

多重局所トーンマッピングオペレータを用いたHDR画像の顕著性マップ

岡崎, 大暉
九州大学大学院芸術工学府

原, 健二
九州大学大学院芸術工学研究院

井上, 光平
九州大学大学院芸術工学研究院

浦浜, 喜一
九州大学大学院芸術工学研究院

<https://hdl.handle.net/2324/1959209>

出版情報 : The IEICE transactions on information and systems. J101-D (11), pp.1494-1498, 2018-11. 電子情報通信学会
バージョン :
権利関係 : ©一般社団法人電子情報通信学会2018

研究速報

多重局所トーンマッピングオペレータを用いた HDR 画像の顕著性マップ

岡崎 大暉^{†a)} (学生員) 原 健二[†] (正員)
井上 光平[†] (正員) 浦浜 喜一[†] (正員)

Saliency Maps of HDR Images Using Multiple Local Tone Mapping Operators

Daiki OKAZAKI^{†a)}, Student Member, Kenji HARA[†], Kohei INOUE[†], and Kiichi URAHAMA[†], Members

[†]九州大学大学院芸術工学研究院, 福岡市

Department of Communication Design Science, Kyushu University, 4-9-1 Shiobaru, Minami-ku, Fukuoka-shi, 815-8540 Japan

a) E-mail: 2DS17097P@s.kyushu-u.ac.jp

DOI:10.14923/transinfj.2018JDL8009

あらまし 本論文では, HDR 画像から異なる輝度領域が強調された複数枚の LDR 画像を仮想的に生成し用いることで, LDR 画像を対象とする既存の顕著性マップ生成手法を HDR 画像にも適用できるように拡張する. 実際の HDR 画像を用いた実験を通して, 提案手法の有効性を示す.

キーワード 顕著性マップ, HDR 画像, トーンマッピングオペレータ

1. まえがき

顕著性マップは, 注視の集まりやすさを画素ごとに数値化したもので, 物体検出, カラーイメージセグメンテーション, 視線解析, ロボットビジョンなど様々な分野への応用が期待されており, 盛んに研究が行われてきている [1]~[4]. 本論文では, 高ダイナミックレンジ (High Dynamic Range, HDR) 画像の顕著性マップを推定する問題を扱う. HDR 画像は, 人間の視覚特性 (Human Visual System, HVS) と同程度のダイナミックレンジを保持した画像のことで, CG 分野における高品質なレンダリングに始まり, 車載カメラや監視カメラ, 医用画像など, 汎用性の高さから様々な分野で利用されている. HDR 画像を一般的な低ダイナミックレンジ (Low Dynamic Range, LDR) ディスプレイ上に表示する際には, トーンマッピング処理, すなわちトーンマッピングオペレータ (Tone Mapping Operator, TMO) を用いて HDR 画像のダイナミックレンジを LDR 画像の 1 画素あたり RGB 各 8 ビット・256 階調に圧縮する処理が必要になる. これまで顕著性マップ生成手法が多数提案されているが, これらの既存手法の多くは, 一般的な LDR 画像を対象としており, 10 ビット以上の高い階調数をもつ HDR 画像に

対しては性能が発揮できない恐れがある. また, 微細なコントラストを表現するための 1 チャンネル 10 ビット以上の階調データに対応可能な表示機器は十分に普及しておらず, HDR 画像を用いた視線計測による学習データの収集は難しい. そのため, HDR 画像を対象とする場合, 深層学習などのデータ分析技術の利用は困難である.

本論文では, 通常の LDR 画像を対象とする既存の顕著性マップ生成手法を HDR 画像にも適用できるように拡張する. 提案手法では, まず HDR 画像に位相の異なる複数の TMO を用いてトーンマッピング処理を施して複数枚の異なる輝度領域が強調された LDR 画像を生成する. 次に, 各 LDR 画像に既存の顕著性マップ生成手法を適用して得られた複数枚の多重露光顕著性マップを凸最適化で 1 枚の顕著性マップに統合する. 本手法の利点は, 任意の顕著性マップ生成手法の既存コードを修正することなくそのまま再利用できることである. 提案手法を実際の HDR 画像に適用し, その有効性を示す.

2. 関連研究

これまで多くの顕著性マップ生成手法が報告されている. 例えば, Itti らの特徴理論に基づく顕著性マップの計算モデル [1] に始まり, 近年提案されている学習ベースのアプローチによるモデル [2] に至るまで数多くの手法が存在する [3], [4]. しかし, これらの手法はいずれも LDR 画像を対象としたものであり, HDR 画像にそのまま直接適用すると, HDR 画像の広いダイナミックレンジに対応できず, 正確な顕著性マップが得られないことが多い [5].

近年, Bremond らは, Itti らの手法 [1] を HDR 画像に拡張し, 10 ビット以上の階調数をもつ HDR 画像に対しても正確な顕著性マップを生成する手法を提案している. しかし, 彼らの手法は, HDR 画像の輝度値の高い明部領域における顕著領域検出に主眼が置かれており, 暗部領域における推定精度は高くない. これに対し, 本論文では, HDR 画像の明部領域に加えて暗部領域においても顕著領域を正確に検出する手法を提案する.

3. 提案手法

提案手法の詳細について述べる. 図 1 に本手法の概要を示す. 提案手法は, 図 1(a) に示すような HDR 画像を入力とする.

3.1 多重局所トーンマッピング

まず, 入力として与えられた HDR 画像に対し, 複

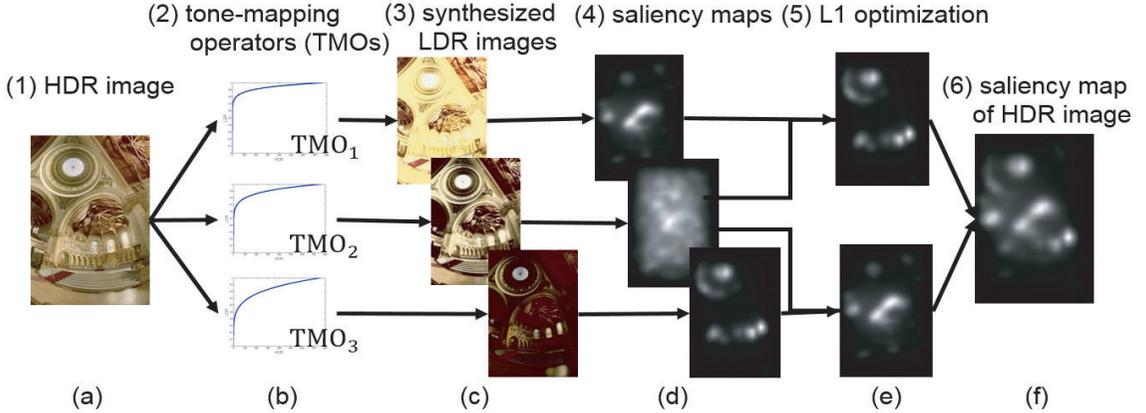


図1 提案手法の手順
Fig. 1 Flow of our method.

数の TMO によるトーンマッピング処理を行う。HDR 画像の代表的な生成方法として、露光量を変えて撮影された複数枚の画像から 1 枚の HDR 画像を合成する手法がある。例えば、低い露光量で撮影された低露光画像、高い露光量で撮影された高露光画像、その中間である中露光画像の 3 枚を 1 枚の HDR 画像に合成する。これに対し提案手法では、この HDR 画像生成の入出力を逆にして、階調変換により強調される輝度範囲が互いに異なる三つのトーンマッピングオペレータ TMO₁, TMO₂, TMO₃ (図 1 (b)) を用いて、仮想的な高露光画像 I_1 , 中露光画像 I_2 , 低露光画像 I_3 をそれぞれ生成する (図 1 (c))。これらの多重露光画像 I_1, I_2, I_3 の各々に対して、LDR 画像を対象とする既存の顕著性マップ生成手法を適用して顕著性マップ V_1, V_2, V_3 をそれぞれ得る (図 1 (d))。各 TMO の推定は、次式の Drago らの手法 [6] を用いる。

$$L_d = \frac{L_{d_{max}} \cdot 0.01}{\log_{10}(L_{max} + 1)} \cdot \frac{\log_{10}(L + 1)}{\log_{10}\left(2 + 8\left(\left(\frac{L}{L_{max}}\right)^{\frac{\log_{10}(b)}{\log_{10}(0.5)}}\right)\right)} \quad (1)$$

ここで、 L は HDR 画像の輝度値、 L_{max} は HDR 画像の最大輝度値、 L_d はトーンマッピング処理後の LDR 画像の輝度値、 $L_{d_{max}}$ は LDR 画像の最大輝度値、 b は明度領域のレンジ圧縮性と暗部領域でのコントラストを制御するパラメータである。式 (1) のパラメータ b を適切に設定することで、各 TMO の推定を精度よく行う。パラメータ b の設定方法は、トーンマッピング処理により変換した LDR 画像における白とび領域

や黒潰れ領域が本来の顕著領域を含まないように決定される。例えば TMO₁ の場合、はじめに初期値 $b^{(0)}$ を用いて生成した高露光画像 I_1 の顕著性マップ $V_1^{(0)}$ を生成する。次に、この高露光画像 I_1 の白とび領域における顕著性マップ $V_1^{(0)}$ の顕著度の平均値 $\mu^{(0)}$ を計算する。更に、パラメータ b を次式に基づき更新する。

$$b^{(k+1)} = \begin{cases} b^{(k)} & \text{if } \mu^{(k)} \leq \frac{S^{(k)}}{N} \\ b^{(k)} - \Delta b & \text{if } \mu^{(k)} > \frac{S^{(k)}}{N} \end{cases} \quad (2)$$

ここで、 $k = 0, 1, \dots$ は更新ステップ、 $S^{(k)}$ は更新ステップ k における顕著性マップ $V_1^{(k)}$ の画素値の合計、 N は画素数である。更新ステップ k における平均値 $\mu^{(k)}$ が、 $S^{(k)}/N$ より大きい場合、高露光画像の白とびさせた領域が顕著性を多く含んでいるため、 $b^{(k)}$ を $b^{(k)} - \Delta b$ とおいて白とびの領域を拡大させ、 $S^{(k)}/N$ 以下になるまで $k \leftarrow k + 1$ として式 (2) の計算を収束するまで繰り返すことで TMO₁ のパラメータ b を決定する。TMO₃ においては、式 (2) で平均値 $\mu^{(k)}$ が $S^{(k)}/N$ より大きい場合、 $b^{(k)}$ を $b^{(k)} + \Delta b$ とおいて黒潰れの領域を拡大させること以外は同様に行う。提案手法では、 $\Delta b = 0.01$ とした。TMO₂ のパラメータ b は、TMO₁ のパラメータ値と TMO₃ のパラメータ値の中間値として決定する。

3.2 凸最適化を用いた多重顕著性マップ統合

顕著性マップ V_1, V_2, V_3 を 1 枚の顕著性マップに統合する。高露光画像 I_1 の顕著性マップ V_1 は HDR 画像の暗部領域の顕著性、低露光画像 I_3 の顕著性マップ V_3 は HDR 画像の明部領域の顕著性を正しく検出している点に着目し、中露光画像 I_2 の顕著性マップ V_2

の顕著度レベルを基に V_1 と V_3 をトーンカーブ (Tone Curve, TC) で階調補正することを考える. それぞれの TC を TC_{high} , TC_{low} とおいて, 次式の $L1$ ノルムの最適化問題を解くことにより TC_{high} と TC_{low} を推定する.

$$\begin{aligned} \min_{\alpha_1, \alpha_2} & \| \mathbf{V}'_2 - \alpha_1 \mathbf{V}'_1 - \alpha_2 \mathbf{1} \|_1 \\ \text{s.t.} & \alpha_1 > 0, 0 \leq \alpha_2 \leq t \end{aligned} \quad (3)$$

$$\begin{aligned} \min_{\beta_1, \beta_2} & \| \mathbf{V}'_2 - \beta_1 \mathbf{V}'_3 - \beta_2 \mathbf{1} \|_1 \\ \text{s.t.} & \beta_1 > 0, 0 \leq \beta_2 \leq t \end{aligned} \quad (4)$$

ここで, $\mathbf{V}'_1 \in \mathbb{R}^L$, $\mathbf{V}'_3 \in \mathbb{R}^L$ は各露光画像の顕著性マップ V_1, V_3 の一定値以上の顕著性をもつ L 次元ベクトル, $\mathbf{V}'_2 \in \mathbb{R}^L$ は V_1, V_3 の一定値以上の顕著性をもつ位置と同じ位置の V_2 の L 次元ベクトル, $\mathbf{1}$ は要素が全て 1 の L 次元ベクトル, L は V_1, V_2, V_3 の一定値以上の顕著性をもつ画素数, α_1, β_1 はそれぞれ TC_{high} , TC_{low} の一次係数, α_2, β_2 はそれぞれ TC_{high} , TC_{low} の切片である. ここで, 一次係数 α_1, β_1 が正の値となるような制約を課しているのは, 各 TC が単調増加関数でなければならないためである. また, 切片 α_2, β_2 が負の値や極端に大きくなることを防ぐために, 0 より大きく, あるしきい値 t よりも小さい値となるような制約を与える. 次章の実験では $t = 0.1$ とおいた. 式 (3), (4) をパラメータ $\alpha_1, \alpha_2, \beta_1, \beta_2$ に関して解くことで, V_1, V_3 を最適に補正することが可能である. このようにして得られた TC_{high} , TC_{low} を図 2 に示す. これらの TC を用いて, 顕著性マップ V_1, V_3 を補正した後, 画像内の最も小さい画素値が 0, 最も大きい画素値が 1 となるように画像の正規化を行う (図 1(e)). 最後に, 得られた 2 枚の顕著性マップの和をとり, 同様に画像の正規化を行うことで HDR 画像の顕著性マップが生成される (図 1(f)).

4. 実験結果

提案手法を実際の HDR 画像に適用した結果を示す. 本実験では, Radiance の RGBE フォーマット (拡張子.hdr) を使用した. 入力に用いた RGBE フォーマットの HDR 画像を図 3(a), (h), (o) に示す. ただし, HDR 画像にトーンマッピング処理を施して, 8 ビットに変換した LDR 画像を表示している. 提案手法のトーンマッピング処理により得られた各入力画像に対して直接 (1) Itti らの手法 [1] を適用した場合 (以下, Itti) と (2) Bremond らの手法 [5] を適用した場合 (以下, Bremond) を従来手法として, 提案手法における

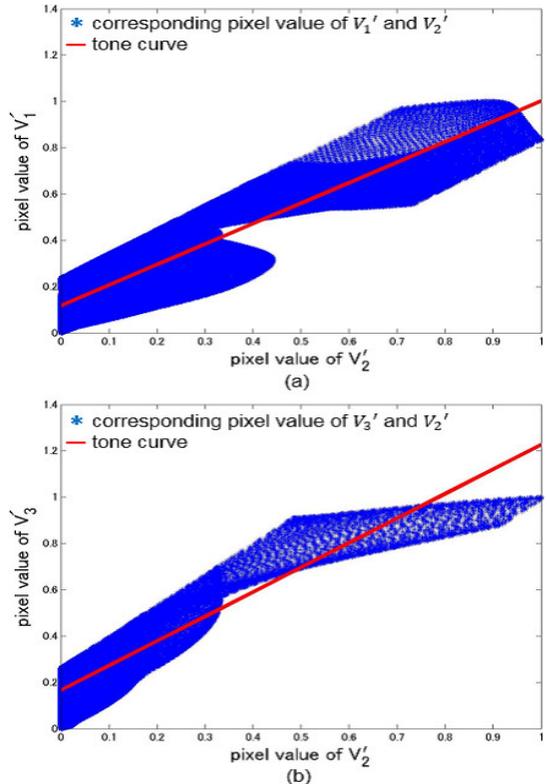


図 2 顕著度補正のためのトーンカーブ. (a) TC_{high} , (b) TC_{low}

Fig. 2 TC for saliency correction. (a) TC_{high} , (b) TC_{low} .

各仮想露光画像に (3) Itti らの手法 [1] を適用した場合 (以下, ours-Itti), (4) Harel らの手法 [2] を適用した場合 (以下, ours-Harel), (5) Fang らの手法 [3] を適用した場合 (以下, ours-Fang), (6) Tavakoli らの手法 [4] を適用した場合 (以下, ours-Tavakoli) との性能比較を行った.

4.1 定性的比較

従来手法で得られた顕著性マップをそれぞれ図 3 (b), (c), (i), (j), (p), (q) に示す. 提案手法を適用して得られた顕著性マップをそれぞれ図 3 (d)~(g), (j)~(n), (r)~(u) に示す. 従来手法では, HDR 画像内における明部領域の顕著性が正しく評価されている一方で, 暗部領域の顕著性は低く評価されていることが分かる. 例えば, 図 3 (i), (j) では, 背景の明るい光の影響を受けてしまっており, 暗部領域にある車の顕著性を正確に評価できていない. これは, 背景の光と車では輝度値の差が大きく, 単純な明度やコントラストの特徴

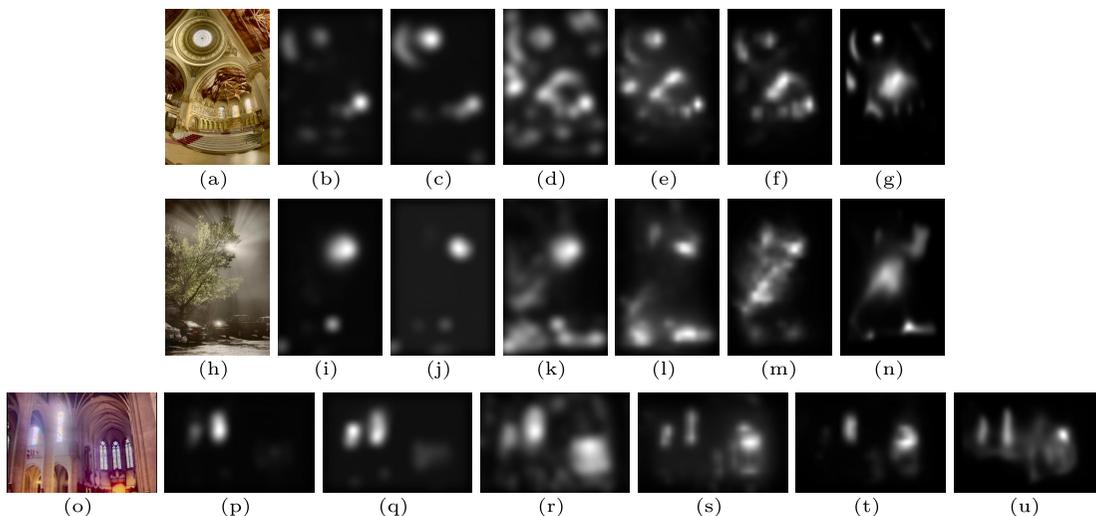


図3 顕著性マップ生成. (a) HDR 画像, (b) Itti, (c) Bremond, (d) ours-Itti, (e) ours-Harel, (f) ours-Fang, (g) ours-Tavakoli, (h) HDR 画像, (i) Itti, (j) Bremond, (k) ours-Itti, (l) ours-Harel, (m) ours-Fang, (n) ours-Tavakoli, (o) HDR 画像, (p) Itti, (q) Bremond, (r) ours-Itti, (s) ours-Harel, (t) ours-Fang, (u) ours-Tavakoli

Fig. 3 Saliency map generation. (a) HDR image, (b) Itti, (c) Bremond, (d) ours-Itti, (e) ours-Harel, (f) ours-Fang, (g) ours-Tavakoli, (h) HDR image, (i) Itti, (j) Bremond, (k) ours-Itti, (l) ours-Harel, (m) ours-Fang, (n) ours-Tavakoli, (o) HDR image, (p) Itti, (q) Bremond, (r) ours-Itti, (s) ours-Harel, (t) ours-Fang, (u) ours-Tavakoli.

量では輝度値の大きい方に影響されるためと考えられる。一方、提案手法では暗部領域の顕著性も他の領域同様に正確に評価できている。

4.2 定量的比較

提案手法の定量評価の結果を示す。比較手法として Itti [1] と Bremond [5]，提案手法として ours-Itti を用いる。実験では、カラー画像と被験者の視線方向分布をもとに得られる視線の停留マップのペアからなる公開データセット [7] を用いる。ここで、データセットの画像は LDR 画像であるため、その LDR 画像から輝度レベルの異なる複数枚の画像を生成し、これらを MATLAB の makehdr 関数で統合することにより生成された HDR 画像を入力とする。また、評価指標には、ランダムな視線移動に対する有意差を測定する尺度である NSS (Normalized Saliency Scanpath) [8] を用いる。入力画像と停留マップの 20 個のデータセットを用いて、各手法の NSS を評価した結果を表 1 に示す。ここで、表中の ± の前の値は NSS の平均、± の後の値は NSS の標準偏差である。表 1 より、従来手法と比べて提案手法が優位な結果が得られていることが分かる。

表 1 定量的比較

Table 1 Quantitative comparison.

Criterion	Method		
	Itti [1]	Bremond [5]	ours-Itti
NSS	1.09±0.0054	1.23±0.0060	1.54±0.0078

5. む す び

複数の異なるトーンマッピングオペレータを用いて既存の顕著性マップ生成手法を HDR 画像に適用できるように拡張する手法を提案した。実験では、従来手法と比べて広ダイナミックレンジに対応した良好な結果が得られ、提案手法の有効性が示された。

謝辞 本研究は、科学研究費補助金（課題番号：15K00239）によって行われたものです。

文 献

- [1] L. Itti, C. Koch, and E. Niebur, “A model of saliency based visual attention for rapid scene analysis,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.20, no.11, pp.1254–1259, 1998.
- [2] J. Harel, C. Koch, and P. Perona, “Graph-based visual saliency,” *NIPS*, vol.19, pp.545–552, 2006.
- [3] S. Fang, J. Li, and Y. Tian, “Learning discriminative subspaces on random contrasts for image saliency

- analysis,” *IEEE Trans. Neural Netw. Learning Syst.*, vol.28, no.5, pp.1095–1108, 2016.
- [4] H.R. Tavakoli, E. Rahtu, and J. Heikkila, “Fast and efficient detection using sparse sampling and kernel density estimation,” *Proc. SCIA*, pp.666–675, 2011.
- [5] R. Bremond, J. Petit, and J.P. Tarel, “Saliency maps of high dynamic range images,” *Media Retargeting Workshop in Conjunction with ECCV ’10*, pp.118–130, 2010.
- [6] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, “Adaptive logarithmic mapping for displaying high contrast scenes,” *Computer Graphics Forum*, vol.3, pp.419–426, 2003.
- [7] T. Judd, K. Ehinger, F. Durand, and A. Torralba, “Learning to predict where humans look,” <http://people.csail.mit.edu/tjudd/WherePeopleLook/index.html>, 2009.
- [8] S. Marat, A. Rahman, D. Pellerin, N. Guyader, and D. Houzet, “Improving visual saliency by adding ‘face feature map’ and ‘center bias’,” *Cognitive Computation*, vol.5, pp.63–75, 2013.
- (平成 30 年 3 月 16 日受付, 6 月 25 日再受付,
7 月 23 日早期公開)