

高ダイナミックレンジ画像のための顕著性マップ推定手法

岡崎, 大暉
九州大学大学院芸術工学研究院

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/2324/1956517>

出版情報：情報科学技術フォーラム講演論文集．17（第3分冊），pp.151-152，2018-09-19．電子情報通信学会
バージョン：
権利関係：利用は著作権の範囲内に限られます



高ダイナミックレンジ画像のための顕著性マップ推定手法 Saliency Map Generation of High Dynamic Range Images

岡崎 大暉†
Daiki Okazaki

原 健二†
Kenji Hara

井上 光平†
Kohei Inoue

浦浜 喜一†
Kiichi Urahama

1. はじめに

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、視線解析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。本報告では、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、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。これまで多くの顕著マップ生成手法が報告されている [1, 2]。しかし、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持つ 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、deep learning などの学習アプローチは困難である。

本報告では、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. 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。

2.1. 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、複数の TMO によるトーンマッピング処理を行う。提案手法では、階調変換により強調される輝度範囲が

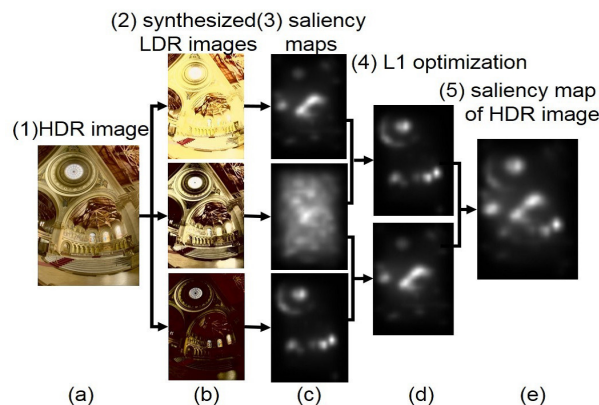


図 1: 提案手法の手順

互いに異なる 3 つのトーンマッピングオペレータ TMO_1, TMO_2, TMO_3 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(b))。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(c))。各 TMO の推定は、次式の Drago らの手法 [5] を用いる。

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

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 TMO_1 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。さらに、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 $b^{(k)} - \Delta b$

†九州大学大学院芸術工学研究院

とに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 TMO_1 のパラメータ b を決定する。 TMO_3 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 $b^{(k)} + \Delta b$ とに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。 TMO_2 のパラメータ b は、 TMO_1 のパラメータ値と TMO_3 のパラメータ値の中間値として決定する。

2.2. 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_2 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_3 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 $\text{TC}_{\text{high}}, \text{TC}_{\text{low}}$ とにおいて、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。

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

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

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 $\text{TC}_{\text{high}}, \text{TC}_{\text{low}}$ の一次係数、 α_2, β_2 はそれぞれ $\text{TC}_{\text{high}}, \text{TC}_{\text{low}}$ 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、ある閾値 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、画像の正規化を行う (図 1(d))。最後に、得られた 2 枚の顕著性マップの和をとり、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(e))。

3. 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。入力に用いた HDR 画像を図 2(a),(f) 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 [4] 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) との性能比較を行った。

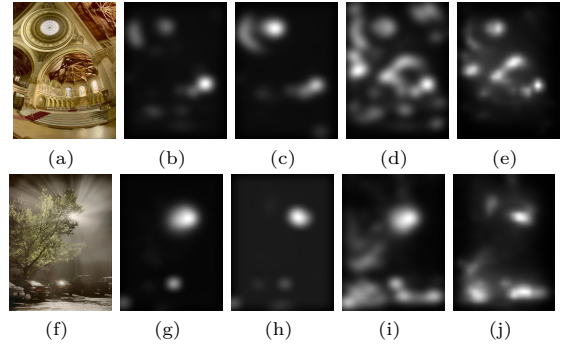


図 2: 顕著性マップ生成. (a)HDR 画像, (b)Itti, (c)Bremond, (d)ours-Itti, (e)ours-Harel, (f)HDR 画像, (g)Itti, (h)Bremond, (i)ours-Itti, (j)ours-Harel

従来手法で得られた顕著性マップをそれぞれ図 2(b),(c),(g),(h) 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 2(d),(e),(i),(j) 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 2(g),(h) では、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。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. おわりに

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。

参考文献

- [1] L. Itti, C. Koch, and E. Niebur, “A model of saliency based visual attention for rapid scene analysis”, IEEE Trans. PAMI, 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] Spivak, A. Belenky, A. Fish and O. Yadid-Pecht, “Wide-dynamic-range CMOS image sensors - comparative performance analysis”, IEEE Trans. On Electron Devices, 56, 11, pp.2446–2461, Nov (2009).
- [4] R. Bremond, J. Petit, J.P. Tarel, “Saliency maps of high dynamic range images”, Media Retargeting Workshop in Conjunction with ECCV’10 (2010).
- [5] 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).