Tesis Fin de Máster
Máster en Ingeniería de Sistemas e Informática
Curso 2009-2010

Reverse Tone Mapping for Suboptimal Exposure Conditions

Belén Masiá Corcoy

Septiembre de 2010

Director: Diego Gutiérrez Pérez

Departamento de Informática e Ingeniería de Sistemas
Centro Politécnico Superior
Universidad de Zaragoza

Gracias...

...a mis padres y hermana, por aguantarme todo, por su eterno e incondicional apoyo, y por todo lo que he aprendido de y gracias a ellos.

...a Diego, por su ayuda, por su infinita paciencia y sus infinitos ánimos, y por tantísimas cosas que he aprendido de él.

...a José María Martínez Montiel, por darme la oportunidad que me dio y por su contribución a que yo esté en el mundo de la investigación.

...a la gente del lab, porque trabajar con personas así hace que todo sea mucho más fácil.

...a Óscar, porque sin su ayuda yo no estaría ahora escribiendo esto, y por supuesto por los buenos ratos.

...a mis amigos, por las risas cuando estaba de bajón (y cuando no lo estaba :)), por las juergas para desconectar, por las horas escuchándome hablar (y lamentarme) de funciones sigmoidales, linealizaciones o experimentos psicofísicos, y de lo duro que es esto de la investigación, por los innumerables cafés, por las cervezas a las ocho de la tarde, y sobre todo, por estar ahí siempre.

RESUMEN

La mayor parte de las imágenes y videos existentes son de bajo rango dinámico (generalmente denominado LDR por las siglas del término en inglés, *low dynamic range*). Se denominan así porque, al utilizar sólo 8 bits por canal (R,G,B) para almacenarlas, sólo son capaces de reproducir dos órdenes de magnitud en luminancia (mientras que el sistema visual humano puede percibir hasta cinco órdenes de magnitud simultáneamente). En los últimos años hemos asistido al nacimiento y expansión de las tecnologías de alto rango dinámico (HDR por sus siglas en inglés), que utilizan hasta 32 bits/canal, permitiendo representar más fielmente el mundo que nos rodea.

Paulatinamente el HDR se va haciendo más presente en los *pipelines* de adquisición, procesamiento y visualización de imágenes, y como con el advenimiento de cualquier nueva tecnología que sustituye a una anterior, surgen ciertos problemas de compatibilidad. En particular, el presente trabajo se centra en el problema denominado *reverse tone mapping*: dado un monitor de alto rango dinámico, cuál es la forma óptima de visualizar en él todo el material ya existente en bajo rango dinámico (imágenes, vídeos...). Lo que hace un operador de *reverse tone mapping* (rTMO) es tomar la imagen LDR como entrada y ajustar el contraste de forma inteligente para dar una imagen de salida que reproduzca lo más fielmente posible la escena original. Dado que hay información de la escena original que se ha perdido irreversiblemente al tomar la fotografía en LDR, el problema es intrínsecamente *ill-posed* o mal condicionado.

En este trabajo, en primer lugar, se ha realizado una serie de experimentos psicofísicos utilizando un monitor HDR Brightside para evaluar el funcionamiento de los operadores de *reverse tone mapping* existentes. Los resultados obtenidos muestran que los actuales operadores fallan -o no ofrecen resultados convincentes- cuando las imágenes de entrada no están expuestas correctamente. Los rTMO existentes funcionan bien con imágenes bien expuestas o subexpuestas, pero la calidad percibida se degrada sustancialmente con la sobreexposición, hasta el punto de que en algunos casos los sujetos prefieren las imágenes originales en LDR a imágenes que han sido procesadas con rTMOs.

Teniendo esto en cuenta, el segundo paso ha sido diseñar un rTMO para esos casos en los que los algoritmos existentes fallan. Para imágenes de entrada sobreexpuestas, proponemos un rTMO simple basado en una expansión gamma que evita los errores introducidos por otros métodos, así como un método para fijar automáticamente un valor de gamma para cada imagen basado en el *key* de la imagen y en datos empíricos.

En tercer lugar se ha hecho la validación de los resultados, tanto mediante experimentos psicofísicos como utilizando una métrica objetiva de reciente publicación.

Por otro lado, se ha realizado también otra serie de experimentos con el monitor HDR que sugieren que los artefactos espaciales introducidos por los operadores de *reverse tone mapping* son más determinantes de cara a la calidad final percibida por los sujetos que imprecisiones en las intensidades expandidas.

Adicionalmente, como subproyecto menor, se ha explorado la posibilidad de abordar el problema desde un enfoque de más alto nivel, incluyendo información semántica y de saliencia.

La mayor parte de este trabajo ha sido publicada en un artículo publicado en la revista *Transactions on Graphics* (índice JCR 2009 2/93 en la categoría de *Computer Science, Software Engineering*, con un índice de impacto a 5 años de 5.012, el más alto de su categoría). Además, el *Transactions on Graphics* está considerado como la mejor revista en el campo de informática gráfica. Otra publicación que cubre parte de este trabajo ha sido aceptada en el *Congreso Español de Informática Gráfica 2010*. Como medida adicional de la relevancia del trabajo aquí presentado, los dos libros existentes hasta la fecha (hasta donde sabemos) escritos por expertos en el campo de HDR dedican varias páginas a tratar el trabajo aquí expuesto (ver [2, 3]). Esta investigación ha sido realizada en colaboración con Roland Fleming, del *Max Planck Institute for Biological Cybernetics*, y Olga Sorkine, de *New York University*.

Contents

List of Figures

1. Prologue

When I entered this project, a collaboration between Roland Fleming from the Max Planck Institute in Tübingen, Olga Sorkine from New York University, and Diego Gutierrez from Universidad de Zaragoza had already been established, and the whole of the project has been carried out within this collaboration. The aim was to work on the field of reverse tone mapping, and the starting point, the intuition that current algorithms would fail if the image was not correctly exposed. The first steps were taken with a previous publication, by Miguel Martin and the three already cited authors, where image exposure, and its possible correlation with different image statistics, was analyzed [1]. I entered the project at that point, and we had the double objective of: a) formally evaluating the current algorithms, and b) if there was room for improvement, devising a new reverse tone mapping operator.

Two publications have spawned from this work, the initial objectives being widely covered. The first of them, *Evaluation of Reverse Tone Mapping Through Varying Exposure Conditions*, which covers most of the contents of this master thesis, was presented at *SIGGRAPH Asia 2009* in Yokohama, Japan. A total of 70 out of 275 papers were accepted for the conference, the acceptance rate being 25%. Articles accepted for this conference are also published in the journal *Transactions on Graphics*, whose JCR 2009 index is 2/93 in the category of Computer Science, Software Engineering (5-year impact factor 5.012, highest of its category). *Transactions on Graphics* is regarded as the top journal in the field of Computer Graphics. The second publication, *Selective Reverse Tone Mapping*, a smaller sub-project which covers the work described in Chapter 7, has been accepted for publication in *Congreso Español de Informática Gráfica 2010*.

As another representative measure of the impact of the work here presented, the only two technical/scientific books (i.e. in academia, written by researchers in the field; there are many divulgative books on HDR for photographers) which, to our knowledge, have been published on high dynamic range imaging, dedicate several pages to the explanation of our work and findings on reverse tone mapping. These books are *High Dynamic Range Imaging, Second Edition: Acquisition, Display, and Image-Based Lighting* [2] and *Advanced High Dynamic Range Imaging Theory and Practice* [3].

The present master thesis contains essentially the same information as both papers, adding an explanation on the problem and its background and certain data and information which did not appear in the articles, and merging them into one single work.

My interest in what the emergent high dynamic range technologies would bring lead me to begin research in this field. With this master thesis I put an end to my work on reverse tone

mapping, but my interest on image capture, processing and editing has continued to increase, and I will be carrying on with research in the field of computational photography, hopefully in successful collaboration with the Camera Culture Group from the MIT Media Lab (lead by Ramesh Raskar).

2. Introduction

2.1 High Dynamic Range Imaging: background

The dynamic range of a scene is defined as the difference between the highest and the lowest luminance values of the scene. The human visual system (HVS) can cope with lighting conditions that range over up to 10 orders of magnitude in luminance. As an example, a dark night with no moon and the stars as the only light would typically have luminance values of around $10^{-3}cd/m^2$, while sunlight at noon on a clear bright day is around $10^5 cd/m^2$. This is achieved through *adaptation*, that is, the HVS can allocate its dynamic range at will. Adaptation is the reason for well known phenomena like the fact that when we come out of the cinema or out of a dark place into sunlight it takes us a few seconds to see properly. In a certain instant, within the same scene, we can perceive up to five orders of magnitude in luminance. However, our HVS does not adapt equally well to all conditions: in dark scenarios we perceive luminance differences well, and color badly, while the opposite holds for bright scenarios.

Even though we can perceive up to ten orders of magnitude in luminance (five within the same scene), traditional capture and display systems can deal with up to two log-units of luminance, since they use 8 bits per channel. As a consequence, many times traditional images fail to faithfully recreate the real world. An example of this is shown in Figure 2.1. None of the images succeed in reproducing the real scene as a human observer would see it. Through the adjusting of exposure, range can be allocated either to the foreground (*left*) or to the background (*right*), but our HVS would be able to perceive both foreground and background simultaneously.

As a solution to this, Debevec and Malik proposed their multibracketing technique [4]. An image of high dynamic range (HDR) can be obtained by combining several single low dynamic range (LDR) photographs taken with different exposures (see Figure 2.2). The resulting images, which are HDR, use 16 or 32 bits per channel. After this pioneer work, and during the last decade, we have assisted to a great expansion of HDR technologies, not only for capturing images, but also for displaying, storing and processing them.

It is now common believe that HDR technology will be commonplace in the near future. However, as with any other big change in technology, there is a transition time during which both technologies, LDR and HDR, coexist. The first issue that arised was how to display HDR images in the LDR off-the-shelf displays that everyone had at home. This gave origin to the problem of tone mapping, and dozens of algorithms have been developed in the last decade which try to solve this problem in the best possible way. The opposite problem is called reverse tone mapping: once HDR displays become common, what do we do with all legacy material,

Figure 2.1: The problem of traditional LDR imaging. Being able to deal with only two orders of magnitude in luminance, LDR images fail to recreate the real world as we perceive it. Through the adjusting of exposure the dynamic range of the camera is either allocated to the foreground (*left*) or to the background (*right*), but both cannot be faithfully captured at the same time.

our old photographs and videos? How do we display them in HDR monitors? The problem of reverse tone mapping has been addressed by only a few works in the last years (see Chapter 3).

Figure 2.2: The multibracketing technique. Several low dynamic range photographs taken with different exposures are combined to produce a high dynamic range image. Images from *www.cambridgeincolour.com*.

2.2 Reverse Tone Mapping: the problem

As previously stated, high dynamic range display devices are becoming increasingly common [5], yet very large amount of existing low dynamic range legacy content and prevalence of 8-bit photography persist. This presents us with the problem of reverse tone mapping. The aim of reverse tone mapping operators (rTMOs) is to endow low dynamic range imagery with the appearance of a higher dynamic range without introducing objectionable artifacts. Ideally, an rTMO should take a standard LDR image as input and reconstruct as accurately as possible the true luminance values of the original scene. As depicted in Figure 2.3, this is an ill-posed problem. For most scenes and imaging devices, the image data is irreversibly distorted by unknown nonlinearities, sensor noise, lens flare, blooming, and perhaps most importantly, sensor saturation, which clips high intensities to a constant value. Reverse tone mappers must somehow reconstruct the missing data, or boost the contrast in a way that does not cause the clipped

regions to appear visually unpleasant.

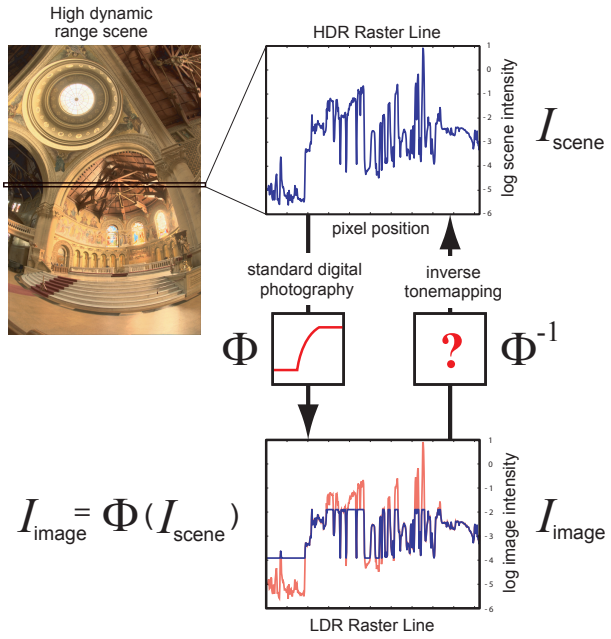


Figure 2.3: The reverse tone mapping problem. Standard imaging loses data by transforming the raw scene intensities I_{scene} through some unknown function Φ, which clips and distorts the original scene values to create the I_{image} , shown in the bottom panel (values clipped from the original are shown in red). The goal of an rTMO is to invert Φ to reconstruct the original scene data, or to convincingly "fake" it.

Existing rTMOs tackle this ill-posed problem in different ways, leading them to succeed and fail in different conditions. For example, some reverse tone mapping strategies may handle small clipped highlights well, but cause large saturated regions to appear unnatural. Conversely, other rTMOs may avoid introducing artifacts in over-exposed conditions, but fail to enhance under-exposed images sufficiently. The key is to understand which strategies produce the best possible visual experience, for which a number of user studies have recently been conducted [6, 7, 8, 9]. These experiments have yielded many valuable insights which may guide future rTMO and even HDR display design. However, they have been applied only to subjectively *correctly exposed* images, usually with knowledge of the dynamic range of the original, real-world scene. A key challenge in rTMO design is how to handle non-optimal LDR content, particularly images that are incorrectly exposed.

2.3 Approach and contributions of this work

Our research is dedicated to finding non-intrusive ways to take advantage of the higher dynamic range of the display medium, irrespective of the dynamic range of the original image. Reverse tone mapping also sheds light on a general problem in signal processing: taking partial, distorted or corrupted data and reconstructing the original as faithfully as possible. Here our quality criterion is perceptual faithfulness rather than physical accuracy.

The vast amount of LDR legacy content spans a large range of exposures. Under- or over-exposure may be due to different reasons, including bad choices by the photographer or pure

artistic intentions. Legacy professional material may have been shot to make the most appropriate use of the dynamic range available at the time, very different from what is currently available. Additionally, the information about the dynamic range of the real scene is typically not recorded. It is therefore crucial to extend previous studies by taking into consideration varying exposure conditions for a set of images without additional information.

We have performed a series of psychophysical studies assessing how rTMOs handle images across a wide range of exposure levels (Chapter 4). We have found that, while existing rTMOs perform sufficiently well for dimmer (under-exposed) images, their performance systematically decreases for brighter (over-exposed) input images. This suggests that there is a need for an rTM method that effectively deals with over-exposed content. We show that simply boosting the dynamic range by means of an adaptive γ curve achieves good results that outperform the current rTMOs, and propose a simple method to obtain a suitable value of γ for each image (Chapter 6).

We additionally observe that artifacts produced by some rTMOs are also visible in low dynamic range renditions of the images (Chapter 5). This is because many artifacts are not simply due to inappropriate intensity levels, but also have a spatial component. We perform a second user study to shed light on which type of inaccuracies introduced by reverse tone mapping most hamper our perception of the final image. This information can further help future rTMO design.

Finally, we explore reverse tone mapping strategies which allow the user to control dynamic range expansion based on her own preferences or intended goal. We present an interactive higher-level approach to reverse tone mapping (Chapter 7). Inspired by the Zone System used in photography, it can also be used as an artistic tool where both the tonal balance and the mood of the final depiction can be adjusted by the user.

3. Previous Work

3.1 Reverse tone mapping

Dynamic range expansion, along with related subsequent problems such as contour artifacts, has been initially addressed by bit-depth extension techniques [10] and decontouring methods [11]. However, these techniques are designed for extension to bit-depths much lower than that of HDR displays. More recently, a few works have looked at the problem of reverse tone mapping for the display of LDR images and videos on HDR displays. The general approach of these reverse tone mapping techniques has been to identify the bright areas within the image, and in particular areas that have been clamped due to sensor saturation, such as light sources. Those areas are typically significantly expanded, while the rest is left unchanged or mildly expanded, to prevent noise amplification. We offer here a brief discussion on reverse tone mapping techniques, and refer the reader to the work by Banterle and colleagues [12] for a comprehensive review on the topic.

Banterle et al. [13, 14] apply the inverse of Reinhard's tone mapping operator [15] to the LDR image and detect areas of high luminance in the resultant HDR image. They then produce a so-called *expand-map* by density estimation of the bright areas, and use this map to interpolate between the LDR image and the initial inverse tone mapped HDR image, thus modulating the expansion range. This framework has been extended to video by designing a temporally-coherent version of the expand-map [16]. The LDR2HDR framework of Rempel et al. [17] is similar in spirit, but their expand-map (which they term *brightness enhancement function*) can be computed in real time using the GPU. The image intensity is first linearized, and a binary mask is computed by thresholding the saturated pixels; the brightness enhancement map is computed as a blurred version of the binary mask, combined with an edge stopping function to retain contrast of prominent edges. The contrast of the LDR image is then scaled according to the enhancement map. Note that the expansion is affected by the size of the bright objects: larger objects may receive more brightness boost. Recently, Kovaleski and Oliveira [18] presented a reverse tone mapping technique which is also based on real-time computation of a brightness enhancement function, but substitutes a bilateral filter for the combination of a Gaussian blur and an edge stopping function used by Rempel et al. [17].

Meylan et al. [19, 20] explicitly focus on specular highlight detection and apply a steep linear tone mapping curve to the presumably clamped areas, whereas the rest of the image is expanded by a mild linear curve. A more sophisticated segmentation and classification of bright areas in the image is done in the work of Didyk and colleagues [21]: they segment the bright image areas and label them as diffuse surfaces, light sources, specular highlights and reflections using

a trained classifier. Different expansion functions are designed for each class to reproduce the dynamic range more accurately (in particular, the luminance of light sources and highlights is expanded more than that of reflections, while bright diffuse surfaces are not expanded). The method is suitable for high-quality video enhancement thanks to the temporal coherence of the segmentation and the expansion function. Finally, Wang et al. [22] propose to fill in the texture information of the clamped bright areas by transferring texture from other (well exposed) areas, although the method may not be viable if a suitable region for transferring detail is not found elsewhere. Both methods [21, 22] rely on user assistance to guide the process, whereas we are interested in more automatic approaches.

3.2 User studies

It is now generally accepted that HDR displays provide a richer visual experience than their LDR counterparts. However, different parameters such as luminance, contrast or spatial resolution influence our visual experience, which makes it difficult to come up with an ideal combination. Additionally, image content probably also affects our preferences. In computer graphics, several researchers have performed a series of user studies, the findings of which may even influence future hardware development.

Yoshida et al. [6] judged subjective preference (without a reference image) and fidelity (by comparing to a real world scene) for a series of tone mapped images. Users could adjust brightness, contrast and saturation for each individual image. Although their work was geared towards the design of a forward tone mapping operator, their conclusions are also useful for rTMO development: they found that, in general, brighter images were preferred over dimmer ones. Interestingly, however, in certain cases users would break this tendency and keep a significant portion of the image dark, reducing overall brightness and giving more importance to contrast.

Influence of luminance, contrast and amplitude resolution of HDR displays, was analyzed by Seetzen et al. [7] to guide future display designs. Their studies show that the preferred luminance and contrast levels are related: for a given contrast, perceived image quality increases with peak luminance, reaches a maximum and then slowly decreases.

Akyüz and colleagues [8] performed a series of psychophysical studies which revealed that a linear range expansion of the LDR image could surpass the appearance of a true HDR image, suggesting that simple solutions may suffice for reverse tone mapping. Recently, Banterle et al. [9] have presented a psychophysical evaluation of existing reverse tone mapping techniques, the results of which indicate that nonlinear contrast enhancement may yield better results overall.

These previous studies provide useful insight into the desirable behavior of tone mapping operators. A key difference with our work is that they were performed on correctly exposed images, whereas we are interested in analyzing reverse tone mapping across varying exposure conditions. In this work, we define over-exposed pixels as those with values ≥ 254, and under-exposed pixels as those with null values [17, 1].

4. Experiment One: rTMO Evaluation

4.1 Introduction

To assess the overall performance of an rTMO, it is important to evaluate it across a range of different imaging conditions. To this end, we have performed a user study in which subjects directly compared the output of three reverse tone mapping schemes (plus standard LDR visualization) across a range of exposures, from clearly under-exposed to clearly over-exposed images. We asked subjects to rate the appearance of the reverse tone mapped images on a calibrated Brightside DR37-P HDR monitor (32.26″ wide and 18.15″ high), with a black level of 0.015 cd/m^2 and a peak luminance of over 3000 cd/m^2. Calibration of the Brightside monitor was performed to confirm linearity and stable performance during the experiment and to enable comparison to specific intensities in cd/m^2 should the need have arisen in the analysis, as per standard practice in psychophysics. Temperature compensation was turned off to avoid changes in intensity (this was possible thanks to the air conditioning in the room). The LDR versions of the images were displayed by approximately matching the contrast to a typical desktop TFT (Dell).

Ambient luminance was kept at about 20 cd/m^2, and the participants were seated approximately one meter away from the monitor. Based on the subjects' ratings, we can infer which rTMOs are most effective at recreating the experience of an HDR scene without visually objectionable side-effects. As opposed to other studies, we do not provide a ground truth HDR image for direct comparison, since it is almost always unavailable in the case of legacy content.

4.2 Stimuli and subjects

The stimuli consist of photographs of nine scenes with different lighting conditions, captured with a Nikon D200 at an original resolution of 3872 by 2592 (down-sampled for visualization purposes on the Brightside monitor, which has a 1920 by 1080 pixel resolution). Each scene was captured with four different exposure times. Five scenes were made up of bright images (from approximately correct exposure to clearly over-exposed), and the remaining four were made up of dark images (from clearly under-exposed to approximately correct). Figure 4.1 shows a representative image of each scene, while Figure 4.2 shows the four exposures for two example scenes. The stimuli (please refer to Appendix A for the complete series of all the scenes) have been obtained from a previous study on exposure perception [1], where the authors analyze basic image data to try to obtain a correlation between image statistics and the perception of under-

Figure 4.1: Representative samples of the stimuli used in our tests. Top: bright images (*Building, Lake, Graffiti, Strawberries, Sunset*), each showing a certain degree of over-exposure. Bottom: dark images (*Car, Flowers, Crayons, Pencils*), with varying degrees of under-exposure.

Figure 4.2: The complete bracketed sequence for the *Building* and *Flowers* scenes.

and over-exposure.

From each exposure in the bracketed sequence, we obtained three candidate renditions for display on the HDR monitor using a representative subset of reverse tone mapping algorithms: LDR2HDR [17], Banterle's operator [13] and linear contrast scaling [8]. Except for the straightforward linear scaling (in Yxy color space, and thus performed on linearized values) we obtained the images from the authors of the original algorithms, in order to ensure accuracy in the implementation. For the LDR2HDR algorithm the parameters used were 150 pixels for the standard deviation of the large Gaussian blur applied to the mask, a brightness amplification factor $\alpha = 4$ and a gradient image baseline width for divided differences of 5 pixels, plus a 9×9-pixel kernel for the antialiasing blur and a 4-pixel radius for the *open* operator used to clean up the final edge stopping function (please refer to the original paper for a detailed explanation of these parameters). In the case of Banterle's operator, when generating the expand-map, the parameters of the density estimation were a radius ranging from 16 to 42 pixels (smaller radius for lower exposures) and a threshold of 1 to 4 light sources (lower threshold for higher exposures), being 2048 the number of generated light sources for Median Cut sampling. In both cases, Banterle's operator and LDR2HDR, images were linearized using gamma correction ($\gamma = 2.2$). We also added a fourth LDR rendition in which the original images are presented within a luminance range matched to a typical desktop TFT monitor. The goal of this fourth image is to study whether the established assumption that visual preference is given to HDR holds over a range of exposures.

A gender-balanced set of twelve subjects with normal or corrected-to-normal acuity and normal color vision were recruited to participate in the experiment. All subjects were unaware of the purpose of the study, and were unfamiliar with HDR imaging.

4.3 Procedure

Participants viewed the stimuli on the Brightside HDR display in a dark room. On each trial, subjects were presented with all four renditions of a given exposure of a given scene in a 2×2 array (a *stimulus quadruple*). The positions of the four renditions within the array were random across trials, and the order of the trials was random with the constraint that consecutive trials did not present the same scene. The subjects' task was to rate the quality of the four renditions on a scale from 1 to 7, according to how accurately the images depicted how the scene would appear to the subject if they were actually present in the scene. Thus the key criterion for comparison was the subjective fidelity of the renditions. Subjects were given unlimited time for each trial and could modify their rating of any of the renditions on a given trial before proceeding to the next trial. Additionally, they were free to assign the same values to all four renditions on a given trial, although they were instructed to try to use as much of the 1-7 scale as possible within the experiment as a whole. To aid them in setting their scale, and to accustom them to the experimental procedure, the subjects were presented with a number of practice trials before the start of the experiment.

4.4 Results and Discussion

Several conclusions can be drawn from this test. First, for our images, there was a clear difference in perceived quality between the bright and the dark series: subjects clearly preferred the reverse tone mapped depictions of darker images over brighter ones. This can be seen in Figure 4.3: not only is the overall mean value significantly higher in the former case, but it is relatively stable across exposure as well. In contrast, for the bright images, there is a general downward trend in ratings across the four exposure levels.

Note that this gradual decrease in performance does not correlate with the subjective perception of quality of the original LDR image: in a previous pilot study, users picked different exposures for each series as the subjective best, not necessarily the same as the objective best (defined as the one with the smallest proportion of under- and over-exposed pixels [8]). The trend instead correlates with the proportion of over-exposed pixels and the mean luminance, which do increase with exposure.

Secondly, we can observe *systematic* differences between the rTMOs. On average, subjects rated the LDR2HDR and the Linear rTMOs best (the difference between the two failed to reach statistical significance), followed by the LDR images, and finally the output of the Banterle's rTMO (see Figure 4.3). Pairwise Wilcoxon rank sum tests (similar to a non-parametric version of the t-test) reveal that these differences were significant to $p < 0.05$, except for LDR2HDR vs. Linear in the bright series and Banterle's operator vs. the LDR depiction in the dark series (see Table 4.1 for the complete results).

It is important to note, however, that this ordering does not hold for all conditions. For instance, the LDR depiction was systematically ranked lower than two of the rTMOs, suggesting that indeed HDR visualization is still preferred over LDR, even for under- and over-exposed images. Surprisingly, though, it ranked higher on average than Banterle's rTMO for bright images. The poor overall performance of Banterle's rTMO with this data set is probably due

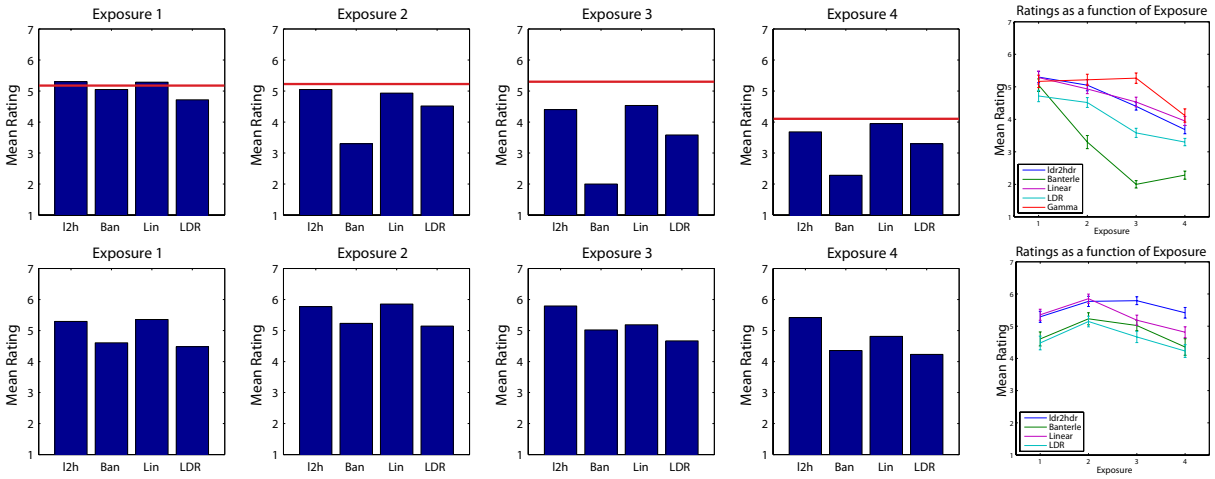


Figure 4.3: Top: bright images series. The blue bars represent the mean ratings across subjects for the four rTMOs (LDR2HDR, Banterle's, Linear and LDR) with increasing exposure levels. The last chart clearly shows the downward trend in perceived image quality. Error bars represent standard errors on the mean. The red line in the first four charts represents the mean ratings for our proposed γ-curve expansion (see Chapter 6). It can be seen that it rates generally higher and is more stable. Bottom: same information for the dark images series, showing higher overall means and a more stable perceived quality across exposures.

(i-j)	$p_b(i,j)$	$p_d(i,j)$
LDR2HDR - Banterle's	2.0532e-21	2.8633e-7
LDR2HDR - Linear	0.5734	0.0283
LDR2HDR - LDR	1.7762e-6	1.4976e-11
Banterle's - Linear	1.1739e-22	0,0013
Banterle's - LDR	4.4489e-11	0.1938
Linear - LDR	1.4697e-7	2.0538e-6

Table 4.1: Results of the Wilcoxon rank sum tests for the bright and dark series (denoted by subindices b and d respectively). Values of $p < 0.05$ are considered to indicate statistically significant differences between rTMOs. Thus, all differences were significant except for LDR2HDR vs. Linear in the bright series and Banterle vs. LDR in the dark series.

to the fact that it often exaggerates the errors in poorly exposed images, resulting in intrusive artifacts. This becomes clear when we measure the extent to which each rTMO yields outlier rating values for each image. We calculate the median rating for each image across rTMOs. We then obtain the *outlier index* as the difference in rating for each rTMO relative to this median value:

$$(Outlier\ index)_{rTMO_i} = (Rating)_{rTMO_i} - (Median\ rating\ across\ rTMOs) \tag{4.1}$$

When an rTMO is *neutral*, simply reflecting the overall quality of the exposure of the image, then the outlier index tends to be close to zero. However, when an rTMO stands out relative to the others (for example due to the introduction of artifacts), then the outlier index tends to deviate from zero. In Figure 4.4, we plot the histogram of the outlier index values for the three rTMOs and the LDR depiction. It is notable that for LDR2HDR, Linear and LDR, the distribution tends to be relatively tightly tuned, while for Banterle's the spread is much broader.

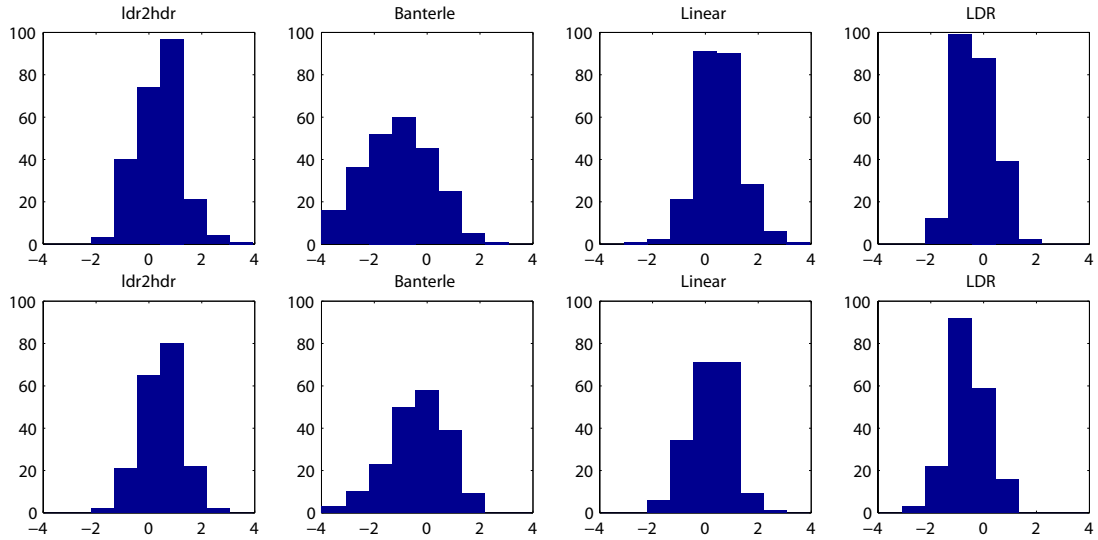


Figure 4.4: Distribution of outlier indices for all four rTMOs. Top: bright series. Bottom: dark series.

This means that on the one hand, when it performs well, it tends to equal or exceed the others. However, it sometimes introduces substantial artifacts that cause the images to look worse than if they were not reverse tone mapped at all.

Although this seems to contradict a recent study where Banterle's operator actually outperformed other rTMOs [9], it is important to note that the experiments carried out in both cases differ significantly: first of all, in the work by Banterle et al. [9] the LDR source images were again well exposed, which is the regime within which Banterle's rTMO performs well, as we also found. However, when the source material is less flattering, we found that the algorithm sometimes produces clearly visible artifacts, which leads to lower ratings. Second, in [9] the authors used a 2AFC paradigm with direct ground truth comparison, whereas we propose a rating approach, which allows users to report their relative subjective preferences. Both tasks are valid ways of assessing fidelity. However, ours has the advantage that it is closer to the real usage scenario: in general the ground truth is unknown and is not presented for comparison.

5. Experiment Two: HDR vs. LDR Monitor

5.1 Introduction

We noticed that artifacts produced by LDR2HDR and Banterle's rTMOs are typically visible in low dynamic range renditions of the images. This is because they generally have a spatial component: they are not simply due to inappropriate intensity levels for certain features, but they also include fringes, visibly boosted noise and other artifacts. To analyze this, we performed a second experiment with seven new subjects, which was identical to the first experiment, except that on each trial, the 2×2 stimulus array was tone mapped using histogram adjustment[1] [23]. The array was then presented on a standard TFT monitor (note that this means that the LDR control condition now appears much darker than on a normal TFT).

5.2 Results and Discussion

In Figure 5.1, we plot the average ratings for each image in the LDR control condition against the average ratings in the HDR condition. As can be seen from the scatter plot, the ratings in the LDR control condition correlated extremely strongly with the ratings in the original experiment on the HDR monitor ($r^2 = 0.9018$). We found no significant difference between bright and dark images.

This result does not imply that the images look *the same* in LDR as in HDR: the subjects were not asked to compare these conditions directly, and previous studies have confirmed that HDR depictions are preferred over LDR [8]. Indeed, none of the subjects saw both renditions. However, it does demonstrate that the pattern of preferences is extremely well conserved. In other words, the images that were less preferred on the HDR monitor were also less preferred when tone mapped back down to LDR. This has two important implications. First, the strong correlation found suggests that a reasonably predictive evaluation of a rTMO could be made without directly testing on an HDR monitor. Second, as noted, the subjective ratings of HDR images that have been generated from LDR images seem to depend more on the presence or absence of disturbing spatial artifacts than on the exact intensities of different features. A similar observation (confirmed by our test) was made by Aydin et al. [24]: they noted that the key issue in image reproduction is to accurately maintain the important features while preserving overall structure, whereas achieving an optical match becomes relatively less important. This becomes

[1] We have used the *pcond* program in *Radiance* to tone map the stimuli.

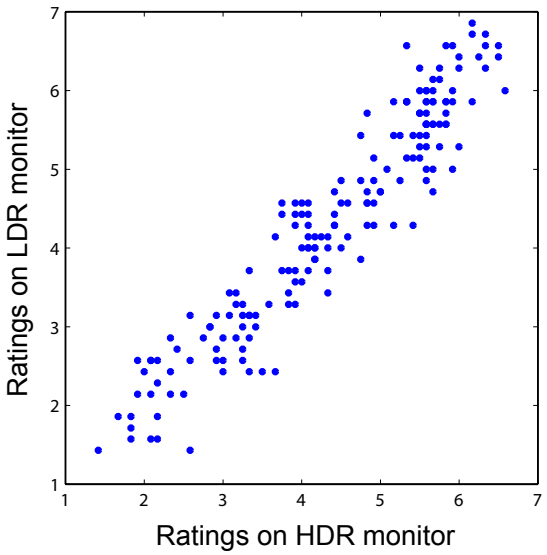


Figure 5.1: Scatter plot showing a strong correlation between ratings on an HDR monitor and ratings when the images were tone mapped back down to LDR and presented on a standard TFT monitor.

even more salient given that the dark-adaptation state of the observer is typically unknown, making absolute intensities meaningless to the user.

The design philosophy that emerges from these considerations is that it is generally better to apply simpler, less-aggressive rTMO schemes if the original image is imperfect. Failing to fully recreate the HDR experience is less disturbing to users than unintended artifacts that can occur when poorly-exposed images are adjusted too aggressively. In the following section we present a simple and robust approach to boosting the dynamic range of over-exposed images, and show that it is less prone to artifacts than other rTMOs.

6. Expanding Over-exposed Content

6.1 Introduction

Our experiments have shown that the danger with computationally sophisticated reverse tone mapping schemes is the potential to make the image appear worse than before processing, through the introduction of objectionable artifacts. However, the goal of a rTMO is to make the image content look better in general and avoid, under any circumstances, making it look worse. Simple global reverse tone mappers, such as linear scaling and gamma boosting, never cause polarity reversals, ringing artifacts or spuriously boost regions well beyond their context. Our first experiment clearly indicates that there is room for improvement in devising an rTMO for bright input images with large saturated areas, whilst darker images turn out much better. We thus focus on the former in this section.

6.2 Gamma expansion

Examining the bright sequence in Figure 4.2 we observe that as exposure increases, more detail is lost as pixel values become saturated, and colors fade to white. It thus seems reasonable to attempt to depict the image in a way that the remaining details become more prominent, as opposed to boosting saturated areas as existing rTMOs do. Note that we do not aim to recover information lost to over-exposure, for which existing hallucination techniques may work [22], but rather to increase perceived quality.

We make the following key observations, which have been confirmed by previous studies on reverse tone mapping: on the one hand, darker HDR depictions are usually preferred for bright input LDR images [19]; on the other hand, in many cases contrast enhancements improve perceived image quality [17]. These suggest expansion of the linearized luminance values following a simple γ curve, which has the desired effect of darkening the overall appearance of the images while increasing contrast. Figure 6.1 shows how the expansion is performed, how the final HDR luminance values relate to the input LDR luminances.

Linearization of the luminance values prior to the dynamic range expansion was done with a gamma curve ($\gamma = 2.2$), following the findings by Rempel et al. [17] which note that simple gamma correction can be used for linearization instead of the inverse of the camera response without producing visible artifacts. To avoid amplifying noise, a bilateral filter [25] can be used prior to expansion [17]. Gamma expansion may potentially boost noise; however, over-exposed

images tend to be significantly less noisy than under-exposed ones. Our psychophysical tests confirmed that noise amplification did not affect the final perceived quality.

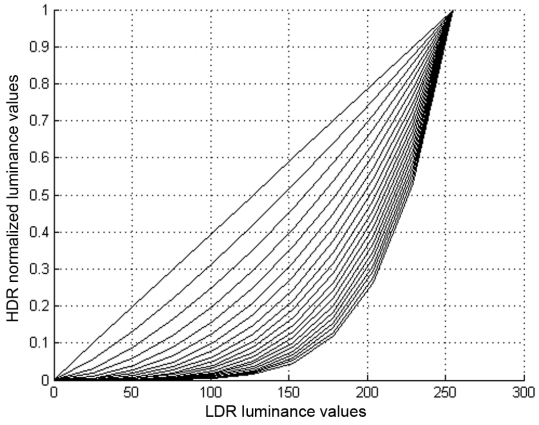


Figure 6.1: Expansion following a gamma curve. The x-axis shows luminance values in low dynamic range, while the y-axis shows normalized high dynamic range luminance values. The curves show different gamma expansion functions, with the value of γ ranging from 1.0 (straight line) to 6.0, at intervals of 0.25.

After the luminance expansion chromaticities are recovered according to the following equations [2]:

$$R_{hdr} = L_{hdr}(\frac{R_{ldr}}{L_{ldr}})^{(1/s)} \tag{6.1}$$

$$G_{hdr} = L_{hdr}(\frac{G_{ldr}}{L_{ldr}})^{(1/s)} \tag{6.2}$$

$$B_{hdr} = L_{hdr}(\frac{B_{ldr}}{L_{ldr}})^{(1/s)} \tag{6.3}$$

where L accounts for luminance, and R, G, B for intensities in each of the channels. s is a factor which accounts for saturation and can be manually adjusted (it ranges between 0 and 1) to obtain the best depiction.

Obviously, the problem with the proposed expansion lies in automatically obtaining an image-dependent suitable γ value, to avoid the cumbersome manual readjustment of the display settings for each individual image to be shown. For this, we first obtain a measure of image brightness, for which we compute its key value; this key acts as an indicator of whether the scene is subjectively dark or light. Since overall brightness can be approximated with log-luminance [26, 15], we estimate the key of an image as [27]:

$$k = \frac{\log L_{avg} - \log L_m}{\log L_M - \log L_m} \tag{6.4}$$

where $\log L_{avg} = (\sum_{x,y} \log(L(x,y) + \delta))/n$. L_m and L_M are the minimum and maximum image luminances respectively, n is the number of pixels and $L(x,y)$ is the pixel luminance. The small

offset δ prevents singularities when $L(x, y) = 0$. We exclude 1% of the highest and lowest pixel values following the suggestion in [27], to make the estimation less sensitive to outliers. We asked users in a pilot study to manually adjust the value of γ in a set of images, and fitted empirical data with a linear regression $\gamma = a \cdot k + b$ (with $a = 10.44$ and $b = -6.282$), which relates γ as a function of the image key ($r^2 = 0.82$). We have used this expression in this work to compute the reverse tone mapped results in this paper. Table 6.1 shows the key and γ values used for all the stimuli.

	1	**2**	**3**	**4**
Building	0.697 / 1.22	0.762 / 1.5	0.816 / 1.75	0.845 / 2.6
Lake	0.7714 / 1.1	0.7453 / 1.2	0.7487 / 1.5	0.7830 / 2.25
Graffiti	0.7666 / 1.2	0.8193 / 1.35	0.8738 / 1.5	0.9184 / 1.75
Strawberries	0.6696 / 1.22	0.7218 / 1.35	0.7218 / 1.55	0.8479 / 1.9
Sunset	0.7022 / 1.1	0.8103 / 1.35	0.8016 / 1.4	0.8713 / 1.75

Table 6.1: Key and γ values for the five scenes and the four exposure levels.

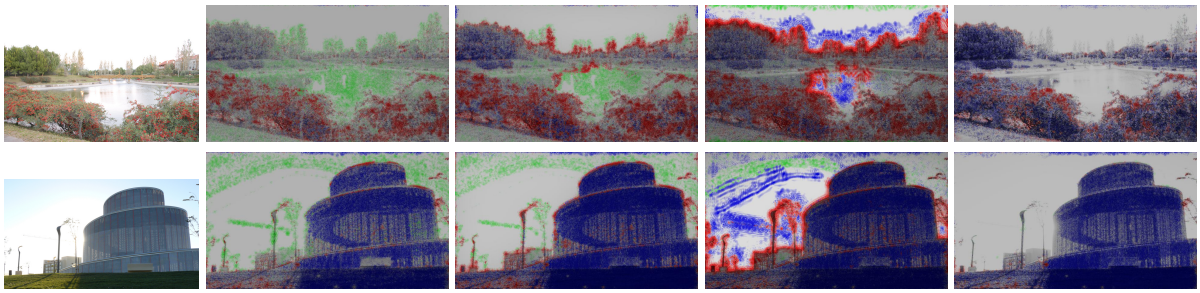
6.3 Validation

To provide a subjective evaluation of the performance of this strategy, we repeated Experiment One (Section 4), substituting the LDR depiction with our γ-expanded versions in order to maintain the 2×2 stimulus array. The red line in Figure 4.3 shows the results.

Experiment One provides useful information about the *subjective* perception of image quality. However, we are also interested in evaluating our approach from an *objective* point of view. The problem is the fact that the intended comparison needs to be performed between an LDR and an HDR image. Recently, Aydin and colleagues [24] have presented a novel image quality metric which identifies visible distortions between two images, independently of their respective dynamic ranges. The metric uses a model of the human visual system, and classifies visible changes between a reference and a test image. The authors identify three types of structural changes: loss of visible contrast (when contrast visible in the reference image becomes invisible in the second one), amplification of invisible contrast (when invisible contrast in the reference image becomes visible in the second one), and reversal of visible contrast (when contrast polarity is reversed in the second image with respect to the reference). It is important to remember that, as Rempel and colleagues noted [17], contrast enhancement tends to increase perceived quality, and therefore is a desired outcome of the rTMO.

Figure 6.2 shows the results of this metric[1] comparing two of the original LDR images (reference images) with the corresponding outputs using linear expansion, LDR2HDR, Banterle's operator and our proposed γ curve. Our method reveals more detail, shows no loss of contrast and minimizes gradient reversals. Note that while our approach may fail to utilize the dynamic range to its full extent in some cases, it has the important and experimentally validated advantage of avoiding objectionable and unpredictable artifacts.

[1]We have used the online implementation provided by the original authors of the paper: http://drim.mpi-inf.mpg.de/generator.php



LDR image | Linear expansion | LDR2HDR | Banterle's operator | our γ curve

Figure 6.2: Objective validation. Comparing the results of several rTMOs with the image quality metric from Aydin et al.[24]. The reference LDR images are *Lake* (top) and *Building* (bottom) as depicted in Figure 4.1 (which correspond to the third and second exposure levels in the series. Please refer to Appendix B for all the exposures in all the scenes). Green, blue and red identify loss of visible contrast, amplification of invisible contrast and contrast reversal respectively. Our γ expansion does not lose any contrast, while minimizing gradient reversals. More importantly, it reveals more detail in the most significant areas of the images (trees, grass, bushes and buildings in the images shown).

7. Selective Reverse Tone Mapping

7.1 Introduction

As we have seen in the previous chapters, to produce a pleasant HDR image from LDR input, existing rTMOs work under the general assumption that highly saturated pixels need to be expanded much more than the rest. As a result, bright image areas representing features like highlights, or the sun in the sky, are largely boosted, thus counter-parting the clamping of information in the LDR image and better representing the real-world experience.

Even though these techniques can produce appealing results for a wide range of LDR content, there are some cases in which the general approach of boosting bright areas may not be the best way to proceed, as demonstrated by our previous experiments (see Chapter 4). These cases include images -such as those shown in Figure 7.1- which contain large saturated areas, either because of artistic purposes or due to a bad exposure.

Figure 7.1: Examples of images containing large saturated areas.

As a response to this, in Chapter 6 we have presented an expansion operator based on a simple gamma curve. In the current chapter we will explore another approach to the problem, an interactive reverse tone mapping technique. We will show that a tailored, higher-level approach to dynamic range expansion can be a good alternative in those cases which are unfavorable for existing rTMOs.

We present two different techniques, one based in Ansel Adams's Zone System [28] (Section 7.2), and another based on detection of salient features (Section 7.3), which allow the user to control dynamic range expansion based on her own preferences or intended goal. The techniques can also be used in combination with each other. This provides a new method for reverse tone mapping and an artistic tool where tonal balance and mood of the final HDR image can be adjusted by the user (in a similar manner to existing tools for LDR or HDR images [29, 30, 31, 32, 33]).

7.2 Using the Zone System for rTM

The so-called Zone System was introduced by Ansel Adams as a guide to produce good photographs with correct tonal values [28]. Exposure is the main factor which determines the way in which the luminance values of the scene are finally mapped to the limited range of values which can be reproduced by the photograph; choosing the right exposure is therefore one of the most important concerns of a photographer. Common exposure meters are designed to aid in this task by measuring luminance values of the scene (or object of interest) and suggesting the lens aperture and shutter speed values. However, irrespective of the scene -its lighting or content-, the values provided by an exposure meter are always such that the object of interest will appear as middle gray in the final image, which in many cases will not be the adequate election. A simple example which illustrates this problem is that of photographing a black and white checkerboard and a scene which is all black except for a white square: the same exposure settings should be used in both of them, yet the reading of an exposure meter would give very different exposure settings for each one. Ansel Adams's Zone System provides a simple way of, using this middle gray reading of exposure meters, choosing the best exposure settings.

Figure 7.2: Division of luminance in zones according to Ansel Adams's System.

This system is not only a tool for photographers still widely in use today [34, 35], but also a formalization of sensitometry principles which provides deep insight into how mapping of tonal values works. Reinhard et al. [15] already rely on it as a basis for their well-known tone mapper, and posterior works on interactive tone management have also built on this system [32]. Following Adams' technique the luminance values in a scene can be divided into ten different luminance zones (0 through IX, see Figure 7.2) according to the equation given by Koren [36]:

$$p = \left(\left(\exp(v \sin\left(\pi \frac{zone - 1}{16}\right) - 1 \right) / \left(\exp(v) - 1\right) \right)^{\psi}, \tag{7.1}$$

where p represents the zone limits in normalized pixel luminances and ψ is the encoding function responsible for non-linearities in the LDR values (usually the inverse of a γ function). The value $v = 5.25$ is set so that zone V on a properly calibrated monitor appears as middle gray [28], defined as 21% of the maximum screen brightness level (this is similar to 18% reflectance referenced to 90% white, which is pure white on good photographic paper). Equation 7.1 is designed so that input values of zero and one map to zones 0 and IX respectively, while the sine function is responsible for the compression required at high pixel levels. By working in XYZ or Yxy space, color information remains unchanged [8].

Once the luminance range of the LDR input image is divided in zones according to Equation 7.1 (see Figure 7.3) the reverse tone mapping process is done by assigning different expansion functions to the different zones. Although in theory these functions could be as complex as desired, we choose to use linear functions for each zone, as they offer a good balance between simplicity and control over the expansion. Thus, the resulting rTM function is piece-wise linear.

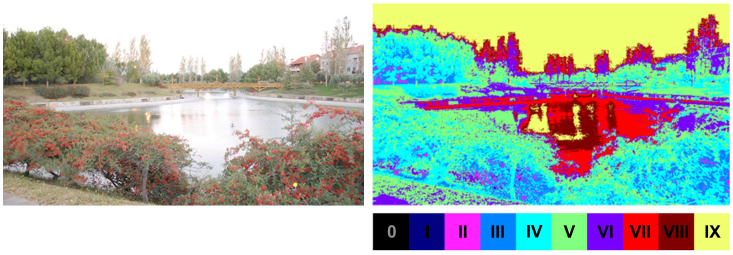


Figure 7.3: Left: Input LDR image. Right: The result of luminance decomposition for zone-based reverse tone mapping.

The darkest and the brightest zones (0 and IX, respectively) of the LDR image are mapped to the lowest and the highest luminance values of the HDR display. A second constraint is that the rTM function must be monotonically increasing, as otherwise gradient reversals may appear that spoil the final depiction. Adjusting the slopes of each of the zones may seem like an involved process; however, in the end it somehow resembles what photographers constantly do, as it translates to assigning ranges of the HDR image luminance to each zone of the LDR input image. Besides, the calculation of the resulting HDR image is almost immediate, thus allowing the user to try different curves before choosing the final one. As an example, Figure 7.4, right shows an HDR image obtained by using a piece-wise linear curve on which only three values were specified: Zone IV being assigned 10% of the HDR image luminance range, Zone VI 40% of that range, and Zone VII 60% of it, which translates to adjusting three points of the LDR–HDR curve shown. We can also appreciate how this simple tuning of the rTM function yields a more appealing depiction than the linear scaling (shown to be on par in subject preference with the HDR image itself by Akyüz and colleagues [8]). Additionally, this zone-based expansion can also be used as part of a bigger rTM framework, as examples in Section 7.4 show.

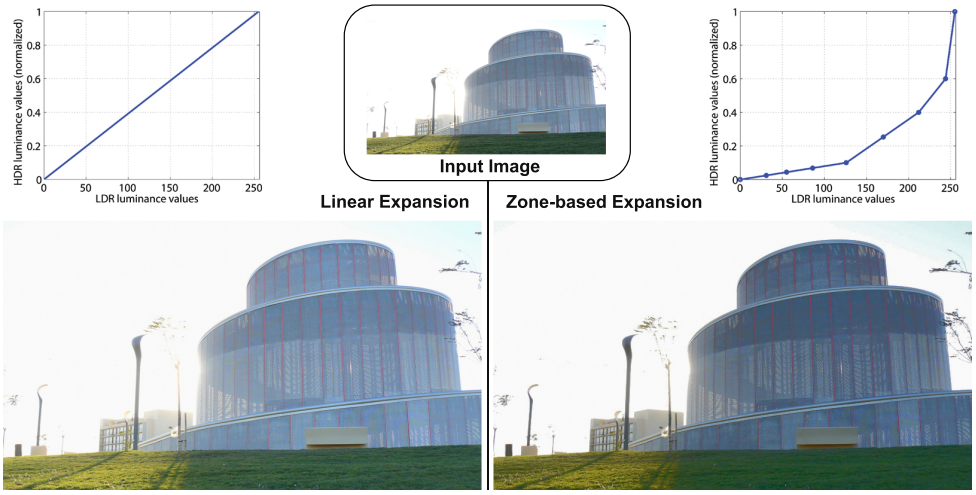


Figure 7.4: Zone-based reverse tone mapping. Left: HDR image obtained by linearly expanding luminance values, and corresponding expansion function. Top center: Original LDR image. Right: HDR image obtained with a piece-wise linear expansion function based on the Zone System, and corresponding graph showing this expansion function.

7.3 Content-aware rTM

As noted before, the general approach in rTM is to allocate most of the additional dynamic range that an HDR display offers to saturated areas in the scene. However, this may not always be the optimal choice. To our knowledge, none of the previous techniques have taken into consideration the *semantics* of the scene. In an image where a large region of it is saturated, such as the leaf in the snow in Figure 7.1, treating in a different way the object of interest (in this case the leaf) and the saturated background (the snow) can lead to more visually appealing results than boosting the saturated area while leaving the leaf nearly untouched. The same reasoning applies to the rest of the images in Figure 1, and in general to images which, either as a result of the artist's choice, or because of wrong exposure, contain large saturated areas. Moreover, when dealing with these type of images, linearly expanding the dynamic range (which in general terms is the other rTM alternative offered by the literature) would result in a significant loss of visible contrast, which is a crucial characteristic of these type of images.

We therefore propose to use a higher-level approach in these cases, taking into account the content of the scene and detecting the object of interest in order to use different reverse tone mapping functions for it and for the background. To separate the region of interest from the background a saliency detector can be used.

7.3.1 Detecting salient features

Saliency detection techniques pursue the objective of detecting those regions where the viewer's attention concentrates when looking at the image. Even though it is an active field where research continues to offer new and improved methods, a series of detectors exist which are able to offer convincing results in a wide variety of images. In general, saliency detection is performed by developing more or less complex models of the human visual system and using them in combination with image metrics to obtain a saliency map, as does the well-known work of Itti et al. [37]. However, for many purposes it is necessary to perform a third stage in which object segmentation is applied to extract salient *objects* instead of just a map of salient locations. In our case the need for this third stage in the saliency detection is obvious, as we look for an accurate separation between the object of interest and the background. From within the saliency detection techniques developed in the last years, we found two of them to meet our needs and applied them to our content-aware reverse tone mapping framework. They are both outlined below, while Appendix D contains a brief description of both methods.

The first of them, learning-based saliency detection [38], formulates the problem of, given an input image I, obtaining a binary saliency mask A, as a Conditional Random Field, in which the probability $P(A|I)$ is inferred using a combination of salient features. Learning using a large training database is used to determine the optimal linear combination of the computed salient features. Figure 7.5 (bottom row) shows an example of the feature functions used and the final saliency mask obtained from them.

The second method, namely Saliency Cuts [39], is essentially a combination of two techniques: the use of graph cuts for object segmentation [40, 41] and the spectral approach to saliency detection of Hou et al. [42]. Graph Cuts is a well-known segmentation algorithm which poses

image segmentation into foreground and background as a minimal graph cuts problem. Seeds for the background and the foreground have to be given as input. The idea behind Saliency Cuts is that the seeds be provided by a saliency map previously obtained with the technique by Hou and colleagues. This way, an accurate segmentation into object of interest and background in the form of a binary saliency map is obtained. This binary saliency map, together with the seeds used for the background and the salient object when performing the segmentation, can be seen in Figure 7.5 (top right).

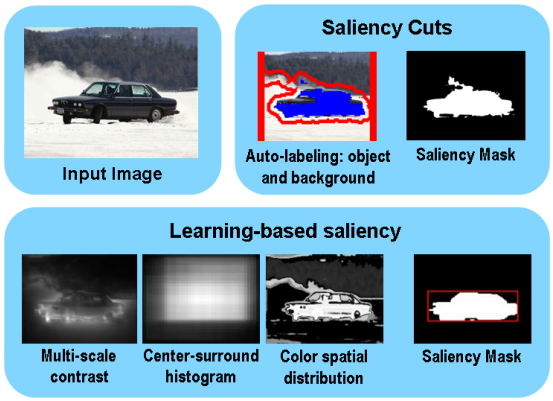


Figure 7.5: Saliency detection with the different methods. Top left: Input image. Top right: Saliency detection using the Saliency Cuts algorithm. Bottom row: Saliency detection using the learning-based saliency detection approach (images from the saliency database publicly available at http://research.microsoft.com/∼jiansun/). Further details can be found in the text.

7.3.2 Expanding the dynamic range

Once the division in object of interest and background has been performed, different expansion functions can be used for each. These expansion functions can be of any type. Given that we are focusing on an interactive approach where the user guides the reverse tone mapping process, we choose again to use piece-wise linear functions after a separation in luminance zones as explained in Section 7.2. Resulting HDR images obtained with this rTM framework and the corresponding saliencies and expansion functions are shown in the Results section.

7.4 Results and Discussion

Figure 7.6 shows an example of a complete pipeline using our rTM approach, combining the two techniques described in the previous sections. The original image is segmented yielding a binary mask containing the salient object, and a division in luminance zones of the input image is performed. Next, the user can adjust the range of luminance in the HDR final image that will be assigned to each zone, both for the seals and for the background independently. This allows the user to easily manipulate the tonal balance of the image to get the best depiction. In this case a non-linear curve (shown in blue in the graph) has been applied to the seals, thus increasing their contrast and making them more salient; the snow has been just linearly expanded. Segmentation has been performed using Saliency Cuts (seeds used for the foreground and background are shown in blue and red, respectively). Even though both of the saliency

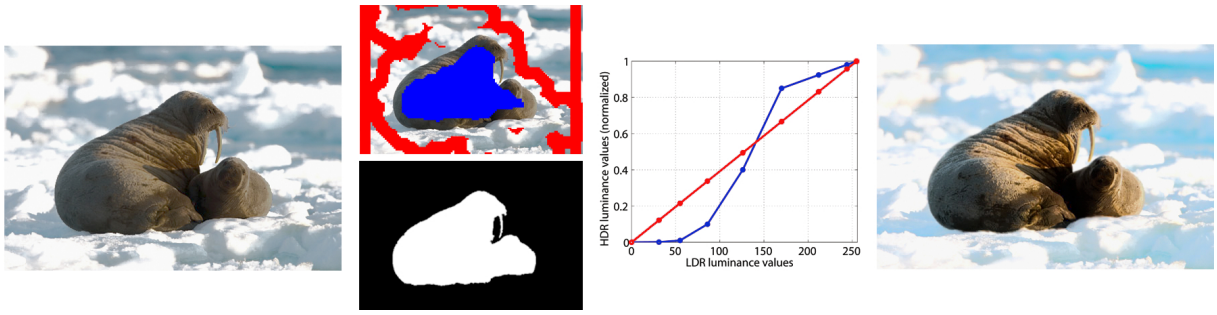


Figure 7.6: Complete pipeline using our interactive rTM approach. From left to right: Input LDR image, auto-labeling of salient object (blue) and background (red) and binary saliency mask, expansion functions for the salient object (blue) and the background (red), and final HDR image. Original image copyright of National Geographic.

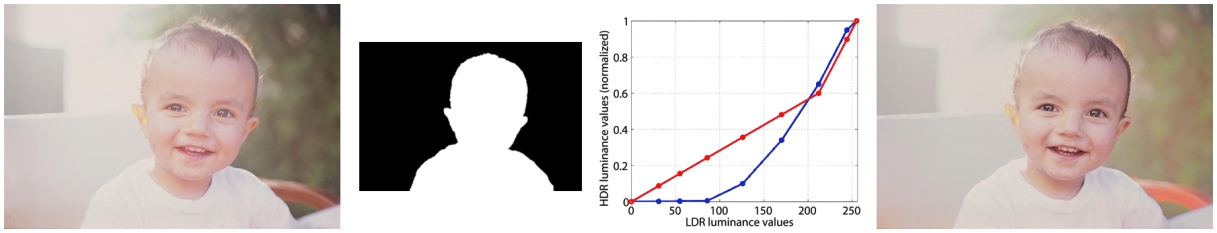


Figure 7.7: Reverse tone mapping using different zone-based expansion functions for the salient object and the background. From left to right: Input LDR image, manually obtained saliency mask, expansion functions for the salient object (blue) and the background (red) and final HDR image. Original image courtesy of Leandro Fessia, all rights reserved.

detection methods presented produce segmentations accurate enough for our purposes given input images which are not excessively complex, for increasing complexity (either morphological or related to luminance values) manual segmentation may be necessary. The presence of more than two salient objects in the image also requires a manual segmentation, as the methods discussed cannot segment more than one object. In the results presented in Figure 7.7 the object of interest was segmented manually and, again, different zone-based piece-wise linear expansion functions were used for the salient object and for the background.

The interactive nature of the approach presented implies that the functions for reverse tone mapping, which determine how the high dynamic range image will look, are adjusted and tuned with low dynamic range renditions of the images as feedback. This is reasonable due to the fact that recent psychophysical experiments have demonstrated that the subjective quality of HDR images that have been generated from LDR images depends more on the presence of absence of spatial artifacts than on the exact luminance values, and thus a reasonably predictive evaluation of an HDR image can be done with an LDR depiction of it [43].

8. Conclusions and future work

8.1 Conclusions

Previous works on the perception of HDR images and rTM design have assumed that the input images were, in general, correctly exposed. While these provide valuable knowledge that could guide the development of both HDR display hardware and reverse tone mapping algorithms, existing LDR legacy content actually covers a wide range of exposures, including material that suffers from bad exposure. As currently designed, existing rTMOs tend to boost over-exposed areas more than the rest of the image. The strategy works well for small areas such as light sources or highlights if the rest of the image is correctly exposed, but no performance evaluation on generally over-exposed imagery had been performed.

Experiment One shows that performance of rTMOs decreases for input images containing a large number of over-exposed pixels, while they seem to perform significantly better for darker images. This suggests that for bright images the consensual approach of boosting bright areas could be improved. We have shown that a simple rTMO based on γ expansion, without the need for explicitly detecting saturated areas, outperforms existing rTMOs in these cases, and propose an empirical expression to automatically find a suitable γ as a function of the image's key, without user interaction. This rTMO has the desired properties of boosting contrast and detail in non-saturated areas of the image, visually compensating for the lack of information in the saturated ones.

We have performed two validation studies, both subjective and objective. The first one has confirmed that our approach increases the perceived image quality for these kind of images. Pairwise Wilcoxon rank sum tests revealed that the differences in rating were statistically significant with respect to all other rTMOs tested. Given that it produces darker overall images with increased contrast, this result is in accordance with previous suggestions [19, 17]. The second evaluation uses a recently published image quality metric which operates with arbitrary dynamic ranges [24]. The metric concludes that our method reveals more detail in non-saturated areas, does not reduce contrast and shows less gradient reversals than the other rTMOs tested. Thus, the artists' original intentions are better preserved.

In both experiments we used typical numbers of subjects for a within-subject design in psychophysics, and the results were highly coherent across subjects. In Experiment One the reported results are statistically significant to the $p < 0.05$ level, meaning that the chances that the outcome of the pairwise comparisons would change after running more subjects from the same population is less than 5%. Indeed, for many of the results, the probability is many orders

of magnitude lower than this, which implies that the qualitative pattern of the results is well conserved across subjects. Likewise, data from Experiment Two exhibit a correlation coefficient of 0.9018, notably conclusive in statistical terms.

Our findings seem to indicate that superior rTMOs should take into account global statistics about the image, and not just individual pixel values. We have derived a simple strategy based on the key value of the images, but more sophisticated strategies could also be devised, possibly including high-level semantics.

We also ran the same expansion on the images from the dark series: as expected, we found no significant improvements over the tested rTMOs, given that our expansion is designed for bright images (Figure C.1 in Appendix C shows the results of this evaluation).

The results from our second experiment confirm that spatial artifacts are more disturbing than inaccuracy in reproduced intensity levels [24]. We found a very strong correlation in the pattern of preferences when viewing images on HDR and LDR displays. This does not mean that the images looked the same, but it does suggest that the artifacts that emerge with poorly-exposed input images are spatial in nature and severe enough that HDR evaluation is not necessary: they can also be clearly seen in LDR.

Our results complement those in the work by Akyüz et al. [8], where the authors show that, for *correctly exposed* imagery, a simple linear expansion works well and suggest that sophisticated treatment of LDR data may not be necessary. In fact, our work is consistent with that of Akyüz et al. [8] in the sense that our proposed γ curves approach linear scaling when the image is approximately correctly exposed.

In a second part of the work, we have presented an interactive approach to reverse tone mapping which can be useful for a wide variety of images, especially those containing large saturated areas. The basis of our method is inspired by photographer Ansel Adams's well-known Zone System, which allows us to divide the luminance range of the image into zones. With the aid of this division in zones, and in an interactive process, a piece-wise linear function to expand the LDR image can be provided by the user. Furthermore, our technique includes the possibility of using higher-level information as a guide for the expansion, segmenting the image in the object of interest and the background and using different expansion functions for each. This interactive approach offers a tool to expand the dynamic range of a scene with significant yet intuitive control over the final result. Besides, being able to freely adjust the luminance ranges of the zones makes it possible to obtain very different HDR depictions of the same input image, potentially providing an artistic tool for photographers and artists in general.

8.2 Future work

The conclusions drawn aim to be valuable for further development of HDR display technology, HDR imaging in general and the development of future LDR expansion algorithms in particular. However, further tests on LDR expansion are desirable. As the community investigates this issue further, this and similar studies will surely be extended and updated. Future reverse tone mapping strategies could involve the design of a contrast-based rTMO, following the findings of the work by Mantiuk et al. [44], which shows promising results in the field of contrast processing

of HDR images, working in visual response space.

Similarly, reverse tone mapping for video content is a key challenge in this field. In order to develop operators that gracefully handle changes in exposure over time, it is crucial to first understand how they fail in the static case, for which we hope this work stimulates future research.

Regarding future work, adding a fitting step of the piece-wise linear rTM functions proposed to smoother ones would be desirable. In the same sense, when dealing with content-aware rTM, taking care of the luminance transitions in the boundary between the objects of interest and the backgrounds would be necessary, either by somehow smoothing the binary mask or by placing constraints to the relationship between both -the object's and the background's- expansion functions. Besides, thorough comparison between the proposed rTM technique and existing reverse tone mapping operators by means of psychophysical experiments would certainly be interesting for the field. Additionally, salient object detection is an open field of research, and our approach would definitely benefit from future advances in this field.

Bibliography

[1] Miguel Martin, Roland Fleming, Olga Sorkine, and Diego Gutierrez. Understanding exposure for reverse tone mapping. In *Congreso Español de Informática Gráfica*, pages 189–198, 2008.

[2] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski. *High Dynamic Range Imaging, Second Edition: Acquisition, Display and Image-Based Lighting.* Morgan Kaufmann Publishers, 2010.

[3] F. Banterle, A. Artusi, K. Debattista, and A. Chalmers. *Advanced High Dynamic Range Imaging Theory and Practice.* A.K. Peters Ltd., 2010.

[4] Paul E. Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. *Computer Graphics*, 31(Annual Conference Series):369–378, 1997.

[5] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs. High dynamic range display systems. *ACM Trans. Graph.*, 23(3):760–768, 2004.

[6] Akiko Yoshida, Rafal Mantiuk, Karol Myszkowski, and Hans-Peter Seidel. Analysis of reproducing real-world appearance on displays of varying dynamic range. *Computer Graphics Forum*, 25(3):415–426, 2006.

[7] H. Seetzen, H. Li, L. Ye, G. Ward, L. Whitehead, and W. Heidrich. Guidelines for contrast, brightness, and amplitude resolution of displays. In *Society for Information Display (SID) Digest*, pages 1229–1233, 2006.

[8] Ahmet Oğuz Akyüz, Roland Fleming, Bernhard E. Riecke, Erik Reinhard, and Heinrich H. Bülthoff. Do HDR displays support LDR content?: a psychophysical evaluation. *ACM Trans. Graph.*, 26(3):38, 2007.

[9] Francesco Banterle, Patrick Ledda, Kurt Debattista, Marina Bloj, Alessandro Artusi, and Alan Chalmers. A psychophysical evaluation of inverse tone mapping techniques. *Computer Graphics Forum*, 28(1):13–25, 2009.

[10] Scott Daly and Xiaofan Feng. Bit-depth extension using spatiotemporal microdither based on models of the equivalent input noise of the visual system. In *Proc. of Color Imaging VIII: Processing, Hardcopy, and Applications*, volume 5008, 455. SPIE, June 2003.

[11] Scott Daly and Xiaofan Feng. Decontouring: prevention and removal of false contour artifacts. In *Proc. of Human Vision and Electronic Imaging IX*, volume 5292, 130. SPIE, June 2004.

[12] Francesco Banterle, Kurt Debattista, Alessandro Artusi, Sumanta Pattanaik, Karol Myszkowski, Patrick Ledda, Marina Bloj, and Alan Chalmers. High dynamic range imaging and LDR expansion for generating HDR content. Annex Eurographics 2009, April 2009.

[13] Francesco Banterle, Patrick Ledda, Kurt Debattista, and Alan Chalmers. Inverse tone mapping. In *GRAPHITE '06*, pages 349–356, New York, NY, USA, 2006. ACM.

[14] Francesco Banterle, Patrick Ledda, Kurt Debattista, Alan Chalmers, and Marina Bloj. A framework for inverse tone mapping. *Vis. Comput.*, 23(7):467–478, 2007.

[15] Erik Reinhard, Michael Stark, Peter Shirley, and James Ferwerda. Photographic tone reproduction for digital images. *ACM Trans. Graph.*, 21(3):267–276, 2002.

[16] Francesco Banterle, Patrick Ledda, Kurt Debattista, and Alan Chalmers. Expanding low dynamic range videos for high dynamic range applications. In *Proc. of the Spring Conference on Computer Graphics*, New York, NY, USA, 2008. ACM.

[17] Allan G. Rempel, Matthew Trentacoste, Helge Seetzen, H. David Young, Wolfgang Heidrich, Lorne Whitehead, and Greg Ward. Ldr2Hdr: on-the-fly reverse tone mapping of legacy video and photographs. *ACM Trans. Graph.*, 26(3):39, 2007.

[18] Rafael Pacheco Kovaleski and Manuel M. Oliveira. High-quality brightness enhancement functions for real-time reverse tone mapping. *The Visual Computer*, 25(5-7):539–547, April 2009.

[19] Laurence Meylan, Scott Daly, and Sabine Süsstrunk. The reproduction of specular highlights on high dynamic range displays. In *IS&T/SID 14th Color Imaging Conference*, 2006.

[20] Laurence Meylan, Scott Daly, and Sabine Süsstrunk. Tone mapping for high dynamic range displays. In *Proc. IS&T/SPIE Electronic Imaging: Human Vision and Electronic Imaging XII*, volume 6492, 2007.

[21] Piotr Didyk, Rafal Mantiuk, Matthias Hein, and Hans-Peter Seidel. Enhancement of bright video features for HDR displays. *Computer Graphics Forum*, 27(4):1265–1274, 2008.

[22] Lvdi Wang, Li-Yi Wei, Kun Zhou, Baining Guo, and Heung-Yeung Shum. High dynamic range image hallucination. In *Eurographics Symposium on Rendering*, pages 321–326, 2007.

[23] G. Ward, H. Rushmeier, and C. Piatko. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Trans. on Visualization and Computer Graphics*, 3(4):291–306, 1997.

[24] Tunç Ozan Aydin, Rafal Mantiuk, Karol Myszkowski, and Hans-Peter Seidel. Dynamic range independent image quality assessment. *ACM Trans. Graph*, 27(3):69, 2008.

[25] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *ICCV*, pages 839–846, 1998.

[26] Jack Tumblin and Holly E. Rushmeier. Tone reproduction for realistic images. *IEEE Computer Graphics and Applications*, 13(6):42–48, 1993.

[27] Ahmet Oğuz Akyüz and Erik Reinhard. Color appearance in high dynamic range imaging. *SPIE Journal of Electronic Imaging*, 15(3), 2006.

[28] Ansel Adams. *The Print*. The Ansel Adams Photography series. Little, Brown and Company, 1983.

[29] Peter J. Burt and Edward H. Adelson. The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31(4):532–540, 1983.

[30] Anat Levin, Dani Lischinski, and Yair Weiss. Colorization using optimization. *ACM Trans. Graph.*, 23(3):689–694, August 2004.

[31] Soonmin Bae, Sylvain Paris, and Frédo Durand. Two-scale tone management for photographic look. *ACM Trans. Graph.*, 25(3):637–645, 2006.

[32] Dani Lischinski, Zeev Farbman, Matt Uyttendaele, and Richard Szeliski. Interactive local adjustment of tonal values. *ACM Trans. Graph.*, 25(3):646–653, 2006.

[33] Zeev Farbman, Raanan Fattal, Dani Lischinski, and Richard Szeliski. Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Trans. Graph.*, 27(3):67, 2008.

[34] Chris Johnson. *The practical zone system*. Focal Press, 3rd edition, 1999.

[35] Carson Graves. *The zone system for 35mm photographers*. Focal Press, second edition, 1997.

[36] Norman Koren. A simplified zone system. www.normankoren.com/zonesystem.html.

[37] L. Itty, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(11):1254–1259, 1998.

[38] T. Liu, J. Sun, N.-N. Zheng, X. Tang, and H.-Y. Shum. Learning to detect a salient object. In *IEEE Computer Vision and Pattern Recognition*, 2007.

[39] Yu Fu, Jian Cheng, Zhenglong Li, and Hanqing Lu. Saliency cuts: An automatic approach to object segmentation. In *International Conference on Pattern Recognition (ICPR)*, 2008.

[40] Y. Y. Boykov and M. P. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In *ICCV*, pages I: 105–112, 2001.

[41] Carsten Rother, Vladimir Kolmogorov, and Andrew Blake. "grabcut": interactive foreground extraction using iterated graph cuts. *ACM Trans. Graph*, 23(3):309–314, 2004.

[42] X. D. Hou and L. Q. Zhang. Saliency detection: A spectral residual approach. In *Computer Vision and Pattern Recognition*, 2007.

[43] Belen Masia, Sandra Agustin, Roland W. Fleming, Olga Sorkine, and Diego Gutierrez. Evaluation of reverse tone mapping through varying exposure conditions. *ACM Transactions on Graphics (Proc. of SIGGRAPH Asia)*, 28(5):160:1–160:8, 2009.

[44] Rafal Mantiuk, Karol Myszkowski, and Hans-Peter Seidel. A perceptual framework for contrast processing of high dynamic range images. *ACM Trans. Graph.*, 3(3):286–308, 2006.

[45] Y. Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1124–1137, 2004.