

THUẬT TOÁN CHUYỂN ĐỔI ẢNH DẢI ĐỘNG CAO SANG ẢNH DẢI ĐỘNG THẤP DỰA TRÊN HỌ PHÂN GIẢI TUYẾN TÍNH TÁCH BIỆT VỚI NỘI SUY ĐA THỨC

TONE MAPPING OPERATORS BASED ON SEPARABLE NON-LINEAR MULTIRESOLUTION FAMILIES WITH POLYNOMIAL INTERPOLATION

Thái Bá Chiến¹, Phạm Quốc Thái¹

¹Trường Đại học Bách khoa - Đại học Đà Nẵng; tbchien@dut.udn.vn; pqthai@dut.udn.vn

(Nhận bài: 30/11/2020; Chấp nhận đăng: 19/01/2021)

Tóm tắt - Bài viết này đề xuất các toán tử ánh xạ tông màu hình ảnh (TMO). Một bức ảnh dải động cao sẽ tìm được các hệ số xấp xỉ, nhờ toán tử lấy mẫu hoặc toán tử trung bình, và các hệ số chi tiết. Các hệ số chi tiết thì có liên quan đến phép nội suy đa thức mà có lợi thế để tính đến các điểm kỳ dị, chẳng hạn như các điểm cạnh của hình ảnh, trong mô hình toán học do đó bảo toàn thông tin cấu trúc của hình ảnh dải động cao (HDR). Ngoài ra, hiện tượng Gibbs, có hại trong hình ảnh ánh xạ tông màu, được tránh. Đánh giá chất lượng của hình ảnh ánh xạ tông màu được đo theo số liệu TMQI. Kết quả mô phỏng cho thấy các TMO được đề xuất cung cấp kết quả tốt so với các chiến lược TMO truyền thống.

Từ khóa - Ảnh dải động cao; toán tử ánh xạ tông màu hình ảnh; nội suy đa thức; chỉ số chất lượng tông màu.

Abstract - This paper proposes High Dynamic Range (HDR) Image Tone Mapping Operators (TMOs). An HDR image will be separate into approximation parts, thanks to downsampling or averaging operator, and detail parts. The later part uses polynomial interpolation, which has the advantages of the singularities such as, firstly, edge points of the image thus preserving the structural information of the HDR image. Secondly, solving the Gibbs phenomenon, which is harmful in tone mapped images. The quality assessment of the tone mapped images is measured by the TMQI metric. Simulation results show that the proposed TMOs provide good results compared to traditional TMO strategies.

Key words - High dynamic range (HDR) image; Tone Mapping Operator (TMO); polynomial interpolation; Tone Mapped Quality Index (TMQI).

1. Đặt vấn đề

Hệ thống thị giác con người (HVS - Human Visual System) có thể cảm nhận các cảnh thực với nhiều màu sắc và cường độ khác nhau. Để trung thành với HVS, việc thu nhận cảnh trong thế giới thực bao gồm chụp nhiều hình ảnh dải động thấp (LDR - Low Dynamic Range) với các mức độ phơi sáng khác nhau của cảnh. Những hình ảnh này sau đó được hợp nhất xây dựng một hình ảnh trung thực như cảnh trong thế giới thực nơi các khu vực rất tối và sáng của nó được hiển thị cùng một lúc. Những hình ảnh như vậy, với chất lượng hình ảnh tuyệt vời, được gọi là hình ảnh dải động cao (HDR) tránh bị chặn dưới và trên các khu vực phơi sáng có thể được cảm nhận trên hình ảnh LDR. Tuy nhiên, những hình ảnh HDR này không thể được hiển thị trên các thiết bị hiển thị LDR tiêu chuẩn do phạm vi dải động của chúng nhỏ hơn so với hình ảnh HDR. Hơn nữa, các thiết bị hiển thị HDR hiện vẫn còn quá đắt. Do đó, nhiều toán tử ánh xạ tông màu hình ảnh (TMO) đã được đề xuất [1]. Mỗi quan tâm chính là giảm phạm vi dải động (độ tương phản, gam màu, chi tiết ...) của hình ảnh HDR xuống phạm vi động của thiết bị hiển thị LDR trong khi duy trì sự xuất hiện của cảnh đã chụp về độ tương phản và ấn tượng chung về độ sáng và màu sắc.

Trong những năm gần đây, một số lượng lớn các nghiên cứu về chủ đề hình ảnh TMO đã được phát triển. Mặc dù mỗi phương pháp TMO có chiến lược cơ bản riêng, nhưng bài báo này không thể xem lại tất cả các công việc đã được phát triển. Tuy nhiên, một danh sách TMO khá đầy đủ trong tài liệu tham khảo [1] trong đó phân loại TMO thành các phương pháp có tính chất như cục bộ, toàn cục, phân vùng và tần số.

Một số TMO được chọn sẽ sử dụng để đánh giá hiệu suất phương pháp, như: Trong [2], Durand và Dorsey đã đề xuất một TMO làm giảm độ tương phản HDR trong khi vẫn giữ được các chi tiết hình ảnh. Công việc này sử dụng bộ lọc song phương bảo toàn cạnh để phân tách hình ảnh HDR thành hai lớp: Lớp cơ sở mã hóa các biến thể quy mô lớn và lớp chi tiết. Độ tương phản sau đó chỉ được giảm ở lớp đầu tiên trong khi các chi tiết được giữ nguyên. Sự kết hợp của các lớp TM này tạo ra hình ảnh LDR. TMO được thực hiện trên miền logarit, xem xét rằng sự khác biệt giữa cường độ điểm ảnh logarit có thể đề cập đến một biện pháp tương phản. Trong [3], Drago và cộng sự trình bày một phương pháp ánh xạ logarit thích nghi của các giá trị độ chói. Nó liên quan đến sự điều chỉnh thích nghi của cơ sở logarit tùy thuộc vào độ rọi của các pixel. Một tập hợp các hàm logarit từ $\log_2$ đến $\log_{10}$ đã được sử dụng để bảo toàn chi tiết cảnh và để cải thiện độ tương phản kết xuất. Hàm trọng lượng được sử dụng để đảm bảo nội suy trơn tru giữa các cơ sở logarit khác nhau. Trong [4], Li và cộng sự đề xuất một kiến trúc băng con liên quan đến một đại diện kim tự tháp Haar. Các hệ số băng con được điều chỉnh lại bằng cách sử dụng chức năng điều khiển khuếch đại làm giảm cường độ tần số cao và tăng cường độ thấp. Các băng con được sửa đổi sau đó được kết hợp với các bộ lọc tổng hợp Haar và được tổng hợp để tái tạo lại hình ảnh LDR cuối cùng. Trong [5], Duan và cộng sự đề xuất một phương pháp tối ưu hóa dựa trên sự điều chỉnh biểu đồ giữa ánh xạ tuyến tính và ánh xạ biểu đồ cân bằng. Trong [6], Fattal và cộng sự đề xuất một thể hệ sóng con thứ hai dựa trên nội dung cạnh của hình ảnh tránh việc có các pixel từ cả hai phía của một cạnh. Cách tiếp cận này sau đó được khai thác để ánh xạ hình ảnh HDR thành hình ảnh LDR.

¹ The University of Danang - University of Science and Technology (Ba Chien Thai, Pham Quoc Thai)

Bài viết này đánh giá hiệu suất của các TMO liên quan đến các họ đa biến phân tách phi tuyến tính với phép nội suy đa thức sử dụng phép nội suy phụ thuộc dữ liệu. Theo hiểu biết của nhóm tác giả, các nhóm đa năng này chưa được khai thác trong chủ đề ánh xạ tông màu hình ảnh. Sự lựa chọn định hướng này đã được thực hiện do các nhóm này có khả năng đưa vào các mô hình toán học liên quan đến các điểm kỳ dị hình ảnh.

2. Phương pháp đề xuất dựa trên nội suy dữ liệu phụ thuộc

Các TMO được đề xuất dựa trên các cách tiếp cận đa biến phi tuyến tính có thể tách rời (1D) dựa trên phép nội suy phụ thuộc dữ liệu. Các cách tiếp cận khai thác chiến lược nội suy cơ bản không dao động (ENO - Essentially Non-Oscillatory) được phát triển bởi Harten [7], [8], [9], [10], trong đó họ đột biến biến đổi giá trị điểm (Point-Value, PV) và nhóm đa biến đổi tế bào (Cell-Average, CA) xem xét. Phần này đưa ra các toán tử dự báo ENO 1D sẽ được khai thác bởi các toán tử ánh xạ tông màu hình ảnh của nhóm tác giả. Các nhóm này có khả năng xem xét trong mô hình toán học các điểm kỳ dị biệt lập như các điểm cạnh trong ảnh, do đó tránh hiện tượng Gibbs đặc biệt có hại trong các hình ảnh ánh xạ tông màu.

2.1. Họ đa phân giải giá trị điểm ENO

Trong họ này, ký hiệu là ENO-PV, vector rời rạc $v^j := (v_k^j)_{k \in \mathbb{Z}}$ được coi là giá trị điểm của hàm liên tục v trên lưới $\Gamma^j := (2^{-j}k)_{k \in \mathbb{Z}}$, tức là $v_k^j = v(2^{-j}k)$. Toán tử chiếu D_j^{j-1} là toán tử lấy mẫu xuống (downsampling), tức là:

$$v_k^{j-1} = v_{2k}^j \quad (1)$$

Để điều chỉnh dự đoán gần với các điểm kỳ dị của dữ liệu, Harten đề xuất sử dụng kỹ thuật nội suy ENO. Ở cấp độ phân giải j , các giá trị dự đoán $(\hat{v}_{2k+1}^{j-1})_{k=1..Nj}$ được tính gần đúng bằng cách sử dụng các giá trị v_k^{j-1} từ khuôn mẫu dự đoán (prediction stencil) có độ dài $2M - 1$ được xác định là:

$$S(k, r) := \{(k + 1 - r - M)2^{j-1}, \dots, (k - r + M)2^{j-1}\} \quad (2)$$

với, r là một số nguyên trong $[-M + 1, M - 1]$, tương ứng với vị trí của khuôn mẫu đối với k . Một đa thức $p_{k,r}$ bậc $2M - 1$ được định nghĩa để nội suy giá trị của v trên $S(k, r)$ trong số các tập đa thức $p_{k,r}$. Toán tử phỏng đoán sau đó được định nghĩa là toán tử nội suy phụ thuộc dữ liệu. Các giá trị dự đoán sau đó được đưa ra bởi:

$$\hat{v}_{2k+1}^{j-1} = p_{k,r^*}((2k + 1)2^{j-1}) \quad (3)$$

Trong đó, tham số r^* được liên kết với đa thức $p_{k,r}$ giới thiệu dao động nhỏ nhất xung quanh vùng lân cận của k theo hàm chi phí $C^{j-1}(S(k, r)) = \sum_{l \in S(k, r)} |v_{l+1} - v_l|$ cho bởi nghiệm phương trình:

$$r^* = \arg \min_{-M+1 \leq r \leq M-1} (C^{j-1}(S(k, r))) \quad (4)$$

Các giá trị dự đoán, $\hat{v}_{2k+1}^{j-1}$ cho $r = [-M + 1, M - 1]$, là cung cấp cho đa thức Lagrange $(2M - 1)$ ngũ phân, thất phân hoặc cửu phân, tương ứng với $M = 3, 4$ hoặc 5 .

2.2. Họ đa phân giải trung bình ô ENO

Trong họ này, được ký hiệu là ENO-CA, vector rời rạc $v^j := (v_k^j)_{k \in \mathbb{Z}}$ là được coi là giá trị trung bình của một hàm liên tục gấp khúc v trên lưới $\Gamma_k^j := \{[2^{-j}k, 2^{-j}(k + 1)], k \in \mathbb{Z}\}$. Các lưới ở đây được xác định bằng cách sử dụng các khoảng dyadic có dạng $I_k^j = [2^{-j}k, 2^{-j}(k + 1)]$. Toán tử D_j^{j-1} trở thành trong ngữ cảnh này là toán tử tính trung bình:

$$\hat{v}_k^{j-1} = \frac{1}{2}(v_{2k}^j + v_{2k+1}^j) \quad (5)$$

Như trong trường hợp PV, khuôn mẫu dự đoán $S(k, r)$ cho bởi:

$$S(k, r) := \{I_{k+1-r-M}^{j-1}, \dots, I_{k-1-r+M}^{j-1}\} \quad (6)$$

Chiều dài khuôn mẫu dự đoán vẫn là $2M - 1$. Một đa thức $p_{k,r}$ bậc $2M - 2$ được xác định để nội suy các giá trị trung bình trên $S(k, r)$. Theo cùng một chiến lược ENO được xác định trong trường hợp PV, giữa tập đa thức $p_{k,r}$ với $r = [-M + 1, M - 1]$, với p_{k,r^*} được chọn là đa thức dao động nhỏ nhất tương ứng với r^* . Các giá trị dự đoán sau đó được suy ra từ: $\hat{v}_{2k+1}^{j-1} = \int_{I_{2k+1}^j} p_{k,r^*}(t) dt$ (7)

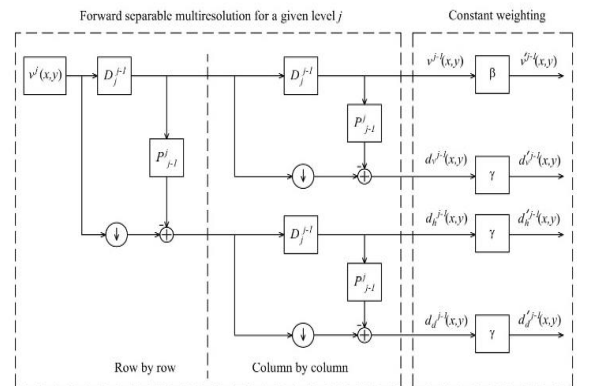
Đa thức nội suy Lagrange $(2M - 2)$ tứ phân, lục phân hoặc bát phân được sử dụng ($M = 3, 4$ hoặc 5).

2.3. Toán tử ánh xạ tông màu sử dụng phép nội suy phụ thuộc dữ liệu

Phần này trình bày cách tiếp cận được đề xuất ánh xạ hình ảnh HDR sang hình ảnh LDR. Toán tử được chọn dựa trên phần mở rộng của các họ đa phân giải phi tuyến tính 1D, được mô tả trong các phần trước, sử dụng phương pháp tiếp cận sản phẩm tensor cổ điển. Giả sử rằng hình ảnh HDR có kích thước $N \times M$ pixel. Ký hiệu T là TMO tương ứng với ENO5-PV, ENO5-CA, ENO7-PV, ENO7-CA, ENO9-PV hoặc ENO9-CA. Nó được thực hiện trên phép biến đổi logarit của các giá trị Luma L_w của hình ảnh HDR. Thật vậy, mục tiêu của phép biến đổi logarit là tính đến tất cả các biến thể nhỏ trong cảnh. Trong số các phép biến đổi logarit có thể có, chức năng sau được chọn:

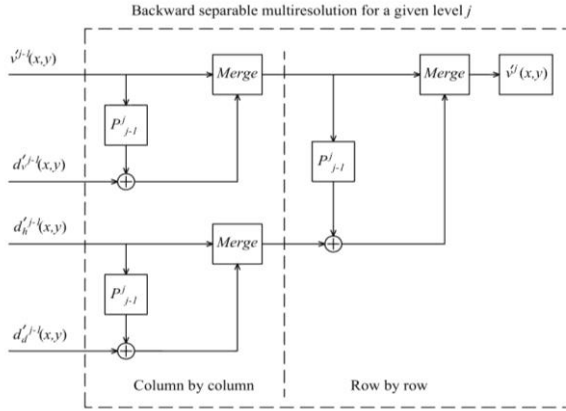
$$D = \log_{10}(L_w + \varepsilon); \quad (8)$$

ε là giá trị dương được thêm vào để tránh xa các trường hợp kỳ dị logarit. Luma được biến đổi sau đó được phân rã theo sơ đồ được cung cấp bởi Hình 1.



Hình 1. Sự phân hủy tách biệt (ENO-PV, ENO-CA) và hiệu chỉnh các hàm trọng lượng hằng số

Quy trình được áp dụng đầu tiên trên các dòng và sau đó các cột của hình ảnh. Ở mức độ phân giải J , tập hợp các hệ số xấp xỉ v^0 và các hệ số chi tiết $\{d^0, d^1, d^2, \dots, d^{j-1}, \dots, d^{J-1}\}$ thu được. Lưu ý rằng, d^{j-1} bao gồm 3 khối (d_v^{j-1} , d_h^{j-1} , d_d^{j-1}) ở dạng cách khôi phục không dư thừa. Các hệ số này được nhân tỷ lệ như sau: $\beta \times v^0$ với $0 \leq \beta \leq 1$; và $\gamma \times \{d^0, d^1, d^2, \dots, d^{j-1}, \dots, d^{J-1}\}$ với $0 \leq \gamma \leq 1$.



Hình 2. Sự khôi phục tách biệt (ENO-PV, ENO-CA)

3. Kết quả mô phỏng

Phần này so sánh chất lượng của ảnh xạ tổng màu hình ảnh bằng cách sử dụng các họ phân giải được mô tả trong các phần trước. Chất lượng ảnh xạ tổng màu hình ảnh được đo với chỉ số chất lượng hình ảnh ảnh xạ tổng màu (Tone Mapped image Quality Index, TMQI) bao gồm 2 chỉ số đánh giá đơn cộng lại: Đánh giá về sự tin cậy về cấu trúc ảnh (Structural Fidelity) và đánh giá có tính thống kê về tính tự nhiên của bức ảnh (Statistical Naturalness) được phát triển trong [13]. Chỉ số này đánh giá hình ảnh LDR bằng hình ảnh HDR gốc. Lưu ý rằng, số đo TMQI có giới hạn trên giá trị 1. Mô phỏng đã được tiến hành phần mềm Matlab môi trường sử dụng hộp công cụ HDR [1], với các tên hình ảnh HDR thử nghiệm như "BottleSmall", "Light", "SmallOffice" và "AtriumNight". Các tham số khác nhau được chọn để đưa ra kết quả tốt nhất về chỉ số TMQI trong tất cả các phương pháp.

Sáu toán tử TMO ENO5-PV, ENO5-CA, ENO7-PV, ENO7-CA, ENO9-PV và ENO9-CA được sử dụng với các tham số sau $\beta = 0,3$; $\gamma = 0,3$; $J = 2$. Chúng được so sánh đến :

- Fattal [6] sử dụng phương pháp WCDF và RBW với các tham số sau $\alpha = 0,8$; $\beta = 0,3$; $\gamma = 0,7$; $J = 2$; $upd = 0$;
- Drago [3], Reinhard [11], Ward [12], Durand [2] với các tham số mặc định như đã cho trong Hộp công cụ HDR và Duan [5] sử dụng $\beta = 0,3$.

Bảng 1 cung cấp các chỉ số chất lượng ảnh xạ tổng màu hình ảnh TMQI theo các phương pháp khác nhau này. Hiệu suất của các họ đa phân giải phi tuyến tính là cạnh tranh với Fattal [8]. Cũng lưu ý rằng, các giải pháp của nhóm tác giả ít tốn kém hơn về mặt tính toán so với WCDF yêu cầu tính toán và lưu trữ hệ số trọng lượng cho bước tái tạo. Thử nghiệm với tên ảnh HDR "Light" sử dụng các toán tử RBW, WCDF, ENO5-CA, ENO5-PV, ENO7-CA, ENO7-PV,

ENO9-CA và ENO9-PV lần lượt được đưa ra trong Hình 3, Hình 4, Hình 5, Hình 6, Hình 7, Hình 8, Hình 9 và Hình 10. Những hình ảnh này có chất lượng hình ảnh tương tự.

Bảng 1. Chỉ số chất lượng tổng màu hình ảnh (TMQI)

Phương pháp TMO	Tên ảnh HDR thử nghiệm			
	Bottle Small	Light	Small Office	Atrium Night
Drago [3]	0,801	0,801	0,814	0,799
Reinhard [11]	0,807	0,826	0,789	0,801
Ward [12]	0,783	0,775	0,817	0,797
Durand [2]	0,892	0,825	0,814	0,929
Duan [5]	0,915	0,955	0,986	0,964
FattalWCDF [6]	0,969	0,920	0,914	0,941
FattalRBW [6]	0,972	0,920	0,916	0,944
ENO5-CA	0,941	0,925	0,989	0,962
ENO5-PV	0,926	0,914	0,976	0,928
ENO7-CA	0,947	0,912	0,956	0,927
ENO7-PV	0,989	0,945	0,952	0,91
ENO9-CA	0,960	0,940	0,964	0,913
ENO9-PV	0,991	0,943	0,941	0,959



Hình 3. Fattal's RBW ($\alpha = 0,8$; $\beta = 0,3$; $\gamma = 0,7$; $J=2$; $upd=0$)



Hình 4. Fattal's WCDF ($\alpha = 0,8$; $\beta = 0,3$; $\gamma = 0,7$; $J=2$; $upd=0$)



Hình 5. ENO5-CA ($\beta = 0,3; \gamma = 0,3; J=2$)



Hình 9. ENO9-CA ($\beta = 0,3; \gamma = 0,3; J=2$)



Hình 6. ENO5-PV ($\beta = 0,3; \gamma = 0,3; J=2$)



Hình 10. ENO9-PV ($\beta = 0,3; \gamma = 0,3; J=2$)



Hình 7. ENO7-CA ($\beta = 0,3; \gamma = 0,3; J=2$)



Hình 11. Drago



Hình 12. Reinhard



Hình 8. ENO7-PV ($\beta = 0,3; \gamma = 0,3; J=2$)



Hình 13. ENO5-CA ($\beta = 0,3; \gamma = 0,3; J=2$)



Hình 14. ENO5-CA ($\beta = 0,3; \gamma = 0,3; J=2$)



Hình 15. Khả năng khôi phục cạnh ngang của cầu thang bằng phương pháp ENO5-CA ($\beta = 0,3; \gamma = 0,3; J=2$)

4. Kết luận

Bài báo này đánh giá hiệu suất của các toán tử ánh xạ tông màu bắt nguồn từ các họ đa phân giải phi tuyến tính tách biệt. Các nhóm này, dựa trên phép nội suy phụ thuộc vào dữ liệu, đưa vào mô hình toán học của chúng các điểm kỳ dị của hình ảnh HDR. Hiện tượng Gibbs ảnh hưởng đến chất lượng hình ảnh của hình ảnh ánh xạ tông màu sau đó bị giảm xuống. Hơn nữa các cạnh dọc và ngang cũng được hiển thị tốt như Hình 15. Các kết quả mô phỏng

có tính cạnh tranh với các toán tử ánh xạ tông màu hình ảnh hiện có.

Lời cảm ơn: Nghiên cứu này được tài trợ bởi Quỹ Phát triển Khoa học và Công nghệ - Đại học Đà Nẵng trong đề tài có mã số B2019-DN02-51.

TÀI LIỆU THAM KHẢO

- [1] Banterle, F., Artusi, A., Debattista, K., và Chalmers, A., Advanced High Dynamic Range Imaging, CRC Press, ISBN: 978-156881-719-4, 2011.
- [2] Durand, F., and Dorsey, J., "Fast bilateral filtering for the display of high-dynamic-range images", *ACM Transactions on Graphics*, 2002, 257-266.
- [3] Drago, F., Myszkowski, K., Annen, T., and Chiba, N., "Adaptive logarithmic mapping for displaying high contrast scenes", *Computer Graphics Forum* 22, 2003, 419-426.
- [4] Li, Y., Sharan, L., and Adelson, E., "Compressing and companding high dynamic range images with subband architectures", *ACM Transaction on Graphics* 24, 2005, 836-844.
- [5] Duan, J., Bressan, M., Dance, C., and Qiu, G., "Tone-mapping high dynamic range images by novel histogram adjustment", *Pattern Recognition*, 2010, 1847-1862.
- [6] Fattal, R., "Edge-Avoiding Wavelets and their Applications", *ACM Transaction on Graphics*, 2009.
- [7] Harten, A., "Multiresolution representation of data: a general framework", *SIAM J. Numerical Analysis*. 33, 1996, 1205-1256.
- [8] Harten, A., Engquist, B., Osher, S., and Chakravarthy, S., "Uniformly high order accurate essentially non-oscillatory schemes III", *Journal of Computer*, 1987, 231-303.
- [9] Amat, S., Arandiga, F., Cohen, A., Donat, R., Garcia, G., and Oehsen, M.V., "Data compression with ENO schemes : A case study", *Applied and Computational Harmonic Analysis*, 2001, 273-288.
- [10] Amat, S., Donat, R., Liandrat, J. and Trillo; J.C. "A fully adaptive PPH multiresolution scheme for image processing", *Mathematical and Computer Modelling*, 2006.
- [11] Reinhard, E., and Devlin, K., "Dynamic range reduction inspired by photoreceptor physiology", *IEEE Transactions on Visualization and Computer Graphics* 11, 2005, 13-24.
- [12] Ward, G., Rushmeier, H., and Piatko, C., "A visibility matching tone reproduction operator for high dynamic range scenes", *IEEE Transactions on Visualization and Computer Graphics* 3, 1997, 291-306.
- [13] Yeganeh, H. and Wang, Z., "Objective quality assessment of tonemapped images", *IEEE Transaction on Image Processing*, 2013, 657-667.