

FlavorLUT: 大規模言語モデルを活用した写真編集手法の提案

小澤 知夏* 皆川 達也† 落合 陽一*†

概要. 本研究では、大規模言語モデル (LLM) を活用した写真編集手法を実現する Web アプリ「FlavorLUT」を提案する。専門的なパラメータ操作が困難な初心者を対象とし、パラメータの代わりに「フレーバーノート」という直感的な語彙を選択するだけで、編集パラメータを生成する。さらに、LLM が関連語彙を提案して表現の幅を広げたり、XY パッド UI でニュアンスを直感的に調整したりする機能を使用できる。専門知識がなくても容易に深い写真表現を追求できる環境の提供を目指す。

1 はじめに

デジタル写真の普及に伴い、撮影後のレタッチは写真表現において不可欠なプロセスとなっている。Adobe Lightroom¹などの高機能な編集ソフトウェアは、多岐にわたるパラメータ操作を提供するが、これらを初心者を使いこなす、望む印象を再現することは困難である。

この課題に対し、AI を用いた自動写真編集の研究も存在するが [3][2][1]、その多くは露出や色かぶりを補正する「機能的レタッチ」を主目的としている。一方で、写真全体の色味やトーンを変え、独自の世界観を付与する「装飾的レタッチ」は、未だユーザーの試行錯誤に大きく依存しており、初心者を取り組みにくい。

特に、自然言語を用いた編集手法は直感的だが、ユーザーが「夕焼けのように」といった具体的な言葉を持たなければ、思い通りの編集結果を得ることが難しいという「語彙力依存」の課題があった。

そこで本研究では、この装飾的レタッチのプロセスに着目し、LLM を活用して語彙力依存の問題を解決するパラメータレスな写真編集手法「FlavorLUT」を提案する。

2 提案手法

提案手法「FlavorLUT」は、ブラウザ上で動作する Web アプリケーションとして実装されており (図 1,2)、「フレーバーノート」「AI による調整支援」の 2 要素で構成される。

まず、ユーザーは編集指針として、写真の特徴に関する語彙群であるフレーバーノートから好みの印象に関する単語を選ぶ。語彙は以下の 3 カテゴリーで

構成される。

- **環境的語彙:** 時間、天候、場所など。
- **写真の印象に関する語彙:** 暖かい、ノスタルジックなど、感覚的な印象。
- **写真表現に関する語彙:** ハイキー、ソフトフォーカスなど、技術・表現的な側面。

これらを組み合わせ、編集のレシピを作成する。

次に、LLM を活用し、より深い編集を支援する。ユーザーの選択語彙に基づき、LLM が「黄昏時」「エモい」といった関連語彙を提案し、表現の幅を広げる (AI による語彙提案)。さらに、生成されたパラメータに対して「シネマティック-フィルムライク」などの 4 印象を軸とする XY パッドを操作し、LLM が座標を解釈して編集の効き具合を直感的に微調整する (XY パッドによるニュアンス調整)。

3 結果

提案手法の有効性を評価するため、いくつかの異なる入力写真を用いて実験を行った。

3.1 多様な編集スタイルの生成

同じ入力写真に対し、3 つの異なるフレーバーノートの組み合わせを与え、生成された LUT (パラメータ) を適用した。その結果、入力からは指定した語彙に基づき 3 種類の異なるパラメータ群が出力され、それらを適用することで、意図した通りの異なる写真表現が確認できた。この傾向は、複数の異なる入力写真を用いても同様に確認された。

3.2 AI による微調整の効果

XY パッドによるニュアンス調整の実験を行った。例えば、あるベース設定から「フィルムライク」の方向に調整すると、シャドウ部に青みがかかった特徴的なテイストが付与された。また、「レトロ+シネマティック」の方向に調整すると、全体的にコントラストが抑えられ、柔らかい雰囲気の写真が生成された。これらのニュアンス調整の効果は、特定の入力

Copyright is held by the author(s). This paper is non-refereed and non-archival. Hence it may later appear in any journals, conferences, symposia, etc.

* 筑波大学

† デジタルネイチャー開発研究センター

¹ <https://lightroom.adobe.com/>

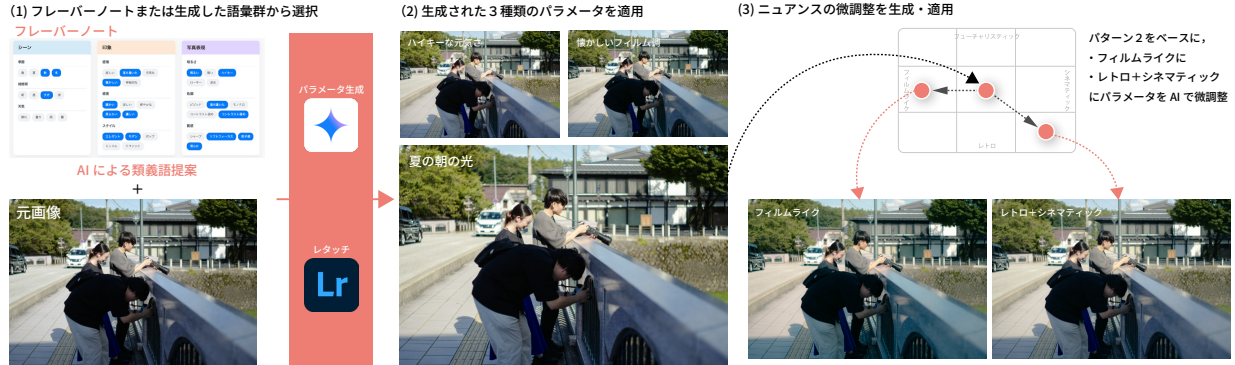


図 1. FlavorLUT の概要. (1) (3) の流れで写真編集用のパラメータを得て、そのパラメータは Adobe Lightroom で使用することを想定している. 類義語の生成・パラメータの生成・微調整後のパラメータの生成には Gemini2.5Lite を使用した.

写真に依存せず、複数の写真で共通して確認することができた.

4 今後の展望

本研究で提案した FlavorLUT は、専門知識がなくとも多様な編集を可能にしたが、いくつかの課題が残る.

今後の展望として、第一に、Fine-tuning などを用いてユーザー個人の編集スタイルや癖を反映させる機能の検討が挙げられる. 第二に、フレーバーノートの基底語の体系化である. 現在は著者の判断で語彙を選定しているが、調査に基づき、編集空間を網羅する普遍的な語彙セットを定義する研究が必要である.

さらに、生成されたパラメータが色空間に与える影響、すなわち潜在空間を可視化し、システムの透明性を高めることも目指す.

最後に、WISS (インタラクティブシステムとソフトウェアに関するワークショップ) などの場で多くの被験者に本システムを評価してもらい、フィードバックを収集することで、さらなる改良を進めていきたい.

参考文献

- [1] T. Brooks, A. Holynski, and A. A. Efros. InstructPix2Pix: Learning To Follow Image Editing Instructions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 18392–18402, June 2023.
- [2] V. Bychkovsky, S. Paris, E. Chan, and F. Durand. Learning photographic global tonal adjustment with a database of input/output image pairs. In *CVPR 2011*, pp. 97–104, 2011.
- [3] N. S. Dutt, D. Ceylan, and N. J. Mitra. Monet-GPT: Solving Puzzles Enhances MLLMs' Image

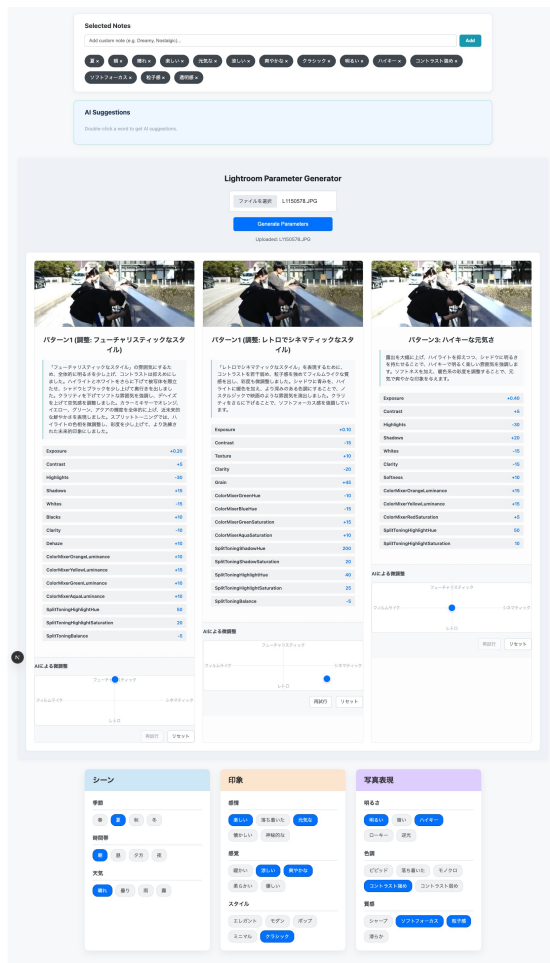


図 2. web アプリケーション全体の UI.

Retouching Skills. *ACM Trans. Graph.*, 44(4),
July 2025.

未来ビジョン

本研究の根底には、写真編集における「装飾的レタッチ」のプロセスを根本から見直し、より多くの人々が自らの表現を探究できる環境を構築したいというビジョンがある。

写真レタッチは、被写体の欠陥修正やノイズ除去といった「機能的レタッチ」と、写真全体の色味やトーン（調子）を調整し、撮影者の意図や独自の世界観を反映させる「装飾的レタッチ」に大別できる。近年、AI技術の進展は目覚ましく、前者の機能的レタッチは多くのソフトウェアで高度に自動化されつつある。

一方で、後者の装飾的レタッチは、未だ属人的な領域に留まっている。現在主流となっている手法は、プロのクリエイターが作成したプリセットやLUTをダウンロードし、自らの写真に適用することである。これは、手軽に高品質な結果を得られる反面、本質的には「他者の表現のコピー」に過ぎない。このプロセスで

は、自らの感性に基づいて試行錯誤したりする機会が失われがちである。結果として、個々のユーザーが「自分らしさ」とは何かを深く探究し、独自の表現スタイルを確立するプロセスが阻害されているのではないだろうか。

我々は、この「装飾的レタッチ」のプロセスを、専門知識を持つ一部の層から解放し、誰もが直感的に扱えるものへと「民主化」することを目指している。パラメータという専門知識の壁を取り払い、人間が最も得意とする「印象の言語化（フレーバー）」によって写真の意図を伝達する本手法 FlavorLUT は、そのビジョンの実現に向けた重要な第一弾であると位置づけている。