CEIG'08, Barcelona, Sept. 3–5 (2008)
L. Matey and J. C. Torres (Editors)

Understanding exposure for reverse tone mapping

Miguel Martin[1], Roland Fleming[2], Olga Sorkine[3] and Diego Gutierrez[1]

[1]Universidad de Zaragoza, Spain
[2]Max Planck Institute for Biological Cybernetics, Germany
[3]New York University, USA

Abstract

High dynamic range (HDR) displays are capable of providing a rich visual experience by boosting both luminance and contrast beyond what conventional displays can offer. We envision that HDR capture and display hardware will soon reach the mass market and become mainstream in most fields, from entertainment to scientific visualization. This will necessarily lead to an extensive redesign of the imaging pipeline. However, a vast amount of legacy content is available, captured and stored using the traditional, low dynamic range (LDR) pipeline. The immediate question that arises is: will our current LDR digital material be properly visualized on an HDR display? The answer to this question involves the process known as reverse tone mapping (the expansion of luminance and contrast to match those of the HDR display) for which no definite solution exists.

This paper studies the specific problem of reverse tone mapping for imperfect *legacy still images, where some regions are under- or overexposed. First, we show the results of a psychophysical study compared with first-order image statistics, in an attempt to gain some understanding in what makes an image be perceived as incorrectly exposed; second, we propose a methodology to evaluate existing reverse tone mapping algorithms in the case of imperfect legacy content.*

Categories and Subject Descriptors (according to ACM CCS): I.4.0 [Image Processing and Computer Vision]: General–Image Displays I.3.3 [Computer Graphics]: Picture/Image Generation H.1.2 [Models and Principles]: User/Machine Systems Human factors–Human Information Processing

1. Introduction

High dynamic range imagery allows a broad range of physically-accurate photometric values to be stored per pixel, mimicking the ranges that can perceived by the human visual system [RWPD05]. The well-known process of tone mapping [DCWP02] deals with the problem of strong contrast reduction of the stored HDR radiance values to fit the low dynamic range of traditional display technology, typically trying to preserve image details and/or color appearance.

The problem of tone mapping is expected to progressively fade away when HDR displays reach the mass market [SHS*04]. However, during the logical transition period, there will be a need to display conventional low dynamic range (LDR) imagery on HDR displays. Although this need may decline over time (once HDR capture becomes mainstream), 8-bit photography will most likely still be used for a long time. This means that display algorithms will have to scale up luminance and contrast, instead of compressing them. This brings about the problem of *reverse tone mapping*†, to which currently no definite solution exists. Recently, Seetzen et al. [SLY*06] and Yoshida et al. [YMMS06] showed that the subjective perception of image quality increases when both brightness and contrast are increased simultaneously. Besides, Čadík and colleagues [ČWNA06] also suggest that the global appearance of an image seems to depend much more on brightness and contrast than other attributes, as shown in their OIQ (overall image quality) equation. This result indicates that merely emulat-

† Some authors [BLDC06, AFR*07] refer to the process as *inverse* tone mapping, while others [RTS*07] use the term *reverse* instead. Given that the field is still in its infancy, a fixed nomenclature has not been chosen yet. We opt to use *reverse* since the term *inverse* can also refer specifically to mathematically inverting a tone mapping operator, not to the whole process.

ing LDR characteristics on an HDR display is probably not the best option, as suggested in [RTS*07].

Very few works exist that deal with the problem of reverse tone mapping. Banterle and colleagues [BLDC06, BLD*07] propose a method by first inverting Reinhard's tone mapping operator [RSSF02]. The authors then find areas of high luminance and apply density estimation techniques to produce an *expand-map*, which guides the range expansion of the images. In the work by Meylan et al. [MDS06] the user first selects which pixels in the image can be considered highlights and then two different linear scaling functions are applied according to this classification. Rempel et al. [RTS*07] present a real-time reverse tone mapper operator (rTMO) based on a linearization of pixel values and contrast scaling, followed by a brightness enhancement function similar in spirit to the expand-map. In a series of psychophysical tests, Akyüz and co-workers [AFR*07] come up with a surprising conclusion: LDR data might not require sophisticated treatment prior to its visualization on an HDR display. By merely linearly scaling the range of the LDR input image to fit the range of the HDR display the results are considered as good as (or better than) an original HDR image. Unfortunately, they base their tests solely on correctly exposed images, and the outcome is unclear if that assumption is broken. In fact, while some of the above works present solutions to minimize noise expansion [BLDC06, RTS*07], none deal specifically with the problem of bad exposure in *imperfect*, legacy content, where the image is either under- or overexposed. Highlights in [MDS06] are in fact defined as overexposed pixels above a certain threshold value; however, the method seems to work better if these are localized to small regions of the image. It is unclear whether the algorithm would provide a pleasant solution by boosting large areas (such as an overexposed sky) the way it boosts small highlights.

We need a method to deal with imperfect content as well, but how to expand its dynamic range is not obvious. Clearly, under- and overexposure effects have been consciously used for decades, and have become standard artistic expressions, not just the result of a faulty capture process (Figure 1). Common dodge and burn techniques, for instance, are usually employed to apply local adjustments to aid tonemapping; however, they can be used for exactly the opposite reasons, to actually simulate the effects of incorrect exposure. In other words, sometimes what we call bad exposure is a deliberate decision based on artistic and aesthetic issues, and then we are facing the additional problem of carrying over the *mood* to an HDR display when reverse tone mapping is applied.

This paper aims at shedding some light onto reverse tone mapping for imperfect digital photography. We first show the results of a psychophysical test, where the subjects were presented a series of images with increasing exposures within each image set, and were asked to tag each individual image (exposure) as underexposed, correctly exposed or overexposed merely by visual inspection. We analyze the results comparing with four luminance statistics in the image: histogram, mean, median and percentage of under- and overexposed pixels. We then propose a methodology to evaluate four existing reverse tone mapping algorithms for incorrectly exposed content, also based on psychophysics. To our knowledge, this is the first time that such study is performed, and the reasons to do it are twofold: on the one hand, the fact that, as argued, a lot of the current digital content is *not* properly exposed (and complete backward compatibility is a must for HDR displays to succeed). On the other hand, before a working reverse tone mapping algorithm can be developed, it is necessary to understand all the aspects of the problem, both technical and psychophysical.

Figure 1: *Using exposure as artistic expression (Jill, by Joseph Szymanski)*

The rest of the paper is organized as follows: the next section introduces the concepts of under- and overexposure, and justifies the psychophysical approach to the following tests. In Section 3 we present the stimuli, methodology and results for our test on the perception of exposure. Section 4 explains the proposed methodology to evaluate four existing reverse tone mapping algorithms. Finally, Section 5 presents conclusions and future work.

2. Under- and overexposure

Exposure in photography can be defined as the total amount of light allowed to fall on the photographic medium during the process of taking a photograph [Kel06]. Under- or overexposure can then be loosely defined as having allowed too little or too much light. But according to what? Let us imagine the following "text-book" example: a scene made up of a green landscape, a red car and a man driving it. If the photographer wants the red car to have correct exposure then he has to measure the light reflecting off of it and sub-expose the photometer reading between one and two stops. However, if he wants the (pale) driver to be correctly exposed, he will have to over-expose one and a half stops, and if he wants the grass to be correctly-exposed he will use the exact measuring of the photometer. So, even if the camera were able to interpret such high-level components of the scene as the green landscape, the red car and the pale driver, it still could not guess the intention of the photographer.

If the images' exposure correctness could be objectively assessed using only image data (with no human interpretation), the digital cameras' firmware could in theory automatically obtain the proper exposure for every scene. Whilst most consumer cameras do offer an estimation that works well for a sufficiently large number of cases, sometimes skilled human intervention is necessary, especially at professional levels.

We thus argue that high-level semantics and human interpretation of the image are necessary in the process of determining whether an image is under- or overexposed. This is further backed by the experiments performed by Akyüz and colleagues [AFR*07]. The authors use LDR bracketed sequence as proposed in [DM97] to create the HDR images. The participants were asked to determine which single exposure was the best among the exposures used. Their results (not included in the paper, but available in [Aky]) show that participants do not always choose the image with the fewest under- or overexposed number of pixels, nor simply the middle exposure of the bracketed sequence. A high-level (and probably individual) interpretation of the scene seems to take place in the decision-making process. The design of our psychophysical tests is in part motivated by these findings.

3. Psychophysical test: exposure perception

As we have argued, under- and overexposure have apparently not yet been defined in objective terms[‡]. This suggests that there is no correlation between apparent correct exposure and objective image data, such as luminance histogram, mean, median or percentage of under- or overexposed pixels (see Figure2), which holds for a sufficiently large number of images. It would be possible in theory to detect a subset of cases, for instance when the histogram shows null values above or below certain thresholds. But even then, false detections would happen, as in the case of low-contrast images with uniformly lit surfaces. For some applications, a useful approach may be to define a threshold under which pixels will be considered underexposed, and a second one over which overexposure is defined (which is how Meylan and colleagues define highlights in [MDS06]). However, these are operations performed at pixel level, and provide no information about the aspect of the image as a whole. More complicated cases include the possibility of an image being under- and overexposed at the same time in different areas (see Figure 3).

[‡] This has been confirmed by interviews with professional photographers and cinematographers

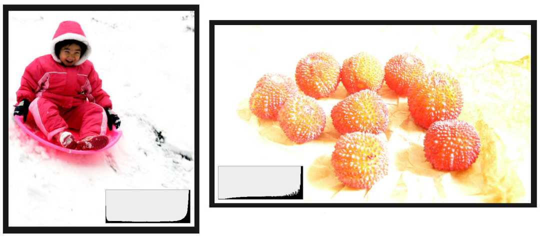

Figure 2: *Two different photographs with very similar luminance histogram, mean, median and percentage of saturated pixels. However, taking into account high-level semantics, the photograph on the left can be considered correctly exposed, while the one on the right is clearly overexposed.*

Figure 3: *Under- and overexposure in the same photograph. Not enough light reaches the corner of the wall, while there is too much light in the window area.*

It thus seems that to properly classify an image as under- or overexposed we need to rely on context-dependent, high-level image semantics, as suggested in previous studies [AFR*07, Aky] and shown in Figure 2. We put this assumption to the test, by comparing the subjective perception of exposure in images with first-order image statistics: luminance histogram, mean, median and percentage of under- and overexposed pixels (defined with reference to a certain threshold). Gaining insight on this matter seems crucial for the problem of reverse tone mapping for imperfect legacy content.

3.1. Stimuli

We use images taken from 10 different scenes. The stimuli images were captured with a Nikon D200 at a resolution of 3872 by 2592 and then down-sampled to 1920 by 1080 for visualization purposes. The scenes were chosen to cover a broad range of lighting conditions and environment types. We shot a bracketed series of five exposures for each scene, ranging from clearly underexposed (labeled as 1 in the paper) to highly overexposed (labeled as 5), giving a total of 50 images used. For each scene, a tone-mapped sample is shown in Figure 4 for visualization purposes .

Luminances values for the experimental stimuli are obtained from their (R,G,B) pixel values according to $L = 0.213R + 0.715G + 0.072B$, as proposed in [RWPD05]. Figure 5 shows the complete bracketed sequence for the *sunset* scene, along with the respective histograms. Tables 1 and 2 present the luminance mean and median respectively; Table 3 shows the percentage of pixels above a given luminance threshold of 254. This value is chosen since it has been found to work well discriminating overexposed areas in photographs [RTS*07]. Finally, Table 4 shows the percentage of pixels with null luminance value, which represents our underexposed pixel threshold.

Sequence / Exposure	1	2	3	4	5
Building	126.66	158.96	183.77	211.39	232.45
Car	14.72	23.79	38.41	58.58	85.06
Indoor flower	19.67	30.29	50.28	60.41	60.51
Lake	70.77	86.47	138.04	170.17	201.54
Pencils	26.68	43.37	68.65	104.5	138.95
Computers	14.93	20.55	24.05	34.98	52.51
Waxes	29.48	50.41	78.21	111.85	148.84
Sunset	91.86	108.56	139.87	194.62	229.99
Graffitti	127.06	169.69	202.56	226.57	242.94
Strawberries	90.39	130.18	168.60	199.79	223.45

Table 1: *Pixel-luminance mean for the bracketed sequence of each scene.*

tinySequence / Exposure	1	2	3	4	5
Building	85	137	189	240	255
Car	6	14	27	50	86
Indoor flower	7	15	33	45	45
Lake	43	61	119	164	210
Pencils	9	18	39	83	136
Computers	0	3	5	11	26
Waxes	23	46	80	120	167
Sunset	75	97	136	218	255
Graffitti	133	189	233	255	255
Strawberries	90	140	198	246	254

Table 2: *Pixel-luminance median for the bracketed sequence of each scene.*

3.2. Experimental design

The design of the psychophysical experiment follows the scheme sometimes referred to as the *method of constant stimuli* [DBW08]: the fifty images are shown one by one, in random order, thus mixing both exposures and scenes. The

Sequence / Exposure	1	2	3	4	5
Building	4.60	19.97	40.03	44.33	47.78
Car	0.02	0.03	0.21	0.60	1.90
Indoor flower	0.64	0.79	1.30	1.95	1.98
Lake	0	0	18.71	23.93	34.49
Computers	0.14	0.54	0.90	2.32	7.36
Waxes	0	0	0	0.01	2.26
Sunset	0.01	4.45	8.35	26.16	51.72
Pencils	0	0	0	0	1.94
Graffitti	0.01	1.00	20.19	49.20	61.46
Strawberries	0	0.01	5.97	22.16	38.39

Table 3: *Percentage of pixels with luminance values 254 and 255.*

Sequence / Exposure	1	2	3	4	5
Building	0	0	0	0	0
Car	30.66	19.65	9.23	4.08	1.16
Indoor flower	26.11	17.10	5.99	3.62	3.50
Lake	0.10	0	0	0	0
Computers	58.06	23.53	14.11	2.70	0.12
Waxes	13.13	4.56	0.45	0	0
Sunset	0	0	0	0	0
Pencils	12.51	8.38	5.04	1.02	0.06
Graffitti	0	0	0	0	0
Strawberries	0	0	0	0	0

Table 4: *Percentage of pixels with null luminance values.*

participants are requested to classify each image in one of these groups: (1) underexposed, (2) correct, (3) overexposed. There is no fixed time for every image to be shown. The participant can move forward (to the next photograph) whenever they are done judging the current image. To ensure the validity of the data, a brief learning task is performed prior to the real test as suggested by [Ken75]: the participants are invited to judge a few images before they start classifying until they feel confident and understand the concepts. These previous images come from extra scenes and are not part of the test itself. The display used was a 24-inch FP241VW model from BenQ. The experiment was set up in a darkened room in order not to reduce the perceived contrast ratio of the display (measured at 60:1). Ambient luminance measured from the wall was 26 cd/m_2.

A gender-balanced group of 24 participants took part in the experiment. Half of them had some photographic skills, whilst all reported normal or corrected-to-normal vision. They sat at a viewing distance of approximately a half meter from the display.

3.3. Significance of the results

Visual inspection of the results of the test (Figure 6) shows the expected logical diagonal distribution of perceived exposure. Strong backlighting of the main objects in some scenes has been mostly interpreted as under- (*indoor, car*) or overexposure (*building, sunset*), although it could be that the photographer's intention was to achieve that effect. This again indicates the need for high-level semantics and possibly human intervention when judging exposure. Some kind of machine learning or classification method, such as Support Vector Machines [Vap95] would be interesting to optimally separate images perceived as under- or overexposed, or even correctly or incorrectly (both under and over) exposed. Four of the five images with strongest gradients (the four previously mentioned plus *computers*) obtained the least number of "correct exposure" votes, suggesting that second-order statistics could provide additional insight into this topic. As expected, the histogram by itself does not provide enough information about an image's exposure.

To analyze correlations in the data, we rely on the Pearson correlation coefficient $\rho_{X,Y}$, defined as:

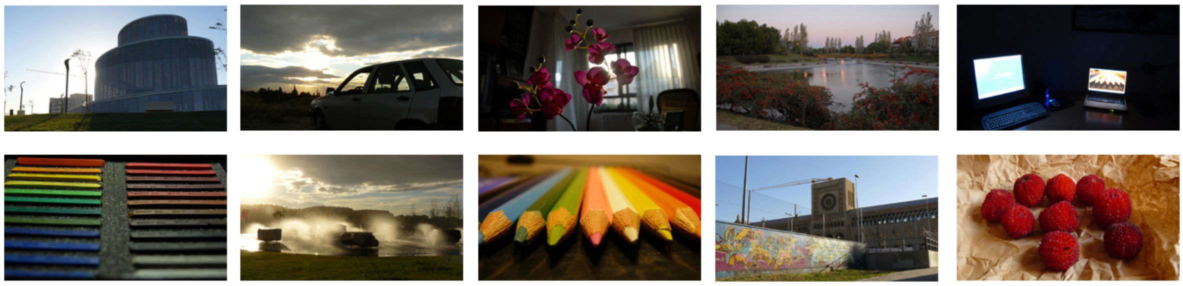

Figure 4: *Tone-mapped samples of each stimuli scene.*

Figure 5: *The complete bracketed sequence for the* sunset *scene, along with the respective histograms.*

$$\rho_{X,Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)}\,\sqrt{E(Y^2) - E^2(Y)}} \qquad (1)$$

where E is the expected value operator, X is the results of the phychophysics evaluation and Y represents the objective parameter being under study (mean, media or the percentage of under- or overexposed pixels). For luminance mean and overexposure, this Pearson coefficient is $\rho^o_m = 0.869$. This is a relatively high value for psychological research, according to Cohen [Coh88]. Similar correlation exists for the luminance median and overexposure ($\rho^o_{md} = 0.846$). This correlation is logically negative for perceived underexposure but, maybe surprisingly, not so strong ($\rho^u_m = -0.726$ and $\rho^u_{md} = -0.691$).

A similar behavior can be observed for the percentage of badly exposed pixels. There is a strong positive correlation between perceived overexposure and saturated pixels ($\rho^o_p = 0.890$) but it becomes lower again for perceived underexposure and pixels with null values ($\rho^u_p = 0.675$). Although this is nothing but mere speculation at this point, these results may suggest some correlation between perceived exposure and the well-known asymmetry of the human visual system under photopic and scotopic conditions [Liv02]. We believe this is an interesting result which we plan to investigate further. Figure 7 shows these results for the case of mean and underexposure. Figure 8 shows the relation between perceived overexposure and the percentage of overexposed pixels. These two cases represent the most-correlated cases for under- and overexposure respectively. Finally, it could be thought that perceived correct exposure may be related to the low occurrence of badly exposed pixels in the image. We found evidence of this, as indicated by its low correlation coefficient ($\rho^c_{sum} = -0.676$).

In conclusion, the two key ideas learned from this experiment, at least for the images shown and the statistics analyzed, are:

- The results seem to confirm the hypothesis that high-level semantics are needed for a proper classification of exposure. This is interesting since it apparently clashes with the notion that visual appeal is based on low-level attributes of an image [AFR*07].
- We found an asymmetry in under- and overexposure perception which may be deeply rooted in the behavior of our visual system. To confirm this, more research needs to be conducted.

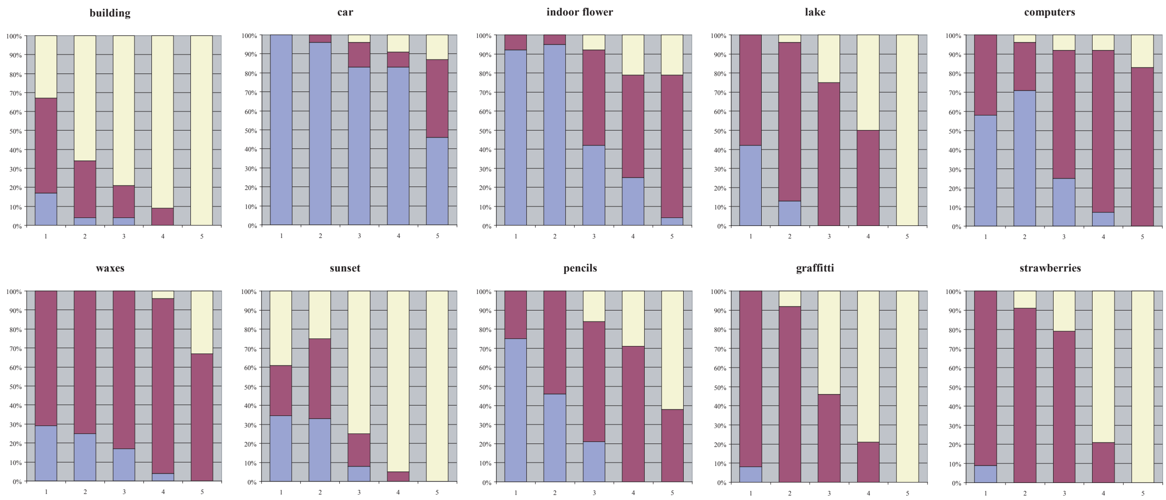

Figure 6: *Results of psychophysics test: participants' stimuli taxonomy. X-axis represents the five exposures for each scene; Y-axis represents the percentage of agreement in classification (blue for underexposure, yellow for overexposure and red for correct exposure).*

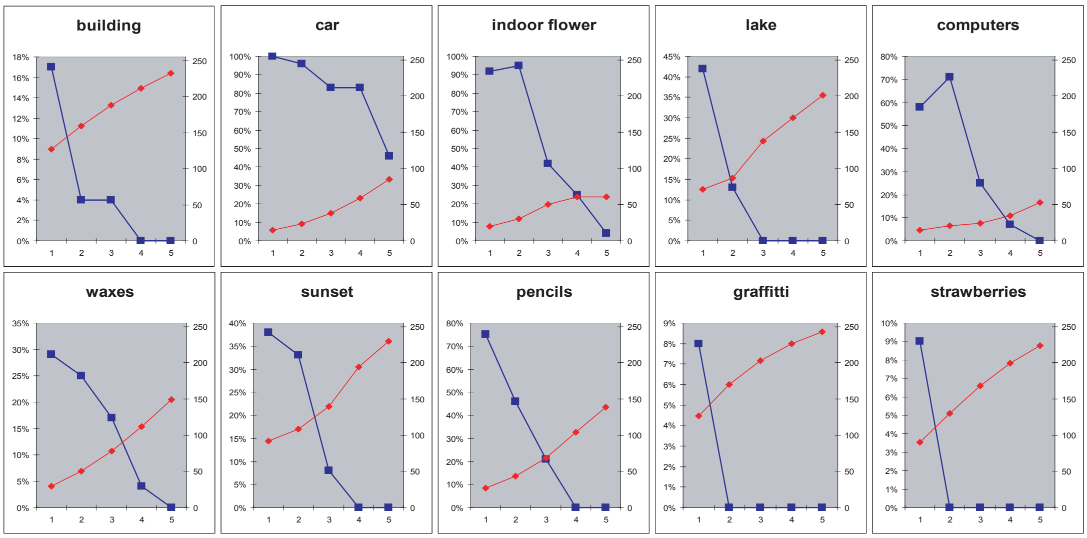

Figure 7: *Inverse correlation between psychophysics results for perceived underexposure (blue) and pixel luminance mean (red). X-axis represents the five exposures for each scene; Y-axis represents the percentage of subjects who perceived the stimulus as underexposed (left) and mean luminance values (right). Note the changing scale in the Y-axis.*

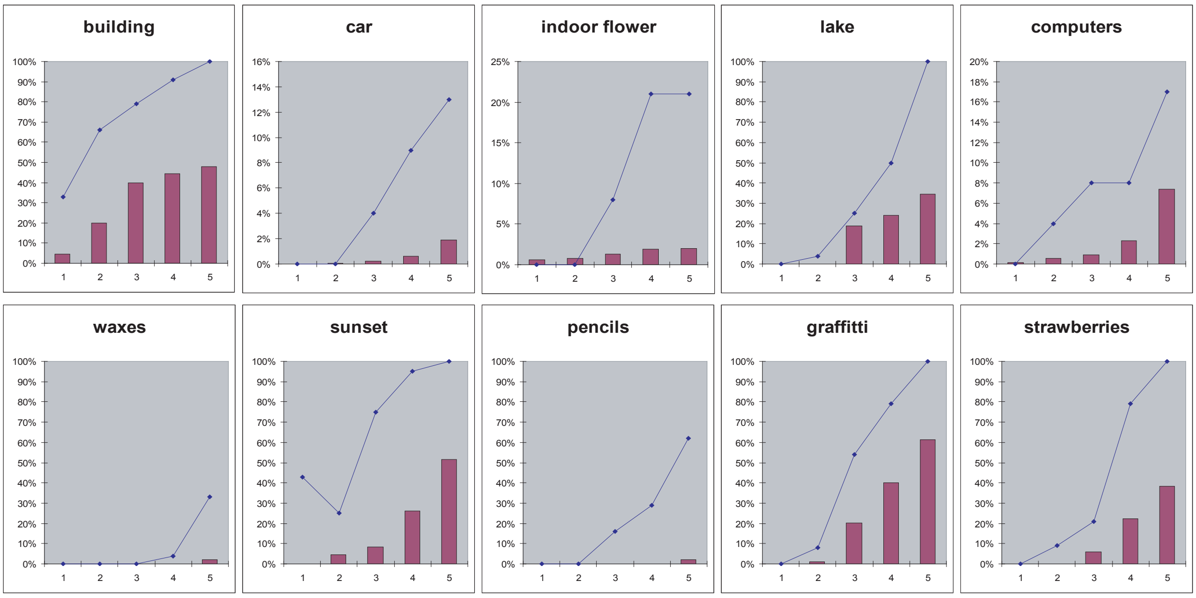

Figure 8: *Correlation between psychophysics results for overexposure (blue lines) and percentage of overexposed pixels (bars). X-axis represents the five exposures for each scene; Y-axis represents the percentage of subjects who perceived the stimulus as overexposed.*

4. Evaluating rTMO's with incorrect exposures

The results of the previous experiment provide us with a systematic labelling of images as under-, correctly-, and over-exposed. Given this labelling, a key question is how well the existing reverse tonemapping techniques can handle incorrectly-exposed LDR data. The aim of reverse tonemapping is to take LDR content and 'boost it' to HDR without introducing objectionable artifacts. Do any of the existing techniques achieve this goal? Which reverse tonemapping schemes are most appropriate for each level of exposure? To test these questions, we are currently conducting an experiment in which we ask subjects to compare the appearance of reverse tonemapped images on a Brightside DR37-P monitor. The design of the experiment is as follows.

Our goal is to perform a side-by-side comparison of the following four reverse tonemapping schemes:

1. **LDR**: the original LDR image shown on the HDR monitor,
2. **Linear**: the contrast of the original LDR image is linearly scaled to match the displayable range, as described in [AFR*07],
3. **Map**: rTMO based on expand-maps introduced by [BLDC06],
4. **Fly**: the 'on-the-fly' rTMO introduced by [RTS*07].

Stimuli were created as follows. For all 5 exposures of each of the10 scenes (i.e. 50 images), we apply these four rTMOs to the image, to yield four alternative HDR renditions. On each trial, subjects are presented with the four renditions of a given image simultaneously in a randomized 2x2 grid (a 'stimulus quadruple'). Subjects are asked to rank the four images according to how 'visually appealing and compelling' they appear. Subjects are instructed that this is a subjective judgment and that there is no correct answer, they should simply indicate the ordering of their personal preference. Given that previous studies showed that different judgment criteria (such as 'realism', and 'attractiveness') correlate strongly [AFR*07, SLY*06], we decided a single subjective criterion was sufficient.

Blocks of trials consist of all 50 stimulus quadruples in pseudo-random order, with the constraint that consecutive trials cannot feature images from the same scene. Subjects are given unlimited time to respond to each trial. The entire experiment consisted of three blocks of trials. Between blocks, subjects are instructed to take a short pause before continuing with the experiment.

Once the data is analyzed, the results will provide a mean ranking score for each rTMO applied to each exposure level of each scene. This will allow us to determine which rTMO is most effective for each exposure level, and whether there is a general consensus across subjects and across scenes, or whether current rTMOs have to be selected on a case-by-case basis.

5. Conclusions and Future Work

Reverse tone mapping is a process for which no definite solution exists. With the increasing availability of HDR displays, the question of how to display the huge amount of LDR legacy content becomes an important issue. In this paper we have focused on *imperfect* legacy content, more specifically on under- and overexposed material. Rather than attempting to come up with a new reverse tone mapping algorithm, we first have looked into the crucial topic of how exposure is perceived, so that an algorithm can be devised that keeps the look and feel of the original LDR image when viewed on an HDR display. We argue that preliminary steps in this direction are necessary, in order to avoid a proliferation in a near future of multiple co-existing rTMO's, representing partial, incomplete solutions to the problem. According to Google Scholar, there is more than 900 papers written on the topic of tone mapping, which amount to at least a few dozen different algorithms [MS08]. This is a situation we would like to avoid for reverse tone mapping.

From our psychophysical tests, two conclusions are drawn: first, the results seem to confirm that high-level semantics are probably needed for a reliable classification of exposure in images. It could be argued, though, that for some extreme cases this assumption would fail: for instance, a badly washed-out image will most likely be tagged as overexposed even in the absence of any recognizable features (and probably *due to* this absence of recognizable features). However, we believe our assumption holds for a sufficiently large number of cases. Second, we have found a clear tendency for asymmetric exposure perception, which may be related to the functioning of the human visual system.

In any case, both conclusions need to be further investigated, and in that sense we believe there is potential for lots of future research in this area. It could be argued, for instance, that the thresholds chosen for the experiments introduce bias, a topic worth looking into. We are also aware that there is an intrinsic correlation in our chosen parameters (histogram, mean, media and pixel percentages); our results should thus be seen just as a first attempt at providing a taxonomy of visual stimuli for reverse tone mapping research. Nevertheless, we hope to confirm our conclusions with additional tests which will de-correlate these parameters. Higher-order statistics will be analyzed as well, given that visual inspection of the results suggests a correlation with luminance gradients. Finally, more advanced analysis techniques need to be employed.

The psychophysical experiment proposed in Section 4 to evaluate four existing reverse tone mapping algorithms is already being performed by the authors, using a BrightSide DR37-P (display area of 32.26 by 18.15 inches, contrast ratio in excess of 200.000 : 1, black level of 0.015 cd/m^2 and peak luminance of 3000 cd/m^2). We hope to be able to report the results soon in a subsequent publication.

6. Acknoweldgements

This research has been funded by the project UZ2007-TEC06 (University of Zaragoza) and TIN2007-63025 (Spanish Ministry of Science and Technology). Olga Sorkine was partially funded by the Alexander von Humboldt Foundation, while Diego Gutierrez was additionally supported by a mobility grant by the Gobierno de Aragon (Ref: MI019/2007). The authors would like to express their gratitude to Karol Myszkowski and Matthias Ihrke (Max Planck Institute for Informatics), for their help and advices in the early stages of this work. We also thank Francesco Banterle (University of Warwick) and Allan Rempel, Matthew Trentacoste and Wolfgang Heidrich from Dolby Canada for applying their reverse tone mapping algorithms to our stimuli, and Choss, who gave us invaluable photography information and lots of hints to guide our work.

References

[AFR*07] AKYÜZ A. O., FLEMING R., RIECKE B. E., REINHARD E., BÜLTHOFF H. H.: Do hdr displays support ldr content?: a psychophysical evaluation. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers* (New York, NY, USA, 2007), ACM, p. 38.

[Aky] AKYÜZ A. O.: Homesite and additional materials (http://www.coolhall.com/homepage /pubs/hdrdisp_eval/hdrdisp_project.html).

[BLD*07] BANTERLE F., LEDDA P., DEBATTISTA K., CHALMERS A., BLOJ M.: A framework for inverse tone mapping. *Vis. Comput. 23*, 7 (2007), 467–478.

[BLDC06] BANTERLE F., LEDDA P., DEBATTISTA K., CHALMERS A.: Inverse tone mapping. In *GRAPHITE '06: Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia* (New York, NY, USA, 2006), ACM.

[Coh88] COHEN J.: *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.

[ČWNA06] ČADÍK M., WIMMER M., NEUMANN L., ARTUSI A.: Image attributes and quality for evaluation of tone mapping operators. In *Proceedings of Pacific Graphics 2006 (14th Pacific Conference on Computer Graphics and Applications)* (Oct. 2006), National Taiwan University Press, pp. 35–44.

[DBW08] D. BARTZ D. CUNNINGHAM J. F., WALLRAVEN C.: The role of perception for computer graphics. EUROGRAPHICS State of the Art Reports, 2008.

[DCWP02] DEVLIN K., CHALMERS A., WILKIE A., PURGATHOFER W.: Star: Tone reproduction and physically based spectral rendering. In *State of the Art Reports, Eurographics 2002* (September 2002), Fellner D., Scopignio R., (Eds.), The Eurographics Association, pp. 101–123.

[DM97] DEBEVEC P. E., MALIK J.: Recovering high dynamic range radiance maps from photographs. *Computer Graphics 31*, Annual Conference Series (1997), 369–378.

[Kel06] KELBY S.: *The Digital Photography Book*. Peachpit Press, Berkeley, CA, USA, 2006.

[Ken75] KENDALL M.: *Rank Correlation Methods*. Charles Griffin & Co. Ltd, 1975.

[Liv02] LIVINGSTONE M.: *Vision and Art: The Biology of Seeing*. Harry N. Abrams., 2002.

[MDS06] MEYLAN L., DALY S., SÃIJSSTRUNK S.: The Reproduction of Specular Highlights on High Dynamic Range Displays. In *IS&T/SID 14th Color Imaging Conference* (2006).

[MS08] MANTIUK R., SEIDEL H.-P.: Modeling a generic tone-mapping operator. *Computer Graphics Forum (Proceedings of Eurographics 08) 27*, 2 (2008), 699–708.

[RSSF02] REINHARD E., STARK M., SHIRLEY P., FERWERDA J.: Photographic tone reproduction for digital images. *ACM Trans. Graph. 21*, 3 (2002), 267–276.

[RTS*07] REMPEL A. G., TRENTACOSTE M., SEETZEN H., YOUNG H. D., HEIDRICH W., WHITEHEAD L., WARD G.: Ldr2hdr: on-the-fly reverse tone mapping of legacy video and photographs. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers* (New York, NY, USA, 2007), ACM, p. 39.

[RWPD05] REINHARD E., WARD G., PATTANAIK S., DEBEVEC P.: *High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting*. Morgan Kaufmann Publishers, 2005.

[SHS*04] SEETZEN H., HEIDRICH W., STUERZLINGER W., WARD G., WHITEHEAD L., TRENTACOSTE M., GHOSH A., VOROZCOVS A.: High dynamic range display systems. *Proceedings of ACM Transactions on Graphics 23*, 3 (2004), 760–768.

[SLY*06] SEETZEN H., LI H., YE L., WARD G., WHITEHEAD L., HEIDRICH W.: Guidelines for contrast, brightness, and amplitude resolution of displays. In *In Society for Information Display (SID) Digest* (2006), pp. 1229–1233.

[Vap95] VAPNIK V.: *The Nature of Statistical Learning Thoery*. Springer-Verlag, Berlin, 1995.

[YMMS06] YOSHIDA A., MANTIUK R., MYSZKOWSKI K., SEIDEL H.-P.: Analysis of reproducing real-world appearance on displays of varying dynamic range. In *EUROGRAPHICS 2006 (EG'06)* (Vienna, Austria, September 2006), Gröller E., Szirmay-Kalos L., (Eds.), vol. 25 of *Computer Graphics Forum*, Eurographics, Blackwell, pp. 415–426.