

Exposure Fusion Algorithm Based on Perceptual Contrast and Dynamic Adjustment of Well-Exposedness

Pablo Martínez-Cañada and Marius Pedersen

Gjøvik University College, Gjøvik, Norway
{pablo.canada,marius.pedersen}@hig.no

Abstract. The luminance of natural scenes frequently presents a high dynamic range and cannot be adequately captured with traditional imaging devices. Additionally, even if the technology to capture the scene is available, the image to be displayed on conventional monitors must be compressed by a tone mapping operator. Exposure fusion is an affordable alternative which blends multiple low dynamic range images, taken by a conventional camera under different exposure levels, generating directly the display image. In this paper, the Retinal-like Sub-sampling Contrast metric has been adapted to work with the original version of the exposure fusion algorithm in the CIELAB color space. In addition, saturation and well-exposedness metrics have been reformulated in this color space, adding a dynamic adjustment mechanism to the latter one which avoids amplification of invisible contrast. Results based on objective evaluation show that the proposed algorithm clearly outperforms the original exposure fusion technique and most of the state-of-the-art tone mapping operators for static images.

Keywords: Exposure Fusion, Tone Mapping, Perceptually Based Image Processing, Contrast, Saturation, Well-Exposedness.

1 Introduction

The range of luminances in natural scenes is often of five orders of magnitude. However, contrast ratio provided by conventional displays is not higher than two orders of magnitude. Tone Mapping Operators (TMOs) and, more recently, Exposure Fusion (EF) techniques are used to compress the High Dynamic Range (HDR) of luminances into the Low Dynamic Range (LDR) of the target display, generating images which are visually similar to the original scene. TMOs assume the HDR radiance map from the scene is available to calculate in a second step the corresponding LDR mapped image. EF directly provides the display LDR image with no necessity of the HDR radiance map. Instead EF uses a bracketed exposure sequence of LDR images which can be captured by a conventional camera. These images are blended together using some quality metrics, which guide the process of selecting those regions which contain the most relevant information. In addition, EF presents other advantages compared to TMOs, such

us the simplification of the pipeline, the non-necessity of a camera calibration process and the decrease of damage in scene details due to the one-step approach versus the two-step approach [1,2]. On the downside, EF takes more time when acquiring the set of input LDR images and camera shake or objects moving during the long-exposure captures might cause blurred images [3].

The EF algorithm proposed in this paper is based on the multiscale fusion approach by Mertens et al. [1]. The original contrast metric guided by a Laplacian filter has been replaced by a perceptually based contrast metric, Retinal-like Sub-sampling Contrast (RSC) [5], adapted to work in CIELAB color space. In addition, saturation and well-exposedness metrics have been reformulated in this color space, adding a dynamic adjustment mechanism to the latter one which avoids amplification of invisible contrast.

The rest of the paper is organized as follows; first a review of HDR mapping, then the proposed algorithm is presented. Further, the experimental setup and results are presented, at last conclusions are drawn.

2 Review of HDR Mapping

2.1 Tone Mapping Operators

Tone mapping algorithms can be classified according to the type of processing as global or local. Global operators are simple and barely time-consuming, but cannot deal well with huge contrast ratios. Local operators generally achieve greater contrast reduction allowing significant compression of the dynamic range of a scene. However, a major concern with them is the presence of halos or artifacts around edges. One of the simplest global operator is the linear scale factor by Ward [6], while one of the most representative local operator is the gradient based by Fattal et al. [7], to give some examples. A comprehensive study of TMOs can be found in [2,8,9,10]. Below we provide a short description of the algorithms selected to carry out our objective comparison.

In general, TMOs attempt to imitate the perceptual processing of the Human Visual System (HVS). Drago et al. [12] implement a logarithmic compression of the HDR values, simulating the photoreceptors response to light. In the same way, Reinhard et al. [13] use a global transformation by means of a sigmoid function which allows users to adjust some function parameters to preserve different characteristics of the HDR image. The operator proposed by Pattanaik et al. [14] includes a perceptually-based temporal processing of the stream of input frames simulating the changes in visual appearance caused by variations in scene luminance. In [15] Ashikhmin presents a TMO which mimics two relevant functions of HVS: reproduction of absolute brightness and preservation of local contrast. A TMO is applied locally, aiming to capture overall brightness of each region in the scene, then lost details and contrast are put back into the final image. Mantiuk et al. [16] proposed a framework for perceptual contrast processing of HDR images based on gradient domain, like Fattal et al. [7].

Reinhard et al. proposed another operator [17] which simulates the dodging and burning technique used in traditional photography allowing different exposures

across the image to be printed. The TMO of Durand et al. [18] carries out a two-scale decomposition of the image into a base layer, encoding large-scale variations, and a detail layer.

2.2 Exposure Fusion Algorithms

To the best of our knowledge, the original idea of using exposure fusion to perform the mapping of the HDR image was formulated initially by Mertens et al. [1]. Their algorithm computes an image quality metric for each pixel in the multi-exposure sequence, which encodes perceptually desirable qualities, i.e. contrast, saturation and well-exposedness. Guided by this set of quality metrics, the algorithm selects the best pixels from the sequence and combine them into the final result. The blending of input LDR images is driven by a multiscale approach in which a Laplacian pyramid of the input images is multiplied by the Gaussian pyramid of the quality weight maps. This multiscale processing is used to avoid seams in the final LDR mapped image.

Other authors [2,19] use probabilistic models to fuse LDR images. In [2] visible contrasts and scene gradients are computed first for the Y channel of the YUV color space and then a Maximum a Posteriori framework is considered in the fusion step to preserve these visible contrasts and suppress the gradient reversals between the input sequence and the corresponding LDR image. On the other hand, in [19] it is proposed to use a second-order derivative to compute contrast based on a central difference approach on a 3×3 neighborhood. In addition, a color consistency measure is applied, which is not considered in previous methods. A generalized random walks framework is used to calculate a globally optimal solution.

Other EF techniques have been proposed for some specific applications, such as mobile phones. In [3] longer exposed input images are corrected to avoid blurred images produced by camera and object motion. A parallel GPU and FPGA implementation is proposed in [4].

3 The Proposed Algorithm

The full processing pipeline of the algorithm by Mertens et al. [1] has been adapted to work in the CIELAB color space, preserving the multiscale fusion and the multiplicative weighting combination of the original method. By introducing a color space perceptually more correlated with the HVS we expect that the final image further resembles the visual scene. Further, multiscale fusion is a simple technique which removes fairly well artifacts in the final image, i.e. halos and seams, so that it has been included in our algorithm without modification. The pseudo code of the proposed algorithm is shown below. Contrast, Saturation and Well Exposedness quality weighting maps are calculated for the CIELAB input set of images and combine in a multiplicative fashion. The pyramid computation, following the algorithm by Mertens et al. [1] combines the Gaussian pyramid of the quality weighting map and the Laplacian pyramid of the CIELAB images.

Finally this pyramid is collapsed to obtain the tone-mapped image, which is transformed back to the RGB space.

```

Perceptual Exposure Fusion (LDR_Images, W_threshold)
begin
  Lab_Im = CIELAB(LDR_Images);
  Contrast = C(Lab_Im);
  Saturation = S(Lab_Im);
  Well_Exposedness = W(Lab_Im, W_threshold);

  Quality_Measure = Contrast*Saturation*Well_Exposedness;
  normalize(Quality_Measure);

  initialize(pyr);

  for i = 1:Number_LDR_images
    pyrW = gaussian_pyramid(Quality_Measure);
    pyrI = laplacian_pyramid(Lab_Im);
    for l = 1:pyr_number_levels
      pyr[l] = pyr[l] + pyrW[l]*pyrI[l];
    end
  end

  R = reconstruct_laplacian_pyramid(pyr);
  LDR_output = RGB(R);
end

```

The quality metrics are key components in the processing pipeline. We propose to introduce modified versions of these metrics.

Contrast. The Retinal-like Sub-sampling Contrast (RSC) metric [5] was proposed to measure contrast in digital images, and has shown to correlate well with perceived contrast. RSC uses a multilevel approach with Tadmor and Tolhurst’s [20] Difference-of-Gaussians (DoG) to measure the contrast in each pixel. In the original formulation [5] three DoG contrast maps, $DoG_i^{c_k}$, are calculated in each level i , halving the size of the image, which correspond to the three color channels c_k : L^* , a^* and b^* (c_1, c_2, c_3) and then these three maps are linearly combined to obtain one number of contrast at the end. However, in the EF algorithm a weight map is required so that we consider the DoG contrast map of the first level, see Eq. 1, as the final contrast map $W_{contrast}^{c_k}$ for each color channel.

$$W_{contrast}^{c_k} = DoG_{i=1}^{c_k}, \quad (1)$$

The width of center and surround Gaussian components are set to 1 and 2, following the recommendation by [21]. It was observed experimentally that RSC

produces some artifacts overamplifying small undetectable contrasts in underexposed images. The normalization factor when computing the DoG (see equation 2 in [5] for more details) tends to highlight the contrast even when pixel values are very low. For example, we can think of a very dark region of the image with receptive-field values for the center and surround of 0.002 and 0.001. In this case, according to equation 2 in [5] the output of the DoG is 0.33, a third of the maximum value. However, this difference between center and surround is barely perceived. A simple solution is proposed based on a set of fixed thresholds for each $L^*a^*b^*$ component. Considering a normalized range of values for each color component in the range [0,1] if the value of a pixel in the LDR image is below 0.05 automatically the RSC metric is not computed and its contrast value is set to 0.

Saturation. Saturation is reformulated in CIELAB color space. In [22] image colorfulness is correlated with experimental data using a linear combination of the standard deviation and mean value in the a^*b^* plane. We base our metric in the same concept and compute for each pixel the standard deviation within the a^* and b^* channels according to Eq. 2.

$$W_{saturation} = 10\sqrt{\frac{(a^* - ((a^* + b^*)/2))^2 + (b^* - ((a^* + b^*)/2))^2}{2}}, \quad (2)$$

A multiplicative scale factor of 10 is added to set the output of the saturation metric in the same range of the other quality metrics.

Well-Exposedness. The function based on the Gauss curve of the original EF algorithm is applied to the luminance component, L^* , to determine how close is each pixel to the considered gray value, i.e. 0.5. Then the same weight map is applied to the three color components.

We observed a susceptibility of this metric to overamplify contrast if all overexposed images from the input sequence were included in the computation. Therefore, a dynamic control mechanism has been added to this quality metric, based on the global luminance level of each LDR input image, resulting in a considerable decrease of the amplification of invisible contrast in the final results. We estimate the global luminance level as the average of the L^* value for each pixel in each LDR image. Then if this luminance level is above a threshold the well-exposedness metric of the LDR image is set to 0, which means that this input image is not considered for the fusion procedure because of the multiplicative nature of the weighting in the original EF algorithm (see [1]).

4 Experimental Setup

In the literature, we find that most of the TMOs and EF algorithms are often perceptually assessed by visual comparisons of subjects who rate or rank the different tone mapped LDR images [8,23,24,25] and whose decisions are later

statistically analyzed. Conditions of each experiment and metrics used vary from one to another and there is no standard methodology for performing such studies. In this paper, we contrast the performance of some representative TMOs and EF techniques widely used for static images by means of one of the most popular objective quality metric used in tone mapping of HDR images. The Dynamic Range Independent Metric (DRIM) [26], using a model of the HVS, is capable of comparing images which have different dynamic ranges. This metric has been amply validated and results show high correlation with visual ratings. For a recent comparison of TMOs including a temporal model see [27].

The output of DRIM consists of three quality maps: contrast reversal, contrast lost and contrast amplification. To obtain a single number from these maps a Minkowski's pooling of each one is computed with exponent 4. This exponent is considered by Wang et al. [28] to further penalize low quality scores in the image. The average of the three pooling numbers associated to each quality map of the DRIM metric is considered as the final result. Then a sign-test at the 5 % significance level is applied comparing results between the proposed algorithm and each one of the evaluated algorithms.

The code of the evaluated TMOs can be downloaded from the web page of the *HDR lab* in the *Max Planck Institute* [11]. All these algorithms and the EF technique run with default parameters, except *Pattanik00* whose parameter m has been modified to 10 to produce reasonably fair results. Nevertheless, a fine tune of these parameters might improve results of the TMOs considered. For the proposed algorithm, considering normalized LDR input images in the range $[0,1]$, the threshold of the well-exposedness metric has been fixed to 0.5 in all images. DRIM runs online with default parameters [29] and each quality map, i.e. the contrast reversal, contrast lost and contrast amplification, is saved independently through the interactive viewer.

The ten images selected are commonly used in HDR Tone Mapping and can be referenced as: *Bristol Bridge*, *Tinterna*, *Atrium Night*, *Belgium House*, *Stanford Memorial Church*, *Girl in Lit Room*, *Clock Building*, *Cornell Box*, *Tahoe* and *Rosette*.

5 Experimental Results

Fig. 1 shows a comparison between the output of the original EF technique and the proposed algorithm for the image *Stanford Memorial Church*. The algorithm by Mertens et al. tends to increase the brightness and overamplify contrast of the tone mapped image. The DRIM metric marks in blue these overamplified regions of the image corresponding to the wooden structure of the ceiling. It can be argued that an increase of contrast usually enhances details in the darkest areas of the image, as can be seen in the upper left corner of both images. However, if contrast is too amplified some colors may be distorted such as the color of wood in this example. In fact, this distortion has been observed in other images, specially in dimly lit areas of the scene, when contrast is overamplified causing a similar effect of noise in images. An excessive increase of brightness

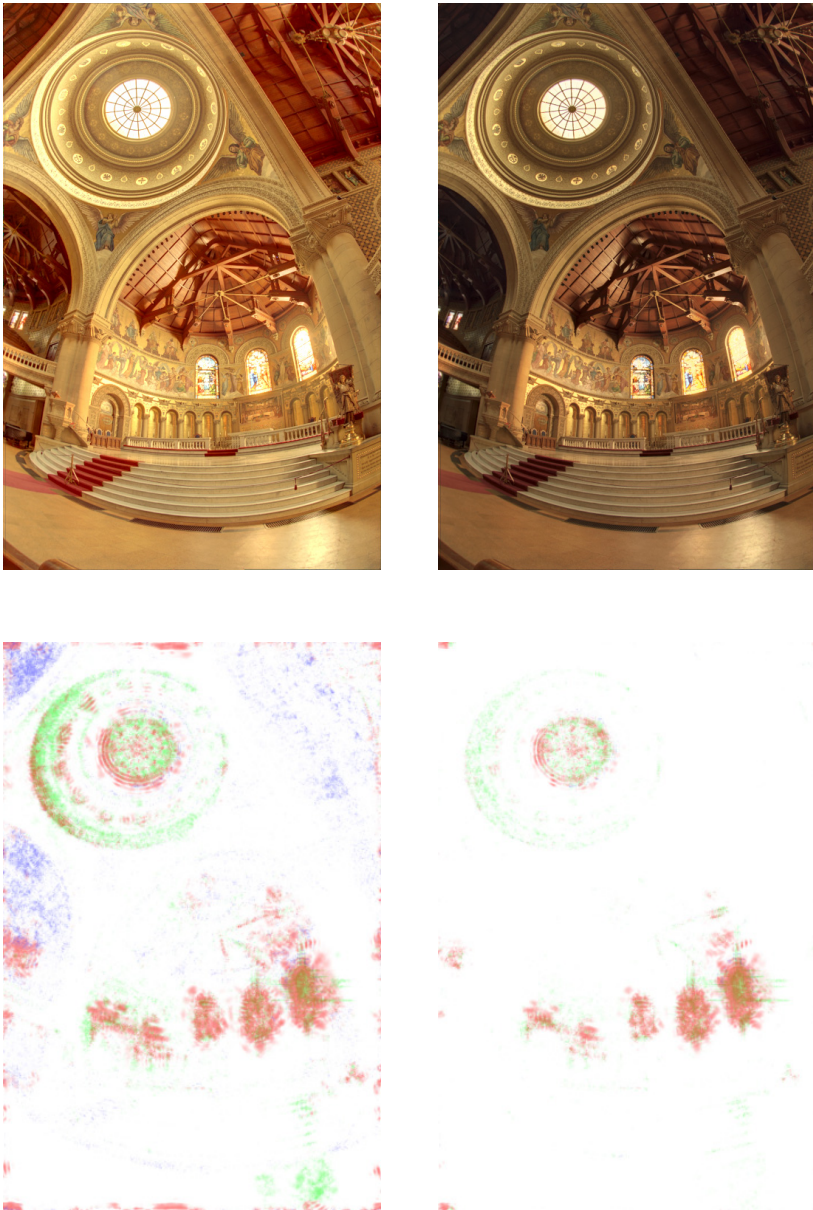


Fig. 1. On top: output images of the original EF (left) technique and the proposed EF algorithm (right). In the bottom: DRIM evaluation of both images respectively. Contrast reversal is marked in red, contrast loss in green and contrast amplification in blue.

also causes losses of contrast in over-exposed regions of the HDR image, marked in green in the dome ring.

Table 1. Minkowski’s pooling with exponent 4 of DRIM images and average of the three values. Results shown have been multiplied by 1000 to better viewing. *PA* is the proposed algorithm, *EF* is the original Exposure Fusion [1], *Dr03* is Drago03 [12], *Du02* is Durand02 [18], *Ma06* is Mantiuk06 [16], *Pa00* is Pattanaik00 [14], *Re02* is Reinhard02 [17] and *Re05* is Reinhard05 [13].

	Algorithm							
	PA	EF	Dr03	Du02	Ma06	Pa00	Re02	Re05
Bridge	0.06	44.36	27.24	0.30	13.78	28.84	7.21	74.10
Tinterna	0.02	22.28	7.99	0.02	11.51	0.00	0.52	3.19
Atrium	3.76	12.51	10.08	5.48	4.46	5.69	6.24	18.79
House	1.44	11.01	16.07	1.90	1.85	31.01	31.01	38.10
Memorial	6.65	16.24	13.86	10.26	7.53	20.69	14.95	39.02
Girl-Room	16.34	50.24	69.41	20.30	26.96	109.70	45.13	139.41
Clock Building	13.37	79.79	28.55	17.20	38.79	1.59	30.37	44.25
Cornell Box	8.31	8.53	65.03	19.67	12.37	12.86	15.11	173.92
Tahoe	27.62	5.77	99.69	64.25	57.22	114.82	69.59	301.11
Rosette	16.00	32.81	22.48	2.83	6.391	48.14	44.05	63.67

The objective evaluation conducted reflects that our algorithm produces the best result in six out of ten images, see Table 1. When the proposed algorithm is outperformed the pooling obtained is the closest to the best one. In general, highly textured images, such as *Belgium House* and *Stanford Memorial Church*, and very under-exposed images, such as *Cornell Box*, result in the best performance of our algorithm. The statistical sign-test at the 5% significance level of Table 2 shows a reliable improvement of the algorithm proposed with regard to the rest of algorithms evaluated except for *Pattanaik00*.

Table 2. Sign-test at the 5% significance level. The proposed algorithm is compared to each one of the evaluated algorithms.

	Algorithm						
	EF	Dr03	Du02	Ma06	Pa00	Re02	Re05
p-value	0.022	0.002	0.040	0.022	0.110	0.002	0.002
Reject null hypothesis	yes	yes	yes	yes	no	yes	yes

6 Conclusions

A HDR mapping algorithm is proposed based on the EF technique. The RSC metric has been adapted to work with the original version of the EF in the CIELAB color space. Saturation and well-exposedness metrics have been reformulated in this color space, adding a dynamic adjustment mechanism to the latter one which avoids amplification of invisible contrast.

Ten HDR images, commonly used in tone mapping, are considered to perform an objective quality evaluation using the DRIM metric. Our algorithm is compared to the original EF technique and six of the most representative TMOs for static images. Parameters of our algorithm and other algorithms are fixed for all images. A statistical sign-test at the 5% significance level shows that the proposed algorithm clearly outperforms the original EF algorithm and five out of six of the TM operators. More images are going to be considered in future tests to reliably show the improvement of the algorithm proposed. In addition, a perceptual subjective experiment could support results presented in this paper.

References

1. Mertens, T., Kautz, J., Van Reeth, F.: Exposure fusion. In: IEEE 15th Pacific Conference on Computer Graphics and Applications, pp. 382–390. IEEE (2007)
2. Song, M., Tao, D., Chen, C., Bu, J., Luo, J., Zhang, C.: Probabilistic exposure fusion. In: 21th IEEE Transactions on Image Processing, pp. 341–357. IEEE (2012)
3. Tico, M., Gelfand, N., Pulli, K.: Motion-blur-free exposure fusion. In: 17th IEEE International Conference on Image Processing, pp. 3321–3324. IEEE (2010)
4. Ureña, R., Martínez-Cañada, P., Gomez-Lopez, J.M., Morillas, C., Pelayo, F.: Real-time tone mapping on GPU and FPGA. *EURASIP Journal on Image and Video Processing*, 1–15 (2012)
5. Rizzi, A., Simone, G., Cordone, R.: A modified algorithm for perceived contrast measure in digital images. In: Conference on Colour in Graphics, Imaging, and Vision, pp. 249–252. Society for Imaging Science and Technology (2008)
6. Ward, G.: A contrast-based scalefactor for luminance display. In: 4th Graphics Gems, pp. 415–421 (1994)
7. Fattal, R., Lischinski, D., Werman, M.: Gradient domain high dynamic range compression. *ACM Transactions on Graphics*, 249–256 (2002)
8. Ledda, P., Chalmers, A., Troscianko, T., Seetzen, H.: Evaluation of tone mapping operators using a high dynamic range display. *ACM Transactions on Graphics*, 640–648 (2005)
9. Salih, Y., Malik, A.S., Saad, N.: Tone mapping of HDR images: A review. In: 4th IEEE International Conference on Intelligent and Advanced Systems, pp. 368–373. IEEE (2012)
10. Čadík, M., Wimmer, M., Neumann, L., Artusi, A.: Evaluation of HDR tone mapping methods using essential perceptual attributes. *J. Computers & Graphics* 32, 330–349 (2008)
11. PFStmo package from the Max Planck Institute, <http://www.mpi-inf.mpg.de/resources/tmo/>
12. Drago, F., Myszkowski, K., Annen, T., Chiba, N.: Adaptive logarithmic mapping for displaying high contrast scenes. In: Computer Graphics Forum, pp. 419–426. Blackwell Publishing, Inc. (2003)
13. Reinhard, E., Devlin, K.: Dynamic range reduction inspired by photoreceptor physiology. *IEEE Transactions on Visualization and Computer Graphics*, 13–24 (2005)
14. Pattanaik, S.N., Tumblin, J., Yee, H., Greenberg, D.P.: Time-dependent visual adaptation for fast realistic image display. In: 27th Proceedings of the Annual Conference on Computer Graphics and Interactive Techniques, pp. 47–54. ACM Press/Addison-Wesley Publishing Co. (2000)

15. Ashikhmin, M.: A tone mapping algorithm for high contrast images. In: 13th Proceedings of the Eurographics Workshop on Rendering, pp. 145–156. Eurographics Association (2002)
16. Mantiuk, R., Myszkowski, K., Seidel, H.P.: A perceptual framework for contrast processing of high dynamic range images. *ACM Transactions on Applied Perception*, 286–308 (2006)
17. Reinhard, E., Stark, M., Shirley, P., Ferwerda, J.: Photographic tone reproduction for digital images. *ACM Transactions on Graphics*, 267–276 (2002)
18. Durand, F., Dorsey, J.: Fast bilateral filtering for the display of high-dynamic-range images. *ACM Transactions on Graphics*, 257–266 (2002)
19. Shen, R., Cheng, I., Shi, J., Basu, A.: Generalized random walks for fusion of multi-exposure images. *IEEE Transactions on Image Processing*, 3634–3646 (2011)
20. Tadmor, Y., Tolhurst, D.: Calculating the contrasts that retinal ganglion cells and LGN neurones encounter in natural scenes. In: 40th Vision Research, pp. 3145–3157 (2000)
21. Simone, G., Pedersen, M., Hardeberg, J.Y.: Measuring Perceptual Contrast in Digital Images. *Journal of Visual Communication and Image Representation*, 491–506 (2012)
22. Hasler, D., Süsstrunk, S.: Measuring colourfulness in natural images. In: Proc. SPIE, pp. 87–95 (2003)
23. Drago, F., Martens, W.L., Myszkowski, K., Seidel, H.P.: Perceptual evaluation of tone mapping operators. In: ACM SIGGRAPH 2003 Sketches & Applications, p. 1. ACM (2003)
24. Yoshida, A., Blanz, V., Myszkowski, K., Seidel, H.P.: Perceptual evaluation of tone mapping operators with real-world scenes. In: Electronic Imaging, pp. 192–203. International Society for Optics and Photonics (2005)
25. Yoshida, A., Mantiuk, R., Myszkowski, K., Seidel, H.P.: Analysis of Reproducing Real-World Appearance on Displays of Varying Dynamic Range. In: Computer Graphics Forum, pp. 415–426. Blackwell Publishing, Inc. (2006)
26. Aydin, T.O., Mantiuk, R., Myszkowski, K., Seidel, H.P.: Dynamic range independent image quality assessment. *ACM Transactions on Graphics*, 69 (2008)
27. Eilertsen, G., Wanat, R., Mantiuk, R.K., Unger, J.: Evaluation of Tone Mapping Operators for HDRVideo. In: Computer Graphics Forum, pp. 275–284. Wiley (2013)
28. Wang, Z., Shang, X.: Spatial pooling strategies for perceptual image quality assessment. In: IEEE International Conference on Image Processing, pp. 2945–2948. IEEE (2006)
29. DRIM online version, <http://driiqm.mpi-inf.mpg.de>