The Reproduction of Specular Highlights on High Dynamic Range Displays

Laurence Meylan [1]***, Scott Daly*** [2] ***and Sabine Süsstrunk*** [1]
(1) Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland; (2) Sharp Laboratories of America, Camas WA, USA

Abstract

Recent advances in the design of high dynamic range (HDR) monitors enable the display of images having a large dynamic range, close to that encountered in the real world. As their usage will increase, we will be confronted with the problem of re-rendering images that have been mapped to standard dynamic range (SDR) displays so that they look natural on HDR monitors. We address this issue for SDR images representing original HDR scenes. We propose a tone scale function that takes advantage of the increase in dynamic range of HDR monitors to recreate the brightness of specular highlights, which were clipped or compressed by the capturing and rendering process to SDR. We validate the use of such functions with a psychovisual experiment conducted on an HDR display, where the observers' task was to judge pairs of tone-scaled images. The result of the experiment shows that using part of the extension of dynamic range provided by HDR displays to enhance the brightness of specular highlights leads to more natural looking images.

Introduction

HDR monitors capable of displaying simultaneously bright highlights and dark shadows have just started to come on the market. The development of these monitors raise new questions about how to re-render the large amount of legacy images that are already mapped to SDR displays.

Specular highlights are often badly reproduced in images rendered to SDR displays. This is due to a strong luminance compression and/or clipping taking place during the image capturing and rendering process. As they offer important visual cues about three dimensional shapes and increase the sense of realism [1, 4], it would be beneficial to use part of the extended dynamic range of HDR displays to enhance their representation. The presence of specular highlights in an image suggests that the original scene had a high dynamic range, as specular highlights can be several orders of magnitude brighter than diffuse highlights [7]. We suggest the use of a tone scale function that expands the luminance range allocated to the specular parts of an image with the goal of recovering the natural look of the original HDR scene.

In a psychovisual experiment, we test different tone scale functions by varying the display luminance range allocated to specular highlights. We prove that allocating some of the additional display range provided by an HDR monitor to specular highlights leads to a more natural displayed image than using a simple linear scaling of code values. In addition, the proposed tone scale prevents the re-rendered image to look too bright, which is likely to happen when applying just a linear scaling (illustrated in Figure 1).

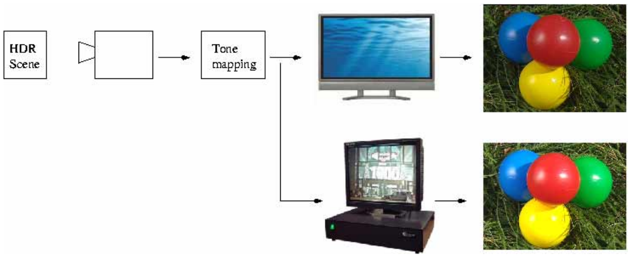

Figure 1. *Problem of re-rendering SDR images to HDR displays. If a simple linear scaling is applied, the image can appear too bright.*

This article is structured as follows: First, we describe the construction of the tone scale function. Then, we explain the generation of the stimuli used in the psychovisual experiment. The experiment procedure is presented followed by a statistical analysis of the collected data. A discussion of the results concludes the article.

The Tone Scale Function

The tone scale function is applied to the luminance channel of a linearly-encoded input image. It is a piece-wise linear function composed of two slopes (Figure 2). Here, we only describe aspects of the tone scale function that are necessary to understand the psychovisual experiment. Implementation details are published elsewhere [5].

The shape of the tone scale is entirely defined by ω, the normalized code value of the maximum diffuse white in the image, and ρ, the percentage of the maximum display luminance allocated to ω.

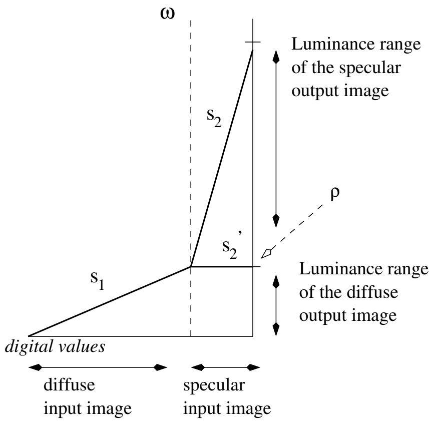

Figure 2. *Piece-wise linear tone scale function.*

ω is determined by segmenting the image into its diffuse and specular components, which we call “diffuse image” and “specular image,” respectively. The specular image is composed of the parts of the image that contain specular highlights. The diffuse image can include glossy and non glossy objects and is composed of the rest of the image that is not part of the specular image. Figure 3 gives an example of a segmentation. The minimum digital value of the specular image defines the maximum diffuse white ω. The segmentation was done manually for each image prior to the experiment. A way to automatically segment the image and compute ω is described in [5].

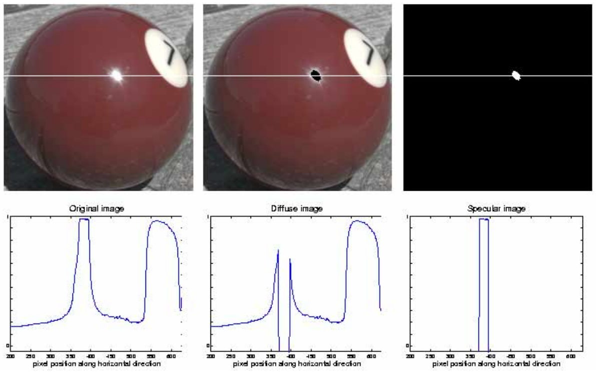

Figure 3. *Example of an image segmentation into its specular and diffuse components. The white line in the top three images represents the position of the traces in the bottom graphs.* **Top left:** *Original image.* **Top center:** *Diffuse image. The specular part of the image is filled with black.* **Top right:** *Specular image. The diffuse part of the image is filled with black.* **Bottom left:** *Horizontal trace in the original image.* **Bottom center:** *Horizontal trace in the diffuse image.* **Bottom right:** *Horizontal trace in the specular image.*

ρ is the parameter tested in the experiment. It varies for each tone-scaled image. The tone scale function f is defined as follows:

$$f(\Lambda(p)) = \begin{cases} s_1 \cdot \Lambda(p) & \text{if } \Lambda(p) \leq \omega \\ s_1 \cdot \omega + s_2 \cdot (\Lambda(p) - \omega) & \text{if } \Lambda(p) > \omega \end{cases}, \quad (1)$$

where

$$s_1 = \frac{\rho}{\omega}, \quad (2)$$

$$s_2 = \frac{1-\rho}{\Lambda_{max} - \omega}. \quad (3)$$

Λ is the normalized luminance and p is a pixel in the image. The maximum digital value of the input image is noted as Λ_{max}. By using $\Lambda_{max} = 1$, we make the method independent of the digital code value range.

The shape of the tone scale (Figure 2) allows the allocation of more dynamic range to the specular image than that allocated in the SDR input (horizontal axis). All pixels of the input image whose normalized code values are smaller than ω are considered being part of the diffuse image and are scaled by s_1. s_2 has a steeper slope and is used to scale the specular image defined by pixels having a value greater than ω.

We added a clipped version of the tone scale where the specular highlight maximum value is not matched to the maximum display luminance ($s_2\prime$ in Figure 2). This enables us to test if participants preferred specular highlights clipped or enhanced given a particular overall image brightness.

Stimuli Preparation

We chose to focus on the re-rendering of images representing HDR scenes and containing specular highlights. The set of images used in the experiment is shown in Figure 4.

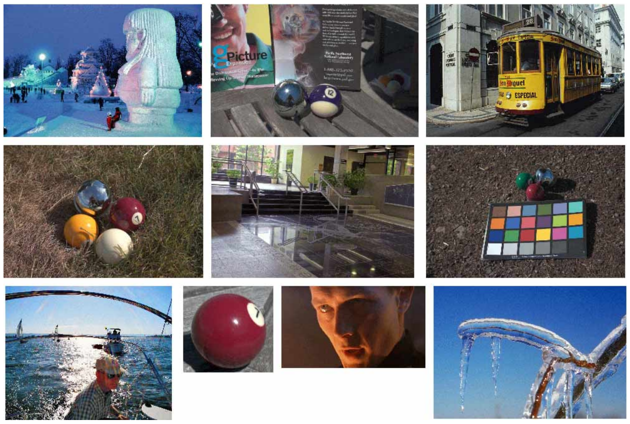

Figure 4. *Set of images used in the experiment.*

For each tested scene, different tone-scaled images are constructed by varying the luminance allocated to the diffuse white. We tested four different values of ρ varying from 20% to 67% of the maximum display luminance (Ψ_{max}) using logarithmic increments, as well as a linear scaling. For the monitor used in the experiment (Brightside 37”), $\Psi_{max} = 2500\ cd/m^2$. This value is reached when measuring a large white patch. With smaller areas such as specular highlights, Ψ_{max} tends to have a lower value. However, the effect of our tone scale function remains valid as long as its general shape is conserved. This is the case for all but extremely small specular highlights, as discussed in the measurement section.

The tone scale functions used in the experiment are shown in Figure 5, for an example ω value. Table 1 shows the corresponding ρ values. For tone scales 1 to 4, ω is matched to 20, 30, 47, and 67 percent of Ψ_{max}, while the maximum code value of the input image is matched to Ψ_{max}. Tone scale 5 corresponds to linear scaling. For one of these tone scales ($\rho = 0.47$), we constructed a clipped version, where the maximum code value of the input image is matched to $\rho \cdot \Psi_{max}$.

Table 1: Tone scales used in the experiment.

	1	2	3	4	5	6
ρ	0.2	0.3	0.47	0.67	lin.	0.47 clipped

For the non-clipped tone scales (1 to 5), changing the value of ρ affects both the image global brightness and the reproduction of specular highlights. The more luminance range is allocated to the diffuse image, the brighter the image appears while simultaneously decreasing the range allocated to specular highlights. A smaller luminance range allocated to the diffuse image causes the image to look dimmer and the specular highlights to look brighter. Figure 6 illustrates these two cases.

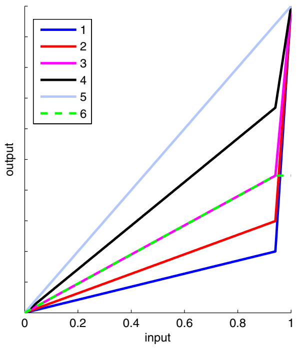

Figure 5. *Illustration of the 6 tone scale functions used in the psychovisual experiment.*

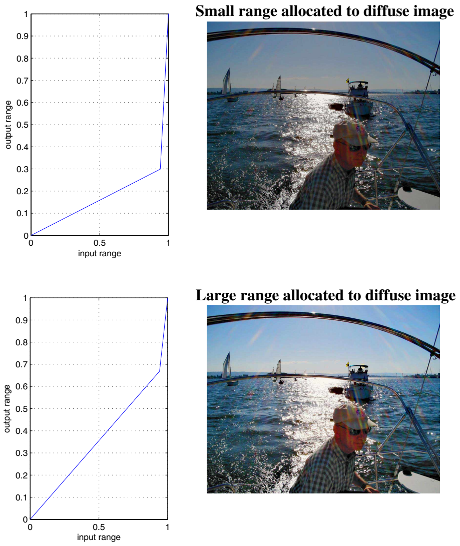

Figure 6. *Example of tone scale functions for two different input parameters. The top image corresponds to the case where a small range is allocated to the diffuse image. The bottom image corresponds to a larger range.* ω *for this image was* 0.94.

A Smoothing Technique to Remove Unnatural Contours

The discontinuity in the tone scale function may produce unnatural contours, which influence the participants' judgment in an undesirable way. We added a smoothing step to our algorithm to overcome this problem. Our solution is to introduce a slight blur around each specular highlight, thus removing unnatural contours [5].

The Generation of Pairs of Tone-Scaled Images

The images thus processed are presented in pairs to the observers. Each image in the pair is computed by a different tone scale. Prior to the experiment, all possible combinations of pairs of images generated with the tone scale functions are computed.

The number of possible pairs N_{pair} generated by T number of tone scale is given by

$$N_{pair} = \frac{T \cdot (T-1)}{2} \tag{4}$$

In our case, $T = 6$ and $N_{pair} = 15$ for each tested image.

The two tone-scaled images composing a pair are scaled and stored as another image having the resolution of the HDR display (1980×1280). A black border of 80 pixels (1.3 degree of visual angle) separates them. We experimented with different border sizes and empirically found that 1.3 degree was sufficient to prevent the brightness of one image from influencing the color of the other one, which would influence the observer judgment in an uncontrolled way. The left/right position of the tone-scaled images is chosen randomly. Examples of stimuli pairs are shown in Figure 7.

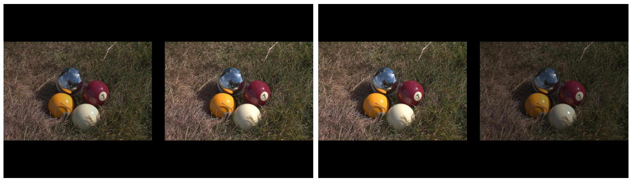

Figure 7. *Example of stimuli shown in pairs.*

The Psychovisual Experiment

Procedure

A computer program displayed pairs of scaled images in random order. For each image of the test set, 15 pairs were presented. Then, the 15 pairs of the next image were shown until all images from the test set have been used.

The process was repeated once with a different image sequence. The pairs of one image were still displayed randomly. The left and right position of the tone-scaled images, which was random for the first sequence, was swapped.

Each time a pair was displayed, the observer used the keyboard to select an image according to the following question, which they could read on the information sheet:

Which image looks more natural (i.e. more like a real scene, like real lighting)? Focus on the tone reproduction; try not to be influenced by other factors (contouring, noise, etc).

Observers

20 observers participated in the experiment, 2 of them had some knowledge about the purpose of the experiment. 14 were naive observers, and 6 were experts in judging image quality. Each of them saw 330 images, which took about 25 minutes.

Viewing Conditions

The experiment was set up in a room with no window. The lights were on, which created an ambient luminance of 22 cd/m^2 on the wall surrounding the display. The images were displayed on a Brightside's 37" HDR monitor. Observers sat at a viewing distance of three times the display height, which resulted in a total viewing angle of 33 degrees.

Measurements Performed on the HDR Display

The maximum displayed luminance of our HDR monitor was obtained by displaying and measuring a large white patch. However, for very small bright areas, this measured value can not be reached. This is due to the characteristics of the HDR display and to the software that provides the conversion between the ideal tone-scaled image (input to the HDR monitor) and the image displayed at the screen. Here we provide a brief summary of the HDR display's hardware and software. The reader is referred to [6] for a detailed explanation.

The HDR display is composed of an array of LEDs providing the backlight, and a LCD panel. A software provides the conversion between the ideal tone-scaled image and the images that are sent to the LED array, and to the LCD panel, respectively. The displayed image is the multiplication of these two images. One important part of this software is the cross-talk correction, which computes the values that drive the LEDs. The goal of the cross-talk correction is to compensate for the fact that the luminance measured at one LED physical position is not only due to one LED but also to remaining light emitted by neighboring LEDs. The model used for cross-talk is limited to six LEDs, which are direct neighbors, but the light contribution of further surrounding LEDs is not zero. Therefore, a small bright area on a dim background suffers from the fact that there is not enough contribution coming from surrounding LEDs and can not reach the maximum displayed value. Consequently, the measured luminances at the screen differ from what is intended by the tone scale function applied to the image.

To understand this behavior better, we measured white patches of varying sizes (simulating specular highlights) using a spectrophotometer (PR650). We used patches of 8, 16, 32, and 64 pixels corresponding to 0.14, 0.27, 0.55, and 1.1 degrees of visual angle. Backgrounds of varying gray levels were used to simulate the luminance allocated to the maximum diffuse white. Example of generated images are shown in Figure 8.

Figure 8. *Generated images for measurements.* **Left:** *Background:* 50% *of* Ψ_{max}*, Specular highlight size: 1.1 degree.* **Right:** *Background:* 10% *of* Ψ_{max}*, Specular highlight size: 0.55 degree.*

Measurements are plotted in Figure 9. We observe that the smaller the specular highlight is, the lower is the display luminance. Moreover, the luminance of the area surrounding the specular highlight also influences the actual measured value. The darker it is, the lower the specular highlight measured value is.

Consequently, the actual applied tone scale varies locally and depends on the size of the specular highlights as well as on the luminance value allocated to the maximum diffuse white.

Figure 10 gives an example of a theoretical tone scale and the two corresponding actual tone scale functions for different specular highlight sizes. We assume that the image contains a large diffuse white area. With large specular highlights, the actual tone scale approaches the intended behavior. However, with small specular highlights, it is possible that the measured luminance of a large diffuse white area exceeds that of a specular area, despite the behavior intended by the tone scale function.

This display limitation has an influence on the type of images that need to be chosen for the psychovisual experiment. Images with small specular highlights can not be used to validate our proposed tone scale. Based on our measurements, we consider that the diameter of a specular highlight must be more than 16 pixels for the results to be meaningful.

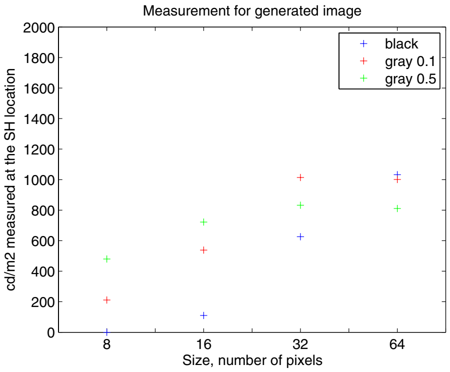

Figure 9. *Measurements of simulated specular highlights. The horizontal axis shows four different sizes of specular highlights. Each color corresponds to a different background luminance value.*

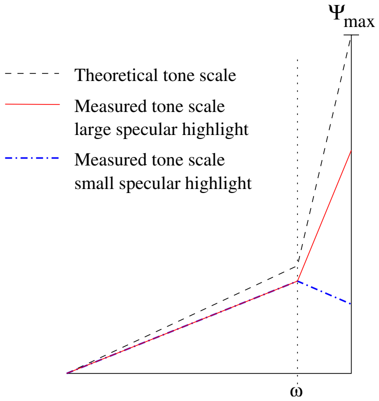

Figure 10. *Example of a theoretical tone scale and the two corresponding actual tone scale functions. Diffuse white area is assumed to be large. With large specular highlights, the actual tone scale approaches the intended behavior. With small specular highlights, the measured luminance of a large diffuse white area exceeds that of a specular area.*

Results

Statistical Analysis

Thurstone's law of comparative judgment Case V [2] was applied to convert the paired comparison observer data into an interval scale of preferences. We used the toolbox provided in [3], which calculates the z-scores and confidence intervals from such data.

With Thurstone's law of comparative judgment, unanimous judgments (i.e., when a stimuli is preferred by all observers or no observer) are problematic as corresponding z-value are undefined. This problem is referred as "zero proportion matrix problem." It is solved by substituting missing z-values using a linear regression technique.

The interval scale of preferences along with 95 % confidence intervals are shown in Figure 11. For two tone scales to be considered significantly different, their errors bars must not overlap. The display luminance allocated to maximum diffuse white increases from tone scale 1 to 5. Tone scale number 6 is the clipped version of tone scale number 3. Their diffuse luminances are the same but specular highlights of tone scale 6 are not boosted up.

In Figure 11, we also included the percentage of specular pixels in each image. It is denoted by r and computed as follows:

$$r = \frac{N_{specular}}{N}, \tag{5}$$

where $N_{specular}$ is the number of specular pixels given by the segmentation of the image in its diffuse and specular components (Figure 3) and N is the total number of pixels in the image.

The six plots represent six different images that we selected from our set to give representative results. Indoor and outdoor scenes containing various specular highlight sizes are shown. In the discussion that follows, the term "prefer" is used to describe observer choice. However, it is important to remember that it relates to a sensation of naturalness.

For the two images at the top of Figure 11, participants significantly preferred tone scale 4 over a simple linear scaling (5). At equal brightness (tone scales 3 and 6), they selected the tone scale with bright specular highlights (3) significantly more than the clipped one (6).

For images (c), (d), and (e), our tone scale is slightly preferred than linear scaling but not statistically different. At equal brightness, the images with bright specular highlights are statistically judged to be better than the ones with clipped highlights. This can be explained by the fact that these three images represent outdoor scenes and thus participants expected a very bright scene. Similarly, image (a) benefits from a low luminance allocated to diffuse white probably because observers recognized it as an indoor scene and expected a dim overall impression. Concerning the boat image (b), the lower luminance preference despite the fact that this is an outdoor scene can be explained by the size of the specular highlight, which is quite larger than in image (c), (d), and (e). In this case, the large size of the specular highlight changes the overall impression of brightness, which influences the observer's preferences.

Image (f) is an example of a problematic image, i.e. containing few small specular highlights (< 16 pixels). We showed with the measurements performed on the HDR monitor that small specular highlights were not scaled as much as predicted due to some display limitations. Consequently, in image (f), the increase in luminance of the specular highlights performed by the tone scale function could not be displayed on the screen. This explains why tone scale 3 and 6 are statistically equivalent.

Discussion

The results of this experiment show that the preferred luminance range allocated to the diffuse image varies with the image content. Different tendencies can be observed for indoor scenes and outdoor scenes. For outdoor scenes, observers tend to select images where only a small part of the dynamic range is allocated to specular highlights. However, with images of equal diffuse brightness, they select the image with bright highlights.

For indoor scenes, the participants clearly prefer to allocate more range to the specular highlights instead of a linear scaling, which would result in an unnaturally bright image. When comparing images of equal diffuse brightness, the image with bright specular highlights is also significantly preferred.

It was also shown that the percentage of specular pixels in the image plays a role in the observer's judgment. In the case of large specular highlights, the overall impression of brightness is changed and observers tend to prefer dimmer images. Very small specular highlights appeared to be problematic due to the display characteristics.

To conclude, the use of a tone scale that boosts the specular highlights instead of rendering a globally brighter image is validated for indoor scenes. Most importantly, the results of the comparison between clipped and non-clipped specular highlights in images of equal diffuse brightness confirmed that bright specular highlights lead to a more natural impression, for all tested images.

Conclusion

The recent marketing of HDR displays opens new research opportunities in the field of HDR imaging as well as related applications. This article focuses on the conversion of SDR images (whose original scenes were HDR) into images that can be displayed on an HDR monitor. We present a tone scale function whose goal is to improve the realism of specular highlights. The use of such a tone scale is justified by a psychovisual experiment.

This experiment suggests that when using an HDR display, it is preferable not to use the entire dynamic range for the diffuse component of the input image despite the reduction in mean brightness. Instead, part of the dynamic range could be used to provide a better reproduction of specular highlights and thus increase the realism of the displayed image. More importantly it confirms that at equal diffuse brightness observers significantly prefer images with brighter specular highlights.

References

[1] Andrew Blake and Heinrich Bülthoff. Shape from specularities: computation and psychophysics. *Philosophical Transactions of the Royal Society of London, B: biological sciences*, 331(1260):237–252, February 1991.

[2] Peter G. Engeldrum. *Psychometric scaling: A toolkit for imaging systems development*. Imcotek press, Winchester, MA, 2000.

[3] Phil. J. Green. A colour engineering toolbox. http://www.digitalcolour.org/toolbox.htm, 2003.

[4] Victoria Interrante, Henry Fuchs, and Stephen M. Pizer. Conveying the 3d shape of smoothly curving transparent surfaces via texture. *IEEE Transactions on Visualization and Computer Graphics*, 3(2):98–116, April-June 1997.

[5] Laurence Meylan. *Tone Mapping for High Dynamic Range Images*. PhD thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL), 2006.

[6] Helge Seetzen, Wolfgang Heidrich, Wolfgang Stuerzlinger, Greg Ward, Lorne Whitehead, Matthew Trentacoste, Abhijeet Ghosh, and Andrejs Vorozcov. High dynamic range display systems. *ACM Transactions on Graphics (special issue SIGGRAPH 2004)*, 23(3):760–768, August 2004.

[7] Lawrence Wolff. On the relative brightness of specular and diffuse reflection. In *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 1994)*, pages 369–376, Seattle, WA, June 1994.

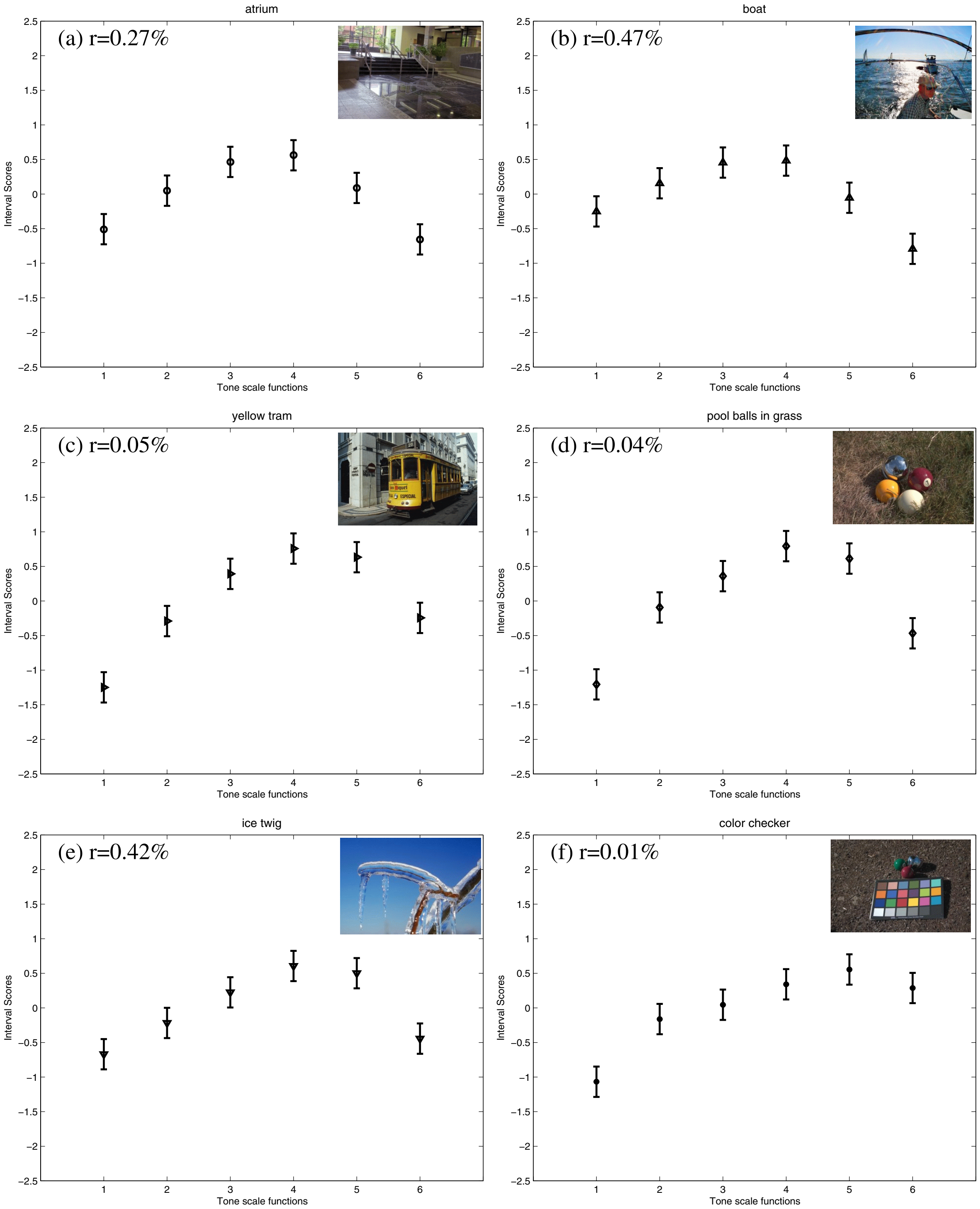

Figure 11. **Top left (image a):** *Tone scale 4 is significantly preferred than linear scaling but not significantly preferred than 3. At equal brightness, the non-clipped version (3) is significantly preferred than the clipped tone scale (6).* **Top right (image b):** *Tone scale 4 is significantly preferred than linear scaling but statistically equivalent to 3 and 2. At equal brightness, 3 is significantly preferred than 6.* **Middle left (image c):** *Tone scale 3,4,5 were significantly preferred than 1,2,6. At equal brightness, the non-clipped version (3) is significantly preferred than the clipped tone scale (6).* **Middle right (image d):** *Tone scale 4,5 are statistically better than 1,2. At equal brightness, the non-clipped version (3) is statistically better than the clipped version (6).* **Bottom left (image e):** *Tone scale 4 and 5 are preferred over 1,2,6. At equal brightness, tone scale 3 is statistically better than 6.* **Bottom right (image f):** *Tone scale 5 (linear) is statistically preferred over 1,2,3. Clipped (6) and non-clipped tone scales (3) are equivalent. These results are due to the small size of specular highlights. r gives the percentage of specular pixels in the images.*