Eurographics Symposium on Rendering - Experimental Ideas & Implementations (2015)
J. Lehtinen and D. Nowrouzezahrai (Editors)

Color Clipping and Over-exposure Correction

Mekides Assefa Abebe*[1,2], Tania Pouli[1], Jonathan Kervec[1] and Chaker Larabi[2]

[1]Technicolor Research & Innovation, Cesson-Sévigné, France
[2]Université de Poitiers, Poitiers, France

Abstract

Limitations of the camera or extreme contrast in scenes can lead to clipped areas in captured images. Irrespective of the cause, color clipping and over-exposure lead to loss of texture and detail, impacting the color appearance and visual quality of the image. We propose a new over-exposure and clipping correction method, which relies on the existing correlation between RGB channels of color images to recover clipped information. Using a novel region grouping approach, clipped regions are coherently treated both spatially and temporally. To reconstruct over-exposed areas where all channels are clipped we employ a brightness profile reshaping scheme, which aims to preserve the appearance of highlights, while boosting local brightness. Our method is evaluated using objective metrics as well as a subjective study based on an ITU standardized protocol, showing that our correction leads to improved results compared to previous related techniques. We explore several potential applications of our method, including extending to video as well as using it as a preprocessing step prior to inverse tone mapping.

Categories and Subject Descriptors (according to ACM CCS): I.4.4 [Image Processing and Computer Vision]: Restoration—I.4.3 [Image Processing and Computer Vision]: Enhancement—

1. Introduction

Pictures are worth a thousand words. Yet often, when capturing a sunset or a brightly colored object, the brightest and most saturated parts of the scene appear clipped and flat in the resulting image. Our goal in this work is to recover lost information due to clipping, improving the visual quality of images impaired in this way. We address this problem in two parts: *over-exposure*, where all three channels are clipped and *color clipping*, where only one or two channels are clipped and some information remains.

Over-exposure occurs when camera sensors receive more light than their maximum capacity, resulting in washed-out image regions where all three color channels reach the maximum allowed intensity. This, for instance, may occur when a scene is captured with incorrect camera settings or when the dynamic range of the scene is higher than that of the camera. On the other hand, color clipping occurs when intensity information for one or two of the color channels is outside the intended device gamut or beyond the capacity of the camera sensor, e.g. when capturing a bright and saturated object or near over-exposed areas. This results in loss of image details and color information, potentially altering the color appearance and reducing the visual appeal of images.

In addition to decreasing the visual quality of images, the loss of image detail in clipped areas complicates image analysis and computer vision [ASI91,WIH10,Has13]. For example, many computer vision applications expect linear input data, an assumption broken in over-exposed and clipped image areas. To add to the issue, display capabilities are growing at fast rates, with high dynamic range (HDR) and wide gamut displays starting to become available to consumers. Given the lack of native HDR and wide gamut content, this is creating an unparalleled demand for ways of preparing legacy content so that it takes advantage of such displays. However, to maximize the visual appeal of legacy content on new displays, it is not sufficient to expand the dynamic range, but it is also desirable to recover the details and color information lost due to over-exposure or color clipping.

Although the correction of these artifacts has been addressed before, the quality of results often comes at the cost of high computational and algorithmic complexity, making most existing methods not well suited for interactive implementations. The proposed method relies on conceptually and

* e-mail:Mekides.Abebe@technicolor.com

DOI: 10.2312/sre.20151169

algorithmically simple processing steps, allowing us to correct clipping and over-exposure in a wide variety of content at interactive speeds. We handle color clipping first, relying on the correlation between RGB color channels. Over-exposed regions are then reconstructed using the color clipping corrections and a novel brightness profile reshaping. To avoid artifacts due to spatial processing, color clipped and over-exposed pixels are segmented and categorized according to a novel color and connectivity based scheme, which is evaluated both spatially and temporally. This allows us to handle complex scenarios, such as sky shown behind tree branches or bright areas separated by shadows.

We assess the quality of our results in relation to HDR ground truth imagery and to alternative methods through a subjective study that follows an ITU recommended protocol [ITU12]. Additionally, we evaluate the visual quality of our processing using HDRVDP2 [MKRH11] and show many results on a variety of scenes. Our analysis indicates that our technique offers a significant improvement over related methods, while allowing for corrections to be applied interactively. We explore the application of our correction method as a preprocessing step for inverse tonemapping and in a simple video extension. Finally we demonstrate an implementation of our method in the context of a professional color grading environment, allowing for interactive corrections with optional user guidance.

2. Related Work

The problem of over-exposure and clipping can either be solved directly on the camera side [HS11], or after the image is captured on the display side. Our work belongs to the latter case. Even though lost information due to clipping can be reconstructed by transferring information from similar images contained in online photo databases [JMKA10, ZGW*14], our work only considers information within the input itself, as in most cases it provides sufficient information to recover clipped areas, obviating the need for external data sources. Bright or clipped parts in images can either be enhanced as part of dynamic range extension, or can be directly corrected to improve image quality. We discuss related methods in both these scenarios.

Inverse Tone Mapping. The advent of HDR displays has given rise to solutions for expanding the dynamic range of existing content. The expansion typically employs inverse tone mapping operators (ITMO), which increase the contrast and dynamic range of the image, either globally or with the addition of local processing of highlight areas. To emphasize the appearance of bright image regions, the expansion process can be applied with different weights depending on the brightness of each pixel [BLD*07]. Alternatively, to better approximate the luminance distribution and perception of the original scene, the image can be separated in diffuse and specular components, which are then processed separately to construct the HDR result [MS06, DMHS08, MFSG10]. In all these cases however, although the highlight regions might better approximate the appearance of real highlights, they remain flat, unlike real specularities.

To account for this limitation, Rempel et al. [RTS*07] apply a Gaussian brightness enhancement to bright regions to give shape to the inverse tone mapped specularities. Other approaches opt for reconstructing missing information based on available well-exposed regions in the image, either by manually annotating similar image regions [WWZ*07], or by automatically detecting appropriate texture information in the image [SMZ*11]. Although the latter two solutions can effectively reconstruct information in textured areas, they are less suited to general cases of over-exposure as they require a suitable texture to be present in the image.

Over-exposure Correction. To address more general cases of over-exposure, where no information remains in the clipped areas, many approaches propagate information from unclipped nearby regions. Rather than using image values directly, gradient information can be used to recover the structure in the over-exposed regions [RLH12]. Guo et al. [GCZS10] operate in the CIELab color space, and compress pixels in the image non-linearly according to a probability map encoding over-exposure confidence. Non-clipped pixels are compressed using the tone mapping operator of Fattal et al. [FLW02] to create space for the over-exposed pixels. At the same time, chromatic information is propagated to the over-exposed areas according to neighborhood similarity. Such interpolation and extrapolation techniques heavily rely on the assumption that the original content of over-exposed areas is similar to its under-exposed neighbors. As such, they often lead to over-smoothed results, where specularities are lost or colors are propagated beyond the object boundaries. There are also methods [HJS13] with a bit more complex wavelet based inpaininting approach.

Color Clipping Correction. When only one or two channels are clipped, information in the same spatial location within the remaining channels can be used for reconstruction, exploiting the correlation between channels. At the simplest case, this can be done at the scanline level, however in this case only one and two channel information can be recovered [APKL14]. In the YCbCr color space, strong correlation is present between the two chromatic channels. Xu et al. [XDN11] rely on this correlation and interpolate chromatic information into the clipped regions through a normalized convolution. Instead, Fu et al. [FPCM13] observe that Cb and Cr values of over-exposed regions lie at the boundaries of the chromatic plane, and therefore move these values towards the center of the plane to reduce saturation.

The known correlation in the RGB channels of color images [RP11] has also been exploited for this purpose. To recover clipped information, Zhang et al. [ZB04] use a Bayesian model of the global correlation between RGB channels, which is extended by Masood et al. [MZT09] to estimate ratios between RGB channels of clipped pixels by using the relationship between pixels, at the same spatial lo-

cation and their neighborhood. Although their use of neighborhood information allows this method to propagate chromatic information into the clipped regions, color bleeding may occur into connected but semantically different regions.

More complex color statistics have also been employed to guide the reconstruction of clipped areas. Mansoor et al. [MSNW10] formulate this problem in terms of sparse signal reconstruction, while Elboher et al. [EW10] detect *Color Lines* in the color distribution and expand them along a monotonically increasing function. Although this works well for images with distinct, colorful objects with accurate color segmentation, Color Lines may not be as easily distinguishable in complex scenes with multiple objects.

Our method relies on cross-channel correlation to correct regions where one or two channels are clipped, similar to Masood et al. [MZT09]. However, to ensure that even disconnected but similar regions are corrected in a coherent manner, we propose a novel segmentation and region grouping scheme that considers both chromatic differences and distance between regions. At the same time, we reconstruct fully over-exposed areas using a novel brightness profile reshaping approach that relies on a combination of inpainting and edge-aware filtering. In contrast to the majority of previous work, we additionally explore the application of our method in the context of video, by extending the concept of region similarity in the temporal domain as well. This is achieved with only a single frame delay, forgoing costly multiple passes over the video, making our solution potentially well-suited for interactive or broadcasting applications.

3. Single Frame Correction

Our method addresses both color clipping, where one or two channels are saturated, and over-exposure, where all three channels are saturated. To address the former, we rely on the fact that in clipped image areas, the RGB components tend not to be clipped at the same spatial position. Because of that, clipped data in one channel can overlap with well-exposed information in another. At the same time, values in the three RGB channels tend to be highly correlated. Based on these two observations, we use the local variation in unclipped channels at the same spatial location to reconstruct missing information in areas of color clipping. To correct the remaining over-exposed areas where all channels are clipped, we opt for brightness profile reshaping step that relies on the one and two channel corrections to determine how each over-exposed region should be expanded.

The input to our algorithm is an 8-bit RGB video or image, linearized by removing gamma correction. The corrected images are encoded using linear RGB values and often require more than 8 bits, as their chrominance and luminance range is expanded beyond 255. As such, our corrected results can be further compressed to an 8-bit range or expanded to a wider dynamic range depending on the target application as shown in Section 6. All results shown, unless otherwise specified, are linearly scaled according to the maximum output intensity and gamma corrected.

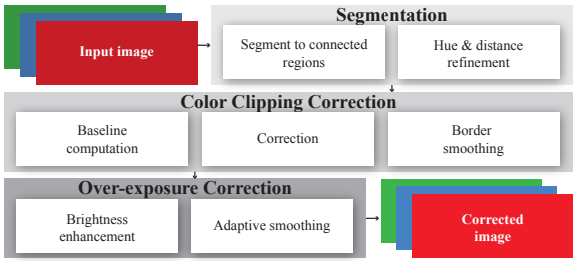


Figure 1: *An overview of the steps of our correction method.*

The steps of our method are outlined in Figure 1. First, clipped image regions are detected and segmented, producing a map of segments for one, two or three channel clipping cases (Section 3.1). The segmentation is further refined based on chromatic and spatial information, grouping segments into coherent regions to be treated together. For each of these regions, a baseline intensity is determined and our color clipping correction is then applied (Section 3.2). Finally, over-exposed segments are reconstructed according to the corrections of color clipped regions (Section 3.3). Our correction process is applied in 3 steps. Single channel clipping is addressed first, followed by correcting areas where two channels are clipped. Finally, using these intermediate corrections, over-exposed areas can be corrected.

3.1. Segmentation and Region Grouping

Images often consist of multiple similar objects and surfaces, which, if clipped, should be treated in a consistent manner. As such, we first organize clipped pixels into a set of regions $\mathbf{R}$ for each channel C in the image I, so that pixels within each region have similar characteristics and likely belong to similar objects. This ensures that the applied correction is coherent across the image and similar objects or parts of the same object obtain the same appearance. Given an input 8-bit RGB video frame or image, we first construct a mask M marking clipped areas in each channel according to a threshold τ_{clip}.

An initial estimation of the set of regions $\mathbf{R}_{\text{con}}$ is determined from the mask M for each channel by dividing the mask into connected components based on 8-neighbor connectivity. Even though the maximum value encoded in 8-bit imagery is 255, similar to previous methods [MZT09], we set τ_{clip} to a lower value ($\tau_{\text{clip}} = 235$) to compensate for noise and further processing that may have been applied to the image, such as color grading that might reduce the peak brightness. Although higher threshold values are more conservative, we find that lower values of τ_{clip} lead to more visible details recovered in the image.

Hue and Distance Refinement. To ensure that similar

a. Input b. No region merging c. Region merging

Figure 2: *Sky regions showing between the branches (a) should be consistently treated. If these regions are processed separately, their appearance becomes inconsistent (b). Our hue and distance refinement step treats these regions similarly, producing a more consistent result (c).*

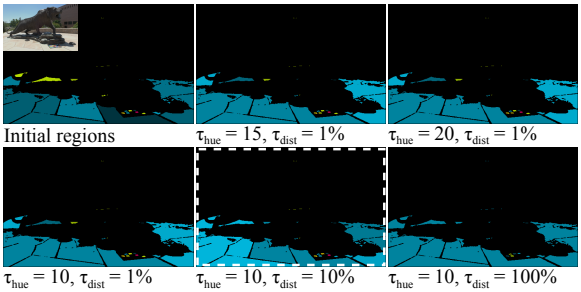


Figure 3: *The effect of different values of the hue and distance thresholds in Eq. (2) on the region grouping step. The image with the dashed border represents our default settings.*

clipped regions are treated in a consistent way, the initial connected regions $\mathbf{R}_{\text{con}}$ are further refined according to chromatic and spatial information. Intuitively, nearby areas of similar color, are likely to belong to the same object, even if they are not actually connected. An example showing the effect of the region merging step is shown in Figure 2. To measure chromatic similarities, one can choose any of the common color spaces. For this particular work, we rely on CIELab hue. For a region $R_{\text{con},i} \in \mathbf{R}_{\text{con}}$ we compute its hue difference Δh with each other region $R_{\text{con},j}$ as follows:

$$\Delta h(i,j) = \sqrt{(\bar{a}(R_{\text{con},i}) - \bar{a}(R_{\text{con},j}))^2 + (\bar{b}(R_{\text{con},i}) - \bar{b}(R_{\text{con},j}))^2} \tag{1}$$

where $\bar{a}$ and $\bar{b}$ denote the average values for each region in the CIELab chromatic channels a and b respectively. Note that this process is necessary instead of directly computing the average hue to bypass wrap-around issues due to the circular nature of the hue channel. The spatial distance d between regions $R_{\text{con},i}, R_{\text{con},j}$ is computed as the smallest distance between the bounding boxes for each region. We opted for this approach to simplify computations although it may reduce accuracy in the distance calculation. In practice, we found

To obtain the final set of regions $\mathbf{R}$ for each channel, two regions $R_{\text{con},i}$ and $R_{\text{con},j}$ are merged depending on their distances:

$$R_i = \begin{cases} R_{\text{con},i} \cup R_{\text{con},j} & \text{if} \quad \Delta h(R_{\text{con},i}, R_{\text{con},j}) < \tau_{\text{hue}} \quad \wedge \\ & \quad d(R_{\text{con},i}, R_{\text{con},j}) < \tau_{\text{dist}} \\ R_{\text{con},j} & \text{otherwise} \end{cases} \tag{2}$$

where τ_{hue} and τ_{dist} are thresholds for hue difference and distance between regions, set to 10 and 1% of the largest image dimension respectively. Figure 3 shows the effect of these parameters for one example image. This process is computationally cheap since comparisons are done on a relatively small number of regions, rather than directly on pixels.

To handle cases where two channels are clipped at the same spatial location, a similar process is followed. However, in contrast to the segmentation and refinement procedure described above, we consider pairs of channels against the threshold τ_{clip} when constructing the mask M.

3.2. Color Clipping Correction

Once the image is segmented into coherent regions, the next step in our algorithm is to correct one and two channel clipping. For each clipped region R_i of a channel C_{clip}, a reference channel C_{ref} is selected and a baseline intensity ρ_i value is determined for that channel. Using this information, local variation is reconstructed in the clipped channel(s). An illustration of this process is shown in Figure 4.

Reference Channel Selection. Before computing the baseline intensity or reconstructing clipped information, we need to select a reference channel. If only one channel is saturated the average value of the remaining unsaturated channels will be used as a reference channel for our correction. As we rely on the assumption that color channels are correlated, using both channels improves the robustness of the reference information against chromatic features that might be present only in one channel. In cases where two channels are saturated, we have to use the only well-exposed channel remaining as a reference channel for the correction of the two others. Here for clarity, we will describe the correction process for only one region R_i of a single-channel saturation case. The same process is applied to all regions of each one and two channel clipping case.

Baseline Intensity Computation. As a consistency constraint, we determine a baseline intensity value ρ_i for each region R_i, which remains constant for all pixels of the region. This value serves to separate local variation within the channel from the global intensity level of that region. Once ρ_i is determined, we can replicate variation within a reference channel that is above that value into the clipped channel. The baseline intensity ρ_i is computed as the minimum pixel intensity value from all the pixels p in the segment R_i in the reference channel, $\rho_i = \min(p \in R_i)$. Note that the hue and distance refinement step in Section 3.1 ensures that the baseline intensity ρ_i is the same for disconnected regions that semantically belong together.

Correction. Using the baseline intensity ρ_i, for each pixel p within a clipped region R_i, the corresponding corrected

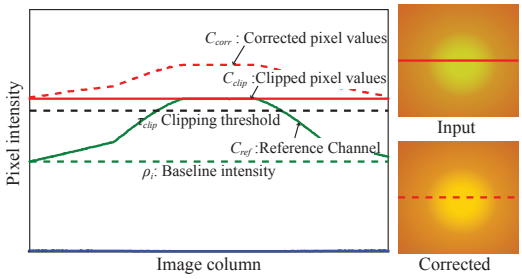


Figure 4: *Visualization of our correction process for a sample one channel clipping case. In this particular case, the red channel is clipped in all pixels of the line profile of an image and the green channel is used as a reference channel.*

pixel value will be computed as follows:

$$C_{\text{corr}}(p) = C_{\text{clip}}(p) + C_{\text{ref}}(p) - \rho_i \quad (3)$$

This process (Figure 4) is repeated for each of the one/two channel clipping cases using the corresponding reference channel. Although the correlation between color channels would suggest that a multiplicative solution should be employed, in practice we have found that this can lead to an over-saturated appearance in the corrected area. Instead an additive correction step as described in Equation 3 in practice leads to a gradual decrease in saturation towards the center of the region, as intensity increases, which may appear more natural particularly for specular surfaces.

Smoothing Discontinuities. To prevent discontinuities and contouring artifacts at the corrected region boundaries, a smoothing step is necessary. In the case of single channel clipping, the borders of clipped regions are feathered against the input image. When correcting 2 channel clipping (cases c4 to c6), these regions are likely to border areas with one channel clipping that have been already corrected. As a surface becomes brighter, clipping artifacts become progressively worse. As such, areas where two channels are clipped likely border already partially clipped regions.

To smooth discontinuities, a mask is extracted for the current region R_i, such that $M_i(p) = 1$ if $p \in R_i$. It is then filtered with a simple averaging filter obtaining the smoothed mask $M'_{\text{smr},i}$. Due to the filtering, the mask $M'_{\text{smr},i} = 1$ for pixels inside R_i and between 0 and 1 close to the borders of R_i, allowing us to smoothly interpolate between the input and corrected values of those pixels. A pixel p in the corrected channel C_{corr} is finally smoothed as follows:

$$C'_{\text{corr}}(p) = (1 - M'_{\text{smr},i}(p))C_{\text{ref}}(p) + M'_{\text{smr},i}(p)C_{\text{corr}}(p), \quad (4)$$

where C'_{corr} encodes the final clipping correction for that channel.

3.3. Over-exposure Correction

Over-exposed areas where all three channels are clipped require different treatment, since no information remains in any of the channels. Similar to color clipping, the first step in this case too is to detect connected regions where all channels are clipped, and the connected component analysis as described in the previous section. Then, correction of these areas follows, as described below.

Brightness Enhancement. After the color clipping correction described in the previous section, pixels affected by one and two channel clipping may obtain brightness values beyond 255. However, the pixels within over-exposed regions will still be limited to 255. As such, the brightness of each over-exposed region $R_{\text{ov},i}$ is first raised to the local maximum value, determined from the color clipping correction step. To compute this local maximum, we need to consider pixels outside the border of the over-exposed area, since this area has not been corrected yet. Therefore, a dilated mask $M_{\text{ext},i}$ is computed by dilating the clipped region by a small number of pixels. In practice, we found that a 3 pixel dilation is sufficient. Based on this mask, for each region $R_{\text{ov},i}$, a maximum pixel intensity z_i is computed as $z_i = \max_{p \in M_{\text{ext},i}}(I(p))$. Pixels within $R_{\text{ov},i}$ are then set to the maximum z_i across channels, so that the region remains achromatic.

Adaptive Smoothing. The brightness enhancement step pushes over-exposed regions to a brightness level corresponding to the surrounding corrected regions, but is likely to create discontinuities at region borders. At the same time, the brightness profile within over-exposed regions is still flat. To reshape the brightness profile and smooth region boundaries we are looking for a filter that can respect actual image edges (e.g. the edges of objects), smooth discontinuities created at the borders of over-exposed regions, and create a smooth brightness profile that preserves the appearance of specularities.

We use the bilateral filter [TM98] for this purpose, as it naturally respects image edges. Although it would be possible for a given image to find filtering parameters so that the second condition is also satisfied, this is likely to be image specific and therefore not robust. Instead, we opt to remove the discontinuities at the region borders through an inpainting step. First, we compute a representation of the region borders with varying thickness by a combined dilation and erosion operation on the region mask (see Figure 5(b-d). The identified border pixels are then inpainted through a least squares interpolation step [Gar10], obtaining $R_{\text{inp},i}$ (Figure 5(e). Depending on the border thickness κ, more or less of the region is inpainted. We set $\kappa = 3$ in the case of specular highlights. Although the inpainting result removes or at least reduces discontinuities at regions borders, using the inpainted region directly might reduce the specular appearance of highlights. Instead, to fulfill the third condition for our filter, we use the inpainted region $R_{\text{inp},i}$ as an edge

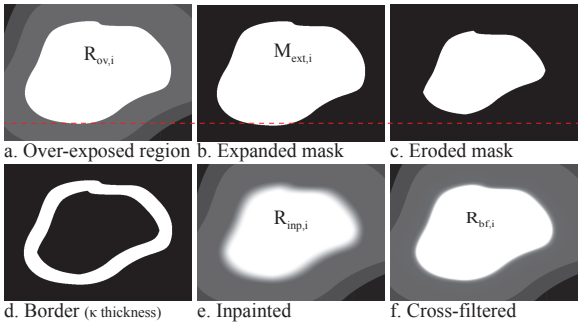


Figure 5: *An illustration of the different masks used in the adaptive smoothing step of our over-exposure correction. The red dotted line shows how the border of the different masks relate to each other.*

map in a cross bilateral filtering step, where $R_{\text{ov},i}$ is filtered to obtain $R_{\text{bf},i}$:

$$R_{\text{bf},i}(p) = \frac{\sum_{q\in\Omega} G_{\sigma_s}(||p-q||)G_{\sigma_r}(R_{\text{inp},i}(p)-R_{\text{inp},i}(q))R_{\text{ov},i}(p)}{W_p} \tag{5}$$

where G_{σ_s} and G_{σ_r} are Gaussians in the spatial and range domain respectively, and $W_p = \sum_{q\in\Omega} G_{\sigma_s}(||p-q||)G_{\sigma_r}(R_{\text{inp},i}(p)-R_{\text{inp},i}(q))$ is a normalization term.

Depending on the parameters used in the filtering, the brightness profile of the region will be more or less sharply peaked. Since $R_{\text{ov},i}$ is achromatic, the peak of the profile will remain white, but it will smoothly transition into the chromatic areas. We adapt the spatial and range σ to the region properties. Specifically, σ_r is set to 0.5 of the range of $R_{\text{in},i}$ so that any residual discontinuities after the filling-in step will be filtered but without crossing high contrast edges. For the results in the paper, the spatial parameter is set to $\sigma_s = (cN_i)^\gamma$, where N is the number of pixels in $R_{\text{ov},i}$. The parameters $c = 0.01$ and $\gamma = 0.8$ were set empirically such that over-smoothing in the region is avoided and the profile of the specularity is preserved. Lower γ leads to a more pronounced intensity profiles in the corrected regions, while higher values of c lead to smoother but flatter results. The effect of these parameters is shown in the supplemental materials.

Finally, to composite the filtered region into the image, we smooth the region mask according to Eq. (5), obtaining $M_{\text{bf},i}$, to further emphasize the specular profile of the region and normalize it between 0 and 1. The final corrected result uses the normalized $M_{\text{bf},i}$ as a weight and is computed as:

$$R'_{\text{ov},i} = M_{\text{bf},i}R_{\text{bf},i} + (1-M_{\text{bf},i})R_{\text{in},i}. \tag{6}$$

4. Psychophysical Evaluation

To evaluate whether our algorithm improves the visual quality of images and to assess its performance against other methods, we have performed a psychophysical experiment. We based our experiment on the 'Subjective Assessment of Multimedia Video Quality' (SAMVIQ) protocol, recommended in ITU-R BT.1788 for assessing video quality in multimedia applications [ITU12].

4.1. Experimental Protocol

The goal of this experiment is to assess (a) the relative performance of our method compared to alternative solutions, (b) the degree of improvement that our method affords relative to the uncorrected, clipped input, and (c) the quality of our results compared to the ground truth, unclipped image.

To achieve this, participants were shown images of the same scene processed in 5 different ways, as well as the reference unclipped image, and were asked to rate *how similar* each image was to the corresponding reference. The images shown included the clipped, uncorrected LDR input (Input), the ground truth HDR image (Ref), which was the same as the reference, the result of our method (Ours) as well as results from Masood et al.'s method produced using their source code (Masood) and Guo et al.'s method processed by the authors themselves (Guo).

Procedure. For a given scene, the participant first viewed the reference HDR image and then chose one of the five alternatives using a specially labeled keyboard. Each alternative was scored on a scale from 1 (not similar) to 5 (very similar). Participants were allowed to view the reference image whenever desired and could alter the scores for each alternative within a series, as described by SAMVIQ. Once all scores were entered for a given scene, the participant saved their decisions and proceeded to the next scene. A gray screen was shown between any alternatives or the reference.

Images. We used a set of 15 HDR images as our ground truth and extracted LDR images from each as described in Section 5. The LDR images were processed with our method as well as those of Masood et al. [MZT09] and Guo et al. [GCZS10]. Images were scaled according to the HDR ground truth, such that areas that were not clipped would obtain the same values (all images are shown in the supplemental materials). As the input HDR images were not given in absolute radiometric values but rather relative linear values, all HDR images were initially scaled to a range of 0 to 2000, to match the peak luminance of the HDR display used.

Setup. A SIM2 HDR© display was used for showing the stimuli (42' diagonal size, HD resolution, peak luminance limited to 2000 cd/m^2). Viewing distance was approximately 1.5m and participants were free to adjust their chair and move their head so that they were approximately centered on the display. The resolution of the stimuli was limited to a square of 600x600 pixels, so that only the central portion of the display was used to minimize directional viewing issues. **Participants.** A total of 23 participants (18 M/5 F), took part in the experiment, all reporting normal or

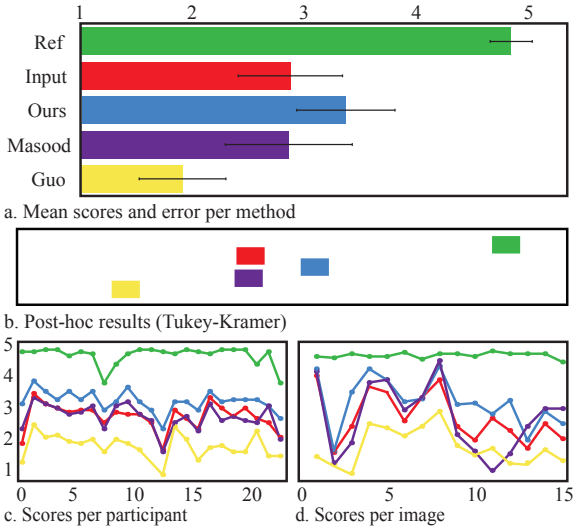


Figure 6: *Experiment results over all images and participants. (a) Mean scores for each method and corresponding errors. A score of 1 means 'not similar' while a score of 5 means 'very similar'. (b) Post-hoc analysis using the Tukey-Kramer test. (c,d) Average scores for each participant and image respectively.*

corrected to normal vision. Participant ages ranged from 23 to 55 years old (μ : 39.09, σ : 10.28).

4.2. Results & Analysis

In total, each participant saw 75 images, totalling 1725 trials. Figure 6(a) shows the mean scores and corresponding errors for each method pooled across all participants. These results were analyzed for significance using one-way ANOVA (ANalysis Of VAriance), showing that our method was deemed to be significantly closer to the HDR reference compared to the LDR input or either of the other methods ($F(4,1720) = 363.1$, $p < 0.001$). Masood et al. [MZT09] was scored similar to the uncorrected input images ($F(1,688) = 0.023$, $p = 0.88$). Individual comparisons of our method with the two other techniques also indicate that results were significant, as summarized in Table 1. Post-hoc analysis was also performed using the Tukey-Kramer test, shown in Figure 6(b), further supporting the conclusion that our method leads to a significant improvement over the uncorrected input and the two methods tested.

Results were consistent across participants in general, but indicating overall differences in scaling, as can be seen in Figure 6(c). Significant differences in scores were found across images, ($F(14,1710) = 39.793$, $p < 0.001$). However upon further analysis only one image was found to have significantly different behavior to all others, with a much lower score across all methods, likely due to the presence of large fully clipped areas (see ACES image in supplementals). Interestingly, the mean score for our method was equal or higher to that of the clipped LDR input for all images, indicating that our method never reduced the quality of the image. Mean scores per image for each method are shown in Figure 6(d).

Comparison	F	p-value
Overall	$F(4,1720) = 363.130$	$p < 0.001$
Ours vs Ref	$F(1,688) = 552.515$	$p < 0.001$
Ours vs Input	$F(1,688) = 34.610$	$p < 0.001$
Ours vs Masood	$F(1,688) = 29.246$	$p < 0.001$
Ours vs Guo	$F(1,688) = 358.560$	$p < 0.001$
Masood vs Guo	$F(1,688) = 109.914$	$p < 0.001$
Masood vs Input	$F(1,688) = 0.023$	$p = 0.88$

Table 1: *One-way ANOVA results comparing the different methods.*

5. Results

In the previous section, we presented a psychophysical study comparing our method with related techniques and the ground truth. Here, we show our results for a variety of images and present visual and quantitative comparisons.

Ground Truth Dataset. In an ideal case, the information reconstructed by our method should match what was in the scene, so to evaluate the quality of our results, we construct a ground truth dataset using HDR images. Given a reference HDR image I_{hdr}, we create a stack of simulated LDR exposures and select the one that minimizes over and under-exposed pixels (i.e. pixels above the 95th or below the 5th percentile). This LDR exposure, referred to as I_{clip}, is then gamma corrected and converted to 8-bit range to serve as input to the different methods tested. Since we know the maximum value s_{clip} of the reference scene encoded within I_{clip}, both the LDR image and the output of each correction method can be re-scaled according to the HDR reference range to enable comparisons. This approach was used to create the input images for our experiment and the ground truth evaluation described below.

Visual Comparison. We have processed several images using the proposed method and two other closely related saturation correction methods, namely the approach of Guo et al. [GCZS10] and Masood et al. [MZT09]. The authors' implementation was used for the method of Masood et al. Results for the method of Guo et al. were kindly produced by the authors using images we provided. Note however that the latter method includes a non-linear tonemapping step and as such their results are only qualitatively comparable with the ground truth or to the other methods. We have applied the same tone mapping operator to all other results with the same parameters, as mentioned in [GCZS10] to provide fair comparisons with either method.

Figure 7 shows detailed results produced by our method

for several images and their corresponding inputs. Our method recovers color and detail in clipped areas (e.g. see the yellow flower and grass in Figure 7), while enhancing the appearance of specularities and light sources (e.g. water and car highlights in Figure 7). Note also that our method preserves skin appearance (e.g. see second column of Figure 7).

A visual comparison between our method, as well as the alternative methods discussed, is shown in Figure 11. In this case, HDR images are used as reference and all images are scaled with respect to the HDR values. For visualization, the tone mapping operator of Fattal et al. [FLW02] was used with parameters set according to [GCZS10]. Due to our hue and distance-based grouping of clipped regions, our method avoids wrong color propagations from adjacent regions observed with both other methods (e.g. see the sky in Figure 11).

Finally, visual differences were quantitatively assessed using HDRVDP2 [MKRH11], showing that our correction reduces distortions with respect to the reference image (see rows 2 and 4 of Figure 11). The average HDRVDP2 quality over our image set was 93.9 for the clipped input images, 96.7 for our method, 94.9 for Masood et al. and 92.0 for Guo et al., withh 100 representing a perfect match.

6. Applications

In the previous sections we demonstrated the use of our correction method in a variety of still images. Here, we briefly explore some additional applications of our technique.

6.1. Tone Reproduction

Our correction method recovers clipped and over-exposed details, and as shown in Sections 4 and 5 improves the visual quality of images and videos compared to their clipped counterparts. We briefly explore different applications of our method in the context of tone reproduction. The correction applied by our method locally extends the dynamic range of images. To visualize the added information, it is possible to combine our method with any existing tone mapping operator (TMO) [RHD*10] to locally increase detail without expanding the range of the image. Figure 8 shows example results using different operators. See also Figure 7 for linear compression examples and Figure 11 for more examples using the TMO by Fattal et al. [FLW02]. At the same time, our correction can serve as a pre-processing step when the goal is to significantly expand the dynamic range of the image, by combining it with inverse tone mapping operators (ITMO) [YHB14, MFSG10, BLD*07]. Several ITMOs aim to reconstruct details in clipped areas, as discussed in Sec. 2, but their focus is on brightness enhancement only, ignoring color clipping. We have applied our method in conjunction with different ITMOs to produce HDR images. Please refer to the supplemental materials for HDR images produced in this way.

6.2. Video Extension

In Section 3 we described our correction method for single images. However, in order to correct clipping in videos, additional considerations are necessary. Although a full video solution is not the main focus of this work, we propose a simple extension to our single image approach to handle video content. Both the segmentation and the correction steps described in the previous section can potentially lead to temporal discontinuities if applied directly to video content. The local variation copied into clipped regions from the unclipped reference channel may lead to differences in global statistics in consecutive frames, potentially causing flickering. Additionally, the baseline intensity ρ_i of a region i can vary between consecutive frames as the object moves, due to changes in lighting or occlusions.

The first issue can simply be addressed through temporal smoothing of the maximum intensity value of each frame by re-normalizing a given frame with a combination of the current and previous maximum. To address the latter issue, we propose a simple extension to the region grouping and baseline computation process to the temporal domain, through a 2-frame temporal window. To compute the baseline ρ_i for a given region i in a frame f in this context, we also consider the regions in the previous frame, $f-1$. The key idea is that regions corresponding to the same clipped object are likely to overlap between consecutive frames. As such, instead of actively tracking the motion of regions or objects, we rely on this overlap to compute the baseline intensity for the current frame.

For each region $R_{i,f}$, we first compute its baseline intensity $\rho_{i,f}$ as described in Section 3.2. Then we detect overlapping regions in $f-1$ and compute an average baseline intensity $\overline{\rho}_{i,f-1}$ from them. This is in turn used to compute the final baseline intensity for the region as follows:

$$\rho'_{i,j} = \alpha\rho_{i,f} + (1-\alpha)\overline{\rho}_{i,f-1}. \quad (7)$$

Here, α determines how adaptive we are to changes in the video and is set to 0.2. If there is no region in the previous frame which spatially overlaps with a region in the current frame, then the region's baseline intensity $\rho_{i,f}$ value will be used directly for the correction. Example video results using this temporal smoothing solution are given in the supplemental materials.

7. Discussion

In our evaluation we have observed that different causes of over-exposure can lead to different behavior of the corresponding regions in the image. Consequently, a single set of parameters for the different steps of our method is not ideal for all images. Specifically, the parameters used for over-exposure correction, described in Section 3.3 are well-suited to over-exposure due to specular reflections, light sources or for 'soft' objects and surfaces such as clouds or water. Diffuse surfaces with well-defined geometry on the other hand,

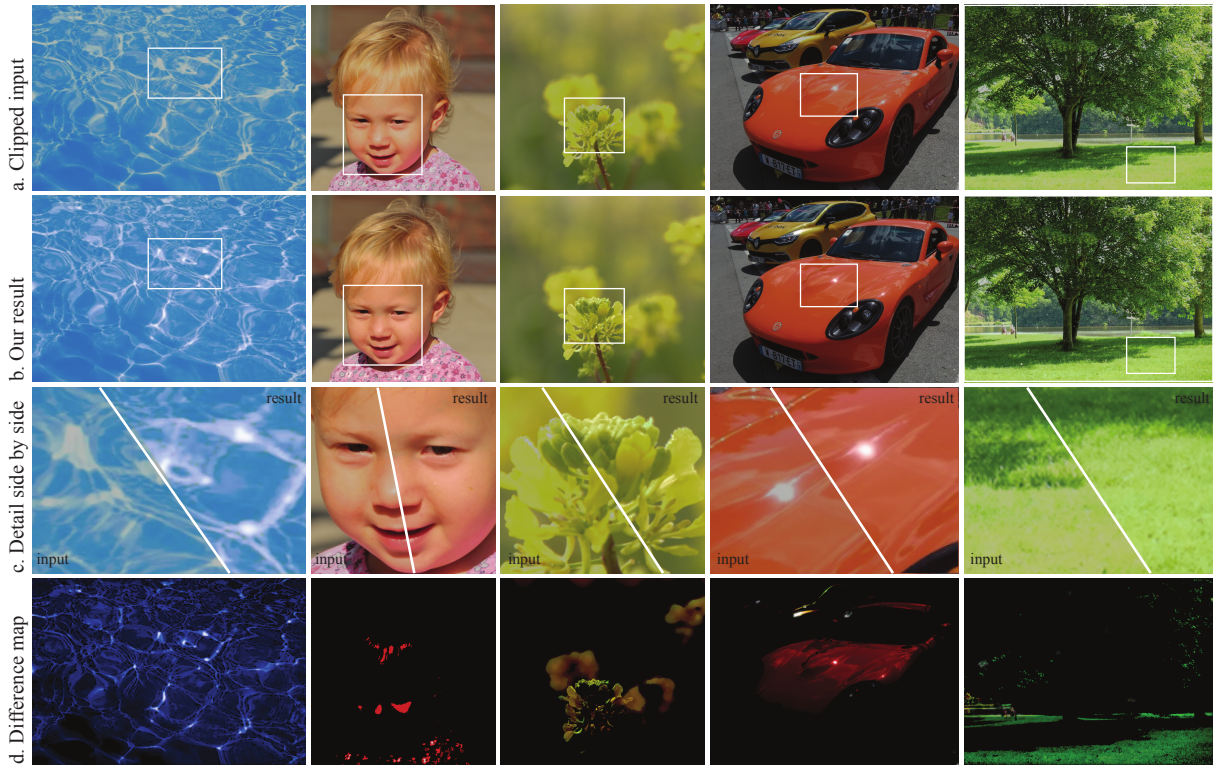


Figure 7: *Example results on still images using our method. Detailed views of the clipped input (a) and the corrected result (b) are shown in (c), normalized difference maps between the input and output are shown in (d) to illustrate the effect of the correction.*

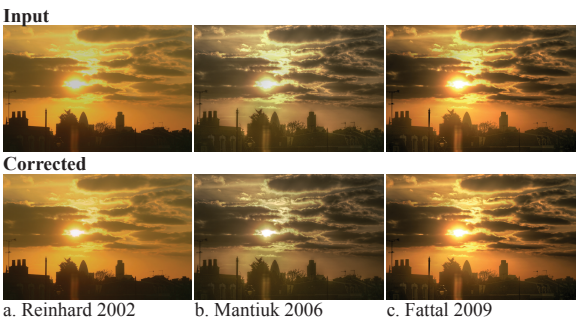


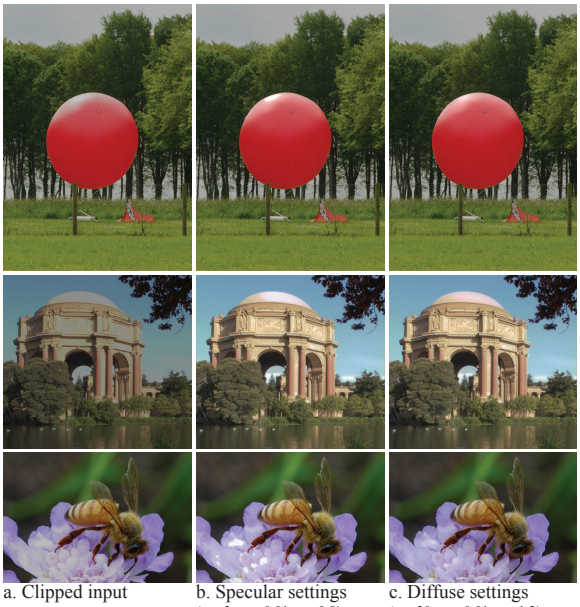
Figure 8: *The clipped input and our result compressed with different TMOs ([RSSF02, MMS06, Fat09]) to achieve different visualization styles.*

particularly in the case of textured surfaces, require more aggressive smoothing since there is no highlight that should be preserved.

Figure 9 illustrates some example images of this type, corrected both with our default parameters and with an alternative set of parameters better suited to this type of over-exposure. Currently, detection of diffuse clipping cases is manual, however better detection and classification of clipped regions is an interesting problem in itself, meriting future work.

In some images, over-exposure can be severe, with large areas having no information. Our method cannot recover details and color information in such over-exposed regions — at best, it can partially inpaint and enhance the brightness of those areas in a way consistent to the rest of the corrections as described in Section 3.3. Combining our approach with more advanced inpainting and texture synthesis solutions that are able to reconstruct texture in over-exposed areas would help recover some of this lost information. However, in this paper, we have opted for a simpler solution to improve robustness and to allow for interactive implementations.

In Section 6.2 we showed a simple extension of our approach for handling video content. Although we found this to perform well in the videos tested (see supplemental materials), to ensure robustness in video processing, particularly with rapidly varying content, a motion estimation step would be necessary to accurately track the movement of different



a. Clipped input
b. Specular settings ($\kappa = 3, c = 0.01, \gamma = 0.8$)
c. Diffuse settings ($\kappa = 30, c = 0.01, \gamma = 1.2$)

Figure 9: *Diffuse surfaces require more aggressive smoothing when correcting over-exposure. The settings used for specularities and highlights lead to still visible discontinuities. We propose a set of diffuse settings for such images.*

clipped regions. We believe this is an interesting avenue for further work.

Implementation. We used a non-optimized MATLAB implementation of the proposed approach to produce the results in the paper. However, we have also tested a C implementation of our method, which is integrated as a plugin within Autodesk's Lustre color grading system. The plugin processes HD resolution images in 3 to 4 seconds on the Lustre system, allowing for interactive control of the parameters of the method. Additionally, this enables our method to be combined with existing color grading tools, and therefore to be employed in a professional post-production environment. Figure 10 shows the color grading setup and plugin.

8. Conclusions

Recent trends towards brighter, more colorful content and displays have increased the need for robust and efficient solutions to 'upgrade' legacy content to these specifications. Color clipping and over-exposure correction is a crucial step in this process and forms a particularly challenging problem, since reconstructed information needs to be coherent across the image, with often insufficient information. To that end, we have proposed a novel color clipping and over-exposure correction method, which relies on the correlation existing between RGB channels to recover clipped information from unclipped channels. To ensure spatial consistency

Figure 10: *Our method is implemented as a color grading plugin, allowing for interactive modification of parameters and fast visualization.*

across similar regions, we employ a region grouping scheme considering both distance and color similarity between regions. To correct over-exposed areas where no information is present, we explore ideas similar to inverse tone mapping, locally reshaping their brightness profile and smoothly integrating them with the color clipping corrections.

The validation of our solution in comparison to HDR ground truth and related methods is twofold: a rigorous psychophysical experiment following the ITU recommended SAMVIQ protocol and the use of HDR-VDP2, providing predicted quality scores and a map showing spatial error distributions. Our evaluation showed that our method outperforms previous similar solutions and can improve the visual quality of the clipped input. Beyond the scope of the approach, we demonstrated the usefulness of our approach through an interactive color grading implementation as well as a simple video extension. Our work can also be naturally combined with further tone reproduction steps, preparing content for future displays and getting us a step closer to recovering the full one thousand words.

References

[APKL14] ASSEFA M., POULI T., KERVEC J., LARABI M.: Correction of over-exposure using color channel correlations. In *GlobalSIP: Perception Inspired Multimedia Signal Processing Techniques* (2014). 2

[ASI91] ABEL J. S., SMITH III J. O.: Restoring a clipped signal. In *Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference on* (1991), IEEE, pp. 1745–1748. 1

[BLD*07] BANTERLE F., LEDDA P., DEBATTISTA K., CHALMERS A., BLOJ M.: A framework for inverse tone mapping. *The Visual Computer 23*, 7 (2007), 467–478. 2, 8

[DMHS08] DIDYK P., MANTIUK R., HEIN M., SEIDEL H.-P.: Enhancement of bright video features for HDR displays. In *EGSR* (2008). 2

[EW10] ELBOHER E., WERMAN M.: Recovering color and details of clipped image regions. *Proc. CGVCVIP* (2010). 3

[Fat09] FATTAL R.: Edge-avoiding wavelets and their applications. *ACM Transactions on Graphics (SIGGRAPH '09) 28*, 3 (2009), 22. 9

[FLW02] FATTAL R., LISCHINSKI D., WERMAN M.: Gradient domain high dynamic range compression. *ACM Transactions on Graphics 21*, 3 (2002), 249–256. 2, 8, 12

[FPCM13] FU J., PENG H., CHEN X., MOU X.: Correcting saturated pixels in images based on human visual characteristics. In *IS&T/SPIE Electronic Imaging* (2013), International Society for Optics and Photonics, pp. 866009–866009. 2

[Gar10] GARCIA D.: Robust smoothing of gridded data in one and higher dimensions with missing values. *Computational statistics & data analysis 54*, 4 (2010), 1167–1178. 5

[GCZS10] GUO D., CHENG Y., ZHUO S., SIM T.: Correcting over-exposure in photographs. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on* (2010), IEEE, pp. 515–521. 2, 6, 7, 8, 12

[Has13] HASINOFF S. W.: Saturation (imaging). In *Encyclopedia of Computer Vision*. Springer, 2013. 1

[HJS13] HOU L., JI H., SHEN Z.: Recovering over-/underexposed regions in photographs. *SIAM J. Imaging Sciences 6*, 4 (2013), 2213–2235. 2

[HS11] HIRAKAWA K., SIMON P.: Single-shot high dynamic range imaging with conventional camera hardware. In *Computer Vision (ICCV), 2011 IEEE International Conference on* (Nov 2011). 2

[ITU12] ITU-R: Recommendation ITU-R BT.500-13: Methodology for the subjective assessment of the quality of television pictures. 2, 6

[JMKA10] JOSHI N., MATUSIK W., KRIEGMAN D., ADELSON E. H.: Personal photo enhancement using example images. *ACM Transactions on Graphics* (March 2010). 2

[MFSG10] MASIA B., FLEMING R., SORKINE O., GUTIERREZ D.: Selective reverse tone mapping. *Proc. CEIG* (2010). 2, 8

[MKRH11] MANTIUK R., KIM K. J., REMPEL A. G., HEIDRICH W.: HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. *ACM Trans. Graph. 30*, 4 (July 2011), 40:1–40:14. 2, 8, 12

[MMS06] MANTIUK R., MYSZKOWSKI K., SEIDEL H.: A perceptual framework for contrast processing of high dynamic range images. *ACM Transactions on Appliced Perception 3*, 3 (2006), 286–308. 9

[MS06] MEYLAN L., SUSSTRUNK S.: High dynamic range image rendering with a retinex-based adaptive filter. *IEEE Transactions on Image Processing 15*, 9 (2006), 2820 –2830. 2

[MSNW10] MANSOUR H., SAAB R., NASIOPOULOS P., WARD R.: Color image desaturation using sparse reconstruction. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on* (2010), IEEE, pp. 778–781. 3

[MZT09] MASOOD S. Z., ZHU J., TAPPEN M. F.: Automatic correction of saturated regions in photographs using cross-channel correlation. In *Computer Graphics Forum* (2009), vol. 28, pp. 1861–1869. 2, 3, 6, 7, 12

[RHD*10] REINHARD E., HEIDRICH W., DEBEVEC P., PATTANAIK S., WARD G., MYSZKOWSKI K.: *High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting*, 2nd ed. Morgan Kaufmann, 2010. 8

[RLH12] ROUF M., LAU C., HEIDRICH W.: Gradient domain color restoration of clipped highlights. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on* (2012), IEEE, pp. 7–14. 2

[RP11] REINHARD E., POULI T.: Colour spaces for colour transfer. In *IAPR Computational Color Imaging Workshop*, vol. 6626 of *Lecture Notes in Computer Science*. 2011, pp. 1–15. 2

[RSSF02] REINHARD E., STARK M., SHIRLEY P., FERWERDA J.: Photographic tone reproduction for digital images. *ACM Transactions on Graphics 21*, 3 (2002), 267–276. 9

[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. *ACM Transactions on Graphics 26*, 3 (2007), 39. 2

[SMZ*11] SHEN Y., MO R., ZHU Y., WEI L., GAO W., PENG Z.: Over-exposure image correction with automatic texture synthesis. In *Image and Signal Processing (CISP), 2011 4th International Congress on* (2011), vol. 2, IEEE, pp. 794–797. 2

[TM98] TOMASI C., MANDUCHI R.: Bilateral filtering for gray and color images. In *ICCV '98: Proceedings of the 1998 IEEE International Conference on Computer Vision* (Washington, DC, USA, 1998), IEEE Computer Society, pp. 839–846. 5

[WIH10] WETZSTEIN G., IHRKE I., HEIDRICH W.: Sensor saturation in fourier multiplexed imaging. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on* (2010), IEEE, pp. 545–552. 1

[WWZ*07] WANG L., WEI L.-Y., ZHOU K., GUO B., SHUM H.-Y.: High dynamic range image hallucination. In *Proceedings of the 18th Eurographics conference on Rendering Techniques* (2007), Eurographics Association, pp. 321–326. 2

[XDN11] XU D., DOUTRE C., NASIOPOULOS P.: Correction of clipped pixels in color images. *Visualization and Computer Graphics, IEEE Transactions on 17*, 3 (2011), 333–344. 2

[YHB14] YONGQING HUO F. Y., BROST V.: Inverse tone mapping based upon retina response. *The Scientific World Journal* (2014). 8

[ZB04] ZHANG X., BRAINARD D. H.: Estimation of saturated pixel values in digital color imaging. *JOSA A 21*, 12 (2004), 2301–2310. 2

[ZGW*14] ZHANG C., GAO J., WANG O., GEORGEL P., YANG R., DAVIS J., FRAHM J., POLLEFEYS M.: Personal photograph enhancement using internet photo collections. *IEEE Trans. Vis. Comput. Graph. 20*, 2 (2014), 262–275. 2

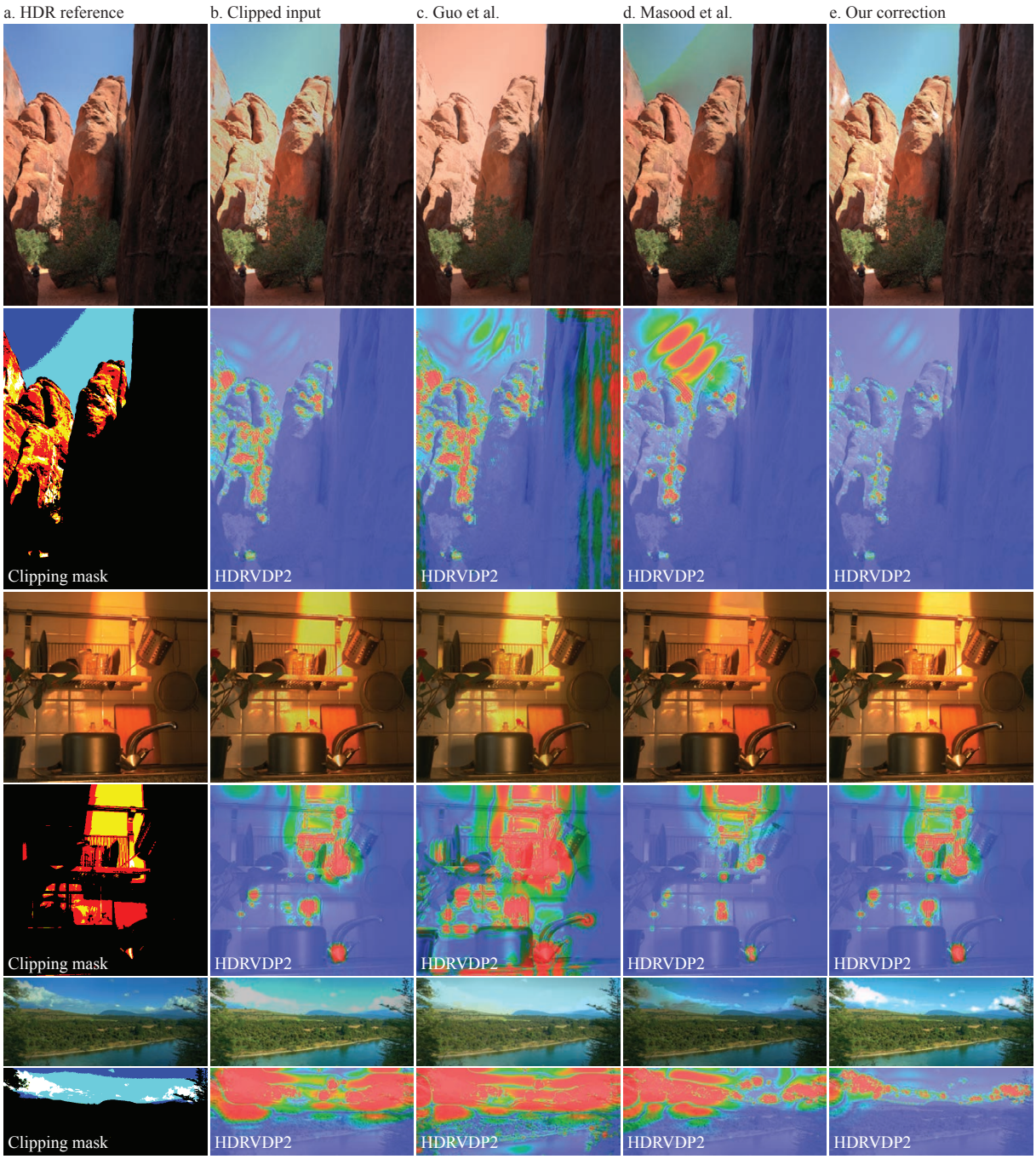


Figure 11: *Comparison between the HDR reference (a), extracted LDR clipped image (b), Guo et al. [GCZS10] (c), Masood et al. [MZT09] (d), and our results (e). Maps illustrating the differences between each image and the HDR reference using the HDRVDP2 [MKRH11] are shown in the 2nd and 4th row. The first image in these rows shows a map of clipped pixels per channel. All images were tonemapped for visualization using the TMO by Fattal et al. [FLW02].*