

High Dynamic Range Image Based on Multiple Exposure Time Synthetization

Annamária R. Várkonyi-Kóczy

Dept. of Measurement and Information Systems, Budapest University of
Technology and Economics, Budapest, Hungary,
Integrated Intelligent Systems Japanese-Hungarian Laboratory
koczy@mit.bme.hu

András Rövid

Dept. of Measurement and Information Systems, Budapest University of
Technology and Economics, Budapest, Hungary,
Integrated Intelligent Systems Japanese-Hungarian Laboratory
Dept. of Electrical and Electronics Engineering, Shizuoka University, Japan
tarovid@ipc.shizuoka.ac.jp

Szilveszter Balogh

Dept. of Measurement and Information Systems, Budapest University of
Technology and Economics, Budapest, Hungary,
Integrated Intelligent Systems Japanese-Hungarian Laboratory

Takeshi Hashimoto, Yoshifumi Shimodaira

Dept. of Electrical and Electronics Engineering, Shizuoka University, Japan

Abstract: High dynamic range of illumination may cause serious distortions and other problems in viewing and further processing of digital images. In this paper a new tone reproduction preprocessing algorithm is introduced which may help in developing hardly or non-viewable features and content of the images. The method is based on the synthetization of multiple exposure images from which the dense part, i.e. regions having the maximum level of detail are included in the output image. The resulted high quality HDR image makes easier the information extraction and effectively supports the further processing of the image.

Keywords: high dynamic range images, image reproduction, image processing

1 Introduction

Digital processing can often improve the visual quality of real-world photographs, even if they have been taken with the best cameras by professional photographers in carefully controlled lighting conditions. This is because visual quality is not the same thing as accurate scene reproduction. In image processing most of the recently used methods apply a so-called preprocessing procedure to obtain images which guarantees – from the point of view of the concrete method – better conditions for the processing. A typical example is noise elimination from the images which yields much better results as else. There are many kinds of image properties to which certain methods are more or less sensitive [1]. The different image regions usually have different features. The parameters of the processing methods in many of the cases are functions of these image features.

The light intensity at a point in the image is the product of the reflectance at the corresponding object point and the intensity of illumination at that point. The amount of light projected to the eyes (luminance) is determined by factors such as: the illumination that strikes visible surfaces, the proportion of light reflected from the surface and the amount of light absorbed, reflected, or deflected by the pre-vailing atmospheric conditions such as haze or other partially transparent media [2].

An organism needs to know about meaningful world properties, such as color, size, shape, etc. for the interpretation of the view. These properties are not explicitly available in the retinal image and must be extracted by visual processing.

In this paper we will deal with the reproduction of images when the high dynamic range of the lightness causes distortions in the appearance and contrast of the image in certain regions e.g. because a part of the image is highly illuminated looking plain white or another is in darkness.

The dynamic range in photography describes the ratio between the maximum and minimum measurable light intensities. High dynamic range (HDR) imaging covers a set of techniques that allow a far greater dynamic range of exposures than normal digital imaging techniques [3]. HDR images enable to record a wider range of tonal detail than the cameras could capture in a single photo.

Recently HDR imaging techniques have come into the focus of research because of their high theoretical and practical importance. The application of HDR capable sensors in cars may get an important role in traffic safety, because in many of the cases the environment around the car has high dynamic range (dark and bright) illumination. (Just consider the case when a car leaves a (dark) tunnel and enters the (bright) sunshine.) An HDR sensor, in contrast to a linear sensor, can detect details that bright environment washes out and it misses fewer details in dark environment [3].

Another example can be the application of HDR techniques in the preprocessing phase of images, by which the performance of different image processing algo-

rithms like corner and edge detectors or scene reconstruction algorithms can be improved.

Several methods can be found in the literature addressing the problem. Each of them tries to compress the high dynamic range of luminance values to a displayable range keeping as much amount of information as possible. Just to mention some characteristic approaches, Rubinstein and Brooks's method in [4] is based on fusion in the Laplacian pyramid domain. The core of the algorithm is a simple maximization process in the Laplacian domain. Wide dynamic range CMOS image sensors play also very important role in HDR imaging (see e.g. Kawahito's work in [5]). The sensor introduced in [5] uses multiple time signals and such a way extends the image sensor's dynamic range. Reinhardt in [6] applies a so-called zone system. The authors of this paper also have introduced new fuzzy-based tone reproduction algorithms in [7], [8].

2 Gradient Based Multiple Exposure Time Synthetization Algorithm

2.1 The Basic Concept

Given N images of a static scene obtained at different exposures using a stationary camera. The introduced method combines the images into a single image in which each of the input image information is involved without producing noise. The main idea is the following:

Each image is segmented into small local regions, each of the same size. The shape of these regions is rectangular. The proposed method selects the most informative image for each local image region. The main task is to identify the image that contains the highest density of information within a local region.

When the dynamic range of the scene is high taking just one photo using a normal camera is not enough for producing a HDR image. In such cases several pictures are needed to capture all the scene details. These images should be merged together such a way, that all the involved information should be preserved.

If the scene contains regions with high luminance values, then in that highly illuminated region it is necessary to take a picture with low exposure time for the visualization of the details. On the other hand, if the scene contains very dark areas, then the exposure time should be possibly much higher. Approaching from the opposite site, if we have images with different exposures we have to decide somehow which exposure contains the maximum level of information in case of a certain region. There are existing methods, which use statistical elements for se-

lecting the most informative image region while others apply the histogram of the luminance values of the processed region.

According to our experiments, the level of the details in a region can be measured based on the sum of gradient magnitudes of luminance in that region. The complexity of this approach is lower than that of the other ones, thus the processing time can also be reduced. As detailed is the information as higher the sum of the gradient values is in that region.

The main step is to find the segments with the highest sum of gradients of luminance values. After merging the segments, as the output of this task we get an image, which contains the maximum amount of information in each region.

Another important step is smoothing, because we have to eliminate the sharp transitions, which arise at the borders of the regions. A monotonically decreasing blending function is centered over the selected regions of the input images and the pixel intensities are weighted according to the corresponding blending function values. The blending function assigns the maximum weight to the pixel located at the center of the considered region. To the other pixels, weights inversely proportional to the distances from the center of the region are assigned. After this procedure we get the output image which will not contain discontinuities along the image regions.

2.2 Measuring the Level of the Detail in an Image Region

For extracting all of the details involved in a set of images of the same scene made with different exposures, it is required to introduce a factor for characterizing the level of the detail in an image region. For this purpose we propose the gradient of the intensity function corresponding to the processed image and a linear mapping function, which is applied for setting up the sensitivity of the measurement of the detail level. In the followings the description of the estimation of the mentioned factor is introduced.

Let $I(x,y)$ be the pixel luminance at location $[x, y]$ in the image to be processed. Let us consider the group of neighboring pixels which belong to a 3x3 window centered on $[x, y]$. For calculating the gradient of the intensity function in horizontal ΔI_x and vertical ΔI_y directions at position $[x, y]$ the luminance differences between the neighboring pixels are considered:

$$\begin{aligned}\Delta I_x &= |I(x+1, y) - I(x, y)| \\ \Delta I_y &= |I(x, y-1) - I(x, y)|\end{aligned}\tag{1}$$

For the further processing the maximum of the estimated gradient values should be chosen, which solves as the input of the normalized linear mapping function P defined as follows:

$$P(v) = v / I_{\max}, \quad (2)$$

where $I_{\max}$ is the maximum luminance value. (For 8 bit grayscale images it equals 255.)

Let $\mathbf{R}$ be a rectangular image region of width rw and height rh , with upper left corner at position $[x_r, y_r]$. The level of the detail inside of region $\mathbf{R}$ is defined as

$$M_D(\mathbf{R}) = \sum_{i=0}^{rw} \sum_{j=0}^{rh} P(\max(\Delta I_x(x_r + i, y_r + j), \Delta I_y(x_r + i, y_r + j))) \quad (3)$$

As higher is the calculated M_D value as detailed the analyzed region is. In the followings we will use this parameter for characterizing the measure of the image detail.

2.3 HDR Image Synthetization

Let I_k denote the intensity function of the input image with index k , where $k=1, \dots, N$ and N stands for the number of images to be processed, each of them taken with different exposure. Each image contains regions, which are more detailed as the corresponding regions in the other $N-1$ images. Our goal is to produce an image, which is the combination of the N input images and contains all details involved in all of the images without producing noise. Using such detailed image, the most of the feature detection methods can be improved and can effectively be used even if the lighting conditions are not ideal.

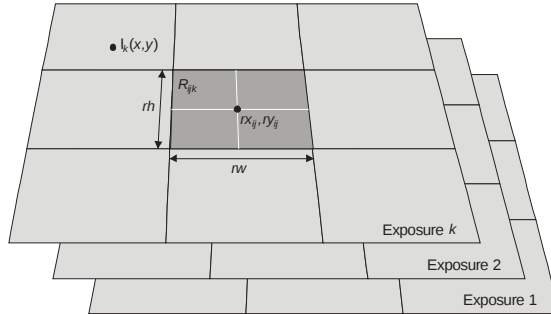


Figure 1
Image regions

The first step of the processing is to divide the pictures into n rows and m columns, which yields $n \times m$ rectangular image regions. The regions in the images are of the same size with height rh and width rw (see Figure 1). (In Figure 1 you can see a 3×3 division.) Let $\mathbf{R}_{ijk}$ denote the region in the i th row and j th column of the image with index k . Let rx_{ij} , ry_{ij} denote the horizontal and vertical coordinates of the center of the region in the i th row and j th column. $I_k(x,y)$ stands for the intensity of the pixel at position (x,y) in the image with index k .

For each image, the level of the detail has to be estimated inside every region $\mathbf{R}_{ijk}$. This information helps us to select the most detailed regions among the corresponding image parts (indexed by the same i and j values).

Let $\mathbf{D}$ denote the matrix of regions with the highest level of detail. Let d_{ij} be the element in the i th row and j th column of $\mathbf{D}$, which stands for the index of the image, which has the most detailed region in the i th row and j th column, i.e.

$$M_D(R_{ijl}) > M_D(R_{ijk}) \mid k=1, \dots, l-1, l+1, \dots, N; l = d_{ij}. \quad (4)$$

The next step is to merge the regions $\mathbf{R}_{ijl}$ together, where $l=d_{ij}$ $i=1..n$, $j=1..m$. Merging the selected regions together results in an image that contains every detail involved in the N input images. Unfortunately, the resulted image usually contains sharp transitions along the borders of the regions. These sharp transitions should be eliminated. For this purpose the Gaussian blending function [4] can be applied advantageously, which has the form of

$$B_{ij}(x, y) = \frac{e^{-\left(\frac{(x-rx_{ij})^2}{2\sigma_x^2} + \frac{(y-ry_{ij})^2}{2\sigma_y^2}\right)}}{\sum_{p=1}^m \sum_{q=1}^n e^{-\left(\frac{(x-rx_{pq})^2}{2\sigma_x^2} + \frac{(y-ry_{pq})^2}{2\sigma_y^2}\right)}}. \quad (5)$$

Here i and j stands for the row and column indices of the region over which the Gaussian hump $G_{ij}(x, y)$ is centered (see Figure 2), m denotes the total number of columns and n the total number of rows in the input images. σ_x and σ_y stand for the standard deviation of the 2D Gaussian function. The values rx_{pq} and ry_{pq} represent the coordinates of the center of the region in p th column and q th row.

Let U be a function defined as

$$U(x, y) = \begin{cases} 1 & |x| \wedge |y| \leq \varepsilon \\ 0 & \text{else} \end{cases}. \quad (6)$$

Function U is used for eliminating the influence of those segments, whose center points fall outside a predefined ε environment of the actually processed pixel.

Using the blending function (eq. (5)) and function U the luminance value of the output image can be evaluated according to

$$I_{out}(x, y) = \sum_{i=1}^n \sum_{j=1}^m B_{ij}(x, y) U(x - rx_{ij}, y - ry_{ij}) I_{d_{ij}}(x, y). \quad (7)$$

The output luminance can be influenced by changing the size of the regions and the standard deviation of the Gaussian functions. As smaller is the standard deviation as higher influence the regions with low detail level onto the result have.

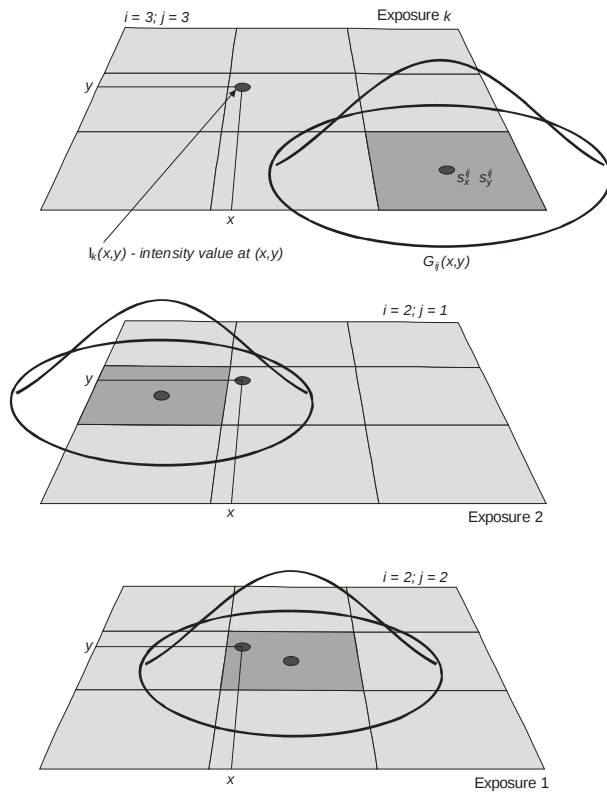


Figure 2

Illustration of the regions covered by Gaussian hump

Using such detailed image the edges can also be effectively extracted and advantageously be used by further processing, by object recognition, scene reconstruction etc.

In Figure 2 in case of exposure 1 the region in the 2nd row and 2nd column has the maximum level of detail compared to the other images, in case of exposure 2 the region in the 2nd row and 1st column while in case of exposure k the region in the 3rd row and 3rd column is the most detailed. The Gaussian functions are centered at (rx_{ij}, ry_{ij}) of the maximum level regions.

3 Illustrative Examples

The effectivity of the proposed algorithm is illustrated by two examples. In both of the examples the width and height of the regions are chosen to be the 1/15 part of the image width and height, respectively. Deviations σ_x and σ_y have the same value as the width and height of the regions, respectively.

Example 1: In Fig. 3 the original under-exposed image and its edge map can be seen. Fig. 4 represents the original over-exposed image and its edge map. In Fig. 5 the resulted HDR image using the proposed method and its edge map can be followed.

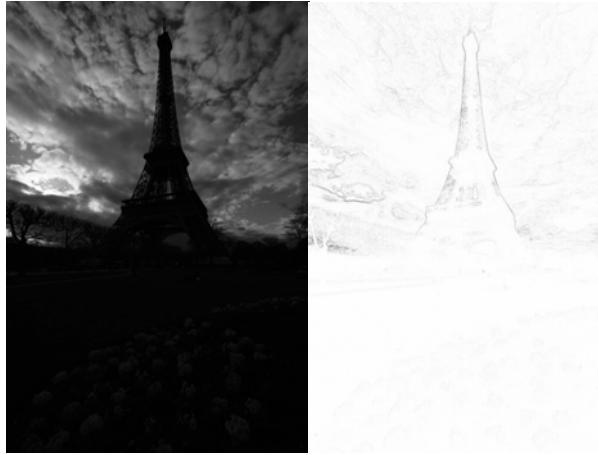


Figure3

Under-exposed image: Original image (left) and its edge map (right)



Figure 4

Over-exposed image: Original image (left) and its edge map (right)



Figure 5

Obtained HDR image: The resulted high dynamic range image (left) and its edge map (right)

Example 2: In Figs. 6(a)-(e) five different exposure images can be seen. Fig. 7 stands for the resulted HDR images using the proposed method.



(a)



(b)



(c)



(d)



(e)

Figure 6

Multiple exposure images



Figure 7

The obtained output HDR image (a) output image

Conclusions

In this paper a new gradient based approach is introduced for extracting image details. The method uses multiple exposure images of the same scene as input data. During the processing the images are divided into regions and always the most detailed region is chosen from the different exposure versions for the output image. With the help of this technique it becomes possible to produce a good quality HDR image from a set of bad quality photos taken by a range of exposures.

Acknowledgment

This work was sponsored by the Hungarian Fund for Scientific Research (OTKA T049519) and the Structural Fund for Supporting Innovation in New Knowledge and Technology Intensive Micro- and Spin-off Enterprises (GVOP-3.3.1-05/1.2005-05-0160/3.0).

References

- [1] Russo, F.: Recent Advances in Fuzzy Techniques for Image Enhancement, IEEE Transactions on Instrumentation and Measurement, Vol. 47, No. 6, Dec. 1998, pp. 1428-1434
- [2] Adelson, E. H., A. P. Pentland: The Perception of Shading and Reflectance, In D. Knill and W. Richards (eds.), Perception as Bayesian Inference, New York: Cambridge University Press, 1996, pp. 409-423
- [3] Li, Y., L. Sharan, E. H. Adelson: Perceptually-based Range Compression for High Dynamic Range Images, Journal of Vision, Vol. 5, No. 8, August 2005, p. 598
- [4] R. Rubinstein, A. Brook: Fusion of Differently Exposed Images, Final Project Report, Israel Institute of Technology, p. 14, 2004
- [5] M. Sasaki, M. Mase, S. Kawahito, Y. Wakamori: A Wide Dynamic Range CMOS Image Sensor with Multiple Exposure Time Signals and Column-Parallel Cyclic A/D Converters, IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors, 2005
- [6] Reinhard, E., R. E. Stark, M. Shirley, P. J. Ferwerda: Photographic Tone Reproduction for Digital Images, In Proc. of the 29th Annual Conference on Computer Graphics and Interactive Techniques, 2002, San Antonio, Texas, pp. 267-276
- [7] Várkonyi-Kóczy, A. R., A. Rövid, P. Várlaki: Fuzzy-based Brightness Compensation for High Dynamic Range Images, In Proc. of the 9th International Conference on Intelligent Engineering Systems, INES 2005, Cruising on Mediterranean Sea, Greece, Turkey, Sept. 16-19, 2005, pp. 245-248
- [8] Várkonyi-Kóczy, A. R., A. Rövid: High Dynamic Range Image Reproduction Methods, In CD-Rom Proc. of the 2006 IEEE Instrumentation and Measurement Technology Conference, IMTC/2006, Sorrento, Italy, April 24-27, 2006