Real-time noise aware tone mapping

Supplementary material

Gabriel Eilertsen
Linköping University, Sweden

Rafał K. Mantiuk
Bangor University, UK &
The Computer Laboratory,
University of Cambridge, UK

Jonas Unger
Linköping University, Sweden

1 Introduction

In this supplemental material we present some of the details that did not fit in the main manuscript. In Section 2 we describe our method to compare the proposed tone mapping method to previous work, were a qualitative evaluation is executed. To motivate the number of iterations needed for the detail extraction filter presented in the paper, Section 3 describes an experiment to motivate this selection. Finally, in Section 4 we perform an analysis of the tone reproduction over time, and how the different parts of the tone mapping affect the final outcome.

2 Subjective evaluation

To evaluate the visual quality of our tone mapping approach, we compared it to six current state-of-the-art video tone mapping operators:

- Mal-adaptation TMO, [Irawan et al. 2005]
- Display-adaptive TMO, [Mantiuk et al. 2008]
- Virtual exposures TMO, [Bennett and McMillan 2005]
- Temporal coherence TMO, [Boitard et al. 2012]
- Zonal temporal coherence TMO, [Boitard et al. 2014]
- Motion path filtering TMO, [Aydin et al. 2014]

The selected operators are: the two best performing TMOs (Mal-adaptation and Display adaptive), and the best performing local TMO (Temporal coherence), from the evaluation in [Eilertsen et al. 2013]. We also included the Virtual exposures TMO since it is conceptually similar to our approach. Finally, we included the Zonal temporal coherence and Motion path filtering TMOs as they represent the state-of-the-art in the field and were recently published. For the comparison we used four sequences from [Froehlich et al. 2014]. Gamma mapped example frames from the sequences, together with histograms calculated over all the frames, are shown in Figure 2. The results of applying the different operators on these sequences are demonstrated in Figure 3, 4, 5 and 6. For demonstrations of the tone mappings over time, we refer to the supplemental video.

In Figure 3, the magnified parts of the images show examples of the tone mappings in low and high luminance areas, respectively, in order to highlight differences of the operators. The previous TMOs either show problems in compressing the dynamic range, or in reproducing details without artifacts. In contrast to this, our local tone curves successfully compress the dynamic range without sacrificing contrast. Also, using our approach image details are rendered with no visible artifacts.

Qualitative analysis In order to evaluate the performance of our TMO and how well we addressed the different challenges in tone mapping (image noise, large contrasts, ringing, ghosting, temporal flickering, and display adaptivity), we also conducted a qualitative analysis using the TMOs listed above. The TMOs were evaluated in a rating experiment using the following attributes: overall *brightness*, overall *contrast*, overall color *saturation*, temporal color *consistency*, temporal *flickering*, *ghosting* and excessive *noise*. In addition to these, we also added an attribute for *detail* reproduction to assess local contrast in the tone mapped sequences.

The experiment was carried out in a controlled environment: all clips were viewed on a 24" 1920×1200 colorimetric LCD (Nec PA241W), with a peak luminance of 200 cd/m^2. The observers were placed at approximately 3 display heights (97 cm) distance from the display, in a dim room.

The tone mapped sequences were displayed in a random order, and 10 observers experienced in digital imaging rated the visibility/level of the image attributes for each operator and sequence. The final result of the ratings, averaged over the observers, is illustrated in Figure 1. Overall it is clear that the proposed noise aware TMO consistently produces results that show image characteristic at about just the right level without visible artifacts.

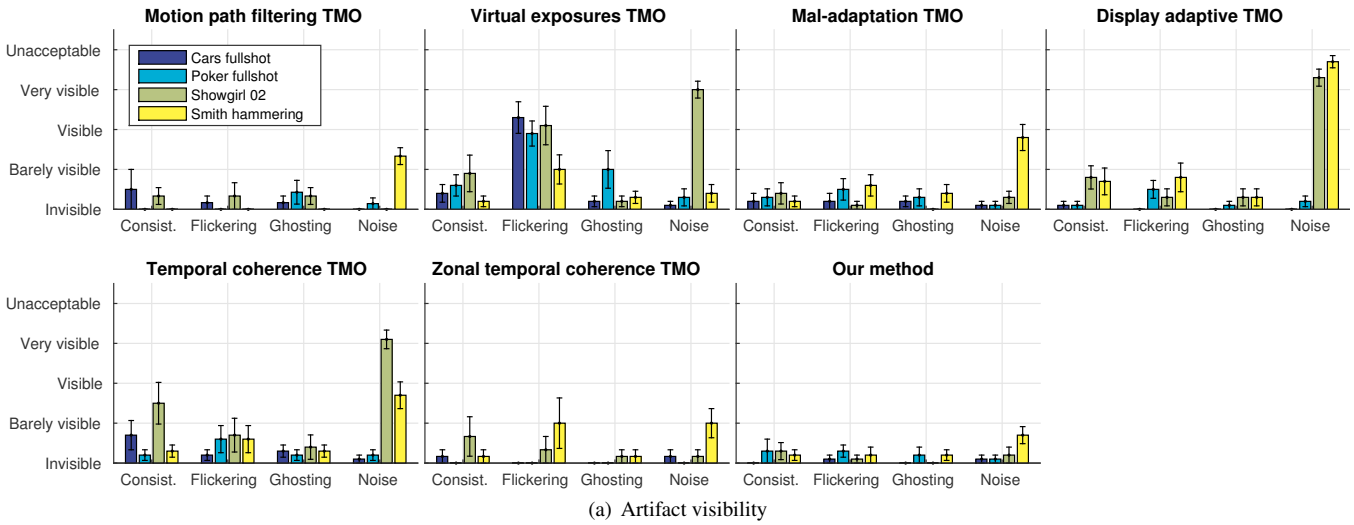
In Figure 1, the most problematic sequence, showing highest rating of a single artifact for our TMO, is the *Smith hammering* video. The sequence exhibits a high level of noise, which is very difficult to completely hide in this dark scene. However, the only operator that produces a result with noise less visible is the Virtual exposures TMO, which essentially is targeted at noise reduction and performs a computationally expensive filtering over time.

Discussion Even though some previous methods show good results in the qualitative analysis, conceptual differences and details difficult to evaluate in this way still motivate the need for the real-time and noise-aware tone mapping approach. Below, we summarize the analysis, to put our TMO into context to the other operators, and present a short discussion on the features of each operator:

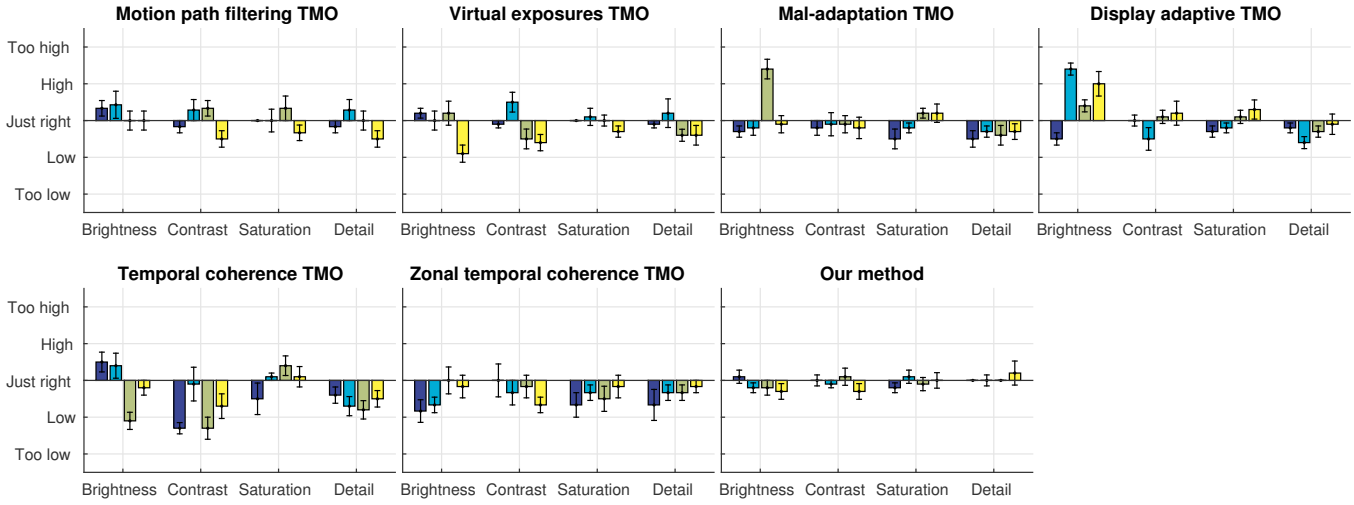
The **Mal-adaptation TMO** [Irawan et al. 2005] produced results with overall small amounts of artifacts. However, utilizing a global tone curve, there are inevitable problems with detail reproduction and maintaining contrast for compression of footage exhibiting a large dynamic range.

The **Display-adaptive TMO** [Mantiuk et al. 2008] relies on a global tone curve and have problems in maintaining the contrast if the dynamic range in the input is too large. This is the only one of the previous methods which adapts to the viewing environment. However, the tone curve calculations are relatively expensive, and real-time performance is difficult.

The **Virtual exposures TMO** [Bennett and McMillan 2005] uses a bilateral filter to preserve or enhance details. However, as demonstrated in the paper, this may generate visible artifacts in some situations, especially for detail enhancements. There are also problems in adapting the tone curve over time, resulting in disturbing flickering artifacts. Since a bilateral filtering operates also in the temporal domain, the method is time-consuming and not suited for real-time applications.



(a) Artifact visibility



(b) Image characteristics

Figure 1: *Qualitative analysis result, showing the average ratings from the conducted experiment, with error bars for the standard errors.*

The **Temporal coherence TMO** [Boitard et al. 2012] ensures temporal brightness coherency for arbitrary TMOs in a pre-processing step. However this comes at the cost of reduced contrast and time-consuming pre-processing.

The **Zonal temporal coherence TMO**, [Boitard et al. 2014] uses local temporal coherence calculations to better preserve the overall contrast compared to the Temporal coherence TMO. However, extensive pre-processing is still required, making the method unsuitable *e.g.* for online applications.

The **Motion path filtering TMO** [Aydin et al. 2014] is essentially a filtering algorithm for tone mapping. In terms of the qualitative analysis, the method produces results without visible artifacts and manage to reproduce image characteristic quite well. However, the method does not consider the actual tone reproduction (tone curve) and temporal behavior of the tone curve, but relies on previous TMOs. As demonstrated in the main paper, the spatial filter in some cases produces artifacts which may become disturbing if detail manipulation is required. Also, the optical flow filtering may not work well in complicated image transitions (see *e.g.* the noise level in the *smith hammering* sequence in Figure 1). Finally, the method requires a time-consuming filtering process, i.e. the method is computationally expensive and not suited for real-time applications.

In contrast to these methods, our video tone reproduction can handle large variations in dynamic range due to the local tone curves used. Even for extreme dynamic range compression, details are effectively preserved or enhanced using the stable spatial filtering approach, which is specifically tailored for use in the context of tone mapping. Furthermore, noise is effectively hidden with the noise-aware capabilities of the tone curves, and the tone curves adapt to the specific display and viewing environment. Finally, all computations run in frame rates above real-time, even for high resolution (1080p) material, making it suitable for real-time and online applications.

3 Convergence analysis of detail extraction

To estimate and motivate the appropriate number of iterations for the diffusion simplification presented in the paper, we here provide a numerical demonstration on the convergence rate of the fast de-

(a) Cars (b) Poker

(c) Showgirl (d) Smith hammering

(e) Cars (f) Poker (g) Showgirl (h) Smith hammering

Figure 2: *HDR video sequences used in the qualitative evaluation (a-d), together with histograms calculated over all frames in the sequences (e-h), to show the spatial as well as temporal dynamic range. Sequences from [Froehlich et al. 2014].*

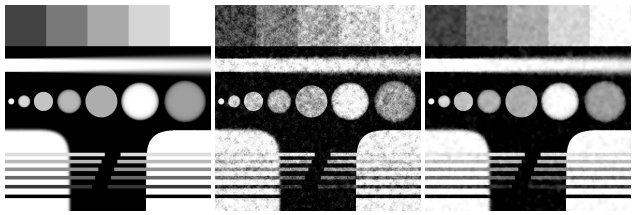
tail extraction diffusion filter. This is carried out using an artificial 2D signal with a range of different edge properties, see Figure 7(a). To this signal we add details, Figure 7(b). The details are represented by normally distributed noise, at three different scales; per-pixel varying, and with size corresponding to approximately 3% and 10% of the image width. This should by no means be considered a case generalizing to all situations, since natural images may show different statistics of the image details, and since the quality of the detail extraction in the end is a subjective measure. We rather present this to give an indication on performance and the best choice of diffusion iterations.

In Figure 7(d) the mean squared error of the detail extraction is plotted as a function of the number of diffusion iterations. The three curves show the result for anisotropic diffusion, diffusion with locally isotropic behavior, and the final approximation used, with gradient formulation and kernel sizes described above. For the anisotropic diffusion and its isotropic modification the diffusion rate (α in the diffusion equation) is decreased with increasing iterations, so that the only difference with increasing iterations is the distance of the diffusion. For the full detail extraction diffusion filter, the size of the local neighborhood is chosen such that the same total filter size is achieved, that is, σN is the same for all N.

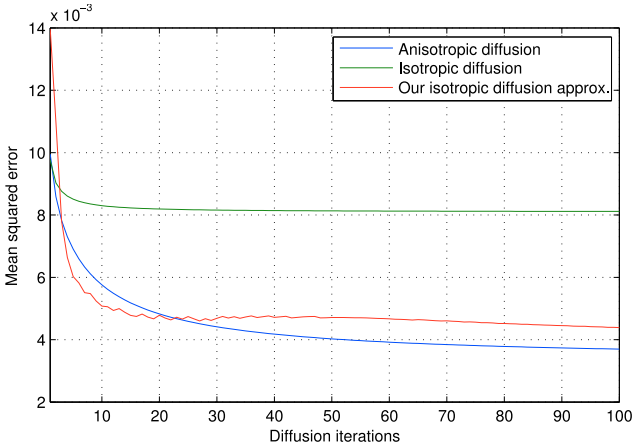
As can be seen, the error drastically decrease at first, and then flattens out. The optimal ratio between performance and approximation lies somewhere around 10-15 iterations. This varies slightly with the input material and the parameter settings used. However, about 10-15 iterations usually is a good choice, and we fix this parameter to 12 for all examples we show.

4 Behavior of temporal tone curve

To highlight the characteristics of the tone curves, and how they adapt over time, we analyze the tone mapping over time in a scene with changing lighting conditions. The sequence used in this demonstration, show a static scene where a light bulb is switch on and off, Figure 8(a-b). The output of the TMO at two pixel positions



(a) Signal (b) Signal + details (c) Filtered (Our)



(d) Detail extraction error

Figure 7: *Mean squared error for different number of diffusion iterations used. The jaggedness of the isotropic diffusion approximation is due to the discrete gradient evaluations for the edge-stop function.*

are plotted in Figure 8(c-d), showing the effect of using the different features of the tone curve. The extracted details are strongly enhanced before being added back to the tone mapped image in order to highlight the influence of the noise-dependent local contrast control.

Adding the noise-aware capabilities, the tone mapping output at both pixels show reduced temporal fluctuations. Also the tone reproduction is shifted towards lower intensities to better conceal the noise.

Using local tone curves, the dynamic range can be further reduced. This is apparent analyzing the brighter pixel, for which the tone curve adapts to map it to a darker tone. The effect is that the intensity of the two pixels are at about the same level after the local curves have been applied.

Finally, utilizing a temporal edge-stopping filter there is a rapid transition between the different lighting conditions, preventing the overshoot created by the IIR low-pass filter.

References

AYDIN, T. O., STEFANOSKI, N., CROCI, S., GROSS, M., AND SMOLIC, A. 2014. Temporally coherent local tone mapping of HDR video. *ACM Trans. Graphics 33*, 6, 1–13.

BENNETT, E. P., AND MCMILLAN, L. 2005. Video enhancement using per-pixel virtual exposures. *ACM Trans. Graphics 24*, 3, 845–852.

BOITARD, R., BOUATOUCH, K., COZOT, R., THOREAU, D., AND GRUSON, A. 2012. Temporal coherency for video tone mapping. In *Proc. SPIE 8499, Applications of Digital Image Processing XXXV*, 84990D–84990D–10.

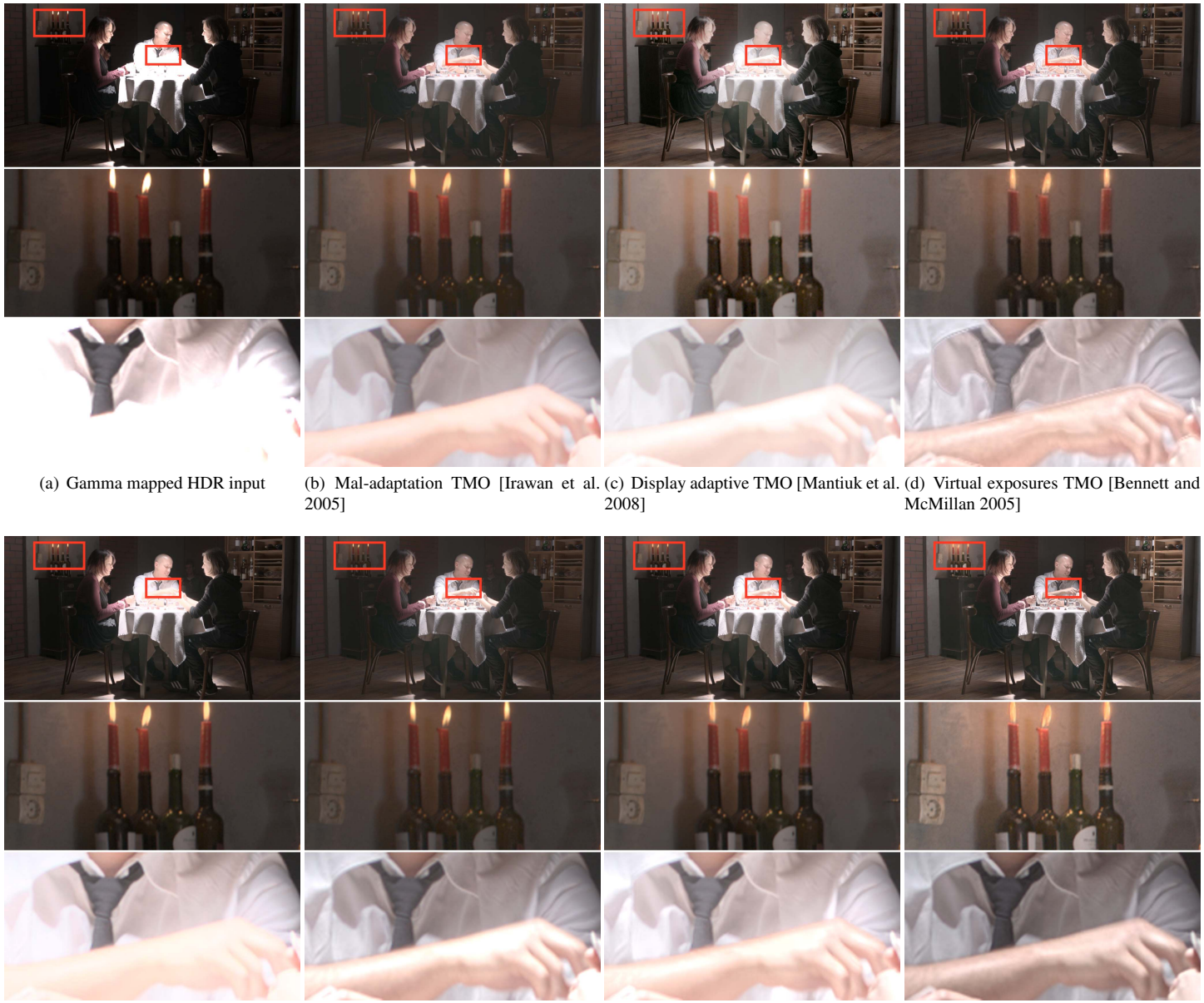
BOITARD, R., COZOT, R., THOREAU, D., AND BOUATOUCH, K. 2014. Zonal brightness coherency for video tone mapping. *Signal Processing: Image Communication 29*, 2, 229–246.

EILERTSEN, G., WANAT, R., MANTIUK, R. K., AND UNGER, J. 2013. Evaluation of Tone Mapping Operators for HDR-Video. *Computer Graphics Forum 32*, 7, 275–284.

FROEHLICH, J., GRANDINETTI, S., EBERHARDT, B., WALTER, S., SCHILLING, A., AND BRENDEL, H. 2014. Creating Cinematic Wide Gamut HDR-Video for the Evaluation of Tone Mapping Operators and HDR-Displays. In *Proc. SPIE 9023, Digital Photography X*, 90230X–90230X–10.

IRAWAN, P., FERWERDA, J. A., AND MARSCHNER, S. R. 2005. Perceptually based tone mapping of high dynamic range image streams. In *Proc. Eurographics Conference on Rendering Techniques 16*, 231–242.

MANTIUK, R., DALY, S., AND KEROFSKY, L. 2008. Display adaptive tone mapping. *ACM Trans. Graphics 27*, 3, 68:1–68:10.



(a) Gamma mapped HDR input

(b) Mal-adaptation TMO [Irawan et al. 2005]

(c) Display adaptive TMO [Mantiuk et al. 2008]

(d) Virtual exposures TMO [Bennett and McMillan 2005]

(e) Temporal coherence TMO [Boitard et al. 2012]

(f) Zonal temporal coherence TMO [Boitard et al. 2014]

(g) Motion path filtering TMO [Aydin et al. 2014]

(h) Our method

Figure 3: *Tone mapping results of the* Poker *sequence, with highlighted details. To compare our method with the Virtual exposures TMO in terms of detail reproduction, the operators have been set to enhance details slightly (scaling 1.5), to facilitate comparison. Video frame from [Froehlich et al. 2014].*

(a) Gamma mapped HDR input

(b) Mal-adaptation TMO [Irawan et al. 2005]

(c) Display adaptive TMO [Mantiuk et al. 2008]

(d) Virtual exposures TMO [Bennett and McMillan 2005]

(e) Temporal coherence TMO [Boitard et al. 2012]

(f) Zonal temporal coherence TMO [Boitard et al. 2014]

(g) Motion path filtering TMO [Aydin et al. 2014]

(h) Our method

Figure 4: *Tone mapping results of the* Cars *sequence. Video frame from [Froehlich et al. 2014].*

(a) Gamma mapped HDR input

(b) Mal-adaptation TMO [Irawan et al. 2005]

(c) Display adaptive TMO [Mantiuk et al. 2008]

(d) Virtual exposures TMO [Bennett and McMillan 2005]

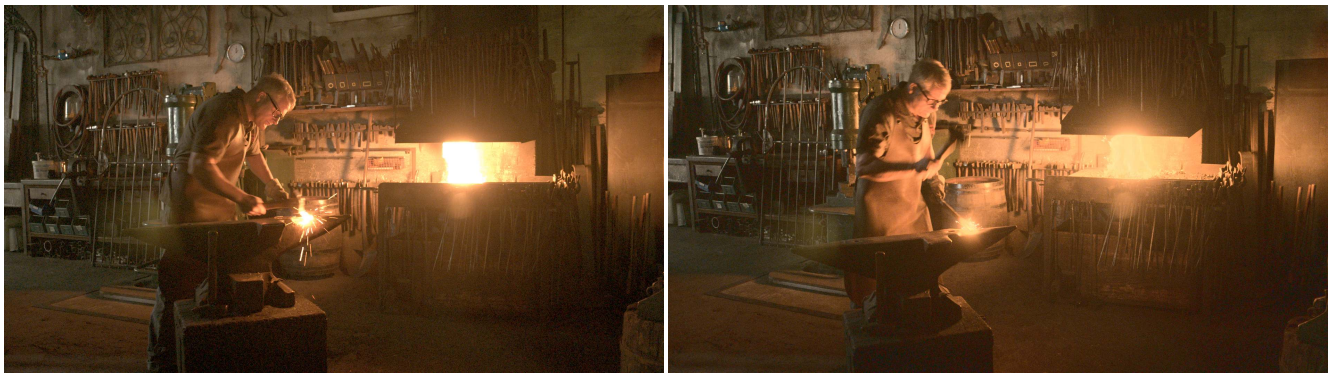
(e) Temporal coherence TMO [Boitard et al. 2012]

(f) Zonal temporal coherence TMO [Boitard et al. 2014]

(g) Motion path filtering TMO [Aydin et al. 2014]

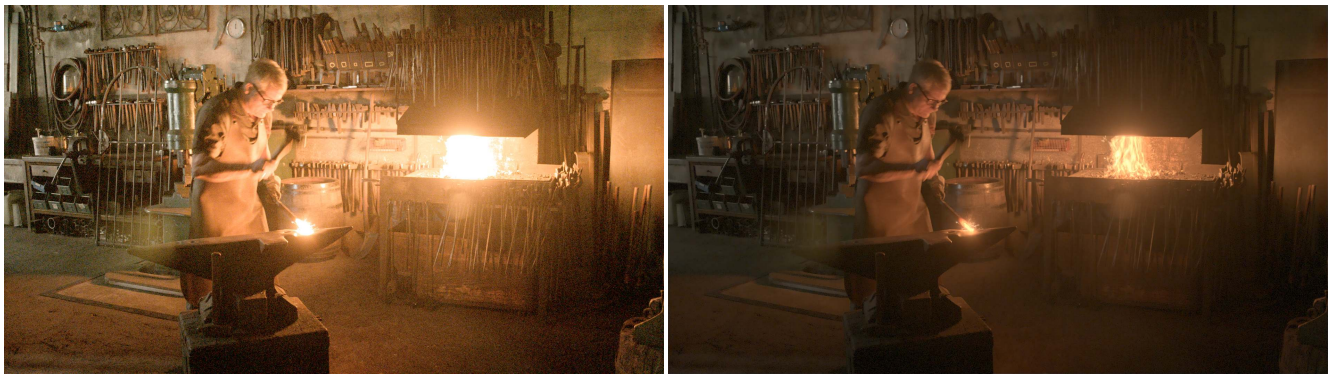
(h) Our method

Figure 5: *Tone mapping results of the* Showgirl *sequence. Video frame from [Froehlich et al. 2014].*



(a) Gamma mapped HDR input

(b) Mal-adaptation TMO [Irawan et al. 2005]

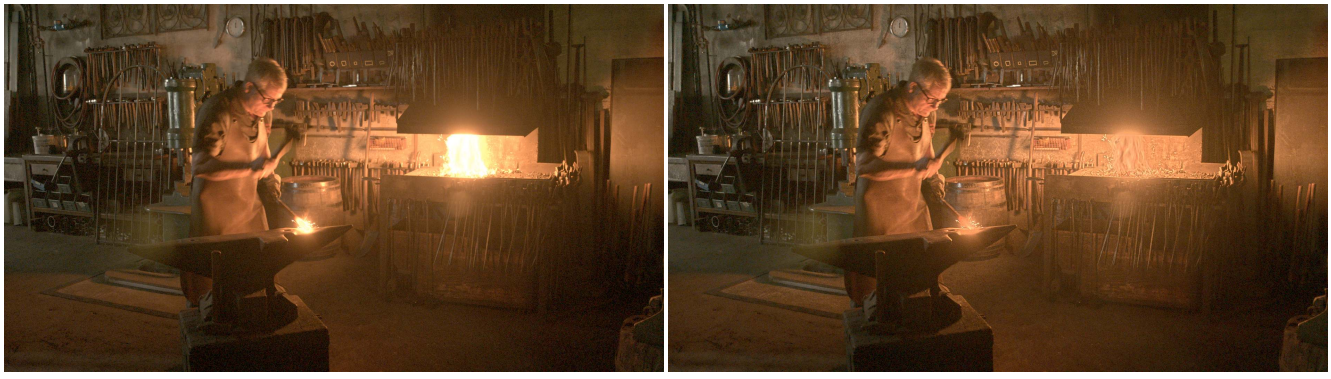


(c) Display adaptive TMO [Mantiuk et al. 2008]

(d) Virtual exposures TMO [Bennett and McMillan 2005]

(e) Temporal coherence TMO [Boitard et al. 2012]

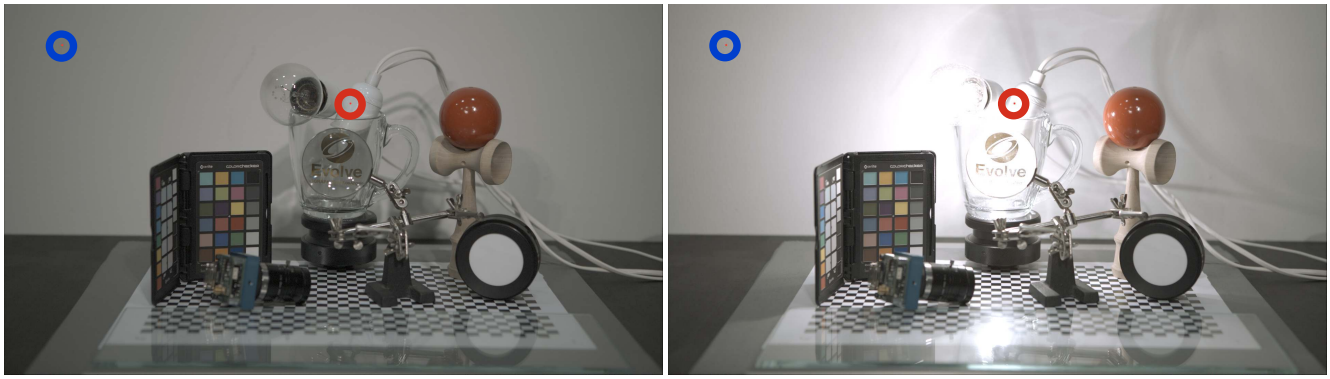
(f) Zonal temporal coherence TMO [Boitard et al. 2014]



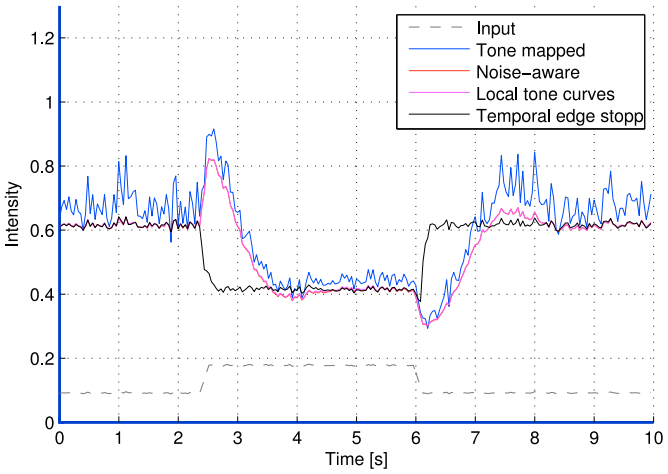
(g) Motion path filtering TMO [Aydin et al. 2014]

(h) Our method

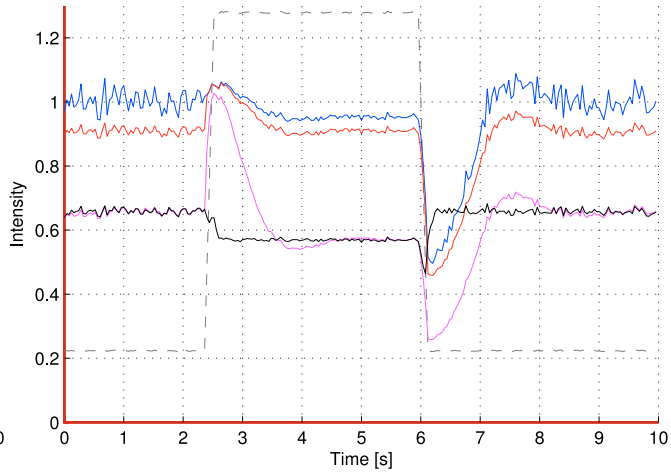
Figure 6: *Tone mapping results of the* Smith hammering *sequence. Video frame from [Froehlich et al. 2014].*



(a) (b)



(c) Lower temporal dynamic range (blue point)



(d) Higher temporal dynamic range (red point)

Figure 8: *The plots show the tone mapping result at the marked blue and red points in the top images. The different curves are produced by adding different properties in the tone curve evaluation; noise-aware tone curve and details scaling, local tone curves, and a temporal edge-stopping filter. The temporal filter used before applying the edge-stop function is a low-pass IIR filter.*