

# A NON LOCAL VARIATIONAL FORMULATION FOR THE IMPROVEMENT OF TONE MAPPED IMAGES \*

PRAVEEN CYRIAC , THOMAS BATARD , AND MARCELO BERTALMÍO †

**Abstract.** Due to technical limitations, common display devices can only reproduce images having a low range of intensity values (dynamic range). As a consequence, the dynamic range of images encoding real world scenes, which is large, has to be compressed in order for them to be reproduced on a common display, and this technique is called tone mapping. Because there is no ground truth to compare with, evaluation of a tone mapped image has to be done by comparing with the original high dynamic range image. As standard metrics based on pixel-wise comparisons are not suitable for comparing images of different dynamic range, non local perceptual based metrics are commonly used. We propose a general method for optimizing tone mapped images with respect to a given non local metric. In particular, if the metric is perceptual, i.e. it involves perceptual concepts, we provide an adequate minimization strategy. Experiments on a particular perceptual metric tested with different tone mapped images provided by several tone mapping operators validate our approach.

**Key words.** Tone mapping, non local variational problem, perceptual distance, just noticeable difference

## 1. Introduction.

### 1.1. On tone mapping operators.

**1.1.1. The tone mapping challenge.** The vast range of light intensities of the real world span many orders of magnitude. Even though the Dynamic Range (DR) of the Human Visual System (HVS) is only 3-4 orders of magnitude, nonetheless it is capable to handle intensities from about  $10^{-6}$  to  $10^8$   $cd/m^2$  [21], due to the fact that it continuously adjusts to the light in any viewed scene. As the DR of most cameras is only 2-3 orders of magnitude, they tend to fail in capturing the details and contrast that we perceive with the naked eye.

The most popular approach to capture the real world luminance is to create High Dynamic Range (HDR) images through the fusion of multiple Low Dynamic Range (LDR) images generated by a standard camera shooting the scene with varying exposure time [6]. However, HDR images can not be directly reproduced on common displays since their ranges usually span only two orders of magnitude, meaning that they can only reproduce LDR images.

Hence, a range compression of HDR images has to be performed in order for them to be reproduced on common displays, and this technique is called tone mapping. In particular, a good tone mapping operator (TMO) should produce in anyone watching the display a perception of details as close as possible to the one he/she would have had by observing the original scene directly [29].

**1.1.2. Overview of tone mapping operators.** Tumblin and Rushmeir [27] formally introduced the problem of tone mapping to the computer graphics field. Their TMO aims at transforming the real world luminance into the luminance generated by the display device. Since then, many TMOs have been proposed, and can be classified as global or local approaches.

---

\*This work was supported by European Research Council, Starting Grant ref. 306337, and by Spanish grants ref. TIN2011-15954-E and ref. TIN2012-38112. Reduced version of this work appeared as a conference paper in [4].

†Department of Information and Communication Technologies, University Pompeu Fabra, Barcelona, Spain(praveen.cyriac@upf.edu, thomas.batard@upf.edu, marcelo.bertalmio@upf.edu).

Most of global TMOs consist in applying a compression curve to the image levels, based on psychovisual laws. Besides Tumblin and Rushmeir [27] who use Stevens' law, the Naka-Rushton formula is used in ([21],[24],[13]), Ferwerda's model in [12], and Weber-Fechner's law in ([2],[29]) to name a few. In particular, Reinhard et al. [24] developed a TMO based on the idea that tone mapping is similar to the adaptation process in the HVS, and used a modification of the Naka-Rushton equation. Drago et al. [7] introduced an adaptive logarithmic curve using a collection of logarithmic functions ranging from  $\log_2$  to  $\log_{10}$ , the choice of the logarithm base depending on the luminance values. Mantiuk et al. [18] developed a piece-wise linear tone curve to achieve DR compression, whose parameters are chosen so as to minimize the difference between the estimated response of the HVS for the resultant image and the original image. Global TMOs are in general very fast and do not introduce halos or artifacts, but tend to produce low contrast images.

Local TMOs achieve DR compression by modifying each pixel based on its neighborhood. Even though they are computationally more expensive than global approaches, they produce higher contrast images. However they have the tendency to produce artefacts and halos [8][10][16].

More recently, Ferradans et al. [11] proposed a two stage TMO combining both approaches. The first stage is a global tone mapping method that implements visual adaptation by combining the Naka-Rushton equation and Weber-Fechner's law. The second stage performs local contrast enhancement, based on a variational model inspired by colour vision phenomenology.

**1.1.3. Evaluation of tone mapping operators.** The most straightforward way to evaluate a TMO is to perform a subjective evaluation, where observers rate the tone mapped image by comparing with the reference HDR image displayed on a HDR screen [14], or they simply evaluate the tone mapped image by itself, without any reference [31]. Subjects are asked to consider image attributes such as brightness, contrast, colors, and naturalness, as well as overall quality. But subjective evaluation is limited in many ways: firstly, it is often time consuming and expensive; secondly, it is difficult to incorporate it in the design of tone mapping algorithms. This points out the importance of objective evaluation of TMOs.

An accurate objective evaluation should mimic the subjective evaluation described above, so it requires the use of a perceptual metric between images of different dynamic range. An objective tone mapping evaluation tool has been proposed by Smith et al. [26], based on the measure of suprathreshold contrast distortion between the source HDR image and its tone mapped LDR version. However, the contrast measure is local, meaning that its sensitivity is limited to high frequency details. More recently, Aydin et al. [3] proposed a dynamic range independent metric (DRIM) whose contrast measure is not limited to the values of neighboring pixels. Moreover, the metric predicts three type of distortions at each pixel between the two images it compares: loss of visible features, amplification of invisible features, and reversal of contrast polarity (see Sect. 4.1 for a more detailed description of the metric DRIM).

**1.2. Contribution.** We propose in this paper a new approach for the improvement of the tone mapped images that takes into account the metric that will be used for the evaluation. In such a way, we are able to improve an existing TMO with respect to a given metric. The aim of this paper is two-fold:

1. We develop a general framework for improving any tone mapped image by reducing the distance with the corresponding HDR image. More precisely, assuming that the metric *met* is designed to compare images of different dynamic range, our

approach for reducing the distance between an HDR image  $H: \Omega \rightarrow \mathbb{R}$  and its tone mapped version  $L_0 \in LDR(\Omega)$  is to minimize functionals of the form

$$(1.1) \quad \arg \min_{L \in LDR(\Omega)} \int_{\Omega} \Phi(\text{met}(L, H)(x)) dx$$

where  $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^+$ , through a gradient descent algorithm of initial condition  $L_0$ . Then the main task to implement the algorithm is to construct a discrete approximation of the functional derivative

$$(1.2) \quad \frac{\partial \text{met}(L, H)(x)}{\partial L(y)}$$

2. We apply the general framework described in (1.1) in the context of the perceptual metric DRIM [3]. More precisely, we consider the following particular cases of (1.1)

$$(1.3) \quad \arg \min_{L \in LDR(\Omega)} \int_{\Omega} \|DRIM(L, H)(x)\|_k dx$$

where  $k$  is a strictly positive constant. We evaluate our method by comparing in a two-fold way the initial condition with the steady-state of the gradient descent algorithm. Our tests show that we are able to improve the tone mapping results of several state of the art approaches by a significant amount.

This paper is organized as follows. In Sect. 2, we present our general framework for improving any tone mapped image as a gradient descent algorithm associated to a non local variational problem. In particular, we give the expression of the gradient of the functional to minimize. In Sect. 3, we consider the case of perceptual metrics and detail the computation of the discrete functional derivative (1.2) from which derives the discrete gradient of the functional (1.1). At last, in Sect. 4, we test the corresponding gradient descent algorithm on several tone mapped images dealing with the perceptual metric DRIM.

## 2. General framework for improving tone mapping results.

**2.1. Distance between images as a non local operator.** Many tasks in image processing and computer vision require a validation by comparing the result with the original data, e.g. optical flow estimation, image denoising, tone mapping. Whereas measures based on pixel-wise comparisons (e.g. MSE, SNR, PSNR) are suitable to evaluate algorithms for problems such as image denoising and optical flow estimation, they are not relevant to evaluate tone mapping results, because the ground truth is not known. Indeed, tone mapping results are evaluated by comparisons with the original HDR images, where pixel-wise comparisons are not suitable since the images compared are of different dynamic range. This leads us to the following definitions, where we first remind the reader of the concept of functional differentiation.

**DEFINITION 2.1** (functional derivative). *Let  $X$  be a Banach space of scalar-valued functions defined on some domain  $\Omega$ . The derivative of  $F: X \rightarrow \mathbb{R}^n, n \geq 1$ , with respect to the function  $I \in X$  at the point  $y \in \Omega$  is the quantity*

$$(2.1) \quad \frac{\partial F(I)}{\partial I(y)} = \lim_{\epsilon \rightarrow 0} \frac{F(I + \epsilon \delta_y) - F(I)}{\epsilon}$$

where  $\delta_y$  is the Dirac delta function concentrated at the point  $y$ .

The differential of  $F$  in the direction  $\varphi \in X$  is then

$$(2.2) \quad \delta F(I; \varphi) = \int_{\Omega} \varphi(y) \frac{\partial F(I)}{\partial I(y)} dy$$

From now on, we consider the set of grey-level images defined on a domain  $\Omega \subset \mathbb{R}^2$  as the Banach space  $X = C^\infty(\Omega; [0, 1])$ .

**DEFINITION 2.2 (metric).** Let  $Y$  be the Banach space of vector-valued smooth functions defined on  $\Omega$ . We call **metric** an operator  $\mathbf{met}: X \times X \rightarrow Y$ .

For  $x \in \Omega$ , we consider  $F: X \times X \rightarrow \mathbb{R}^n, n \geq 1$ , defined by  $F(L, H) = \mathbf{met}(L, H)(x)$  and the partial derivatives of  $F$  with respect to  $L$  and  $H$  at the point  $y \in \Omega$ , i.e.

$$(2.3) \quad \frac{\partial \mathbf{met}(L, H)(x)}{\partial L(y)} = \lim_{\epsilon \rightarrow 0} \frac{\mathbf{met}(L + \epsilon \delta_y, H)(x) - \mathbf{met}(L, H)(x)}{\epsilon}$$

$$(2.4) \quad \frac{\partial \mathbf{met}(L, H)(x)}{\partial H(y)} = \lim_{\epsilon \rightarrow 0} \frac{\mathbf{met}(L, H + \epsilon \delta_y)(x) - \mathbf{met}(L, H)(x)}{\epsilon}$$

where  $\delta_y$  is the Dirac delta function concentrated at the point  $y \in \Omega$ .

We say that  $\mathbf{met}$  is

- (i) **pixel-wise** if  $\forall L, H \in X, \forall x, y \in \Omega, y \neq x$ , the quantities (2.3), (2.4) vanish.
- (ii) **non local** if  $\forall L, H \in X, \forall x \in \Omega, \exists \Omega \supset \mathcal{N}(x) \ni x, \mathcal{N}(x) \neq \{x\}, \exists y \notin \mathcal{N}(x)$  s.t. the quantities (2.3), (2.4) do not vanish.

**DEFINITION 2.3 (distance).** A **distance** associated to the metric  $\mathbf{met}$  is a functional  $E: X \times X \rightarrow \mathbb{R}^+$  of the form

$$(2.5) \quad E: (L, H) \mapsto \int_{\Omega} \Phi(\mathbf{met}(L, H)(x)) dx$$

for some map  $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^+$ .

We say that the distance  $E$  is pixel-wise, resp. non local, if the associated metric is pixel-wise, resp. non local.

Note that we only require the distance (2.5) to satisfy the non-negativity property among the classical properties of a distance function (identity for  $L = H$ , symmetry, triangle inequality).

In this context, we can classify image quality measures into several categories. The set of pixel-wise distances includes MSE, PSNR and SNR measures, where the function  $\Phi$  in (2.5) is the Euclidean norm  $\|\cdot\|_2$ . Indeed, the corresponding terms  $\mathbf{met}(L + \epsilon \delta_y, H)(x) - \mathbf{met}(L, H)(x)$  and  $\mathbf{met}(L, H + \epsilon \delta_y)(x) - \mathbf{met}(L, H)(x)$  in formulae (2.3), (2.4) vanish for  $y \neq x$  since such quantities only depend on the values  $L(x)$  and  $H(x)$ .

The second category gathers the metrics that are non local. This category can actually be divided into two sub-categories: the set of metrics that compare images

of same dynamic range (see e.g. [5] for LDR images and [15],[19] for HDR images), and the set of metrics comparing images of different dynamic range (see e.g. [26]). The metric DRIM [3] belongs to both sub-categories since it is independent of the dynamic range of the images it compares.

At last, there exist image quality measures that are neither metrics nor distances according to the definitions 2.2 and 2.3. This includes for instance the SSIM metric [28].

**2.2. Reduction of the distance between an HDR image and its tone mapped version.** In practice, LDR images are usually encoded in 8 bits whereas HDR images are often encoded in 16 bits. This leads us to define the functional spaces  $LDR(\Omega)$  and  $HDR(\Omega)$  as follows

$$(2.6) \quad HDR(\Omega) = \{f \in C^\infty(\Omega; [1/65536, 1])\}$$

$$(2.7) \quad LDR(\Omega) = \{f \in C^\infty(\Omega; [1/256, 1])\}$$

In particular, we have  $LDR(\Omega) \subset HDR(\Omega)$ .

We consider the following problem: given an HDR image  $H$  and a metric  $met$ , we aim at constructing the LDR image  $L_*$  minimizing the distance with  $H$ , i.e. we aim at solving the following variational problem

$$(2.8) \quad L_* := \arg \min_{L \in LDR(\Omega)} E(L, H)$$

**PROPOSITION 2.4.** *Assuming that  $met$  is continuous, bounded, and  $\Phi$  is continuous, the variational problem (2.8) has a solution.*

*Proof.* Under the assumption of the Proposition, the energy  $E$  is bounded since the domain  $\Omega$  of an image is a compact subset of  $\mathbb{R}^2$ . Moreover, the set  $LDR(\Omega)$  being closed, we deduce that there exists a function  $L^* \in LDR(\Omega)$  (not necessarily unique) solution of the variational problem (2.8).  $\square$

**PROPOSITION 2.5.** *Assuming that  $met$  and  $\Phi$  are differentiable, the functions  $L$  which are critical points of the energy*

$$(2.9) \quad E_H: L \longmapsto E(L, H)$$

*satisfy*

$$\int_{\Omega} \delta \Phi \left( met(L, H)(x); \frac{\delta met(L, H)(x)}{\delta L(y)} \right) dx = 0 \quad \forall y \in \Omega$$

*Proof.* The functional  $E_H$  being differentiable on the whole set  $LDR(\Omega)$ , its critical points are the functions  $L$  where its gradient  $\nabla E_H$  vanishes.

Let  $\psi: \Omega \longrightarrow \mathbb{R}$  be a compact support function. We compute the differential  $\delta E_H$  of the energy  $E_H$  at a function  $L$  in the direction  $\psi$ . We have

$$\delta E_H(L; \psi) = \delta E((L, H); (\psi, 0))$$

$$\begin{aligned}
&= \int_{\Omega} \delta(\Phi \circ \text{met})((L, H); (\psi, 0))(x) dx \\
&= \int_{\Omega} \delta\Phi(\text{met}(L, H)(x); \delta \text{met}((L, H); (\psi, 0))(x)) dx \\
&= \int_{\Omega^2} \delta\Phi\left(\text{met}(L, H)(x); \frac{\delta \text{met}(L, H)(x)}{\delta L(y)}\right) \psi(y) dx dy
\end{aligned}$$

The last equality results from formula (2.2). At last, as  $\psi$  has compact support,  $\delta E_H(L; \psi) = 0 \implies$

$$(2.10) \quad \int_{\Omega} \delta\Phi\left(\text{met}(L, H)(x); \frac{\delta \text{met}(L, H)(x)}{\delta L(y)}\right) dx = 0 \quad \forall y \in \Omega$$

□

The gradient of the functional  $E_H$  at the function  $L$  is the map

$$(2.11) \quad \nabla E_H(L): y \longmapsto \int_{\Omega} \delta\Phi\left(\text{met}(L, H)(x); \frac{\delta \text{met}(L, H)(x)}{\delta L(y)}\right) dx$$

Due to the complexity of the metrics that compare images of different dynamic range, the non local operators  $\text{met}$  encoding the metrics lack mathematical properties like convexity. As a consequence, it is hard to prove the uniqueness or not of the solution and establish accurate numerical schemes for reaching the solution(s) of the variational problem (2.8).

For this reason, we adopt in this paper the following approach: instead of constructing the best tone mapped image(s) of a given HDR image  $H$  with respect to the metric  $\text{met}$  by constructing solution(s) of the variational problem (2.8), we improve existing tone mapping results by the use of a gradient descent algorithm where the initial condition  $L_0$  is a tone mapped version of  $H$ .

Even if the gradient descent algorithm might converge towards critical points  $L_*$  that are not global minima, we have  $E_H(L_*) < E_H(L_0)$  meaning the LDR image  $L_*$  is an improvement of the tone mapping result  $L_0$ .

The main task to compute a discrete approximation of the gradient (2.11) in order to perform the gradient descent algorithm is to compute accurate discrete approximations of the functional derivative

$$\frac{\partial \text{met}(L, H)(x)}{\partial L(y)} : = \lim_{\epsilon \rightarrow 0} \frac{\text{met}(L + \epsilon \delta_y, H)(x) - \text{met}(L, H)(x)}{\epsilon}$$

where  $\delta_y$  is the Dirac delta function concentrated at the point  $y \in \Omega$ .

Our proposal is to make use of central differences of the form

$$(2.12) \quad \frac{\partial \text{met}^D(L, H)(a, b)}{\partial L(i, j)} : = \frac{\text{met}(L + \epsilon_1 \delta_{(i, j)}, H)(a, b) - \text{met}(L - \epsilon_2 \delta_{(i, j)}, H)(a, b)}{d(L + \epsilon_1 \delta_{(i, j)}, L - \epsilon_2 \delta_{(i, j)})}$$

for some well-chosen  $\epsilon_1, \epsilon_2$ , and where  $d(L + \epsilon_1 \delta_{(i, j)}, L - \epsilon_2 \delta_{(i, j)})$  measures a difference between the two images  $L + \epsilon_1 \delta_{(i, j)}$  and  $L - \epsilon_2 \delta_{(i, j)}$ .

A straightforward choice would be to impose  $\epsilon_1, \epsilon_2$  to be constant and independent of both the intensity values  $L(i, j)$  and the metric  $\text{met}$ .

**3. Discrete functional derivatives in the case of perceptual metrics.** We argue that the increments  $\epsilon_1, \epsilon_2$  in (2.12) should be dependent on the nature of the metric. In particular, dealing with perceptual metrics, we claim that  $\epsilon_1, \epsilon_2$  should be chosen to make the difference between the images  $L$  and  $L + \epsilon_1 \delta_{(i,j)}$  be perceived equally to the difference between the images  $L$  and  $L - \epsilon_2 \delta_{(i,j)}$ .

Our proposal in this paper is to make use of the concept of Just Noticeable Difference (JND) in order to compute perceived differences between images. However, perceptual metrics do not necessarily compare images in perceptually uniform spaces, meaning that we need to express the JND into the space in which the metric is operating, which is the luminance space in most of cases.

**3.1. Approximation of the JND in the luminance space.** In the domain of psychophysics, the JND is the smallest difference  $\Delta I$  in the intensity of a stimulus at which a human is able to perceive a difference between a uniform intensity  $I$  and a superimposed intensity  $I + \Delta I$ .

Weber's law, named after the German physician E.H. Weber, relates the JND with the intensity of the stimulus according to the formula

$$(3.1) \quad \frac{\text{JND}}{I} = k$$

for some constant  $k > 0$ . Note that  $k$  varies with the nature of the stimulus. Regarding perception, visual experiments conducted later on by others showed that Weber's law holds for a large range of light intensity  $I$ . However, the relation does not hold for low-intensity values [23].

The **luminance**, denoted by  $Y$ , is a measure of the light stimulus and is defined as the radiance (light intensity reaching the retina) weighted by spectral sensitivity functions. However the space  $Y$  is not perceptually uniform in the sense that the difference between two luminance values is not proportional to the difference of light intensity observed. In 1976, the CIE introduced a measure of the perception of the light, called **lightness**, as the quantity

$$(3.2) \quad L^* = \begin{cases} 903.3 \frac{Y}{Y_n} & \text{if } \frac{Y}{Y_n} \leq 0.008856 \\ 116 \left( \frac{Y}{Y_n} \right)^{1/3} - 16 & \text{if } \frac{Y}{Y_n} > 0.008856 \end{cases}$$

where  $Y_n$  is the highest luminance value of the scene. The quantity  $Y/Y_n$  is called **relative luminance**. **The space  $L^*$  is perceptually uniform** since a difference of 1 in the space  $L^*$  approximates pretty well 1 JND (see the book of Poynton [22] for details about this and other of the following technical concepts).

Formula (3.2) makes the lightness have a range of 0 to 100. When scaled to the range  $[0, 1]$ ,  $L^*$  can be approximated by the 0.4-power of the relative luminance, i.e. we have the relation

$$(3.3) \quad 0.01 L^* \simeq \left( \frac{Y}{Y_n} \right)^{0.4}$$

When shooting a scene, a digital camera captures the light intensity from which it encodes  $R, G, B$  values. Standard digital cameras also perform **gamma correction**:

$$(3.4) \quad R' = R^{\frac{1}{\gamma}} \quad G' = G^{\frac{1}{\gamma}} \quad B' = B^{\frac{1}{\gamma}}$$

Assuming that these values are encoded in the sRGB color space (which is the standard color space used in the broadcast and computer industries),  $\gamma$  is approximately 2.2.

Then the relative luminance  $Y/Y_n$  of an image displayed by a monitor is given by

$$(3.5) \quad \frac{Y}{Y_n} \simeq 0.2126 R'^\gamma + 0.7152 G'^\gamma + 0.0722 B'^\gamma$$

From formulae (3.5) and (3.3), we end up with the formula

$$(3.6) \quad L^* \simeq (0.2126 R'^\gamma + 0.7152 G'^\gamma + 0.0722 B'^\gamma)^{0.4}$$

(assuming that the lightness is normalized to the range  $[0,1]$ ).

Finally, using the formula (3.6) and the perceptual uniformity of the lightness  $L^*$ , we can express the JND for the light intensity of a color  $(R', G', B')$  displayed at a screen. An alternative approach to approximate the JND in luminance space can be found in [25] (chapter 10, section 10.4.2)

**3.2. Computation of the terms  $\epsilon_1, \epsilon_2, d$ .** We assume that the metric *met* in formula (2.12) is operating in the luminance space (e.g. the metric DRIM [3]).

In order to relate the increments  $\epsilon_1, \epsilon_2$  as well as the measure  $d$  in formula (2.12) to the JND, we have to face the following issue: the JND is determined on a uniform background whereas expression (2.12) deals with image pixels. In order that the computation of the JND at an image pixel makes more sense, our proposal is to compute it using a smoothed version of the image  $L$

$$(3.7) \quad \tilde{L} := L * G_\sigma$$

where  $G_\sigma$  is the Gaussian kernel associated to some variance  $\sigma$ . Indeed, convolving an image with a Gaussian kernel reduces its variations, making the smoothed image be locally closer to a uniform patch than the original image.

We can then express the JND for the light intensity perceived at the pixel  $(i, j)$  of the image  $L$  using the formula (3.3) and the smoothed version  $\tilde{L}$  of  $L$  as follows.

As a difference of 0.01 in the (normalized) lightness space approximates pretty well 1 JND, we aim at solving the two functional equations

$$(3.8) \quad (\tilde{L} + \epsilon_1)^{0.4} - \tilde{L}^{0.4} = 0.01$$

$$(3.9) \quad \tilde{L}^{0.4} - (\tilde{L} - \epsilon_2)^{0.4} = 0.01$$

whose solutions are

$$(3.10) \quad \epsilon_1 = (\tilde{L}^{0.4} + 0.01)^{2.5} - \tilde{L}$$

$$(3.11) \quad \epsilon_2 = \tilde{L} - (\tilde{L}^{0.4} - 0.01)^{2.5}$$

However, defining  $\epsilon_1, \epsilon_2$  as in (3.10), (3.11) yields an extra issue: we might have  $L(i, j) + \epsilon_1(i, j) > 1$  or  $L(i, j) - \epsilon_2(i, j) < 0$ . In such a case, the image  $L + \epsilon_1(i, j)\delta_{i,j}$  or  $L - \epsilon_2(i, j)\delta_{i,j}$  and consequently formula (2.12) would not be defined.



We then relax this perceptual uniformity paradigm in order to stay in the range  $[0, 1]$ . We end up with the following two extensions of equations (3.8) and (3.9)

$$(3.12) \quad (\tilde{L} + \epsilon_1)^{0.4} - \tilde{L}^{0.4} = \min(0.01, 1 - \tilde{L}^{0.4})$$

$$(3.13) \quad \tilde{L}^{0.4} - (\tilde{L} - \epsilon_2)^{0.4} = \min(0.01, \tilde{L}^{0.4})$$

whose solutions are

$$(3.14) \quad \epsilon_1 = (\tilde{L}^{0.4} + \min(0.01, 1 - \tilde{L}^{0.4}))^{2.5} - \tilde{L}$$

$$(3.15) \quad \epsilon_2 = \tilde{L} - (\tilde{L}^{0.4} - \min(0.01, \tilde{L}^{0.4}))^{2.5}$$

Hence, the image  $L + \epsilon_1(i, j)\delta_{i,j}$  never exceeds 1 since  $\tilde{L} \leq L$  by construction. For the same reason,  $L - \epsilon_2(i, j)\delta_{i,j}$  never gets negative.

Finally, we set

$$(3.16) \quad d = \min(0.01, 1 - \tilde{L}^{0.4}) + \min(0.01, \tilde{L}^{0.4})$$

where  $\min(0.01, 1 - \tilde{L}^{0.4})$  can be viewed as a perceptual distance between the images  $L$  and  $L + \epsilon_1(i, j)\delta_{i,j}$  and  $\min(0.01, \tilde{L}^{0.4})$  a perceptual distance between the images  $L$  and  $L - \epsilon_2(i, j)\delta_{i,j}$ . The summation of the two terms makes the quantity  $d$  represent a perceptual distance between the images  $L + \epsilon_1(i, j)\delta_{i,j}$  and  $L - \epsilon_2(i, j)\delta_{i,j}$ .

#### 4. Tests with the perceptual metric DRIM.

**4.1. Dynamic range independent perceptual distances.** As mentioned above, the metric DRIM of Aydin et al. [3] compares in a perceptual manner images of any dynamic range. It aims at predicting details (contrast) changes between two images. More precisely, the purpose of this metric is to consider the perception that a viewer would have of both images relying on psychophysical data, and to estimate at each pixel the probabilities that distortions between the two images appear. The metric estimates three types of distortions between the two images:

- Loss of Visible Contrast (LVC), meaning that contrast is visible in one image (called the reference image) and not in the second one (called test image).
- Amplification of Invisible Contrast (AIC), when details that were not in the reference image appear in the test image.
- Contrast reversal (INV), meaning that contrast is visible in both reference and test images but with different polarity (i.e. the same contour has the gradient in opposite directions in the two images.)

The metric first considers the two input images separately and estimates the pixel-wise probability of contrast to be visible ( $P_{./v}^{k,l}$ ) and the probability of contrast not to be visible ( $P_{./i}^{k,l}$ ) for each image for each of several bands of orientation  $l$  and spatial frequency  $k$ , using models for processes in the human visual system [15, 30, 5].

Then the distortions LVC, AIC and INV are estimated for each pair of bands as:

$$(4.1) \quad P_{LVC}^{k,l} = P_{r/v}^{k,l} \times P_{t/i}^{k,l} \quad P_{AIC}^{k,l} = P_{r/i}^{k,l} \times P_{t/v}^{k,l} \quad P_{INV}^{k,l} = P_{r/v}^{k,l} \times P_{t/v}^{k,l} \times R^{k,l}.$$

The subscript  $r/$  and  $t/$  indicate reference and test images. If the polarity of contrast in the test and reference images is the same then  $R$  is set to 1, else to 0.

The final LVC output of the metric is the probability of detecting LVC distortion in any band, and is obtained by combining the probabilities over all the bands, according to the formula

$$LVC = 1 - \prod_{k=1}^N \prod_{l=1}^M (1 - P_{LVC}^{k,l})$$

where  $N$  and  $M$  are the number of orientation and spatial bandwidth. AIC and INV are computed in a similar way.

Note that for applications to tone mapping evaluation, the reference image is the luminance map of an HDR image and the test image is the luminance map of its LDR tone mapped version.

The metric also provides a “distortion map” (a color image) to visualize pixel-wise distortions with the following color code: green hue represents LVC, blue hue stands for AIC and red hue indicates INV, the saturation encoding the magnitude of the corresponding distortion, whereas the intensity corresponds to the intensity of the reference image (up to rounding). At each pixel, the maximum of the three distortions is computed, and the corresponding color is displayed. If the maximum is lower than 0.5, then the saturation is set to 0.

These three types of distortions estimated by the metric DRIM are illustrated on Fig. 4.1. On the left column, we show two LDR grey-level images. We apply some filters to these images in order to create distortions. To the image located in the top row, we apply a Gaussian smoothing on its top-right part and unsharp masking on its top-left part in order to obtain respectively a contrast reduction and a contrast enhancement. To the image located in the bottom row, we apply some contrast reversal technique on the pattern in the chair. The images resulting of these filters are shown in the middle column. Then, we compute the metric DRIM where the reference images are the original images on the left column and the test images are their distorted versions in the middle column. The distortion maps, shown in the right column, are consistent with the artificial distortions we created.

As the output of the metric DRIM encodes at each pixel the probabilities of detecting the distortions LVC,AIC,INV between two images, we reinterpret the metric DRIM as follows.

**DEFINITION 4.1** (Dynamic range independent perceptual metric). *Let  $L, H: \Omega \rightarrow \mathbb{R}$  be two images of any dynamic ranges, the perceptual metric DRIM between  $L$  and  $H$  is the map*

$$(4.2) \quad \begin{aligned} \text{DRIM}(L, H): \quad \Omega &\longrightarrow [0, 1]^3 \\ x &\longmapsto (LVC(L, H)(x), AIC(L, H)(x), INV(L, H)(x)) \end{aligned}$$

From the perceptual metric (4.2) we define a set of perceptual distances between images of any dynamic range.

**DEFINITION 4.2** (Dynamic range independent perceptual distances). *Let  $k$  be a strictly positive number. The perceptual distance  $E_k(L, H)$  between  $L$  and  $H$  is*

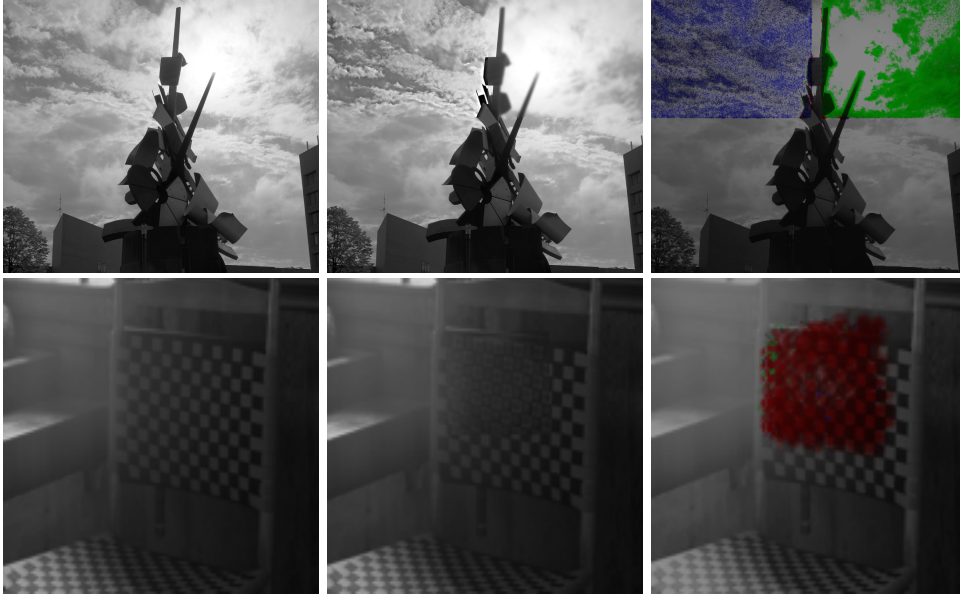


FIG. 4.1. *Distortion map illustration. Left: reference images. Middle: test images. Right: distortion maps.*

$$(4.3) \quad E_k(L, H) := \|DRIM(L, H)\|_{L^k}$$

In others words, we have

$$(4.4) \quad E_k(L, H) = \int_{\Omega} (LVC^k(L, H)(x) + AIC^k(L, H)(x) + INV^k(L, H)(x))^{1/k} dx$$

Note that both definitions of metric and distance are coherent with the definitions 2.2 and 2.3 in Sect. 2.

Moreover, the distances  $E_k$  are symmetric since  $LVC(L, H) = AIC(H, L)$  and  $INV(L, H) = INV(H, L)$  according to formula (4.1).

**4.2. Reducing a dynamic range independent perceptual distance between an HDR image and its tone mapped version.** In this Section, we detail the gradient descent algorithm mentioned in the end of Sect. 2 in the context of the distances (4.3).

**4.2.1. Expression of the gradient of the distance.** We first show the existence of a solution of the minimization problem related with the distances (4.3).

**PROPOSITION 4.3.** *Let  $H$  be an HDR image. Assuming that the metric  $DRIM$  of Definition 4.1 is continuous, the variational problem*

$$(4.5) \quad \arg \min_{L \in LDR(\Omega)} E_k(L, H)$$

*has a solution.*

*Proof.* Each distortion being in the range  $[0, 1]$ , the  $L^k$  norms being continuous, and the space  $LDR(\Omega)$  being closed, Proposition 4.3 appears to be a particular case of Proposition 2.4.  $\square$

PROPOSITION 4.4. *Assuming that  $H$  is non constant and the metric DRIM is differentiable on the space  $LDR(\Omega)$ , the critical points of the energy*

$$(4.6) \quad E_{kH}: L \longrightarrow E_k(L, H)$$

*satisfy*

$$(4.7) \quad \int_{\Omega} \frac{LVC(L, H)^{k-1} \frac{\partial LVC(L, H)}{\partial L(y)} + AIC(L, H)^{k-1} \frac{\partial AIC(L, H)}{\partial L(y)} + INV(L, H)^{k-1} \frac{\partial INV(L, H)}{\partial L(y)}}{(LVC(L, H)^k + AIC(L, H)^k + INV(L, H)^k)^{1-1/k}} d\Omega = 0$$

$\forall y \in \Omega$ .

The proof of the Proposition relies upon the following postulate: **the metric DRIM only vanishes when the inputs are two identical constant images.**

Even if the postulate is counterintuitive, it holds because the metric DRIM first treats the two images separately. More precisely, it estimates the probabilities that the contrast is visible or not in each image and then combines the results in order to determine the distortions  $LVC, AIC, INV$  between the two images. Hence, even if the images are identical, the fact that the images are first treated separately makes the distortions be 0 only if the probability that the contrast is visible is 0 at each pixel of both images, i.e. if both images are constant. All the experiments we have run confirm this analysis.

*Proof.* The energy  $E_{kH}$  is differentiable on the whole set  $LDR(\Omega)$  since the metric DRIM does not vanish for non constant images (postulate). Then, according to Prop. 2.5, the critical points of (4.6) satisfy

$$(4.8) \quad \int_{\Omega} \delta \| \|_k \left( DRIM(L, H)(x); \frac{\delta DRIM(L, H)(x)}{\delta L(y)} \right) dx = 0 \quad \forall y \in \Omega$$

Finally, expression (4.7) follows from (4.8) and

$$\delta \| \|_k((u_1, u_2, u_3); (v_1, v_2, v_3)) = \frac{u_1^{k-1}v_1 + u_2^{k-1}v_2 + u_3^{k-1}v_3}{\|(u_1, u_2, u_3)\|_k^{k-1}}$$

$\square$

The term  $\nabla E_{kH}(L): y \mapsto$

$$(4.9) \quad \int_{\Omega} \frac{LVC(L, H)^{k-1} \frac{\partial LVC(L, H)}{\partial L(y)} + AIC(L, H)^{k-1} \frac{\partial AIC(L, H)}{\partial L(y)} + INV(L, H)^{k-1} \frac{\partial INV(L, H)}{\partial L(y)}}{(LVC(L, H)^k + AIC(L, H)^k + INV(L, H)^k)^{1-1/k}} d\Omega$$

is the gradient of the energy  $E_{kH}$  at the function  $L$ .

**4.2.2. The discrete gradient descent algorithm.** The discrete approximation of the gradient descent algorithm in the context of the distances (4.3) is of the form

$$(4.10) \quad L_{n+1} = L_n - \alpha_n \nabla^D E_{kH}(L_n), \quad L_0 = TMO(H)$$

In what follows, we detail the expressions of the discrete gradient  $\nabla^D E_{kH}$ , the initial condition  $L_0$  and the steps  $\alpha_n$ .

*The discrete gradient operator.* We express the discrete gradient  $\nabla^D E_{kH}$  as

$$(4.11) \quad \nabla^D E_{kH}(L): (i, j) \mapsto \sum_{(a,b) \in \Omega} \frac{LVC(L, H)^{k-1} \frac{\partial LVC^D(L, H)}{\partial L(i, j)} + AIC(L, H)^{k-1} \frac{\partial AIC^D(L, H)}{\partial L(i, j)} + INV(L, H)^{k-1} \frac{\partial INV^D(L, H)}{\partial L(i, j)}}{(LVC(L, H)^k + AIC(L, H)^k + INV(L, H)^k)^{1-1/k}}$$

Note that we have omitted the dependence of the terms on the pixels  $(a, b)$  for the sake of shortness.

The discrete functional derivatives of the distortions  $LVC, AIC, INV$  with respect to  $L$  at the pixels  $(i, j)$  are computed using formula (2.12). For instance, dealing with the distortion  $LVC$ , we have

$$(4.12) \quad \frac{\partial LVC^D(L, H)(a, b)}{\partial L(i, j)} : = \frac{LVC(L + \epsilon_1(i, j) \delta_{(i, j)}, H)(a, b) - LVC(L - \epsilon_2(i, j) \delta_{(i, j)}, H)(a, b)}{d(L + \epsilon_1(i, j) \delta_{(i, j)}, L - \epsilon_2(i, j) \delta_{(i, j)})}$$

and the values  $\epsilon_1, \epsilon_2, d$  are determined by (3.14), (3.15) and (3.16) respectively.

*The step.* The images produced at each iteration of the gradient descent (4.10) are not necessarily LDR images, but as the metric DRIM is designed for images of any dynamic range, the numerical scheme is well defined if the values of  $L_{n+1}$  are in the range  $[0, 1]$  at each iteration which can be achieved if  $\alpha_n$  is chosen small enough.

We then determine  $\alpha_n$  through a line search strategy that will force the values of the image to be in the range  $[0, 1]$ . The pseudo-code for the computation of  $\alpha_n$  is the following

```

Initializations:  $\lambda = 0, E_{new} = E_{kH}(L_n), E_{old} = E_{kH}(L_n) + 1, L_{temp} = L_n$ 
while  $E_{new} < E_{old}$  and  $L_{temp}(i, j) \in [0, 1] \forall (i, j) \in \Omega$  do
 $\lambda = \lambda + 0.001$ 
 $E_{old} = E_{new}$ 
 $L_{temp} = L_n - \lambda \nabla^D E_{kH}(L_n)$ 
 $E_{new} = E_{kH}(L_{temp})$ 
end while
 $\alpha_n = \lambda - 0.001$ 

```

*Initial condition and output image of the gradient descent.* As the perceptual metric involved is the metric DRIM, the function  $L_0$  should be the light intensity produced by a screen when displaying the tone mapped image  $TMO(H)$ , which can be obtained as follows.

Denoting by  $R_0, G_0, B_0$  the inverse gamma corrected components (see Sect. 3.1) of the tone mapped image  $TMO(H)$  in the sRGB color space, we transform  $TMO(H)$  into XYZ color coordinates with the formula

$$(4.13) \quad \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} = \begin{pmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9502 \end{pmatrix} \begin{pmatrix} R_0 \\ G_0 \\ B_0 \end{pmatrix}$$

As mentioned in Sect. 3.1, the component  $Y_0$  is an accurate approximation of the light intensity produced when displaying a color image on a monitor. Hence, we take  $Y_0$  as the initial condition  $L_0$  of the gradient descent (4.10).

In order to make the final output of our algorithm be an LDR image, we project the steady-state  $L_*$  of the gradient descent onto the discrete space of LDR images  $LDR(\Omega)$  as follows.

Assuming that the domain  $\Omega$  of the HDR image  $H$  is of size  $M \times N$ , we define the discrete spaces  $HDR(\Omega)$  and  $LDR(\Omega)$  as matrices of the form

$$(4.14) \quad HDR(\Omega) = \{A \in \mathcal{M}(M, N), A_{i,j} = \frac{k}{65536}, k \in [1, 65536] \cap \mathbb{N}\}$$

$$(4.15) \quad LDR(\Omega) = \{A \in \mathcal{M}(M, N), A_{i,j} = \frac{k}{256}, k \in [1, 256] \cap \mathbb{N}\}$$

Then, the projection  $P^D$  onto the set  $LDR(\Omega)$  consists, for  $A \in \mathcal{M}(M, N)$  with values in the range  $[0, 1]$ , in clipping and quantizing into the usual 8 bits domain as

$$(4.16) \quad P^D A_{i,j} = \begin{cases} \frac{1}{256} & \text{if } A_{i,j} \leq \frac{1}{256} \\ \frac{\lfloor 256 \times A_{i,j} \rfloor}{256} & \text{if } \frac{1}{256} \leq A_{i,j} \leq 1 \end{cases}$$

where  $\lfloor \cdot \rfloor$  is the floor function.

Finally, the output of the algorithm is the color image of components  $(R_*, G_*, B_*)$  defined by

$$\begin{pmatrix} R_* \\ G_* \\ B_* \end{pmatrix} = \begin{pmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9502 \end{pmatrix}^{-1} \begin{pmatrix} X_0 \\ P^D L_* \\ Z_0 \end{pmatrix}$$

**4.2.3. Preprocessing.** To increase the chance that the gradient descent algorithm (4.10) does not stop at a local minimum of the energy  $E_{kH}$  too close to the initial condition  $L_0$ , we apply a preprocessing on  $L_0$  in order to get an initial condition  $L_{new}$  of the algorithm (4.10) that is closer (in terms of perceptual distance) to  $H$  than  $L_0$ . The method we propose relies on the intuition that high values of LVC might be reduced by application of local sharpening whereas high values of AIC might be reduced by local Gaussian blurring. Hence we perform local Gaussian blurring and unsharp masking (4.17) (see [1] for details) to  $L_0$  depending on the values of the function  $LVC(L_0, H) - AIC(L_0, H)$ . More precisely, denoting by  $L_{smooth}$  a blurred version of  $L_0$  and defining  $L_{sharp}$  as

$$(4.17) \quad L_{sharp} = L_0 + \alpha (L_0 - L_{smooth})$$

for some constant  $\alpha$ , we define the image  $L_{new}$  as

$$(4.18) \quad L_{new}(i, j) = \begin{cases} L_{sharp}(i, j) & \text{if } LVC(L_0, H) - AIC(L_0, H)(i, j) > 0 \\ (1 - \beta)L_0(i, j) + \beta L_{smooth}(i, j) & \text{if } LVC(L_0, H) - AIC(L_0, H)(i, j) < 0 \end{cases}$$

Experiments on different tone mapped images described in the next Section show that the distance with  $H$  is indeed reduced as well as the errors in the distortions map.

**4.2.4. Numerical scheme.** We summarize in this Section the different steps of our algorithm.

1. Let  $H$  be an HDR image and  $L_0$  be a tone mapped version of  $H$ . As the metric DRIM only takes into account the luminance information of the input images, we first convert them into luminance maps: we apply the transformation (4.13) on the LDR image  $L_0$  and extract the luminance channel  $Y_0$ , and use the function *pfs\_read\_luminance*[17] to extract the luminance information of  $H$ .
2. We perform the preprocessing described in Sect. 4.2.3 on the image  $Y_0$  with the following parameters: the variance  $\sigma$  of the Gaussian smoothing kernel is set to 0.62, and the constants  $\alpha, \beta$  are respectively 0.7 and 0.5. These values provide good results and have been fixed for all the experiments in this paper.
3. We apply the gradient descent algorithm (4.10) where the initial condition is the output of the preprocessing. We test different values for  $k$  and different domains for the summation in the expression of the discrete gradient operator (4.11). Indeed, because the variational problem we propose is non local, the gradient operator is an integral operator, meaning that its computation might be very time-consuming. In order to decrease the execution time of the algorithm, we adopt the following two approaches: we consider  $50 \times 50$  neighborhoods, and make use of computers equipped with multiple cores to compute the gradient operator over the whole image domain. The algorithm stops when the energy does not decrease anymore, i.e. when we reach a local or global minimum.

The pseudo code of the gradient descent algorithm is given below.

```

while  $E_{new} < E_{old}$  do
  for each pixel  $p$  do
     $L_n^+ = L_n + \epsilon_1(p)\delta_p$ 
     $L_n^- = L_n - \epsilon_2(p)\delta_p$ ,  $\triangleright \epsilon_1, \epsilon_2$  from (3.14) and (3.15)
     $(LVC^{+/-}, AIC^{+/-}, INV^{+/-}) = DRIM(L_n^{+/-}, H)$ ,
     $Diff_{type} = (type^+ - type^-)/d(p)$ ,  $type \in Type = \{LVC, AIC, INV\}$   $\triangleright d$  from (3.16)
     $gradient(p) = \sum_X (\sum_{Type} type^{k-1} \times Diff_{type}) / (\sum_{Type} type^k)^{1-1/k}$ 
     $\triangleright X$  is either  $50 \times 50$  or full image domain
  end for
   $L_{n+1} = L_n - \alpha_n \times gradient$ ,  $\triangleright \alpha_n$  determined in Sect. 4.2.2
   $(LVC, AIC, INV) = DRIM(L_{n+1}, H)$ 
   $E_{old} = E_{new}$ 
   $E_{new} = \sum_X (LVC^k + AIC^k + INV^k)^{1/k}$ 
end while
 $L_* = L_n$   $\triangleright L_*$  steady-state of the gradient descent
Return  $P^D(L_*)$   $\triangleright P^D$  from (4.16)

```

4. The final output LDR color image is then obtained by combining the  $X_0$  and  $Z_0$  components of the output of the preprocessing with the output  $P^D L_*$  of the gradient descent algorithm.

**4.3. Experiments.** We test our algorithm dedicated to optimize tone mapped images on different HDR images taken from the MPI [20] and Fairchild [9] databases, and tone mapped images produced by the TMOs of Ferradans et al. [11], Drago et al. [7], Reinhard et al. [24], and Mantiuk et al. [18].

The evaluation of our algorithm is two-fold: global and pixel-wise. As a pixel-wise measure, we compare the distortion maps (see details in Sect. 4.1) of the initial condition and output. As a global measure, we compare their averaged perceptual distance with  $H$ , making use of the following energy

$$(4.19) \quad E(L, H) = \frac{1}{|\Omega|} \sum_{(a,b) \in \Omega} \|DRIM(L, H)(a, b)\|_2$$

The first experiment we propose consists in evaluating the preprocessing. In Fig. 4.2, we show the input tone mapped images, output color images of the preprocessing, as well as their distortion maps. We have applied the formula (4.18) to the luminance channel of tone mapped images obtained with the method of Ferradans et al. [11] (top row image) and Drago et al. [7] (bottom row image). The HDR source images are taken from the MPI database [20]. We observe that the LVC distortion has been reduced, whereas the INV distortion has increased a bit. As we can see in formula (4.18), the preprocessing is only devoted to reduce the LVC and AIC distortions, and does not take into account the INV distortion. Hence, some choices of the parameters  $\alpha, \beta$  might yield an increase of the INV distortion.

On Table 4.1, we present results of the preprocessing (amongst other results) tested on images of the Fairchild database [9] for the TMOs aforementioned. Note that the images have been rescaled to  $200 \times 200$  pixels in order to speed up the gradient descent algorithm. Average results have been computed over 10 images of the dataset. The results confirm that the preprocessing reduces the perceptual distance with respect to the HDR source image.

In the second experiment, we evaluate the final output of our method described above for different perceptual distances (4.3) parametrized by a strictly positive number  $k$ . Table 4.1 shows the distance (4.19) of the initial tone mapped images and output images with a given HDR image for the following values of the parameter  $k$  : 0.8, 1, 1.2, 2, 5 and 50. The summation for the computation of the gradient operator (4.11) has been done on  $50 \times 50$  neighborhoods. The results show that the minimum distance is always reached for  $k$  close to 1 ( $k = 0.8$  in most of the cases), and the distance increases when  $k$  gets away from this optimal  $k$  value. Note that the same behavior occurs when  $k$  is lower than the optimal value but we do not show the result in this paper. We would like to point out that these results were not expected.

In Fig 4.3, we show a region of the output color images obtained with  $k = 0.8, 1$  and 5 along with their distortion maps. The HDR source is the image “snow” from the MPI database [20]. The input tone mapped image is provided by the TMO of Mantiuk et al. [18]. Taking a close look, we can see in the distortion maps that there is an increase in AIC (blue patches) and INV (red patches) with the increase in  $k$  value, which is coherent with the fact that the distance increases with the value of  $k$ . However, visual differences are almost impossible to perceive for such close values of  $k$ .





FIG. 4.2. *Evaluation of preprocessing stage for Ferradans et al.[11] (top row) and Drago et al. (bottom row) TMOs. From left to right: input tone mapped image, output of preprocessing stage, distortion map of the input, distortion map of the output.*



FIG. 4.3. *Comparison of output with different  $k$  values. Top left: final output with  $k = 0.8$ . Top middle: final output with  $k = 1.2$ . Top right: final output with  $k = 5$ . Bottom row: corresponding distortion maps*

In the third experiment, we compare the output of the preprocessing stage with the final output of our method using  $50 \times 50$  neighborhoods and parameter  $k = 0.8$ , as well as their distortions maps (see Fig 4.4). The HDR source is the image “Peck-Lake” from the Fairchild database, and the input tone mapped image is provided by the TMO of Drago et al. [7]. Its corresponding distortion map reveals a great loss of contrast (green patches). We observe that the preprocessing stage reduces such a distortion, and the gradient descent algorithm applied to the output of the preprocessing reduces it to even greater extent. These improvements are confirmed when computing the perceptual distances (4.19) at each stage with the HDR source image: initial (0.639), preprocessing (0.509), and final (0.389). There is a 23% reduction in distance in the final image compared to the preprocessed image. From this result and the ones shown in Table 4.1 we can claim that applying the gradient descent algorithm to the output of the preprocessing provides much better results than applying only the preprocessing.

Fig 4.5 shows some regions of images on Fig. 4.4 and “BarHarborPresunrise” from the Fairchild database. We observe that the preprocessing and the final output have

TABLE 4.1  
Distance with different  $k$  values.

Image \ TMO		Drago et al.[7]	Reinhard et al.[24]	Mantiuk et al.[18]	Ferradans et al.[11]
AmikBeav DamPM1	Initial	0.726	0.744	0.763	0.745
	Preprocess	0.573	0.611	0.66	0.702
	Final ( $k = 0.8$ )	0.464	0.515	0.59	0.546
	Final ( $k = 1$ )	0.469	0.517	0.59	0.545
	Final ( $k = 1.2$ )	0.474	0.519	0.59	0.546
	Final ( $k = 2$ )	0.494	0.534	0.6	0.558
	Final ( $k = 5$ )	0.543	0.57	0.624	0.559
Barharbor Presun	Initial	0.662	0.665	0.533	0.561
	Preprocess	0.53	0.552	0.493	0.511
	Final ( $k = 0.8$ )	0.417	0.454	0.446	0.46
	Final ( $k = 1$ )	0.416	0.452	0.442	0.457
	Final ( $k = 1.2$ )	0.418	0.453	0.444	0.454
	Final ( $k = 2$ )	0.433	0.465	0.447	0.462
	Final ( $k = 5$ )	0.469	0.508	0.47	0.490
Average (10 images)	Initial	0.645	0.657	0.573	0.590
	Preprocess	0.527	0.575	0.515	0.536
	Final ( $k = 0.8$ )	0.417	0.460	0.457	0.460
	Final ( $k = 1$ )	0.420	0.461	0.457	0.460
	Final ( $k = 1.2$ )	0.426	0.464	0.457	0.462
	Final ( $k = 2$ )	0.442	0.476	0.467	0.471
	Final ( $k = 5$ )	0.475	0.511	0.481	0.494
	Final ( $k = 50$ )	0.490	0.529	0.497	0.511

higher contrast than the original tone mapped images which means that the LVC distortion has been reduced. Note that the results on top-row are coherent with the distortion maps in Fig. 4.4.

At last, we analyze the impact of modifying the neighborhood size used for the summation in the expression of the gradient operator (4.11). We test our algorithm for  $50 \times 50$  as well as  $200 \times 200$  neighborhoods (the whole image domain). The parameter  $k$  has been set to 0.8, which is the value giving the best results in the case of  $50 \times 50$  neighborhoods according to the second experiment. In Fig 4.6, we show the output color images obtained with the two neighborhood sizes along with their distortion maps. The HDR source images are “BarHarborPresunrise” and “PeackLake” from the Fairchild database. The input tone mapped image is provided by the TMO of Reinhard et al. [24] and Drago et al. [7]. The distortion maps show that using the whole image domain substantially reduces the LVC distortion (less green patches appear in the distortion map), and reduces (in less proportion) the INV distortion (less red patches). This improvement is due to the fact that the metric DRIM uses the information of the whole image to estimate the distortions at each pixel (see details in Sect. 4.1). Then, restricting the domain where we compute the gradient operator yields some information loss.

Table 4.2 compares the distance between the final output images (using the aforementioned domains) with the HDR source image. We can see that the results obtained by using a  $200 \times 200$  neighborhood have smaller error than the results achieved by using a  $50 \times 50$  neighborhood, and this numerical behavior is consistent with the visual distortion maps presented in Fig 4.6.

In Fig 4.7, we show some results of our algorithm with a  $200 \times 200$  neighborhood

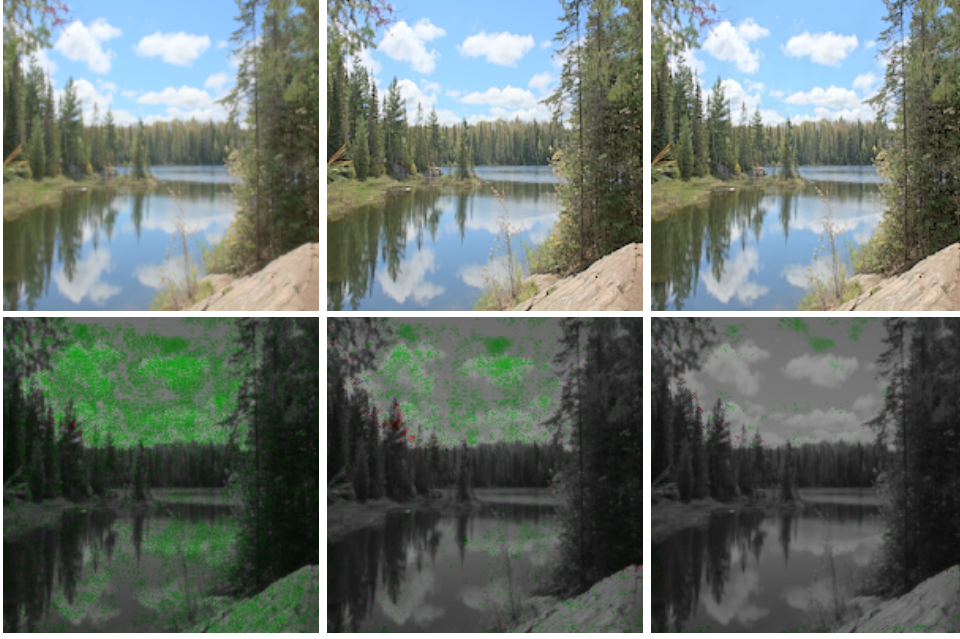


FIG. 4.4. Comparison between the output of preprocessing stage and the final stage. Top left: input tone mapped image of Drago et al. [7]. Top middle: output of preprocessing stage. Top right: final output with  $k = 0.8$ . Bottom row: corresponding distortion maps.



FIG. 4.5. Visual comparisons between the output images of the different stages. Top left: “PeacLake” tone mapped by Drago et al. [7]. Bottom left: “BarHarborPresun” tone mapped by Drago et al. [7]. Middle: output of preprocessing stage. Right: final output with  $k = 0.8$ .

tested on the different TMOs mentioned above applied to the same HDR image. By a close observation of the output images, we can notice an enhancement of details of the initial tone mapped images which is confirmed by comparing the corresponding distortion maps (reduction of LVC distortion).

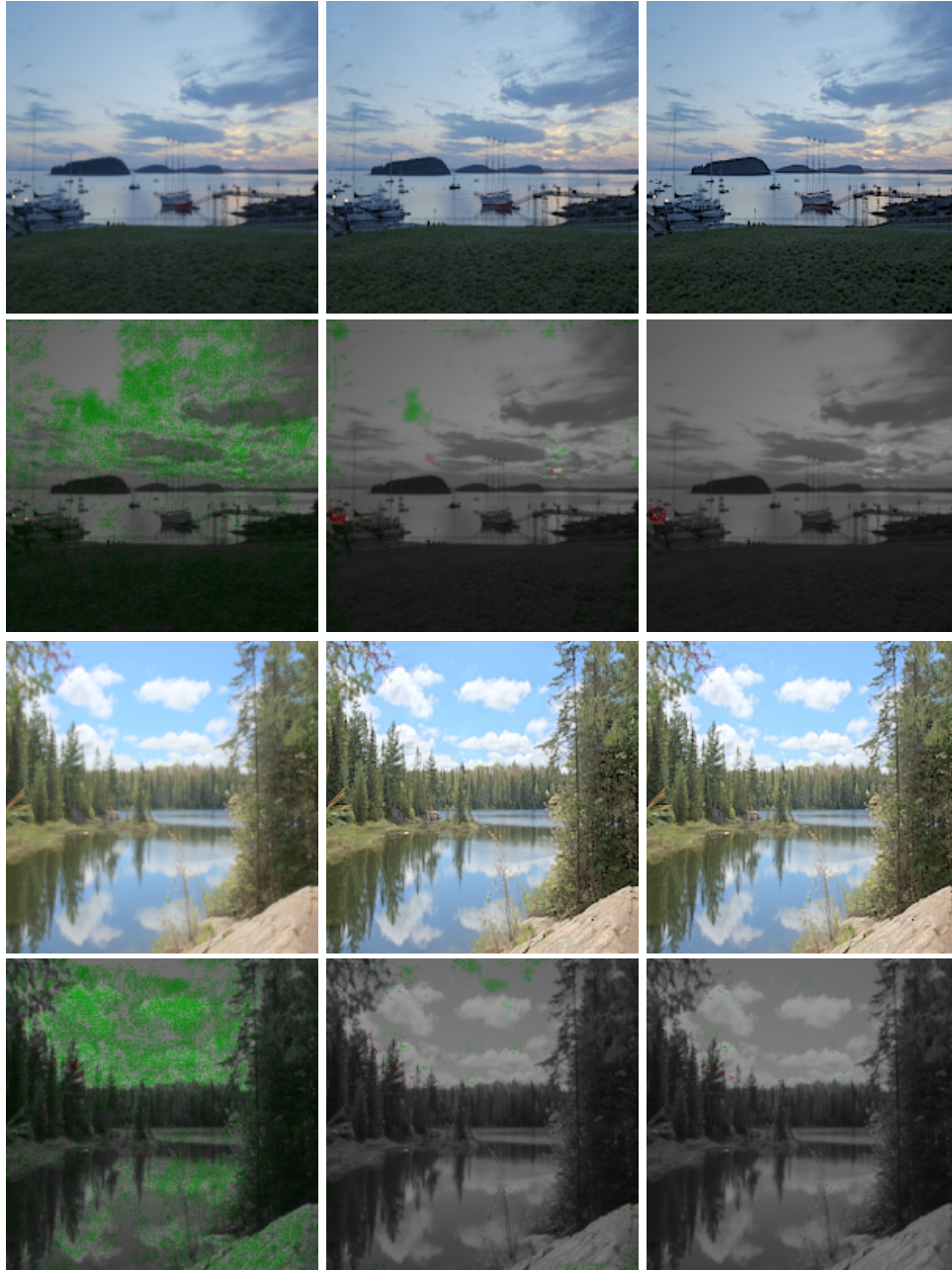


FIG. 4.6. Comparison between the final output with  $50 \times 50$  and  $200 \times 200$  neighborhoods. First row left: input tone mapped image of Reinhard et al. [24], third row left: input tone mapped image of Drago et al. [7]. Middle: final output with  $50 \times 50$  neighborhood and their distortions maps. Right: final output with  $200 \times 200$  neighborhood and their distortions maps.



TABLE 4.2

Distance at the final stage of our method with  $50 \times 50$  and  $200 \times 200$  neighborhoods. Percentile improvement at each stage with respect to the original tone mapping result is given in brackets.

Image \ TMO		Drago et al.[7]	Reinhard et al.[24]	Mantiuk et al.[18]	Ferradans et al.[11]
AmikBeav DamPM1	Initial	0.726	0.744	0.763	0.745
	Preprocess	0.573 (21%)	0.611 (13%)	0.66 (14%)	0.702 (6%)
	Final ( $50 \times 50$ )	0.464 (36%)	0.515 (31%)	0.59 (23%)	0.546 (27%)
	Final ( $200 \times 200$ )	0.264 (64%)	0.356 (52%)	0.47 (38%)	0.388 (48%)
Barharbor Presun	Initial	0.662	0.665	0.533	0.561
	Preprocess	0.53 (20%)	0.552 (17%)	0.493 (7%)	0.511 (9%)
	Final ( $50 \times 50$ )	0.417 (37%)	0.454 (32%)	0.446 (16%)	0.46 (18%)
	Final ( $200 \times 200$ )	0.206 (69%)	0.248 (63%)	0.352 (34%)	0.34 (39%)

**5. Conclusion.** Based on perceptual metrics that measure distortions between images of different dynamic range, we propose in this paper a method to improve tone mapped images based on a non local variational problem. We tested this approach with the metric DRIM [3] for different tone mapped images provided by several tone mapping operators. The experiments show that our approach improves the tone mapped images tested in the sense that it reduces their perceptual distance to their HDR source. Our method provides an average reduction of the distance by more than 25%.

Further work will be devoted to apply the proposed framework to contexts where minimization of a perceptual distance could also be useful. One such application may be to optimize gamut mapping methods by considering a perceptual metric that measures color distortions between images. We are also investigating improvements of our mathematical model. Indeed, we were using in this paper a gradient descent approach in order to reduce perceptual distances between images. However, this approach suffers from two main issues when dealing with tone mapping: first, the gradient descent algorithm might stop at local minima that are not global and that are close to the initial condition; second, the representation of LDR and HDR images as smooth functions is not very realistic since we are not taking into account the specificity of their dynamic ranges: the use of other functional spaces might be more appropriate and lead to better algorithms.

**6. Acknowledgment.** The authors would like to thank the anonymous reviewers for their helpful comments and suggestions to improve the final version of the paper. Many thanks to Tunc Aydin, Karol Myszkowski, Rafal Mantiuk and Hans-Peter Seidel for their help.

## REFERENCES

- [1] J.E. ADAMS, A.T. DEEVER, E.O. MORALES, B.H. PILLMAN, *Perceptually based image processing algorithm design. Perceptual Digital Imaging: Methods and Applications*, (2012), pp. 123-166.
- [2] M. ASHIKHMIN, *A tone mapping algorithm for high contrast images*, Proc. Eurographics Workshop Rendering, (2002), pp. 1-11.
- [3] T.O. AYDIN, R. MANTIUK, K. MYSZKOWSKI, AND H.-P. SEIDEL, *Dynamic range independent image quality assessment*, Proc. ACM SIGGRAPH 2008, (2008), pp. 1-10.
- [4] P. CYRIAC, T. BATARD, AND M. BERTALMÍO, *A variational method for the optimization of tone mapping operators*, Proc. 6th Pacific-Rim Symposium on Image and Video Technology PSIVT 2013, R. Klette et al. Eds., LNCS 8333, (2014), pp. 505-516.

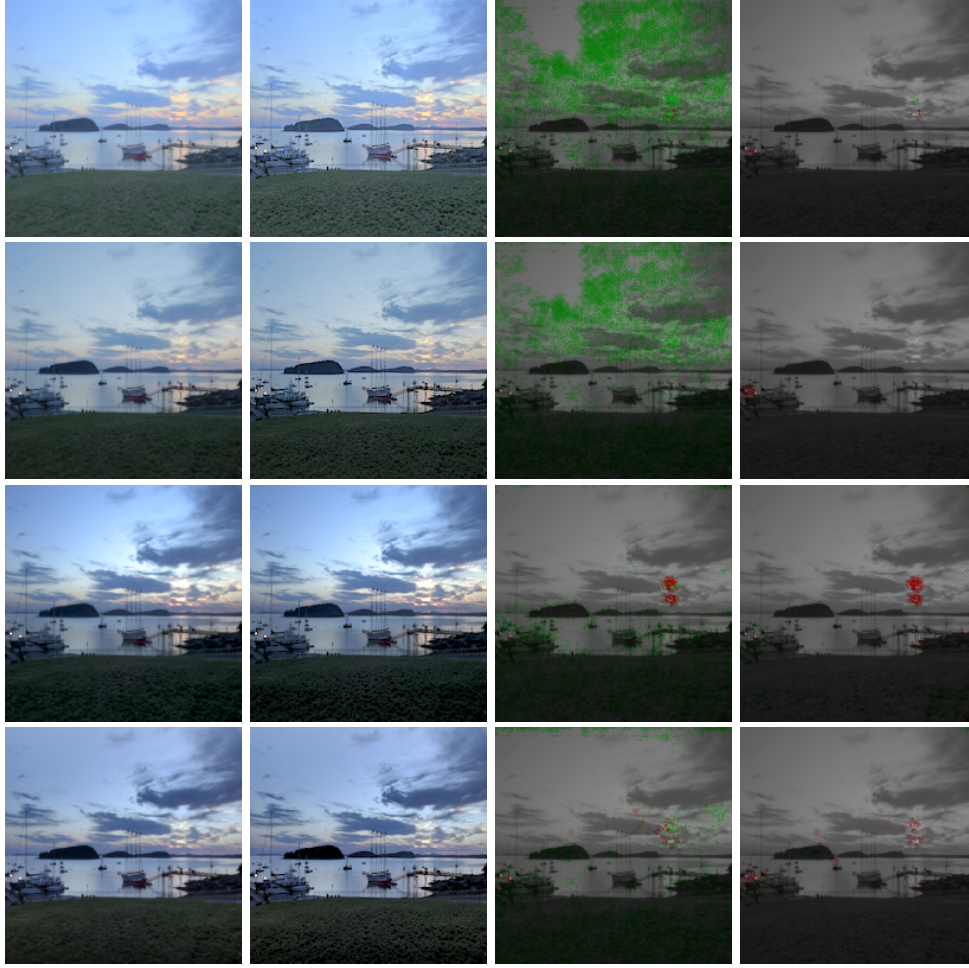


FIG. 4.7. The final output. First column: input tone mapped images (TMOs from top to bottom: Drago et al.[7], Reinhard et al.[24], Mantiuk et al.[18], Ferradans et al. [11]). Second column: final output images with a  $200 \times 200$  neighborhood. Third column: distortion maps of input tone mapped images. Fourth column: distortion maps of final outputs. See Table. 4.2 (image “BarharborPresun”) for the corresponding distances.

- [5] S. DALY, *The visible differences predictor: an algorithm for the assessment of image fidelity*, Digital Images and Human Vision. MIT Press, A.B. Watson, Eds., (1993), pp. 179-206.
- [6] P. E. DEBEVEC, J. MALIK, *Recovering high dynamic range radiance maps from photographs*, ACM SIGGRAPH, (1997), pp. 369-378.
- [7] F. DRAGO, K. MYSZKOWSKI, T. ANNEN, N. CHIBA, *Adaptive logarithmic mapping for displaying high contrast scenes*, Computer Graphics Forum, 22(3), (2003), pp. 419-426.
- [8] F. DURAND, J. DORSEY, *Fast bilateral filtering for the display of high dynamic range images*, Proc. ACM SIGGRAPH, (2002), pp. 257-266.
- [9] M.D. FAIRCHILD, *Color Appearance Models*, Wiley and Sons, 3rd Edition (2013).
- [10] R. FATTAL, D. LISCHINSKI, M. WERMAN, *Gradient domain high dynamic range compression*, Proc. ACM SIGGRAPH, (2002), pp. 249-256.
- [11] S. FERRADANS, M. BERTALMÍO, E. PROVENZI, AND V. CASELLES, *An analysis of visual adaptation and contrast perception for tone mapping*, IEEE Trans. Pattern Anal. Mach. Intell., 33(10) (2011), pp. 2002-2012.
- [12] J. FERWERDA, S. PATTANAIK, P. SHIRLEY, D. GREENBERG, *A model of visual adaptation for realistic image synthesis*, Proc. ACM SIGGRAPH, (1996), pp. 249-258.

- [13] J. KUANG, G. M. JOHNSON, M. FAIRCHILD, *iCAM06: A refined image appearance model for HDR image rendering*, J. Visual Comm. and Image Representation, 18, (2007), pp. 406-414 .
- [14] P. LEDDA, A. CHALMERS, T. TROSCIANKO, H. SEETZEN, *Evaluation of tone mapping operators using high dynamic range display*, ACM SIGGRAPH, (2005), pp. 640-648.
- [15] R. MANTIUK, S. DALY, K. MYSZKOWSKI, AND H.-P SEIDEL, *Predicting visible differences in high dynamic range images - model and its calibration*, Human Vision and Electronic Imaging X, SPIE Proceedings Serie, 5666, (2005), pp. 204-214.
- [16] R. MANTIUK, K. MYSZKOWSKI, H.-P. SEIDEL, *A perceptual framework for contrast processing of high dynamic range images*, ACM Trans. Applied Perception, 3(3), (2006), pp. 286-308
- [17] R. MANTIUK, G. KRAWCZYK, R. MANTIUK, H.-P. SEIDEL, *High dynamic range imaging pipeline: perception-motivated representation of visual content.*, Human Vision and Electronic Imaging XII, SPIE Proceedings, 6492, (2007)
- [18] R. MANTIUK, S. DALY, L. KEROFISKY, *Display adaptative tone mapping*, Proc. ACM SIGGRAPH, (2008), 68.
- [19] R. MANTIUK, K.J. KIM, A.G. REMPEL, AND W. HEIDRICH, *HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions*, ACM Trans. Graphics, 30(4) (2011), 40.
- [20] MPI database: <http://www.mpi-inf.mpg.de/resources/hdr/gallery.html>
- [21] S. PATTANAIK, J. TUMBLIN, H. YEE, D. GREENBERG, *Time-dependent visual adaptation for fast realistic image display*, Proc. ACM SIGGRAPH, (2000), pp. 47-54.
- [22] C. POYNTON, *Digital Videos and HDTV: Algorithms and Interfaces*, Morgan Kaufmann, (2003).
- [23] W. PRAT, *Digital Image Processing.*, John Wiley & Sons, (2000).
- [24] E. REINHARD, K. DEVLIN, *Dynamic range reduction inspired by photoreceptor physiology*, IEEE Trans. Visualization and Computer Graphics, 11(1), (2005), pp. 13-24.
- [25] E. REINHARD, G. WARD, S.N. PATTANAIK, P.E. DEBEVEC, W. HEIDRICH, AND K. MYSZKOWSKI, *High Dynamic Range Imaging - Acquisition, Display, and Image-Based Lighting*. Morgan Kaufmann, (2010).
- [26] K. SMITH, G. KRAWCZYK, K. MYSZKOWSKI, AND H.-P. SEIDEL, *Beyond tone mapping: enhanced depiction of tone mapped HDR images*, Computer Graphics Forum, 25(3) (2006), pp. 427-438.
- [27] J. TUMBLIN, H. RUSHMEIR, *Tone reproduction for computer generated images*, IEEE Computer Graphics and Applications, 13(6), (1993), pp. 42-48.
- [28] Z. WANG, AND A.C. BOVIK, *A universal image quality index*, IEEE Signal Processing Letters, 9(3), (2002), pp. 81-84.
- [29] G. WARD, H. RUSHMEIER, C. PIATKO, *A visibility matching tone reproduction operator for high dynamic range scenes*, IEEE Trans. Visualization and Computer Graphics, 3(4), (1997), pp. 291-306
- [30] A. WATSON, *The Cortex transform: rapid computation of simulated neural images*, Comp. Vision Graphics and Image Processing, 39, (1987), pp. 311-327.
- [31] A. YOSHIDA, V. BLANZ, K. MYSZKOWSKI, H.P SEIDEL, *Perceptual evaluation of tone mapping operators with real-world scenes*, International Society for Optics and Photonics, (2005), pp. 192-203.