High Dynamic Range Image Watermarking Robust Against Tone-Mapping Operators

Fabrizio Guerrini*[1], Masahiro Okuda[2], Nicola Adami[1] and Riccardo Leonardi[1]

[1]University of Brescia, Italy

Signals & Communications Laboratory - Department of Information Engineering

Email: {firstname.surname}@ing.unibs.it

[2]University of Kitakyushu, Japan

Department of Information and Media Sciences

Email: okuda-m@env.kitakyu-u.ac.jp

EDICS: WAT-BINM, WAT-OTHA

High Dynamic Range Image Watermarking Robust Against Tone-Mapping Operators

Abstract

High Dynamic Range (HDR) images represent the future format for digital images since they allow to accurately render a wider range of luminance values. However, nowadays special types of pre-processing collectively known as tone-mapping (TM) operators are needed to adapt HDR images to currently existing displays. Tone-mapped images, although of reduced dynamic range, have nonetheless high quality and hence retain some commercial value. In this paper we propose a solution to the problem of HDR image watermarking, *e.g.*, for copyright embedding, that should survive tone-mapping. Therefore, the requirements imposed on the watermark encompass imperceptibility, a certain degree of security and robustness to tone-mapping operators. The proposed watermarking system belongs to the blind, detectable category, it is based on the QIM paradigm and employs a higher order statistics as feature. Experimental analysis shows positive results and demonstrates the system effectiveness with current state-of-art TM algorithms.

EDICS: WAT-BINM, WAT-OTHA

I. Introduction

High Dynamic Range (HDR) images represent the radiance of scenes captured by a device or generated by an artificial rendering system. Due to the wide dynamics of visible luminance, their pixels usually take floating point values and thus HDR images cannot be directly rendered by present-day displays. In fact, the original values must be adapted in order to fit the dynamic and color gamut of the target device. This adaptation, which transforms a "scene referred" image into an "output referred" one, can be obtained by applying a so called tone-mapping (TM) process. Tone-mapping techniques are all non-linear in nature and could vary depending on the algorithm: their complexity can range from something as simple as some kind of global transformation (*e.g.* clipping followed by histogram equalization) to more complex non-linear local processing.

Tone-mapped images keep some commercial and intellectual value because of their high visual quality despite their limited dynamic range and hence they are possible targets for misappropriation or misuse by malevolent entities. This is an unavoidable problem which will have to be faced as soon as HDR images and their tone-mapped versions reach widespread diffusion - and they certainly will in the near future.

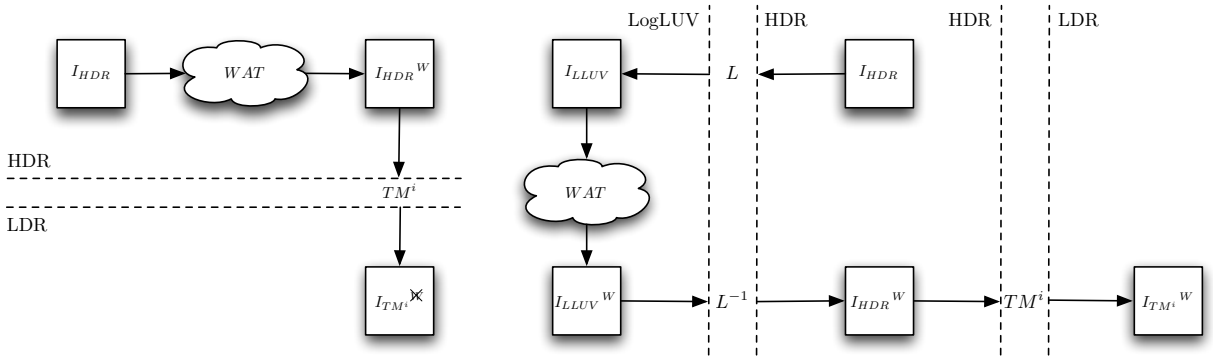
A possible solution is represented by data hiding techniques, such as digital watermarking [1] [2], which try to embed into any given HDR image some kind of information that should have to stay therein even after a tone-mapping processing. To the authors' knowledge, no major work concerning the problem of HDR image watermarking robust to tone-mapping has been widely studied yet.

In this paper, we therefore propose to embed a watermark to enforce or simply communicate property rights of a given HDR image to any potential user. The watermark needs only to survive to those processing that preserve perceptual value, as in the case of TM operators, while as soon as the tone-mapped image loses its high quality, due to some other kind of manipulation, its commercial value is greatly diminished and therefore watermark loss could be afforded. This obviously implies that the watermark must be highly imperceptible. Depending on the application framework, the watermarking system could be either blind or non-blind, according to the availability of the original image during the watermark recovery; here we conduct a preliminary analysis by embedding a detectable (*i.e.* only the watermark presence must be assessed, but the watermark by itself bears no information), blindly recoverable, imperceptible watermark robust to non-linear value-metric attacks, using an improved version of the algorithm presented in [3]. The latter was originally proposed for 8 bpp grayscale images; however its structure and the particular nature of the problem at hand makes it suitable for its adaptation to the HDRI context too.

The remainder of the paper is organized as follows. In Section II, the system framework is proposed; in particular, the description of several overall design considerations constitutes the focus of this Section. In Section III a thorough, step-by-step description of the watermarking system is reported, highlighting how it fulfills the requirements in the HDR context. Then, Section IV shows some experimental results for a varied set of HDR images, both in terms of detection performances and imperceptibility. Finally, Section V draws some conclusive remarks.

II. System Framework

For watermarking purposes, it is reasonable to work in the luminance domain since tone-mapping is likely to strongly tamper with the chrominance components in an unpredictable way (*e.g.*, not monotonically). Therefore, the first step of our system must be luminance component extraction from the HDR image. It will carry the watermark, while chrominance components will be left untouched. This consideration is also compatible with the pre-existent LDR image watermarking system; for the sake of this Section, let us assume that we have at our disposal a reliable watermarking system, imperceptible and robust to both linear and non-linear pixel value-metric attacks, to use as a block box. The discussion on the watermarking



(a) First method: watermark does not survive TM.

(b) Second method: using a LogLuv reference image.

Figure 1: Sketch of HDR watermarking possible methods.

system is postponed to Section III. Figure 1 schematically shows how the watermarking system could be conceived. The most straightforward approach, depicted in Figure 1a, would be to watermark directly the HDR original image I_{HDR}, assuming the watermarking algorithm could be somehow ported in the HDR domain, and then hope that the watermark embedded in the resulting ${I_{HDR}}^W$ survives any subsequent TM process, denoted by TM^i. However, estimating the perceptual impact that the watermark is going to produce on the tone-mapped versions of the image is very difficult. Conservative approaches would surely imply that the watermark would not survive the TM process. Moreover, the watermarking algorithm may not be ported in the HDR domain without losing its nice robustness properties.

The solution adopted here is instead illustrated by Figure 1b. In [4] the LogLuv color space is shown; the log transform applied to the luminance guarantees that distortions of the same magnitude on different starting log-luminance values will be perceived approximately in the same way. In fact, in [5] it is argued that any generic tone-mapping method will replicate the Weber Fechner law which makes a luminance log transform an essential component of a TM process. Therefore, the transformation RGB-to-LogL, abbreviated in Figure 1b by L, leads to a reasonable domain in which to perform robust watermark embedding. Since the inverse log transform (or L^{-1}) is readily available, it allows to obtain back in a simple way the HDR watermarked image.

Hence, our proposed system works as follows. Initially, the original HDR image I_{HDR} is transformed to a reference image I_{LL} by applying RGB-to-LogL transformation as described in [4]. The log-luminance values range is set to $[0, 255]$ as for 8 bpp LDR images. Then, the algorithm embeds the watermark in I_{LL}, obtaining ${I_{LL}}^W$. Lastly, the reverse log transform is applied obtaining the watermarked version

of the given original HDR image. When a generic tone-mapping TM^i is applied to the watermarked HDR image ${I_{HDR}}^W$, the watermark is still present as long as the assumption of similarity between the log transform and tone-mapping operators in the luminance component holds, that is it is assumed that ${I_{TM^i}}^W$ can be obtained from ${I_{LL}}^W$ through a mild non-linear transformation against which the watermarking process is robust. The effectiveness of this approach will be demonstrated in Section IV.

III. Watermarking System

Digital watermarking systems comprise several possible frameworks. In particular, still image watermarking is a huge field, with a plethora of solutions proposed for a variety of different applications. The selection of the particular technique needed in a watermarking system is strongly dependent on its requirements, which in turn depend on the intended application. The main requirements of a watermarking system are robustness, security, capacity and imperceptibility. Robustness is the ability of the watermark to resist to non-malicious content manipulations, such as compression, editing, etc. Such manipulations are likely to occur during the normal lifetime of the watermarked object. On the other hand, security is concerned with intelligent attackers interested in removing or disabling the watermark by exploiting possible weaknesses in the watermarking system; in particular, an attacker could also take advantage of poor robustness against a particular transformation. Security usually requires that the system relies on some secret key shared between the embedder and the recovery entities as seen in private cryptography. Capacity refers to the quantity of information conveyed by the watermark itself; when this quantity is only 1 bit, that is to say we are only interested in assessing whether the watermark is present in a given image, the watermark is called detectable, otherwise it is called decodable or multi-bit. Last, imperceptibility refers to the obvious fact that the watermark embedding should not alter the perceptual quality of the original content. Another important classification in watermarking is between the blind and the non-blind watermark recovery, in which the latter process needs the original (unwatermarked) data to be performed whereas the former does not. All these requirements, plus a number of marginal others such as computational complexity of either the embedder or the recovery block or both, blend in different ways to form a watermarking system suitable for an application at hand.

The watermarking technique that is at the core of the proposed HDR image watermarking is explained in [6] and [3]. To summarize, the system proposed in this work is based on the QIM paradigm [7] and is designed to embed a blindly detectable one bit watermark in HDR images. The watermark must be specifically robust against TM operators (as well as against some other common image processing) that are thoroughly described in Section IV. Watermark security is guaranteed by a number of randomization

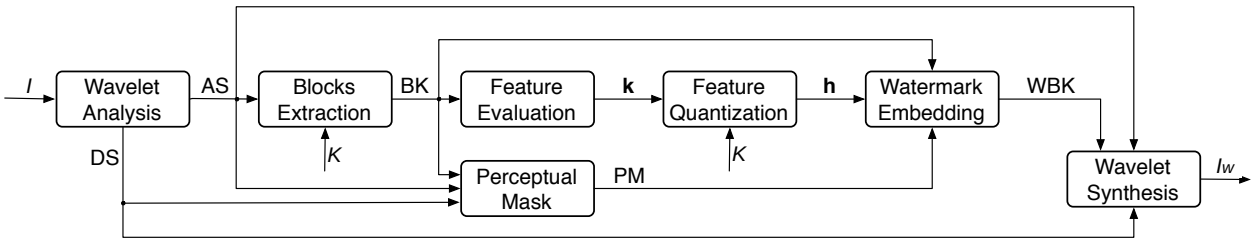


Figure 2: Watermark embedder flowchart.

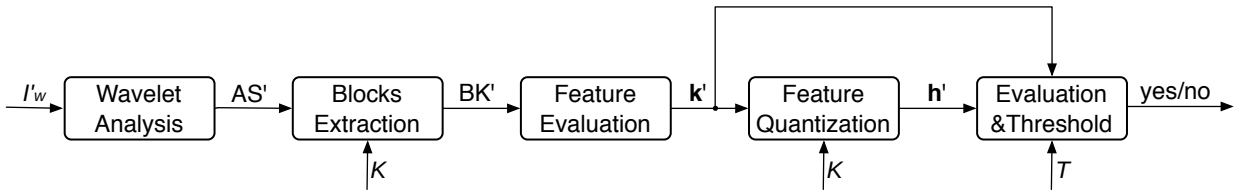


Figure 3: Watermark detector flowchart.

steps in the embedding process. Last, high-quality imperceptibility is an imperative requirement, pursued by employing a powerful perceptual mask during the watermark embedding process.

This Section first points out in Subsection III-A the requirements we have imposed on the HDR image watermarking system proposed in this work. Then, the system structure we have implemented to achieve them is described in the following Subsections. The feature selected as the watermark medium is the kurtosis of the approximation subband wavelet coefficients extracted in N blocks, and bit repetition coding is employed by embedding the same bit in every block. The feature extraction process is described in Subsection III-B. The perceptual mask necessary to guarantee that the imperceptibility requirement is met is illustrated in Subsection III-C. Finally, the watermark embedding and the recovery procedures are explained in Subsection III-D. The watermark embedder and detector respective flowcharts are shown in Figures 2 and 3; they are described in turn in what follows.

A. Watermarking Requirements

In our case, the main requirement which is imposed on the detectable watermarking system is imperceptibility, since there is strong interest in retaining the high perceptual quality of the original HDR image. The other important requirement is robustness against any TM operator which could be applied to the original HDR image. This means that after the watermark has been embedded in the original HDR image, all of its tone-mapped versions should also contain the watermark, regardless of the particular TM

algorithm used as long as the quality has been preserved to a sufficient extent. As already mentioned, there exist several tone-mapping algorithms, so that tailoring the approach to every one is impracticable. However, it can be easily assumed that all tone-mapped versions of a given HDR image are perceptually similar. In fact, in [5] it is argued that all tone-mapping operators could be described by a single model with an appropriate fitting of some model parameters.

All tone-mapped versions of a given HDR image can be derived one from the other by means of a global linear transformation (*e.g.*, a brightness adjustment) followed by some mild non-linear transformation. So, the original requirement of robustness against TM operators translates into robustness against this mix of linear and mild non-linear transformations. If this condition is verified we will show that when a particular TM image has been watermarked, the other TM versions will also contain the watermark. The watermarking system proposed in [6] is robust against pixel value-metric non-linear attacks. It is also invariant against linear value-metric scaling and thus it is well suited for the problem at hand. However, it needs some adjustments since it was originally proposed for LDR image watermarking. We choose to embed a detectable watermark, which means that its capacity is single bit. Security can be achieved by suitable randomization of the embedding algorithm steps (see K in Figures 2 and 3) thanks to the use of a secret key. Imperceptibility it is also a concern for the LDR domain; even if imperceptible changes on the TM images are likely to introduce significant distortion in the HDR domain because of the reverse log, this fact can be ignored by recognizing that TM operators will mimic the response of the HVS to real light stimulus, the same that is contained in a HDR image. Thus, when a HDR image is rendered on a normal display, the tone mapping reduces contrast in the same way the HVS would normalize the input stimulus. On a lower dynamic range display, the effect would be even stronger so that the watermark induced distortion in the HDR domain can be masked provided that imperceptibility in the LDR domain is achieved.

B. Feature Extraction

This Subsection explains the feature extraction process, which is a necessary step for both the watermark embedder and the detector. The input of the embedder is assumed to be the reference log-luminance image I. The first step of the feature extraction process is to apply a wavelet transform on the host image I. We use the orthogonal, 2-level Daubechies DWT, with 8 tap filter impulse responce. This DWT, aside from guaranteeing perfect reconstruction, has the additional advantage of providing a sparser representation, given by the high number of vanishing moments on the analysis filters, which in turn has benefits in the embedding process (see below).

Robustness and imperceptibility are by their nature conflicting requirements and a suitable tradeoff has to be reached in the watermarking system design. This tradeoff is evident even in the feature domain selection. In our case, the feature is extracted from the approximation subband, abbreviated with AS in Figure 2. It represent the actual watermark domain. The 6 detail subbands, at all levels, are left untouched and are collectively indicated by DS. Any small change, including watermark embedding, applied to the deep-level detail subbands (the "high-frequency" part of the wavelet decomposition, that in our case are the LLHL, LLLH, and LLHH subbands), may be imperceptible; however, these subbands have poor robustness properties since their content is easily destroyed by the TM transformations with virtually no visual implication. On the other hand, the AS of a 2-level wavelet decomposition is a low-pass filtered version of the host image, therefore there is more correlation between the AS coefficients than in the pixels of the original image so that robustness is inherently higher in this subband. Using more decomposition levels increases the coefficients correlation and thus their robustness; however, this has also detrimental effects on the block size and their number (which affects the the block selection process). We observed that using 2 levels of decomposition is a good tradeoff between AS subband robustness properties and block parameters selection. The AS accounts for the most perceptible part of the image data, hence achieving imperceptibility using this subband as the watermark domain is challenging; however, to obtain the desired level of robustness it is mandatory to embed the watermark in the AS domain. Imperceptibility must therefore be granted by adopting a complex perceptual mask (see Subsection III-C) and by carefully tuning the system parameters.

Now, a N-length feature vector $\mathbf{k}$ is to be obtained by evaluating the kurtosis of the AS coefficients contained in N blocks; let us defer the description of the blocks selection process and first discuss about the properties of the kurtosis feature. Let $\mathbf{x}_i, 1 \leq i \leq N$ be the coefficients vectors (the scan order is without particular significance) forming a block.

The sample kurtosis of an input coefficient vector of length L (indicating its j-th component as $x_i^{(j)}$), with sample mean $\overline{x_i}$ is defined as:

$$k(\mathbf{x}_i) = \frac{L \cdot \sum_{j=1}^{L} \left(x_i^{(j)} - \overline{x_i} \right)^4}{\left(\sum_{j=1}^{L} \left(x_i^{(j)} - \overline{x_i} \right)^2 \right)^2} \tag{1}$$

which is a biased estimator of the statistical kurtosis β_2 of a probability density function (pdf), defined as:

$$\beta_2 = \frac{\mu_4}{\sigma^4} = \frac{\mu_4}{{\mu_2}^2} \tag{2}$$

where, μ represent the pdf mean, $\mu_i = E[(X-\mu)^i]$ represents the i-th order central statistical moment and σ is the standard deviation. Therefore, the kurtosis β_2 is the fourth-order, central, normalized statistical moment. In informal terms, the kurtosis measures the "peakedness" of the distribution around its mean and how much the pdf tails are "fat". For instance, normal pdfs have $\beta_2 = 3$, uniform distributions have $\beta_2 = 1.8$, Laplacian distributions have $\beta_2 = 6$ and p-Bernoulli distributions have $\beta_2 = \frac{1}{p} + \frac{1}{1-p} - 3$ (for $p = 0.5$ we have the unique minimum for the kurtosis, $\beta_2 = 1$ and so $\beta_2 \geq 1$). The precision of the chosen k is adequate, since we are not trying to infer any underlying statistical property of the coefficients in $\mathbf{x}_i$ but we are instead interested in affirming that the feature is somewhat tied to the local perceptual content of the image, since it describes to some extent the shape of the local wavelet coefficients distribution in a compact way.

The sample kurtosis $k(\mathbf{x})$ of a vector $\mathbf{x}$ is invariant to affine transformations of the kind $p\mathbf{x}+q$ (where both product and sum are component-wise), that is $k(\mathbf{x}) = k(p\mathbf{x}+q), \forall\, p, q \in \Re$; the proof is immediate from inspection of Equation 1. So, the kurtosis feature is invariant to any CGA (constant gain attacks) applied to the image, since they are linearly propagated to the AS coefficients by the wavelet filters. In addition, experiments showed the robustness of the kurtosis feature; in particular, the *relative* variation experienced by this feature after the image from which it was extracted has been attacked is with high probability in the order of 10% of the original feature value.

Once we have defined the feature, we discuss the blocks extraction process. The block size is a key factor for the robustness properties of the kurtosis feature. Blocks that are too small fail to capture any perceptual content because of their excessive localization; moreover, and even more importantly, robustness is seriously harmed by evaluating the feature on small blocks given how easy it is to radically change the sample distribution with even minimum value changes. On the other hand, large blocks, being too global, lose their information uniqueness, that is they tend to contain the same information as the area of the image involved is greater. Furthermore, the complex watermark embedding procedure (Section III-D) is very sensitive to the number of coefficients involved; in particular, we will discuss further ahead on how the probability of not being able to correctly embed the watermark increases with the number of coefficients in the block. These are due to the possible non-convergence of the embedding algorithm and will be discussed in Subsection III-D. Note that the starting distribution of the AS coefficients is important in this respect too: a certain degree of sparsity in the initial pdf is advisable for the algorithm to converge. The complexity of the embedding algorithm, which in turn depends on the highly non-linear behavior of the kurtosis feature, also dictates that the blocks must be strictly non-overlapping. This fact puts an upper bound on the number of blocks N, but it has to be noted that the

perceptual mask would also have limited N.

The block shape has its importance too. If the block boundary is represented by straight lines as in square blocks, blocking artifacts are likely to appear (note that vertical and horizontal gradients are particularly perceptible for a human observer). Finally, some kind of randomization in the blocks selection is advisable for security concerns.

For all these reasons, the procedure for the N blocks selection on the AS subband is as follows; all random steps are driven by the secret key K. Assuming for simplicity of notation that the AS subband is square, first a uniform, $m \times m$ square cells grid is constructed. Each basic square cell side has dimension D, that drives the actual block size, so it has to be selected carefully. The cell boundaries are then randomized, using a low-pass filtered, zero-mean pseudo-random sequence to represent their deviation from the standard square cell border. This way, each cell has a generally square but with random boundary shape. The outer cells are discarded and then the whole grid of $M = (m-2) \times (m-2)$ remaining cells is randomly shifted both horizontally and vertically by at most $D/2$. Last, N out of M cells are randomly selected as the embedding blocks, provided that they satisfy a certain minimum area criterion, expressed by the fact that the number of coefficients in any given cell must be greater than ρD^2, where ρ is a design parameter fixed at 0.9 in this work. Although the blocks are not completely random, using this extraction procedure still achieves a high degree of security (given by the random shaping, shifting and selection) while guaranteeing both that wavelet coefficients spatial redundancy is adequately exploited and a sufficient number of non-overlapping blocks can be selected.

The number of blocks N is an important factor to determine both watermark imperceptibility and detection performances. It is obtained as follows:

$$N = \lfloor \beta \cdot M \rfloor \tag{3}$$

where β is a design parameter controlling the coverage by the blocks of the AS subband. Given the above area constraint on the blocks, there exists a maximum value β_{max} beyond which it is impossible to obtain the requested number of blocks satisfying said criterion. The value of β_{max} depends on both the host image and the secret key, and is usually in the range $[0.6, 0.7]$.

In this work, we have found that using $\rho = 0.85$, $D = 6$ and letting β vary in the range $[0.4, 0.65]$ for the extraction of the blocks on the AS subband of a 2-level decomposition provides a good tradeoff between all the requirements expressed above.

The output of the blocks selection process is therefore a structured set of blocks, indicated by BK, which does not constitute a matrix because the blocks have different size and is formed by the position

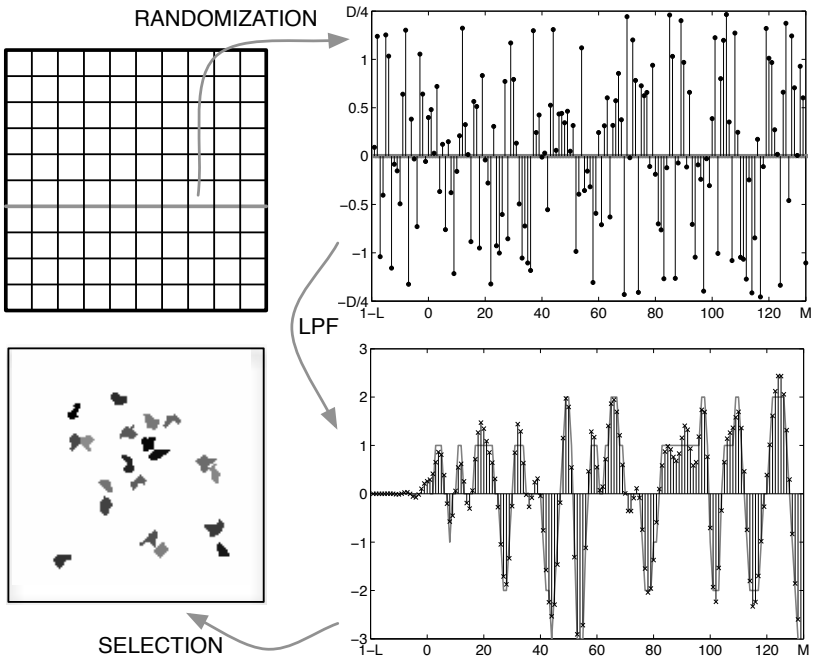


Figure 4: Blocks selection process.

in the AS subband of the coefficients pertaining to each one of the N blocks and a set of N vectors $\mathbf{x}_i, 1 \leq i \leq N$, each containing the wavelet coefficients in the i-th block. This information is also used to construct the perceptual mask, as explained in the next Subsection III-C. The feature is evaluated on the blocks by applying Equation 1, thus returning the original, unwatermarked feature vector $\mathbf{k}$ which constitutes, together with the perceptual mask, the input to the QIM embedding algorithm.

C. Perceptual mask

In this Subsection, the construction of the perceptual mask (PM) is discussed. The perceptual mask allows to control the watermark perceptibility. Since it is directly evaluated in the same wavelet domain which the feature belongs to, the PM can be applied as a simple constraint in the watermark embedding process, limiting the coefficients change during the watermark embedding.

The perceptual mask is actually calculated for every coefficient of the block coefficients vectors $\mathbf{x}_i$ hence PM is evaluated on the same structured set of blocks BK: a set of N masking vectors $\mathbf{m}_i$, $1 \leq i \leq N$, that contain the perceptual mask values for each coefficient location in BK. However, for the sake of simplicity, in the following we will evaluate PM on the entire AS subband.

The basis of our perceptual mask is directly derived from [8] and references therein, in which an additive spread-spectrum detectable watermarking algorithm is applied to the first-level detail subbands

of a wavelet decomposition. The algorithm is also applicable to the other subbands as well. There, the masking effect is achieved by multiplying the pseudo-random binary sequence constituting the watermark with a perceptual mask, which we call V. This was derived from previous work on the calculation of optimal quantization step sizes for wavelet domain image compression. The perceptual mask V is evaluated using all decomposition levels up to 4 and all of the subbands and represents the maximum distortion that is perceptually tolerable on a coefficient-by-coefficient basis; that means that coefficients corresponding to higher values of the perceptual mask are allowed to vary more to accommodate the watermark without violating the imperceptibility requirement.

The originally proposed perceptual mask V is the product of 3 terms, as in:

$$V_d^\theta(r,c) = \Theta(d,\theta) \cdot \Lambda(d,r,c) \cdot \Xi(d,r,c)^{0.2} \tag{4}$$

where (r,c) represents the row and column position in a given subband. The first term is a scalar value $\Theta(d,\theta)$ which depends on the particular considered subband, corresponding to decomposition level d and orientation θ. The second term $\Lambda(d,r,c)$ is derived from the luminance (that is the coefficient value itself), following the principle that very dark and very bright regions of the image are less perceptually sensitive to distortions than medium luminance ones. The last term $\Xi(d,r,c)$ is an activity measure, since the human eye is more tolerant to noise in textured areas of the image. It is calculated as the product of the local mean square value of the detail subbands and the local variance in the approximation subband, both evaluated in a local neighborhood of the considered coefficient, where the local neighborhood is defined as unitary for the highest decomposition level and doubling in size as the level decreases.

Our perceptual mask is slightly different and is expressed as:

$$PM(r,c) = \alpha \cdot \Lambda(r,c) \cdot \Xi(r,c)^{0.1} \cdot \Gamma(r,c)^{-2} \tag{5}$$

where (r,c) now represents the row and column position in the AS subband, which is the watermark domain, so the d and θ indexes could be dropped. Since we always embed in the AS subband, the term Θ is discarded; instead, we introduce a weight α as an overall strength factor. The log-luminance term $\Lambda(r,c)$ scans the AS subband. Given the variable, real-valued luminance range of the log-luminance image I, as opposed to the standard luminance range $[0,255]$ found in LDR images, to properly identify dark and bright areas of the image the following expression for $\Lambda(r,c)$ is adopted, based on log-luminance distribution percentiles:

$$\Lambda(r,c) = \begin{cases} 1 & \text{if } l \le l_{10} \ \vee \ l > l_{90} \\ 1 - \frac{1}{2} \cdot \frac{l - l_{10}}{l_{50} - l_{10}} & \text{if } l > l_{10} \ \wedge \ l \le l_{50} \\ 1 - \frac{1}{2} \cdot \frac{l_{90} - l}{l_{90} - l_{50}} & \text{if } l > l_{50} \ \wedge \ l \le l_{90} \end{cases} \tag{6}$$

where l is the coefficient value in the position individuated by (r, c) in the AS subband (which is a low-pass filtered version of the actual log-luminance value) and l_{10}, l_{50} and l_{90} are the 10-percentile, 50-percentile and 90-percentile values respectively of the AS coefficients distribution. The term $\Lambda(r, c)$, therefore, assumes the value 1 for the extreme parts of the distribution and then decreases linearly towards its median value where it assumes the value 0.5, in accordance with the luminance principle expressed above.

The expression for the activity term Ξ, using Matlab notation for the sake of simplicity, is as follows:

$$\begin{aligned} \Xi(r,c) &= Var\left\{AS(r-1:r+1, c-1:c+1)\right\} \cdot \\ &\quad Msv\left\{DS_2(r-1:r+1, c-1:c+1);\right. \\ &\quad \left. DS_1(2r-9:2r+2, 2c-9:2c+2)\right\} \end{aligned} \tag{7}$$

where $Var(\cdot)$ and $Msv(\cdot)$ stand for variance and mean square value respectively. The local neighborhood is enlarged to a 3×3 square centered on (r, c) in the AS; the position of the local area for other subbands is derived from the decimation and filtering steps in the wavelet decomposition. The DS are considered to compute the PM values since their coefficient values affects perceptibility of the reconstructed data.

The exponent for the activity term in Equation 5 is different from that in Equation 4 since in the former an additional term $\Gamma(r, c)$ has been introduced to take into account contour information. Among the areas with high activity, individuated by the activity term, the ones characterized by a sharp contour dividing two relatively flat areas are the less perceptually tolerant to noise. Therefore a dedicated term is necessary to reduce the perceptual mask in their vicinity. To construct the term $\Gamma(r, c)$, first we run a Canny edge detector on the AS to extract the most significant edges. Then, starting from a binary map (1 represents the edge points), we apply a moving average filter in a 7 coefficients sized neighborhood to spread the edges. Then, the result $\gamma(r, c)$ is transformed into the final value by: $\Gamma(r, c) = 10^{\gamma(r,c)}$. This way, the coefficients without edges retain the multiplicative value 1 in Equation 5, while this value is decreased exponentially as they get nearer the contours (recall the negative exponent of the term $\Gamma(r, c)$).

For illustrative purposes, Figure 5 provides an example of a perceptual mask and its constitutive parts; all figures are rendered in a normalized gray scale version of the corresponding function, so that the gray values are only representative of the relative strength and not of the absolute values of the various terms. The original AS coefficients of the "Tree" HDR image and the closely associated Λ component, that favors medium valued coefficients, are depicted in Figure 5a and 5b respectively; note how boundaries, which present artifacts caused by the border symmetry condition in the wavelet transform, are neglected in the perceptual mask computation. The activity term Ξ is reported in Figure 5c; note how the product

of the log-luminance term and Ξ attains high values in the proximity of edges, as illustrated in Figure 5d (edges visibility is reduced in Figure 5a due to the low-pass filtering). With the introduction of the additional edge term Γ of Figure 5e (inverted to improve readability), the overall perceptual mask, which is depicted in Figure 5f, is much less permissive around the edges and tends to reach its maximum values in the textured, medium luminance areas of the image. Subsection IV-C shows how the various components coordinate to achieve imperceptibility while retaining robustness.

D. Watermark embedding and detection

This Subsection briefly describes in turn the embedding process and the detector used in the watermarking process of the reference log-luminance image. As already mentioned, the watermarking paradigm employed in this work is the well-known QIM method, first presented in [7], in its detectable form. At the detector side, we are only interested in assessing whether the watermark individuated by the secret key K is present or not in the considered HDR image. Since our feature is 1-dimensional, that is to say scalar, there is no need for vector quantization techniques thus leading to an efficient computational implementation of the algorithm.

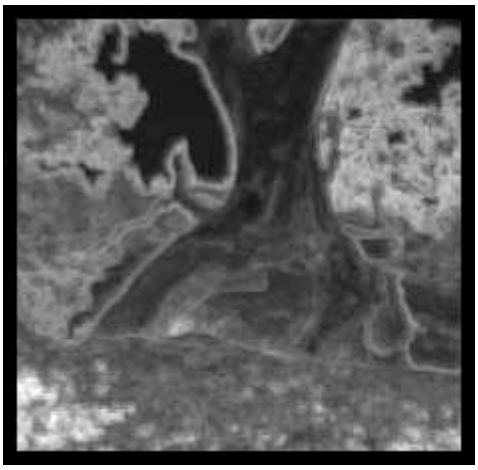
The watermark embedding consists of the quantization of the feature value using one of two possible scalar quantizers, one shifted with respect to the other by $\Delta/2$ (where Δ is the quantization step in the case of uniform quantizers). Each quantizer is identified by an embedding bit value, so if multiple feature values are available in a feature vector it is possible to embed a watermark codeword by selecting the appropriate quantizer for each feature vector component; the bit associated with every component (or block as in our case) is embedded and decoded separately. In the case of detectable watermarking scenario, the watermark payload is only 1 bit (the watermark presence itself), hence bit repetition is the most effective coding for its transmission. We can assume without loss of generality that the quantizer associated to value 1 is always used. Moreover, to enable security, QIM techniques generally rely on shifting the codebook by a random quantity. In our case it is a scalar quantity d, whose application does not affect either imperceptibility or robustness given that it is derived exclusively from the secret key K and it is independent from the host feature value. Given a kurtosis feature value, the corresponding watermarked feature value is expressed by:

$$h = Q_{\Delta}(k - d\Delta) + d\Delta \tag{8}$$

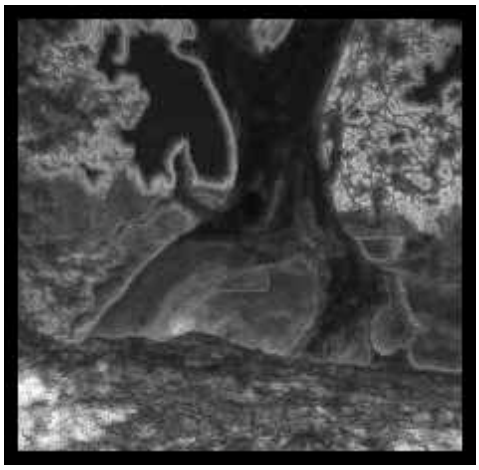
where k is the original feature value, h is the quantized value, Δ is the quantization step, $Q_{\Delta}(\cdot)$ is the quantization operator, with quantization step Δ as parameter, and d is the codebook shift. The shift d is

(a) Original AS of the log luminance image I.

(b) Luminance perception term: Λ.

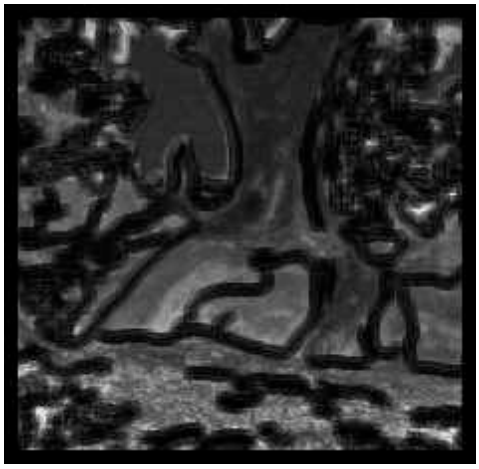


(c) Activity perception term: Ξ.



(d) The product $\Lambda \cdot \Xi$.

(e) Edges perception term: Γ.



(f) The final perceptual mask: $\text{PM} = \Lambda \cdot \Xi \cdot \Gamma$.

Figure 5: The perceptual mask of the "Tree" figure and its product terms. Brighter values in (f) represent less perceptible areas to changes in the kurtosis feature.

extracted from a uniform distribution in $[-1/2, 1/2]$ and then multiplied by Δ. Previous counter-shifting of the original feature value is necessary to ensure that h is the nearest reconstruction point. Therefore, the secret key K associates with each block a codebook shift uniformly distributed in the quantization interval.

To decode the bit embedded in a block using the QIM paradigm, the received feature value k' (which is potentially different from h, *e.g.*, due to attacks on the watermarked image) is re-quantized based on a codebook that is the union of all possible embedding codebooks. The bit sequence associated with the codebook to which the resulting reconstruction value belongs is the decoded string. In our case a single bit is embedded in every block, by always employing only one of the two scalar quantizers shifted of half quantization step one with respect to the other, so that it is simply necessary to evaluate whether the reconstruction value h' is associated or not to the quantizer corresponding to the bit value 1. From another perspective, typical of the QIM detectable watermarking embedding process [7], the detection of the bit embedded in every block consists in assessing whether the received feature k' is within $\Delta/4$ to the closest reconstruction value of the sole quantizer associated with the bit value 1 (otherwise the bit value 0 would be detected). Therefore, using again the notation presented above and in Equation 8, the decoded bit b' out of a single received block is recapped in the following pair of expressions:

$$\begin{aligned} h' &= Q_\Delta(k' - d\Delta) + d\Delta \\ b' &= \begin{cases} 1 & \text{if} \quad |k' - h'| \leq \Delta/4 \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{9}$$

In this work we do not use uniform quantization; instead, as we mentioned earlier, due to the fact that the feature shows its robustness properties by bounding its relative (and non absolute) variation following attacks on the image a non-uniform quantization with higher quantization steps as the feature value increases is more appropriate in this context. The quantization scheme is illustrated in Figure 6, where reconstruction values are identified by crosses and the circles stand for the reconstruction values of the dual quantizer associated with the bit value 0, which is never used. For kurtosis values $k < 15$, three quantization intervals I_q, $q \in \{1, 2, 3\}$ are employed. Each interval hosts a single reconstruction value, that is every interval is long exactly one corresponding quantization step; they are respectively $I_1 = [1, 2]$ with $\Delta_1 = 1$, $I_2 = [2, 5]$ with $\Delta_2 = 3$ and $I_3 = [5, 15]$ with $\Delta_3 = 10$. For those rare values beyond I_3, namely $k > 15$, the quantization step Δ_3 is used too. The case $d = 0$ in which the codebook is not shifted is illustrated in Figure 6. If $d \neq 0$, the reconstruction values are no more in the center of its corresponding quantization interval, as depicted in Figure 7 (observe how the quantization

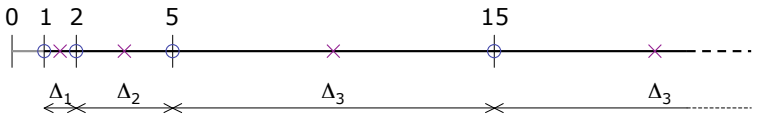


Figure 6: Quantization with $d = 0$.

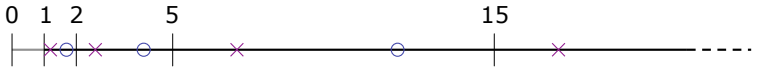


Figure 7: Quantization with $d = -0.3$.

intervals remain fixed). This in turn implies that a segment of length $|d| \cdot \Delta_q/2$, corresponding to the part associated to the dual reconstruction value (the circles), is included in any given quantization interval; if the feature value belongs indeed to these segments, its corresponding reconstruction value would lie in the adjacent quantization interval. Therefore, the quantization step associated to the latter interval is in these cases adopted to minimize unavoidable decoding errors near the quantization interval boundaries; in particular, with respect to Eq. 9, the correct decoding region when $d < 0$ is less than $\Delta_q/4$ for $k' < h$ and when $d > 0$ is more than $\Delta_q/4$ for $k' > h$. Such errors can be expected to be compensating each other and concern only the less robust blocks anyway. The actual watermark embedding is performed by applying an optimization algorithm to $\mathbf{x}_i,\ 1 \leq i \leq N$, the AS coefficients in each of the N blocks contained in BK, in a way such that the corresponding watermarked coefficients vectors $\mathbf{y}_i$ in WBK have the quantized feature values vector $\mathbf{h}$. The perceptual mask described in Subsection III-C is used as a $\mathcal{L}_\infty$-norm constraint in this process, in the sense that $PM(r, c)$ acts as the upper bound of the absolute admissible distortion of the AS coefficient in the position (r, c) for the watermark to be imperceptible. The embedding procedure is described as follows: dropping as usual (r, c) and the block index, given a coefficient vector $\mathbf{x}$ of length L and the associated PM vector $\mathbf{m}$, we first define the coefficients subspace Ω in $\Re^L$ (more precisely, in $[1, \infty]^L$) representing the region around the point $\mathbf{x}$ within the perceptual mask:

$$\Omega = \left\{ \mathbf{z} \in [1, \infty]^L : |\mathbf{z} - \mathbf{x}| \leq \mathbf{m} \wedge \|\mathbf{z} - \mathbf{x}\|_1 \leq \tau \cdot \|\mathbf{m}\|_1 \right\} \tag{10}$$

The $\mathcal{L}_1$ constraint is necessary to support the coefficient-by-coefficient based construction of the perceptual mask, that is to say the PM is constructed neglecting the case where more than one adjacent coefficients is changed at the same time. For this reason, the second inequality is introduced to avoid that too many coefficients simultaneously experience their maximum variation in any given block; it is driven by the

parameter τ which expresses the ratio of the sum of actual coefficient modifications with respect to the maximum allowed, namely the $\mathcal{L}_1$-norm of $\mathbf{m}$. The fine tuning of τ is an important design choice, given that its value greatly influences not only the imperceptibility requirement satisfaction but also the size of the subspace Ω and therefore the convergence of the embedding procedure, as described below. Experimentally, a good value for τ is 0.3.

Hence, the mathematical procedure embodying the embedding process has to move the original feature value to its quantized value, or at least as close as possible, while not moving $\mathbf{y}$ out of Ω:

$$\mathbf{y} = \arg\min_{\mathbf{z}\in\Omega} |k(\mathbf{z}) - h| \tag{11}$$

where the function $k(\cdot)$ is the kurtosis evaluation as in Equation 1. If this procedure converges to 0, then $k(\mathbf{y}) = h$ as required. Notably, there could be many such solutions in the subspace Ω. This does not matter as long as the individuated $\mathbf{y}$ satisfies both the PM constraints and the above condition on its kurtosis.

There is a distinct probability that the procedure expressed by Equation 11 does not converge to 0. Less likely is the non-convergence of the minimization process; in particular, this especially happens when trying to change low kurtosis starting values k to distant h target values, which is a further justification of the non-uniform, increasing step quantization adopted in this work. A much more likely possibility is that the perceptual mask prevented the reaching of a suitable solution, namely the feature value can not be quantized without introducing more distortion than allowed in the coefficient vector, which implies that the perceptual constraints must not be stricter than necessary. In fact, when Equation 11 does not converge, the kurtosis value of the watermarked vector is not h and this fact decreases the robustness of the considered block against attacks since its feature value is closer than intended to the correct decoding threshold. In extreme situations, it could be impossible to move the feature value out of the wrong quantization interval, in this way triggering a decoding error even in the absence of attacks. Obviously, it is imperative that not too many blocks experience this non-convergence situation because they hurt the miss probability P_M. The analysis on P_M is carried out in Section IV.

Once the watermarked feature vectors $\mathbf{y}_i,\ 1 \leq i \leq N$, are obtained, the last step that the embedder performs is the inverse wavelet transform of the log-luminance watermarked image I_w using both the AS subband modified according to WBK and the detail subbands DS.

The watermark recovery block, that is the watermark detector, is depicted in Figure 3. Given a potentially watermarked log-luminance image I'_w, it follows in the steps of the embedder, using the secret key K, by extracting the received block information BK', composed of attacked coefficients vectors

$\mathbf{y}'_i, 1 \leq i \leq N$, and then by evaluating their kurtosis as in Equation 1, until a received feature vector $\mathbf{k}'$ is obtained. Then, each block is decoded separately according to Equation 9. Last, a detection decision is taken using a threshold T on the number of correctly decoded blocks (that is to say with decoded bit value $b' = 1$) out of the total number of embedding blocks N.

This detection structure allows to control in a simple way the probability of false alarm P_{FA} as a function T. In fact, if the image is unwatermarked, the received feature value is completely uncorrelated with the reconstruction value associated to the bit 1, hence its probability of correct detection is 0.5 given the uniform distribution of the shift d. Furthermore, since the blocks are non-overlapping, we can safely assume that the blocks detection probabilities are independent one from the others. Given these premises, the occurrence of a false alarm consists in the contemporaneous verification of at least T events with probability $p = 0.5$ out of N. If we call E the latter event for a generic p, we have the following formula based on the binomial distribution:

$$P(E) = \sum_{l=T}^{N} \binom{N}{l} p^l (1-p)^{N-l} \tag{12}$$

and then substituting $p = 0.5$ we obtain the false alarm probability as:

$$P_{FA} = \sum_{l=T}^{N} \binom{N}{l} \left(\frac{1}{2}\right)^{l+N-l} = \left(\frac{1}{2}\right)^{N} \sum_{l=T}^{N} \binom{N}{l} \tag{13}$$

which is easily computable and is used in Section IV to construct the ROC curves.

IV. Experimental Results

This Section shows the effectiveness of the watermarking system while showing an example workflow for the parameters setting and performance evaluation. First, to properly assess how the system performs in terms of imperceptibility and detection, it is important to note that there is a number of variables which affect the outcome, both external and internal (that is to say design parameters). The former ones include the host HDR image, signifying both its content (and thus features) and size; the secret key K, which drives all the pseudo-random steps in the algorithm and the tone-mapping operators (or any other attack) considered for robustness requirements. Aside from those suitably fixed throughout Section III, the design parameters that the system leaves free to vary at this point are: the PM overall strength factor α (see Equation 5) and the blocks coverage percentage β (see Equation 3).

These parameters should be chosen to guarantee optimal performance depending on the external variables; however, determining a good-for-all set of design parameter is not simple in this case. In this work, we have operated as follows. First, in Subsection IV-B a single image is considered. Here,

some experimental tests are run to evaluate both detection performance and imperceptibility, using a small set of secret keys and of TM operators while varying the α and β parameters, to infer some guidelines on how to choose the latter for any image by studying their behavior for the considered image. Then, having fixed said design parameters (which are not the optimal ones for every image, every secret key and every TM operator but represent a reasonable choice), in Subsection IV-C we conduct a brief evaluation on the functionality of the perceptual mask and in Subsection IV-D we run the robustness tests on the whole set of images and finally report the results for each TM operator. Subsection IV-E assesses the robustness of the system to mask-driven noise addition. Finally, Subsection IV-F compares the system to an hypothetical spread-spectrum system applied adopting the same experimental conditions and TM processes. The experimental Subsections are preceded by Subsection IV-A, in which the HDR image database and the set of tone-mapping operators considered has been described.

A. Experimental Conditions

The HDR image database used in our test is an heterogeneous set of 32 bit RGBe encoded images collected from two sources: the Munsell Color Science Laboratory's [9] and Greg Ward's [10] website repositories. The selection was performed so as to provide some variety in terms of subjects being represented as well as image sizes and dynamic ranges. Table Ia reports the images list, complete with the relevant data.

The tone-mapping operators considered in this work are those made available by the `pfstmo` library, part of the `pfstools` package [18]; they are listed in Table Ib. Additionally, a tone-mapping algorithm recently proposed by some of the authors of this work is also considered. Most of the algorithms need some calibration efforts such as image-dependent parameter tuning and/or gamma correction to improve the output tone-mapped image. Each image has undergone this process after it has been watermarked (recall that the watermark is imperceptible); without discussing the details at length, the images have been calibrated manually with the aim to provide the best possible result from a perceptual point of view with reasonable effort, *i.e.*, without exhaustive manipulation.

B. Design Parameters Tests

The well-known "Tree" image is selected as the benchmark image for these tests due to its varied visual content, making it a challenging testbed.

First, imperceptibility is evaluated through the HDR-Visual Difference Predictor (HDR-VDP) tool [19], which is a full-reference visual difference metric (meaning there is a so-called mask image to compare

Image short name	*Source*	*Size*	*Notes*
Apartment	[10]	2048×1536	Natural, indoor
AtriumNight	[10]	760×1016	Natural, indoor
Desk	[10]	644×874	Natural, indoor
Montreal	[10]	2048×1536	Natural, outdoor
Chair	[9]	2000×1312	Natural, indoor
Atrium2	[9]	2000×1312	Natural, indoor
Fog	[10]	751×1130	Natural, outdoor
Colorcube	[9]	2000×1312	Natural, indoor
Dandelion	[9]	1312×2000	Synthetic
DaniBelgium	[10]	1024×768	Natural, indoor
Hallway	[9]	2000×1312	Natural, indoor
Memorial	[10]	512×768	Natural, indoor
Rend01	[10]	1024×1024	Synthetic
Splitcube	[9]	2000×1312	Natural, outdoor
Tree	[10]	928×906	Natural, outdoor

(a) List of the HDR images composing the test database.

Authors and year	*Abbreviation*	*Reference*
Mantiuk et al., 2008	M08	[11]
Reinhard et al., 2005	R05	[12]
Drago et al., 2003	D03	[13]
Fattal et al., 2002	F02	[14]
Pattanaik et al., 2000	P00	[15]
Durand et al., 2002	D02	[16]
Boschetti et al., 2010	B10	[17]

(b) List of the tone-mapping operators considered throughout this work.

Table I: Database details.

a target image with). Given two similar images, the output of the HDR-VDP is the percentage of pixels that, according to its model, a human observer would perceive as different with a given probability (respectively 75% and 95%). In this context, it is used to evaluate the watermarked HDR image quality with respect to the HDR original image in a way much more precise and tied to the HVS than purely objective metrics like the PSNR.

In Table II it is reported a run with a random secret key. The factors α and β are allowed to vary respectively in the sets $\{4, 6, 8, 10, 12\}$ and $\{0.4, 0.45, 0.5, 0.55, 0.6, 0.65\}$. The number of extracted blocks N is also reported for reference. The output of the HDR-VDP is expressed as the pixel percentages at the 75% and 95% work points for all the combinations of the two design parameters. Obviously, the perceptibility of the watermark increases with higher α (which means looser PM). However, this is not always true column-wise, that is increasing the coverage percentage β. This counter-intuitive result is due to the fact that changing β has been implemented as to lead to a complete re-extraction of the blocks position, so that their "embeddability" properties vary from one coverage percentage to the next. This was a design choice to increase security, since this way it is impossible to guess the position of some blocks (and so possibly the secret key) by observing two watermarked images using different β but the

β	N	HDR-VDP %	α 4	6	8	10	12
0.5	561	75	0.103	0.246	0.407	0.541	0.667
		95	0.044	0.115	0.202	0.273	0.343
0.55	617	75	0.110	0.272	0.387	0.515	0.673
		95	0.051	0.139	0.195	0.273	0.342
0.6	673	75	0.130	0.281	0.444	0.609	0.803
		95	0.057	0.141	0.224	0.313	0.444
0.65	729	75	0.135	0.263	0.460	0.652	0.872
		95	0.063	0.130	0.218	0.338	0.460

Table II: Imperceptibility test: relative frequency of perceived changes in terms of number of pixels, expressed in percentage.

same key. Although in principle the same process could be applied with α, for which the blocks are not re-extracted, this latter factor is not meant to be changed after its selection for a specific image, while β is intended as the primary imperceptibility/robustness tradeoff parameter. If it is planned to securely watermark the same image with both parameters variable (as in this demonstrative test), then the blocks should be re-extracted for every (α, β) pair. However, though perhaps not monotonically for a given secret key as in Table II, imperceptibility generally worsen with increasing β as it should be.

Next, decoding performance have to be taken into account to properly select the embedding parameters. The watermarked images obtained through the previous run used to construct Table II are decoded both as they are and after tone-mapping using the methods listed in Table Ib. The number of missed blocks (not correctly decoded as in Equation 13) with respect to the total number of blocks N are shown in Table III, which has the same setting as Table II. The third column represents T^*, the complement of the threshold T ($T^* = N - T$), which is the maximum number of allowed missed blocks for an overall watermark miss not to occur. The threshold T is calculated using Equation 13 so that $P_{FA} < 10^{-6}$. The rows labeled "None" represent the decoding results of watermarked images without any specific TM (clearly the decoding is still operating after the log-luminance transform) and so account for some of the non-convergent blocks (those not correctly decoded) as explained in Subsection III-D. As foreseeable, the amount of these blocks decreases as the PM is looser (that is with higher α); column-wise, the percentage tends to remain invariant, with some variability induced by the secret key. Interestingly, the latter variability is reflected in Table II (the difference being that in the latter there is an underlying linear

β	N	T^*	TM	$\alpha = 4$	$\alpha = 6$	$\alpha = 8$	$\alpha = 10$	$\alpha = 12$
0.5	561	224	None	129	102	88	84	79
			M08	158	154	149	141	131
			D03	195	201	205	198	194
			R05	**226**	**227**	**232**	**237**	**233**
			F02	**241**	**240**	**241**	**238**	**242**
			P00	**225**	**232**	**233**	**227**	**230**
			D02	**242**	**235**	**234**	**227**	**227**
			B10	188	181	177	183	174
0.55	617	250	None	141	113	100	90	82
			M08	189	170	163	157	156
			D03	201	193	194	189	188
			R05	245	240	244	**251**	**256**
			F02	**275**	**267**	**266**	**263**	**262**
			P00	234	228	226	231	230
			D02	233	231	233	230	223
			B10	219	208	206	197	205
0.6	673	274	None	165	134	119	111	101
			M08	201	184	175	174	168
			D03	233	227	229	229	222
			R05	**275**	267	266	263	262
			F02	**300**	**290**	**294**	**299**	**292**
			P00	272	268	**277**	**277**	**281**
			D02	**280**	**278**	271	272	**277**
			B10	228	221	223	227	217
0.65	729	300	None	149	113	99	96	92
			M08	218	190	182	175	164
			D03	233	225	220	222	223
			R05	296	286	280	281	281
			F02	**307**	296	289	288	287
			P00	275	268	268	267	265
			D02	287	272	269	271	266
			B10	245	243	236	238	232

Table III: Decoding test: number of missed blocks.

increase), suggesting that non-convergent blocks are also those that impact on perceptibility the most. However, given that the secret key could not be chosen freely and that in blind watermarking systems there is no side channel, it is impossible to avoid using these blocks in this framework. It has to be noted that, as the number of blocks N increases along with the coverage factor β, the threshold accounts for a higher percentage of incorrectly decoded blocks, as it is more and more difficult to get the number of 0.5-probability events biased away from half of the overall occurrences as the number of events increases.

Therefore, there is a distinct advantage in using more blocks from a decoding performance point of view, although the monotonicity of this advantage is dependent on the secret key.

The other rows, labeled with the corresponding abbreviation, depict the decoding results after the application of the considered TMs. As expected, more blocks are wrongly decoded; this results in some overall misses (highlighted in bold), that in turn reflect the loss of the watermark. The (α, β) combination for which a miss occur for any TM should be discarded, at least for this image, unless some particular reason prevents it (*e.g.*, unsatisfactory tone-mapped image visual quality due to algorithm inappropriateness). It is notable that, as α increases, the number of incorrectly decoded blocks lowers much less than its non-TM counterpart, suggesting that those blocks which converge to the correct decoding interval only by loosing the PM are inherently less robust than the others. This reasoning implies that, at similar HDR-VDP scores, it is more convenient from a robustness point of view to have more and perceptually tighter blocks.

It is also possible to observe a sharp increment in decoding errors in the column $\alpha = 4$. This indicates the lower limit of the strength of the PM, as it hampers the convergence of too many blocks, including those with good robustness and perceptibility properties (with $\alpha = 0$ there would be no watermarking process).

In conclusion, the best watermarking strategy would be to choose the perceptual mask strength factor α to let imperceptibility and robustness lie in a suitable interval and then tuning the coverage factor β according to the robustness-imperceptibility trade-off as required. Decoding performance, however, are bound to improve as the size of the image increases, both because of the higher number of blocks used for watermarking purposes and of the higher quality of the watermarked images.

By inspection of Table III, a reasonable parameter choice would be $\alpha = 8$ and $\beta = 0.65$, which does not show misses and is still very good in terms of perception. Following our previous discussion, this choice maximizes the number of blocks while not widening excessively the perceptual mask. This combination is therefore adopted for the remainder of the tests.

C. Perceptual mask components evaluation

Table IV proves the importance of each of the components of the mask for achieving high imperceptibility, while at the same time ensuring sufficient robustness. In this test, the “Tree” image has been watermarked with the same secret key as before, but using only some of the mask components in turn and normalizing the mask energy with the α parameter so that all the alternative masks have the same energy as the proposed one (note that it does not imply that the watermark will have the same energy). The HDR-VDP

Components	HDR-VDP %		TM		DWR_{dB}
	75	95	None	M08	
Λ	0.123	0.047	171	208	34.8
Ξ	1.730	1.045	20	99	22.4
$\Lambda \cdot \Xi$	1.373	0.820	24	105	22.8
$\Lambda \cdot \Xi \cdot \Gamma$	0.460	0.218	99	182	32.2

Table IV: Perceptual mask components performance.

is reported together with the Data-to-Watermark Ratio (DWR), expressed in dB, in the logL domain. To assess robustness, the number of decoding errors is also reported, for the cases of no additional manipulation and the M08 tone-mapping algorithm.

As it can be seen, using only the luminance component Λ achieves high imperceptibility, but this is because the total watermark energy, expressed by the DWR, is quite low; this also causes an appreciable loss in robustness. On the other hand, using only the activity component Ξ shows how the watermark is robust when it is preferably embedded in areas of the image with high variance, but also how perceptible it is, since high activity, either dark or bright areas of the image are heavily modified. Combining the edge component Γ with the activity component helps to improve the perceptibility issue without impacting too much on the system robustness performance, showing that dealing specifically with sharp contours is advisable. Joining the three components as proposed is a satisfying tradeoff between the high degree of imperceptibility necessary for high quality HDR images and the requested robustness.

D. Robustness against tone-mapping

Tables V and VI report the decoding results with this parameter combination, for large and small images respectively; the results for the "Tree" image are repeated here for comparison. As it can be seen, the total number of decoding errors is always lower than T^*, so that none of the images results in a miss (except in a pair of cases, where proper image post-tone-mapping calibration was not satisfactory). Therefore, the watermark proves to be widely robust to tone-mapping operators as expected. With such a small set of images it is impossible to draw ROC curves based on actual misses and false alarms. To approximate them, we have evaluated the sample mean of the block decoding error probability p_e for each of the tone-mapping algorithms. Then, the miss probability P_M is estimated as follows:

$$P_M = 1 - \sum_{l=T}^{N} \binom{N}{l} (1 - p_e)^l \, p_e^{N-l} \tag{14}$$

		Apartment	Montreal	Chair	Atrium2	Colorcube	Dandelion	Hallway	Splitcube
N		3106	3106	2567	2567	2567	2567	2567	2567
T		1686	1686	1404	1404	1404	1404	1404	1404
T^*		1420	1420	1163	1163	1163	1163	1163	1163
TMO	None	518	882	611	508	544	811	567	474
	M08	652	957	731	587	599	824	597	541
	D03	734	936	779	579	599	811	588	528
	R05	951	1000	696	588	653	840	600	551
	F02	1278	1106	910	843	994	856	920	811
	P00	803	1009	711	703	717	820	601	603
	D02	1131	931	779	667	818	869	594	630
	B10	1017	950	1030	666	895	**1219**	761	677

Table V: Decoding test for the largest images: number of missed blocks.

		AtriumNight	Desk	Fog	DaniBelgium	Memorial	Rend01	Tree
N		666	457	737	666	298	938	729
T		394	279	433	394	190	542	429
T^*		272	178	304	272	108	396	300
TMO	None	102	119	101	175	64	231	129
	M08	161	151	144	195	96	259	182
	D03	130	149	127	201	81	245	220
	R05	142	**189**	118	223	98	257	280
	F02	256	176	279	263	97	341	289
	P00	182	168	173	212	103	269	268
	D02	204	163	203	224	98	332	269
	B10	154	171	195	239	98	295	243

Table VI: Decoding test for the smallest images: the number of missed blocks.

The ROC for 4 values of the number of blocks N are reported in Figure 8. Although the results are already satisfactory for smaller images, they are excellent for larger images. It has to be noted that the number of block decoding errors in a given image is strongly dependent on how good is the tone-mapping algorithm for the specific image (as the misses in Tables V and VI demonstrate). There are other factors of variability, as already stated at the beginning of this Section, embodied by the image itself and the secret key. However, the fluctuation around the mean error probability given by these two parameters are negligible when the tone-mapping result possesses sufficient visual quality; hence, the low

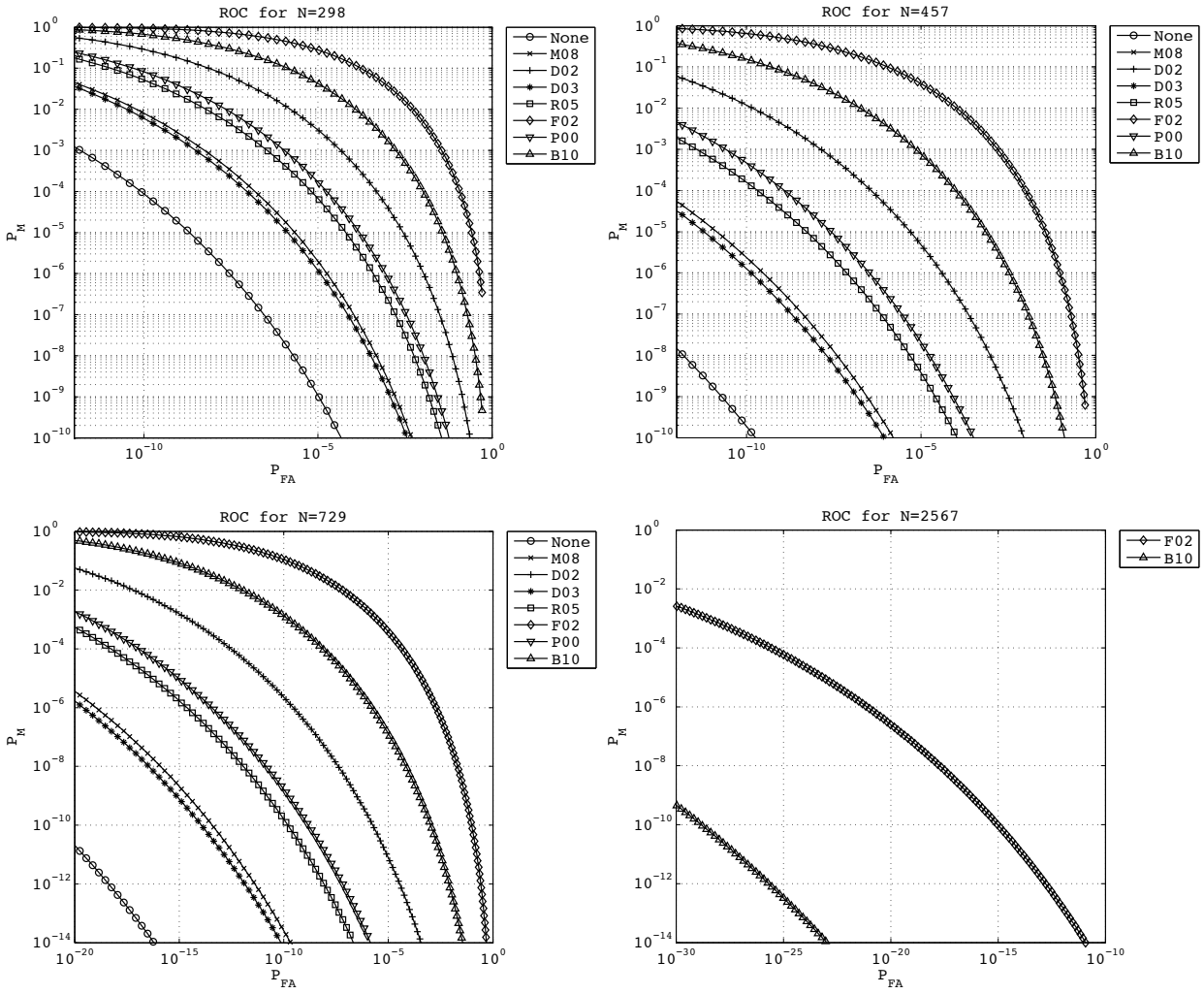
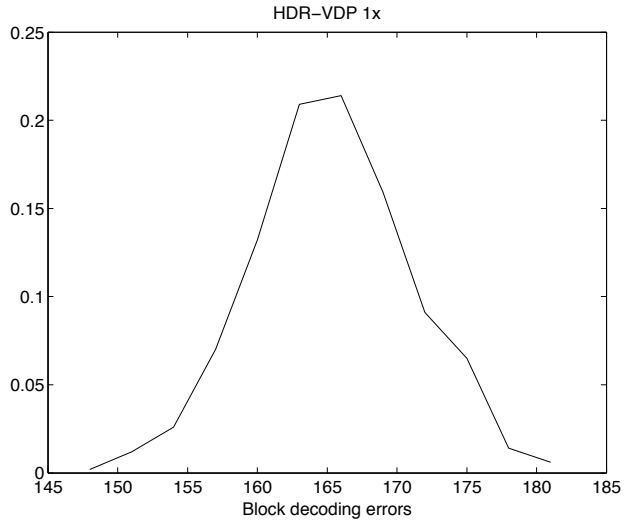


Figure 8: Estimated ROC for each of the TMOs, plotted for various values of the number of blocks N.

block decoding error probability obtained in these cases guarantees the excellent performance shown in Figure 8, especially for larger images.

E. Robustness against noise addition

Next, we also consider robustness against a more basic manipulation: noise addition. To comply with the high quality imagery paradigm of this work, noise addition is performed in such a way that its power is maximized while guaranteeing that its perceptibility remains low. To this aim, the perceptual mask is used and the noise is added directly in the embedding wavelet domain. First, a pseudo-random noise matrix with uniform, zero-centered distribution is extracted. The noise is 2-D low-pass filtered to increase its



(a) Histogram for the "Tree" image.

Image	P_M	DNR_{dB}
Apartment	$< 10^{-14}$	28.3
Montreal	$< 10^{-14}$	38.2
Chair	$< 10^{-14}$	27.8
Atrium2	$< 10^{-14}$	31.1
Colorcube	$< 10^{-14}$	31.3
Dandelion	$< 10^{-14}$	43.9
Hallway	$< 10^{-14}$	43.8
Splitcube	$< 10^{-14}$	35.6
AtriumNight	$< 10^{-14}$	38.5
Desk	$7 \cdot 10^{-14}$	30.5
Fog	$< 10^{-14}$	34.9
DaniBelgium	$6 \cdot 10^{-12}$	33.4
Memorial	$< 10^{-14}$	31.4
Rend01	$4 \cdot 10^{-11}$	35.7
Tree	$< 10^{-14}$	31.1

(b) P_M and mean DNR for all the images.

Figure 9: Robustness against mask-driven noise addition.

correlation and also locally scaled so as to evenly distribute its energy in the approximation coefficients subband. Then, the resulting pattern is multiplied by the perceptual mask and by a global strength factor; finally, it is added to the corresponding watermarked subband.

The results of this test are shown in Figure 9. The global strength parameter is manually set for each watermarked image such that the perceptual distortion in terms of HDR-VDP is about the same as that introduced by the watermarking process itself with respect to the original images. Each image underwent 1000 noise additions; the noise obviously causes more blocks to be incorrectly decoded. The histogram of the number of incorrectly decoded blocks is reported in Figure 9a for the "Tree" image as an example. To estimate the miss probability P_M, one could approximate the histogram with a normal distribution, thus computing its mean and variance, and then calculating the area of the distribution that goes beyond the threshold, fixed by $P_{FA} < 10^{-6}$, using the complementary error function. Table 9b shows such estimated P_M, along with the mean Data-to-Noise Ratio (DNR) of the noise attack for that image, expressed in dB. Both the DNR and P_M largely depend on the structure of the image through the mask. However, given the usually high noise energy, the worst miss probability is still quite satisfying, considering also the fact that the system is optimized for the "Tree" image.

F. Comparison with spread-spectrum watermarking

Last, we have run a comparison test using spread-spectrum watermarking, the other major paradigm in digital watermarking (see [1]). The test is confined to the "Tree" image and the embedding is performed in the same wavelet domain as the proposed method employing the algorithm discussed in [8]. The proposed log-luminance transform step is still applied to convert the originally LDR based algorithm to the HDR domain.

Since now the watermark is embedded in all of the AS subband, the column feature vector $\mathbf{x}$ indicates all of the coefficients. After having generated the watermark vector $\mathbf{z}_K$, which is extracted using the secret key K (in this case the watermark is a pseudo-random sequence composed of +1 and -1 of the same size of $\mathbf{x}$), the watermarked subband is:

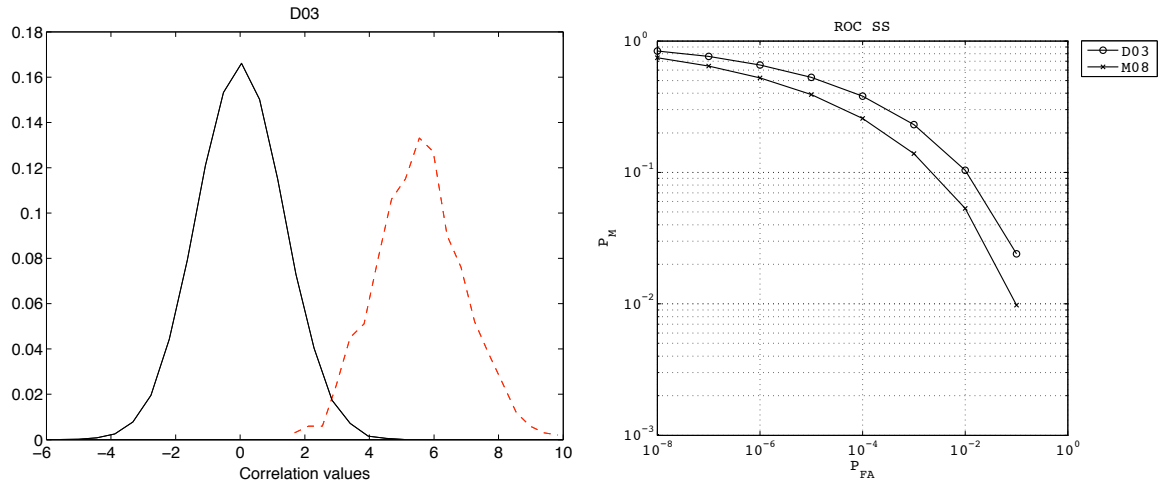
$$\mathbf{y} = \mathbf{x} + \gamma \cdot \mathbf{z}_K \cdot \mathbf{w} \tag{15}$$

where γ is a strength parameter, which as before is set such that the perceptual quality of the watermarked image is about the same as that obtained with our method, and $\mathbf{w}$ is the masking sequence obtained as in [8]. The watermark detection is correlation-based; given a received image, its embedding subband $\mathbf{y}'$ is extracted and the correlation is computed as:

$$\rho_K = \frac{\mathbf{z}_K^T \mathbf{y}'}{V} \tag{16}$$

where V is the size of the vectors. Then the correlation is checked against a threshold T_{ss} to assess the watermark presence. How the threshold is set depends on the system: for example, in [8] it is theoretically set according to the Neyman-Pearson criterion and by exploiting the expected normal distributions of the correlation under all the decoding hypothesis and some independence assumptions.

In this work, the Neyman-Pearson criterion is still used but an experimental framework is adopted instead. Given a "Tree" image watermarked with the secret key K and then tone-mapped using a given operator, the correlation $\rho_{K'}$ with a high number of extraneous watermarks (that is generated by different secret keys) is computed: an example histogram is plotted as the solid line in Figure 10a. As expected, a zero-mean normal distribution is obtained. By estimating its standard deviation, one could fix $T_s s$ such that $P_{FA} = p(\rho_{K'} > T_{ss})$. Then, 1000 other watermarked images using different keys are generated and subsequently tone-mapped. Eq. 16 is finally applied to compute the correlations of the latter images, using consistent keys, to obtain the correlation values in case of detection with matching keys. The histogram of these values are plotted as the dotted line in Figure 10a which report the D03 TM case; as expected again, it is a non zero-mean normal distribution. By estimating the standard deviation of the latter distribution, it



(a) Example with the D03 TM.

(b) ROC for D03 and M08 cases.

Figure 10: Comparison with spread-spectrum watermarking.

is possible to compute a projected miss probability P_M by evaluating the probability that the correlation value is less than T_{ss}, using the standard complementary error function. Figure 10b depicts the ROC curves for such estimations. Remarkably, all TM perform similarly in the spread-spectrum context; for this reason, only the worst and best cases are reported (D03 and M08 respectively). As it can be observed by comparison with Figure 8, the results of spread-spectrum watermarking applied in the same conditions as our method are considerably worse than those reported in Figure 8.

V. Conclusions

In this paper we presented an algorithm for a HDR image detectable watermarking system with the requirements of imperceptibility and robustness against tone-mapping operators as well as security. A previously developed watermarking system for grayscale LDR images has been employed in the LogLuv domain.

Experimental results have proven to be very good, especially considering how the design parameters have been set on a single image and then employed for the entire dataset. Automatic ways of setting these parameters will in fact be considered in our future work. The watermarks embedded using our system have always been detected, with the exception of the cases where the tone-mapping algorithms have given visually unsatisfactory output images. The system has also been compared to a basic spread-spectrum watermarking algorithm operating in the same domain.

As a final note, it could be interesting to switch to non-blind watermarking, which will probably be

another likely applicative scenario for HDR image watermarking. This would allow to choose which blocks to use for embedding, avoiding those difficult to watermark; to this aim, an extensive study of feature variability could be of great aid in determining in which zone of the image the watermarking system is more effective.

REFERENCES

[1] I.Cox, M. Miller, and J. Bloom, *Digital Watermarking*. Morgan Kaufman, 2001.

[2] M. Barni and F. Bartolini, *Watermarking Systems Engineering: Enabling Digital Assets Security and Other Applications*. CRC Press, 2004.

[3] F. Guerrini, R. Leonardi, and M. Barni, "Image watermarking robust to non-linear value-metric scaling based on higher order statistics," in *Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Toulouse, France, May 2006.

[4] G. W. Larson, "Logluv encoding for full-gamut, high-dynamic range images," *Journal of Graphics Tools*, vol. 3, no. 1, pp. 15–31, 1998.

[5] R. Mantiuk and H.-P. Seidel, "Modeling a generic tone-mapping operator," *Graphics Forum (Proc. of EUROGRAPHICS)*, vol. 27, no. 2, pp. 699–708, 2008.

[6] F. Guerrini, "Digital image watermarking robust against non-linear attacks," Ph.D. dissertation, University of Brescia, 2008.

[7] B. Chen and G. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding," *IEEE Trans. on Information Theory*, vol. 47, no. 4, pp. 1423–1443, 2001.

[8] M. Barni, F. Bartolini, and A. Piva, "Improved wavelet-based watermarking through pixel-wise masking," *IEEE Trans. on Image Processing*, vol. 10, no. 5, pp. 783–791, 2001.

[9] Munsell Color Science Laboratory Database. [Online]. Available: http://www.cis.rit.edu/mcsl/icam/hdr/rit_hdr

[10] Anyhere Software Database. [Online]. Available: http://www.anyhere.com/gward/hdrenc/pages/originals.html

[11] R. Mantiuk, S. Daly, and L. Kerofsky, "Display adaptive tone mapping," *ACM Trans. on Graphics*, vol. 27, no. 3, 2008.

[12] E. Reinhard and K. Devlin, "Dynamic range reduction inspired by photoreceptor physiology," *IEEE Trans. on Visualization and Computer Graphics*, vol. 11, pp. 13–24, 2005.

[13] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," *Eurographics*, 2003.

[14] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," *ACM Trans. on Graphics*, 2002.

[15] S. Pattanaik, J. Tumblin, H. Yee, and D. Greenberg, "Time-dependent visual adaptation for fast realistic image display," *Proc. of ACM SIGGRAPH*, 2000.

[16] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," *ACM Trans. on Graphics*, 2002.

[17] A. Boschetti, N. Adami, R. Leonardi, and M. Okuda, "High dynamic range image tone mapping based on local histogram equalization," in *Proc. IEEE Int. Conf. on Multimedia and Expo (ICME)*, Singapore, Singapore, July 2010.

[18] Pfstools website. [Online]. Available: http://pfstools.sourceforge.net/

[19] K. Myszkowski, R. Mantiuk, and H.-P. Seidel, "Visible difference predictor for high dynamic range images," *Int. Conf. on Systems, Man and Cybernetics*, pp. 2763–2769, 2004.