IMAGE TONE MAPPING APPROACH USING ESSENTIALLY NON-OSCILLATORY BI-QUADRATIC INTERPOLATIONS COMBINED WITH A WEIGHTING COEFFICIENTS STRATEGY
Ba Chien Thai, Anissa Mokraoui, Basarab Matei, Mahdia Tabal

To cite this version:
Ba Chien Thai, Anissa Mokraoui, Basarab Matei, Mahdia Tabal. IMAGE TONE MAPPING APPROACH USING ESSENTIALLY NON-OSCILLATORY BI-QUADRATIC INTERPOLATIONS COMBINED WITH A WEIGHTING COEFFICIENTS STRATEGY. International Symposium on Signal Processing and Information Technology, Dec 2017, Bilbao, Spain. 10.1109/ISPIT.2017.8388688 . hal-02428266

HAL Id: hal-02428266
https://hal.science/hal-02428266
Submitted on 5 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
ABSTRACT
This paper proposes a Tone Mapping Operator (TMO) to convert High Dynamic Range (HDR) images into Low Dynamic Range (LDR) images. It is related on two consecutive stages. The first one decomposes the HDR image on several resolution levels according to a non-separable multiresolution approach using Essentially Non-Oscillatory (ENO) strategy where adapted bi-quadratic interpolations are exploited. The underlying idea of this decomposition is to better represent the details of the complex HDR image preserving as much as possible the HDR image quality. In the second stage, a weighting coefficients process is deployed to reduce judiciously the dynamic range of each subband’s coefficient. The weights are deduced from Gaussian filtering operations combined with a monotonic decreasing function where its parameters depend on the statistical properties of each HDR subband, neighborhood and resolution level. Simulation results show that the proposed image TMO has the ability to better represent the details compared to other operators.

1. INTRODUCTION
High Dynamic Range (HDR) images are considered as high-quality images on which different levels of exposure areas of the real-world scene are rendered at the same time faithfully to the Human Visual System (HVS). However these images cannot be viewed on standard Low Dynamic Range (LDR) display devices since their dynamic range is smaller than those of HDR images. Unfortunately HDR display devices are still too expensive to replace all LDR devices. To overcome this problem, Tone Mapping Operators (TMOs) have been developed [1]. Their main purpose is to reduce the dynamic range (contrast, color gamut, details...) of HDR images to LDR dynamic range device while preserving as much as possible the captured scene appearance. For more details, a complete state of the art is available in reference [1] where a classification into local, global, segmentation, frequency and perceptual operators is proposed. Although each TMO approach has its own specificity, it is not possible to review the wide range of the developed work in this paper. Only those selected (according to their good performance) to be compared to our approach are summarized below.

In [5], Durand and Dorsey proposed an edge-preserving bilateral filter to decompose the HDR image into two layers: a base layer encoding large-scale variations and a detail one. Contrast is then reduced only in the first layer while the details are kept unchanged. The combination of these tone mapping layers produces an LDR image. In [2], Drago et al. presented an adaptive logarithmic mapping method of luminance values. It concerns the adaptive adjustment of the logarithmic basis depending on the radiance of the pixels. A set of logarithmic functions ranging from $\log_2$ to $\log_{10}$ were used to preserve scene details and improve the rendering contrast. A bias power function was used to ensure a smooth interpolation between the different logarithm bases. In [11], Li et al. proposed a subband architecture related on an oversampled Haar pyramid representation. Subband coefficients are rescaled using a gain control function reducing the high frequency magnitudes and boosting low ones. The modified subbands are then convolved with the Haar synthesis filters and summed to reconstruct the final LDR image. In [8], Duan et al. proposed an optimization approach based on a histogram adjustment between linear mapping and equalized histogram mapping. In [9], Fattal et al. proposed a second generation of wavelets based on the edge content of the image avoiding having pixels from both sides of an edge. This approach was then exploited to map an HDR image into an LDR image. In [10], the authors proposed non-linear separable multiresolution families using data dependent interpolation where a set of operators have been derived. These latter have the ability to consider in their mathematical models the image singularity points. However the weakness of this approach is that the detail and approximation coefficients are weighted in the same way without taking into account the statistical properties of the coefficients. To improve the performance of this TMO, this paper investigates on non-separable and non-linear Essentially Non-Oscillatory (ENO) multiresolution where each subband’s coefficient is weighted according to the statistical properties of the subband.

This paper is organized as follows. Section 2 intro-
duces the basic concepts of the non-linear multiresolution approach using data dependent interpolation. Section 3 presents the non-separable and non-linear multiresolution image TMO. Section 4 evaluates the TMO approach performance. Section 5 concludes this paper.

2. BASIC CONCEPTS OF HARTEN’S NON-LINEAR MULTIRESOLUTION

This section introduces the non-linear multiresolution approach based on Essentially Non-Oscillatory (ENO) interpolation strategy developed by Harten [12], [13]. Cell-Average (CA) multiresolution is considered since it has the ability to introduce in its mathematical model the isolated singularities such as edge points in the image avoiding the Gibbs phenomenon particularly harmful in HDR tone mapped images.

Denote $j$ the resolution level. Harten’s multiresolution strategy is based on two discrete interlevel operators, called "projection" $D^{j+1}$ and "prediction" $P^{j-1}$ operators. The pixel value $v^j_k$, located at position $k = (k_1, k_2)$ and resolution level $j$, is defined by the cell average of the underlying bidimensional function $v$ modeling the image, as follows:

$$v^j_k = 2^{2j} \int_{C^j_k} v(x,y)dx\,dy,$$  

where $C^j_k$ is the cell defined as $[2^{-j}k_1, 2^{-j}(k_1 + 1)] \times [2^{-j}k_2, 2^{-j}(k_2 + 1)]$ with $0 \leq k_1, k_2 \leq 2^j - 1$.

The projection operator $D^{j+1}$, acting from the fine to coarse level, computes the dataset $v^{j+1} = D^{j+1}v^j$ corresponding to the coarser version of $v$. In the cell average context, the operator $D^{j+1}$ is the averaging operator and is always linear:

$$v^{j+1}_{lk} = \frac{1}{4}(v^j_{2k} + v^j_{2k+e_1} + v^j_{2k+e_2} + v^j_{2k+e_1+e_2}),$$

where $e_1$ and $e_2$ are unit vectors oriented to the right and upward, respectively.

The prediction operator $P^{j-1}$ acts from the coarse to fine levels by computing an approximation $\hat{v}^j$ of $v^j$ from $v^{j-1}$: $\hat{v}^j = P^{j-1}v^{j-1}$. This operator may be non-linear. Both prediction and projection operators satisfy the following consistency condition: $D^{j+1}P^{j-1} = I$, where $I$ is the identity operator. The null space of $D^{j+1}$ represents the detail space of dimension $N_j - N_{j-1}$. The prediction error $e^{j} = v^j - \hat{v}^j$ belongs to the detail space and then expanding it on a basis of that space getting the detail vector $(d^{j}_{k_1})_{k_1=1,...,N_j-N_{j-1}}$. Therefore, one can equivalently write $v^j$ in the form $(v^{j-1}, d^{j-1})$. Iterating this process, one obtains a multiresolution representation of the representation $v^j$ at the scale $J$ into $(v^J, d^J, d^J, ..., d^J)$.

3. NON-SEPARABLE MULTIRESOLUTION IMAGE TMO USING BI-QUADRATIC INTERPOLATIONS

This section deals with the proposed HDR image TMO given by the scheme of Figure 1. The first stage is based on the non-separable decomposition developed in [14], [15]. The prediction operator is related on the smoothness properties of the function $v$, considered as a bi-dimensional smooth function and represented by a piecewise polynomial function of degree 2, modeling the complex HDR image (i.e. smooth regions separated by edges). To adapt the prediction near the singularities, ENO interpolation techniques are then deployed.

3.1. Prediction step

Consider a shift vector defined as $r \in \{-e_1 - e_2, -e_1, -e_1 + e_2, -e_2, 0, e_2, e_1 - e_2, e_1 + e_2\}$. Define for a given cell $C^j_k$ and a shift $r$ the set of cells:

$$S^j(k, r) = \{C^j_{k+r} : l = (l_1, l_2), -1 \leq l_1, l_2 \leq 1 \text{ and } k = (k_1, k_2), -1 \leq k_1, k_2 \leq 1\},$$

called stencil of the cell $C^j_k$ associated to the shift $r$.

The stencil $S^j(k, r)$ has $3 \times 3$ cells and contains the cell $C^j_k$. An example of such stencil is the center stencil obtained for $k = 0$ and shift $r = 0$, i.e.:

$$S^j(0, 0) = \begin{bmatrix} C^j_{(-1,-1)} & C^j_{(-1,0)} & C^j_{(-1,+1)} \\ C^j_{(0,-1)} & C^j_{(0,0)} & C^j_{(0,+1)} \\ C^j_{(+1,-1)} & C^j_{(+1,0)} & C^j_{(+1,+1)} \end{bmatrix}.$$  

Therefore 9 stencils of size $3 \times 3$, containing the cell $C^j_k$ associated to the different shifts $r$, exist. To emulate the ENO strategy, the cost function $\zeta$ for the stencil $S^j(k, r)$ is defined as:

$$\zeta(S^j(k, r)) = \sum_{l \in S^j(k,r)} |\Delta v^j_l|,$$

where $\Delta v^j_l = v^j_{l+1} - v^j_l$ and $l = (l_1, l_2)$.

This cost function measures the sum of oscillations in $S^j(k, r)$. The least oscillatory stencil, for the cell $C^j_k$, is retained as follows:

$$r^* = \arg\min_{-M+1 \leq r_1, r_2 \leq M-1} \zeta(S^j(k, r)).$$

The least oscillatory stencil $S^j(k, r^*)$ is then retained. The predicted values $\hat{v}^j$ are deduced from the values at the coarse level $v^j_{2k}$ on $S^j(k, r^*)$. For this, a bi-quadratic polynomial $p_{k, r^*}$, defined by its 9 unknown coefficients $p_{m,n}$, is selected:

$$p(x, y) = \sum_{0 \leq m,n \leq 2} p_{mn} x^m y^n,$$

to predict the values as follows:

$$\hat{v}^j_{2k+w,r} = 2^{2j} \int_{P^j} p_{k, r^*} dxdy,$$
Forward non-separable multiresolution for a given level $j$

![Diagram showing forward non-separable multiresolution](image)

Adaptive weighting

![Diagram showing adaptive weighting](image)

3.2. Detail step

The prediction errors, with respect to $v^j$, are by construction:

$$e^j_w = v^j_w - ar{v}^j_w \quad \text{with} \quad w \in \{0, e_1, e_2, e_1 + e_2\}. \tag{10}$$

Due to the consistency property $D_{j-1}^{-1}P^{j-1}_i = I$, we get that $\sum_{w \in \{0, e_1, e_2, e_1 + e_2\}} e^j_w = 0$. These errors can be written in a non-redundant way by representing them in a basis of of $KerD_{j-1}^{-1}$. The matrix of this change of basis $H$ has the columns $h_{lo} = [\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}], \quad h_{h} = [\frac{1}{4}, -\frac{1}{4}, \frac{1}{4}, -\frac{1}{4}], \quad h_{v} = [\frac{1}{4}, \frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}], \quad$ and $h_{d} = [\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, \frac{1}{4}]$. Therefore the details vectors are deduced from $e_{2k}^j, e_{2k+1}^j, e_{2k+2}^j, e_{2k+3}^j$ as follows:

$$d_{h}^{j-1} = \{h_{h} \times [e_{2k}^j, e_{2k+1}^j, e_{2k+2}^j, e_{2k+3}^j]^t\}, \tag{11}$$

for horizontal detail block;

$$d_{v}^{j-1} = \{h_{v} \times [e_{2k}^j, e_{2k+1}^j, e_{2k+2}^j, e_{2k+3}^j]^t\}, \tag{12}$$

for vertical detail block; and

$$d_{d}^{j-1} = \{h_{d} \times [e_{2k}^j, e_{2k+1}^j, e_{2k+2}^j, e_{2k+3}^j]^t\}, \tag{13}$$

for diagonal detail block.

Note that the size of the block is half of each direction of $v^j$. The multiresolution representation is then obtained $v^j := (v^{j-1}, (d_{h}^{j-1}, d_{v}^{j-1}, d_{d}^{j-1}))$. Iterating the process until resolution level $J$ involves a set of approximation and detail coefficients : $(v^0, (d_{h}^0, d_{v}^0, d_{d}^0), (d_{h}^1, d_{v}^1, d_{d}^1), ..., (d_{h}^{j-1}, d_{v}^{j-1}, d_{d}^{j-1}))$.

3.3. Adaptive weighting coefficients strategy

This section proposes to independently weight each subband’s coefficient to reduce in an appropriate way the HDR image dynamic range.

The proposed weight is a positive value between 0 and 1 depending on the statistical properties of each subband and the neighborhood. The main operations are presented below.

First, the absolute value of all subband coefficients is taken : $A_{h}^0 = |d_{h}^0|, A_{h}^1 = |d_{h}^1|, A_{v}^j = |d_{v}^j|, A_{d}^j = |d_{d}^j|$ for $j = 0, ..., J - 1$.

Denote $A_{sb}^j(x, y)$, with $sb = 0, d, h, v$ and $j = 0, ..., J - 1$, the positive coefficient at position $(x, y)$ of the subband $sb$ at resolution level $j$. Define $g(x, y)$ the Gaussian filter given by :

$$g(x, y) = \frac{1}{2\pi\sigma_{sb}^2} e^{-\frac{x^2 + y^2}{2\sigma_{sb}^2}}, \tag{14}$$

where $\sigma_{sb}$ is the standard deviation of the Gaussian distribution of the subband $sb$ deduced from the median component as in [14] :

$$\sigma_{sb} = \frac{\text{median}(A_{sb}^j)}{0.6745}. \tag{15}$$
This filter is used to compute the weighting coefficient \( a^j_{sb}(x, y) \) associated to each coefficient \( A^j_{sb}(x, y) \) in the corresponding domain as follows:

\[
a^j_{sb}(x, y) = A^j_{sb}(x, y) \ast g(x, y),
\]

where * is the convolutional product.

After this, a monotonic nonlinear decreasing function \( F \) (as in [11]) is applied on the coefficients \( a^j_{sb} \), resulting in the positive weighting coefficient \( q^j_{sb}(x, y) \) between 0 and 1, as follows:

\[
q^j_{sb}(x, y) = \frac{1}{\left(\frac{a^j_{sb}(x, y) + \varepsilon}{\delta_j}\right)^{1-\gamma}},
\]

where \( \gamma \in [0, 1] \) is the compressive factor; \( j \) the resolution level (i.e. \( j = 0, ..., J - 1 \)) associated to the subband \( sb \). 

\( \varepsilon \) is used to avoid singularities when the power \( 1 - \gamma \) is a positive value. \( \delta_j \) is considered as a gain control stability level depending on the resolution level \( j \) and the subband coefficients:

\[
\delta_j = [1 - \frac{j(1 - \xi)}{J}] \times \bar{A}^j_{sb},
\]

where \( \bar{A}^j_{sb} \) is the average of \( A^j_{sb} \).

Finally, the subband coefficients are weighted as follows:

\[
v^0 = q^0_{sb} \otimes v^0 \quad \text{and} \quad d^j_{sb} = q^j_{sb} \otimes d^j_{sb},
\]

with \( sb = d, h, v, j = 0, ..., J - 1 \); and \( \otimes \) is the product term by term.

\section{4. SIMULATION RESULTS}

This section evaluates the performance of the proposed TMO using the TMQI (Tone-Mapped image Quality Index) metric developed in [16]. It evaluates the quality of the HDR image tone mapped compared to the original HDR image. Note that this metric is upper-bounded by 1.

Simulations have been conducted under Matlab environment using the HDR Toolbox with its HDR test images : "Anturium", "Bottle Small" renaming "Bottle", "Office", "Oxford Church", "Memorial" and "Light" [1]. The different parameters are chosen so as to give the best results in terms of TMQI metric in all methods. These latter are performed on the logarithmic transformation of the Luma values \( L_w \) (i.e \( \log_{10}(L_w + \varepsilon) \), \( \varepsilon \) avoiding the logarithmic singularities).

Simulation results are compared to competitive methods : (i) SEP [10] with \( \beta = 0.3, \gamma = 0.7, J=2, upd=0, b = e \); (ii) Li TMO [11] with Haar multiscale ; (iii) Fattal [9] using RBW method with \( \alpha = 0.8, \beta = 0.3, \gamma = 0.8, J=2, upd=0 \); (iv) Dragi [2], Reinhard [3], Ward [4], Durand [5], Tumblin [6], Schlick [7] with the default parameters as given in the HDR Toolbox ; and (v) Duan [8] using \( \beta = 0.5 \). For the proposed method, denoted "Proposed", simulations are conducted with the following parameters : \( \varepsilon = 10^{-9}, J = 2, \gamma = 0.6, \xi = 0.1, \sigma_i \) deduced from equation (15).

Figure 2 clearly shows the impact of the second stage on the visual quality of the LDR image.

To further improve the visual quality of some HDR tone mapped images, it is possible to add the histogram adjustment, developed in [8], as a post-processing. This post-processing can be applied on the HDR tone mapped image in a local way on equal non-overlapping blocks of size \( 32 \times 32 \) or in a global way. Indeed this approach adjusts the contrast between linear mapping and histogram equalized mapping thanks to a control parameter \( \beta \in [0, 1] \) (we set \( \beta = 0.3 \)). The two approaches are named "Proposed\_G" for global and "Proposed\_L" for local post-processing.

Table 1 provides the TMQI of the HDR tone mapped images according to the above listed methods. One can observe that our approach is competitive to those developed in the literature.

Figure 3 shows the "Bottle" HDR tone mapped images according to the "Proposed", "Proposed\_G" and "Proposed\_L" approaches. Figure 4 presents the "Bottle" HDR tone mapped images using [8] with local ("Duan\_L") and global ("Duan\_G") methods. The visual rendering of our tone mapped images is better than those developed in [8], even when the post-processing is not used. This is observed even if the TMQI is higher for "Duan\_G" (0.91) and "Duan\_L" (0.92) than our "Proposed" (0.87) (see Table 1) where "Proposed\_L" TMQI remains high (0.93). Note that the TMQI metric does not always reflect the visual quality.

Figures 5, 6, 7 and 8 provide the visual quality of the "Memorial" HDR tone mapped images. One can observe that the details are well rendered in favor of our method (e.g. tiles, rosette). In Figure 6, the tiles on the top right on the tone mapped image are better rendered in favor of our method.

Based on these simulation results, the proposed operator has the ability to better represent the details compared to other operators. Its strength points are inherent not
only to the non-separable multiresolution decomposition but also to the ENO strategy. It suffices to compare its performance with that of separable multiresolution ENO strategy developed in [10].

<table>
<thead>
<tr>
<th>TMOs</th>
<th>Anturium</th>
<th>Bottle</th>
<th>Office</th>
<th>Church</th>
<th>Memorial</th>
<th>Light</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drago[2]</td>
<td>0.874</td>
<td>0.801</td>
<td>0.801</td>
<td>0.814</td>
<td>0.800</td>
<td>0.800</td>
</tr>
<tr>
<td>Reinhard[3]</td>
<td>0.778</td>
<td>0.807</td>
<td>0.826</td>
<td>0.789</td>
<td>0.792</td>
<td>0.794</td>
</tr>
<tr>
<td>Ward[4]</td>
<td>0.806</td>
<td>0.783</td>
<td>0.775</td>
<td>0.817</td>
<td>0.793</td>
<td>0.789</td>
</tr>
<tr>
<td>Durand[5]</td>
<td>0.811</td>
<td>0.892</td>
<td>0.825</td>
<td>0.929</td>
<td>0.814</td>
<td>0.800</td>
</tr>
<tr>
<td>Tumblin[6]</td>
<td>0.715</td>
<td>0.713</td>
<td>0.735</td>
<td>0.675</td>
<td>0.761</td>
<td>0.750</td>
</tr>
<tr>
<td>Schlick[7]</td>
<td>0.770</td>
<td>0.835</td>
<td>0.926</td>
<td>0.970</td>
<td>0.788</td>
<td>0.780</td>
</tr>
<tr>
<td>Duan_G[8]</td>
<td>0.964</td>
<td>0.916</td>
<td>0.955</td>
<td>0.986</td>
<td>0.935</td>
<td>0.969</td>
</tr>
<tr>
<td>Duan_L[8]</td>
<td>0.910</td>
<td>0.924</td>
<td>0.927</td>
<td>0.906</td>
<td>0.839</td>
<td>0.924</td>
</tr>
<tr>
<td>Fattal[9]</td>
<td>0.889</td>
<td>0.928</td>
<td>0.928</td>
<td>0.889</td>
<td>0.927</td>
<td>0.971</td>
</tr>
<tr>
<td>SEP[10]</td>
<td>0.888</td>
<td>0.931</td>
<td>0.928</td>
<td>0.885</td>
<td>0.903</td>
<td>0.960</td>
</tr>
<tr>
<td>Li[11]</td>
<td>0.964</td>
<td>0.954</td>
<td>0.854</td>
<td>0.877</td>
<td>0.834</td>
<td>0.888</td>
</tr>
<tr>
<td>Proposed</td>
<td>0.938</td>
<td>0.873</td>
<td>0.948</td>
<td>0.820</td>
<td>0.846</td>
<td>0.932</td>
</tr>
<tr>
<td>Proposed_G</td>
<td>0.975</td>
<td>0.927</td>
<td>0.828</td>
<td>0.968</td>
<td>0.896</td>
<td>0.946</td>
</tr>
<tr>
<td>Proposed_L</td>
<td>0.920</td>
<td>0.930</td>
<td>0.861</td>
<td>0.939</td>
<td>0.881</td>
<td>0.909</td>
</tr>
</tbody>
</table>

Table 1. Tone Mapped Image Quality Index (TMQI)

5. CONCLUSION

This paper focused on HDR image tone mapping operator. As a first stage, the operator decomposes the HDR image into several resolution levels according to non-separable orthogonal transform using ENO bi-quadratic interpolation. This latter has the advantage to take into its mathematical model the possible singularities in all directions. Therefore the details are well represented in the subbands. As a second stage, a weighting strategy adapted to each subband's coefficient is developed in order to reduce adequately the dynamic range of the HDR image while preserving the more important details in the HDR image. Simulation results show good performance achieved not only in terms of TMQI metrics but also in terms of visual quality.

REFERENCES


Fig. 3. (Up) Proposed (ε = 10^-9; J = 2; γ = 0.6; ξ = 0.1); (Middle) Proposed with global post-processing; (Down) Proposed with local post-processing.

Fig. 4. (Up) Duan with global processing; (Down) Duan with local processing.
Fig. 5. (Left) Duan ($\beta = 0.5$, local); (Right) Proposed with local post-processing.

Fig. 6. (Left) Fattal (RBW, $\alpha = 0.8$, $\beta = 0.3$, $\gamma = 0.8$, $J=2$, $\text{upd}=0$); (Right) Proposed with local post-processing.

Fig. 7. (Left) Li (Haar multiscale) with post-processing; (Right) Proposed with local post-processing.

Fig. 8. (Left) SEP ($\beta = 0.3$, $\gamma = 0.7$, $J=2$, $\text{upd}=0$, $b = e$); (Right) Proposed with local post-processing.


