HIGH DYNAMIC RANGE IMAGE PROCESSING USING MANIFOLD-BASED ORDERING

Olivier Lézoray

To cite this version:
Olivier Lézoray. HIGH DYNAMIC RANGE IMAGE PROCESSING USING MANIFOLD-BASED ORDERING. International Conference on Pattern Recognition, Dec 2016, Cancun, Mexico. hal-01411026

HAL Id: hal-01411026
https://hal.archives-ouvertes.fr/hal-01411026
Submitted on 6 Dec 2016

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
Very few research works have addressed the problem of directly manipulating raw HDR vectors for general HDR image processing. In this paper a framework is proposed towards this aim and is based on a new representation of HDR images in the form of an ordering of vectors and an index image. This enables to formulate vector-preserving image processing methods dedicated to HDR images. The ordering relies on three steps: dictionary learning, manifold learning, and out of sample extension. The performance of the proposed approach is illustrated with innovative examples of HDR image filtering and enhancement.

I. INTRODUCTION
When a scene with high light contrast is captured by a camera, either the dark or the bright areas will be saturated in the output image. It is therefore difficult to simultaneously get details in shadows and highlights. This is due to the physical limitations on the maximal variations of luminosity that a camera sensor can capture and above a given quantity of incident photons, the sensor saturates. Since conventional cameras cannot capture the full dynamic range of a real world scene, defined as the range between the largest and the smallest light intensities in the scene, high dynamic range (HDR) imaging has been proposed to address this problem [1].

The principle consists in mixing several LDR (Low Dynamic Range) images of the scene taken under different exposure times with a conventional camera (see Figure 1). Today many camera devices have built in functionality for acquiring HDR images. Therefore, high dynamic range imaging has become more and more popular in recent years. It attracts many professional and amateur photographers and finds utility in many applications such as visual effects production.

Unfortunately, an HDR image cannot be displayed directly on conventional LDR display devices due to hardware limitations similar to those of LDR cameras. Thus, many research works have been conducted on how to compress an HDR image into an LDR one that can be displayed on conventional display devices. This problem is referred to as HDR tone mapping and global or local operators have been proposed [2], [3], [4], [5].

However the processing of HDR images does not reduce only to tone mapping, and one could be interested in processing raw HDR images (and not the tone-mapped LDR ones) to filter, simplify, or enhance them. Most algorithms (including tone-mapping) for HDR images operate on the logarithm of HDR pixel luminance values rather than on the original HDR pixel values. This is an easy way to adapt classical LDR image processing algorithms to HDR images and the logarithm domain is related to how the Human Visual System is sensitive to light. Nevertheless, the manifold where HDR pixel live is obviously much more nonlinear than a simple projection onto a log-luminance curve. Therefore, we propose another approach that consists in learning the manifold of HDR image pixels to build a new HDR image representation in the form of an index image associated with an ordering of the HDR pixels’ vectors.

To perform this, we embark from our earlier works as reported in [6] and apply a similar scheme to obtain an ordering of HDR vectors. Then, we demonstrate how the proposed ordering leads to a new representation of HDR images and how it can be used to perform classical image processing tasks such as filtering, simplification an enhancement. We focus in this paper on discrete algebraic methods (morphological methods and median processing) because they cannot introduce new vectors in a processed image and therefore they are not subject to the production of halos, which is a common problem when processing HDR images.

The paper is organized as follows. In Section II we introduce a learned ordering of the HDR vectors of an image. In Section III we explain how a new representation can be derived from this ordering and used for discrete vector-preserving HDR image processing. Section IV presents how tone-mapping can be performed with the proposed frame-
work, and demonstrates its effectiveness with respect to the state-of-the-art.

II. LEARNING AN ORDERING OF HDR VECTORS

II-A. Notations

We consider HDR images encoded with the Open EXR standard that associates three 16-bits floating point values to pixels, after merging several LDR images. An HDR image is represented by the mapping \( f : \Omega \subset \mathbb{Z}^2 \rightarrow \mathcal{T} \subset \mathbb{R}^3 \) where \( \mathcal{T} \) is a non-empty set of image multivariate HDR vectors (RGB values of linear radiance). To each pixel \( \mathbf{p}_i \in \Omega \) of an image is associated a vector \( \mathbf{v}_i = f(\mathbf{p}_i) \). The set \( \mathcal{T} = \{ \mathbf{v}_1, \ldots, \mathbf{v}_m \} \) denotes all the HDR vectors associated to all pixels in the image, i.e., the vectorization of the matrix \( \mathbf{F} \in \mathbb{R}^{N_1 \times N_2} \), with \( N_1 \times N_2 = m \) that represents an image \( f \). We will use the notation \( \mathcal{T}[i] = \mathbf{v}_i \) to denote the \( i \)-th element of a set.

II-B. Complete lattice

To process HDR images, we build a new HDR image representation in the form of an index image associated with an ordering of the HDR pixels’ vectors \( \mathcal{T} \) of an image. Ordering all the values of the set \( \mathcal{T} \) can be done with the use of an ordering relation within HDR vectors. This amounts to dispose of a complete lattice \( (\mathcal{T}, \leq) \), a key item for the definition of mathematical morphology operators [7]. Unfortunately there is no universal method for ordering vectorial data [8]. Many works have proposed specific orderings for color vectors, mainly with the use of lexicographic orderings [9], [10], but they cannot be used directly for HDR vectors since they use specific assumptions on the ordering of channels. Other recent works consider space-filling curves in the image domain for image denoising using patches [11], [12], but they are not suited for HDR images. Our proposal is to build on the ideas we presented in [6].

Indeed, the framework of \( h \)-orderings [13] can be used to define an ordering relation between HDR vectors of a set \( \mathcal{T} \). The latter corresponds to constructing a surjective mapping \( h \) from \( \mathcal{T} \) to \( \mathcal{L} \) where \( \mathcal{L} \) is a complete lattice equipped with the conditional total ordering [13]. We refers to \( \leq_h \) as the \( h \)-ordering given by:

\[
h : \mathcal{T} \rightarrow \mathcal{L} \text{ and } \mathbf{v} \rightarrow h(\mathbf{v}), \forall (\mathbf{v}_i, \mathbf{v}_j) \in \mathcal{T} \times \mathcal{T} \quad \mathbf{v}_i \leq_h \mathbf{v}_j \iff h(\mathbf{v}_i) \leq h(\mathbf{v}_j) .
\]

Then, by means of \( h \) [14], the ordering of \( \mathcal{T} \) can be induced upon \( \mathcal{L} \) and \( \mathcal{T} \) is no longer required to be a complete lattice. When \( h \) is bijective, this corresponds to defining a space filling curve [15] or equivalently a rank transform [16]. It is obvious that the projection \( h \) cannot be linear [17] since a distortion of the space topology is inevitable. As a consequence, we choose to focus our developments on learning the HDR manifold to construct \( h \) and deduce the complete lattice \( (\mathcal{T}, \leq_h) \).

II-C. Manifold-based ordering

Our approach consists in learning the manifold of HDR vectors from an image with a nonlinear mapping and to define the ordering from this projection. To learn the manifold, we use Laplacian EigenMaps (LE), a technique for non-linear dimensionality reduction [18]. Computationally, performing LE on the whole space of HDR vectors is not tractable in reasonable time (especially for large images). So, an efficient strategy is needed to construct a \( h \)-ordering. We summarize its principle in the sequel.

Given an HDR image that provides a set \( \mathcal{T} \) of \( m \) vectors in \( \mathbb{R}^3 \), a dictionary \( \mathcal{D} = \{ \mathbf{x}_1, \ldots, \mathbf{x}_p \} \) of \( p \ll m \) vectors is build by Vector Quantization [19].

Manifold learning by Laplacian EigenMaps is performed on this dictionary. One starts by computing a similarity matrix \( \Phi \) that contains the pairwise similarities

\[
K_D(i, j) = \exp \left( -\frac{||\mathbf{x}_i - \mathbf{x}_j||^2}{\sigma^2} \right)
\]

between all the dictionary vectors \( \mathbf{x}_i \). To have a parameter-free algorithm, we consider \( \sigma = \max_{\mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}} ||\mathbf{x}_i - \mathbf{x}_j||^2 \). The normalized Laplacian matrix \( \mathbf{L} = \mathbf{I} - \mathbf{D}^{-\frac{1}{2}} \mathbf{K}_D \mathbf{D}^{-\frac{1}{2}} \) is then computed with \( \mathbf{D} \) the degree diagonal matrix of \( \mathbf{K}_D \). Then, Laplacian Eigenmaps Manifold Learning consists in searching for a new representation \( \Phi \) obtained by minimizing

\[
\frac{1}{2} \sum_{i,j} ||\Phi(\mathbf{x}_i') - \Phi(\mathbf{x}_j')||_2^2 K_D(i, j) = Tr(\Phi^T \mathbf{L} \Phi)
\]

under the constraint \( \Phi^T \mathbf{D} \Phi = \mathbf{I} \). This cost function encourages close input vectors to be mapped to close output vectors. The solution is obtained [20] by finding the eigenvectors \( \Phi_D \) of \( \mathbf{L} \). Therefore, the decomposition \( \mathbf{L} = \Phi_D \Phi_D^T \mathbf{L} \Phi_D^T \mathbf{L} \) is computed with corresponding eigenvectors \( \Phi_D = [\Phi_1, \ldots, \Phi_p] \) and eigenvalues \( \lambda_D = \text{diag}[\lambda_1, \ldots, \lambda_p] \).

Then, a new representation \( h_D(D) \) is obtained for each element \( x_i' \) of the dictionary \( D \):

\[
h_D : x_i' \rightarrow (\phi_D(x_1'), \ldots, \phi_D(x_p'))^T \in \mathbb{R}^p .
\]

where \( \phi_D(x_i') \) denotes the \( i \)-th coordinate of the \( h \)-th eigenvector. Such a strategy of modeling the manifold from a patch dictionary was also explored in [17]. The projection operator \( h_D \) corresponds to constructing an \( h_D \)-ordering from the sole data of the dictionary \( D \). This correspond to the construction of the complete lattice \( (\mathcal{D}_D, \leq_{h_D}) \) with a \( h_D \)-ordering, and this ordering is only valid for the set of vectors of the dictionary. Since we need the complete lattice \( (\mathcal{T}, \leq_h) \), the reduced dictionary lattice is extended to all the vectors of the initial lattice \( \mathcal{T} \) by Nyström extrapolation [21] of \( h_D \) on \( \mathcal{T} \). To do so, we compute the similarity matrix \( \mathbf{K}_D \) between sets \( \mathcal{T} \) and \( \mathcal{T} \) and the associated degree diagonal matrix \( \mathbf{D}_D \). The extrapolated eigenvectors are then obtained by

\[
\tilde{\Phi} = \mathbf{D}_D^{-\frac{1}{2}} \mathbf{K}_D \mathbf{D}_D^{-\frac{1}{2}} \Phi_D (\text{diag}[\mathbf{I}] - \mathbf{Pi}_D)^{-1} .
\]
Finally, the projection operator \( h : \mathcal{T} \subset \mathbb{R}^3 \rightarrow \mathcal{L} \subset \mathbb{R}^p \) on the manifold is defined as \( h(\mathbf{x}) = (\Phi^1(\mathbf{x}), \ldots, \Phi^p(\mathbf{x}))^T \), and the complete lattice \((\mathcal{T}, \leq_h)\) is obtained by using the conditional ordering on this new representation.

III. HDR IMAGE REPRESENTATION AND DERIVED PROCESSING

III-A. HDR Image Representation

Once the complete lattice \((\mathcal{T}, \leq_h)\) is available, a new HDR image representation can be defined. Let \( \mathcal{P} \) be a sorted permutation of the elements of \( \mathcal{T} \) according to the manifold-based ordering \( \leq_h \), one has:

\[
\mathcal{P} = \{ \mathbf{v}_1, \ldots, \mathbf{v}_m \} \quad \text{with} \quad \mathbf{v}_i \leq_h \mathbf{v}_{i+1}, \forall i \in [1, (m - 1)].
\]

This can also be written as \( \mathcal{P} = \mathcal{P}^{\mathcal{T}} \) with \( \mathcal{P} \) a permutation matrix of size \( m \times m \). From this ordered set of HDR vectors, an index image can be defined. Let \( I : \Omega \subset \mathbb{Z}^2 \rightarrow [1, m] \) denote this index image. Its elements are defined as:

\[
I(\mathbf{p}_i) = \{ k \mid \mathbf{v}_k = f(\mathbf{p}_i) = \mathbf{v}_i \}.
\]

Therefore at each pixel \( \mathbf{p}_i \) of the index image \( I \), one obtains the rank of the original HDR vector \( f(\mathbf{p}_i) \) in \( \mathcal{P} \), the set of sorted HDR vectors. Given \((I, \mathcal{P})\), a new representation of the original HDR image \( f \) is obtained. When an HDR image is encoded in this way, HDR information is not directly used for any image processing. However, we have to be able to reconstruct a processed HDR image from a processed index image, and this necessitates that each value of the processed index image corresponds to one vector of the manifold-based ordering \( \mathcal{P} \). This means that a processed index image must have its values that are kept within the integer range \([1, m]\). In other words, this means that we can only consider vector-preserving image processing methods. Therefore, given a specific processing \( g \), the corresponding reconstructed processed HDR image is obtained by

\[
g(f(\mathbf{p}_i)) = \mathcal{P}[g(I(\mathbf{p}_i))].
\]

Typical vector preserving methods are morphological and median-based ones, we consider these ones in this paper. We will mainly focus on algebraic (and vector-preserving) morphological methods, so we define the two main operators, erosion and dilation of an HDR image \( f \) at pixel \( \mathbf{p}_i \in \Omega \) by a structuring element \( B \subset \Omega \) as:

\[
\epsilon_B(f)(\mathbf{p}_i) = \{ \mathcal{P}[\wedge I(\mathbf{p}_j)], \mathbf{p}_j \in B(\mathbf{p}_i) \}
\]

\[
\delta_B(f)(\mathbf{p}_i) = \{ \mathcal{P}[\vee I(\mathbf{p}_j)], \mathbf{p}_j \in B(\mathbf{p}_i) \}
\]

One sees that these morphological operators do work on the index image \( I \), and the processed HDR image is reconstructed through the sorted vectors \( \mathcal{P} \) that represent the learned complete lattice. Given their formulation, the proposed operators do inherit the standard algebraic properties of morphological operators [22]. From these basic operators, we can obtain all standard morphological filters for HDR images such as openings \( \gamma_B \) and closings \( \delta_B \).

Figure 3 illustrates the benefits of the proposed approach on two HDR images with three different morphological processings: erosion, dilation and opening \( \gamma_B \). Since HDR images cannot be directly shown, the result is presented after tone-mapping using the reference method of [3]. As illustrated, with the proposed representation, we recover the classical aspects of morphological operators: erosion extends structures close to first vector of the ordering. Dilation provides the dual effect and extends structures close to last.
vector of the ordering. With the opening, details smaller than the structuring element, and close to the first vector of the ordering, are removed. Here a very large structuring element has been chosen to better illustrate the effects of the processing. To the best of our knowledge, this is the first time that such processing is presented for HDR images. As

\[
\kappa_{h,B}(f)(p_i) = \begin{cases} 
\delta_B(f)(p_i) & \text{if } \Delta_{1B}^1(f)(p_i) \leq \Delta_{2B}^2(f)(p_i) \\
\varepsilon_B(f)(p_i) & \text{if } \Delta_{1B}^1(f)(p_i) > \Delta_{2B}^2(f)(p_i)
\end{cases}
\]

(11)

with

\[
\Delta_{1B}^1(f)(p_i) = \|f(p_i) - \delta_B(f)(p_i)\|_2 \\
\Delta_{2B}^2(f)(p_i) = \|f(p_i) - \varepsilon_B(f)(p_i)\|_2
\]

(12)

Usually, for detail manipulation, edge-aware transforms are considered [24]. We adopt the strategy of [23] that consists in decomposing an image into a base layer and a detail layer. We apply only one level of decomposition and replace their decomposition filter by a morphological Open Close Close Open (OCCO) filter defined as pixelwise average of open-close and close-open [9]:

\[
OCCO_B(f) = \frac{\gamma_B(\phi_B(f)) + \phi_B(\gamma_B(f))}{2}
\]

(13)

The structuring element is a square of side 5 pixels. The detail layer is boosted with a factor 1.5 and the image is recomposed without the presence of halos since the decomposition is a vector-preserving one. This shows that our framework can also be advantageously used for such edge-aware image detail manipulation, and this has never been explored before with raw HDR images.

**IV. TONE MAPPING**

We have seen in the previous section that with the new HDR image representation we proposed, it is now possible to apply morphological filters for e.g., obtaining a coarse version of raw HDR images. This was considered for detail manipulation. A morphological contrast mapping is iteratively applied ten times with a square of side 5 pixels. This enhances the local contrast of \( f \) by sharpening its edges with the following transformation, similarly to a shock filter:

\[
\kappa_{h,B}(f)(p_i) = \begin{cases} 
\delta_B(f)(p_i) & \text{if } \Delta_{1B}^1(f)(p_i) \leq \Delta_{2B}^2(f)(p_i) \\
\varepsilon_B(f)(p_i) & \text{if } \Delta_{1B}^1(f)(p_i) > \Delta_{2B}^2(f)(p_i)
\end{cases}
\]

(11)

with

\[
\Delta_{1B}^1(f)(p_i) = \|f(p_i) - \delta_B(f)(p_i)\|_2 \\
\Delta_{2B}^2(f)(p_i) = \|f(p_i) - \varepsilon_B(f)(p_i)\|_2
\]

(12)

Usually, for detail manipulation, edge-aware transforms are considered [24]. We adopt the strategy of [23] that consists in decomposing an image into a base layer and a detail layer. We apply only one level of decomposition and replace their decomposition filter by a morphological Open Close Close Open (OCCO) filter defined as pixelwise average of open-close and close-open [9]:

\[
OCCO_B(f) = \frac{\gamma_B(\phi_B(f)) + \phi_B(\gamma_B(f))}{2}
\]

(13)

The structuring element is a square of side 5 pixels. The detail layer is boosted with a factor 1.5 and the image is recomposed without the presence of halos since the decomposition is a vector-preserving one. This shows that our framework can also be advantageously used for such edge-aware image detail manipulation, and this has never been explored before with raw HDR images.
manipulation. However, image decomposition into both base and detail layers is at the core of state-of-the-art tone-mapping algorithms. So far, we have presented the final processed HDR images with the tone-mapping approach of Durand and Dorsey [3], but we can make the most of our proposal to dispose of a new morphological tone-mapping operator. First we recall the principle of the approach presented in [3].

This method computes the log-luminance \( L_i = \log(L(f)) \) from an HDR image \( f \) with

\[
L(f) = \frac{20 f_R + 40 f_G + f_B}{61}
\]  

(14)

where \( f_* \) denotes a given channel (subscript \( * \)) of the HDR image. The log-luminance \( L_i \) is then separated into a base layer \( B_i \), obtained by blurring the image, and a detail layer \( D_i = L_i - B_i \) which is the difference between the log-luminance and the base layer. The base layer, assumed to contain most of the high contrast that needs to be reduced, is then compressed using gamma correction and finally added to the unmodified detail layer. This is expressed by:

\[
f'_* = \left( \frac{f_*}{L(f)} \exp(\tau B_i + (L_i - B_i)) \right)^\gamma
\]

(15)

with \( f' \) the RGB tone mapped image, \( \tau \) a contrast parameter and \( \gamma \) a gamma correction parameter (fixed to 50 and 2.2 as recommended in [3]).

Using a simple low-pass filter (such as Gaussian filtering) for the layer separation would cause haloing effects: some high-contrast edges appear both into the base and the detail layer. To avoid this problem, the tone mapping operator of [3] uses Bilateral Filtering instead \( B_i = BF(L_i) \), that smoothes only regions with similar brightness while preserving edges.

In our case we can modify this tone-mapping method by replacing the bilateral filter with a morphological filter. Since our framework operates directly on the raw HDR image and not on the log-luminance, the image is filtered first and then its log-luminance is computed. Our proposed base layer is therefore \( B_i = \log(L(MF(f))) \) where \( MF(f) \) expresses a morphological filtering of the HDR image \( f \). The rest of the tone-mapping algorithm of [3] is kept the same. Figure 5 presents tone-mapping results on the HDR image of Figure 1. The morphological filter we have considered for base layer computation is a morphological Open Close Close Open \( MF(f) = OCCO_B(f) \) with \( B \) a square of side 5 pixels. Figure 5 presents a comparison between the state-of-the-art approach of [3] and ours. As it can be seen, some details are still present in the base layer with the bilateral filter (in the stained glass) and the edges are not always sharp (book and deck). This is not the case with our base layer decomposition which is much more piecewise constant. This results in a more contrasted and sharper tone-mapped image.

This is assessed by the values of the log-PSNR values (a quality metric for HDR images [25]).

**V. CONCLUSION**

This paper has detailed an approach for the processing of raw HDR images. To address this, a new HDR image representation has been proposed in the form of an ordering of vectors and an index image. The ordering of vectors is available online at http://driiqm.mpi-inf.mpg.de
interpreted as the construction of a complete lattice and a manifold-based approach is proposed towards this. To be efficient, a strategy based on dictionary learning, manifold learning and out of sample extension has been devised. From this new representation, vector-preserving image processing have been proposed. Morphological methods have then been studied in this paper for HDR image simplification, detail enhancement but also tone-mapping, and they are a competitive alternative to the actual state-of-the-art methods. Future works will consider non vector-preserving algorithms with the proposed representation.

ACKNOWLEDGMENT

This work received funding from the Agence Nationale de la Recherche, ANR-14-CE27-0001 GRAPHSIP.

VI. REFERENCES


