Dynamic range expansion based on image statistics

Belen Masia[1,2] · **Ana Serrano**[1] ·
Diego Gutierrez[1]

Abstract As the dynamic range of displays keeps increasing, there is a need
for reverse tone mapping methods, which aim at expanding the dynamic range
of legacy low dynamic range images for viewing on higher dynamic range
displays. While a number of strategies have been proposed, most of them are
designed for well-exposed input images and are not optimal when dealing with
ill-exposed (under- or over-exposed) content. Further, this type of content
is more prone to artifacts which may arise when using local methods. In this
work, we build on an existing, automatic, global reverse tone mapping operator
based on a gamma expansion. We improve this method by providing a new
way for automatic parameter calculation from the image statistics. We show
that this method yields better results across the whole range of exposures.

1 Introduction

The vast majority of the existing visual media, both static images and videos,
is coded in low dynamic range (LDR) format. This means that they can only
convey a limited contrast range, far from the rich contrasts found in real-
world scenes. For years, this was not much of an issue since existing displays
could only show very limited contrast ratios as well, so there was no need
for more sophisticated formats. In the last decade or so, however, there has
been a fast development in the field of high dynamic range (HDR) imaging.
Images can store increased contrast values now, orders of magnitude larger
than traditional LDR formats [5,23]. In order to visualize HDR images on
LDR displays, the data range must be compressed to fit the limited capabilities
of the display. This is the process known as *tone mapping*, for which a large

1 Universidad de Zaragoza. Maria de Luna 1. 50018 Zaragoza (Spain).
2 MPI Informatik. Campus E1 4. 66123 Saarbruecken (Germany).
E-mail: bmasiac@mpi-inf.mpg.de; {aserrano — diegog}@unizar.es

Fig. 1 Different examples of common images showing different degrees of overexposure. Contrary to what most existing rTMOs do, it is not convenient to expand the dynamic range of these images by boosting large, saturated areas

number of algorithms and strategies have been proposed over the years, see, e.g., [9] for a review in images and [12] for a review in video.

Modern displays can now show much higher dynamic ranges than before [25,19]. This has given rise to the inverse problem: how to correctly visualize legacy LDR content on these modern displays, for which the data range needs to be *expanded*. This is a much less explored field, generally called *reverse tone mapping* (or alternatively, *inverse* tone mapping).

Reverse tone mapping techniques aim at increasing the dynamic range of legacy low dynamic range images, for viewing on high dynamic range displays, generally to produce a result that matches real-world luminances as closely as possible [4]. Almost all existing reverse tone mapping operators (rTMOs) follow the same strategy: they first compute some sort of *expand map* based on detecting high-luminance areas of the LDR input images; pixels included in such expand map are then aggressively expanded, while the rest are either left unchanged or only slightly expanded. This strategy works well under the assumption that only a small percentage of pixels are saturated and their values need to be boosted; these usually represent light sources visible in the scene, or bright highlights on shiny surfaces.

However, an underlying assumption of most methods based on this strategy is that images are correctly exposed to begin with. This is often not the case, especially for content captured by amateurs with aim-and-shoot or cell-phone cameras (see Figure 1). Intuitively, while it makes sense to increase the value of a few saturated pixels in an image containing a bright light bulb, for instance, applying the same strategy to larger areas (such as a bright sky clamped to pure white by the sensor) may produce suboptimal, overly bright results. In those cases, it would be better to make visible details more prominent instead. Conversely, if the image is very dark, assigning most of the available range to the saturated or very bright areas (which will be minimal or non-existent) might not be the best strategy.

In previous work, Masia and colleagues [17] analyzed existing rTMOs by means of several user studies, and made the key observation that their performance did in fact decrease significantly when the input images were overexposed. Based on these findings, they proposed a proof-of-concept global rTMO based on a simple gamma expansion, which was shown to outperform more sophisticated algorithms across a range of overexposure levels. To obtain a suitable γ value for each image, they proposed computing the *key* of the input image, as an indicator of whether the scene is subjectively dark or light. This

computation, however, was based on a simple linear regression with the key of the images. While it was designed for overexposed (high-key) images, it can yield negative values for key values low enough, which need to be clamped [5].

In this work, we aim to obtain an rTMO with the following desired characteristics: i) it does not rely on the assumption that the input images are correctly exposed; ii) it is simple enough that modern displays can compute it on the fly; and iii) it is a global operator, to avoid possible artifacts that arise from local approaches, especially prone to cause problems in the case of poorly exposed content [17]. We choose a γ-based approach as in Masia et al.'s work, since it has been shown to have beneficial characteristics for our particular problem: γ functions tend to darken the overall appearance of images while increasing contrast, which has been shown to be preferred by viewers [20,24]. We thoroughly explore simple image statistics, and propose new, robust models that are straightforward to compute, while showing a higher correlation with the data gathered in previous user studies [17]. In particular, based on the new statistical analysis in this work, different parametric regressions are obtained together with various metrics, to assess how well they fit the data. Both multilinear regression based on linear least squares, and a robust regression based on iteratively reweighted least squares are tested, always seeking to obtain the best trade-off between accuracy and simplicity of the model. Based on these, we propose an expansion that yields consistent results over a range of exposures, and show that it outperforms that of Masia et al., especially for lower exposures.

2 Previous Work

As stated before, most previous rTMOs follow the same basic approach: first, the brightest areas of the image are identified (usually by simple thresholding), yielding some sort of *expansion map*. This includes all pixels clamped to pure white due to limitations in the sensor. Those bright areas are then significantly expanded by applying different dynamic range expansion functions, whereas the rest of the pixels of the image are left untouched or only slightly modified.

One of the pioneering works in the field is that of Banterle et al. [6,8], who define their expansion map by density estimation of the bright areas in the input image, and use the inverse of Reinhard's tone mapping operator [22] to boost the dynamic range. A further extension of the method allows handling video by ensuring temporal coherence of the expansion map between frames [7]. Meylan and colleagues [20,21] opt for a similar yet simpler approach, applying a steep linear expansion to saturated areas, and a less steep linear expansion function to the rest. Rempel et al. [24] define a brightness enhancement function (BEF), similar to an expansion map, computed by blurring a mask of all saturated pixels. Carefully selected edge-stopping functions ensure that contrast is maintained along strong edges in the input image. In another work, a bilateral grid is used to compute this edge-preserving BEF in real time [14];

this work is later improved to be able to handle a wide range of exposures of the input image, and to be compatible with any bilateral filtering strategy [15].

While all the methods discussed above are automatic, other works exist that rely on user input. In the work by Didyk and colleagues [10], a trained classifier labels bright image areas as belonging to one of four different categories: diffuse surfaces, light sources, specular highlights and reflections. This allows the system to apply different expansion functions to different parts of the image, even though their saturated values may be the same. Masia et al. [18] present two different approaches. In the first, the image is classified according to Ansel Adams' Zone System [1], while in the second saliency detection models predict the most important parts of the image. In both cases, the user can then control different expansion functions to specific areas of the image.

The work of Masia et al. [17] evaluated a number of existing rTMOs for varying degrees of exposure of the input images, and formally showed that their performance decreased for increasingly overexposed imagery, to the extent that sometimes the original LDR image was preferred over an expanded HDR version. The problem of exposure in rTM was first identified by Martin et al. [16], where the authors correlate user preferences for different scenes and exposures with first order image statistics. The authors do not to develop a new rTMO, and instead conclude that deeper statistical analysis should be carried out for that purpose. Masia et al. do propose a content expansion method: They make the observation that a simple, global rTMO can outperform more complex strategies in these situations, and propose an expansion based on a simple gamma curve. In this work, we do opt for a global method to avoid the risk of artifacts, which are more prone to appear in suboptimal input content such as the one we are dealing with, and in particular we build on the proposal of Masia et al. [17], based on a gamma curve, since it proved to work well on ill-exposed content. However, we improve on the automatic computation of the gamma value (γ) for this curve. To obtain the specific value of γ^1 for each image, Masia et al. simply employ its *key* value, which gives an intuition of its overall brightness; however, as mentioned before, it is intended for overexposed content and thus its performance decreases for dark (low key) images. We propose a better computation of the γ value that works well across a wide range of exposures while ensuring the absence of artifacts thanks to being a global operator.

3 Multilinear Models for Gamma Expansion

Our goal, as mentioned, is that of building an automatic, global rTMO based on a gamma expansion; this section aims at finding the best method for automatically obtaining the γ value for the expansion on a per-image basis. This

---

[1] From this point on, we will use the symbol γ to refer to the parameter of the expansion curve, and the word *gamma* when referring to the expansion method or the type of curve.

is done by fitting manually-obtained values of γ available for a set of images to a function of image statistics using multilinear models.

3.1 Input data

We use the image database provided by Martin et al. in their study on the perception of exposure [16] (see Figure 2) for learning image statistics for our model. This database was also used in Masia et al.'s work [17], so it allows us to directly compare our results to theirs in terms of which model fits better the data.

Our goal is to find a gamma-based rTMO, for which we need a way to compute the specific γ value for each given image. We follow a similar approach as previous works, and use manually-adjusted values of γ for each image as ground truth: These were obtained in a pilot study, asking participants to tweak the γ curve until they found the subjective best rendition of the image [17]. Table 1 shows the results of this manual process. Columns 1 to 4 indicate increasing exposure, while rows correspond to the different scenes captured. Images corresponding to this table are shown in Figure 2.
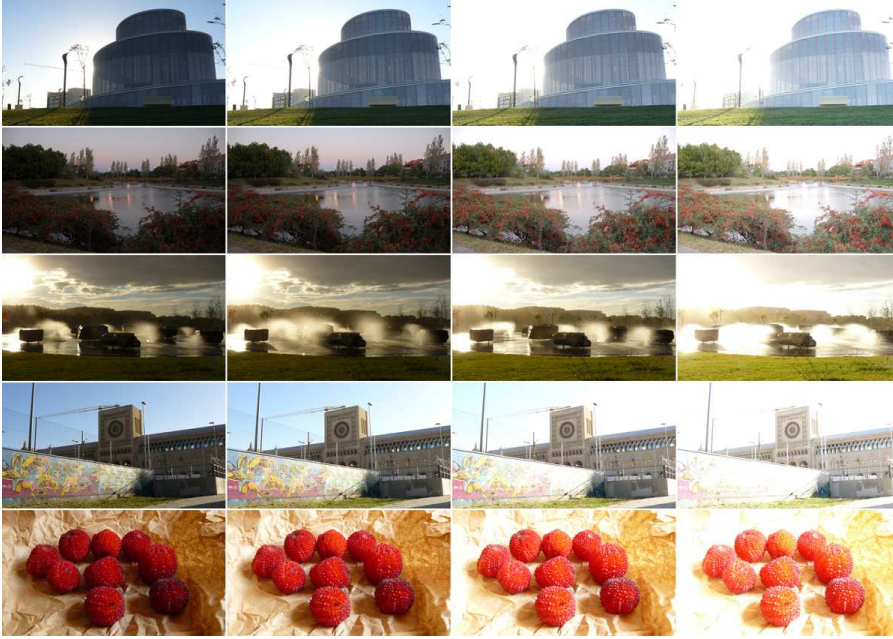

Fig. 2 The images used in this work for training the model, showing increasing degrees of overexposure. From top to bottom: *building, lake, sunset, graffiti* and *strawberries* [16]

To compute image statistics from which to derive an expression to obtain γ for each image, we first linearize sRGB values and then compute luminance [23]

Table 1 Manually-adjusted γ values (adapted from [17])

Scene	1	2	3	4
Building	1.22	1.5	1.75	2.6
Lake	1.1	1.2	1.5	2.25
Sunset	1.1	1.35	1.4	1.75
Graffiti	1.2	1.35	1.5	1.75
Strawberries	1.22	1.35	1.55	1.9

L as:

$$L = Y = 0.2126R + 0.7152G + 0.0722B, \tag{1}$$

where L is in the range [0..1].

The first statistical values we compute are both the arithmetic and the geometric mean luminance (referred to as L_{avg} and L_H, respectively). The arithmetic mean is simply obtained by averaging the luminance value of all pixels ($L_{avg} = 1/N \sum_{i=1}^{N} L(i)$, with N being the total number of pixels in the image); the geometric mean, known to reduce the contribution of outliers, is obtained as follows [5]:

$$L_H = exp\left(\frac{1}{N} \sum_{i=1}^{N} log(L(i) + \varepsilon) \right), \tag{2}$$

where ε is a very small positive number to prevent singularities in black pixels. We additionally compute the logarithm of this quantity, simply $logL_H$.

The key of the images is also obtained, using the following equation [2]:

$$k = \frac{logL_H - logL_{min}}{logL_{max} - logL_{min}}. \tag{3}$$

In this equation L_{max} and L_{min} are the maximum and minimum luminance values, respectively, once a percentage of outlier pixels (both on the dark and bright sides) has been eliminated. We calculate two key values, k_5 and k_1, considering 5% or 1% of the pixels as outliers, respectively.

Additionally, both the median, L_{med}, and a series of central moments, are computed for the luminance of the images. These include variance V_L (and standard deviation σ_L), skewness ($skew_L$) and kurtosis ($kurt_L$). Finally, we compute the percentage of overexposed pixels for each of the images, defining overexposed pixels as those with $L \cdot 255 \geq 254$; we will refer to it as p_{ov}.

Tables 2 and 3 include the values obtained for each of the aforementioned statistics for the images of our dataset. These values are the ones used for the regressions explored in the following subsections.

3.2 Multilinear Regression

The term multilinear regression refers to a linear regression with multiple variables as predictors. Restricting ourselves to linear regressions was decided

Table 2 Statistics for the images in the dataset (1/2)

Image	γ	L_{avg}	L_H	$logL_H$	k_5	k_1	L_{med}
Building01	1.22	0.3493	0.1182	-2.1352	0.5743	0.6019	0.0762
Building02	1.5	0.4853	0.2485	-1.3922	0.6472	0.6775	0.2076
Building03	1.75	0.5792	0.4052	-0.9033	0.6865	0.7265	0.4176
Building04	2.6	0.7105	0.6196	-0.4787	0.7392	0.7806	0.7399
Lake01	1.1	0.1188	0.0338	-3.3881	0.4963	0.5612	0.0248
Lake02	1.2	0.1662	0.0570	-2.8651	0.5151	0.5721	0.0473
Lake03	1.5	0.3689	0.2020	-1.5996	0.5545	0.6418	0.1827
Lake04	2.25	0.4977	0.3613	-1.0182	0.5975	0.6874	0.3668
Sunset01	1.1	0.2088	0.0719	-2.6319	0.4857	0.5622	0.0784
Sunset02	1.35	0.2633	0.1143	-2.1688	0.5252	0.6170	0.1361
Sunset03	1.4	0.3930	0.2259	-1.4875	0.5884	0.6907	0.2675
Sunset04	1.75	0.6633	0.5505	-0.5969	0.6891	0.8009	0.7168
Graffiti01	1.2	0.2891	0.1568	-1.8525	0.6213	0.6726	0.2744
Graffiti02	1.35	0.5020	0.3335	-1.0981	0.6659	0.7405	0.5732
Graffiti03	1.5	0.6796	0.5449	-0.6071	0.7201	0.8074	0.8761
Graffiti04	1.75	0.8091	0.7415	-0.2991	0.7761	0.8591	0.9949
Strawberries01	1.22	0.1718	0.0954	-2.3501	0.5646	0.6074	0.1075
Strawberries02	1.35	0.3381	0.2240	-1.4962	0.5713	0.6335	0.2544
Strawberries03	1.55	0.5304	0.4098	-0.8921	0.6235	0.6884	0.5048
Strawberries04	1.9	0.6961	0.6046	-0.5033	0.6942	0.7524	0.8147

Table 3 Statistics for the images in the dataset (2/2)

Image	γ	V_L	σ_L	$skew_L$	$kurt_L$	p_{ov}
Building01	1.22	0.1357	0.3684	0.5261	1.6310	5.9764
Building02	1.5	0.1774	0.4212	0.1879	1.1346	20.04220
Building03	1.75	0.1527	0.3908	0.0275	1.1750	40.1717
Building04	2.6	0.0912	0.3020	-0.4495	1.7895	44.3981
Lake01	1.1	0.0256	0.1601	1.1686	2.6714	0.0003
Lake02	1.2	0.0450	0.2121	1.1497	2.6493	0.0017
Lake03	1.5	0.1316	0.3628	0.8654	2.1098	17.8714
Lake04	2.25	0.1225	0.3500	0.4138	1.5494	22.8514
Sunset01	1.1	0.0786	0.2803	1.6711	4.7589	4.0931
Sunset02	1.35	0.0888	0.2980	1.3072	3.5468	5.0731
Sunset03	1.4	0.1152	0.3394	0.6207	1.9193	9.4525
Sunset04	1.75	0.1025	0.3202	-0.3523	1.6076	29.5133
Graffiti01	1.2	0.0557	0.2361	0.3878	1.8997	0.0039
Graffiti02	1.35	0.1110	0.3332	-0.0818	1.4351	1.0967
Graffiti03	1.5	0.1158	0.3403	-0.5197	1.5827	21.1744
Graffiti04	1.75	0.0692	0.2631	-1.0777	2.8062	51.2861
Strawberries01	1.22	0.0283	0.1681	1.2238	3.6194	0.0000
Strawberries02	1.35	0.0729	0.2700	0.7483	2.3469	0.0781
Strawberries03	1.55	0.1037	0.3220	0.1416	1.5362	8.9808
Strawberries04	1.9	0.0934	0.3057	-0.4674	1.6204	26.5756

to keep the model as simple as possible; if a good model could not be found assuming a linear relationship, we would move on to more complex fittings. We initially used ordinary least squares to do the fittings. This implies a series of assumptions over the errors, mainly that they are normally distributed, with constant variance, and independent of each other. It also implies that the independent variables are free of error, or that their error is insignificant compared to the error of the dependent variable.

Once the type of model (i.e., linear) has been chosen, the problem which arises when working with multiple predictors is knowing which of the possible predictors (i.e., the independent variables, in our case the calculated image statistics) should be included in the model and which should be left out. The way in which we deal with this is performing F-tests over the possible models. Computing the R^2 value or another goodness-of-fit metric and comparing their values for both models is typically not enough. The reason for this is that given two models, A and B, with p_A and p_B terms, respectively, if $p_A > p_B$, model A will always fit the data at least as good as model B. Thus, what has to be found out is if the addition of that extra parameter(s) to model A gives a *significantly* better fitting; as mentioned, we make use of F-statistics to assess that. In the following we describe this process in detail, while Appendix A describes the use of F-tests in the construction of multiple variable models.

3.2.1 Building multiple variable models

A stepwise regression is used to build the possible multilinear models [26]. The idea is to start from an initial model, typically with the single variable showing the highest correlation with the data. Then, the steps described below are performed iteratively until no more variables can be added.

- **Step 1:** Test all variables independently with partial F-statistics. The variable with the lowest p-value is added to the model (as long as the p-value is lower than the significance level, p-value $< \alpha$). This step is repeated until no more variables can be added.
- **Step 2:** Since the addition of a new variable can cause a previously introduced value to be non-significant, a partial F-statistic is computed for each variable currently in the model as if it were the one introduced in the last place. The one with the largest p-value is removed (as long as the p-value is larger than twice the significance level, p-value $> 2\alpha$), and a new iteration begins with Step 1.

The process ends when, in Step 2, no variables are removed because there are no variables whose p-value is larger than the significance level (which implies that no variables can be added either). Please note that even though some combinations of variables may not have much sense (e.g. having k_5 and k_1 both included in a regression) we initially make exhaustive tests without taking into consideration these constraints, which will be evaluated over the final model selected.

It is clear from the procedure above that the initial model selected and the order in which variables are selected to be introduced or removed from the model determines the final model reached. For this reason, we repeated the algorithm for stepwise regression starting from every possible single variable initial model. This yields different final models, for which a series of metrics are computed in order to evaluate the accuracy of the fitting. In particular we compute the RMSE and the *overall* F-statistic for each model obtained (see Appendix B for a definition of these parameters). Based on them we select the best model, which yields the following equation for the regression:

$$\gamma = 3.8872 + 0.3752 log L_H - 2.9941 k_1 + 0.0160 p_{ov}. \qquad (4)$$

Figure 3 shows the observed γ values against the γ values predicted by this model. Additionally, it compiles different metrics assessing the accuracy of the previous regression, including R^2 and $\tilde{R}^2$ (see Appendix B for a definition of $\tilde{R}^2$) as well as the already mentioned RMSE and F-statistic with its associated p-value.
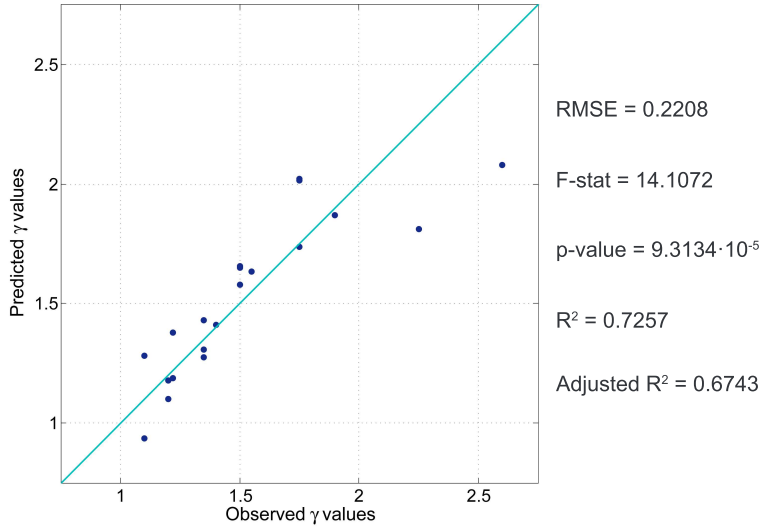

RMSE = 0.2208

F-stat = 14.1072

p-value = $9.3134 \cdot 10^{-5}$

R^2 = 0.7257

Adjusted R^2 = 0.6743

Fig. 3 Predictive accuracy of the regression shown in Equation 4. The x-axis shows observed γ values, while the y-axis depicts the values predicted by the regression. The cyan line shows the quadrant bisection (i.e., y=x)

3.3 Robust regression

When performing an outlier analysis over the residual of the model obtained in Section 3.2, there are two observations which systematically appear as outliers in the fitted models (the 95% confidence interval for the mean of their residuals

does not include zero). One option is therefore to remove these two values (*Building04* and *Lake04*) from the data and re-do the multilinear regression fitting.

An alternative to eliminating the observations regarded as outliers is to retain all the observed data but weight their influence when computing the regression. To do this, we perform a new regression with the same predictors as in Equation 4 using iteratively reweighted least squares. The weight function used is a bisquare function. The new regression is thus given by the following equation:

$$\gamma = 2.4379 + 0.2319 logL_H - 1.1228k_1 + 0.0085p_{ov}. \tag{5}$$

Figure 4 shows the predictive accuracy of the model obtained by robust regression compared to Ordinary Least Squares, both with all the observed data (*left*) and after outlier removal (*right*). Additionally, if we compute a robust RMSE estimate for this last regression [11], we obtain an estimate of 0.0962 (while estimates for the previous ones, OLS with and without oultiers, were 0.2208 and 0.0664 respectively).
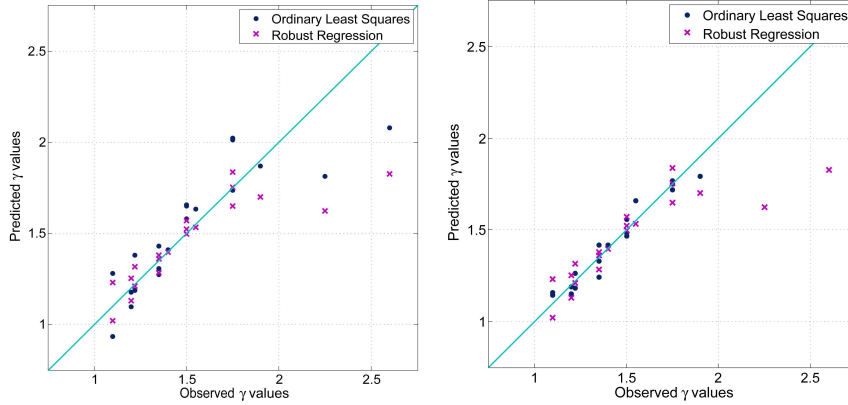

Fig. 4 Predictive accuracy of the model obtained by robust regression against the ones obtained by ordinary least squares, with all the observed data (*left*) and after outlier removal (*right*). The abscissae show observed γ values, while the y-axes depict the values predicted by the regression. The cyan lines mark the quadrant bisection

4 Results

We evaluate the results of our multilinear model with a collection of images from the HDR Photographic Survey [13], *different* from the set of images we used to learn our model. For a collection of LDR images with different levels of exposure we perform the expansion with the γ estimated with our multilinear robust regression model (Equation 5) as well as with the γ prediction from the

linear model proposed by Masia et al. [17]. The images used for this evaluation can be seen in Figure 5.
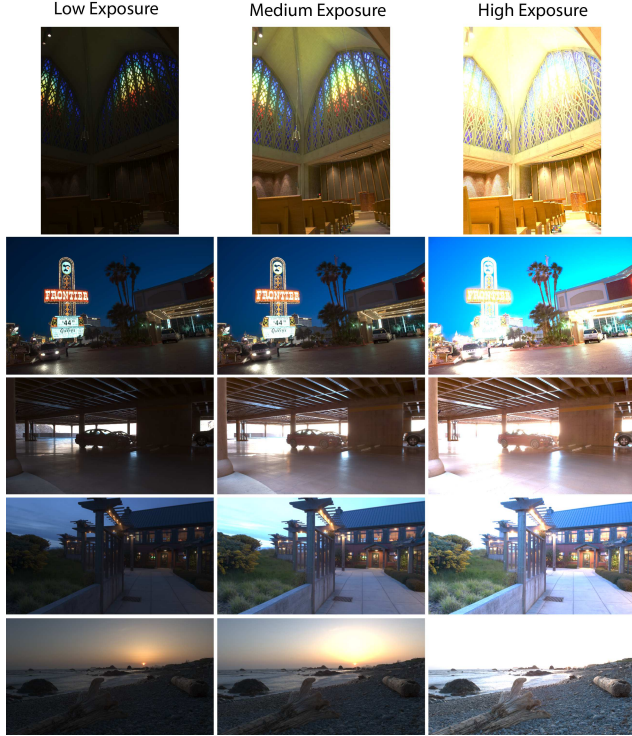

Fig. 5 Images used in this work for the evaluation of the model showing, from left to right, increasing degrees of exposure. From top to bottom: *Chapel, Frontier, Hoovergarage, McKeesPub,* and *Redwoodsunset* [13]

We first compute the γ values for all the test images and compare them with the γ values predicted in Masia et al.'s work. Figure 6 reveals that, while estimated γ values in bright images (higher *key*) are very close in both models, when images become darker, the two different predictions tend to differ. To find out which of these models yields better results we perform an expansion of the set of LDR images with the γ values obtained from Masia et al.'s linear model and those obtained with our multilinear model (Equation 5), and then evaluate the visual improvement of the expansion with respect to the original LDR image. For this purpose we use the image quality metric proposed by Aydin et al. [3], which identifies visible distortions between two images independently of their respective dynamic ranges. The metric uses a model of the human visual system, and visible changes between a reference and a test image are classified into three types of structural changes: loss of visible contrast (when contrast visible in the reference image becomes invisible
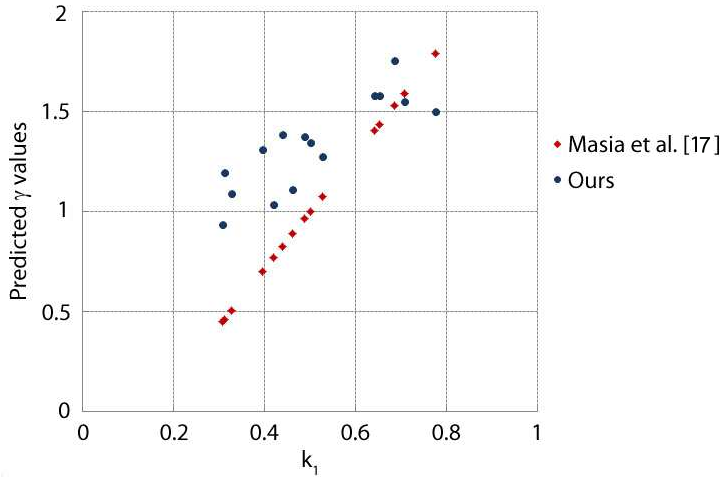
Fig. 6 Estimated γ values with Masia et al.'s linear model and with our multilinear model. The x-axis depicts the *key* of the image with 1% of the darker and brighter pixels removed. It can be seen that, the lower the *key* is, the more the two models tend to differ. Figure 7 later shows how the γ values obtained by our regression yield comparable results to those of Masia et al. for brighter images while clearly performing better for dark images

in the test one), amplification of invisible contrast (when invisible contrast in the reference image becomes visible in the test one), and reversal of visible contrast (when contrast polarity is reversed in the test image with respect to the reference). It is important to remember that, as noted by Rempel and colleagues [24], contrast enhancement usually increases perceived quality and is therefore a desired output of the rTMO. The results of this metric for two of the images of the test collection for a variety of exposures can be seen in Figure 7 (for the results corresponding to the rest of the images tested please refer to the supplementary material[2]). The metric reveals that, while performing similarly good in bright images, our multilinear model performs better in dark images. We have observed that, the lower the *key* of the image, the better our performance is compared with the model proposed by Masia et al. [17].

Finally, we show in Figures 8 and 9, for the lower exposures of our test set, the tonemapped HDR resulting from the expansion with the γ calculated with our multilinear model as well as with Masia et al.'s linear model and we confirm the visual improvement, especially in terms of improved contrast, of our model over Masia et al.'s work, specifically for under-exposed content.

---

[2] Available at:
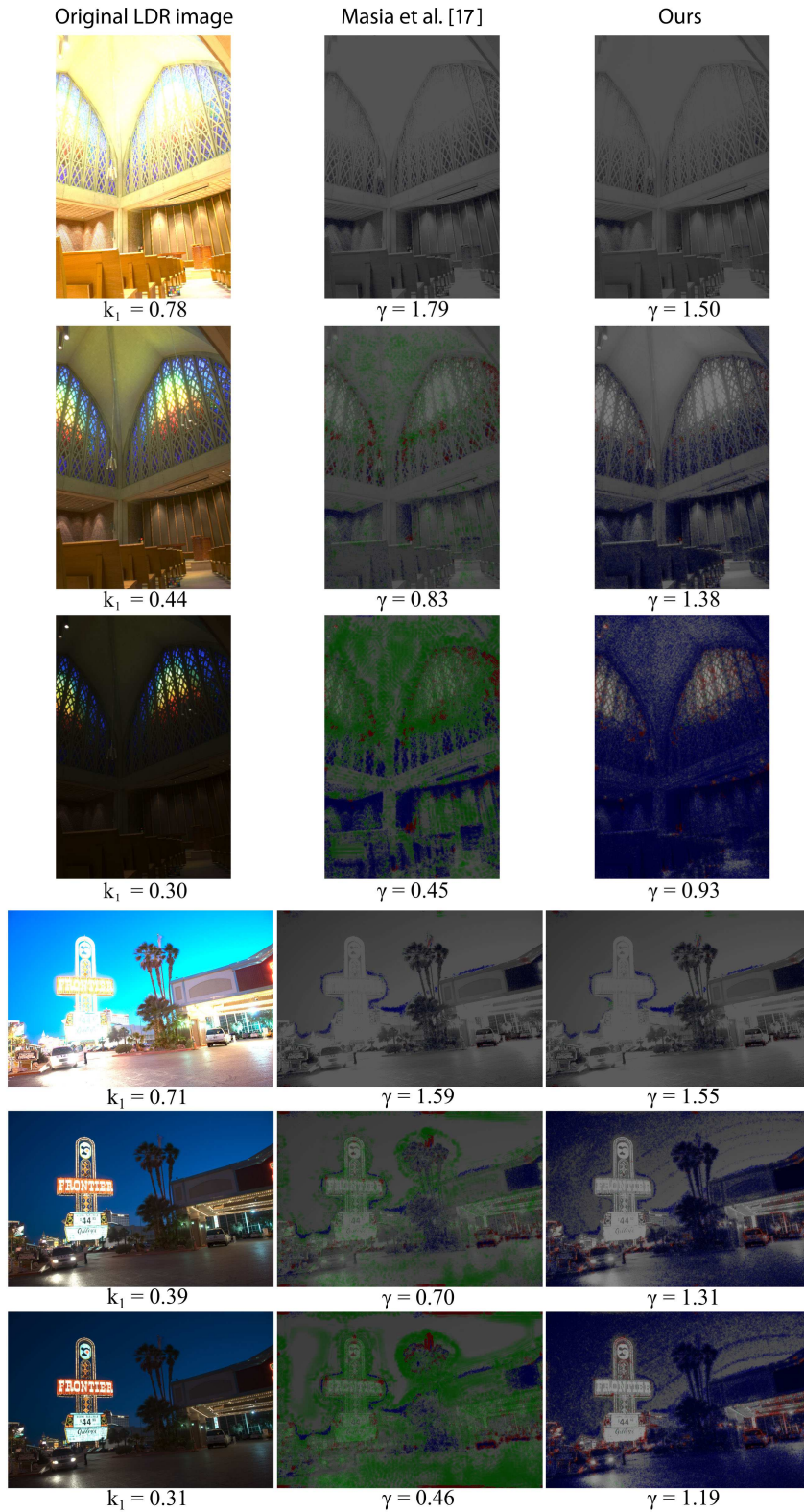http://webdiis.unizar.es/~bmasia/downloads/MMTA_Masia_SupplementaryMaterial.zip

Original LDR image	Masia et al. [17]	Ours
$k_1 = 0.78$	$\gamma = 1.79$	$\gamma = 1.50$
$k_1 = 0.44$	$\gamma = 0.83$	$\gamma = 1.38$
$k_1 = 0.30$	$\gamma = 0.45$	$\gamma = 0.93$
$k_1 = 0.71$	$\gamma = 1.59$	$\gamma = 1.55$
$k_1 = 0.39$	$\gamma = 0.70$	$\gamma = 1.31$
$k_1 = 0.31$	$\gamma = 0.46$	$\gamma = 1.19$

Fig. 7 Comparing the results with the image quality metric [3]. Expansion with the γ obtained in [17] and our γ, for the images *Chapel* and *Frontier* [13]. For each image we show the original LDR for different degrees of exposure together with its key k_1 and the result of the metric together with the γ value used for the expansion. Green, blue and red identify loss of visible contrast, amplification of invisible contrast and contrast reversal respectively. Our gamma expansion shows a better performance the darker (lower *key*) the image is (recall that amplification of contrast is desirable [24]), while yielding comparable results in high-key images. Please refer to the digital version of the article for full appreciation of the details
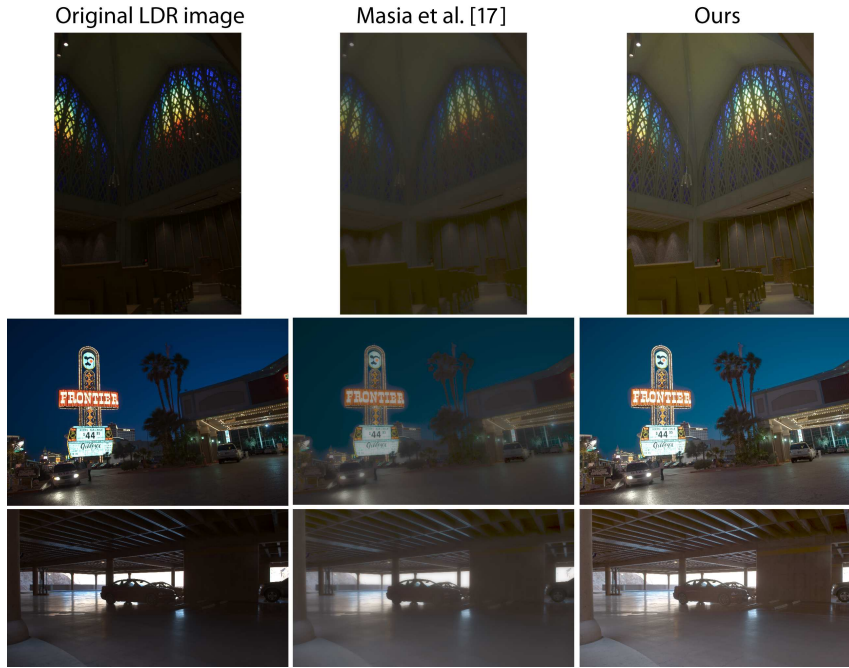
Fig. 8 Original LDR image and tonemapped result of the expansion with the γ obtained in [17] and our γ. From top to bottom: *Chapel, Frontier,* and *Hoovergarage* [13]. Note the increased contrast yielded by our method. Please refer to the digital version of the article for full appreciation of the details

Fig. 9 Original LDR image and a cropped region of the tonemapped result of the expansion with the γ obtained in [17] and our γ. Top row: *McKeesPub*; bottom row: *Redwood-sunset* [13]. Note the increased contrast yielded by our method. Please refer to the digital version of the article for full appreciation of the details

5 Conclusions

In this work we have presented an automatic, global rTMO, fast enough to be computed on the fly and based on a simple gamma expansion following insights by Masia et al. [17] about user preferences on expanded content. We have computed image statistics from a set of varied images and have derived a multilinear model that fits this data and estimates the γ value to be used to expand the LDR content.

We have validated our model on a collection of images different from the ones we used to fit the model. We have demonstrated that the γ value obtained with our robust multilinear model yields results comparable to Masia et al.'s work for overexposed content, while outperforming their linear model in underexposed images. Across a wide range of exposures, we have shown greater amplification of contrast invisible in the original LDR image while avoiding loss and reversal of contrast.

6 Acknowledgements

References

1. ADAMS, A. *The Print.* The Ansel Adams Photography series. Little, Brown and Company, 1983.
2. AKYUZ, A. O., AND REINHARD, E. Color Appearance in High Dynamic Range Imaging. *Journal of Electronic Imaging 15*, 3 (2006).
3. AYDIN, T. O., MANTIUK, R., MYSZKOWSKI, K., AND SEIDEL, H.-P. Dynamic range independent image quality assessment. *ACM Trans. Graph 27*, 3 (2008), 69.
4. BANTERLE, F., ARTUSI, A., AYDIN, T., DIDYK, P., EISEMANN, E., GUTIERREZ, D., MANTIUK, R., AND MYSZKOWSKI, K. Multidimensional image retargeting. In *ACM Siggraph ASIA 2011 Courses* (December 2011), ACM Siggraph ASIA, ACM.
5. BANTERLE, F., ARTUSI, A., DEBATTISTA, K., AND CHALMERS, A. *Advanced High Dynamic Range Imaging: Theory and Practice.* A.K. Peters, Ltd., 2011.
6. BANTERLE, F., LEDDA, P., DEBATTISTA, K., AND CHALMERS, A. Inverse tone mapping. In *GRAPHITE '06* (New York, NY, USA, 2006), ACM, pp. 349–356.
7. BANTERLE, F., LEDDA, P., DEBATTISTA, K., AND CHALMERS, A. Expanding low dynamic range videos for high dynamic range applications. In *Proceedings of the Spring Conference on Computer Graphics* (New York, NY, USA, 2008), ACM.
8. BANTERLE, F., LEDDA, P., DEBATTISTA, K., CHALMERS, A., AND BLOJ, M. A framework for inverse tone mapping. *Vis. Comput. 23*, 7 (2007), 467–478.
9. ČADÍK, M., WIMMER, M., NEUMANN, L., AND ARTUSI, A. Evaluation of hdr tone mapping methods using essential perceptual attributes. *Computers & Graphics 32* (2008), 330–349.
10. DIDYK, P., MANTIUK, R., HEIN, M., AND SEIDEL, H. Enhancement of bright video features for HDR displays. *Computer Graphics Forum 27*, 4 (2008), 1265–1274.

11. DUMOUCHEL, W., AND O'BRIEN, F. *Integrating a robust option into a multiple regression computing environment.* Springer-Verlag New York, Inc., New York, NY, USA, 1991, pp. 41–48.

12. EILERTSEN, G., WANAT, R., MANTIUK, R., AND UNGER, J. Evaluation of tone mapping operators for hdr-video. *Computer Graphics Forum (Proc. of Pacific Graphics 2013) 32* (2013).

13. FAIRCHILD, M. D. The HDR Photographic Survey. `http://rit-mcsl.org/fairchild/HDR.html`.

14. KOVALESKI, R. P., AND OLIVEIRA, M. M. High-quality brightness enhancement functions for real-time reverse tone mapping. *The Visual Computer 25*, 5-7 (April 2009), 539–547.

15. KOVALESKI, R. P., AND OLIVEIRA, M. M. High-quality reverse tone mapping for a wide range of exposures. In *Proceedings of the Conference on Graphics, Patterns and Images (SIBGRAPI 2014)* (2014), pp. 49–56.

16. MARTIN, M., FLEMING, R., SORKINE, O., AND GUTIERREZ, D. Understanding exposure for reverse tone mapping. In *Congreso Español de Informática Gráfica* (2008), pp. 189–198.

17. MASIA, B., AGUSTIN, S., FLEMING, R. W., SORKINE, O., AND GUTIERREZ, D. Evaluation of reverse tone mapping through varying exposure conditions. *ACM Transactions on Graphics (Proc. of SIGGRAPH Asia) 28*, 5 (2009), 160:1–160:8.

18. MASIA, B., FLEMING, R. W., SORKINE, O., AND GUTIERREZ, D. Selective reverse tone mapping. In *Congreso Español de Informatica Grafica* (2010), Eurographics.

19. MASIA, B., WETZSTEIN, G., DIDYK, P., AND GUTIERREZ, D. A Survey on Computational Displays: Pushing the Boundaries of Optics, Computation, and Perception. *Computers & Graphics 37*, 8 (2013), 1012 – 1038.

20. MEYLAN, L., DALY, S., AND SÜSSTRUNK, S. The reproduction of specular highlights on high dynamic range displays. In *IS&T/SID 14th Color Imaging Conference* (2006).

21. MEYLAN, L., DALY, S., AND SÜSSTRUNK, S. Tone mapping for high dynamic range displays. In *Proc. IS&T/SPIE Electronic Imaging: Human Vision and Electronic Imaging XII* (2007), vol. 6492.

22. REINHARD, E., STARK, M., SHIRLEY, P., AND FERWERDA, J. Photographic tone reproduction for digital images. *ACM Trans. Graph. 21*, 3 (2002), 267–276.

23. REINHARD, E., WARD, G., PATTANAIK, S., DEBEVEC, P., HEIDRICH, W., AND MYSZKOWSKI, K. *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting. 2nd Edition.* Morgan Kaufmann, 2010.

24. REMPEL, A. G., TRENTACOSTE, M., SEETZEN, H., YOUNG, H. D., HEIDRICH, W., WHITEHEAD, L., AND WARD, G. Ldr2Hdr: on-the-fly reverse tone mapping of legacy video and photographs. *ACM Trans. Graph. 26*, 3 (2007), 39.

25. SEETZEN, H., HEIDRICH, W., STUERZLINGER, W., WARD, G., WHITEHEAD, L., TRENTACOSTE, M., GHOSH, A., AND VOROZCOVS, A. High dynamic range display systems. *ACM Trans. Graph. 23*, 3 (2004), 760–768.

26. STEIGER, J. H. Introduction to multiple regression. `http://www.statpower.net/Content/312/Lecture%20Slides/MultipleRegressionIntro.pdf`. Online, last accessed 08-August-2015.

A F-tests for assessing the appropriateness of adding new predictors to a model

An F-test is typically performed to decide whether or not a certain null hypothesis can be rejected. To do this, a test statistic (the F-statistic) is needed which under the null hypothesis follows an F-distribution. In our case, the null hypothesis is that, given two models, A and B, with a number of predictors p_A and p_B ($p_A > p_B$), the two models fit equally well the data. The F-statistic is then given by:

$$F_{p_A - p_B, n - p_A} = \frac{(SS_B - SS_A)/(p_A - p_B)}{SS_A/(n - p_A)}, \tag{6}$$

where SS_i, $i = \{A, B\}$, is the sum of squared residuals of model i, and n is the number of data values [26]. It must be noted that in Equation 6, and throughout the document, p_i as a measure of the number of terms in the regression includes the constant term (i.e. the intercept).

For the particular case of creating model A by adding one variable to a model B that has p terms, and expressing the formula in terms of R^2, the F-statistic becomes:

$$F_{1,n-p-1} = \frac{R_A^2 - R_B^2}{(1 - R_A^2)/(n - p - 1)} \tag{7}$$

As it is well known, given a value for F in an F-test, the p-value is the probability of obtaining a value as extreme as the F obtained, assuming that the null hypothesis is true. As a consequence, the null hypothesis is typically rejected if the p-value is lower than the significance level α (which, in this work, has the usual value of $\alpha = 0.05$).

B Goodness of fit in multilinear regressions

This appendix includes the description of a series of metrics which are typically used in regression analysis to measure the accuracy of the fitting of a certain model.

RMSE. For a multilinear regression, RMSE is computed as shown in Equation 8, where Y_i are the observed data (i.e. the given γ values) and $\hat{Y}_i$ the data predicted by the model.

$$RMSE = \sqrt{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2/(n - p)}, \tag{8}$$

where, n is the data size and p the number of terms in the regression. Please recall that in this formulation the intercept is included in p. This metric provides an intuition on the error we would incur in when using a certain regression to estimate the value of a variable.

Overall F-statistic. The *overall* F-statistic is simply an F-test in which the null hypothesis is that the data can be explained by a constant (which would be the mean of the observed data), versus the hypothesis that the data can be explained by the selected model. Therefore, a high F-statistic and, specially, a low associated p-value indicate that the hypothesis that our model explains the data (vs. the hypothesis that a constant explains them) is clearly correct.

R^2 and adjusted R^2. Typically used to assess how well the values predicted by a model will adjust to the real values, in the case of linear regressions R^2 is simply the square of the correlation coefficient between the observed and the predicted data.

However, in the case of multilinear regression, the R^2 value will always increase as new variables are added to the model. For this reason sometimes the *adjusted* R^2 is used, which corrects for the number of explanatory variables in the model. As a result, the adjusted R^2 value will only increase if the new term improves the regression more than would be expected by chance. The adjusted R^2 value is usually denoted by $\tilde{R}^2$ and computed as follows:

$$\tilde{R}^2 = 1 - (1 - R^2)\frac{n - 1}{n - p} \tag{9}$$

where, again, n is the data size and p the number of terms in the regression. Please recall that in this formulation the intercept is included in p. It is well-known that the higher the R^2 and the adjusted R^2 values, the higher the correlation between the values predicted by the model and the values actually observed.