OPTIMIZATION OF THE PYRAMID HEIGHT IN THE PYRAMID-BASED EXPOSURE FUSION ALGORITHMS[1]

Tomislav Kartalov, Zoran Ivanovski, Ljupcho Panovski,
Faculty of Electrical Engineering and Information Technologies – Skopje, Republic of Macedonia …

Abstract - *Some of the most successful exposure fusion algorithms are based on image decomposition in Gaussian/Laplacian pyramids. However, the visual quality of the fused result is highly dependent on the number of decomposition levels, i.e. the pyramid height. In this paper, the analysis of that dependence is performed, and an automated algorithm for optimization of the pyramid height is proposed. The algorithm is computationally inexpensive, as it uses results and parameters already computed in the pyramid decomposition process. This algorithm is intended to be a part of a fully automated, low processing power exposure fusion solution, designed for implementation on mobile platform devices.*

Key Words— *exposure fusion, Gaussian/Laplacian pyramid, mobile platform*

1. INTRODUCTION

High Dynamic Range (HDR) imaging was introduced in the photographic community a long time ago, as a method to increase the luminance dynamic range of the image, when a dynamic range of the recorded scene is higher than the dynamic range which the acquisition device is capable of recording. With this method, instead of one, two or several input images are recorded from the scene, each with a different exposure setting. Afterwards, a combination of the input images is made to obtain an output image with higher visual quality. The HDR imaging can be implemented by two different concepts, the true HDR, and the exposure fusion (EF). The true HDR tries to reconstruct the actual radiance map of the recorded scene, and for the result to be presented on a low dynamic range devices, it has to undergo a process of tone mapping. The EF, which is more common, employs the usable parts of the input images as they are or with little processing, not concerning the actual radiance interdependencies between the objects in the scene (brighter, darker…). It generally produces visually pleasant output images, although often with mild surrealistic appearance, which is sometimes used to reach some artistic expression.

Limited only to the professionals with a lot of knowledge and expensive equipment few decades back, nowadays the EF is available almost to everyone holding a camera. Various software tools for image processing support it, and the scientific community constantly provides new and better solutions for the procedure.

Usually the EF procedure is build from two main parts, the selection part, which locally decides about the best representative(s) among all the input images, and the blending part, whose task is to seamlessly join the selected parts from different input images into one result.

One class of the EF algorithms shows particularly high robustness to the local errors in the selection part, [1], [2]. The EF in both algorithms is performed in a similar manner, using Gaussian/Laplacian pyramid decomposition, as explained in detail in [3], [4], on multiple input images. In the selection part, Mertens in [2] presents more complex scheme for computing the local weights of participants in the output image, while Rubinstein, in [1], uses very simple selection model, producing binary selection maps, and thus locally choosing only one of the submitted images. Mertens and Rubinstein both employ Gaussian and Laplacian pyramid decomposition of the input images, and of the selection (weight) maps, afterwards building the output image from the useful parts of the pyramid decompositions.

Both [1] and [2] algorithms perform well, however multiple parameters have to be adjusted by the user, depending on the content and statistics of the input images. Also, they lack automatic decision for the pyramid height (the final number of levels in which the pyramid decomposition is performed), which can change the quality of the results by a great scale.

In this paper, we propose an algorithm for optimization of the pyramid height in pyramid based exposure fusion algorithms, in order to achieve the best quality result without end user intervention.

The paper is organized as follows: in Section 2 the dependence between the number of pyramid decomposition levels and the fused image quality is explained. The Section 3 contains the details about the proposed algorithm. The paper is concluded in Sections 4 and 5 with the obtained results and discussions.

2. THE PYRAMID HEIGHT VS. THE QUALITY OF THE FUSED IMAGE

The visual appearance of the fused results obtained with the reconstruction of the Gaussian/Laplacian pyramid in the EF procedure is very dependent on the number of the levels

[1] The results were obtained in the course of a research project commissioned and funded by NXP Software B.V., Eindhoven

of the pyramid decomposition (the pyramid height). The exact number of levels of the decomposition is not especially informative, as it must be observed together with the dimensions (resolution) of the whole image. Larger images would have to be processed in a higher number of pyramid levels. A more informative parameter for the pyramid height is the resolution of the last (smallest) image in the Gaussian series. For that reason, those resolutions are used as level numbers in the further text, e.g. the notation “level 2×2” would mean that the image is decomposed until its Gaussian pyramid representative reaches resolution of a 2 pixels wide, and 2 pixels tall image. Of course, for the pyramid decomposition to be able to reach levels with low resolutions such as 2×2 or 4×4, the input image should be square, and the number of its rows and columns should be power of two (2^B, B is integer). This is required in order to obtain dimensions divisible by two (for the decimation process) in all levels of the pyramid. The most of the real images do not satisfy these conditions. However, usually the standard resolutions of the images allow the first three levels of the pyramid decomposition without appearance of the odd numbered dimension. Thus, the pyramid decomposition is performed regularly in the first three levels (to save memory space), and then the third member of the Gaussian pyramid is expanded to the square dimensions of the next power of two. This allows the decomposition to continue as far as it is required. The weight map also has to be expanded, since its decomposition into Gaussian pyramid continues too. The expansion is performed by padding the images with their values in a mirror fashion, in order to maintain the statistical properties of the images.

For the purposes of the analysis performed in this paper, a simple pyramid based EF algorithm is developed. It uses only two images at its input, the overexposed and the underexposed, Figure 1. In its selection part, simple thresholding of the luminance values in the overexposed image is implemented. All pixels in the overexposed image having luminance values above 80% of the Maximum Possible Luminance - MPL value (e.g. 255 in an 8 bit luminance representation) are declared as white-saturated pixels, and they are discarded from the image. At their locations, the respective pixels from the underexposed image are included in the fused result. The pixels from the overexposed image having luminance values lower than 80% of the MPL, are considered as best representatives for their locations, regardless of the luminance values of their respective pixels from the underexposed image. All decisions about the selected best representatives are stored in a weight map, according to predefined rule. E.g. the value of a certain pixel of the weight map could be a percentage of the contribution of the overexposed image in the fused image, at that particular location – the value 100 would indicate that at that location the luminance of the fused result would be equal to the luminance of the underexposed image. The blending part of the EF algorithm is designed similarly to the blending part proposed in [2], however the smoothing filters implemented in our algorithm are computationally cheaper, operating only on integer numbers and including only additions and binary shifts in their design.

For the input images shown in Figure 1, the developed EF algorithm is used to produce fused results at different pyramid heights, from 32×32 up to the last possible decomposition level 1×1. The results are presented in Figure 2. There can be noticed that a great amount of the fused image is subject to change under the influence of the pyramid height.

If the pyramid is too low, the interdependency between separate image parts could be weak, and this can lead to visible transitions in the luminance values from one region of the image to the other, known as halos, observable in Figure 2 a), b) and c).

If the pyramid is too high, the global averaging of the image features at a higher pyramid levels could lead to over-compensation of the luminance dynamic range, decreasing of the contrast and de-saturation of the colors. This effect is barely noticeable in the Figure 2 e), and very prominent in the Figure 2 f).

The halo effect or the color de-saturation effect can occur at different pyramid heights for different input images, depending on image contents and statistics. The optimal pyramid height which can produce artifact free output image is different for various input images. Our visual experiments showed that the images with best quality are always obtained with pyramid decompositions up to the levels with resolutions of 8×8, 4×4, or 2×2 pixels. However, a procedure for selection of the best image from these three is necessary for designing a fully automated EF algorithm.

a) overexposed image b) underexposed image

Figure 1. Input images

a) 32×32　b) 16×16　c) 8×8

d) 4×4　e) 2×2　f) 1×1

Figure 2. Fused results obtained with different pyramid heights

3. OPTIMIZATION OF THE PYRAMID HEIGHT

The most straight-forward way to obtain the resulting image with the best quality in the pyramid based EF algorithms is to generate the fused images with various pyramid heights and to choose the best through their comparison. This method is inapplicable in the practice, because the generation of multiple fused images would result in unjustifiably high computational complexity of the algorithm. Thus, the pyramid decomposition must be stopped at the level which would produce the best fused image, and for the stopping decision only the data available in the decomposition process should be used, with as low as possible increase of its complexity.

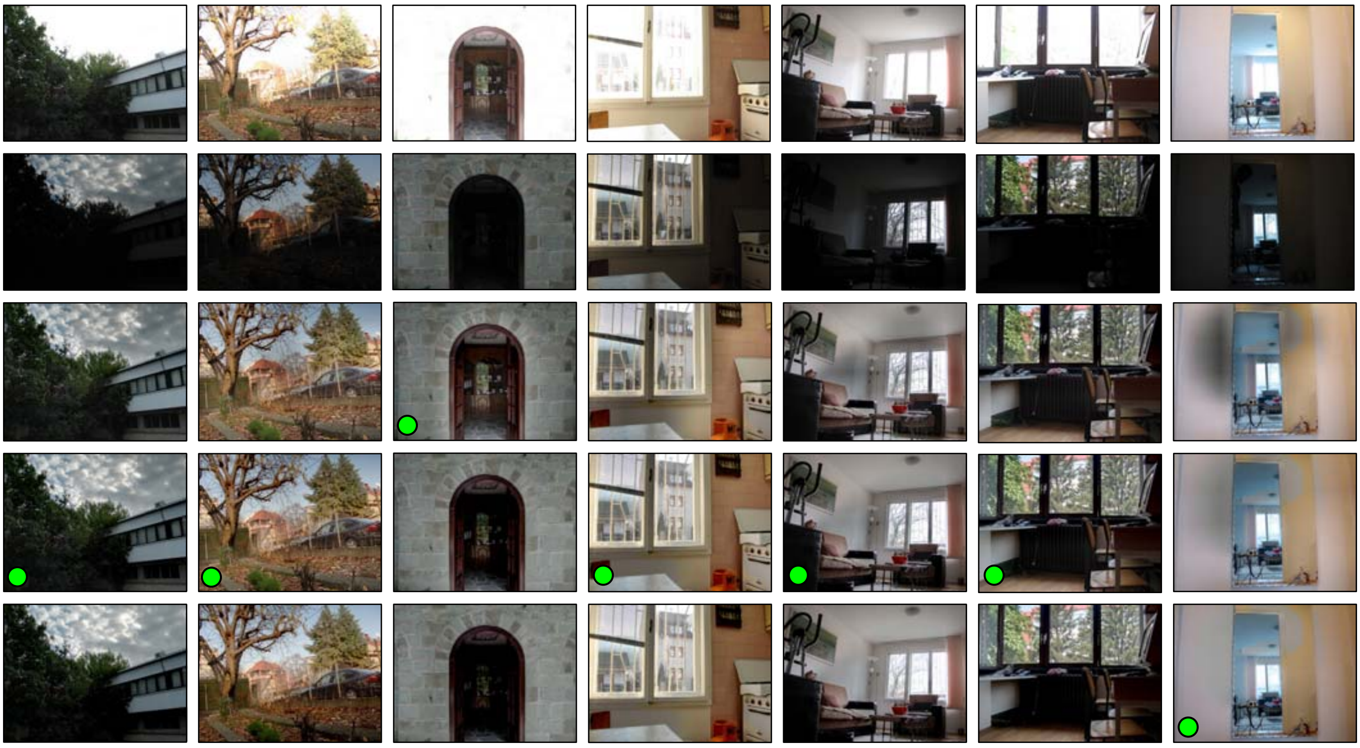
Two parameters are observed as criteria for stopping the pyramid decomposition. The first parameter is the image activity (presence of transitions in luminance values), which could mask the halo effect. The halo effect is hardly noticeable in images with high activity (many objects, edges and details in the image), so in such images the pyramid decomposition could be stopped at the lower level of the pyramid without the concern about visible halo artifacts. The image activity is calculated based on the Laplacian pyramid members, available during the pyramid decomposition. The Laplacian pyramid contains the image features (details, edges, luminance transitions); therefore the information about the quantity of the image activity can be extracted from it. The activity of the image ACT_L, for the level L of the Laplacian pyramid, is calculated as absolute average value of the Laplacian member $\mathbf{D}_L$ at that level.

The second observed parameter that can stop the pyramid decomposition is the smoothness of the weight map $\mathbf{GM}_L$ at a particular Gaussian pyramid decomposition level L. If the weight map is smooth, with generally uniform values, the two input images participate to the fused images with shares that are not changed to a large extent throughout the fused image. If the weight map contains high difference elements, that means that some parts of the fused image is build mainly, or even exclusively from the features of only one of the input images. This situation could introduce the halo artifacts, and in such case the pyramid decomposition should continue to the next higher level, until a relatively smooth weight map is obtained. The smoothness of $\mathbf{GM}_L$ is calculated testing its extreme values (minimum and maximum) against the adaptive interval $[I_L, I_H]$. If both of the extremes have values within the tested interval, then $\mathbf{GM}_L$ is **smooth-order 2**. If only one of the extremes is within the interval, then $\mathbf{GM}_L$ is **smooth-order 1**. Otherwise $\mathbf{GM}_L$ is **non-smooth**. The endpoints of the interval for pyramid level L depend on the image activity, as shown in (1), where T_{H1} is empirical threshold and M_V is the maximum possible value of the elements in the weight map.:

$$\left[I_L,\, I_H\right]\Big|_{\text{Level}=L} = \left[M_V - (ACT_L + T_{H1})\ ,\ ACT_L + T_{H1}\right] \qquad (1)$$

The criteria that can stop the pyramid decomposition are too high activity, or too smooth map. The implementation of those criteria is being observed starting at the level with resolution 8x8. If none of the criteria is not meet until the level with resolution 2x2 is reached, the pyramid decomposition is stopped at that level, without checking any further criteria. The 1×1 resolution level is never used. For 8×8 and 4×4 levels the decomposition is stopped if any of the following three conditions is met:

- The map is **smooth-order 2**.
- The map is **smooth-order 1** AND $ACT_L > T_{H3}$.
- $ACT_L > T_H(L)$..



a) b) c) d) e) f) g)

Figure 3. The results of the implemented pyramid height optimization algorithm. Top row: the overexposed images. Second row: the underexposed images. Third row: the fused images obtained with pyramid decomposition up to the level 8×8. Fourth row: the fused images obtained with pyramid decomposition up to the level 4×4. Bottom row: the fused images obtained with pyramid decomposition up to the level 2×2. The stopping decisions of the proposed algorithm are indicated with green points.

The used threshold values are experimentally adjusted to obtain optimal stopping of the pyramid decomposition, and for 8 bit luminance image their values are: T_{H1}=225; T_{H3}=11; $T_H(L)$=15 for level 8×8; and $T_H(L)$=10 for level 4×4.

4. THE EXPERIMENTAL RESULTS

The performance of the explained pyramid decomposition stopping scheme can be studied observing the Figure 3. The first two rows present various examples of the overexposed and the underexposed input images. The rows 3-5 present the fused images obtained by manually adjusting the pyramid height, forcing the pyramid decomposition to the level 8×8 (row 3), 4×4 (row 4) and 2×2 (the bottom row). The decision obtained by the proposed algorithm for automatic pyramid height adjustment for every input pair is indicated by green points. The exposure fusion algorithm directly generates the indicated result, without calculation of the rest of the choices, and without any end-user involvement in the process.

In most of the cases (the columns a) and b) as outdoor examples, and the columns d), e) and f) as indoor examples), the algorithm chooses the level 4×4 as final level of the pyramid decomposition. It seems that this level represents fine balance between the level 8×8, where generally halo artifacts are prominent, and the level 2×2, where the fused result begins to lose the color saturation and many objects become unrecognizable. These examples have moderate image activity and well balanced smoothness of the weight map. In terms of the visual quality, a decision of a human observer given all the results to pick the best would generally concur with the choice of the algorithm.

Figure 3 – column c) shows an example where the weight map has very high variance (it is not smooth), due to the highly accented complementary properties of the two input images. It is very clear here which parts of the fused image should be imported from which input image. The extreme values in the weight map can be a source of the halo artifacts. However, this image also has high spatial activity throughout its whole area, so the algorithm chooses to stop the decomposition at the 8×8 level. That is obviously the best choice, because a continuation of the decomposition to the higher levels would only make the objects and details inside the room less visible. A visible halo artifact can be spotted around the door; however every observer would agree that this artifact affects the image quality less than the loss of the details caused by further decomposition.

The column g) contains an example with relatively low spatial activity. Large flat areas and surfaces are observable in that scene, in which even a weak halo artifact, (such as that in the level 4×4 result), would be easy noticeable. For that reason, the algorithm chooses to continue the decomposition to the level 2×2, until all the luminance transitions caused by the halos are successfully flattened.

5. CONCLUSIONS

In this paper, an analysis of the dependence between the pyramid height and the image quality is performed. An automated algorithm for adjustment of the optimal pyramid height is proposed. The proposed algorithm is computationally not demanding, and it uses input parameters that are already calculated in the pyramid decomposition process. This algorithm is intended to be used as a part of fully automated exposure fusion algorithm designed for mobile platform devices.

The experimental results show that the proposed algorithm performs well, obtaining the best quality fused result under various statistical and contextual circumstances present in the processed images.

The future work should include testing the algorithm's performance versus the mean opinion score for the fused image quality, acquired from a large group of non-expert observers.

6. REFERENCES

[1] Ron Rubinstein, "Fusion of Differently Exposed Images", Technion, Israel Inst. of Technology, Project Report, Oct 2004.

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

[3] Peter J. Burt, Edward H. Adelson, "The Laplacian Pyramid as a Compact Image Code", *IEEE Trans. On Communications*, vol. com-3l, no. 4, pp. 532-540, April 1983.

[4] J. M. Ogden, E. H. Adelson, J. R. Bergen, P. J. Burt, "Pyramid-based computer graphics", *RCA Engineer*, vol. 30, no. 5, 1985.