Xim Cerdá, C. Alejandro Párraga, Xavier Otazu
Dept. Ciències de la Computació – Centre de Visió per Computador
Univ. Autònoma de Barcelona, Spain.
http://www.cvc.uab.cat

# Which tone-mapping is the best?
# A comparative study of tone-mapping perceived quality

## Motivation

High-dynamic-range (HDR) imaging refers to the methods designed to increase the dynamic range present in standard digital imaging techniques. This increase is achieved by taking the same picture under different exposure values and mapping the resulting HDR intensity levels into a single Low-dynamic-range (LDR) image by way of a tone-mapping operator (TMO).

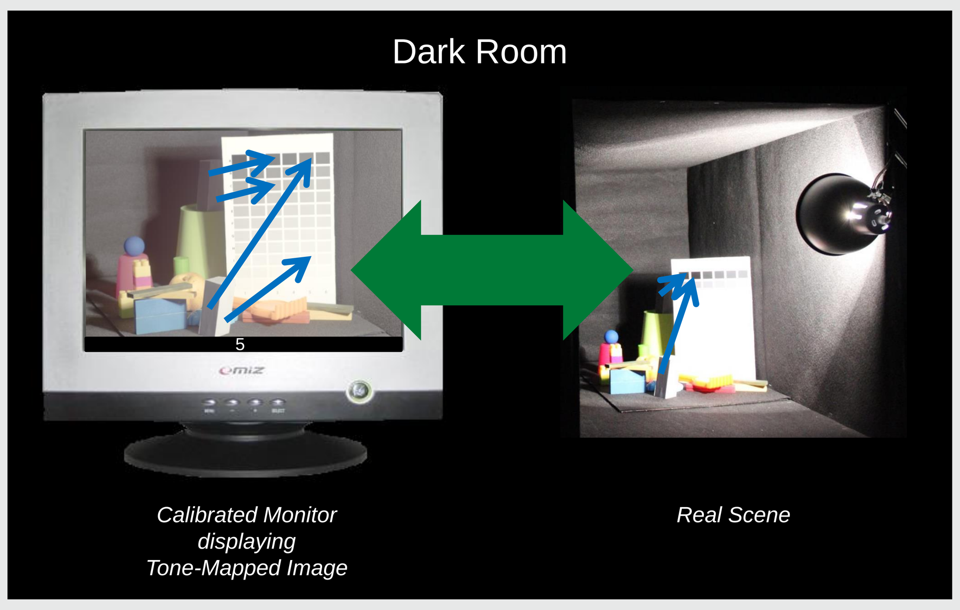
Currently, there is no agreement on how to evaluate the quality of different TMOs. In this work we psychophysically evaluate 15 different TMOs obtaining rankings based on the perceived properties of the resulting tone-mapped images.

Our criteria is that the best TMO should be the one that perceptually reproduces the real scene best.

*LDR image of HDR scene, without using tone-mapping operator (left) and using one (right).*

## Laboratory Setup

- Calibrated CRT Monitor: Mitsubishi Diamond-Pro 2045u
- ViSaGe MKI Stimulus Generator
- Sigma SD10 Camera
- Sigma Photo Pro
- HDR Toolbox for Matlab




- 10 participants (4 females, 6 males, age between 17-54 y.o.)
- 3 different man-made scenes with grey and colored objects
- Reference Table (65 grey-level patches)
- Grey objects: 15 different surfaces (5 in each of the 3 scenes)

## Methods & Results

### Exp 1: Grey-Levels Matching

**Objective**

To study the **INTERNAL** relationships among grey-levels in the tone-mapped image and in the real scene and to construct a ranking of tone mapping operators according to these relationships.

**Procedure**

1. A printed grey-level **reference table** was created and included in all scenes. The luminance of its patches was measured using a spectroradiometer.
2. Subjects were asked to **match (in the real scene) the grey-levels of object's surfaces to the grey-levels of the reference table**.
3. Subjects conducted the **same task** using the **tone-mapped** image presented on the monitor screen.
4. Results were converted to cd/m$^2$ using Table 1.

**Table 1: Reference Table Luminance (cd/m$^2$)**

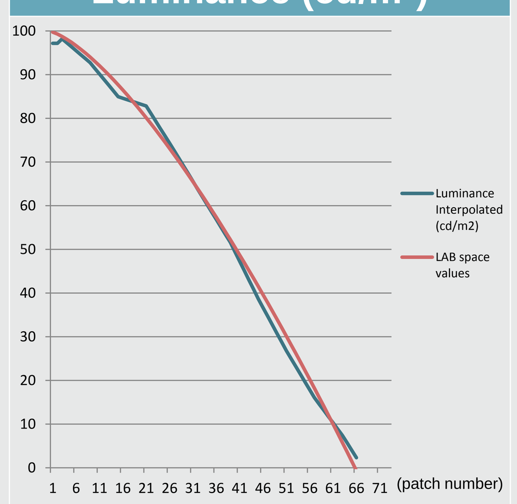



*Table 1: luminance of each patch in the reference table as measured by the spectroradiometer.*

**Results**

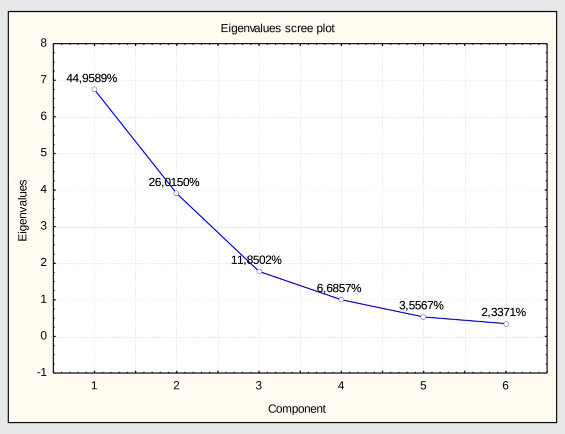
1. We obtained 16 values (corresponding to the 15 TMO algorithms plus the real scene) for each of the 15 surfaces considered.
2. 15 different **ANOVAs** were calculated, (one per grey-level surface and scene) and **Fisher's Least Significant Difference (LSD)** post-hoc analysis was applied to obtain a ranking.
3. All observer's results were averaged (one value for each surface and TMO) and a **PCA** was applied: 15 dimensions were reduced to 6 and another ranking was obtained by measuring the **Euclidean distance** in the new space from every TMO to the real scene.
4. A **Spearman's correlation** between the two rankings was calculated ($p<0.05$, Table 2)

**ANOVA**

| TMO | Score |
|---|---|
| iCAM | 14 |
| Ferradans | 11 |
| KimKautz | 10 |
| Durand | 9 |
| Fattal | 9 |
| Krawczyk | 9 |
| Mertens | 9 |
| Reinhard-Devlin | 9 |
| Li | 9 |
| Drago | 8 |
| Meylan | 8 |
| Otazu | 8 |
| Reinhard | 8 |
| Ashikhmin | 7 |
| Ferwerda | 6 |

**PCA**

| TMO | Eucl.Dist |
|---|---|
| iCAM | 2.26 |
| Durand | 4.26 |
| Fattal | 4.52 |
| Li | 4.85 |
| Mertens | 4.89 |
| KimKautz | 4.96 |
| Krawczyk | 5.16 |
| Reinhard | 5.24 |
| Meylan | 5.27 |
| Reinhard-Devlin | 5.29 |
| Ferwerda | 5.66 |
| Ferradans | 5.91 |
| Drago | 6.02 |
| Otazu | 6.22 |
| Ashikhmin | 9.00 |




*PCA components, corresponding eigenvalues and percentage total data represented by every component. The 95.40% of data could be represented in 6 components.*

| Spearman's Correlation | Ranking PCA | Ranking ANOVA |
|---|---|---|
| Ranking PCA | 1.000 | 0.616 |
| Ranking ANOVA | 0.616 | 1.000 |

*Table 2: Spearman's correlation of PCA and ANOVA rankings is significant at p<0.05.*

### Exp 2: Pairwise Comparison

**Objective**

To study the **GLOBAL** characteristics that determine whether a TMO image is more or less similar to the real scene and to rank the best perceptual tone-mapping operators.

**Procedure**

We performed a **pairwise comparison** between 15 tone-mapped images obtained applying the different algorithms. Each pair was presented besides the real scene and we asked subjects to choose the **most realistic image**.

**Results**

Using the **5th case of Thurstone's law**, we obtained a ranking for each scene and a global

**SCENE 1**

| TMO | Score |
|---|---|
| KimKautz | 6.92 |
| Reinhard | 6.85 |
| Krawczyk | 6.74 |
| Ferwerda | 6.72 |
| Drago | 6.01 |
| Ferradans | 5.52 |
| Li | 5.41 |
| Otazu | 4.50 |
| iCAM | 4.50 |
| Durand | 4.09 |
| Meylan | 1.74 |
| Reinhard-Devlin | 1.53 |
| Ashikhmin | 1.24 |
| Mertens | 0.14 |
| Fattal | 0 |

**SCENE 2**

| TMO | Score |
|---|---|
| Krawczyk | 7.50 |
| KimKautz | 7.35 |
| Reinhard | 7.06 |
| Ferwerda | 6.45 |
| Ferradans | 6.28 |
| Drago | 5.82 |
| Li | 5.38 |
| Otazu | 4.15 |
| iCAM | 3.73 |
| Durand | 3.59 |
| Mertens | 1.25 |
| Reinhard-Devlin | 0.92 |
| Meylan | 0.83 |
| Ashikhmin | 0.43 |
| Fattal | 0 |

**SCENE 3**

| TMO | Score |
|---|---|
| Krawczyk | 6.93 |
| Ferradans | 6.71 |
| KimKautz | 6.57 |
| Reinhard | 6.39 |
| Drago | 6.19 |
| Ferwerda | 5.70 |
| Durand | 4.97 |
| Li | 4.65 |
| Otazu | 4.12 |
| iCAM | 4.03 |
| Mertens | 1.42 |
| Reinhard-Devlin | 1.12 |
| Meylan | 1.12 |
| Ashikhmin | 1.04 |
| Fattal | 0 |

We obtained the **Spearman's rank correlation coefficients** ($p<0.05$) between the rankings of the scenes.

Since correlations were higher or equal than 0.90, we generated a global ranking, based on the different scene's rankings.

| Spearman's Correlation | Scene 1 | Scene 2 | Scene 3 | Global |
|---|---|---|---|---|
| Scene 1 | 1.00 | 0.96 | 0.90 | 0.98 |
| Scene 2 | 0.96 | 1.00 | 0.95 | 0.97 |
| Scene 3 | 0.90 | 0.95 | 1.00 | 0.94 |
| Global | 0.98 | 0.97 | 0.94 | 1.00 |

**GLOBAL**

| TMO | Score |
|---|---|
| KimKautz | 6.73 |
| Krawczyk | 6.61 |
| Reinhard | 6.56 |
| Ferwerda | 5.96 |
| Drago | 5.85 |
| Ferradans | 5.64 |
| Li | 4.84 |
| Durand | 4.50 |
| iCAM | 4.24 |
| Otazu | 4.19 |
| Reinhard-Devlin | 1.70 |
| Meylan | 1.63 |
| Mertens | 1.38 |
| Ashikhmin | 1.12 |
| Fattal | 0 |

## Conclusions

In Experiment 1 the best algorithm was ICAM by J.Kuang et al (2007) and in Experiment 2 the three top algorithms (which differed by less than a jnd) were KimKautz (Kim and Kautz, 2008), Krawczyk (Krawczyk et al, 2005) and Reinhard (Reinhard et al, 2002). Our results also show no correlation between the rankings produced by these two experiments. With the possible exception of KimKautz (3rd, 6th, 1st), no single algorithm comes near the top of the ranking in all these metrics or is capable of scoring high for both the global and local criteria analysed.

Our results also suggest that these algorithms may have been defined using different criteria, depending on the aim of their authors. For example, it might be possible that iCAM was defined taking into account a local criterion, while Krawczyk was defined using a global one. We conclude that an agreed standard criteria is needed for defining tone-mapping operators and this method should take into account some local and global characteristics of the image. As a corollary, we consider there is ample room for improvement in the future development of TMO algorithms.

### References

Ashikhmin, M. , "A Tone Mapping Algorithm for High Contrast Images", Thirteenth Eurographics Workshop on Rendering, 2002.
Banterle, F. et al., "Advanced High Dynamic Range Imaging: Theory and Practice", AK Peters, CRC Press, 2011.
Drago, F. et al., "Adaptive Logarithmic Mapping For Displaying High Contrast Scenes", Proc. of EUROGRAPHICS, vol. 22 of Comp. Graph. 419-426, 2003.
Durand, F. and Dorsey, J., "Fast Bilateral Filtering for the Display of High-Dynamic-Range Images", In Proc. of ACM SIGGRAPH, ACM Press, 257-266, 2002.
Fattal, R. et al., "Gradient Domain High Dynamic Range Compression", In Proc. of ACM SIGGRAPH , ACM Press, 249-256, 2002.
Ferwerda, J.A. et al., "A Model of Visual Adaptation for Realistic Image Synthesis", In Proc. of ACM SIGGRAPH, ACM Press, 249-258, 1996.
Ferradans, S. et al., "An Analysis of Visual Adaptation and Contrast Perception for Tone Mapping", TPAMI, vol. 33, no. 10, 2011.
Kuang, J. et al., "iCAM06: A refined image appearance model for HDR image rendering", J. Vis. Commun. image R. 18, 406-414, 2007.
Kim, M.H. and Kautz, J., "Consistent Tone Reproduction", Proc. Computer Graphics and Imaging, 2008.
Krawczyk, G. et al., "Lightness Perception in Tone Reproduction for High Dynamic Range Images", Proc. of EUROGRAPHICS, vol. 24(3), 2005.
Li, Y. et al., "Compressing and Compading High Dynamic Range Images with Subband Architectures", ACM Trans. Graph., vol. 24, no.3, pp. 836- 844, 2005.
Mertens, T. et al., "Exposure Fusion", 15th Pacific Conf. on Comp. Graph. and Applications, pp. 382- 390, 2007.
Meylan, L. et al., "Model of retinal local adaptation for the tone mapping of color filter array images", J. Opt. Soc. Am. A, vol. 24, no. 9, 2007.
Otazu, X., "Perceptual tone-mapping operator based on multiresolution contrast decomposition", Perception, vol.41, pp. 86 Suppl, 2012.
Reinhard, E. and Devlin, K. , "Dynamic Range Reduction inspired by Photoreceptor Physiology", IEEE Trans. on Vis. Comp. Graph, vol. 11 (1), pp. 13-24, 2005.
Reinhard, E. et al., "Photographic Tone Reproduction for Digital Images", ACM Trans.on Graphics, vol. 21, no.3 pp. 267-276, 2002.

### Acknowledgements

CAP and XO are partially supported by the Spanish Ministry of Science and Innovation through research project TIN2010-21771-C02-01 and TIN2013-41751.