Non-Local Pairwise Energy Based Model For The HDR Image Compression Problem

Max Mignotte

Abstract—We present a new energy based compression model for the display of high dynamic range images. The proposed tone mapping method tends to exploit the biologically-inspired dynamic retina concept which is herein mathematically expressed via an image representation based on the specification of the statistical distributions of the non-local gradient magnitude. In this framework which also operates the notion of non-local gradient recently put forward by Gilboa and Osher, the detailpreserving contrast reduction problem is therefore expressed by a energy based model with non-local pairwise pixel interactions defined on a complete graph whose cost function is locally minimized by a conjugate gradient descent procedure. The experiments reported in this paper demonstrate that the proposed compression method is efficient and provides pleasing results on various images with different scene contents and performs competitively compared to the best existing state-of-the-art tone mapping methods recently proposed in the literature.

Index Terms—High dynamic range (HDR) image display, compression, tone mapping, contrast reduction, Gibbs energy based model, conjugate gradient minimization/optimization, non-local gradient distribution, dynamic retina concept.

I. INTRODUCTION

B Y their ability to directly store the amount of light measured by the camera and thus to acquire the whole dynamic range of radiance that can be observed in the realworld scenes (which is close to a contrast ratio with an upper bound of 10^5 : 1), High Dynamic Range (HDR) images are nowadays becoming increasingly popular and important in computer graphics (CG) and computer vision (CV) applications. HDR images, which can be easily constructed [1], [2], [3] by compiling/combining different standard photographs of the same scene with an increasing time range of exposure, have many advantages over standard Low Dynamic Range (LDR) images. They can more accurately represent the wide range of intensity levels found in real scenes ranging from direct sunlight to faint starlight or deep shadows, simultaneously preserving fine textural details of an object both in highlight and shadow areas: a characteristic which remains impossible to capture in a single photograph or with traditional LDR devices (monitors, projector, printers, etc.) displaying a traditional digital 8-bits image per color channel (i.e., a 24-bits per pixel

¹Max Mignotte is with the Département d'Informatique et de Recherche Opérationnelle (DIRO), Université de Montréal, C.P. 6128, Succ. Centre-ville, H3C 3J7, Montréal (Québec). E-MAIL: MIGNOTTE@IRO.UMONTREAL.CA

http://www.iro.umontreal.ca/~mignotte/

color image) thus using a contrast ratio of 256 : 1. This is why HDR images are often called *scene-referred*, in contrast to traditional digital images called *device-referred* or *output referred*, because they are specifically made for common LDR display devices.

In the past decade, there have been a number of research initiatives undertaken by the CG community concerned with displaying HDR images on LDR display devices (i.e., tone mapping methods). They can be classified into two broad groups (and an overview is given in [4]), namely global (spatially invariant) or local (spatially variant) mappings. DiCarlo et al. [5] refer to the former as tone reproduction curves (TRC) and to the latter as tone reproduction operators (TRO). Amongst the existing (simple and computational efficiency) TRC methods, which use the same mapping function for all pixels, we can cite the simple stretching and the histogram equalization techniques which both suffer from severe loss of contrast and visibility on the final LDR display. An interesting TRC model is the global histogram adjustment proposed by Ward Larson et al. [6] that can be viewed as an improvement of the classical histogram equalization technique. In their method, the empty portions of the histogram are cleverly disregarded in order to save more local contrast. This technique is nevertheless limited since it does not work efficiently if the luminance distribution is almost uniform. Another TRC model is the approach proposed in [7] in which the tone reproduction is achieved by histogram equalization of macro edges.

Spatially variant or local TROs are more interesting (and perhaps more appropriate), since they take local spatial context (i.e., a spatial neighborhood) into account thus trying to mimic the Human Visual System (HVS) that is mainly sensitive to local contrast. In our opinion, these TRO methods (some of them have been compared in terms of visual differences in [8]) can be classified into four main categories.

The first category is based on the interesting multi-scale image decomposition based technique initially proposed by Tumblin *et al.* in [9]. This approach consists of: decomposing an image into a piecewise smooth base layer or image *profile* (capturing large scale variations in intensity) and a residual detail layer (capturing the details in the image), strongly compressing the contrasts of this profile to fit it into the LDR display range and finally adding back all small details with little or no compression. To this end, Tumblin *et al.* use partial differential equations inspired by anisotropic diffusion in order to extract details from the image at various spatial scales. Durand *et al.* [10], Farbman *et al.* [11] and Xu *et al.* [12] exploit respectively an edge-preserving bilateral filter and the level set method in order to obtain the abovementioned image profile. Another multiscale decomposition based technique using symmetrical analysis-synthesis filter banks and automatic gain control has been proposed in [13].

The second category is based on a strategy combining a global and a local tone mapping technique. Examples of this strategy include the approach proposed by Reinhard *et al.* in [14] which tends to mimic the photographic process and whose local method is also based on a edge-preserving filter and the recent work of Shan *et al.* [15] that performs (in a final global optimization procedure) a set of local linear adjustments on small overlapping windows over the entire input image.

The third category is based on methods that physically attempt to model the local sensitivity-adjusting process of the HVS for high contrast scenes. Amongst these HVS-based models of local adaptation, Pattanaik [16] et al. have used a multiscale decomposition of the image according to comprehensive psycho-physically-derived filter banks and extensive psycho-physical data. The retinex model has also recently inspired several research projects [17], [18], including the multiscale center/surround retinex model of color perception of human vision proposed by Rahman et al. [19], [20] and the retinex-model-based adaptive filter where the luminance channel is defined by the first component of a principal component analysis [21], [22]. Other interesting approaches include the local-eye adaptation method [23] that compresses the dynamic range of the luminance channel by simulating the photoreceptor responses in the retina and the simple functional model of human luminance perception proposed by Ashikhmin [24] (whose method can also be categorized in the first TRO group since his tone mapping process tends to *infuse* details back into a scene), the iCAM image-appearance model [25] that has been extended to render HDR images for display or the Perceptually inspired TRO model proposed by Gatta et al. [26].

Finally, the fourth category can be considered as a *symbolic* compression technique. This group of techniques consists of the HDR image compression in a transformed space, representation or domain, in which the tone mapping problem is easier (or more rightly expressed) to solve. An image inverse transform leads back into the original (generally luminance) domain. We classify the excellent work of Fattal *et al.* [27] into this category since the authors use, as transformed space, the gradient domain and as inverse transform, the solution of a Poisson equation on the modified (i.e., compressed) gradient field.

Our work can be classified into the third and/or the fourth categories of the tone mapping method, i.e., one that attempts to model the local adaptation process of the HVS for high contrast scenes and one that tries to solve the problem in a transformed representation/space.

Let us now justify our model for the third category. First let us mention that constructing a local adaptation visual model is not easy due to our insufficient understanding of the HVS. Nevertheless, we know that the local adaptation of the HVS is based on two important characteristics described in the well-known Munsell experiments $[28]^2$. In addition to the classical non-linear compression ability of the HVS, the second characteristic is that this non-linear compression is done locally and pixel pairwise (or more precisely panel pairwise in the case of the Munsell experiments). In fact, our subjective perception of the luminance of a fine and subtle detail (both in a deep shadow and in highlighted areas) or the contour perception by the HVS can only be done comparatively, after multiple pixel pairwise comparisons. Biologically, this phenomena is explained by the fact that our eyes are never still, even during fixation³ [29], [30]. The presence of random vibrations (which follows a Gaussian distribution centered on the target [31]) in the human eye means that the pairwise comparison process is not only locally restricted to a first order neighborhood (i.e., the four nearest neighbors) as it would be in a computer estimating the contour magnitude with a local gradient by mathematically using a classical finite central difference formula. This leads to a straightforward model that can be easily adapted for HDR compression. This model consists of finding a mapping which replaces the value of all (local and spatially Gaussian distributed) pairwise pixel differences, existing in the image, close to their respective compressed difference value.

Let us now justify our model for the fourth category. The image model which is commonly used for digital image processing, CV or CG applications is generally based on the wellknown pixel (or picture element) notion. In this paper, and as previously noted, in order to more rightly express the HDR compression problem with the above-mentioned biologicallyinspired considerations, we have considered a new image representation in which the new fundamental and smallest representative element in the image is no longer the pixel but rather the pixel pair. In this transformed model, the compression problem is more rightly expressed (because somewhat biologically justified). After the compression process of the dynamic range using the transformed image model, our application resorts to a non-linear optimization problem in order to lead back into the original image representation domain using the classical pixel, as fundamental picture element.

The remainder of this paper is organized as follows : Section II describes the proposed energy based HDR compression model. Section III describes the optimization strategy used to minimize the objective cost function related to our model. Finally, Section IV presents the set of experimental results and comparisons with existing dynamic range reduction models.

²In this experiment, an observer is shown a set of n panels, of which the leftmost panel is black and the rightmost panel is white and is then invited to adjust the luminance of each of the in-between panels until he is satisfied that the panels form steps of equal-brightness (the brightness is our subjective perception of the luminous intensity). The observer will invariably set the panel luminances so that the ratio between each pair of panel luminances is equal, thus defining a logarithmic serie.

³The involuntary small vibrations of the eyes and the natural saccadic eye movements process (now well-known under the name of *eye tremor* or *microsaccades* under the *dynamic retina* concept) play this important role in luminance compression but also in fine detail and edge detection and robust (to noise) restoration process [29], [30], [31]. It is clearly established that without these microscopic movements the photo-receptors do not compress, saturate and the retinal images disappear [31].

II. PROPOSED COMPRESSION MODEL

A. A Pairwise Energy-Based Model

As already noted and previously justified (see Section I), given an HDR image with luminance X to be compressed, our tone mapping model consists simply of locally searching a new luminance mapping \hat{X} in which all the local pairwise pixel differences (between sites s and t), existing in the luminance image, are close to their respective compressed difference value ($\beta_{s,t}$). If this mapping is estimated in a minimal mean square sense, then \hat{X} is the solution image that minimizes the following objective function:

$$\hat{X} = \arg\min_{X} \sum_{s, t_{s \neq t}} g_{s,t} \left(\beta_{s,t}^2 - (X_s - X_t)^2 \right)^2$$
(1)

where X_s denotes the luminance mapping at site (or pixel) s and the summation $\sum_{s,t_{s\neq t}}$ is over all the pair of sites (i.e., for all sites s and for all the pair of sites including s) existing in X. As already said, the set of $\beta_{s,t}$ stores the set of compressed difference values related to each existing pairwise pixel difference. The $g_{s,t}$ positive weighting factor thus simulates the (spatial) decreasing (somewhat Gaussian) distribution effect of the pairwise comparison process by simply giving more importance for pairwise comparisons involving pixels that are close together. In our application, $g_{s,t} = d_{s,t}^{-1}$ where $d_{s,t}$ is the L_{∞} norm of the distance vector between sites s and t. This decreasing function acts as a degree of locality since it controls the number of considered pairwise pixels taken into account for each site and allowed us to not consider a complete graph but a square neighbourhood window of fixed size N_s pixels (i.e., by just considering the (N_s^2-1) nearest neighbors of each site s). In our application $\beta_{s,t}$ and $g_{s,t}$ are pre-computed off line (i.e., before the optimization process).

B. Compression Model

Let us underline that the HDR compression model is, in fact, completely expressed by the set of $\beta_{s,t}$ which will correspond, as close as possible (after the local optimization process), to the value of the pairwise pixel difference $|X_s - X_t|$ in the new luminance mapping \hat{X} to be recovered.

The common approach generally used in tone compression methods consists of the redistribution of the radiance values with a sigmoidal transformation (somewhat justified by electrophysiological studies which have found that the neurones in the vertebrate retina have an intensity-response functions with a sigmoid shape [32]). Equivalently in the gradient domain [27], tone compression techniques consist of the attenuation of large magnitude gradients (i.e., greater than the mean magnitude) along with the amplification of small magnitude gradients. These compression strategies usually involve considering an arbitrary value of the mean radiance or gradient magnitude (which will remain unchanged by the compression scheme) and above all, without considering the inherent statistical properties of the resulting output LDR image (compared to the statistical properties of any natural and real-world images). Indeed, if the statistical distribution of the luminance value of any real images may be very



Fig. 1. Repartition of the mean gradient magnitude (absolute value of the first order difference, averaged over the image) on the Berkeley image database. This histogram shows us what must be a reasonable value of $\sigma^2_{\nabla x}$ which is ultimately also the stopping criterion of our energy-based compression algorithm related to an image with a photo realistic look and displaying a lot of details.

different, the statistical distribution of the gradient magnitude follows a (well-known in the denoising community) longtail distribution⁴ mathematically expressed by a two-parameter density function of the form $P(k) \propto \exp(-|k/\lambda|^p)$ [33] where p and λ are respectively a shape and a scale factor.

A rapid study on the Berkeley real image database [34] shows that the shape factor p varies approximately between the interval [0.1:0.5], depending on the image content (exhibiting natural or man-made objects). The variance of the gradient magnitude (absolute value of the first order difference) varies between [0.0005:0.009], respectively ranging from very coarse to highly detailed image (with a nearly quasi flat gradient magnitude distribution) and a variance distribution shown in Fig. 1. These two properties remains true for the distribution of the absolute value of difference involving sites far away from more than one pixel (with the variance slightly increasing as the distance between sites increases).

In contrast to these natural LDR images belonging to the Berkeley database, the gradient magnitude of a HDR image exhibit approximately a very sharped distribution (which also belongs to the above-mentioned two-parameter density function but with a shape factor p >> 1) and therefore a variance of the gradient magnitude close to zero.

In order to take into account these statistical remarks, and the richness of details that any HDR image is able to render, our compression scheme will be designed to finally obtain an LDR output image with a nearly flat gradient magnitude distribution, whose flattening factor (which controls the richness of details desired in the output LDR image), is adjusted via a simple control parameter namely the previously mentioned variance value of the gradient magnitude. To this end, before the local optimization process, we set the $\beta_{s,t}$ as being the values given by a histogram specification method of the gradient magnitude with the uniform distribution (for capturing both the high contrast appearance of the scene and its

⁴This is due to the intrinsic stationary property of real-world images, containing smooth areas interspersed with occasional sharp transitions (i.e., edges). The smooth regions produce small amplitude gradient magnitude and the transitions produce sparse large-amplitude gradient magnitude [33].

small low-contrast details) and also by considering a classical 256 LDR of bin values (for the luminance and the gradient magnitude values, thus ensuring a usual contrast ratio for the output LDR image). In order to overcome the error due to the classical discrete implementation of this specification algorithm, we have used the recent histogram specification method described in [35] which is based on the definition of an ordering relation on the values to be histogram-specified.

More precisely, the set of $\beta_{s,t}$ values are computed using the method described in [35] in the following way. Let I be an HDR image with $N \times M$ (length \times width) pixels and let $W = 8 \cdot N \cdot M$ be the number of absolute values of the first order difference $|X_s - X_t|$ in the original HDR image and let also $H = \{h_0, h_1, ..., h_{Z-1}\}$ be the (*a priori* imposed) target non-normalized uniform distribution (i.e., $h_i = W/Z \forall 0 \le i < Z$) with Z = 256 bins. Let finally \prec be a strict ordering relation, defined among the $|X_s - X_t|$. The $\beta_{s,t}$ estimation then proceeds as follows

• Order the W = 8NM pairwise pixel absolute differences:

$$|X_s - X_t \prec |X_u - X_v| \prec \ldots \prec |X_x - X_y|$$

- Split this pixel absolute difference ordering relation from left to right in Z groups, such as group j has h_j elements, i.e., h_j couples of pixels.
- For all pair of pixels or pair of sites (s,t) whose the absolute difference is in a group j, assign β_{s,t} = j.

C. Discussion

To summarize, our local search (or minimization)-based compression scheme aims at flattening, *via* an iterative gradient descent procedure, the gradient magnitude distribution of an input HDR image, as long as the desired level of details in the output image is reached. This desired level of details of the output image content is controlled by an internal parameter which is simply the variance of the gradient magnitude. Fig. 1 shows that highly detailed images are related to variance parameters between [0.005 : 0.009]. This provides a good indication of a range of variance related to images displaying a lot of details.

There are three points worth mentioning concerning the local search-based HDR compression model expressed by Eq. (1).

- Since $g_{s,t}$ is a symmetric weight $(g_{s,t} = g_{t,s})$ and is defined here as a decreasing function of a positive measure defined between sites s and t, the objective function expressed by Eq. (1) thus exploits the interesting concept of non-local gradient (herein, $g_{s,t}^{1/4} (X_s X_t)$) recently put forward by Gilboa and Osher in [36] as a generalization of the gradient operator.
- Secondly, the objective functions to be locally minimized can be easily viewed as a Gibbs energy field related to a nonstationary (and non-local) Markovian model defined on a graph with possibly long-range pairwise interactions, i.e., binary cliques $\langle s, t \rangle$ (or pairwise of pixels). Each binary clique of this MRF model is associated to a nonstationary potential since this model

is spatially variant and depends on the distance between the sites s and t.

• Thirdly, it is possible to view this model as (somewhat) the generalization of the compression model proposed by Fattal *et al.* [27] in that it also uses the first order gradient domain to do the HDR compression task. Let us also mention that our method also consists of the amplification and attenuation of non-local magnitude gradients. Nevertheless, this crucial step is achieved by considering the statistical properties of the non-local gradient magnitude of any real images and without considering an ad-hoc empirical compression formula (and/or an arbitrary value of the mean radiance or gradient magnitude which will remain unchanged by this ad-hoc empirical compression formula).

III. OPTIMIZATION STRATEGY

In our framework, our tone mapping model is thus cast as an optimization problem of a complex (non-convex) Gibbs energy function E. The simplest local minimization technique is probably the Iterative Conditional Modes (ICM) introduced by Besag [37]. This method, which is simply a gradient descent alternating the directions, i.e., that selects a variable while keeping all other variables fixed, is deterministic and simple (it does not even require an analytical expression of the derivative of the energy function to be optimized). Nevertheless, it requires a good initialization of the image to be recovered (sufficiently close to the optimal solution). Otherwise it will converge towards a bad local minima (i.e., in our application, an image solution which does not flatten enough the gradient magnitude distribution). In order to be less dependent on the initial guess, and since an analytical expression of the derivative of the energy function to be minimized is easily available, we have herein used a conjugate gradient procedure with derivative

$$\nabla E(X) = -4 g_{s,t} \sum_{s,t_{s \neq t}} (X_s - X_t) \left(\beta_{s,t}^2 - \left[X_s - X_t \right]^2 \right) \quad (2)$$

For the initial guess, we use a simple redistribution of the radiance value with a histogram specification method [35] with a Gaussian distribution with mean $\hat{\mu}_{\text{Berk}} = 0.5$ and variance $\hat{\sigma}_{\text{Berk}}^2 = 0.05$ and 256 bin values (μ_{Berk} and σ_{Berk}^2 denoting approximately the maximum likelihood estimation of the Gaussian distribution parameter vector fitting the average of the luminance distribution within the Berkeley image database). This initial image allows us to simply both ensure an initial LDR of bin values and an initial solution close to an optimal (i.e., with the adequate level of details) tone mapping result (see Fig. 2).

For the conjugate gradient, the step size is fixed to 0.5 and adaptively decreased by a factor of two if the energy to be minimized increases between two iterations. We stop the procedure if either a fixed number of iterations (IterMax=10) or the variance of the gradient magnitude (controlling the desired level of details) in the output image is reached or if the energy of the gradient descent does not decrease any



Fig. 2. From top to bottom: initial luminance map obtained after the redistribution of the radiance value with a histogram specification method [35] (with a Gaussian distribution with parameters $\mu_{\text{Berk}} = 0.5$ and $\sigma_{\text{Berk}}^2 = 0.05$) and luminance map obtained after the conjugate gradient optimization routine on the HDR image STANFORD MEMORIAL CHURCH (768 × 512 radiance map, courtesy of P. Debevec). Evolution of the energy function *E* and the variance of the magnitude gradient along the iterations of the gradient descent.

more. Concretely, if the variance criterion is not met after the maximal number of iterations it means that no more details can be extracted in the image with our method (local minima). Fig. 2 shows an example of the result of our optimization process on our initial luminance map.

A. Color Treatment

After the optimization process, we must now assign colors R,G,B to the pixels of the compressed HDR image from the luminance map \hat{X} . To this end, we use the conversion procedure proposed by [9] and also exploited by [27] in which⁵

$$R_{\text{out}}, G_{\text{out}}, B_{\text{out}} = \left(\frac{R_{\text{in}}, G_{\text{in}}, B_{\text{in}}}{X}\right)^{\epsilon} \cdot \hat{X}$$
(3)

where we recall that X and \hat{X} denote the luminance respectively before and after HDR compression and ϵ is a parameter controlling the color saturation of the resulting image.

The final R,G,B maps are stretched between 0 and 255, but instead of linearly scaling, we authorize that 0.25% of the pixels at the beginning (i.e., channel level 0) and/or at the end

⁵Alternatively, we can use the following conversion formula [38] which gives approximately the same output result

$$R_{\text{out}}, G_{\text{out}}, B_{\text{out}} = \left(\left[\frac{R_{\text{in}}, G_{\text{in}}, B_{\text{in}}}{X} - 1
ight] \epsilon + 1
ight) \cdot \hat{X}$$



Fig. 3. HDR compression with different values of N_s ($\sigma_{\nabla x}^2 = 0.005$). From top to bottom, magnified regions extracted from the HDR image shown in this paper with respectively $N_s = 3$, and $N_s = 7$. We can notice some aliasing effects or artifacts in color transitions or unrealistic too sharp edges with a small neighborhood $N_s = 3$.

(i.e., channel level 255) of the color channel are saturated in order to enhance the contrast when this is necessary (color maps whose number of pixels are greater than 0.25% at each side of its color channel range remain unchanged).

IV. EXPERIMENTAL RESULTS

A. Set Up and HDR Image Used

In all the experiments, we have thus considered the following set-up. $\beta_{s,t}$ and $g_{s,t}$ are pre-computed before the optimization process. For the model, we have thus considered $g_{s,t} = d_{s,t}^{-1}$ where $d_{s,t}$ is the L_{∞} norm of the distance vector between sites s and t. For each site s, a square neighbourhood window of fixed size $N_s = 7$ pixels (i.e., the 48 nearest neighbors) have been considered for HDR images with less than 2×10^6 pixels and (for computational reasons), a square neighbourhood window of fixed size $N_s = 5$ pixels (i.e., the 24



Fig. 4. From top to bottom HDR radiance maps respectively called the STANFORD MEMORIAL CHURCH (768×512 radiance map, courtesy of P. Debevec [39] dynamic range exceeding 250.000:1) and STREETLIGHT ON A FOGGY NIGHT (1130×751 radiance map, courtesy of J. Tumblin [9] dynamic range exceeding 100.000:1) with from left to right, the method of Fattal *et al.* [27] and Durand *et al.* [10] and our method.

nearest neighbors) for large images with greater than 2×10^6 pixels. We treat the image as toroidal (i.e., wrapping around at the edges). Experiments have shown that the compression results are better when the neighbourhood size is large but at the expense of computational time and memory space. We have used 256 bins for the histogram of the magnitude gradient (ranging from [0.0:0.5] and for the radiance values of an HDR image ranging from [0.0:1.0]). Experiments have shown that there is a relationship between the upper bound of the histogram of the magnitude gradient (herein set to 0.5 and corresponding, in fact, to the maximal value of the gradient magnitude that will belong to the output mapping) and the level of details desired in the output image. For the initial guess of the gradient descent, we use a simple redistribution of the radiance value with an exact histogram specification method [35] with a Gaussian distribution with mean $\hat{\mu}_{\text{Berk}} = 0.5$ and variance $\hat{\sigma}_{\text{Berk}}^2 = 0.05$ (experiments

have shown that the output result is not very sensitive to these parameters). For the conjugate gradient, the step size, the maximal number of iterations and the variance of the gradient magnitude (controlling the desired level of details) in the output image is set respectively to $\gamma = 0.5$, IterMax=10 and $\sigma_{\nabla x}^2 = 0.005$. Finally, for the color treatment, we have set ϵ to 0.6 and when necessary, we use a final stretching of the R, G and B map between 0 and 255 with a maximum of 0.25% pixels saturating at the beginning and at the end of its color channel.

We have tested our model on popular and publicly available HDR radiance images (with dynamic range exceeding 100.000 : 1) already compressed by the state-of-the-art tone mapping methods recently proposed in the literature [40], namely the method of Fattal *et al.* [27] and Durand *et al.* [10].



Fig. 5. Some magnified regions extracted from some HDR compression results obtained by our reduction model using, from top to bottom; at top row, a larger neighborhood system, namely $N_s = 13$ pixels (for comparisons with the results given in the second column of Fig. 3) and at bottom row; a slower decreasing weighting function $g_{s,t}(=d_{s,t}^{-1})$ obtained by setting $d_{s,t} = \text{constant} = 1.0$ for all $\langle s, t \rangle$ (with a neighborhood size setting to the default value $N_s = 7$).

B. Discussion

As previously mentioned, the most important internal parameter is the neighbourhood size. A value greater than 8 $(N_s = 3)$ makes the compression results better with less artifacts (such as halos, aliasing effects or artifacts in color transitions or unrealistic too sharp edges for some HDR images). $N_s = 7$ pixels (i.e., the 48 nearest neighbors) is a good compromise between high quality of the compression results and computational time (and memory requirement). Fig. 3 shows different HDR compression results with different values of N_s . The second most important internal parameter is our stopping parameter, i.e., the variance of the gradient magnitude $\sigma_{\nabla r}^2$ which controls the desired level of details in the output compressed image. Fig. 7 shows different HDR compression results with different values of $\sigma_{\nabla x}^2$. Experiments have also shown that our decreasing positive weighting factor $g_{s,t}$ decreases the halo effect around light sources. Fig. 5 illustrates the effect of a larger neighborhood ($N_s = 13$ pixels) and a slower decreasing weighting factor function $q_{s,t}$ on some (magnified regions extracted from some) compression results. We can notice that a larger neighborhood does not sensibly improve the results (while increasing the computational time) and the weighting factor $g_{s,t}$ should not decrease too slowly. Inversely, as already said, a weighting factor $g_{s,t}$ decreasing faster, and thus equivalent to a small neighbourhood size N_s , is not also recommended (see Fig. 3 with $N_s = 3$ pixels).

In our opinion, our method seems to yield a more photographic look than either the gradient domain method [27] or the toning reduction methods based on the multi-scale image decomposition based techniques such as [10] for which the blacks are not very deep (and the lights, as the sun, are not so bright) and for which almost all the shadows are either removed (cf. Fig. 4) or sometimes omnipresent (cf. Fig. 8). With our method, there is a good balance between shadow

and bright light. In addition, the blacks are deep and the whites are brilliant. Besides, the light is more diffuse without artifacts showing light rays in our HDR radiance map (see image FOGGY NIGHT). The details of the image appear more visible and the image seems to be more contrasted (see Fig. 6). For example, with our method, one can almost read the registration number of the leftmost car in the FOGGY NIGHT image, and we can say, almost surely, that there is a "24" and a "3" in this registration number. Identification after reading is not really possible with the other methods. In addition, in the MEMORIAL CHURCH image, the interior of the church, showing some gilding (using gold leaf) is sensibly closer to the actual color of gold and the white marble color of the stairs is more white in our compression result, compared to the results given by Durand and Fattal's method, in which any color seems orange. The photographic appearance provided by our method is certainly mainly due to our stopping criterion, i.e., the control of the gradient magnitude variance $(\sigma_{\nabla x}^2)$ which is adjusted in order to produce a classical digital photographic look. Nevertheless, a less photographic appearance and a resulting compressed image with more details can be obtained if we specify a higher value for $\sigma_{\nabla x}^2$ thus specifying more details in the output compressed image. In our tests, the exponent ϵ which controls the color saturation of the resulting image is constantly set to 0.6. For some images, a slightly higher value for this parameter (e.g., 0.7 or 0.8) would have produced more intense colors and a visually more pleasant image, although possibly at the expense of a less photographic look.

C. Evaluation

In [40], a thorough evaluation of image preference and rendering accuracy for six (previously published) HDR rendering algorithms has been conducted through several psycho-



Fig. 6. Magnified region extracted from the images shown in Fig. 4. From left to right, the method of Fattal *et al.* [27] and Durand *et al.* [10] and our method.

physical and paired-comparison experiments and over several scenes. In this latter evaluation, the results have shown that the bilateral filter significantly and consistently outperforms with significantly higher rating scale (in regard to overall contrast, sharpness and colorfulness) other test algorithms for both preference and accuracy, making it a good candidate for an obligatory or default algorithm that could be included in future algorithm evaluation experiments [40]. That is why, in that spirit, an experiment is herein conducted to judge the performance and/or preference of our tone mapping algorithm compared to the so-called bilateral filter [10] in a blind subjective paired-comparison paradigm. In our experiment, 31



Fig. 7. HDR compression with different variance values of the gradient magnitude $\sigma_{\nabla x}^2$ (and with $N_s = 7$). From top to bottom, magnified regions extracted from the HDR image (see Fig. 4) called STANFORD MEMORIAL CHURCH (courtesy of P. Debevec with respectively $\sigma_{\nabla x}^2 = 0.02$, $\sigma_{\nabla x}^2 = 0.05$ and $\sigma_{\nabla x}^2 = 0.10$.

participants took part in the experiment and all were naive about its purpose. Each image pair, i.e., the tone mapping result given by our method and the one given by the bilateral filter, was shown in random order to a participant, who had to select the one that he or she preferred (with possibly a "no difference" option) in terms of overall image quality (i.e., overall contrast, colorfulness, sharpness, brilliance, aesthetics aspect and lack of artifacts). These well known images, publicly available, covers a wide range of image content types and are called; MEMORIAL (Fig. 4 above), BIGFOGMAP (Fig. 4 below), BELGIUM, SYNAGOGUE (Fig. 10), SMALL-DESIGN-CENTER (Fig. 8), SMALL-OFFICE (Fig. 9), ATRIUM-NIGHT, FOYER, INDOOR and VENICE images (compression results are given in this paper or in our website at http address www.iro.umontreal.ca/~mignotte/ResearchMaterial/pagetm).

The results are summarized in Fig. 12 and show that, for 4 images out of ten, the participants have preferred the tone mapping result given by our algorithm (i.e., MEMORIAL, BIGFOGMAP, INDOOR and VENICE). For 3 images, they have preferred the compression result given by the bilateral filter (i.e., BELGIUM, SMALL-DESIGN-CENTER and SMALL-OFFICE) and finally for 3 images, they have visually noticed no significant difference between these two results or they have considered them of similar quality (i.e., SYNAGOGUE, ATRIUM-NIGHT and FOYER). On average, for this set of 310 paired-comparisons (10 pictures multiply by 31 observers),



Fig. 8. Interior scene (1000×656 radiance map, courtesy of Marsh Partnership, Perth, Australia, picture called SMALL-DESIGN-CENTER) displayed with respectively (from top to bottom) the method of Durand *et al.* [10] and our method.

42% percent of the selected images came from the bilateral filter algorithm, compared to 48% percent for our compression method whereas 10% percent of the images was considered of similar quality. This experiment demonstrates that the proposed compression method performs competitively compared to the best existing state-of-the-art tone mapping method proposed in the literature. Let us finally add that our compression model is also perfectible since more specific or different distributions shapes can be given for each n-order non-local gradient magnitudes.



Fig. 9. Window scene (656×1000 radiance map, courtesy of Durand [10], picture called SMALL-OFFICE) displayed with respectively (from top to bottom) the method of Durand *et al.* and our method.

D. Algorithm

The tone mapping procedure takes, on average, approximately between 20 and 40 seconds for a 1025×769 HDR image with an AMD Athlon 64 Processor 3500+, 2.2 GHz, 4435.67 bogomips and non-optimized code running on Linux. It is undoubtedly the slowest method amongst the two best state-of-the-art mapping methods presently existing. Nevertheless, it is worth noting that the full multigrid algorithm used in [27] is a very optimized algorithm and our code is clearly non-optimized. Besides, it must be noted than our energy minimization can be efficiently implemented by using the parallel abilities of a graphic processor unit (GPU) (embedded on most graphics hardware currently on the market) and can be greatly accelerated (up to a factor of 200) with a standard NVIDIA©GPU (2004) as indicated in [41].

The source code (in C++ language) of our algorithm with the set of presented compressed images and other images are publicly available at the following http address www.iro.umontreal.ca/~mignotte/ResearchMaterial/pagetm in order to make possible eventual comparisons with future tone reduction algorithms and visual comparisons.

V. CONCLUSION

In this paper, we have presented a new compression model for the display of HDR images. This tone mapping method exploits both the involuntary dynamic retina phenomena, which was recently used in image processing for the difficult edge detection problem [31] and the knowledge of the variance of the statistical distributions of the non-local gradient magnitude related to any natural and real world highly-detailed



Fig. 12. Performance/preference (in terms of percentage) of the proposed algorithm versus the bilateral filter on some 10 well known test images with 31 participants (visual comparison for all these 10 pairs of images are available in our website at http address www.iro.umontreal.ca/mignotte/ResearchMaterial/pagetm).



Fig. 10. SYNAGOGUE scene (769×1025 radiance map, courtesy of Fattal [27]) displayed with respectively (from top to bottom) the method of Fattal *et al.* and our method.

LDR images. This enables the proposed HDR compression problem to be expressed as a Gibbs energy based model or as a nonstationary Markovian model with non-local pairwise interactions. Alternatively, this model can be viewed as a local search or a local optimization problem combined with a new image model, whose fundamental and smallest representative element is the pair of pixels and which also exploits the interesting concept of non local gradient recently put forward by Gilboa and Osher in [36] as a generalization of the gradient operator. Numerically, our detail-preserving contrast reduction model is simply ensured by a conjugate gradient descent based local search, starting from a good initial guess, given by the initial HDR image whose gradient magnitude histogram has been specified beforehand, and stopped when the desired level of details in the output image is reached. While being simple to implement, and also perfectible (e.g., more specific or different distributions shapes can be given for each n-order non-local gradient magnitudes), the proposed procedure performs competitively among the state-of-the-art tone mapping methods recently proposed in the literature.

ACKNOWLEDGMENT

The author would like to thank the anonymous reviewers for their many valuable comments and suggestions that helped to improve both the technical content and the presentation quality of this paper. The author would also like to thank Prof. C. Hélou and his students (IFT1935, Fall 2011) at Montreal University (DIRO) for having kindly accepted to participate to the (blind subjective) paired-comparison experiment of the Section IV-C.

REFERENCES

- G. Joffre, W. Puech, F. Comby, and J. Joffre, "High dynamic range images from digital cameras raw data," *Proc. of ACM SIGGRAPH '05 Posters, ACM Transactions on Graphics*, p. 72, 2005.
- [2] N. Barakat, A. N. Hone, and T. E. Darcie, "Minimal-bracketing sets for high-dynamic-range image capture," *IEEE Trans. Image Processing*, vol. 17, no. 10, pp. 1864–1875, 2008.
- [3] L. Xinqiao and A. E. Gamal, "Synthesis of high dynamic range motion blur free image from multiple captures," *IEEE Trans. Circuits Syst. I*, vol. 50, no. 4, pp. 530 – 539, April 2003.
- [4] S. Battiato, A. Castorina, and M. Mancuso, "High dynamic range imaging for digital still camera: An overview," *SPIE Journal of Electronic Imaging*, vol. 12, no. 3, pp. 459–469, July 2003.
- [5] J. DiCarlo and B. Wandell, "Rendering high dynamic range images," in Proc. of the SPIE:Image Sensors, vol. 3965, 2001, pp. 392–401.



Fig. 13.

INDOOR and FOYER radiance maps and compression results obtained by our method.

- [6] G. W. Larson, H. Rushmeier, and C. Piatko, "A visibility matching tone reproduction operator for high dynamic range scenes," *IEEE Trans. on Visualization and Computer Graphics*, vol. 3, no. 4, pp. 291–3066, Oct-Dec. 1997.
- [7] W.-C. Kao, L.-Y. Chen, and S.-H. Wang, "Tone reproduction in color imaging systems by histogram equalization of macro edges," *IEEE Trans. Consumer Electron.*, vol. 52, no. 2, pp. 682 – 688, May 2006.
- [8] A. Yoshida, V. Blanz, K. Myszkowski, and H.-P. Deidel, "Testing tone mapping operators with human-perceived reality," *SPIE Journal* of *Electronic Imaging*, vol. 16, no. 1, p. 013004 1 14, 2007.
- [9] J. Tumblin and G. Turk, "LCIS: A boundary hierarchy for detailpreserving contrast reduction," *Proc. of SIGGRAPH'99, ACM Transactions on Graphics*, pp. 83–90, 1999.
- [10] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of highdynamic-range images," *Proc. of ACM SIGGRAPH'02, ACM Transactions on Graphics*, vol. 21, no. 3, pp. 257–266, 2002.
- [11] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," *Proc. of* ACM SIGGRAPH'08, ACM Transactions on Graphics, 2008.
- [12] R. Xu and S. N. Pattanaik, "High dynamic range image display using level set framework," in *International Conferences in Central Europe* on Computer Graphics, Visualization and Computer Vision, WSCG'99, 2003.
- [13] Y. Li, L. Sharan, and E. H. Adelson, "Compressing and companding high dynamic range images with subband architectures," *Proc. of ACM SIGGRAPH'05, ACM Transactions on Graphics*, pp. 836–844, 2005.
- [14] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *Proc. of ACM SIGGRAPH'02, ACM Transactions on Graphics*, vol. 21, no. 3, pp. 267–276, 2002.
- [15] Q. Shan, J. Jia, and M. S. Brown, "Globally optimized linear windowed tone-mapping," *IEEE Trans. on Visualization and Computer Graphics*, vol. To appear, 2010.
- [16] S. N. Pattanaik, J. A. Ferwerda, M. D. Fairchild, and D. P. Greenberg, "A multiscale model of adaptation and spatial vision for realistic image display," *Proc. of ACM SIGGRAPH'98, ACM Transactions on Graphics*, pp. 287–298, 1998.
- [17] L. Meylan, "Tone mapping for high dynamic range images," Ph.D. dissertation, Lausanne, 2006.
- [18] J. Zhang and S.-I. Kamata1, "An adaptive tone mapping algorithm for high dynamic range images," in *Lecture Notes in Computer Science, Computational Color Imaging*, 2009, pp. 207–215.
- [19] Z. Rahman, D. Jobson, and G. Woodell, "A multiscale retinex for color rendition and dynamic range compression," SPIE International Symposium on Optical Science, Engineering, and Instrumentation, Conference on Signal and Image Processing, 1996.
- [20] D. J. Jobson, Z. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Trans. Image Processing*, vol. 6, no. 7, pp. 965–976, 1997.
- [21] L. Meylan and S. Susstrunk, "High Dynamic Range image rendering using a Retinex-based adaptive filter," *IEEE Trans. Image Processing*, vol. 15, no. 9, pp. 2820–2830, 2006.

- [22] L. Meylan, D. Alleysson, and S. Ssstrunk, "A Model of Retinal Local Adaptation for the Tone Mapping of Color Filter Array Images," *Journal* of the Optical Society of America A (JOSA A), vol. 24, no. 9, pp. 2807– 2816, 2007.
- [23] P. Ledda, L. Santos, and A. Chalmers, "A local model of eye adaptation for high dynamic range images," in *Proc. of the 3rd International Conference on Computer Graphics, Virtual reality, Visualization and Interaction in Africa (Proceedings of ACM AFFRIGRAPH'04)*, 2004, pp. 151–160.
- [24] M. Ashikhmin, "A tone mapping algorithm for high contrast images," in Proc. Eurographics. Workshop on Rendering, 2002, pp. 145–156.
- [25] G. Johnson and M. Fairchild, "Rendering HDR images," in IS&T/SID 11th Color Imaging Conference, Scottsdale, 2003, pp. 36–41.
- [26] C. Gatta, A. Rizzi, and D. Marini, "Perceptually inspired hdr images tone mapping with color correction," *Journal of Imaging Systems and Technology*, no. 17, pp. 285–294, 2007.
- [27] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," *Proc. of ACM SIGGRAPH'02, ACM Transactions on Graphics*, vol. 21, no. 3, pp. 249–256, 2002.
- [28] A. Munsell, L. Sloan, and I. Godlove., "Neutral values scales. i. munsell neutral value scale," *Journal of the optical Society of America*, vol. 23, pp. 394–411, 1933.
- [29] A. Roka, A. Csapo, B. Resko, and P. Baranyi, "Edge detection model based on involuntary eye movements of the eye-retina system," Acta Polytechnica Hungarica, Journal of Applied Sciences, Special Issue on Computational Intelligence, and Applied Machine Intelligence and Informatics, vol. 4, no. 1, pp. 31–46, 2007.
- [30] S. Zozor, P. Amblard, and C. Duchene, "Does eye tremor provide the hyperacuity phenomenon?" 2009. [Online]. Available: http://infoscience.epfl.ch/record/130710
- [31] M.-O. Hongler, Y. L. de Menesse, A. Beyeler, and J. Jacot, "The resonant retina: Exploiting vibration noise to optimally detect edges in an image," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 25, no. 9, pp. 1051–1062, September 2003.
- [32] J. Shen, H. Sun, H. Zhao, and X. Jin, "Ram-based tone mapping for high dynamic range images," in *IEEE International Conference on Multimedia and Expo*, 2009. *ICME 2009.*, New York City, USA, July 2009, pp. 1110–1113.
- [33] E. Simoncelli, "Statistical models for images: compression, restoration and synthesis," in *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems & Computers*, Pacific Grove, CA, USA, Feb. 1997, pp. 673–678.
- [34] D. Martin and C. Fowlkes, "The Berkeley segmentation database and benchmark," image database and source code publicly available at address http://www.cs.berkeley.edu/projects/vision/grouping/segbench/.
- [35] D. Coltuc, P. Bolon, and J. M. Chassery, "Exact histogram specification," *IEEE Trans. Image Processing*, vol. 15, no. 5, pp. 1143–1152, 2006.
- [36] G. Gilboa and S. Osher, "Nonlocal linear image regularization and supervised segmentation, multiscale model," *Multiscale Model. Simul.*, vol. 6, no. 2, p. 595 630, 2007.
- [37] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society*, vol. B-48, pp. 259–302, 1986.



Fig. 11. NANCY CHURCH (France) (radiance map courtesy of R. Mantiuk) results obtained by our method.



Fig. 14. VENICE and ATRIUMNIGHT radiance maps and compression results obtained by our method.



Max Mignotte received the DEA (Postgraduate degree) in Digital Signal, Image and Speech processing from the INPG University, France (Grenoble), in 1993 and the Ph.D. degree in electronics and computer engineering from the University of Bretagne Occidentale (UBO) and the digital signal laboratory (GTS) of the French Naval academy, France, in 1998. He was an INRIA post-doctoral fellow at University of Montreal (DIRO), Canada (Quebec), from 1998 to 1999. He is currently with DIRO at the Computer Vision & Geometric Modeling Lab as

an associate Professor at the University of Montreal. He is also a member of LIO (Laboratoire de recherche en imagerie et orthopedie, Centre de recherche du CHUM, Hopital Notre-Dame) and researcher at CHUM. His current research interests include statistical methods, Bayesian inference and hierarchical models for high-dimensional inverse problems.

[38] R. Mantiuk, R. Mantiuk, A. Tomaszewska, and W. Heifrich, "Color correction for tone mapping," in *Computer Graphics Forum (Proc. EUROGRAPHICS)*, 2009, pp. 193–202.

- [39] P. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs." *Proc. of ACM SIGGRAPH'97, ACM Transactions on Graphics*, pp. 369–378, 1997.
- [40] J. Kuang, H. Yamaguchi, C. Liu, G. M. Johnson, and M. D. Fairchild, "Evaluating HDR rendering algorithms," ACM Transactions on Applied Perception (TAP), vol. 4, no. 2, July 2007.
- [41] P.-M. Jodoin and M. Mignotte, "Markovian segmentation and parameter estimation on graphics hardware," *Journal of Electronic Imaging*, vol. 15, pp. 033015–1–15, 2006.