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.

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/m² using Table 1.

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 Euclidian 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).

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 Thurstonian’s law, we obtained a ranking for each scene and a global ranking.

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, 2000) 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, 8th, 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.

Acknowledgements
This work was partially supported by the Spanish Ministry of Science and Innovation through research project TIN2015-67777-C2-1-R and TIN2015-67777-C2-2-R.

References

Acknowledgements

This work was partially supported by the Spanish Ministry of Science and Innovation through research project TIN2015-67777-C2-1-R and TIN2015-67777-C2-2-R.

References


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.

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/m² using Table 1.

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 Euclidian 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).

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 Thurstonian’s law, we obtained a ranking for each scene and a global ranking.

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, 2000) 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, 8th, 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.