camera will be able to produce accurate relative luminance of the objects in the scene. The objects with luminance higher than the average will be reconstructed as bright objects, and the objects with luminance lower than average as dark objects.

B. The Capturing Part

Based on the information provided by the metering part, the camera (or the photographer) chooses the set of shooting parameters in order to obtain optimal exposure in the photography. The amount of exposure is directly proportional to the intensity of the light that enters the camera, and to the time interval in which the camera sensor is exposed to the light, [4]. These two parameters are controlled by the lens aperture setting (f-stop), and shutter speed of the camera, respectively. In the further text, to address these two settings simultaneously, we will use the EV (Exposure Value) numbers traditionally established in the photographic community as a way to refer to specific combinations of f-stops and shutter speeds. EV number locks together the two settings in such a way that changing either one automatically also changes the other to compensate, [4], thus maintaining the amount of exposure of the camera sensor. The physical representation of EV numbers is that increasing the amount of exposure by one EV unit indicates doubling of the intensity of the light, [4], [5], [6].

The capturing part of the digital image acquisition process transfers the distribution of the light from the scene to the camera sensor, and digitizes it. The digitization is performed by dividing the scene elements into finite resolution of the camera (the image is divided into pixels) and coding the quantity of light recorded by each pixel with certain number of bits. The most spread standard in the consumer photography is coding the luminance with 8 bits per pixel. That gives the values for luminance in the resulting digital image from 0 (black level) to 255 (white level), placing the center of the interval to the value of 127 (middle gray level). Hence, in order to obtain optimal utilization of the luminance range, the amount of the exposure (EV number) should be set in that manner, placing the average luminance from the scene as middle gray level in the photography.

The transfer function between the light intensity from the scene (real luminance) and the perceived brightness by the observer is not linear, [7], [8], [9]. The field tests with great number of observers showed that the surface that reflects only 18% of the incoming light is perceived by the human as half bright compared to the surface with 100% reflectance. On the other hand, the digital luminance in the photography is designed to be linearly perceived by the human, e.g. the level 127 is perceived as half bright compared to the level 255. For mapping all the real luminance levels from the scene to perceived brightness values, the equation given by the CIE (Commission Internationale de l'Eclairage) standard for calculating lightness L^* in the $L^*a^*b^*$ color space is used, equation (1). This equation gives the relationship between the perceived brightness by the human observer (the lightness), and the real luminance in the scene Y. The parameter Y_n is one of the three (X_n, Y_n, Z_n) tristimulus values of a specific white colour (referent white) calculated using the colour-matching functions of the CIE 1931 standard colorimetric system [7], [8].

$$L^* = 116 \cdot f\left(\frac{Y}{Y_n}\right) - 16 \qquad (1)$$

where

$$f\left(\frac{Y}{Y_n}\right) = \left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} \qquad \text{for } \left(\frac{Y}{Y_n}\right) > \left(\frac{6}{29}\right)^3$$

$$f\left(\frac{Y}{Y_n}\right) = \left(\frac{841}{108}\right) \cdot \left(\frac{Y}{Y_n}\right) + \frac{4}{29} \qquad \text{for } \left(\frac{Y}{Y_n}\right) > \left(\frac{6}{29}\right)^3$$

The following simplifications can be implemented to the equation (1). The unity value of luminance can be declared for the referent white ($Y_n = 1$), thus the value of the real luminance in the scene Y will be in the interval from 0 (for black object) to 1 (for white objects). The lightness L^* can be translated (mapped) linearly to the digital luminance values (0-255) in the photography. Calculating the limits and taking into account the respective value of the real luminance for the unity value of the digital luminance, the equation (1) can be rewritten as

$$DLL = 2.55 \cdot \left(116 \cdot \sqrt[3]{Y} - 16\right) \qquad \text{for } Y > 8.9 \cdot 10^{-3} \qquad (2)$$

Where DLL is the value of the digital luminance level (0-255) used in the digital image. The equation (2) gives simplified transfer function for digital camera operation, accurate for relative luminance values above 0.89%.

C. The Post Processing Part

This part performs post processing tasks such as white balance, gamma correction, lens imperfection compensation etc. All these processes should make the digital image closer to the display of the original scene and more acceptable for the end user. The explanation of these processes is beyond the scope of this paper.

III. THE OPTIMAL EXPOSURE VALUE DIFFERENCE IN HDR

Previously explained definitions give good starting point for making an analysis of the optimal EV number difference between two input images in the HDR algorithm. The analysis can be performed if the equation (2) is rewritten to express the real scene relative luminance levels as a function of digital luminance levels (DLL)

$$Y = \left(\frac{DLL + 40.8}{295.8}\right)^3 \qquad \text{for } DLL > 19 \qquad (3)$$

Calculated for all valid digital luminance levels (20 – 255), the graph for this function is shown in Fig.2. As can be observed, the digital middle gray level (DLL = 127) corresponds to 18% reflectance in the real scene.

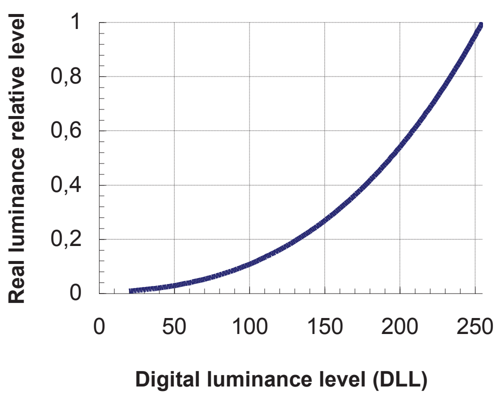

Figure 2: Real luminance vs. digital luminance

In the process of capturing digital images from the scene that has higher luminance dynamic range than the possible dynamic range of the camera, if a proper light metering is performed, the middle gray level of the camera still would be set to correspond to the average luminance in the scene. If an image is captured in that manner, that image will suffer from major white saturated and black saturated areas, matching up with the bright and the dark objects in the scene, respectively. So, the objects that have higher luminance than the average luminance in the scene, may end up as white objects (DLL=255) in the digital image, and similarly the objects with lower luminance may be reproduced as black objects (DLL = 0). The proper course of action to resolve this is to shoot two images instead of one, overexposed and underexposed image, both with exposure values deliberately shifted from the calculated optimum value (COV) for that scene. In this paper the analysis is performed for finding the optimal shift in EV numbers from the COV for capturing the overexposed and the underexposed image. Because the equation (3) is more accurate for higher luminance values (not valid for *DLL*<20), the analysis is given only for the underexposed image, assuming that statistically the white saturated and the black saturated areas have near symmetric distribution around the average of the luminance.

In order to estimate the necessary shift in EV numbers from COV for capturing the underexposed image, the relative luminance difference of the white saturated objects and middle gray level objects should be calculated:

$$Y_{WHI.SAT} = \left(\frac{255+40,8}{295.8}\right)^3$$
$$Y_{MID.GRAY} = \left(\frac{127+40,8}{295.8}\right)^3 \tag{4}$$

As stated above, increase or decrease in the EV number of one point corresponds to doubling or halving the relative luminance. Thus, to set the camera to capture the white saturated objects as middle gray level objects, the camera EV number should be decreased by

$$EV_{decrease} = \log_2\left(\frac{Y_{WHI.SAT}}{Y_{MID.GRAY}}\right) = \log_2\left(\frac{255+40.8}{127+40.8}\right)^3 = 2.454 \tag{5}$$

The closest standard values are -2.33 EV or -2.66 EV. The image captured with one of these exposure values will be generally darker, but the objects that were white saturated in the image captured with COV, here will be placed in the middle gray luminance range. This process should not interfere with the real white areas in the scene, because the white saturation is many to one mapping procedure. Every part of the scene with luminance higher than the darkest object that appears white saturated in the image captured with COV exposure, will also appear as white saturated. In the underexposed image captured with decreased EV some of those parts are possibly white saturated again, depending on the total luminance dynamic range in the scene. The similar explanation holds for black saturated regions in the overexposed image.

IV. Visual confirmation

For visual confirmation of the given analysis, the HDR procedure proposed in [3] is implemented on various combinations of the input images shown in Fig.1. The results are presented in the Fig.3. As can be seen, the result obtained with the input pair ± $^6/_3$ EV is visually comparable with the result obtained with all input images, while the results obtained with the other pairs with lower EV difference are obviously darker and less satisfactory. The small difference between the results shown in Fig. 3 c) and d) is a consequence of the fact that ± $^6/_3$ EV is not the estimated optimal EV shift, but is nearest to the optimal. The problem is the fact that ± $^6/_3$ EV is maximum EV difference achievable by the non-professional class of equipment, and in the lack of better equipment, the theoretically optimal case ± $^7/_3$ EV could not be tested for this paper.

V. Conclusions

In this paper, we performed mathematical analysis of the exposure value shift practiced in the capturing of the input images for HDR procedures. We explained the reasons for taking only two instead of multiple input images, and we estimated the optimal EV shift of these two images.

The obtained value for the EV shift is optimal in terms of maximum utilization of the luminance range in the recorded scene.

Unfortunately, the most of low and middle end imaging devices have the possibility to change the exposure value number up to 2 EV, and the values above 2 EV are only achievable in the professional and semi-professional equipment. That leaves the conclusion that for low and middle end equipment the maximum setting for EV number should be always used in HDR procedures.

In the scenes where the HDR procedure is not really needed, but still implemented, taking several input images with different exposures will yield a better result than taking only two input images with highest EV difference. Hence, the necessity occurs for developing an algorithm that decides whether in the recorded scene the implementation of HDR would bring benefit.

VI. Acknowledgement

The results presented in this paper were obtained in the course of a research project commissioned and funded by NXP Software B.V., Eindhoven.

REFERENCES

[1] Ardeshir Goshtasby, "Fusion of multi-exposure images", *Image and Vision Computing,* vol. 23, pp. 611-618, 2005.

[2] Ron Rubinstein, "Fusion of Differently Exposed Images", Technion, Israel Institute of Technology, Final Project Report, Oct 2004.

[3] Tom Mertens, Jan Kautz, Frank Van Reeth, "Exposure Fusion", *Proc. of the 15th Pacific Conference on Computer Graphics and Applications,* Maui, Hawaii, Oct/Nov 2007, pp. 382-390.

[4] Ansel Adams, *The Negative*, Little, Brown and Company, New York, 2005. ISBN 0-8212-2186-8 (PB)

[5] Jeff Conrad, *Exposure Metering – Relating Subject Lighting to Film Exposure*, 2003, pdf article published on internet: http://www.largeformatphotography.info/articles/conrad-meter-cal.pdf and also: http://cybered.swu.ac.th/content/images/flash/conrad-meter-cal.swf

[6] Kodak, *Basic Sensitometry and Characteristics of Film*, Eastman Kodak Company, 2006

[7] J. Schanda, "CIE Colorimetry and Colour Displays", Final Program and Proceedings of IS&T/SID Fourth Color Imaging Conference: Color Science, Systems and Applications, pp. 230-234, 1996.

[8] CIE Draft Standard, *Colorimetry – Part 4: CIE 1976 L*a*b* Colour Space*, CIE DS 014-4.3/E: 2007

[9] ISO International Standard, *ISO 12232: Photography – Digital still cameras – Determination of exposure index, ISO speed ratings, standard output sensitivity, and recommended exposure index*, ISO Geneva, 2006

a) Using images captured with ± $^2/_3$ EV

b) Using images captured with ± $^4/_3$ EV

c) Using images captured with ± $^6/_3$ EV

d) Using all input images

Figure 3: Results from the HDR algorithm proposed in **[3]**, for various sets of input images